

Detection of dim point targets in cluttered maritime backgrounds through multisensor image fusion

Alexander Toet*

TNO Human Factors, Kampweg 5, 3769 DE Soesterberg, The Netherlands

ABSTRACT

Multispectral IR imaging techniques are frequently deployed in maritime operations, for instance to detect floating mines or to find small dinghies and swimmers during search and rescue operations. However, maritime backgrounds usually contain a large amount of clutter that severely hampers the detection of dim point targets. Here we present a simple algorithm that deploys the correlation between target signatures in two different (3-5 and 8-12 μm) IR frequency bands to reduce the amount of clutter. First, both individual IR bands are filtered with a morphological opening top-hat transform to extract small details. Second, the resulting detail images are thresholded to produce binary detail images, representing potential target areas. Third, a fused detail image is obtained by taking the intersection (logical AND) of both binary IR detail images. Details that appear in both IR bands remain in this fused detail image, whereas a large fraction of uncorrelated noise details is filtered out. Remaining noise details can be removed by taking into account the temporal characteristics of the target signatures and by using a priori knowledge of structure of the scene and the size of potential targets. The method is tested on two image sequences showing a maritime scene with three kayaks approaching from far away. The scenario was registered in the 3-5 μm and 8-12 μm IR frequency bands, and in the visual range. The results show that the proposed multispectral processing technique has the potential to improve the detection of dim point targets in cluttered maritime backgrounds.

Keywords: Image fusion, infrared, detection, mathematical morphology

1. INTRODUCTION

This paper discusses the problem of small low observable (dim) target detection in cluttered maritime infrared images. This problem is of interest in many applications such as ocean surveillance (e.g. oil spills), search (e.g. speedboats, dinghies) and rescue (e.g. swimmers), remote sensing, floating mine detection, etc.

In the literature there are many useful algorithms for the detection and tracking of targets of significant size. However, in maritime infrared images, targets of interest are usually small and rather dim in a relative dark and cluttered sea surface background. Sea surface structure, reflection and emission changes related to incident angle variations and surface effects are standard features governing the clutter behavior¹⁰. Sun glint and horizon effects also contribute to clutter. The existence of scanline disturbance and noise further increases the difficulty in proper detection.

Non-linear filtering is a powerful technique to detect small dim targets^{2,4,8,9}. It has also been observed that co-registered multispectral imagery may result in improved target detection^{1,8,12}. As result, multispectral IR imaging techniques for target detection have recently received increased attention. Maritime surveillance systems consisting of combined visual and IR sensor systems have already been developed¹³.

The aim of the present study is to reduce the false alarm rate of a combined visual and IR surveillance camera system to an acceptable level. The camera system used here registers the 3-5 and 8-12 μm IR bands and the visual band. The combination of a morphological top-hat transform and a thresholding operation is used to detect potential target areas in the raw individual IR band images. The amount of clutter is reduced by taking the intersection of the potential target areas thus found. Finally, the outlines of the remaining potential target areas are projected onto the visual display. Experiments on a large set of sea going kayak images prove the effectiveness of the method presented in this paper.

* E-mail: toet@tm.tno.nl

2. IMAGE REGISTRATION

This section describes the experimental equipment, the location, the scenario, and the registration procedures that were deployed to collect the imagery that is used in the rest of this study.

2.1. Equipment

The following cameras were used to register the imagery that is used in the rest of this study:

- A Radiance HS IR camera (Raytheon), sensitive for 3-5 μm .
- An AIM 256 μLW camera (AEG), sensitive for 8-10 μm .
- A Philips LTC500 CCD camera (f/1.2 50dB s/n at 0.4 Lux) sensitive from visual to near IR (400-900 nm).

The Radiance HS IR focal plane array camera was used to register the 3-5 μm middle wavelength band of the infrared spectrum. The Radiance HS produces an image consisting of 256×256 pixels, with a FOV of 4.4×4.4 degrees², and an IFOV of 0.3 mrad. A lens with a focal length of 100 mm was used. The experimentally determined noise equivalent temperature difference (NETD) was 0.045 K.

The AIM 256 μLW focal plane array camera was used to register the 8-10 μm long wavelength band of the infrared spectrum. This camera produces an image consisting of 256×256 pixels, with a FOV of 5.9×5.9 degrees², and an IFOV of 0.4 mrad. A lens with a focal length of 100 mm was used. The experimentally determined noise equivalent differential temperature (NETD) of 0.033 K was little above specification (<0.025 K).

The Philips LTC500 CCD camera was used to register the visual part of the electromagnetic spectrum. This camera is equipped with a zoom lens and produces an image consisting of 752 x 582 pixels. In the narrow angle (tele) zoom setting the FOV is 2×1 degrees², and in the wide angle zoom setting the FOV is 14.7×7.4 degrees².

The field of view (FOV) of the cameras roughly matched, and was about 5×5 degrees² (varying from 4.5 to 6 degrees).

The cameras were rigidly mounted on a common base plate, that was placed on a pan and tilt unit, which was in turn supported by a tripod. Their optical axes were aligned and placed within 60 cm from each other. Since the viewing distances were relatively large, parallax can be neglected for all further purposes.

At the start of each recording session, the non-uniformity corrections (NUC) were determined for the IR cameras.

The camera outputs were continuously recorded using Panasonic AG-7700 SVHS video recorders. The resolution of the stored images is bandwidth limited by these videorecorders. However, video recording allows a full back-up and enables the comparison of dynamic scenes. The videorecorders were set-up in a master-slave configuration, providing each frame with a common time stamp. This enables digitisation of the videotapes with a temporal resolution of less than one frame length.

2.2. Location

The image recording session was held at the fourth floor of a hotel in Den Helder, The Netherlands. The hotel was about 20 m from the coastline, facing the sea. The cameras were placed outside on a balcony overlooking the sea. This condition was chosen to simulate the view from the bridge of a tall ship.

2.3. Environmental conditions

During the image registration period the sky was partly overcast (about 50% cloud cover). The temperature was 7.1 degrees Celsius (Fahrenheit). The wind was blowing from the south, with a windspeed of 9 m/s. Meteorological

visibility was about 7 km. Air pressure was 1020 hPa. These conditions resulted in a moderate amount of clutter (breaking waves) in the maritime environment.

2.4. Scenario

The scenario represented a part of the North Sea coastline, with three kayaks approaching the viewing location from far away. As a result, their corresponding image size varies from less than 1 pixel to almost the entire field of view. In this study we will only use images in which the kayaks are not larger than a few pixels.

2.5. Image warping

As mentioned before, the optical axes of all cameras were aligned such that the common area of their FOV was maximised. However, due to imperfections in the alignment procedure and differences in the optics of the cameras, an affine geometric transform is usually required to achieve one-to-one (pixel wise after digitisation) correspondence between the different image modalities. The parameters of such a transform can be determined from a set of fiducial points that are (a) well defined and distinctly represented in all image modalities, and (b) evenly distributed over the area of the common FOV. At the beginning of each recording session a well defined fiducial point was created by placing a large plastic jerry can on the beach near the coastline. The can was bright (white) and filled with hot water. As a result it was clearly visible in both the visual and infrared image modalities. It was located at such a distance in the scene that its image size was about 2x2 pixels. Images were registered with the jerry can at 9 different positions, evenly distributed over the common FOV of the cameras. This procedure resulted in 9 fiducial points that were later used to compute the parameters of an affine warping transform that maps the images to a common underlying reference grid.

3. IMAGE PROCESSING

This section first presents a brief introduction to mathematical morphology in general and the top-hat transform in particular. Then we describe how the top-hat transform can be employed to detect dim point targets in the individual 3-5 and 8-12 μm IR image bands. Finally, it is argued that uncorrelated noise details can be filtered out by taking the intersection of the potential target areas detected in each of the individual IR bands.

3.1. Mathematical morphology

In this section we briefly define the basic morphological transformations for discrete (e.g. sampled) functions. For an extensive introduction to mathematical morphology we refer to the literature¹¹.

In the sequel \mathbb{Z} will denote the set of integers and \mathbb{R} the set of real numbers. Discrete (i.e. sampled) images and structuring elements will be denoted by capital letters A, B, C, \dots , their domains and ranges by script letters $\mathcal{A}, \mathcal{B}, \mathcal{C}, \dots$. Let $F(x)$ be a function defined on a finite discrete d -dimensional domain $\mathcal{F} \subset \mathbb{Z}^d$ ($d=1,2,\dots$) and with a discrete and continuous amplitude range $\mathcal{R} \subset \mathbb{R}$ or \mathbb{Z} . A *structuring element* is any function $G(x)$ whose support \mathcal{G} is a compact and connected subset of \mathbb{Z}^d . In the sequel we will only use flat (binary or bivalued structuring elements $B(x)$, with $B(x)=0$ for $x \in \mathcal{B}$). The extension of our results to multivalued (grayscale) structuring elements is straightforward^{3,5,11}. The basic morphological operations on discrete functions with binary structuring elements are defined in Table I. To implement these operations we assume that $F(x)=-\infty$ for $x \notin \mathcal{B}$.

3.2. Top-hat transform

The opening transform removes bright details from an image that are smaller than the size of the structuring element. The residual image, obtained by subtracting the opened image from the original, comprises of only those image features that have been removed by the structuring element in the opening operation. This filter operation is called the *top-hat transform*⁶ and is defined as

$$(F - (F \circ B))(x) = (F - (F \ominus B) \oplus B)(x)$$

It provides an excellent tool for extracting bright features smaller than a given size from an uneven background.

Table I. Basic morphological transformations for discrete functions.

Name	Definition	$(x, y, p \in \mathbb{Z}^d, \mathcal{B} \subset \mathbb{Z}^d, d=1,2,\dots)$
set translation of \mathcal{B} by p	$\mathcal{B}_p = \{b+p : b \in \mathcal{B}\}$	
symmetric set of \mathcal{B}	$\tilde{\mathcal{B}} = \{-b : b \in \mathcal{B}\}$	
symmetric function of B	$\tilde{B}(x) = B(-x)$	
Minkowski addition of F and B	$(F \oplus B)(x) = \max\{F(y) : y \in \tilde{\mathcal{B}}_x\}$	
Minkowski subtraction of B from F	$(F \ominus B)(x) = \min\{F(y) : y \in \tilde{\mathcal{B}}_x\}$	
dilation of F by B	$(F \oplus \tilde{B})(x) = \max\{F(y) : y \in \mathcal{B}_x\}$	
erosion of F by B	$(F \ominus \tilde{B})(x) = \min\{F(y) : y \in \mathcal{B}_x\}$	
closing of F by B	$F \bullet B(x) = (F \oplus \tilde{B}) \ominus B(x)$	
opening of F by B	$F \circ B(x) = (F \ominus \tilde{B}) \oplus B(x)$	

3.3. Multiband detection

The top-hat transform can be employed to detect dim point targets in the individual 3-5 and 8-12 μm IR image bands. However, in maritime environments with a high degree of clutter the top-hat transform will yield many false alarms. It is a priori likely that small targets like boats, mines and persons, will be represented in both bands. In contrast, noise details will mostly be restricted to one particular band. Therefore, uncorrelated noise details may be filtered out by taking the intersection of the potential target areas detected in each of the individual IR bands.

4. RESULTS

Figures 1-7 show a set of typical maritime images, representing three kayaks (approximately in the middle of the scene) at high sea. The composition of each of these figures is identical. The upper left and right images correspond respectively to the original Radiance 3-5 μm and AIM 8-12 μm IR images. The middle row in each figure shows the potential target areas detected in the individual IR bands. These areas are obtained by thresholding the top-hat transform of the IR bands. The potential target areas are enlarged by dilation with a disk shaped structuring element of size 3 (radius 1). This is done to compensate for possible offsets in the registration between the individual IR bands. Dilation increases the chances that the potential target areas in the individual IR bands will have at least some degree of physical overlap. By taking the intersection (implemented as a logical AND on both binary IR target area images), a large amount of irrelevant noise details are eliminated and the potential target areas remain. The result of this operation is shown in the lower left image in Figures 1-7. Note that (a) the number of false alarms (see middle row of Figures 1-9) is significantly reduced, and (b) the correct targets (three kayaks in the middle of the scene) are detected each time. The lower right image in Figures 1-7 shows the enlarged outlines of the potential target areas projected over the corresponding visual CCD image.

5. CONCLUSIONS

The morphological top-hat transform can be used to detect dim point targets in cluttered maritime backgrounds registered in the 3-5 and 8-12 μm IR bands. The number of false alarms (noise details) can be reduced significantly by

taking the intersection of the potential target areas (alarms) in both bands. Thus, the proposed multispectral processing technique can improve the detection of dim point targets in cluttered maritime backgrounds. In the scenario used in this study, the dim target moves very slowly ($v < 0.5$ pixel/frame). The signal to noise ratio can therefore be improved (the definition of the target enhanced) by summation of successive frames. This will result in an increase in target energy and a decrease in clutter energy. Considering the low speed of the target, the summing operation may be implemented by directly adding N consecutive frames under the assumption that the point target stays at a pixel for at least N frames. The detection probability can be improved even further by the use of a priori knowledge. Tuning the structuring element used in the top-hat transform to the size and shape of the targets will significantly reduce the amount of false alarms⁷. Information on the distance and possible target locations may also increase the detection probability.

ACKNOWLEDGEMENTS

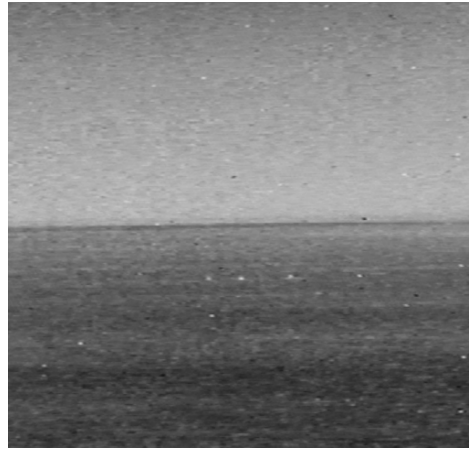
This material is based upon work supported by the European Office of Aerospace Research and Development, Air Force Office of Scientific Research, Air Force Research Laboratory, under contract No. F61775-01-WE026, and by Senter, Agency of the Ministry of Economic Affairs of the Netherlands.

REFERENCES

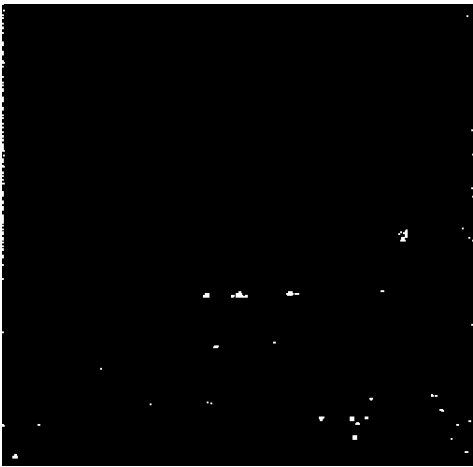
1. Abu-Tahnat, M. and Thompson, M.W., Frequency-band comparison using adaptive filter for multispectral imagery, In: A.C. Dubey, R.L. Barnard, C.J. Loe & J.E. McFee (Ed.), *Detection and remediation technologies for mines and minelike targets*, pp. 36-45, The International Society for Optical Engineering, Bellingham, WA, 1996.
2. Deshpande, S.D., Er, M.H., Ronda, V. and Chan, P., Max-mean and max-median filters for detection of small targets, In: O.E. Drummond (Ed.), *Signal and Data Processing of Small Targets 1999*, pp. 74-83, The International Society for Optical Engineering, Bellingham, WA, 1999.
3. Haralick, R.M., Sternberg, S.R. and Zhuang, X., Image analysis using mathematical morphology, *IEEE Transactions on Pattern Analysis and Machine Intelligence PAMI*, 9(4), pp. 532-550, 1987.
4. Kemper, P.J., Mathematical morphology enhancement of maximum entropy thresholding for small targets, In: D.P. Casasent & A.G. Tescher (Ed.), *Hybrid Image and Signal Processing VI*, pp. 84-91, The International Society for Optical Engineering, Bellingham, WA, 1998.
5. Maragos, P., Pattern spectrum of images and morphological shape-size complexity, In: *IEEE International Conference on Acoustics, Speech and Signal Processing ICASSP 87*, pp. 241-244, IEEE Computer Society Press, Washington, USA, 1987.
6. Meyer, F., Contrast feature extraction, In: J.L. Chermant (Ed.), *Quantitative analysis of microstructure in material sciences, biology and medicine*, Riederer Verlag, Stuttgart, GE, 1978.
7. Moon, V.-S., Zhang, T., Zuo, Z. and Zuo, Z., Detection of sea surface small targets in infrared images based on multilevel filter and minimum risk Bayes test, *International Journal of Pattern Recognition and Artificial Intelligence*, 14(7), pp. 907-918, 2000.
8. Perez-Jacome, J.E. and Madisetti, V.K., Target detection via combination of feature-based target-measure images, In: I. Kadar & V. Libby (Ed.), *Signal processing, sensor fusion, and target recognition VIII*, pp. 345-356, The International Society for Optical Engineering, Bellingham, WA, 1999.
9. Sang, N., Zhang, T. and Shi, W., Characteristics of contrast and application for small-target detection, In: O.E. Drummond (Ed.), *Signal and Data Processing of Small Targets 1998*, pp. 123-129, The International Society for Optical Engineering, Bellingham, WA, 1998.
10. Schwering, P.B., Maritime infrared background clutter, In: W.R. Watkins & D. Clement (Ed.), *Targets and backgrounds: characterization and representation II*, pp. 255-266, The International Society for Optical Engineering, Bellingham, WA, 1996.
11. Serra, J., *Image analysis and mathematical morphology*, Academic Press, New York, USA, 1982.
12. Singer, P.F. and Sasaki, D.M., Multispectral detection of dim slightly extended targets in heavy clutter, In: O.E. Drummond (Ed.), *Signal and Data Processing of Small Targets 2000*, pp. 96-103, The International Society for Optical Engineering, Bellingham, WA, 2000.
13. Yamamoto, K., Yamada, K. and Kiriya, N., System for maritime surveillance aid, In: S.G. Ackleson (Ed.), *Ocean Optics XIII*, pp. 815-820, The International Society for Optical Engineering, Bellingham, WA, 1997.



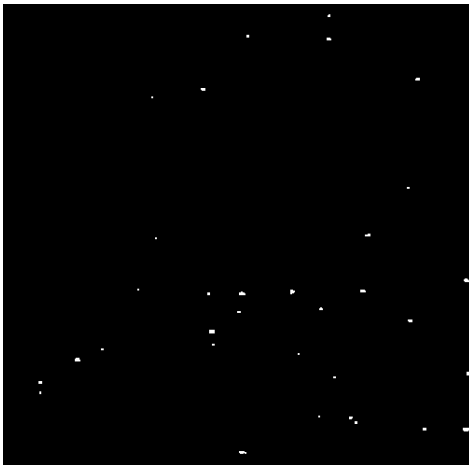
Radiance 3-5 μm



AIM 8-12 μm



Radiance Alarms



AIM Alarms



Multiband Alarms

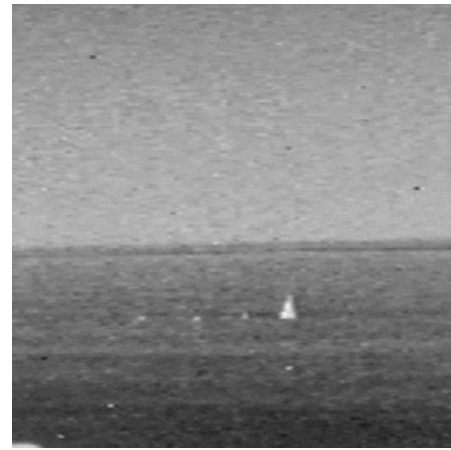


CCD image with potential target areas

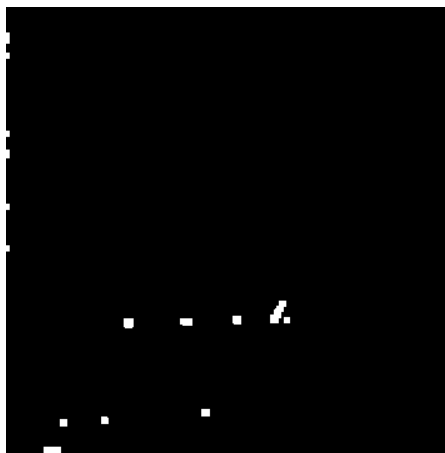
Figure 1 Upper row: corresponding original Radiance 3-5 μm and AIM 8-12 μm IR images. Middle row: potential targets detected in the individual IR bands. Lower row: potential targets detected in both IR bands (left) and target areas projected over the corresponding visual CCD image (right).



Radiance 3-5 μm



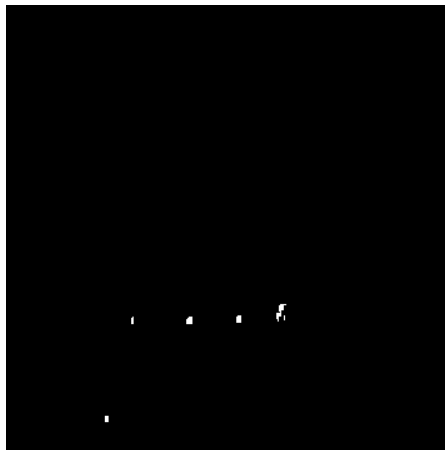
AIM 8-12 μm



Radiance Alarms



AIM Alarms



Multiband Alarms



CCD image with potential target areas

Figure 2 Upper row: corresponding original Radiance 3-5 μm and AIM 8-12 μm IR images. Middle row: potential targets detected in the individual IR bands. Lower row: potential targets detected in both IR bands (left) and target areas projected over the corresponding visual CCD image (right). The large triangular object on the right is a buoy.

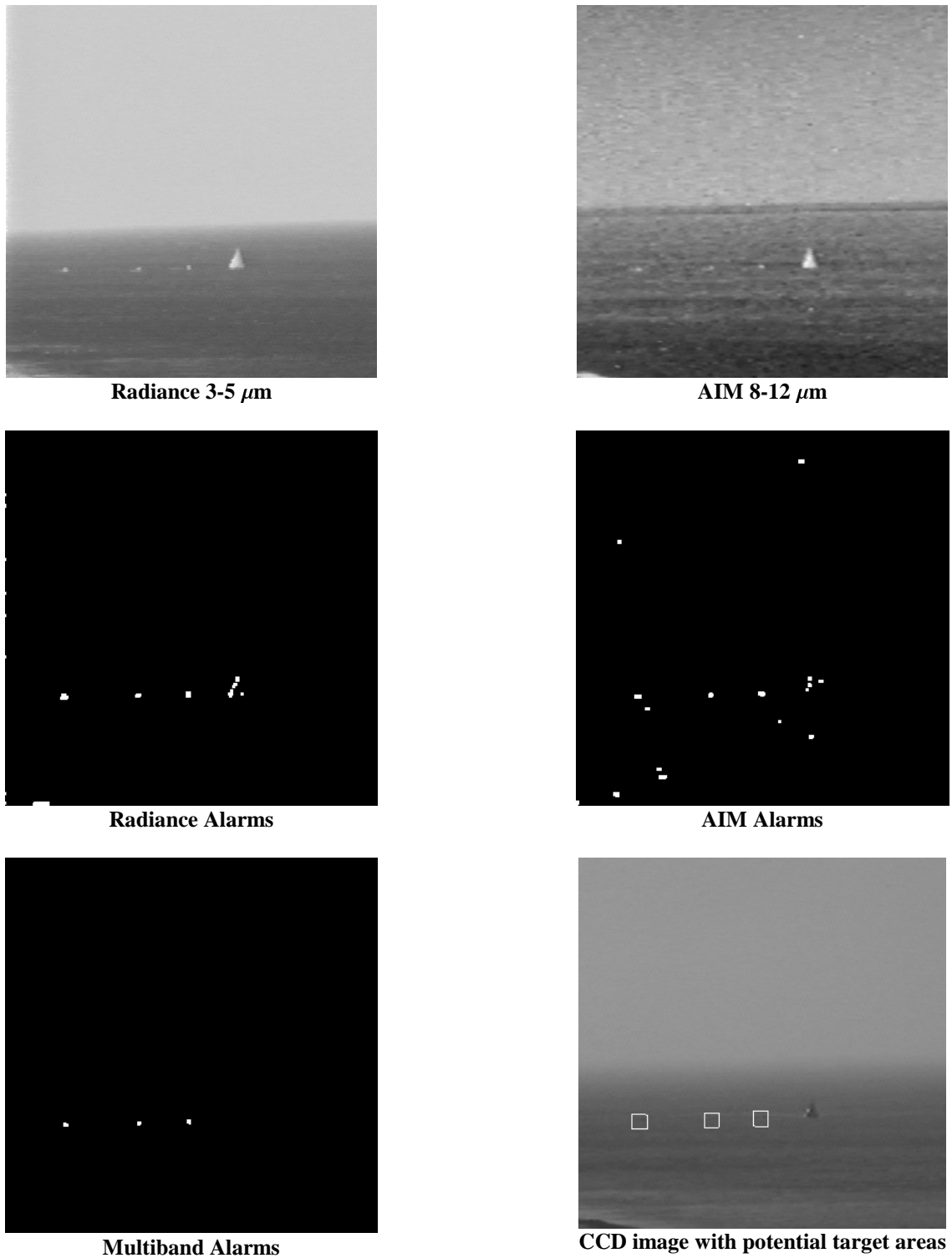


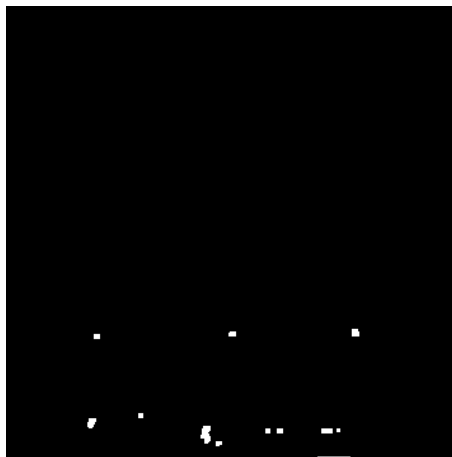
Figure 3 Upper row: corresponding original Radiance 3-5 μm and AIM 8-12 μm IR images. Middle row: potential targets detected in the individual IR bands. Lower row: potential targets detected in both IR bands (left) and target areas projected over the corresponding visual CCD image (right).



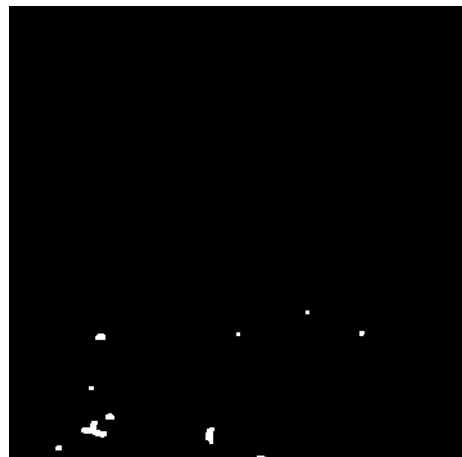
Radiance 3-5 μm



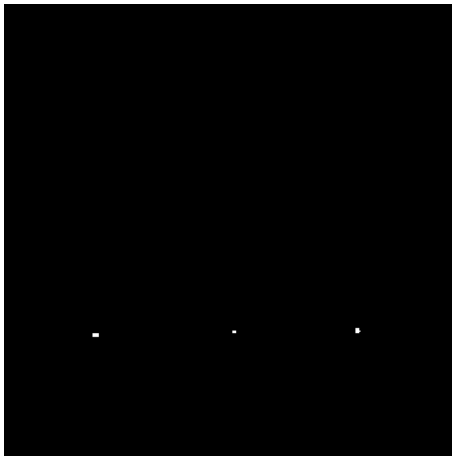
AIM 8-12 μm



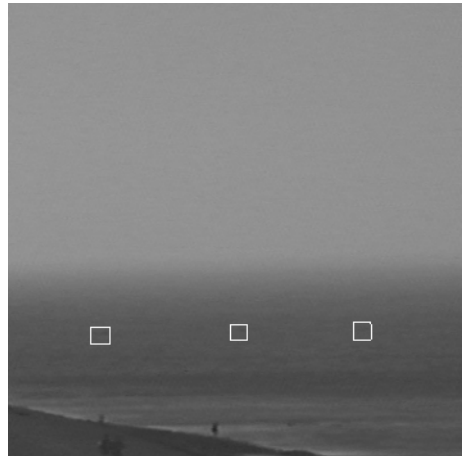
Radiance Alarms



AIM Alarms



Multiband Alarms

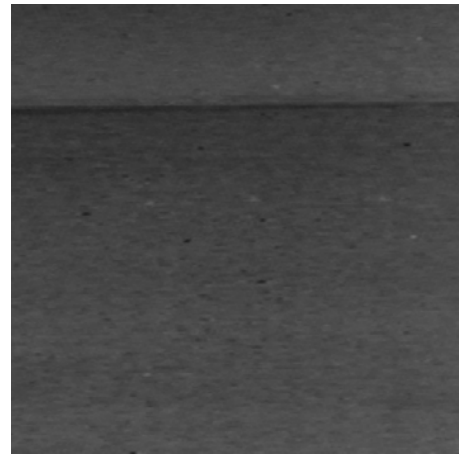


CCD image with potential target areas

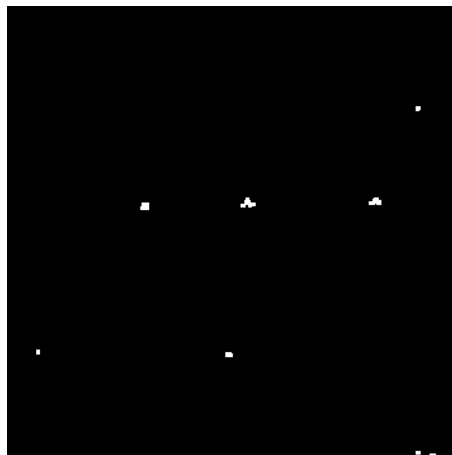
Figure 4 Upper row: corresponding original Radiance 3-5 μm and AIM 8-12 μm IR images. Middle row: potential targets detected in the individual IR bands. Lower row: potential targets detected in both IR bands (left) and target areas projected over the corresponding visual CCD image (right).



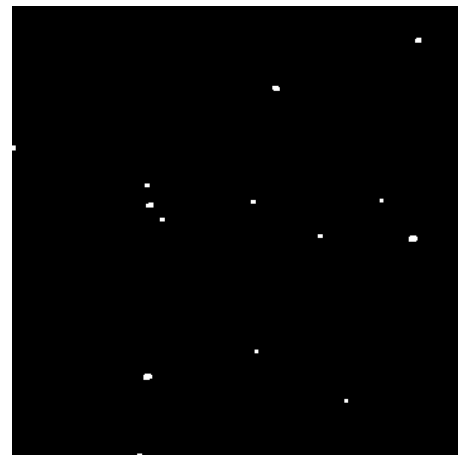
Radiance 3-5 μm



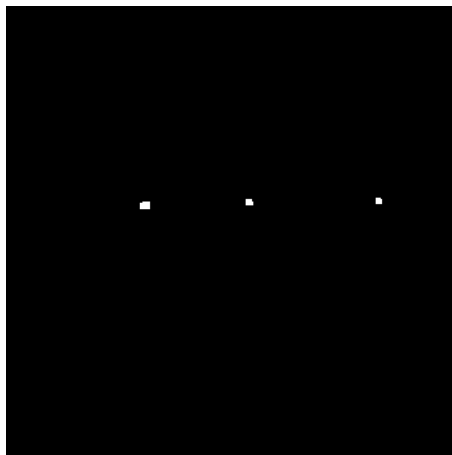
AIM 8-12 μm



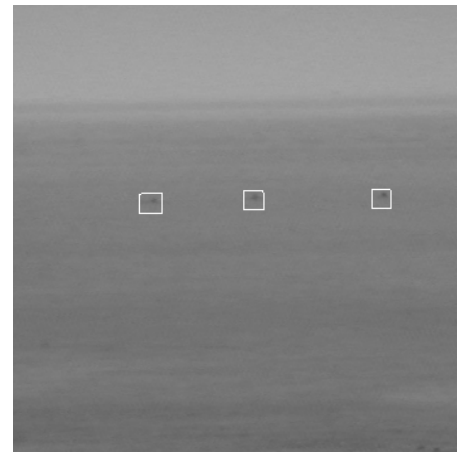
Radiance Alarms



AIM Alarms



Multiband Alarms



CCD image with potential target areas

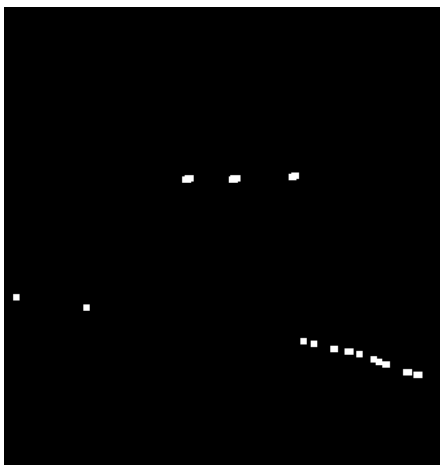
Figure 5 Upper row: corresponding original Radiance 3-5 μm and AIM 8-12 μm IR images. Middle row: potential targets detected in the individual IR bands. Lower row: potential targets detected in both IR bands (left) and target areas projected over the corresponding visual CCD image (right).



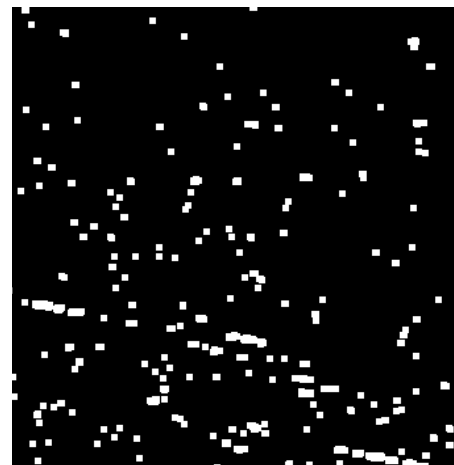
Radiance 3-5 μm



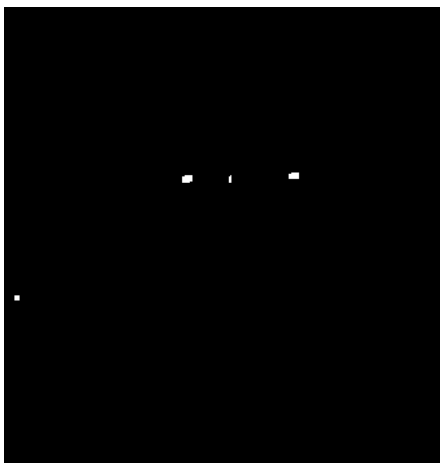
AIM 8-12 μm



Radiance Alarms



AIM Alarms



Multiband Alarms

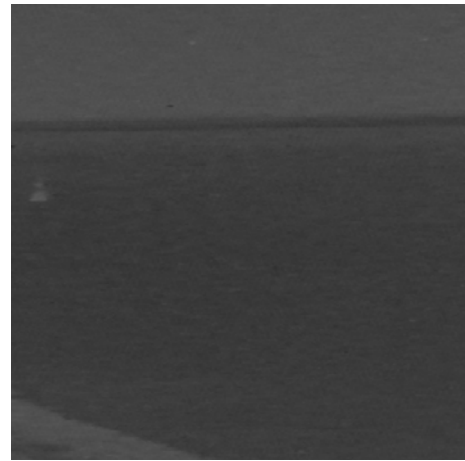


CCD image with potential target areas

Figure 6 Upper row: corresponding original Radiance 3-5 μm and AIM 8-12 μm IR images. Middle row: potential targets detected in the individual IR bands. Lower row: potential targets detected in both IR bands (left) and target areas projected over the corresponding visual CCD image (right).



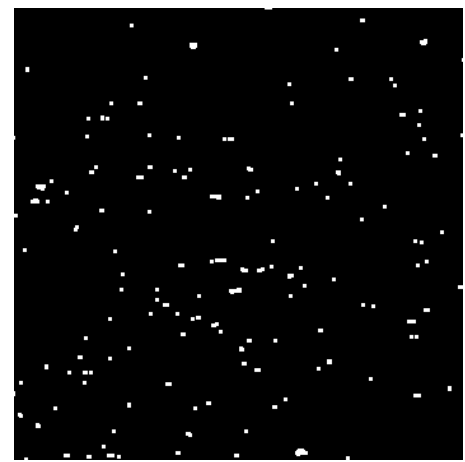
Radiance 3-5 μm



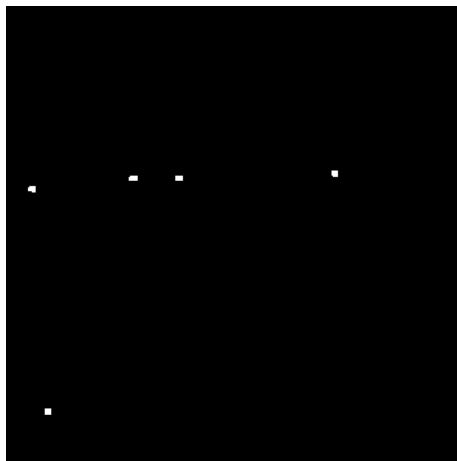
AIM 8-12 μm



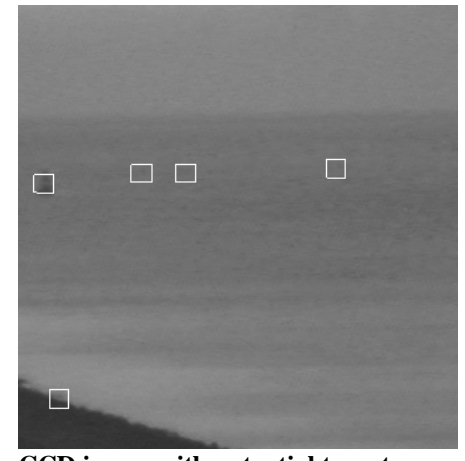
Radiance Alarms



AIM Alarms



Multiband Alarms



CCD image with potential target areas

Figure 7 Upper row: corresponding original Radiance 3-5 μm and AIM 8-12 μm IR images. Middle row: potential targets detected in the individual IR bands. Lower row: potential targets detected in both IR bands (left) and target areas projected over the corresponding visual CCD image (right).