<u>Some first attempts at feature extraction from 32 metre resolution satellite images</u>.

<u>Introduction</u>.

Some 10 years ago a book entitled 'Computer Vision - a unified, biologically-inspired approach' was published by the present author (Ref. 1), this book collecting together a powerful & versatile selection of software processes which had been shown to be capable, in pixel-related terms, of emulating a wide range of levels of threshold performance known to exist in human vision. Amongst other things, this suite of software had been shown to be capable of sensing position & orientation of ***individual*** points on brightness or chrominance discontinuities in images to better than 0.1 pixels in radial position from the pixel centre and to better than 1 degree in local orientation of the edge. Additionally, where two or more images were available which had been captured at slightly different times or from slightly different viewpoints, it had been shown possible, in addition to the local edge position & orientation, to determine - again for ***every*** pixel on a brightness or chrominance discontinuity - the local component of motion of the edge orthogonal to the local edge orientation (which subsequently can be used in groups to determine the vector motion of individual segmented regions in the image) or the local stereo disparity (from which the differential distance with respect to the 'fusion plane' for the stereo views can be directly derived).

Over the intervening years it has been possible to refine these & many related processes and to implement them all for use with MS Windows so that large images (up to at least 1000 x 1000 pixels) can be totally analysed to provide tabulations of the characteristics of all edge (or profile) points and a segmentation of the scene into regions based on the profile points sensed & selected. Typically these tabulations can run into many tens or even hundreds of thousands of profile points and several hundred regions, even for input images of moderate size.

Supplementary software is then available which can generate 'cartoon' reconstructions of the original image - in full colour if the original was in colour - and/or sub-pixel 'edge maps' showing the actual sub-pixel position, orientation and (if desired) edge strength of selected sub-groups of profile points.

The entire processing for a given image can be carried out in at most a few tens of seconds - often just a matter of a few seconds - (once appropriate conditions have been keyed into command files from which the processes are run).

<u>Initial considerations</u>.

The purpose of the present exercise was to take a few 'typical' 32 metre resolution satellite images as kindly provided by SSTL and to explore what sort of 'features' could be extracted with what fidelity. Since the author had been unable to determine just what 'features' might be of interest, it was decided, for this first exploration, to select what seemed to him to be a ***variety*** of possibly interesting features and to show what could be done with them.

As already implied, a typical ***total*** analysis of a full satellite image could be expected to yield tabular data for tens or hundreds of thousands of profile points and at probably up to several thousand segmented regions. Furthermore, with the very coarse ground resolution available, one part of the processing suite which was known to be of considerable value in helping to derive 'intelligent' data from such coarse resolution involved an initial image scaling before offering image data to the ***main*** processes. For the ***full*** satellite images this inevitably meant a large increase of image area over & above the already large image size of the original images, with attendant further large increase in both number of profile points sensed and regions segmented. All in all, this was considered grossly excessive for an initial study (and possibly beyond the memory handling capacity of a personal computer). Instead, therefore, it was decided to 'crop' fairly small portions out of some of the provided images, which nevertheless contained a wealth of potentially interesting features.

Since one of the original images was a monochrome image of the London area (ut21e-london-msi-nir.gif), it was deemed useful to use part of this which covered Central London, in particular including Westminster and the area around Buckingham Palace - which therefore contained a considerable quantity of ***well-known*** and varied features - Fig. 1.

Fig. 1. View of central London cropped from original image 'ut21e-london-msi-nir.gif'.

A second monochrome image containing a wealth of detail appeared to be the central part of of one of the images of New York (New York - U012.bmp). This monochrome image was chosen, in preference to the alternative false colour image (newyorkcity-msi-irg.bmp), because an initial visual inspection suggested that it was rather 'cleaner'. For this, since the author has been unable to date to access ***detailed*** ground maps of New York, the feature extraction had to be carried out somewhat 'blind'. So for the present he has concentrated on bridges, bridge approaches & coastline features in the cropped portion shown as Fig. 2.

Fig. 2. Small portion cropped from original image 'New York - UO12.bmp'-

Staying with New York images, the first of the group UT347600.bmp to UT347603.bmp was used to crop out a portion clearly including an airport - see Fig. 3. For this portion, attention was primarily focussed on the features associated with the airport - viz. runways & perimeter tracks, with some attention also to a nearby bridge.

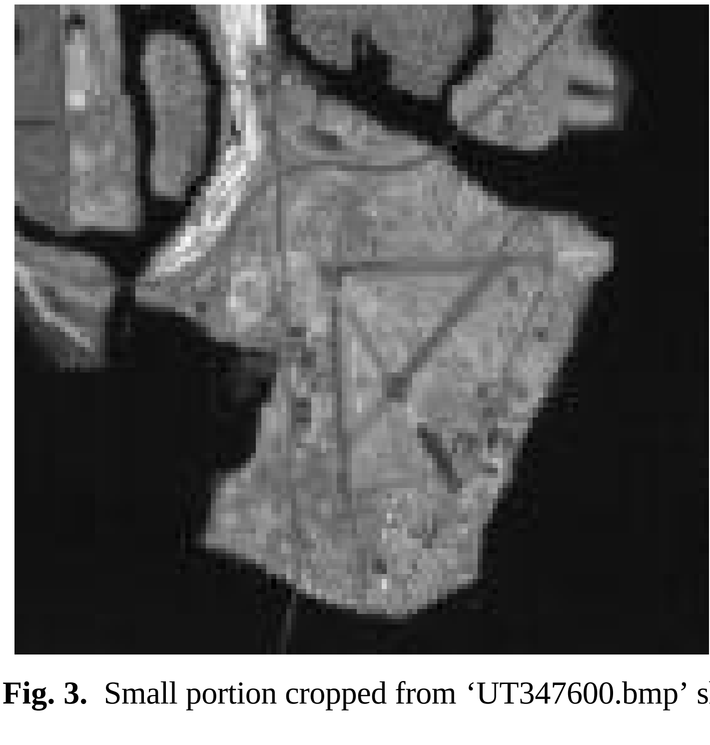


Fig. 3. Small portion cropped from 'UT347600.bmp' showing mainly airport.

A fourth image selected - this time to demonstrate at least some of the capabilities of the software for handling ***colour*** images - was that of Aukland, N.Z. (Aukland-msi-irg.bmp). From this a portion was cropped as shown in Fig. 4 For this cropped portion, apart from demonstration of a complete 'cartoon' regeneration (in full colour) of the entire portion of cropped image, again he has concentrated on the extraction of bridges and coastline features.

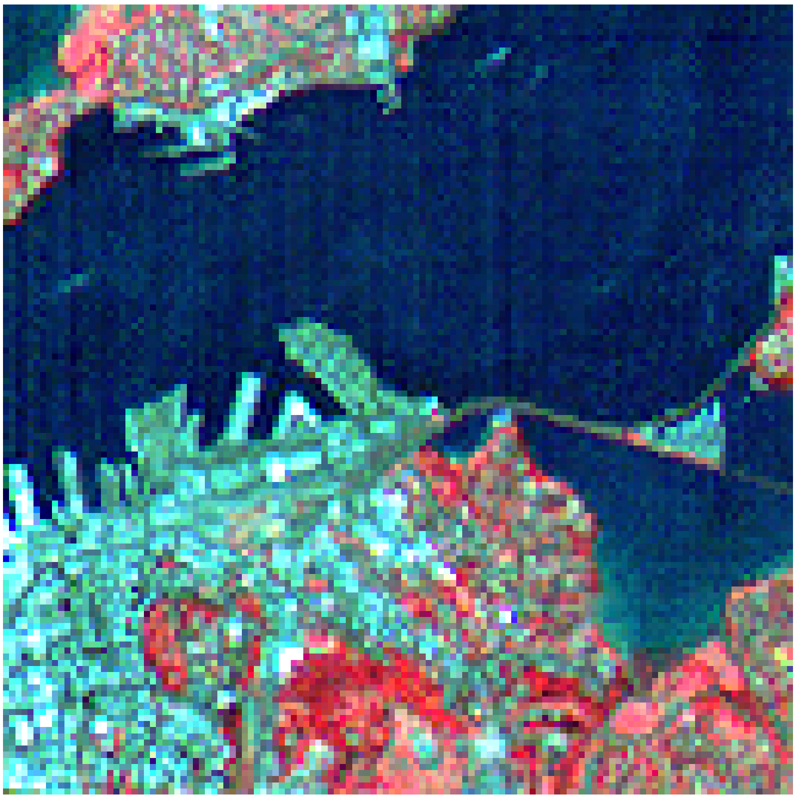


Fig. 4. Small portion cropped from 'Aukland-msi-irg.bmp' showing harbour, island & bridges

<u>Central London image analysis</u>.

An initial inspection of a zoomed up screen image showed two things very clearly. Firstly, although one could ***visually*** perceive the ***existence*** of familiar things like St. James' Park, Buckingham Palace grounds, Hyde Park, the lakes in St. James' Park & Buckingham Palace grounds, the Serpentine, major bridges across the Thames etc. ***when viewing at 1:1***, on zooming up most of these 'features' became nothing more than clusters of varied grey 'blobs' (see Fig. 5a, which is an enlargement of a rather tiny portion of Fig. 1 of size just 64 x 64 pixels, but within this covering most of Westminster). Whilst this was no surprise to the author - indeed, it is what one must inevitably expect from such a low ground resolution - it was immediately evident that one important immediate task in using the biologically-inspired image processing should be to endeavour to recover some measure of 'perception' of what the eye could see in the original at 1:1 scale. Past experience had shown that the way to handle such a situation was first to carry out some 'intelligent' interpolation of the input data and only ***then*** to attempt the ***main*** visual simulation processing.   By so doing, one effectively converts very small, apparently random, 2D clusters of 'blobs' which have any semblance of an effective 'centre of mass' into much more substantial groupings which begin to take on a semblance of recognisable 2D shape. At the same time, initially single pixel 'line features' (as the Thames bridges appear when zoomed up for screen viewing) are broadened, so that they present themselves as ***bar*** features which have two distinct edges and therefore have at least a potential chance of being segmented as individual regions in their own right. On the other hand, larger and more prominent features, such as the river Thames itself, can be expected to be adequately resolved at the ***original*** image scale and, as such , one can expect to be able to segment such features satisfactorily by direct processing at the original image scale.

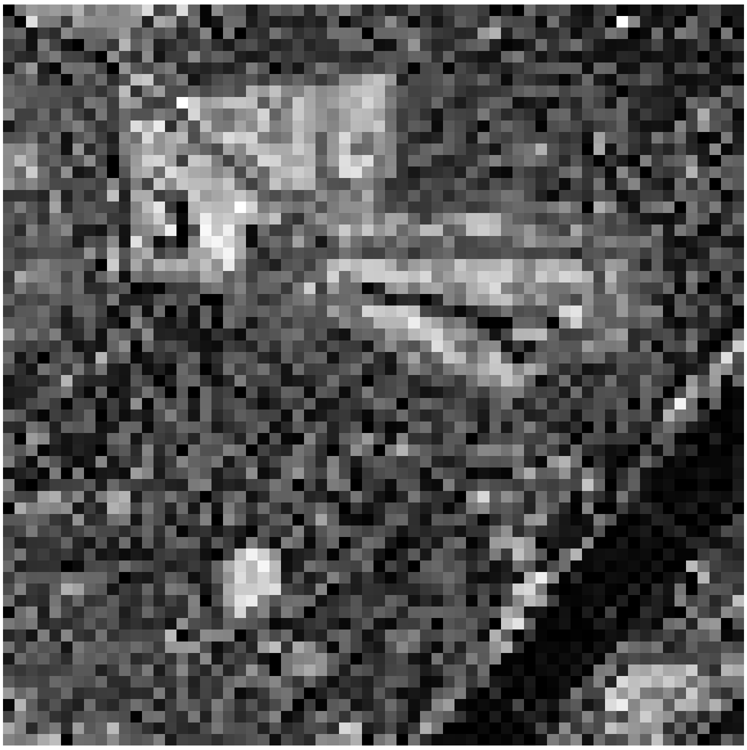


Fig. 5a. 64 x 64 pixel portion of Fig. 1 showing Westminster area.

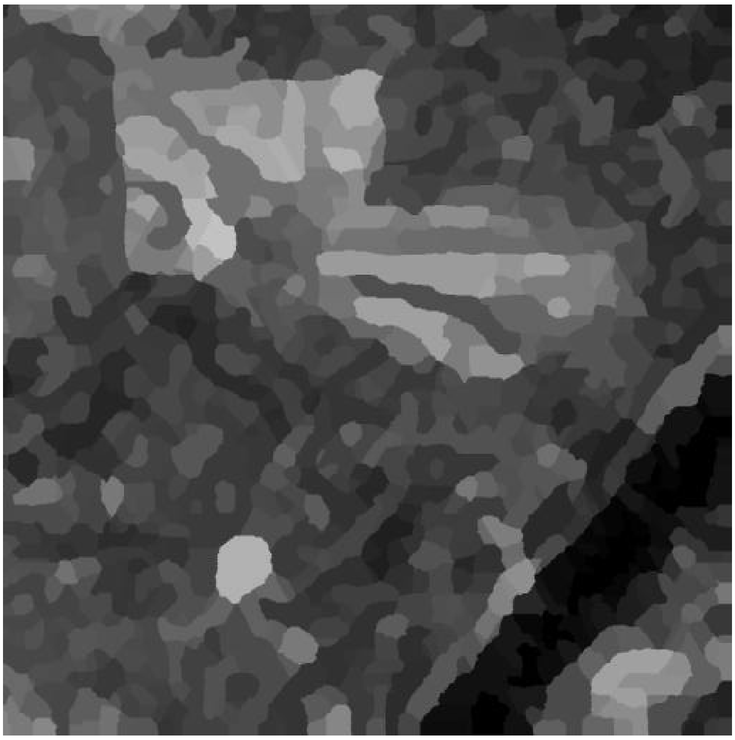


Fig. 5b. 'Cartoon' reconstruction of Fig. 5a from segmented regions.

With these things in mind, an attempt was made to process a portion of the original image containing the ‘snake’ of the river Thames at 1:1 scale and a second portion cropped out as shown in Fig. 5a after a X8 ‘intelligent’ 2D interpolation, in both cases carrying out a preprocessing of substantial 2D image blurring to soften edges as known to be exist at the level of image sensing in the retina of the human eye. From the X8 processing a ‘cartoon’ reconstruction was carried out which resulted in the image shown in Fig. 5b. From the 1:1 process the (single) region number appertaining to the ‘snake’ of the river Thames was determined and a sub-pixel reconstruction of the banks of the river was produced from the profile fragment tabulation for the said region. This is shown in Fig. 6a, where the individual ‘dipole’ bars provide a close representation of the sub-pixel position and orientation of the river banks at each pixel along the profile (although the graphical representation shown here is itself resolution limited by the plotting processes). For comparison, a X8 representation of the relevant portion of the original input image is shown at Fig. 6b.

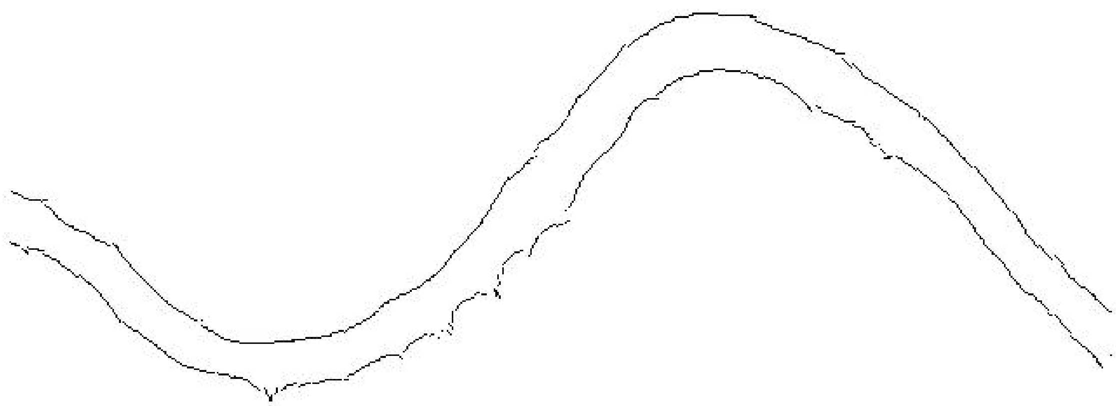


Fig. 6a. Sub-pixel plot of the banks of the Thames recovered from profile/region data files.

Fig. 6b. Zoomed representation of portion of input image used to derive Fig. 6a.

As an illustration of the additional capabilities of the overall processes, Fig. 6c shows the same sub-pixel profile output as Fig. 6a, but with "T's" plotted for each profile point, where the stems of the "T's" display the relative local edge strength at each point. It will be observed, as might be expected, that the edge contrast along the banks of the river varies substantially from location to location.

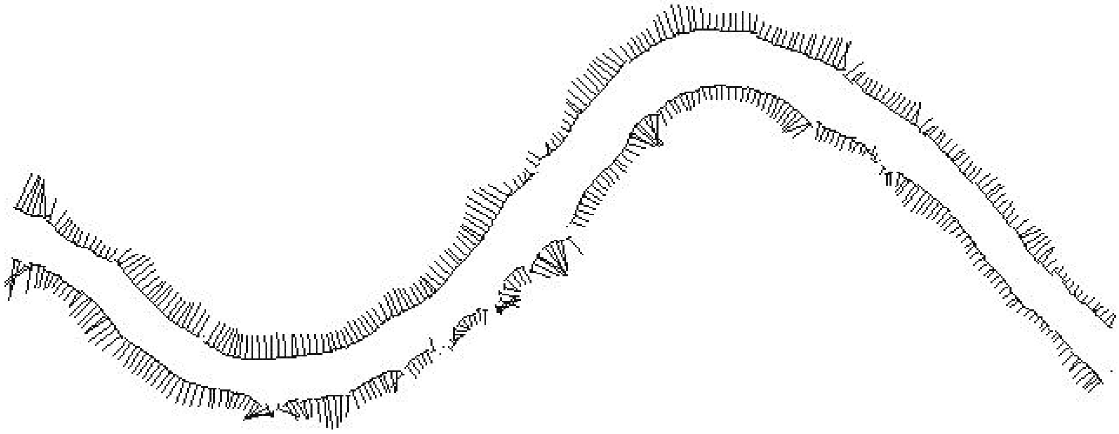


Fig. 6c. As Fig. 6a, but showing progressive variations of local relative edge contrast as "T's"

Finally, as an attempt to illustrate the association between the profile data and the region segmentation, Fig. 6d shows a 'cartoon' reconstruction of the area covered by Fig. 6b, with the actual profile defined region boundaries shown as an overlay. The types of display shown as Fig. 6a & b will not be repeated for the rest of the analyses in this report for sake of compactness.

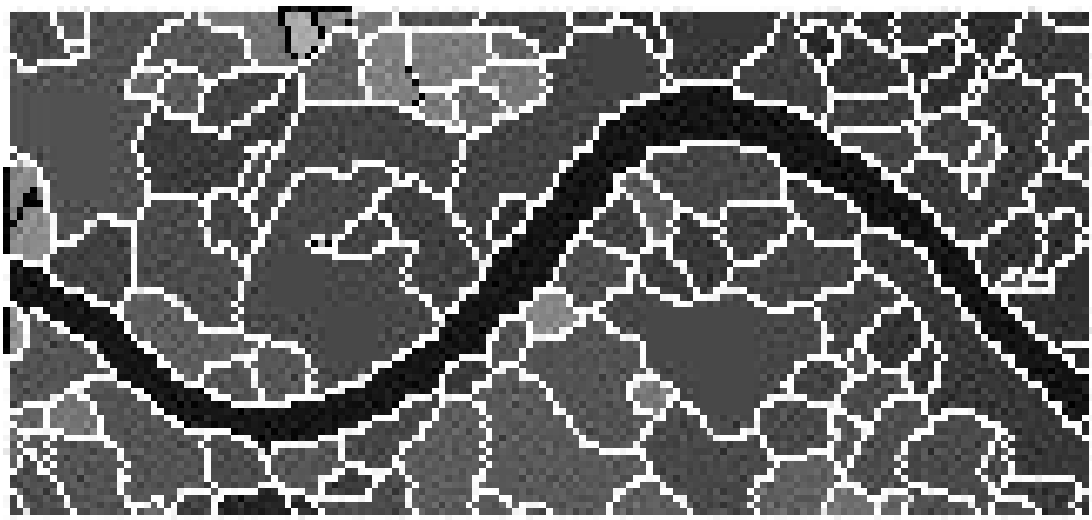
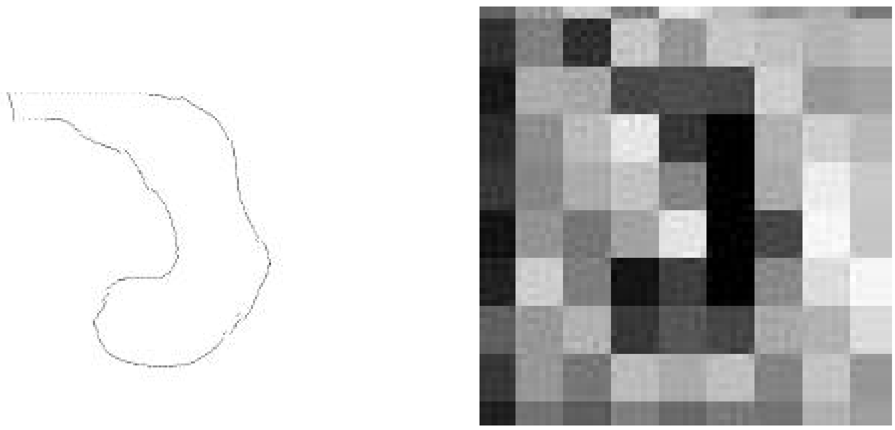


Fig. 6d. 'Cartoon' reconstruction from processing of Fig. 6b, showing overlayed profile-derived region boundaries.

Referring back to the X8 processing shown in Fig. 5b, it will be seen that, apart from the rather smoother appearance due to the sets of irregular shapes instead of individual scaled up pixels of Fig. 5a, the general structure of the scene is well retained.   It is possibly worth recording that this reconstruction is based on 35,600 profile points defining & segmenting 1164 regions, all from a 64 x 64 pixel portion of the original!

From the Fig. 5b ‘cartoon’ a number of features recognisable with the aid of local geographical & map knowledge were selected and for each of these the appropriate region numbers were determined. Then for each of the selected regions a sub-pixel profile map was produced in the same way as for the sub-pixel image of Fig. 6a. Features selected were the lake in the grounds of Buckingham Palace, the lake in St. James’ Park, the lawns behind Buckingham Palace (as widely publicised during the Queen’s Jubilee celebrations), the Houses of Parliament and Westminster Bridge. The resulting plots for each of these, together with scaled up representations of the relevant fragments of the original satellite image, are presented as Figs. 7 to 11.



(a) (b)

Fig. 7.   Sub-pixel profile plot of Buckingham Palace lake (a) and the original fragment (b).

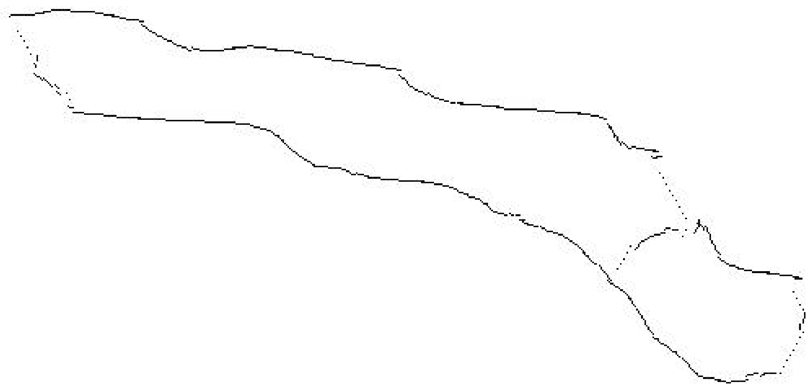


Fig. 8a. Sub-pixel profile plot of St. James' Park lake.

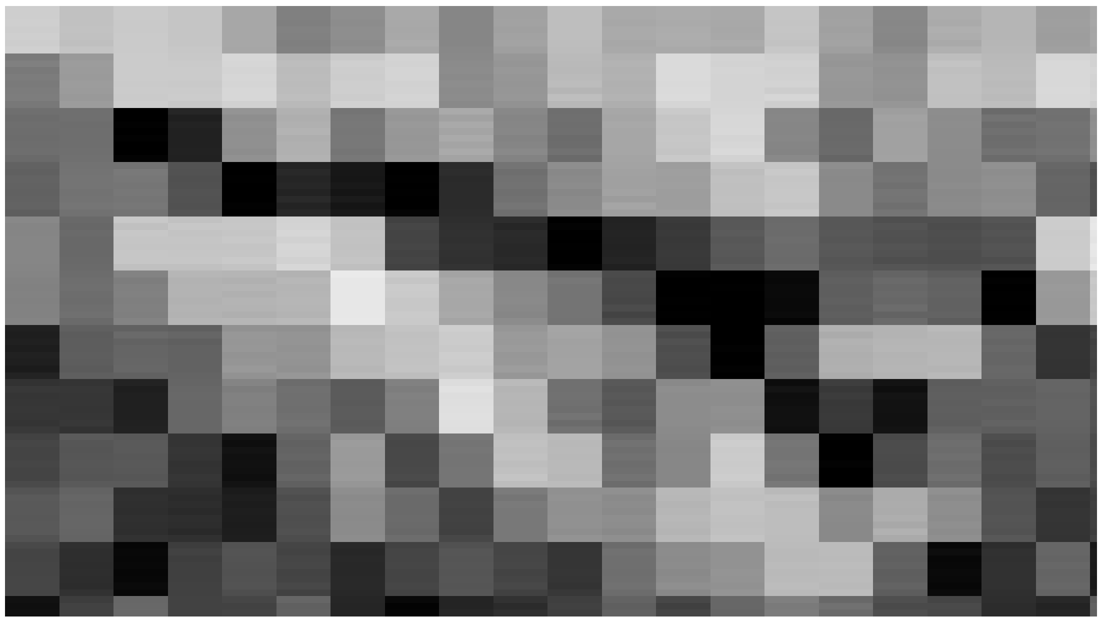
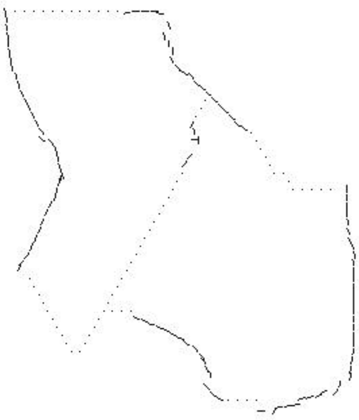
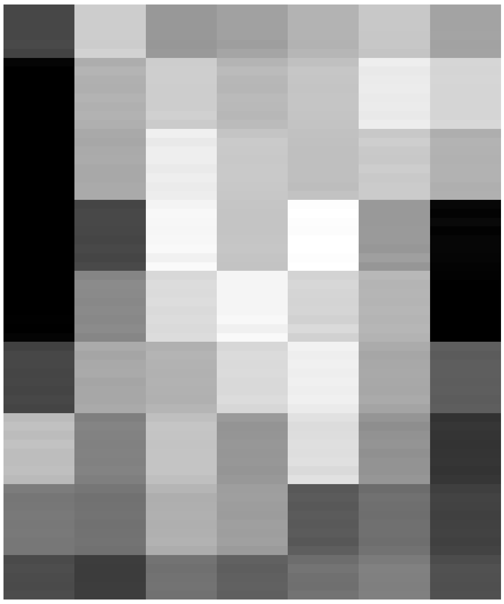


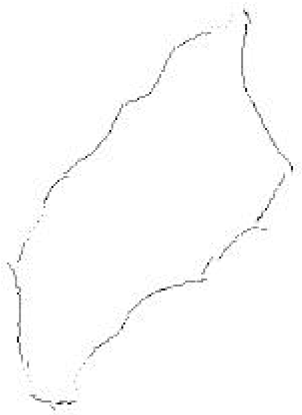
Fig. 8b. Enlargement of relevant area of input image for Fig. 8a.

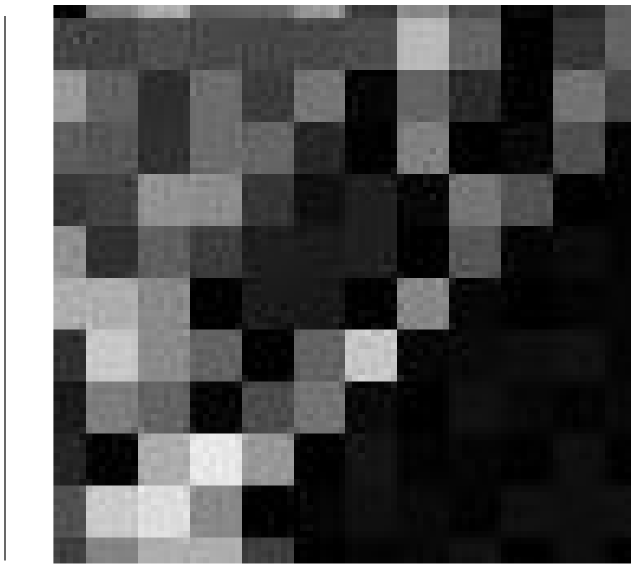




(a) (b)

Fig. 9. Buckingham Palace Lawns. Sub-pixel profile plot (a) and the approximate relevant area of the original image (b).

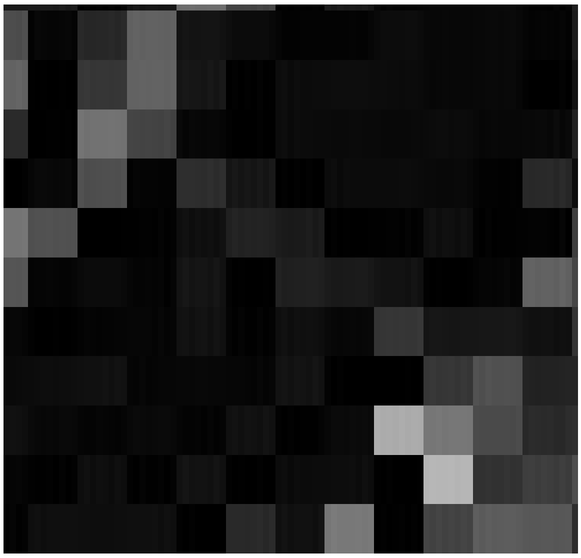




(a) (b)

Fig. 10. Houses of Parliament. Sub-pixel profile plot (a) and the approximate relevant area of the original image (b).

(a)



(b)

Fig. 11. Westminster Bridge. Sub-pixel profile plot (a) and approximate relevant area of original image (b).

In figures 7 & 8 it is considered that there is a very fair agreement between the sub-pixel profile plots and the actual forms of the two lakes as shown on street maps of London. In figures 9 to 11, however, although it is considered that a reasonable ***segmentation*** has been achieved, it is equally apparent that the resolution of the original image was not really sufficiently fine to be able to produce a a good profile ***form*** plot. It is believed that an increase of ground resolution of no more than 2 or 3 linear could result in good profile form plots of features such as these.

New York analysis.

Initially a portion of the New York image was selected which included parts of the waterways, Manhatten Island and three major bridges (see Fig. 2). It was hoped then to be able, by use of intelligent interpolation followed by standard processing, to segment all three bridges and some rather prominent & obvious approaches to the bridges, as well as producing a good reconstruction of the coastline. The ***basic*** results of this attempt are shown as a 'cartoon' at Fig. 12, this reconstruction being based on 46100 profile points defining & segmenting 1805 regions. It will be seen that, by and large, all the main image details have been succesfully segmented and reproduced in the 'cartoon'. When, however, an attempt was made to separate out the individual bridges & their approaches, it was found that, although the visual ***appearance*** of the said bridges & approaches in Fig. 12 looks like single regions, in actual fact they tend to be comprised of groups of several regions which have ***visually indistinguishable grey levels, but are nevertheless*** discretely different as far as the computer is concerned. This is a well understood problem of region segmentation which also applies to human vision in the limit. Any sensing system (biological systems included) must, of necessity, employ some small but finite threshold in initially selecting profile points during processing. But in the real world there is no such thing as a ***hard*** threshold situation - rather, all profiles tend to have subtly changing strength along them. Raising the threshold will reduce the tendency for generation of lots of regions with closely similar grey levels, but with the danger that parts of some ***wanted*** profiles may be lost, resulting in desirable segmentation not being achieved. For really faithful scene reconstruction, therefore, it is very important to use a very low threshold for initial selection of potential candidates for profile points. Intelligence or other prior knowledge must then be used to ‘merge’ adjacent regions which are ‘similar’ (according to some criteria). The present author has available a supplementary program which is intended to take over-rich region maps and apply semi-intelligent comparison testing of adjacent pairs of regions iteratively in an attempt to ‘thin down’ the number of regions existing ***without losing important structure***. However, use of this supplementary program is considered to be beyond the scope of this initial study.

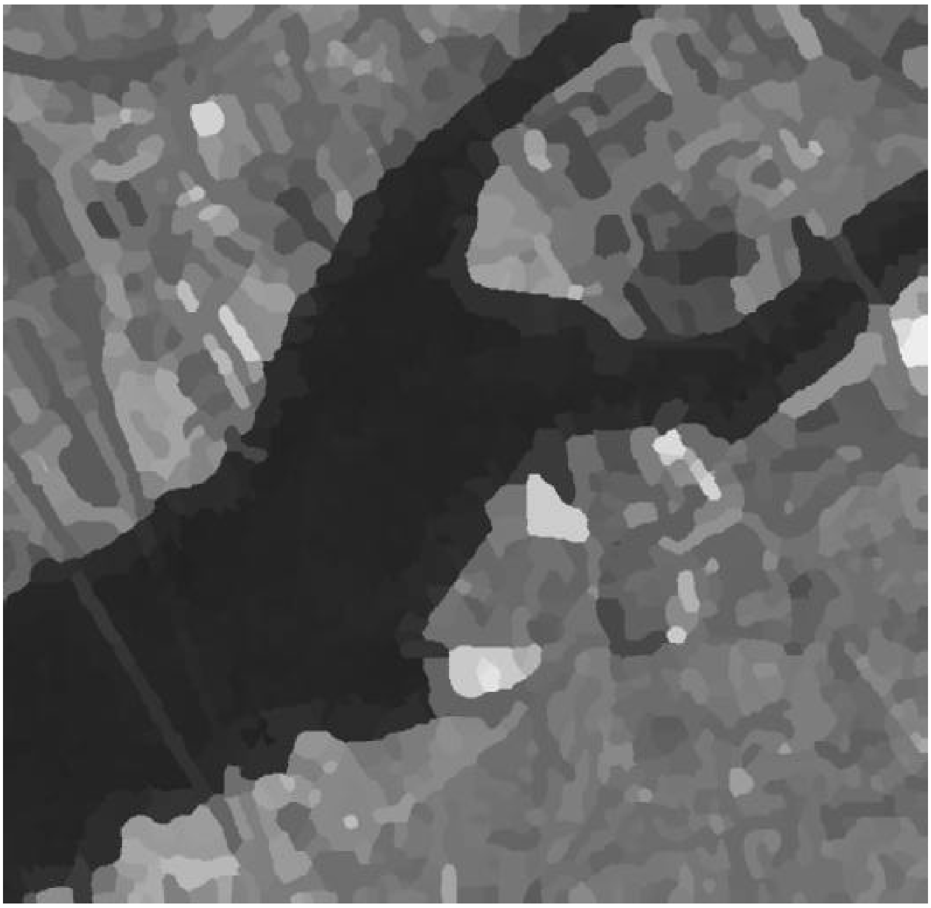
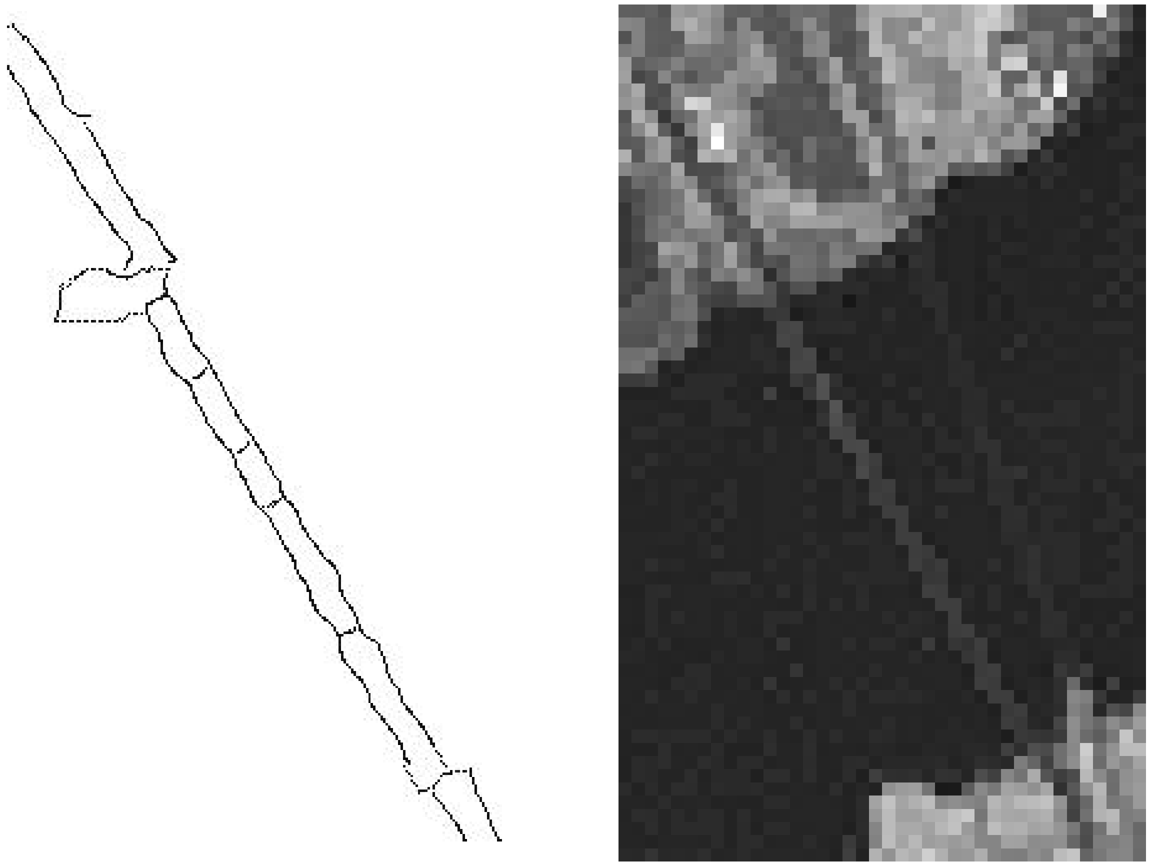


Fig. 12. ‘Cartoon’ reconstruction from the intelligently scale processing of Fig. 2.

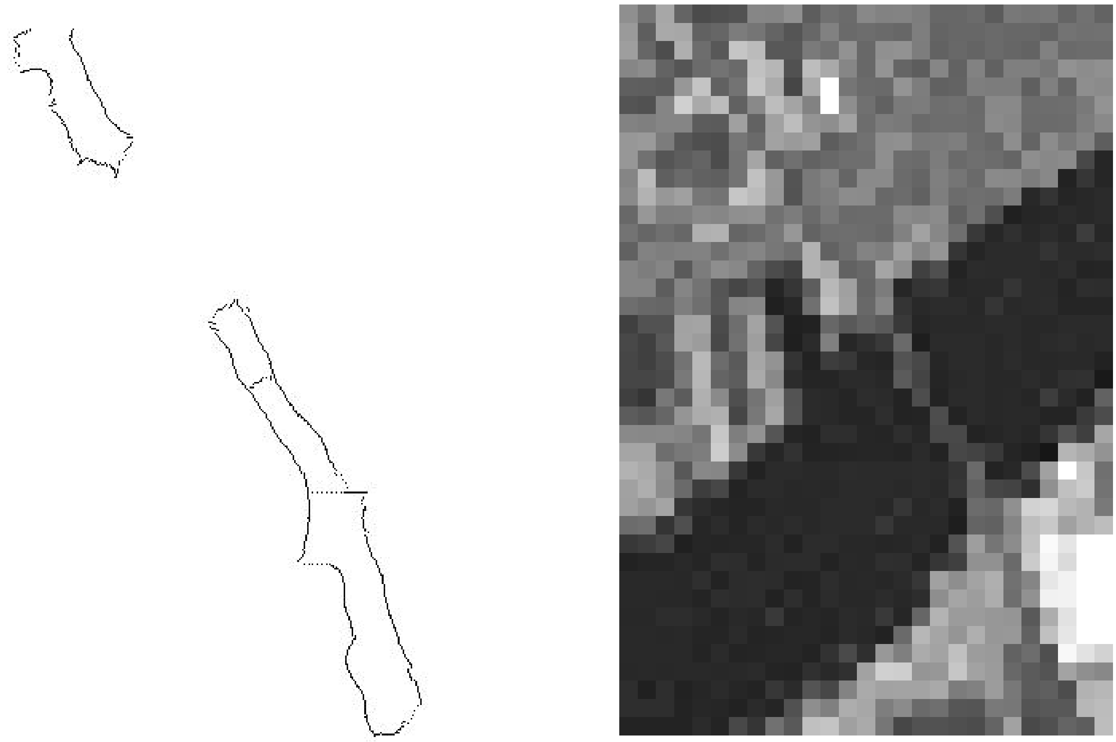
Despite the fact that the bridges & approaches were comprised of several regions, it was considered worth trying to generate sub-pixel maps of the said bridges etc. by selecting groups of associated region numbers to enter into the sub-pixel plotting program. To aid with such region number selection, an optional alternative ‘reconstruction’ facility exists which generates, instead of the ‘cartoon’ of recovered grey levels, a grey level map where the grey levels themselves code the local region number. For two of the three bridges this approach proved to be readily possible, but for the third - which is of ***very*** low contrast on the original image - it proved virtually impossible to select the appropriate group of region numbers to effect a successful profile map, despite the fact that the said bridge is visually obvious in the reconstruction. This is solely because the regions which make up this bridge in the reconstruction are somewhat broken & irregular - a matter which the human eye & brain together are very clever at ‘filling in’ as part of the ***perception*** processes. This must be taken as some indication of the limitations of successful feature extraction for extremely low contrast input features.

The sub-pixel profile plots for the two successful bridge & approach extractions are shown in Figs. 13 & 14, together with appropriate enlarged fragments of parts of the original image.



(a) (b)

Fig. 13. Sub-pixel profile plot of the first New York bridge & its approaches (a) and the approximate portion of the original image relevant (b).



(a) (b)

Fig. 14. Sub-pixel profile plot of a second New York bridge & approaches (a) and the approximate relevant portion of the original image (b).

In figure 14 it is not surprising that there is a break in the upper approach to the bridge, since this portion of the approach is very obscure in the original!

For the coastline features, processing of a 1:1 scale portion of the original image was resorted to, partially to limit the image size for processing, partially to reduce the likelihood of excessive region segmentation and partially to demonstrate the sub-pixel profile generation capabilities of the main processes for well separated, large regions. Fig. 15 shows the results of this exercise and a scaled up version of the appropriate input image portion for comparison. As with the image of the Thames in Fig. 6, the graphics resolution of the plotting program again limits the ***obvious*** sub-pixel vector representation.

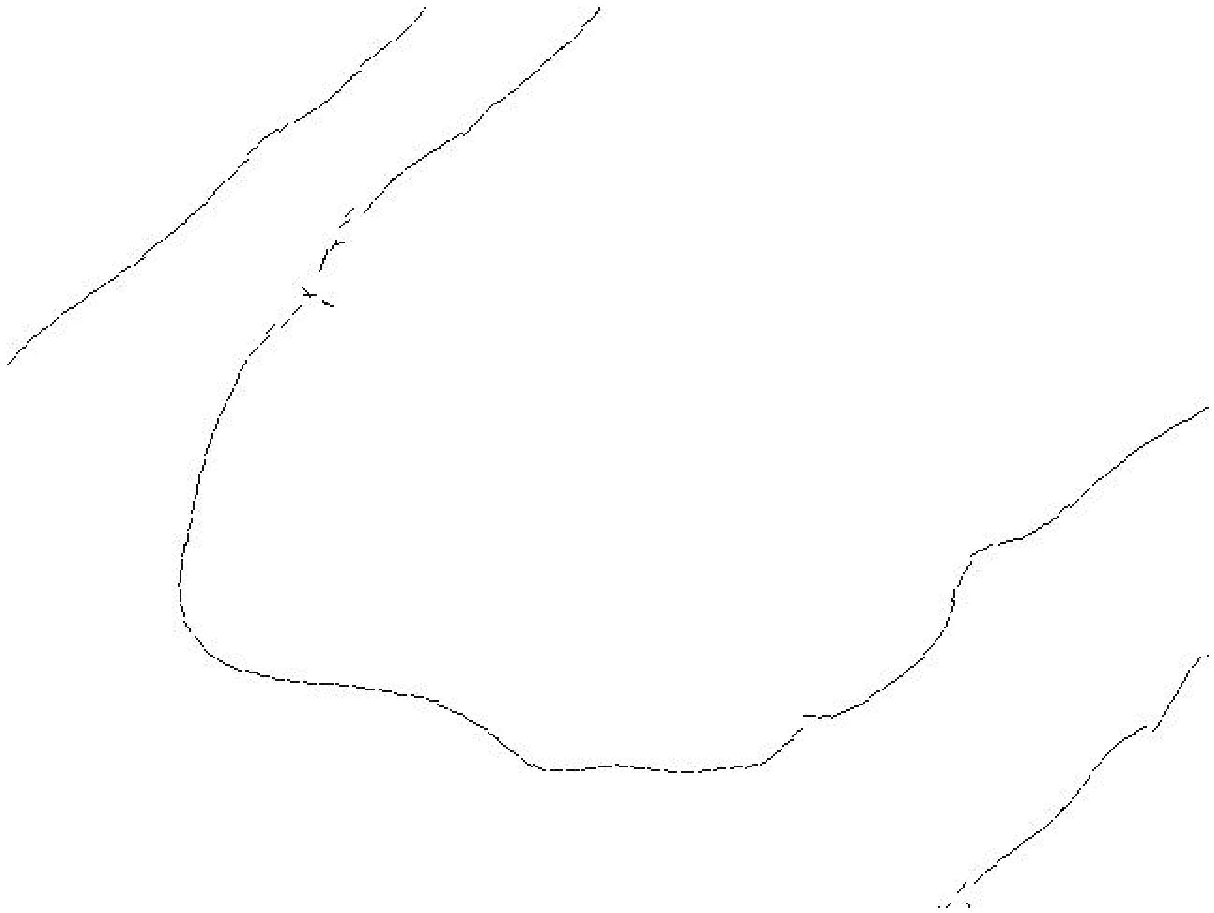


Fig. 15a. Sub-pixel plot of the coastline of part of figure 2 after processing at 1:1.

Fig. 15b. Enlargement of the relevant portion of Fig. 2 used to derive Fig. 15a.

Turning now to the other cropped image portion from the New York set of images - that shown in Fig. 3 - a X4 intelligently interpolated image was processed and a 'cartoon' reconstruction again produced from the tabulated results. This reconstruction is shown as Fig. 16, this being produced from 44130 profile points defining & segmenting 1841 regions. After the initial processing it was realised that not only should it be possible to extract the regions solely related to the runways & taxiways, but also there was potential for extracting just the local bridge and its approaches. These two separate extractions are presented as Figs. 17 & 18, together with an appropriate scaled up fragment of the original. For the bridge area

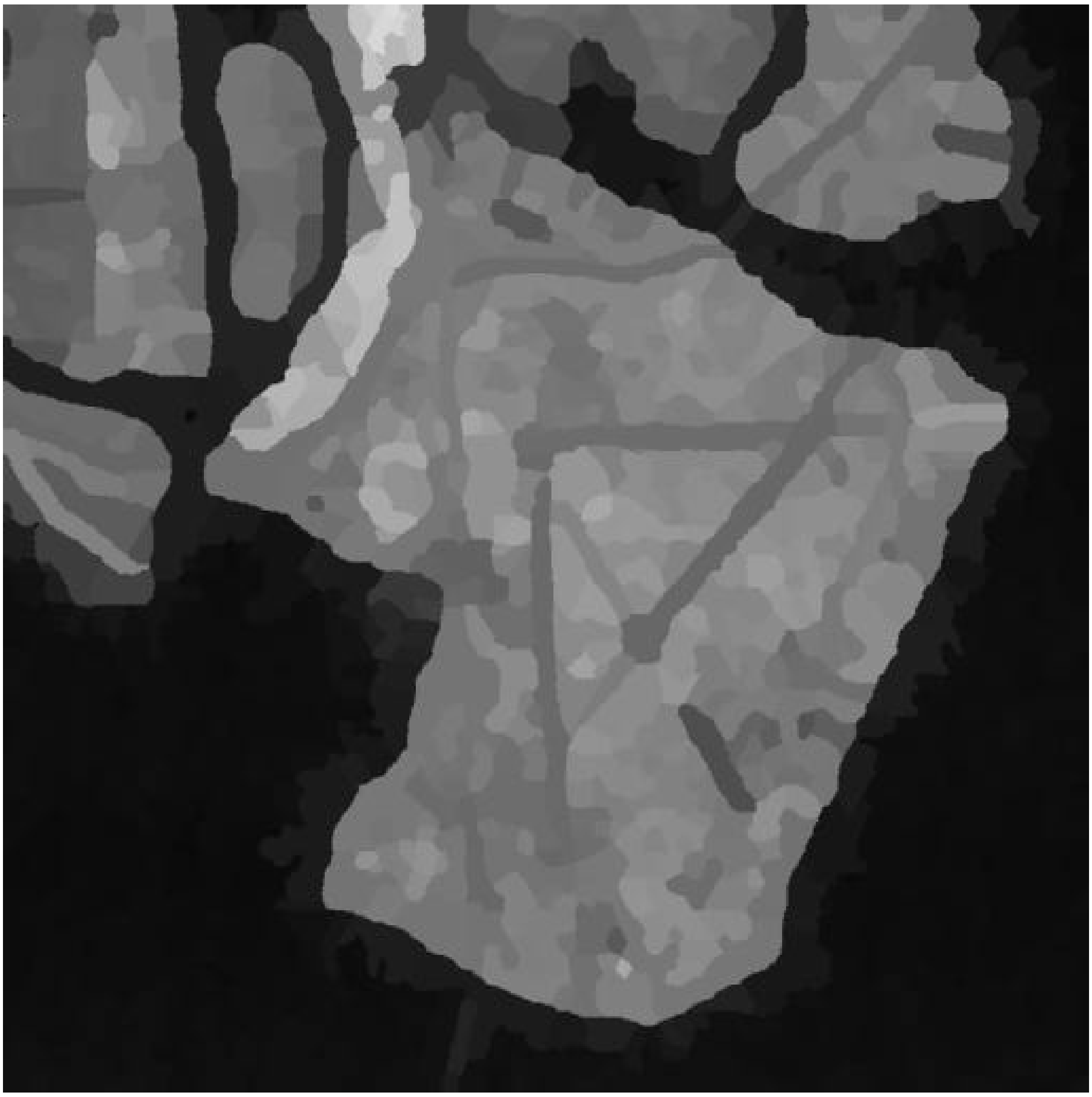


Fig. 16. 'Cartoon' reconstruction of Fig. 3 from profile-driven region segmentation.

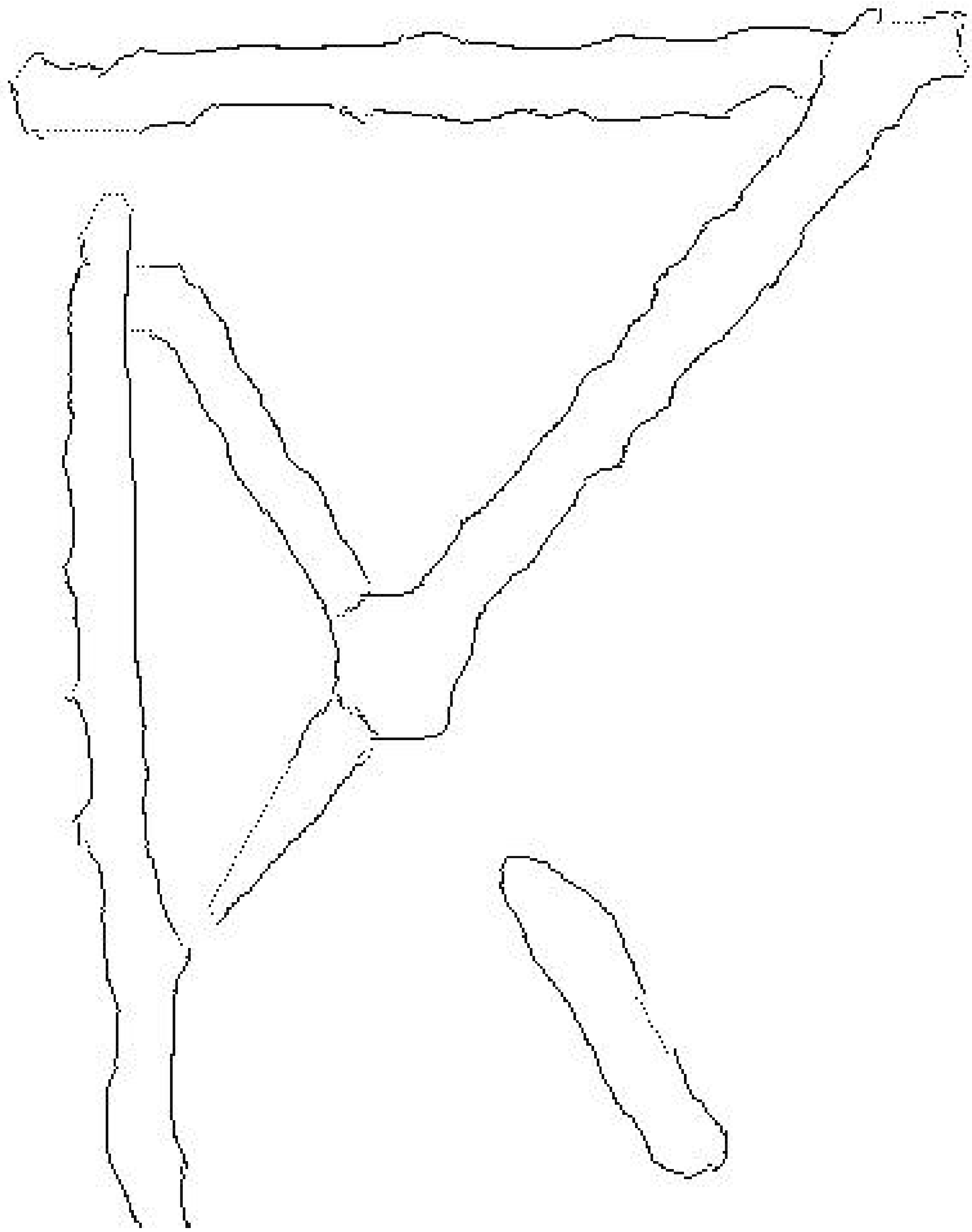


Fig. 17. Sub-pixel profile plot of runways & perimeter tracks for the airport in Fig.s. 3 &16.

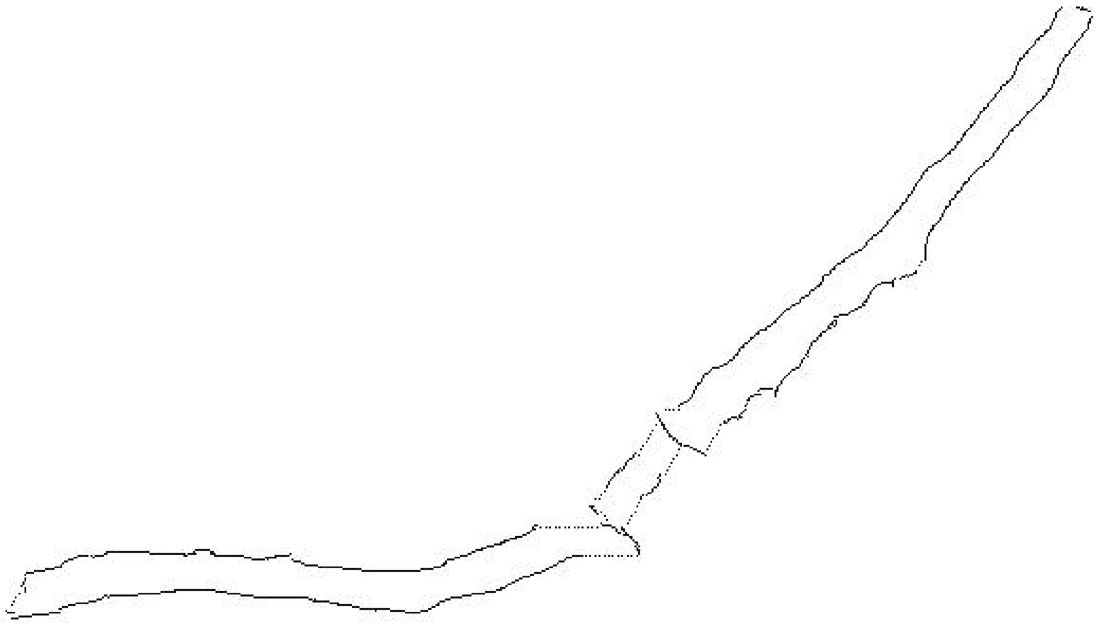


Fig. 18a. Sub-pixel profile plot of bridge & approaches from Fig. 3.

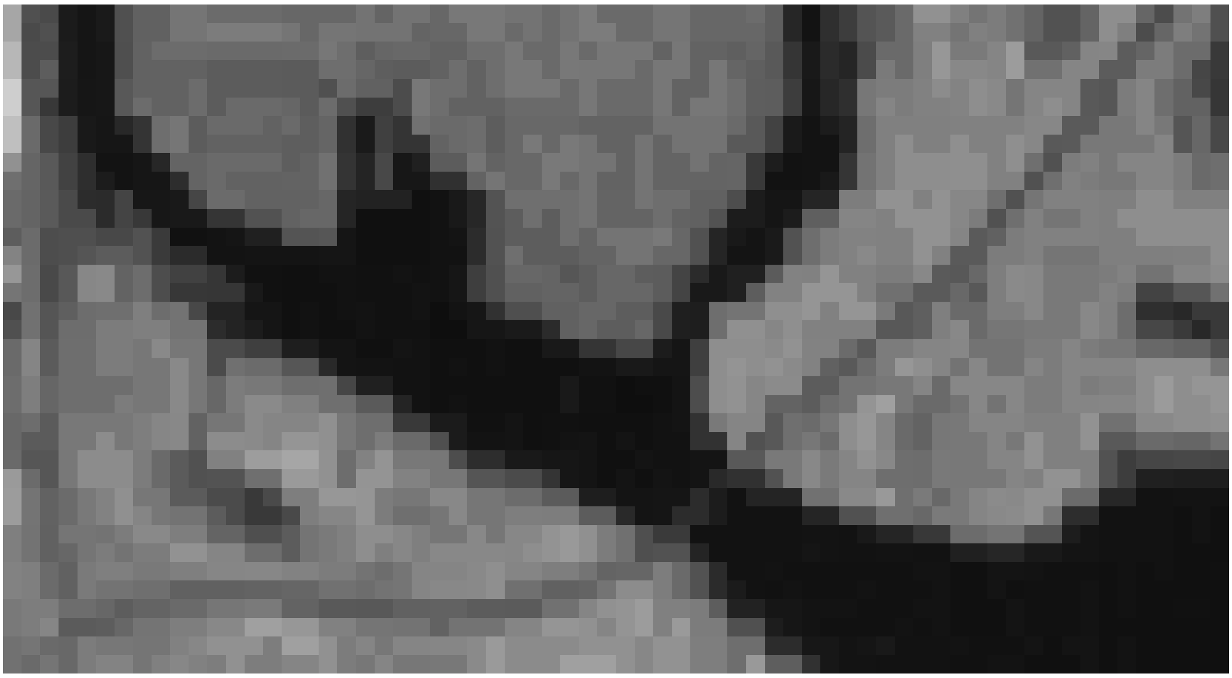


Fig. 18b. Enlargement of relevant portion of Fig. 3

Aukland analysis.
Again, if the whole original image had been used, the processing and the attendant data extracted would have been excessive. So again a portion of the image was cropped, this time covering what appeared to be an interesting portion of coastline including the Aukland central harbour area, plus a small island served by three bridges (see Fig. 4). As with the initial New York analysis, attention was focussed primarily on the bridges and the coastline.

An initial X4 intelligent interpolation was carried out before applying the main processing. The resulting (full colour) region segmentation data were then used to reconstruct a 'cartoon' of the original in full colour (Fig. 19). Again it is evident that the region segmentation - this time including the chroma data - is very good, this being produced from 46600 profile points defining & segmenting 1874 regions.

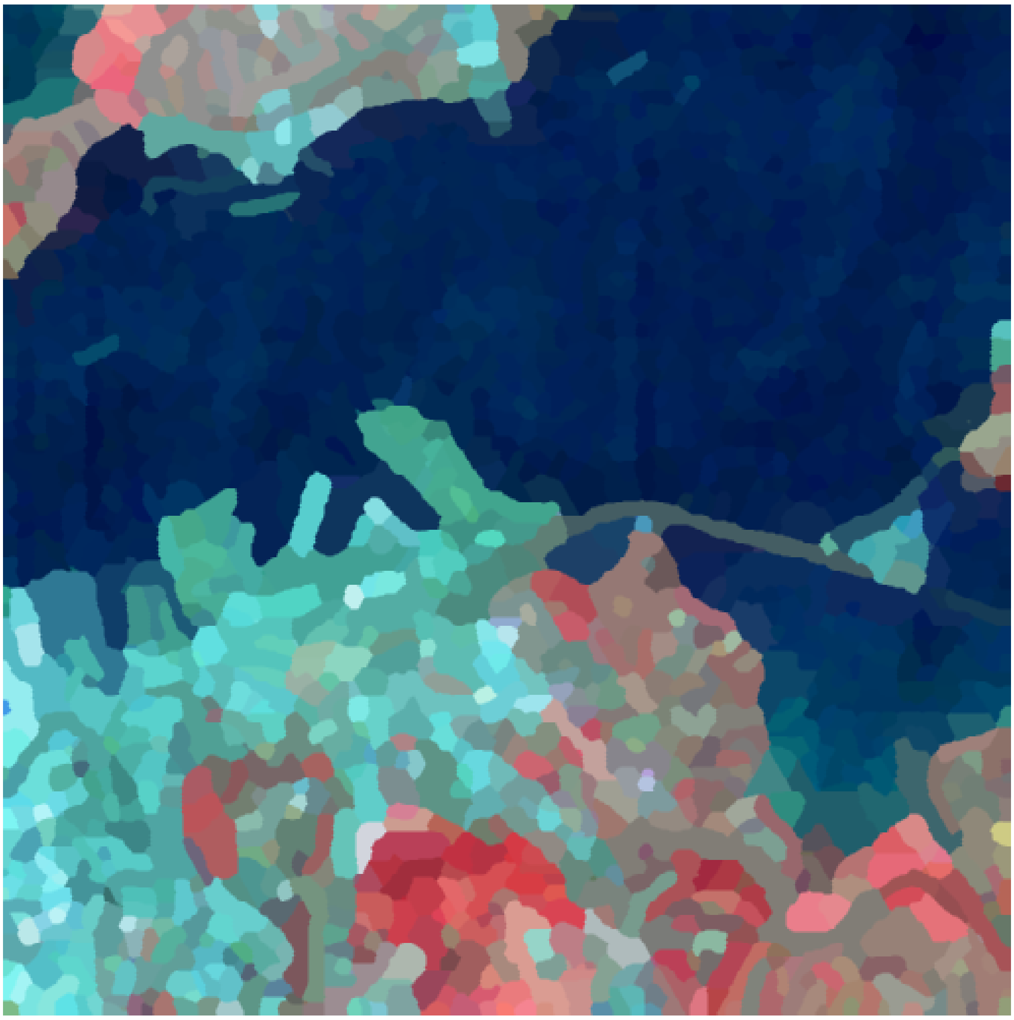


Fig. 19. 'Cartoon' full colour reconstruction of Fig. 4 from profile-driven region segmentation.

To assess further the capabilities for recovery of selected region data for bridge features etc. it was decided to attempt a sub-pixel profile plotting for the three bridge complex together with the small island in the middle of them. The results of this are shown in Fig. 20, together with an enlarged scale reproduction of the original fragment relevant to this plot. By and large the results are again considered to be good.

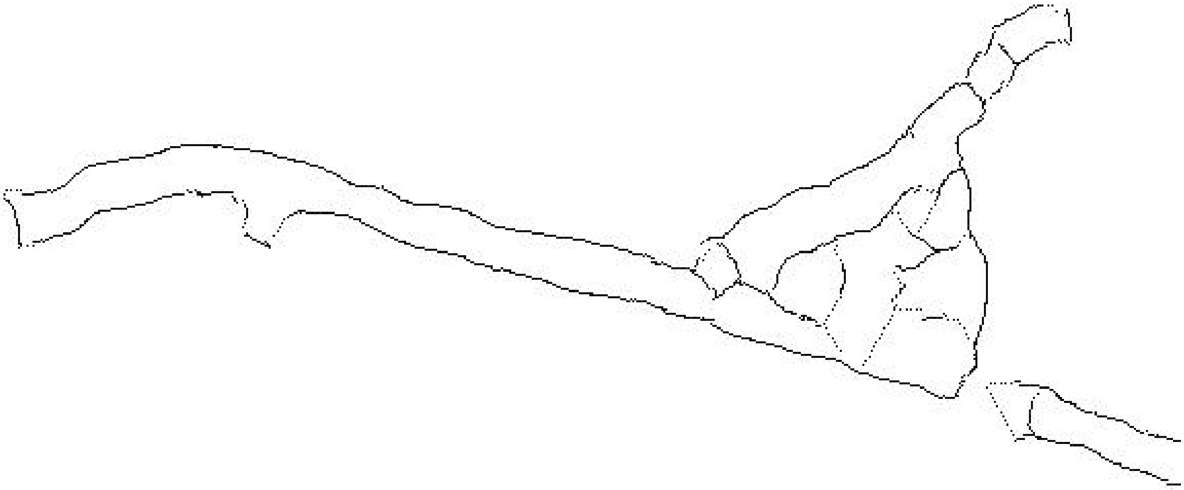


Fig. 20a. Sub-pixel profile plot of three bridege & island complex in Fig. 4.

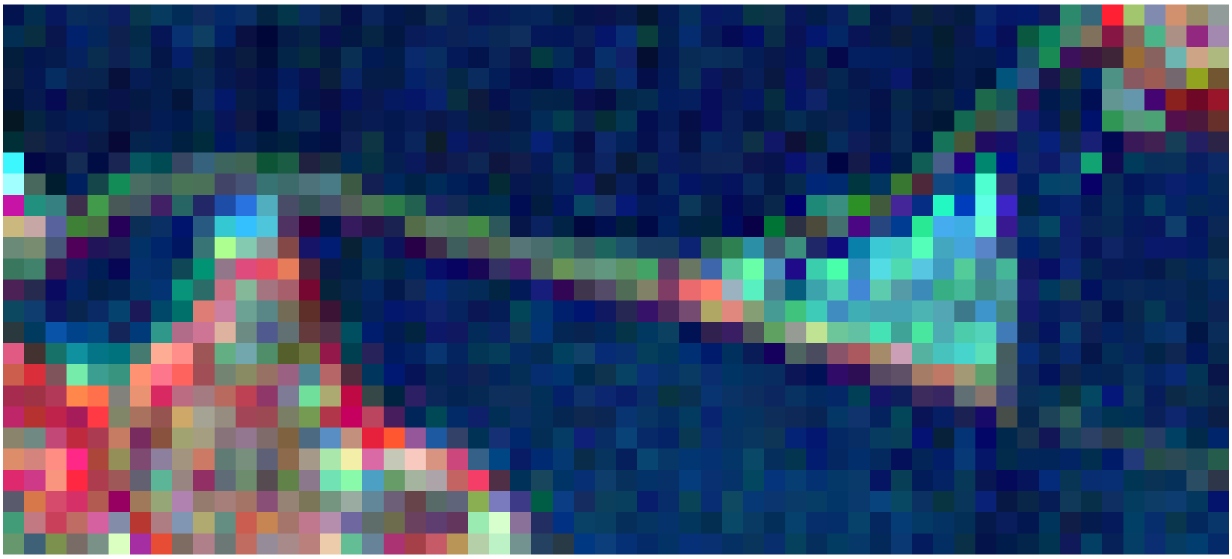


Fig 20b. Enlargement of the relevant portion of Fig. 4.

<u>Other thoughts</u>.
Whilst carrying out the study reported here, it has come to the author's notice that one potential application of feature extraction from satellite images which apparently has large potential is that of sensing the exact location and direction of such things as buried pipelines (given away by tell-tale ***lines*** of different ground surface or vegetation coloration). A very recently reported approach suggested for this (Ref. 2) essentially comprises the sensing of a series of points in the image roughly along the anticipated line of the pipe using what is essentially a tuned 1D second difference sensing filter, the statistically best line then being determined by regression analysis. Whilst such a technique is acknowledged to be essentially sound - and whilst claims made by the authors of the said paper that other techniques fail to work satisfactorily because of problems such as awkward influence of crossing edge features etc. are generally accepted - it immediately dawned on the ***present*** author that the processes incorporated within the currently being explored main sofware computations essentially

contain both first & second difference filters operating in parallel at ***six*** orientations (i.e. every 30 degrees of orientation). Whilst the present author was at pains throughout the development of the current software to ***suppress*** any ***line*** features so that they did not introduce false cues for region segmentation purposes, nevertheless in the ***initial*** phases of the processing both edges ***and*** lines are sensed! There is then an alternative option built into the command file structure for running the software which permits the processing to completion of the ***raw*** profile data instead of carefully edge-limited profile data, to the exclusion of region formation. Provided that some supplementary intelligence is available to define roughly the location of the anticipated line feature, it should then be readily possible to isolate the appropriate string of fragmentary profile points relating to the desired line feature, these being, as always, stored in terms of their specific sub-pixel position and local orientation ***for every profile point sensed***. This collected data set are then virtually complete in themselves, not only providing basic line direction simply by averaging the orientation values for the set of data points, but also permitting the determination of any local ***curvature*** in the line feature (this latter computation being carried out by a well tried, but currently 'mothballed' supplementary program). The ***only*** significant restriction on potential use of this optional facility is that the line feature needs to be of at least several (possible greater than 15 or 20) pixels long, since the only parts of such line features which are (of necessity) lost in the processing are the two or three pixels at each end of the line where it interacts with crossing features. This latter restriction is not a limitation of the present processes alone, but is rather a basic physical limitation related to the properties of difference signals in 2 dimensional imagery. This limitation, as with many other possibly unusual facts which are hinted in this report, is discussed at some length in Ref. 1.

I. Overington.
24th August, 2002.

References.

1. Overington I., 1992, 'Computer Vision - a unified, biologically-inspired approach ', Elsevier North Holland.
2. Petrou M. & Gracia I., 2002, 'Buried pipeline mapping from satellite images', Electronics & Communication Engineering Journal, Vol. 14, No. 4.