International Journal of Computer Vision manuscript No.
(will be inserted by the editor)

Detection and Tracking of Multiple Metallic Objects in Millimetre-wave Images

C. D. Haworth[1], Y. De Saint-Pern[1], D. Clark[1], E. Trucco[1], Y. R. Petillot[1]

Heriot-Watt University, Edinburgh, EH14 4AS, United Kingdom

Received: date / Revised version: date

Abstract In this paper we present a system for the automatic detection and tracking of metallic objects concealed on moving people in sequences of millimetre-wave (MMW) images. The millimetre-wave sensor employed has been demonstrated for use in covert detection because of its ability to see through clothing, plastics and fabrics.

The system employs two distinct stages: detection and tracking. In this paper a single detector, for metallic objects, is presented which utilises a statistical model also developed in this paper. The second stage tracks the target locations of the objects using a Probability Hypothesis Density filter. The advantage of this filter is that it has the ability to track a variable number of targets, estimating both the number of targets and their locations. This avoids the need for data association techniques as the identities of the individual targets are not required. Results are presented for both simulations and real millimetre-wave image test sequences demonstrating the benefits of our system for the automatic detection and tracking of metallic objects.

Key words millimetre-wave, detection, tracking, metallic objects

1 Introduction

1.1 Millimetre-wave Imagery in Security and Surveillance

We present a system for the automatic detection and tracking of metallic objects concealed on moving people in sequences of millimetre-wave (MMW) images, which can penetrate clothing, plastics and fabrics.

MMW imaging is emerging as an important modality for security and surveillance thanks to recent advances in MMW sensing technology. Providing head-to-toe monochrome images highlighting concealed threats opens the possibility to locate threats on the body and analyse their shape, which is far beyond the reach of conventional metal detection portals. A recently demonstrated proof-of-concept sensor [1] developed by QinetiQ provides video-frame sequences with near-CIF resolution (320×240 pixels) and can image through clothing, plastics and fabrics. The combination of image data and through-clothes imaging offers huge potential for automatic covert detection of weapons concealed on human bodies via image processing techniques. Previous trials of the QinetiQ MMW sensor, some involving the Department for Transport and British Airport Authority, showed potential for passengers screening at airports [2], public event security [1] and detection of illegal passengers in lorries [3]. All trials involved human operators.

The sequences in this paper are generated by an electro-optic sensor working between Infra-Red (IR) and microwave wavelengths. The sensor forms an image of the temperature received from the scene, which is a standing human subject turning around slowly. Figure 1 shows examples of frames from a typical sequence considered in our work. A person turns around by 360° in front of the sensor and is captured at video rate (12 frames per second). The temperature (and therefore the pixel intensity) is a function of the reflectivity, emissivity and transmissivity of the scene surfaces. At the wavelength used, metallic objects are highly reflective and tend to appear bright, the human body is partially reflective and appears less bright and clothes are partially transparent. An illumination chamber is required for indoor operation [4] but does not expose the subject to harmful levels of radiation.

To our best knowledge, very little work has been reported on the automatic analysis of MMW sequences or images. Basic image segmentation [5,6] has been reported with some success. Shape identification on the segmented images [7,8] has been investigated and suitable shape descriptors proposed. However, the image quality is poor with a small field-of-view and cannot be gathered at video frame-rate. Due to these limita-

Fig. 1 Example MMW image showing a human subject. Notice the Speckle noise pattern particularly apparent on the torso and the substantial smoothing which is applied during the image formation process to minimise visual artifacts.

tions only a highly constrained stationary scene is considered. The proposed shape descriptors prove reliable under the operational constraints. In comparison, the QinetiQ sensor used in this paper is a real-time, head-to-toe sensor with considerably improved image quality. More recently, some work on the detection and segmentation of metallic objects has been proposed [9] for the QinetiQ sensor. Alternatively, image fusion [10,11] has been employed on MMW / IR. While this work has produced visually appealing images for human operators, no work has been reported on automatic detection in fused MMW-IR images and it is not clear that there would be any benefit. The main contribution of our work is therefore to apply advanced image processing for the automatic detection of concealed weapons to a new video imaging technique of high potential for many applications in public security.

1.2 Target Detection and Tracking

The aim of the target tracking is to improve the robustness of the estimates of target properties, minimise the number of detection errors and, importantly for an operational system, to provide a robust estimate of the number of threats on a subject. The MMW images have a relatively low spatial and colour resolution combined with considerable speckle noise, resulting in a a poor signal-to-noise ratio (SNR). For these reasons, a significant number of errors is anticipated within any target detection with clutter a particular problem. Therefore, target tracking is required to operate in an environment with an unknown and changing number of visible targets and in the presence of clutter. Future work will concentrate on providing a complete characterisation of the object, including areas such as shape identification, for which accurate target tracking will be important.

Tracking objects in visible-wavelength sequences is a well-studied problem in image processing and computer vision [12–17]. Traditional multi-target tracking is based on coupling trackers such as Kalman filters, extended Kalman filters or particle filters with a data association technique (see for example [18]). The aim of the data association process is to interpret which measurements are due to the targets and which are due to false alarms.

Particle filter approaches to multiple target tracking have continued to use data association techniques [19–21]. This can be partly attributed to well established techniques for tracking and partly due to a lack of efficient techniques for modelling multiple targets with particle filters. In contrast, the PHD Filter is a method of propagating a multi-modal measure within a unified framework without associating the measurements and has the ability to estimate the number and position of targets in data with clutter. If the identities of the targets are not required then avoiding data association has a significant computational advantage. Furthermore, the PHD Filter provides a natural fusion framework to combine multiple detectors therefore increasing reliability and robustness in the detection of concealed weapons.

Particle filter methods for the PHD filter have been devised by Vo [22] and Zajic [23]. Practical applications of the filter include tracking vehicles in different terrains [24], tracking targets in passive radar located on an ellipse [25] and tracking a variable number of targets in forward scan sonar [26].

In this paper the PHD Filter is presented in a system with a single detector, which provides the position, area and scene intensity of any likely metallic objects in the scene. The combined detection and tracking of metallic objects is tested across a range of different MMW sequences of real people carrying concealed objects. An implementation of a particle filter tracking system is also presented for comparison on these real sequences.

1.3 Paper Organisation

This paper is organised as follows: Section 2 presents a statistical mixture model. The mixture model is then used in Section 3 to develop a classification strategy for millimetre wave images. The classification first examines the sequence, then individual frames and finally regions of a frame to identify possible threats. In Section 4 target tracking is presented that addresses the problem of estimating the number of targets and tracking multiple targets simultaneously, with the aim of improving the robustness of the threat detection and providing extended information on the position of the threat on the subject. Results are presented in Section 4.8 on real MMW sequences.

2 Mixture Models for MMW Images

MMW images offer good data for material discrimination as different materials yield, generally speaking, different image properties. In analysing the image statistics it would be desirable to have an understanding of the

physical process which could be incorporated in a model for the MMW image formation process. However, given the complexity of the MMW imager and the extensive amount of hardware calibration, software equalisation and interpolation undertaken to produce a MMW image, this is a non-trivial task.

In this paper we adopt an approach modelling the differences in image properties statistically, using a weighted mixture model in which each PDF, f_i, is associated to a specific material:

$$f_{mix} = \sum_{i=1}^{N} \alpha_i f_i(\boldsymbol{\theta}) \quad (1)$$

where α_i is a weight and $\boldsymbol{\theta}$ a vector of parameters.

To identify the best-fitting PDF for each material (incl. background, i.e., non-figure pixels), we built a number of mixture models made by combinations of standard distributions (e.g. Gaussian, Rayleigh, Laplacian), optimised the parameters with a standard Maximum Likelihood (ML) algorithm and picked the best fitting combination for the observed image histograms using a Chi-Square test. We started with background-only sequences (no subject) to identify the background distribution. We then moved to sequences of scenes with a subject but no threats, then with a subject carrying threats (metallic objects). The final result is a best-fitting mixture model for each material (types of component distributions, and parameters). As an example, Figure 2 shows histograms and results of the ML distribution fit for a scene containing a subject carrying no threats. Here, a two-component mixture model is used: two Gaussians, leading to poor fit, and Laplacian-Rayleigh, showing good fit and little overlap between component distributions.

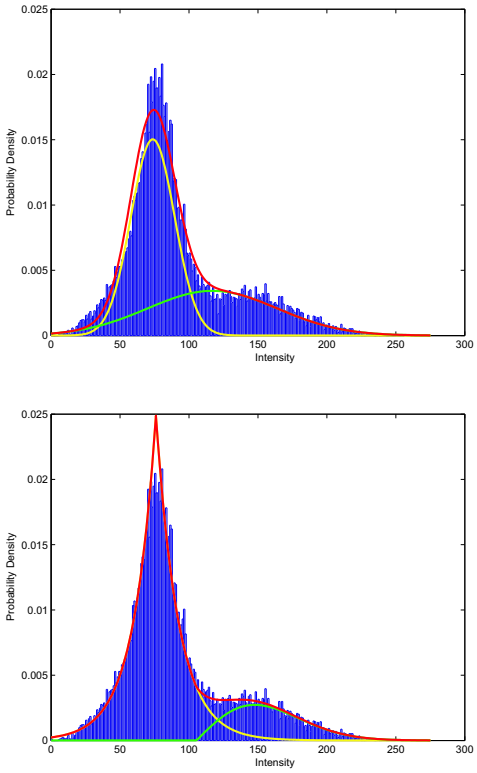

Fig. 2 Example of PDF fitting with a subject but no objects in the scene. The top picture shows the output of the ML algorithm for a mixture of 2 Gaussians while the bottom picture is obtained for a Laplacian and a Rayleigh.

3 Classification for MMW Images

3.1 Identifying sequences containing threats

The presence of metallic objects changes the maximum temperature recorded significantly, providing a good criterion to identify frames containing threats. Within a sequence, the range of variation of the maximum image temperature provides a reliable measure of the presence of a threat when compared to a normalised threshold. However, detecting which frames in the sequence contain objects is more difficult.

3.2 Identifying frames containing threats

To solve the problem of identifying individual frames containing metallic objects we trained a standard Hidden Markov Model (HMM) to detect significant changes in maximum temperatures (i.e., image intensities). The advantage to employing a HMM in contrast to relying on a set of defined rules is that it will detect any type of anomaly between the observed data and the training set. The training set can be extended to deal with any specific problem or threat encountered. Ideas such as motion compensated filtering from the visible-wavelength image processing domain would not be appropriate due to the high levels of image speckle noise and shallow depth of field in the MMW images.

In our HMM implementation, the data is first quantised into 10 levels and the hidden field is composed of 2 states (threat, no threat). A Baum-Welch algorithm [27] is used for parameter estimation, and a Viterbi algorithm to determine the optimal state sequence. As an example, Figure 3 shows the maximum temperature signal for a sequence of 180 frames, and the corresponding frame classification.

3.3 Locating threat regions within frames

We now turn to the problem of locating the image region corresponding to a metallic object in frames classified as containing threats. We use Expectation-Maximisation (EM) to perform the necessary unsupervised clustering. The EM algorithm [28] uses ML to iteratively compute the PDF parameters until a convergence criterion is met. We initialise the mixture model to the one containing the

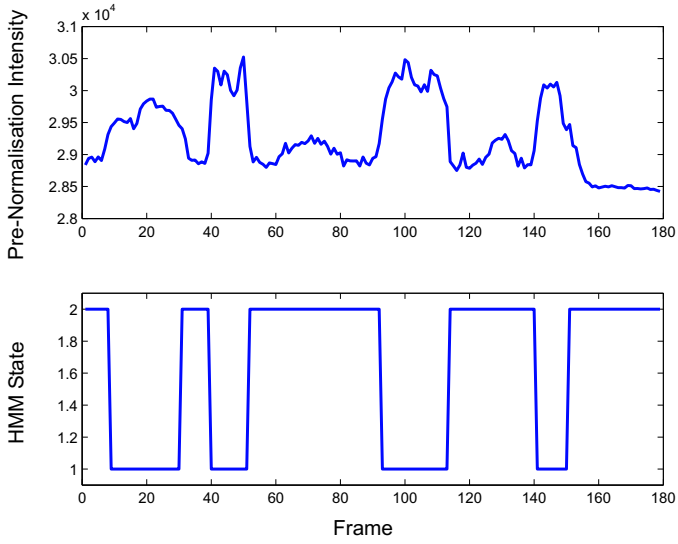

Fig. 3 An example of the HMM model being applied to a sequence of 180 frames. In the top row the maximum image intensity is shown for each frame in the sequence. In the bottom row the HMM state (1=object present, 2=no object) is shown across the sequence.

Fig. 4 Expectation-Maximisation classification (right) of a scene containing potential targets in a MMW image (left). The subject is carrying three small metal objects and the classification performs well in this example.

best-fitting distributions for the background-body-metal case (as defined in Section 2) with default parameters. Notice that this is not necessary for the EM algorithm, but improved the convergence speed significantly in our experiments. An example of threat location is shown in Figure 4, where the estimated threat regions are highlighted in white.

4 Tracking Threat Regions

4.1 Problem Definition

To make inference about a dynamic system, two models are needed: the system model which describes the evolution of state with time and the measurement model which relates the measurements to the state.

In the case where a target estimate is required every time a measurement is received, a recursive filtering approach is taken where these two models correspond to a prediction stage and a data update stage respectively. The prediction stage uses the system model to predict the state probability density function in the next time step and the update stage uses the measurement model to modify this density function using Bayes' Law.

4.2 Single Target Tracking

Let $x_{0..t}$ be the state sequence (x_t is a random vector representing the target state at time t) and $z_{1..t}$ be the sequence of measurements obtained. The tracking problem is governed by two functions:

$$x_t = F_t(x_{t-1}, v_{t-1}) \tag{2}$$

$$z_t = H_t(x_t, n_t) \tag{3}$$

where $v_{1..t}$ is the process noise sequence from the system model and $n_{1..t}$ is the measurement noise sequence. The process noise reflects the unknown target motion and the measurement noise reflects the sensor errors. The process and measurement noises are uncorrelated. Function F_t is a Markov Process on the state of the system and H_t is a function related to observing x_t. In Bayesian terms, the problem is to recursively calculate the belief of state x_t at time t given observations $z_{1..t}$.

The prior distribution $p_{t|t-1}(x_t|z_{1:t-1})$ of the target being in state x_t based on previous observations is:

$$p_{t|t-1}(x_t|z_{1:t-1}) = \int f_{t|t-1}(x_t|x_{t-1}) p_{t-1|t-1}(x_{t-1}|z_{1:t-1}) dx_{t-1}, \tag{4}$$

where $f_{t|t-1}(x_t|x_{t-1})$ represents the motion of the target and $p_{t-1|t-1}(x_{t-1}|z_{1:t-1})$ is the posterior distribution at time $t-1$.

When z_t has been observed, the posterior distribution at time t is obtained by Bayes' Law:

$$p_{t|t}(x_t|z_{1..t}) \propto g_t(z_t|x_t) p_{t|t-1}(x_t|z_{1..t-1}), \tag{5}$$

where $g_t(z_t|x_t)$ is the likelihood of observing z_t given target state x_t.

4.3 Multiple Target Tracking

The multiple target tracking problem is to estimate the positions of an unknown varying number of targets, based on observations which may also include false alarms due to erroneous measurements.

The multiple target state at time t is represented by random-set $X_t = \{x_{t,1}, ..., x_{t,T_t}\}$, where $x_{t,i}$ represents the state of an individual target and T_t is the number of targets at time t. The multiple target measurement at time t is given by $Z_t = \{z_{t,1}, ..., z_{t,m_t}\}$, where $z_{t,j}$ represents a single target measurement or false alarm and m_t is the number of observations at time t.

The tracking problem is then to estimate the unobserved states of the targets $X_{0:t} = \{X_0, ..., X_t\}$ based on observations $Z_{1:t} = \{Z_1, ..., Z_t\}$, i.e. to obtain $\hat{X}_t = \{\hat{x}_{t,1}, ..., \hat{x}_{t,\hat{T}_t}\}$, where $\hat{x}_{t,i}$ are the individual target estimates and $\hat{T}_t$ is the estimate of the number of targets at time t.

Analogous equations to the single-target tracking recursion can be found for the multiple-target model by using Finite Set Statistics [17].

4.4 PHD Filter

Instead of propagating the posterior distribution, as in Bayesian target tracking, the PHD (Probability Hypothesis Density) filter propagates the first-order moment of a multiple-target posterior distribution [17]. This property represents the expectation, whose integral over the state space is the expected number of targets. This has a significant computational advantage over propagating the multiple-target posterior as the time-complexity of calculating multiple-target likelihoods [29] grows exponentially with the number of targets whereas only single-target likelihoods are computed with the PHD filter. The PHD posterior is a multi-modal distribution whose modes represent the targets; these can be found by extracting the peaks of this distribution. In a similar manner to Bayesian tracking, we have prediction and update equations. These are given by

$$D_{t|t-1}(x|Z_{1:t-1}) \tag{6}$$
$$= \gamma_t(x) + \int \phi_{t|t-1}(x, x_{t-1}) D_{t-1|t-1}(x_{t-1}|Z_{1:t-1}) dx_{t-1},$$

$$D_{t|t}(x|Z_{1:t}) \tag{7}$$
$$= \left[\nu(x) + \sum_{z \in Z_t} \frac{\psi_{t,z}(x)}{\kappa_t(z) + \langle D_{t|t-1}, \psi_{t,z} \rangle}\right] D_{t|t-1}(x|Z_{1:t-1}),$$

where $\phi_{t|t-1}(x, x_{t-1}) = P_S(x_{t-1}) f_{t|t-1}(x|x_{t-1}) + b_{t|t-1}(x|x_{t-1})$, $\nu(x) = 1 - P_D(x)$, $\kappa_t(z) = \lambda_t c_t(z)$ and $\psi_{t,z} = P_D(x) g(z|x)$.

In the prediction equation, γ_t is the PHD for spontaneous birth of a new target at time t, $b_{t|t-1}$ is the PHD of a spawned target, P_S is the probability of target survival and $f_{t|t-1}(x_t|x_{t-1})$ is the single target motion distribution. In the data update equation, g is the single target likelihood function, P_D is the probability of detection, λ_t is the Poisson parameter specifying the expected number of false alarms and c_t is the probability distribution over the state space of clutter points. The $\langle ., . \rangle$ notation is defined as the inner product $\langle D_{t|t}, \varphi \rangle = \int D_{t|t}(x_t|Z_{1:t}) \varphi(x_t) dx_t$.

4.5 Particle PHD Filter Algorithm

The implementation of the PHD Particle filter used is an adaptation of the method described by Vo [22] based on a sequential Monte Carlo algorithm for multi-target tracking.

The algorithm can be informally described by the following stages: In the initialisation stage, particles are uniformly distributed across the field of view. The particles are propagated in the prediction stage using the dynamic model with added process noise and, in addition, particles are added to allow for incoming targets. When the measurements are received, weights are calculated for the particles based on their likelihoods determined by the distance of the particles to the set of observations (the sum of the weights gives the estimated number of targets). Particles are then resampled from the weighted particle set to give an unweighted representation of the PHD. Target locations are found by clustering the data using the estimated number of targets as the number of clusters and taking the centroids of the clusters.

The technical description of the algorithm is given here. The algorithm is initialised in Step 0 and then iterates through Steps 1 to 4.

At time $t = 0$:

Step 0:*Initialisation*

The filter is initialised with N_0 particles drawn from a prior distribution. The number of particles is adapted at each stage so that it is proportional to the number of targets, let N be the number of particles per target. The mass associated to each particle is $\hat{T}_0/N$, where $\hat{T}_0$ is the expected initial number of targets (this will be updated after an iteration of the algorithm).

•$\forall i = 1, .., N_0$ sample $x_0^{(i)}$ from $D_{0|0}$, set $N_1 = N_0$ and $t = 1$.

At time $t \geq 1$:

Step 1:*Prediction Step*

In the prediction step, the previous particles are moved using the transition function.

•$\forall i = 1, .., N_t$, sample $\tilde{x}_t^{(i)}$ from proposal density $q_t(.|x_{t-1}^{(i)}, Z_t)$.

•$\forall i = 1, .., N_t$, evaluate the predicted weights $\tilde{\omega}_{t|t-1}^{(i)}$:

$$\tilde{\omega}_{t|t-1}^{(i)} = \frac{\phi_t(\tilde{x}_t^{(i)}, x_{t-1}^{(i)})}{q_t(\tilde{x}_t^{(i)}|x_{t-1}^{(i)}, Z_t)} \tilde{\omega}_{t-1}^{(i)} \tag{8}$$

M_t new-born particles are also introduced from the spontaneous birth model to detect new targets entering the

state space.
•$\forall i = N_t+1,..,N_t+M_t$, sample $\tilde{x}_t^{(i)}$ from proposal density $p_t(.|Z_t)$.
•$\forall i = N_t+1,..,N_t+M_t$, compute the weights of new born particles $\tilde{\omega}_{t|t-1}^{(i)}$:

$$\tilde{\omega}_{t|t-1}^{(i)} = \frac{1}{M_t}\frac{\gamma_t(\tilde{x}_t^{(i)})}{p_t(\tilde{x}_t^{(i)}|Z_t)} \tag{9}$$

Step 2: *Update Step*
After the new measurements are obtained, the weights are recalculated using the likelihood function $g(.|.)$ to update the distribution based on new information.
Let $R_t = N_t + M_t$.
•$\forall z \in Z_t$, compute:

$$\langle \tilde{\omega}_{t|t-1}, \psi_{t,z} \rangle = \sum_{i=1}^{R_t} \psi_{t,z}(\tilde{x}_t^{(i)})\tilde{\omega}_{t|t-1}^{(i)}. \tag{10}$$

•$\forall i = 1,..,R_t$, update weights:

$$\tilde{\omega}_t^{(i)} = \left[\nu(\tilde{x}_t^{(i)}) + \sum_{z\in Z_t} \frac{\psi_{t,z}(\tilde{x}_t^{(i)}))}{\kappa_t(z) + \langle \tilde{\omega}_{t|t-1}, \psi_{t,z}\rangle}\right]\tilde{\omega}_{t|t-1}^{(i)}. \tag{11}$$

Step 3: *Resampling Step*
The particles are resampled to obtain an unweighted representation of $D_{t|t}$.
• Compute mass of particles:

$$\hat{T}_t = \sum_{i=1}^{R_t} \tilde{\omega}_t^{(i)} \tag{12}$$

• Resample $\left\{\frac{\tilde{\omega}_t^{(i)}}{\hat{T}_t}, \tilde{x}_t^{(i)}\right\}_{i=1}^{R_t}$ to get $\left\{\frac{\omega_t^{(i)}}{\hat{T}_t}, x_t^{(i)}\right\}_{i=1}^{R_t}$.

Step 4: *Target Extraction*
To find the target locations, an EM algorithm is used to fit a Gaussian Mixture Model of $\hat{T}_t$ components to the particle data. The estimated target locations are taken to be the means of the $\hat{T}_t$ Gaussians and $N_{t+1} = N\hat{T}_t$.

4.6 Time Complexity of PHD filter Tracker

The algorithm is initialised with N_0 particles drawn from a prior distribution, which requires $O(N_0)$ calculations. In the prediction step, N_t particles are sampled from one proposal distribution and M_t from a birth proposal distribution, which requires $O(N_t + M_t)$ calculations. In the update step, the weights are recalculated, requiring $O((N_t + M_t)|Z_t|)$ calculations. The resampling step requires $O(N_{t+1})$ calculations, where $N_{t+1} = N\hat{T}_t$ which depends on the estimated number of targets $\hat{T}_t$. We have used the EM algorithm here to estimate the target locations, which has quadratic time complexity in the estimated number of targets: $O(\hat{T}_t^2 Nn)$, where n is the number of iterations for the EM algorithm. Our experiments have shown that the k-means algorithm provides reliable results for target extraction [30] with a time complexity of $O(\hat{T}_t Nn)$, which is linear in the number of targets. The overall complexity of the algorithm we have used in this paper at each iteration is then $O(|Z_t|^2 Nn)$, as the estimated number of targets is bounded by the number of measurements. Mathematical proofs of convergence for the particle PHD filter indicate that the empirical measures represented by the particles converge to their true measures [31,32].

In comparison, the time complexity for $\hat{T}_t$ particle filters working independently would require $O(\hat{T}_t N)$ calculations plus the additional complexity of a pre-filter data association technique. Assuming the number of iterations for the PHD clustering algorithm, is not too large then the complexity of the PHD filter and multiple Particle Filters are theoretically comparable. Furthermore, the PHD filter represents a significant improvement on calculating the joint multi-target likelihoods whose complexity increases exponentially with the number of targets [29]. Therefore, from the perspective of time complexity, the PHD filter is a good choice for our application which, if deployed, would require real-time operation.

4.7 Data Association

The advantage of the PHD filter is that it has the ability to track a variable number of targets, estimating both the number of targets and their locations. It avoids the need for data association techniques as the identities of the individual targets are not required. Our purpose here is to identify the possible threat regions and not necessarily the same target throughout a sequence, although techniques for data association with the PHD filter have been developed recently and could be incorporated into this application. The first of these, by Panta *et al.* [33], used the PHD filter for pre-filtering the data as input to a Multiple Hypothesis Tracker, which is known to have a high time and space complexity. Another technique, by Lin [34], associated the estimated targets between frames by using a Kalman filter to estimate the target state in the next iteration and only considers existing targets but this relies on the linear Gaussian assumptions of the Kalman filter. New techniques have been developed specifically for the particle PHD filter [35] which require only the output of the PHD filter and the clustering algorithms to determine the target locations and do not rely on the linear assumptions of the Kalman filter, these will be presented in full in the near future.

4.8 Dynamic and Observation models

Using the current classification scheme (Section 3) the state vector can contain any of the following: centroid (position) ($\mathbf{x}$), velocity ($\dot{\mathbf{x}}$), area (ϕ) and average gray-level (intensity) ($\mathcal{I}$). The selection of the state vector is investigated in Section 5 but as an example, assuming use of position and area, the state vector contains the position, velocity and area of the target:

$$x_t = \left(\mathbf{x_t}\ \dot{\mathbf{x}}_\mathbf{t}\ \mathbf{y_t}\ \dot{\mathbf{y}}_\mathbf{t}\ \phi_t \right)^T. \tag{13}$$

For simplicity, we use a linear Gaussian dynamic model using the following standard state space model for a constant velocity model and random walk model for the area (see Bar-Shalom *et al* [18], for example):

$$x_t = \begin{pmatrix} 1 & T & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & T & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} x_{t-1} + \begin{pmatrix} T^2/2 & 0 & 0 \\ T & 0 & 0 \\ 0 & T^2/2 & 0 \\ 0 & T & 0 \\ 0 & 0 & T \end{pmatrix} v_{t-1}, \tag{14}$$

and observation model:

$$z_t = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} x_t + n_t. \tag{15}$$

v_t and n_t are the process and measurement noises respectively, which are uncorrelated. Noises are assumed Gaussian for simplicity but this is not a restriction of the algorithm as a Monte-Carlo Method is used. These models are used during both simulation and real MMW test sequences.

4.9 Simulation Results

We tested 3 sequences of 120 frames using the PHD filter with 2000 particles (ρ). All components of the process noise and observation noise were assumed to be independent with a process noise variance of 2 and observation noise variance of 60 employed in the filter. The results are shown in Figure 5, in which lines are real trajectories, dots measurements and circles estimated targets. Due to the nature of the simulations only positions and velocities are accounted for in the state vector which becomes 4-dimensional. In each case the Root Mean Squared Error (RMSE) and the error in the estimated number of targets are calculated. We also assume that the data association is always successful (in RMSE calculations, estimated targets are then compared to the nearest ground truth).

Numerical results can be seen in Table 1. In the first sequence the PHD Filter is used as a Particle Filter, i.e. there is one measurement per frame and Probability of Birth (P_B) of 0 and Probability of Survival (P_S) of 1. Deaths and births are then considered. A target vanishes 3 times for 10 iterations. $P_B = 0.1$ and $P_S = 0.9$ are chosen. In the third sequence targets are always detected and in the presence of clutter (3 false measurements on average generated by a Poisson process). $P_B = 0$ and $P_S = 1$.

Table 1 Results of PHD Filter tracking on simulated data

Sequences	RMSE	$< T_t >$	$< T_t - \hat{T}_t >$
Single, no clutter	6.64	1	0
Multi, with clutter	8.82	2	0.02
Single, deaths & births	7.16	0.75	0.77

It can be seen from the results (Table 1) that the PHD filter performs obviously well in very different situations. However, the tracker needs to be tuned carefully. In the three examples shown settings were chosen to optimise the performances. A trade-off that can cope with any tracking situation and be reasonably accurate must then be chosen.

5 Results on MMW Images

5.1 MMW Application

The PHD Filter developed in Section 4 is now applied to the output of the MMW classification process developed in Section 3. The classification process identifies, using the statistical model, all possible metallic objects in the scene. The centroid, area and average grey-level of each metallic object is then passed on to the PHD tracker. Due to the motion of the subject in the sensor the metallic objects are visible for an undefined but generally short period of time. It is, therefore, vital that the tracker is capable of rapidly detecting births and deaths within the set of tracked targets.

To evaluate our system, eleven real test sequences were employed, four with a subject without a threat (Clean), four with a subject carrying a single threat (Single) and three with a subject carrying multiple threats (Multi); providing a total of 3697 frames, including 656 frames where a threat was visible. Table 2 summarises the details of the test sequences. To allow comparison of accuracy for the PHD Filter, the ground truth for each sequence was manually tracked to an accuracy of approximately ±2-pixels. A simple nearest-neighbour data association stage has been employed for the accuracy calculations. For inclusion in a complete system, a more complex data association stage would be required. This will be investigated in future work.

The evaluation of results is divided into 5 subsections looking at classification (5.2), PHD tuning (5.3), PHD state vector selection (5.4) and finally PHD robustness (5.5).

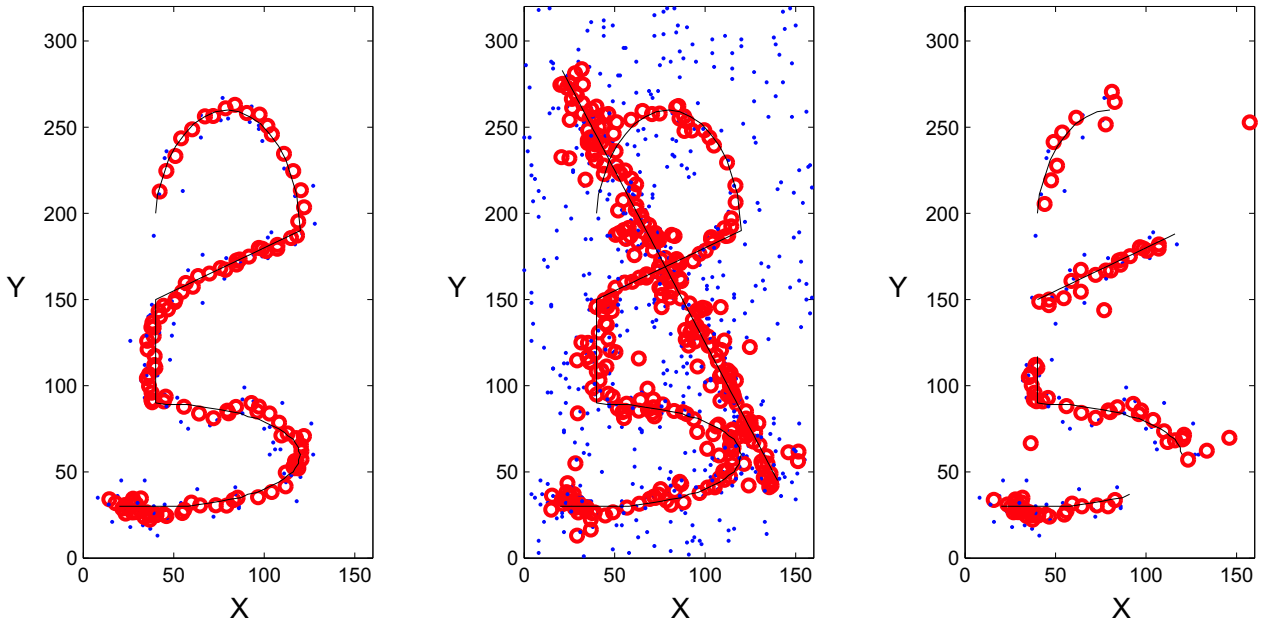

Fig. 5 Simulated tracking of targets using the PHD Filter. Lines are real trajectories, the dots are the measurements and the circles are the estimated targets. The left hand figure shows a simple case of a single target with no clutter. The centre figure shows two targets being tracked with clutter. The right hand target shows a single target disappearing and reappearing over time.

Table 2 Test Sequences Employed

Sequence	Frames	Threats	Threat Frames
Clean01	211	No	—
Clean02	252	No	—
Clean03	218	No	—
Clean04	236	No	—
Single01	242	1	24
Single02	155	1	27
Single03	179	1	56
Single04	136	1	30
Multi01	752	2	132
Multi02	731	2	206
Multi03	585	3	181
Total	3697	7/11	656

Table 3 Threat Identification

Sequence	Threat?	*Error*	E_{false}	E_{miss}
Clean01	No	—	—	—
Clean02	No	—	—	—
Clean03	No	—	—	—
Clean04	No	—	—	—
Single01	Yes	8%	0%	100%
Single02	Yes	3%	0%	100%
Single03	Yes	5%	22%	78%
Single04	Yes	8%	0%	100%

5.2 Classification Results

Table 3 shows the results of the sequence and frame threat identification algorithms described in Subsections 3.1 and 3.2, giving percentage error in classified frames (*Error*) with a breakdown of target frames missed (E_{miss}) compared to false alarms (E_{false}). The results clearly show that both stages of the threat identification perform very effectively. The missed target frames were primarily in situations where the target was identified through a combination of shape and intensity rather intensity alone. Shape-based targets can not be detected with the single intensity-based detector employed. However, future work is planned to address this, with the introduction of multiple detectors.

It should be noted that threat identification also worked for all three multi-target sequences. However, given multiple targets there is a much greater chance of at least one target presenting a metallic face perpendicular to the image plane, which produces the maximum image intensity. The task of identifying target frames is, therefore, considerably simpler and calculation of the *Error* metrics does not yield useful information.

After review of the classification output, illustrated by Figure 4, it is possible to identify a number of noise sources within the classification process. This is important to allow accurate PHD Filter set-up. The first noise source is inaccuracy within the classification with the low SNR of the system a major cause of this error. Further complexity is added by the rotating subject who

Fig. 6 Weapon tracking on a MMW image. The centroids of all possible metallic objects are shown as a "+" and estimated targets from the PHD Filter are shown as circles. Notice the considerably amount of clutter from the single-frame classification process.

produces dynamic lighting effects from the illumination chambers spot-lighting (discussed in Section 1). The second noise source is clutter, where bright spots caused, again, by the spot-lighting effect are mistaken for metallic objects and segmented. The position of this clutter moves unpredictably. An example of the clutter can be seen in Figure 6.

5.3 PHD Filter Parameter Tuning

The selection of the state vector and tuning of the subsequent noise covariances for the specific application is very important. Reviewing the dynamics and noise sources from the classification stage, presented in previous subsections, the PHD Filter is operating in a system with:

- Frequent target birth and death
- Average target lifespan of a few seconds (<50 frames)
- Uniform and slow target motion (<3 pixels / frames)
- Rapid and unpredictable changes in target area
- Significant false alarm rate

Using the current classification scheme the state vector can contain any of the following: centroid (position) ($\mathbf{x}$), velocity ($\dot{\mathbf{x}}$), area (α) and average grey-level (intensity) ($\mathcal{I}$). Due to the difficulty in assessing accurate ground-truth values, these components are investigated empirically in the following subsection (5.4). The noise covariance settings also have strong effects on both accuracy and flexibility of the filter. To summarise:

- The prediction covariance (C_{vt}) determines how far particles are propagated on average thanks to the dynamics equation. The goal is to keep all particles in bounded regions around measurements but nevertheless to avoid the phenomenon of degeneracy.
- The observation covariance (C_{nt}) determines the selectivity of the particle filter since its values appears in the computation of particle likelihoods. If the filter is too discriminative (i.e., a few particles get very high weights compared to others) degeneracy may occur.

The final covariance matrices were chosen, through detailed testing, to provide optimal performance across all test sequences. However, as a starting point for position and velocity using the knowledge of slow, uniform motion (<3 pixels / frame) the corresponding prediction variances were set to 10. Similarly, it is important that the peaks of targets do not overlap; so the minimum separation between targets was assumed to be 7 pixels and the corresponding observation variances were set to 50. The intensity changes were minimal and a constant state assumption proved sufficient. Finally, it is important to allow for the rapid and unpredictable area changes, forcing the related covariance components to be significantly greater than those of position and velocity.

As an example, suitable covariances matrices C_{vt} and C_{nt} of noises v_t and n_t for a state vector $(\mathbf{x}, \dot{\mathbf{x}}, \alpha)^T$ in our experiments:

$$C_{vt} = \begin{pmatrix} 10 & 0 & 0 \\ 0 & 10 & 0 \\ 0 & 0 & 100 \end{pmatrix} \tag{16}$$

$$C_{nt} = \begin{pmatrix} 50 & 0 & 0 \\ 0 & 50 & 0 \\ 0 & 0 & 1000 \end{pmatrix} \tag{17}$$

It was also empirically determined that the following parameters were suitable for the PHD Filter: $\rho = 2000$, $P_B = 0.1$ and $P_S = 0.9$. The robustness of the PHD tracking is crucial for the target application. For this reason, the effect on the PHD Filter to changes within the tuning parameters is investigated in Subsection 5.5.

An initial evaluation into the performance of the PHD Filter was carried out through comparison with a similarly set-up Particle Filter (PF). The results are shown for the four single-target sequences in Table 4. While it is possible to operate the PHD identically to the PF (Section 4) this has not been done in this case; this can be seen by the slight degradation in results. However, the PHD Filter is tracking accurately across the four single target scenes. At this stage no further comparison of the PHD filter's tracking accuracy can easily be presented. Any alternative multi-target tracking filter would be heavily dependent on the target estimation and data association steps, which would need to be specially developed for the target system. However, recent literature [24] has demonstrated the significant advantages

Table 4 PF / PHD Comparison

Sequences	PF RMSE	PHD RMSE
Single01	8.1	9.3
Single02	8.5	8.7
Single03	5.1	7.2
Single04	5.5	7.1

of the PHD Filter for both target estimation and tracking accuracy. It is hoped that an alternative method of estimating the number of targets can be presented for comparison.

5.4 Investigation of the PHD State Vector

The output from the current detection algorithm is the position, area and average intensity for each segmented region. It is important to determine the value of each of these pieces of information to the tracking filter. Table 5 shows four different combinations employed in the PHD state vector, giving the RMSE (in pixels). While it can be seen that using position and velocity leads to the most accurate results overall it also produces greatest error in the predicted number of targets. Introduction of area constrains the systems producing a minor reduction in accuracy but a significant improvement in target estimation. Figure 7 shows an example of four frames where area can improve target estimation. Finally, it can be seen that the inclusion of all three measures (fourth column of results) over-constrains the system leading a number of frames where no target is predicted, resulting in a negative $\overline{E_T}$ value.

For future practical applications, an accurate estimate of the number of targets will be extremely important to the overall function and computational performance. This is not apparent in our RMSE calculation using nearest-neighbour data association. It would become apparent in a threat decision / identification stage. For this reason, the chosen state vector at this stage is $(\mathbf{x}, \dot{\mathbf{x}}, \alpha)^T$. In the future, with the inclusion of multiple detectors, the choice of state vector will have to be re-examined.

5.5 PHD Robustness to Parameter Variation

To test the robustness of the PHD tracking filter to parameter values two experiments are run on the observation and prediction covariance matrices, respectively. The sequence Single04 has been chosen as a representative example containing some typical difficulties. It should be noted that the observation and prediction covariances used in other experiments were selected to perform well across all test sequences but are not optimal for this particular sequence.

Table 6 shows the tracking accuracy (RMSE) and target identification accuracy (i.e. percentage of frames

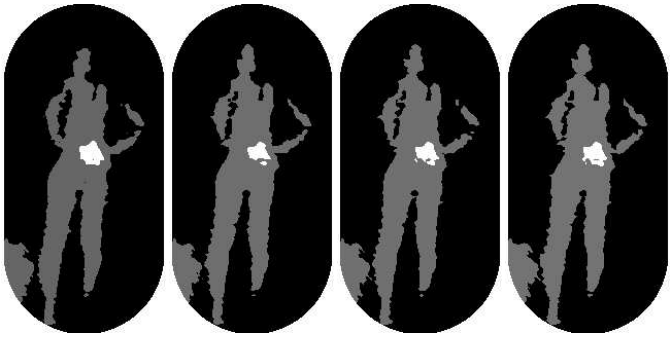

Fig. 7 Example frames demonstrating a situation where inclusion of area (α) in the state vector can provide a significant advantage.

Table 6 PHD Robustness - Observation Covariance

RMSE		var **x**			
Frames%		25	50	100	200
var α	500	7.1 *60%*	7.3 *80%*	8.4 *83%*	8.4 *93%*
	1000	6.1 *73%*	7.1 *97%*	8.0 *90%*	8.3 *97%*
	2000	5.0 *90%*	6.1 *90%*	5.7 *97%*	6.7 *97%*
	4000	3.9 *73%*	4.0 *90%*	4.7 *97%*	5.5 *97%*

Table 7 PHD Robustness - Prediction Covariance

RMSE		var **x**			
Frames%		5	10	20	40
var α	50	7.1 *97%*	7.3 *97%*	7.2 *97%*	6.5 *97%*
	100	6.7 *97%*	7.1 *97%*	7.0 *97%*	6.3 *97%*
	200	6.7 *97%*	6.7 *97%*	6.1 *97%*	5.8 *97%*
	400	6.1 *97%*	5.9 *97%*	5.9 *97%*	5.7 *97%*

the target was correctly identified in - *Frames%*) for the PHD Filter using the state vector $(\mathbf{x}, \dot{\mathbf{x}}, \alpha)^T$ across a range of values for the observation covariance components. It can be seen that the tracking accuracy remains excellent throughout despite a 64-fold change in the ratio of the observation covariance components. The effect on target identification accuracy shows a greater variation but has minimal effect on the tracking accuracy. This demonstrates that the overall tracking accuracy is good even in situations where the PHD filter is not optimally tuned and is missing targets in some frames.

Similarly, Table 7 shows the tracking accuracy and target identification accuracy for the PHD Filter across a range of values for the prediction covariance components. It can be seen that both the tracking accuracy and target identification accuracy remain excellent throughout with very little change in overall performance.

5.6 MPEG Video Results

Three MPEG video sequences showing the completed system operating on test sequences Single03, Single04 and Multi03 have been provided. The video sequences show the incoming MMW video sequence, on the right, and the classified image sequence, on the left. The border

Table 5 Selection of the PHD state vector, where $\mathbf{x}$ is the position, $\dot{\mathbf{x}}$ is the velocity, α is the area and $\mathcal{I}$ is the intensity. The RMSE is in pixels and is calculated between the ground truth and the estimated target. The average error in the estimated number of targets (T_t) is given by $\overline{E_T}$.

	$(\mathbf{x}, \dot{\mathbf{x}})^T$		$(\mathbf{x}, \dot{\mathbf{x}}, \alpha)^T$		$(\mathbf{x}, \dot{\mathbf{x}}, \mathcal{I})^T$		$(\mathbf{x}, \dot{\mathbf{x}}, \alpha, \mathcal{I})^T$	
Sequence	RMSE	$\overline{E_T}$	RMSE	$\overline{E_T}$	RMSE	$\overline{E_T}$	RMSE	$\overline{E_T}$
Single01 *t1*	6.9	*1.4*	9.3	*0.0*	7.0	*1.3*	9.3	*-0.3*
Single02 *t1*	7.5	*1.4*	8.7	*-0.3*	6.8	*1.3*	8.4	*-0.4*
Single03 *t1*	7.3	*0.2*	7.2	*0.0*	7.4	*0.2*	7.6	*0.0*
Single04 *t1*	4.1	*0.1*	7.1	*0.0*	4.1	*0.1*	7.4	*-0.1*
Multi01 *t1*	6.0	*0.4*	4.3	*0.2*	5.9	*0.4*	4.6	*0.2*
t2	5.3		4.7		5.5		4.7	
Multi02 *t1*	4.7	*0.1*	5.2	*0.0*	4.7	*0.1*	5.1	*0.0*
t2	2.0		1.6		2.1		1.7	
Multi03 *t1*	4.2	*0.3*	4.6	*-0.1*	4.2	*0.3*	4.7	*-0.1*
t2	5.2		8.1		5.1		8.9	
t3	3.0		5.1		3.1		5.7	
$\sum$RMSE, $\sum \|\overline{E_T}\|$	5.6	*0.5*	7.0	*0.1*	5.6	*0.4*	7.3	*0.1*

colour changes from black to white to indicate when a possible target has entered the scene. All material classified as as metal has the centroid marked as a blue cross (+). The output from the PHD Filter is then overlaid back onto the original MMW video sequence, with each detected target being shown as a red circle.

The video sequences demonstrate the excellent performance of the system and are indicative of the complete data test set. The round object present in the classified images (left) for sequences Single03 and Single04 is part of the MMW sensor accidentally placed close to the field-of-view.

For these MPEG video sequences, the average processing time per frame, calculated across all sequences containing threats, was 0.90 seconds per frame on a Intel Pentium 4 (2.66GHz) running Fedora Core 4. This breaks down to 0.79 seconds for classification and 0.11 seconds for PHD filter tracking. For comparison, the Regularised PF took an average of 0.04 seconds for a single target. The complete system was written in C++ but no attempt has been made to optimise the computational performance of the code. Real-time operation of this system, with the MMW imager working at 12 fps, would be possible given suitable optimisation of the software and selection of an appropriate hardware platform.

6 Conclusions and future work

We have presented a novel system for the automatic detection and tracking of metallic objects concealed under clothes using MMW sequences. The recent emergence of MMW video imaging for security and surveillance applications makes our work very timely. To the best of our knowledge, the only previous system attempting employ image processing for automatic threat detection in MMW [8] used static images in a highly restricted environment. We have demonstrated a system employing advanced image processing on images from a real-time head-to-toe MMW imager. The system is capable of automatically detecting the presence of metallic objects within the scene, classifying the materials in the scene and tracking any targets detected in the scene.

Our metallic object detection, employing a HMM, has proved reliable on the current data test set. While the material classification has performed accurately and robustly across the data test set, with no target missed due to classification errors. The PHD multi-target tracking filter has proved reliable both in detecting the number of targets in the scene and tracking the targets position. The PHD Filter also provides a natural fusion framework to combine multiple detectors therefore increasing reliability and robustness in the detection of concealed weapons. The selection of the PHD filter has been justified from a theoretical time complexity viewpoint and its suitability for real-time software implementation demonstrated.

Future work will extend our approach to multiple-detectors designed to operate on a wider range of target materials and to incorporating human body models to improve tracking and provide 3-D visualisation. There are significant concerns over privacy protection issues related to the use of MMW imagers and similar technologies. The combination of human body model and the automatic system presented in this paper could enable public use of the system in the future. Preliminary work has also been undertaken [9] on a model of MMW image formation and further work will be reported on this in the future.

Acknowledgements The authors would like to acknowledge the support of QinetiQ, in particular Roger Appleby, Christina Brooks and Peter Coward. This work is supported by EPSRC Research Grant GRS/68088 “ATRIUM” (Automatic Threat Recognition and Identification Using MMW).

References

1. G. N. Sinclair, R. N. Anderton, and R. Appleby, "Outdoor passive millimetre wave security screening," in *Proceedings of the 35th International Carnahan Conference on Security Technology*, (London, UK), pp. 172–179, IEEE, October 2001.
2. K. S. J. Murphy, R. Appleby, G. Sinclair, A. McClumpha, K. Tatlock, R. Doney, and I. Hutcheson, "Millimeter wave aviation security scanner," in *Proceedings of the 36th International Carnahan Conference on Security Technology*, (London, UK), pp. 162–166, IEEE, October 2002.
3. G. N. Sinclair, P. R. Coward, R. N. Anderton, R. Appleby, T. Seys, and P. Southwood, "Detection of illegal passengers in lorries using passive millimeter wave scanner," in *Proceedings of the 36th International Carnahan Conference on Security Technology*, (London, UK), pp. 167–170, IEEE, October 2002.
4. P. Coward and R. Appleby, "Development of an illumination chamber for indoor millimetre-wave imaging," in *Passive Millimeter-wave Imaging Technology VI and Radar Sensor Technology VII* (R. Appleby, D. A. Wikner, R. Trebits, and J. L. Kurtz, eds.), vol. 5077 of *Proceedings of SPIE*, pp. 54–61, SPIE, August 2003.
5. M.-A. Slamani, P. K. Vershney, R. M. Rao, M. G. Alford, and D. Ferris, "Image processing tools for the enhancement of concealed weapon detection," in *Proceedings of the International Conference on Image Processing*, vol. 3, (Kobe, Japan), pp. 518–522, IEEE, October 1999.
6. P. Keller, D. L. McMakin, D. M. Sheen, A. D. McKinnon, and J. W. Summet, "Privacy algorithm for cylindrical holographic weapons surveillance system," *Aerospace and Electronic Systems Magazine*, vol. 15, pp. 17–24, February 2000.
7. M.-A. Slamani and D. D. Ferris Jr., "Shape-descriptor-based detection of concealed weapons in millimeter-wave data," in *Optical Pattern Recognition XII* (D. P. Casasent and T.-H. Chao, eds.), vol. 4387 of *Proceedings of SPIE*, pp. 176–185, SPIE, 2001.
8. M.-A. Slamani, P. K. Varshney, and D. D. Ferris Jr., "Survey of image processing techniques applied to the enhancement and detection of weapons in mmw data," in *Passive Millimeter-Wave Imaging Technology VI*, vol. 4719B, pp. 296–305, SPIE, 2002.
9. C. D. Haworth, B. G. González, M. Tomsin, R. Appleby, P. Coward, A. Harvey, K. Lebart, Y. Petillot, and E. Trucco, "Image analysis for object detection in millimetre-wave images," in *Passive Millimetre-wave and Terahertz Imaging and Technology* (R. Appleby, J. M. Chamberlain, and K. A. Krapels, eds.), vol. 5619, pp. 117–129, SPIE, 2004.
10. P. K. Varshney, H.-M. Chen, L. C. Ramac, M. Uner, D. Ferris, and M. Alford, "Registration and fusion of infrared and millimeter wave images for concealed weapon detection," in *Proceedings of the International Conference on Image Processing*, vol. 3, (Kobe, Japan), pp. 532–536, IEEE, October 1999.
11. Z. Xue and R. S. Blum, "Concealed weapon detection using color image fusion," in *Proceedings of the 6th International Conference on Information Fusion*, vol. 1, (Queensland, Australia), pp. 622–627, IEEE, July 2003.
12. E. Trucco and K. Plakas, "Video tracking: a concise survey," *IEEE Journal of Oceanic Engineering*, vol. 30, p. (to appear), 2005.
13. M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking," *IEEE Transactions on Signal Processing*, vol. 50, pp. 174–188, February 2001.
14. T. Tommasini, A. Fusiello, E. Trucco, and V. Roberto, "Making good features track better," in *Proceedings of the International Conference on Computer Vision and Pattern Recognition*, pp. 178–183, IEEE, June 1998.
15. D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 24, pp. 603–619, May 2002.
16. D. Comaniciu, V. Ramesh, and P. Meer, "Kernel-based object tracking," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 25, pp. 564–577, May 2003.
17. R. P. S. Mahler, "Multitarget Bayes filtering via first-order multitarget moments," *IEEE Transaction on Aerospace and Electronic Systems*, vol. 39, pp. 1152 – 1178, October 2003.
18. Y. Bar-Shalom and T. E. Fortmann, *Tracking and Data Association*. Academic Press, 1988.
19. C. Hue, J.-P. L. Cadre, and P. Perez, "Tracking multiple objects with particle filtering," in *IEEE Transactions on Aerospace and Electronic Systems*, vol. 38, pp. 791 – 812, IEEE, July 2002.
20. R. Karlsson and F. Gustafsson, "Monte Carlo data association for multiple target tracking," *Target Tracking: Algorithms and Applications (Ref. No. 2001/174), IEE*, vol. 1, pp. 13/1– 13/5, 2001.
21. D. Schulz, W. Burgard, D. Fox, and A. B. Cremers, "People tracking with a mobile robot using sample-based Joint Probabilistic Data Association Filters," *International Journal of Robotics Research*, pp. 99–116, 2003.
22. B. Vo, S. Singh, and A. Doucet, "Sequential Monte Carlo Implementation of the PHD filter for Multi-target Tracking.," *Proc. FUSION 2003*, pp. 792–799, 2003.
23. T. Zajic and R. Mahler, "A particle-systems implementation of the PHD multitarget tracking filter," *SPIE Vol. 5096 Signal Processing, Sensor Fusion and Target Recognition*, pp. 291–299, 2003.
24. H. Sidenbladh, "Multi-target particle filtering for the probability hypothesis density," in *Sixth International Conference on Information Fusion*, (Cairns, Australia), pp. 800–806, 2003.
25. M. Tobias and A. D. Lanterman, "A Probability Hypothesis Density-based multitarget tracker using multiple bistatic range and velocity measurements," in *Proceedings of the Thirty-Sixth Southeastern Symposium on System Theory*, pp. 205–209, March 2004.
26. D. Clark and J. Bell, "Bayesian Multiple Target Tracking in Forward Scan Sonar Images Using the PHD Filter," *IEE Radar, Sonar and Navigation*, vol. 152, pp. 327–334, October 2005.
27. L. R. Rabiner, "A tutorial on Hidden Markov Models and selected applications in speech recognition," *Proceedings of the IEEE*, vol. 77, pp. 257–285, February 1989.

28. K. N. Choi, M. Carcassoni, and E. R. Hancock, "Recovering facial pose with the EM algorithm," *Pattern Recognition*, vol. 35, pp. 2073–2093, 2002.
29. D. Hall and J. Llinas, eds., *Handbook of Multisensor Data Fusion*, ch. 7. CRC Press, 2001.
30. D. E. Clark, J. Bell, Y. de S.-Pern, and Y. Petillot, "Phd Filter Multi-target Tracking in 3D Sonar," in *Oceans Europe Conference*, (Brest), IEEE, June 2005.
31. D. E. Clark and J. Bell., "Convergence Results for the Particle PHD Filter," *Accepted to appear in IEEE Transactions on Signal Processing*.
32. A. M. Johansen, S. S. Singh, A. Doucet, and B.-N. Vo, "Convergence of the SMC implementation of the PHD filter," *Technical Report CUED/F-INFENG/TR-517, University of Cambridge*, 2005.
33. K. Panta, B. Vo, S. Singh, and A. Doucet, "Probability hypothesis density filter versus multiple hypothesis tracking," in *Signal Processing, Sensor Fusion, and Target Recognition XIII* (I. Kadara, ed.), vol. 5429, pp. 284–295, SPIE, August 2004.
34. L. Lin, "Parameter estimation and data association for multitarget tracking," *PhD Thesis, The University of Connecticut*, 2004.
35. D. E. Clark and J. Bell., "Data Association for the PHD Filter," *Accepted to appear in ISSNIP 2005, 5-8 December, Melbourne*.