

# Underwater Target Detection Using Multichannel Subband Adaptive Filtering and High-Order Correlation Schemes

Chunhua Yuan, Mahmood R. Azimi-Sadjadi, *Senior Member, IEEE*, JoEllen Wilbur, *Member, IEEE*, and Gerald J. Dobeck

**Abstract**—In this paper, new pre- and post-processing schemes are developed to process shallow-water sonar data to improve the accuracy of target detection. A multichannel subband adaptive filtering is applied to preprocess the data in order to isolate the potential target returns from the acoustic backscattered signals and improve the signal-to-reverberation ratio. This is done by estimating the time delays associated with the reflections in different subbands. The preprocessed results are then beamformed to generate an image for each ping of the sonar. The testing results on both the simulated and real data revealed the efficiency of this scheme in time-delay estimation and its capability in removing most of the competing reverberations and noise. To improve detection rate while significantly minimizing the incident of false detections, a high-order correlation (HOC) method for postprocessing the beamformed images is then developed. This method determines the consistency in occurrence of the target returns in several consecutive pings. The application of the HOC process to the real beamformed sonar data showed the ability of this method for removing the clutter and at the same time boosting the target returns in several consecutive pings. The algorithm is simple, fast, and easy to implement.

**Index Terms**—Adaptive filtering, high-order correlation, sonar, target detection, wavelet transform.

## I. INTRODUCTION

SEA mine hunting and detection pose notoriously challenging problems due to the vastness of the ocean, small sizes of the mines, the acoustically reverberant environment, and the frequent occurrence of biologicals or magnetic clutter. It is easy to lay a minefield but very dangerous, costly, and time consuming to localize and clear it. Considerable attention has recently been focused on this area and various signal processing schemes have been developed to aid in the detection of underwater mines [1]–[3]. The research efforts have primarily been concentrated on three major topics: preprocessing, beamforming and postprocessing.

*Preprocessing* is applied to the acoustic backscattered signals to get rid of most of the clutter and reverberation effects and thus improve the signal-to-noise ratio (SNR). The most commonly

used preprocessing method is *matched filtering* [4], [5] which gives the maximum likelihood (ML) [4] estimate of the signal amplitude and the arrival time. However, in sonar applications, the assumption of a deterministic signal is rarely valid. Due to the strong reverberation effects and the characteristics of the medium, the received signal generally contains distorted, delayed, amplitude weighted, and overlapped replicas of the transmitted or incident signal. In [4], in order to overcome the degradation of the classical matched filter in the presence of overlapped signals, an inverse filter is used to improve the resolution of the time-delay estimation. The price paid for this high resolution is a reduction of SNR. Moreover, careful design is needed to avoid the instability of the inverse filter. In [6] and [7], algorithms were developed to perform multipath time-delay estimation. In [6], to deal with a long received signal and a short transmitted signal, the long data record was broken into overlapped short segments and each segment was modeled individually without misfitting or truncating any path at the end of each segment. The estimates in each segment were then combined to model the entire data record. In [7], a least-square (LS)-based algorithm was developed to estimate the arrival time of the overlapping acoustic signals that consist of attenuated and delayed replicas of a known transmitted signal. To avoid minimization of a highly oscillatory error function, complex amplitude was allowed which resulted in a much smoother error function that was easier to minimize using the gradient-based schemes. In [8]–[12], various kinds of adaptive filtering approaches were developed to jointly estimate the time delays and filter weights. Instead of modeling the received acoustic signal as the sum of delayed and attenuated incident signals as in [6] and [7], a more general model allowing frequency attenuation in the delay path was used. Various higher order statistics-based algorithms [13]–[16] were also developed for time-delay estimation and signal detection. In [5] and [13], different schemes combining matched filter and higher order statistics were proposed in order to come up with a detection scheme that is tolerant to signal shift and deals with the additive noise with unknown spectral characteristics. These schemes construct a hypothesis using the zeroth lag of the higher order correlations of the matched filter output. The signals can be either deterministic or random non-Gaussian, and the noise is assumed to be zero-mean Gaussian with an unknown covariance sequence. In many underwater acoustic applications, an estimate of the power spectral density (PSD) of the received backscattered data is often employed for signal

Manuscript received June 20, 1998; revised August 14, 1999. This work was supported by the Office of Naval Research (ONR 321TS) under Contract N61331-94-K-0018.

C. Yuan and M. R. Azimi-Sadjadi are with the Department of Electrical and Computer Engineering, Colorado State University, Fort Collins, CO 80523 USA.

J. Wilbur and G. J. Dobeck are with NSWC/DD, Coastal Systems Station, Panama City, FL 32407-7001 USA.

Publisher Item Identifier S 0364-9059(00)00749-4.

detection. PSD is essentially a sum of the estimates of the second-order moments for each frequency component in the frequency domain. Higher order spectra contain information not present in the PSD, especially in detecting non-Gaussian and nonlinear transient signals. Frequency domain *Kurtosis* estimation and *Bispectral* analysis were used in [15] and [16] as aids in detecting randomly occurring signals. Several wavelet-based algorithms [17]–[21] were also developed to perform linear system modeling, adaptive filtering, signal detection, and time-delay estimation for underwater target detection. In [18], wavelet transform was used in an adaptive filtering structure and the coefficients of the adaptive filter were updated by the least mean squares (LMS) algorithm. In [20], a wavelet transform-based cross-correlation method was proposed for time-delay estimation in pulsed wave ultrasound. The scheme has a unique feature of using varying time–frequency windows in processing compared with the conventional windowed cross-correlation method where a fixed size window is used. The wavelet shrinkage method [21] was also used to *denoise* the noisy acoustic backscattered signal and greatly improve the SNR. The underlying assumption is that the noise contributes to most of the wavelet coefficients while the signal contributes to only a few coefficients. Therefore, by setting the smaller coefficients to zero, in a statistically guided manner, noise can be optimally eliminated while preserving the signal.

In *postprocessing*, the beamformed output is further processed to find additional clues for distinguishing targets from the background noise/clutter. It will alleviate the burden of sonar operator, help to provide a high target detection rate, and reduce the incident of false reports. In [22], statistical measures, i.e., *mean*, *standard deviation*, *skewness*, and *kurtosis* were calculated from the digitized beamformed sonar data. The combination of these statistical measures provides an additional clue about the presence of a target versus background. In [23], several methods including energy detector, sliding matched filter, skewness matched filter, and dispersion-based reconditioning were applied to the Toroidal Volume Search Sonar (TVSS) beamformer output in order to increase the signal-to-reverberation ratio (SRR). It was reported that, while the skewness matched filter offers some improvements, the dispersion-based reconditioning provides considerable improvement. However, the common drawback of the above *postprocessing* schemes is that the spatial-temporal correlational information of the target in several consecutive pings of the sonar is neglected.

In this paper, a preprocessing scheme based on *multichannel subband adaptive filtering* and a postprocessing scheme based on the *high-order correlation* (HOC) method [24], [25] was proposed to improve the SRR before and after the beamforming process. Both schemes were designed to overcome the shortcomings of the existing methods while combining their advantages.

## II. TIME-DELAY ESTIMATION USING MULTICHANNEL SUBBAND ADAPTIVE FILTERING

In the shallow-water environment, strong reflections caused by the bottom and surface of the sea and those of biologics in

the water column present dominant reverberation effects. Thus, the primary challenge is to process the data and detect targets in such a highly reverberant environment. This may be done by processing each sonar channel separately and estimating the time delays associated with the reflections and then isolating them from the background noise prior to beamforming. The subband adaptive filtering scheme has been applied [11] to the problem of underwater target detection. Time delays are estimated iteratively in each subband using two different adaptation mechanisms that minimize the mean of squared error (MSE) between the reference and primary signals in the corresponding subband and level. The localization of the minima of the MSE curves at different levels and subbands are used in order to arrive at the time-delay estimates. However, as mentioned before, target detection in the shallow-water environment presents special problems due to the dominant reverberation effects in which the target returns may be buried. It is difficult to accurately estimate time delays using only one sonar channel. Therefore, a multichannel method is introduced which exploits the interdependency of adjacent sonar stave data.

### A. Modeling of the Backscattered Signal in the Shallow-Water Environment

In active sonar sensing systems [26], a known waveform of finite duration is generated which propagates through the medium and is reflected by some objects back to the point of origin. The transmitted signal is usually modified both in amplitude and phase depending on the target characteristics, which by themselves might be changing with time and position in space. These disturbances give rise to a random return or backscattered signal. In addition to these target-generated direct signals, there may also be spurious returns such as clutter from biologics and reverberation from the ocean surface or bottom, especially in the shallow-water environment. One of the characteristics of the shallow-water medium is the presence of strong *multipath* effects. Signals from a target can undergo reflections, creating multipath returns that are delayed, distorted, amplitude-weighted replicas of the direct signal to the sensor array. These signals can generate *coherent interference* which greatly increase the difficulty for accurate target detection.

For accurate time-delay estimation, the choice of a practical and reasonable model is critical. Usually, the backscattered signal is modeled as a sum of delayed and attenuated versions of the incident signal [26]. In this paper, however, we use a more generalized model which allows time-varying attenuation in the delay path. Let us consider the following model for each backscattered signal (or stave) which consists of the target returns, surface and volume reverberations, and additive background noise, i.e.,

$$y(n) = \sum_{i=1}^M z_i(n) + e(n) \quad (1a)$$

where

$$z_i(n) = h_i(n) * x(n - \Delta_i) \quad (1b)$$

and  $y(n)$  represents the backscattered signal,  $h_i(n)$  represents the unknown linear operation for the  $i$ th reflection

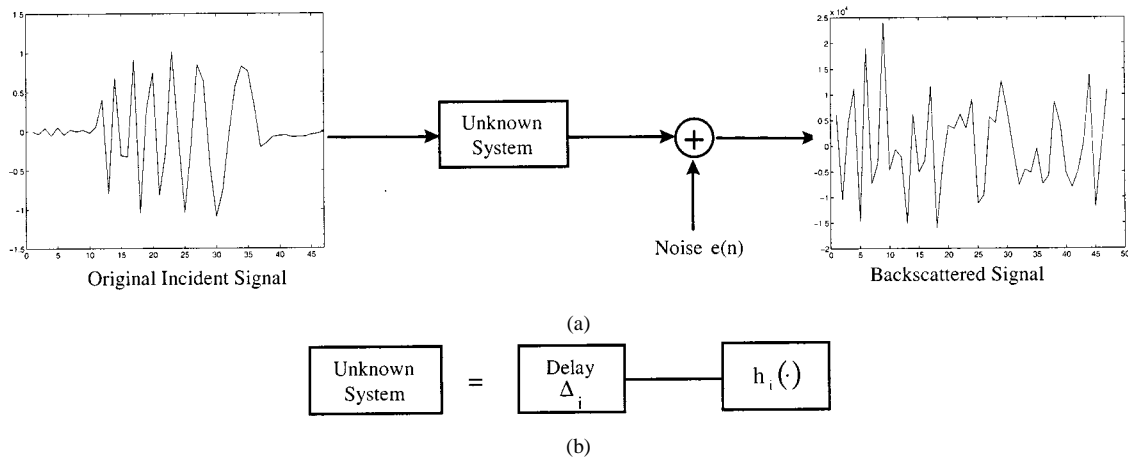


Fig. 1. (a) Effects of media and the target on the incident signal. (b) Unknown system represented by an unknown time delay  $\Delta_i$  and impulse response  $h_i(\cdot)$ .

corresponding to the effects of media and the scattering properties of the target on the incident signal,  $x(n)$  represents the incident signal,  $e(n)$  is the additive ambient noise, and the symbol  $*$  denotes the convolution operation. The unknown system consists of a pure time delay  $\Delta_i$  followed by a filtering operation with the impulse response  $h_i(n)$  [see Fig. 1(a)]. The reflection signals are modeled as the distorted or modified version of the incident signal and they may or may not be overlapping in time. Note that, in the above model, we don't distinguish the target returns from those of other reflections. By doing this, we assume that the target returns have similar characteristics as those of multipath and reverberations. However, one should bear in mind that the target mine is stationary and thus appears more consistently in the sensor data while the reverberation does not have this characteristic. Fig. 1 portrays the effects of the media and the scattering properties of the target on the incident signal.

Having modeled the process using (1), now in order to separate the specular returns from the backscattered signal, the first step is to accurately estimate the time delays  $\Delta_i$  associated with each return. This can be done by “undoing” the effects of the media by approximating the inverse response of the unknown linear system. The inverse filtered backscattered signal can then be compared with the delayed incident signal, and the time delays associated with the returns can be found by examining the local minima of the MSE curve. The adaptive filter in this case performs the inverse modeling of the unknown function  $h_i(n)$ . However, since the backscattered signal is quite noisy, accurate time-delay estimation based on simple adaptive filtering becomes very difficult. In order to overcome this problem, discrete wavelet transform (DWT) is employed in conjunction with adaptive filters to perform time-delay estimation in subbands.

### B. Time-Delay Estimation Using Subband Adaptive Filtering

The structure of the subband adaptive system designed for time-delay estimation is shown in Fig. 2. The process involves two steps. The first step is to apply signal decomposition using DWT to both the backscattered and incident signals. The reason for using the DWT decomposition is that the information about the time delays is common to all the levels and subbands. Ad-

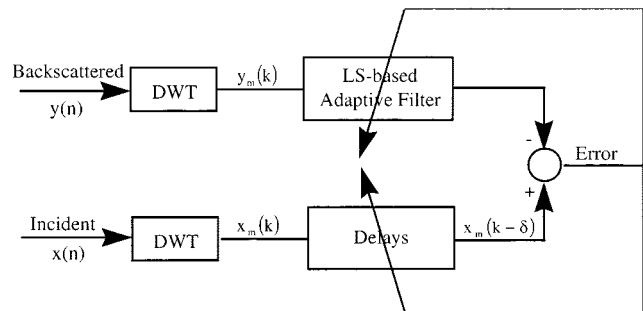


Fig. 2. Subband adaptive filtering.

ditionally, as a result of the filtering and the decimation processes at each level, the signals are less correlated than those at the original signal domain. Consequently, this method generates multiple “looks” of the signals at different subbands/levels and at the same time provides signal decorrelation and noise removal for better time-delay estimation [11].

Once the incident and backscattered signals are decomposed using the DWT, an adaptive filtering scheme is used in each subband to yield estimates of the time delays. This is done in each subband by applying the decomposed backscattered signal to the reference input of the adaptive filter while using a delayed version of the decomposed incident as a desired signal. Fig. 2 demonstrates the entire process in one subband. The delay  $\delta_i$  in each subband is incremented to provide an estimate of  $\Delta_i$ , i.e., the time delay associated with the  $i$ th specular return, where  $\Delta_i = 2^L \delta_i$ , and  $L$  is the number of DWT levels. For every new  $\delta_i$ , the function of the adaptive filter is to modify its weights in order to generate an output signal which minimizes the MSE at the output. In this way, the weights of the adaptive filter tend to capture the inverse response of the process,  $h_i(n)$ , in the relevant subband. At the right delay, the output of the adaptive filter provides the best match, hence leading to a minimum in the MSE curve. Thus, the analysis of the minima of the MSE curve provides the locations of all the possible time delays  $\delta_i^l$ 's. This can be done by using a simple thresholding operation of the MSE curve after the complete curve is generated. To account for the time lag introduced by the finite impulse response (FIR) adaptive

filter [27] at these positions, the lag associated with the largest weight is subtracted from the delay estimation in each level. The final delay value is then calculated using  $\Delta_i = 2^L(\delta_i - q)$ , where  $\delta_i$  is the estimated delay in the subband and  $q$  is the lag associated with the weight having the largest magnitude.

Let  $x_m(k)$  and  $y_m(k)$ ,  $m \in [1, L]$ ,  $k \in [1, N]$ , be subband decompositions of the incident and backscattered signals, respectively, where  $L$  is the total number of subbands at the lowest level of decomposition and  $N$  is the size of the signal in each subband. To estimate the time delays, the following cost function is minimized over all subbands:

$$\varepsilon(\delta) = \sum_{m=1}^L \sum_{k=1}^N \left( x_m(k - \delta) - \sum_{j=1}^P w_{mj} y_m(k - j + 1) \right)^2 \quad (3)$$

where the first term in the bracket is the delayed incident signal, the second term is the output of the adaptive filter, and  $P$  is the filter order. Minimizing  $\varepsilon(\cdot)$  with respect to  $w_{ir}$  yields the solution [27] shown in (4), at the bottom of the page, where  $R_{y_m y_m}(j)$  is the autocorrelation of the backscattered signal in subband  $m$  and  $R_{y_m x_m}(j)$  is the cross correlation between the backscattered and incident signals. Substituting this LS solution into the cost function (3) gives the expression for the overall minimum estimation error for each block of data as (5), shown at the bottom of the page, where  $R_{x_m x_m}(\cdot)$  is the autocorrelation of the incident signal and  $\varepsilon_m(\delta)$  is the MSE in the  $m$ th subband. The overall estimation error is the sum of all MSE's in each subband. Since the time-delay information is common to all subbands/levels, after the summation, the local minima in  $\varepsilon(\delta)$  corresponding to the possible targets will be emphasized while all the others will be deemphasized. The resulting curve  $\varepsilon(\delta)$  is then thresholded and the local minima below the threshold are picked as possible time-delay estimates.

### C. Multichannel Subband Adaptive Filtering

The Toroidal Volume Search Sonar (TVSS) uses a circular transducer array geometry to form a number of narrow beams from the 120 nearly omnidirectional elements where the center of each element is separated by  $3^\circ$  [1]–[3]. The incident was a 1-ms linear frequency modulated (LFM) signal. Each ping of TVSS data, representing one second in time, encompasses  $3^\circ$  along track out to a range of 675 m to port and starboard sides and to the ocean surface and bottom. The beams emanating from the transducer array are vertically narrow. Since the sensors in the array are interdependent, the targets appear consistently at about the same range cell in several adjacent staves with perhaps a minor shift which can be estimated using the geometry of the sensor array. Consequently, it is more appropriate to model the system as a multichannel process. In this case, in contrast to processing each sensor data independently, the interdependency among staves is fully exploited, hence accentuating the target returns while suppressing the effects of reverberations.

It has been shown [28] that simple phase-compensating and adding operations among the sensor array outputs result in significant improvement in the SNR by a factor equal to the number of sensors used. This is achieved by combining the desired signals coherently and the noise incoherently. The first step in our multichannel scheme is the phase compensation among adjacent sensors. The goal is to steer the sensor array toward a direction of interest. By summing these aligned signals, one achieves an enhancement of the desired signal while suppressing reverberations.

When the source is far away from the sensor array, the received signals can be considered to arrive at the sensors as plane waves. As depicted in Fig. 3, in the direction of interest, the plane wave arrives earlier at sensor  $j$  than sensor  $i$ . The distance between sensors  $j$  and  $i$  with respect to the direction of interest is  $d = r(\sin \theta_j - \sin \theta_i)$ , where  $r$  is the radius of the array of sensors. Then, the phase difference  $\phi$  among the adjacent staves is  $\phi = (2\pi d)/\lambda$ , where  $\lambda$  is the wavelength. With

$$\begin{bmatrix} w_{m1} \\ w_{m2} \\ \vdots \\ w_{mP} \end{bmatrix} = \begin{bmatrix} R_{y_m y_m}(0) & R_{y_m y_m}(1) & \cdots & R_{y_m y_m}(P-1) \\ R_{y_m y_m}(1) & R_{y_m y_m}(0) & \cdots & R_{y_m y_m}(P-2) \\ \cdots & \cdots & \cdots & \cdots \\ R_{y_m y_m}(P-1) & R_{y_m y_m}(P-2) & \cdots & R_{y_m y_m}(0) \end{bmatrix}^{-1} \begin{bmatrix} R_{y_m x_m}(\delta) \\ R_{y_m x_m}(\delta-1) \\ \vdots \\ R_{y_m x_m}(\delta-P+1) \end{bmatrix}, \quad m \in [1, L] \quad (4)$$

$$\begin{aligned} \varepsilon(\delta) &= \sum_{m=1}^L R_{x_m x_m}(0) - \sum_{m=1}^L \begin{bmatrix} R_{y_m x_m}(\delta) \\ R_{y_m x_m}(\delta-1) \\ \vdots \\ R_{y_m x_m}(\delta-P+1) \end{bmatrix}^t \\ &\quad \times \begin{bmatrix} R_{y_m y_m}(0) & R_{y_m y_m}(1) & \cdots & R_{y_m y_m}(P-1) \\ R_{y_m y_m}(1) & R_{y_m y_m}(0) & \cdots & R_{y_m y_m}(P-2) \\ \cdots & \cdots & \cdots & \cdots \\ R_{y_m y_m}(P-1) & R_{y_m y_m}(P-2) & \cdots & R_{y_m y_m}(0) \end{bmatrix}^{-1} \begin{bmatrix} R_{y_m x_m}(\delta) \\ R_{y_m x_m}(\delta-1) \\ \vdots \\ R_{y_m x_m}(\delta-P+1) \end{bmatrix} \\ &= \varepsilon_1(\delta) + \varepsilon_2(\delta) + \cdots + \varepsilon_L(\delta) \end{aligned} \quad (5)$$

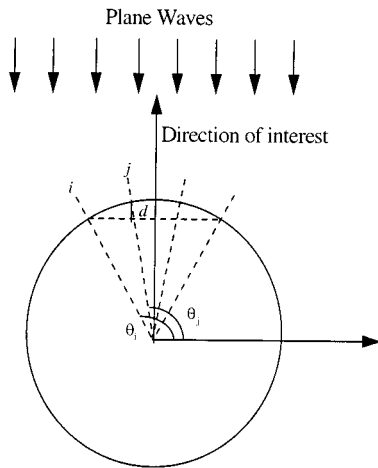


Fig. 3. Phase correction between adjacent sensors.

the knowledge of the geometry of the sensor array, one can exactly decide  $d$  between adjacent sensors and the corresponding phase difference  $\phi$  using the above two equations. By compensating for the phase difference, it is possible to combine the data of several sensors together along a direction of interest. In order to do this, the data in each stave are first preprocessed to form the corresponding analytic signal or *pre-envelope* or *baseband reduced complex envelope*,  $s(t) = s_r(t) + j\hat{s}_r(t)$ , where  $s_r(t)$  is the backscattered signal and  $\hat{s}_r(t)$  is the Hilbert transform of  $s_r(t)$ . The phase-compensated signal is  $s^{\text{comp}}(t) = s(t)e^{j\phi}$ . As the real part of the pre-envelope corresponds to the time-domain signal, the stave after the phase correction/compensation is  $s_r^{\text{comp}}(t) = \text{Re}\{s^{\text{comp}}(t)\}$ .

After the phase compensation among the adjacent staves with respect to the direction of interest, the related staves are weighted and summed up to form a combined stave for more accurate time-delay estimation using subband adaptive filtering. For the TVSS, since the beams are relatively narrow, it was found that only three or four sensors may contain the same target returns. Consequently, only four channel outputs are used.

### III. TEST RESULTS OF MULTICHANNEL SUBBAND ADAPTIVE FILTERING

Both simulated and real TVSS data were used in this testing. The TVSS data were collected in the Gulf of Mexico in November 1994 [1]–[3]. The testing on the simulated data demonstrates the accuracy of the proposed schemes and was also used to fine tune the algorithm for the testing on the real data. The testing on real data sets represents the real-life use of the method for operation in conjunction with TVSS in the shallow-water environment. The testing results are achieved in three steps. First, the multichannel phase compensation is applied to get a combined signal with enhanced target returns and reduced clutter. Note that this step only applies to multichannel processing. For single-channel processing, we skip this step. Second, we apply the subband adaptive filtering scheme to the combined signal to obtain accurate time-delay estimation and generate the output with target returns. Finally,

the beamforming process is applied to the output data to generate the final image to identify the possible targets.

#### A. Parameter Selection

For processing the TVSS data, the backscattered signal is divided into overlapping blocks with the overlap equal to the length of the incident signal. DWT is then applied to both the backscattered and incident signals. After the signals are decomposed, the time delays are estimated in each subband using the LS-based subband adaptive filter described before. Once the time delays are estimated in each block, the associated returns are separated from the backscattered signal by simply extracting the corresponding block of the same length as the incident signal.

There are several parameters in the adaptive filtering process, such as block size, number of levels of DWT decomposition, and adaptive filter order that need to be optimized in order to achieve the best performance.

1) *Block Size*: The proposed scheme is applied to blocks of the backscattered data to increase the processing speed. The block-based method also allows detection of low amplitude or weak returns which will be quite insignificant in a long data record. Due to the nonstationarity of the backscattered signal, the block size cannot be too large in comparison with the length of the incident signal (which is less than 50 samples in the TVSS case). Thus, the block-based approach not only increases the processing speed but also takes into account the nonstationarity of the backscattered signal, leading to a more accurate time delay estimation. The block size of 1024 was empirically determined to be the optimal choice and was used throughout the testing process.

2) *DWT Decomposition Levels*: As mentioned earlier, the benefit in using DWT decomposition is that it provides multiloops of the same signal at different levels. It is well understood that increasing the number of decomposition levels not only makes the signals more decorrelated but also suppresses the noise to a greater extent. Nonetheless, as the number of levels increases, the time resolution is halved at each level. This can result in deterioration of the accuracy of time-delay estimation. This is especially true for the TVSS data because the incident signal is only 47 samples long. It was empirically determined that  $L = 2$  is the optimal number of levels for the 1-ms LFM incident chosen in these cases. In addition, downsampling is not used in our decomposition to avoid inaccuracies in estimating delays that are not multiples of 2. Daubechies wavelet [29] of order 6 was used to perform DWT operation. Daubechies wavelets are orthogonal, compactly supported with frequency response which is maximally flat at  $\Omega=0$  and  $\Omega=\pi$ .

3) *Adaptive Filter Order*: The adaptive filter order is the most important parameter in the whole process as it determines the target detectability and the accuracy of time delay estimation. Several experiments were conducted in order to find the optimal filter order. It was found that, by increasing the order of the FIR filter, the detectability of the target peak improved significantly. However, increasing the order ( $>6$ ) widens the regions in which the minima occur, hence reducing the detectability of these minima, which in turn reduces the

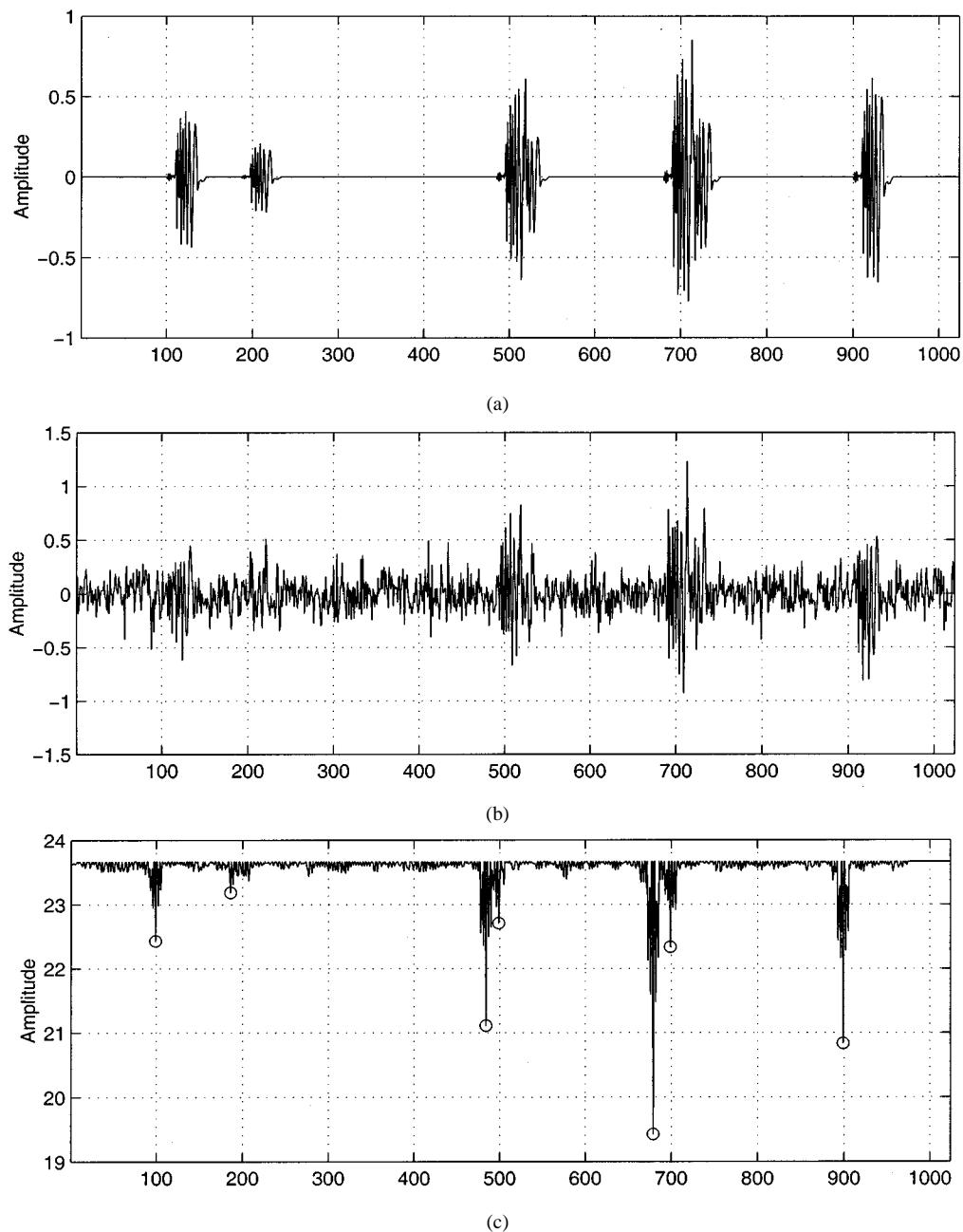


Fig. 4. (a) Clean simulated backscattered signal. (b) Noisy simulated backscattered signal. (c) Combined MSE curve and picked local minima.

accuracy of delay estimation. Additionally, large-order filters require more computational efforts. In processing the TVSS data, a filter of order 6 was used.

4) *Adaptive Threshold*: The choice of the MSE threshold is also crucial in this algorithm. If the threshold is too high, a large number of local minima are picked which results in less reverberation removal and clutter rejection in the processed results; while, for very low threshold, the local minima corresponding to the target might be missed altogether. In our proposed scheme, the threshold is calculated adaptively using  $\text{Threshold} = (1 - \alpha) \max(\epsilon) + \alpha \min(\epsilon)$  where  $\epsilon$  is the combined MSE,  $\alpha$  is a ping-dependent constant, and its optimal value can be determined empirically.

#### B. Simulations on Synthesized Backscattered Data

To show the accuracy of this scheme in time-delay estimation, the subband adaptive filtering method was applied to a simulated backscattered signal. For simplicity, a signal of length 1024 (one block) was generated with delayed and attenuated versions of the incident signal [Fig. 4(a)] to simulate different returns. The actual values of the time delays are shown in Table I. As can be seen, the third and fourth returns, fifth and sixth returns are overlapping in time. White Gaussian noise was added to simulate the effects of the channel noise. Fig. 4(b) shows the noisy simulated backscattered signal. The first two returns can barely be seen after adding the noise. The SNR

TABLE I  
SNR VALUES OF DIFFERENT RETURNS AND THEIR LOCATIONS IN THE SIMULATED DATA.

Targets	1	2	3	4	5	6	7
SNR (dB)	3.2637	-2.7569	5.2019	0.7649	8.1244	3.2637	6.7855
True Locations	100	187	485	500	680	700	900
Estimated Locations	100	187	485	500	680	700	900

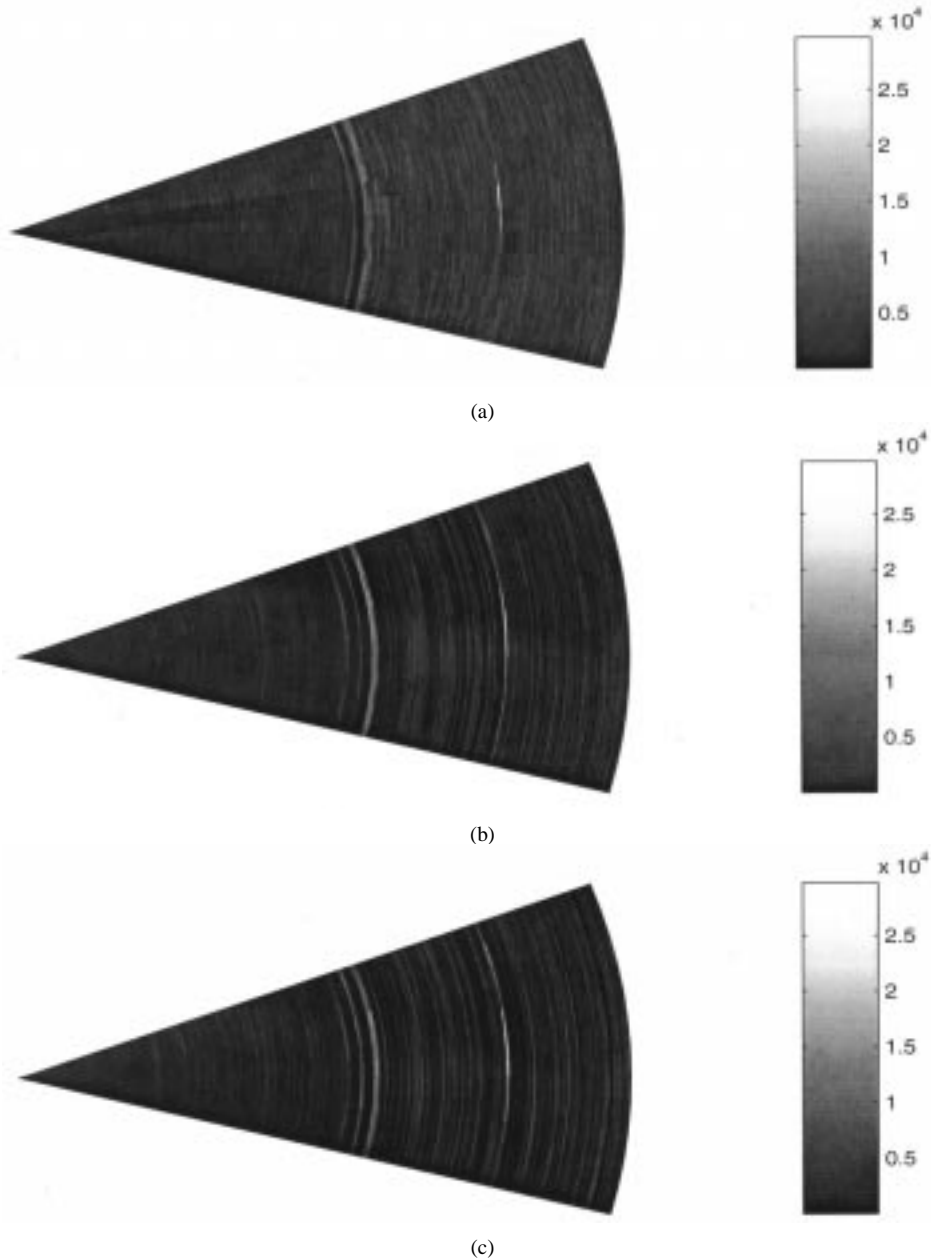


Fig. 5. Beamforming results for ping 0631. (a) Original data. (b) Single-channel processed data. (c) Multichannel processed data.

values associated with the returns are also shown in Table I. Fig. 4(c) shows the combined MSE resulting from the subband adaptive filtering scheme. Those local minima marked with a “o” show the exact locations of the returns. Table I shows the estimated locations. From the results, we can see that, even though the SNR for some returns were very low, e.g., the SNR for the second return was  $-2.7569$  dB, the scheme was still

able to correctly estimate its location. This is true even when the returns are overlapping in time. One may also notice that the values of the local minima depend on the amplitudes of the backscattered signal. The higher the amplitude, the more obvious the local minima in the overall combined MSE curve. The threshold needs to be chosen properly so that all the local minima corresponding to returns can be picked up.

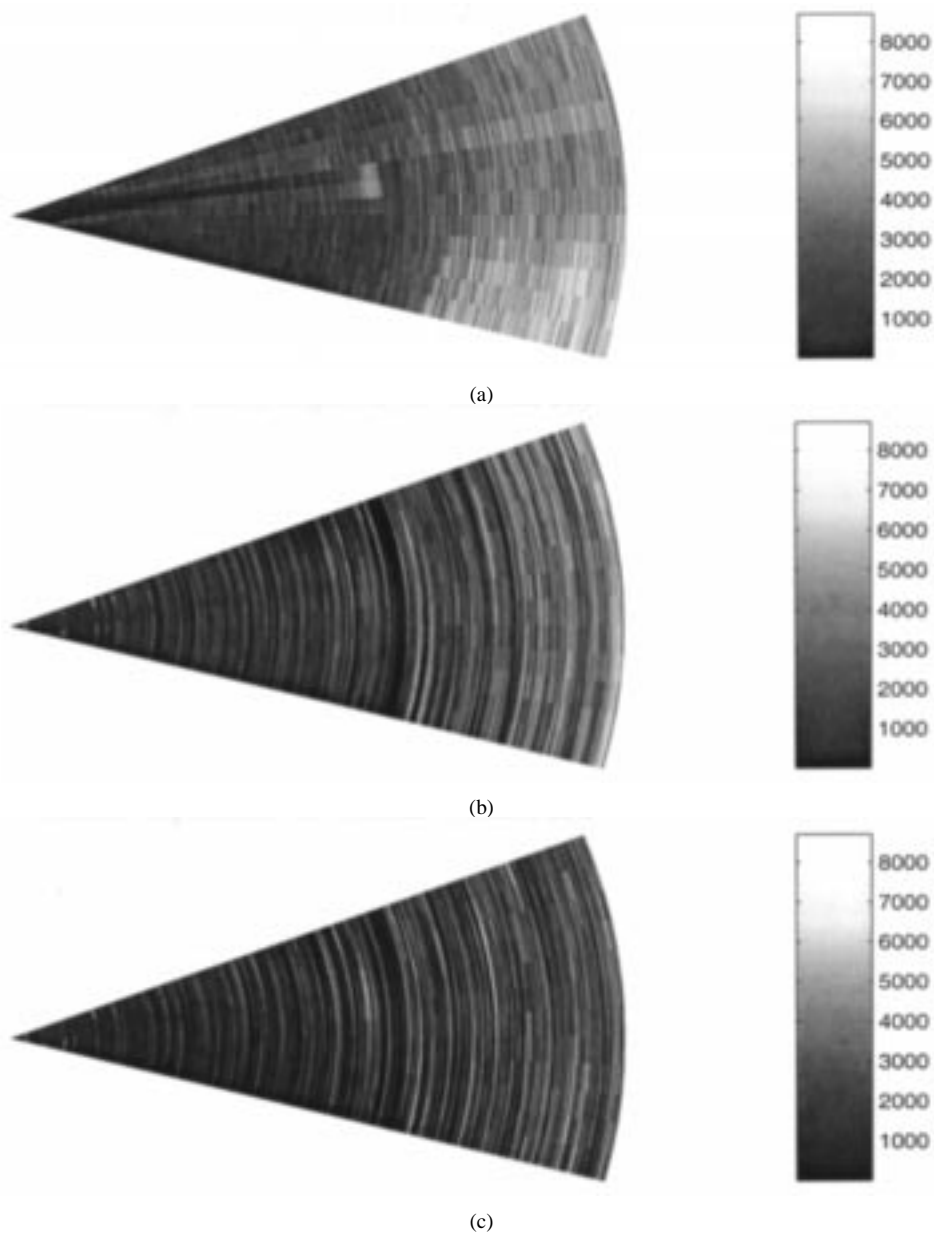


Fig. 6. Beamforming results for ping 0625. (a) Original data. (b) Single-channel processed data. (c) Multichannel processed data.

### C. Processed versus Unprocessed TVSS Beamformed Results

The 18 pings of TVSS data are divided into three groups according to the significance of the reverberation effects. In addition, biologics, bottom features like coral or rock outcropping, and man-made litter of all descriptions can cause difficulties in analyzing the backscattered signals. Some typical images of the processed versus the unprocessed beamformed results are shown in Figs. 5–7 for different levels of reverberation. Part (a) in each figure shows the beamformed image of original or unprocessed data; part (b) shows the image of single-channel sub-band adaptive filtering results, while part (c) shows the multi-channel processed result. Only those peaks that appear consistently in several consecutive pings are considered as potential target candidates.

1) *Low Reverberation Cases:* In these cases, the targets can be clearly seen in the beamformed images even without the pre-

processing, and they are generally the brightest returns in the images. Nonetheless, in the unprocessed beamformed results of some pings, the targets are only prominent in one beam while in the processed beamformed result the target peaks are quite obvious in several adjacent beams. Fig. 5(a), (b) and (c) show the original, single-channel, and multi-channel processed beamformed images for one ping, respectively. For this ping, the true target appears on the starboard side at range cell 380, beams 92–95. Note that, if the circular array with 120 elements is traversed counterclockwise, the sensor at the 12 o'clock position (looking up) is labeled as the first element while the ones at the 9 o'clock (Port Side) and 3 o'clock (Starboard Side) positions, i.e., the 31st and 91st, are looking into the water volume. As can be seen, for this case, the single-channel processed result is comparable to that of the original beamformed result while the target peak is more prominent in the multichannel processed results. From these images, one can see that both the single



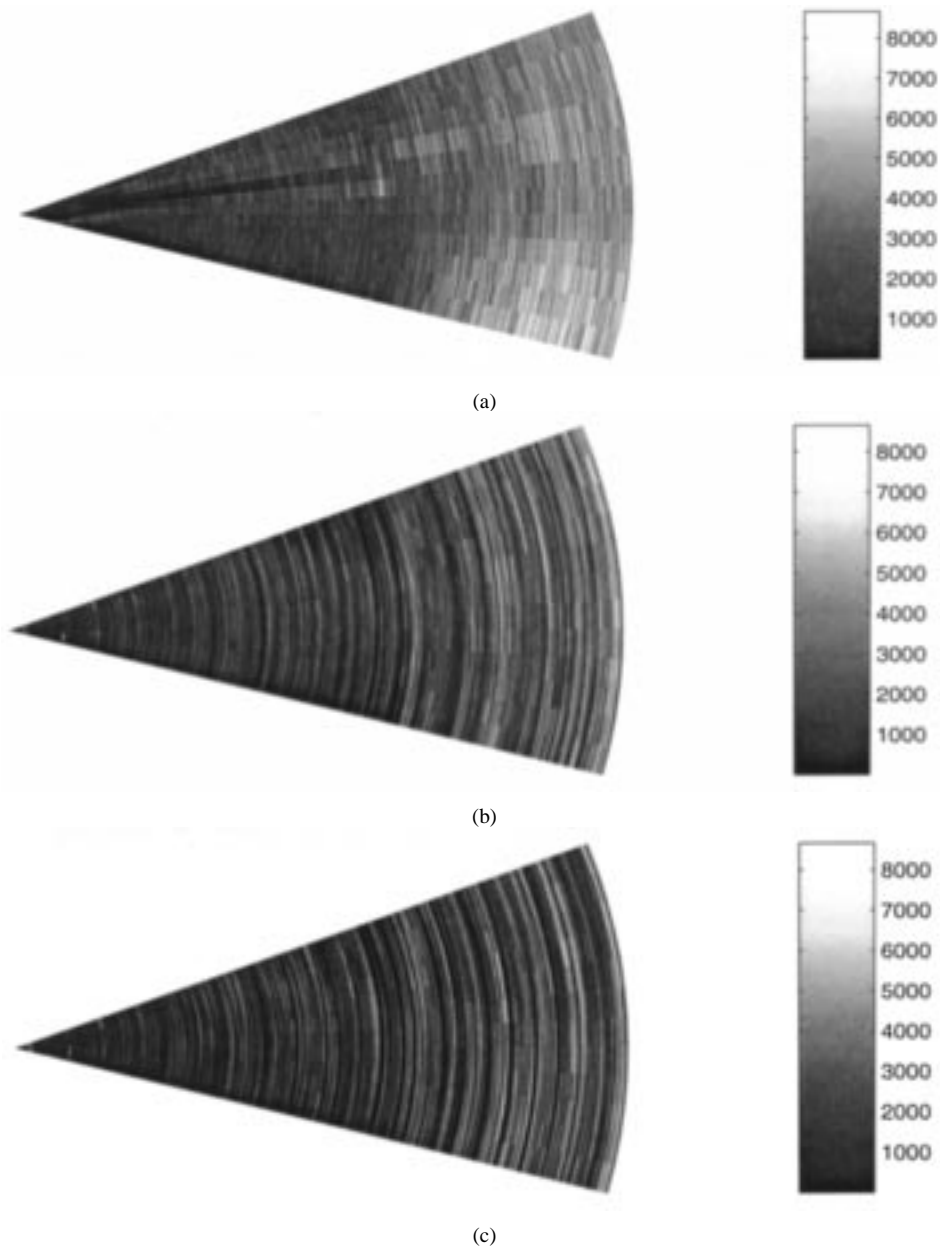


Fig. 7. Beamforming results for ping 0627. (a) Original data. (b) Single-channel processed data. (c) Multichannel processed data.

and multichannel processed results have enhanced the target returns while greatly reducing the clutter and reverberation effects at a farther range. Thus, in these cases, single-channel processing provides comparable or better results than the unprocessed beamformed results while the multichannel processing provides the best results.

2) *Moderate Reverberation Cases:* The reverberation effects are much higher in these cases than in the previous cases and the targets can barely be observed in the original beamformed results. Fig. 6(a), (b) and (c) show the original, single-channel, and multichannel processed beamformed results of ping 0625, respectively. This ping has the same target as shown previously in Fig. 5. However, due to the motion of the vehicle, the target position is slightly shifted in the beam while the range is approximately the same. The target peak cannot easily be detected in the original beamformed image.

In the single-channel result, one can see from the image that the target return has been enhanced. This is more obvious when comparing with the original beamformed results, but there is still substantial clutter which makes the correct target detection difficult. In Fig. 6(c), the beamformed result of the multichannel subband adaptive filtering exhibits significantly less reverberation effects and the target peak can easily be identified. Again, both the single-channel and multichannel processed results show significant improvements in the target detectability while the multichannel processing is more effective than the single-channel processing.

3) *High Reverberation Cases:* In these cases, the reverberation effects are very high and the beamformed results of the unprocessed data are extremely noisy. Fig. 7(a), (b), and (c) show the original, single-channel, and multi-channel processed results, respectively. From Fig. 7(a), it is virtually impossible to

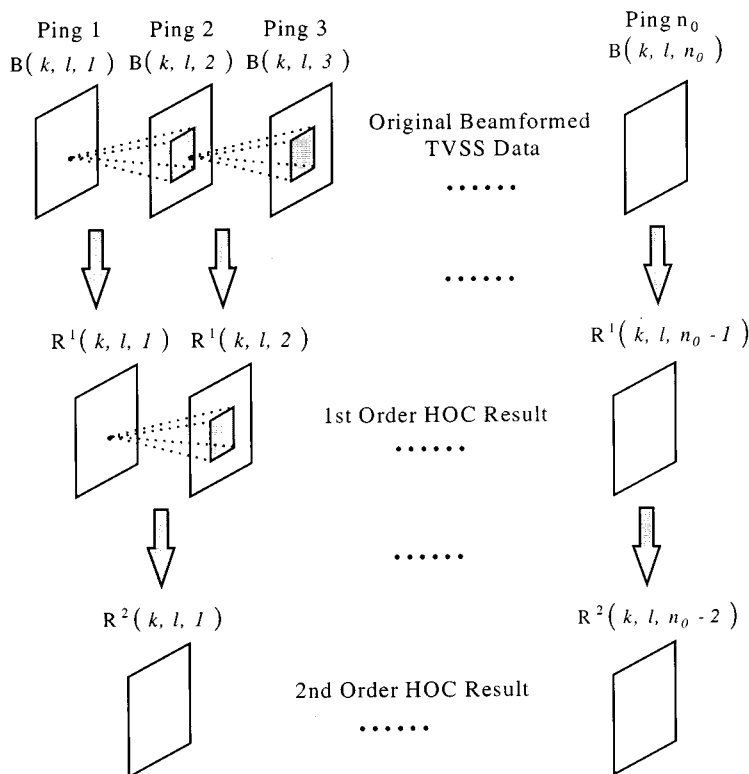


Fig. 8. Computing HOC results of different orders.

determine where the target location might be from the original beamformed image. The single-channel result in Fig. 7(b) shows no visible improvement. The multi-channel processing, however, made the target return observable. Though image background is still noisy, the target detectability is greatly improved. Thus, in the high reverberation cases, the multichannel processing demonstrates much better performance than the single-channel processing.

In summary, the test results of the subband adaptive filtering scheme on the real TVSS data show improvements in the overall target detectability. Single-channel processing performs well in the low and moderate reverberation cases but fails to enhance the target returns in the high reverberation cases. Multichannel processing, on the other hand, performs well in all the scenarios.

#### IV. HIGH-ORDER CORRELATION FOR POSTPROCESSING

Clearance and avoidance of mines and mine-like objects rely on accurate target detection. Generally, to cover a large area, TVSS needs to collect 800–900 pings of data per run [22]. This workload is beyond the capabilities of human operators. In order to postprocess the TVSS beamformed results automatically and detect all the possible targets while maintaining a false alarm rate as low as possible, a method based on high-order correlation (HOC) is developed. The original HOC method was proposed [24], [25] to detect multiple dim target tracks in heavily cluttered background from infrared (IR) satellite data. This method exploits the fact that a moving target builds a spatial-temporal track in the three-dimensional (3-D) space and that there exists

spatial-temporal correlations between consecutive target observation points. This method does not make any *a priori* assumption about the number of targets, the dynamical information of the target, initial conditions, or background clutter. It allows recursive computation of the *spatio-temporal* cross correlation among consecutive data sets.

The use of HOC as a postprocessor for beamformed TVSS data was initiated by the following facts. First, the along-track coverage of the TVSS widens with distance, hence causing a target to appear in multiple consecutive pings. HOC is capable of determining the temporal and spatial dependencies of consecutive pings of data. Second, no *a priori* information is available about the targets and clutter/noise in the TVSS data. The goal of the HOC approach in postprocessing the TVSS beamformed data is to determine the *consistency* in the target locations in several consecutive pings of the beamformed results. The basic assumptions behind this approach are: 1) targets should consistently appear in a limited area (*beam*  $\times$  *range*) in several consecutive pings and 2) there might be more than one target in the searching region. In addition, the targets may be barely observable or even missed in certain pings due to the movement of the vehicle on which the TVSS is deployed.

##### A. HOC Algorithm for TVSS Shallow-Water Application

In order to apply the HOC method, the beamformed results in the information bearing portion were first converted to *binary* values. The information bearing portion corresponding to beams 26–37 and 86–97 is also called the along-track direction *port side* and *starboard side*, respectively.

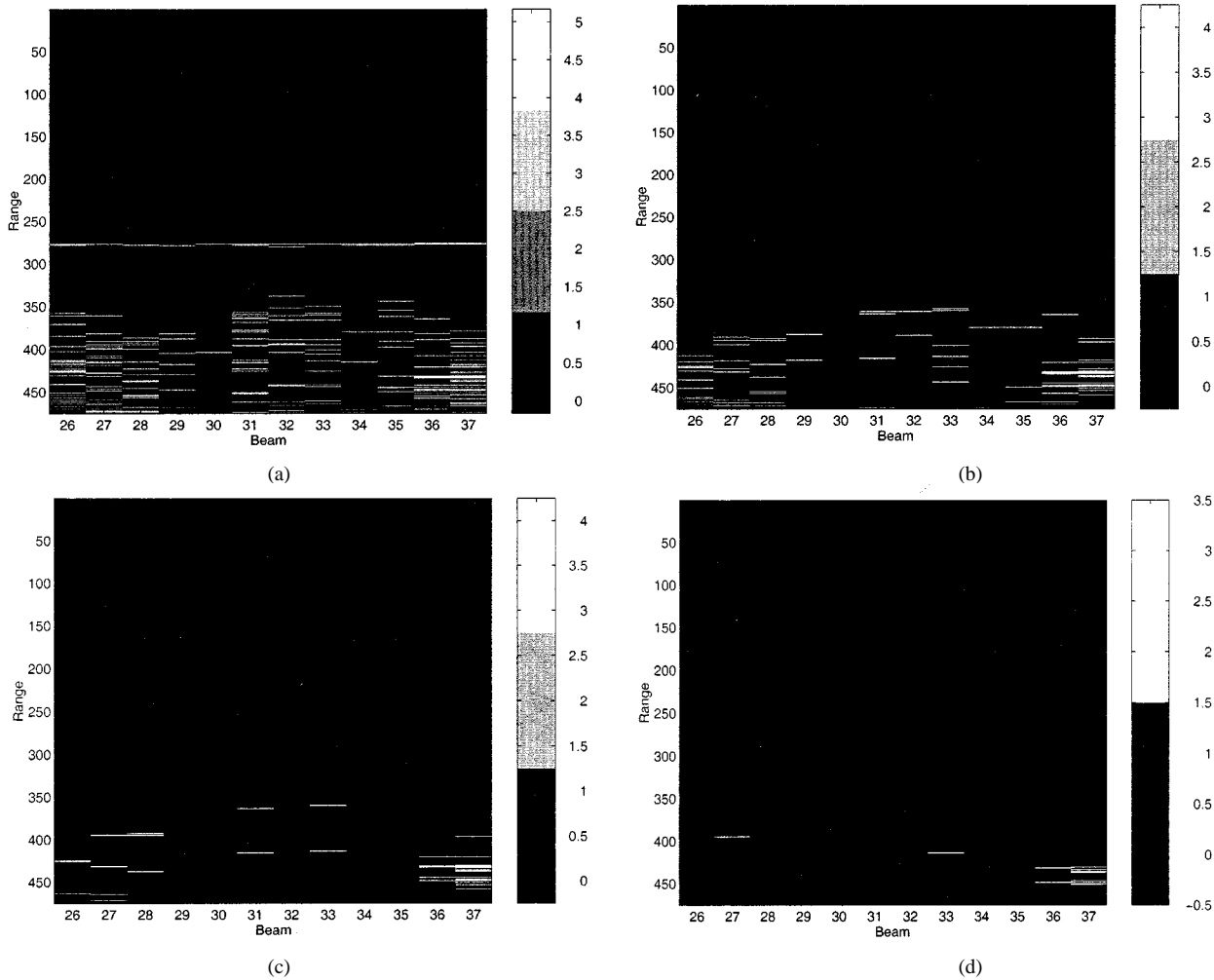


Fig. 9. HOC results—port side. (a) Target consistency count. (b) First-order correlation. (c) Second-order correlation. (d) Third-order correlation.

1) *Target Consistency Count*: The target consistency count (TCC) is a measure of consistency in the occurrence of a point in the beamformed image. It indicates how frequently a certain point might be detected as a possible target location in several consecutive pings. TCC can be obtained by simply stacking and summing up the corresponding points of the beamformed results in several consecutive pings in the region of interest. TCC can be viewed as the zeroth-order HOC. TCC can provide some information about the consistency in the occurrence of targets if the movement of the vehicle is stable. However, the target location on the beamformed image changes from ping to ping, especially when the vehicle yaws and rolls. When this happens, TCC provides less reliable information about the true target location. Thus, we need to calculate HOC of higher orders to overcome this problem.

2) *HOC of Different Orders*: Since the target is stationary in the water column and the vehicle moves toward certain directions, there is spatial-temporal dependency between the adjacent pings of data. To find such dependencies, cross correlations between adjacent pings can be calculated. However, this computes only the spatio-temporal correlation between two consecutive pings and, hence, no memory is built into the process. The HOC method solves this problem by allowing one to calculate the cor-

relations among several consecutive pings recursively [24] as opposed to only two pings.

The first-order correlations can be calculated using

$$R^1(k, l, n) = g \left[ \sum_{i=-M}^M \sum_{j=-N}^N B(k, l, n) B(k+i, l+j, n+1) \right] \quad (8a)$$

where  $g(\cdot)$  is the standard hard limiter thresholding function, i.e.,

$$g(x) = \begin{cases} 1, & x > 0 \\ 0, & x \leq 0 \end{cases} \quad (8b)$$

$B(k, l, n)$  is the beamformed result of ping  $n$  at position  $(k, l)$ , i.e., beam  $k$  and range  $l$ . Thus,  $R^1(k, l, n)$ , which represents the result of the first-order correlation, provides information on how the points of  $B(k, l, n)$  at ping  $n$  are correlated to their neighboring points  $B(k+i, l+j, n+1)$  at ping  $n+1$ . The correlation is evaluated in a window of size  $(2M+1) \times (2N+1)$ . This window size is chosen under the assumption that the target location changes (due to the vehicle's movement) from one ping to the next do not exceed certain limits. Clearly, if  $R^1(k, l, n) = 1$ , then there is a two-point spatio-temporal sequence initiated

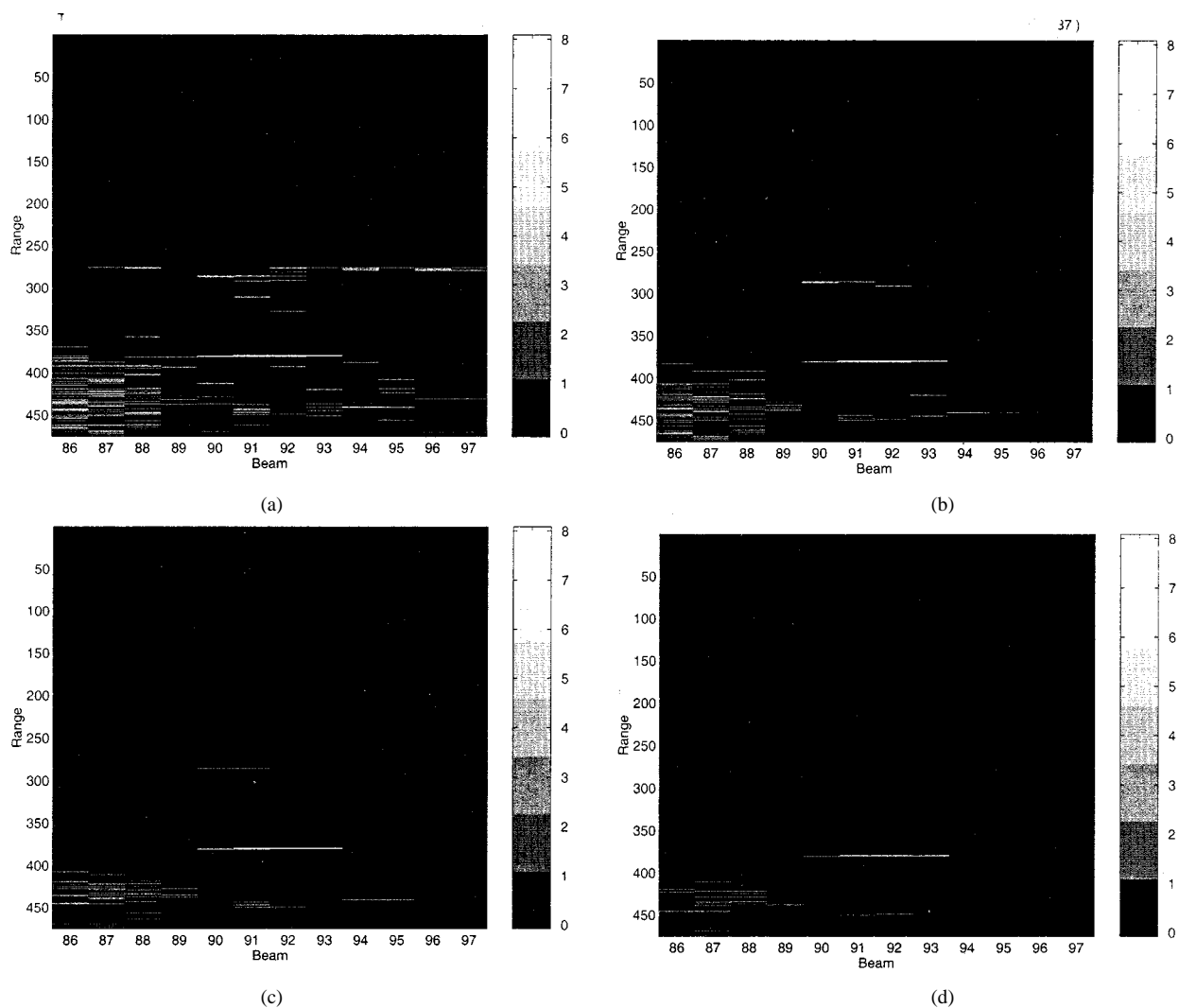


Fig. 10. HOC results—starboard side. (a) TCC. (b) First-order correlation. (c) Second-order correlation. (d) Third-order correlation.

at location  $(k, l)$  of ping  $n$  to location  $(k + i, l + j)$  of ping  $n + 1$ . Although this process can be repeated to identify all such two-point sequences, it cannot provide correlational information of more than two pings of data. However, in order to declare a detection, we need to verify if the target shows consistently in several consecutive pings of data. To calculate the correlation in more than two consecutive pings, the HOC approach builds memory into the process and computes correlations of  $R$  recursively using

$$R^p(k, l, n) = g \left[ \sum_{i=-M}^M \sum_{j=-N}^N R^{p-1}(k, l, n) R^{p-1}(k+i, l+j, n+1) \right] \quad (9)$$

where  $p$  is the order of the correlation and  $n$  is the index of pings which varies from 1 to  $n_{\max}$ . When the order  $p$  increases,  $n_{\max}$  should decrease such that  $p + n_{\max} = n_0$  for computing correlations among  $n_0$  consecutive pings. The initial condition for the above recursive equation is  $R^0(k, l, n) = B(k, l, n)$ . If  $R^p(k, l, n) = 1$ , that means a possible target location appears

consistently in  $(p + 1)$  consecutive pings [24]. The process of calculating different orders of HOC results is shown in Fig. 8.

The choice of  $p$  is critical in the HOC process since it presents a tradeoff between clutter removal capability of the HOC and its sensitivity to missing target points at certain pings. More specifically, large  $p$  gives better clutter rejection but at the same time increases the likelihood of losing the target point in the final HOC result when there are several missing target points in some of the pings. The effect of  $p$  can be seen clearly in the following test results.

#### B. Results of Applying HOC to TVSS Beamformed Data

To test the performance of the HOC method, several sequences of pings of the TVSS beamformed data were studied (see Figs. 9 and 10). The processed results of the HOC process are images in which the points satisfy the recursive correlation criterion and the background clutter and reverberation are substantially removed. The horizontal axis represents the beams in the region of interest (26–37 for the port side and 86–97 for the starboard side), while the vertical axis shows the range cell (after 1 out of 47 peak-picking process) from 1 to 475 for both sides. Generally, the targets appear in the

center of the along-track direction, i.e., beams 28–32 for the port side and beams 88–92 for the starboard side. A detection is generally declared if a point appears consistently in three out of four consecutive pings.

The results for one sequence of pings for both the port side and starboard side are presented here. For this group of data, there is no target on the port side. On the starboard side, there is a target located at range 380, beams 91–93. Fig. 9(a) shows the result of TCC on the port side. Figs. 9(b)–(d) show the HOC results for  $p = 1, 2$ , and  $3$ , respectively. Close examination of these results indicates that there is nothing prominent (showing consistently in the beam and range) for any order ( $p = 1, 2$ , and  $3$ ) of HOC results. This implies that there is no target present on the port side. Although there is a peak at about range cell 420 and beams 36–37, it is not in the along-track direction; hence, it is not considered to be a target. This result does indeed correspond to the ground truth information. Consequently, the HOC process for this sequence of pings does not result in any false detection.

Fig. 10(a) shows the results of TCC for the starboard side. We can see that there are many peaks indicating possible target locations. The two most prominent ones are around the range cells 380 (beams 90–95) and 286 (beams 90–92). According to the ground truth information, the true target location should be in the range cell 380 (beams 90–95) while the clutter at range cell 286 was believed to be associated to a biologic. Fig. 10(b)–(d) show the HOC results for  $p = 1, 2$ , and  $3$  on the starboard side, respectively. The result of the first-order HOC in Fig. 10(b) shows much fewer peaks than that in Fig. 10(a). Note that, though the competing clutter around range cell 286 was still there, it became much dimmer than that in the TCC result and no longer competitive with the peak around range cell 380. In the second-order HOC result (Figs. 10(c)), more clutter was removed. In the third-order HOC result (Fig. 10(d)), the competing clutter was totally removed and the only prominent peak at the range cell 380, which corresponds to the true target position, was kept. In this set of pings, the HOC process led to correct detection with no false detection.

In conclusion, the simulation results showed the promise of the HOC method as a postprocessor for TVSS shallow-water data for removing the competing clutter and at the same time boosting the target returns in several consecutive pings. The algorithm is simple, easy to implement, and fast. For example, processing all 31 pings took only about 18 s of CPU time on a HP 700 series workstation. Apart from the window size, which limits the maximum deviation of the vehicle, no other constraints pertaining to the dynamical information of the vehicle, such as velocity, are incorporated into the process. For the TVSS shallow-water application, if the dynamical information of the vehicle is available, the HOC results may be significantly improved by building additional constraints into the process.

## V. CONCLUSION

In this paper, a preprocessing multichannel subband adaptive filtering scheme was developed that exploits the interdependency among the signals recorded by adjacent sensors in the toroidal array and forms a combined signal with improved SNR.

Taking advantage of this improved SNR, subband adaptive filtering achieves very accurate time-delay estimation. The performance of this scheme was demonstrated on simulated data as well as real TVSS shallow-water data containing several targets at moderate to far range with moderate to extremely high reverberation. The HOC-based postprocessing method was also developed to find the consistency of the target appearance in the beamformed output. This process greatly enhances the target returns, removes the competing clutter and reverberation, and provides the exact locations (range and beam) of the targets. It provides an efficient and simple way to detect low observable targets in extremely high clutter.

## REFERENCES

- [1] J. D. Lathrop, "High area rate reconnaissance (HARR) and mine reconnaissance/hunter (MR/H) exploratory development programs," in *Proc. SPIE*, vol. 2496, Orlando, FL, Apr. 1995, pp. 350–356.
- [2] S. F. Castelin *et al.*, *Exploratory Development Minehunting Sensors for Unmanned Vehicles* [Online]. Available: <http://www.ncsc.navy.mil/css/papers/uuv96.html>
- [3] R. J. McDonald and J. Wilbur, "Adaptive beamforming algorithm for mine detection in shallow water," in *Proc. SPIE*, vol. 2765, Apr. 1996, pp. 549–556.
- [4] E. J. Tacconi and J. J. Carasi, "Inverse filtering approach for improving sonar signal resolution and amplitude estimation," *Appl. Acoustics*, vol. 32, pp. 297–309, 1991.
- [5] P. A. Delaney and D. O. Walsh, "Performance analysis of the incoherent and skewness matched filter detectors in multipath environments," *IEEE J. Oceanic Eng.*, vol. 20, pp. 1284–1296, Jan. 1995.
- [6] T. G. Manickam and R. J. Vaccaro, "Multipath time-delay estimation for long data record," in *1995 IEEE Int. Conf. on Acoustics, Speech and Signal Processing*, 1995, pp. 3143–3146.
- [7] T. G. Manickam, R. J. Vaccaro, and D. W. Tufts, "A least-squares algorithm for multipath time delay estimation," *IEEE Trans. Signal Processing*, vol. 42, no. 11, pp. 3229–3233, Nov. 1994.
- [8] D. Boudreau and P. Kabal, "Joint time-delay estimation and adaptive recursive least squares filtering," *IEEE Trans. Signal Processing*, vol. 41, pp. 592–601, Feb. 1993.
- [9] T. J. Lim and M. D. Macleod, "Adaptive algorithm for joint time delay estimation and IIR filtering," *IEEE Trans. Signal Processing*, vol. 43, pp. 841–851, Apr. 1995.
- [10] M. R. Azimi-Sadjadi, J. Wilbur, and G. J. Dobeck, "Isolation of resonance in acoustic backscatter from elastic targets using adaptive estimation schemes," *IEEE J. Oceanic Eng.*, vol. 20, pp. 346–353, Oct. 1995.
- [11] M. R. Azimi-Sadjadi, S. Charleston, J. Wilbur, and G. J. Dobeck, "A new time delay estimation in sub-bands for resolving multiple specular reflections," *IEEE Trans. Signal Processing*, vol. 46, pp. 3398–3403, Dec. 1998.
- [12] M. Hasan and M. R. Azimi-Sadjadi, "A modified block FTF adaptive algorithm with applications to underwater target detection," *IEEE Trans. Signal Processing*, vol. 44, pp. 2172–2185, Sept. 1996.
- [13] C. L. Nikias and A. P. Petropulu, "Higher-order spectra analysis," in *Signal Processing Series*. Englewood Cliffs, NJ: Prentice-Hall, 1993.
- [14] G. B. Giannakis and M. K. Tsatsanis, "Signal detection and classification using matched filtering and higher order statistics," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, July 1990.
- [15] R. F. Dwyer, "Use of the kurtosis statistic in the frequency domain as an aid in detecting random signals," *IEEE J. Oceanic Eng.*, vol. OE-9, pp. 85–92, Apr. 1984.
- [16] A. M. Richardson and W. S. Hodgkiss, "Bispectral analysis of underwater acoustic data," *J. Acoust. Soc. Amer.*, vol. 96, no. 2, pp. 828–836, Aug. 1994.
- [17] M. I. Doroslovacki and H. Fan, "Wavelet-based linear system modeling and adaptive filtering," *IEEE Trans. Signal Processing*, vol. 44, no. 5, pp. 1156–1167, May 1996.
- [18] N. Erdol and F. Basbug, "Wavelet transform based adaptive filters: Analysis and new results," *IEEE Trans. Signal Processing*, vol. 44, no. 9, pp. 2163–2171, Sept. 1996.
- [19] S. D. Marco and J. Weiss, "Improved transient signal detection using a wavepacket—Based detector with an extended translation-invariant wavelet transform," *IEEE Trans. Signal Processing*, vol. 45, no. 4, pp. 841–850, Apr. 1997.

- [20] X. L. Xu, A. H. Tewfik, and J. F. Greenleaf, "Time delay estimation using wavelet transform for pulse-wave ultrasound," *Annals Biomed. Eng.*, vol. 23, pp. 612–621, 1995.
- [21] A. C. Dubey, C. Yuan, and M. R. Azimi-Sadjadi, "Comparison of two different wavelet-based approaches for target detection in active sonar data," in *Proc. SPIE*, vol. 3079, Apr. 1997, pp. 225–230.
- [22] C. J. Robertson, "Statistical analysis of sonar data for target detection," in *Proc. SPIE*, vol. 2765, Apr. 1996, pp. 130–136.
- [23] R. J. McDonald and J. Wilbur, "Detection processing of complex beamformer output data, a new dispersion-based reconditioning algorithm," in *Proc. SPIE*, vol. 2765, Apr. 1996, pp. 157–166.
- [24] R. J. Liou and M. R. Azimi-Sadjadi, "Dim target detection using high order correlation method," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 29, pp. 841–856, July 1993.
- [25] R. J. Liou and M. R. Azimi-Sadjadi, "Multiple target detection using modified high order correlation," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 34, pp. 553–568, Apr. 1998.
- [26] W. S. Burdick, *Underwater Acoustic System Analysis*. Englewood Cliffs, NJ: Prentice-Hall, 1984.
- [27] S. Haykin, *Adaptive Filter Theory*, 3rd ed. Englewood Cliffs, NJ: Prentice-Hall, 1996.
- [28] S. U. Pillai, *Array Signal Processing*. Berlin, Germany: Springer-Verlag, 1989.
- [29] M. Vetterli and J. Kovacevic, *Wavelets and Subband Coding*. Englewood Cliffs, NJ: Prentice-Hall, 1995.



**Chunhua Yuan** received the B.S. degree from Tsinghua University, Beijing, China, in 1992 and the M.S. degree from Colorado State University, Fort Collins, in 1997, both in electrical engineering.

She is currently working as a Systems Engineer at Conexant Systems, Inc.



**Mahmood R. Azimi-Sadjadi** (S'81–M'81–SM'89) received the M.S. and Ph.D. degrees from the Imperial College of Science and Technology, University of London, U.K., in 1978 and 1982, respectively, both in electrical engineering with specialization in digital signal/image processing.

He is currently a Full Professor at the Electrical and Computer Engineering Department, Colorado State University (CSU), Fort Collins. He is also serving as the director of the Multi-sensory Computing Laboratory at CSU. His main areas of interest include digital

signal and image processing, target detection, classification and tracking, adaptive filtering and system identification, and neural networks. His research efforts in these areas resulted in over 100 journal and refereed conference publications. He is the co-author of the book *Digital Filtering in One and Two Dimensions* (New York: Plenum, 1989).

Dr. Azimi-Sadjadi is the recipient of the 1999 ABELL Teaching Award, the 1993 ASEE-Navy Senior Faculty Fellowship Award, the 1991 CSU Dean's Council Award, and the 1984 Dow Chemical Outstanding Young Faculty Award. He served as an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING. Presently, he is an Associate Editor of the IEEE TRANSACTIONS ON NEURAL NETWORKS.

**JoEllen Wilbur** (S'79–M'82) received the B.S. degree in electrical engineering from Virginia Polytechnic Institute, Blacksburg, in 1982 and the M.S. and Ph.D. degrees in electrical engineering from the University of Florida, Gainesville, in 1984 and 1987, respectively.

She was an Assistant Professor in the Department of Electrical and Computer Engineering at Clemson University, Clemson, SC, from 1987 to 1989. She has taught graduate and undergraduate courses in communications, digital signal processing, linear systems, and stochastic processes. She has been employed as part of the Signal and Image Processing Branch at the Coastal Systems Station, Panama City, FL, since 1989. Her research has involved underwater acoustic signal design/synthesis, LPI design/interception, sonar array processing, image enhancement, detection and classification, and broad-band sonar processing. Her current research interests include statistical signal processing, digital signal processing, and their applications. She has numerous publications in the form of journal papers, conference proceedings, and a book chapter. She holds one U.S. patent and has several patents pending.

Dr. Wilbur is a member of Eta Kappa Nu. She was a member of the IEEE Statistical Signal and Array Processing Technical Committee from 1992 to 1995. She is the recipient of the 1991 and 1996 Paper Awards under the Navy In-House Independent Research Program. She held a CAD/CAM Institutional Graduate Fellowship in 1984–1985 and an Army Research Office Graduate Fellowship in 1986–1987.



**Gerald J. Dobeck** received the B.S. degree in physics from the University of Massachusetts, Amherst, in 1970 and the M.S. and Ph.D. degrees in electrical engineering from the University of South Florida, Tampa, in 1973 and 1976, respectively.

Since 1976, he has been employed at the Coastal Systems Station, Naval Surface Warfare Center, Dahlgren Division, Panama City, FL. His current research interests include automatic detection and classification of targets in cluttered environments from synthetic/real aperture sonar imagery, the echo structure of acoustic returns, underwater electrooptic imagery, and gradiometer/magnetometer signals. He is project leader of the Sensor Signal and Image Processing Project under the Office of Naval Research 6.2 Mine Countermeasures program. In this, he is Technical Leader on the development of automated detection and classification algorithms of naval sea mines from sonar imagery, broad-band acoustic backscatter data, and magnetic gradiometer data. He has authored or co-authored more than 60 technical reports and papers.

Dr. Dobeck received the 1981 and 1996 Commanding Officer/Executive Director award for Science and Technology. He is a reviewer for IEEE, ASME, and SPIE journals and the *Journal of Underwater Acoustics* and has been Session Chair at past IEEE and SPIE conferences.