Adaptive Beamformer Simulating a Highly Directive Microphone

Christof Faller
Illusonic GmbH, 1025 St-Sulpice, Schweiz, Email: christof.faller@illusonic.com

Abstract

High quality directive sound capture ideally requires a microphone with a frequency-independent directional response shape, to prevent coloration of off-look-direction sources. This also implies a frequency-independent relative diffuse field response. High quality directional gradient microphones nearly achieve this goal. On the other hand, in practice, beamformers fail regarding this requirement and are thus not used when high audio quality is required. We describe an adaptive beamformer with the target of a frequency-independent directional and diffuse response. The proposed beamformer can be measured like a microphone and enables high quality highly directive sound capture.

Introduction

An ideal microphone has a frequency invariant directional response. The directional response also defines the microphone's response to a diffuse soundfield. The more directive a microphone's directional response is, the more diffuse sound is attenuated relative to omni-directional sound capture.

There are many applications where it is desired that a microphone is directive. An ideal directive microphone captures a desired source located at the direction towards which the microphone is oriented, while other sources at other directions and diffuse sound are attenuated frequency invariantly. In film, broadcast, telephony, and tele-conferencing applications it is desired to capture the desired sound source, while attenuating undesired noise sources and reverberation.

First order directive differential microphones [1], such as cardioid and super-cardioid microphones, have a relatively frequency-invariant directional response within a large part of audible frequencies. Compared to non-directional microphones, their sensitivity at low frequencies is compromised, but still usable even for professional applications. More directivity would be desired for many applications, but higher order differential microphones have a too compromised sensitivity at low frequencies.

Beamformers [2], combining the signals of several microphones, have been proposed to achieve higher directivity. But in order to achieve reasonable sensitivity and a directional response invariant over a large frequency range, the aperture of the microphone array has to be large and the microphones need to be densely spaced. This implies the need for many microphones, making beamformers impractical due to the cost of microphones and cost of processing many microphone signals.

Adaptive beamformers [3] are more effective than static beamformers in terms of signal-to-noise ratio (SNR) improvement for a source in a given direction. This is so because adaptive beamformers frequency dependently SNR-optimize their directional response to the specific acoustic scene that is being captured. An adaptive beamformer's directional and diffuse field responses are not pre-determined and in general frequency variant.

Proposed Beamformer

We propose an adaptive beamformer with the goal of simulating a highly directive microphone with a pre-determined directional and diffuse field response, similar to [4], but with improved parameter estimation (5). The beamformer uses an array of two coincident cardioid microphones, oriented towards forward (the desired beam direction) and backwards. The thin solid and dotted lines in Figure 1 show the directional responses of two such cardioid microphones. The signals are processed in a short-time Fourier transform (STFT) domain. The STFT representations of the forward and backwards pointing cardioid signals are denoted $X_1(k,i)$ and $X_2(k,i)$, respectively, where k is the spectrum (time) index and i the frequency index.

We assume that the microphone signal $X_1(k,i)$ can be written as

$$X_1(k,i) = a(k,i)X_2(k,i) + N_1(k,i)\,, \tag{1}$$

where $a(k,i)$ is a time and frequency dependent gain factor determining the cross-talk between both microphone signals $X_1(k,i)$ and $X_2(k,i)$. It is assumed that all signals are zero mean and that $X_2(k,i)$ and $N_1(k,i)$ are un-correlated.

The goal of the proposed adaptive beamformer is to modify the directional and diffuse response of the cardioid signal $X_1(k,i)$. The beamformer output signal is

$$Y_1(k,i) = c(k,i)\left(X_1(k,i) - w(k,i)X_2(k,i)\right)\,, \tag{2}$$

where $w(k,i)$ and $c(k,i)$ relate to the directional and diffuse response of $Y_1(k,i)$, respectively. The predictor $w(k,i)$ is computed as

$$w(k,i) = \min\left\{\mathrm{Re}\left\{\frac{\mathrm{E}\{X_1(k,i)X_2(k,i)^\star\}}{\mathrm{E}\{X_2(k,i)X_2(k,i)^\star\}}\right\},\ q\right\}\,, \tag{3}$$

where * denotes complex conjugate, Re{.} is the real part of a complex number, and q is a limit applied to the predictor magnitude. The beamformer's directional responses for different values of q are shown in Figure 1.

The diffuse field response can be freely specified, e.g. $c_0 = 0.25$, corresponding to a diffuse response 12 dB below the

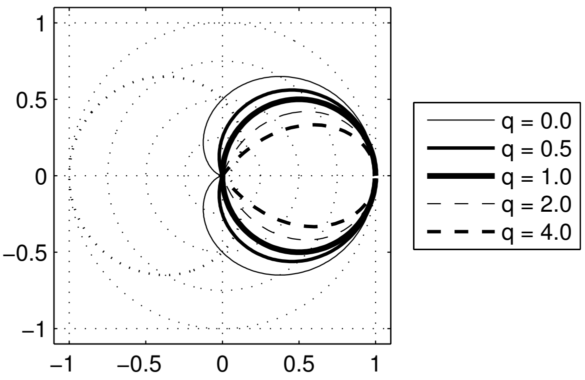

Figure 1: Directional responses of the beamformer for different predictor magnitude limits q.

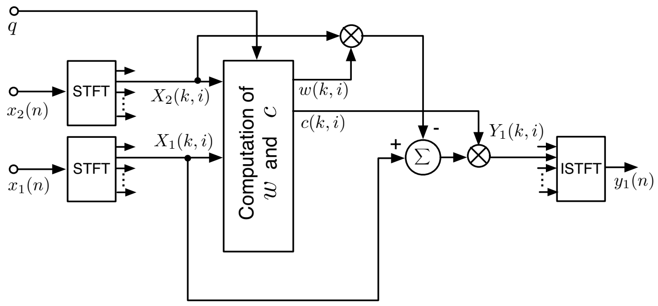

Figure 2: Scheme illustrating the processing of the beamformer.

free-field forward response. When there is only direct sound, the post-scaling factor is $c(k,i) = 1$ and if there is only diffuse sound it is $c(k,i) = c_0$. A post-scaling factor $c(k,i)$, imposing a diffuse response for any mix of direct and diffuse sound, is

$$c(k,i) = (1 - \alpha(k,i)) + \alpha(k,i)c_0\,, \qquad (4)$$

where $\alpha(k,i)$ is a diffuseness indicator, e.g.

$$\alpha = \frac{\frac{1-\max\{\Phi_{12},\ \Phi_{\text{diff}}\}}{1-\Phi_{\text{diff}}} \min\{\mathrm{E}\{X_1X_1^\star\},\ \mathrm{E}\{X_2X_2^\star\}\}}{\frac{1}{2}\mathrm{E}\{X_1X_1^\star\} + \frac{1}{2}\mathrm{E}\{X_2X_2^\star\}} \qquad (5)$$

with $\Phi_{12}(k,i)$ and Φ_{diff} being the estimated and diffuse field cross-correlation coefficient, respectively. For brevity of notation the time and frequency indices have been ignored in (5).

The time constants for computing the statistics, i.e. E{.} and $\Phi_{12}(k,i)$, are between 20 and 80 ms. The beamformer adapts fast enough that, when measured with microphone measurement systems, it yields responses similarly as claimed. Examples of such measurements are given in [5]. The various processing stages of the proposed beamformer are illustrated in Figure 2.

The top part of Figure 3 shows the responses of an ideal cardioid microphone to sound fields with different relative amounts of direct sound, where DTR is the direct sound power to total sound power ratio. For DTR = 1 we get the free-field directional response and for DTR = 0 the diffuse field response (omni-directional, 4.8 dB below unity).

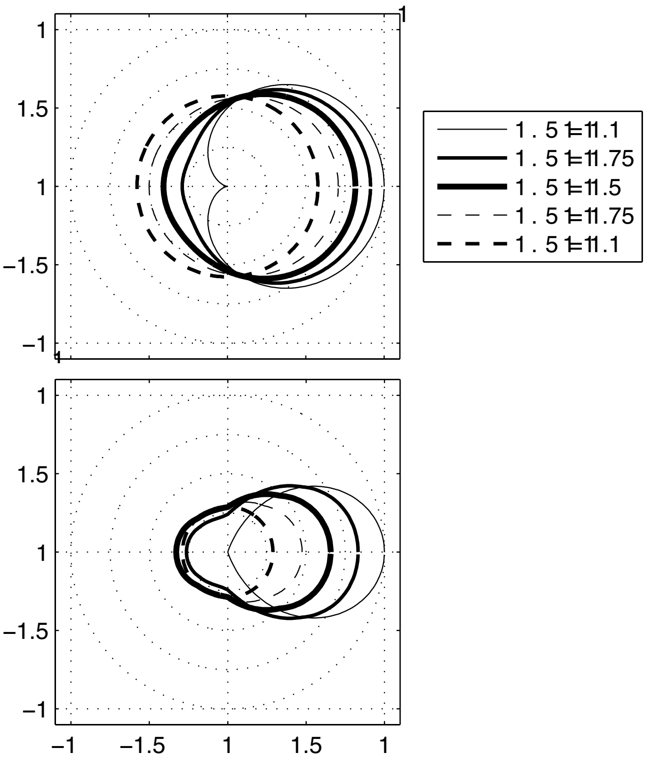

Figure 3: Direct sound and diffuse field responses of a cardioid (top) and the proposed beamformer (bottom).

The bottom part of the figure shows the corresponding responses, obtained with a steady state simulation, for the proposed beamformer with $q = 2$ and $c_0 = \frac{1}{3}$, corresponding to a microphone with a directivity index of about 10 dB. The behavior of the beamformer for different mixes of direct and diffuse sound are very similar to that of the cardioid microphone, but providing higher directivity.

References

[1] Harry F. Olson, "Gradient microphones," *J. Acoust. Soc. Am.*, vol. 17, no. 3, pp. 192–198, January 1946.

[2] B. D. Van Veen and K. M. Buckley, "Beamforming: A versatile approach to spatial filtering," *IEEE ASSP Magazine*, pp. 4–23, Apr. 1988.

[3] L. J. Griffiths and C. W. Jim, "An alternative approach to linearly constrained adaptive beamforming," *IEEE Transactions on Antennas and Propagation*, vol. 30, pp. 27–34, January 1982.

[4] C. Faller, "A highly directive 2-capsule based microphone system," in *Preprint 123rd Conv. Aud. Eng. Soc.*, Oct. 2007.

[5] H. Wittek, C. Faller, C. Langen, A. Favrot, and C. Tournery, "Digitally enhanced shotgun microphone with increased directivity," in *Preprint 129th Conv. Aud. Eng. Soc.*, Nov. 2010.