DRIVEABLE REGION SEGMENTATION USING A PEARL BAYES NETWORK

P. G. Ducksbury, Defence Research Agency, Malvern, Worcs., UK.

ABSTRACT. This paper describes a parallel implementation of a texture segmentation algorithm. The algorithm uses a Pearl Bayes Network (PBN) to combine evidence for the location of driveable regions in autonomous land vehicle imagery. A multilevel PBN approach is introduced and followed by an example which is used to illustrate the derivation of the propagation and fusion equations. A parallel implementation is then described with results demonstrating its effectiveness[1].

INTRODUCTION. The vast majority of papers in the literature which deal with the fusion of knowledge for applications in evidential reasoning have concentrated on the theoretical aspects. The papers have dealt with the structure of the networks themselves and how they can be used to represent and manipulate knowledge. The Bayesian approach for reasoning is described in a series of papers by Pearl in particular (4) where he describes the basics of Bayesian networks and belief functions. There are a number papers which are now beginning to address the problems of using evidential reasoning in the area of image understanding.

THE PROBLEM. The problem is generically defined as the location of some region in an image. This problem will be approached by taking several statistical measures from small patches of an image these are treated as a set of judgements (*virtual evidence in Pearl's notation*) about the content of the patches. This evidence is then combined into a belief of the patch belonging to the defined region. The belief is improved by using information from a higher level. An example of this being given in figure 1 which shows how a multilevel approach can be applied. The lowest level in the pyramid contains the raw pixels at full image resolution whilst the top level is the overall belief in a region.
In the next section we concentrate on a simple case of just a single level network as shown in figure 2. It is shown how the equations for the belief and propagation of information in a PBN can be derived.

NETWORK AND EQUATION CONSTRUCTION. Consider the network in figure 2 and the definition given below.
*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 $y \in V$.*
The equations for computing the belief and propagation of information are derived in the following sections.
If we consider the link from node B to A then the graph G consists of the two subgraphs G^+_{BA} and G^-_{BA}. These two subgraphs contain the datasets D^+_{BA} and D^-_{BA} respectively.

BELIEF EQUATIONS. From figure 2 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 :

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

by using Bayes rule the belief in A_i can be written as

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

[1]The work was originally developed for the location of urban regions in airborne infra-red imagery as reported in Ducksbury (2).

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 (*from the incoming links*).
- Diagnostic support λ (*from the outgoing links*).
- A fixed conditional probability matrix (*which 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 :

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

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

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

Secondly the diagnostic support equation is given by

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

Finally the conditional probability matrix is defined to be

$$P(A|B,C,E) \tag{7}$$

Equation 2 can now be rewritten in order to obtain the belief at node A.

$$BEL(A_i) = \alpha\lambda_F(A_i).\sum_{j,k,l} P(A_i|B_j,C_k,E_l).\pi_A(B_j).\pi_A(C_k).\pi_A(E_l) \tag{8}$$

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

$$BEL(B_j) = \alpha.\pi_A(B_j).\lambda_A(B_j) \tag{9}$$

$$BEL(C_k) = \alpha.\pi_A(C_k).\lambda_A(C_k) \tag{10}$$

$$BEL(E_l) = \alpha.\pi_A(E_l).\lambda_A(E_l) \tag{11}$$

<u>PROPOGATION EQUATIONS</u> The propagation equations for the network are derived as follows, firstly the diagnostic ones. From an analogy with equation 6 we can write

$$\lambda_A(B_i) = P(D^-_{BA}|B_i) \tag{12}$$

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

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

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

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

and

$$\lambda_A(E_k) = \alpha \sum_{i,j} \left[\pi_A(B_i).\pi_A(C_j).\sum_l \lambda_F(A_l).P(A_l|B_i,C_j,E_k) \right] \tag{15}$$

The Causal equations can be derived in a similar way. An important point to realise is the fact that equations (13) - (15) and those for the causal equations 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.

<u>IMPLEMENTATION</u> The application is for the detection of driveable regions in a sequence of forward looking images from a moving vehicle. A mesh is placed over an image and for each of the windows a set of statistics (*which provide strong texture discrimination*) are computed. The statistics are the number of edges, the number of extrema and gray level distribution type. These statistics are quantized down into a smaller number of levels. The number of edges and extrema are both reduced to 5 levels, whilst the distribution type remains with its 4 possibilities.

The statistics are then used to produce a set of judgements, for example an expert might upon looking at a particular window issue a report of the form $(0.0, 0.7, 0.9, 0.6, 0.0)$. This means that he believes there is a 70% chance that level 2 describes the number of edges, 90% chance that its level 3 and 60% for level 4. But he believes there to be no chance of it being levels 1 or 5.

For the Belief at nodes Bf, Bc and B in figure 1 it was decided to have 3 variables which denote the possible values that the region can have, namely (low, medium, high).

The fixed conditional probability matrices (eg $P(Bf|s1, s2, s3)$ etc) which are the prior information and relate the given node with its causal information are created along similar lines to that used in (2) and which originally came from (1). They are based upon the assumption that the probability of an event at a given node should be greater if its causal information is tightly clustered together than it should be if the causal information is further apart. For the $P(B|Bf, Bc)$ matrix (*which relates the beliefs from the fine and coarse resolutions*) slightly more emphasis is given to the causal information received from the coarse resolution belief.

If the application remains of a broadly similar nature (ie classifying (*or clustering*) regions) then the only change necessary would perhaps be a new set of statistics which more accurately describe the detail required in the image. In addition to this if the number of input nodes alters then the prior knowledge in the fixed conditional probability matrix will need to change, however the set of basic equations given in references (1) and (2) can be used to automatically generate this type of information.

<u>PARALLEL PROCESSING.</u> The architecture is based upon a transputer array called CHIP (Conceptual Hierarchical Image Processor). It is a real-time image processing system which is intended to be used as a test-bed for developing new algorithms, and prototyping of new image processing architectures.

Video data is communicated throughout using two digital video busses. CHIP provides for video input to an acquisition/display board which is connected to a video crossbar switch, this in turn is connected to the main processing unit of CHIP. Various DSP devices can be flexibly interconnected by the video crossbar for the pre-processing of images. The result can then be passed to all processors in the transputer array. Each module in the array being a T805 transputer with 4MByte of DRAM and 2MByte of VRAM. The latter forms two banks of 4 (512×512) framestores. Two video input and output busses are common to each module, each bank of framestores being connected to one of the input and one of the output busses. The individual transputer links of each module are connected to a link crossbar switch allowing for different network topologies. For details of the high level system design for CHIP refer to (5).

There are a number of ways in which the algorithm could actually be parallelised. However geometric parallelism was chosen as likely to be the most suitable and is described as follows. Each processor handles a small section of the data space but has a complete copy of the algorithm with one processor allocated as a master controller. This approach was chosen as all stages of the algorithm are totally deterministic and since once the information regarding the PBN tree structure and probabilities has been communicated to each processor and the tree built no further communication between processors is required. This eliminates the need for any (possibly) expensive communications and opens the way for expectation of significant speedups as the number of processors is increased. The physical processors are arranged in a pipeline.

RESULTS AND CONCLUSIONS. Figure 3 shows the actual probability surface corresponding to the driveable region. This belief output labelling is post processed and presented as an outline overlayed on the original image in figure 4.

The algorithm runs with a video as the input signal and currently has a total processor time per frame for a 500×500 pixel image with $20 \times$ T805 transputers of approximately 0.95 seconds[2].

The approach has been applied two two different driveable sequences giving good results on both. Once initialisation has been performed the algorithm is ideally suitable for asynchronous parallel implementation for which a linear speedup obtainable. Realistically the degree of parallelism is only limited by the number of processors that are available, each of the windows could theoretically have been computed in parallel. This approach compares extremely favourably with a previous one described in (3) which used a Hidden Markov Mesh Random Field for the texture region segmentation of the statistical data, and was not as suitable for a parallel implementation.

REFERENCES

1. Devijver P.A., *'Real-time modeling of image sequences : Based on Hidden Markov Mesh Random Field Models'* in *'Decision making in context'* a Course on Statistical Pattern Recognition by P.A.Devijver and J.Kittler, Surrey University, 1989.

2. Ducksbury P.G., *'Parallel Texture Region Segmentation using a Pearl Bayes Network'*, British Machine Vision Conference BMVC-93, University of Surrey, Sept 21-23, vol 1, pp 187-196, 1993.

3. Ducksbury P.G., *'Parallel Model-Based Segmentation using a* $3^{\rm rd}$ *order Hidden Markov Model'*, IEE 4^{th} Int. Conf. Image Processing and its Applications, Maastricht, The Netherlands, 7-9 April, 1992.

4. Pearl J., *'Fusion, Propagation, and structuring in Belief Networks'*, AI 29, 1986, pp 241-288.

5. Smith R.W., *'Conceptual Hierarchical Image Processor (CHIP) : System Design'*, Issue 1.0, Oct 92, CSE1 Division, DRA Malvern.

[2]It is estimated that with some of the initialisation performed in hardware and an array of T9000 transputers the algorithm should be able to perform at the rate of at least 12-15 frames per second.

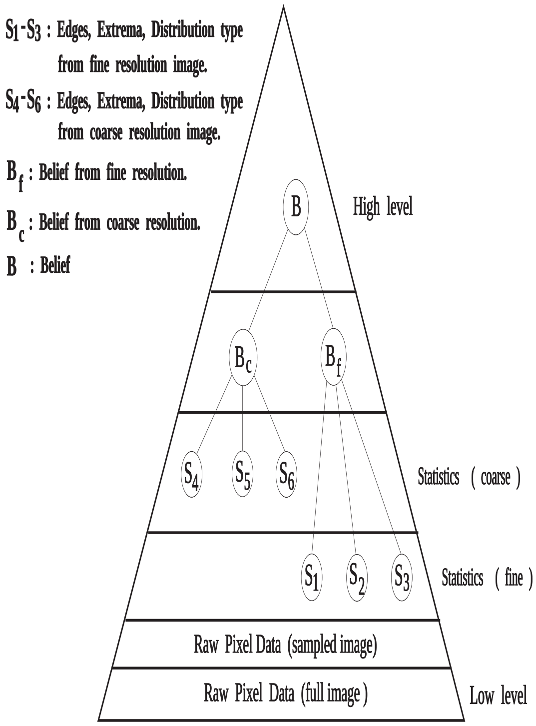

Figure 1: Multi-level Pyramid.

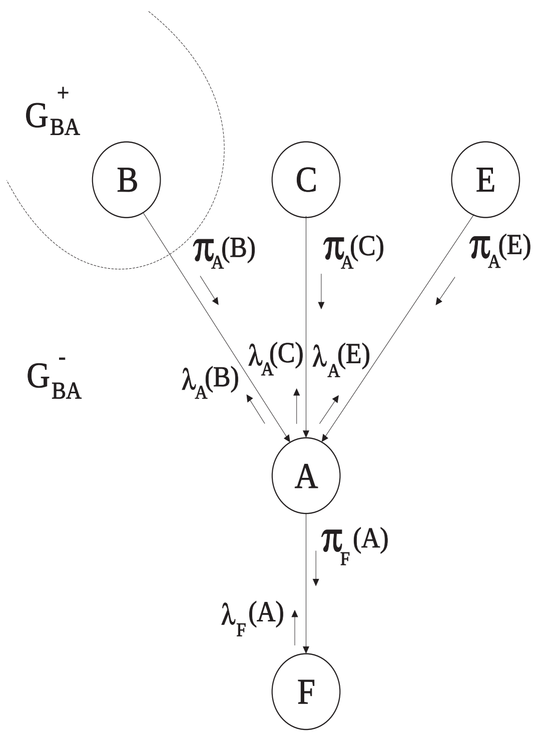

Figure 2: Pearl Bayes Network : Nodes B, C and E represent statistical information, Whilst node A represents overall belief.

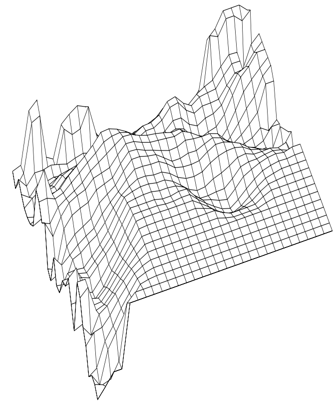

Figure 3: Probability Surface.

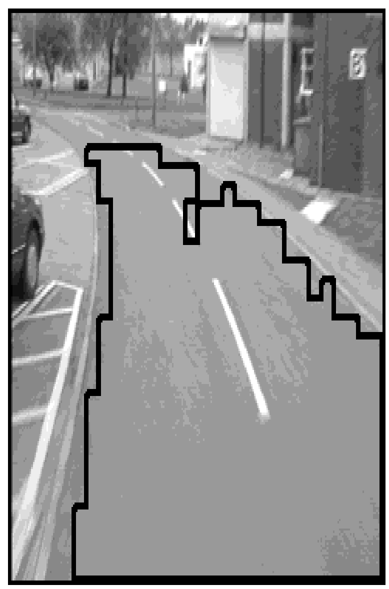

Figure 4: Driveable Region.

Acknowledgments : This work was partially funded by the UK, DTI IED Programme : IED-1936 *'Vision by Associative Reasoning'*.