Proceedings of the Oceans '05 Europe meeting in Brest, France, June 2005

Pair-wise Processing of Spectrograms for Localization of Multiple Broadband CW Sources

Eva-Marie Nosal and L. Neil Frazer
School of Ocean and Earth Sciences
Department of Geology and Geophysics
University of Hawaii at Manoa
1680 East-West Road POST 813, Honolulu, HI 96822
nosal@hawaii.edu and neil@soest.hawaii.edu

Abstract - **A pair-wise processing algorithm has been developed to localize broadband sources in shallow water. A simple sparse hydrophone array with number of elements roughly equal to the maximum number of sources is used. The sources can be continuous-wave (i.e. no onset times), and no previous knowledge of source signatures is required. The processor is spatially coherent and partially frequency coherent. Simulations show considerable improvement over conventional (i.e., frequency incoherent) matched field techniques under realistic noise conditions, with environmental mismatch and multiple sources. Spectrograms have been incorporated into the algorithm to make use of higher frequencies at greater ranges. Our work is motivated by the problem of localizing multiple singing humpback whales, and recordings of humpback whale song from the Pacific Missile Range Facility in Hawaii are being analyzed.**

I. INTRODUCTION

The goal of our research is to extend and implement passive acoustic localization algorithms for use in tracking vocalizing humpback whales on winter breeding grounds. Acoustical techniques have advantages over visual and tagging techniques since they are non-invasive and unobtrusive, they are not interrupted by poor weather conditions or lack of daylight, they enable continuous and remote sensing, and they are cost and time efficient.

Although localization methods for underwater sources have made great progress over the last 25 years (see [1-3] for overviews and references), their application to humpback whale localization is problematic because of their need for large numbers of hydrophones, e.g., vertical line arrays, assets that are seldom available to scientists studying whales. Accordingly, acoustical methods for locating whales have often relied on simple assumptions, such as constant sound-speeds and straight-line propagation, that are not satisfied[4] by the shallow water, complex bathymetry environments in which humpbacks are usually found. Model-based methods (i.e., those that use computer models of acoustic propagation) are desirable in this problem, but available model-based methods can be difficult to apply, mainly due to the characteristics of humpback vocalizations. These include:

1) Unknown waveforms: the whale's song is not known (or, technically, how far the whale is into the song is unknown).
2) Continuous waveforms: song units typically consist of up-sweeps, down-sweeps, and constant-frequency contours[5].
3) Multiple sources: singers tend to space themselves about 4-6 km apart[6], although the spacing becomes tighter with increasing density of whales.
4) Broadband, mid-frequencies: 30 Hz – 8kHz[5].

As model-based algorithms depend on the agreement of measured signals with synthetic signals, they have difficulty with high-frequency sources, i.e., sources located many wavelengths from receivers. For source-receiver offsets many wavelengths long, fluctuations and uncertainties in sound speed profile and bathymetry distort actual signals to the point where they no longer agree with signals synthesized under the assumption of a constant environment. Thus popular model-based techniques have been limited to low frequencies (well below 1 kHz) where such environmental mismatch is less harmful. Only recently have mid and high frequencies begun to be explored for use in source localization[2, 7].

Another limit to existing techniques is that most require additional assumptions about the source. For example it is often assumed that there is only one source, or that the source waveform is known, impulsive, or narrow-band. In addition, as noted above, some techniques rely on line arrays. We consider arrays with a few hydrophones separated by many source wavelengths (sparse arrays) because they are usually the only type of array available to whale researchers. They are also, in general, the simplest and least costly tools for acoustic research[8]. Of further interest is that various sparse arrays currently in operation (e.g., AUTEC, PMRF, and the Southern California Offshore Acoustic Range) can be used to gather marine mammal data.

In this work we address the problem of low spatial resolution (of arrays with relatively few hydrophones) by utilizing the frequency coherence of the source signal as well as its spatial coherence; we do this without the usual requirement that the source signal be known. We address the problem of lowered coherence at high frequencies by processing spectrograms instead of waveforms.

II. ALGORITHMS

To deal with unknown, continuous-wave sources, a pair-wise waveform (PWW) processor is used. It is an extension of the pair-wise inversion technique of Frazer and Sun[9], with application of ideas from Westwood's broadband processor[10]. To understand the PWW processor, consider the received signals at two

hydrophones, $R_1(\omega)$ and $R_2(\omega)$. Let $G_1(\omega)$ and $G_2(\omega)$ denote the channel impulse responses from the source to the first and second hydrophones, respectively. The received spectra (measured) are the products of the source spectrum, $W(\omega)$, with the impulse responses, i.e.

$$R_n(\omega) = W(\omega)G_n(\omega) \qquad n = 1,2$$

Now, let $\hat{G}_n(\bar{x},\omega)$ denote the modeled impulse response between receiver *n* and candidate source location $\bar{x}$. We introduce the following two products:

$$H_{12}(\bar{m},\omega) = R_1(\omega)\hat{G}_2(\bar{x},\omega)$$
$$H_{21}(\bar{m},\omega) = R_2(\omega)\hat{G}_1(\bar{x},\omega)$$

Denote the correct source location by $\bar{s}$. Then $\hat{G}_n(\bar{s},\omega) \approx G_n(\omega)$ (approximately since the propagation model cannot be perfect). This leads to:

$$H_{12}(\bar{s},\omega) \approx W(\omega)G_1(\omega)G_2(\omega) \approx H_{21}(\bar{s},\omega)$$

We define our PWW processor (a probabilistic indicator of source location) as:

$$\varphi(\bar{x}) = \left[\frac{2\,\mathrm{Re}\left(\sum_{\omega} \left(H_{12}(\bar{x},\omega)\right)^* H_{21}(\bar{x},\omega) \right)}{\sum_{\omega}|H_{12}(\bar{x},\omega)|^2 + \sum_{\omega}|H_{21}(\bar{x},\omega)|^2} \right]^2$$

where * denotes conjugation. The reason for this definition of $\varphi(\bar{x})$ can be understood as follows. Think of H_{12} and H_{21} as two complex column vectors. Concatenate them twice, once with H_{12} sitting on H_{21}, then vice-versa, to make two longer vectors. Then $\varphi(\bar{x})$ is just the normalized inner product of these two longer vectors. The definition above is preferable to just taking the inner product between H_{12} and H_{21} directly because it adds symmetry to the algorithm; it does not matter which receiver is named 1 and which is named 2. By the Cauchy-Schwartz inequality, the processor reaches its maximum value (unity) when $H_{12} = H_{21}$. In particular, $\varphi(\bar{x})$ is maximized at the true source location, i.e. for $\bar{x} = \bar{s}$.

To generalize the PWW processor to more than two receivers, we sum coherently over receiver pairs:

$$\varphi_{pww}(\bar{x}) = \left[\frac{2\,\mathrm{Re}\left(\sum_{\omega} \sum_{i=1}^{N-1} \sum_{j>1}^{N} \left(H_{ij}(\bar{x},\omega)\right)^* H_{ji}(\bar{x},\omega) \right)}{\sum_{\omega} \sum_{i=1}^{N-1} \sum_{j\neq 1}^{N} \left|H_{ij}(\bar{x},\omega)\right|^2} \right]^2$$

It can be shown that a frequency incoherent version of the PWW processor is equivalent to the Bartlett (conventional) matched field processor[3]. To get a Bartlett processor from the equation above, the sums over frequency are taken out of the square brackets in both the numerator and denominator; thus each frequency is processed separately and an average is taken over all frequencies.

To address the problem of incoherence at long ranges, we process spectrograms instead of waveforms. We call this the pair-wise spectrogram (PWS) processor. Spectrograms are less sensitive to mismatch and fluctuations in the ocean wave-guide, particularly at higher frequencies. Our use of spectrograms is in the spirit of envelope processing[2] in which signal envelopes are processed instead of waveforms. Compared to envelope processing, PWS processing uses more of the frequency characteristics of the signal, and retains the advantages of the pair-wise algorithm.

Denote the complex spectrogram of $H_{ij}(\bar{x},\omega)$ by $S_{ij}(\bar{x},t,f)$, where *t* and *f* are time and frequency respectively. Above the crossover frequency f_c (to be determined), only the envelope of each channel is processed, and the mean is removed from each envelope because a constant offset holds no information. The formula for the PWS processor is analogous to that of the PWW processor:

$$\varphi_{pws}(\bar{x}) = \left[\frac{2\,\mathrm{Re}\left(\sum_{t}\sum_{f} \sum_{i=1}^{N-1} \sum_{j>1}^{N} \left(S_{ij}(\bar{x},t,f)\right)^* S_{ji}(\bar{x},t,f) \right)}{\sum_{t}\sum_{f} \sum_{i=1}^{N-1} \sum_{j\neq 1}^{N} \left|S_{ij}(\bar{x},t,f)\right|^2} \right]^2$$

Once again, intuition into this processor is gained by thinking of the spectrograms as vectors and taking normalized inner products. Of course, a weighting over frequencies and/or times may be introduced in either of the processors above to emphasize or de-emphasize certain aspects of the signal. For a slowly drifting source, for example, it may be advantageous to put more weight on more recent times.

III. SIMULATIONS

The Bartlett, PWW, and PWS processors were implemented in MATLAB[11]. Bellhop[12] was used to model impulse responses. Simulations were run for a 700m by 700m by 200 m (constant depth) area. The sound speed profile used was typical of that seen in Hawaiian winter waters. It was based on historical values taken from the Generalized Digital Environmental Model[13]. Simulated whale signals, 40 s long and sampled at 2 kHz, were propagated (by convolution with modeled impulse responses) from several source locations to several receiver locations within the range. The numbers and positions of sources and receivers were varied throughout the simulations. Our noise was of the worst-case type: many noises sources with source signatures identical to that of the actual source, except for their randomized strengths and

start times. “Noise whales” (sources with source waveforms identical to those of the “real” whales) were placed at every grid point in the search area. The signals from these “noise whales” were propagated to the receivers and summed in time to give the background noise. The power of the “noise whales” was adjusted to give a specified average signal-to-noise ratio (SNR) over all receivers.

After generation of the noisy, synthetic data, the PWW, PWS and Bartlett algorithms were used to try to locate the simulated singers. The grid used in the localization was at a single depth (60 m) and grid spacing was 4 m (ultimately, searches will be conducted over several depths). Spectrograms were generated using a 256 point FFT. Signals were Hanning windowed prior to computing each spectrum, and there was a 128 point overlap between successive time windows of the spectrogram. Environmental mismatch was introduced in the form of incorrect water depth; all inversions shown are for a depth of 204 m rather than of 200 m. SNRs were gradually decreased to explore the localization error due to noise. In a typical simulation with one unknown source at (252, 304) and three receivers, all processors localized the source at 5dB SNR (Fig. 1). The Bartlett processor began to lose the source at 0 dB SNR (Fig. 2). The PWW lost the source at –5 dB SNR (Fig.3). The PWS processor could successfully localize the source even with worse than –10 dB SNR. In another simulation with two sources, the PWS processor could localize both sources in 0 dB noise, but 4 receivers were required (see Fig. 4). The other processors only found one source.

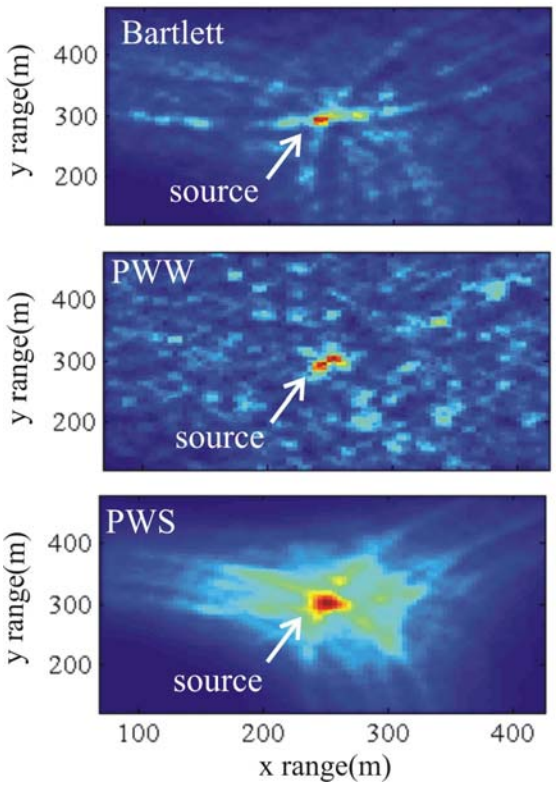


Fig. 1. SNR 5 dB. 1 source, 3 receivers. All three processors successfully localize the source.

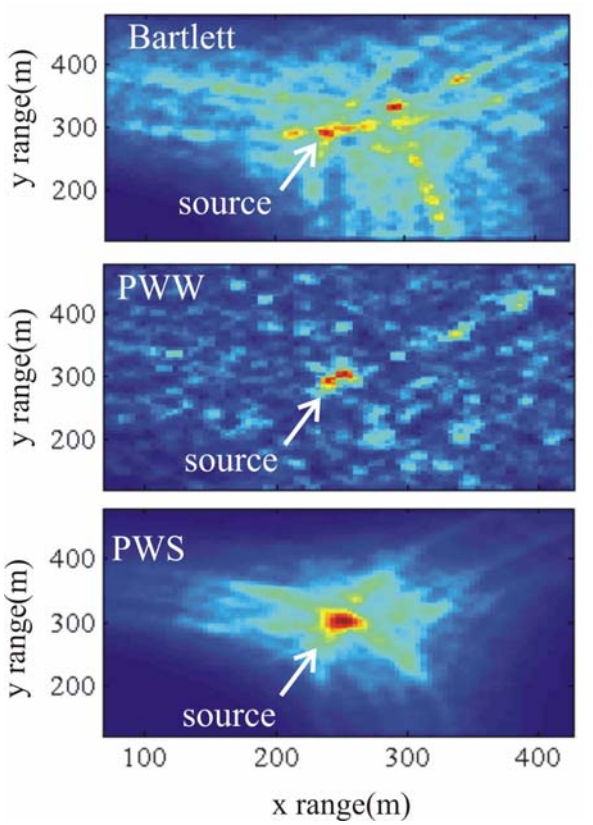


Fig. 2. SNR 0 dB. 1 source, 3 receivers. The PWW and PWS processors successfully localize the source which is lost to the Bartlett processor.

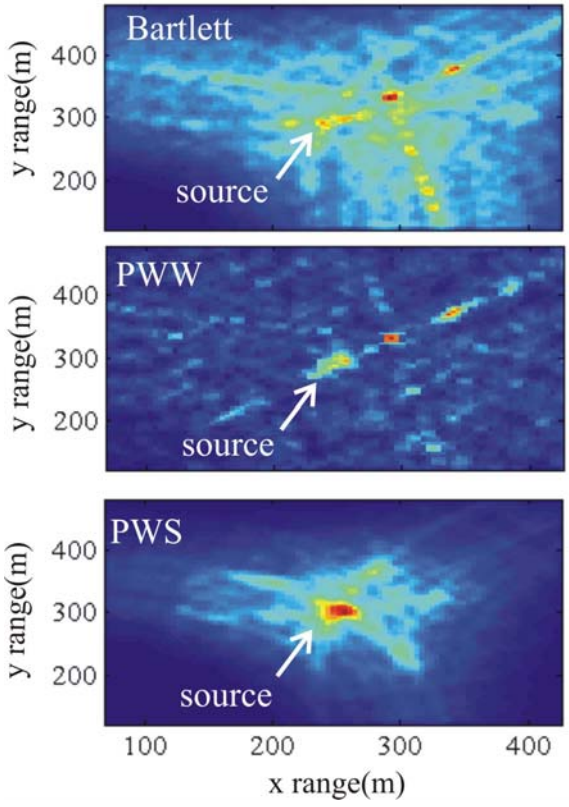


Fig. 3. SNR -5 dB. 1 source, 3 receivers. Only the PWS processor successfully localizes the source.

IV. CONCLUSIONS

A pair-wise spectrogram (PWS) processor has been proposed for the localization of multiple broad-band unknown continuous-wave sources in shallow water. It appears robust with respect to mismatch and noise; with only three receivers, a single source could be localized under conditions of both environmental mismatch and signal to noise ratios worse than –10 dB. Two identical (but unknown) sources could be localized in mismatch and 0 dB SNR. Many aspects of these processors remain to be explored though simulations and analysis of real data. The processor is being applied to humpback whale song recorded at the Pacific Missile Range Facility, but results are not yet available.

Of interest for future work is the use of cochleagrams instead of spectrograms in the PWS processor. These are auditory representations of sound (as heard by whales) whereby a cochlear filter-bank[14, 15] is applied to the signal. Being biologically more relevant than spectrogram processing, cochleagrams may prove to aid in localization – perhaps the whales know something we don’t.

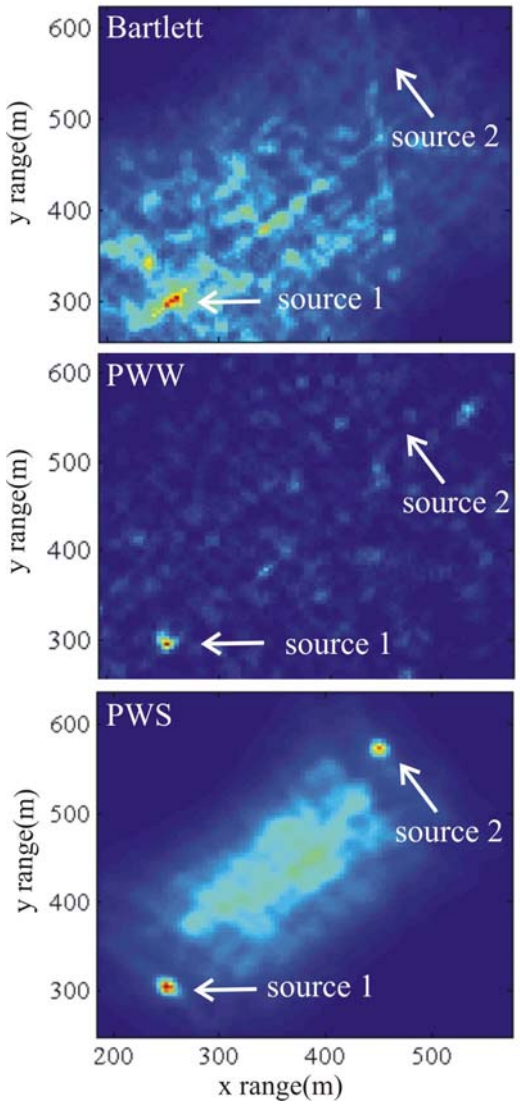


Fig. 4. SNR 0 dB. 2 sources, 4 receivers. Only the PWS processor localizes both sources (4 receivers).

Acknowledgments

This project is currently funded by the Maui High Performance Computing Center (MHPCC) and by ONR. Computing resources were also provided by MHPCC. E.-M. N. is grateful for the support of an NSERC scholarship.

References

[1] A. B. Baggeroer, W. A. Kuperman, and P. N. Mikhalevsky, "An overview of matched field methods in ocean acoustics," *IEEE J. Ocean. Eng.*, vol. 18, pp. 401-424, 1993.

[2] P. Hursky, M. B. Porter, and M. Siderius, "High-frequency (8-16 kHz) model-based source localization," *J. Acoust. Soc. Am.*, vol. 115, pp. 3021-3032, 2004.

[3] A. Tolstoy, *Matched field processing*. Singapore: World Scientific, 1993.

[4] D. M. F. Chapman, "You can't get there from here: Shallow water sound propagation and whale localization," *Canadian Acoustics*, vol. 32, pp. 167-171, 2004.

[5] L. K. Winn and H. E. Winn, *Wings in the sea: the humpback whale*. Hanover, NH: University Press of New England, 1985.

[6] A. S. Frankel, "Acoustic and visual tracking reveals distribution, song variability and social roles of humpback whales in Hawaiian waters." Honolulu: University of Hawai'i at Manoa, 1994, pp. 142.

[7] W. S. Hodgkiss, W. A. Kuperman, J. J. Murray, G. L. D'Spain, and L. P. Berger, "High frequency matched field processing," in *High frequency acoustics in shallow water*, N. G. Pace, E. Pouliquen, O. Bergem, and A. P. Lyons, Eds. La Spezia: NATO SCALANT Undersea Research Center, 1997, pp. 229-234.

[8] S. A. Hayes, D. K. Mellinger, D. A. Croll, D. P. Costa, and J. F. Borsani, "An inexpensive passive acoustic system for recording and localizing wild animal sounds," *Journal of the Acoustical Society of America*, vol. 107, pp. 3552-3555, 2000.

[9] L. N. Frazer and X. Sun, "New objective functions for waveform inversion," *Geophysics*, vol. 63, pp. 213-222, 1998.

[10] E. K. Westwood, "Broadband matched-field source localization," *Journal of the Acoustical Society of America*, vol. 91, pp. 2777-2789, 1992.

[11] "MATLAB," The MathWorks, Inc.

[12] P. Hursky, M. B. Porter, M. Siderius, and V. K. McDonald, "High-frequency (8-16 kHz) model-based source localization," *Journal of the Acoustical Society of America*, vol. 115, pp. 3021-3032, 2004.

[13] "Generalized Digital Environmental Model": https://128.160.23.42/gdemv/gdemv.html.

[14] D. A. Helweg, D. S. Houser, and P. W. B. Moore, "An integrated approach to the creation of a humpback whale hearing model," San Diego, Technical 1835, September 2000.

[15] B. C. J. Moore, *An introduction to the psychology of hearing*, 5th ed. San Diego: Academic Press, 2003.