International Journal of Remote Sensing, 2015
Vol. 36, No. 1, 318–342, http://dx.doi.org/10.1080/01431161.2014.995272

Taylor & Francis
Taylor & Francis Group

Airborne hyperspectral discrimination of tree species with different ages using discrete wavelet transform

A. Ghiyamat[a]*, H.Z.M. Shafri[b], G.A. Mahdiraji[c], R. Ashurov[d], A.R.M. Shariff[e], and S. Mansor[b]

[a]*Institute of Gerontology, Universiti Putra Malaysia, Selangor, Malaysia;* [b]*Department of Civil Engineering, Universiti Putra Malaysia, Selangor, Malaysia;* [c]*Department of Electrical Engineering, University of Malaya, Kuala Lumpur, Malaysia;* [d]*Institute of Mathematics, National University of Uzbekistan, Tashkent, Uzbekistan;* [e]*Geospatial Information Science Research Centre (GISRC), Faculty of Engineering, Universiti Putra Malaysia (UPM), 43400 Selangor, Malaysia*

In this article, the capability of discrete wavelet transform (DWT) to discriminate tree species with different ages using airborne hyperspectral remote sensing is investigated. The performance of DWT is compared against commonly used traditional methods, i.e. original reflectance and first and second derivatives. The hyperspectral data are obtained from Thetford forest of the UK, which contains Corsican and Scots pines with different ages and broadleaved tree species. The discrimination is performed by employing three different spectral measurement techniques (SMTs) including Spectral Angle Mapper (SAM), Spectral Information Divergence (SID), and a combination of SAM and SID. Five different mother wavelets with a total of 50 different orders are tested. The wavelet detail coefficient (CD) from each decomposition level and combination of all CDs plus the approximation coefficient from the final decomposition level (C-All) are extracted from each mother wavelet. The results show the superiority of DWT against the reflectance and derivatives for all the three SMTs. In DWT, C-All provided the highest discrimination accuracy compared to other coefficients. An overall accuracy difference of about 20–30% is observed between the finest coefficient and C-All. Amongst the SMTs, SID provided the highest accuracy, while SAM showed the lowest accuracy. Using DWT in combination with SID, an overall accuracy up to around 71.4% is obtained, which is around 13.5%, 14.7%, and 27% higher than the accuracies achieved with reflectance and first and second derivatives, respectively.

1. Introduction

Tree species identification is a key element in monitoring forest species distribution and the changes in biodiversity with time for the goal of designing meaningful conservation strategies. Due to a higher number of spectral bands and spectral resolution, hyperspectral remote-sensing technology can provide significant enhancement of spectral measurement capabilities compared with multispectral remote-sensor systems (Ghiyamat and Shafri 2010) and this can be useful for forest tree species identification.

Reflectance spectra (or original reflectance) are the main information obtained from remote-sensing imagery. So far, researchers have been able to classify tree species based on airborne hyperspectral reflectance (Dibley, Turner, and Skidmore 1997; Martin et al. 1998; Clark, Roberts, and Clark 2005; Cho et al. 2010). However, tree species discrimination using reflectance spectra is challenging due to the influence

of environmental factors such as soil characteristics, precipitation, and soil moisture (Portigal et al. 1997; Lee et al. 2010). These environmental factors can affect chlorophyll content (Zarco-Tejada et al. 2003), water content (Lee et al. 2010), and transmission properties of leaves and wood, which might lead to increased dissimilarity among individuals of the same species.

The derivative of reflectance spectra is another method commonly used in different remote-sensing applications to mitigate the influence of background and illumination effects (Demetriades-Shah, Steven, and Clark 1990; Gong, Pu, and Miller 1992; Xiao-chen et al. 2008). The capabilities of derivative analysis to recognize different vegetation types (Yamano, Chen, and Tamura 2003; Kalluri, Prasad, and Bruce 2010), estimation of leaf area index (LAI) (Xiao-chen et al. 2008; Gong, Pu, and Miller 1992), stress (Smith, Steven, and Colls 2004), leaf chlorosis (Adams, Philpot, and Norvell 1999), chlorophyll (Zarco-Tejada et al. 2003), water content (Zhang, Jianjun, and Zhou 2010), red-edge position (REP) estimation (Shafri et al. 2006), and tree species discrimination (Affendi et al. 2006) using airborne hyperspectral remote sensing have been reported. Also, airborne hyperspectral discrimination of tree species with different ages using derivative spectra over three different classifications has been examined (Ghiyamat et al. 2013). They have shown that derivative spectra outperformed reflectance spectra in simple classification schemes. In more challenging classification schemes, however, no advantage has been observed when using spectral derivative compared to reflectance spectra (Ghiyamat et al. 2013).

Recently, wavelet transform (WT) has been used as an efficient approach for extracting more detailed spectral information at different scales in analysing hyperspectral data. The fine-scale and large-scale (i.e. the first and the last WT levels/coefficients) information of hyperspectral signals can be simultaneously investigated by projecting signals onto a set of wavelet bases with various scales (Li 2004). WT can provide analysis of signals either across continuum scales, referred to as continuous WT (CWT), or over a discrete set of scales, known as discrete WT (DWT).

Several case studies have been reported in the literature that applied WT on hyperspectral data for forest monitoring purposes, including chlorophyll quantification (Blackburn and Ferwerda 2008), denoising (Chen and Qian 2011), and forest LAI estimation (Banskota, Wynne, Serbin, et al. 2013; Banskota, Wynne, Thomas, et al. 2013). The performance of WT in different applications has also been compared against several traditional methods. Misman, Shafri, and Ahmad (2010) used the support vector machine (SVM) algorithm to compare the performance of spectral reflectance against DWT, CWT, and the first derivative in discriminating five land-cover classes, namely water, rooftop, road, concrete, and vegetation, using airborne hyperspectral data. In that study, the highest and the lowest discrimination accuracies were obtained using DWT and first derivative analysis, respectively. Pu and Gong (2004) assessed the performance of DWT against band selection and principal component analysis (PCA) in mapping the forest crown closure and leaf area index using satellite hyperspectral data. Li (2004) has shown that the abundance estimation deviation can be reduced by 30% to 50% on average by using DWT-based features compared to using conventional PCA- and discrete-cosine-transform-based features or the original hyperspectral signals. The results of these studies demonstrate the capability of WT to analyse hyperspectral data for different applications. However, only a few studies have employed WT for species discrimination using airborne hyperspectral data. Zhang et al. (2006) investigated within and between tropical tree species variations measured by the Spectral Angle Mapper (SAM) over airborne hyperspectral data. They showed that by using DWT, the spectral separability between tree species can be increased compared to the reflectance and first derivative spectra.

Banskota, Wynne, and Kayastha (2011) applied DWT with linear discriminant analysis (LDA) on original reflectance to discriminate three pine species (loblolly pine, Virginia pine, and shortleaf pine) utilizing airborne visible/infrared imaging spectrometer (AVIRIS) data. Their results showed the outperformance of DWT by 7.5% higher accuracy compared to the original reflectance in pine tree species discrimination. The capability of DWT to discriminate tree species has been emphasized by both Zhang et al. (2006) and Banskota, Wynne, and Kayastha (2011). However, further study is required to examine the capability of different wavelet families for different tree species.

Discriminating tree species with different ages has been shown to be quite challenging (Ghiyamat et al. 2013) due to the high similarity or small spectral difference between species from the same age category. Since DWT is a frequency sensitive technique, it can enlarge small spectral variations using different mother wavelets and/or in different wavelet coefficients. This study aims to investigate whether DWT can be useful in highlighting the small spectral differences in such tree species.

For discriminating different spectral signatures, a good spectral measurement technique (SMT) is required to determine spectral similarity and dissimilarity. The spectral angle mapper (SAM) is one of the most commonly used SMTs in remote-sensing applications (Yuhas, Goetz, and Boardmann 1992; Chang 2000; Cho et al. 2010; Vyas et al. 2011; Hillnhutter et al. 2011; Dudeni et al. 2009; Baoxin, Levesque, and Ardouin 2008; Zhang et al. 2006; Mundt et al. 2005). In addition, spectral information divergence (SID) (Chang 2000) and the combination of SID and SAM (Du et al. 2004) are the other SMTs that have been proposed to improve the performance of SAM.

The work presented in this article is a continuation of our recent study reported by Ghiyamat et al. (2013). In that article, it was shown that the performance of original reflectance in discriminating tree species with different ages can be improved by using multiple endmembers (MEMs). In this study, the main aim is to test the ability of DWT to discriminate tree species with different ages. The study was carried out in Thetford Forest, a planted forest in eastern England. The first objective of this study is to investigate how DWT with different mother wavelets can be discriminated among six vegetation cover types in the study area, namely three tree species (Corsican pines, Scot pines, and broadleaved species) with three age classes. The second objective is to examine the suitability of the three SMTs (SAM, SID, and a combination of SAM with SID) with DWT for discriminating tree species with different ages. Since the use of reflectance and first and second derivatives in discriminating vegetation types is frequently reported in the literature, they are considered to be the benchmark against which DWT is compared in this study. For this purpose, the performances of five different mother wavelets with a total of 50 different orders are compared against reflectance, first, and second derivative spectra using the three SMTs.

2. Methodology

2.1. Study site and data

The study was conducted in the Thetford Forest in East Anglia (0° 41′ 40.89″ to 0° 43′ 50.96″ N and 52° 26′ 40.43″ to 52° 25′ 14.72″ E). It is a conservation forest, considered to be the largest man-made lowland pine forest in Britain, which occupies an area of approximately 22,000 ha. The forest area consists mainly of planted and managed Corsican and Scots pines of different age classes. There are six different vegetation covers including old Scots pine (OSP), young Scots pine (YSP), mature Corsican pine (MCP), young Corsican pine (YCP),

old Corsican pine (OCP), and broadleaved (BL), where the young, mature, and old trees had ages of around 16, 34, and 70 years, respectively.

The HyMap (Hyperspectral Mapper) image data as shown in Figure 1 were acquired on 17 June 2000 using the HyMap sensor as part of the Synthetic Aperture Radar and Hyperspectral Airborne campaign run by the Natural Environment Research Council and the British National Space Centre. The hyperspectral data have a spatial resolution of 5 m and an average spectral resolution of 15 nm, consisting of 126 bands from 0.45 μm to 2.48 μm. Detailed spectral characteristics of the HyMap data are shown in Table 1. The imagery was

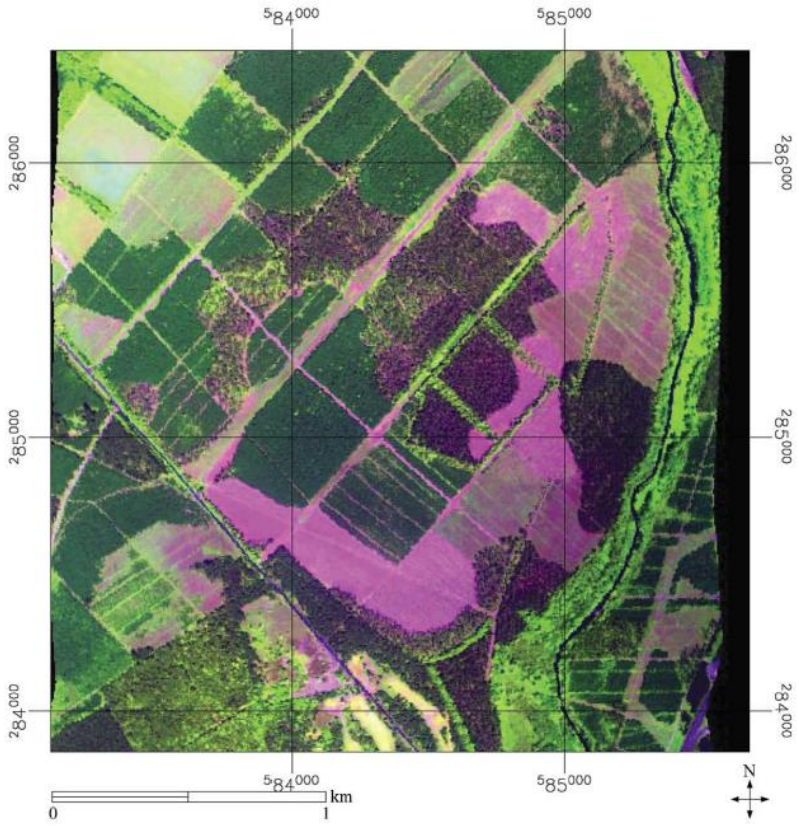

Figure 1. Hyperspectral data acquired from Thetford forest of the UK (red, 1647.8 nm; green, 829.2 nm; blue, 661.7).

Figure 1. Hyperspectral data acquired from Thetford forest of the UK (red, 1647.8 nm; green, 829.2 nm; blue, 661.7).

Table 1. Spectral characteristics of the HyMap sensor.

	Spectral configuration		
Module	Spectral range (μm)	No. of bands	Average spectral sampling interval or spectral resolution (nm)
VIS	0.45–0.89	31	15
NIR	0.89–1.35	31	15
SWIR1	1.40–1.80	32	13
SWIR2	1.95–2.48	32	17

atmospherically corrected by using the hyperspectral correction algorithm. The overlapping scenes were also georectified, mosaicked, and normalized to minimize the effect of the sensor look angle. The HyMap sensor provides an excellent image quality with a signal-to-noise ratio (SNR) of >500:1. The SNR analysis was conducted for removing noisy atmospheric water absorption bands from the original data set (Patenaude et al. 2008).

For the Thetford hyperspectral data, ground reference data generated from the UK Forestry Commission's geographical information system (GIS) vector data and stock map are available, in which the scanned image of the digitized and vectorized ground reference data is shown in Figure 2. Each vegetation cover in the hyperspectral data is accordingly labelled in the ground reference data. As mentioned in the introduction, there are six different vegetation covers in the ground reference data, namely OSP, YSP, MCP, YCP, OCP, and BL. Several groups of pixels in the form of the region of interest (ROI) were selected from each vegetation cover. Since the individual tree crowns in the hyperspectral image are not distinguishable due to the low spatial resolution (5 m), each ROI might contain pixels from more than one tree in the same vegetation cover. Therefore, the analysis in this study is based on pixel detection. The number of ROIs and the total number of pixels per vegetation cover that are used in this study are shown in Table 2. Each ROI in this study contains about 50 pixels. The selected ROIs for each vegetation cover are presented in Figure 2. For

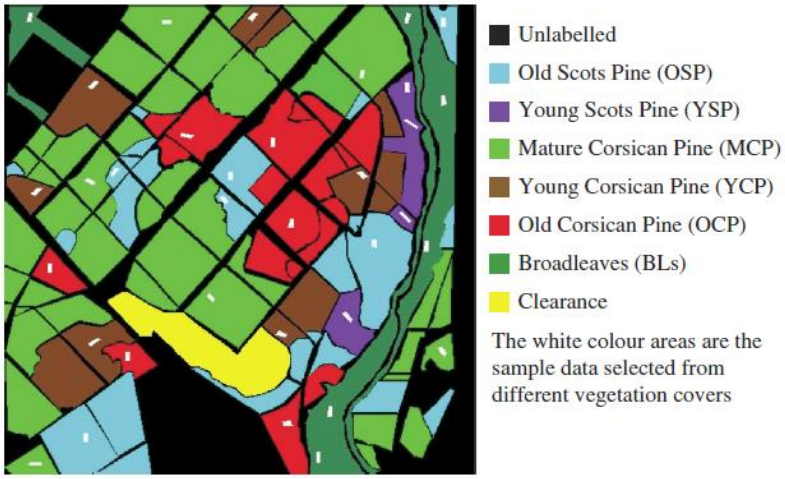

Figure 2. Ground reference image of the hyperspectral data. The selected ROIs for this study are presented in the ground reference image.

Table 2. The number of ROIs and total pixels per tree species selected from hyperspectral data.

Species	BL	MCP	OCP	YCP	OSP	YSP
No. of ROIs	6	6	6	6	6	4
Total no. of pixels	321	308	306	318	328	216
No. of testing pixels	231	218	216	228	238	156

comparison purposes, the same sets of ROIs used in our previous research (Ghiyamat et al. 2013) are used in the present study. This can be helpful for comparing the performance of WT with MEMs in discriminating the same sets of data with similar sample pixels.

2.2. *Wavelet-transform*

In the past two decades, WT has been developed as a powerful analytical tool for signal processing and is now being used in remote-sensing applications (Bradshaw and Spies 1992; Bruce, Li, and Huang 2002; Koger et al. 2003). Bruce and Li (2001) indicated that the wavelet-based method is feasible and practical for the analysis of hyperspectral signatures, and specifically for computing scale-space images and spectral fingerprints.

As noted above, with CWT, multidimensional signals such as image cubes can be analysed across a continuum of scales. DWT, on the other hand, analyses signals over a discrete set of scales, and typically contains $\log_2 n$ segments of various transforming lengths 2^n (2^n, $n = 1, 2, 3, \ldots$). DWT has an advantage over the CWT, where it can be implemented using a variety of fast algorithms, and thus has fewer computational requirements (Bruce, Morgan, and Larsen 2001).

The WT decomposes the hyperspectral signal into sets of coefficients. Each set is associated with a spectral scale and each element in a set associated with a particular wavelength location. A set of wavelet basis functions, $\{\psi_{a,b}(\lambda)\}$, can be formed from the mother wavelet, $\psi(\lambda)$, by a series of scaling and shifting operations using:

$$\psi_{a,b}(\lambda) = \frac{1}{\sqrt{a}}\psi\left(\frac{\lambda - b}{a}\right), \tag{1}$$

where a and b are real numbers in which the variable a is greater than 0, and it indicates the scale (or width) of a particular basis function, and the variable b specifies its shifted position. The wavelet coefficients for the DWT, denoted by $W_{j,k}$, are defined by the scalar product of the hyperspectral signal as a function $f(\lambda)$ and the scaling function (i.e. the wavelet basis function or mother wavelet) $\psi(\lambda)$, which can be obtained by

$$W_{j,k} = \left\langle f(\lambda), \psi_{j,k}(\lambda) \right\rangle, \tag{2}$$

where the wavelet basis function $\psi_{j,k}(\lambda)$ can be calculated by

$$\psi_{j,k}(\lambda) = 2^{\frac{-j}{2}}\psi\left(2^{-j}\lambda - k\right), \tag{3}$$

where j is the jth decomposition level or step and k is the kth wavelet coefficient at the jth level. The scales of DWT are $a = 2, 4, 8, \ldots, 2^j, \ldots, 2^p$, where p indicates the maximum decomposition level. Theoretically, the maximum level of decomposition is referred to a level such that the set of detailed coefficients corresponding to this level consists of only one element. Therefore, if the initial signal length is N, then the maximum level of decomposition cannot exceed $\log_2(N)$. In practice, however, the maximum number of decomposition levels also depends on the choice of the mother wavelet (Bruce, Morgan, and Larsen 2001), which can be calculated as

$$p = \mathrm{fix}(\log_2(N) - \log_2(M-1)), \tag{4}$$

where M is the length of the mother wavelet, which depends on the type and order of mother wavelet, and *fix*(.) is to round the value in the parentheses to the nearest integers towards zero. For example, the maximum decomposition level for a signal with a length of $N = 21$ and the mother wavelet of Daubechies-3, which has a length of 6, is equal to 2.

The DWT has been extensively used in the development of fast wavelet algorithms. The most common implementation of DWT is the well-known dyadic filter tree algorithm developed by Mallat (1989). The term dyadic refers to the resolution between two scales of WT, which is decreased by a factor of two. Figure 3 illustrates the dyadic filter tree algorithm for implementing DWT. The input to the filter tree, $f(\lambda)$, is the hyperspectral signal (e.g. reflectance spectra), and the signal is then passed through a series of high-pass and low-pass filters (HPF and LPF). After passing the input signal through the filters in each scale, the signal length is degraded by a factor of two. The outputs of the HPF and LPF at scale a are called the wavelet coefficient detail and approximation (CD_a and CA_a), respectively. At each scale or stage, the wavelet approximation coefficients from the previous scale are used as the input to the next stage.

The final result of the DWT decomposition of a spectrum is a set of wavelet coefficients that are represented as a vector: $\boldsymbol{W} = [\mathrm{CA}_p, \mathrm{CD}_p, \mathrm{CD}_{p-1}, \ldots, \mathrm{CD}_2, \mathrm{CD}_1]$, where p is the coarsest decomposition level. At a particular scale, each wavelet coefficient is directly related to the amount of energy in the signal. It should be noted that $\boldsymbol{W}$ is dependent on the selection of the type and order of mother wavelet (Zhang et al. 2006).

A variety of parameters can be computed from the DWT decomposition, such as the wavelet coefficients, their energy, and any combination of the two. While the wavelet energy feature represents the energy distribution of the original spectrum across different scales, the wavelet detail coefficient reveals the spectral information in hyperspectral signals at a specific spectral scale. In this study, each wavelet CD at different decomposition levels individually, and $\boldsymbol{W}$ (combination of all CDs plus the last approximate coefficient (hereafter denoted C-All)) are used as the main features for discriminating tree species. In addition, five different wavelet families such as Haar, Daubechies (db), Coiflet (coif), Symlet (sym), and Biorthogonal (bior) with different orders are used in this study. These five common mother wavelets are used

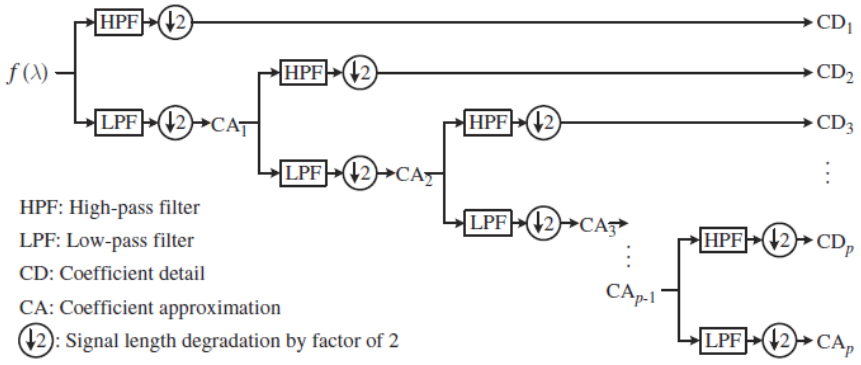

Figure 3. Dyadic filter tree algorithm for DWT implementation.

Table 3. The maximum number of decomposition levels for different mother wavelets and orders.

Wavelet family and orders	Maximum decomposition levels based on 128 bands
Haar	7
db 2	5
db 3 to db 4	4
db 5 to db 8	3
db 9 to db16	2
sym 2	5
sym 3 to sym 4	4
sym 5 to sym 8	3
sym 9 to sym 16	2
coif 1	4
coif 2	3
coif 3 to coif 5	2
bior 1.3	4
bior 1.5	3
bior 2.2	4
bior 2.4 and 2.6	3
bior 2.8	2
bior 3.1	5
bior 3.3	4
bior 3.5 and 3.7	3
bior 3.9	2
bior 4.4 and 5.5	3
bior 6.8	2

in this study as examples to represent how DWT can be helpful in discriminating tree species with different ages.

To apply WT, the sample data (here that are reflectance spectra) were directly used for wavelet analysis using Matlab software. Since 126 spectral bands are available in the hyperspectral data, two artificial bands were added into the end of each of the sample spectra to have the maximum possible decomposition level of seven ($2^7 = 128$) rather than six. The reflectance value for the new (or artificial) bands was assigned by calculating the average of bands 123 and 124, and 125 and 126. In this study, the maximum decomposition levels for the five different wavelet families and different orders were determined using Equation (4) as shown in Table 3.

From the five different wavelet families, a total of 50 different orders and 195 coefficients (as shown in Table 3) are used for discriminating tree species. Since db1, sym1, and bior1.1 are the same as Haar, they are not considered in this study. Except for the detail coefficient in the seventh decomposition level in the Haar mother wavelet, which has only one sample, all other detail coefficients from different mother wavelets and orders are considered. As mentioned above, in addition to all CDs per wavelet order, C-All is also used for discriminating tree species.

2.3. *Classification approaches*

SMTs are useful tools for discriminating different spectral signatures. As mentioned in Section 1, three different SMTs, namely SAM, SID, and a combination of SID with a