SNR ESTIMATION FOR CLIPPED AUDIO BASED ON AMPLITUDE DISTRIBUTION

Xiaoqing Liu, Jia Jia and Lianhong Cai

Key Laboratory of Pervasive Computing, Ministry of Education
Tsinghua National Laboratory for Information Science and Technology (TNList)
Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China

ABSTRACT

Waveform clipping is a common problem in digital audio signal processing. Clipping effect changes the audio amplitude distribution and brings noise, which results in the decrease of audio signal-to-noise ratio (SNR). To estimate the degree of clipping effect, we propose a method for SNR estimation of clipped signals base on audio amplitude distribution. We first locate the clipping intervals of audio signal with high precision based on the differences between the amplitude probability density functions (PDFs) of normal and clipped audio signals. Then, the clipped amplitude PDF is restored using Gamma distribution to estimate the non-clipped amplitude distribution. Finally, based on the restored PDF, the estimated SNR of distorted audio signals are calculated, which can indicate the degree of clipping effect. Experiments on 3200 audios show that our approach can achieve a low error of 0.017 in clipped values locating, and an error less than 0.5dB in SNR estimation.

Index Terms—amplitude PDF, Gamma distribution, clipping effect, SNR estimation

1. INTRODUCTION

To record sounds precisely, the amplitude of the input signal needs to be large enough to increase signal-to-noise ratio (SNR), where the noise comes from electric hum, quantization noise etc. However, when the amplitude exceeds the limit of the recording devices, it will be cut off to the highest receivable level. This phenomenon is called *clipping effect*, and of the highest receivable level is called the *clipping value*. Sometimes the audio is not clipped by a constant value but by a value varies in a certain interval. In this case, the interval is called the *clipping interval*, and the lower bound of the clipping interval is treated as the clipping value. SNR estimation of clipping audio is helpful to estimate the severity of clipping effect in audio quality evaluation.

In this paper, we focus on the problem of SNR estimation of clipped audio signals. The SNR of clipped audio signals can be estimated through waveform restoration, which can restore the clipped samples. There are several existing waveform restoration algorithms. Approximate waveform substitution restores the clipped signals by double side periodic substitution methods [1-2] while approximate spectrum substitution method is similar method in the spectrum domain [3]. Linear prediction method [4] predicts the amplitudes of the clipped samples with the last few samples. Recursive Vector Projection is used to fit the clipped samples with the surrounding un-clipped samples [6]. The SNR of clipped audio signal can be calculated with the clipped and restore signal. The SNR of clipped signals can also be estimated directly without waveform restoration using general SNR estimation methods, such as the Maximum-Likelihood Estimator of SNR [7-9] or Second- and Fourth-Order Moments Estimator [10-12]. In these methods, there are two key technological problems:

1) Clipping values, which is essential in waveform restoration, cannot be easily obtained as it is generally in a state of flux rather than a constant. The sample comparison approach used in existing methods would lead to some detection error.

2) Most general SNR estimating methods would lead to large estimation errors for clipping noise.

To solve the above problems, we propose an SNR estimation method for clipped audio based on amplitude estimation. First, the clipping value is obtained using the amplitude PDF difference between clipped and normal signals. Then, Gamma distribution is used to restore the clipped PDF. Finally SNR of the clipped signal is estimated according to the clipped and restored amplitude PDF. The estimated SNR could be further used in audio quality evaluation or recording equipment adjustment.

2. CLIPPING VALUE LOCATING BASED ON AMPLITUDE PDF

In this section, we analyze the clipping effect of audio signals, and propose a method for locating clipping intervals using waveform amplitude PDF.

2.1. Characteristics of clipped audio signals

A clipped audio signal is a mixture of the original signal and the clipping noise, with high-amplitude samples clipped to

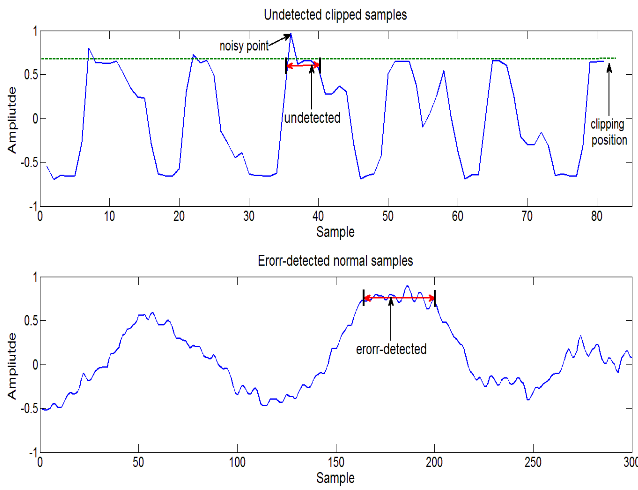

Figure 1. The upper figure shows the undetected clipped samples due to the noisy local maximum point; the lower figure shows the error-detected normal samples for the smooth changing at waveform peak.

low level. Let the original audio signal be x_i, the positive and negative clipping values C^+ and C^-, then clipped audio signal $y_i$is:

$$y_i = \begin{cases} C^+ & x_i > C^+ \\ C^- & x_i < C^- \\ x_i & \text{otherwise} \end{cases} \quad (1)$$

And the clipping noise n_i can be calculated as:

$$n_i = y_i - x_i \, , i = 1, \dots, N \quad (2)$$

In many existing methods clipped segments are located simply by comparing amplitudes of adjacent samples [3]. If the amplitudes of some adjacent samples are approximately equal to each other and are close to the maximum of the entire waveform, they will be marked as clipped, and the minimum amplitude of the clipped samples will be treated as the clipping value. However, the clipping value is generally in a state of flux rather than a constant that the high-amplitude samples are cut to the clipping interval. And the clipping value might be much smaller than the maximum of the whole waveform when the maximum value comes from some random noise. The above mentioned method would result in some undetected clipped segments and error-detected normal segments, as is shown in Figure 1. The undetected clipped segments are the results of a random noise point with high amplitude. And the error-detected ones are usually caused by some normal adjacent samples whose amplitude changes smoothly near the local peak of the waveform.

2.2. Locate clipping intervals using amplitude PDF

As all samples whose amplitude is larger than the clipping value would be clipped to the clipping value, and the

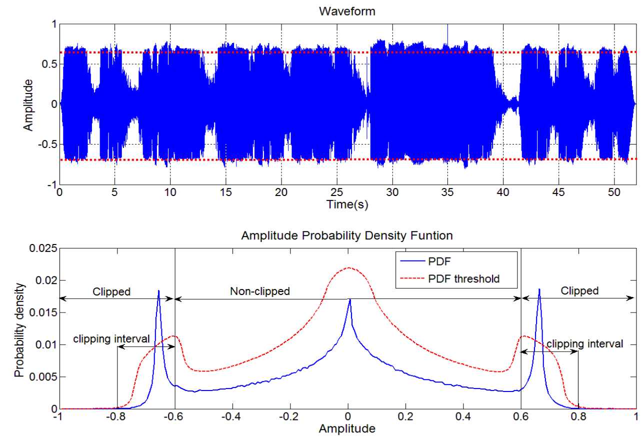

Figure 2. The upper figure is a wave clipped at [0.6, 0.8] and [-0.8, -0.6]; the lower one shows amplitude PDF and PDF threshold vector.

probability of the amplitudes in the clipping interval in the amplitude PDF are significantly larger than other points in neighborhood. Therefore, the amplitude distribution could be used to locate the clipping intervals. Figure 2 shows the amplitude PDF of a clipped signal with blue line. The probability densities in clipping intervals [0.6, 0.8] and [-0.8, -0.6] are much larger than the values around to them.

To automatically detect the clipping intervals in the amplitude PDF, a threshold function is generated from the smoothed PDF to locate the sharp peaks in amplitude PDF caused by clipping:

$$PDF_i^* = \frac{D}{2l+1} \sum_{k=i-l}^{i+l} PDF_k \, , i = -M, \dots, M \quad (3)$$

where D is the weight of detecting threshold, l is smooth step length and M is the maximum possible amplitude value of the digital audio signal. PDF_i^* is the detection threshold for amplitude i and PDF_i is the probability density of amplitude i.

If the probability density of a positive amplitude value, PDF_i, is larger than the threshold PDF_i^*, then the local minimum value of PDF between position 0 and i, namely $C^+ = argmin_{0<j\leq i} PDF_j$, will be lower bound of the clipping interval, which is also the clipping value. The whole clipping interval will extend to where the probability of amplitude goes to zero. As is shown in Figure 2, the upper figure shows the clipped audio waveform and the red dotted line is the approximate clipping value. The lower figure shows the amplitude PDF of the audio signal in blue and the PDF threshold vector in red dotted line with weight $D = 2$. It is obvious that the samples whose amplitude are greater than 0.6 are clipped into an interval of [0.6, 0.8] while those ones less than -0.6 are clipped into [-0.8, -0.6].

In the next section, Gamma distribution is used to restore the amplitude PDF based on the obtained clipping value, and then SNR is estimated.

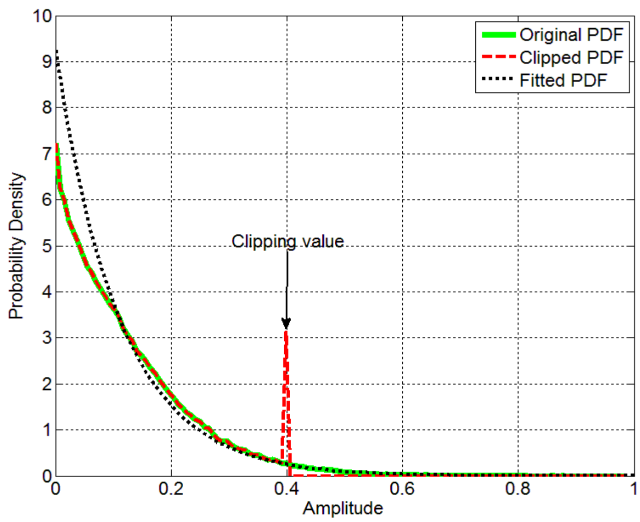

Figure 3. Comparison of the original, clipped, and the restored amplitude PDF

3. SNR ESTIMATION BASED ON GAMMA DISTRIBUTION

As the clipping intervals are located, Gamma distribution is utilized to restore the non-clipped PDF, and then SNR can be estimated with the restored distribution in the whole amplitude range.

3.1. Estimate amplitude PDF with Gamma distribution

It is widely known that the symmetric gamma distribution is a good approximation to the amplitude distribution of a large speech corpus [13]. The probability density function, $f(x)$ of clean audios can be represented by the following function:

$$f(x;\alpha,\beta) = \frac{\beta}{2\Gamma(\alpha)}(\beta|x|)^{\alpha-1}e^{-\beta|x|} \tag{4}$$

where x is amplitude; α and β are the shaping and rate parameter of gamma distribution.

Therefore, we use the values in the non-clipped area of the amplitude PDF to estimate the real amplitude distribution in positive and negative part respectively. To estimate the values of the parameters α,β in Gamma distribution and avoid computational complexity, the maximum likelihood estimators of the two parameters are first calculated, and then their values are refined with the least mean square (LMS) criterion.

The maximum likelihood estimators of α,β are as follows [14]:

$$\hat{\beta}^{+} = \frac{1}{\hat{\alpha}^{+}C^{+}}\sum_{i=1}^{C^{+}} PDF_i \tag{5}$$

$$\hat{\alpha}^{+} = \frac{3 - s^{+} + \sqrt{(s^{+}-3)^2 + 24s^{+}}}{12s^{+}} \tag{6}$$

$$s^{+} = ln\left(\frac{1}{C^{+}}\sum_{i=1}^{C^{+}} PDF_i\right) - \frac{1}{C^{+}}\sum_{i=1}^{C^{+}} ln(PDF_i) \tag{7}$$

where $\hat{\alpha}^{+}$ and $\hat{\beta}^{+}$ are the preliminary estimators of α and β respectively. The values of α and β are further refined with LMS criterion in their neighborhood. The formula is shown as follows:

$$(\bar{\alpha}^{+},\bar{\beta}^{+}) = \underset{\substack{\alpha\epsilon(\hat{\alpha}^{+}-\delta_\alpha,\hat{\alpha}^{+}+\delta_\alpha)\\ \beta\epsilon(\hat{\beta}^{+}-\delta_\beta,\hat{\beta}^{+}+\delta_\beta)}}{argmin}\left(\frac{1}{C^{+}}\sum_{i=1}^{C^{+}}(f(i;\alpha,\beta) - PDF_i)^2\right) \tag{8}$$

where $\bar{\alpha}^{+}$ and $\bar{\beta}^{+}$ are the final estimated values of α and β for the positive half of the PDF, δ_α and δ_β are length parameters of calculation neighborhood.

Similar process is carried out to get $(\bar{\alpha}^{-},\bar{\beta}^{-})$ in the negative half of the amplitude PDF. With the estimated values of α and β, the amplitude PDF of the original audio signal in positive and negative part can be restored respectively because C^{+} and C^{-} maybe not symmetric.

3.2. Amplitude PDF restoration and SNR estimation

As both of the two parameters of estimated amplitude PDF are obtained, the PDF of the clipped audio signal can be estimated, and the number of estimated values should let the summarize of the estimated probability close to the probability of the positive clipped PDF as much as possible, shown as:

$$\overline{PDF_i} = \frac{\bar{\beta}}{2\Gamma(\bar{\alpha})}(\bar{\beta}|i|)^{\bar{\alpha}-1}e^{-\bar{\beta}|i|}, i = C^{+},\ldots,L \tag{9}$$

$$L = \underset{L>C^{+}}{argmin}\left|\sum_{i=C^{+}}^{L}\overline{PDF_i} - \sum_{i=C^{+}}^{M} PDF_i\right| < \delta \tag{10}$$

where $\overline{PDF_i}$ is the estimated probability density of amplitude i, L is the maximum amplitude where PDF needs to be restored, and δ is a small constant. An example of original amplitude PDF, clipped PDF and estimated PDF is shown in Figure 3.

Then, the estimated SNR of the clipped audio signal could be calculated. The estimated signal energy is calculated with the non-clipped and restored PDF, and the estimated noise energy is calculated with the restored PDF in clipped part. As is shown in the following formula:

$$n_i = i - C^{+} + 1, i = C^{+},\ldots,L \tag{11}$$

$$\widehat{SNR} = 10log_{10}\frac{\sum_{i=1}^{C^{+}-1}PDF_i i^2 + \sum_{i=C^{+}}^{L}\overline{PDF_i}i^2}{\sum_{i=C^{+}}^{L}\overline{PDF_i}{n_i}^2} \tag{12}$$

where n_i is the noise brought by the clipped samples with amplitude of i, $\widehat{SNR}$ is the estimated SNR of the clipped audio signal.

4. EXPERIMENTS

Two experiments are conducted to test the accuracy and efficiency of the proposed method. The first experiment illustrates the accuracy of our method in clipping value estimation; and the second experiment is carried out to test the efficiency and effectiveness of the proposed clipped audio SNR estimation method.

4.1. Clipping value estimation based on amplitude distribution

The dataset for clipped interval locating experiment contains audio files downloaded from the internet. The files are manually labeled and 3200 audio files are chosen for the experiment, while all clipped audio files are kept and the normal ones are chosen randomly. Among the 3200 audio files, 2000 are clipped while the other 1200 are normal. The audio length is between 50s and 500s. All audio files are converted to the sampling rate of 8kHz and windows PCM format. The bit depths of the audio may be either 8 or 16. There are total 2000 audio files clipped at a value between 0.2 and 0.9, among which 500 audio files are clipped by a clipping interval while the other 1500 are clipped by a constant value.

In the experiment, parameter D and l in equation (3) are set to 2.2 and 10. Then clipping values of positive and negative amplitude is respectively obtained by the method proposed in section 2.2. The common method used in waveform restoration is implemented and used for comparison [3].

As is shown in Table 1, the proposed method of locating clipping values achieves a mean error of 0.017 while the estimated maximum amplitude L in equation (10) is normalized to 1.0, significantly better than the method from [3]. And the results from signals clipped by interval prove that our method is barely affected by the fluctuating clipping value.

Table 1. Results of the proposed and original method

Mean error	Clipped by fluctuating value	Clipped by constant	Total
The method in [3]	0.124	0.032	0.063
The proposed method	**0.031**	**0.011**	**0.017**

4.2. SNR estimation based on Gamma distribution

After clipping values are obtained, the following experiment is carried out to test the accuracy of SNR estimation method. Because clipped audio signals with high SNR don't significantly affect the perceived quality, 768 clipped audio files whose SNRs are smaller than 30dB are selected from the above dataset. The estimated SNR is calculated by equation (12) with δ_α and δ_β set to 2 and 4.

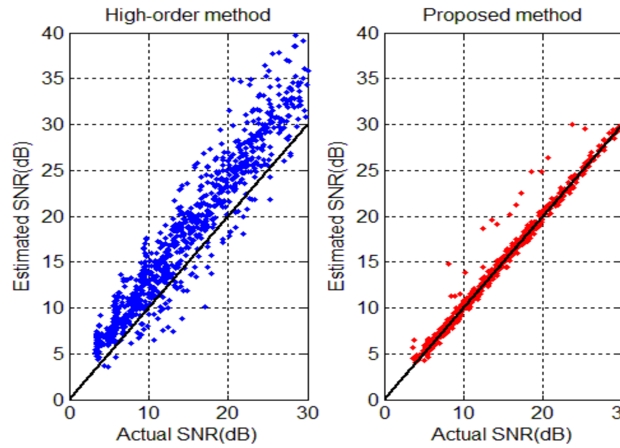

Figure 4. The results of SNR estimation. The points are the estimated results while the black line is theoretical value.

Two methods are used in the comparison experiment with the proposed method. One is SNR estimation with waveform restoration using the approximate waveform substitution method [5], which has the least computation complication in the discussed methods. The other is the fourth-order moments method for general SNR estimation [9]. The estimating results by the proposed method give a mean error of 0.47dB, while the SNR estimation method based on fourth-order moments gives a mean error of 4.28dB and the approximate waveform substitution gives 0.61dB. The estimation results for the proposed method and the fourth-order moments method are shown in Figure 4.

Computation complexity is also test for the three methods. All the methods are implemented in JAVA. The test in carried out on a computer with a CPU of Core i5 760 and memory of 4GB. An audio signal which lasts for 100s and sampled at 16kHz is used in the test. The proposed method takes 0.72s to estimate the SNR. While the approximate waveform substitution method takes 7.4s and fourth-order moments estimator method takes 2.3s. It is clearly that the proposed SNR estimating method for clipped signal is efficient and effective for audio quality evaluation.

5. CONCLUSION

Unlike existing clipped signal restoration and SNR estimation methods, we propose a method to estimate the SNR of clipped audio signals based on characteristics of clipping effect and the regularity of amplitude distribution. The clipping value is obtained based on the difference between clipped and normal amplitude PDFs. Gamma distribution is utilized to restore the clipped amplitude PDF, which is then used to estimate the SNR of the clipped audio signal. Experiment results show that out method can reach a high accuracy in clipping interval detection and SNR estimation. The computational complexity is also much lower than other SNR estimation algorithms.

ACKNOWLEDGMENT

This work is partially supported by the National Basic Research Program (973 Program) of China (2013CB329304), and the National High Technology Research and Development Program (863 ” Program) of China (2012AA011602), and the National Natural and Science Foundation of China (61003094).

REFERENCES

[1] D. Zhang, C. Bao, F. Deng, B. Xia, H. Chen, “A Restoration Method of the Clipped Audio Signals based on MDCT”, ISSPIT-2011, pp. 253-257, December 2011.
[2] Y. Ren and M. T. Johnson, “An Improved SNR Estimator for Speech Enhancement”, ICASSP-2008, pp. 4901-4904, 2008.
[3] Abel, J. S. and Smith, J. 0, III, “Restoring a Clipped Signal”, ICASSP-91, vol. 3, pp. 1745-1748, 1991.
[4] W. Yang and Y. Ben-Zion, “An Algorithm for Detecting Clipped Waveforms and Suggested Correction Procedures”, Seismological Research Lettersvol. 81 no. 1, pp. 53-62, January/February 2010.
[5] A. Dahimene, M. Noureddine, A. Azrar, “A Simple Algorithm for the Restoration of Clipped Speech Signal”, Informatica, vol. 32, no. 2, pp. 183-188, 2008.
[6] S. Miura, H. Nakajima,S. Miyabe, S. Makino, T. Yamada, K. Nakadai, “Restoration of Clipped Audio Signal Using Recursive Vector Projection,” TENCON-2011, pp. 394-397, 2011.
[7] A. Wiesel, J. Goldberg, H. Messer, “Non-Data-Aided Signal-to-Noise-Ratio Estimation”, ICC-2002, vol. 1, pp. 197-201, 2002.
[8] P. Gao and C. Tepedelenlioglu, “SNR Estimation for Non-Constant Modulus Constellations”, WCNC-2004, vol. 1, pp. 24-29, March 2004.
[9] R. Matzner and F. Euglberger, “An SNR Estimation Algorithm Using Fourth-Order Moments”, ISIT-1994, pp. 119, 1994.
[10] E. Nemer, R. Goubran, S. Mahmoud, “SNR Estimation of Speech Signals Using Subbands and Fourth-Order Statistics”, Signal Processing Letters, vol. 6, pp. 171-174, 1999.
[11] D. R. Pauluzzi and N. C. Beaulieu, “A Comparison of SNR Estimation Techniques for the AWGN Channel”, IEEE Transactions on Communications, vol. 48, pp. 1681-1691, 2000.
[12] R. Matzner, “An SNR Estimation Algorithm for Complex Baseband Signals Using Higher Order Statistics”, Facta Universitatis , vol. 6, no. 1, pp. 41-52, 1993.
[13] C. Kim and R. M. Stern, “Robust Signal-to-Noise Ratio Estimation Based on Waveform Amplitude Distribution Analysis”, InterSpeech2008, pp. 2598-2601, 2008.
[14] S. Intajag, S. Chitwong, “Speckle Noise Estimation with Generalized Gamma Distribution,” SICE-ICASE2006, pp.1164-1167, 2006.