

Segmenting Multispectral Landsat TM Images Into Field Units

Carolyn Evans, Ronald Jones, Imants Svalbe, and Mark Berman

Abstract—This paper presents a procedure for the automated segmentation of multispectral Landsat TM images of farmland in Western Australia into field units. The segmentation procedure, named the canonically-guided region growing (CGRG) procedure, assumes that each field contains only one ground cover type and that the width of the minimum field of interest is known.

The CGRG procedure segments images using a seeded region growing algorithm, but is novel in the method used to generate the internal field markers used as “seeds.” These internal field markers are obtained from a multiband, local canonical eigenvalue image. Before the local transformation is applied, the original image is morphologically filtered to estimate both between-field variation and within-field variation in the image. Local computation of the canonical variate transform, using a moving window sized to fit just inside the smallest field of interest, ensures that the between- and within-field spatial variations in each image band are accommodated. The eigenvalues of the local transform are then used to discriminate between an area completely inside a field or at a field boundary.

The results obtained using CGRG and the methods of Lee [1] and Tilton [2] were numerically compared to “ideal” segmentations of a set of sample satellite images. The comparison indicates that the results of the CGRG are usually more accurate in terms of field boundary position and degree of over-segmentation and under-segmentation, than either of the other procedures.

Index Terms—Image segmentation, multispectral images, multivariate statistics.

I. INTRODUCTION

AGROWING proportion of agricultural regions in Western Australia are becoming salt-affected and infertile. Salinity maps are being produced using bands 1–5 and 7 of Landsat Thematic Mapper (TM) images to assist farmers to manage their land [3]. These salinity maps contain a ground cover label for every pixel in the Landsat TM image. Currently the spectral properties of a single pixel, together with those of its 8 nearest neighbors, are used to identify the most likely ground cover type of that pixel. However classification results tend to be unreliable near field boundaries where ground cover type is less predictable. Roads or trees sometimes border fields and crop types tend to mix near field boundaries.

Manuscript received August 23, 2001; revised March 1, 2002. This work was supported in part by the Australian Research Council and the Division of Mathematical and Information Sciences, CSIRO Australia.

C. Evans and M. Berman are with the Division of Mathematical and Information Sciences, CSIRO, New South Wales, Australia (e-mail: carolyn.evans@csiro.au; mark.berman@csiro.au).

R. Jones is with Proteome Systems, Ltd., Sydney 2113, Australia (e-mail: ronald.jones@proteomesystems.com).

I. Svalbe is with the Department of Physics and Materials Engineering, Monash University, Melbourne 3800, Australia (e-mail: imants.svalbe@spme.monash.edu.au).

Publisher Item Identifier S 0196-2892(02)05277-4.

Given the research of Kettig and Landgrebe [4] and Janssen and Molenaar [5], it is thought that the accuracy of salinity maps could be improved by replacing the pixel-based classification procedure with a region-based procedure. If the field boundaries were known a priori, the spectral properties of an entire field could be used to identify the most likely ground cover type of the field. Higher-level knowledge about the shape and texture of the field may also be used in a region-based classification process. These possibilities motivate our research into automatic segmentation of Landsat TM images into field units.

A. The Segmentation Problem

Fig. 1 shows bands 1–5 and 7 of a sample multispectral Landsat TM image. (Each band has been histogram equalized for display purposes.) Landsat TM images of fields, such as this example, are difficult to segment for several reasons. Information about objects in these images is distributed differently over multiple image bands. For example, in Fig. 1 band 1 has relatively more within-field variation than band 7. This variability must be taken into account to optimize segmentation results. In addition, each field may not contain a uniform ground-cover type. Small regions, corresponding to clumps of forest, bare soil or salt-affected vegetation, occur inside fields regularly throughout these images. Some of these regions have been circled in Fig. 1(d). These regions are spectrally atypical with respect to the majority of values inside the local field.

Fields are also not uniform in shape. It has been estimated that only 60% of the fields in Western Australia are rectangular [6] so applying a rectangular constraint on the field is generally not appropriate. A size constraint is more appropriate. The smallest field unit of interest for the production of salinity maps has been estimated to be 330 m wide, or 11 pixels [6] (each pixel represents a 30 m by 30 m area).

B. Previous Approaches to Segmenting Multispectral Satellite Images

Similar problems to the one described in Section I-A have been considered previously [1], [2] and [7]–[11]. The automatic segmentation procedures designed for segmenting fields, comparable in size to those found in Western Australia, can be described as either boundary-based, region-based or a hybrid of the two.

Boundary-based segmentation methods (e.g., Ji [9]) depend on pixel values changing rapidly at region boundaries in an image. Ji applies a (manually determined) threshold to an image gradient to accurately locate region boundaries, but the boundaries are often incomplete and must be closed using

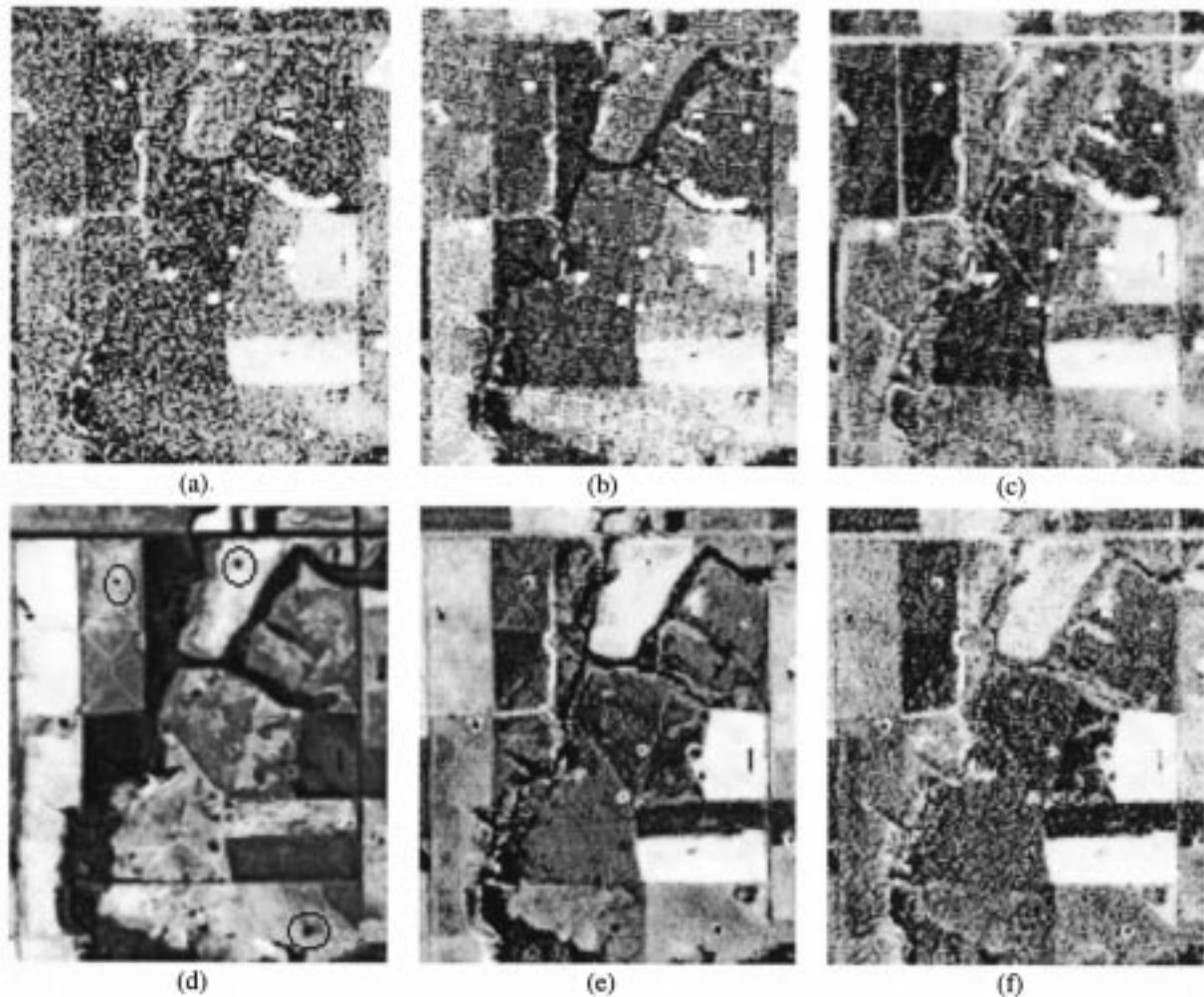


Fig. 1. Sample multispectral Landsat TM image of farmland near Esperance in Western Australia. (a)–(e) Bands 1–5, and (f) band 7. Each band has been histogram equalized for display purposes. Locally atypical within-field variation has been circled in (d).

other means before spatially connected regions correspond to individual fields.

Region-based segmentation methods partition images into a set of regions with closed boundaries, working under the assumption that pixels within the same region have a certain spectral homogeneity. Often, an image is first over-segmented. Neighboring regions are iteratively merged if they satisfy a homogeneity criterion. Kettig and Landgrebe [4], Barbosa *et al.* [7] and Lobo [11] start by defining each pixel to be its own region. Lee [1] uses evenly-tiled, 3×3 “seeds,” spaced closer than the width of a typical field in the image, as starting points for the seeded region growing (SRG) technique [12] to produce the initial over-segmentation. He then merges regions according to the minimum description length (MDL) principle [13].

In Tilton’s procedure [2] the homogeneity criterion is determined from the image data itself and is adaptively varied. A hierarchy of fine-to-coarse segmentation levels are produced by relaxing the homogeneity criterion that controls region merging. The user must still decide, however, which segmentation level is the most appropriate for their particular application.

Making an appropriate choice of homogeneity criterion is critical to the success of any region-based segmentation proce-

dure, especially for multispectral images, where the criterion is highly dependent on the choice of “spectral closeness” metric. Many definitions of multispectral distance (gradient) exist [14]. For example, the Euclidean distance metric, used in [2], [10], [9], gives each band equal weight, and thus implicitly assumes that they all have the same quality (“signal-to-noise”), which is clearly not true of the images in Fig. 1. Segmentation results based on the Euclidean distance metric, or on band ratios [15], will also vary if one or more of the satellite image bands has undergone radiometric re-scaling during pre-processing. In contrast, the Mahalanobis distance metric [15] is insensitive to linear re-scalings of individual image bands and takes into account any between-band correlations in image data.

Hybrid segmentation methods, such as those proposed by Soille [10] and Janssen and Molenaar [5], segment images using region and boundary information with the aim of producing a segmentation containing closed regions with accurately positioned boundaries. (Soille’s also uses knowledge of the number of ground-cover types in the images in an unsupervised pixel classification stage of his procedure. This information is not available in our case.) These results are expected to be more accurate because they are based on all the available spatial and spectral information. Hybrid methods, whilst using more infor-

mation, must deal with the problem of how to weight and combine different kinds of information.

C. Organization of the Paper

In this paper, we attempt to overcome the difficulties described in Section I-A and present a fully automated hybrid, boundary/region-based procedure for segmenting multispectral Landsat TM satellite images into field units. This procedure is called canonically-guided region growing (CGRG). It assumes that each field unit contains only one ground-cover type. However a filtering stage is included in the procedure so that it is robust to locally atypical spectral values inside any field. CGRG uses a local Canonical transform to take into account the distribution of between-field and within-field information in the different image bands. The Mahalanobis distance metric is used to define “spectral closeness.” No assumptions are made about the shapes of fields, but the procedure requires the width of the smallest field of interest as an input parameter.

The paper proceeds as follows. In Section II, we outline the new CGRG segmentation procedure, before discussing it in detail. The segmentation procedure is then evaluated and compared with Lee’s and Tilton’s methods in Section III. Conclusions are drawn in Section IV.

II. OVERVIEW OF THE SEGMENTATION PROCEDURE

The CGRG procedure relies on automatically extracting internal field markers and partial field boundaries from the original TM image. The image is then segmented using SRG, with the extracted field markers being used as the “seeds.” The SRG algorithm is constrained to include the partial field boundaries. The novelty of our procedure lies in how we extract internal field markers and partial field boundaries from the multispectral images. These are extracted from an intermediate image, a local canonical eigenvalue image, calculated using all bands of the multispectral image. This should be more reliable than Lee’s evenly-tiled seeds. Rather than correcting over-segmentation by region merging, we focus on obtaining a reliable set of field markers to try to avoid over-segmentation.

The local canonical eigenvalue image is based on the eigenvalues of canonical transformations of local image data. We now discuss this transformation in detail and explain how the local canonical eigenvalue image can be used to find internal field markers and partial field boundaries.

A. Local Canonical Transformation

The canonical transform [15] is a generalized eigen-analysis transform, similar to the principal component (PC) transform and the maximum noise fraction (MNF) transform [16], [17], based on data-dependent covariance matrices. The canonical transform finds linear combinations of the image bands which successively maximize the ratio of the between-class variance to the within-class variance, subject to certain orthogonality constraints. In order to compute this transform, we need to know the between-class covariance, S_B , and the within-class covariance, S_W , or at least have good estimates of them. Normally, the class membership of each pixel is known or estimated from training

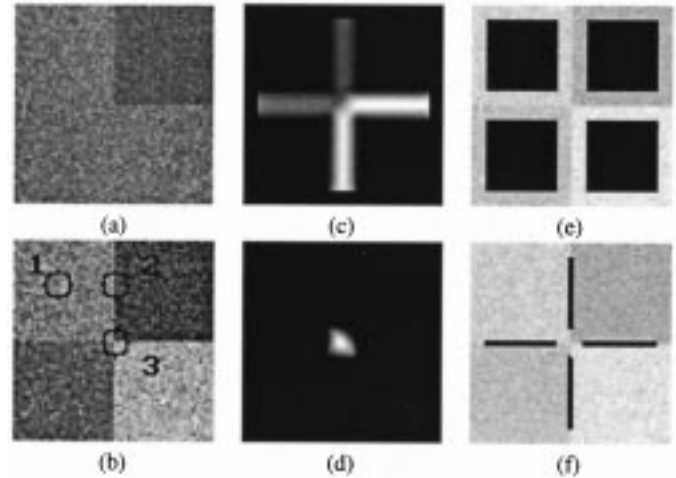


Fig. 2. (a) Second band of a synthetic three band image. (b) Different positions of the moving window of diameter 11, used to calculate canonical eigenvalue image. Refer to text for details. (c) First and (d) second canonical eigenvalue images calculated with moving window. (e) Region markers determined from where the first eigenvalue is small. (f) Partial region boundaries obtained from the ridge points of the first eigenvalue image where the second eigenvalue is small.

data. In this context, estimation of S_B and S_W is straightforward.

In the present context, a class corresponds to all the pixels in a field. This presents several challenges. First, field boundaries are not known a priori; these are what we are trying to estimate. This complicates the estimation of both S_B and S_W . Second, variation within a field may arise from the effect of salinity on crop growth, natural variation in crop growth, or rocky outcrops. Some of the latter are circled in Fig. 1(d). This means that S_W will need to be estimated robustly.

A saving grace here is that we do not need to obtain maximal discrimination between all fields in an image for segmentation—we only need to obtain maximal discrimination between adjacent fields that share a common boundary. This means we are not interested in performing a global transformation but rather a local transformation based on data contained inside a moving window. We choose to perform the local canonical transformation using a window sized to be large enough to capture robust statistics, yet small enough to fit completely within all fields of interest in the image.

To help us understand how this local canonical transform should work in theory, consider the simulated example in Fig. 2(a). It shows the second band of a color image, 75×75 pixels in size. The image contains four different colored regions, with randomly distributed, independent Gaussian noise in each of the three image bands. The ratio of between-to-within field variances is different in each band. Fig. 2(b) shows that as such a window moves across a satellite image it will sometimes only contain data from one field (e.g., position 1), whereas other times it will contain data from two fields (e.g., position 2) or more (e.g., position 3). The local canonical transform, along with its associated eigenvalues, will vary significantly between each of these positions because of their dependence on the local covariance matrix S_B .

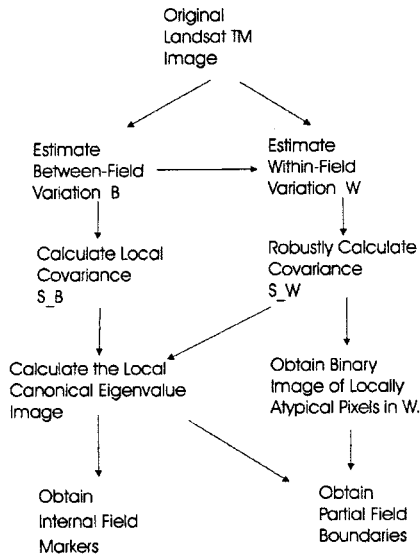


Fig. 3. Detailed description of the proposed segmentation procedure.

If S_W is constant throughout the image and is small relative to S_B [as it is in the example of Fig. 2(a)], then the eigenvalues of S_B and the canonical transform will behave similarly. The number of significantly large eigenvalues should correspond to the number of fields in the window minus 1. Therefore, when the window is located completely inside one field, all the eigenvalues should be small. When the window straddles the boundary between two fields, only the first eigenvalue should be significantly large. It can be shown that this eigenvalue is maximized when the moving window contains equal proportions of pixels from either field. (See [18] for a proof.) If the field boundary is locally straight, the first eigenvalue reaches a local maximum when the window is centred on the field boundary.

When the window straddles the boundary between more than two fields, at least the first two eigenvalues will be significantly large. However, the first eigenvalue no longer reaches a maximum when a certain proportion of each field is locally present, but depends on the mean spectral value of each field. (See [18] for a proof.)

Therefore, in theory, locally straight partial field boundaries can be determined by finding those locations where the second eigenvalue is sufficiently small and the first eigenvalue peaks and forms linear ridges along field boundaries. Field markers can be determined by finding where the sum of the local eigenvalues is small. This is demonstrated in Fig. 2(c)–(f).

It is easily shown that the sum of the local canonical eigenvalues equals the sum of the diagonal elements of $S_B S_W^{-1}$ (pg. 38 of [19]). This means that an image containing the local sum of the canonical eigenvalues can be computed without calculating the eigenvectors and associated eigenvalues of the local matrix $S_B S_W^{-1}$. The benefit of computing the *individual* eigenvalues is that the first eigenvalue can then be used to accurately locate the position of boundaries between two fields.

B. The CGRG Procedure in Detail

The flow chart in Fig. 3 shows in detail how internal field markers and partial field boundaries are extracted in the CGRG procedure.

First, the between-field variation, B , and within-field variation, W , in the original satellite image are estimated. Next the covariance matrices S_B and S_W are robustly computed from the estimated images B and W . A binary image of pixels with atypical within-field variations is produced simultaneously with robust computation of S_W . The local canonical eigenvalue image is then calculated using a moving local window. Internal field markers and partial field boundaries are determined by filtering the local eigenvalue image and the binary image of atypical pixels.

For the remainder of this section we will explain each stage of the CGRG procedure in detail and in the same order as shown in Fig. 3. We begin with the estimation of between-field and within-field variation in the images.

C. Estimating Between- and Within-Field Variations in the Landsat TM Images

When estimating the between-field variation in a multispectral image, we would like such a field to be preserved, but any smaller scale variation contained inside the field should appear in the estimate of within-field variation instead. We can try to separate within-field variations from the between-field variations in the TM images by filtering them using linear or morphological image filters. Morphological image filtering is preferred because, unlike linear filters, (a) they smooth by replicating local image values of a selected rank [20] and so are robust to locally atypical pixel values, such as those highlighted in Fig. 1(d), and (b) they do not blur information across image region boundaries.

By selecting appropriate windows (called structuring elements), morphological filters are sensitive to the size and shape of objects in an image. For example, the morphological filters denoted here by ϕ_{11} and γ_{11} are performed with a set of 20 oriented lines of length 11. (Twenty orientations are sufficient to cover the span of a digital circle of diameter 11.) ϕ_{11} is known as an intersection of closings [20]. It removes all dark image variations shorter than 11 pixels from an image. Its dual, the union of openings, γ_{11} , removes all bright image variations shorter than 11 pixels from an image. The length of 11 pixels was chosen to match the width of the smallest field of interest in the satellite images.

Similarly the closing and opening filters $\phi_{3 \times 3}$ and $\gamma_{3 \times 3}$ are performed using a square, 3×3 structuring element. The closing, $\phi_{3 \times 3}$, removes dark variations smaller than 3×3 pixels from an image, while its dual, $\gamma_{3 \times 3}$, removes bright variations smaller than 3×3 pixels. We make the conservative assumption that image variations narrower than 3 pixels in the image domain correspond to within-field variation. We do not impose a stricter width criterion on between-field variation to allow for nonrectangular fields.

The four filters, ϕ_{11} , γ_{11} , $\phi_{3 \times 3}$, $\gamma_{3 \times 3}$, can be applied sequentially to remove the within-field variations from each band of the original Landsat TM image, as, for example, the operation $\gamma_{3 \times 3} \phi_{3 \times 3} \gamma_{11} \phi_{11}(I_i)$ or the operation $\phi_{3 \times 3} \gamma_{3 \times 3} \phi_{11} \gamma_{11}(I_i)$, where I_i denotes the i th band of the image. Note that the first sequence has a bias toward returning high image values, because the closing ϕ_{11} is applied first. Similarly the second sequence is biased toward returning low image values because

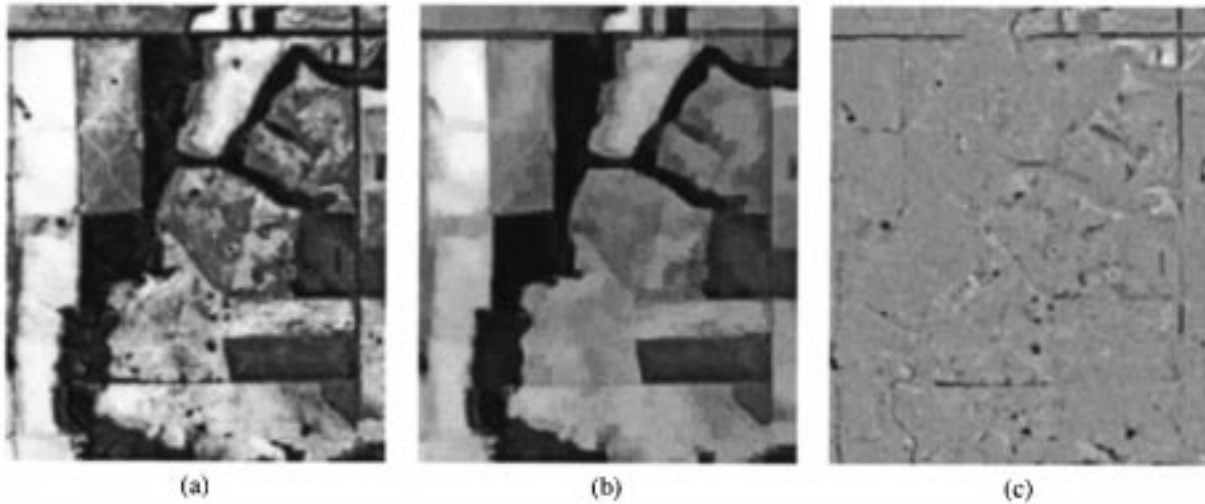


Fig. 4. (a) Band 4 of sample Landsat TM image. (b) Estimate of band 4 for image B made using the morphological filter in (1). (c) Estimate of band 4 of image W made by subtracting (b) from (a).

the opening γ_{11} was used first. To avoid any bias, we form a *self-dual* operator from the two sequences. The morphological operation used here to estimate the between-field variations, B , and the within-field variations W , for the i th band of the Landsat image is defined by

$$B_i = (I_i \vee \phi_{3 \times 3} \gamma_{3 \times 3} \phi_{11} \gamma_{11}(I_i)) \wedge \gamma_{3 \times 3} \phi_{3 \times 3} \gamma_{11} \phi_{11}(I_i) \quad (1)$$

$$W_i = I_i - B_i, \quad i = 1, \dots, p \quad (2)$$

where p is the number of bands (in our case $p = 6$), and \vee and \wedge represent the operations of pixel-wise maximum and minimum, respectively.

Fig. 4(b) and (c) show the estimated images B and W respectively, for band 4 of the original Landsat TM image. While the exact between-field variations of Fig. 1 are unknown, visual comparison of Fig. 4(a) and (b) indicates field boundaries have been preserved whilst removing locally atypical variations within fields.

D. Estimating the Covariance Matrices S_B and S_W

The covariance matrix S_B is calculated locally from the estimated image, B , using a moving local window. To eliminate orientational bias, we used a digital circular window with a diameter of 11 pixels so that it would just fit inside the smallest field of interest in the image. Updating formulae can be used to rapidly calculate the covariance as the window moves across the image, given that the window elements are uniformly weighted. Although the localization of field boundaries in the final eigenvalue image can be improved using a window with Gaussian weighting [21], updating can no longer be applied so that the local versions of S_B are much slower to compute.

The covariance matrix S_W is estimated robustly to prevent atypical within-field variation, such as evident in Fig. 4(c), from affecting the overall estimate. We wish the estimate of S_W to reflect the majority of within-field variation. The biweight S -estimator [22] is therefore used.

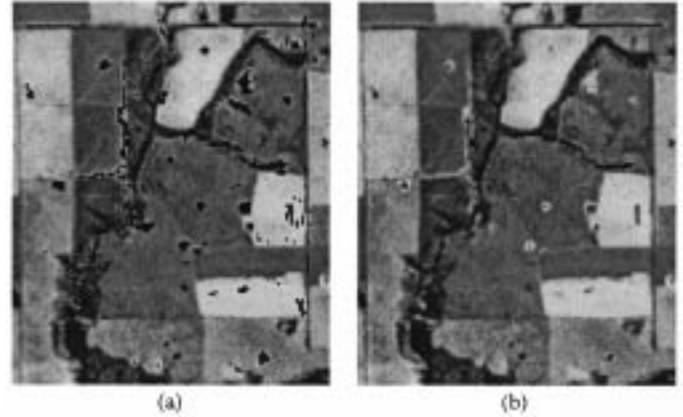


Fig. 5. (a) Pixels identified as outliers in the image W , overlaid on band 4 of the original Landsat TM image. (b) The result of filtering the binary outlier image of (a) using (7).

In S -estimation, initial estimates of S_W are computed based on multiple sets of $p + 2$ randomly selected pixels from image W . An initial sample size of at least $p + 1$ data points is required to obtain a meaningful estimate of the mean and all terms in the $p \times p$ covariance matrix, S_W . From each starting point an iterative estimation process begins. At each iteration, pixel values that are atypical, given the current estimate of S_W and assuming a multivariate Normal distribution, are either down-weighted or rejected from the next estimate of S_W . Arbitrarily we consider the iterative process to have converged when there is less than 0.01% change in the diagonal elements of S_W from one iteration to the next. Of the estimates of S_W obtained from each of the different starting points, the one with the smallest determinant is selected as the final robust estimate. As a by-product of the robust estimation process, a binary image of the atypical pixels rejected is formed. Fig. 5(a) shows the rejected, atypical pixels, A , from Fig. 4(c).

Fig. 6 show the range of variance terms from nonrobust and robust S -estimates of S_W calculated from data from 8 different sample images. The set of sample images was chosen (from one

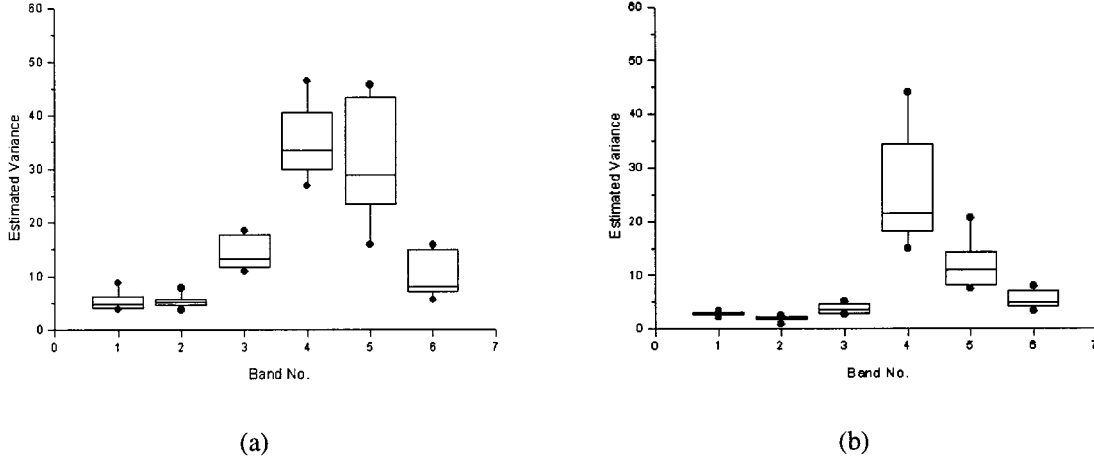


Fig. 6. Estimated variance terms from the covariance matrix S_W made from 8 sample satellite images: (a) non-robust variance estimates and (b) robust variance estimates.

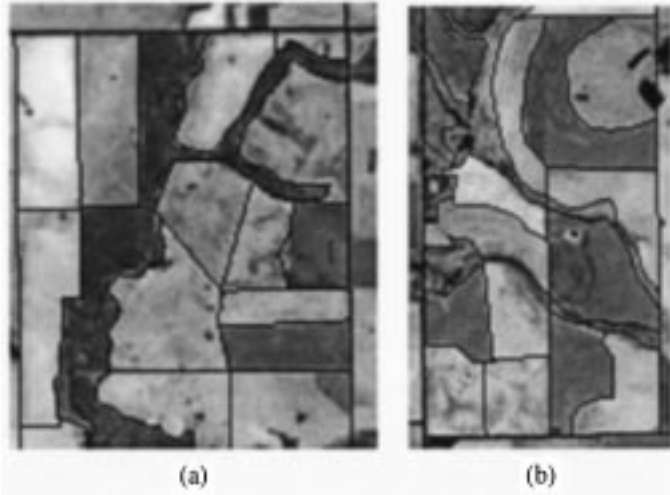


Fig. 7. Ideal segmentation of two sample Landsat TM images produced manually.

larger image) by a remote sensing expert [6] as being representative of the different types of field distributions usually found in the agricultural region. Two of the sample images are shown in Fig. 7, together with “ideal” segmentations produced manually by the remote sensing expert. The box plots [23] indicate that there is much more variation in the nonrobust estimates of the variance terms in S_W , where the contribution of atypical pixels have not been down-weighted or rejected, than in the robust estimates.

The range of the robust estimates of S_W over different sample images confirms that the within-field variations are not constant over the larger satellite image. Thus, to obtain an accurate estimate of the local canonical eigenvalue image, robust estimates of S_W should be computed locally. However, computationally this is too expensive. For example, to estimate S_W for a sample image such as those shown in Fig. 7(a), we randomly chose five different starting estimates of S_W and between six to eight iterations were required from each starting point before convergence. If S_W were to be computed using a local moving window, the same level of computation would be required every time the window moved.

Instead, we assume that within-field variation in the Landsat TM images varies much more slowly than between-field variation and one global estimate of S_W is sufficient. In this paper, we use the robust estimate of S_W made from the sample image shown in Fig. 1 as the global estimate because its variance terms were the median of the ranges plotted in Fig. 6(b). We compensate for resulting inaccuracies in the eigenvalue image by filtering, as described in Section II-E.

E. Obtaining Markers of Fields and Field Boundaries

The first and second canonical eigenvalue images, λ_1 and λ_2 , and the sum of all the local canonical eigenvalues $\Lambda = \sum_{i=1}^p \lambda_i$ for Fig. 1, are shown in Fig. 8(a), (b), and (c), respectively. Because the canonical eigenvalues in this real-world image do not behave quite like those in the synthetic image Fig. 2(a), further pre-filtering is required before internal field markers are extracted.

Spurious, small scale minima and maxima in the eigenvalue image are removed from the eigenvalue images λ_1 , λ_2 and Λ , by sequentially applying an opening, $\gamma_{5 \times 5}$, and closing, $\phi_{5 \times 5}$, using a 5×5 structuring element. (The diameter of this structuring element was chosen to be half the diameter of the window used to compute the local eigenvalue image.) The eigenvalue peaks are then separated from the background using a morphological top hat [20]. The sharpness of the peaks is limited by the 11 pixel diameter of the moving window used to calculate the eigenvalue image. The morphological top hat performed with a circular structuring element 13 pixels in diameter will retain the bright peaks of the eigenvalue image and remove the more slowly varying background. Internal field markers were then obtained where the top hat of the smoothed eigenvalue image, Λ , equals zero, that is

$$\gamma_{5 \times 5} \phi_{5 \times 5} \gamma_{5 \times 5}(\Lambda) - \gamma_{13} \gamma_{5 \times 5} \phi_{5 \times 5} \gamma_{5 \times 5}(\Lambda) = 0. \quad (3)$$

The internal field markers obtained in this way for Fig. 1 are overlayed in Fig. 8(d). It is these markers that are used as seeds for seeded region growing, as described in Section II-F.

Given the response of the first and second local canonical eigenvalues at the boundaries of two or more adjacent fields

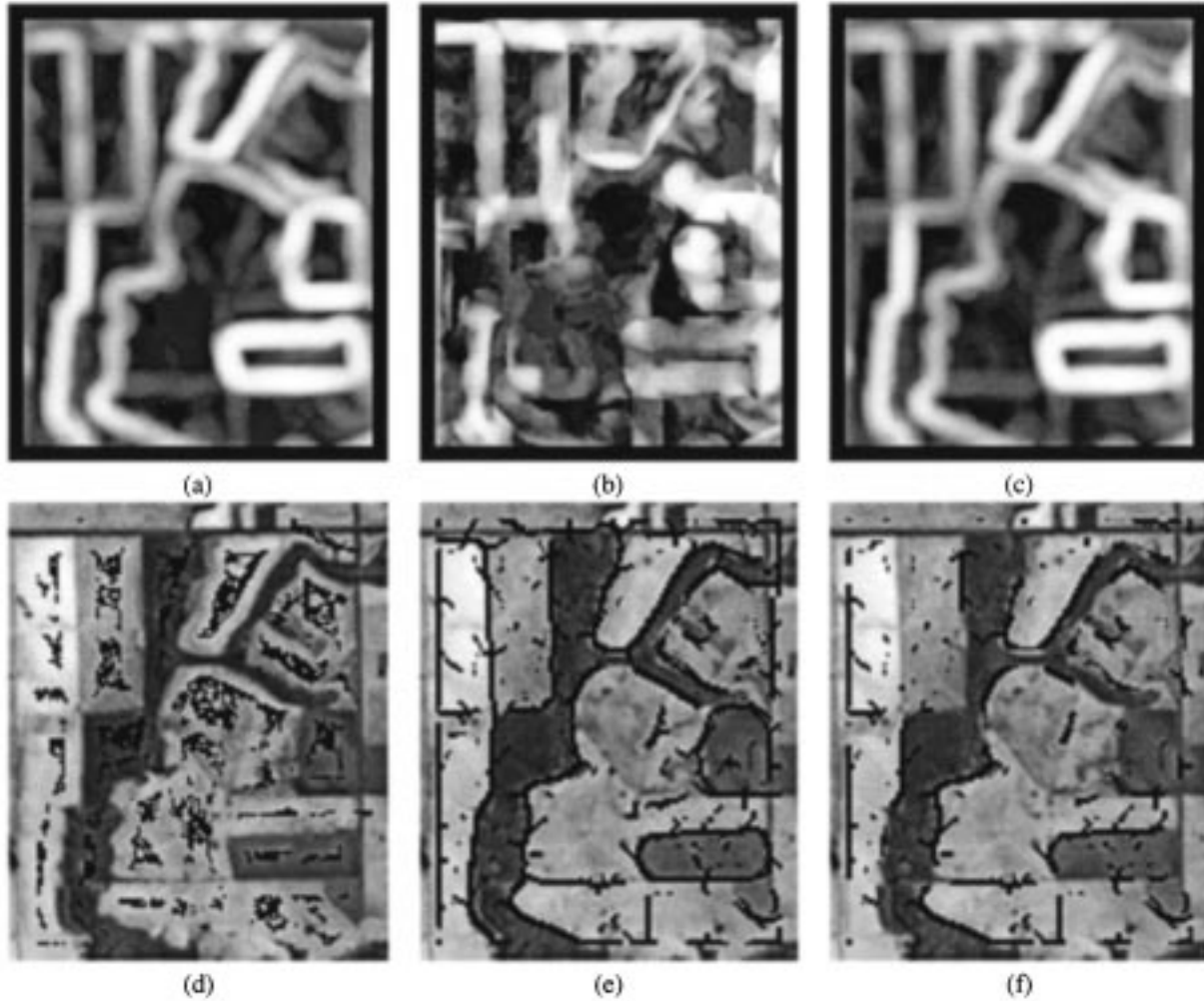


Fig. 8. (a) λ_1 , (b) λ_2 , and (c) Λ eigenvalue images of Fig. 1. (d) Field markers obtained from (c) as detailed in (3). (e) Ridge points (a). (f) Ridge points satisfying constraints (4) and (5). Fig. 8(d)–(f) have been overlaid on Fig. 1.

(Section II-A), partial field boundaries are obtained from the images λ_1 and λ_2 . Candidate boundary pixels, separating only two fields, satisfy the constraints

$$\gamma_{5 \times 5} \phi_{5 \times 5} \gamma_{5 \times 5}(\lambda_1) - \gamma_{13} \gamma_{5 \times 5} \phi_{5 \times 5} \gamma_{5 \times 5}(\lambda_1) \geq 1.0 \quad (4)$$

$$\gamma_{5 \times 5} \phi_{5 \times 5} \gamma_{5 \times 5}(\lambda_2) - \gamma_{13} \gamma_{5 \times 5} \phi_{5 \times 5} \gamma_{5 \times 5}(\lambda_2) < 1.0 \quad (5)$$

in the morphologically filtered versions of λ_1 and λ_2 . The thresholds in the above equations have been set to 1.0 instead of zero, to reduce the detection of spurious boundary pixels due to within-field variation. Boundary pixels must be those where the estimated between-field variation is greater than the within-field variation, i.e., where the ratio of them is greater than 1.0.

To locate the boundary positions more accurately, we also add the constraint that the 3×3 top hat of the original image λ_1 must be nonzero at partial boundaries, i.e.

$$\lambda_1 - \gamma_{3 \times 3}(\lambda_1) > 0. \quad (6)$$

This limits the width of partial boundaries to a maximum of 2 pixels which is reasonable at this scale of imaging. A

3×3 structuring element is used in (6) so the ridge points will form lines no wider than 2 pixels and localize the partial field boundaries more accurately. The partial field boundaries obtained from the first and second eigenvalue images are shown in Fig. 8(e).

Further morphological filtering of these partial field boundaries may be required to prune very short partial boundaries, as they probably will not be constructive additions to the segmentation process. If we choose not to obtain partial field boundaries from the canonical eigenvalue image, the individual eigenvalues do not need to be calculated. Only the image containing the sum of the local eigenvalues (i.e., the trace of the local matrix $S_B S_W^{-1}$) needs to be calculated to determine the internal field markers.

Partial field boundaries can also be obtained from the binary image of atypical pixels (AP) computed in Section II-D. Because some of these pixels correspond to roads or trees following field boundaries. Boundaries at least 11 pixels long and no wider than two pixels are extracted from A by

$$\gamma_{11}(\text{AP} - \gamma_{3 \times 3}(\text{AP})) > 0. \quad (7)$$

Fig. 5(b) shows the results of filtering the images of Fig. 5(a).

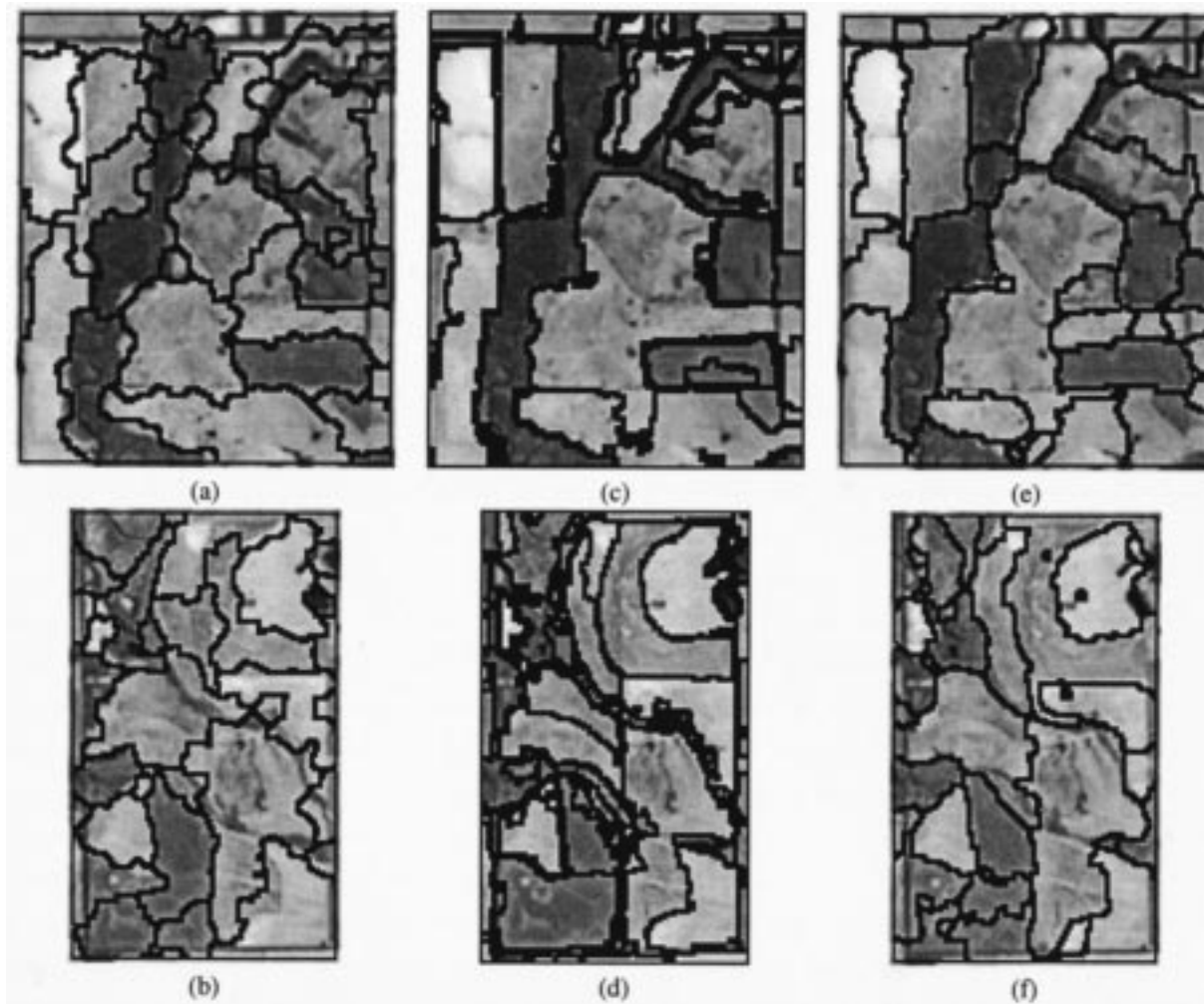


Fig. 9. Result of segmenting the two sample Landsat images shown in Fig. 7 using Lee's segmentation procedure [Fig. 9(a) and (b)], Tilton's segmentation procedure [Fig. 9(c) and (d)] and the CGRG segmentation procedure [Fig. 9(e) and (f)].

F. Seeded Region Growing

Finally the image is segmented by SRG using the set of internal field markers as seeds. Partial boundary pixels, obtained in Section II-E, are excluded from the set of field markers to constrain the results of the SRG. The result of SRG on Fig. 1 using the internal field markers of Fig. 8(d) as seeds is shown in Fig. 9(e). For comparison purposes discussed in the next section, we do not include partial boundary pixels in the final set of segmented field boundaries. Normally, however, partial boundaries and known roads and property boundaries would be included to make use of all the available information. The seeds then grow using the pixel values in the multispectral estimated image B free of any locally atypical regions inside fields. Spectral "closeness" is defined by the Mahalanobis distance metric [see (8.13) of [15]] based on the global, robust estimate of S_W .

For those areas of Western Australia where the position of some roads and properties boundaries are known, the SRG algorithm can incorporate these constraints in the final segmentation, along with the partial field boundaries. However, such information is not always available and so is not used in this paper.

III. COMPARISON OF SEGMENTATION RESULTS

We now compare the performance of the CGRG procedure with the performance of Lee's and Tilton's procedures (see Section II-B). Each procedure is applied to a set of Landsat TM images and the discrepancies between the results of each procedure and the ideal segmentation for each image are measured. Unfortunately, the ideal segmentations had to be produced manually by a remote sensing expert [6], so we were limited to working with a small set of eight sample Landsat TM images. (Two ideal segmentations are shown in Fig. 7.)

The measures we use to quantify the discrepancy of segmentation procedure results and the ideal are Levine and Nazif's M_2 metric [24]; and Baddeley's $\Delta^2(t_0 = 5)$ metric [25]. M_2 is a normalized metric that measures the degree of over- and under-segmentation of regions in a segmented image compared with the ideal. Δ^2 is a normalized metric that measures the discrepancy between the position of the region boundaries in a segmented image compared with the ideal. The better the segmentation achieved, the closer both of these discrepancy measures are to zero. We also compare the number of regions

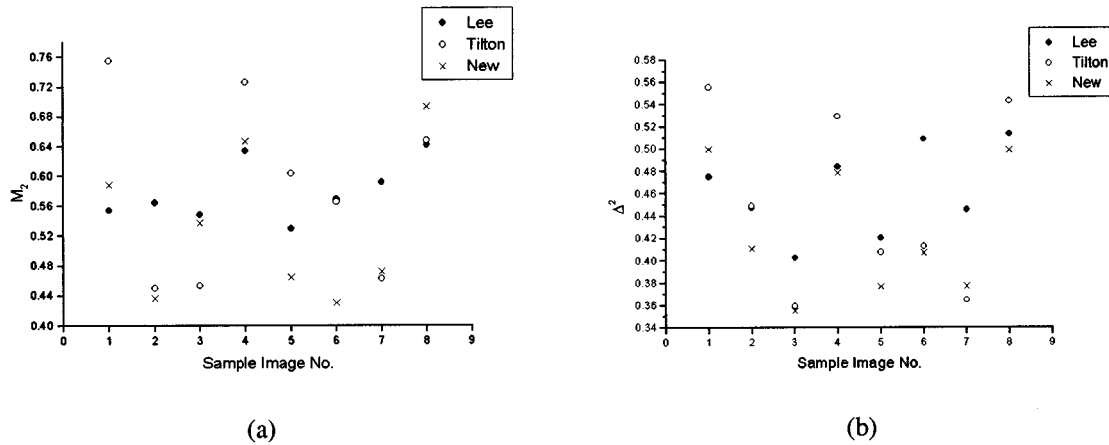


Fig. 10. Comparison of segmentation results using Lee's method, Tilton's method and the new CGRG method: (a) plot of M_2 values and (b) plot of Δ^2 values.

TABLE I
COMPARISON OF THE NUMBER OF REGIONS IN THE SAMPLE IMAGES SEGMENTED USING THE THREE DIFFERENT PROCEDURES

Sample Image No.	Ideal	Lee's Segmentation	Tilton's Segmentation	CGRG Segmentation
1	18	52	76	36
2	18	43	59	25
3	21	70	59	33
4	16	51	70	38
5	32	64	95	36
6	14	26	15	14
7	30	72	75	39
8	25	68	155	58

produced with the ideal segmentation. This is a simple but informative measure.

To ensure a fair comparison between the segmentation methods, the CGRG procedure was applied as described in Section II, but for one minor change. Every pixel was assigned to a region by the SRG algorithm in the CGRG procedure, and no partial boundary information was included, since Tilton's and Lee's procedures assigned no pixels to region boundaries. The positions of field boundaries in all segmentation results were later determined using a 3×3 Laplacian edge detector.

Lee's segmentation procedure was applied using evenly tiled seeds, 3×3 pixels in size, spaced nine pixels apart vertically and horizontally. Since the smallest field of interest is 11 pixels wide, this ensured that every field would contain at least one seed. SRG was then performed in the same way as for the CGRG procedure (see Section II-F), except for a different set of initial seeds. The set of images, B , was transformed by the global canonical transform and an arbitrary amount of independent, random Gaussian noise (variance = 9.0) was added to each band of the multispectral image to satisfy the image model assumed by the MDL criterion for region merging. The results of the MDL region-merging do vary with the amount of within-field variation in the image. If different levels of random noise had been added to image B , the MDL region-merging results may have changed. However, it is not clear how much random noise should be added to obtain optimal results.

Tilton's segmentation of each sample image was based on the between-field variation, B , associated with each sample image. Dr. Tilton kindly processed the sample images using his segmentation method with the L_1 distance metric [26] to define

"spectral closeness." Connected regions in Tilton's results did not always have a unique label as his procedure allows non-connected regions to merge. We relabeled the regions in each segmentation result because unique labels are required to allow computation of the degree of over-segmentation and under-segmentation. The hierarchical segmentation level that minimized M_2 was chosen to compare to the results of the other segmentation methods. This level varied between the sixth and the 16th level (out of approximately 20 segmentation levels) for different sample images in the set.

Before calculating the discrepancy of any of the segmentation results from the ideal, the results were cropped by a margin of five pixels to allow for a moving window 11 pixels in diameter being used to calculate the local eigenvalues in the CGRG procedure.

A. Performance Results

The performance of the three segmentation procedures can be compared by considering Figs. 9 and 10. Fig. 9 shows the segmentation results for the two images manually segmented in Fig. 7. The pixels marked in black are at region borders. Plots of the values of the segmentation evaluation measures M_2 and Δ^2 for the eight sample Landsat TM images are given in Fig. 10(a) and (b), respectively. According to M_2 , the CGRG procedure gives more accurate segmentations than do the procedures of either Lee or Tilton for five out of the eight sample images. Fig. 10 suggests that Lee's procedure oversegments because it produces too many small regions. This is confirmed by the results shown in Table I, which lists the number of spatially-

connected regions produced by each segmentation procedure. [Sample image number 3 corresponds to the image segmented in Fig. 9(a), (c), and (e), and sample image number 7 corresponds to the image segmented in Fig. 9(b), (d), and (f).] Both Lee's procedure and Tilton's procedure segmented each sample image into significantly more regions than the CGRG procedure and more than the ideal. According to Δ^2 , the CGRG procedure is more accurate than either of the other methods for seven of the eight sample images.

It is thought that the size, shape and position of the internal field markers must help steer region growth in the SRG algorithm to a greater extent than evenly-tiled 3×3 seeds, to lead to improved segmentation results. This improvement comes at a cost, however. More computation is required to determine the internal field markers from the local canonical eigenvalue image. Currently it takes 14.5 s to calculate the local canonical eigenvalue image of a 150×150 6-band satellite image, on a 4-processor Sun Sparc 1000, running Solaris 2.5.1. Morphological filtering of the eigenvalue images takes less than 1% of the transformation time. The computational load can be reduced by not computing individual eigenvalue images λ_1 and λ_2 to extract partial field boundaries from, but just calculating the sum of the local eigenvalues, Λ , to obtain reliable internal field markers using (3).

Table I indicates that Tilton's procedure oversegments the sample satellite image to a much greater extent than the CGRG procedure. This is because Tilton's procedure does not impose a minimum field-width criteria so that many very small regions are included in the final segmentation. Fig. 10(a) and (b) show that the median and average Δ^2 and M_2 are both higher for Tilton's segmentation results than for the CGRG procedure, as well as the results being more highly variable. This variability reflects the adaptive nature of the homogeneity criterion used in Tilton's procedure.

An advantage of the CGRG procedure over Tilton's procedure is that the CGRG procedure relies on one main parameter, the width of the minimum region of interest in the image. Although Tilton's hierarchical segmentation method does not require this parameter to be specified, the best level of segmentation must eventually be selected from the hierarchy, and this segmentation level may still contain many regions too small to be fields. It is more tractable to select a minimum field size than to select an optimal level of segmentation in Tilton's procedure.

IV. CONCLUSION

A new method of segmenting multispectral images into homogeneous regions, the CGRG procedure, has been proposed and applied to segment multispectral Landsat TM images of agricultural areas of Western Australia into field units. The CGRG procedure involves the calculation of a local canonical eigenvalue image using a moving window. The canonical eigenvalue image allows for different amounts of between-field and within-field variation across the image bands and combines the multispectral information to obtain better discrimination between adjacent field units. For the majority of sample images that were segmented, the CGRG results are usually more accurate in terms of over- and under-segmentation and

field-boundary position than results given by either Lee's segmentation procedure or Tilton's procedure. Further evaluation is still required to show whether the CGRG procedure will give consistently better results over a large dataset of Landsat TM images.

ACKNOWLEDGMENT

The authors would like to thank Dr. J. Tilton, NASA's Goddard Space Flight Centre, MD, for providing segmentations of the sample Landsat TM images, and S. Furby and Dr. N. Campbell, Division of Mathematical and Information Sciences, CSIRO Australia, for providing the reference segmentations of the sample Landsat TM images and answering queries about robust covariance estimation.

REFERENCES

- [1] T. C. M. Lee, "Some models and methods in image segmentation," Ph.D. thesis, Macquarie Univ., Sydney, NSW, Australia, 1997.
- [2] J. C. Tilton, "Image segmentation by region growing and spectral clustering with a natural convergence criterion," in *Proc. 1998 Int. Geosci. Remote Sensing Symp. (IGARSS'98)*, Seattle, WA, July 6–10, 1998.
- [3] G. A. Wheaton, J. F. Wallace, D. J. McFarlane, S. L. Furby, N. A. Campbell, and P. Caccetta, "Mapping and monitoring salt-affected land in Western Australia," in *Proc. Res. Technol. Conf.*, Melbourne, Victoria, Australia, 1994, pp. 531–543.
- [4] R. L. Kettig and D. A. Landgrebe, "Classification of multispectral image data by extraction and classification of homogeneous objects," *IEEE Trans. Geos. Electron.*, vol. GE-14, pp. 19–26, Jan. 1976.
- [5] L. L. F. Janssen and M. Molenaar, "Terrain objects, their dynamics and their monitoring by the integration of GIS and remote sensing," *IEEE Trans. Geos. Remote Sensing*, vol. 33, pp. 749–758, Mar. 1995.
- [6] private communication Personal correspondence with S. Furby, a remote sensing expert from the Division of Mathematical and Information Sciences.
- [7] V. C. Barbosa, R. J. Machado, and F. dos D. Liporace, "A neural system for deforestation monitoring on landsat images of the amazon region," *Int. J. Approximate Reasoning*, vol. 11, pp. 321–359, 1994.
- [8] J. Le Moigne and J. C. Tilton, "Refining image segmentation by integration of edge and region data," *IEEE Trans. Geosci. Remote Sensing*, vol. 33, pp. 605–615, Mar. 1995.
- [9] C. Y. Ji, "Delineating agricultural field boundaries from TM imagery using dyadic wavelet transforms," *ISPRS J. Photogramm. Remote Sensing*, vol. 51, pp. 268–283, 1996.
- [10] P. Soille, "Morphological partitioning of multispectral images," *J. Electron. Imaging*, vol. 5, no. 3, pp. 252–265, 1996.
- [11] A. Lobo, "Image segmentation and discriminant analysis for the identification of land cover units in ecology," *IEEE Trans. Geosci. Remote Sensing*, vol. 35, pp. 1136–1145, May 1997.
- [12] R. Adams and L. Bischof, "Seeded region growing," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 16, pp. 641–647, June 1994.
- [13] J. Rissanen, *Stochastic Complexity in Statistical Inquiry*, Singapore: World Scientific, 1989.
- [14] M. Cebrián, M. P. Pérez-Luque, and G. Cisneros, "Edge detection alternatives for multispectral remote sensing images," in *Proc. 8th Scandinavian Conf. Image Anal.*, vol. 2, Tromsø, Norway, 1993, pp. 1047–1054.
- [15] J. A. Richards, *Remote Sensing Digital Image Analysis: An Introduction*. Berlin, Germany: Springer-Verlag, 1986.
- [16] A. A. Green, M. Berman, P. Switzer, and M. D. Craig, "A transformation for ordering multispectral data in terms of image quality with implications for noise removal," *IEEE Trans. Geosci. Remote Sensing*, vol. 26, pp. 65–74, Jan. 1988.
- [17] J. B. Lee, A. S. Woodyatt, and M. Berman, "Enhancement of high spectral resolution remote-sensing data by a noise-adjusted principal components transform," *IEEE Trans. Geosci. Remote Sensing*, vol. 28, pp. 295–304, Mar. 1990.
- [18] C. J. Evans, "Segmenting Multivariate Images," Ph.D. thesis, Monash Univ., Melbourne, Victoria, Australia, 1999.
- [19] K. Fukunaga, *Introduction to Statistical Pattern Recognition*. London, U.K.: Academic, 1972.

- [20] J. Serra, Ed., *Image Analysis and Mathematical Morphology, Vol. 2: Theoretical Advances*. London, U.K.: Academic, 1988.
 - [21] C. Evans, R. Jones, and I. Svalbe, "Field segmentation in multispectral satellite images," in *Proc. ISPRS Commission III Symp.*, Columbus, OH, 1998, pp. 213–217.
 - [22] N. A. Campbell, H. P. Lopuhaä, and P. J. Rousseeuw, "On the calculation of a robust S -estimator of a covariance matrix," *Stat. Med.*, vol. 17, no. 23, pp. 2685–2695, 1998.
 - [23] R. A. Johnson and G. K. Bhattacharyya, *Statistics: Principles and Methods*, 2nd ed. New York: Wiley, 1992.
 - [24] M. D. Levine and A. Nazif, "An experimental rule based system for testing low level segmentation strategies," in *Multicomputers and Image Processing: Algorithms and Programs*, K. Preston and L. Uhr, Eds. New York: Academic, 1982, pp. 149–160.
 - [25] A. J. Baddeley, "Errors in binary images and an L^p version of the Hausdorff metric," *Nieuw Archief voor Wiskunde*, vol. 10, pp. 157–183, 1992.
 - [26] M. Barni, V. Cappellini, and A. Mecocci, "The use of different metrics in vector median filtering: Application to fine arts and paintings," in *Signal Processing VI: Theories and Applications*, J. Vandewalle, R. Boite, M. Moonen, and A. Oosterlinck, Eds. New York: Elsevier, 1992, pp. 1485–1488.
- Carolyn Evans**, photograph and biography not available at the time of publication.
- Ronald Jones**, photograph and biography not available at the time of publication.
- Imants Svalbe**, photograph and biography not available at the time of publication.
- Mark Berman**, photograph and biography not available at the time of publication.