Vehicle detection using multimodal imaging sensors from a moving platform

Christopher N. Dickson,[a,b] Andrew M. Wallace,[a] Matt Kitchin,[b] and Barry Connor[b]

[a]School of Engineering & Physical Sciences, Heriot Watt University, Riccarton, EH14 4AS, UK;

[b]Thales UK, 1 Linthouse Road, Glasgow, G51 4BZ, UK

ABSTRACT

A modular vehicle detection system, using a two-stage hypothesis generation (HG) and hypothesis combination (HC) approach is presented. The HG stage consists of a set of simple algorithms which parse multi-modal data and provide a set of possible vehicle locations. These hypotheses are subsequently fused in a combination stage. This modular design allows the system to utilise additional modalities where available, and the combination of multiple information sources is shown to reduce false positive detections. The system uses Thales' high-resolution long wave infrared polarimeter and a four-band visible/near infrared multispectral system. Vehicle cues are taken from motion flow vectors, thermal intensity hot spots, and regions with a locally high degree of linear polarisation. Results using image sequences gathered from a moving vehicle are shown, and the performance of the system is assessed with Receiver Operator Characteristics.

Keywords: thermal infrared, polarisation, multimodal, detection

1. INTRODUCTION

Multi-sensor fusion can be used to integrate complementary information from a variety of modalities to enhance the performance of tasks beyond the capabilities of a single sensor or modality.[1] Visible imagery is natural to the user, and has been the focus of many vehicle detection systems.[2,3] High resolution visible cameras are available at minimal cost and can provide further discriminatory information to the system. In addition, it has been shown that Near-Infrared imagery can be used to enhance standard visible imagery enabling processing techniques such as haze penetration,[4] and vegetation segmentation.[5] Jai's AD-080 GE 2-CCD multi-spectral imaging colour camera provides access to both the visible and near infrared (NIR) bands (1024×768 pixels, NIR at 750 nm to 1000 nm, 30 frames/s) through a single optical system.

Many Defence applications favour long-wave infrared sensors (LWIR, 8 µm to 12 µm) at which wavelength emission dominates the observed radiation, allowing 24-hour use and makes such sensors highly suited to locating thermally distinct objects from the background – even if visually camouflaged.

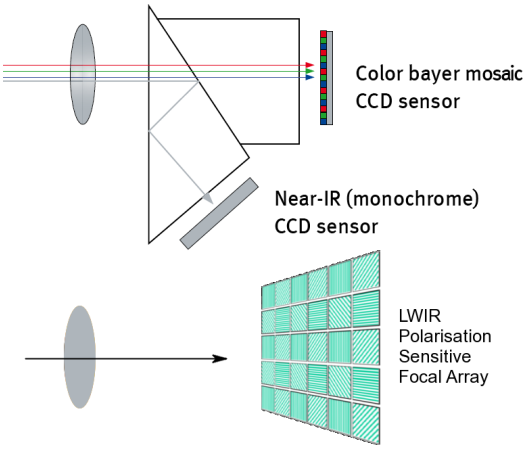
Standard imagery measures the intensity of radiation at certain wavelengths as a function of spatial coordinate. There is increasing interest in extending this to include measurements of the polarisation state of observed radiation, which has the potential to provide more information about the scene.[6–9] The complete polarisation state of a plane monochromatic wave can be characterised by the four Stokes parameters:[10]

$$\mathbf{S} = \begin{bmatrix} I \\ Q \\ U \\ V \end{bmatrix} = \begin{bmatrix} \langle a_1^2 \rangle + \langle a_2^2 \rangle \\ \langle a_1^2 \rangle - \langle a_2^2 \rangle \\ 2\langle a_1 a_2 \cos\delta \rangle \\ 2\langle a_1 a_2 \sin\delta \rangle \end{bmatrix} \tag{1}$$

where a_1 and a_2 are the instantaneous amplitudes of the two orthogonal components E_x, E_y of the electric vector, and $\delta = \phi_1 - \phi_2$ is their phase difference. The parameters I, Q, U and V may be interpreted as the total intensity, the amount of linear polarisation in the horizontal direction, the amount of linear polarisation in a plane rotated 45° from horizontal, and the amount of circular polarisation respectively.

Further author information: E-mail: cnd3@hw.ac.uk, Telephone: +44 (0)131 458 4168

(a) Jai's AD-080 GE (upper) and Thales' Catherine MP (lower) mounted on a pan & tilt head.



(b) Optical diagram showing two different camera systems.

Figure 1: Demonstration of sensor setup on the trials van: the Visible and NIR sensors come through a single optical system to two CCD arrays, while the LWIR polarimeter is separate, using a polarisation sensitive focal array.

They Stokes parameters may be determined by a set of simple experiments, if $I(\theta, \epsilon)$ represents the intensity of light in the direction making an angle θ with the positive direction of the x-axis, when the y-component is subjected to a retardation ϵ with respect to the x-component, then the following relations hold:

$$\left.\begin{array}{l} I = I(0°, 0) + I(90°, 0) \\ Q = I(0°, 0) - I(90°, 0) \\ U = I(45°, 0) - I(135°, 0) \\ V = I\left(45°, \frac{\pi}{2}\right) - I\left(135°, \frac{\pi}{2}\right) \end{array}\right\} \tag{2}$$

Thales' high-resolution long wave infrared (LWIR) imager, Catherine MP (640×512 pixels, 8 µm to 12 µm, 100 frames/s),[11] can be modified to be sensitive to radiation at various angles of polarisation,[12] allowing access to the first three Stokes parameters in real time.

Both sensors are securely fixed to a pan & tilt head which is mounted on top of the trials van, this was positioned to look at the road ahead which the vehicle was in motion, and was not moved while data was captured. The sensor setup is demonstrated in Figure 1. To register the images from each sensor, the parameters of the perspective transformation function were optimised to maximise the mutual information between the different images. The use of mutual information is common in the related field of medical image registration,[13] where input images capture similar underlying structure, but the intensities of points across different modalities are not directly related.

Thermal emission can be modelled through the process of refraction, where electromagnetic waves transition from one medium to another.[14] The refracted wave is partially polarised in a direction perpendicular to the plane of incidence respectively, the components of which are governed by the Fresnel transmissivity equations:

$$\left.\begin{array}{l} \mathfrak{T}_{\parallel} = \frac{\sin 2\theta_i \sin 2\theta_t}{\sin^2(\theta_i+\theta_t)\cos^2(\theta_i-\theta_t)} \\ \mathfrak{T}_{\perp} = \frac{\sin 2\theta_i \sin 2\theta_t}{\sin^2(\theta_i+\theta_t)} \end{array}\right\} \tag{3}$$

where θ_i is the angle of incidence and θ_t is the angle between the surface normal and the line of sight of the sensor, related by Snell's law ($n_i \sin\theta_i = n_t \sin\theta_t$). Man-made objects, particularly those with smooth surfaces, are found to emit strongly polarised radiation compared to natural objects,[15,16] making measurements of the degree of polarisation in a scene useful for vehicle detection. See for example, Figure 2, which shows the Q Stokes image of two vehicles alongside the visible image.

Figure 2: Emitted radiation is partially polarised in a plane perpendicular to the surface, and polarisation is strongest for surfaces observed near grazing angle. An image of two vehicles is presented as observed in the Q Stokes image (upper) and the visible (lower). Notice that the roof is present with maximum intensity, and the sides are present with minimum intensity.

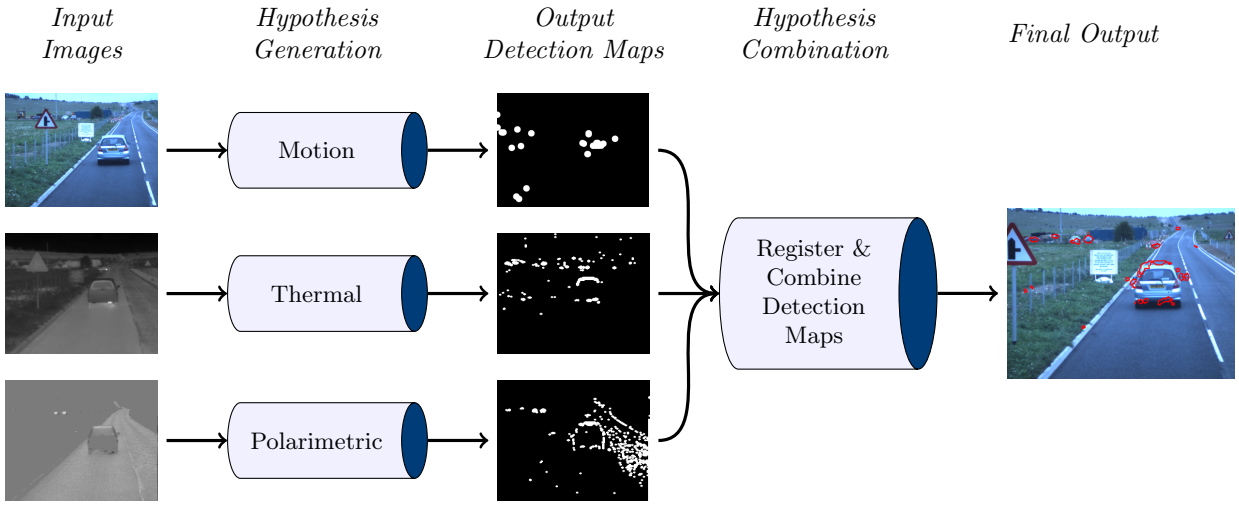


Figure 3: Flowchart demonstrating operation of algorithm. In this example, at least two detection maps must agree for a final detection to be marked.

It is hypothesised that through combining data collected from multiple sensing modalities, the performance of simple detection algorithms will be improved. Utilising data collected from multiple sensors requires fusion to combine the measurements. There are multiple levels at which data can be combined:[17] pixel level fusion, where a datacube is formed with the raw (or preprocessed) imagery; feature-level fusion, where features are extracted from the inputs and combined before being classified; and decision-level fusion, where each input stream produces a decision (binary or otherwise) which are subsequently combined. Fusion at lower levels of abstraction requires accurate registration methods to perform well and is more computationally intensive; fusion at higher levels is more readily adaptable to different sensor sets, or different operating conditions, but discards information earlier in the system.

This paper explores how a vehicle detection system performs when using different combinations of input imagery, and different types of detectors. The vehicle detection system is described in more detail in Section 2. The experiments carried out are described in Section 3 and discussed in Section 4. Finally, Section 5 concludes the paper with some closing remarks and paths for further investigation.

2. VEHICLE DETECTION SYSTEM

The vehicle detection system consists of a set of hypothesis generators (HGs) which process the input imagery to form a set of output detection maps which are subsequently combined in a fusion engine to provide a final

set of detections for that frame. This process is summarised in Figure 3. Each of n HGs operate on a single image I_j from the set of all m available images, $\mathbf{I} = [I_0, I_1, \ldots, I_m]$. In the current implementation j can be one of visible, near infrared (NIR), LWIR Intensity (TI), or LWIR Q Stokes (TQ):

$$y_{i,j} = f_i\left(I_j\right), i \in \{0 \ldots n\} \tag{4}$$

with each HG, $f_i(.)$, producing one component in the sparse set of detection maps $\mathbf{y} = [y_{0,0}, \ldots, y_{n,m}]^{\mathrm{T}}$. Each map is a 2 dimensional image of the same size as the image processed to obtain it, I, where each pixel in the map, $p(u, v) \in \{0 \ldots 1\}$ represents the probability of finding a vehicle at that pixel. Once generated, these maps are then registered to a common frame (that of the visible image) using a perspective transformation:

$$y'_{i,j}\left(\mathbf{x}'\right) = y_{i,j}\left(t_x\left(\mathbf{x}\right), t_y\left(\mathbf{x}\right)\right) \tag{5}$$

which maps coordinates $\mathbf{x}' \leftarrow \mathbf{x}$, $\mathbf{x} = (x_0, x_1)^{\mathrm{T}}$, as:

$$x'_0 = t_x\left(x_0, x_1\right) = \frac{a_{00}x_0 + a_{01}x_1 + a_{02}}{a_{20}x_0 + a_{21}x_1 + a_{22}} \tag{6}$$

$$x'_1 = t_y\left(x_0, x_1\right) = \frac{a_{10}x_0 + a_{11}x_1 + a_{12}}{a_{20}x_0 + a_{21}x_1 + a_{22}} \tag{7}$$

The perspective transformation matrix, $\mathbf{a}$, is calculated offline using a mutual information similarity criterion in an iterative registration process. These transformed maps, $\mathbf{y}' = \left[y'_{0,0}, \ldots, y'_{n,m}\right]^{\mathrm{T}}$, are then combined in the hypothesis combination (HC) stage using a certain fusion engine, $g(.)$:

$$z = g\left(\mathbf{y}'\right) \tag{8}$$

where z is the final set of detections. Details about specific HGs, $f\left(.\right)$, and the method used in the HC stage, $g\left(.\right)$, are described in the following sections.

2.1 Hypothesis generation

This stage consists of a selection of hypothesis generators (HGs) which each transform a particular input into a detection map. This has been done in such a way to allow any number of HGs to process and provide detections from a wide range of input data, but provide output in a form which can be easily combined. This section details three different HGs, two fairly low level methods consisting of simple thresholds which are used to process the thermal intensity and the thermal Q Stokes inputs, and one higher level HG which aims to detect moving objects from the visible input.

2.1.1 Motion

The motion HG makes the assumption that the camera is moving through the environment and, by tracking feature points, attempts to detect regions where the tracks deviate from tracks associated with the background, these outlying points are then further processed to identify any secondary consistent motion which could represent a moving object.

In the first image of a sequence, the HG is initialised by detecting FAST (Features from Accelerated Segment Test[18]) feature points, which are then tracked using the Pyramidal implementation of the Lucas Kanade Optical Flow algorithm,[19,20] where correspondences are sought within a 10×10 pixel window over four pyramid levels. Every n frames the HG is reinitialised by detecting new FAST feature points, where currently $n = 10$.

The points which are successfully tracked are then used in a RANSAC (RANdom SAmple Consensus[21]) framework to obtain the Fundamental matrix.[22] The RANSAC method is particularly attractive as it provides an estimation of both the best model given noisy data, and the set of inliers and outliers to this model based on whether they are within a threshold distance, p_{M}, between the epipolar line associated with the original point and the matching point. Outliers are then filtered with a second RANSAC Fundamental Matrix estimation to identify any consistent motion within the outlying points. Tracks which are outliers to global motion, but have consistent motion, are marked as detections.

The ultimate detection map produced by the Motion HG is formed around the set of k tracks marked as detections, $\mathbf{X}_{\mathrm{o}} = [\mathbf{x}_{\mathrm{o}1}, \mathbf{x}_{\mathrm{o}2}, \ldots, \mathbf{x}_{\mathrm{o}k}]^{\mathrm{T}}$, using a top-hat function:

$$f_{\mathrm{M}}(\mathbf{x}) = \begin{cases} 1 & \min_i ||\mathbf{x} - \mathbf{x}_{\mathrm{o}i}||_2 \leq q \\ 0 & \text{otherwise} \end{cases} \tag{9}$$

2.1.2 Thermal Polarimetric

Due to their nature, vehicles can often be identified as local heat sources, and also as regions of high linear polarisation. This is exploited by both the thermal HG and the polarimetric HG which identify vehicles by means of a threshold. Both local (adaptive) and global thresholds were tested.

The global threshold method is controlled by a single parameter, p_{Global}:

$$f_{\mathrm{TGlobal}}(\mathbf{x}) = \begin{cases} 1 & I(\mathbf{x}) \geq p_{\mathrm{Global}} \\ 0 & \text{otherwise} \end{cases} \tag{10}$$

The local threshold method computes the threshold for each pixel, $t(\mathbf{x})$, depending on the intensity within a local neighbourhood:

$$f_{\mathrm{TLocal}}(\mathbf{x}) = \begin{cases} 1 & I(\mathbf{x}) \geq t(\mathbf{x}) \\ 0 & \text{otherwise} \end{cases} \tag{11}$$

where the local threshold is equal to the mean intensity of pixels within a small window of the current pixel, plus a constant c. For the current study a window size of 3×3 pixels was used, and the constant was used to control the sensitivity of the HG.

2.2 Hypothesis combination

At this stage the vector of detection maps generated at the HG stage, $\mathbf{y}'(\mathbf{x})$, is combined into a single detection map, $z(\mathbf{x})$:

$$z(\mathbf{x}) = g(\mathbf{y}'(\mathbf{x})) \tag{12}$$

This process is known as decision-level fusion (as this process is combining decisions which have already been made at the HG stage). The fusion engine, $g(.)$, should be able to recognise agreement between different hypotheses while eliminating false detections. This is achieved using the relatively simple voting method where detection maps are combined additively and a threshold is then applied which allows the user to select how many HGs must agree:

$$g(\mathbf{y}'(\mathbf{x})) = \begin{cases} 1 & \sum_i^n \alpha_i y_i'(\mathbf{x}) \geq t \\ 0 & \text{otherwise} \end{cases} \tag{13}$$

Here, α_i is a weighting coefficient applied to the output of detector i, and t is the ultimate threshold for determining a detection. The value for α_i can be chosen from knowledge about how the individual detectors perform, although in this work $\alpha_i = 1$ for all HGs.

3. EXPERIMENTS

3.1 Procedure

The experimental goal is to determine the detection performance of the system, specifically how each HG performs independently, and how effectively the HC stage eliminates false positive detections. The system was tested using an image sequence obtained with sensors as described above while travelling along a rural road. The sequence extends for a total of 700 image frames, during which a number of vehicles are present at a variety of distances. Ground truth data was created for the sequence using vatic (the Video Annotation Tool from Irvine, California[23]) in offline mode. Example frames from the dataset with ground truth annotations are shown in Figure 4. Note the combination of a close range vehicle on the road, some parked medium distance vehicles to one side, and some distant vehicles on the road ahead.

(a) 100 (b) 300 (c) 500 (d) 700

Figure 4: Example frames from test data with ground truth annotations, the numbers are represent the frame number. Note the presence of vehicles at a range of distances.

Experiments were performed to evaluate each HG independently, and then test how different combinations of HGs altered the individual detection rates. From the sequence of 700 frames, every 10th frame was used to test the system, making a total of 71 test frames. Performance was assessed using receiver operating characteristic (ROC) curves which plot the true positive rate (tpR) against the false positive rate (fpR) over a number of sensitivity levels. An ideal curve will pass through the point $(0, 1)$, and so a metric was also calculated as the smallest distance between this point and the ROC curve. The rates were defined as:

$$tpR = \frac{v_d}{v_t} \tag{14}$$

$$fpR = \frac{fp}{N} \tag{15}$$

where v_d are the number of vehicles which were detected, defined as vehicles with at least 10% of their area marked as detected, v_t are the total number of unoccluded vehicles in the scene, fp are the number of pixels incorrectly identified as belonging to a vehicle respectively, and N are the total ground truth number of non-vehicle pixels.

The ROC curves were generated by collecting these statistics from detection maps produced for all 71 of the test frames.

3.2 Individual HG Results

Each hypothesis generator was tested in isolation, and the results are presented in this section. These results are discussed in Section 4, and in Section 4.4 the we explore the outcomes when different HGs are combined.

3.2.1 Motion HG

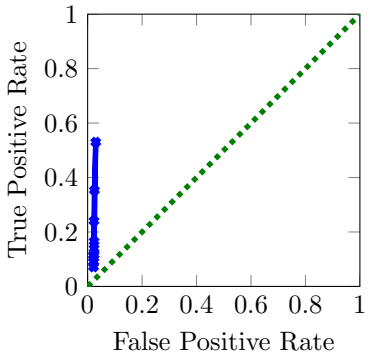
The motion hypothesis generator, introduced in Section 2.1.1, makes predictions about the presence of moving vehicles based on tracked points from the visible (RGB) input. In order to allow a reasonable distance to have been travelled between input images, every 5th frame was selected from the original set of 700 frames. The motion HG ROC curve is shown in Figure 5 alongside example tracked points and resultant detections. Note that in Figure 5b tracks are denoted by white lines terminated by green circles for inliers to global motion, red circles for outliers to global motion, but possess locally consistent motion, and blue circles for outliers to both global and local motion.

3.2.2 Thermal HG

The thermal hypothesis generator input takes the input from the LWIR I image and tests both the local and global threshold HGs introduced in Section 2.1.2. Results are shown alongside example output in Figure 6.

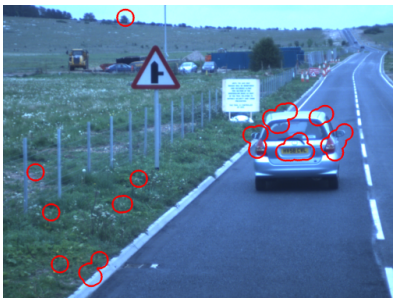
3.2.3 Polarimetric HG

The polarimetric hypothesis generator takes input from the LWIR Q image and tests both the local and global threshold HGs introduced in Section 2.1.2. Results are shown alongside example output in Figure 7.



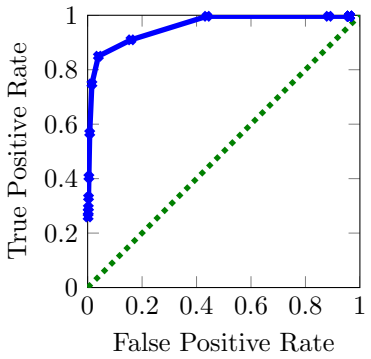
(a)

(b)



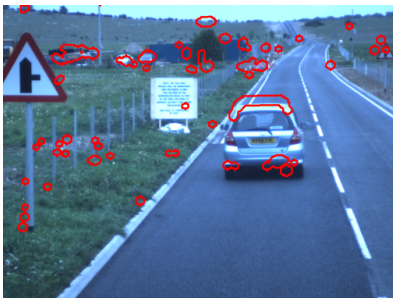
(c)

Figure 5: The ROC curve generated using the Motion HG is shown in (a), alongside an example showing the tracks from which the output was generated. Tracked feature points are indicated by white lines terminated by large red circles if they are outliers to global motion and have consistent motion. Figure c then shows the output produced from this HG.

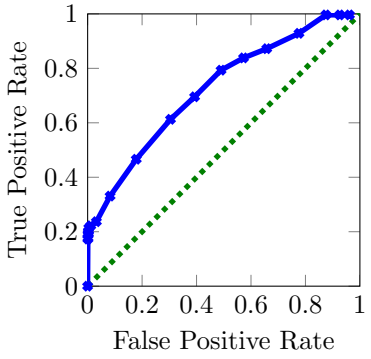


(a)

(b)

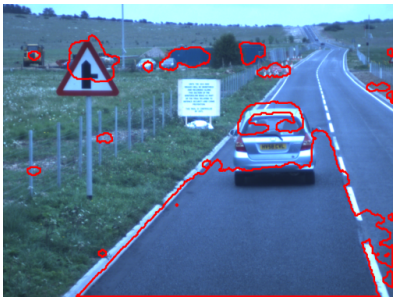


(c)



(d)

(e)



(f)

Figure 6: ROC curves for Thermal Intensity HG when using (a) local and (d) global thresholds with corresponding example input (b), (e) and output (c), (f) respectively.

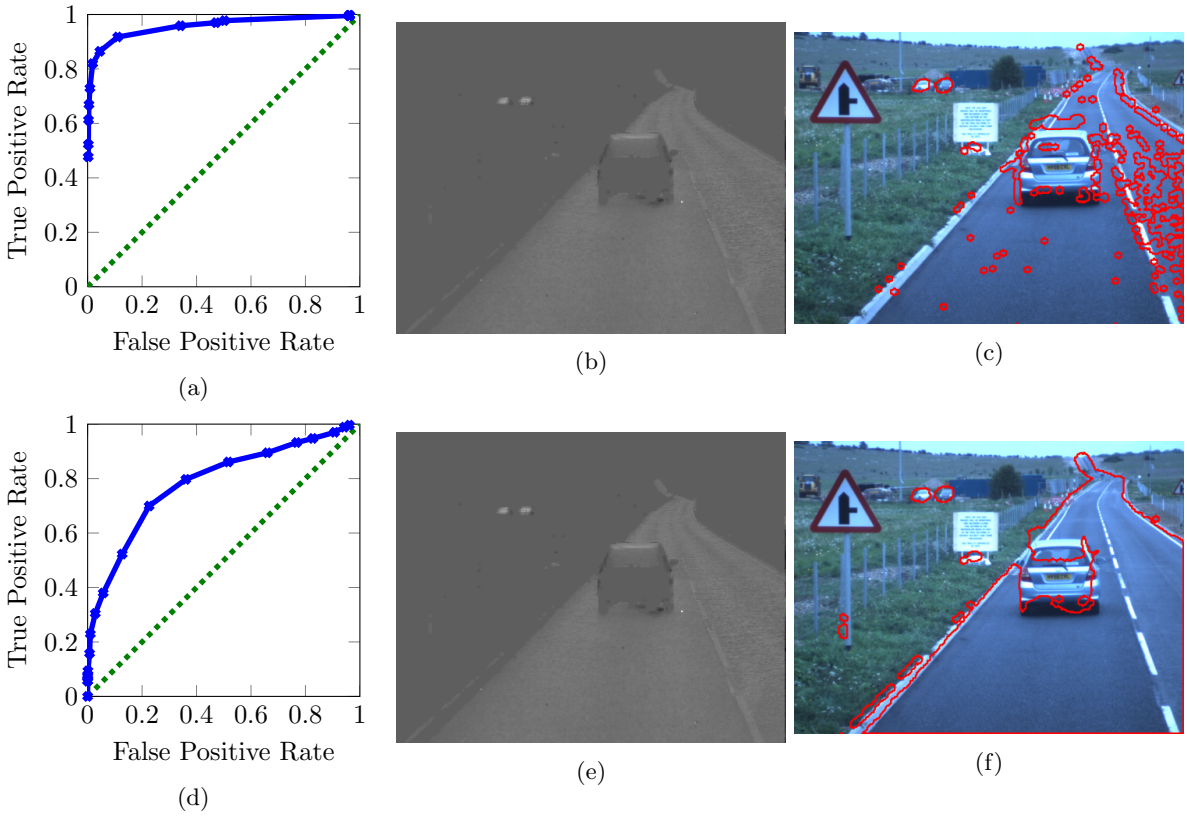


Figure 7: ROC curves for Polarimetric HG when using (a) local and (d) global thresholds with corresponding example input (b), (e) and output (c), (f) respectively.

4. DISCUSSION

In this section we consider the performance of each HG when working independently, then in Section 4.4 an overall discussion is presented comparing each HG and studying how different combinations of HGs alters the detection results.

4.1 Motion HG

The Motion HG attempts to identify moving vehicles by tracking feature points identifying any points which deviate from the global motion in a consistent manner. Results from this HG were shown in Figure 5.

These results show that this HG successfully identifies the large moving vehicle in the foreground, as a number of the tracks on this vehicle are seen to deviate from the global motion. Additionally a number of the tracks that are inconsistent with global motion, but also not conforming to a secondary motion (lines terminating in a blue dot in Figure 5b) are identified and discarded — these will typically correspond to tracking errors. However, a number of false positives are still present in the detection map.

The ROC curve seems to show a varying true positive rate at approximately constant false positive rate for this HG, but the curve is not increasing monotonically. For each feature in the first image, the best fit projection matrix allows an epipolar line to be drawn in the second image, on which the matching point in the second frame is expected to lie. The distance between the matched point and this line is compared to a threshold to determine whether the match is declared an inlier or outlier for this projection. This threshold was chosen as the controlling parameter for this HG. However, the ROC curve suggests that this parameter has little effect on the number of false positives recorded, but does alter the number of true positives recorded.

Although this HG is fairly simple and will only identify one non-global motion in the scene, even if it were extended to a multi-body segmentation it would still only ever detect moving objects. This means that other

PLocal	TLocal	Motion	t	tpR	fpR	metric
y	-	-	1	0.9173	0.1103	0.1379
-	y	-	1	0.8496	0.0396	0.1555
-	-	y	1	0.5301	0.0298	0.4709
y	y	-	2	0.8045	0.0038	0.1955
y	y	y	3	0.3421	0.0003	0.6579
y	y	y	2	0.8722	0.0078	0.1281

Table 1: This table shows the true and false positive rates, alongside the metric, for system runs consisting of different combination of detectors, and combination thresholds. PLocal is the Polarisation HG with local threshold, TLocal is the Thermal HG is local threshold, Motion is the Motion HG, t defines how many HGs must agree for a final detection to be marked, tpR, fpR, and metric are the true positive rate, false positive rate, and metric as defined in the text. The ideal metric is 0.0. The values shown are for runs with optimal parameters.

HGs will have to be responsible for detecting any stationary vehicles, and this HG can only provide additional confidence to any moving vehicle detections.

4.2 Thermal HG

The results obtained with the Thermal HG were shown in Figure 6. It is evident that the local threshold is much more suited to this modality, and produces a much improved ROC curve. This can be explained by observing for example the road near to the camera. This area is present at relatively large intensity which triggers the global threshold into reporting a detection, while as the area is consistent, the adaptive threshold is better suited to picking out the smaller objects of interest from the scene.

One issue with this detector, in both global and local modes, is that while the medium- and long-range vehicles are identified successfully, only sections (the exhaust, tyres, and roof) of the short-range vehicle are identified. Improvements to this HG may be realised by performing some post-processing to group vehicle part detections in order to expand the detected region to encompass the whole vehicle.

4.3 Polarimetric HG

The results obtained with the Polarimetric HG were shown in Figure 7. The Stokes Q image used for this HG can be seen to be much more noisy than the thermal intensity image, and this is evident from the adaptive threshold which shows many local false detections in the road region. However, compared to the global example, where the entire road region is highlighted as a detection, the local version is found to perform more favourably. This can be ascribed to the strong signature of the road in this modality, which is marked along its entirety as a detection in the global threshold method, while only in the more speckled areas in the local method.

One method of improving this HG could be to apply a median filter to the input, this would remove much of the speckled appearance, however this would also remove the ability to detect any distant vehicles present over a small number of pixels. As such it would be preferable to filter the detection map by another means.

4.4 Overall comparison

A comparison of the optimum detection rates for each of the three HGs, along with different combinations is provided in Table 1. The metric shown is the distance between the optimum (fpR, tpR) point listed for that system, and the ideal (0, 1) point (an ideal metric would be 0.0).

Comparing the HGs operating independently, it can be seen that the Polarimetric detector provides the largest true positive rate of all the detectors, with a correspondingly small metric. However of the individual HGs it also exhibits the largest false positive rate. This is confirmed qualitatively by observing the large number of false positive detections in the road region of Figure 7c.

The thermal HG used alone performs similarly to the polarimetric HG, with both the true and false positive rates slightly lowered, and similar overall metric.

However, the motion HG when used on its own is quite interesting in that while it has a comparably poor true positive rate, detecting around half of the vehicles present, it has the smallest false positive rate of the

(a) (b) (c)

Figure 8: Detection maps from (a) thermal and (b) polarimetric HGs alongside and combined system output (c).

individual HGs. The small true positive rate is to be expected as this HG is not designed to detect any vehicles which are not in motion, but the small false positive rate is encouraging.

The lower half of the table shows different combinations of HGs. Line four shows the results obtained when combining the polarimetric and thermal HGs run with optimum parameters, and only generating an output detection if both HGs agree. This has the result of lowering the true positive rate below both of the individual HGs, but also reduces the false positive rate by an order of magnitude. This demonstrates that this combination does effectively eliminate much of the noise present in both of these HGs, as can be seen in Figure 8 which shows the output from each HG individually, and the fused result after combination.

Two methods of combining all three HGs were trialled where a detection was marked if any two or all three HGs agree. Requiring a unanimous decision was found to have the effect of reducing the true positive rate somewhat, but significantly reducing the false positive rate by two orders of magnitude compared to any individual HG, or one order of magnitude below the combined Polarimetric-Thermal system. Relaxing the constraint to only require agreement between any two of the three HGs improves the true positive rate to a level comparable to the best performing individual HG, while also significantly reducing the false positive rate.

5. CONCLUSIONS & FUTURE WORK

This paper introduces a new vehicle detection system which combines a set of simple detectors operating on multiple input modalities. It has been shown that combining different types of data is an effective method to significantly reducing false positive detections, without significantly hampering the number of true positive detections. This work also highlights how discriminative polarimetric imaging can be, especially in the thermal infrared waveband where emission dominates and observed radiation can reveal significant information about object shape and location.

For future work, it is thought that combining the data at an earlier stage may improve performance further: combining the polarimetric and intensity data into a single image in a lower-level fusion system and then processing could be beneficial, as combining processed data inherently involves loss of information. Additionally the current method does not attempt to group information from the different HGs: for example, the polarimetric HG detects the vehicle roof while the thermal HG detects the tyres. Without grouping these measurements will be considered independent and will not contribute to an increased confidence of detection, even though they may both be identifying the same vehicle.

Additionally, the system only currently uses temporal data in the Motion HG, while if detections were filtered throughout the sequence, instantaneous detections from the current frame could be fed into a more probabilistic detection framework which makes detection predictions based on information from multiple frames.

ACKNOWLEDGMENTS

The author would like to acknowledge the support of both academic and industrial supervisors, and the financial support of the EPSRC.

REFERENCES

[1] Connor, B., Letham, J., Robertson, N., and Iain Carrie, "Scene understanding and task optimisation using multimodal imaging sensors and context: a real-time implementation," in [*Proc. SPIE 8012*], 80120A (2011).

[2] Buch, N., Orwell, J., and Velastin, S., "Detection and classification of vehicles for urban traffic scenes," in [*Proc. 5th Int. Conf. Visual Information Engineering*], 182–187 (2008).

[3] Breckon, T. P., Barnes, S. E., Eichner, M. L., and Wahren, K., "Autonomous Real-time Vehicle Detection from a Medium-Level UAV," in [*Proc. 24th Int. Conf. Unmanned Air Vehicle Systems*], 29.1–29.9 (2009).

[4] Schaul, L., Fredembach, C., and Süsstrunk, S., "Color image dehazing using the near-infrared," in [*Proc. 16th IEEE Int. Conf. Image Processing*], 1629–1632, IEEE (2009).

[5] Tucker, C. J., "Red and photographic infrared linear combinations for monitoring vegetation," *Remote Sensing of Environment* **8**(2), 127–150 (1979).

[6] Cremer, F., Jong, W. D., and Schutte, K., "Processing of polarimetric infrared images for landmine detection," in [*Proc. 2nd Int. Wksp. Adv. Ground Penetrating Radar*], (2003).

[7] Bartlett, B. D., Schlamm, A., Salvaggio, C., and Messinger, D. W., "Anomaly detection of man-made objects using spectropolarimetric imagery," in [*Proc. SPIE*], **8048**, 80480B (2011).

[8] Lavigne, D. A., Breton, M., Fournier, G., Charette, J.-F., Pichette, M., Rivet, V., and Bernier, A.-P., "Target discrimination of man-made objects using passive polarimetric signatures acquired in the visible and infrared spectral bands," in [*Proc. SPIE*], **8160**(1), 816007 (2011).

[9] Sadjadi, F. and Sadjadi, F., "Passive Polarimetric Information Processing for Target Classification," in [*Augmented Vision Perception in Infrared: Algorithms and Applied Systems*], Hammoud, R. I., ed., ch. 2, 37–61, Springer (2008).

[10] Tompkins, H. G., Irene, E. A., Hill, C., and Carolina, N., [*Handbook of Ellipsometry*], William Andrew Publishing & Springer-Verlag GmbH & Co. KG (2005).

[11] Crawford, S., Craig, R., Haining, A., Parsons, J., Costart, E., Bois, P., Gauthier, F.-H., and Cocle, O., "Thales long-wave advanced IR QWIP cameras," in [*Proc. SPIE 6206*], 62060H (2006).

[12] Parsons, J. and Craig, R., "A LWIR polarimetric imager," in [*Proc. 5th EMRS DTC Conf.*], (2008).

[13] Hill, D. L., Batchelor, P. G., Holden, M., and Hawkes, D. J., "Medical Image Registration," *Phys. Med. Biol.* **46**(3), R1–R45 (2001).

[14] Wolff, L., Lundberg, A., and Tang, R., "Image understanding from thermal emission polarization," in [*Proc. IEEE Conf. Computer Vision and Pattern Recognition*], 625–631 (1998).

[15] Connor, B., Carrie, I., Craig, R., and Parsons, J., "Discriminative imaging using a LWIR polarimeter," in [*Proc. SPIE*], **7113**, 71130K (2008).

[16] Carrie, I. D., Connor, B., and Christie, A., "Difficult Target Discrimination Using Polarimetric Imaging," in [*Proc. 7th EMRS DTC Conf.*], B11 (2010).

[17] Steinberg, A. and Bowman, C., "Revisions to the JDL Data Fusion Model," in [*Handbook of Multisensor Data Fusion*], Hall, D. L. and Llinas, J., eds., ch. 2, CRC Press (2001).

[18] Rosten, E. and Drummond, T., "Machine learning for high-speed corner detection," in [*Proc. Euro. Conf. Computer Vision*], 430–443 (2006).

[19] Bouguet, J.-Y., "Pyramidal implementation of the lucas kanade feature tracker description of the algorithm," (2000).

[20] Lucas, B. D. and Kanade, T., "An iterative image registration technique with an application to stereo vision," in [*Proc. Int. Joint Conf. Artificial Intelligence*], 674–679 (1981).

[21] Fischler, M. and Bolles, R., "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM* **24**(6), 381–395 (1981).

[22] Hartley, R. and Zisserman, A., [*Multiple View Geometry in Computer Vision, 2nd ed*], Cambridge University Press (2004).

[23] Vondrick, C., Patterson, D., and Ramanan, D., "Efficiently Scaling Up Crowdsourced Video Annotation," *Int. J. Comput. Vision* (2012).