Journal of Computer and Communications, 2016, 4, 46-53
Published Online March 2016 in SciRes. http://www.scirp.org/journal/jcc
http://dx.doi.org/10.4236/jcc.2016.43007

Joint Noise Reduction and l_p-Norm Minimization for Enhancing Time Delay Estimation in Colored Noise

Jingxian Tu, Youshen Xia

College of Mathematics and Computer Science, Fuzhou University, Fuzhou, China
Email: 568205127@qq.com

Received 24 December 2015; accepted 26 February 2016; published 2 March 2016

Abstract

Time delay estimation (TDE) is an important issue in signal processing. Conventional TDE algorithms are usually efficient under white noise environments. In this paper, a joint noise reduction and l_p-norm minimization method is presented to enhance TDE in colored noise. An improved subspace method for colored noise reduction is first performed. Then the time delay is estimated by using an l_p-norm minimization method. Because the clean speech signal form the noisy signal is well extracted by noise reduction and the l_p-norm minimization method is robust, the TDE accuracy can be enhanced. Experiment results confirm that the proposed joint estimation method obtains more accurate TDE than several conventional algorithms in colored noise, especially in the case of low signal-to-noise ratio.

Keywords

Time Delay Estimation, Speech Enhancement, Noise Reduction, Subspace

1. Introduction

Estimating the time delay from two received signals at spatially separated sensors is of important significance in signal processing [1]. It has many practical applications, such as multichannel speech enhancement, echo cancellation and wireless communications. The basic problem of the time delay estimation (TDE) is estimating accurately the time delay of interfering signals, aiming to exclude the influence of noise and interference. Many TDE approaches have been proposed. They mainly include the generalized correlation method [2], the statistical method [3], the parametric estimation method [4], the adaptive estimation method [5] [6], the combinational estimation method [7], and the l_p-norm minimization-based estimation method [8]. Among them, the l_p-norm minimization-based estimation method can find the time delay by minimizing an l_p-norm objective function. It was reported in [8] that this method can obtain more robust results than other several conventional approaches against impulse noise. However, these conventional TDE approaches do not consider the influence of noise, specially under low SNR conditions. Moreover, in general they are only efficient in white noise.

Speech enhancement techniques have been applied in speech recognition and voice communication. They can

How to cite this paper: Tu, J.X. and Xia, Y.S. (2016) Joint Noise Reduction and l_p-Norm Minimization for Enhancing Time Delay Estimation in Colored Noise. *Journal of Computer and Communications*, **4**, 46-53.
http://dx.doi.org/10.4236/jcc.2016.43007

recover the clean speech signal from the noisy signal by noise reduction. Speech enhancement algorithms can be classified as single channel and multichannel speech enhancement algorithms. The multichannel speech enhancement algorithms usually mix multiple noisy signals for noise reduction. By contrast, the single speech enhancement algorithms utilize only one noisy signal and thus do not change the time delay of the noisy signal at each channel. So, the single speech enhancement algorithms have the potential to improve the performance of the TDE algorithm. At present, the single-channel speech enhancement algorithms mainly include the spectral subtraction-based methods [9], the Kalman filtering-based parametric method [10], the statistic-based approach [11] [12], and the subspace-based method [13]. These conventional algorithms are suitable for white noise reduction. To deal with colored noise, one conventional approach is that the noisy speech signal is multiplied by the square root of the noise covariance matrix's inverse [14]. Another conventional approach is the prewhitening covariance matrix of the colored noise. These prewhitening approaches all require to estimate the covariance matrix of the colored noise in advance. Recently, to avoid disadvantage of estimating the covariance matrix of the colored noise, an improved subspace method was proposed in [15]. It was reported that the improved subspace method outperforms conventional speech enhancement methods in colored noise reduction.

In this paper, a new method for enhancing time delay estimation (TDE) in colored noise is presented by joining noise reduction and l_p-norm minimization. We first perform the improved subspace method for enhanced signals corrupted by colored noise and then we use the l_p-norm minimization based TDE method to estimate the time delay from the enhanced signals. Experiment results show that the proposed joint algorithm can obtain more accurate TDE than several conventional algorithms in colored noise, especially in the case of low signal-to-noise ratio.

2. TDE Signal Model and Estimation

2.1. Signal Model

Consider the following TDE signal model:

$$x_1(n) = s(n) + v_1(n), x_2(n) = \beta s(n - D^*) + v_2(n) \\ (n = 1, \cdots, N) \tag{1}$$

where $s(n)$ is the unknown random source signal, $\beta \in (0,1)$ is the attenuation factor, $D^* \in \mathbb{R}$ is the time delay to be estimated, $v_1(n)$ and $v_2(n)$ are uncorrelated noise observations which are independent of $s(n)$, and N is the sampling length of the noisy signals. The goal of TDE is to estimate the delay D^* from N noise observations.

2.2. l_p-Norm Minimization-Based Estimation

For robust TDE, Ma and Nikias introduced [16] the following l_p-norm cost function about the delay D and the attenuation factor β:

$$f_p(D, \beta) = \sum_{n=M+1}^{N-M} | x_2(n) - \beta \sum_{i=-M}^{M} x_1(n-i) sinc(i-D) |^p \tag{2}$$

where $1 \le p \le 2$, M is the order parameter and $sinc(t) = \frac{sin(\pi t)}{\pi t}$ is the sinc function. (2) can be written in a matrix form as:

$$f_p(D, \beta) = \left\| \beta \mathbf{b}(D) - \mathbf{x}_2 \right\|_p^p \tag{3}$$

where $\mathbf{x}_2 \triangleq [x_2(M+1), x_2(M+2), \cdots, x_2(N-M)]^T$, $\mathbf{b}(D) \triangleq \mathbf{X}_1 \mathbf{c}(D)$, $\mathbf{c}(D) \triangleq [sinc(-M-D), \cdots, sinc(M-D)]$, and

$$\mathbf{X}_1 \triangleq \begin{bmatrix} x_1(2M+1) & x_1(2M) & \cdots & x_1(1) \\ x_1(2M+2) & x_1(2M+1) & \cdots & x_1(2) \\ \vdots & \vdots & \ddots & \vdots \\ x_1(N) & x_1(N-1) & \cdots & x_1(N-2M) \end{bmatrix}.$$

To minimize $f_p(D,\beta)$, Zeng *et al*. presented an efficient two-steps procedure [8]. In the first step, the global optimum β is estimated for each given *D*. The estimation for the global optimum β has the following three cases.

Case 1: $p=2$. The l_p cost function is a one-dimensional quadratic function of β. Its optimal solution is given by

$$\beta^* = \frac{\mathbf{b}^T(D)\mathbf{x}_2}{\|\mathbf{b}(D)\|^2}. \tag{4}$$

Case 2: $p=1$. The l_p cost function is the least absolute deviation function:

$$f_1(D,\beta) = \|\beta\mathbf{b}(D)-\mathbf{x}_2\|_1 = \sum_{n=M+1}^{N-M} |b_n(D)| \left|\beta - \frac{x_2(n)}{b_n(D)}\right| \tag{5}$$

where $b_n(D)$ is the $(n-M)$th element of $\mathbf{b}(D)$. Let $y(n) = \dfrac{x_2(n)}{b_n(D)}, n = M+1,\ldots,N-M$, then the optimal solution of the cost function is the weighted median of the sequence $\{y(n)\}_{n=M+1}^{N-M}$ with the weights $\{|b(n)|\}_{n=M+1}^{N-M}$. The procedure of computation of β^* is listed in **Algorithm 1**.

Algorithm 1 Computation of β^*

Step 1) Compute the threshold $b_0 = \frac{1}{2}\sum_{n=M+1}^{N-M}|b_n(D)|$.

Step 2) Sort the sequence $\{y(n)\}_{n=M+1}^{N-M}$ in ascending order along with the corresponding weights $\{|b(n)|\}_{n=M+1}^{N-M}$.

Step 3) Let m increase from $M+1$ to $N-M$, β^* is $y(m)$ whose weight hold the inequality $\sum_{n=M+1}^{m}|b_n(D)| \geq b_0$ first.

Case 3: $1<p<2$. The l_p cost function has derivative. Thus the following fixed-point iteration is used to find the optimal solution β^*:

$$\hat{\beta}^{k+1} = \frac{\mathbf{b}^T(D)\mathbf{W}(\hat{\beta}^k)\mathbf{x}_2}{\mathbf{b}^T(D)\mathbf{W}(\hat{\beta}^k)\mathbf{b}(D)} \tag{6}$$

where $\hat{\beta}^k$ is the estimate of β in the *k*th iteration, and

$$\mathbf{W}(\beta) = diag\{|r_{M+1}|^{(p-2)/2},\cdots,|r_{N-M}|^{(p-2)/2}\}$$

where r_n is the $(n-M)$th element of $\beta\mathbf{b}(D)-\mathbf{x}_2$. In the second step, a search range $[D_{min}, D_{max}]$ and a step size ΔD are first determined, then the value of D increases from D_{min} to D_{max} with the step size being ΔD, after that the delay profile is computed by substituting each given D and the corresponding $\beta^*(D)$ into the l_p cost function, finally the minimum $\hat{D}^* = \arg\min_D f_p(D,\beta^*(D))$ is used to estimate the D^*.

3. Proposed TDE Algorithm

3.1. Improved Subspace Method for Colored Noise Reduction

Without the loss of generality, we consider the following noise signal model:

$$x_1(n) = s(n) + v_1(n) \quad (n = 1,\cdots,N) \tag{7}$$

Our goal is to restore $s(n)$ from $x_1(n)$ by colored noise reduction.

Recently, for colored noise reduction, an improved subspace method was presented in [15]. Let the colored noise be modeled as the *p*th order autoregressive signal process

$$v_1(n) = \sum_{i=1}^{p} a_i v_1(n-i) + w(n) \tag{8}$$

where $\{a_i\}$ are the AR noise model parameters, $w(n)$ is the drive noise which is assumed to be white with variance σ_w^2. Let K denote the length of one frame signal, and let $x_1 = [x_1(n-K+1), \ldots, y(n-1), y(n)]^T$, $s = [s(n-K+1), \ldots, s(n-1), s(n)]^T$, and $v_1 = [v_1(n-K+1), \ldots, v_1(n-1), v_1(n)]^T$. Then (7) can be written in a vector form:

$$\mathbf{x}_1 = \mathbf{s} + \mathbf{v}_1 \tag{9}$$

and (8) can be written in a vector form:

$$\mathbf{A}\mathbf{v}_1 = \mathbf{w}$$

where $\mathbf{w} = [w(n-K+1), \ldots, w(n-1), w(n)]^{\mathrm{T}}$ and A is the $K \times K$ whitening matrix:

$$A = \begin{bmatrix} 1 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 & 0 \\ -a_1 & 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ -a_p & -a_{p-1} & \cdots & 1 & \cdots & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & -a_p & \cdots & -a_1 & 1 \end{bmatrix}$$

Multiplying (9) by A, we have

$$\mathbf{z} = \mathbf{s}_0 + \mathbf{w}. \tag{10}$$

where $\mathbf{z} = A\mathbf{x}_1$ and $\mathbf{s}_0 = A\mathbf{s}$. Since $w(n)$ is the white noise, the conventional subspace method for white noise reduction can be directly used to estimate $\mathbf{s}_0$, given as

$$\widehat{\mathbf{s}_0} = \mathbf{H}_0\mathbf{z} \tag{11}$$

where $\mathbf{H}_0 = \mathbf{U}\Lambda(\Lambda + \mu\sigma_w^2\mathbf{I})^{-1}\mathbf{U}^{\mathrm{T}}$, μ is the Lagrangian multiplier, $\mathbf{U}^{\mathrm{T}}\mathbf{R}_z\mathbf{U} = \mathbf{\Lambda}$, $\mathbf{R}_z$ is the $K \times K$ covariance matrix of the whitening signal vector $\mathbf{z}$, and $\mathbf{U}$ and $\mathbf{\Lambda}$ consist of the eigenvector and eigenvalue of $\mathbf{R}_z$, respectively. Then the clean speech signal $\mathbf{s}$ can be estimated by

$$\hat{\mathbf{s}} = \mathbf{A}^{-1}\widehat{\mathbf{s}_0} = \mathbf{A}^{-1}\mathbf{U}\mathbf{\Lambda}(\mathbf{\Lambda} + \mu\sigma_w^2\mathbf{I})^{-1}\mathbf{U}^{\mathrm{T}}\mathbf{A}\mathbf{x}_1. \tag{12}$$

For each signal frame, the improved subspace algorithm (denoted as ISS) is summarized as:

Algorithm 2 Improved Subspace Algorithm

Step 1) Update the AR coefficients a_i and variance σ_w^2 of noisy signal during the speech-absent frame.
Step 2) Construct the whitening matrix $\mathbf{A}$ using the estimated AR coefficients, and whiten the noisy signal by using $\mathbf{A}$.
Step 3) Compute the covariance matrix of the whitening noisy signal $\mathbf{R}_z$.
Step 4) Compute the eigenvalue and eigenvector of $\mathbf{R}_z$.
Step 5) Estimate the clean speech signal by using (12).

3.2. Proposed TDE Algorithm

In this section, we introduce a new method for enhancing time delay estimation (TDE) in colored noise, based on joint noise reduction and l_p-norm minimization. An improved subspace method for colored noise reduction is first performed. The time delay is then estimated by using the enhanced signal, based on the l_p-norm minimization. The proposed TDE algorithm is listed in **Algorithm 3**. Compared with conventional TDE algorithms, the proposed TDE algorithm can greatly reduce the interference of colored noise such that the TDE accuracy is enhanced.

Algorithm 3 The Proposed TDE Algorithm

Input: $x_1(n), x_2(n), p, M, D_{min}, D_{max}, \Delta D$.
Output: Time delay estimate $\hat{D}^*$.
Estimate $s(n)$ and $s(n-D^*)$ from $x_1(n)$ and $x_2(n)$ by performing **Algorithm 2**.
$x_1(n) \leftarrow \hat{s}(n)$ and $x_2(n) \leftarrow \hat{s}(n-D), n = 1, \cdots, N$.
Select a segment of the enhanced signal of each microphone where the energy of the signal is biggest and the length is 500 for time delay estimation.
for $D = D_{min} : \Delta D : D_{max}$ **do**
  Compute $\mathbf{b}(D) \triangleq \mathbf{X}_1\mathbf{c}(D)$.
  if $p = 1$ **then**
    Compute the optimal $\beta^*(D)$ by performing **Algorithm 1**.
  else if $p = 2$ **then**
    Compute the optimal $\beta^*(D)$ according to (4).
  else
    Compute the optimal $\beta^*(D)$ using the fixed-point iteration of (6).
  end if
  Compute the cost function $f_p(D, \beta^*(D))$ using (3).
end for
Find the minimum of the delay profile by $\hat{D}^* = \arg\min_D f_p(D, \beta^*(D))$.

4. Experimental Results

In this section, we conduct numerical simulations to demonstrate the effectiveness of the proposed algorithm. We compare the proposed TDE algorithm with the TDE algorithm based l_p-norm minimization without noise reduction. We also compare with other two TDE algorithms based l_p-norm minimization with noise reduction, where the minimum mean square error(MMSE) and maximum a posterior(MAP) estimators of the magnitude-squared spectrum(denoted as MMSE-MSS and MAP-MSS) are used for noise reduction, respectively. The source signal and noisy signal are taken from the NOIZEUS [17] and NOISEX [18] corpora, respectively. We randomly select twenty different speech sentences from the NOIZEUS corpora. Babble and factory noises are selected from the NOISEX corpora. Each speech sentence is corrupted by these two noises with different input SNRs. The true delay is set to $D^* = 3$, the attenuation factor is $\beta = 0.9$, the approximation order parameter is $M = 10$, the delay search range is $[D_{min}, D_{max}] = [-10, 10]$, and the search step size is $\Delta D = 10^{-3}$. To evaluate the performance of our proposed methods, we use the root mean square error(RMSE), which is defined as:

$$RMSE = \sqrt{\frac{1}{M_c}\sum_{m=1}^{M_c}(\hat{D}_m^* - D^*)^2} \tag{13}$$

where $M_c = 20$ is the number of speech sentences and $\hat{D}_m^*$ is the delay estimate of the *m*th speech sentence. By the l_p-norm minimization method we see that $p = 1.4$ and $p = 1.6$ are the best choice in Babble noise and factory noise, respectively.

In the first test, we perform the four algorithms for different values of the input SNRs. **Figure 1** and **Figure 2** display the RMSE results of the four algorithms with different values of the input SNRs (From 0 dB to 10 dB) in factory noise and babble noise, respectively. From the two figures, we first see that the four algorithms obtain higher value of RMSE when increasing the input SNRs. This indicates that the noise decrease the performance of TDE. Second, we see that the proposed algorithm can outperform the TDE algorithm based l_p-norm minimization without noise reduction for all input SNRs in terms of RMSE. Third, the proposed TDE algorithm can get a lower value of RMSE than the other two TDE algorithms with noise reduction, based on the MMSE-MSS and MAP-MSS estimators, respectively. This also indicates that the proposed algorithm can obtain the best accurate TDE in terms of RMSE.

In the second test, we perform the four algorithms with the input SNRs being 5 dB via different values of D^*. **Figure 3** and **Figure 4** display their RMSE results of the four algorithms in factory noise and babble noise,

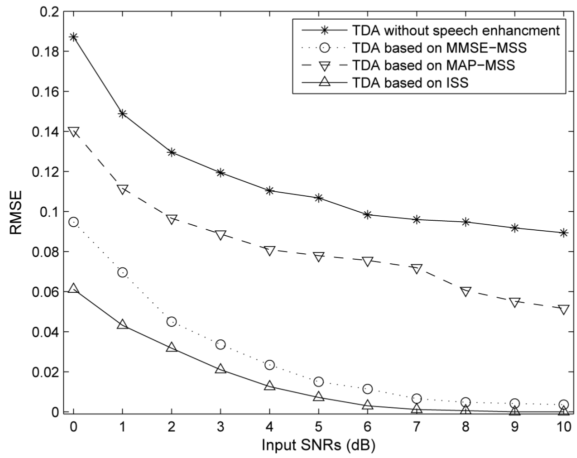

Figure 1. RMSE of TDE based on four algorithms with different input SNRs and $p = 1.6$ in factory noise.

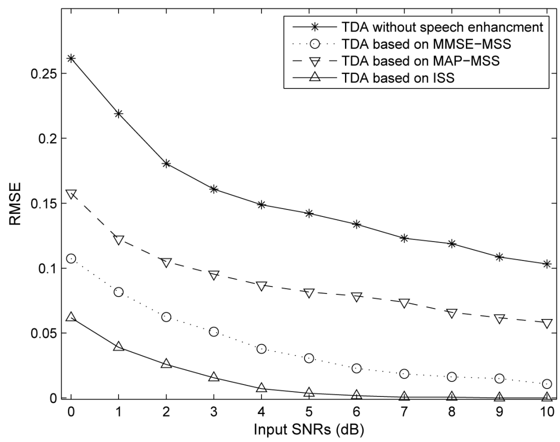

Figure 2. RMSE of TDE based on four algorithms with different input SNRs and $p = 1.4$ in babble noise.

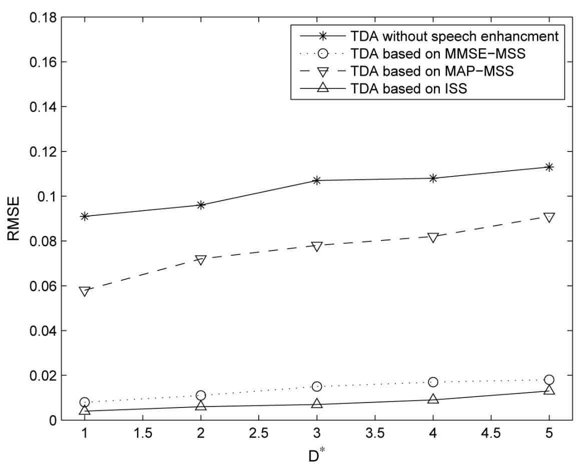

Figure 3. RMSE of TDE based on four algorithms via different values of D^* with the input SNRs being 5 dB and $p = 1.6$ in factory noise.

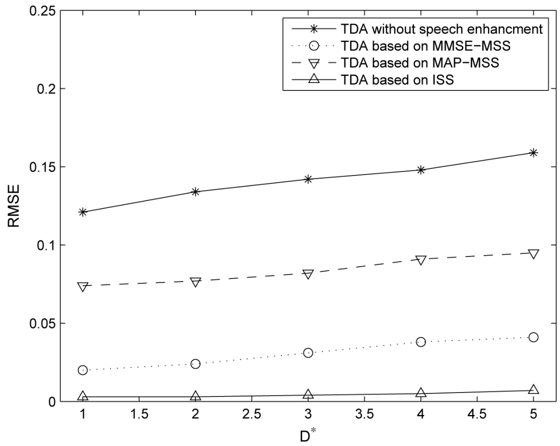

Figure 4. RMSE of TDE based on four algorithms via different values of D^* with the input SNRs being 5 dB and $p = 1.4$ in babble noise.

respectively. From the two figures, we first see that the proposed algorithms outperform TDE without speech enhancement in any value of D^*. Second, the proposed TDE algorithm can get a lower value of RMSE than the other two TDE algorithms with noise reduction, based on the MMSE-MSS and MAP-MSS estimators, respectively. This indicates that the proposed algorithm can obtain the best accurate TDE in any value of D^*.

References

[1] Xing, H.Y. and Tang, J. (2008) Analysis and Survey of Algorithms for Time-Delay Estimation. *Technical Acoustics*, **27**, 110-114.

[2] Knapp, C.H. and Carter, G.C (1976) The Generalized Correlation Method for Estimation of Time Delay. *IEEE Trans. Acoust. Speech Signal Process*, **24**, 320-327. http://dx.doi.org/10.1109/TASSP.1976.1162830

[3] Torrieri, D.J. (1984) Statistical Theory of Passive Location Systems. *IEEE Transactions on Aerospace and Electronic Systems*, **20**, 183-197. http://dx.doi.org/10.1109/TAES.1984.310439

[4] Kenneth, W.K., Lui, B., Frankie, K.W. and So Chan, H.C. (2009) Accurate Time Delay Estimation Based Passive Localization. *Signal Processing*, **89**, 1835-1838. http://dx.doi.org/10.1016/j.sigpro.2009.03.009

[5] Palanisamy, P. and Kalyanasundaram, N. (2010) Sonar Target Detection by Modified Adaptive Noise Cancellation Using Correlating Filter. *International Journal of Electronics*, **98**, 41-60. http://dx.doi.org/10.1080/00207217.2010.497670

[6] Widrow, B. and Stearn, S.D. (1993) Adaptive Signal Processing. Prentice-Hall, Englewood Cliffs, NJ.

[7] Lotf, S., Mourad, T. and Sabeur, A.A.C. (2013) Performance of Wavelet Analysis and Neural Networks for Pathological Voices Identification. *International Journal of Electronics*, **98**, 1129-1140.

[8] Zeng, W.J., So, H.C. and Abdelhak, M.Z. (2013) An l_p*-norm* Minimization Approach to Time Delay Estimation in Impulse Noise. *Digital Signal Processing*, **23**, 1247-1254. http://dx.doi.org/10.1016/j.dsp.2013.03.013

[9] Boll, S. (1979) Suppression of Acoustic Noise in Speech Using Spectral Subtraction. *IEEE Trans. Acoust., Speech, Signal Process*, **27**, 113-120. http://dx.doi.org/10.1109/TASSP.1979.1163209

[10] So, S. and Paliwal, K.K. (2011) Suppressing the Influence of Additive Noise on the Kalman Gain for Low Residual Noise Speech Enhancement. *Speech Communication*, **53**, 355-378. http://dx.doi.org/10.1016/j.specom.2010.10.006

[11] Lu, Y. and Loizou, P.C. (2011) Estimators of the Magnitude-Squared Spectrum and Mehthods for Incorporating SNR Uncertainty. *IEEE Transactions on Audio, Speech and Language Processing*, **19**, 1123-1137. http://dx.doi.org/10.1109/TASL.2010.2082531

[12] Martin, R. (2005) Speech Enhancement Based on Minimum Mean-Square Error Estimation and Supergaussian Priors. *IEEE Transactions on Speech Audio Processing*, **13**, 845-856. http://dx.doi.org/10.1109/TSA.2005.851927

[13] Ephraim, Y. and Van Trees, H.L. (1995) A Signal Subspace Approach for Speech Enhancement. *IEEE Transactions on Speech Audio Processing*, **3**, 251-266. http://dx.doi.org/10.1109/89.397090

[14] Lev-ari, H. and Ephraim, Y. (2003) Extension of The Signal Subspace Speech Enhancement Approach to Colored Noise. *IEEE Signal Processing Letters*, **10**, 104-106. http://dx.doi.org/10.1109/LSP.2003.808544

[15] Wei, Q. and Xia, Y.S. (2013) A Novel Prewhitening Subspace Method for Enhancing Speech Corrupted by Colored Noise. *Proc. IEEE CISP*, **3**, 1282-1286. http://dx.doi.org/10.1109/cisp.2013.6743870

[16] Ma, X. and Nikias, C.L. (1996) Joint Estimation of Time Delay and Frequent Delay in Impulsive Noise Using Fractional Lower Order Statistics. I*EEE Trans. Signal Process.*, **44**, 2669-2687. http://dx.doi.org/10.1109/78.542175

[17] Loizou, P.C. (2007) Speech Enhancement: Theory and Practice. CRC, Boca Raton, FL, USA.

[18] Varga, A. and Steeneken, H.J.M. (1993) Assessment for Automatic Speech Recognition: II, NOISEX-92: A Database and An Experiment to Study The Effect of Additive Noise on Speech Recognition Systems. *Speech Communication*, **12**, 247-251. http://dx.doi.org/10.1016/0167-6393(93)90095-3