Automatic detection of hostile behaviour

Bert van den Broek*, Gertjan Burghouts, Sebastiaan van den Broek, Arthur Smith, Ronald Hagen, Laura Anitori, Wim van Rossum
TNO Defence, Security and Safety, P.O. Box 96864, 2509 JG, The Hague, The Netherlands

ABSTRACT

In current military operations threats should be monitored accurately. The use of sensors is indispensable for this purpose, for example with camera and radar systems. Using data from such systems we have studied automated procedures for extracting observable behavioral features of persons and groups, which can be associated with threats. We have analysed algorithms for identifying animals versus humans, and for determining the activity of detected humans. Secondly, geospatial algorithms are studied to determine people in suspicious places.

Keywords: target recognition, sensor fusion, hostile behaviour

1. INTRODUCTION

For military out of area operations such as in Afghanistan situation awareness is one of the key issues in dealing with threats such as IEDs, mortar attacks and ambushes. For situational awareness, the area needs to be monitored continuously to get an overview of the scene. For this purpose sensors are indispensable. Cameras, FMCW radars and acoustic sensors are more and more deployed, since these are becoming compact and affordable. These sensors can be placed on the ground, attached to masts or placed on board of an aerostat in order to improve the line of sight. The sensor data need to be processed for obtaining information in order to assess the threats. Human interpretation of the data is usually quite effective, but is less suitable in case of simultaneous signals from multiple sources. Also, in general, the availability of analysts is limited and hence human interpretation is a bottleneck. To support and to limit the workload for the interpreters automated procedures should be used as much as possible.

TNO Defense, Security and Safety has started a programme on the topic Hostile Intent. Within this programme a study is defined which deals with automatic procedures for the detection of hostile behaviour from multi-source data. For the study, a scenario has been chosen for monitoring threats near compounds or operating bases in current out of area operations such as in Afghanistan. In the scenario, threats are identified such as deployment of IEDs, preparation for mortar attacks, ambushes and riots. For these threats indicators are determined which can be associated with observable behavioral features of persons and groups. An example is a person stopping for a while a few kilometers from a compound at night and possibly deploying mortar equipment. Another example is a group approaching an operating base in an unusual way, possibly intending an attack.

To study the feasibility for monitoring threats with sensor systems we have performed several measurements of human behaviour in a relevant setting using a wide- and narrow-field of view camera system and a stereo range-Doppler radar. On basis of the measured data we have studied several algorithms to extract behavioural features. Studied are automatic target recognition (ATR) algorithms for classifying detected objects in the data (i.e. are we able to determine human beings versus e.g. animals such as dogs, horses etc.) and for determining the activity of detected humans (walking, carrying objects, digging etc.). In addition, geospatial algorithms are studied to determine people in suspicious places. An analysis is made how the obtained behavioral features can be linked with the threat indicators in combination with scene and context information.

In the paper, we report results of the measurements and the analysis and draw conclusions about the feasibility for detecting hostile behaviour using an advanced radar and camera system.

*bert.vandenbroek@tno.nl; phone 31 70 3740430; fax 31 70 3740654; www.tno.nl

Unmanned/Unattended Sensors and Sensor Networks VI, edited by Edward M. Carapezza,
Proc. of SPIE Vol. 7480, 74800R · © 2009 SPIE · CCC code: 0277-786X/09/$18 · doi: 10.1117/12.829468

2. OBSERVING HOSTILE BEHAVIOUR

We first define here what we mean with hostile behaviour. We can consider activities in the world as composed of various scenarios. For example supporters going to a football game, go out, come together, have fun, watch the game, have fun and go home again. Such a scenario does not contain any unwanted activities. On the other hand if supporters come together before the game and start a fight this is something which is obviously not wanted and we describe the associated behavior as hostile. Observation and comprehension of the situations which precede the fight is required in order to be able to take actions in order to prevent the fight.

In general we consider a scenario as a sequence of situations. A hostile scenario will lead to an unwanted (i.e. hostile) situation. The goal of surveillance is to recognize situations which precede the unwanted situation so that action can be taken to prevent the hostile situation. Recognition of a situation is closely related to human awareness of a situation. Endsley [1] defines situation awareness as the perception of the elements in the environment within a volume of time and space, the compression of their meaning and the projection in the near future. Situation awareness is also a well-known topic for sensor fusion where various levels are discriminated. These range from sensor signals, object awareness, situation awareness to threat awareness [2]. In order to recognize pre-hostile situations we need to observe objects, activities and behaviour that are indicative for such a situation. This is depicted by the diagram in Figure 1.

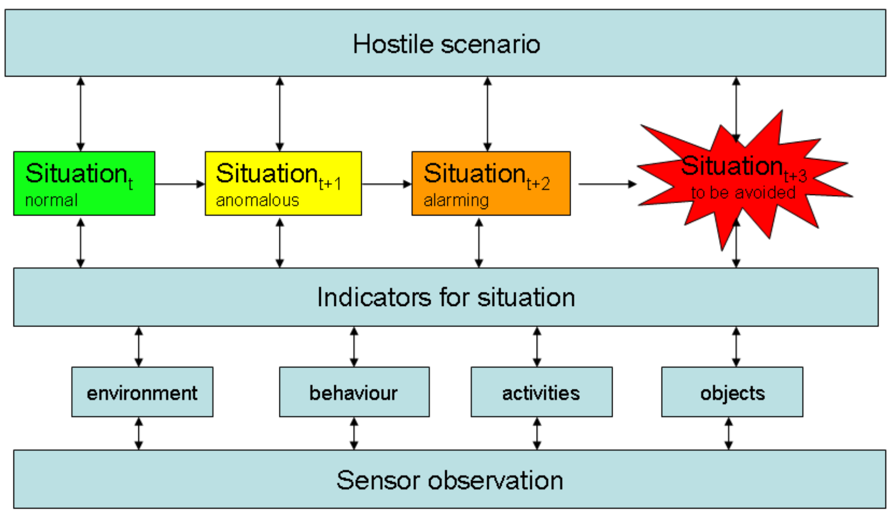

Fig. 1. Diagram showing the relations between the scenarios and the sensor observations.

We discriminate between 4 situations: normal, anomalous, alarming and the hostile situation itself (see Figure 1). For recognition we want to observe the indicators for the threatening or anomalous/alarming situation, which are associated with the kind of objects, their position, their activities and behaviour. Since we focus here on indicators for pre-hostile situations these indicators are concisely called here *threat indicators*.

We are interested to study the sensor observation of the threat indicators. The role of the sensors above visual monitoring is to enhance the observation capability for surveillance. Sensors make it possible to observe continuously, overcome larger distances and provide overview. The use of sensors is steadily increasing. This is illustrated by the use of monitoring cameras, for which it is common to display the results from several surveillance cameras on one monitor, without any further processing. This results in information overkill which is very confusing, so that interpreters are likely to miss many details. In order to be effective, the use of sensors for surveillance should not require too many people and too much time. Bottleneck is here the processing and interpretation of the sensor data. Therefore the sensor data and information need to be processed by automatic procedures as much as possible.

We consider here two types of sensors: an advanced radar system and a camera system. The radar has the advantage that it can monitor with a high precision at large distances and with a wide field of view. It can monitor during adverse weather conditions (e.g. fog) and at night, since it is an active system. Detection of motion can be done easily and automatically. On the other hand detailed description of an object is difficult with radar data and for this purpose camera data is unrivalled. In the following we study the radar for providing automatically information for the detection and locations of persons, their activities such as walking and running, and geospatial information. The camera is studied for automatically providing information about the identity of an object and behavioral features, such as carrying objects.

3. METHOD

In order to provide an outline for the study we have defined hostile scenarios near an operating base. We describe two vignettes, which give a detailed description of key elements and events in a situation for the hostile scenario. The context used is Afghanistan near an operating base where human beings on the road are mixed with animals, e.g. horses and dogs. For the investigation of these vignettes we have made measurements of human and animal behaviour in order to study the feasibility for recognizing the threat indicators.

3.1 Rocket attack

In this hostile scenario an operating base is attacked by short range rockets, e.g. an RPG or a 107-milimetre missile, which on April 6 2009 killed one person in the camp Holland. Rockets and mortars are usually fired from an area which cannot be directly observed from the camp, for example because the line of sight is obstructed by relief or trees. Such suspicious places sometimes called hot spots. People deploying activities near a hot spot, which in a *normal situation* is desolated, this indicates an *alarming situation*. Such activity can be various people moving from one place to the other in the hot spot for carrying and displacing equipment.

3.2 People attacking

In a peace keeping mission an operating base is often close to a city for logistics reasons but also to be able to contact the local people. In a *normal situation* people behave disciplinary, quietly and with clear intentions near the base and in the Afghanistan context also animals are present. An *anomalous situation* occurs when objects are detected that are approaching the base rapidly without warning and without clear intentions. Such a situation can be part of a non-hostile or a hostile scenario. A hostile scenario can be an attack where a person is running towards the camp to draw the attention of guards so that others can fire on the camp unnoticed. A non-hostile scenario can be a horse which is galloping in a training exercise near the camp. In this case it obviously important to know what is moving as early as possible and how many people are approaching.

3.3 Threat Indicators

On basis of above mentioned vignettes we can infer various threat indicators, for which we can formulate the following topics for automatic processing of sensor data:

Table 1. Threat indicators and automatic processing

	Threat indicator	Topic for automatic processing
1	Human presence	Detect humans and discriminate humans from animals
2	Running person	Determine a running person
3	One or more persons	Determine one person or more persons are approaching
4	Person in a suspicious place	Determine people in a suspicious place (hot spots)
5	Person carrying object	Determine a person carrying an object

In the following section we describe processing techniques which can be used for abovementioned topics.

4. PROCESSING TECHNIQUES

4.1 Radar Range Doppler processing

The basic capability of a radar is its ability to measure distances (ranges) to the reflecting object with an accuracy given by the range resolution (δr), which is determined by the bandwidth (B) following:

$$\delta r = \frac{c}{2B} \tag{1}$$

and where c is the speed of light. For example, a bandwidth of 400 Mhz implies a range resolution of 40 cm. In addition radar measurements are useful for automatically discriminating moving persons and animals from the stationary background. This is possible since motion causes a velocity dependent Doppler shift in the reflected radar signal. We consider here radars which transmit a frequency modulated continuous wave (FMCW) signal. This implies that the frequency changes as function of time resembling a saw tooth pattern (frequency sweeps). The total frequency change determines the bandwidth (B). In range-Doppler processing [3] range and velocity information of moving objects is retrieved by applying a double Fourier transform to the received signal. The first transform is applied to the received signal from a transmitted upsweep in order to produce a high resolution range line. The second Fourier transform is applied across a number of range lines choosing an appropriate integration interval (T_{int}) in order to produce Doppler or velocity values for each range cell. The velocity resolution δv is obtained following:

$$\delta v = \frac{c}{2T_{\text{int}}f} \tag{2}$$

where f is the carrier frequency. Typical numbers are here T_{int} ~ 0.1 sec, f= 5.7 GHz (C-band), so that δv ~ 0.25 m/s).

In this way so-called range-Doppler maps are obtained where one axis represents range and the other axis represents velocity (see Figure 2).

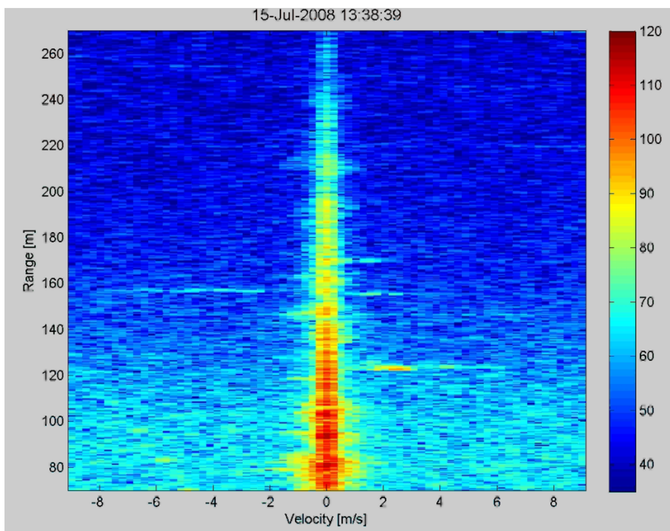

Fig. 2. Range-Doppler map showing to the left and right from the bright clutter band (zero velocity) moving persons at ranges of 120, 155, and 170 meter. Intensity values are arbitrary.

The range-Doppler maps show a bright band near zero velocity which is due to the refection from the stationary background (clutter). In order to obtain range, and velocity information from moving objects the reflection has to be isolated from the clutter background. A suitable technique is to reduce the clutter by subtracting a so-called clutter map (recent-time average of range-Doppler maps without the moving object) before applying a threshold or CFAR detector [4].

4.2 Micro-Doppler signature analysis

The object (target) to be recognized will be recorded during a certain time interval (typically 5 seconds) so that an analysis can be made of its movement [5]. Doppler signatures are obtained by tracking objects in the range-Doppler diagram during consecutive measurements. In this way range-Doppler information centered on the object can be obtained and analysed.

Features can be extracted in two ways. One way for feature extraction is to integrate the range-Doppler maps centered on the object over a certain time. Another way is to extract features from the repetitive patterns or micro-Doppler signatures, e.g. due to human walking and running. Also galloping horses and running dogs produce such patterns. These micro-Doppler signatures can be visualized in so-called Doppler spectrograms where the Doppler spectrum is plotted as a function of time.

4.3 Geospatial analysis

Most available surveillance radars are scanning radars, where the azimuth accuracy is proportional to the beam width, so that for a good azimuth resolution a small beam width has to be used. This has the disadvantage that analysis of micro-Doppler signatures related to activities like walking, running etc are more difficult to observe, since the time the target is illuminated is very short. In order to overcome this problem a staring radar can be used in an interferometric stereo mode. For such a stereo radar two antennas are used. One antenna is transmitting and two antennas are receiving. In the interferometric stereo processing the phase differences ($\delta\varphi$) between the two receiving antennas are related to the azimuth angle θ following [6]:

$$\sin\theta = \frac{c\partial\varphi}{2\pi f L} \tag{3}$$

where L is the baseline, c the speed of light and f the carrier frequency, so that for each detected (moving) object next to range also the azimuth angle can be determined. The distance between the two receiving antennas, the baseline L, determines the accuracy of the azimuth position, which can be 0.5 degrees generally much better than the azimuth resolution of a scanning radar. A stereo radar system has a wide-field of view angle depending on the beam width of the radar which can be much larger than for a scanning system. In this case this was 40 degrees. Of course when 360 degrees of azimuth has to be monitored several radar systems have to be used. Since a stereo radar provides both azimuth and range information, the location of the detected moving object is known when also the position and height of the radar and the relief of the terrain is taken into account. From the range-Doppler processing also the radial velocity is measured. This information can be used to reconstruct a track of the moving object. The track can then be used to determine if the object enters a hot spot. In order to do this automatically we have to define how we consider a hot spot in a geospatial context. Referring to the vignette in section 3.1, hot spots are places which we consider to be the more likely places for firing a mortar or an RPG. These places are determined by the terrain. For example places which can not be directly observed from an operating base are more suspicious since the attacker can hide there. Also there is a favorable distance for firing RPGs and mortars. Other aspects are the vicinity of access roads. On basis of these criteria a map can be constructed which indicates the likeliness for positions from where RPGs or mortars are fired. We call such a map a Observable Probability Map (OPM) referring to what we want to observe: presence of persons in a hot spot.

4.4 Video data analysis based on segmentation

Moving objects in a video stream can be detecting from differences to a static background, which has to be estimated first. The resulting segment can be used to extract a number of features, such as the area, the major axis length, perimeter, eccentricity, etc. These features can be collected in a vector, so that this vector is assigned to an individual video frame. The information content of such a vector reflects information about the pose of the observed object. Monitoring moving objects for a certain time will result in a sequence of vectors, which contain information about the changes in the poses, which are related to the motion of the object. By analyzing this sequence with autocorrelation techniques, information about the repetition rate in the motion (e.g. running and walking) can be determined. This information can then be used for to discriminate between differently behaving persons, and between persons and animals.

4.5 Video data analysis using patches

Visual details are very informative of a person's behavior, such as movements and details such as carrying an object. A person can be considered as a dynamic, non-rigid object consisting of many parts. Hence, in a video feed, we collect image patches for these parts from the object. The patches are detected from strong motions. Motion is detected by using an online Gabor filter. The Gabor filter yields high responses to parts with large dynamics. An example of a sequence with a moving person is shown in Figure 3a. The maxima of local responses provide us with detections of moving video patches as illustrated in Figure 3b.

Fig. 3. a. Motion sequence (left), b. Extracted patches from sequence (right)

The detected patches are described by features. These behavioral features comprise both the *appearance* and *dynamics*. *Appearance* is modeled by Hu-moments [7], which are statistical features. Advantages are that they are compact (we keep the most important 3 values), efficient to compute and invariant under translation, changes in scale and rotation. To capture also the dynamics of the patches, 3 time samples are collected. Hence, the feature describing both *appearance* and *dynamics* is a vector consisting of 3x3 = 9 values.

To compare moving objects to each other, a similarity measure is needed from one set of Hu-feature vectors to another. In our application, the number of patches may vary and thus the number of features. Hence the similarity function should be able to compare two sets of patches that are not of the same size. The Earth Mover's Distance (*EMD*) is able to express dissimilarity between two sets of different sizes, based on individual feature-to-feature dissimilarities [8]. The *EMD* dissimilarity is the result of a minimization of feature-to-feature assignment, resulting in larger dissimilarities if the features from the two sets are less similar. An example of the *EMD* dissimilarity measure has been applied to two persons and a dog. The *EMD* for a person and the dog is large. The *EMD* for the two persons is small, despite the fact that the posture, poses and the colors of the clothes for the two persons are very different. This implies that the *EMD* similarity measure provides a robust means to compare objects and their behavior.

Classification is done on basis of learning using input from an expert. Given the expert's specification of the part of the object that is of interest, the video features are extracted from that particular region. They are the input for learning of a classifier. An example here is to decide whether a person carries an object or not: we learn the *dynamics* of both cases and search for the differences. In this case the expert is thus interested in two cases between which a distinction has to be made. Two-class classification implies that for both classes examples have to be provided by the expert. The two-class classifier assigns the current input simply to the class to which it finds the best match based on the smallest *EMD* distance (i.e. nearest neighbor classifier). A more extended description can be found in Burghouts et al. [9].

4.6 ATR system

To facilitate the automatic processing, an automatic procedure for moving object classification has been developed at TNO, which will be implemented in a prototype system. In the system object classes depending on threat indicators are specified by the operator. For the classification, the object classes are correlated with extracted sensors features. The correlation is done on basis of empirical relations for sensor features versus object classes, theoretical models or by training (learning) using the input from experts and probabilistic models such as a Bayesian network and (hidden) Markov models. Also a priori information about the object classes can be taken into account. The operator can indicate the relevance for detecting an object. For example detection of humans is often more relevant than animals and the system deals with this, so that false alarms are reduced as much as possible. The system allows free specification of the object classes as long as the information for correlation is provided. Sensor features are automatically extracted using the techniques described in the previous section. Moving objects are detected by the (stereo) radar from which the location and speed is determined. Features used for classification comprise micro-Doppler and optical *appearance* and *dynamic* features.

5. MEASUREMENTS

For the study we have performed various measurements with cameras and radars. In one series of measurements we have recorded moving persons and group activities in a scene which expressed a possible pre-hostile situation near an operating base. Another series of measurements focused on the recognition of objects such as persons and animals.

5.1 Situation measurements

In these experiments we used a 5.7 GHz (C-band) stereo radar next to a wide-field of view and a narrow-field of view camera. The sensors were mounted on the measurement tower of TNO at a height of 38 meter. This also simulates placement on an aerostat in order to improve the line of sight. The stereo radar consisted of one transmitting and two receiving antennas where the latter two were separated by an 8 cm baseline. The beam width of the radar is about 40 degrees. The bandwidth is 400 Mhz resulting in a range resolution of 40 cm and the maximum range is slightly more than 200 m. See for the layout of the experiment and the position of the radar Figure 4.

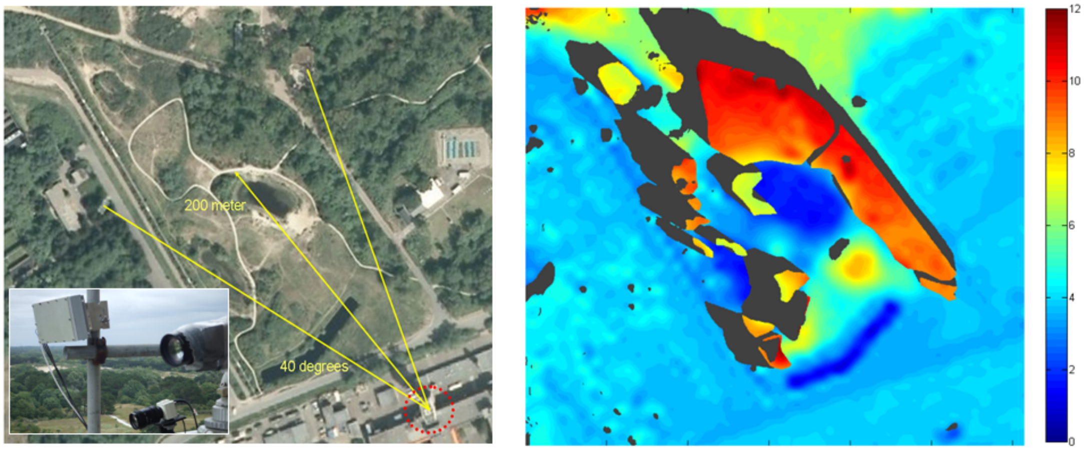

Fig. 4. a. Experiment layout with the position of the tower (red dotted circle), the radar beam (yellow lines). Inset shows mounted stereo radar, a wide and narrow field of view camera (left), b. Digital elevation model, including tree height which is assumed to be at a constant height of 5 meters. The radar shadow is indicated in grey. Height is given in meters above average sea level (right).

Measurements were made of individual people walking or running on the paths (white lanes, see Figure 4a). Also measurements of people walking together and walking on the grass area were recorded. Since the radar is staring in the background (grazing angles of 10-15 degrees) a significant contribution from the clutter can be expected.

5.2 Object measurements

In these measurements we have recorded a horse, a dog and various walking and running persons with the radar. To minimize the contribution of the clutter, we have measured the objects in clutter poor environments (i.e. an empty road and a landing strip of an airfield). The radar consisted of a frontend with the following specifications: carrier frequency 9.4 GHz and 600 Mhz bandwidth so that a range resolution of 25 cm is obtained. Next to the radar measurements also optical and infrared videos were taken.

6. RESULTS AND ANALYSIS

Using the measurement results we apply for the various topics the processing techniques described in the previous section (see Table 2).

Table 2. Topics and analysis techniques

	Topic	**Analysis Technique for automatic processing**
1	Detect humans and discriminate humans from animals	Video analysis Micro-Doppler signatures
2	Determine a running person	Video analysis Range-Doppler Micro-Doppler signatures
3	Determine one person or more persons are approaching	Micro-Doppler signatures
4	Determine people in a suspicious place (hot spots)	Geospatial analysis
5	Determine a person carrying an object	Video analysis

6.1 Detect humans and discriminate humans from animals

6.1.1 Video analysis

Both the *appearance* and the *dynamics* from the whole object are exploited to distinguish persons from no-persons. Figure 5 shows a classification result. The no-person is a dog, which has moving legs, like persons. The persons considered here have significantly different *appearances*, which makes the classification challenging. Yet the features are able to discriminate all persons from no-persons, such as the dog.

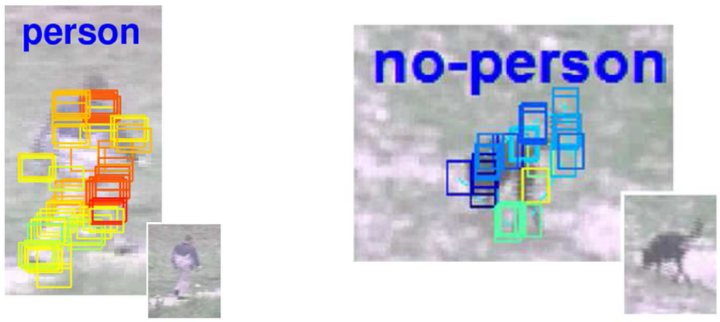

Fig. 5. Discrimination of person and no-person. Image patches which are significant for the classification are coloured red to yellow, for non-significant patches the colouring is towards blue.

6.1.2 Range-Doppler analysis

Four examples of micro-Doppler signature analysis are shown in Figure 6. The figure shows the spectra for an approaching horse and human, centered in range at the estimated body position and averaged over 0.2 seconds. The determination for running and walking will be explained in the next paragraph 6.2. The distinction between humans and horses are based on three features: one, the overall extend in range (humans about 1 m, horse about 2.5 m); two, the velocity (on average a running horse has a much higher velocity than a human running), and three, the presence of two motion centers (at -0.5 m and + 0.75 m) for the horse as compared to the single motion center (0 m) for the human.

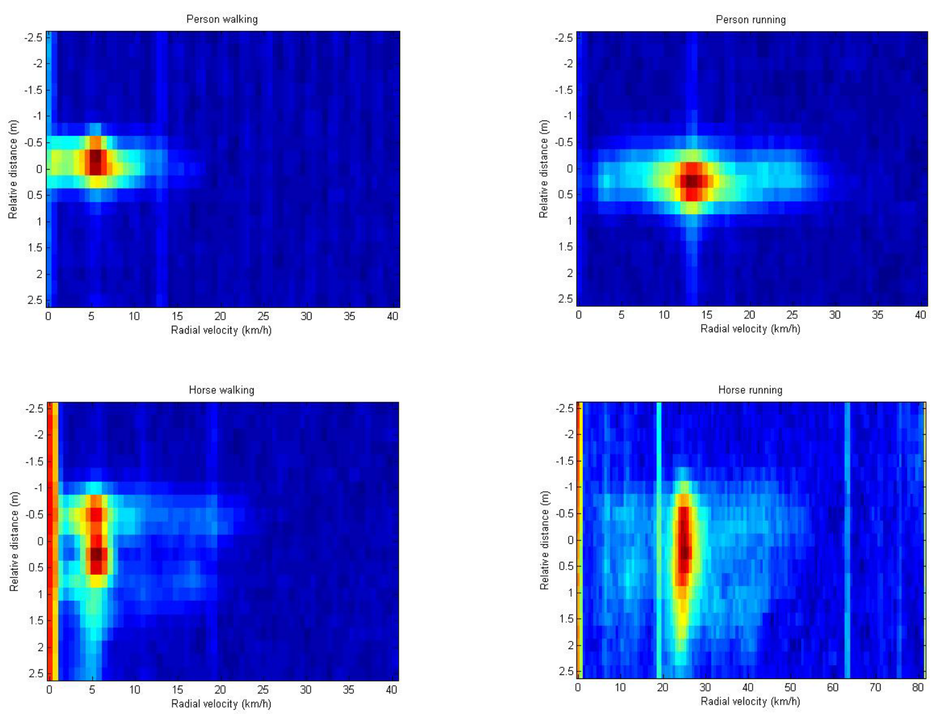

Fig. 6. Integrated RD images of a person walking (top left), a person running (top right), a horse walking (middle left), a horse running (middle right) and a dog running (bottom right).

6.2 Determine a running person

6.2.1 Video analysis

Using video data from the object measurements (see section 5.2) the motion of walking and running persons has been analysed using the processing method described in section 4.4. In Figure 7 we show so-called power spectral density plots which are the results of processing data from a walking (Figure 7a) and a running person (Figure 7b). We find the first peaks at 0.07 Hz and 0.1 Hz in Figure 7a and 7b respectively. The frequency is per video frame, where the frame repetition rate is 15. Human gait repeats itself exactly after 2 steps. Using a step length of about 0.8 meter this corresponds to a speed of 6.0 km/h and 8.6 km/h, which confirms walking and running. Although somewhat smaller, a second peak is found at double frequencies which reflects the repetition rate of poses between one step. Although these poses are not identical, they are similar and only mirrored along the direction of motion.

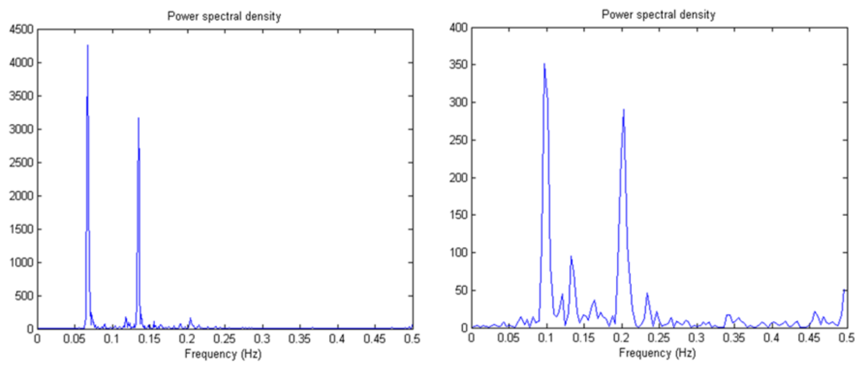

Fig. 7. a. Power spectral density plots for a walking person (left), b. For a running person (right).

6.2.2 Micro-Doppler signatures

Range-Doppler Data from the situation measurements (see section 5.1) of a walking and running person were processed for a total length of about 35 seconds, of which the first 20 seconds (which do not include the target) are used to build up the clutter map. In the last 15 seconds, which include the target, the clutter is subtracted using the clutter map. The data were processed with an integration time of 0.20 seconds (in order to achieve the desired Doppler resolution) and with an overlapping factor of 0.15 seconds, so that we get an updated range-Doppler map every 0.05 seconds. Oversampling is required in order to visualize target spectral features in the Doppler spectrogram. In Figure 8a Doppler spectrograms of a walking person and in Figure 8b a running person are shown. A clear pattern (periodic features) can be observed in both cases. The time axis (y) is given in frames, where the time interval between the frames is 0.05 seconds.

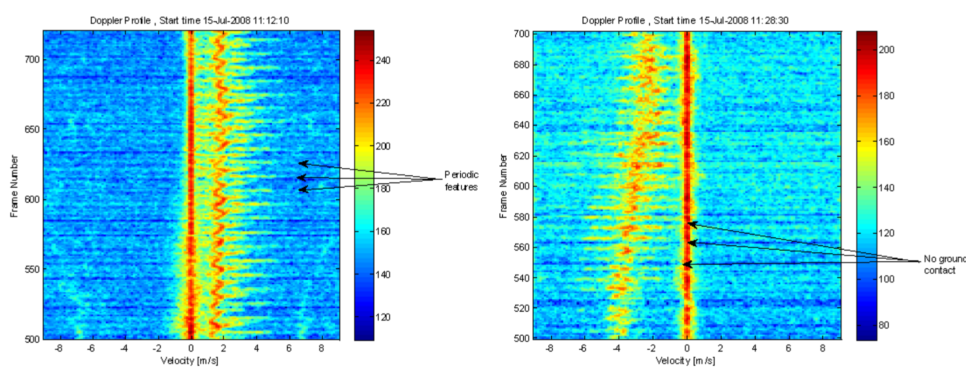

Fig. 8. a. Doppler spectrogram for a walking person (left), b. for running person (right).

For discriminating a walking person from a running person we use two features. Beside the difference in speed, walking and running persons can be distinguished by looking at the Doppler spectrum in the region close to the zero velocity (ground return). For a walking person this region is more filled due to the stationary contact of always one leg with the ground (corresponding to zero Doppler, see Figure 8a), whereas for a running person both legs are moving most of the time (see Figure 8b).

6.3 Determine one person or more persons are approaching

Range Doppler data from the situation measurements of a walking person (same as in the previous Section) are compared here with range-Doppler data of two persons walking together. The data were processed analogously to the data in the previous section. In Figure 9 the Doppler spectrograms are shown for the single person walking (left) and the two persons walking together (right). We cannot use the speed to discriminate between the two cases since the target velocity is the same (walking speed of ~1.5 m/s). Therefore analysis of the Doppler spectrum has to be used to distinguish between one and two (or more) persons. For the case of a single person periodic features in the Doppler spectrum are clearly visible (Figure 9a) while they are not distinguishable when there are two or more people walking together (Figure 9b). The latter is expected when two or more persons are walking asynchronously.

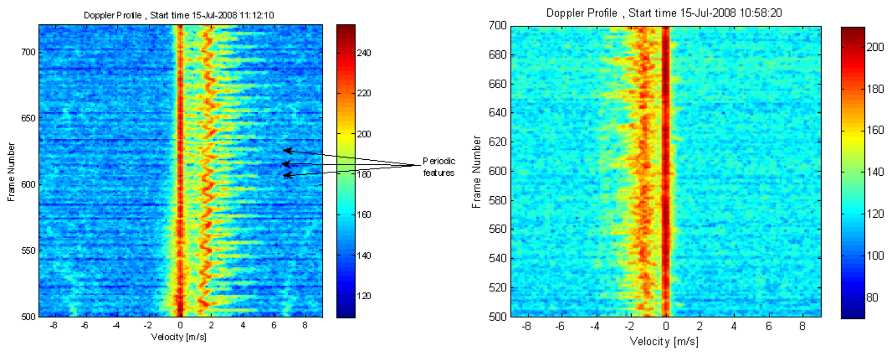

Fig. 9. a. Doppler spectrogram for a single person walking (left), b. For two persons walking together (right).

6.4 Determine people in a suspicious place (hot spots)

Detected moving objects from stereo range-Doppler maps are also called contacts and have information about the range, angle and radial velocity. With Kalman filtering [10] contacts of successive updates can be transformed into tracks, if the changes from one contact to another follow the laws of motion that apply to humans. Contacts that do not follow these laws of motion are considered to be false contacts and are filtered out. A Kalman filter recursively estimates the current state based on the estimated state from previous update and the current measurement. In order to make an accurate prediction for the next contact the filter also updates internally an error covariance matrix. The covariance matrix holds for each state the dispersion on the range, angle and radial velocity and reflects therefore the position error. This measure, the position error measure (PEM) can be modeled by (normalized) multivariate normal distributions (see Figure 11c). In Section 4.3 we introduced the observable probability map (OPM) as a map containing probabilities for the likeliness of a suspicious place or hot spot. An example was given for the rocket attack vignette described in 3.1, where the likeliness was based on no-line-of-sight from an operating base and a critical shooting range for e.g. RPGs. Figure 10 shows how the OPM is obtained by a convolution of no-line-of sight areas and the critical zone.

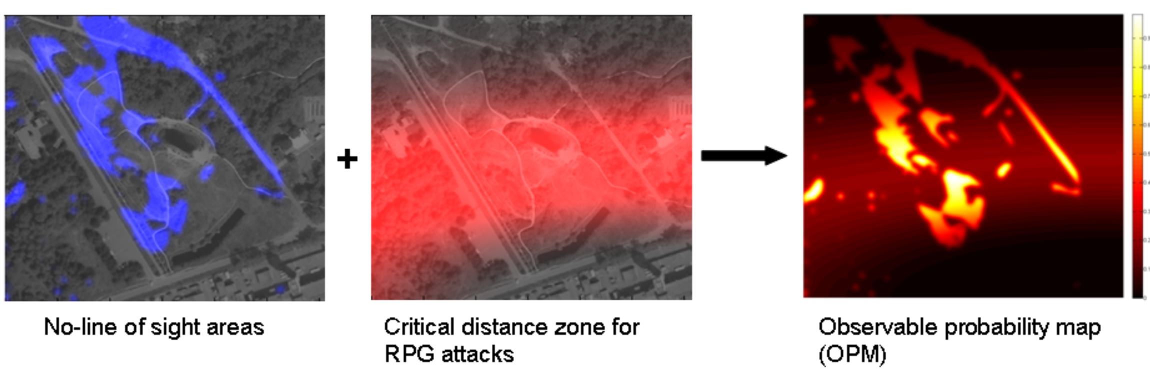

Fig. 10. OPM on basis of thematic maps for line of sight from an operating base and fictive firing distance for RPG, scale for probabilities ranges from 0 to 1.

Note that we assume that for hot spots there is no line-of-sight from the base itself, but that these spots can be observed from the elevated position on board of an aerostat. In order to determine the probability for presence in a hot spot, the

OPM is convoluted with the PEMs, so that also the position uncertainty is taken into account. In Figure 11a a red and a blue track obtained by the stereo radar of two persons moving in the terrain are shown. In Figure 11b these tracks are overlaid on the OPM shown in Figure 10. In Figure 11c PEMs for the track elements are shown, where the inset shows one of the individual PEMs. The larger gaps, visible between the individual elements (contacts) of the tracks are due to the fact that the Kalman filter did not find a contact. In these cases elements including an associated PEM where added by linear interpolation.

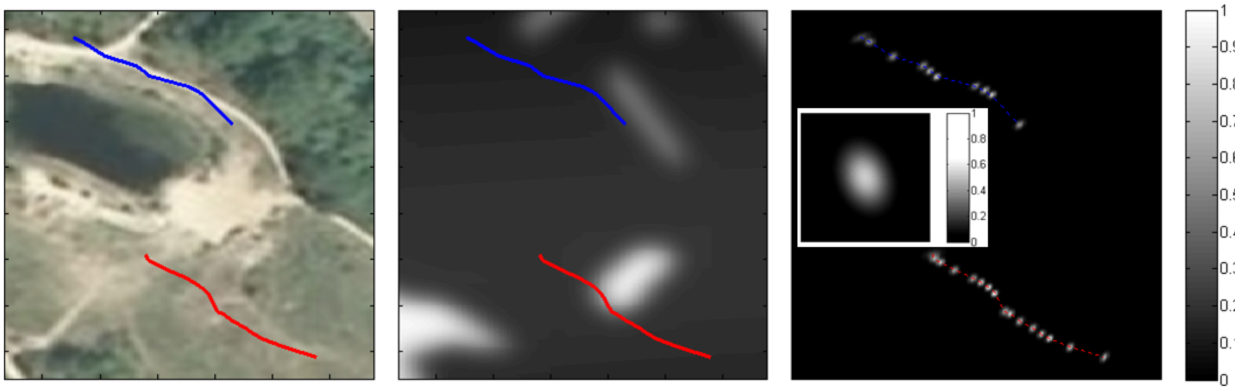

Fig. 11. a. Tracks (running from left to right) of two walking persons shown in the scene (left). b. Tracks overlaid on the OPM (middle). c. Individual PEMs for the two tracks (right).

The results of the convolution of the PEMs with the OPM, which indicates the probability for presence in a hot spot, are shown in Figure 12. The red track shows a peak were the track is overlaid on the bright area of the OPM while the blue tracks comprises only low values increasing to the end when the tracks approaches a brighter area. We can use a threshold (chosen here 0.4) to decide whether a person is present in a hotspot. In this case the person associated with the red track is present in a hot spot while the person associated with the blue tracks is not. By determining the time a person is present in a hot spot, an alarm can be given when the time spent is sufficient to prepare a RPG or mortar attack.

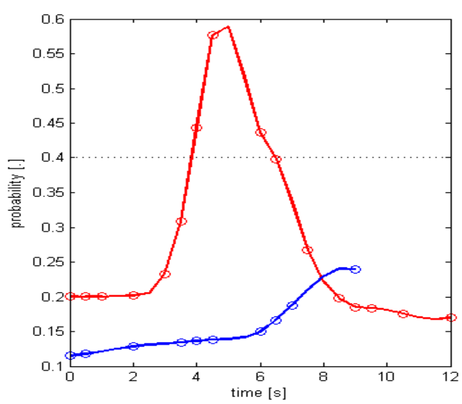

Fig. 12. Probability of presence in a hot spot as function of time for the blue and red track. The black horizontal line is a manually selected threshold.

6.5 Determine a person carrying an object

To distinguish whether persons are carrying an object or not, the behavioral features focus on the *dynamics* only. *Dynamics* are derived from the image patches that are beside the lower part of the person. The rationale is that persons who carry an object have lower *dynamics* around the carried object. Figure 13 shows a classification result. The *dynamics* feature is able to discriminate whether persons carry an object, so that persons getting rid of an object can be identified. Note that the *dynamics* feature is also able to specify the position of the carried object, as can be concluded from the red and orange patches. Interestingly, the high contrast of the carried object is not an informative detail since the hands of the person who is not carrying also show a high contrast.

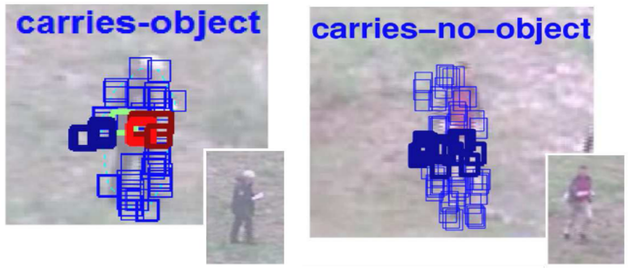

Fig. 13. Discrimination of a person carrying an object and a person carrying no object. Image patches which are significant for the classification are coloured red to yellow, for non-significant patches the colouring is towards blue.

7. DISCUSSION & SUMMARY

The increasing availability of more advanced, more compact and cheaper sensors and sensor suites is a promising development for surveillance. Human involvement, needed for the sensor operation and data interpretation, is expensive and should therefore be minimized as much as possible. This is a challenge for sensor management and automatic processing of sensor data. In this paper we have studied the feasibility of automatic data processing techniques for surveillance purposes, which are coupled to threat scenarios. These threat scenarios always depend on the context and are therefore principally not well-defined. This means that automatic procedures for detection of threat indicators (in this case associated with human behaviour) always have to be tuned to specific scenarios. In this paper we have chosen two exemplary scenarios for which we have studied relevant automatic processing techniques using radar and camera data. We have studied procedures which are quite generally applicable and which will often be a component of a tuned procedure. The results confirm the feasibility of detecting human behaviour that is associated with threat indicators using these automatic techniques. In a next stage of the study methods how to generally relate threat scenarios and automatic processing techniques will be studied using probabilistic models such as Bayesian networks and (hidden) Markov models.

ACKNOWLEDGEMENTS

The authors kindly thank Ardjan de Jong, Philip van Dorp, Bram Alefs, Richard den Hollander, Eric den Breejen, and Klamer Schutte for discussions and contributions to the project.

REFERENCES

[1] Endsley, M.R., "Toward a theory of situation awareness in dynamic systems", Human Factors 37(1), 32-64 (1995).
[2] Steinberg, A., Bowman, C.L., "Rethinking the JDL data fusion levels", Proceedings of the Military Sensing Symposia(MSS) National Symposium on Sensor and Data Fusion (NSSDF) Colombia, (2004).
[3] Barrick, D.E., "FMCW Radar Signals and Digital Processing", National Oceanic and Atmospheric Administration Boulder Technical Report AD-774-829, (1973).
[4] Nitzberg, R., "Clutter Map CFAR Analysis", IEEE TAES Vol. 22 No.4, 419-421 (1986).
[5] van Dorp, P., Groen, F.C.A, "Human walking est.imation with radar", IEE Proceedings, Radar and Sonar Navigation., Vol. 150 (No.5), 356-365 (2003).
[6] de Jong, A.N., "FMCW radar observation in buildings during calamities", TNO report TNO-DV2006A509, (2006).
[7] Hu, M.K., "Visual pattern recognition by moment invariants", IRE Trans. on Information. Theory vol. 8, 179-187 (1962).
[8] Rubner, Y., Tomasi, C., Guibas, L.J., "A metric for distributions with applications to image databases", IEEE conference on Computer Vision and Pattern Recognition, (1998).
[9] Burghouts, G.J., van den Broek, A.C., Alefs, B.G., den Breejen, E., Anitori, L., Boekema, R.B., Schutte, K., "Hostile intent, automatic behavioral indicators", to be published, (2009).
[10] Roweis, S., Ghahramani, Z., "A unifying review of linear Gaussian models", Neural Comput. Vol. 11 No. 2, 305-345 (1999).