3 Post-Filtering Techniques

K. Uwe Simmer[1], Joerg Bitzer[2], and Claude Marro[3]

[1] Aureca GmbH, Bremen, Germany
[2] Houpert Digital Audio, Bremen, Germany
[3] France Télécom R&D, Lannion, France

Abstract. In the context of microphone arrays, the term post-filtering denotes the post-processing of the array output by a single-channel noise suppression filter. A theoretical analysis shows that Wiener post-filtering of the output of an optimum distortionless beamformer provides a minimum mean squared error solution. We examine published methods for post-filter estimation and develop a new algorithm. A simulation system is presented to compare the performance of the discussed algorithms.

3.1 Introduction

What can be gained by additional post-filtering if the Minimum Variance Distortionless Response (MVDR) beamformer already provides the optimum solution for a given sound field?

Assuming that signal and noise are mutually uncorrelated the MVDR beamformer minimizes the noise power (or variance) subject to the constraint of a distortionless look direction response. The solution can be shown to be optimum in the Maximum Likelihood (ML) sense and produces the best possible Signal to Noise Ratio (SNR) for a narrowband input [1]. However, it does not maximize the SNR for a broadband input such as speech. Furthermore, the MVDR beamformer does not provide a broadband Minimum Mean Squared Error (MMSE) solution. The best possible linear filter in the MMSE sense is the multi-channel Wiener filter. As shown below the broadband multi-channel MMSE solution can be factorized into a MVDR beamformer followed by a single-channel Wiener post-filter. The multi-channel Wiener filter generally produces a higher output SNR than the MVDR filter. Therefore, additional post-filtering can significantly improve the SNR, which motivates this chapter.

The squared error minimized by the single-channel Wiener filter is the sum of residual noise and signal distortion components at the output of the filter. As a result, linear distortion of the desired signal cannot be avoided entirely if Wiener filtering is used. Additional Wiener filtering is advantageous in practice, however, because signal distortions can be masked by residual noise and a compromise between signal distortion and noise suppression can be found. Using MVDR beamforming alone often does not provide sufficient noise reduction due to its limited ability to reduce diffuse noise and reverberation.

M. Brandstein et al. (eds.), *Microphone Arrays*

The first concept of an electronic multi-microphone device to suppress diffuse reverberation was proposed by Danilenko in 1968 [2]. His research was motivated by Békésy's [3] observation that human listeners are able to suppress reverberation if sounds are presented binaurally. In Danilenko's reverberation suppressor a main microphone signal is multiplied by a broadband gain factor that is equal to the ratio of short-time cross-correlation and energy measurements. Two auxiliary microphones were used to measure correlation and energy. Danilenko already noted that such a system would also suppress incoherent acoustic noise. However, the proposed analog, electronic tube version of this system was not realized at that time. Another proposal in [2] was to evaluate squared sum and differences of two microphone signals, an idea that later was developed independently by Gierl and others in the context of digital multi-channel spectral subtraction algorithms [4], [5], [6], [7], [8].

According to Danilenko, his correlation-based concept was first realized during Blauert's stay at Bell Labs. In [9], Allen *et al.* presented a digital, two-microphone algorithm for dereverberation based on short-term Fourier-Transform and the overlap-add method. In 1984, Kaneda and Tohyama extended the application of the correlation based post-filters to noise reduction [10]. The first multi-microphone solution was published by Zelinski [11], [12]. Simmer and Wasiljeff showed that Zelinski's approach does not provide an optimum solution in the Wiener sense if the noise is spatially uncorrelated, and developed a slightly modified version [13]. A deeper analysis of the Zelinski and the Simmer post-filter can be found in [14], [15].

In the last decade, several new combinations and extensions of the post-filter approach were published. Le-Bouquin and Faucon used the coherence function as a post-filter [16], [17] and extended their system by a coherence subtraction method to overcome the problem of insufficient noise reduction at low frequencies [18], [19]. The problem of time delay estimation and further improvement of the estimation of the transfer function was independently addressed by Kuczynski *et al.* [20], [21] and Drews *et al.* [22], [23]. Fischer and Simmer gave a first solution by associating a post-filter and a generalized sidelobe canceler (GSC) to improve the noise reduction in case the noise field is dominated by coherent sources [24], [25]. Another system for the same task was introduced by Hussain *et al.* [26] and was based on switching between algorithms. The same strategy of switching between different algorithms, where the decision is based on the coherence between the sensors, can be found in [27], [28]. Furthermore, Mamhoudi and Drygajlo used the wavelet-transform in combination with different post-filters to improve the performance [29], [30]. Bitzer *et al.* [31], [32] proposed a solution with a super-directive array and McCowan *et al.* used a near-field super-directive approach [33].

Reading these papers we find that a theoretical basis for post-filtering seems to be missing. Therefore, an analysis based on optimum MMSE multi-channel filtering is presented in the following section.