Evidence for Acoustic Imaging Capability in a Bottlenose Dolphin

R. A. Altes
Chirp Corp.
8248 Sugarman Drive
La Jolla, CA, 92037, USA
altes@att.net

L.A. Dankiewicz
Science Applications International Corp.
3990 Old Town Avenue, Suite 105A
San Diego, CA, 92152, USA
loisd@spawar.navy.mil

P.W. Moore
SSC San Diego, Code 23502
53560 Hull Street
San Diego, CA, 92152, USA
pmoore@spawar.navy.mil

David A. Helweg
Pacific Island Ecosystems Research Center
3190 Maile Way, Rm 408
Honolulu, HI, 96822, USA
dhelweg@usgs.gov

Abstract*- A necessary condition for acoustic imaging (e.g., synthetic aperture sonar processing) is the capability to sum echo samples from the same point in the environment over different signal-echo pairs. To decide whether a dolphin has this capability, a limited number *N* of electronically simulated echoes with constant delay were transmitted back to an echolocating dolphin. The experiment was performed in San Diego Bay, which has a large indigenous population of snapping shrimp. The signal-to-noise ratio (SNR) of the simulated echoes was controlled by adding artificial Gaussian noise and by varying echo amplitude. If the dolphin is capable of SAS-like imaging, then the SNR required for detection should decrease as the number of available echoes *N* is increased. This phenomenon was indeed observed. The best receiver model for describing the dependence of SNR on *N* uses binary summation (an M-out-of-N detector). Binary summation is robust against the strong impulsive interference produced by snapping shrimp. The effect of binary summation on SAS-like processing is assessed by creating SAS images from binary-quantized data at the output of a broadband, dolphin-like sonar, and comparing these images to those obtained without binary quantization.

I. INTRODUCTION

Sonar signals of the Atlantic Bottlenose dolphin (*Tursiops truncatus*) usually have very broad bandwidth (30 to 150 kHz) and small time-bandwidth product [1]. Many such signals typically are used to interrogate an object of interest as the animal moves. Physical sonar dimensions are limited by the size of the animal, and transmitter/receiver beam widths are thus large in comparison with many human-made sonar systems.

The utilization of multiple signals by a moving, wideband sonar system with small physical aperture size is suggestive of synthetic aperture sonar (SAS). In SAS, the different locations of the moving transmitter/receiver constitute element positions in a synthetic array. The array can be focused on an object of interest by a delay-and-sum operation applied to received echoes. Bearing resolution is determined by the dimension of the synthetic array, rather than by the physical sonar dimensions.

* This work was partially sponsored by the Office of Naval Research and the Defense Research Projects Agency (DSO).

If the synthetic array is sufficiently long, its cross-range resolution (bearing resolution multiplied by range) is comparable to the range resolution of the wideband transmitted waveform (~ 1 cm for *Tursiops*). In this case, echo samples can be used to form a two-dimensional image with similar resolution in both dimensions.

Does *Tursiops* utilize an imaging process similar to SAS? A necessary condition for SAS-like processing is the capability to implement a delay-and-sum operation on echoes from multiple signals. The simplest version of the delay-and-sum operation occurs when the echo delay is the same for each transmitted signal. The corresponding test is to determine whether a dolphin can add *N* echoes from a stationary target. This paper describes the design of such a test, an interpretation of the results, and some implications for biomimetic acoustic imaging.

II. EXPERIMENTAL DESIGN

A dolphin was trained to report detection of synthetic echoes generated in response to the dolphin's clicks, placing the number of echoes available to the dolphin under experimental control. The dolphin's signal detection performance in additive noise was assessed when 1,2,4,8, and 16 echoes were made available. Although the available number of echoes (*N*) was preset, the number of emitted dolphin clicks was not controlled and could be much larger than *N*. If the dolphin adds echoes from a stationary target, then the echo SNR required for detection should decrease when the available number of echoes is increased.

The NO-GO (signal absent) stimulus consisted of Gaussian noise (white over the dolphin echolocation bandwidth) over the 4 sec trial duration, and the GO (signal present) stimulus consisted of noise plus simulated echoes. Both the GO and NO-GO stimuli were presented to the dolphin through a wideband (Reson TC4013) projector located 1.34 m from the dolphin's melon, directly in front of the dolphin. The simulated echoes were triggered by the dolphin's click, as received by another TC4013 transducer located 0.7 m in front of the echo projector. An 8ms delay was inserted between the trigger event and the simulated echo, resulting in a total simulated range of 7 m.

The simulated echo was a triangle-windowed sinusoid with 50 kHz center frequency and 80 μsec duration (Fig. 1). The computer-generated noise power was fixed at 95 dB SPL re: $1\mu Pa^2/Hz$ between 10 kHz and 150 kHz. At 50 kHz, the ambient noise level was approximately 70 dB re: $1\mu Pa^2/Hz$ in the area of San Diego Bay where the experiment was performed. The simulated Gaussian noise level was thus 25 dB above the time-averaged ambient noise level, and 38 dB above the ambient noise level when the dolphin's directivity index is considered [1].

The 4 sec white noise burst defined the trial duration for the dolphin, which was stationed in a hoop 1.35 m below the water surface. To report a signal present condition, i.e., a decision that the signal-plus-noise hypothesis H1 was true (GO response), the dolphin immediately moved to a nearby paddle and touched it with her rostrum. To indicate the absence of a signal, i.e., a decision that the noise-only hypothesis H0 was true (NO-GO response), she remained stationary in the hoop for the 4sec trial duration. Tone and fish rewards were given for every correct response. An equal number of GO and NO-GO responses was presented in a randomized Gellerman series [2]. The dolphin's motivation to perform reliably was assessed by ten warm-up trials before every session, with an 80% correct response rate required in order for a test session to ensue. No more than one experimental session was conducted in a day.

Thresholds were estimated for both training and test phases by using a signal amplitude titration method (up/down staircase) [3]. The GO signal amplitude was held constant at an easily discernable level for the first ten trials of the session. After the first ten trials, 0.2 V decrements in signal amplitude were made until the dolphin responded incorrectly to the GO stimulus. Amplitude was raised in 0.1 V increments until the dolphin detected the signal again. Subsequent amplitude adjustments were in 0.1 V steps, with decrements after every correct GO response and increments after every incorrect GO response. After the dolphin became experienced with the task and echo amplitudes were reduced, 0.05 V increments and decrements were used. A change in the direction of amplitude adjustment constituted a "reversal," and a threshold estimate was calculated from the mean decibel level at ten reversal points.

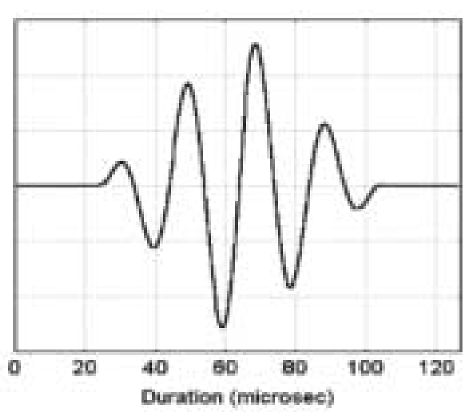

Fig. 1. Simulated echo projected back to dolphin in response to transmitted click, with total delay corresponding to a range of 7m.
Center frequency = 50 kHz.

During training, a generous number of synthetic echoes were provided (*N*=256) and SNR was held at approximately 40 dB (1.0 V). Stimuli were presented in three to six same-trial blocks, e.g., four NO-GO trials followed by six GO trials. Approximately four sessions were required before appropriate responses were observed. Trial type was then randomized, and *N* was gradually reduced after responding was stable. No further training was undertaken after the dolphin demonstrated stable performance at the minimum *N*-level (*N*=1). All experimental conditions occurred in the training phase so that test-phase thresholds were free of novelty effects.

During the test phase, the titration procedure was used with constant *N* to obtain two final threshold measurements for *N*=1,4,8, and 16, and three final threshold measurements for *N*=2. These thresholds were estimates of the required echo amplitudes *A(N)* for detection with N available echoes in artificial and ambient noise.

III. EXPERIMENTAL RESULTS

In order to convert the threshold measurements *A*(*N*) into SNR values, the relevant noise power must be computed from the noise power spectral density and the effective bandwidth of the dolphin's receiver. Noise masking experiments indicate that dolphin hearing, like that of other mammals, can be modeled by passing auditory data through a parallel set of overlapping band-pass filters. The bandwidth of each of these filters depends on center frequency and is called the critical band. At the center frequency of the synthetic echo (50 kHz), critical band measurements indicate that the effective noise bandwidth is approximately 22.72 kHz [4]. The 10 dB bandwidth of the simulated echo in Fig. 1 is approximately the same as the dolphin critical bandwidth at the echo center frequency, and the echo is approximately matched to a critical band. The SNR for each echo thus was computed using *A*(*N*) together with a noise bandwidth of 22.72 kHz.

The detection threshold SNR's that were measured during the test phase of the experiment are shown in Fig. 2. The two measured threshold SNR values for *N*=1 were identical, and are thus represented by a single point in the figure. The mean false alarm rate was 0.034. SNR required for detection decreases monotonically as the number of available echoes (*N*) is increased.

The number of transmitted clicks always exceeded the number of available echoes, thus assuring that all the available echoes were utilized by the dolphin. The number of transmitted clicks was significantly higher for *N*=1 relative to all other cases, implying that the *N*=1 condition is the most difficult.

IV. INTERPRETATION OF THE RESULTS

When N echoes are available to the dolphin, Fig. 2 shows a monotonic decrease in the SNR required for detection as N increases, as expected for echo summation. The three receivers in Fig. 3 (among others) incorporate echo summation. Which receiver is the best model for the results in Fig. 2?

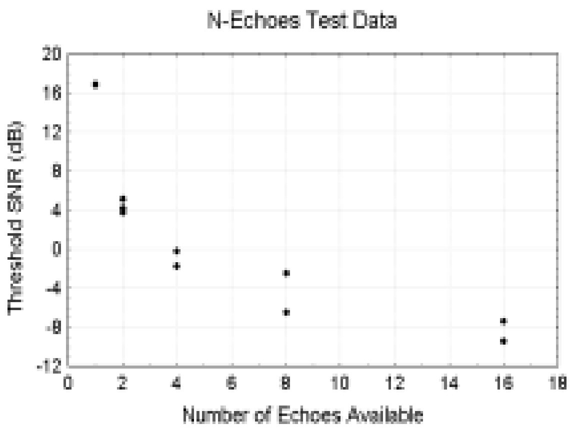

Fig. 2. Final detection thresholds from test session data. Two threshold measurements were made for N=1, but only one point is shown because the measured threshold values are identical.

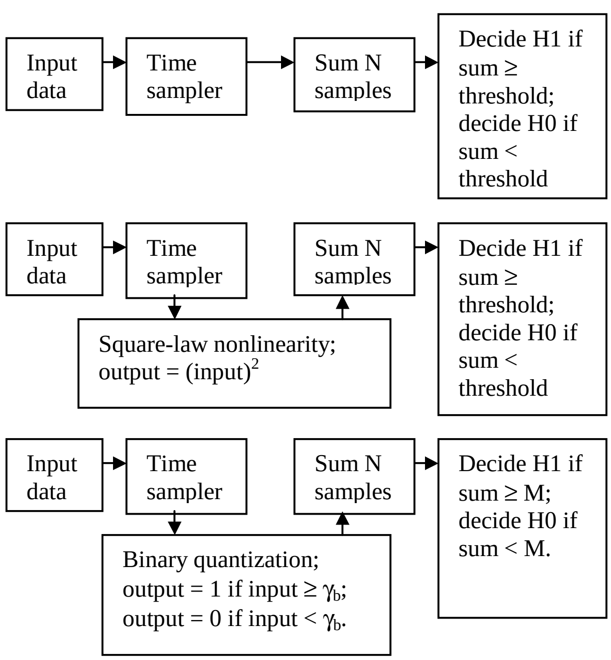

Fig. 3. Three receiver models: Linear summation (top), energy detection (center), and binary summation (bottom).

The N-echo detection performance of the three receiver models is predicted by the analysis in the Appendix. For the linear summation, energy detection, and binary M-out-of-N receivers, the required echo amplitudes $A(N)$ for detection with N available echoes, divided by the rms noise level σ, are

$$[A(N)/\sigma]_{\text{lin}} = c_l / \sqrt{N} \tag{1}$$

$$[A(N)/\sigma]_{\text{egy}} = (c_e / \sqrt{N})\left[1 + \sqrt{1 + (2N / c_e^2)}\right]^{1/2} \tag{2}$$

$$[A(N)/\sigma]_{\text{bin}} = erfc_*^{-1}(p_0) - erfc_*^{-1}(p_1) \tag{3}$$

where c_l and c_e are constants, $erfc_*^{-1}(\bullet)$ is the inverse complementary error function as defined in the Appendix, p_0 is the probability that the threshold level of the binary quantizer is exceeded when only noise is present (the H0 hypothesis) and p_1 is the probability that the binary quantizer threshold is exceeded when both signal and noise are present (the H1 hypothesis).

In the binary summation model, the function $erfc_*^{-1}(p)$ is related to the probit transformation [5]

$$erfc_*^{-1}(p) = probit(1-p) \tag{4}$$

and p_1 depends upon p_0 , N, and a constant c_b:

$$p_1 = \frac{(2p_0 + c_b/N) + \sqrt{(2p_0 + c_b/N)^2 - 4p_0(1 + c_b/N)[p_0(1 + c_b/N) - c_b/N]}}{2(1 + c_b/N)}. \tag{5}$$

For a pre-specified value of p_0, all the $A(N)/\sigma$ expressions in (1)-(3) depend on a constant (c_l, c_e, or c_b) and on the number of available echoes, N. Pre-specifying p_0 is equivalent to choosing a threshold γ_b for binary quantization. The most appropriate choice for the binary quantization threshold in Fig. 3 is zero. In this case, p_0 = 0.5 for all symmetric, zero-mean noise distributions, independent of the noise power σ^2.

To compare the receiver models with dolphin detection data, the parameters c_l, c_e and c_b (with $p_0 \equiv 0.5$) are adjusted to provide a minimum mean-square error (MMSE) fit between the values of $[A(N)/\sigma]_{\text{model}}$ in (1)-(3) and the average experimental value of $A(N)/\sigma$ for each N-value. The correlation coefficients between the average data points and their theoretical counterparts are computed for each model. For visual comparison, the data points and theoretical curves are plotted together in Fig. 4 on a decibel scale, showing $20\log_{10}[A(N)/\sigma]$ vs. N.

Fig. 4 illustrates that the best fit (by far) is obtained with the binary M-out-of-N receiver. The best-fit parameters and data-model correlation coefficients r are as follows:

Linear summation: c_l=4.58, r_l=0.9055 (6)

Energy detection: c_e=3.12, r_e=0.9095 (7)

Binary summation: p_0=0.5, c_b=0.9999993, r_b=0.9997. (8)

The more accurate specification of c_b is necessitated by receiver operation on a steep part of the probit function when N=1, as discussed in the Appendix.

The detectability index is a measure of receiver performance. It is the distance between the mean values of the detector output distributions for the H1 and H0 hypotheses, normalized by the square root of the average variance of the two distributions. The constants c_l, c_e, and c_b are related to the corresponding detectability indices by the equations

$$c_l = d_l;\ c_e = d_e;\ c_b = d_b^2/2. \qquad (9)$$

The detectability indices corresponding to the MMSE estimates of c_l, c_e, and c_b in (6-8) are

$$d_l = 4.58;\quad d_e = 3.12;\quad d_b = 1.41 \qquad (10)$$

for the linear summation, energy detection, and binary M-out-of-N receivers, respectively. The receiver model with the best fit to the data has the worst performance in zero-mean Gaussian noise.

The linear receiver is optimum for Gaussian noise with known variance, but the binary M-out-of-N processor has distribution tolerant (nonparametric) properties. The false alarm rate of the binary M-out-of-N receiver is insensitive to time-varying noise power and to the shape of any symmetric, zero-mean noise distribution. Fig. 4 and (10) imply that the dolphin has traded optimality in Gaussian noise with specified noise power for robustness with respect to the distribution and power of the noise. The performance disparity between binary and linear summation is not large if many echoes are used. Fig. 4 implies that large N is associated with small SNR for all of the models. For large N and small SNR, it is shown in the Appendix that $d_l \approx 1.25 d_b$.

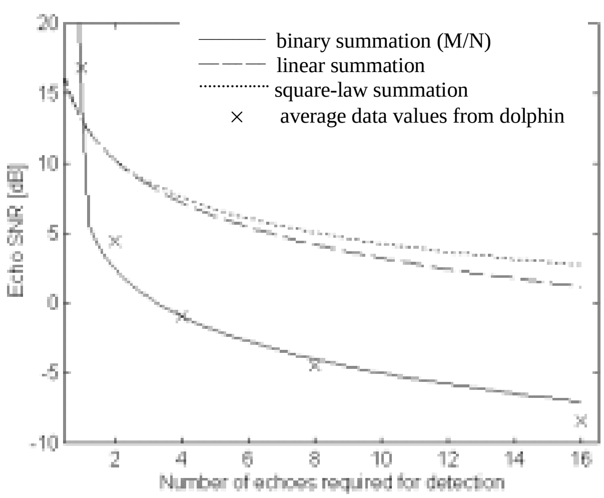

Fig. 4. MMSE fits between three receiver models and dolphin N-echo detection data. The MMSE algorithm compares models to noise-normalized data amplitudes $A(N)/\sigma$, but the curves are shown on a dB scale corresponding to SNR[dB] = $20\log_{10}[A(N)/\sigma]$.

The power spectrum of artificially added Gaussian noise (white over the dolphin's echolocation bandwidth) was 95 dB re: 1μPa^2/Hz. When an extra 13 dB is added to account for the dolphin's directivity index, the artificial Gaussian noise was 108 dB above spatially uniform noise with a power spectral density of 1μPa^2/Hz and 38 dB above the average ambient noise level of 70 dB re: 1μPa^2/Hz. The ambient level included average snapping shrimp interference and the sounds of other dolphins in the area. Despite the 38 dB difference, the analysis indicates that the dolphin's receiver did not adapt to operation in Gaussian noise as opposed to strong transient interference.

This apparent lack of adaptability can be explained by the large disparity between peak and average levels of transient sounds. At a distance of one meter, the power spectral density of a single snapping shrimp pulse is between 105 and 111 dB re: 1μPa^2/Hz [6,7]. Spherical spreading decreases this level by 20 log r if the shrimp is r meters from the receiver. A spatially averaged interference power spectral density level of 108 dB re: 1μPa^2/Hz is required to make the interference power spectrum equal to the power spectrum of the artificially added Gaussian noise at the input to the dolphin's receiver. The possible presence of nearby snapping shrimp thus could have constrained the dolphin's receiver design. The experiment was performed in an area with floating walkways to support trainers and equipment. Snapping shrimp appear to congregate in such areas [8]. The binary summation (M-out-of-N detection) process limits the effect of high power transient interference such as that introduced by snapping shrimp.

All the artificial echoes had the same amplitude $A(N)$ in a given N-echo trial. The dolphin thus could discard any cues that are associated with echo-to-echo amplitude variation for N-echo detection. When changes in amplitude from one echo to the next are important, the dolphin is sensitive to them [9]. Binary quantization is associated with insensitivity to amplitude. The introduction of such changes could affect the current receiver model.

The next experiment to test for SAS-like imaging in dolphins will be to move the simulated target in a predictable way, e.g., by gradually changing its range. If multi-echo delay-and-sum beam forming can be performed by the dolphin, then predictable movement of the simulated target should not greatly affect detection performance. Another interesting experiment would be to repeat the current procedure with better acoustic isolation of the dolphin from snapping shrimp sounds.

V. IMPLICATIONS FOR ACOUSTIC IMAGING

Assuming that SAS-like images can be formed by an echolocating dolphin, it is relevant to consider the effects of binary quantization on such images. Two easily predicted effects are: (1) A reduction in masking of weak point scatterers by strong ones, and (2) a contrast stretch that accentuates many of the relevant pixels in an amplitude normalized image. The first effect is predicted from the SAS range, cross-range ambiguity function for binary-quantized data. The second effect follows from the amplitude limit that is imposed by binary quantization.

Each pixel value in a SAS-like image is calculated by summing echoes from the same scattering point, as seen from multiple aspect angles. At each aspect, the contribution from a given scattering point forms a line in the image plane. The magnitude of the line is the echo sample value, and the line is orthogonal to the direction of propagation. Delay-and-sum synthetic aperture processing yields a sum of lines with different orientations, intersecting at the location of the desired scattering point. This sum of lines corresponds to the range, cross-range ambiguity function of the SAS sonar system and to the point spread function of the SAS image. For 180 degree rotation at constant aspect increments, the sum of unit-amplitude lines is as shown in Fig. 5 [10].

If *N* echoes from different aspects are used to create a SAS image, then the peak-to-sidelobe ratio of the point spread function equals *N*. Suppose that a weak scattering point is being imaged at the center of Fig. 5, and a strong scattering point is located on one of the sidelobes. If the reflectivity of the strong scattering point is larger than the central point by a factor of *N*, then the strong point will mask the weak point, and a spurious line through the strong scattering point will be seen in the image.

The pixel value at the center point in Fig. 5 is obtained from a weighted sum of the lines passing through the center point, with weights equal to the measured reflectivity of the imaged point at each aspect angle. Binary summation replaces the weighted sum with a sum of values that are either zero or one, depending upon whether the measured reflectivity at a given aspect exceeds threshold. A very strong scattering point on one of the sidelobes makes a unit contribution to the sum, thus preventing the strong interferer from completely masking the weak scattering point at the image center.

An advantage of binary summation for SAS is to limit the effect of interference caused by strong scattering points, just as the effect of snapping shrimp interference is limited in a binary M-out-of-N detector. The disadvantage of binary summation is the relative suppression of a scattering point with echoes that are not consistently above threshold as the aspect varies. A SAS image constructed with binary summation represents the aspect invariance (aspect tolerance) of each supra-threshold scattering point.

For an amplitude-normalized image, a particularly large scattering point can suppress the pixel values in the rest of the image. This effect is mitigated by binary quantization. When zero-one binary quantization is used, the maximum value of the central point in Fig. 5 is limited to *N*, the number of different aspects that are used to form the image. Most of the medium-to-large pixel values are accentuated in an amplitude normalized image constructed from binary-quantized data, while a few extremely large pixels are reduced but not eliminated. The result is a contrast-stretched image.

Figs. 6a and 6b illustrate typical effects of binary quantization on an acoustic image. Each pixel value in Fig. 6a was formed by non-coherent summation of relevant

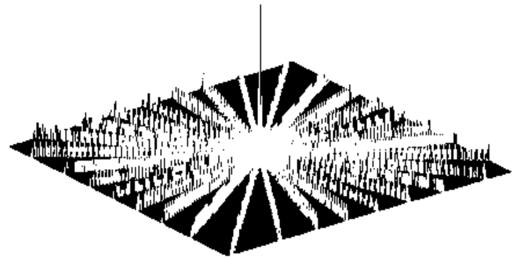

Fig. 5. Example of a SAS range, cross-range ambiguity function (point spread function) in the image plane.

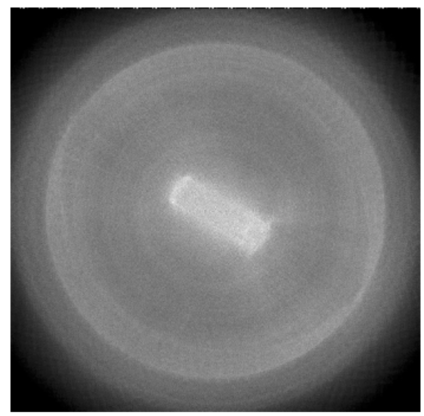

Fig. 6a. Biologically inspired acoustic image of a culvert (an open ended concrete cylinder) resting on a textured turntable. Total width of image is 7.7 m.

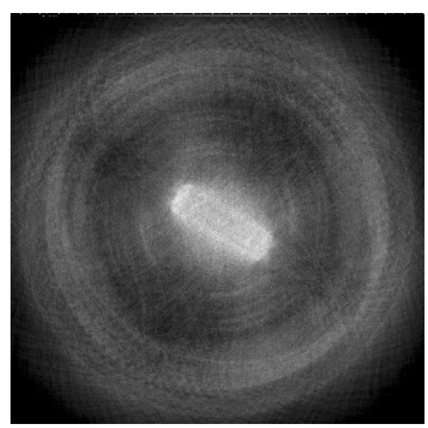

Fig. 6b. Effect of binary quantization of the echo sample values that are summed over different aspect angles to form the image in Fig. 6a. Both images are amplitude normalized.

samples (one sample from each echo) from pulse-compressed outputs of the two "ears." The echoes were obtained from a wideband sonar system that emulates the dolphin's transmitter beam width (~12 deg), receiver beam width (~24 deg), projector/receiver bandwidth (~100 kHz), and interaural distance (~12 cm) [11]. An open-ended concrete cylinder (a culvert) was placed on a textured surface that covered a turntable with 6.1 m diameter. The sonar was 22.9 m from the center of the turntable, at a grazing angle of 3.5 deg. The image was formed from echoes that were sampled at 3.3 deg aspect intervals over 360 deg. The effect of binary summation was assessed by using a quantization threshold at the average value of the echo samples that were used to construct Fig. 6a. Samples above this threshold were set equal to one, and samples below the threshold were set to zero. The resulting image is shown in Fig. 6b.

In Fig. 6b, weak scatterers are suppressed by the quantization threshold. Scattering points that are consistently above threshold are accentuated relative to strong transients or "glints" that occur as the bearing is changed. The amplitude limitation imposed by binary quantization causes a relative increase in the amplitude of relevant pixel values that are less than the maximum value.

Fig. 7a is an acoustic image of a truncated cone, obtained under the same conditions and with the same signal processing as in Fig. 6a. The corresponding image for binary-quantized data is shown in Fig. 7b. A point on an edge of the truncated cone exhibits high variation in backscatter amplitude as the aspect is changed. Such points are suppressed in an acoustic image formed from binary-quantized data.

VI. CONCLUSION

Characteristics of dolphin sonar systems are suggestive of multi-echo combining for acoustic imaging. A necessary condition for such imaging is the capability to utilize information about the same scattering point obtained from multiple echoes. This condition was tested and verified via a behavioral experiment. Three receiver models for echo combining are considered, and the best model uses binary summation (M-out-of-N detection). This receiver model is relevant to the experimental environment, which includes high amplitude, impulsive interference from snapping shrimp. Binary summation is used to form acoustic images, and the results are compared to more conventional images. The comparison shows that a binary summation image experiences contrast stretching and edge suppression.

VII. ACKNOWLEDGMENTS

We thank Justine M. Zafran and Michael D. Massimi for their assistance during the behavioral data collection phase of this study. This research was made possible in part thanks to support from Dr. John Tague and Dr. Robert Gisiner (Office of Naval Research 321US and 342), and from Dr. Allen Rudolph of the Defense Sciences Office of the Defense Research Projects Agency.

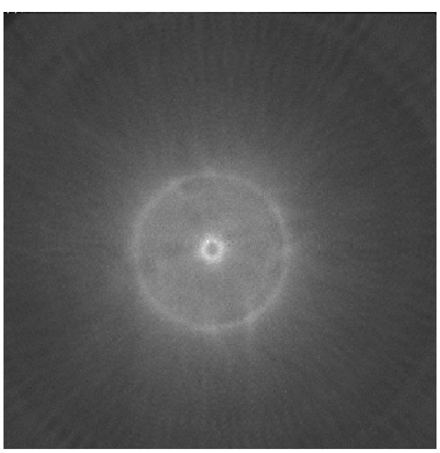

Fig. 7a. Biologically inspired acoustic image of a truncated cone resting on a textured turntable.

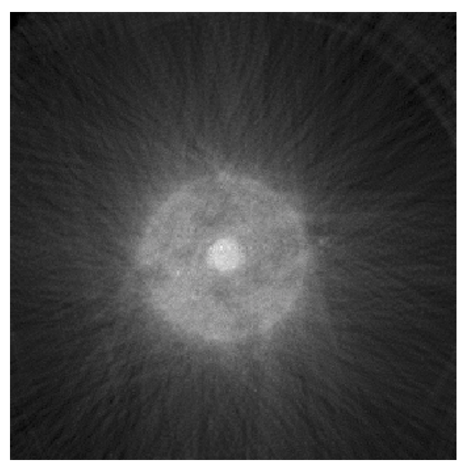

Fig. 7b. Effect of binary quantization of the echo samples that are summed over different aspect angles to form Fig. 7a.

VIII. APPENDIX

Detection performance varies monotonically with the detectability index, which is defined as the difference between the output mean value when hypothesis H1 (signal plus noise) is true and the output mean value when H0 (noise alone) is true, divided by the square root of the average output variance for the two hypotheses. In the following analysis, the detectability index of each receiver model is assumed to be constant for all values of *N*, the number of echoes available for detection. This assumption implies that each receiver model (as well as the dolphin) uses a consistent performance criterion (detection, false, alarm, and error probability) for decision-making at all *N*-values.

The linear summation model computes the function

$$h(\underline{x}) = (1/N)\sum_{i=1}^{N}(x_i + A) \tag{A.1}$$

where A is the sampled signal value and $\{x_i;\ i=1,\ldots,N\}$ are independent identically distributed noise samples (one from each echo) with zero mean and variance σ^2; $E(x_i)=0$ and $E(x_i x_j)=\sigma^2$ if $x_i=x_j$ and zero otherwise. $E(x)$ is the ensemble expected value of x, and $\underline{x}$ is the set of noise samples $x_1,x_2,\ldots,x_N$. It follows that

$$E[h(\underline{x})] = (1/N)\sum_{i=1}^{N}E(x_i + A) = A \tag{A.2}$$

and

$$E[h^2(\underline{x})] = (\sigma^2/N) + A^2. \tag{A.3}$$

The variance of the averaged, linearly transformed data is then

$$Var[h(\underline{x})] = E[h^2(\underline{x})] - E^2[h(\underline{x})] = \sigma^2/N. \tag{A.4}$$

When hypothesis H1 is true, the data consist of signal plus noise with $A\neq 0$. If H0 is true, the data consist of noise alone ($A=0$). The corresponding detectability index is

$$d_l = \frac{|E[h(\underline{x})|H1] - E[h(\underline{x})|H0]|}{\left[(1/2)\{Var[h(\underline{x})|H0] + Var[h(\underline{x})|H1]\}\right]^{1/2}}$$

$$= \sqrt{N}A/\sigma. \tag{A.5}$$

In the psychophysical literature, d' is used instead of d. The prime is omitted here in order to simplify notation in the equations.

The linear summation model is compared with data by adjusting the constant $c_l=d_l$ in the equation

$$[A(N)/\sigma]_{\text{lin}} = c_l/\sqrt{N} \tag{A.6}$$

to obtain a minimum mean-square error (MMSE) fit of $[A(N)/\sigma]_{\text{lin}}$ to $[A(N)/\sigma]_{\text{expt}}$ for the experimental values of N. The resulting best fit is then evaluated via the correlation coefficient between $[A(N)/\sigma]_{\text{lin}}$ and $[A(N)/\sigma]_{\text{expt}}$.

For an average of squared data samples (energy detection),

$$h(\underline{x}) = (1/N)\sum_{i=1}^{N}(x_i + A)^2 \tag{A.7}$$

and

$$E[h(\underline{x})] = (1/N)\sum_{i=1}^{N}E(x_i + A)^2 = \sigma^2 + A^2. \tag{A.8}$$

For a zero-mean Gaussian random variable x_i with $E(x_i)=0$, $E(x_i^2)=\sigma^2$, $E(x_i^3)=0$, $E(x_i^4)=3\sigma^4$, and $E(x_i x_j)=\sigma^2$ if $x_i=x_j$ and zero otherwise,

$$E[h^2(\underline{x})] = (1/N)(3\sigma^4 + 6\sigma^2A^2 + A^4) + [(N-1)/N](\sigma^2 + A^2)^2$$

and (A.9)

$$Var[h(\underline{x})] = (2\sigma^2/N)(\sigma^2 + 2A^2). \tag{A.10}$$

The detectability index is then

$$d_e = (A/\sigma)^2\sqrt{(N/2)/[(1+(A/\sigma)^2]} \tag{A.11}$$

and

$$[A(N)/\sigma]_{\text{egy}} = (c_e/\sqrt{N})\left[1+\sqrt{1+2(c_e/\sqrt{N})^{-2}}\right]^{1/2} \tag{A.12}$$

where $c_e=d_e$. To evaluate the general square-law model, the constant c_e in (A.12) is adjusted to obtain a MMSE fit of $[A(N)/\sigma]_{\text{egy}}$ to $[A(N)/\sigma]_{\text{expt}}$ for the experimental values of N.

For the M-out-of-N receiver model, the binomial distribution [12] describes the probabilities of various numbers of ones and zeros at the output of the binary quantizer for N echoes. Let p_1 equal the probability that the binary random variable equals one (the sampled data value is greater than the binary quantization threshold, γ_b) when an echo is present (the signal plus noise condition, H1). Let p_0 equal the probability that the binary random variable is one when the echo is absent (the noise alone condition, H0). For one binary sample from each of N echoes,

The expected number of "ones" with echo present (H1) equals $N p_1$.

The expected number of "ones" with echo absent (H0) equals $N p_0$.

The variance of the distribution of the number of "ones" given H1 equals $N p_1 (1- p_1)$.

The variance of the distribution of the number of "ones" given H0 equals $N p_0 (1- p_0)$.

The detectability index is the difference in means divided by the square root of the average variance:

$$d_b = \frac{Np_1 - Np_0}{\sqrt{(1/2)[Np_1(1-p_1) + Np_0(1-p_0)]}} \tag{A.13}$$

If a threshold value γ_b is used to convert echo samples into binary data, the probability that the binary random variable equals one when the signal is absent (noise alone) is

$$p_0 = erfc_*(\gamma/\sigma), \tag{A.14}$$

and the probability that the binary random variable equals one when the signal is present is

$$p_1 = erfc_*\{[\gamma - A(N)]/\sigma\}, \tag{A.15}$$

where σ is the rms noise power and the complementary error function *erfc*(x)* is the integral of a zero-mean, unit variance normal distribution between *x* and infinity.

From (A.14), the threshold level is

$$\gamma = \sigma \; erfc_*^{-1}(p_0) \qquad \text{(A.16)}$$

where $erfc_*^{-1}(p_0)$ is the inverse complementary error function of p_0. Similarly, (A.15) can be solved for $A(N)$:

$$A(N) = \gamma - \sigma \;\; erfc_*^{-1}(p_1)\cdot \qquad \text{(A.17)}$$

Substituting (A.16) into (A.17),

$$[A(N)/\sigma]_{\text{bin}} = erfc_*^{-1}(p_0) - erfc_*^{-1}(p_1)\,. \qquad \text{(A.18)}$$

Letting $c_b = (1/2)d_b^2$ and solving (A.13) for p_1 yields

$$p_1 = \frac{(2p_0 + c_b/N) \pm \sqrt{(2p_0 + c_b/N)^2 - 4p_0(1 + c_b/N)[p_0(1 + c_b/N) - c_b/N]}}{2(1 + c_b/N)} \qquad \text{(A.19)}$$

where the plus-or-minus operation must always be plus in order for p_1 to be non-negative. For a given p_0 value in (A.18), the parameter c_b, in (A.19) can be adjusted to obtain a MMSE fit of $[A(N)/\sigma]_{\text{bin}}$ to $[A(N)/\sigma]_{\text{expt}}$. The resulting values of $[A(N)/\sigma]_{\text{bin}}$ yield a much better fit to the dolphin data than can be obtained via linear summation or energy detection models.

The binary M-out-of-N receiver can be easily compared with linear summation for small SNR, which corresponds to a large number *N* of available echoes. For zero binary quantization threshold ($\gamma_b = 0$) and for small SNR,

$$p_0 = 1/2 \qquad \text{(A.20)}$$

and

$$\begin{aligned} p_1 = erfc_*[(\gamma_b - A)/\sigma] &= \int_{-A/\sigma}^{\infty} (2\pi)^{-1/2} \exp(-y^2/2)dy \\ &= (1/2) + \int_0^{A/\sigma} (2\pi)^{-1/2} \exp(-y^2/2)dy \\ &\approx p_0 + (A/\sigma)[(d/dx)\int_0^{x} (2\pi)^{-1/2} \exp(-y^2/2)dy]_{x=0} \\ &= p_0 + [A/(\sqrt{2\pi}\sigma)]\,. \end{aligned} \qquad \text{(A.21)}$$

Substituting (A.20) and (A.21) into (A.13) yields

$$d_b \approx \frac{\sqrt{N}A/\sigma}{\sqrt{\pi[(1/2) - (A/\sqrt{2\pi}\sigma)^2]}} \approx (2/\pi)^{1/2} d_l \cdot \qquad \text{(A.22)}$$

At low SNR (large *N*), $d_b \approx 0.8 d_l$. The performance of the binary M-out-of-N receiver is slightly worse than that of the linear summation receiver for a large number of echoes, and a slightly larger SNR should be required for detection. This comparison pertains to additive zero-mean Gaussian noise with known variance (known expected noise power). For non-Gaussian and/or nonstationary noise, the binary M-out-of-N receiver may be superior to linear summation.

For a small number of echoes, performance prediction of the binary M-out-of-N model involves a nonlinear transformation that greatly increases the required SNR. The binary summation data-fit illustrated in Fig. 4 corresponds to p_0=1/2 and c_b=0.9999993 in (A.19). The parameter c_b is written with high accuracy because a small change in c_b is associated with a large change in $erfc_*^{-1}(p_1)$ when p_1 is close to unity, a condition that occurs when p_0=1/2, *N*=1, and $c_b \approx 1.0$ in (A.19). The extremely nonlinear behavior of the inverse complementary error function under these conditions allows the binary M-out-of-N model to closely approximate the large SNR required by the dolphin when only one echo is available (*N*=1).

REFERENCES

[1] W.W.L. Au, *The Sonar of Dolphins.* New York: Springer-Verlag, 1993, pp. 48, 78-79, and 120.

[2] L.W. Gellermann, "Chance orders of alternating stimuli in visual discrimination experiments," J. Gen. Psy., vol. 42, pp. 206-208, 1933.

[3] P.W.B. Moore and R.J. Schusterman, "Audiometric assessment of Northern fur seals, Callorhinus ursinus," Marine Mammal Sci., vol. 3, pp. 31-53, 1987.

[4] W.W.L. Au, and P.W.B. Moore, "Critical ratio and critical bandwidth for the Atlantic Bottlenose dolphin," J. Acoust. Soc. Am., vol. 88, pp. 1635-1638, 1990.

[5] D. Collett, *Modeling Binary Data*, 2nd ed. Boca Raton: Chapman and Hall/CRC, 2003, pp. 56-57.

[6] W.W.L. Au and K. Banks, "The acoustics of the snapping shrimp *Synalpheus parneomeris* in Kaneohe Bay," J. Acoust. Soc. Am., vol. 103, pp. 41-47, 1998.

[7] M. Versluis, B. Schmitz, A. von der Heydt, and D. Lohse, "How snapping shrimp snap: through cavitating bubbles," Science, vol. 289, pp. 2114-2117, 2000.

[8] B.G. Ferguson and J.L. 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., vol. 109, pp. 3031-3027, 2001.

[9] L.A. Dankiewicz, D.A., Helweg, P.W. Moore, and J.M. Zafran, "Discrimination of amplitude-modulated synthetic echo trains by an echolocating bottlenose dolphin," J. Acoust. Soc. Am., vol. 112, pp. 1702-1708, 2002.

[10] R.A. Altes, "Synthetic aperture and image sharpening models for animal sonar," in *Echolocation in Bats and Dolphins*, J. Thomas, C. Moss, and M. Vater, Eds. Chicago: Univ. of Chicago Press, in press.

[11] The Dolphin Biomimetic Sonar (DBS) was built by the SPAWAR Biosonar Program Office (Code D3501) in San Diego, with assistance from the Applied Research Laboratory (ARL), Univ. of Texas at Austin. DBS data were collected by SPAWAR with ARL assistance, using the facilities at the ARL Lake Travis Test Station.

[12] A. Papoulis, *Probability, Random Variables, and Stochastic Processes.* New York: McGraw-Hill, 1965, pp.102 and 158.