# Ambient noise imaging in warm shallow waters; robust statistical algorithms and range estimation

Mandar Chitre,<sup>a)</sup> Subash Kuselan, and Venugopalan Pallayil

Acoustic Research Laboratory, Tropical Marine Science Institute, 18 Kent Ridge Road, National University of Singapore, Singapore 119227

(Received 28 January 2012; revised 15 June 2012; accepted 19 June 2012)

The high frequency ambient noise in warm shallow waters is dominated by snapping shrimp. The loud snapping noises they produce are impulsive and broadband. As the noise propagates through the water, it interacts with the seabed, sea surface, and submerged objects. An array of acoustic pressure sensors can produce images of the submerged objects using this noise as the source of acoustic "illumination." This concept is called ambient noise imaging (ANI) and was demonstrated using ADONIS, an ANI camera developed at the Scripps Institution of Oceanography. To overcome some of the limitations of ADONIS, a second generation ANI camera (ROMANIS) was developed at the National University of Singapore. The acoustic time series recordings made by ROMANIS during field experiments in Singapore show that the ambient noise is well modeled by a symmetric  $\alpha$ -stable (S $\alpha$ S) distribution. As high-order moments of S $\alpha$ S distributions generally do not converge, ANI algorithms based on low-order moments and fractiles are developed and demonstrated. By localizing nearby snaps and identifying the echoes from an object, the range to the object can be passively estimated. This technique is also demonstrated using the data collected with ROMANIS. © 2012 Acoustical Society of America. [http://dx.doi.org/10.1121/1.4733553]

PACS number(s): 43.60.Cg, 43.60.Lq, 43.30.Wi [NPC]

Pages: 838-847

# I. INTRODUCTION

The possibility of using ambient noise in the ocean as the source of acoustic "illumination" to form images of submerged objects was first explored by Flatté and Munk,<sup>1</sup> and subsequently developed into a concept of acoustic daylight by Buckingham and colleagues.<sup>2</sup> The acoustic daylight technique was first demonstrated experimentally in 1994 using ADONIS, an acoustic "camera" developed at the Scripps Institution of Oceanography.<sup>3</sup> Although the acoustic daylight concept has an analog in optical vision and therefore easy to understand, it relies on averaging out the statistical variations in the acoustic noise field. Rather than average away the variation, Potter and Chitre explored the possibility that these variations contain useful information that can be used for imaging.<sup>4</sup> They showed that tracking these variations allowed images to be produced in some cases when acoustic daylight failed to produce discernible images. The family of imaging algorithms, now known as ambient noise imaging (ANI), use ambient noise to produce images of submerged objects; these include acoustic daylight and other statistical techniques based on spatial and temporal noise statistics.

Although select data from ADONIS was used to successfully produce images of submerged objects at up to 40 m range,<sup>3,5</sup> much of the data yielded no recognizable images.<sup>6</sup> This was believed to be due to the statistical variation in ambient noise producing favorable acoustic illumination at times, and unfavorable illumination at other times. The primary source of acoustic illumination in the ADONIS experiments was believed to be snapping shrimp which produce

a)Author to whom correspondence should be addressed. Electronic mail: mandar@nus.edu.sg loud "snaps" episodically. Potter and Chitre developed a model-based tracking algorithm which reduce the problems associated with episodic illumination.<sup>6</sup> However, the tracking algorithm works on the output of the ANI algorithms and hence produces better images when the underlying ANI algorithm is more robust to the illumination conditions.

Snapping shrimp dominate the high frequency ambient noise in warm shallow waters.<sup>7,8</sup> The acoustic pressure variation as a result of a large number of snapping shrimp can be modeled using a symmetric  $\alpha$ -stable (S $\alpha$ S) distribution.<sup>9</sup> Except for the Gaussian distribution (which belongs to the family of S $\alpha$ S distributions), all other S $\alpha$ S distributions are impulsive. The second and higher order statistics of these impulsive S $\alpha$ S distributions are theoretically infinite and therefore their estimates do not converge.<sup>10</sup> This may have contributed to the high variability in the images produced by high-order statistical ANI algorithms.<sup>6</sup> We explore the use of fractional low-order statistics and fractile measures of the distribution for ANI and show that algorithms based on these measures demonstrate robust imaging performance.

A limitation of ADONIS is that it records only energy estimates but no phase information for each of its 126 receive beams, thus limiting the signal processing to incoherent techniques.<sup>6</sup> ADONIS effectively records data representing about 48 ms of incoming acoustic energy every second; roughly 95% of the incoming acoustic signal is therefore unavailable for processing. To overcome these limitations, a second generation ANI camera known as ROMANIS was developed at the Acoustic Research Laboratory (ARL) of the National University of Singapore (NUS).<sup>11</sup> This camera consists of 508 pressure sensors forming a two-dimensional planar array approximately 1.3 m diameter. Each sensor is sampled at a rate of 196 kSa/s and the resulting time series recorded. Digital beamforming on the recorded data allows 288 beams to be formed over a field of view of about  $18^{\circ} \times 9^{\circ}$  and a frequency range of 25–85 kHz.<sup>12</sup> Preliminary tests of ROMANIS in 2003 yielded ambient noise images of a submerged object at a range of 70 m using previously established ANI techniques.<sup>11</sup> A more recent ROMANIS field experiment in 2010 produced a rich ANI dataset that we use to test our new ANI algorithms.

When a snap arrives at the ANI camera, we can estimate the direction towards the snapping shrimp that produced it. Assuming that snapping shrimp mostly live near the seabed and on underwater structures,<sup>13,14</sup> we can estimate the location of the source shrimp. The shrimp can be then treated like a deterministic source in a bi-static sonar system with the ANI camera as the receiver. This not only allows an underwater object to be detected or imaged, but also its range to be estimated passively. Although Potter explored this idea theoretically,15 many theoretical and practical difficulties remain unaddressed. We have previously shown that carefully selected snaps can indeed be used for passive ranging,<sup>16</sup> but the selection process used was not automated. We outline a practical automated algorithm to use snapping shrimp as sources of opportunity for determining range to a previously imaged object, and demonstrate it using the ROMANIS 2010 dataset.

The rest of this paper is organized as follows. In Sec. II, we introduce the statistical properties of snapping shrimp noise and use them to derive two families of algorithms for ANI. These algorithms rely on fractional low-order moments and fractile measures, respectively. We demonstrate the efficacy of these algorithms using the ROMANIS 2010 dataset in Sec. III. We systematically develop and demonstrate the idea of deterministic processing for passive range estimation using snapping shrimp noise in Sec. IV. We end with some concluding remarks in Sec. V.

## **II. STATISTICAL IMAGING WITH IMPULSIVE NOISE**

# A. Statistical properties of snapping shrimp noise

Snapping shrimp dominate the high frequency ambient noise in warm shallow waters.<sup>7,8</sup> The snaps produced by these animals are compact in time and broadband. With a large number of snapping shrimp producing these noises, the resulting acoustic pressure variation received at a hydrophone can be modeled using the family of  $S\alpha S$  distributions.<sup>9</sup> This family of distributions arises out of the generalized central limit theorem which states that the sum of a number of independent and identically distributed random variables with finite or infinite variance will tend to a stable distribution as the number of variables grows.<sup>17</sup> Stable distributions are characterized by four parameters-characteristic exponent  $\alpha$ , location parameter  $\mu$ , scale parameter c, and skewness parameter  $\beta$ . Characteristic exponent  $\alpha$  controls the heaviness of the tails. Location parameter  $\mu$  controls the location of the peak of the distribution. Since we are dealing with dynamic (high-pass filtered) acoustic pressure signals, the distributions of interest are centered around zero ( $\mu = 0$ ). Scale parameter c controls the width or spread of the distribution. It is also common to use the dispersion  $\gamma$  to describe the spread of the distribution; the scale parameter *c* and dispersion  $\gamma$  are directly related through the expression  $c = \gamma^{1/\alpha}$ . Skewness parameter  $\beta$  controls the skewness of the distribution. The special case of  $\beta = 0$  yields the family of S $\alpha$ S distributions. Further, when  $\alpha = 2$ , the distribution reduces to a Gaussian distribution, while  $0 < \alpha < 2$  yield heavier tailed distributions. Except for the special cases of the Gaussian ( $\alpha = 2$ ) and the Cauchy ( $\alpha = 1$ ) distributions, the family of S $\alpha$ S distributions does not have a known closed-form probability density function (pdf). Instead, the distributions are expressed in terms of their characteristic function

$$\phi_{\alpha}(\omega) = e^{-|c\omega|^{\alpha}}.$$
 (1)

The S $\alpha$ S pdf  $f_{\alpha}(x)$  is given in terms of the characteristic function

$$f_{\alpha}(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \phi_{\alpha}(\omega) e^{i\omega x} \mathrm{d}\omega.$$
 (2)

A sample of 25–70 kHz bandpass filtered data collected from the hydrophones in ROMANIS during the 2010 experiment in Singapore is shown in Fig. 1. This and other such samples pass  $\chi^2$  goodness-of-fit tests for S $\alpha$ S distribution with typical  $\alpha \approx 1.5$  at a 5% level of significance. During beamforming, a conventional wideband beamformer delays and sums data from all 508 hydrophones in ROMANIS. Samples of time series data at the output of the beamformer also pass  $\chi^2$ goodness-of-fit tests for S $\alpha$ S distribution with  $\alpha$  between 1.5 and 2.0 depending on the beam direction. Beams pointed in directions where large number of shrimp can be heard are likely to be more impulsive (lower  $\alpha$ ) than beams pointed other directions.

Let  $x_{b,t}$  be the band-limited acoustic pressure arriving at the ANI camera in a given receive beam *b* sampled at time index *t*. Since acoustic intensity is proportional to the square of the acoustic pressure, the energy in beam *b* arriving within one sampling period is  $Cx_{b,t}^2$  for some constant *C*. The average energy  $y_{b,k}$  in frame *k* is given by



FIG. 1. One-second sample of 25–70 kHz bandpass filtered data collected from ROMANIS hydrophone No. 1 during the 2010 experiment in Singapore showing the impulsive nature of the noise.

$$y_{b,k} = \frac{C}{N} \sum_{t=kN}^{Nk+N-1} x_{b,t}^2,$$
(3)

where N is the length of the frame in samples.

The average energy for beams pointing in different spatial directions can be used to produce the pixels in an ambient noise image.<sup>3,6</sup> Given a stationary ergodic noise process and a static environment, one would expect the pixel corresponding to each beam to converge to its expected value by time averaging the pixel value. Since the pixel value is given by the average beam energy, the expected pixel value is

$$\mathbb{E}[y_{b,k}] = \frac{C}{N} \sum_{t=kN}^{Nk+N-1} \mathbb{E}[x_{b,t}^2].$$
 (4)

Since  $x_{b,t}$  is modeled as an S $\alpha$ S random variable,  $\mathbb{E}[x_{b,t}^2]$  is generally infinite (except for the special case of  $\alpha = 2$ ).<sup>10</sup> Hence  $\mathbb{E}[y_{b,k}]$  is also generally infinite. Although each pixel takes a finite value as the sampled  $x_{b,t}$  are finite, the infinite expected value implies that pixel values simply do not converge irrespective of the averaging time. Any higher order statistical measure of the beam energy also suffers from the same problem.

Although the sample mean and standard deviation of the beam energy were used to produce good images from selected ADONIS data segments, for much of the data no recognizable image was produced at all.<sup>6</sup> We suspect that this inconsistent performance may be partially attributed to this lack of theoretical convergence of the statistical measures used. To get around this problem, we next propose two statistical measures that exhibit good convergence properties in  $S\alpha S$  noise.

## B. Fractional low-order moment imaging

For any S $\alpha$ S random variable X with  $\alpha < 2$ ,  $\mathbb{E}[|X|^p]$  is finite for  $p < \alpha$  and infinite otherwise.<sup>10</sup> The statistical measure  $\mathbb{E}[|X|^p]$  is known as a fractional low-order moment (FLOM) of order p < 2. It is also known that all the low-order moments of an S $\alpha$ S random variable are equivalent, i.e., any two of the lower order moments differ by a fixed constant which is independent of the random variable itself.<sup>10</sup> As variance  $\mathbb{E}[|X|^2]$  is a measure of spread for Gaussian random variables, FLOM can be used a measure of spread for S $\alpha$ S random variables. Since FLOMs of order  $p < \alpha$  are always finite, we define the FLOM pixel value for a beam as

$$y_{b,k} = \left(\sum_{t=kN}^{Nk+N-1} |x_{b,t}|^p\right)^{2/p}$$
(5)

for some  $p < \alpha$ . Equation (5) is motivated by 3 but replaces  $x_{b,t}^2$  by  $|x_{b,t}|^p$  to keep the expected value of the pixel finite. The multiplicative constant C/N is dropped without any effect, since the final image is scaled to fit the dynamic range of the display pixels. The power 2/p is introduced to keep

the units consistent with other statistical measures and thereby allowing direct comparison of the resulting images.

## C. Fractile imaging

In FLOM imaging, we used FLOM to measure the spread of the S $\alpha$ S random variable. However, a more natural measure of the spread is the scale parameter *c*. We therefore consider the use of  $c^2$  as the pixel value rather than the sample variance in 3. To do this, we need a good estimator for the scale parameter *c*. Although FLOM based parameter estimation methods exist,<sup>18</sup> the information captured using them would essentially be the same as the FLOM imaging method outlined earlier. Fama and Roll developed a fractile based estimation technique for the scale parameter.<sup>19</sup> Using this estimator,

$$c = \frac{1}{1.654} \left( \mathbb{F}_{0.72}[X] - \mathbb{F}_{0.28}[X] \right), \tag{6}$$

where  $\mathbb{F}_{f}[X]$  refers to the *f*th fractile of the random variable *X*. We therefore have

$$c^{2} = \frac{1}{1.654^{2}} (\mathbb{F}_{0.72}[X]^{2} + \mathbb{F}_{0.28}[X]^{2} - 2\mathbb{F}_{0.72}[X]\mathbb{F}_{0.28}[X]).$$
(7)

Since the S $\alpha$ S distribution of interest is symmetric around zero,  $\mathbb{F}_{0.28}[X] = -\mathbb{F}_{0.72}[X]$ . Moreover, the 0.44 fractile of  $X^2$ ,  $\mathbb{F}_{0.44}[X^2] = \mathbb{F}_{0.72}[X]$  since squaring folds the negative half of the distribution over the positive half. Therefore,

$$c^{2} = \frac{4}{1.654^{2}} \mathbb{F}_{0.44}[X^{2}].$$
(8)

Replacing the generic random variable  $X^2$  with the beam energy  $x_{b,t}^2$  and dropping the multiplicative constant, we get a fractile based pixel value

$$y_{b,k} = \mathbb{F}_{0.44}[x_{b,Nk}^2 \cdots x_{b,Nk+N-1}^2],$$
(9)

where the operator  $\mathbb{F}_{f}[\cdots]$  operates on a set of samples  $\{x_{b,Nk}^2 \cdots x_{b,Nk+N-1}^2\}$  to produce the *f*th sample fractile. We generalize this statistical measure further by allowing any fractile *f* to be used:

$$y_{b,k} = \mathbb{F}_f[x_{b,Nk}^2 \cdots x_{b,Nk+N-1}^2].$$
 (10)

To implement this, the *N* beam energy values  $\{x_{b,Nk}^2 \cdots x_{b,Nk+N-1}^2\}$  are sorted in ascending order and the  $(\lfloor fN \rfloor + 1)$ th value is chosen as  $y_{b,k}$ .

## **D.** Post-processing

At low frequencies, the beam width of the ANI camera may be larger than the pixel spacing. This causes the energy from one pixel to spill into the adjacent pixels, effectively blurring the image. To counteract this effect, at low frequencies, we optionally apply a classical image processing technique called *unsharp mask* using MATLAB's fspecial and imfilter functions,<sup>20</sup> to sharpen the image. The technique applies a two-dimensional finite impulse response filter with a  $3 \times 3$  kernel given by

$$\frac{1}{\zeta+1} \begin{bmatrix} -\zeta & \zeta-1 & -\zeta \\ \zeta-1 & \zeta+5 & \zeta-1 \\ -\zeta & \zeta-1 & -\zeta \end{bmatrix},$$
 (11)

where  $\zeta = 0.2$ . To allow various degrees of sharpening, we define a sharpening factor  $\lambda$  between 0 and 1, to combine the original image with the sharpened image. If *A* is the original image and **B** is the sharpened image, the output image is given by  $\lambda \mathbf{B} + (1 - \lambda)\mathbf{A}$ .

Typical images from ANI cameras have poor spatial resolution and only a few hundred pixels. To present them to the viewer, they have to be interpolated to a larger size. In the results presented here, we use MATLAB's interp2 function<sup>21</sup> for cubic interpolation to increase the image size by a factor of 4.

Finally we display the resulting interpolated and optionally sharpened image using a suitable high contrast pseudocolor map. The image is scaled such that the lowest pixel value in the image is mapped to 0 in the pseudo-color map, and the highest value is mapped to 1.

## **III. ROMANIS 2010 EXPERIMENT**

### A. Ambient noise imaging camera (ROMANIS)

The ROMANIS sensor array (see Fig. 2) consists of 508 acoustic pressure sensors forming a two-dimensional planar array approximately 1.3 m diameter.<sup>11</sup> Each sensor is 50 mm  $\times$  50 mm in size and has good receive sensitivity over a



FIG. 2. (Color online) The ROMANIS ambient noise imaging camera being deployed at sea in Singapore waters during the 2010 field experiment. The black circular surface is the sensing surface which is populated with 508 acoustic pressure sensors. Each sensor is sampled at 196 kSa/s and the data is channeled to the surface via fiber optic cables visible in the photograph on the top-left and top-right of the array.

25-85 kHz frequency band. The sensors are simultaneously sampled at 196 kSa/s and the data is recorded as a pressure time series for each sensor. The sensors are arranged compactly with each row of sensors offset from the previous row for optimum beamforming performance,<sup>12</sup> as shown in Fig. 3. Since the entire array surface is populated with sensors, when beamforming in the broadside direction, the array performance is similar to a piston hydrophone of 1.3 m diameter. However, as the beamformer steers the beam away from broadside, the performance drops. The beampattern of each  $50\,\text{mm}$  imes  $50\,\text{mm}$  sensor remains fixed and cannot be steered. The array beampattern is steered but as the array is sparse, the beampattern has grating lobes. These grating lobes are perfectly cancelled by the nulls in the sensor beampattern when the beam is pointed broadside, but move out of the nulls as the beam is steered.<sup>12</sup> The desired main-to-grating lobe ratio limits the achievable field of view (FOV) to approximately  $18^{\circ} \times 9^{\circ}$ . The beamwidth at 85 kHz is about  $0.8^{\circ}$  and therefore we digitally beamform to produce a 24  $\times$ 12 grid of 288 beams with 0.8° angular spacing between steering angles in the azimuth and elevation directions.

#### B. Experimental setup

The ROMANIS 2010 experiment was conducted in Singapore waters at the Selat Pauh anchorage (1° 12.9670'N, 103° 44.3814' E) in an area where the average water depth is about 15 m. ROMANIS was positioned on the seabed looking southwards and a target frame was deployed approximately at a range of 65 m south of ROMANIS. 1 m  $\times$  1 m closed-cell Neoprene foam reflector panels with Aluminum backing were attached to the target frame to form an inverted L-shape as seen in Fig. 4. The seabed in the area is mostly sandy with some muddy areas. The closest island and associated shallow areas were to the south of ROMANIS, directly behind the target frame, at a distance of about 400 m.



FIG. 3. The 508 acoustic pressure sensors on ROMANIS form a twodimensional planar array of approximately 1.3 m diameter. Each sensor is  $50 \text{ mm} \times 50 \text{ mm}$  in size. The sensors are arranged compactly with each row of sensors offset from the previous row for optimum beamforming performance.



FIG. 4. (Color online) A target frame with  $1 \text{ m} \times 1 \text{ m}$  closed-cell Neoprene foam aluminum-backed reflector panels in an inverted L-shaped arrangement being deployed in Singapore waters during the 2010 field experiment.

A cluster of islands to the north was about 1 km away with a shipping channel between ROMANIS and the islands. A long-term mooring buoy was in the FOV of ROMANIS, slightly east of the target frame at a range of about 135 m. The mooring buoy and the marker buoys marking the target frame location can be seen in Fig. 5. The buoy and retaining walls of the islands may provide a conducive habitat for snapping shrimp to colonize.<sup>13,14</sup> Snapping shrimp are also likely to be found scattered on the seabed around the area, although perhaps at a lower density.

The orientation of ROMANIS was adjusted to ensure that the target frame was in the FOV with the help of a 38 kHz pinger attached to the target frame. Once the orientation was confirmed, the pinger was removed and several ambient noise illuminated datasets were collected. We primarily use one such representative 30-s dataset collected on April 11, 2010 at 21:04:55 local time for the results and analysis presented here; similar results were obtained from other datasets collected during the experiment.

## C. Imaging results

Figure 6 shows an ambient noise image produced from the first second of the dataset using the fractile imaging



FIG. 5. (Color online) A photograph taken from the barge supporting ROMA-NIS, looking south towards the target frame. A mooring buoy is visible to the left of the target frame location marked by the two orange marker buoys.



FIG. 6. An ambient noise image of the target in the 25–50 kHz band from a 1 s data segment using fractile imaging with f = 0.44. The image is  $4 \times$  cubic interpolated and sharpened with  $\lambda = 0.25$ . The gridlines demarcate each pixel in the image before interpolation. The estimated position and size of the target based on active insonification is marked in black for reference.

algorithm. The target is clearly visible in the right half of the image. An estimated location and size of the target based on active insonification is superimposed to verify that the object visible in the image is indeed the target we deployed. The image shows higher noise levels near  $0^{\circ}$  elevation where the beams receive noise from distant sources in relatively larger spatial regions.

Figure 7 shows the ambient noise images produced by various algorithms using 200 ms datasets at three arbitrarily chosen times in the dataset. This allows us to observe the robustness of each algorithm to the statistical variations in the illumination due to the episodic snaps from the shrimp. The images in this figure are cropped to only show the right half of the image where the target is expected to appear for easier presentation. The top row uses the acoustic daylight algorithm where the mean energy in a beam is used to determine the pixel value. No discernible image is formed in the three chosen segments using this algorithm. Although we do not show the results from higher order statistical imaging in this figure, the results were qualitatively similar to the acoustic daylight images. The FLOM imaging algorithm produces some images of the target, but they show significant variability. The FLOM imaging with p = 0.5 performs slightly better than p = 1.0. The fractile imaging algorithm produces the best images and has lesser variability in performance. The fractile imaging with f = 0.44 seems to produce images that are slightly qualitatively better than f = 0.10. Similar results were obtained from other datasets recorded by ROMANIS during the experiment.

We only present results in the lower part of the frequency band (25–50 kHz) of ROMANIS as the robustness of images at higher frequencies was found to be significantly lower. This is perhaps because the higher frequencies are rapidly absorbed with range and therefore dominated by local snaps. Since the number of snapping shrimp in nearby areas is small, one would expect higher statistical variability in the resulting pixel estimates. Due to less absorption at lower frequencies, higher number of snapping shrimp from farther ranges contribute to the lower frequency noise and provide a more stable illumination for ANI.





Fractile(0.44) Time: 0.0-0.2s













Fractile(0.44) Time: 10.0-10.2s



# IV. SOURCE LOCALIZATION AND DETERMINISTIC PROCESSING

When a snap arrives at the ANI camera, we can estimate the direction towards the snapping shrimp that produced it, using the time difference of arrival across the sensors. Snapping shrimp mostly live near the seabed and on underwater

Mean Energy Time: 25.0-25.2s



FLOM(0.5) Time: 25.0-25.2s



FLOM(1.0) Time: 25.0-25.2s



Fractile(0.10) Time: 25.0-25.2s



Fractile(0.44) Time: 25.0-25.2s



structures.<sup>13,14</sup> Combining this knowledge with the local bathymetry, we can estimate the location of the source shrimp. The shrimp can be then treated like a deterministic source in a bi-static sonar system with the ANI camera as the receiver. This not only allows an underwater object to be detected or imaged, but also its range to be estimated passively. A key challenge in the practical implementation of this approach lies in data association. With multiple snaps arriving at the ANI camera from different directions simultaneously, how does one identify which snap detected on one sensor corresponds with a snap detected on another sensor? Since the snaps have very similar waveforms, how does one identify the echo of a specific snap from an underwater object? Potter explored this idea theoretically and suggested imposing some constraints to reduce the challenge of echo association,<sup>15</sup> but many theoretical and practical difficulties remain unsolved. We have previously shown that carefully selected snaps can indeed be used for passive ranging,<sup>16</sup> but the selection process used was manual and not readily automated.

In this section, we outline a practical automated algorithm to use snapping shrimp as sources of opportunity for determining range to a previously imaged object. An implementation of the algorithm is then successfully demonstrated using experimental data acquired by ROMANIS. The key steps in the algorithm are as follows.

- Detect strong snaps at each sensor. Associate corresponding snaps on all sensors in the ANI camera and determine time of arrival at each sensor.
- (2) Estimate the direction of arrival of the snap, and the location of the source.
- (3) Identify candidate echoes from the underwater object, and compute the range corresponding to each echo. Assuming that the object is static, combine the information from multiple snaps over a period of time to reject false echoes and identify true echoes. Fuse the range estimates from the true echoes to produce a final range estimate.

Each step is described in detail in the following subsections. Throughout our discussion below, we use a righthanded Cartesian coordinate system with the origin at the center of the ANI camera, the *x*-axis pointing along the broadside axis of the camera and the *z*-axis pointing upwards as shown in Fig. 10. We also use a spherical coordinate system with the origin coinciding with the origin of the Cartesian coordinate system, the azimuth axis coinciding with the *x*-axis and the zenith coinciding with the *z*-axis.

#### A. Snap detection and association

Let  $u_i(t)$  be the acoustic pressure signal at time t from sensor i of the ANI camera. The envelope of this signal is computed as

$$\tilde{u}_i = \sqrt{u_i^2 + \mathbb{H}[u_i]^2},\tag{12}$$

where  $\mathbb{H}[\cdots]$  is the discrete Hilbert transform operator. A threshold  $\mathbb{F}_{0.9995}[\tilde{u}_i]$  is determined as the 0.9995 fractile of the sensor signal envelope. All peaks exceeding this threshold are identified as strong snaps and their arrival times  $\{\tau_{i,j}\}$  are recorded, where *j* is the snap index.

To solve the snap association problem across sensors, we assume that the snapping shrimp producing the snaps are in the far field of the ANI camera. This implies that the wavefront of the snap is planar when it arrives at the ANI camera. Given position vector  $s_i$  of each sensor i and the

arrival times  $\{\tau_{i,j}\}$ , we apply a three-dimensional Hough transform<sup>22</sup> over azimuth, elevation, and arrival time at the origin. The Hough transform output has peaks when a large number of snaps lie close to the plane represented by the Hough transform coordinates. We find the peaks in the Hough transform domain; each peak is assigned a snap number k and associated with a 3-tuple  $(\tilde{\phi}_k, \tilde{\theta}_k, \tilde{\Gamma}_k)$  containing coarse estimates of azimuth, elevation, and arrival time, respectively. The arrival angle estimates are coarse as the bin sizes used in the Hough transform are intentionally large to allow some errors in the arrival time estimates and to limit memory requirement for the Hough accumulator. We use  $3^{\circ}$ bins in azimuth and elevation, extending to  $\pm 60^{\circ}$  in both angular axes. We use time bins at the ROMANIS sampling rate of 196 kSa/s. The coarse estimates are refined in the next step.

#### B. Snap direction and location estimation

Given a coarse arrival angle and time estimate  $(\phi_k, \tilde{\Gamma}_k)$ , we next refine the estimate to give an accurate azimuth, elevation, and arrival time estimate  $(\phi_k, \theta_k, \Gamma_k)$ . This is done by first identifying the snap on each sensor that is consistent with the coarse arrival estimate:

$$\tau_{i}^{'} = \arg \min_{\tau \in \{\tau_{i,i} \forall j\}} \left| \tau - \tilde{\Gamma}_{k} + \frac{s_{i}^{\mathrm{T}} \mathbf{D}(\tilde{\phi}_{k}, \tilde{\theta}_{k})}{c} \right|,$$
(13)

where *c* is the speed of sound and

$$\mathbf{D}(\phi, \theta) = \begin{pmatrix} \cos \phi \cos \theta \\ \sin \phi \cos \theta \\ \sin \theta \end{pmatrix}.$$
 (14)

If the timing error minimized in Eq. (13) is larger than a preset threshold (five samples), the timing from that sensor is not used in the refinement process described next.

The refinement consists of a local iterative search to reduce the mean square error between the observed and estimated arrival times:

$$(\phi_k, \theta_k, \Gamma_k) = \arg\min_{(\phi, \theta, \Gamma)} \sum_i \left| \tau_i' - \Gamma + \frac{s_i^T \mathbf{D}(\phi, \theta)}{c} \right|^2.$$
(15)

Since we wish to locate snaps from shrimp on the seabed, we only consider snaps with  $\theta_k < 0$ . Snaps arriving from positive  $\theta_k$  are ignored, as they most likely correspond to surface reflections. Assuming a flat bathymetry with  $z = z_{\text{seabed}}$ , we compute the location vector  $\Lambda_k$  of snap k as

$$\mathbf{\Lambda}_{k} = \frac{z_{\text{seabed}}}{\sin \theta_{k}} \mathbf{D}(\phi_{k}, \theta_{k}).$$
(16)

Applying the snap detection algorithm to our 30-second ROMANIS dataset, we identify 448 strong snaps in the Hough transform domain. After applying direction refinement to each of the snap, we plot the angle of arrival of the snaps in Fig. 8. Of the 448 snaps, 217 snaps arrive have



FIG. 8. A plot showing the direction of arrival of strong snaps detected in the dataset. Most snaps arrive from the seabed. Some snaps arrive from the sea surface—these are probably surface reflections of snaps originating on the seabed. At an azimuth angle of  $2.4^{\circ}$ , a large number of snaps arrive from the mooring buoy.

negative elevation angles and originate on the seabed. Some snaps clearly arrive from the sea surface; these are probably surface reflections of snaps originating on the seabed. At an azimuth angle of 2.4°, a large number of snaps arrive from various elevation angles; these sounds are believed to originate from the mooring buoy seen in Fig. 5. Whether these sounds are made by the rubbing motion of the mooring cables or by snapping shrimp living on the heavily fouled mooring is not known. We compute the location vectors for each of the 217 snaps arriving from the seabed. The locations of snaps originating within a 100 m × 100 m area in front of ROMANIS are shown in Fig. 9.

## C. Echo identification and range estimation

Having already imaged one or more objects using methods described in Sec. II, we identify the target object to range by selecting a beam pointing towards the dominant part of the object. We use a complex frequency-domain



FIG. 9. A plot showing the location of seabed originating strong snaps within a 100 m  $\times$  100 m area in front of ROMANIS.



FIG. 10. (Color online) An illustration showing the Cartesian coordinate system used and the geometry for range computation using snapping shrimp as a source of opportunity. Not to scale.

beamformer output  $B_f(t)$  for frequency bin f and the selected beam to produce a time series x(t) for echo detection:

$$x(t) = \sum_{f} |B_f(t)|, \tag{17}$$

where the summation is taken over selected frequency bins. We use frequency bins spanning the lower frequency band of 25–37 kHz, as we expect the reflections from the target to be less specular at these frequencies. This allows sources located a wider area to produce detectable echoes from the object.

Let  $\phi_T$  and  $\theta_T$  be the azimuth and elevation angles of the selected beam. We set a maximum range of interest  $R_{max}$ and assume that the object of interest is within this range. We use  $R_{max} = 120$  m as our  $1 \text{ m} \times 1$  m target panels are only expected to be detectable by ROMANIS up to this range; beyond this range, the panels subtend an angle much smaller than the beam resolution of ROMANIS.

For each snap source located on the seabed, we locate peaks in x(t) for  $\Gamma_k < t < \Gamma_k + 2R_{\max}/c$  as candidate echoes from the target. We select a threshold  $\mathbb{F}_{0.95}[x]$  for peak detection to limit the number of candidate echoes identified. For each candidate echo *m* at time  $e_m$ , the time difference between the direct arrival and the target-reflected arrival is computed based on the geometry shown in Fig. 10. Thus the range estimate  $\hat{r}_{k,m}$  to the target must satisfy

$$|\hat{r}_{k,m}\mathbf{D}(\phi_{\mathrm{T}},\,\theta_{\mathrm{T}})-\mathbf{\Lambda}_{k}|+\hat{r}_{k,m}-|\mathbf{\Lambda}_{k}|=c\delta,\tag{18}$$

where  $\delta = e_m - \Gamma_k$ . The positive solution to this equation is given by

$$\hat{r}_{k,m} = c\delta \frac{|\mathbf{\Lambda}_k| + c\delta/2}{|\mathbf{\Lambda}_k| + c\delta - \mathbf{D}(\phi_{\mathrm{T}}, \theta_{\mathrm{T}})^{\mathrm{T}}\mathbf{\Lambda}_k}.$$
(19)

Since the object reflecting the snap must be further from the ANI camera than the snap source, and we are interested only in a range of up to $R_{\text{max}}$ . Any range estimates that do not satisfy  $|\Lambda_k| < \hat{r}_{k,m} < R_{\text{max}}$  are discarded.

For each snap k, we now have a set of ranges  $\{\hat{r}_{k,m}\}$  associated with candidate echoes from the target. We use a voting system to fuse the information from all snaps to find

the range estimate that is most consistent between snaps. To do this, we divide the interval between 0 and  $R_{\text{max}}$  into N range bins with range center  $\rho_b$  for bin b given by

$$\rho_b = \frac{R_{\max}}{N} \left( b - \frac{1}{2} \right) \quad \forall b \in \{1 \cdots N\}.$$
(20)

We associate a width  $\varepsilon$  with each bin such that all ranges in the closed interval  $[\rho_b - \varepsilon/2, \rho_b + \varepsilon/2]$  are considered to fall into that bin. The width  $\varepsilon \ge R_{\text{max}}/N$  and bins may overlap to allow for some error in range estimation.

Let *v* represent an accumulator array to count the votes  $\mathbf{v}_b$  for each range bin *b*. Initially we set  $\mathbf{v} = 0$ . For each snap *k*, we update the votes in all bins *b* that satisfy  $|\mathbf{\Lambda}_k| \leq \rho_b \leq R_{\text{max}}$  since the range estimates outside this range are discarded in a previous step. The update adds a vote for each bin that has one or more candidate echoes in its range interval. For all other bins, it subtracts a fractional penalty vote  $\nu$  for lack of evidence of echoes from that range. After processing all snaps, the range bin with the largest voting score is picked as the estimate for range to the target. This algorithm is summarized below.

Algorithm 1 Voting algorithm to estimate range to target.

Require: Candidate echo ranges $\{\hat{r}_{k,m}\}$
$1: v_b \leftarrow 0 \forall  b \in \{1 \cdots N\}$
2: for all snaps <i>k</i> do
3: for all <i>b</i> s.t. $ \Lambda_k  \le \rho_b \le R_{\text{max}}$ do
4: if $\exists m \text{ s.t. } \rho_b - \epsilon/2 \leq \hat{r}_{k,m} \leq \rho_b + \epsilon/2$ then
5: $v_b \leftarrow v_b + 1$
6: else
7: $v_b \leftarrow v_b - \nu$
8: end if
9: end for
10: end for
11: return $\rho_{\arg \max v_b}$

Since the lowest updated range bin changes with snap location while the highest range bin remains constant, higher numbered range bins generally get updated more frequently. Noisy estimates of range therefore are likely to contribute



FIG. 11. (Color online) Voting scores show a peak at target range of 67 m. No systematic trend in scores with respect to the range bins validates appropriate tuning of penalty  $\nu$ .

more significantly to higher numbered range bins. To counteract this bias, the penalty  $\nu$  is set to be the probability of false echo candidates. This ensures that the expected value of each range bin is zero. In practice the probability of false echo candidates may not be known *a priori* but can be estimated from the data. Moreover, when  $\nu$  is appropriately tuned, the voting scores  $v_b$  show no systematic trend with respect to *b*. This property can be used to tune  $\nu$  when no probability estimate is available.

In our analysis, N = 120,  $\varepsilon = 1.5$ m, and  $\nu = 0.4$ . The resulting voting scores are shown in Fig. 11. There is clear peak in the voting scores, and we obtain a corresponding target range estimate of 67 m. This is consistent with the GPS range measurements ( $65 \pm 5$  m) and active acoustic range measurements (65-70 m) that we performed during the experiment to verify the estimated range.

#### **V. CONCLUSIONS**

As a consequence of the generalized central limit theorem, the pressure time series for snapping shrimp dominated ambient noise is modeled well by a  $S\alpha S$  distribution. When  $\alpha < 2$ , the second and high-order moments become infinite and algorithms based on estimation of these moments suffer from lack of robustness. However, FLOMs and fractiles remain well-defined for  $S\alpha S$  distributions, and can be used in place of high-order moments in ambient noise imaging algorithms. Data collected by the ROMANIS acoustic camera shows that these new FLOM and fractile based algorithms perform robustly in snapping shrimp dominated waters around Singapore. The performance of fractile algorithms is qualitatively better than the FLOM algorithms. Although the statistical methods are developed based on intuitions from  $S\alpha S$  noise distribution, the methods are not critically dependent on the distribution; the methods are likely to work well in most impulsive noise environments.

The source of loud snaps can be localized in space and time using time of arrival information from various sensors in the acoustic camera and the knowledge of the local bathymetry. Echoes of each loud snap from an object of interest can be processed to yield range estimates for the object without having to insonify the object actively. The key challenges involved are those of data association and fusion. They can be solved using a combination of a threedimensional Hough transform and a voting algorithm. Target range estimates obtained from the ROMANIS dataset using this approach are consistent with GPS and acoustic range measurements.

In warm shallow waters, snapping shrimp provide a distributed set of episodic sources that are well suited as sources of acoustic illumination for ambient noise imaging. Ambient noise imaging is shown to be a powerful technique that not only images submerged underwater objects passively, but is also able to estimate range to these objects.

# ACKNOWLEDGMENTS

We would like to thank Amogh Raichur, Tan Soo Pieng, and Naveen Chandhavarkar for helping rebuild ROMANIS and making it a robust imaging camera. We would also like to thank Mohan Panayamadam and numerous others at ARL for their support in field experiments and dive operations. A special thanks to John Potter, who first set us on the path of ANI, and championed as well as led the development of the first version of ROMANIS. Finally, we would like to thank the anonymous reviewers whose constructive feedback has helped strengthen this paper.

- <sup>1</sup>S. Flatté and W. Munk, "Submarine detection: Acoustic contrast versus acoustic glow," JASON Report No. JSR-85-108, MITRE corporation, McLean, Virginia, 1985.
- <sup>2</sup>M. J. Buckingham, B. V. Berkhout, and S. A. L. Glegg, "Imaging the ocean with ambient noise," Nature (London) **356**, 327–329 (1992).
- <sup>3</sup>M. J. Buckingham, J. R. Potter, and C. L. Epifanio, "Seeing in the ocean with background noise," Sci. Am. **274**, 86–90 (1996).
- <sup>4</sup>J. Potter and M. Chitre, "ADONIS imaging with a Kalman filter and higher-order statistics," in *3rd European Conference on Underwater Acoustics*, Greece (1996), Vol. 1, pp. 349–354.
- <sup>5</sup>C. L. Epifanio, J. R. Potter, G. B. Deane, M. L. Readhead, and M. J. Buckingham, "Imaging in the ocean with ambient noise: the ORB experiments," J. Acoust. Soc. Am. **106**, 3211–3225 (1999).
- <sup>6</sup>J. R. Potter and M. Chitre, "Ambient noise imaging in warm shallow seas; second-order moment and model-based imaging algorithms," J. Acoust. Soc. Am. **106**, 3201–3210 (1999).
- <sup>7</sup>M. Readhead, "Snapping shrimp noise near Gladstone, Queensland," J. Acoust. Soc. Am. **101**, 1718–1722 (1997).
- <sup>8</sup>J. R. Potter, T. W. Lim, and M. A. Chitre, "Ambient noise environments in shallow tropical seas and the implications for acoustic sensing," in *Oceanology International*' 97, Singapore (1997), pp. 1–9.
- <sup>9</sup>M. A. Chitre, J. R. Potter, and S. H. Ong, "Optimal and near-optimal signal detection in snapping shrimp dominated ambient noise," IEEE J. Oceanic Eng. **31**, 497–503 (2006).

- <sup>10</sup>M. Shao and C. Nikias, "Signal processing with fractional lower order moments: Stable processes and their applications," Proc. IEEE **81**, 986–1010 (1993).
- <sup>11</sup>V. Pallayil, M. A. Chitre, E. T. Tan, J. Potter, T. B. Koay, S. B. Ruiz, and S. P. Tan, "Ambient noise imaging—First deployments of ROMANIS and preliminary data analysis," in *OCEANS 2003 IEEE/MTS*, San Diego, (2003), pp. 1–7.
- <sup>12</sup>M. A. Chitre and J. R. Potter, "Optimisation and beamforming of a twodimensional sparse array," in *High Performance Computing 1998* (1998), Vol. 1, pp. 452–459.
- <sup>13</sup>M. Johnson, F. Everest, and R. Young, "The role of snapping shrimp (Crangon and Synalpheus) in the production of underwater noise in the sea," Biol. Bull. **93**, 122–138 (1947).
- <sup>14</sup>B. Ferguson and J. Cleary, "*In situ* source level and source position estimates of biological transient signals produced by snapping shrimp in an underwater environment," J. Acoust. Soc. Am. **109**, 3031 (2001).
- <sup>15</sup>J. R. Potter, "Ambient noise imaging techniques and potential in warm shallow water," in *Proceedings of the ASW Asia Workshop* (2003), pp. 1–11.
- <sup>16</sup>S. Kuselan, M. Chitre, and V. Pallayil, "Ambient noise imaging through joint source localization," in OCEANS 2011 Kona (IEEE/MTS) (2011), pp. 1–5.
- <sup>17</sup>G. Samorodnitsky and M. S. Taqqu, *Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance* (Chapman and Hall, New York, 1994), pp. 1–632.
- <sup>18</sup>X. Ma and C. Nikias, "Parameter estimation and blind channel identification in impulsive signal environments," IEEE Trans. Signal Process. 43, 2884–2897 (1995).
- <sup>19</sup>E. F. Fama and R. Roll, "Parameter estimates for symmetric stable distributions," J. Am. Stat. Assoc. 66, 331–338 (1971).
- <sup>20</sup>The MathWorks, Inc., Image Processing Toolbox User's Guide, R2012a edition (2012), pp. 1–614.
- <sup>21</sup>The MathWorks, Inc., *MATLAB Mathematics*, R2012a edition (2012), pp. 1–580.
- <sup>22</sup>J. Illingworth and J. Kittler, "A survey of the Hough transform," Comput. Vis. Graph. Image Process. 44, 87–116 (1988).