

KNOWLEDGE-BASED SEGMENTATION FOR REMOTE-SENSING

A.M. Taylor*, D.G. Corr*, A. Cross*, D.C. Hogg**, D.H. Lawrence**,
D.C. Mason*, M. Petrou*

* Systems Designers Scientific, Pembroke House, Pembroke Broadway, Camberley,
GU15 3XD.

+ Natural Environment Research Council Unit for Thematic
Information Systems, Reading University, PO Box 227, RG6 2AB.

** University of Sussex, School of Social Sciences, Falmer, Brighton,
BN1 9QN.

ABSTRACT

In order to cope with the large volume of remotely-sensed data available now and expected in the future, efficient automatic processing techniques are required. A particular problem in automatic interpretation of this data is the identification of relevant connected regions in the image, i.e. segmentation. This can generally only be achieved to a required degree of accuracy if performed manually. This paper describes the current implementation of a system for automatic segmentation of multi-temporal remotely-sensed images which exploits prior knowledge to isolate the regions of interest. The system is directed principally towards the applications of crop and environmental monitoring.

I INTRODUCTION

The volume of quantitative data from remote-sensing platforms has increased considerably in the last decade and will continue to increase in the future with the planned launch of Europe's ERS-1 satellite and the polar platforms in NASA's Earth Observation System. The problems to which these images are applied have also increased in scope and magnitude with time. Present-day problems using remotely-sensed images cover applications as widely varying as urban land use studies and crop inventories (see, for example, Landgrebe 1981 for typical applications).

There is a continuing need for efficient, machine-implemented analyses of the data. Unfortunately, the image generation techniques have far outstripped the image analysis techniques, even though the latter have advanced significantly since the launch of the first LANDSAT.

The extraction of information from remotely-sensed data may be thought of as a two stage process of segmentation and classification. The segmentation stage, which partitions the image data into connected homogenous regions, is traditionally performed by a photointerpreter.

A problem is encountered in the definition of the required homogeneous regions. The optimum regions are governed by the application, i.e. what the end-user perceives as a classifiable object on the ground. Therefore, in this project we have necessarily limited investigations to a subset of possible applications: those of renewable resources, e.g. crop monitoring, and environmental monitoring, e.g. the limits of natural fenland. Within this category of applications, segmentation to some arbitrary level is attempted.

*This work is supported, in part, within the UK Alvey Man Machine Interface developments.

Automatic segmentation techniques, whether region or edge based, are generally inadequate or unreliable, and result in an ambiguously segmented image. Classification errors result as a direct consequence of this. The segmentation errors may be due to practical issues, for example the sensing conditions, but the principal source of error is the data driven aspect of techniques. There is generally insufficient information in the image itself to completely discriminate between regions. Even if the information is in the image, the exact properties and values to use are not known a priori. The work described in this paper is directed towards reliable automatic segmentation techniques which can be adapted for a range of remotely-sensed images and a broad category of applications. In order to achieve this, recourse to a priori knowledge about the scene must be made. This knowledge is normally only available to an experienced photointerpreter.

The project is divided into two major phases: the study of candidate systems for knowledge based segmentation, and the implementation and evaluation of a final system based on the results of the first phase. The next section details the types of information which may be exploited in order to improve segmentation. An overview of the basic system used for evaluation in phase 1 in terms of its features and the knowledge it exploits, is given in Section III. Section IV gives an evaluation of the system strategies developed in the first phase and details the data used, the tests performed and the results obtained. The design for the final system is summarized in Section V, by detailing the changes and additions made to the basic system. The final section presents a summary of the system.

II EXPLOITING KNOWLEDGE

For more accurate segmentation, the information in the image, given by the spectral, spatial and temporal variations of the pixels, can be supplemented by a priori knowledge. This knowledge, which imposes constraints on the regions to be expected in the data, is external to the image.

The a priori knowledge can take the form of either additional data sets or domain data. The additional data sets may be maps or other sensor data of the same scene. However some aspects may differ from those of the original data, for example the data representation or the spatial resolution. The domain data, using the expertise of an analyst, is generally in the form of conditions on the regions in the data, for example, fields generally have straight boundaries.

The available knowledge may act on the properties of regions at the segmentation level or on the identities of regions at the classification level. The information on the properties may, in turn be two- or three-dimensional.

Standard region and edge segmentation techniques, such as clustering (e.g. Townshend and Justice, 1980, Seddon and Hunt 1985), split-and-merge (e.g. Cross and Mason 1985), edge detection and linking (e.g. Nevatia and Babu 1979), generally only use limited spectral and spatial information from the image. Techniques which exploit other forms of image information include relaxation, which employs local relationships between pixels (see, for example, Peleg 1984), and signature extension methods, which use the temporal signature of objects to aid classification (see, for example, Henderson 1976).

The exploitation of information external to the image requires more sophisticated techniques for data integration and imposition of constraints. A review of such techniques can be found in Taylor et al 1986. In particular, Image Understanding Systems (Matsuyama 1986) may be built to improve aspects of the interpretation process. Ideally, the knowledge would be used as a model of the state on the ground. However, the complex nature of the data means the knowledge is rarely available in the right form or is not accurate enough for it to be used this way. More often, the knowledge is used as rules in a rule-based system to impose constraints on the regions. Depending on the nature of the information, most systems combine the modelling and the rule-based approaches. Examples are the systems of Nazif and Levine (1984) for segmentation of aerial imagery, Goldberg et al (1985) for updating forestry maps and McKeown et al (1985) for the interpretation of airport scenes in aircraft data.

III KNOWLEDGE-BASED SEGMENTATION SYSTEM

This section describes the functionality of the basic knowledge-based system which was developed in phase 1 for segmentation of multi-temporal, high altitude aircraft and satellite data. The emphasis is on general segmentation techniques, as in the Nazif and Levine (1984) system, rather than on interpretation of more specific information, as in Goldberg et al (1985) and McKeown et al (1985). However, compared to the system of Nazif and Levine (1984), we include a modelling, or goal directed, approach to supplement the data driven approach, together with the use of temporal information.

The system uses cartographic map data as a crude model of the situation on the ground, together with domain data in the form of rules, to supplement the information in a multi-temporal aircraft data set.

The basic system consists of a set of modules for integrating the available information and then refining an initial segmentation of the data, based on that information. For each image in the time sequence, the two major stages of processing are:

a. Initial segmentation.

The current image is first segmented using a Split and Merge region segmentation and a Sobel edge detection procedure. The lines in the map are then imposed onto the region segmentation and become in themselves thin regions or parts of region boundaries. Strongly parallel edges are joined and also imposed as thin regions. Relationships are then specified between the image regions and edges, the line information from the map, and the regions from the previously segmented image in the time sequence, if available.

b. The refinement of the initial segmentation.

Based on the evidence from the first stage and the domain constraints imposed on the region properties, the initial segmentation is updated. The domain data represents the analysts expertise and is embodied as production rules in a rule base (see below). By focussing attention onto ambiguities in the segmentation, and accessing the rules, decisions can be made.

Once each time sequence image has been processed, the output is the current best segmentation, as both a two-dimensional array and a list of the properties of each segment.

The rules in the rule base were designed to use limited knowledge about the forms and types of regions to be expected in the data. This knowledge stems from perceptual and common-sense constraints on the following region properties:

- the variance of the region
- the number of image edge and map line pixels interior to the region
- the number of image edge and map line pixels on the region boundary
- the contrast between the region and its neighbours
- the consistency with the previous time image segmentation.

Additional simple domain constraints were placed on the straightness of a regions boundary in agricultural areas and the size of a region. More complex domain constraints, using a photointerpreter's knowledge, were reserved for the final system. The knowledge used is all two-dimensional: no elevation data was included at this stage.

The decision strategy at the refinement stage consists of four major steps:

- assign a measure of confidence to each region
- weigh the evidence for each region in the rule base and make a decision
- modify any regions, if the evidence strongly supports a change
- update the overall segmentation in the light of any modifications.

The confidence measure takes a symbolic value, e.g. high, medium, low. It is chosen based either on the region properties, for the first time image, or on the degree of match with previous regions, for subsequent images. Modification of regions takes the form of either merging, splitting or making an adjustment of a boundary.

The system has been implemented in POP-11, a powerful general purpose programming language, integrated into the POPLOG environment. POPLOG has been developed at Sussex University and is marketed commercially by Systems Designers.

IV EVALUATION

A. Data.

The basic system described in Section III has been tested on multi-temporal image data from the NERC simulated Thematic Mapper aircraft data. Each image has 11 spectral channels and 10m spatial resolution.



Figure 1 Aircraft Image Data of Barton Broads, Spring 1986. The figure shows a first principal component, 256 x 256 pixel image at 10m resolution.

The data were transformed using a principal component transform to decorrelate the spectral information. Two areas were selected for study, both with little elevation. Figure 1 shows a 256 x 256 pixel, first principal component image of one of the areas, that over Barton Broads, Norfolk. The second area selected is over Blewbury, near Reading. The areas have a three and a two image time sequence, respectively.

Figure 2 shows a 64 x 64 pixel section of figure 1, taken from the middle left. Figure 2 consists of a combination of unstructured, natural fenland, in the top-right corner, and structured fields with well-defined boundaries.

Figure 3 shows an area of digitised Ordnance Survey 1:10000 map data which corresponds to the image of figure 2. The map band shown is the first (linear features) of three band digitised data. The further two bands (polygonal features and contours)

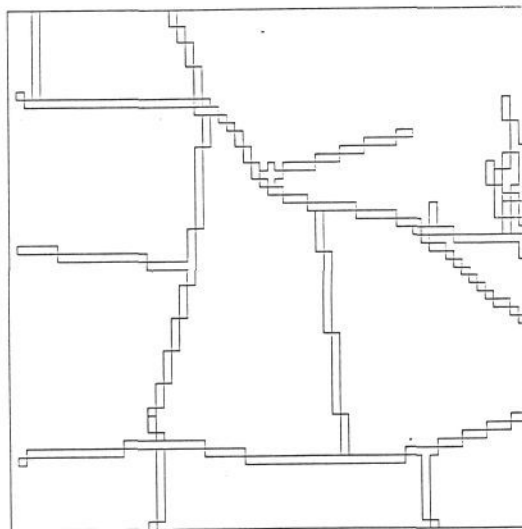


Figure 3 Digital Ordnance Survey 1:10000 Map Data corresponding to the area of figure 2.

were not used in this first phase. Each of the lines in figure 3 is 1 pixel wide. The two lines running from the left to right across the map are roads; the incomplete lines running into the fenland are footpaths, while the remainder are field boundaries.

A two band hand segmentation has been produced for each image, which is used as a reference for validation of the segmentation. The two bands separate the scene into areal and linear segments. Figure 4 shows the hand segmentation into areal regions corresponding to figure 2. The hand segmentations have been drawn up by a photointerpreter using information on the identity of regions from the image, the map and ground truth.

The images, the maps and the hand segmentations have each been registered to the British National Grid to within a 2 pixel accuracy.

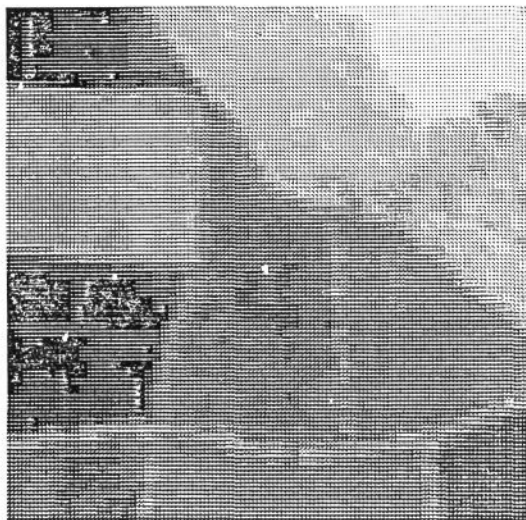


Figure 2 64 x 64 pixel sub image of Figure 1.

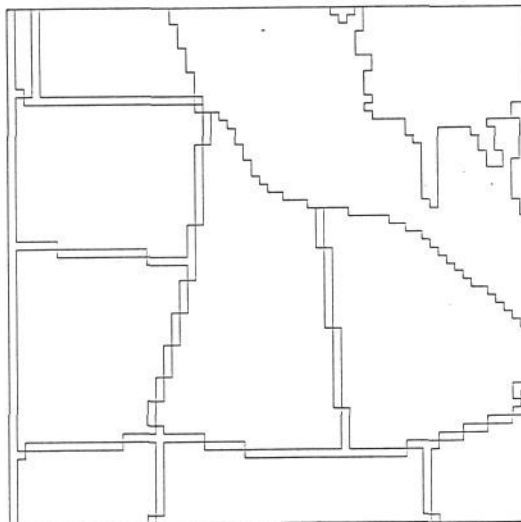


Figure 4 Hand Segmentation of the scene into regions

B. Tests.

A number of strategies were devised for manipulating the system modules in order to investigate the importance of the knowledge used. Each strategy used a combination of image region, image edge, map line and temporal information. More detail can be found in Taylor et al (1987).

Each of the two time sequences of images was segmented using each of the strategies. Only the first principal component of each image was used. In addition, a standard segmentation by the AMOEBA spectral-spatial clustering package (Drake 1986) was performed on the last image in each time sequence, for comparison. The resultant segmentations were validated against the hand segmentation. The validation gives two performance parameters; the overmerge (OM) and undermerge (UM) parameters, which are the percentage of overmerged (undersegmented) and undermerged (oversegmented) pixels respectively (Levine and Nazif 1982).

As an additional test, each segmentation was classified using a region-based maximum-likelihood classifier on the first four principal components of the last image in each sequence. The classifications, together with a per-pixel maximum-likelihood classification of the last image in each time sequence, were compared against a hand classification of the data. The percentage of correctly classified pixels was calculated for each classification. The hand segmentation was also classified to give an indication of the amount of classification accuracy achievable.

C. Results.

1. Segmentation

Table 1 gives the segmentation parameter values for a selection of the segmentations of the scene shown in figure 2. If we can assume that a classification can cope more readily with oversegmented regions, by classifying them as one region, then the overmerge parameter becomes most significant.

The results for the standard segmentation by clustering (figure 5) and for the knowledge-based segmentation using only image region information are comparable. The evidence in the rule-base for splitting a region is dependent only on the line and edge information. Consequently, using only region information, we have, in effect, a split and merge segmentation (figure 6) with a refinement stage consisting of rules for region merging based on contrast. An additional constraint is the minimum region size. This results in the undermerging error being less than that for clustering. Notable deficiencies in both these segmentation results were found in the ill detection of high variance regions and the inexact positioning of boundaries.

Table 1: Segmentation Performance Parameters

	UM	OM
Clustering	0.191	0.154
Image Regions	0.124	0.15
Image Regions and Edges	0.202	0.081
Image Regions and Edges and Map Lines	0.221	0.065
One time image	0.142	0.214

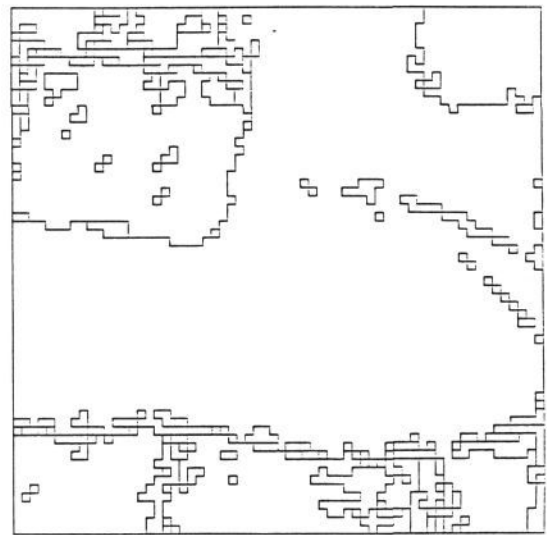


Figure 5 Segmentation of the scene using a standard clustering procedure.

The edge segmentation of the image is shown in figure 7. Incorporating this information into the segmentation reduces the overmerge parameter significantly: splitting of regions is allowed. However, an increase in the undermerge parameter is also evident due to spurious edges either splitting regions incorrectly or not merging regions which should be merged.

The line information again improves the overmerging error, due to there being more support for splitting regions. In addition, thin regions and boundaries are better defined. The final segmentation of the scene using all information can be seen in figure 8.

The undermerging error is increased with line information incorporated. However the majority of this increase reflects the imposition of thin regions onto the segmentation which are not included in the region hand segmentation. Other undermerging errors

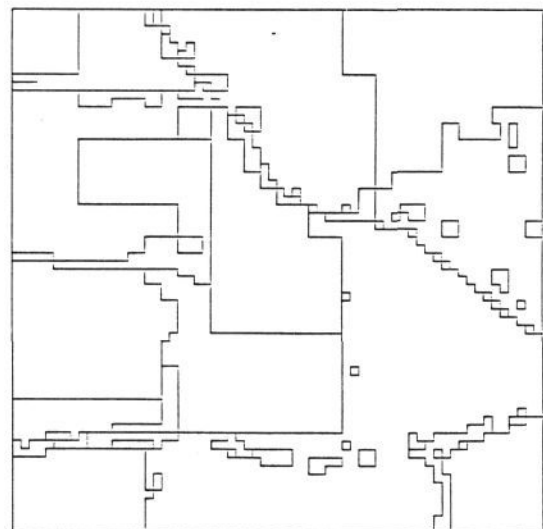


Figure 6 Split and Merge Segmentation of the scene.

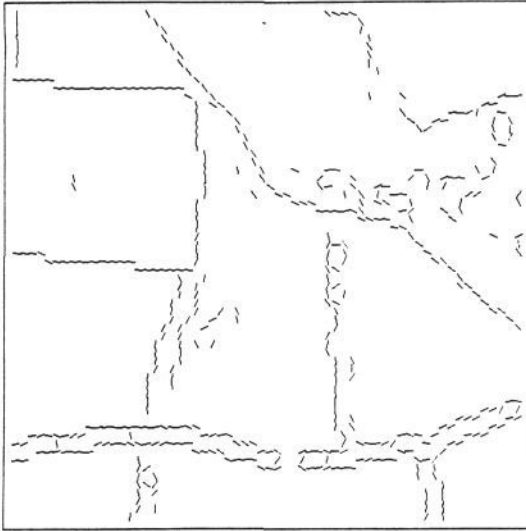


Figure 7 Edge Segmentation of the Scene.

result from doubly defined boundaries due to a lack of complete registration between the image and map data. As an example, the small region in the top left corner of the image which had been segmented into 2 regions using the edge information, is segmented into 3 regions by incorporation of the misregistered line information.

The last entry in table 1 shows the performance parameters for a segmentation of only the last image in the time sequence of the scene in figure 2, using image region, image edge and map line information. This serves to illustrate the importance of the temporal information for improved segmentation.

In summary, the knowledge-based segmentation gives an improved segmentation over traditional segmentation by clustering. Problems occur of undermerging in high variance areas, such as textured regions and boundaries, and overmerging where the spectral contrast in only the one principal component is insufficient for boundary discrimination. The

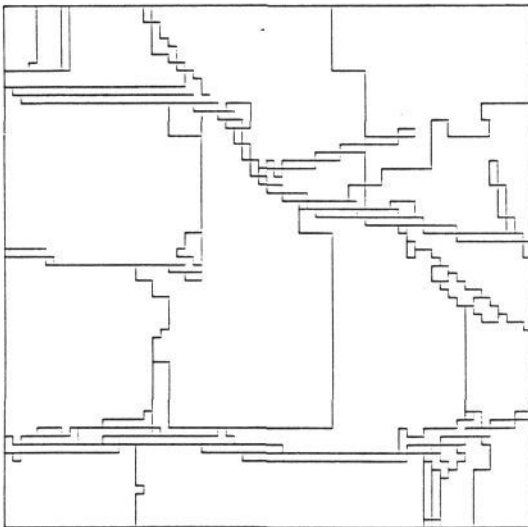


Figure 8 Result of the Knowledge-based Segmentation, using all information.

Table 2: Percentage of Correct Classification

Per-pixel	86.2
Hand Segmentation	96.5
Clustering	81.2
Image Regions	80.1
Image Regions and Edges	81.1
Image Regions and Edges and Map Lines	81.4

latter problem is alleviated somewhat when using the image edge data and the map line information; however, additional undermerging problems are caused by spurious edges and misregistration. The undermerge performance parameter suffered somewhat by not having the linear regions explicit in the hand segmentation into regions used for validation.

2. Classification

The percentage of correctly classified pixels for each of the classifications using the first four principal components are given in table 2.

The poorer performance of the per-pixel classifier compared to the classification of the hand segmentation serves to indicate the benefit of segmenting the image before classification, provided the segmentation is accurate. The knowledge-based segmentation is comparable to the clustering in terms of classification accuracy, although both give inferior classifications to the per-pixel classifier.

The majority of the classification errors result directly from the undersegmentation, or overmerging, of the segmentation process. This in turn stems directly from the lack of spectral discrimination in the one principal component used for segmentation. The large segment in the top right corner of figure 8 comprises both a lake area and a wood area with no global or local contrast in that principal component. This area caused the majority of classification errors for that image sequence. These errors do not occur in either the hand segmentation or the per-pixel classifier because, in the first the regions are correctly aligned with the hand classification, and in the second four principal components were used throughout.

It is noted that some of the classification error when using the map information, results from the map lines not being included in the hand classification. In particular, one of the roads in the segmentation was classified correctly, but was not evident in the hand classification.

V FINAL SYSTEM DESIGN

A. Major Differences to the Phase 1 System.

In the light of the evaluation of the initial system, which demonstrates the importance of each type of information used, a system has been designed which uses the information in a more efficient manner, while also using additional knowledge. A flow diagram for this final system is shown in Figure 9. The system is currently being implemented.

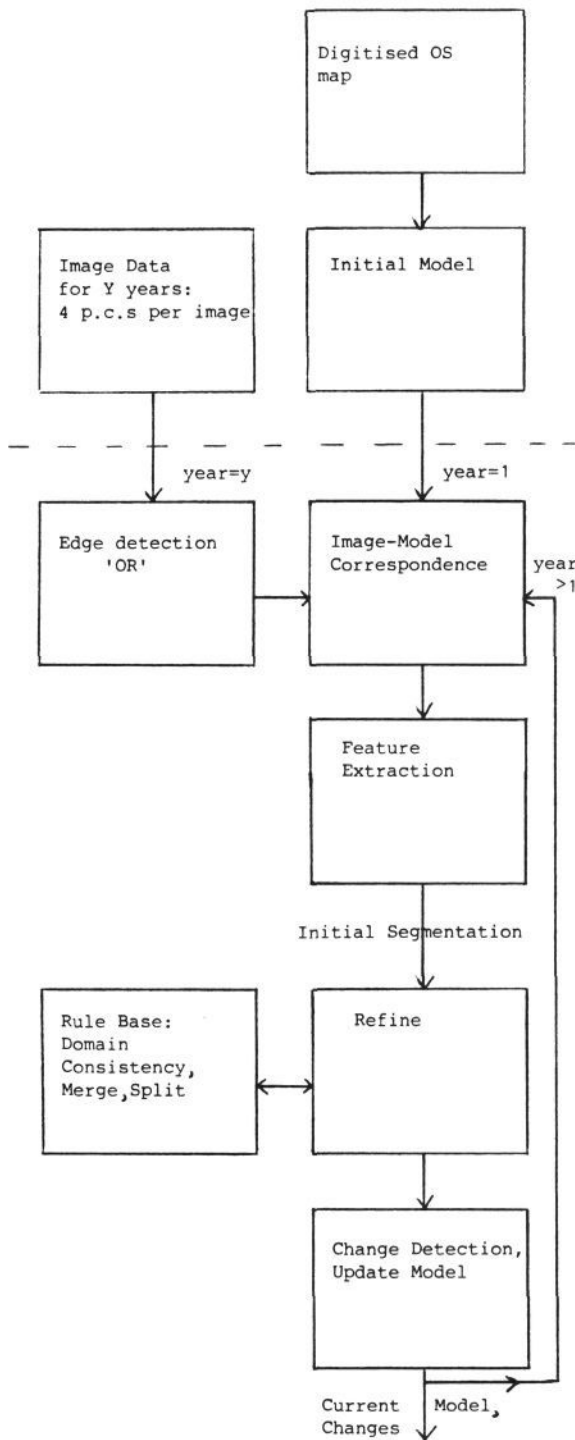


Figure 9 Flow Chart for the final system.

Two major areas were isolated where the information was not being used to its full potential. These areas are concerned with the image edge and the map line data, which were both supplementing the region segmentation, even though they are visibly superior. The edge data is now used instead of the region segmentation to initialize the current segmentation. The map is now used explicitly to define an initial model of the situation on the ground, which may or may not be changed depending on the new information in the remotely-sensed data.

An additional system function was required to cater for high variance or textured regions. Texture is considered by isolating possible texture edges from possible region boundary edges. Oversegmentation is not allowed in texture edge areas unless there is more than one distinct textural distribution.

Additional information is included in the system in three forms:

a) The first four principal components are used for each input image, rather than one. This guarantees the use of most of the significant spectral information.

b) The further two bands of the digitised map data are used. Band 2, containing the map polygons is used with the first band of map lines to initialise the model. The third band of map contours is referenced by the rules to resolve class consistency, e.g. a water area will have zero slope.

c) The rule base is supplemented by domain knowledge about the forms and properties of the expected classes of objects in the scene.

B. System Flow.

The system takes as input a time sequence of remotely-sensed image data and an optional digitised OS map. The image data is assumed to be arranged into individual years images, in order to cope with seasonal variations in object types. For example different crop fields will be indistinct from each other at certain times of the year, but will differ at other times in the growing season. The input image and map data is assumed to be geocoded. The major deviation of the input data from this assumption is in the mis-registration. If the mis-registration error is greater than that expected from random fluctuations due to sensing conditions, orientation etc., then the error will be output as a change on the ground.

If map data is available, then the first two bands, consisting of linear and areal regions, respectively, are combined to form an initial model of the situation on the ground. The region classes are retained in the model.

The remainder of the system, shown below the dotted line in Figure 9, is repeated for each year in the time sequence of images.

An initial segmentation is constructed for the current year, based on a combination of what is observed in the images and what is resident in the current model. For each band in each image of the current year, edge points are detected. A simple Sobel operator is used. The edge images are combined using an OR condition. An edge array for this year's images is formed by thinning and omitting those edge pixels which are likely to be interior to textured regions. Those edge pixels which remain are used either to support existing model region boundaries by adjusting their confidence, or to form additional boundaries. A full representation of this initial segmentation is constructed by filling in regions and calculating features for each region and boundary.

This initial segmentation is then refined in the light of the knowledge resident in the rule base. For each region, its possible class, or classes, are adjusted based on the expected properties of the class and the classes of its neighbours. The possible classes of each region each have an associated confidence and are determined from the current model. If there is no model, i.e. this is the first year's images and there is no map data, then all classes are possible. After this domain consistency checking, possible merges and splits are assessed using the same rules in the phase 1 system, but adjusted to take into account the probable region class.

A post segmentation stage is performed if there is already a model. This consists of updating the current model based on the changes which have arisen in the processing of the images of the current year. Those changes which are retained are output for change detection purposes.

After processing of all input images, the current model should represent a truer state of the situation on the ground and may be used for subsequent processing.

Due to the domain class information being included, a partial classification results, so we intend to use a modified maximum likelihood classifier in the evaluation of this system. The modified classifier will only consider those classes which are likely for each region, as output from the segmentation.

VI SUMMARY

An initial system for knowledge based segmentation of a time sequence of remotely-sensed images has been presented. Evaluation of this system has shown its potential for improved automatic segmentation of such images. All knowledge used was found to benefit segmentation, however, a number of drawbacks to this initial system have resulted in a re-design to produce a final system for this project. This system is currently being implemented.

This final system uses knowledge of the types and forms of objects to be expected in the scene to improve an initial segmentation. The approach taken is a model driven one, where the initial segmentation is formed from the correspondence between the boundaries located in the current time images and those in the current best model of the scene. Knowledge in the form of a digitised Ordnance Survey map is used to initialize the model, if such map data is available. In this way a model of the scene segmentation is built-up and successively improved through time.

As a consequence of the approach to segmentation, the system has a number of additional features. A partial classification is produced which may be used to aid subsequent processing. Changes through time are registered and may be used for change detection applications. However, without the assumption of geocoded data, these changes will include mis-registration errors if these errors are greater than those expected from random fluctuations between images. Finally, the current best model, as output from the system, may be used for updating or topologically structuring cartographic map data.

ACKNOWLEDGEMENTS

The authors would like to thank Nigel Brown and David Norris of the Institute of Terrestrial Ecology who have been responsible for digitising the Ordnance Survey maps.

REFERENCES

1. Cross, A. and Mason, D.C., 1985, Segmentation of remotely-sensed images by a Split and Merge process. Proc. Int'l Conf. on Advanced Technology for Monitoring and Processing Global Environmental Data, London, UK.
2. Levine, M.D., and Nazif, A.M., 1982, An experimental rule-based system for testing low level segmentation strategies. in Preston, K. and Uhr, L. (editors). Multicomputers and Image Processing : Algorithms and Programs. (Academic Press)

3. Goldberg, M., Goodenough, D.G., Alvo, M., and Karam, G.M., 1985, A hierarchical expert system for updating forestry maps with LANDSAT data. Proc. IEEE, 73, 1054.
4. Henderson, R.G., 1976, Signature Extension using the MASC algorithm. IEEE Trans. GE, 14, 34
5. Landgrebe, D.A., 1981, Analysis Technology for Land Remote Sensing. Proc. IEEE, 69, 628
6. Drake, N.A., 1986, A description of AMOEBA : A spectral-spatial clustering program. Tech. Memo NUTIS, Reading University, England.
7. Matsuyama, T., 1986, Knowledge-based Aerial Image Understanding Systems. Proc. Int'l Geoscience and Remote Sensing Symp. (IGARSS' 86), 277
8. McKeown, D.M., Harvey W.A. and McDermott, J., 1985, Rule-based Interpretation of Aerial Imagery. IEEE Trans. PAMI, 7, 570.
9. Nazif, A.M., and Levine, M.D., 1984, Low level Image Segmentation : an Expert System. IEEE Trans PAMI, 6, No 5, 555.
10. Nevatia, R. and Babu, K.R., 1979, Linear feature extraction and Description. Proc. Int'l Joint Conf. on AI, 6, 639.
11. Peleg, S., 1984, Classification by Discrete optimization. Comp. Vis. Graph & IP, 25, 122.
12. Seddon, A.M. and Hunt, G.E., 1985, Segmentation of Clouds Using Cluster Analysis. Int'l J. Remote Sensing, 6, 717.
13. Tailor, A.M., Cross, A., Hogg, D.C., and Mason, D.C., 1986, Knowledge-based Interpretation of Remotely-Sensed Images. Image and Vision Computing, 4, 67.
14. Tailor, A.M., D.G. Corr, P. Cosoli, A. Cross, D.C. Hogg, D.H. Lawrence, D.C. Mason, M. Petrou, R.D. Vango. 1987 'A System for Knowledge-Based Segmentation of Remotely-Sensed Images' Proc. IGARSS '87., 18-21 May, Ann Arbor, Michigan.
15. Townshend, J.R.G and Justice, C.O., 1980, Unsupervised Classification of MSS Landsat Data for mapping spatially complex vegetation. Int. J. Remote Sensing, 1, 105.

