Abstract—Oil-spill and furtive vessel detection are among the most required real-time monitoring tasks for environmental and government authorities. In Argentina, whose sea jurisdiction is huge (about one million square km), human assisted monitoring is unfeasible, and therefore remote sensing technologies, for instance satellite imagery, is the only operative and economically feasible solution. Recent associated efforts between the National Aerospace Activities Council of Argentina (CONAE) and its Italian counterpart to operate together the SAOCOM and COSMO satellite constellations enabled the possibility of developing a system capable of providing accurate, real-time space-born SAR imagery. In this work, we describe the theoretical foundations and implementation details of AMTO (Ambiente de Monitoreo Terrestre-Oceánico), a system specifically designed to take advantage of the SAR imagery delivered by those two satellite constellations, to provide real-time detection of vessels and oil spills.

Keywords—SAR IMAGERY - OIL SPILL DETECTION - VESSEL DETECTION

1 INTRODUCTION

Currently, oil spills on the sea surface constitute one of the major sources of marine environmental damage. According to [16], 48% of oil pollution in the oceans is comprised of fuels and 29% of crude oil. In addition, recent observations correlate oil spills with the most frequent shipping routes [12]. Together, this two facts tend to confirm the observation made by [15] that deliberate oil spills are considerably more frequent than the ones produced by reported ship accidents. This implies that as of today, intentionally produced regular oil spills constitute a more important threat to marine ecosystems than larger but less frequent spills produced by accidents.

In Argentina, the impact oil spills is presently unknown, but the main environmental impact is assumed to be huge given the recent spills that affected different species of seabirds, including penguin colonies (see for instance www.ens-newswire.com/ens/jan2008/2008-01-02-03.asp). The Argentine ocean littoral is too large to be surveilled by human-assisted means with adequate revisiting times and geographical precision.

Another related monitoring requirement is vessel detection. In Argentina’s case, vessel tracking is not only important to correlate with oil spills (and to spot the responsible), but also for the early detection and action upon the frequent violation of the sea jurisdiction by furtive fishing factory boats. In both requirements, remote sensing appears to be the only technologically feasible solution.

There are several sources of satellite imagery, including multispectral and hyperspectral imagery provided by missions like LANDSAT, MODIS and Ikonos, among many others. Even though multispectral or hyperspectral imagery may be suited for vessel and oil detection, there are several associated monitoring features in those sensing modalities that make them inadequate for a regular monitoring service. Among them, we may mention that passive sensors working in the visible and NIR regions of the spectrum are susceptible to fail due to climatic and daylight conditions. Also, in most cases the spatial and temporal resolution (i.e., the pixel size and the revisit rate) are insufficient for the required processing. The LANDSAT revisit rate is 14 days, and the MODIS resolution is 250m at best, making both missions unfit for the purpose at hand. Ikonos imagery has both a finer resolution (about 1m) and a daily revisit rate, but the fact that the service is provided by private companies makes it extremely expensive for the wide required areas to survey.

Synthetic-Aperture RADAR (SAR), on the other hand is an active, coherent monitoring technique that takes advantage of the relative motion between an antenna and its target region. The use of millimeter wavelengths has the advantage of being able to produce consistent imagery notwithstanding daylight and weather conditions. SAR satellite imagery, then, is among the best technologies for oil spill monitoring due to its wide area coverage, weather and daylight conditions obliviousness, and the fact that most space agencies nowadays operate constellations of satellites that are able to produce high revisiting rates like
the required for a good temporal precision in real-time monitoring.

This increasing amount of available SAR imagery implies a growing load for the operators at analysis centers. Also, even extensively trained personnel are prone to different, incongruent confidence levels in their assessment, and are also subject to errors due to distractions and fatigue [5]. Furthermore, in usual satellite SAR imagery the pixel resolution is not fine enough for detecting small targets, and then small-to-medium size vessels are hardly distinguishable from the speckling present in the images. Therefore, supervised or semi-supervised algorithms for detection can help in screening the images, and triggering alarms that capture the operators’ attention only for non-trivial situations.

Research on unsupervised vessel and oil spill detection in SAR imagery has been a relatively active field in the last two decades (for instance, see [1] for a brief description of current research with worldwide available SAR satellite imagery). In this paper we present a system for real-time vessel and oil spill detection using distribution parameter estimation as a local feature descriptor in the imagery available through the Argentine SAOCOM and Italian COSMO constellations. These results are incorporated in a software product that is being deployed at the National Aerospace Activities Council of Argentina (CONAE).

2 MATERIALS AND METHODS

The final development will be operative together with the SAOCOM Argentine SAR satellite constellation. These satellites are provided with an L-band full polarimetric SAR sensor. The constellation will also be operative together with the Italian COSMO constellation, provided with similar sensors but in the X-band. In addition, under specific agreements, the Argentine Space Agency also exchanges imagery from the Canadian RADARSAT-1 satellite, whose sensor provides C-band unpolarized higher resolution imagery. For this reason, the development and testing of our processing system will be targeted specifically at imagery of these two constellations.

2.1 SAR Imaging

Space-borne SAR imaging is usually implemented by mounting a single beam-forming antenna on the satellite, from which the Earth is repeatedly illuminated with pulses, usually on the millimeter wavelength range. The obtained coherent signal variations may be processed to obtain a finer spatial resolution than the one possible with conventional beam-scanning means at the used wavelengths [18]. The backscatter received at different antenna positions is further processed (usually in batch, at a ground station) to produce coherent imagery of the target region. Also, specifically designed antennas are able to discriminate the polarization of the backscatter, and the direct and cross-directional backscatter magnitudes and phases reveal properties about the anisotropy of the surface, which provide useful information in image understanding [14].

Since SAR imaging is the result of evaluating the backscatter of a coherent electromagnetic source, it is subject to the speckle noise effect, which is produced when several different backscatterers partially reflect the same electromagnetic pulse within the range of the same pixel, therefore breaking the initial coherence of the source. Speckle noise is inherently a multiplicative process [8], and therefore all usual image processing techniques fail in segmenting and recognizing specific targets in the images. Speckle reduction techniques, like Wiener or Lee filters [9] or wavelets [2], make images more intelligible at large scales, but tend to obliterate or diffuse small details in the small resolution scale, and for this reason the use of these filters is not adequate for vessel and small oil spills detection.

Like any multiplicative process, SAR imagery can be modeled using multiplicative density functions. The usual Gamma distribution model appears to work merely roughly, since it can only model the scale and shape parameters of the backscatterer, but antenna parameters cannot be modeled. Therefore, the K distribution was developed as a more specific density function for SAR imagery that also takes into account the antenna parameters [17]. Basically, the K distribution is a compound probability model that can be regarded as the product of two Gamma density distributions, the first modeling the backscatterer with a scale parameter $\mu$ and shape parameter $\nu$, and the second modeling the antenna width, with a scale parameter set to one, and a shape parameter $L$ ($0 \leq x, \nu, L \leq 1$):

$$f_{\nu}^L(x) = 2 x \left(\frac{L \nu x}{\mu}\right)^{\frac{\nu-2}{2}} \frac{1}{\Gamma(L) \Gamma(\nu)} K_{\nu-L} \left(2 \left(\frac{L \nu x}{\mu}\right)\right),$$

where $K_{\nu-L}$ is the second kind Bessel function and $\Gamma$ is the Gamma function. In multilook SAR images it is possible to reduce the relationship between standard deviation and mean of the image due to the speckle [11], relating the amount $k$ of looks to the antenna shape parameter. In Fig. 1 we show part of a RADARSAT-1 image taken at the San Jorge gulf (45°59 S, 66° 41 W), and the histogram of the normalized modulus. Several vessels are present, but are unnoticeable since the visual clutter of the speckle makes them difficult to spot.

2.2 Parameter Estimation

Since most of the usual image processing techniques are bound to fail at detecting vessels and oil spills in SAR imagery, we applied a different technique based on local estimation of the parameters of the underlying distribution function. The mean and variance of the K distribution of a given pixel $X = x_1, \ldots, x_k$ is:

$$E(X) = \mu,$$
$$\text{var}(X) = \mu^2 \frac{\nu + L + 1}{L \nu},$$

where $E(X)$ is the expectation of $X$, and $\text{var}(X)$ is the variance of $X$. The parameters $\mu$, $\nu$, and $L$ are estimated using maximum likelihood estimation. The likelihood function is given by:

$$L(\mu, \nu, L | X) = \prod_{i=1}^{k} f_{\nu}^L(x_i),$$

where $f_{\nu}^L(x)$ is the density function defined in (1). The log-likelihood function is:

$$\ln L(\mu, \nu, L | X) = \sum_{i=1}^{k} \ln f_{\nu}^L(x_i).$$

Taking the partial derivatives with respect to $\mu$, $\nu$, and $L$, and setting them to zero, we obtain the following equations:

$$\frac{\partial \ln L}{\partial \mu} = 0,$$
$$\frac{\partial \ln L}{\partial \nu} = 0,$$
$$\frac{\partial \ln L}{\partial L} = 0.$$
where $k$ is the amount of looks over the same pixel. The likelihood of $X$ is then

$$
\hat{f}(X) = \prod_{i=1}^{k} f_{\nu}^{\mu}(x_i). \quad (4)
$$

Given a fixed amount of looks $k$, and also a fixed antenna shape parameter $L$, then it is possible to consider that the K distribution with parameters $\mu, \nu, 1, L$ will behave closely to a Gamma distribution. Since evaluating the true parameters of this new Gamma distribution is an underdetermined problem, we expect to find the parameters of a putative Gamma distribution at a given pixel $p$, that can be evaluated from populations of a small window of contiguous pixels centered at $p$. Our purpose is to use the shape and scale parameters of these local distributions as features for image segmentation, i.e., classifying the image in this local 2D feature space, where the features are the parameters of the local Gamma distribution around a given pixel (methods for parameter estimation are well known for the Gamma distribution [3], while in the case of the K distribution much less is known).

For simplicity, we will refer to the shape and scale parameters of this new (putative) Gamma distribution as $\mu, \nu$, even though they are not necessarily the same to the original $\mu, \nu$ of the original K distribution. Given a pixel population $P = p_1, \ldots, p_n$, the likelihood of $P$ can be specified as:

$$
\hat{f}(X) = \prod_{i=1}^{n} \frac{1}{\Gamma(\nu)\mu^\nu} x_i^{\nu-1} e^{-\frac{x_i}{\mu}}, \quad (5)
$$

which in turn allows to evaluate the $\mu$ parameter:

$$
\hat{\mu} = \frac{1}{n\nu} \sum_{i=1}^{n} x_i.
$$

On the other hand, the estimation of the $\nu$ parameter has no known closed form, and therefore must be numerically evaluated. In this case the techniques proposed in [13] for estimating $\nu$ can be applied, which consist on regarding the logarithm of Eq. 4 and approximate it to a simpler function:

$$
\log (f(P)) = (\nu - 1) \sum_{i=1}^{n} \ln(x_i) - \sum_{i=1}^{N} \frac{x_i}{\mu} - n \nu \ln(\mu) - n \ln(\Gamma(\mu)) 
\sim a_0 + a_1 \nu + a_2 \nu^2.
$$

In this way, the log-likelihood can be numerically estimated with a generalized Newton’s method that converges very fast to a final $\nu$ estimated parameter. In numerical simulations of this procedure, the estimated parameter was always within 2% to the actual shape parameter. In Fig. 2 we show the histogram of the normalized modulus of the backscatter, and superimposed the resulting parameter estimation, at pixels centered in three different regions (resp. sea surface, vessel, and oil) that were selected under an expert’s supervision. In the first case, the population (i.e., the amount of pixels or ”window size”) is large enough (typically $11 \times 11$ or larger) for having a large population that makes a robust parameter estimation. In the second and third cases, smaller windows were taken (typically between $5 \times 5$ and $9 \times 9$). Larger windows would be mainly dominated by the sea-surface parameters, obliterating the features of vessels and oil spills, and making fuzzier the location of the actual features. On the other hand, smaller windows will have a pixel population not large enough for producing significant parameter estimation.

### 2.3 Classification

Since the feature space for classification (in this case, the shape vs. scale parameters of the local Gamma distributions) is not a metric space, we used the Kullback-Leibler divergence (KLD) [10] for establishing a distance notion between two distributions, given their parameters. Given distribution $G_p$ and $G_q$, with parameters (resp.) $\mu_p, \nu_p$ and $\mu_q, \nu_q$, the KLD is:

$$
\text{KLD} = (\mu_p - \mu_q)\psi(\mu_p) - \log(\Gamma(\mu_p)) + \log(\Gamma(\mu_q)) + \mu_q(\log \nu_p - \log \nu_q) + \mu_p \frac{\nu_q - \nu_p}{\nu_p},
$$

where $\psi(.)$ is the digamma function.

A large amount of samples with sea surface, vessels, and oil, were manually selected, and their respective distribution parameters were computed and averaged to find the three representative parameter pairs for the three

---

*Figure 1: Part of a RADARSAT-1 image, and the associated histogram of the normalized modulus.*
classes. These patterns exhibited a remarkable invariance in different images of the same satellite, especially in varying sea conditions, which many times triggers false alarms in commercial vessel detection software. The classification of every pixel $p$ in an image then proceeds by computing the $\mu, \nu$ of a neighborhood around $p$, and associating that pixel with the class with which the KLd is smaller. Other more sophisticated classification schemata were also tried (clustering, for instance), with similar or only negligibly better results.

3 RESULTS

The full classification workflow was implemented and tested with a series of sample images provided by the Argentine Space Agency. In every case, the main detection targets were successfully found, in several cases outperforming other commercial products used by the Agency. In Fig. 3 we show a magnification of a crop in Fig. 1, where many factory ships are found, and the result after applying our processing pipeline. The resulting image is shown using the KLd to sea surface as a pseudocolor. In Fig. 4 the segmentation of an oil spill is shown, where also the detection is robust with respect to the sample cases. For establishing comparisons with other methods in the literature, we based our analysis on confusion matrices and ROC curves.

3.1 Confusion matrices and ROC curves

The effectiveness of our system can be understood in terms of binary classification, or two-class prediction, in Fig. 5 the point corresponding with the confusion matrix given a certain threshold. We have $N = 100$ events, $N = 50$ and $N = 20$. The classifier correctly predicts 40 of the 50 positive events ($TP = 40$) and therefore has 10 misses ($FN = 10$). Also the classifier correctly predicts 30 of the negative events ($TN = 30$) and has thus 20 false alarms ($FP = 20$). In this situation, the $TPR = \frac{40}{50} = 0.8$ and the $FPR = \frac{20}{20} = 0.4$. Changing the threshold will change the $TPR$ and $FPR$. The behaviors with all possible thresholds can be represented then in a $TPR$ vs. $FPR$ space, which defines the ROC space of the classifier.

<table>
<thead>
<tr>
<th>$P$</th>
<th>$N$</th>
</tr>
</thead>
<tbody>
<tr>
<td>40 (TP)</td>
<td>20 (FP)</td>
</tr>
<tr>
<td>10 (FN)</td>
<td>30 (TN)</td>
</tr>
</tbody>
</table>

Table 1: Confusion matrix given a certain threshold.

accepting a negative event is a false positive (FP) (or false alarm, or type I error).

Binary classification can be parametrized by an acceptance threshold, in our case the KLd of a pixel distribution to the plain sea distribution. Events with an associated value above (resp. below) the threshold are considered positive (resp. negative) observations, or the other way around depending the case. More lenient thresholds have a better sensitivity, since they tend to have less misses, while stricter thresholds have higher specificity, since they reduce or eliminate false alarms.

The general goal in this two-class prediction model is to find the best possible threshold that minimizes the wrong predictions. This can be stated regarding the contingency matrix (or confusion matrix), where the set of outcomes of a given classification problem is quantified in terms of the amount of TPs, TNs, FPs, and FNs (see Table 1). In any instance of a confusion matrix for $N$ events we have that $TP+FP+TN+FN= N$, and the goal of the best possible classification can be understood as having the largest possible values along the main diagonal.

Another way to characterize the behavior of a two-class predictor is in terms of the true and false positive rate (resp. $TPR$ and $FPR$). The $TPR$ is defined as the ratio between true positives and actual positives ($TPR = \frac{TP}{TP+FP}$) i.e., the sensitivity of the classifier. The $FPR$ is defined as the ratio between false positives and actual negatives ($FPR = \frac{FP}{FP+TN}$). If the ratio between $TN$ and $N$ is the specificity of the classifier, then the $FPR$ is 1 minus the specificity.

Suppose now that we have a confusion matrix like the one in Table 1. We have $N = 100$ events, $P=50$ and $N = 50$. The classifier correctly predicts 40 of the 50 positive events ($TP = 40$) and therefore has 10 misses ($FN = 10$). Also the classifier correctly predicts 30 of the negative events ($TN = 30$) and has thus 20 false alarms ($FP = 20$). In this situation, the $TPR = \frac{40}{50} = 0.8$ and the $FPR = \frac{20}{50} = 0.4$. Changing the threshold will change the $TPR$ and $FPR$. The behaviors with all possible thresholds can be represented then in a $TPR$ vs. $FPR$ space, which defines the ROC space of the classifier.

A ROC curve represents all the possible outcomes as a function of the threshold parameter [6]. For this example, in Fig. 5 the point corresponding with the confusion matrix.
Figure 3: At the left an interest area in Fig. 1 is highlighted. The area is enlarged and shown in the center. At the right, the KLd to the sea surface distribution is shown as a pseudocolor (black meaning closer, and white meaning farther).

Figure 4: An oil spill close to the seashore, and the respective segmentation.

Figure 5: An example of a ROC curve and some landmarks in ROC space.

matrix in Table 1 is marked with an A. A perfect classifier would be like point B, i.e., TPR = 1 with FPR = 0. A random classifier will tend (for large N) to behave close to point C, i.e., TPR = 0 and FPR = 0.5. Any classifier behaving along the main diagonal is making no gain respect to a random classifier (for this reason this is called the no discrimination line).

Back to our hipotetical example, varying the threshold parameter will move the A point along the ROC curve of the classifier. A lower threshold will enhance the specificity (higher TPR) at the expense of a higher sensitivity (lower FPR), and the other way around. The quality of a classifier can be derived from the ROC curves in several ways, though no general consensus exists. The most popular quality indicator is the area below the curve (AOC) [7], larger areas implying better classifiers. Also, the TPR can be associated to the benefit of the classifier, and the FPR to the cost.

3.2 Comparison with traditional methods
As mentioned above, Wiener and Lee filtering are among the most widespread despeckling techniques used in SAR imagery. We implemented a higher order non-linear adaptive Wiener despeckling filter to assess the quality of our methodology. In Fig. 6 the same crop of a Cosmo SAR image is processed using both methods. The thresholding was adjusted to have a 85% TPR in both cases. It can be seen that the Wiener filtered image has a much higher FPR (false alarms).

For quantifying the quality of the results, we computed the ROC curves of the KLd- and Wiener-based detectors using a supervised training set of 428 vessel contacts (true positives). The KLd-based detector performed better than the Wiener-based in all conditions. An optimal threshold was found for this training set, where the achieved sensi-
tivity was 89.4% with a specificity of 96.8%. This means that almost 90% of the vessels are detected, causing less than a 4% of false alarms. For the Wiener-based detector, the optimal threshold achieves a sensitivity of 70.1%, with a specificity of 72.8%. These results are shown in Fig. 7.

3.3 Additional features

As an extra asset, the Agency considered the possibility of identifying different kind of vessels. For that goal we applied a super-resolution algorithm presented in [4] that is able to find accurate vectorizations from grayscale images. The underlying idea is to try to recover the information in the mixed pixel (in our case, partially covered by sea and partially by a ship), associating the backscatter to the neighboring classes.

In our case, for every pixel in the neighborhood of a pixel p classified as vessel, we computed the normalized relative distance in KLd among the sea surface class and the vessel class (distance 0 means ship, distance 1 means sea surface, and distances in between are roughly correlated to the percentage of sea surface in the pixel). With this information, the algorithm is able to compute super-resolution vectorizations that resemble the actual shape of the ships (see Fig. 8).

Several other issues are being considered for the final deployment of our application. One of them is that the whole system should be operative using royalty-free, open-source software, and standard, off-the-shelf hardware that requires no specific training to be used. In this way, the system may be operated freely and autonomously, and can also be easily cloned to be operative under similar conditions by other Agencies. Second, the system should be able to compute the whole pipeline on the fly as the images are downloaded from the satellite.

If the pace of images to be processed is increased, there should be easy ways to adapt the throughput of the system accordingly. For this reason we chose GPGPU technology (see http://gpgpu.org) for numerically heavy algorithms. In this way, the main server of the application can be easily strengthened by plugging in additional graphic cards, without requiring software maintenance.

The complete system was developed using Qt libraries (see http://qt-project.org) for the core and user interface, and OpenCL (www.khronos.org/opencl) for the GPGPU programming. The underlying technology works under a client-server model, allowing platform independent, multi-user connections for different responsibilities. The whole pipeline can be monitored and accessed by authorized personnel, and the status of the process and the required alarms are triggered by automated emailing.

4 CONCLUSIONS AND FUTURE WORK

We presented a software product for real-time vessel and oil spill detection using distribution parameter estimation as a local feature descriptor in the imagery available through the Argentine SAOCOM and Italian COSMO constellations. The results show better detection performance than most of the available commercial software. The system is currently being deployed at the National Aerospace Activities Council of Argentina (CONAE). Our implementation uses entirely royalty-free and open-source platforms, and off-the-shelf hardware, providing the users the possibility to migrate, enlarge, or reproduce the processing capabilities on their own.

Several ideas are currently under exploration in agreement with the Argentine Space Agency. On the detection side, the most important one is to be able to identify and track vessel routes, to cross check that information with the available from the Naval authority. Also, the study and elaboration of transport models based on sea currents and wind information would be useful for predicting the advection of oil spills. On the operative side, we are currently implementing mobile-based access to the system, in a way such that the users can access to the processing pipeline using mobile devices, and the status of the system and the required alarms are also delivered by phone.
Figure 7: ROC curves of the detectors using a supervised training set. In blue the KLD-based detector, where the optimal thresholding (marked with point A, sensitivity = 89.4% and specificity 96.8%). In red, the Wiener-based detector, where the optimal threshold is marked with point B (sensitivity 70.1%, specificity 72.8%).

Figure 8: Three different vessels, and superimposed their respective super-resolution vectorizations.
Acknowledgements
This project was developed under grant AO-SAOCOM Nr. 10, provided by the INNOVA-T Foundation. Further funding was provided by the SECyT-UNS under grant 24K/046. All the imagery in this paper was provided by the National Aerospace Activities Council (CONAE) and cannot be reproduced without written authorization. We are grateful to Ing. (Capt.) Roberto Gonzalez and his team at CONAE for their helpful and patient guide in the specification and validation of the functionality of the software product presented in this work. Finally, we are specially grateful to Mg. José Luis Pasciaroni for introducing us to the possibility of this development, and participating in its initial steps.

References


