5^{th} Int. Conf. Image Proc. and Applications, Heriot-Watt University, Edinburgh, 3-5 July 95.

VEHICLE DETECTION IN INFRARED LINESCAN IMAGERY USING BELIEF NETWORKS

P.G.Ducksbury, D.M.Booth, C.J.Radford

Defence Research Agency, Malvern, UK.

This paper describes a system for detecting vehicles in airborne downward looking infrared linescan imagery, and in particular, the use of a Pearl-Bayes Network (PBN) to combine disparate sources of evidence. Here the primary source of evidence is a vehicle detection algorithm with supporting evidence being provided by vehicle track and shadow detectors. The spatial arrangement of the vehicles also provides useful contextual evidence since vehicles often move in convoy or are clustered into small groups when encamped. This observation is the basis for allowing neighbouring detections to re-enforce one another and for incorporating a feedback loop with which to increase the sensitivity of the vehicle detection algorithm within areas of suspected activity.

INTRODUCTION

A system is described for detecting vehicles in airborne downward looking infrared linescan imagery as a precursor to analysis by a photographic interpreter. The imagery being studied consists of only natural terrain, such as heathland, since urban areas are in themselves sufficient to attract the attention of an interpreter, and algorithms already exist for identifying such regions (2). The vehicles have a typical resolution of around 10-15 pixels across.

This paper focuses on the use of a PBN to combine sources of evidence which reinforce the classification of a vehicle and in doing so examines the techniques which generate the prior knowledge required by the PBN. A detailed survey of related work is contained within (3), in particular, we exploit an IED project concerned with the identification of urban and driveable regions in various types of imagery (2).

The primary source of evidence is a vehicle detection algorithm which extracts potential targets by filtering, and then assigns each of them a probability based on the 'vehicleness' of their shape and grey level distribution etc. Supporting evidence is provided by vehicle track and shadow detectors. The spatial arrangement of the vehicles also provides useful evidence, since vehicles often move together or are clustered into small groups. This latter feature is important because it suggests that detections could re-enforce one another, and potentially, that a number of the marginally missed detections could be corrected. Spatial relationships are also exploited by allowing a combined belief in the support of a vehicle to feed back to the detectors and influence their sensitivity within the local neighbourhood.

VEHICLE - FEATURE DETECTION

The vehicle detector (1) consists of 5 stages : preprocessing; vehicle detection; segmentation; feature extraction and classification.

Preprocessing incorporates scanline noise removal and a rectilinearisation procedure for correcting across heading sec-squared geometrical distortion that is intrinsic to the scanning mechanism.

Vehicle detection is achieved using two filters : a matched filter which detects localised bright regions that contrast greatly with their surroundings; and a difference of two medians filter (also known as a co-median filter) which emphasises small objects by subtracting large scale background structure. Vehicle candidates are segmented by thresholding the co-median filter output on the basis of local variance statistics. In the default mode of operation the matched filter acts as an enabler to the co-median filter. Thus the desirable properties of both filters are utilised, most notably, the matched filter's robustness to false alarms, and the retention of shape information by the co-median filter.

Four types of feature are extracted from the local neighbourhood of a potential vehicle : shape; object grey level distribution; texture and object to background contrast. Classification as vehicle or as background clutter is achieved by a Fisher linear discriminant (4).

EVIDENCE ACCUMULATION

The sources of supporting evidence are : shadows; tracks; and the spatial relationships that exist between potential vehicles.

Vehicle Shadow Detection

Vehicle shadows exhibit a characteristic wedge shape and grey level distribution which can be recognised by a detection-classification strategy similar to that used for the detection of vehicles.

Most notably, the choice of feature set was heavily constrained by the limited availability of training samples and a requirement to keep inputs to the PBN as independent as possible.

Vehicle Track Detection

Vehicle tracks are detected by the application of a Marr-Hildreth filter, followed by non-maximal suppression, hysteresis thresholding (*to remove small relatively unconnected line segments*), line filtering (*to remove small jagged spurs connected to main line segments*), and finally, parallel sets of lines are identified by a matching process based on the Hough Transform. To ensure that tracks are detected for a range of different vehicles and ground conditions, the line detection process is performed at several spatial resolutions with the results being combined using a logical OR.

Vehicle Group Detection

The orientation of each vehicle detected by the classifier is computed and a binormal distribution fitted to all neighbouring vehicles of similar orientation. The result is a set of interleaving binormal distributions. More specifically, the process is achieved by examining a small patch of the image around a vehicle. A gradient image is computed for this patch and the angle is obtained from arctan of the ratio of the x and y components at each edge point. These angles are then accumulated into a histogram covering all 360 degrees. We utilise reflection to consider just 4 main compass points (*0, 45, 90 and 135 degrees*). The resulting angles are a good representation of a vehicles orientation.
Once similarly orientated vehicles have been identified, clustering is applied using the Median and Median Absolute Deviation (*normalised to an ideal model, Huber (5)*) to fit one or more binormal distributions to each orientation cluster . The distributions being rotated to align with the cluster itself. The angle being obtained by a simple linear regression ($y = ax + b$) on the coordinate points of the vehicles in relation to the principal axis of the cluster.

Evidence Combination - Theoretical

Bayesian networks are directed acyclic graphs, such that a graph G is a pair of sets (V,A) for which V is non-empty. The elements of V are vertices (*nodes*) and the elements of A are pairs (x,y) called arcs (*links*) with $x \in V$ and $y \in V$.
In our network (Figure 3) node A has several predecessors, and consequently the belief generation and propagation equations are more complex than might be expected. The equations are derived along similar lines to those of Pearl in (6).
If we just consider the link from nodes B to A then graph G consists of two subgraphs G^+_{BA} and G^-_{BA} which contain the datasets D^+_{BA} and D^-_{BA} respectively.

Belief equations

From Figure 3 we can see that node A separates the two subgraphs $G^+_{BA} \cup G^+_{CA} \cup G^+_{EA}$ and G^-_{AF}. Given this fact we can write the equation :

$$P(D^-_{AF}|A_i, D^+_{BA}, D^+_{CA}, D^+_{EA}) = P(D^-_{AF}|A_i) \quad (1)$$

by using Bayes rule the belief in A_i is given by

$$\begin{aligned} BEL(A_i) &= P(A_i|D^+_{BA}, D^+_{CA}, D^+_{EA}, D^-_{AF}) \\ &= \alpha P(A_i|D^+_{BA}, D^+_{CA}, D^+_{EA}).P(D^-_{AF}|A_i) \\ &= \alpha P(D^-_{AF}|A_i). \sum_{j,k,l} [\; P(A_i|B_j, C_k, E_l). \\ &\quad P(B_j|D^+_{BA}).P(C_k|D^+_{CA}), P(E_l|D^+_{EA})] \quad (2) \end{aligned}$$

where α is taken to be a normalizing constant.
It can be seen that equation 2 is computed using three types of information

- Causal support π (*incoming links*).
- Diagnostic support λ (*outgoing links*).
- Fixed conditional probability matrix (*relates A with its immediate causes B,C and E*).

The equations which form the above information are given as follows. Firstly the causal support equations :

$$\pi_A(B_j) = P(B_j|D^+_{BA}) \quad (3)$$

$$\pi_A(C_k) = P(C_k|D^+_{CA}) \quad (4)$$

$$\pi_A(E_l) = P(E_l|D^+_{EA}) \quad (5)$$

Secondly the diagnostic support equation is given by

$$\lambda_F(A_i) = P(D^-_{AF}|A_i) \quad (6)$$

Finally the conditional probability matrix is defined to be $P(A|B, C, E)$.
Equation 2 can now be rewritten in order to obtain the belief at node A.

$$\begin{aligned} BEL(A_i) = \alpha\lambda_F(A_i). \sum_{j,k,l} P(A_i|B_j, C_k, E_l). \\ \pi_A(B_j).\pi_A(C_k).\pi_A(E_l) \quad (7) \end{aligned}$$

The belief at nodes B,C and E can be obtained from the equations

$$BEL(B_j) = \alpha.\pi_A(B_j).\lambda_A(B_j) \quad (8)$$

$$BEL(C_k) = \alpha.\pi_A(C_k).\lambda_A(C_k) \quad (9)$$

$$BEL(E_l) = \alpha.\pi_A(E_l).\lambda_A(E_l) \quad (10)$$

Propagation equations

These are derived firstly for the diagnostic case. From an analogy with equation 6 we can write

$$\lambda_A(B_i) = P(D^-_{BA}|B_i) \quad (11)$$

by partitioning the D^-_{BA} into its component parts, namely A, D^-_{AF}, D^+_{CA}, D^+_{EA} we can obtain

$$\lambda_A(B_i) = \alpha \sum_{j,k} [\ \pi_A(C_j).\pi_A(E_k). \sum_l \lambda_F(A_l).P(A_l|B_i, C_j, E_k)\ ] \quad (12)$$

likewise for $\lambda_A(C_j)$ and $\lambda_A(E_k)$

$$\lambda_A(C_j) = \alpha \sum_{i,k} [\ \pi_A(B_i).\pi_A(E_k). \sum_l \lambda_F(A_l).P(A_l|B_i, C_j, E_k)\ ] \quad (13)$$

and

$$\lambda_A(E_k) = \alpha \sum_{i,j} [\ \pi_A(B_i).\pi_A(C_j). \sum_l \lambda_F(A_l).P(A_l|B_i, C_j, E_k)\ ] \quad (14)$$

Similarly for the causal case, an analogy with equation 3 allows us to write

$$\pi_F(A_i) = P(A_i|D^+_{BA}, D^+_{CA}, D^+_{EA}) \quad (15)$$

and then to derive the equation

$$\pi_F(A_i) = \alpha \sum_{j,k,l} [\ P(A_i|B_j, C_k, E_l). \pi_A(B_j).\pi_A(C_k).\pi_A(E_l)] \quad (16)$$

Note that equations 12 - 14 and equation 16 demonstrate that the parameters λ and π are orthogonal to each other ie. perturbation of one will not affect the other. Hence evidence propagates through a network and there is therefore no reflection at boundaries.

Evidence Combination - Applied

As was described earlier, a Pearl-Bayes Network is a Bayesian approach to reasoning (6). In this application the fusion of evidence is achieved using the relatively simple tree structured network, shown in Figure 1, the actual PBN being contained in the large dashed box. The input level in the lower half of the diagram contains the supporting evidence, i.e. that which relates to vehicle shadows, tracks and groupings, this being combined at the next level to form a single measure of belief in all the supporting evidence. This is then combined at the next level with evidence supplied by the vehicle detector itself to give an overall belief in the existence of a vehicle given all the available knowledge.
Concerning the fusion process : concurring evidence should not generate a high belief if it is not supportive of a vehicle; a vehicle lacking supporting evidence should not imply that a vehicle does not exist, and similarly, strong supporting evidence on its own must not be allowed to infer the existence of a vehicle. However, the latter has been exploited in a feedback loop to the vehicle detector, as although the detector has proved reasonably robust, the initial filtering process occasionally fails to highlight potential vehicles. The feedback loop indicates regions most likely to contain vehicles. In view of the possible pay-off, the filtering procedure then boosts its sensitivity at the expense of increased false alarms.
One of the difficulties that has prevented people from using the Pearl-Bayes Network approach is the construction of the prior (*knowledge*) probabilities that relate the input evidence to the output belief. One solution is to treat the evidence essentially as a clustering problem and this has proved successful in the past (2), currently it is based on using the statistical median and median absolute deviation as basic measures when assigning these probabilities.
The probabilities for combining all of the available evidence form a 4-dimensional set comprising three inputs and one output. The median and Median Absolute Deviation (MAD) normalised to an ideal model are used as measures. Firstly the medians and MAD's are computed from the labels and table 2 is used to obtain the probabilities which are then scaled to (0,1). The resulting set of probabilities are then converted into a multivariate set of 5 labels according to the following rule.

$$P(_q|_{r,s,t}) = \begin{cases} (0.1, 0.1, 0.1, 0.1, 0.6) & if\ P > .8 \\ (0.1, 0.1, 0.1, 0.6, 0.1) & if\ P > .6 \\ (0.1, 0.1, 0.6, 0.1, 0.1) & if\ P > .4 \\ (0.1, 0.6, 0.1, 0.1, 0.1) & if\ P > .2 \\ (0.6, 0.1, 0.1, 0.1, 0.1) & otherwise \end{cases}$$

such that $\sum_q P(_q|_{r,s,t}) = 1 \forall$ r,s,t and q = 0,..,4.

Combining evidence with target prior

The label set of 5 labels is denoted as low (L), lower medium (LM), medium (M), upper medium (UM) and high (H). We generally try to ensure that the output belief corresponds to the input unless the evidence strongly suggests otherwise. The probabilities are a 3-dimensional set that can be interpreted using Table 1. Evidence is shown on the vertical axis and target probability on the horizontal. Since the probabilities are 3-dimensional each entry in Table 1 has an associated vector ranging from $L = (0.6, 0.1, 0.1, 0.1, 0.1)$ through to $H = (0.1, 0.1, 0.1, 0.1, 0.6)$. The following condition must apply on the probabilities $\sum_q P(_q|_{r,s}) = 1\ \forall$ r,s and q = 0,..,4.

RESULTS

Over the terrain being considered, vehicle detection and classification have been reasonably successful. As expected, evidential reasoning does not have a major effect when the vehicle classification results are reliable (and certain) but can come into its own when the classifier decisions are more marginal. In such cases, the vehicle classifications have been reinforced by the supporting evidence.

Figure 2 shows typical input imagery; note the wedge shaped shadows and the presence of tracks. The vehicle under highlight is comparatively weak and as a consequence was missed by the first pass of the vehicle detection filter. In this case, evidence of neighbouring vehicles was fed back to the detector, its local sensitivity increased, and the vehicle located. On a small set of 9 images containing a total of 79 vehicles, 67 vehicles (85%) and 4 false alarms were detected initially. After implementation of the PBN and the feedback loop, vehicle detection had increased to 75 vehicles (95%) with 12 false alarms.

Most of the missed detections resulted, at least in part, from vehicles being either unusually small or partially obscured by vegetation. One miss was caused by a vehicle being adjacent to a transition in background texture and an inability of the detection algorithm to adapt accordingly.

The increased false alarm rate is high in comparison with the previous level but not regarded as a serious problem as it is likely that vehicles have already been detected in the same areas and, therefore, the additional overhead to a photographic interpreter is not significant. In any case, the feedback loop could be configured to particular operational requirements.

DISCUSSION

An evidential reasoning approach such as a Pearl-Bayes Network offers a flexible framework into which alternative or more reliable sources of evidence could be incorporated should the need arise. Even so, it appears that more comprehensive systems will be required (*i.e. making the transition from image processing to image understanding*) if it were seen desirable to extend this work into a more general image interpretation system. This and previous work has shown that we can successfully detect vehicles, urban regions, road networks etc, what is now required is bringing all these components together into a more general based image understanding system.

REFERENCES

1. Booth D.M. and Radford, C.J., Oct 1994, *'The detection of vehicles in downward looking infrared linescan imagery'*, 12^{th} ICPR, Jerusalem, Israel.

2. Ducksbury P.G., Sept 1993, *'Parallel Model-Based Segmentation using a Pearl Bayes Network'*, British Machine Vision Conference, Sept 1993, University of Surrey.

3. Ducksbury P.G., Booth D.M., Radford C.J., March 1995, *'Improved vehicle detection in infrared linescan imagery using belief networks'*, DRA internal Milestone report.

4. Duda R.O., Hart P.E., 1973, *'Pattern Classification and scene analysis'*, Wiley and Sons.

5. Huber P.J., 1981. *'Robust Statistics'*, Wiley and Sons.

6. Pearl P., 1987, *'Fusion, Propagation, and structuring in Belief Networks'*, Artificial Intelligence 29 (1986) pp 241 - 288.

		Target Probability				
		L	LM	M	UM	H
Evidence	L	L	LM	M	UM	H
	LM	LM	LM	M	UM	H
	M	M	M	M	UM	H
	UM	H	H	H	H	H
	H	H	H	H	H	H

Table 1: *Target/Evidence Contextual knowledge.*

Prob	Condition	Explanation
0.8	$m \geq x/2 \wedge M \leq x/4$	Supporting, close
0.6	$m \geq x/2 \wedge M > x/4$	Supporting but not close
0.4	$m < x/2 \wedge M \leq x/4$	Not Supporting, close
0.1	$m < x/2 \wedge M > x/4$	Not Supporting and not close

where m = median of (s,t,u),
M = Median Absolute Deviation of (s,t,u),
x = range of labels,
s = Shadow prior label,
t = Track prior label,
u = Group prior label

Table 2: *Rules for contextual knowledge.*

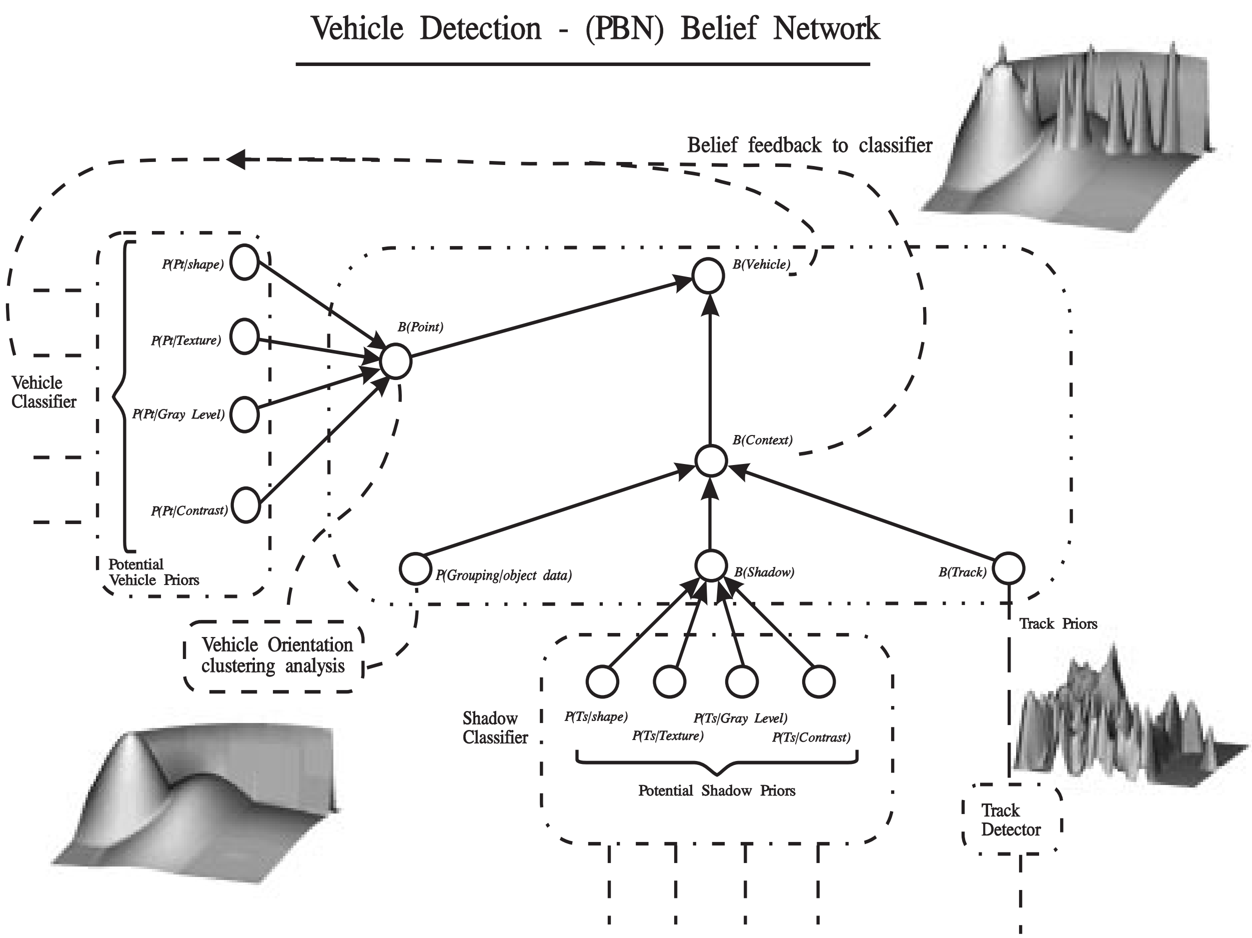

Figure 1: *Vehicle Detection Belief Network.*

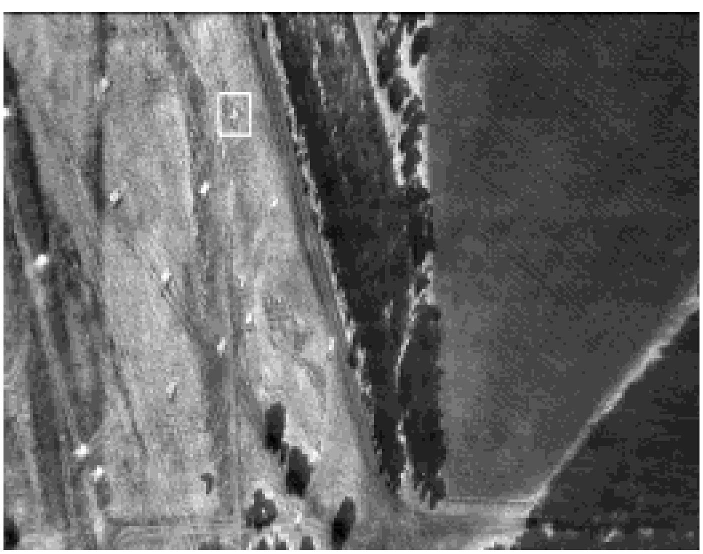

Figure 2: *Example of vehicle reinforcement.*

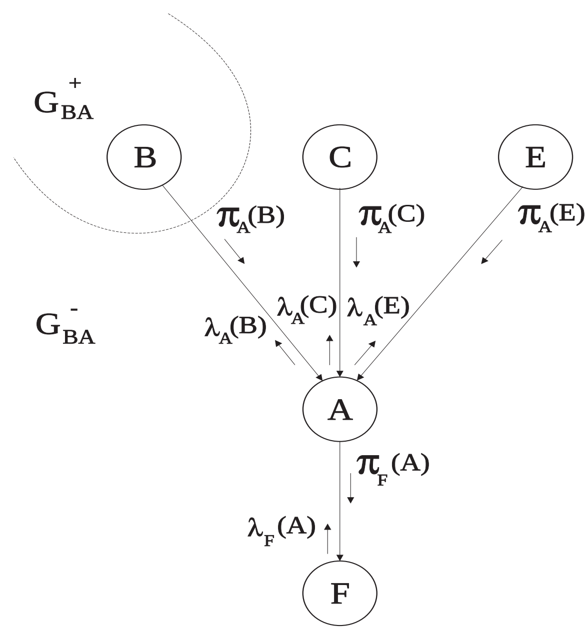

Figure 3: *PBN - belief computation.*