LOUDSPEAKER EQUALIZATION BASED ON MULTI-LOCATION OBSERVATION WITH RELIABLE TIME-FREQUENCY REGION SELECTION AND ITS EVALUATION USING SOUND PROPAGATION MEASUREMENT

Masanori Morise and Hideki Kawahara

Faculty of Systems Engineering Wakayama University
930 Sakaedani, Wakayama, Wakayama 640-8510, Japan
phone: +81 73 457 8461, fax: +81 73 457 8112, email: s055068@sys.wakayama-u.ac.jp
web: www.sys.wakayama-u.ac.jp/~kawahara/index-e.html

ABSTRACT

A new loudspeaker equalization method that is robust against listening environment and listener's movement is proposed based on a new criteria for making use of multi location measurement and reliable time-frequency region selection. The time-frequency region selection is implemented by low-pass filtering operation on an nonlinearly stretched time axis for computational efficiency. The proposed criteria for equalization is based on average energy distribution and average temporal spread of equalized multiple responses. A high-density measurement of wave propagation radiated from the loudspeaker under study was used for evaluating and demonstrating effectiveness of the proposed method.

1. INTRODUCTION

In spite of recent advancement of sound reproduction based on digital signal processing, sound coloring due to loudspeakers is still a difficult problem. Primary difficulty is due to limited degree of freedom in control. Usual loudspeakers only have one driving port while reproduced sound can be observed at various spatial locations at the same time. The other difficulty is zeros close to unit circle in the discrete-time transfer functions. Number of procedures using multiple loudspeakers and multiple observation points were proposed to implement desired transfer functions [3, 7]. They are handling loudspeaker characteristics and room acoustics as a whole and reduced the target problem into linear algebra. While this approach is powerful and conceptually simple to solve both problems at the same time, their effectiveness is limited within frequency region where dimensions of source and observation point spacing are relatively small to wavelength. The approach proposed in this article tries to divide sound reproduction problem into two disjoint problems; loudspeaker equalization and sound field control. This article focuses on the first problem. The solution to the first problem can be general enough to be applicable to conventional sound field control methods and enables to improve their effectiveness.

2. CRITERIA FOR LOUDSPEAKER EQUALIZATION

The problem to be solved is equalization of a SIMO (single input and multiple output) system. For solving this problem, it is necessary to establish dependable criteria for gathering multiple output and designing cost function. Equalization of a SISO (single input and single output) system is basically an inverse filtering problem. It doesn't have any gathering problem. The basic LMS (least mean square) criterion can be used to derive the necessary cost function. Major difficulties in this framework are transfer function zeros close to unit circle and causality. MINT [3] was proposed to solve this problem by transforming the original SISO problem into MIMO (multiple input and multiple output) or MISO (multiple input and single output) problem taking advantage of spatial nature of acoustics.

The ideal target for SIMO equalization is to make observed impulse responses at any locations be impulse with propagation delay. However, this target cannot be attained in real world situations, because transfer functions to different observation points generally cannot be the same. Application of SISO oriented criteria to one specific location in SIMO problem introduces severe distortions in responses at other observation points. In other words, there is a conflict between observation points about equalization target. The proposed criteria was designed to embody arbitration mechanism for this conflict.

The criteria consists of two requirements; after equalization, averaged energy at each frequency has to distribute uniformly across frequencies, and averaged duration, standard deviation around the temporal energy centroid, of equalized responses has to be minimum. The requirement on duration introduces other problem to be solved prior to solve this SIMO problem, because duration is highly susceptible to background noise and reflections. The problem is how to measure reliable impulse responses at multiple observation points in real world situations.

3. RELIABLE IMPULSE RESPONSE MEASUREMENT

Various generalized stretched pulse-based methods have been proposed for impulse response measurement in real world situations [5, 1, 4, 6]. Those methods dramatically improve the signal to noise ratio. However, it is not enough when applied for estimating duration of impulse responses of a loudspeaker system in an realistic environment, because of parabolic weighting on response components distant from the primary response. If we have an access to prior knowledge about the time-frequency region where direct response is dominant, recovering the impulse response only using components residing within such reliable region. This time-frequency selection and reconstruction of the impulse response can be implemented using wavelet transform or other joint time-frequency representations, there exist an efficient implementation using nonlinear time stretching.

3.1 Boundary frequency and time stretching

Assume that a reliability index is defined in the time-frequency domain $r(\omega,t)$. Then, by setting the threshold of desired reliability (for example α), the boundary of the reliable time frequency region is implicitly defined by the following equation:

$$\alpha = r(\omega,t). \tag{1}$$

It is also possible to represent this boundary as a function of time $\omega_B(t)$. Let call the upper boundary of this function as $\omega_{UB}(t)$. Defining the new time axis τ by stretching the original time axis by the amount of $\omega_{UB}(t)/\omega_L$, the boundary in this new time axis is given by $\omega_{UB}(\tau) = \omega_L$, a constant. In other words, the desired

This research is supported partly by Grant in Aids for Scientific research (B) 14380165 and the Telecommunications Advancement Organization of Japan. The second author is also an invited researcher of ATR, Japan.

selection is implemented by time axis stretching and low-pass filtering. The new time axis is explicitly represented by the following equation:

$$\tau(t) = \int_{t_0}^{t} \frac{\omega_{UB}(\lambda)}{\omega_L} d\lambda. \quad (2)$$

Practically, the constant is calculated from the lowest audible frequency $f_L = 20$Hz; ($\omega_L = 2\pi f_L$).

3.2 Designing time stretching function

An exemplar design process of a time stretching function is illustrated here. First assumption is that disturbance to the measurement is only due to background noise. The next assumption is that the target impulse response consists of damped sinusoid, impulse response of a second order system, having the same damping factor ζ and different characteristic frequencies f_k. The initial position of each damped sinusoids distribute in $(0, t_s)$. These are first order approximation of loudspeaker measurement in realistic environment. The last assumption represents multi-path effect due to the size of the sound radiating body. One reasonable selection of the reliability index in this case is S/N. By assuming S/N threshold to r, the new stretched time axis $\tau(t)$ to implement the desired time-frequency selection is yielded as follows.

$$\begin{aligned} \tau(t) &= \frac{\log(r)}{2\pi\zeta f_L}(\log(t-t_s) - \log(t_1)) + \frac{f_s(t_1+t_s)}{2f_L} \quad (3) \\ t_1 &= \frac{\log(r)}{\pi f_s \zeta} \end{aligned}$$

where f_s represents the sampling frequency and t_s represents the maximum propagation lag proportional to the size of the sound radiating body. This equation holds for $t > t_1 + t_s$. For $0 < t < t_1 + t_s$, it is reasonable to set effective cut-off to the Nyquist frequency. The inverse function $t_r(\tau)$ of this warping is easily represented in a closed form.

3.3 Experiments

Effects of the proposed post processing parameters ζ and t_s are illustrated by impulse response measurement of a compact loudspeaker system (JBL CM-40). A miniature omni-directional condenser microphone was placed at 40 cm in front of the woofer cone. The impulse response was measured using a cyclic M-sequence [5] with 131071 samples for one period. The microphone output was amplified and sampled at 44.1 kHz 16 bit. Measurements were conducted in an sound proof room (YAMAHA AVITECS-AMG, volume: 17 m^3). The background noise level was 20 dB (A weighting) when air conditioning is turned off.

Figure 1 shows effects of ζ on amplitude response. Note that the top plot in the figure without post processing has noisy peaks and dips due to background noise, reverberation and reflections. Larger ζ values introduce stronger smoothing in the frequency response. This effect has an important role in the optimization of SIMO problem as shown in the next section.

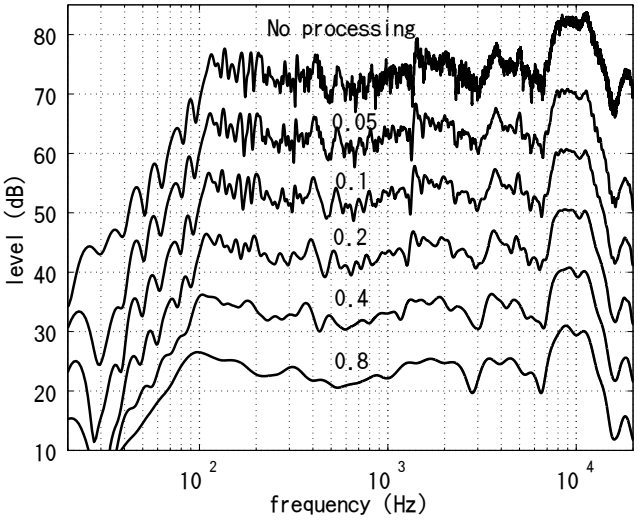

Figure 1: Effects of assumed dumping rate ζ on amplitude response. The other parameters were kept constant. ($t_s = 1$ms and $r = 1000$). The vertical location is arbitrary. Each plot is separated by 10 dB.

4. MULTI LOCATION EQUALIZATION

The criteria proposed in the previous section is formulated in the following equations. The goal for this SIMO problem is to design a linear time invariant filter $F(\omega)$ to shape multiple outputs to optimize the criteria. The first requirement on energy distribution in the frequency domain is rewritten taking physical constraints into account. The goal is to minimize the cost function $L_A^2(\omega)$.

$$L_A^2(\omega) = G(\omega)^2 - \frac{|F(\omega)|^2}{N}\sum_{k=1}^{N} B_k(\omega)^2 \quad (4)$$

$$B_k(\omega)e^{j\psi_k(\omega)} = \mathscr{F}[s_k(t)] \quad (5)$$

where $s_k(t)$ is a post processed impulse response measured at k-th observation location and N represents the number of measuring points. $\mathscr{F}$ also represents Fourier transform. Physical constraint is built into the target energy distribution $G(\omega)^2$, for example, by restricting effective bandwidth to avoid overloading to speakers.

The second design goal is to minimize the averaged response duration $\bar{\sigma}$ defined in the following equations.

$$\bar{\sigma} = \sqrt{\frac{1}{N}\sum_{k=1}^{N}\sigma_k^2} \quad (6)$$

$$\sigma_k^2 = \int (t - \langle t \rangle_k)^2 |y_k(t)|^2 dt \quad (7)$$

where $\langle t \rangle_k$ is the energy centroid of the equalized response $y_k(t)$ at the k-th observation point. Note that normalization factor and the boundary of integration in this weighted integration is omitted for avoiding unnecessary complexity. The following notation is also introduced.

$$A_k(\omega)e^{j\phi_k(\omega)} = F(\omega)B_k(\omega)e^{j\psi_k(\omega)} = \mathscr{F}[y_k(t)] \quad (8)$$

The duration $\sigma_k(t)$ can also be calculated in the frequency domain using group delay $-\phi_k'(\omega)$ ("$'$" represents frequency derivative) and the amplitude response $B_k(\omega)$ [2].

$$\begin{aligned} \sigma_k^2 &= \int \left(\frac{A_k'(\omega)}{A_k(\omega)}\right)^2 A_k^2(\omega) d\omega \\ &\quad + \int (\phi_k'(\omega) - \langle t \rangle_k)^2 A_k^2(\omega) d\omega \quad (9) \end{aligned}$$

This representation and Eq. 4 is used to derive the optimum equalization filter $F(\omega)$. For the first step of the optimization process, $\psi_R(\omega)$ is directly calculated as the weighted average of each group delay based on the second term of Eq. 9. Similarly, optimization of Eq. 4 provides the weighted root mean square average as the optimum amplitude equalization $|F(\omega)|$. Those values can be calculated independently in terms of ω. The second term of Eq. 9 introduces additional constraint on the smoothness of the amplitude response of the equalizer $|F(\omega)|$. This term is indirectly controlled by measuring parameters r, t_s and ζ.

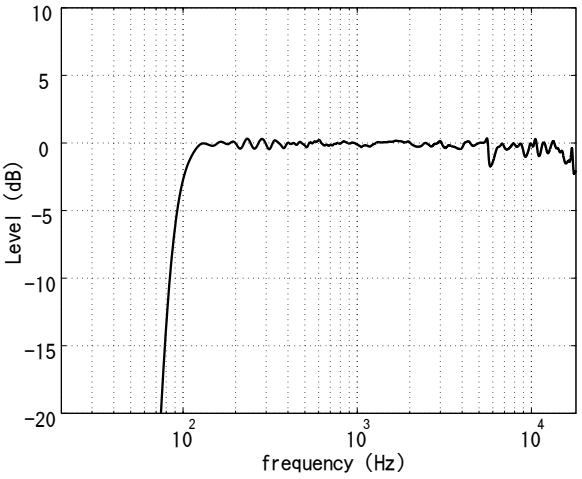

Figure 2: Averaged amplitude frequency response after equalization using 12 measuring points for equalizer design.

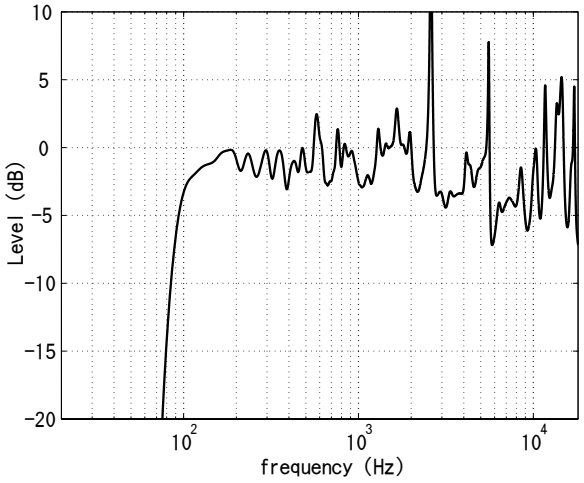

Figure 3: Averaged amplitude frequency response after equalization using only one measuring point for equalizer design.

Based on these preparations, the compensating filter $F(\omega)$ is calculated by the following equation.

$$F(\omega) = \frac{G(\omega)}{\sqrt{\frac{1}{N}\sum_{k=1}^{N} B_k^2(\omega)}} e^{-j\psi_R(\omega)} \quad (10)$$

$$\psi_R(\omega) = \int \tau_g(\omega) d\omega$$

$$\tau_g(\omega) = -\frac{\sum_{k=1}^{N} \psi_k'(\omega) B_k^2(\omega)}{\sum_{k=1}^{N} B_k^2(\omega)}$$

4.1 Experiments

The proposed equalization method was tested using the same compact loudspeaker system. A robot controlled measuring stage (Nittobo Onkyo MT-2000-DC2) was used for positioning a miniature microphone (Sony ECM-77S). The other conditions were the same as the pervious experiments.

In the first experiment, the microphone was positioned on two axes (X axis: horizontal, and Y axis: vertical) on the plane parallel to the front panel of the loudspeaker system at a distance of 40 cm. The origin of the coordinate was set just in front of the center of the woofer cone. Measuring points were placed in 8 mm steps and were spanning 0 mm to 400 mm on each axis. The total number of measuring points was 101.

4.1.1 Amplitude characteristics

By using all measuring points, $G(\omega)$ was attained precisely. By removing measuring points in designing the equalization filter, its effect on the construction $L_A^2(\omega)$, that is calculated using all 101 measuring points, was investigated. It was found that the averaged amplitude frequency responses after equalization were virtually the same as the design target even when a relatively small number of measuring points for design were used. Figure 2 shows the equalized averaged response using 12 measuring points. The rolling off characteristics around 100 Hz and 20 kHz are due to $G(\omega)$ design. However, when only one measuring point was used to design the equalizing filter, severe distortion was introduced to the averaged characteristics. Figure 3 shows an example using one measuring point, the origin in this example, for design. The distortion is dependent on the measuring point used for the equalizer design and varies very much. This is a good demonstration indicating how harmful SISO formulation for loudspeaker equalization is.

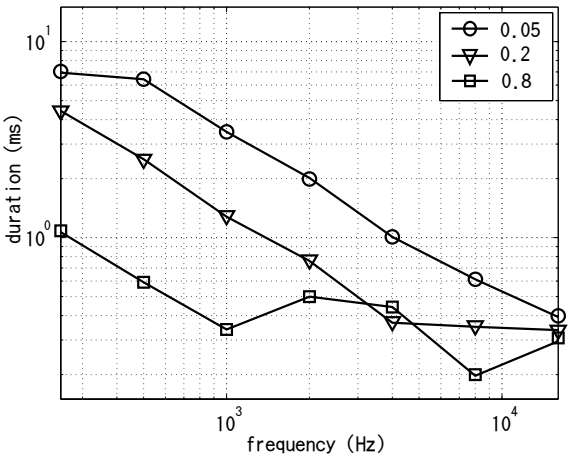

Figure 4: Squared average duration for each octave band. The damping factor ζ was annotated to each line.

4.1.2 Band-wise duration

For duration equalization, the squared average duration after equalization $\bar{\sigma}^2$ was investigated as a function of the number of measuring points for equalizer design. It was indicated that by using only 4 point, the squared average duration attained its asymptotic value when $\zeta = 0.2$ was used.

To investigate duration equalization characteristics in terms of post processing parameters, it is better to rewrite the squared average duration explicitly.

$$\bar{\sigma}_B^2(\omega_c) = \sum_{k=1}^{N} \int_{\omega_L(\omega_c)}^{\omega_H(\omega_c)} \left(\frac{A_k'(\omega)}{A_k(\omega)}\right)^2 A_k^2(\omega) d\omega + \sum_{k=1}^{N} \int_{\omega_L(\omega_c)}^{\omega_H(\omega_c)} (\phi_k'(\omega) - \langle t \rangle_k)^2 A_k^2(\omega) d\omega, \quad (11)$$

where ω_c represents the center frequency of a frequency band $[\omega_L(\omega_c), \omega_H(\omega_c)]$ for evaluation.

Figure 4 shows dependencies of $\bar{\sigma}_B^2(\omega_c)$ on ζ in each octave band. Note that the averaged durations are decreasing inversely proportional to the center frequency in lower frequency portion of each

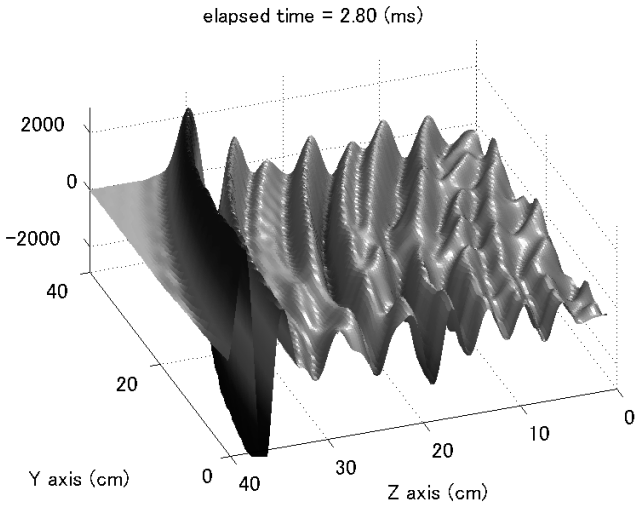

Figure 5: Three dimensional view of the pressure waveform of the virtual impulse drive at 2.80 ms from the stimulation.

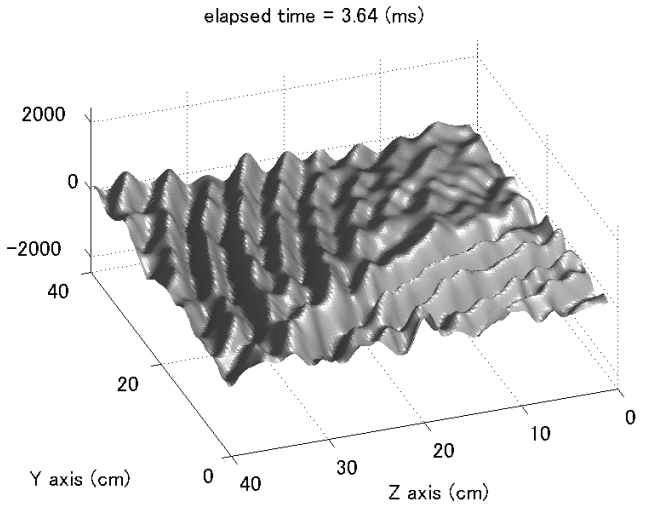

Figure 6: Three dimensional view of the pressure waveform of the virtual impulse drive at 3.64 ms from the stimulation.

line. This indicates that the truncation due to ζ effectively suppresses effects of background noise and other uncorrelated signals in those frequency region. This trend disappears in the higher frequency portion of each line and the boundary is dependent on ζ. This indicates that the optimum ζ has to be frequency dependent.

4.1.3 Wave propagation

To investigate the source of these behavior and the physical meaning of the proposed optimization, next experiments were conducted. The loudspeaker system tested here was another compact system, BOSE 101.

The measuring points this time were placed on the Y-Z plane, where Z axis is normal to the front panel and starting from the center of the woofer cone. The origin of the measurement was placed 40 cm from the front panel and the positive direction was defined outward. Measuring points were placed on two dimensional grid points with 8 mm steps spanning 400 mm for each direction. The total number of measuring points was 2500.

Two channel recording was done for each measuring point; one is for microphone output and the other is driving signal of the loudspeaker system. Using this set up, impulse responses were aligned on the common time coordinate and converted into a three dimensional movie visualization of the pressure waveform of the virtual impulse drive. Two dimensional cubic spline interpolation was employed to remove visual aliasing effects due to spatial sampling.

Figures 5 and 6 show two snapshots from the movie at two instants; 2.80 ms and 3.64 ms from the virtual stimulation. As shown in these figures, direct sounds only be observed at the very beginning and easily are contaminated by reflected sounds. It was also observed in the movie that clearly identifiable reflections are still existing even after 30 ms from the stimulation. The time-frequency region selection effectively eliminates contributions of these reflections on high-frequency equalization.

This wave propagation experiment and its scientific visualization lead to deeper understanding of the loudspeaker equalization problem based on SIMO formulation. It also suggests alternative verification tests for further study.

5. DISCUSSION

The band-wise duration experiments revealed that the boundary derived from the constant ζ assumption is not appropriate. The results indicated that effective values of ζ that introduce different behavior in the equalized band-wise duration are 0.8 around 1 kHz, 0.2 around 4 kHz and 0.05 around 16 kHz. In other words, the effective ζ should be reduced when the frequency is increased. This observation would be utilized to provide the objective method for determining the best stretching function depending on the data.

6. CONCLUSION

New criteria for loudspeaker equalization based on multi location observation was proposed with a post processing method for reliable impulse response measurements. The method is straightforward and found to be able to attain asymptotic performance with relatively small numbers of observation points. The proposed method is transparent and is applicable to other MIMO techniques for sound field control. It may also be applicable to other acoustic equalization where SIMO formulation is relevant such as headphone equalization. Subjective evaluation of the proposed method is for the further study and to be reported elsewhere.

REFERENCES

[1] Nobuharu Aoshima. Computer-generated pulse signal applied for sound measurement. *J. Acoust. Soc. Am.*, 65(5):1484–1488, 1981.

[2] L. Cohen. *Time-frequency analysis*. Prentice Hall, Englewood Cliffs, NJ, 1995.

[3] N. Miyoshi and Y. Kaneda. Inverse filtering of room acoustics. *IEEE Trans. on ASSP*, 36(2):145–155, 1988.

[4] D. D. Rife and J. Vanderkooy. Transfer function measurement with maximum-length sequences. *J. Aud. Eng. Soc.*, 37(6):419–444, 1989.

[5] M. R. Schroeder. Integrated impulse method measuring sound decay without impulses. *J. Acoust. Soc. Am.*, 66(2):497–500, 1979.

[6] Y. Suzuki, F. Asano, H. Kim, and T. Sone. Computer-generated pulse signal suitable for the measurement of very long impulse responses. *J. Acoust. Soc. Am.*, 92(2):1119–1123, 1995.

[7] H. Tokuno, O. Kirkeby, P. A. Nlson, and H. Hamada. Inverse filter of sound reproduction system using regularization. *IEICE Trans. Fundamentals*, E80-A(5):809–820, 1997.