Simultaneous Region and Edge Segmentation of Infrared Images using Non-Maximal Suppression for Edge Thinning

J.F. Haddon* & **J.F. Boyce**†
*Royal Aerospace Establishment, Farnborough, Hampshire, GU14 6TD
†Wheatstone Laboratory, King's College, Strand, London, WC2R 2LS

A cooccurrence space is defined by utilising the combinations of pixel strengths defined by a Canny edge operator. A region and boundary segmentation derived from this space is first edge thinned by non-maximal suppression and then hysteresis is used as a post-processing step to improve the edges. The distributions in cooccurrence space define the thresholds employed in the hysteresis post-processing.

It will be shown that edge detection, including non-maximal suppression and hysteresis post-processing, may be combined naturally with region segmentation. Examples will be given which demonstrate that this technique performs better than conventional segmentation or edge detection techniques.

The cooccurrence matrices [1] of an image exhibit second order statistical properties. Hence they may be utilised to analyse images which may be characterised by their second order statistics; the matrices being interpreted as feature spaces. In particular, the nearest-neighbour cooccurrence matrix of an image composed of regions of distinct intensity which are separated by boundaries one pixel wide may have elliptical classification regions defined within it, and the original image may then be segmented by mapping pixel pairs from image space to cooccurrence space[2]. The result of such a mapping is that every pixel is assigned to a class of region and is also identified as interior or edge, relative to the cooccurrence direction. Since the classification involves only pairs of pixels, the effect of noise is to create disconnections between edge elements and isolated misclassifications. Both errors are aggravated if the region boundaries are more than one pixel in width.

The use of the intensities of a pair of adjacent pixels as a boundary indicator is analogous to applying a first order finite difference operator as an edge detector, since in both cases the difference of the intensities is taken as a measure of the likelihood of the existence of an edge or boundary. Conversely, an optimised edge detector, such as a Canny operator[3], provides the optimal combination of strengths of extended sets of pixels in terms of which a region-edge segmentation may be achieved. An implementation of such a segmentation algorithm is described in section 2.

The use of extended pixel combinations effectively results in the removal of isolated pixel region misclassifications; however it has the drawback that it may cause the appearance of thick boundaries. These may be thinned using the same non-maximal suppression and hysteresis post-processing as employed in realisations of Canny edge detectors[4]. As shown in section 3, the extent and degree of overlap of the region and edge distributions in the cooccurrence space provide natural, and intrinsic, definitions of the upper and lower thresholds used in the edge thinning by hysteresis.

1 Cooccurrence Matrices

Consider an image of dimensions $L \times L$ whose intensity, $i(\mathbf{x})$ is sampled at $2N \times 2N$ equidistant points indexed by

$$\mathbf{x} = (l - 1/2)\Delta x \; \hat{\mathbf{i}} + (m - 1/2)\Delta y \; \hat{\mathbf{j}} \qquad (1)$$

with $-(N-1) \leq l, m \leq N$, and $\Delta x = \Delta y = L/(2N)$.

A Canny edge operator forms an edge strength, $e_{\mathbf{\Delta}}(\mathbf{x})$, from an intensity image by a $(2n+1)$ element convolution[5].

$$e_{\mathbf{\Delta}}(\mathbf{x}) = \sum_{k=-n}^{n} E(k) i(\mathbf{x} + k\mathbf{\Delta}) \qquad (2)$$

where Δ defines the direction normal to the edge. An edge image may then be formed by combining the edge strengths from orthogonal $\mathbf{i}$ and $\mathbf{j}$ directions.

Since the Canny operator has been optimised for step edge detection at any given resolution we infer that the left and right combinations of intensity defined by it are appropriate for the formation of a cooccurrence feature space for region and edge segmentation. The corresponding cooccurrence matrix is defined by

BMVC 1990 doi:10.5244/C.4.55

$$S_{\mathbf{\Delta}}(i,j) = \sum_{\mathbf{x}} \delta\left(i; \sum_{k=1}^{n} E(k) i(\mathbf{x} - k\mathbf{\Delta})\right) \delta\left(j; \sum_{k=1}^{n} E(k) i(\mathbf{x} + k\mathbf{\Delta})\right) \quad (3)$$

where $\delta(i;j)$ is a Kronecker delta function and $\sum_{k=1}^{n} E(k)i(\mathbf{x} + k\mathbf{\Delta})$ is the intensity arising from one lobe of the Canny operator.

The general structure of a cooccurrence matrix consists of a set of distributions close to the leading diagonal of the matrix, these being characteristic of regions, and an accompanying set of off-diagonal distributions, which are characteristic of edges. The region distributions are centred on the mean intensities of the regions; those of the edges are centred on the intensity pairs corresponding to the mean of the intensities of the contiguous regions. Consider a simple square image of uniform grey level α with a smaller embedded square of uniform grey level β. The nearest horizontal neighbour grey level cooccurrence matrix of this image will contain delta functions at (α, α) and (β, β) which correspond to the mean intensities of the two regions and whose magnitude is proportional to the size of the regions. There will also be much smaller delta functions at (α, β) and (β, α) corresponding to the boundaries between the regions, the magnitude of these delta functions will be proportional to the length of the vertical boundary between the two regions. If the image was corrupted by additive gaussian noise of standard deviation σ then the delta functions in the cooccurrence matrix would become gaussian.

The cooccurrence matrix can be labelled to indicate those parts of the matrix which correspond to particular parts of the image. For example, an ellipse drawn at a radius of 3σ from the centres of the distributions could be used as a decision boundary on whether contributing pixel pairs were from region α, β or the boundary between the two regions. This example of a labelled cooccurrence matrix is shown in figure 1, with two region distributions centred on (α, α) and (β, β) and two boundary distributions centred on (α, β) and (β, α). A cooccurrence matrix may be identified as a feature space, with classification regions defined therein by fitting binormal distributions to the region and boundary distributions, the ellipses so obtained defining the decision boundaries.

For images which are corrupted by uniform additive gaussian noise, the region distributions are well approximated by elliptical Gaussian distributions centred on, or near, the leading diagonal of the matrix, with major axes along the leading diagonal. The minor axes, normal to the diagonal, are proportional to the standard deviation of the noise in the image and are therefore constant and may be determined from the noise statistics[6], while the along diagonal axes may be determined by analysing the histogram formed from the leading diagonal of the cooccurrence matrix. The edge distributions can be approximated by circular gaussian distributions centred off the diagonal and on the paired mean intensities of two different regions.

2 Segmentation

An image is segmented relative to a given cooccurrence direction by mapping the pairs of convolved intensities, $(\sum_{k=1}^{n} E(k)i(\mathbf{x} - k\mathbf{\Delta}), \sum_{k=1}^{n} E(k)i(\mathbf{x} + k\mathbf{\Delta}))$ to a location within the cooccurrence matrix, $S_{\mathbf{\Delta}}(i,j)$, and then assigning the pixel at $\mathbf{x}$ to the region class or edge defined by the decision region within which (i,j) lies in cooccurrence space.

The result is that each pixel of the image has been assigned either to a distinct region class or is classified as an edge pixel. A separate classification is made for each of two orthogonal cooccurrence directions. Inconsistency of region classification may occur between the two cooccurrence directions. Such ambiguity is removed by assignment to the class with mean closest to the mapped intensity pair. A pixel is classified as edge if an edge strength exceed a given threshold.

Should the mean intensities of contiguous regions be 'similar' then there may be a significant overlap between the region and edge classification ellipses in cooccurrence space. In this case, the cooccurrence space is labelled as edge where the distributions overlap.

3 Edge Thinning

The edges are of widths of the order of the extent of the Canny operator. Each edge pixel has associated with it the edge strengths in the two orthogonal directions. Hence a non-maximal suppression algorithm[4] may be employed. In conventional non-maximal suppression, an edge point would be eliminated if it is not a local maxima within the range of influence of the Canny operator. In this work, the edge point is relabelled as region, the actual label being determined by the region of which the local maxima is a boundary.

So far, the cooccurrence space has been labelled in a simplistic way. By extending this labelling, thresholds which can be used in the hysteresis post-processing can be built into the labelled cooccurrence space.

If the matrix is summed parallel to the leading diagonal

$$H(k) = \sum_{i=-N}^{N} \sum_{j=-N}^{N} S_{\mathbf{\Delta}}(i,j) \qquad k = 1 + j, \quad i,j = 1,2...N \quad (4$$

then the resultant histogram is essentially the histogram of edge strengths in the convolution of the image and the canny edge operator.

In conventional hysteresis processing, edge strengths above an upper threshold T_u are considered edge points and are used as seeds. Pixels which have an edge strength above a lower threshold T_l are considered to be edge points if they are connected to a seed point (possibly via other non-seed points which have an edge strength above T_l).

In our labelling of the cooccurrence space, T_u corresponds to the cross-diagonal spread of the region ellipses. Where region ellipses overlap, T_u may be locally reduced. The lower threshold T_l corresponds to overlaying a second labelling of the matrix using ellipses with smaller cross-diagonal spreads. These ideas are illustrated in figure 2 which shows the basic labelling of the feature space for a high threshold T_u (bold lines) and the labelling for the lower threshold T_l (dotted lines). The upper threshold has been locally changed near the intersection of the two region distributions.

Although the hysteresis post-processing may appear similar to that of reference 4, there is the distinct advantage that the thresholds are being determined from the data rather than supplied externally and that both the thresholds T_u and T_l are not necessarily uniform throughout the image. Both T_u and T_l may be locally reduced if there is evidence for this. Furthermore, this technique also results in a segmentation of the major regions in the image.

4 Examples and Comparisons

A typical forward looking infrared (FLIR) image is shown in Figure 3: the picture is of a bridge over a river with a hillside and cliff in the background. Hot (white) parts of the cliff are reflected in the river, as is the sky, parts of the bridge and hillsides. A seven pixel resolution realisation of the Canny operator as derived by Spacek [7] was utilised to segment the image as described above. The resulting segmentation into four distinct region classes and boundaries is shown in Figure 4: the boundaries are shown as black. The non-maxima suppression and hysteresis processing described in the previous section has been applied and the results are shown in figure 5. Note how many of the fine lines have been obtained along the edge of the river below the base of the cliff. This is a part of the image in which it is very difficult to discern the river by eye and on which other segmentation and edge detection techniques generally fail. These lines are particularly obvious in figure 6 which shows only the edge map.

5 Conclusions

In this paper a technique has been presented for combining a simultaneous image segmentation and edge detection technique with a standard post- processing technique for improving edge detection. The majority of the proposed technique is inherently parallel and is believed to be, by its formalism, better than comparable edge detection techniques. In addition, the technique results in both a segmentation and an edge map, a characteristic missing from the majority of techniques.

References

[1] **Haralick, R.M., Shanmaugan, K. and Its'Hak Dinstein** "Texture Features for Image Classification" *IEEE Trans. Sys. Man. Cyber.* Vol. **SMC-3** (1973) pp 610-621.

[2] **Haddon, J.F. and Boyce, J.F.** "Simultaneous Image Segmentation and Edge Detection" *Third International Conference on Image Processing and its Applications* 18-20 July 1989 The Institution of Electrical Engineers, pp 411-415.

[3] **Canny, J.F.** "A Computational Approach to Edge Detection", *IEEE Patt. Anal. Machine Intell.* Vol. **PAMI-8** (1986) pp 679-698.

[4] **Canny, J.F.** "Finding Edges and Lines in Images" *MIT Report* MIT AI-TR-720 (1983) .

[5] **Haddon, J.F. and Boyce, J.F.** " Image Segmentation by Unifying Region and Boundary Information" *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Sept 1990.

[6] **Haddon, J.F.** " Generalised Threshold Selection for Edge Detection" *Pattern Recognition*, Vol 21, No 3, pp195-203, 1988

[7] **Spacek, L.A.** "Edge Detection and Motion Detection" *Image and Vision Computing* Vol. 4 (1986) pp 43-56.

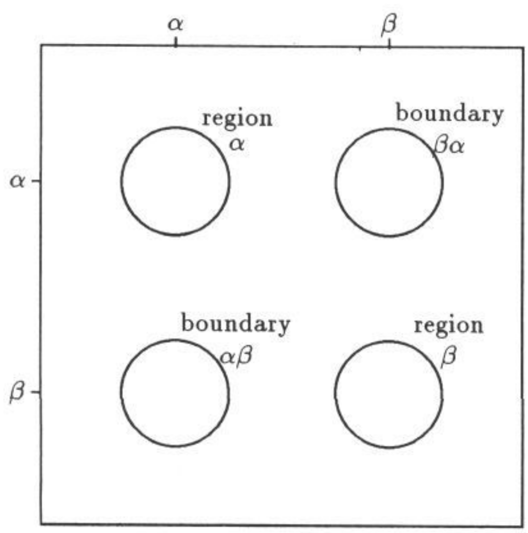

Figure 1 Labelled cooccurrence space

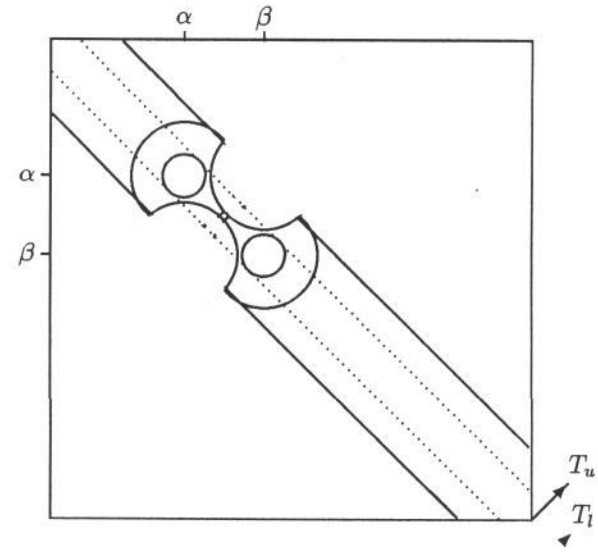

Figure 2 Labelled cooccurrence space for hysteresis post-processing

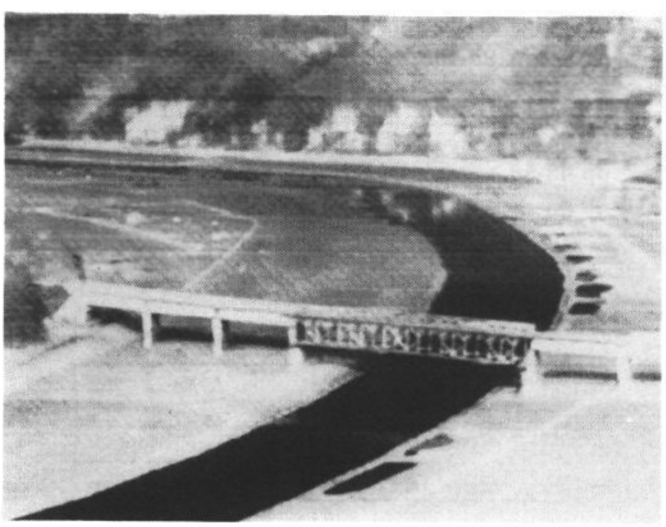

Figure 3 A forward look infrared image of a bridge over a river.

Figure 4 Segmentation and edge image of the FLIR image using a seven point Canny cooccurrence matrix

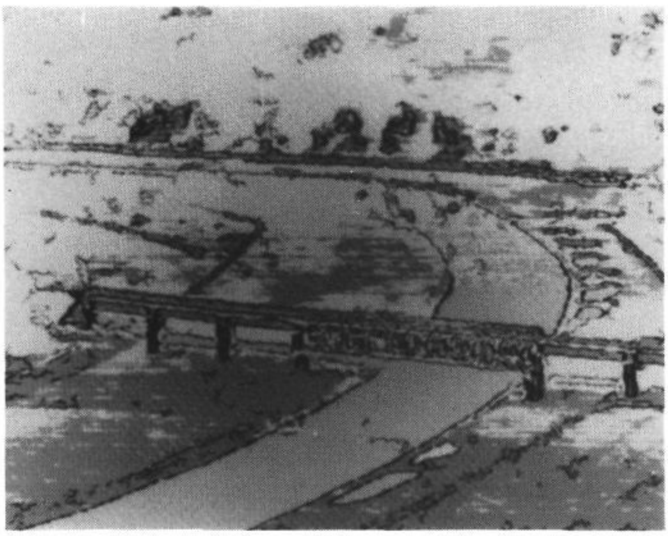

Figure 5 The result of non-maximally suppressing the edges in figure 4 and applying hysteresis post processing.

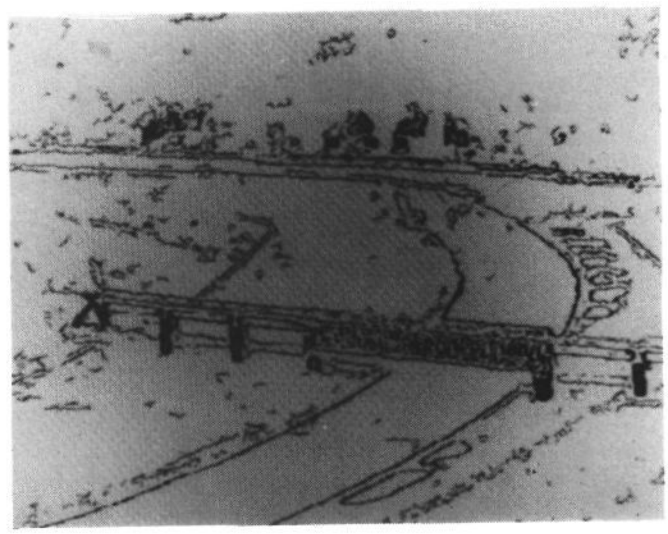

Figure 6 The edge map of figure 5.