REGION BASED IMAGE CONTENT DESCRIPTORS AND REPRESENTATION[1]

P.G.Ducksbury, M.J.Varga

Defence Research Agency, UK

This paper addresses some of the difficulties involved in the automatic generation of image content descriptors within a specific domain of airborne reconnaissance. The initial stages of the work concentrated on the development/collation of a set of feature detectors covering points (*corners*), lines (*edges, ridges*) and regions which operate on a variety of image sensors (eg SPOT, IRLS - *Infra-red Linescan*, dd5 - *Russian satellite data).* We also consider the questions of combining several feature detectors to improve their performance and reliability on a wider variety of images. The descriptors once generated and represented in some suitable way could be used for the indexing, matching and searching as well as retrieval of data in an image database. These operations on the image database depend on a fundamentally reliable detection stage. For this reason the initial objective of our work was not to depend purely upon just one algorithm but to have several algorithms, the output of which may then be combined in a novel manner into a composite estimate of the feature(s) that are required.

FEATURE DETECTION

Feature detectors used in this work currently cover the detection of points (Smith (1)), edges (Canny (2), Smith (1)), ridges, (Petrou (3)) and regions such as Fractal and Wavelet, (Xie (4)) and Pearl Bayes (*belief*) Networks, (Ducksbury (5)). In this paper we will concentrate on the detection of regions and their subsequent representation in an image database.

Region Detection - Belief Network

The belief network is a directed acyclic graph which was originally proposed by Pearl (6) in the application area of medical diagnosis. Previous work Ducksbury (5) considered the use of belief networks in image processing for the detection of urban regions in IRLS imagery together with the detection of driveable regions in autonomous land vehicle imagery. This paper contains the full derivation of the equations as well as examples of a parallel implementation. Other work using the same technique but for improved target detection in airborne imagery is described in Ducksbury et al (7). The technique has now successfully been applied to images from the different image sensors mentioned above.

A Belief Network is an acyclic graph an example being shown in figure 1. Here π represents the causal support (*from the incoming links*), λ the diagnostic support (*from the outgoing links*), α is a scaling constant and *P()* is a conditional probability matrix which is essentially the prior knowledge (*which relates A with its immediate causes B,C and E*). The belief at node A can be dervied as

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

The idea behind using the belief network approach for region detection is one of combining multiple sources of information in a network to study the changes in beliefs for a set of events. In this application the image is decomposed into a set of small windows and a set of suitable statistical measures which give some indication of the type of texture within the window are computed. The belief network is used to combine these measures (*or evidence*) over multiple scales into a belief that represents the content of the region. Here, two scales proved sufficient together with three states (*low, medium and high*) for the evidence variable that represented the required region, the actual network used being a simple causal tree, figure 2.

The prior knowledge *P()* is a set of multi-dimensional conditional probabilities which relate the inputs (*causal information*) to the outputs of a given node within the graph. A set of rules are used which are based on the assumption that casual information that is tightly clustered should naturally generate a higher belief than that which is more dispersed. This is a reasonable assumption since if we consider segmentation of gray level imagery, then neighbouring pixels from different regions which have intensity values lying at either end of the gray level spectrum have low probability and for those that belong to homogenous regions, high probability.

We have also used other techniques such as fractal and wavelet approaches, which are now briefly described.

[1] Presented at 6th IEE IPA-97, Trinity College, Dublin, July 1997.

Region Detection - Fractal

The work on fractal analysis Xie (4) concentrated on texture discrimination and assumes that texture can be modelled as fractals (*a reasonable assumption for the type of imagery and application domain of interest*) and discusses the estimation of suitable fractal parameters. A fractal has several properties, namely irregular structures and scale invariance (*or self similarity of the structures over scale space*). Examples of fractal sets could be snow flakes, ferns, coast lines etc. For a self-similar fractal a measure is used called the fractal dimension. This does however only measure the topological properties of a fractal and research has shown that two textures can have very different appearances even if they have the same fractal dimension.

Further discrimination can be achieved by using a measure called the D-dimensional measure which is equivalent to length, area or volume for 1, 2 or 3 dimensions respectively, Falconer (8). This measure being independent of dimension. The work developed by Oxford consists of a multiscale analysis algorithm for texture discrimination.

Region Detection - Wavelet

The wavelet analysis was also developed by Xie (5). The work provides a detailed discussion on wavelet theory and application, also describing how this can be extended to two dimensions. A scheme was developed that employs the two dimensional local energy of the wavelet transform for texture segmentation. The approach basically consists of four stages as follows

Convolution of the image with a set of dilated wavelets to generate a set of wavelet detail images at given orientations and scale.

Generation of a set of local energy images (which removes phase dependency from the detail images).

Scale space fusion which includes inter-scale clustering (to minimise the spatial localisation problem) and inter orientation fusion.

Texture feature clustering and boundary detection.

FEATURE COMBINATION

One of the difficulties that has existed in computer vision is that researchers have had a tendency to concentrate on developing single algorithms (*eg edge and region detectors*) for specific types of images. It is only recently that papers are appearing which consider the questions of combining detectors to improve their performance and reliability on a wider variety of images.

Consideration was given to a variety of algorithms for combining feature detectors such as regions and regions, ridges and edges as well as regions and edges. Here we use the Belief Network approach for combination of all three region algorithms (*belief network, fractal, wavelet*) into a single composite region. This is effectively using the same approach but at a higher level of abstraction (*ie processing results of algorithms rather than images*). This approach has also been evaluated for combining the multiscale output of the wavelet algorithm into a single feature image.

Figure 3 shows a feature image resulting from the combining of the three region finding algorithms, this would then be thresholded prior to outline extraction. The example shown in figure 4 is the result of the most likely urban area located using the belief network (*the intention here being for speed of location for a 'focus of attention' operation rather than a highly accurate outline*). This outline can in fact be improved considerably by utilising additional evidence such as that obtained from a corner detector, Smith(1). This is understandable due to the fact that corners will be more tightly clustered around the urban areas.

The advantage of our use of the Belief Network approach is that it is independent of the application (*eg. either combining low level pixel level data or higher level algorithmic data*).

FEATURE REPRESENTATION

Once a region is obtained then there is the issue of how to represent its shape. We have chosen to use the curvature scale space approach by Mokhtarian (9) which provides a very concise representation for objects. We initially explored this approach for the representation of objects such as aircraft and ships in an image database, Ducksbury (10).

The technique requires some explanation and for this we refer to figure 6. For a given object (*in this illustration an aircraft line drawing*[2]) the raw outline is obtained as a simple (x,y) co-ordinate file (*lower left*). By iteratively convolving the (x,y) data with a gaussian function, the outline will be successively smoothed, this being shown along the left hand side of figure 6. This gaussian convolution is repeated with larger values of scale parameter σ, at each iteration the zero crossings of curvature are located and marked on

[2] Aircraft outline data obtained by processing information supplied by Janes Information Group, Brighton road, Coulsdon, Surrey.

the u-σ plane. The curvature scale space (*CSS*) image itself, as shown plots contour position u against scale space σ. If a line is drawn horizontally across this image then this represents some particular scale, the points on this line effectively marking the zero crossings (*points of inflexion*) on the object contour. Hence the first line near the bottom has numerous points (*corresponding to the many small features on the object*) whilst the next line has just 6, which mark the three concavities (*two major ones either side of the rotor blades, one small one below the tail*). The iterative process terminates when no more zero crossings are located (*essentially no more concavities are indicated*) and the object has reduced to a convex shape, illustrated at the top left of the figure.

From the curvature scale space image the major peaks are located. The (x,y) location of each peak marks the concavities contour position and scale, it is these peaks that are used as the shape description for subsequent searching. Hence for the line drawing in figure 6 the six largest peaks are chosen, typically therefore just a handful of pairs of integers are used to encode the objects shape.

Figure 5 shows a curvature scale space image on the left for the urban region outline that is shown in figure 4, this also shows on the right one of the smoothed outlines (*in this case taken at scale* $\sigma = 2.0$) together with points of zero-crossing.

APPLICATIONS

The main applications of this work will be in the area of image databases. The amount of imagery that is currently being collected is going to require some form of automatic (*or even just semi-automatic*) processing prior to insertion into an image database. It is unreasonable to expect an operator to have to view each image and to select the image content even if done via a graphical user interface.

From our research the curvature scale space approach could in principle be used as a means to either search for an image containing a particular shape of region or to prevent multiple insertions of the same image into a database.

Research is continuing into other types of feature detectors and ways of combining these to obtain more accurate results.

ACKNOWLEDGEMENTS

This work has resulted from collaboration with the University of Oxford and the University of Surrey. We wish to thank Zhi-Yan Xie for providing the fractal and wavelet analysis software and Farzin Mokhtarian for a version of the curvature scale space software.

REFERENCES

1. Smith S.M., Brady J.M., 1992, 'A new class of Corner Finder', British Machine Vision Conference, BMVC-92, 139-148

2. Canny J., 1986, 'A Computational Approach to Edge Detection', IEEE Trans PAMI, 8

3. Petrou M., 1990, 'Optimal Convolution Filters and an Algorithm for the Detection of Linear Features' Report VSSP-TR-21/90, Dept of Electronic and Electrical Engineering, University of Surrey.

4. Xie Zhi-Yan, 1994, 'Multi-scale Analysis and Texture Segmentation', PhD Thesis, Dept of Eng Science, University of Oxford.

5. Ducksbury P.G., 1993, 'Parallel Texture Region Segmentation using a Pearl Bayes Network', British Machine Vision Conf., BMVC-93, 187-196

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

7. Ducksbury P.G., Booth D.M., Radford C.J., 1995, 'Vehicle detection in infrared linescan imagery using Belief Networks', IEE IPA-95, 415-419

8. Falconer K., 1990, 'Fractal Geometry', John Wiley Sons,

9. Mokhtarian F., 1995, 'Silhouette-Based Isolated Object Recognition through Curvature Scale Space', IEEE Trans PAMI, 17, 5

10. Ducksbury P.G., 1997 'Image Content Descriptors : Feature Combination and Representation', (unpublished) DRA technical report, UK

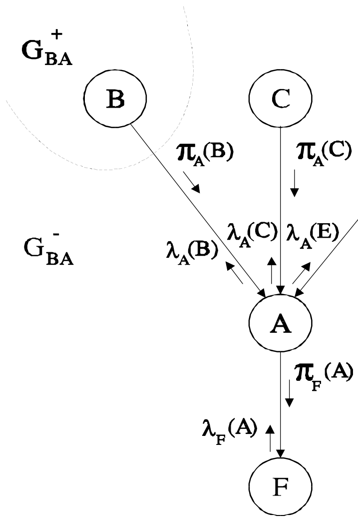

Figure 1: Belief Network

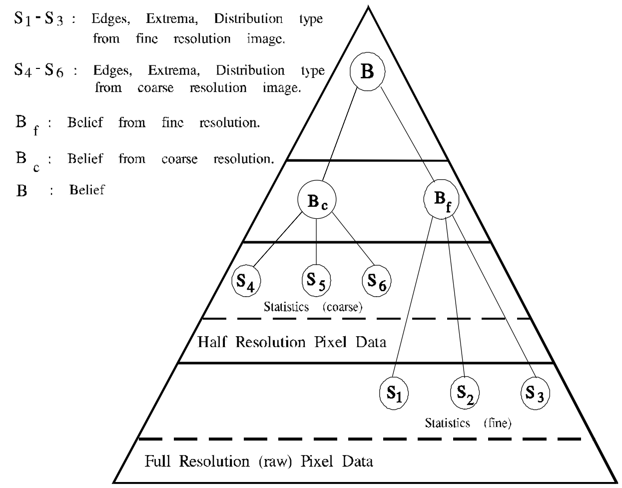

Figure 2: Causal Belief Network

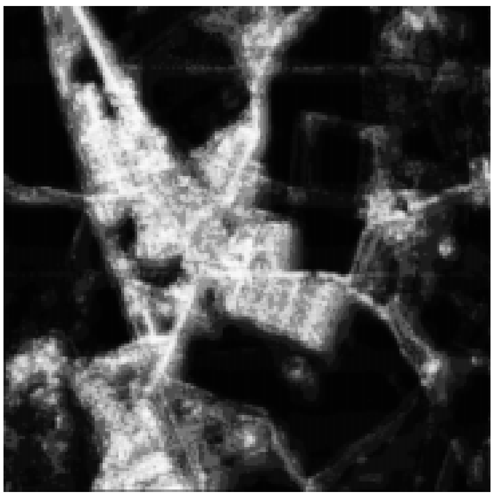

Figure 3 : Combined feature image

Figure 4 : Most likely urban area outline.

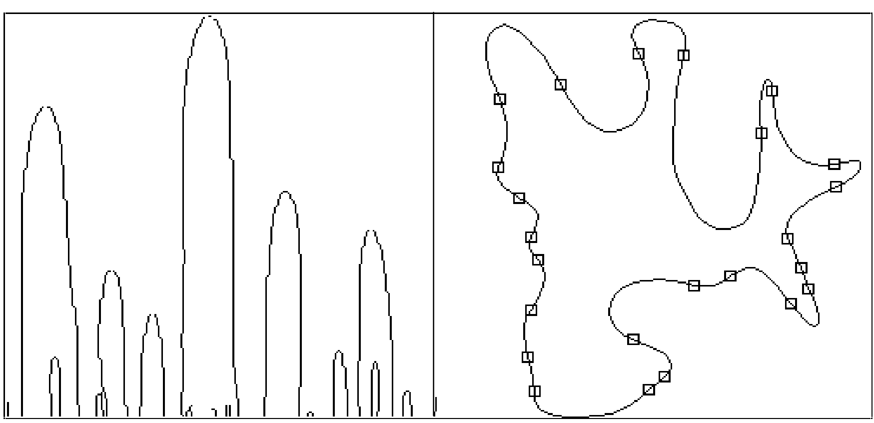

Figure 5 : Left - Curvature scale space image;
Right - Outline representation at scale 2.0

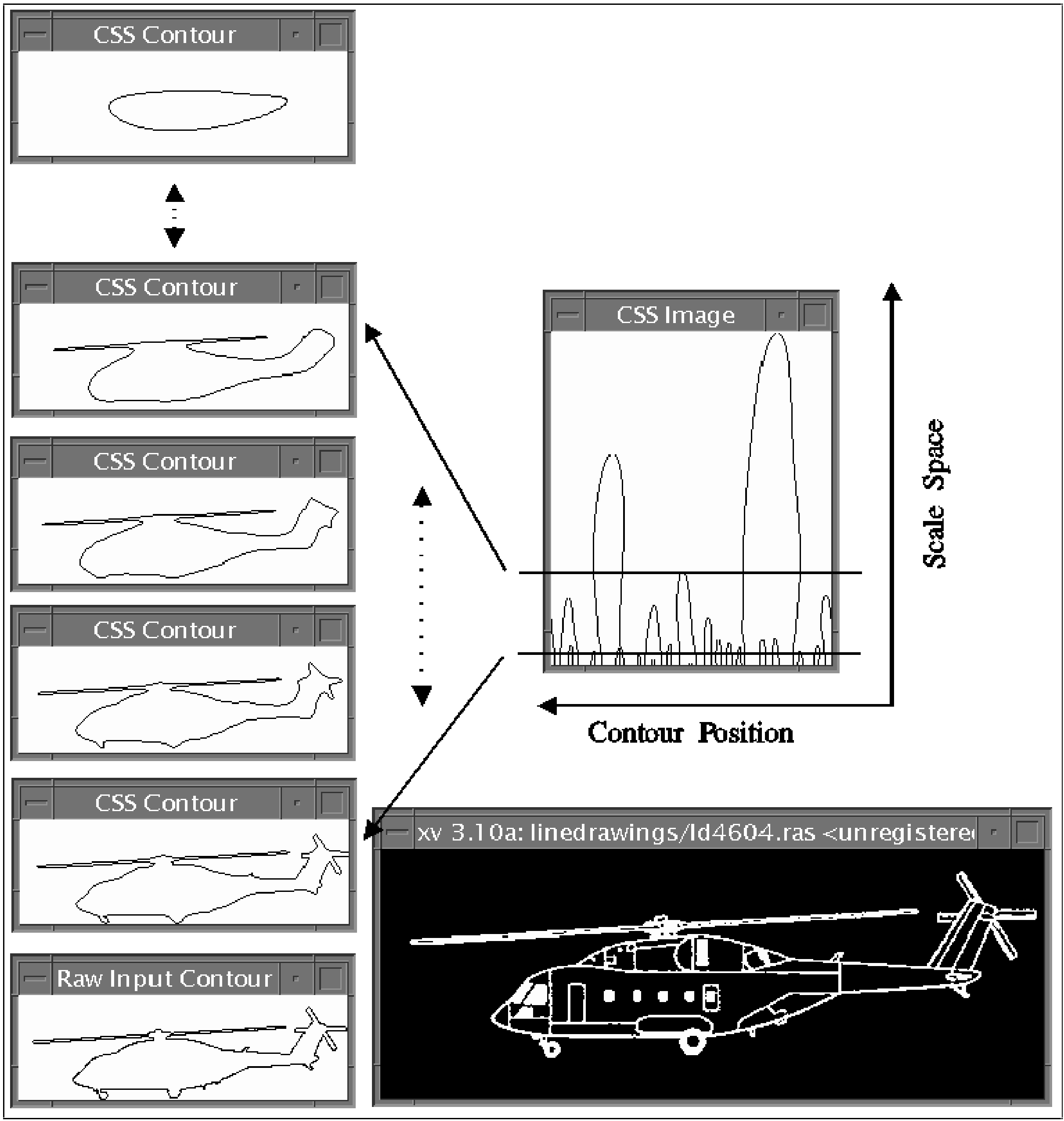

Figure 6: Curvature Scale Space Representation