TIME FREQUENCY ANALYSIS AND PARAMETRIC APPROXIMATION OF ROOM IMPULSE RESPONSES

John C. Sarris and George E. Cambourakis

National Technical University of Athens,
Department of Electrical and Computer Engineering,
Heroon Polytecniou 9, 157 73, Athens, Greece

ABSTRACT

Rooms are modeled as linear time invariant systems, where the acoustic transmission characteristics between a specific source receiver pair are described by the room impulse response (RIR). A recursive realization of the short time Fourier transform is employed to decompose the RIR in the time frequency domain, where the different subband signals are approximated by parametric models. The pole-zero models, evaluated by the Steiglitz-McBride iteration, perform perfect reconstruction of the early reflections, whereas the decay rate of the reverberant part is sufficiently approximated. The RIR of a real room is studied as an example case.

1. INTRODUCTION

Rooms are traditionally modeled as linear time invariant systems, where the room impulse response (RIR) and the corresponding room transfer function (RTF) describe the transmission characteristics for a specific source receiver pair. In the time domain the RIR reveals the arrival times of the various reflections and the global decay rate, whereas in the frequency domain the RTF presents the overall frequency response. However, the different decay properties of different frequencies are not clearly shown from these representations.

Therefore time frequency analysis is employed to give insight on the different characteristics of different frequencies. Time frequency methods analyze the two domains jointly and thus give the ability to describe time variable and frequency dependent characteristics of non-stationary signals. RIRs are highly non-stationary signals with different decay properties at different frequencies and hence are well suited for time frequency analysis.

Methods of time frequency analysis with application to room acoustics that have appeared in the literature include the short time Fourier transform and the wavelet transform [1], [2]. In this work a recursive realization of the short time Fourier transform is used for the decomposition of RIRs in the time frequency domain.

In order to relax the need for handling large amount of data, different frequency bins are modeled using parametric models. In the literature all-zero finite impulse response (FIR) and all-pole infinite impulse response (IIR) models have been used to approximate either the RIR or the RTF [3], [4], [5]. In this work the Steiglitz-McBride algorithm is employed in order to evaluate the coefficients of pole-zero models that best match the different frequencies subbands. Since different frequencies have different decay rates and hence different complexities, models of different orders are used for the low, mid, and high frequencies. The evaluated models give an exact replica of the early reflections, whereas the decay rate of the reverberant part is sufficiently approximated.

The paper is structured as follows. In section 2 the windowed running z transform is presented and is implemented recursively using filter banks. In section 3 the Steiglitz-McBride algorithm is employed in order to evaluate the coefficients of pole-zero models that best match the different frequencies subbands. An application of the proposed methodology to a real room RIR is demonstrated in section 4. The paper is concluded in section 5.

2. RECURSIVE FILTERING FOR TIME FREQUENCY ANALYSIS

2.1. The windowed running z transform

Linear time frequency analysis is used, as it is simple to implement and provides means of analysis and synthesis of non-stationary signals. In particular we generalize and employ a recursive algorithm for the evaluation of the short time Fourier transform, which is faster compared to non-recursive realizations. The method is implemented recursively using filter banks.

The running z-transform is defined as the short time z transform of a delayed signal [6]. Hence, for a sequence $x(n)$, the running z-transform is:

$$\Phi(n,z)=\sum_{k=0}^{N-1}x(n-k)z^{-k} \tag{1}$$

For fixed n, $\Phi(n,z)$ is the z-transform in the variable k of the segment $x(n-k),\ 0\le k\le N-1$ of $x(n)$. For simplicity the above formulation assumes a rectangular window function applied to the signal. This definition is chosen because it leads to a recursive evaluation of the discrete Fourier transform coefficients.

Here we extend the above definition of the running z-transform to include windows, other than the rectangular, applied to the signal. The windowed running z-transform of a sequence $x(n)$ is defined as:

$$\Phi(n,z)=\sum_{k=0}^{N-1}x(n-k)p(k)z^{-k} \tag{2}$$

where $p(k)$ is the window function.

Substituting $k=l+1$, in equation (2) we have

$$\Phi(n,z)=z^{-1}\sum_{l=0}^{N-1}x(n-1-l)p(l+1)z^{-l}+ \\ +x(n)p(0)-x(n-N)p(N)z^{-N} \tag{3}$$

If we request $p(l+1)=Cp(l)$, where C is complex, then from equation (3) it follows that $\Phi(n,z)$ satisfies the first order recursion equation:

$$\Phi(n,z)=C\Phi(n-1,z)z^{-1}+x(n)p(0)- \\ -Cx(n-N)p(N-1)z^{-N} \tag{4}$$

and with the substitution $z=w^{-m}$, function $\Phi(n,w^{-m})$, i.e. the discrete Fourier transform coefficients, follows the simple recursion:

$$\Phi(n,w^{-m})-Cw^{m}\Phi(n-1,w^{-m})=x(n)p(0)- \\ -Cx(n-N)p(N-1) \tag{5}$$

Equation (5) defines a discrete recursive system with input $x(n)$, output $\Phi(n,w^{-m})$ and system function

$$S(m,z)=\frac{p(0)-Cp(N-1)z^{-N}}{1-Cw^{m}z^{-1}} \tag{6}$$

The system consists of one shift register with output $x(n-N)$, one delay element and two multipliers (Fig. 1). Connecting N such systems together in parallel, results in a running Discrete Fourier Series (DFS) spectrum analyser which can be realised using filter bank structure (Fig. 2). In Figure 2 it is $F_m=\Phi(n,w^{-m})$.

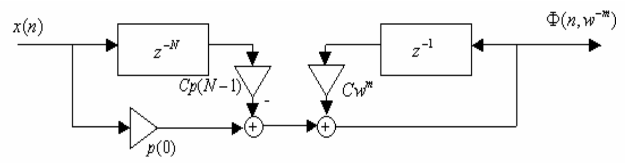

Figure 1: Elementary filter structure.

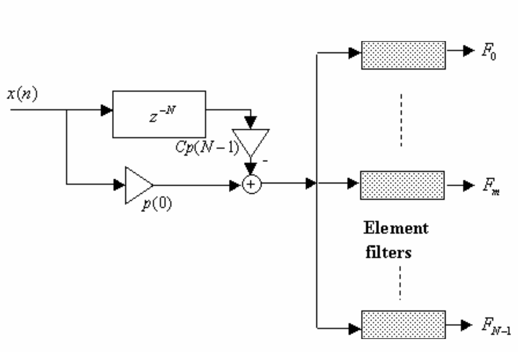

Figure 2: Recursive spectrum analyzer.

2.2. Candidate windows

From the equation $p(l+1)=Cp(l)$ we see that the windows considered are of the form:

$$p(l)=C^{l}p(0) \tag{7}$$

Substituting in equation (7) variable C with its complex representation $C=re^{j\theta}$ we have the following form for the candidate windows:

$$p(l)=\alpha r^{l}e^{j\vartheta l} \tag{8}$$

where $a=p(0)$ is complex.

Equation (8) shows that the candidate windows have an exponential form and hence equation (4) can be extended to include windows that can be decomposed in exponential functions [7]. Therefore periodic functions that are expanded in a sum of exponentials, using the Fourier series decomposition, can be used as windows. Assuming that a function p can be expanded into exponential functions $r_k^{l}e^{j\theta_k l}$, i.e.

$$p(l)=\sum_{k=0}^{K-1}a_k r_k^{l}e^{j\theta_k l} \tag{9}$$

where a_k is complex, then this function p can be used as a window and equation (5) is generalized according to

$$\Phi(n,w^{-m})=\sum_{k=0}^{K-1}\Phi_k(n,w^{-m}) \tag{10}$$

where

$$\Phi_k(n,w^{-m})-C_k w^{m}\Phi_k(n-1,w^{-m})= \\ =x(n)p_k(0)-C_k x(n-N)p_k(N-1) \tag{11}$$

In this work the Hamming function is adopted because it owns good properties for the processing of audio and acoustics signals [8]. The Hamming function is described in the time domain by the equation

$$p(n)=0.54-0.46\cos\left(\frac{2\pi}{N}n\right),\ 0\le n\le N-1 \tag{12}$$

and hence can be expanded into exponential functions as

$$p(n) = \sum_{k=0}^{2} p_k(n) \quad (13)$$

with

$$p_0(n) = 0.54,$$
$$p_1(n) = -0.46e^{j\frac{2\pi}{N}n}, \quad (14)$$
$$p_2(n) = -0.46e^{-j\frac{2\pi}{N}n}$$

and by using the representation $C_k = r_k e^{j\theta_k}$ we have

$$C_0 = 1, C_1 = e^{j\frac{2\pi}{N}}, C_2 = e^{-j\frac{2\pi}{N}} \quad (15)$$

3. PARAMETRIC APPROXIMATION OF FREQUENCY SUBBANDS

Time frequency analysis of RIRs gives insight on the time variation of their spectral content but generates large amount of data, which are difficult to handle. In order to relax this restriction the subband signals are approximated by parametric models. Since different frequencies have different decay rates and hence different complexities, models of different orders are used for the low, mid, and high frequencies.

ARMA models that yield pole-zero filters are used, since they can sufficiently approximate signals like the RIRs that exhibit non-minimum phase behavior. The linear filter

$$H_m(z) = \frac{\sum_{k=0}^{Q} b_k z^{-k}}{1 + \sum_{k=1}^{P} a_k z^{-k}} \quad (16)$$

with finite poles and zeros in the z-plane is used to model the m^{th} subband signal. The problem of obtaining the parameters of the subband models from the subband reference impulse responses is a classical non-linear problem of identification, which can be solved using the structure depicted in Figure 3.

For the evaluation of the coefficients of the pole-zero models that best match the different frequencies subbands the Steiglitz-McBride algorithm is employed [9]. The Steiglitz-McBride is a powerful iterative algorithm, with independent numerator and denominator orders, which perform perfect reconstruction of the initial part by the numerator order and sufficient approximation of the reverberation part by the denominator order.

4. APPLICATION RESULTS

The acoustic characteristics of a rectangular small-size real room, of approximate dimensions 4 by 3 by 3.5 m, are presented, as an example case. The RIR was measured using a binary maximum-length sequence excitation technique and is shown in Figure 4. The RIR is decomposed in the time frequency domain, using the recursive realization of the short Fourier transform with a hamming window of 128 points, where it can be seen that the first reflections are distributed over the whole frequency range, while mainly low and mid frequencies dominate the reverberant part (Fig. 5).

Parameters of pole-zero models to approximate the subband responses are evaluated using the Steiglitz-McBride iteration. Since most of the reverberant energy is concentrated in the low and mid frequencies the models that approximate these subbands are of higher orders than the models that approximate the high frequencies. The order of the models that approximate the low frequencies (up to 500 Hz) is P=200, Q=400, while models of order P=150, Q=300 and P=100, Q=200 are used for the mid (between 500 and 2500 Hz) and high (up to 8000 Hz) frequencies respectively.

In Figure 6 the overall RIR approximation in the time frequency domain is shown. Detailed comparison between the 36^{th} frequency bin impulse response, corresponding to 4500 Hz, and the model approximation is presented in Figure 7. It is clear that the model performs an exact reconstruction of the initial part, determined by the model's numerator order, while the decay rate of the reverberant part is sufficiently approximated.

5. CONCLUSION

The short time Fourier transform proves to be a powerful tool for the analysis of RIRs. In this work the short time Fourier transform is evaluated recursively, using a methodology, which is faster when compared to non-recursive realizations and provides the potential of real time implementation in hardware. The methodology is generalized to include as windows functions that can be expanded as a sum of exponentials.

In order to relax the need for handling large amount of data, different frequency bins are modeled using pole-zero models, whose parameters are evaluated by the Steiglitz-McBride iteration. The models perform perfect reconstruction of the initial part by the numerator order and sufficient approximation of the reverberation part by the denominator order. The original broadband signal can be reconstructed by combining the approximation subband signals.

6. REFERENCES

[1] S. K. Tang, "On the time-frequency analysis of signals that decay exponentially with time", *Journal of Sound and Vibration*, Academic Press, vol. 234, no. 2, pp. 241-258, 2000.

[2] M. Schonle, N. Fliege, and U. Zolzer, "Parametric approximation of room impulse responses by multirate systems", *in Proc. ICASSP-93*, Minneapolis, vol. 1, pp. 153-156, 1993.

[3] J. Pongsiri, P. Amin, and C. Thompson, “Modeling the acoustic transfer function of a room”, *in Proc. 12th Intern. Conf. Mathematical Computer Modeling & Scientific Computing*, Chicago, Illinois, 1999.

[4] J. Mourjopoulos, and M. A. Paraskevas, “Pole and zero modeling of room transfer functions”, *Journal of Sound and Vibration*, Academic Press, vol. 146, no. 2, pp. 281-302, 1991.

[5] T. Paatero, and M. Karjalainen, “New digital filter techniques for room response modeling”, *in Proc. 21st Audio Eng. Soc. Intern. Conf.*, St. Petersburg, Russia, pp. 173-182, 2002.

[6] A. Papoulis, *Signal Analysis*, McGraw Hill Inc., 1977.

[7] C. Richard, and R. Lengelle, “Recursive implementation of time-frequency representations and their modified version by the reassignment method”, *Signal Processing*, Elsevier, vol. 60, pp. 163-179, 1997.

[8] F. Harris, “On the use of windows for harmonic analysis with the discrete Fourier transform”, *Proc. Of the IEEE*, vol. 66, no. 1, pp. 51-83, 1978.

[9] K. Steiglitz, and L. E. MCBride, “A technique for the identification of linear systems”, *IEEE Trans. On Automatic Control*, vol. AC-10, pp. 461-464, 1965

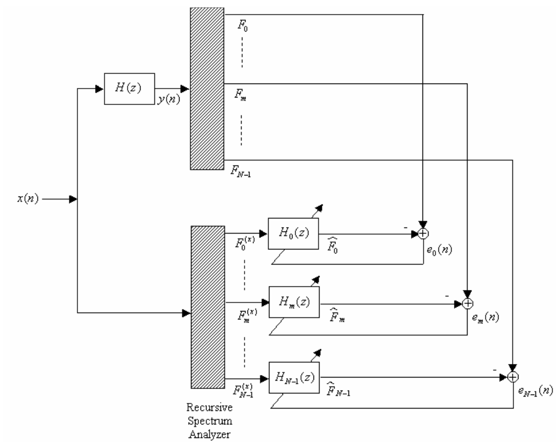

Figure 3: Structure for subband approximation.

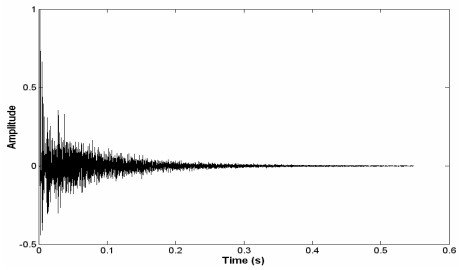

Figure 4: Room impulse response (fs=16000 Hz).

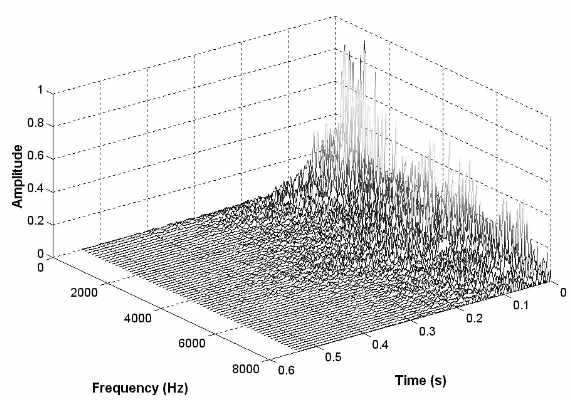

Figure 5: Time frequency analysis of RIR.

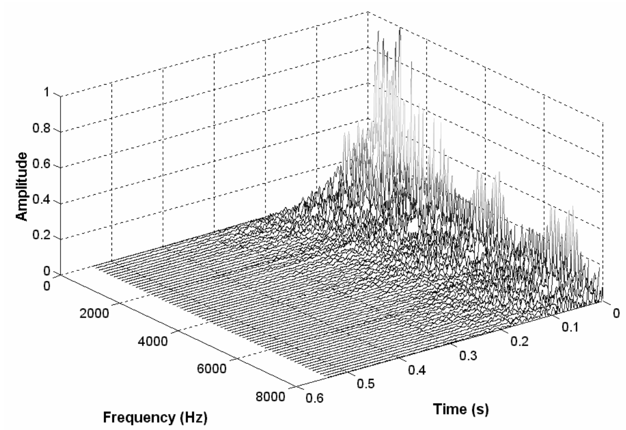

Figure 6: Time frequency approximation of RIR.

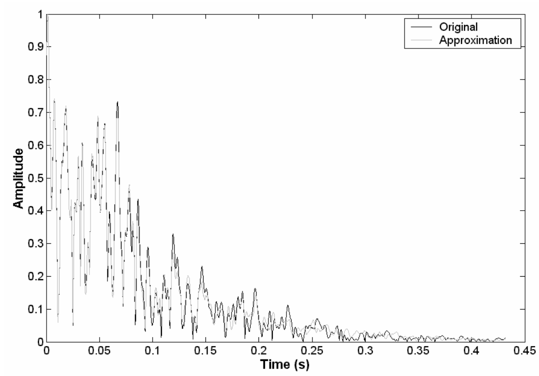

Figure 7: Detailed comparison for frequency bin 36.