Multiscale remote sensing data segmentation and post-segmentation change detection based on logical modeling: Theoretical exposition and experimental results for forestland cover change analysis

Yashon O. Ouma\textsuperscript{a,}\textsuperscript{*}, S.S. Josaphat\textsuperscript{b}, Ryutaro Tateishi\textsuperscript{b}

\textsuperscript{a}CEReS, Tateishi Lab, Graduate School of Science and Technology, Chiba University, Japan
\textsuperscript{b}Center for Environmental Remote Sensing (CEReS), Chiba University, 1-33 Yayoi-cho, Inage-ku, Chiba 263-8522, Japan

Received 18 November 2005; received in revised form 7 May 2007; accepted 11 May 2007

Abstract

Quantification of forestland cover extents, changes and causes thereof are currently of regional and global research priority. Remote sensing data (RSD) play a significant role in this exercise. However, supervised classification-based forest mapping from RSD are limited by lack of ground-truth- and spectral-only-based methods. In this paper, first results of a methodology to detect change/no change based on unsupervised multi-resolution image transformation are presented. The technique combines directional wavelet transformation texture and multispectral imagery in an anisotropic diffusion aggregation or segmentation algorithm. The segmentation algorithm was implemented in unsupervised self-organizing feature map neural network. Using Landsat TM (1986) and ETM+ (2001), logical-operations-based change detection results for part of Mau forest in Kenya are presented. An overall accuracy for change detection of 88.4%, corresponding to kappa of 0.8265, was obtained. The methodology is able to predict the change information \textit{a-posteriori} as opposed to the conventional methods that require land cover classes \textit{a priori} for change detection. Most importantly, the approach can be used to predict the existence, location and extent of disturbances within natural environmental systems.

\textcopyright 2008 Elsevier Ltd. All rights reserved.

Keywords: Multispectral anisotropic diffusion (MAD); Wavelets transformation; Object-oriented segmentation; Logical modeling; Unsupervised change detection; Forestland cover

1. Introduction

Remote sensing data (RSD) from spaceborne and airborne platforms continue to provide valuable data for different applications. These include among others: mapping, environmental monitoring, disaster management, military intelligence. The analysis of RSD has however been predominantly carried out through per-pixel multispectral classification at single spatial scales, also termed as monoscale multispectral data classification. This approach does not accurately represent the inherent or true nature of features on the Earth's surface, which are varied in size, shape, color, texture as well as degrees of compactness.
During the last few years, the concept of object-oriented image analysis integrated low-level image processing techniques, such as multiresolution segmentation procedures and algorithms (Baatz and Schäpe, 2000), with high-level methods, such as artificial intelligence (knowledge-based expert systems and fuzzy systems) and pattern recognition methods. Within this approach, the low-level image analysis produces primitive image objects, while the high-level processing classifies these primitives into meaningful domain objects (Benz et al., 2004). The main advantage of this new approach is that the digital image is no longer considered as a grid of pixels, but as a group of primitives and homogeneous regions, called image objects. Using this kind of representation, one can tackle the problem of using local criteria for classification and/or change detection, as these features can be computed within the boundaries of an image object. As well, context information can easily be represented through topologic features between the objects (eCognition™, 2005). Arguably, objects can be more intelligent than pixels, in a sense of knowing their neighbors and the spatial or spectral relations within and amongst them. Because primitive objects are knowledge free, to accomplish thematic classification, one has then to use higher-level (knowledge-based) techniques. The main purpose of segmentation in object-oriented image analysis is not the extraction of semantic objects, but the extraction of image primitives.

Various segmentation algorithms have been proposed in the literature with promising results (Burnet and Blaschke, 2003; Benz et al., 2004; Hay et al., 2005; eCognition™, 2005). The decision for a segmentation technique to be included into an object-oriented image analysis system is based on the particular use of the system and the problem(s) it has to deal with. Remote sensing applications deal with data of various resolutions and complexity; consequently, a generic and less computation-intensive segmentation algorithm is a better choice (Baatz and Schäpe, 2000). Thus in RSD processing, a preferable choice would be a multiscale image segmentation algorithm (Chen et al., 2002; Hall and Hay, 2003; Fauzi and Lewis, 2003).

According to eCognition™ Software (2005), multiresolution-based segmentation of RSD is one of the latest and most suitable techniques in image segmentation. The approach involves the use of scale, color (spectral), smoothness or roughness (texture) and/or compactness in object-oriented image primitives’ extraction, resulting in the creation of homogenous image objects prior to classification.

1.1. Aims and flow of this research

For forestland cover change detection, different attempts have been made in the direction of multiresolution-based unsupervised change detection (Sgrenzaroli et al., 2002; Bruzzone and Cossu, 2003; Peddle et al., 2003; Hay et al., 2005). However, the challenge in large-scale afforestation, reforestation and/or deforestation (ARD) RS-based data acquisition is the inability of the existing systems to automatically determine, within considered time intervals, (i) whether there is change or no change; (ii) where change has occurred and the spatial extent of the change; and (iii) precisely the kind of change. Since forestland consists of many similarly reflecting features, it makes the unsupervised change detection task extremely difficult (Hall and Hay, 2003; Hay et al., 2005). This is coupled with the fact that obtaining ground-truth or training data is nearly impossible.

Following the above challenges, the aim of the current work is on the derivation of a general framework for the identification of changes in forested areas, with minimal a-priori ARD information. In particular, the technique is to be applicable to pairs of multitemporal–multispectral optical-sensor imagery. The specific objectives are to determine whether the corresponding(s) area was: (i) covered with forest in both dates, (ii) not forested in both dates, (iii) cleared between the two considered dates or (iv) vice versa.

The proposed approach is composed of three main steps: (1) 2D-discrete wavelets transform (DWT) decomposition of the band selected for multiscale–multidirectional texture extraction; (2) integrated textural and spectral segmentation by means of anisotropic diffusion. This step aims at reducing noise without blurring inter-region edges as well as creating the desired multiscale low-level primitives, and is implemented using the Self-Organizing Feature Map (SOFM) Neural Network. This is followed by the change detection step, which is accomplished through (i) classification of each image obtained from the segmentation or smoothing step, into forest/non-forest thematic maps via thresholding, and

---

(ii) identifying the areas and types of change by comparing the achieved classification in (i) through logical modeling (LM). The outlined steps above mark the differences between the proposed study and other state-of-the-art research in the same direction as done by Baatz and Schäpe (2000), Burnet and Blaschke (2003), Benz et al. (2004), Hay et al. (2005), and eCognition™ (2005). Finally, the reliability of the results is assessed through the computation of the overall accuracy and kappa index and class-producers’ and users’ accuracy. Fig. 1a presents a summary of the workflow for this research.

In Fig. 1a, an analysis of divergence measures computed at different scales enables one to identify the decomposition level that contains the relevant information for the specific application-desired primitives’ extraction. A logical comparison between the multi-temporal thematic maps enables one to derive the desired change detection map. TDM in Fig. 1 refers to the transformed divergence measure for class separability. The proposed approach is tested using 15.36 km × 15.36 km (512 pixels × 512 pixels) of Landsat TM and Landsat ETM+ datasets dated 1986 and 2001, respectively. In this paper, the terms smoothing, aggregation and segmentation are used interchangeably to mean the same thing. The same applies to the terms multiscale and multiresolution.

1.2. The rationale of investigation

In image texture transformation, different physical structures of the scene can be characterized at different resolutions based on the details of the

![Diagram](image-url)
image from fine-to-coarse. This is useful as a pattern recognition strategy within the spectral image feature space. The orientation of texture elements and their frequency contents are important clues for feature discrimination. This specifically motivated early researchers to study the representation of texture energy in the Fourier domain (Haralick et al., 1973; Eklundh, 1979). For segmentation purposes, it is however necessary to localize texture measurements over neighborhoods of varying sizes.

The rationale for the multiscale–multidirectional texture derivation is to detect sharp variations within an image \( I(x,y) \) in the direction of maximum change of the surface. Texture-based multispectral segmentation then aims at “dividing” the image into homogeneous regions where the local texture properties are approximately invariant. The goal then is to find the minimum number of measurements that can discriminate textures that are perceived to be inhomogeneous. Similarly, the measurements should also be approximately constant in a region where the texture is considered to be homogeneous.

Multiresolution or multiscale image decomposition provides the possibility of a scale-invariant interpretation of an image. In multiscale-wavelet texture-integrated segmentation, the difficulty is to find an algorithm that aggregates the wavelet-texture responses at all scales and orientations in order to find the boundaries of homogeneous textured regions. In this study, we suggest the anisotropic diffusion smoothing or segmentation approach. The use of anisotropic diffusion enables one to smooth homogeneous regions and at the same time enhance edges (Weickert et al., 1998), hence creating the desired low-level image primitives. The concept is illustrated in Fig. 1b, whereby at some level or scale of the segmentation, feature edges and noise are enhanced and eliminated, respectively.

1.3. Paper organization

For ease in flow, the paper is divided into three main parts: part A is on the introduction, study area, data and the results of pre-processing (geometric and radiometric normalization). Part B presents the theoretical domain of this work including methodology-based experimental results. Finally, part C is on the change detection based on LM and results validation. Note that since this study is largely experimental, the discussions are presented with the corresponding results in each phase.

2. Data and study area

2.1. Research test area

The test area for this study is situated in the Mau complex forests, which lies in the Kenyan Rift Valley. In the last 25 years there has been a rapid increase in human settlement and urban-related activities in this region. This has resulted into severe deforestation, mostly in the eastern part of the escarpment (Fig. 2). The consequence of which is the severe environmental degradation that has called for rapid monitoring, identification and quantification of the resulting land use changes and consequences. Obviously, conventional per-pixel multispectral-based approaches like supervised post-classification change detection, which requires ground-truth information prior to the change detection procedures, may not be efficient for the assessment of this area due to the economic logistics associated with remoteness and data acquisition.

2.2. Experimental data

Fig. 3 shows the test site, which is part of Fig. 2. Fig. 3 is subset from a full Landsat scene (path 169, row 60). Summary of the experimental data is presented in Table 1.

2.3. Data preprocessing—geometric correction and radiometric normalization

The test data were geometrically rectified to the Universal Transverse Mercator map projection system, zone 37-east, using GCPs collected mostly around the accessible lower parts of the mountain using Differential Global Positioning System. A 1:50 000 topographic map of 1997 and the GCPs of the study area were used to geometrically rectify the Landsat ETM+. The RMSE was 0.2 pixel. The Landsat TM was then rectified to the ETM+ through image-to-image registration with RMSE of 0.25 pixel. Both the images were resampled to 30 m x 30 m using the nearest-neighbor resampling method, and for both images the transformations were accomplished using the first-order polynomial. Through visual inspection of the resampling results,
the nearest-neighbor resampling conserved the original values of the images much better, than the bilinear interpolation and cubic convolution.

For change detection, a common radiometric response is required for accurate quantitative analysis of imagery of a scene acquired on different dates with different sensors. The technique of radiometric rectification, as described in Hall et al. (1991) and Jensen et al. (1995), was used to normalize or “radiometrically rectify” the geometrically corrected multi-temporal Landsat images. The radiometric rectification algorithm identifies “radiometric control sets”, i.e. sets of scene landscape elements with a mean reflectance that is expected to change very minimally with time, i.e. temporarily pseudo-invariant features. The average digital count values of these radiometric control sets are used to calculate linear transforms relating digital count values between images. Within the Landsat scene, the following features (pixels) were used as the control points: (a) five well-distributed Rift Valley lakes within the scene—Baringo, Bogoria, Nakuru, Elementaita and Naivasha, (b) central business district of Nakuru Town and (c) old-permanent tarmac roads as well as road intersections.
2.4. Texture and spectral bands selection

In order to select the most informative band for forest texture extraction, five classes of different but related spectral characteristics representing the forest area, and three classes representing non-forest areas were selected as separability training sets from the TM image, since it was the oldest. The two-dimensional (2-D) class separability feature-space and within-class variance measures were used to determine the most informative bands.

To determine the optimal bands for texture mapping and spectral separation of forest from non-forest, the average of the variance within each class (five forest classes and three non-forest classes) was computed. The results for the average variance comparison are presented in Table 2, and that of the 2-D feature-space separability signatures are presented in Fig. 4. In Fig. 4 as well as in Table 2, the five forest classes were combined to a single representative class, and the three non-forest representative classes were also combined to one class so as to clearly visualize the overall separability between the two main classes (forest/non-forest).

From Table 2 it is deducible that bands 3, 4 and 5 are the most informative for forest mapping. Thus we only plotted the 2-D scatter-plots for these bands.

---

**Table 1**

Summary of experimental data

<table>
<thead>
<tr>
<th>Test site</th>
<th>Data (satellite and reference)</th>
<th>Size</th>
<th>Resolution</th>
<th>Date of acquisition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mau escarpments in Kenya</td>
<td>Landsat ETM+</td>
<td>512 x 512</td>
<td>30 m x 30 m</td>
<td>3 Apr 2001</td>
</tr>
<tr>
<td></td>
<td>Landsat TM</td>
<td>512 x 512</td>
<td>30 m x 30 m</td>
<td>28 Mar 1986</td>
</tr>
<tr>
<td></td>
<td>1: 50,000 (TOPO)</td>
<td>1 Sheet</td>
<td>Scale: 1:50,000</td>
<td>1997</td>
</tr>
<tr>
<td></td>
<td>1:250,000(VEGETATION)</td>
<td>1 Sheet</td>
<td>Scale: 1:250,000</td>
<td>1986</td>
</tr>
</tbody>
</table>

**Table 2**

Average variance for forest and non-forest classes in visible (red, green and blue), near-infrared (NIR) and mid-infrared (MIR) bands

<table>
<thead>
<tr>
<th>Bands and wavelengths</th>
<th>Blue (0.45–0.52)</th>
<th>Green (0.52–0.60)</th>
<th>Red (0.63–0.69)</th>
<th>NIR (0.76–0.90)</th>
<th>MIR1 (1.55–1.75)</th>
<th>MIR2 (2.08–2.35)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forest</td>
<td>12.10</td>
<td>9.06</td>
<td>66.42</td>
<td>262.31</td>
<td>227.58</td>
<td>20.61</td>
</tr>
<tr>
<td>Non-forest</td>
<td>33.18</td>
<td>18.75</td>
<td>43.17</td>
<td>46.92</td>
<td>204.12</td>
<td>106.92</td>
</tr>
</tbody>
</table>

Fig. 4. Forest/non-forest feature space. Green (lower-ellipse) depicts forest classes, while bluish (upper-ellipse) depicts non-forested areas for: (a) band 3 vs. band 4, and (b) band 5 vs. band 4.
in Fig. 4. Band 4 was selected for the texture extraction as it gave the highest variance or best separability. Bands 3, 4 and 5 were selected as spectral inputs for the multispectral anisotropic diffusion (segmentation) since they depicted the least correlation.

3. Multispectral–multitemporal imagery smoothing: theoretical insight

3.1. Multiresolution wavelet decomposition (MWD)

It can be argued that the spatial dependence within a land cover class may show the following three peculiar texture features: (a) a characteristic scale or scales of variability that can be defined as the minimum distance over which a comparison must be made between two pixels if the difference between them will represent, on average, all the variability within the class; (b) directional dependence (anisotropy), where variation in DN values may be seen in one direction more than another; and (c) spatial periodicity, where spatial variation in DN values may show periodicity characteristics of a particular land cover in an image. Exploitation of these texture characteristics is important in low-level primitive information detection.

Deviating from the occurrence and co-occurrence textures, Mallat (1989) developed a theory for multiresolution signal decomposition using the orthonormal wavelet basis. The multiresolution wavelet transform decomposes a signal into low-frequency approximation and high-frequency detail information at successively coarser resolutions. The resultant approximation is then decomposed into the second level of approximation and detail, iteratively (Fig. 5). MWD in discrete-time corresponds to successive band filters decomposing the image into details and overall pattern. It separates high from low frequencies recursively using the same transform at the new scale (Carvalho et al., 2001). The result is multiscale image texture represented in the vertical, diagonal and horizontal directions.

MWD texture plays an important role in pattern recognition, image interpretation and segmentation in a variety of applications. A detailed theoretical description of MWD and its application to RSD analysis can be found in Mallat (1989), Ouma et al. (2006). In this study, the MWD results are to be used in the next phase of the multispectral anisotropic diffusion (segmentation).

3.2. Multispectral anisotropic diffusion (MAD)

A fundamental objective in object/feature recognition is the definition of the feature boundary. This recognition can be best achieved if the object/feature of interest is captured at its representative or inherent spatial resolution. Sharp contrasts in images generally correspond to object boundaries or phenomena such as noise and shadow. To remove noise, which generally shows up much more in low contrast areas, a smoothing filter can be used. A smoothing filter can also be used for the purpose of data reduction when employing a hierarchical scale-space approach. For the choice of a smoothing filter, the following considerations are necessary: first, the operator must reduce the range of scales that the image represents, meaning that it must be localized in the frequency domain. Second, the operator should not generate spurious detail when smoothing from finer to coarser scales.

The Gaussian operator comes close to satisfying the above conditions for a smoothing filter. The smoothed images \( I(x, y, \sigma) \) are generated by convolving the original image, \( I_0(x, y) \) with a Gaussian kernel of variance \( \sigma^2 \). The variance determines the amount of smoothing and is sometimes referred to as the scale-space parameter. The Gaussian while being popular has the following limitations: (1) it delocalizes and blurs edges, as greater amounts of smoothing are performed, and (2) it does not adequately consider inhomogeneity or anisotropy when smoothing, and this sometimes causes spectrally different regions to be inappropriately merged.

There have been several attempts to use texture transform to improve the results of spectral analysis of remotely sensed data, leading to the development of...
of many algorithms in spectral–textural analysis, e.g. Perona and Malik (1990), Deng and Liu (2000). In this paper, we utilize the anisotropic diffusion scheme, first proposed by Perona and Malik (1990), to tackle the problem of noise reduction and smoothing or segmentation via edge preservation and homogenous land cover derivation in multitemporal forestland cover data.

The anisotropic diffusion scheme is a modification of the heat equation or linear diffusion, and the continuous anisotropic diffusion is given by

\[
\frac{\partial I_t(x,y)}{\partial t} = \text{div}[c_t(x,y)\nabla I_t(x,y)],
\]

where \(I_t(x,y)\) is the image at time \(t\), \(\text{div}\) the divergence operator, \(\nabla I_t(x,y)\) the gradient of the image and \(c_t(x,y)\) the diffusion coefficient. If \(c_t(x,y)\) is constant, then Eq. (1) is reduced to the isotropic diffusion equation, and is equivalent to convolving with a Gaussian kernel. The idea of anisotropic diffusion is to adaptively choose \(c_t\) such that intra-regions become smooth while edges of inter-regions are preserved. The diffusion coefficient \(c_t\) is generally selected to be a non-negative function of gradient magnitude, so that small variations in intensity such as noise or shading can be well smoothed, and edges with large intensity transition are distinctly retained. In-depth analysis of the behavior of the anisotropic diffusion model can be found in Weickert et al. (1998) and Barash (2002).

While anisotropic diffusion approach has grown to become an useful tool for edge detection (Chen and Barcelos, 2001); image enhancement (Sole’ and Lo’pez, 2001); image smoothing (Tsuji et al., 2002); image segmentation (Bakalexis et al., 2002); and texture segmentation (Deng and Liu, 2000) in other fields of digital image processing, its applications in land cover remote sensing data processing and analysis is still not well established.

In this paper, we use anisotropic diffusion to differentiate homogeneous and inhomogeneous land covers. Theoretically, the anisotropic diffusion is to act as a selective smoothing filter, whereby it triggers the smoothing process in homogeneous areas and stops the diffusion process at the edges to preserve sharp edges of different land cover types, hence acting as a segmentation algorithm. The two diffusion coefficient functions with an adaptive-function parameter are evaluated for the specific application of forestland cover segmentation. The adaptive-function parameter is annealed over time so that the diffusion process will effectively smooth irregular background spectral–textural information, and yet distinctly preserve edges.

### 3.2.1. Anisotropic diffusion derivation and parametrization

If we let \(I_0(x,y)\) be a gray-level image at coordinates \((x,y)\) of a digital image at theoretical time \(t\), and \(I_0(x,y)\) be the original input image, the continuous anisotropic diffusion in Eq. (1) can be discretely implemented using 4-nearest neighbors (Fig. 6a) and the Laplacian operator (Torkamani-Azar and Tait, 1996)

\[
I_{i+1}(x,y) = I_i(x,y) + \frac{1}{4} \sum_{i=1}^{4} [c_i(x,y)\nabla I_i(x,y)],
\]

where \(\nabla I_i(x,y), i = 1, 2, 3\) and 4 represent the gradients of the 4-neighbors (Fig. 6a) in the north, south, east and west directions, respectively, and is computed as

\[
\nabla I_i^1(x,y) = I_i(x,y - 1) - I_i(x,y),
\]

\[
\nabla I_i^2(x,y) = I_i(x,y + 1) - I_i(x,y),
\]

\[
\nabla I_i^3(x,y) = I_i(x + 1,y) - I_i(x,y),
\]

\[
\nabla I_i^4(x,y) = I_i(x - 1,y) - I_i(x,y),
\]

and \(c^i(x,y)\) is the diffusion coefficient associated with \(\nabla I_i^i(x,y)\), and can be considered as a function of the magnitude of gradient \(\nabla I_i^i(x,y)\), i.e.,

\[
c^i(x,y) = g(\nabla I_i^i(x,y)).
\]

For the sake of simplicity, \(\nabla I_i(x,y)\) is denoted by \(\nabla I\), and \(g(\nabla I)\) has to be a non-negative monotonically decreasing function with \(g(\nabla I) = 1\) and \(\lim_{|\nabla I| \to \infty} g(\nabla I) = 0\) (Barash, 2002). The selection of \(g(\nabla I)\) is to have a low coefficient value within image regions, so that unwanted noise are thoroughly smoothened and inter-region edges of defects are preserved. Two possible diffusion coefficient functions are

\[
g(\nabla I) = c_1(x,y,t) = \exp\left(\frac{|\nabla I(x,y,t)|}{\kappa}\right)^2
\]

and

\[
g(\nabla I) = c_2(x,y,t) = \frac{1}{1 + \left(\frac{|\nabla I(x,y,t)|}{\kappa}\right)^2}
\]

In the anisotropic diffusion model by Perona and Malik (1990), the parameter \(\kappa\) is a constant, and must be fine-tuned for a particular application. \(\kappa\) is a threshold parameter that influences the anisotropic
smoothing process. It is also called the diffusion constant or the flow constant, and acts as an edge strength threshold.

If the $k$ value is an overly small constant in all diffusion iterations, the diffusion will stop in early iterations and the background or truly homogeneous areas cannot be sufficiently smoothed. This may cause false rejection of faultless homogeneous land cover patches. Reversely, if the $k$ value is a large constant, the diffusion process will over-smooth in early iterations and both the background texture and edges will be removed, and this may cause false acceptance. The plots in Fig. 6b depict the diffusion coefficient functions of Eqs. (5) and (6).

If $\phi(\nabla I)$ is the flux defined as $\phi(\nabla I) = g(\nabla I)\nabla I$, which can be expressed as $\Phi_1(k)$ and $\Phi_2(k)$, for $c_1$ and $c_2$, respectively, then the following plots (Fig. 6c) are obtained. Simplifying further the equations above for graphical representation, we denote $h = |\nabla I(x,y,t)|$, and $s = h/k$, and obtain the following expressions for the conductivity functions: $c_1(x,y,t) = \exp(-s^2)$ and $c_2(x,y,t) = 1/(1 + s^2)$.

A large flux value indicates a strong effect on smoothness. Fig. 6c presents the graphical representations of the flux functions of the respective diffusion coefficient functions in Eqs. (5) and (6). For a given $k$ value, it can be seen from Fig. 5 that the diffusion coefficient function of Eq. (5) drops dramatically and approximates to zero when the gradient magnitude $|\nabla I|$ is larger than $2k$, i.e. the diffusion stops as soon as $|\nabla I| > 2k$. The maximum smoothness occurs at $|\nabla I| = 0.75k$ as shown in the corresponding flux function. The diffusion coefficient function of Eq. (6), instead, decreases more gradually even when $|\nabla I| > 2k$. Its corresponding flux function shows that the maximum smoothness is at $|\nabla I| = 1k$.

Compared to Eq. (6), it can be concluded that Eq. (5) privileges high-contrast edges over
low-contrast edges. Therefore in the application of forestland cover segmentation, diffusion coefficient function of Eq. (6) is more desirable than that of Eq. (5), as it will favor even low-contrast edges and adequately considers the large homogeneous patches. Natural forestland cover surfaces consist of inhomogeneous textures and some homogeneous regions may contain noise. The diffusion coefficient function of Eq. (5) may cause the segmentation or smoothing process to stop in the early iterations, and the background spectral and textural information will not be sufficiently removed.

Given that the gradient threshold $k$ is a constant, the selection of a best $k$ value becomes extremely crucial. A large $k$ value will over-smooth both background information and edges. An overly small $k$ value disables the diffusion process and the unwanted background texture will be preserved.

In order to alleviate the limitations of the use of a constant $k$, we used the annealing $n$th root function for the gradient threshold $k$, to determine the suitable thresholds. Its value will be reduced as the diffusion iteration increases. In each diffusion iteration, the gradient magnitude, that is intensity-contrast texture, will be reduced in the filtered image. A constant $k$ will eventually smooth out the irrelevant information like noise. However, as the gradient threshold adaptively decreases with the increment of iterations, the diffusion process has no effect on the inhomogeneous regions, while it can gradually remove the background textures as long as the decrement of gradient magnitude in homogeneous regions is competitive with the decrement of the $k$ value.

3.2.2. Segmentation threshold parameter ($k$) determination

The annealing $n$th root function used in this study is defined by

$$k(t) = k(0)t^{-1/n},$$

where $k(t)$ is the gradient threshold at iteration $t$, $k(0)$ is the initial value and $n$ is a positive integer. Figs. 7(a)–(d) present the plots of four root functions with $n = 1, 2, 3$ and 4, respectively.

Note how the shape of the function is affected as the value of $n$ is changed. The graphs show that a small $n$, such as $n = 1$, will make the $k$ value drop rapidly and cause the diffusion process to stop at a small number of iterations. As $n$ increases, the $k$ value will decrease gradually and result in fast smoothness in a small number of iterations. An overly large value of $n$ may over-smooth both background textures and subtle defects in early iterations. From the results in Fig. 7, $n = 3$ gave the optimal solution, whereby maximum diffusion

---

![Fig. 7. (a)–(d). Plots for first four root functions from (a)–(d) for $n = 1, 2, 3$ and 4, respectively. $t$ represents number of iterations; $n$ is level or scale of diffusion and $k$ represents gradient threshold.](image-url)
occurs at approximately $1 \kappa$. Thus the used diffusion function was defined by
\[
g(\nabla I) = c_2(x, y, t) = \frac{1}{1 + \left(\frac{|\nabla I(x, y)|}{\kappa}\right)^2}. \tag{8}
\]

Since the gradient threshold value $\kappa$ is adaptively decreased as the iteration number increases, the selection of the initial value $\kappa(0)$ is not as crucial as that of a constant $\kappa$ in the Perona and Malik’s model. As seen in Fig. 6c, the flux function of the diffusion coefficient in Eq. (6) shows that the maximum smoothness is given by $|\nabla I| = 1\kappa$. We therefore set the initial value $\kappa(0)$ of the annealing cubic-root function to the average gradient magnitude of the whole image used to control the multispectral segmentation (Eq. (9)), and compared this with other possible image spatial parameters like spectral and textural gradient information:
\[
\kappa(0) = \frac{1}{4MN} \sum_i \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} |\nabla I(x, y)|. \tag{9}
\]

For example, for the implementation of an image $\{I(x, y)\}, 0 \leq x, y < N$, the wavelet transform produces four components $\{LL, LH, HL, and HH\}$, where $LL$ corresponds to the smoothed image, $HL$ contains the horizontal edges, $LH$ contains the vertical edges and $HH$ contains edges not parallel to the axes, and the estimation of the conduction parameter $\kappa(x, y)$ based on wavelet transform upon experimentation is given by
\[
\kappa(x, y) = \frac{\kappa_0}{1 + \alpha|HL(x, y)|' + \beta|LH(x, y)|' + \gamma|HH(x, y)|'}, \tag{10}
\]

where $\kappa_0$ is the 90% value of average value of $|\nabla I(x, y)|'$; the weights satisfy the conditions: $0 \leq \alpha, \beta, \gamma \leq 1; \alpha + \beta + \gamma = 1$, and $|HL(x, y)|'$, $|LH(x, y)|'$, $|HH(x, y)|'$ are the wavelet transform texture gradients in the horizontal, vertical and diagonal directions, respectively, at $t$ optimal iterations. A thorough experimentation showed that the selected $\kappa(0)$ works successfully for this application. This was determined purely based on empirical testing and may vary from one application to another. Thus Eq. (9) may be considered as a more universal approach.

In related utility of anisotropic diffusion, Barash (2002) showed that an iteration of adaptive smoothing:
\[
I_{t+1}(x, y) = \frac{\sum_i \sum_j I(x + i, y + j)w_t(x + i, y + j)}{\sum_i \sum_j w_t(x + i, y + j)}
\]
is an implementation of the discrete version of the anisotropic diffusion equation if weight $(w_t)$ is taken as same as the diffusion coefficient $c$. This weight factor can be modeled in the case of RSD as the input texture features at the corresponding levels of diffusion or scale, which is represented by the conduction parameter $\kappa$, as in our case (Eq. (10)). At each segmentation level, each band is smoothed based on the homogeneity of the input directional textures.

3.3. Self-organizing feature map (SOFM)

The multispectral–multitextural segmentation, MAD, was implemented using the Kohonen’s unsupervised SOFM neural network, as it is conceptually analogous to the anisotropic diffusion process. SOFM is a special type of neural network that can learn from complex, multi-dimensional data and transform them into visually decipherable clusters. SOFM imitates the function of ‘grouping by categories’ operated by the human brain. SOFM operates on the concept of unsupervised cluster analysis. Cluster analysis is a technique for grouping subjects into clusters of similar elements. In cluster analysis, similar elements are identified and grouped according to their attributes.

The main function of SOFM networks is to map the input data from an $n$-dimensional space to a lower-dimensional (usually 1-D or 2-D) plot while maintaining the original topological relations. The physical location of points on the map shows the relative similarity between the points in the multi-dimensional space. Unlike other neural network approaches, the SOFM network performs unsupervised training. That is, during the learning process, the processing units in the network adjust their weights primarily based on the lateral feedback connections. Unsupervised learning does not require the knowledge of target values. The nodes in the network converge to form clusters to represent groups of entities with similar properties. The number and composition of clusters can be visually
determined based on the output distribution generated by the training process.

SOFM networks combine competitive learning with dimensionality reduction by smoothing the clusters with respect to an a-priori grid and provide a powerful tool for data visualization. The SOFM, through the self-organization process, configures the output units into a topological representation of the original data, positioning the prototype vectors on a regular low-dimensional grid in an ordered fashion, making the SOFM a powerful visualization tool.

The essence of Kohonen’s SOFM is that it substitutes a simple geometric computation for more detailed properties of the Hebb-like rule and lateral interactions. There are three basic steps involved in the application of the algorithm after initialization, namely: sampling, similarity matching and updating. These three steps are repeated until the map formation is complete. More theoretical details on SOFM can be found in Kohonen (2001), and therefore only a summary is presented in this study. The generalized SOFM algorithm can be summarized as follows:

Generalized SOFM algorithm.

Step 1: Initialize the synaptic weights of the network, $V_j(0)$, to small, different, random numbers at iteration $k = 0$.

Step 2: Draw a sample $y$ from the input set.

Step 3: Find the best-matching (winning) neuron $r(y)$ at iteration $k$, using the minimum distance Euclidean criterion

$$r(y) = \min(||y - V_j|| | j = 1, 2, \ldots, L).$$

Step 4: Update the synaptic weight vectors using the update formula

$$V_{r(y)}^{k+1} = V_{r(y)}^k + \eta^k(y - V_{r(y)}^k),$$

$$V_j^{k+1} = V_j^k + (\eta^k)^2(y - V_j^k),$$

$\forall j \in \Omega_{r(y)}(k),$

where $\Omega_{r(y)}(k)$ is the neighborhood of $r(y)$.

Step 5: Increment $k$ by 1, go to step 2, and continue until the synaptic weights $V_j$ reach their steady-state values.

The similarity between Figs. 6a and 8 is rather obvious, and given the stated advantages of SOFM, it was used in the implementation of MAD, hence the SOFM–MAD acronym used in this paper.

4. Methodology implementation

The implementation is principally in two phases: first computation of the wavelets texture and then the integration of the texture with multispectral data in an unsupervised SOFM neural network. The procedures were implemented in MATLAB WaveLab Toolbox 4.0.5 and PCI Geomatica version 9.1 software for wavelets texture and MAD–SOFM implementation, respectively.

4.1. DWT implementation

The db8 (Daubechies, 1992) wavelet filter coefficients as in MATLAB WaveLab were chosen based on earlier experiments (Ouma et al., 2006) to create multiscale texture maps, resulting in a set of detailed images at different scales. Detail images occur in groups of three (horizontal, vertical and diagonal) at each scale. The first group of detail images is of size $P/2$ pixels and $L/2$ lines from the initial image of $P$ pixels and $L$ lines. The $n$th group will be of size $P/(2^n)$ pixels and $L/(2^n)$ lines. A detailed implementation of DWT can be found in Ouma et al. (2006).
4.2. **MAD–SOFM implementation**

Based on the SOFM concept, the segmentation process using the multiscale anisotropic diffusion is accomplished by constructing multiscale or multilevel structure with successive anisotropic diffusion of the original image and then linking nodes at each level to the final node at root level as shown on Fig. 9a. The node value at \((i, j)\) of level \(l\) is initialized by anisotropic diffusion result in the 4-directions or nearest-neighbors (Fig. 6a) level of level \(l/2\). The root level can be selected for segmentation, and thus the root level determines the number of final segments. Generally, the linking process where node values are updated and each node is linked to one of the root nodes is iteratively performed; each node is linked to the father that has the root value closest to its value, the node value is updated based on the node values of its sons, and finally, root values are assigned to the descendent nodes successively.

Fig. 9b presents the basic structure of the self-organizing neural network procedure for implementing the segmentation process, which is a translation of Fig. 9a.

![Diagram](image)

**Fig. 9.** (a) Aggregation or segmentation process using multiscale anisotropic diffusion for each level or scale. (b) Structure of SOFM neural network consisting, in principal, of input layer—image wavelets textural and spectral information, and output (Kohonen) layer—smoothened multiscale visible, NIR and MIR bands. H, V and D represent horizontal, vertical and diagonal DWT subbands, and WT-level \(i\) refers to wavelet transform level.
Summarily, MA–DSOFM algorithm modifies the pixel values of the input image by consecutive weighted averaging with neighboring pixels. The rate of smoothing is a function of the sum of the local spectral and textural similarity (PCI Geomtica v9.1). The similarity values are scaled by the spectral and textural rates. The following are the main characteristics of the applied MAD–SOFM algorithm (Fig. 9b):

(i) Simultaneous segmentation or smoothing of all input spectral bands–multispectral anisotropic diffusion.

(ii) Operation at five levels of scale/resolution of the multispectral image. Each network level is a 2-D array of nodes containing the current modified image at a given scale and location. The bottom network level consists of one node for each pixel in the full scale (initial image). This level is initialized using the input multispectral image and the successive levels undergo a 2:1 downsampling.

(iii) Both spectral and texture information of images are used to determine the amount of modification at a given scale and location. Two kinds of data are used in this step: (1) multispectral images of two dates—three bands (3,4,5) for each date, and (2) five successive levels of texture maps (consisting of three detail images for each level), resulting from wavelet transform. Considering the largest detail image, which contains half of the pixels and lines of the initial image in wavelet transform, the window size of the initial image for segmentation was 256 pixels and 256 lines (equal to the largest detail image resulting from wavelet transform).

Fig. 10. (a) TM band 4 and the first group of DWT detail images: (b) vertical, (c) horizontal and (d) diagonal information images.
The algorithm was run for multitemporal images separately at five levels using the same parameters.

5. DWT and MAD results

5.1. Wavelets transformation (DWT) results

As already stated above, the DWT was carried out at 5 scales for both TM and ETM+ data sets. The first groups of detail images are of size 256 P x 256 L. The second, third, fourth and fifth groups were 128 P x 128 L, 64 P x 64 L, 32 P x 32 L and 16 P x 16 L, respectively. In multilevel segmentation using anisotropic diffusion technique, the first 5 groups of detail images are used as input texture maps. Fig. 10 and 11 shows TM and ETM+ band 4 and their first group of detail images, respectively.

5.2. MAD–SOFM results

In this section, the performance of the MAD–SOFM is first evaluated followed by details on the experimental results.

5.2.1. Testing of the MAD–SOFM performance

For the MAD–SOFM performance evaluation, a subset of band 4 (NIR-band) of the latest Landsat data, that is 2001 ETM+, was selected as the test area. The results for the selection of optimal parameters were evaluated via boundary overlay of the manually delineated subset and the segmentation output boundaries. Next, the deviation of the segmented feature from the original feature was
numerically evaluated. For this, the two features were overlapped and their boundaries were compared. Then, the numbers of pixels in the interior and exterior of the original feature that is not matched in the segmented feature were counted as the interior and exterior error. The results of this step are presented in Table 3 for the original imagery against the five-segmentation levels.

It was found that level 3 gave minimal boundary errors, as compared with the manually derived portions. Level 1 and level 5 gave the poorest results, of course owing to the observations already discussed above from Fig. 6c. The results in Table 3 are comparable to those of Fig. 7. Implying that for $n = 1$, there is an obvious pre-mature segmentation effects, irrespective of the number of iterations.

5.2.2. MAD–SOFM output results

The output results for this phase are 5 levels or scale with level 1 being of size $256 \times 256$; level 2—$128 \times 128$; level 3—$64 \times 64$; level 4—$32 \times 32$; and level 5—$16 \times 16$. Each level contains three bands corresponding to the selected bands 3, 4 and 5. Not all the levels are relevant in mapping forest and non-forest information. Therefore, by computing the TDM separability between the two classes (forest and non-forest) at each level for each band, we statistically determined the optimal or most informative level amongst the five. The average variances for the test classes were also computed in the respective output bands.

The TDM average was computed as follows: first, for each forest class, its separability (TDM) with the three non-forest classes was computed. This resulted in three TDM measures for each forest class, for every band in a given level. This was then averaged to give a single TDM measure per forest class such that for band $i$ in level $j$, five TDM measures (corresponding to the five forest classes) were obtained and the average taken. The average variances were also computed in a similar manner for the forest classes only. The TDM average ($TDM_{\text{AVERAGE}}$) and average variances ($VAR_{\text{AVERAGE}}$) were computed as summarized in Table 4. Fig. 12 shows the TDM results for TM and ETM+ datasets.

Fig. 12 shows that TM and ETM+ results are different, depicting the fact that some land cover changes had occurred over this temporal interval. Overall, level 3 gave the highest TDM, meaning highest separability between the forest and non-forest classes. Level 1 had the least separability with

<table>
<thead>
<tr>
<th>Segmentation level</th>
<th>Interior</th>
<th>Exterior</th>
<th>Total error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level 1</td>
<td>911</td>
<td>843</td>
<td>1754</td>
</tr>
<tr>
<td>Level 2</td>
<td>209</td>
<td>120</td>
<td>329</td>
</tr>
<tr>
<td>Level 3</td>
<td>99</td>
<td>107</td>
<td>206</td>
</tr>
<tr>
<td>Level 4</td>
<td>112</td>
<td>143</td>
<td>255</td>
</tr>
<tr>
<td>Level 5</td>
<td>459</td>
<td>302</td>
<td>761</td>
</tr>
</tbody>
</table>

Table 4

<table>
<thead>
<tr>
<th>Forest ($F_i$)</th>
<th>Non-forest $(nF_i)$</th>
<th>TDM between $(F_i$ and $nF_i)$</th>
<th>Variance for $F_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$F_1$</td>
<td>$nF_1$</td>
<td>$TDM_{11}$ $TDM_{12}$ $TDM_{13}$</td>
<td>$\sigma^2_{F_1;level1;band1}$</td>
</tr>
<tr>
<td>$F_2$</td>
<td>$nF_2$</td>
<td>$TDM_{21}$ $TDM_{22}$ $TDM_{23}$</td>
<td>$\sigma^2_{F_2;level1;band1}$</td>
</tr>
<tr>
<td>$F_3$</td>
<td>$nF_3$</td>
<td>$TDM_{31}$ $TDM_{32}$ $TDM_{33}$</td>
<td>$\sigma^2_{F_3;level1;band1}$</td>
</tr>
<tr>
<td>$F_4$</td>
<td></td>
<td>$TDM_{41}$ $TDM_{42}$ $TDM_{43}$</td>
<td>$\sigma^2_{F_4;level1;band1}$</td>
</tr>
<tr>
<td>$F_5$</td>
<td></td>
<td>$TDM_{51}$ $TDM_{52}$ $TDM_{53}$</td>
<td>$\sigma^2_{F_5;level1;band1}$</td>
</tr>
</tbody>
</table>

Example: For level 1, band 1 (L1, B1)

Utility of TDM and class average variance. Example illustration with band $x$ and level $y$, whereby $y = 1, 2, 3, 4, 5$ and each. $y$ has three bands ($x = 1, 2, 3$ for each level). TDM_AVERAGEs are plotted as in Fig. 11. To avoid multiple variances, we only state VAR_AVERAGE performances. (Similar concept of average variances is used in Table 2.)

$TDM_{11}$ is derived from $F_1$ and $nF_1$
$TDM_{12}$ is derived from $F_1$ and $nF_2$ and
$TDM_{13}$ is derived from $F_1$ and $nF_3$
and so on for the rest of $F_i$-classes.
TDM values lower than 0.500, followed by level 5. Levels 2 and 4 were the only levels that were comparable with level 3.

Results for the first three segmentation levels are presented in Figs. 13 and 14 for TM and ETM+, in RGB composite of the resulting bands 3, 4 and 5. From the first three levels, it is noticed that the first level has different information from the second and the third. The last two levels have very coarse resolutions and thus do not present very informative results.

As the levels increase from 1 to 5, the nature of the aggregation of neighboring pixels also changes. The results of the different levels are influenced by the scale or resolution of the texture information. The result of level 1 corresponds to the initial network output. As $n$ increases from level 1 to level 5, a logarithmic pattern of termination of the segmentation process is observed, resulting in the image seen in level 1. Stability or best smoothing is gained at level 3 and over-smoothing is observed past level 3. Thus from level 4, there is evident onset of over-aggregation such that the well-aggregated features at level 3 begin to be disintegrated. These facts are also depicted by the shape of Fig. 12, with a maxima of TDM observed at level 3.

It is also observed that level 3 results, corresponding to 240 m spatial resolution, depict the true forest/non-forest scenario in the test area as compared with the higher or lower segmentation resolutions. The fact here is that noise, which is in most cases of lower spatial size than the scene forest patches, is minimized at the optimal spatial resolution, but is integrated into the forest patches at lower and higher spatial scales. Coarser spatial resolutions, corresponding to 480 m (level 4) and 960 m (level 5), over-aggregated the landscape patterns, resulting in the emergence of undesirable or unclear results about the scene. Logically speaking, the forest patches in this scene could not be adequately mapped at the resolutions of 480–960 m, and instead new segmentation patterns are seen to emerge like in the first two levels. It is appropriate to conclude from the segmentation results presented that a pre-implementation performance strategy is useful in deriving the possible expected behavior of the diffusion or segmentation results.

6. Multiscale change detection based on logical modeling

This phase of change detection is also termed the high-level primitives' derivation stage and is derived from the low-level image processing results, i.e., the segmented imagery. In this study, we introduce the use of LM concept in the change-detection process. LM is considered as a formalism for reasoning under uncertainty, with special advantages in its treatment of ambiguous data and the “ignorance” arising from or about it as is demonstrated below.

Implementing the logical theory in a specific problem generally involves solving two related problems. First, the uncertainties in the problem must be sort into $a$-priori independent items of

![Fig. 12. TDM between forest and non-forest classes for TM and ETM+ for five levels of diffusion. $B_1$, $B_2$ and $B_3$ refer to TM and ETM+ bands 3, 4 and 5, respectively.](image-url)
Secondly, the logical rule is then computationally carried out. These two problems and their solutions are closely related. Sorting the uncertainties into independent items leads to a structure involving items of evidence that bear on different but related questions, and this structure can be used to make computations feasible. In our case, we have two hypotheses or subsets \{forest (F), non-forest (nF)\}, the plausible combinations upon change/no-change are \{F, F\}, \{nF, nF\}, \{F, nF\}, \{nF, F\}. These sets imply that between two temporal intervals, it is possible to have no changes \{F, F\} and \{nF, nF\}, or changes \{F to nF\} and \{nF to F\}.

Fig. 15 presents the stepwise logic of the implementation of the LM. Polygons or theme maps are generated for reasoning from \{F\} and \{nF\}. Logical operations are performed on the database \{F, F\}, \{nF, nF\}, \{F, nF\}, \{nF, F\} of bitmap segments with the following “boundary” condition (Eq. (11)):

$$L([F, F],[nF, nF],[F, nF],[nF, F]) = 1. \quad (11)$$

This boundary condition states that within the two time periods, the expected probability for change and no change irrespective of any other conditions is constant and equal to 1.

6.1. Theme map generation

In order to quantify the changes, allocation of a unique value to each smoothed area (representing one class) is required. Because the forest/non-forest
areas have been isolated via the diffusion smoothing procedure, it is direct and easy to determine the threshold DN values that define these forest and non-forest regions. This was achieved by comparing: \(k\)-means clustering, density slicing and histogram for the forest/non-forest regions. \(k\)-means clustering and density slicing gave similar results, which were then used in this study. Note that level 3 results consisting of VIS, NIR and MIR were used in this stage.

Thresholding technique was used to allocate a unique value to the range of values representing forest areas. Then the logical operation \(NOT\) was applied to separate the forest class from non-forest areas. **Fig. 16a and b** shows the results of this

---

**Fig. 14.** Anisotropic diffusion results form ETM+ image presented as RGB for levels 1, 2 and 3, respectively.

**Fig. 15.** Logical modeling (LM) concept.
6.2. Multiscale change detection

Change/no-change theme maps resulting from segmentation are then extracted using the LM logical operation on the already created theme maps. The logical AND binary operator separated no-change areas of forest and non-forest areas, and logical SUB operation was used to separate change areas (forest \( F \) to non-forest \( nF \), and non-forest \( nF \) to forest \( F \)). With this approach, areas with different properties (unknown) are mapped out automatically. A composite of the oldest image (TM bands 4 and 5), and recent image channel 1-level3 of ETM+ (Fig. 17a), is used to visualize the change/no-change. This visualization approach is synonymous to the multi-temporal hyper-clustering, whereby a post-classification comparison is carried out between the first time \( t_1 \) data and \( t_1 \) plus second time \( t_2 \) data sets. The composite in Fig. 17a presents a quick and rough indication of the predominant change/no-change areas, but cannot be wholly relied upon especially for small areas of change or no-change. Fig. 17a is cross-correlated to the “color” information in Fig. 17b, and thus acts as complementary information to the derived change/no-change information depicted in Fig. 17b.

The resultant change/no-change map in Figs. 17b represented the following four thematic classes: (1) no-change for forest areas, (2) no-change for non-forest areas, (3) change of forest to non-forest areas and (4) change of non-forest to forest areas.

A summary on the proportion statistics is illustrated in Fig. 17c. The legend of the classes in Fig. 17c is explained as follows: (i) non-forest—soil, pasture and crops; (ii) forest-unchanged—primary forest (preserved tropical rainforest); (iii) forestation—needleleaf old and young woody secondary forest cover; and (iv) deforestation—same as non-forest. It was verified that reforestation took place with different species of trees, i.e. from natural broadleaves forest to needleleaves pine forest. Thus the observed difference is as a result of leaf structure and tree height textures of the different tree ages and species. It is also possible to conclude that part of the forested land cover had been cleared in-between 1986 and 2001, and then reforestation took place. While the excised forest cover is predominantly used for subsistence farming, the other parts of the forest are illegally logged for timber and charcoal burning (fuel). Detailed investigation into the causes of change is not within the scope of the current study.

6.3. Change detection results accuracy assessment

To estimate the accuracy of the proposed method, ground reference data and aerial photos of this scene were used as reference information (Table 1). Where ground surveys could not be achieved, aerial photos were relied upon, and where both were not available...
the knowledge from the forest master and near-true false color composites were used. For each region of homogeneous pixels, stratified sampling formulae were applied to estimate the error matrix cell proportions, from which the estimates of overall accuracy, kappa index and class-specific users’ and producers’ accuracy were derived. The results for the four classes (refer to Fig. 17c) are presented in Table 5.

The overall accuracy achieved was 88.4% (kappa = 0.8265). These results can be said to be fairly high or accurate owing to two factors: (i) the respective class accuracies were generally high as depicted by the users’ and producers’ accuracies (Table 5), and (ii) the collection of the ground-truth information is not easy in such forested land cover.

The overall accuracy achieved was 88.4% (kappa = 0.8265). These results can be said to be fairly high or accurate owing to two factors: (i) the respective class accuracies were generally high as depicted by the users’ and producers’ accuracies (Table 5), and (ii) the collection of the ground-truth information is not easy in such forested land cover.
7. Conclusions

The experimental results have shown that the proposed anisotropic diffusion scheme can effectively remove background non-informative features, and yet maintain sharp edges of anomalies in the filter image. The change detection and recognition system presented in this study could reliably detect and identify deforested, forested and no-changes as per the accuracy assessment results. Derivation of the optimal scale is observed to result in the detection of homogeneous forest patches, which are much easier to process. This observation may however vary depending on the scene characteristics. Post-analysis of the derived forest patches is much easier and more accurate, due to the reduced search space, than in the original bands.

A number of choices have to be made, which in effect influences the reliability of the results. These are (a) selection of the appropriate bands for texture and spectral information extraction; (b) selection of the suitable or optimal detail level to be used for change detection; and (c) threshold selection, from the smoothened optimal level bands, to discriminate between forested and non-forested areas. From the theoretical standpoint, in comparison with other change detection methods, the presented technique differs from the traditional methods presented in Singh (1989) because no final classification is produced. Instead the LM concept is adapted to model possible changes from no-changes. The results obtained here are considered satisfactory for further analysis or implementation into wider related geographic areas. The inherent limitation of the anisotropic diffusion model is that the convergence of the diffusion process is time-consuming, thus a faster convergence strategy is being investigated.

The following specific and general conclusions and insights are derived from the results of this study:

(a) for a given scene, the derivation of an optimal scale for deriving and representing candidate features/objects is more appealing for the subsequent applications;

(b) the multilevel segmentation in this case is an unsupervised process, so there is no need to select training data. This implies rapidity, cost and time effectiveness as compared with supervised methods;

(c) the simultaneous independent change detection using the LM approach frees the procedure from classification and post-classification processing cumulative errors;

(d) the explicit delineation of forestland cover by the presented approach can be used for baseline mapping and/or for updating existing geo-information databases. Different habitat maps can also be determined from the same scene based on specific cover derivation, and can be used in monitoring illegal activities in forests and for disaster monitoring in areas where there is little or no information available a priori;

(e) this approach offers a valuable and complementary approach that is similar to human operators, by creating regions instead of pixels or points as feature “carriers”. Each of these regions corresponds exactly to one and only one class. This object-oriented capability is much more appealing than the per-pixel analysis;

(f) further research needs to be carried out to give an in-depth knowledge about the inter-band-level characteristics, and about how they not only influence the change detection process but also impact on the classification after the anisotropic diffusion-based segmentation process; and finally

(g) future attempts to integrate other data like land cover probability and ancillary information will be tested for a more robust change information mapping, since the system has the capability to integrate multiple datasets about land cover through the MAD–SOFM system.

Acknowledgments

The authors wish to thank the Japanese Ministry of Education, within the framework of MOM-BUSHO for sponsoring this study.

References


