A PSYCHOACOUSTICAL PREPROCESSING TECHNIQUE FOR VIRTUAL BASS ENHANCEMENT OF THE PARAMETRIC LOUDSPEAKER

Chuang Shi, Hao Mu, and Woon-Seng Gan

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

ABSTRACT

The parametric loudspeaker is a novel type of loudspeaker that can project a directional sound beam. It is commonly used in creating personal sound zone and projecting private messages to a targeted audience. However, the parametric loudspeaker possesses a very poor bass (or low-frequency) response due inherently to the nonlinear acoustic principle generating sound from ultrasound in air. A psychoacoustic signal processing method known as "virtual bass" has been successfully implemented in some consumer electronics with miniature or flat loudspeaker unit, aiming to enhance their bass performances. In this paper, we adapt this "virtual bass" approach for parametric loudspeakers. Unlike conventional loudspeakers, the parametric loudspeaker brings in an added degree of complexity in "virtual bass" enhancement due to its inherent nonlinear acoustic property. Accordingly, a new preprocessing technique is proposed for the parametric loudspeaker to psychoacoustically reproduce the low-frequency components within an octave below its cut-off frequency.

Index Terms— Psychoacoustics, nonlinear acoustics, digital signal processing, nonlinear distortion

1. INTRODUCTION

The parametric loudspeaker can generate a super-directional sound beam by utilizing such a parametric array effect in air [1]. When two primary waves at 40 kHz and 42 kHz travel on the same path, the difference-frequency wave at 2 kHz is gradually generated due to the nonlinearity of air [2]. Since the 2 kHz wave is a result of the interaction between the two primary waves, it inherits their sharp directivities. Thus, an ultrasonic wave beyond human hearing range is used in the parametric loudspeaker to carry audible sounds to a targeted location. Due to its sharp and controllable radiation pattern [3, 4], the parametric loudspeaker can be applied in various areas, such as digital signage, privacy messaging, personal announcement [5], active noise control [6], interactive sound visualization [7], and even landmine detection [8]. Recently, parametric loudspeakers are also found to be advantageous in reproducing immersive 3D soundscapes for 3D gaming and entertainment [9].

However, there are two major physical limitations of the parametric loudspeaker due to its sound generation principle [2]. Firstly, during the self-demodulation process caused by the parametric array effect, higher harmonic components of the original sound are generated as by-products. Secondly, this self-demodulation process shows a high-pass filtering effect, resulting in a very poor bass quality of the parametric loudspeaker. Research over the last two decades [3-11] have mainly been focused on reducing harmonic distortions using different preprocessing techniques and controlling the beam patterns of the parametric loudspeaker.

Till now, there have been only two studies carried out in addressing the bass reproduction limitation of the parametric loudspeaker. Croft and Norris [12] stated that the parametric loudspeaker's poor bass quality is physically unsolvable. In order to boost up the low-frequency component levels of the reproduced sound, the total energy output of the parametric loudspeaker has to be increased significantly, which results in less conversion efficiency because of the saturation in air. Alternatively, a psychoacoustic phenomenon was previously investigated for the parametric loudspeaker [13]. It is known as the “missing fundamental” [14], which states that human auditory system can perceive the fundamental frequency merely from its higher harmonics. For instance, by hearing a harmonic series of 200, 300 and 400 Hz, human brain can perceptually obtain the sensation of the common difference (fundamental) frequency at 100 Hz. Accordingly, harmonic generators, which are adapted from nonlinear functions, are studied to generate the harmonic series in the psychoacoustic bass enhancement [15-17]. In Figure 1, the audio input is split into two bands through a pair of low-pass and high-pass filters. The lower band is processed by a harmonic generator and the artificially generated higher harmonics of the low-frequency components create the virtual bass enhancement. In comparison to the physical method that simply boosts the low-frequency components, the psychoacoustic enhancement has an advantage in preventing overload and damage to the excitation unit of the loudspeaker.

The psychoacoustic bass enhancement can be applied to the parametric loudspeaker as well. In Figure 2, the sum of the preprocessed lower band and the higher band is amplitude modulated with an ultrasonic carrier before being

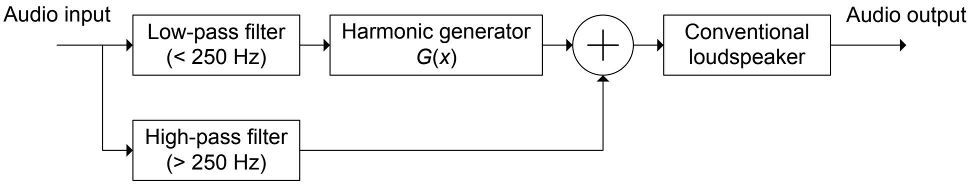

Figure 1. Processing diagram of psychoacoustic bass enhancement in the conventional loudspeaker.

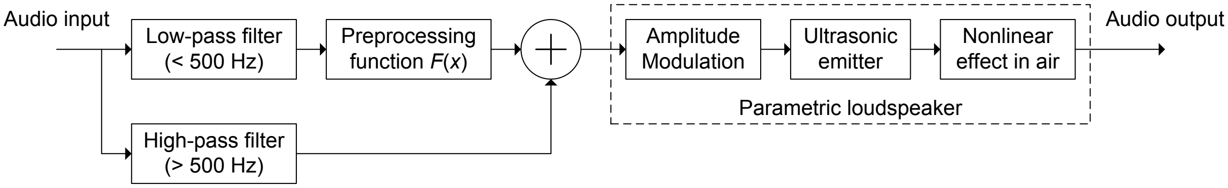

Figure 2. Processing diagram of psychoacoustic bass enhancement in the parametric loudspeaker.

transmitted into air. In [13], a power function with a base constant greater than 0.5 was suggested as the preprocessing function. Although the power function was commonly used as harmonic generators, the fundamental difference between the conventional and parametric loudspeakers should not be neglected. The conventional loudspeaker can be assumed to have a linear response, whereas the parametric loudspeaker is undoubtedly a nonlinear system. So the approach reported in [13] may not result in desired low-frequency enhancement and could generate unwanted intermodulation distortions.

For the aforementioned reason, this paper aims to derive a unified approach in linking the preprocessing function of the parametric loudspeaker to a selected harmonic generator that has been evaluated to be effective in virtually enhancing bass performance of the conventional loudspeaker [16]. In other words, with the consideration of the nonlinear acoustic model [18], the preprocessing function is specially designed to shape the harmonics of the parametric loudspeaker to be equivalent as the output of the known harmonic generator. Instead of treating all the harmonics as distortions, this new preprocessing approach seeks to retain those harmonics that contribute to the "missing fundamental" effect. As a result, the bass quality limitation of the parametric loudspeaker is solved by psychoacoustically reproducing the low-frequency components within one octave below its cut-off frequency.

In this paper, we shall highlight the processing approach in achieving this virtual bass enhancement for the parametric loudspeaker and validate its performance through subjective tests. This paper is organized as follows. Section 2 presents the derived approach that links the preprocessing function to a given harmonic generator. Section 3 gives examples of the preprocessing functions equivalent to two types of harmonic generators, which previously were designed and verified for conventional loudspeakers. Section 4 provides the subjective testing results to validate proposed preprocessing functions in this paper. Lastly, Section 5 concludes this paper.

2. THEORY

The nonlinear acoustic model of the parametric loudspeaker can be described by the Berktay's far-field solution [18] as

$$p_d\left(r,0\right)=K\frac{\partial^2}{\partial t^2}\left\{E^2\left[\sin\omega_d\left(t-r/c_0\right)\right]\right\}, \qquad (1)$$

where r is the distance from the ultrasonic emitter to the observation point, ω_d is the angular frequency of the audio wave; $E\left(x\right)$ is the envelope function with unity amplitude that varies slowly; K is a constant related to several acoustic parameters of air and the speed of sound c_0.

Furthermore, the harmonic generator for conventional loudspeaker is denoted as $G\left(x\right)$. Hence, the preprocessing function $F\left(x\right)$ of the parametric loudspeaker can be designed to generate the same harmonic series as the given harmonic generator $G\left(x\right)$. When the audio input is a sine tone wave, *i.e.* $x=\sin\omega_d\left(t-r/c_0\right)$, this process is described as

$$K\frac{d^2F^2\left(x\right)}{dt^2}=G\left(x\right). \qquad (2)$$

Based on the chain rule [19], we note that

$$\frac{d^2x}{dt^2}=-\omega_d^2\sin\omega_d\left(t-r/c_0\right)=-\omega_d^2x \qquad (3)$$

and

$$\frac{dx}{dt}\frac{dx}{dt}=\omega_d^2\cos^2\omega_d\left(t-r/c_0\right)=\omega_d^2\left(1-x^2\right). \qquad (4)$$

Thus, the left-hand side of (2) is manipulated as

$$K\frac{d^2F^2\left(x\right)}{dt^2}=-K\omega_d^2x\frac{dF^2\left(x\right)}{dx}+K\omega_d^2(1-x^2)\frac{d^2F^2\left(x\right)}{dx^2}. \qquad (5)$$

By substituting (5) into (2) and treating the expression in the square bracket (*i.e.* the derivative of the square of $F\left(x\right)$ with respect to x) as the dependent variable (unknown function), a first-order linear ordinary differential equation is given as

$$\frac{d}{dx}\left[\frac{dF^2\left(x\right)}{dx}\right]-\frac{x}{(1-x^2)}\left[\frac{dF^2\left(x\right)}{dx}\right]=\frac{G\left(x\right)}{K\omega_d^2(1-x^2)}. \qquad (6)$$

Equations of this form can be solved based on the product rule [20]. Thus, the solution to (6) is given by

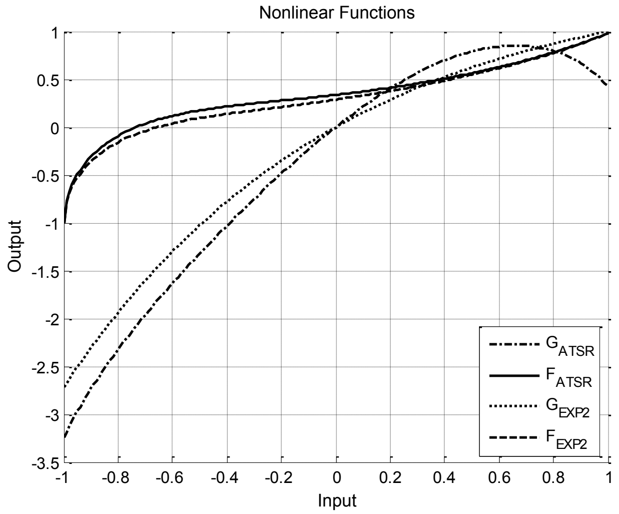

Figure 3. Curves of the nonlinear and equivalent preprocessing functions.

$$\frac{dF^2(x)}{dx} = \frac{\int\left[\frac{G(x)}{K\omega_d^2(1-x^2)} e^{-\int\frac{x}{1-x^2}dx}\right]dx + C}{e^{-\int\frac{x}{1-x^2}dx}}, \tag{7}$$

where C is a constant to be determined later by the boundary conditions. Subsequently, the preprocessing function $F(x)$, which generates the equivalent harmonic series as the given nonlinear function $G(x)$, is solved as

$$F(x) = \frac{Sgn(x)}{\sqrt{K}\omega_d}\sqrt{\int\left[\frac{\int\frac{x-e^{x^2/2}}{1-x^2}G(x)dx + \left(K\omega_d^2\right)C}{x-e^{x^2/2}}\right]dx}. \tag{8}$$

It is noted in (8) that there is a reciprocal factor of the angular frequency ω_d. In practice, when the audio input is a broadband signal instead of a sine tone, this reciprocal factor need be carried out by a low-pass filter. However, there is a built-in low-pass filter in the parametric loudspeaker, which matches the audio input's frequency range with the limited bandwidth of the ultrasonic emitters [21]. Thus, this built-in low-pass filter can be adapted to account for the reciprocal factor.

3. PROPOSED PREPROCESSING FUNCTIONS

In the previous studies of the harmonic generator [16, 22], two forms of nonlinear functions, namely ATSR and EXP2, are recommended based on the subjective testing results. Their expressions are respectively given as

$$G_{ATSR}(x) = 2.5\tan^{-1}(0.9x) + 2.5\sqrt{1-(0.9x)^2} - 2.5 \tag{9}$$

and

$$G_{EXP2}(x) = \frac{e-e^{1-x}}{e-1}. \tag{10}$$

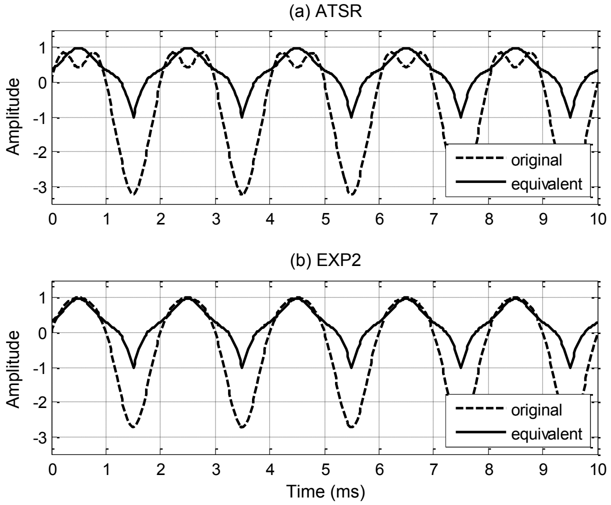

Figure 4. Curves of the nonlinear and equivalent preprocessing functions after the nonlinear acoustic model.

To get the equivalent preprocessing functions of ATSR and EXP2, (9) and (10) are substituted into (8), where the double-integrals are difficult to solve analytically. Therefore, in each integral, we expand the integrand into a Taylor's series [19], and the integral of the resulted polynomial can be calculated easily. The intermediate results are truncated to the 10*th* order, and the final result is further truncated to the 6*th* order, which is suggested to be sufficient in a pitch perception research [23]. Furthermore, the aforementioned constant C is determined to make the output range of $f(x)$ similar to the input range. Thus, the preprocessing functions equivalent to ATSR and EXP2 are respectively proposed as

$$\begin{aligned} F_{ATSR}(x) = &-0.9679x^6 + 0.6669x^5 + 0.7828x^4 \\ &-0.1838x^3 - 0.0859x^2 + 0.3826x + 0.3613 \end{aligned} \tag{11}$$

and

$$\begin{aligned} F_{EXP2}(x) = &-0.8632x^6 + 0.6476x^5 + 0.6871x^4 \\ &-0.2309x^3 - 0.0470x^2 + 0.4596x + 0.3050, \end{aligned} \tag{12}$$

Figure 3 shows curves of the nonlinear and equivalent preprocessing functions, and Figure 4 compares the outputs when the input is a sine tone at 500 Hz. It is observed that the outputs of ATSR and EXP have wider ranges than their inputs. Moreover, the parametric loudspeaker usually has a narrower dynamic range than the conventional loudspeaker. Thus, clipping distortions may occur when the preprocessed audio inputs are played back by the parametric loudspeaker. This is because that the ATSR and EXP2 are designed for the conventional loudspeaker with a cut-off frequency lower than 250 Hz. However, the parametric loudspeaker's cut-off frequency is much higher and normally around 500 Hz. The lower band of the conventional loudspeaker consists of less low-frequency components and has lower amplitude than the lower band of the parametric loudspeaker. Therefore, using ATSR and EXP2 in the parametric loudspeaker may cause unpleasant distortions and notable degradation in the sound

Table 1. List of stimuli.

Title	Artist	Length	Genres
Ferry (Du Kou)	Tsai Chin	12 sec	Folk
Everybody	Backstreet Boys	9 sec	Pop
The phantom of the opera	Andrew Webber	8 sec	Opera

Table 2. List of processing methods.

Index	Description
1	High-pass filter (> 500 Hz)
2	Proposed envelop function F_{ATSR}
3	Nonlinear function ATSR G_{ATSR}
4	Nonlinear function EXP2 G_{EXP2}
5	Proposed envelop function F_{EXP2}
6	Unprocessed, same as reference

quality. Moreover, when the output of the nonlinear function is scaled down, the generated harmonics may not be strong enough to exceed the unmasked threshold [24]. In this case, the nonlinear function still contributes to the intermodulation distortions but no enhancement of the virtual bass.

Generally, the proposed preprocessing functions behave differently from the original nonlinear functions and show obvious superiority in overflow handling, on account of the added attention paid to the nonlinear acoustic model in air.

4. SUBJECTIVE TESTING RESULTS

A subjective test was conducted to compare the virtual bass enhancement performance of the nonlinear functions (G_{ATSR} and G_{EXP2}) and the equivalent preprocessing functions (F_{ATSR} and F_{EXP2}) proposed based on (8). Three snapshots of music with duration of 8-12 seconds as listed in Table 1 were used as the reference (unprocessed) stimuli. The reference stimuli and their high-pass filtered versions were also compared in the subjective test. All the processing methods were indexed and listed in Table 2. A commercial parametric loudspeaker (AS050AW3PF1 [25]) was set up in a quiet room and the subjects sit at 1.5 meters away. In total, 10 subjects from 20 to 50 years old were required to grade their perceived bass intensity and sound quality of the stimuli (processed and unprocessed) from 0 (poor) to 100 (excellent). The subjects were briefed that the reference stimuli have standard scores of 50 for bass intensity and 100 for sound quality. The bass intensity refers to the perceived quantity of lower-frequency components, and the audio quality concerns with noises and distortions. The subjective testing results were compiled and plotted in Figure 5, where both the average scores and 95% confidence intervals were shown.

From the subjective testing results, the high-pass filtered stimuli achieved almost identical low-frequency and sound quality performances to the unprocessed references, which indicated that the cut-off frequency of the tested parametric loudspeaker is at least 500 Hz. Both the nonlinear functions ATSR and EXP2 resulted in moderate improvements of bass perception but severe losses of sound quality, mainly due to the clipping distortions discussed in the previous section. In contrast, the proposed preprocessing function, which takes into account of the entire nonlinear acoustic process of the parametric loudspeaker, preserved a relatively good sound quality and resulted in an improved bass performances. The higher scores of sound quality indicates that the proposed preprocessing functions successfully prevent overflow that may be caused by directly using the conventional nonlinear functions. Among the four methods in the comparison, the preprocessing function F_{ATSR} is conclusively recommended for virtual bass enhancement of the parametric loudspeaker.

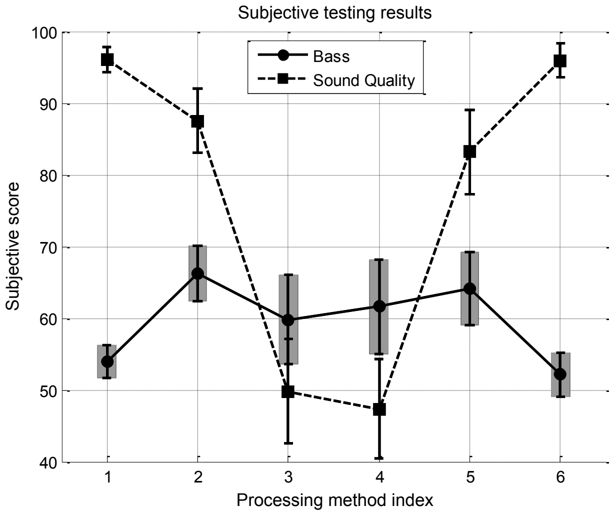

Figure 5. Subjective testing results.

5. CONCLUSIONS

In this paper, a psychoacoustic bass enhancement technique was investigated for the parametric loudspeaker. Because of the nonlinear nature of the parametric loudspeaker, harmonic generators designed for the conventional loudspeaker were not effective for the parametric loudspeaker and may cause clipping distortions. Thus, a new approach was proposed by incorporating with the nonlinear acoustic principle of the parametric loudspeaker. From two well-evaluated nonlinear functions ATSR and EXP2 that were recommended for virtual bass processing of the conventional loudspeaker, two equivalent preprocessing functions were proposed for the parametric loudspeaker. This psychoacoustic preprocessing approach for the parametric loudspeaker had been validated through subjective tests in terms of perceived bass intensity and sound quality. The parametric loudspeaker has been enabled to psychoacoustically reproduce the low-frequency components within an octave below its cut-off frequency. At the same time, an acceptable level of distortion with minimal computational increment has been achieved.

6. ACKNOWLEDGMENTS

This work is supported by Singapore Ministry of Education Academic Research Fund Tier-2 (MOE2010-T2-2-040).

7. REFERENCES

[1] M. Yoneyama, J. Fujimoto, Y. Kawamo, and S. Sasabe, "The audio spotlight: An application of nonlinear interaction of sound waves to a new type of loudspeaker design," *Journal of the Acoustical Society of America*, vol. 73, no. 5, pp. 1013-1020, 1983.

[2] P. J. Westervelt, "Parametric acoustic array," *Journal of the Acoustical Society of America*, vol. 35, no. 4, pp. 535-537, 1963.

[3] C. Shi and W. S. Gan, "Grating lobe elimination in steerable parametric loudspeaker," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 58, no. 2, pp. 437-450, 2011.

[4] C. Shi and W. S. Gan, "Product directivity models for parametric loudspeakers," *Journal of the Acoustical Society of America*, vol. 131, no. 3, pp. 1938-1945, 2012.

[5] W. S. Gan, E. L. Tan, and S. M. Kuo, "Audio Projection," *IEEE Signal Processing Magazine*, vol. 28, no. 1, pp. 43-57, January 2011.

[6] N. Tanaka and M. Tanaka, "Active noise control using a steerable parametric array loudspeaker," *Journal of the Acoustical Society of America*, vol. 127, no. 6, pp. 3526-3537, 2010.

[7] K. Nakamura, H. Ogura, and T. Sugimoto, "Direct visualization of high-intensity focused ultrasonic field using light-emitting diodes and piezoelectric elements," *Acoustical Imaging*, vol. 29, no. 5, pp. 309-316, 2009.

[8] R. W. Haupt and K. D. Rolt, "Standoff acoustic laser technique to locate buried land mines," *Lincoln Laboratory Journal*, vol. 15, no. 1, pp. 3-22, 2005.

[9] S. Aoki, M. Toba, and N. Tsujita, "Sound localization of stereo reproduction with parametric loudspeakers," *Applied Acoustics*, vol. 73, no. 12, pp. 1289-1295, 2012.

[10] F. J. Pompei, "The use of airborne ultrasonics for generating audible sound beams," *Journal of the Audio Engineering Society*, vol. 47, no. 9, pp. 726-731, September 1999.

[11] P. Ji, E. L. Tan, W. S. Gan, and J. Yang, "A comparative analysis of preprocessing methods for the parametric loudspeaker based on the Khokhlov-Zabolotskaya-Kuznetsov equation for speech reproduction," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 19, no. 4, pp. 937-946, 2011.

[12] J. J. Croft and J. O. Norris, "Theory, history, and the advancement of parametric loudspeakers: A technology review," American Technology Corporation, White paper, Part # 98-10006-1100 Rev. E, 2001.

[13] F. A. Karnapi, W. S. Gan, and M. H. Er, “Method to enhance low frequency perception from a parametric array loudspeaker”, in *the 112th Convention of the Audio Engineering Society*, Munich, Germany, 2002.

[14] J. C. R. Licklider, "A duplex theory of pitch perception," *Experientia*, vol. 7, no. 4, pp. 128-134, 1951.

[15] E. Larsen and R. M. Aarts, "Reproducing low-pitched signals through small loudspeakers, *Journal of the Audio Engineering Society*, vol. 50, no. 3, pp. 147–164, 2002.

[16] N. Oo, W. S. Gan, and M. O. J. Hawksford, “Perceptually-motivated objective grading of nonlinear processing in virtual-bass systems”, *Journal of the Audio Engineering Society*, vol. 59, no. 11, pp. 804–824, 2011.

[17] H. Mu, W. S. Gan, and E. L. Tan, “A psychoacoustic bass enhancement system with improved transient and steady-state performance,” in *the 37th IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Kyoto, Japan, 2012, pp. 141 –144.

[18] H. O. Berktay, "Parametric amplification by the use of acoustic nonlinearities and some possible applications," *Journal of Sound and Vibration*, vol. 2, no. 4, pp. 462-470, 1965.

[19] T. M. Apostol, *Mathematical Analysis*. Boston, United States: Addison Wesley, 1974.

[20] G. Birkhoff and G. C. Rota, *Ordinary Differential Equations*. New York, United States: Wiley, 1978.

[21] C. Shi and W. S. Gan, "Development of a parametric loudspeaker: A novel directional sound generation technology," *IEEE Potentials*, vol. 29, no. 6, pp. 20-24, 2010.

[22] N. Oo, W. S. Gan, and W. S. Lim, "Generalized harmonic analysis of arc-tangent square root (ATSR) nonlinear device for virtual bass system," in *the 35th IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Dallas, United State, 2010, pp. 301-304.

[23] C. J. Plack, A. J. Oxenham, and R. R. Fay, and A. N. Popper, *Pitch: Neural Coding and Perception*. New York, United States: Springer, 2005.

[24] E. Ambikairajah, A. G. Davis, and W. T. K. Wong, "Auditory masking and MPEG-1 audio compression," *Electronics Communication Engineering Journal*, vol. 9, no. 4, pp. 165-175, 1997.

[25] Nippon Ceramic Co. LTD., "Parametric speaker," [Online]. Available: http://www.nicera.co.jp/pro/ut/ut-04e.html