GENERALIZED HARMONIC ANALYSIS OF ARC-TANGENT SQUARE ROOT (ATSR) NONLINEAR DEVICE FOR VIRTUAL BASS SYSTEM

Nay Oo, Woon-Seng Gan and Wee-Tong Lim

Digital Signal Processing Lab, School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore

ABSTRACT

Nowadays, portable devices demand small-sized and low-end loudspeaker, however, the physical acoustic bass (or low-frequency) reproduction is usually poor. Bass enhancement by equalization is not feasible, and may even overload or damage the loudspeaker. Bass enhancement for such low-end loudspeaker can be psychoacoustically accomplished by exploiting the missing fundamental phenomenon. These systems are known as virtual bass systems (VBS), and generally use nonlinear device (NLD) as the main processing block to generate harmonics for virtual pitch perception. One of the recently developed NLD is the Arc-Tangent Square Root (ATSR) function, which can be used to control the pitch perception by a set of parameters. Mathematical relation between the input and output of the NLD is derived under a single-tone analysis framework. A detailed study on how the parameters of this NLD affect the harmonic’s decay pattern and intensity is presented in this paper.

Index Terms— Low-frequency signal processing, Bass perception, Portable audio system, Psychoacoustics.

1. INTRODUCTION

The problem with small-sized loudspeakers is that they cannot reproduce good bass due to their poor low-frequency responses. The virtual bass system (VBS) [1], [2] enhances such poorly reproduced bass by tricking human auditory system to perceive low-frequency. This psychoacoustics phenomenon is known as the "missing fundamental", which states that our auditory system can reconstruct the fundamental frequency from their higher harmonics, even if the fundamental frequency is missing [1], [2]. As the small-sized loudspeakers generally are better equipped to reproduce the higher harmonics that lie in the mid-band frequency, bass perception can be psychoacoustically enhanced by these harmonic components.

Nonlinear device (NLD) can be used in VBS to generate harmonics, which in turn creates the virtual pitch perception. There are different types of NLDs [1], [2], [3] available that are suitable to produce a set of harmonics that can be tuned for human perception. In particular, the static memoryless NLDs have been found to achieve good low-frequency perception [1].

Therefore, controlling the harmonics’ intensity by tuning the parameters of the NLD is of interest, and more importantly, how these parameters and nature of NLD can influence the psychoacoustic bass perception is the question that we are attempting to answer. Hence, we propose a particular type of static NLD called the Arc-Tangent Square Root (ATSR) function, which can generate both even and odd harmonics, and these harmonics can be easily controlled by a set of parameters. We derive the closed-form mathematical relation between input single tone amplitude and the output harmonics’ amplitudes with reference to the parameters of the ATSR. From these mathematical results, we further relate the harmonic characteristics to the existing psychoacoustic findings for pitch perception [2], and in turn, provide guidelines on how to tune the parameters of ATSR for VBS.

This paper is organized as follows. Section 2 presents the proposed VBS algorithm based on the ATSR function. Section 3 describes the detailed derivation of the closed-form harmonic analysis equations of the proposed NLD. Section 4 uses the derived analysis equations to study the effect of parameters on harmonics’ decay pattern and their intensities, followed by some discussions on preserving timbre while varying harmonics’ intensities. Section 5 concludes this paper.

2. PROPOSED NONLINEAR DEVICE

Our VBS algorithm follows a similar structure proposed by Larsen and Aarts [2], as shown in Fig.1. $x_L(t)$ and $x_R(t)$ denote the left and right channel of the stereo input signals. Since low-frequency signal generally do not consist of directional cues, these two signals are combined into a mono signal of $x_M(t) = 0.5[x_L(t) + x_R(t)]$ for low-frequency bass processing. Therefore, low-pass filter (LPF) is used to extract the low-frequency content and feeds into a nonlinear device based on the ATSR function to generate the necessary harmonics for the VBS.
The upper and lower branch of the VBS, representing the left and right channels are fed into a high-pass filter (HPF), and these channels are mixed with the harmonic generated in the middle branch to produce the psychoacoustics-enhanced stereo signal. The cutoff frequencies of these LPF and HPF are generally dependent on the loudspeakers’ cutoff frequency so as to customize the VBS for the loudspeakers.

In this paper, we propose a NLD that can generate both even and odd harmonics that are essential for bass perception. The proposed NLD has an acronym of ATSR and is defined as

\[
\text{ATSR} \triangleq f(x) = \alpha \tan^{-1}(\beta x) + \psi \sqrt{1 - (\zeta x)^2} - \psi ,
\]

(1)

where \(\alpha, \beta, \psi, \zeta\) are the four parameters that control the harmonics generation pattern and intensity. The mathematical relation between these four parameters and harmonic generation will be considered next.

By using Taylor’s series, (1) can be expressed as follows:

\[
f(x) = \alpha \beta \sum_{n=0}^{N_e} \left( \frac{(-1)^n}{2n+1} \right) x^{2n+1} + \psi \sum_{n=0}^{N_o} \frac{(-\zeta)^n (2n)!}{(1-2n)n!} 4^n x^{2n} - \psi .
\]

(2)

Limiting the highest order of polynomials to \(N\), the highest even order as \(N_e\), and the highest odd order as \(N_o\), which can be stated as

\[
\{N_e, N_o\} = \begin{cases} N_e = N, N_o = N_e - 1, & \text{if } \text{mod}(N, 2) = 0 \\ N_o = N, N_e = N_o - 1, & \text{if } \text{mod}(N, 2) = 1 \end{cases}.
\]

(3)

An estimated polynomial approximation of the ATSR is written as

\[
j(x) = \alpha \beta \sum_{n=0}^{N_e-1} \left( \frac{(-1)^n}{2n+1} \right) x^{2n+1} + \psi \sum_{n=0}^{N_o-1} \frac{(-\zeta)^n (2n)!}{(1-2n)n!} 4^n x^{2n} - \psi .
\]

(4)

The input-output transfer function plot is depicted in Fig. 2. \(f_j(x)\) and \(j_j(x)\) are the odd-symmetric and even-symmetric functions that generate odd and even harmonics, respectively. These odd and even harmonic blocks are

combined in parallel and subtract its combined result by a constant \(\psi\). The parameters \(\alpha\) and \(\beta\) are used to control the intensity and harmonic patterns of the odd harmonics; while \(\psi\) and \(\zeta\) are used to control the even harmonics.

\[
\text{ATSR is a non-homogeneous NLD, i.e.,} \quad f(\xi x) \neq \xi f(x) ,
\]

where \(\xi\) is a constant term. In other words, the operation of the NLD depends on the amplitude of the input signal. This implies that the ratio of the level of each harmonic component to its fundamental component changes when the input level varies. The harmonic components of a natural musical instrument correspond to the transient level of the input signal [3]. For the strong transient signal, the harmonics’ energies are stronger, and vice versa. Hence, polynomial NLDs mimic the natural musical instrument’s harmonic generation pattern, by creating stronger harmonics’ energy when the input level is strong, and vice versa.

In the next section, we present a closed-form mathematical relationship between the four parameters of the ATSR NLD, and its output harmonics’ amplitudes.

3. GENERALIZED HARMONIC ANALYSIS OF ATSR

To derive the generalized harmonic analysis equation for ATSR, let the input signal be

\[
x(t) = A \cos(2\pi f_o t) ,
\]

(6)

where \(A\) is the input single tone’s amplitude, \(f_o\) is the fundamental frequency, and \(t\) denotes the time index. The output signal of the ATSR function can be defined as

\[
y(t) = \frac{c_0}{2} + \sum_{k=1} c_k \cos(2\pi kf_o t) ,
\]

(7)
where \( c_0/2 \) is the DC component, \( c_i, k \in \mathbb{N} = \{1, 2, 3, \ldots\} \) are harmonics’ amplitudes respectively. By ignoring the time factor due to the memoryless nature of the NLD, the power series expression of the ATSR function is stated as
\[
f(x) = f_1(x) + f_2(x) - \psi = \sum_{i=0}^{\infty} h_i x^i,
\]
where \( h_i, i \in \mathbb{N} = \{0, 1, 2, 3, \ldots\} \) are power series coefficients.

Each and every individual harmonic’s amplitude, \( c_k \) can be related to the series coefficients, \( h_i \) and the input single tone amplitude, \( A \) by using the Schaefer-Suen equations \([4],[5]\) as follows
\[
c_k = \left( \frac{1}{2^j} \right)^j \sum_{j=0}^{\infty} A^{j+2j} \frac{h_{j+2j}}{2^{j(j+1)}} \binom{k+2j}{j},
\]
where
\[
\binom{n}{k} = \frac{n!}{k!(n-k)!},
\]
are the binomial coefficients. The output of the ATSR function in (1) can be modeled as
\[
y(t) = \sum_{l=22 \omega=1}^{N} d_l \cos(2\pi f_l t) + \frac{c_0}{2} + \sum_{m=2 \omega=1}^{N} e_m \cos(2\pi mf_l t) - \psi.
\]

By truncating summations in (10) to the highest order \( N \), (8) becomes
\[
\tilde{f}(x) = \tilde{f}_1(x) + \tilde{f}_2(x) - \psi = h_0 + h_1 x + \cdots + h_N x^N,
\]
where \( \tilde{f}_1(x) = h_0 + h_1 x^0 + \cdots + h_N x^N \), and \( \tilde{f}_2(x) = h_0 + h_1 x^2 + \cdots + h_N x^N \). Using (11), (10) becomes
\[
\tilde{y}(t) = \frac{\tilde{c}_0}{2} + \sum_{l=1}^{N} \tilde{c}_l \cos(2\pi f_l t) + \frac{\tilde{c}_0}{2} + \sum_{m=2}^{N} \tilde{e}_m \cos(2\pi mf_l t) - \psi
\]
\[
= \sum_{l=1}^{N} \tilde{d}_l \cos(2\pi f_l t) + \frac{\tilde{c}_0}{2} + \sum_{m=2}^{N} \tilde{e}_m \cos(2\pi mf_l t) - \psi
\]
\[
, l \in \{1, 2, 3, \ldots, N\}, t \in \{1, 3, \ldots, N_o\}, m \in \{2, 4, \ldots, N\}.
\]

Equation (12) is the output of the polynomial-approximated expression of the ATSR function, which consists of both even and odd harmonics. The DC magnitude is given as
\[
\text{DC} \approx \frac{\tilde{c}_0}{2} - \frac{\psi}{2}.
\]

Equations (12) consists of odd harmonics terms (up to \( N_o \)) in \( \tilde{y}_1(t) \), even harmonics terms (up to \( N_o \)) in \( \tilde{y}_2(t) \), and a DC term. Using Schaefer-Suen equation in (9), the generalized closed-form formulae for both even and odd harmonics’ amplitudes are
\[
\tilde{d}_l = \frac{\alpha \beta}{2} \sum_{j=0}^{N_o - 1} A^{j+2j} \binom{l+j+1}{j} \frac{\beta^{j+1}}{2^j(l+2j)} \binom{l+2j}{j}, l \in \{1, 3, \ldots, N_o\},
\]
\[
\tilde{c}_l = \frac{\psi}{2} \sum_{j=0}^{N_o - 1} (-A^j)^{l+2j} (m+2j)! (m+2j) \binom{m+2j}{j}, l \in \{0, 2, \ldots, N_o\},
\]
\[
m \in \{0, 2, \ldots, N_o\}.
\]

Equations (14) and (15) generalize the single-tone harmonic analysis of the ATSR function. It is noted that in (14), odd harmonics’ amplitudes, \( \tilde{d}_l \), are influenced by the two parameters of ATSR function, \( \alpha, \beta \), and the input single tone amplitude, \( A \). Likewise in (15), even harmonics’ amplitudes, \( \tilde{c}_l \), are influenced by \( \psi, \zeta \), and \( A \).

In the next section, we discuss how these parameters affect the harmonic generation and intensity, and their relation with psychoacoustic bass perception.

### 4. Simulation Results and Discussions

Using (14), (15), and letting \( N = 12 \), the output harmonics are set as \( \tilde{c}_1 = \tilde{c}_2 = \tilde{d}_2, \ldots, \tilde{c}_{N-1} = \tilde{c}_{N-1}, \tilde{c}_N = \tilde{d}_N \). If \( \tilde{c}_i < 0 \) , the output harmonics are 180 degree out-of-phase with respect to the input single tone in (6). If \( \tilde{c}_i > 0 \) , the output harmonics are at the same phase as the input signal tone. We use a total harmonic richness (THR) metric to quantify the amount of harmonic introduced into the original input signal. The THR is defined as
\[
\text{THR} = \sqrt{\frac{\tilde{c}_1^2 + \tilde{c}_2^2 + \cdots + \tilde{c}_{N-1}^2 + \tilde{c}_N^2}{\tilde{c}_1^2 + \tilde{c}_2^2 + \cdots + \tilde{c}_{N-1}^2 + \tilde{c}_N^2}} \times 100\%.
\]

where \( \tilde{c}_k \), \( k \in \{1, 2, \ldots, N\} \) are the root mean square values of the harmonics, which are computed as \( \tilde{c}_k = \sqrt{\tilde{c}_k^2} \).

To quantify the harmonics’ decay pattern, we define
\[
\Delta_{x,y} = 20 \log_{10} \left( \left| \tilde{c}_1 \right| \right) - 20 \log_{10} \left( \left| \tilde{c}_y \right| \right) = 20 \log_{10} \left( \frac{\tilde{c}_1}{\tilde{c}_y} \right),
\]
where \( \tilde{c}_1 \) and \( \tilde{c}_y \) are harmonics’ amplitudes of harmonic number \( x \) and \( y \) respectively. Hence \( \Delta_{x,y} \) is the absolute decibel level difference between the harmonic number \( x \) and \( y \). For the analysis purpose, the input is bounded between -1 and 1. To ensure Taylor’s series convergence, we set \( \beta x \leq 1, \zeta x \leq 1, \beta = 0.9 \) and \( \zeta = 0.9 \). We vary only \( \alpha \) and \( \psi \) from 1.3 to 3.1 with the step size of 0.1. The amplitude of the input tone is varied from -100 dB to 0 dB.

Figure 3 shows that ATSR is level-dependent. For large input signal, the harmonics’ energies are richer, whereas, for very small input signal, the harmonics’ amplitudes are insignificant. Hence, the harmonics’ richness is not significant for the very small input signal. This effect is desirable to mimic the naturalness of the musical
and for the left, \( E \), and the even harmonics is small, whereas for the large input level, the difference is large. Therefore, the results show that the harmonic decay rate is faster for small input signal, and decay rate is slower for large input signal.

From Fig. 5, the NLD variable parameters, \( \alpha \) and \( \psi \) have no influence on the decay rate at all, i.e., the level difference between harmonics is the same when the parameters are varied. However, we can control the harmonics’ intensity by varying the parameters, \( \alpha \) and \( \psi \) in (14) and (15), while preserving the decay rate. Moreover, we can separately tune \( \alpha \) or \( \psi \) or both to adjust the intensities of even and/or odd harmonics. This means that we can increase/decrease the harmonics’ energies while preserving the spectral envelope shape or preserving the timbre perception.

5. CONCLUSIONS

In this paper, we proposed the ATSR function to be used as the NLD in the VBS, and derived closed-form formulae to compute the individual harmonics’ amplitudes and phases. Using these formulae, simulations were performed to study the effect of varying the parameters of the ATSR function for psychoacoustic bass perception. From the above results, we conclude that we can increase or decrease the harmonic intensity by varying the ATSR’s parameters, without affecting the harmonics’ decay rate. From informal listening tests, the ATSR can preserve the timbre characteristics of any sound tracks, while altering the harmonic intensity. Without destroying the naturalness of the original musical instruments’ harmonic decay pattern [3], as mentioned in Section 2.

Figure 4 shows that the input level also has direct influence on the harmonics’ decay rate. For the small input level, the level difference between two subsequent odd/even harmonics is small, whereas for the large input level, the difference is large. Therefore, the results show that the harmonic decay rate is faster for small input signal, and decay rate is slower for large input signal.

From Fig. 5, the NLD variable parameters, \( \alpha \) and \( \psi \) have no influence on the decay rate at all, i.e., the level difference between harmonics is the same when the parameters are varied. However, we can control the harmonics’ intensity by varying the parameters, \( \alpha \) and \( \psi \) in (14) and (15), while preserving the decay rate. Moreover, we can separately tune \( \alpha \) or \( \psi \) or both to adjust the intensities of even and/or odd harmonics. This means that we can increase/decrease the harmonics’ energies while preserving the spectral envelope shape or preserving the timbre perception.

6. REFERENCES