NON-LINEAR DISTORTION REDUCTION FOR A LOUDSPEAKER BASED ON RECURSIVE SOURCE EQUALIZATION

Hirofumi NAKAJIMA, Naoto SAKATA, and Kihiro HASHINO

Kogakuin University

ABSTRACT

This paper proposes a method of reducing nonlinear distortion in sound reproduction systems by recursively equalizing input signals. The proposed method has lower computational costs and modeling errors compared with conventional methods. A computer simulation was undertaken to confirm these advantages. An experiment for the proposed method was conducted using multi-loudspeakers as multi-sound sources. As a result, a reduction level of about 2 to 40 dB was achieved using the proposed method. In addition, it was found that the proposed method was particularly effective for frequency ranges lower than 8 kHz.

Index Terms**—** Non-linear distortion, Loudspeaker, Recursive equalization, Sound quality equalization

1. INTRODUCTION

Sound reproduction systems such as a loudspeakers and headphones may, depending on distortions in the system, reproduce sounds that are different from the input signal [1]. There are two kinds of distortions: "linear distortion" and "nonlinear distortion". Nonlinear distortion is a distortion that is linearly unrelated between inputs and outputs. Therefore, even if a sinusoidal signal is given as an input, a distorted sinusoidal signal with harmonic distortion would be obtained as an output. Reducing this distortion is desirable as it is a major cause of degradation of sound quality. Since it is impossible for harmonic distortion to be modeled using linear functions, it is more difficult to reduce nonlinear distortion than linear distortion. This paper proposes an algorithm for reducing nonlinear distortion by recursively equalizing the input signal and without modeling the nonlinearity of the system.

2. RELATED WORKS AND A PROPOSED APPROACH

Several methods of reducing nonlinear distortion of a loudspeaker have been proposed [2-10]. These methods are fundamentally based on the modeling and/or equalizing of nonlinear systems using Volterra filter [2-5], mirror filter [6-8] or control theory [9-10]. Since the computational costs of using the Volterra filter [2-4] increases exponentially according to the order of the nonlinearity, it is inadequate to reduce the higher order distortion. Although other methods [5-10] can reduce the computational costs, the overall performance is inferior to the Volterra filter due to the error in modeling the behavior of the loudspeaker approximately. Moreover, an adaptive method of reducing nonlinear distortion [11] has been reported recently, however, it requires a prior knowledge of the target system, therefore the issue of the modeling error remains.

In conventional methods, the main reason for the limited performance of the reduction of distortion is the modeling error of the system. To overcome this limitation, nonlinear systems are not modeled in this paper, but nonlinear distortion is reduced by measuring nonlinear distortion using a real system and adding a cancelling signal for the distortion. The proposed method has the advantage of being able to reduce distortion accurately with a small amount of computational cost. However, as the proposed method measures the distortion dependent on the sound source, it is necessary to reproduce the sound source and record the output at the same time for each sound source. Therefore, the proposed method can be applied in cases where that the sound source is known in advance. The proposed method is applicable for reproducing a limited kind of sound sources such as a toy that has a sound function, automated voice announcements or alert sounds at railway stations.

3. A PROPOSED SYSTEM MODEL AND PRINCIPLE

3.1. System model

With n as a discrete time index, the input signal as $x(n)$ and the output signal as $y(n)$, an acoustic system including nonlinear distortion is expressed as

$$\begin{aligned} y(n) &= \sum_{m=0}^{N-1} h(m)x(n-m) + e(n,\mathbf{x}) \\ &= h(n) * x(n) + e(n,\mathbf{x}), \end{aligned} \tag{1}$$

where $h(n)$ is the impulse response of the system (linear component), N is the length of the impulse response, $*$ stands for the convolution operation, $\mathbf{x}$ is to make $x(n)$ as the argument of the function, $e(n,\mathbf{x})$ is the nonlinearly-distorted signal dependent on $\mathbf{x}$. Transforming both sides of Eq. (1) by the FFT to the frequency domain,

$$Y(\omega) = X(\omega) \cdot H(\omega) + E(\omega,\mathbf{X}) \tag{2}$$

where $H(\omega)$ is the transfer function, ω is the discrete angular frequency, $\mathbf{X}$ is to make $X(\omega)$ as the argument of the function, $E(\omega,\mathbf{X})$ is the frequency response of the nonlinearly-distorted signal dependent on $X(\omega)$. Here, it is assumed that $E(\omega,\mathbf{X})$ can be modeled as a smooth continuous function with respect to $X(\omega)$.

3.2. The principle of the proposed method

The proposed method is composed of two processes: the extraction of the nonlinear distortion and the equalization of the sound source. First, the extraction of nonlinear distortion

978-1-4673-6997-8/15/$31.00

is explained. Nonlinear distortion $E(\omega,\mathbf{X})$ can be written as

$$E(\omega,\mathbf{X}) = Y(\omega) - X(\omega)\cdot H(\omega). \quad (3)$$

This means that nonlinear distortion $E(\omega,\mathbf{X})$ is obtained by subtracting the linear component $X(\omega)\cdot H(\omega)$ from the output of the reproduction system. $H(\omega)$ is obtained by taking the FFT for the impulse response $h(n)$ measured in advance. Similarly, $X(\omega)$ and $Y(\omega)$ are obtained by taking the FFT for $x(n)$ and $y(n)$. Second, the equalization of the sound source is explained. The equalized sound source $\tilde{X}(\omega)$ is calculated as

$$\tilde{X}(\omega) = X(\omega) - E(\omega,\mathbf{X})/H(\omega). \quad (4)$$

The equalized sound source in the time domain $\tilde{x}(n)$ is given by taking the IFFT for $\tilde{X}(\omega)$. The reduction for the nonlinear distortion can be proven by reproducing $\tilde{x}(n)$. The output signal $\tilde{Y}(\omega)$ is expressed as

$$\begin{aligned}\tilde{Y}(\omega) &= \tilde{X}(\omega)\cdot H(\omega) + E(\omega,\tilde{\mathbf{X}})\\ &= (X(\omega) - E(\omega,\mathbf{X})/H(\omega))\cdot H(\omega) + E(\omega,\tilde{\mathbf{X}})\\ &= X(\omega)\cdot H(\omega) + D(\omega,\tilde{\mathbf{X}},\mathbf{X}), \quad (5)\end{aligned}$$

where $D(\omega,\tilde{\mathbf{X}},\mathbf{X}) = E(\omega,\tilde{\mathbf{X}}) - E(\omega,\mathbf{X})$. If the energy of the nonlinear distortion is much smaller $\|E(\omega)\|_2 \ll \|X(\omega)H(\omega)\|_2$, $\|\tilde{X}(\omega) - X(\omega)\|_2 \ll \|X(\omega)\|_2$ is given. Therefore, $E(\omega,\mathbf{X})$ becomes very close to $E(\omega,\tilde{\mathbf{X}})$. $\|D(\omega,\tilde{\mathbf{X}},\mathbf{X})\|_2 \ll \|E(\omega,\mathbf{X})\|_2$ should be given.
Here, $\|A(\omega)\|_2 = \sum_\omega |A(\omega)|^2$.

3.3. Recursive source signal equalization

Including the recursive process, nonlinear distortion can be further reduced. In the following explanation, the argument ω is omitted for simplicity. The 2nd equalized sound source X is calculated so that the 1st equalized distortion $D(\tilde{\mathbf{X}},\mathbf{X})$ is reduced, such as

$$\tilde{\tilde{X}} = \tilde{X} - D(\tilde{\mathbf{X}},\mathbf{X})/H. \quad (6)$$

The output $\tilde{\tilde{Y}}$ for the 2nd equalized sound source $\tilde{\tilde{X}}$ is described as

$$\begin{aligned}\tilde{\tilde{Y}} &= \tilde{\tilde{X}}H + E\left(\tilde{\tilde{\mathbf{X}}}\right) = \tilde{X}H - D(\tilde{\mathbf{X}},\mathbf{X}) + E\left(\tilde{\tilde{\mathbf{X}}}\right)\\ &= XH - E(\mathbf{X}) - E(\tilde{\mathbf{X}}) + E(\mathbf{X}) + E\left(\tilde{\tilde{\mathbf{X}}}\right)\\ &= XH + D\left(\tilde{\tilde{\mathbf{X}}},\tilde{\mathbf{X}}\right) \quad (7)\end{aligned}$$

Similarly, the 3rd equalized sound source $\tilde{\tilde{\tilde{X}}}$ is calculated so that the 2nd equalized distortion $D\left(\tilde{\tilde{\mathbf{X}}},\tilde{\mathbf{X}}\right)$ is reduced. If the 1st, 2nd, 3rd input signals (original sound source, and the equalized sound sources) $X,\tilde{X},\tilde{\tilde{X}}$ are redefined as X_0, X_1, X_2, and if the 1st, 2nd, 3rd output signal (original output, and the equalized outputs) $Y,\tilde{Y},\tilde{\tilde{Y}}$ are redefined as Y_0, Y_1, Y_2, the generalized relationship can be expressed as

$$X_L = X_{L-1} - D(\mathbf{X}_{L-1},\mathbf{X}_{L-2})/H \quad (8)$$

$$Y_L = X_0 H + D(\mathbf{X}_L,\mathbf{X}_{L-1}) \quad (9)$$

where $L(L=1,2,\cdots)$ is the number of the equalization, $X_{-1}=0$ and $E(0)=0$. Theoretically, the recursive equalization can reduce the nonlinear distortion. The details of the proof are omitted due to limitations of space.

4. A FLOW OF THE PROPOSED METHOD

The proposed method has two processes: "Initialization process (STEP 1-3)" and "Recursive process (STEP 3-4)".

Step. 1 Calculation of the transfer function
The transfer function $H(\omega)$, which is the model of the linear component, is obtained by measuring the impulse response $h(n)$ and taking the FFT for the response. Several methods [12][13] can be applied to measure the linear response. In this paper, the Warped-TSP method [12] is employed in order to separate the linear component from the harmonic distortion.

Step. 2 Calculation of the ideal output
In order to reduce the computational cost, the output $Y_{Ideal}(\omega)$ of the ideal system only having the linear component is calculated in advance as $Y_{Ideal}(\omega) = X(\omega)\cdot H(\omega)$. $H(\omega)$ is obtained at Step 1, and $X(\omega)$ is obtained by taking the FFT for the original sound source $x(n)$.

Step. 3 Design of the inverse filter
Eq. (4) would be diverged if $H(\omega)$ is close to 0. To prevent this, the inverse filter $H^{-1}(\omega)$ is designed by the following equation as

$$H^{-1}(\omega) = \frac{H^*(\omega)}{|H(\omega)|^2 + a}, \quad (10)$$

where * stands for the complex conjugate, a is the regularization constant to prevent the divergence, which is

$$a = \max\left\{\left|\tilde{H}(\omega)\right|^2\right\}\cdot r, \quad (11)$$

where $\max\{x\}$ is the maximum value of x, r is a coefficient to decide the ratio for x.

Step. 4 The extraction of the nonlinear distortion
While the sound source $x(n)$ is reproduced, $y(n)$ is recorded at the same time. The nonlinear distortion $E(\omega,\mathbf{X})$ can be calculated as

$$E(\omega,\mathbf{X}) = Y(\omega) - Y_{Ideal}(\omega) \quad (12)$$

where $X(\omega)$ and $Y(\omega)$ are obtained by taking the FFT of $x(n)$ and $y(n)$ respectively. $H(\omega)$ and $Y_{Ideal}(\omega)$ are given from Step.1 and Step.2, respectively.

Step. 5 Equalization of the input signal
Using the inverse filter $H^{-1}(\omega)$ obtained at Step.3 and the nonlinear distortion $E(\omega,\mathrm{X})$ obtained at Step.4, the equalized sound source $\tilde{X}(\omega)$ is calculated as

$$\tilde{X}(\omega) = X(\omega) - E(\omega,\mathbf{X})\cdot H^{-1}(\omega) \quad (13)$$

The equalized sound source in the time domain $\tilde{x}(n)$ is given by taking the IFFT of $\tilde{X}(\omega)$. The recursive process can be iteratively performed from Step.4 using the equalized sound source $\tilde{x}(n)$ as a new $x(n)$.

5. A PROCESS IN THE TIME DOMAIN

If it is difficult to take the FFT for a long period of the sound source at one time, a process in the time domain is available. The process in the time domain can be realized by replacing Eq. (12) and (13) with the following equation, such as

$$e(n,\mathbf{X}) = y(n) - y_{Ideal}(n) \quad (14)$$

$$\tilde{x}(n) = x(n) - e(n, \mathbf{X}) * h^{-1}(n) \quad (15)$$

where $y_{Ideal}(n)$ is calculated by the convolution $y_{Ideal}(n) = x(n) * h(n)$, and the inverse filter $h^{-1}(n)$ is obtained by taking the IFFT of $H^{-1}(\omega)$ in Eq. (12). Note that the circular shift is required since the inverse filter $h^{-1}(n)$ has an noncausal component. Computational cost for the convolution is drastically saved using the FFT based on "Overlap-add" or "Overlap-save" technique [14].

6. EVALUATION BY A COMPUTER SIMULATION

6.1. Verification experiment for the principle

A computer simulation was conducted in order to verify if the proposed method could work properly. The sampling frequency of the system was 44.1 kHz. The output signal of the system $y(n)$ is numerically simulated for the input signal $x(n)$ as

$$y(n) = f(x(n)) * h(n) + w(n) \quad (16)$$

where $f(x)$ is a nonlinear function given as

$$f(x) = x + 0.2x^2 + 0.5x^3 + 0.1x^4 + 0.25x^5. \quad (17)$$

$h(n)$ is the measured impulse response shown in Fig. 1. $w(n)$ is an external noise that is defined as 0 in this chapter. The four sound sources $x(n)$ (10 seconds) are prepared, (1) 1 kHz sinusoidal signal, (2) 1 kHz rectangular signal, (3) music (a vocal), (4) white noise. The sound source (3) was prepared from a part of No. 17 in RWC-DB [15]. The transfer function when the proposed method is executed was given as the signal shown in Fig. 1. The coefficient of the inverse filter r was 0, and the number of the iteration was 10 times. Fig. 2 shows the nonlinear distortion level EL for each sound source, and E_L was $E_L = 10\log_{10}\sum_{\omega}|E(\omega, \mathbf{X})|^2$. According to the figure, the distortion for all sound sources decreased largely. This means that the proposed method is effective in principle.

6.2. A comparison of the accuracy between the proposed method and the conventional method

The performance of the proposed method was compared with that of the conventional method (Volterra filter). Although the condition in this experiment was the same as the condition in the experiment 5.1, $w(n)$ was a white noise whose SNR was 60 dB, and $h(n)$ was changed to the delta function for simplicity. In addition, 3 types of the nonlinear system were prepared, (a) Eq. (17), (b) Eq. (17) but truncated at the 3rd order of Eq. (17), (c) the following function (the order was infinite) as

$$f(x) = \tan^{-1} x \quad (18)$$

$$\approx x - \frac{1}{3}x^3 + \frac{1}{5}x^5 - \frac{1}{7}x^7 \dots \quad (19)$$

In the conventional method, it was assumed that the 2nd or 3rd order coefficients of the nonlinear filter could be estimated without any error. Practically, the coefficients of Eq. (17) or (19) were utilized. Fig. 3 shows the amount of the distortion reduction for each method in the nonlinear order of the system. Inf in the figure corresponds to the nonlinear function by Eq. (18), and P is the order of Volterra filter. It can be observed that the proposed method showed the better performance than the conventional method in the case of higher nonlinearity system ($P > 3$).

6.3. A comparison of the computational cost between the proposed method and the conventional method

Fig. 4 shows the number of multiplications that is required in the process with respect to the length of the filter. M in the figure means the number of iterations in the proposed method. The number of multiplications was defined as the number required in the distortion reduction, excluding the number required in the initialization. The sampling frequency of sound sources was set as 44.1 kHz and the data length was 60 seconds. According to the figure, the number of multiplications in the proposed method was less than that in the conventional method. The difference was particularly remarkable when the filter length was longer. Since the number of multiplications is almost proportional to the required memory size, the proposed method has an advantage in terms of the memory size.

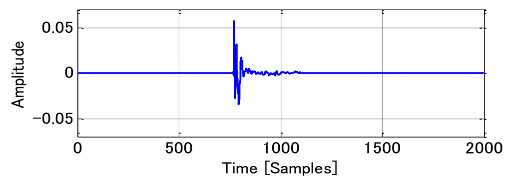


Fig. 1. Impulse response for numerical simulation

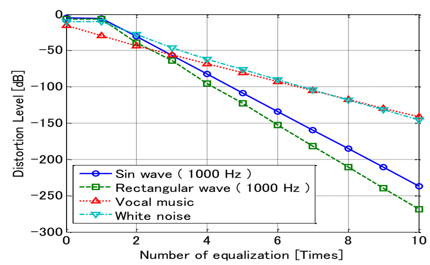


Fig. 2. Distortion level vs. number of iterates

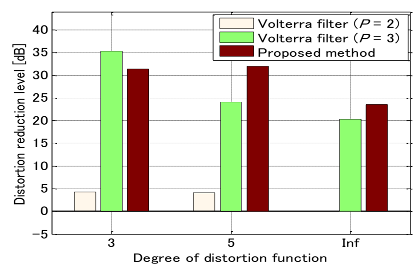


Fig. 3. Distortion reduction level

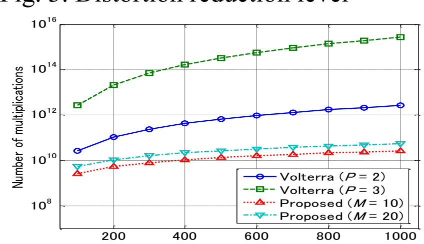


Fig. 4. Required number of multipliers

7. EVALUATION USING SETS OF LOUDSPEAKERS

7.1. Conditions of the experiment

Fig. 5 shows the environment of the experiment in an anechoic room. The distance between the loudspeaker and the microphone was 150 mm. The equipment used in the experiment was an audio interface (M-Audio Fast TrackPro), a precision sound level meter used as a microphone (Onosokki LA-4440), a PC for analytical use (Apple MacBook Air in Fig. 6) and three loudspeakers (A: Elecom MS-P01WH, B: Creative GigaWorks T20-II, C: Yamaha MSP5 STUDIO). The sound sources were three sinusoidal signals (3 seconds of 0.5, 1, 2 kHz) and some part of 3 music data in RWC-DB [15] (15 seconds of No.10: Heavy Metal, No.21: Techno , No.56: Classic). The sampling frequency was 44.1 kHz and the number of the equalizations was 10 times. The filter in the sound level meter was set as A-weighting. The output level of the loudspeaker was adjusted so that 1 kHz sinusoidal signal gave 85 dB at the measuring point.

7.2 Results and discussion

Fig. 8 shows the distortion level E_L (dB) for the number of the equalization. E_L was expressed by the ratio of the distortion in the output using the original output level as a reference, such as

$$\tilde{E}_L = 10\log_{10}\frac{\|E_L(\omega)\|_2\|Y_0(\omega)\|_2}{\|Y_L(\omega)\|_2} \qquad (20)$$

Loudspeaker B was used here. It was found that the distortion was consistently reduced for every sinusoidal signals as shown in Fig. 8(a). Remarkably, a reduction level of about 24 dB was achieved in 2 kHz sinusoidal signal. Similarly, it was observed that there were some cases of the distortion level increasing after it was reduced once, as shown in Fig. 8(b). The maximum levels of the distortion reduction were 9.9 dB for Classic, 7.3 dB for Heavy Metal and 1.7 dB for Techno. Fig. 9 shows the distortion spectra of 2 kHz sinusoidal signal as the best performance and Techno music as the worst performance. The upper graph is "before the equalization" and the lower graph is "after 10 times equalization". According to Fig. 9(a), although the distortions at 4, 6, 8 kHz was largely reduced, the distortions at 12, 16, 20 kHz increased. The peaks at 2 kHz in Fig. 9(a) are caused by the amplitude error of the source signal, which are regarded as nonlinear distortion in this algorithm. As the level of the original source was 85 dB, the levels of the peaks about 65 and 40 dB are much smaller than the level of the original source. As for Fig. 9(b), although the distortion lower than 8 kHz was reduced about 5-15 dB, the distortion higher than 12 kHz increased. The reason for the increasing distortion at the higher frequency could be the low accuracy of the inverse filter due to the larger fluctuation of the transfer function affected by the short wavelength. All levels of reduction for every sound source and loudspeaker are summarized in Fig. 10. It can be said that the proposed method is effective for any loudspeaker and the reduction levels for the nonlinear distortion ranged from 2-40 dB. Subjective impressions by the authors are as follows: It was obvious that all sinusoidal waves became milder pure tone by the equalization (less harmonics). For Classic, the instruments having some low frequency such as base strings sounded clearer by the equalization. For Techno and Heavy metal, it was hard to recognize the difference by the equalization except for the slight emphasis in the higher frequency.

8. CONCLUSION AND FUTURE WORK

This paper proposed a method of reducing nonlinear distortion in sound reproduction systems by recursively equalizing input signals. As a result of using several sound sources, levels of reduction in nonlinear distortion ranging 2-40 dB were achieved with little computational cost. The proposed method was especially effective at frequencies lower than 8 kHz. To prevent any feedbacks of noise into the equalized input signals, noise-free environment is required. This issue in the feedbacks of noise should be mentioned in a part of the future work. In addition, utilizing this proposed method to the loudspeaker in a smart-phone would also be an important application with an eye to an industrial prospect.

Fig. 5. Experimental setting

Fig. 6. Laptop

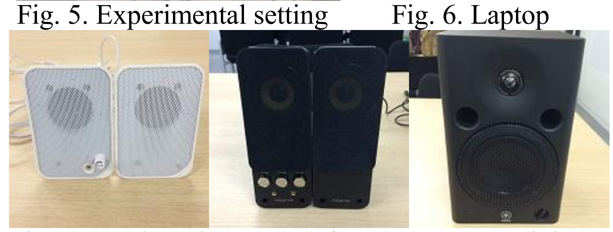


Fig. 7. Loudspeakers (A: Left, B: Center, C: Right)

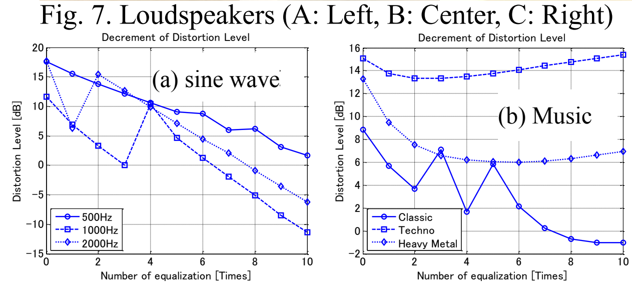
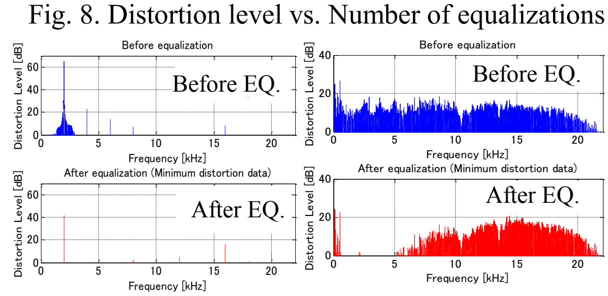


Fig. 8. Distortion level vs. Number of equalizations



(a) 2kHz sine (b) Techno music

Fig. 9. Amplitude spectrum with/without equalization

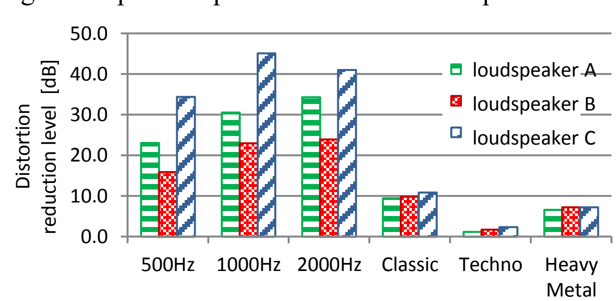


Fig. 10. Final summary of distortion reduction level

9. REFERENCES

[1] W. Klippel, “Loudspeaker nonlinearitics - Causes, prameters, symptoms,” *J. Audio Eng. Soc.*, 54, pp. 907-939, 2006.
[2] F. Rumsey, “DSP in loudspeakers,” *J. Audio Eng. Soc.*, 56, pp. 65-72, 2008.
[3] Y. Kajikawa, “Linearization Method Based on Multiple Loudspeaker Systems,” Acoustical Science and Technology, vol. 32, no. 5, pp. 220-223 , 2011.
[4] V.J. Mathews, “Adaptive polynomial filters,” *IEEE Signal Process. Mag.*, 8, pp. 10-26, 1991.
[5] P. M. S. Burtand and J. H. M. Goulart , “Evaluating the potential of Volterra-PARAFAC IIR models,” ICASSP2013, pp. 5745-5749, 2013.
[6] W.A. Frank, “An efficient approximation in realtime loudspeaker linearization,” Signal Process., 45, pp. 97-113, 1995.
[7] W. Klippel, “The Mirror filter-A new basis for reducing nonlinear distortion and equalizing response in woofer systems,” *J. Audio Eng. Soc.*, 40, pp. 675-691, 1992.
[8] K. Iwai and Y. Kazikawa, “Linearization Of Dynamic Loudspeaker System Using Third-Order Nonlinear IIR Filter,” EUSIPCO2012, pp. 1970-1974, 2012.
[9] J. Suykens, et. al., “Feedback Linearization of Nonlinear Distortion in Electrodynamic Loudspeakers,” JAES, vol. 43, 9, pp. 690-694, 1995.
[10] D. Franken, et. al., “Observer-based feedback linearization of dynamic loudspeakers with Ac amplifiers,” IEEE T-SAP, vol. 13, no. 2, pp. 233-242, 2005.
[11] M. Hotz and C. Vogel, “Linearization of Time-Varying Nonlinear Systems Using A Modified Linear Iterative Method,” IEEE Trans on Signal Processing, vol. 62, no. 10, 2014.
[12] M. Morise, et. al., “Warped-TSP: An acoustic measurement signal robust to background noise and harmonic distortion,” Electro. and Comm. in Japan, III, vol. 90, no. 4, pp. 18-26, 2007.
[13] F. Heinle, et. al., “A measurement method for the linear and nonlinear properties of electro-acoustic transmission systems,” Signal Processing, 64, pp. 49-60, 1998.
[14] A. V. Oppenheim and R. W. Schafer, ”Discrete-Time Signal Processing,” Prentice Hall, 1989.
[15] M. Goto, et. al., “RWC Music Database: Popular, Classical, and Jazz Music Databases,” Proc. of the 3rd Int'l Conf. on Music Information Retrieval (ISMIR 2002), pp.287-288, 2002.