Architecture of an operational ship detection system using high spatial resolution optical satellite imagery: Application to shrimp boats in French Guiana

Christina Corbane, Fabrice Marre, Philippe Vendeville, François Marques, Michel Petit

aESPACE Unit, Institut de Recherche pour le Développement, Montpellier, France
bESPACE Unit, IRD, French Guiana
cIFREMER, French Guiana
dNev@nthropic, French Guiana
christina.corbane@ird.fr

Abstract

Within the framework of the IBIS project, an algorithm for automatic ship detection from high spatial resolution optical imagery was developed to complement existing fishery control measures, in particular the Vessel Monitoring System. The algorithm focused on feature-based analysis of satellite imagery. Genetic algorithms and neural networks were used to deal with the feature-borne information. Based on the described approach, a first prototype was designed to classify small targets such as shrimp boats and tested on panchromatic SPOT5, 5-m resolution imagery taking into account the environmental and fishing context. The proposed algorithm provided a detection rate of 80%, proving its potential for operational use in maritime surveillance.

1. Introduction

Ship detection is a key requirement for monitoring traffic, fisheries and for associating ships with oil discharges. Automatically recognizing and extracting ship patterns from remote sensed imagery has often formed part of major research efforts in the fields of pattern recognition and computer vision. Ship pattern detection by synthetic aperture radar (SAR) has been studied extensively and most recently reviewed in [1]. Even though SAR imagery is advantageous due to its ability to scan large areas and its independence from cloud and light conditions, individual identification and classification of vessels at a higher detail level remains a difficult task. Far less research and development activity has taken place in automatic detection and classification of vessels using optical imagery than using SAR imagery [2]. This is a consequence of the novelty of the 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 since it is most suitable for ship classification and it permits the detection of wooden and fibreglass boats, which are difficult to detect with radar [3]. In this context, there is an imperative need for a system that automatically detects ship patterns from high spatial resolution imagery in an operational framework.

This paper proposes an approach for the detection and classification of ship patterns from high spatial resolution optical imagery. It was developed in the framework of the IBIS project (Implementation of a Boat Information System) to complement existing fishery control measures, in particular the Vessel Monitoring System. The algorithm focuses on feature-based analysis of satellite imagery. Genetic algorithms (GAs) and neural networks are used to deal with the feature-borne information. Based on the described approach, a first prototype was designed to classify small targets such as shrimp boats and tested on panchromatic SPOT5, 5-m resolution imagery.

2. Methodology

2.1. Method Overview

The algorithm is a three-step object detection task consisting of the following stages:
- segmentation or predetection of ship patterns,
- feature extraction,
- classification.

The segmentation aims at detecting potential ship targets. It involves a pre-processing stage for the purpose of removing noise. Then, image objects corresponding to potential ship targets are created by means of a region-growing algorithm. During the feature extraction stage, image objects are characterized by spectral, shape and textural features. The feature extraction step is concerned with finding transformations to map features to a lower dimensional space for enhanced class separability and optimized performance. Because of their robustness, GAs are considered a suitable tool to address the optimization problem [4]. The GA-driven selection procedure provides a vector of feature values corresponding to a series of feature combinations that is passed to the subsequent classification stage.

In the third and last step, artificial neural networks are used for the classification of image objects. A neural network architecture is created according to the optimal feature combinations and optimal number of hidden...
nodes. Fig. 1 shows an overview of the approach which consists of two phases: a learning phase and an operational phase. In the learning phase, genetic algorithms are used to train (evolve) a feed-forward neural network based on reference samples. An objective function is used to calculate fitness that is equal to the inverse of classification error rate. In the operational phase, the best low dimensional neural network architecture is selected as a classifier in the three-step ship detection and classification algorithm.

In what follows, the three main stages of the algorithm are described in detail.

### 2.2. Segmentation (Predetection of ship patterns)

Pre-screening of possible ship patterns is based on the contrast between sea (noise-like background) and target (a cluster of bright pixels). The contrast depends on the sea conditions, the ship’s detailed shape, and its position relative to the satellite beam. The proposed algorithm applies a simple moving window adaptive threshold to the image pixel values to discriminate bright pixels [5]. The threshold used for the detection of intensity peaks is based on the mean and the standard deviation of the sea background in the moving window.

\[ \frac{X_{ij} - \mu_{oc}}{\sigma_{oc}} \geq \text{Threshold} \]  

(1)

Noise resulting from image thresholding is removed using a morphological opening operation with a 2 × 2 pixels structural element. Indeed, isolated pixels cannot belong to a ship object, which is usually characterized by a cluster of several bright pixels. The resulting thresholded image is then segmented into coherent image objects by means of the region-growing segmentation. Shrimp boats have a characteristic texture, usually consisting of two regions of high intensity related by a region of lower intensity as shown in fig. 2. The region-growing operator allows the grouping of the regions of a ship that may be detected separately during the thresholding operation.

![Figure 2. On the left: Panchromatic SPOT5 (5m) image representing a shrimp boat in pseudo-colours. The middle image shows the two regions of the same boat detected separately during the thresholding operation. The image on the right represents the image object obtained by the region-growing operator.](image)

### 2.3. Feature extraction

According to [6], image characteristics such as shape and texture are the most useful features in visual interpretation of images acquired at a high spatial resolution. Notwithstanding their importance, it is difficult to successfully automate the recognition of ships solely based on quantified shape and texture features. Using them in combination with spectral features might result in a better discrimination of ships. Hence, based on a priori knowledge of ships’ characteristics, we screen out spectral, shape and textural features that most likely characterize ship objects in a unique way, bearing in mind that rotation-position invariance is requisite.

A ship can be generally described by the following characteristics:
- bright pixels,
- large length to width ratio,
- symmetry between its head and tail, like a long narrow ellipse,
- a regular and compact shape,
- ship wakes which have a linear texture.

Accordingly, 28 spectral, shape and texture features were derived for the image objects. Table 1 lists the 28 features calculated for an image object representing a shrimp boat (image object 5, in the example shown on fig. 3).

Ship detection can be considered as a 28-dimensional classification problem with two classes: the first class corresponds to ship objects including moving and stationary ships and the second class corresponds to all
non-ship objects such as clouds. For classification of such a high-dimensional data set, a large training sample is required. In the case of shrimp boat detection by optical remote sensing, a limited amount of ground truth information is available concerning ship position. The learning performance may not be good in small-sample conditions and with high-dimensional data. For this reason, it is desirable to reduce input dimensionality so as to improve generalization capability and to obtain a network that performs well in terms of both training and test classification accuracies [7]. This underscores the relevance of feature extraction for NNs; i.e., finding the best combination of features in a lower dimensional space that does not lead to a significant decrease in the overall classification accuracy. One way to deal with dimensionality reduction is to use a GA.

A GA is inspired by biological evolution, and is widely believed to be an effective global optimization algorithm. A genetic algorithm consists of a population of genetic strings, referred to as chromosomes, which are evaluated using a fitness function. Chromosomes consist of variables or genes. The fittest chromosomes are then regenerated at the expense of the others. Furthermore, genetic operations such as crossover and mutation are defined. The mutation operator changes individual elements of a chromosome, the crossover operation interchanges parts between strings. The combination of these operations is then repeated during several generations. The intrinsic parallelism of a genetic algorithm, i.e., the ability to manipulate large numbers of chromosomes in parallel, and the crossover operation whereby good portions of different strings are combined, both make the technique a very effective optimization technique [8]. The usefulness of GAs in pattern recognition and image processing has been demonstrated [9]. Our approach consists of using a GA to train neural networks by evolving learning parameters and input features [10]. The problem is coded in a binary chromosome with a length of 28 bits (one for each object feature), where the genes have values of 0 or 1. Starting with an initial population of 100 individuals, the selection process selects the healthier ones, directed by the survival-of-the-fittest concept of natural genetic systems. Fitness computation is based on an objective function that is the inverse of classification error rate. For a judgement on their fitness to be made, the individuals have to be decoded to serve as inputs for the NN classifier. The back-propagation algorithm is used for training the network with a learning rate = 0.25 and a momentum = 0.10. A tournament selection procedure and uniform crossover are adopted to select the new population. The probability of crossover is set to 0.5, while the mutation probability is set to 0.05. Within each successive generation, the individuals yielding the highest fitness value, corresponding to the lower classification error rate as measured on a validation set, are enriched in number. The evolutionary process for network refinement is terminated when the number of generations reaches 28. All final offspring individuals thus represent a combination of spectral, shape and texture features that should result in a high classification accuracy. Finally, the best performing low dimensional NN architectures are selected for subsequent classification.

<table>
<thead>
<tr>
<th>Number of pixels</th>
<th>39</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>189.41</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>57.05</td>
</tr>
<tr>
<td>Sum of squares</td>
<td>1,52E+06</td>
</tr>
<tr>
<td>Mn</td>
<td>95</td>
</tr>
<tr>
<td>Max</td>
<td>255</td>
</tr>
<tr>
<td>Variation</td>
<td>3254.30</td>
</tr>
<tr>
<td>Asymmetry coefficient</td>
<td>0.49</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>0.42</td>
</tr>
<tr>
<td>Perimeter</td>
<td>23.90</td>
</tr>
<tr>
<td>Area</td>
<td>39.00</td>
</tr>
<tr>
<td>Compactness</td>
<td>0.86</td>
</tr>
<tr>
<td>Elongation</td>
<td>2.44</td>
</tr>
<tr>
<td>M1 *</td>
<td>0.23</td>
</tr>
<tr>
<td>M2</td>
<td>0.03</td>
</tr>
<tr>
<td>M3</td>
<td>6.59E+14</td>
</tr>
<tr>
<td>M4</td>
<td>7.33E+13</td>
</tr>
<tr>
<td>Major axe</td>
<td>11.14</td>
</tr>
<tr>
<td>Minor axe</td>
<td>4.54</td>
</tr>
<tr>
<td>Ratio Major/Minor</td>
<td>2.46</td>
</tr>
<tr>
<td>Eccentricity</td>
<td>357.84</td>
</tr>
<tr>
<td>GLCM* mean</td>
<td>150.00</td>
</tr>
<tr>
<td>GLCM variance</td>
<td>1.08E+06</td>
</tr>
<tr>
<td>GLCM uniformity</td>
<td>404.00</td>
</tr>
<tr>
<td>GLCM inertia</td>
<td>4.97E+05</td>
</tr>
<tr>
<td>GLCM correlation</td>
<td>5.01E+05</td>
</tr>
<tr>
<td>GLCM entropy</td>
<td>99.55</td>
</tr>
<tr>
<td>GLCM homogeneity</td>
<td>22.00</td>
</tr>
</tbody>
</table>

Table 1: List of spectral, shape and texture features calculated for image objects. An example of feature values is provided for image object 5 of figure 3 (*M = momentum of inertia, *GLCM= Gray Level Co-occurrence Matrix)
2.4. NN image classification

The single hidden layer NN used for the classification is the result of a GA selection procedure employing mutation, different initial weight conditions and uniform crossover. Since it consequently represents the fittest and best performing individuals, its expected error rate is low.

Fig. 4 illustrates the framework of evolving a three-layered neural network using a GA. In the decoding procedure, all the selected object features valued by 1 are represented by a fully connected input neuron. All other neurons, valued by 0 correspond to non-selected features and are then disconnected. There are two nodes in the output layer, which account for the two classes into which image objects have to be classified. As for the number of hidden nodes in the hidden layer, it is evolved using the GA and follows the rule: number of hidden nodes = number of selected object features/2.

Figure 4. Framework of evolving neural network with a GA (modified from Van Coillie et al., 2007).

3. Method implementation

In this section, we present some experimental results obtained from the application of the proposed algorithm to the detection of shrimp boats. To illustrate the methodology, 6 SPOT5 images, with a high resolution panchromatic band (5m) were acquired over the Exclusive Economic Zone of French Guiana. The following picture is an example of a shrimp boat targeted in this application. As already mentioned, the methodology was applied in two distinct phases: a learning phase and an operational phase.

Figure 5: Example of a shrimp boat in the French Guiana area (ranging from 20 to 25 m in length).

3.1 Learning phase

A training set consisting of 200 sample objects, among which were 68 represent shrimp boats, was used during the learning phase. This training set was obtained from 5 SPOT5 images. The remaining image, acquired on 13 August 2003, was used for the evaluation of the algorithm’s performance. The network was trained for up to 4000 epochs with the GA. The output of the optimization procedure is represented in fig. 6.

Figure 6: The best fitness corresponding to each generation during the evolutionary training of the NN.

On this figure, we can see the variation of best fitness represented by the best chromosome with the number of generations of the GA. We notice that the best chromosome is located at the 22nd generation with a maximum fitness of 0.56.

Hence we can determine what features played a significant role for the classification, and what features were useless for, or even disturbed, the NN classifier. Among the 28 initial features calculated for the image objects, only 8 features were extracted: Number of pixels, Mean, Standard deviation, Minimum, Maximum, Variance, Ratio Major/Minor and Texture uniformity. Hence, the optimal NN architecture...
consisted of 4 nodes in the hidden layer as determined by the GA. An average value of 0.1377 (95% confidence interval: 0.0854; 0.1899) was obtained for the generalization error of the optimal NN, estimated by means of 10-fold cross-validation.

The effect of the GA-driven feature selection on the detection of shrimp boats was evaluated in the operational phase.

3.2 Operational phase

The image acquired on 13 August 2003 (fig. 7) was used to evaluate the algorithm’s performance in the operational phase.

When applying the algorithm for the entire scene, a land mask is needed so not to mistake land for ship objects. A global coastline database with a high accuracy was therefore a necessary element of the operational system. Once the land mask was applied, the image was submitted to the three processing stages that constitute the proposed algorithm. In a network population of 100 individuals, 8 fully connected neurons resulting from feature extraction using the GA were used for the classification.

The results for shrimp boat detection on SPOT5 imagery using the developed algorithm are represented in fig. 8. 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 Identification System (AIS) 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.

Performance was measured by detection rate (DR) and false alarm rate (FAR). DR is the number of shrimp boats correctly detected as a percentage of the total number of real shrimp boats and FAR is the number of shrimp boats incorrectly reported as a percentage of total number of real shrimp boats. In this experiment, it was found that the 16 shrimp boats detected by the system perfectly matched the operator-reported ships’ positions. Hence, FAR for class ‘ship’ was equal to 0. According to the operators’ report, a total of 31 ships were identified in the entire scene. Among them, 10 ships were less than 14 m long, and there was one moored ship. This means that there were possibly 20 real shrimp boats. Detection rate as referenced in this way to the operators’ report was thus found to be equal to 80%, which is quite high.

The output of the system is a list of detected ships’ positions and ancillary information related to the ships’ lengths and headings (for moving targets). In its current state, the system does not allow speed extraction; instead, the detected ships are categorized into two categories: ‘in motion’ or ‘static’.

Of particular interest are those detected shrimp boats for which no corresponding VMS (Vessel Monitoring System) is reported, highlighting potential unreported fishing activity. Close inspection of these targets by patrol aircraft, for example, may be required in some cases. Consequently time delay becomes a requirement for an operational system for the automatic detection of ships – the total delay should be below 1 hour to allow for meaningful follow-up action [12]. Therefore another aspect taken into consideration in performance evaluation was the timeliness. When applied on an entire SPOT5 image, the system allowed delivery of end results within 1 hour from image acquisition, thus proving its suitability for near real-time monitoring of fishing activities.

4. Discussion

The implementation of the proposed algorithm on a 5-m resolution panchromatic image showed that the system can provide reliable detection of shrimp boats with minimal operator intervention and practically without any false alarms. Failure to detect targets of
less than 14 m length is mainly due to the lack of such small ships in the learning set used for training the NN classifier. Even though these small targets were correctly screened-out as being possible ships during the segmentation stage, they could not be correctly assigned to the ‘shrimp boat’ class in the classification stage. For most of the detected ships, the algorithm overestimated ships’ lengths compared to visually extracted sizes. This is a result of particular image conditions: almost all the ships were in motion. On optical imagery, the moving ship and its near wake are difficult to separate because they are connected and can have similar brightness, so wake and (moving) ship detection often amount to the same. Therefore during segmentation, a ship and its wake were considered as being part of the same image object, resulting in an overestimation of ship size.

Even though the overall results are encouraging, there still remain several issues that need improvement or refining in order to render ship detection from optical satellite imagery fully operational. The algorithm was tested in conditions of calm sea and almost cloudless sky, thus limiting the probability of false alarms. In the near future, more work has to be done into testing the algorithms’ performance under different weathering conditions. Depending on the amount of false alarms that would be obtained following these tests, it would be possible to integrate weather and oceanographic data to reduce false alarms rates. The present work needs to be extended to implement an automatic wake detection approach so as to improve ship length classification and speed estimation. A great deal of effort is currently being undertaken to improve validation procedures and control efficiency by i) introducing information from other maritime monitoring systems, such as VMS, ii) cross-cueing to other sensors, such as the Synthetic Aperture Radar (SAR) sensor, for obtaining or confirming classification.

5. Conclusion

This paper has demonstrated the following:
- the feasibility of ship pattern recognition from high spatial resolution satellite imagery,
- the appropriateness of a feature-based approach for ship detection and
- the viability of utilizing neural networks evolved by genetic algorithms in classifying shrimp boats.

From an application point of view, the most remarkable benefit is the great contribution to the detection of illegal fishing activities, especially in areas where AIS information is unavailable.

Future research will include tests with other datasets for various ships and environmental conditions; study of sea state effects, antenna gain and ship motion on detection performance; and evaluation of the algorithm using very high spatial resolution optical imagery (sub-metric pixel resolution).

6. Acknowledgements

This work was conducted within the framework of the IBIS project (Implementation of a Boat Information System, 2003-2005) in collaboration with SPOTimage and IFREMER. It was financed by the French Ministry of Research (RTE program). The authors acknowledge Seas Guyane, Affaires maritimes (AFMAR), COMAR, SPOTimage, IFREMER and Guillaume Jubelin for their valuable support.

7. References


