

# Automatic Sea-Surface Obstacle Detection and Tracking in Forward-Looking Sonar Image Sequences

Imen Karoui, Isabelle Quidu, and Michel Legris

**Abstract**—Automatic sea-surface object detection and tracking for safe autonomous underwater vehicle and submarine surfacing is a critical issue in relation to the accidents reported in the last decades. Here, we propose an efficient tool to detect and track sea-surface obstacles by processing forward-looking sonar images. The proposed method can detect either still or moving objects with and without wake. For each image sequence, a sequential procedure is proposed to detect various obstacle signatures. Then, target positions and velocities are estimated in Cartesian coordinates using the debiased converted measurement Kalman filter and the joint probabilistic data association filter. Detection and tracking stages exchange information in order to reduce the number of false alarms. Promising results are obtained using real data collected at sea with various objects and scenarios.

**Index Terms**—Multitarget tracking, obstacle avoidance, sea-surface obstacle detection.

## I. INTRODUCTION

OBSTACLE avoidance sonar systems are usually used in a forward-looking mode to detect and avoid obstacles either floating in the sea column or lying on the sea bottom [1], [2]. Automatic detection and tracking of sea-surface obstacles from forward sonar images is undoubtedly an interesting application that may help enable safe submarine and autonomous underwater vehicle (AUV) surfacing. In fact, in the last decades, many accidents have occurred. A famous accident was the collision between the Japanese vessel *Ehime Maru* and the submarine USS *Greenville* in 2001. Within minutes of the collision, *Ehime Maru* sank. Nine of her crewmembers were killed, including four high school students.

To our knowledge, no former work has been published on safe AUV surfacing, at least in open literature. However, some surface object detection and classification methods exist for harbor surveillance. However, most of these methods are based on passive acoustic detection [3], [4]. Active sonar systems dedicated to the ocean surface characterization mainly focus

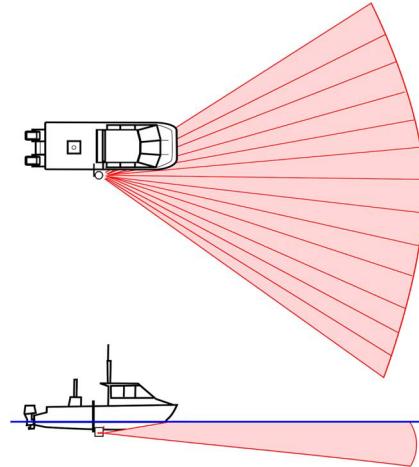


Fig. 1. Sonar configuration.

on experimental and theoretical analysis of ship wake structure through the study of bubble dynamics and backscattering [5]–[7], and only a few recent methods have been proposed for automatic ship detection using active sonar images [8]–[10]. These methods are mainly suited to a static configuration for harbor surveillance purposes, and they concern single ship detection.

Here, we propose a pattern analysis procedure for automatic detection and tracking of multiple surface objects from forward-looking sonar images. The proposed method can detect still and moving vessels with or without self-noise and wake.

This paper is organized as follows. In Section II, the experiment setup is described. In Section III, the different target signatures and the associated detection methods are detailed. In Section IV, the association and tracking procedure are depicted. In Section V, some results are given using real data gathered at sea. In Section VI, the conclusion and future work are outlined.

## II. EXPERIMENTAL SETUP

Obstacle detection is usually accomplished using forward-looking sonar. Such a system is an active sonar that emits a short acoustic pulse forward on a large horizontal sector (of about 120°) and a medium vertical sector (from 15° to 20°) (see Fig. 1). Acoustic echoes are recorded on a horizontal linear array of transducers. A beamforming process [11] (consisting in phase shifts and amplitude summation) is applied on the recorded data to sort the acoustic energy in time (related to the

Manuscript received March 12, 2014; revised July 16, 2014 and November 21, 2014; accepted January 23, 2015. This work was supported by Direction Générale de l'Armement through the “Rapid” Sponsorships.

The authors are with Ecole Nationale Supérieure de Techniques Avancées Bretagne, Lab-STICC UMR CNRS 6285, 29806 Brest, France, and also with SPARTE iXBlue & ENSTA Bretagne Joint Laboratory, 29806 Brest, France (e-mail: imen.karoui@ensta-bretagne.fr; isabelle.quidu@ensta-bretagne.fr; michel.legris@ensta-bretagne.fr).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TGRS.2015.2405672



Fig. 2. (a) ENSTA survey boat and (b) inflatable boat involved in the experiments.

distance of the echoes) and angle of arrival (named acoustic beams). The envelope of the resulting signals is then displayed as an image (see Fig. 3). That image could be represented in polar coordinates (range and bearing) or Cartesian coordinates (the acoustic image then appears in a fan shape).

Here, we use the RESON 8101 sonar with the image sector mode. Experiments were conducted by ENSTA Bretagne in Brest, Marina Du Château, in June 2013. The RESON 8101 forward-looking sonar was mounted on the ENSTA Bretagne survey boat at roughly 2 m below the sea surface. The sonar looks forward but intercepts the sea surface (see Fig. 1).

This sonar works at a central frequency of 240 kHz and a bandwidth of approximately 15 kHz. The horizontal transmitting aperture is wide and covers 150° horizontally. The vertical aperture is about 15°. However, the recorded data only covers a 60°-wide sector in the broadside direction due to electronic bandwidth constraints. The beamwidth for each of the received beams is 1.5°, and the range resolution is about 0.05 m.

Our survey boat was moving slowly in the zone under investigation, and data were recorded according to various scenarios with different sea-surface objects: static man-made objects laid on the sea surface, and moving ferries and sailboats arriving at or leaving the military harbor nearby. Some scenarios include a cooperative inflatable boat (see Fig. 2).

### III. SEA-SURFACE TARGET DETECTION

In this paper, the sea-surface obstacles considered are man-made objects: buoys, boats (motorboats or sailboats), containers, etc. According to the object type and state (still or moving), the surface acoustic signature can be as follows:

- strong-intensity beam(s) in the sonar image, (radial strips in Cartesian representation of sonar data) due to the stationary ship self-noise. This kind of signature indicates the ship bearing (it is the strongest among all surface object signatures and the easiest to detect, see Fig. 3);
- a high-contrast-intensity feature in the case of noise-free objects such as buoys, sailboats, and static ships such as fishing vessels (see Fig. 4);
- some high-intensity lines due to the wake in the track of moving vehicles (see Fig. 4). The wake echo level strongly depends on parameters such as the sensor distance to the ship, the aspect angle, and the center frequency of the transmitted pulse [7].

A detection and tracking procedure is proposed whatever the signature. The proposed approach is depicted in the flowchart of Fig. 5. It is a hierarchical detection procedure in order to deal

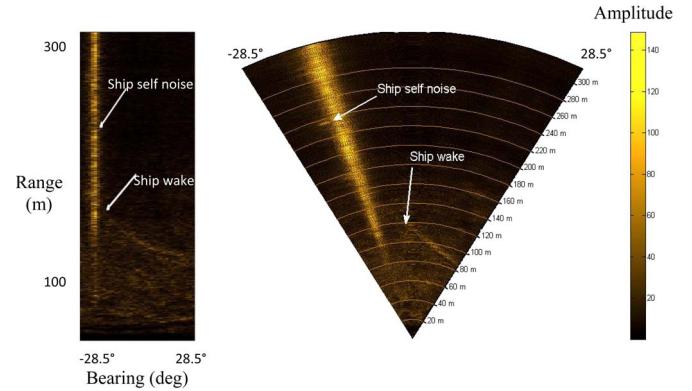


Fig. 3. Stationary-ship self-noise signature. (Left) Polar coordinates ( $x$ -axis corresponds to bearing and  $y$ -axis to range values). (Right) Cartesian coordinates.

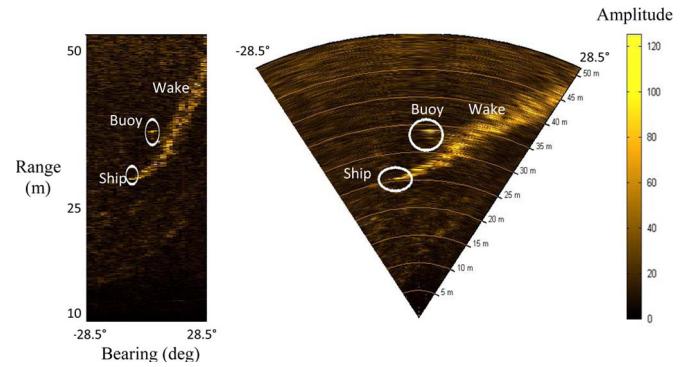


Fig. 4. Stationary-ship self-noise signature. (Left) Polar coordinates ( $x$ -axis corresponds to bearing and  $y$ -axis to range values). (Right) Cartesian coordinates.

with the various target signatures: stationary-ship self-noise, noise-free obstacles, and finally, wakes if they exist. In addition, as will be explained later on, object detection and tracking are carried out together: Information about the scene extracted from the last recurrences is taken into account to improve the detection rate for the image under process.

The proposed method is organized as follows: As the stationary-ship self-noise is strong and is spread on one or several adjacent beams, it could mislead the detection of other features. Therefore, we first try to detect this strongest signature. In the case of detection, the strong intensity beams corresponding to the ship bearing are ignored during the second step that consists in detecting the other types of obstacles. The third and final detection step consists in analyzing the area around every detected position in order to state if it corresponds to a reliable obstacle or to wakes only.

#### A. Detection of Ships With Self-Noise

The detection of the stationary-ship noise is often based on the evaluation of the mean echo level along each image beam. The beam of the highest energy is selected to give ship bearing [8]. As we assume neither the existence nor the uniqueness of vessel, we search for the more energetic beam(s) with respect to their neighbors. Instead of using the average of the beam amplitude as an indicator of ship self-noise presence,

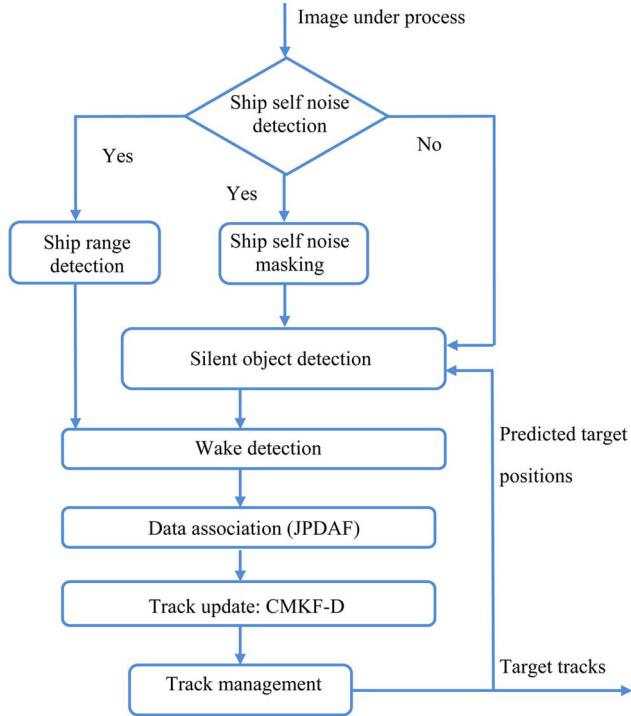


Fig. 5. Flowchart of the overall method.

we use the median amplitude value of each image beam that is more suitable for the multiple-target case. Indeed, the average of the high amplitude of a given beam may correspond to several targets without stationary noise and not to a single ship self-noise.

Formally, the indexes of the bearing measurements associated to ship self-noise at each image sequence  $\mathbf{I}_k$  are given by

$$\{\tilde{n}_k\} = \left\{ j \mid \mathbf{d}(j) > \frac{\xi}{2(n-n_g)} \left[ \sum_{k=-n}^n \mathbf{d}(j-k) - \sum_{k=-n_g}^{n_g} \mathbf{d}(j-k) \right] \right\} \quad (1)$$

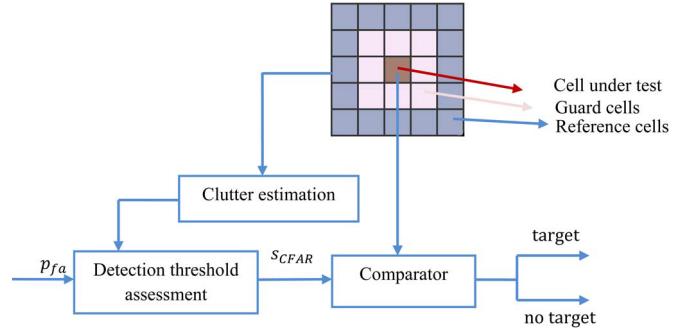
where  $\mathbf{d}(j)$  is the median of the image beam  $j$ .  $\xi$  is a preselected threshold higher than one, which is here set to 1.5.

$2(n-n_g)$  stands for the number of adjacent beams used to estimate the detection threshold. The  $2n_g$  closest neighbors are excluded from the threshold assessment because the signal is spread on several neighboring beams. This spreading is the result of two factors. First, no beamforming can achieve an ideal directivity pattern: a small part of the acoustic energy leaks on adjacent beams (sidelobe effects [11]). Second, the sampling interval is generally smaller than the beam resolution. For our sonar, we observed that a good value of  $n_g$  is

$$n_g = 3 \left\lfloor \frac{\Delta_\varphi}{\Delta\delta} \right\rfloor \quad (2)$$

where  $\lfloor \cdot \rfloor$  is the smaller integer inside the brackets,  $\Delta_\varphi$  stands for the beamwidth, and  $\Delta\delta$  for the angular sampling rate (for the RESON 8101,  $\Delta_\varphi = \Delta\delta = 1.5^\circ$ ).

Ship echoes are usually strong, and thus, ship range detection consists in finding the highest intensity pixels in the detected beams. Because of the time-varying gain or the automatic gain amplifier usually applied to sonar images, the ship noise

Fig. 6. CFAR detector:  $p_{fa}$  is the probability of false alarm, and  $s_{CFAR}$  is the detection threshold.

level increases with range, and the target could be masked particularly if it occurs at far ranges. To reduce this effect, as in [8], we normalize the detected image beams by a moving window median filter as follows:

$$\tilde{\mathbf{I}}_k(m, n) = \frac{\mathbf{I}_k(m, n)}{\text{med}(\mathbf{I}_k((m-n_b), n), \dots, \mathbf{I}_k((m+n_b), n))} \quad \forall n \in \{\hat{n}_k\} \quad (3)$$

where  $\mathbf{I}_k(m, n) \equiv \mathbf{I}_k(m\Delta r, n\Delta\delta)$  stands for the image amplitude at polar coordinates  $(m\Delta r, n\Delta\delta)$ ;  $\Delta r$  and  $\Delta\delta$  are, the range and bearing sampling rates, respectively;  $\text{med}(\cdot)$  stands for the median value of the quantity inside the brackets; and  $2n_b + 1$  is the length of the moving window median filter. For our experiments, we set  $n_b = 50$  (about 2.5 m).

We then pick the first ten strongest peaks of each normalized detected beam. If ship wakes are visible, their positions indicate ship ranges (see Fig. 3). In such cases, only the range value selected after the wake detection stage is kept (see Section III-C).

### B. Detection of Noise-Free Targets

To our knowledge, there is no work dealing with the detection of ships without self-noise for active sonar images. In SAR imagery, the topic of ship detection has been extensively studied [12]. The most powerful detector widely used in SAR imagery for ship detection is the constant false-alarm rate (CFAR) algorithm with an adaptive threshold [13]. This threshold is computed for each image cell by taking into account the clutter power estimated from a set of reference cells surrounding the cell under investigation, as shown in Fig. 6.

Guard cells are the nearest neighbors of the cell under investigation and are not used in the local clutter estimation in order to ensure that pixels belonging to an extended target cannot be included in the clutter statistics estimation.

Adaptive radar detection for several clutter probability density functions (pdfs) (Gaussian, Weibull,  $K$  distribution, Pearson, etc.) have been widely studied in the literature [14]–[16].

Here, we deal with amplitude sonar images, and the distribution of pixel levels is usually modeled by a Rayleigh distribution [17]:  $f(x) = (2x/b^2) \exp(-(x^2/b^2))$ . We assume that clutter samples are independent and identically distributed, and we estimate the parameter  $b$  according to the maximum-likelihood method [15]:  $\hat{b} = ((1/J) \sum_{j=1}^J x_j^2)^{1/2}$ ,

where  $J$  is the number of reference cells, and  $x_j$  the  $j$ th cell amplitude. The CFAR detector is then based on the following threshold [14]:

$$s_{\text{CFAR}} = \left[ \left( p_{\text{fa}}^{-\frac{1}{J}} - 1 \right) \sum_{j=1}^J x_j^2 \right]^{\frac{1}{2}}. \quad (4)$$

For our data, two events can increase the false-alarm rate or induce target missing: the presence of wake and other targets in the immediate vicinity of pixels under test. Indeed, every interfering target or wake that appears in the reference window associated to the target under investigation may raise the CFAR threshold extemporaneously. To avoid a wrong estimation of the detection threshold, we allow information exchanges between the tracker and the different detectors. Practically, at the  $k$ th image of the sequence, we exclude every cell from the reference cells, which are used in the clutter parameter estimation ( $\hat{b}$ ), belonging to the following:

- detected beams corresponding to ship self-noise if they exist;
- previously detected wakes;
- target positions predicted by the Kalman filter.

### C. Wake Detection

Moving ships generate bubbly turbulent wake in their track. The turbulence can persist for several minutes and extend over hundreds of meters [9]. Commonly, wake width is five to ten times the ship's width, and it increases slowly with the distance in relation to the ship. The exact wake structure depends, however, on the ship propeller, hull design, ship speed, and maneuver [18].

When wake cell amplitudes are high enough to be detected by the CFAR detector, detections associated to the wake drastically increase the amount of false alarms and hence mislead ship localization and tracking.

Here, we do not need an exact wake detection and characterization. Our goal is only to detect the reliable position of the ship among all the detections generated by wakes. To do so, we first search for wake ends, and we assume that ship position is given by the wake endpoint with the highest echo level. However, for our experiments, as we work at quite short maximum ranges (50 m for a typical sequence for which the wake is visible), we notice that, for some cases, ship echo is the wake endpoint with the lowest energy. A pragmatic solution is to keep both wake ends if both of them are detected, and the spurious end could easily be removed after tracking.

The proposed wake end detection consists in analyzing the area around each detection within oriented strips as shown in Fig. 7. We assume that if the amount of detections for one orientation is much higher than the detection number in the other orientations, the detection is likely to be a wake end. In order to reduce the computation time, the search for wake ends is not carried out for each detection but only around cluster centers: closely spaced detections are gathered into clusters, and each cluster center is given by the averaged position of all detections belonging to the cluster.

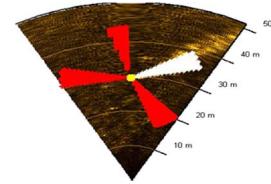


Fig. 7. Wake detection procedure for (yellow) a given detection. (Red) Three strips of 15-m length and about 20° width for three different orientations. (White) detected strip.

Formally, we jointly use two strip lengths  $l_1$  and  $l_2 = l_1/2$ . We consider that a given detection could correspond to a wake end if these two conditions are satisfied.

- 1) The number of detections in one of its associated strip of length  $l_1$  for a given orientation is higher than five times the mean detection number of all different strips of length  $l_1$ , and this number exceeds  $n_w$  detections.
- 2) The number of detections in its associated strip of length  $l_2$  in the same orientation is higher than five times the mean detection number of all different strips of length  $l_2$ , and this number exceeds  $2n_w/3$  detections.

Condition 2 avoids the detection of isolated detections near the wake by taking into account a high number of detections near the ship. The minimum detection number  $n_w$  ensures that the wake is dense. We set  $n_w = 50$ .

For our data, we noticed that turbulent wakes have a maximum width of 20°; hence, we use strips of 20° in width, and we set  $l_1 = 15$  m.

Many detections belonging to the same wake can satisfy these two conditions. The clustering of wake detections inside the same wake is carried out as follows: All pixels belonging to overlapped strips and having similar detected orientations (modulo  $2\pi$  because we keep both wake ends) are clustered, and only the end position is kept.

All the other detections inside the detected wakes are removed from the measurement list.

## IV. OBJECT ASSOCIATION AND TRACKING

### A. Target Tracking

Once the detection is carried out and potential target coordinates are estimated, we aim at finding reliable targets and tracking them. Target tracking is performed in the Cartesian frame according to the classic near constant velocity dynamical model coded in matrices  $\mathbf{F}$  and  $\mathbf{G}$  of the following state equation:

$$\mathbf{x}_k^t = \mathbf{F}\mathbf{x}_{k-1}^t + \mathbf{G}\mathbf{v}_{k-1} \quad (5)$$

where

—  $\mathbf{x}_k^t = [p_x, p_y, \dot{p}_x, \dot{p}_y]$ ,  $[p_x, p_y]$ , and  $[\dot{p}_x, \dot{p}_y]$  are the position and the velocity vectors in the  $xy$  plane relative to the sonar, respectively;

$$-\mathbf{F} = \begin{bmatrix} 1 & 0 & \Delta t & 0 \\ 0 & 1 & 0 & \Delta t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{G} = \begin{bmatrix} (\Delta t^2/2) & 0 \\ 0 & (\Delta t^2/2) \\ \Delta t & 0 \\ 0 & \Delta t \end{bmatrix}$$

and  $\Delta t$  refers to the time sampling;

—  $\mathbf{v}_{k-1}$  is a 2-D zero-mean independent and identically distributed white Gaussian noise with standard deviation  $\sigma_v$  for each component.

At ping  $k$ , measurements that originated from target  $t$  are related to the state vector  $\mathbf{x}_k^t$  according to the following measurement equation:

$$\mathbf{y}_k = \mathbf{H}\mathbf{x}_k^t + \mathbf{w}_k \quad (6)$$

where  $\mathbf{H}$  is a known matrix, and  $\mathbf{w}_k$  is a known measurement noise.

Target positions are reported in polar coordinates (range and bearing) with respect to the sonar position. In order to deal with a linear measurement equation (6), we convert polar coordinates to Cartesian coordinates, and the bias due to the nonlinearity of measurement conversion is tackled by using the converted measurement Kalman Filter with debiased conversion (CMKF-D) [19].

Let us consider  $\{(d_m, \delta_m)\}$  the polar measurements corrupted by a white and Gaussian noise with covariance matrix

$$\mathbf{R}_{\text{polar}} = \begin{bmatrix} \sigma_d^2 & 0 \\ 0 & \sigma_\delta^2 \end{bmatrix}$$

where  $\sigma_d$  and  $\sigma_\delta$  are the error standard deviations in range and bearing, respectively.

The conversion of the polar measurement to debiased Cartesian measurement is as follows:

$$\mathbf{y}_k^c = \begin{bmatrix} p_x^c \\ p_y^c \end{bmatrix} = \begin{bmatrix} d_m \cos(\delta_m) \\ d_m \sin(\delta_m) \end{bmatrix} - \boldsymbol{\mu}_c \quad (7)$$

where  $\boldsymbol{\mu}_c$  is the bias compensation term given by (see [19] for details)

$$\boldsymbol{\mu}_c = \begin{bmatrix} d_m \cos(\delta_m) \left( \exp^{-\sigma_\delta^2} - \exp^{-\sigma_\delta^2/2} \right) \\ d_m \sin(\delta_m) \left( \exp^{-\sigma_\delta^2} - \exp^{-\sigma_\delta^2/2} \right) \end{bmatrix}. \quad (8)$$

The converted measurement coordinates are not mutually independent, and their error covariance matrix  $\mathbf{R}_c = \begin{bmatrix} r_c^{11} & r_c^{12} \\ r_c^{21} & r_c^{22} \end{bmatrix}$  depends on the target bearing and range and on their error standard deviations [19], i.e.,

$$\begin{aligned} r_c^{11} = & d_m^2 \exp^{-2\sigma_\delta^2} [\cos^2(\delta_m) (\cosh(2\sigma_\delta^2) - \cosh(\sigma_\delta^2)) \\ & + \sin^2(\delta_m) (\sinh(2\sigma_\delta^2) - \sinh(\sigma_\delta^2))] \\ & + \sigma_d^2 \exp^{-2\sigma_\delta^2} [\cos^2(\delta_m) (2 \cosh(2\sigma_\delta^2) - \cosh(\sigma_\delta^2)) \\ & + \sin^2(\delta_m) (2 \sinh(2\sigma_\delta^2) - \sinh(\sigma_\delta^2))] \end{aligned} \quad (9)$$

$$\begin{aligned} r_c^{22} = & d_m^2 \exp^{-2\sigma_\delta^2} [\sin^2(\delta_m) (\cosh(2\sigma_\delta^2) - \cosh(\sigma_\delta^2)) \\ & + \cos^2(\delta_m) (\sinh(2\sigma_\delta^2) - \sinh(\sigma_\delta^2))] \\ & + \sigma_d^2 \exp^{-2\sigma_\delta^2} [\sin^2(\delta_m) (2 \cosh(2\sigma_\delta^2) - \cosh(\sigma_\delta^2)) \\ & + \cos^2(\delta_m) (2 \sinh(2\sigma_\delta^2) - \sinh(\sigma_\delta^2))] \end{aligned} \quad (10)$$

$$\begin{aligned} r_c^{12} = r_c^{21} = & \sin(\delta_m) \cos(\delta_m) \exp^{-4\sigma_\delta^2} \\ & \times [\sigma_d^2 + (d_m^2 + \sigma_d^2) (1 - \exp^{-\sigma_\delta^2})] \end{aligned} \quad (11)$$

We notice that under some conditions ( $d\sigma_\delta^2/\sigma_d < 0.4$  and  $\sigma_\delta < 23^\circ$ ) a simpler covariance form, obtained through the linearization of the Cartesian coordinate errors, could be used [19].

The CMKF-D consists in iterating the two following standard Kalman filter steps for each target  $t$ .

Prediction step:

— The predicted state

$$\hat{\mathbf{x}}_{k|k-1}^t = \mathbf{F}\hat{\mathbf{x}}_{k-1|k-1}^t. \quad (12)$$

— The predicted state error covariance

$$\mathbf{P}_{k|k-1}^t = \mathbf{F}\mathbf{P}_{k-1|k-1}^t\mathbf{F}^T + \mathbf{Q}. \quad (13)$$

— The predicted measurement

$$\hat{\mathbf{y}}_{k|k-1}^t = \mathbf{H}\hat{\mathbf{x}}_{k|k-1}^t. \quad (14)$$

— The predicted measurement covariance matrix

$$\mathbf{S}_k^t = \mathbf{H}\mathbf{P}_{k|k-1}^t\mathbf{H}^T + \mathbf{R}_c^t \quad (15)$$

where  $\mathbf{Q} = \sigma_v^2 \mathbf{G}\mathbf{G}^T$  is the process noise covariance,  $T$  stands for the matrix transpose, and  $\mathbf{R}_c^t$  is the debiased converted measurement error covariance for target  $t$ . The components of  $\mathbf{R}_c^t$  are computed according to (9), (10), and (11).

Update step: This step is carried out for each existing target track using its associated measurement(s) and is done according to the following standard manner:

$$\hat{\mathbf{x}}_{k|k-1}^t = \hat{\mathbf{x}}_{k|k-1}^t + \mathbf{K}_k^t \vartheta_k^t \quad (16)$$

$$\mathbf{P}_{k|k}^t = \mathbf{P}_{k|k-1}^t + \mathbf{K}_k^t \mathbf{S}_k^t (\mathbf{K}_k^t)^T \quad (17)$$

where  $\vartheta_k^t$  is the innovation vector,  $\mathbf{S}_k^t$  is the associated covariance matrix, and  $\mathbf{K}_k^t$  is the conventional Kalman filter gain

$$\mathbf{K}_k^t = \mathbf{P}_{k|k-1}^t \mathbf{H}^T (\mathbf{S}_k^t)^{-1}. \quad (18)$$

## B. Data Association

A specific procedure is needed to assign each target to its associated measurement(s). This procedure is designated as data association. Many data association algorithms have been developed [20]–[22]. Among them, the global nearest neighbor, the joint probabilistic data association (JPDA), and multiple hypothesis tracking algorithms are considered the most common and the most efficient techniques. Here, we chose to use the JPDA filter because of its relatively low computational complexity, its suitability to multitarget tracking, and its robustness to noisy environments. In JPDA-filter-based association, all measurements lying in the target validation gate are combined to assess the innovation used to update its state (16).

Formally, a validation gate for a target  $t$  is an ellipsoid defined as  $\{\mathbf{y}_i | (\mathbf{y}_i - \hat{\mathbf{y}}_{k|k-1}^t)(\mathbf{S}_k^t)^{-1}(\mathbf{y}_i - \hat{\mathbf{y}}_{k|k-1}^t)^T \leq \gamma\}$ .  $\gamma$  ensures that the correct measurements will lie within the gate with a specified probability, and it is determined from the Chi-square distribution [20]. This region is computed at every sampling time and is used to select believable correct measurements.

The combined (weighted) innovation vector is given by

$$\boldsymbol{\vartheta}_k^t = \sum_{j=1}^{m_k} \beta_{jt}^k (\mathbf{y}_k^j - \hat{\mathbf{y}}_{k|k-1}^t) \quad (19)$$

where

- $\mathbf{y}_k^j$  is the measurement  $j$ ,  $j = 1, \dots, m_k$ .  $m_k$  is the measurement number at ping  $k$ ;
- $\beta_{jt}^k$  is the posterior probability that the measurement  $\mathbf{y}_k^j$  truly originated from target  $t$ ,  $t = 0, \dots, t_k$ .  $t = 0$  stands for false alarm, and  $t_k$  is the target number at ping  $k$ .

The probability  $\beta_{jt}^k$  of the event “measurement  $j$  originated from target  $t$ ” is obtained by evaluating the probabilities of the joint events conditioned on all measurements assuming that a measurement cannot be generated by more than one target. Formally, if we denote these joint events  $\theta_k = \cap_{j=1}^{m_k} \theta_{jt_j}$ , where  $\theta_{jt_j}$  is the event that the measurement  $\mathbf{y}_k^j$  is originated from target  $t_j$ , we have

$$\beta_{jt}^k \triangleq p(\theta_{jt}| \mathbf{y}_{1,\dots,k}) = \sum_{\theta_k | \theta_{jt} \in \theta_k} p(\theta_k | \mathbf{y}_{1,\dots,k}) \quad (20)$$

$$\beta_{0t}^k \triangleq 1 - \sum_{j=1}^{m_k} \beta_{jt}^k \quad (21)$$

where

- $\mathbf{y}_{1,\dots,k} = \{\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3, \dots, \mathbf{y}_k\}$  is the measurement set from time one up to time  $k$  and  $\mathbf{y}_k = \{\mathbf{y}_k^j\}_{j=1,\dots,m_k}$  the measurement set at ping  $k$ .
- $\beta_{0t}^k$  is the probability that none of the measurements is originated from target  $t$ .

The evaluation of the conditional joint event probabilities  $p(\theta_k | \mathbf{y}_{1,\dots,k})$  is detailed in [20]. For a parametric JPDA with Poisson distributed false alarms, we obtain the following:

$$p(\theta_k | \mathbf{y}_{1,\dots,k}) = \frac{1}{c} \prod_{j=1}^{m_k} \left\{ \lambda^{-1} N \left( \hat{\mathbf{y}}_{k|k-1}^{t_j}, \mathbf{S}_k^{t_j} \right) \left( \mathbf{y}_k^j \right) \right\}^{\tau_j(\theta_k)} \times \prod_{t=1}^{t_k} (p_d^t)^{\delta_t(\theta_k)} (p_d^t)^{1-\delta_t(\theta_k)} \quad (22)$$

where

- $c$  is a normalization term;
- $N$  is the normal distribution;
- $\lambda$  is the expected density of false measurement;
- $\tau_j(\theta_k)$  is a binary measurement association indicator equal to one if the measurement is associated to a target (not considered as a false alarm) in the joint event  $\theta_k$  and zero otherwise;
- $\emptyset(\theta_k) = \sum_{j=1}^{m_k} [1 - \tau_j(\theta_k)]$  is the number of false measurements in event  $\theta_k$ ;
- $p_d^t$  stands for the detection probability for target  $t$ ;
- $\delta_t$  is the target detection indicator equal to zero if no measurement is assigned to target  $t$  and to one, otherwise.

The state error covariance matrix update (17) expressed according to  $\{\beta_{jt}^k\}$  is

$$\mathbf{P}_{k|k}^t = \beta_{0t}^k \mathbf{P}_{k|k-1}^t + (1 - \beta_{0t}^k) \mathbf{P}_{k|k}^{t,c} + \tilde{p}_k \quad (23)$$

where

$$\begin{aligned} \mathbf{P}_{k|k}^{t,c} &= \mathbf{P}_{k|k-1}^t - \mathbf{K}_k^t \mathbf{S}_k^t \mathbf{K}_k^{t,T} \\ \tilde{p}_k &= \mathbf{K}_k^t \left( \sum_{j=1}^{m_k} \beta_{jt}^k \mathbf{y}_k^{t,j} (\boldsymbol{\vartheta}_k^{t,j})^T - \boldsymbol{\vartheta}_k^t (\boldsymbol{\vartheta}_k^t)^T \right) (\mathbf{K}_k^t)^T. \end{aligned}$$

## B. Track Management

JPDAF does not allow for automatic track initiation and termination. In our approach, tracks are managed according to the following rules.

- If there are  $n_1$  detections falling outside all already existing target gates in the last  $m_1$  recurrences below a given distance from the most recent detection (difference between  $x$ -coordinates below  $x_{\max}$  and between  $y$ -coordinates below  $y_{\max}$ ), a new track is initiated at that position.
- If an existing track is not updated at least  $n_2$  times during the  $m_2$  later recurrences, it is interrupted.
- In order to decrease the processing time and to avoid likely false track maintenance, we remove tracks with a position error standard deviation exceeding a given value (20 m for our data).

## V. RESULTS

The proposed approach is applied to several sequences with still and moving objects for mainly two maximum range values: 200 and 50 m. In this section, we show detections and tracking results for four recorded sequences.

- Sequence I: 600 images of 4000 pixels in range and 39 beams (equivalent to 5 min) recorded as our survey boat was slowly approaching a large buoy (of nearly 6 m diameter), two smaller buoys behind it (of nearly 3 m diameter), and concrete blocks near the pier (see Fig. 8);
- Sequence II: 280 images of 1000 pixels in range and 39 beams (equivalent to about 2.5 min) where the inflatable boat was running on its own momentum crossing behind an emerged buoy (see Fig. 9);
- Sequence III: 480 images of 1000 pixels in range and 39 beams recorded while the inflatable boat was running on its own momentum and crossing very close to a sailboat (see Fig. 10);
- Sequence IV: very short sequence of 20 images of 6000 pixels in range and 39 beams, which corresponds to the crossing of a passenger boat with a self-noise visible in the acoustic images (see Fig. 3). Given our sensor frequency (240 kHz), the boat noise specifically corresponds to the cavitation noise.

Standard deviation of the polar measurement error  $\sigma_d$  and  $\sigma_\delta$  are set according to the sonar range and bearing sampling rates:  $\sigma_d = \Delta_d = 0.05$  m,  $\sigma_\delta = \Delta_\delta = 1.5^\circ$ , and sampling time  $\Delta t = 0.5$  s. The noise process standard deviation is set low but nonzero:  $\sigma_v = 10^{-3}$  m/s<sup>2</sup>. Target detection probability and clutter density parameters are set to  $p_d^t = 0.9 \forall t$  and  $\lambda = p_{fa}$ . Tracks are initiated if a detection is observed at least 16 times during the last 20 pings within a region of 5 m by 5 m for sequences of maximum range of 50 and 10 m by 10 m for



Fig. 8. Sequence I, known surface objects: (from left to right) a photograph of the surface objects, sonar data at ping 550 in polar coordinates, and sonar data at ping 550 in Cartesian coordinates.



Fig. 9. Sequence II, known surface objects: (from left to right) a photograph of the surface objects, sonar data at ping 250 in polar coordinates, and sonar data at ping 250 in Cartesian coordinates.

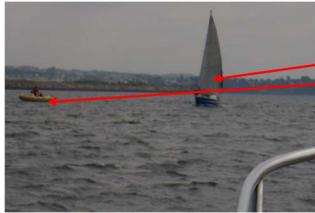


Fig. 10. Sequence III: the known surface objects: (from left to right) a photograph of the surface objects, sonar data at ping 450 in polar coordinates, and sonar data at ping 450 in Cartesian coordinates.

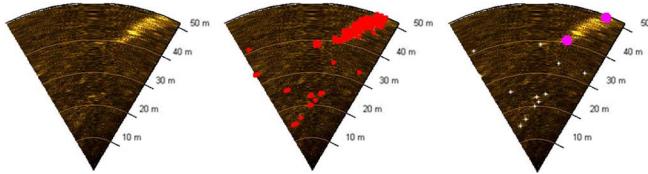


Fig. 11. Sequence II: Detection results at ping 47: (from left to right) raw data in Cartesian coordinates, all detected positions ( $p_{fa} = 10^{-6}$ ), remaining measurements after clustering, and wake detection removal (in white are the cluster center positions and in magenta the detected wake ends).

sequences of maximum range of 200 m. A track is terminated if no measurements are associated to it at least eight times during the last 20 pings ( $m_1 = m_2 = 20$ ,  $n_1 = 2n_2 = 16$ ),  $x_{\max} = y_{\max} = 5$  m for sequences of maximum range of 50 m, and  $x_{\max} = y_{\max} = 10$  m for sequences with maximum range of 200 m.

Initial state estimate error covariance  $\mathbf{P}_{0|0}^t$  for a track associated to a target  $t$  is set as follows:

$$\mathbf{P}_{0|0}^t = 5 \begin{bmatrix} \mathbf{R}_0^t & 0 & 0 \\ 0 & (x_{\max}/m_1 \Delta t)^2 & 0 \\ 0 & 0 & (y_{\max}/m_1 \Delta t)^2 \end{bmatrix}$$

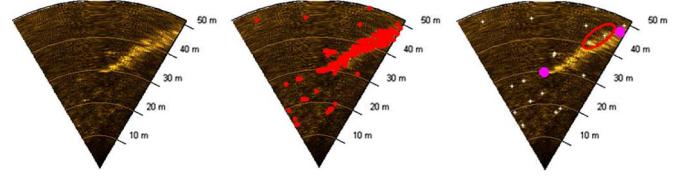


Fig. 12. Sequence II: Detection results at ping 161: (from left to right) data in Cartesian coordinates, all detections ( $p_{fa} = 10^{-6}$ ), remaining measurements after wake detection and detection clustering (in white are the cluster centers, and in magenta are the detected wake ends; the red ellipse shows two detections inside the wake that are not removed by the proposed method).

where  $\mathbf{R}_0^t$  is the measurement error covariance matrix at the initial target position computed according to (9), (10), and (11).

The proposed method was tested for various  $p_{fa}$  values and various reference cell numbers  $J$ . Detection capabilities were assessed visually by the analysis of the detection results and according to our knowledge of the sea-surface environment during the sequence acquisition. To allow this analysis, we estimated the theoretic average false-alarm number as the product of the theoretical  $p_{fa}$  by the number of image pixels. The number of false alarms is computed as the difference between the total detection number and ‘true’ detections that are visually identified. We noticed that, in the absence of wakes, we globally obtain false-alarm numbers consistent with the theoretical  $p_{fa}$  for a number of reference cells below 50.

In the presence of wake, the number of false alarms can be very high compared to the theoretical false-alarm number. For example at ping 47 of Sequence II, even for low  $p_{fa}$  values ( $p_{fa} = 10^{-6}$ ), the number of false alarms is high (about 300 false alarms). Most of these false alarms are related to the inflatable boat wake (see Fig. 11). The amount of detections drastically increases as the ship wake becomes longer (see Fig. 12). For such cases, removal of wake cells is essential. Figs. 11 and 12 show the detection results, respectively, for ping 47 and ping 161 of Sequence II with  $p_{fa}$  equal to  $10^{-6}$ . In the middle, we show all detections using the CFAR detector (4). On the right, we draw the measurements used for tracking: Most of detections belonging to the wake are removed, and the number of detections drops from 350 to 15 after clustering (detections within 2 m are replaced by their average position) and wake pixel removal.

Wake detection works well in most of cases and particularly in the case of a strong wake behind the ship. However, in some cases, as the inflatable boat changed its heading frequently, the wake signature could appear distorted, and the method did not succeed in removing all the detections originated from the wake (see Fig. 12). Despite these failures, prior wake detection has the advantage of reducing the overall method computation time while maintaining robust ship tracking (see Figs. 16 and 17). Cavitation noise is easily detected in all pings of Sequence IV, and the ship range is well deduced from the wake detection step (see Fig. 13).

Tracking results can only be assessed in a qualitative way because our targets are not equipped with Global Positioning System. We notice that the proposed procedure succeeds in tracking the known targets. Figs. 14, 16, and 17 show the tracks associated to the known targets in Sequences I, II and III, respectively, for three different pings. We superimpose a confidence ellipse on each target track that describes the Kalman

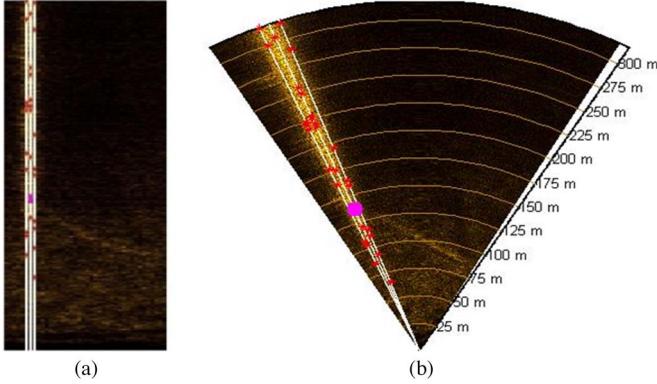


Fig. 13. Sequence IV: Detection results: (white lines) cavitational noise detection, (red) candidate range detections, and (magenta) the final range deduced from wake detection ( $p_{fa} = 10^{-6}$ ), in (a) polar coordinates and in (b) Cartesian coordinates.

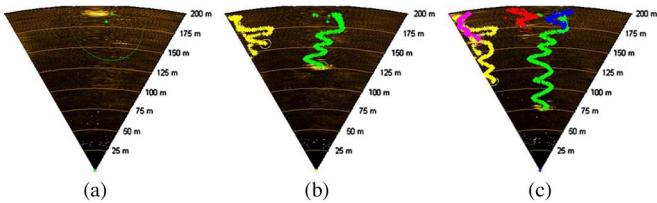


Fig. 14. Sequence I: The known object tracks and their corresponding confidence ellipse up to ping: (a)  $k = 27$ , (b)  $k = 350$ , and (c)  $k = 590$ . In green is the track associated to the large buoy, yellow and magenta tracks correspond to the two concrete blocks, and in red and in blue are the tracks associated to the smaller buoys ( $p_{fa} = 10^{-5}$ ,  $\gamma = 9$ ).

filter own assessment of tracking accuracy, expressed by the covariance matrix  $\mathbf{P}_{k|k}^t$ . This ellipse is only shown for the predicted target position and it is defined as follows:

$$\left\{ \mathbf{y} = [p_x, p_y] \mid (\mathbf{y} - \mathbf{H}\tilde{\mathbf{x}}_{k|k}^t)(\mathbf{L}\mathbf{P}_{k|k}^t\mathbf{L}^T)^{-1}(\mathbf{y} - \mathbf{H}\tilde{\mathbf{x}}_{k|k}^t)^T \leq \lambda \right\} \quad (24)$$

where  $\mathbf{L} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix}$ .

At the beginning of Sequence I, only the largest buoy is in the field of view of the sonar, and the concrete blocks and the two smaller buoys are seen later. The proposed method succeeds in automatically initiating every target track since they appear in the field of view of the sonar. However, the first concrete block was initiated by a ‘false’ track (the track in yellow in Figs. 14 and 15). Indeed, for this sequence, despite using low  $p_{fa}$  value ( $p_{fa} = 10^{-5}$ ), ‘false’ tracks were created. These false tracks were related to pier pillars present at the edge of the image but not ‘seen’ by the sonar at every ping because of the survey boat yaw (see Fig. 15).

For Sequence I, the targets are still. The shape of the tracks (see Fig. 14) was due to the survey boat yaw. In an operational system, the sonar image will be stabilized according to carrier movement, and uniform target trajectory will appear as straight lines; hence, we expect better performances.

For Sequences II and III, we show three tracks in Figs. 16 (Sequence II) and 17 (Sequence III): one associated to the fixed buoy and two tracks associated to both wake

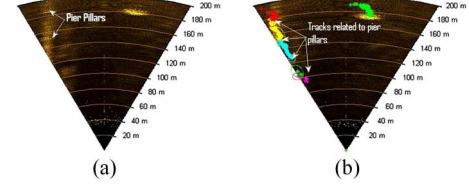


Fig. 15. Sequence I: (a) Image at ping:  $k = 77$ . (b) Existing tracks at ping  $k = 111$ : All tracks on the left are initiated by pier pillar associated detections.

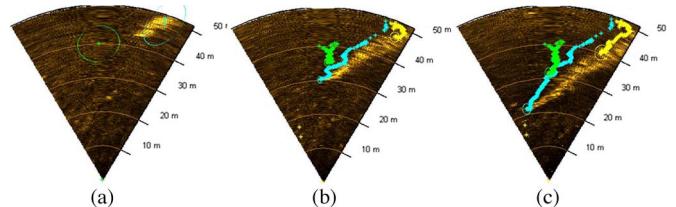


Fig. 16. Sequence II—The known object tracks and their confidence ellipses up to ping (a)  $k = 22$ , (b)  $k = 150$ , and (c)  $k = 275$ . In green is the track associated to the buoy, in cyan is the track associated to the inflatable boat, and in yellow is the track associated to the second wake end ( $p_{fa} = 10^{-6}$ ,  $\gamma = 4$ ).

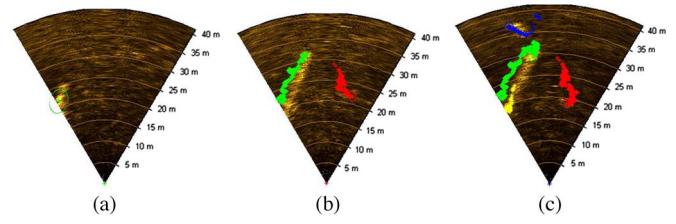


Fig. 17. Sequence III—The known object tracks and their confidence ellipses up to (a) ping  $k = 21$ , (b)  $k = 330$ , and (c)  $k = 460$ . In green is the track associated to the inflatable boat, in yellow is the track associated to the second wake endpoint, in red is the track associated to the fixed buoy, and in blue is the track related to a sailboat ( $p_{fa} = 10^{-6}$ ,  $\gamma = 4$ ).

endpoints. At the end of Sequence III, we can see the track related to the sailboat.

In Sequence II, when the zodiac approaches the buoy then goes past it, the algorithm succeeds in maintaining both tracks (see Fig. 16).

## VI. CONCLUSIONS AND DISCUSSIONS

Sea-surface detection and tracking by means of forward-looking sonar is a useful application for safe AUV and submarine surfacing. In this paper, we have proposed an automatic method for detecting and tracking various sea-surface targets: Still and moving objects with or without wake and cavitational noise. Promising results have been obtained with real data gathered at sea.

Future improvements will first concern the integration of additional features in the target state vector, e.g., the object size and the wake orientation. Indeed, in the proposed method, wake detection was mainly used to improve the CFAR detector performance and to reduce the false-alarm rate due to the removal from the measurement list of unwanted detection inside the wakes. In addition, in the operational system, the data are stabilized from the sensor yaw, pitch, and roll; hence, target trajectories will appear straighter than the tracks presented here, and the wake orientations will be consistent with target

trajectories. For such case, wake orientation could be used as an additional target feature, which may improve target tracking. However, adding wake orientation to the target state vector will lead to a highly nonlinear measurement system, and the CMKF-D will no longer be suitable; for example, an extended Kalman filter or a JPDAF particle filter should be used.

The proposed method has been carried out on images without any preprocessing, and image denoising may also improve the detection performances.

## REFERENCES

- [1] Y. Petillot, I. T. Ruiz, and D. M. Lane, "Underwater vehicle obstacle avoidance and path planning using a multi-beam forward looking sonar," *IEEE J. Ocean. Eng.*, vol. 26, no. 2, pp. 1194–1199, Apr. 2001.
- [2] I. Quidu, A. Hétet, Y. Dupas, and S. Lefèvre, "AUV (REDERMOR) obstacle detection and avoidance experimental evaluation," in *Proc. IEEE OCEANS Conf.*, Aberdeen, Scotland, Jun. 2007, pp. 18–21.
- [3] L. Fillinger, A. Sutin, and A. Sedunoy, "Acoustic ship signature measurements by cross-correlation method," *J. Acoust. Soc. Amer.*, vol. 129, no. 2, pp. 774–778, 2011.
- [4] G. L. Ogden, L. M. Zurk, M. E. Jones, and M. E. Peterson, "Extraction of small boat harmonic signatures from passive sonar," *J. Acoust. Soc. Amer.*, vol. 129, no. 2, pp. 3768–3776, 2011.
- [5] M. V. Trevoror, S. Vagle, and D. M. Farmer, "Acoustical measurements of microbubbles within ship wakes," *J. Acoust. Soc. Amer.*, vol. 95, no. 4, pp. 1922–1930, 1994.
- [6] S. Vagle and H. Burch, "Acoustic measurements of the sound-speed profile in the bubbly wake formed by a small motor boat," *J. Acoust. Soc. Amer.*, vol. 117, no. 1, pp. 153–163, 2005.
- [7] A. Sutin, A. Benilov, H. S. Roh, and Y. I. Nah, "Acoustic measurements of bubbles in the wake of ship model in tank," *J. Acoust. Soc. Amer.*, vol. 123, no. 5, pp. 8651–8656, 2008.
- [8] K. W. Lo and B. G. Ferguson, "Automatic detection and tracking of a small surface watercraft in shallow water using a high-frequency active sonar," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 40, no. 4, pp. 1377–1388, Oct. 2004.
- [9] Ç. Öner, "Acoustic tracking of ship wakes," M.S. thesis, Graduate School Natural Appl. Sci., Middle East Tech. Univ., Ankara, Turkey, 2013.
- [10] S. Jeong, S. W. Ban, S. Choi, and D. Lee, "Surface ship-wake detection using active sonar and one-class support vector machine," *IEEE J. Ocean. Eng.*, vol. 37, no. 3, pp. 456–466, Jul. 2012.
- [11] X. Lurton, *An Introduction to Underwater Acoustics: Principles and Applications*. New York, NY, USA: Springer-Verlag, 2002.
- [12] D. J. Crisp, "The state-of-the-art in ship detection in synthetic aperture radar imagery," Intell., Surveillance Reconnaissance Div. Inf. Sci. Lab., Cambridge, MA, USA, DSTO-RR-0272, 2004.
- [13] H. M. Finn and R. S. Johnson, "Adaptive detection mode with threshold control as a function of spatially sampled clutter-level estimates," *RCA Rev.*, vol. 29, no. 3, pp. 414–464, 1968.
- [14] R. Ravid and N. Levanon, "Maximum-likelihood CFAR for Weibull background," *Proc. Inst. Elect. Eng. Radar Signal Process.*, vol. 139, no. 3, pp. 256–264, Jun. 1992.
- [15] P. J. Grieve, "The optimum constant false alarm probability detector for relatively coherent multichannel signals in Gaussian noise of unknown power," *IEEE Trans. Inf. Theory*, vol. IT-23, no. 6, pp. 708–721, Nov. 1977.
- [16] A. Cetin, "CFAR Detection in k-distributed sea clutter," Ph.D. dissertation, Electron. Eng. Dept., Middle East Tech. Univ., Ankara, Turkey, 2008.
- [17] I. Quidu, L. Jaulin, A. Bertholom, and Y. Dupas, "Robust multitarget tracking in forward looking sonar image sequences using navigational data," *IEEE J. Ocean. Eng.*, vol. 37, no. 3, pp. 417–430, Jul. 2012.
- [18] M. Trevoror, "High frequency acoustic scattering and absorption effects within ship wakes," Defence Res. Develop. Canada, Ottawa, ON, Canada, DRDC-ATLANTIC-SL-2005-125, Aug. 2005.
- [19] D. Lerro and Y. Bar-Shalom, "Tracking with debiased consistent converted measurements versus EKF," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 29, no. 3, pp. 1015–1022, Jul. 1993.
- [20] Y. Bar-Shalom and T. Fortmann, *Tracking and Data Association*. New York, NY, USA: Academic, 1988.
- [21] D. Schulz, W. Burgard, D. Fox, and A. B. Cremers, "Tracking multiple moving targets with a mobile robot using particle filters and statistical data association," in *Proc. IEEE ICRA*, 2001, vol. 2, pp. 1665–1670.
- [22] I. Cox and S. Hingorani, "An efficient implementation of Reid's multiple hypothesis tracking algorithm and its evaluation for the purpose of visual tracking," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 18, no. 2, pp. 138–150, Feb. 1996.



**Imen Karoui** received the Engineer's degree from Ecole Polytechnique de Tunis, Tunis, Tunisia, in 2002 and the Ph.D. degree in signal processing and telecommunications from Telecom Bretagne, Plouzané, France, in 2007.

She is currently a Research Engineer with Ecole Nationale Supérieure de Techniques Avancées Bretagne, Brest, France. Her main research interests include texture segmentation, detection, and tracking applied to sonar and synthetic aperture radar images.



**Isabelle Quidu** received the Ph.D. degree in electronics from Université de Bretagne Occidentale, Brest, France, in 2001.

From 2002 to 2005, she was a Research Engineer with Thales Underwater Systems in the field of mine warfare. Since 2005, she has been an Associate Professor with Ecole Nationale Supérieure de Techniques Avancées Bretagne, Brest, and a member of the Ocean Sensing and Mapping Team. She is also a member of Lab-STICC UMR CNRS 6285. Her research interests include detection and classification of underwater targets in sonar images.



**Michel Legris** received the Engineer's degree from Ecole Centrale de Paris, Paris, France, in 1985 and the Ph.D. degree in electrical engineering from the University of Grenoble, Grenoble, France, in 1996.

From 1986 to 2001, he was an engineer with the French Ministry of Defense, where he did research on low magnetic fields, hydroacoustics noise and high-frequency sonar technology. Since 2001, he has been an Associate Professor with Ecole Nationale Supérieure de Techniques Avancées Bretagne, Brest, France. His current research interests include underwater acoustics systems and innovative sonar.