Some considerations about the integration of r-ACF

Dario D'Orazio[a*], Massimo Garai[a]

[a] DIENCA – University of Bologna, Bologna, Italy

*corresponding author: dario.dorazio3@unibo.it

(Received 30 November 2009; accepted 7 December 2009)

The integration of the autocorrelation function (r-ACF) over finite temporal windows presents well known signal analysis problems. This work shows how it is possible to perform the computation taking advantage from acknowledged methods and at the same time avoiding common problems. The aim of this work is to suggest how to obtain the most accurate results, optimizing the computation algorithms. In particular, this study investigates how much different r-ACF integration methods affect the computed values of τ_e. The main results of previous r-ACF studies are discussed, then different integration solutions are applied to two reference sample tracks, analyzing the corresponding significant factors. Finally, different optimization techniques of the integration parameters are presented and compared, with particular attention to their computational load.

Key words: ACF, Temporal window

1. INTRODUCTION

In Ando's subjective preference theory for the sound field in concert halls a central role is played by the so called running autocorrelation function, r-ACF of music [1]. In particular, the effective duration of r-ACF, τ_e, has to be calculated. In existing literature different methods can be found to perform such calculation, having elaborate definitions and leading to different values of τ_e for the same music piece.

2. METHOD

2.1 Test sample tracks

To compare the results of this work with previous studies on the same subject, two motifs widely used in the literature have been selected: the motif A (Gibbons: Royal Pavane) and the motif B (Malcolm: Sinfonietta) [2].

2.2 Running analysis

The normalized r-ACF is a function of time lag (τ) that can be formulated as [5]:

$$\phi(\tau;t,2T) = \frac{\Phi(\tau;t,2T)}{\Phi(0;t,2T)} \tag{1}$$

$$\Phi(\tau;t,2T) = \int_{t-T}^{t+T} p(\xi)p(\xi+\tau)d\xi \tag{2}$$

In expression (1) $p(\xi)$ represents an A-weighted signal and 2T is the temporal window amplitude centered at time t. The effects of time window amplitude $2T$ on the r-ACF are examined in [2, 6]. Precisely, in [2] the long-time r-ACF ($2T$=30-35 s) is studied; in [4] a iterative method for fitting $2T$ by calculation of effective duration τ_e is proposed.

The choice of 2T is not standardized but often depending on the genre of music or type of acoustic signal [4].

Recent works in the domain of neurophysiology attribute to the auditory system a $2T$=200-250 ms asymmetric temporal window [3]. In the r-ACF analysis different kind of temporal windows are possible, but in this work only rectangular windows are considered, while the use of different window shapes (also considered in [5]) is left to future studies.

2.3 Evaluation of τ_e

The effective duration of the r-ACF is defined as the time lag at which the envelope of the absolute value of r-ACF is fallen by -10 dB [1]. In previous studies [2, 5] different methods for determining τ_e are presented.

The peak detecting method considers as linear the initial part of the logarithm of r-ACF amplitude and calculate τ_e by interpolation of the local peaks found in the first 5 dB of decay. Another way to find τ_e makes use of the Schroeder backward integration technique [2]. The similarity of the results of this method with the subjective impression is criticized in [2]. In [5] it is stated that the accuracy of this method is strictly related to the initial time of the backward integration.

τ_e can also be derived from Hilbert transform. In fact, the envelope of r-ACF corresponds to the absolute value of Hilbert transform [2].

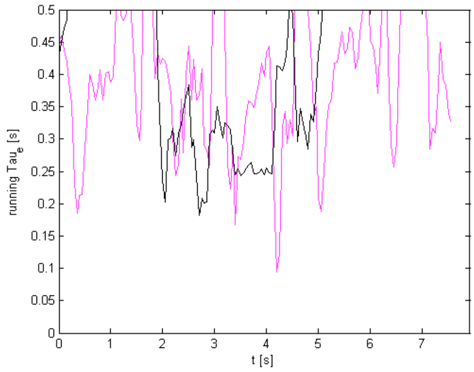

Fig. 1. Motif A, Peak detection method, 2T=250ms (magenta), 2s (black)

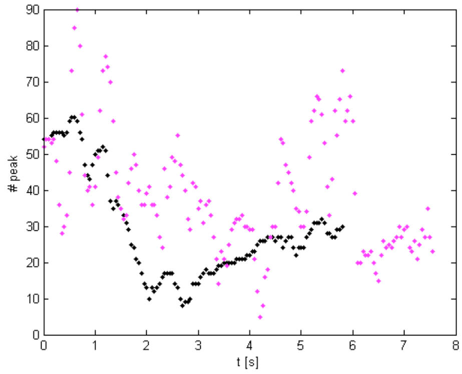

Fig. 2. Motif A, Peak detection method, 2T=250ms (magenta), 2s (black)

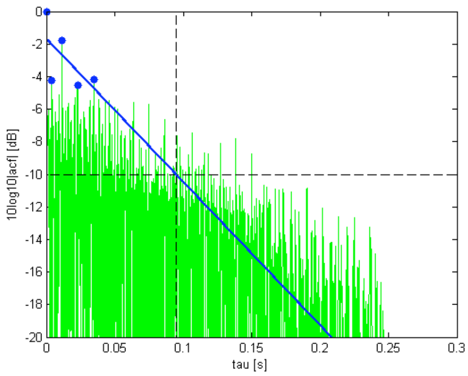

Fig. 3. $(\tau_e)_{min}$ Motif A, Peak detection method, 2T=250ms

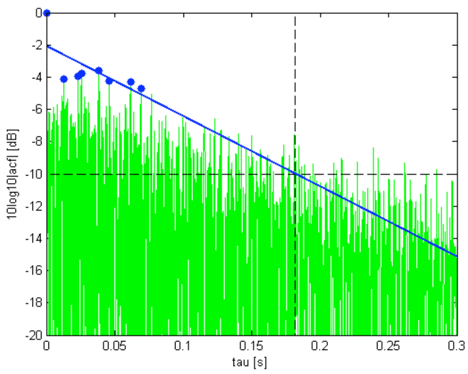

Fig. 4. $(\tau_e)_{min}$ Motif A, Peak detection method, 2T=2s

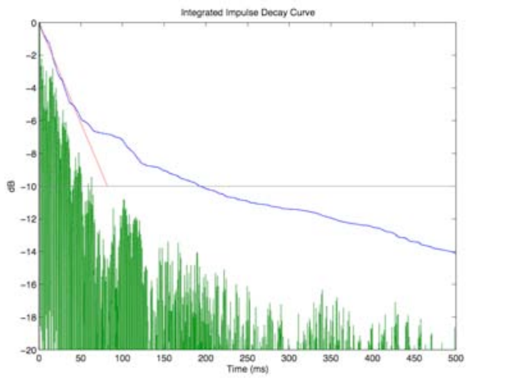

Fig. 5. Motif A, Schroeder integral method, 2T=2s

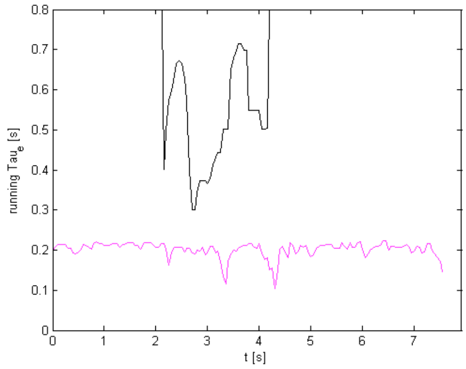

Fig. 6. Motif A, Hilbert transform m., 2T=250ms (magenta), 2s (black)

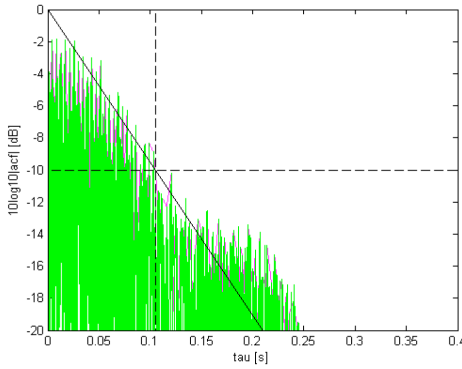

Fig. 7. $(\tau_e)_{min}$ Motif A, Hilbert transform m., 2T=250ms

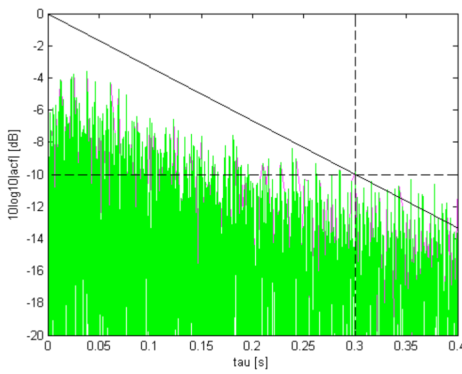

Fig. 8. $(\tau_e)_{min}$ Motif A, Hilbert transform m., 2T=2s

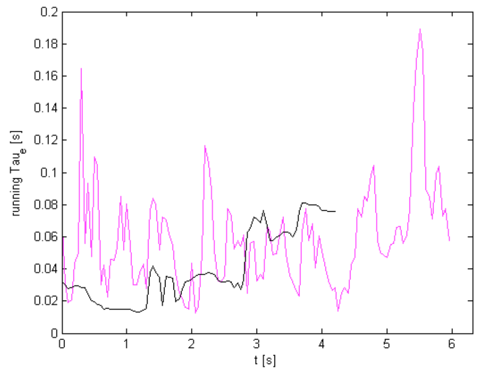

Fig. 8. Motif B, Peak detection method, 2T=250ms (magenta), 2s (black)

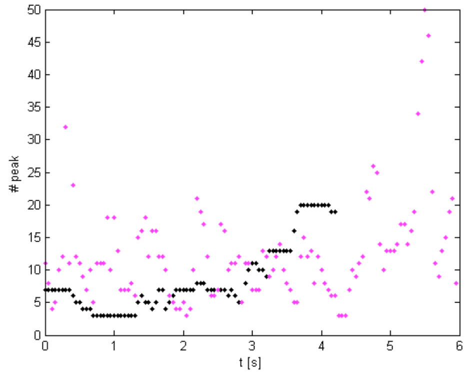

Fig. 9. Motif B, Peak detection method, 2T=250ms (magenta), 2s (black)

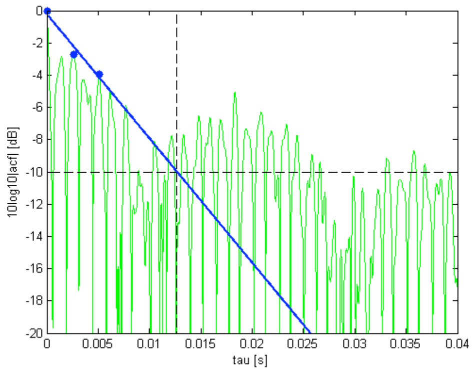

Fig. 10. Motif B, Peak detection method, 2T=250ms

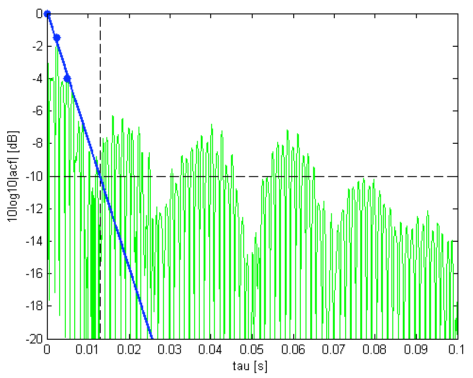

Fig. 11. $(\tau_e)_{min}$ Motif B, Peak detection method, 2T=2s

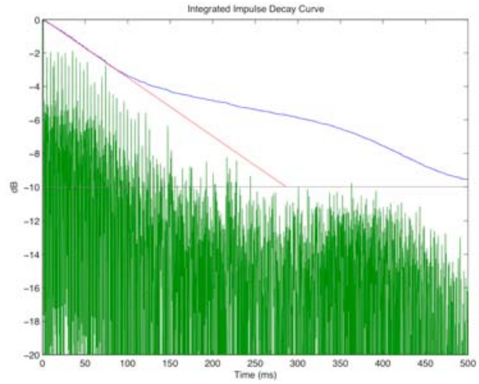

Fig. 12. Motif B, Schroeder integral method, 2T=2s

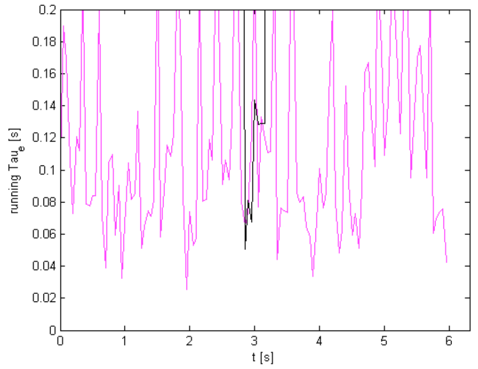

Fig. 13. Motif B, Hilbert transform m., 2T=250ms (magenta), 2s (black)

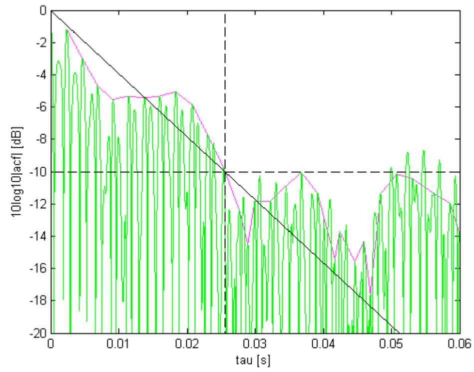

Fig. 14. $(\tau_e)_{min}$ Motif B, Hilbert transform m., 2T=250ms

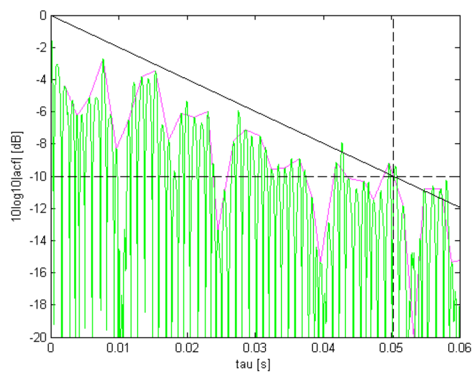

Fig. 15. $(\tau_e)_{min}$ Motif B, Hilbert transform m., 2T=2s

3. RESULTS

The peak detection method is reliable in the case of motif A ($2T$=2s) but not for motif B and motif A (case $2T$=250ms). (Fig. 3, 4, 10, 11). One possible solution to test the validity of this method can be realized drawing the number of peaks detected in the first 5 dB (Fig. 2, 9).

There are one or two peaks for motif A in the first 5 dB. If we fix the threshold at 10 dB, the regression curve may intercept too many peaks in the secondary 'lobe' of the autocorrelation function. These peaks could be too relevant in the interpolation. A temporal threshold of 50 ms has been proposed [5], that generally cuts all the peaks located in the secondary 'lobes'.

For the Motif A, it can be seen that starting the backward Schroeder integration from 250 ms there is a good approximation of the envelope of ACF, while using 2s produces an over-estimation of the r-ACF (Fig. 5). In a certain cases the values of τ_e are directly related to the choice of $2T$ and initial time of integration. We can observe something similar to what reported in [5], but with different values. According to [3] 200-250 ms can be considered a good compromise.

Unlike the peak method, the Hilbert transform is influenced by some secondary lobes only for the temporal windows where τ_e is large anyway (Fig. 6, 13). Knowing the envelope (the absolute value of Hilbert transform) it is easy to filter the secondary lobes with a negligible energetic content.

4. CONCLUSIONS

This comparison of existing literature on the determination of τ_e shows that some proposed methods are quite elaborate and difficult to be used. This work presents the analysis of the envelope of r-ACF through Hilbert transform.

It can be concluded that:

- The amplitude of the temporal window, $2T$, influences the results in all proposed methods. In abstract signal theory an amplitude as large as possible could seem advisable, while in acoustical applications related to the perception of human beings it may be interesting to relate the amplitude of temporal window analysis to the amplitude of auditory system's temporal window [7].

- The determination of τ_e by interpolation of the peaks located in the first 5 dB may result in a loss of accuracy.

- The determination of τ_e by Schroeder integration is to be performed choosing as the initial point of the backward integration a carefully selected value, otherwise the result will be overestimated. On the other hand, this is the method most suitable for real-time computation.

- The determination of τ_e by Hilbert transform turns out to be the most accurate and is now practically feasible; therefore it is proposed for future applications.

4. ACKNOWELEDGEMENTS

The authors would like to thank Dr. Ryota Shimokura for providing the test sample tracks and Ms. Simona De Cesaris for her contribution to programming.

REFERENCES

[1] Ando, Y. (1998). Architectural Acoustics-Blending Sound Sources, Sound Fields, and Listeners-. AIP Press/Springer-Verlag, New York.

[2] Ando, Y., Okano, T., and Takezoe, Y. (1989). The running autocorrelation function of different music signals relating to preferred temporal parameters of sound fields. J. Acoust. Soc. Am. 86(2), 644-649.

[3] Yabe, H., Tervaniemi, M., Sinkkonen, J., M. Huotilainen, R., Ilmoniemi, J., and Näätänen, R. (1998). The temporal window of integration in auditory information in the human brain. Psychophysiol. 35, 615-9.

[4] Mouri, K., Akiyama, K., and Ando, Y. (2001). Preliminary study on recommended time duration of source signals to be analyzed, in relation to its effective duration of the auto-correlation function, J. Sound and Vib. 241(1), 87-95.

[5] Kato, K., Hirawa, T., Kawai, K., Yano, T. and Ando, Y. (2006). Investigation of the Relation Between (τe)min and Operatic Singing with Different Vibrato Styles. J. Temporal Des. Arch. Environ. 6(1), 35-48.

[6] Takebayashi, Y., and Kido, K. (1980). Effects of time window on the estimation of coherence function for the system with delay. J. Acoust. Soc. Jpn. (E) 1(1), 17-23.