Synthesis of Polynomial-Based Nonlinear Device and Harmonic Shifting Technique for Virtual Bass System

Wee-Tong Lim, Nay Oo and Woon-Seng Gan
Digital Signal Processing Lab, School of EEE
Nanyang Technological University
Singapore
Email: {limw0049, e070001, ewsgan}@ntu.edu.sg

***Abstract*— Low frequency bandwidth limitation is a common problem faced by miniature speakers and highly-directional speakers that have relatively high cut-off frequency. As such, low frequencies cannot be effectively reproduced. One of the current methods in addressing this problem is to employ psychoacoustic signal processing based on the "missing fundamental phenomenon". Nonlinear device is generally used to induce virtual pitch that enhances the low frequency perception. However, some of the difficulties in using nonlinear device include the need of precise adjustment of harmonics' magnitudes, harmonic order and its decay rate to achieve good perceived bass. In this paper, we propose two techniques on synthesis of polynomial-based nonlinear device and harmonic shifting by modulation in an attempt to overcome these difficulties. Real-Time implementation and listening tests were conducted to verify the effectiveness of the proposed algorithm.**

I. INTRODUCTION

Rapid evolution of technology advancements with the consumers' demand in extreme portable and low-cost devices over the years have led to the tremendous reduction in size of mobile devices, as well as portable speakers. An associated implication of these requirements is the size reduction of practically all components used in mobile phones such as the micro-speaker unit. As a result, these speakers usually have a much higher cut-off frequency compared to hi-end speakers used in home entertainment. This problem is also seen in highly-directional speaker, with cut-off frequency above 700 Hz [1]. As such, the perceived sound usually lacks bass.

Traditional method in addressing the poor-frequency problem is to simply boost up (or amplify) the weak signals. However, this method comes at an increase power consumption, distortion, and possibly speaker overload. A recent approach is to use psychoacoustic signal processing to enhance bass perception based on the "missing fundamental phenomenon", when fundamental frequency, f_0 can still be perceived from its associated set of harmonics even if it is absent. This is also known as virtual pitch or residue pitch [2].

The method of virtual pitch can generally be divided into two categories: time-domain method and frequency-domain method [3]. In this paper, a time-domain approach employing nonlinear devices (NLDs) is used to generate the required harmonics. A desirable characteristic of an NLD is the ability to accurately control the generated harmonics spectrum in terms of their (even and odd) order and magnitudes. The importance of these requirements was reported in [2] - [4]. Arora *et al* [2] reported three essential properties of a harmonic generator, namely, (i) the harmonic generator should be able to produce all (even and odd) harmonics, (ii) relative magnitudes of the harmonics should be controllable, and (iii) the method for harmonics generation should be low in complexity. Larsen *et al* [5] suggested that having stronger harmonics at lower frequency bands and weaker harmonics at increasing high frequency bands is beneficial in bass enhancement. Their findings also emphasized the need for accurate control over the magnitudes and order of the generated harmonics.

Some of the current reported NLDs used in Virtual Bass Systems (VBSs) [6] include half-wave and full-wave rectifier that generate only even order harmonics. Although, full-wave integrator can generate even and odd harmonics, the decay rate of harmonics is generally too slow to induce good bass perception [5]. Hence, these limitations lead to the motivation of exploring other techniques of harmonics generators which have better control over their harmonic spectrum.

This paper is organized as follows. Section II presents a proposed method that uses polynomial-generated harmonics and frequency shifting. Section III discusses the results from subjective evaluation and Section IV concludes this paper.

II. PROPOSED METHOD

Previously, we approximated static memoryless NLDs using polynomials [6]. Hence, with a set of appropriate coefficients, a polynomial-based NLD can be used to represent different NLDs. In this paper, we use the Chebyshev polynomial of the first kind relationship to control the magnitudes of the generated harmonics. Our work extends the generalized formula on polynomial analysis and synthesis, derived by R.A. Schaefer [7], into a closed-form equation for harmonic synthesis.

However, as some harmonics (depending on the fundamental frequency) may lie in the non-reproducible frequency band (see Fig. 1) of speaker, we attempt to counter this problem by splitting the low frequencies (0~770 Hz) into 3 filter bands and introduce a harmonic shifting technique to shift the affected subbands to the reproducible frequency band.

This research is supported by the Singapore National Research Foundation Interactive Digital Media R&D Program, under research grant **NRF2007IDM-IDM002-086**.

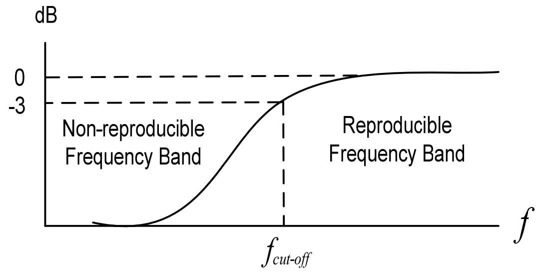

Figure 1. Typical frequency response of a speaker

A. Synthesis of polynomial-based nonlinear device

A polynomial-based NLD can be described as

$$y = h_1 x + h_2 x^2 + \ldots + h_n x^n \,, \tag{1}$$

where $h_1, h_2, \ldots, h_n$ are coefficients of the polynomials and n is the highest order of the polynomials. x and y denote the input and output signals, respectively. The polynomial-based NLD is a memoryless nonlinear system, i.e. the current output sample depends on the current input sample and superposition principle does not hold true. This system can be depicted as shown in Fig. 2 below.

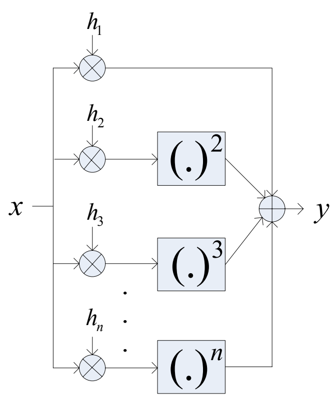

Figure 2. Block diagram of a Polynomial-based NLD.

The generated harmonics (by a single tone), including DC term and the fundamental frequency can be described as

$$y = 0.5c_0 + c_1 \cos\omega_0 + c_2 \cos 2\omega_0 + \ldots + c_n \cos n\omega_0 \,, \tag{2}$$

where c_0 is a DC term, and $c_1, \ldots, c_n$ are the magnitudes of the harmonics. In frequency domain, they can be viewed as shown in Fig. 3.

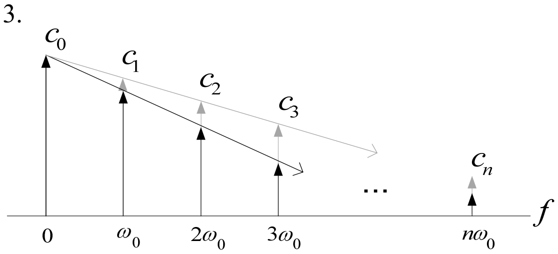

Figure 3. Magnitudes generated harmonics and DC-term of NLD in frequency domain.

The process of NLD analysis is defined as finding the magnitudes of the harmonics, given the polynomials' coefficients. Conversely, NLD synthesis obtains the polynomial coefficients from the magnitudes of the harmonics. R. A. Schaefer derived the generalized polynomial analysis and synthesis equations using Chebyshev polynomials of the first kind [7]. Schaefer's harmonics synthesis equation is given as

$$h_p = \frac{2^{p-1}}{p!} \sum_{i=0}^{\infty} (-1)^i c_{p+2i} (p+2i) \left[\frac{(p+i-1)!}{i!} \right], \tag{3}$$

where h_p are the polynomial coefficients and c_{p+2i} are the magnitudes of the harmonics.

Based on psychoacoustic research findings [8], [9], the first five harmonics are critical for bass perception. Therefore, in our analysis, we set $n = 5$. In order to generate up to the fifth order harmonics, the highest order of polynomials should be five, i.e. $h_p, p \in \{1,2,3,4,5\}$, are to be synthesized, given $c_q, q \in \{1,2,3,4,5\}$. Table I shows the derivations of the simple algebraic closed form formulas of polynomial coefficients from (3). Individual harmonic's magnitude can be set and the polynomial-based NLD's five coefficients can be synthesized easily.

TABLE I. DERIVATION TABLE

h_1	$= c_1 - 3c_3 + 5c_5$
h_2	$= 2c_2 - 8c_4$
h_3	$= 4c_3 - 5c_5$
h_4	$= 8c_4$
h_5	$= 16c_5$

However, a problem associated with the complexity and harmonic wastage may arise. Assuming the case of a 5th order polynomial-based NLD and a single tone with fundamental frequency of 100 Hz from band 1 (see Fig. 4), the synthesized harmonics of up to 5th order are 100, 200, 300, 400 and 500 Hz. In the case of a highly-directional speaker with cut-off frequency of 700 Hz, none of the harmonics can be reproduced. Higher harmonics may need to be reproduced, but with reduced impact on the bass perception.

Hence, our next proposed technique on harmonic shifting attempts to address this issue by shifting the synthesized harmonics in the non-reproducible frequency band to the reproducible frequency band.

B. Harmonics shifting using double side band suppressed carrier modulation scheme

Double sideband suppressed carrier (DSB-SC) modulation technique can be used to shift specific low-frequency subbands that lie in the non-reproducible frequency band of the speaker to the frequency of the carrier. The carrier frequency can be suitably chosen to lie in the reproducible frequency band of the speaker that faithfully reproduces the harmonics. The frequency spectrum produced by amplitude modulation is symmetrically spaced above and below the carrier frequency. One of the reasons for using DSB-SC modulation technique over Singe-Side Band (SSB) modulation is the lower complexity. Furthermore, subharmonics (lower side band) can also be produced that may enhance the bass perception [2].

Assuming the case of a single tone, f_0 with unity amplitude, the DSB-SC equation can be expressed as:

$$y(t) = \underbrace{V_m \cos(\omega_m t)}_{f_0} \times \underbrace{V_c \cos(\omega_c t)}_{f_c} \times 2$$

$$= \frac{V_m V_c}{2}\left[\cos\left((\omega_c + \omega_m)t\right) + \cos\left((\omega_c - \omega_m)t\right)\right] \times 2$$

$$= \cos((\omega_c + \omega_m)t) + \cos((\omega_c - \omega_m)t), \quad (4)$$

where $y(t)$ is the modulated spectrum, V_m is the amplitude of the single tone, V_c is the amplitude of unmodulated carrier, $\omega_m = 2\pi f_m$ is the single tone angular frequency in radians/s; and $\omega_c = 2\pi f_c$ is the unmodulated carrier angular frequency in radians/s [10]. Equation (4) has been multiplied by 2 in order to get unity output.

C. Case Study: Combining Polynomial Synthesis and Harmonic Shifting

This subsection illustrates an example in combining polynomial synthesis and harmonic shifting. A f_0 of 100 Hz from band 1 (see Fig. 4) is fed into a 5th order polynomial-based NLD with f_c of 700 Hz.

Coefficients for the polynomial-based NLD are computed based on the desired harmonics' magnitudes. The following presents a guideline for selection of the magnitudes. c_n can be set based on 2 known parameters: desired decay rate and magnitude of the first harmonic. Assuming $c_1 = 0.9$, a linear decay rate and the difference between 2 harmonics, Δ of 0.1 dBFS, the other 4 magnitudes can be computed based on (5) below:

$$\Delta = c_k - c_{k+1}, \quad (5)$$

where $c_k, k \in \{1,2,3,4,\}$. Table II shows the computed magnitudes.

TABLE II. COMPUTED MAGNITUDES

c_1	0.9
c_2	0.8
c_3	0.7
c_4	0.6
c_5	0.5

After application of the harmonic shifting technique, the synthesized harmonics (100, 200, 300, 400, and 500) Hz are now shifted to (800, 900, 1000, 1100 and 1200) Hz with additional harmonics (200, 300, 400, 500 and 600) Hz being generated as a result of the DSB-SC modulation. Assuming the case of a highly-directional speaker with cut-off frequency of 700 Hz, all the frequencies in the upper side band (above the carrier), lie in the reproducible region, while most frequencies in the lower side band (below the carrier) are greatly attenuated. However, these attenuated frequencies may still be perceivable to a certain extend. It is noted that the magnitude of these shifted harmonics (dependent on the frequency band) and the f_c (dependent on the speaker's cut-off frequency) must be carefully selected so as to produce desirable results.

Equation (6) below can be used to shape the harmonic spectrum in 2 steps if necessary. The first step is to select a suitable decay rate by adjusting α. $\alpha > 1$ results in a steeper decay curve while $\alpha < 1$ results in a more gradual decay curve. Once a suitable α is selected, step 2 is to select a suitable gain, G where $G > 1$ causes amplification while $0 \le G < 1$ leads to attenuation. The result is a new set of harmonics, $y_n, n \in \{1,2,3,4,5\}$. We can set $\alpha = 1$ and $G = 1$ to retain the initial set of c_n since it may not be necessary to further modify the spectrum for bands that do not require the process of harmonic shifting.

$$y_n = \begin{cases} G.c_n^{\ \alpha} & \alpha > 1 \\ G.c_n^{\ \alpha} & 0 < \alpha < 1 \\ G.c_n & \alpha = 1 \\ G & \alpha = 0 \end{cases}. \quad (6)$$

Fig. 4 below illustrates an example of a simplified VBS where frequencies below 770 Hz are split into filter bands and each is processed by a NLD. As most of the synthesized harmonics from band 1 fall into non-reproducible frequency band, DSB-SC modulation technique is used to shift frequencies in band 1 to the reproducible frequency band before mixing. This combined technique can further enhance the bass perception.

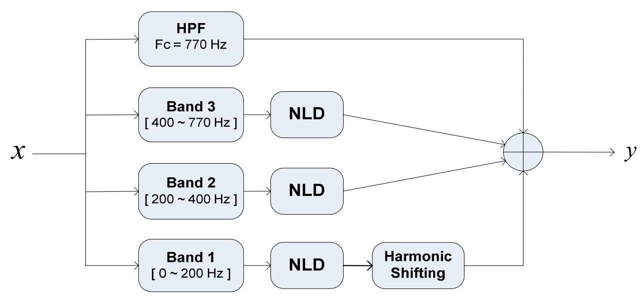

Figure 4. Block diagram of a simplfied VBS

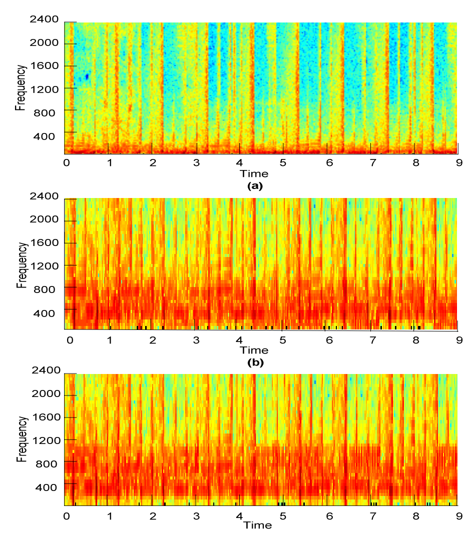

Figure 5. Spectrogram plots with (a) orginal input signal, (b) processed signal without modulation, and (c) processed signal with modulation.

Fig. 5 above depicts the spectrogram plots of (a) original input signal, (b) processed signal without modulation, and (c) processed signal with modulation. Clearly, it can be observed that both Fig. 5b and Fig. 5c showed an increased in energies across the frequencies (y-axis) illustrating the effects of the synthesized harmonics. On the other hand, Fig. 5c shows an increased in energies from 700 to 1100 Hz and a reduction of energies from 0 to 200 Hz signifying the effects of the shifted harmonics.

III. SUBJECTIVE EVALUATION AND RESULTS

Subjective testing was carried out to examine the effectiveness of the proposed method over the unprocessed sound track. The bass enhancement algorithm was implemented using Analog Device's BlackFin 533 Ez-Lite Kit. Appropriate values for the harmonics' magnitudes were selected. In order to demonstrate the problem of high cut-off, a highly-directional speaker [1] with cut-off of approximately 1000 Hz was used in the test. Twenty-four subjects aged between 20 to 40 years old participated in the tests. Arrangement of the speaker, position of subject and the conditions of listening room follows the ITU standard ITU-R BS.1116 [11]. Multi-stimuli with hidden reference and anchor (MUSHRA) of ITU-R BS.1534-1 [12] were used as the test procedure. Rating scales were based on ITU-T P.1534 continuous scale and ITU-R 5-point continuous impairment scale [13]. Three stimuli with appropriate amount of bass energies were used for our subjective tests. They are:

- Stimulus 1: Puff Daddy – I'll Be Missing You
- Stimulus 2: Britney Spears – I love Rock 'N' Roll
- Stimulus 3: Empire Brass Quintet – Hopper Dance

Two models are evaluated in the listening test. Model 1 (M1) uses the polynomial-based NLD without DSB-SC modulation and Model 2 (M2) uses the same polynomial-based NLD but with DSB-SC modulation technique. Data are compiled and plotted in a bar diagram. Figure 8 below contains the compiled scores for both models on two attributes, bass and impairment. Scores for anchor (high-passed signal) and reference signal are not included here. Scores were computed by taking the mean of the scores from 24 subjects.

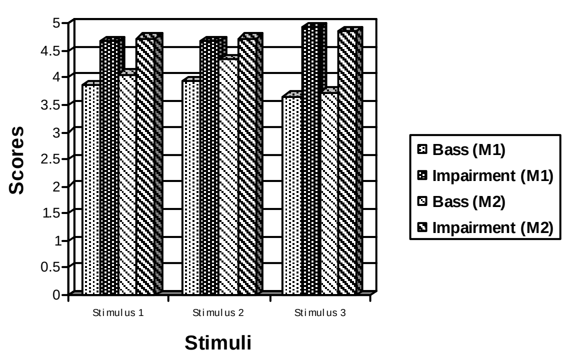

Figure 6. Listening test results

It is clear from Fig. 6 that both models showed a significant enhancement of bass perception especially on stimulus 2 with M1 and M2 having a score of 3.9 and 4.1, respectively. It can also be observed that impairment was generally imperceptible with a score of 4.5 and higher for each stimulus. Feedbacks from subjects mentioned the better quality in terms of naturalness and sensation of the processed sounds. Results also show that M2 employing DSB-SC modulation outperformed M1 with at least half a unit difference in all stimuli and a maximum of 1.2 in stimulus 2.

IV. CONCLUSION

This paper proposes two techniques on synthesis of polynomial-based NLD and DSB-SC modulation technique for harmonic shifting. The relation between coefficients of polynomial equation and magnitude of harmonics were discussed through presentation of the closed form algebraic equations. Next, a harmonic shifting technique employing DSB-SC modulation was used to shift the harmonics in the subbands that lie in the non-reproducible frequency band to the reproducible frequency band of the emitting device. This approach results in an increase in bass perception without increasing the computation complexity. Since we are able to have an accurate control over the magnitudes and order of the synthesized harmonics, a desired NLD can be designed. The proposed algorithm was implemented on a fixed-point embedded processor and applied to a highly-directional speaker with high cut-off frequency. Subjective evaluation was carried out to test the effectiveness of the algorithms and it was found that majority of the subjects preferred M2 which employed the harmonic shifting technique.

REFERENCES

[1] American Technology Corporation, "Hyper Sonic Sound H460 Product Sheet," September 2007.

[2] M. Arora, H.-G. Moon, and S. C. Jang, "Low Complexity Virtual Bass Enhancement Algorithm for Portable Multimedia Device," in AES 29th International Conference Seoul,Korea, September 2006.

[3] E. Larsens and R. M. Aarts, Audio Bandwidth Extension, Wiley, West Sussex, UK, September 2004.

[4] W. S. Gan, S. M. Kuo, and C. W. Toh, "Virtual bass for home entertainment, multimedia PC, game station and portable audio systems," IEEE Trans. Consumer Electronics, vol. 47, no. 4, pp. 787-794, November 2001.

[5] E. Larsen, and R. M. Aarts, "Reproducing low-pitched signals through small loudspeakers," J. Audio Eng. Soc., vol. 50, no. 3, pp. 147-164, March 2002.

[6] N. Oo and W. S. Gan, "Harmonic and Intermodulation Analysis of Nonlinear Devices used in Virtual Bass Systems," in AES 124th Convention, Amsterdam,The Netherlands, May 2008.

[7] R. A. Schaefer, "Electronic Musical Tone Production by Nonlinear Waveshaping," J.Audio Eng. Soc., vol. 18, pp. 413-417, August 1970.

[8] H. Dai, "On the relative influence of individual harmonics on pitch judgement", J. Acoust. Soc. Am., vol 107, no. 2, pp. 953-959, February 2000.

[9] B. C. J. Moore, B. R. Glasberg, and R. W. Peters, "Relative dominance of individual partials in determining the pitch of complex tones", J. Acoust. Soc. Am., vol. 77, no. 4, pp. 1853-1860, May 1995.

[10] S. Haykin and M. Moher, Introduction to Analog and Digital Communications, Wiley, 2007.

[11] ITU-R, "Recommendation BS.1116-1, Methods for subjective assessment of small impairments in audio systems including multi-channel sound systems," International Telecommunications Union Radiocommunication Assembly, 1997.

[12] ITU-R BS.1534-1, "Method for the Subjective Assessment of Intermediate Sound Quality (MUSHRA)," International Telecommunications Union, Geneva, Switzerland, 2001.

[13] ITU-R, "Recommendation BS.562-3, Subjective assessment of sound quality,"International Telecommunications Union Radiocommunication Assembly, 1990.