IMAGE CONTENT DESCRIPTORS - THE DETECTION STAGE[†]

P G Ducksbury[1] , M J Varga[1], Z Y Xie[2] , J M Brady[2]

1 Introduction

This work starts to address the difficulties involved in the automatic generation of image content descriptors. These descriptors can be used for the indexing, matching and searching of data in an image database. The current stage of the work has concentrated on the development/collation of a set of feature detectors covering points (*corners*), lines (*edges, ridges*) and regions which will operate on a variety of image sensors (eg SPOT, IRLS - *Infra-red Linescan,* dd5 - *Russian satellite data).*

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

All subsequent stages such as searching or matching of descriptions 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 which may then be combined into a composite estimate of the feature(s) that are required.

2 Preprocessing

For some detectors it may be desirable to preprocess the imagery to filter out noise prior to the main feature detection stage. The two filtering algorithms that we describe in this section are adaptive morphology, Wu[1] and the SUSAN smoothing filter, Smith[2].

2.1 Adaptive Morphology

Mathematical Morphology is a non-linear filtering technique in which the operations are set-to-set transformations of an image by a structuring element. The operations of opening (*erosion followed by dilation*) and closing (*dilation followed by erosion*) are idempotent however they are not very good at preserving image structure. The operations preserve parts of image structures that the structuring element boundary can reach but destroy those parts that it cannot reach.

Adaptive morphology has the advantage of smoothing the image yet preserving important structure within the image and retaining the property of idempotency (*using two or more structuring elements in sequence, regardless of order is identical to just using the largest structuring element*). If we imagine a gray level intensity image as a surface in 3-dimensions then in adaptive morphology the shape of a structuring element can alter in any direction to adapt to the image surface. However it may not grow upward or downward or shrink into its original shape.

For adaptive opening, the structuring element is moved under the image surface and can stretch itself to fit spatial corners within the surface. However if a part of the image surface is smaller than the structuring element then that part of the image surface will be removed *(as the structuring element cannot shrink into itself*). Adaptive closing is the dual of adaptive opening. Result of these operations are shown at the end of the next section.

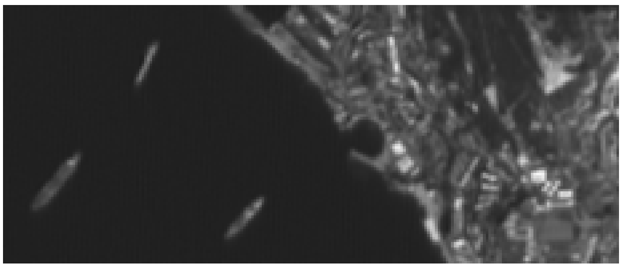
[†] Presented at an IEE Colloquium on 'Intelligent Image Databases', London, 22-May-1996.
[1] Defence Research Agency, St Andrews Road, Malvern, WR14 3PS, UK, email : ducksbury@signal.dra.hmg.gb
[2] Department of Engineering Science, University of Oxford, Oxford, UK.

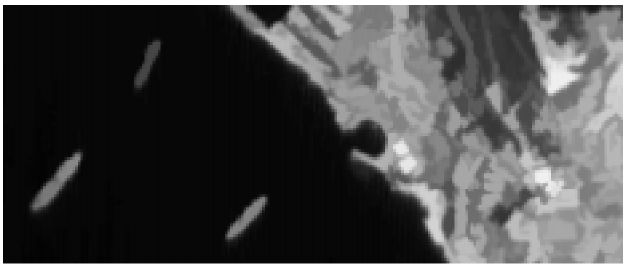
2.2 SUSAN smoothing filter

The SUSAN (Smallest Univalue Segment Assimilating Nucleus) smoothing (*noise reduction*) filter preserves image structure by only applying the smoothing operator over a given pixels neighbours that form part of a homogenous region. Homogenous regions being taken here as those for which pixel values are (*roughly)* similar, these regions being refered to as USAN's (Univalue Segment Assimilating Nucleus). The filter works by forming the average over all the pixels which fall within the USAN but not those outside it. As a direct result of this the fundamental image structure should remain unaltered.

The figures below illustrate the results of filtering the raw image (*left*), with the adaptive morphology using the mean of a close-open and an open-close (COMOC) (*center*) and with the SUSAN smoothing filter (*right*).

The figures below illustrates another example of the adaptive morphology in this case being ship detection in SPOT imagery, the requirement was to filter out the bright peaks on the decks of the ships (*left - raw image*) but not to destroy the fundamental shape, as would tend to happen in basic morphology (*right - processed image*).





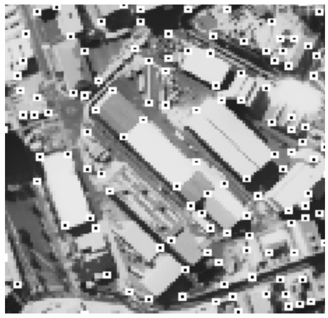
3 Feature Detection

The feature detectors used cover the detection of points, lines and regions. Point detection (Smith[2]) edge detection (Canny[3], Smith[2]), ridge detection (Petrou[4]), region detectors such as Fractal and Wavelet (Xie[5]) and Pearl Bayes (*belief*) Networks (Ducksbury[6]).

3.1 Point (corner) Detection

The SUSAN corner detector is a different option of the same suite of software as was mentioned previously for the smoothing filter (*and will be mentioned again for edge/line detection*). The detector makes no assumptions about local image structure around a well localised point. It also does not look for 'points of interest'. It analyses different regions separately, using direct local measurements to locate places where individual region boundaries have high curvature (*ie finds corners formed by single regions*). For junctions involving more than two regions such as 'T' junctions, more than one region may contribute to the detection of the 'corner'.

The detector operates by placing the kernal over a given pixel and computing the USAN which is then used to generate a corner strength image prior to testing for false positives. Non-maximal suppression is used for localisation of the corner. The figure on the right shows (*an enlarged section*) the result of applying the corner detection to a dd5 image covering the area around Plymouth docks.



3.2 Edge Detection - Canny

Canny[3] developed a detection and localisation criteria for a class of edges. His edge detector relies upon the first derivative of a gaussian and is noted as being almost identical to the 1-D Marr-Hildreth[7] operator. However in two dimensions the directional properties of his detector enhance its detection and localisation when compared with the Laplacian used by Marr-Hildreth. The Marr-Hildreth technique does not require thresholding but this can be advantageous for a first derivative operator.

After applying a convolution of an image with the first derivative of a gaussian, non-maximal suppression is applied prior to a hysteresis threshold tracking stage. In addition to this standard algorithm we have considered an additional stage. Here for each ‘acceptable’ edge, compute the ratio of the number of pixels in the edge to the number of line segments (*obtained from a recursive linear approximation*) that would be required to represent the edge to some given degree of accuracy. This ratio can then be described as a sort of ‘fractal’ descriminent for removing small unwanted ‘wriggly’ edges.

3.3 Edge Detection - SUSAN

The SUSAN edge detector[2] is a different option of the same suite of software as has mentioned previously for the smoothing filter and corner detection. The detector operates by placing the kernal over a given pixel and computing the USAN which is then used to generate a corner strength image. Moment calculations are then applied to the USAN in order to obtain edge direction. Finally non-maximal suppression, thinning and sub-pixel estimation, are applied.

3.4 Ridge Detection - Petrou

Most edge detectors that have been developed are based around the assumption of detecting step edges and have been optimised with this goal in mind. The work by Petrou (*on a previous DTI-IED (1936) project*) presents a theory for the development of optimal convolution filters for the identification of wide linear features, such as, roads, canals, hedges and rivers. Undoubtedly conventional edge detectors that have been optimised for step edge detection will yield some response upon encountering wide linear structures but this is unlikely to be as good as from an purposely designed filter.

3.5 Region Detection - Belief Networks

The idea behind the belief network approach is one of combining multiple sources of information in a network to study the changes in beliefs for a set of events. The idea was originally proposed by Pearl[8] in the application area of medical diagnosis. Previous work at DRA Malvern[6] had considered the use of belief networks in image processing for the detection of urban regions in IRLS imagery, followed by the detection of driveable regions in autonomous land vehicle imagery. This work proved successful (*it was also demonstrated as having considerable parallel processing potential*) and has been reassessed for all three image types chosen for this project.

The idea behind the use of belief networks in this application is to decompose the image into a set of small windows (*typically 8x8 or 16x16 pixels*) and then to compute a set of suitable statistical measures which give some indication of the type of texture in the window. The belief network is a directed acyclic graph which is then used to combine these measures over multi-scale into a belief that represents the content of the region. In this work, two scales where used and three states (*low, medium and high*) for the belief variable that represented the required region.

Prior knowledge is required in terms of a set of fixed multi-dimensional conditional probabilities which relate the inputs (*causal information*) to the outputs of a given node within the graph. A set of rules where developed which follow similar lines to those used by Devijver[9] for his work on segmentation using Markov Models. They 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.

3.6 Region Detection - Fractal

The work on fractal analysis Xie[5] 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[10]). This measure being independent of dimension. The work developed by Oxford consists of a multiscale analysis algorithm for texture discrimination.

3.7 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.

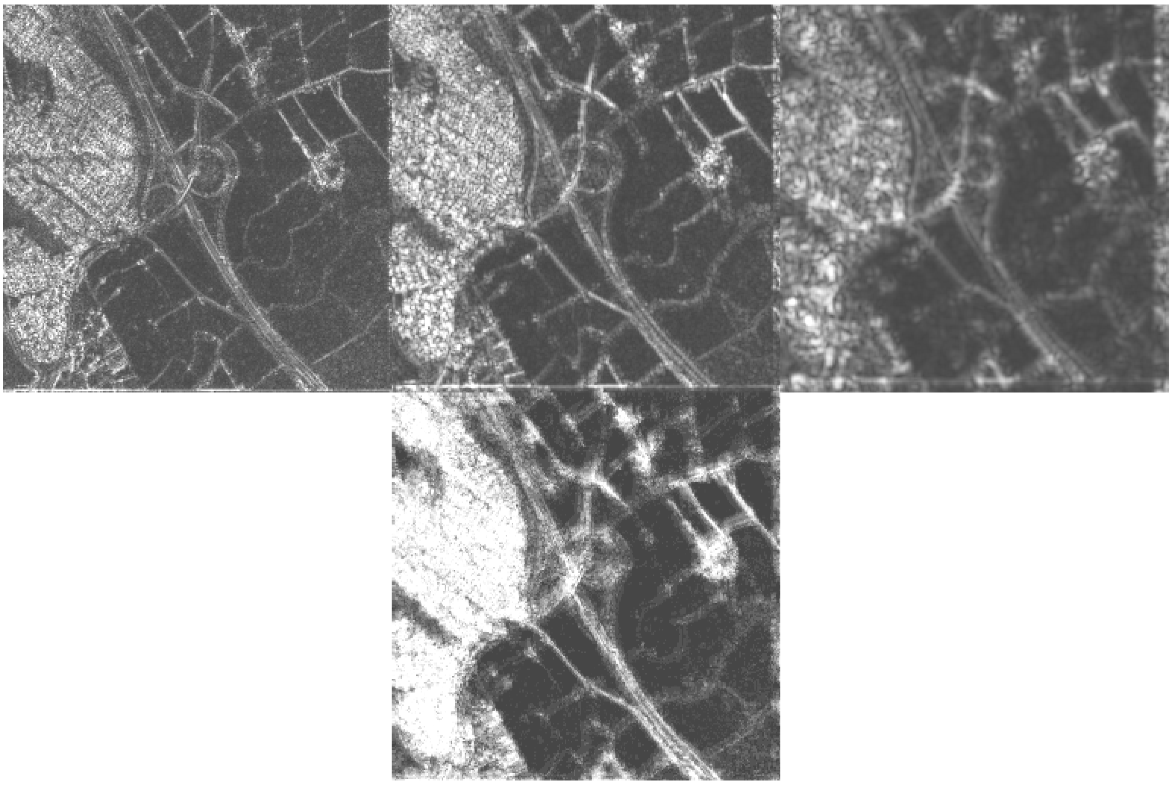
Generation of a set of local energy images (which removes phase dependency from the detail images).

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

Texture feature clustering and boundary detection.

4 Feature Combination

Some initial work was carried out into using the Belief Network approach for combination of all three region algorithms (*belief network, fractal, wavelet*) into a single composite region, as shown in the figure directly below[†]. This is effectively using the same approach but at a higher level of abstraction (*ie processing algorithms rather than images*). This approach has also been evaluated for combining the multiscale output of the wavelet algorithm into a single feature image *(top row - wavelet scalespace output, bottom row - combined image*).



5 Conclusions

The work described in this paper consists of a fairly comprehensive set of feature detectors covering points, lines and regions together with related utilities for filtering (*although only one point detector is currently used the structure is such that other algorithms can readily be incorporated).* The algorithms have all been shown to produce good results on the three classes of image type (*SPOT, IRLS, dd5*) that are being considered. This

† Outline is the result of combining the three surface images, no postprocessing on result outline

set of algorithms will form the baseline for future work on the project. The algorithms have all been incorporated into the first version of a demonstrator which runs on a Sun Unix Sparc 20 (Solaris 2.4) platform.

The future work falls into a number of distinct categories, namely, continuation with efforts into feature detection, confidence measures, combination of feature detectors both similar (*eg edge with edge*, *edge with region*) and dissimilar (*eg region with edge*), region competition[11] , shape representation and finally the problem of matching.

6 Acknowledgements

Work described in this paper forms part of a collaborative project between DRA Malvern and Oxford University. In addition to this we wish to thank Dr S Smith from DRA Chertsey for making available the SUSAN image processing software.

7 References

1 Wu DeQuan., 'Morphological filters in Image Analysis', PhD Thesis, Dept of Engineering Science, University of Oxford, 1994.

2 Smith S.M., Brady J.M., 'A new approach to low level image processing', DRA Technical Report TR95SMS1, DRA Chertsey, 1994.

3 Canny J., 'A Computational Approach to Edge Detection', IEEE Trans PAMI, Vol. 8, No. 6, Nov. 1986.

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

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

6 Ducksbury P.G., 'Parallel Texture Region Segmentation using a Pearl Bayes Network', British Machine Vision Conference, BMVC-93, University of Surrey, Sep 1993, pp 187-196.

7 Marr D., Hildreth E., 'Theory of Edge Detection', Proc. R. Soc. London, B 207:187-217, 1980.

8 Pearl J., 'Fusion, Propagation and structuring in Belief Networks', Artificial Intelligence 29, pp 241-288, 1986.

9 Devijver P.A., 'Real-time modelling 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.

10 Falconer K., 'Fractal Geometry', John Wiley & Sons, 1990.

11 Zhu S.C., Lee T.S., Yuille A.L., 'Region Competition : Unifying Snakes, Region Growing, Energy/Bayes/MDL for Multi-band Image Segmentation', IEEE ICCV-95, MIT, Boston, pp 416-423, June 1995.