

A Noise Reduction Method Based on LMS Adaptive Filter of Audio Signals

Yang Liu¹, Mingli Xiao and Yong Tie

Abstract. Noise reduction of audio signals is a key challenge problem in speech enhancement, speech recognition and speech communication applications, etc. It has attracted a considerable amount of research attention over past several decades. The most widely used method is optimal linear filtering method, which achieves clean audio estimate by passing the noise observation through an optimal linear filter or transformation. The representative algorithms include Wiener filtering, Kalman filtering, spectral restoration, subspace method, etc. Many theoretical analysis and experiments have been carried out to show that the optimal filtering technique can reduce the level of noise that is present in the audio signals and improve the corresponding signal-to-noise ratio (SNR). However, one of the main problems for optimal filtering method is complexity of the algorithm which based upon SVD-decompositions or QR-decompositions. In almost real signal applications it difficult to implement. In this paper, a method for reducing noise from audio or speech signals using LMS adaptive filtering algorithm is proposed. The signal is filtered in the time domain, while the filter coefficients are calculated adaptively by steepest-descent algorithm. The simulation results exhibit a higher quality of the processed signal than unprocessed signal in the noise situation.

Keywords: Audio signals · Noise reduction · Adaptive signal processing · LMS filter

¹ Y. Liu (✉)

College of Electronic Information Engineering, Inner Mongolia University,
010021, Hohhot, China
e-mail: yangliuimu@163.com

Y. Liu

Faculty of Electronic Information and Electrical engineering, Dalian University of
Technology, Dalian, China

M. Xiao

College of Electronic Information Engineering, Inner Mongolia University, Hohhot, China

Y. Tie

College of Electronic Information Engineering, Inner Mongolia University, Hohhot, China

1 Introduction

Since we live in a natural environment where noise is inevitable and ubiquitous, speech signals can seldom be recorded in pure form and are generally contaminated by acoustic background noise. As a result, the speech signals have to be cleaned up with digital signal processing tools before they are stored, transmitted, or played out [1]. Noise reduction algorithms and systems for speech enhancement have received considerable interest in the past, primarily because the reduced speech intelligibility under noisy conditions is one of the major complaints in hearing impaired subjects. Recent years, noise reduction has been in great demand for an increasing number of audio applications, such as automatic speech recognition systems and cellular telephone.

A variety of noise reduction algorithms have been proposed in the literature [2-4]. The noise reduction process, which is often referred to as rather noise reduction or speech enhancement, can be achieved in many different ways, such as beamforming, adaptive filtering, temporal filtering, spatial-temporal filtering, etc. Generally speaking, all of these algorithms can be classified into two categories: single-channel technique and multi-channel technique according to the number of sensors they needed. Compared to the single-channel technique, the multi-channel technique is substantially superior in reducing noise and enhancing speech, due to its spatial filtering capability of suppressing the interfering signals arriving from arbitrary direction other than the specified direction [5]. Among multi-channel noise reduction systems, post-filtering is normally needed to improve the entire performance in practical applications [6]. A multi-channel post-filter is first presented by Zelinski with assumption of zero cross-correlation between noise and signals on different microphones. The linearly constrained adaptive beamformer, first presented by Frost, keeps the signals arriving from the desired look-direction distortionless while suppressing the signals from other directions by minimizing the output power of the beamformer. A small-scale subtractive beamformer based noise reduction algorithm has been proposed in [7]. Its superiority lies in the fact that no adaptive signal processing is adopted and high performance in reducing sudden noise. And its weakness lies in the assumption that only localized noise exist in the environment.

One of the most widely used techniques is the single-channel optimal linear filtering approach. A variety of such algorithms have been developed, including Wiener filter, spectral restoration, subspace method, and parametric method. Subspace approaches have been used in speech enhancement, which project a noisy signal onto two orthogonal subspaces using Karhunen-Loève transform, one is composed of both speech and noise and the other consists of noise only. Noise reduction is achieved by retraining only the signal subspace projection, usually modified by filtering or speech prior. In the frequency domain, a speech signal can be factorized into spectral amplitude and phase components. From perceptual point of view, the spectral amplitude is more important than the phase. Therefore,

the spectral-restoration technique recovers only the spectral amplitude of the clean speech from that of the corruption speech [8]. Additionally, audio signals can be modelled as an AR process. Thus, noise reduction can be formulated as a parameter estimation problem with its objective to estimate the AR model parameter of the clean audio signal from the noisy observations. All these techniques require prior knowledge about the noise, typically, the noise variance or power spectral density, or the instantaneous signal-to-noise ratio (SNR), at all times [9]. When the required noise-related statistics are not available, they are predicted by using neighboring observations without significant speech content based on voice activity detection. Most noise estimation algorithms work well for stationary or slowly varying noise, but less so for heavily non-stationary noise. Without specific knowledge about the noise, it can be difficult to the speech from the noise for these methods.

In this paper, we propose a noise reduction method based on the least-mean-square (LMS) adaptive filter of audio signals. It restores the desired audio signal by passing the noisy speech through a FIR filter whose coefficients are estimated by minimizing the mean square error (MSE) between the clean signals [10]. In many applications, LMS adaptive filtering algorithms are widely used, partly because they require less calculation and are simple to implement. It can also be delineated in the frequency domain, resulting in various derivative techniques [11]. For obtaining faster convergence, this paper will derive a normalized least square (NLMS) algorithm and the associated extended algorithm under Gaussian noise assumption. Simulation results indicate a higher quality of the processed speech signal than original observed signal.

2 Signal Model and Problem Formulation

2.1 Signal Model

The noise reduction problem considered in this paper is to recover a speech signal of interest (SOI) $x(n)$ from the observation signal $y(n)$ which is corrupted by the noise $v(n)$,

$$y(n) = x(n) + v(n) \quad (1)$$

where $v(n)$ is the additive noise, which is a Gaussian random process. It is assumed that the noise $v(n)$ is uncorrelated with the SOI signal $x(n)$. By applying the discrete Fourier transform (DFT) to equation (1), we can get the relationship of the signal model in discrete frequency domain.

2.2 The Adaptive LMS Algorithm

The least-mean-square (LMS) algorithm is widely used in adaptive signal processing for its robustness and simplicity. It is known for its simplicity and its good steady-state performance in stationary context [12]. Consider in general an N -tap filter, with the weight vector $\mathbf{w}(n)$ at time instant n denoted by,

$$\mathbf{w}(n) = [w_1(n) \ w_2(n) \ \cdots \ w_N(n)]^T \quad (2)$$

Let $\{x(n)\}$ be the input sequence and $\mathbf{x}(n) = [x(n) \ x(n-1) \ \cdots \ x(n-N+1)]^T$ be its vector representation containing the immediate past N samples of $\{x(n)\}$. The filter output $\mathbf{y}(n) = \mathbf{w}^T(n)\mathbf{x}(n)$ aims to follow a desired signal $d(n)$, and the estimation error $e(n)$ is defined by

$$e(n) = d(n) - y(n) \quad (3)$$

An adaptive filtering algorithm adjusts the filter tap weight $\mathbf{w}(n)$ at each time instant according to the measured value of $e(n)$. The standard LMS algorithm updates as [10]

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu e(n)\mathbf{x}(n) \quad (4)$$

where μ is defined as the step-size parameter which affects the convergence behavior of the filter weights.

3 Noise Reduction Algorithm Based on LMS Filter

We propose a noise cancelling scheme based on LMS filtering algorithm of its optimum performance. The block diagram of the noise reduction method is shown in Fig. 1. Most audio signals are time varying signals, in order to achieve effective noise reduction with LMS filtering method, the input signal must be segmented. The unprocessed noisy signal is segmented every 40ms. Let $\mathbf{y} = \{y_t : t = 1, 2, \dots, T\}$ be a noisy test signal with T frames and being the frame at time t .

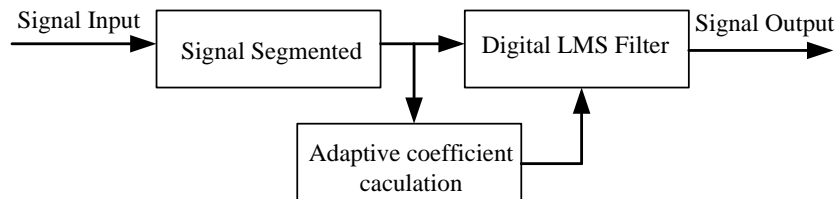


Fig. 1 Block diagram of the proposed noise reduction method

In the proposed method, the problem of noise cancellation can be stated as identifying for each noisy frame y_t a matching weight vector \mathbf{w}_t . Since a whole unit with several frames, when treated as a segment of consecutive speech frame, can be identified more accurately from noise than the whole frames. In the LMS filter we employ the normalized least mean square (NLMS) algorithm. The NLMS is a version of the well-known LMS algorithm that normalizes the weight vector updates with respect to the squared norm of the regressor. This normalization makes the NLMS algorithm less sensitive to variations in the input power of the adaptive filter. Therefore, the NLMS algorithm is a good candidate for applications with a high degree of uncertainty about the filter input power. The statistical analysis of the NLMS algorithm is complicated by the normalized weight update.

The NLMS algorithm can be reviewed as a special case of the LMS algorithm with a time-varying step size, in which the step size varies with the input signal strength. The tap-weight adaptation equation of the NLMS algorithm is given by

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \frac{\hat{\mu}}{\mathbf{x}^T(n)\mathbf{x}(n)} e(n)\mathbf{x}(n) \quad (5)$$

where $\hat{\mu}$ is a parameter to be chosen, and $\hat{\mu}/\mathbf{x}^T(n)\mathbf{x}(n)$ is the actual step size. It can be seen from Fig. 1 that the output signal of the proposed method is

$$z_t(k) = \sum_{k=0}^{N-1} y_t(k-n)w(k) \quad (6)$$

A more reliable implementation of the NLMS algorithm in practice requires the assistance of a regularization parameter ε , and it results in the ε -NLMS algorithm [10] with the improved step size $\hat{\mu}/(\mathbf{x}^T(n)\mathbf{x}(n) + \varepsilon)$. Although having the tap-weight adaptation equations in a similar form, LMS-type algorithms are usually not compared with NLMS-type algorithms. It has been shown that NLMS-type algorithms provide a potentially faster rate of convergence.

4 Simulation Results

In this section, we apply an NLMS adaptive algorithm to noise reduction in speech applications, and examine the performance of the proposed noise reduction method. The audio data which sampled at 48 kHz was used to evaluate the proposed method, and the recorded time is 4.5 s. The waveform is shown in Fig. 2. For some audio signals such as speech signal, music signal are typical non-stationary signals, in order to achieve a good performance the speech signals must be segmented. The audio signal is segmented every 40 ms and a frame includes 1920 sampled data. In the experiments, zero mean white Gaussian noise is added to the audio signal, and the SNR is 0 dB.

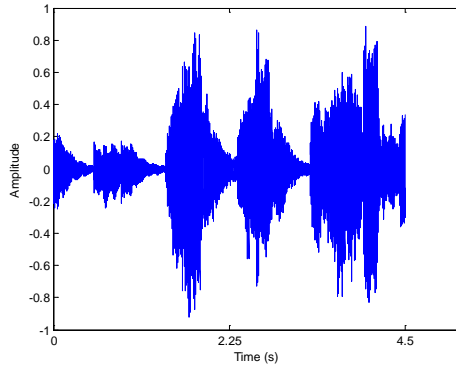


Fig. 2 The audio signal waveform

To implement the NLMS algorithm that a step size can be chosen to minimize the sum of the squares of the measured estimation errors and to stabilize the filter, we propose a suboptimum realization. The step size of the NLMS algorithm adjusted as the power of the input signals. If $\mathbf{x}(n)$ is the input vector, the correlation matrix of the input signal is $\mathbf{C}_{\mathbf{xx}}$. Let λ_m is the maximum eigenvalue of $\mathbf{C}_{\mathbf{xx}}$, then the regularization parameter is $\varepsilon = 1/\lambda_m$. Without loss of generality, we assume that the adaptation starts at time instant 0, where the input signal is $\mathbf{y}(0) = [y(0) \ y(-1) \ \dots \ y(-N+1)]^T$ ($N = 1920$), and the weight vector $\mathbf{w}(0)$ is chosen to be the zero vector.

The audio signal noise reduction performance of the proposed method is shown in Fig. 3. The proposed approaches improve noise audio signal quality. The noise canceled signal has a time domain waveform that is nearly equal to original. In this figure we have plotted the dissimilarity between the signals waveform of the clean audio signal and the noisy signal and then the noise canceled signal. It is noted that since the iteration of the adaptive filter is 100, the first 100 sampled points of the output signal is zero. An objective indication of the improvement afforded by the proposed adaptive noise cancelling scheme is illustrated in Fig. 4 and Fig. 5. It can be seen from Fig. 4 that the noise level in the noise canceled signal is reduced clearly. Fig. 5 indicates that the adaptive noise reduction method result in a noise variance that is more nearly equation to zero. The variance of the noise from the clean signal and the noisy signal nearly equals to 0.008. By contrast, the variance of the noise from the clean signal and the noise canceled signal nearly equals to 0.0015. From Fig. 4 and Fig. 5, we see that the proposed method is able to remove noise from the noisy speech signal for Gaussian noise and thus significantly enhance the speech signals.

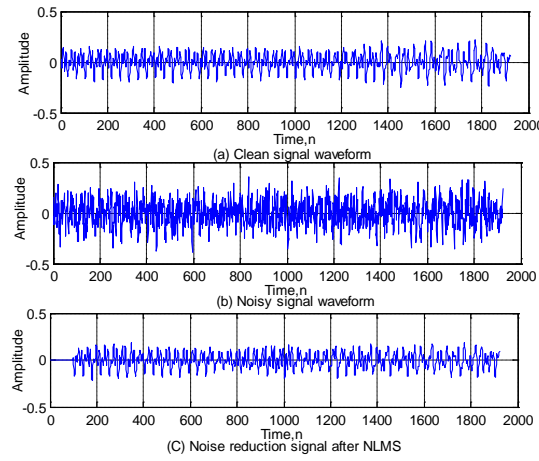


Fig. 3 The proposed noise reduction algorithm performance for audio signal

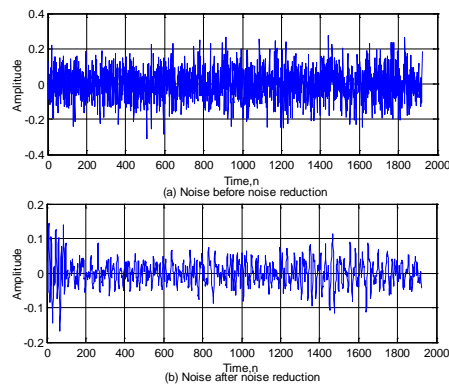


Fig. 4 Illustration of improvement in noise representation of audio signal

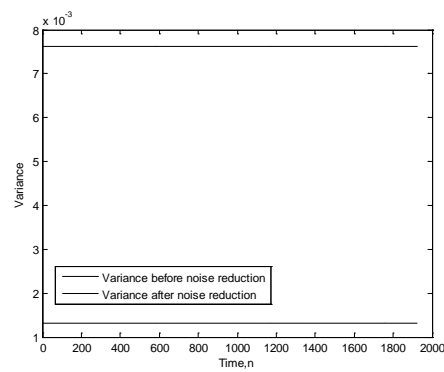


Fig. 5 Variance of the noise

5 Conclusions

In this paper, we have presented a noise reduction method for audio and speech signals by applying adaptive linear filtering technique. The noise reduction problem has been formulated as a filtering problem which is efficiently solved by using the NLMS method. In addition, the method pays attention to the non-stationary nature of some audio signal. Simulation results indicate that the proposed method can improve the performance the quality of noisy audio signal. Through computer simulations, we have demonstrated that the proposed method is quite effective in noise reduction, especially in the case of stationary white Gaussian noise.

6 References

1. Jingdong, C., Jacob, B., Arden, Huang. (2007). On the optimal linear filtering techniques for noise reduction. *Speech Communications*, 49(2), 305-316.
2. Akagi, M., Mizumachi, M. (1997). Noise reduction by paired microphones. *In Proceedings of EUROSPEECH, 1997*, 335-338.
3. Cohen, I., Berdugo, D. (2001). Speech enhancement for non stationary noise environments. *Signal Processing*, 81(11), 2403-2418.
4. Hamid, H. (2008). A time-frequency approach for noise reduction. *Digital Signal Processing*, 18(5), 728-738.
5. Bitzer, J., Simmer, K. U. (2001). Super directive microphone arrays. In: *Microphone arrays signal processing techniques and applications*. Springer, Berlin, 19-38.
6. Simmer, K. M., Bitzer, J., Marro, C. (2001). Post-filtering techniques. In: *Microphone arrays signal processing techniques and applications*. Springer, Berlin, 39-60.
7. Akagi, M., Kago, T. (2002). Noise reduction using a small-scale microphone array in multi noise source environment. *In Proceedings of ICASSP, 2002*, 909-912.
8. Ming, J., Srinivasan, R., Crookes, D. (2011). A Corpus-based approach to speech enhancement from nonstationary noise. *IEEE Transactions on Audio, Speech, and Language Processing*, 19(4), 822-836.
9. McAulay, R. J., Malpass, K. L. (1980). Speech enhancement using a minimum mean-square error short time spectral amplitude estimator. *IEEE Transactions on Acoustic, Speech, Signal Processing*, 28(2), 137-145.
10. Haykin, S (2001). Minimum mean square error adaptive filter. In: *Adaptive Filter Theory*, 4th ed. Prentice Hall, Upper Saddle River, 183-228.
11. Lotter, T., Vary, P. (2005). Speech enhancement by MAP spectral amplitude estimation using a super-Gaussian speech model. *EURASIP Journal in Applied Signal Processing*, 2005(1), 1110-1126.
12. Alouane, M., Jaïdane, M. (2006). A new nonstationary LMS algorithm for tracking Markovian time varying systems. *Signal Processing*, 86(1), 50-70.