COMPUTATION OF COMMON ACOUSTICAL POLES IN SUBBANDS BY MEANS OF A CLUSTERING TECHNIQUE

(1)Pedro Zuccarello, (1)Alberto González, (2)Juan Domingo and (3)Guillermo Ayala

(1)Audio and Communications Signal Processing Group, Universidad Politécnica de Valencia
Valencia, Spain
email: {pzuccar,agonzal}@dcom.upv.es
web: www.gtac.upv.es
(2)Institute of Robotics, Universitat de Valencia
Valencia, Spain
email: juan.domingo@uv.es
(3)Department of Statistics and Operations Research, Universitat de Valencia
Valencia, Spain
email: guillermo.ayala@uv.es

ABSTRACT

This paper presents a new and simple method for calculating the common poles of a room. A set of impulse responses is measured inside a room. ARMA models are computed for each of these measured impulse responses. Since the measured linear systems are stable, the poles of the ARMA models are located inside the unit circle of the Z plane. A clustering technique can be used to group these points into clusters. The centroids of these clusters can be interpreted as the common poles of the room. The analysis was carried out in a subband basis considering only the poles inside the pass-band of the bandpass filter. The clustering technique shows a considerable improvement compared to the averaging method proposed by Haneda et al., and the stability of the new common acoustical poles models can be theoretically assured.

1. INTRODUCTION

According to [1], the acoustic path between any pair of points inside a given room can be described by an ARMA model where the AR coefficients are common to any pair of points inside that particular room. From this point of view, a room behaves as any linear network where the resonances are a characteristic of the network, while the zeros depend on the particular input-output pair being considered.

Because of the zeros of the room, not every acoustical pole can be observed in a single impulse response. Therefore, to find the common acoustical poles of a room the first step always consists in measuring a set of impulse responses between different pairs of points inside the room. The outcome of these measurements is a set of FIR filters. We can mention three different methods that can be used to obtain ARMA models with common poles out of these FIR filters: least squares, averaging, and clustering. The first and the second were proposed by Haneda et al. [1] and the third one is proposed as a novelty into this paper.

The least squares method, based on the minimization of a cost function associated to the error between the measured FIR filters and the common acoustical poles ARMA models, requires much more computational effort than the others.

The averaging method is based upon averaging the AR coefficients of the ARMA models obtained after different impulse response measurements. This method requires much less computational effort than the first one, but the stability of the resulting common poles models can not be theoretically ensured.

The third and novel approach tackles the common poles problem not as AR polynomial coefficients but as the roots of the denominator of the ARMA models inside the Z plane. The poles of ARMA models corresponding to the impulse response measurements are located into the Z plane; then a clustering technique is used to group these poles in different clusters. The centroids of these clusters are assumed to be the common poles of the room.

ARMA models of acoustic systems is a very well motivated problem with applications ranging from room and loudspeakers responses analysis, or acoustic systems equalization, to modelling of instrument sounds [2]. ARMA models of room impulse responses using common acoustical poles principles can be very useful, leading to simplified filtering and processing structures [1], thus saving computational effort. For instance, it becomes a meaningful approach to be used in equalization of multichannel reproduction systems with large arrays of loudspeaker such as those used in wavefield synthesis systems [3]. As it was already briefly explained, previously published methods shown several problems such as computational effort (least squares), or stability of common poles models (averaging). The proposed method overcomes both problems: computational effort approach remains much lower than in the least squares case, and the stability of the resulting common poles models can be theoretically ensured.

As suggested by Karjalainen et al. [2] the accuracy of the analysis can be improved performing a sub-band filtering and downsampling of the measured impulse responses. This is the approach followed throughout this paper. Since a bandpass filter is implemented, only the poles that fall within the filter's pass-band are considered in the analysis.

A comparison between the results obtained with averaging and clustering methods is also carried out.

This paper is organized as follows. In section 2, a description of the experimental setup used for impulse responses measurement is carried out. Section 3 contains a brief description of the bandpass filtering and downsampling process. Individual ARMA models by means of the Shanks' method [4] are calculated in section 4. After that, the common acoustical poles analysis is described, implemented and evaluated in section 5. A statistical analysis of the results is carried out in section 6. Finally conclusions are exposed in section 7.

2. IMPULSE RESPONSE MEASUREMENT

A set of 4 arrays of 8 loudspeakers each was used to measure the acoustical impulse response between 32 loudspeakers and a total of 196 points distributed into a 14x14 square. Therefore a total of 6272 impulse responses were obtained. The location of the loudspeakers and the capture points can be seen in Fig. 1. The walls of the room have absorbing material, and its dimensions are 561cm x 460cm x 308cm. The measurements were simultaneously performed [5] using frequency sweeps, measuring 32 acoustical channels each time. The sampling frequency was 8000Hz. After the measurement process, impulse responses were lowpass filtered and downsampled to 4000Hz.

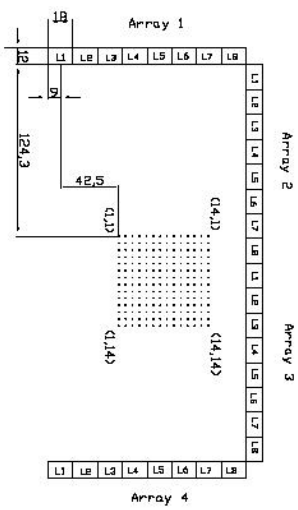

Fig. 1. Loudspeaker and microphone position for impulse response measurements. Distance between microphones is 5cm in vertical and horizontal directions. All distances are in cm.

3. IMPULSE RESPONSE FILTERING AND DOWNSAMPLING

As we already mentioned a bandpass filtering and downsampling stage was introduced. We used a bandpass filter that is a modulated version of a uniform filter bank basic window of width $2\pi/32$ implemented according to the principles described by Cvetkovic in [6]. With a sampling frequency of 4000Hz the window width corresponds to 125Hz. The window was modulated to the frequency range of 312.5Hz to 437.5Hz. Only positive frequencies were filtered, thus the filtered impulse responses are complex valued.

As the filter bandwidth is $2\pi/32$ the filtered impulse responses can be downsampled by a factor of 32. To diminish aliasing artifacts a factor of 16 was used. The filtered and downsampled impulse responses are 150 samples length.

4. INDIVIDUAL ARMA MODELS FOR MEASURED IMPULSE RESPONSES

For implementing averaging and clustering methods, firstly, individual ARMA models must be computed. The Shanks' method [4] was very useful at this point, and, as will be seen, shown to have enough accuracy to model the acoustical measured channels. Individual ARMA models can be mathematically expressed as follows:

$$\hat{H}_i(z) = \frac{\sum_{k=0}^{q} b_i(k) z^{-k}}{1 + \sum_{k=1}^{p} a_i(k) z^{-k}}, \quad (1)$$

where z is the independent complex variable of the Z-transform, $\hat{H}_i(z)$ is the ARMA model of the i-th measured impulse response, q and p are the numerator and denominator orders of the model respectively, and $b_i(k)$ and $a_i(k)$ are the numerator and denominator coefficients respectively corresponding to the i-th impulse response. We used $p = 60$ and $q = 20$. Once the 6272 models were calculated the error was evaluated according to the following expression:

$$\varepsilon_{dB} = 10 \log_{10} \left(\frac{\sum_{n=0}^{L-1} |h(n) - \hat{h}(n)|^2}{\sum_{n=0}^{L-1} |h(n)|^2} \right), \quad (2)$$

where $h(n)$ is the measured impulse response after filtering and downsampling, $\hat{h}(n)$ is the time response of the ARMA model obtained with the Shanks' method, L is the length of $h(n)$ and $\hat{h}(n)$, and ε_{dB} is the normalized energy of the error signal between the measured and modelled acoustic channel in dB. A histogram of the error values can be seen in Fig. 2. The energy error values lead us to consider that the models are accurate enough.

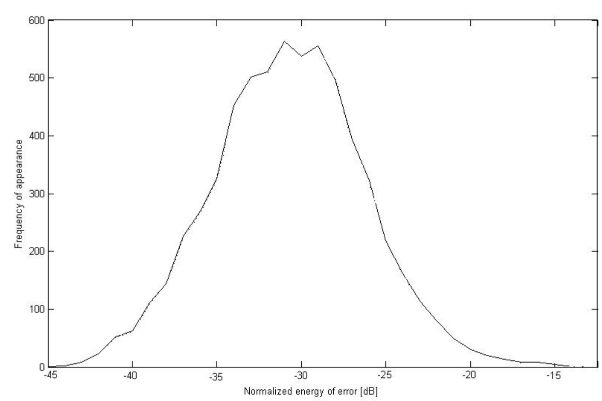

Fig. 2. Histogram of the values of the normalized energy of the error signal.

5. COMMON ACOUSTICAL POLES MODELLING

5.1 Averaging

Averaging method states that common AR coefficients can be found by averaging AR coefficients of the individual ARMA models computed in section 4 according to the following equation:

$$a_{CAP}(k) = \frac{1}{M} \sum_{i=1}^{M} a_i(k), \quad (3)$$

where M is the number of measured impulse responses, and $a_{CAP}(k)$ is the coefficient of order k of the common denominator.

We applied Eq. 3 to the individual ARMA models of the 196 acoustic channels between loudspeaker 1 of array 1 and all the points of the measurement grid.

We can go one step further to reduce the order and complexity of our common acoustical poles ARMA models. As, to reduce aliasing artifacts, the downsampling factor was 16 instead of 32, the pass-band of the downsampled bandpass filter does not occupy the whole frequency range $[0,2\pi]$. Therefore, we can reduce the common poles of the model to the ones that have more influence into the filter pass-band. Figs. 3 and 4 show a graphical representation of this concept.

It should be noticed that the frequency response of the common poles model fades outside the pass-band of the filter. This characteristic of our models should not be considered a drawback, but in fact is a strong advantage. The sub-band processing is usually performed in several contiguous frequency bands, thus, the stop band of one filter overlaps with the pass-band of the contiguous one. The less the energy falls into the stop band of the filters, the more accurate the results will be after the reconstruction filter bank is applied.

The order of the common acoustical poles ARMA models restricting the poles to the pass-band of the filter is $p = 38$ and $q = 20$. Using the common poles into the pass-band and the Shanks algorithm, new MA coefficients were found for the 6272 ARMA models. As the accuracy of the model is restricted to the pass-band frequencies, the error should be evaluated only in this frequency range. Thus, Eq. 2 should be modified to:

$$\varepsilon_{BP} = 10\log_{10}\left(\frac{\sum_{k=k_0}^{k_1}\left|H(k)-\hat{H}(k)\right|^2}{\sum_{k=k_0}^{k_1}\left|H(k)\right|^2}\right), \qquad (4)$$

where k_0 and k_1 are the lower and upper limits respectively of the pass-band of the filter, $H(k)$ is the frequency response of the measured acoustic channel after filtering and downsampling, $\hat{H}(k)$ is the frequency response of the common acoustical poles model with pole restriction, and ε_{BP} is the normalized energy in dB of the error signal restricted to the filter pass-band.

Fig. 6 shows a histogram of the values of error according to Eq. 4 together with other error histograms that will be analyzed latter in this paper together with the clustering method results.

5.2 Clustering

As we already mentioned, the clustering approach uses the location of the poles in the Z plane instead of the AR coefficients of the Z transform. As for the averaging case, we used the individual ARMA models of the 196 acoustic channels between loudspeaker 1 of array 1 and all the points of the measurement grid. Following the principles stated in the described implementation of the average method, the poles of the 196 ARMA models have been restricted to the ones with more influence into the pass-band of the filter. Fig. 5 shows the result of this operation.

A clustering technique can be used to group the poles of Fig. 5 into different clusters. We used fuzzy c-means algorithm [7] [8]. The clustering approach gives an extra degree of freedom to solve our mathematical problem. Different values for the number of clusters can be used. The final selected value would be the one that shows lower error values according to Eq. 4. We used 37, 35, 32 and 28 clusters,

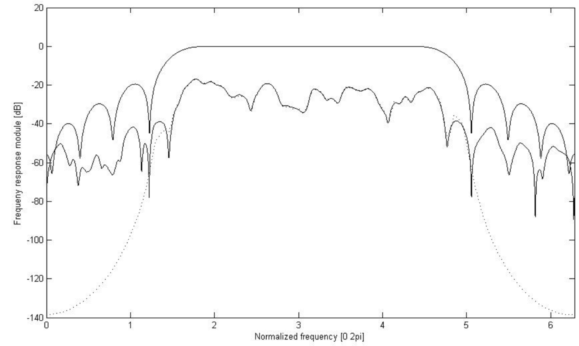

Fig. 3. Frequency domain representation of downsampled bandpass filter (upper continuous curve), filtered and downsampled acoustic channel between loudspeaker 5 of array 2 and point X=6 and Y=14 of the measurement grid (lower continuous curve), and common acoustical poles ARMA model restricting the poles to the ones that have more influence into the pass-band of the filter (dotted).

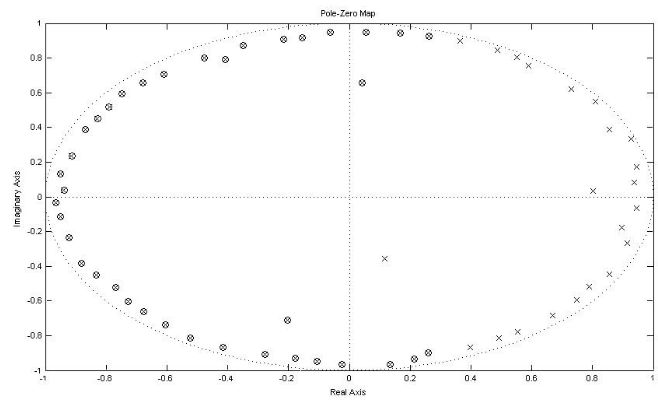

Fig. 4. Representation into the Z plane of the averaging common acoustical poles. Whole band poles (x), and poles restricted to the ones that have more influence into the pass-band of the bandpass filter (o).

calculating each time common poles ARMA models with the corresponding centroids of the clusters. As in the averaging case, new MA coefficients were found for the 6272 ARMA models. The histograms of the frequency of appearance of the error values for the 6272 models is shown in Fig. 6.

6. RESULTS ANALYSIS AND COMPARISON

The results obtained as the outcome of the different approaches followed in this paper are statistically analyzed in this section.

A one-way anova was performed showing that the means are significatively different at a level of 0.0001. Multiple comparisons amongst the means were performed. The means of the groups obtained with 35 and 28 centroids clustering are not significatively different between them. However, the other three groups 32 centroids clustering, averaging, and 37 centroids clustering have means different amongst them and different of the two means 35 and 28 centroids clustering.

Errors obtained with 32 centroids clustering have, with a significance level of 0.0001, the lowest mean value compared to the errors obtained with the other methods implemented in

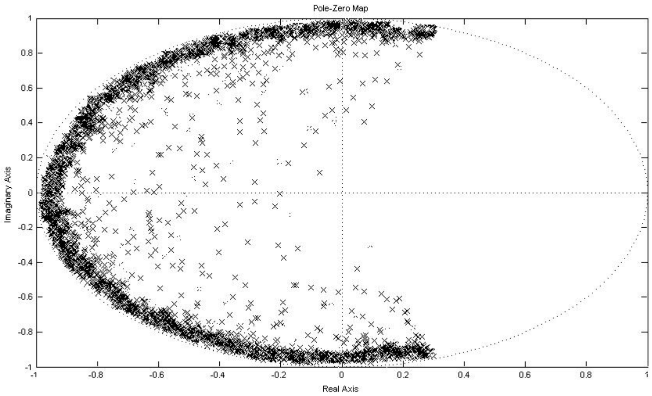

Fig. 5. Location into the Z plane of the poles of the 196 individual ARMA models of the acoustic responses between loudspeaker 1 of array 1 and all the points of the measurement grid. The poles are obviously restricted to the ones that more influence the systems' response into the pass-band of the implemented bandpass filter.

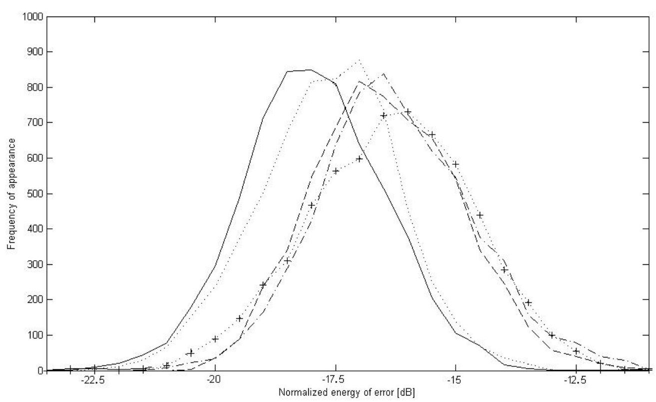

Fig. 6. Histograms of normalized error values according to Eq. 4 for averaging and posterior restriction of poles to the ones with more influence into the filter pass-band (..), restriction of poles to the ones with more influence into the filter pass-band and posterior application of fuzzy c-means clustering algorithm with 37 (- -), 35 (.+.), 32(continuous line) and 28 (-.-) centroids.

this work (see table 1).

7. CONCLUSIONS

In this paper a new approach to find the common acoustical poles of a room was proposed. Several acoustic channels of the room must be measured. Individual (non common poles) ARMA models must be calculated for the measured channels. As the analysis is performed in a sub-band basis, only the poles that fall within the pass-band of corresponding bandpass filter are considered in the process. Using the location of the poles of these ARMA models into the Z plane a clustering algorithm is used to group them into different clusters. The centroids of the clusters are assumed to be the common acoustical poles of the room in the considered frequency range.

Simulations with 37, 35, 32 and 28 centroids have been done and the results have been compared with the results obtained with the averaging technique proposed by Haneda in [1]. The histograms of error values in Fig. 6 show that the

	Mean error values [dB]			
Method	Subset 1	Subset 2	Subset 3	Subset 4
32 centroids clustering	-17.6955			
Averaging		-17.3993		
37 centroids clustering			-17.0115	
28 centroids clustering				-16.0206
35 centroids clustering				-15.9346

Table 1. Results of one-way anova test using a significance level of 0.0001 carried out with error values obtained with Eq. 4 for the 6272 measured impulse responses. The test points out that significantly different mean values can be grouped in 4 subsets.

accuracy attained by the common acoustical poles ARMA models is high enough for all the cases considered and simulated in this paper.

The statistical analysis shows that the mean of the errors obtained with the novel approach using 32 centroids is the lowest amongst all error mean values obtained in our simulations. Furthermore, the mean obtained with 32 centroids clustering is significatively different (with a level of 0.0001) to the other means. Therefore, 32 centroids clustering shows the better performance compared to the other simulated methods.

If all the poles used as starting point for the clustering technique are inside the unit circle, the resulting centroids will be inside the unit circle too. This way, the stability of the common poles models is guaranteed, in contrast to the averaging case where the stability of the resulting common poles models could not be theoretically demonstrated.

The two main drawbacks (computational effort and possible unstable solutions) of previously published methods for acoustical common poles computation are overcome with the new implemented approach.

8. ACKNOWLEDGEMENTS

This work was supported in part by grants TIC2003-08238-C01 from the Ministry of Science and Technology of Spain, and Grupos04-08 from the Valencian Agency of Science and Technology (Generalitat Valenciana).

REFERENCES

[1] Y.Haneda, S.Makino, and Y.Kaneda, "Common acoustical pole and zero modeling of room transfer functions," *IEEE Transactions on Speech and Audio Processing*, vol. 2, no. 2, pp. 320–328, April 1994.

[2] Matti Karjalainen, Paulo A.A.Esquef, Poju Antsalo, Aki Makivirta, and Vesa Valimaki, "Frequency-zooming arma modeling of resonant and reverberant systems," *Journal of the Audio Engineering Society*, vol. 50, no. 12, pp. 1012–1029, December 2002.

[3] Laura Fuster, José-Javier López, Alberto González, and Pedro Zuccarello, "Room compensation using multichannel inverse filters for wave field synthesis systems," in *118th Audio Engineering Society Convention*,

Barcelona, Spain, May 2005, Audio Engineering Society.

[4] Monson H. Hayes, *Statistical digital signal processing and modeling*, John Wiley & Sons, Inc., 1996.

[5] A.González, P.Zuccarello, G.Pinero, and M.de Diego, "Simultaneous measurement of multichannel acoustic systems," *Journal of the Audio Engineering Society*, vol. 52, no. 1/2, pp. 26–42, January/February 2004.

[6] Zoran Cvetkovic and James D.Johnston, "Nonuniform oversampled filter banks for audio signal processing," *IEEE Transactions on Speech and Audio Processing*, vol. 11, no. 5, pp. 393–399, September 2003.

[7] J.C.Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*, Plenum Press, 1981.

[8] F.Hoppner, F.Klawonn, R.Kruse, and T.Runkler, *Fuzzy Cluster Analysis*, Wiley, Chichester, 1999.