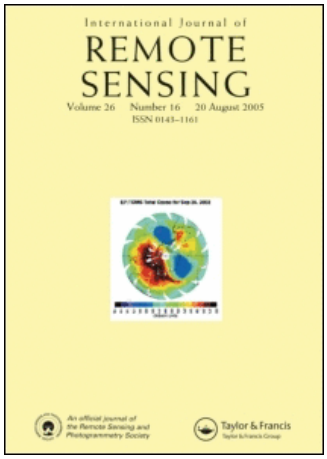

International Journal of Remote Sensing
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713722504

Analysis of co-occurrence and discrete wavelet transform textures for differentiation of forest and non-forest vegetation in very-high-resolution optical-sensor imagery

Yashon O. Ouma [a]; J. Tetuko [b]; R. Tateishi [b]
[a] Tateishi Lab, CEReS, Graduate School of Science and Technology, Chiba University, Chiba, 263-8522, Japan
[b] Center for Environmental Remote Sensing (CEReS), Chiba University, Chiba, 263-8522, Japan

Online Publication Date: 01 June 2008

To cite this Article: Ouma, Yashon O., Tetuko, J. and Tateishi, R. (2008) 'Analysis of co-occurrence and discrete wavelet transform textures for differentiation of forest and non-forest vegetation in very-high-resolution optical-sensor imagery', International Journal of Remote Sensing, 29:12, 3417 — 3456

To link to this article: DOI: 10.1080/01431160701601782
URL: http://dx.doi.org/10.1080/01431160701601782

International Journal of Remote Sensing
Vol. 29, No. 12, 20 June 2008, 3417–3456

Analysis of co-occurrence and discrete wavelet transform textures for differentiation of forest and non-forest vegetation in very-high-resolution optical-sensor imagery

YASHON O. OUMA*†, J. TETUKO‡ and R. TATEISHI‡

†Tateishi Lab, CEReS, Graduate School of Science and Technology, Chiba University, 1–33 Yayoi, Inage-ku, Chiba, 263-8522, Japan

‡Center for Environmental Remote Sensing (CEReS), Chiba University, 1–33 Yayoi, Inage-ku, Chiba, 263-8522, Japan

(*Received 13 October 2005; in final form 27 May 2007*)

The extraction of texture features from high-resolution remote sensing imagery provides a complementary source of data for those applications in which the spectral information is not sufficient for identification or classification of spectrally similar landscape features. This study presents the results of grey-level co-occurrence matrix (GLCM) and wavelet transform (WT) texture analysis for forest and non-forest vegetation types differentiation in QuickBird imagery. Using semivariogram fitting, the optimal GLCM windows for the land cover classes within the scene were determined. These optimal window sizes were then applied to eight GLCM texture measures (mean, variance, homogeneity, dissimilarity, contrast, entropy, angular second moment, and correlation) for the scene classification. Using wavelet transformation, up to five levels of macro-texture were computed and tested in the classification process. Comparing the classification results, (1) the spectral-only bands classification gave an overall accuracy of 58.69%; (2) the statistically derived 21×21 optimal mean texture combined with spectral information gave the best results among the GLCM optimal windows with an accuracy of 73.70%; and (3) the combined optimal WT-texture levels 4 and 5 gave an accuracy of 63.56%. The combined classification of these three optimal results gave an overall accuracy of 77.93%. The results indicate that even though vegetation texture was generally measured better by the GLCM-mean texture (micro-textures) than by WT-derived texture (macro-textures), the results show that the micro–macro texture combination would improve the differentiation and classification of the overall vegetation types. Overall, the results suggests that computer-assisted classification of high-spatial-resolution remotely sensed imagery has a good potential to augment the present ground-based forest inventory methods.

1. Introduction

There is the evidence of an ever-increasing need for accurate and cost-effective forest information acquisition for operational and strategic applications in ecological sustainability, forest exploitation, and rural development (Hyyppa *et al.* 2000, Hyyppa and Hyyppa 2001). Tropical rain forests (TRFs) are of special concern.

*Corresponding author. Email: yashon@graduate.chiba-u.jp

International Journal of Remote Sensing
ISSN 0143-1161 print/ISSN 1366-5901 online
http://www.tandf.co.uk/journals
DOI: 10.1080/01431160701601782

TRFs contain a substantial fraction of terrestrial biodiversity (Reaka-Kudla *et al.* 1997), their soils and biomass account for a large percentage of the total world terrestrial carbon pool, and they are estimated to account for roughly one-third of net primary productivity on land (Dixon *et al.* 1994). Further, understanding the distribution of different tree species in TRFs and hence distribution of different rates of transpiration would, for instance, give a clue to the availability and management of ground water.

The potential of applying remote sensing to forestry resource mapping and management has been recognized by several authors summarized in the following review articles (Woodcock *et al.* 1994, Trotter *et al.* 1997). The use of ground-based survey methods to obtain detailed inventory data, though still used, is both too expensive and time-consuming (Schneider 1989, Woodcock *et al.* 1994). The need for a less expensive and more comprehensive approach has prompted much research on extracting forest information from aerial and satellite remote sensing (Trotter *et al.* 1997).

Remotely sensed spectral data have been used to identify broad categories of forest cover, for example, conifer versus deciduous stands (Nelson *et al.* 1985, Shen *et al.* 1985, Hodgson *et al.* 1988, Lathrop *et al.* 1994). Woodcock *et al.* (1994) mentions the costs which can be saved, both in the long term and per unit area, and the high degree of consistency between locations that is offered by remote-sensing data. A number of studies have used broadband instruments like Landsat-Thematic Mapper (TM), Multispectral Scanner (MSS), SPOT HRV, and Very High-Resolution Radiometer (AVHRR) instruments to classify forest types with varying degrees of success (Frank 1988, Skidmore 1989, Franklin 1994, Schriever and Congalton 1995, White *et al.* 1995, Franklin *et al.* 2001a). Vegetation-species classifications have also been made with an airborne multispectral scanner (Rohde and Olson 1972) and video imagery (Everitt *et al.* 1987, Thomasson *et al.* 1994). Most related studies have been carried out on conifers, though reasonable accuracies have been demonstrated with non-coniferous trees (Thomasson *et al.* 1994). In the next sub-sections, the objectives of the present research are stated, and a review on the utility of the most current very-high-resolution spatial imagery in vegetation mapping is presented.

1.1 *Extraction of vegetation from very-high-resolution imagery*

The utility of very-high-spatial-resolution imagery (HRI) for semi-automated/automated vegetation composition classification needs to be evaluated. The utility of HRI has gained popularity in urban-related applications. However, there has not been as much work in detailed vegetation mapping (Benediktsson *et al.* 2003), and this needs to be investigated (Ehlers *et al.* 2003). This preference for urban areas is partly due to the proximity of the spectral signatures for different species and the difficulties in capturing texture features for vegetation (Carleer and Wolff 2004). While high-spatial-resolution remote sensing provides more information than coarse-resolution imagery for detailed observation of vegetation, an increasingly smaller spatial resolution does not necessarily benefit classification performance and accuracy. With the increase in spatial resolution, single pixels no longer capture the characteristics of classification targets. The increase in intra-class spectral variability causes a reduction of statistical separability between classes with traditional pixel-based classification approaches. Consequently, classification accuracy is reduced, and the classification results show a salt-and-pepper effect, with individual pixels classified differently from their neighbours (Yu *et al.* 2006).

1.1.1 Texture and vegetation extraction in HRI. A texture measure should describe both the primitives from which the feature texture is composed and the spatial relationships. However, the weakness of the commonly used grey-level co-occurrence matrix (GLCM) texture measures is that they do not capture the shape, here defined as the macro-texture, of the topographic objects of the grey-level primitives. This implies that valuable information is lost if the co-occurrence method is used by itself for derivation of texture measures. Thus, the significant questions posted in this research are: (1) can a texture analysis approach be modelled in such a way that the two basic texture characteristics, micro- and macro-texture features, are completely and simultaneously satisfied, and (2) is it possible to establish robust approaches for selecting optimal micro- and macro textures, i.e. texture measure source and type, texture window size for TRFs?

In this research, we demonstrate that a useful addition to the GLCM texture is to perform multiscale-multichannel decomposition for the extraction of the macro-texture elements for combination with micro-texture and spectral elements as cues in vegetation species (types) differentiation. The aim is to acquire shape and size information through wavelet transformation.

Given that the size and shape of vegetation types can best be derived from very-high spatial resolution data, the main objective of this study is to identify and map natural and planted tree species or types on part of Mt. Kenya from very-high spatial resolution (QuickBird) imagery, using the proposed texture-analysis approach based on the combination of micro- and macro-vegetation texture and in combination with vegetation spectral information. The specific objectives of this study are to assess the:

1. potential synergy of texture and spectral data from high-resolution satellite images, in order to classify spectrally complex landscapes like forest environments, by comparison and analysis;
2. ability of combined micro- and macro-texture analysis techniques in digital discrimination of different vegetation types from high-spectral-resolution optical multispectral data.

1.1.2 Review: HRI and vegetation mapping. The use of space-borne remote-sensing data has been hindered by the coarse-spatial-resolution satellite imagery as mentioned in §1. Until recently, forest managers had continued to use almost solely aerial photographs and field surveys (Gillis and Leckie 1996, Muinonen *et al.* 2001). The advent of very-high-resolution multispectral as well as panchromatic imagery appears particularly promising as a source of information for forest inventory survey tools. With the launch of QuickBird in 2001, sub-metre-resolution panchromatic and less than 2.5-m multispectral imagery has become commercially available from satellite-based sensors. Image analysis at this resolution enables the identification of parts of trees and hence individual trees. In practice, however, it has been proved to be difficult to classify these very-high-resolution images on a pixel-by-pixel basis due to the high level of information captured by these images (Puissant *et al.* 2005, Yu *et al.* 2006). With regards to automatic analysis of such data, actual cases of practical application are still hardly established (Leckie and Gougeon 1999, Wulder *et al.* 2000, Culvenor 2002, Ehlers *et al.* 2003, Yu *et al.* 2006).

Considering well-established approaches for analysing classical optical remote-sensing data from Landsat TM and SPOT HRV, very-high-spatial-resolution data like those from QuickBird or IKONOS have a range of particularities. First, Toutin

and Cheng (2000) showed that an accurate geometric correction of IKONOS-2 data requires the use of a rigorous model or at least the rational polynomial model and a DEM, and that the sub-pixel accuracy which may be obtained with satellite sensors such as SPOT HRV and Landsat TM/ETM + will not be achievable for IKONOS-2 data, even for flat terrain.

Second, as the spatial resolution increases, the internal radiometric variability within meaningful objects for land-cover mapping increases (Aplin *et al.* 1999). At the spatial resolution of 1 m or sub-metre, the per-pixel classification of forested land cannot be conducted without taking into account the spatial context information. The radiometric value of an isolated pixel provides little information, since the objects of interest are much larger than the pixel size. Thus, a per-object or per-parcel classification is required. Even for imagery with a lower resolution, the per-object classification may give better accuracy than the broadly used per-pixel approach (Kilpelainen and Tokola 1999, Lobo *et al.* 1996, Aplin and Atkinson 2001). This approach assumes that object boundaries are known. They can be derived from image-segmentation results (Lobo 1997, St. Onge and Cavayas 1997), from manual digitizing or from existing GIS data (Janssen and Molenaar 1995, Aplin *et al.* 1999, Smith and Fuller 2001). There is indeed an increasing use of GIS technology for forest management and topographic mapping in temperate zones so that parcel boundaries would be more easily available in the future. On the other hand, the extraction of parcel boundaries of image segmentation are not yet well established, even if there are significant improvements with using morphological tools to deal with the spatial context (Definiens 2000).

Third, one must pay particular attention to texture analysis of very-high-spatial-resolution imagery. At a spatial resolution of about 1 m or less, texture analysis becomes as important as spectral analysis for land-cover classification (Gougeon 1996, Bakker *et al.* 2001, Atzberger 2004). Franklin *et al.* (2001b) showed that a higher accuracy was obtained using texture data alone than using spectral data alone when classifying forest stands in CASI imagery of a spatial resolution of about 1 m. Similar results pointing out the significance of the texture of such data were reported by Franklin *et al.* (2000). Fundamental to the discrimination of tree species from HRI is the extraction of signatures from the crown (Lucas *et al.* 2004). Lewis (1994) exposed a methodology for vegetation classifications to spectral classifications in order to map variation in species composition within natural vegetation of the Australian spinifex hummock grassland.

In the analysis of forest HRIs, Pinz (1989) suggested that it is no longer applicable to work with pixels as the basic units. The natural approach is to detect visible single trees as light image objects and use them as individuals in the subsequent analysis. The estimated tree crown dimensions are appropriate to predict the stem dimensions (Minor 1951, Jakobsons 1970). Further measures of the image objects might be used as a complement.

Several other methods have been developed in an attempt to delineate trees from HRI. Some methods have been based on local maxima or also referred to as the valley-following approach (Gougeon 1998, Warner *et al.* 1998) and regions (Brandtberg and Walter 1998, Culvenor *et al.* 1998). However, these methods are most robust for detecting trees which are partly in shadow and pre-processing is necessary to avoid detecting light patches in lower density areas. Pollock (1994) presented a model for matching, training, and spatial information to locate tree crowns with the advantage of avoiding improbable regions. The disadvantage of this approach is that it is in part manually based.

1.2 *Theoretical background on vegetation-type differentiation*

Tree canopy spectral response and texture provide important clues in tree species mapping. Canopy roughness refers to the variation in vertical structure of the canopy as illustrated in figure 1, and reflects the tree surface texture. As seen in the hypothetical scenarios in figure 1(*a*), it is possible to have rough and smooth texture over the canopies, and these canopy-texture characteristics correspond to different vegetation types. In figure 1(*b*), the geometric and spectral patterns over the coniferous pine and deciduous (broadleaf) tree types are hypothesized, as in the case of Pollock (1994). The deciduous broadleaf trees may, for example, take a near-uniform over the surface reflectance pattern, as opposed to coniferous trees. It is argued here that different trees closely exhibit similar spectral patterns, but with respect to texture, the patterns are species-biased, and consequently a single texture model may not suffice for the different texture patterns exhibited by different tree species. Patterns are also dependent on the age and condition of the trees. However, this may be difficult to derive for wide-natural areas.

Generally, the registered spectral reflectance patterns determine the textural aspects of every tree species. That is, a stand with a large variance in tree height has a rougher upper canopy than a stand with a small variance in tree height. Young and even-aged stands of pure species tend to have smoother upper canopies than mixed-aged and mixed-wood stands. Larger shadows of trees are cast in stands of uneven height than in stands of even height. Furthermore, in stands of varying tree heights, the shadows are less uniform in their spatial distribution. These differences in tone and texture between smooth and rough canopy surfaces can be distinguished

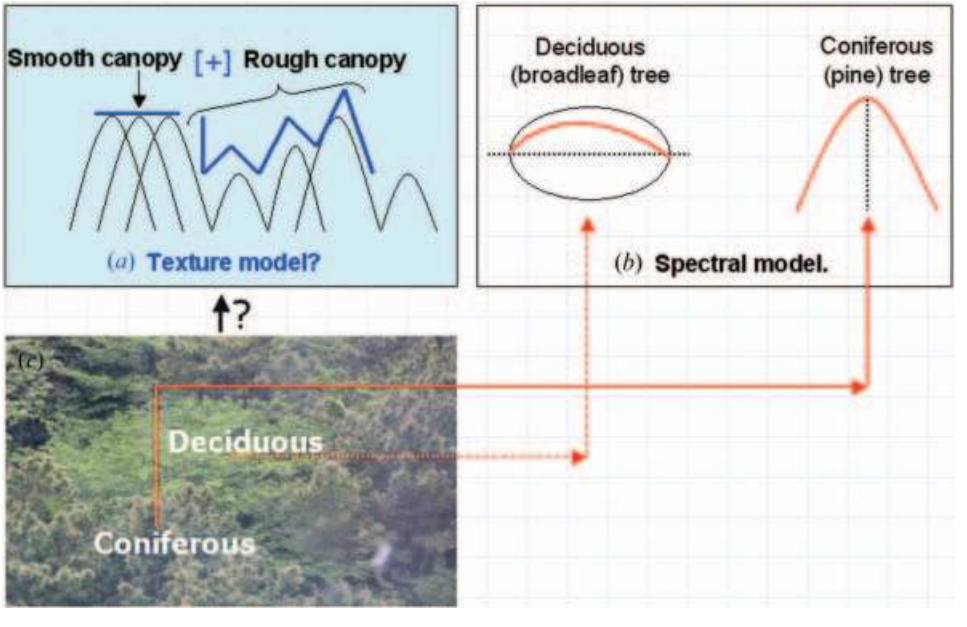

Figure 1. Theoretical plan view of: (*a*) possible tree type(s) and canopy texture-whether coniferous or deciduous; (*b*) spectral patterns/intensity for coniferous (pine) and deciduous (broadleaf) trees; and (*c*) oblique photo of deciduous and coniferous trees. The question mark (?) between (*a*) and (*c*) denotes the lack of models for comprehensive textural analysis in vegetation mapping, more so in new very-high-spatial-resolution optical satellite sensors like QuickBird.

in high-spatial-resolution remotely sensed imagery by using their texture patterns and can adequately aid in discriminating tree types. Conclusively, the overall shape of the tree species or type may also significantly contribute in tree-species discrimination.

While micro-textures can be used to represent a texture sample, they are spatially localized and do not characterize global attributes of global textures. For instance, consider the textures in figure 1(*c*); the visually observed textural differences could be as a result of the arrangement of the different sizes of tree leaves or tree branches referred to in this study as macro-texture. This provides one way by which different vegetation types or tree species can be differentiated. For classification, a better texture model may be derived from the combination of micro- and macro-model parameters than from micro-textures alone. The conceptual differences represented in figure 1 suggest that it may not always be simple and direct to integrate the spectral and co-occurrence textural information and obtain good classification results in tree species mapping. This may be attributed to the fact that some canopy structures exhibit micro-textural behaviours, others exhibit macro-textural characteristics, while others exhibit both the textural characteristics. Other complexities arise from spacing of tree crowns. Crowns with interlocking branches are challenging to delineate, and the complexity of stand composition and understorey vegetation makes the background complicated.

The fundamental question then is how to define, identify, and integrate the suggested micro- and macro-textural patterns so that we do not lose information or include irrelevant information in discriminating among vegetation types. In general, even though combining the per-pixel approach with spatial-or-texture processing approaches has been shown to provide additional accuracy, the reduction of the accuracies observed in some cases may be attributed to the following factors:

1. Insufficient utility of the relevant spatial features or structural information in the classification process.
2. Insufficient sensor spatial resolution for deriving the informative spatial attributes.
3. As identified in this study, the Gray Level Co-occurrence Matrix (GLCM) (Haralick *et al.* 1973) textures are restricted to the analysis of spatial interactions over relatively small neighbourhoods on a single scale. As a consequence, their performance is best suited for the analysis of micro-textures only.
4. The general lack of the utility of both spatial (texture) and topographic-feature-based shape information in high spatial resolutions images, as identified in this study. Shape may be viewed theoretically as an element of macro-texture. Shape derived by manual digitization or Geographic Information Systems (GIS) has been included in other studies but not related to vegetation species mapping in HRI.

In both practical and theoretical modelling of tree types, the shape and the size of the tree are significant components of the formulations. For example, if we model the basic tree crown as a generalized ellipsoid that in the (*x, y, z*) Cartesian coordinates has the surface defined by the following equation (1):

$$\left[\frac{\left(z^2\right)^{n/2}}{a^n}+\frac{\left(x^2+y^2\right)^{n/2}}{b^n}=1\right], \tag{1}$$

the n in equation (1) corresponds to the *shape* parameter. The other variables of course correspond to the standard ellipsoid definition parameters.

In this study, vegetation-type discrimination is based on the integration of spectral and optimized GLCM and wavelet-derived texture measures in the classification process. A strategy for GLCM selection is used to determine the most suitable micro-texture measure and window size. A wavelet filter bank is used to split the input image into different directional sub-bands to derive the macro-texture. Then the optimal texture features based on co-occurrence matrices and wavelet are determined using semivariance geostatistical fitting and classification accuracy reports. The organization of the rest of the paper is as follows: §2 describes the study area and data used in this research. §3 introduces the background theory for the co-occurrence matrices textures, the wavelet transformation and the proposed combined texture approach (methodology). Implementation details, experimental results and analysis of the results are summarized in §4. §5 presents the discussions and research conclusions.

2. Study area and data

A section of Mt. Kenya forest in Kenya was selected for this study. The study site (figure 2) has a rich vegetation diversity, including the natural afromontane forests, mostly camphor trees, and planted vegetation like coniferous pine trees and cash crop trees like tea plantations (non-forest vegetation). Other land covers found

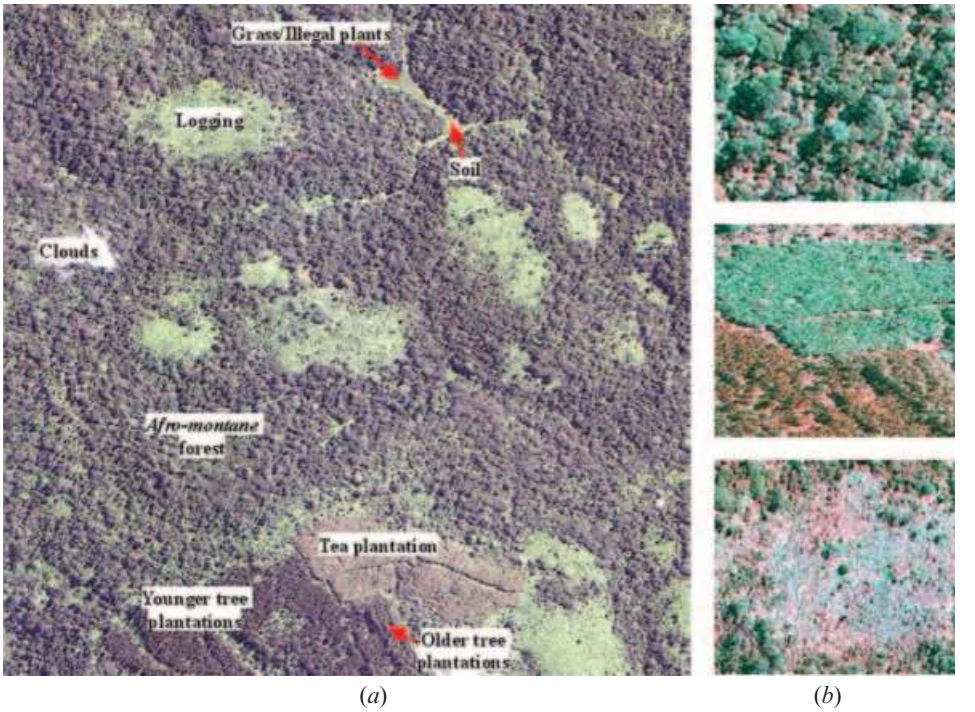

Figure 2. (*a*) Different vegetation types within the study area. (*b*) Afromontane forest camphor trees (top); tea and pine (middle); and logged areas (bottom). The scene comprises vegetated (natural and planted) and non-vegetated land cover. Not to scale.

within this site are logged areas consisting of understorey non-forest vegetation like grass and shrubs and bare soil surfaces (non-vegetated land cover). This scene, apart from being rich in vegetation diversity, was also chosen due to its accessibility for field validation. Figures 2(*a*) and 2(*b*) detail some of the test-scene vegetation covers. The QuickBird image in figure 2 was acquired on 5 September 2004 using the QuickBird (QB02) satellite sensor.

The entire location was approximately 5 km × 5 km in size with the image central coordinates located at −0.3300° (S) in latitude and 37.5600° (E) in longitude. Only part of the whole imagery shown in figure 2 was used in this study. The entire scene was re-projected to the local UTM coordinate system (zone 37, South) on the WGS-84 datum, and resampled to 1-m spatial resolution using nearest-neighbour convolution before sub-setting the 2048 pixels × 2048 pixels scene shown in figure 2(*a*). The resampling to 1 m is because of the utility of the dyadic-orthogonal discrete wavelet transforms discussed later in the paper.

The QuickBird data used in this study consisted of the pan-sharpened multispectral channels. The availability of a 0.61-m panchromatic band, in conjunction with 2.44-m IR bands, affords the opportunity to create an effective 0.61-m IR pan-sharpened image through a fusion technique. The PANSHARP algorithm (PCI Geomatics' exclusive pan-sharpening algorithm to QuickBird data) 'fuses' the high-resolution panchromatic and lower-resolution multispectral imagery to create a high-resolution colour image. The high-resolution colour image preserves the original colour fidelity and allows for a better visualization and interpretation. Comparatively better visual results are obtained from this fusion technique than from other transformations especially in tree crown delineations, and the results are preferred to the original bands (PCI 2001, Keeletsang 2004).

Other data used in this research comprised Landsat ETM+ acquired on 21 February 2000, covering the same area as the QuickBird data, and were used for general reconnaissance. The coarser resolution of ETM+ compared with QuickBird was useful in the rapid depiction of the general distribution of different vegetation patches. This substantially eased the ground data sampling process in time economy, especially in such difficult topography. Using aerial photography and differential global positioning systems (DGPS) data from the Kenya Wildlife Services (KWS), locations of ground data points including specific land cover within the scene were captured.

2.1 *Training data analysis*

Scene-representative training data were identified for the investigation and understanding of the spectral characteristics of the features within the test area. In figure 3, the spectral signatures of the 'training samples' are presented for blue, green, red, NIR bands and NDVI (=(NIR−red)/(NIR+red)), of the pan-sharpened QuickBird image. These 'training samples' represented vegetation/non-vegetation land cover within the study area as marked out in figure 2: clouds; afro-montane forest; planted pine trees–older; planted pine trees–younger; tea plantations, logged areas; soil surface and grass cover. Clouds and bare soil surfaces are used in this study only to analyse the behaviour of '100% pure' or homogenous surface signatures.

The differences and similarities between the 'training samples' are illustrated in figure 3. All the curves in figure 3 appear to be similar or to overlap in the compared wavebands except for clouds and bare soil, which of course is because of the absence

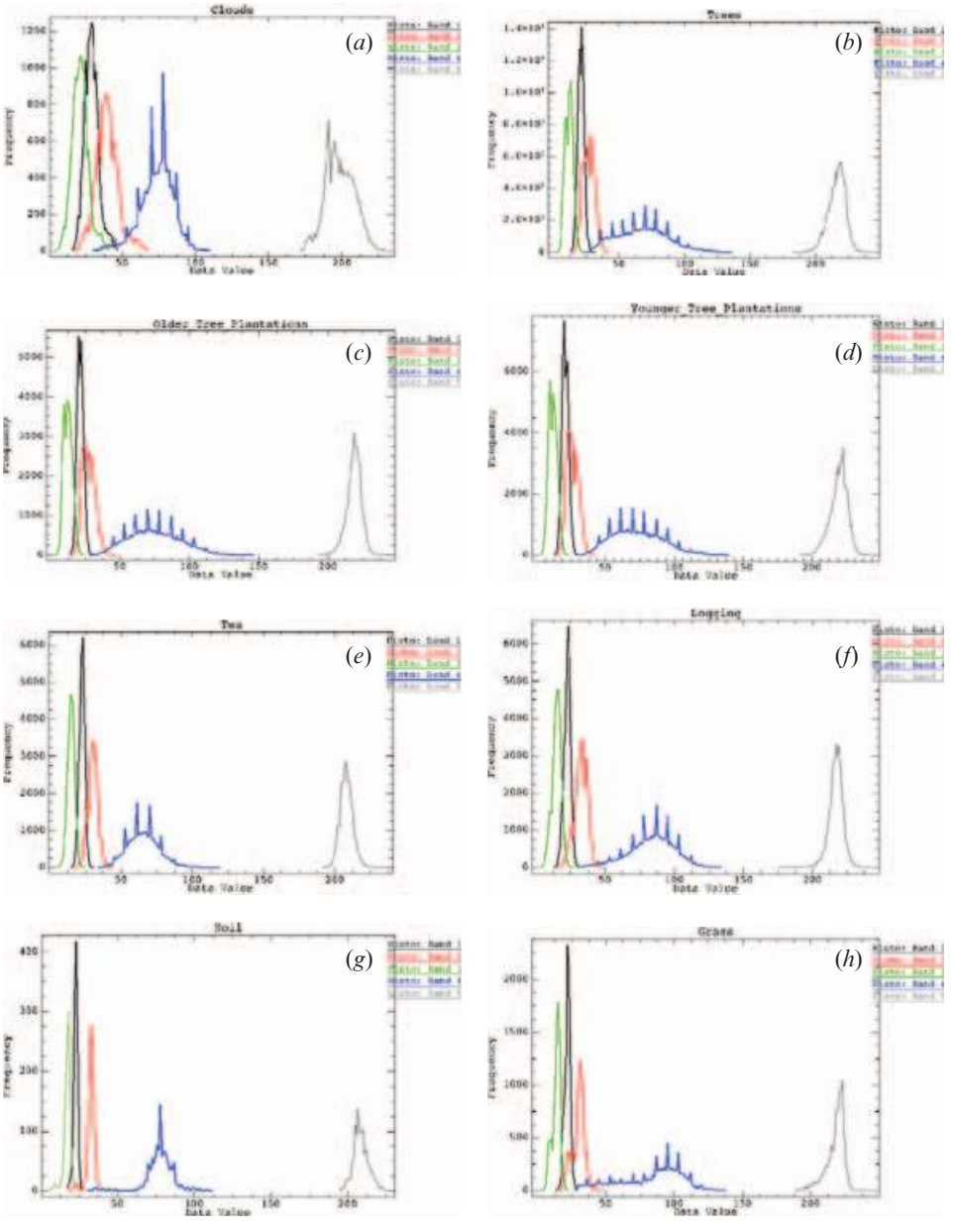

Figure 3. Histograms of the training areas as recorded in pan-sharpened QuickBird imagery: band 1 (blue), band 2 (green), band 3 (red), band 4 (NIR), and NDVI (represented as band 5) for the different training samples.

of green matter (chlorophyll). The degree of confusion in the visible bands is very high, thus showing that the visible bands are by themselves not suitable for the separation of vegetation types. It is noted that even in the NIR and NDVI, where vegetation is supposed to be most sensitive, there is a high degree of spectral confusion, as observed from the pattern similarities, thus reasserting that spectral information alone may not be wholly relied upon for accurate and complete

vegetation differentiation. NDVI measures the amount of green vegetation and consequently does not capture the unique features of specific vegetation conditions and characteristics and/or structures. These preliminary investigations suggest that neither the NIR band nor NDVI can be relied upon as significant measures for vegetation spectral differentiability.

For reliable forest texture extraction, suitable or relevant band(s) must be identified so as to minimize the inclusion of non-tree textural primitives within the scene like soil, cut-trees stems, etc. To determine which of the QuickBird bands available for this research is suitable for the spatial information (texture) extraction, the regressions between the NIR-band versus the blue, green and red bands were first compared. The following regressions with the NIR band were obtained: (a) 0.71 (blue), (b) 0.85 (green), and (c) 0.79 (red). All the visible bands showed a high correlation (R^2>0.70) with the NIR band. Second, NIR gave the highest average of the variance within each training sample class, by comparing similar sizes of the training samples in each of the four bands. This is in part due to the fact that the location and narrower wavelength of the NIR band capture better vegetation structural information than the other compared wavelengths and the wider panchromatic band, respectively (Ouma *et al.* 2006). Other studies have used a combination of the VIS and NIR for vegetation texture extraction from HRI (e.g. Zhang 2001). In vegetation analysis, the wavelength of the panchromatic band is argued to be too wide, and the panchromatic band image contains extra information that can be viewed as noise for vegetation texture mapping. Thus, the optimal or suitable multispectral band is derived for vegetation texture extraction. It is worth pointing out here that the selection of a suitable textural band is an open research question and varies from scene to scene (Zhang 2001, Asner *et al.* 2002). In this study, it was found that the texture or spatial pattern structure of the vegetation classes was resident in the NIR band. Detailed experimentation on the derivation of textural band(s) for vegetation textural primitives' extraction can be found in Ouma *et al.* (2006).

3. Methodology

Texture analysis has a history of almost three decades. During the 1970s and early 1980s, the algorithms have been mainly based on first- and second-order statistics of the image pixel grey-level values as a spatial domain grey-level co-occurrence matrix (SDCM) and neighbouring grey-level dependence matrix (GLCM). In the mid-1980s, model-based methods such as Markov Random Fields (MRF), simultaneous autoregressive models, and the Gibbs distribution appeared as alternatives. From the late 1980s, based on the theoretical impact of the works of Daubechies (1988), who provided the discretization of the wavelet transforms (WTs), and Mallat (1989), who established the connection between the WT and the multiresolution theory, signal-processing methods based on the Gabor transform and the WT rapidly competed with the former two in the fields of computer vision and image processing. By representing signals, or images as in our case, in multiple resolutions by the WT, it is thought that one can extract more powerful features than the single scale case.

3.1 *Mathematical framework*

3.1.1 Discrete wavelet transformation. The wavelet transform provides a multi-resolution decomposition of an image in an orthonormal basis and results in a

non-redundant image representation. Wavelet transforms extract information from an image at different scales (sub-band). A family of basis functions can be generated by translating and dilating the *mother wavelet* corresponding to that family (Mandal and Aboulnasr 1996). The wavelets form a family, and the basic form is called the mother wavelet. All the daughter wavelets are derived from this wavelet (Ψ) according to equation (2):

$$\Psi_{s,\tau}(t) = \frac{1}{\sqrt{s}}\Psi\left(\frac{t-\tau}{s}\right) \tag{2}$$

The two variables, s and τ, are the scale and translation of the daughter wavelet, and t is the theoretical time element. The term $s^{-1/2}$ normalizes the energy for different scales, whereas the other terms define the width and translation of the wavelet. The Continuous Wavelet Transform (CWT) is defined in equation (3), whereby the asterisk denotes a complex conjugate function.

$$\gamma(s,\tau) = \int f(t)\Psi^*_{s,\tau}(t)\mathrm{d}t. \tag{3}$$

As good as the theory of the CWT is, it still has three major problems. These problems make the continuous wavelets difficult to implement for solving any real problem. These are:

1. Redundancy—the basis functions for the continuous wavelet transform are shifted and scaled versions of each other. It is clear that these cannot form a very orthogonal basis;
2. Infinite solution space—the result holds an infinite number of sub-spaces. This makes it hard to solve and even harder to find the desired results out of the transformed data; and
3. Efficiency—most of the transforms cannot be solved analytically. The solutions have to be calculated numerically, which takes an incredible amount of time. In order to use the wavelet transform for rational application(s), we must find very efficient algorithms.

The redundancy in the continuous wavelets can be fixed. First, the continuous variables s and τ are discretized, hence a discrete wavelet transform (DWT), so that the mother wavelet can only be scaled and translated in discrete steps. The discretized values of τ are made dependent on s in such a manner that low-frequency components are sampled less often. As a result of the discretization, we obtain equation (4) for deriving daughter wavelets:

$$\Psi_{j,k}(t) = \frac{1}{\sqrt{s_0^j}}\Psi\left(\frac{t-k\tau_0 s_0^j}{s_0^j}\right). \tag{4}$$

s_0^j is usually assigned a value of *2* and t_0 a value of 1. This results in dyadic sampling, hence the multiresolution products of the wavelets transformation process, represented in figure 4.

This characterization of the wavelet transform allows the study of an image from fine to coarse resolution and the extraction of information in any of the levels of decomposition. Further details on the mathematical representation and applications of the DWT can be found in Mallat (1989) and Ouma *et al.* (2006) respectively.

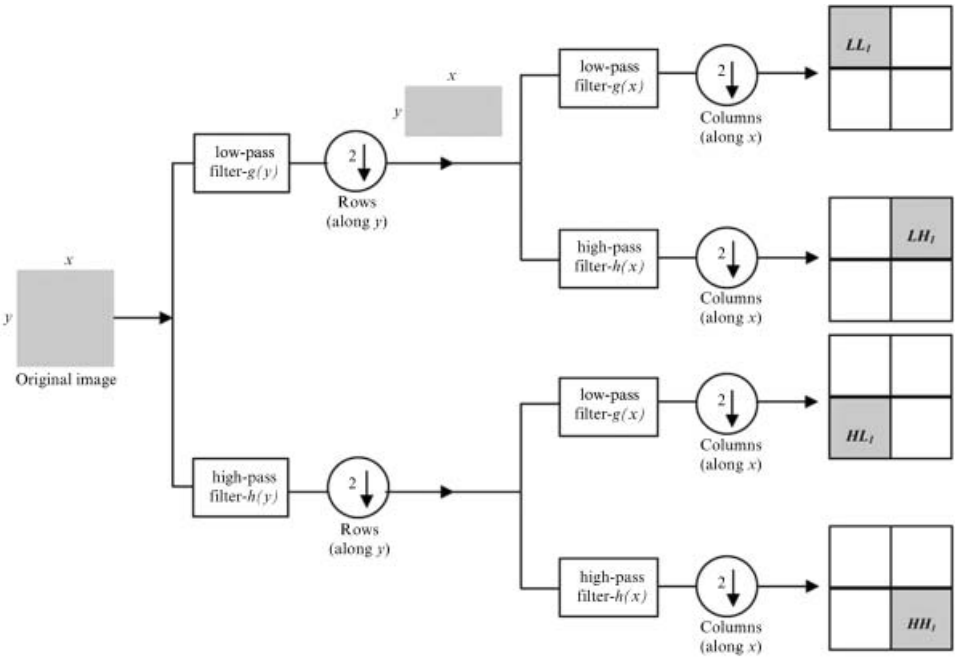

Figure 4. Generalized computational approaches for wavelet transform (WT).

The underlying idea in the use of the multiresolution analysis is first to obtain multiple evidences from the same feature, and search for the sub-band(s) that correctly, with respect to resolution, shape, and size, defines the features of interest. Second, our approach follows the paradigm of fusion that utilizes multiple evidences of the same feature(s) resident in multiple sub-bands or scales. A useful property of the wavelet transform is that the wavelet coefficients are sensitive to spatial variations of a signal and can isolate details at a given scale. This is justified by the fact that a signal with a high variation has high wavelet coefficients and vice versa. Compared with other frequency-domain-based transformations like the Windowed Fourier Transform (WFT) (e.g. Gabor), WTs offer superior image-computational abilities such as: the resulting mutually orthogonal coefficients, reversibility, and the fact that the filters used remain the same at all scales (van de Vower *et al.* 1999). Therefore, in this study, we use the 2D discrete wavelet transform to extract multiple sub-band vegetation with macro-texture information.

An algorithm for implementing the wavelet decomposition follows the Mallat (1989) approach, with a two-channel filter bank using quadrature mirror filters. The algorithm applies a one-dimensional high- and low-pass filtering step to both the rows and columns to the input image. Each filtering step is followed by sub-sampling, which results in a change in scale. The filter bank to implement the wavelet transform is shown in figure 4. At each decomposition level, there are four different output images.

The information contained in the output sub-bands of the wavelet transform is the approximation of the input image and three detail images conventionally named LL_1, an approximation of the input image, and HL_1, a detail image containing information on edges in the horizontal. A high value indicates the presence of a vertical edge. With LH_1, the detail image containing information on edges in the

vertical, a high value, indicates the presence of a horizontal edge. With HH_1, the detail image containing information on edges, a high value indicates the presence of a corner point. In the second level of decomposition, the low-pass sub-band, LL_1, is further decomposed into one low-pass and three high-pass sub-bands (horizontal, vertical, and diagonal direction). The process can be repeated on the low-pass sub-band to form a higher-level wavelet decomposition. In this study, the *Daubechies* eight-tap filters (*db8*)-DWT (Daubechies 1988) were chosen, since they are compactly supported and well localized in both time and space.

3.1.2 Grey-level co-occurrence matrix (GLCM). Statistical relations between pairs of pixels are called second-order statistics. The statistical behaviour of two pixels in an image can be described by the joint probability density function, which fully defines the second-order statistics. Nearest-neighbour grey-tone spatial-dependence matrices or grey level co-occurrence matrices (GLCM) as proposed by Haralick *et al.* (1973) provide measures from which the texture features are derived. The GLCM has grown to be a standard technique for extracting texture characteristics. The technique works by forming a window on the image and then calculating the frequency of co-occurrence for the pixel values (DN). If the pixels' grey values range from $m \times n$, then the matrix will be of dimensions $m \times n$, and the frequency of co-occurrence with pixel value i and j will be put into the entry (i, j). The GLCM $P(i, j)$ is defined by specifying a displacement vector and counting all pairs of pixels separated by distance d and direction ϕ having a grey level (i, j).

Once the counting for the co-occurrence frequencies of all the bi-pixels within the window has been completed, one can then design statistical measures to extract the characteristics of the matrix. The resulting measures reflect the grey-values variation (i.e. texture) within the prescribed window. For example, if the area covered by the window is relatively smooth, the resulting GLCM will hold peaks along the main diagonal. Likewise, if the pixels' values within the window are nearly random, it will form a GLCM with a similar frequency for all the entries. The GLCM has been shown to be very successful in capturing image textures (Anys *et al.* 1994, Treitz *et al.* 2000, Arzandeh and Wang 2002).

The window size, directions, and lag value selected for GLCM will vary according to the problem at hand and the spatial scale for the features of interest. Normally, the lag value of 1 is chosen (i.e. neighbouring pixels), and four directions, namely horizontal, vertical, left diagonal, and right diagonal, are used for forming GLCM. In our study, after the GLCM is generated for each direction, eight statistical measures are used for texture extraction. The four directions are then averaged to remove directional effects, hence omnidirectional texture derivation. Details on the eight-texture measures analysed in this work are summarized in table 1. A more complete theoretical description of the most commonly used co-occurrence measures can be found in Haralick *et al.* (1973) and Soares *et al.* (1997). To evaluate the window sizes, the selected 'training samples' were used. To carry out the training samples window evaluation, the eight GLCM texture measures (table 1) are computed for 20-window sizes ranging from 3×3 to 41×41.

3.2 *Texture feature extraction*

A block diagram of the proposed combined spectral and macro–micro texture classification approach is shown in figure 5. The significant steps in figure 5 are described in subsequent subsections.

Table 1. GLCM texture measures tested in this study.

GLCM texture measure	Description
Mean: $\text{Mean}_{\Delta x \Delta y} = \sum_i ip(i)$	Mean is the average grey level in the local window.
Variance: $\text{Var}_{\Delta x \Delta y} = \sum_i \sum_j (i-\mu_i)^2 P(i,j)$	Grey-level variance in the local window. High when there is a large grey-level standard deviation in the local region.
Entropy: $\text{ENTH}_{\Delta x \Delta y} = -\sum_i \sum_j P(i,j)_{\Delta x \Delta y} \log P(i,j)_{\Delta x \Delta y}$	Entropy is a measure of the degree of disorder in an image. Entropy is larger when the image is texturally non-uniform or heterogeneous and approaches its maximum when all GLCM entries have similar contents, indicating an image with completely random pixel values. High entropy when GLCM have relatively equal values and low when the elements are close to either 0 or 1.
Angular Second Moment (ASM): $\text{ASM}_{\Delta x \Delta y} = \sum_i \sum_j P(i,j)^2_{\Delta x \Delta y}$	Also called energy, angular second moment and uniformity is a measure of textural uniformity or pixel-pair repetitions. When the pixels of the image window under consideration have similar grey levels, the energy reaches its maximum (equal to or close to 1). Therefore, a constant or periodic distribution of grey levels over the window will produce high values for energy. It is high when GLCM has few entries of large magnitude, when all entries are almost equal. This is the measure of the local homogeneity. Entropy and ASM are inversely correlated.
Homogeneity: $\text{Hom}_{\Delta x \Delta y} = \sum_i \sum_j \frac{P(i,j)_{\Delta x \Delta y}}{1+(i-j)^2}$	Also called inverse difference moment, homogeneity is a measure of lack of variability or the amount of local similarity in the scene. High homogeneity values suggest a small grey tone differences in pair elements. In this case, the associated GLCM will present elements around the main diagonal. Homogeneity is high when GLCM concentrates along the diagonal. This occurs when the image is locally homogenous in the scale of the length of spatial
Contrast: $\text{Cont}_{\Delta x \Delta y} = \sum_i \sum_j (i-j)^2_{\Delta x \Delta y} P(i,j)_{\Delta x \Delta y}$	Contrast is a measure of the degree of spread of the grey levels or the average grey level difference between neighbouring pixels. The contrast values will be higher for regions exhibiting large local variations. The GLCM associated with these regions will display more elements distant from the main diagonal, than regions with low contrast. Contrast is high when the local regions have a high contrast in the scale of spatial. Local statistics contrast and GLCM contrast are strongly correlated. Contrast and homogeneity are inversely correlated.

Table 1. (*Continued.*)

GLCM texture measure	Description
Dissimilarity: $\mathrm{Dis}_{\Delta x\Delta y} = \sum_i \sum_j P(i,j)_{\Delta x\Delta y} \lvert i-j \rvert$	Similar to contrast. High when the local region has a high contrast.
Correlation: $\mathrm{Cor}_{\Delta x,\,\Delta y} = \frac{\sum_i \sum_y (i-\mu_i)(j-\mu_j) P(i,j)_{\Delta x\Delta y}}{\sigma_i \sigma_j}$	Correlation is a measure of grey-level linear dependencies in the image. High correlation values denote a linear relationship between the grey levels of pixel pairs. A completely homogeneous area is a limiting case of linear dependency, for which the correlation reaches its maximum (equal to 1). Correlation is uncorrelated to entropy and energy, i.e. to pixel pair repetitions.

Where $p_x(i) = \sum_j p(i,j)$ and $p(i,j) = \frac{P(i,j)}{\sum_i \sum_j P(i,j)}$. Each element $P(i,j)_{\Delta x\Delta y}$ represents the relative frequency with which two neighbouring pixels separated by a distance of Δx columns and Δy lines occur (Soares *et al.* 1997).

3.2.1 GLCM feature selection. Traditionally, texture methods have been evaluated over windows of a single size, the latter being commonly defined on an experimental basis, e.g. in Puissant *et al.* (2005). Although many studies regarding the performance of the different families of texture feature extraction methods have been carried out in the past, only a few have dealt with the issue of determining

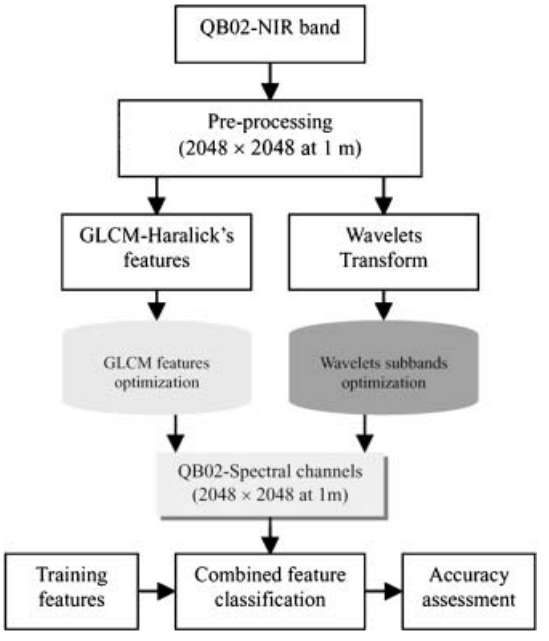

Figure 5. Generalized flow of the proposed combined texture classification approach. The term ''optimization'' is used here to infer the process of selecting and/or deriving the suitable band(s) or subband(s) accordingly. QB02 denotes QuickBird sensor-2.

optimal window sizes, especially in tree or vegetation-type differentiation. Within the scope of pixel-based vegetation-type discrimination in high-resolution imagery, a technique for determining the optimal window size for each ground cover that leads to the best possible estimator among different texture models of interest is necessary. Determining the optimal GLCM window via trial-and-error methods is time-consuming and may not always be reliable.

An optimal window size for calculating GLCM texture measures is a compromise between providing enough spatial information to characterize the land cover and limiting overlapping textures between different land covers (Ulaby *et al.* 1986). No rules have been recommended for the texture-measures selection. The most appropriate combination of texture features depends very much on the surface properties of the land-cover types of interest to a user. Since unique texture patterns are hypothesized to discriminate different land-cover types, a proper window size that approximately matches the patch size can extract the textural pattern of the particular landscape feature. A large window size can capture the spatial patterns of each land-cover type better but may contain more than one land category, which could introduce systematic errors. The window should then be small enough to keep the variance low and to maximize the potential for class separability.

With the assumption that the 'training samples' are homogenous and accurately represent the scene under investigation, it can theoretically be argued that the best GLCM texture should have the least variance over the training sample for the given image. Thus, the variance for each of the tested GLCM texture type is computed over the test windows and compared for each training sample. This texture measure, with the minimum variance, is then used to determine the optimal window size for each of the 'training sample' classes using the variogram fitting geostatistical method. This optimal window is assumed to be the same for the rest of the GLCM texture measures. The resultant optimal window texture types are then tested in the classification process to determine the suitable texture measure through accuracy assessments.

3.2.2 Wavelet transforms sub-bands selection. Suitable wavelet transforms sub-band selection is focused on evaluating the variations of parameters related with multiscale texture from wavelet decomposition like: influence of decomposition level (scale) and direction on the classification results. However, many features result from WT processing, and considerations as to whether the features are independent and discriminating must be carried out to avoid unnecessary and redundant computations.

For example, there are C_k^n possible combinations of k sub-bands from a total of n sub-bands. It is not practical to employ a brute force approach, which finds the best combination by trying out each one. This calls for deriving an efficient way of determining which of these combinations of sub-bands are optimal in terms of being best at discriminating different textures for subsequent classification. A reduction in the dimensionality of the problem may result in not only a better accuracy but also a faster classification (Ouma *et al.* 2006).

Theoretically, natural vegetation and other vegetation cover, unlike urban structures, exhibit multidirectional or omnidirectional textural primitives that cannot be represented accurately in a single direction. Thus, in this study, a faster and reliable approach to determine the relevance of the sub-bands is to compare the classification results for the sum of the three output channels per decomposition

level. This approach was adopted based on the conclusions from our earlier experimentation and experiences presented in Ouma *et al.* (2006).

3.2.3 Combined features classification and accuracy assessment. For classification, we chose a classifier which is already implemented, is well tested, and has shown a good performance in many practical applications. One such classifier is the Gaussian maximum-likelihood classifier (MLC) (Yu *et al.* 2006). Several other authors have applied maximum likelihood due to its more exhaustive and rigorous statistical formulation compared with other classifiers based on a statistical computational approach. The MLC is easy to implement, and its formulation is more robust than the other decision-rule-based classifiers. Based on the MLC, a supervised classification was adopted whereby groups of contiguous pixels were selected as training, testing, and validation samples in the class signatures. Note that it is not the aim of this study to investigate the performances of different classification algorithms.

The results of the GLCM and wavelet transform sub-bands selection phase are integrated with the spectral channels in the classification process, as shown in figure 5, and the classification results analysed for accuracy for the following combinations: (1) spectral-only classification; (2) spectral + GLCM; spectral + wavelet transform (WT); (3) and spectral + GLCM + wavelet transform (WT).

Fifty per cent of the collected ground data (test data set) were used for the accuracy assessment. The test points were carefully chosen to ensure that the test, and the training data were equally spread geographically. Each classified image was then crossed with the test data to generate a confusion matrix. The respective confusion matrices were then used to calculate the different accuracy measures, i.e. class producer and user accuracy, and the overall accuracy. Kappa statistics and its variance were also calculated to test the significance of difference in accuracy. The significance of difference test between the confusion matrices was done using the Z-test with $\alpha = 0.05$. A qualitative assessment of the classification results was carried out with field-data comparisons.

4. Results: texture measures, classification results and comparative analysis

This section presents the results of this study as well as a detailed comparison and analysis thereof in the following order: (a) scene texture derivation, for input texture band selection; and (b) a comparison of the classification results for spectral-only and spectral combined with the optimal texture band(s).

4.1 *Texture derivation results*

4.1.1 GLCM texture features. All the eight-texture measures showed differing results for the 'training samples'. There were, however, some common observations for all the 'training samples': (a) the texture measure with the least variance was consistently recorded as the angular second moment (ASM or S-Moment), and (b) the second least variance was recorded by the homogeneity texture. The rest of the textures were mixed with respect to the variance-based ranking. Thus, as hypothesized in the methodology, the ASM results were utilized to model or determine the optimal window size using Variowin 2.2 Geostatistics software (Pannatier 1996). Variowin 2.2 software is based on the variogram geostatistical fitting method for deriving the optimal widow sizes (Pannatier 1996, Curran and

Atkinson 1998). The results from this variogram fitting were then applied to the rest of the GLCM texture measures.

A summary of the ASM variance over the 20-moving windows is shown in figure 6. It is observed that apart from the clouds and soil, the rest of the 'training samples' had a lower variance and reached a saturation point after around the first nine or so windows i.e. around the 19×19 window size. Clouds and bare soil are 'pure' homogeneous surfaces whose spatial scale may be explained in terms of patch or grain sizes and not spatial scale. For this reason, it may be difficult to directly determine or measure the optimal scales for their texture mapping, unless they are considered as landscape patches. Surfaces with such characteristics require some modelling like the proposed semivariance fitting.

For each of the curves in figure 6, Spherical, Gaussian, and Exponential models were fitted to the variograms using Variowin 2.2 software. In table 2, the 'range' column depicts the optimal texture window size as derived from the semivariogram fit, 'sill' is proportional to object (class) variance, and the 'nugget variance' is the variance unexplained by the fitted curve (Pannatier 1996, Curran and Atkinson 1998). The quality of the semivariogram fit to the data 'goodness of fit (F)' was determined using regression (R^2) and an F-test calculated as in equation (5):

$$F = \frac{R^2}{1-R^2} \times \frac{N-k}{k-1}, \tag{5}$$

where N is the number of samples, and k is the number of variables in the regression model. The Spherical fitting model estimator gave the best fit, as shown in table 2 for all the 'training samples'.

The results in table 2 show that clouds, trees, and logged areas required two optimal windows. For these classes, even-numbered windows were approximated to be best, i.e. by taking the average of the minimum and maximum as illustrated in the approximation (approx.) column. Thus, for the afromontane trees, for example, the 12×12 window size was approximated as the optimal for texture mapping. Clouds and bare soil had the largest optimal window sizes compared with afromontane trees

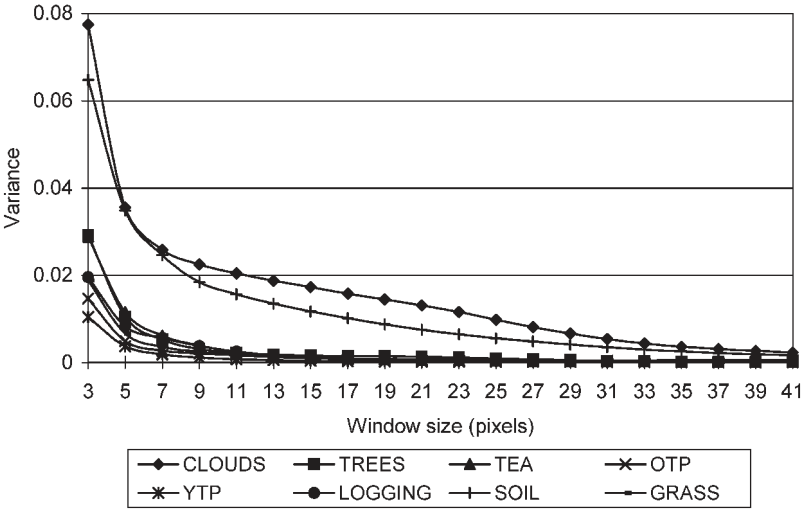

Figure 6. ASM-variance patterns for the training samples over the 20-test windows.

Table 2. Descriptive statistics for the training samples based on the standard semivariance fitting (OTP and YTP refer to old and young tree plantations, respectively).

Training sample	Range	Sill	Nugget variance	Best-fit model	Goodness of fit found (F)	Selected square windows		
						Min	Approx.	Max
Clouds	22.006	1.0000	0.1199	Spherical	2.3836×10^{-3}	21×21	22×22	23×23
Trees	12.123	0.9399	0.1300	Spherical	1.2491×10^{-3}	11×11	12×12	13×13
OTP	14.924	0.9900	0.1499	Spherical	7.2938×10^{-3}		15×15	
YTP	13.074	0.9699	0.1399	Spherical	6.4794×10^{-4}		13×13	
Tea	13.342	1.0000	0.0900	Spherical	9.7408×10^{-4}		13×13	
Soil	19.093	1.0000	0.0050	Spherical	8.9166×10^{-3}		19×19	
Logging	13.612	0.9999	0.0600	Spherical	5.4590×10^{-3}	13×13	14×14	15×15
Grass	12.415	0.9800	0.1000	Spherical	5.0525×10^{-3}		13×13	

with the smallest window size. Old and young trees, tea, soil, and grass could be mapped using approximated window sizes, as shown in table 2. If one were to empirically determine the optimal texture bands, from the variance of the ASM plots directly as shown in figure 6, then the multiple window concept for some 'training samples' would probably be used. The drawback with this empirical selection, however, is that it is time-consuming, and there is a high likelihood of errors due to subjective estimations.

4.1.2 Wavelet transforms texture selection. In this study, 'wavelet transforms texture selection' refers to the determination of the suitable or optimal band(s) for the classification of the land-cover features within the study area. Five levels of the wavelet transformation were generated from the original 2048×2048 image at 1-m spatial resolution. The results were five levels of WT sub-bands corresponding to 2 m, 4 m, 8 m, 16 m, and 32 m, respectively. Higher levels were ignored, since resolutions of more than 32 m were considered not to have any relevance to the capturing of the scene feature(s)-texture details.

Figure 7(*a*) shows the histogram plots of the different WT levels. Level 1 exhibits the highest frequency within the first 50-DN values, followed by levels 5, 2, 3, and 4. Levels 3 and 4 exhibit closely related (coinciding) histogram information. One of the most striking observations here is that level 5, which might theoretically be considered to have the least significance, does have a higher DN frequency than levels 2, 3, and 4. This implies that for this particular scene, some interesting textural information component(s) is resident in level 5. To further confirm this observation, the standardized energy is plotted for the five levels and presented in figure 7(*b*).

The energy plots show that level 1 has the highest amount of energy, followed by level 5 and finally levels 2, 3, and 4 in successive order. Levels 3 and 4 have the same energy, and if this graph is viewed between levels 1 and 4 only, an exponential shape is observed, with a saturation starting somewhere after level 3. It may thus be misleading not to compute level 5 based on the threshold observed in the first four levels. Critically, level 5 should be computed, as it represents the maximum possible resolution or size of individual features within the scene, which are camphor trees.

The significance of the contribution of the WT levels is further investigated by classification of these bands in combination with the spectral information. The classification results from independent WT bands in combination with the spectral data and different levels combinations were also tested. The best results for these combinations are presented. The same approach was applied to the GLCM results.

4.2 *Classification results*

The accuracy of the classification results was tested using the overall accuracy (OA) and kappa index methods derived from the evaluation of the confusion matrix or contingency table. In this matrix or table, classification is given as rows, and verification (ground data or reference data) is given as columns for each sample point. The difference between the overall accuracy measure and the kappa index is that the kappa index measures the relationship between beyond chance agreement and expected disagreement, and uses all elements in the matrix, not just the diagonal elements.

The first results of the classification involved the QB02 spectral bands only. For the eight training classes, the overall accuracy (OA) and kappa index results were 58.69% and 0.4851, respectively. This implies that just about 50% of the entire scene

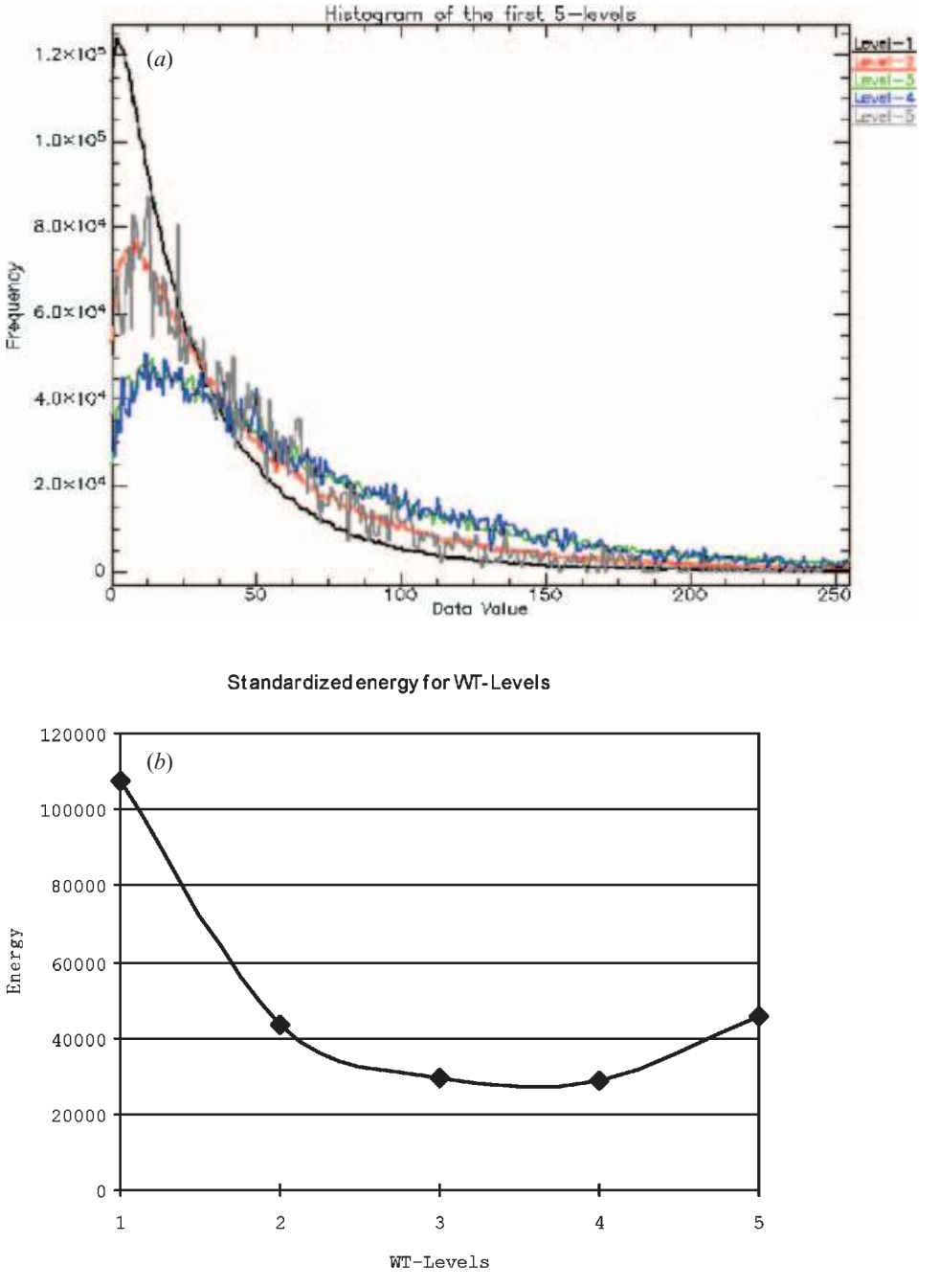

Figure 7. (*a*) Histogram and (*b*) standardized energy plots for the five WT levels. Note that the histogram in (*a*) represents the frequency of occurrences in a continuous format for each level, while the energy plots in (*b*) reflect the distribution of energy along the frequency axis over scale and orientation.

was properly classified into their correct classes. Of course, this is a low accuracy, implying that the thematic information derived from such data may not be very useful in the detection of all the vegetation types. In the next subsections, the results for spectral and texture integrated classifications are presented.

4.2.1 Classification of spectral bands and the GLCM texture. The optimal GLCM textures presented in table 2 were combined with the spectral bands and classified using MLC. The results for the following eight-GLCM texture window sizes (11×11; 12×12; 13×13; 15×15; 19×19; 21×21; 23×23; 25×25) are presented in figure 8. For all tested optimal windows and GLCM texture types, the mean GLCM measure consistently emerged as the best texture type with respect to the overall accuracy reports. The OA results indicated that the mean texture was consistently, by at least 3%, better than the rest of the texture measures for the selected window sizes. The eight texture measures and eight window sizes resulted in a huge database of textures to choose from. The first step to reduce the dimensionality of this database was to select the consistent and most accurate texture type(s). However, as depicted in figure 8, the differences in OA between the window sizes are very marginal from one texture measure to the other. Thus, a window size with the best overall Producer Accuracy (PA) and User Accuracy (UA) for the classes was used to make the final selection decision.

The outliers observed in figure 8 are from the 12×12 window and correspond to homogeneity, entropy, and angular second moment texture types. These GLCM measures recorded very low OAs for this window, implying that the averaging concept adopted in table 1 did not apply for some of the texture types. The results for the original windows used in the averaging concurred well in the other remaining textures types/windows.

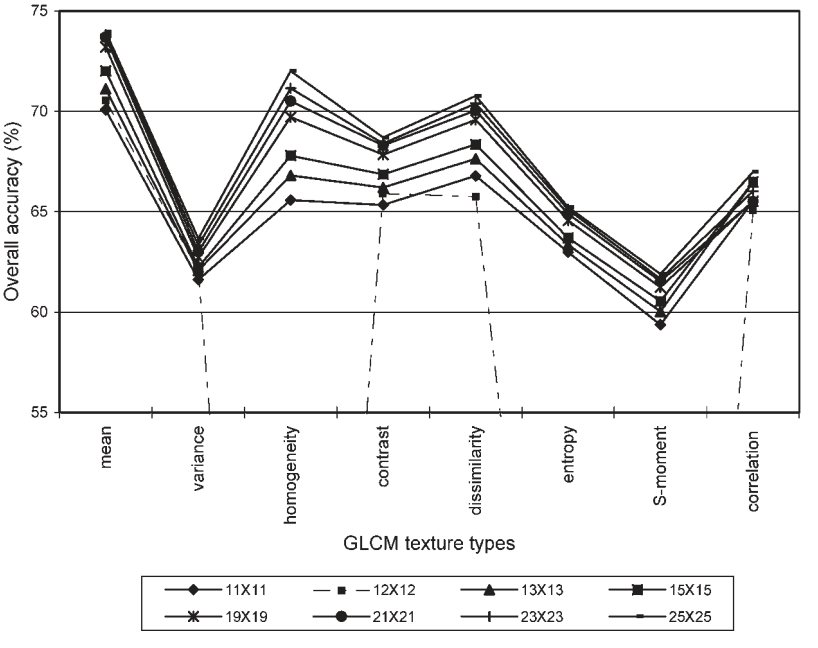

Figure 8. Overall classification accuracy results for GLCM-texture measures versus window size. Note that the significantly very low values for the 12×12 window, are not shown in the graph but are stated and explained accordingly in the text.

Ranking all the textures with respect to their OA and kappa indices, the mean texture measure gave the overall best results followed by homogeneity (70.51%), dissimilarity (70.05%), contrast (68.30%), correlation (65.51%), entropy (64.89%), variance (63.00%), and lastly second angular moment (61.53%). Homogeneity and dissimilarity had remarkably close results for this window size, as did correlation and entropy. Even with the optimal band, the most difficult class to classify was the older planted pine trees. Thus, in the selection of the optimal windows, the PA and UA for this particular class was closely considered such that these two measures (PA and UA) were not only the highest possible but also balanced out.

From figure 9, the kappa index results for spectral bands, and the spectral bands combined with GLCM-mean texture classification, are presented for all the optimal windows. It is observed that from the 21×21 window onwards, the kappa index is as good as being constant (i.e. a maximum difference of 0.0033). Because of the observed constancy, we rely on the PAs and UAs to select the best window size. These results show that the 21×21 window is probably the most optimal single window choice for the scene.

Following the above observations, the window choice for the mean GLCM was the 21×21 window size with an OA of 73.70% and kappa index of 0.6618. This optimal band was named GLCM21-Mean, meaning that the optimal texture window size corresponds to 21×21, and Mean is the optimal GLCM texture measure. In overall, the differences between the best texture (mean) and the least informative texture (second angular moment) for the 21×21 window size was 12.27%. Despite the fact that the 25×25 window had a marginally higher kappa (by 0.0032) than the 21×21 window, the former had a poorer PA and UA than the latter. For example, for the old pine trees, the PA was only 2.38% from the 25×25 window size as compared with 42% for the 21×21 window size.

Comparing the spectral-only and spectral-GLCM texture integrated classification results in figure 9, a significant increase is observed between the spectral-only and upon combination with 11×11-mean texture window of about 11.38%. This is followed by a progressively gradual increase in the OA or kappa, as the window size increases. Between the 11×11 window and the optimal 21×21 window, an increase

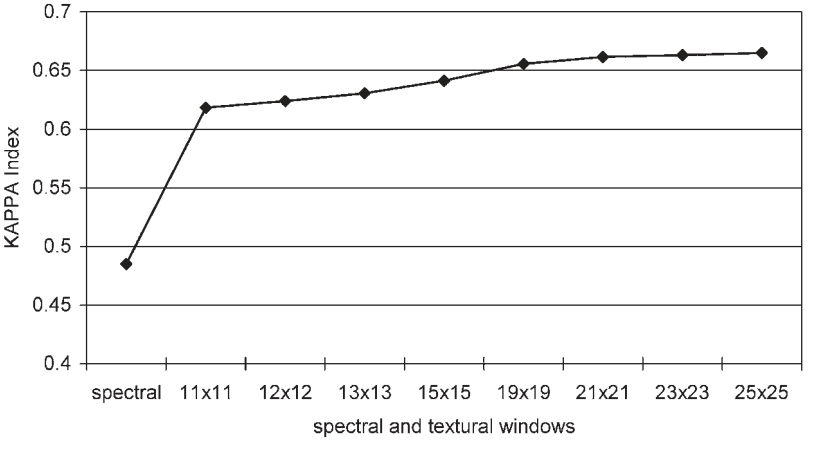

Figure 9. Spectral bands and GLCM-mean texture measure versus window size classification accuracy results.

in classification accuracy of 3.63% is observed. An overall accuracy increase of 15.01% is recorded between the spectral and the optimal GLCM21-Mean texture.

4.2.2 Classification of spectral bands and the WT-texture. The results for the classification of the spectral bands alone and their integration with the WT derived textures are presented in figure 10. From figure 10, it can be seen that the results for levels 1, 2, and 3 are marginally lower than those for spectral-only classification. This shows that the information contained in these levels may not be suitable for capturing the macro-texture information that is needed for the vegetation-type discrimination. However, as the levels increase to 4 and 5, the classification accuracy improves from 58.89% (spectral-only) to 59.32% and 61.42%, respectively. This implies that the significance of macro-texture is slightly noticeable from level 4 by a mere 0.5% and at level 5 by about 2.5%.

When different WT levels are combined into the integrated classification process: (L1 + L2), (L1 + L2 + L3), (L1 + L2 + L3 + L4), and (L1 + L2 + L3 + L4 + L5); where *Li* refers to level *i*, the results as shown in figure 10 are obtained. An increase in OA is noticed only when the first four and five levels are combined; otherwise, for the first two and three levels, the accuracy of the classification is reduced. From these observations, it can justifiably be concluded that levels 1, 2, and 3 are not significant in the macro-texture discrimination of the features in this particular scene. Even on combination with the levels 4 and 5, the lower levels did not produce results that were better than the levels 4 and 5 when considered independently. As independent WT bands, level 5 gave the best results followed by level 4. This means that at a resolution of 32 m, the macro-texture of the scene features is better captured than at the lower resolutions.

The final WT texture combination considered the integration of levels 4 and 5 (16 m and 32 m), since they independently improved the classification results. This combination is referred to in this study as WT (L4 + L5). The OA for all the levels combined and levels 4 and 5 combined, and integrated with the spectral information, showed no significant difference at all, as illustrated in figure 10. In conclusion, to reduce the computational costs and time, and based on the fact that the accuracy improvement contributions of the first three levels were insignificant,

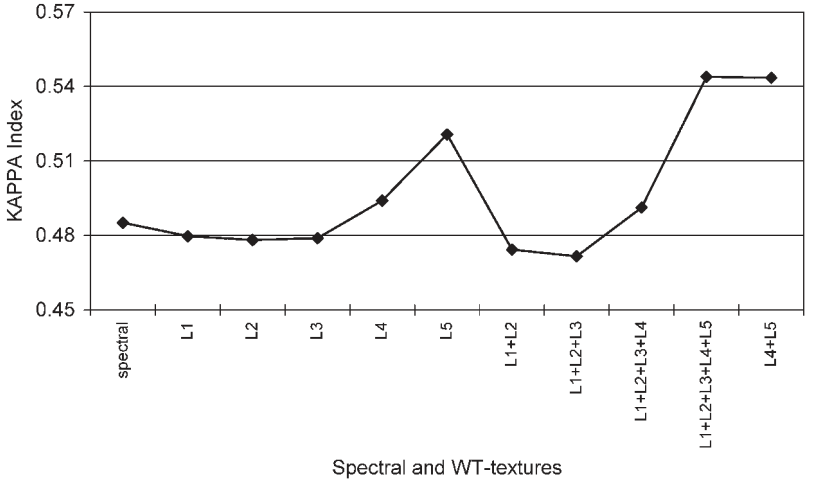

Figure 10. Spectral bands and WT texture classification accuracy results.

the last two levels (L4 + L5) were chosen as the most optimal WT textures, i.e. optimal macro-textures, for the scene classification. With this combination, the OA was increased by 4.86% only, which is approximately three times less than with the GLCM21-Mean improvement of 15.01%, over the spectral-only classification (58.69%).

4.2.3 Spectral bands, GLCM, and WT texture classification. The bands with the highest OA (or kappa index) as determined in the previous subsections were combined in the next classification step to determine the influence of the proposed micro- and macro-texture combinations in vegetation-type classification.

A summary of the OA and kappa indices for the following band combinations is presented in table 3: spectral-only (Spec); Spec combined with GLCM21-Mean texture; Spec combined with WT-levels 4 and 5; Spec combined with GLCM21-Mean and WT-level 4; Spec combined with GLCM21-Mean and WT-level 5; and finally Spec combined with GLCM21-Mean and WT-levels 4 and 5.

For the GLCM texture combination with the spectral data, the results show an increase of 15.01% in the OA with a significant kappa increase by 0.1766. Translating this into the respective class PA and UA results presented in figures 11 and 12, it is observed that both the PA and UA increased significantly, i.e. by more than 10% for most of the classes except for the clouds, which increased only marginally (0.53% of PA and 3.03% for UA). As expected, clouds do not exhibit any spectral or textural confusion with the rest of the 'training samples' and are the most homogenous surface. Hence, equal and high PA and UA scores are expected, as observed in figures 11 and 12. The most remarkable PA/UA increment was that of the old pine tree plantations (OTP). With the spectral bands alone, the PA and UA for the OTP were as low as 30.15% and 28.08%, respectively. However, on integrating the spectral with the GLCM21-Mean, the PA and UA measures increased to 42.36% and 42.53%, respectively.

When WT-levels 4 and 5 are combined with the spectral data, the overall classification accuracy increased only marginally by 4.86%. This is approximately three times less than with the impact of the GLCM21-Mean texture. In fact, the PA for some classes, i.e. soil, decreased by about 10%, as compared with spectral-only, while for other classes like young planted pine trees, the PA and UA remained unchanged (figure 13). Only in the case of grass is the UA more than doubled. Overall, the impact of the WT is not as significant as that of the GCLM21-Mean texture for this particular scene. The probable explanation for this observation is that the structural setting or orientations of features within this scene, which are

Table 3. Comparison of the classification accuracy results for the different band combinations.

	Accuracy measure	
Band	Overall accuracy (%)	Kappa index
Spectral-only (four bands) (Spec)	58.69	0.4851
Spec + GLCM21-Mean	73.70	0.6617
Spec + WT (L4 + L5)	63.55	0.5435
Spec + GLCM21-Mean + WT (L4)	74.96	0.6768
Spec + GLCM21-Mean + WT (L5)	76.26	0.6933
Spec + GLCM21-Mean + WT (L4 + L5) ('proposed method')	77.93	0.7131

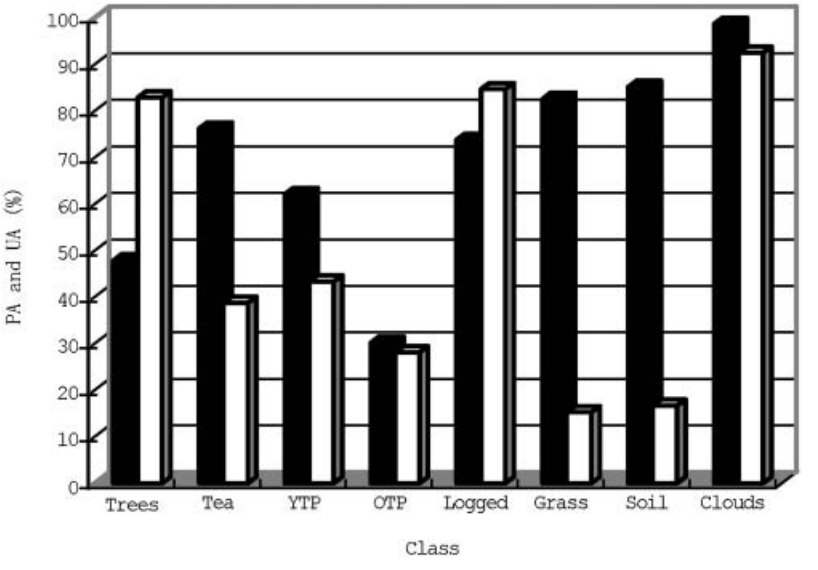

Figure 11. PA and UA result for spectral bands classification.

predominantly vegetation, are weak or not easily captured, i.e. not as sensitive to macro-texture as they are to micro-texture.

The next set of data compared is the combination of WT-L4 with GLCM21-Mean and spectral data (results shown in figure 14). Compared with the spectral bands only, the OA is increased by 16.27% (table 3), which translates into a kappa increment of about 0.1917. This implies that the combination of the GLCM21-Mean with the level 4 WT had more impact in the classification accuracy than using

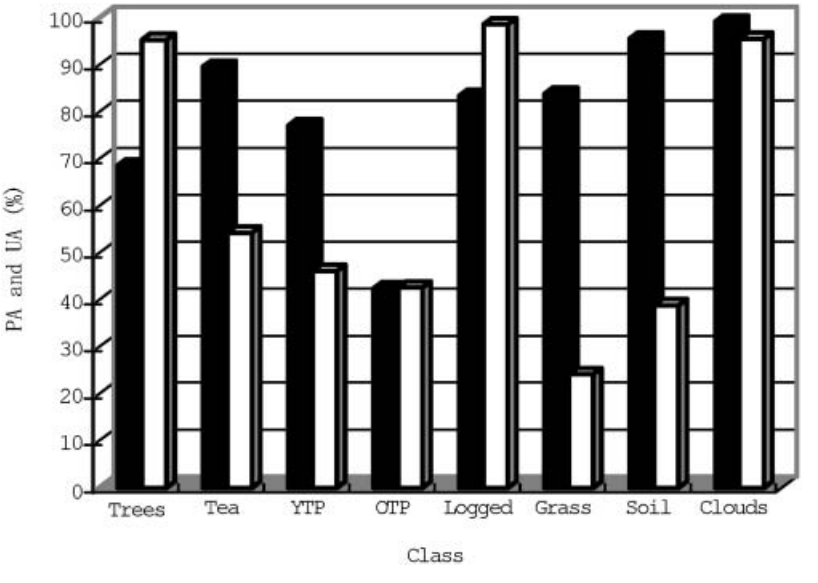

Figure 12. PA and UA results for combined spectral and GLCM21-Mean band classification.

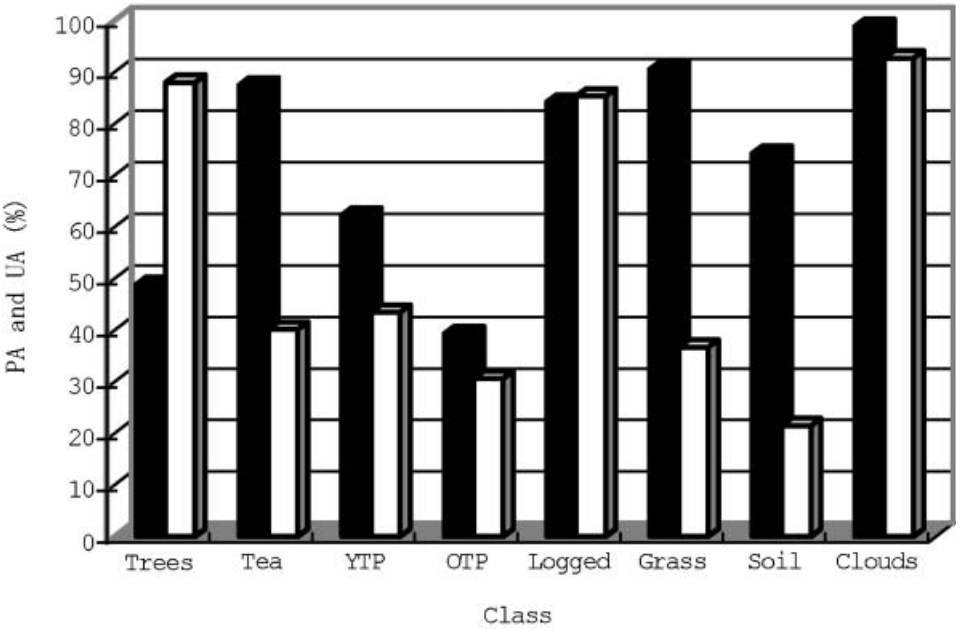

Figure 13. PA and UA results for combined spectral and WT-levels 4 and 5 band classification.

the GLCM21-Mean or even L4 textures combined with spectral bands. The PA and UA for the scene classes also increased by differing amounts. Some of the results were closer to the spectral and GLCM21-Mean combination, i.e. for the young and old pine trees, logged areas, and grass, while better results were obtained for the rest of the classes.

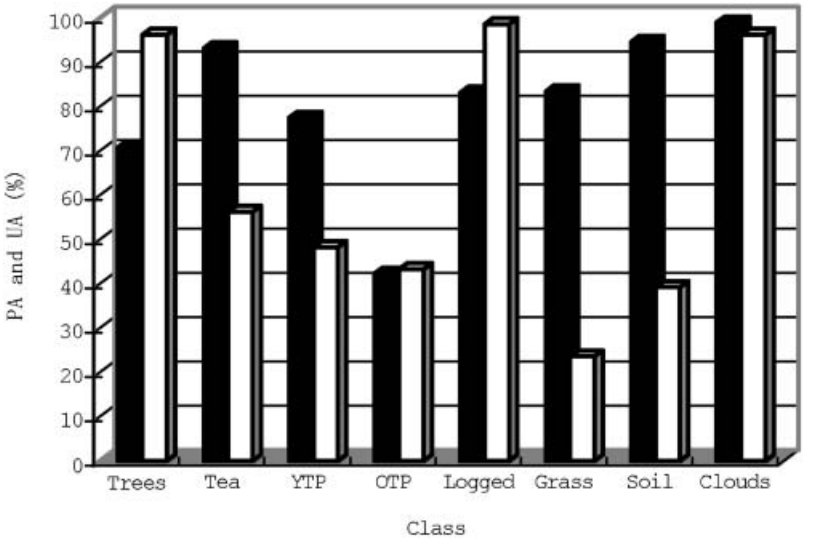

Figure 14. PA and UA results for combined spectral, GLCM21-Mean and WT-level 4 band classification.

When the spectral WT-level 4 is replaced with level 5, the OA increased by 17.57% (table 3). This implies that the level-5 texture is more related or suitable for the scene macro-texture than with level 4. This is approximately 2.56% better than GLCM21-Mean texture alone with spectral data. With respect to the classes and compared with the band combinations already considered above, this particular combination gave the highest PAs and UAs (figure 15). For example, for tea, logged areas, grass, soil, and clouds, the PAs were higher than 91%. Even for old pine tree plantations, the PA was 49.50%.

In figure 16, the overall results for all the optimal textures and spectral data are combined. This is represented as the 'proposed method' results in table 3. The results are termed 'proposed method' results, since part of the objective of this study was to determine the suitable micro–macro texture combination for vegetation species mapping. These results are also termed accordingly, since the rest of the results are only comparative intermediary results, and the proposed method yielded the highest classification accuracy (OA=77.93% and kappa coefficient=0.7131). This is approximately a 20% increment in the scene classification accuracy as compared with spectral-band-only classification results. It is worth noting that even a combination of all the WT bands with the GLCM21-Mean still gave a lower OA (77.46%) and in some cases lower PAs and UAs in comparison with the results of the 'proposed method' bands.

Compared with the previous results, all the PAs and UAs were slightly increased for all the classes except for clouds where there was no change (figure 16). All the classes are mapped with a PA of above 93.5%, except for camphor trees with a PA of 70.66%, younger pine trees with a PA of 72.35%, and older pine trees with a PA of 49.35%. The UAs were significantly high (above 95%) for camphor trees, logged areas, and clouds.

The PAs and UAs are summarised in figure 17(*a*) and 17(*b*), respectively. Figure 17 also presents a clearer format for deciding which band combination is most reliable for mapping a particular class type. Clouds, used here more or less as a benchmark to observe the characteristics of truly homogeneous surface, show that

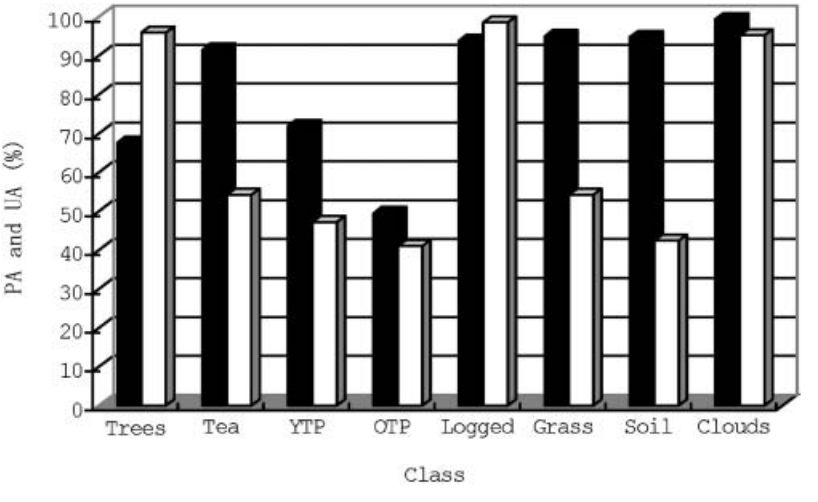

Figure 15. PA and UA results for combined spectral, GLCM21-Mean and WT-level 5 band classification.

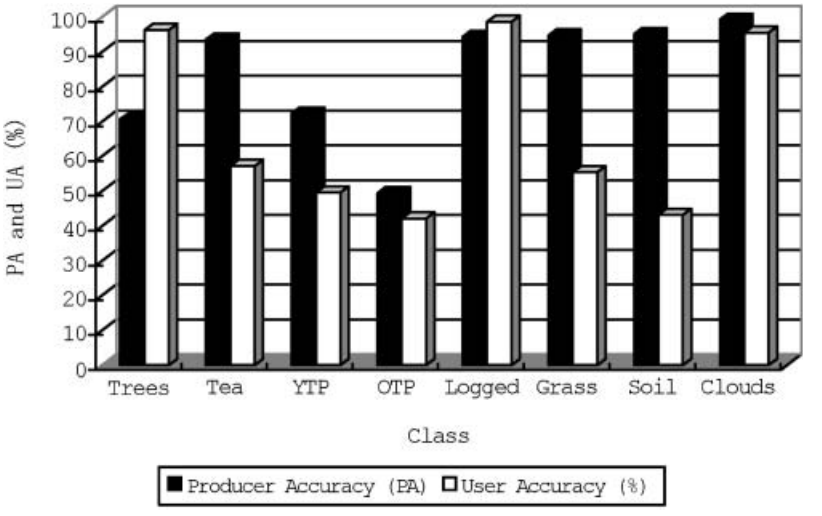

Figure 16. PA and UA results for combined spectral, GLCM21-Mean, and WT-levels 4 and 5 band classification.

PA and UA should have similar values in any band combination considered. Comparing the PA and UA plots, apart from the clouds, the last band combination (spectral, GLCM texture, and WT) gave the best results not only with respect to the overall classification accuracy but also for the PA and UA for most of the classes. The UA results in figure 17(*b*) show a difference between two groups of classes: those above 80% like camphor trees, logged areas, and clouds, and the rest of the classes lying between 15 and 60%. This is quite different from the PA results (figure 17(*a*)), which are more or less widely spread for all the classes. It is not possible to explain why this difference is observed.

In figure 18, some of the results from different band combinations are presented for visual or empirical comparison. Figure 18(*a*) shows the spectral-only classification results, and figures 18(*b*) and 18(*c*) are the results of the spectral combined with GLCM21-Mean, and GLCM21-Mean together with the WT-levels 4 and 5 textures, respectively. The results of spectral-only classification presented in figure 18(*a*) retained the 'salt-and-pepper' like pixel appearance, even after the 3×3 or 5×5 majority filter smoothing kernel was applied to try and smooth the classification results for visualization purposes. From all the results, soils and clouds were well mapped in all the classification results. The problem is with the rest of the vegetated classes, as they all possess green matter, and thus their differentiation is critical. In the spectral-only classification results (figure 18(*a*)), it is empirically observed that apart from the excessive 'salt-and-pepper' appearance, older pine trees were confused or replaced with grass in most cases, and the logged areas are not homogenously mapped as directly visible in the original imagery.

When the GLCM is introduced in the classification process figure 18(*b*), the previously dominant 'salt-and-pepper' appearance is greatly reduced, especially in the highly homogeneous land cover areas like camphor trees, tea, young pine trees, and logged areas. This is an interesting phenomenon and reflects the contribution of the GLCM mean texture measure with corresponding window size of 21×21. Combining the GLCM results with the spectral and WT-levels 4 and 5 (figure 18(*c*))

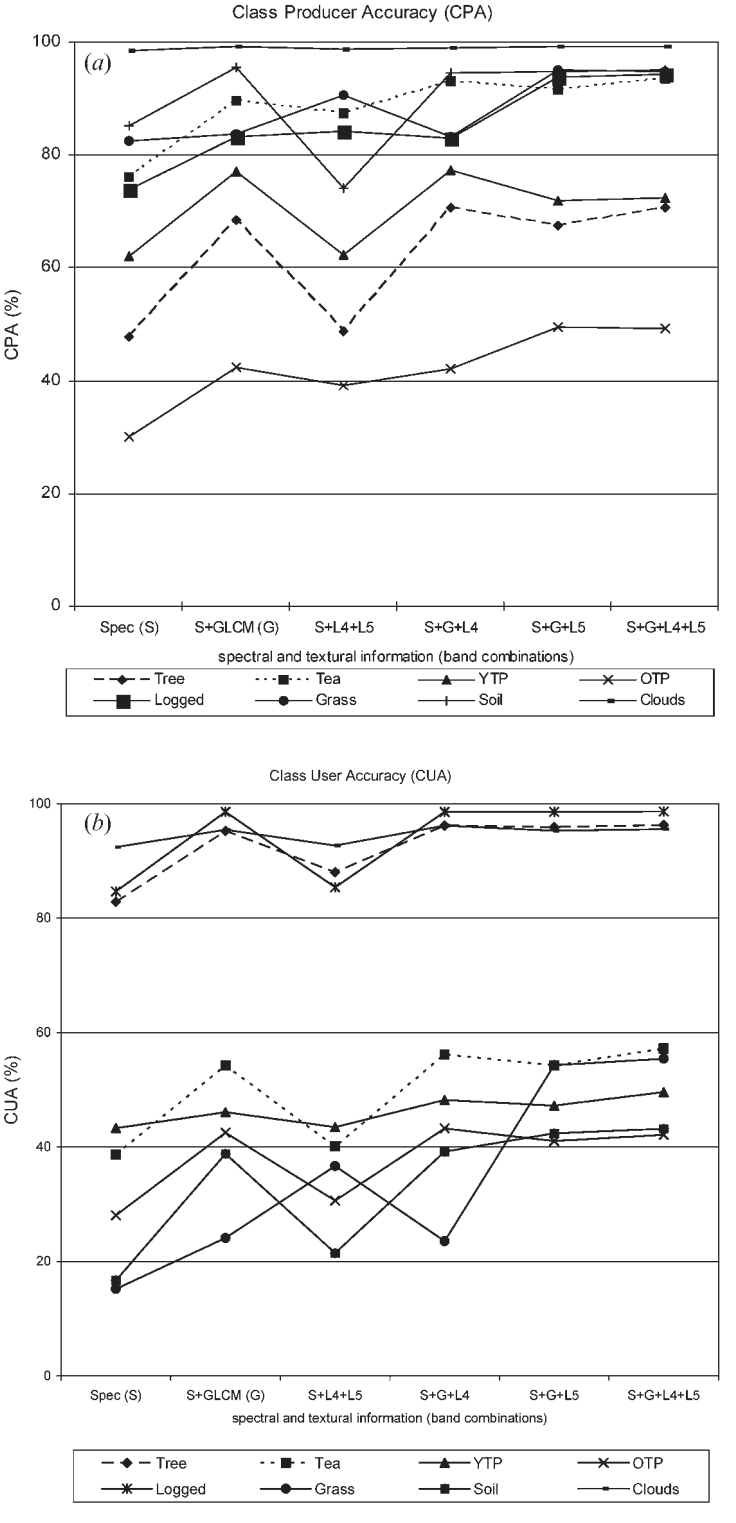

Figure 17. (*a*) Comparison of PA of the eight classes from the selected band combinations. (*b*) Comparison of UAs of the eight classes for the selected band combinations.

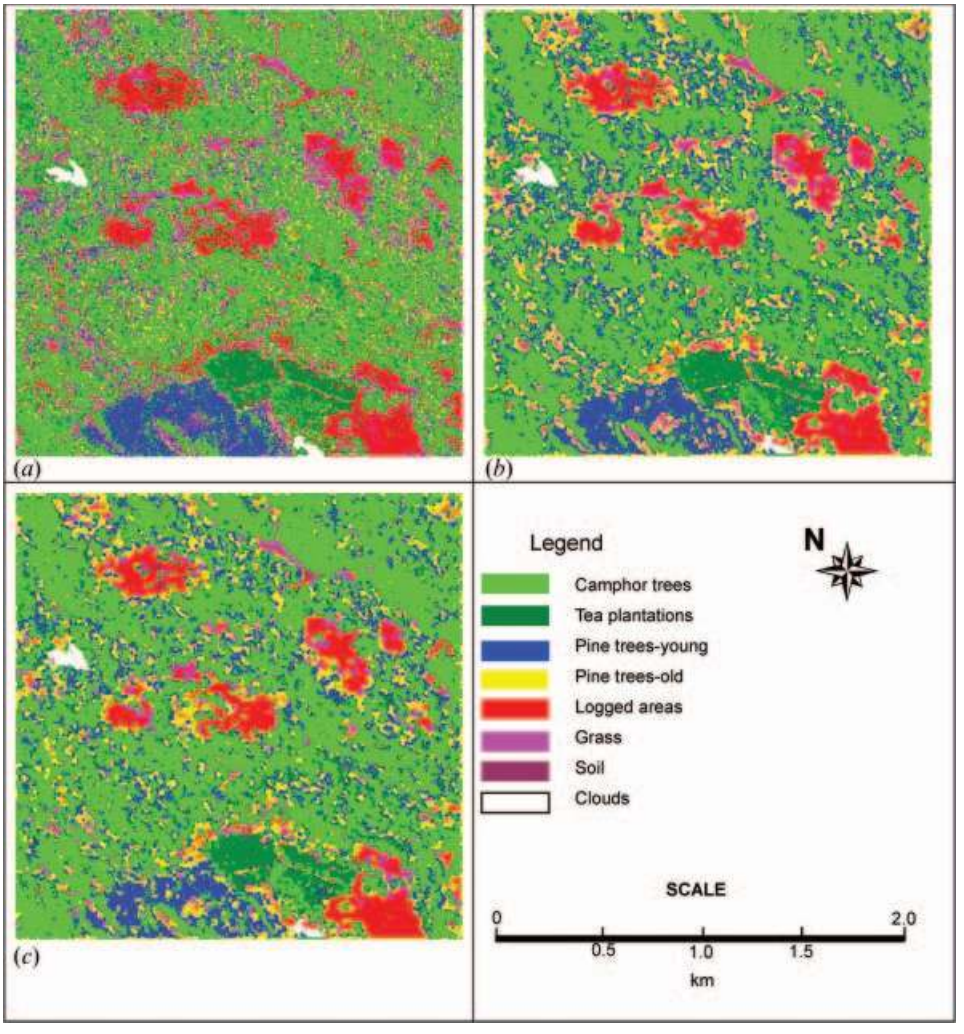

Figure 18. Classification results for: (*a*) spectral bands only; (*b*) spectral and GLCM21-Mean texture bands; and (*c*) spectral, GLCM21-Mean texture, and WT-levels 4 and 5 bands.

appears to further improve the discrimination for some classes like camphor trees, logged areas, and grass. However, the introduction of the WT texture appears to reduce the discrimination of some parts of tea and planted pine trees slightly, especially around the edges. This observation can however be argued to be discrete or occurs in isolated incidences and is overridden not only by the increased PAs and UAs of these classes but also by the improved overall accuracy of the classification results.

5. Discussions and conclusions

What human interpreters cannot possibly or adequately do with respect to vegetation types differentiation, apart from being slower than digital processors, is to successfully isolate sparsely located vegetation types within a scene. This is

where automated processing methods are required, and one of the main cues apart from the spectral information is the texture. Depending on the scene and the task, different texture measures are necessary for capturing the true texture patterns. For example, for texture measure definition in forest applications, the type and density of vegetation may influence the texture measure. While, in urban applications, this may involve the analysis of heterogeneity and symmetry of the urban features, one also ought to appreciate here that we cannot always obtain or acquire other cues to image interpretation like height, etc. It is therefore imperative to maximize the possible available cue sources before investing other data sources, which may also be quite expensive.

For the human visual-perception system, six combined textural features may be most responsible for this judgement: coarseness, contrast, directionality, shape/line-likeness, regularity and roughness. Not all of them can be or have been adequately formularized or reprinted in the digital computing system. Coarseness has a direct relationship to scale and repetition rates, and is a fundamental texture feature. An image will contain textures at several scales; coarseness aims to identify the largest size at which a texture exists, even where a smaller micro-texture exists. Contrast aims to capture the dynamic range of grey levels in an image, together with the polarization of the distribution of black and white. Directionality is a global property over a region, hence macro-texture. The feature directionality texture described does not aim to differentiate between different orientations or patterns but measures the total degree of directionality. This functional reasoning works well in the complicated human visual-based pattern recognition and may change or adapt itself with changes in the scene the characteristics.

There is a wide range of texture-analysis techniques that are used with different criteria for feature extraction: statistical methods (grey-level co-occurrence matrix, semivariogram analysis); filter techniques (wavelet decomposition filters, Gabor filters), etc. The combination of parameters that optimize a method for a specific texture application should be decided upon when these techniques are used. The combination of parameters and the texture method used is expected to be crucial in the success and efficiency of these techniques for a particular application. In this study, we analysed a combinational approach towards micro- and macro-texture application in the classification of vegetation cover from high-resolution remote-sensing images for different types of vegetation.

As reported in Coburn and Roberts (2004), the standard approach to texture analysis is to process a single band of spectral information with a fixed processing window. The results of such an analysis may be profoundly misleading, based on the fact that landscapes and landscape features inherently exist in different scales. This suggests that a single window size or texture may not always yield the desired results. For this reason, in this study, the combinational approach to texture analysis for vegetation species or types classification has been investigated. The results of this study indicate that a combination of micro-texture and macro-texture yields better results. The degree of improvement, however, depends on the texture window size and texture measure for the GLCM textures, and on the level and sum of the wavelet textures.

Nevertheless, the fundamental question that remains elusive is whether the presented approach was able to capture the six fundamental human visual-perception textural features. The answer to this question forms a basis for one of the crucial future areas of further research.

5.1 *On the optimal window size determination*

The results indicate that without some kind of optimization like the semivariance window size modelling, it might not be possible to directly select an optimal window for the case of GLCM textures (micro-textures). However, for macro-textures, the window size may be referred to as the spatial resolution(s) at which the classification rate(s) are maximum, with respect to overall accuracies and the class producers and user accuracies. This observation is supported by the fact that a single window may not always represent the real-world landscape features or components. For example, the same tree species might exist at different growth stages. This may result in different leaf textures based on the age-biophysical characteristics.

The results from this study indicate that the most homogeneous and smooth surfaces (clouds and soil) possessed the highest optimal GLCM texture window sizes, compared with the rest of the classes that were relatively heterogeneous, and do possess rougher canopy structures. This observation implies that the smoothness or roughness of the 'canopy' indeed influences its texture measure window size. However, for practical application purposes, it may be most desirable to determine the window size that yields the maximum information with respect to classification accuracy and with correct edge/shape representation of the original features. The possibility of determining the effectiveness of a common window for the scene features is illustrated in this study. We do not rule out the fact that for any other scene(s), it is possible to combine different texture windows and obtain better results.

5.2 *Optimal texture measure*

The choice of the appropriate texture measure for a given task or scene is essential for any feature(s) discrimination-based application. It is illustrated in this study that with a large GLCM texture database (eight texture measures, each with up to 20 windows), the first step is to reduce this dimensionality by determining the optimal window sizes through semivariogram fitting followed by an assessment of the classification accuracies for the test texture measures over statistically derived optimal windows. The results of this phase yielded a single texture measure, mean, for further analysis. This approach is relevant to GLCM, but for the WT textures, the relevant levels are combined with the spectral bands and the accuracies compared. By doing so, a set of bands will emerge that are suitable for the classification of the scene features. The fact that the mean gave the best GLCM texture measure results shows that this texture measure captured the fundamental vegetation micro-texture properties like coarseness, contrast, and roughness. For the WT textures, the results show that the level-energy measure alone may be misleading if used in deciding on the suitable level(s) for incorporation into the classification process. This is because the features possess textural information that is relevant to their respective nominal scales or patch sizes. Rather, the classification error rates ought to be investigated for independent and combined level selection.

5.3 *Influence of scale in vegetation classification*

While, for the GLCM textures, the influence of scale may not be directly inferred, the influence of scale is quite important from the wavelet texture analysis. With respect to GLCM scale representation, we argue that the optimal windows may be

used as pointers towards the texture scales at which the respective features can be mapped. This should however not be interpreted literally to imply that the spatial resolution for mapping camphor for example trees is 12×12, but rather to depict the scale or window size for mapping the co-occurrence texture patterns for the camphor trees.

The wavelet texture, on the other hand, gives a direct interpretation of the scale range within which the scene features can be best mapped. From the results, it is observed that the optimal WT textures scale lies somewhere between 16 m and 32 m, which corresponds to levels 4 and 5, respectively. First, comparing this scale range and that of the optimal GLCM texture, we have some clue. That is, the average scale for the wavelet texture is roughly around 24 m, while the optimal GLCM texture window size is 21×21 (21 m). Comparing the difference between these windows (3 m) and assessing the scene feature spatial characteristics, we can say with confidence that the scale and window size are very closely related.

Second, apart from the clouds and soils, the sizes of the rest of the features within this scene are quantifiable, either in terms of patch size, e.g. for grass, or in terms of spatial size for camphor trees. An empirical derivation of this scale–space relationship may not be direct unless an approach like that presented in this study is adopted. Third, as supported by the results of this study, an optimal scale may not be resident in a single level, but rather in the combination of two or more levels. Fourth, it is observed that the macro-texture information is not captured at the leaves or branches level but rather over the entire canopy or patch, as indicated by the optimal scale sizes, i.e. 16–32 m. It can be said that the scale–space response depends on the characteristics of the feature of interest. Vegetation exhibits fairly weaker directional structural patterns, hence a low response to wavelet textures in comparison with GLCM textures. Built-regular urban environments with well-structured directional patterns are likely to exhibit very strong wavelet textures. However, their response to GLCM textures may not be that obvious.

5.4 *Classification results*

The main difficulty with respect to GLCM and WT texture integration is that while the former measure is localized and point-based, the latter measure more or less picks the global image characteristics. Arguably, localized texture will combine or integrate well with the spectral information irrespective of the window size, but the global texture or macro-texture appears to slightly over-aggregate some of the classes past some scale or level. For example, at level 5, the tea class appears to attain the stationarity or crisp boundary delineation problem, whereby part of the edges are misclassified, yet the classification accuracies are overestimated, and a distortion in the overall size is observed. This effect is described by Csillag and Kabos (1996) as the dilation of the feature boundaries to encompass the ground data and was also noted later by Coburn and Roberts (2004).

Notably, the findings in this study are subject to the test site as well as the available data, and may thus vary depending on the characteristics of the scene features under investigation. The following observations were made with respect to individual class differentiations in terms of their response to micro- and macro-textures:

1. The camphor trees responded very well to GLCM (micro-) texture with WT (macro-texture). A better accuracy was observed when the two texture types were combined, especially at WT-level 4.

2. Tea exhibited good results with both micro- and macro-textures independently. Level 4 and combination of levels 4 and 5 gave the best results overall.
3. Young pine trees were best discriminated using micro-texture. The impact of macro-texture was insignificant, even on combination with micro-texture and spectral data.
4. The old pines were the most difficult to discriminate. It was only possible to discriminate the older pine trees from the younger ones when the WT textures were introduced into the classification process. This means that the WT scale aspect responded to the age–size differences between these two tree type classes. Macro-texture was observed to have more significance than micro-texture in this particular class.
5. Grass responded well to the combined WT-levels 4 and 5 (macro-textures) than to micro-textures. However, the best results were obtained from both textures combined.
6. In the logged areas, both the micro- and macro-textures had similar impacts. A combination of the two significantly improved the overall discrimination possibility.
7. For clouds and soil, macro-texture was observed to play no role at all. They both exhibit a close-to-perfectly homogeneous surface. These surfaces can only be represented with respect to some optimal patch size and are not captured by wavelet textures.

From the above class results, it can be concluded that the most homogeneous surfaces (clouds and soils), also possessing the largest optimal window sizes, were the least influenced by the macro-texture. For the rest of the classes, mostly vegetation, macro-texture had some impact, particularly for tea and grass vegetation. It is conclusive that the size and density of vegetation leaf size/structural arrangement may influence the texture (micro- or macro-) measure for its discrimination from other vegetation types. The observed 'salt-and-pepper' in the spectral-only classification results confirms that the spectral distribution among different vegetation types depicted in figure 1 is indeed true. For the deciduous camphor trees, for example, the lower spectral reflectance strength between the interlocked tree-canopies is mapped as different classes from the trees, hence the observed 'salt-and-pepper' in figure 18(*a*). The introduction of texture successively captures this unresolved difference through the analysis of neighbourhood patterns giving a better differentiation and aggregation of the vegetation types as shown in figures 18(*b*) and 18(*c*).

It is generally observed that the lowest PA and UA are from spectral-only and spectral plus WT (levels 4 and 5) combinations. This confirms two facts: (1) that spectral information alone is not suitable for vegetation species isolation in this scene using the QB02 channels, and (2) that for vegetation-type differentiation, macro-texture should not be wholly relied upon, as vegetation does not exhibit strong macro-texture properties (shape, direction, definite size). If a user is interested in a particular class, then the selection of the highest PA and UA for the class is a very significant statistic. For an entire scene mapping, however, some slight compromise might have to be made for some class(s). Notably, each class exhibits a unique response to the types of spectral and/or textural information used in the extraction process. However, as reported in the results of this study, an appropriate combination(s) of the micro- and macro-textures will result in better results than if individual textures are used in combination with spectral information.

Finally, it is suggested that different classification approach(s) be compared against the maximum-likelihood as this may also significantly improve the classification output results as observed by Meyer *et al.* (1996) regarding vegetation-type differentiation.

The results of this research conclusively indicate that to improve on the feature recognition and extraction from satellite imagery through classification, a careful selection and integration strategy of spectral, scale-based macro-texture and co-occurrence (micro-) texture cues is inevitable. Our future work, apart from comparing different classification approaches, would be to derive results for other scenes/features so as to be able to obtain generalized models or theories. This is supported by observations in Wilkinson (2005) pointing to the fact that satellite-image classification accuracy has not improved in the past 15 years, and as a consequence more robust approaches are required.

Acknowledgements

The authors would like to acknowledge the fellowship grant from the Japanese Ministry of Science and Education. We also acknowledge the Kenya Wildlife Services (KWS) for the ancillary data supplied for this research.

References

ANYS, H., BANNARI, A., HE, D.C. and MORIN, D., 1994, Texture analysis for the mapping of urban areas using airborne MEIS-II images. In *Proceedings of the First International Airborne Remote Sensing Conference and Exhibition, ERIM*, 12–15 September, Strasbourg, France, pp. 231–245.

APLIN, P. and ATKINSON, P.M., 2001, Sub-pixel lands cover mapping for per-field classification. *International Journal of Remote Sensing*, **22**, pp. 2853–2858.

APLIN, P., ATKINSON, P.M. and CURRAN, P.J., 1999, Fine spatial resolution simulated satellite sensor imagery for land cover mapping in the United Kingdom. *Remote Sensing of Environment*, **68**, pp. 206–216.

ARZANDEH, S. and WANG, J., 2002, Texture evaluation of RADARSAT imagery for wetland mapping. *Canadian Journal of Remote Sensing*, **28**, pp. 653–666.

ASNER, G.P., KELLER, M., PEREIRA JR, R. and ZWEEDE, J.C., 2002, Remote sensing of selective logging in Amazonia: assessing limitations based on detailed field observations, Landsat ETM+ and textural analysis. *Remote Sensing of Environment*, **80**, pp. 483–496.

ATZBERGER, C., 2004, Object-based retrieval of biophysical canopy variables using artificial neural nets and radiative transfer models. *Remote Sensing of Environment*, **93**, pp. 53–67.

BAKKER, W.H., et al. 2001, *Principles of Remote Sensing*, ITC Educational Textbook Series 2 (Enschede, Netherlands: ITC).

BENEDIKTSSON, J.A., PESARESI, M. and ARNASSON, K., 2003, Classification and feature extraction for remote sensing images from urban areas based on morphological transformations. *IEEE Transactions on Geoscience and Remote Sensing*, **41**, pp. 1940–1949.

BRANDTBERG, T. and WALTER, F., 1998, An algorithm for delineation of individual tree crowns in high spatial resolution aerial images using curved edge segments at multiple scales. In *Automated Interpretation of High Spatial Resolution Digital Imagery for Forestry*, pp. 41–54 (Victoria, Canada: Natural Resources Canada, Canadian Forest Service).

CARLEER, A. and WOLFF, E., 2004, Exploitation of very high resolution satellite data for tree species identification. *Photogrammetric Engineering and Remote Sensing*, **70**, pp. 135–140.

COBURN, C.A. and ROBERTS, A.C.B., 2004, A multiscale texture analysis procedure for improved forest stand classification. *International Journal of Remote Sensing*, **25**, pp. 4287–4308.

CSILLAG, F. and KABOS, S., 1996, Hierarchical decomposition of variance with applications in environmental mapping based on satellite images. *Mathematical Geology*, **28**, pp. 385–405.

CULVENOR, D.S., COOPS, N., PRESTON, R. and TOLHURST, K.G., 1998, A spatial clustering approach to automated tree crown delineation. In *Automated Interpretation of High Spatial Resolution Digital Imagery for Forestry*, pp. 67–80 (Victoria, Canada: Natural Resources Canada, Canadian Forest Service).

CULVENOR, D.S., 2002, TIDA: an algorithm for the delineation of tree crowns in high spatial resolution remotely sensed imagery. *Computers and Geosciences*, **28**, pp. 33–44.

CURRAN, P. and ATKINSON, P., 1998, Geostatistics and remote sensing. *Progress in Physical Geography*, **22**, pp. 61–78.

DAUBECHIES, I., 1988, Orthonormal bases of compactly supported wavelets. *Communications on Pure and Applied Mathematics*, **41**, pp. 909–996.

DEFINIENS, 2000, *eCognition—Object Oriented Image Analysis. User's Manual*. Published by Definiens.

DIXON, R.K., BROWN, S., HOUGHTON, R.A., SOLOMON, A.M., TREXLER, M.C. and WISNIEWSKI, J., 1994, Carbon pools and flux of global forest ecosystems. *Science*, **263**, pp. 185–190.

EHLERS, M., GAHLER, M. and JANOWSKY, R., 2003, Automated analysis of ultra high-resolution remote sensing data for biotope mapping: New possibilities and challenges. *ISPRS Journal of Photogrammetry and Remote Sensing*, **57**, pp. 315–326.

EVERITT, J.H., ESCOBAR, D.E., ALANIZ, M.A. and DAVIS, M.R., 1987, Using airborne middle-infrared (1.45–2.0 μm) video imagery for distinguishing plant species and soil conditions. *Remote Sensing of Environment*, **22**, pp. 423–428.

FRANK, T.D., 1988, Mapping dominant vegetation communities in the Colorado Rocky Mountain Front Range with Landsat Thematic Mapper and digital terrain data. *Photogrammetric Engineering and Remote Sensing*, **54**, pp. 1727–1734.

FRANKLIN, S.E., 1994, Discrimination of subalpine forest species and canopy density using CASI, SPOT, PLA, and Landsat TM data. *Photogrammetric Engineering and Remote Sensing*, **60**, pp. 1233–1241.

FRANKLIN, S.E., HALL, R.J., MOSKAL, L.M., MAUDIE, A.J. and LAVIGNE, M.B., 2000, Incorporating texture into classification of forest species composition from airborne multispectral images. *International Journal of Remote Sensing*, **21**, pp. 61–79.

FRANKLIN, S.E., MAUDIE, A.J. and LAVIGNE, M.B., 2001a, Using spatial co-occurrence texture to increase forest structure and species composition classification accuracy. *Photogrammetric Engineering and Remote Sensing*, **67**, pp. 849–855.

FRANKLIN, S.E., WULDER, M.A. and GERYLO, G.R., 2001b, Texture analysis of IKONOS panchromatic data for Douglas-fir forest age class separability in British Columbia. *International Journal of Remote Sensing*, **22**, pp. 2627–2632.

GILLIS, M.D. and LECKIE, D.G., 1996, Forest inventory update in Canada. *The Forestry Chronicle*, **72**, pp. 138–156.

GOUGEON, F.A., 1996, Télédétection numérique. In *Manuel de foresterie*, J. Bérard (Ed.), pp. 540–569 (Sainte-Foy, Canada: Les presses de l'Université de Laval).

GOUGEON, F.A., 1998, Automatic individual tree crown delineation using a valley-following algorithm and a rule-base system. In *Automated Interpretation of High Spatial Resolution Digital Imagery for Forestry*, pp. 115–125 (Victoria, Canada: Natural Resources Canada, Canadian Forest Service).

HARALICK, R.M., SHANMUGAN, K. and DINSTEIN, I., 1973, Texture features for image classification. *IEEE Transactions Systems, Man and Cybernetics*, **3**, pp. 610–621.

HODGSON, M.E., JENSEN, J.R., MACKEY, J.E., J.R. and COULTER, M.C., 1988, Monitoring wood stork foraging habitat using remote sensing and geographic information systems. *Photogrammetric Engineering and Remote Sensing*, **54**, pp. 1601–1607.

HYYPPA, H.J. and HYYPPA, J.M., 2001, Effects of stand size on the accuracy of remote sensing-based forest inventory. *IEEE Transactions on Geoscience and Remote Sensing*, **39**, pp. 2613–2621.

HYYPPA, J.M., HYYPPA, H.J., INKINEN, M., ENGDAHL, M., LINKO, S. and ZHU, Y.H., 2000, Accuracy comparison of various remote sensing data sources in the retrieval of forest stand attributes. *Forest Ecology and Management*, **128**, pp. 109–120.

JAKOBSONS, A., 1970, *Sambandet mellan trädkronans diameter och andra trädfaktorer, främst brösthöjdsdiametern. Rapport 14, Inst. f. skogstaxering* (Stockholm: Skogshögskolan) (in Swedish with English summary).

JANSSEN, L.L.F. and MOLENAAR, M., 1995, Terrain objects, their dynamics and their monitoring by the integration of GIS and remote sensing. *IEEE Transactions on Geoscience and Remote Sensing*, **33**, pp. 749–758.

KEELETSANG, M., 2004, Assessment of dry season transpiration using IKONOS images: Serowe case study, Botswana. MSc thesis, ITC, Enschede.

KILPELAINEN, P. and TOKOLA, T., 1999, Gain to be achieved from stand delineation in LANDSAT TM image-based estimates of stand volume. *Forest Ecology and Management*, **124**, pp. 105–111.

LATHROP, R.G., ABER, J.D., BOGNAR, J.A., OLLINGER, S.V., CASSET, S. and ELLIS, J.M., 1994, GIS development to support regional simulation modeling of northeastern (USA) forest ecosystems. In W. Michener, J.W. Brunt and S. Stafford (Eds). *Environmental Information Management and Analysis*, pp. 431–451 (London: Taylor & Francis).

LECKIE, D.G. and GOUGEON, F.A., 1999, An assessment of both visual and automated tree counting and species identification with high spatial resolution multispectral imagery. *Automated Interpretation of High Spatial Resolution Digital Imagery for Forestry International Forum*, pp. 141–152.

LEWIS, M.M., 1994, Species composition related to spectral classification in an Australian spinifex hummock grassland. *International Journal of Remote Sensing*, **15**, pp. 3223–3239.

LOBO, A., 1997, Image segmentation and discriminant analysis for the identification of land cover units in ecology. *IEEE Transactions on Geoscience and Remote Sensing*, **35**, pp. 1136–1145.

LOBO, A., CHIC, O. and CASTERAD, A., 1996, Classification of Mediterranean crops with multisensor data: Per-pixel versus per-object statistics and image segmentation. *International Journal of Remote Sensing*, **17**, pp. 2385–2400.

LUCAS, R.M., ROWLANDS, A., NIEMANN, O. and MERTON, R., 2004, Hyperspectral Sensors and Applications. In *Advance Image Processing Techniques for Remotely Sensed Hyperspectral Data* (New York: Springer).

MALLAT, S.G., 1989, A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **11**, pp. 674–693.

MANDAL, M.K. and ABOULNASR, T., 1996, Image indexing using moments and wavelets. *IEEE Transactions on Consumer Electronics*, **42**, pp. 57–565.

MEYER, P., STAENZ, K. and ITTEN, K.I., 1996, Semi-automated procedures for tree species identification in high spatial resolution data from digitised colour infrared-aerial photography. *ISPRS Journal of Photogrammetry and Remote Sensing*, **51**, pp. 5–16.

MINOR, C.O., 1951, Stem-crown diameter relations in southern pine. *Journal of Forestry*, **49**, pp. 490–493.

MUINONEN, E., MALTAMO, M., HYPPANEN, H. and VAINIKAINEN, V., 2001, Forest stand characteristics estimation using a most similar neighbor approach and image spatial structure information. *Remote Sensing of Environment*, **78**, pp. 223–228.

NELSON, R.F., LATTY, R.S. and MOTT, G., 1985, Classifying northern forests using Thematic Mapper Simulator data. *Photogrammetric Engineering and Remote Sensing*, **50**, pp. 607–617.

OUMA, Y.O., NGIGI, T.G. and TATEISHI, R., 2006, On the optimisation and selection of wavelet texture for feature extraction from high-resolution satellite imagery with application towards urban-trees delineation. *International Journal of Remote Sensing*, **27**, pp. 73–104.

PANNATIER, Y., 1996, *Variowin: Software for Spatial Data Analysis in 2D (Software)* (New York: Springer).

PCI 2001, *Using PCI software*, V. 8.0 (Richmond Hill, Canada: PCI).

PINZ, A., 1989, Final results of the vision expert system VES: Finding trees in aerial photographs. In A. Pinz (Ed.). *Wissensbasierte Mustererkennung (Knowledge-Based Pattern Recognition)*, pp. 90–111 (Vienna: Oldenbourg).

POLLOCK, R.J., 1994, A model-based approach to automatically locating tree crowns in high spatial resolution images. *Image and Signal Processing for Remote Sensing*, **2315**, pp. 526–537.

PUISSANT, A., HIRSCH, J. and WEBER, C., 2005, The utility of texture analysis to improve per-pixel classifications for high to very high spatial resolution imagery. *International Journal of Remote Sensing*, **26**, pp. 733–745.

REAKA-KUDLA, M.L.,WILSON, D.E., and WILSON, E.O. (Eds), 1997, *Biodiversity II: Understanding and Protecting Our Biological Resources* (Washington, DC: Joesph Henry Press).

ROHDE, W.G. and OLSON, C.E. JR., 1972, Multispectral sensing of forest tree species. *Photogrammetric Engineering and Remote Sensing*, **38**, pp. 1209–1215.

SCHNEIDER, W., 1989, Forest damage mapping in Austria with the aid of Landsat TM image data. In *Monitoring the Earth's Environment: A Project Campaign on Landsat/TM Applications*, pp. 223–231 (Noordwijk: ESA SP-1102).

SCHRIEVER, J.R. and CONGALTON, R.G., 1995, Evaluating seasonal variability as an aid to cover-type mapping from Landsat Thematic Mapper data in the northeast. *Photogrammetric Engineering and Remote Sensing*, **61**, pp. 321–327.

SHEN, S.S., BADHWAR, G.D. and CARNES, J.G., 1985, Separability of boreal forest species in the Lake Jennette area, Minnesota. *Photogrammetric Engineering and Remote Sensing*, **51**, pp. 1775–1783.

SKIDMORE, A.K., 1989, An expert system classifies eucalypt forest types using Thematic Mapper data and a digital terrain model. *Photogrammetric Engineering and Remote Sensing*, **55**, pp. 1449–1464.

SMITH, G.M. and FULLER, R.M., 2001, An integrated approach to land cover classification: an example in the Island of Jersey. *International Journal of Remote Sensing*, **22**, pp. 3123–3142.

ST. ONGE, B.A. and CAVAYAS, F., 1997, Automated forest structure mapping from high resolution imagery based on directional semivariogram estimates. *Remote Sensing of Environment*, **61**, pp. 82–95.

SOARES, J.V., RENNO, C.D., FORMAGGIO, A.R., YANASSE, C.C.F. and FRERY, A.C., 1997, An investigation of the selection of texture features for crop discrimination using SAR imagery. *Remote Sensing of Environment*, **59**, pp. 234–247.

THOMASSON, J.A., BENNETT, C.W., JACKSON, B.D. and MAILANDER, M.P., 1994, Differentiating bottomland tree species with multispectral videography. *Photogrammetric Engineering and Remote Sensing*, **60**, pp. 55–59.

TOUTIN, T. and CHENG, P., 2000, Demystification of IKONOS. *Earth Observation Magazine*, **9**, pp. 17–21.

TROTTER, C.M., DYMOND, J.R. and GOULDING, C.J., 1997, Estimation of timber volume in a coniferous plantation forest using Landsat TM. *International Journal of Remote Sensing*, **18**, pp. 2209–2223.

TREITZ, P.M., HOWARTH, P.J., ROTUNNO, F.O. and SOULIS, E.D., 2000, Agricultural crop classification using SAR tone and texture statistics. *Canadian Journal of Remote Sensing*, **26**, pp. 18–29.

ULABY, F.T., KOUYATE, F., BRISCO, B. and WILLIAMS, T.H.L., 1986, Textural information in SAR images. *IEEE Transactions on Geoscience and Remote Sensing*, **GE-24**, pp. 235–245.

VAN DE VOWER G., SCHEUNDERS, P. and VAN DYCK, D., 1999, Statistical texture characterization from discrete wavelet representations. *IEEE Transactions on Image Processing*, **8**, pp. 592–598.

WARNER, T.A., LEE, J.Y. and MCGRAW, J.B., 1998, Delineation and identification of individual trees in the eastern deciduous forest. In *Automated Interpretation of High Spatial Resolution Digital Imagery for Forestry*, pp. 81–91 (Victoria, Canada: Natural Resources Canada, Canadian Forest Service).

WHITE, J.D., KROH, G.C. and PINDER, J.E., 1995, Forest mapping at Lassen Volcanic National Park, California, using Landsat TM data and a geographical information system. *Photogrammetric Engineering and Remote Sensing*, **61**, pp. 299–305.

WILKINSON, G.G., 2005, Results and implications of a study of fifteen years of satellite image classification experiments. *IEEE Transactions on Geoscience and Remote Sensing*, **43**, pp. 433–440.

WOODCOCK, C.E., COLLINS, J.B., GOPAL, S., JAKABHAZY, V.D., LI, X., MACOMBER, S., RYHERD, S., HARWARD, V.J., LEVITAN, J., WU, Y. and WARBINGTON, R., 1994, Mapping forest vegetation using Landsat TM imagery and a canopy reflectance model. *Remote Sensing of the Environment*, **50**, pp. 240–254.

WULDER, M., NIEMANN, K.O. and GOODENOUGH, D.G., 2000, Local maximum filtering for the extraction of tree locations and basal area from high spatial resolution imagery. *Remote Sensing of Environment*, **73**, pp. 103–114.

YU, Q., GONG, P., CLINTON, N., BIGING, G., KELLY, M. and SCHIROKAUER, D., 2006, Object-based detailed vegetation classification with airborne high spatial resolution remote sensing imagery. *Photogrammetric Engineering and Remote Sensing*, **72**, pp. 799–811.

ZHANG, Y., 2001, Texture-integrated classification of urban treed areas in high-resolution color-infrared imagery. *Photogrammetric Engineering and Remote Sensing*, **67**, pp. 1359–1367.