Impulsive Noise Suppression Using Interpolated Zero Phase Signal

Ryo Takehara* Arata Kawamura* and Youji Iiguni*
* Osaka University, Osaka, Japan
E-mail: takehara@sip.sys.es.osaka-u.ac.jp

***Abstract*—This paper proposes an impulsive noise suppression method from a noisy speech signal. The proposed method is based on a zero phase (ZP) signal. The ZP signal converted from the impulsive noise has large values only around the origin. On the other hand, the ZP signal of a voiced speech signal is periodic. Hence, the impulsive noise suppression can be achieved by replacing the ZP signal around the origin with the signal around the start position of the second period. Unfortunately, it is no guarantee to get the exact start position of the second period, since the ZP signal is defined only on discrete time indexes. In this paper, we propose a method to more accurately detect the start position of the second period by using the sinc interpolation or the cubic spline interpolation. Simulation results show that the proposed method can improve the impulsive noise suppression capability in comparison to the conventional method.**

I. Introduction

In recent years, with the development of speech communication and speech recognition technologies, devices such as smart phones and car navigation systems are rapidly spreading. Since these devices are used in various environments, noise signals generated from ambient noise sources or in a transmission process are added to a desired speech signal. Since the noise signal causes degradation of speech quality and recognition accuracy, the noise suppression technology is required.

Stationary noises can be suppressed by using spectral processing technologies such as spectral subtraction [1], Wiener filtering [2] and statistical estimation methods based on a speech spectral distribution [3]-[5]. On the other hand, it is difficult to suppress impulsive noises generated from switching, mouse clicking, clap, etc., because the impulsive noise disappears in a short duration and we do not know when it occurs. Miyake et al. classify noise features based on Gaussian mixture model (GMM) and achieve to suppress various types of noise [6]. Basically, unclassified noises cannot be suppressed. Duan et al. use non-negative spectral decomposition to separate the speech part and the noise part to suppress noise [7]. It is necessary to learn the component of noise as a dictionary beforehand. Unfortunately, noise suppression performance is changed depending on type of the training data, and appropriate noise suppression is not always achieved.

A method based on a zero phase (ZP) signal [8] is more attractive which do not rely on the classified noise or training. The ZP signal is defined as the inverse discrete Fourier transform (IDFT) of the amplitude spectrum [8], [9]. Thus, the ZP signal can be interpreted as DFT analysis of the amplitude spectral shape. The ZP noise suppression methods [8], [9] utilize periodicity of the ZP signal of the desired speech signal. Since the ZP signal of an impulsive noise has values around the origin, replacing ZP samples around the origin with corresponding samples in the second period achieves impulsive noise suppression [8], [9]. In this process, it is necessary to search for the start position of the second period of the ZP signal. Unfortunately, it is no guarantee to get the exact start position of the second period, since the zero phase signal is defined only on discrete time indexes.

In this research, we propose a method to replace ZP signal around the origin with a more appropriate samples. In general, the true starting position of the second period is located between the discrete indexes of the ZP signal. Hence, we interpolate the ZP signal, and get a more accurate start position of the second period. Replacing the ZP signal around the origin with the ZP signal around the interpolated second period, we have improved the noise suppression capability.

In this study, two kinds of interpolation methods are used. The first is sinc interpolation using sinc function as basis function [10]. The second is cubic spline interpolation [11]. Although the interpolation accuracy of the cubic spline is inferior to sinc interpolation, the cubic spline requires less amount of calculation than sinc interpolation. Simulation results show that the start position of the second period of the desired ZP signal can be more accurately estimated with the interpolated ZP signal in comparison to the conventional method. Although the proposed method increases computational cost, the noise suppression capability is superior to the conventional ZP method.

II. Impulsive noise suppression using zero phase signal

A. Definition of zero phase signal

Let $x(n)$ be an observed signal which is segmented and windowed by a window function whose frame size is N, and n denotes time index $(n = 0, 1, \dots, N-1)$. The N-points DFT of $x(n)$ given as

$$\begin{aligned} X(k) &= \sum_{n=0}^{N-1} x(n) e^{-j\frac{2\pi k}{N}n} \\ &= |X(k)| e^{j\angle X(k)}, \end{aligned} \tag{1}$$

where $|\cdot|$ and $\angle\{\cdot\}$ denote the amplitude spectrum and the phase spectrum, respectively, and k $(0, 1, \dots, N-1)$ denotes the frequency index. The ZP signal $x_0(n)$ for $x(n)$ is defined as [8]

$$x_0(n) = \frac{1}{N}\sum_{k=0}^{N-1} |X(k)|^q e^{j\frac{2\pi n}{N}k}, \tag{2}$$

where q is a real value. Conversely, it is possible to obtain $|X(k)|$ by taking the DFT of $x_0(n)$ as

$$|X(k)| = \left\{\sum_{n=0}^{N-1} x_0(n)e^{-j\frac{2\pi k}{N}n}\right\}^{\frac{1}{q}}. \tag{3}$$

Through this paper, we put $q = 1$.

B. Zero phase signal of impulsive noise

We consider a typical impulsive noise such that only one component of $d(n)$ $(n = 0, 1, \ldots, N-1)$ has a value and the others are zero given as

$$d(n) = \alpha_0\delta(n-p), \tag{4}$$

where $\delta(\cdot)$ is unit impulse function, and $0 \le p < N$. The DFT amplitude spectrum of $d(n)$ is given as

$$|D(k)| = |\alpha_0|. \tag{5}$$

Taking IDFT for (5), we have

$$d_0(n) = |\alpha_0|\delta(n). \tag{6}$$

The ZP signal becomes an impulse function with $p = 0$.

C. Zero phase signal of periodic signal

We assume that voiced speech is a periodic signal consisting of fundamental frequency k_c $(\le N/2)$ and its harmonics [12]. In this case, the amplitude spectrum of the voiced speech $s(n)$ has values at equal intervals as

$$|S(k)| = \sum_{i=1}^{\lfloor \frac{N}{2k_c} \rfloor} \frac{\alpha_i}{2}\{\delta(k - ik_c) + \delta(k + ik_c - N)\}, \tag{7}$$

where the nonnegative real value α_i $(i = 1, 2, \ldots, \lfloor N/2k_c \rfloor)$ is the ith harmonic amplitude of $s(n)$ and $\lfloor\cdot\rfloor$ represents a floor function. Substituting $|X(k)| = |S(k)|$ into (2), we have the ZP signal of $s(n)$ given as

$$s_0(n) = \sum_{i=1}^{\lfloor \frac{N}{2k_c} \rfloor} \frac{\alpha_i}{N}\cos\left(\frac{2\pi i k_c}{N}n\right). \tag{8}$$

The ZP signal $s_0(n)$ is a periodic signal having a cycle of N/k_c.

D. Impulsive noise suppression using ZP signal

Let $s(n)$ be the desired speech signal whose amplitude spectrum is represented by (7). Also, let $d(n)$ be an impulsive noise. It is assumed that each segmented frame does not include two or more impulsive noise [8], [9], [12]. Then, we assume that an observed signal is given as

$$x(n) = s(n) + d(n). \tag{9}$$

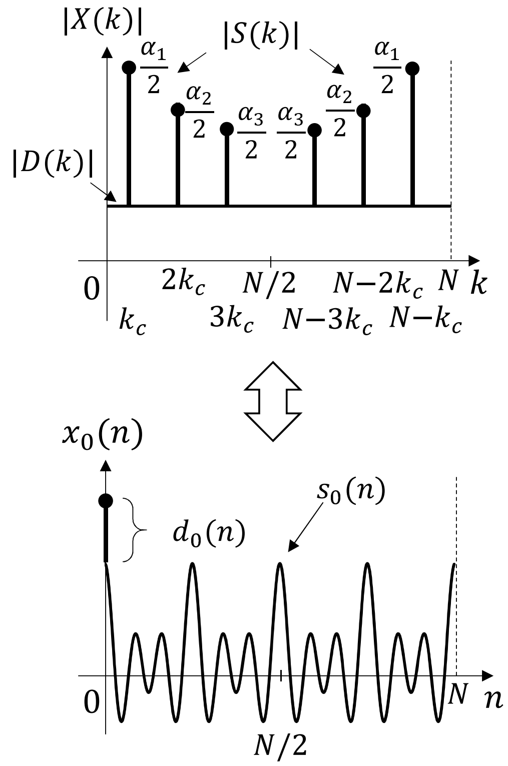

Fig. 1. Linearity of amplitude spectrum and ZP signal.

Taking DFT for (9), we obtain

$$X(k) = S(k) + D(k). \tag{10}$$

In the same manner as the conventional noise reduction methods [1]-[12], we assume that the spectral phase of the estimated speech signal is equal to that of the observed signal. It means that

$$|X(k)| = |S(k)| + |D(k)|. \tag{11}$$

Taking IDFT for (11), we obtain the ZP signal as

$$x_0(n) = s_0(n) + d_0(n). \tag{12}$$

Figure 1 shows the relationship of (11) and (12). On the upper side, vertical axis shows the amplitude spectrum $|X(k)|$ and the horizontal axis shows frequency index k. Here, k_c and α_1 denotes the fundamental frequency of $s(n)$ and its amplitude, respectively. In addition, α_2 and α_3 are the amplitudes of second and third harmonics, respectively. On the lower side, vertical axis shows the ZP signal $x_0(n)$ given as $s_0(n)$ plus $d_0(n)$, and the horizontal axis shows time index. Note that the ZP signal has the dimension of time. The impulsive noise exists only at the origin in the ZP signal.

We obtain the estimated ZP speech signal by using the replacement technique given as [9]

$$\hat{s}_0(n) = \begin{cases} g(n,T)\cdot x_0(n+T), & 0 \le n < L \\ x_0(n), & L \le n \le \frac{N}{2} \end{cases}, \tag{13}$$

$$g(n,T) = \frac{1+\cos\frac{2\pi}{N}n}{1+\cos\frac{2\pi}{N}(n+T)}, \quad (14)$$

$$T = \underset{t_L \le n \le t_H}{\operatorname{argmax}} \{x_0(n)\}, \quad (15)$$

where L denotes the number of replacement samples. $g(n,T)$ is a weighting factor to compensate the effect of the window function, where (14) is obtained from Hanning window [9]. T denotes the estimated start position of the second period of $x_0(n)$ in a range from t_L to t_H. Taking IDFT of $|\hat{S}(k)|e^{j\angle X(k)}$, we have the estimated speech signal in the time domain as $\hat{s}(n)$.

As pointed out in the literature [8], in the speech model shown in (7), we can get the exact start position of the second period if and only if an integer I satisfies

$$I = \frac{N}{k_c}. \quad (16)$$

III. Zero phase noise suppression using interpolation

When N/k_c is not an integer, interpolating the ZP signal can provide an accurate start position of the second period. It is useful to suppress noise more effectively. In this section, we introduce two types of interpolation methods, i.e., sinc interpolation and cubic spline interpolation.

A. Interval to be interpolated

We explain the interval to be interpolated with the Figs. 2(a) and (b). where we put $L = 3$. As shown in Fig.2(a), when T is less than the true start position T^* of the second period, i.e., $T < T^*$, candidate samples for replacement exist within $[T, T+L]$. As shown in Fig.2(b), when $T > T^*$, candidate samples for replacement exist within $[T-1, T+L-1]$. Since it can not be determined which $T > T^*$ or $T < T^*$, we use the interval $[T-1, T+L]$ to be interpolated.

B. Noise suppression using interpolated ZP signal

We express the ZP signal of the observed signal after the interpolation as

$$\tilde{x}_0(m) \quad m \in [r(T-1), r(T+L)], \quad (17)$$

where m denotes the integer index of the interpolated ZP signal, r denotes the number of samples existing between $x_0(n)$ and $x_0(n+1)$ and we call r as a ratio of interpolation. An example of interpolation of ZP signal is shown in Fig.3. The upper side shows an observed ZP signal, and the lower side shows the interpolated ZP signal when $L = 10$ and $r = 100$. These figures illustrate that 11 samples increases to 1100 samples by the interpolation.

After the interpolation procedure, we estimate $\hat{s}_0(n)$ from the interpolated ZP signal $\tilde{x}_0(m)$ as

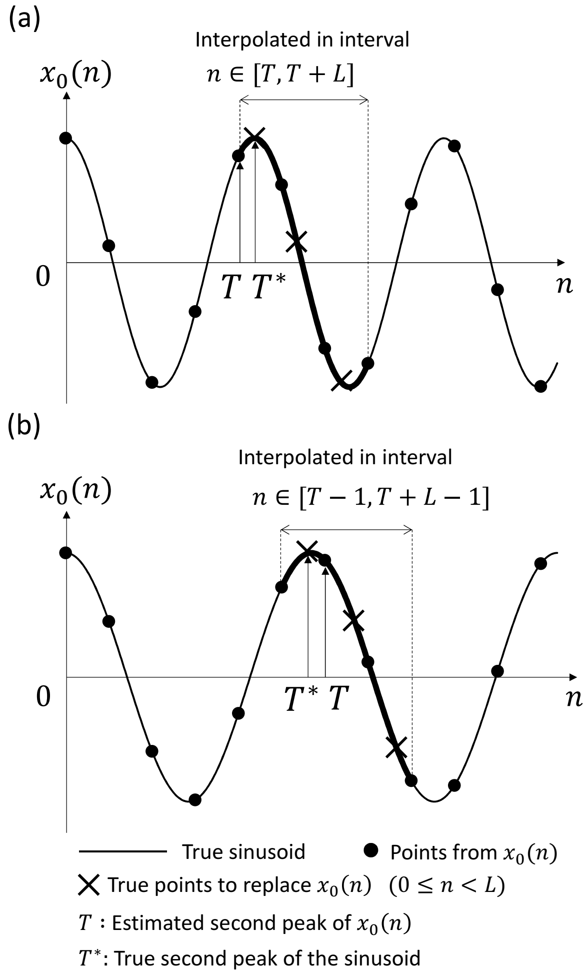

Fig. 2. Two cases of the interval for interpolation ($L = 3$). (a)$T < T^*$, (b)$T > T^*$.

$$\hat{s}_0(n) = \begin{cases} \tilde{g}(n,\tilde{T}) \cdot \tilde{x}_0(rn+\tilde{T}), & 0 \le n < L \\ x_0(n), & L \le n \le \frac{N}{2} \end{cases}, \quad (18)$$

$$\tilde{g}(n,\tilde{T}) = \frac{1+\cos\frac{2\pi}{N}n}{1+\cos\frac{2\pi}{N}(n+\frac{\tilde{T}}{r})}, \quad (19)$$

$$\tilde{T} = \underset{m \in [rt_L, rt_H] \cap [r(T-1), r(T+L)]}{\operatorname{argmax}} \{\tilde{x}_0(m)\}. \quad (20)$$

where, $\tilde{g}(n,\tilde{T})$ denotes a weighting factor corresponding to the interpolated ZP signal, and $\tilde{T}$ denotes the estimated start position of the second period. In the subsequent processing, the estimated speech signal $\hat{s}(n)$ is obtained in the same way described in section II-D.

There are some kinds of methods of interpolation. For example, linear interpolation is one of the simplest methods of interpolation. Unfortunately, applying to linear interpolation,

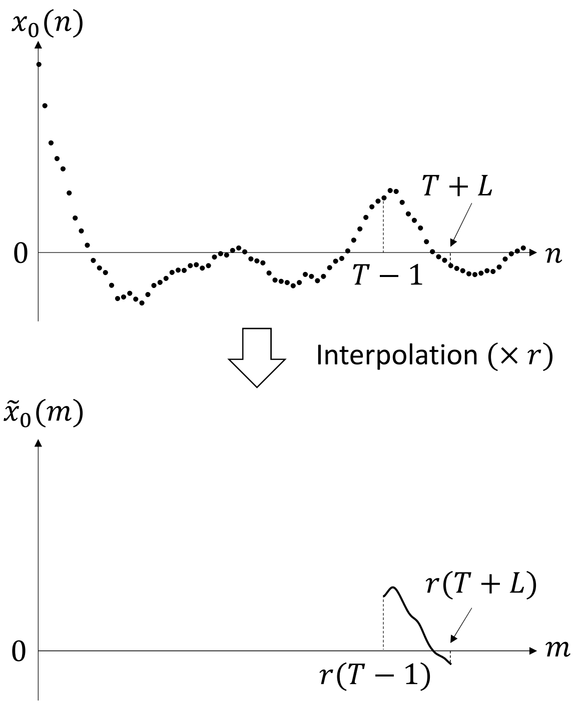

Fig. 3. Example of interpolation of $x_0(n)$, where $\tilde{x}_0(m)$ denotes interpolated signal $(L = 10)$.

we obtain $\tilde{T} = rT$. It implies that we find the same start position of the second period from $x_0(n)$ and $\tilde{x}_0(m)$. Then, we obtain the same $\hat{s}_0(n)$ of the conventional method. In the following sections, we will explain more effective methods of interpolation for the proposed method.

C. Sinc interpolation

Sinc interpolation is a well-known interpolation method to convert discrete signals into continuous signals [10]. The ZP signal $\tilde{x}_0(m)$ after sinc interpolation with r is given as [10]

$$\begin{aligned}\tilde{x}_0(m) &= \sum_{n=L}^{\lfloor m/r+N/2 \rfloor} x_0(n)\mathrm{sinc}\left(\frac{m}{r}-n\right) \\ &+ \sum_{n=\lfloor m/r+N/2 \rfloor+1}^{N-1} x_0(n)\mathrm{sinc}\left(\frac{m}{r}-n+N\right),\end{aligned} \tag{21}$$

where $\mathrm{sinc}(\cdot)$ denotes sinc function expressed as

$$\mathrm{sinc}(t) = \frac{\sin(\pi t)}{\pi t}. \tag{22}$$

Equation (21), shows that all $x_0(n)$ $(n = 0, 1, \ldots, N-1)$ are required to get one sample of $\tilde{x}_0(m)$.

D. Cubic spline interpolation

Addition to the sinc interpolation, we investigate about cubic spline interpolation which is an interpolation method with a small amount of computation load [11]. An interpolated ZP signal $\tilde{x}_0(m)$ with cubic spline interpolation is given as [11]

$$\tilde{x}_0(m) = \begin{cases} v_0(m), & m \in [r(T-1), rT-1] \\ v_1(m), & m \in [rT, r(T+1)-1] \\ \vdots & \\ v_l(m), & m \in [r(T+l-1), r(T+l)-1] \\ \vdots & \\ v_L(m), & m \in [r(T+L), r(T+L)-1] \end{cases}, \tag{23}$$

with

$$v_l(m) = a_l\left(\frac{m}{r}\right)^3 + b_l\left(\frac{m}{r}\right)^2 + c_l\left(\frac{m}{r}\right) + d_l \quad (l = 0, 1, \ldots, L). \tag{24}$$

Here, $a_l, b_l, c_l, d_l \quad (l = 0, \ldots, L)$ are given by

$$a_l = \frac{1}{6}(u_{l+1} - u_l), \tag{25}$$

$$b_l = \frac{1}{2}u_l, \tag{26}$$

$$c_l = x_0(T+l-1) - x_0(T+l) - \frac{1}{6}(u_{l+1} + 2u_l), \tag{27}$$

$$d_l = x_0(T+l-1), \tag{28}$$

where $u_0, \ldots, u_L$ are the second derivatives of $v_0(m), v_1(m), \ldots, v_L(m)$ at $m = T-1, T, \ldots, T+L$, respectively [11].

E. Estimation accuracy for single sinusoid

As a simple example, we add single impulsive noise $d_0(n)$ to a single sinusoid $s_0(n)$ of period N/k_c. The ZP signal of the observed signal is expressed as

$$x_0(n) = \cos\left(\frac{2\pi k_c n}{N}\right) + \delta(n). \tag{29}$$

The estimation accuracy of the start position of the second period is evaluated by using the relative error given as

$$E_{r0} = \frac{|\hat{s}_0(0) - s_0(0)|}{|s_0(0)|} \times 100 \quad (\%). \tag{30}$$

We evaluate E_{r0} for various periods N/k_c. We put $L = 1, N = 512$, and $r = 100$. Figure 4 plots E_{r0}, where the horizontal axis shows k_c from 3 to $N/2$. The dashed line shows the conventional ZP method [8], the solid line shows the proposed method with sinc interpolation and the dotted line shows the proposed method with cubic spline interpolation. When (16) is satisfied, i.e., $k_c = 4, 8, 16, \ldots$, we have $E_{r0} = 0$ because the true start position of the second period exists on $x_0(n)$. When k_c is small, E_{r0} becomes small. Because the ZP values around the true start position in the second period are almost the same to the true value due to its long period, i.e. the interpolation effect may be negligible. On the other hand, as k_c increases, E_{r0} increases. In this case, the interpolation effectively works.

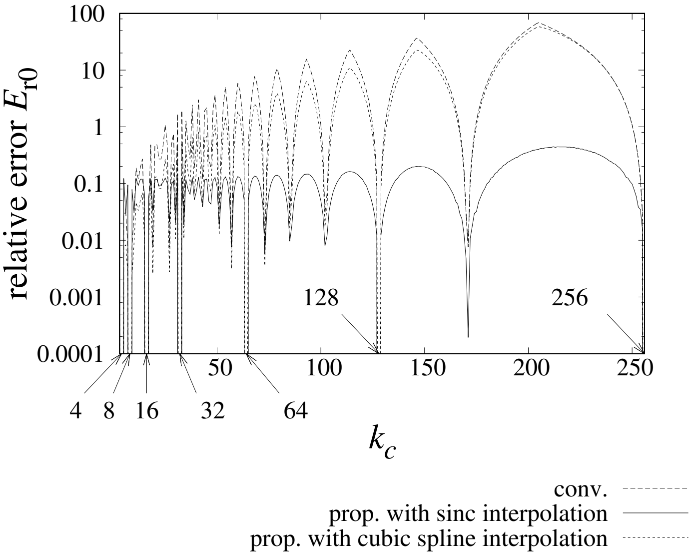

Fig. 4. Estimation error E_{r0} for single sinusoid plus unit impulse function

In the proposed method with sinc interpolation, E_{r0} is less than 0.45% at each k_c. On the other hand, the proposed method with cubic spline interpolation decreases the estimation accuracy in comparison to sinc interpolation, althogh it gives E_{r0} less than the conventional ZP method.

F. Speech reconstruction accuracy by replacement

When the ZP signal is periodic and does not include noise, the replacement method should not change the original ZP signal. To confirm this capability, we applied the conventional ZP method [8] and the proposed method (19)-(20) to the voiced speech portion taken from the newspaper article reading speech corpus [15]. Here, we cut out from 6 to 16 frames of voiced speech from a speech signal. We used speech signals of 100 people, 50 male and 50 female. The sampling frequency was 16 kHz, and $N = 512$, Hanning window was used for signal segmentation. Figure 5 shows an example of the male voiced speech signal used for this simulation. We see from Fig. 5 that the signal has periodicity, and gradually changes. We put $t_L = 53, t_H = 160$ that covers the averaged male and female pitch periods [14]. To evaluate the reconstruction capability for $s(n)$ we use $\mathrm{SNR_{out}}$ expressed by

$$\mathrm{SNR_{out}} \quad = \quad 10\log_{10}\frac{\sum_{n=0}^{M-1}s^2(n)}{\sum_{n=0}^{M-1}\{\hat{s}(n)-s(n)\}^2}, \qquad (31)$$

where M denotes the number of samples of the entire speech signal. The number of samples used for replacement was $L = 1, \ldots, 64$. Figures 6(a) and (b) show the reconstruction results to male and female voiced speech, respectively. In each figure, the dashed line, the solid line, and the dotted line show the conventional ZP method, the proposed methods with sinc interpolation, and with cubic spline interpolation, respectively. We see from Figs.6(a) and (b) that the proposed methods made the reconstruction accuracy of voiced speech higher

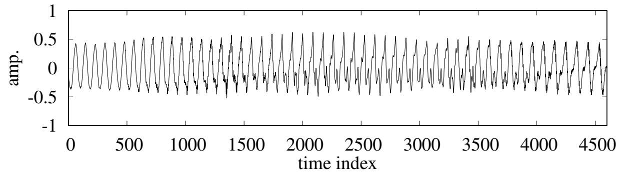

Fig. 5. Voiced sound part of speech signal (male voice).

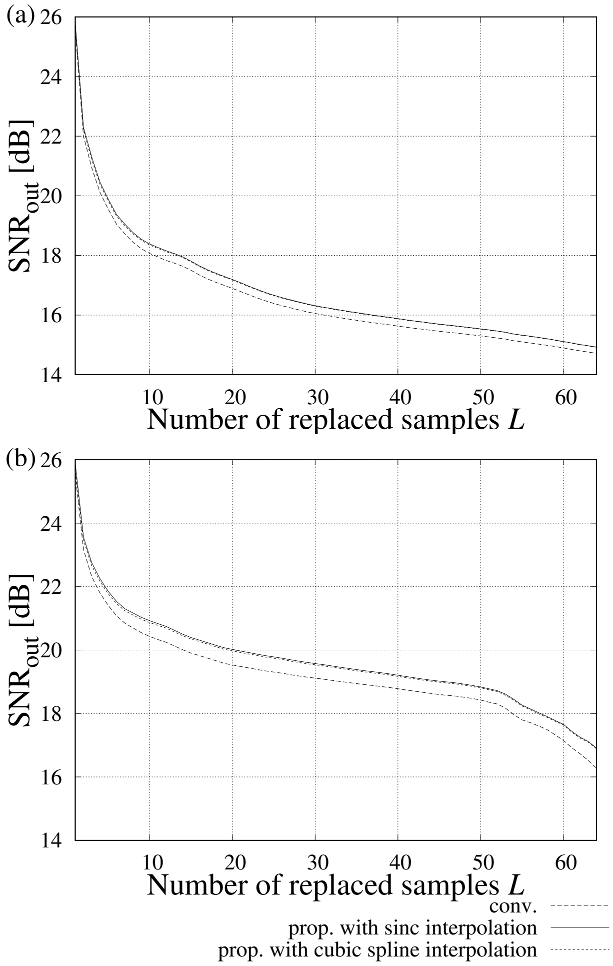

Fig. 6. Comparison of voiced sound replacement accuracy. (a)Results of application to male voiced speech. (b)Results of application result to female voiced speech.

than the conventional one. When L increases, the amount of improvement by interpolation also increases. This effect is more obvious in the results for the female voiced speech compared with the male voiced speech. We improved 0.63 dB for sinc interpolation and 0.61 dB for cubic spline interpolation at $L = 64$.

IV. Noise suppression result

To evaluate the impulsive noise suppression capability of the proposed method, we performed computer simulations. The sampling frequency was 16 kHz, Hanning window was used for signal segmentation with $N = 512$. We put $t_L = 53, t_H = 160$. As desired speech signals, we used 50 male and 50 female speech signals taken from the newspaper article reading speech corpus [15]. The signals are the shortest 3.1 seconds (49080 samples), the longest 8.4 seconds (135168 samples), and the average 5.2 seconds (83398 samples).

A. Impulsive noise suppression

Impulsive noise whose duration is 1 sample was added to a speech signal. Here, the impulsive noise was constant for magnitude, and added every 2000 samples with inverted sign so that $\mathrm{SNR_{in}} = 10\mathrm{dB}$. $\mathrm{SNR_{in}}$ is given as

$$\mathrm{SNR_{in}} = 10 \log_{10} \frac{\sum_{n=0}^{M-1} s^2(n)}{\sum_{n=0}^{M-1} d^2(n)}. \tag{32}$$

We put $L = 1$ as the number of replacement samples suitable for suppressing impulsive noise. An example of the noise suppression results is shown in Figs.7(a)-(e). Here, (a) is the original speech, (b) is the observed signal, (c) shows the output signal without interpolation, (d) and (e) show the output signals of applying the proposed method with sinc interpolation and cubic spline interpolation, respectively. Comparing the waveforms shows that the impulsive noise is suppressed by the conventional and the proposed methods. The difference between (c), (d), and (e) is not large enough to be visible in the waveform.

Next, we evaluate the $\mathrm{SNR_{out}}$ of the proposed method when $\mathrm{SNR_{in}}$ is changed from 0 to 20 dB. The results are shown in Fig.8, where the solid line and the dotted line represent the proposed method with sinc interpolation and cubic spline interpolation, respectively. We see from Fig.8 that $\mathrm{SNR_{out}}$ of the proposed methods seems to be almost the same.

In addition, Fig.9 shows improvement of $\mathrm{SNR_{out}}$ that is the difference between the proposed methods and the conventional method. Here, the solid line and the dotted line represent the proposed method with sinc interpolation and cubic spline interpolation, respectively. The maximum difference of 0.09 dB for sinc interpolation and 0.06 dB for cubic spline interpolation. Although the difference is small, the proposed method exactly improved the impulsive noise suppression capability in comparison to the conventional method for all tested signals.

In this simulation, the fundamental frequency of the voiced speech belongs to the low frequency band was less than 1000 Hz which is corresponding to $k_c = 32$. As shown in Fig.4, the proposed method may be more effective for a higher fundamental frequency sound.

B. Computational load and processing time

We discuss about the computational load increased by interpolation. The number of calculations per frame is shown in TABLE I. While sinc interpolation requires additional 460800 operations, cubic spline interpolation requires 4018 operations. Compared to the sinc interpolation, amount of operations of cubic spline interpolation is only 0.87%. The processing time when using the conventional method and the proposed method

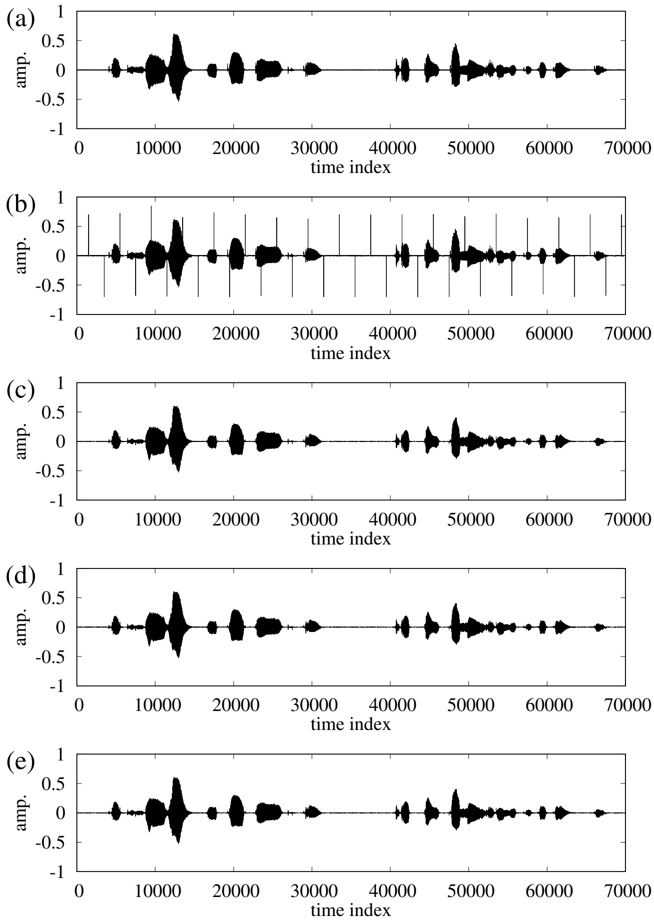

Fig. 7. Impulsive noise suppression result: (a) original speech, (b) signal observed, (c) output by conventional method, (d) output by proposed method with sinc interpolation, (e) output by proposed method with cubic spline interpolation.

in a computer with a 4-core 3.4 GHz CPU is shown in TABLE II, where we displayed on averaged processing time for 50 male and 50 female speech signals whose average length is 5.2 seconds. We see from TABLE I, sinc interpolation takes about 10 times as long as the conventional method, cubic spline interpolation is slightly increased the processing time compared with the conventional method. We can expect that the proposed method with cubic spline interpolation is reasonable to use the impulsive noise suppression.

C. Suppression of real recorded sound

We simulated to suppress mouse clicking sound, which are considered as impulsive noises existing in the real environment. We used two kinds of mouse clicking sound (down and up) sampled at 44.1 kHz taken from Freesound.org [16] and downsampled to 16 kHz as the impulse noise mouse click sound. The ZP signal of the mouse clicking sound is shown in Fig. 10. We set $L = 20$ [9], [12]. We added the mouse clicking sounds where alternately added to the speech signal so that $\mathrm{SNR_{in}} = 5$ dB. An example of the mouse clicking noise suppression results is shown in Figs.11(a)-(d). Here, (a)

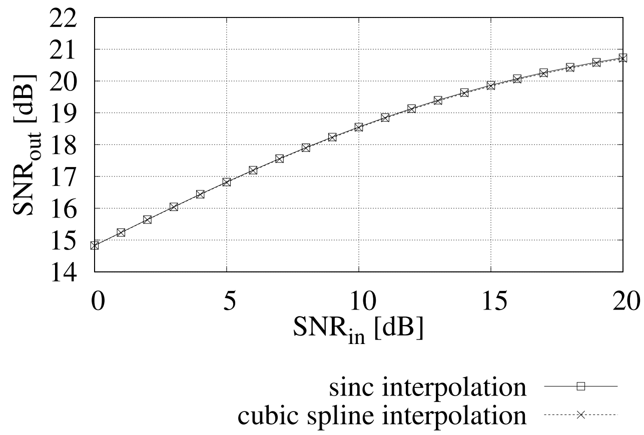

Fig. 8. Output SNR of proposed method.

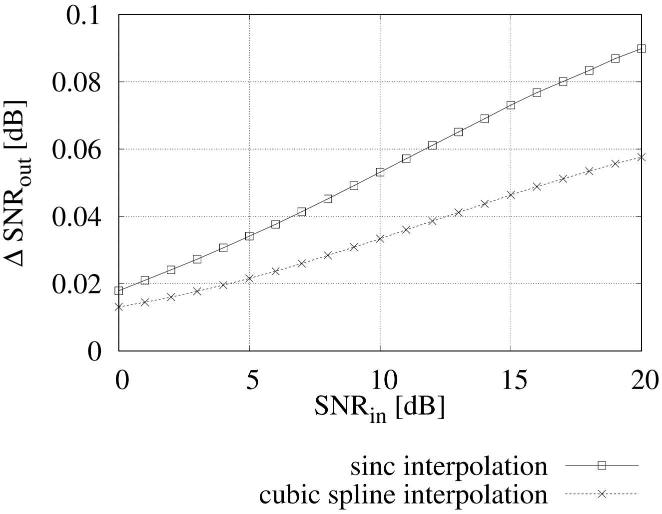

Fig. 9. Comparison between proposed method and conventional method.

TABLE I
THE CALCULATION AMOUNT PER FRAME

	conv.	prop. sinc	prop. cubic spline
addition(+), subtraction(-)	0	153600	2110
multiplication (*)	0	153600	1204
division (/)	0	153600	704
total	0	460800	4018

TABLE II
THE PROCESSING TIME FOR THE IMPULSIVE NOISE SUPPRESSION

	conv.	prop. sinc	prop. cubic spline
processing time [ms]	61.1	634.8	63.9

is the observed signal, (b) shows the output signal without interpolation, (c) and (d) show the output signal of applying the proposed method with sinc interpolation and cubic spline interpolation, respectively. Comparing the waveforms shows that the clicking noise is suppressed by conventional and

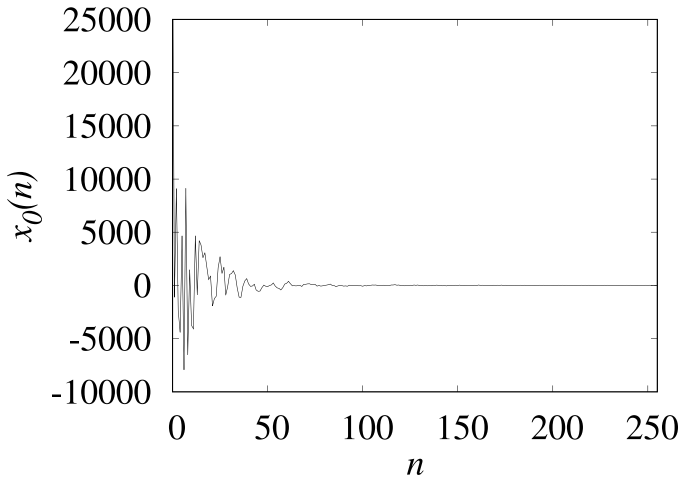

Fig. 10. $x_0(n)$ of mouse clicking sound

proposed methods. Similar to the result of impulse noise suppression, difference between (b), (c), and (d) is not large enough to be visible in the waveform.

Next, we evaluate the $\mathrm{SNR_{out}}$ of the proposed method when $\mathrm{SNR_{in}}$ is changed from -5 to 15 dB. In the condition of $\mathrm{SNR_{out}} = -5$, the noise sound is very annoying to hear. The results are shown in Fig.12, where the solid line and the dotted line represent the proposed method with sinc interpolation and cubic spline interpolation, respectively. $\mathrm{SNR_{out}}$ of the proposed methods seems to be almost the same. Here, the improvement is less than the artificial impulsive noise suppression results shown in Fig. 8.

To more accurately show the improvement of $\mathrm{SNR_{out}}$, the difference between the proposed methods and the conventional method is shown in Fig. 13. Here, the solid line and the dotted line represent the proposed method with sinc interpolation and cubic spline interpolation, respectively. The maximum difference of 0.15 dB for sinc interpolation and 0.17 dB for cubic spline interpolation. These improvements are higher than the results shown in Fig. 9. It implies that the proposed method is more effective to suppress the noise when L is much larger than 1 in comparison to the conventional method.

V. CONCLUSION

In this paper, we proposed an impulsive noise suppression method using interpolated ZP signal. The proposed method utilizes sinc or cubic spline interpolation. The interpolated ZP signal improves an estimation accuracy of the start position of the second period of the ZP signal, and noise reduction capability. Simulation results show that the proposed method achieved the output SNR of around 18.5 dB when the input SNR is 10 dB, where the improvement is about 0.03-0.05 dB in comparison to the conventional method. The improvement is slight, but exactly obtained. Finally, we carried out the simulation to suppress mouse clicking sound as a real environmental

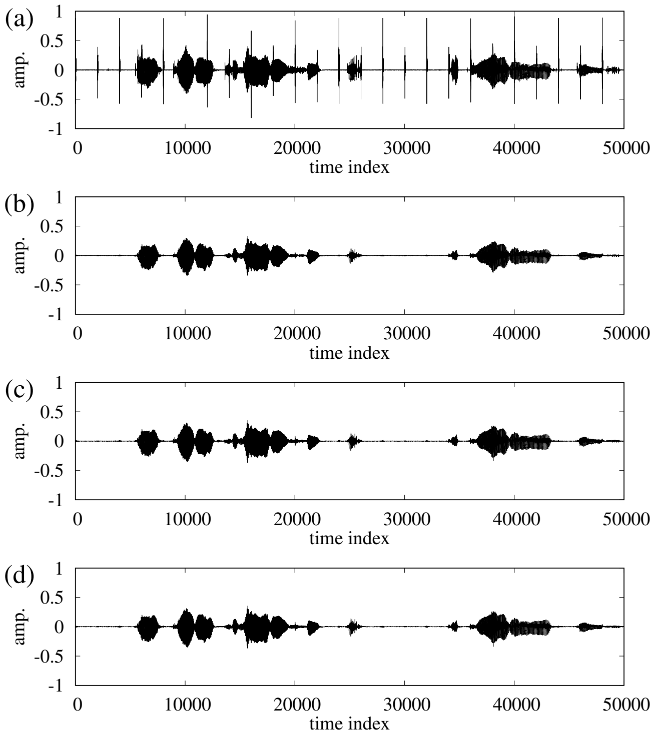

Fig. 11. Mouse clicking sound suppression result: (a) signal observed, (b) output by conventional method, (c) output by proposed method with sinc interpolation, (d) output by proposed method with cubic spline interpolation.

impulsive noise. When input SNR is 15 dB, the improvement was about 0.15-0.17 dB in comparison to the conventional method.

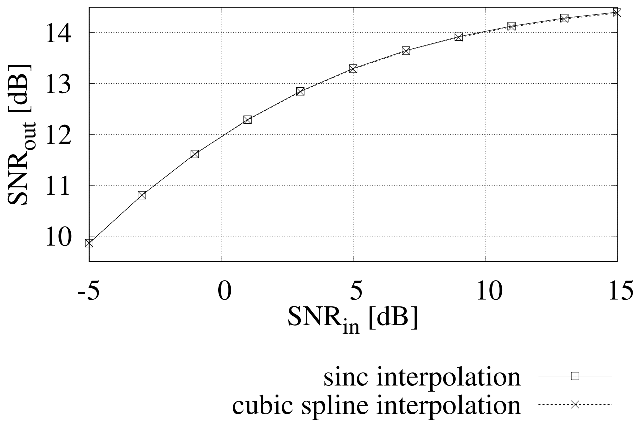

Fig. 12. Output SNR of proposed method.

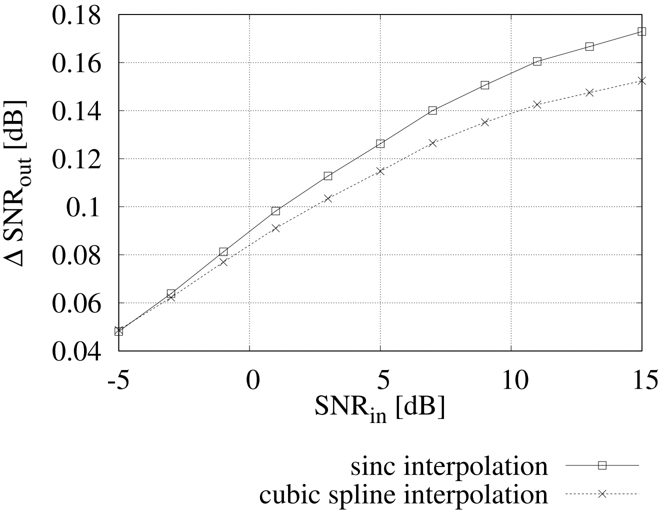

Fig. 13. Comparison between proposed method and conventional method.

REFERENCES

[1] S. F. Boll, "Suppression of acoustic noise in speech using spectral subtraction," *IEEE Trans. Acoust. Speech Signal Process.*, vol.ASSP-27, no.2, pp.113-120, April 1979.

[2] N. Wiener, The Extrapolation, Interpolation and Smoothing of Stationary Time Series, John Wiley & Sons, Inc., New York, N.Y., 1949.

[3] Y. Ephraim and D. Malah, "Speech enhancement using a minimum mean-square error short-time spectral amplitude estimator", *IEEE Trans. Acoust. Speech Signal Process.*, vol.ASSP-32, no.6, pp.1109-1121, Dec. 1984.

[4] P. J. Wolfe and S. J. Godsill, "Efficient alternatives to the Ephraim and Malah suppression rule for audio signal enhancement," *EURASIP Journal on Applied Signal Processing*, vol.10, pp.1043-1051, Oct. 2003.

[5] T. Lotter and P. Vary, "Speech enhancement by MAP spectral amplitude estimation using a super-Gaussian speech model," *EURASIP J. Adv. Signal Process.*, vol.2005, no.7, pp.1110-1126, July 2005.

[6] N. Miyake, T. Takiguchi and Y. Ariki, "Sudden noise reduction based on GMM with noise power estimation," *Proc. Interspeech 2008*, pp.403-406, April 2008.

[7] Z. Duan, G. J. Mysore, and P. Smaragdis, "Speech enhancement by online non-negative spectrogram decomposition in non-stationary noise environments," *Proc. Interspeech 2012*, pp.595-598, Sep. 2012.

[8] Y. Kamamori, A. Kawamura, Y. Iiguni, "Zero Phase Signal Analysis and Its Application to Noise Reduction" *IEICE Trans. Fundamentals (Japanese Edition)*, Vol. J93-A, No. 10, pp.658-666, Oct. 2010.

[9] W. Thanhikam, Y. Kamamori, A. Kawamura, and Y. Iiguni, "Stationary and Non-stationary Wide-Band Noise Reduction Using Zero Phase Signal," *IEICE Trans. Fundamentals*, Vol. E95-A, No. 5, pp.843-852, May. 2012.

[10] C. E. Shannon. "Communication in the presence of noise," *Proceedings of the IRE* 37.1, pp.10-21, Jan. 1949.

[11] C. de Boor, A Practical Guide to Splines, New York: Springer-Verlag, 1978.

[12] S. Kohmura, A. Kawamura, and Y. Iiguni, "A Zero Phase Noise Reduction Method with Damped Oscillation Estimator," *IEICE Trans. Fundamentals*, Vol.E97-A, No.10, pp.2033-2042, Oct. 2014.

[13] A. Tanaka, Y. Amma, A. Kawamura, "Noise Suppression by Modified Zero Phase Signal Analysis" *IEICE Technical Report*, Vol. SIP2015-141, pp.153-157, Mar. 2016.

[14] S. Furui, Digital speech processing, Tokai University Press, Tokyo, 1985.

[15] "ASJ Japanese Newspaper Article Sentences Read Speech Corpus (JNAS)," Speech Resources Consortium, http://research.nii.ac.jp/src/en/JNAS.html

[16] "mouse click.wav," Freesound.org, https://www.freesound.org/people/THE_bizniss/sounds/39562/

[17] R. Takehara, A. Kawamura, Y. Iiguni, "A Nonlinear Spectral Subtraction Using Zero Phase Signal" *Proceedings of the 2016 IEICE General Conference*, A-8-10, p.121, Mar. 2016.

[18] R. Takehara, A. Kawamura, Y. Iiguni, "Zero Phase Noise Suppression Using sinc-Interpolation" *Proceedings of the 31th IEICE SIP Symposium*, pp.305-310, Nov. 2016.