Adaptive dynamic range compression for improving envelope-based speech perception: Implications for cochlear implants

Ying-Hui Lai, Fei Chen and Yu Tsao

Abstract The temporal envelope is the primary acoustic cue used in most cochlear implant (CI) devices for eliciting speech perception in implanted patients. Due to biological constraints, a compression scheme is required to adjust the wide dynamic range (DR) of input signals to a desirable level. Static envelope compression (SEC) is a well-known strategy used in CI speech processing, where a fixed compression ratio is adopted to narrow the envelope DR. More recently, a novel, adaptive envelope compression (AEC) strategy has been proposed. In contrast to the SEC strategy, the AEC strategy more effectively enhances the modulation depth of the envelope waveforms to make the best use of the DR, in order to achieve higher intelligibility of envelope-based speech. In this chapter, we first introduce the theory of and implementation procedures for the AEC strategy. Then, we present four sets of experiments that were designed to evaluate the performance of the AEC strategy. In the first and second experiments, we investigated AEC performance under two types of challenging listening conditions: noisy and reverberant. In the third experiment, we explore the correlation between the adaptation rate using the AEC strategy and the intelligibility of envelope-compressed speech. In the fourth experiment, we investigated the compatibility of the AEC strategy with a noise reduction (NR) method, which is another important facet of a CI device. The AEC-processed sentences could provide higher intelligibility scores under challenging listening conditions than the SEC-processed sentences. Moreover, the adaptation rate was an important factor in the AEC strategy for producing envelope-compressed speech with optimal intelli-

Ying-Hui Lai
Department of Electrical Engineering, Yuan Ze University, Chung Li, Taiwan. e-mail: yhlai@ee.yzu.edu.tw

Fei Chen
Department of Electrical and Electronic Engineering, Southern University of Science and Technology, Shenzhen, China. e-mail: fchen@sustc.edu.cn

Yu Tsao
Research Center for Information Technology Innovation, Academia Sinica, Taipei, Taiwan. e-mail: yu.tsao@citi.sinica.edu.tw

gibility. Finally, the AEC strategy could be integrated with NR methods to enhance speech intelligibility scores under noisy conditions further. The results from the four experiments imply that the AEC strategy has great potential to provide better speech perception performance than the SEC strategy, and can thus be suitably adopted in CI speech processors.

1 Introduction

A cochlear implant (CI) is currently the only electronic device that can facilitate hearing in people with profound-to-severe sensorineural hearing loss (SNHL). In a report published in December 2012, the US Food and Drug Administration estimated that approximately 324,200 people worldwide have received cochlear implants, and predicted that this number will continue to increase in the near future [1]. Although CI devices can considerably improve the hearing capabilities of individuals with profound-to-severe SNHL in quiet environments, their efficacy markedly degrade in a noisy and/or reverberating environment [2, 3]. In a CI device, the input sound signal is received via a microphone and fed into a speech processor. The speech processor captures the multi-channel temporal envelopes of the input signal, and then generates electric stimulations that directly excite the residual auditory nerves [4, 5]. Due to biological constraints, the dynamic range (DR) of stimulation generated by a speech processor in a CI is much smaller than that of a real speech signal. Hence, a compression scheme is required to compress the DR of the input signal to a desirable level.

In the past, several compression strategies have been proposed. Among them, static envelope compression (SEC) is a popular method that is widely adopted in current CI devices. The SEC strategy uses a fixed compression ratio (CR) [6] to convert the acoustic amplitude envelope to an electric current signal [7, 8]. Although the SEC strategy can confine the overall electrical current signal applied to the CI to within a preset DR, the CR is not optimized to make the best use of the DR for speech perception. Yet, the DR of a temporal envelope is an important factor in speech intelligibility for CI users [9, 10, 11], particularly under noisy conditions. Therefore, designing a satisfactory, adaptive compression strategy that can perform compression effectively while maintaining a satisfactory DR is important.

More recently, a novel, adaptive envelope compression (AEC) strategy has been proposed [12, 13, 14]. The AEC strategy aims to provide deeper modulation than the SEC strategy to endow CI recipients with improved speech intelligibility. The AEC strategy shares a concept with the adaptive wide-dynamic range compression (AWDRC) amplification scheme, which is designed for hearing aids [15]. The AEC strategy aims to optimize the CR while confining the compressed amplitude of the speech envelope within a preset DR. To this end, the AEC strategy dynamically updates the CR in real time, so that the local DR of the output envelope waveform can be effectively increased, resulting in a larger modulation depth and improved intelligibility.

In this chapter, we present four sets of experiments to show the effects of the AEC strategy. We first evaluated the performance of AEC under challenging listening conditions. Then, we investigated the effects of the adaptation rate in the AEC strategy on the intelligibility of envelope-compressed speech. Finally, we investigated the compatibility of the AEC strategy with noise reduction (NR) methods. We show that the AEC strategy has better speech perception performance than the SEC strategy, and can be suitably adopted in a CI speech processor.

The remainder of this chapter is organized as follows. Section 2 reviews speech processing in a CI device. Section 3 introduces vocoder-based speech synthesis, which is an important tool and is popularly used in the field of CI research. Section 4 introduces the compression scheme, including the conventional SEC strategy and the proposed AEC strategy. Section 5 presents the experimental setup and four sets of experimental results. Finally, Section 6 provides the concluding remarks and summarizes this chapter.

2 Speech processor in CI devices

A CI device consists of four fundamental units: (1) a microphone that picks up the sound, (2) a signal processor that converts the sound signals into electrical signals, (3) a transmission system that transmits the electrical signals to implanted electrodes, and (4) an electrode array that is implanted into the cochlea [16, 17]. The combination of the first two units comprises the speech processor in a CI device. Figure 1 shows the design of a speech processor with eight channels. As shown in the figure, sound signals are first picked up by a microphone and then passed through a pre-emphasis filter to reduce the energy of the low-frequency components. The emphasized signals are then processed through a set of bandpass filters (eight filters in Fig. 1) to generate a set of bandpass signals. These bandpass signals are then processed by rectifiers and low-pass filters to create a set of temporal envelopes, each representing a specific frequency band. In general, the electrical DR of a CI recipient, between the threshold current (T-level) and the maximum comfortable current (C-level), is 5-20 dB [18, 19]. This DR is markedly less than that of sound levels encountered in real-world environments. For example, the DR of speech signals for a single speaker is 40-50 dB [9, 20]. Therefore, the speech processor of a CI device requires a compression (or automatic gain-control [17, 21]) strategy. The compression stage (i.e., COMP. in Fig. 1) is used to compress the wide temporal envelope to output magnitudes within the ranges of 0 to 1 [17]. The base level is the envelope level that produces a magnitude of 0, which yields a current at T-level. The saturation level is the envelope level that produces a magnitude of 1, which yields a current at C-level. Finally, the temporal envelopes are multiplied by non-simultaneous, biphasic pulse trains delivered as an electrical current through the cochlea via the electrode array.

Speech perception by CI users is primarily facilitated via temporal envelopes. The compression stage narrows down the envelope DR and, accordingly, may re-

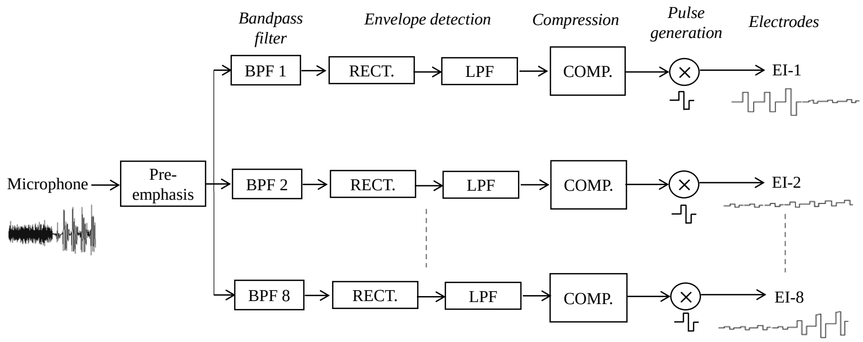

Fig. 1 Block diagram of an eight-channel speech processor. [4].

duce speech comprehension by CI recipients, especially under challenging listening conditions (e.g., in noise and/or reverberation). Therefore, a suitable compression scheme is requisite so that CI users can obtain additional temporal envelope information.

3 Vocoder-based speech synthesis

Although the number of CI recipients has greatly increased in recent years, one key challenge remains in the CI research field: it is difficult to conduct experiments on real CI recipients. To address this challenge, vocoder simulations derived from CI speech processing strategies have been presented to normal-hearing (NH) listeners in an attempt to predict the intelligibility of CI speech processing [12, 22]. Many studies have shown that vocoder simulations could predict the pattern of the performance observed by CI users, including the effects of background noise [23], the type of speech masker [24], and the number of electrodes [2, 25, 26].

Figure 2 shows the block diagram of an eight-channel tone-vocoder. The signal processing units of the tone-vocoder are similar to those of the CI speech processor (refer to Fig. 1). The input signals are first processed through the pre-emphasis filter (with a 3-dB/octave roll-off and 2000-Hz cutoff frequency). The bandpass filters (sixth-order Butterworth filters) are then used to filter the emphasized signal into eight frequency bands between 80 and 6000 Hz (with cutoff frequencies of 80, 221, 426, 724, 1158, 1790, 2710, 4050, and 6000 Hz). The temporal envelope of each spectral channel is extracted by a full-wave rectifier, followed by a low-pass filter. The envelope of each band is then compressed using a compression strategy (COMP in Fig. 3). Here, we implemented both the SEC and AEC strategies for COMP (refer to Figs. 3 and 4). The SEC strategy uses a fixed CR to confine the DR of the amplitude of the entire envelope to within a preset value; the AEC strategy, on the other hand, dynamically varies the CR value in a frame-by-frame manner (e.g., 2.5

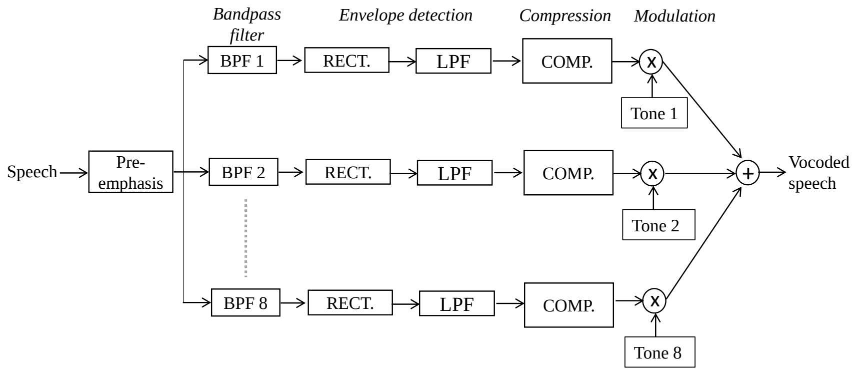

Fig. 2 Block diagram of an eight-channel tone-vocoder.

ms in this study), with the maximum and minimum values of the compressed amplitude limited to a preset range. The compressed envelopes are then modulated using a set of sine waves (i.e., tone *i* in Fig. 2), where the frequencies for these sine waves are equal to the center frequencies of the bandpass filters. Finally, the amplitude-modulated sine waves of the eight bands are summed, and the level of the summed signal is adjusted to yield a root-mean-square (RMS) value equal to that of the original input signal. Notably, the vocoder simulations are not expected to predict the absolute results, but rather the performance trends noted by CI users. In the present chapter, we adopted the tone-vocoder, as shown in Fig. 2 to conduct four speech recognition tests on NH subjects.

4 Compression scheme

The goal of the compression scheme in a CI device is to convert the sound signal into an electrical signal within the preset DR. In this section, we first review a popular compression scheme, namely the SEC strategy. Then, we present the recently proposed AEC strategy.

4.1 The static envelope compression strategy

The SEC strategy uses a linear transformation to compress the DR of the input signal to a desirable level [6]. Figure 3 shows the block diagram of the SEC-based speech processor in one channel. In Fig. 3, x and y denote the envelopes of input and output envelope signals, respectively. The output-compressed amplitude envelope signal y is computed as:

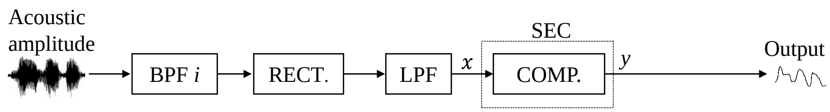

Fig. 3 Block diagram of the *i*channel speech processor using the SEC strategy.

$$y = \alpha \times (x - \bar{x}) + \bar{x}, \tag{1}$$

where $\bar{x}$ is the mean of the input amplitude envelope x, and α is a scaling factor (SF), which is determined in order to ensure that the output amplitude envelope falls within a desirable DR:

$$UB = LB \times 10^{\frac{DR}{20}}. \tag{2}$$

In Eq. (2), UB and LB denote the upper bound (i.e., the maximum) and lower bound (i.e., the minimum) of the output amplitude values, respectively. From Eq. (2), it can be seen that the mean value of the output amplitude envelope equals that of the input amplitude envelope (i.e., $\bar{y} = \bar{x}$).

Notably, using a small SF in Eq. (1) induces a large CR, and vice versa. When α equals 0, the compressed amplitude envelope becomes a direct current (DC) signal with a constant value of $\bar{y}$ (i.e., $\bar{y} = \bar{x}$), and the DR is 0 dB; when α equals 1, the output amplitude envelope has the same DR as the input amplitude envelope. For the SEC strategy, α is set by audiology, and a fixed value is applied to the whole amplitude envelope so as to confine its DR to a preset value. A previous study [11] showed that by setting α to 1/3, 1/5, and 1/13 in Eq. (1), the DR of multi-channel amplitude envelopes of the Mandarin version of sentences for the Hearing in Noise Test (MHINT) [27] were adjusted to 15, 10, and 5 dB, respectively.

4.2 The adaptive envelope compression strategy

For the SEC strategy as shown in Fig. 3, a fixed α is applied to the whole amplitude envelope to confine its DR to a preset value. Although fixed mapping can effectively confine the DR, the SEC strategy does not make optimal use of the DR for speech perception. Recently, the AEC strategy was proposed to confine the amplitude envelope of speech signal within a fixed DR while continuously adjusting the SF for short-term amplitude [12, 13, 14]. Using the AEC strategy, the local DR approaches that of the uncompressed amplitude envelope, and thus the AEC-processed amplitude envelopes yield higher intelligibility than do the SEC-processed envelopes [12, 13, 14]. Figure 4 demonstrates the block diagram of the AEC-based speech processor in one channel. Compared with the SEC strategy (in Fig. 3), the AEC strategy uses a feedback unit to calculate the bounds and to apply the AEC rules and peak clipping unit in each channel. In Fig. 4, x, z, and y denote the input, compressed, and output envelopes, respectively, and $\Delta\alpha$ is the step size. Instead of using a fixed SF, the AEC strategy adjusts the SF value for each speech frame to compute the compressed amplitude envelope:

$$z_t = \alpha_t \times (x_t - \bar{x}) + \bar{x}, \tag{3}$$

where z_t and x_t are the compressed and original envelopes, respectively, at the t-th speech frame, and α_t is the adaptive SF that is updated with $\Delta\alpha$:

$$\alpha_{t+1} = \alpha_t + \Delta\alpha. \tag{4}$$

In Fig. 4, the boundary calculation and the AEC rules units are used to determine $\Delta\alpha$ in Eq. (4). Given an input envelope, the boundary calculation unit computes the DR of the compressed envelope by estimating two bounds: the upper bound, UB; and the lower bound, LB, represented as:

$$\begin{aligned} UB &= \bar{x} + \alpha_0 \times (max(x) - \bar{x}) \\ LB &= \bar{x} + \alpha_0 \times (min(x) - \bar{x}) \end{aligned}, \tag{5}$$

where $max(x)$ and $min(x)$ are the maximum and minimum values of input amplitude envelope x, and α_0 is an initial SF for the AEC strategy. Generally speaking, the fixed compression rate used for the SEC strategy, α in Eq. (1), can be used as the initial SF α_0 for the AEC strategy the fixed compression rate used for the SEC strategy, such as α in Eq. (1), can be used as the initial SF α_0 for the AEC strategy in Eq. (3) and (5).

With the estimated UB and LB based on Eq. (5), $\Delta\alpha$ in Eq. (4) is determined by two AEC rules, as:

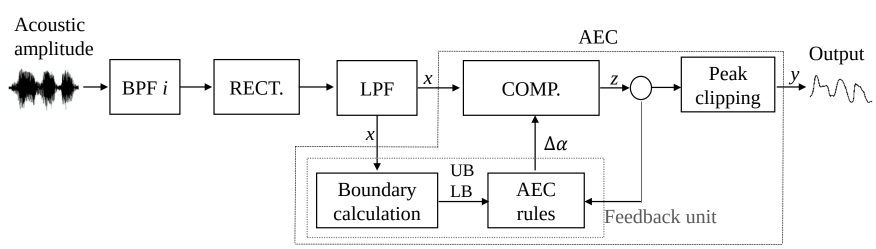

Fig. 4 The block diagram of the i-th channel speech processor with the AEC strategy.

1. *Increasing-envelope rule*: This rule aims to keep the compression process as close to linear as possible, such that $\alpha = 1$. By doing so, fewer input signals will be perturbed by compression when producing the output signal. When z_t lies between UB and LB, the AEC strategy will increase α_t by using a positive $\Delta\alpha$ in Eq. (4), accordingly increasing the SF value. This increasing-envelope rule stops when α_{t+1} reaches 1, which corresponds to the absence of compression application. In this scenario, the original signal is used as the output signal.

2. *Decreasing-envelope rule*: This rule aims to ensure that the amplitude of the output envelope will not fall outside the preset DR. When z_t in Eq. (3) becomes larger than UB, or lower than LB, the AEC strategy will decrease t by using a negative $\Delta\alpha$ in Eq. (4), which accordingly reduces the SF value. This decreasing-envelope rule stops when α_{t+1} reaches the initial value, α_0.

The AEC strategy follows the above two rules to dynamically adjust the SF values so as to compress the input amplitudes to fit the DR of the electrical current applied to the CI. However, when sudden changes in the input envelope happen, overshooting or undershooting may occur. To overcome such issues, the peak clipping unit is applied, as shown in Fig. 4, to ensure that the output envelope falls between the maximum and minimum levels, where UB and LB in Eq. (5) can be used as maximum and minimum values, respectively. Finally, the compressed amplitude envelope is computed as:

$$\begin{cases} y_t = z_t, & if\ LB < z_t < UB \\ y_t = UB, & if\ z_t \geqslant UB \\ y_t = LB, & if\ z_t \leqslant LB \end{cases} . \tag{6}$$

Figures 5 (a) and (b) show examples of the SEC- and AEC-processed amplitude envelopes of the *6*-th channel (i.e., flow = 1790 Hz, fhigh = 2710 Hz) in a noisy environment, when masked by speech-shape noise (SSN) at 0 dB for the signal-to-noise ratio (SNR). The UB and LB in this example are 125.3 and 70.6, respectively, yielding a DR of approximately 5 dB. As seen in Fig. 5 (a) and (b), both the SEC and AEC strategies effectively compress the DR of the amplitude envelope within the preset DR; however, the local (e.g., around 1.1 s) DR is larger for the envelope processed using the AEC strategy, which had a DR of 4.3 dB in Fig. 5 (b), than for that processed by the SEC strategy, which had a DR of 2.3 dB in Fig. 5 (a). Figure 5 (c) shows the SF values used in the SEC and AEC strategies for compression of the sentences in Fig. 5 (a) and Fig. 5(b), respectively. From Fig. 5 (c), the SEC strategy applies a fixed SF of α = 1/13, while the AEC strategy continuously adjusts its SF α_t. It has been noted that the SF α_t for the AEC strategy is generally larger than the fixed SF of α = 1/13 employed in the SEC strategy. These results indicate that the AEC strategy can modulate the SF, based on the characteristics of the input signals so as to utilize the usable DR optimally. Moreover, the AEC strategy generates an amplitude envelope with a larger DR and, consequently, a larger modulation depth.

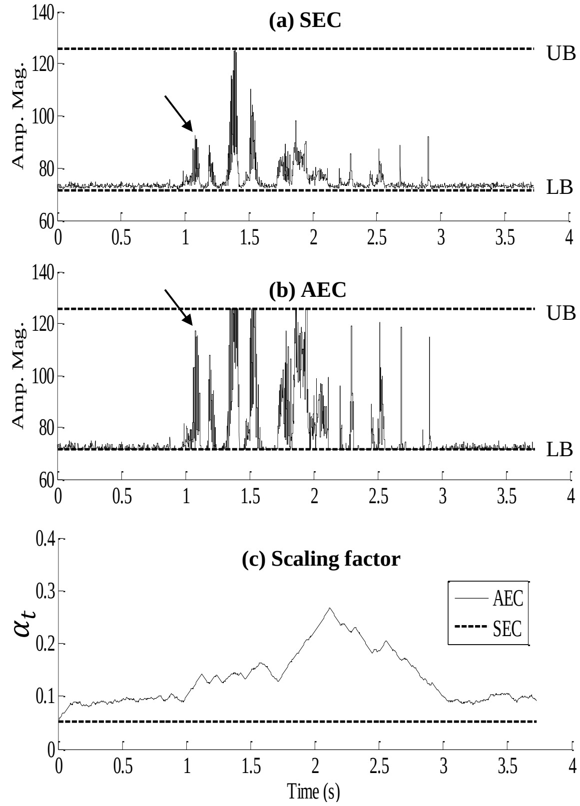

Fig. 5 Examples of the amplitude envelope processed using the (a) SEC and (b) AEC (i.e., $\Delta\alpha$ was 0.001 per 2.5 ms). The envelope waveforms were extracted from the *6*-th channel (i.e., f_{low} = 1790 Hz, f_{high} = 2710 Hz) of a testing sentence masked by an SSN masker at 0 dB, and compressed to a DR of 5 dB. In (c), the solid line shows the SF α_t used in the AEC strategy for the compressed amplitude envelope in (b); the dashed line indicates the fixed SF used in the SEC strategy in (a).

5 Experiments and results

In this section, we present four sets of experiments used to verify the effectiveness of the AEC strategy. In Experiment-1 and Experiment-2, we tested the performance of the AEC strategy in noise and in reverberation, respectively. In Experiment-3, we explored the effect of the adaptation rate in the AEC strategy. Experiment-4 was designed to investigate whether the advantage of a front-end noise reduction scheme could be preserved upon integration with a subsequent AEC strategy, and how this advantage would be influenced by the factors of input SNR, type of NR, and type of noise.

5.1 Experiment-1: The speech perception performance of AEC in noise

CI recipients have limited hearing DR for speech perception. This may partially account for their poor speech comprehension, particularly under noisy conditions. The proposed AEC strategy aims to maximize the modulation depth for CI recipients, while confining the compressed amplitude envelope to the preset DR. The purpose of Experiment-1 was to compare speech recognition synthesized by the proposed AEC strategy and by the SEC strategy under noisy conditions.

5.1.1 Subjects and materials

Eleven (age range: 18-24 years; six females and five males) NH native-Mandarin speakers were recruited to participate in the listening tests. Sentence lists from the MHINT database were used to prepare the testing materials [27]. All sentences were pronounced by a male native-Mandarin speaker, with a fundamental frequency ranging from 75 to 180 Hz, and recorded at a sampling rate of 16 kHz. Two types of maskers: an SSN and two equal-level interfering male talkers (2T), were used to corrupt the testing sentences at two SNR levels (5 and 10 dB), which were chosen to avoid the ceiling/floor effects.

5.1.2 Procedure

The listening tests were conducted in a soundproof booth. The stimuli were played to listeners through a set of Sennheiser HD headphones at a comfortable listening level. The speech was compressed to an envelope DR of 5 dB in a vocoder simulation, which was done by using the SF $\alpha = 1/13$ and $\alpha_0 = 1/13$ in Eq. (1) and Eq. (4), respectively. Each subject participated in a total of eight [2 SNR levels $\times$ 2 types of maskers $\times$ 2 envelope compression strategies, i.e., the SEC and AEC strategies, respectively] testing tasks. Each task contained 10 sentences, and the order of the eight tasks was randomized across subjects. None of the 10 sentences were repeated across testing tasks. Subjects were instructed to repeat what they heard, and were allowed to listen to each stimulus twice. The sentence recognition score was used to evaluate the performance, which was calculated by dividing the number of words correctly identified by the total number of words in each testing task. During testing, each subject was given a 5-minute break every 30-minutes during the test.

5.1.3 Results and discussion

Figure 6 shows the listening test results in terms of mean sentence recognition scores at different SNR conditions. As shown in Fig. 6, the AEC strategy yielded higher

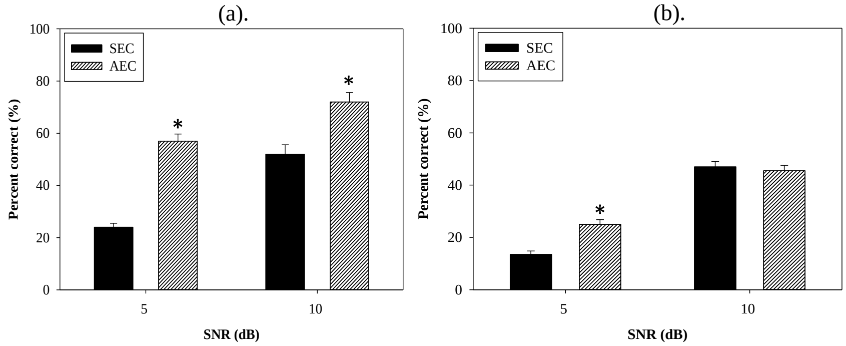

Fig. 6 The mean recognition scores of SEC and AEC in (a) SSN and (b) 2T maskers. The DR of the envelope amplitude is confined to 5 dB. The error bars denote the standard errors of the mean (SEM) values. An asterisk indicates a statistically significant ($p<0.05$) difference between SEC and AEC scores.

speech recognition performance than did the SEC strategy in the three noisy testing tasks (5 dB and 10 dB for SSN, and 5 dB for 2T). To confirm the significance of the improvements further, two-way analysis of variance (ANOVA) and post-hoc comparisons were used to analyze the results of the two strategies under the four noisy conditions.

For the SSN results in Fig. 6, the ANOVA measures indicated significant effects of SNR level (F[1, 10]=36.03, $p<0.005$), compression strategy (F[1, 10]=33.41, $p<0.005$), and the interaction between SNR level and compression strategy (F[1, 10]=5.52, $p=0.041$). The post-hoc analyses further confirmed that the score differences between the SEC-processed sentences and AEC-processed sentences were significant ($p<0.05$). For the 2T results, shown in Fig. 6 (b), the ANOVA measures indicated the significant effect (F[1, 10]=77.87, $p<0.005$) of SNR level, a non-significant effect of compression strategy (F[1, 10]=3.14, $p=0.107$), and a significant interaction between SNR level and compression strategy (F[1, 10]=12.63, $p=0.005$). The post-hoc analyses further showed that the score difference at 5 dB SNR was significant ($p<0.05$) while that at 10 dB SNR was non-significant ($p=0.46$) in Fig. 6 (b).

To analyze the advantage of the AEC strategy further, it is worthwhile to revisit the examples shown in Fig. 7. From this figure, it can be seen that the AEC-processed envelope around 1 second and 2 second in Fig. 7 (a) has a larger DR than that processed by the SEC strategy in Fig. 7 (b). Moreover, Fig. 7 (c) demonstrates that the SEC strategy employs a fixed compression factor, where α = 1/13, while the AEC strategy uses a small amplitude compression ratio, or a large SF α_t, for the frames around 1 second and 2 second, thus yielding a large modulation depth for the amplitude envelope and improved speech intelligibility.

In summary, the results of this experiment showed that the intelligibility of the AEC-processed sentences was significantly better than their SEC-processed counterparts under noisy conditions. This suggests that the proposed AEC strategy holds

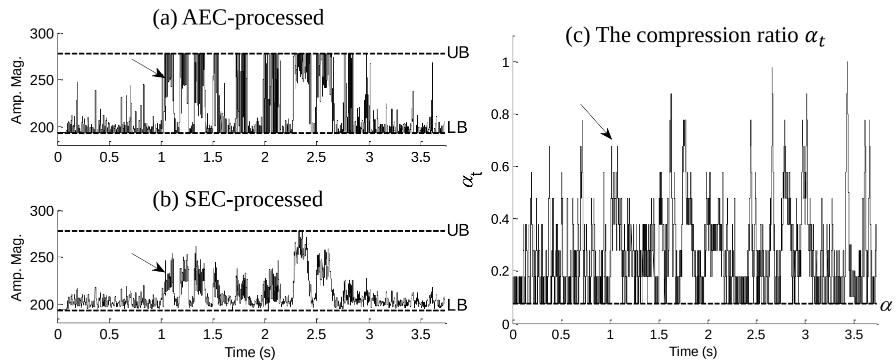

Fig. 7 Examples of amplitude envelope processed by (a) AEC (i.e., $\Delta\alpha = 0.1$) and (b) SEC strategies, and (c) the compression ratio α used in the AEC and SEC (dashed line). The envelope waveforms were extracted from the *6*-th channel of a testing sentence masked by SSN at 5 dB SNR, and compressed to 5 dB DR within [LB, UB].

promise for improving speech perception performance under noisy listening conditions for patients with CI.

5.2 *Experiment-2: The speech perception performance of AEC in reverberation*

Reverberation, which results from multiple reflections of sounds from objects and surfaces in an acoustic enclosure, causes spectro-temporal smearing of speech [28]. Previous studies have indicated that a reverberating environment may reduce a CI recipients ability to identify words [29, 30]. In Experiment-2, we intended to assess the effects of the AEC strategy in reverberation.

5.2.1 Subjects and materials

Nine (age range: 19-27 years, four females and five males) NH native-Mandarin speakers were recruited to participate in the listening tests. As with Experiment-1, sentence lists from the MHINT [27] were used to prepare the testing materials. The reverberant conditions were simulated by head-related transfer functions recorded in a 5.5 m $\times$ 4.5 m $\times$ 3.1 m (length $\times$ width $\times$ height) room with a total volume of 76.8 m^3 [31]. The average reverberation time of the experimental room (T_{60} = 1.0 s) was reduced to T_{60} = 0.6, 0.4 and 0.2 seconds by adding floor carpeting, absorptive panels on the walls, and a ceiling, respectively. Additional details on simulating reverberant conditions can be found in previous studies [31, 32].

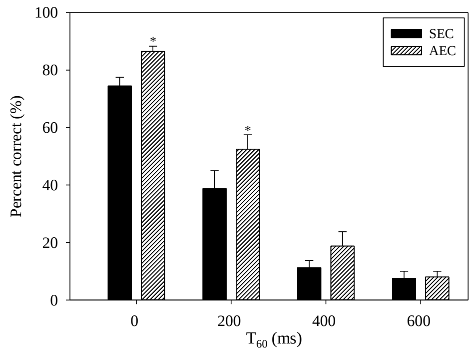

Fig. 8 The mean recognition scores of SEC and AEC. The DR of envelope amplitude is confined to 5 dB. The error bars denote the SEM values. An asterisk indicates a statistically significant ($p<$ 0.05) difference between SEC and AEC scores.

5.2.2 Procedure

The listening tests were conducted in a soundproof booth. The stimuli were played to listeners through a set of Sennheiser HD headphones at a comfortable listening level. The speech signals were compressed to the envelope DR of 5 dB in the vocoder simulation. This was done by using the SF α=1/13 and α_0=1/13 in Eq. (1) and Eq. (3), respectively. Each subject participated in a total of eight [= 4 reverberant conditions (i.e., T_{60} = 0, 200, 400, and 600 ms) $\times$ 2 envelope compression strategies (i.e., SEC and AEC)] testing tasks. Each task contained 10 sentences. The order of the eight tasks was randomized across subjects, and none of the 10 sentences were repeated across testing tasks. The subjects had repeated what they had heard during the experiments, and were allowed to listen to each stimulus twice. The sentence recognition score was used to compare speech recognition performance.

5.2.3 Results and discussion

Figure 8 shows the listening test results in terms of mean sentence recognition scores for all testing tasks. The two-way ANOVA measures were computed by using the recognition score as the dependent variable. The reverberant condition, the T_{60} value, and the compression strategy were considered as the two within-subject factors. ANOVA results indicated significant effects in the reverberant condition (F[3, 24] = 208.62, $p < 0.05$), the compression strategy (F[1, 8] = 35.01, $p < 0.05$), and in the interaction between the reverberant condition and compression strategy (F[3, 24] = 8.05, p = 0.001). Post-hoc analyses showed that the score differences between the SEC-processed sentences and AEC-processed sentences were significantly different ($p < 0.05$) under the reverberant conditions of T_{60}=0 and 200 ms in Fig. 8.

In accordance with the intelligibility advantage observed in Experiment-1, the results of Experiment-2 showed that the amplitude envelope processed by the AEC strategy yielded higher intelligibility scores for vocoded sentences under reverberation compared to those processed using the SEC strategy. This makes the AEC strategy a highly promising way of enhancing speech comprehension in implanted listeners under reverberant conditions in the future. Interestingly, the above intelligibility advantage was not observed for all reverberant conditions. This indicated that there was no significant improvement in intelligibility found under the reverberant conditions of T_{60}= 600 ms in Fig. 8. This may be partially attributed to the usage of initial compression parameters (e.g., α_0 and $\Delta\alpha$) in the experiment. It is worthwhile to investigate optimal configuration of compression parameters so as to achieve the best performance for the AEC-based speech processing for CI recipients under reverberant listening conditions in future studies.

5.3 *Experiment-3: The effect of adaptation rate on the intelligibility of AEC-processed speech*

As mentioned in Section 4.2, the AEC strategy specifies the rate at which the SF value should be updated, using the adaptation rate outlined in Eq. (4), which is determined based on the two AEC rules. The adaptation rate used in the AEC strategy is similar to the attack time (AT) and release time (RT) of the wide-dynamic-range compression (WDRC) amplification scheme that is widely used in hearing aids [33]. The AT and RT describe the duration required for a hearing aid device to respond to a changing input signal [34]. When setting an inadequate time constant value for AT/RT, the gain will fluctuate rapidly and thus generate an undesirable pumping effect. Conversely, when setting an excessive time constant value, a lag in perception will be induced. Previous studies have explored the effects of AT/RT values on the intelligibility and satisfactory sound quality of a hearing aid for its users [35, 36, 37, 38]. Their results indicated that AT and RT values should be carefully optimized in order to achieve satisfactory performances in speech intelligibility, listening comfort, and sound quality. For this reason, the adaptation rate is an important parameter in the AEC strategy. In this section, we investigate the effect of the adaptation rate in Eq. (4) on the intelligibility of the AEC-processed speech.

5.3.1 Subjects and materials

Eight NH, native-Mandarin listeners (age range: 19-26 years, four females and four males) listeners were recruited to participate in the listening experiment. Sentences from the MHINT were used as the testing materials [27]. Two types of maskers, SSN and 2T, were used to prepare the noisy testing sentences at SNR levels of 5 and 10 dB, which were chosen based on a pilot study to avoid ceiling and floor effects.

5.3.2 Procedure

The listening tests were conducted in a sound-proof room, and stimuli were played to listeners through a set of Sennheiser HD headphones at a comfortable listening level. The DR of the envelope waveforms were compressed to 5 dB in tone-vocoder simulations, where the values of SEC and α_0 of the AEC strategy were set to 1/13 [11]. Four different envelope compression methods were used in this experiment, i.e., SEC, AEC with $\Delta\alpha$ = 0.001, AEC with $\Delta\alpha$ = 0.01, and AEC with $\Delta\alpha$ = 0.1. The last three are referred to as AEC.001, AEC.01, and AEC.1, respectively. Each subject participated in a total of 16 (2 SNR levels $\times$ 2 types of maskers $\times$ 4 envelope compression methods) listening tasks. Each task contained 10 sentences, and the order of the 16 tasks was randomized across subjects. None of the 10 sentences were repeated across the listening task. Subjects were instructed to repeat what they heard, and they were allowed to listen to each stimulus twice. The sentence recognition score was used to evaluate the performance.

5.3.3 Results and discussion

Figure 9 demonstrates the speech recognition scores in terms of the mean recognition rates for all of the testing tasks. From Fig. 10 (a), we noted that the three AEC strategies (AEC.001, AEC.01, and AEC.1) produced notably higher intelligibility scores than the SEC strategy in the SSN test condition. From Fig. 9 (b), we noted that AEC.001 yielded notably higher recognition scores than did the SEC strategy in the 2T masker condition. One-way ANOVA and Tukey post-hoc comparisons were conducted to analyse the results of the four compression strategies in the four testing conditions further. These analyses are summarized in Table 1. In the table, each mean score represents the corresponding recognition score in Fig. 10, and n denotes the sample size. For the results of the SSN masker, the ANOVA measures confirmed that the intelligibility scores differed significantly across the four groups, with (F = 10.25, $p < 0.001$) and (F = 5.80, $p = 0.003$) at SNR levels of 5 dB and 10 dB, respectively. The Tukey post-hoc results further verified the significant differences for the following group pairs at both SNR levels of 5 and 10 dB: SEC with AEC.001; SEC with AEC.01; and SEC with AEC.1. Moreover, the ANOVA results for the 2T masker confirmed that the intelligibility scores differed significantly across the four groups, with (F = 3.00, $p = 0.048$) and (F = 3.49, $p = 0.029$) at SNR levels of 5 and 10 dB, respectively. The Tukey post-hoc comparisons verified the significant differences between the group pairs of SEC and AEC.001 at both SNR levels (5 and 10 dB).

These results confirm that the value of adaptation rate, $\Delta\alpha$, indeed affects the intelligibility of AEC-processed speech. The use of an inappropriate value of $\Delta\alpha$ may diminish the benefits affecting speech intelligibility, especially under testing conditions using an interfering masker. To demonstrate the effect of the adaptation rate further, Fig. 10 highlights the results obtained by the traditional SEC strategy; the AEC strategy with a very slow adaptation rate, where $\Delta\alpha$ = 0.0001; and the AEC

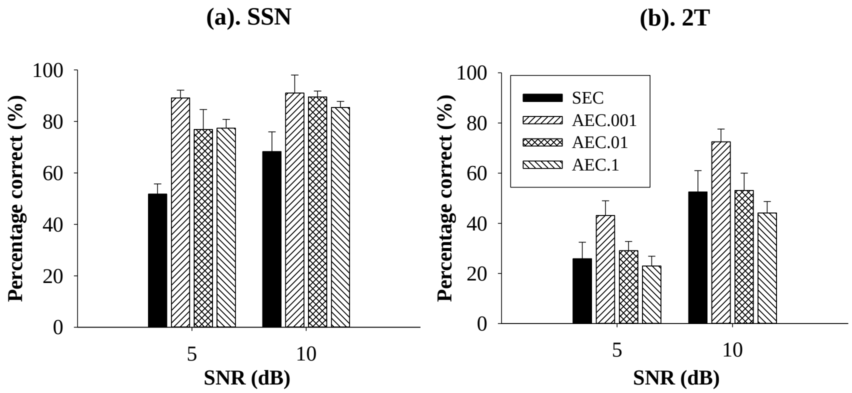

Fig. 9 The mean recognition scores for Mandarin sentences with (a) an SSN masker and (b) a 2T masker at SNR levels of 5 and 10 dB. The error bars indicate SEM values.

strategy with a fast adaptation rate, where $\Delta\alpha = 0.1$. The examples showed that, when using an inadequate $\Delta\alpha$ value, such as 0.0001 as shown in Fig. 10 (b), the benefits of the AEC strategy are limited, as the available DR is not effectively used. Conversely, when using a similarly inadequate $\Delta\alpha$ value, such as 0.1, as shown in Fig. 10 (c), a pumping effect may occur. Consequently, some envelopes will fall within the range of peak clipping, accordingly causing speech signal distortions. In addition, when the envelope waveform varies too rapidly, speech intelligibility will also be decreased. From the results of Figs. 9 and 10, when $\Delta\alpha$ is equal to 0.001, the optimal balance between benefits, the pumping effect, and the distortion in the AEC strategy becomes evident. In this experiment, we only selected these three values of $\Delta\alpha$ (slow, moderate, and fast adaptation rate) to investigate the performance of the AEC strategy. Future studies should further investigate the effect of $\Delta\alpha$ while considering the wider characteristics of language, such as the band importance function or the tone of Mandarin, in addition to the noise types.

5.4 *Experiment-4: The effect of joint envelope compression and noise reduction*

An NR method plays a crucial role in the improvement of sound quality/intelligibility in noisy conditions [39, 40, 41, 42]. Chung [43] found that NR methods greatly enhanced the modulation depth of noise-suppressed signals, but these benefits were somehow eliminated by the compression stage. On the other hand, the benefits of NR approaches can be maintained or even further improved by using a suitable compression strategy. Previous studies of NR have found that integrating the NR approach with a fixed compression ratio strategy benefits speech perception by CI

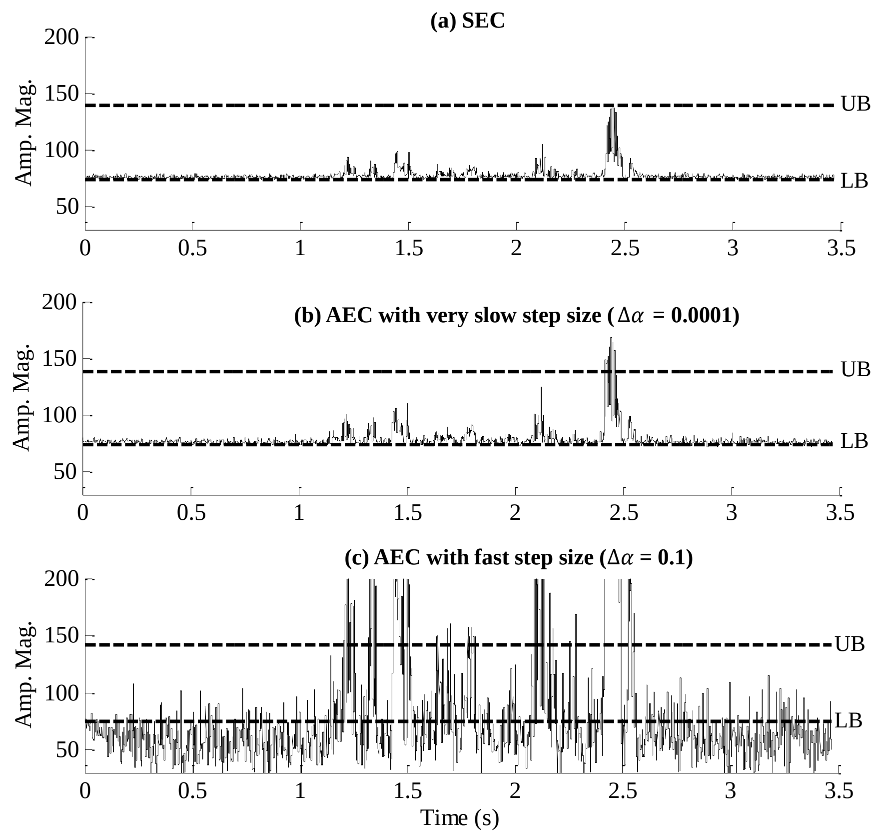

Fig. 10 Examples of the amplitude envelope processed by (a) SEC, (b) AEC with a very slow adaptation rate, $\Delta\alpha = 0.0001$, and (c) AEC with a fast adaptation rate, $\Delta\alpha = 0.1$. The envelop waveforms were extracted from the 6th channel of a testing sentence masked by SSN at 5 dB, and compressed to 5 dB DR.

recipients [44]. The aim of this experiment was to evaluate how the AEC strategy interacted with the NR approaches in the handling of noisy speech.

5.4.1 Subjects and materials

Eight NH, native-Mandarin listeners (age range: 19-26 years, four females and four males) were recruited to participate in the listening test. The MHINT sentences were used to test the performance, with SSN and 2T maskers used to corrupt testing sentences. The test speeches included 0, 5, and 10 dB SNR levels.

Table 1 The mean recognition scores for different strategies, where each factor was included in the one-way ANOVA and Tukey post-hoc testing.

Test condition	**Strategy**	***n***	**Mean score**	***F***	***p***	**Post-hoc comparison* (group$_i$, group$_j$)**
				10.25	<0.001	
SSN (SNR = 5 dB)	SEC	8	51.8			(SEC, AEC.001)
	AEC.001	8	89.1			(SEC, AEC.01)
	AEC.01	8	76.9			(SEC, AEC.1)
	AEC.1	8	77.4			
				5.80	0.003	
SSN (SNR = 10 dB)	SEC	8	68.3			(SEC, AEC.001)
	AEC.001	8	91.0			(SEC, AEC.01)
	AEC.01	8	89.5			(SEC, AEC.1)
	AEC.1	8	85.4			
				3.00	0.048	
2T (SNR = 5 dB)	SEC	8	25.9			
	AEC.001	8	43.2			(SEC, AEC.001)
	AEC.01	8	29.2			
	AEC.1	8	23.0			
				3.49	0.029	
2T (SNR = 10 dB)	SEC	8	52.5			
	AEC.001	8	72.5			(SEC, AEC.001)
	AEC.01	8	53.1			
	AEC.1	8	52.5			

5.4.2 Signal processing with NR and envelope compression

In this set of experiments, an NR process was implemented before the AEC strategy in an eight-channel tone-vocoder. Figure 11 shows the block diagram of the NR- and AEC-based tone-vocoder in one channel. In the figure, the noisy speech signal was first processed by an NR method and then fed into a standard tone-vocoder, where the SEC strategy or the AEC strategy could be used as the compression scheme. We adopted two types of NR approaches, namely the Wiener filtering approach [45] and the Karhunen Loeve theorem (KLT) algorithm [46] in this set of experiments. The Wiener filtering approach utilizes a priori SNR statistics to design a gain function to filter out noise components from the noise input. For the KLT method, the KLT algorithm is first applied to the noisy signal. The KLT components that represent the signal subspace were modified by a gain function, while the remaining KLT components that represent the noise subspace were nulled. An enhanced signal was obtained by applying the inverse KLT of the modified components. The techniques used in these two algorithms have been detailed in previous studies [45, 46].

Following the NR stage in Fig. 11, the envelope of the noise-suppressed signal was extracted via bandpass filtering and waveform rectification. Next, in the CR estimation (CRE) stage, the appropriate SF for the SEC strategy or the initial SF (i.e., α_0) for the AEC strategy, was determined based on the input envelope (x in Fig. 11). This SF was then used to transform the output envelope to generate a compressed signal (z in Fig. 11). The final peak clipping stage was used to confine the compressed envelope to within the expected DR (y in Fig. 11). The compressed envelopes were then modulated by a set of sine waves (i.e., tone i) with frequencies equal to the center frequencies of the bandpass filters. Finally, the envelope-modulated sine waves of the eight bands were combined, and the level of the combined signal was adjusted to produce an RMS value equal to that of the original input signal.

Figure 12 shows examples of an amplitude envelope processed by the Wiener filtering and KLT methods, followed by the SEC and AEC strategies, under same testing conditions (SSN masker at SNR 5 dB). The envelope was extracted from the sixth channel, and compressed to a 5-dB DR, where the initial SF was computed by the CRE stage. We noted two findings: (1) the KLT performance was similar to that of the Wiener filtering approach when integrated with the same compression strategy, and (2) the AEC strategy can provide better modulation depth than the SEC strategy when it is integrated with the same NR algorithm.

5.4.3 Procedure

As NR methods were used in this experiment, the SF should differ from the values used in Experiment-1, in order to ensure that the output envelope falls within the desirable DR, which was 5 dB in this experiment. As shown in Fig. 11, the CRE stage computes an α value for the SEC strategy and an α_0 for the AEC strategy, based on the NR-processed signals. Four different signal processing methods were used: (1) Wiener+SEC, (2) Wiener+AEC, (3) KLT+SEC, and (4) KLT+AEC. The adaptation rate ($\Delta\alpha$) of the AEC strategy was 0.001 in this experiment.

The listening tests were conducted in a soundproof booth. Each subject participated in a total of 24 (3 SNR levels $\times$ 2 types of maskers $\times$ 4 types of signal processing) testing tasks. Each task contained 10 sentences, and the order of these 24 tasks was randomized across subjects. None of the 10 sentences were repeated

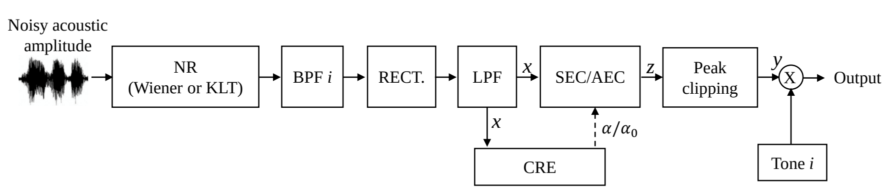

Fig. 11 Block diagram for obtaining the output tone-vocoded speech for speech enhancement methods followed by a compression strategy (either AEC or SEC) in the i-th channel.

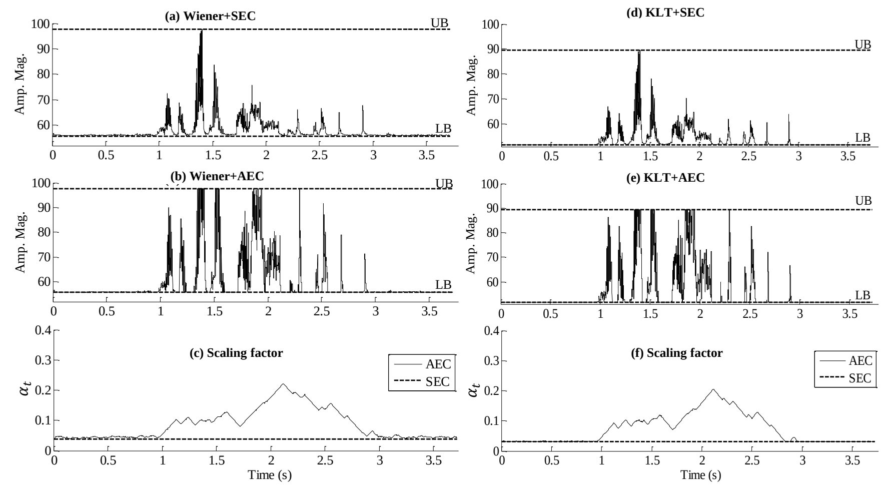

Fig. 12 Examples of the amplitude envelope for (a) Wiener+SEC, (b) Wiener+AEC, (d) KLT+SEC, and (e) KLT+AEC processing. The envelop waveforms were extracted from the *6*-th channel of a testing sentence masked by the SSN masker at SNR 5 dB, and compressed to a 5-dB DR. In (c) and (f), the solid lines show the SF used in the AEC strategy for (b) and (e); the dashed lines show the fixed SF in the SEC strategy for (a) and (d).

across the testing conditions. Subjects were instructed to repeat what they heard, and they were allowed to listen to each stimulus twice.

5.4.4 Results and discussion

Figure 13 shows the mean speech recognition scores for all of the testing tasks. For the SSN results indicated in Fig. 13 (a), the mean recognition rates for Wiener+SEC, Wiener+AEC, KLT+SEC, and KLT+AEC, respectively, were: 18.4%, 33.6%, 15.8%, and 41.1% for 0 dB SNR; 23.3%, 53.9%, 14.1%, and 53.0% for 5 dB SNR; and 15.7%, 52.6%, 24.1%, and 64.1% for 10 dB SNR. For the 2T results in Fig. 13 (b), the mean recognition rates were: 2.3%, 3.1%, 1.1%, and 2.6% for 0 dB SNR; 4.3%, 15.6%, 6.5%, and 14.4% for 5 dB SNR; and 17.1%, 40.6%, 17.3%, and 41.3% for 10 dB SNR. Three-way ANOVA measures was used to analyze these data for the following three factors: the type of masker (masker), SNR level (SNR), and processing method (F1). The results indicated that all of the main effects and second-order interaction were significant (refer to Table 2). Tukey's post-hoc analysis showed significant differences for the following group pairs: Wiener+SEC and Wiener+AEC; Wiener+SEC and KLT+AEC; Wiener+AEC and KLT+SEC; and KLT+SEC and KLT+AEC.

The results of the three-way ANOVA and Tukey post-hoc comparisons indicated that the AEC strategy can provide higher speech recognition scores than the SEC strategy when integrated with a Wiener filter or KLT. The reason for the poorer per-

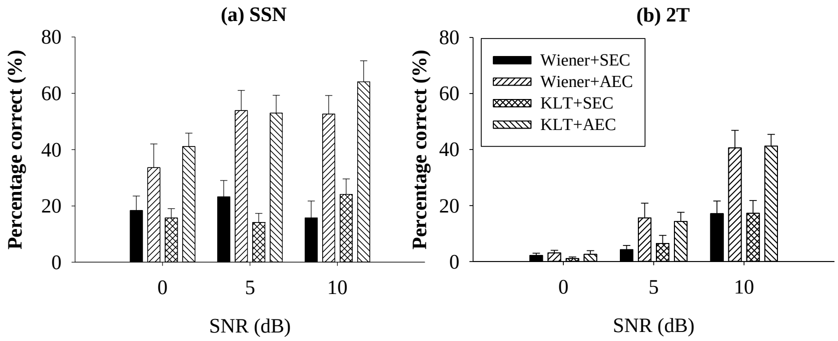

Fig. 13 Mean recognition scores for Mandarin sentences with (a) an SSN masker and (b) a 2T masker, at SNR levels of 0, 5, and 10 dB. The error bars indicate SEM values.

Table 2 The mean recognition scores for different strategies, where three factors (types of masker [masker], SNR levels [SNR], and processing method [F1]) were included in three-way ANOVA and Tukeys post-hoc testing.

Source of variance	Type III sum of squares	*df*	Mean square	*F*	*p*	Post-hoc comparison* (group$_i$, group$_j$)
Corrected model	65113.9[a]	17	3830.2	20.03	<0.001	
Intercept	110544.0	1	110544.0	578.15	<0.001	
SNR × F1	3800.57	6	633.43	3.31	0.004	
masker × F1	4892.81	3	1630.94	8.53	<0.001	
masker × SNR	2508.07	2	1254.04	6.56	0.002	
masker	19784.38	1	19784.38	103.47	<0.001	
SNR	12065.76	2	6032.88	31.55	<0.001	
F1	22062.35	3	7354.12	38.46	<0.001	(1,2), (1,4), (2,3), (3,4)
Error	33269.05	174	191.20			
Total	208927.00	192				
Corrected total	98382.99	191				

F1 group variable: 1, Wiener+SEC; 2, Wiener+AEC; 3, KLT+SEC; 4, KLT+AEC.

Dependent variable: speech intelligibility scores.

[a] $R^2 = 0.669$ (adjusted $R^2 = 0.624$)

formance of the SEC strategy compared to the AEC strategy in noisy conditions may be similar to that applicable in hearing aids [47, 48], where the static compression processing will increase the low-level noise during the pauses in the speech signal and thereby decrease the SNR performance for the NR algorithm. Therefore, the SEC strategy did not benefit as much from the NR algorithm as the AEC strategy. Moreover, the results in Table 2 show that different NR algorithms, such as the Wiener filter and KLT, did not produce significant differences when integrated with the same compression strategy (i.e., SEC or AEC). In contrast, different compression strategies produced significantly different results when integrated with the same NR algorithm, where the AEC strategy consistently outperformed the SEC strategy. NR algorithms probably did not yield significant differences due to the very narrow electrical DR (5 dB) used for CIs. Since noisy signals processed by an NR algorithm will generate speech with increased DR, a smaller SF has to be used to ensure that the sound signals remain within the audible range (i.e., between UB and LB). From the examples of Fig. 5 (c) and Fig. 12 (c), the initial SF, α_0, was larger without than with NR, under the same testing conditions, and the speech intelligibility scores were higher for a larger SF. This provides further evidence that the compression strategy is more important than the NR method in noisy conditions for compressed speech perception.

In summary, this experiment investigated the performance of the AEC strategy when integrated with different NR algorithms. The results indicate that integration of NR algorithms with the AEC strategy provided better speech intelligibility than with the SEC strategy, implying that the AEC strategy, integrated with NR methods, is useful for further improving the speech perception performance in CI recipients.

6 Summary

A non-linear, compressive mapping function is normally used in CI devices to convert an acoustic amplitude envelope to an electric current signal (with a narrow DR). The present chapter assessed the performance of compression strategies (i.e., static vs. adaptive) relative to that of CI speech processing by vocoded simulation. More specifically, we used a simple compression function, as shown in Eq. (3) and Eq. (6), called the AEC strategy, to compress the amplitude envelope into a preset DR. Note that most of the present acoustic-to-electric conversions in CI devices use a fixed mapping function. It is reasonable to foresee that the adaptive mapping (from acoustic to electric) function will improve the speech comprehension of implanted patients. In addition, the signal processing in the AEC strategy is similar to that used by the SEC strategy, but is characterized by additional boundary calculations (for UB and LB). Furthermore, the AEC rules optimally and continuously adjust for the compression ratio on a frame-by-frame basis. Since these two additional units are rather simple, the computation load for the AEC strategy is reasonable when compared with the conventional SEC strategy. This enables for the practical implementation of the AEC strategy by means of microprocessors.

The results of these experiments showed that the amplitude envelope processed by the AEC strategy yielded significantly higher intelligibility scores for vocoded sentences in noisy and in reverberation conditions than when it was processed by the SEC strategy. Moreover, integration of NR methods with the AEC strategy outperformed integration of NR methods with the SEC strategy under noisy conditions. This makes the proposed AEC strategy a highly promising approach for the enhancement of speech comprehension in noisy conditions for listeners with CIs.

Acknowledgements This work was supported by the Ministry of Science and Technology of Taiwan under Project MOST 104-2221-E-001-026-MY2. This work was also supported by the National Natural Science Foundation of China (Grant No. 61571213). We thank Dr. Dao-Peng Chen of the Institute of Biomedical Sciences, Academia Sinica, for help with the statistical analysis.

References

1. NIDCD, "Cochlear implants," *NIH Publication*, vol. 116, no. 11-4798, 2013.
2. L. M. Friesen, R. V. Shannon, D. Baskent, and X. Wang, "Speech recognition in noise as a function of the number of spectral channels: comparison of acoustic hearing and cochlear implants," *The Journal of the Acoustical Society of America*, vol. 110, no. 2, pp. 1150–1163, 2001.
3. B. L. Fetterman and E. H. Domico, "Speech recognition in background noise of cochlear implant patients," *Otolaryngology–Head and Neck Surgery*, vol. 126, no. 3, pp. 257–263, 2002.
4. P. C. Loizou, "Introduction to cochlear implants," *Engineering in Medicine and Biology Magazine, IEEE*, vol. 18, no. 1, pp. 32–42, 1999.
5. F. G. Zeng, "Trends in cochlear implants," *Trends in Amplification*, vol. 8, no. 1, pp. 1–34, 2004.
6. P. C. Loizou, M. Dorman, and J. Fitzke, "The effect of reduced dynamic range on speech understanding: implications for patients with cochlear implants," *Ear and Hearing*, vol. 21, no. 1, pp. 25–31, 2000.
7. D. K. Eddington, W. Dobelle, D. Brackmann, M. Mladejovsky, and J. Parkin, "Auditory prostheses research with multiple channel intracochlear stimulation in man.," *The Annals of otology, rhinology, and laryngology*, vol. 87, no. 6 Pt 2, pp. 1–39, 1977.
8. F. G. Zeng and R. V. Shannon, "Loudness balance between electric and acoustic stimulation," *Hearing Research*, vol. 60, no. 2, pp. 231–235, 1992.
9. F. G. Zeng, G. Grant, J. Niparko, J. Galvin, R. Shannon, J. Opie, and P. Segel, "Speech dynamic range and its effect on cochlear implant performance," *The Journal of the Acoustical Society of America*, vol. 111, no. 1, pp. 377–386, 2002.
10. R. van Hoesel, M. Böhm, R. D. Battmer, J. Beckschebe, and T. Lenarz, "Amplitude-mapping effects on speech intelligibility with unilateral and bilateral cochlear implants," *Ear and Hearing*, vol. 26, no. 4, pp. 381–388, 2005.
11. F. Chen, L. L. Wong, J. Qiu, Y. Liu, B. Azimi, and Y. Hu, "The contribution of matched envelope dynamic range to the binaural benefits in simulated bilateral electric hearing," *Journal of Speech, Language, and Hearing Research*, vol. 56, no. 4, pp. 1166–1174, 2013.
12. Y. H. Lai, Y. Tsao, and F. Chen, "Effects of adaptation rate and noise suppression on the intelligibility of compressed-envelope based speech," *Plos One*, vol. 10, no. 7, p. e0133519, 2015.
13. Y. H. Lai, F. Chen, and Y. Tsao, "Effect of adaptive envelope compression in simulated electric hearing in reverberation," in *2014 14th International Symposium on Integrated Circuits (ISIC)*, pp. 204–207, IEEE, 2014.
14. Y. H. Lai, F. Chen, and Y. Tsao, "An adaptive envelope compression strategy for speech processing in cochlear implants.," in *Interspeech*, pp. 481–484, 2014.
15. 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, no. 8, pp. 671–683, 2013.
16. R. S. Tyler, S. Waltzman, and S. Bankoski, *Cochlear implants: Audiological foundations*. Singular Publishing Group, 1993.
17. P. P. Khing, B. A. Swanson, and E. Ambikairajah, "The effect of automatic gain control structure and release time on cochlear implant speech intelligibility," *Plos One*, vol. 8, no. 11, p. e82263, 2013.

18. F. G. Zeng and J. J. Galvin III, “Amplitude mapping and phoneme recognition in cochlear implant listeners,” *Ear and Hearing*, vol. 20, no. 1, pp. 60–74, 1999.
19. K. Kasturi and P. C. Loizou, “Use of s-shaped input-output functions for noise suppression in cochlear implants,” *Ear and Hearing*, vol. 28, no. 3, pp. 402–411, 2007.
20. A. Boothroyd, F. N. Erickson, and L. Medwetsky, “The hearing aid input: a phonemic approach to assessing the spectral distribution of speech.,” *Ear and Hearing*, vol. 15, no. 6, pp. 432–442, 1994.
21. C. J. James, P. J. Blamey, L. Martin, B. Swanson, Y. Just, and D. Macfarlane, “Adaptive dynamic range optimization for cochlear implants: a preliminary study,” *Ear and Hearing*, vol. 23, no. 1, pp. 49S–58S, 2002.
22. R. V. Shannon, F. G. Zeng, V. Kamath, J. Wygonski, and M. Ekelid, “Speech recognition with primarily temporal cues,” *Science*, vol. 270, no. 5234, pp. 303–304, 1995.
23. Q. J. Fu, R. V. Shannon, and X. Wang, “Effects of noise and spectral resolution on vowel and consonant recognition: Acoustic and electric hearing,” *The Journal of the Acoustical Society of America*, vol. 104, no. 6, pp. 3586–3596, 1998.
24. G. S. Stickney, F. G. Zeng, R. Litovsky, and P. Assmann, “Cochlear implant speech recognition with speech maskers,” *The Journal of the Acoustical Society of America*, vol. 116, no. 2, pp. 1081–1091, 2004.
25. M. F. Dorman, P. C. Loizou, and D. Rainey, “Simulating the effect of cochlear-implant electrode insertion depth on speech understanding,” *The Journal of the Acoustical Society of America*, vol. 102, no. 5, pp. 2993–2996, 1997.
26. M. F. Dorman, P. C. Loizou, and D. Rainey, “Speech intelligibility as a function of the number of channels of stimulation for signal processors using sine-wave and noise-band outputs,” *The Journal of the Acoustical Society of America*, vol. 102, no. 4, pp. 2403–2411, 1997.
27. 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, no. 2, pp. 70S–74S, 2007.
28. O. Hazrati, S. O. Sadjadi, P. C. Loizou, and J. H. Hansen, “Simultaneous suppression of noise and reverberation in cochlear implants using a ratio masking strategy,” *The Journal of the Acoustical Society of America*, vol. 134, no. 5, pp. 3759–3765, 2013.
29. O. Hazrati, J. Lee, and P. C. Loizou, “Blind binary masking for reverberation suppression in cochlear implants,” *The Journal of the Acoustical Society of America*, vol. 133, no. 3, pp. 1607–1614, 2013.
30. O. Hazrati and P. C. Loizou, “Reverberation suppression in cochlear implants using a blind channel-selection strategy,” *The Journal of the Acoustical Society of America*, vol. 133, no. 6, pp. 4188–4196, 2013.
31. T. Van den Bogaert, S. Doclo, J. Wouters, and M. Moonen, “Speech enhancement with multichannel wiener filter techniques in multimicrophone binaural hearing aids,” *The Journal of the Acoustical Society of America*, vol. 125, no. 1, pp. 360–371, 2009.
32. F. Chen, O. Hazrati, and P. C. Loizou, “Predicting the intelligibility of reverberant speech for cochlear implant listeners with a non-intrusive intelligibility measure,” *Biomedical Signal Processing and Control*, vol. 8, no. 3, pp. 311–314, 2013.
33. T. Venema, *Compression for clinicians*. Delmar Pub, 2006.
34. P. E. Souza, “Effects of compression on speech acoustics, intelligibility, and sound quality,” *Trends in Amplification*, vol. 6, no. 4, pp. 131–165, 2002.
35. A. C. Neuman, M. H. Bakke, C. Mackersie, S. Hellman, and H. Levitt, “Effect of release time in compression hearing aids: Paired-comparison judgments of quality,” *The Journal of the Acoustical Society of America*, vol. 98, no. 6, pp. 3182–3187, 1995.
36. A. C. Neuman, M. H. Bakke, C. Mackersie, S. Hellman, and H. Levitt, “The effect of compression ratio and release time on the categorical rating of sound quality,” *The Journal of the Acoustical Society of America*, vol. 103, no. 5, pp. 2273–2281, 1998.
37. M. Hansen, “Effects of multi-channel compression time constants on subjectively perceived sound quality and speech intelligibility,” *Ear and Hearing*, vol. 23, no. 4, pp. 369–380, 2002.
38. S. Gatehouse, G. Naylor, and C. Elberling, “Linear and nonlinear hearing aid fittings–1. patterns of benefit,” *International Journal of Audiology*, vol. 45, no. 3, pp. 130–152, 2006.

39. R. Van Hoesel and G. M. Clark, “Evaluation of a portable two-microphone adaptive beamforming speech processor with cochlear implant patients,” *The Journal of the Acoustical Society of America*, vol. 97, no. 4, pp. 2498–2503, 1995.
40. V. Hamacher, W. Doering, G. Mauer, H. Fleischmann, and J. Hennecke, “Evaluation of noise reduction systems for cochlear implant users in different acoustic environment.,” *Otology & Neurotology*, vol. 18, no. 6, pp. S46–S549, 1997.
41. J. Wouters and J. V. Berghe, “Speech recognition in noise for cochlear implantees with a two-microphone monaural adaptive noise reduction system,” *Ear and Hearing*, vol. 22, no. 5, pp. 420–430, 2001.
42. P. C. Loizou, A. Lobo, and Y. Hu, “Subspace algorithms for noise reduction in cochlear implants,” *The Journal of the Acoustical Society of America*, vol. 118, no. 5, pp. 2791–2793, 2005.
43. K. Chung, “Challenges and recent developments in hearing aids part i. speech understanding in noise, microphone technologies and noise reduction algorithms,” *Trends in Amplification*, vol. 8, no. 3, pp. 83–124, 2004.
44. F. Chen, Y. Hu, and M. Yuan, “Evaluation of noise reduction methods for sentence recognition by mandarin-speaking cochlear implant listeners,” *Ear and Hearing*, vol. 36, no. 1, pp. 61–71, 2015.
45. P. Scalart *et al.*, “Speech enhancement based on a priori signal to noise estimation,” in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 2, pp. 629–632, IEEE, 1996.
46. Y. Hu and P. C. Loizou, “A generalized subspace approach for enhancing speech corrupted by colored noise,” *IEEE Transactions on Speech and Audio Processing*, vol. 11, no. 4, pp. 334–341, 2003.
47. 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, no. 2, pp. 43–51, 2013.
48. 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, no. 3, pp. 161–171, 2009.