

AN ADAPTIVE MULTIPLE POSITION ROOM RESPONSE EQUALIZER IN WARPED DOMAIN

Stefania Cecchi †, Alberto Carini ‡, Andrea Primavera †, Francesco Piazza †

† DIBET - A3Lab, Università Politecnica delle Marche, Via Brecce Bianche 1, 60131 Ancona, Italy

‡ DiSBef, Università di Urbino, Piazza della Repubblica 13, 61029 Urbino, Italy

ABSTRACT

In a recent paper, the authors proposed an adaptive method for multiple position room response equalization. In that paper, the room responses at different positions in the zone to be equalized are estimated in the frequency domain and their common trend is extracted with efficient statistical operators. Then, the resulting prototype response is used to design the equalizer. In this paper, the approach is extended by estimating the equalizer in warped domain to improve the equalization performance at low frequencies. The effect of warping is analyzed and compared with objective and subjective measurements.

Index Terms— Room response equalization, Adaptive system, Multipoint equalization

1. INTRODUCTION

Cinema theaters, home theaters, and car Hi-Fi systems use room response equalizers in order to improve the quality of the sound reproduction. Room response equalizers are designed trying to compensate the room transfer function (RTF) from the sound reproduction system to the listener [1].

Single position and multiple position room equalizers are used in common practice. A single position room equalizer estimates the equalization filter on the basis of the measurement in a single location of the room impulse response (IR) [2] and it is effective only on a reduced zone around the measurement point (of the size of a fraction of the acoustic wavelength). In reality, the room IR varies significantly with the position in the room [3] and with time [4] as the room can be considered a “weakly non-stationary” system. To enlarge the equalized zone and to contrast the room response variations, multiple-position equalizers have been proposed [1]. A multiple position room equalizer uses multiple measurements of the room IR at different locations in order to design the equalizer. Moreover, complex spectral smoothing and short equalization filters are often used to contrast the audible distortions caused by equalization errors due to the IRs variations [4].

Different techniques have been proposed for multiple position room equalization [1, 5, 6, 7, 8, 9]. A least-square approach for inverting room responses was presented in

[5]. An exact equalization technique based on the multiple-input/multiple-output inverse theorem was proposed in [6]. An equalization filter using the common acoustical poles of RTFs was introduced in [8]. An equalization system based on a k -means with splitting clustering algorithm applied to all-pole RTF measures was presented in [9]. An equalization technique based on fuzzy c -means clustering and frequency warping was introduced in [1, 10] and was elaborated and improved in [11, 12].

While the approaches considered so far use fixed equalizers, the room is a time varying environment which changes with temperature, pressure, and with movement of people or other obstacles within the enclosure. Thus, it is important to develop also adaptive solutions, capable to track and to adapt to these room response variations.

A first adaptive equalizer was proposed in [7], where the sum of the squared errors between the equalized responses and a delayed version of the signal is adaptively minimized. Unfortunately, the approach of [7] is very sensitive to peaks and notches in the room response, to the room response variations at different positions, and suffers pre-echo problems. An adaptive frequency domain single-channel room equalizer was proposed in [13]. Working in the Fast Fourier Transform (FFT) domain, the loudspeaker signal and the measured microphone signals are split in subbands. Equalization is performed by updating the filter weights in these subbands. The approach of [13] is interesting for its simplicity, robustness towards peaks and notches of the room response, and for its ability to track the room response variations. In [14] we elaborated and improved the approach of [13] by developing a multiple position room response equalizer. In this paper, we further elaborate it by estimating the equalizer in warped domain to improve the equalization performance at low frequencies. Specifically, the room responses at different positions in the zone to be equalized are estimated in warped domain and the common trend of these responses is extracted using efficient statistical operators studied in [12, 15]. The common trend of the room responses is used to design an FIR equalizer in the frequency domain. The proposed approach is simple, computationally efficient and it is able to adapt to slow time variations in the room response.

The rest of the paper is organized as follows. Section 2

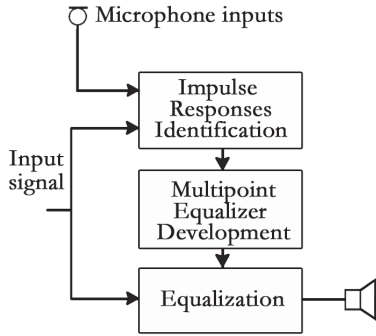


Fig. 1. Overall scheme of the proposed approach.

provides description of the proposed equalizer. Section 3 details some experimental results that illustrate the performance and the quality of the proposed approach. The effect of warping is analyzed and compared with objective and subjective measurements. Eventually, Section 4 provides some concluding remarks.

2. PROPOSED ROOM RESPONSE EQUALIZER

The proposed equalizer is based on an adaptive approach that iteratively estimates the IRs and at the same time generates the equalizer, as shown in Fig.1. R microphones positioned in the zone to be equalized are used to capture the signal reproduced by the loudspeaker system and altered by the room transfer function. The loudspeaker input signal and the microphones signals are used for identifying the corresponding R room responses. The latter are then employed for designing the equalizer used for compensating the room response. The next subsections illustrates the two steps of the algorithm.

2.1. Room response identification

Fig. 2 schematizes the approach used for the identification of the room response from an input signal – microphone signal pair with a frequency domain Least Mean Square optimization approach. The input signal reproduced by the loudspeaker system and the microphone signal, which captures the room dynamics, are FFT transformed on a block by block basis with a length K frame size and the magnitude spectrum of the signals is extracted. The FFT spectra are subsampled and smoothed at M frequencies chosen in order to frequency warp the room responses with an approximate Bark scale. Identification of the room responses is performed at the M frequency components with an adaptive filter having weights $H_n(m)$, with $m = 0, \dots, M - 1$ and n the frame index. Let us call $d_n(m)$ and $x_n(m)$, for $m = 0, \dots, M - 1$, the M values at frame n of the magnitude spectrum of microphone signal and of the loudspeaker signal, respectively. The magnitude values are transformed in dB and the error $e_n(m)$ is estimated as follows:

$$e_n(m) = d_n(m) - x_n(m) \quad [dB]. \quad (1)$$

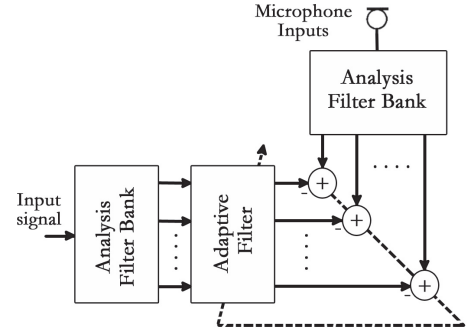


Fig. 2. Scheme for the IR identification: for each microphone, the adaptive filter is composed by M weights $H_n(k)$.

The weights $H_n(m)$ are updated frame by frame according to the following rule:

$$H_{n+1}(m) = H_n(m) + \mu(m)E \{e_n(m)\} \quad [dB], \quad (2)$$

where $\mu(m)$ is the frequency dependent adaptation step size and E is the expectation operator. For each subband m , the expectation is computed as a length- L time average of $e_n(m)$,

$$E \{e_n(m)\} = 20 \log_{10} \left(\frac{1}{L} \sum_{i=n-L+1}^n 10^{e_i(m)/20} \right) \quad [dB]. \quad (3)$$

At the end, from the weights $H_n(m)$, a length K room response is derived in the unwrapped domain interpolating the M values of each band and considering a new linear frequency vector of length K through the use of a cubic spline interpolation function [16].

2.2. Equalizer design

All the equalizer design operations are performed in the frequency domain in order to reduce the computational cost and to obtain an algorithm capable of working in real-time. A prototype filter is extracted with the mean of the identified room magnitude responses,

$$|\hat{H}_p(k)| = \frac{1}{R} \sum_{r=1}^R |\hat{H}_r(k)|. \quad (4)$$

Several approaches for prototype design were studied in [12, 15]. Compared with the other approaches, the mean in (4) is able to reduce the influence of peaks and notches of the room magnitude responses and it is often capable to obtain a better estimation of the common component of the room magnitude responses [17].

Frequency deconvolution with regularization [18] is used to design the equalizer in warped frequency domain, as follows:

$$H_{\text{inv}}(k) = \frac{|H_p(k)| e^{-j2\pi\theta \frac{k}{K}}}{|H_p(k)|^2 + \beta} \quad (5)$$

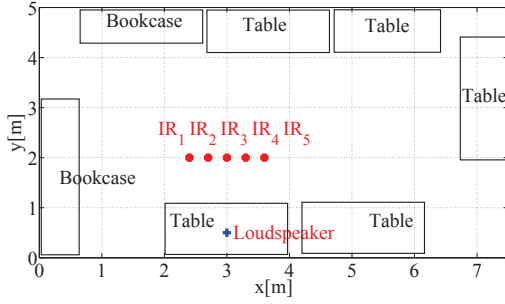


Fig. 3. Loudspeaker and microphone positions.

with θ a suitable delay experimentally determined, β a regularization factor and $k = 0, \dots, K - 1$. The regularization allows to avoid excessive gains, especially at high frequencies. From (5), by computing the FFT inverse and truncating the resulting signal in order to preserve high resolution for low frequency, a FIR equalizer is derived in time domain. The computational complexity of the frequency deconvolution method is essentially that of the inverse FFT which is an $\mathcal{O}(K \log_2 K)$ algorithm [18], therefore it is a valid approach for a real-time application.

3. EXPERIMENTAL RESULTS

In this section some experimental results are presented in order to test the effectiveness of the proposed approach. Furthermore, a comparison with the method proposed in [14] and in [19] is provided.

3.1. Objective results

Several tests have been conducted considering IRs recorded in different rooms; the adaptation procedure has been simulated with real IRs in order to have a realistic term of comparison. For sake of brevity, just the results of one room are reported here, however other tests have been performed providing similar results. Loudspeaker and microphones positions are shown in Fig. 3 together with room size: the distance between each microphone has been set to 30 cm. Measurements have been performed with professional equipment and following the procedure described in [14].

For the adaptation procedure we have considered 240 frequency components (i.e., $M = 240$) with a step size of 0.01 for both algorithms. The frequency range of equalization is 50 – 16000 Hz with a frame size of 4096 samples and a sample frequency of 48kHz.

Fig. 4 shows the results of the adaptation procedure after 10 s for one IR. It is clear that, using the proposed approach, the identified IR follows better the low frequencies behavior of the real IR than the first algorithm of [14].

Fig. 5 shows the magnitude responses of the smoothed identified IRs at the different microphone positions, of the

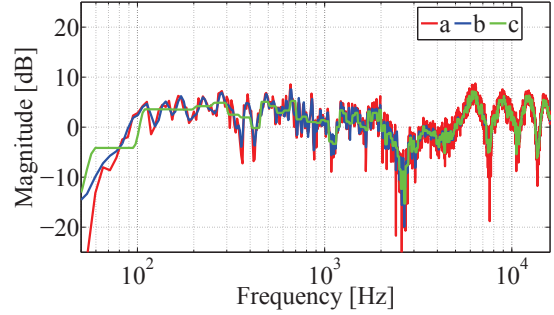


Fig. 4. (a) Real room magnitude response, (b) identified room magnitude response with the proposed approach, (c) identified room magnitude response with the approach of [14].

prototype function, and of the equalizer derived using the on-line procedure described in Section 2.2.

Fig. 6 depicts IR magnitude spectra resulting after equalization procedures at the different positions using the proposed equalizer and that of [14]. The equalization procedure should ideally lead to a flat curve around zero. Obviously, this cannot be achieved since the equalizer is derived from a set of IRs. However, Fig. 6 shows better results with the proposed equalizer than with the equalizer of [14]. In particular, an improved behavior has been obtained at low frequencies thanks to the frequency warping, which has increased the low frequencies resolution. The bad behavior of the method of [19] at low frequencies is caused by the absence of frequency warping and by the use of a single microphone signal.

Tab. 1 shows a comparison of the performance in terms of mean spectral deviation measures, as described in [14]. The mean spectral deviation is quite similar with the proposed approach and that of [14] because the improvement in the low frequency band is balanced from a less precise equalization at higher frequency. Anyway, the improvement at the low frequencies results in a broader equalized zone [17].

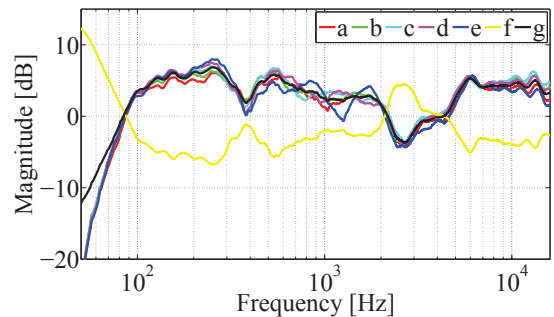


Fig. 5. (a-e) Real room magnitude responses, (f) resulting equalizer, and (g) prototype filter.

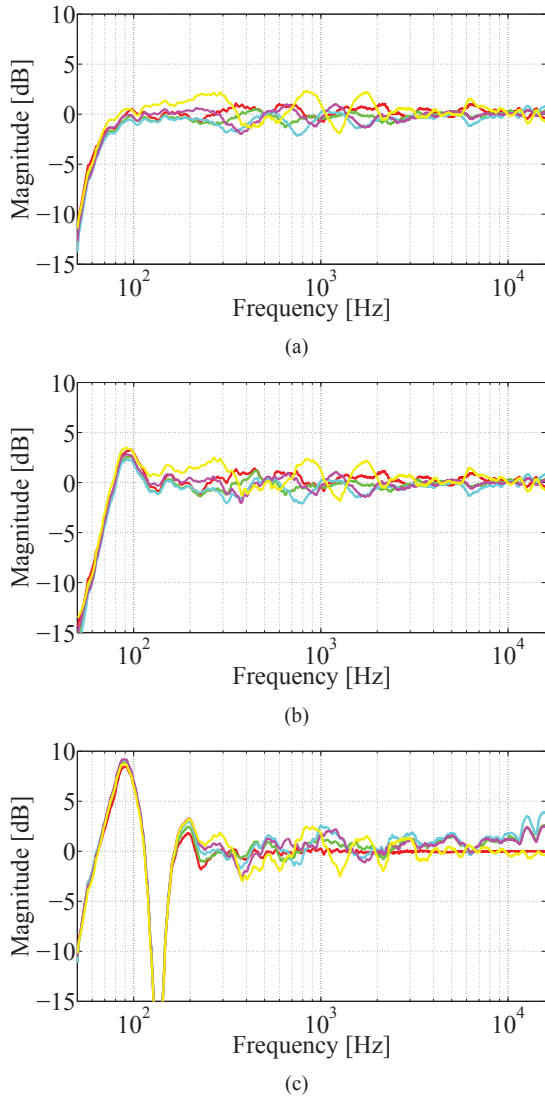


Fig. 6. Room magnitude response after equalization: (a) proposed method, (b) method of [14], (c) method of [19].

3.2. Subjective results

To assess the overall audio quality perception of the equalization procedure, subjective listening tests were performed according to the ITU-R BS.1284-1 [20] which provides guidelines for the general assessment of sound quality. Following

Table 1. Spectral Deviation Measures calculated for each IR: (a) not equalized, (b) proposed method, (c) method of [14], and (d) method of [19].

		IR1	IR2	IR3	IR4	IR5	Mean
S_{in}	(a)	2.05	2.00	2.03	2.11	2.13	2.06
	(b)	1.13	0.90	1.12	1.06	1.43	1.13
S_{fin}	(c)	1.14	0.96	1.19	1.12	1.42	1.16
	(d)	1.78	1.62	1.93	1.79	1.64	1.63

Table 2. List of sound tracks used for the listening tests.

Genre	Author	Sound Track
Rock	ZZtop	Concrete and steel
Popular	George Michael	Amazing
Classical	Chopin	Valzer Op.69 n.2
Soul	Isaac Hayes	Never can say goodbye

the recommendations, the subjects involved in the listening tests were 10 expert listeners (8 males and 2 females, ages from 21 to 35) with a technical background in acoustics. The subjective tests have been performed in the room of Fig. 3.

Before the listening test, a training test was presented to the listeners to familiarize them with the test procedure, the test materials, and the test environment, as suggested in [20]. All the subjects had the possibility to listen to each audio item in all conditions under evaluation. Care was taken not to disclose to the listeners the different methodologies employed for equalizing the audio track they were listening. Only the original not equalized signal was labeled.

For the proposed test session, three signals have been considered and compared with the original signal without equalization (i.e., the reference signal): a signal equalized using the proposed approach, a signal equalized with the magnitude equalization procedure of [14], and a signal equalized using the single position equalization of [19]. For the listening test, the subjects were instructed to make a paired comparison between the original sound track (i.e., the reference signal) and one of the equalized stimuli under test, scoring the basic audio quality. According to [20], the score was given using the continuous quality scale divided into six equal intervals (with a recommended resolution of 1 decimal place)[17]. Each stimulus was 20 s long, as suggested in the recommendation, and the listeners had the possibility to repeat the presentation. The same listening level has been used for all processed items, setting the SPL of reproduction to 70dB. As suggested in [20], the subjective data have been processed to derive the mean values and the confidence intervals. Fig. 7 shows the obtained results taking into consideration four different musical genres as shown in Table 2: the mean grade values, computed over

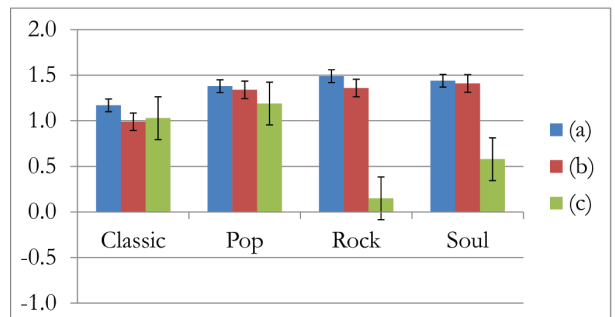


Fig. 7. Results of listening test: (a) proposed method (b) method of [14] and (c) method of [19]

all involved subjects are plotted with vertical colored blocks, while the confidence intervals calculated with a significance level of 0.05 are given by vertical black lines. The results show an improvement of the proposed approach performance compared with the unequalized signal. Also for the method of [14] the obtained results are good even if in some cases they are slightly worse compared with the proposed approach. The proposed method has superior performance also with respect to the single point equalization of [19], especially with the rock music where the signal presents rich spectral content. Therefore, the proposed approach shows good results considering both the subjective evaluation and the comparison with other approaches.

4. CONCLUSIONS

An adaptive multiple position room response equalizer has been discussed in the paper. The equalizer has been obtained by combining an adaptive procedure to determine the IRs with a technique capable to develop a multipoint equalizer. Both approaches have been designed in warped frequency domain with simple and computationally efficient techniques. Objective and subjective experimental results comparing the proposed approach with that of [14] and of [19] have been discussed. The experimental results show that the proposed adaptive multipoint equalizer is capable of improving the environment listening performance.

5. REFERENCES

- [1] S. Bharitkar and C. Kyriakakis, *Immersive Audio Signal Processing*, Springer, New York, 2006.
- [2] S.T. Neely and J.B. Allen, "Invertibility of a room impulse response," *Journal of Acoustical Society of America*, vol. 66, pp. 165–169, 1979.
- [3] J. Mourjopoulos, "On the variation and invertibility of room impulse response functions," *J. Sound Vibr.*, vol. 102, no. 2, pp. 217–228, Sept. 1985.
- [4] 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.
- [5] J. Mourjopoulos, P. Clarkson, and J. Hammond, "A comparative study of least-squares and homomorphic techniques for the inversion of mixed phase signals," in *Proc. ICASSP 1982*, Paris, May 1982, pp. 1858–1861.
- [6] M. Miyoshi and Y. Kaneda, "Inverse filtering of room acoustics," *IEEE Trans. Signal Processing*, vol. 36, no. 2, pp. 145–152, Feb. 1988.
- [7] 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.
- [8] Y. Haneda, S. Makino, and Y. Kaneda, "Multiple-point equalization of room transfer functions by using common acoustical poles," *IEEE Trans. Speech and Audio Processing*, vol. 5, no. 4, pp. 325 – 333, July 1997.
- [9] J. Mourjopoulos, "Digital equalization of room acoustics," *J. Audio Eng. Soc.*, vol. 42, no. 11, pp. 884–900, Nov. 1994.
- [10] 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.
- [11] I. Omicciolo, A. Carini, and G. L. Sicuranza, "Multiple position room response equalization with frequency domain fuzzy c-means prototype design," in *Proc. IWAENC 2008*, Seattle, Washington, USA, Sept. 2008.
- [12] A. Carini, I. Omicciolo, and G. L. Sicuranza, "Multiple position room response equalization: frequency domain prototype design strategies," in *Proc. ISISPA 2009*, Salzburg, Austria, Sept. 2009, pp. 633–638.
- [13] A. J. S. Ferreira and A. Leite, "An improved adaptive room equalization in the frequency domain," in *Proc. 118th AES Convention*, Barcelona, Spain, May 2005.
- [14] S. Cecchi, A. Primavera, F. Piazza, and A. Carini, "An Adaptive Multiple Position Room Response Equalizer," in *Proc. EUSIPCO 2012*, Barcellona, Spain, May 2011.
- [15] S. Cecchi et al., "Evaluation of a Multipoint Equalization System based on Impulse Responses Prototype Extraction," in *Proc. 127th AES Convention*, New York, NY, USA, Oct. 2009.
- [16] C. de Boor, *A Practical Guide to Splines*, Springer-Verlag, 1978.
- [17] A. Carini, S. Cecchi, I. Omicciolo, F. Piazza, and G. L. Sicuranza, "Multiple Position Room Response Equalization in Frequency Domain," *IEEE Trans. Speech and Audio Processing*, vol. 2, no. 1, pp. 122–135, Jan. 2012.
- [18] O. Kirkeby, P.A. Nelson, H. Hamada, and F. Orduna-Bustamante, "Fast deconvolution of multichannel systems using regularization," *IEEE Trans. Speech and Audio Processing*, vol. 6, no. 2, Mar. 1998.
- [19] S. Cecchi, L. Palestini, P. Peretti, F. Piazza, F. Bettarelli, and R. Toppi, "Automotive audio equalization," in *Proc. 36th AES Conference*, Deaborn, MI, USA, Jun. 2009.
- [20] ITU-R BS. 1284-1, "General methods for the subjective assessment of sound quality," Geneva, 2003.