K. Chiao, N. Harris, and C. Kyriakakis
Integrated Media Systems Center
University of Southern California
Los Angeles, CA 90089-2564, USA

Presented at the 109th Convention 2000 September 22-25 Los Angeles, California, USA

A NEW APPROACH TO SPEAKER/ROOM EQUALIZATION

K. Chiao, N. Harris, and C. Kyriakakis

Integrated Media Systems Center
University of Southern California
3740 McClintock Ave., EEB 432
Los Angeles, California 90089-2564, USA

and

ABSTRACT

In this paper we examine methods for digital room correction and loudspeaker equalization as they apply to Distributed Mode flat-panel loudspeakers. We present a method that combines LPC inverse filtering and tunable DWT octave-band equalization. Furthermore, we discuss frequency response and time-domain smearing/spread issues and considerations for real-time implementation.

1. INTRODUCTION

Equalization of loudspeakers to correct for the effects of room acoustics is critical for many multimedia applications such as telepresence, home entertainment, and gaming. In this paper we examine the performance of current equalization methods and also present a new approach that combines parametric and non-parametric methods. Our method is based on a tunable non-parametric equalizer that can be used for real-time equalization. The main application of interest is multi-participant telepresence using life-size video projection screens arrayed in front of the local participant. One of the problems with such systems in the past has been the mismatch between picture and sound that arises when the sound is rendered using traditional loudspeakers placed at the side of the projection screen. In our approach, we use a novel Distributed Mode flat-panel loudspeaker that also serves as the projection screen (Fig. 1). The Distributed Mode Loudspeaker (DML) has a wide directivity and a somewhat diffuse radiation [1][2]. While this solves the mismatch problem, it is still necessary to perform digital equalization in order to achieve high-fidelity sound at the listening position.

Figure 1. Three-screen immersistation for multi-participant conferencing. Each of the three screens is a distributed mode loudspeaker, thus achieving a perfect spatial match between picture and sound.

2. DESIGN ISSUES

Parametric equalization methods that combine constant Q boost and variable Q cut equalizers are commonly used for sound reinforcement, studio recording, and movie theater systems. These are typically based on critical band 1/3 octave analysis that is supported by psychoacoustics evidence [3]. Recent developments of audio DSP methods have also led to non-parametric processing of audio signals that allow more control and flexibility in the frequency-domain.

The problems associated with inverting a room impulse response were examined by Neely and Allen [4], and methods for achieving this by making corrections to the allpass part of the target impulse response were proposed by Mourjopoulos [5]. Methods using frequency warping techniques can also be used [6, 7].

The performance of any loudspeaker is compromised because of room effects, although these effects may be less audible for the DML than for conventional loudspeakers, due to their more diffuse radiation [8]. Room modes can cause very audible distortion particularly in the lower frequency band response. Therefore it is of little use to invert the anechoic impulse response if the loudspeaker is to be used in a reverberant environment.

For near field applications such as the one described earlier, the impulse response of the direct sound dominates. However, in the far field situation, the direct response combines with early reflections to determine tonal balance. The effects of such reflections on sound quality has been studied extensively [9, 10, 11, 12] and it has been shown that they are the dominant source of monitoring non-uniformities. These non-uniformities appear in the form of colorations (frequency response anomalies) in rooms with an early reflection level that exceeds –15 dB spectrum level relative to the direct sound for the first 15 ms [12,13]. Such a high level of reflected sound gives rise to comb filtering in the frequency domain that in turn causes severe changes in timbre. The perceived effects of such distortions were not quantified until psychoacoustic experiments demonstrated their importance.

In our work we examine digital inverse filtering by LPC all-pole modeling for room modes correction. Filter lengths must be chosen to be long enough to allow good resolution at low frequencies. Another issue is time-domain smearing that arises when the equalization filters are convolved with audio signals.

Too long an equalization filter can introduce echoes in high frequencies and comb filter effects at lower frequencies, even though the frequency-domain correction is perfect. In order to overcome that, we use minimum phase inverse filters to minimize delay.

3. IMPLEMENTATION USING DISTRIBUTED MODE LOUDSPEAKERS

In our case study, we used a distributed-mode loudspeaker (DML) instead of a conventional loudspeaker. We measured the room impulse response (RIR) from the DML to the listening position. The length of the RIR is 7226 points (about 164ms) at a 44.1 kHz sampling rate. For the telepresence application described above in which several participants communicate in the same shared environment, the main program material is speech. As such, we focused our efforts in the 100 Hz to 8 kHz frequency range.

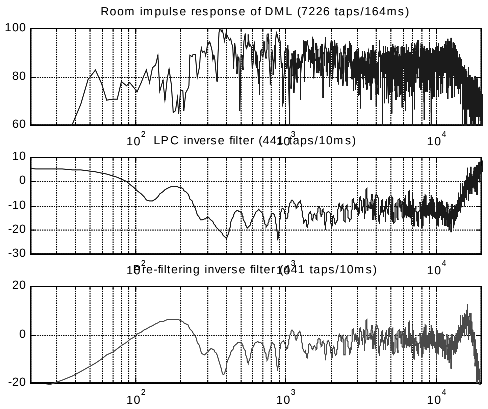

Figure 1

In Fig. 1, the top plot is the room impulse response of the DML loudspeaker, and it can be seen that there is a resonance dip around 200Hz, as well as a resonance peak around 52Hz. Both are DML resonance modes and are not due to the room modes. The 52Hz resonance is due to the exciter, and the 200Hz resonance is due to the panel being mounted in a shallow box. In order achieve good performance of the inverse LPC FIR filter we used a 441-tap inverse filter that allowed us to shape the frequency response down to 100 Hz, but this did not affect the 52 Hz peak. That was handled using a separate 441-tap band-pass linear phase filter. The middle subplot in Fig. 1 shows the frequency response of the inverse FIR filter derived using the Linear Predictive Coding (LPC) method. The third subplot in Fig. 1 shows the corrected frequency response that filters out the 52Hz resonance of the DML loudspeaker without introducing any time-domain smearing.

After the pre-filtering stage, we feed our 441-tap band-limited FIR inverse filter into a multi-resolution system to achieve the desired gain control in each subband. We chose the maxflat (Daubechies) set of orthogonal filters that have maximum flatness which is optimal for audio signal processing. In other words, we can decompose the band-limited inverse filter by convolving it with these orthogonal filters and then downsampling to get the wavelet coefficients for each subband [14]. To find the decomposed filter components (subfilters), we convolve each wavelet coefficient with the same orthogonal filter and then upsample by the same factor [15] since we use orthogonal filter banks. We choose the Dyadic form of the Discrete Wavelet Transform (DWT) to achieve octave-band gain control similar to parametric equalizers. In order to achieve good octave-band gain control it is desired to have steep attenuation to minimize the overlapping area of adjacent subfilters. By choosing 32-point maxflat orthogonal filters, the errors can be kept within ±5dB.

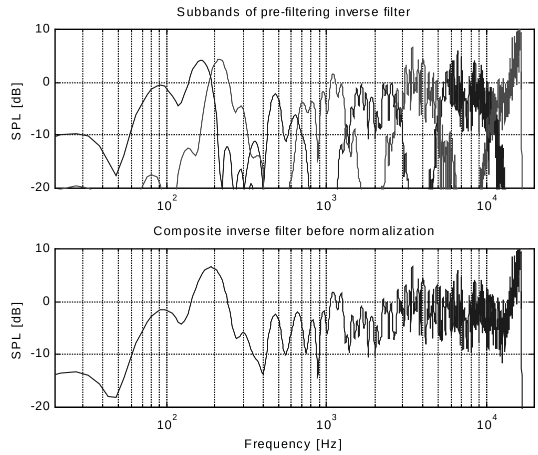

Figure 2.

The band-limited inverse filter is decomposed into 9 subfilters consisting of an eight-level DWT and a one-level DWT for the highest octave band:

D1D1 (detail 1): 16.5375 kHz-22.050 kHz

D1A1 (detail 1): 11.025 kHz-16.5375 kHz.

D2 (detail 2): 5.513 kHz-11.025 kHz.

D3 (detail 3): 2.756 kHz-5.513 kHz.

D4 (detail 4): 1.378 kHz-2.756 kHz.

D5 (detail 5): 689 Hz-1.378 kHz.

D6 (detail 6): 345 Hz-689 Hz.

D7 (detail 7): 172 Hz-345 Hz.

D8 (detail 8): 86 Hz-172Hz.

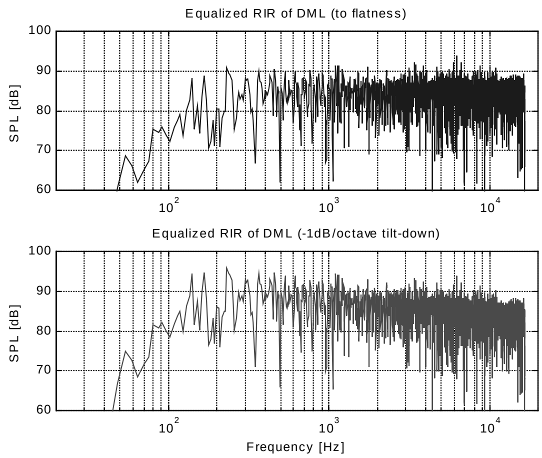

Figure 3.

A8 (approximation 8): DC-86Hz.

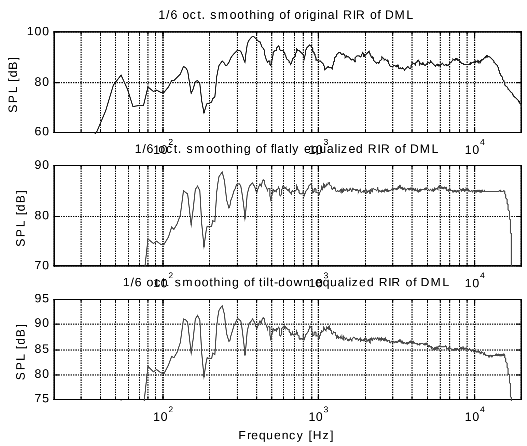

Figure 4.

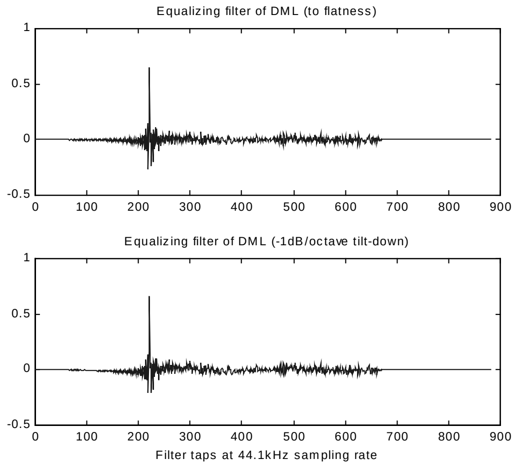

Figure 5.

We can discard out-of-band D1D1 and A8 (multiply them with gain zero respectively) and thus end up with 8 octave band subfilters with independent gain control in each band.

In Fig. 2, the upper subplot shows seven decomposed octave bands, and the lower subfigure is the synthesized inverse filter with good attenuation below 100Hz. Also note the good attenuation for each band. Even so, if we have an inverse filter with large magnitude differences in between two adjacent bands, the "tuning error" will increase accordingly.

We choose two target responses as examples:

Target 1: a flat response (normal equalization).

Target 2: a tilt-down (-1dB/octave) response.

In most cases a flat response tends to be perceived as too bright, so a tunable inverse filter can offer more flexibility in shaping the overall response. In Fig. 3, the upper subplot is the equalized response for target 1 (flat response), and the lower subfigure is the equalized response for target 2.

In Fig. 4, the upper subplot is the original measured room impulse response without any equalization (1/12 octave smoothed). The second subplot shows the RIR equalized to the flat target. In Fig. 5, the upper subfigure is the flat equalizing filter and the lower subfigure is the tilt-down (-1dB/octave) equalizing filter. In summary, we have developed a method that combines inverse filtering and parametric-tuning to achieve better control when dealing with loudspeaker/room equalization.

4. CONCLUSIONS

We present a complete solution on loudspeaker/room equalization with inverse filtering and tunable octave gain control combined together by using LPC & DWT techniques. (Future expansion to 1/3 octave tuning is possible.)

Frequency equalization based on LPC modeling is good at higher bands than lower bands, and it's generally good enough for peak rather than dip equalization (preferred by us).

Time-domain smearing/spread of a lengthy filter causes echo/comb filter effects –dynamic range is reduced too by observing ETC and CDS. From informal listening tests, fortunately, our filters are acceptable showing little aforementioned effects in time domain.

Even so, based on 3, this equalization system is not recommended to be cascaded with other digital filtering systems. That may let time-domain smearing audible and that's why we combine inverse filtering and octave gain control together.

For better low frequency equalization, frequency warping with individual subband gain controls is a better way and is our next move together with optimal Wavelet decomposition according to pychoacoustic model.

Real-time implementation is in progress, too.

5. REFERENCES

[1] Neil Harris, Malcolm Omar Hawksford; "*The Distributed-mode loudspeaker (DML) as a broad-band acoustic radiator*"; 103rd AES Convention, New York 1997; pre-print #4526.
[2] Graham Bank, Neil Harris; "*The Distributed Mode Loudspeaker - Theory And Practice*"; AES UK Conference "The Ins and Outs of Audio"; London, March 1998.
[3] Zwicker E. and Fastl H., *Psycho-acoustics Facts and Models*, Springer, Berlin, 1999.
[4] Neely S.J. and Allen J.B., "Invertibility of a Room Impulse Response", J. Acoust. Soc. Am., vol. 66, No. 1, pp. 165-169, 1979 July.
[5] Mourjopoulos J.N., "Digital Equalization of Room Acoustics", J. Audio Eng. Soc., Vol. 42, No. 11, pp. 884-900, 1994 November.
[6] Oppenheim A.V. and Johnson D.H., "Discrete Representation of Signals", Proc. IEEE 60, pp. 681-689, 1972.
[7] Karjalainen M., Piirila E., Jarvinen A., and Huopaniemi J., "Comparison of Loudspeaker Equalization Methods Based on DSP Techniques", JAES, Vol. 47, No. ½, pp. 14-30, 1999 Jan./Feb.
[8] Harris N., Flanagan S., and Hawksford O.J., "Stereophonic Localization in Rooms, Comparing Conventional and Distributed-

Mode Loudspeakers", AES 105th Convention, Preprint No. 4794, 1998.
[9] Toole, F. E., "Loudspeaker measurements and their relationship to listener preferences," Journal of the Audio Engineering Society, 34, 227-235, (1986).
[10] Bech, S., "Perception of timbre of reproduced sound in small rooms: influence of room and loudspeaker position," Journal of the Audio Engineering Society, 42, 999-1007, (1994).
[11] Holman, T., "Monitoring Sound in the One-Person Environment," submitted to the SMPTE Journal, (1997).
[12] Holman, T., "Report on Mixing Studios Sound Quality," Journal of the Japan Audio Society, (1994).
[13] Toole, F. E., "Subjective measurements of loudspeaker sound quality and listener performance," Journal of the Audio Engineering Society, 33, 2-32, (1985).
[14]Vetterli M. and Kovacevic J., Wavelets and Subband Coding, Prentice Hall, 1995.
[15] Oppenheim A.V. and Schafer R.W., Discrete-Time Signal Processing, Prentice-Hall, 1989.