

A Robust Fast LMM/Newton Algorithm for Generalized Side-lobe Canceller

Xianxiang Wang, Yi Zhou, Wenyi Yuan and Hongqing Liu
Chongqing University of Posts and Telecommunications
2# Chongwen Road, Nan'an District
Chongqing, 400065

s150131128@stu.cqupt.edu.cn, zhouy@cqupt.edu.cn, 1106550603@qq.com, hongqingliu@cqupt.edu.cn

ABSTRACT

For the conventional generalized side-lobe canceller (GSC), least square (LS) type adaptive filters have poor convergence performance when the speech enhancement system is corrupted by impulsive noise. In this paper, a robust GSC algorithm based on Fast-LMM/Newton algorithm is proposed. Experiments on dual-microphone system proved the proposed algorithm can achieve better robustness performance under the impulsive noise.

KEYWORDS

GSC, Fast-LMS/Newton, Fast-LMM/Newton, M- estimator

1 INTRODUCTION

With many years' development, speech enhancement has become an indispensable technique in various applications ranging from speech communication to speech recognition. Compared with the single-microphone based enhancement scheme, the dual-microphone technology [1] reveals its advantage by making full use of signal spatial information. When compared with the multi-microphone array, the dual-microphone solution is relatively simple and easy to be integrated into portable devices.

Speech enhancement technique based on dual-microphone is widely used due to its low complexity and high implementability. Various multi-microphone based speech enhancement techniques have been proposed, like delay-and-sum beamforming, minimum variance distortionless response (MVDR) beamforming, adaptive beamforming, and so on. MVDR beamforming retains the signal in the desired direction and minimizing the variance of the output noise [2]. The generalized side-lobe canceller (GSC) using adaptive filter proposed in [3] has been widely used due to its simple structure. It transforms the constraint problem of the linear constraint minimum variance (LCMV) beamformer into none constraint problem to eliminate the noise and interference. It is well known conventional LS-type adaptive algorithms will be severely affected by the impulse noise [7]. As a result, GSC system based on them will not be robust to impulsive interference. In [5] a Fast-LMS/Newton algorithm was proposed which explores the linear predictability of the speech signal to achieve fast convergence and with low complexity. Its robustness version, the Fast-LMM/Newton, was studied in [9] in the impulsive

interference scenario. In this paper, a new GSC adaptive robust algorithm based on [9] is proposed, which improves the robust performance under the non-stationary condition with impulsive interference. Compared with the traditional NLMS and Fast-LMS/Newton algorithms, the proposed algorithm has better robustness performance under the impulsive noise, and significantly improves the robustness of the GSC algorithm. The rest of the paper is organized as follows. In section 2, the GSC and conventional algorithms are introduced. Section 3 derives the proposed robust algorithm. Simulation results are given in section 4 and conclusions are drawn in section 5.

2 GSC Based on Fast-LMS/Newton Algorithm

The block diagram of GSC is shown in Fig.1. It has two distinct substructures which are shown as the upper and lower processing paths, and consists of three parts which are fixed beamformer (FBF), blocking matrix (BM), and adaptive noise canceller (ANC) [3]. The basic idea of GSC is to obtain the enhanced signal by using the output of the FBF in the upper path to subtract the signal filtered by the BM in the lower path adaptively with the ANC.

The fixed beamformer is typically delay-and-sum or filter-and-sum beamformer. The former is used in this paper. After the beam direction of the fixed beamformer is steered to the desired direction, the delay-and-sum signal. The signal received by the two microphones can be expressed as

$$x_i(n) = s(n - \tau_i) + v(n) \quad (1)$$

where $x_i(n)$ is the noisy speech signal, $s(n)$ denotes the pure speech signal from the direction θ , $v(n)$ is the additive noise from the direction θ_1 , l is the distance between the microphone, τ_i denotes the time delay of the i th microphone signal. Applying short time Fourier transform (STFT) as a fixed beamformer, $x(n)$ can be weighted as follows

$$\mathbf{d}(k) = \mathbf{x}_i^T(k) \mathbf{w}_{qi}(k) \quad (2)$$

$$\mathbf{w}_{qi}(k) = \frac{1}{2} [1, e^{-j\xi(k)}] \quad (3)$$

$\xi(n)$ represents the phase shift of the signal at the frequency bin k

$$\xi(k) = \frac{2\pi l F_c(k)}{c} \sin(\theta) \quad (4)$$

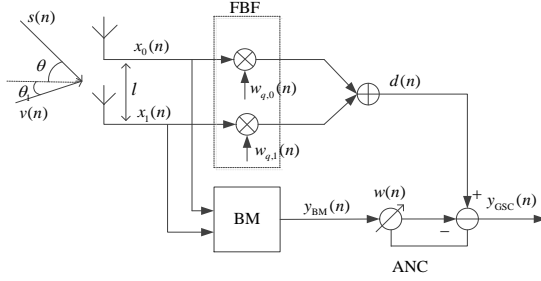


Figure 1: Block diagram of GSC.

where $F_c(k)$ is the center frequency and $c = 340m/s$.

The role of the blocking matrix is to block the desired signal in the lower path and let output only contains the interfering noise. It is possible to ensure that the output signal does not contain the desired direction signal as long as the matrix have $N - 1$ linearly independent rows which sum up to zero, where N denotes the number of the microphones. When the incident angle is 90 degrees, it can be concluded that the blocking matrix can be chosen as

$$\mathbf{B} = [1 \quad -1] \quad (5)$$

$$\mathbf{y}_{BM}(n) = \mathbf{B}\mathbf{x}(n) \quad (6)$$

In the conventional GSC scheme, the ANC module usually adopts the normalized least mean square (NLMS) algorithm. However, its convergence rate is too slow to suit for practical application. LMS/Newton type algorithms [5] bear improved convergence performance, which are formulated as

$$\mathbf{w}(n+1) = \mathbf{w}(n) + 2\mu e(n)\hat{\mathbf{R}}_{xx}^{-1}\mathbf{x}(n) \quad (7)$$

Where $\hat{\mathbf{R}}_{xx}$ is the covariance matrix for the input signal, μ is the step-size. For the above LMS/Newton algorithm, $\hat{\mathbf{R}}_{xx}$ is usually assumed known or can be approximated by its linear expectation. In [6] Boroujeny employed the linear prediction (LP) technique to approximate $\hat{\mathbf{R}}_{xx}$. Defining the vector $\mathbf{b}(n) = \mathbf{L}\mathbf{x}(n)$ as the reverse prediction error so its covariance matrix $\hat{\mathbf{R}}_{bb}$ is a diagonal matrix, which is reversible can be further written as

$$\mathbf{R}_{bb} = E\{\mathbf{L}\mathbf{x}(n)(\mathbf{L}\mathbf{x}(n))^T\} = \mathbf{L}\mathbf{R}_{xx}\mathbf{L}^T \quad (8)$$

Defining $\mathbf{u}(n) = \hat{\mathbf{R}}_{bb}^{-1}\mathbf{x}(n)$ and using (8) yields

$$\mathbf{u}(n) = \mathbf{L}^T\hat{\mathbf{R}}_{bb}^{-1}\mathbf{b}(n) \quad (9)$$

Substituting (9) into (7), the Fast-LMS/Newton algorithm can be obtained as

$$\mathbf{w}(n+1) = \mathbf{w}(n) + 2\mu e(n)\mathbf{L}^T\hat{\mathbf{R}}_{bb}^{-1}\mathbf{b}(n) \quad (10)$$

The Fast-LMS/Newton algorithm based on LP technique shows a significantly faster convergence performance over the NLMS algorithm [6] and thus is a promising solution to speech applications.

3 Improved GSC Algorithm

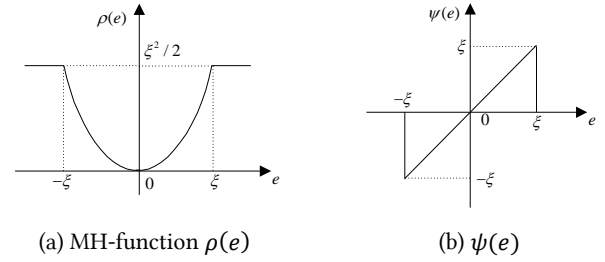


Figure 2: Modified Huber estimation function.

As introduced in the previous section, the Fast-LMS/Newton algorithm can be used in GSC to speed up convergence. However, like other LMS-type adaptive algorithms which are based on the LS estimation, when the noise is of impulsive nature, its convergence performance will be seriously affected. In this section, a robust GSC scheme based on Fast-LMM/Newton algorithm is proposed, which has improved stability in face of impulsive interference.

The commonly used robust statistical methods include the Order Statistic Least Mean Square (OSLMS) [8] algorithm, the LMS/Newton algorithm based on M-estimation (Least Mean M-estimate, LMM) [9] and the recursive M-estimate (RLM) algorithm [10]. The LMM and RLM algorithms are based on the M-estimation principle, which shows improved robustness to impulsive interference. Performance comparison of these two algorithms and other related algorithms can be referred to [9] and [10]. As shown in Fig.2, the cost function of LMM algorithm with M-estimation $J_\rho = E\{\rho[e(n)]\}$ replaces that of conventional LMS-type algorithm $J_{LS} = E[e^2(n)]$, in which $\rho[e(n)]$ is a modified Huber (MH) function with the evaluation function $\psi(e) = \partial\rho(e)/\partial e$. More specifically,

$$\rho_{MH}(e) = \begin{cases} e^2/2 & 0 \leq |e| < \xi \\ \xi^2/2 & \xi \leq |e| \end{cases} \quad (11)$$

where ξ is a threshold parameter that is used to compare with the real value of $e(n)$. When the latter becomes greater than the threshold, $\psi(e(n))$ is set to zero to halt the coefficient update so as to protect the algorithm from being divergent. When the $e(n)$ is smaller than the threshold, the algorithm is equivalent to the conventional Fast-LMS/Newton algorithm. When $\rho(e) = e^2/2$, the MH function will reduce to the traditional MSE cost function. It can be seen from Fig.2 that the smaller ξ is, the stronger suppression will be imposed on the impulsive interference.

Due to the time-varying nature of the processed signals, the value of ξ needs to be adaptively estimated. The optimal vector equation for the M-estimation equation is:

$$\mathbf{R}_\rho\mathbf{w}_\rho^* = \mathbf{P}_\rho \quad (12)$$

where $\mathbf{R}_\rho = E[q(e(n))\mathbf{x}(n)\mathbf{x}^T(n)]$, $\mathbf{P}_\rho = E[q(e(n))d(n)\mathbf{x}(n)]$ are respectively the M-estimation auto-correlation matrix and the cross-correlation vector $q(e) = (\partial\rho(e)/\partial\rho)/e = \psi(e)/e$. Recalling the optimal LS solution

$$\mathbf{w}_{LS}^* = \mathbf{R}^{-1}\mathbf{P} \quad (13)$$

The gradient vector can be expressed as

$$\nabla_w = 2\mathbf{R}_{xx}\mathbf{w} - 2\mathbf{P} \quad (14)$$

Multiplying the both sides of eqn. (14) by $\frac{1}{2}\mathbf{R}_{xx}^{-1}$ yields

$$\mathbf{w}_{LS}^* = \mathbf{w} - \frac{1}{2}\mathbf{R}_{xx}^{-1}\nabla_w \quad (15)$$

Incorporating a step size factor in (15) generates the LMS/Newton algorithm

$$\mathbf{w}(n+1) = \mathbf{w}(n) - \frac{1}{2}\mu\mathbf{R}_{xx}^{-1}\nabla_{w(n)} \quad (16)$$

where $\nabla_{w(n)} \approx \nabla_w\{e^2(n)\} = -e(n)\mathbf{x}(n)$ is similar to its instantaneous gradient value. In the M-estimation, the \mathbf{R}_{xx}^{-1} estimate in eqn. (16) is replaced by \mathbf{R}_ρ^{-1} , and $\nabla_{w(n)}$ is replaced by the M-estimated instantaneous gradient vector:

$$\hat{\nabla}_{w_\rho} = \nabla_{w_\rho}\{\rho(e(n))\} = q(e(n))e(n)\mathbf{x}(n) \quad (17)$$

$$q(e) = \frac{\varphi_{MH}(e)}{e} = \begin{cases} 1 & 0 \leq |e| < \xi \\ 0 & \text{others} \end{cases} \quad (18)$$

The threshold parameters are adjusted by the Adaptive Threshold Selection (ATS) [11] method. By introducing the formula (17) into the Fast-LMS/Newton algorithm, a new Fast-LMM/Newton robust algorithm can be obtained:

$$e(n) = d(n) - \mathbf{x}^T(n)\mathbf{x}(n) \quad (19)$$

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu q(e(n))e(n)\mathbf{u}_a(n) \quad (20)$$

$$\mathbf{u}_a(n) = \mathbf{L}^T \mathbf{R}_{bb}^{-1} \mathbf{b}(n) \quad (21)$$

The Fast-LMM/Newton algorithm can thus be introduced into the ANC module of GSC algorithm. In order to dynamically estimate the threshold parameter ξ , the algorithm needs to estimate the variance:

$$\hat{\sigma}_e^2(n) = \lambda_\sigma \hat{\sigma}_e^2(n-1) + c_1^2(n - \lambda_\sigma) \text{med}(A_e(n)) \quad (22)$$

where $A_e(n) = [e^2(n), \dots, e^2(n - N_w + 1)]$, $\text{med}(\cdot)$ is the median filter, λ_σ is the forgetting factor, c_1 is the finite length correction factor, N_w is the median filter window length, $\xi = k_\xi \hat{\sigma}_e(n)$ where k_ξ is the constant controlling algorithm convergence and the degree of impulse interference suppression.

4 Simulation Results

In this section, the NOISEX-92 database [12] is used for experimental testing. The spacing between the two microphones of the dual-microphone is 2.26 cm. It is assumed that the desired and interference sources are both located in the far field. The direction of arrival of the speech signal and the interference signal are 90 and 30 degrees. The test uttered speech is "blue sky, white clouds and green sea". The background noises are white, pink, and babble noise and speaker interference with various SNR being 0dB, 5dB and 10dB respectively. The impulsive interference are imposed to the desired signal respectively at the

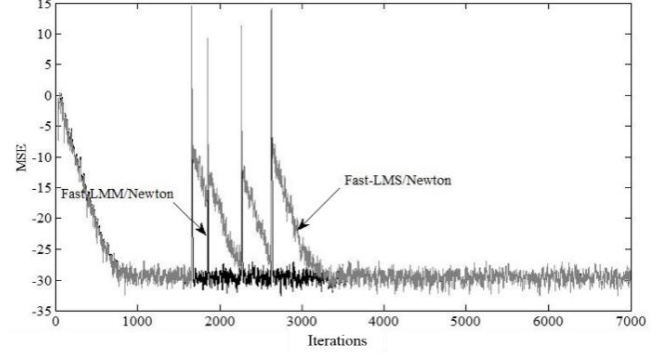


Figure 3: MSE learning curve of two algorithms.

beginning of the speech segment, in the speech active period and in the noise-only period. The sampling rate is 8 kHz. The order of adaptive FIR filter in GSC is $L = 256$. Other parameters of Fast-LMM/Newton algorithm are set as: $\mu_{LMM} = 0.8$, $\alpha_{LMM} = 0.9$, $\beta_{LMM} = 0.95$, $N_w = 64$, $c_1 = 2.13$, $\lambda_\sigma = 0.99$, $k_\xi = 1.96$.

In order to verify the robustness of the GSC based on Fast-LMM/Newton algorithm, three algorithms, the conventional GSC algorithm based on NLMS algorithm, the LMS/Newton method proposed in [2], and the proposed algorithm are compared.

Fig.3 depicts the MSE learning curves for the Fast-LMS/Newton and the Fast-LMM/Newton algorithms in identifying an $L = 256$ FIR system. The signals used are randomly generated with SNR being 30dB. The probability of the impulsive noise occurrence $P_r = 0.005$. The position of the impulsive noises are respectively at the time instants $P_{im1} = 1648$, $P_{im2} = 1841$, $P_{im3} = 2253$, $P_{im4} = 2620$. λ_e and λ_p are 0.99. $\xi = 1.96\hat{\sigma}_e(n)$.

As can be seen from Fig.3, the two algorithms reach the same steady state after about 1000 iterations. When the impulse interference occurs, the Fast-LMS/Newton algorithm experiences a divergence and re-convergence state whereas the Fast-LMM/Newton algorithm remains stable, which shows improved robustness. Fig.4 depicts the results when interference source is the white noise plus impulse noise. Figures (b)-(d) are the waveforms of the enhanced speech with the NLMS, Fast-LMS/Newton and Fast-LMM/Newton algorithms in the GSC algorithm. It is clear that the convergence performance of NLMS algorithm is weaker than that of the other fast algorithms. The NLMS and Fast-LMS/Newton algorithms diverge immediately after being disturbed by the impulse interference. However, the Fast-LMM/Newton algorithm shows significantly improved robustness.

The Global Signal-to-Noise Ratio (Global SNR) [10] is calculated as follows

$$SNR_i = 10 \log \frac{\sum_{n=0}^{M-1} x_i^2[n]}{\sum_{n=0}^{M-1} n_i^2[n]} \quad (23)$$

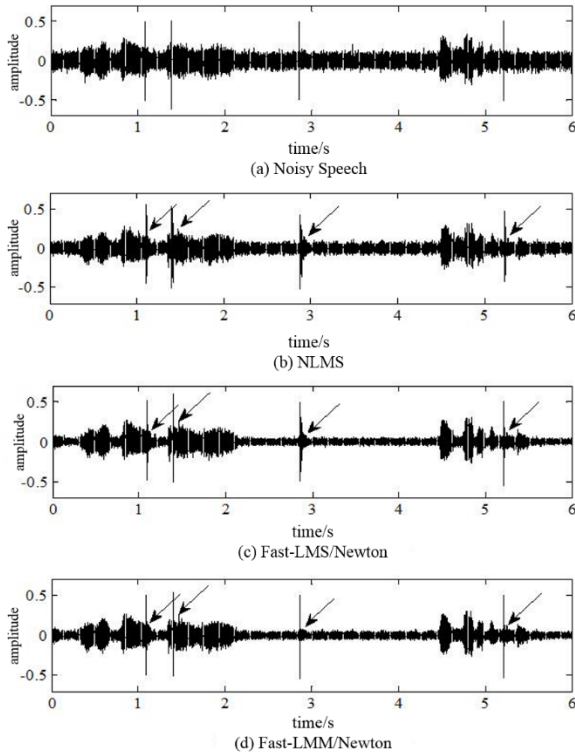


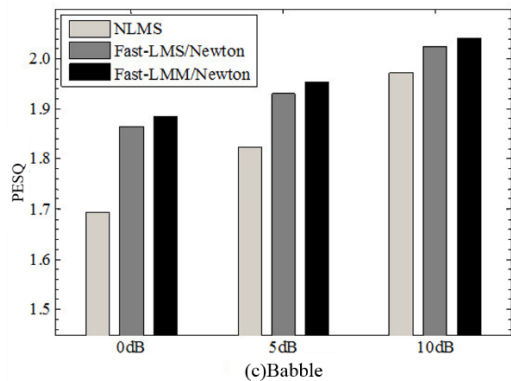
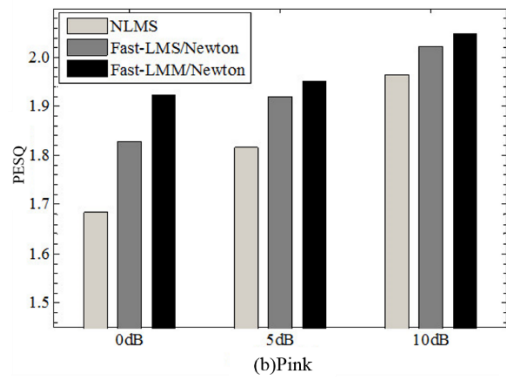
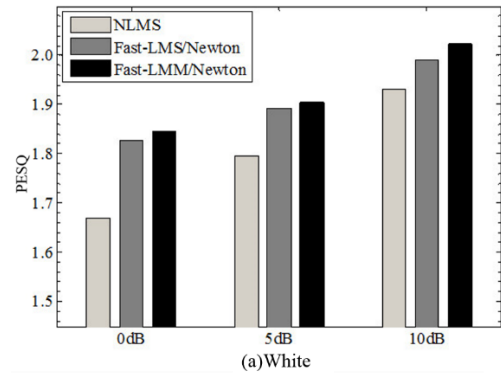
Figure 4: Time domain waveform of the enhanced speeches by three different algorithms.

where $x_i(n)$ is the enhanced speech of the i -th frame. $n_i(n)$ denotes the noise signal of the i -th frame, M denotes the speech signal frame index. Higher global SNR value indicates that the enhanced speech is closer to the pure speech and the speech enhancement algorithm is better. Table 1 shows the average SNR of the improved Fast-LMM/Newton algorithm is higher than that of the NLMS and the Fast-LMS/Newton algorithms, especially in the case of low SNR. The improved speech SNR increases for about 0.1dB. And under the low SNR condition, the speech SNR is improved for about 1dB, which indicates that the proposed algorithm has better anti-impulsive noise ability.

Then we will evaluate the speech quality based on the perceptual evaluation of speech quality (PESQ) of ITU-TP.862 scores. Fig.5 shows the objective quality measurements of enhanced speech for the NLMS, the Fast-LMS/Newton and the proposed algorithms in the GSC algorithm under three different SNRs and four different background noise types, respectively. Under the condition of high SNR, the PESQ values of the three algorithms are only slightly different. However, when the SNR is low, the PESQ value of the proposed algorithm under the white, pink and babble noises has increased at average of 0.3, while under the speaker interference condition, the PESQ value is slightly lower, which obtains improvement for about 0.2. The proposed algorithm is superior in impulsive interference suppression under the low SNR environment.

Table 1: The comparison of Global SNR

Noise type	Algorithm	SNR/dB		
		0	5	10
White	NLMS	4.87	8.63	13.2
	Fast-LMS	6.44	10.1	14.8
	Fast-LMM	7.3	10.4	14.9
Pink	NLMS	3.15	7.48	12.2
	Fast-LMS	5.27	9.25	13.3
	Fast-LMM	6.63	9.82	13.5
Babble	NLMS	4.53	8.59	13.1
	Fast-LMS	7.12	9.92	14.6
	Fast-LMM	7.81	11.0	14.7
Speaker interference	NLMS	4.69	9.01	13.6
	Fast-LMS	5.47	10.1	14.1
	Fast-LMM	6.61	10.4	14.2



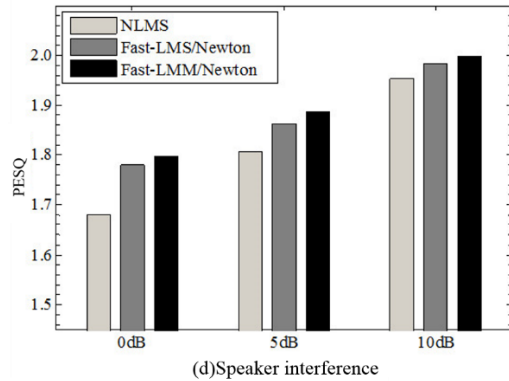


Figure 5: The objective quality measures of enhanced speech by three different algorithms.

CONCLUSIONS

In this paper, a Fast-LMM/Newton robust speech enhancement algorithm for GSC is proposed by incorporating the M-estimation function in the conventional Fast-LMS/Newton algorithm. Used in dual-microphone GSC scenario, the proposed algorithm shows superior robustness to the impulsive interference when compared with its non-robust counterpart algorithms.

ACKNOWLEDGMENT

This work is supported by the Chongqing Science and Technology Commission (cstc2015jcyjA40027) and National Science Foundation of China (No. 61501072).

REFERENCES

- [1] Valente, Michael, David Fabry, and Lisa G. Potts. "Recognition of speech in noise with hearing aids using dual microphones." *Journal of the American Academy of Audiology* 6.6 (1995).
- [2] Benesty, Jacob, Jingdong Chen, and Yiteng Huang. *Microphone array signal processing*. Vol. 1. Springer Science & Business Media, 2008.
- [3] Griffiths, Lloyd, and C. W. Jim. "An alternative approach to linearly constrained adaptive beamforming." *IEEE Transactions on antennas and propagation* 30.1 (1982): 27-34.
- [4] Sunitha, T., and R. Suja Mani Malar. "Nonlinear acoustic echo cancellation using adaptive filters." *Control, Instrumentation, Communication and Computational Technologies (ICCICCT)*, 2014 International Conference on. IEEE, 2014.
- [5] Farhang-Boroujeny, Behrouz. *Adaptive filters: theory and applications*. John Wiley & Sons, 2013.
- [6] Farhang-Boroujeny, Behrouz. "Fast LMS/Newton algorithms based on autoregressive modeling and their application to acoustic echo cancellation." *IEEE Transactions on Signal Processing* 45.8 (1997): 1987-2000.
- [7] Huber, Peter J., and Elvezio M. Ronchetti. "Robustness of design." *Robust Statistics*, Second Edition (2009): 239-248.
- [8] Haweel, Tarek I., and Peter M. Clarkson. "A class of order statistic LMS algorithms." *IEEE Transactions on Signal Processing* 40.1 (1992): 44-53.
- [9] Zhou, Yi, Shing-Chow Chan, and Ka-Leung Ho. "A new LMS/Newton algorithm for robust adaptive filtering in impulsive noise." *Signal Processing Conference, 2004 12th European*. IEEE, 2004.
- [10] Zou, Yuexian, S. C. Chan, and T. S. Ng. "A recursive least M-estimate (RLM) adaptive filter for robust filtering in impulse noise." *IEEE Signal Processing Letters* 7.11 (2000): 324-326.
- [11] Koike, Shiin'ichi. "Adaptive threshold nonlinear algorithm for adaptive filters with robustness against impulse noise." *IEEE Transactions on Signal Processing* 45.9 (1997): 2391-2395.
- [12] Loizou, Philipos C. *Speech enhancement: theory and practice*. CRC press, 2013.

- [13] Vondrasek, Martin, and Petr Pollak. "Methods for speech SNR estimation: Evaluation tool and analysis of VAD dependency." *Radioengineering* (2005).
- [14] Rix, Antony W., et al. "Perceptual evaluation of speech quality (PESQ)-a new method for speech quality assessment of telephone networks and codecs." *Acoustics, Speech, and Signal Processing, 2001. Proceedings.(ICASSP'01). 2001 IEEE International Conference on*. Vol. 2. IEEE, 2001.