Recognition of ships for long-term tracking

Sebastiaan P. van den Broek [*], Henri Bouma, Henny E.T. Veerman, Koen W. Benoist, Richard J.M. den Hollander, Piet B.W. Schwering

TNO, Oude Waalsdorperweg 63, 2597 AK The Hague, The Netherlands

ABSTRACT

Long-term tracking is important for maritime situational awareness to identify currently observed ships as earlier encounters. In cases of, for example, piracy and smuggling, past location and behavior analysis are useful to determine whether a ship is of interest. Furthermore, it is beneficial to make this assessment with sensors (such as cameras) at a distance, to avoid costs of bringing an own asset closer to the ship for verification. The emphasis of the research presented in this paper, is on the use of several feature extraction and matching methods for recognizing ships from electro-optical imagery within different categories of vessels. We compared central moments, SIFT with localization and SIFT with Fisher Vectors. From the evaluation on imagery of ships, an indication of discriminative power is obtained between and within different categories of ships. This is used to assess the usefulness in persistent tracking, from short intervals (track improvement) to larger intervals (re-identifying ships). The result of this assessment on real data is used in a simulation environment to determine how track continuity is improved. The simulations showed that even limited recognition will improve tracking, connecting both tracks at short intervals as well as over several days.

Keywords: track recognition, persistent tracking, maritime, situational awareness, electro-optical, infrared, ship

1. INTRODUCTION

In naval operations, such as piracy prevention, surveillance takes place for a longer time, in large areas, with irregular intervals of observation of ships. Knowing if an observed ship has been seen before is helpful for different reasons. If it can be known from a distance whether a ship was earlier assessed as non-threatening, it does not require further observation time. Furthermore, if the history of a ship is known, this can help to predict its intent, for example based on whether the ship has been seen in a suspect location, or if its behavior in time makes it suspicious. When intent of ships can be predicted, it may help planning the use of assets such as UAVs to more efficiently cover a (larger) surveillance area. In the Netherlands MoD research program V1114, Maritime Situational Awareness (MSA), these aspects are examined in different projects. This paper describes part of the research on determining the use of ship's features obtained from electro-optical systems, in aiding the (long term) tracking of small ships. This research is a continuation of work on using features from electro-optical imagery for discriminating between types of ships (where one image is compared to a single description of a type from multiple images) [9], to recognition of single recordings. In this earlier work, the use of scale-invariant feature transform (SIFT) keypoints in infrared recordings was found to be useful even with low detail imagery of ships. A difficulty with the use of keypoints is how to account for different numbers of keypoints in two images under comparison. [17] compared different types of keypoint descriptions, and added a comparison of relative keypoint locations as further improvement. Instead of matching individual descriptors, there are ways to aggregate the local descriptors in a single object/image level description. For instance, Fisher vector encoding [13][14] is a recent method for making compact image representations while retaining sufficient discriminative power. In this paper, we use Fisher vectors in combination with SIFT keypoints for recognizing ships in infrared images.

In the MSA research program, we also use simulated ship traffic for the testing of persistent tracking, intent estimation and asset planning. This allows the evaluation of the effect of recognition in a realistic scenario with respect to ships encounters and observation intervals.

[*] bas.vandenbroek@tno.nl; phone +31 888 66 4086; http://www.tno.nl

S.P. van den Broek, H. Bouma, H.E.T. Veerman, K.W. Benoist, R.J.M. den Hollander, P.B.W. Schwering, "Recognition of ships for long-term tracking," Proc. SPIE 9091, (2014). http://dx.doi.org/10.1117/12.2053295

The outline of the paper is as follows. The different recognition methods are presented in Section 2. The experiments and results are shown in Section 3. Finally, conclusions and recommendations are presented in Section 4.

2. RECOGNITION METHOD

2.1 Overview of the method

The recognition method consists of two main steps (Figure 1). First, detection and segmentation is performed to localize the targets (Sec. 2.2). Second, matching is performed to recognize the targets. In this paper, we compare matching approaches based on central moments (Sec. 2.3), localized SIFT (Sec. 2.4) and SIFT with Fisher vectors (Sec. 2.5). Matching one reference image to all other images, representing a search for an earlier observation of the same ship, could improve the tracking of a ship. The benefit of recognition for tracking in a more realistic scenario of ships and observation intervals, is examined in a simulation environment (Sec. 2.7).

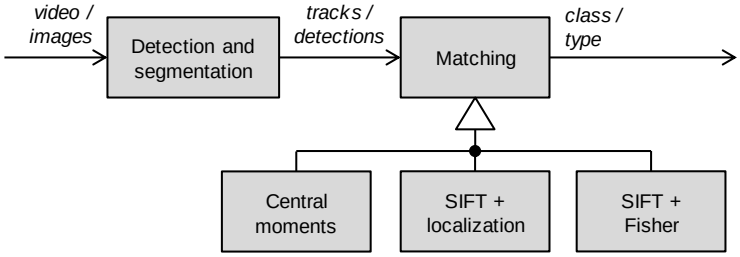

Figure 1: System overview.

2.2 Detection and segmentation

To separate foreground objects (e.g., ships) from background (sea, sky), we used the detection method described in [6]. In this method, the background is assumed to have properties that are constant per image line (corrected for rotation) and to vary only smoothly across the lines. From infrared statistical background analysis work of Schwering [20][21][22] we know that this assumption is sufficiently accurate for various levels of sea state, in open ocean as well as in bay environments. Pixels that are statistically unlikely to be background are clustered in a segmentation of possible ships. This results in a binary mask and bounding boxes around the target, which are manually annotated.

2.3 Matching with normalized central moments

The first type of matching is based on normalized central moments [8][9], for the binary mask of detected pixels. These moments are statistical features that describe properties such as symmetry and skew, which are informative for the general shape of the detected object.

The zero- to second-order central moments result in 9 values, but three of those contain no information. For example, the 0-0 moment gives the number of pixels, which is used to normalize the moments, resulting in a value of 1. Similarly, the moments with order 0 in x or y direction are related to width and height. The remaining six vary in different amounts for ships, and can therefore be assumed to be discriminative to a different extend. In order to obtain a relevant distance measure, the reciprocal value of the standard deviation over all images is used as a weight for each of these features, similar to a statistical distance. The similarity distance of one image to another is then the root mean square of difference in moments, times their weights.

2.4 Matching with SIFT and localization

The second type of matching is based on scale-invariant feature transform SIFT [15][16] and keypoint localization [17]. Often when using SIFT descriptors for matching, only the number of matches (that are similar enough) is used as a match criterion, sometimes weighed to the number of keypoints. However, this does not take into account the locations

of these keypoints relative to each other. The performance improves when this relative location is taken into account [17]. The relative location can be computed by comparing locations relative to the center-of-mass of the thresholded detection image (see description in Section 2.2), scaled by the width (to compensate for slight differences in aspect angle) and height of the detection (to compensate for observation distance). Matching will be demonstrated with and without the localization constraint in order to demonstrate the effect of this condition. We will use a somewhat different distance measure than in [17], namely:

$$1 - \frac{\#\text{matches}}{\#\text{keypoints}} \tag{1}$$

where only the number of keypoints in the observation image is used in the denominator. Since matching is performed for each of the 'observed' keypoints, there will be at maximum this number of matches and the distance lies in the range [0,1].

2.5 Matching with SIFT and Fisher vectors

We use sparse SIFT to describe the local structure in images. For each keypoint in the image, a descriptor of length 128 is computed. This results in a variable number of descriptors per image. There are several ways to obtain a more compact representation for the image. One way is by using a Bag-of-words (BOW). The BOW approach first computes cluster centers (e.g., with k-means clustering) from the descriptors. Subsequently, the variable number of SIFT-points are transformed to a histogram of fixed length by making a hard assignment of each SIFT descriptor to the closest cluster center. One fixed-length vector is computationally less intensive for matching than a variable number of SIFT-descriptors. A more advanced approach to obtain a compact representation is the Fisher Vector (FV) [13][14], which is based on soft-assignment to cluster centers. This also reduces the variable number of descriptors to one fixed-length descriptor.

The method consists of the following steps:

- Computation of SIFT descriptors.
- Principal component analysis (PCA) on the SIFT-descriptors.
- Gaussian mixture model on the SIFT after PCA mapping (GMM).
- Computation of the Fisher Vector.
- Power-normalization and L2-normalization.
- Optionally another PCA mapping of Fisher Vectors.
- Computation of nearest neighbors (NN), which results in a dissimilarity distance.

The method contains two important parameters: the number of Gaussians in the GMM (*g*), and the number of principal components to keep with PCA of the SIFT (*p*). Optionally, another PCA can be applied to the resulting Fisher Vector (parameter *p2*). We did not perform this second PCA. Note that the length of the Fisher Vector is *p*g* (or optionally *p2*).

2.6 Robust and symmetric recognition

All three matching methods result in a dissimilarity distance between one 'observation' image, and all other images (representing the historical observations). If recognition is correct, the most similar image, with smallest dissimilarity distance, is an image of the same ship. A confusion matrix can be made, showing ground truth (GT) against the predicted best fit, counting the percentages that a ship is recognized correctly, or that one of the other ships has the smallest dissimilarity distance. This is called the rank-1 result. An overall accuracy is obtained by taking the average of the diagonal as measure of correct recognition. With the limited data set we are using, and sometimes small number of images of a ship, the rank-1 result may not be the best estimate. Therefore a more robust measure is used, where the matching ship is the one that, when sorting images by the dissimilarity measure and increasing the number of selected images one by one, has a given (interpolated) percentile of its images in this selection, before any other ship. The percentile is computed over the ranks of each class separately. The 0-percentile result corresponds to the rank-1 result of the class and the 50-percentile is the median rank of the class. The class with the best percentile value is selected. Note that for 0-percentile, this is identical to selecting the best rank-1 result. Both confusion matrices of the rank-1 result and graphs of the percentile results will be presented for the different matching methods.

The methods have an a-symmetry, in that a horizontally mirrored version of an image would not match, for example because odd-ordered moments are negative, and keypoints have a different description and location. To compensate for this without changing the feature computation and localization, horizontally flipped versions of all images are added to the image set. When computing dissimilarity distances between an image and all others, both the original image and its flipped version are discarded.

2.7 Simulation environment

In the MSA (Maritime Situational Awareness) research program, persistent tracking is one of the topics of research. Other topics are intent estimation based on the observed tracks (and other information) [4] and asset planning, where assets (such as helicopters and UAVs controlled from a frigate) are planned in time based on updated risks from intent and commercial shipping. These three parts are also built into a simulation environment, where ship traffic is simulated and observations are created based on sensor ranges and vessel appearance. An impression of the simulated traffic is shown in Figure 2, where white marks indicate co-operative (commercial) vessels, and other marks are (smaller) vessels, such as skiffs, dhows, trawlers, etc. The squares indicate the three observation areas, each covered by a frigate with helicopter and UAV.

Figure 2: Simulation environment.

This simulation chain (see Figure 3) provides a feeling for the kind of recognition performance that would be helpful in persistent tracking, given a more realistic distribution of (small) ships, and intervals of observation. The latter depends on the platform planning and (a-priori) intent risk estimation. There is a feedback loop from tracking, via asset planning, back to observations. The implementation of persistent tracking is as follows:

- For commercial (cooperative) shipping, it is assumed that positions are known. These ships are not taken into account in the further examination.
- As long as ships are in radar range, they are assumed to be tracked by the observing platform.
- Only limited kinematic tracking is implemented, based on last observations and a maximum speed.
- Recognition is simulated by an single value for each vessel, which is only observed by certain sensors and at limited range. A difference in this value is an indication of the possibility of the observation matching a track. Setting a maximum allowed difference is a way of indicating how probable a confusion between ships is. I.e., if values are allowed to differ by 10%, one in ten ships looks alike, at 1% only one in a hundred.

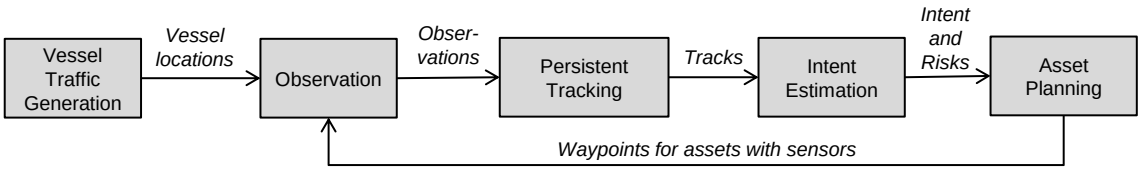

Figure 3: Modules in the MSA simulation chain.

As a measure of performance, the average and maximum number of track IDs given to each ship are counted. Recognition is only possible when recognition information becomes available. As at that moment there will already be a (new) radar track, the overall number of IDs would be the same whether there is a correct recognition or not. In order to measure an improvement when recognition is applied, in cases a ship (with a radar only track) first has a recognition observation (i.e., is in camera range), an artificial track break is applied, by not using the radar ID for identification.

By changing the limits on recognition (i.e., the confusion between ships), the effect of recognition can be examined, and compared with the situation where no recognition is used.

3. EXPERIMENTS AND RESULTS

3.1 Description of the imagery

For assessing the different methods of matching, recordings of five small ships are used, recorded during the international SMARTEX trial, which took place over two days in June 2012 in co-operation with the US Coast Guard. An example of MWIR (Mid Wave Infrared, wavelength band 3-5 μm) data is given in Figure 4, where a recording of different small ships is shown. Target location from GPS is available for several targets, which is used in annotating images and providing distance and aspect angle estimations. Currently, five targets are used in these tests for recognition. In Figure 4, four of those are visible in the MWIR image: the sailing boat (center), two 24-feet cabin boats (called 24A and 24B, on two sides of the sailing boat) and a flat boat of 35 feet (left). In addition, a 74 feet flat boat is used.

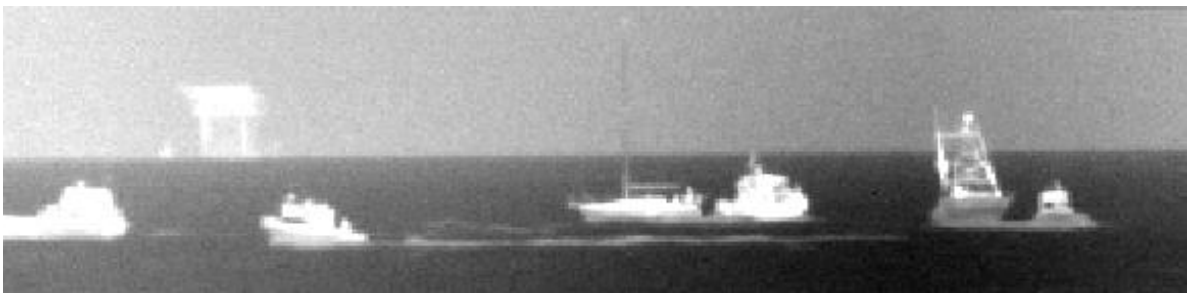

Figure 4: SMARTEX MWIR image containing small targets, including a sail boat, cabin boats, and other boats of similar size.

From the ship positions, a distance and aspect angle relative to the camera can be estimated, as well as expected pixel resolution on the ship. From this, images are automatically retrieved over a range of resolutions and aspect angles. Most images (95%) range in resolution from 0.1 to 0.5 meters per pixel, with a maximum of 1 meters per pixel (see examples in Figure 5). 80% of images are within 20 degrees of side view, with 10% being closer to frontal/back views. Images are taken at least one minute apart, in the few cases recordings fit the selection criteria for a longer period. After applying the automatic detection and segmentation, images are manually checked to see if the ship was actually in the image, and detected. This annotation results in 186 recordings of the five different ships, in both mid-wave and long-wave infrared. Table 1 gives the number of images per ship.

Table 1: Ships and numbers used in the experiment.

Ship label	Description	Number of samples	Fraction
24A	24-feet cabin boat	54	29%
24B	24-feet cabin boat	67	36%
35	35-feet flat boat	21	11%
74	74 feet flat boat	4	2%
Sail	sail boat	40	22%
TOTAL	-	186	100%

The 35 and 74 (feet) targets are flat boats that are somewhat alike and viewed at a distance they are not dissimilar from the sail boat, of which its mast is often not seen. The 74 flat boat only had 4 examples in this set, which makes the result for this target less accurate.

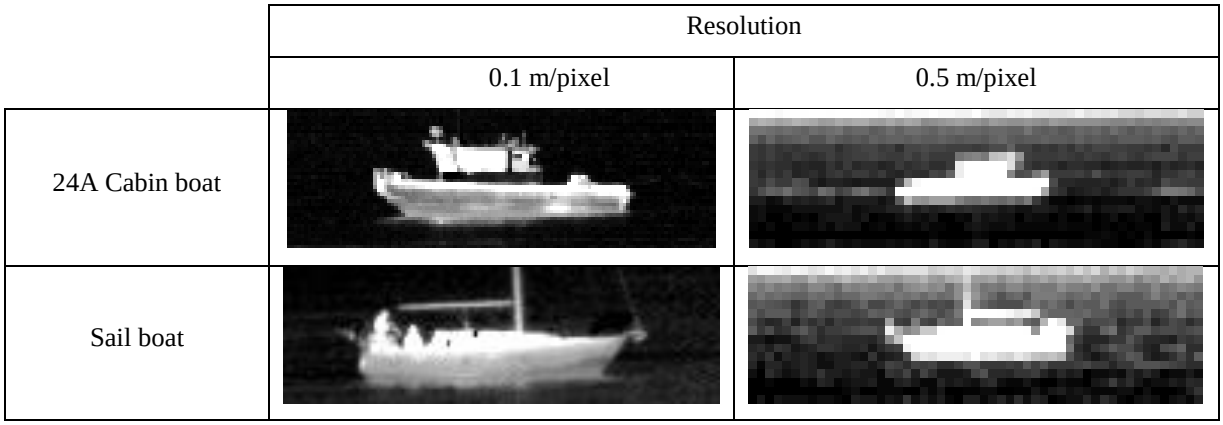

Figure 5: Examples of MWIR images of ships in the data set, with high and low resolution.

3.2 Recognition results with central moments

The confusion matrix for central moments is shown in Table 2. The data set includes flipped images to make the matching symmetric. Even with these very simple features and matching method, most often the correct target is recognized, distinguishing for example between the two cabin boats well above their a-priori probabilities.

Table 2: Confusion matrix using central moments in percent based on the best rank-1 fit. The accuracy is 67%.

		Predicted				
		24A	**24B**	**35**	**74**	**Sail**
GT	**24A**	**72**	24	4	0	0
	24B	13	**78**	7	0	1
	35	10	29	**52**	0	10
	74	0	0	0	**50**	50
	Sail	3	3	8	5	**83**

When combining the two 24-feet cabin boats (i.e., counting a correct match when the other cabin boat was indicated), their correct recognition increases to 93% (with an overall accuracy of 70%). Another interesting test is to compare each image only with images of the other day (in this two day recording). In this case, recognition values are very similar for the three ships that were recorded on both days. This indicates that similarity is not just based on recordings that are made with the same camera, at similar distance and aspect angle, at a time near the reference image.

Table 3: Confusion matrix using central moments in percent based on the best rank-1 fit, where the classes 24A and 24B are merged. The accuracy is 70%.

		Predicted			
		24A+B	**35**	**74**	**Sail**
GT	**24A+B**	**93**	6	0	1
	35	38	**52**	0	10
	74	0	0	**50**	50
	Sail	5	8	5	**83**

Table 4: Confusion matrix using central moments in percent based on the best rank-1 fit on another day. The accuracy is 71%. Classes 35 and 74 were only present on one of the days.

		Predicted				
		24A	**24B**	**35**	**74**	**Sail**
GT	**24A**	**69**	30	2	0	2
	24B	21	**69**	7	0	3
	35	-	-	-	-	-
	74	-	-	-	-	-
	Sail	13	3	10	0	**75**

Figure 6 shows the accuracy for different percentiles, according to the method described in Sec. 2.6. This shows that rank-1 (0 percentile) – which is based on the best match only – gives better results than higher percentiles – which are based on the top (5 to 50) percentiles of best matches. The decrease is however not very steep.

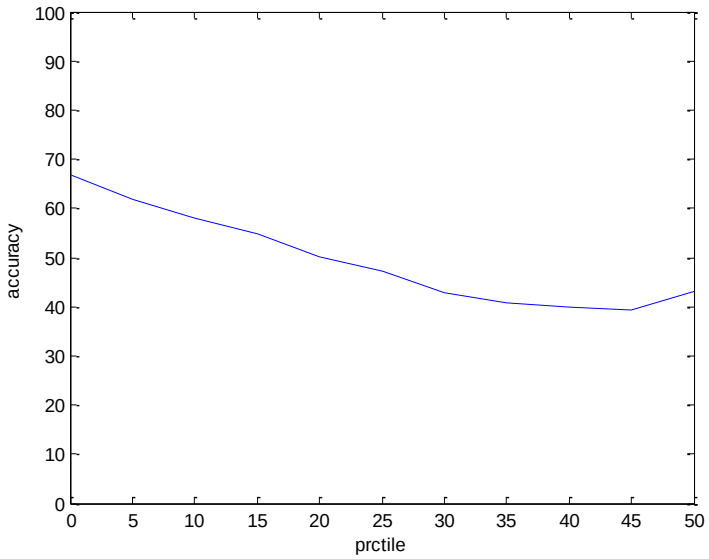

Figure 6: Accuracy for central moments for varying percentiles.

3.3 Recognition results with SIFT matching

We first evaluated the results for SIFT matching with and without the localization constraint. The confusion matrix for the rank-1 result without localization is shown in Table 5 and with localization in Table 6. It is clear that the localization of the matching keypoints is a powerful additional constraint, since the accuracy improves from 59% to 68%.

Table 5: Confusion matrix using SIFT matching in percent based on the best rank-1 fit. The accuracy is 59%.

		Predicted				
		24A	**24B**	**35**	**74**	**Sail**
GT	**24A**	**61**	20	19	0	0
	24B	9	**72**	6	0	13
	35	14	19	**52**	0	14
	74	0	25	25	**25**	25
	Sail	0	13	0	3	**85**

Table 6: Confusion matrix using SIFT matching with the localization constraint in percent based on the best rank-1 fit. The accuracy is 68%.

		Predicted				
		24A	**24B**	**35**	**74**	**Sail**
GT	**24A**	**74**	22	2	0	2
	24B	24	**64**	4	1	6
	35	19	14	**67**	0	0
	74	0	25	25	**50**	0
	Sail	15	0	0	0	**85**

The addition of flipped images in the evaluation is expected to improve the classification result further. However, as shown in the percentile graphs of Figure 7, it turns out that there is no clear improvement in classification accuracy for any percentile threshold. This could be due to the number of images of (most) ships, making it likely there is already one in both directions in the set.

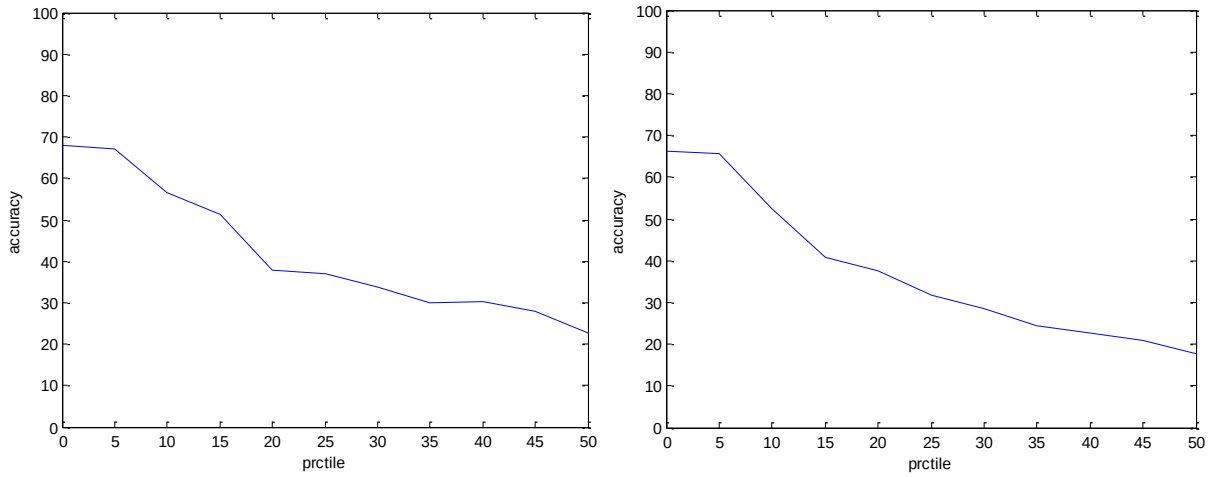

Figure 7: Accuracy for SIFT matching for varying percentiles in the configuration similar to Table 6 with only the original images (left) and including the flipped images (right).

3.4 Recognition results with SIFT and Fisher Vector matching

First we tested two parameters of the Fisher Vector without flipped images and with flipped images. We used 10 iterations, because the clustering inside the GMM contains a stochastic component. The results are shown in Table 7, which shows that the highest accuracy is 76%, although the variation is not very large between the different parameter settings. The corresponding confusion matrix for 5-percentile is shown in Table 8 and the accuracies for different percentiles in Figure 8. For completeness, we also inserted the confusion matrix of the 0-percentile (rank-1) in Table 9. Note that the 5% performs poorly in ship 24B and the 0% in ship 74. The poorer performance of ship 74 in the 0% case can be explained by the low number of samples and hence the lower probability of obtaining a rank-1 result. We suspect that the poorer performance of ship 24B in the 5% case can be explained by the high variation in these images. The results in Figure 8 show that the performance with flipped images is slightly better than without flipped images.

Table 7: Accuracies using Fisher Vectors for different parameters (*g* for GMM and *p* for PCA) over 10 iterations (mean ± standard deviation). The highest accuracy is 76%.

		No Flip			**Flip**		
		p			***p***		
		32	**64**	**128**	**32**	**64**	**128**
g	**32**	70.6 ±2.8	72.5 ±1.5	71.5 ±2.7	70.2 ± 3.5	71.6 ± 2.3	75.9 ± 3.3
	64	73.1 ±2.4	74.9 ±3.3	72.4 ±2.0	73.1 ± 2.5	75.1 ± 2.3	73.6 ± 2.3
	128	74.0 ±2.9	74.5 ±1.6	74.1 ±2.0	74.7 ± 1.9	**76.4 ± 1.8**	75.1 ± 2.8

Table 8: Confusion matrix using Fisher Vectors in percent based on the best 5% fit (g=128, p=64) with flipped images. The accuracy is 76%.

		Predicted				
		24A	**24B**	**35**	**74**	**Sail**
GT	**24A**	**77**	9	6	6	2
	24B	22	**43**	15	6	13
	35	5	9	**74**	6	5
	74	0	0	1	**99**	0
	Sail	2	0	8	7	**83**

Table 9: Confusion matrix using Fisher Vectors in percent based on the best rank-1 (0 percentile) fit (g=128, p=64) with flipped images. The accuracy is 72%.

		Predicted				
		24A	**24B**	**35**	**74**	**Sail**
GT	**24A**	**77**	16	4	0	3
	24B	12	**74**	5	0	9
	35	14	12	**71**	1	1
	74	21	11	13	**46**	9
	Sail	1	2	4	0	**93**

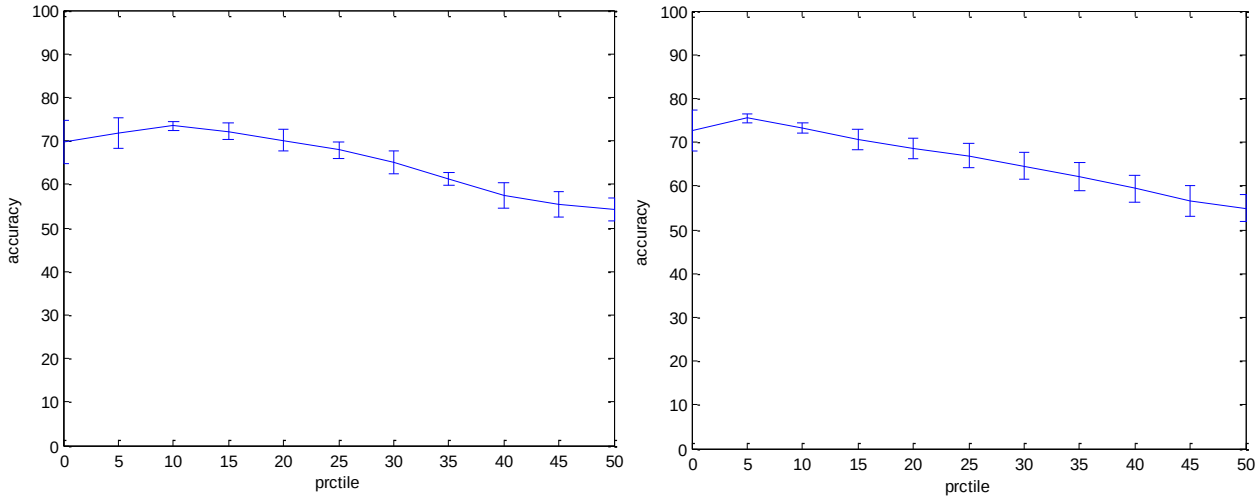

Figure 8: Accuracy for Fisher Vectors with varying percentiles (parameters *g=128*, *p=64*) without flipped images (left) and with flipped images (right). The error bars indicate the standard deviation of 10 iterations.

3.5 Comparison of matching methods and discussion

The comparison of accuracy values for the different matching methods is shown in Table 10.

Table 10: Comparison of matching methods.

Method	Accuracy
Central moments	67%
SIFT matching	59%
SIFT + localization	68%
SIFT + Fisher Vector (0%)	72%
SIFT + Fisher Vector (5%)	**76%**

The values of Table 10 show that overall, SIFT in combination with Fisher Vectors performs best.

The fact that the simple central moments, and the SIFT with localization are able to distinguish between the two cabin boats (see Table 2 and Table 6) is an interesting result, as to a human observer there is no clear difference between the two. The Fisher vectors seem to be better in describing the different types of boats, as the sail boat and the 35 and 74 flat boats are better recognized and not so often confused with the cabin boats.

The data set poses some limitations, in the small number of different ships, and the small number of images for some of them. The percentile results compensate to some extend for the latter, indicating an overlap in sorted matches. These indicate that the central moments and SIFT keypoint matching methods show a larger drop than the Fisher vector matching, when increasing the percentile of images. This indicates that overall, the Fisher vectors provide a better separation of the images of the different ships.

3.6 Effect of recognition in tracking in simulations

For this test, a single recognition value, indicated by confusion probability CP, is used to simulate how well ships can be distinguished from each other when a camera observation is available. In the association from observations to existing tracks, this is used to eliminate unlikely matches. A value of one indicates no recognition improvement, as each ship can appear the same as all others. Lower values indicate some exclusion is possible. For example, with CP=0.1, of possible wrong associations, 90% will be excluded. A value of 10^{-6} indicates (practically) perfect recognition, i.e., no wrong matches occur.

Table 11 shows the results of a simulation of one day of traffic. Note that only 79 smaller ships are observed (by sensors) in this time, in the limited area of operation of the three frigates and their assets (see Figure 2), and recognition can only help when multiple observations are done with a camera where the ship was out of view in between observations. Table 11 shows that even limited recognition in this data already gives almost as much gain as perfect recognition.

Table 11: Results of persistent tracking for different confusion probability (CP) values, from 1 (only kinematics in the tracking algorithm) to 10^{-6} (very precise recognition).

Confusion probability	Measures of performance	
	Average tracks per ship	Max. tracks per ship
1	3.910	12.70
0.1	3.068	10.30
0.02	3.009	10.75
0.005	3.000	11.00
10^{-6}	2.996	10.00

Table 12: Results of persistent tracking for a confusion probability (CP) value of 1 (only kinematics in the tracking algorithm), and for a ship type dependent recognition.

Confusion probability	Measure of performance	
	Average tracks per ship	Max. tracks per ship
1 (no recognition)	18.1	69
Type specific, from 1/4 to 1/50	9.1	51

Table 12 shows the results over five simulated days, increasing the number of observed ships to 290, as well as increasing how often they are revisited by a sensor. This results in an average of 18.1 track IDs, without recognition. For recognition CP is set at fixed values for different types of ships: 0.25 for skiffs, 0.05 for dhows and 0.02 for (less common) larger vessels. In that case, it shows that even this limited recognition of the smaller ships (skiffs, dhows) gives a major improvement. Note also that for measuring performance, extra track breaks were introduced (see section 2.7), so both values have a (similar) offset, and total number of tracks would be lower.

4. CONCLUSIONS

In this paper, we presented our research on the use of feature extraction and matching methods for recognizing ships from electro-optical imagery. An indication of discriminative power is obtained on infrared imagery of ships. Simple central moment features already show some capability for recognition, even distinguishing between two nearly similar ship types. The use of SIFT keypoints gives a similar result, if the relative location of the keypoints is taken into account. Fisher vectors to describe SIFT keypoints is an efficient way of indicating the difference between a current observation and many images in a database. The overall performance using SIFT keypoints with Fisher vectors is an improvement over the other methods.

Although the recognition is not perfect, continuing improvement of methods increases the possibilities of semi-automatic recognition, especially when combining different features, such as moments and SIFT keypoints, with for example contours or other ship characteristics.

The simulations showed that even limited recognition will improve tracking, connecting both tracks at short intervals as well as over several days. This will help in higher level assessment of the observed ships, estimating intent based on a ships history and behavior.

ACKNOWLEDGEMENT

The work for this paper was supported by the Netherlands MoD program V1114, Maritime Situational Awareness.

REFERENCES

[1] Alves, J., Herman, J., Rowe, N., "Robust recognition of ship types from an infrared silhouette," Thesis Monterey Naval Postgraduate School California USA, (2004).
[2] Bouma, H., Lange, D.J. de, Broek, S.P. van den, Kemp, R., Schwering, P., "Automatic detection of small surface targets with electro-optical sensors in a harbor environment," Proc. SPIE 7114, (2008).
[3] Bouma, H., Dekker, R., Schoemaker, R., Mohamoud, A., "Segmentation and wake removal of seafaring vessels in optical satellite images," Proc. SPIE 8897, (2013).
[4] Broek, A.C. van den, Broek, S.P. van den, Heuvel, J.C. van den, Schwering, P., Heijningen, A.W. van, "A multi-sensor scenario for coastal surveillance," Proc. SPIE 6736, (2007).
[5] Broek, A.C., van den, Hanckmann, P., Smith, A., Bolderheij, F., "Vessel intent recognition in maritime security operations," Proc. NATO SCI 247, (2012).
[6] Broek, S.P. van den, Bakker, E.J., Lange, D.J.J. de and Theil, A., "Detection and classification of infrared decoys and small targets in a sea background," Proc. SPIE 4029, 70-80 (2000)
[7] Broek, S.P. van den, Schwering, P., Liem, K., Schleijpen, R., "Persistent maritime surveillance using multi-sensor feature association and classification," Proc. SPIE 8392, (2012).
[8] Broek, S.P. van den, Bouma, H., Degache, M., "Discriminating small extended targets at sea from clutter and other classes of boats in infrared and visual light imagery," Proc. SPIE 6969, (2008).
[9] Broek, S.P. van den, Bouma, H., Degache, M., Burghouts, G., "Discrimination of classes of ships for aided recognition in a coastal environment," Proc. SPIE 7335, (2009).
[10] Dekker, R., Bouma, H., Breejen, et. al., "Maritime situation awareness capabilities from satellite and terrestrial sensor systems," Proc. Maritime Systems and Technologies MAST Europe, (2013).
[11] Gray, G. J., Aouf, N., Richardson, M. A., Butters, B., Walmsley, R., Nicholls, E., "Feature-based recognition approaches for infrared anti-ship missile seekers," Imaging Science Journal 60(6), 305-320 (2012).
[12] Ergula, M., Alatana, A., "An automatic geo-spatial object recognition algorithm for high resolution satellite images," Proc. SPIE 8897, (2013).
[13] Jegou, H., Douze, M., Schmid, C. "Hamming embedding and weak geometry consistency for large scale image search," Proc. ECCV, (2008).
[14] Jegou, H., Perronnin, F., Douze, M. Sanchez, J. Perez, P., Schmid, C., "Aggregating local image descriptors into compact codes," IEEE Trans. Pattern Analysis and Machine Intelligence 34(9), 1704 - 1716 (2012).
[15] Lowe, D., "Object recognition from local scale-invariant features," IEEE ICCV, (1999).
[16] Mikolajczyk, K., Schmid, C., "A performance evaluation of local descriptors," IEEE Trans. Pattern Analysis and Machine Intelligence 27(10), 1615-1630 (2005).
[17] Mouthaan, M.M., Broek, S.P. van den, Hendriks, E.A., Schwering, P., "Region descriptors for automatic classification of small sea targets in infrared video," Optical Engineering 50(3), (2011).
[18] Perronnin, F., Sánchez, J., Mensink, T., "Improving the fisher kernel for large-scale image classification," ECCV, 143-156 (2010).
[19] Schwering, P., Lensen, H., Broek, S.P. van den, Hollander, R. den, Mark, W. van der, Bouma, H., Kemp, R., "Application of heterogeneous multiple camera system with panoramic capabilities in a harbor environment," Proc. SPIE 7481, (2009).
[20] Schwering, P.B.W., "Infrared clutter measurements of marine backgrounds," Proc. SPIE 1486, 25-36 (1991).
[21] Schwering, P.B.W., "Maritime Infrared Background Clutter," Proc. SPIE 2742, 255-266 (1996).
[22] Schwering, P.B.W., Bezuidenhout, D.F., Gunter, W.H., le Roux, F.P.J., Sieberhagen R.H., "IRST Infrared Background analysis of Bay environments," Proc. SPIE 6940, (2008).
[23] Steinvall, O., Elmqvist, M., Karlsson, K., Larsson, H., Axelsson, M., "Laser imaging of small surface vessels and people at sea," Proc. SPIE 7684, (2010).
[24] Withagen, P., Schutte, K., Breuers, M., Vossepoel, A., "Automatic classification of ships from infrared (FLIR) images," Proc. SPIE 3720, (1999).