The coherence estimate function and its dependency on the room acoustic situation

Roman Scharrer und Michael Vorländer

Institut für Technische Akustik, 52066 Aachen, Deutschland, Email: rsc@akustik.rwth-aachen.de

Introduction

It is possible to use the coherence estimate function (CEF) for the blind estimation of the reverberation time from binaural signals [3] as well as to optain further information on the acoustic environment of two or more spatially distributed acoustic sensors. This can be of interest in mobile devices such as hearing aids or mobile phones where this information is important a priori knowledge for the signal processing strategies, but can also be used in sensor array applications without a controlled excitation. To get a better idea of the possibilities and limitations of a sound field analysis using the CEF, the dependencies of the CEF on the reverberation time and the signal to noise ratio (*SNR*) will be evaluated further. At this stage, the focus will be on approximately diffuse sound fields, limiting the application to scenarios with sound sources far outside the critical distance r_h.

The Coherence Estimate Function

The real coherence of a system is a fixed property. In typical applications only a coherence estimation is possible due to the limited measurement duration, fixed analysis block size or other constraints. Even though an influence of the analysis block size (or frequency resolution) on the coherence estimate is discussed in some publications [1][2], the exact influence factors in a room acoustic context are not investigated at all.

The coherence estimate should reflect the real coherence of a system, but it is biased by other factors such as the block size and number of blocks used for the estimation [1]. The coherence estimate function (CEF) reflects this dependency as it is defined as the coherence estimate of a system as a function of estimation parameters. It thus is a function depending on frequency f as well as the block size bs and number of blocks n.

$$\mathrm{CEF}(f, \mathrm{bs}, n) = \frac{\langle S_{xy}(f)\rangle_{\mathrm{bs}}^{n}}{\sqrt{\langle S_{xx}(f)\rangle_{\mathrm{bs}}^{n}\,\langle S_{yy}(f)\rangle_{\mathrm{bs}}^{n}}} \qquad (1)$$

Where $\langle\cdot\rangle_{\mathrm{bs}}^{n}$ indicates averaging over n time segments of the length bs. The block size can also be expressed as a time t_{bs}, depending on the sampling rate f_s of the signal.

The magnitude squared mean coherence estimate function (mCEF), which will be used in the following plots, is the frequency averaged squared magnitude of the CEF, calculated with a sufficiently high number of blocks to exclude a bias from n.

$$\mathrm{mCEF}(\mathrm{bs}) = \frac{1}{f_h - f_l}\cdot\int_{f_l}^{f_h} \left|\mathrm{CEF}\left(f, \mathrm{bs}, n\right)\right|^2 \mathrm{d}f \qquad (2)$$

Variances of the CEF

The characteristics of the CEF are influenced, among other things, by the room acoustic situation and the *SNR*. A Monte-Carlo simulation using the stochastic simulation method described in [3] was performed to get an idea of typical variances of the CEF. Figure 1 shows the mCEF for 100 simulations of the same room acoustic situation. Room volume and reverberation time T as well as source to receiver distance and *SNR* were kept constant.

The mCEF for large block sizes is very similar for all simulations. For low block sizes there are some deviations between the simulations. This corresponds to the assumption that the mCEF for low block sizes is dominated by the magnitude and time delay of the early reflections whereas the mCEF for big block sizes is more dominated by the energy decay of the impulse response.

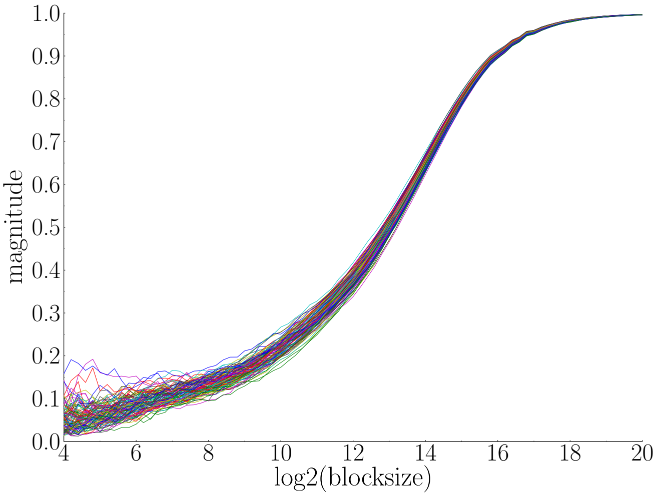

Figure 1: The mCEF for 100 simulations of the same room acoustic situation

Reverberation Time Estimation

[3] demonstrates the possibility of the estimation of the reverberation time T of a room using the CEF as indicator and a neural network as estimator. As the CEF is not identical for rooms with the same reverberation time, there must be a limit of the estimation accuracy due to the deviations of the CEF. Figure 2 shows the mCEF for a set of different reverberation times. The other room acoustic influences were kept constant. For every reverberation time 50 simulations were performed. Displayed are the mean and standard deviation values. For the higher reverberation times, the mCEFs are close to each other, but the standard deviations do not overlap. Accordingly an estimation precision in the range of 0.1 s to 0.2 s seems feasible using the CEF as indicator.

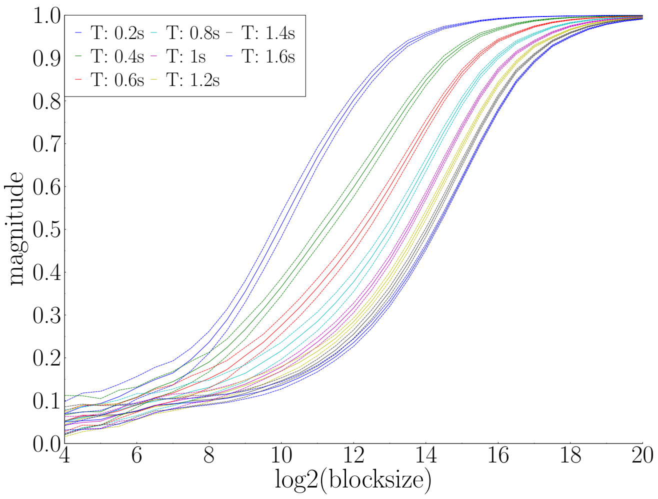

Figure 2: The mCEF for rooms with different reverberation times

SNR Estimation using the CEF

Using the CEF values at big block sizes it is possible to calculate the signal-to-noise ratio by assuming that the room is a time invariant system, source and receiver are stationary, and the noise is uncorrelated for the two sensors. This means that every acoustic signal is considered a signal, without determining whether or not it is a wanted signal. The noise only consist of sensor- and electronic noise, wind noise and quantization noise. Accordingly, the magnitude of the coherence of the signal $\gamma_{\mathrm{pp,s}}$ is unity and the coherence of the noise $\gamma_{\mathrm{pp,n}}$ is zero. As long as noise and signal are uncorrelated, the combined coherence of the signal with the energy E_{s} and the noise with the energy E_{n} can then be written as:

$$\gamma_{\mathrm{pp}} = \gamma_{\mathrm{pp,s}} \cdot \frac{E_{\mathrm{s}}}{E_{\mathrm{s}}+E_{\mathrm{n}}} + \gamma_{\mathrm{pp,n}} \frac{E_{\mathrm{n}}}{E_{\mathrm{s}}+E_{\mathrm{n}}} \tag{3}$$

$$|\gamma_{\mathrm{pp}}| = \frac{E_{\mathrm{s}}}{E_{\mathrm{s}}+E_{\mathrm{n}}} \tag{4}$$

This can be used to calculate the *SNR* from the CEF of two signals at a sufficiently high block size. In contrast

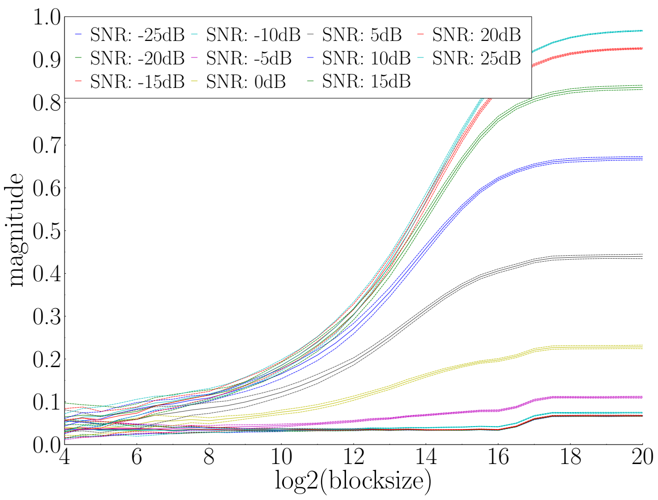

Figure 3: The mCEF for situations with different SNRs

to a single a coherence estimation, the CEF is a better choice as the characteristics of the CEF indicate which block size is needed to optain a valid coherence estimation.

An example of the *SNR* estimation is illustrated by figure 3 where the mCEF is shown for situations with one room but different *SNR*. Obviously the CEF differs for high block sizes whereas it is almost identical for the low block sizes. Figure 4 shows the result of an *SNR* estimation using a measured impulse response of a room with an artificially decreased *SNR* by adding uncorrelated noise to both channels. The *SNR* estimation is almost perfect, only for ratios below -10 dB the estimation deviates significantly from the simulated *SNR* because the limited evaluation block size results in a correlation of the noise that is not completely zero.

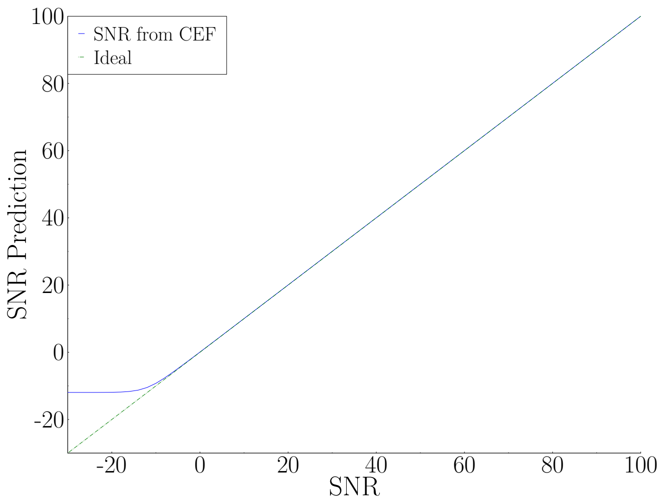

Figure 4: *SNR* estimation from the CEF vs. true *SNR*

Conclusion

The CEF of two or more spatially distributed sensors includes some information on the room acoustic situation such as the reverberation time, as well as the *SNR*. In contrast to the typical coherence estimation with a fixed block size, the CEF shows a saturation effect, allowing a robust estimation of the real system coherence, whereas a coherence estimation with a fixed block size could return a biased, typically lower, value.

Acknowledgements

The authors would like to thank Phonak for supporting this work.

References

[1] G. Carter. Coherence and Time Delay Estimation. *Proceedings of the IEEE*, 75(2):236–255, 1987.

[2] A. Piersol. Use Of Coherence And Phase Data Between Two Receivers In Evaluation Of Noise Environments. *Journal of Sound and Vibration*, 56(2):215–228, 1978.

[3] R. Scharrer and M. Vorländer. Blind Reverberation Time Estimation. In *International Congress on Acoustics, ICA*, Sydney, Australia, 2010.