SEGMENTATION OF SPOT AND LANDSAT SATELLITE IMAGERY

Eija Parmes
Technical Research Center of Finland
Instrument laboratory
Espoo, Finland

Abstract

The paper describes a graph-based segmentation method for digital imagery. The principle is to point links from pixels to the direction of minimum edge intensity. The pixels of the same tree, ie. linked to the same root pixel, form a segment. The method has been validated to handle large multichannel Spot and Landsat satellite images. A method for evaluation and comparison of different segmentation results is proposed.

1 INTRODUCTION

The segmentation is here defined as the partition of the image area into a number of homogeneous units composed of contiguous lattice points.

The segmentation of satellite or aerial imagery can be used in many ways. The elimination of boundary or mixed pixels /Gru87/ ameliorates the separability of the land cover classes in the multidimensional histogram. The segmentation makes possible to calculate and use the geometrical, contextual and topological properties of the fields in the interpretation of the satellite image. After pixel by pixel classification of the satellite image, the result can be smoothened by defining the mode class of the segment to the final class of its pixels.

The segmentation result can also be used as such in the terrain verification and be a base print for remapping and updating of maps, and the correction of interpretation results, to minimize the amount of manual digitalization of area boundaries.

Most of the methods used in segmentation are grouped to region merging/splitting, clustering and edge detection methods (/Fu81/). The critical point of most of the methods is the use of a global threshold that determines the location of the boundary, which is especially sensitive to the threshold and the direction of approach. Narendra et al /Nar80/ have described a graph-based method which should position the boundary to a locally optimal position. The idea is to point a link from every pixel to the direction where the local measure of variation is minimum, i.e. in the ideal case to the direction of the center of a homogeneous region. After pointing a link from every pixel the pixels related by the links to the same tree are labeled to the same region. The method supposes, like most of the other methods that the pixel values of one region differ from those of the neighbouring regions : however, the advantage of this method is that there exists no lower limit for the absolute difference. The algorithm produces closed boundaries.

This graph-based segmentation method was choosed for further validation for Spot and Landsat satellite imagery. The following properties were added to the method described in /Nar80/.

The original images of different channels are calibrated to the same mean and standard deviation to set equal weight to the different channels. Different edge filters were modelled for different edge runs to quantify the optimal edge intensity over different edge runs and directions. For multichannel remote sensing imagery, the edge intensity values are computed for each channel separately and the maximum value among channels is choosed for pixel linking phase.

The method needs global processing in the sense that the homogeneous regions can intersect the blocks that can be processed in central memory. The implementation uses relative row and column coordinates for the linking of the pixels and block overlapping to allow large imagery processed without extra edges between the blocks.

2 THE SEGMENTATION METHOD

2.1 Calibration of channels

The purpose of this calibration is to weigth the responses of the different channels to be treated equally in the edge intensity calculation. The different channel images of the satellite data are calibrated to the same mean and standard deviation.

$$I_k(x,y) = \frac{Io_k(x,y) - mo_k}{so_k} * s + m \tag{1}$$

where Io = original image, k = channel number, mo, so = mean and standard deviation of the original channel, m, s = resulting mean and standard deviation.

The original mean and standard deviations are calculated from all the pixel values of the satellite image channel. In practice, this calibration gives more weigth to the visible channels in relation to the infrared channels than the original values.

2.2 Edge intensity calculation

The edge intensity for each pixel is calculated from the channel calibrated satellite image values. The edge intensity value G for each pixel is defined by convolution of 3 x 3 filters M for different edge directions.

$$G_k(x,y) = \max_{m=1}^{nm} \left| \sum_{i=-1}^{1} \sum_{j=-1}^{1} (M_m(i,j) * I_k(x+i, y+j)) \right| \tag{2}$$

$$G(x,y) = \max_{k=1}^{nk} [G_k(x,y)] \tag{3}$$

where nk = number of channels and nm = number of edge filters.

The value of the results is the maximum over the different filters and over the different spectral channels. The difference between water and forest is clearest in near infrared channel, so this channel result is implicitly chosen as the edge intensity value for water and forest boundaries. The edge intensity is thus defined to each pixel from the channel (spectral or textural) producing the maximal edge intensity value. (the implementation allows also the use of the sum over the channels in equation 3 : this seems to be more appropriate for forest segments).

2.3 Pixel linking

The linking of the pixels is done from the edge intensity image.

The linking of the pixels is based on the chaining of the pixels by the directed trees method described for instance in /Nar80/. The pixels are linked to the neighbouring pixel with minimum edge intensity value. The pixels linked to the same root pixel form a segment. The method needs a threshold e, which hinders the linking of the pixels towards a too great edge intensity. The increase of the threshold only reduces the amount of boundaries, or segments, without changing the position of the remaining boundaries, so the result is stable in this sense.

The whole image is processed three times in the pixel linking phase :

1. The defining of evident root pixels and evident edge pixels by computing the maximum difference between edge intensity values of the central pixel and its neighbouring pixels.

$$D(x,y) = \max_{i=-1,j=-1}^{1,1} [G(x,y) - G(x+i,y+j)] : (i,j) \neq (0,0) \qquad (4)$$

- if $D(x,y) \leq -e$: defined as root pixel, not to be linked to neighbouring pixel (all the edge intensity values around the central pixel exceed by e the central pixel's edge intensity)

- if $D(x,y) \geq e$: defined as edge pixel, linked to the neighbouring pixel with minimum edge intensity

2. The linking of the remainding pixels to the neighbouring pixel with minimum edge intensity and without creation of cycles.

3. The tracing of the chains to link the pixels directly to their root pixel.

2.4 Renumbering of the segments

The result of the previous step is a two channel image, where the first channel image is for the relative column coordinate of the pixel to its root pixel and the second channel image for the relative row coordinate of the pixel to its root pixel. The renumbering of the segments by unique numbers or labels is done by first numbering the root pixels. This can be done line by line. In the second step, the label for the other pixels is then found by their link to these root pixels.

2.5 Handling of large images

After the linking of the pixels to the neighboring pixels, the chain following to root pixels for images of 3000 to 3000 pixels would need at least 36 MB central memory, two bytes for input relative coordinates per pixel and two bytes for resulting relative coordinates to root pixel. With compression and bit operations the data amount could be lessened to 22.5 MB.

The use of the virtual memory system to allow such matrices was not choosed, because the direction of processing is not predictable, and the pixel chains can lead to any direction of the image matrice. The other solution is to divide the image into blocks to fit the central memory. With block processing, the proper handling of segments intersecting the block boundaries is problematic. With the following methods, the blocks can be processed independently without the need to transfer the data to and fro between the central memory and mass storage.

The problem was solved by using a data structure of two channel image for the chain following result (after step three in pixel linking). The first channel value is the pixel's relative image column coordinate to its root pixel and the second channel value is the pixel's relative image row coordinate to the root pixel. This together with using block overlap eliminates the discontinuities of segments otherwise resulting from the block processing. In practice 512 to 512 block size have been used with block overlap of 100 pixels (3 MB central memory allocation).

3 RESULTS

Figure 1 shows an example of a segmentation result of a Spot satellite image with 20 m spatial resolution. The segmentation used the channels 1, 2 and 3 with a threshold 15. The area is about 10 km x 7 km.

The cpu-times for instance for a 1200 x 1600 pixel image with 4 channels is about 1 hour (DEC Vaxstation 3100 with Disimp interface) with about 3 MB central memory allocated.

To minimize the effect of noise in segmentation the edge preserving smoothing was tested as a smoothening step before edge intensity calculation. This caused for instance small islands to merge into water areas. However,in forest areas the smoothening homogenizes the segments and reduces the number of extra boundaries.

The evaluation and comparison of different segmentation results by quantitative measures is difficult. A method to compress the amount of data to be verified or compared could be the filling of the segments by the mode or median value of the original pixel values in the segment and the visual evaluation of the difference image between the original image and the mode or median filtered segment image. Larger connected areas with great differences show segments where the segmentation method has failed. Highly textured areas would appear as salt and pepper. As an example, figure 2 shows the segment boundaries on the original image, with the corresponding difference image (below) with segment boundaries. The black areas in the difference images are such areas where the difference between the mode value in the segment and the original pixel value is ≥ 3. There exists 5-10 connected areas, whose segments should be checked (if one pixel wide areas are left unnoticed).

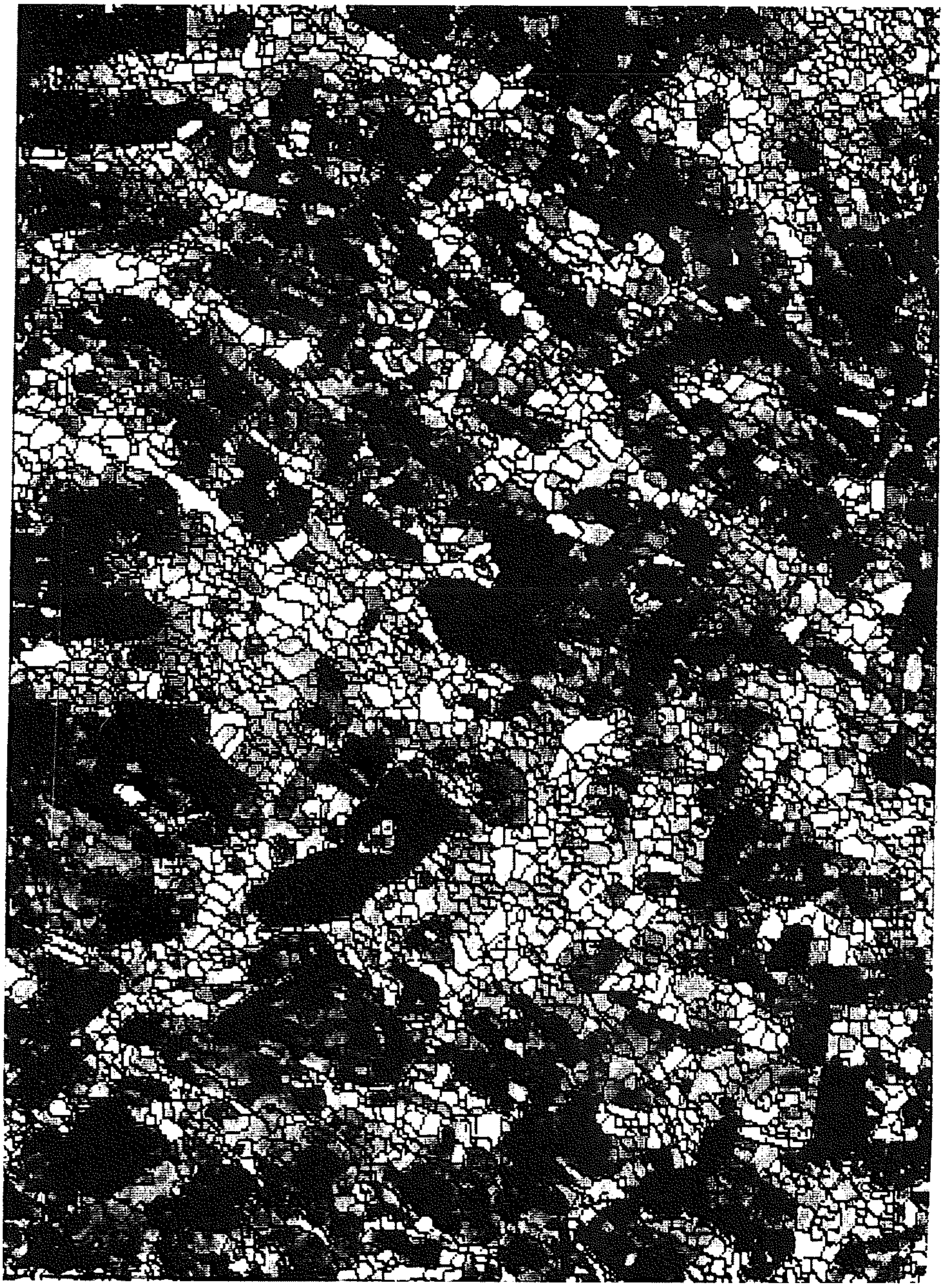

Figure 1: The segmentation result of a Spot satellite image. Scale 1 : 50 000.

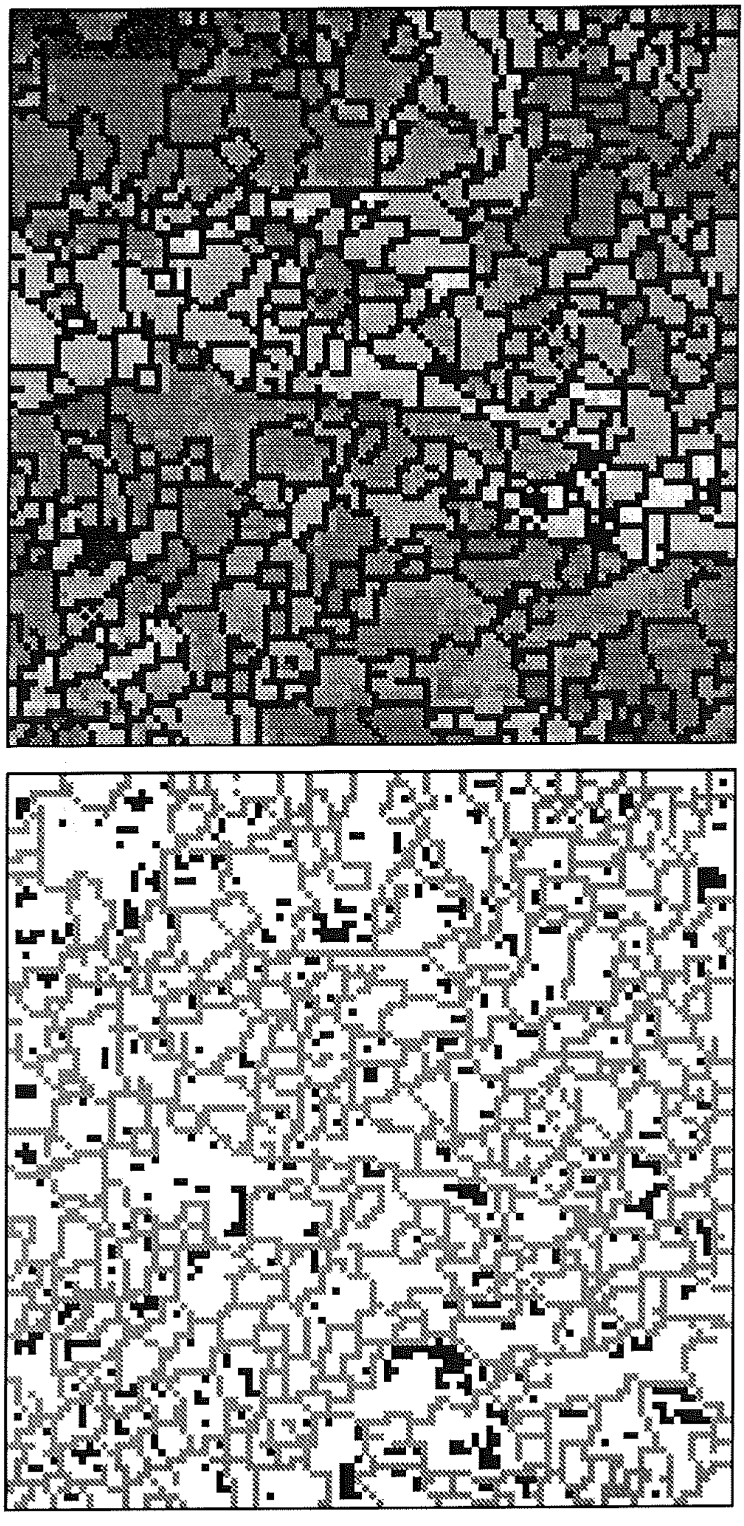

Figure 2: The segmentation result of a satellite image with the corresponding difference image.

4 DISCUSSION

The segmentation method has been applied to the analysis of watershed areas /Kui86/, delineation of forest stands /Tom87/, /Häm92/, the elimination of boundary pixels from multidimensional histogram clustering and the production of the land use map over Finland in The National Board of Survey.

The following aspects should be further evaluated :

The segmentation method has been used only for the original radiometric values of the satellite image. It works well in the delineation of quite homogeneous areas (waters, agricultural areas, open areas etc). For textured areas like forest areas and transitional areas with changing vegetation, the inclusion of texture channels as input to the segmentation could ameliorate the results.

The method does not locate the boundaries with subpixel accuracy. It gives, however, a good basis for the interpolation of subpixel boundaries.

The insertion of new edge filters covering a larger amount of different edge runs is a way to further optimize the results.

References

[Fu81] K.S. Fu and J.K. Mui : A survey on image segmentation. Pattern Recognition, vol. 13, nr. 1, 1981, pp. 1-16.

[Gru87] J. Grunblatt : An MTF analysis of Landsat classification error at field boundaries. Photogrammetric engineering and remote sensing, vol. 53, nr. 6, 1987, pp. 639-643.

[Häm92] T. Häme : Spectral interpretation of changes in forest using satellite scanner images. Acta Forestalia Fennica 222, 1991.

[Kui86] R. Kuittinen, E. Parmes, Y. Rauste, Y. Sucksdorff: Watershed area interpretation method using satellite imagery. Method description. Technical Research Center of Finland,Section for Remote sensing. Espoo, 1986. In Finnish.

[Nar80] P.M. Narendra, M. Goldberg : Image segmentation with directed trees. IEEE Transactions on pattern analysis and machine intelligence, vol. PAMI-2, nr. 2, march 1980.

[Tom87] E. Tomppo : An application of a segmentation method to the forest stand delineation and estimation of stand variates from satellite images. Proceedings of the 5th Scandinavian conference on image analysis of IAPR, Stockholm 1987.