

Restoration of Clipped Audio Signal Using Recursive Vector Projection

Shin Miura^{1,2}, Hirofumi Nakajima³, Shigeki Miyabe¹, Shoji Makino¹, Takeshi Yamada¹, Kazuhiro Nakadai²

¹ Graduate School of Systems and Information Engineering, University of Tsukuba

1-1-1 Tennoudai, Tsukuba, Ibaraki, 305-8573, Japan

Email: miura@mmlab.cs.tsukuba.ac.jp

² Honda Research Institute Japan Co., Ltd., Japan

³ Department of Computer Science, Kogakuin University, Japan

Abstract—This paper proposes signal restoration from clipping effect without prior knowledge. First, an interval of signal including clipped samples is analyzed by recursive vector projection. By analyzing the neighboring samples of the clipped interval and excluding the clipped interval in the analysis of similarity, signal estimation in the clipped interval is estimated as a by-product of the analysis. Since the estimation holds consistency with the neighboring samples, the restored signal does not suffer from click noise. Evaluation of the clipping restoration with various audio signal ascertained that the proposed method improves signal-to-noise ratio.

I. INTRODUCTION

To record sounds precisely, the input gain level should be high to reduce noise such as electric hum, quantization noise etc. However, the high input gain often causes an over-level distortion in recorded signal. In the instant when the amplitude of the signal exceeds the limit of the recording devices, the amplitude of the recorded signal is cut off to the highest acceptable level and the signal shape is distorted to be flat as shown in Fig. 1. This phenomenon is called clipping effect, and the distortion is harmful in various applications, such as speech recognition, source separation, direction estimation, and so on.

Since the waveform of the clipped signal is deteriorated severely seriously from the original, perfect restoration is difficult. However, it is possible to estimate the original amplitude partially by using the prior knowledge on the target signal. If the signal is band-limited, the property of oversampling can be utilized for the restoration [1]. Several works on the restoration specified to speech are reported [2], [3]. Statistical assumption on signal can also be used [4].

Clipping restoration for generic signal without prior knowledge on the target signal is much harder problem, and there have been few works on this problem. Most of the conventional works, e.g., [5]–[7], are based on signal substitution, which suffers from clicking noise.

In this paper we propose clipping restoration based on recursive vector projection (RVP) [8], which is a signal approximation by selecting prepared candidates of basis vectors. By fitting the bases to the neighboring samples of the clipped ones by RVP, signal restoration of the clipped samples is obtained as a by-product of the fitting. Since the restoration is

estimated to be consistent with the neighboring samples, the restored signal is free from the clicking noise.

II. THE CONVENTIONAL METHOD

In this section we review the conventional restoration methods which do not require prior knowledge of the target audio signal.

A. The approximate waveform substitution method

The approximation waveform substitution restores the clipped signal by double sides periodic substitution (DSPS) method [5], [6], which is originally developed for the restoration of packet loss. First, an interval of signal is detected by measuring similarity to the both preceding and proceeding samples of the clipped samples. Second, the estimation of the lost samples is synthesized as the scaled and biased version of the signal in the detected interval. After that, the clipped samples and the neighbors are substituted by the estimation with smoothing by cross fade.

B. The approximate spectrum substitution method

The approximate spectrum substitution method is frequency domain version of the approximate waveform substitution method. First, the clipped signal is transformed into the frequency domain. Second, the frame similar to the clipped frame is detected measuring cosine distance. Finally, the clipped frame is substituted by the amplitude modification of the detected frame. By transforming the substituted frames into the domain, signal restoration is obtained.

C. The method based on linear prediction

The predictive value of the true amplitude of clipped part is obtained by the linear prediction with the last several samples of clipped part. If two or more successive samples are clipped, the prediction of each sample is obtained recursively using the previous prediction.

D. The problems of the conventional methods

The problems of the above discussed conventional methods could be summarize to computational complexing of the signal restorations and to introduced discontinuity at the clipped signal edges, thus those methods applicability is low.

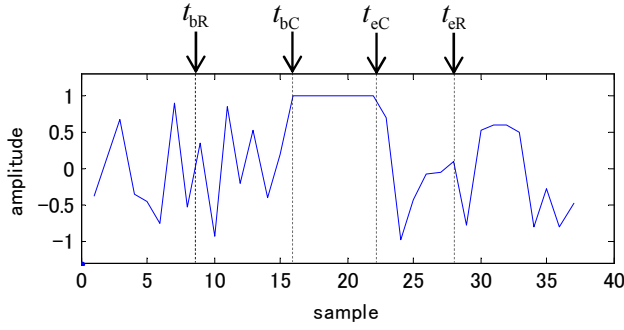


Fig. 1. An example of clipped signal

III. THE PROPOSED METHOD

A. The method based on RVP

We propose a clipping restoration method based on the recursive vector projection (RVP). We formulate the problem to approximate a vector as linear sum of priory prepared bases.

Here, suppose we approximate an N -dimensional vector \mathbf{y} using RVP. We prepare M candidates of bases $\mathbf{a}_k (k = 1, 2, \dots, M)$. Each $\mathbf{x}(k)$ are coefficients to approximate vectors. The number of vectors which can be selected is $L < M$, and \mathbf{y} is described as

$$\mathbf{y} = x_1 \mathbf{a}_1 + x_2 \mathbf{a}_2 \dots + x_M \mathbf{a}_M + \mathbf{e} = \mathbf{A} \mathbf{x} + \mathbf{e}, \quad (1)$$

where \mathbf{x} is M -dimensional vector having $x(k)$ as the basis, \mathbf{A} is $N \times M$ matrix whose columns are \mathbf{a}_k . \mathbf{e} is error of approximation whose Euclidean norm $\|\mathbf{e}\|$ is minimized during the iteration. and, \mathbf{x} and \mathbf{A} are described as

$$\mathbf{x} = [x_1 x_2 \dots x_M]^T, \\ \mathbf{A} = [\mathbf{a}_1 \mathbf{a}_2 \dots \mathbf{a}_M].$$

Also, RVP constrains the coefficient vector \mathbf{x} to have less than or equal to L nonzero elements. RVP optimizes the coefficients \mathbf{x} in the following recursive algorithm, and the flowchart of procedure is shown Fig. 2

[step 1] Initialization:

Parameters are initialized as

$$u \leftarrow 1, \mathbf{x} \leftarrow \mathbf{0}, \mathbf{e} \leftarrow \mathbf{y}, \quad (2)$$

where u is iteration count of the recursive optimization, \mathbf{x} is coefficients, and \mathbf{e} is error vector.

[step 2] Projection of most similar basis:

Projections \mathbf{p}_k of the error \mathbf{e} onto the basis candidate \mathbf{a}_k are calculated as

$$\mathbf{p}_k \leftarrow \frac{(\mathbf{e}, \mathbf{a}_k)}{(\mathbf{a}_k, \mathbf{a}_k)} \mathbf{a}_k \quad (k = 1, 2, \dots, n), \quad (3)$$

and the index m of the coefficient to be updated is determined as

$$m \leftarrow \arg \max_k \|\mathbf{p}_k\| \quad (k = 1, 2, \dots, n), \quad (4)$$

where, $\arg \max_k [\cdot]$ shows the argument k which is the maximum element of $[\cdot]$. Equation (3) expresses obtaining \mathbf{p}_k by

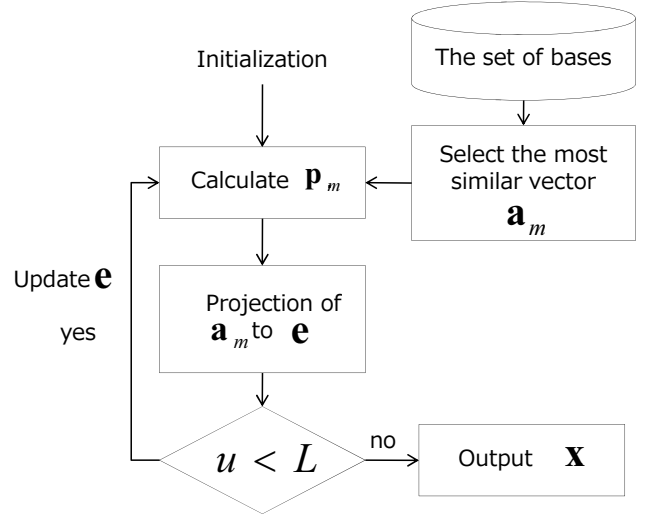


Fig. 2. Flowchart of RVP

projecting \mathbf{e} to \mathbf{a}_k . Equation (4) corresponds to select the column vector of \mathbf{A} which has the nearest angle to error vector.

$$x_m \leftarrow \frac{(\mathbf{e}, \mathbf{a}_m)}{(\mathbf{a}_m, \mathbf{a}_m)}, \quad (5)$$

Equation (6) expresses updating the solution and error vector.

[step 3] Updating solution and iteration:

If the iteration count u reaches the limit of the number of the basis L , the recursive update is terminated. Otherwise the error vector is updated as

$$\mathbf{e} \leftarrow \mathbf{e} - \mathbf{p}_m, \quad (6)$$

the iteration count is incremented as

$$u \leftarrow u + 1, \quad (7)$$

and the procedure goes back to Step 2.

B. Estimation of clipped samples using RVP

In this section, we propose estimation of lost samples based on consistency with the clean signal in the neighboring samples without clipping effect. We apply RVP to signal of an interval including samples with and without clipping effect. In this RVP analysis, we estimate the basis and the coefficients by fitting to the clean samples without clipping. By reconstructing the signal with the basis and coefficients, a waveform appears even in the samples which is not took into account in the fitting. As a result, the obtained signal in the clipped samples of the recover is the estimation consistent with the non-clipped samples. By repeating this procedure with all the clipped interval, entire the length of the signal is restored.

First, we determine the interval of the signal to be analyzed by RVP. Note that we assume clipping can be detected easily by checking the amplitude. Suppose the clipping effect is

found in the interval from t_{bC} -th to t_{eC} -th samples. The beginning sample t_{bR} of the interval to be analyzed is determined as α samples before t_{bC} as

$$t_{bR} = t_{bC} - \alpha. \quad (8)$$

Note that if previous clipped interval is included in the analyzed interval, those interval should be substituted the previous restoration. Similarly, the end sample t_{eR} of the interval to be analyzed is determined as β samples after t_{bC} as

$$t_{eR} = t_{eC} + \beta. \quad (9)$$

The value β does not exceed β' so that $t_{eR} = t_{eC} + \beta$ is kept to be before the next clipped sample. Now, we set the vector \mathbf{y} to be the signal in the interval to be analyzed with the amplitudes of the clipped samples set to be zero as

$$\mathbf{y} = [s(t_{bR}), \dots, s(t_{bC} - 1), 0, \dots, 0, s(t_{eC} + 1), \dots, s(t_{eR})], \quad (10)$$

where s is the target signal, $s(t)$ is the t -th element of s . Next, prepare bases \mathbf{A} contains complex sinusoidal waves of various frequency. Its row vector \mathbf{a}_k is described below

$$\mathbf{a}_k = \exp \left[\frac{2\pi j f t}{N} \right], \quad (11)$$

$$f = Fk, \quad (12)$$

where, F is parameter of frequency step range. Furthermore, to reconstruct the \mathbf{y} 's amplitude to true amplitude, \mathbf{a}_k is needed to amplify as described below

$$\mathbf{h} = b - \cos[2\pi t/N] \quad (13)$$

$$\mathbf{a}_k \leftarrow \hat{\mathbf{a}}_k * \mathbf{h} \quad (14)$$

where, $*$ is Hadamard product, and b (≥ 2) is a parameter to adjust the shape of the window function and set experimentally. By analysis of \mathbf{y} using RVP, the restoration of the clipped interval is estimated as a by-product of fitting to the non-clipped intervals. With the amplitudes of the clipped interval substituted by 0, the fitting does not take the clipped interval into account because the inner product of the error e and each basis \mathbf{a}_k are effected by the only the non-clipped intervals as

$$\begin{aligned} (\mathbf{a}_k, e) &= \sum_{t=t_{bR}}^{t_{eR}} \mathbf{a}_k(t)e(t) \\ &= \sum_{t=t_{bR}}^{t_{bC}-1} \mathbf{a}_k(t)e(t) + \sum_{t=t_{bC}}^{t_{eC}} 0e(t) \\ &\quad + \sum_{t=t_{eC}+1}^{t_{eR}} \mathbf{a}_k(t)e(t) \\ &= \sum_{t=t_{bR}}^{t_{bC}-1} \mathbf{a}_k(t)e(t) + \sum_{t=t_{eC}+1}^{t_{eR}} \mathbf{a}_k(t)e(t), \quad (15) \end{aligned}$$

where $\mathbf{a}_k(t)$ or $e(t)$ is the t -th element of each vector.

IV. EXPERIMENT OF RESTORATION

We conducted experiments to evaluate the performance of the proposal method. We prepared various sounds that are clipped deliberately. Evaluation is done by comparing the process and unprocessed signals.

A. Data used in evaluation

We used various sound data (voice, music, etc...). The detail of data are shown in Table I. The music data are from RWC Music Database [9].

TABLE I
SOUNDS FOR EVALUATION

name	details
male speech	sentences a few seconds
female speech	sentences a few seconds
driving sound	recorded by microphone put on car registration plate
heavy metal music	a part of "Lost in My Dreams"
jazz music	a part of "Jive(piano solo)"
classical music	Haydn Sym. No94 "Surprise" 1st mov.

B. Experimental condition

We normalized the sounds shown in previous chapter, and clip 4-level (0.05, 0.10, 0.15, 0.20). The parameter setting of RVP is listed in Table II. on experimental ground. The parameters are shown at Table 2.

TABLE II
PARAMETER SETTING OF PROPOSED METHOD

step of frequency F'	0.2 Hz
sample number of α	50
sample number of β'	30
number of basis vector L	same as clipping length

C. evaluation criteria

We conducted an objective evaluation SNR and subjective evaluation. The evaluation criteria are as follows.

1) *SNR*: The higher SNR the better sound quality. It is given by the following equation.

$$\text{SNR [dB]} = 10 \log \frac{\sum (x(t))^2}{\sum (x(t) - \hat{x}(t))^2}, \quad (16)$$

where, $x(t)$ is the reference (original) signal, and $\hat{x}(t)$ is the signal to be evaluated.

2) *Subjective assessment*: The subjective assessment is held with 4 listeners, but no noticeable difference was found.

D. Result of experiment and consideration

Restored the clipping with each methods introduced at section 3. The result of SNR is shown Table III. Figs. 3–4 show examples of spectrogram and a waveform that are restored with proposed method. In the conventional method, SNR score is degraded. The reasons needs to be further analyzed, but we suspect that the approximate waveform substitution

TABLE III
IMPROVEMENT OF SNR SCORE

method	SNR
approximate waveform substitution	-1.95
approximate spectrum substitution	-0.28
linear prediction	-9.91
proposed	+2.00

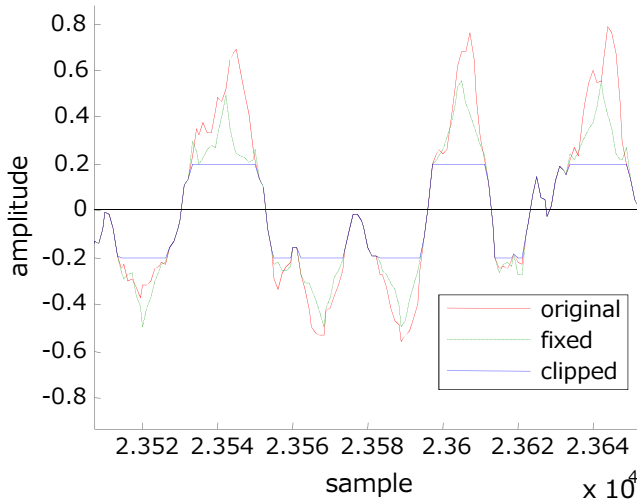


Fig. 3. An example of waveform restored with proposed method

method destroys phase continuity in the vicinity of restored interval, and the linear prediction method accumulates the errors when the clipped interval is too long. On the other hand, the proposed method improved SNR without annoying artifact such as clicking noise.

V. CONCLUSIONS

We first proposed a restoration algorithm based on RVP for generic audio signal, which does not use prior knowledge of the target signal. RVP analysis without clipped samples produces the restoration waveform in the clipped interval. Second, we conducted experiments to evaluate the performance of the proposed method. Although the difference was not noticeable the proposed method improved SNR. In future research our approach is to improve the restoration performance by considering the criteria which is effective for evaluating clipping restoration.

REFERENCES

- [1] J. S. Abel and J. O. Smith, "Restoring a clipped signal," in Proc. ICASSP, Toronto, Canada, 1991.
- [2] A. Janssen, R. Veldhuis, and L. Vries, "Adaptive interpolation of discrete-time signals that can be modeled as autoregressive processes," IEEE Trans. Acoustics, Speech and Sig. Proc., vol. 34, no. 2, pp. 317-330, 1986.

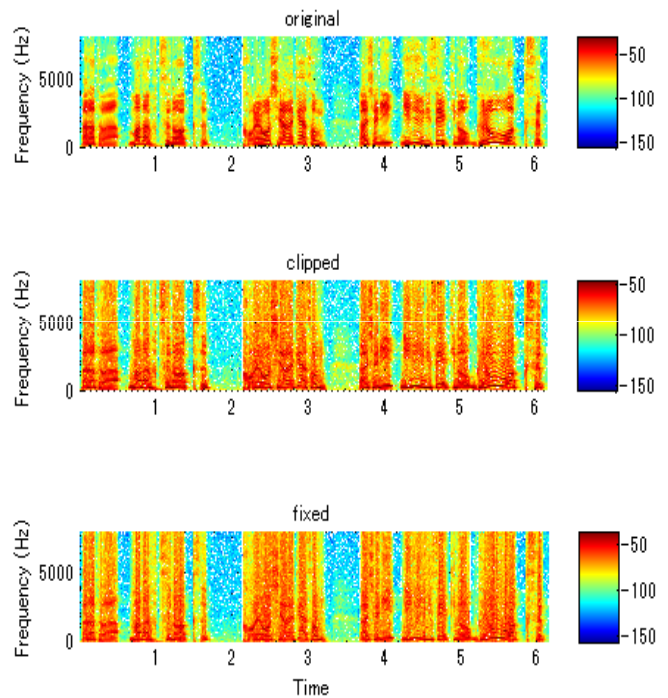


Fig. 4. An example of spectrogram restored with proposed method

- [3] A. Dahimene, M. Noureddine, and A. Azrar, "A simple algorithm for the restoration of clipped speech signal," Informatica, vol. 32, pp. 183-188, 2008.
- [4] S. J. Godsill, P. J. Wolfe, and W. N. W. Fong, "Statistical model-based approaches to audio restoration and analysis," Journal of New Music Research, vol. 30, no. 4, 2001.
- [5] J. Tang and F. Itakura, "Double sides periodic substitution (DSPA) method for recovering missing speech," Proc. ISSPA, pp. 544-549, 1987.
- [6] J. Tang and F. Itakura, "A new method for recovering missing speech," Proc. ICCT pp. 14-387, 1987.
- [7] J. Deller, J. Hansen and J. Proakis, *Discrete-time processing of speech signals*, Macmillan Pub. Co., New York, 1993.
- [8] H. Nakajima, M. Tohyama, M. Tanaka, "A recursive vector projection method and its applications," IEICE Technical Report, vol. J83-A no. 4, pp. 353-360, 2000.
- [9] M. Goto, H. Hashiguchi, T. Nishimura, and R. Oka, "RWC Music Database: popular, classical, and jazz Music Databases, Proc. ISMIR, pp. 287-288, 2002.