MULTIPLE POSITION ROOM RESPONSE EQUALIZATION WITH FREQUENCY DOMAIN FUZZY C-MEANS PROTOTYPE DESIGN

Ivan Omiciuolo[1], *Alberto Carini*[2] *and Giovanni L. Sicuranza*[3]

[1,3] DEEI - University of Trieste - Via A. Valerio, 10 - 34127 Trieste, Italy
[2] STI - University of Urbino - Piazza della Repubblica, 13 - 61029 Urbino, Italy

ABSTRACT

In this paper we elaborate and improve the multiple position room response equalization technique based on fuzzy c-means clustering of Bharitkar and Kyriakakis. Differently from their approach, we perform most of operations in the frequency domain and we apply the fuzzy c-means clustering to the amplitude room frequency responses at different positions. We show in the paper that working in the frequency domain can avoid many onerous operations performed in the approach of Bharitkar and Kyriakakis and thus a computationally efficient algorithm for multiple position room response equalization can be obtained. In all our experiments, the equalization performances of the proposed technique have always been similar to those of the technique of Bharitkar and Kyriakakis.

Index Terms— Room response equalization, amplitude equalization, fuzzy c-means, clustering, pattern recognition.

1. INTRODUCTION

Room response equalization (or room equalization) has been applied in theory and practice for improving the objective and subjective quality of sound reproduction systems in cinema theaters, home theaters, and car HiFi systems [1]. Room response equalization systems act by shaping the room transfer function (RTF) from sound reproduction system to listener with a suitably designed equalizer. Both minimum-phase and mixed-phase room equalizers have been proposed in the literature [2]. Minimum-phase room equalizers can be used in order to shape the RTF magnitude response and they can act on the minimum-phase part of the RTF phase response. Mixed-phase room equalizers can correct also the non-minimum-phase part of the RTF phase response. In principle, they can remove some of the room reverberation [3], but particular care must be taken to avoid "pre-echoes" due to errors in the acausal part of the equalizer. Modal equalizers have also been proposed in order to control the behavior of acoustic resonances at low frequencies (below 200 Hz) so that the decay time of the associated modes can be reduced [4].

Room equalizers can be divided in single position and multiple position equalizers. In single position room equalizers, the equalization filter is designed on the basis of a measurement of the room impulse response in a single location. These equalizers can achieve the room equalization only in a reduced zone around the measurement point (of the size of a fraction of the acoustic wavelength). Indeed, the room impulse response varies significantly with the position in the room. Moreover, it was shown in [3] that the room impulse response varies also with time and that the room should be considered a "weakly nonstationary" system.

Multiple position room equalizers are capable to enlarge the equalized zone by measuring the room impulse response in multiple locations. Different techniques for multiple position room equalization have been proposed in the literature [1, 5, 6, 7, 8]. An exact equalization technique for multiple positions based on MINT (multiple-input/multiple-output inverse theorem) was proposed in [5]. With this technique, exact equalization can be obtained provided that the number of equalization points is lower than the number of sound sources (loudspeakers) and provided that the room responses have uncommon zeros among them. A multiple-channel adaptive equalization system was proposed in [6]. This system adaptively minimizes the sum of the squared errors between the equalized responses and a delayed version of the played signal. In [7] a multiple-point equalization filter using the common acoustical poles of RTFs was proposed. Wave domain adaptive filtering for the equalization of massive multichannel sound reproduction systems was discussed in [9]. A multiple position room response equalization technique based on fuzzy c-means clustering and frequency warping was introduced in [1, 8]. Specifically, given a set of room impulse responses measured at different positions, the technique in [1] applies a fuzzy c-means algorithm for clustering these room responses on the basis of their similarity. A prototype impulse response, obtained by combining the cluster centroids, is then used for designing the low order equalization filter by means of Linear Predictive Coding (LPC) analysis. In order to obtain a better fit of the LPC model to the room response in the low frequency region, the measured room responses are frequency warped [10] using a psychoacustically motivated Bark scale [11]. In [12] the method of [1] was further developed by introducing a weighted fuzzy c-means clustering algorithm that allows to introduce different weights on the room impulse response samples. Although the technique in [1] is able to obtain only a magnitude equalization of the room response, it is effective and robust against displacement effects.

In this paper, we elaborate and improve the technique of [1] by performing most of its operations in the frequency domain. Differently from [1], the fuzzy c-means clustering is applied to the room amplitude frequency responses at different positions. We show in the paper that working in the frequency domain can avoid many onerous operations performed in the approach of [1] and thus a computationally efficient algorithm for room equalization can be obtained. In all our experiments, the equalization performances of the proposed technique have always been equivalent to those of [1].

The paper is organized as follows. Section 2 provides a brief overview of the multiple position room response equalization of [1]. Section 3 describes the proposed approach. Section 4 presents some experimental results that compare the performances of the proposed room response equalization with those of [1]. Conclusive remarks are given in Section 5.

Throughout the paper, small boldface letters are used to denote vectors and italic letters are used for scalar quantities.

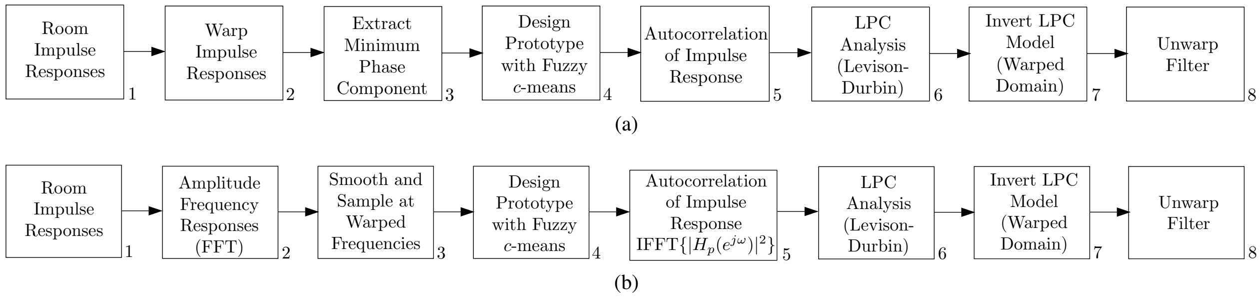


Fig. 1. Multiple position room response equalization: (a) method of [1] and (b) proposed method.

2. ROOM RESPONSE EQUALIZATION USING PATTERN RECOGNITION

Figure 1.(a) describes the operations performed in the room equalization technique of [1]. The operations done by this room equalizer are the following:

1. M room impulse responses of memory length N are measured at different positions in the zone we want to equalize.

2. The room responses are frequency warped with an approximate Bark scale. Frequency warping was introduced in [10] and it involves the replacement of every delay element z^{-1} with an all pass filter of order 1,

$$D(z) = \frac{z^{-1} + \lambda}{1 + \lambda z^{-1}}. \tag{1}$$

Frequency warping is used in order to improve low order modelling at low frequencies. Positive values of the warping parameter λ takes to an higher resolution in the lowest frequencies and to a lower resolution to the highest frequencies. By properly choosing the λ parameter, the frequency resolution can approximate the Bark scale [11]. It should be noted that the frequency warping operation performed in the time domain is a computational heavy operation that requires an order N^2 of multiplications and additions.

3. The minimum-phase part of the warped room impulse responses is extracted. In [13], the authors suggested to compute the minimum-phase sequences from cepstrum. The operation requires the computation of two FFTs and two IFFTs on N samples, N logarithms and N exponentiations for each of the M impulse responses. The minimum-phase extraction is important for the clustering operation, in order to avoid undesirable time and frequency domain effects, due to incoherent linear combination [1].

4. Fuzzy c-means clustering in applied to the minimum-phase impulse responses. Clustering is used to extract the common patterns of the impulse responses by means of c centroids. Let us indicate with $\mathbf{h}_k$, for $k = 1, \ldots, M$, the M warped minimum-phase room impulse responses, with $\mathbf{h}_i^*$, for $i = 1, \ldots, c$, the c centroids, and with $\mu_i(\mathbf{h}_k)$ the cluster membership functions. Fuzzy c-means clustering is performed by iteratively applying the following equations [1]:

$$\mathbf{h}_i^* = \frac{\sum_{k=1}^{M} \mu_i^2(\mathbf{h}_k)\mathbf{h}_k}{\sum_{k=1}^{M} \mu_i^2(\mathbf{h}_k)}, \tag{2}$$

$$\mu_i(\mathbf{h}_k) = \left[\sum_{j=1}^{c} \frac{d_{ik}^2}{d_{jk}^2}\right]^{-1} = \frac{\frac{1}{d_{ik}^2}}{\sum_{j=1}^{c} \frac{1}{d_{jk}^2}} \tag{3}$$

with $d_{ik}^2 = \|\mathbf{h}_k - \mathbf{h}_i^*\|^2$, $i = 1, \ldots, c$, and $k = 1, \ldots, M$. Eq. (2) and (3) are iterated until the variation of the objective function that is minimized,

$$J = \sum_{k=1}^{M} \sum_{i=1}^{c} \mu_i^2(\mathbf{h}_k) \|\mathbf{h}_k - \mathbf{h}_i^*\|^2, \tag{4}$$

is lower than a small constant ϵ. The cluster membership functions $\mu_i(\mathbf{h}_k)$ are typically initialized with random numbers between 0 and 1 such that $\sum_{i=1}^{c} \mu_i(\mathbf{h}_k) = 1$ for all k, and c is an integer close to $\sqrt{M}$. A final prototype filter of length N, $\mathbf{h}_p$, is then constructed with a weighted mean of the cluster centroids as follows:

$$\mathbf{h}_p = \frac{\sum_{j=1}^{c} \left(\sum_{k=1}^{M} \mu_j^2(\mathbf{h}_k)\right) \mathbf{h}_j^*}{\sum_{j=1}^{c} \left(\sum_{k=1}^{M} \mu_j^2(\mathbf{h}_k)\right)}, \tag{5}$$

The prototype filter represents the room response component at the different positions we want to equalize.

5. The autocorrelation function (ACF) of the prototype impulse response is computed. The ACF of the prototype impulse response is needed by the Levinson-Durbin algorithm used in LPC analysis. The ACF can be computed in time domain using time averages, or it can be obtained from the inverse of the squared amplitude frequency response. This second approach is often the most computationally efficient, and it requires the computation of an FFT on N samples for estimating the frequency response of the prototype filter and of an IFFT on N samples for estimating the ACF.

6. The Levinson-Durbin algorithm is applied to derive a low order all-pole LPC model of the prototype filter, with order P. The use of a low order equalizer is beneficial not only for computational complexity reasons, but also for improving the equalizer robustness toward displacement effects and slow time variations of the room response [3].

7. The all-pole LPC model is inverted in order to obtain an FIR equalizer in warped domain with memory length P.

8. The frequency response of the equalizer is unwarped. This operation can be performed by replacing the FIR warped equalizer, with the IIR filter of Fig. 2, where $p(i)$ are the coefficients of the warped prototype filter, and λ is the same parameter used for warping the impulse responses.

3. THE PROPOSED TECHNIQUE

Figure 1.(b) describes the operations performed with the proposed technique. In contrast to the approach of [1], we perform the fuzzy c-means clustering and most of the operations in the frequency domain. In this way, we can avoid most of the computational burden of the technique in [1]. The operations performed by the proposed equalizer are the following:

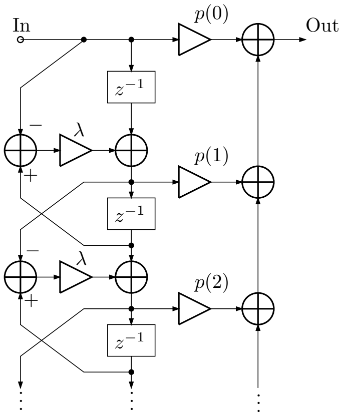


Fig. 2. IIR unwarped filter

1. M room impulse responses of memory length N are measured at different positions in the zone we want to equalize. Alternatively, we can directly estimate the amplitude frequency responses of the room at the different positions.
2. The room amplitude frequency responses $|H_i(e^{j\omega_k})|$ are estimated by means of M FFTs on N samples. By discarding the phase information, we avoid the extraction of the minimum phase component of the impulse responses, which was needed for implementing the fuzzy c-means algorithm in time domain.
3. Warping is performed in the frequency domain by sampling the room amplitude frequency responses. Given W equally spaced frequencies in the warped frequency domain (with $W \ll N$), we determine the corresponding frequencies in the linear frequency domain according to the rule in (1). By sampling the room amplitude responses at these frequencies, the warped amplitude responses $|\hat{H}_i(e^{j\omega_k})|$ are obtained. A $1/3$ octave smoothing of the amplitude response can optionally be applied before the frequency warping. The $1/3$ octave smoothing allows to remove notches in the frequency response that could disturb the LPC modelling operation. Moreover, the spectrum smoothing makes the room equalizer more robust to displacement errors and time variations of room response [3].
4. The fuzzy c-means clustering is used again to extract the common patterns of the room amplitude responses. The algorithm can by applied by iterating equation (2) and (3), where now $\mathbf{h}_k = [|\hat{H}_k(e^{j\omega_1})|, |\hat{H}_k(e^{j\omega_2})|, \ldots, |\hat{H}_k(e^{j\omega_W})|]^T$, for $k = 1, \ldots, M$ and the computed cluster centroids $\mathbf{h}_i^*$ represents the common patterns of the room amplitude responses. A prototype amplitude response, $\mathbf{h}_p$, can be obtained from the weighted mean of the cluster centroids using (5). The prototype amplitude response represents the room response component we want to equalize.
5. The autocorrelation function of the prototype impulse response is estimated by inverting the squared amplitude spectrum, without the need of estimating the prototype impulse response itself.

The other steps, 6., 7., and 8., are identical to those of Section 2.

With the proposed technique we can avoid many heavy operations performed in the approach of [1]. Specifically, the frequency warping, which in the time domain is a very computationally intensive operation, with the proposed approach is done by sampling the amplitude frequency responses. In our experiments, we implemented this operation at zero cost by simply selecting the closest samples $|H_i(e^{j\omega_k})|$ to the desired warp frequencies. The sampling operation allows also to obtain a more compact representation of the amplitude frequency responses, with a reduced number of samples W. Moreover, with the proposed approach the extraction of the minimum phase component of the impulse responses is avoided and the computation the autocorrelation of the impulse response of prototype filter is done simply with an IFFT on W squared-amplitude frequency response samples. The computational complexity of the fuzzy c-means clustering algorithm strongly depends on the number of iterations necessary for the algorithm convergence. In our experiments, the computational complexities of the fuzzy c-means algorithm performed in the frequency domain and in the time domain were similar.

Table 1. Computational complexity comparison of the method of [1] and of proposed method (I number of iteration of fuzzy c-means algorithm).

	Method of [1]		Proposed Method	
Step	Mult	div	Mult	div
2.	$4MN^2$		$MN(\log_2 N + 2)$	
3.	$4MN\log_2 N$		0	
4.	$2MNcI$	$2McI$	$2MWcI$	$2McI$
5.	$2N(\log_2 N + 1)$		$W(\log_2 W + 1)$	
6.	P^2	P	P^2	P

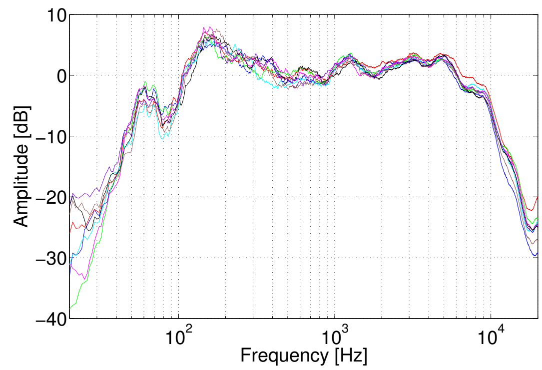


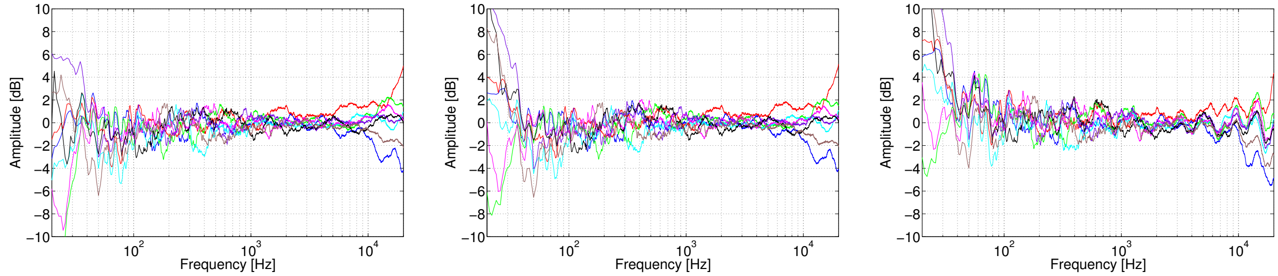
Fig. 3. Room amplitude frequency responses.

Table 1 provides the computational complexity of the method of [1] and of the proposed method in terms of multiplications and divisions. In the proposed method we have not taken into account the $1/3$-rd octave spectral smoothing, which cost around $0.04NW$ multiplications. In the method of [1] we have also to compute MN logarithms and MN complex exponentiations, and in the proposed method we have also MN square-roots.

4. EXPERIMENTAL RESULTS

In this section, we provide some experimental results that compare the performance of the proposed approach with those of [1].

Eight room impulse responses were recorded in a room of 6 m × 9 m × 3 m with an omnidirectional microphone using an injected pink noise with a sampling frequency of 96 kHz and length $N = 65,536$. The room responses were recorded in four separate positions equispaced by 90 cm on a line at the room center and at the heights of 1 m and 1.30 m. The loudspeaker was positioned in one room corner at 1 m from the walls and at an height of 90 cm. Fig. 3 shows the room amplitude frequency responses in the audible band after a $1/3$-rd octave smoothing. In both approaches, we set the length of the LPC model order $P = 512$, the number of centroids $c = 3$ and the small constant $\epsilon = 10^{-5}$, and the warping parameter $\lambda = 0.82108$. In our approach, we set the warped amplitude frequency response length $W = 4096$. The mean value of iterations



(a) (b) (c)

Fig. 4. Room amplitude frequency responses after equalization

of the fuzzy c-means algorithm over 100 experiments with the same room responses was 22 iterations with the method of [1], 32 iterations with the proposed method without $1/3$-rd octave smoothing of the amplitude frequency responses before warping, and 28 iterations with the the proposed method applying also a $1/3$-rd octave smoothing of the amplitude frequency responses before warping. The computational cost of the equalizer design was of $1.38E^{+11}$ multiplications for the method of [1], $1,71E^{+7}$ multiplications for the proposed method without $1/3$-rd octave smoothing of the amplitude frequency responses before warping, and $2.70E^{+7}$ for the proposed method with $1/3$-rd octave smoothing. The cost in terms of divisions, square roots, logarithms and exponentiations was always negligible compared with the multiplication cost. Fig. 4 shows the amplitude frequency responses after equalization (a) with the method of [1], (b) with the proposed method without $1/3$-rd octave smoothing of the amplitude frequency responses before warping, and (c) with the the proposed method applying also a $1/3$-rd octave smoothing of the amplitude frequency responses before warping. Compared with Fig. 3, we have reduced the vertical scale in order to enhance details visibility. In these plots, the mean value of the spectral deviation measure [8, 12] in the band 100 Hz - 10 kHz was 2.8249 for the equalization method of [1], was 2.8240 for the proposed method without $1/3$-rd octave smoothing, and was 2.9204 for the proposed method with $1/3$-rd octave smoothing, while it was 5.1267 for the eight room responses of Fig. 3. By comparing Fig. 4.(a) and (b) , we see that the proposed equalizer provides performances similar to those of [1]. Thus, the experimental results shows that the clustering algorithm applied in the frequency domain is capable to extract the common pattern of the room responses as well as the algorithm applied in the time domain. From Fig.4.(c), we see that also the proposed algorithm with $1/3$-rd octave smoothing provides results similar to those of [1], with the performances that degrades a little at high frequencies and improve at low frequencies.

5. CONCLUDING REMARKS

In this paper we have elaborated and improved the multiple position room response equalization technique based on fuzzy c-means clustering of [1]. Differently from the approach of [1], most of operations are performed in the frequency domain. In this way, many onerous operations performed in the approach of [1] can be avoided. In all our experiments, the equalization performances of the proposed technique have always been similar to those of the technique of [1]. Listening tests are currently being performed with the proposed equalization technique using both speech and music. The preliminary results of these tests have shown that the equalizer is able to compensate for the room resonances without introducing any audible artifact in the audio signals.

6. REFERENCES

[1] S. Bharitkar and C. Kyriakakis, *Immersive Audio Signal Processing*, Springer, New York, 2006.

[2] M. Karjalainen et alt., “About room response equalization and dereverberation,” in *Proc. 2005 IEEE Work. on Appl. of Signal Process. to Audio and Acoustics*, New Paltz, NY, Oct. 2005, pp. 183–186.

[3] P. D. Hatziantoniou and J. N. Mourjopoulos, “Errors in real-time room acoustics dereverberation,” *J. Audio Eng. Soc.*, vol. 52, no. 9, pp. 883–899, 2004.

[4] A. Mäkivirta et alt., “Modal equalization of loudspeaker-room responses at low frequencies,” *J. Audio Eng. Soc.*, vol. 51, no. 5, pp. 324–343, May 2003.

[5] M. Miyoshi and Y. Kaneda, “Inverse filtering of room acoustics,” *IEEE Trans. Signal Processing*, vol. 36, no. 2, pp. 145–152, Feb. 1988.

[6] S. J. Elliot and P. A. Nelson, “Multiple-point equalization in a room using adaptive digital filters,” *J. Audio Eng. Soc.*, vol. 37, no. 11, pp. 899–907, Nov. 1989.

[7] Y. Haneda, S. Makino, and Y. Kaneda, “Multiple-point equalization of room transfer functions by using common acoustical poles,” *IEEE Trans. Speech and Audio Process.*, vol. 5, no. 4, pp. 325 – 333, July 1997.

[8] S. Bharitkar and C. Kyriakakis, “A cluster centroid method for room response equalization at multiple locations,” in *Proc. 2001 IEEE Work. on Appl. of Signal Process. to Audio and Acoustics*, New Paltz, NY, Oct. 2001, pp. 55 – 58.

[9] S. Spors et alt., “Active listening room compensation for massive multichannel sound reproduction systems using wave-domain adaptive filtering,” *J. Acoust. Soc. Amer.*, vol. 122, no. 1, pp. 354–369, July 2007.

[10] A. Oppenheim, D. Johnson, and K. Steiglitz, “Computation of spectra with unequal resolution using the fast fourier transform,” *Proc. IEEE*, vol. 52, no. 2, pp. 299–301, Feb. 1971.

[11] J. O. Smith and J. S. Abel, “Bark and ERB bilinear transforms,” *IEEE Trans. Speech and Audio Process.*, vol. 7, no. 6, pp. 697–708, Nov. 1999.

[12] S. Turmchokkasam and S. Mitaim, “Effects of weights in weighted fuzzy c-means algorithm for room equalization at multiple locations,” in *Proc. IEEE Int. Conf. on Fuzzy Systems*, Vancouver, Canada, July 2006, pp. 1468–1475.

[13] S. Bharitkar and C. Kyriakakis, “Perceptual multiple location equalization with clustering,” in *Proc. 36th Asilomar Conf. on Signals, Systems, and Computers*, Pacific Grove, CA, Oct. 2002, vol. 1, pp. 179 – 183.