

Results of Optical Detection Trials in Harbour Environment

Rob. A.W. Kemp¹, Johannes F. de Groot², Sebastiaan P. van den Broek¹, Dirk-Jan J. de Lange¹,
Judith Dijk¹ and Piet B. W. Schwering¹,

¹TNO Defence, Security and Safety, P.O. Box 96864, NL-2509 JG The Hague, Netherlands, rob.kemp@tno.nl

²Thales Naval, P.O. Box 42, NL-7550 GD Hengelo, The Netherlands, hans.degroot@nl.thalesgroup.com

ABSTRACT

In harbour environments operators should perform tasks as detection and classification. Present-day threats of small objects, as jet skis etc, should be detected, classified and recognized. Furthermore threat intention should be analysed. As harbour environments contain several hiding spaces, due to fixed and floating neutral objects, correct assessment of the threats is complicated when detection tracks are intermittently known. For this purpose we have analysed the capability of our image enhancement and detection technology to assess the performance of the algorithms in a harbour environment. Data were recorded in a warm harbour location. During these trials several small surfaces targets were used, that were equipped with ground truth equipment. In these environments short-range detection is mandatory, followed by immediate classification. Results of image enhancement and detection are shown. An analysis was made into the performance assessment of the detection algorithms.

Keywords: Electro-optical systems, system concepts, video enhancement, detection.

1. INTRODUCTION

Apart from the naval scenario, algorithms have been developed for detection of threats in harbour scenarios. For detection the algorithms make use of different kinds of target features, such as hot spots, motion, and structure. Algorithms for adaptive sensor management are being tested using live recordings from naval operations. Incorporating these algorithms into new types of Electro-optical (EO) system concepts, based on infrared and visible light systems as well as integrated systems, is foreseen. At TNO a research program for the Dutch Ministry of Defence is examining several aspects of the use of electro-optical systems, from (near)-future system concepts to the image processing chain.

Thales Naval is under contract by the Dutch Ministry of Defence to demonstrate the feasibility of an electro-optical Security System. Under this contract a Sensor Head and processing were developed.

A generic overview of this processing chain can be found in Figure 1. Sensor data processing, as laid out in this paper provides a vast contribution to the Common Operational Picture.

In this paper we present first results from a trial, held in the Rotterdam Harbour. Further more, the advantages of the possible integration of the described sensor and image enhancement are discussed.

2. SYSTEM AND ALGORITHM APPROACH

In general, a modern electro optics (EO) observation system can be represented with the block diagram in Figure 1. The sensor used in this paper has 3 High-resolution TV cameras and 3 uncooled 8-14 μm IR cameras.

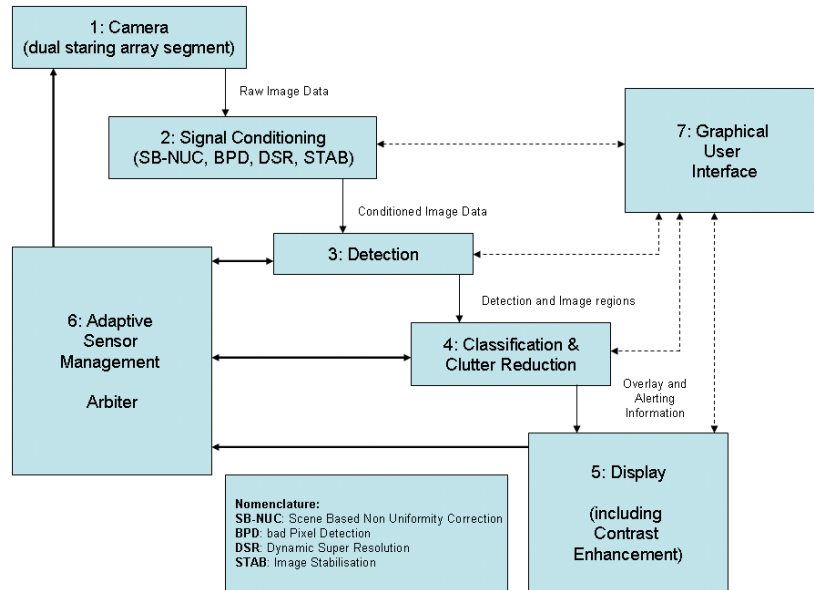


Figure 1 Block schematic of sensor image

The measurements are described in Section 3, while the sensor and processing are described in Section 4 and 5. We have applied Image Enhancement on the sensor data, the results can be found in Section 6. Results on classification and clutter reduction are discussed in Section 7.

Concerning Figure 1, the sensor and processing are discussed in section 4 and 5, signal conditioning (blocks 2 and 5 in Figure 1) is discussed in Section 6. Block 4 is discussed in Section 7. Blocks 3, 6, and 7 are not discussed in this paper.

When the sensor processing is organized in this way, processing can be performed as a pipeline.

3. MEASUREMENTS

Live recordings were used for developing, testing and tuning of algorithms for new types of Electro-optical (EO) systems, based on infrared and visible light. For this purpose recordings were made at 2007 April 26, at the Harbour of Rotterdam. The equipment was situated at the "Kop van Katendrecht", on which the sensor was located.

Use has been made of the following targets (see Figure 2):

- Water taxi
- RHIB from "Korps Commando Troepen" (KCT, Special Forces)
- Jet ski from (KCT)
- Targets of opportunity



Figure 2 Water taxi, RHIB and Jet Ski

During the trials the RHIB was equipped with AIS and GPS, used for ground truth reference data. The full data rate of the uncompressed raw digital camera data has been recorded for both 3x TV as well as 3x IR, covering a total of 120 degrees in azimuth. Figure 3 shows a Google Earth Map with plotted ground truth GPS/AIS data.

For all targets the following scenarios were performed:

- Straight Incoming with radial separation: at a speed of 40 km/hr.
- Straight Incoming with lateral (100m) separation: at a speed of 40 km/.
- Sailing circles (low speed) at various distances , for signature analyses at different aspect angles.
- Zigzag scenario with large amplitude, low frequency for analysing detection and tracking performance.
- Zigzag scenario with small amplitude, high frequency for analysing detection and tracking performance.
- Secret Passage, for analysing the influence of masking.
- Hiding behind large vessels, attacking at short distance, for analysing the influence of masking and reaction time.



Figure 3 Google Earth Map with plotted ground truth GPS/AIS data, red arrow indicates sensor position

4. DATA RECORDING AND PROCESSING EQUIPMENT

Sensor Head / Processing:

In Figure 4 the major hardware components of the Sensor, used during the recording, is illustrated. One sensor head covers 120 degrees in azimuth.

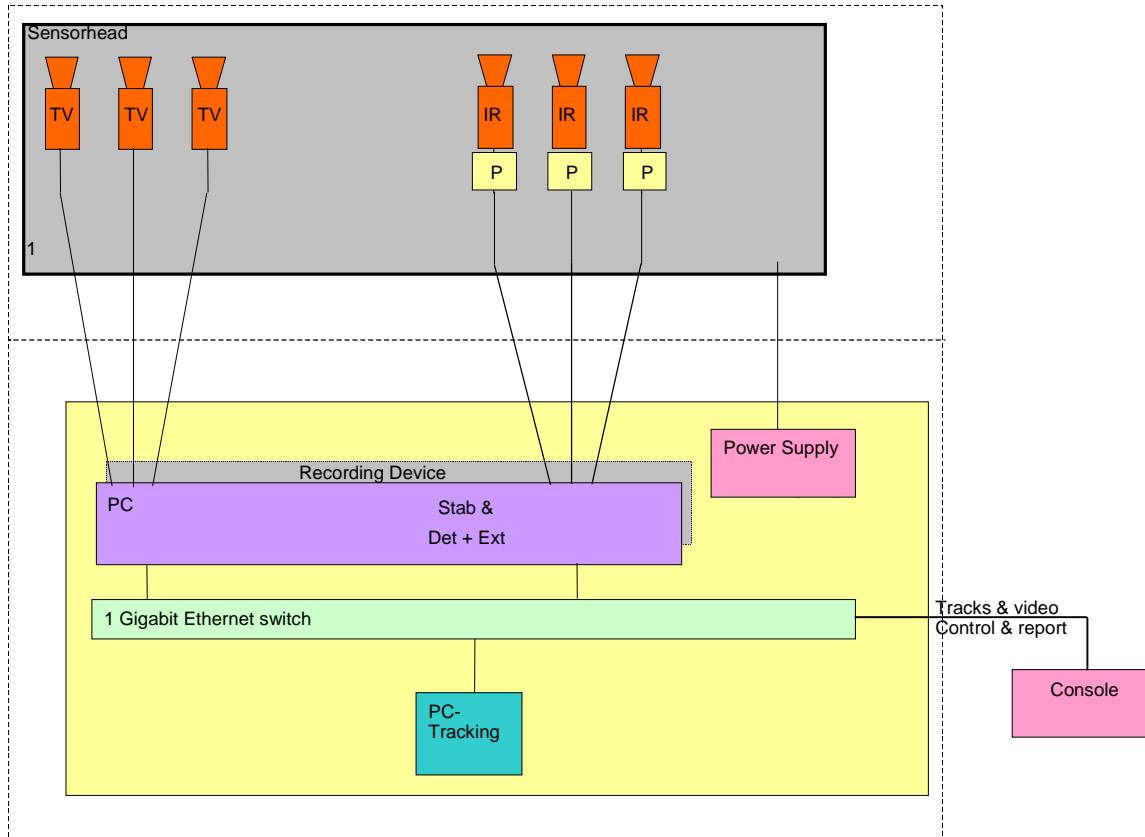


Figure 4 Recording system with one PC per sensor head and a central Tracking Computer

The Sensor head (Figure 5) used for the trial recordings did contain the following components:

- 3x High-resolution TV camera.
- 3x uncooled LW-IR camera.



Figure 5 The Sensor head used for the trial recordings

The Processing is built on high performance COTS PC. Mainly because of the high performance GPU, currently available, the complete image processing for one 120-degr unit is performed on one single PC.

This processing performs at the full frame rate and resolution of the 6 cameras and includes the following functionality:

- Camera correction TV/IR
- Stabilisation TV/IR
- Stitch processing TV/IR
- Detection/extraction IR/Tracking
- Data/Video distribution TV/IR for HMI

The console is capable of displaying aligned panoramic TV/IR views, with track overlay. The Operator can select dedicated section views.

Data Recording:

All raw camera data can be recorded at the full frame rate. The recorded data can be replayed in real time in the real processing, which makes offline tuning of the system very efficient. Terabytes of representative data is recorded from the Rotterdam Harbour.

5. SENSOR DATA PROCESSING

This section discusses the processing, as done inside the sensor. Topics like signal processing, data processing and sensor calibration can be found in this section. Post processing, like image enhancement can be found in section 6.

TV sensor processing:

The TV cameras are equipped with a Bayer CCD. The sensor processing converts the Bayer coded image into RGB image. TOD measurements [ref 8] are used to optimise the parameters for this decoding.

IR sensor processing: The full dynamic range digital data of the IR cameras is processed. Flat Field correction is performed in the cameras by means of a periodic shutter process.

Stabilisation and Stitching for IR and TV: A generic approach enables the use of a wide blend of cameras. Each camera is fully defined by a unique set of intrinsic and extrinsic parameters. Those parameters are determined during the calibration test set-up, as shown in Figure 7. The parameters fully define the elevation/azimuth of each camera pixel, i.e. how the camera sees the world. On the other hand the processing requires a north horizontal stabilised view, with a certain resolution and field of view and update rate, which is not equal to camera output. Different applications require different view parameters. Different applications can be either a) to support the HMI (e.g. panoramic views for situation awareness, or Section views for classification and identification purposes) or b) To optimise automatic detection performance. In Section 6 Image Enhancement Algorithms are described.

The following processing steps are required for generating a stabilised view:

- The stabilisation process makes use of INS data.
- The views are spherical projections of the camera data.
- The stitch process assures seamless transitions in the camera overlap areas, by adjusting inter camera gain, offset and colour differences (for example see Figure 6).
- The views (only for HMI) shall be enhanced in a way, which is dependent of the application and the role of the operator.



Figure 6 Example of blending and stitching for the IR cameras of one camera unit

IR Signal Processing Stages:

IR detection processing is performed on the full panoramic North horizontal stabilised view at the camera resolution and frame rate

Adaptive Kalman filtering: In case of a stationary sensor suite, this Kalman filter pre-integrates the IR data. This pre-integration results in increased sensitivity and enables motion detection. The adaptive parameters are a function of the measured temporal statistics of the scene. In case of a moving sensor, the integration time constant is a function of the relative motion in the image. The use of this filtering for motion detection is of course less effective on a moving platform.

Extended/Point Target Detection: A two-stage detection is performed on the pre-integrated data, firstly the large objects are extracted from the image, secondly the smaller extended and point target objects are detected without background disturbances caused by the large objects.

Clustering and Feature extraction and data reduction: Multi-frame Detections are clustered to plots, shape and detection features are preserved in the plot information and are later used in the track classification.

IR Data Processing:

Multiple Hypotheses Tracking (MHT): is used to find the best associations of plots to tracks.

Track Classification: Using spatial and temporal features to classify tracks. Because of the high update rate and relatively long reaction time for this kind of application, the False Alarm Rate reduction for this processing stage is very high.

Camera and Sensor Calibration method:

In the Sensor Unit: cameras are relative coarsely mechanically aligned assuring sufficient overlap between cameras.

All individual fine alignment parameters are determined with a calibrated test set-up with a target for IR as well as TV. Figure 7 shows the Calibration test set-up. The positions of the crossing wires are measured very accurately. From the observed camera image and the ground truth data all required parameters could be calculated. Applying all parameters in the processing assures pixel accurate alignment in overlap areas. Figure 8 illustrates the projection of the ground truth data on the observed IR camera image, after applying the inverse transformation with the estimated parameters.



Figure 7 Test Set-Up for Camera Calibration

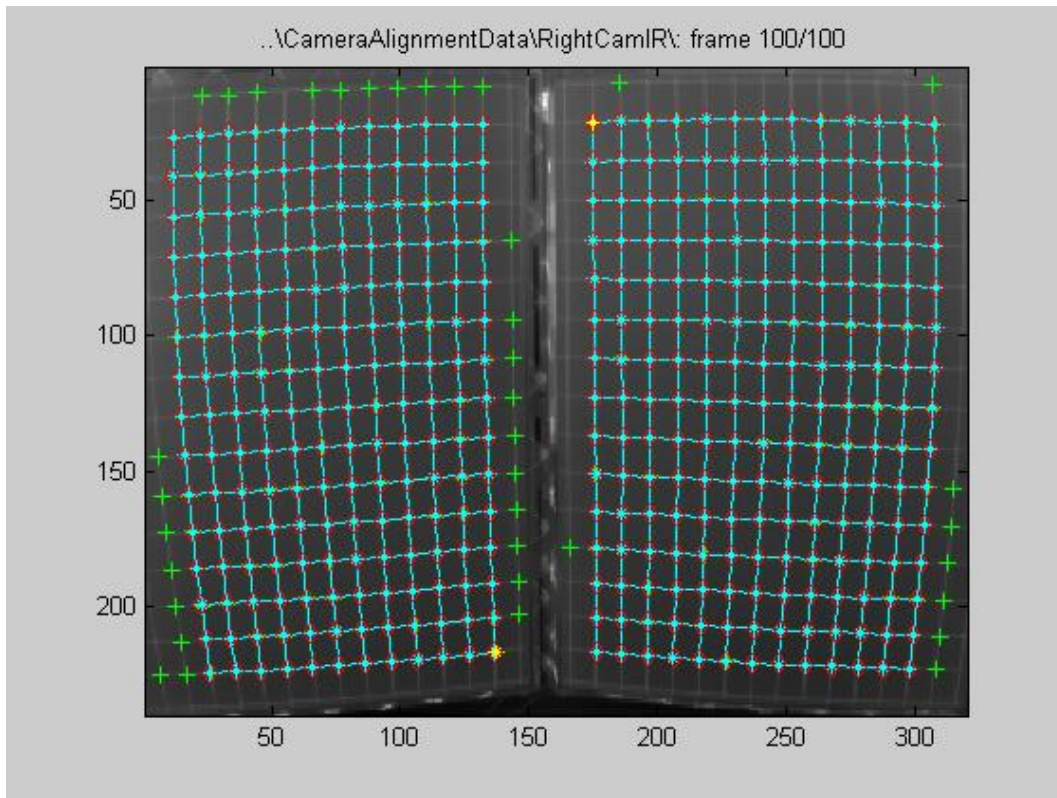


Figure 8 The projection of the ground truth data on the observed IR camera image

6. IMAGE ENHANCEMENT

As a way of post processing, image Enhancement techniques can be applied. Using Image Enhancement techniques will significantly increase the level of details in the imagery and therefore increase detection ranges. Three kinds of techniques are described here:

- Noise Reduction
- Resolution enhancement
- Local Adaptive Contrast Enhancement (LACE)

Noise reduction and Resolution Enhancement

The basic mechanism behind Resolution Enhancement is that temporal resolution of image sequences is exchanged for spatial resolution while improving the signal-to-noise ratio. This means that Noise Reduction and Resolution Enhancement are combined into a single algorithm. In Dynamic Super Resolution (DSR) several recordings of the same scene are combined in a High-Resolution Image. The differences between the recordings are used to calculate a new, improved, High-Resolution image. For the algorithm to be effective, the low resolution imagery needs to be registered with sub-pixel precision. Figure 9 gives a visualization of the mechanism behind DSR. The algorithm will increase the resolution by a factor of typically 2-4. This means that for a resolution enhancement factor of 2, there will be $2 \times 2 = 4$ High-Resolution pixels for each Low-Resolution camera pixel.

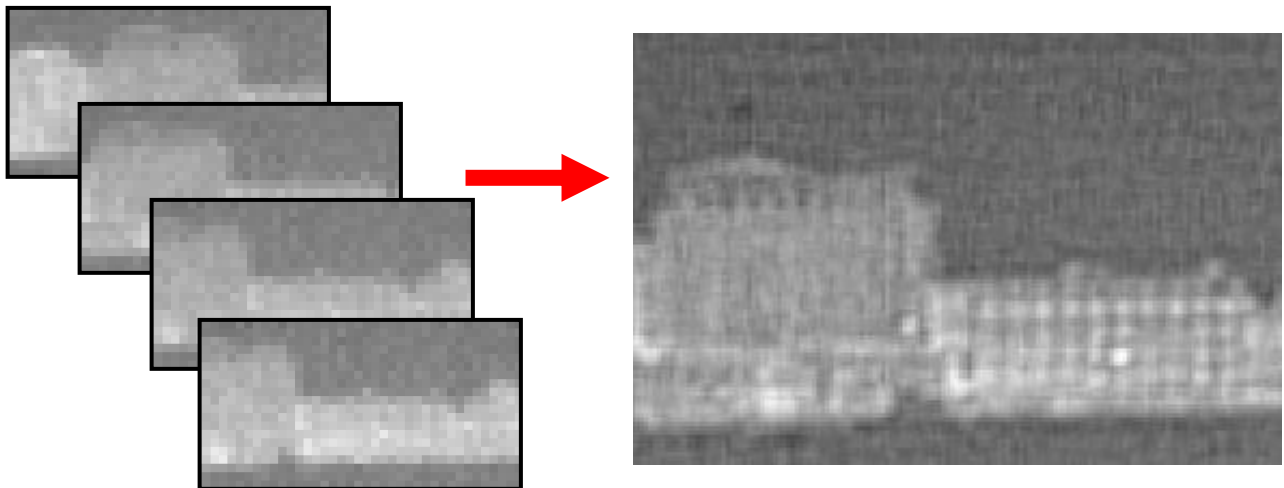


Figure 9, IR recordings of a beach Hotel from a distance of about 15km. Example of Resolution and Sensitivity Enhancement, where temporal resolution is exchanged for spatial resolution. The resolution enhancement factor was 4

Local Adaptive Contrast Enhancement:

Local Adaptive Contrast Enhancement (LACE) [ref 10] is used only for displaying. It is effective way to map the whole dynamic range of the camera (typical 8-14 bits) on a display, while visualising as much details as possible LACE will allow seeing details in light and dark areas without the need for dedicated gain and offsetting controls. LACE requires two input parameters, the input noise and the level of enhancement. The noise level in the input image can be supplied by the Noise Reduction algorithm. The level of enhancement depends on the task of the operator. For Situational Awareness a mild version of LACE is needed. For this task the large details are more important than the small scale details. For a detection task a medium version of LACE is needed, which enhances more small details, at the risk of reduce the general scene understanding.

Mild LACE will result in a natural image, while medium LACE will be less natural with lots of details

Results:

Image enhancement has been applied on the raw sensor data from Figure 5. Figure 10 show the raw image data of image 50 from the sequence (left hand side) and the same image, after applying a mild version Local Adaptive Contrast Enhancement (LACE).



Figure 10, Original image (left), as acquired from the sensor, after application of mild LACE (right)

The application of LACE clearly shows more details, even reflections of the building in the water. For this input resolution LACE runs real-time of a Graphics Processing Unit (GPU). Both images in Figure 10 have a resolution of 320x240 pixels.

When we apply Dynamic Super Resolution (DSR) on the input image, we use the sensor motion to obtain more detail in a reconstructed set image images. Sensor motion in this case was a manually induced ship-motion.

Figure 11 shows a detail of the enlarged version of the input image (image 50, magnified to the same size of the DSR image) and the same area of the image, after applying DSR (factor 2) and Contrast Enhancement.



Figure 11, Original Image (left) Dynamic Super resolution with a resolution enhancement factor of 2 applied, including a mild version of Local Adaptive Contrast Enhancement (right)

Figure 11 clearly shows the advantage of the combination of different image enhancement techniques. Dynamic Super Resolution was achieved solely by using image information. The total size of the DSR image is 640x480 pixels. Being acquired with a camera of 320x240 pixels, the image in Figure 11 looks similar to an image, acquired with a camera of 640x480 pixels.

Since a super resolved images may even contain more detail, we also applied DSR4 to the input sequences. Results of the application of Dynamic Super Resolution with a factor 4 (DSR4) are shown in Figure 12. For comparison, the input image (image 180 from the sequence) has been magnified to the same size as the DSR image and a mild version of LACE was applied.

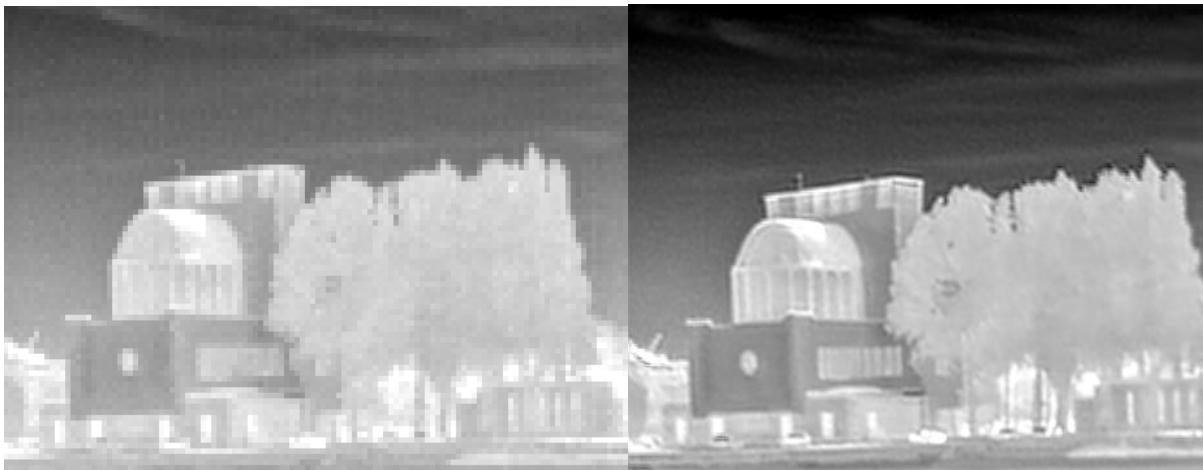


Figure 12, Detail from original image (left, mild LACE applied) and DSR4 with mild LACE applied (right)

Being acquired with a camera of 320x240 pixels, the image in Figure 12 looks similar to an image, acquired with a camera of 1280x960 pixels. Remarkable fact is that in the super resolved image details, hardly visible in the original image, are clearly visible in the super resolved image. Because of this, target identification can be done faster.

In addition, we have applied a medium version of LACE, Figure 13 shows these results. Again, we have used image 180 from the recorded sequence.

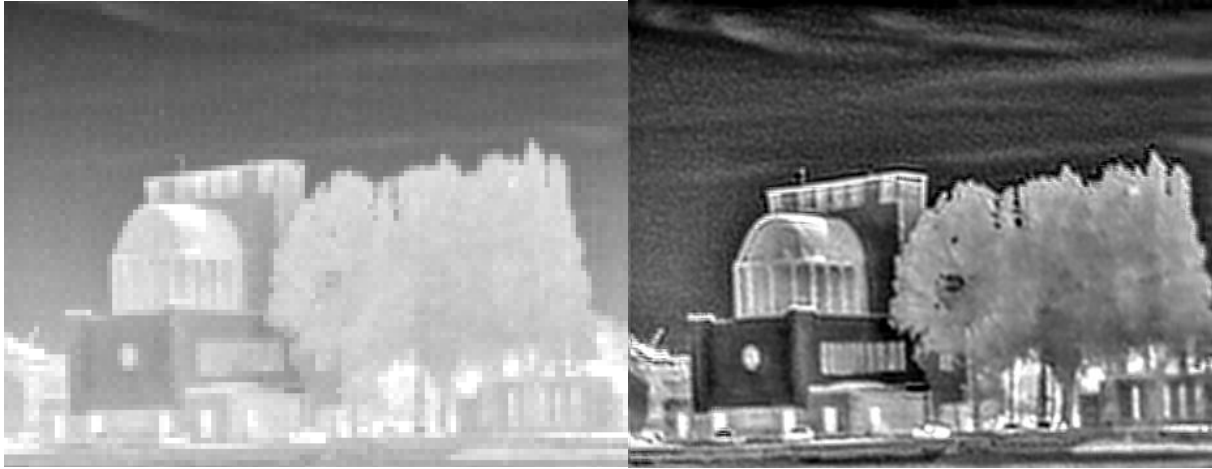


Figure 13, Detail from original image (left), mild LACE applied and DSR4 with medium version of LACE applied (right)

Though making the picture looking somewhat more artificial, medium LACE reveals even more details. Note for example the cars in front of the building in the right-hand picture of Figure 13, they were not recognizable in the left hand picture.

7. CLASSIFICATION AND CLUTTER REDUCTION

Part of the TNO research in the program “Electro-optical systems” is on extracting information from detected objects, in a process towards classification as target of interest. This includes what information can be extracted, and how this can be used to discriminate between different classes, at different levels (such as object or clutter, big or small, cabin boat or RHIB). This work relates to the steps between detection and showing object information to an operator. The information can assist an operator in the classification when there are many targets, such as in a harbour scenario. The ongoing work is described in [ref 9], here we only discuss the basics. This includes processing of data recorded with the TNO cameras in the Rotterdam trial, which was used to examine what general information about size, shape or type can be obtained from simple features, and if this is possible when the target is only few pixels high. The latter occurs for example when targets are at larger distance, and a system more aimed at detection, with corresponding smaller instantaneous field of view, provides the image of the target. An example of such a system is the sensor described in Section 4 of this paper.

Features that were examined are simple ones (e.g., based in dimensions) and moments (which describe more of the shape). Both are influenced by target distance and aspect angle. The results for simple features and moments show that it is possible to use them to discriminate between the three targets, even when their segmentation is only a few pixels high. A single feature will not be enough, but a combination of features allows distinguishing between targets in many cases. An example of this is shown in Figure 14, where two moments are plotted for the side views of the targets.

An important fact is that most features depend on aspect angle. Where distance dependencies can partly be corrected for and introduce mostly inaccuracies, different aspect angles may result in entirely different values. It is clear that in order to be able to distinguish between targets, having an estimate for the aspect angle is a necessity. Such an estimate may be obtained from track information from other sensors (such as radar).

It should be noted that since only three targets are examined, it is not known how much other small boats are similar to one of these, or quite different. However, it does show that discrimination between classes of small boats (e.g., ‘similar to a water taxi’ and ‘similar to a RHIB’) is possible, even for segmentations that are only several pixels high. An extensive description can be found in [ref 9].

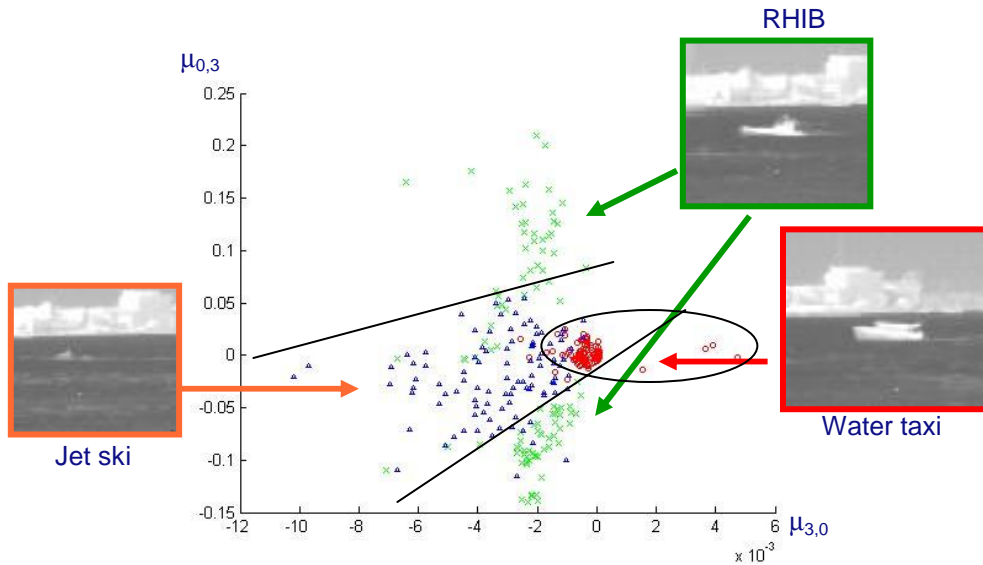


Figure 14, Two moments computed for three small targets. Areas can be defined so that most detections of the RHIB are outside the two lines, those for the jet-ski are between the lines, and for the water taxi inside the ellipse.

8. CONCLUSIONS

Thales Naval is under contract by the Dutch MOD to demonstrate the feasibility of an electro-optical Security System. Under this contract a Sensor Head and processing were developed. In a warm harbour scenario, Thales Naval and TNO carried out a recording experiment. Targets have sailed predefined patterns in a harbour environment and detection algorithms for those targets have been evaluated. Beside the overview of the recording and processing architecture, Thales Naval has demonstrated a camera calibration method, which enables perfect blending and stitching of adjacent and overlapping imagery.

Sensor image enhancement has been applied by TNO to the recorded data and results were presented in this paper. Image Enhancement, as used for the processing of imagery shown in this paper, can be performed real-time on a standard COTS PC. Image Enhancement provides a more detailed image to the observer, thus enabling detection and classification to be faster and earlier. Contribution to a more detailed common operational picture makes Image Enhancement a must for future sensors. Its flexibility for integration makes it a serious option for retrofitting into current electro-optical sensor systems.

Since stand-off distances can be larger, applying Image Enhancement on streams from uncooled camera enables the use of cheaper cameras for e.g. Special Forces or patrol boats in –for example- littoral scenarios.

9. ACKNOWLEDGEMENTS

The work for this paper was supported by the Netherlands MoD under the programme V602 “Electro-Optical sensor systems” and the work, as carried out by Thales Naval for the Pagode National Technology Project. The authors wish to thank for the support of the Netherlands MoD and specially the KCT for providing the crew, the RHIB and the Jet-Ski. The authors wish to thank the Pagode team of Thales and Peter Fritz of TNO for his work in preparation and execution of the measurement campaign.

10. REFERENCES

1. Schwering, P.B.W., van den Broek, S.P., van Iersel, M., *EO System Concepts in the Littoral*, SPIE Proc. Vol. 6542 paper 100, 2007
2. Schwering, P.B.W., *IRST evaluation methodologies: Maritime infrared background simulator*, SPIE Proc. Vol. 6206, 620621, 2006
3. Schutte, K., de Lange, D.J.J., van den Broek, S.P., *Signal conditioning algorithms for enhanced tactical sensor imagery*, SPIE Proc. Vol. 5076, 92-100, 2003.
4. van den Broek, S.P., Bakker, E.J., de Lange, D.J., Theil, A; *Detection and classification of infrared decoys and small targets in a sea background*, 2000, SPIE Vol. 4029, 70-80, 2000
5. Theil, A., Huizing, A.G., van Heijningen, A.W.P., *Single Picture Integration for Territorial Waters Surveillance (SPITS): An initiative to improve situational awareness in littoral waters*, MAST (Maritime Systems and Technology) Conference, Nice France, 4-6 September 2006
6. Arie N. de Jong, Alexander M.J. van Eijk, Marcel M. Moerman, Leo H. Cohen, *Investigation of aerosol particle size distribution in the San Diego Bay by means of multi-band transmissometry*, SPIE Proc. Vol. 6303, August 2006
7. Arie N. de Jong, Alexander M.J. van Eijk, Peter J. Fritz, Leo H. Cohen, Marcel M. Moerman, *The use of multi-band transmission data, collected at Scripps pier in November 2007, for the investigation of aerosol characteristics*, SPIE Proc. Vol. 6708, August 2007
8. Bijl, P., Valeton, J.M. (1998), *TOD, the alternative to MRTD and MRC*, Optical Engineering 37, 7, 1976 – 1983.
9. van den Broek, S.P., Bouma, H. and Degache, M.A.C., *“Discriminating small extended targets at sea from clutter and other classes of boats in infrared and visual light imagery, “*, Proc. SPIE Vol. 6969, to be published, (2008)
10. K. Schutte, *“Multi-Scale Adaptive Gain Control of IR Images”*, Proc. SPIE Vol. 3061, pp 906-914, 1997