Robust Layer-based Boat Detection and Multi-target-tracking in Maritime Environments

Wolfgang Krüger
Fraunhofer Institute of Optronics,
System Technologies
and Image Exploitation IOSB
Email: wolfgang.krueger@iosb.fraunhofer.de

Zigmund Orlov
Fraunhofer Institute of Optronics,
System Technologies
and Image Exploitation IOSB
Email: zigmund.orlov@iosb.fraunhofer.de

***Abstract*—In this paper we present a robust solution to detect and track small and distant boats from thermal images captured by a camera mounted on an autonomous platform (buoy or patrol vessel). It is characterized by a multiple-layer and multiple-algorithm architecture, which uses a combination of algorithms relying on complementary image cues to generate detections that are robust with respect to variations of boat appearance, image quality, and environmental conditions. The core component of the image exploitation is a detection layer which provides the results of several detection algorithms in a motion-stabilized scene coordinate frame aligned with the estimated horizon line. In the autonomous system, detections are used to trigger alarms and to facilitate multi-target tracking.**

I. Introduction

Criminal activities at sea such as illegal immigration, piracy or trafficking of drugs, weapons and illicit substances have been a reality for past years. Often, small and midsize maritime vessels, which are difficult to detect, are used in such activities. Until now, border agencies observe and protect the critical maritime areas by ships, planes or helicopters, which is very expensive and full coverage is difficult to obtain.

To improve this situation, the European research project AMASS (Autonomous Maritime Surveillance System) investigates to use a network of unmanned surveillance platforms located a considerable distance from shore. The platforms are equipped with different sensors, the optical sensors being uncooled thermal imagers. In order to exploit the data delivered by the thermal imager, detection and tracking algorithms which are able to work with a moving sensor under a variety of weather and visibility conditions are required.

In this paper, we present a robust solution for boat detection and multi-target tracking in maritime environments that has been developed at Fraunhofer IOSB. It is characterized by a multiple-layer and multiple-algorithm architecture, which makes robust implementations of maritime video surveillance systems possible, where each module can be developed individually or can be easily substituted by another module.

The paper is organized as follows: Section II gives an overview of related work. Section III explains the hardware and software architecture and the algorithms developed for image exploitation. In Section IV some detection and tracking results are presented. Conclusions can be found in Section V.

II. Related Work

A variety of tools from machine vision has been applied to detect ships and boats in colour and infrared imagery.

Template matching is used in [1] to detect ships in infrared images. Silhouettes of different ships from different aspect angles are correlated with a region of interest at the horizon line.

In [2] it is proposed to detect ships as groups of adjacent image regions. Mean-shift segmentation is applied to find regions in infrared images and a minimum spanning tree is used to cluster the segmented regions.

A fast radial blob detector which is motivated by blob detection with difference of Gaussians [3] is presented in [4] and applied to maritime target detection. Blobs are detected on multiple scales and clustered using a minimum spanning tree. Tracking is employed to find temporally salient clusters.

Background models are used in [5] and [6] to detect maritime targets that differ from the background. [5] proposes to estimate a functional relationship between the index of an image row and the mean image intensity and standard deviation on that row. In [6] statistical image properties such as entropy are computed in overlapping windows of variable size. These properties are used as features to build a background model which is estimated as the main cluster in feature space.

There are also approaches which use edge information to detect ships. [7] characterizes small infrared targets by combining horizontal edge energy computed on different levels of image resolution. [8] describes an algorithm to detect marine vessels from a colour video camera mounted on a buoy. The Canny edge detector [9] is used to find long edges for vessels with extensive contour lines. In a subsequent work [10] multiple hypothesis tracking is added to further reduce false detections by exploiting temporal continuity of valid detections.

Most approaches require the position and orientation of the horizon line as input. Existing approaches to detect horizon lines in images are based on colour segmentation [11],[12] or edge detection [7].

There is only little information in the literature about visual surveillance from moving autonomous platforms deployed at sea. An example is [13] where the authors describe an

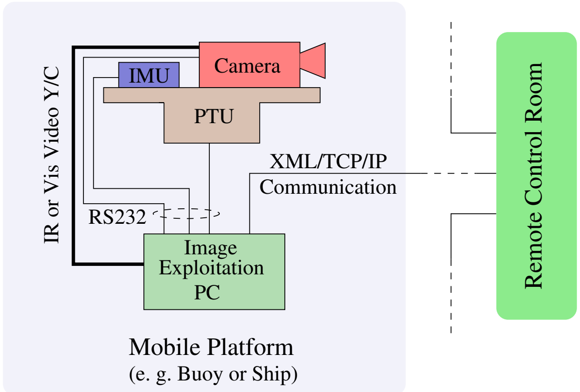

Fig. 1. Overall architecture

un-tethered autonomous buoy that stations itself on the sea floor and is able to ascend to the surface when needed. The installed optical surveillance unit is based on a low-power minicomputer and processes colour images from a web-camera. The camera is more or less at sea-level and no physical stabilization is available. Ship detection and tracking is based on previous work of the authors [8], [10]. Due to limitations in computing power, video data is collected in an online phase and image exploitation has to be done offline afterwards.

In this paper we describe a layered software architecture to detect and track small and distant boat/ships from thermal images captured by a camera mounted on an autonomous platform (buoy or patrol vessel). Compared to other approaches, a combination of algorithms relying on complementary image cues is proposed to generate detections that are robust with respect to variations of boat appearance, image quality, and environmental conditions. The core component of the software architecture is a detection layer which provides the results of several detection algorithms in a motion-stabilized scene coordinate frame aligned with the estimated horizon line. In the autonomous system, detections are used to trigger alarms and to facilitate multi-target tracking.

III. Architecture and Algorithms

The task of image exploitation developed by Fraunhofer IOSB for maritime surveillance systems is to automatically detect and track distant vessels in the images generated by the thermal or visual imager on an autonomous mobile platform such as buoy or ship. In the following, the image-exploitation hardware and software architecture on the mobile platform is described.

A. Image-Exploitation Hardware Architecture

Fig.1 depicts the hardware architecture of the image exploitation on an autonomous platform. There are four components: camera, pan-tilt unit (PTU), the inertial measurement unit (IMU) and the image-exploitation computer (IE-PC)

The IMU and the camera are located together on the turret of the platform. Both components are mounted on the PTU. The task of the IMU is to measure the angular orientation of the camera with respect to a world coordinate frame. The measured roll and pitch angles are needed by the image exploitation software in order to estimate position and slope of the horizon line for processed camera images.

The IE-PC is mounted in a waterproof rack. It is connected to the ethernet switch, the IMU, the PTU and the camera. Its task is to run image exploitation algorithms using received IMU data, to control camera and pan-tilt unit, and to perform communication with a remote control room by using XML-based messages over TCP/IP.

An operator in the remote control room has to evaluate the information received from sensor system and to make decisions how to act.

B. Image-Exploitation Operating Modes

There are three operation modes supported by image exploitation: manual scan, miniscan, tracking.

In *manual scan mode*, the operator will see the current image from the camera looking to the desired direction. No image exploitation algorithms are executed.

In *miniscan mode*, an automatic observation of a certain angle range is initiated by a command from the remote control room. Image exploitation algorithms are executed for an internally chosen number of camera positions covering the desired angle range. Upon detection of small vessels alarms are generated and sent to the remote control room, where the operator has to evaluate them.

In *tracking mode*, a boat at an user-selected direction will be tracked automatically. This mode is useful after a vessel detection and a positive evaluation by the operator. This mode is initiated by a command from the remote control room.

C. Image-Exploitation Software Architecture

The image exploitation software consists of four processes communicating with each other:

- *Communication and Control Process (CCP)*. It is the image exploitation interface to the remote control room and to the hardware components of the image exploitation on the buoy. It coordinates and controls each of the image exploitation operating modes described above.
- *Inertial Measurement Unit Communication Process (IMUCP)*. The task of IMUCP is to communicate IMU measuring data to CCP in real-time. Data is send with a frequency of 100 Hz.
- *Optronic Process (OP)*. The OP controls the PTU and the camera. It executes appropriate operations upon commands from CCP. Additionally, it communicates the current compass data to the CCP periodically and in real-time.
- *Image Exploitation Process (IEP)*. The IEP contains algorithms for boat detection, classification, and tracking that are executed upon commands from CCP. Results are communicated to the CCP for further analysis and notification of the remote control room.

1) Communication and Control Process (CCP): CCP is the master multithread-process in the IE-PC initiating, coordinating, and controlling the operating modes and the communication with the image exploitation system (Fig.2).

CCP has four communication partners (remote control room, IMUCP, IEP, OP). The communication protocol is XML-based. The messages are delivered over TCP/IP.

The remote control room sends its requests to CCP, where they are processed and transformed to the commands for OP and IEP. The images generated by the thermal imager are delivered to IEP in a shared memory area. Every data entry in the shared memory contains an image and appropriate IMU data, received from IMUCP in real-time. IEP sends its results to CCP over XML/TCP/IP channels. The CCP analyzes them and sends e.g. alarms to the remote control room, where they have to be evaluated by the operator.

According to the described operating modes of the image exploitation, CCP implements a state machine controlling the execution of operating modes (Fig.3) .

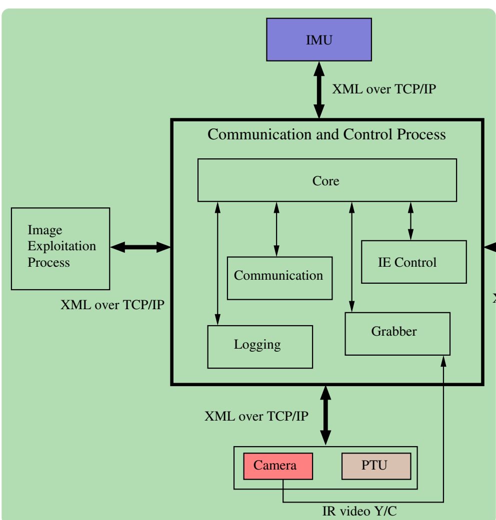

Fig. 2. CCP architecture

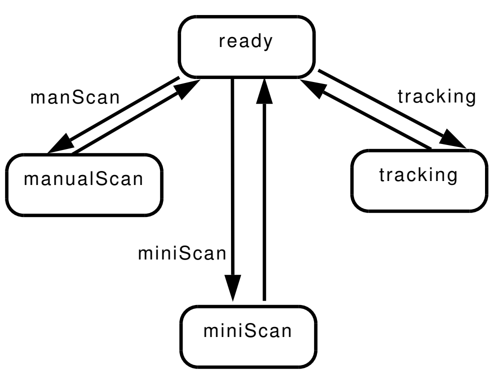

Fig. 3. CCP IE state machine

There are four states: *ready*, *manualScan*, *miniScan*, and *tracking*. Only from state *ready*, all IE operating modes are directly reachable, i.e. at one point of time only one IE operating mode is possible. If an operating mode is finished or an abort-command was received from the control room, the CCP is in state *ready* again.

2) Image Exploitation Process (IEP) : The image exploitation process (IEP) implements and runs the image exploitation algorithms. The algorithms are structured according to the layer model in Fig.4. Each layer builds on the results of the previous layer. Data processing is done from bottom to top.

The first and lowermost layer is the IMU-based estimation of camera orientation. Input for this layer is the measurement data generated by sensors (accelerometers, gyroscopes, and magnetometers) inside the IMU. An extended Kalman filter is used to estimate the time-varying pitch and roll angles of the camera to which the IMU is attached.

The next layer uses the estimated pitch and roll angles to find and improve the localization of the horizon line in the captured camera images. The horizon line is determined by a robust fit to edge features extracted from the images. Pitch and roll angles from the IMU-layer are used to narrow the search areas for feature extraction. The image-based horizon localization also detects situations (e.g. bad visibility due to fog) when the image quality is too low to find the correct horizon line. In such a case the localization of the horizon line is estimated from pitch and roll angles alone. Results of this layer are demonstrated in Fig.5 and Fig.6.

The third software layer is the boat detection layer. Since the aim is to detect distant boats and those boats will appear near the horizon line, the boat detection layer uses the information about the horizon line in an image to set up search areas for the implemented detection algorithms. The search areas are fixed relative to the horizon line. Since uncompensated pitch motion of the buoy on which the camera is mounted will cause the horizon line to move up and down in the images, this amounts to an additional software-based image stabilization.

The common idea of the boat detection algorithms is to search for temporally stable image features to separate detections at boats from those at sea clutter. To enhance robustness with respect to variations of boat appearance, image quality, and environmental conditions, a combination of detection algorithms using different image features and search strategies has been implemented in the detection layer. The first algorithm is a track-before detect algorithm [14]

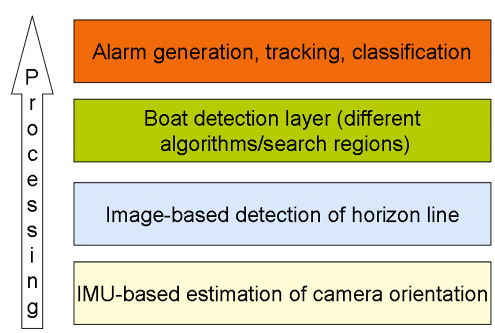

Fig. 4. Layer structure of the image exploitation algorithms.

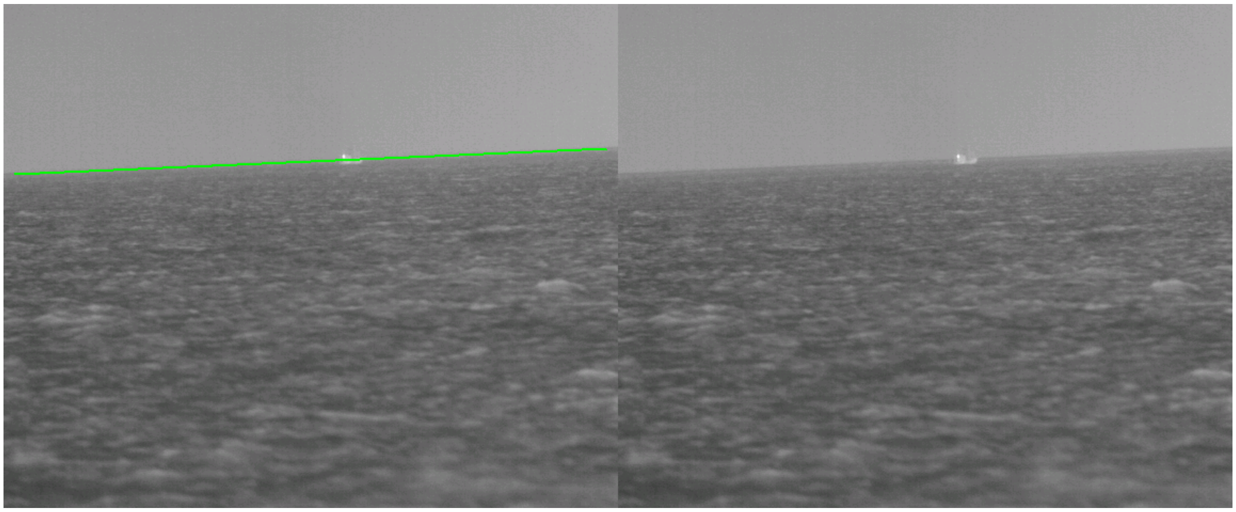

Fig. 5. Example of the estimated horizon line (left, green). Corresponding input images on the right.

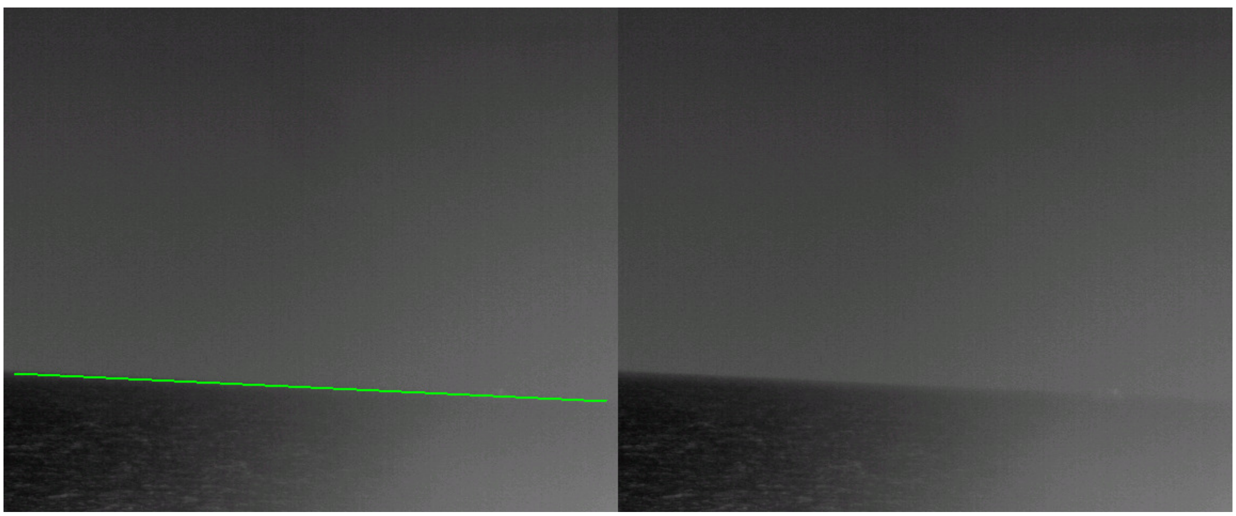

Fig. 6. Example of the estimated horizon line (left, green). Corresponding input images on the right.

using spatio-temporal integrated blob strength, the second one exploits stable image regions [15], and the third algorithm is based on tracking salient image points [16].

The detection layer is able to run in parallel several of those detection algorithms on different search areas in an image. A global scene data structure has been designed to collect results generated from the detection layer. Detection results are transformed from image coordinates to a scene frame given by horizon line and initial point of the compass. The transformation is based on localization of the horizon line, inertial measurements, and compass data. Transforming from image to scene coordinates amounts to a compensation of camera motion. The scene data structure is thus able to provide a motion-stabilized interface to the detection results. Results of the detection layer are demonstrated in Fig.7. Fig.8 shows the stabilized scene coordinate system with a detected object.

The final and topmost software layer uses the generated detections to compute internal alarms or to perform tracking. In order to do this, results from the detection layer are fused and classifiers are used to separate relevant detections at boat/ships from false or irrelevant detections. The alarms sent to the remote control room are produced in the CCP based on the internal alarms received from the IEP.

Fusion of detection results, classification and alarm generation are described in [17].

The tracking algorithm uses the spatial and temporal information from the robust multi-algorithm detection layer. The combination and processing of both types of data in the scene data structure makes it possible to perform a robust multi-target tracking. Tracking results are shown in Fig.9 to Fig.14.

During the tracking process, the PTU will be repositioned, if necessary. The IEP sends tracking images periodically to CCP. They can be requested from remote control center over the XML-based communication channel.

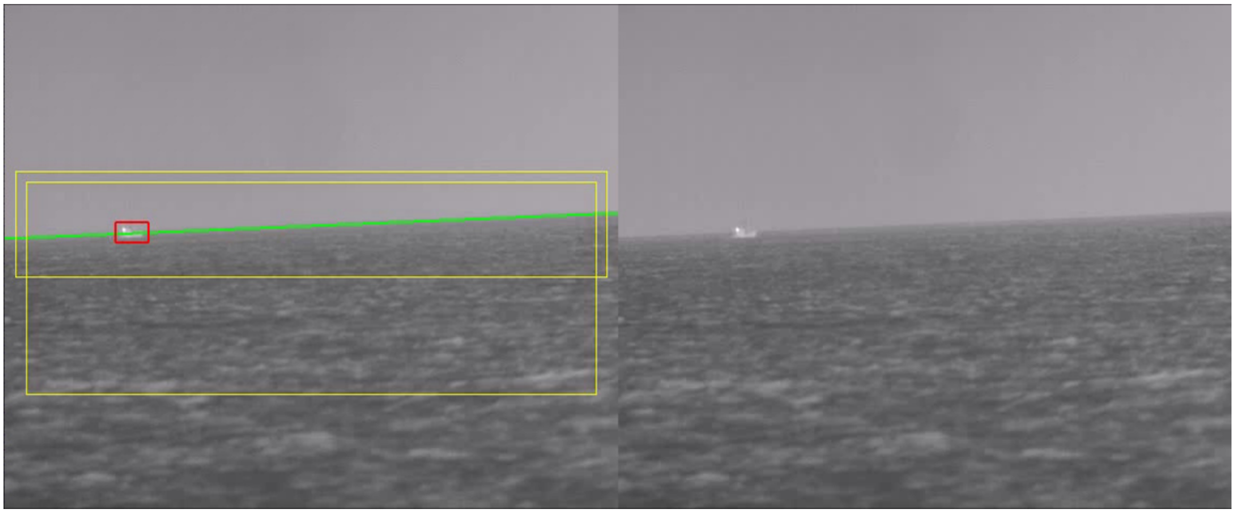

Fig. 7. Estimated horizon line (green) with two search areas (yellow) and detection result (red) for input image on the right.

IV. Results

The results demonstrated here were achieved using an extensive collection of image sequences captured by a thermal imager from land and ships in the North Sea. From the data, a number of relevant sequences were chosen and used for the development of algorithms.

Fig.5 and Fig.6 show results of the horizon detection layer. On the left side of each image, the detected horizon line is drawn. On the right side, the original image is shown.

It can be seen, that the horizon detection performs well under different visibility conditions. If the conditions are unfavorable, the horizon line can be detected using the information about the roll and pitch angle, as mentioned in Section III-C2.

Boat detection results are demonstrated in Fig.7. On the right, the source image is presented. On the left, we see the detected horizon line and two search areas for two detection algorithms running in parallel. A small object was detected and marked with a red bounding box.

The stabilized scene coordinate system with detection results and the corresponding original images are shown in Fig.8 The horizontal axis of the scene coordinate frame corresponds to the estimated horizon line in the original images.

Fig.9 to Fig.12 show a tracking sequence with two objects: a large vessel and a small one. Below the bounding boxes identifiers are drawn which correspond to the tracks observed by the tracker. We see, that the tracker is distinguishing and tracking both objects in parallel, although they are crossing and almost merging (Fig.10, 11)

Fig.13 and Fig.14 show a tracking sequence with a small object. Due to the alignment of the scene coordinate scene with the estimated horizon line, the system tracks the object although there is significant pitch movement in the pictures.

V. Conclusion

A vision-based maritime surveillance system has been developed and tested on video sequences captured with a thermal imager from shore and ships in the North Sea. The achieved image exploitation results for detection and tracking of small boats and ships are promising. The detection results have also been successfully used for classification and alarm generation [17]. Further field trials are planned and the software will be adapted based on the gained experience.

The XML/TCP/IP interfaces of the control and communication part of the software allows to use it in customer-specific environments, where the software in a remote control room only has to implement a system-specific XML-based protocol to communicate with the autonomous video surveillance platforms.

Acknowledgment

This work has been done in cooperation with Carl Zeiss Optronics GmbH and was supported with funds from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement No. SP1-Cooperation-218290.

References

[1] A. Ondini, G. Marsiglia, L. Fortunato, G. Balzarotti, *Techniques for detection of multiple, extended, and low contrast targets in infrared maritime scenarios*, Optical Engineering, Vol. 45, 2006.

[2] W. Tao, H. Jin, J. Liu, *Unified mean shift segmentation and graph region merging algorithm for infrared ship target segmentation*, Optical Engineering, Vol. 46, 2007.

[3] D. G. Lowe, *Distinctive image features from scale-invariant keypoints*, International Journal of Computer Vision, 60(2), pp. 91-110, 2004.

[4] C. Bibby, I. Reid, *Fast Feature Detection with a Graphics Processing Unit Implementation*, http://www.robots.ox.ac.uk/∼cbibby/pubs/paper994.pdf.

[5] S. P. van den Broek, E. J. Bakker, D. -J. de Lange, A. Theil, *Detection and classification of infrared decoys and small targets in a sea background*, Proc. SPIE Vol. 4029, Targets and Backgrounds VI: Characterization, Visualization, and the Detection Process, pp. 70-80, 2000.

[6] P. Voles, A. A. W. Smith, M. K. Teal, *Nautical Scene Segmentation Using Variable Size Image Windows and Feature Space Reclustering*, Lecture Notes in Computer Science, Volume 1843, ECCV 2000, pp. 324-335, 2000.

[7] Y. Wei, Z. Wei, H. Yu, *Wavelet analysis based detection algorithm for infrared image small target in background of sea and sky*, Proceedings of the 3rd International Symposium on Image and Signal Processing and Analysis, ISPA 2003, Vol. 1, pp. 23-28, 2003.

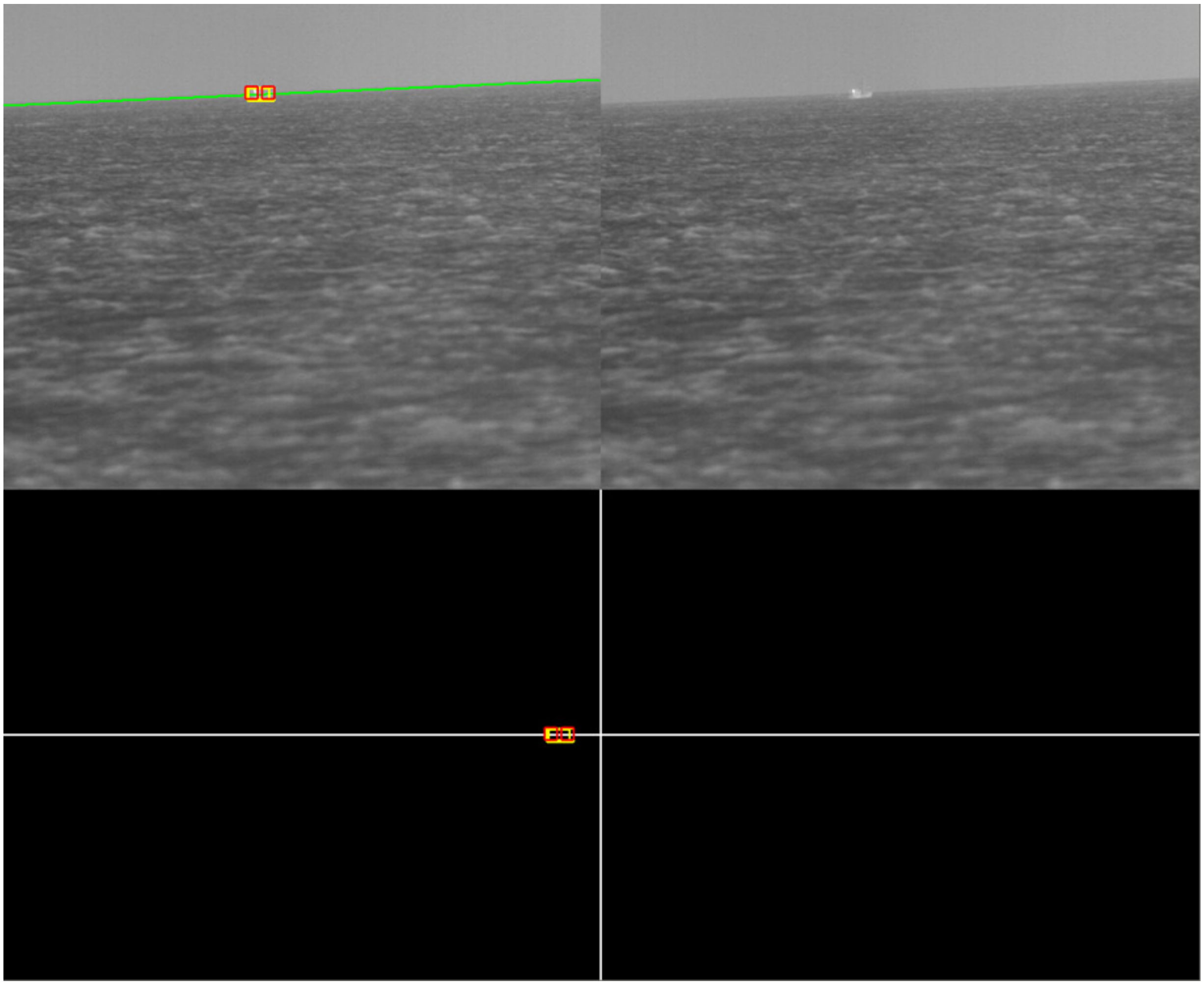

Fig. 8. Detection results (red, yellow) from different algorithms in an image (top, left) and corresponding scene data structure (bottom). The horizontal axis of the scene coordinate frame corresponds to the estimated horizon line (green, top) in the original images.

[8] S. Fefilatyev, D. B. Goldgof, L. Langebrake, *Towards detection of marine vehicles on horizon from buoy camera*, Proc. SPIE, Vol. 6736, Unmanned/Unattended Sensors and Sensor Networks IV, 2007.

[9] J. Canny, *A computational approach to edge detection*, IEEE Trans. Pattern Analysis and Machine Intelligence 8 (6), pp. 679-698, 1986.

[10] S. Fefilatyev, D. B. Goldgof, *Detection and tracking of marine vehicles in video*, 19th International Conference on Pattern Recognition, ICPR 2008, pp. 1-4, 2008.

[11] S. M. Ettinger, M. C. Nechyba, P. G. Ifju, M. Waszak, *Vision-guided flight stability and control for micro air vehicles*, IEEE/RSJ International Conference on Intelligent Robots and Systems 2002, Volume 3, pp. 2134-2140, 2002.

[12] T. G. McGee, R. Sengupta, K. Hedrick, *Obstacle Detection for Small Autonomous Aircraft Using Sky Segmentation*, Proceedings of the 2005 IEEE International Conference on Robotics and Automation, ICRA 2005, 18-22 April 2005, pp. 4679-4684, 2005.

[13] S. Fefilatyev, D. B. Goldgof, C. Lembke, *Autonomous Buoy Platform for Low-Cost Visual Maritime Surveillance: Design and Initial Deployment*, Proc. SPIE, Vol. 7317, Ocean Sensing and Monitoring, 2009.

[14] L. D. Stone, C. A. Barlow, T. L. Corwin, *Bayesian multiple target tracking*, Artech House, 1999.

[15] J. Matas, O. Chum, M. Urban, T. Pajdla, *Robust Wide Baseline Stereo from Maximally Stable Extremal Regions*, Proc. British Machine Vision Conference BMVC-90, Sept. 24-27, 1990, Oxford, England, pp. 384-393, 2002.

[16] J. Shi, C. Tomasi, *Good Features to Track*, Proc. Conference on Computer Vision and Pattern Recognition CVPR ’94, June 21-23, Seattle, Washington, pp. 593-600, 1994.

[17] M. Teutsch, W. Krüger, *Classification of small Boats in Infrared Images for maritime Surveillance*, Proceedings of the 2010 NURC Waterside Security (WSS), November 3-5, Carrara, Italy, 2010.

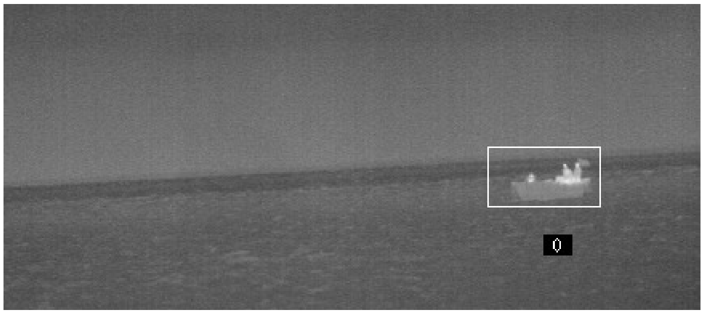

Fig. 9. Tracking sequence 1, frame number 55: Below the bounding boxes identifiers are drawn which correspond to the tracks observed by the tracker.

Fig. 10. Tracking sequence 1, frame number 428.

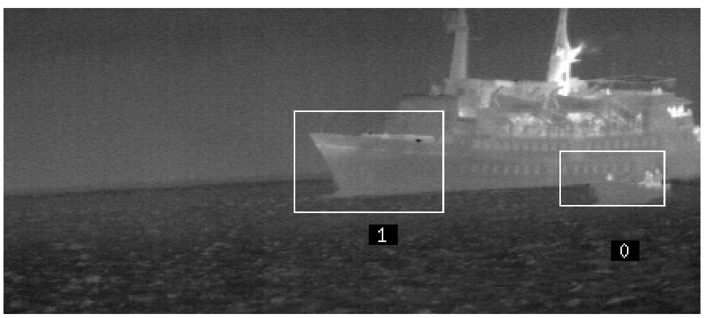

Fig. 11. Tracking sequence 1, frame number 569.

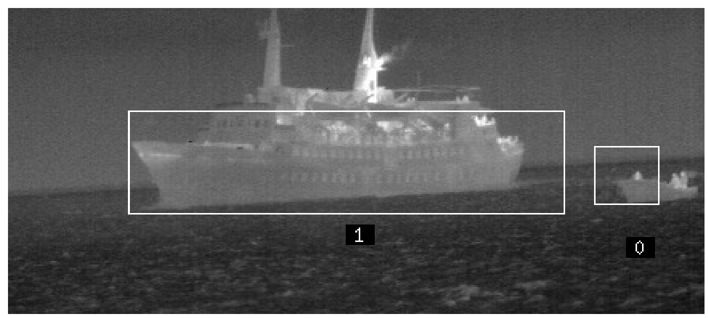

Fig. 12. Tracking sequence 1, frame number 721.

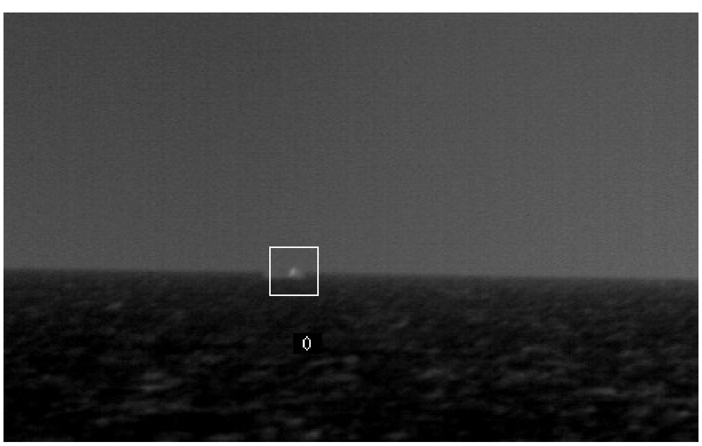

Fig. 13. Tracking sequence 2, frame number 50.

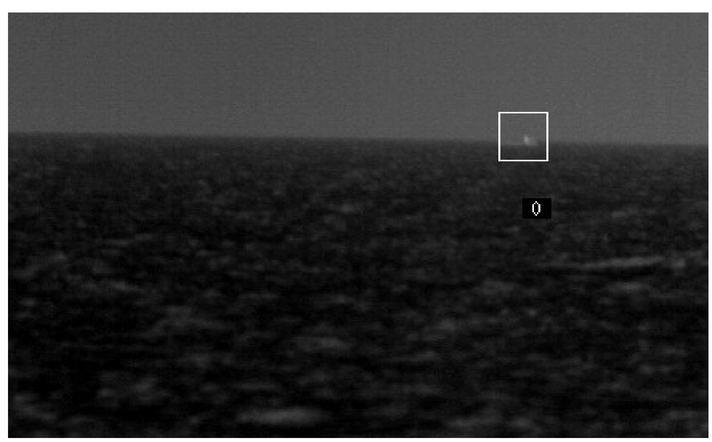

Fig. 14. Tracking sequence 2, frame number 625.