PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A
MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

Development of a Knowledge-Based Segmentor for Remotely Sensed Images [and Discussion]

A. M. Tailor, D. G. Corr, A. Cross, D. C. Hogg, D. H. Lawrence, D. C. Mason, M. Petrou, R. D. Vango and D. Lane

Phil. Trans. R. Soc. Lond. A 1988 **324**, 437-446
doi: 10.1098/rsta.1988.0031

Phil. Trans. R. Soc. Lond. A 324, 437–446 (1988)
Printed in Great Britain

Development of a knowledge-based segmentor for remotely sensed images

By A. M. Tailor†, D. G. Corr†, A. Cross‡, D. C. Hogg§, D. H. Lawrence§, D. C. Mason‡, M. Petrou‡ and R. D. Vango†

† *Systems Designers Scientific, Pembroke House, Pembroke Broadway, Camberley, Surrey GU15 3XD, U.K.*

‡ *Natural Environment Research Council Unit for Thematic Information Systems, Reading University, P.O. Box 227, Reading RG6 2AB, U.K.*

§ *University of Sussex, Falmer, Brighton BN1 9QN, U.K.*

To interpret remotely sensed data accurately, a variety of information is required about the sensor, the sensing conditions, the scene and the application. At present a good interpretation can only be achieved by a photointerpreter, unless the problem is small and can be well defined. To exploit to the full the wealth of satellite data available, efficient automatic techniques are required.

This paper describes the progress of a U.K. Alvey Man–Machine Interface Project to develop a system for the knowledge-based segmentation of multitemporal remotely sensed images. The knowledge used includes both domain-dependent knowledge and knowledge contained in maps and previous segmentations. The domain-dependent knowledge includes information about the types of objects to be expected in the scene and their relations. The knowledge contained in the maps and previous segmentations could be regarded as a crude model of the situation on the ground which, although probably only partly correct, should still greatly facilitate segmentation.

A major objective is to increase classification accuracy in images of the type used in the environmental sciences. A further use could be to update the existing map data as a result of segmentation.

Introduction

The volume of quantitative data from remote-sensing platforms has increased considerably in the last decade. This has resulted in a corresponding increase in the scope and magnitude of the problems to which the images are applied. Present-day problems that use remotely sensed images cover applications as widely varying as urban land use studies and crop inventories (see, for example, Landgrebe (1981) for typical applications).

Traditionally, interpretation of remotely sensed data is performed by a photointerpreter using information in the image, together with his expertise and any available data, such as maps and ground truth. A full interpretation of the image is still best performed this way, although it is a lengthy process. There is a need for efficient machine-implemented analyses of the data.

The automatic techniques used for image interpretation are traditionally those of pattern recognition. The stages of a typical pattern-recognition system are shown in figure 1. The input to such a system for remote sensing is a digital image. The image is first preprocessed as necessary; this may not only include standard corrections, for example, registration to a reference grid or correction to reduce sensor-induced noise or distortions, but may also include

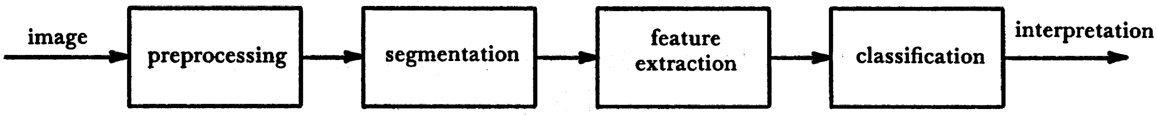


FIGURE 1. Block diagram of a traditional pattern-recognition system.

image transformations to enhance or suppress specific image features. The preprocessed image is then partitioned into segments, or regions, at the segmentation stage. Features, or properties, of the regions are extracted which describe each region uniquely and enable the correct class to be assigned to the region at the classification stage. A full interpretation of the scene can then be performed.

Most automatic analyses of remotely sensed imagery still rely only on the supervised spectral analysis of pixels on an individual basis. This process omits the automatic segmentation stage: a photointerpreter trains the classifier by selecting typical examples of the regions of interest. The spectral properties of the example regions are used as identifiers of the relevant class. Each new pixel is assigned to the class whose spectral properties are most similar to those of the pixel. Extensive training is required and the method is generally only satisfactory for highly constrained applications, where the spectral properties of the specific regions are sufficient for the problem to be adequately defined.

More recently, unsupervised methods have become more common. Here, no *a priori* assumptions are made about the classes of regions in the data and automatic techniques are used to divide the image. In addition, methods that use information other than the spectral values of pixels are being used. These considerations will be discussed in the sections below.

SEGMENTATION

Definition of segmentation

There is no absolute definition of segmentation in the literature. Segmentation may be said to be a partitioning of an image into segments, each of which is a group of adjacent pixels defined by properties which lie within a certain range. This reduces the problem to finding properties which, when their values are limited, define similarity within spatially connected areas of the image.

Segmentation is application dependent in that different levels of partition will be required by different applications. For example, an agriculturist requires segmentation to the individual field level, whereas a forester is content with a larger agricultural category.

Segmentation is also sensor dependent: two images of the same scene taken at different spatial resolutions, or that use different spectral wavebands, will produce a different segmentation. In addition, variations in the sensing conditions across an image and between images taken at different times will affect the segmentation: not only the values, but also the relevant properties required for isolation of an object may vary.

Goals of segmentation

The purpose of a segmentation is to create a low-level representation of the image which retains as much information as is necessary for classification, while presenting that information in a straightforward manner. To this end, there are two principal goals of a system for segmentation, namely: to produce a reliable segmentation; to improve classification accuracy.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

For investigation of general segmentation methods, the influences of the application, the sensor and the day-to-day variations must be reduced. The dependence on the application cannot be eliminated, but segmentation to some arbitrary level can be attempted. However, the techniques must be parametrized to cater for all input image dependencies.

Segmentation techniques

Segmentation techniques may be region-based or edge-based. Region-based techniques locate regions by using a measure of similarity between the given pixel properties. Edge-based techniques locate the region boundaries by detection of large discontinuities in the given properties. All such techniques are data-driven in that they rely on the properties in the input image to define the segments.

Region-based techniques do not cope well with thin regions and the boundaries are not normally located accurately. Examples of such techniques used for remotely sensed data are clustering (see, for example, Townshend & Justice 1980; Seddon & Hunt 1985) and split-and-merge (see, for example, Cross & Mason 1985). Clustering iteratively merges pixels until the similarity criterion is satisfied for each region. Split-and-merge techniques successively split and merge blocks of adjacent pixels to satisfy the criterion.

Edge detection techniques are principally concerned with locating points of large gradient and linking these points into connected boundaries (see, for example, Nevatia & Babu 1979). Consequently, such techniques are more successful in locating the region boundaries and thin regions, provided that the contrast is strong enough.

Errors in segmentation

None of the standard segmentation techniques is adequate for the segmentation of complex images from aircraft or satellite. Even if the technique performs to the best of its ability, segmentation errors and ambiguities still occur. These errors manifest themselves as over-segmentation, in break-up of a region, and undersegmentation, in either full or part loss of a region.

These errors may be due to practical issues, for example the sensing conditions, but the principal source of error is the data-driven aspect of the techniques. There is generally insufficient information in the image itself to discriminate completely between regions. Even if the information is in the image, the exact properties and values to use are not known *a priori*. Recourse to information outside that in the input image is necessary.

Exploiting knowledge

Knowledge available

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, which represent the expertise of an analyst, are generally in the

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

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 classes of regions at the classification level. The spatial information on the properties may be two- or three-dimensional.

Techniques for incorporating knowledge

The standard region and edge-segmentation techniques mentioned previously generally only use limited spectral and spatial information from the image. Techniques that exploit other forms of image information include relaxation, which uses local relations 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 Tailor *et al.* (1986). In particular, image understanding systems (Matsuyama 1986) may be built up to improve aspects of the interpretation process. Ideally, the knowledge would be used as a model of the state on the ground, for example, the system of McKeown *et al.* (1985) for the interpretation of airport scenes in aircraft data. However, the complex nature of the data means that 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, for example, the system of Nazif and Levine (1984) for segmentation of aerial imagery. Depending on the nature of the information, most systems combine the modelling and the rule-based approaches.

The remaining sections of the paper are concerned with a knowledge-based system which we have developed for segmentation of multitemporal, high-altitude aircraft and satellite data. The emphasis is on general segmentation techniques, as in the Nazif & Levine (1984) system, rather than on interpretation of more specific information, as in Goldberg *et al.* (1985) and McKeown *et al.* (1985). However, compared with the system of Nazif & Levine (1984), we include a modelling, or goal-directed, approach to supplement the data-driven approach, together with the use of temporal information.

Knowledge-based segmentation system

The preliminary system for knowledge-based segmentation developed in this project uses map data, together with domain data in the form of rules, to supplement the information in a multitemporal aircraft data set. Segmentation is to the lower order of segment, e.g. a field, rather than the higher order of segment, e.g. agriculture.

The rules in the rulebase have been designed to use as little knowledge of the classes of the regions in the data as possible, other than that in the map. The classification information that is known *a priori* is generally either specific to an application or depends on the sensing conditions. As the system improves it is the intention to add classification information as necessary.

The knowledge used is currently all two dimensional; no elevation data are included.

The approach to segmentation is a combination of the data-driven and the goal-directed approaches, the latter being evident due to the use of the map as a crude model of the situation on the ground.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

The basic system is modular and is shown in figure 2. An initial segmentation of the current image is formed by imposing lines, for example roads, from the map onto a region segmentation of the image. An edge segmentation of the image is produced to supplement the initial segmentation. Features of the regions are calculated which are used to satisfy constraints in the rulebase. These features include the uniformity of the region, the contrast between the region and a neighbour, and the support to the region boundary from edge and line pixels.

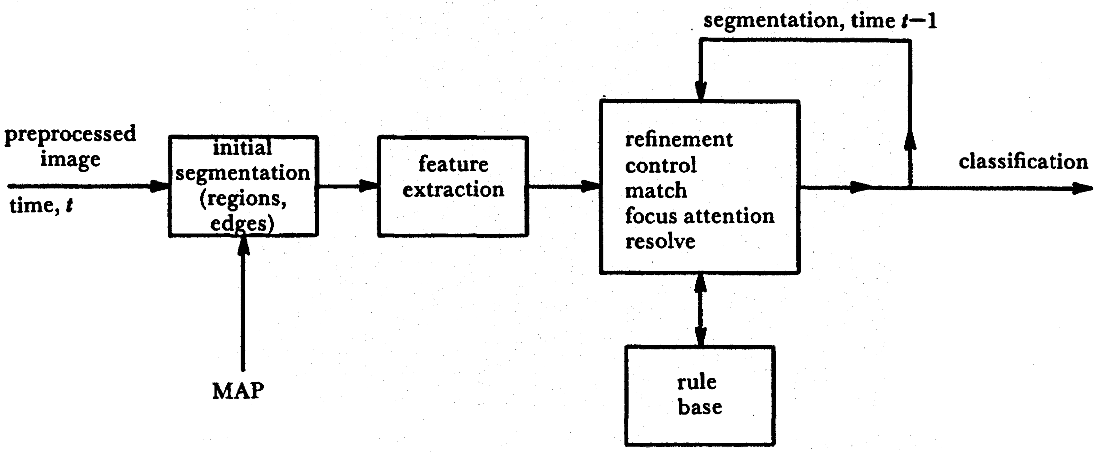


FIGURE 2. Block diagram of the knowledge-based segmentation system.

The initial segmentation is refined, based on the evidence from the region features and the constraints imposed on the regions in the rulebase. At this stage, the segmentation of the previous image in the time sequence, if available, is matched to the initial segmentation to assess temporal consistency. A measure of confidence is assigned to each region based on the compatibility of the evidence. Attention is focused onto ambiguity in the segmentation, as represented by low-confidence regions. If the evidence strongly supports a change, a region may be modified by merging with a neighbour, splitting, or adjusting a portion of the boundary.

After processing all the images in the time sequence, the segmentation may be classified.

DATA

The system has been tested on multitemporal image data from the NERC simulated thematic mapper aircraft data. Each image has 11 spectral channels and spatial resolution of 10 m. The data were transformed by using a principal component transform to decorrelate the spectral information.

Figure 3 shows a 64×64 pixel, first principal component section of one image in a three-image time sequence. The image consists of a combination of unstructured, natural fenland in the top-right corner and structured fields with well-defined boundaries. There is little elevation.

Figure 4 shows an area of digitized Ordnance Survey 1:10000 map data which corresponds to the image. Each of the lines in the map is 1 pixel wide. The two lines running from left to right across the map are roads; the incomplete lines running into the fenland are footpaths; the remainder are field boundaries.

A hand segmentation has been produced for each image, which is used as a reference for validation of the segmentation. Figure 5 shows the hand segmentation into regions corresponding to figure 3. There is a corresponding hand segmentation into lines which includes the

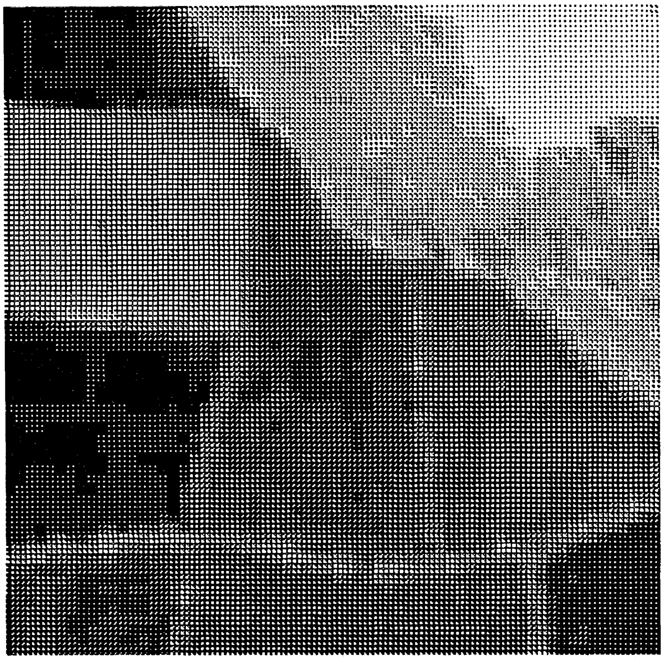


FIGURE 3. Aircraft image data of Barton Broads, spring 1986. The figure shows a first principal component, 64 × 64 pixel image at resolution of 10 m.

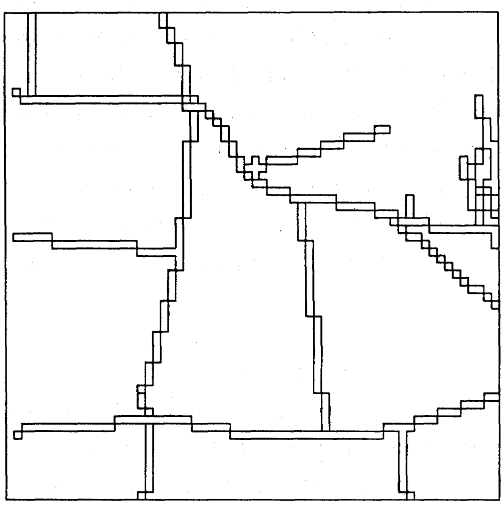


FIGURE 4. Digital Ordnance Survey 1:10000 map data corresponding to the area of figure 3.

thin regions as suggested by the map. The hand segmentation has been drawn up by a photointerpreter using information on the identity of regions from the image, the map and the ground truth.

The image, the map and the hand segmentation have each been registered to the British National Grid.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

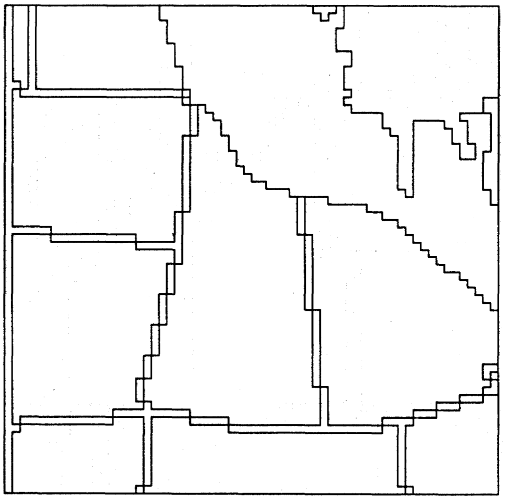


FIGURE 5. Hand segmentation of the scene into regions.

TESTS

The system was tested by using the image and map data described above. Various configurations of the system were tested to assess the usefulness of the knowledge. For example, the system was run with and without the map data.

In addition, a segmentation by a clustering procedure was performed on the last image in the time sequence for comparison. Both segmentations were validated against the hand segmentation.

To assess the ability of the system to improve classification, the segmentation was classified by using a region classifier. The classifier assigns a region to the class whose spectral characteristics are most similar to those of the region. The classes are defined by a photointerpreter labelling the regions in the hand segmentation. The classification of the segmentation was compared with that obtained by a standard per-pixel classifier.

The percentage of correctly classified pixels was calculated for both the region and the per-pixel classifications.

RESULTS

Figure 6 shows the result of the knowledge-based segmentation of the image in figure 3. In general there is good correspondence between the regions and those of the hand segmentation in figure 5.

The boundaries of regions are not detected well: they are often misplaced, doubled up or represented by a number of small regions. This is due to a number of factors.

Firstly, the line information, after being imposed on the initial segmentation, is allowed to change at the refinement stage. This is consistent with the assumption that the information may be incorrect. The condition for line accuracy is however too strict, resulting in the field boundaries in the centre of the image being modified.

Double boundaries result in general from lack of good registration between the image and map data: both strong edges and map lines may reinforce a boundary, but at different positions.

The break up of a boundary into small regions is due to the large number of high-contrast edges present. A similar effect occurs in high variance or textured regions such as woods or

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

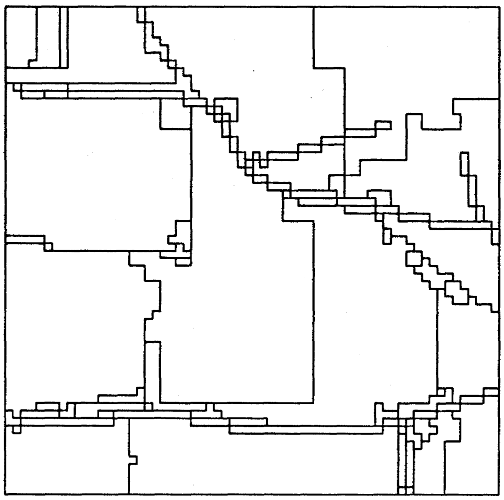


FIGURE 6. The result of knowledge-based segmentation of the scene.

urban areas. The system at present has no provision for combining these small regions into a consistent interpretation. However, this is principally a problem for the higher level stages of classification and interpretation.

Undersegmentation is a problem in the top-right corner of the image where the spectral values of the pixels are insufficient for discrimination of the two regions in the fen area. This is due to the segmentation being performed on only the first principal component of the data. The area is too uniform in the one component for any evidence to be gathered for splitting the resultant region.

Figure 7 shows the result of the segmentation by clustering of the image. Oversegmentation of high-variance regions is also a problem here, not only on the boundaries, but also interior to regions. In addition, undersegmentation is much more evident than in the knowledge-based segmentation; there is no means by which the clustering can divide similarly valued pixels. On this basis, the knowledge-based segmentation performance is superior to that of the clustering.

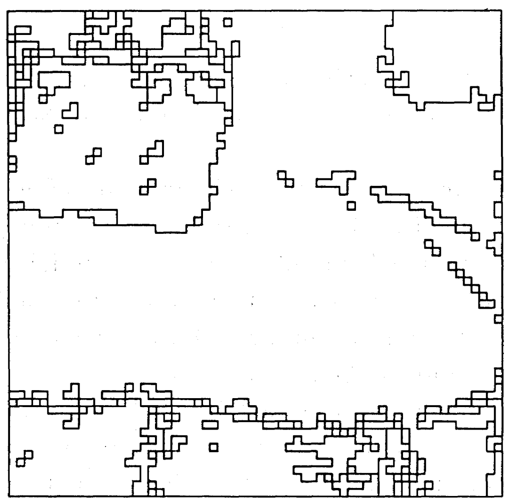


FIGURE 7. Segmentation of the scene by a standard clustering procedure.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

The classification of the knowledge-based segmentation of the image gave 81.4% of pixels correctly classified.

Many of the segmentation ambiguities due to oversegmentation at the boundaries and in textured regions were resolved at the classification stage, because of the regions having spectral values sufficiently similar. However, the undersegmented area in the fen, because it was made up of pixels from two different distributions, was classified as neither class. This contributed greatly to the classification error.

The presence of the roads from the map in the segmentation caused the majority of the remaining classification error. These line features were not included in the region hand segmentation, but are present in the line hand segmentation and are treated separately.

The accuracy of the per-pixel classifier was 86.2%. The greater accuracy was due to the two regions in the fen area being correctly classified when examined on a pixel-by-pixel basis with more than one spectral component. This classifier, however, suffers from high variance in the image in a similar way to the segmentation: many small regions result. The manifestation of the small regions at this level of the interpretation process is much more difficult to amend.

Discussion

The preliminary system for knowledge-based segmentation described here compares favourably with the standard automatic techniques. However, there have been a number of inadequacies isolated in the knowledge-based system.

The initial segmentation is not adequate for successful refinement to take place. Over-segmentation at this stage in high variance regions causes problems: a separate system module is required for dealing with these areas. In addition, the initial region and edge segmentation procedures used cause many spurious features in noisy areas. This is due partly to the inferior techniques and partly to the lack of a full multispectral capability.

The refinement stage is deficient in one major area: the treatment of regions or their boundaries which have resulted from the imposition of map lines. Greater emphasis on these areas is required to avoid losing the information. This should include a mutual enforcement of map lines and image edges.

It is planned to improve and extend the system in the light of the results to produce a demonstration operational system. This system will be designed principally for the land use applications of environmental and agricultural monitoring. However, we envisage that it will have general applicability to the remote-sensing community. This is due to the modularity of the system which will allow its adaptation to other sources of knowledge or other requirements. This final system will also provide explicit information on map updates and image change detection. Future tests will include the processing of other sensor data, such as that from the French *SPOT* satellite and synthetic-aperture radar data.

The final stages of the current project will include studies of the hardware required by a full operational system, the augmentation of the knowledge-base to include classification information and pointers to potential areas of research.

We thank Nigel Brown and David Norris of the Institute of Terrestrial Ecology who have been responsible for digitizing the Ordnance Survey Maps.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

References

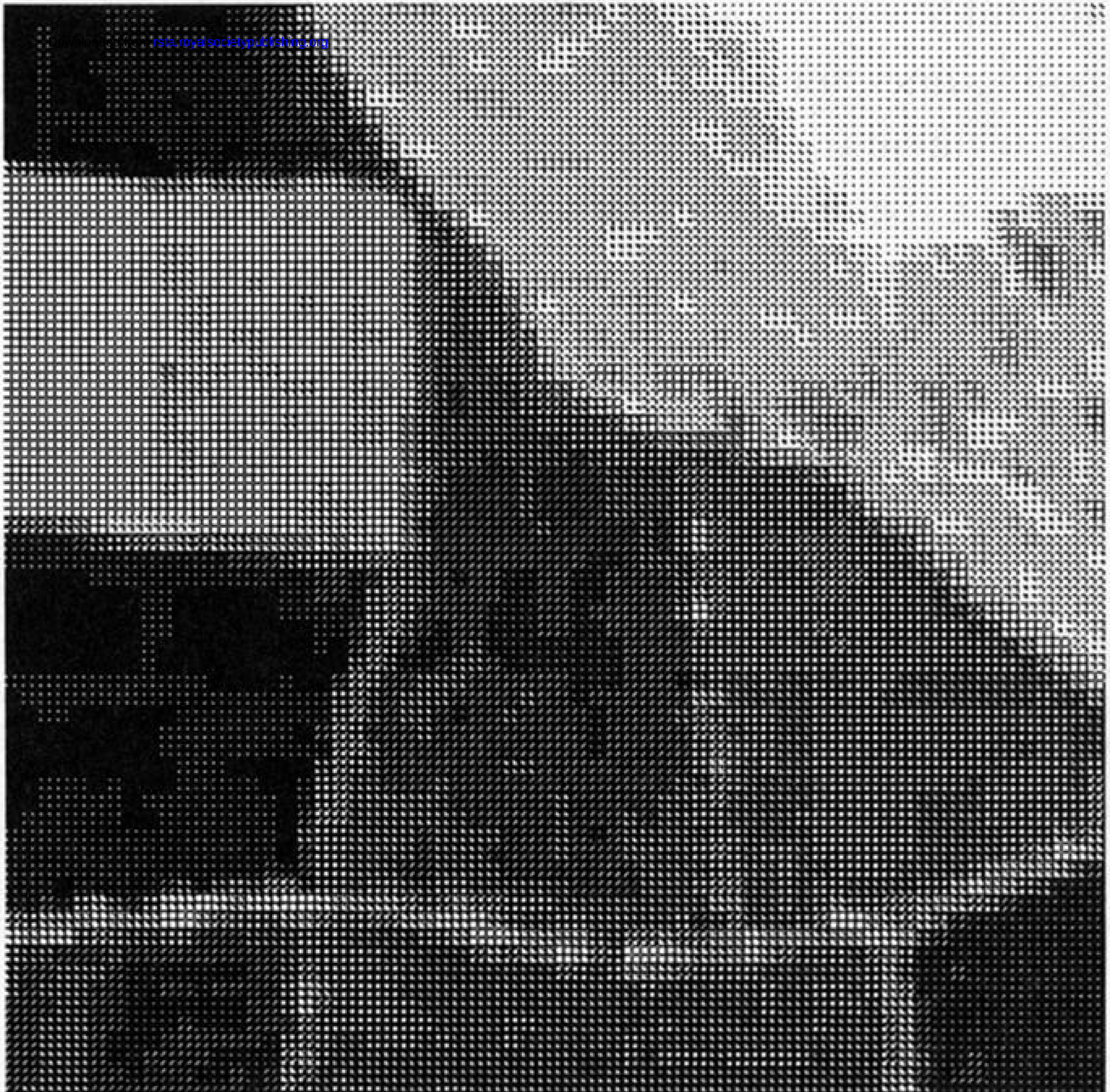
Cross, A. & Mason, D. C. 1985 Proc. Int. Conf. of the Remote Sensing Society and the Center for Earth Resources Management on Advanced Technology for Monitoring and Processing Global Environmental Data, London, 9–12 September 1985.
Goldberg, M., Goodenough, D. G., Alvo, M. & Karam, G. M. 1985 *Proc. IEEE* **73**, 1054–1063.
Henderson, R. G. 1976 *IEEE Trans.* **GE 14**, 34–37.
Landgrebe, D. A. 1981 *Proc. IEEE* **69**, 628–642.
Matsuyama, T. 1986 In *Proc. Int. Geoscience and Remote Sensing Symp, Zürich, 8–11 September 1986*, pp. 277–282. ESA Publications Division.
McKeown, D. M., Harvey, W. A. & McDermott, J. 1985 *IEEE Trans.* **PAMI 7**, 570–585.
Nazif, A. M. & Levine, M. D. 1984) *IEEE Trans.* **PAMI 6**, 555–577.
Nevatia, R. & Babu, K. R. 1979 In *Proc. 6th Int. Joint Conf. on Artificial Intelligence, Tokyo, 20–23 August 1979*, vol. 6. pp. 639–641.
Peleg, S. 1984 *Comput. Vision Graph. Image Process.* **25**, 122–130.
Seddon, A. M. & Hunt, G. E. 1985 *Int. J. Remote Sensing* **6**, 717–731.
Tailor, A. M., Cross, A., Hogg, D. C. & Mason, D. C. 1986 *Image Vis. Comput.* **4**, 67–84.
Townshend, J. R. G. & Justice, C. O. 1980 *Int. J. Remote Sensing* **1**, 105–120.

Discussion

D. Lane (*Intelligent Automation Laboratory, Department of Electrical and Electronic Engineering, Heriot-Watt University, Edinburgh, U.K.*). Would Dr Tailor briefly describe the hardware and software environments that she is using for this work, and comment upon the advantages and disadvantages that have resulted.

A. M. Tailor. The system has been implemented in POP 11, a general-purpose programming language, integrated into the POPLOG environment. POPLOG has been developed at Sussex University and is marketed commercially by Systems Designers. The system is running both in a VAX environment and on a SUN 3 workstation.

Major advantages have been found using POP 11; namely the ease of development and flexibility in implementation. However, for an operational system, both some recoding into a language such as PASCAL or C, and use of special-purpose hardware would be required, because of the current long execution times.



GURE 3. Aircraft image data of Barton Broads, spring 1986. The figure shows a first principal component, 64 × 64 pixel image at resolution of 10 m.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES