CASCADED FIR FILTERS FOR MULTIPLE LISTENER LOW FREQUENCY ROOM ACOUSTIC EQUALIZATION

Sunil Bharitkar, Chris Kyriakakis

Audyssey Labs., Inc., and University of Southern California
350 S. Figueroa St., Ste. 196,
Los Angeles, CA 90071.

ABSTRACT

Low frequency room acoustic equalization is an extremely challenging problem to solve with realizable digital equalization filters. In this paper we propose a cascaded FIR solution, using a pattern recognition technique, for multiposition room equalization at low frequencies. The cascaded FIR solution allows for a real-time implementation of realizable filters whereas the pattern recognition approach captures the variabilities in room acoustics between different listening positions for multiposition equalization. We compare the results with the popular warping with low order spectral modeling approach, as well as with the multirate approach. The proposed method outperforms both the warping based and the multirate methods as can be evidenced through published data.

I. INTRODUCTION

Room equalization has traditionally been approached as a classic inverse filter problem. Although this may work well in simulations or highly-controlled experimental conditions, once the complexities of real-world listening environments are factored in, the problem becomes significantly more difficult. This is particularly true for small rooms in which standing waves at low frequencies cause significant variations in the frequency response at the listening position. A typical room is an acoustic enclosure that can be modelled as a linear system whose behavior at a particular listening position is characterized by an impulse response, $h(n); n \in \{0, 1, 2, ...\}$. This is generally called the room impulse response and has an associated frequency response, $H(e^{j\omega})$, which is a function of frequency. Generally, $H(e^{j\omega})$ is also referred to as the room transfer function (RTF). The impulse response yields a complete description of the changes a sound signal undergoes when it travels from a source to a receiver (microphone/listener). The signal at the receiver consists of direct path components, discrete reflections that arrive a few milliseconds after the direct sound, as well as a reverberant field component. In the frequency domain, the RTF shows significant spectral peaks and dips in the human range of hearing (i.e., 20 Hz-20 kHz) causing audible distortions at the listener position. In addition, it is well established that room responses change with source and receiver locations in a room [1], [2]. In other words, a room response can be uniquely defined by a set of spatial co-ordinates $l_i \triangleq (x_i, y_i, z_i)$. This assumes that the source is at origin and the receiver i is at the spatial co-ordinates, x_i, y_i and z_i, relative to a source in the room.

Clearly, the magnitude response variations in the RTF at a location need to be compensated (equalized). In addition, the variations in the RTF's, at different locations, in a room also need to be compensated. Accordingly, in our previous papers [3], [4], [5], we proposed several approaches for designing a multiple location/listener (viz., multiposition) equalization filter. The approaches ranged from using psychoacoustic warping where the equalization filter was designed on a warped frequency axis (i.e., the perceptual Bark scale) of the room response function with a lower order model (viz., LPC), to a multirate approach designed for low frequency equalization. In this paper we present an alternative approach for multiposition low frequency equalization and demonstrate its improvement over the warping/low-order modeling approach, as well as the multirate approach.

In the next section, we briefly describe the concept of multichannel and multiple listener room response equalization. We will also review the concept of fuzzy c-means clustering and its use for modeling the variabilities of room acoustics between multiple locations (viz., multiple listener positions). We also present some basic information of prior approaches for equalization (viz., the warping/low-order modeling and the multirate). In Section 3 we present the proposed cascaded FIR and clustering approach for low frequency multiple listener room response equalization. Results, including comparisons to the warping/low-order modeling and the multirate approach, are presented in Section 4. Section 5 concludes the paper and proposes future directions.

II. ROOM ACOUSTICS AND MULTIPLE LISTENER EQUALIZATION

II-A. Room Acoustics

Thus, audio signals delivered in each channel to an appropriate loudspeaker are radiated into the environment via the loudspeaker. However, the signal delivered to a listener's ear is not a faithful reproduction of the signal delivered to each of the loudspeaker. In fact, the audio signal is first affected by the acoustic properties of the loudspeaker (i.e., the loudspeaker frequency response) and the room acoustical properties (viz., the absorption characteristics of walls and floors, multi-path reflections from walls, the position of the loudspeakers and the listeners, room standing wave modes governed by the dimension of the rooms and the wavelength of sound). Thus, the non-ideal loudspeaker response (viz., non-flat magnitude response) and the effects of the room introduce distortion to the the audio signal. The problem is further compounded since different listeners are seated at different positions, and differences in the arrival of sound at different listeners will introduce different distortions at these listeners.

II-B. Multiple Listener Equalization

The position of the multiple-listener equalization filters in the playback chain is shown in Fig. 1. The goal of the equalization filter for each channel is such that the response at all listening positions is near flat in the frequency domain in a region of interest (e.g., between 20 Hz and the crossover frequency for the subwoofer channel).

II-C. Pattern Recognition for Multiposition Response Modeling

In previous papers [3], [4], [6] we have a proposed a pattern recognition technique, the fuzzy c-means clustering algorithm, for capturing the variabilities of loudspeaker-room acoustics between multiple positions. Broadly speaking, clustering procedures yield a data description in terms of clusters having centroids or *prototypes*. The clusters are formed from data points (room responses in the present case) having strong *similarities*. Clustering procedures use a criterion function, such as a sum of squared distances from the prototypes, and seek a grouping (cluster formation) that extremizes a criterion function.

A cluster room response *prototype* is a generalized representation of the room responses that are grouped in the cluster, and the prototype plays a fundamental role in the proposed multiple-listener equalization technique.

In fuzzy clustering, a room response $\underline{h}_j$ may belong to more than one cluster by different "degrees". This is accomplished by a continuous membership function- $\mu_i(\underline{h}_j) \in [0,1]$. The motivation for using the fuzzy c-means clustering approach can be best understood from Fig. 2, where the direct path component of the response associated with listener 3 is similar (in the Euclidean sense) to the direct path component of the response associated with listener 1 (since listener 1 and 3 are at same radial distance from the loudspeaker). Furthermore, it is likely that the reflective component of listener 3 response may be similar to the reflective component of listener 2 (due to their proximity). Thus, it is clear that if responses 1 and 2 are clustered separately, then response 3 should belong to both clusters to some degree. Thus, this clustering approach permits an intuitively reasonable model for prototype formation.

It can be shown that the centroids (prototypes) and membership functions are given by

$$
\begin{aligned}
\underline{\hat{h}}_i^* &= \frac{\sum_{k=1}^{N}(\mu_i(\underline{h}_k))^2\underline{h}_k}{\sum_{k=1}^{N}(\mu_i(\underline{h}_k))^2} \\
\mu_i(\underline{h}_k) &= [\sum_{j=1}^{c}(\frac{d_{ik}^2}{d_{jk}^2})]^{-1} = \frac{\frac{1}{d_{ik}^2}}{\sum_{j=1}^{c}\frac{1}{d_{jk}^2}}; \\
d_{ik}^2 &= \|\underline{h}_k - \underline{\hat{h}}_i^*\|^2 \qquad (1) \\
i = 1,2,...,c; &\qquad k = 1,2,...,N
\end{aligned}
$$

where $\underline{\hat{h}}_i^*$ denotes the i-th cluster room response prototype, and the room responses are of length P. In this paper the responses were measured in a reverberant room with a reverberation time $T_{60} = 0.125$ seconds and hence, at a 48 kHz sampling rate, the room responses were of length 8192 samples so as to capture the reverberant portion of the response.

Once the prototypes are formed, it is required that they be combined to form a single final prototype which can be inverted to form the equalization filter. One approach to do this is by using the following non-uniform weighting model:

$$\underline{h}_{final} = \frac{\sum_{j=1}^{c}(\sum_{k=1}^{N}(\mu_j(\underline{h}_k))^2)\underline{\hat{h}}_j^*}{\sum_{j=1}^{c}(\sum_{k=1}^{N}(\mu_j(h_k))^2)} \qquad (2)$$

II-D. Warping/LPC-based Lower-order Equalization Filters

The prototype is of length 8192 samples and hence needs to be modeled by a lower-order response at which point this lower-order response can be inverted and used as an equalization filter for real-time applications.

We have presented a warping/fuzzy-clustering/LPC method for designing filters with low order [3] based on psychoacoustic warping. The algorithm that was used for designing the lower order equalization filters is shown in Fig. 3. The method works very well for frequencies above 100 Hz with 512 taps of an FIR filter. However, subwoofer-room responses at low frequencies (viz., below 100 Hz) cannot be equalized satisfactorily with this method. For example, Fig. 4 shows five subwoofer and room responses, for speaker A, measured in a reverberant room (dashed lines), whereas the resulting equalized responses after applying a 512-tap FIR filter, designed using the approach of Fig. 3, are shown as solid lines. Clearly, a limited performance can be expected with this approach as the plots are not sufficiently flat over the low-frequency region of interest. Hence a multirate approach was proposed in [5] for equalizing the low-frequency subwoofer-room responses.

II-E. Multirate-based Low-frequency Equalization Filters

The clustering algorithm was used for low frequency equalization by utilizing multirate signal processing techniques. Specifically, each of the low frequency room responses $\underline{h}_k (k = 1, 2, ..., 6)$, of length 8192 samples at 48 kHz sampling rate, is filtered by a decimation filter having a sampling rate of about 400 Hz for targeting the low frequency region (utpo about 200 Hz). Each of the room responses had a length of $8192 \times 400/48000 \approx 64$ taps. The fuzzy c-means clustering algorithm was then applied after decimation to give a prototype. Finally, the minimum phase part of the 64 coefficient prototype was inverted to give the equalization filter. The resulting equalized responses, for speaker A, are shown in Fig. 5 and are relatively flatter in the low-frequency region.

III. CASCADED FIR FILTERS FOR LOW FREQUENCY EQUALIZATION

A sufficiently high order LPC technique provides an effective low frequency response model. An example of a $p = 1024$ order LPC model of a full-range subwoofer response (i.e., a response measured without any bass-management) is shown in Fig. 6. Because, the number of coefficients that form the equalization filter is large (viz., 1025 for a 1024 order LPC), it is possible to "compress" the number of coefficients to target modelling of only the low frequency region of the subwoofer-room response. This can be achieved by first determining the 1024 poles of the LPC model of the cluster centroid $\underline{h}_{final}$, then determining *only* those Q poles that contribute to the low-frequency model, forming the complex conjugate pair of each of the low-frequency poles to form a second-order denominator polynomial for each low-frequency pole, forming the second-order numerator polynomial from the denominator polynomial corresponding to each of the stable poles, and forming the Q-cascade of these second-order numerator polynomials or FIR coefficients. The factorization of such high-order polynomials, into roots, involves optimization techniques that are computationally intensive. Fox *et al.* [9] present the Horner's method of evaluating and deflating polynomials which is a much faster technique than the eigenvalue method of a companion matrix (for e.g., the *roots* function in Matlab). We have used this method for finding the roots of the polynomial associated with the 1024-order LPC. The flow chart for the algorithm is shown in Fig. 7, whereas the block diagram for the cascade FIR equalization filters, $H_i(z)$ is shown in Fig. 8, where $H_i(z) = z^2 - 2\Re\{z_i\}z + |z_i|^2$. Fig. 9 shows one set of results using this method with the subwoofer-room response for speaker A, and Fig. 10 shows another set of results for speaker B with $Q = 15$. Figs. 11 and 12 show results obtained for the subwoofer-room response of speaker B using the warping/LPC and multirate method, respectively. The proposed cascade FIR method gave the best results as the equalization curves are flatter over the multirate and warping/LPC methods.

IV. CONCLUSIONS

In this paper, we presented a cascaded FIR method for low frequency subwoofer-room response multiple listener equalization. The proposed method outperforms the multirate and warping/LPC based multiple listener equalization methods at low frequencies. Future work will involve a combination of warping, pole-modeling, and only unwarping the low-frequency poles. This will allow faster realizations, due to lesser number of cascaded filters, as it is a well-established fact that warping significantly reduces the number of LPC coefficients required to represent the spectrum.

V. REFERENCES

[1] H. Kuttruff, *Room Acoustics,* Elseiver Applied Science, 3rd ed., New York. 1991.

[2] J. Mourjopoulos, "On the variation and invertibility of room impulse response functions," *Journal of Sound and Vibration,* vol. 102(2), pp. 217–228, 1985.

[3] S. Bharitkar and C. Kyriakakis, "A Cluster Centroid Method for Room Response Equalization at Multiple Locations," *Proc. 2001 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA '01),* New York. Oct. 2001.

[4] S. Bharitkar and C. Kyriakakis, "Perceptual Multiple Location Equalization with Clustering," *Proc. 36th IEEE Asilomar Conf. on Signals, Systems, & Computers.,* Pacific Grove. CA. Nov. 2002.

[5] S. Bharitkar and C. Kyriakakis, "Multirate signal processing for multiple listener low frequency room acoustic equalization," *Proc. 38th IEEE Asilomar Conf. on Signals, Systems, & Computers.,* Pacific Grove. CA. Nov. 2004.

[6] S. Bharitkar and C. Kyriakakis, "Visualization of multiple listener room acoustic equalization with the Sammon map," *to appear i IEEE Trans. Speech & Audio Proc., 2006.*

[7] J. Bezdek, *Pattern Recognition with Fuzzy Objectiver Function Algorithms,* Plenum Press, New York. 1981.

[8] P. P. Vaidyanathan, *Multirate Systems and Filter Banks,* Prentice Hall, New Jersey. 1993.

[9] G. A. Sutton, C. S. Burrus, J. W. Fox, and S. Treitel, "Factoring very-high-degree polynomials," *IEEE Sig. Proc. Mag.,* pp. 27-42, Nov. 2003.

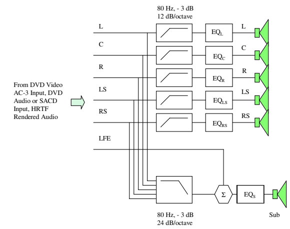

Fig. 1: Equalization filters EQ_i for channel i in a 5.1 setup.

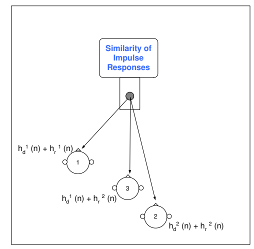

Fig. 2: Motivation for using fuzzy c-means clustering.

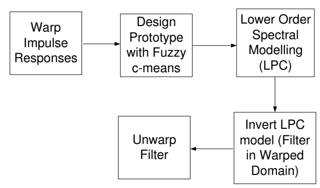

Fig. 3: Lower order filter design with warping and LPC.

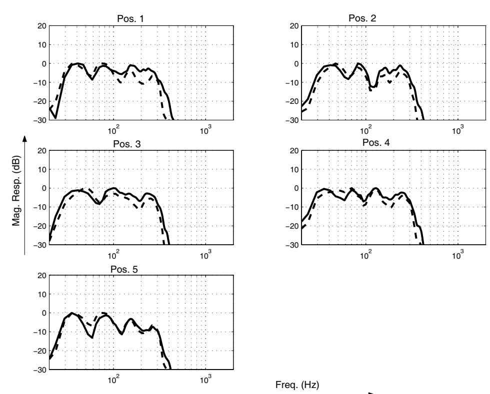

Fig. 4: Warping/LPC based subwoofer A equalization.

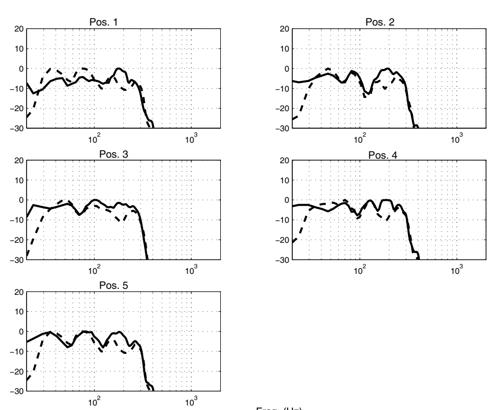

Fig. 5: Multirate subwoofer A equalization.

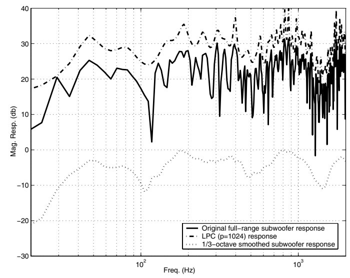

Fig. 6: LPC model of order 1024.

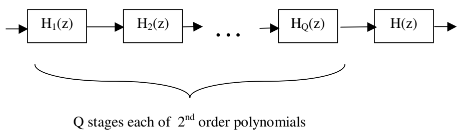

Fig. 8: System-level depiction for cascade FIR structure.

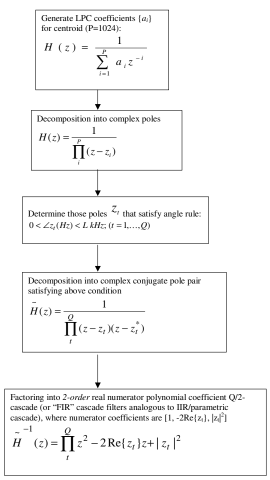

Fig. 7: Cascade FIR filtering flow-chart.

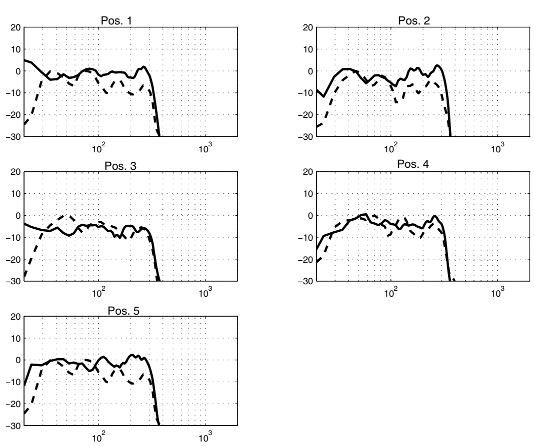

Fig. 9: Cascade filtering results for subwoofer A.

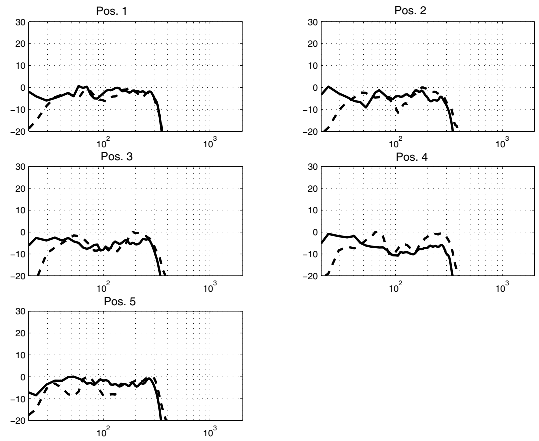

Fig. 10: Cascade filtering results for subwoofer B.

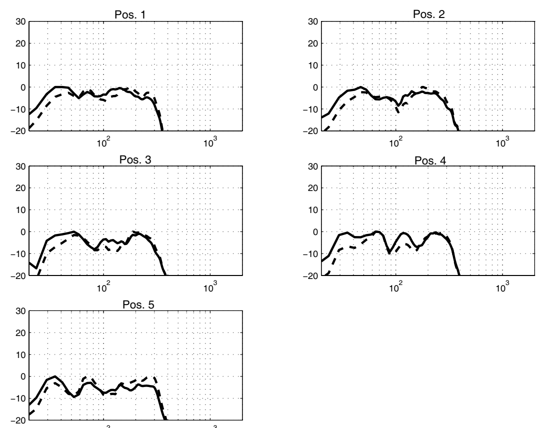

Fig. 11: LPC/warping results for subwoofer B.

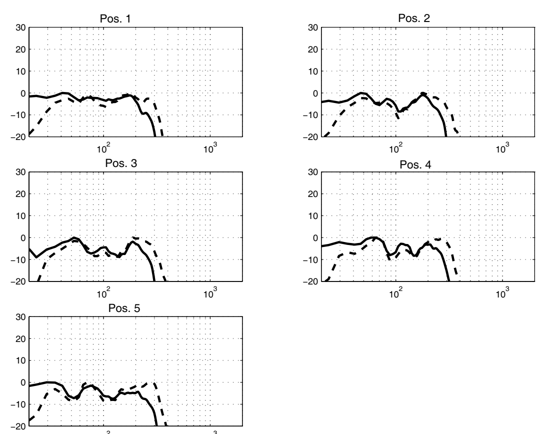

Fig. 12: Multirate results for subwoofer B.