

## **A complete processing chain for ship detection using optical satellite imagery**

CHRISTINA CORBANE\*†‡, LAURENT NAJMAN§, EMILIEN PECOUL¶,  
LAURENT DEMAGISTRI† and MICHEL PETIT†

†ESPACE Unit, Institut de Recherche pour le Développement, Montpellier, France

‡Current affiliation: European Commission, Joint Research Center, Institute for the  
Protection and Security of the Citizen, Ispra, Italy

§Laboratoire d'Informatique Gaspard-Monge, Université Paris-Est, Equipe A3SI,  
ESIEE Paris, France

¶IUP Génie Physiologique et Informatique, Faculté des Sciences Fondamentales et  
Appliquées, Université de Poitiers, Poitiers, France

Ship detection from remote sensing imagery is a crucial application for maritime security, which includes among others traffic surveillance, protection against illegal fisheries, oil discharge control and sea pollution monitoring. In the framework of a European integrated project Global Monitoring for Environment and Security (GMES) Security/Land and Sea Integrated Monitoring for European Security (LIMES), we developed an operational ship detection algorithm using high spatial resolution optical imagery to complement existing regulations, in particular the fishing control system. The automatic detection model is based on statistical methods, mathematical morphology and other signal-processing techniques such as the wavelet analysis and Radon transform. This article presents current progress made on the detection model and describes the prototype designed to classify small targets. The prototype was tested on panchromatic Satellite Pour l'Observation de la Terre (SPOT) 5 imagery taking into account the environmental and fishing context in French Guiana. In terms of automatic detection of small ship targets, the proposed algorithm performs well. Its advantages are manifold: it is simple and robust, but most of all, it is efficient and fast, which is a crucial point in performance evaluation of advanced ship detection strategies.

### **1. Introduction**

Ship detection from satellite imagery is a valuable tool for the identification of illegal oil spills and monitoring maritime traffic in the fisheries, and the commercial transportation sector. Fishing, shipping and export of oil and natural gas are some of the world's largest industries. To ensure a sustainable development and the safety of people, a control system must be in place. The vessel monitoring system (VMS) that relies on a ship-borne component provides the authorities with a continuous monitoring of vessels' location and movements in real time. However, many ships are not equipped with these systems, for example smaller fishery vessels and passenger boats do not have to apply with the existing directives (e.g. European Commission (EC) directive 2002/59/EC). Remote sensing using Earth Observation can potentially detect all vessels, that is those with shipboard VMS units, those without VMS units and

---

\*Corresponding author. Email: [Christina.corban@jrc.ec.europa.eu](mailto:Christina.corban@jrc.ec.europa.eu)

those with faulty VMS units. Satellite imaging might be also useful for remote monitoring of areas or fisheries where traditional surveillance methods and VMS are not feasible. Also, the combined use of a VMS and satellite imaging could be more effective than a VMS alone.

Space-based imaging for ship detection and maritime traffic surveillance has often formed part of major research efforts in the fields of automatic target detection and recognition. Basically two remote sensing techniques have been employed for ship detection: synthetic aperture radar (SAR) with capacity to image day and night under most meteorological conditions became the state-of-the-art technique for ship detection (Crisp 2004). Several papers in the open literature treat methods for the detection of ship targets on SAR data. Inggs and Robinson (1999) investigated the use of radar range profiles as target signatures for the identification of ship targets and neural networks for the classification of these signatures. Tello *et al.* (2004) used wavelet transform by means of the multiresolution analysis to analyse multiscale discontinuities of SAR images and hence detected ship targets in a particularly noisy background. The relative improvement in ship detection performance of polarimetric SAR was evaluated in comparison with single-channel SAR by Liu *et al.* (2005). Given the long history and ongoing interest, there is an extensive literature on algorithms for ship detection in the literature. However, in terms of operational performance, Zhang *et al.* (2006) reported limitations of SAR in identifying smaller ships in inland waters. Besides, because of the presence of speckle and the reduced dimensions of the targets compared with the sensor spatial resolution, the automatic interpretation of SAR images is often complex even though vessels undetected are sometimes visible to eye. The second technique for ship detection lies on optical remote sensing, which has been explored since the launch of Landsat in the 1970s. McDonnell and Lewis (1978) demonstrated the possibility to detect ships of 100-m length using Landsat Multi-Spectral Scanners (MSS). Burgess (1993) applied Landsat Thematic Mapper (TM) and Satellite Pour l'Observation de la Terre (SPOT) data to identify smaller ships. In a recent work, Corbane *et al.* (2008) developed an approach based on genetic algorithms and neural networks for the detection and classification of small fishing boats on 5-m resolution SPOT 5 imagery. Increased false alarm rates (FARs) were obtained when using this approach on particular types of images with a high percentage of cloud cover and a cluttered sea background.

Compared to the large amount of investigations on the feasibility of satellite-based SAR for ship detection purposes, far less research and development activity has taken place in automatic detection and classification of vessels using optical imagery than using SAR imagery. This is a consequence of the novelty of high-resolution optical satellite sensors, the problem of clouds and the fact that the swath of high-resolution imagery is relatively small, making it less suitable for surveillance over the oceans. However, high spatial resolution can complement SAR because it is most suitable for ship classification and it permits the detection of small ships and wooden and fibreglass boats, which are more difficult to detect with radar (Greidanus *et al.* 2004). This topic is at the heart of operational capabilities targeted by several French and European projects (Implementation of Boat Information System (IBIS), Detection and Classification of Marine Traffic from Space (DECLIMS), Land and Sea Integrated Monitoring for European Security (LIMES)). Thus, in the framework of a European integrated project Global Monitoring for Environment and Security (GMES)/LIMES, we have developed a pre-operational tool for automatic ship detection on high spatial resolution optical imagery to complement existing fishery control measures, in particular the VMS. The

particularity of the system is its ability to deal with images with high percentages of cloud cover and cluttered background arising in particular meteorological situations. The approach consists of three main steps: (1) a preprocessing stage involving *cloud masking* and *local contrast enhancement*, (2) a prescreening stage including *automatic threshold estimation* and *connected filtering* using component trees for the detection of potential ship targets and (3) a postprocessing stage where membership probabilities to the ‘ship’ category are estimated using a *logistic model* based on the variables obtained from *wavelet transform* and *Radon transform* (RT). The following section of this article is dedicated to the algorithm’s description. The third section addresses the validation of the system on SPOT 5, 5-m resolution images. The validity of the developed procedure for operational ship detection is discussed in the last section along with some directions for future work.

## 2. Detection algorithm

The ship detection problem can be considered as a simple detection of bright point targets against a noisy background. However, the reality is more complicated because of possible confusions associated with small clouds and wave crests that could be falsely detected as ships. An optimum detector for this situation should maximize the probability of detection while minimizing the probability of false alarm. We therefore adopted a strategy that aims at maximizing the detection of ship targets in a pre-screening phase and then in minimizing the false alarms by assigning membership probabilities to the results of the detection in a postprocessing phase. Hence, the final decision is left to the operator who can then validate the results of the detection based on his experience. The basic structure of the detection algorithm is summarized in figure 1. The following sections describe in detail the three major steps of the algorithm: preprocessing, prescreening and postprocessing.

### 2.1 Preprocessing

The aim of preprocessing is to facilitate the subsequent detection stages by first decomposing the image into rectangular segments or tiles, and then by using statistical characteristics of the data in these segments to extract localized information for the cloud-masking operation.

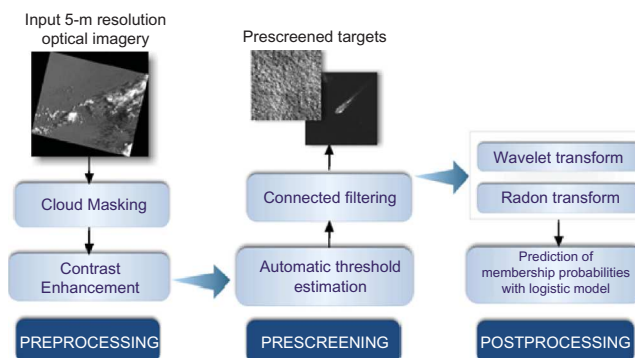


Figure 1. Flowchart of the three-phase algorithm for automatic ship detection.

- (a) *Image tiling*: The average size of SPOT 5 (5-m resolution) image data is around  $15,000 \times 15,000$  pixels. The original scene is divided into 25 tiles of  $3000 \times 3000$  pixels. This operation results not only in a faster processing time but also in a more effective handling of local information especially in images with heterogeneous sea clutter background.
- (b) *Cloud masking*: The intensity of SPOT 5 image data is coded in 8 bits (256 grey levels). The presence of brightly reflecting clouds over the sea surface can detract from the application of contrast enhancement techniques. Besides, clouds can contaminate the prescreening approach as small clouds can be easily mistaken for potential ship targets. For the removal of cloud pixels, we developed a routine based on the assumptions that (i) an image segment has a Gaussian distribution and (ii) cloud pixels correspond to the brightest pixels. The modal value of the grey-level histogram corresponds to sea background pixels. Cloud masking is then tackled using threshold information determined through histogram method. Typically, the threshold value is empirically derived from a dataset consisting of 15 SPOT 5 images acquired between 2003 and 2007 at different dates and thus under different sun elevation angles. This demonstrates that the threshold is not sensitive to the illumination conditions. The cloud threshold value is determined according to the following equation:

$$\text{Cloud threshold} = \text{Modal value} + 150(\text{DN}) \quad (1)$$

where DN represents digital number. All pixels greater than or equal to the cloud threshold value are then masked following this principle.

With this empirically derived value, the cloud masking performed well for bright and thick clouds. Although thin clouds could not be easily masked, this is not problematic for the following processing steps that are not affected by the presence of thin clouds (i.e. the latter are easily differentiable from ship targets). We recommend, however, the adjustment of this threshold value for other types of optical images.

- (c) *Contrast enhancement*: A linear contrast stretching is applied on each image segment. However, instead of computing the lowest pixel intensity, the modal value of the Gaussian distribution is considered as the minimum value. Hence the highest ( $x_{\max}$ ) and the modal intensity values ( $x_{\min}$ ) are set to 255 ( $y_{\max}$ ) and 0 ( $y_{\min}$ ), respectively, and all other pixel intensities are scaled accordingly. This yields considerable improvement for subsequent ship target detection.

## 2.2 Prescreening

At the heart of the algorithm, the prescreening phase consists in searching the image segments for potential ship pixels. Several classical segmentation techniques have been attempted, none of which were particularly successful. For example, standard thresholding techniques did not perform well because of the difficulties in automatically determining threshold values for the various sea state conditions. Even local thresholding methods did not yield satisfactory results because of the non-bimodal histograms and because both the ship targets and the sea background assume some broad range of grey-level values. Neural networks have been attempted for the detection and classification of ship targets (Corbane *et al.* 2008); however, because

of the complexity of the training phase and the high number of false alarms in high sea clutter situations, this approach was subsequently abandoned. Finally, the approach that was adopted for the prescreening relies on filtering tools derived from mathematical morphology and particularly from connected component transforms (Salember and Serra 1995, Jones 1999, Breen *et al.* 2000). Using mathematical morphology, image data can be filtered to either preserve or remove features of interest, sizing transformations can be constructed and information relating to shape, size and form can be easily applied (Breen *et al.* 2000). Besides, the large choice of morphological operators and their easy implementation offer the possibility of combining them into more complex operators that can solve a broad variety of sophisticated applications in image analysis and, in the case of this study, the ship detection problem.

The morphological operators implemented in our application belong to a specific class of operators called *connected operators*. These filtering tools interact with the signal by means of specific regions called *connected components* of the space where the image is constant: they eliminate the connected component that would be totally removed by an erosion with a given structuring element and they leave the other components unchanged (Salember and Garrido 2000). This filtering approach offers the advantage of simplifying the image, because some components are removed, as well as preserving the contour information, because the components that are not removed are perfectly preserved. In that sense, the connected operators are then different from the classical morphological operators, such as erosion and dilation with selected structuring element, that act on the individual pixels. A theoretical definition of connected operators can be found in Salember *et al.* (1998).

**2.2.1 The filtering strategy.** The filtering strategy adopted in this article is illustrated by figure 2. The image is considered as a 3D relief and the first step is to construct a tree structure of the image, called the component tree. The component tree is a representation of a grey-level image that contains information about each image component and the links that exist between components at sequential grey levels in the image (Najman and Couprie 2006). The nodes of the tree represent the connected components resulting from the thresholding of the original image at all possible grey-level values. The leaves of the tree correspond to the maxima of the image. The links between the nodes describe how the connected components may be merged. The tree structure is defined by the absolute grey levels of the connected components. Finally,

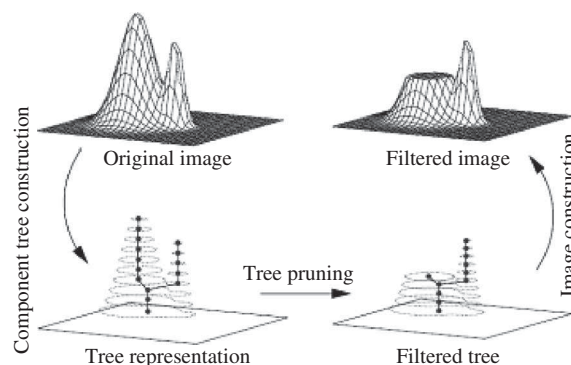


Figure 2. Connected operator: filtering strategy (Salember and Garrido (2000), with permission).

the root node corresponds to the lowest grey-level value. Once the component tree is constructed, the next stage is to introduce mechanisms to filter the component tree to decide which components are to be preserved and which are to be removed by the image filter. Filtering the tree or tree pruning is a decision-making process that classifies nodes into *active* and *non-active* nodes. A tree node will be referred to as *active* if it represents a component that is to be preserved by the filter. The filtering criterion, denoted by  $T$ , is based on one or more of the component's attributes, for example the area or perimeter of the component. The criterion requires that the values of these attributes lie within certain given thresholds. The notion of a connected filter may be formalized as follows:

*Definition 1.* If  $X \subseteq \mathbb{R}^2$  is a binary set and  $T$  is some filtering criterion, we define filter  $\Phi_T$  by

$$\Phi_T(X) = \cup\{C \subseteq X : C \text{ is a component satisfying criterion } T\} \quad (2)$$

*Definition 2.* If  $f$  is a grey-level image,  $\chi_s(f)$  is a threshold set at grey level  $s \in \Pi$ , where  $\Pi$  is a chain of integer grey levels, and  $\Phi_T$  is a binary connected filter; then we define a grey-level connected filter  $\Phi_T$  by

$$\Phi_T(f)(x, y) = \max \left\{ t : (x, y) \in \bigcap_{s < t} \Phi_T[X_s(f)] \right\} \quad (3)$$

At the end of the pruning, the filtered image is reconstructed by stacking the connected components corresponding to the remaining nodes.

In our application, a quasi-linear and fast algorithm for the construction of the tree on each image segment is used. The algorithm was developed by Najman and Couprie (2006) and it consists in computing the component tree on symmetric graphs based on Tarjan's (1975) union-find procedure. The success of filtering based on component trees is very dependent on the type of attribute used. In the case of detection of ship targets, the purpose is to discriminate pixels belonging to a ship from the rest of the image segment. More importantly, the idea is to avoid the problem of some remaining unmasked clouds and ocean homogeneities (i.e. transitions between regions with different wind conditions, low wind spiral marks, etc.) that give rise to a large number of false alarms. The features (small vessels) we are interested in are characterized by their brightness and by their small size compared with other features present in the scene. Therefore, among the numerous attributes that can be computed from the component tree (i.e. volume, perimeter, eccentricity, etc.), we selected the two most discriminating for the filtering criteria, the height and the area (figure 3), which are defined as follows: Let  $[k, c] \in C(f)$ :

$$\text{height}([k, c]) = \max\{f(x, y) - k + 1(x, y) \in c\} \quad (4)$$

$$\text{area}([k, c]) = \text{card}(c)$$

where  $C(f)$  is the component tree of image  $f$ ,  $c$  is a component of  $f$ ,  $[k, c]$  is a level  $k$  component of  $F$  and  $\text{card}$  stands for cardinal of  $c$ , which is the number of elements in the set  $c$ .

First bright targets are extracted using the height criteria and the result ( $\text{out}_{\text{height}}$ ) is submitted to a second filtering based on the area criteria. Because of the high sensitivity

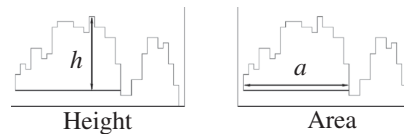


Figure 3. Illustration of the height- and area-filtering criteria of a component (Najman and Couprie (2006), with permission).

of the system to the height criteria, we propose to analyse this parameter in a separate paragraph (section 2.2.2). Because we are mainly interested in small fishing boats the threshold for the area-filtering criteria is set to 20. In other words, all components in the image that have an area of 20 pixels (=100 m) or less are removed by the transform; the remaining components are preserved in their entirety. Finally the resulting image ( $out_{area}$ ) is subtracted from the previous one ( $out_{height}$ ) allowing to mark only those pixels that correspond to potential ship targets (figure 4). Several tests undertaken with 4-connected and 8-connected components showed that the results were not very sensitive to this parameter. Therefore, for both the height- and the area-filtering criteria, we use 8-connected components for the connectivity parameter.

Determining the threshold value ( $t_h$ ) for the height-filtering criteria is critical in our application. Small modifications of the height criterion threshold involve drastic changes on the output and similar images may produce quite different results. Therefore, we have developed an adaptive threshold module that allows an automatic

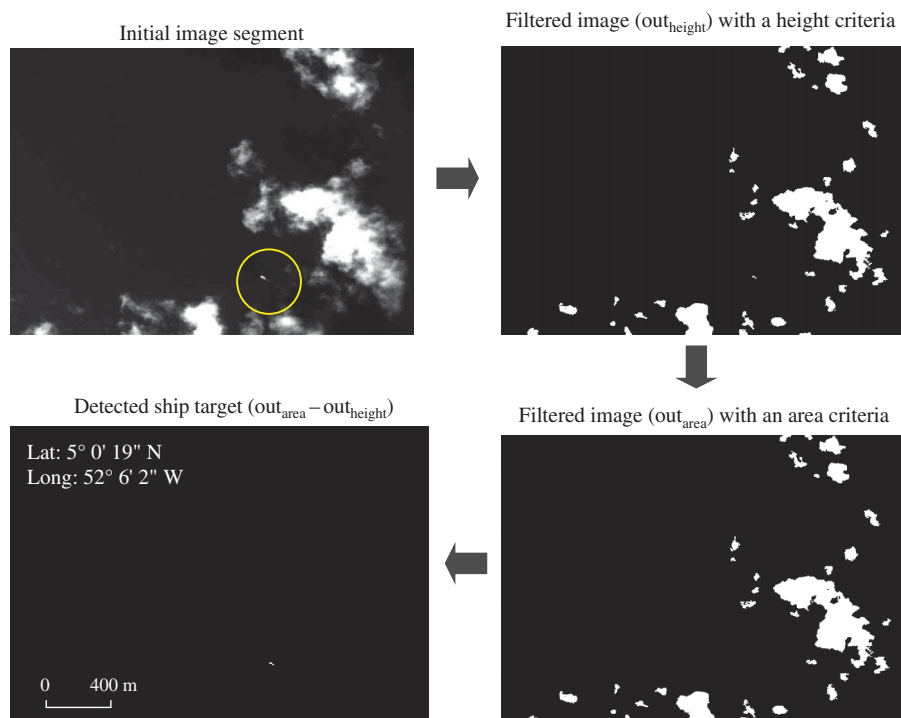


Figure 4. Illustration of the filtering strategy applied on image segment. For illustration purposes, the cloud masking was not applied prior to image filtering, allowing a better visualization of the results. The circle highlights the ship target to be extracted in the prescreening phase.

estimation of the threshold for the height attribute. It is based on local statistics of each image segment and is empirically calculated as a function of the width ( $W$ ) of the stretching range, defined in section 2.1:

$$W = y_{\max} - (x_{\max} - x_{\min}) \quad (5)$$

The threshold for the height criteria ( $t_h$ ) is then obtained by the following model:

$$t_h = Wa + \sigma + \bar{x}b \quad (6)$$

where  $\sigma$  is the standard deviation of the image segment,  $\bar{x}$  its mean, and  $a$  and  $b$  are the weighting factors that take alternately values 0.5 and 0.75 depending on the values of  $W$  and  $\sigma$  as follows:

- If  $W < 200$  and  $\sigma < 40$  then  $a = 0.5$  and  $b = 0.75$ ;
- If  $W \geq 200$  and  $\sigma < 40$  then  $a = 0.75$  and  $b = 0.75$ ;
- If  $W < 200$  and  $\sigma \geq 40$  then  $a = 0.5$  and  $b = 0.75$  and  $\sigma$  is multiplied by 2;
- If  $W \geq 200$  and  $\sigma \geq 40$  then  $a = 0.75$  and  $b = 0.75$  and  $\sigma$  is multiplied by 2.

These rules are dependent on the object scales and may therefore vary with the image's spatial resolution. They should, hence, be adapted to the spatial resolution of the images under process. Because, in our study, only SPOT 5, 5-m resolution images were available, the model was then calibrated for this resolution. Besides, the method was calibrated for small ship targets that present relatively homogeneous intra-target DN values. Consequently, the rules may need also to be adapted for targets with very heterogeneous DN values, such as very big ship targets.

### 2.3 Postprocessing

In the previous stage, we purposely used component operators based on rather unrestrictive criteria to avoid missed detections. Hence, a large number of false-positive detections may be expected as a result of the prescreening step. The purpose of the postprocessing stage is to assign membership probabilities to the potential ship targets obtained from the preceding stage. It is intended to supply the human expert with a first quantitative assessment of detection results, giving him the control of the final false alarm discrimination.

Assigning membership probabilities to the results of the prescreening phase can be considered as a dichotomous classification task where the class labels are either 'ship' or 'other'. Among the different types of data classification approaches, we chose the binary logistic regression, which is commonly used in statistical pattern recognition.

The logistic regression model calculates the class membership probability for one of the two categories 'ship' or 'other' given the information of explanatory variables:

$$P(Y = 1|X) = \frac{1}{1 + \exp[-\{\beta_0 + \beta_1 X_1 + \dots + \beta_i X_i\}]} + \varepsilon \quad (7)$$

where  $\beta_0$  is a constant term,  $\beta_i$  terms are the derived coefficients and  $X_i$  terms are the values of the variables used to determine the case classification (0 or 1 for dichotomous type).

The explanatory variables that are believed to have an effect on the dependent variable have to be identified. The idea is to look for some typical characteristics that



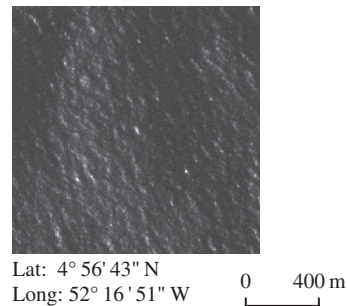


Figure 5. Example of false positive detection related to wind-wave crests.

allow us to further differentiate the true ship targets from wind-wave crests that result in false alarms during the prescreening phase.

As shown in figure 5, wind-wave crests may be visually identified as false alarms because of the presence of a relatively large number of uniformly oriented elongated features. Moreover, by managing the contextual information a human operator would be capable of focusing and relating the different features at different scales in the image. This intuitive visual perception allows the differentiation of bright spots and lines corresponding respectively to small and moving ship target from the surrounding oceanic sea surface turbulences. This supports the statement that software-based algorithms for ship detection may not be as good as a human operator, who is better at dealing with complex clutter situations (Greidanus and Kourti 2006). Because multi-scale processing is able to model the operation of the human vision, it seems interesting to analyse the detected targets by means of time–frequency methods and, in particular, by means of the wavelet tools (Suhling *et al.* 2004). Consequently, in the last stage, we propose to further process the candidate targets detected by the prescreening phase and to mimic the human vision by using the wavelet transform for multiscale analysis of the signal and RT to accentuate linear features.

### 2.3.1 Wavelet transform

Signal processing with wavelets is just one among the other time–frequency methods but it presents clear advantages. The short-time Fourier transform and the Wigner–Ville transform are not always suitable for transient phenomena. Besides, wavelet transforms have been used successfully for the detection/estimation in non-stationary environment (Qiang *et al.* 2005). Moreover, wavelet tools are especially well suited for their use in the processing of natural scenes because they are well adapted to analyse multifractal properties (Tello *et al.* 2006a). Wavelet analysis decomposes an image into a hierarchical set of approximation and detail wavelet maps. The approximation map contains the image's low-frequency information, whereas the detail maps contain the high-frequency information. At each level, the wavelet transform is applied to the approximation map, breaking it down into further approximation and detail maps (Mallat 1989, Mallat and Hwang 1992).

The wavelet transform starts with a mother wavelet. The mother wavelet is an irregular, asymmetric waveform of limited duration. There are many different mother wavelets, the choice of which depends on the application. The mother wavelet can be thought of as a 'window' that is shifted along the original signal. At each location, or

translation, along the signal the wavelet is correlated with the signal at that particular point. Once the wavelet has been translated to every point along the signal, the process is repeated. This time the wavelet is stretched, or dilated, to a larger scale. The wavelet scale is inversely related to frequency. A large scale corresponds to a low frequency, whereas a short scale corresponds to a high frequency. The final result of the process is a map of correlation values, called wavelet coefficients, corresponding to each translation (time) and scale (frequency).

The discrete wavelet transform (DWT) is employed in our algorithm. It allows a signal to be sampled at discrete points, resulting in efficient computation. Discrete wavelets are scaled and translated in discrete steps [13]. This is achieved using scaling and translation integers instead of real numbers. The following is the DWT equation:

$$\Psi_{j,k}(t) = \frac{1}{\sqrt{s_0^j}} \Psi\left(\frac{t - k\tau_0 s_0^j}{s_0^j}\right) \quad (8)$$

where  $j$  and  $k$  are integers with  $j$  determining the scale and  $k$  the translation. The scale describes the time domain width of the wavelet and the translation identifies the position of the wavelet with respect to the dataset. The rate of scale dilation is  $s_0$  and the translation step magnitude is  $\tau_0$ . The rate of scale dilation, together with the size of the dataset, governs the number of scales generated. The choice of the number of decomposition levels is clearly a trade-off between the size of the targets to detect and the presence of noise. In the case of images resulting from the prescreening phase, a visual inspection of the information at different scales showed that most of the vessels appear in the first level ( $j = 1$ ), even if their presence is transmitted over higher scales.

Among the different families of mother wavelets, the Haar was chosen because it is quite appropriate to spot detection according to Tello *et al.* (2006b). A 2D wavelet transform of an original image  $f(x, y)$  at a scale  $j$  produces four images at a scale  $j + 1$ : three detailed images in the horizontal ( $D_j^h$ ), vertical ( $D_j^v$ ) and diagonal ( $D_j^d$ ) directions and an approximation image ( $A_j^f$ ) (figure 6). The approximation image contains the image's low-frequency information, whereas the detailed maps contain the high-frequency information.

Figure 7 shows the results of the application of a 2D DWT, on three images obtained from the prescreening phase. Figure 7(a) corresponds to DWT of a real ship target in a flat sea state, whereas figure 7(c) corresponds to DWT of a ship target in a high sea state and figure 7(d) to DWT of a false alarm representing ocean turbulences. It is obvious that the presence of a ship target is noticeably enhanced by the DWT. Compared to figure 7(d), which corresponds to ocean turbulences, the ship target is appreciable in the approximation component ( $A_j^f$ ) represented in the upper left corner of the DWT in figures 7(a) and (c) where the peak intensity is mostly significant. A closer inspection of the approximation component (containing

$A_j^f$	$D_j^h$
$D_j^v$	$D_j^d$

Figure 6. The result of 2D DWT decomposition.

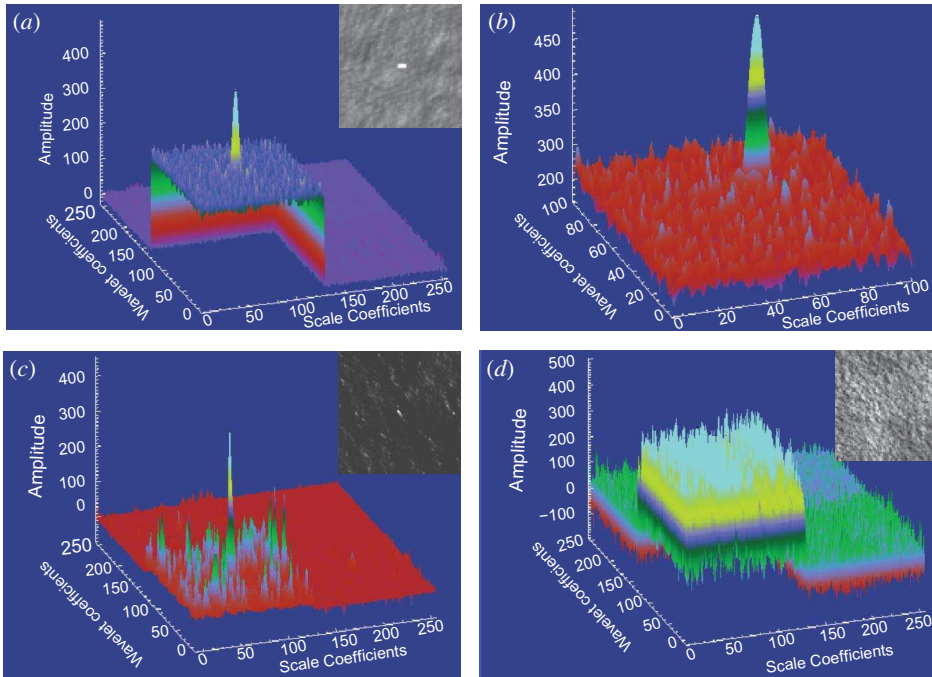


Figure 7. The result of 2D DWT decomposition applied on (a) a prescreened ship target in a flat sea surface, (c) a prescreened ship target in a high sea state and (d) a false positive corresponding to wind-wave crests. (b) represents a close view of the approximation component ( $A_j^1$ ) visible in the upper left corner of (a).

low-frequency information) shows that the central peak is significantly higher than the surrounding ones (figure 7(b)). Therefore, by counting the number of peaks in the approximation component and by calculating  $H_{DWT}$ , which is the difference in DWT values between the central peak and the surrounding ones, it is possible to detect the presence of a ship target. The effect of the variable  $H_{DWT}$  extracted from the DWT on the discrimination performances of the logistic model is studied in section 2.3.3.

**2.3.2 Radon transform.** In addition to the application of the DWT for an extended detection of prescreened targets, a RT was also applied to the results of the prescreening phase. RT has certain advantages as regards its computing efficiency for linear features detection compared with Hough transforms (Götz and Druckmüller 1996) or directional morphological operators (Couloigner and Zhang 2007).

The RT of a 2D function  $f(x,y)$  is the set of projections along angles  $\theta$ ,

$$Rf \equiv p(\rho, \theta) = \iint_{-\infty}^{+\infty} f(x,y) \delta(x \cos \theta + y \sin \theta - \rho) dx dy \quad (9)$$

$$= \int_{-\infty}^{+\infty} f(\rho \cos \theta - l \sin \theta, \rho \sin \theta + l \cos \theta) dl \quad (10)$$

$$\text{where } \begin{bmatrix} \rho \\ l \end{bmatrix} = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \text{ (coordinate rotation)}$$

where  $\delta(x)$  is the Dirac function,  $\rho \in \pi(-\infty, +\infty)$  and  $\theta \in [0, \pi]$ . The RT performs the integration of the image along each possible straight line of the image with polar parameters  $(\rho, \theta)$ . The RT of an image containing a segment will therefore exhibit a prominent peak of coordinates  $(\rho_0, \theta_0)$  such that  $\rho_0 = x \cos \theta_0 + y \sin \theta_0$  is the equation of the straight line along which the segment lies (Magli *et al.* 1999). For the continuous RT, back-projection is the adjoint operator to the transform,

$$R^* p \equiv b(x, y) = \int_{-\infty}^{+\infty} \int_0^\pi p(\rho, \theta) \delta(x \cos \theta + y \sin \theta - \rho) d\rho d\theta \quad (11)$$

$$= \int_0^\pi p(x \cos \theta + y \sin \theta, \theta) d\theta \quad (12)$$

Here we propose to apply the RT by integrating image intensity along all lines starting from the centre of each ship candidate obtained in the preceding prescreening phase. The image is then reconstructed by simply taking the inverse transform of the projection. This involves two steps; the image is back-projected and then filtered using a Laplacian of Gaussian filter, allowing further to highlight edges in the reconstructed image. Figure 8 displays the result of filtered back-projections applied to two detected targets, one representing a real ship (figure 8(a)) and one representing a false alarm (figure 8(b)) caused by the presence of wind waves. Clearly the ship target is well localized. The ship and its wake are enhanced and sharpened by the RT. They are represented by a central peak significantly higher than its surrounding. Conversely, the peak corresponding to a wave crest in the centre of figure 8(b) is faintly visible in the centre of the 3D representation of the back-projection. The significance of the central peak is tested by calculating its relative height. This is a ratio of the central peak to the mean peak value as used and defined in Hill *et al.* (2000). It is defined by the variable  $H_{RT}$ , which is thought to be relevant for the separation of the prescreened targets into two categories ‘ship’ and ‘other’.

**2.3.3 Logistic regression model.** The predictive logistic model, used for assigning membership probabilities to each detected target, was built based on a dataset consisting of 15 SPOT 5 images with 186 targets among which 54 correspond to ship

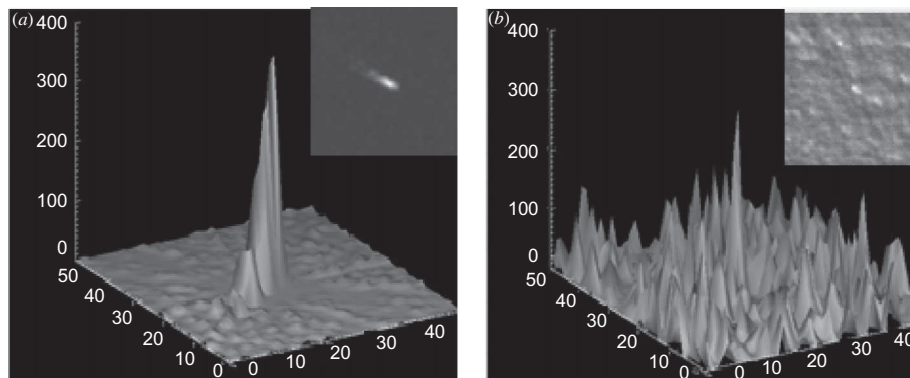


Figure 8. The result of Radon Transform applied on two detected targets: a real ship (a) and a false positive (b).

Table 1. Estimates of parameters (statistical significance  $p < 0.05$ ) and goodness-of-fit test measures for the fitted logistic model.

Predictor	Estimate	Standard error	$p$ -level
Constant $\beta_0$	-2.65	0.50	0.000
$\beta_1 (H_{RT})$	0.045	0.09	0.000
$\beta_2 (H_{DWT})$	0.0067	0.002	0.013
Test	Degrees of freedom (Df)	Statistics	Df/statistics
<i>Goodness-of-fit</i>			
Deviance	174	130.62	1.33
Pearson $\chi^2$	174	172	1.01
Loglikelihood		-16.57	

samples. Variable selection was performed with  $H_{DWT}$  and  $H_{RT}$  (obtained respectively from the analysis of DWT and RT of the data) as explanatory variables. We tested the statistical significance of the coefficients with an automatic stepwise *forward selection* procedure that starts with a simple model and add terms sequentially until further additions do not significantly improve the fit. Both  $H_{DWT}$  and  $H_{RT}$  were significant predictors ( $p < 0.05$ ). The test of the intercept (the constant  $\beta_0$  in table 1) merely suggests that an intercept should be included in the model ( $p < 0.05$ ).

Goodness-of-fit statistics were also calculated for assessing the fit of the logistic model against actual outcomes. The ratios of the various statistics (deviance, Pearson  $\chi^2$ ) to the respective degrees of freedom are close to 1.0. Thus, there is no evidence of overdispersion. This suggests that the values of the parameters' estimate for variables  $H_{RT}$  and  $H_{DWT}$  are appropriately scaled. The fitted logistic model is used in the processing chain to assign predicted membership probabilities to each prescreened target during the final postprocessing stage.

### 3. Experimental results

For evaluating the performance of the developed procedure, 37 SPOT 5 images with a high-resolution panchromatic band (5 m) were acquired over the Exclusive Economic Zone (EEZ) of French Guiana. These images were not used in the model building process. They were provided by the Direct Receiving Station of SPOT 5 satellite, operating under the SEAS-Guyane (Survey of Environment of the Amazonia Assisted by Satellites) program. For the detection of ship targets, panchromatic imagery was preferred over multi-spectral, because additional bytes (bandwidth) of information are better spent on increased resolution than on additional colour. On SPOT 5 optical images of 5-m resolution, ships are easy to detect with the human eye; their size is readily estimated and details on the superstructure can easily be discerned. Some of the larger vessel types can be immediately recognized, such as container ships, oil tankers and bulk carriers. Intermediate vessels such as shrimp boats that range from 20 to 25 m in length still show details, but their interpretation is not so straightforward: it is difficult for an untrained interpreter to discern, for example, a fishing vessel from a patrol boat.

The results for small ship detection on the 37 SPOT 5 images using the developed algorithm are represented in table 2. It is generally difficult to correctly cross-check the results of automatic ship detection because only limited ground truth information is available concerning ship positions. Moreover, unavailability of Automatic

Table 2. Good and false detection results of algorithm validation on 37 images represented in the form of intervals of *membership probabilities* ( $Mp$ ). The latter are predicted from the logistic model during the postprocessing phase. DR and FAR are calculated according to probability threshold values  $Mp > 30\%$  and  $Mp > 70\%$ .

Membership probabilities ( $Mp$ ) to 'ship' category (%)	Good detections				False detections			
	<30	30–50	50–70	>70	<30	30–50	50–70	>70
Number of detected targets	2	3	18	50	918	30	50	27
	DR				FAR			
For $Mp > 30\%$	89.8%				135%			
For $Mp > 70\%$	63.3%				34.1%			

Number of ships reported by human operator = 79.

Identification System data in French Guiana precluded a correct validation of the algorithm's performance. Nevertheless, in our case, visual interpretation by trained human operators was used to help assess performance. Seventy-nine ship targets were identified by human operators. Performance was measured by detection rate (DR) and FAR. DR is the number of ships correctly detected as a percentage of the total number of real ships and FAR is the number of ships incorrectly reported as a percentage of total number of real shrimp boats. As a reminder, the algorithm output is a detection bulletin with membership probabilities ( $Mp$ ) assigned to each detected target. Hence, detection results in table 2 were represented in the form of intervals of  $Mp$  and separated into good detections and false detections for a more detailed performance analysis.

A total of 2000 possible targets were detected by the algorithm. The specific distribution of these positives is as follows:

- 73 good detections;
- 1027 false detections.

The classification of good detections and false detections into intervals of membership probabilities allows the refinement of the evaluation criteria of the results. If we set the probability threshold value to 30%, the total number of good detections falls to 70 and the false detections to 107. Consequently the DR would be 89.8% whereas the FAR would be 135%. If the probability threshold is set to 70%, the total number of good detections would be only 50 with a DR of 63.3%. However, the total number of false detections would be far lower with only 27 false positives and a FAR of 34.1%.

From the foregoing results analysis, it is manifest that the detection performances of the algorithm are strongly related to membership probabilities automatically assigned to the detected targets but most of all to the detection threshold fixed by the human operator. Thus exists the classical battle between sensitivity and false alarms; that is, the desire to increase the DR is offset by the resulting increase in the FAR. To avoid this dilemma, the system presented in this article does not include an automated FAR. Instead it provides a graphical user interface (GUI) displaying the original image with detection overlaid together with a detailed detection bulletin and quick looks of the detected targets enabling the intervention of an experienced operator for final false alarm discrimination.

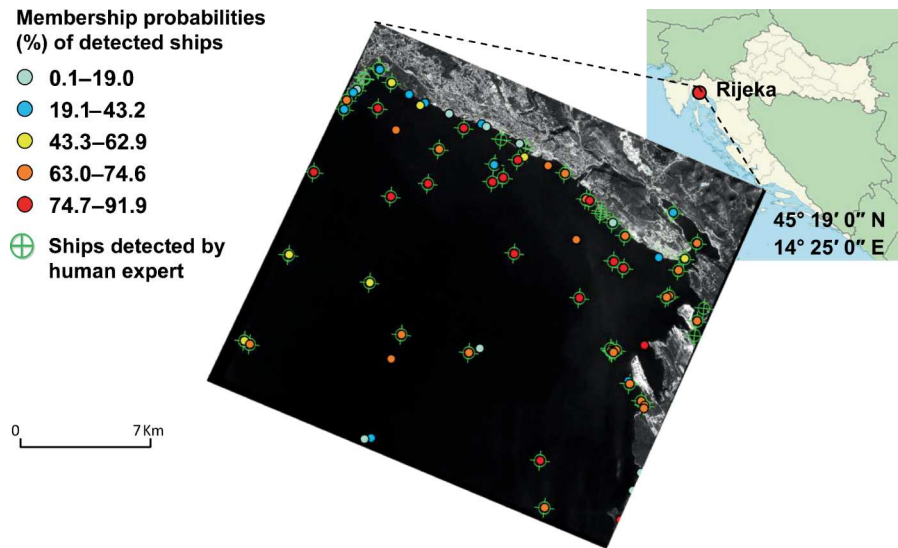


Figure 9. Result of automatic ship detection during the operational validation of the system over Rijeka harbour in Croatia.

In addition to the performance and qualification tests performed over the EEZ of French Guiana, the system was also validated over the Adriatic Sea in a pre-operational context. A pilot exercise was conducted on 14 October 2008 over Rijeka Harbour with the purpose of assessing the real potentials and limitations of space-borne (SAR and optical data) and airborne technologies (hyperspectral aerial photos) for the surveillance of oil spills and for identifying the position of probable pollution culprit. One-eighth of a SPOT 5 scene covering an area of  $20 \times 20$  km was acquired for the demonstration case study. The acquired image was available a few hours after its acquisition allowing to run the validation exercise in almost operational conditions. The result of automatic ship detection, shown in figure 9, was obtained in less than 3 minutes after image reception. The coloured circles represent the membership probabilities of the detected ship targets, whereas the cross circle corresponds to those targets identified by a human expert. It can be noticed that there are six missed targets and only five false alarms with membership probabilities greater than 70%. This satisfactory result illustrates the suitability of the system for operational maritime surveillance. It is a first step towards the definition of conditions and interrelations for an information fusion approach that combines data provided from different observation technologies for monitoring oil spills.

#### 4. Discussion and conclusion

We presented a system designed for automatic ship detection on high spatial resolution optical satellite imagery. The procedure is completely unsupervised and takes few minutes (5–10 minutes), depending on hardware characteristics. The heart of the algorithm is the filtering strategy (prescreening phase) for the detection of potential ship targets. It lies on connected operators derived from mathematical morphology. This promising technique for ship detection was augmented by a preprocessing and a postprocessing to reduce both the computation time and the number of false alarms.

The preprocessing involves cloud masking and local contrast enhancement of image tiles. The postprocessing involves the application of wavelet transform and RT. Using these two techniques, it is possible to refine the results of the prescreening phase by assigning membership probabilities to the 'ship' category predicted from a fitted logistic model. The algorithm was calibrated on 15 SPOT 5, 5-m resolution, panchromatic images and validated on an independent dataset comprising 37 images, all acquired over the EEZ of French Guiana. The results are promising. There were no missed detections and 73 detected targets for a total of 79 ship targets visually identified by human operators. However, in an absolute way, the total number of false detections (1027) for 37 processed images is unacceptably high for satellite ship surveillance applications. Most of these false alarms are due to local wind turbulences, which increase background noise. Some other false alarms are related to confusions between moving targets with characteristic ship wakes and other naturally produced linear features such as internal waves or wind-wave crests. Some large boats with their wakes were considered as several small ship targets and hence resulted in multiple detections that augmented the total number of false positives.

When we take into account membership probabilities, we notice that most of these false-positive detections are assigned low probabilities (i.e. 918 false positives have membership probabilities less than 30%). This evinces the potential of wavelet transform and RT for dealing with complex clutter situations and for the discrimination between ship wakes and other naturally occurring linear features. This suggests that an automatic wake detection module would obviously enhance the ship detection system. The presence of a wake confirms the presence of a ship and further it can be used to estimate the ship heading and speed. As for multiple detections of large ships, they may be reduced or eliminated by applying bounds on target area (size), which can be implemented using morphological filters (i.e. dilation that connects the neighbouring ship pixels and thereby clusters them). Compared to the detection system developed by Corbane *et al.* (2008) and based on non-parametric techniques such as neural networks and genetic programming, the current approach provides a much lower FAR. For instance, when tested on the same image with a highly cluttered background and high percentage of cloud cover, the algorithm based on neural network clustering resulted in a FAR of 5700% whereas the current approach yielded only one false positive for a DR of 100%.

The promising detection performances and the modest computational cost of the algorithm (the CPU time for an average  $15,000 \times 15,000$  pixels image is around 5 minutes on a standard 2 GHz Pentium IV-class machine) are a proof of its potential for providing a recognition system to a variety of users such as coast guards, search and rescue teams and harbour masters. The algorithm is currently used in a pre-operational framework. A great deal of effort is being undertaken to improve the validation procedures and control efficiency by (i) introducing information from other maritime monitoring systems, such as VMS, and (ii) cross-cuing to other sensors, such as SAR sensor, for obtaining or confirming detections. Further experiments are required to render the system fully operational. For instance, it would be interesting to adapt the algorithm to various ship sizes and types and to test it on very high spatial resolution optical imagery (sub-metric pixel resolution). Other possible extensions to this work could be the automatic determination of threshold values, mainly the cloud screening and the height criteria thresholds. Current research efforts are directed towards the identification of additional characteristics that enhance the discrimination of true targets from wind-wave crests. Mainly, directional morphological filters with line



segments as structuring elements are being investigated for a more noise-robust extraction of ships.

### Acknowledgements

This work was conducted within the framework of LIMES project (funded by EU). The authors acknowledge SEAS-GUYANE Project for the valuable support.

### References

- BREEN, E.J., JONES, R. and TALBOT, H., 2000, Mathematical morphology: A useful set of tools for image analysis. *Statistics and Computing*, **10**, pp. 105–120.
- BURGESS, D.W., 1993, Automatic ship detection in satellite multispectral imagery. *Photogrammetric Engineering & Remote Sensing*, **59**, pp. 229–237.
- CORBANE, C., MARRE, F. and PETIT, M., 2008, Using SPOT-5 HRG data in panchromatic mode for operational detection of small ships in tropical area. *Sensors*, **8**, pp. 2959–2973.
- COULOIGNER, I. and ZHANG, Q., 2007, Iterative and localized Radon transform for road center-line detection from classified imagery. In *Proceedings of the International Conference on Image Analysis and Recognition*, 22–24 August 2007, Montreal, Canada, M. Kamel and A. Campilho (Eds.), pp. 1093–1104 (Berlin: Springer-Verlag).
- CRISP, D.J., 2004, *The State-of-the-Art in Ship Detection in Synthetic Aperture Radar Imagery*, pp. 115 (Australian Government, Department of Defence).
- GÖTZ, W.A. and DRUCKMÜLLER, H.J., 1996, A fast digital Radon transform – an efficient means for evaluating the Hough transform. *Pattern Recognition*, **29**, pp. 711–718.
- GREIDANUS, H. and KOURTI, N., 2006, Findings of the DECLIMS project – detection and classification of marine traffic from space. In *SEASAR*, 23–26 January 2006. Available on CD-ROM.
- GREIDANUS, H., CLAYTON, P., INDREGARD, M. and STAPLES, G., 2004, Benchmarking operational SAR ship detection. In *Geoscience and Remote Sensing Symposium*, Anchorage, Alaska (Piscataway, NJ: IEEE International).
- HILL, K.L., ROBINSON, I.S. and CIPOLLINI, P., 2000, Propagation characteristics of extratropical planetary waves observed in the ATSR global sea surface temperature record. *Journal of Geophysical Research C: Oceans*, **105**, pp. 21927–21945.
- INGGS, M.R. and ROBINSON, A.D., 1999, Ship target recognition using low resolution radar and neural networks. *IEEE Transactions on Aerospace and Electronic Systems*, **35**, pp. 386–393.
- JONES, R., 1999, Connected filtering and segmentation using component trees. *Computer Vision and Image Understanding*, **75**, pp. 215–228.
- LIU, C., VACHON, P.W. and GELING, G.W., 2005, Improved ship detection with airborne polarimetric SAR data. *Canadian Journal of Remote Sensing*, **31**, pp. 122–131.
- MAGLI, E., OLMO, G. and PRESTI, L.L., 1999, Pattern recognition by means of the Radon transform and the continuous wavelet transform. *Signal Processing*, **73**, pp. 277–289.
- MALLAT, S. and HWANG, W.L., 1992, Singularity detection and processing with wavelets. *IEEE Transactions on Information Theory*, **38**, pp. 617–643.
- MALLAT, S.G., 1989, Theory for multiresolution signal decomposition: the wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **11**, pp. 674–693.
- MCDONNELL, M.J. and LEWIS, A.J., 1978, Ship detection from LANDSAT imagery. *Photogrammetric Engineering and Remote Sensing*, **44**, pp. 297–301.
- NAJMAN, L. and COUPRIE, M., 2006, Building the component tree in quasi-linear time. *IEEE Transactions on Image Processing*, **15**, pp. 3531–3539.
- QIANG, Q., JILIU, Z., KUN, H. and JIAN, L., 2005, Noise reduction based on wavelet transform under non-stationary environments. *Mechatronics and Automation, 2005 IEEE International Conference*, **4**, pp. 2123–2129.

- SALEMBIER, P. and GARRIDO, L., 2000, Binary partition tree as an efficient representation for image processing, segmentation, and information retrieval. *IEEE Transactions on Image Processing*, **9**, pp. 561–576.
- SALEMBIER, P. and SERRA, J., 1995, Flat zones filtering, connected operators, and filters by reconstruction. *IEEE Transactions on Image Processing*, **4**, pp. 1153–1160.
- SALEMBIER, P., OLIVERAS, A. and GARRIDO, L., 1998, Antiextensive connected operators for image and sequence processing. *IEEE Transactions on Image Processing*, **7**, pp. 555–570.
- SUHLING, M., JANSEN, C., ARIGOVINDAN, M., BUSER, P., MARSCH, S., UNSER, M. and HUNZIKER, P., 2004, Multiscale motion mapping: A novel computer vision technique for quantitative, objective echocardiographic motion measurement independent of Doppler. First clinical description and validation. *Circulation*, **110**, pp. 3093–3099.
- TARJAN, R.E., 1975, Efficiency of a good but not linear set union algorithm. *Journal of Association for Computing Machinery*, **22**, pp. 215–225.
- TELLO, M., LÓPEZ-MARTÍNEZ, C. and MALLORQUI, J.J., 2006b, Automatic vessel monitoring with single and multidimensional SAR images in the wavelet domain. *ISPRS Journal of Photogrammetry and Remote Sensing*, **61**, pp. 260–278.
- TELLO, M., LÓPEZ-MARTÍNEZ, C. and MALLORQUI, J., 2006a, A novel algorithm for automatic ship and oil spill detection based on time-frequency methods. In *Advances in SAR Oceanography from Envisat and ERS Missions, Proceedings of SEASAR 2006*. (Frascati, Italy: European Space Agency).
- TELLO, M., LOPEZ-MARTINEZ, C., MALLORQUI, J. and GREIDANUS, H., 2004, A novel algorithm for ship detection in Envisat SAR imagery based on the wavelet transform. In *Proceedings of the 2004 Envisat & ERS Symposium*, L. Ouwehand (Ed.), Salzburg, Austria (The Netherlands: European Space Agency).
- ZHANG, F., WU, B., ZHANG, L., HUANG, H. and TIAN, Y., 2006, Illicit vessel identification in inland waters using SAR image. In *International Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 3144–3147 (Piscataway, NJ: IEEE International).