


國立交通大學

機械工程學系

碩士論文

以人工類神經網路連結揚聲器客觀量測指標與
主觀聽覺屬性之間的相關性



**Correlation of Objective Nonlinear Measures and
Subjective Timbral Attributes of Loudspeakers using
Artificial Neural**

指導教授：白明憲

研究生：廖國志

中華民國九十九年六月

以人工類神經網路連結揚聲器客觀量測指標與
主觀聽覺屬性之間的相關性

Correlation of Objective Nonlinear Measures and Subjective Timbral Attributes of
Loudspeakers using Artificial Neural

研究生：廖國志

Student : Chun-Jen Wang

指導教授：白明憲

Advisor : Mingsian R.Bai

國立交通大學

機械工程學系

碩士論文



Submitted to Department of Mechanical Engineering

College of Engineering

National Chiao Tung University

In Partial Fulfillment of Requirements

For the Degree of

Master of Science

In

Mechanical Engineering

June 2010

HsinChu, Taiwan, Republic of China.

中華民國九十九年六月

以人工類神經網路連結揚聲器客觀量測指標與 主觀聽覺屬性之間的相關性

研究生：廖國志

指導教授：白明憲 教授

國立交通大學 機械工程學系 碩士班

摘要

揚聲器在大訊號的情形下工作時，許多揚聲器的非線性效應便會顯現，造成聲音的失真進而破壞音質。本研究的目的主要是連結動圈式揚聲器客觀量測指標與主觀聽覺屬性之間的相關性。利用大訊號揚聲器模型製造出數個非線性模型，並分別對非線性模型做客觀指標量測以及主觀聽覺測試。再使用迴歸分析、多變異數分析來分析數據，並使用 Fisher's LSD 檢測主觀聽覺測試結果是否在統計上有顯著的差異。再來我們使用類神經網路來連結客觀量測指標與主觀聽覺屬性之間的相關性。類神經網路對於非線性失真所造成的音質檢測，提供了一個不需要做主觀聽覺測試，而且更有效率的方法。

Correlation of Objective Nonlinear Measures and Subjective Timbral Attributes of Loudspeakers using Artificial Neural

Student: Guo-Zhi Liao

Advisor : Dr. Mingsian R.Bai

Department of Mechanical Engineering

National Chiao-Tung University

Abstract

As a loudspeaker operates in the large-signal domain, nonlinear distortion may arise and impair the sound quality. This work aims to correlate various subjective audio attributes and the objective nonlinear measurements for moving-coil loudspeakers. Several nonlinear models of loudspeaker are created, based on a large-signal loudspeaker model. The data of subjective listening test were processed by the regression analysis, the multivariate analysis of variance (MANOVA), and the least significant difference method (Fisher's LSD) as a *post hoc* test to justify the statistical significance of the results. The objective and subjective indices are correlated with the aid of an artificial neural network (ANN). The network proved effective in assessing subjectively the sound quality impairment due to nonlinear distortions of loudspeakers based on only objective measurements, without having to conduct listening tests.

誌謝

短短兩年的研究生生涯轉眼即逝。在此感謝白明憲教授的諄諄教誨與照顧，在白明憲教授的指導期間，深刻的感受到教授對於追求學問的熱忱，更是佩服教授淵博的學問與解決問題的方法。在教授豐富的專業知識以及嚴謹的治學態度下，使我能夠順利完成學業與論文，在此致上最誠摯的謝意。

在論文寫作方面，感謝本系呂宗熙教授和陳宗麟教授在百忙中撥冗閱讀，並提出寶貴的意見與指導，使得本文的內容更趨完善與充實，在此學生致上無限的感激。

在這兩年的研究生生涯中，承蒙博士班林家鴻學長、李雨容學姊、陳勁誠學長、劉志傑學長，以及已畢業的王俊仁學長、郭育志學長、何克男學長、艾學安學長、劉冠良學長在研究與學業上的適時指點，並有幸與廖士涵同學、陳俊宏同學、張濬閣同學、桂振益同學、曾智文同學、劉嫻婷同學互相切磋討論，讓我獲益甚多。此外學弟王俊凱、吳俊慶、許書豪、衛帝安、馬瑞彬在生活上的朝夕相處與砥礪磨練，亦值得細細回憶。因為有了你們，讓實驗室裡總是充滿歡笑。能順利取得碩士學位，要感謝的人很多，上述名單恐有疏漏，在此一併致上我最深的謝意。

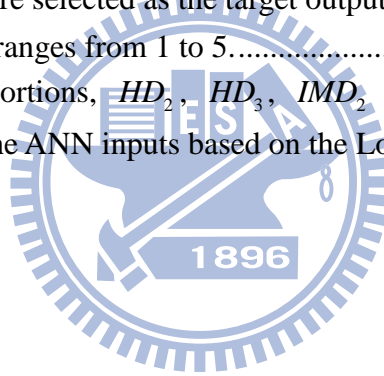
最後僅以此篇論文，獻給我摯愛的家人，父親廖世涼先生、母親程月招女士、大妹廖怡婷以及小妹廖怡茹，這一路上，因為有你們的付出與支持，給了我最大的精神支柱，也讓我有勇氣面對更艱難的挑戰。

TABLE OF CONTENTS

摘要.....	i
Abstract.....	ii
誌謝.....	iii
TABLE OF CONTENTS.....	iv
LIST OF TABLES.....	v
LIST OF FIGURES.....	vi
1. INTRODUCTION.....	1
2. THE LARGE-SIGNAL MODEL OF MOVING-COIL LOUDSPEAKERS..	3
2.1 Nonlinearities of Moving-Coil Loudspeakers	3
2.2 Large-Signal Model	4
3. OBJECTIVE EVALUATION BY NUMERICAL SIMULATION.....	5
3.1 Measures of Nonlinear of Symptoms	5
3.2 Relations between Nonlinear Causes and Symptoms of Loudspeakers	6
4. SUBJECTIVE LISTENING TESTS.....	7
4.1 Experimental Arrangement.....	7
4.2 Results of the Listening Test.....	9
5. ARTIFICIAL NEURAL NETWORKS.....	10
6. CONCLUSIONS.....	13
7. APPENDIX.....	14
Design of a MEMS microphone using SA method.....	14
7.1. Quasi-Static model.....	14
7.2 Linear Dynamic Model.....	16
7.3 SA method.....	18
REFERENCES.....	21

LIST OF TABLES

Table 1. The relations between nonlinear causes and symptoms of moving-coil loudspeakers.....	25
Table 2. The MANOVA output of the listening test for the fifteen nonlinear loudspeaker models. Cases with significance value p below 0.05 indicate that statistically significant difference exists among all cases.	26
Table 3. Multiple regression of Total Preference in relation to Fidelity, Fullness and Artifacts.....	27
Table 4. Multiple regression analysis of Total Preference in relation to HD_2 , HD_3 , IMD_2 and IMD_3	28
Table 5. Fifteen nonlinear models with variations on speaker parameters of loudspeaker A. The first ten cases of objective index were selected as the input for training the ANN.....	29
Table 6. The target output (listening test result) of the ANN. The first ten cases of objective index were selected as the target output data for training the ANN. The grading scale ranges from 1 to 5.....	30
Table 7. Four nonlinear distortions, HD_2 , HD_3 , IMD_2 and IMD_3 , of the out-group test as the ANN inputs based on the Loudspeaker model B.	31



LIST OF FIGURES

Fig. 1	Electroacoustic analogous circuit of a moving-coil loudspeaker. (a) The equivalent circuit. (b) The circuit of the thermal model.	32
Fig. 2	The flowchart of the modeling and verification procedure for the ANN-based loudspeaker assessment system	33
Fig. 3	The result of the listening test for the fifteen nonlinear loudspeaker models. The means and spreads (with 95% confidence intervals) of the grades are indicated in the figure. The x-axis and y-axis correspond to the nonlinear models and the grades, respectively.....	34
Fig. 4	The schematic of the neuron. The symbol a_n is the input, w_n is the weight, b is the bias, and o is output.	35
Fig. 5	The structure of the back-propagation ANN. The symbol a is the hidden layer, a hidden layer of fifteen neurons, x_i is the input layer of the nonlinear distortion measures HD_2, HD_3, IMD_2 and IMD_3 , w_{ij} and w_k are the weights, b_1 and b_2 are the bias units, and y is the output of the subjective attributes: Fidelity, Fullness, Artifacts and Total Preference.	36
Fig. 6	Comparison of Total preference, Fidelity, Fullness and Artifacts predicted by the ANN and the subjective data obtained from the listening test for the last five cases (in-group verification). $Error = \frac{Output - Target}{5}$	37
Fig. 7	The out-group test result of the listening test for the four nonlinear loudspeaker models. The means and spreads (with 95% confidence intervals) of the grades are indicated in the figure. The x-axis and y-axis correspond to the nonlinear models and the grades, respectively.....	38
Fig. 8	The results of the out-group test of the ANN based on the Loudspeaker model B. Total preference, Fidelity, Fullness and Artifacts predicted by the ANN and the subjective data obtained from the listening test are compared. $Error = \frac{Output - Target}{5}$	39
Fig. 9	The flow chart of the iterative procedure from FDM method for the coupled equation system.....	40
Fig. 10	Acoustical system includes the sound radiation, air gap, back chamber and acoustical holes influence.	41
Fig. 11	Combination of mechanical and electrical system.	41

Fig. 12 The comprehensive system combines the acoustical, mechanical and electrical system.....42



1. INTRODUCTION

Loudspeaker evaluation has been a long-standing issue in that it involves not only physical aspects but also psychoacoustic features which are largely subjective. More often than not, audiophiles trust more on the experience of human experts than the frequency response on the data sheet. This is why sometimes it is regarded as art rather than science to assess loudspeaker quality and the price at the marketplace can vary drastically without bound. Many factors can contribute to the overall sound quality of a loudspeaker. Among these factors, nonlinear distortions have profound impact on the timbral quality of a mono-channel loudspeaker. To address the issue, this paper aims at exploring the correlation between the objective measures of nonlinear distortions and the subjective perception of timbral quality associated with these distortions. If this correlation can be found, akin to the PEAQ [1] for audio codec assessment, then it is possible to develop an automatic system for loudspeaker evaluation, without having to conduct human listening tests.

Loudspeakers can be evaluated with objective measurements and subjective listening tests. The latter are generally time consuming and tedious to carry out. Lavandier et al. investigated subjective dimensions based on underlying perceived differences between loudspeakers [2]. Multidimensional scaling technique was employed to analyze temporal and spectral data of listening tests. Two principal perceptual dimensions, the bass/treble balance and the medium emergence, were identified for loudspeaker evaluation. Liu et al. [3]-[4] attempted to evaluate loudspeakers using sound quality metrics [5] suggested by Zwicker and Fastl. Metrics including loudness, sharpness, fluctuation strength and roughness that are widely used in assessing “noise” quality of products are employed for quantifying the timbral quality of loudspeakers. Listening tests were conducted in that study.

This paper seeks to establish the correlation between objective measurement and

subjective attributes with special regards to nonlinear distortion of moving-coil loudspeakers. Traditionally, many objective indices such as sensitivity, efficiency, directivity pattern, nonlinear distortions, etc., can be used for loudspeaker evaluation. Among these indices, nonlinear distortions have direct impact on the perception of timbral quality produced by loudspeakers. As loudspeakers operate in the large-signal regime, nonlinear distortions may arise and can strongly impair sound quality. Metrics such as DC-displacement, harmonic distortion, inter-modulation, etc., can be used to quantify loudspeaker nonlinearities [6]-[7]. In this paper, four types of nonlinear metrics, HD_2 , HD_3 , IMD_2 , and IMD_3 are used in the objective measurements. Nonlinear distortions of a loudspeaker may arise as a result of various causes such as nonlinear compliance, nonlinear force factor, nonlinear inductance, etc [8]-[9]. In order to produce a sufficient database of nonlinear distortion, a large-signal model [10]-[13] is employed in this study to simulate large-signal responses of loudspeakers.

With the database, an artificial neural network (ANN) [14]-[15] was trained to correlate the objective nonlinear indices and the subjective attributes obtained in listening tests. Two loudspeaker models A and B were used in this study for in-group and out-group verifications, respectively. An ANN was established on the basis of fifteen nonlinear models of loudspeaker A and the human perceptions of nonlinear distortions in listening tests. Fifteen nonlinear models of loudspeaker A were then used for in-group verification. Among the models, ten nonlinear models are used to train the neural network, and the remainder is used to verify the ANN. For out-group verification, four nonlinear models derived from loudspeaker B were used. The assessment predicted by the ANN was compared to a listening test conducted for the loudspeaker B. With robustness so justified, the network serves for an automatic loudspeaker evaluation system, without resorting to listening tests.

2. THE LARGE-SIGNAL MODEL OF MOVING-COIL LOUDSPEAKERS

Sufficient amount of training data are required to establish a reliable ANN. Instead of gathering data from a large number of real loudspeakers, we opt for a more practical approach by using a large-signal moving-coil loudspeaker model, as described in this section. At low frequencies where the wavelength is large in comparison to the geometric dimensions, the state of a loudspeaker can be described by a lumped parameter model. In Fig. 1(a), the model is shown in terms of electroacoustic equivalent circuits coupled in the electrical, mechanical and acoustical domains [10]-[12]. The lumped parameters in the circuits such as the force factor $Bl(x)$, the mechanical compliance $C_{MS}(x)$ and the voice coil inductance $L_E(x)$ can vary with the displacement of the voice coil.

2.1 Nonlinearities of Moving-Coil Loudspeakers

Four types of nonlinearity of moving-coil loudspeakers employed in the large-signal model are summarized as follows:

1. Loudspeaker stiffness $K_{MS}(x)$: the mechanical stiffness of driver suspension which is also defined as the inverse of mechanical compliance $C_{MS}(x)$. The restoring force $F = K_{MS}(x)x$ of the suspension system can cause nonlinear distortions.
2. Force factor $Bl(x)$: instantaneous electrodynamic coupling factor between the mechanical and electrical domains, where B is flux density and l is the effective length of the voice coil. The force factor $Bl(x)$ is not a constant and is a function of diaphragm displacement x . Two nonlinear effects force factor $Bl(x)$ can arise via: (a) back electromotive force (back-EMF) $e = Bl(x)u$, where u is cone velocity, and (b) Lorentz force $F = Bl(x)i$, where i is current.
3. Voice-coil inductance $L_E(x)$: the effective inductance of voice coil. $L_E(x)$ is function of displacement x and is hence a cause of loudspeaker nonlinearity.

2.2 Large-Signal Model

A large-signal model is exploited to simulate nonlinear responses of loudspeaker in the large-signal regime. The loudspeaker model is represented using the equivalent circuit with nonlinear parameters, as shown in Fig. 1(a). The temperature increase in the voice coil is described by a separate thermal model shown in Fig. 1(b) [11]. The DC resistance R_E is dependent of the ambient temperature T_A and the temperature increase of voice coil ΔT_V .

$$R_E(T_A + \Delta T_V) = R_E(T_A)(1 + \delta \Delta T_V), \quad (1)$$

where $\delta = 0.00393 \text{ degK}^{-1}$ for copper and $\delta = 0.00393 \text{ degK}^{-1}$ for aluminum. All parameters can be identified by a distortion analyzer [16]. The dynamics of the system depicted by Fig. 1(a) are governed by the following simultaneous ordinary differential equation system:

$$u = iR_E(T_V) + \frac{d(L_E(x)i)}{dt} + \frac{d(L_2(x)i_2)}{dt} + Bl(x)v, \quad (2)$$

$$\frac{d(L_2(x)i_2)}{dt} = (i - i_2)R_2(x), \quad (3)$$

$$Bl(x)i - F_m(x, i, i_2) = M_{MS} \frac{d^2 x}{dt^2} + R_{MS} \frac{dx}{dt} + K_{MS} x, \quad (4)$$

Choosing $y_1 = x$, $y_2 = v$, $y_3 = i$ and $y_4 = i_2$ as the state variables enables us to rewrite Eqs. (2)-(4) into the following state-space equation:

$$\begin{bmatrix} \dot{y}_1 \\ \dot{y}_2 \\ \dot{y}_3 \\ \dot{y}_4 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ \frac{-1}{M_{MS}C_{MS}(x)} & \frac{-R_{MS}}{M_{MS}} & \frac{Bl(x) + \frac{1}{2} \frac{dL_E(x)}{dx} y_3}{M_{MS}} & \frac{\frac{1}{2} \frac{dL_2(x)}{dx} y_4}{M_{MS}} \\ 0 & \frac{-Bl(x) - \frac{dL_E(x)}{dx} y_3}{L_E(x)} & \frac{-(R_E(T_V) + R_2(x))}{L_E(x)} & \frac{R_2(x)}{L_E(x)} \\ 0 & 0 & \frac{R_2(x)}{L_2(x)} & \frac{-R_2(x) - \frac{dL_2(x)}{dx} y_2}{L_2(x)} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} u \quad (5)$$

The Runge-Kutta [17] numerical integration algorithm can readily be applied to solve

the state-space equation for the nonlinear responses. Now that the velocity $y_2 = v$ is obtained, the time-domain farfield sound pressure response can be calculated using a baffled point source model

$$p(t, r) = \frac{\rho_0 S_D}{2\pi r} \frac{dv}{dt}, \quad (6)$$

where r is the distance between the diaphragm and the listening position, ρ_0 is the density of air and S_D is the area of the diaphragm. Note that pressure calculation of Eq. (6) requires numerical differentiation.

3. OBJECTIVE EVALUATION BY NUMERICAL SIMULATION

For the objective and subjective evaluations to be conducted next, nonlinearly distorted signals were generated by the preceding large-signal model and reproduced by a high-quality linear loudspeaker (iPod HiFi A1121).

3.1 Measures of Nonlinear of Symptoms

Various objective measures are adopted in the study to characterize loudspeaker nonlinearity. The first basic measure is harmonic distortion (HD) that employs a single-tone stimulus according to IEC standard 60268-5 [18]. The n th harmonic distortion associated with f_1 is defined as

$$HD_n = \frac{|P(nf_1)|}{P_t} \times 100\%, \quad (7)$$

where $P(nf_1)$ is the complex spectrum of sound pressure P at the n th harmonic, and P_t is the rms-value of the total signal within the averaging duration T

$$P_t = \sqrt{\frac{1}{T} \int_0^T p(t)^2 dt}. \quad (8)$$

Another important nonlinear measure is the Inter-modulation Distortion (IMD). In this measure, two separate tones are used as stimuli, the first tone f_1 is set to be the resonance frequency of the loudspeaker and the second tone f_2 is set to be

higher than $8.5f_1$. Furthermore, the amplitude U_1 of the input voltage of the first tone should be 4 times larger than the amplitude U_2 of the second tone. In the IMD test, the following two-tone excitation signal is used

$$u(t) = U_1 \sin(2\pi f_1 t) + U_2 \sin(2\pi f_2 t). \quad (9)$$

IMD accounts for extra frequency components due to intermodulations occurring at $f_2 \pm nf_1 (n=1, 2, \dots)$. The n th-order IMD is defined as

$$IMD_n = \frac{|P(f_2 - (n-1)f_1)| + |P(f_2 + (n-1)f_1)|}{|P(f_2)|} \times 100\%. \quad (10)$$

3.2 Relations between Nonlinear Causes and Symptoms of Loudspeakers

A rule-based logic is established according to the relations between the nonlinear causes and symptoms of loudspeaker nonlinearities, as summarized in Table 1. Nonlinear distortions with various types and levels can be synthesized by varying loudspeaker parameters, $Bl(x)$, $C_{MS}(x)$, and $L_E(x)$, which will be used in the subsequent objective and subjective experiments.

The nonlinear causes in a loudspeaker's response are subdivided into the following two categories: critical nonlinearity variation and asymmetric nonlinearity [6]-[7]:

1. Critical nonlinearity variation: "coil height" and "symmetrical limiting of suspension" A symmetric curve usually produces the 3rd- and other odd-order distortion components. To simulate the "coil height" and the "suspension limiting" defect, the 2nd-degree term of the Bl - and the C_{MS} -series are usually multiplied by a factor $\beta > 1$ to increase their variations.

$$Bl'(x) = b_0 + b_1 x + \beta b_2 x^2 + \sum_{i=3}^n b_i x^i, \quad (11)$$

$$C'_{MS}(x) = c_0 + c_1 x + \beta c_2 x^2 + \sum_{i=3}^n c_i x^i \quad (12)$$

2. Asymmetric nonlinearity: “coil offset,” and “asymmetry in $C_{MS}(x)$ and $L_E(x)$ ”
 Asymmetric $Bl(x)$ and $C_{MS}(x)$ curves generally result in the 2nd- and other even-order distortion components. These defects can be modeled by shifting the $Bl(x)$ and $C_{MS}(x)$ curves on the x -axis by a small constant ε .

$$Bl'(x) = Bl(x - \varepsilon) = \sum_{i=0}^n b_i (x - \varepsilon)^i \quad (13)$$

$$C'_{MS}(x) = C_{MS}(x - \varepsilon) = \sum_{i=0}^n c_i (x - \varepsilon)^i . \quad (14)$$

The inductance $L_E(x)$ without shorting ring also exhibits asymmetric characteristics. To model this, we multiply the linear term of the L_E -series by a factor $\beta > 1$ to yield

$$L'_E(x) = l_0 + \beta l_1 x + \sum_{i=2}^n l_i x^i \quad (15)$$

4. SUBJECTIVE LISTENING TESTS

Figure 2 illustrates the overall architecture of the experimental procedure. The real time signal was used as the input for both in-group and out-group verifications. Fifteen nonlinear models of loudspeaker A were created using the preceding large signal model. These models will be used in the following objective and subjective tests, separately. The objective nonlinear measures and subjective timbral attributes obtained from these models serve as the input and the output to the ANN in the training phase. After the ANN is trained, another four nonlinear models of loudspeaker B were used for out-group verification in which the output of the ANN and the listening test results will be compared.

4.1 Experimental Arrangement

Initially, subjective tests are conducted to obtain the relations between the objective measures and the subjective attributes. There are fifteen experienced subjects between the age of 22 and 31 years (13 males, 2 females) participating in this

test. Fifteen participants in the listening tests were instructed with definitions of the subjective attributes and procedures before the test began. The participants were asked to respond in a questionnaire after listening. The listening tests were carried out in a standard listening room (ITU-R BS. 1116-1 [19]). The listeners sat 1m away from the monophonic loudspeaker (iPod HiFi A1121). The listening test complies with the MUSHRA procedure (ITU-R BS. 1534-1 [20]), a modified double-blind Multi-Stimulus test with a hidden reference and a hidden anchor. The grading scale ranges from 1 to 5, indicating bad, poor, fair, good and excellent, respectively. Nonlinear loudspeaker models were created using the large-signal model with reference to the nonlinear cause-symptom logic summarized in Table 1. Fifteen nonlinear models of loudspeaker A, alongside the reference and anchor signals were created for the in-group listening tests. The test stimulus is a music clip of "Hotel California." The linear speaker model was used as the reference (the grading scale of the attributes is 5). The hidden anchor is a high-pass filtered input signal. The subjective attributes employed in this subjective test are the following four indices:

- (1) Fidelity: clarity of the voice and the music.
- (2) Fullness: quality of low-frequency sound.
- (3) Artifacts: any extraneous disturbances to the signal are considered as artifact.
- (4) Total Preference: global attribute for listeners to judge the sound based on their own likes or dislikes.

Every subject is asked to grade the attributes above in sequence with the index Total Preference being the last item. It took approximately thirty minutes to finish a listening experiment.

4.2 Results of the Listening Test

A listening test was conducted using the preceding subjective attributes. The cases with significance levels (p -values) of the test results processed using Multivariate ANalysis Of VAriance method (MANOVA) are shown in Table 2 and Fig. 3. The nonlinear models of loudspeaker A with p -values below 0.05 indicate that statistically significant difference exists among methods. The p -values of the attributes, Fidelity, Fullness, Artifacts and Total Preference are all zeros. Since significant difference is present in all attributes, we conduct a *post hoc* Fisher's LSD test as multiple paired comparisons. There is neither statistically significant difference among Cases 3, 6, 10, 11 and 14, nor among Cases 9, 12 and 15. Apart from these cases, the results of the *post hoc* test indicate significant difference in pairs.

It can be observed in Fig 3 that Cases 4 and 7 received highest grades because of its low HD_2 , HD_3 , IMD_2 and IMD_3 values. (<10%) Cases 6, 10, 11 and 14 received relatively higher grade due to its lower HD_2 and HD_3 values. Conversely, Cases 9, 12 and 15 received the lowest grades due to its higher HD_2 and HD_3 values.

Figure 3 and Table 2 summarizes the influence of nonlinear distortions on the subjects. In this experiment, the nonlinear speaker models seem to result in marked difference ($p < 0.05$) in all subjective indices.

In order to examine how Total Preference is subjectively related to Fidelity, Fullness and Artifacts, multiple regression analysis was conducted for the listening test results, as shown in Table 3. The linear equation obtained from the regression analysis is given as

$$\text{Total Preference} = -0.0333 + 0.1141 \times \text{Fidelity} + 0.2345 \times \text{Fullness} + 0.6750 \times \text{Artifacts} \quad (16)$$

The correlation coefficient of the attribute Artifacts is significantly larger than those of the other indices. Fidelity is the lowest one. This is due to the fact that the loudspeakers parameters as $Bl(x)$, $C_{MS}(x)$ and $L_E(x)$ are all functions of displacement. As a loudspeaker undergoes large excursions at the low frequency range, nonlinear distortions are perceived as low-frequency artifacts. The sound quality of the test signals were not compromised at the high frequency range because the nonlinear distortions influence only the low frequencies. Overall, the attribute Artifacts is the most influential index to the Total Preference, followed by Fullness and Fidelity.

5. ARTIFICIAL NEURAL NETWORKS

Artificial neural networks (ANN) consist of a large number of simple processing units called neurons (Fig. 4). The a_n is the input, w_n is the weight, b is the bias and o is output. A neuron is a multiple-input-single-output unit in which the output signal is usually a non-linear function of the input vector and a weight vector. Analogous to a human brain, an ANN is capable of learning, recalling and generalizing from the training data by assigning or adjusting the connection weights. An ANN is a multi-layer nonlinear filter, which lends itself very well to nonlinear modeling or correlation. In this paper, we use the back-propagation algorithms to update network parameters. In the first place, the ANN is trained by supervised learning rules, where a set of input-output pair are required for the network training. The input is propagated to the output layer to be compared with the target output. Then the error signals are then back-propagated from the output layer back to the intermediate hidden layers to update the weight coefficients. The preceding procedure is repeated until the error converges to a small value.

Traditionally, many objective indices such as sensitivity, efficiency, directivity

pattern, nonlinear distortions, etc., can be used for loudspeaker evaluation. Among these indices, nonlinear distortions have direct impact on the perception of timbral quality produced by loudspeakers. Four types of nonlinear distortions (HD_2 , HD_3 , IMD_2 , and IMD_3) were used in this study. In the following, we shall use an ANN to correlate the objective nonlinear measures and the subjective attributes. Four nonlinear distortion measures (HD_2 , HD_3 , IMD_2 , and IMD_3) serve as the inputs to the ANN, whereas four subjective attributes (Fidelity, Fullness, Artifacts and Total Preference) serve as the outputs to the ANN. Next, we applied multiple regression analysis to examine the influence of the HD_2 , HD_3 , IMD_2 and IMD_3 on Total Preference (Table 4). The correlation coefficient of HD_3 is -0.8864, implying that HD_3 is a more prominent index than the other indices. However, the significance values of HD_2 , IMD_2 and IMD_3 are 0.7887, 0.1123 and 0.8429, respectively, which are greater than 0.05. This indicates that the difference among these three indices is not statistically significant. The ANN is used to correlate the objective nonlinear measures and the subjective attributes. The four nonlinear distortions (HD_2 , HD_3 , IMD_2 and IMD_3) were all considered in the ANN. A three-layered feedforward network used in the work is shown in Fig. 5. The network is comprised of an input layer of four neurons, a hidden layer of fifteen neurons, and an output layer of four neurons, respectively. The activity functions at the hidden layer and output layer are the logsig function and the poslin function, as defined by

$$\text{logsig: } f(x) = \frac{1}{1 + e^{-x}} \quad (17)$$

$$\text{poslin: } \begin{cases} f(x) = x, x \geq 0 \\ f(x) = 0, x < 0 \end{cases} \quad (18)$$

where x is the input and $f(x)$ is the output. In Fig. 5, the ANN operations are defined by

$$a = \sum_i^4 \sum_j^{15} \text{logsig}(w_{ij} \times x_i + b_1) \quad (19)$$

$$y = \sum_k^{15} \text{poslin}(w_k \times a + b_2) \quad (20)$$

where a is the hidden layer, x_i is the input of the nonlinear distortions HD_2, HD_3, IMD_2 and IMD_3 , w_{ij} and w_k are the weights, b_1 and b_2 are the bias units, and y is the output of the subjective attributes, Fidelity, Fullness, Artifacts and Total Preference. Fifteen nonlinear loudspeaker models (Table 5) were used to generate stimuli for the listening tests. The results of the listening tests are summarized in Table 6. Ten groups are selected to train the ANN, whereas the remaining five groups serve to verify the network. The output of the ANN is compared with the target listening test data ($Error = \frac{Output - Target}{5}$) in Fig 6. The network yields satisfactory accuracy since the error between the prediction and the target output is quite small ($< \pm 10\%$).

To justify the ANN, we conducted an out-group test using another loudspeaker B. The experimental arrangement of the in-group and the out-group test are the same. There were fifteen experienced subjects participating in this test. The participants were instructed with definitions of the subjective attributes and procedures before the test. Loudspeaker B is a 10cm woofer with the fundamental resonance frequency 60 Hz. The objective test procedure follows the IEC 60268-5, $U_1 = 4U_2$, $f_1 = f_s$, and $f_2 = 8.5f_s$. The bass tone input voltage U_1 was chosen according to the rated power of the loudspeaker. Four nonlinear models were generated using loudspeaker B and were subjected to another out-group listening test along with the reference and anchor signals as defined previously. The objective measurement (HD_2, HD_3, IMD_2 and IMD_3) of the four nonlinear models serve as the input, whereas the results of the listening test serve as the target output to the ANN, as

summarized in Table 7 and Fig. 7. The result of comparison is summarized in Fig. 8 ($Error = \frac{Output - Target}{5}$). We observed in Fig. 7 that Case3 received higher grades than the other cases because of its low HD_2 , IMD_2 and IMD_3 values. Conversely, the Case 4 received the lowest grades due to its highest HD_2 and IMD_2 values. The errors of prediction are mostly below 10% except Case 4. From the in-group and out-group test results, it is also observed that the high-grade cases usually yield low nonlinear distortion level (<10%), while the low-grade cases usually yield high HD_2, HD_3 levels.

6. CONCLUSIONS

This paper presents an ANN-based inference engine for correlating objective nonlinear measures and subjective timbral attributes. Fifteen nonlinear loudspeaker models of loudspeaker A created using the large-signal model served as the stimuli in the objective and subjective tests, respectively. The objective data and the subjective data obtained using these fifteen nonlinear models served as the inputs and outputs in the training and the in-group verification of the ANN. The errors of the in-group test between the predicted output and the target output are quite small ($< \pm 10\%$). Another out-group test was undertaken using loudspeaker B to further verify the ANN. The results inferred by the neural network were in good agreement with those obtained using numerical prediction. The errors of the out-group test between the predicted output and target output are mostly below $\pm 10\%$ except Case 4. From the results, high-grade cases usually have low nonlinear distortion level (<10%), while low-grade cases usually have high HD_2, HD_3 levels. This ANN-based loudspeaker assessment system provides a cost-effective solution for loudspeaker diagnostics without have to conduct listening tests.

7. APPENDIX

Design of a MEMS microphone using SA method

This work aims to design of a MEMS microphone using SA method. The first part of appendix introduces the MEMS microphone models: the quasi-static model and the linear dynamic model. The second part of appendix is SA method of optimal design.

7.1. Quasi-Static model

Quasi-Static Analysis is used to calculate the deflection of the plate and the maximum DC bias voltage can be applied. Firstly, a simple structure of MEMS condenser microphone is considered [21]-[23].

Considering the bending moments on a small part of the plate $dx dy$, the equation of equilibrium reads [24]

$$\frac{\partial^2 M_x}{\partial x^2} - \frac{\partial^2 M_{xy}}{\partial x \partial y} + \frac{\partial^2 M_y}{\partial y^2} = -p \quad (21)$$

where $p(x, y)$ is the externally applied loads, M_x , M_{xy} and M_y are bending moment given by

$$M_x = -\frac{h_d^3}{12} \left(C_{11} \frac{\partial^2 w_d}{\partial x^2} + C_{12} \frac{\partial^2 w_d}{\partial y^2} \right) \quad (22)$$

$$M_y = -\frac{h_d^3}{12} \left(C_{21} \frac{\partial^2 w_d}{\partial x^2} + C_{22} \frac{\partial^2 w_d}{\partial y^2} \right) \quad (23)$$

$$M_{xy} = \frac{h_d^3}{6} C_{44} \frac{\partial^2 w_d}{\partial x \partial y} \quad (24)$$

$w_d(x, y)$ is the plate thickness, $C_{11} = C_{12}$, $C_{12} = C_{21}$, and C_{44} are material constants of the plate given by

$$C_{11} = \frac{E}{1-\nu^2}, \quad C_{12} = \frac{E\nu}{1-\nu^2}, \quad C_{44} = \frac{E}{2(1+\nu)} \quad (25)$$

where E and ν are the Young's modulus and the Poisson's ratio, σ_d is the built-in stress. The plate equilibrium equation is written as

$$\begin{aligned} & C_{11} \frac{\partial^4 w_d}{\partial x^4} + 2(C_{12} + 2C_{44}) \frac{\partial^4 w_d}{\partial x^2 \partial y^2} + C_{11} \frac{\partial^4 w_d}{\partial y^4} \\ &= \frac{12}{h_d^3} \left[p + \sigma_d h_d \left(\frac{\partial^2 w_d}{\partial x^2} + \frac{\partial^2 w_d}{\partial y^2} \right) \right] \end{aligned} \quad (26)$$

The electrostatic loads between the diaphragm and the back-plate has been derived as force per unit area

$$p_{el}(x, y) = K_{holes} \frac{\epsilon_d \epsilon_0}{2(h_d + \epsilon_d (h_a - w_d))} V_{ba}^2 \quad (27)$$

where ϵ_d and ϵ_0 are the relative permittivity and vacuum permittivity of the diaphragm material, h_a and h_d are the thickness of air gap and the thickness of the diaphragm, K_{holes} is the scalar factor.

The whole system is considered as

$$\begin{aligned} & C_{11} \frac{\partial^4 w_d}{\partial x^4} + 2(C_{12} + 2C_{44}) \frac{\partial^4 w_d}{\partial x^2 \partial y^2} + C_{11} \frac{\partial^4 w_d}{\partial y^4} \\ &= \frac{12}{h_d^3} \left[p_{sp} + p_{el} + \sigma_d h_d \left(\frac{\partial^2 w_d}{\partial x^2} + \frac{\partial^2 w_d}{\partial y^2} \right) \right] \end{aligned} \quad (28)$$

where p_{sp} is the sound pressure. Only a quarter of the diaphragm should be evaluated since the structure is assumed to be squared and symmetric. The boundary conditions on the edges is given by

$$w_d(x, y) = \frac{\partial w_d(x, y)}{\partial x} = 0 \quad (29)$$

Finally, we arrange the linear equations of comprehensive discrete points and boundary conditions into a matrix form as follows

$$\mathbf{A}_w \mathbf{w}_d = \mathbf{p}_{sp} + \mathbf{p}_{el} \quad (30)$$

The procedure is started with initializing \mathbf{w}_d as a zero vector. Then, p_{el} is

calculated. Once p_{el} is evaluated, Eq. (30) is calculated again to acquire w_d . If the center deflection of the diaphragm is greater than the thickness of the air gap, the diaphragm and back-plate are attracted each other. If the structure is not collapsed, to determine the solution is converged or not. Until the solution is converged, this procedure is terminated. Flow chart of the iterative procedure of FDM method is shown in Fig 9.

7.2 Linear Dynamic Model

The basic structure of the condenser microphone consists of a diaphragm, a perforated back plate, an air gap and a back chamber. The dynamic behavior is very complex in that it involves actions and interactions in and between three different. The different fields include the acoustical, mechanical and electrical domains.

1. The acoustical radiation impedance of a baffle piston

$$R_{A1} = \frac{0.441\rho_0 c}{L_m^2} \quad (31)$$

$$C_{A1} = \frac{5.94L_m^3}{\rho_0 c^2 \pi \sqrt{\pi}} \quad (32)$$

$$R_{A2} = \frac{\rho_0 c}{L_m^2} \quad (33)$$

$$M_{A1} = \frac{8\rho_0}{3\pi\sqrt{\pi}L_m} \quad (34)$$

where ρ_0 is the density of the air, c is the sound speed in air, L_m is the side length of diaphragm.

2. Modeling of the air gap

The resistance and mass in the mechanical domain has been derived by Skvor [25]. Skvor considered the air gap is a non-compressible laminar flow because the compliance is small enough to be neglected.

$$R_a = \frac{1.22\eta\pi b^2}{h_a^3 L_m^2} B \quad (35)$$

$$M_a = \frac{0.102\rho_0\pi b^2}{h_a L_m^2} B \quad (36)$$

$$B = \frac{1}{4} \ln\left(\frac{0.16b^2}{a_h^2}\right) - \frac{3}{8} + \frac{a_h^2}{0.16b^2} - \frac{a_h^4}{0.204b^4} \quad (37)$$

where $\eta = 1.86 \times 10^{-5}$ Ns/m² is the dynamic viscosity of air, a_h is the half side length of the square holes.

3. Modeling of the air in the acoustical holes

To consider the holes are like narrow slits [26]. The mechanical resistance and mass of the air in the acoustical holes are

$$R_h = \frac{12\eta h_b}{b^2 L_m^2} \quad (38)$$

$$M_h = \frac{24\rho_0 h_b a_h^2}{5b^2 L_m^2} \quad (39)$$

4. Modeling of the air in the back chamber

Because the air in the back chamber is compressible, it is considered as acoustical compliance.

$$C_{bc} = \frac{V}{\rho_0 c^2} \quad (40)$$

where V is the volume of the back chamber.

By the above equivalent circuits, we can illustrate the acoustical system in the fig. 10.

5. Modeling the diaphragm

Assume that the diaphragm is clamped at the boundary. We only consider the mechanical mass and compliance of the diaphragm.

$$C_{MD} = \frac{32}{\pi^6 \sigma_d h_d L_m^2} \quad (41)$$

$$M_{MD} = \rho_d h_d L_m^2 \quad (42)$$

where σ_d is the built-in stress and ρ_d is the density of the diaphragm material.

Fig. 11 illustrates the analogous circuit of the diaphragm and electrical circuit, the mechanical-electrical coupling factor as follow

$$\phi_T = \frac{C_{E0}}{C_{EM}} \quad (43)$$

$$C_{E0} = K_{holes} \frac{\epsilon_b \epsilon_d \epsilon_0 L_m^2}{h_a \epsilon_b \epsilon_d + h_b \epsilon_d + h_d \epsilon_b} \quad (44)$$

$$C_{EM} = \frac{h_a - w_d}{V_{ba}} \quad (45)$$

C_{E0} is the capacitance of the plate caused by DC bias voltage,

The mechanical compliance has to be modified by

$$K_f^2 = \frac{C_{E0} C_{MD}}{C_{EM}^2} \quad (46)$$

and

$$C'_{MD} = \frac{C_{MD}}{1 - K_f^2} \quad (47)$$

where K_f^2 is called electromechanical coupling factor. Finally we combine acoustical, mechanical and electrical system. The whole analogous circuit is illustrated in Figure. 12 .

7.3 SA method

In order to improve the performance of a condenser microphone, SA method [27]-[29] is utilized to optimize the main parameter of diaphragm length (L_d), diaphragm thickness (T_d), the number of the acoustic holes (N) and holes side length (L_a). Then the performance indices of sensitivity and bandwidth can be improved.

In addition, the design variable and the constraints are given in the following

inequalities:

$$\begin{cases} 750 \text{ (um)} < L_d < 1500 \text{ (um)} \\ 1 \text{ (um)} < T_d < 2.5 \text{ (um)} \\ 64 < N < 196 \\ 15 \text{ (um)} < L_a < 30 \text{ (um)} \end{cases} \quad (48)$$

The SA algorithm is a generic probabilistic meta-algorithm for the global optimization problem, namely, locating a good approximation to the global optimum of a given function in a large search space. The major advantage of the SA is the ability of avoid becoming trapped in the local minima. In the SA method, each state in the search space is analogous to the thermal state of the material annealing process. The objective function G is analogous to the energy of the system in that state. The purpose of the search is to bring the system from the initial state to a randomly generated state with the minimum objective function. An improve state is accepted in two conditions. If the objective function is decreased, the new state is always accepted. If the objective function is increased and the following inequality holds, the new state will be accepted: [29]

$$P = \exp\left(-\frac{\Delta G}{T}\right) > \gamma, \quad (48)$$

where P is the acceptance probability function, ΔG is the difference of objective function between the current and the previous states, T is the current system temperature, and γ is a random number which is generated in the interval (0,1). In the high temperature T , there is high probability P to accept a new state that is “worse” than the present one. This mechanism prevents the search from being trapped in a local minimum. As the annealing process goes on and T decreases, the probability P becomes increasingly small until the system converges to a stable solution. The annealing process begins at the initial temperature T_i and proceeds with temperature that is decreased in steps according to

$$T_{k+1} = \alpha T_k, \quad (49)$$

where $0 < \alpha < 1$ is a annealing coefficient. The SA algorithm is terminated at the preset final temperature T_f . In the MEMS microphone optimization, we choose $T_i = 1000$, $T_f = 1 \times 10^{-9}$, and $\alpha = 0.95$.

In our problem, we wish to maximize the sensitivity and bandwidth of a MEMS condenser microphone. The goal is set up for the design optimization. It is hoped that the SPL in the working range is maximized, and the cut-off frequency is on 19-20k Hz. The compound objective function G can be written as

$$G = |20000 - BW| \times w_1 + |SEN| \times w_2 \quad (50)$$

where w_1 and w_2 are the weighting constant ($w_1 = 0.01$ and $w_2 = 0.05$), BW is the bandwidth and SEN is the sensitivity. With the SA procedure, the optimal solutions of diaphragm length, diaphragm thickness, acoustic holes number and the side length of the holes are 770um, 2.2um, 196 and 25um. The sensitivity and bandwidth of the MEMS microphone are -55.896V/Pa (dB) and 20.24kHz.

REFERENCES

- [1] ITU-R Recommendation BS.1387, “Draft Revision to Recommendation ITU-R BS.1387 - Method for Objective Measurements of Perceived Audio Quality”, ITU Radio communication Study Group 6, (1998).
- [2] M. Lavandier, S. Meunier and P. Herzog, “Identification of Some Perceptual Dimensions Underlying Loudspeaker Dissimilarities,” *J. Acoust. Soc. Am.*, vol. 123, pp. 4186 (2008).
- [3] X. H. Liu and B. S. Huang, “Sound Quality Analysis and Test to the Loudspeakers,” presented at The 16th National Conference on Sound and Vibration, Taipei, May 24, (2008).
- [4] X. H. Liu and H. X. Huang, “The influence of the audio processing to the sound quality parameters,” presented at The 17th National Conference on Sound and Vibration, Taipei, June 6, (2009).
- [5] E. Zwicker and H. Fastl, *Psychoacoustics: Facts and Models* (Springer, NY, 1999).
- [6] W. Klippel, “Loudspeaker Nonlinearities – Causes, Parameters, Symptoms,” presented at 119th Convention of the Audio Engineering Society, New York, 7-10 October, 2005.
- [7] W. Klippel, “Assessment of Voice-Coil Peak Displacement X_{max} ,” *J. Audio Eng. Soc.*, vol. 51(5), pp. 307-324 (2003).
- [8] J. Borwick, *Loudspeaker and Headphone Handbook* (Focal Press, Oxford, UK, 1994).
- [9] R. H. Small, “Direct-Radiator Loudspeaker System Analysis,” *J. Audio Eng. Soc.*, vol. 20(5), pp. 383-395 (1972).
- [10] W. Klippel, “Prediction of Speaker Performance at High Amplitudes,” presented

at the 111th Convention of the Audio Engineering Society, New York, 21-24 September, 2001.

- [11] W. Klippel, “Nonlinear Modeling of the Heat Transfer in Loudspeakers,” *J. Audio Eng. Soc.*, vol. 52(1/2), pp. 3-25 (2004).
- [12] E. R. Olsen and K. B. Christensen, “Nonlinear Modeling of Low Frequency Loudspeakers – A more Complete Model,” The 100th Convention Audio Engineering Society, Copenhagen, 11-14 May, 1966.
- [13] M. S. Bai, C. M. Huang, “Expert Diagnostic System for Moving-Coil Loudspeakers using Nonlinear Modeling,” *J. Acoust. Soc. Am.*, vol. 125, pp. 819 (2009).
- [14] T. Kohonen, “An Introduction to Neural Computing,” *Neural Networks*, 1, 3-16 (1988).
- [15] C. T. Lin and C. S. G. Lee, *Neural Fuzzy Systems* (Prentice-Hall, Englewood Cliffs, NJ, 1966).
- [16] W. Klippel, “Distortion Analyzer – a New Tool for Assessing and Improving Electrodynamic Transducer,” presented at the 108th Convention of the Audio Eng. Soc., Paris, February 19-22, (2000).
- [17] C. F. Juang and C. T. Lin, “An On-Line Self-Constructing Neural Fuzzy Inference Network and its Applications,” *IEEE Trans. Fuzzy Systems*, vol. 6(1), pp. 12-13 (1998).
- [18] J. H. Mathews and K. K. Fink, *Numerical Methods Using Matlab* (Prentice-Hall, Upper Saddle River, New Jersey, 1966).
- [19] ITU-R Recommendation BS.1116-1, “Method for the Subjective Assessment of Small Impairments in Audio Systems Including Multichannel Sound Systems”

- International Telecommunications Union, Geneva, Switzerland, (1994-1997).
- [20] ITU-R Recommendation BS.1534-1, “Method for the Subjective Assessment of Intermediate Sound Quality (MUSHRA)” International Telecommunications Union, Geneva, Switzerland, (2001).
- [21] D. Hohm and G. Hess, “A Subminiature Condenser Microphone with Silicon Nitride Membrane and Silicon Back Plate,” *J. Acoust. Soc. Am.*, vol. 85, pp. 476-480 (1989).
- [22] W. Kühnel and G. Hess, “A Silicon Condenser Microphone with Structured Back Plate and Silicon Nitride Membrane,” *Sensors and Actuators A*, vol. 30, pp. 251-258 (1992).
- [23] M. Pederson, W. Olthuis and P. Bergveld, “On the Electromechanical Behaviour of Thin Perforated Backplates in Silicon Condenser Microphones,” The 8th International Conference on Solid-State Sensors and Actuators, and Eurosensors IX., June 1995.
- [24] S. P. Timoshenko and S. Woinowsky-Krieger, *Theory of Plates and Shells*, (McGraw-Hill, New York, USA, 1959).
- [25] Z. Skvor, “On the Acoustical Resistance due to Viscous Losses in the Air Gap of Electrostatic Transducers,” *ACOUSTICA*. vol. 19, pp. 295-299 (1967/1968).
- [26] L. L. Beranek, *Acoustics* (McGraw-Hill, New York, USA, 1954).
- [27] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, “Equations of State Calculations by Fast Computing Machines,” *J. Chem. Phys.* vol. **21**(6), pp. 1087-1092 (1953).
- [28] A. Das and B. K. Chakrabarti, (Eds.), *Quantum Annealing and Related Optimization Methods* (Springer, Heidelberg, 2005).
- [29] J. De Vicente, J. Lanchares, and R. Hermida, “Placement by Thermodynamic

Simulated Annealing,” Phys. Lett. vol. **317**,pp. 415-423 (2003).



Table 1. The relations between nonlinear causes and symptoms of moving-coil loudspeakers.

Physical cause	HD_2	HD_3	IMD_2	IMD_3
Coil offset	X		X	
Coil height		X		X
Asymmetry in $C_{MS}(x)$	X			
Symmetrical limiting of suspension		X		
Asymmetry in $L_E(x)$			X	

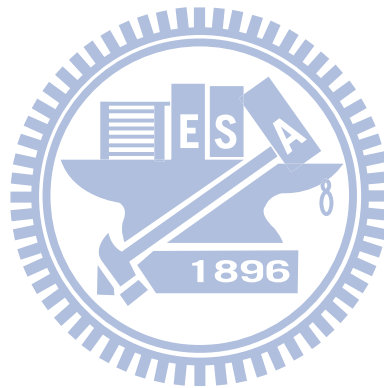


Table 2. The MANOVA output of the listening test for the fifteen nonlinear loudspeaker models. Cases with significance value p below 0.05 indicate that statistically significant difference exists among all cases.

	Significance value p			
	Fidelity	Fullness	Artifacts	Total preference
Cases	0.0000	0.0000	0.0000	0.0000

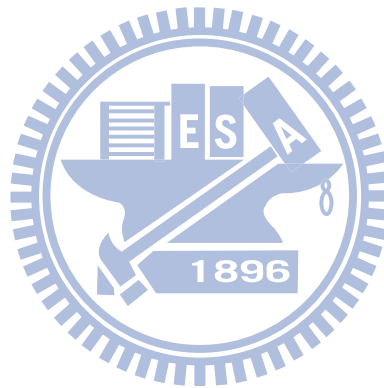


Table 3. Multiple regression of Total Preference in relation to Fidelity, Fullness and Artifacts.

	Fidelity	Fullness	Artifacts
Correlation coefficient	0.1147	0.2215	0.6717
Significance value p	0.0207	0.0000	0.0000

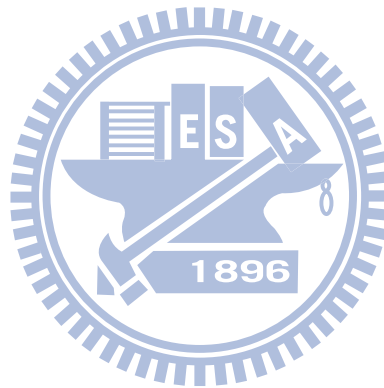


Table 4. Multiple regression analysis of Total Preference in relation to HD_2 , HD_3 , IMD_2 and IMD_3 .

	HD_2	HD_3	IMD_2	IMD_3
Correlation coefficient	-0.0416	-0.8864	0.1619	-0.0272
Significance value p	0.7887	0	0.1123	0.8429



Table 5. Fifteen nonlinear models with variations on speaker parameters of loudspeaker A. The first ten cases of objective index were selected as the input for training the ANN.

Cases	Four types of nonlinear distortion			
	HD_2 (%)	HD_3 (%)	IMD_2 (%)	IMD_3 (%)
Case1 $Bl(x) + C_{MS}(x)$	26	35	11	8
Case2 $C_{MS}(x)$	21	31	0	1
Case3 $Bl(x), \beta = 2$	14	23	16	21
Case4 $L_E(x)$	2	2	6	0
Case5 $C_{MS}(x), \varepsilon = 0.5$	45	25	0	1
Case6 $Bl(x), \varepsilon = 0.4$	22	12	24	9
Case7 $Bl'(x)$	4	2	0	0
Case8 $C_{MS}(x), \varepsilon = 0.2$	32	28	0	1
Case9 $Bl(x) + (C_{MS}(x), \beta = 2)$	21	54	9	5
Case10 $Bl(x), \varepsilon = 0.2$	18	12	19	9
Case11 $Bl(x)$	15	11	14	9
Case12 $C_{MS}(x) + L_E(x)$	20	52	6	2
Case13 $Bl(x) + (C_{MS}(x), \beta = 1.5)$	24	46	10	7
Case14 $Bl(x) + L_E(x)$	14	11	17	9
Case15 $L_E(x) + (C_{MS}(x), \beta = 1.5)$	21	43	6	1

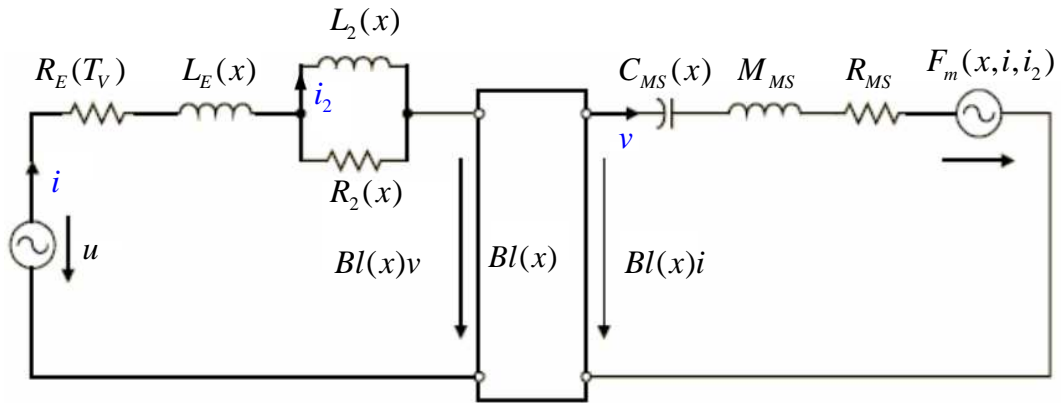
Table 6. The target output (listening test result) of the ANN. The first ten cases of objective index were selected as the target output data for training the ANN. The grading scale ranges from 1 to 5.

Case	Subjective indices			
	Total preference	Fidelity	Fullness	Artifacts
Case1 $Bl(x) + C_{MS}(x)$	2.82	3.00	2.45	2.91
Case2 $C_{MS}(x)$	2.00	2.27	1.82	1.91
Case3 $Bl(x), \beta = 2$	3.55	3.91	3.18	3.73
Case4 $L_E(x)$	4.18	4.09	3.82	4.00
Case5 $C_{MS}(x), \varepsilon = 0.5$	2.18	2.36	1.91	2.09
Case6 $Bl(x), \varepsilon = 0.4$	3.73	3.82	3.55	3.64
Case7 $Bl'(x)$	4.18	4.36	4.00	4.09
Case8 $C_{MS}(x), \varepsilon = 0.2$	2.55	3.00	2.55	2.55
Case9 $Bl(x) + (C_{MS}(x), \beta = 2)$	1.00	1.18	1.09	1.00
Case10 $Bl(x), \varepsilon = 0.2$	3.91	3.91	3.73	3.82
Case11 $Bl(x)$	3.91	3.82	3.36	3.91
Case12 $C_{MS}(x) + L_E(x)$	1.00	1.27	1.09	1.00
Case13 $Bl(x) + (C_{MS}(x), \beta = 1.5)$	1.73	1.83	1.45	1.64
Case14 $Bl(x) + L_E(x)$	3.18	3.55	3.27	3.36
Case15 $L_E(x) + (C_{MS}(x), \beta = 1.5)$	1.09	1.36	1.18	1.09

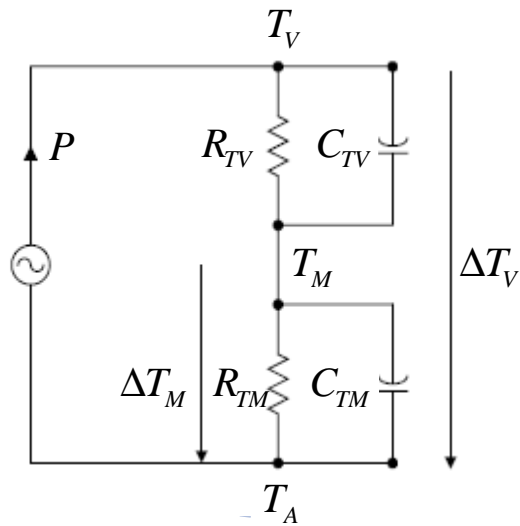
Table 7. Four nonlinear distortions, HD_2 , HD_3 , IMD_2 and IMD_3 , of the out-group test as the ANN inputs based on the Loudspeaker model B.

	Case1 $Bl(x)$	Case2 $le(x)$ $+(bl(x), \beta = 3))$	Case3 $Cms(x)$	Case4 $le(x)$ $+(Cms(x), \varepsilon = 2)$
HD_2 (%)	7	3	2	06
HD_3 (%)	19	20	1	32
IMD_2 (%)	6	0	14	10
IMD_3 (%)	1	0	1	5





(a)



(b)

Fig. 1 Electroacoustic analogous circuit of a moving-coil loudspeaker. (a) The equivalent circuit. (b) The circuit of the thermal model.

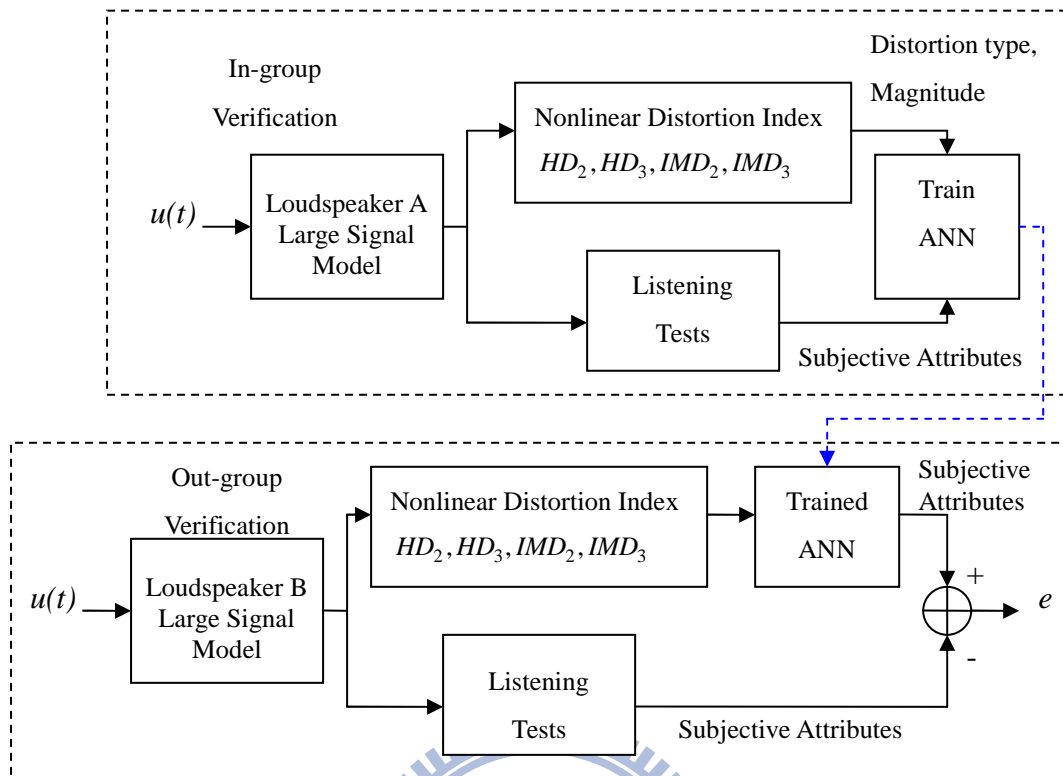


Fig. 2 The flowchart of the modeling and verification procedure for the ANN-based loudspeaker assessment system

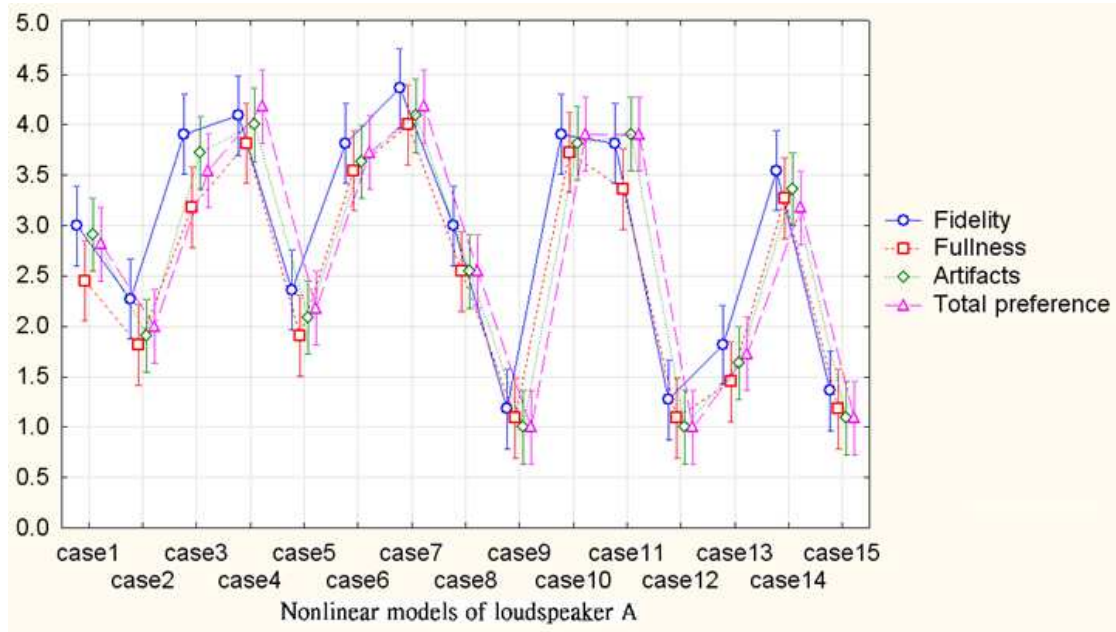


Fig. 3 The result of the listening test for the fifteen nonlinear loudspeaker models. The means and spreads (with 95% confidence intervals) of the grades are indicated in the figure. The x-axis and y-axis correspond to the nonlinear models and the grades, respectively.

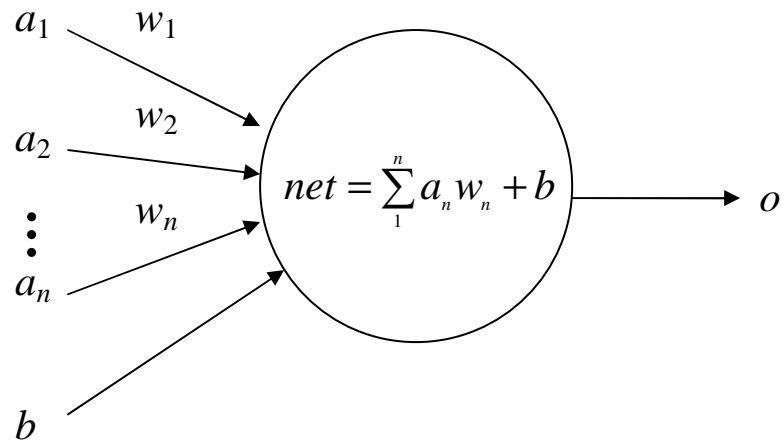


Fig. 4 The schematic of the neuron. The symbol a_n is the input, w_n is the weight, b is the bias, and o is output.

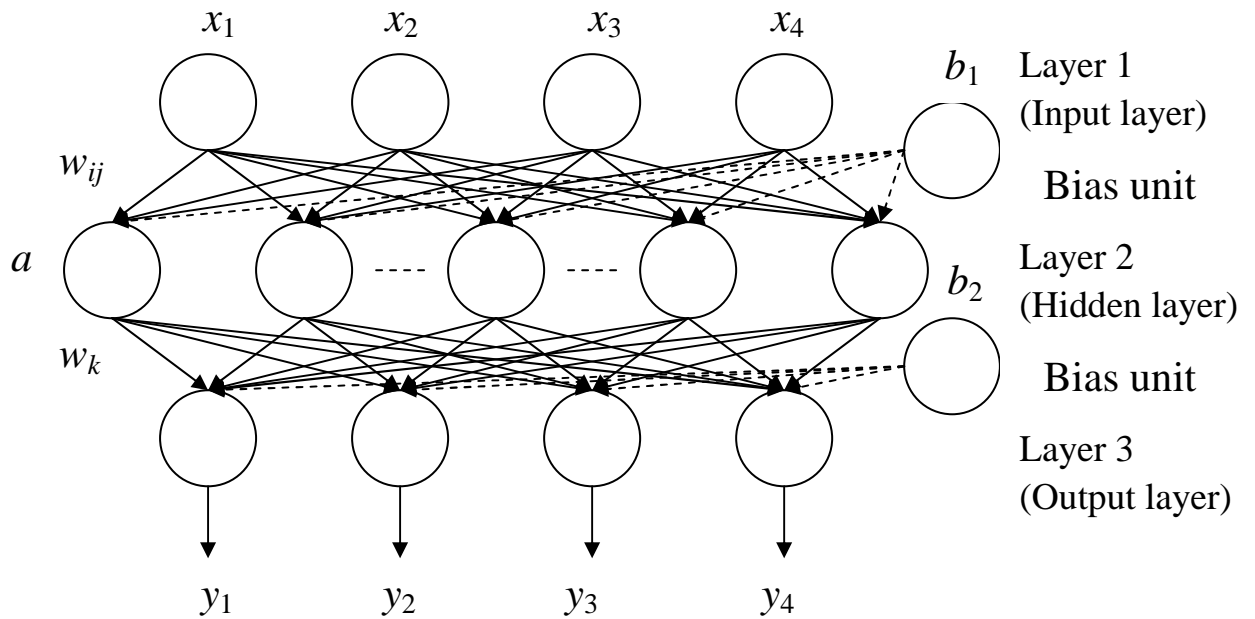


Fig. 5 The structure of the back-propagation ANN. The symbol a is the hidden layer, a hidden layer of fifteen neurons, x_i is the input layer of the nonlinear distortion measures HD_2, HD_3, IMD_2 and IMD_3 , w_{ij} and w_k are the weights, b_1 and b_2 are the bias units, and y is the output of the subjective attributes: Fidelity, Fullness, Artifacts and Total Preference.

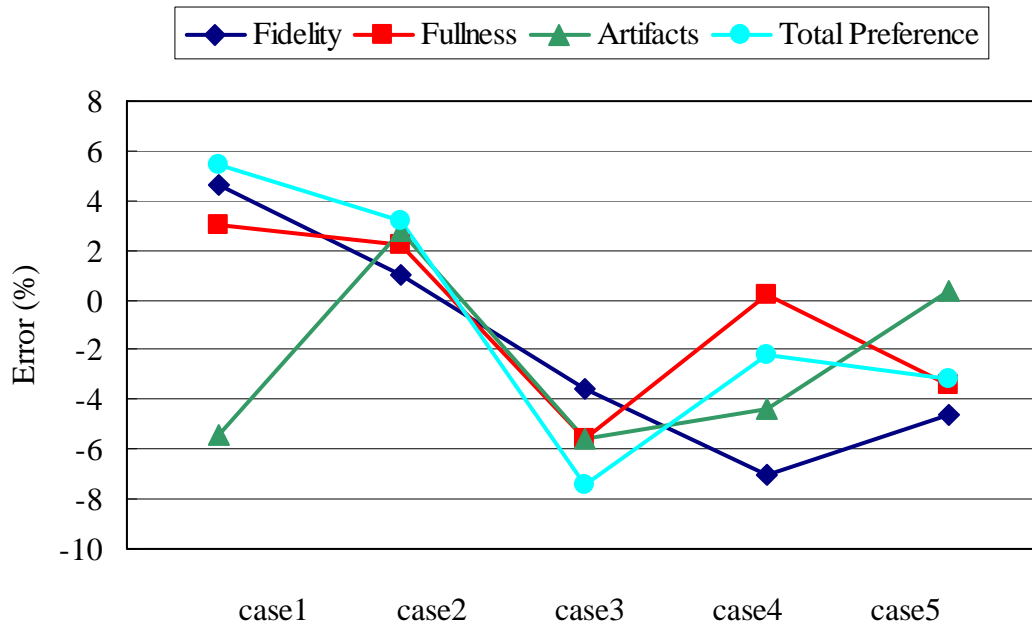


Fig. 6 Comparison of Total preference, Fidelity, Fullness and Artifacts predicted by the ANN and the subjective data obtained from the listening test for the last five cases (in-group verification).

$$Error = \frac{Output - Target}{5}$$

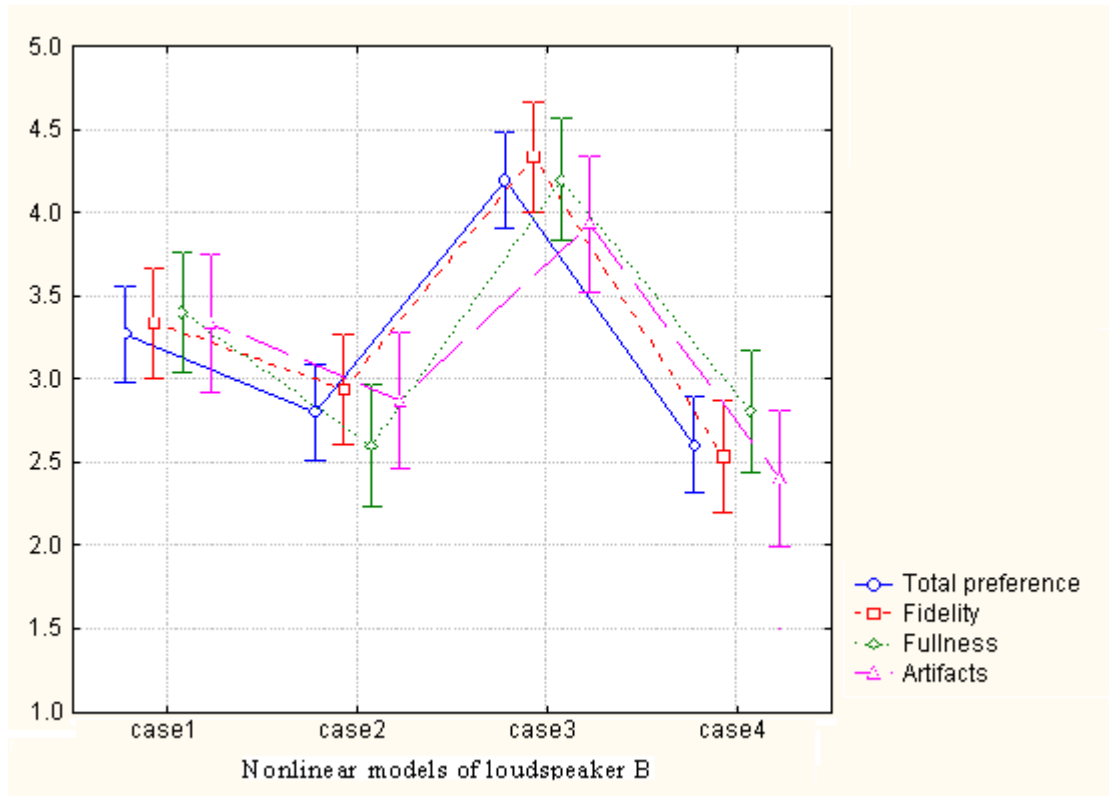


Fig. 7 The out-group test result of the listening test for the four nonlinear loudspeaker models. The means and spreads (with 95% confidence intervals) of the grades are indicated in the figure. The x -axis and y -axis correspond to the nonlinear models and the grades, respectively.

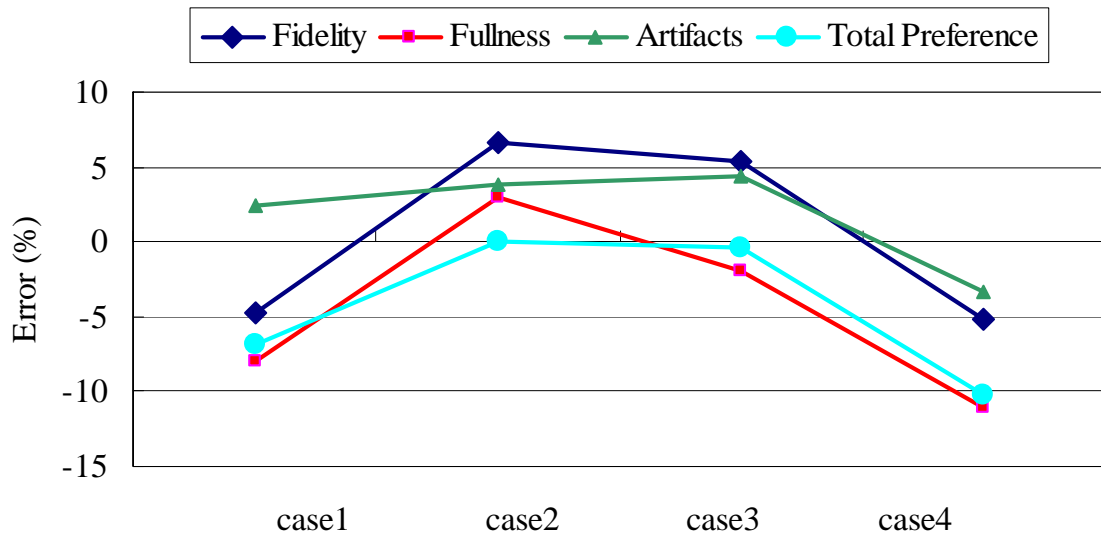


Fig. 8 The results of the out-group test of the ANN based on the Loudspeaker model B. Total preference, Fidelity, Fullness and Artifacts predicted by the ANN and the subjective data obtained from the listening test are compared.

$$Error = \frac{Output - Target}{5}$$

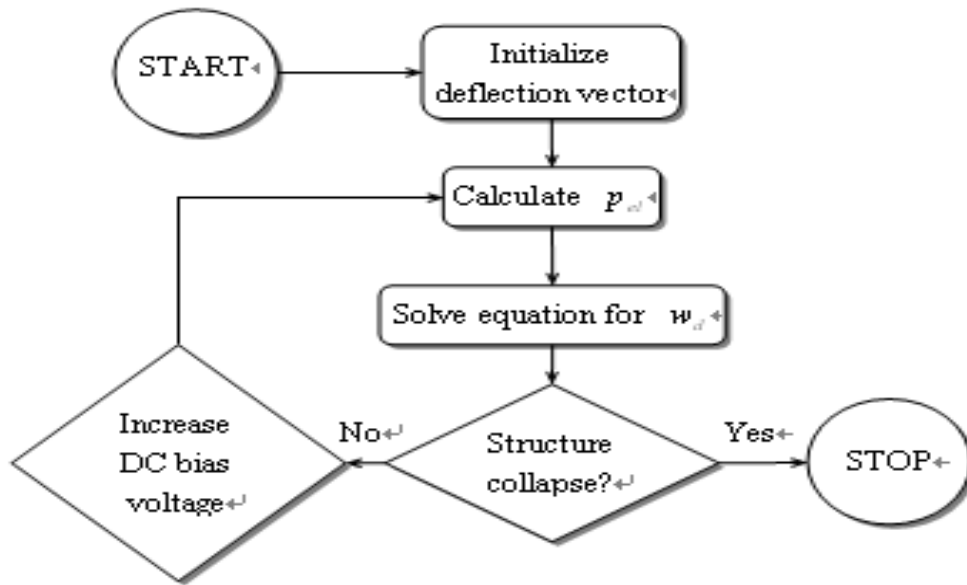


Fig. 9 The flow chart of the iterative procedure from FDM method for the coupled equation system



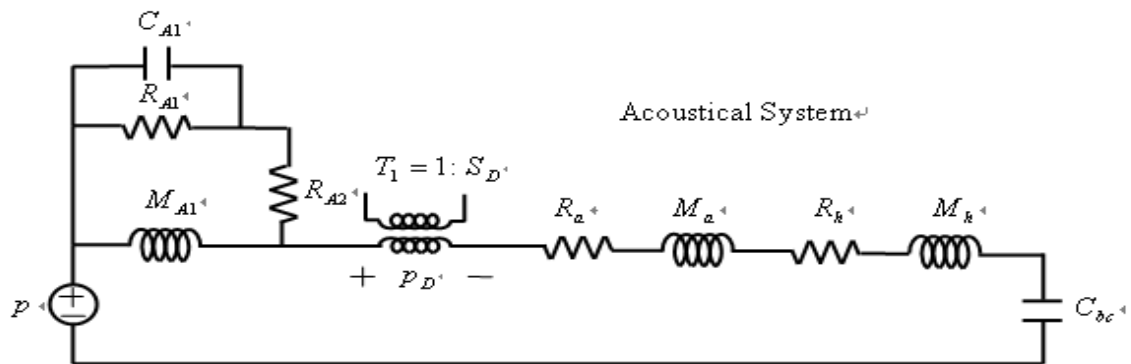


Fig. 10 Acoustical system includes the sound radiation, air gap, back chamber and acoustical holes influence.



Mechanical System

Electrical System

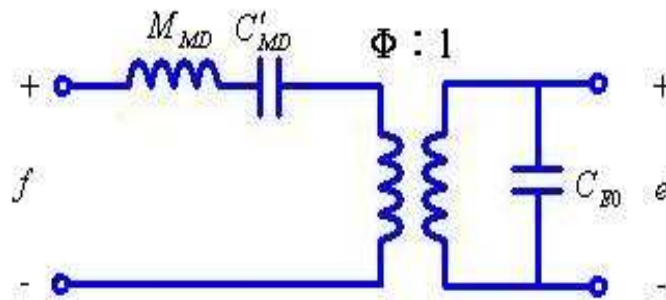


Fig. 11 Combination of mechanical and electrical system.

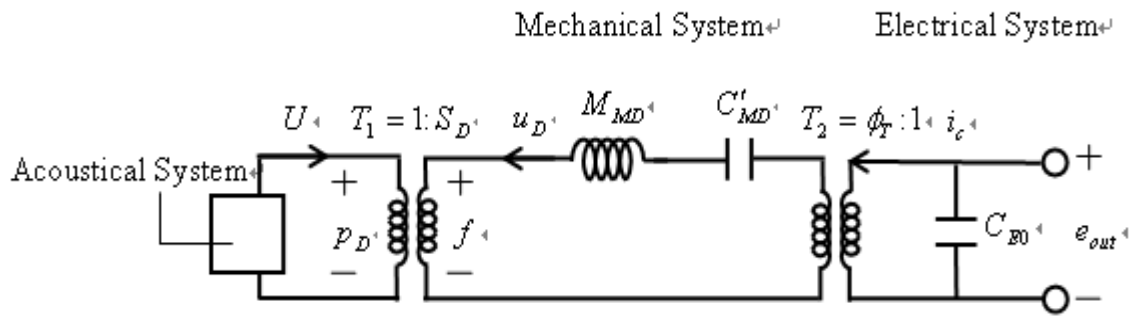


Fig. 12 The comprehensive system combines the acoustical, mechanical and electrical system.

