

Low Complexity Formant Estimation Adaptive Feedback Cancellation for Hearing Aids Using Pitch Based Processing

Yi FanChiang, Cheng-Wen Wei, Yi-Le Meng, Yu-Wen Lin, Shyh-Jye Jou, *Senior Member, IEEE*, and Tian-Sheuan Chang, *Senior Member, IEEE*

Abstract—This paper proposes a novel algorithm and architecture for the adaptive feedback cancellation (AFC) based on the pitch and the formant information for hearing aid (HA) applications. The proposed method, named as Pitch based Formant Estimation (PFE-AFC), has significantly low complexity compared to Prediction Error Method AFC (PEM-AFC). The proposed PFE-AFC consists of a forward and a backward path processing. The forward path processing includes a low complexity pitch based formant estimator for decorrelation filter coefficients update and a pitch based voice activity detector for speech detection, which facilitates the feedback cancellation filter in the backward path to reduce feedback component and maintain speech quality. From system point of view, the PFE-AFC has low complexity overhead since it is easy to share computation resource with other components in the HA system, such as noise reduction and auditory compensation. In addition, the PFE-AFC is suitable for hardware implementation owing to its regular structure. Complexity evaluations show that the PFE-AFC has four orders lower complexity than the PEM-AFC. Simulation results show that the PFE-AFC and the PEM-AFC can achieve similar PESQ (perceptual evaluation speech quality) and ASG (added stable gain). Moreover, the proposed PFE-AFC can outperform the conventional AFC.

Index Terms—Adaptive feedback cancellation, hearing aid application, voice activity detection.

I. INTRODUCTION

ACOUSTIC feedback, mainly caused by the closed loop resulted from the venting of hearing aids (HAs), yields annoying howling (or oscillation) and greatly reduces the willingness to use HAs. To solve this problem, many solutions have been proposed [1][2], such as subband gain suppression, frequency or phase modulation control, and adaptive feedback cancellation (AFC). The subband gain suppression reduces the gain of subbands where oscillations are detected, hence immediately eliminating howling. However, reducing gain usually

degrades the auditory prescription and speech quality may be downgraded as well. Frequency and phase modulation control adjust the frequency or phase of signal in the subbands where oscillations are present, thereby avoiding oscillation caused by the gain or phase margin of the Nyquist stability criterion [3]. This method can provide better performance compared to gain suppression methods. However, it also degrades speech quality, since the signal phase and magnitude spectrum are changed. In addition to the degradation of prescription and quality, both the subband gain suppression and frequency or phase modulation control may start their operations after oscillation is large enough to be detected. Therefore, these methods are not able to avoid the annoying feeling for HA users.

In contrast to above methods, the AFC tries to prevent system oscillation by using an adaptive algorithm to mimic feedback component and then subtract it from the input signal. The AFC cancels feedback without directly modifying subband gain and signal spectrum. Therefore, the AFC methods, according to [2], can provide better performance and speech quality compared to subband gain suppression and frequency or phase modulation methods. However, the AFC in HA system confronts bias problem. Tonal and speech signal lead to correlation between microphone and speaker signal [2][4]. The bias problem causes the adaptive filter tend to minimize tone or speech input rather than feedback components resulted from feedback channel.

To eliminate the bias problem, several methods have been proposed. For instance, a constraint adaptation [5] is proposed to directly constraint feedback filter such that its transfer function is not updated too far away from a pre-defined one. This method is effective for system robustness; however, this method is not able to deal with the conditions if feedback channel is changed a lot, such as a phone handset is placed near the aided ear. Thus, a more effective solution, named as filtered-x method [6][7] which is shown in Fig. 1, is proposed to remove speech component in error signal before adaptive updating by using a pre-defined bandstop filter $A[k]$. The $F[k]$ and $\hat{F}[k]$ are feedback channel and feedback cancellation filter in the backward path of HA, respectively. Both the $F[k]$ and $\hat{F}[k]$ are modeled as finite impulse response filters. The forward path in HA is modeled as a frequency dependent gain $G[k]$. The transfer function $A[n]$ generally adopts the inverse of long-term average speech spectrum, hence being called as decorrelation filter. However, it is reasonable that the decorrelation filter should be adaptively adjusted according to speech variation. Therefore, a prediction error method based AFC (PEM-AFC), as shown in Fig. 2, is

Manuscript received October 06, 2013; revised January 29, 2014; accepted May 21, 2014. Date of publication May 29, 2014; date of current version June 25, 2014. This work was supported by the National Chip Implementation Center, and National Science Council, R.O.C., under Grant NSC101-2220-E-009-004. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. DeLiang Wang.

The authors are with the Department of Electronics Engineering and Institute of Electronics, National Chiao Tung University, Hsinchu 30050, Taiwan (e-mail: d9711829.ee97g@nctu.edu.tw; jackiewei.ee95g@nctu.edu.tw).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TASLP.2014.2327300

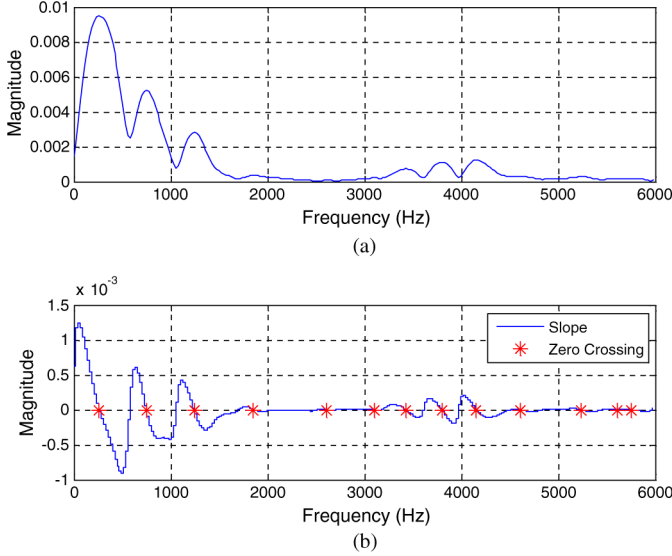


Fig. 4. (a) Magnitude of speech spectrum and (b) the slope of (a).

being the main part of HA, includes six blocks which can be divided into two groups. The first group is a pitch computation processing (marked by dotted line) which is proposed for speech discrimination and the pitch information calculation. The pitch information is used by the decorrelation filter coefficients update (in the backward path). The second group can be viewed as conventional blocks in HA and is designed to reduce background noise and provide the hearing prescription. The second group includes frequency decomposition and reconstruction, acoustic noise reduction and auditory compensation. This group may enhance speech based on the pitch information provided by the first group, or adopt conventional methods [11]–[13] for noise suppression and speech intelligibility improvement.

In this paper, we use 256 data sample as a frame with zero padding technique to perform length-1024 fast Fourier transform (FFT) to provide 23 Hz frequency resolution at 24 kHz sampling rate. The using of 256 data sample as a frame has 10.67 ms time delay and can meet HA application. 50% overlap between two successive FFTs is exploited for the continuity of speech signal. The 23 Hz resolution is derived from the minimum resolution of ANSI S1.11 filter bank [14].

B. The Peak and Harmonic Detection in the Forward Path

The peak and harmonic detection are used to estimate the speech spectrum for the decorrelation filter and to provide speech information for the pitch based VAD. For the peak detection in the forward path, an efficient slope based method is proposed and is illustrated in Fig. 4. Fig. 4(a) depicts an example of magnitude spectrum for speech signal, and Fig. 4(b) draws its slope and corresponding zero crossings which mark spectral peaks. For clarity, only the bandwidth from 0 Hz to 6 kHz is plotted. To find peaks, slope $Slp[k]$ between each two samples can be calculated as:

$$Slp[k] = M[k+1] - M[k] \quad k = 0, 1, 2, \dots, N/2 - 2 \quad (1)$$

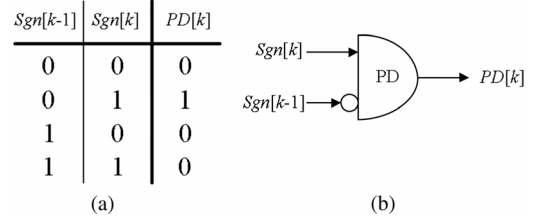


Fig. 5. Peak detection, (a) truth table and (b) corresponding PD circuit.

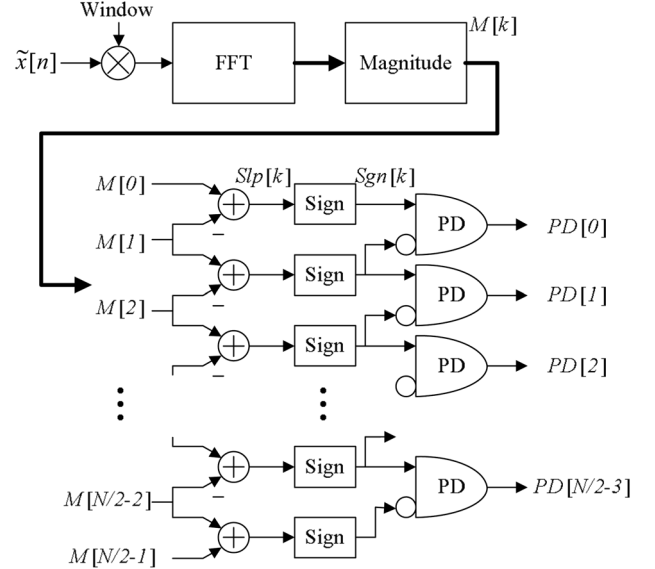


Fig. 6. The circuit design of peak detection.

where k is the frequency index, $M[k]$ is the magnitude response of input signal $\tilde{x}[n]$. N is FFT length and is set to 1024 in our case.

With $Slp[k]$, the peaks can be detected by simply tracking $Slp[k]$ along frequency axis from 0 Hz to 12 kHz and finding the locations where they change from positive value to negative value. This method can be simplified to detect the change of slope's sign bit in a 2's complement number system. Fig. 5 shows its truth table and corresponding detection circuit, named as peak detection (PD). The PD is simply composed by a NOT gate and an OR gate shown as (2).

$$PD[k] = Sgn[k+1] \bullet (\overline{Sgn[k]}) \quad k = 0, 1, 2, \dots, N/2-3 \quad (2)$$

where $Sgn[k]$ is the sign bit of $Slp[k]$. For the length- N FFT, only $N/2-2$ PDs are required as illustrated in Fig. 6 owing to the symmetry feature of magnitude spectrum. Peaks can be found by detecting which the PD outputs are high, as shown in (3).

$$\mathbf{P}_{PD}[n] = \arg[PD[k] == 1] \quad n = N/2, N, 3N/2, \dots \quad (3)$$

where $\mathbf{P}_{PD}[n]$ is a vector containing the frequency index k of all detected peaks at time n . $N/2$ means 50% overlap of length N . $\mathbf{P}_{PD}[n]$ is computed once for each frame. Fig. 7 illustrates an example of $\mathbf{P}_{PD}[n]$ result for Chinese 2-characters "Shi-Yong." For clarity, only the bandwidth from 0 Hz to 6 kHz is drawn.

The peak detection circuit finds all peaks in magnitude spectrum. However, some of detected peaks are not the harmonics,

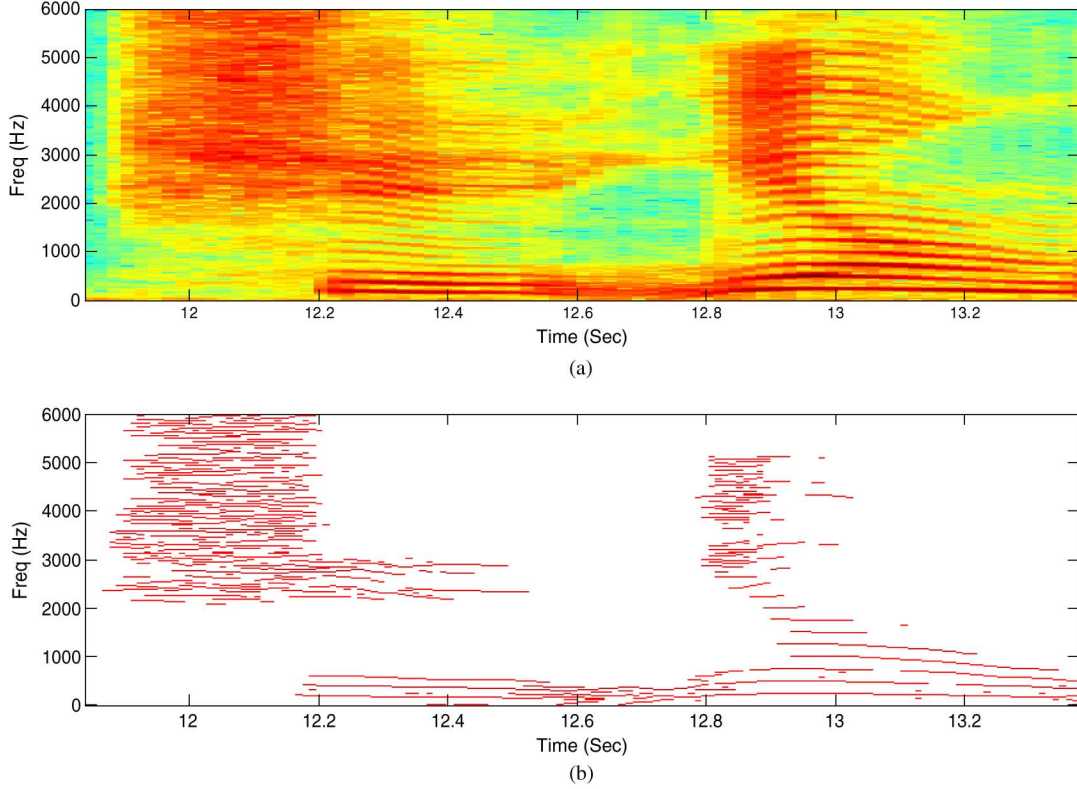


Fig. 7. An example of peak detection, (a) the spectrogram of input signal which is a Chinese 2-characters “Shi-Yong” and (b) the corresponding result of peak detection, $PD[k]$.

especially in high frequency. A two-steps method is proposed to select the appropriate harmonics. The first step selects candidate peaks for speech harmonics based on an adaptive threshold as below:

$$Thr[n] = \max(E(M[\mathbf{P}_{PD}[n]]), B) \quad n = N/2, N, 3N/2, \dots \quad (4)$$

where $M[\mathbf{P}_{PD}[n]]$ is the magnitude of each element in $\mathbf{P}_{PD}[n]$, $E(\cdot)$ is the expectation (mean) operator. The lower bound B is an empirical constant and is set to 0.003 in our case. The lower bound scheme is designed for speech absent condition. Using $Thr[n]$, the candidate of peaks can be detected as below.

$$\mathbf{P}_{Cand}[n] = \arg_{k \in \mathbf{P}_{PD}[n]} [M[\mathbf{P}_{PD}[n]] > Thr[n]] \quad n = N/2, N, 3N/2, \dots \quad (5)$$

where $\mathbf{P}_{Cand}[n]$ contains the frequency index k at time n for all peaks whose magnitude are greater than $Thr[n]$. Fig. 8 depicts the detection results of Fig. 4, where $Thr[n]$ and the detected peaks are marked by dash line and circle, respectively.

Based on the candidate of peaks, the second step discriminates vowel harmonics by exploiting two features, namely harmonic structure and frequency modulation (FM). The harmonic feature indicates the vowel energy is primarily distributed at frequency f_0 and its multiples as below:

$$F = \{f_0, 2f_0, 3f_0, \dots\} \quad (6)$$

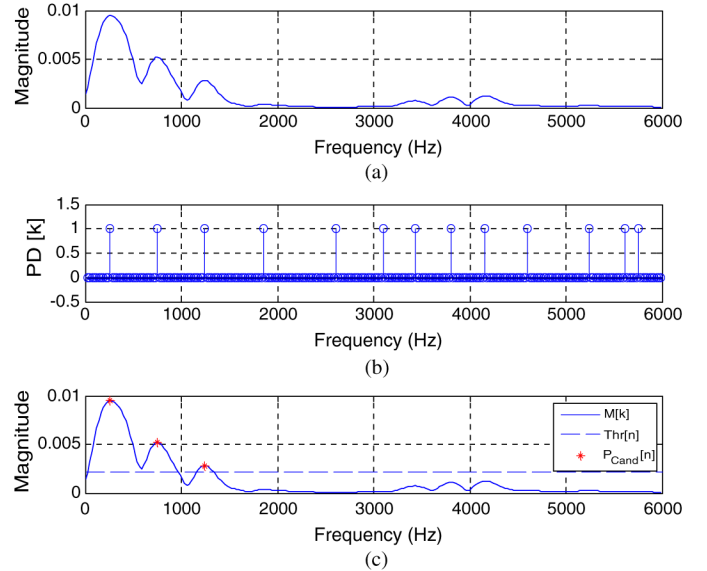


Fig. 8. Detection results of Fig. 4, (a) magnitude of speech spectrum, (b) $PD[k]$ and (c) candidates of peaks.

where f_0 is the fundamental frequency. Therefore, vowel can be detected if three candidates satisfy the harmonic features defined in (6). With respect to discrete frequency domain, (6) can be modified as (7):

$$K = \{k_0, 2k_0 \pm B_1, 3k_0 \pm B_2\} \quad (7)$$

where B_1 and B_2 are integer constants. B_1 and B_2 are designed to provide tolerance range for discrete frequency, since

the fundamental frequency, second harmonics, and third harmonics may not exactly be located at k_0 and its multiples. Here, B_1 and B_2 are set to 1 and 3, respectively. The detected peak index to fulfill (7) is written in (8).

$$\mathbf{P}_{\text{Harmo}}[n] = \{\text{Pitch}_0[n], \text{Pitch}_1[n], \text{Pitch}_2[n]\}$$

$$n = N/2, N, 3N/2, \dots \quad (8)$$

$$\text{Pitch}_0[n] = \arg_{k \in \mathbf{P}_{\text{PD}}[n]} [\mathbf{P}_{\text{PD}}[n] == k_0]$$

$$\text{Pitch}_1[n] = \arg_{k \in \mathbf{P}_{\text{PD}}[n]} [\mathbf{P}_{\text{PD}}[n] \in 2k_0 \pm B_1]$$

$$\text{Pitch}_2[n] = \arg_{k \in \mathbf{P}_{\text{PD}}[n]} [\mathbf{P}_{\text{PD}}[n] \in 3k_0 \pm B_2] \quad (9)$$

where $\text{Pitch}_0[n]$, $\text{Pitch}_1[n]$ and $\text{Pitch}_2[n]$ contain the harmonic index k at time n . The harmonic features described above can provide a good speech discrimination [10]–[16]. However, for several cases there may be more than one candidate peak in the neighborhood of f_0 , $2f_0$ and $3f_0$. Therefore, a tracking scheme based on the FM feature of speech harmonics is adopted to improve the discrimination accuracy. The FM features indicate that the vowel harmonics of spectrogram prolong and slowly change, as shown in Fig. 7. Therefore, the tracking scheme treats peaks as speech harmonics if it meets prolonging and difference requirement shown in (10). In which, the harmonic peak index k in current frame is limited to less than one index difference from the harmonic index in the previous frames.

$$\text{Pitch}'_m[n] = \arg_{k \in \text{Pitch}_m[n]} \left[\left| \text{Pitch}_m[n] - \text{Pitch}_m \left[n - \frac{N}{2} \right] \right| \leq 1 \right]$$

$$m \in \{0, 1, 2\} \quad (10)$$

For the case that at least two peaks meet (10), the detector chooses the smallest k for $\text{Pitch}'_m[n]$. Hence, the final detected result $\mathbf{P}_{\text{Final}}[n]$ is written as:

$$\mathbf{P}_{\text{Final}}[n] = \{\text{Pitch}'_0[n], \text{Pitch}'_1[n], \text{Pitch}'_2[n]\}$$

$$n = N/2, N, 3N/2, \dots \quad (11)$$

The detected frequency index in (11) and corresponding magnitude $M[\mathbf{P}_{\text{Final}}[n]]$ are exploited in the PFE-AFC, or in the noise reduction and the auditory compensation for further processing, as depicted in Fig. 3(b). The peak detection also leads to an efficient VAD algorithm, which will be described in the pitch based VAD subsection.

C. The Decorrelation Filter Coefficients Update

Fig. 9 depicts the proposed algorithm in the forward path for the coefficients update of the decorrelation filter. For low computational complexity, the proposed method exploits the results of $M[k]$ in (1) and $\mathbf{P}_{\text{Final}}[n]$ in (11). In addition, we propose an efficient algorithm by using the segmental average and constant multiplications instead of conventionally adopted linear prediction or RLS algorithm. The proposed method can be divided into two steps, inverse formant estimator and frequency sampling filter design.

The inverse formant estimator is proposed to estimate speech energy distribution based on the segmental average, as shown

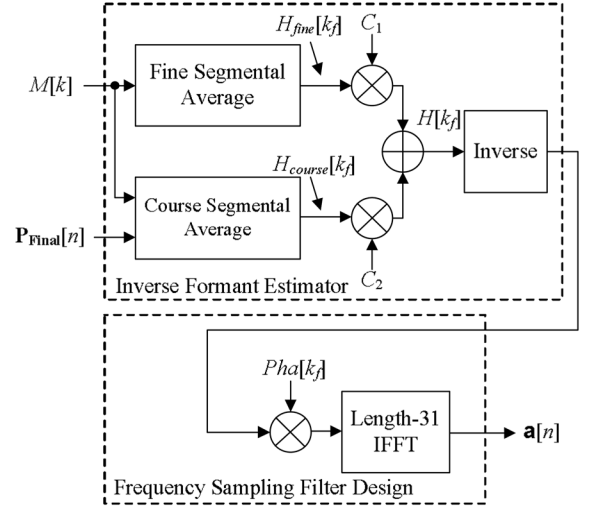


Fig. 9. The block diagram of the proposed decorrelation filter coefficients update in the forward path.

in Fig. 9. The upper part is a fine segmental average, as shown in (12), to estimate the shape of speech formant.

$$H_{fine}[k_f] = \begin{cases} \frac{1}{16} \sum_{l=0}^{15} M[l] & k = 0 \\ \frac{1}{32} \sum_{l=k-16}^{k+15} M[l] & k = 32, 64, \dots, 480 \end{cases} \quad (12)$$

where k_f is a frequency index and is set to $k/32$. $H_{fine}[k_f]$ can be viewed as a decimation-by-32 version of magnitude spectrum $M[k]$. The decimation ratio is adopted empirically to provide compromise between the speech quality and the computational complexity. Furthermore, for real-valued data, only the computations from 0 Hz to half sampling rate are needed to be carried out owing to the even symmetry of magnitude spectrum. Hence, only 16 samples ($k = 0, 32, 64, \dots, 480$) are calculated for each operation.

Although $H_{fine}[k_f]$ can provide an estimation of the formant shape, the formant estimation can be improved by using a nonuniform coarse segmental average, as depicted in the lower path of Figs. 9 and 10. The coarse segmental average, as shown in (13), adopts the similar average scheme as the fine segment average except that the coarse segment average only uses $M[\mathbf{P}_{\text{Final}}[n]]$ for average calculation.

$$H_{coarse}[k_c] = \sum M[\mathbf{P}_{\text{Final}}[n] \in B_c[k_c]] \quad k_c=0, 1, 2, 3$$

$$B_c[k_c] = \begin{cases} \{0 \text{ Hz} \sim 2 \text{ kHz}\} & \text{if } k_c == 0 \\ \{2 \text{ kHz} \sim 4 \text{ kHz}\} & \text{if } k_c == 1 \\ \{4 \text{ kHz} \sim 8 \text{ kHz}\} & \text{if } k_c == 2 \\ \{8 \text{ kHz} \sim 12 \text{ kHz}\} & \text{if } k_c == 3 \end{cases} \quad (13)$$

where $B_c[k_c]$ is designed based on a nonuniform scheme which has two 2 kHz segments ($B_c[0]$ and $B_c[1]$) in low frequency and two 4 kHz segments ($B_c[2]$ and $B_c[3]$) in high frequency. The nonuniform scheme is used to achieve an efficient estimation, since speech energy mainly distributes in low frequency. For high frequency, there is almost no speech harmonic, hence making the coarse segment average in those bandwidths tend to zero.

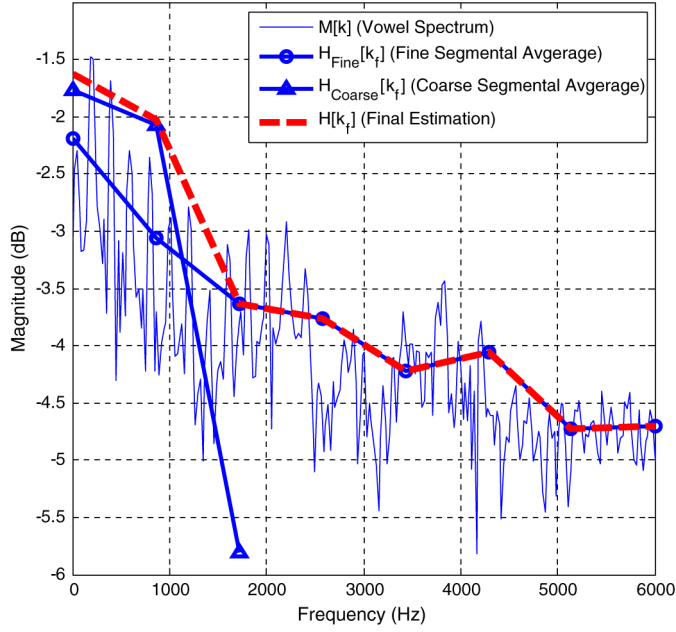


Fig. 10. An example for vowel spectrum and corresponding estimation.

$H_{coarse}[k_c]$ is interpolated to 16 uniform samples by repeating $H_{coarse}[k_c]$ to obtain $H_{coarse}[k_f]$ and then linearly combined with $H_{fine}[k_f]$ to generate the estimated formant $H[k_f]$ as (14).

$$H[k_f] = C_1 H_{fine}[k_f] + C_2 H_{coarse}[k_f] \quad (14)$$

where C_1 and C_2 are constants. Then, $H[k_f]$ is used to generate the inverse formant estimation for the second step, as follows.

$$A[k_f] = \left\{ \frac{1}{H[0]}, \frac{1}{H[1]}, \dots, \frac{1}{H[15]}, \frac{1}{H[15]}, \frac{1}{H[14]}, \dots, \frac{1}{H[1]} \right\} \quad (15)$$

The second step is the frequency sampling filter design which computes filter coefficients based on IFFT. The frequency sampling method [17][18] consists of two blocks, a phase multiplication and a length-31 IFFT as shown in (16).

$$\mathbf{a}[n] = IFFT(A[k_f] \times Pha[k_f]) \quad (16)$$

where $Pha[k_f]$ is a constant phase defined as (17).

$$Pha[k_f] = \exp\left(-j \frac{2\pi \times k_f}{31}\right) \quad k_f = 0, 1, \dots, 30 \quad (17)$$

The length of IFFT is fixed and thus can be implemented with the Winograd [19] technique, which consumes 160 multiplications [20] for each frame. The phase multiplication costs 30 multiplications for each frame because $A[k_f]$ is even symmetry and $Pha[k_f]$ is complex conjugate. Since all of those multiplications are constant operations, several low power techniques, such as shift-and-add transformation [21], are easy to be applied to reduce the power consumption. The low power methods adopted by [14][22] are also good choices to significantly reduce power consumption.

Fig. 11 depicts the circuit design of the decorrelation filter coefficients update, where the decorrelation filter is plotted to show entire architecture. ‘‘Average’’ blocks consist of simple

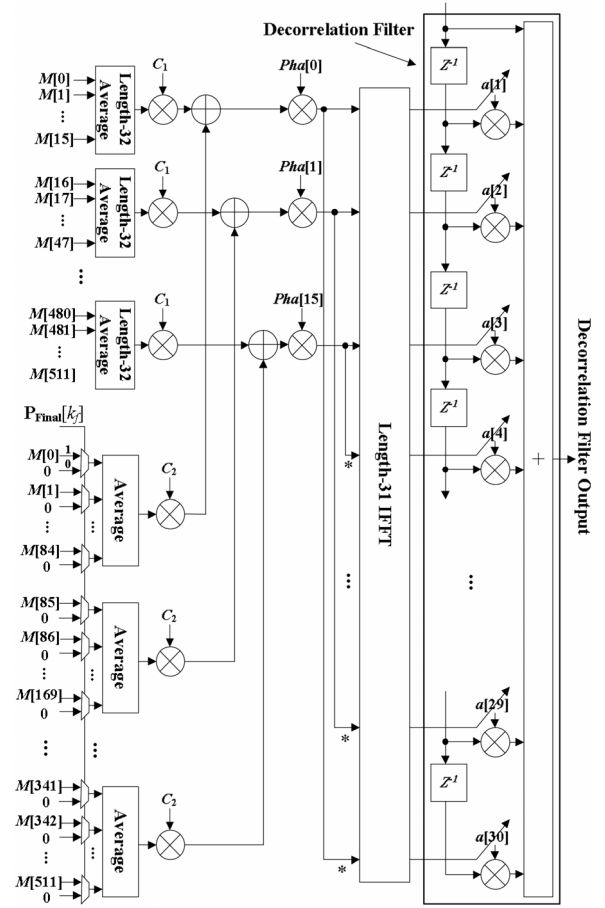


Fig. 11. The circuit design of the proposed decorrelation filter coefficients update, and the decorrelation filter, where asterisk means conjugate.

shifters and accumulators. The accumulators as well as the decorrelation filter can be implemented by using the carry save adder [21] for high speed and regular architecture. Therefore, the proposed decorrelation filter algorithm has a regular architecture and thus is suitable for hardware implementation.

D. The Adaptive Feedback Cancellation in Backward Path

Estimated coefficients $\mathbf{a}[n]$ of the decorrelation filter is exploited in the backward path for speech filtering, to aid feedback cancellation filter to be more insensitive to input speech energy. The feedback cancellation filter $\hat{F}[k]$ adopts conventional normalized least mean square (NLMS) method as shown in (18) for feedback channel approximation.

$$\hat{\mathbf{f}}[n+1] = \hat{\mathbf{f}}[n] + \frac{2\mu}{\mathbf{y}_d^T[n]\mathbf{y}_d[n]} e_d[n]\mathbf{y}_d[n] \quad (18)$$

where $\hat{\mathbf{f}}[n]$ is the coefficients of feedback cancellation filter, μ is the step size, and $e_d[n]$ is the error signal which is generated by filtering feedback cleaned input $\tilde{x}[n]$ through decorrelation filter $\mathbf{a}[n]$ as (19):

$$e_d[n] = \{\tilde{x}[n - N_a + 1], \tilde{x}[n - N_a + 2], \dots, \tilde{x}[n]\} * \mathbf{a}[n] \quad (19)$$

where $*$ is a convolution operator. $\tilde{x}[n]$ is shown as below:

$$\tilde{x}[n] = \hat{x}[n] - \hat{x}_f[n] = \hat{x}[n] - \mathbf{y}[n] * \hat{\mathbf{f}}[n] \quad (20)$$

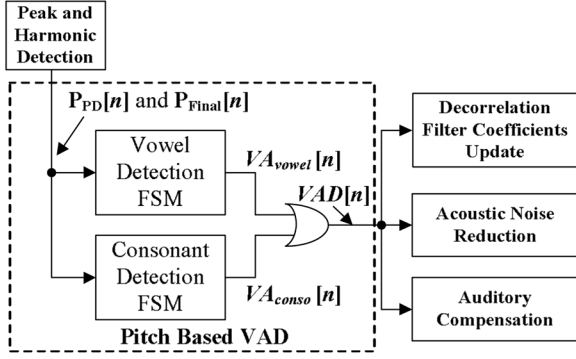


Fig. 12. The block diagram of pitch based VAD.

$y_d[n]$ is the result of $y[n]$ filtered by $a[n]$ as below:

$$y_d[n] = y[n] * a[n] \quad (21)$$

Starting from system initialization, the feedback cancellation continuously operates no matter it is speech or noise segment. To further improve speech quality, a VAD-aid- μ -control, shown in (22), is adopted for speech frames.

$$\mu = \begin{cases} \mu_0 \times ATT & \text{if } VAD[n] == 1 \\ \mu_0 & \text{if } VAD[n] == 0 \end{cases} \quad (22)$$

where μ_0 is a preset constant and is set to 0.001. ATT is set to 0.25 and can be implemented by a 2-bits shifter. It is an attenuation gain to slow down update speed if speech is present. $VAD[n]$ is the VAD detection result which will be described in the next subsection.

E. Pitch Based VAD in Forward Path

VAD is an important block in HA system, since it facilitates the operation of noise reduction and speech enhancement [28]. The proposed PFE-AFC also utilizes the VAD results to improve speech quality as shown in (20). The proposed pitch based VAD detects voice activity based on peak detection results. The pitch based VAD, as shown in Fig. 12, can be divided into two parts, namely vowel detection and consonant detection.

The vowel detection discriminates speech vowel by detecting harmonic structure based on (1)-(11). Using those results, a finite state machine (FSM), depicted in Fig. 13, is designed to perform vowel detection. The FSM denotes signal frame as four states, namely, silence, semi-vowel, vowel and hold state. If no tone is detected, the frame is denoted as silence frame. The silence state lasts until speech harmonic structure is detected. However, the VAD treats those frames having two tones harmonic structure (one fundamental and one harmonic tone) as semi-vowel frames. For the PFE-AFC, the semi-vowel is also treated as the vowel and thus VAD-aid- μ -control is enabled. During semi-vowel frames, the VAD does not denote the frame as the vowel state until the harmonic requirements described in (8) and (9) are confirmed. Additionally, in the vowel state, harmonic tracking presented in (10) and (11) is also utilized to assist the vowel detection. The FSM departs from the vowel state, if (8)-(11) are not confirmed. Since several frames next to vowel frames may still contain weak speech energy without satisfying (8)-(11), a hold state is designed for those frames to improve

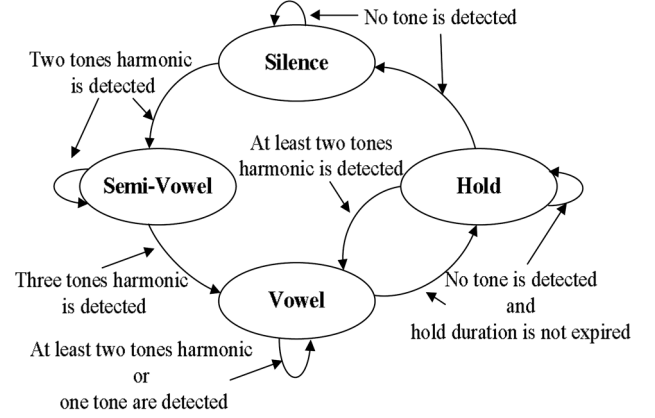


Fig. 13. The FSM of vowel detection.

vowel detection. During the hold frames, the VAD continuously searches for harmonic structure. If (8) and (9) are confirmed again, the FSM directly enters the vowel state without entering the semi-vowel. On the other hand, the FSM backs to the silence state if a preset duration for hold state is expired. For the PFE-AFC, the hold state is also viewed as vowel to activate the VAD-aid- μ -control for quality improvement. The summary of the vowel detector is shown in (23).

$$VA_{vowel}[n] = \begin{cases} 1 & \text{if (state == Semi-Vowel) or} \\ & \text{(state == Vowel) or (state == Hold)} \\ 0 & \text{if (state == Silence)} \end{cases} \quad (23)$$

The vowel detection cannot find consonant which does not have harmonic structures. Fortunately, the proposed peak detection can also achieve the consonant detection. As depicted in Fig. 7, the consonant between 11.9 s and 12.2 s has highly concentrated peaks in high frequency. By exploiting this feature, a consonant detector is proposed with consonant state and empty state. The consonant state is detected by checking if the number of peaks ($P_{PD}[n]$) above 3 kHz is larger than a preset constant. Otherwise, the state is denoted as empty state. The consonant detector is summarized as (24).

$$VA_{conso}[n] = \begin{cases} 1 & \text{if (state == Consonant)} \\ 0 & \text{if (state == Empty)} \end{cases} \quad (24)$$

$VA_{vowel}[n]$ and $VA_{conso}[n]$ are computed once every frame and are ORed to obtain a final result $VAD[n]$, as shown in Fig. 12.

III. COMPLEXITY AND PERFORMANCE EVALUATION

A. Complexity Comparison

Although power consumption may rely on implementation techniques, the computational complexity of the algorithm still plays an important role for power dissipation. This section provides the complexity of the proposed PFE-AFC in terms of the number of additions (ADD), multiplications (MUL) and divisions (DIV). In addition, the PEM-AFC using the RLS method [23] is selected for comparison because of its good performance. The comparison focuses on the complexity of decorrelation filter coefficients update, since we can assume that both the PFE-AFC and the PEM-AFC use the same AFC.

TABLE I
THE COMPLEXITY OF DECORRELATION FILTER COEFFICIENTS
UPDATE FOR THE PFE-AFC

Equations	MUL	DIV	ADD
Fine average (12)	0	0	$N/2-1$
Coarse average (13)	0	0	12
Linear combination (14)	0	0	16
Inverse formant estimation (15)	0	16	0
Frequency sampling method (Including IFFT) (16), (17)	160+30	0	984
Total	190	16	$N/2+1001$
Total (If each multiplication of IFFT and phase multiplication are substituted by eight shift-and-adds)	0	16	$N/2+2521$

Table I shows the complexity of decorrelation filter coefficients update for the PFE-AFC, where N is the FFT length and is 1024 in our case. The complexity in (16) is caused by phase multiplications (30 MULs) and length-31 Winograd IFFT (160 MULs and 984 ADDs) [20]. Those multiplications are constant operations, and thus are easy to be implemented by dedicated hardware to significantly reduce complexity. In the comparisons, we replace the complexity of each multiplication for IFFT and phase multiplication by eight shift-and-adds, assuming multiplier width is 16-bits. Furthermore, the binary DIV complexity is the same order as the binary MUL. Therefore, the total complexity is equivalent to 16 MULs and 3033 ADDs when N is 1024. Moreover, the decorrelation filter coefficients update is a frame based method and updates once for each frame. Hence, from the sample based point of view, its complexity can be divided by 128 and thus is much lower than that of the PEM-AFC.

Table II provides the complexity of the RLS algorithm for comparisons, where L_a is the length of decorrelation filter which is 21 in our case. In addition, the divisions by λ are not counted since they are constant operations and do not dominate complexity. The RLS can perform fast convergence at the expense of $O(L_a^2)$ complexity, which is much more than the $O(L_a)$ of NLMS adopted by feedback cancellation filter. Therefore, the RLS decorrelation filter dominates the computational complexity of the PEM-AFC. However, low complexity design for RLS is relatively difficult than that of the proposed method, as its huge number of non-constant multiplications.

Table III summaries the comparisons of total complexity, where N and L_a are substituted by 1024 and 21, respectively. The PFE-AFC significantly reduces the complexity of decorrelation filter coefficients update. This complexity is further reduced by 128 from the sample based point of view, because the PEM-AFC is a sample based algorithm. As a result, the decorrelation filter of the PFE-AFC has about four orders lower complexity compared to that of PEM-AFC.

B. Performance Evaluation

This section provides performance comparisons for speech quality and AFC performance. The speech quality is evaluated in terms of the PESQ [24]. PESQ is being used as an accepted standard, even though it shows weakness with tonal languages.

TABLE II
THE COMPLEXITY OF DECORRELATION FILTER COEFFICIENTS
UPDATE USING RLS METHOD

Equations	MUL	DIV	ADD
$\mathbf{\varepsilon}[n] = [\varepsilon[n] \ \varepsilon[n-1] \ \dots \ \varepsilon[n-L_a+1]]^T$	0	0	0
$e[n] = \varepsilon_p - \mathbf{\varepsilon}^T[n] \mathbf{a}[n-1]$	L_a	0	L_a
$\mathbf{g}[n] = \frac{\mathbf{R}^{-1}[n-1] \mathbf{e}^*[n]}{\{\lambda + \kappa_{prior} \mathbf{e}^*[n]\}}$ $\kappa_{prior} = \mathbf{e}^T[n] \mathbf{R}^{-1}[n-1]$	$L_a + 2L_a^2$	L_a	$2L_a^2 - 1$
$\mathbf{R}^{-1}[n] = \lambda^{-1} (\mathbf{R}^{-1}[n-1] - \mathbf{g}(n) \kappa_{prior})$	$2L_a^2$	0	L_a^2
$\mathbf{a}[n] = \mathbf{a}[n-1] + e[n] \mathbf{g}[n]$	L_a	0	L_a
Total	$4L_a^2 + 3L_a$	L_a	$3L_a^2 + 2L_a - 1$

TABLE III
THE COMPARISON OF COMPLEXITY FOR THE DECORRELATION FILTER
UPDATE OF THE PEM-AFC AND THE PFE-AFC

Methods	MUL	DIV	ADD
PEM-AFC (for each sample)	1827	21	1364
PFE-AFC (for each sample)	0	0.13	23.69
PFE-AFC (for each frame)	0	16	3033

The PESQ ranges from 1.0 to 4.5, where higher score implies a better perceptual quality. With respect to the AFC performance, added stable gain (ASG) [8] is applied for robust comparisons. Both of the adopted parameters are derived from the estimated feedback channel $\hat{F}[k]$ and real one $F[k]$. The ASG, defined as (25), is the maximum additional gain that HA can provide.

$$\text{ASG} = \text{MSG} - 20 \log_{10} \left(\min_k \frac{1}{|F[k]|} \right) \quad (25)$$

where maximum stable gain (MSG) is the maximum allowable gain of HA [25], if the HA's frequency response is assumed to be flat. The MSG is defined as follows.

$$\text{MSG} = 20 \log_{10} \left(\min_k \left(\frac{1}{|F[k] - \hat{F}[k]|} \right) \right) \quad (26)$$

The MSG can be viewed as the largest remainder of feedback component that the AFC cannot cancel, which decides the maximum achievable gain of HA. From another perspective, the MSG represents the minimum distance between 0 dB gain and the feedback cancellation remainder. HA gain larger than the MSG makes the loop gain exceed 0 dB and may thus cause oscillation. The $1/|F[k]|$ term in (25) can be treated as a special case of (26) when the AFC is not available. Therefore, the ASG represents the added gain provided by the AFC and is an important parameter for HA application.

In this paper, we adopt the feedback channel measured by [26] as channel model for simulations. The channel is modeled as a FIR filter with length of 75 at 24 kHz sampling rate. Fig. 14 depicts its frequency response. For comparisons, three AFC methods are adopted, the conventional AFC without

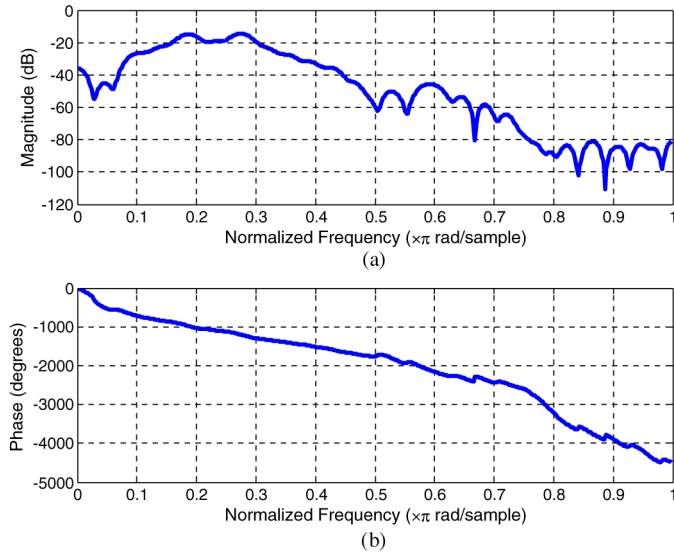


Fig. 14. The frequency response of feedback channel, (a) magnitude response and (b) phase response.

decorrelation filter (C-AFC), the PEM-AFC and the proposed PFE-AFC. The PEM-AFC uses the RLS for decorrelation filter coefficients updating [23]. With respect to the feedback cancellation filter update, all of the methods adopt NLMS method with the same setting for fair comparisons. In addition, many simulation parameters are from [8]. For example, the filter length of feedback cancellation filter is set to 75, which is the same as the feedback channel length. The decorrelation filter length for the PEM-AFC is 21, whereas the one for the PFE-AFC is 31. The forgetting factor λ of the RLS is set as $1-(1/240)$, whereas the step size μ of NLMS is set to 0.001.

1) *PESQ*: In simulations, 27 2-characters Chinese recording chosen from Academia Sinica Balanced Corpus is adopted [27]. The length of 27 2-characters Chinese is 46.12 seconds. The gender of the speaker is female. White noise, babble noise, car noise and factory noise provided by NOISEX-92 are used as noise source. The simulation results are evaluated by PESQ [28] at six SNRs (0 dB, 3 dB, 5 dB, 7 dB, 10 dB, clean speech) for comparisons.

Table IV shows comparisons of PESQ for three AFC methods. Simulation results show that the PEM-AFC and the PFE-AFC provide the better speech quality, while the C-AFC performs the worst speech quality. The C-AFC, without the aid of decorrelation filter, tends to attenuate input speech when speech is present, hence being easy to cause distortion. On the other hand, the proposed PFE-AFC reduces speech disturbance with decorrelation filter, thereby can achieve the PESQ score of 3.96. To further analyze the contribution of VAD control, the PESQ of PFE-AFC with perfect VAD and without VAD control is provided. It shows the PESQ of PFE-AFC without VAD control has slight degradation and the PFE-AFC is not sensitive to VAD performance due to the decorrelation filter. Moreover, the PESQ is decreased when the SNR of input signal is increased, which is caused by the more serious bias problem.

Fig. 15 and Fig. 16 show the simulation results of the input signal $x(t)$, which is the clean speech, and the output $y[n]$ of the PEM-AFC, the PFE-AFC and the C-AFC. The temporal

TABLE IV
THE PESQ OF FIVE AFC METHODS

SNR (dB)	Accuracy Rate (%)				
	White	Babble	Factory	Car	Average
0	85.05	57.55	71.17	83.25	74.26
3	85.92	70.15	81.57	86.05	80.92
5	86.69	80.80	84.90	86.29	84.67
7	86.64	83.35	85.92	86.46	85.59
10	86.99	85.92	86.13	87.07	86.53
Average	86.26	75.55	81.94	85.83	82.39

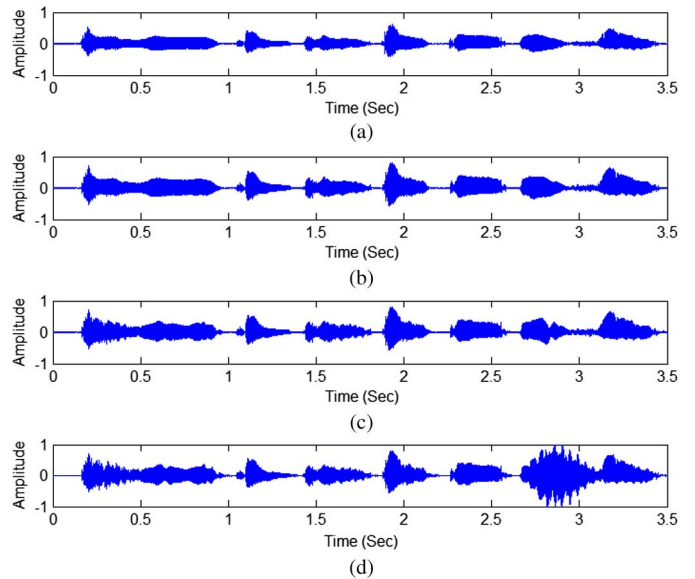


Fig. 15. The temporal waveform of (a) input speech $x(t)$ and the output signal $y[n]$ of (b) PEM-AFC, (c) PFE-AFC and (d) C-AFC.

waveform in Fig. 15 reveals that all methods cause speech distortion. However, the distortion caused by the C-AFC is much more obvious and serious. On the contrary, the proposed PFE-AFC and the PEM-AFC significantly reduce speech distortion by the decorrelation filtering. The spectrogram shown in Fig. 16 also reveals similar results. The PEM-AFC and the proposed PFE-AFC have similar output spectrogram with that of input. In contrast, the C-AFC spectrogram evidently illustrates difference between the others.

Listening to the processed sound provides another perspective for speech distortion. The speech processed by the PEM-AFC can provide good speech quality, while the PFE-AFC sometimes has small and high frequency artifacts in the beginning of each sentence. This is because the decorrelation filter coefficients update of the PFE-AFC is frame based and cannot immediately respond to the onset of each sentence. For the C-AFC, the artifacts become even worse since the feedback filter update is continuously impacted by speech signal. Thus, low frequency artifacts sometimes appear after high frequency ones occur. The low frequency artifacts cause smeared and obscure feeling which reduces speech quality. The artifacts can be used to prove that without the protection of

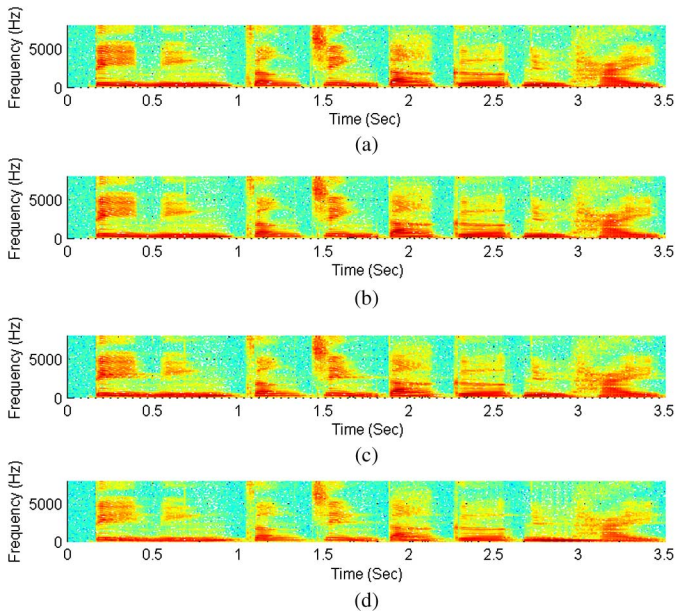


Fig. 16. The spectrogram of (a) input speech $x(t)$ and the output signal $y[n]$ of (b) PEM-AFC, (c) PFE-AFC and (d) C-AFC.

decorrelation filtering, the feedback cancellation filter tends to eliminate the input speech.

2) *ASG*: Fig. 17 shows the ASG simulations in SNR 0 dB white noise, 0 dB babble noise, and clean speech environments. In the beginning of the simulation, the feedback cancellation filters are set to zero. At the 23 seconds of the simulation, the feedback path gain is increased by 10 dB, which is to simulate the scenario of bringing a telephone handset to the aided ear, to investigate the convergence behavior. During 0 to 23 seconds, the feedback path is denoted as original feedback path. Therefore, the ASG progressively increases from 0 dB to its steady range, i.e. the PFE-AFC is about 11 dB for original feedback path in 0 dB white noise.

Fig. 17(a) shows the clean speech and the rest figures of Fig. 17 show the ASG results of the three methods in different scenarios. Generally, in the speech segment, the ASG gradually increases or decreases owing to the adaptive filter update. Among three AFC methods, the PEM-AFC and the PFE-AFC could perform the stable ASG and there are about 2 dB differences in white noise. In babble and clean speech environments, the PEM-AFC and the PFE-AFC achieve comparable results. However, the ASG of C-AFC has larger variation caused by speech disturbance. In addition to ASG variation, the maximum ASG is also an important system parameter, which is limited by the minimum value of each curve. For example, in Fig. 17(b), during the steady range in original feedback path, the PEM-AFC has the best ASG performance (about 13 dB), while the PFE-AFC has similar performance (about 11 dB). The C-AFC has the worst ASG performance (about 0.3 dB).

3) *VAD Accuracy*: In the VAD evaluations, the speech signals used in PESQ simulation are adopted. VAD accuracy in terms of percentage (%) is provided to show the VAD accuracy rate. Table V lists the simulation results of the VAD accuracy. The average accuracy of the proposed VAD is 74.26%, 84.67%, and 86.53% at SNR of 0 dB, 5 dB, and 10 dB, while the overall average VAD accuracy is 82.39%. General speaking, the VAD

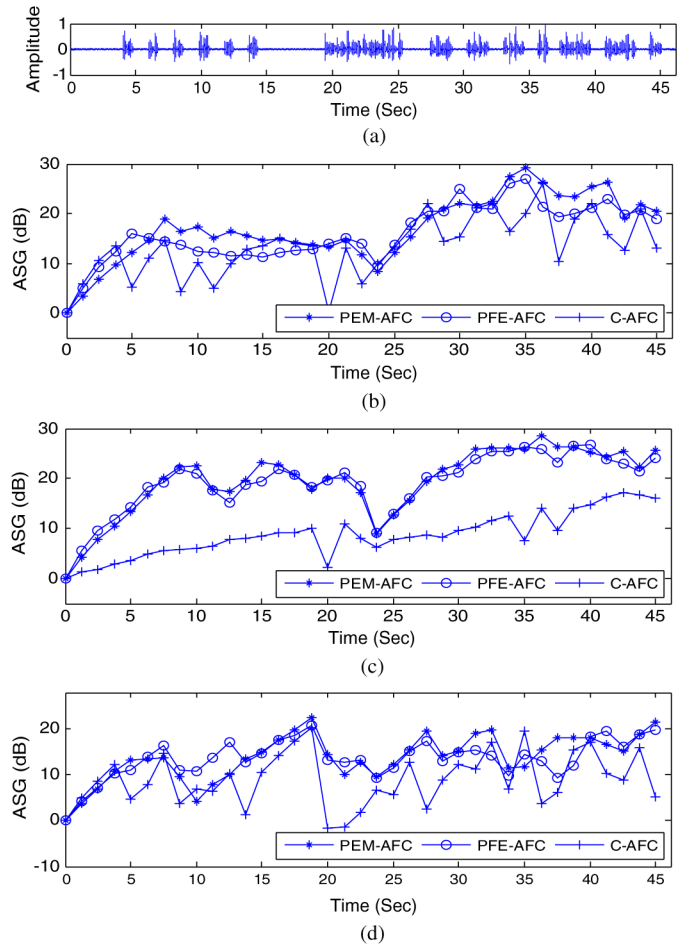


Fig. 17. (a) the clean speech signal, (b) the results of ASG simulation in 0 dB white noise, (c) the results of ASG simulation in 0 dB babble noise, and (d) the results of ASG simulation for clean speech.

TABLE V
THE EVALUATION OF VAD PERFORMANCE

SNR (dB)	Accuracy Rate (%)				
	White	Babble	Factory	Car	Average
0	85.05	57.55	71.17	83.25	74.26
3	85.92	70.15	81.57	86.05	80.92
5	86.69	80.80	84.90	86.29	84.67
7	86.64	83.35	85.92	86.46	85.59
10	86.99	85.92	86.13	87.07	86.53
Average	86.26	75.55	81.94	85.83	82.39

accuracy rate increases when SNRs are increased. It can be observed that the VAD accuracy can achieve similar performance at SNR of 10 dB and its accuracy in white and car noise is higher than in babble and factory noise. The features of the proposed VAD can be described by two aspects. First, the vowel detector in the pitch based VAD discriminates speech according to harmonic structure which is one of the most important speech features. Second, although the consonant detection is easily impacted by large noise, the hold state designed for speech protection still reduces VAD error. It is because that the hold state

covers most of the highly polluted consonants, except the onset of first Chinese character in each sentence. Therefore, the accuracy does not significantly degrade in low SNR scenarios.

IV. CONCLUSION

In this paper, a low complexity PFE-AFC based on the pitch and formant information has been proposed for HA applications. Based on the pitch and the harmonic detection, a high accurate pitch based VAD has also been proposed to assist the PFE-AFC. According to simulation results, the PFE-AFC can achieve the comparable PESQ and ASG compared with the high complexity PEM-AFC. In addition, the PFE-AFC outperforms the C-AFC in all evaluations. Note the complexity of the proposed decorrelation filter coefficients update which dominates the PEM-AFC complexity is significantly reduced by about four orders. From the system point of view, the proposed PFE-AFC has a regular architecture and low overhead since it can share computation resource with other HA processing units. Thus, the proposed algorithm is suitable for hardware implementations and HA applications.

ACKNOWLEDGMENT

The authors would like to thank Prof. S. T. Young for providing speech database as well as professor T. S. Chi for constructive discussions.

REFERENCES

- [1] J. M. Kates, *Digital hearing aids*. San Diego, CA, USA: Plural, 2008.
- [2] T. V. Waterschoot and M. Moonen, "Fifty years of acoustic feedback control: state of the art and future challenges," *Proc. IEEE*, vol. 99, no. 2, pp. 288–327, Feb. 2011.
- [3] H. Nyquist, "Regeneration theory," *Bell Syst. Tech. J.*, vol. 11, pp. 126–147, 1932.
- [4] J. Hellgren and U. Forssell, "Bias of feedback cancellation algorithms in hearing aids based on direct closed loop identification," *IEEE Trans. Speech Audio Process.*, vol. 9, no. 7, pp. 906–913, Nov. 2001.
- [5] J. M. Kates, "Constrained adaption for feedback cancellation in hearing aids," *J. Acoust. Soc. Amer.*, vol. 106, no. 2, pp. 1010–1019, Aug. 1999.
- [6] J. Hellgren, "Analysis of feedback cancellation in hearing aids with filtered-x LMS and direct method of closed loop identification," *IEEE Trans. Speech Audio Process.*, vol. 10, no. 2, pp. 119–131, Feb. 2002.
- [7] J. M. Kates, "Feedback cancellation in a hearing aid with reduced sensitivity to low-frequency inputs," U.S. patent 6,831,986, Dec. 14, 2004.
- [8] A. Spriet, I. Proudler, M. Moonen, and J. Wouters, "Adaptive feedback cancellation in hearing aids with linear prediction of the desired signal," *IEEE Trans. Signal Process.*, vol. 53, no. 10, pp. 3749–3763, Oct. 2005.
- [9] T. V. Waterschoot, "Design and evaluation of digital signal processing algorithms for acoustic feedback and echo cancellation," Ph.D. dissertation, Dept. Elect. Eng., Leuven Univ., Leuven, Belgium, 2009.
- [10] Y. J. Chen, C. W. Wei, Y. FanChiang, Y. L. Meng, Y. C. Huang, and S. J. Jou, "Neuromorphic pitch based noise reduction for monosyllabic hearing aid system application," *IEEE Trans. Circuits Syst. I*, vol. 61, no. 2, pp. 463–475, Feb. 2014.
- [11] C. W. Wei, C. C. Tsai, T. S. Chang, and S. J. Jou, "Perceptual multi-band spectral subtraction for noise reduction in hearing aids," in *Proc. Asia Pacific Conf. Circuits Syst.*, Kuala Lumpur, Malaysia, 2010, pp. 692–695.
- [12] Y. T. Kuo, T. J. Lin, W. H. Chang, Y. T. Liu, and C. W. Liu, "Complexity-effective auditory compensation for digital hearing aids," in *Proc. Int. Symp. Circuits Syst.*, May 2008, pp. 1472–1475.
- [13] K. C. Chang, Y. T. Kuo, T. J. Lin, and C. W. Liu, "Complexity-effective dynamic range compression for digital hearing aids," in *Proc. IEEE Int. Symp. Circuits Syst.*, Paris, France, Jun. 2010, pp. 2378–2381.

- [14] Y. T. Kuo, T. J. Lin, Y. T. Li, and C. W. Liu, "Design and implementation of low-power ANSI S1.11 filter bank for digital hearing aids," *IEEE Trans. Circuits Syst. I*, vol. 57, no. 7, pp. 1684–1696, Jul. 2010.
- [15] T. Nakatani and H. G. Okuno, "Harmonic sound stream segregation using localization and its application to speech stream segregation," *Speech Commun.*, vol. 27, no. 3–4, pp. 209–222, Apr. 1999.
- [16] K. Nakadai, K. Hidai, H. G. Okuro, and H. Kitano, "Real-time speaker localization and speech separation by audio-visual integration," in *Proc. IEEE Int. Conf. Robotics Autom.*, Washington, D.C., USA, 2002, pp. 1043–1049.
- [17] A. V. Oppenheim, R. W. Schaffer, and J. R. Buck, *Discrete-time signal processing*, 2nd ed. Upper Saddle River, NJ, USA: Prentice-Hall, 2009.
- [18] V. K. Ingle and J. G. Proakis, *Digital signal processing using Matlab*, 4th ed. Boston, MA, USA: PWS, 1997.
- [19] S. Winograd, "On computing the discrete fourier transform," *Math. Comput.*, vol. 32, pp. 175–199, 1978.
- [20] I. Selesnick, Apr. 10, 2010, Winograd's short DFT algorithms, [Online]. Available: <http://cnx.org/content/m16333/latest/>
- [21] B. Parhami, *Computer arithmetic: algorithms and hardware designs*, 2nd ed. New York, NY, USA: Oxford Univ. Press, 2009.
- [22] A. Wang and A. Chandrakasan, "A 180-mV subthreshold FFT processor using a minimum energy design methodology," *IEEE J. Solid-State Circuits*, vol. 40, no. 1, pp. 310–319, Jan. 2005.
- [23] T. V. Waterschoot, G. Rombouts, and M. Moonen, Oct. 2005, Software for double-talk robust acoustic echo cancellation, [Online]. Available: <ftp://ftp.esat.kuleuven.be/pub/SISTA/vanwaterschoot/abstracts/05-162.html>
- [24] "Perceptual evaluation of speech quality (PESQ), and objective method for end-to-end speech quality assessment of narrowband telephone networks and speech codecs," ITU-T Recommendation, 2000, p. 862.
- [25] J. M. Kates, "Room reverberation effects in hearing aid feedback cancellation," *J. Acoust. Soc. Amer.*, vol. 109, no. 1, pp. 367–378, Jan. 2001.
- [26] J. Yang, M. T. Tan, and J. S. Chang, "Modeling external feedback path of an ITE digital hearing instrument for acoustic feedback cancellation," in *Proc. IEEE Int. Symp. Circuits Syst.*, Kobe, Japan, 2005, vol. 2, pp. 1326–1329.
- [27] [Online]. Available: <http://rocling.iis.sinica.edu.tw/CKIP/engversion/20corpus.htm>
- [28] P. C. Loizou, *Speech Enhancement, Theory and Practice*. Boca Raton, FL, USA: CRC, 2007.



Yi FanChiang received the B.S. degree in electrical engineering from National Central University, Taoyuan, Taiwan, in 2006 and M.S. degree from National Tsing Hua University, Hsinchu, Taiwan in 2008. He is currently working toward the Ph.D. degree in electronic engineering from National Chiao Tung University (NCTU,) Hsinchu, Taiwan, since 2008. His research interests include digital signal processing, speech processing, and bio-electronic circuit.



Cheng-Wen Wei received the B.S. and M.S. degrees in electrical engineering from Yuan Ze University, Taoyuan, Taiwan, in 1998 and 2000. He received the Ph.D. degree in electronic engineering from National Chiao Tung University (NCTU,) Hsinchu, Taiwan in 2013.

From 2000 to 2006, he was an engineer and worked on Delta Sigma data convertor, speech signal processing and VLSI design, with the Product Development Division/Digital Circuit Design Dept. (DCD,) Elan Microelectronics Corporation (EMC.)

Hsinchu, Taiwan. He is currently a research and design manager in the Digital Circuit Division of ZEITEC semiconductor, Hsinchu, Taiwan, where he is designing algorithms and digital circuits for low power signal enhancement and noise reduction. His research interests include digital signal processing, speech processing, data conversion and low power VLSI design.

Dr. Wei received the Gold Medal Award at the 2013 Merry Electroacoustic Thesis Award.



Yi-Le Meng received the B.S. and M.S. degrees in electronic engineering from National Chiao Tung University (NCTU,) Hsinchu, Taiwan, in 2008 and 2011. His research interests include low-power digital IC design and speech processing.



Yu-Wen Lin received the B.S. electronic engineering from National Chiao Tung University (NCTU,) Hsinchu, Taiwan, in 2012. He is currently working toward M.S. degree in electronic engineering from National Chiao Tung University (NCTU,) Hsinchu, Taiwan, from 2012.

His research interests include low-power digital IC design and speech processing.



Shyh-Jye Jou (S'86–M'90–SM'97) received his B.S. degree in electrical engineering from National Chen Kung University in 1982, and M.S. and Ph.D. degrees in electronics from National Chiao Tung University in 1984 and 1988, respectively.

He joined Electrical Engineering from 1990 to 2004 and became a Professor in 1997. Since 2004, he has been a Professor of National Chiao Tung University and became the Chairman of Department of Electronics Engineering from 2006 to 2009.

From 2011, he becomes the Vice President, Office of International Affairs, National Chiao Tung University. He was a visiting research Professor in the Coordinated Science Laboratory at University of Illinois, Urbana-Champaign, during 1993-1994 and 2010 academic years. In

the summer of 2001, he was a visiting research consultant in the Communication Circuits and Systems Research Laboratory of Agere Systems, USA. He received Outstanding Engineering Professor Award, Chinese Institute of Engineers and Chinese Institute of Electrical Engineering in 2011 and 2013, respectively.

He is an active contributor in IEEE, especially, CAS Society. He served as Chapter Chair of IEEE Circuits and Systems Society Taipei Chapter, 2006, Distinguished Lecturer of CAS society, 2009-2010, Guest Editor of IEEE JOURNAL OF SOLID STATE CIRCUITS, Nov. 2008, and was Track Chair of Nanoelectronics and Gigascale Systems and Technical Committee member of VLSI Systems and Applications, 2011-2013. He served as the Conference Chair of IEEE International Symp. on VLSI Design, Automation and Test (VLSI-DAT) and International Workshop on Memory Technology, Design, and Testing. He also served as Technical Program Chair or Co-Chair in IEEE VLSI-DAT, International IEEE Asian Solid-State Circuit Conference, IEEE Biomedical Circuits and Systems, and other international conferences. He has published more than 100 IEEE journal and conference papers. His research interests include design and analysis of high speed, low power mixed-signal integrated circuits, communication and Bio-Electronics integrated circuits and systems.



Tian-Sheuan Chang (S'93–M'06–SM'07) received the B.S., M.S., and Ph.D. degrees in electronic engineering from National Chiao Tung University (NCTU,) Hsinchu, Taiwan, in 1993, 1995, and 1999, respectively.

He was a Deputy Manager with Global Unichip Corporation, Hsinchu, from 2000 to 2004. In 2004, he joined the Department of Electronics Engineering, NCTU, where he is currently a Professor. In 2009, he was a Visiting Scholar in IMEC, Heverlee, Belgium.

His current research interests include system-on-a-chip design, VLSI signal processing, and computer architecture.

Dr. Chang was a recipient of the Excellent Young Electrical Engineer from Chinese Institute of Electrical Engineering in 2007, and the Outstanding Young Scholar from Taiwan IC Design Society in 2010. He has been actively involved in many international conferences as an organizing committee or technical program committee member. He is currently an Editorial Board Member of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY.