Millimetre-Wave Personnel Scanners for Automated Weapon Detection

Beatriz Grafulla-González[1], Christopher D. Haworth, Andrew R. Harvey, Katia Lebart, Yvan R. Petillot, Yves de Saint-Pern, Mathilde Tomsin, Emanuele Trucco

Heriot-Watt University, Edinburgh, UK

Abstract

The ATRIUM project aims to the automatic detection of threats hidden under clothes using millimetre-wave imaging. We describe a simulator of realistic millimetre-wave images and a system for detecting metallic weapons automatically. The latter employs two stages, detection and tracking. We present a detector for metallic objects based on mixture models, and a target tracker based on particle filtering. We show convincing, simulated millimiter-wave images of the human body with and without hidden threats, including a comparison with real images, and very good detection and tracking performance with eight real sequences. (International Workshop on Pattern Recognition for Crime Prevention, Security and Surveillance)

1 Introduction

The ATRIUM project (Automatic Threat Recognition and Identification Using Millimetre-waves) emerges from the necessity to protect public environments such as airports, train stations and other public buildings. The project investigates the use of a millimetre-wave (henceforth MMW) imaging sensor combined with image processing techniques for detecting threats hidden under clothes.

We present our current work and results on two image processing approaches: an *image-based strand*, whereby image intensities are analysed without reference to the physics of image formation; and a *physics-based strand*, whereby the image is analysed on the basis of a physical model of MMW image formation.

Specifically, we report a system for the automatic detection and tracking of metallic objects concealed on moving people in sequences of MMW images (image-based strand), and a complete model for the formation and simulation of MMW images (physics-based strand).

We adopt QinetiQ's recently demonstrated, proof-of-concept sensor, providing video-frame sequences with near-CIF resolution (320 $\times$ 240 pixels). It can image through clothing, plastics and fabrics. Together, through-clothes imaging and current video sequence analysis offer huge potential for automatic, covert detection of weapons concealed under clothes.

Existing MMW simulation packages, e.g., PMWCM or Speos, are not designed for indoors scenes and do not model specular reflections or bulk emission, which makes them unsuitable for ATRIUM. We have therefore developed a novel simulator combining a ray-tracer (Zemax) and low-level Matlab modules to process ray-tracing data. The system models both the geometry *and* the physical parameters of the scene objects, allowing full simulation of MMW images and sequences of indoors scenes.

To our best knowledge, very little work has been reported on the automatic analysis of MMW sequences or images. Most authors focus on very basic segmentation [1] or image fusion. In a related application, shape identification on segmented images [1] has been investigated and suitable shape descriptors proposed. More recently, basic work on object detection has appeared [2]. The main contribution of our work is therefore to apply advanced image processing techniques to a new video imaging technology of high potential for public security.

This paper is divided into three parts: first, an introduction to MMW images focusing on image formation and simulation; second, a description of our metallic-threat detection algorithm; third, experimental results for both simulation and detection.

[1] E-mail: bg5@hw.ac.uk, Telephone: +(44) 0131 451 3299

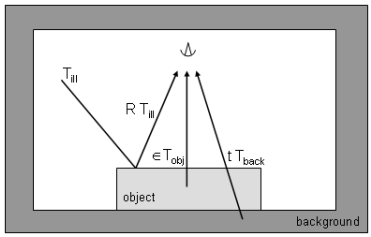


Figure 1: Illustration of the combination of power from various scene components.

2 Millimetre-wave images

2.1 Formation of millimetre-wave images

Two different phenomena influence the formation of MMW images; (1) the combination of signal power from various components of the scene and (2) the modification of the recorded signal by the instrument response, including the impulse response of the imaging sensor and noise artifacts such as, for instance, those due to scanning and interpolation.

The radiation frequency used is $f = 35$ GHz or equivalently $\lambda \approx 9$ mm in wavelength. The focal length of the imager is 0.8 m resulting in a diffraction-limited spot-size of ~ 2 cm. We consider only short-range indoor scenes with mostly incoherent illumination [3].

The temperature of the objects in the scene is above absolute zero, so that scene objects radiate power in the MMW range with an emissivity ϵ compared to the radiation of an ideal black body. Since the surfaces of body and threats are flat at the scale of the illumination wavelength, reflections are considered specular [4]. This implies that scattering effects are small and light propagation within the scene obeys ray optics approximations. The intensity of MMW radiation at each pixel is determined by three different contributions: self-emission by scene components, reflections from illumination source and background radiation. We consider illumination as ambient background black body radiation with temperature around 290 K and by extended diffuse sources with high equivalent temperatures of, typically, 800 K. The reflectivity, R, the emissivity, ϵ, and the transmissivity t are related through Equation (1) [5].

$$R + \epsilon + t = 1 \tag{1}$$

These three coefficients depend on the physical characteristics of materials and geometrical aspects of the scene defined via the dielectric constant ε, the permeability μ, the angle of incidence θ_i, the angle between the electric field and the plane of incidence α, and the polarization p (horizontal or vertical). Note that these coefficients can be expressed as a sum of their "projections" on the horizontal and vertical polarization planes, as described in Equation (2) for the power reflectivity [6]:

$$R(\varepsilon, \mu, \theta, \alpha) = R_p(\varepsilon, \mu, \theta)\ \cos^2 \alpha + R_s(\varepsilon, \mu, \theta)\ \sin^2 \alpha, \tag{2}$$

where $R_p(\varepsilon, \mu, \theta)$ is the power reflectivity in P-polarisation and $R_s(\varepsilon, \mu, \theta)$ the power reflectivity in S-polarisation. Similar equations can be obtained for transmissivity and emissivity.

Since the source is incoherent, the three intensity coefficients in Equation (1) are added (Fig. 1), obtaining Equation (3). This describes the temperature received by the sensor [4]:

$$T_{rec}(\varepsilon, \mu, \theta, \alpha) = R\ T_{ill} + \epsilon\ T_{obj} + t\ T_{back}, \tag{3}$$

where $T_{rec}(\varepsilon, \mu, \theta, \alpha)$ is the received temperature, T_{ill} the temperature of the illumination, T_{obj} the temperature of the object and T_{back} the temperature of the background. T_{ill}, T_{obj} and T_{back} are constant values.

Finally, we assume a highly incoherent illumination source to allow Equation (3). However, it is important to remark that the source has some residual spatial and temporal coherence, resulting in low-level speckle noise (see Subsection 4.1).

2.2 Simulation of millimetre-wave images

As we consider indoors scenes with incoherent illumination, MMW image formation is modeled as superposition of the components in Equation (3). To generate synthetic MMW images we must model these components as well as relevant sensor effects.

In envisaged security applications, the scene is composed by a person possibly carrying one or several threats (weapons, knives, explosives, etc). The relevant physical characteristics of scene objects are the dielectric constant ε_{obj}, the permeability μ_{obj} (which is unity for non-magnetic materials) and the physical temperature T_{obj}. The value of the dielectric constant depends on the type of material: for instance, flesh is well approximated by salty water with a dielectric constant of $\varepsilon_{body} \approx 28 + i\,34$. The body geometry is modeled by a triangular mesh, and threats as metallic or dielectric patches located on the body. Table 1 shows the reference dimensions of a typical scene.

Table 1: Dimensions of the scene

Description	**Length (m)**
Range object - closest part of the imager	1.6
Range object - apeture	2.9
Focal length	0.8
Lens diameter	1.6
Height of the body	1.8
Dimensions of the threat	0.075×0.075

The MMW-image simulator is composed of two parts. A ray-tracing programme, *Zemax*, is used to propagate rays back from each detector pixel via reflections from scene components to the source (either hot or background). All reflections within the scene are characterised using *Zemax* and the history of every ray is stored as a text file. By repeated application of Equation (3) at each intercept of a ray with a scene component, we can calculate the intensity of MMW radiation incident on each pixel detector. This component is executed by MATLAB code reading the *Zemax* output text file. The code calculates the equivalent temperatures at each pixel. Convolution of these images with the imager's point spread function (Airy disk) and the addition of random, low-level speckle noise to images yields the final, simulated MMW image.

The ray-optics model used here is strictly valid only for large scene dimensions compared to the wavelength (a few centimeters); more accurate models are requird otherwise, e.g., electromagnetic methods used with unbounded problems such as the Boundary Element Method. When imaging people, the ray-optics model will provide accurate results for large body parts with large radii of curvature, but some inaccuracies might be expected for smaller features such as fingers and details of the face. Although these inaccuracies may become noticeable with future improvements in detector technology, they are currently unobservable due to noise levels and discrepancies between model parameters and real values. The salient advantage of the ray optics model is computation speed, crucial when simulating the large numbers of video sequences required for training automated detection algorithms.

2.3 Mixture models for MMW images

MMW images offer good data for material discrimination as different materials yield, generally speaking, different image properties. We model such differences statistically using a weighted mixture model, in which each pdf, f_i, is associated to a specific material:

$$f_{mix} = \sum_{i=1}^{N} \alpha_i f_i(\boldsymbol{\theta}). \tag{4}$$

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

To identify the optimal 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-Squared 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 an optimal mixture model for each material (types of component distributions, and parameters).

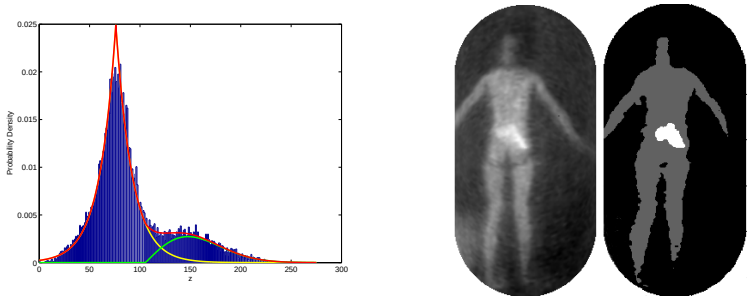


Figure 2: Example of threat location showing an example PDF classification (left) with a MMW image containing a potential target (centre) and the Expectation-Maximisation classification (right).

3 Automatic detection of metallic objects

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

We trained a standard Hidden Markov Model (HMM) to detect significant changes in maximum temperatures (i.e., image intensities). The data is first quantised into 10 levels and the hidden field is composed of two states (threat, no threat). A Baum-Welch algorithm [7] is used for parameter estimation, and a Viterbi algorithm to determine the optimal state sequence.

3.3 Locating threat regions within frames

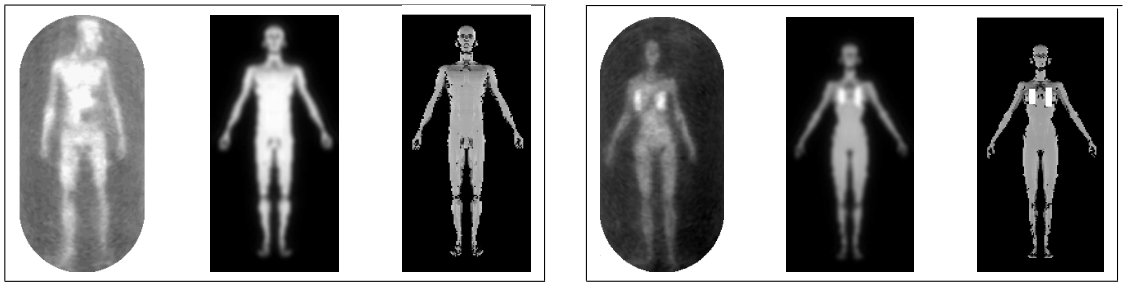
In frames classified as containing threats, we use Expectation-Maximisation (EM) to perform the necessary unsupervised clustering. The EM algorithm uses ML to recompute the pdf parameters until a convergence criterion is met. We initialise the mixture model to the one containing the optimal distributions for the background-body-metal case (as defined in Subsection 2.3) with default parameters. Although not strictly necessary, this improves the convergence speed significantly. An example of threat location is shown in Fig. 2, with an example PDF classification (left), original image (centre) and classified image (right).

3.4 Tracking threat regions

The previous classification stage yields two results: a set of frames showing metal threats, plus, in each such frame, the regions corresponding to threats. Such regions are characterised by frame number, centroid location, and area. The problem is now to track such regions throughout a sequence for as long as the region remains visible, with frequent births, deaths and temporary occlusions. The problem is made more difficult by the noisy nature of the MMW images, making accurate segmentation difficult.

Tracking objects in visible-wavelength sequences is a well-studied problem in image processing and computer vision [8]. Particle Filters (PF) [9] are a powerful class of algorithms removing the Gaussian constraint typical of Kalman filters. They also provide robustness against clutter, a significant problem in MMW images. Common to PF is the *degeneracy problem* whereby all but a few particles have negligible weights after several iterations. For this reason we employ a Regularised PF (RPF) [9] which has an improved re-sampling stage, helping to avoid the degeneracy problem.

The tracking filter was used with a state vector containing the position, velocity and area of the target: $(x, \dot{x}, y, \dot{y}, \phi)^T$. Suitable values for the prediction and observation covariance matrices were determined experimentally. Due to the nature of the segmentation, it is necessary to allow greater variance within the area measurements than for the position estimate.



(a) (b)

Figure 3: Real and synthetic images: (a) man carrying three objects; (b) woman carrying two metal patches. For each box: (left) real image, (middle) simulated image including convolution with point spread function and (right) intensity image (power distribution of the scene, i.e. the received temperature at the input of the sensor for each pixel).

Table 2: Threat Identification & Target tracking

Sequence	Frames	Threat?	$Error$	E_{false}	E_{miss}	Average Targets	RMSE
Plain01	211	No	—	—	—	—	—
Plain02	252	No	—	—	—	—	—
Plain03	218	No	—	—	—	—	—
Plain04	236	No	—	—	—	—	—
Threat01	242	Yes	8%	0%	100%	2.4	8.1
Threat02	155	Yes	3%	0%	100%	2.1	11.6
Threat03	179	Yes	5%	22%	78%	1.3	5.1
Threat04	136	Yes	8%	0%	100%	1.1	5.5

4 Results

4.1 Simulation of MMW images

Fig. 3 presents the images generated by the MMW simulator. As can be observed, the differences in terms of grey levels range between synthetic and real images are minimal except for specific situations. For instance, the man in Fig. 3(a) carries three different objects: one below the chest, one on the abdomen and one on the (left) knee. These objects are not represented in the simulated scene and therefore no shading areas appear in the synthetic image.

In the case of the woman in Fig. 3(b), no big difference can be appreciated in terms of grey level range and shape, even if the mesh model does not fit perfectly the real body. However, as can be seen, images are not exactly the same. This is due to imperfections in the real sensor, the simulator and the illumination which is slightly different in real (highly incoherent) and synthetic (completely incoherent) scenes. But the noticeable difference is noise as it has not been included in the simulator yet. The study of noise distributions in real images as well as its addition into synthetic images remain as a future task.

4.2 Automatic detection of metallic objects

We tested our system with eight real sequences, four with subjects without a threat and four with subjects carrying a threat, giving a total of 1629 frames of which 137 frames where a threat was visible (see summary in Table 2, columns 1-3).

Table 2 (columns 4-6) shows the results of the sequence and frame threat identification described in Section 3, 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. Missed target frames occurred primarily when targets were identified through shape rather than intensity.

Table 2 (columns 7 & 8) shows results for EM classification and RPF target tracking, giving the average number of targets (true target + clutter) per frame for the sequence and RMSE of the tracked position. The ground truth for the target position was established manually and is accurate

to ±2 pixels. It can be seen that very good target tracking has been achieved, even in the sequences with considerable clutter (Threat01, Threat02). The comparatively poorer tracking results seen in Threat02 are due to the very short time span over which the threat is visible (approx. 9 frames on each occasion compared to an average of 15 frames for other sequences). In this instance, the particle filter does not have enough time to converge.

5 Conclusion

We have described the formation and simulation of MMW images as well as an automatic system to detect and track metallic threats concealed on people. Initial results show that the geometric and physical models deployed yield good-quality MMW sequences compared to real ones. We have also demonstrated an automatic system for metallic threat detection, showing good performance with eight real sequences in field conditions. Key future work will address the inclusion of low-level speckle noise, non-uniform sampling and post-processing of images for the simulator, and of a wider range of materials, more complex tracking scenarios, and human body models for tracking. Further work will concern 3-D visualisation techniques preserving privacy.

6 Acknowledgements

The authors would like to acknowledge the support of Qinetiq. Beatriz Grafulla-González and Christopher D. Haworth are supported by EPSRC Research Grant GRS/68088 "ATRIUM" under the Think Crime programme.

References

[1] M.-A. Slamani, P. K. Varshney, and D. D. Ferris, "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. SPIE, 2002, pp. 296–305.

[2] 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. SPIE, 2004, pp. 117–129.

[3] P. Coward and R. Appleby, "Development of an illumination chamber for indoor millimetre-wave imaging," in *SPIE Proceedings, Passive Millimeter-Wave Imaging Technology VI and Radar Sensor Technology VII*, vol. 5077, 2003, pp. 54–61.

[4] G. N. Sinclair, R. Appleby, P. Coward, and S. Price, "Passive millimetre wave imaging in security scanning," in *SPIE Proceedings, Passive Millimeter-Wave Imaging Technology IV*, vol. 4032, 2000, pp. 40–45.

[5] N. A. Salmon, R. Appleby, and S. Price, "Scene simulation of passive millimetre wave images of plastic and metal objects," *SPIE Passive Millimetre Wave Imaging Technology VI*, April 2002.

[6] M. Born and E. Wolf, *Principles of Optics*, 6th ed. Pergamon Press, 1987.

[7] L. R. Rabiner, "A tutorial on hidden markov models and selected applications in speech recognition," *Proceedings of the IEEE*, vol. 77, no. 2, pp. 257–285, February 1989.

[8] E. Trucco and K. Plakas, "Video tracking: a concise survey," *IEEE Journal of Oceanic Engineering*, vol. 30, no. 1, January 2005.

[9] 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, no. 2, pp. 174–188, February 2001.