Multipoint Room Response Equalization with Group Delay Compensation

A. Carini
Department MFI - Università di Urbino “Carlo Bo”
Piazza della Repubblica 13, 61029, Urbino, Italy
Email: alberto.carini@uniurb.it

S. Cecchi and L. Romoli
A3Lab - DIBET - Università Politecnica delle Marche
Via Brecce Bianche 1, 60131, Ancona, Italy
Email: s.cecchi@univpm.it

***Abstract*—The paper copes with the problem of mixed-phase room response (RR) equalization. A minimum-phase multiple position RR equalizer recently proposed in the literature for room amplitude equalization is combined with a FIR group delay equalizer. The group-delay equalizer is designed in the frequency domain with a simple and computationally efficient but also effective and robust technique. Experiments performed on impulse responses acquired in different real environments have shown that the proposed equalizer is capable of improving the Clarity index in listening positions without introducing artifacts. Preliminary subjective tests have confirmed the improvement in the perceived audio quality.**

I. Introduction

A still open problem in the field of room response (RR) equalization is the derivation of effective, perceptually useful, mixed-phase room equalizers. RR equalizers (or room equalizers) aim to improve the objective and subjective quality of sound reproduction systems by compensating with a suitably designed equalizer the room transfer function (RTF) from the sound reproduction system to the listener [1]. 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 compensate the RTF magnitude response but they can act only on the minimum-phase part of the RTF phase response. In contrast, mixed-phase room equalizers can correct the non-minimum-phase part of the RTF phase response too. In principle they can remove also some of the room reverberation [3].

While the importance of phase equalization or of group delay compensation for improving perceived audio quality of sound reproduction systems has been recognized [4], [5], many of the mixed phase equalizers proposed in the literature [4], [6]–[9] suffer from annoying distortions, also in the form of pre-echoes (pre-ringing) effects. The first cause of degradation in equalizer performance is the variation of the room impulse response at different positions [10] and with time [3]. The use of complex spectral smoothing and short equalization filters was proposed in [3] to contrast these effects. But while the room amplitude equalization can benefit from short equalization filters, the RR group delay often varies by thousand of taps in the audio band (as shown for example in Fig. 5), and the use of long filters is often mandatory in phase or group delay equalization. Since the RTF is normally non-minimum phase, the acausal nature of the equalizer and the long filter length cause pre-ringing effects in the room response. The audibility and annoyance of these effects depend on their relative length referred to the time constant of the ear (whose value ranges between 30-200 ms) [4]. As a rule of thumb, if the length of the acausal part of the equalizer is lower than the time constant of the ear, the pre-ringing artifacts are negligible.

Since the room amplitude equalization and the group delay compensation have contrasting needs, this paper deals with the two tasks separately: we consider a well known minimum-phase RR equalization technique and we combine it with a suitably designed RR group delay equalizer. In particular, the minimum-phase multiple position RR equalizer based on fuzzy c-means clustering and frequency warping of Bharitkar and Kyriakakis [1] have been considered. In [1] a fuzzy c-means clustering algorithm is applied to extract the common trend of the room responses at different positions and frequency warping of the room responses is used to improve the equalization performances in the low frequency region. The technique of [1] was elaborated and improved in [11]–[13]. First, the fuzzy c-means clustering and frequency warping were implemented in the frequency domain [11] and, later, the fuzzy c-means clustering was replaced with different (but equally effective) averaging techniques [12], [13]. For simplicity, in this paper we do not consider the frequency warping (which will be introduced in a future paper) and we develop a group delay equalizer for the multiple position RR equalization technique presented in [13]. The proposed RR group delay equalizer is designed in the frequency domain using the same strategy of the room amplitude equalizer, i.e., a prototype group delay is derived by averaging and smoothing the group delay responses at different positions. The prototype group delay represents the common trend of the group delay response in the zone we want to compensate and it is used to design an all-pass FIR group delay equalizer. Averaging and smoothing the group delay responses allow to reduce the effect of peaks in these responses and to reduce the length of the group delay equalizer, avoiding in this way perceivable pre-ringing effects. The proposed design of this equalizer is simple, computationally efficient, but also effective, and it could be used in self-adjusting systems. The experiments performed on impulse responses acquired in different environments have shown that the proposed equalizer is able to improve the Clarity index [14] in the listening positions without introducing any meaningful artifacts. Preliminary subjective tests have confirmed the improvement in the perceived audio quality.

The rest of the paper is organized as follows. Section II describes in detail the proposed approach. Section III discusses some experimental results and some objective measurements that compare the proposed approach with that of [1]. Section IV gives concluding remarks.

II. The mixed-phase room response equalizer

Here we consider the equalization of a single channel sound reproduction system but the proposed procedure can be applied also to multi-channel systems by designing a different equalizer for each channel. Fig. 1 describes the proposed approach for the design of a mixed-phase room equalizer. Steps 1–5 are used to estimate a room amplitude equalizer as already described in [13]. On the contrary, steps 6–10 design a group delay equalizer. In particular, the following operations are performed:

1. M impulse responses of N samples length are measured at different positions in the zone to be equalized.

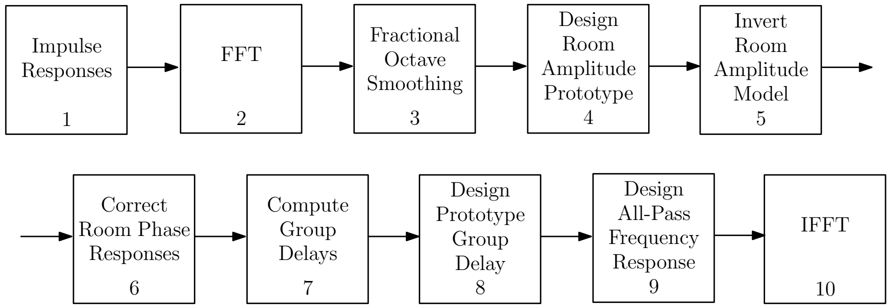


Fig. 1. Block diagram of the proposed approach.

2. The frequency responses at the M positions are computed with M FFTs of length K.
3. Complex fractional octave smoothing is performed on the M frequency responses using the methodology of [15]. This technique performs magnitude as well as phase spectrum smoothing simultaneously. Let us indicate the frequency response $H_i(k) = |H_i(k)|\, e^{j\phi_i(k)}$ with $1 \leq i \leq M$ and $\phi_i(k)$ the unwrapped phase response, the complex smoothed response $H_{cs,i}(k) = |H_{cs,i}(k)|\, e^{j\phi_{cs,i}(k)}$ is given by the following equations:

$$|H_{cs,i}(k)| = \sum_{l=0}^{K-1} W_{sm1}(m(k), l)\, |H_i((k-l) \bmod K)| \tag{1}$$

$$\phi_{cs,i}(k) = \sum_{l=0}^{K-1} W_{sm1}(m(k), l)\, \phi_i((k-l) \bmod K) \tag{2}$$

where $W_{sm1}(m(k), k)$ is a zero-phase window function and $m(k)$ is the half-window length, which is a monotonically increasing function of the frequency index k. This method simulates a well-known property of the auditory system which presents a poorer frequency resolution at higher frequencies. In this way it is possible to consider a non-uniform resolution, which decreases by increasing the frequency to obtain a less precise equalization at higher frequencies. Complex smoothing improves the robustness of the equalizer, reducing the displacement effects and increasing the equalized zone [16].
4. A prototype room amplitude response is derived taking into account all smoothed IRs. The prototype room amplitude response should represent the common trend of the room responses. In the original method of [1] the prototype response was derived with a fuzzy c-means clustering algorithm. It was shown in [12] that, without altering the performance of the room amplitude equalizer, the common trend of the room responses can be estimated from the arithmetic mean of the smoothed frequency responses. Thus, we estimate the prototype room amplitude response as follows,

$$H_p(k) = \frac{1}{M} \sum_{i=1}^{M} |H_{cs,i}(k)| \qquad k = 0, \cdots, K-1. \tag{3}$$

5. An inverse model $h_{inv}(n)$ for the prototype room amplitude response is obtained. As in [1], using the Levinson-Durbin algorithm a low order all-pole LPC model is extracted from the prototype. The inverse of the all-pole LPC model provides the FIR room amplitude equalizer $h_{inv}(n)$ of length P.
6. The phase response of $h_{inv}(n)$, $\phi_{inv}(k)$, is computed and is used to correct the smoothed phase responses $\phi_{cs,i}(k)$ in order to account for the amplitude equalizer effect:

$$\overline{\phi}_{cs,i}(k) = \phi_{cs,i}(k) + \phi_{inv}(k). \tag{4}$$

7. The group delay responses are estimated. These can be computed from the partial differences of the smoothed phase responses,

$$GD_i(k) = -\frac{K}{2\pi}\left(\overline{\phi}_{cs,i}(k) - \overline{\phi}_{cs,i}(k-1)\right) \tag{5}$$

with $1 \leq k \leq K/2+1$
8. A prototype group delay is computed by averaging the group delay at the different positions,

$$GD_p(k) = \frac{1}{M}\sum_{i=1}^{M} GD_i(k). \tag{6}$$

The prototype group delay is also smoothed with a fixed window,

$$GD_{sp}(k) = \sum_{l=0}^{K-1} W_{sm2}(l)\, GD_p((k-l) \bmod K). \tag{7}$$

The choice of a fixed window function $W_{sm2}(l)$ was determined experimentally from the observation that, after the fractional octave smoothing, the group delay is almost constant at high frequencies while it varies considerably at low frequencies. By averaging and smoothing the group delay responses, we extract the common trend of the group delay responses and we reduce the influence of the peaks, reducing in this way the length of the group delay equalizer. Therefore, in a frequency band B of interest, we want to compensate the positive group delay defined as follows

$$\overline{GD}_{sp}(k) = GD_{sp}(k) - \min_{k \in B} GD_{sp}(k), \tag{8}$$

taking into account that

$$M_L = \max_{k \in B} \overline{GD}_{sp}(k) \tag{9}$$

gives an index for the minimum length of the group delay equalizer.
9. An all-pass frequency response $H_{ap}(k') = e^{j\phi_{ap}(k')}$, suitable for compensating the prototype group delay in band B, is computed. The all-pass frequency response has length L equal to the desired length of the group delay equalizer. First, the phase response of the filter with group delay $\overline{GD}_{sp}(k)$ is computed,

$$\phi_{GD}(k) = \begin{cases} \phi_{GD}(k-1) - \overline{GD}_{sp}(k)\frac{2\pi}{K} & k \in B \\ 0 & \text{elsewhere.} \end{cases} \tag{10}$$

Then, $\phi_{GD}(k)$ is subsampled on L points and $\phi_{ap}(k')$ is estimated,

$$\phi_{ap}(k') = \phi_{GD}(Sk') - D\frac{2\pi k'}{L} \tag{11}$$

with $1 \leq k' \leq \frac{L}{2}+1$, $S = K/L$ (for simplicity, an integer), and D a delay with $D \geq M_L$. For $k' > \frac{L}{2}+1$ the frequency response $H_{ap}(k')$ is extended by conjugate symmetry.
10. The impulse response of the all-pass FIR filter is obtained by computing the IFFT of $H_{ap}(k')$.

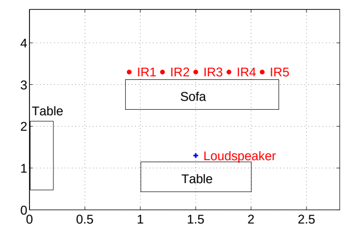


Fig. 2. Loudspeaker and microphones positions in the room.

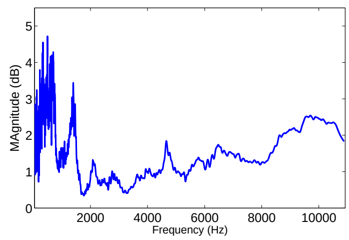


Fig. 3. Frequency response of the magnitude equalization filter.

TABLE I
MEAN SPECTRAL DEVIATION MEASURES

IR	Not Equalized	Method of [1]	Proposed Approach
1,2,3,4,5	3.2798	2.3940	2.4093

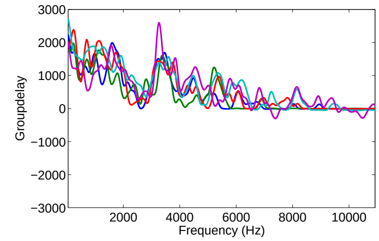
III. EXPERIMENTAL RESULTS

In this section some experimental results are provided in terms of performance comparison (the proposed technique is compared with the original method in [1]) and of quality evaluation considering both objective and subjective measures. Several tests have been conducted on a standard room of $2.8\text{m} \times 4.8\text{m} \times 2.8\text{m}$. Loudspeaker and microphones have been arranged as shown in Fig. 2. The distance of loudspeaker and microphones from the floor has been set to 1.2m. Measurements have been performed using a professional ASIO sound card and professional microphones with an omni directional response. A personal computer running NU-Tech platform has been used to manage all the I/Os [17]. The IRs have been derived using a logarithmic sweep signal excitation [18] at a 48 kHz sampling frequency. With reference with the algorithm of Section II, the following parameters have been considered in the experimental results: $N = 8760$, $M = 5$, $K = 32768$, $P = 512$, $L = 4096$, $B = 60 - 16000Hz$. The window functions $W_{sm1}()$ and $W_{sm2}()$ used for smoothing were Hanning windows. $W_{sm2}()$ had a width of 400 samples.

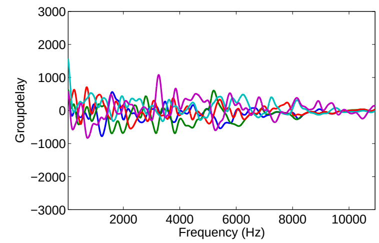
A. Performance comparison

As mentioned above, the proposed technique has been compared with the method proposed in [1] that is completely performed in time domain. Tests have considered five IRs on a line (i.e. IR1, IR2, IR3, IR4, IR5), as shown in Fig. 2. Just for comparison, Fig. 4(a) shows the time domain behavior of one IR (i.e. IR1) while Fig. 4(d) shows the five not equalized room amplitude responses.

First of all, amplitude equalization is reported and analyzed. Fig. 3 diagrams the frequency response of the amplitude equalization filter. Table I reports the mean spectral deviation values averaged over the set of measured IRs, while Fig. 4(e) and 4(f) depict their magnitude spectra obtained by applying the equalization techniques in the equalization range B. The spectral deviation gives a measure of the deviation of the magnitude frequency response from a flat one [1]. We can see that, after equalization, the proposed approach provides very good performances that are comparable with those obtained in [1]. Considering the spectral deviation, the performance seems to be very similar but it is important to underline that the proposed approach is better in the medium-high part of the spectrum while it shows worse performances in the range of $0 - 500$ Hz. This is due to the fact that the method in [1] uses a frequency warping technique to improve resolution at low frequencies.



(a)



(b)

Fig. 5. Group delay of the measured responses (a) before and (b) after the group delay equalization.

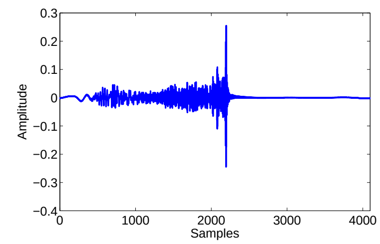


Fig. 6. Time response of the group delay equalization filter.

As for the group delay compensation, Fig. 5(a) compares the group delay behavior of the IRs (smoothed with a Hanning window of 400 samples length) before and after equalization with the proposed approach. The objective of the compensation is to have a group delay as flat as possible. Considering Fig. 5(b) it is evident that the proposed approach is able to reduce the different peaks, allowing a uniform decay rate in the frequency range of interest. The impulse response of the group delay equalizer is reported in Fig. 6. The samples of the impulse response before the peak could originate pre-echo effects but their length is short compared with the time constant of the ear and, thus, their effect is perceptually negligible.

B. Quality evaluation

In order to asses the quality of the results, objective measures have been considered. As reported in [19], the quality of an audio signal can be evaluated considering some objective quality measures based on the impulse response. In our approach we have considered the Clarity, which is defined as the logarithmic ratio of energy of the first 80ms after the main peak to the remaining energy of the IR (C80). Fig. 7 shows the frequency behavior of C80 as defined in [14]. As suggested in [20], the Clarity index should achieve a value very close to -3dB; the proposed approach reduces the value of

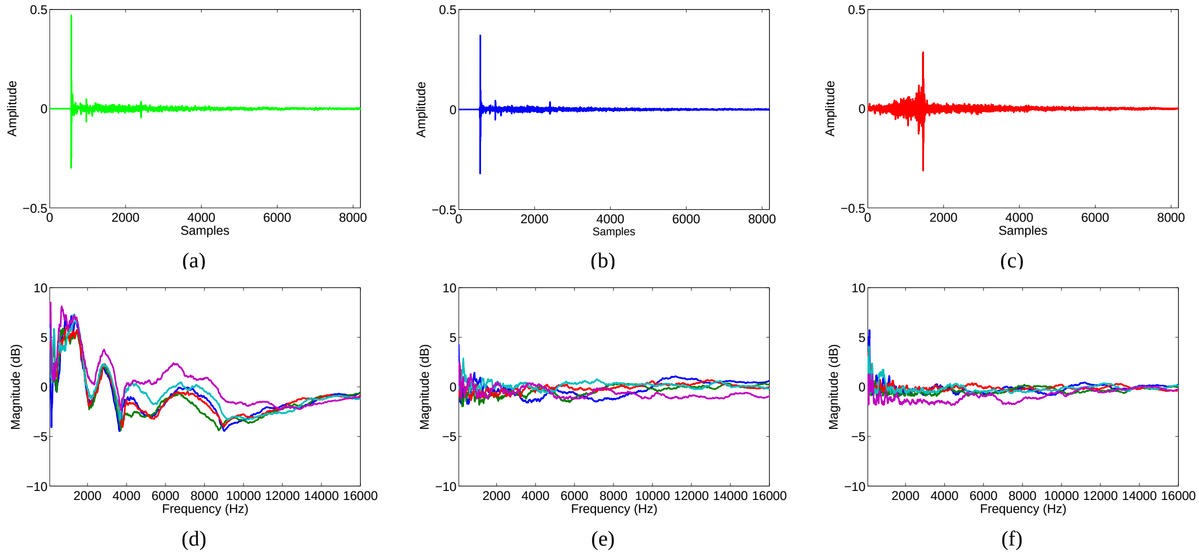


Fig. 4. Impulse response IR1 (a) before equalization, (d) equalized with [1], and (c) equalized with the proposed approach. Room amplitude frequency responses (d) before equalization, (e) equalized with [1], and (f) equalized with the proposed approach.

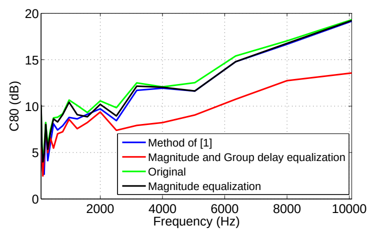


Fig. 7. Frequency behavior of Clarity index.

C80 achieving values close to the desired one, especially for the low - medium frequencies.

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.

IV. Conclusion

A multiple position mixed-phase RR equalizer has been discussed in the paper. The equalizer has been obtained by combining a room amplitude equalizer with a group delay compensator. Both filters have been designed in the frequency domain with simple and computationally efficient techniques. Averaging and smoothing the room amplitude responses and the group delay responses at different positions have been used for extracting the common trend of the responses, reducing the memory length of the equalizer and improving the robustness towards displacement effects. The proposed mixed-phase RR equalizer results capable of improving the Clarity index without introducing artifacts: informal subjective listening tests have confirmed these results.

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. of 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] B. D. Radlović and R. A. Kennedy, “Nonminimum-phase equalization and its subjective importance in room acoustics,” *IEEE Trans. Speech and Audio Processing*, vol. 8, no. 6, pp. 728–737, Nov. 2000.

[5] S. Bharitkar, C. Kyriakakis, and T. Holman, “Time alignment of Multi-way Speakers with Group Delay Equalization,” in *124th AES Convention*, Amsterdam, The Netherlands, May 2008.

[6] S. Neely and J. Allen, “Invertibility of a room impulse response,” *Journal of Acoustical Society of America*, vol. 66, pp. 165–169, 1979.

[7] 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.

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

[9] 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.

[10] J. Mourjopoulos, “On the variation and invertibility of room impulse response functions,” *J. Sound Vibr.*, vol. 102, no. 2, pp. 217–228, Sep. 1985.

[11] 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.

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

[13] S. Cecchi, L. Palestini, P. Peretti, L. Romoli, F. Piazza, and A. Carini, “Evaluation of a multipoint equalization system based on impulse responses prototype extraction,” in *Proc. of the AES 127th Convention*, 2009.

[14] International Organization of Standardization, *ISO 3382:1997 - Measurement of the reverberation time of rooms with reference to other acoustical parameters*. [Online]. Available: www.iso.org/

[15] 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.

[16] R. Genereux, “Signal processing considerations for acoustic environment correction,” in *Proc. of UK 7th Conference: Digital Signal Processing (DSP)*, Sep. 1992.

[17] A. Lattanzi, F. Bettarelli, and S. Cecchi, “NU-Tech: The Entry Tool of the hArtes Toolchain for Algorithms Design,” *Prof. of 124th AES Convention*, May 2008.

[18] S. Müller and P. Massarani, “Transfer-Function Measurement with Sweeps,” *J. Audio Eng. Soc.*, vol. 49, no. 6, 2001.

[19] S. Goetze et al., “Quality assessment for listening-room compensation algorithms,” in *Proc. of ICASSP 2010*, Dallas, Texas, USA, Mar. 2010, pp. 2450–2453.

[20] H. Kuttruff, *Room Acoustics*. London: Spoon Press, 2000.