AMSE JOURNALS-2016-Series: Advances B; Vol. 59; N°1; pp 164-176
Submitted July 2016; Revised Oct. 15, 2016, Accepted Dec. 10, 2016

A New Subspace Based Speech Enhancement Algorithm with Low Complexity

Qi Sun[1, 2], Xiaohui Zhao[1]

[1] College of Communication Engineering, Jilin University, Changchun 130012, China (sq227@126.com)
[2] School of Electronics and Information Engineering, Shenyang Aerospace University, Shenyang 110136, China

Abstract

A subspace based speech enhancement algorithm with low complexity is proposed by using our improved subspace iteration method. Due to the use of the iteration, the proposed algorithm can greatly reduce its number of calculation or its complexity with satisfied performance in speech enhancement in the case of both white and color noise, which proves that it is more suitable for real time implementation. Simulation results show that the proposed algorithm can achieves better speech enhancement compared with classical subspace based algorithm in low input SNR.

Key words

Speech enhancement, subspace iteration, eigenvalue and eigenvector estimation

1. Introduction

Speech enhancement is to improve quality of speech communication system when speech is polluted by different kinds of noise interference in real environment by suppressing bad influence of the noise to extract pure primitive speech signal as clean as possible for speech intelligibility and integrity. However, speech enhancement algorithms can suppress noise in the noisy speech and inevitably bring speech distortion to the original clean speech. This implies the contradiction and balance of how to develop a speech enhancement algorithm with better performance of speech enhancement but little destruction for the original speech. Various speech enhancement algorithms have been proposed to find the solutions in consideration of these two aspects with simple and easy implementation for practical situations.

There are two major classes of speech enhancement algorithms: the time domain methods, for example subspace speech enhancement algorithms [1-6], and the frequency domain methods, such as the spectral subtraction algorithms [7-9]. With the rapid development of speech processing technology, many new methods have emerged, for instance wavelet transforms method [10], and auditory masking method [11].

The subspace based speech enhancement methods normally have the advantage of low distortion in enhanced speech and small speech residual noise. Thus these methods are widely used in speech processing. First subspace speech enhancement algorithm to deal with white noise in speech enhancement is proposed in [1]. A suboptimal estimation of speech signal through approximate diagonalization of noise covariance matrix using KLT is given in [2]. Hu Y and Loizou proposed a subspace speech enhancement algorithm by optimal speech estimation for both white noise and colored noise is provided in [3]. This algorithm, considered as a classical one, simultaneously conducts diagonalization of clean speech and noise covariance matrices through non-unitary transform using eigen decomposition, so that the noisy speech signal space is decomposed into two orthogonal subspaces, namely the signal subspace and the noise subspace. The decomposition in the algorithm requires $O(K^3)$ computational complexity where K is the frame length of the sampled speech data, which means that this classical algorithm needs more calculation when K is large.

To reduce heavy calculation burden and improve performance of the speech enhancement, a new subspace based speech enhancement algorithm with low complexity is proposed in this paper. This algorithm uses the subspace iteration method, widely used in engineering applications, which is effective to solve large generalized eigenvalue problem [12] and [13]. By repeatedly using one-dimensional subspace iteration, all eigenvalues and corresponding eigenvectors of speech data matrix can be iteratively estimated in order so as to realize speech enhancement.

In this paper, Section 2 presents the classical subspace speech enhancement algorithm, which is the basis of our proposed algorithm. Section 3 describes the improved subspace iteration method, a new iterative estimation for eigenvalues and corresponding eigenvectors of speech data matrix is derived for the proposed speech enhancement algorithm. Section 4 gives computer simulation and its analysis. Finally, the conclusion of the paper is provided.

2. Classical Subspace Speech Enhancement Algorithm

Assuming the noisy speech signal at time n can be expressed as

$$\boldsymbol{y}(n) = \boldsymbol{x}(n) + \boldsymbol{d}(n) \tag{1}$$

where $\boldsymbol{y}(n)$ represents the K-dimensional noisy speech vector, $\boldsymbol{x}(n)$ represents the K-dimensional clean speech vector, $\boldsymbol{d}(n)$ represents the K-dimensional additive noise vector uncorrelated with the clean speech vector, namely

$$E\left[x(i)d(j)\right] = 0,\ \forall i, j \tag{2}$$

$$\boldsymbol{R}_y(n) = E\left[\boldsymbol{y}(n)\boldsymbol{y}(n)^T\right] = E\left[\boldsymbol{x}(n)\boldsymbol{x}(n)^T\right] + E\left[\boldsymbol{d}(n)\boldsymbol{d}(n)^T\right] = \boldsymbol{R}_x(n) + \boldsymbol{R}_d(n) \tag{3}$$

where $\boldsymbol{R}_y(n)$, $\boldsymbol{R}_x(n)$ and $\boldsymbol{R}_d(n)$ are the $K\times K$ covariance matrices of $\boldsymbol{y}(n)$, $\boldsymbol{x}(n)$ and $\boldsymbol{d}(n)$ respectively.

Assuming the estimation of the clean speech signal is as follows

$$\hat{\boldsymbol{x}}(n) = \boldsymbol{H}(n)\boldsymbol{y}(n) \tag{4}$$

$\boldsymbol{H}(n)$ represents a $K\times K$ linear estimation matrix, the optimal estimation matrix can be obtained by using the Lagrange multiplier method, that is

$$\boldsymbol{H}_{opt}(n) = \boldsymbol{R}_x(n)(\boldsymbol{R}_x(n) + \mu(n)\boldsymbol{R}_d(n))^{-1} \tag{5}$$

where $\mu(n)$ is the Lagrange multiplier.

Eq. (5) can be rewritten by the eigen decomposition of $\boldsymbol{R}_x(n) = \boldsymbol{U}(n)\boldsymbol{\Delta}_x(n)\boldsymbol{U}^T(n)$ as

$$\boldsymbol{H}_{opt}(n) = \boldsymbol{U}(n)\boldsymbol{\Delta}_x(n)(\boldsymbol{\Delta}_x(n) + \mu(n)\boldsymbol{U}^T(n)\boldsymbol{R}_d(n)\boldsymbol{U}(n))^{-1}\boldsymbol{U}^T(n) \tag{6}$$

where $\boldsymbol{U}(n)$ is the eigenvector matrix of $\boldsymbol{R}_x(n)$, $\boldsymbol{\Delta}_x(n)$ is the diagonal eigenvalue matrix of $\boldsymbol{R}_x(n)$.

When $\boldsymbol{d}(n)$ is white noise with variance $\sigma_d^2(n)$, substituting $\boldsymbol{R}_d(n)=\sigma_d^2(n)\boldsymbol{I}$ into Eq. (6), we have [1]

$$\boldsymbol{H}_{opt}(n)=\boldsymbol{U}(n)\boldsymbol{\Lambda}_x(n)(\boldsymbol{\Lambda}_x(n)+\mu(n)\sigma_d^2(n)\boldsymbol{I})^{-1}\boldsymbol{U}^T(n) \tag{7}$$

When $\boldsymbol{d}(n)$ is color noise, simulation experiments show that $\boldsymbol{U}^T(n)\boldsymbol{R}_d(n)\boldsymbol{U}(n)$ is not usually a diagonal matrix. But we can use an approximated diagonal matrix $\boldsymbol{\Lambda}_d(n)$ and substitute $\boldsymbol{\Lambda}_d(n)$ into Eq. (6), thus suboptimal estimation matrix can be obtained as follows [2]

$$\boldsymbol{H}(n)=\boldsymbol{U}(n)\boldsymbol{\Lambda}_x(n)(\boldsymbol{\Lambda}_x(n)+\mu(n)\boldsymbol{\Lambda}_d(n))^{-1}\boldsymbol{U}^T(n) \tag{8}$$

For the color noise, an optimal estimation method is proposed in [3], which proves the existence of the matrix $\boldsymbol{V}(n)$ such that

$$\boldsymbol{V}^T(n)\boldsymbol{R}_x(n)\boldsymbol{V}(n)=\boldsymbol{\Lambda}_\Sigma(n),\boldsymbol{V}^T(n)\ \boldsymbol{R}_d(n)\boldsymbol{V}(n)=\boldsymbol{I} \tag{9}$$

where $\boldsymbol{V}(n)$ is the eigenvector matrix and $\boldsymbol{\Lambda}_\Sigma(n)$ is the corresponding eigenvalue diagonal matrix of $\boldsymbol{\Sigma}(n)=\boldsymbol{R}_d^{-1}(n)\boldsymbol{R}_x(n)=\boldsymbol{R}_d^{-1}(n)(\boldsymbol{R}_y(n)-\boldsymbol{R}_d(n))=\boldsymbol{R}_d^{-1}(n)\boldsymbol{R}_y(n)-\boldsymbol{I}$ respectively. The optimal estimation matrix can be obtained under the background of the color noise by Eq. (5) and Eq. (9) [3], that is

$$\boldsymbol{H}_{opt}(n)=\boldsymbol{V}^{-T}(n)\boldsymbol{\Lambda}_\Sigma(n)(\boldsymbol{\Lambda}_\Sigma(n)+\mu(n)\boldsymbol{I})^{-1}\boldsymbol{V}^T(n) \tag{10}$$

Therefore we obtain the classical subspace speech enhancement algorithm with Eq. (4) and Eq. (10).

In fact, Eq. (7) is a special form of Eq. (10) which is a common expression of the subspace methods. The useful advantage of the methods is that it is not necessary to estimate noise characteristics, either for white noise or color noise.

We can see that the key to solve the problem of the classical algorithm is the eigen decomposition of the matrix $\boldsymbol{\Sigma}(n)$. Since the computational complexity of this decomposition is $O(K^3)$, the classical algorithm has large computation burden and is not practical or suitable to its real-time implementation. In order to reduce the computational complexity of the algorithm, improve the real-time capability of the algorithm, a new iterative estimation for the eigenvalues and the eigenvectors of $\boldsymbol{\Sigma}(n)$ is derived and applied to the classical subspace speech enhancement algorithm.

3. New Speech Enhancement Algorithm

3.1 Iterative Estimation for Eigenvalues and Eigenvectors of $\boldsymbol{\Sigma}(n)$

In this section, we will derive a matrix estimation by the subspace iteration method which is an effective way to obtain the largest eigenvalues and the corresponding eigenvectors of large sparse matrix with high reliability.

For the *K*-dimensional noisy speech vector $\boldsymbol{y}(n)$, let p and r be a positive integer, and $r < p < K$, the specific steps of the subspace iteration method to calculate r largest eigenvalues and the corresponding eigenvectors of the speech data matrix $\boldsymbol{\Sigma}(n)$ are provided in [12] and [13]. The summary of the method is as follows:

(1) In initialization, we randomly select p vectors, get $K \times p$ column orthonormal initial matrix $\boldsymbol{A}(0) = [\boldsymbol{a}_1(0), \boldsymbol{a}_2(0), \cdots, \boldsymbol{a}_p(0)]$ through orthogonalization.

(2) We assume that the iterative matrix $\boldsymbol{A}(n-1)$ is obtained after n-1 iterations, then the nth iteration is

a. $\boldsymbol{A}(n) = \boldsymbol{\Sigma}(n)\boldsymbol{A}(n-1)$, $\boldsymbol{A}(n) = [\boldsymbol{a}_1(n), \boldsymbol{a}_2(n), \cdots, \boldsymbol{a}_p(n)]$.

b. *QR* decomposition, $\boldsymbol{A}(n) = \boldsymbol{Q}(n)\boldsymbol{R}(n)$, where $\boldsymbol{Q}(n)$ is $K \times p$ column orthonormal matrix, $\boldsymbol{R}(n)$ is upper triangular matrix of order p.

c. Construction of $\boldsymbol{B}(n) = \boldsymbol{Q}^T(n)\boldsymbol{\Sigma}(n)\boldsymbol{Q}(n)$.

d. Through eigen decomposition, the eigenvalue $\lambda_i(n)$ and the corresponding eigenvector $\boldsymbol{w}_i(n)$ of $\boldsymbol{B}(n)$ are calculated, then we can define $\boldsymbol{D}(n) = \mathrm{diag}(\lambda_1(n), \lambda_2(n), \cdots, \lambda_p(n))$, $\boldsymbol{W}(n) = [\boldsymbol{w}_1(n), \boldsymbol{w}_2(n), \cdots, \boldsymbol{w}_p(n)]$.

e. Calculation of the vector matrix $\boldsymbol{V}(n) = \boldsymbol{Q}(n)\boldsymbol{W}(n) = [\boldsymbol{v}_1(n), \boldsymbol{v}_2(n), \cdots, \boldsymbol{v}_p(n)]$.

f. Convergence evaluation, if $\|\boldsymbol{\Sigma}(n)[\boldsymbol{v}_1(n), \cdots, \boldsymbol{v}_r(n)] - [\boldsymbol{v}_1(n), \cdots, \boldsymbol{v}_r(n)]\boldsymbol{D}_1(n)\|_F < \varepsilon$, the end, otherwise continue. When $\boldsymbol{\Sigma}(n)$ is a symmetric matrix, then $\boldsymbol{A}(n) = \boldsymbol{V}(n)$, when $\boldsymbol{\Sigma}(n)$ is a asymmetric matrix, $\boldsymbol{A}(n)$ can be obtained by orthogonal normalization for $\boldsymbol{V}(n)$.

Where $\boldsymbol{D}_1(n) = \text{diag}(\lambda_1(n), \lambda_2(n), \cdots, \lambda_r(n))$, and $\lambda_1(n), \lambda_2(n), \cdots, \lambda_r(n)$ are the r largest eigenvalues of $\boldsymbol{\Sigma}(n)$, $\boldsymbol{v}_1(n), \boldsymbol{v}_2(n), \cdots, \boldsymbol{v}_r(n)$ are the corresponding eigenvectors. $\|\bullet\|_F$ is the Frobenius norm. For the special case of $p = 1$, the results of the iteration are the largest eigenvalue and the corresponding eigenvector of $\boldsymbol{\Sigma}(n)$.

In the iterative process of the subspace iteration method, the amount of calculation for the eigenvalues and the corresponding eigenvectors of $\boldsymbol{B}(n)$ is large. In this process, we do not make the eigen decomposition of $\boldsymbol{\Sigma}(n)$ and directly estimate the eigenvectors, but we get a set of standard orthogonal basis of $\boldsymbol{\Sigma}(n)$.

In order to solve the above problems, through the improvement of the subspace iteration method, a new iterative estimation for the eigenvalues and the eigenvectors of $\boldsymbol{\Sigma}(n)$ is derived in [14-19]. In this estimation, the subspace iteration with $p = 1$ is performed, the result of the iteration is the largest eigenvalue and the corresponding eigenvector of $\boldsymbol{\Sigma}(n)$ known as the largest eigenvector of $\boldsymbol{\Sigma}(n)$. Then we remove the projection of $\boldsymbol{\Sigma}(n)$ onto this eigenvector from $\boldsymbol{\Sigma}(n)$ itself. Since the second eigenvalue and the corresponding eigenvector now become the largest eigenvalue and the corresponding eigenvector in the updated matrix $\boldsymbol{\Sigma}(n)$, they can be extracted by the continued application of the subspace iteration method with $p = 1$. By repeatedly applying this procedure, all the eigenvalues and the corresponding eigenvectors of $\boldsymbol{\Sigma}(n)$ can be estimated. This algorithm is as follows:

Initialize $\boldsymbol{a}_i(0)$

For each time step, do

$$\boldsymbol{\Sigma}_1(n) = \boldsymbol{\Sigma}(n) \tag{11}$$

FOR $i = 1, 2, \cdots, K$

$$\boldsymbol{a}_i(n) = \boldsymbol{\Sigma}_i(n)\boldsymbol{a}_i(n-1) \tag{12}$$

$$\boldsymbol{v}_i(n) = \boldsymbol{a}_i(n) / \|\boldsymbol{a}_i(n)\| \tag{13}$$

$$\lambda_i(n) = \boldsymbol{v}_i^T(n)\boldsymbol{\Sigma}_i(n)\boldsymbol{v}_i(n) \tag{14}$$

$$e_i(n) = \boldsymbol{\Sigma}_i(n)\boldsymbol{v}_i(n) - \boldsymbol{v}_i(n)\lambda_i(n) \tag{15}$$

$$\boldsymbol{a}_i(n) = \boldsymbol{v}_i(n) \tag{16}$$

$$\boldsymbol{\Sigma}_{i+1}(n) = \boldsymbol{\Sigma}_i(n) - \boldsymbol{v}_i(n)\boldsymbol{v}_i^T(n)\boldsymbol{\Sigma}_i(n) \tag{17}$$

END

where $\lambda_i(n)$ and $\boldsymbol{v}_i(n)$ are the i^{th} eigenvalue and the corresponding eigenvector of $\boldsymbol{\Sigma}(n)$. We define $\boldsymbol{\Lambda}_\Sigma(n) = diag\{\lambda_1(n), \lambda_2(n), \cdots, \lambda_K(n)\}$ and $\boldsymbol{V}(n) = [\boldsymbol{v}_1(n), \boldsymbol{v}_2(n), \cdots, \boldsymbol{v}_K(n)]$ for the later use. Eq. (17) shows that each iteration removes the projection of $\boldsymbol{\Sigma}_i(n)$ onto the i^{th} eigenvector from $\boldsymbol{\Sigma}_i(n)$ itself. The initial vector $\boldsymbol{a}_i(0)$ can be set to the K-dimensional unit vector to avoid orthogonal initialization in the original algorithm, compared with the original algorithm, the improved algorithm has the following advantages:

(1) $\boldsymbol{A}(n)$ is derived through iteration instead of *QR* decomposition in step (b).

(2) $\boldsymbol{B}(n)$ is derived by the transformation of directly obtaining $\lambda_i(n)$ in the improved algorithm instead of eigen decomposition in step (d).

(3) Convergence depending on the evaluation of $e_i(n)$ instead of calculation of the Frobenius norm in step (f), which reduces the computation complexity.

3.2 New Speech Enhancement Algorithm

Through the improved subspace iteration method, the eigenvalues and the eigenvectors of $\boldsymbol{\Sigma}(n)$ can be obtained and applied to the subspace speech enhancement presented in the above section. For the K-dimensional noisy speech vector $\boldsymbol{y}(n)$, the new proposed improved speech enhancement algorithm is given in the following by introducing new iterative method given from Eq. (11) to Eq. (17):

(1) Computation of $\boldsymbol{R}_y(n)$, and estimation of $\boldsymbol{\Sigma}(n) = \boldsymbol{R}_d^{-1}(n)\boldsymbol{R}_y(n) - \boldsymbol{I}$, where $\boldsymbol{R}_d(n)$ is computed by using noise samples collected during speech-absent frames.

(2) Estimation of $\boldsymbol{V}(n)$ and the corresponding $\boldsymbol{\Lambda}_\Sigma(n)$.

(3) Since the obtained eigenvalues are in descending order, i.e. $\lambda_1(n) \geq \lambda_2(n) \geq \cdots \geq \lambda_K(n)$, the dimension estimation of the speech signal subspace is

$$M(n) = \arg\max_{1 \leq i \leq K} \{\lambda_i(n) > 0\} \tag{18}$$

(4) Computation of $\mu(n)$ as

$$\mu(n) = \begin{cases} \mu_0 - (\mathrm{SNR}_{\mathrm{dB}}(n))/s, & -5 < \mathrm{SNR}_{\mathrm{dB}}(n) < 20 \\ 1 & \mathrm{SNR}_{\mathrm{dB}}(n) \geq 20 \\ 5 & \mathrm{SNR}_{\mathrm{dB}}(n) \leq -5 \end{cases} \tag{19}$$

where $\mu_0 = 4.2$, $s = 6.25$, $\mathrm{SNR}_{\mathrm{dB}}(n) = 10\log_{10}\left(\sum_{i=1}^{M(n)} \lambda_i(n) \Big/ K \right)$.

(5) Calculation of the optimal linear estimation matrix

$$\boldsymbol{H}_{opt}(n) = \boldsymbol{V}^{-T}(n) \begin{bmatrix} \boldsymbol{G}_1(n) & 0 \\ 0 & 0 \end{bmatrix} \boldsymbol{V}^T(n) \tag{20}$$

Where

$$\boldsymbol{G}_1(n) = \mathrm{diag}\{g_{11}(n), g_{22}(n), \cdots, g_{M(n)M(n)}(n)\} \tag{21}$$

$$g_{ii}(n) = \begin{cases} \dfrac{\lambda_i(n)}{\lambda_i(n) + \mu(n)}, & i = 1, 2, \ldots, M(n) \\ 0, & i = M(n)+1, \ldots, K \end{cases} \tag{22}$$

(6) Then the enhanced speech signal is

$$\hat{\boldsymbol{x}}(n) = \boldsymbol{H}_{opt}(n)\boldsymbol{y}(n) \tag{23}$$

From the above algorithm, we can see that the computational complexity of this speech enhancement algorithm by the derived iterative estimation for the eigenvalues and the eigenvectors of $\boldsymbol{\Sigma}(n)$ reduces to $O(K^2)$. When K is large, the superiority of its low complexity is more the obvious. In addition, the characteristic of the background noise is not required in the proposed algorithm, so that the proposed algorithm is suitable and effective for different type of noise.

4. Simulation Results and Analysis

In this section, the performance comparison of the proposed speech enhancement algorithm and the algorithm given in [3] is provided by computer simulation experiments. In these experiments, we use 24 seconds clean speech consisting of 6 male and 6 female test speakers, white noise and buccaneer noise taken from the NOISEX-92 database, and 16kHz sampling frequency. The noisy speech signal is with SNR of -10,-5,0,5,10 dB respectively, and divided into frames under the frame overlap of 50%. The frame length of the classical algorithm is K=40, while K=40,80 for the proposed algorithm for the performance analysis. The log spectral distance (LSD), the segmental signal-to-noise ratio (SegSNR), and the perceptual evaluation of speech quality (PESQ) are used to evaluate the performance of the speech enhancement [20].

The LSD of the two algorithms is shown in Fig.1, in which we can see that the LSD of the proposed algorithm is smaller than that of the classical algorithm in low SNR (-10dB, -5dB, 0dB). This means that the enhanced speech by the proposed algorithm is more close to the clean speech with little distortion. If the input SNR is more than 5dB, the LSD of both algorithms increases slightly. In addition, if we increase the frame length of the proposed algorithm to K=80, the LSD of this algorithm is even better, but the calculation number is much more less than that of the classical algorithm.

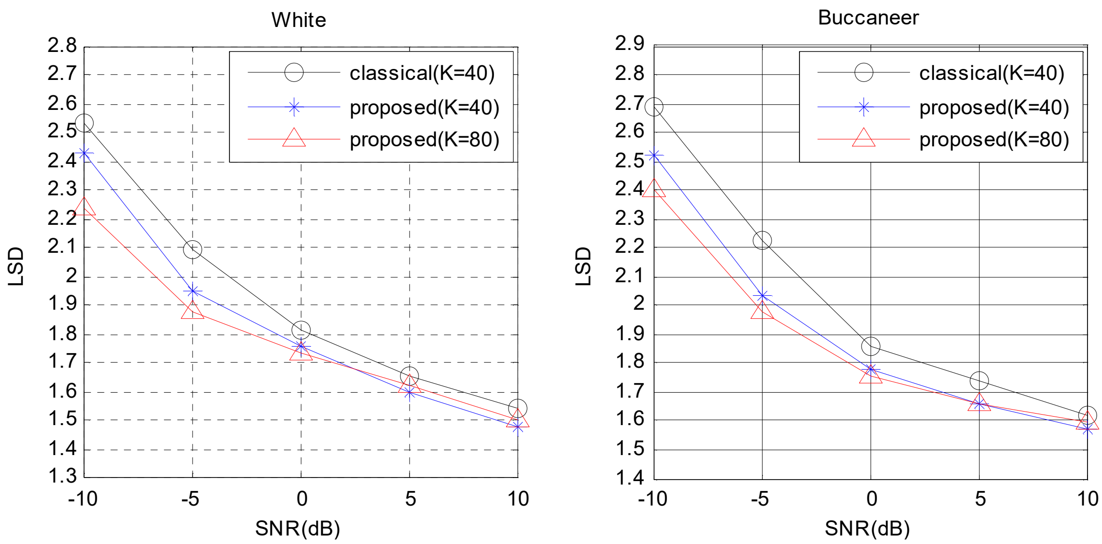

Fig.1. LSD of two algorithms

Table 1 gives the experiment results of the SegSNR of two different algorithms. From this table, we find that the proposed algorithm is better than the classical algorithm by about 0.25dB increased SegSNR in white noise with K=40. While in the case of buccaneer noise, the proposed algorithm is also better than the classical algorithm by approximate 0.2dB increased SegSNR.

Specifically, the SegSNR improvement of the proposed algorithm is significant in low SNR (-10dB, -5dB, 0dB), for example when SNR is -10dB of buccaneer noise, there is about 0.6dB increase of SegSNR. If we increase the frame length to *K*=80, the SegSNR of the proposed algorithm increase 0.2~0.3dB on the basis of the original improvement.

The PESQ of these two algorithms is given in Fig.2, in which the PESQ score ranges from -0.5 to 4.5, and high score represents better speech quality. From Fig.2, we can see that the PESQ score of the proposed algorithm is very close to that of the classical algorithm when *K*=40, but this score of the proposed algorithm significantly improves when *K*=80, which implies that it increases with the increase of *K*. The low calculation number of our proposed algorithm gives large space for higher *K* to improve the PESQ performance. Therefore, this simulation results mean that the subjective listening for the enhanced speech by the proposed algorithm and the clean speech is quite similar in auditory perception.

Table .1. SegSNR of two algorithms

Noise type	SNR(dB)	Classical *K*=40	proposed *K*=40	proposed *K*=80
White	-10	-2.3636	**-2.006**	**-1.7664**
	-5	0.1083	**0.4086**	**0.7132**
	0	2.7632	**3.0172**	**3.3622**
	5	5.5425	**5.7520**	**6.0915**
	10	8.3562	**8.5007**	**8.7289**
Buccaneer	-10	-3.0615	**-2.4869**	**-2.2841**
	-5	-1.0129	**-0.6473**	**-0.4828**
	0	1.3421	**1.4256**	**1.6259**
	5	3.7138	**3.7208**	**3.9409**
	10	6.4844	**6.4872**	**6.6739**

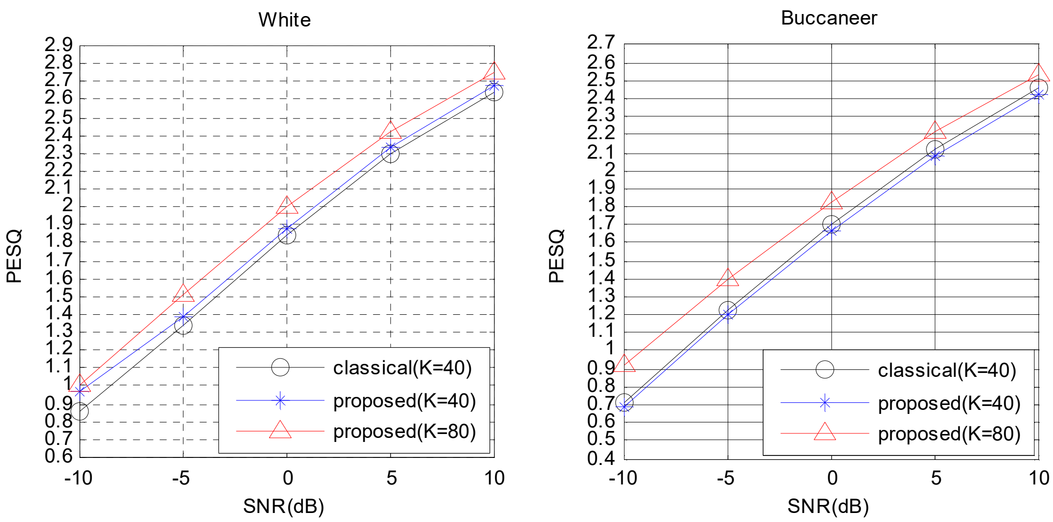

Fig.2. PESQ of two algorithms

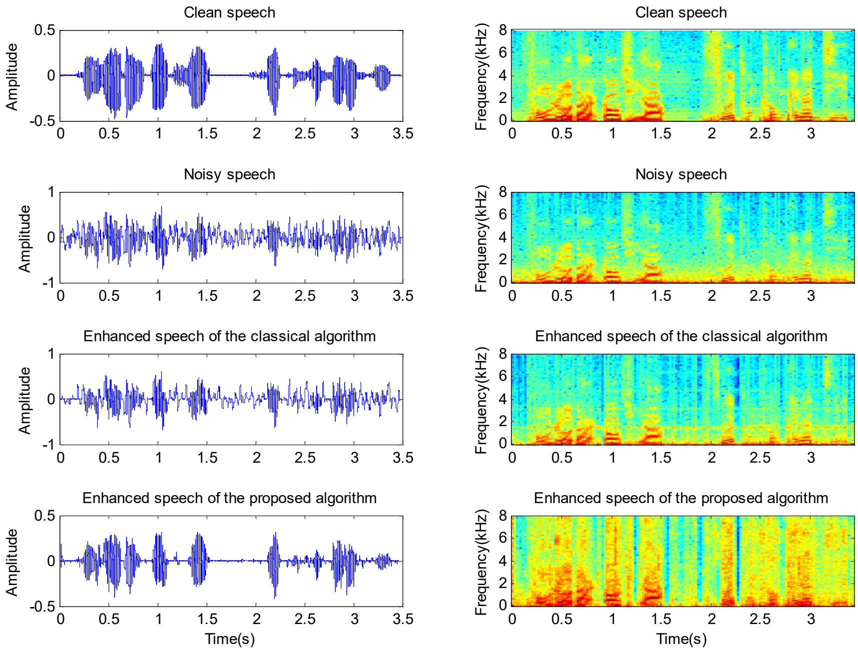

Fig.3. Time-domain waveforms and spectrograms of two algorithms with -5dB Volvo noise

In order to intuitively show the differences between the enhanced speech and the clean speech, we present time-domain waveforms and spectrograms of the clean speech and the noisy speech with the Volvo noise of -5 dB SNR under K=80 for the proposed algorithm in Fig.3. We find that the enhanced speech is more close to the clean speech. But the classical algorithm basically cannot well recover detail frequency component of the original clean speech and also makes a lot of energy loss.

Discussion

The proposed algorithm does not need make the eigenvalue decomposition, which has the advantage of a small amount of computation, low complexity and simple implementation. It is effective for both white noise or colored noise and is an optimal estimation algorithm. The proposed algorithm will have an important practical value in speech signal processing.

Conclusion

A new low calculation complexity speech enhancement algorithm based on subspace approach in iterative form is proposed without eigen decomposition. It is valid for both white and color noise with better performance in several speech quality evaluation criteria. The low complexity provides more capability for practical implementation in real speech applications, and the increase of the frame length can improve the quality of speech enhancement in less number of iterative calculations for the classical speech enhancement algorithm. In low SNR, the proposed algorithm performs better and proves its superiority if we consider the balance between SNR and the choice of the frame length to obtain acceptable enhanced speech quality.

References

1. Y. Ephraim, H. L. Van Trees, A signal subspace approach for speech enhancement, 1995, IEEE Transactions on Speech and Audio Processing, vol. 3, no. 4, pp. 251-266.
2. A. Rezayee, S. Gazor, An adaptive KLT approach for speech enhancement, 2001, IEEE Transactions on Speech and Audio Processing, vol. 9, no. 2, pp. 87-95.
3. Y. Hu, P. C. Loizou, A generalized subspace approach for enhancing speech corrupted noise by colored noise, 2003, IEEE Transactions on Speech and Audio Processing, vol. 11, no. 4, pp. 334-341.
4. A. Saadoune, A. Amrouche, S. A. Selouani, Perceptual subspace speech enhancement using variance of the reconstruction error, 2014, Digital Signal Processing: A Review Journal, vol. 24, no. 1, pp.187-196.
5. J. R. Jensen, J. Benesty, M. G. Christensen, Jingdong Chen, A class of optimal rectangular filtering matrices for single-channel signal enhancement in the time domain, 2013, IEEE Transactions on Audio, Speech, and Language Processing, vol. 21, no. 12, pp.2595-2606.
6. A. Saadoune, A. Amrouche, S. A. Selouani, MCRA noise estimation for KLT-VRE-based speech enhancement, 2013, International Journal of Speech Technology, vol. 16, no. 3, pp. 333-339.
7. S. F. Boll, Suppression of acoustic noise in speech using spectral subtraction, 1979, IEEE Trans ASSP, vol. 27, no. 2, pp. 113-120.
8. S. Mosayyebpour, M. Esmaeili, T. A. Gulliver, Single-microphone early and late reverberation suppression in noisy speech, 2013, IEEE Transactions on Audio, Speech, and Language Processing, vol. 21, no. 2, pp. 322-335.
9. S. Singh, M. Tripathy, R. S. Anand, Single channel speech enhancement for mixed non-stationary noise environments, 2014, Advances in Signal Processing and Intelligent Recognition Systems, vol. 264, pp. 545-555.
10. T. F. Sanam, C. Shahnaz, Noisy speech enhancement based on an adaptive threshold and a modified hard thresholding function in wavelet packet domain, 2012, Digital Signal Processing, vol. 23, no. 3, pp. 941-951.

11. R. Yu, Speech enhancement based on soft audible noise masking and noise power estimation, 2013, Speech Communication, vol. 55, no. 10, pp. 964-974.
12. K. J. Bathe, Solution methods for large generalized eigenvalue problems in structural engineering, 1971, U C SESM Report 71-20 Department of Civil Engineering, University of California, Berkeley, CA.
13. G. H. Golub, C. F. Van Loan, Matrix Computations, The 4th Editon, 2013, The Johns Hopkins University Press, Baltimore, MD.
14. B. Yang, Projection approximation subspace tracking, 1995, IEEE Transaction on Signal Processing, vol. 43, no. 1, pp. 95-107.
15. Y. H. Choi, Robust adaptive beamforming method using principal eigenpairs with modification of PASTd, 2013, Digital Signal Processing: A Review Journal, vol. 23, no. 2, pp. 595-600.
16. O. O. Oyerinde, S. H. Mneney, Subspace tracking-based decision directed CIR estimator and adaptive CIR prediction, 2012, IEEE Transactions on Vehicular Technology, vol. 61, no. 5, pp. 2097-2107.
17. S. C. Chan, H. C. Wu, K. M. Tsui, Robust recursive eigendecomposition and subspace-based algorithms with application to fault detection in wireless sensor networks, 2012, IEEE Transactions on Instrumentation and Measurement, vol. 61, no. 6, pp. 1703-1718.
18. T. Duong Nguyen, I. Yamada, Necessary and sufficient conditions for convergence of the DDT systems of the Normalized PAST algorithms, 2014, Signal Processing, vol. 94, no. 1, pp. 288-299.
19. B. Liao, S. C. Chan, K. M. Tsui, Recursive Steering Vector Estimation and Adaptive Beamforming under Uncertainties, 2013, IEEE Transactions on Aerospace and Electronic Systems, vol. 49, no. 1, pp. 489-501.
20. Y. Hu, P. C. Loizou, Evaluation of objective quality measures for speech enhancement, 2008, IEEE Transactions on Audio, Speech, and Language Processing, vol. 16, no. 1, pp. 229-238.