to appear in European Symp. on Satellite Remote Sensing, 1994

Parcel-based change detection

Paul L. Rosin

Institute for Remote Sensing Applications
Joint Research Centre, I-21020 Ispra (VA), Italy
email: paul.rosin@jrc.it

ABSTRACT

Various methods for automatic change detection in multi-temporal LANDSAT-TM images are described. In contrast to most previous work in change detection, which has operated at a pixel level, we operate at a parcel level (within a minimum size of 25 Ha). This makes it easier to employ structural measures (e.g. based on edges, corners, and texture) as well as correlation methods since these approaches cannot be calculated at each pixel independently. A neural network is trained to combine the different change measures in an appropriate manner.

1 INTRODUCTION

One of the major applications of remotely-sensed data is change detection, due to the repetitive and frequent coverage, and low cost compared to traditional methods such as ground survey and analysis of aerial photographs. Change detection is used to aid planning and management of land resources,[2] and has been employed in many applications such as the monitoring of desertification,[20] forest defoliation,[15] and urban development.[8,12,26]

This paper describes methods for change detection to enable the automation (or at least semi-automation) of the revision of land cover maps that have been previously generated manually by photo-interpreters for the CORINE programme of the European Commission.[25] Such a large scale project (involving approximately 2.25 Mha) is slow and costly. Hence automated updating of the maps would be beneficial for both economic and efficiency reasons. Furthermore, manual revision is prone to subjectivity which can lead to widely varying estimates of land cover change statistics. Automation provides a means for reproducability and consistency of results over the complete data set.

Although there has been considerable research into automatic techniques for change detection they have tended to be at the individual pixel level.[23] Since the CORINE land cover maps only contain land cover parcels with a minimum size of 25 ha we wish to perform change detection at the parcel level. There is considerable difference between parcel and pixel based change detection. The higher spatial and spectral resolution of Landsat TM data compared to MSS data increases class heterogeneity which is problematic both for per–pixel classifiers[13] and change detection. The variations occurring at individual pixels will be smoothed out over the parcel, potentially making parcel analysis more reliable. On the other hand, parcel analysis is less sensitive to local changes much smaller than the parcel size. For instance, depending on the change measure used, if half the parcel changes significantly this can produce a similar overall change value as if the total parcel had changed slightly. Therefore it may be necessary to keep some account of the variation in the change measure within the parcel.

2 DIFFICULTIES IN CHANGE DETECTION

When analysing image sequences there are many factors that can produce temporal variations that can be potentially confused with significant changes (due to land cover change). Examples of sources of spurious change are differences in atmospheric conditions, sun angle, season, ground moisture, and sensor calibration, and the misregistration of the images. Many methods have been developed for change detection, some of which attempt to overcome the above problems. However, it is not clear which is the best since different comparisons produce conflicting results.[23] Indeed, the most common method of change detection is still the simple differencing of individual pixel spectral values, which in many cases works almost as well as more complex methods.[23] However, one of the problems with directly differencing spectral values is that the resulting difference does not take into account the location of the spectral values in feature space. Thus a large change in spectral value may not signify land cover change if both the new and old spectral values are within range of the class's spectral signature. Methods that first transform the image such as vegetation indexes or texture measures overcome this problem to some extent, but are liable to their own problems. Alternatively, an obvious approach that directly takes the location of class signatures in feature space into account would be to first classify both images and then flag as areas of change those pixels whose classification changes between dates. In practise this approach performs relatively poorly since it is sensitive to errors in the classification process - an error in the classification of either image is likely to produce an error in the estimated change.[17,23]

Rather than attempt to find a general change measure that is universally applicable, it is more practical to combine several different change measures so that the limitations of each are offset by the others' strengths. For instance, when attempting to detect rural to urban land cover changes, it was found that combining spectral and textural information was beneficial.[8,12] Each measure in itself was ambiguous; for instance some fields exhibited a similar coarse texture to urban land cover, while direct classification of urban land cover proved problematic due to its spectral homogeneity. Combining both types of information produced a higher accuracy than either one on its own. A similar improvement in detecting change in vegetation cover by combining albedo and texture was noted by Frank.[7]

2.1 Image Registration

In his survey on change detection Singh[23] noted that one of the limitations of most current techniques for change detection is that they require precise registration between the images being compared. Since it is often difficult to obtain the necessary accuracy in registration, the change detection methods are liable to produce errors, particularly in areas of rapid change such as edges.

We solve this problem by allowing for some pre-specified amount of spatial variation between the images. The images are searched within this neighbourhood for the best match. This variation between the images A and B can be constrained to an X and Y translation, so that the average pixel difference D_k within region R_k is

$$D_k = \frac{1}{|R_k|} \min_{(i,j)\in W} \sum_{(x,y)\in R_k} |A_{x,y} - B_{x+i,y+j}|$$

where $A_{x,y}$ is the pixel value of image A at (x,y), $|R_k|$ is the size (number of pixels) of R_k, and W is the neighbourhood of allowable spatial variation (e.g. $W = \{0, \pm 1\}$ specifies a 3×3 window). Alternatively, since other more general deformations may be present between images, each pixel is allowed to move independently of its neighbouring pixels. Although the resulting estimated spatial offsets are unlikely to be realistic, this method is only intended as a simple approximation to solving the image misregistration. With this alternative definition the average pixel difference within R_k becomes

$$D_k = \frac{1}{|R_k|} \sum_{(x,y)\in R_k} \min_{(i,j)\in W} |A_{x,y} - B_{x+i,y+j}|$$

3 CHANGE MEASURES

In section 2 we discussed some of the limitations of measures for change detection. Instead of restricting attention to a single measure we shall use a combination of several. Some of these are standard measures often used in change detection (e.g. difference in vegetation index) whereas others are more novel (e.g. average mutual information).

It is noteworthy that parcels enable various global measurements and comparison techniques to be applied that are not available or are inappropriate at a pixel level. For instance, whereas we will use the entropy of a parcel as a means of assessing change, the entropy of a single pixel is fixed and therefore provides no useful information. Likewise, the rank correlation between two single pixels is uninformative. Thus we can divide the types of change measures into two categories: 1) global, such as the change in parcel entropy, and 2) local differences of pixel values. The global measures are only applicable to parcels while the local ones are calculated at each pixel independently and then some function of these local values (e.g. average) is taken over the parcel.

Although many of the following measures could be applied to the full multi-band Landsat TM images, for computational convenience we have restricted most comparisons to the first principal component.

3.1 Change in entropy

One of the texture measures calculated by Haralick[9] using his gray-level co-occurence method is entropy. Entropy quantifies the amount of information or uncertainty present in a data set. It can also be considered as reflecting the complexity of the data and therefore should distinguish between textured and untextured areas. The gray-level co-occurence method requires a pre-set square window as well as (possibly multiple) pre-set offsets and orientations between pixel pairs. Instead we shall calculate entropy directly from the pixel values within the region so that the measure uses exactly those values within the region and avoids unnecessary parameters. The entropy of region R_k in image A is calculated as

$$H_A(R_k) = \sum_{i=1}^{N} p_i^A \log\left(p_i^A\right)$$

where there are N distinct pixel values within R_k in image A, and p_i^A is the proportion of the i'th pixel value in R_k in image A. The absolute change in entropy in R_k between images A and B is then:

$$\Delta E(R_k) = |H_A(R_k) - H_B(R_k)|$$

3.2 Correlation

A common method for assessing the similarity of two data sets is to determine the correlation between them. There are many different methods for calculating correlation coefficients. These typically vary in terms of the type of correlation being tested (e.g. linear) as well as their robustness in the presense of noise. We will try several of them. The first is the cross-correlation coefficient which is commonly used in image processing for template matching, and is defined a:

$$r(R_k)_{m,n} = \frac{\sum_{(x,y)\in R_k} A_{x,y} B_{x+m,y+n}}{\sqrt{\sum_{(x,y)\in R_k} A_{x,y}^2 \sum_{(x,y)\in R_k} B_{x+m,y+n}^2}}$$

where (m,n) is the (X,Y) offset between the two images. Thus, a rigid translation of the regions of up to $\pm w$ pixels can allowed by choosing the offset which produces the best correlation:

$$r(R_k)_{\max} = \max\left(\sum_{m=-w}^{w}\sum_{n=-w}^{w} r(R_k)_{m,n}\right)$$

An alternative method for calculating correlation is Spearman's rank correlation. It does not assume any *a priori* parametric model of the data or noise and compares the ranks of corresponding values relative to each data set rather than their actual values. This can greatly improve the robustness of the technique since the correlation coefficient is independent of offsets and scalings applied uniformly to either (or both) data set, as well as other more complex transformations that preserve the ordering of the data sets. The consequent disadvantage of such non-parametric methods is their low efficiency, i.e. the estimated correlation value has a relatively high variance. The Spearman's rank correlation is calculated as

$$1 - \frac{6(D^2 + T)}{n^3 - n}$$

where $n = |R_k|$ is the size of the region, D^2 is the summed squared difference in ranks, and T is a correction factor to take tied ranks into account.

$$D^2 = \sum_{(x,y)\in R_k} (G_{x,y} - H_{x,y})^2$$

where $G_{x,y}$ and $H_{x,y}$ are the ranks of the pixel values at (x,y) in R_k in images A and B respectively, and

$$T = \sum_j \left(\frac{j^3 - j}{12}\right) t_j$$

where t_j is the number of ties in the ranking involving j values.

The standard version of the Spearman rank correlation assumes correct correspondence between the data sets, i.e. perfect image registration. We can augment the correlation method to include either of the two corrections for misregistration described in section 2.1. Thus, when allowing a rigid shift between images the rank correlation is calculated as

$$D^2 = \min_{(i,j)\in W}\left(\sum_{(x,y)\in R_k} (G_{x,y} - H_{x+i,y+j})^2\right)$$

and the unconstrained distortion as

$$D^2 = \sum_{(x,y)\in R_k}\left(\min_{(i,j)\in W} (G_{x,y} - H_{x+i,y+j})^2\right)$$

3.3 Average Mutual Information

Although the Spearman rank correlation is unaffected by many complex transformations it still expects the ranks of corresponding values to be the same. Here we suggest a method for comparing two data sets that relaxes even that expectation. The average mutual information (AMI) index quantifies the similarity between two data sets based on the degree that one set of values predicts the other,[1,6] so that change is inversely proportional to the AMI.

$$AMI = \sum_{i=1}^{M}\sum_{j=1}^{N} p(b_i, a_j) \log\left[\frac{p(b_i \mid a_j)}{p(b_i)}\right]$$

where there are $\{a_i; i = 1 \ldots N\}$ and $\{b_i; i = 1 \ldots M\}$ distinct pixel values in R_k in images A and B respectively; $p(a_i)$ and $p(b_i)$ are the proportion of pixel value a_i in image A, and b_i in image B. Thus, if the spectral values within an area consistently change over time, this will produce a high AMI value, equivalent to no change. An example of such an event in which it might be desirable to consider spectral differences as suggesting non change is a field imaged at different seasons. Spectral differences would arise from phenological changes; the extreme case would be to compare a crop field before and after harvesting.

However, a limitation of the technique is that it is based on nominal values. Since the spectral values are continuous (although quantised) this means that AMI ignores much of the available information. For instance, even if there are many mappings between the images of a pair of values such as $5 \rightarrow 138$, if a similar mapping like $5 \rightarrow 137$ was infrequent it would register as a substantial change.

3.4 Structural Information

A limitation of most classification methods is that they analyse each pixel independently based on its spectral values, without incorporating any spatial information. A similar criticism can be applied to most methods for change detection with the exception of those that measure differences in textural properties.[8,12] Wang and Howarth suggested using various structural measures such as edge and junction density for the classification of remotely sensed images.[24] We take a related approach and incorporate some structural information into the change detection process.

This is done by first detecting spatial features such as edges and corners, and then propagating the influence of these features over the whole image using the multi-scale salience distance transform (MSSDT).[21] The MSSDT is an extension of the standard distance transform (DT)[3] in which each pixel in the image is assigned the (approximate) distance to the nearest feature in the image. The DT has several weaknesses which the MSSDT rectifies. First, the DT requires a binary feature map. However, when dealing with features such as edges and corners it is generally difficult to automatically choose a good threshold value. Second, an appropriate scale for feature detection must be known beforehand. The salience DT overcomes the first restriction by operating directly on the unthresholded feature map. It weights the distance to the feature by the strength of the feature, so that a short distance to a weak feature may be considered less significant than a longer distance to a strong edge. This effectively reduces the influence of noise without requiring an arbitrary cut-off. The problem of scale is overcome by performing feature detection at several scales. The salience DT is applied to each feature map, and the results at multiple scales are then combined by addition.

Efficient algorithms are available for calculating the DT which only use local operations. Similar algorithms are also available for the salience DT and MSSDT. The iterative parallel version is described below, and the serial ("chamfer") algorithm operates in a similar fashion. The initial distance map $d^0_{x,y}$ is created containing values of 0 at feature pixels and ∞ (or an appropriate large value) elsewhere. The initial feature map $m^0_{x,y}$ contains the magnitude of the feature pixels. At each iteration, at every pixel, weighted distances within a 3×3 neighbourhood are calculated using the mask $w_{a,b}$ centred on the pixel, and the minimum distance is propagated to the centre pixel. The most common weights are 3 for horizontal and vertical weights, 4 for diagonal weights, and 0 for the central value. In the salience DT this weighted distance is converted into a salience value by dividing by the corresponding feature magnitude. This process of propagating salience values from feature pixels can be described as

$$\begin{aligned} a', b' &= \arg\left(\min_{(a,b)=-1\ldots1}\left(\frac{d^{i-1}_{x+a,y+b} + w_{a,b}}{m^{i-1}_{x+a,y+b}}\right)\right) \\ d^i_{x,y} &= d^{i-1}_{x+a',y+b'} \\ m^i_{x,y} &= m^{i-1}_{x+a',y+b'} \end{aligned}$$

where a', b' is the location in the window producing the minimum salience, $m^i_{x,y}$ and $d^i_{x,y}$ are the edge magnitude

and distance values propagated to (x, y) at iteration i, and the final salience map at iteration f (after which additional iterations make no further changes to $m^i_{x,y}$ and $d^i_{x,y}$) is $s_{x,y} = \frac{d^f_{x,y}}{m^f_{x,y}}$.

The result of the salience DT is that every pixel in the image is assigned a salience value. This has the advantage over methods for quantifying structure based on averaging feature values within a pre-set window[8,10,24] that it does not require any parameters. Moreover, propagation of values is preferable to blurring since it does not distort fine features. Structural measures are calculated for each region by averaging the salience values of the pixels within the region, and change is estimated by the difference in the structure measures between images.

We apply the MSSDT to features maps of the image consisting of edges and corners. Edges are generated using the Canny edge detector.[4] The first derivative is calculated for each image band independently, and then the responses of all the bands are combined by performing a maximum operation at each pixel. Finally, non-maximal suppression is performed on the combined edge response image. Corners are detected using the Kitchen/Rosenfeld operator.[14] Although a multi-band version could have been implemented in a similar method to the Canny edge detector, for simplicity it was applied to the 1st principal component instead. Both edge and corner detection was performed at four scales of Gaussian smoothing with $\sigma = \{1, 2, 4, 8\}$.

3.5 Pixel Differencing

Finally, we also include standard pixel spectral differencing and vegetation index differencing. There have been many different vegetation indices proposed in the literature. However, it has been shown that most of them are highly correlated with each other.[19] We use the Normalised Vegetation Index (NVI) which is defined for each pixel in a Landsat TM image as

$$NVI = \frac{\text{band } 4 - \text{band } 2}{\text{band } 4 + \text{band } 2}$$

4 COMBINING THE CHANGE MEASURES

Having defined the set of change measures in the previous section (listed in table 1) it now has to be determined how they can be combined in the best manner. To perform this task we use a multi-layer perceptron neural network.[22] A set of training data containing examples of change measures derived from areas of change and no change is provided, and the network learns the best combination of the measures that discriminates the change using the back-propagation algorithm. This is a similar approach to the direct multidate classification methods except that they generally operate on the spectral vales directly using, for instance, a maximum likelihood classifier,[17] or apply cluster analysis to discriminate between change and no change areas in feature space.[16]

TYPE OF CHANGE MEASURE	METHOD
structural information	entropy difference
	MSSDT edge difference
	MSSDT corner difference
correlation	cross correlation
	Spearman rank correlation
	average mutual information
pixel differencing	spectral difference
	vegetation index difference

Table 1. Change measures

We previously mentioned that some changes in the image may be due to external factors such as sensor calibration, atmospheric conditions, etc. The spectral variations associated with these factors should be ignored, although their magnitude may be greater than that of real (i.e. significant) change. Most previous work in change detection has been restricted to a binary classification of change or no change.[23] There are also other types of changes, more directly associated with land cover change, that it may also be desirable to ignore for certain applications. For instance, as discussed previously, certain changes in vegetation may be unimportant. On the other hand, it may be of interest to specify more precisely some particular types of change, so that for instance, forest to cleared ground can be distinguished from rural to urban, etc. Using simple pixel differencing it would be difficult to distinguish between different types of change. However, with the large feature vector composed of the change measures described in section 3 it is now feasible. In the extreme case, n classes are detected in each image, and all combinations of possible land cover change are calculated, forming an $n \times n$ change matrix.[18] However, a more restricted set of interesting types of change is probably sufficient (and more reliably detected).[17] In training the network we have distinguished the following changes:

- Vegetation $\rightarrow$ vegetation. This class includes areas of vegetation with no change as well all types of unimportant vegetation change that are not to be distinguished.
- Urban $\rightarrow$ urban (1). This class includes urban areas containing no (significant) change.
- Urban $\rightarrow$ urban (2). This class covers significant changes (development) in urban areas.
- Vegetation $\rightarrow$ urban.
- Urban $\rightarrow$ vegetation.

5 RESULTS

Although all the techniques described above are operational, only preliminary results can be described. This is because we currently have little multi-temporal ground truth. The parcel based approach requires substantial areas of image change for training since each parcel becomes just a single training element. Moreover, the need for large amounts of training data is compounded by the many change measures being used, since increasing the dimensionality of the feature space increases the amount of training data required.[11] Therefore it is difficult both to train the network with any confidence in its generalisation abilities, and to test the system.

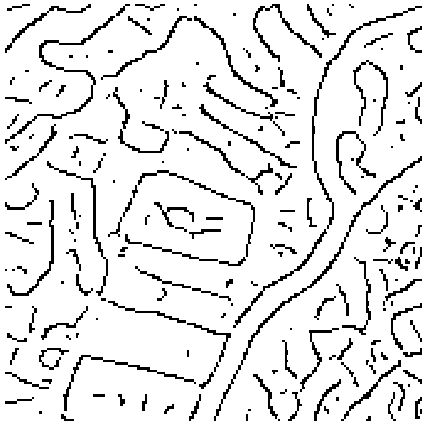
The application of the structural salience measures are demonstrated on a 200×200 portion of the image shown in figure 1. The first principle component is shown in figure 1a, and the edges detected at two scales ($\sigma = 1, 4$) are shown in figure 1b&c. It can be seen that there is considerable noise at the fine scale. Although this is suppressed at the coarser scale, fine detail has also been removed, and boundaries have been distorted. The MSSDT calculated at the four scales is shown in figure 1d (log mapped for visualisation). The main features are emphasised while a reasonable spatial fidelity has been maintained. Figure 1e shows the log mapped MSSDT of the corners which overcomes the problems of noise and scale in a similar manner.

6 CONCLUSIONS

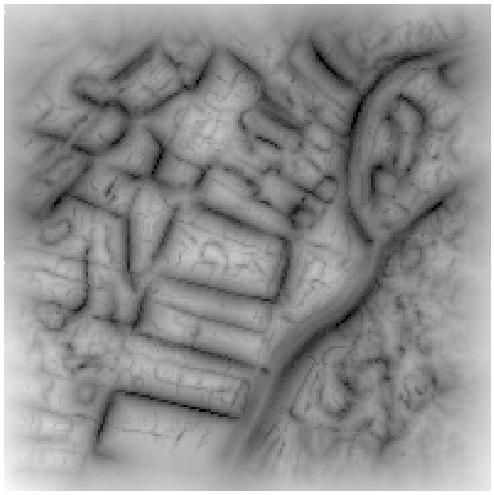
We have suggested using a large number of measures to increase the reliability and discriminative power of change detection procedures. This would enable a variety of classes of change to be identified rather than just the usual binary change/no change classification provided. A number of new change detection measures and comparison techniques have been proposed. The former incorporate spatial information in the form of propagated structural measures and region complexity measured by its entropy. The various comparison techniques differ in

(a) 1st principal component

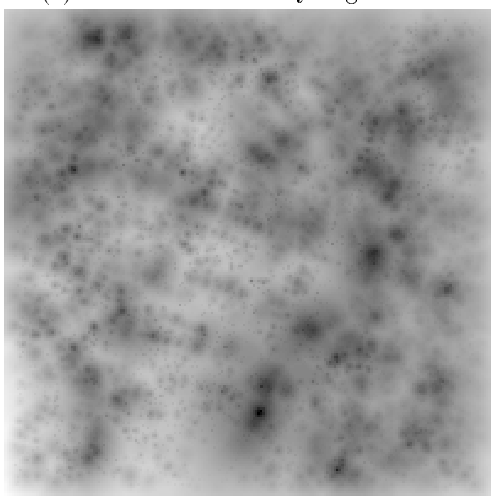
(b) Thresholded Canny edges $\sigma = 1$



(c) Thresholded Canny edges $\sigma = 4$



(d) MSSDT of edges



(e) MSSDT of corners

Figure 1: MSSDT of structural features

the expected similarity in spectral values between images. These range from invariance to uniform scaling of the data to any consistent mapping between the two sets of values.

Currently all the measures are combined using a neural network which is trained on examples of the different classes of change. At least some of the proposed change measures described in the previous section are likely to be correlated. Therefore the full set is probably redundant, and for reasons of efficiency could be reduced to a smaller subset of measures that contain most of the information. There are several ways this could be done; standard methods for feature selection are described by Devijver and Kittler.[5]

7 REFERENCES

[1] N. Abramson. *Information theory and coding.* McGraw-Hill, 1963.

[2] J.R. Anderson. Land use and land cover changes. A framework for monitoring. *J. Research USGS*, 5:143–153, 1977.

[3] G. Borgefors. Distance transforms in digital images. *Computer Vision, Graphics and Image Processing*, 34:344–371, 1986.

[4] J. Canny. A computational approach to edge detection. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 8:679–698, 1986.

[5] P.A. Devijver and J.V. Kittler. *Pattern Recognition: A Statistical Approach.* Prentice-Hall, Englewood Cliffs, New Jersey, 1982.

[6] J.T. Finn. Use of the average mutual information index in evaluating classification error and consistency. *Int. J. Geog. Inf. Sys.*, 7:349–366, 1993.

[7] T.D. Frank. The effect of change in vegetation cover and erosion patterns on albedo and texture of Landsat images in a semiarid environment. *Ann. Assoc. Am. Geog.*, 74:393–407, 1984.

[8] T. Fung and E. LeDrew. Land cover change detection with Thematic Mapper spectral/textural data at the rural-urban fringe. In *Proc. 21st Int. Symp. on Remote Sensing of Environment*, pages 783–789, 1987.

[9] R.M. Haralick. Statistical and structural approaches to texture. *Proc. IEEE*, 67:768–804, 1979.

[10] C.A. Hlavka. Land-use mapping using edge density texture measures on Thematic Mapper simulation data. *IEEE Trans. on Geoscience and Remote Sensing*, 25:104–108, 1987.

[11] G.F. Hughes. On the mean accuracy of statistical pattern recognizers. *IEEE Trans. on Information Theory*, 14:55–63, 1968.

[12] J.R. Jenson and D.L. Toll. Detecting residential land use development at the urban fringe. *Photogrammetric Engineering and Remote Sensing*, 48:629–643, 1982.

[13] S. Khorram, J.A. Brockhas, and H.M. Cheshire. Comparison of Landsat MSS and TM data for urban land-use classification. *IEEE Trans. on Geoscience and Remote Sensing*, 25:238–243, 1987.

[14] L. Kitchen and A. Rosenfeld. Gray level corner detection. *Pattern Recognition Letters*, 1:95–102, 1982.

[15] D.G. Leckie, P.M. Teillet, D.P. Ostaff, and G. Fedosejevs. Sensor band selection for detection of current defoliation caused by spruce budworm. *Remote Sensing of Environment*, 26:31–50, 1988.

[16] W.A. Malila. Change vector analysis. In *Proc. 6th Ann. Symp. Machine Processing of Remotely Sensed Data*, pages 326–335, 1980.

[17] L.R.G. Martin and P.J. Howarth. Change-detection accuracy assessment using SPOT multispectral imagery of the rural-urban fringe. *Remote Sensing of Environment*, 30:55–66, 1989.

[18] N.M. Mattikalli. An Integrated Geographical Information System's approach to land cover change assessment. In *Proc. IGARSS*, pages 1204–1206, 1994.

[19] C.R. Perry and L.F. Lautenschlager. Functional equivalence of spectral vegetation indices. *Remote Sensing of Environment*, 14:169–182, 1984.

[20] V.B. Robinson, J.C. Coiner, and T.H. Barringer. Dynamic modeling of vegetation change in arid lands. In *Proc. 1st Thematic Conf. Remote Sensing of Arid and Semi-Arid Lands*, pages 121–131, 1982.

[21] P.L. Rosin and G.A.W. West. Multi-scale salience distance transforms. In *Proc. British Machine Vision Conf.*, pages 579–588, 1993.

[22] D.E. Rumelhard and J.L. McClelland, editors. *Parallel Distributed Processing (Volume 1)*. MIT Press, 1986.

[23] A. Singh. Digital change detection techniques using remotely-sensed data. *Int. J. Remote Sensing*, 10:989–1003, 1989.

[24] J.F. Wang and P.J. Howarth. Structural measures for linear feature pattern recognition. *Canadian J. Remote Sensing*, 17:294–303, 1991.

[25] G. Wilkinson, S. Folving, K. Fullerton, and J. Megier. A study on the automatic revision of the European Community's Corine land cover database using satellite data. In *Int. Archives of Photogrammetry and Remote Sensing*, volume XXIX, Part B4, pages 543–548, 1992.

[26] S. Yafang, C. Shupeng, and Z. Yuanghong. Monitoring urban development of Hangzhou city by using multitemporal TM data. *GIS*, 2:8–13, 1992.