Adaptive Beamforming for Sonar Audio

Chaoying Bao

Defence Science Technology Organisation, A-51, HMAS Stirling, Rockingham, WA 6958, AUSTRALIA

ABSTRACT

Sonar audio is a major tool used by sonar operators to assist in classifying acoustic contacts. In this paper we discuss some issues that arise when adaptive beamforming is used for sonar audio. Frequency domain beamforming is used to reduce computational cost, and diagonal loading of the cross-spectral matrix is used to obtain the best quality output. The effectiveness of a robust Capon beamformer (RCB) with nonuniform loading is compared with that of a minimum variance distortionless response beamformer with uniform loading (MVDRUL). For the tests described in this paper, which involve a signal together with a strong interference and background noise, RCB produced the highest quality output and was more robust than MVDRUL. On the other hand, conventional beamforming (CBF) failed to provide a satisfactory output for any test.

INTRODUCTION

Sonar audio is an important tool that sonar operators use when classifying acoustic contacts (Barger 1997). Beamforming is almost always applied to increase signal-to-noise ratio (SNR). Because of the advantages of high bearing resolution and low sidelobes, it is preferable to use adaptive beamforming over conventional. Adaptive beamforming in the time domain (ABFTD) is unrealistic for many sonar audio applications, because its implementation has a high computational cost (Van Trees 2002). The problem is that there is the requirement to invert matrices of very large dimension: typically of the order of many thousands (the dimension equals the number of array elements multiplied by the number of coefficients of the adaptive filters (Van Trees 2002)). On the other hand, adaptive beamforming in the frequency domain (ABFFD) subdivides the large processing job into many smaller independent jobs, one for each frequency bin. Consequently, the requirement to invert a matrix of very large dimension in ABFTD, is replaced by the inversion of many matrices of smaller dimension (the dimension now equals the number of array elements) in ABFFD. This greatly reduces the computational cost and makes ABFFD implementation feasible for sonar applications. In this paper we compare the effectiveness of several beamforming algorithms for sonar audio, and discuss some issues that arise.

FORMULATION

In frequency domain beamforming, the time series output of each array element is divided into blocks. Each block of data is Fourier transformed, and the results for each array element in the same frequency bin f are combined into an array output vector $\mathbf{x}(f)$. The beamformer output at a steering angle q and at frequency bin f is then given by

$$y(q,f)=\mathbf{w}(q,f)^{H}\,\mathbf{x}(f)\,, \tag{1}$$

where $\mathbf{w}(q,f)$ is a vector of complex weights given by the beamforming algorithm, and superscript ^{H} denotes the conjugate transpose. When $y(q,f)$ has been computed at each frequency bin, the time domain signal $y(q,t)$ is obtained by inverse Fourier transformation.

One method of calculating $\mathbf{w}$ (for simplicity we henceforth drop indices q and f) is the Robust Capon Beamforming (RCB) scheme proposed by Li etc. (Li 2003). Let $\mathbf{x}_n$ denote the array output vector for the n^{th} block, and $\hat{\mathbf{R}}$ the sample estimate of the cross spectral matrix,

$$\hat{\mathbf{R}}=\frac{1}{N}\sum_{n=1}^{N}\mathbf{x}_n\mathbf{x}_n^H\,. \tag{2}$$

Let $\mathbf{v}$ denote the theoretical signal response vector obtained from the signal model of the array; this is the array output vector when a signal arrives from beamsteer in the absence of noise. The RCB algorithm anticipates possible errors in the signal model. The actual signal response vector, $\mathbf{v}_a$, is estimated by solving the following quadratic problem,

$$\min_{\mathbf{v}_a}\mathbf{v}_a^H\hat{\mathbf{R}}^{-1}\mathbf{v}_a \quad \text{subject to} \left\|\mathbf{v}_a-\mathbf{v}\right\|^2\le\varepsilon\,, \tag{3}$$

where ε is a small positive number proportional to the signal mismatch. Once $\mathbf{v}_a$ is obtained (see (Li 2003) for details), the RCB solution for $\mathbf{w}$ can be computed using the formula:

$$\mathbf{w}_{RCB}=\frac{\hat{\mathbf{R}}^{-1}\mathbf{v}_a}{\mathbf{v}_a^H\hat{\mathbf{R}}^{-1}\mathbf{v}_a}\,. \tag{4}$$

As explained in (Li 2003), the RCB approach belongs to the class of (extended) diagonally loaded minimum variance distortionless response (MVDR) beamformers. The degree of loading is determined by and proportional to ε. However, unlike the widely used MVDR beamformer with uniform diagonal loading, the loading in the RCB method is nonuniform and is adaptive to the data. It should be noted that the original purpose of the RCB method is to make the algorithm robust to errors in the signal model, and the appropriate value for ε is determined by the anticipated signal mismatch. For the audio application, however, loading may be introduced for another purpose. Some degree of diagonal loading is required in audio processing to limit the so-called white noise gain (WNG) (Cox 1987) of the beamformer. The WNG must not be too high because the weights of the beamformer $\mathbf{w}$ are obtained through averaging N blocks whereas time domain audio signals are calculated using individual blocks (i.e., no averaging). Within each block the noise may be stronger in different bearings. With a high WNG, this noise might be amplified and interfere with the target signal. The problem of determining the appropriate degree of loading for a specified WNG is quite difficult to solve.

In contrast to RCB, the MVDR beamformer with uniform loading (MVDRUL) has **w** vector

$$\mathbf{w}_{MVDRUL} = \frac{(\hat{\mathbf{R}} + \delta\mathbf{I})^{-1}\mathbf{v}}{\mathbf{v}^H(\hat{\mathbf{R}} + \delta\mathbf{I})^{-1}\mathbf{v}}, \quad (5)$$

where $\delta \geq 0$ is the loading coefficient and **I** the identity matrix. Finally, the uniformly shaded conventional beamformer (CBF) has a non-adaptive **w** vector given by

$$\mathbf{w}_{CBF} = \mathbf{v}. \quad (6)$$

SIMULATIONS

The aim of the simulations is to compare the ability of RCB, MVDRUL and CBF, to produce sonar audio in the presence of strong interference and background noise. The performance metric is based on the correlation coefficient, defined in a way that accounts for a time-delay between the signal and the time-series output from the beamformer,

$$\gamma = \max_{\tau} \frac{|\,\mathrm{COV}\{s(t)y(t-\tau)\}\,|}{\sqrt{\mathrm{COV}\{s^2(t)\}\mathrm{COV}\{y^2(t-\tau)\}}}. \quad (7)$$

Here $\mathrm{COV}\{s(t)y(t-\tau)\}$ denotes the covariance between signal $s(t)$ and the delayed beamformer output $y(t-\tau)$, etc. Note that γ has a value between 0 and 1, with value 0 indicating statistical independence between the signal and beamformer output, and value 1 indicating that the two waveforms are proportional when an appropriate time-delay is included.

Table 1. Performance Metric, γ

	RCB*	MVDRUL*	MVDRUL(0)	CBF
Scenario 1	0.959	0.935	0.822	0.113
Scenario 2	0.976	0.963	0.836	0.135
Scenario 3	0.985	0.978	0.838	0.453
Scenario 4	0.986	0.980	0.839	0.704

* With optimal loading

The sonar array considered in the simulation was a uniform linear array with 16 elements and inter-element spacing of 0.75 meters. Sound speed in the water was taken to be 1500 meters/sec. The sampling frequency was 4000 Hz, the length of the FFT (fast Fourier transform) was 256, and the number of blocks (N) used to estimate matrix $\hat{\mathbf{R}}$ was 64. Four scenarios were considered. Results are summarised in Table 1 and discussed in the sub-sections that follow. It was determined experimentally that a γ of at least 0.950 was required for the audio to be of high quality from the viewpoint of a listener. This subjective criterion was use to judge if the audio is satisfactory or not.

Scenario 1

The acoustic signal arrives at a bearing of 96° defined relative to the axis of the array, and consists of four tonal components (720 Hz, 780 Hz, 890 Hz and 1000 Hz). The interference is from the broadside direction (90°) and consists of a broadband (200 Hz to 1000 Hz) component and two tonal components (265 Hz and 310 Hz), and its power is 20 dB relative to the signal power. Independent and identically distributed (IID) noise of 0 dB relative to the signal power is also added. This represents a scenario in which the signal is completely masked by the main beam of the interference if CBF is used (the broadband beamwidth of the main lobe at the signal level is about 50°).

The results summarised in Table 1 were obtained by computing the performance metric γ by averaging over 400 independent runs of the simulation process, where the phase relations among the tonal components were randomly chosen for each run. It should be noted that for RCB and MVDRUL the metric γ is dependent on the degree of loading applied to the beamformers. A small degree of loading results in a high white noise gain (WNG), and a signal distorted by IID noise. A large degree of loading broadens the main lobe of the interference, and leads to the signal being distorted by the interference. In Table 1, the values of γ for RCB and MVDRUL represent the maximum values obtained by using optimal loading. 'MVDRUL(0)' is the case of MVDR without loading (δ=0).

Seen from the 2nd row in Table 1, the signal is highly distorted when CBF is used (γ = 0.113). This is to be expected as the signal is completely masked by the main lobe of the interference. For MVDR without loading (MVDRUL(0)), the signal is noticeably distorted (γ = 0.822), although the signal is now outside of the main lobe of the interference due to the high bearing resolution of the adaptive beamforming. In this case the distortion is due to IID noise passing through the system as a result of high WNG. In the case of MVDRUL, optimal loading gives a γ value of 0.935, which was still of unsatisfactory quality when listening to the audio. Only RCB was able to give a satisfactory audio output, with γ = 0.959.

We can define a 'loading ratio' as the ratio between the highest and the lowest values of ε (in the case of RCB) or δ (in the case of MVDRUL) that keeps the signal satisfactory (γ > 0.950). The larger the ratio, the less sensitive the performance of the beamformer is to the degree of loading. For RCB the loading ratio is 66, which indicates that the performance of RCB is not very sensitive to the value of ε. It should be noted that we did not introduce any signal mismatch in the simulation; i.e., $\mathbf{v_a}=\mathbf{v}$. Thus in (3), ε can be chosen to be almost zero. However, the aim here is to produce the best quality audio, and the value of ε cannot be chosen to be too small or the WNG will lead to excessive distortion of the signal with IID noise.

Scenario 2

All of the parameters in this scenario are the same as those in Scenario 1, except that the signal now arrives at bearing 100°. The signal is now separated from the interference by a larger angle (10°), but is still located inside its main lobe if CBF is used.

The results are summarised in the 3rd row in Table 1. Again the signal is highly distorted if CBF is applied (γ = 0.135), because the signal is completely masked by the main lobe of the interference. For MVDRUL(0), the signal is noticeably distorted (γ = 0.836) by IID noise, due to high WNG. The best value of γ achievable by MVDRUL is up to 0.963 and the audio is now satisfactory. However, the loading ratio is only 7, so the loading coefficients are required to remain in a very narrow range to keep the audio output satisfactory. This indicates that performance is quite sensitive to the loading parameter. On the other hand, the best value of γ achievable by RCB is 0.976, and the loading ratio is 1250. Thus the performance of RCB is insensitive to the degree of loading.

Scenario 3

All the parameters in Scenario 3 are the same as those in Scenario 2 except that the signal is now at bearing 130°. The signal is now 40° from the interference, and outside of its main lobe if CBF is used.

The 4th row in Table 1 summarises the results. The signal is still severely distorted when CBF is used ($\gamma = 0.453$). In this case the signal is masked by a strong sidelobe of the interference, rather than by the main lobe. For MVDRUL(0), the signal is again highly distorted ($\gamma = 0.838$) by IID noise due to high WNG. The audio output is satisfactory for MVDRUL (γ = 0.978) and RCB (γ = 0.985), and both algorithms are insensitive to the degree of loading applied. The loading ratio is 96 for MVDRUL and 6300 for RCB.

Scenario 4

All parameters in Scenario 4 are the same as in Scenario 3, except that the signal is now at bearing 180°. In this case, the signal is separated from the interference by 90°.

The results are summarised in the last row in Table 1. The values of γ for MVDRUL(0), MVDRUL and RCB are similar to those obtained for scenario 3. However, γ for CBF is substantially increased, although the audio is still unsatisfactory. Once again, only MVDRUL and RCB have satisfactory audio, and both algorithms are quite insensitive to the degree of loading applied. The loading ratio is 240 for MVDRUL and 7000 for RCB. The higher value of γ for RCB and its significantly larger loading ratio indicates superior performance.

SONAR DATA

In order to demonstrate the effectiveness of RCB in a more realistic situation, we compare the performance of RCB and CBF using experimental data collected by a passive sonar array. The data used consists of two sonar recordings superimposed on each other. One recording contains a broadband contact moving from bearing 22° to bearing 20° in 45 seconds. This contact is taken to be the signal for the sake of this study. The other recording contains a different broadband contact moving from bearing 42° to bearing 32° in the same period of time. This second contact is taken to be interference. The signal to interference ratio (SIR) varies between –6 dB and –12 dB in the data.

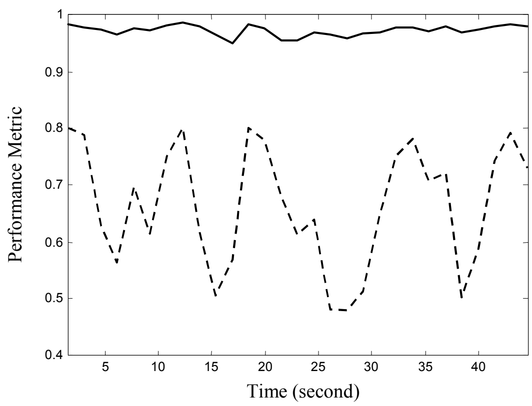

Figure 1. Performance metric, γ, for RCB (solid line) and CBF (broken line) for sonar data.

Figure 1 plots the metric γ (calculated about every 1.5 seconds) for both RCB and CBF. Note that γ for RCB lies above 0.95 over the entire time-period, whereas γ for CBF lies below about 0.8. This demonstrates the superior performance of the RCB algorithm.

CONCLUSIONS

We have examined adaptive beamforming for sonar audio. To reduce computational cost, frequency domain beamforming was considered. Simulations were carried out that demonstrate (i) in the presence of strong interferences, conventional beamforming (CBF) may be inadequate to produce a satisfactory, undistorted audio signal, even if the signal is well-separated in bearing from the interferers, and (ii) adaptive beamformers are often capable of producing satisfactory audio, and in the case of our simulations, the robust Capon beamformer (RCB) consistently outperformed other adaptive beamformers that were tested. RCB outperformed in regard to the quality of audio (higher value of γ) and robustness (larger loading ratio). Testing was also carried out on a data set obtained by superimposing actual sonar recordings of a contact and an interferer, and the good results obtained by the RCB algorithm were verified for this somewhat more realistic scenario

ACKNOWLEDGEMENTS

I thank Derek Bertilone for assistance in the preparation of this paper.

REFERENCES

Barger, J. E. (1997) *Sonar systems. in Encyclopedia of acoustics. Vol.1*, Crocker, M. J., Ed. Wiley.

Cox, H., Zeskind, R., and Owen, M. (1987) *Robust adaptive beamforming*, IEEE Trans. ASSP, Vol.35 (10), pp.1365-1376.

Li, J., Stoica, P., and Wang, Z. (2003) *On robust capon beamforming and diagonal loading,* IEEE Trans. Signal Processing, Vol.51 (7), pp.1702-1715.

Van Trees, H. L. (2002) *Optimum array processing, part IV of detection, estimation and modulation theory.* Wiley-Interscience.