AN ADAPTIVE MULTIPLE POSITION ROOM RESPONSE EQUALIZER

S. Cecchi, A. Primavera, F. Piazza

DIBET - A3Lab
Università Politecnica delle Marche
Via Brecce Bianche, 1
60131 Ancona Italy
s.cecchi@univpm.it

A. Carini

DiSBeF
Università di Urbino "Carlo Bo"
Piazza della Repubblica, 13
61029 Urbino Italy
alberto.carini@uniurb.it

ABSTRACT

The paper deals with an adaptive method for multiple position room response equalization. The proposed method works in the frequency domain. An adaptive estimation of the room responses at different positions is performed in the zone to be equalized and the common trend of these responses is extracted with efficient statistical operators. Then this prototype response is used to design the equalizer. The proposed approach is simple, computationally efficient and it is able to adapt to slow time variations in the room response.

1. INTRODUCTION

This paper discusses an adaptive multiple position room response equalizer. Room response equalizers improve the objective and subjective quality of sound reproduction systems by compensating the room transfer function (RTF), characterizing the path from the sound reproduction system to the listener, with a suitably designed equalizer [1]. Minimum-phase and mixed-phase room equalizers have been proposed in the literature [2]. Minimum-phase room equalizers act on the minimum-phase part of the RTF phase response and can compensate only the RTF magnitude response. In contrast, mixed-phase room equalizers cope also with the non-minimum-phase part of the RTF phase response. In principle, these equalizers can remove also some of the room reverberation [3], even though "pre-echoes" problems caused by errors in the non-causal part of the equalizer may occur.

Room equalizers are also categorized as single position or multiple position. Single position room equalizers design the equalization filter on the basis of a measurement of the room impulse response in a single location [4] and they can achieve room equalization only in a reduced zone around the measurement point (of the size of a fraction of the acoustic wavelength). It should be noted that the room impulse response varies significantly with the position in the room [5] and with time [3]. The room can be considered a "weakly non-stationary" system. To contrast audible distortions caused by equalization errors due to the impulse responses variations, the use of complex spectral smoothing and short equalization filters is often adopted [3].

Multiple position room equalizers design the equalization filter on the basis of measurements of the impulse response at different locations and are able to suitably enlarge the equalized zone. Different multiple position room equalization techniques have been proposed [1, 6, 7, 8, 9, 10, 11, 12, 13, 14, 17, 18].

A least-square approach for inverting mixed-phase room responses was presented in [6]. An exact multiple position equalization technique based on MINT, the multiple-input/multiple-output inverse theorem was proposed in [7]. A multiple-point equalization filter using the common acoustical poles of RTFs was discussed in [9]. A room response equalization system based on a k-means with splitting clustering algorithm applied to all-pole RTF measures was presented in [10]. A multiple position room response equalization technique based on fuzzy c-means clustering and frequency warping was introduced in [1, 12] and was elaborated and improved in [13, 14]. Wave domain adaptive filters [11] for the equalization of massive multichannel sound reproduction systems have also been investigated.

So far, all cited multi-point approaches employ fixed equalizers: the equalizer is designed on the basis of measurements of the room impulse response and subsequently used without any adaptation. Actually, as we mentioned before, the room is a time varying environment. The room response changes with time due to temperature and pressure variations or movements of people or other obstacles within the enclosure. An adaptive equalizer, capable to track and adapt to the room response variations, on the long term will provide better performance than a fixed equalizer.

A first adaptive multiple positions equalizer was proposed in [8], 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 [8] is very sensitive to peaks and notches in the room response and to the room response variations at different positions, and it suffers from pre-echo problems.

More recently, an adaptive frequency domain room equalizer was proposed in [15, 16]. The input sound and the measurement microphone signals are split in subbands through the use of the Fast Fourier Transform (FFT), and then the equalization is performed updating the filter weights in subbands derived from the FFT bins. Despite [15, 16] design a minimum-phase single position room equalizer, the proposed approach 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 [18] we have improved the approach of [15, 16] applying the results to a car environment. In this paper we develop a multiple position room response equalizer extending the approach of [18]. In particular we estimate the room responses at different positions in the zone to be equalized and we extract the common trend of these responses using efficient statistical operators we studied in [14, 17]. The common trend of the room responses is used to design the equalizer. 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 provides a description of the proposed algorithm, considering first the impulse response identification (Section 2.1) and then the multipoint equalizer development (Section 2.2). Section 3 reports some experimental results that illustrate the performance and the quality of the proposed approach. Finally, Section 4 contains some concluding remarks.

2. ALGORITHM DESCRIPTION

The proposed algorithm is based on an adaptive approach to iteratively estimate the impulse responses and at the same time to generate the equalizer in order to improve the audio quality in the involved environment. The approach is described in Fig. 1: R microphones positioned in the zone to be equalized capture an input signal reproduced by the loudspeaker and altered by the room transfer function. The original undistorted input signal and the microphone signals are used to characterize the environment by identifying the corresponding R room impulse responses. Then, these impulse responses are employed to generate the equalizer used to enhance the

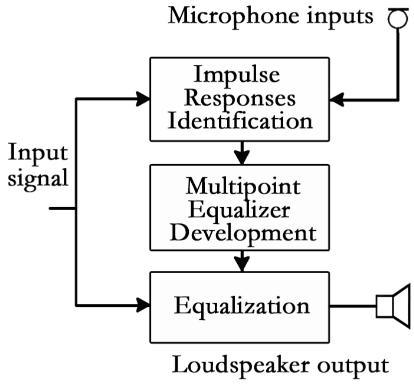


Figure 1: Overall scheme of the proposed approach.

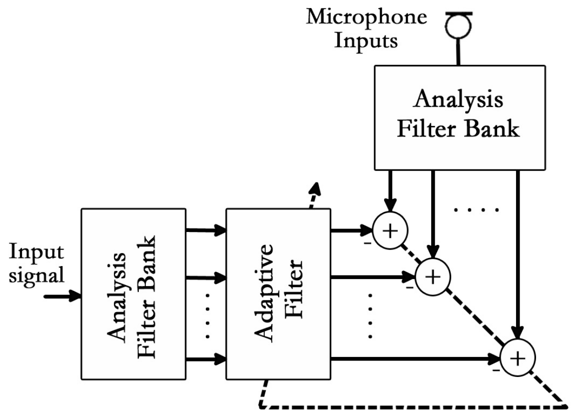


Figure 2: Scheme for the impulse response identification: for each microphone, the adaptive filter is composed of M weights $H_n(m)$.

input signal by compensating the room response. The next subsections illustrates the two steps of the algorithm.

2.1 Impulse response identification

The proposed approach is based on the work presented in [18] which was adapted from [15, 16]. The approach derives essentially from the Least Mean Square optimization applied in the frequency domain [19]. Fig. 2 schematizes the approach used for the identification of the room response from an input signal – microphone signal pair. In particular, the input signal, i.e., the signal reproduced by the loudspeaker, and the microphone signal, which captures the room dynamics, are Fourier transformed on a block by block basis with a length K frame size and the magnitude spectrum of the signals is extracted. The signals are then divided in M subbands and the average magnitude spectrum in these subbands is computed. Identification is performed in these M subbands in the frequency domain with an adaptive filter. In particular, at frame n, the adaptive filter is defined by M weights $H_n(m)$, $m = 0,\cdots,M-1$, which represent the impulse response in the frequency domain. The adaptation of $H_n(m)$ involves both the microphone signal and the input signal. Let us call $d_n(m)$ and $x_n(m)$, for $m = 0,\cdots,M-1$, the M values at frame n of the magnitude spectrum of microphone signal and of the input signal, respectively. These magnitude values are transformed in dB and the error $e_n(m)$ is computed as follows [18]:

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

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

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

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

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

In order to prevent artifacts and excessive gain values, two types of control are applied to the weights, before equalization:

1. *Band Control*. It is used to avoid excessive gains below and above suitably predetermined frequency bounds, which would lead to a longer equivalent inverse filter with the possible introduction of aliasing, or excessive boost in frequency ranges [18].
2. *Gain Control*. As in the static case, the weights level cannot exceed a predetermined threshold to avoid excessive dips and peaks.

It is worth noting that in the adaptation equation the weights do not undergo such modifications. These controls are applied only to the final weights used for the equalization. From the weights $H_n(m)$, a length K room response is reconstructed by considering a constant magnitude response in the M bands.

In order to avoid artifacts due to aliasing and signals non-stationarity, an overlap and save implementation of the system has been considered: a Short Time Fourier Transform (STFT) with a overlap of 25% has been used to manage the microphones and the loudspeaker signals. For the simulations, a frame size K of 4096 samples with an overlap of 1024 samples was adopted.

2.2 Multipoint equalizer development

Fig. 3 describes the steps of the multipoint equalizer development. All the operations are performed in the frequency domain: in this way, the computational cost is reduced and the algorithm is capable to run in real-time. The operations performed by the proposed equalizer are the following:

- A fractional octave smoothing of the magnitude responses is applied in order to attenuate peaks and notches from the room response, obtaining the R responses $\hat{H}_{n,r}(k)$, with $r = 1,\cdots,R$ and $k = 0,\cdots,K-1$ (for notational simplicity, in what follows the frame index n is dropped). The fractional octave smoothing can be performed on the magnitude spectrum using the methodology of [20].
- The prototype filter is computed from the mean of the room magnitude responses

$$|\hat{H}_{\text{p}}(k)| = \frac{1}{R}\sum_{r=1}^{R}|\hat{H}_r(k)|. \tag{4}$$

  Several approaches for prototype design were studied in [21, 22]. 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 was found that it is often capable to obtain a better estimation of the common component of the room magnitude responses [21, 22].
- In order to obtain the frequency domain inverse filter, frequency deconvolution with regularization [23] is applied to the proto-

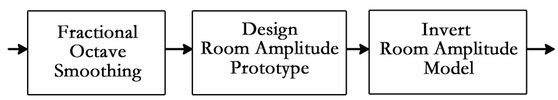


Figure 3: Flow diagram of the multipoint equalizer approach

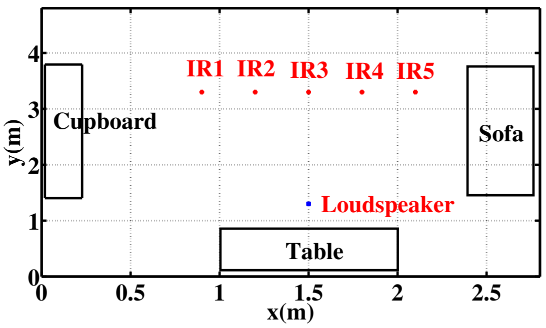


Figure 4: Loudspeaker and microphone positions.

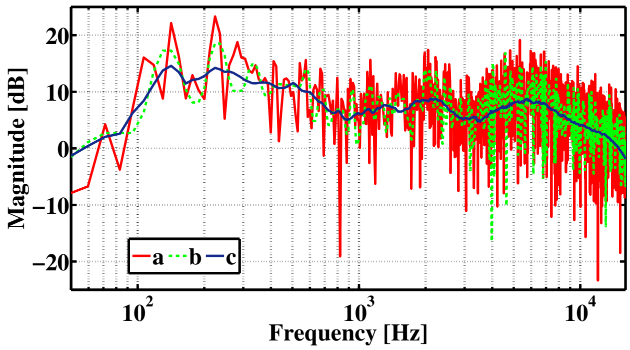


Figure 6: (a) Real room magnitude response (b) Identified room magnitude response (c) Smoothing of the identified room magnitude response.

type as follows:

$$H_{inv}(k) = \frac{H_p^*(k)}{|H_p(k)|^2 + \beta} \tag{5}$$

where β is the regularization factor, $k = 0, \cdots, K-1$, and $(^*)$ represents the complex conjugate. The regularization allows to avoid excessive gains, especially at high frequencies. For the results presented in Section 3, a small regularization factor with value 0.00001 is considered. From (5), the equalizer is determined by computing the FFT inverse and truncating the resulting impulse response; therefore, in the experimental results, the length of the equalizer was 1024 samples.
The computational complexity of the frequency deconvolution method is essentially that of the inverse FFT which is an $O(K \log K)$ algorithm [23], 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 least-square method of [6], with the adaptive least-square equalizer of [8], and with the adaptive approach of [18] is provided, in terms of spectral deviation measurement. The spectral deviation, S_D, of a frequency response $E(k)$ can be expressed as

$$S_D = \sqrt{\frac{1}{Q_h - Q_l + 1} \sum_{i=Q_l}^{Q_h} (10 \log_{10} |E(i)| - D)^2}, \tag{6}$$

where

$$D = \frac{1}{Q_h - Q_l + 1} \sum_{i=Q_l}^{Q_h} (10 \log_{10} |E(i)|), \tag{7}$$

Q_l and Q_h are the lowest and the highest frequency indexes, respectively, of the equalized band. In the experimental results an initial spectral deviation S_{in}, calculated with $E(k) = H(k)$, and a final spectral deviation S_{fin}, computed after equalization by considering $E(k) = H(k) \cdot H_{inv}(k)$ where $H_{inv}(k)$ represents the designed equalizer, are provided. A Mean Spectral Deviation Measure (MSDM) that represents the mean value of the final spectral deviation measures over the entire set of the considered impulse responses has also been considered.

Several tests have been conducted considering impulse responses recorded in different rooms; the adaptation procedure has been simulated with real impulse responses in order to have a realistic term of comparison. For sake of brevity, just the results of one room are reported here: loudspeaker and microphones positions are shown in Fig. 4 together with room size. For each measurement, the distance of loudspeaker and microphones from the floor has been set to 1.2 m. Measurements have been performed using a professional ASIO sound card and microphones with an omnidirectional response. A personal computer running NU-Tech platform has been used to manage all I/Os [24]. The impulse responses have been derived using a logarithmic sweep signal excitation [17] at 48 kHz sampling frequency. Fig. 5 shows the time behavior of the impulse response measured in the first position (i.e., IR1), and Fig. 6 the corresponding magnitude response. Fig. 7 shows the reverberation time as a function of the frequency: it shows a regular behavior at medium and high frequencies, while there is a decrease at low frequencies, due to the room size and the furniture composition.

For the test we have considered a frame size of 4096 samples and a sample frequency of 48kHz. For the adaptation procedure we have considered 240 frequency bands (i.e., $M = 240$) with a step size of 0.01. The frequency range of equalization is $50 - 16000$ Hz. Fig. 8 displays the learning curves of adaptive algorithm for

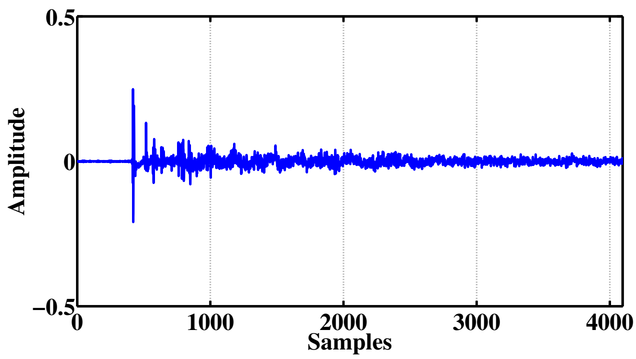


Figure 5: The room impulse response

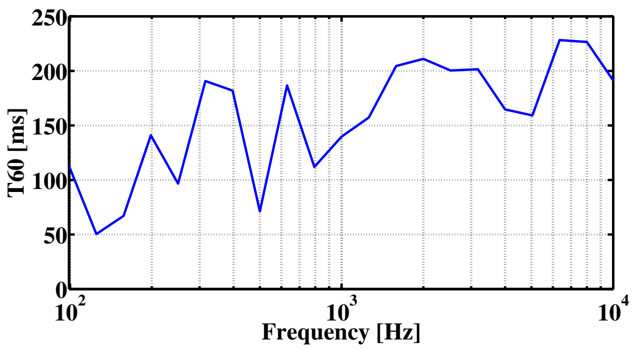


Figure 7: Room reverberation time vs. frequency.

Table 1: Spectral Deviation Measures calculated for each IR.

		IR1	IR2	IR3	IR4	IR5	MSDM
S_{in}	Not Equalized	2.94	2.89	2.77	2.83	3.05	2.89
S_{fin}	Proposed Method	2.61	2.59	2.51	2.61	2.78	2.62
	Method of [6]	2.81	2.89	2.76	2.88	3.02	2.88
	Method of [8]	2.79	2.91	2.78	2.90	3.01	2.87
	Method of [18]	2.44	3.16	3.05	3.09	3.26	3.00

the five microphone signals: after 100 iterations the residual errors are drastically reduced. Fig. 6 shows the results of the adaptation procedure after 10 s for one impulse response: it is clear that the identified impulse response follows the behavior of the real impulse response, as also demonstrated by its smoothed version.

Fig. 9 shows the magnitude response of the smoothed identified impulse responses at the different microphone positions and the equalizer derived using the on-line procedure described in Section 2.2. Fig. 10(a) depicts impulse response magnitude spectra resulting after equalization procedures at the different positions using the equalizer shown in Fig. 9. The equalization procedure should ideally lead to a flat curve around zero considering that the target curve is flat. Obviously, this is not achievable in practice considering that the equalizer is derived from a set of impulse responses. However, it is clear that the results are good and comparable with those obtained in [14, 18, 21, 22]. The behavior at low frequencies could be improved considering a non-uniform filter bank giving more resolution to those frequencies (i.e., with a warping operation).

Tab. 1 shows a comparison of the performance in terms of mean spectral deviation measures, as described in Eq. (6). It is clear that the results are good in comparison with the techniques proposed by [6] and [8]. The graphs of the magnitude responses confirm this result as reported in Figs. 10(b) and 10(c). The single position approach of [18] has also been considered as a term of comparison computing the equalizer in the single position IR1. The approach of [18] provides a worst MSDM than that obtained with the proposed approach even though a better spectral deviation is obtained in IR1, i.e., in the position used for the equalizer design. This is clearly shown in Fig. 10(d) where just one equalized impulse response tends towards the flatness.

Informal listening tests have been conducted by reproducing audio material to evaluate the perceptive effect of the equalization. The results seem to confirm the validity of the proposed approach since all involved subjects have reported positive comments and impressions on the global perceived sound image.

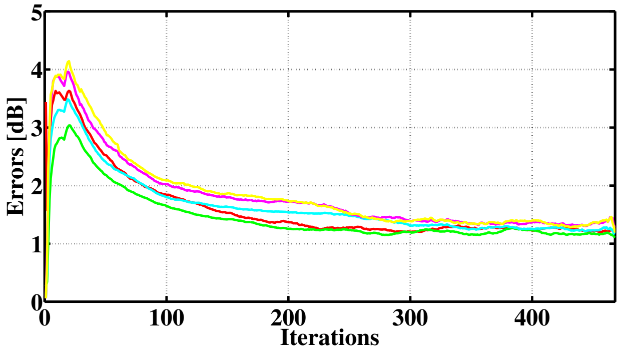


Figure 8: Learning curves of the adaptive filter.

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 impulse responses with a technique capable to develop a multipoint equalizer. Both approaches have been designed in the frequency domain with simple and computationally efficient techniques. First of all, the impulse responses are identified considering an adaptive subband structure. Then the multipoint equalizer is developed applying an inversion method to the prototype obtained by averaging the room magnitude responses. Several results have been proposed, comparing the proposed approach with well-known techniques. The proposed adaptive multipoint equalizer results capable of improving the environment listening performance in terms of objective measures.

Future works will be oriented to the extension of the proposed work considering more than one loudspeaker. A possible solution could be based on the transparent insertion of an acoustic probe signal, so that each impulse response could be separately identified.

REFERENCES

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

[2] M. Karjalainen et al., "About room response equalization and dereverberation," in *Proc. WASPAA 2005*, 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] S. Neely and J. Allen, "Invertibility of a room impulse response," *Journal of Acoustical Society of America*, vol. 66, pp. 165–169, 1979.

[5] J. Mourjopoulos, "On the variation and invertibility of room

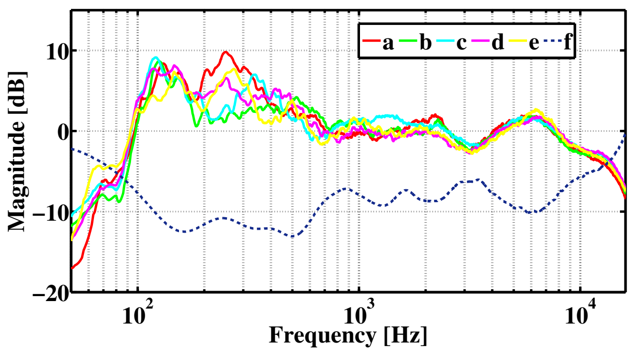
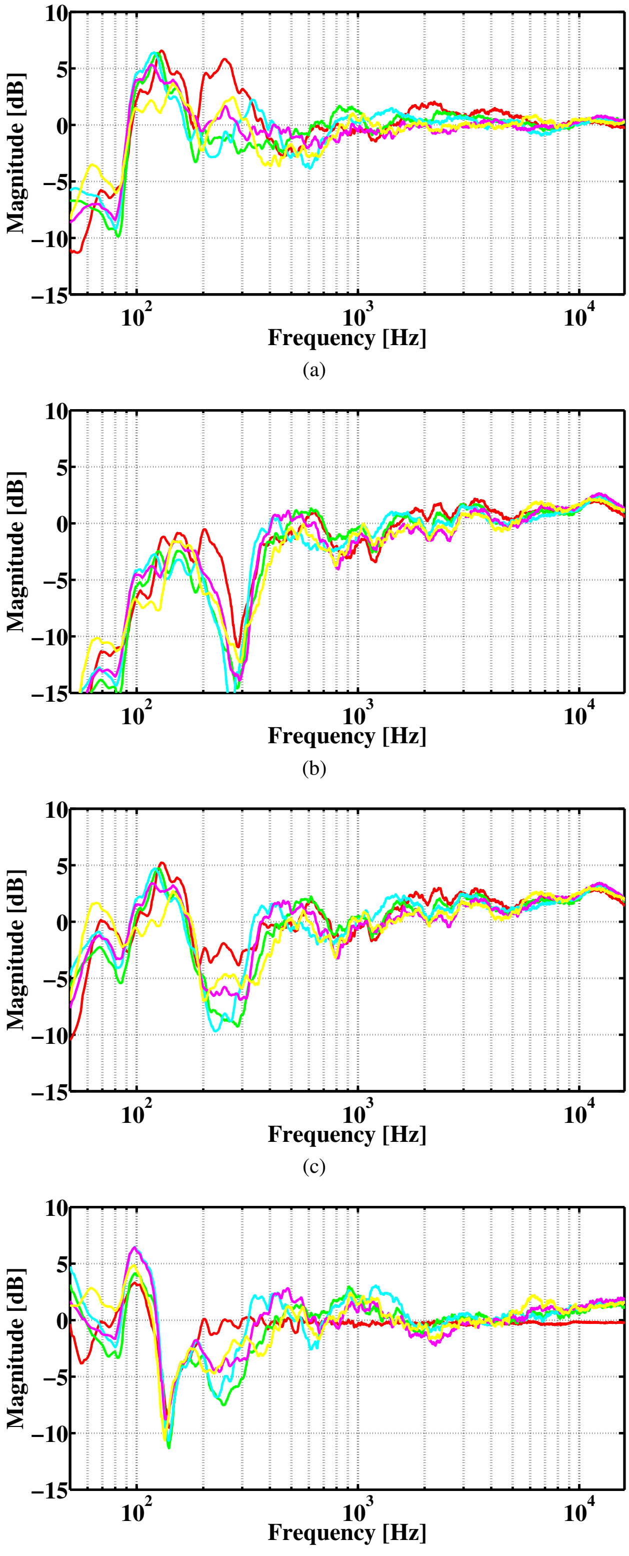


Figure 9: (a-e) Real room magnitude response and (f) resulting equalizer.



(a) (b) (c) (d)

Figure 10: Room magnitude response after equalization.(a) Proposed method. (b) Method of [6]. (c) Method of [8]. (d) Method of [18].

impulse response functions," *J. Sound Vibr.*, vol. 102, no. 2, pp. 217–228, Sep. 1985.

[6] 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, France, May 1982, pp. 1858– 1861.

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

[8] 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.

[9] 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, Jul. 1997.

[10] J. Mourjopoulos, "Digital equalization of room acoustics," *J. Audio Eng. Soc.*, vol. 42, no. 11, pp. 884–900, Nov. 1994.

[11] S. Spors et al., "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, Jul. 2007.

[12] 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.

[13] I. Omiciuolo, 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, Sep. 2008.

[14] A. Carini, I. Omiciuolo, and G. L. Sicuranza, "Multiple position room response equalization: frequency domain prototype design strategies," in *Proc. ISISPA 2009*, Saltzburg, Austria, Sep. 2009, pp. 633–638.

[15] J. Leitao, G. Fernandes, and A. J. S. Ferreira, "Adaptive room equalization in the frequency domain," in *Proc. 116th AES Convention*, Berlin, DE, May 2004.

[16] 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.

[17] 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.

[18] ——, "Automotive audio equalization," in *Proc. 36th AES Conference*, Deaborn, MI, USA, Jun. 2009.

[19] S. Haykin, *Adaptive Filter Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1996.

[20] P. D. Hatziantoniou and J. N. Mourjopoulos, "Generalized fractional octave smoothing of audio and acoustic responses," *J. Audio Eng. Soc.*, vol. 48, no. 4, pp. 259–280, Apr. 2000.

[21] S. Cecchi et al., "Multipoint equalization of digital car audio systems," in *Proc. ISPA09*, Salzburg, Austria, Sep. 2009.

[22] A. Carini, S. Cecchi, and L. Romoli, "Multipoint room response equalization with group delay compensation," in *Proc. IWAENC'10*, Tel Aviv, Israel, Sep. 2010.

[23] O. Kirkeby et al., "Fast deconvolution of multichannel systems using regularization," *IEEE Trans. Speech and Audio Processing*, vol. 6, no. 2, Mar. 1998.

[24] A. Lattanzi, F. Bettarelli, and S. Cecchi, "NU-Tech: the entry tool of the hArtes toolchain for algorithms design," *Proc. 124th AES Convention*, May 2008.