AUTOMATIC SHIP DETECTION IN SPACE-BORNE SAR IMAGERY

F. Meyer [a], S. Hinz [b]

[a] University of Alaska Fairbanks, Geophysical Institute, 903 Koyukuk Dr., Fairbanks, Alaska-99775
[b] Institute for Photogrammetry and Remote Sensing, Universität Karlsruhe, Kaiserstrasse 12, 76128 Karlsruhe
fmeyer@gi.alaska.edu ; stefan.hinz@ipf.uni-karlsruhe.de

ABSTRACT:

Synthetic aperture radar (SAR) imagery has proven to be a promising data source for the surveillance of maritime activity, and its application for automatic ship detection has been the focus of many research studies. Apart from the well-known CFAR detector, there has emerged a novel method for automatic ship detection, based on the wavelet transform. Since the underlying principles for both methods are fundamentally different, their advantages and disadvantages concerning various image features also differ.

Within this paper we will present a prototype ship detection system that attempts to combine the benefits yielded by the two aforementioned techniques, thus gaining both sensitivity for weak targets and robustness against false alarms in inhomogeneous areas. For this, a wavelet-based prescreening stage is applied, which is followed by an object analysis, and a final adaptive-threshold test. The prototype has been tested and assessed on ALOS PALSAR and RADARSAT-1 data, especially with respect to the behavior toward sea-ice areas and irregularities such as beam seams in ScanSAR imagery. The results indicate a compensation of the intrinsic drawbacks held by the individual detection methods, producing a reliable and versatile detection system.

1. INTRODUCTION

The observation of maritime activity has been a field of research ever since synthetic aperture radar (SAR) imagery of the ocean surface became available for the first time. However the large amounts of image data modern SAR systems generate are capable of quickly overburdening a human observer. This creates an obvious need for automated detection systems, guaranteeing both reliability and predictability for the detection result.

For the purpose of an operational ship surveillance system, it is crucial to maximize the spatial coverage. The use of ScanSAR data can help to achieve this goal, however, it comes along with a sacrifice of image resolution. This means that the typical input imagery for an automatic ship detection system for surveillance purposes is expected to have a resolution range from 25 to 100m, causing some ships to be represented only by very few pixels, as illustrated by Figure 1. The task of ship detection under these circumstances can equally be described as spot detection in a noisy environment.

The reminder of this paper is organized as follows: Section 2 introduces the procedure for ship detection on a conceptual basis, while Section 3 outlines important implementation issues. A performance evaluation and discussion of the system is provided in Section 4, including an assessment of challenges and potential of ship detection in the new class of high-resolution spaceborne SAR data provided by TerraSAR-X, CosmoSkyMed, etc.

2. DETECTION STRATEGY

The detection of ships relies on three individual modules that are run independently, and whose results are combined in a subsequent fusion step. Initial ship detection is based on wavelet-analysis (Section 2.1), which is combined with a modified and enhanced Constant False Alarm Rate (CFAR) spot detector (Section 2.2). As the information provided by the raw SAR signal is rather limited due to the relatively coarse resolution of the used PALSAR and RADARSAT-1 imagery, we include automatically delineated context information in form of land-/watermasks, which are derived by texture-based classification. The focus is put on properties that are relevant

for the performance of the prototype. For a more general overview we refer to the respective technical literature.
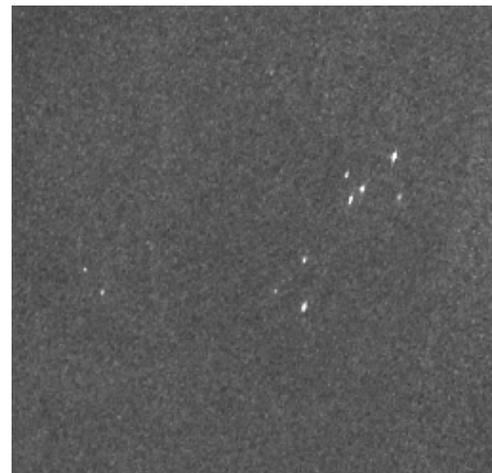
Figure 1: Group of ships in ALOS WB1 scene near Dutch Harbor, Alaska (100m spatial resolution).

2.1 Wavelet Analysis

The discrete Wavelet transformation has numerous applications in the field of SAR data, be it general processing, compression or pattern identification (Chan 1993, DeGrandi 2007). Recently, there also has been published extensive and promising research on its applicability on the detection of ships in artificial and real imagery (Tello et al. 2005). The first stage of the prototype ship detection system presented here is heavily inspired by the findings published in these papers.

The behavior of any transformation is determined by the choice of the appropriate transformation kernel, in case of the wavelet transformation simply referred to as "the wavelet". For the purpose of ship detection, it has been shown that the utilization of the so-called Haar wavelet (Figure 2) provides good results. This setup has several basic properties which make it very suitable for ship detection (Tello et al. 2006):

- **Edge detection**

  The Haar wavelet enhances steps in contrast of a signal, very much like an edge-detection filter. Ships can be regarded as singular structures in a noisy background, therefore generating conspicuous response to an edge-detection method.

- **Spatial locality**

  Unlike the kernels of other common transformations like trigonometric functions, wavelets are highly localized in both the spatial and frequency domain. Since in the application of ship detection not only their mere occurrence, but also their position is of prior interest, this property is of vital function.

- **Multi-resolution**

  Wavelets can be applied in different scales, in each case emphasizing only structures corresponding to the current scale. By varying the decomposition level, or respectively the number of iterations, a selective multi-scale analysis can be performed.
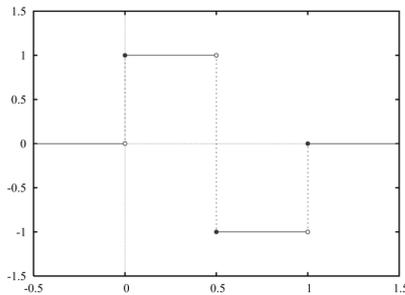
Figure 2: 1st order Haar wavelet. Higher order Haar wavelets are spatially scaled versions of the 1st order kernel, thus only extracting details which resemble their respective size.

The wavelet decomposition of a two-dimensional image creates sets of sub-band coefficients, which contain the extracted information with respect to the different directions H (horizontal), V (vertical) and D (diagonal). Ships show up as deterministic signals in all of the sub-bands, while noise effects decorrelate due to the orthogonality between the different dimensional channels. For ship detection, after the decomposition a spatial correlation is applied to the several sub-bands, by which only spot-like deterministic features become enhanced, while noise and arbitrarily-shaped irregularities are suppressed (Figure 3).

It is obvious that not only ships display as roughly spot-like objects in SAR imagery. Heavily inhomogeneous structures like sea-ice borders or very irregularly-shaped beam seams still can cause similar deterministic features, causing a significant number of false alarms (Figure 4). To address this problem by raising the threshold would only shift the problem to a loss of valid targets, thus increasing the false negative errors.

Also, the method's intrinsic property of edge detection implies that its success partly relies on a sharp-cut edge of the feature that is to be extracted. This can impose severe problems especially for lower resolution (>50m) imagery, since here targets often appear diluted and do not show a clean edge, resulting in false negative decisions.

A simple adaption of the decision threshold at this stage would preserve these targets, but at the same time cause a dramatic increase of false alarms for other structures found in inhomogeneous regions.


(a)   (b)
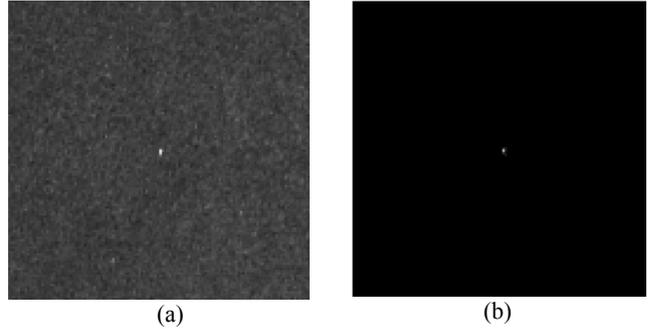
Figure 3: a) Detail crop (100x100 pixels) of ALOS PALSAR WB1 scene with one easily detectable target. The spatial resolution is 100m. b) Spatial correlation of the reconstructed horizontal, vertical and diagonal single-level sub-bands. The clutter is diminished, while the target as the only deterministic feature remains.

2.2 Spot-Detection

The most common approach to ship detection in SAR images certainly is the adaptive-threshold or CFAR detector, on which several current operational systems rely. These algorithms are designed to find pixels of unusual brightness by comparing the encountered intensity with the statistical properties of the local background, obtaining a significance value S for each pixel or pixel cluster analyzed (Crisp 2004 p.27ff).

To accurately set the threshold and to achieve a constant false alarm rate, these systems in fact have to know the background statistics or apply local distribution fits. In practice it has turned out that it is a more feasible approach is to model the background distribution parametrically, adjusting it with local data samples.
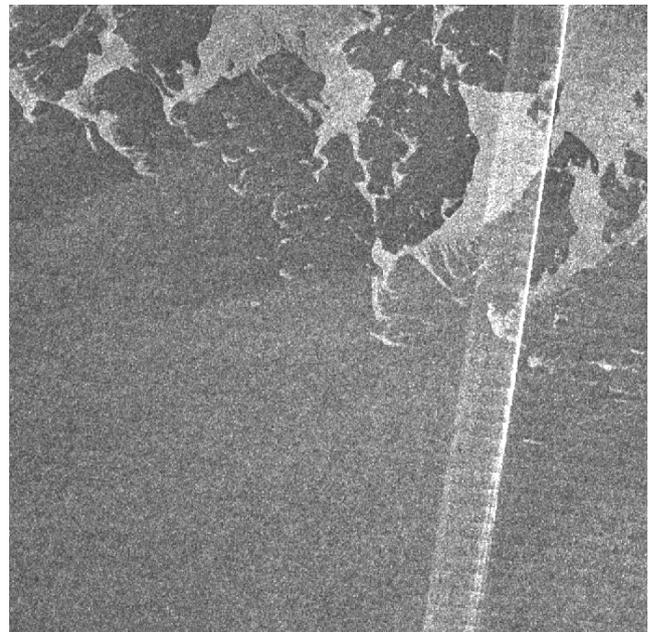
Figure 4: Radarsat1 SWB image fragment of a sea-ice border northern of Dutch Harbor, Alaska. The scene shows a predominant beam seam and a field of scattered ice, characteristic for the border region of an ice shelf.

For this, the most established setup consists of three concentric windows (Figure 5). A target window extracts the current pixel(s) under test, and a bigger background window defines the area that is considered to contain the associated background information. Between these windows, a guard region is assigned to prevent the target from leaking into the background area, which would distort the estimation of the underlying statistic properties. An important limitation to such a system is the window size, which can hardly be varied for the reason of computational efficiency, but has to satisfy all expected targets sizes (Crisp 2004 ibid.).
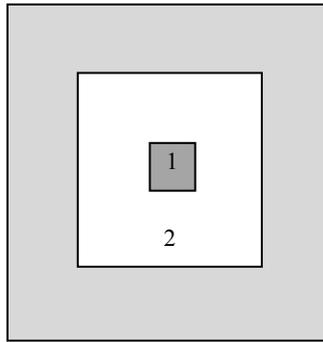
Figure 5: Typical setup for an adaptive-threshold detector: a target area consisting of one or more pixels (1) is separated by a guard area (2) from the background ring (light gray).
The comparison of the target area intensity with the data cropped by the background area supplies a significance degree on which the detection decision is based.

Another general problem of the CFAR method as such is that the characterization of the background distribution despite extensive research is always a factor of uncertainty. It displays large fluctuations within even the same image, caused by imhomogeneities such as sea-clutter, beam-seams, or further non-stationary effects (Figure 6). It is worth mentioning that the term CFAR for this kind of detector can be misleading, since the constancy of the false-alarm rate is entirely dependent on the correct statistical estimation of the background distribution of the data. Furthermore, the prototype system described in this paper applies its adaptive threshold detection on an extremely limited non-random *selection of samples* from the original data, hence violating the statistical frame for which a calculated constant false alarm rate is meaningful. Therefore, without loss of generality the application of CFAR-like techniques in the following will be referred to as "adaptive-threshold (AT)" rather than "CFAR" detector.

2.3   Land- and watermask classification

While the aforementioned detection schemes perform reasonably well in open water areas, their quality decreases significantly when applied to coastal areas, within river deltas. Image structure along rugged coastlines, on undulated land surfaces, or urban areas may appear as small, ship-like blobs when observed only locally, i.e. in a small neighbourhood like shown in Figure 5. Hence the above procedures tend to over-segment coastal regions and land surfaces.
To overcome this problem, we perform an image segmentation of land and water areas in the detection scheme, which is based on context information derived from the SAR image itself. The most important step of this procedure is the correct selection of representative and distinctive features:
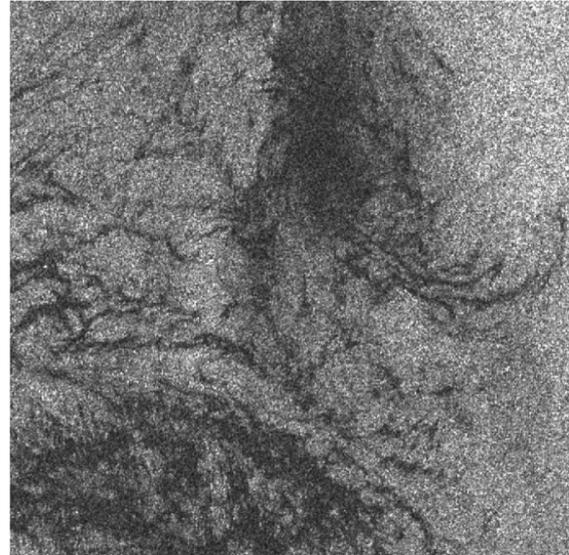
Figure 6: Heterogeneous ocean surface in Radarsat1 SNB mode image (resolution 25m).

- **Brightness**
  For calm weather conditions, the scattering off water surfaces is dominated by specular reflection, directing most of the incident energy away from the sensor. Thus, for these conditions, radar brightness estimated in a local neighbourhood (see e.g. Figure 4) can be used to identify water areas in SAR scenes. We utilize a rank-filter to estimate the local brightness robustly even in the case of heavy speckle-noise.
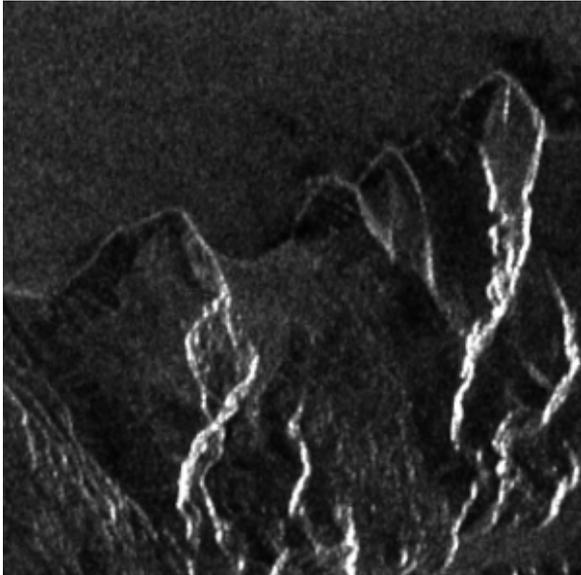
- **Grayvalue variance**
  The same physical argument as above holds also for the grayvalue variance. According to the speckle statistic, the tail of the probability density function of image intensity scales with the radar cross section. Hence the grayvalue variance, again estimated in a local neighbourhood, is an additional feature to separate land from water.
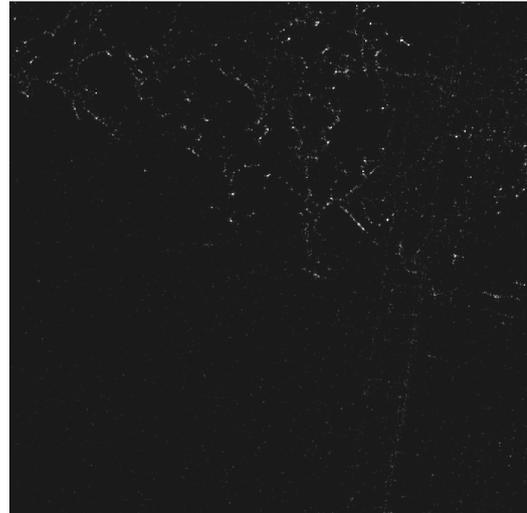
- **Grayvalue structure**
  As can be seen from Figure 6, also water regions can be characterized by strong backscatter and heavy noise. Such situations appear in dependence of the wave characteristics of the ocean, especially in case of the well-known Bragg-scattering effect. The first and second order statistical moments are thus no more distinctive enough for land and water classification. However, the local grayvalue structure is still different. While the grayvalues of ocean surface are more or less randomly organized, significant edge or line structures can be identified for land regions (see Figure 7). To capture these features, gradient orientation histograms weighted by the edge-amplitude are calculated in local neighbourhoods. In these, ocean surface appears as uniform distribution of the orientations while land is characterized by significant modes of the orientation histogram.

The mentioned features are fed into a supervised classification, whose output (after some standard morphological cleaning) is regions for ocean and land.

Figure 7: Similar grayvalue moments but different grayvalue structure of land and ocean surface in Radarsat1 SWB mode image (resolution 50m).

3. IMPLEMENTATION ISSUES

3.1 Wavelet pre-screening

In order to avoid the aforementioned vulnerabilities of a detector exclusively based on wavelet transformed data, the wavelet analysis as described in the previous section is only the first stage of the presented prototype. Its foremost task is to extract the entirety of *possible* targets. This step will sometimes be referred to as "prescreening" in the following. The existence of a consecutive analysis stage allows the discrimination thresholds for the cumulated sub-band image to be set rather low, hence not eliminating targets with a mediocre response in the wavelet domain.

The prescreening system presented in this paper uses the two-dimensional stationary wavelet transform as described in (Nason et al. 1997). After the decomposition the sub-bands are filtered to concentrate the signal energy spatially, leading to more distinctive responses in the cumulated result (Figure 8). For optimal performance, this result is formed out of several decomposition levels, hence encompassing the desired size range of ships. This also prevents oblong or elliptical targets, which show in imagery of 50m or higher resolution, from being unnecessarily suppressed.

3.2 Ship Candidate selection

The cumulative wavelet image assigns large pixel values to ship like features and suppresses objects that don't correspond to the used object model. In this dataset, an object discrimination has to be performed that identifies the ship candidates that will progress to the next processing steps. At this stage, a preliminary general threshold is applied to the cumulated wavelet image, cleaning out the heavily suppressed clutter. Although a relaxed threshold setting is used to guarantee a low number of missed hits, it still successfully isolates compact pixel groups in a mostly zero matrix. These groups are henceforth considered objects, each representing a potential ship.

Figure 8: Cumulated wavelet result of the same image. The manifold sea-ice structures create numerous rather strong responses in the wavelet domain, likely to cause a high number of false alarms. Here the shape-selectivity of the wavelet approach is noticeable: even though the beam seam is the brightest structure, it is suppressed in the wavelet result

In the process of labeling the objects, the object size in pixels is determined. Careful comparison of the object signatures in the cumulated wavelet domain and the original image show that when treated correctly, the estimated sizes match very closely. This information allows several considerations:

Knowledge of the estimated object size considerably contributes to the performance of an adaptive-threshold detector. The crucial sample window sizes can be automatically adjusted to each object, significantly decreasing the probability of a false decision.

Since the underlying physical size of each pixel as well as the pixel spacing are known, a relation between the apparent object size and the corresponding physical object can be easily formulated. By this, size limits for ship candidates can be determined and utilized to purge unwanted objects from the list. To this end, ellipses are fitted to the boundaries of candidate regions and thresholded based on geometrical parameters such as absolute and relative length of major and minor axis, area, and eccentricity. In practice, this test can reduce the number of objects under test by more than 90%.

Since the subsequent processing steps are no longer dependent on the total image size in pixels, but on the number of objects to be analyzed, early identification of the majority of invalid objects results in a considerable speed gain. At the same time, this allows for higher complexity in the following parts of the object analysis, such as advances shape and area examination for the remaining objects.

3.3 Adaptive-threshold detector

For the sake of demonstrating the general benefit of an AT detector following a wavelet pre-screener, the rather uncomplicated and ubiquitous two-parameter detector has been chosen. This detector obtains the target intensity x_t and assumes a Gaussian background distribution, which can be fully parameterized by the local estimation of the first two moments μ_b and σ_b. The obtained significance value S then is compared with an empirically determined threshold value t:

$$S = \frac{(x_t - \mu_b)}{\sigma_b} > t \quad \Rightarrow \text{valid target}$$

Even though the correctness of the Gaussian decision is based on the central-limit theorem and therefore only of asymptotic nature with respect to the number of looks, it still proves to be a useful and simple method to distinguish significant features from their surroundings.

The implementation of the AT detection step for the prototype system is different from conventional setups in many ways, since the detector has been developed to benefit from the previous object discrimination. Nonetheless, the currently included two-parameter detector can easily be switched with a more sophisticated statistical detection mechanism.

Knowledge about the spatial extent of the pixel group under test allows assigning a dynamically tailored target window *for every object*. This is of great importance for the confidence of the adaptive threshold decision, since the pixel group can now be tested as a whole, following the basic principle of the moving-average (MA) CFAR detector. Hence, the target significance term

$$S = \frac{(x_t - \mu_b)}{\sigma_b}$$

can be extended for averaging over the list of target pixels, and has to take into account the number *n* of grouped pixels within a target. This is achieved by weighting the background standard deviation with the multi-looking factor $1/\sqrt{n}$, yielding

$$S_n = \frac{(x_t - \mu_b)}{\sigma_b / \sqrt{n}}$$

This procedure greatly improves the robustness of the AT detection performance especially for larger targets in inhomogeneous areas of the image.

3.4 Selection criteria

To keep the effects of the different methods distinguishable and to maximize the transparency of the detection decision for development and assessment the final selection criterion merely combines the magnitude of the object's response in the wavelet domain M_{HVD} and the significance return value S from the AT detector:

$$S \cdot M_{HVD} > \text{threshold} \quad \Rightarrow \text{valid ship}$$

Although this criterion is rather simplistic, it still maintains the influence of both processing steps on the final detection decision.

Improved selection criteria could of course better exploit the wealth of information available in the detection system. These criteria will include more complex object parameters, such as homogeneity of the object or more advanced shape criteria, to support and strengthen the decision making process. More sophisticated evaluation frameworks, e.g. the application of a fuzzy logic mechanism, will be applied to facilitate the incorporation of heterogeneous information types and to fuse different information channels.

3.5 Land and water classification

The Image parameters described in Section 2.3 are used to train a supervised classification scheme for land masking, which currently is set up as a maximum a-posteriori classifier. The benefit of using other classifiers, e.g. Support Vector Machines, has not been investigated and will be subject of future research. Figure 9 illustrates a typical example of a classification result for a segment of an ALOS-PALSAR FBS scene.
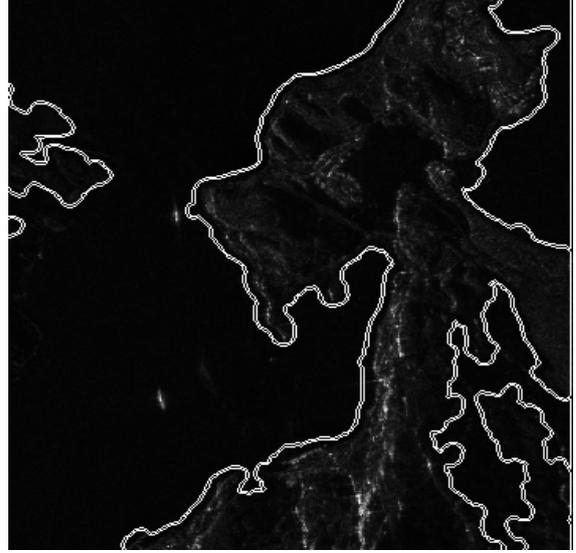
Figure 9: Automatically delineated shoreline by land and water classification for an ALOS PALSAR FBS scene.

4. RESULTS AND DISCUSSION

A set of images has been selected to conduct a first assessment of the system. As high spatial coverage is essential for an operational ship detection system, the testing objective was to cover the typical resolution range of currently available SAR imagery with swath widths of 60km or more. Therefore, the experiments include resolutions between approx. 6.5m for ALOS-PALSAR data and 50-100m for Radarsat-1 SWB. Special focus was put on evaluating the detection performance with respect to

- spatial resolution of different SAR products
- image artifacts (e.g. nadir echoes or beam seams)
- heterogeneous image content (e.g. sea-ice border areas).

After the application of the prototype algorithm, the images also were inspected manually for potential false negative detections. Furthermore, the intermediate products of the detection process were analyzed and compared. A reference ship mask was generated manually from the available SAR imagery and used for evaluating the performance of the different methods applied. For the numerous evaluated test scenes, typical performance characteristics are:

Completeness (71.8%): 23 of 32 ships detected
Correctness (80.0%): 29 total with 23 correct detection

The system shows great robustness with respect to sea-ice features, a consistent problem for ship detection algorithms. This robustness is especially significant in ALOS-PALSAR data, which show lower clutter magnitude compared to the tested Radarsat-1 imagery. Especially beam seams or nadir echoes, whose sometimes spot-like pattern resemble the signature expected from ships, generate strong responses and false detections in the wavelet based detectors (Figures 10a, b). The example in Figure 10 shows that the subsequent object-AT detector is able to remove the spots caused by the beam seams by taking their spatial extent and their statistical significance into account (Figure 10c). Similar behavior can be observed for high-clutter features, as demonstrated by the example case in Figure 11, a Radarsat1 SNB-mode scene. Due to the comparatively high resolution of 25m and the lack of different

looks, the clutter manifests in a coarse pattern. This creates blob-like features which do not become sufficiently suppressed by the wavelet pre-screener, what again makes the additional information provided by the object/AT analysis necessary for a confident decision In an entire WB1 image (100m resolution) of which more than 50% were covered with sea-ice, the system generated a mere of six non-land false alarms. All of these were not on the ice border, but on the main ice body itself.

Although the results of the system are very encouraging, there are numerous improvements possible. Especially under the light of the new high-resolution spaceborne SAR data, it will be possible to include more geometrical object-specific features of ships as well as cues stemming from the motion of a ship. For instance, first results of tracking using sub-apertures have been very successful, same as extraction ship wakes typically seen in the context of ship motion.

REFERENCES

Chan, Andrew K. (1993), "Wavelets for sensing technologies", Artech House

Crisp, D.J. (2004), „The state-of-the-art in ship detection in synthetic aperture radar imagery", DSTO Information Sciences Laboratory, DSTO–RR–0272

DeGrandi, Gianfranco D.; Jong-Sen Lee, Dale L.Schuler (2007), "Target detection and texture segmentation in polarimetric SAR images using a wavelet frame: Theoretical aspects", IEEE Transactions on geoscience and remote sensing, Vol. 45, No. 11

Nason, G.P.; B.W. Silverman (1997), "The stationary wavelet transform and some statistical applications", Lecture Notes in Statistics, 103, pp. 281-299

Tello, Marivi; Carlos López-Martínez, Jordi J. Mallorqui (2005), „A novel algorithm for ship detection in SAR imagery based on the wavelet transform", IEEE Geoscience and Remote Sensing Letters, Vol. 2, No. 2

Tello, Maríví; Carlos López-Martínez, Jordi J. Mallorqui (2006), „Automatic vessel monitoring with single and multidimensional SAR images in the wavelet domain", ISPRS Journal of Photogrammetry & Remote Sensing 61, pp. 260–278
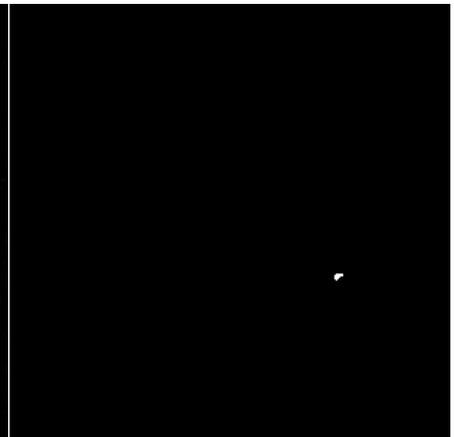
Wackerman, C.C.; K.S. Friedman, W.G. Pichel, P. Clemente-Colon, and X. Li. (2001), „Automatic detection of ships in RADARSAT-1 SAR imagery", Canadian Journal of Remote Sensing, 27(5):568–577, 2001.
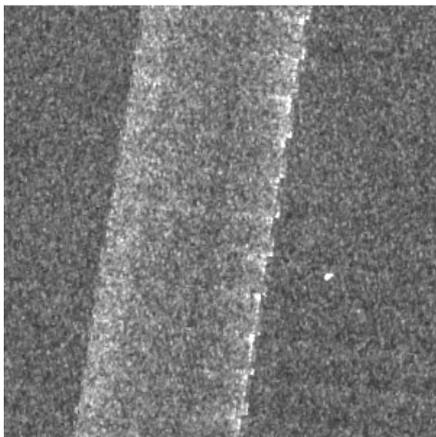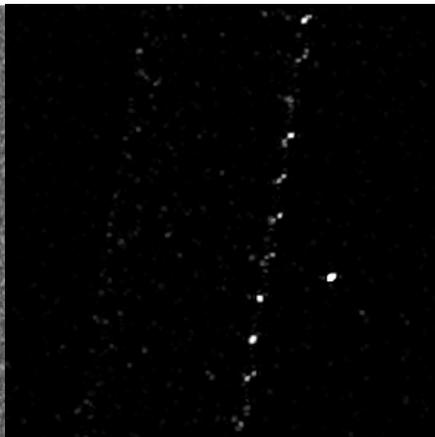
| Figure 10a | Figure 10b | Figure 10c |

Figure 10a: Single ship close to strong beam seam in Radarsat1 SWB mode image with resolution of 50m.
Figure 10b: Wavelet pre-screening result of the scene, no threshold applied. The bright spots at the edge of the seam are strong enough to generate responses comparable to the real ship both in amplitude and shape.
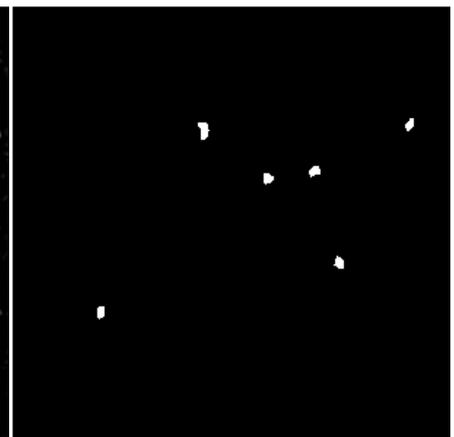**Figure 10c:** Remaining valid objects after the object-discrimination and AT detection. All beam seam responses have been correctly classified as invalid targets.
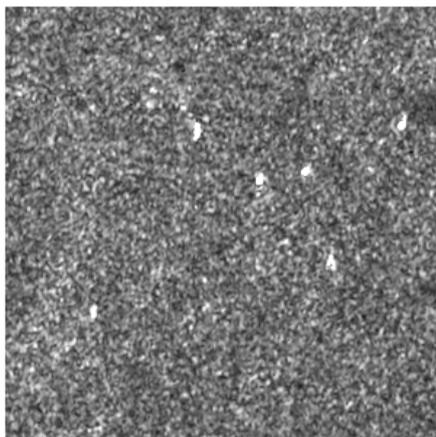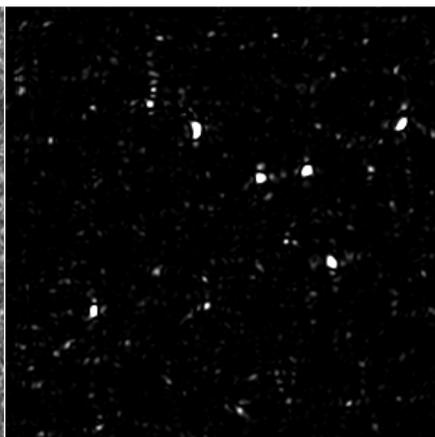
| Figure 11a | Figure 11b | Figure 11c |

Figure 11a: Radarsat1 SNB mode image, 25m resolution. A group of ships in a coarse-clutter region.
Figure 11b: Output of the pre-screener. The ships are enhanced, but also numerous smaller objects emerge from the clutter.
Figure 11c: After the AT-stage has been applied to all objects, the final detection decision only only the ships remain in the binarized image.