

Multiple angle acoustic classification of zooplankton

Paul L. D. Roberts^{a)} and Jules S. Jaffe

Marine Physical Laboratory, Scripps Institution of Oceanography, University of California San Diego,
La Jolla, California 92093-0238

(Received 8 August 2006; revised 23 January 2007; accepted 24 January 2007)

The use of multiple angle acoustic scatter to discriminate between two taxa of fluid-like zooplankton, copepods and euphausiids, is explored. Using computer modeling, feature extraction, and subsequent classification, the accuracy in discriminating between the two taxa is characterized via computer simulations. The model applies the distorted wave Born approximation together with a simple system geometry, a linear array, to predict a set of noisy training and test data. Three feature spaces are designed, exploiting the relationship between the shape of the scatterer and angularly varying scattering amplitude, to extract discriminant features from these data. Under the assumption of uniform random length and uniform three-dimensional orientation distributions for each class of scatterers, the performance of several classification algorithms is evaluated. Simulations reveal that the incorporation of multiple angle data leads to a marked improvement in classification performance over single angle methods. The improvement is more substantial using broadband scatter. The simulations indicate that under the stated assumptions, a low classification error can be obtained. The use of multiple angle scatter therefore holds promise to substantially improve the *in situ* acoustic classification of fluid-like zooplankton using simple observation geometries. © 2007 Acoustical Society of America. [DOI: 10.1121/1.2697471]

PACS number(s): 43.60.Np, 43.30.Sf, 43.60.Fg [EJS]

Pages: 2060–2070

I. INTRODUCTION

Zooplankton play a major role in the global ecosystem and the employment of remote sensing techniques for measuring abundance and behavior continues to be a venerable goal. Compared with optically based methods, acoustic ones have an inherent advantage in that sound is attenuated less than light. This leads to both larger detection distances as well as sampling volumes. Unfortunately, problems associated with a lack of specificity have hindered the use of acoustic techniques on a routine basis. Work by McNaught¹ and the subsequent development of multiple frequency methods by Holliday and colleagues have revealed both the great advantages and also challenges that exist when using this technique.^{2–6} Although additional work with broadband sound to discriminate between three groups of zooplankton^{7,8} indicated that it was possible to correctly classify specific examples from each group with reasonable success (80% overall average correct classification), the goal of robustly mapping acoustic volume scattering to biophysical parameters of zooplankton under various oceanographic conditions has remained elusive. One problem has been the confounding influence of both orientation and material properties on backscatter magnitude. This often prevents investigators from making the necessary link between animal size and backscatter magnitude. Additional complications have been that there is substantial scatter from nonbiological sources such as suspended sand, bubbles, and perhaps even

microstructure.⁹ If a way could be found to discriminate among various taxa acoustically, in spite of these problems, it would be of great value.

In this article the potential increase in classification accuracy that results from observation of reflected sound from multiple angles is considered. It has recently been proposed that sound scattered at multiple angles can be used in order to both size and measure the orientation of fish bladders.¹⁰ The underlying concept is that the spatial structure of the sound field from a single, strong scattering target, has a characteristic pattern related to its size. The success of the method was illustrated with a well-known data set¹¹ and various sampling theorems were proved to obtain unaliased sampling of the scattered sound field. Here, the use of sound scattered at multiple angles in order to discriminate among two zooplankton taxa is explored via forward modeling and subsequent classification.

Many have considered the formulation of acoustic models to predict backscatter as an important component of a program to characterize animals *in situ*. A family of scattering models can be successfully used to predict the acoustic reflectivity of several different types of zooplankton.^{12–15} The situation with respect to crustacean zooplankton is especially good as use of the distorted wave Born approximation (DWBA) has been validated.^{16,17} A website maintained by Benfield¹⁸ provides public access to several zooplankton models and their morphologies.

Outside of the realm of ocean ecology, recent work in the acoustic classification of stationary targets from multiple views has demonstrated that the multiple views can significantly improve target classification when combined with suitable feature extraction and classification algorithms.^{19–22} Applications such as underwater mine detection,^{20,22} airborne

^{a)}Also at: Electrical and Computer Engineering Department, U.C.S.D., La Jolla, California 92093; electronic mail: paulr@mpl.ucsd.edu

target identification,^{23,24} and unexploded ordnance detection²⁵ have been considered. Most algorithms apply a hidden Markov model (HMM) to account for either the unknown sensor-target aspect,^{20,21,26} or the unknown target type.²² One approach decomposes the target reflections into a set of discrete angular regions yielding a set of possible states in the HMM.²⁶ Alternatively, a nonlinear backpropagated neural network has been used to fuse the classification results for multiple views in a wavelet packet based feature space.¹⁹ This formulation demonstrated very good performance in discriminating between mine and nonmine like targets from multiple aspect scattering measurements.

Adaptation of multiple angle scatter techniques to zooplankton classification has promise to confer benefits when used in conjunction with the more traditional backscatter techniques. However, animals are dynamic and therefore require an observation system in which multiple views are obtained almost simultaneously. One solution is to use simultaneous multiple angle scatter measurements. A second issue is related to the feature space used to represent the data. Previous work in target classification considered rigid objects and therefore applied wavelet packets^{19,22,27} or matching pursuit with an elastic scattering based dictionary.^{20,26} However, the resulting feature spaces are not appropriate for the fluid-like weak scatterers considered here. A more appropriate idea for this problem is to exploit the relationship between the shape of the scatterer and the angularly varying scatter amplitude.

This paper explores, through simulation, the use of a one-dimensional array to collect multiple angle scatter and subsequently use these data to discriminate among zooplankton taxa. The case treated is that of differentiating between two taxa of crustacean zooplankton: copepods and euphausiids. The motivation for treating these animals stems from their significance in zooplankton populations of the California Current. As shown here, the large morphological difference between the two groups²⁸ will allow this discrimination.

Section II summarizes the theoretical basis^{14,17} for forward model computations which are used to generate the synthetic data. Section III defines the feature spaces that are applied to reduce the raw data to a small set of discriminant parameters. Section IV describes the nearest neighbor and multilayer perceptron classification algorithms that have been used for classification. Section V discusses the performance of the classifiers as a function of the various parameters that are available. Section VI summarizes the results and identifies future research areas.

II. FORWARD MODELING: THEORY AND NUMERICAL IMPLEMENTATION

In this section, an acoustic forward model for generating synthetic data is proposed using linear system theory and the DWBA. The forward model permits the prediction of the received signal for a known transmit signal using the impulse response of the scatterer. This depends on both the physical properties of the scatterer such as size, shape, and material and also the orientation of the scatterer and the geometry of transmitters and receivers. Under the assumption of linearity, and neglecting effects of spreading and medium attenuation,

the received signal $p(t)$ is given by the convolution of the transmitted signal $s_0(t)$ with the impulse response of the scatterer $s(t, \mathbf{k}_i, \mathbf{k}_s, \theta, \phi, \mathbf{\Gamma})$,

$$p(t) = \int_{-\infty}^{\infty} s_0(\tau) * s(t - \tau, \mathbf{k}_i, \mathbf{k}_s, \theta, \phi, \mathbf{\Gamma}) d\tau, \quad (1)$$

where \mathbf{k}_i and \mathbf{k}_s are the incident and scattered wave vectors, θ and ϕ define the orientation of the scatterer, and $\mathbf{\Gamma}$ is a parameter matrix describing the size, shape, and material properties. Assuming values for these parameters permits the prediction of the impulse response of the scatterer using the DWBA. This model does not include propagation effects, however the effect of scatterer position in the beam is included. A description of the multiple angle DWBA is given in Sec. II A. Section II B defines the size and orientation distributions that are used to generate synthetic data. Section II C describes the procedure for generating synthetic data.

A. Multiple angle DWBA scattering model

The scattering model used to obtain the impulse response of the scatterer is the DWBA.^{14,17} This model relates the size, shape, and material properties of the scatterer to the complex scattering amplitude at a particular frequency. The impulse response of the scatterer can be obtained from the complex scattering amplitude by an inverse Fourier transform. The expression for the complex scattering amplitude $S(\mathbf{k})$ is

$$S(\mathbf{k}) = \frac{k_1^2}{4\pi} \int \int_R (\gamma_\kappa(\mathbf{r}_0) - \gamma_\rho(\mathbf{r}_0) \cos \alpha) e^{i\mathbf{k} \cdot \mathbf{r}_0} dV_0, \quad (2)$$

where

$$\mathbf{k} = k_2(\mathbf{e}_s - \mathbf{e}_i), \quad (3)$$

and

$$\cos \alpha = \mathbf{e}_s \cdot (-\mathbf{e}_i). \quad (4)$$

The scalars $k_1 = 2\pi f/c_1$ and $k_2 = 2\pi f/c_2$ are the wave numbers in the medium and body of the scatterer, respectively, \mathbf{e}_i and \mathbf{e}_s are unit vectors in the direction of the incident and scattered sound waves, and α is the angle between the negative incident wave vector and the scattered wave vector. The term $\gamma_\kappa(\mathbf{r}_0) - \gamma_\rho(\mathbf{r}_0) \cos \alpha$ is the gamma contrast.^{17,29}

The gamma contrast inside of the volume integral is a function of the density and sound speed of the surrounding medium and the body of the scatterer where (omitting the explicit dependence on position in the body) $\gamma_\kappa = (1 - gh^2)/gh^2$ and $\gamma_\rho = (g - 1)/g$. The term $g = \rho_2/\rho_1$ is the ratio of the density of the scatterer to the density of the surrounding medium and $h = c_2/c_1$ is the ratio of sound speed in the scatterer to the sound speed in the surrounding medium. Equation (2) provides the basis for the forward model used in the numerical experiments presented in this paper.



FIG. 1. Cross-sectional view of the scatterer shapes used to define each class. The copepod shape (a) and euphausiid shape (b) are displayed at different scales. The scale is defined in the lower right corner.

Using Eq. (3), the resultant wave vector \mathbf{k} can be written as

$$\mathbf{k} = k_2 \frac{\sin(\pi - \alpha)}{\sin(\alpha/2)} \mathbf{d}, \quad (5)$$

where

$$\mathbf{d} = \frac{\mathbf{e}_s - \mathbf{e}_i}{\|\mathbf{e}_s - \mathbf{e}_i\|_2}, \quad (6)$$

is the unit vector that points in the direction of the difference between scattered and incident wave vectors. It is apparent that the multiple angle DWBA is closely related to the DWBA for backscatter¹⁷ only now with a scaled and rotated wave vector. This important relationship allows the DWBA for multiple angle scatter to be computed using existing numerical methods for backscatter with only minimal modification.

B. Scatterer size and orientation distributions

An important aspect of the simulations is the choice of size and orientation distributions for the ensemble of scatterers. These data have been generated using a single shape for each class, scaled in volume and rotated in three dimensions. The shapes used for the copepod and euphausiid classes were taken from an online database of zooplankton scattering models.¹⁸ Pictures of the cross section of the base shape used for each class are shown in Fig. 1.

The volume scaling is parametrized by a length parameter L , the length of the scatterer from head to tail. In order to simplify the treatment both length classes were drawn from uniform distributions according to $\mathcal{U}[2 \text{ mm}, 4 \text{ mm}]$ for the copepods and $\mathcal{U}[4 \text{ mm}, 15 \text{ mm}]$ for the euphausiids. Note that the length distributions overlap slightly and the distribution for the euphausiids is centered around medium length juveniles rather than the larger adults.

Similarly, for the orientation distributions, a simple approach was taken. Representing the orientation of the scatterer by a θ and ϕ angle where θ is the angle relative to the z axis, and ϕ the angle between the x and y axis (Fig. 2) the orientations were sampled uniformly in three dimensions according to

$$\theta \sim \arcsin(\mathcal{U}[-1, 1]), \quad (7)$$

$$\phi \sim \mathcal{U}[-\pi, \pi]. \quad (8)$$

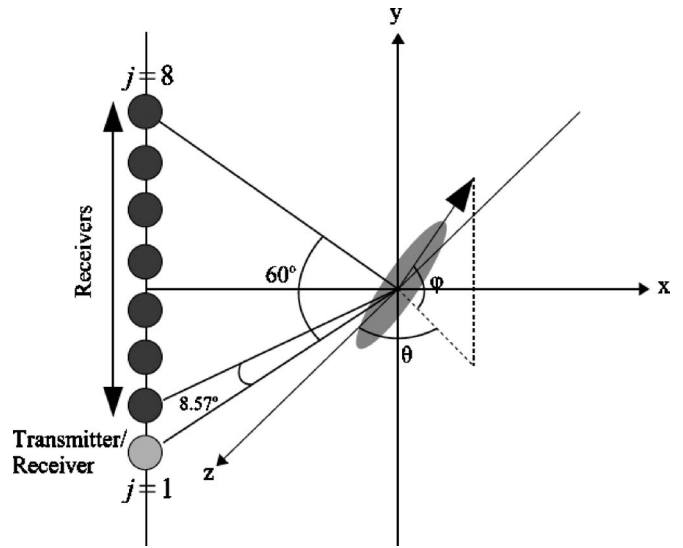


FIG. 2. View of the array configuration used to generate data. The scatterer is represented by the gray ellipse where the nose of the scatterer is directed along the unit vector defined by the angle ϕ in the x - y plane and θ from the z axis. The angular span of the array is 60° with an angular spacing of 8.57° between elements. The $j=1$ element functions as both transmitter and receiver, while the $j=2, \dots, 8$ elements receive only.

C. Creation of model realizations

The simulation of a single realization of received scatter on the array is described in this section. The configuration of the array is shown in Fig. 2. There are eight total elements, $M=8$. The $j=1$ element acts as both transmitter and receiver while the $j=2, \dots, M$ elements act only as receivers. The total angular span of the array is 60° with an angular sampling frequency of one sample per 8.57° . The orientation of the scatterer relative to the array is shown in Fig. 2 and is defined by the angles ϕ and θ as mentioned previously.

The synthetic data are generated by predicting the received pressure signal on each of the eight array elements for a given scatterer orientation. The data generation process is represented graphically in Fig. 3. The first step is the selection of a model: copepod or euphausiid. Each three-dimensional scatterer shape is represented as a series of cylindrical segments of thickness 0.016 mm and location x , y , and z corresponding to the center. The segments have radius a , and relative density and sound speed g and h . The i th segment can be represented as the vector $\boldsymbol{\gamma}_i = (x_i, y_i, z_i, a_i, g_i, h_i)^T$ and the entire model of the body by a matrix $\boldsymbol{\Gamma} = (\boldsymbol{\gamma}_1, \dots, \boldsymbol{\gamma}_S)$ with S being the total number of segments. Ambient sound speed c is fixed at 1500 m/s . For all simulations performed in this work, the values of g_i and h_i are held constant throughout the body of the scatterer such that $g_i = 1.035 \forall i$ and $h_i = 1.027 \forall i$. More information about these parameters, the scattering models, and the algorithm used to compute the DWBA is available from the Acoustic Scattering Models of Zooplankton website.¹⁸

For each realization, a random sample from the distributions for ϕ, θ, L is selected. These parameters are combined with the sound speed c , the incident and scattered wave vectors for each array element: \mathbf{k}_i^j (incident) and \mathbf{k}_s^j (scattered) for $j=1, \dots, M$, and the model for the scatterer $\boldsymbol{\Gamma}$. The

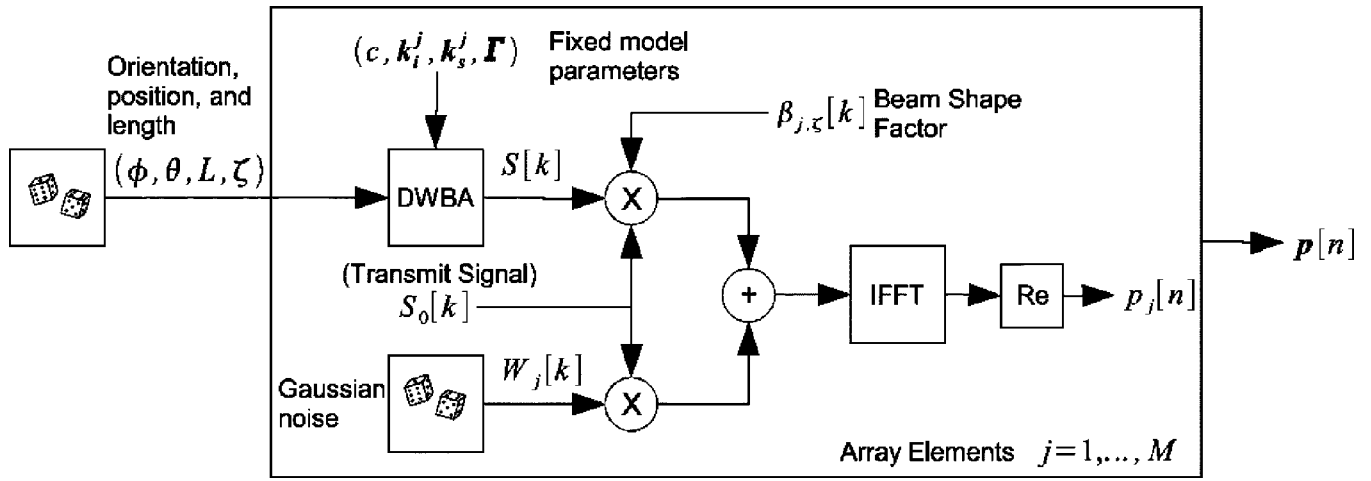


FIG. 3. Block diagram of the creation of a single realization of scattering on the array. Randomness is included in the model through the parameters ϕ, θ, L, ζ , and the Gaussian noise $W[k]$. The model parameters which are constant for all realizations are the sound speed c , the incident and scattered wave vectors for each array element \mathbf{k}_i^j and \mathbf{k}_s^j for $j=1, \dots, M$ and the scatterer model Γ . The beam shape factor $\beta_{j,\zeta}[k]$ is computed for each array element and accounts for the position of the scatterer in the transmit and receive beam pair. The Gaussian noise $W[k]$ is added to the product of the scattering amplitude $S[k]$ and the beam shape factor $\beta_{j,\zeta}[k]$. The result is then multiplied by the FFT transmit signal $S_0[k]$. The real part of the inverse FFT of the product yields the pressure on the j th array element. The pressure is computed for each of the M array elements. This process is repeated 1000 times for each scatterer class to generate a set of training and test data.

DWBA is evaluated for the given model parameters at each frequency yielding the complex scattering amplitude

$$S_j[k] = \text{DWBA}(\phi, \theta, L, c, \mathbf{k}_i^j, \mathbf{k}_s^j, \Gamma), \quad (9)$$

where k represents the index of a particular wave number bin.

To incorporate the effect of a range-dependent sample volume in the simulation, for each realization, the scatterer is assigned a uniformly random three-dimensional position relative to the array. The position is defined by the parameter ζ .

The range-dependent sample volume is studied by calculating the position of the scatterer in the transmit and receive beam. As the beam shape changes with frequency, the incident sound intensity, and received sound intensity will vary in a predictable way. For the simulations considered here, the transducers are assumed to be disk shaped in which case the product of the transmit and receive beam shapes is given by

$$\beta_{j,\zeta}[k] = \left| \frac{2J_1(kr \sin(\eta))}{kr \sin(\eta)} \frac{2J_1(kr \sin(\mu))}{kr \sin(\mu)} \right|, \quad (10)$$

where r is the radius of the transducer, $J_1(x)$ is the Bessel function of the first kind of order 1, and η and μ are the angles between the vector from the transducer to the scatterer position, and the incident and scattered wave vectors, respectively, for a particular transmitter-receiver pair. For a wave vector \mathbf{k} and scatterer position ζ , the angles are given by

$$\eta = \arccos\left(\frac{\mathbf{k}_i^T \mathbf{k}_i - \mathbf{k}_i^T \zeta}{\|\mathbf{k}_i\|_2 \|\mathbf{k}_i - \zeta\|_2}\right), \quad (11)$$

and

$$\mu = \arccos\left(\frac{\mathbf{k}_s^T \mathbf{k}_s - \mathbf{k}_s^T \zeta}{\|\mathbf{k}_s\|_2 \|\mathbf{k}_s - \zeta\|_2}\right). \quad (12)$$

The geometry for the above-presented calculations is shown in Fig. 4. For all of the simulations, the parameter r is set to 12 mm, and the components of ζ selected according to $\mathcal{U}[-5 \text{ mm}, 5 \text{ mm}]$. The horizontal distance from $\zeta=0$ to the array is defined to be 3 m, and thus the majority of scatterer positions are within the -6 dB beam width of the array elements at the highest frequency. To compare the effect of the sample volume on the classification performance, the simulations are performed with and without including the beam shape factor. The case without the beam shape factor is equivalent to setting $\beta_{j,\zeta}[k]=1$ for all realizations.

The received echoes in any practical system will be corrupted by noise due to reverberation, electronics, and other sound sources. Noise due to reverberation will be in the same frequency band as the received echo whereas noise from electronics and other sound sources will have energy in other frequency bands as well as the band of the received echo. Since out of band noise can be reduced by filtering, rever-

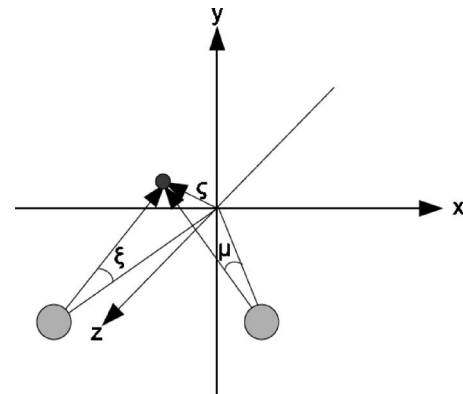
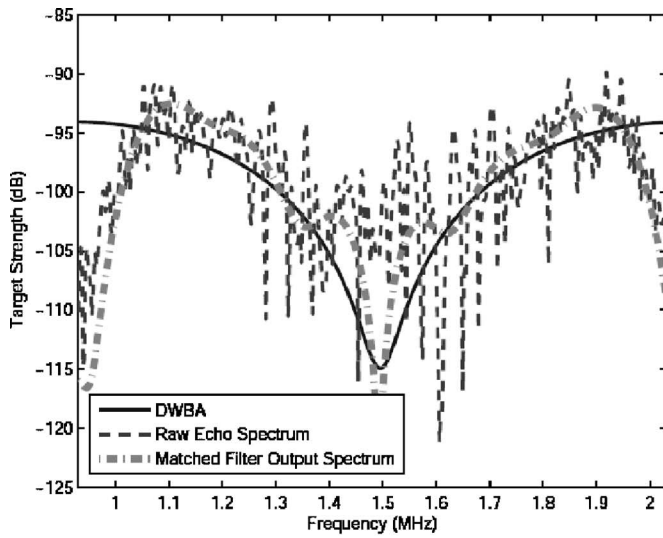
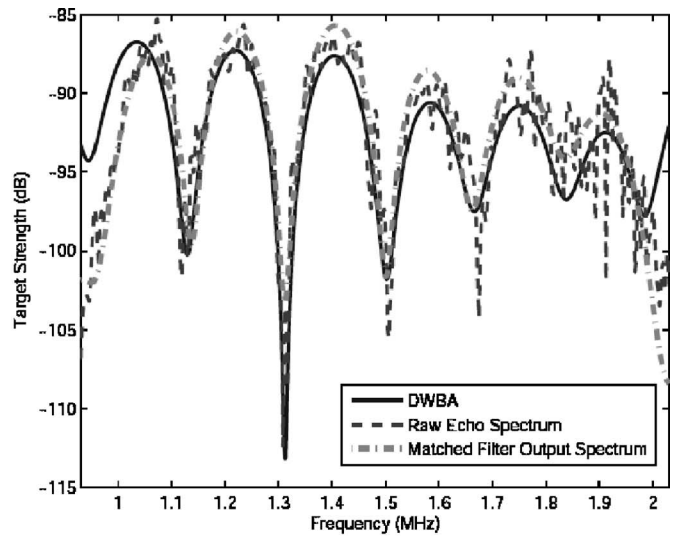


FIG. 4. View of the method for calculating beam shape for a scatterer positioned at ζ and a transmitter and receiver pair. The angles between the scatterer position vector, and the vector to each element are defined by η and μ . The angles are used to compute the change in sound intensity as a function of frequency.



(a) Copepod



(b) Euphausiid

FIG. 5. Typical examples of simulated data for both the copepod and euphausiid class. The estimated target strength using the raw echo, and the matched filter output are plotted on top of the target strength as modeled by the DWBA. The high level of noise is clearly visible, as is the improvement obtained from the matched filter.

beration noise in the same frequency band as the transmit signal is added to the scattering amplitude. The noise is generated by taking the product of the fast Fourier transform (FFT) of the transmit signal $S_0[k]$ with the FFT of a realization of white Gaussian noise $W_j[k]$.

A constant reverberation level is used and thus the signal to noise ratio (SNR) varies as a function of length and orientation of the scatterer. The SNR for the copepod data ranges from -11 to 24 dB whereas the SNR for the euphausiid data ranges from -20 to 35 dB. These ranges for SNR were selected such that the average SNR was close to 7 dB for the copepods and 15 dB for the euphausiids. These values were deemed to be comparable to what is achieved in practical systems^{30,31}. The noise level can also be defined in terms of an equivalent target strength of -110 dB.

In generating the data, the noise term $W_j[k]$ is first added to the product of the scattering amplitude and the beam shape factor at each wave number bin to yield the noisy scattering amplitude

$$\hat{S}_j[k] = S_j[k]\beta_{j,t}[k] + W_j[k]. \quad (13)$$

The convolution defined in Eq. (1) is accomplished in the frequency domain by computing the product of the FFT of the transmit signal $S_0[k]$ with the noisy scattering amplitude $\hat{S}_j[k]$. Using Eq. (13), the pressure signal on the j th element is then obtained by an inverse FFT (IFFT),

$$p_j[n] = \text{Real} \left[\frac{1}{N} \sum_{k=0}^{N-1} S_0[k] \hat{S}_j[k] e^{2\pi i k n / N} \right]. \quad (14)$$

This process is repeated for each of the M array elements building up the vector

$$\mathbf{p}[n] = (p_1[n], \dots, p_M[n])^T. \quad (15)$$

In order to explore the classification success as a function of carrier frequency and bandwidth several different types of signals were used. For the narrowband signals (10% bandwidth) frequencies of 1 and 2 MHz were selected. For the broadband signal, a linear frequency modulated (LFM) chirp was used with a starting frequency of 1 MHz and ending frequency of 2 MHz. The signal duration and energy was kept constant for all signals. The range of frequencies was selected based on past experience with measuring scatter from animals of the size considered here.

To improve the SNR of the data input to the feature extraction algorithms, the raw echo data resulting from the simulation is passed through a matched filter.³² For the transmit signal $s_0[n]$, and a received echo on array element j defined by $p_j[n]$, the output of the matched filter is

$$\mathcal{M}_j[n] = \sum_{p=-\infty}^{\infty} s_0[p] p_j[p-n]. \quad (16)$$

The matched filter output is windowed around the peak in the output with a window size of $W=50$, corresponding to a time of $5 \mu\text{s}$ or a distance of 7.5 mm. For the broadband signal the time-bandwidth product is 120 , yielding a processing gain of roughly 20 dB. For the narrowband signals, the time-bandwidth product is much lower, and thus the processing gain is low as well, around 10 dB.

To visualize the type of noise, and its effect on the estimation of the scattered signal, the model target strength is displayed along with the estimated target strength for the broadband signal type in Fig. 5.

III. FEATURE EXTRACTION

In order to facilitate the classification procedure the data are mapped to a feature space which dramatically reduces the dimensionality of the data while simultaneously highlighting interclass differences. The three feature spaces used in this work are described in the following.

A. Single frequency based feature space

To explore the result of using only a single frequency rather than a broad spectrum of frequencies, a single frequency based feature space is defined in which the FFT of the matched filter output on each array element is computed and its magnitude squared is integrated over a small bandwidth. Only the narrowband data are used in this feature space. The resulting sum squared magnitudes on each array element are then combined to form a single feature vector. Specifically, assuming that the j th array element collects N samples, the power in the narrow bandwidth of the signal (k_{\min} to k_{\max}) is

$$P_j = \sum_{k=k_{\min}}^{k_{\max}} \left| \sum_{n=0}^{N-1} \mathcal{M}_j[n] e^{-2\pi i n k / N} \right|^2. \quad (17)$$

For the simulations presented here, $N=1200$. The single frequency feature vector is then defined as

$$\mathbf{y} = (P_1, P_2, \dots, P_M)^T. \quad (18)$$

The bandwidth (k_{\min} to k_{\max}) is set equal to the transmit signal bandwidth of 10% of the center frequency. The center frequencies used are 1 and 2 MHz.

B. Discrete cosine transform based feature space

The discrete cosine transform (DCT) has numerous qualities that make it attractive as a feature mapping. For one, the coefficients of the DCT are uncorrelated. It can also be shown that the DCT can embed most of the energy in the data into a small number of coefficients. While there is no guarantee that such an embedding will yield a discriminant feature space, this is often the case in practice. The DCT based feature space uses the power spectrum of the matched filter output for the broadband 1–2 MHz data. The power spectrum is computed as

$$P_j[k] = \left| \sum_{n=0}^{N-1} \mathcal{M}_j[n] e^{-2\pi i n k / N} \right|^2. \quad (19)$$

For the results presented here $N=1200$. Having computed $P_j[k]$ for each array element, the DCT of the power spectrum is computed as

$$E_j^{\text{DCT}}[l] = \sqrt{\frac{2}{N}} \beta[l] \sum_{k=0}^{N-1} P_j[k] \cos\left(\frac{\pi l(2k+1)}{2N}\right), \quad (20)$$

where

$$\beta[l] = \begin{cases} \frac{1}{\sqrt{2}}, & l=0 \\ 1, & l=1, \dots, N-1 \end{cases}. \quad (21)$$

The values of the K largest (ordered by magnitude) coefficients in the transform are retained in the feature vector for the j th array element

$$\mathbf{y}_j = (E_j^{\text{DCT}}[l^1], \dots, E_j^{\text{DCT}}[l^K])^T, \quad (22)$$

where the features are arranged such that $l^1 \leq l^2 \leq \dots \leq l^K$. This procedure can be interpreted as an adaptive threshold of the DCT of the power spectrum. Finally, the feature vectors at each array element are combined into a single feature vector,

$$\mathbf{y} = (\mathbf{y}_1^T, \dots, \mathbf{y}_M^T)^T. \quad (23)$$

The feature vector, \mathbf{y} captures the K most energetic wave number bins in the power spectrum of the received signal at each element of the array. A range of values for K were analyzed. It was found that the values $K=1$, $K=2$, and $K=4$ yield the best performance. As K increases, the feature vector is able to capture more of the variability in the frequency response of the scatterer at the cost of a larger feature space dimension. A wave number bin width of $\Delta l=35$ rad/m was used throughout the simulation.

C. Frequency correlation based feature space

One of the major drawbacks of the previous two feature spaces is that they do not naturally combine the multiple angle data. Instead, the features from each angle of the multiple angle data are lumped together as one big feature vector. This can cause problems in the case of the DCT feature space as the dimensionality of the feature space grows as K times the number of angles M . In this section, a feature mapping which combines the multiple angle data systematically while extracting the features is defined. The features are the eigenvalues of the frequency correlation matrix.

The frequency correlation matrix is obtained by computing the correlation between all pairs of received wave forms in the frequency domain. Specifically, the correlation matrix \mathbf{C} is defined as

$$\mathbf{C} = \mathbf{F}^H \mathbf{F}, \quad (24)$$

where

$$\mathbf{F} = (\mathbf{f}_1, \dots, \mathbf{f}_M), \quad (25)$$

and

$$f_j[k] = \sum_{n=0}^{N-1} \mathcal{M}_j[n] e^{-2\pi i n k / N} \quad (26)$$

is the FFT of the matched filter output for the received data on element j . Thus, the individual elements of the frequency correlation matrix are the cross correlations between the Fourier transforms of the data on the array elements

$$C_{ij} = \mathbf{f}_i^H \mathbf{f}_j. \quad (27)$$

The features are extracted from the frequency correlation matrix using an eigenvalue decomposition³³

$$\Lambda = \mathbf{Q}^H \mathbf{C} \mathbf{Q}, \quad (28)$$

where $\Lambda = \text{Diag}(\lambda_1, \dots, \lambda_M)$ is the diagonal matrix of eigenvalues. The feature vector is then formed by taking the diagonal elements of Λ ,

$$\mathbf{y} = (\lambda_1, \dots, \lambda_M)^T. \quad (29)$$

The eigenvalues (features) are not simply related to the data at each angle as in the case of the previous two feature spaces. Each eigenvalue is derived from data at all angles. This is the key benefit over the other two feature spaces. The number of nonzero eigenvalues is upper bounded by the number of angles, yet may be lower depending on the degree to which the echoes received at each angle are correlated with one another in the frequency domain. For example, if

$$\mathbf{f}_i^H \mathbf{f}_j \approx \begin{cases} 0 & \text{for } i \neq j \\ \kappa & \text{for } i = j \end{cases}, \quad (30)$$

the frequency correlation matrix $\mathbf{C} \approx \kappa \mathbf{I}$ and the eigenvalue spread is nearly flat. In contrast, if

$$\mathbf{f}_i^H \mathbf{f}_j \approx \kappa \quad \forall i, j, \quad (31)$$

the frequency correlation matrix is approximately rank one, and the eigenvalue value spectrum will be highly peaked at the first eigenvalue. The first example can be thought of as representing a complex shape, where the spectrum of the received signal varies substantially as a function of angle. This second example corresponds to scattering from an angularly symmetric shape. Therefore, in the presence of noise, the variability at each array element is due only to noise.

IV. CLASSIFICATION OF FEATURES

Given a set of features that have been extracted from these data, the next task is to develop a method for assigning a class label to each feature so as to minimize a particular loss function. As is commonly done in pattern classification, the “0-1” loss function is applied which assigns equal penalties to classification errors made for either class.³⁴ In the zooplankton classification problem considered here, this is a reasonable loss function due to the fact that each class has, in effect, equal significance. It can be shown³⁴ that the classification rule which minimizes the “0-1” loss function is the Bayes decision rule (BDR)

$$i^* = \underset{i}{\text{argmax}} P_{C|Y}(i|\mathbf{y}), \quad (32)$$

where the class i^* , having the maximum *a posteriori* probability given the feature vector \mathbf{y} , is chosen. The BDR can be written in terms of the class conditional density (CCD) $p_{Y|C}(\mathbf{y}|i)$ using Bayes rule, and assuming a prior class distribution $P_C(i)$, as

$$i^* = \underset{i}{\text{argmax}} p_{Y|C}(\mathbf{y}|i) P_C(i). \quad (33)$$

In practice, the prior probability may or may not be known. For the procedure considered here, it is assumed that the priors for each class are equal. As a result, that term drops from the maximization. The remaining task is that of

maximizing the CCD which is equivalent to computing the maximum likelihood estimate of the class label.

Unfortunately, the CCD is almost always unknown. In the best case, only the form of the density is known, but not the parameters that define the actual shape. This is one of the fundamental difficulties encountered in pattern classification and is the point at which *a priori* knowledge or training data must be used to learn about the structure of $p_{Y|C}(\mathbf{y}|i)$.

Here, two popular classifiers are considered: the nearest neighbor (NN) classifier, and the multilayer perceptron (MLP) classifier. The properties of each of these classifiers are briefly reviewed as the implementations used here are standard.

The NN classifier assigns a class label to a new pattern based on the label of the training pattern which is “nearest” to the new pattern according to a particular distance metric. For a given training set $\mathcal{D} = \{(\mathbf{y}_1, i_1), \dots, (\mathbf{y}_N, i_N)\}$ where \mathbf{y}_n is a feature extracted from the data according to the methods defined in Sec. III and i_n is the associated class label, the NN classifier under the 2-norm assigns the label i_k where k is the index of the nearest neighbor,

$$k = \underset{i}{\text{argmin}} (\mathbf{y} - \mathbf{y}_i)^T (\mathbf{y} - \mathbf{y}_i). \quad (34)$$

In contrast to the NN method, the MLP tries to learn the mapping from feature space to class label space using multiple levels of weighted combinations of the components of the features rather than using the training data explicitly to represent the underlying CCDs. In essence, the MLP learns to approximate $P_{C|Y}(i|\mathbf{y})$ via experience gained from analyzing numerous examples. It has been shown that this type of classifier can yield very good results in underwater target classification¹⁹ as it is one of the best methods for approximating a high dimensional function.

Given the two class problem, the MLP has two output nodes. The number of input nodes is the same as the number components of the feature vector \mathbf{y} . A single hidden layer is used with the number of nodes selected to be twice the number of input nodes. The network is compactly expressed³⁵ as

$$c_k = U \left(\sum_{j=0}^{2M} \tilde{w}_{kj} V \left(\sum_{i=0}^M w_{ji} y_i \right) \right), \quad (35)$$

where U and V are nonlinear mapping functions, \tilde{w}_{kj} and w_{ji} are network weights, y_i is the i th component of the feature vector, and c_k is the k th component of the classification vector. The weight matrices of the network are initialized randomly at the start of training, and updated at each iteration so as to minimize the error on the training set. Both mapping functions are selected to be the softmax function,³⁴ and the network is trained using the scaled conjugate gradient method. Prior to training and testing, all inputs to the network are z -scaled in the log domain. The training is implemented in MATLAB (The Mathworks; Natick, MA) using the NETLAB toolbox.³⁶

V. RESULTS AND DISCUSSION

Classifier performance. The classifiers defined in Sec. IV are now evaluated quantitatively on a set of test data

mapped into each of the feature spaces defined in Sec. III. The results are displayed as the absolute probability of error as a function of the number of angles (or array elements) that are combined in the classifier. Specifically, the number of angles is equal to the number of array elements included in order starting from element 1. So, for example, three angles would correspond to using array elements 1, 2, and 3, and four angles would correspond to using elements 1, 2, 3, and 4. The probability of error is computed according to

$$p(\text{error}) = \frac{p(c|e) + p(e|c)}{2}, \quad (36)$$

where $p(c|e)$ is the probability of classifying an euphausiid as a copepod, and $p(e|c)$ is the probability of classifying a copepod as an euphausiid. Here, the fact that each class is equally likely in this simulation has been used. A consequence of the equal representation for each class is that a system which randomly guesses the class would have a probability of error of 50%. Therefore, 50% probability of error can be achieved with no effort, and any classification strategy should have an error below 50%. The classification experiment is performed both for the case where beam shape is neglected from the simulation, and where the beam shape is included. The results of the classification for both cases of data are shown in Fig. 6.

The 1 and 2 MHz curves result from using the single frequency feature space and the respective narrowband data. The DCT and CM curves result from using the DCT and frequency correlation feature spaces with the broadband 1–2 MHz LFM chirp data. The DCT(1), DCT(2), and DCT(4) curves apply the DCT method outlined in Sec. III with $K=1$, $K=2$, or $K=4$, respectively. Figure 6 illustrates that there is a general trend of decreasing probability of error as more angles are used in the classification. The amplitude and frequency response of acoustic scatter from crustacean zooplankton is directly related to the scatterer shape, size, and orientation. The addition of more angles in the classifier can be interpreted physically as observing the scatterer from multiple views. The probability of error is reduced as more angles are used in essence because there is less uncertainty about the shape of the scatterer.

When only one angle is used, observing scatter over a broad range of frequencies (1–2 MHz) reduces the probability of error substantially over the single frequency case. This is a consequence of the fact that single frequency scatter is much more sensitive to scatterer orientation and size than broadband scatter due to the effect of coherent interference at a given frequency. As broadband data excites many frequencies, it is far more robust to changes in scatterer shape and orientation.

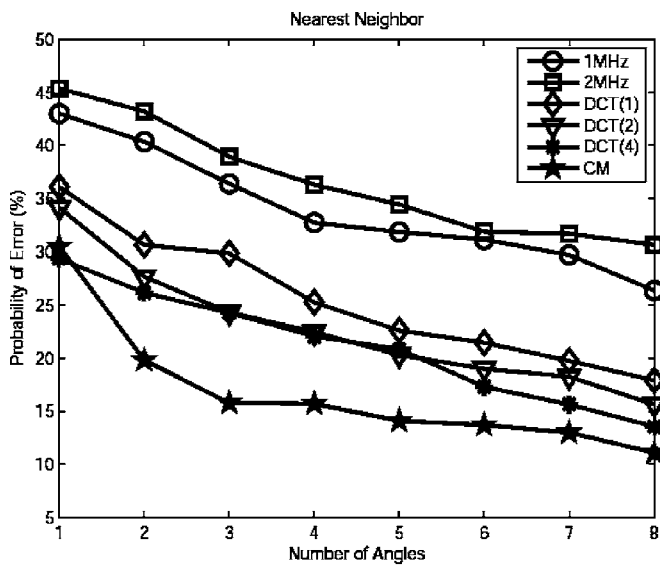
The effect of including beam shape on the probability of error can be clearly seen by comparing the top and bottom rows of Fig. 6. The effect is essentially to shift the curves toward higher probability of error. This is a consequence of the fact that the random position of the scatterer in the beam adds another kind of noise to the data. However, it is possible this that kind of noise can be corrected by exploiting the angular diversity of the array to locate the scatterer in the

beam, and correct for the beam shape. While not considered in this paper, this process will be investigated in future work.

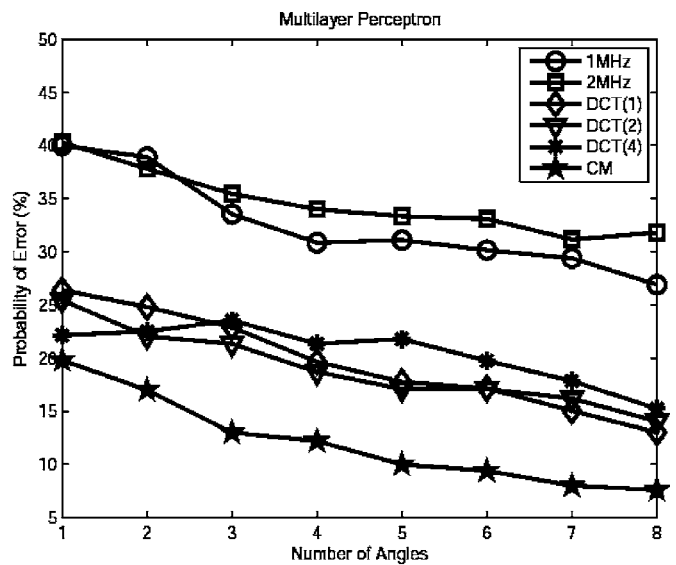
The frequency correlation feature space clearly outperforms all of the other feature spaces for both classifiers. This is to be expected since it is the only feature space which naturally combines the echoes at each angle when extracting features. It may be inferred from the poor performance when only a single angle is used that the variation in the echo as a function of angle is most discriminant between the two classes. The frequency correlation feature space efficiently extracts this information from these data. The performance (in terms of correct and misclassification) of each classifier using all eight angles with, and without including beam shape is displayed in Tables I and II. Interestingly, there is a wide variation in the correct and misclassification results for the different feature spaces. The improvement in classification performance as additional angles, and as a result increased array aperture, are used is a direct consequence of having additional independent views of the scatterer. The additional views reduce the uncertainty in the shape of the scatter by way of the intimate link between scatterer shape, and angularly varying scattering amplitude as defined in Eq (2). Since the two classes of scatterers have distinct shapes, the reduced uncertainty in shape leads to improved classification performance. In general, all of the feature spaces, and classifiers have slightly higher accuracy for the copepod class rather than the euphausiid class except for the frequency correlation feature space which is inconsistent between the two classification algorithms. This is likely a consequence of the large number of orientations for which scattering from the euphausiid is very weak due to the elongated body. There is also a systematic increase in probability of error for the 2 MHz data over the 1 MHz data for both classifiers. This is likely caused by a greater similarity in scattering amplitude at 2 than 1 MHz between the two classes. This could be caused by the fact that at 2 MHz, the scattering is further into the geometric regime, and thus the scattering amplitude is less sensitive to the scatter size. Finally, the best results, are misclassifications of 9.0% and 6.2% for the copepod and euphausiid, respectively. As a result, the total absolute probability of error is 7.6% in the best case. This gives an improvement over random guessing of 84.8%.

VI. CONCLUSIONS AND FUTURE WORK

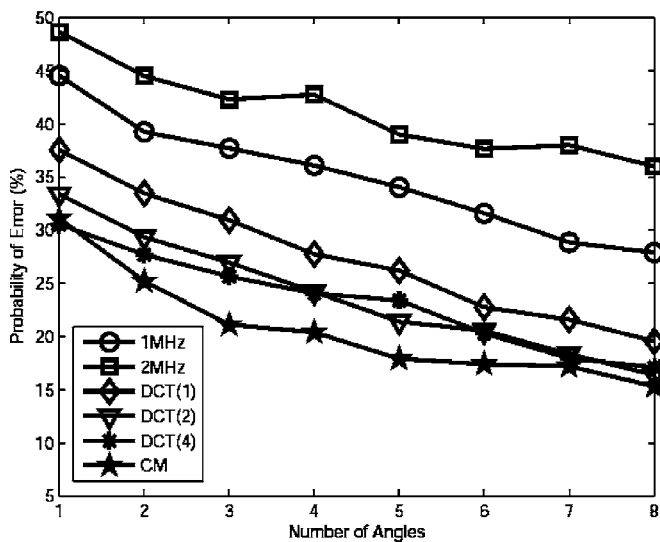
In this paper, the use of multiple angle acoustic scatter to discriminate between two classes of ecologically important zooplankton has been explored using simulations. The research is motivated by the current need for more descriptive acoustic sensors for studying zooplankton *in situ*. Past work in this area has been limited by the inherent ambiguity in discrimination ability due to the sensitivity of acoustic scattering to material properties and scatterer orientation. These difficulties have been confirmed here where it has also been shown that it is possible to use scatter measured over a multiplicity of angles to achieve a higher rate of correct classification. Using synthetic data, generated via the use of the distorted wave Born approximation, two ecologically impor-



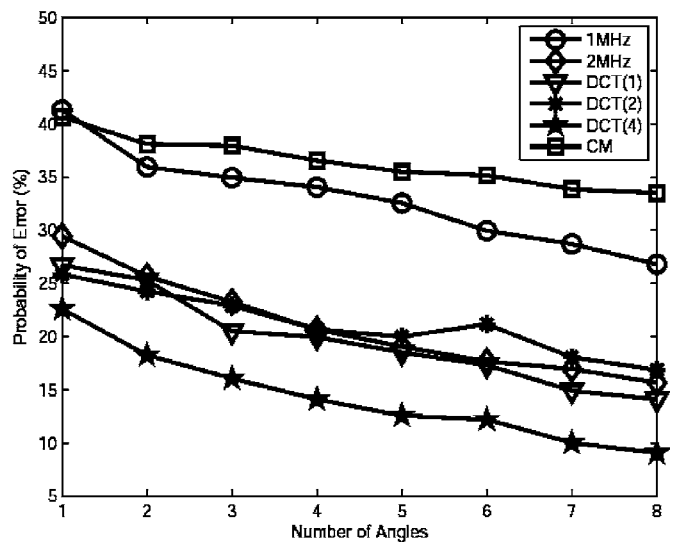
(a) Nearest Neighbor Classifier



(b) Multi-Layer Perceptron Classifier



(c) Nearest Neighbor Classifier
(with beam shape)



(d) Multi-Layer Perceptron Classifier
(with beam shape)

FIG. 6. Comparison between the NN classifier and the MLP classifiers for the case of no beam shape (a), (b) and beam shape (c), (d), in six different feature spaces. The 1 and 2 MHz curves correspond to the single frequency based feature space. The DCT(1), DCT(2), and DCT(4) feature spaces use the DCT method outlined in Sec. III with $K=1$, $K=2$, or $K=4$, bin indices included at each angle. The CM curves result from the frequency correlation feature space outlined in Sec. III.

tant classes of zooplankton—copepods and euphausiids—were classified. The classification performance, measured in terms of probability of error, is dramatically improved over single angle observation methods via the use of additional angles. This improvement is even more substantial when broadband scatter is used.

The simulations performed here were geared toward a practical system which could be deployed in the field. Therefore, constraints were placed on the bandwidth of the trans-

mit signal, and the angular distribution of the receivers in this context. The length distributions for both classes were chosen to be typical of those encountered in the Southern California region.^{37,38} In order to understand the ramifications of the proposed method in the presence of noise, a constant level of noise that resulted in an average SNR of 7 dB for the copepod and 15 dB for the euphausiid, or an equivalent target strength of -110 dB, was used. This noise level is consistent with practical systems that have been used

TABLE I. The correct and misclassification probabilities for each classifier when beam shape is excluded from the simulation.

Classifier	Copepod		Euphausiid		Total <i>p</i> (error) (%)
	Correct (%)	Mis (%)	Correct (%)	Mis (%)	
NN/1 MHz	76.1	23.9	71.2	28.8	26.3
NN/2 MHz	74.1	25.9	64.6	35.4	30.6
NN/DCT (4)	87.5	12.5	85.4	14.6	13.5
NN/CM	90.8	9.2	87.0	13.0	11.1
MLP/1 MHz	76.3	23.7	70.4	29.6	26.6
MLP/2 MHz	70.3	29.7	66.1	33.9	31.8
MLP/DCT (4)	85.0	15.0	84.5	15.5	15.2
MLP/CM	91.0	9.0	93.8	6.2	7.6

in the field.^{30,31} In addition, although a strong effort was expended in order to make the work realistic, the performance of a field system may be limited by issues that have not been considered in this work. Specifically, the models used here, while accurate for weak sound scattering, do not include variability due to individual shape, or body pose. Furthermore, uniform orientation distributions were used here, where as in the field, the distributions may be different. Given the promising results observed here, these additional degrees of freedom certainly warrant further investigation through more complex simulations, as well as observation of live animals.

A curious, but potentially very helpful aspect of our result is that a one-dimensional array is capable of capturing enough information from a random three-dimensional orientation to yield good classification performance. It may therefore be that simple array geometries, which can dramatically reduce the development and deployment cost associated with such systems, constitute a pragmatic solution to the *in situ* classification of zooplankton after all.

ACKNOWLEDGMENTS

The authors would like to thank D. E. McGehee, M. Benfield, D. V. Holliday, and G. Greenlaw for development and maintenance of the Advanced Multifrequency Inversion Methods for Classifying Acoustic Scatters website, two anonymous reviews for helpful comments on the manuscript, and California Sea Grant for funding this research.

TABLE II. The correct and misclassification probabilities for each classifier when beam shape is included in the simulation.

Classifier	Copepod		Euphausiid		Total <i>p</i> (error) (%)
	Correct (%)	Mis (%)	Correct (%)	Mis (%)	
NN/1 MHz	74.1	25.9	70.1	29.9	27.9
NN/2 MHz	67.8	32.2	60.2	39.8	36.0
NN/DCT (4)	83.5	16.5	82.3	17.7	17.1
NN/CM	85.8	14.2	81.5	18.5	16.3
MLP/1 MHz	75.8	24.5	74.2	25.8	25.0
MLP/2 MHz	69.9	30.1	63.1	36.9	33.5
MLP/DCT (4)	85.3	14.7	81.0	19.0	16.8
MLP/CM	91.2	8.8	90.7	9.3	9.0

- ¹D. McNaught, "Acoustical determination of zooplankton distributions," in The 11th Annual Conference on Great Lakes Research, 1968, pp. 76–84.
- ²D. V. Holliday (1977), "Extracting biophysical information from the acoustic signatures of marine organisms," in *Oceanic Sound Scattering Prediction*, edited by N. R. Anderson and B. J. Zahuraned (Plenum, New York), pp. 162–211.
- ³C. Greenlaw, "Acoustic estimation of zooplankton populations," *Limnol. Oceanogr.* **24**, 226–242 (1979).
- ⁴D. Holliday, R. Pieper, and G. Kleppel, "Determination of zooplankton size and distribution with multi-frequency acoustic technology," *J. Cons., Cons. Int. Explor. Mer* **41**, 226–238 (1989).
- ⁵D. McGehee, D. Demer, and J. Warren, "Zooplankton in the Ligurian Sea. I. Characterization of their dispersion, relative abundance and environment during summer 1999," *J. Plankton Res.* **26**, 1409–1418 (2004).
- ⁶M. McManus, O. Cheriton, P. Drake, D. Holliday, C. Storlazzi, P. L. Donaghay, and C. Greenlaw, "Effects of physical processes on structure and transport of thin zooplankton layers in the coastal ocean," *Mar. Ecol.: Prog. Ser.* **301**, 199–215 (2005).
- ⁷L. Martin, T. Stanton, P. Wiebe, and J. F. Lynch, "Acoustic classification of zooplankton," *ICES J. Mar. Sci.* **53**, 217–224 (1996).
- ⁸L. Traykovski, T. Stanton, P. Wiebe, and J. Lynch, "Model-based covariance mean variance classification techniques: Algorithm development and application to the acoustic classification of zooplankton," *IEEE J. Ocean. Eng.* **23**, 344–364 (1998).
- ⁹L. Goodman, "Acoustic scattering from ocean microstructure," *J. Geophys. Res., [Atmos.]* **95**, 11557–11573 (1990).
- ¹⁰J. S. Jaffe, "Using multiple-angle scattered sound to size fish swim bladders," *ICES J. Mar. Sci.* **63**, 1397–1404 (2006).
- ¹¹K. G. Foote, "Rather high frequency sound scattering by swimbladdered fish," *J. Acoust. Soc. Am.* **78**, 688–700 (1985).
- ¹²T. Stanton, D. Chu, and P. Wiebe, "Acoustic scattering characteristics of several zooplankton groups," *ICES J. Mar. Sci.* **53**, 289–295 (1996).
- ¹³T. Stanton, D. Chu, and P. Wiebe, "Sound scattering by several zooplankton groups. II. Scattering models," *J. Acoust. Soc. Am.* **103**, 236–253 (1998).
- ¹⁴T. Stanton and D. Chu, "Review and recommendations for the modelling of acoustic scattering by fluid-like elongated zooplankton: Euphausiids and copepods," *ICES J. Mar. Sci.* **57**, 793–807 (2000).
- ¹⁵D. Reeder and T. Stanton, "Acoustic scattering by axisymmetric finite-length bodies: An extension of a two-dimensional conformal mapping method," *J. Acoust. Soc. Am.* **116**, 729–746 (2004).
- ¹⁶D. McGehee, R. O'Driscoll, and L. Traykovski, "Effects of orientation on acoustic scattering from Antarctic krill at 120 kHz," *Deep-Sea Res., Part II* **45**, 1273–1294 (1998).
- ¹⁷A. Lavery, T. Stanton, D. McGehee, and D. Z. Chu, "Three-dimensional modeling of acoustic backscattering from fluid-like zooplankton," *J. Acoust. Soc. Am.* **111**, 1197–1210 (2002).
- ¹⁸D. E. McGehee, M. Benfield, D. V. Holliday, and C. Greenlaw, "Advanced multifrequency inversion methods for classifying acoustic scatterers," http://zooplankton.lsu.edu/scattering_models/MultifreqInverseMethods.html.
- ¹⁹M. Azimi-Sadjadi, D. Yao, Q. Huang, and G. Dobeck, "Underwater target classification using wavelet packets and neural networks," *IEEE Trans. Neural Netw.* **11**, 784–794 (2000).
- ²⁰N. Dasgupta, P. Runkle, L. Carin, L. Couchman, T. Yoder, J. Bucaro, and G. Dobeck, "Class-based target identification with multispect scattering data," *IEEE J. Ocean. Eng.* **28**, 271–282 (2003).
- ²¹S. Ji, X. Liao, and L. Carin, "Adaptive multispect target classification and detection with hidden Markov models," *IEEE Sens. J.* **5**, 1035–1042 (2005).
- ²²M. Robinson, M. Azimi-Sadjadi, and J. Salazar, "Multi-aspect target discrimination using hidden Markov models and neural networks," *IEEE Trans. Neural Netw.* **16**, 447–459 (2005).
- ²³P. Bharadwaj, P. Runkle, L. Carin, J. Berrie, and J. Hughes, "Multiaspect classification of airborne targets via physics-based HMMs and matching pursuits," *IEEE Trans. Aerosp. Electron. Syst.* **37**, 595–606 (2001).
- ²⁴B. Pei and M. Bao, "Multi-aspect radar target recognition method. Based on scattering centers and HMMs," *IEEE Trans. Aerosp. Electron. Syst.* **41**, 1067–1074 (2005).
- ²⁵Y. Dong, P. Runkle, L. Carin, R. Damarla, A. Sullivan, M. Ressler, and J. Sichina, "Multi-aspect detection of surface and shallow-buried unexploded ordnance via ultra-wideband synthetic aperture radar," *IEEE Trans. Geosci. Remote Sens.* **39**, 1259–1270 (2001).
- ²⁶P. Runkle, P. Bharadwaj, L. Couchman, and L. Carin, "Hidden Markov

- models for multiaspect target classification," *IEEE Trans. Signal Process.* **47**, 2035–2040 (1999).
- ²⁷D. Li, M. Azimi-Sadjadi, and M. Robinson, "Comparison of different classification algorithms for underwater target discrimination," *IEEE Trans. Neural Netw.* **15**, 189–194 (2004).
- ²⁸A. Fleminger, J. D. Isaacs, and J. G. Wyllie, "Zooplankton biomass measurements from calcofi cruises of July 1955 to 1959 and remarks on comparison with results from October, January, and April cruises of 1955 to 1959," Technical Rep., CalCOFI Atlas No. 21, 1974.
- ²⁹P. Morse and K. Ingard, *Theoretical Acoustics* (Princeton University Press, Princeton, 1968).
- ³⁰J. Jaffe, E. Reuss, and G. Chandran, "Ftv: A sonar for tracking macrozooplankton in three dimensions," *Deep-Sea Res., Part I* **42**, 1495–1512 (1995).
- ³¹A. Genin, J. S. Jaffe, R. Reef, C. Richter, and P. J. S. Franks, "Swimming against the flow: A mechanism of zooplankton aggregation," *Science* **308**, 860–862 (2005).
- ³²S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory* (Prentice Hall, Upper Saddle River, NJ, 1998), Vol. **1**.
- ³³T. K. Moon and W. C. Stirling, *Mathematical Methods and Algorithms for Signal Processing* (Prentice Hall, Upper Saddle River, NJ, 2000).
- ³⁴R. Duda, P. Hart, and D. G. Stork, *Pattern Classification* (Wiley Interscience, New York, 2000).
- ³⁵C. M. Bishop, "Neural networks and their applications," *Rev. Sci. Instrum.* **65**, 1803–1932 (1994).
- ³⁶I. Nabney, *Netlab: Algorithms for Pattern Recognition* (Springer, New York, 2001).
- ³⁷M. D. Ohman and B. E. Lavaniegos, "Comparative zooplankton sampling efficiency of a ring net and bongo net with comments on pooling of subsamples," Technical Rep., CalCOFI Rep., Vol. **43**, 2002.
- ³⁸E. Brinton and J. G. Wyllie, "Distributional atlas of euphausiid growth stages off southern California, 1953–1956," Technical Rep., CalCOFI Atlas, No. 24, 1976.