Feature Images from Biologically Inspired Acoustic Imaging

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

***Abstract-* A broadband, broad-beam sonar with two receiving arrays has been constructed at the SPAWAR Systems Center in San Diego. The sonar emulates the transmitted pulse, projector beam width, and receiver beam width of the Atlantic bottlenose dolphin. The sonar output can be used to compute a pixel value in an acoustic image by summing echo samples from a chosen point in the environment for various aspect angles. The echo samples usually correspond to reflectivity, e.g., the sampled output of a pulse compression filter. Other quantities, however, can be summed over multiple aspects in the same way. A sum of squared samples, together with the original sum, can be combined to form an image that represents the aspect-induced variance of the echo from each point on an object. This variance is large for smooth, specular surfaces and small for rough surfaces. Another feature image is obtained from a maximum likelihood bearing estimator, using the outputs of the two "ears." The relevant sample from each echo corresponds to the location of the chosen point on a range-bearing plot. Summation of these samples over different aspects produces an interaural correlation image. Rough/smooth feature images and interaural correlation feature images are useful for object detection and classification.**

I. INTRODUCTION

Animal sonar systems typically are characterized by large bandwidths, motion of the transmitter/receiver, and small aperture (array size) relative to many human-made sonar systems. Small arrays and wideband signals imply that cross-range (azimuth and/or elevation) resolution is worse than range resolution. For bats and dolphins, the theoretical disparity between range and cross-range resolution becomes large for ranges in excess of half a meter.

This resolution disparity can be mitigated by regarding a sequence of sonar locations as element positions in a large, synthetic array. The synthetic array can focus on a particular point by delay-and-sum beam forming, i.e., by forming a spatial matched filter for echoes from a chosen scattering point. Sequential delay-and-sum beam forming obviates the need to store echoes, and parallel-distributed processing can be used to update pixels in an environmental map or object image. The process is further simplified by using semi-coherent summation, i..e., coherent pulse compression filtering [1,2] of individual echoes and summation of envelope detected pulse compression filter outputs for multi-echo combining. In the case of the short duration, broadband echolocation clicks of the Atlantic bottlenose dolphin, temporal pulse compression can be approximated by a band pass filter.

Animal sonar systems often utilize sparse spatial sampling relative to synthetic aperture systems. Images that are robust with respect to sparse aspect sampling and small, uncompensated motion perturbations are obtained with a combination of semi-coherent delay-and-sum beam forming and a biomimetic, nonlinear transformation of pulse-compressed filter outputs, as shown in Fig. 1.

Another characteristic of animal sonar is the ability to detect and classify objects in a heavily cluttered environment. Detection and classification capabilities are enhanced if useful features can be represented in acoustic images. Several such "feature images" are discussed here. The biological counterparts of such images are feature-sensitive neurons in a topographic neuronal map.

The acoustic images in Fig. 1 (and all the images in this paper) are from echo data obtained with a sonar system that emulates the bandwidth, transmitted waveforms, and transmitter/receiver beam widths of the Atlantic bottlenose dolphin (*Tursiops truncatus*). The sonar has two receiving arrays with a separation that emulates the dolphin's interaural distance. The dolphin biomimetic sonar (DBS) was constructed and echo data were collected by the Biosonar Program Office (Code D3501) of the SPAWAR Systems Center in San Diego, with the help of the Applied Research Laboratory, University of Texas at Austin.

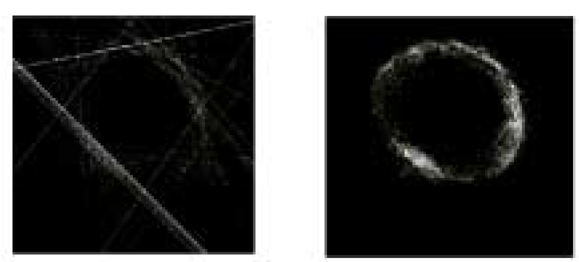

Fig. 1. Left: Conventional SAS image (linear, coherent processing). Right: Biologically inspired acoustic reflectivity image (nonlinear, semi-coherent processing).
In both images, the same rotating object is sampled at 20 deg increments over 360 deg (18 aspects), using a wideband dolphin-like sonar. The object is suspended by thin lines from a rotating platform. The object swings by approximately ± 1 cm (0.6 wavelength at the signal center frequency) relative to the center of rotation. No compensation for this motion is used in constructing the images.

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

0-933957-31-9

II. THE ROUGH-SMOOTH FEATURE IMAGE

At a given aspect angle, the echo sample value corresponding to a desired scattering point is determined by all scattering points that lie within the physical range, cross-range resolution cell of the sonar. The desired scattering point is at the center of this resolution cell, which typically is narrow in range and wide in bearing as shown in Fig. 2A. As the aspect changes, the resolution cell rotates around the desired point, and different scattering points are added to the desired point. Unwanted scattering points are suppressed relative to the desired point by adding relevant samples from pulse-compressed echoes observed at N different aspects, corresponding to differently rotated resolution cells centered on the desired point. The contribution of the desired (center) point to the sum is the center point reflectivity multiplied by N. Each surrounding point ideally contributes only to one term of the sum, since it is contained in only one of the rotated resolution cells. If the surrounding points comprise interference, then the signal-to-interference ratio is proportional to N.

A smooth surface can be modeled by a set of coherent scattering points or Huygens reflectors with specular, mirror-like re-radiation of the acoustic signal in a relatively narrow beam. A rough surface is composed of individual, non-coherent point reflectors that have a wide re-radiation beam width. If a beam of light is focused on a dusty mirror, the specks of dust are visible over a wide set of incidence and reflection angles, while a strong reflection from the mirror itself is observed only when the angles of reflection and incidence are the same. A specular "glint" is observed when many of the smooth-surface point scatterers are contained in a resolution cell, as in Fig. 2B. The measured reflectivity of the center point in Fig. 2B is large at one aspect angle, but is relatively small at other aspects.

A sum of echo samples corresponding to rotated resolution cells is used to estimate the reflectivity of the point at the center of rotation. For a smooth surface, one contribution to the sum is much larger than the others, and

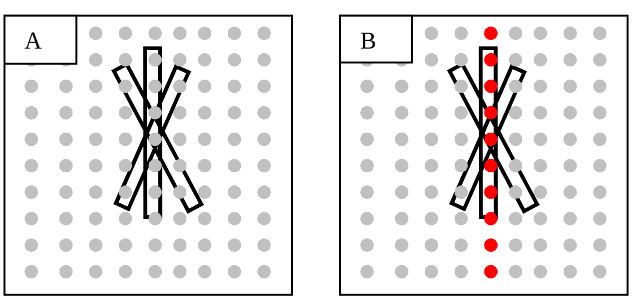

Fig. 2. (A) A set of rotated resolution cells is used to estimate the reflectivity of a scattering point at the center of rotation. (B) A smooth surface results in many scattering points with the same phase, all of which are inside a resolution cell for one rotation angle (one aspect).

the variance of the echo samples that contribute to the sum is relatively large. For a rough surface, the echo samples that contribute to the sum have smaller variance. This variability is observed at the relevant pixel location as the image is sequentially constructed. The aspect-dependent variance of an image pixel can be estimated by summing relevant samples $\{x_i;\ i=1,\ldots,N\}$ from N pulse compressed echoes obtained at different aspects, and by forming another sum of the squared samples $\{x_i^2;\ i=1,\ldots,N\}$. This procedure is equivalent to forming two images, one using envelope detected, pulse-compressed echo samples and the other using the squared values of these samples. The aspect-induced sample variance of a given pixel is then

$$\text{pixel variance} \equiv (1/N)\sum_{i=1}^{N} x_i^2 - [(1/N)\sum_{i=1}^{N} x_i]^2 \,. \qquad (1)$$

If a dolphin is sensitive to aspect-induced pixel variation, then it should be sensitive to the inter-echo amplitude variation of echoes that are identical except for amplitude. This sensitivity has been verified by behavioral experiments [3].

A feature image that accentuates smooth surfaces can be constructed by multiplying a reflectivity image by the aspect-induced standard deviation of each pixel. A corresponding image that accentuates rough surfaces can be obtained by dividing the reflectivity image by the aspect-induced standard deviation of each pixel. A combined rough-smooth feature image is obtained by letting the rough and smooth images compete at each pixel location, and by color-coding the amplitude of the winner (blue for smooth, red for rough).

The rough-smooth feature image can solve challenging search problems that are analogous to finding a smooth needle in a rough haystack. Fig. 3 shows a rough-smooth biologically inspired acoustic image of a 10m×10m area on the bottom San Diego Bay that contains a smooth, buried, bullet-shaped metal object approximately 1 m long. The object was 40 m from the center of a linear path that was followed by the biomimetic sonar (DBS). The sonar obtained wideband echoes at a sequence of positions along the path. At each position, the sonar was scanned in bearing. The total aspect change along the path was approximately 90 deg.

The rough-smooth feature image in Fig. 3 shows a blue patch (evidence of a smooth scattering surface) at the object location, surrounded by a large red (relatively rough) area of the bottom. The blue area is indicative of the size and orientation of the object.

High aspect-dependence also occurs for a scattering point that lies on an edge. Edges thus can be emphasized with a rough-smooth feature image, as shown in Fig. 4. The object in Fig. 4 is a truncated cone resting on a textured turntable. DBS echoes were measured at 3.6 deg intervals over 360 deg, at a grazing angle of 3.5 deg and a range of 22.9 m. The turntable diameter was 6.1 m.

III. THE INTERAURAL CORRELATION IMAGE

Another type of feature image is obtained by replacing the reflectivity estimate for a given point in the environment by the output of an adaptive maximum-likelihood bearing estimator. The inputs to the bearing estimator are the pulse-compressed echoes from the two receiving arrays of the Dolphin Biomimetic Sonar, which are separated by 12 cm. The bearing estimator output is the log-likelihood that the received echo within a given range resolution cell and physical transmitter/receiver beam width is from a hypothesized bearing or azimuth angle. The location of an imaged point determines the delay hypothesis τ and bearing hypothesis θ of a sample point in a range-bearing map at the estimator output.

For a binaural receiver, the log-likelihood ratio for a given delay hypothesis τ and bearing hypotheses θ is [4]

$$\ell(r_1,r_2|\tau,\theta) = w(\tau,\theta)\{r_1[t-\tau+\Delta(\theta)]+r_2[t-\tau-\Delta(\theta)]\}^2 \tag{2}$$

where

$$w(\tau,\theta) = \frac{\sigma_S^2(\tau,\theta)}{\sigma_N^2(\tau,\theta)[\sigma_N^2(\tau,\theta)+2\sigma_S^2(\tau,\theta)]}. \tag{3}$$

In (2), $r_1(t)$ and $r_2(t)$ are pulse-compressed echoes plus noise, which is assumed to be white and Gaussian. The delay τ corresponds to the hypothesized target range. The delays $\pm\Delta(\theta)$ have much smaller magnitude than τ, and they correspond to the hypothesized bearing angle θ. In (3) the signal and noise power estimates $\sigma_S^2(\tau,\theta)$ and $\sigma_N^2(\tau,\theta)$ depend upon the hypothesized range and bearing, and are computed using $r_1(t)$ and $r_2(t)$.

The data dependence of the signal and noise power estimates causes the bearing estimator to be adaptive. The bearing dependence of these power estimates introduces additional angle dependence into the log-likelihood expression. This extra angle dependence causes the estimator to have a smaller predicted error as calculated by the Cramer-Rao bound. The estimator output has a narrow peak as a function of hypothesized bearing, and the acoustic imaging system can resolve target points that are within the usual spotlight in spotlight-mode SAS. This characteristic distinguishes biologically inspired acoustic imaging from conventional spotlight-mode SAS. The extra bearing resolution provided by the adaptive bearing estimator results in less mutual interference between target points, and aspect sampling intervals can be increased.

The interaural correlation image is computed by forming a range-bearing map from the output of the adaptive bearing estimator. A different range-bearing map is formed for each echo, i.e., for each location of the physical sonar. A given point in the environment occupies a specific position in each range-bearing map. At each aspect (physical sonar location), the range-bearing sample corresponding to the position of the given point is extracted and added to corresponding samples from the range-bearing maps at other sonar locations. The resulting sum is a pixel value in the interaural correlation feature image.

Examples of interaural correlation images and corresponding reflectivity images are shown in Figs. 5-10. These images are from DBS turntable data at a grazing angle of 3.5 deg as in Fig. 4. Aspect sampling was at 3.6 deg intervals for reflectivity images and 7.2 deg intervals for interaural correlation images.

Fig. 5 shows reflectivity and interaural correlation images of an object that appears to be trapezoidal when seen from above, with a rounded shape that resembles a clamshell. The images are amplitude normalized and are colorized, with blue denoting small pixel values and red denoting large values. The object appears as a trapezoidal “hole” in the surrounding boundary reflections. The trapezoidal clamshell has very small adaptive bearing estimator responses, but the responses are larger at the boundary between the object and the textured surface.

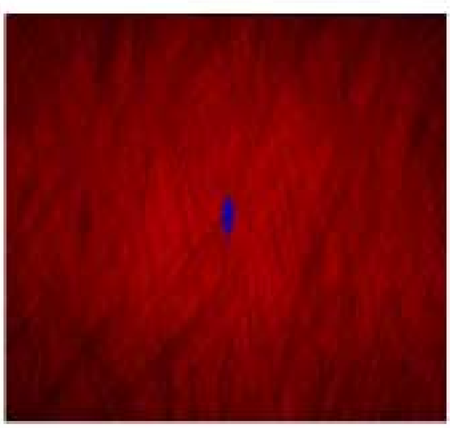

Fig. 3. Rough-smooth feature image of a 10 × 10 m area of a relatively rough bottom containing a smooth, buried, bullet-shaped metal object. The object appears as a blue patch at the center of the image.

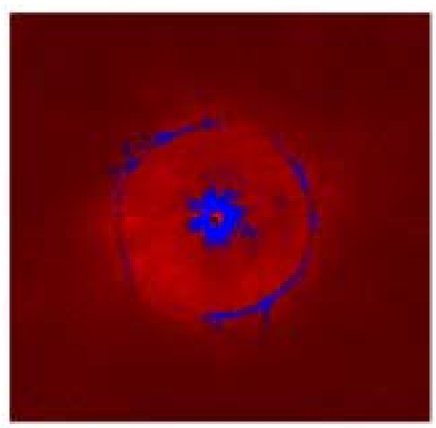

Fig. 4. Rough-smooth feature image of a truncated cone (diameter ~ 1 m) resting on a textured turntable. Points with high aspect variability are colored blue.

Fig. 6 shows reflectivity and interaural correlation images of a truncated cone. The interaural correlation image increases the signal-to-clutter ratio of the object.

Fig. 7 shows reflectivity and interaural correlation images of a buried bullet-shaped metal object (the same object as in Fig. 3, but buried in the simulated bottom on the turntable, rather than in the bottom of San Diego Bay). In this case, the interaural correlation image is far superior to the reflectivity image in both object shape representation and signal-to-clutter ratio.

Fig. 8 shows reflectivity and interaural correlation images of a water-filled drum (a large cylinder with thin metal walls). The interaural correlation image represents the drum's shape as a "hole" in clutter, as in Fig. 5.

Fig. 9 shows reflectivity and interaural correlation images of a culvert (a large, open-ended concrete cylinder). The interaural correlation image appears to be accentuating the boundaries or edges of the object.

Fig. 10 shows reflectivity and interaural correlation images of a waterlogged telephone pole. The reflectivity of the pole is nearly the same as that of the simulated, textured bottom. The adaptive bearing estimator output, however, produces a pole image with high signal-to-clutter ratio.

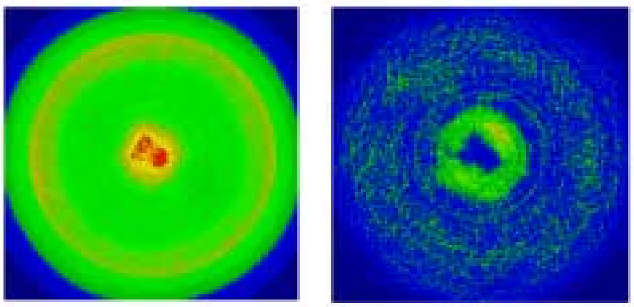

Fig. 5. Reflectivity image (left) and interaural correlation image (right) of a clamshell-shaped object resting on a textured turntable. Image width =7.7 m.

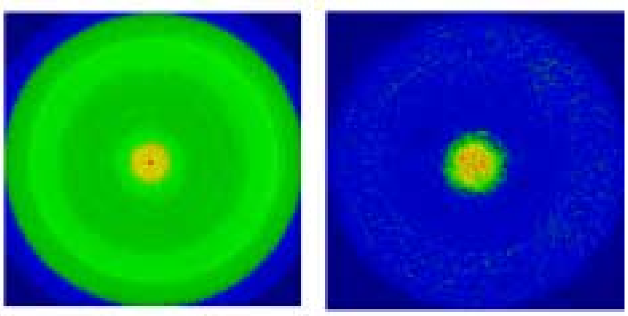

Fig. 6. Reflectivity image (left) and interaural correlation image (right) of a truncated cone resting on a textured turntable. Image width = 7.7 m.

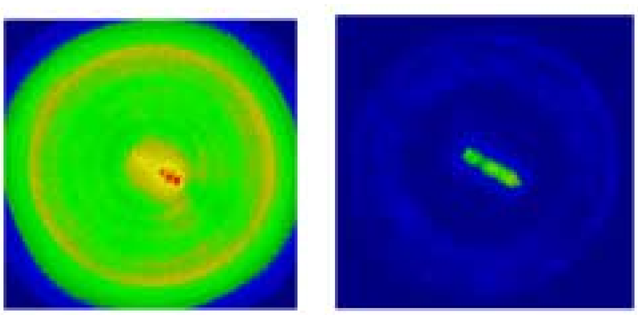

Fig. 7. Reflectivity image (left) and interaural correlation image (right) of a bullet-shaped metal object resting on a textured turntable. Image width = 7.7 m.

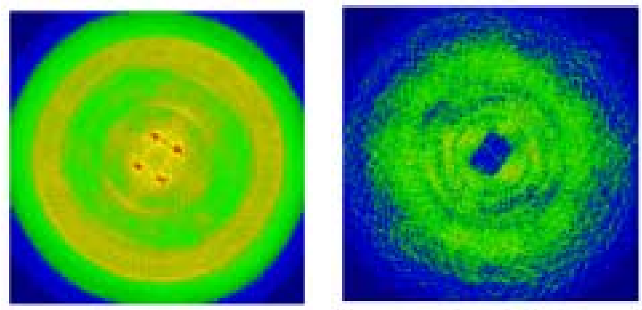

Fig. 8. Reflectivity image (left) and interaural correlation image (right) of a 55 gal water-filled drum resting on a textured turntable. Image width = 7.7 m.

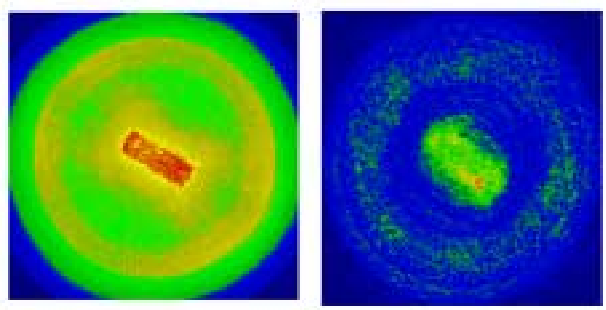

Fig. 9. Reflectivity image (left) and interaural correlation image (right) of a culvert (a large, open-ended concrete cylinder), resting on a textured turntable. Image width = 7.7 m.

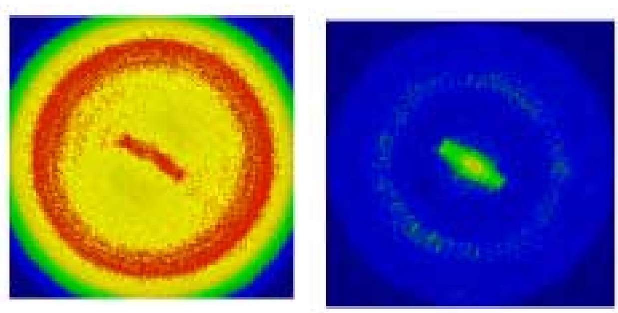

Fig. 10. Reflectivity image (left) and interaural correlation image (right) of a waterlogged wooden telephone pole, resting on a textured turntable. Image width = 7.7 m.

The reflectivity and interaural correlation images of the different objects can be used to represent the objects in a feature space, as shown in Fig. 11. Binary-quantized versions of reflectivity and interaural correlation values relative to the textured turntable determine object locations in quadrants of the two-dimensional feature space. Objects can be further discriminated by adding extra dimensions corresponding to shape and size features.

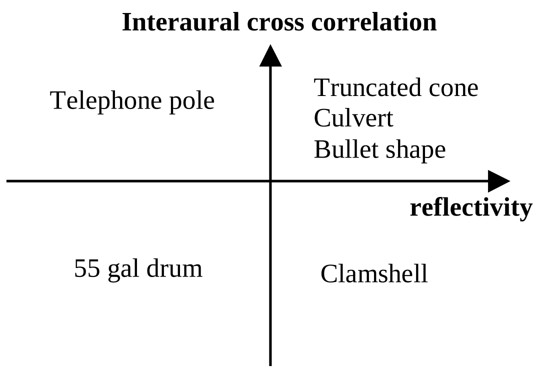

Fig. 11. Positions of objects in two-dimensional feature space, based on Figs. 5 - 10.

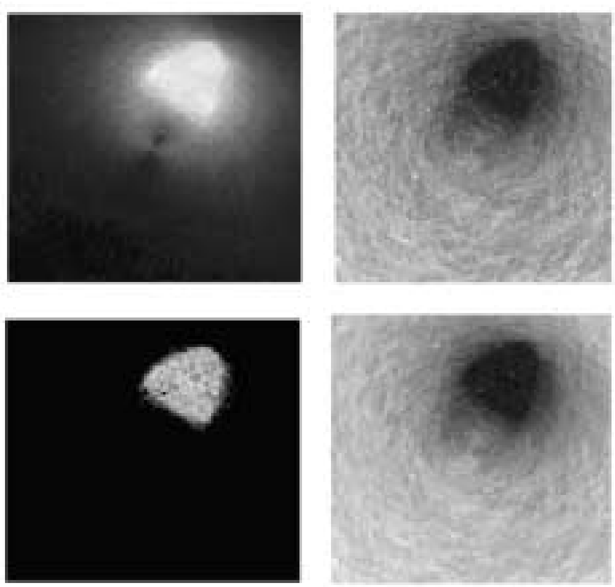

Fig. 12. Top left: Reflectivity image of clamshell-shaped object suspended in volume clutter.
Top right: Volume clutter feature image.
Bottom left: Reflectivity image – (weight × feature image).
Bottom right: Feature image – (weight × reflectivity image).

IV. THE VOLUME CLUTTER FEATURE IMAGE

Figs. 5 and 8 demonstrate that objects can be recognized by the shapes of "holes" in clutter. This technique is well known in medical ultrasound, where volume clutter or "speckle" often is used to form images of organs that have relatively low reflectivity. If an object of interest is best described as the shape of a hole in clutter, then it may be advantageous to design a feature image that represents the clutter distribution and suppresses the target.

A volume clutter feature image [5,6] utilizes the observation that volume clutter usually is composed of many scatterers that are small relative to the average signal wavelength (~1.6 cm for the Atlantic bottlenose dolphin). The feature image accentuates small volume scatterers such as bubbles by using a normalized peak sharpness measure. At a given range value R, the algorithm finds the largest echo peak in a small interval around R. The amplitude of this peak is normalized to unity, and the feature image represents the echo at R by peak sharpness (the negative second derivative of the pulse compression filter output at the peak). Normalization of peak amplitude to unity eliminates reflectivity dependence; tiny scatterers appear to be as strong as large targets. The second derivative measure is equivalent to an ideal receiver for detection of Rayleigh scatterers, i.e., objects that are small relative to a wavelength. The sonar cross section of a Rayleigh scatterer varies as the fourth power of frequency, and reflectivity varies as frequency squared. A frequency-squared filter is a double differentiator, multiplied by -1.

An example of a volume clutter feature image and one of its applications is shown in Fig. 12. The clamshell target is suspended by fine lines and rotated relative to the DBS sonar. The clamshell is partially obscured by volume clutter in the reflectivity image (Fig. 12, top left). The corresponding volume clutter (peak sharpness) feature image (Fig. 12, top right) shows a target-shaped hole in clutter. The signal-to-clutter ratio of the conventional reflectivity image is increased by subtracting a weighted volume clutter feature image from the reflectivity image (Fig. 12, bottom left). Subtracting a weighted reflectivity image from the volume clutter feature image yields a better hole-in-clutter image (Fig. 12, bottom, right).

V. FEATURE IMAGES AND THE SONAR EQUATION

The sonar equation is a logarithmic form of the signal-to-interference ratio (SIR) at the output of a detector [7]. The equation implies that if the SIR is not larger than a reasonable threshold value, then an object cannot be reliably detected. Hole-in-clutter representations, however, demonstrate that objects can be detected and classified at very small SIR (negative dB values).

The problem is that the clutter echo usually is modeled as an additive random process that is independent of the desired signal or object echo. Feature images indicate that this clutter model is not always correct in practice. If an object displaces the clutter, then a drop in the clutter power increases the probability that an object is present. Feature images permit an observer to analyze whether a local decrease in clutter power signifies the presence of an object with a particular shape.

If interference can be estimated, then it can be subtracted from the data. A clutter feature image can be used to estimate a spatial representation of the interference, and this estimate can be subtracted from the reflectivity image as in Fig. 12 (bottom, left). The weighted subtraction operation in Fig. 12 is a form of adaptive interference cancellation [8], where the weight is adaptively determined to maximize SIR. The similarity to adaptive interference cancellation suggests that a further improvement of the object image in Fig. 12 can be obtained by replacing the single weight in the operation

final image = reflectivity image – (weight × feature image)

with an adaptive spatial filter, i..e., a weighted sum of pixels from the clutter feature image.

VI. BIOLOGICAL INSPIRATION AND IMPLICATIONS

Acoustic feature images can supplement reflectivity images for object detection and classification. Such images have counterparts in biological sensory systems. Electric fish, for example, have overlaid, spatially registered neuronal maps from visual and electrosensory receptors in the optic tectum. Some of the neurons in these maps are sensitive to movement. Neurons with inputs from a corresponding location in different sensory maps excite other neurons in a map that controls fish movement [9]. Neurons in the control map combine visual and electrosensory stimuli to locate, detect, and classify objects in the environment.

The rattlesnake optic tectum has corresponding maps for vision and infrared reception [10]. In the owl, the maps are derived from vision and hearing [11]. The mouse superior colliculus (the mammalian counterpart of the optic tectum) has overlaid maps derived from vision, hearing, and somatosensation [12]. In all cases, neurons derived from various sensory modalities and with specific movement sensitivity are adaptively combined (linearly and nonlinearly) to form detector/classifiers at environmental locations corresponding to positions on the maps.

The biological examples suggest a straightforward architecture for object detection, classification, and localization using overlaid, spatially registered versions of acoustic reflectivity images, feature images, and maps derived from non-acoustic sensors.

VII. CONCLUSION

Acoustic images can contain information that is not directly associated with reflectivity (the usual parameter that is represented in synthetic aperture imagery). Feature images that represent such extra information can improve object detection and classification in clutter. Some feature images imply that the additive, signal-independent clutter model used by the sonar equation is not always applicable to high-resolution analysis. Spatially registered neuronal maps in biological sensory systems suggest an architecture for combining acoustic feature images with reflectivity images and inputs from other sensors.

VIII. ACKNOWLEDGMENTS

DBS design, data collection, and data organization involved the efforts of Patrick Moore, Steve Martin, Mike Phillips, Robert Floyd, Eric Bauer, Tim Herrin, David Helweg, and Dorian Hauser. Echo data were supplied by the SPAWAR Biosonar Program Office, Code D3501, San Diego. Support was provided by Drs. John Tague and Robert Gisiner (Office of Naval Research 321US and 342), and by Dr. Allen Rudolph of the Defense Sciences Office of the Defense Research Projects Agency.

VIII. REFERENCES

[1] G.L. Turin, “On the estimation in the presence of noise of the impulse response of a random, linear filter,” IRE Trans. Inf. Theory, vol. IT-3, pp. 5-10, 1957.

[2] R.A. Altes, “Estimation of sonar target transfer functions in the presence of clutter and noise,” J. Acoust. Soc. Am., vol. 61, pp. 1371-1374, 1977.

[3] 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.

[4] R.A. Altes, "Cross correlation and energy detection in multiarray processing," IEEE Trans. Acoust., Speech, Sig Proc., vol. ASSP-33, pp. 493-504, 1985.

[5] R.A. Altes, “Feature imaging and adaptive focusing for synthetic aperture processor,” U.S. Patent No. 6,088,295, July 11, 2000.

[6] 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.

[7] R.J. Urick, *Principles of Underwater Sound*, 2nd ed. New York: McGraw-Hill, 1975.

[8] B. Widrow and S.D. Stearns, *Adaptive Signal Processing*. Englewood Cliffs: Prentice-Hall, 1985, pp. 303-306.

[9] C.E. Carr and L. Maler, “Electroreception in gymnotiform fish,” in *Electroreception*, T.H. Bullock and W. Heiligenberg, Eds. New York: Wiley, 1986, pp. 319-373.

[10] P.H. Hartline, L. Kass, and M.S. Loop, “Merging of modalities in the optic tectum: Infrared and visual integration in rattlesnakes,” Science, vol. 199, pp. 1225-1229, 1978.

[11] E.I. Knudsen, “Auditory and visual maps of space in the optic tectum of the owl,” J. Neurosci., vol. 2, pp. 1177-1194, 1982.

[12] U.C. Drager and D.H. Hubel, “Responses to visual stimulation and relationship between visual, auditory, and somatosensory inputs in the mouse superior colliculus,” J. Neurophysiol., vol. 38, pp. 690-713, 1975.