Exploitation of Very High Resolution Satellite Data for Tree Species Identification

A. Carleer and E. Wolff

Abstract
With the emergence of very high spatial resolution satellite images, the spatial resolution gap which existed between satellite images and aerial photographs has decreased. A study of the potential of these images for tree species in “monoculture stands” identification was conducted. Two Ikonos images were acquired, one in June 2000 and the other in October 2000, for an 11-by-11 km area covering the Sonian Forest in the southeastern part of the Brussels-Capital region (Belgium). The two images were orthorectified using a digital elevation model and 1256 geodetic control points. The identification of the tree species was carried out utilizing a supervised maximum-likelihood classification on a pixel-by-pixel basis. Classifications were performed on the orthorectified data, NDVI transformed data, and principal components imagery. In order to decrease the intraclass variance, a mean filter was applied to all the spectral bands and neo-channels used in the classification process. Training and validation areas were selected and digitized using detailed geographical databases of the tree species. The selection of the relevant bands and neo-channels was carried out by successive addition of information in order to improve the classification results. Seven different tree species of one to two different age classes were identified with an overall accuracy of 86 percent. The seven identified tree species or species groups are Oaks (Quercus sp.), Beech (Fagus sylvatica L.), Purple Beech (Fagus sylvatica purpurea), Douglas Fir (Pseudotsuga menziesii (Mirb.) Franco), Scots Pine (Pinus sylvestris L.), Corsican Pine (Pinus nigra Arn. subsp. iaricio (Poir.) Maire var. corsican), and Larch (Larix decidua Mill.).

Introduction
The spatial resolution of satellite multispectral data has steadily increased in the last 20 years. Since 1982, the Landsat Thematic Mapper (TM) has collected images with a spatial resolution of 30 m. Since 1986, SPOT XS has a resolution of 20 m, and Ikonos-2 has been providing multispectral images at 4 m since 1999. The use of remotely sensed data in forestry has a long history (Bergen et al., 2000; Olson and Weber, 2000). Updating forest cover, mapping of clearcuts (Cohen et al., 1998) and deforestation caused by fire (Sunar and Ozkan, 2001), estimating biomass (Roy and Ravan, 1996), and monitoring the health of forests (Reich and Price, 1999) are just a few examples.

Since the launching of the Ikonos-2 satellite in September 1999, new opportunities for forest mapping exist. The acquisition of very high spatial resolution images of Brussels and the Sonian Forest (Figure 1) made it possible to study the potential of these images for the tree species identification in the particular context of an old growth beech (Fagus sylvatica L.) forest.

Background
Since 1940, aerial photographs have been used in forestry, in particular in the United States (Lachowski et al., 2000). Since then, both the platforms and the sensors have increased, offering varied spatial resolutions and, consequently, prospects for analysis on various scales. The choice of the most suitable data will depend on the relation between the size of the analyzed objects and the spatial resolution and extent of the available images (Key et al., 2001). One should also take into consideration spectral and temporal image characteristics.

The use of high-resolution multispectral sensors like Landsat TM, SPOT, IRS, MOS, and others, which have a spatial resolution from 20 to 60 m, is helpful (Roller, 2000) for identifying several classes of forests. For example, White et al. (1995) carried out an unsupervised classification of bands 1, 2, 3, 4, 5, and 7 of Landsat TM in 20 classes regrouped into four classes (genus Pinus, genus Abies, nonwooded, without vegetation) with an accuracy of 73 percent.

Until 1999 the spatial resolution of available satellite images was incompatible with the geometrical precision and level of detail required by the forest managers for forest mapping on a local scale. The study of such sites at scales ranging from 1:5000 to 1:500 was carried out using photographs or aerial images having a sufficiently high resolution. The fineness of the grain of the aerial photographs ensures the visualization of many details, even with little contrast. They can be scanned for processing and computer-assisted interpretation. They can be viewed stereoscopically to interpret species and size classes (Caylor, 2000). Aerial photography for vegetation mapping remains an effective and widely used tool. The scanning of these aerial photographs for their interpretation and their integration into a geographic information system (GIS) is now commonplace. Images in digital format offer advantages. They can be observed at a great range of scales. Contrast can be improved to ease the interpretation of derived images, such as vegetation indices and classifications. Geometrical distortions can be corrected to create an image which has the geometrical characteristics of a map (Coulter et al., 2000). Meyer et al. (1996) and others used such techniques to carry out the identification of five tree species in the forest of Unterwald in the Canton of Aargau in Switzerland, starting from scanned aerial photographs.

The principal disadvantage of aerial photographs compared to satellite images, in addition to their higher price for the same coverage, is the need to mosaic the photos if the
study area is larger than a single photo. An other disadvantage of photography is that it does not contain the multispectral information in discrete bands that may be transformed. The launching of Ikonos-2 in September 1999 opened new possibilities for direct use of forest mapping by forest managers (MILA, 2001).

**Objective**

The objective of this research was to understand the potential of very high spatial resolution multispectral satellite images for forest type mapping and age class delineation. Currently, for the study area, foresters have maps of parcels of land, each of which is characterized in a database. However, they do not have the location of the tree species within the parcels.

**The Study Area**

This study relates to the Brussels part of the Sonian Forest which is located in the northern part of the forest (Figure 1). The Sonian Forest covers 4383 ha. Since the federalization of the Belgian State in 1983, the forest has been divided between three regions of the country: 56 percent in the Flemish region, 38 percent in the Brussels-Capital region, and 6 percent in the Walloon region. Each region is responsible for the management of its part of the Sonian Forest with sometimes extremely different objectives (van der Ben, 1997). The portion of the Sonian Forest in the Brussels-Capital region covers 1657 ha. This part of the forest alone accounts for 60 percent of the Brussels parks area open to the public. The recreational function of this forest space, representing an extension of Brussels, is very significant.

The Sonian Forest is known for its cathedral aspect of beech areas, consisting of old beeches with slender trunks and almost non-existent underbrush. It covers nearly 65 percent of the forest acreage. This cathedral aspect is not natural. It results from the restoration of the forest undertaken under the Austrian regime in Belgium (1714–1795).

Most tree stands are pure, and the composition of the tree species or species groups in that part of the forest in the Brussels-Capital region is 74 percent beech (Fagus sylvatica L.), 16 percent oak (Quercus sp.), 8 percent coniferous, and 2 percent various other species. The 8 percent of coniferous trees are composed of 47 percent Scots pine (Pinus sylvestris L.), 32 percent larch (Larix decidua Mill.), 13 percent Corsican pine (Pinus nigra Arn. subsp. larcicio (Poir.) Maire var. corsican), 6 percent Douglas fir (Pseudostuga menziesii (Mirb.) franco), and 2 percent spruce (Picea abies L.) (I.B.G.E., 2000).

**Data and Preprocessing**

The acquisition of two Ikonos images was preplanned, and they were acquired on 08 June and 07 October, 2000. The images cover an area of 11 by 11 km. This area extends from north of the historical center of Brussels to the southern boundaries of the Brussels-Capital region, thus covering the Brussels part of the Sonian Forest (Figure 1).

The June image is of excellent quality, with no clouds or haze. The October image was recorded as two scenes, with an overlap of approximately 900 m between the scenes. This image is of poor quality: approximately 15 percent of the area is covered by haze and 2 percent by clouds and their shadows. Usefulness of the imagery will be limited.
The resolution of the images is 1 m in the panchromatic band and 4 m in the multispectral bands. The five spectral bands are 450 to 900 nm for the panchromatic band, 450 to 520 nm (blue), 520 to 600 nm (green), 630 to 690 nm (red), and 760 to 900 nm (near-infrared). These multispectral wavelength intervals are the same as bands 1 to 4 of the Thematic Mapper. For Europe, the images are marketed by Space Imaging Europe. The Imagery images that were acquired were georeferenced by Space Imaging Europe with a horizontal precision of approximately 50 m. This precision was calculated. It is the most common index used to highlight those of October, and on all the bands of the two images.

To identify and map the Sonian Forest tree species, we used a multispectral procedure of supervised classification by pixel. To support the legend, an unsupervised classification was carried out and the resulting classes were interpreted using all the existing benchmark data. This resulted in retaining the common and easiest to discriminate classes, i.e., roads, water, lawns and clearings, young beeches, and old beeches. Also, using the reference data, the four most abundant species of conifers in the Sonian Forest were retained as classes, i.e., Scots pine, larch, Corsican pine, and Douglas fir. We considered oaks as well because they cover surface areas, certainly small, but important. The last selected class was the purple beech, which we could identify visually on the color composite in true or false colors. We could also have taken into account chestnut, poplar, and ash, which are the three other most significant leafy tree species in the Sonian Forest, but they are never present there in pure stands; it is thus impossible to identify training and validation areas for these classes.

Classifications were carried out using a maximum-likelihood classifier on different selections of spectral bands and neo-channels. The 11 selected classes are included in Table 1. To eliminate the salt-and-pepper effect from the result, caused by a small percentage of isolated pixels—generally mis-classified and often located at the limit between two distinct land cover zones—a modal filter was applied to the classified image in a 3- by 3-pixel window. Finally, a classification assessment was carried out by comparing the result to the reference data and by computing a confusion matrix and Kappa coefficient.

TABLE 1. Classes selected for Supervised Classification

<table>
<thead>
<tr>
<th>Legend</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Oak</td>
</tr>
<tr>
<td>2</td>
<td>Scots pine</td>
</tr>
<tr>
<td>3</td>
<td>Corsican pine</td>
</tr>
<tr>
<td>4</td>
<td>Larch</td>
</tr>
<tr>
<td>5</td>
<td>Young Beech</td>
</tr>
<tr>
<td>6</td>
<td>Old Beech</td>
</tr>
<tr>
<td>7</td>
<td>Purple Beech</td>
</tr>
<tr>
<td>8</td>
<td>Lawn and clearing</td>
</tr>
<tr>
<td>9</td>
<td>Road</td>
</tr>
<tr>
<td>10</td>
<td>Water</td>
</tr>
<tr>
<td>11</td>
<td>Douglas Fir</td>
</tr>
</tbody>
</table>
Results

The principal component analysis on the four spectral bands of the June and October images concentrated 97.8 percent of the initial variance of the two images in the first three components. Unfortunately, clouds and haze were present in the second component; this component represented 18.7 percent of the initial variance of the two images. Classification could thus not be carried out on a selection of components. Nevertheless, the first component, not being affected by haze and clouds, could be used. Because of this, little information from the October image could be taken into account. Indeed, the correlation coefficients between the first component and the various spectral bands of the October image were 0.25 for the near-infrared band, 0.025 for the red band, 0.026 for the green band, and 0.011 for the blue band.

Clouds and haze affected all the components of the PCA achieved for the October image. None of them was usable for the study.

A series of supervised classifications was then carried out by successive addition of information in order to improve the results of classification. The best results were obtained with the red, green, blue, and near-infrared spectral bands of the filtered June image, the June NDVI, and the first component of the PCA of the four spectral bands of the June and October images.

The visual assessment of the classified image revealed, first of all, a good identification of species in the Sonian Forest as well as other land-cover types present in the forest. Nevertheless, an edge effect was observable (Figure 2).

This edge effect is due to the application of a mean filter on the spectral bands and neo-channels before classification. The mean filter generates intermediate values at the limit between two homogeneous and well differentiated zones (mixed pixels). This phenomenon is rather recurrent around conifers and water areas. Despite the edge effect due to the mean filter, the mean filter benefit remains important. The quantitative assessment of the classified filtered image showed an overall accuracy of 86 percent and a Kappa coefficient of 0.84 (Table 2) against an overall accuracy of 79 percent and a Kappa coefficient of 0.76 for the classified non-filtered image.

The confusion matrix (Table 2) showed that the classes “purple beech,” “lawn and clearing,” and “road” were very well identified. It also identified omission and confusion errors. We observed a significant confusion between “oak” and “old beech.” This confusion was due to the proximity of the spectral signatures of these leafy tree classes. This is supported by the fact that these classes are not well discriminated in an unsupervised classification. The matrix also showed confusion between conifers and “old beech.” This confusion was undoubtedly due to the edge effects but also to the inclusion of shadows in the training zones.

Confusion between conifers was due to similarities of their spectral signatures. The unsupervised classification that was performed could not discriminate these classes. The conifers were isolated but the distinction between them remained poor.
Conclusions

This study contributes to widening the case studies aimed at interpreting species at the beginning of multispectral image processing of very high spatial resolution imagery. Seven different tree species (Oak, Scots Pine, Corsican Pine, Larch, Beech, Purple Beech, and Douglas Fir) were identified, including one at two different ages (young Beech and old Beech). These seven tree species are the most common species in the Sonian Forest. Other land covers present in the Sonian Forest were also distinguished: “lawn and clearing,” “road,” and “water.” The overall accuracy obtained was 86 percent, which is comparable to other similar studies (M.I.L.A., 2000). The new map of the species of the Sonian Forest achieved on the basis of the very high resolution images brings information supplementing the existing databases.

The introduction of better quality multimodal data than the October images into the spectral procedure of classification would undoubtedly allow a better discrimination between the various species. The comparison of images at several dates could also indicate other types of change such as cuts in the forest, appearance of windfall wood, or the detection of diseases generating defoliation or a chlorosis.

This study also reveals the limits of multispectral classification at the pixel level for the classification of very high resolution images. The use of a mean filter proved to be essential to reduce the intraclass variance, but it generated mixed pixels, themselves sources of error in classification. This problem should be solved by applying classification by region; such an approach seems promising (Debeir et al., 2001). It exploits the spectral and textural characteristics during image segmentation in regions and takes into account higher level parameters such as the shape, the size, and the organization in classification. New methods and techniques have to be tested and used to improve the use of these very high resolution images. This will represent a permanent challenge of adaptation and methodological renewal.

Acknowledgments

The IKONOS images were acquired by the OSTC (Federal Office of Scientific, Technical and Cultural Affairs) within the framework of a feasibility study carried out by the I.G.E.A.T. (Institut de Gestion de l’Environnement et d’Aménagement du Territoire, U.L.B.) within the framework of the TELSAT 4 program. Alexandre Carleer would like to thank the FRIA (Fonds pour la Formation à la Recherche dans l’Industrie et l’Agriculture) for its support.

References


(Received 17 April 2002; accepted 03 January 2003; revised 12 February 2003)