Acoustic Echo Cancellation Using a Vector-Space-Based Adaptive Filtering Algorithm

Yu Tsao, *Member IEEE*, Shih-Hau Fang*, *Senior Member IEEE*, and Yao Shiao

Abstract—A novel vector-space-based adaptive filtering (VAF) algorithm for acoustic echo cancellation (AEC) is presented. The proposed VAF algorithm can be divided into two phases: offline and online. In the offline phase, VAF constructs a vector space to incorporate the prior knowledge of adaptive filter coefficients from a wide range of different channel characteristics. Then, in the online phase, a mapping function is derived to estimate the adaptive filter for the testing condition using the constructed vector space. By using the vector space, VAF can effectively and efficiently estimate the parameters of the adaptive filter for the unknown testing condition. The experimental results for three designed AEC tasks demonstrate that VAF provides notably faster convergence rates compared to conventional adaptive filtering methods.

***Index Terms*—Vector-space-based adaptive filtering algorithm, VAF, prior knowledge, acoustic echo cancellation**

I. Introduction

IN hands-free telephony and teleconferencing systems, the existence of undesired echoes, usually produced by acoustic coupling between the loudspeaker and microphone, may seriously degrade the quality and intelligibility of speech signals. Various adaptive filtering (AF) methods have been developed and are widely employed for acoustic echo cancellation (AEC) in practical communication systems [1-9]. Generally, an AF method involves characterizing an echo-path impulse response using a filter function and removing the predicted echoes from the microphone signals in a step-by-step manner. Figure 1 shows an AF-based AEC system, where $\boldsymbol{x}(n)$, $\boldsymbol{y}(n)$, and $\boldsymbol{e}(n)$ are the input signal, echo signal, and estimated error, respectively. The true and estimated echo paths are represented by $\boldsymbol{h}(n)$ and $\widehat{\boldsymbol{h}}(n)$, respectively. An AEC system removes the undesired echo signal components, $\boldsymbol{y}(n)$, by modeling the echo-path impulse response with an adaptive filter, and subtracting the estimated echo signal components, $\widehat{\boldsymbol{y}}(n)$.

Least mean square (LMS) and normalized LMS (NLMS) are two fundamental AEC approaches [10-18]. Due to their effectiveness and low computation cost, LMS and NLMS have been widely used to perform AEC in real-world applications. Several improved schemes have been proposed more recently. The affine projection algorithm (APA) incorporates a set of input vectors to update the weights rather than only using the current input vector [18-21]. Proportionate APA (PAPA) has been proposed to update the filter according to the proportion of the magnitude of the estimated filter coefficient [22-25]. Several extensions of PAPA have further enhanced performance such as IPAPA [26], simplified PAPA [27], memory IPAPA (MIPAPA), and the approximated MIPAPA algorithms [28, 29]. Moreover, Levenberg–Marquardt regularized APA (LMR-APA) regularized the filter estimation by using the prior knowledge to accelerate the convergence speed [30-32].

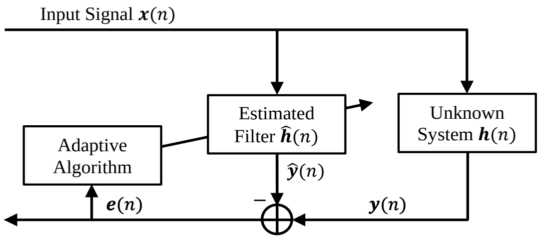

Fig. 1. Flowchart of a conventional AF-based AEC system.

This study proposes a vector-space-based AF (VAF) algorithm for AEC. VAF utilizes prior information of the acoustic conditions to facilitate filter estimations. Different from incorporating the prior information to form the regularization terms (as used in LMR-APA), the proposed VAF structures the prior knowledge into a vector space and uses a mapping function to transform the space for updating the adaptive-filter coefficients. VAF allows the parameters of the adaptive filters to be estimated using the constructed vector space, thereby increasing both the effectiveness and efficiency of the estimation process. Since the concept and procedure of VAF are different from PAPA and LMR-APA, VAF can be suitably integrated with these two approaches. Hereafter, the integrations of VAF with PAPA and LMR-APA are termed vector-space PAPA (VPAPA) and vector-space regularized APA (VRAPA), respectively.

We evaluated the proposed VAF using a designed AEC task. The experimental results demonstrate that VPAPA and VRAPA provide faster convergence and lower convergence mean-square errors compared to PAPA and LMR-APA, respectively. The results demonstrate the ability of VAF to transform traditional methods into enhanced versions for faster convergence and lower errors, thereby confirming the effectiveness of using prior knowledge for AF algorithms.

II. Adaptive Filtering Methods

This section reviews the fundamental aspects of AF and presents PAPA and regularized APA algorithms.

A. Fundamental Aspects of an AF Algorithm

Referring to Fig. 1, $\boldsymbol{x}(n)$ is the input signal, and $\boldsymbol{h}(n)$ is the true echo path, represented as

Yu Tsao is with the Research Center for Information Technology Innovation, Academia Sinica, Taipei, Taiwan (e-mail: yu.tsao@citi.sinica.edu.tw). Shih Hau Fang and Yao Shiao are with the Department of Electrical Engineering and Innovation Center for Big Data and Digital Convergence, Yuan Ze University, Taoyuan, Taiwan (e-mail: shfang@saturn.yzu.edu.tw and s970535@ee.yzu.edu.tw).

$$\boldsymbol{x}(n) = [x(n)\; x(n-1) \cdots x(n-L+1)]^{\mathrm{T}}, \quad (1)$$
$$\boldsymbol{h}(n) = [h_0(n)\; h_1(n)\; \cdots\; h_{L-1}(n)]^{\mathrm{T}}, \quad (2)$$
where L is the length of the filter, $[\;]^{\mathrm{T}}$ represents the transpose operator, and both $\boldsymbol{x}(n)$ and $\boldsymbol{h}(n)$ are L-dimensional vectors. The signal $y(n)$ is the convolution of $\boldsymbol{x}(n)$ and $\boldsymbol{h}(n)$:
$$y(n) = \mathbf{x}^{\mathrm{T}}(n)\boldsymbol{h}(n). \quad (3)$$
The goal of AF methods is to estimate filter $\widehat{\boldsymbol{h}}(n)$, so that
$$\hat{y}(n) = \mathbf{x}^{\mathrm{T}}(n)\widehat{\boldsymbol{h}}(n), \quad (4)$$
producing minimal error, $e(n) = y(n) - \hat{y}(n)$. Below we describe two types of AF methods: PAPA and LMR-APA.

B. The Proportionate APA

The PAPA forms an $L \times p$ matrix, $\boldsymbol{X}(n)$, using the most recent samples of the input signal,
$$\boldsymbol{X}(n) = [\boldsymbol{x}(n)\; \boldsymbol{x}(n-1)\; \cdots \boldsymbol{x}(n-p+1)]. \quad (5)$$
The output signal is formulated as a p-dimensional vector,
$$\boldsymbol{y}(n) = [y(n)\; y(n-1)\; \cdots y(n-p+1)]^{\mathrm{T}}. \quad (6)$$
Each error vector $\boldsymbol{e}(n)$ with p dimensions is formulated as
$$\boldsymbol{e}(n) = \boldsymbol{y}(n) - \boldsymbol{X}^{\mathrm{T}}(n)\widehat{\boldsymbol{h}}(n), \quad (7)$$
where $\boldsymbol{e}(n) = [e(n)\; e(n-1)\; \cdots e(n-p+1)]^{\mathrm{T}}$. The PAPA updates the impulse response of the channel based on [22–25]:
$$\begin{aligned}\widehat{\boldsymbol{h}}(n) = \widehat{\boldsymbol{h}}(n-1) + \mu\boldsymbol{G}(n)\boldsymbol{X}(n)[\boldsymbol{X}^{\mathrm{T}}(n)\\ \boldsymbol{G}(n)\boldsymbol{X}(n) + \delta\mathbf{I}_{p\times p}]^{-1}\boldsymbol{e}(n),\end{aligned} \quad (8)$$
where μ is the step size, δ is a positive parameter, $\mathbf{I}_{p\times p}$ is an $p \times p$ identity matrix, and $\boldsymbol{G}(n)$ is an $L \times L$ diagonal proportionate matrix. The elements of $\boldsymbol{G}(n)$ are denoted as $g_l(n), l = 0, \dots, L-1$, which are related to the magnitudes of the coefficients in $\widehat{\boldsymbol{h}}(n-1)$ and can be expressed as
$$g_l(n) = \frac{k_l(n)}{\sum_{i=0}^{L-1} k_i(n)}, \quad (9)$$
where $k_l(n) = \max\{\rho \times \max[\gamma, |\hat{h}_0(n-1)| \cdots |\hat{h}_{L-1}(n-1)|], |\hat{h}_l(n-1)|\}$, (10)

where $l = 0, \dots, L-1$ being the tap indices. Parameter γ is included in (10) to prevent $\hat{h}_l(n-1)$ from stalling during the initialization stage with $\widehat{\boldsymbol{h}}(0) = \mathbf{0}_{L\times 1}$, while ρ prevents coefficients from stalling when they are much smaller than the largest coefficient [25]. The derivation of the PAPA in (8) can be applied to the two related approaches, APA [18-21] and IPAPA [26], by specifying different forms of $\boldsymbol{G}(n)$.

C. The Regularized APA

The RAPA utilizes prior knowledge to design the regularization term. The LMR-APA estimates the channel characteristic by:
$$\begin{aligned}\widehat{\boldsymbol{h}}(n) = \widehat{\boldsymbol{h}}(n-1) + \mu\boldsymbol{R}_f(n)\boldsymbol{X}(n)\; [\boldsymbol{X}^{\mathrm{T}}(n)\\ \boldsymbol{R}_f(n)\boldsymbol{X}(n) + \boldsymbol{\Sigma}_e(n)]^{-1}\boldsymbol{e}(n),\end{aligned} \quad (11)$$
where $\boldsymbol{R}_f(n)$ is the covariance matrix, and $\boldsymbol{\Sigma}_e(n)$ is the diagonal weighting matrix. More information about preparing $\boldsymbol{R}_f(n)$ and $\boldsymbol{\Sigma}_e(n)$ can be found in [30]. Comparing (8) and (11), PAPA and LMR-APA use similar updating functions with different regularization terms. In [30], it is verified that using a carefully designed $\boldsymbol{R}_f(n)$ and $\boldsymbol{\Sigma}_e(n)$, the regularization techniques effectively improve conventional AF methods.

III. The VAF

The flowchart of the proposed VAF is illustrated in Fig. 2. A comparison of Figs. 1 and 2 reveals that VAF incorporates a vector space structure, $\mathbf{H}$, to adaptively adjust $\widehat{\boldsymbol{h}}(n)$. The proposed VAF algorithm can be divided into two phases: offline and online, which will be described in detail below.

A. Offline Phase of VAF

The goal of the offline phase is to prepare a structure that incorporates the prior knowledge of filter coefficients for a wide range of channel characteristics. By collecting K different channels, we prepare K sets of filter coefficient vectors, where $\boldsymbol{h}^k$ contains L coefficients, namely
$$\mathbf{H} = [\boldsymbol{h}^1 \cdots \boldsymbol{h}^K] = \begin{bmatrix} h_0^1 & \cdots & h_0^K \\ \vdots & \ddots & \vdots \\ h_{L-1}^1 & \cdots & h_{L-1}^K \end{bmatrix}, \quad (12)$$
where $\mathbf{H}$ is the prior-knowledge matrix.

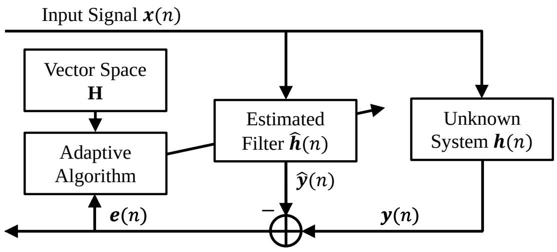

Fig. 2. Flowchart of a VAF-based AEC system.

B. Online Phase of VAF

The matrix $\mathbf{H} = [\boldsymbol{h}^1 \cdots \boldsymbol{h}^K]$ can be considered as the vector space formed by different channel characteristics. With $\mathbf{H}$, we attach an $L \times L$ identity matrix $\mathbf{I}_{L\times L}$ to produce a matrix $\mathbf{S}$:
$$\mathbf{S} = [\mathbf{H}|\mathbf{I}_{L\times L}]. \quad (13)$$
With the prepared $\mathbf{S}$, we compute $\widehat{\boldsymbol{h}}(n)$ using $\widehat{\boldsymbol{w}}(n)$, where $\widehat{\boldsymbol{w}}(n) = [\,\mathbf{a}(n)^{\mathrm{T}}\, \mathbf{b}(n)^{\mathrm{T}}]^{\mathrm{T}}$ is an $(K$+$L)$-dimensional vector, $\mathbf{a}(n) = [a_1(n) \cdots a_K(n)]^{\mathrm{T}}$ is a K-dimensional weighting vector, and $\mathbf{b}(n) = [b_0(n)\; \cdots b_{L-1}(n)]^{\mathrm{T}}$ is an L-dimensional bias vector. Then we estimate the filter $\widehat{\boldsymbol{h}}(n)$ by
$$\widehat{\boldsymbol{h}}(n) = \mathbf{S}\widehat{\boldsymbol{w}}(n) = \mathbf{H}\mathbf{a}(n) + \mathbf{b}(n). \quad (14)$$
From (14), with a well-prepared $\mathbf{H}$, the weighting vector $\mathbf{a}(n)$ facilitates an efficient estimation of $\widehat{\boldsymbol{h}}(n)$. The same technique that incorporates prior information to facilitate online estimation has been confirmed effective in many pattern characterization tasks [33-35]. Notably, by constraining $\mathbf{a}(n)$=0, no prior information is used, and VAF becomes conventional AF.

C. The VPAPA

When applying the vector-space technique, (7) becomes
$$\boldsymbol{e}(n) = \boldsymbol{y}(n) - \boldsymbol{X}^{\mathrm{T}}(n)\mathbf{S}\widehat{\boldsymbol{w}}(n). \quad (15)$$
Similar to the derivation of PAPA as presented in [24], we first define the VPAPA cost function as
$$\begin{aligned}\mathcal{L}(\widehat{\boldsymbol{w}}(n), \boldsymbol{\lambda}) = \|\mathbf{S}\widehat{\boldsymbol{w}}(n) - \mathbf{S}\widehat{\boldsymbol{w}}(n-1)\|^2\\ + [\boldsymbol{y}(n) - \boldsymbol{X}^{\mathrm{T}}(n)\boldsymbol{G}(n)\mathbf{S}\widehat{\boldsymbol{w}}(n)]^{\mathrm{T}}\boldsymbol{\lambda},\end{aligned} \quad (16)$$
where $\boldsymbol{\lambda}$ is a Lagrange multiplier, and $\boldsymbol{G}(n)$ is defined in (9) and (10). By differentiating $\mathcal{L}(\widehat{\boldsymbol{w}}(n), \boldsymbol{\lambda})$ with respect to $\widehat{\boldsymbol{w}}(n)$ and equating the result to zero, we have
$$2\mathbf{S}^{\mathrm{T}}[\mathbf{S}\widehat{\boldsymbol{w}}(n) - \mathbf{S}\widehat{\boldsymbol{w}}(n-1)] - \mathbf{S}^{\mathrm{T}}\boldsymbol{G}(n)\boldsymbol{X}(n)\boldsymbol{\lambda} = 0. \quad (17)$$
By setting $\mathbf{J} = \mathbf{S}^{\mathrm{T}}\mathbf{S}$, $\mathbf{U} = \mathbf{J}^{-1}\mathbf{S}^{\mathrm{T}}$, and $\mathbf{V} = \mathbf{S}\mathbf{U}$, we have
$$\boldsymbol{\lambda} = (\boldsymbol{X}^{\mathrm{T}}(n)\mathbf{V}\boldsymbol{G}(n)\boldsymbol{X}(n))^{-1}2\boldsymbol{e}(n). \quad (18)$$
Finally, the proposed VPAPA updates $\widehat{\boldsymbol{w}}(n)$ by
$$\begin{aligned}\widehat{\boldsymbol{w}}(n) = \widehat{\boldsymbol{w}}(n-1) + \mu'\mathbf{U}\boldsymbol{G}(n)\boldsymbol{X}(n)[\boldsymbol{X}^{\mathrm{T}}(n)\; \mathbf{V}\boldsymbol{G}(n)\boldsymbol{X}(n)\\ +\delta'\,\mathbf{I}_{p\times p}]^{-1}\boldsymbol{e}(n),\end{aligned} \quad (19)$$

where μ' is the step size, and δ' is a positive parameter. With the estimated $\widehat{\boldsymbol{w}}(n)$, we can use (14) to obtain $\widehat{\boldsymbol{h}}(n)$. It should be noted that $\mathbf{S}$, $\mathbf{J}$ (and $\mathbf{J}^{-1}$), $\mathbf{U}$, and $\mathbf{V}$ can be prepared in the offline stage. Therefore, VPAPA does not require additional computation to obtain these matrices in the online phase. Similar to the relation between PAPA and APA, the update function for vector-space APA (VAPA) can be derived using (19), by using an identity matrix for $\boldsymbol{G}(n)$.

D. The VRAPA

Next, we apply the vector-space technique with regularized APA to obtain vector-space regularized APA, namely VRAPA. Similar to the derivation of LMR-APA in [30], VRAPA first formulates a cost function

$$\mathcal{H}\big(\widehat{\boldsymbol{w}}(n)\big) = \mathbf{A}^{\mathrm{T}}(n)\boldsymbol{R}_f^{-1}(n)\mathbf{A}(n) + \mathbf{B}^{\mathrm{T}}(n)\boldsymbol{\Sigma}_e^{-1}(n)\mathbf{B}(n), \qquad (20)$$

where $\mathbf{A}(n) = [\mathbf{S}\widehat{\boldsymbol{w}}(n) - \mathbf{S}\widehat{\boldsymbol{w}}(n-1)]$, and $\mathbf{B}(n) = [\boldsymbol{y}(n) - \boldsymbol{X}^{\mathrm{T}}(n)\mathbf{S}\widehat{\boldsymbol{w}}(n)]$. Based on (20) with some derivations, we can obtain the VRAPA update formulation:

$$\widehat{\boldsymbol{w}}(n) = \widehat{\boldsymbol{w}}(n-1) + \mu'\boldsymbol{Q}(n)\mathbf{S}^{\mathrm{T}}\boldsymbol{X}(n)[\boldsymbol{X}^{\mathrm{T}}(n)\mathbf{S}\boldsymbol{Q}(n) \mathbf{S}^{\mathrm{T}}\boldsymbol{X}(n) + \boldsymbol{\Sigma}_e(n)]^{-1}\boldsymbol{e}(n), \qquad (21)$$

where $\boldsymbol{Q}(n) = [\mathbf{S}^T\boldsymbol{R}_f^{-1}(n)\mathbf{S}]^{-1}$.

IV. Experiments

A. Experimental Setup

The room impulse response (RIR) was simulated using an RIR generator toolkit [36]. The RIR generator allows assignments of the RS, position of the voice source, and position of the microphone. We can also design the reflection coefficients (RCs), sampling rate, and reflection time for a particular testing RIR condition. In this study, we varied the RSs and RCs while keeping other parameters fixed.

The testing data in this study comprised a clean utterance pronounced by a female speaker: “01bc0207.wv1” in the Aurora-4 database [37-39]. The RIR generator was then applied to artificially generate the RIR-affected test utterance. To prepare the vector space for the VAF algorithm, another two utterances made by the same speaker in the Aurora-4 database were selected and concatenated to form the training data. The lengths of training and testing data were around 25 and 13 seconds, respectively, and the sampling rates for both data were 8 kHz.

We simulated 50 RIR effects, with 10 kinds of room size RS and 5 RCs. These 50 RIR effects were applied to the training utterance, thus preparing 50 RIR affected utterances. Table 1 summarizes the RSs and RCs used to prepare the training and testing data. The source and microphone are always placed in (1.0, 1.0, 1.0) and (1.0, 0.4, 0.6), respectively. The sparseness of the impulse responses of the training data range from 0.132 to 0.987 [19]. For the 50 RIR affected utterances, we applied PAPA to compute filters with the length set to 100. Thus, we collected 50 filter vectors (K=50, $\boldsymbol{h}^1 \cdots \boldsymbol{h}^{50}$), each consisting of 100 coefficients (L=100), and used them to form the vector space, $\mathbf{H}$, according to (12). To investigate the effectiveness of VAF, we designed three sets of testing conditions, denoted as test sets A, B, and C. For test set A, we adopted the RS and RC setups that were covered by the training data. For test set B, the RS was included while the RC was not included in the training set. For test set C, neither the RS nor the RC was included in the training set. The sparseness of the impulse responses of test sets A, B, and C are 0.588, 0.438, and 0.387, respectively. In this study, we set the projection order, p=16 in (5).

TABLE I
Configuration of Training and Testing Data Sets

Data Set	Room Size (RS) (x, y, z)	Reflection Coefficient (RC)
Training Set	(1.1, 1.1, 1.1)	0.81
	(1.2, 1.2, 1.2)	0.62
	(1.3, 1.3, 1.3)	0.43
	(1.4, 1.4, 1.4)	0.24
	(1.5, 1.5, 1.5)	0.05
	(1.6, 1.6, 1.6)	
	(1.7, 1.7, 1.7)	
	(1.8, 1.8, 1.8)	
	(1.9, 1.9, 1.9)	
	(2.0, 2.0, 2.0)	
Test Set A	(1.4, 1.4, 1.4)	0.43
Test Set B	(1.4, 1.4, 1.4)	0.5
Test Set C	(1.2, 1.3, 1.4)	0.5

B. Experimental Results

This section presents the experimental results for the PAPA and the proposed VPAPA for the three test sets[1] [40]. In addition to the conventional VPAPA that uses a 100×50 matrix $\mathbf{H}$ to form the vector space, we applied the K-means algorithm [41] on the 50 filter vectors, namely $\boldsymbol{h}^1 \cdots \boldsymbol{h}^{50}$, in $\mathbf{H}$ to obtain an $\mathbf{H}_{\mathrm{C}}$, 100×C matrix (where C is the number of clusters after the K-means process). In our preliminary experiments, we have tested VPAPA using $\mathbf{H}_{\mathrm{C}}$ with various C values, and the performance shows that C=5 already provides satisfactory performance, and thus C=5 is used to prepare $\mathbf{H}_{\mathrm{C}}$ in this study. In the following discussion, we use “VPAPA (O)” and “VPAPA (C)” to denote VPAPA using $\mathbf{H}$ and $\mathbf{H}_{\mathrm{C}}$, respectively, for (13)-(19). Notably since VPAPA (C) uses a 100×5 matrix $\mathbf{H}_{\mathrm{C}}$, the online computation and the storage requirement for the matrix are both lower than VPAPA (O), which uses the original 100×50 matrix $\mathbf{H}$. In Fig. 3, the black, blue, and red lines are for the PAPA, VPAPA (O), and VPAPA (C), respectively. The top, middle, and bottom panels present the results of mean square error (MSE, in dB) of $\boldsymbol{e}(n)$ in (15) with varying iteration numbers for test sets A, B, and C, respectively. From Fig. 3, it is noted that VPAPA (O) gives a higher convergence rate as well as lower convergence errors compared to PAPA consistently over the three test sets. The results confirm the effectiveness of the VPAPA irrespective of whether or not the testing condition is covered by the training condition. Figure 3 also shows that VPAPA (C) provides comparable convergence rates to VPAPA (O) for the three test sets. The results indicate that by applying K-means, we can form a compact vector space for VPAPA while maintaining comparable performance.

Figure 4 presents the echo return loss enhancement (ERLE) and MSE results for VAPA, VPAPA, and VRAPA. Our preliminary experiments also showed that for VAPA and VRAPA, using $\mathbf{H}$ and $\mathbf{H}_{\mathrm{C}}$ achieve similar performances, which agree well with the results of VPAPA in Fig. 3. Thus, we only presented the results of using $\mathbf{H}_{\mathrm{C}}$ (C=5) in this set of experiments, and thus VAPA and VRAPA are denoted as VAPA (C) and VRAPA (C), respectively. Test set C, which was more difficult

[1] The codes are available at: http://www.citi.sinica.edu.tw/pages/yu.tsao/publications_en.html.

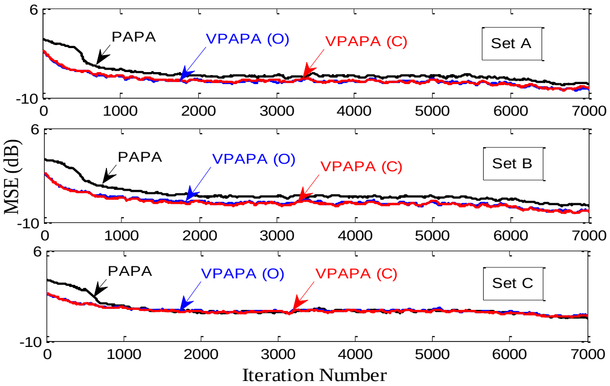

Fig. 3. MSE scores over iteration numbers for PAPA, VPAPA (O), and VPAPA (C) for test sets A, B, and C.

than tests sets A and B for VAF, was selected for evaluation. The results for the APA, PAPA, and LMR-APA (denoted as RAPA for simplicity) are also shown for comparison. Learning from the previous study [30], for RAPA and VRAPA (C), we first prepared a regularization term $\boldsymbol{R}_{f,init}$ by using the mean of the prepared 50 filters from the training data: $\boldsymbol{R}_{f,init} = diag[\bar{h}_0^2, \bar{h}_1^2, \dots \bar{h}_{L-1}^2]$, where $\bar{\boldsymbol{h}}$=mean ($\boldsymbol{h}^1, \cdots \boldsymbol{h}^{50}$). Because $\boldsymbol{R}_{f,init}$ was estimated from the training data, which may not provide the optimal prior information for the unknown test conditions (such as test set C in this study), we proposed to use a scalar $\alpha(n)$ to optimize the regularization term by

$$\boldsymbol{R}_f(n) = [1 - \alpha(n)] \cdot \boldsymbol{R}_{f,init} + \alpha(n) \cdot \mathbf{I}_{L\times L}. \quad (22)$$

In Fig. 4, we presented the results that were obtained using optimal $\alpha(n)$ in the task, where $\alpha(n)$ was set to a small value at the beginning stage and then was set to 1.0 after several iterations. This same setup of $\alpha(n)$ for (22) was used for both RAPA in (11) and VRAPA in (21) to obtain the results shown in Fig. 4 for a fair comparison.

From Fig. 4, the results first demonstrate that PAPA converges faster than APA. It is also noted that the RAPA outperforms APA and PAPA, confirming the effects of using $\boldsymbol{R}_f(n)$, even though the prior information was not directly collected using the data from test C. Finally, comparisons between VAPA(C) and APA, VPAPA (C) and PAPA, and VRAPA (C) and RAPA show that VAF gives notably faster convergences compared to corresponding conventional AF counterparts.

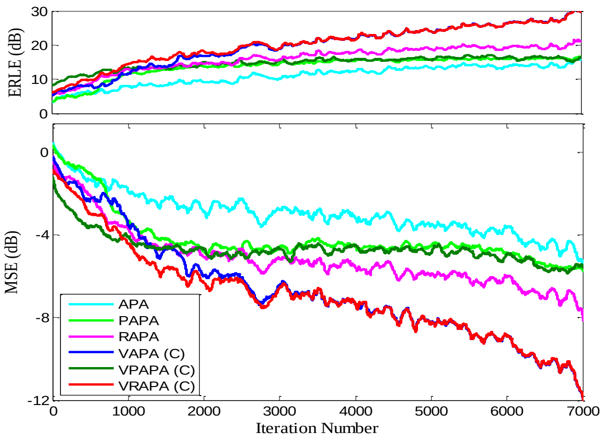

Fig. 4. ERLE (upper panel) and MSE (lower panel) curves of VAPA (C), VPAPA (C), and VRAPA (C) with APA, PAPA, and RAPA for test set C.

These results verify the effectiveness of using the vector space to incorporate prior knowledge for AF estimation in AEC tasks.

In our experiments using different utterances from Aurora-4, VAF methods also outperformed their corresponding AF counterparts, agreeing well with the results shown in Figs. 3 and 4. Additionally, we noted very similar trends for more iteration points in the test results except that performances of PAPA and VPAPA became saturated while APA continued reducing MSE.

C. Discussion

The main purpose of this study is to verify the vector-space technique can enable conventional AF methods to achieve better performance. We tested the proposed technique integrated with APA (no regularization), PAPA (with a proportionate matrix), and RAPA (with a regularization term). Results showed that VAPA, VPAPA, and VRAPA outperform APA, PAPA, and RAPA, respectively, confirming that the vector-space technique can enhance traditional AF methods no matter what kind of regularization term is used.

Since an additional vector-space structure is used, VAF requires extra computation complexity and storage when compared to the conventional AF methods. From Eqs. (8) and (21), VPAPA needs two additional matrix multiplications compared to PAPA. An additional computation of Eq. (14) is also required. When considering storage requirements, VPAPA needs extra $\mathbf{U}$ $((C+L)\times L)$ and $\mathbf{V}$ $(L\times L)$ matrices prepared in the offline phase. The additional complexity and storage are also necessary when extending APA to VAPA and RAPA to VRAPA. Although VAF provides higher convergence speeds over AF counterparts (from Figs. 3 and 4), the additional complexity and storage could limit applicability when resources are limited. Therefore, future studies will investigate effective approaches for simplifying computation and reducing storage requirements.

V. CONCLUSION

The VAF algorithm is proposed for AEC. By using the constructed vector space to incorporate the prior knowledge about the characteristics of various channels, VAF can provide better performance compared to conventional AF algorithms in terms of convergence rates. The experimental results verify that VAPA, VPAPA, and VRAPA achieved notably higher convergence rates compared to the conventional APA, PAPA, and RAPA, respectively. Moreover, it is confirmed that by applying K-means, we can form a compact vector space for VAF while maintaining comparable performance.

The VAF will exhibit good performance when a considerable amount of prior information about speaking conditions is available. Due to the rapid proliferation of mobile devices in recent years, user-specific AEC systems are growing in popularity. The VAF is especially favorable for these personal devices since previous using experience is often accessible. Future research will investigate the complexity consideration and examine techniques for online optimizing the matrix $\mathbf{H}$. Moreover, the optimal step size for the VAF will be investigated, and VAF will be evaluated under more complicated, noisy and time-varying test conditions.

REFERENCES

[1] P. Ahgren, "Acoustic echo cancellation and doubletalk detection using estimated loudspeaker impulse responses,"

IEEE Transactions on Speech and Audio Processing, vol. 13, pp. 1231–1237, 2005.

[2] C. Faller and C. Tournery, "Robust acoustic echo control using a simple echo path model," *in Proc. ICASSP*, pp. 281–284, 2006.

[3] R. Korde and S. Sahare, "Acoustic echo cancellation using independent component analysis," *International Journal of Advances in Engeneering and Technology*, vol.4, pp. 436–442, 1963.

[4] E. Hansler and G. Schmidt, *Topics in Acoustic Echo and Noise Control*, Chapter5, Springer-Verlag, 2006.

[5] T. S. Wada and B.-H. Juang, "Enhancement of residual echo for robust acoustic echo cancellation," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 20, pp. 175–189, 2012.

[6] T. S. Wada and B.-H. Juang, "Acoustic echo cancellation based on independent component analysis and integrated residual echo enhancement," *in Proc.WASPAA*, pp. 205-208, 2009.

[7] S. Haykin, *Adaptive Filter Theory 4th ed.*, Chapter 2. Prentice-Hall, 2003.

[8] P. Vary and R. Martin, *Digital Speech Transmission*, Chapter 9. Error concealment and soft decision source decoding, John Wiley and Sons, 2006.

[9] T. v. Waterschoot and M. Moonen, "Fifty years of acoustic feedback control: state of the art and future challenges," *Proceedings of the IEEE*, vol. 99, pp. 288 - 327, 2011.

[10] B. Widrow and S. D. Stearns, *Adaptive Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1985.

[11] Y.-R. Chien and S.-I Chu, "A fast converging partial update LMS algorithm with random combining strategy," *Circuits, Systems, and Signal Processing*, vol. 33, no. 6, pp. 1883-1898, 2014.

[12] E. Soria, J. Calpe, J. Chambers, M. Martinez, G. Camps, and J. D. M. Guerrero, "A novel approach to introducing adaptive filters based on the LMS algorithm and its variants," *IEEE Transactions on Education*, vol. 47, pp. 127-133, 2008.

[13] A. Tandon, M. O. Ahmad, and M. N. S. Swamy, "An efficient, low-complexity, normalized LMS algorithm for echo cancellation," in *Proc. NEWCAS*, pp. 161-164, 2004.

[14] A. Feuer and E. Weinstein, "Convergence analysis of LMS filters with uncorrelated Gaussian data," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. ASSP-33, pp. 222-230, 1985.

[15] Y.-R. Chien and W.-J. Zeng, "Switching-based variable step-size approach for partial update LMS algorithms," *Electronics Letters*, vol. 49, no. 17, pp.1801-1803, 2013.

[16] J. Benesty, H. Rey, L. R. Vega, and S. Tressens, "A nonparametric VSS NLMS algorithm," *IEEE Signal Processing Letters*, vol. 13, pp. 581–584, 2006.

[17] H.-C. Huang and J. Lee, "A new variable step-size NLMS algorithm and its performance analysis," *IEEE Transactions on Signal Processing*, vol. 60, pp. 2055–2060, 2012.

[18] H.-C. Shin, A. H. Sayed, and W.-J. Song, "Variable step-size NLMS and affine projection algorithms," *IEEE Signal Processing Letters*, vol. 11, pp. 132–135, 2004.

[19] L. Liao and A. W. Khong, "Sparseness-controlled affine projection algorithm for echo cancelation," *in Proc. APSIPA*, pp. 355–361, 2010.

[20] J. M. Gil-Cacho, T. van Waterschoot, M. Moonen, and S. H. Jensen, "Nonlinear acoustic echo cancellation based on a parallel-cascade kernel affine projection algorithm," *in Proc. ICASSP*, pp. 33–36, 2012.

[21] K. Y. Hwang and W. J. Song, "An affine projection adaptive filtering algorithm with selective regressors," *IEEE Transactions on Circuits and Systems*, vol. 54, 2007.

[22] C. Paleologu, J. Benesty, F. Albu, and S. Ciochina, "An efficient variable step-size proportionate affine projection algorithm," *in Proc. ICASSP*, pp. 77–80, 2011.

[23] T. Gansler, J. Benesty, S. L. Gay, and M. M. Sondhi, "A robust proportionate affine projection algorithm for network echo cancellation," *in Proc. ICASSP*, pp. 793–796, 2000.

[24] C. Paleologu, S. Ciochina, and J. Benesty, "An efficient proportionate affine projection algorithm for echo cancellation," *IEEE Signal Processing Letters*, vol. 17, pp. 165–168, 2010.

[25] J. Benesty, C. Paleologu, and S. Ciochina, "Proportionate adaptive filters from a basis pursuit perspective," *IEEE Signal Processing Letters*, vol. 17, pp. 985–988, 2010.

[26] K. Sakhnov, "An improved proportionate affine projection algorithm for network echo cancellation," *in Proc. IWSSIP*, pp. 125–128, 2008.

[27] F. Albu, H. Coanda, D. Coltuc, and M. Rotaru "Intermittently updated simplified proportionate affine projection algorithm," in Proc. ADAPTIVE 2014, pp.42-47, 2014.

[28] F. Albu, C. Paleologu, J. Benesty, and S. Ciochina, "A low complexity proportionate affine projection algorithm for echo cancellation," in Proc. EUSIPCO 2010, pp. 6-10, 2010.

[29] H. Zhao, Y. Yu, S. Gao, X. Zeng, and Z. He, "Memory proportionate APA with individual activation factors for acoustic echo cancellation," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 22, pp. 1047-1055, 2014.

[30] T. v. Waterschoot, G. Rombouts, and M. Moonen, "Optimally regularized adaptive filtering algorithms for room acoustic signal enhancement," *Signal Processing*, vol. 88, pp. 594–611, 2008.

[31] C. Paleologu, J. Benesty, and S. Ciochina, "Practical regularization of the affine projection algorithm," in *Proc. EUSIPCO*, pp. 704-708, 2011.

[32] S. C. Chan, Y. J. Chu, and Z. G. Zhang, "A new variable regularized transform domain NLMS adaptive filtering algorithm—acoustic applications and performance analysis," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 21, pp. 868-878, 2013.

[33] R. Kuhn, J.-C. Junqua, P. Nguyen, and N. Niedzielski, "Rapid speaker adaptation in eigenvoice space," *IEEE Transactions on Speech and Audio Processing*, vol. 8, no. 6, pp. 695–707, 2000.

[34] Y. Tsao and C.-H. Lee, "An ensemble speaker and speaking environment modeling approach to robust speech recognition," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 17, no. 5, pp. 1025–1037, 2009.

[35] P. N. Belhumeur, J. P. Hespanha, and D. Kriegman, "Eigenfaces vs. Fisher-faces: recognition using class specific linear projection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 711–720, 1997.

[36] E. A. Habets, "Room impulse response generator," *Tech. Rep, Technische Universiteit Eindhoven*, vol. 2, pp. 1–21, 2006.

[37] N. Parihar, J. Picone, D. Pearce, and H. G. Hirsch, "Performance analysis of the Aurora large vocabulary baseline system," *in Proc. EUSIPCO*, pp. 553–556, 2004.

[38] G. Hirsch, "Experimental framework for the performance evaluation of speech recognition front-ends on a large vocabulary task," *ETSI STQ Aurora DSR Working Group*, 2001.

[39] N. Parihar and J. Picone, "Aurora working group: DSR front end LVCSR evaluation au/384/02," *Institute for Signal and Information Processing Report*, 2002.

[40] P. Vandewalle, J. Kovacevic, and M. Vetterli, "Reproducible research in signal processing," *IEEE Signal Processing Magazine*, vol. 26(3), pp.37-47, 2009.

[41] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*. New York, NY, USA: Wiley, 2001.