The Removal of Impulsive Noise in Music Signals Using Higher-Order Spectra

3846 (P11.6)

F.L. Hui and W.H. Lau
Department of Electronic Engineering
City Polytechnic of Hong Kong
83 Tat Chee Avenue, Hong Kong

**Presented at
the 96th Convention
1994 February 26 - March 01
Amsterdam**

AES

AN AUDIO ENGINEERING SOCIETY PREPRINT

The Removal of Impulsive Noise in Music Signals Using Higher-Order Spectra[1]

F.L. Hui, and W.H. Lau
Department of Electronic Engineering
City Polytechnic of Hong Kong
83 Tat Chee Avenue, Hong Kong

Abstract

This paper proposes a new method to restore music signals corrupted by impulsive noise. Since the probability density function of the music signals can be approximated by Gaussian distribution, the non-linear nature of the impulsive noises becomes prominent in the cumulant spectra of the music signals. Such feature enables the use of the higher-order spectra to detect the presence of the impulse noises. The music signals can then be restored by extracting the impulse noise.

1. Introduction

Over the past decades, many valuable gramophone and LP records have been corrupted with scratches which produce unpleasant clicking noises during playback. Since these records cannot be recorded again, the only way to preserve them is to reproduce and store them in reliable storage media, e.g. compact disc (CD) or digital audio tape (DAT). However, it is known that both the clicking and hiss noises will be present during the playing of these old records. In this paper, we will only concentrate on dealing with the clicking noises. These noises are caused by the presence of scratches on the groove surface, picked up by the needle of the pick-up cartridge system. The clicking noise may be approximated by spread impulsive noises. Many researchers have worked on the restoration of damaged records by removing the impulsive noises, e.g. [1,2,3]. Most of their methods are based on a model approach or by exploiting the power spectrum of the music signals. In order to restore these corrupted records and preserve them in high-fidelity, the sampling rate is usually very high, e.g. 44.1kHz

---

1 This project is supported by Hong Kong UPGC research grant 904064.

for CD storage. With such a high sampling rate, experimental results have shown that the linear prediction model based approach easily experience difficulty in detecting the impulsive noises unless it is very sharp in duration. In fact it is difficult to find the correct order for a predictor to detect both spread and sharp impulsive noises. In this paper, we present a new method based on the high-order statistics of the music signals to detect and remove the impulsive noises by exploiting its inherent non-linear nature.

2. Definitions and Properties of Higher-Order Statistics

This section will introduce the definitions, properties and computation of higher-order statistics (spectra) [4], i.e. moments and cumulants, and their corresponding higher-order spectra. In recent years higher-order statistics have found applications in many diverse fields. It is because the higher-order statistics contain important information, such as the degree of non-linearity and deviation from normality, about the random processes that are not contained in the power spectrum (order=2). Moreover, the higher-order spectra (order>2) are blind to any kind of Gaussian process [5]. These features make the high-order spectra an ideal tool to extract the non-linear impulsive noises. The introduction is for both stochastic and deterministic signals, however, the emphasis of the discussion is placed on the 3rd-order statistics and its respective Fourier transforms, bispectrum.

Let $x(t)$ be a stationary zero-mean random sequence and its moments up to order k exist. The kth-order cumulant of this random signal is defined as the coefficient of $(v_1\ v_2\ \ldots\ v_\kappa)$ in the Taylor series expansion of the cumulant generating function, $K(v)$, i.e.

$$K(v) = \ln E\{\exp(jv'x)\}\,. \tag{1}$$

And the kth-order moment $M_k^x(\tau_1, \ldots, \tau_{k-1})$ and cumulant $C_k^x(\tau_1, \ldots, \tau_{k-1})$ of a non-Gaussian stationary random signal $x(t)$ are defined as:

$$M_k^x(\tau_1, \ldots, \tau_{k-1}) \cong E[x(t)x(t+\tau_1), \ldots, x(t+\tau_{k-1})] \tag{2}$$

$$C_k^x(\tau_1, \ldots, \tau_{k-1}) = M_k^x(\tau_1, \ldots, \tau_{k-1}) - M_k^G(\tau_1, \ldots, \tau_{k-1}) \tag{3}$$

where E denotes the expectation operator. From the above equations, it can be seen that both the kth-order moment function and cumulant function of the stationary signal depend only on the time differences $\tau_1\ \tau_2, \ldots \tau_{n-1}$, $\tau_i = 0, \pm 1, \pm 2, \ldots$ for all i. For $k \leq 2$, both the cumulants and moments are equal if $x(t)$ is a zero-mean sequence, otherwise, the cumulants and moment are not equal for k>3. The kth-order cumulants may be expressed in terms of the kth-order moment, $M_k^x(\tau_1, \ldots, \tau_{k-1})$, and lower order moments, $M_k^G(\tau_1, \ldots, \tau_{k-1})$.

For a finite length sequence, $\{x(t)\}_{i=0}^{N}$, the kth-order correlation is defined as

$$x_k(\tau_1, \ldots, \tau_{k-1}) \cong \sum_{t=0}^{N} x(t)x(t+\tau_1)\ldots x(t+\tau_{k-1}). \tag{4}$$

And, the third-order cumulant (k=3) is equivalent to the *moment function of order* 3, *also called the tri-correlation*, which is defined as

$$C_3^x(\tau_1, \tau_2) = M_3^x(\tau_1, \tau_2) \cong E[x(t)x(t+\tau_1)x(t+\tau_2)] \ . \tag{5}$$

Thus, given N+1 samples of $x(t)$, the third order cumulant and moment may be consistently estimated. For example, the third order cumulant may be estimated as

$$C_3^x(\tau_1, \tau_2) = M_3^x(\tau_1, \tau_2) = \frac{1}{N+1} x_3(\tau_1, \tau_2) \ , \tag{6}$$

and the kth-order moments may be consistently estimated by

$$M_k^x(\tau_1, \ldots, \tau_{k-1}) = \frac{1}{N+1} x_k(\tau_1, \ldots, \tau_{k-1}) \ . \tag{7}$$

Most important of all, cumulants have several desirable properties. For $k \geq 3$, all cumulants of the Gaussian processes or uniform distributed signal, $x(t)$, are identical to zero, i.e.

$$C_k^x(\tau_1, \ldots, \tau_{k-1}) = 0 \qquad k \geq 3 \ , \tag{8}$$

and the cumulant between $x(t)$ and its circular shifted $x'(t)$ sequence are the same, i.e.

$$x'(t) = x(t+\tau) \tag{9}$$

$$C_k^{x'}(\tau_1, \ldots, \tau_{k-1}) = C_k^x(\tau_1, \ldots, \tau_{k-1}) \tag{10}$$

With the above properties, if we have a signal $y(n)$, comprised with two independent signals $x(n)$ and $d(n)$, defined as

$$y(n) = x(n) + d(n). \tag{11}$$

Its kth-order cumulants can be written as

$$C_k^y(\tau_1, \ldots, \tau_{k-1}) = C_k^x(\tau_1, \ldots, \tau_{k-1}) + C_k^d(\tau_1, \ldots, \tau_{k-1}) \ . \tag{12}$$

If $x(n)$ is a Gaussian signal, $d(n)$ is non-Gaussian or non-uniform distributed signal, and $k \geq 3$, then (12) can be reduced to

$$C_k^y(\tau_1, \ldots, \tau_{k-1}) = C_k^d(\tau_1, \ldots, \tau_{k-1}). \tag{13}$$

This shows that one can use cumulants to extract non-Gaussian signal/noise out of Gaussian signals. In fact this important property forms the basis of the new approach presented in this paper.

Higher-order spectra are defined in terms of either cumulants (e.g., cumulant spectra) or moments (e.g., moment spectra). Generally speaking, higher-order spectra are the multi-dimensional Fourier transform of the higher-order statistics. For the third-order cumulant, its associated spectra is called the *bispectrum* which is defined as follows:

$$B_3^x(\omega_1, \omega_2) = \sum_{\tau_1=-\infty}^{\infty} \sum_{\tau_2=-\infty}^{\infty} C_3^x(\tau_1, \tau_2) e^{-j(\omega_1\tau_1+\omega_2\tau_2)} \tag{14}$$

where $|\omega_1| \leq \pi,\ |\omega_2| \leq \pi,\ |\omega_1 + \omega_2| \leq \pi$.

Since the computation of the cumulants is carried out in multi-dimensions, the symmetry properties of the cumulants are exploited to reduce computations. For third order cumulants, the symmetry regions are as follows:

$$\begin{aligned} C_3^x(\tau_1, \tau_2) &= C_3^x(\tau_2, \tau_1) = C_3^x(-\tau_2, \tau_1 - \tau_2) \\ &= C_3^x(\tau_2 - \tau_1, -\tau_1) = C_3^x(\tau_1 - \tau_2, -\tau_2) \\ &= C_3^x(\tau_1, \tau_2 - \tau_1) \end{aligned} \tag{15}$$

As a consequence, symmetric regions of the bispectrum are also existed and given as follows:

$$
\begin{aligned}
B_3^x(\omega_1,\omega_2) &= B_3^x(\omega_2,\omega_1) \\
&= B_3^{x^*}(-\omega_2,-\omega_1) = B_3^x(-\omega_1-\omega_2,\omega_2) \\
&= B_3^x(\omega_1,-\omega_1-\omega_2) = B_3^x(-\omega_1-\omega_2,\omega_1) \\
&= B_3^x(\omega_2,-\omega_1-\omega_2) \qquad (16)
\end{aligned}
$$

For real processes there are 12 symmetric regions existed in the bispectrum. Hence, the information of the bispectrum within the triangular region $\omega_2 \geq 0, \omega_1 \geq \omega_2, \omega_1+\omega_2 \leq \pi$ completely describes the bispectrum. These are the vital information for reconstructing the non-Gaussian signal and will be covered in more detail in next section.

3. The new method

The task of our excise is to detect the presence of the impulsive noises and to remove them. From (11) to (13), we have made the following two assumptions:

(i) the music signals and the impulsive noises are independent; and

(ii) the distribution of the music signals can be approximated by Gaussian distribution.

It should be noted that assumption (ii) is only true for a long sequence and is required to satisfy (8). For our application, the minimum order required is 3 and the third-order cumulant [5], $C_3^x(\tau_1,\tau_2)$, over N samples of a zero-mean process $\{x(t)\}$ is given by

$$C_3^x(\tau_1,\tau_2) = \frac{1}{N}\sum_{t=0}^{N-1} x(t)x(t+\tau_1)t(t+\tau_2) \qquad (17)$$

A block diagram of the impulsive noise removal system is shown in Figure 1. For high sampling rate, a relative long data frame for analysis is required to satisfy the assumption that the music signals is Gaussian distributed, or otherwise the cumulant may not approach to zero. Unfortunately, the long data record inevitably incurs heavy computational burden in calculating the cumulant and a new *Recursive-Averaging Cumulant Estimation* (RACE) technique has been developed to substantially reduce the computational requirement. The RACE technique is particular suitable for applying to sequential data stream with arbitrary length. Details of the technique is explained in the following.

3.1 The Recursive-Averaging Cumulants Estimation (RACE) technique

Consider a data sequence of length $2N$, the estimated cumulants, $C^x(\tau_1,\tau_2)_{2N}$, can be approximated by splitting the summation into two parts, $C_1^x(\tau_1,\tau_2)_{2N}$ and $C_2^x(\tau_1,\tau_2)_{2N}$,

$$\begin{aligned} C^x(\tau_1,\tau_2)_{2N} &= \frac{1}{2N}\sum_{t=0}^{2N-1} x(t)x(t+\tau_1)x(t+\tau_2) \\ &\approx \frac{1}{2N}\Big[\sum_{t=0}^{N-1} x(t)x(t+\tau_1)x(t+\tau_2) + \sum_{t=N}^{2N-1} x(t)x(t+\tau_1)x(t+\tau_2)\Big] \\ &\approx \frac{1}{2N}\big[N\cdot C_1^x(\tau_1,\tau_2)_N + N\cdot C_2^x(\tau_1,\tau_2)_N\big] \end{aligned} \tag{18}$$

where

$$C_1^x(\tau_1,\tau_2)_N = \frac{1}{N}\sum_{t=0}^{N-1} x(t)x(t+\tau_1)x(t+\tau_2) \tag{19a}$$

$$C_2^x(\tau_1,\tau_2)_N = \frac{1}{N}\sum_{t=N}^{2N-1} x(t)x(t+\tau_1)x(t+\tau_2) \tag{19b}$$

From the above equations, a recursive formula for calculating the third-order cumulants can easily be deduced, i.e.

$$\tilde{C}_k^x(\tau_1,\tau_2)_N = \frac{1}{k\cdot N}\big[N\cdot C^x(\tau_1,\tau_2)_N + (k-1)\cdot N\cdot \tilde{C}_{k-1}^x(\tau_1,\tau_2)_N\big] \tag{20}$$

The recursive-averaging formula is able to modify the cumulants of the current data frame, $C^x(\tau_1,\tau_2)_N$, by accumulating the cumulants of past data frames, $\tilde{C}_{k-1}^x(\tau_1,\tau_2)_N$.

3.2 Imuplsive noises detection

Since the impulsive noises are not necessarily present in the analysis data record, the energy of the cumulant is used as an indicator to detect the presence of the impulse. Since the music signals is assumed to be Gaussian distributed, the cumulant should be very small and approaches to zero, but not zero. This value is dependent upon the statistical distribution of the music signals within the data frame. Hence, we also assume that the energy of the cumulant of the music signals is also Gaussian distributed. If a non-linear component, such as the impulsive noise, is present in the music signals, the energy of cumulant becomes larger and the presence of the impulse or non-linear components can easily be detected. The detection is accomplished by applying a binary hypothesis test with a pre-defined threshold for the energy of the cumulant. In order to improve the accuracy of the threshold detection in the Gaussian distributed model, an adaptive technique [6] is applied to the threshold detection. The principal parameters for the model and the hypothesis testing are the mean and variance of the energy of cumulant. These parameters are adaptively adjusted throughout the detection process.

3.3 Impulse reconstruction

Having detected the presence of the impulsive noise, the next step is to reconstruct the impulse components using the bispectrum. Since the Gaussian and the periodic components of the music signals are suppressed by the cumulant, only the information of the non-linear components of music signals remains in the cumulant. The bispectrum of the impulse can then be obtained through the Fourier Transformation of the cumulant. The Bispectrum estimation [7] is given by:

$$B_3^x(\omega_1,\omega_2)=\sum_{\tau_1=-L_n}^{L_n}\sum_{\tau_2=L_n}^{L_n} C_3^x(\tau_1,\tau_2)w(\tau_1\Delta_n,\tau_2\Delta_n)e^{-j(\omega_1\tau_1+\omega_2\tau_2)} \tag{21}$$

where $w(u_1,u_2)$ is a window function of bounded support and Δ_n is the bandwidth usually taken as $\Delta_n = 1/L_n$. $B_3^x(\omega_1,\omega_2)$ has to be smoothed by a weighting function, $\{K_M(\theta_1,\theta_2)\}$, known as a *spectral window*. The propose of using the spectral window is to obtain a consistent estimate of the bispectrum. In this step, the Optimum Bispectral Window is chosen, and $K_M(\theta_1,\theta_2)$ is given by:

$$K_M(\theta_1,\theta_2)=\begin{cases}\frac{\sqrt{3}}{\pi^3}[1-\frac{1}{\pi^2}(\theta_1^2+\theta_2^2+\theta_1\theta_2)] & \text{, if } (\theta_1,\theta_2)\in G_1 \\ 0 & \text{, otherwise}\end{cases}, \tag{22}$$

where the region G_i is given by the set $\{(\theta_1,\theta_2);\theta_1^2+\theta_2^2+\theta_1\theta_2\le\pi^2\}$.

In signal reconstruction problems using the estimated bispectrum, the relationship between the bispectrum, $B_3^x(\omega_1, \omega_2)$ and the power spectrum of $x(t)$, $X(\omega)$, are given as follow:

$$B_3^x(\omega_1, \omega_2) = \frac{1}{N} X(\omega_1) X(\omega_2) X^*(-\omega_1 - \omega_2), \tag{23}$$

and there are two governing conditions that must be satisfied, i.e.,

$$\beta(\omega_1, \omega_2) = \alpha(\omega_1) + \alpha(\omega_2) - \alpha^*(-\omega_1 - \omega_2) \tag{24}$$

$$A(\omega_1, \omega_2) = \frac{1}{N}\, t(\omega_1) t(\omega_2) t^*(-\omega_1 - \omega_2) \tag{25}$$

where the β and A are functions containing the phase and magnitude information of the impulse in the frequency domain.

By appropriately choosing the boundary values for ω_1 and ω_2, the impulse in the time domain can be reconstructed. The reconstruction algorithm used is the recursive Bartelt-Lohman-Wirnitzer Algorithm [8,9]. The Bartelt-Lohman-Wirnitzer algorithm, although recursive in nature, does not require phase unwrapping methods, and utilizes all bispectral values. Finally, the impulse can then be extracted from the music signals and only a simple subtraction operation is required to remove the impulse from the music signals.

4. Results

Figure 2(a) shows an example of a record of music signals corrupted with an impulse. The sampling rate used is 44.1kHz. By using the new method described above, the impulse has been detected and extracted. The reconstructed signals, as shown in Figure 2(b), is obtained by subtracting the extracted impulse from the corrupted signal.

5. Conclusions

A new method using high-order statistics to detect and remove the impulsive noise from music signals has been proposed. It has been shown this method can detect the presence of the non-Gaussianality nature of the impulsive noises embedded in music signals. The impulsive noises can then be rebuilt and removed from the music signals. The preliminary results obtained from both simulated and real data are very promising. Current effort is to investigate the robustness of the detection technique in various kinds of music signals and improvement to be made in the reconstruction of the music signals.

6. References

[1] S.V. Vaseghi and R. Frayling-Cork, "Restoration of Old Gramophone Recordings," JAES, Vol. 40, No 10, pp. 791-801, 1992.

[2] K.S. Jog and Ajay Ingle, "Sound Restoration in Damaged Accoustical Recording," Proc. of ICASS' 92, pp. II-257-260, 1992.

[3] K. Cisowski, "Efficiency of Impulsive Noise Detection in Musical Recordings Using the Adaptive Filtering Method," Presented at the 94th Convention of the Audio Engineering Society, preprint 3464, 1993.

[4] J.M. Mendel, "Tutorial on Higher-Order Statistics (Spectra) in Signal Processing and System Theory: Theoretical Results and Some Applications," Proc. IEEE, vol.79, No.3, pp.278-305, 1991.

[5] C.L. Nikias, and A.P. Petropulu, "Higher-Order Spectra Analysis: A Nonlinear Signal Processing Framework," Prentice Hall, 1993

[6] M. Yoshikazu, M. Nobuhiro, N. Nobuo, and H. Kozo, "A Speech Analysis Algorithm which Eliminates the Influence of Pittch Using the Model Reference Adaptive System," ASSP, 1982, Vol.30, No.1, pp.88-96.

[7] T.S. Rao and M.M. Gabr, "An Introduction to Bispectral Analysis and Bilinear Time Series Models," Lecture Notes in Statistics, 24, New York:Springer-Verlag, 1984.

[8] H. Bartelt, A.W. Lohmann, and B. Wirnitzer, "Phase and Amplitude Recovery from Bispectra," Applied Optics, 1984, Vol.23, No.18, pp.3121-3129.

[9] M. Toshifumi, and T.J. Ulrych, "Phase Estimation Using the Bispectrum," Proc. IEEE, 1984, Vol.72, No.10, pp1403-1411.

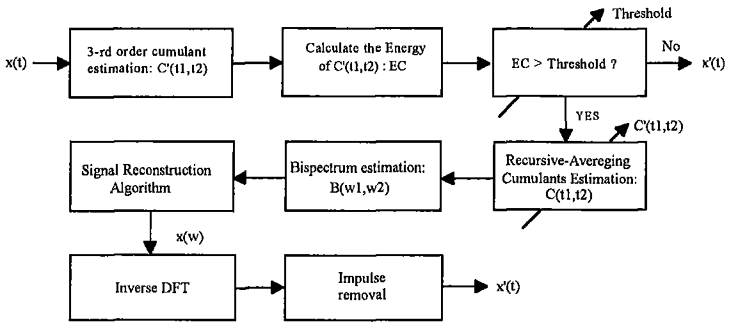

Figure 1: Block diagram of the impulsive noise removal system using high-order spectra

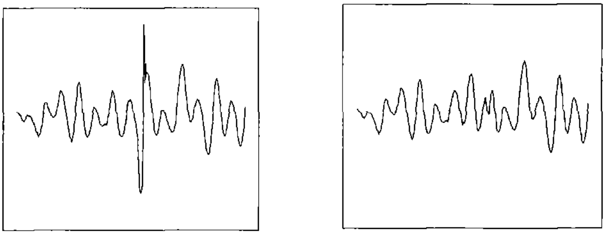

Figure 2: (a) A record of music signals corrupted with an impulse,

(b) the reconstructed music signals with the impulse removed.