ICSV 21
13-17 July, 2014, Beijing/China

The 21st International Congress on Sound and Vibration

13-17 July, 2014, Beijing/China

A VARIABLE COEFFICIENTS ADAPTIVE IIR NOTCH FILTER FOR BASS ENHANCEMENT

Rui Zhu, Feiran Yang, Jun Yang

Key Laboratory of Noise and Vibration Research, Institute of Acoustics, Chinese Academy of Sciences, Beijing, China 100190

e-mail: rzhu89@gmail.com

This paper presents a new variable coefficients adaptive IIR notch filter for bass enchantment. The coefficients of the notch filter are controlled by an onset detector in order to obtain both fast convergence rate and low misadjustment. A single side-band (SSB) modulation is used to enhance the system robustness when the fundamental frequency is close to 0 Hz. We also propose a novel harmonic generation algorithm to generate correct harmonics in the fundamental frequency rapidly changing situation. The effectiveness of the proposed method is verified by both computer simulations and subjective listening test.

1. Introduction

Multimedia loudspeakers, widely used in the consumer portable electronic devices, have poor low-frequency reproduction ability, due to the limitation of small cell size and other physical conditions. Traditional methods such as equalizer boost the energy of bass directly. But this will increase energy consumption and reduce the portability.

To address this problem, virtual bass system (VBS) has been proposed based on the psycho-acoustic principles of missing fundamental. Missing fundamental states that human's ears can reconstruct the fundamental frequency from harmonics even when the fundamental frequency is not physically presented. This is also called virtual pitch theory.[1]

Phase vocoders (PV) and nonlinear devices (NLD) are commonly used for harmonic generation in VBS. PV is implemented in the frequency domain by using a phase-vocoder approach to shift the baseband to higher frequency bands.[2] PV suffers from the unnatural artifact when processing transient signal. The time-domain NLD method is efficient and has low complexity.[3] The drawback of NLD is the unwanted nonlinear distortion, especially intermodulation distortion. Recently, a hybrid method has been proposed which combines PV with NLD to avoid those drawbacks.[4] However, this approach has high computational load and introduces large input-output delay. Meanwhile, frequency tracking (FT) technique has been introduced into VBS.[5,6] Fundamental frequencies are detected and then its harmonics are generated. In this way, the nonlinear distortion can be reduced. The FT based VBS system can be implemented in the time domain with low computational load and small system delay.

This paper presents an improved VBS based on NLD and adaptive notch filter (ANF). We use a variable step-size ANF based on the onset detection to get both fast convergence rate and low misadjustment. Also, a single side-band (SSB) modulation is applied to the system before FT to improve the numerical stability. We also proposed a new harmonic generation method to get a better bass impression and less impairment. The new variable coefficients ANF (VC-ANF) has been shown to provide faster tracking speed by objective and subjective evaluations.

The paper is organized as follows. Section 2 gives a brief view of the conventional method. Section 3 describes proposed virtual bass system. Section 4 shows the results of some simulation experiments and the subjective test's result. Conclusions are presented in Section 5.

2. Conventional method

The VBS based on FT was first proposed by Larsen *et al*.[5] A recursive frequency-tracking algorithm use the equation

$$\hat{r}(n)=\hat{r}(n-1)+x(n-1)\gamma[x(n)+x(n-2)-2x(n-1)\hat{r}(n-1)], \tag{1}$$

where $x(n)$ is the input signal at time index n, and γ controls the convergence speed. The tracked frequency can be calculated by

$$\hat{r}(n)=\cos(\omega_0(n)T_s), \tag{2}$$

where $\omega_0(n)$ is the instantaneous frequency, and T_s is the sampling interval. The harmonic signal $x_h(n)$ is generated as

$$x_h(n)=\sum_{i=M}^{N}A_i\sin(i\omega_0(n)), \tag{3}$$

where M and N denote the minimum and maximum harmonic numbers, A_i is the amplitude of the i th harmonic. Though this method could prevent intermodulation distortion, it can only detect a single frequency. Subsequently, the cascade adaptive notch filters were used to track the fundamental frequency,[6] which can handle a baseband with more than one frequency. The second order lattice notch filter's transfer function is given by

$$H(z)=\frac{1+2k_0z^{-1}+z^{-2}}{1+(1+\alpha)k_0z^{-1}+\alpha z^{-2}}, \tag{4}$$

where k_0 is the adaptive coefficient and α is the pole-zero contraction factor. The adaptation algorithm can be summarized as follows

$$C(n)=\lambda C(n-1)+(1-\lambda)x(n-1)[x(n)+x(n-2)], \tag{5}$$

$$D(n)=\lambda D(n-1)+(1-\lambda)2x(n-1)^2, \tag{6}$$

$$\breve{k}_0(n)=-\frac{C(n)}{D(n)}, \tag{7}$$

where λ is a forgetting factor. The estimated fundamental frequency is

$$f_0(n)=\frac{1}{2\pi}\cos^{-1}\left(-\hat{k}_0(n)\right), \tag{8}$$

where $\hat{k}_0(n)$ is the optimized result of $\breve{k}_0(n)$.

3. Proposed virtual bass system

Figure 1 shows the block diagram of the proposed bass enhancement system based on FT. The input signal is firstly pre-processed by a low pass filter and down-sampling. The pre-processed signal is then processed by a SSB module. A VC-ANF based FT method is used to extract baseline signal. The coefficients of ANF are controlled by the onset detection. After the FT, the extracted baseline signal and the residual signal are processed by different HG modules to generate the harmonic components. At last, the generated harmonics are added to a delayed version of high frequency partial of the input signal.

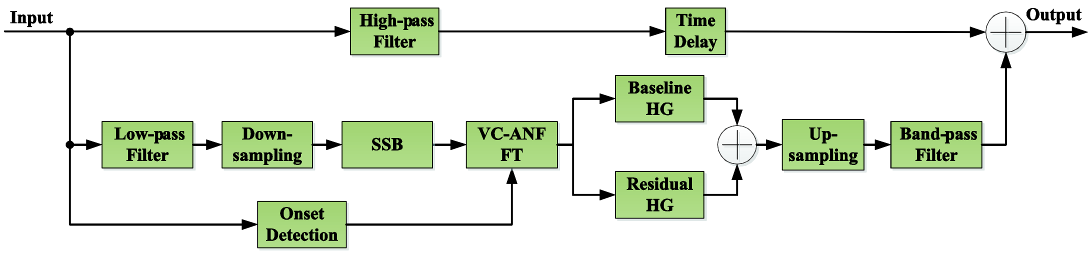


Figure 1. The block diagram of the proposed bass enhancement system.

3.1 Single-sideband modulation of baseband

It is pointed out that an unstable situation may happen when the notch frequency of ANF approaches its extreme values, i.e., 0 and $f_s/2$, where f_s is the sampling rate.[7] One solution to this problem was proposed by Regalia,[8] but it did not work for coloured input signal. We assume that $f_0(n)$ in Eq. (8) can be divided into two parts

$$f_0(n) = f_{cor}(n) + f_{err}(n), \tag{9}$$

where $f_{cor}(n)$ is the correct frequency and $f_{err}(n)$ is the predicted error. In practice, since $\hat{k}_0(n)$ in Eq. (8) is limited between ± 1, it can be deduced that

$$0 \le f_0(n) = f_{cor}(n) + f_{err}(n) \le \frac{f_s}{2}. \tag{10}$$

From Eq. (10) we see that when the input frequency $f_{cor}(n)$ approaches 0 or $f_s/2$, the allowance scale of $f_{err}(n)$ will be very small. While in a fix step-size adaptive algorithm, $f_{err}(n)$ can not be assured to satisfy the Eq. (10), especially in a low signal-to-noise ratio (SNR) situation. Traditionally, clipping and smoothing are used to ensure the correctness of Eq. (10),[9] but this method will bring instability and tracking error.

We propose a new method to address this problem. It can be seen that the allowance scale of $f_{err}(n)$ will be larger if $f_{cor}(n)$ moves closer to $f_s/4$. Thus, the adaptive algorithm will be more stable. A SSB modulator is used to modulate the baseband signal to higher frequency band before FT processing. The modulation frequency is determined by the sampling rate of input signal and down-sampling rate.

3.2 Variable coefficients ANF

In the proposed system, we use a variable coefficients adaptive notch filter to extract the fundamental frequency from the baseband signal. In order to get a fast convergence speed and small tracking error, the coefficients of ANF are controlled by an onset detector.

A time-domain energy-based onset detection method, namely "attack envelope method", is adopted to locate the position of the transient events.[10] This method is based on the fact that onset of the note always coincides with the start of the transient.[11] As long as the onset is detected, we can locate the start point of the transient. In transient part, the frequency content changes rapidly. To avoid error tracking, coefficients of ANF should be variable. The onset point is detected using

$$y(n) = \max_{m=0}^{m=M-1} \left\{ \left| x(nk+m) \right| \right\}, \tag{11}$$

$$P(n) = \max\left\{ y(n), P(n-1) \times k_{DECAY} \right\}, \tag{12}$$

where $x(k)$ is the input sample sequence, $y(n)$ is the local maximum of a block of input, $P(n)$ is the peak follower output, and k_{DECAY} is the decay factor. A onset is detected when

$$\frac{P(n)}{P(n-1)} > T_{D,AE}, \tag{13}$$

where $T_{D,AE}$ denotes the detection threshold.

There is a tradeoff between the convergence rate and steady-state mean square error (MSE) in a fixed step-size adaptive algorithm. Many variable step-size algorithms have been presented.[12] Most of them use the estimated error signal to control the step size. This method performs poorly when the SNR is low. Unfortunately most baseband signals have relatively low SNR. As we have got the onset position of the music notes, an elegant and efficient variable coefficients ANF can be used. In transient part of music piece, a large step size will be used to obtain fast convergence rate. While in steady state, the step size should be decreased to get low misadjustment.

In Eq. (7), the step-size is controlled by the forgetting factor λ. To avoid the distortion brought by the sudden change of λ, a modified sigmoid function can be used to smooth the change. The function is defined as

$$\lambda = 0.6 + 0.3 \cdot \left| \frac{2}{1+e^{-d}} - 1 \right|, \tag{14}$$

where d denotes the length of the detected transient part. The bandwidth of ANF is controlled by the pole-zero contraction factor α. During notch filtering period, α can be changed in the same way as λ to make residual signal mainly contain transient part of the input signal.

3.3 Harmonic generation

The harmonic signal $x_h(n)$ is usually generated by Eq. (3) as mentioned before. However, it is found that when the fundamental frequency changes rapidly over time, equation (3) can't generate harmonics of correct frequency. Equation (3) can be written in the continuous time domain

$$x_h(t) = \sum_{i=M}^{N} A_i \sin\left(i\omega_0(t)t \right), \tag{15}$$

where $\omega_0(t)$ is the instantaneous frequency of baseband signal. The instantaneous frequency of the i th harmonic can be calculated as

$$\omega_h(t) = \frac{d\Phi(t)}{dt} = i \cdot [\frac{d\omega_0(t)}{dt} \cdot t + \omega_0(t)] \neq i \cdot \omega_0(t). \tag{16}$$

It is clear that the i th harmonic with correct instantaneous frequency can't be calculated by Eq. (15), unless $\omega_0(t)$ is constant value. To get correct i th harmonic, the following condition must be satisfied:

$$\omega_i'(t)\cdot t+\omega_i(t)=i\cdot\omega_0(t), \tag{17}$$

where $\omega_i(t)$ denote the radian frequency of the correct i th harmonic. Transforming Eq. (17) to discrete time domain yields

$$[\omega_i(n)-\omega_i(n-1)]\cdot n+\omega_i(n)=i\cdot\omega_0(n). \tag{18}$$

Furthermore Eq. (18) can be written in an iteration form

$$\omega_i(n)=\frac{n}{n+1}\omega_i(n-1)+\frac{i}{n+1}\omega_0(n). \tag{19}$$

Once the instantaneous frequency is estimated, the harmonics of the baseline part can be generated by Eq. (3). The NLD method is adopted to generate the harmonics of transients.[4] The Arctangent Square Root function is used.[3] In this way, the high quality harmonics of baseline and residual can be produced.

4. Objective analysis and subjective experiment

In this section, some experiments are conducted to analyse the tracking and convergence property of the proposed method.

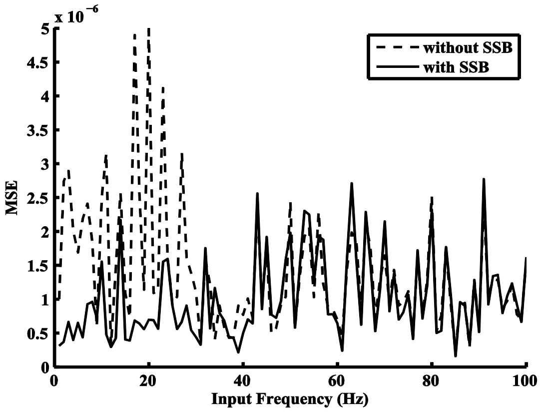


Figure 2. Comparison of MSE for different frequency with and without SSB.

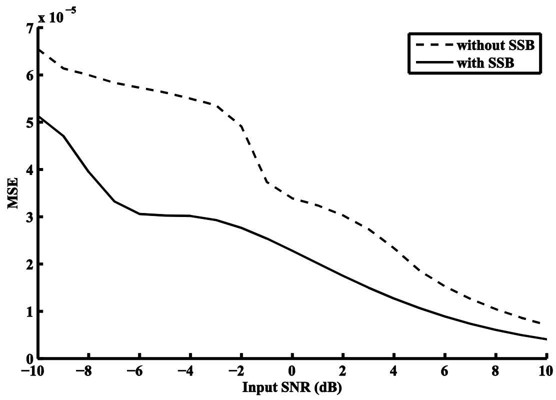


Figure 3. Comparison of MSE for different input SNRs with and without SSB.

The first simulation is conducted to evaluate the tracking behaviour with and without SSB for different input frequencies. The result is showed in Fig. 2. The input signal is a single sinusoid with frequency varies from 1 Hz to 100 Hz. The modulation frequency is set to 400 Hz. The other parameters follow the setting in.[6] We add a white noise to the input signal, and the MSE is calculated by averaging 5000 runs. It shows that the conventional method has a large MSE in low frequency part. After the SSB processing, the MSE has been decreased significantly. It also shows that SSB does not influence the MSE in higher part of baseband.

Another simulation is conducted to estimate the MSE under different SNRs. We add a coloured noise (pink noise) to the 40 Hz sinusoid signal. Figure 3 shows the result. The MSE of FT with SSB is smaller than that without SSB, and the difference becomes more significant as the SNR decreases.

In order to the analysis of VC-ANF's tracking property, a frequency hopping signal is used as input. Figure 4 compares the pitch tracking performance of the conventional method and the proposed method. The proposed method get faster convergence rate than the conventional method.

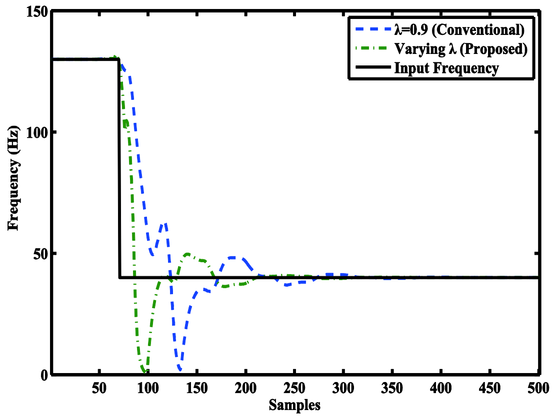
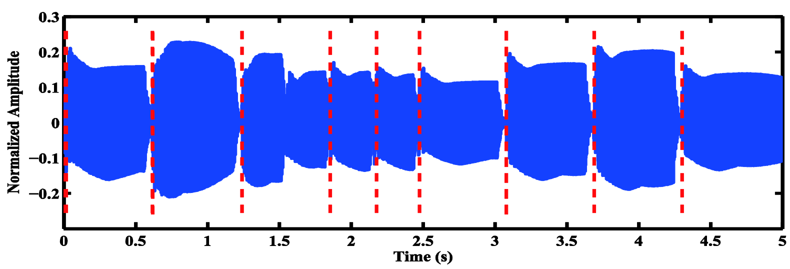
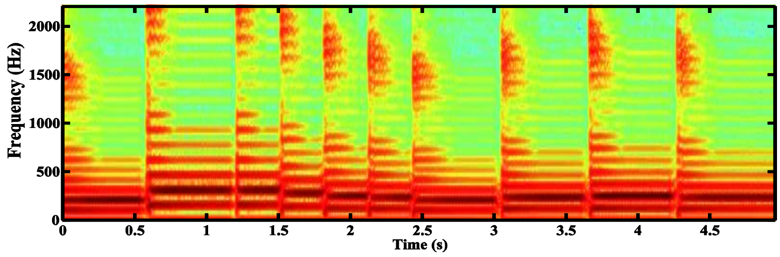


Figure 4. Comparison of tracking result under different forgetting factors.

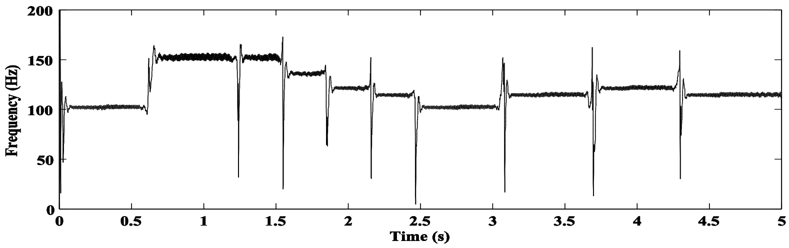
At last a real music piece is selected as experiment material. Results are given in Fig. 5. Fig. 5(a) and Fig. 5(b) show the input signal and its spectrogram, respectively. Fig. 5(c) and Fig. 5(d) show the results of the conventional and the proposed FT systems. The vertical lines of dashes in Fig. 5(a) denote the onset position of every single note. It can be seen that the proposed method has a faster convergence rate and a lower MSE level, thanks to the SSB modulation and variable forgetting factor.



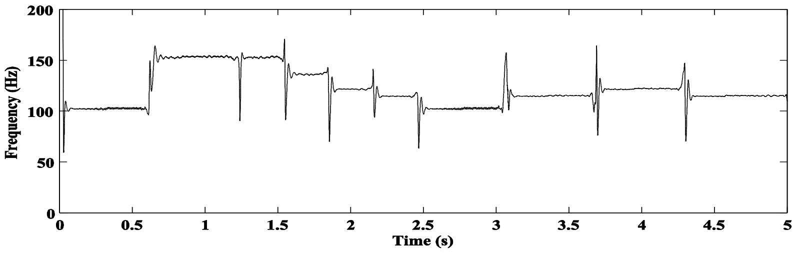
(a) Time domain waveform with onset detection



(b) The corresponding spectrogram



(c) Conventional FT method



(d) Proposed VC-ANF with SSB method

Figure 5. Simulation results of a piece of music

To assess the proposed method's performance, subjective listening experiment is conducted. Ten subjects participate in the experiment. We take multistimuli with hidden reference and anchor (MUSHRA) as the test procedure.[13] Subjects need to grade bass impression and audio quality. Four piece of pop music with duration of about 10 seconds are used as test materials. All of these music contain strong bass line. We adopt NLD, ANF and the proposed method to process the music, respectively. The NLD method's test materials are produced by Adam J. Hill's Virtual Bass Toolbox.[14] A high-pass filtered music without other enhancement is used as the anchor. It can be seen from Fig. 6 that the proposed system produces the strongest bass impression with smallest sound impairment.

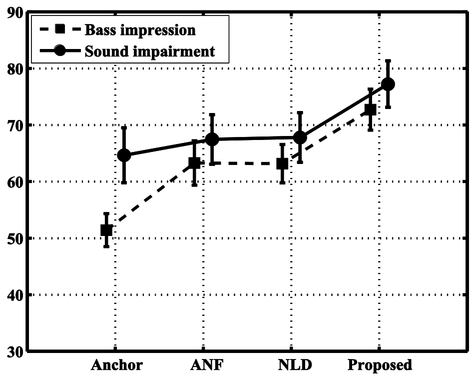


Figure 6. Subjective rating for bass enhancement systems

5. Conclusions

In this paper, we proposed an effective virtual bass enhancement system based on an improved ANF. Simulation results prove that improved ANF obtain both fast convergence rate and low misadjustment. The results of subjective tests show that new method achieves improved bass impression with reduced distortion.

6. Acknowledgements

This work was supported by Strategic Priority Research Program of the Chinese Academy of Sciences under Grants XDA06040501.

REFERENCES

1. Moore, B.C.J., *An Introduction to the Psychology of Hearing*, 4th edition, Academic Press, 177–213, (1997).
2. Bai M.R., Lin W.C., Synthesis and implementation of virtual bass system with a phase-vocoder approach, *The Journal of Audio Engineering Society*, **54** (11), 1077–1091, (2006).
3. Oo N., Gan W.S., Hawksford M.O.J., Perceptually-Motivated Objective Grading of Nonlinear Processing in Virtual-Bass Systems, *The Journal of Audio Engineering Society*, **59** (11), 804–824, (2011).
4. Hill A.J., Hawksford M.O.J., A hybrid virtual bass system for optimized steady-state and transient performance, *Computer Science and Electronic Engineering Conference(CEEC)*, 2010 2nd, (2010).
5. Larsen E.R., Aarts R.M., *Audio bandwidth extension: application of psychoacoustics, signal processing and loudspeaker design*, John Wiley & Sons, Hoboken, NJ, 53–115, (2004).
6. Junho L., Eunjung S., Youngcheol P., Daehee Y., Effective bass enhancement using second-order adaptive notch filter, *IEEE Transactions on Consumer Electronics*, **54** (2), 663–668, (2008).
7. Gil-Cacho P., Waterschoot T.V., Moonen M., Jensen S.H., Regularized adaptive notch filters for acoustic howling suppression, *Proceedings of the 17th European Signal Processing*, Glasgow, Scotland, UK, 24–28 August, (2009).
8. Regalia P., *Adaptive IIR filtering in signal processing and control*, Marcel Dekker, 270 Madison Avenue, NY, (1995).
9. Cho N.I., Lee S.U., On the adaptive lattice notch filter for the detection of sinusoids, *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, **40** (7), 405–416, (1993).
10. Masri P., *Computer Modeling of Sound for Transformation and Synthesis of Musical Signals*, Doctor of Philosophy Thesis, Department of Electrical and Electronic Engineering, University of Bristol, (1996).
11. Bello J.P., Daudet L., Abdallah S., Duxbury C., Davies M., Sandler M.B., A Tutorial on Onset Detection in Music Signals, *IEEE Transactions on Speech and Audio Processing*, **13** (5), 1035–1047, (2005).
12. Mader A., Puder H., Schmidt G.U., Step-size control for acoustic echo cancellation filters – an overview, *Signal Processing*, **80** (9), 1697–1719, (2000).
13. ITU-R BS.1534-1, Method for the Subjective Assessment of Intermediate Sound Quality (MUSHRA), International Telecommunications Union, Geneva, Switzerland, (2001).
14. Hill A.J., Virtual Bass Toolbox, [Online.] available: http://www.adamjhill.com/vb.html.