Subspace Based Speech Enhancement Using Gaussian Mixture Model

Achintya Kundu, Saikat Chatterjee and T.V. Sreenivas

Department of Electrical Communication Engineering
Indian Institute of Science, Bangalore - 560012, India
{achintya,saikat,tvsree}@ece.iisc.ernet.in

Abstract

Traditional subspace based speech enhancement (SSE) methods use linear minimum mean square error (LMMSE) estimation that is optimal if the Karhunen Loeve transform (KLT) coefficients of speech and noise are Gaussian distributed. In this paper, we investigate the use of Gaussian mixture (GM) density for modeling the non-Gaussian statistics of the clean speech KLT coefficients. Using Gaussian mixture model (GMM), the optimum minimum mean square error (MMSE) estimator is found to be nonlinear and the traditional LMMSE estimator is shown to be a special case. Experimental results show that the proposed method provides better enhancement performance than the traditional subspace based methods.

Index Terms: Subspace based speech enhancement, Gaussian mixture density, MMSE estimation.

1. Introduction

In the noisy environment, the estimation of clean speech is a challenging problem and several approaches [1], viz. spectral subtraction, Wiener filtering, statistical model based and subspace based methods, have been explored for speech enhancement (SE). The subspace based SE (SSE) methods are based on the principle that the clean speech signal can be effectively modeled as vectors confined in a lower dimensional subspace[1] of the noisy signal's vector space [2]. Thus, the additive noise can be partially removed by projecting the noisy speech signal onto the signal subspace[2]. The information about the signal subspace is obtained using either the singular value decomposition (SVD) [3]-[5], or the eigen value decomposition (EVD) [2], [6]. Dendrinos *et al.* [3] proposed the SVD based SSE method for speech signal corrupted by white noise. This was extended for colored noise case in [4] using quotient SVD (QSVD).

Ephraim and Van Trees introduced a different SSE method [2] where they minimized the speech distortion subject to the constraint that the residual noise remains under a threshold. In this method, KLT is applied to the noisy signal and the KLT coefficients, which corresponds to the signal subspace, are modified by a gain function. The remaining KLT coefficients, representing noise, are simply nulled. For additive white noise case, the KLT matrix of the clean speech signal is obtained by applying the EVD on the covariance matrix of the noisy speech signal. This approach of Ephraim and Van Trees has been extended for colored noise case in [6]-[9].

The method of projecting the noisy signal onto the signal subspace can only remove the portion of noise that is orthogonal to the signal subspace. Further, to remove the noise component which lies in the signal subspace, the projected noisy speech signal component is enhanced. In the method of Ephraim and Van Trees [2], clean speech KLT coefficients corresponding to the signal subspace are evaluated using the LMMSE estimation approach. The LMMSE estimation based SSE method will provide optimum performance if the respective PDFs of clean speech KLT coefficients and noise KLT coefficients are modeled using Gaussian densities. Recently, it is shown in [10] that the PDF of clean speech KLT coefficients are better modeled using Laplacian distribution; this result leads to the conclusion that the joint PDF of clean speech KLT coefficients is non-Gaussian. For non-Gaussian PDF, the traditional LMMSE estimator is suboptimal. Recently, a super Gaussian density based SSE method is proposed in [11] where the individual KLT coefficients are processed independently. We mention that, for a signal with non-Gaussian PDF, the KLT coefficients are not statistically independent [12]. In this paper, we model the non-Gaussian joint PDF of clean speech KLT coefficients using a multivariate Gaussian mixture (GM) density and develop a vector based estimator unlike the scalar based method of [11]. The motivation for using GM density is attributed to the theoretical result of [13] where it is shown that any multivariate PDF can be approximated arbitrarily closely by a GM density. For the Gaussian mixture model (GMM) based SSE method, we show the optimum MMSE estimator that is nonlinear and the traditional linear MMSE estimator is shown to be a special case. The performance of the developed estimator is shown to be better than the linear estimation based SSE methods of Ephraim and Van Trees [2], and Hu and Loizou [9].

2. GMM based SSE method

We consider the basic model where the clean speech signal is corrupted by additive noise; the noisy speech signal model is given as

$$\mathbf{X} = \mathbf{S} + \mathbf{W}, \tag{1}$$

where $\mathbf{X}$, $\mathbf{S}$ and $\mathbf{W}$ respectively denote $N \times 1$ vectors of noisy speech, clean speech and additive noise. Here, clean speech vector $\mathbf{S}$ and noise vector $\mathbf{W}$ are assumed to be zero mean and statistically independent.

The subspace based speech enhancement (SSE) methods [2] exploit the fact that the correlation matrix of most speech frames (vectors) have certain eigen values that are practically zero. Thus, in the SSE methods, clean speech vectors are assumed to be confined in a K dimensional subspace of the N dimensional Euclidean space, where $K < N$. Let us denote the correlation matrix of clean speech vector $\mathbf{S}$ by $\mathbf{R_s}$. We construct a matrix $\mathbf{U}$ using the orthonormal eigen vectors of $\mathbf{R_s}$, where the eigen vectors are placed column wise in descending order of the eigen values. In the literature, the matrix $\mathbf{U}^T$ is

[1] The subspace occupied by the clean signal is referred to as signal subspace.

[2] Here, the portion of the noise that is orthogonal to the signal subspace is considered.

known as the KLT matrix. Now, we partition $\mathbf{U}$ as

$$\mathbf{U} = [\,\mathbf{U}_1\,\mathbf{U}_2\,], \tag{2}$$

where $\mathbf{U}_1$ and $\mathbf{U}_2$ are respectively of dimension $N \times K$ and $N \times (N-K)$. A set of orthonormal basis vectors of the K dimensional signal subspace is given by the columns of $\mathbf{U}_1$. Therefore, the clean speech signal can be written as

$$\mathbf{S} = \mathbf{U}_1\,\mathbf{\Phi}, \tag{3}$$

where $\mathbf{\Phi}$ is a $K \times 1$ vector that actually contains the first K KLT coefficients of the clean speech vector $\mathbf{S}$. Therefore, it is assumed that the clean speech signal can be effectively reconstructed from the first K KLT coefficients.

Now, combining Eqn. (1) and Eqn. (3), we have the following observation model for the SSE method as

$$\mathbf{X} = \mathbf{U}_1\,\mathbf{\Phi} + \mathbf{W}, \tag{4}$$

where $\mathbf{\Phi}$ and $\mathbf{W}$ are statistically independent as $\mathbf{S}$ and $\mathbf{W}$ are independent. We note that the above model stands as the general Bayesian linear model. Thus, from Eqn. (4), the MMSE estimation of $\mathbf{\Phi}$ can be evaluated as

$$\hat{\mathbf{\Phi}}_{\text{MMSE}} = \mathcal{E}\{\mathbf{\Phi}\,|\,\mathbf{X}\}, \tag{5}$$

and subsequently, the clean speech vector is estimated as

$$\hat{\mathbf{S}}_{\text{MMSE}} = \mathbf{U}_1\;\hat{\mathbf{\Phi}}_{\text{MMSE}} = \mathbf{U}_1\;\mathcal{E}\{\mathbf{\Phi}\,|\,\mathbf{X}\}. \tag{6}$$

In the general Bayesian linear model of Eqn. (4), we use a Gaussian mixture (GM) density with M number of mixture components for modeling the PDF of $\mathbf{\Phi}$ as

$$f_{\mathbf{\Phi}}(\Phi) = \sum_{m=1}^{M} \boldsymbol{\alpha}_m\,\mathcal{N}\left(\boldsymbol{\mu}_{m,\Phi}, \mathbf{\Sigma}_{m,\Phi}\right), \tag{7}$$

where $\boldsymbol{\alpha}_m$ ($\{\boldsymbol{\alpha}_m > 0\}_{m=1}^{M}$ and $\sum_{m=1}^{M}\boldsymbol{\alpha}_m = 1$), $\boldsymbol{\mu}_{m,\Phi}$ and $\mathbf{\Sigma}_{m,\Phi}$ are respectively the prior probability, mean vector and covariance matrix of the mth Gaussian component. On the other hand, the PDF of the noise vector $\mathbf{W}$ is modeled using a multivariate Gaussian density as

$$f_{\mathbf{W}}(\mathbf{w}) = \mathcal{N}\left(\mathbf{0}, \mathbf{C}_{\mathbf{w}}\right), \tag{8}$$

where we make no assumption of whiteness. Now, for evaluating $\mathcal{E}\{\mathbf{\Phi}\,|\,\mathbf{X}\}$, we use our earlier result on GMM based MMSE estimation as given in [14] (see theorem of [14]). Thus, the subspace based MMSE estimator of Eqn. (6) is evaluated as shown in Eqn. (9). From Eqn. (10), we note that $\boldsymbol{\beta}_m(\mathbf{X})$ is a nonlinear function of $\mathbf{X}$ and thus, the estimator of Eqn. (9) is nonlinear. We mention that the traditional linear estimator can be shown as the special case of our estimator as discussed in Appendix.

2.1. Algorithm implementation

For implementing the developed estimator of Eqn. (9), we need to find the signal subspace dimension K and the eigen matrix $\mathbf{U}_1$. Now, to evaluate K and $\mathbf{U}_1$, the clean speech correlation matrix is required to be estimated. From Eqn. (1), we can write the noisy speech correlation matrix as

$$\mathbf{R}_{\mathbf{x}} \triangleq \mathcal{E}\{\mathbf{X}\mathbf{X}^T\} = \mathbf{R}_{\mathbf{s}} + \mathbf{R}_{\mathbf{w}}, \tag{11}$$

where $\mathbf{R}_{\mathbf{s}}$ and $\mathbf{R}_{\mathbf{w}}$ are respectively the correlation matrices of clean speech and noise. Note that the noise covariance matrix $\mathbf{C}_{\mathbf{w}}$ is identical to $\mathbf{R}_{\mathbf{w}}$ as the noise is zero mean. Let us denote the estimates of $\mathbf{R}_{\mathbf{x}}$ and $\mathbf{R}_{\mathbf{w}}$ respectively as $\hat{\mathbf{R}}_{\mathbf{x}}$ and $\hat{\mathbf{R}}_{\mathbf{w}}$. Therefore, using the relationship of Eqn. (11), we get an estimate of $\mathbf{R}_{\mathbf{s}}$ as

$$\hat{\mathbf{R}}_{\mathbf{s}} = \hat{\mathbf{R}}_{\mathbf{x}} - \hat{\mathbf{R}}_{\mathbf{w}}. \tag{12}$$

To ensure the non-negative definiteness of $\hat{\mathbf{R}}_{\mathbf{s}}$, we implement the above relationship in the power spectral domain as

$$\hat{\mathbf{P}}_{\mathbf{ss}}(k) = \max\{\hat{\mathbf{P}}_{\mathbf{xx}}(k) - \hat{\mathbf{P}}_{\mathbf{ww}}(k),\, 0\}, \tag{13}$$

where k is the discrete frequency index and $\hat{\mathbf{P}}_{\mathbf{xx}}(k)$, $\hat{\mathbf{P}}_{\mathbf{ss}}(k)$, $\hat{\mathbf{P}}_{\mathbf{ww}}(k)$ are the power spectral density estimates of the noisy speech, clean speech and noise. For a particular vector, $\{\hat{\mathbf{P}}_{\mathbf{xx}}(k)\}$ is given by the power spectrum of the noisy speech frame. To obtain $\hat{\mathbf{P}}_{\mathbf{ww}}(k)$, any noise estimation algorithm such as [15] can be used. Now, the estimates of clean speech autocorrelations $\{\hat{\mathbf{r}}_{\mathbf{ss}}(l)\}$ are obtained by inverse Fourier transform of $\{\hat{\mathbf{P}}_{\mathbf{ss}}(k)\}$. Finally, using $\{\hat{\mathbf{r}}_{\mathbf{ss}}(l),\, l = 0, 1, \ldots, N-1\}$ an $N \times N$ Toeplitz matrix $\hat{\mathbf{R}}_{\mathbf{s}}$ is formed and the EVD is carried out as

$$\hat{\mathbf{R}}_{\mathbf{s}} = \hat{\mathbf{U}}\hat{\Lambda}_{\mathbf{s}}\hat{\mathbf{U}}^T, \tag{14}$$

where $\hat{\Lambda}_{\mathbf{s}} = \text{diag}\{\hat{\lambda}_{\mathbf{s},1}, \hat{\lambda}_{\mathbf{s},2}, \ldots, \hat{\lambda}_{\mathbf{s},N}\}$ such that $\hat{\lambda}_{\mathbf{s},1} \geq \hat{\lambda}_{\mathbf{s},2} \geq \ldots \geq \hat{\lambda}_{\mathbf{s},N}$. Now, using the standard approach [9], the signal subspace dimension K is determined as

$$K = \arg\max_{i}\,\{\hat{\lambda}_{\mathbf{s},i} > 0\}. \tag{15}$$

After determining K, $\mathbf{U}_1$ can be obtained from $\hat{\mathbf{U}}$ by taking only the first K columns and the clean speech is estimated using the MMSE estimator shown in Eqn. (9).

3. Experiments and results

3.1. Experimental setup

The speech data used in the experiments are taken from the TIMIT database. The original speech signal (sampled at 16 kHz) is first low pass filtered (3.4 kHz cut-off frequency) and then down-sampled to 8 kHz. We have used about 40 minutes of speech data for training purpose and a separate 3 minutes of speech data (6 male and 6 female speakers speaking 5 different sentences each) for testing. The training data are used for estimating the GMM parameters employing expectation-maximization (EM) algorithm. The test speech is generated by adding noise to the clean speech signal at the required signal-to-noise ratio (SNR). We have considered different types of noise taken from NOISEX-92 database: white noise, pink noise, babble noise, F16 cockpit noise, m109 tank noise and HF channel noise. In all our experiments, we have assumed that the noise is stationary and covariance matrix of the noise is estimated only once from the initial 120 msec segment (which contains only noise) of the test speech. For the proposed method, we choose frame size (N) as 160 samples (20 msec at 8 kHz). We fix the maximum dimension of the signal subspace (K) empirically as 80. We use an 80 dimensional GMM with 240 mixture component for modeling the joint PDF of first 80 KLT coefficients of clean speech. Using the property that the marginal densities of multivariate GM density are also GM densities, we obtain the GM density for a K dimensional vector $\mathbf{\Phi}$, where $K \leq 80$. The noisy speech signal is processed as overlapping frames with 50% overlap between successive frames. The enhanced frames

$$\hat{\mathbf{S}}_{\text{MMSE}} = \sum_{m=1}^{M} \boldsymbol{\beta}_m(\mathbf{X}) \; \mathbf{U}_1 \left\{ \boldsymbol{\mu}_{m,\Phi} + \boldsymbol{\Sigma}_{m,\Phi} \mathbf{U}_1^T \left[\mathbf{U}_1 \boldsymbol{\Sigma}_{m,\Phi} \mathbf{U}_1^T + \mathbf{C}_{\mathbf{w}} \right]^{-1} \left[\mathbf{X} - \mathbf{U}_1 \boldsymbol{\mu}_{m,\Phi} \right] \right\}, \tag{9}$$

where $\forall\, m \in \{1, 2, \ldots, M\}$, $\boldsymbol{\beta}_m(\mathbf{X})$ is given by

$$\boldsymbol{\beta}_m(\mathbf{X}) = \frac{\frac{\boldsymbol{\alpha}_m}{\sqrt{\det\left(\mathbf{U}_1 \boldsymbol{\Sigma}_{m,\Phi} \mathbf{U}_1^T + \mathbf{C}_{\mathbf{w}}\right)}} \exp\left\{ -\frac{1}{2}\left[\mathbf{X} - \mathbf{U}_1 \boldsymbol{\mu}_{m,\Phi}\right]^T \left[\mathbf{U}_1 \boldsymbol{\Sigma}_{m,\Phi} \mathbf{U}_1^T + \mathbf{C}_{\mathbf{w}}\right]^{-1} \left[\mathbf{X} - \mathbf{U}_1 \boldsymbol{\mu}_{m,\Phi}\right] \right\}}{\sum_{i=1}^{M} \frac{\boldsymbol{\alpha}_i}{\sqrt{\det\left(\mathbf{U}_1 \boldsymbol{\Sigma}_{i,\Phi} \mathbf{U}_1^T + \mathbf{C}_{\mathbf{w}}\right)}} \exp\left\{ -\frac{1}{2}\left[\mathbf{X} - \mathbf{U}_1 \boldsymbol{\mu}_{i,\Phi}\right]^T \left[\mathbf{U}_1 \boldsymbol{\Sigma}_{i,\Phi} \mathbf{U}_1^T + \mathbf{C}_{\mathbf{w}}\right]^{-1} \left[\mathbf{X} - \mathbf{U}_1 \boldsymbol{\mu}_{i,\Phi}\right] \right\}}. \tag{10}$$

are overlap-added using Hamming window to construct the enhanced speech signal. To measure the speech enhancement performance, we use the following objective measures: SNR, average segmental SNR and perceptual evaluation of speech quality (PESQ) score.

3.2. Speech enhancement performance

For performance comparison, we implement the spectral domain constrained (SDC) estimator of Ephraim and Van Trees [2] and the Wiener estimator of Eqn. (21). For white noise corrupted speech signal, SE performance of these estimators together with that of the GMM based method are shown in Table 1. We can see that the performance of the GMM based method is much better than the simple Gaussian based linear estimator (Wiener estimator) in terms of SNR, average segmental SNR and PESQ score. We also note that the developed method perform better than the SDC estimator of Ephraim and Van Trees at all the three SNR condition. In this context, we mention that the performance of the SDC estimator of Ephraim and Van Trees is close to the performance of the super-Gaussian density based SSE method [11] in terms of average segmental SNR.

In [9], Hu and Loizou have generalized the signal subspace method of Ephraim and Van Trees [2] for colored noise that includes the method of Ephraim and Van Trees as a special case. In Table 2, we present the performance of the generalized subspace based method [9] and the developed method for various kinds of real life noises. We observe that in terms of the objective measures, the proposed method shows better enhancement performance than the generalized subspace based method under all the noise conditions. Informal listening test also confirms the superiority of the proposed method over the method of Hu and Loizou. Thus, the developed subspace based approach using GMM can be regarded as an effective method for speech enhancement.

4. Conclusion

For the subspace based signal model, we develop an optimum MMSE estimator. The use of GMM to model the non-Gaussian PDF of speech KLT coefficients is proposed. Unlike the existing method [11], the developed method does not assume that the speech KLT coefficients are independent. The proposed nonlinear estimator includes the Wiener estimator as special case. For speech enhancement, the subspace based method using GMM provides better result than the generalized subspace based method [9] and the improvement in performance is attributed to the use of GMM and optimum MMSE estimation.

5. Appendix

For subspace based approach, the traditional linear estimator (Wiener estimator) is the MMSE estimator under the assumption that the KLT coefficients of clean speech and noise are Gaussian distributed. Here, we show that the LMMSE estimator can be obtained as a special case of the developed GMM based nonlinear estimator of Eqn. (9). In this case, the noise is assumed to be zero mean and white with variance $\sigma_{\mathbf{w}}^2$ as in [2]. Therefore, the Gaussian PDF of $\mathbf{W}$ is given by

$$f_{\mathbf{W}}(\mathbf{w}) = \mathcal{N}\left(\mathbf{0}, \sigma_{\mathbf{w}}^2 \mathbf{I}\right). \tag{16}$$

Also the PDF of $\boldsymbol{\Phi}$ is modeled as

$$f_{\boldsymbol{\Phi}}(\Phi) = \mathcal{N}\left(\mathbf{0}, \boldsymbol{\Sigma}\right), \tag{17}$$

where $\boldsymbol{\Sigma}$ = *diag*$\{\lambda_1, \lambda_2, \ldots, \lambda_K\}$ and λ_i denotes the variance of the ith KLT coefficient of clean speech. For this Gaussian based model, the MMSE estimator of Eqn. (6) can be evaluated as a special case ($M = 1$) of the Eqn. (9). For $M = 1$, putting $\mathbf{C}_{\mathbf{w}} = \sigma_{\mathbf{w}}^2 \mathbf{I}$, $\boldsymbol{\mu}_{1,\Phi} = 0$ and $\boldsymbol{\Sigma}_{1,\Phi} = \boldsymbol{\Sigma}$ in Eqn. (9), we get

$$\hat{\mathbf{S}}_{\text{MMSE}} = \mathbf{U}_1 \boldsymbol{\Sigma} \mathbf{U}_1^T \left[\mathbf{U}_1 \boldsymbol{\Sigma} \mathbf{U}_1^T + \sigma_{\mathbf{w}}^2 \mathbf{I} \right]^{-1} \mathbf{X}. \tag{18}$$

To simplify the above expression, we define an $N \times N$ matrix $\boldsymbol{\Lambda}$ as

$$\boldsymbol{\Lambda} \triangleq \begin{bmatrix} \boldsymbol{\Sigma} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}, \tag{19}$$

and thus, Eqn. (18) can be expressed as

$$\hat{\mathbf{S}}_{\text{MMSE}} = \mathbf{U} \boldsymbol{\Lambda} \mathbf{U}^T \left[\mathbf{U} \boldsymbol{\Lambda} \mathbf{U}^T + \sigma_{\mathbf{w}}^2 \mathbf{I} \right]^{-1} \mathbf{X}. \tag{20}$$

Now, using Eqn. (2) and the fact that $\mathbf{U}\mathbf{U}^T = \mathbf{U}^T\mathbf{U} = \mathbf{I}$, the above expression can be written as

$$\begin{aligned} \hat{\mathbf{S}}_{\text{MMSE}} &= \mathbf{U} \boldsymbol{\Lambda} \left[\boldsymbol{\Lambda} + \sigma_{\mathbf{w}}^2 \mathbf{I} \right]^{-1} \mathbf{U}^T \mathbf{X} \\ &= \mathbf{U}_1 \boldsymbol{\Sigma} \left[\boldsymbol{\Sigma} + \sigma_{\mathbf{w}}^2 \mathbf{I} \right]^{-1} \mathbf{U}_1^T \mathbf{X} \\ &\triangleq \mathbf{U}_1 \mathbf{G} \mathbf{U}_1^T \mathbf{X}, \end{aligned} \tag{21}$$

where $\mathbf{G}$ is a $K \times K$ diagonal matrix, whose ith diagonal element g_i is given by

$$g_i = \frac{\lambda_i}{\lambda_i + \sigma_{\mathbf{w}}^2},\; i = 1, 2, \ldots, K. \tag{22}$$

We see that the linear estimator given by Eqn. (21) is same as the time domain constrained estimator ($\mu = 1$ case) of [2].

Table 1: For white noise corrupted speech signal, speech enhancement performance of the Wiener estimator (Method-1), SDC estimator [2] (Method-2) and the GMM based proposed method (Method-3) at different input SNR conditions

SNR (dB)				Average Segmental SNR (dB)				PESQ (MOS)			
Input	Method-1	Method-2	Method-3	Input	Method-1	Method-2	Method-3	Input	Method-1	Method-2	Method-3
0	5.57	8.43	10.34	-3.87	-0.31	3.00	4.89	1.59	1.84	2.06	2.14
5	9.80	11.77	13.30	-0.36	3.30	5.91	7.35	1.85	2.19	2.47	2.56
10	13.98	15.31	16.21	3.81	7.25	9.00	10.17	2.16	2.86	2.83	2.90

Table 2: Speech enhancement performance of the generalized subspace based method [9] ($Output_1$) and the GMM based proposed method ($Output_2$) for various types of noise at different input SNR conditions

Noise Type	SNR (dB)			Average Segmental SNR (dB)			PESQ (MOS)		
	Input	$Output_1$	$Output_2$	Input	$Output_1$	$Output_2$	Input	$Output_1$	$Output_2$
m109 tank noise	0	9.51	11.73	-3.27	4.74	6.69	2.37	2.60	2.77
	5	13.12	14.82	0.33	7.85	9.42	2.70	2.98	3.14
	10	16.77	17.78	4.57	11.15	12.05	3.03	3.33	3.43
F16 cockpit noise	0	5.88	7.95	-3.62	0.73	2.35	1.87	2.12	2.23
	5	9.73	11.18	-0.05	3.90	5.50	2.18	2.44	2.53
	10	13.68	14.74	4.14	7.42	8.61	2.50	2.77	2.86
pink noise	0	7.87	9.21	-3.70	3.17	4.39	1.79	2.26	2.39
	5	11.37	12.33	-0.16	5.73	6.74	2.12	2.59	2.69
	10	15.15	16.03	4.01	8.78	9.48	2.46	2.92	2.99
babble noise	0	4.81	6.36	-3.35	0.96	1.81	1.95	2.04	2.18
	5	8.10	9.29	0.25	3.33	4.38	2.27	2.35	2.48
	10	11.96	13.04	4.47	6.39	7.19	2.59	2.68	2.76
HF channel noise	0	7.58	9.32	-3.84	2.90	4.48	1.66	1.94	2.10
	5	11.16	12.48	-0.32	5.69	6.93	1.89	2.35	2.52
	10	14.89	16.10	3.85	8.77	9.58	2.18	2.73	2.81

6. References

[1] P.C. Loizou, "Speech enhancement: Theory and practice," CRC Press, 2007.

[2] Y. Ephraim and H.L. Van Trees, "A signal subspace approach for speech enhancement," *IEEE Trans. Speech Audio Process.*, vol. 3, no. 4, pp. 251-266, July 1995.

[3] M. Dendrinos, S. Bakamidis and G. Carayannis, "Speech enhancement from noise: a regenerative approach, *Speech Commun.*, vol. 10, pp. 45-57, Feb. 1991.

[4] S.H. Jensen, P.C. Hansen, S.D Hansen, J.A. Sorensen, "Reduction of broad-band noise in speech by truncated QSVD," *IEEE Trans. Speech Audio Process.*, vol. 3, no. 6, pp. 439-448, Nov. 1995.

[5] G.H. Ju, L.S. Lee, "A perceptually constrained GSVD-based approach for enhancing speech corrupted by colored noise," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 15, no. 1, pp. 119-134, Jan. 2007.

[6] U. Mittal and N. Phamdo, "Signal/noise KLT based approach for enhancing speech degraded by colored noise," *IEEE Trans. Speech Audio Process.*, vol. 8, no. 2, pp. 159-167, Mar. 2000.

[7] A. Rezayee and S. Gazor, "An adaptive KLT approach for speech enhancement," *IEEE Trans. Speech Audio Process.*, vol. 9, no. 2, pp. 87-95, Feb. 2001.

[8] H. Lev-Ari, Y. Ephraim,"Extension of the signal subspace speech enhancement approach to colored noise," *IEEE Signal Process. Lett.*, vol. 10, no. 4, pp. 104-106, Apr. 2003.

[9] Y. Hu and P.C. Loizou, "A generalized subspace approach for enhancing speech corrupted by colored noise," *IEEE Trans. Speech Audio Process.*, vol. 11, no. 4, pp. 334-341, July 2003.

[10] S. Gazor and W. Zhang, "Speech probability distribution," *IEEE Signal Process. Lett.*, vol. 10, no. 7, pp. 204-207, July 2003.

[11] J. Jensen and R. Heusdens, "Improved subspace-based single-channel speech enhancement using generalized super-Gaussian priors," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 15, no. 3, pp. 862-872, Mar. 2007.

[12] M. Effros, F. Hanying and K. Zeger, "Suboptimality of the Karhunen-Loeve transform for transform coding," *IEEE Trans. Information Theory*, Vol. 50, no. 8, pp. 1605-1619, Aug. 2004.

[13] H.W. Sorenson and D. L. Aspach, "Recursive Bayesian estimation using Gaussian sums," *Automatica*, vol. 7, pp. 465-479, 1971.

[14] A. Kundu, S. Chatterjee, A.S. Murthy and T.V. Sreenivas, "GMM based Bayesian approach to speech enhancement in signal / transform domain," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, pp. 4893-4896, Mar. 2008.

[15] I. Cohen, B. Berdugo, "Noise estimation by minima controlled recursive averaging for robust speech enhancement," *IEEE Signal Process. Lett.*, vol. 9, no. 1, pp. 12-15, Jan. 2002.