Ambient noise imaging in warm shallow seas; second-order moment and model-based imaging algorithms

John R. Potter[a] and Mandar Chitre
Acoustic Research Laboratory, Electrical Engineering Department, National University of Singapore, 10 Kent Ridge Crescent, Singapore 119260

(Received 27 August 1997; revised 15 June 1998; accepted 1 July 1999)

Ambient noise can be used to produce images of submerged objects using the mean intensity of the backscattered energy, a technique coined ''acoustic daylight'' because of its direct analogy to vision. It is suggested that there may be substantial additional information in higher moments of the data. At high frequencies (>10 kHz), absorption suppresses long-range propagation so that a received signal is largely dependent on the local geometry, source characteristics, and the scattering properties of interceding objects. It is shown that for snapping shrimp (*Cragnon, Alpheus, and Synalpheus*) illumination (the primary sources in warm shallow water above a few kHz), significant information is embodied in the second temporal moments of intensity. There is no visual analog to this concept, which suggests a broader imaging approach which may be termed ambient noise imaging (ANI). Another ANI technique explored is the use of spatial cross correlation, which works well and also has no visual analogy. A model-based processor (Kalman filter) is also applied to track targets subject to highly variable illumination such as provided by snapping shrimp. Examples are presented using data provided by Scripps Institution of Oceanography from the initial deployment of the Acoustic Daylight Ocean Noise Imaging System (ADONIS) in San Diego. © *1999 Acoustical Society of America.* [S0001-4966(99)01711-7]

PACS numbers: 43.30.Pc, 43.20.Fn, 43.30.Wi, 43.30.Gv [DLB]

INTRODUCTION

The possibility of using ambient noise in the ocean as the sole source of illumination to form images of submerged objects emerged about a decade ago, beginning with precursory ideas about objects causing ambient noise shadows.[1] Professor Buckingham and colleagues developed this early idea into the more complete concept of ''acoustic daylight'' (AD),[2] which included the possibility of increased illumination from ''front-lighting'' in addition to silhouetting by ''backlighting.'' These ideas are based on an analogy with optical vision, where the visual clues are all spatial, the human eye response being too slow to detect temporal variability in optical illumination under normal circumstances. This analogy is overly restrictive for the ocean acoustic case, where significant variations in the acoustic illumination occur on a time scale long compared to the sensors' response time and sampling interval. Tracking these variations, including a statistical description, offers additional possibilities for imaging. The optical analogy also breaks down when considering the angular resolution that can be achieved, the wavelengths of sound in water being much larger with respect to practical acoustic apertures. The human eye has an aperture about 10^4 times the wavelength of visible light, whereas acoustic apertures are typically 10^2 wavelengths or smaller. Similarly, diffraction is expected to play a greater role as scattering objects are smaller in terms of wavelengths than for visible light.

An ''Acoustic Daylight Ocean Noise Imaging System'' (ADONIS) was constructed at Scripps' Institution of Oceanography (SIO). ADONIS used a 3-m-diam spheroidal reflecting lens to beamform incoming energy over the aperture onto 126 receivers in the focal plane, each of which thus provided an output associated with a unique ''look-direction.'' A largely analog electronics package was developed to estimate the average acoustic amplitude received in 16 frequency bands from 8 to 80 kHz. This estimation was performed for each channel by a switched-capacitor filter whose center frequency was shifted to each of the 16 frequencies in turn, followed by a ''fallow'' settling period before beginning again at the lowest frequency. The pressure signal in each frequency band was rectified and integrated over a 2-ms time window, approximately 30 times a second. Of these 30 ''frames'' each second, an average of 24 were received and recorded by the surface data acquisition and display system as validated corruption-free data frames. The amplitude at each frequency was therefore observed and recorded for a total of some 48 ms each second, representing approximately 5% of the time. Since no phase information was available, and the ambient noise sources are many and presumed incoherent with one another, the recorded acoustic amplitudes have in all cases been squared to give power estimates before further processing. It is these acoustic power values that we shall treat as the raw input data. No raw time-series data are available from ADONIS.

The theoretical possibility of imaging using the AD technique became demonstrated fact with the first deployment of ADONIS in 1994 which yielded images of arrangements of 1×1 m^2 target panels placed at 38-m range.[3] The AD images from that deployment generally display a signal-to-noise ratio of 2–4 dB. A high contrast is unlikely to be observed in AD images due to the reduced angular resolution

[a]Electronic mail: johnp@arl.nus.edu.sg

and diffraction effects which ''fill-in'' geometric shadows. In addition, optical images are often formed from viewing scattered illumination from a few very bright sources, whereas underwater ambient noise is anticipated to be more isotropic, generating milder contrasts. Finally, ambient noise is likely to be statistically complicated both in time and space, and arise from several possible source mechanisms. This poses a problem, since a variation in noise intensity could arise either from scattering from an object moving through the field of view, or by natural variations in the noise generating mechanisms. In the ADONIS deployment which provided the data for this work, the sources were snapping shrimp of unknown species (possibly *Cragnon*, *Alpheus*, or *Synalpheus*), a conclusion we will discuss later. These sources are known to inhabit areas offering some shelter on the bottom or on structures and to ''snap'' episodically,[4,5] thus providing substantial source structure in both time and space.

Rather than consider temporal variations as ''noise,'' to be suppressed by averaging, in late 1995 the Acoustic Research Laboratory (ARL) of the National University of Singapore (NUS) began to explore the possibility that the variability itself could contain information. We show that image information resides in the variance of the acoustic intensity in each pixel, or look direction. Image information can also be extracted from spatial cross correlations of the acoustic intensity data. These are both second-order moments of the acoustic intensity. Furthermore, we demonstrate that a model-based processor is able to estimate the presence of targets more reliably in sporadic and unreliable illumination conditions than a linear estimator based on raw data.

I. AMBIENT NOISE SOURCES RELEVANT TO THE EXPERIMENT

Immediately on entering the water in preparation to deploy ADONIS, the experiment divers heard the characteristic ''crackling'' of snapping shrimp. This sound could also be heard during quiet periods from within the hull of the Offshore Research Buoy (ORB), the floating platform we used to support our surface equipment. The sound of echolocating dolphins from the nearby NOSC pens were also heard intermittently by our divers. ADONIS often received a marked increase in acoustic power at high frequencies if oriented to the right of the target range used in the experiment, towards the pens. The imaging experiments were all conducted along a line free of this interference, and in any event the dolphins became disinterested as time wore on and echolocated on our site less and less.

During the experiment the sea state was mostly 0, and at most 1, with little or no wind. No precipitation fell during data collection. An observer on ORB's deck communicated by VHF to the ADONIS operators to cease data acquisition if any motorized vessels approached. Many of the datasets were taken late at night when boat traffic was at a minimum, offering more opportunity to collect boat-free data and without the possible interference of noise from crews working on vessels tied up or moored nearby.

Although no raw acoustic time series was recorded during the experiment, such data were taken at a subsequent ADONIS deployment at the same site and have generously been made available to us by SIO. A representative excerpt of these data, spanning a period of 4 s, is shown in the upper panel of Fig. 1. The characteristic broadband impulsive ''spikes'' of the shrimp snaps are clearly seen, with the maximum pressure over the 1-s period exceeding the average level by over two orders of magnitude. A spectrogram of this data, shown in the lower panel of Fig. 1, confirms that the sharp impulsive transients dominate the spectral energy distribution over the entire bandwidth of ADONIS, 8–80 kHz. The data in Fig. 1 are similar to those taken from Scripps' Pier, 14 km to the north, shown in Fig. 1 of Buckingham and Potter's ''Sea Surface Sound 94'' paper.[6] The high-frequency ambient noise at Scripps' Pier is also dominated by snapping shrimp, and the individual snaps have been triangulated by short arrays of hydrophones (Jeff Nystuen, personal communication). Indeed, one expects any shallow (less than 60 m) warm ($>11°$ C) coastal waters offering some sheltering materials to be populated by snapping shrimp.[7]

One suspects from the appearance of Fig. 1 that snapping shrimp noise, if considered as a random process, is non-Gaussian. Indeed, it fails the chi-squared test for normal distributions. It has been noted elsewhere[8] that snapping shrimp noise from a variety of locations (including the experimental site) exhibit the robust property of a log-normal distribution of energy with time. As Fig. 1 illustrates, such noise is characterized by relatively quiet periods of many milliseconds, punctuated by energetic narrow pulses of shrimp clicking. A measure of the extreme transient pulse nature of snapping shrimp noise is provided by the scintillation index SI (normalized variance of intensity) given by

$$\mathrm{SI} = \frac{\langle I^2 \rangle - \langle I \rangle^2}{\langle I \rangle^2}, \tag{1}$$

where I is the acoustic intensity (proportional to pressure squared) and $\langle\,\rangle$ indicates the ensemble average.[9] The value of the SI increases as extreme values of acoustic pressure become more common. For a zero-mean normally distributed random pressure variable, the $\mathrm{SI}=2$ irrespective of the mean power level. For the data from the ORB site in Fig. 1, the value of the SI exceeds 500.

Examining ADONIS data from pixels subsequently identified as being oriented towards ''target'' and ''nontarget'' directions reveals that both types display approximately log-normal statistics. The ORB data usually display ''frontlit'' target conditions, we suspect as a result of the geometry of the Marine Facilities Pier to which ORB was tied and which, therefore, was behind ADONIS. The Pier very likely harbored considerable colonies of snapping shrimp. The prevailing ''nontarget'' noise is also shrimp, likely scattered about other structures and on the sea bed, but of a generally lesser intensity.

II. PRINCIPLE OF HIGHER-ORDER TEMPORAL MOMENT IMAGING

We proceed to the reasoning behind the ideas of temporal higher-moment imaging, initially proposed in 1996.[10] Consider the acoustic energy received in a narrow beam corresponding to one of the image pixel ''look'' directions of an

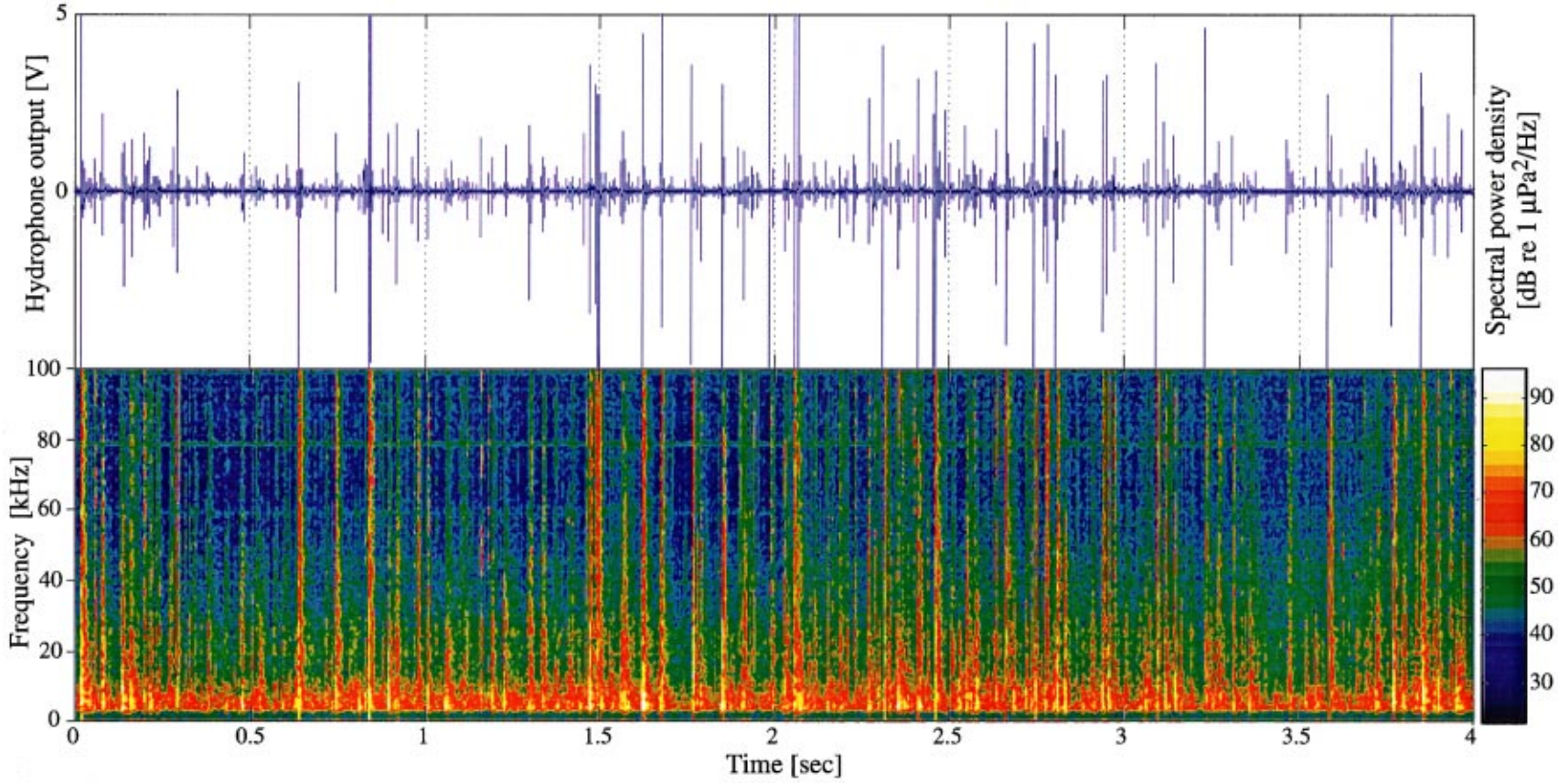

FIG. 1. Hydrophone voltage (proportional to acoustic pressure) versus time for 4 s of raw acoustic data from the experiment site (upper panel) and corresponding calibrated spectrogram evaluated by DWFFT (lower panel).

acoustic imaging system. This energy will be considered the basic data unit, although it is itself a second-order moment of acoustic pressure, taken at zero time and zero spatial lags. We consider a warm, shallow sea, where we anticipate snapping shrimp dominate the noise field at the high frequencies useful for acoustic imaging and an appropriate model for the ORB experiment. For systems that acquire data over time windows of several tens of milliseconds, as does ADONIS, we can therefore expect considerable variations in the illumination intensity from frame to frame. This variability, in particular the method of exploiting it to form images, has no visual analog.

In a time window Δt, the observed time-integrated energy received by a directional acoustic receiver will be one realization from a distribution whose characteristics are determined by the density of snapping shrimp in the beam and their distribution with range. If a target (even a perfectly reflecting one) is present in a receiving beam, this may change the received signal in a number of ways:

(i) A target will scatter energy into the beam from other directions, in addition to partially obscuring energy that would otherwise have been received. If the source field is spatially anisotropic, this may result in a difference in the mean illumination over time, the AD imaging principle. The population density of snapping shrimp is inhomogeneous (they are known to prefer certain bottom types and to colonize surfaces which offer shelter[5]). Different directions will therefore be characterized by different shrimp populations.

(ii) The solid angle of a beam's sensitivity will be different if the target scattering surface is curved. Concave targets will reduce the solid angle over which a receiver is sensitive, and convex targets (most commonly found due to hydrodynamic considerations) will increase the solid angle of a receiving beam. Contrary to one's first impression, this effect does not affect the mean received intensity. The change in solid angle is exactly balanced by the divergence of the acoustic pressure wave after scattering from the curved surface, so that the mean amplitude remains constant (conserving energy).[11] Nevertheless, the second-order temporal statistics are changed. The wider solid angle resulting from scattering from a convex surface reduces the variance, as the receiver is sensitive to a larger population of the (presumed random) sources. This provides one of the methods we refer to as ambient noise imaging (ANI).

There are thus at least two independent mechanisms by which a target is anticipated to change the received acoustic intensity statistics. The first is simply by reflecting energy from different directions. The second is subtler, leaving the mean acoustic amplitude unchanged, but affecting the variance. There is thus good reason to believe that imaging information should be contained in the higher temporal moments of the received acoustic time series, rather than just in the mean. In the (generally unfavorable) condition of isotropic noise, if the target has a curved scattering surface, then second-order moments should still yield images, even when first-order estimates fail to do so. We have not been able to test this hypothesis, since the available data is all of flat targets and anisotropic noise.

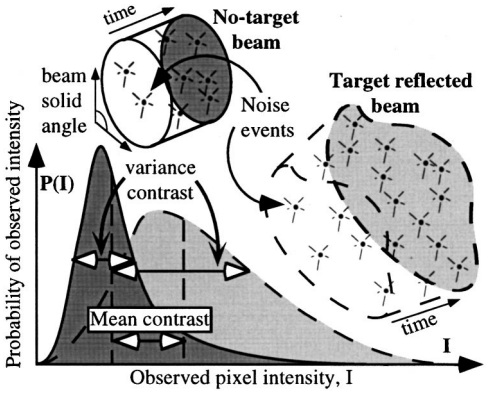

FIG. 2. Geometric basis for expecting differences in the target and nontarget intensity distributions. Distributions may differ not only in mean but also in higher moments, such as variance.

Even in the case of flat targets, such as we have for the ORB experiment, the snapping shrimp distribution requires two parameters to describe it, and the mean provides only one. The other is obtained from the second-order moment of intensity. We thus anticipate that second-order measures will be generally useful in extracting imaging information, and moreso for data involving curved targets.

A schematic illustration of this principle is given in Fig. 2. The different contributing regions and modifications in the solid angle of an imaging beam on reflection from a target give rise to probability distributions for the received energy which are different in both the mean and variance. The cylinder delineated by solid lines depicts regular beam sensitivity, unscattered by targets. The space delineated by a dashed line depicts a beam scattered by a curved target which distorts and expands the solid angle within which source may contribute, in addition to redirecting it. Even if the mean source density is similar, many more sources will be observed in the scattered beam, at correspondingly lower individual intensities. The distributions of beam intensities estimated over many such frames may then appear as at the bottom of Fig. 2, where the solid line shows a distribution curve for unscattered beams, and the dashed line indicates target-scattered beams. An image may be formed by either considering the contrast in mean intensities of the two distributions, or a second-order temporal moment, such as the difference in variance of the two distributions.

III. ORB IMAGES USING FIRST- AND SECOND-ORDER MOMENTS

A disadvantage of the higher statistical moment method is that one generally requires increasingly longer time series to obtain reliable estimates for statistical moments of increasing order (see, for example, Ref. 12). We therefore consider only the first and second temporal moments in this paper, the mean and standard deviation (s.d.) of the received intensity in each beam. This permits us to calculate robust statistics over relatively few (250) frames. The ORB deployment of ADONIS concentrated on two types of target, the

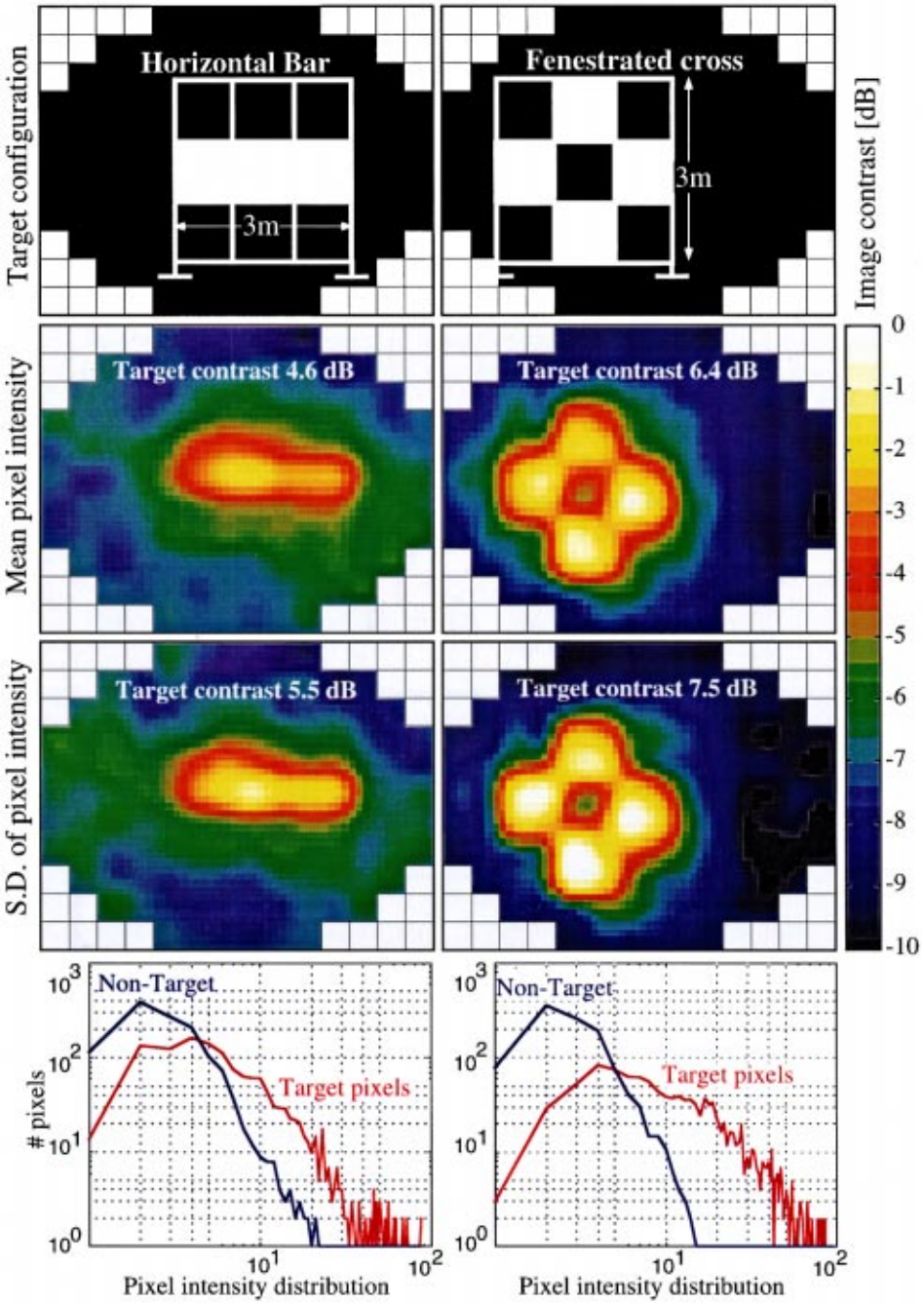

FIG. 3. ''Horizontal bar'' (left) and ''fenestrated cross'' (right) targets, shown schematically in the upper panels, with ambient noise images formed from first (mean) and second-order (s.d.) moments shown below, calculated from 250 frames. Lower panels show the observed probability distributions of target (red) and nontarget (blue) pixel energies for the two target configurations.

''horizontal bar'' and the ''fenestrated cross'' at the maximum available range of 38 m.[13] These targets were composed of flat 1×1 m^2 reflective panels mounted on a ''tic-tac-toe'' frame. For the horizontal bar, three squares were placed in a horizontal row in the middle of the frame. For the ''fenestrated cross,'' additional panels were placed above and below the middle of the bar, and the center panel removed to form a ''hole.'' The 1×1 m^2 ''hole'' subtends an angle of 1.5 degrees at the receiver, and the Rayleigh resolution of a 3-m aperture is expected to correspond to this value at approximately 20 kHz. Therefore, the lower part of the ADONIS bandwidth was expected to provide insufficient resolution to reveal the ''hole,'' and to seriously blur the other parts of the chosen targets. Diffraction of energy impinging on the targets, resulting in less backscattered energy to the receiver, is also expected to reduce imaging contrasts. For this reason only frequency estimates above 26 kHz have been used in this analysis to form images, the average frequency across the bandwidth we have used being 48.3 kHz.

We proceed to show results from the mean intensity (AD) processor in comparison to second-order methods in both the time and spatial domains.

A. Second-order temporal statistical imaging

Some first- and second-order moment images are shown in Fig. 3, where the original receiving pixels have been spatially interpolated as a final processing step using bi-cubic splines, a process which considerably improves the eye's ability to delineate important features. The top panel in

the left-hand column of Fig. 3 shows the configuration for the ''horizontal bar'' target. Below this, a mean intensity (AD) image has been displayed. The next panel shows an ANI image formed from the same 250 data frames using the standard deviation of the intensity. In the bottom panel, we show the probability distribution curves for target (red) and nontarget (blue) pixels, where six of each have been selected from the images above to form the statistics. It is clear that a substantial amount of image information is extracted from the standard deviation of intensity; not only sufficient to produce a very acceptable image, but one with a superior contrast to that obtained from the mean intensity. This is confirmed by the probability distribution curves, which show substantial differences in both the mean and variance of the two distributions.

The right-hand column shows the same panels as for the left, but using 250 frames of ''fenestrated cross'' target data. Again, the second statistical moment proves marginally superior in contrast, though both methods produce excellent images. The differences in the means and characteristic widths of the target and nontarget pixel intensity distributions are very clear, consistent with the conceptual scheme presented in Fig. 2. Despite presenting the more difficult task with regard to resolution, the ''fenestrated cross'' target is for some reason slightly better contrasted with its background than the ''horizontal bar.'' The distributions show that this occurs both because the nontarget pixel distribution is slightly narrower, and the target distribution slightly wider, than their counterparts for the ''horizontal bar'' target case. The reasons for this are unknown; it may simply arise from a small random change in the illumination and be of no particular significance.

Since the estimated mean and variance are independent parameters of the pixel intensity distributions (at least for normal or log-normal distributions), information extracted from these two parameters is independent, or orthogonal. Estimates from orthogonal processors can be combined to form images of greater contrast and statistical stability than using either processor alone.

B. Second-order spatial imaging

The temporal variance of a pixel's energy is equivalent (within an additive constant) to its autocorrelation at zero time lag. A spatial extension of this idea is to consider the spatial cross correlations of pixel intensity at zero time lags, the off-diagonal elements of the correlation matrix formed by cross correlating all 126 receiver channels.

The underlying idea is that not only the mean intensity level and variance (which are statistical estimates), but also the actual observed time history of intensity variations will be correlated for imaging beams which receive energy from neighboring regions. Since reflecting targets are expected to cause very different source regions to illuminate the receiving beams than where there are no targets, it is anticipated that the cross correlations of the 126 channels will indicate how to divide the channels into two sets. One set would be associated with target-reflected look directions, the other with nontarget.

We proceed as follows: First, the (square) normalized correlation matrix of the 126 channels is calculated. The diagonal elements are unity and the (symmetric) nondiagonal elements are the 0-lag cross correlations between channels. Second, the matrix element with the smallest cross correlation is found. The two channels associated with this value are used as the seeds for the target and nontarget sets. The extent to which the other channels correlate to these two seed channels can be used to form the image. We do this by calculating the normalized distance of each of the remaining channels from the two seed channels in the correlation space and assigning a value between 0 and 1 which represents the relative distance to each of the two seeds.

Figure 4 shows images analogous to those in Fig. 3, except taken over the entire data file length for each of the two target types (2323 frames for the bar, and 934 frames for the fenestrated cross), so that image contrast is generally better than for the 250 frames of Fig. 3. Once again, bi-cubic spatial interpolation has been applied as a final step to improve the appearance of the images. The bottom panel of Fig. 4 shows images from the spatial cross-correlation method. The familiar forms are clear, confirming that the spatial cross correlation at zero time lag also contains substantial useful information. Indeed, the image contrast is slightly better both for the spatial and temporal second-moment methods than it is for the traditional mean intensity image. Note that the spatial cross correlation requires no information about the temporal statistics, and would form an image even if there were no information in the mean or second-order temporal statistics of intensity. The spatial cross-correlation method is therefore independent of the first two methods.

IV. MODEL-BASED PROCESSING

While the images of Figs. 3 and 4 are very gratifying, it should be remembered that these have been obtained from the best sequence of 250 frames in the case of Fig. 3 (equivalent to some 10 s in real time) or from long data sets (40–98 s) which suppress noise at the expense of responsiveness. For much of the data, no recognizable image is produced at all. Clearly, this has much to do with the variation in illumination and propagation conditions. There may have been episodic events of particularly serendipitous illumination, of just the right direction and strengths to best reveal the target. In any event, we cannot expect to produce good images on demand and at all times, subject as we are to the vagaries of ambient noise illumination. Nevertheless, targets can be expected to move in smooth, continuous ways through the field of view. Just because it is no longer lit appropriately does not mean that it no longer exists. When the human brain is confronted by sporadic bursts of useful information, it attempts to track the state of the external reality, ready to update its model of the outside world when new ''useful'' information becomes available. We should perhaps consider doing the same with our ANI systems. The usual term for this is ''model-based'' processing, where as much of the physical constraints which define the possible behavior of the system are encoded into a Kalman Filter, which then estimates the

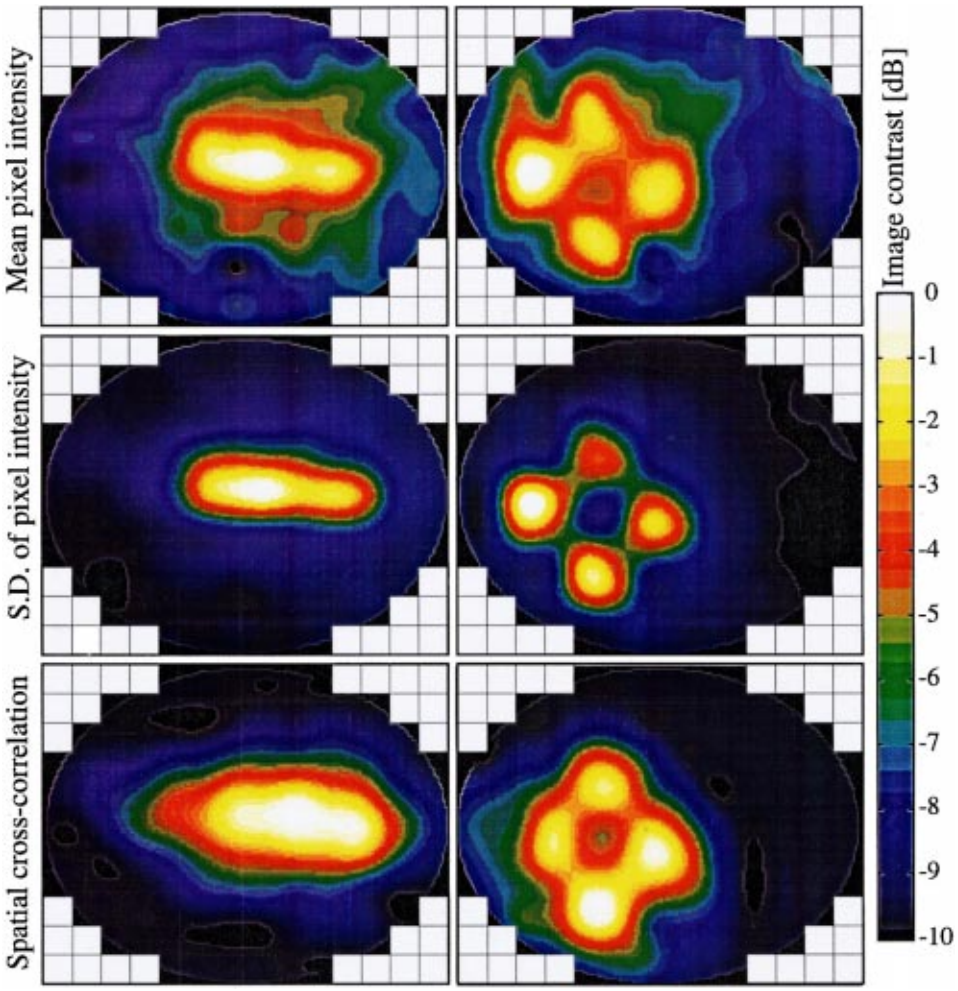

FIG. 4. Images of the horizontal bar and fenestrated cross targets as in Fig. 2, except formed over the entire length of the available datasets (2323 and 934 frames, respectively), and including spatial correlation images in the lowest panel.

current state of the system much more accurately than an uninformed one, even in the presence of substantial noise or total absence of recent useful information.

A. A modified Kalman filter

A Kalman filter is an optimal estimator (with respect to any quadratic function of estimation error) for linear systems perturbed by Gaussian white noise. That our noise is likely to be non-Gaussian should not necessarily deter us from trying this filter, unless an obviously superior processor presents itself. The Kalman filter permits one to encode the physical *a priori* information available about a system. These features make it an obvious candidate for estimating the beam intensity in ADONIS data as a preprocessor to (possibly nonlinear) imaging algorithms.

Our objective is to develop the simplest possible Kalman filter that demonstrates its usefulness by tracking the beam intensity better than simple averaging. To begin, we develop a Kalman filter to be individually applied to each pixel in the image. The state-space description for each pixel is written as

$$r(k)=H(k)x(k)+v(k), \tag{2}$$

where $r(k)$ is the observed energy at discrete time k in each beam at a given frequency. Here $r(k)$ is obtained from ADONIS data (though obviously calibrated and with all possible deterministic distortions removed) and $x(k)$ is the parameter to be estimated. It is the energy that propagates towards the sensing beam from the position of the target frame, i.e., before distortions due to propagation effect and pollution from nontarget-scattered energy entering the beam. The transform $H(k)$ embodies the propagation loss from target to receiver, which can reasonably be assumed constant. Since we are interested in contrasts between target and nontarget pixels rather than absolute values, we set $H(k)$ to unity. Here $v(k)$ is a superimposed zero-mean random component.

The evolution of the state-space over time can be modeled by a one-step auto-regressive process (Markov process):

$$x(k)=M(k-1)x(k-1)+w(k), \tag{3}$$

where M represents the evolution of the object space and $w(k)$ is a zero-mean random component. If we assume that the targets do not move, M can be set to unity. This reduces the state-space model to

$$r(k)=x(k)+v(k), \quad x(k)=x(k-1)+w(k). \tag{4}$$

We now require a physical model for how the illumination is generated and propagated into the beams. Consider the scenario in Fig. 5, which is the least complex model which still embodies some of the essential physical processes. The angular limits of a single pixel's beamwidth combined with the high degree of absorption at the frequencies employed (typically 48 kHz) limit the area in which the sources can be

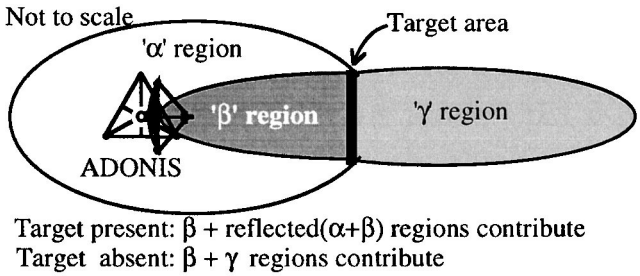

FIG. 5. Schematic of the simple geometric physical model of illumination distribution regions and propagation which contributes to received energy in a beam.

received in the absence of targets to a region which we have indicated consists of the union of areas β and γ. If a target lies in the beam, then energy from the γ region is shadowed, and that arising from some other region, termed α, may be reflected into the receiving beam. Allowing for the β region to contribute both by direct and target-reflected paths, we obtain expressions for the received energy (assumed summed incoherently),

$$r(k)=\Pi[A(k)+\alpha(k)]+(1+\Pi/3)[B(k)+\beta(k)] \quad (5)$$

if there is a reflecting target in the beam, and

$$r(k)=[\Gamma(k)+\gamma(k)]+[B(k)+\beta(k)] \quad (6)$$

if there is none, where the source contributions from the regions α, β, γ are modeled as comprising of mean expected values A, B, Γ with superimposed zero-mean Gaussian perturbations α, β, γ. The factor Π is the reflection coefficient of the target, including geometrical divergence due to curvature. The dividing factor of 3 in the reflected B-region term in Eq. (5) arises from the increased path length associated with the scattered energy from the target compared to directly received energy.

For simplicity, we assume that the target is perfectly reflecting and flat, oriented towards ADONIS, appropriate for the ORB deployment. The absolute amplitudes of the noise terms will depend on the length of the time for which the energy is integrated, the beamwidth of the receiver, the spatial geometry and extent of each region, and the attenuation at the chosen frequency. Since we seek only to estimate the noise terms in a relative sense, these effects can be incorporated into the noise parameters.

Equations (5) and (6) show that any substantial difference in the average amplitude from the reflected α and directly propagating γ regions will reveal the target over a sufficiently long averaging period. This is indeed how the first images were produced. Additionally, even if Π, Γ, A, and B conspire to make the expressions in Eqs. (5) and (6) to be nearly equal, the target might still be revealed by the differences in the variance of the random contributions α, β, and γ.

From Eqs. (5) and (6) and our assumptions,

$$r(k)=x(k)+[\mathrm{B}(k)+\beta(k)], \quad (7)$$

where

$$x(k)=x(k-1)+[[\alpha(k)-\alpha(k-1)]+4/3[\beta(k)-\beta(k-1)]] \quad (8)$$

if there is a target in the beam, and

$$x(k)=x(k-1)+[[\gamma(k)-\gamma(k-1)]+[\beta(k)-\beta(k-1)]] \quad (9)$$

if there is not. Note that the noise term in Eq. (7) is not zero-mean as required unless we estimate B *a priori* and subtract it from the data before Kalman filtering. Since the β region is not expected to contribute as strongly as the other regions due to its small physical size, we anticipate $\mathrm{B} \ll \mathrm{A},\Gamma$ and hence $r(k)$. Numerical experimentation gave results insensitive to the choice of B, so it was set to zero.

Even though two separate models are used for deriving the Kalman filter state-space equations, the state space Eq. (7) is the same for both. The second state space equation for the target and the nontarget scenario [Eqs. (8) and (9)] differ only in their zero-mean random noise term. Thus, the Kalman filter algorithm used for both target and nontarget cases is the same. The Kalman filter to estimate intensities $x(k)$ are then derived in the normal way.[14]

To compare the performance of the Kalman filter with simple averaging, Fig. 6 shows images obtained by the two methods side by side for an example period of some 18 s real time while viewing the fenestrated cross. Both the simple averaging and Kalman filter were initialized with frame 1 of the file. The simple averaging results are shown in the left panels, and the Kalman filter results on the right. The displayed frames proceed from no. 261 (shown at the top), where the image is clear to both algorithms, through no. 344, where it begins to deteriorate and on to frame no. 450, when only one of the four target panels is effectively illuminated. At this point, the simple averaging shows only that panel. The Kalman filter, however, recognizes that the other panels are still present, and that the lack of illumination does not imply lack of target. It interprets the data in the light of the physical propagation model. By frame no. 706, the incoming data is severely polluted by a strong source in the lower right of the field of view, so that the fenestrated cross image has been completely obliterated in the simple averaging image. The Kalman filter retains the target image and is only slightly affected by this noise pollution, which does not fit the state-space model and is therefore largely rejected.

The Kalman filter does not have a fixed averaging time. The effective averaging time is adaptive, becoming shorter when the data matches the model and becoming longer when the data is largely seen as noise. It is therefore not possible to estimate the performance improvement due to a Kalman filter by comparing the averaging time of the Kalman filter with that of simple averaging. However, the performance improvement can be estimated by initializing both the Kalman filter and the simple-averaging algorithm with noise. When actual data is presented to both the filters, they both discard the noisy initial estimate and pick out the target from the data at approximately the same time. This suggests that the performance advantage of a Kalman filter demonstrated

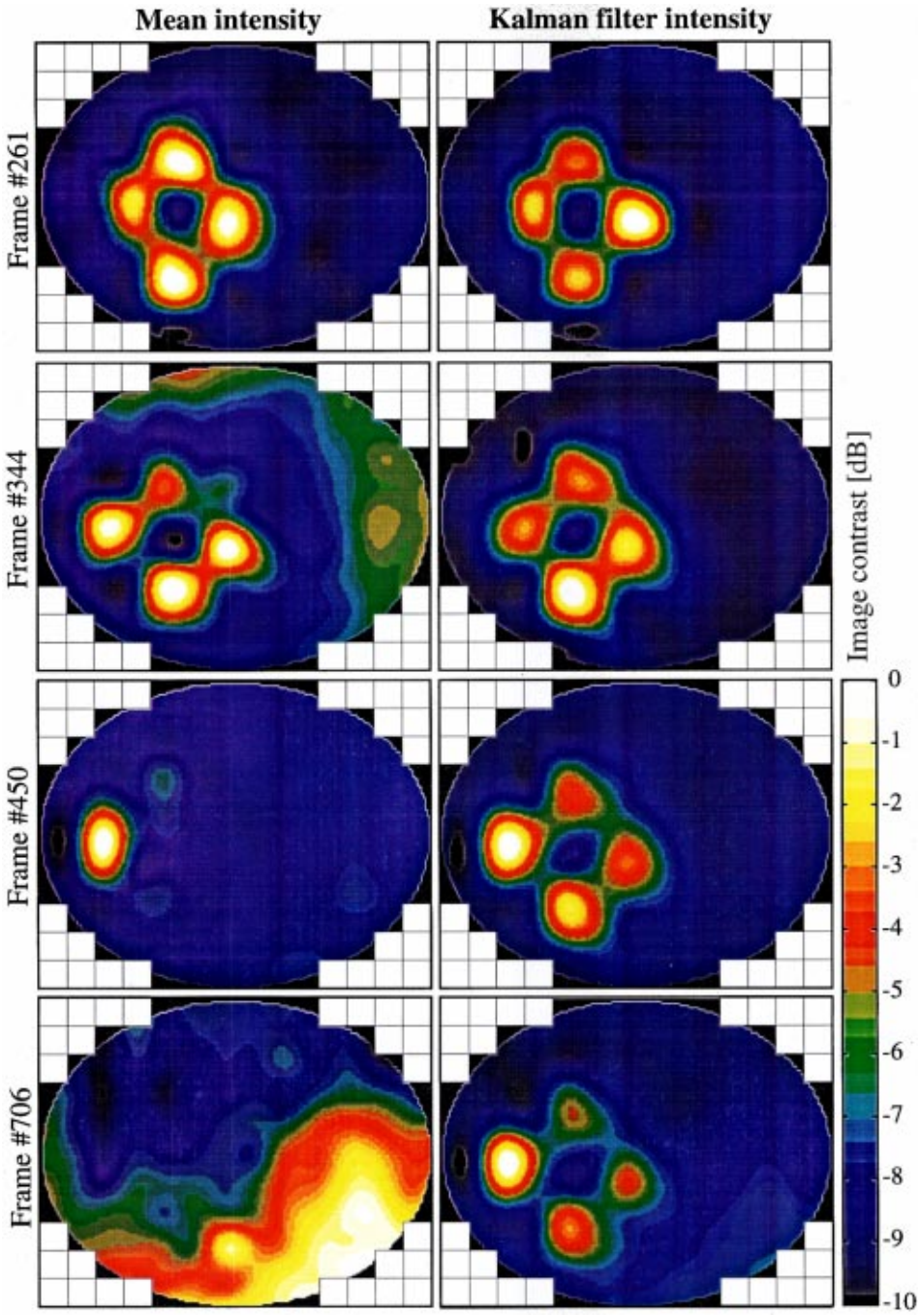

FIG. 6. A sequence of four selected frames over a period of 18 s showing the mean intensity (left) and Kalman filter estimate (right) at each frame.

in Fig. 6 is not due to a longer effective averaging time, but rather because of the adaptive nature of the Kalman filter.

The Kalman filter model developed here is a static model. It can easily be extended to provide the ability to image moving objects by a simple adaptive enhancement. When the innovation sequence of all Kalman filter channels shows that the model matches the data poorly, the estimated variance of each channel can be increased, thus allowing the filter to accept changes in the data due to movement of the target. Such a moving target scenario was tested successfully with actual data from the ORB deployment and seen to be nearly as stable as the traditional Kalman filter.

While the Kalman filter model presented here is a very basic one, it nevertheless demonstrates that the method has great potential power to improve the tracking of targets illuminated by ambient noise.

V. DISCUSSION

Although far from all the data frames are satisfactory, acceptable ambient noise images can certainly often be formed by simple averaging of received intensities, which corresponds most closely to the visual analogy of Acoustic Daylight. Additionally, we have found that we can exploit at least two of the higher statistical moments, processes that have no visual analog. The statistical complexity of ocean ambient noise invites an improved analysis of the source statistics and underlines the potential gains of using a spectral estimation technique which uses all the incoming data,

such as a discrete windowed fast Fourier transform (DWFFT), rather than the sequential stepped analog filter employed in ADONIS, which effectively discards some 94% of the data. It is not yet clear how much of the interframe variability is due to source variation, and how much is due to weak multiple scattering by inhomogeneities in the intervening medium, though if dominated by propagation effects we would not expect such clear images from the spatial cross correlation as shown in the lower panels of Fig. 4. Now that the basic ANI concept has been proven, these and related issues need to be investigated in order to design improved imaging hardware and signal processing approaches. The ANI principle is sufficiently novel (primarily due to its spatially diffuse, random, incoherent source field) that tools developed for optical and radar imaging are as likely to be useful as established marine acoustic propagation techniques.

Because high frequencies are attenuated more rapidly with range, received high-frequency sources are likely to be nearer to the target and receiver than received energy at lower frequencies. This improves the likelihood of observing a strong anisotropy in the ambient noise directionality at high frequencies, compared to noise generated by a spatially similar distribution of lower-frequency sources over a larger volume. As has been demonstrated by simulation, ANI imaging potential is strongly linked to the degree, if not so dependent on actual orientation, of the noise anisotropy.[10] There is thus good reason to believe that in warm shallow waters, which constitute some 70% of the Earth's coastal regions, snapping shrimp will provide excellent ambient noise illumination for imaging.

VI. CONCLUSIONS

We have derived two independent processors (temporal second-moment and spatial second-moment) for ambient noise imaging (ANI) which produce as good or better contrast images compared to the mean intensity processing of the original acoustic daylight (AD) principle. The target information contained in the mean intensity is controled by the spatial inhomogeneity of the sources. The target information in the temporal second moment of intensity is controled both by the source distribution and by the target shape. The temporal mean and variance of the sources themselves are independent parameters of the source distribution, and hence orthogonal. The two methods can therefore be used in concert to improve dynamic contrast and robustness of the image. The spatial cross-correlation method requires no knowledge of the temporal statistics, and is therefore orthogonal to both methods. Indeed, all three processors (and perhaps others, too) would likely be run in parallel in an ideal ANI system.

We have also developed a simple physical model of the illumination and propagation environment which forms the foundation of a Kalman filter to estimate reflected intensities in each beam. Despite the crude assumptions in this physical model, the Kalman outperforms simple averaging, significantly reducing the problems associated with episodic illumination of the targets.

ACKNOWLEDGMENTS

This work was supported by DRD, Singapore and the National University of Singapore. ADONIS data and a raw time series of ambient noise taken from the experiment site were graciously provided by Scripps Institution of Oceanography, UCSD. Special thanks to Michael Buckingham, who first set us on the path, Grant Deane and Chad Epifanio, who contributed freely and substantially to the development of ADONIS and provided the raw time series, and to the many other players who strove to build ADONIS.

[1] S. Flatté and W. Munk, ''Submarine detection: Acoustic contrast versus Acoustic glow,'' JASON Report JSR-85-108 (MITRE, McLean, VA, 1985).
[2] M. J. Buckingham, B. V. Berkhout, and S. A. L. Glegg, ''Imaging the ocean with ambient noise,'' Nature (London) **356**, 327–329 (1992).
[3] M. J. Buckingham, J. R. Potter, and C. L. Epifanio, ''Seeing in the ocean with background noise,'' Sci. Am. **274(2)**, 86–90 (1996).
[4] D. H. Cato and M. J. Bell, ''Ultrasonic ambient noise in Australian shallow waters at frequencies up to 200 kHz,'' Materials Research Laboratory, DSTO, Australia Report MRL-TR-91-23 (1992).
[5] M. L. Readhead, ''Snapping shrimp noise near Gladstone, Queensland,'' J. Acoust. Soc. Am. **101**, 1718–1722 (1997).
[6] M. J. Buckingham and J. R. Potter, ''Observations, theory and simulation of anisotropy in oceanic ambient noise fields and its relevance to Acoustic Daylight imaging,'' in *Sea Surface Sound '94*, Lake Arrowhead, CA (World Scientific, Singapore, 1994), pp. 65–73.
[7] F. A. Everest, R. W. Young, and M. W. Johnson, ''Acoustical characteristics of noise produced by snapping shrimp,'' J. Acoust. Soc. Am. **20**, (1948).
[8] J. R. Potter, T. W. Lim, and M. Chitre, ''Ambient noise in shallow warm waters and implications for imaging,'' in *Sea Surface Sound '97*, Chilworth Manor, University of Southampton, UK (1997), pp. 45–54.
[9] B. J. Uscinski, *The Elements of Wave Propagation in Random Media* (McGraw–Hill, London, 1977), p. 62.
[10] J. R. Potter and M. Chitre, ''ADONIS imaging with a Kalman filter and higher-order statistics,'' in Proceedings of the Third European Conference on Underwater Acoustics (Heraklion, Greece, 1996), pp. 349–354.
[11] J. R. Potter, ''Acoustic imaging using ambient noise: Some theory and simulation results,'' J. Acoust. Soc. Am. **95**, 21–33 (1994).
[12] J. S. Bendat and A. G. Piersol, *Random Data*, 2nd ed. (Wiley, New York, 1986), Chap. 8.
[13] C. L. Epifanio, J. R. Potter, G. B. Deane, M. Readhead, and M. J. Buckingham, ''Imaging in the ocean with ambient noise: the ORB experiments,'' J. Acoust. Soc. Am. **106**, 3211–3225 (1999).
[14] M. S. Grewal and A. P. Andrews, in *Kalman Filtering: Theory and Practice* (Prentice–Hall, Englewood Cliffs, NJ, 1993), Chap. 4, Sec. 4.2.