Noise Reduction Using Fuzzy Median Filter for Audio Signals

Naoyuki Tamaru * and Rui Onizawa *

Abstract

This paper presents a fuzzy median filter for the removal of the Gaussian noises from digital sound signals. The aim of our proposed filter is to remove noises and to gain a clear sound while every detail of the sound is preserved like the original sound. By using the fuzzy membership function our filter is able to correct the detected noisy samples with simple procedures. As the membership function we select the absolute maximum difference in the sound signals between the center sample and the adjacent samples. The mean square error (*MSE*) for the proposed filter is reduced to the small value and the sharpness (*SH*) is kept almost unchanged to that of the original sound signal as compared to those for the conventional median filter.

Key Words : fuzzy filter, median filter, membership function, MSE (mean square error), sound noise

1. Introduction

Recently, the advancement in communication and entertainment devices including smartphones, tablet terminals, in-car navigation systems, portable music players, portable game machines, and so on is tremendous.

In audio signals including music and speech sounds, noise is a large problem because it annoys us and makes it difficult to recognize information. Therefore, it is very important to reduce or cancel or eliminate the noise as large as possible, while the details of the signals are conserved.

Table 1 shows a classification of noise reduction methods that are generally used for sound signals. They are able to be divided into two regions, in case of using a single microphone or multi microphones. Moreover, in the process of noise reduction there are two methods, that are spectrum (frequency) analysis and time-axis one.

In using a single microphone there are a spectrum subtraction (SS) [1, 2, 3], a running spectrum analysis (RSA) [3, 4, 5, 6], a Wiener filter [7, 8], a Kalman filter [9, 10], a lattice filter [11, 12], a median

* *School of Social Information Studies, Otsuma Women's University*

Table 1 The classification of noise reduction methods.

Microphone	Method	Processing	Analysis	References
Single	SS (Spectrum Subtraction)	Eliminate an estimated noise spectrum from received signals	Spectrum	(1), (2), (3)
	RSF (Running Spectrum Filter)	Select large time changes in signal spectrums	Spectrum	(3), (4), (5), (6)
	Wiener filter	Using a least mean square algorithm, estimate a noise free signal	Time	(7), (8)
	Kalman filter	Using a state equation and an observation equation, estimate a noise-free signal	Time	(9), (10)
	Lattice filter	Select a pitch frequency and its haromnic one in a voice signal	Time	(11), (12)
	Median filter	Select the median value among some samples before and behind the terget sample	Time	(13), (14)
Multi	Microphone array	Estimate a speaker's position, and reduct a noise by beam forming	Time	(15)
	ANC (Adaptive noise canceller)	After estimation of noise elements, add the noise element with the revese phase	Time	(16), (17), (18)

filter [(13), (14)] and so on. On the other hand, in using multi microphones, a microphone array [(15)] and an adaptive noise canceller (ANC) [(16), (17), (18)] are proposed.

In the SS method, the extracted noise spectrum is subtracted from the received signal spectrum. The noise is estimated from the signal during non-verbal periods, etc.

About the RSA method, a region of large changes in the time response for the received signal spectrum is selected because the noise spectrum does not largely change in the time response.

The Wiener filter uses a least mean square algorithm, and estimates the noise-free signal from the received signals. It assumes known stationary signal and noise spectra.

The Kalman filter has been applied in many fields such as a radar system, an in-car navigation system, a satellite control, an image enhancement [(19), (20)]. It produces estimates of noise-free signals from time varying observation signals with the noise. In the filter a state equation and an observation equation are modeled for the dynamical system. The state signals are estimated so as to become the estimation error smaller.

The lattice filter selects a pitch frequency and the harmonic one in voice signals.

A famous median filter is a simple smoothing operation with which one calculates a median value of the audio signals inside a moving window around the center samples. This filter has been recognized as a simple and useful audio signal correction technique while the clear and sharp sound is preserved to a certain degree.

The ANC estimates noise elements and adds these elements with the reverse phase to the received signal.

We have studied the fuzzy median filter in order to reduce the impulse noise from color images [(21)]. As the results of the fuzzy median were very effective in spite of the simple algorithm and the calculation, we will also apply the filter to the reduction of audio noises.

As we have effective results from the experiments to some degree, we show the method and the results.

In Section 2, we briefly describe the fuzzy median filter for the reduction of audio noises. We state a definition of the *MSE* (Mean Square Error) and the *SH* (Sharpness) in order to evaluate the corrected signals in Section 3, and the experimental results by the computer are shown there. Finally, the paper is briefly summarized in Section 4.

2. Fuzzy Median Filter

2.1 Gaussian Noise

For the experiments we use a ringing bell sound of cell phones and a musical sound which are coded by the WAVE audio file format with a 16-bit resolution.

The Gaussian noise is computed as follows. We generate twelve fractional random numbers from 0.0 to about 0.9999 by using a library function of "rand" in the C programming language so as to rand( )/ RAND_MAX, where the RAND_MAX in general is 32,768 (=2^{15}). In addition, we calculate the sum S of these twelve numbers, and do n = (S - 6.0) / 12.0. The range of n is from -0.5 to 0.5, and the average is 0.0. The n multiplied by AMP (i.e. n×AMP) is the noise, which is added to the source signals. This additive noise is not real Gaussian, but pseudo-Gaussian exactly. The AMP changes to 1.0E4, 1.0E5, and 1.0E6 in the experiments.

The noise rate is changed as follows. We also generate one fractional random number r from 0.0 to about 0.9999 the same as before, and choose a slice level s, which is from 0.0 to 1.0. If the random number r is greater than the slice level s, the pseudo-Gaussian noises are added to the samples. But the number r is not greater than the level s, the noise is not added. Therefore if the level s is 0.8, the noise rate is 20 %.

2.2 Median Filter

The median filter outputs the median value of the samples in the (2k+1) window W_i, that is centered at X_i, where k is an integer greater than or equal one. The output M_i of the median filter is defined as

$$W_i = \{X_{i-k}, \cdot\ \cdot, X_i, \cdot\ \cdot, X_{i+k}\} \qquad (1)$$

$$M_i = \mathrm{median}\{W_i\}. \qquad (2),$$

where "median" means the calculation procedure of the median value.

2.3 Fuzzy Median Filter

We introduce a fuzzy flag f_i indicating how much the received signal looks like a noisy sample. The flag f_i is the membership function with two parameters, defined as

$$f_i = \begin{cases} 0 & m_i \leqq T_1 \\ (m_i - T_1)/(T_2 - T_1) & T_1 \leqq m_i \leqq T_2 \\ 1 & m_i \geqq T_2, \end{cases} \qquad (3)$$

where T_1 and T_2 are two pre-determined parameters, and m_i, the maximum value of the absolute differences between the values of the center sample X_i and the adjacent samples, is defined as follows:

$$m_i = \max\{|X_i - S_i|\}$$

$$S_i \in W_i \text{ and } X_i \neq S_i .$$

The value m_i shows a measure for detecting noises.

After the membership function f_i is calculated, the filtered output of the signal Y_i, is defined as

$$Y_i = (1 - f_i) \times X_i + f_i \times M_i \qquad (4)$$

From (4), when the function f_i is 1.0, the output Y_i equals the value of the median filter, and if f_i is 0.0, the output Y_i equals the value of the original signal X_i , i.e. the unchanged value.

3. Experimental Results

3.1 MSE

In order to compare quantitatively the performance of these filters we have discussed, the mean square error (MSE) between the input and output signals is evaluated. The MSE is given by

$$MSE = \Sigma\ (X_i - Y_i)^2 / N \qquad (5),$$

where X_i and Y_i are the values for the input and the output signal at sample position i, respectively, N is the number of the total samples. The MSE is able to evaluate the residual noise for the corrected signals.

3.2 SH

In addition, in order to evaluate the sharpness (SH) of sound signals, i.e. the preservation of the detail, we introduce the sharpness (*SH*) as follows:

$$SH = \Sigma \; (X_i - X_{i+1})^2 \, / \, (N - 1) \qquad (6),$$

where X_{i+1} is the sample value just right adjacent sample to X_i. If the *SH* value for the corrected signal is close to that for the original signal (i.e. the noise-free signal), we can estimate that the corrected signals preserve the details and the clear sound without noises.

3.3 Test Sounds

We use two sound sources in the experiments, the bell sound of cell phones, and the vocal music sound. Those signals have a mono channel of the 16-bit resolution with a sample rate of 22.05 kHz. The bell sound is composed with some simple spectrums, and a period of the bell ringing is about 0.95 s, one of the no-ringing is 1.35 s. The duration of the test sound is about 3.6 s. On the other hand, that of the music sound is about 6.4 s.

3.4 Experiment Results

Figure 1 is illustrated an example of a fuzzy membership function f_i with two parameters, T_1 = 15,000 and T_2 = 25,000. The horizontal axis shows m_i, that is the maximum value of absolute differences between the sample value of the center and the adjacent sample values. The vertical axis is the fuzzy membership degree. It is made up of three piecewise linear segments. The ranges of $m_i < T_1$, $T_1 < m_i < T_2$, and $m_i > T_2$ show the sharp

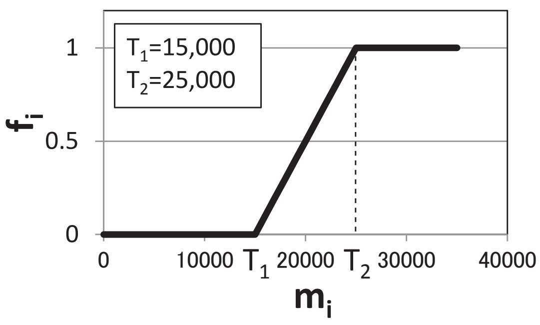

Figure 1 An example of a fuzzy membership function.

preservation of the original signals, the partial augmentation, and the full correlation using the median filter, respectively.

Figure 2 shows the waveforms for the bell sound, (a) is the source signal with noise-free, (b) is the source signal adding the noise signal with the noise amplitude *AMP*=1.0E4 and the noise rate of 20 %, (c) is the median filter signal with the window length *Lw*＝5, and (d) is the proposed filter signal (i.e. the fuzzy median filter signal) with the window length *Lw*＝5, T_1＝5 and T_2＝24,000. The horizontal axis shows the time dimension by the sample number, i.e. one sample period is 45.4 μs (＝1 second/22.05 kHz).

From the Figure 2(c) the every waveform, especially for the beginning region is destroyed heavily, though the additive large noises are almost removed. On the other hand, the proposed filter signal (d) is almost similar to the source signal (a), in spite of existing small noises that are left partially.

Figure 3 shows the *MSE* values obtained by

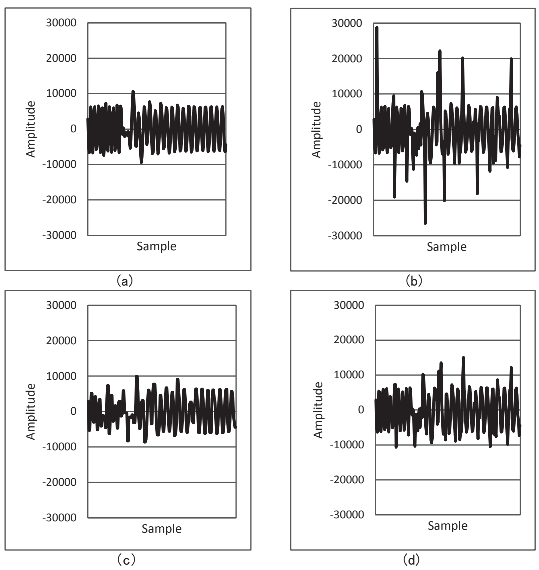

Figure 2 Waveforms for the bell sound, (a) is the source signal with noise-free, (b) is the source signal adding the noise signal, (c) is the median filter signal, and (d) is the proposed filter signal.

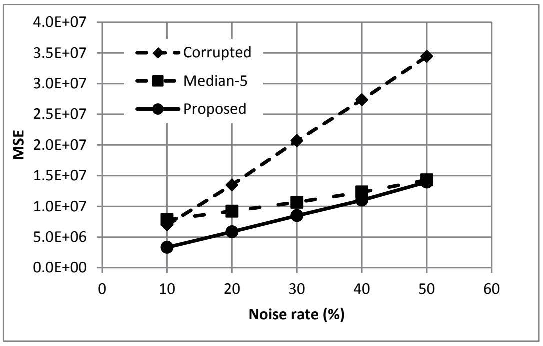

Figure 3 The relationship between the *MSE* and the noise rate for the bell sound.

varying the noise rate for the bell sound. The horizontal axis shows the noise rate within the range from 10 % to 50 %. The dotted line with the diamond plots displays the *MSE* values for the corrupted signal by the noise whose amplitude is 1.0E5, the broken line with the square plots does these for the conventional median filter with the window length of 5, and the solid line with the circular plots does these for the proposed filter with the window length of 5 and $T_1 = 5$, $T_2 = 24{,}000$. If the *MSE* values are smaller than those for the corrupted signal, there is a reduction effect for the noise. From the figure it is seen that the proposed filter is better than the conventional median filter, because the *MSE* for the proposed filter is smaller. Especially on the noise rate below 30 % the proposed filter is more effective for the noise reduction.

Figure 4 is presented the relationship between the *MSE* values and the noise rate for the proposed filter, when $T_1 = 5$ and $T_2 = 24{,}000$. It is found that the results of the window lengths of 3 and 5 are more effective and the residual noises decrease to the smaller values than a half of that for the corrupted signal.

Figure 5 shows the *MSE* values against the window length Lw for the proposed filter. It is seen also that the window lengths of 3 and 5 are optimal for the every range of the noises.

Figure 6 represents a relationship between the *MSE*

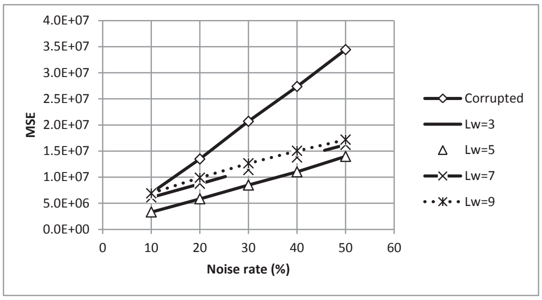

Figure 4 The relationship between the *MSE* and the noise rate against the window length.

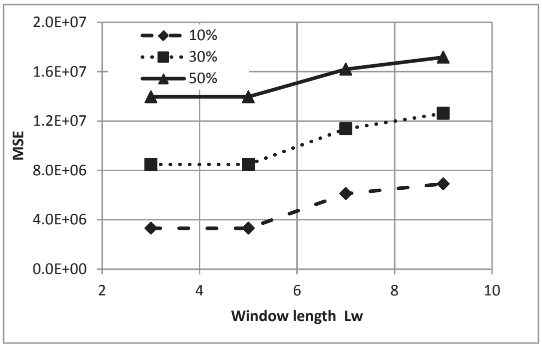

Figure 5 The relationship between the *MSE* and the window length against the noise rate.

values and the parameter T_2 for the membership function, when T_1 is kept 5, the noise amplitude is 1.0E5, and the rate is 20 %. The horizontal axis shows the parameter T_2 by the logarithmic scale. From this figure the optimum value of T_2 is 2.0E4 in case of Lw = 5, and is 8.0E3 in case of Lw =

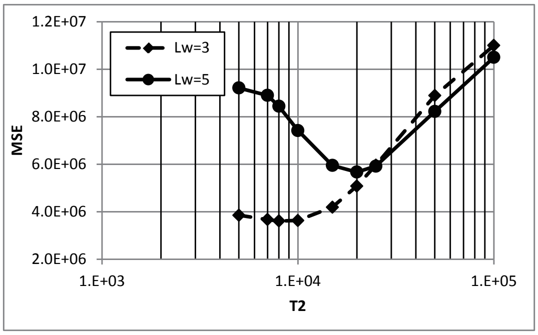

Figure 6 The relationship between the *MSE* and the parameter T_2 against the window length.

3. The optimum value means the parameter T_2 where the *MSE* values become the minimum value. The optimum values of T_2 strongly change by the window length.

On the other hand the optimal value for the parameter T_1 is about 5 in every case, such as against the window length, the noise rate, the noise amplitude and the parameter T_2.

Figure 7 shows the *SH* (sharpness) values obtained by varying the noise rate for a noise amplitude of 1.0E5. The *SH* values for the original sound, i.e. the noise-free sound, is 1.31E7. When the *SH* values for the corrected signal are closer to 1.31E7, the detail preservation for the corrected signal is kept more.

In the figure the broken curve with the circular plots displays the *SH* values for the median filter with the window length of 5, and the solid curve with the square plots does these for the proposed filter with the window length of 5 and T_1 = 5, T_2 = 24,000. The closer to 1.31E7 the *SH* value for the corrected signal is, the better the detail in the corrected signal keeps preserved. Therefore the proposed filter almost does not lose the detail of the signal below the noise rate of 20 %. On the other hand, as the *SH* value for the conventional median filter is approximate 5.0E6, it is worse than the proposed filter at least below the noise rate of 30 %.

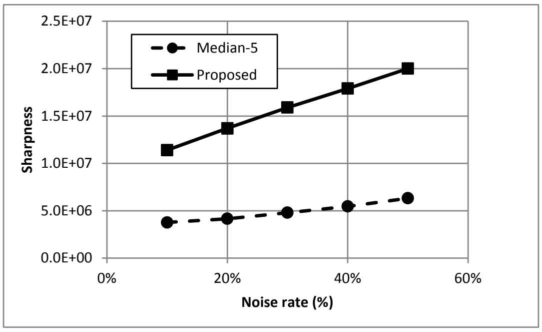

Figure 7 The relationship between the *SH* (sharpness) and the noise rate for the bell sound.

Figure 8 represents the *SH* values against the noise rate for the proposed filter with the noise amplitude of 1.0E5, and T_1 = 5, T_2 = 24,000. From this figure it is shown that the better choice is *Lw* = 5 in case of the noise rates of 10 % and 20 %, but over 30 % that is *Lw* = 7.

Figure 9 displays the *MSE* values against the noise rate in case of the music sound. The dotted line with the diamond plots shows the *MSE* values for the corrupted signal by the noise whose amplitude is 1.0E5, the broken line with the square plots does these for the median filter with the window length of 5, and the solid line with the circular plots does these for the proposed filter with the window length of 5 and T_1 = 5, T_2 = 24,000.

From this figure we can see that the both filters decrease the residual noise smaller to below a third

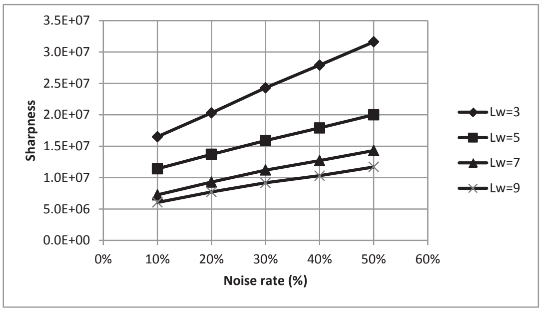

Figure 8 The relationship between the *SH* (sharpness) and the noise rate against the window length.

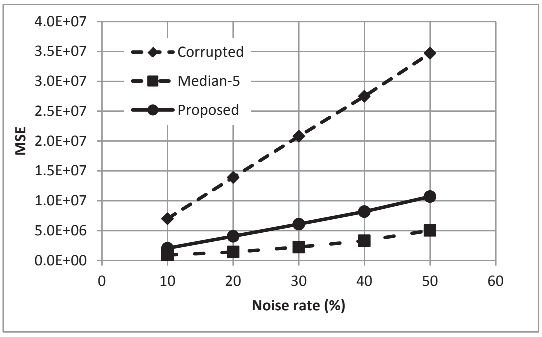

Figure 9 The relationship between the *MSE* and the noise rate for the music sound.

than the corrupted signal. And the median filter is more effective a little than the proposed filter, because it is assumed that there are not much segments for the high frequency in the test sound area.

Figure 10 is presented a relation between the *MSE* values and the noise rate in case of the music sound, when $T_1 = 5$ and $T_2 = 24{,}000$ for the proposed filter. It is found that the window lengths *Lw* of 7 and 9 are more effective and the residual noises decrease to the values smaller than a third of that for the corrupted signal.

Figure 11 shows the *SH* (sharpness) values obtained by varying the noise rate for a noise amplitude of 1.0E5. The *SH* value for the original sound, i.e. the noise-free sound, is 2.1E6. The median filter is more effective than the proposed filter.

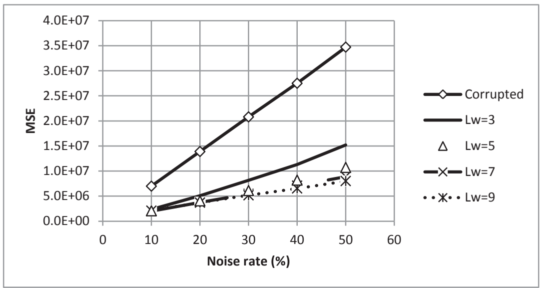

Figure 10 The relationship between the *MSE* and the noise rate against the window length.

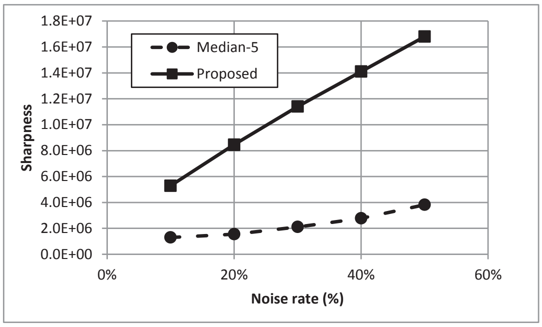

Figure 11 The relationship between the *SH* (sharpness) and the noise rate for the music sound.

3.5 Comparison of the Results to the Image Signals

By using the fuzzy median filter the impulse noise reduction for color images was more effective than the conventional median filter almost in every noise condition [21]. But in case of sound signals its effect is restricted in the specific noise conditions, and is not as large as that for the images. The first reason is because the window size for images is larger than that for sound. For example when the integer *k* for the window in the Equation (1) is one, the window size is 9 (= 3 × 3) for the images, but that is only 3 for the sounds. The second reason is that the signal range for images is from 0 to 255 including the noises, but the sound signals is not limited. Moreover, the image noises are restricted to impulse ones with the levels of about 0 and about 255, but the sound noises are the pseudo-Gaussian noise.

From the listening tests the residual noise in the corrected signal for the median filter does not annoy at all, but the clearness or sharpness of the sound is lost consequently the sound is flat, and dull. On the other hand, for the proposed filter the residual noise is perceived to some extent, but the clearness or sharpness is left greatly. These results are also drawn an analogy from the experimental results shown in Figures 2, 3, 7, 9 and 11.

4. Conclusion

This paper presents a fuzzy median filter for the removal of the additive Gaussian noises in the sound signals. The aim of our proposed filter is to remove the noises and to gain the clear sound signals while every detail of the signals is preserved like the noise-free source. By using the fuzzy membership function, our filter is able to correct the detected noisy signals with simple procedures. As the membership function we select the absolute maximum difference in the sound value between the center sample and the adjacent samples. The mean square error (*MSE*) for the proposed filter is reduced to the smaller value, and the sharpness (*SH*) is kept almost unchanged to that of the original signal as compared to the conventional median filter, but the noise reduction effect is restricted to the condition of the noises, because the filter window for the sound signals is smaller than that for the image signals.

References

(1) Y. Nomura, H. Tozawa, J. Lu, H. Sekiya and T. Yahagi, "Musical noise reduction by spectral subtraction using morphological process", IEICE Trans. on Information and Systems, Vol. J89-D, No. 5, pp. 991-1000 (2006) (in Japanese)

(2) H. Shibayama and H. Hosoi, "Denoising of background noise mixed in chirping of insects", IEICE, Proceedings of general conference, A-10-6, p. 205 (2006) (in Japanese)

(3) N. Ohtsuki and Y. Miyanaga, "A method of musical noise suppression in speech noise reduction", IEICE Technical Report on Smart Info-media System, SIS2007-18, Vol. 107, No. 94, pp. 29-33 (2007) (in Japanese)

(4) N. Ohtsuki, K. Ohnuki and Y. Miyanaga, "The improvement of voice quality in speech noise reduction using RSF", IEICE Technical Report on Smart Info-media System, SIS2005-1, Vol. 105, No. 111, pp. 1-5 (2005) (in Japanese)

(5) K. Fujioka, N. Hayasaka, Y. Miyanaga and N. Yoshida, "A noise reduction method of speech using running spectrum filtering", IEICE Trans. on Information and Systems, Vol. J88-D-II, No. 4, pp. 695-703 (2005) (in Japanese)

(6) N. Ohtsuki, Y. Uchikawa and Y. Miyanaga, "Speech noise reduction using high-precision RSA", IEICE Technical Report on Smart Info-media System, SIS2008-15, Vol. 108, No. 86, pp.7-10 (2008) (in Japanese)

(7) N. Kyoya and K. Arakawa, "A method of impact noise reduction in speech signal with Gaussian-mixed type adaptive Wiener filter", IEICE Technical Report on Smart Info-media System, SIS2009-51, Vol. 109, No. 447, pp. 17-22 (2010) (in Japanese)

(8) K. Ishida and T. Tanaka, "Noise removal of multi-channel speech signals by spatio-temporal Wiener filter", IEICE Technical Report on Signal Processing, SIP2005-155, Vol. 105, No. 635, pp. 75-80 (2006) (in Japanese)

(9) N. Tanabe, T. Furukawa and S. Tsujii, "Robust noise suppression algorithm with the Kalman filter theory for white and colored disturbance", IEICE Trans. on Fundamentals, Vol. E91-A, No. 3, pp. 818-829 (2008)

(10) W. Kim and H. Ko, "Noise variance estimation for Kalman filtering of noisy speech", IEICE Trans. on Information and Systems, Vol. E84-D, No. 1, pp. 155-160 (2001)

(11) A. Kawamura and T. Katahama, "A speech enhancement using lattice filter with correlation control algorithm", IEICE Trans. on Fundamentals, Vol. J92-A, No. 2, pp. 84-93 (2009) (in Japanese)

(12) A. Kawamura, "DSP implementation of a noise suppressor using lattice filter", IEICE Technical Report on Smart Info-media System, SIS2007-65, Vol. 107, No. 374, pp. 41-46 (2007) (in Japanese)

(13) S. Herzog, "Efficient DSP implementation of median filtering for real-time audio noise reduction", Proceedings of 16th Int. Conference on Digital Audio Effects (DAFx-13) Sep. 2-5 (2013)

(14) C. T. Lu, "Reduction of musical residual noise using block-and-directional-median filter adapted by harmonic properties", Journal of Speech Communication, Vol. 58, march, pp. 35-48

(2014)

(15) T. Narita, T. Hanazawa and J. Ishii, "Noise reduction using a small scale microphone array for speech recognition", IEICE Trans. on Information and Systems, Vol. J96-D, No. 11, pp. 2784-2794 (2005) (in Japanese)

(16) T. Tamashiro and T. Wada, "Investigation on harmonic interference canceller (HIC) for AM radio receiver", IEICE Technical Report on Circuits and Systems, CAS2012-140, SIP2012-171, CS2012-146, Vol. 112, No. 485, pp. 253-258 (2013) (in Japanese)

(17) M. kobayashi, T. Chiso, O. Eguchi and Y. Itho, "A study on an adaptive noise canceller", IEICE Technical Report on Signal Processing, SIP2005-87, Vol. 105, No. 293, pp. 13-18 (2005) (in Japanese)

(18) S. A. Hadei and M. Lotfizad, "A family of adaptive filter algorithms in noise cancellation for speech enhancement", International Journal of Computer and Electrical Engineering, Vol. 2, No. 2 (2010)

(19) N. Tamaru and A. Nagayama, "Positioning control on a torsion resonant system using a Kalman filter", Trans. of SICE (Society of Instrument and Control Engineers), Vol. 20, No. 7, pp. 641-647 (1984) (in Japanese)

(20) M. Nikaido and N. Tamaru, "Noise reduction for gray image using a Kalman filter", SICE (Society of Instrument and control engineers), Proceedings of Annual Conference, pp. 92-95 (2003)

(21) N. Tamaru and R. Onizawa, "Impulse noise reduction using Fuzzy median filter for color image", Otsuma Journal of Social Information studies, Vol. 22, pp. 123-128 (2013)