The Photogrammetric Journal of Finland, 20 (1), 2006 Received 13.7.2006, Accepted 13.11.2006

IMPROVING AUTOMATION IN RULE-BASED INTERPRETATION OF REMOTELY SENSED DATA BY USING CLASSIFICATION TREES

Leena Matikainen

Finnish Geodetic Institute, Department of Remote Sensing and Photogrammetry

Leena.Matikainen@fgi.fi

ABSTRACT

The definition of good classification rules for rule-based interpretation of remotely sensed data is a laborious and demanding task. One interesting method that could be used to automate the process is the classification tree method. It can be used to create a tree-structured classification hierarchy and rules automatically from training data. In this study, tests were carried out using the classification tree method in two applications: building detection using laser scanner and aerial image data and land-use classification using E-SAR data. The method was applied to segments with a large number of different attributes. The results were satisfactory and the classification accuracy was near to that obtained in previous studies using manually created classification rules. The most important benefit of the classification tree method was its high level of automation and speed compared with the process of defining the rules manually. A combination of the classification tree method and permanent, up-to-date reference data could be a useful tool in developing new classification applications and testing the feasibility of new remotely sensed datasets. Together with segments and attributes derived from remotely sensed data, it could be used for the rapid construction of classification trees, which could then be directly applied to classification or used as a starting point for further development of the rules.

1. INTRODUCTION

In the analysis of remotely sensed data with automatic methods, it is often advantageous to use more versatile information than the single pixel values of one dataset. This information can include properties of regions (regions derived from the imagery and/or map data), multisource remotely-sensed data, and ancillary data such as maps and geographical information system (GIS) data. Standard image classification methods, however, can encounter difficulties in dealing with such data, which do not necessarily have a certain statistical distribution and which can include both continuous and categorical data with different ranges of values. To overcome the problem, different knowledge-based interpretation and classification methods have been developed (see, for example, Richards and Jia, 1999; Jensen, 2005). These methods are more flexible, but they also have limitations. In particular, they are often based on rules defined by a human interpreter, and the development of good classification rules is usually a difficult and laborious task. When a new dataset becomes available, considerable time and expertise are needed to exploit it efficiently in practical applications. In recent years, many new types of remotely sensed data have become available, such as laser scanner data, digital aerial images, high-resolution optical satellite images, and synthetic aperture radar (SAR) images from airborne and spaceborne sensors. There is plenty of valuable information in these datasets, and it would be useful if this information could be exploited more rapidly.

This article discusses the application of one promising method – classification trees (as described by Breiman et al., 1984) – to the interpretation of remotely sensed data. The basic idea of classification trees (also called decision trees) is to perform stepwise splitting of the data into classes that are arranged in a hierarchical structure (see, for example, Richards and Jia, 1999; Tso and Mather, 2001). Tree structures are often used in knowledge-based classification approaches. The tree structure and classification rules associated with it can be defined manually, but in this study we were interested in automatic methods for creating classification trees. Many different approaches to this have been presented in the literature (see Safavian and Landgrebe, 1991).

Classification or decision trees, typically created with data mining or statistical software tools, have been used increasingly in the classification of remotely sensed data in recent years. They have been used for land-cover/land-use classification (e.g. Hansen et al., 1996; Friedl and Brodley, 1997; Thomas et al., 2003), change detection (Chan et al., 2001; Rogan et al., 2003; Im and Jensen, 2005), detection of buildings from aerial images (Jung, 2004) and from IKONOS data (Tullis and Jensen, 2003), mapping of residential density patterns (McCauley and Goetz, 2004), forest mapping (Huang and Lees, 2004), and many other types of study. Classification trees are well suited to dealing with heterogeneous datasets because they can use continuous and categorical data and do not require assumptions on the distributions of the data (e.g. Hansen et al., 1996; Friedl and Brodley, 1997). Classifications with the classification tree method can be highly automatic, which yields savings in time (see, for example, Thomas et al., 2003). The classification trees are also considered easier to use and understand than artificial neural networks, which are another popular non-parametric classification method (see, for example, Hansen et al., 1996; Chan et al., 2001; Pal and Mather, 2003). The tree structure provides information on the roles and importance of different attributes (features) in the classification problem, which can be very useful, from both the practical and the theoretical perspective (Hansen et al., 1996).

The classification tree method can be applied to pixel-based classification of image data in much the same way as more traditional classification algorithms are applied. On the other hand, it can be used to generate rules for knowledge-based and possibly region-based classification with different types of attributes. Huang and Jensen (1997) used the C4.5 machine learning algorithm (Quinlan, 1993) to create a decision tree and store it as production rules. The application under study was wetland classification using SPOT (Satellite Pour l'Observation de la Terre) multispectral imagery and GIS data. Hodgson et al. (2003) and Tullis and Jensen (2003) applied a similar approach using the newer See5 algorithm (RuleQuest Research, 2006). Hodgson et al. (2003) studied the mapping of urban parcel imperviousness using colour aerial photography and laser scanner-derived height information. Tullis and Jensen (2003) studied the detection of houses from IKONOS data. Lawrence and Wright (2001) used classification and regression tree analysis (Breiman et al., 1984) available in the S-Plus statistical software package to create rules for a land-cover/land-use classification. Landsat Thematic Mapper (TM) images and a digital elevation model (DEM) were used in the study. Other studies using knowledge-based approaches have also been presented (Li et al., 2000; Rogan et al., 2003; Thomas et al., 2003). In some of the studies, region-based classifications were carried out. For example, Hodgson et al. (2003) and Thomas et al. (2003) classified segments created with the eCognition software (Definiens, 2006). As input data for construction of the classification tree using the S-Plus software, Thomas et al. (2003) used a large number of different attributes available for the segments from eCognition. The study was related to land-cover/land-use mapping in an urban environment using high-resolution digital imagery. For actual classification with the rules, the studies have used different software tools, such as the ERDAS Imagine Expert Classifier (e.g. Lawrence and Wright, 2001) or Avenue scripts in ArcView (e.g. Hodgson et al., 2003).

Although the classification tree method has become relatively popular in recent years, its use in some applications is still at an early stage. One such application seems to be the analysis of laser scanner data. As mentioned above, Hodgson et al. (2003) used laser scanner-derived height information in addition to image data in their study. Another study in which the method was applied to laser scanner data was carried out by Ducic et al. (2006). In this study, laser points were classified as vegetation points and non-vegetation points on the basis of full-waveform information. To classify SAR images the classification tree method has been used by, for example, Townsend (2001) and Simard et al. (2002).

In our study, the classification tree method was tested for two different applications and datasets: building detection using laser scanner and aerial image data, and land-use classification using E-SAR data. The results were compared with earlier results obtained using manually defined classification rules (Matikainen et al., 2004a; unpublished results, 2006). Both datasets were from the Espoonlahti suburban study area near Helsinki. The methods used in the study, including the theoretical background of the classification tree tools and the workflow used in the classification tests, are described in Section 2. Section 3 presents the classification tests and results. The results and possible future uses of the method are discussed in Section 4, and the conclusions are presented in Section 5.

2. METHODS

2.1 Classification tree tools

We used the classification (and regression) tree tools available in the Statistics Toolbox of the Matlab software (The MathWorks, 2006). The tools can be used to construct a classification tree with a binary tree structure on the basis of training data, and to apply the tree to the classification of new data (for a detailed description of the classification tree method, see Breiman et al., 1984). A classification tree contains a root node, non-terminal nodes and terminal nodes. Examples of classification trees can be found in Sections 3.1.2 and 3.2.2. The root node and each of the non-terminal nodes contain a question that asks whether a given attribute satisfies a given condition. Each question can be answered as ‘yes’ or ‘no’. The terminal nodes represent individual classes. When an object is classified, the conditions are tested beginning from the root node. From each node, the object goes to the left descendant node if it satisfies the condition and to the right descendant node if it does not satisfy the condition. Finally, the object ends up at one of the terminal nodes and is assigned to the corresponding class. One class can be represented by several terminal nodes. There can thus be several alternative paths (sequences of questions) that lead to the same classification result. (Breiman et al., 1984; The MathWorks, 2003.)

There are two main stages when the classification tree method is used for a given classification task: construction of a tree and application of the tree to classification. The construction of a classification tree can be carried out in Matlab with the ‘treefit’ function, which uses the attributes and classes of training objects as input data. The function then selects the most useful attributes and splits automatically using a splitting criterion. The Gini’s diversity index was used as the splitting criterion in our study. This criterion is a measure of node impurity and is defined as

$$impurity(t) = \sum_{i \neq j} p(i|t)\, p(j|t)\,, \tag{1}$$

where t is the node, and $p(i|t)$ is the proportion of cases $\boldsymbol{x}_n \in t$ which belong to class i ($\mathbf{x}$ is the measurement vector). At each node, a search is made for the split that most reduces node impurity. (Breiman et al., 1984; The MathWorks, 2003.)

When the tree is constructed, it is advantageous to continue the splitting until the terminal nodes contain a small number of training objects and the tree is large, instead of attempting to stop the splitting at the right set of terminal nodes (Breiman et al., 1984). A tree of this size, however, is larger than the data warrant and will have a higher true misclassification rate than a correctly-sized tree. Therefore, the tree must be pruned, which means that a set of smaller subtrees is obtained. The subtree that has the lowest estimated misclassification rate can be selected using, for example, cross-validation. With the 'treefit' function it is possible to compute the full tree and a sequence of pruned subtrees. The best level of pruning can be estimated with the function 'treetest'. The function computes the cost of each subtree in the optimal pruning sequence. The costs are based on the misclassification costs and probabilities of the terminal nodes. When the tree is constructed, a cost matrix C is used to define the costs $C(i,j)$ of classifying an object as class i if its true class is j. In our study, the default values $C(i,j) = 1$ if $i \neq j$ and $C(i,j) = 0$ if $i = j$ were used. The option 'crossvalidate' was used in the 'treetest' function, which means that the training data and 10-fold cross-validation were used to calculate the cost values of the subtrees. The best level of pruning given by the function was the level that produced the smallest tree within one standard error of the minimum-cost subtree. When the tree was initially created with the 'treefit' function, a threshold value of 10 (default) was used for splitting nodes, which means that a node had to contain at least 10 training objects to be split. (Breiman et al., 1984; The MathWorks, 2003.)

Once the tree has been created and the desired pruning level selected, the classification of the full dataset can be carried out with the 'treeval' function. The function takes the tree, the attribute table and the pruning level as input data and produces the class for each object as the result. (The MathWorks, 2003.)

2.2 Workflow used in the classification experiments

In both classification tests, the same basic stages were used:

1. Segmentation of the data into homogeneous regions. The segments were the objects to be classified. The segmentation was carried out with the eCognition software (Definiens Imaging, 2003), which also provides a large number of different attributes for each segment.
2. Exportation of the segments and various attributes for the segments from eCognition.
3. Definition of training segments on the basis of training data (map data or reference points).
4. Construction of a classification tree on the basis of the attributes of the training segments.
5. Classification of all segments on the basis of their attributes and the classification tree.
6. Accuracy estimation.

This basic workflow was similar to that used by Thomas et al. (2003) for land-cover/land-use classification, but we used different software tools for the classification tree construction and classification. In practice, segmentation results available from previous studies (Matikainen et al., 2004a; unpublished results, 2006) were used for stage 1. Stages 3–6 were carried out using Matlab. Simple scripts were written for the classification tree construction and classification. These scripts read the input data, used the classification tree functions available in Matlab and saved the results. The reference data used for accuracy estimation were separate from the training data. Finally, the results were compared with previous classification results, which were obtained using the eCognition software and manually created rules.

3. CLASSIFICATION EXPERIMENTS

3.1 Building detection using laser scanner and aerial image data

3.1.1 Data

In the building detection study, the goal of the classification tree classification was to distinguish building segments from tree segments on the basis of their properties in laser scanner and aerial image data. An area covering about 0.4 km^2 was used for training, and areas covering about 2.1 km^2 were used for testing the accuracy of the classification. The segments were created using a laser scanner-derived digital surface model (DSM). This raster DSM with a pixel size of 30 cm $\times$ 30 cm was formed with the TerraScan software (Soininen, 2005; Terrasolid, 2006), from last pulse laser scanner data acquired with the TopoSys FALCON system. The average point density in the original data was about 17 points per m^2 (includes overlap between adjacent strips). The lowest value within the pixel was assigned to each pixel, and interpolation was used to determine values for pixels without laser points. The original laser points were also classified using the classification routines of the TerraScan software to detect points located over 2.5 m above ground level. This point classification was used to distinguish building and tree segments from ground segments before application of the classification tree method. A segment was classified as 'building or tree' if most of the points within it had a height value of 2.5 m or over. Within each pixel, only the lowest point, which was also used in forming the DSM, was used. In addition to the laser scanner data, an aerial colour ortho image was available. The aerial imagery on a scale of 1:5300 was acquired and scanned by FM-Kartta Oy. The pixel size of the ortho image was 30 cm $\times$ 30 cm. Some more details of the data, segmentation and point classification can be found in Matikainen et al. (2004b). This earlier study used first pulse laser scanner data, but otherwise the same dataset, segmentation method and point classification method for distinguishing ground.

The following 35 attributes were exported for each segment from eCognition to be used as input data for the classification tree method (for a description and formulas of the attributes, see Definiens Imaging, 2003):

- mean value of the segment in the DSM and in each channel of the aerial image (red, green and blue),
- standard deviation of the segment in the DSM and in each channel of the aerial image,
- a texture measure 'Grey Level Co-occurrence Matrix (GLCM) homogeneity' calculated for the segment from the DSM and from each channel of the aerial image,
- shape attributes for the segment: size (area), length, width, length/width, compactness, elliptic fit, rectangular fit, border length, shape index, density, main direction, asymmetry,
- shape attributes calculated for a polygon created from the segment: size excluding inner polygons, size including inner polygons, perimeter, compactness, number of edges, standard deviation of length of edges, average length of edges, length of longest edge, number of inner objects, number of edges longer than 3 pixels, number of rectangular angles with edges longer than 3 pixels (the last two attributes with a threshold value of 10 pixels were also tested, but the resulting classification tree was the same, i.e. these attributes were not selected for classification).

The attribute GLCM homogeneity is one of the texture measures presented by Haralick et al. (1973). The general principle of these texture measures is discussed by, for example, Jensen (2005) and the implementation in eCognition by Definiens Imaging (2003). The measures are calculated from a matrix commonly called a GLCM, and they can take into account grey level variations between neighbouring pixels in different directions. The option 'all directions' was used in our study.

A building map obtained from the city of Espoo and a forest map obtained from FM-Kartta Oy were used as reference data. The map data were also processed into raster format with 30 cm × 30 cm pixels. Buildings and forest areas in the training area were used to determine training segments, and buildings in the test areas were used to test the accuracy of the building detection. A segment was used as a training segment for building or tree if over 80% of its area was classified as building or forest in the map data (some forest areas on the map were excluded because they included a considerable area covered by roads). Segments classified as ground were excluded from the training data. Compared with some ground measurements, the positional accuracy of buildings in the original building map is 0.5 m or higher. It should be noted, however, that there are many differences in the appearance of the buildings between the map and the remotely sensed data. For example, the building outlines on the map represent the ground plans of the buildings rather than the roof edges. Buildings with wide eaves are thus larger in the laser scanner and aerial image data than on the map. Small building polygons (< 20 m^2) were eliminated from the reference map to exclude very small buildings and other constructions from our accuracy estimation. Some small parts of larger buildings, however, were also eliminated in the process.

3.1.2 Classification and results

The attributes and classes of the training segments were contributed as input data to the 'treefit' function (see Section 2.1), which constructed the classification tree. When the 'treetest' function with cross-validation was used, the best pruning level for the tree was determined as 3. In practice, the script created for the construction of the tree was run a few times to find the best level of pruning. The estimated level may vary slightly between the runs because the 'treetest' function selects the subsamples for cross-validation randomly. The tree pruned to level 3 was used in classification and is shown in Figure 1. It can be seen that the tree was very simple and the classification was based on only two attributes: GLCM homogeneity (texture measure, see Section 3.1.1), calculated from the DSM, and mean value in the blue channel of the aerial image. If the GLCM homogeneity of a segment was over 0.629095 or if the mean value of the segment was over 104, the segment was classified as building, otherwise as tree.

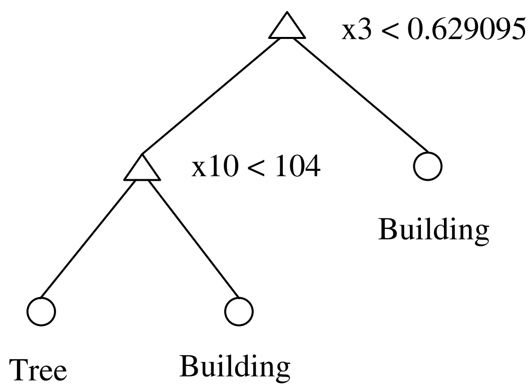

Figure 1. Classification tree for building detection. Attribute x3 is GLCM homogeneity calculated from the DSM and attribute x10 is mean value in the blue channel of the aerial image.

The classification was carried out with the 'treeval' function using the attributes, the tree and the pruning level as input data. Classification of ground was based on prior classification of laser points as described in the previous section. The classification results are shown in Figure 2. The study area can be roughly divided into an industrial area, a high-rise residential area and a low-rise residential area, and the results are shown separately for these areas. Results for two subareas in the low-rise residential area are shown on a larger scale in Figure 3. The accuracy of the results was estimated by comparing them pixel by pixel with the reference map. Accuracy measures i.e. completeness (corresponds to interpretation accuracy or producer's accuracy), correctness (corresponds to object accuracy or user's accuracy) and mean accuracy were calculated for buildings (see Helldén, 1980). The results are shown in Table 1. A small part of the high-rise area was not covered by the reference map and was thus excluded from the accuracy estimation. To allow comparisons, the accuracy of previous results obtained using the same dataset and manually created classification rules is shown in brackets (Matikainen et al., unpublished results, 2006). This previous classification used the rules described in Matikainen et al. (2004b) to distinguish buildings from trees. The ground classification was the same as that used prior to the classification tree classification.

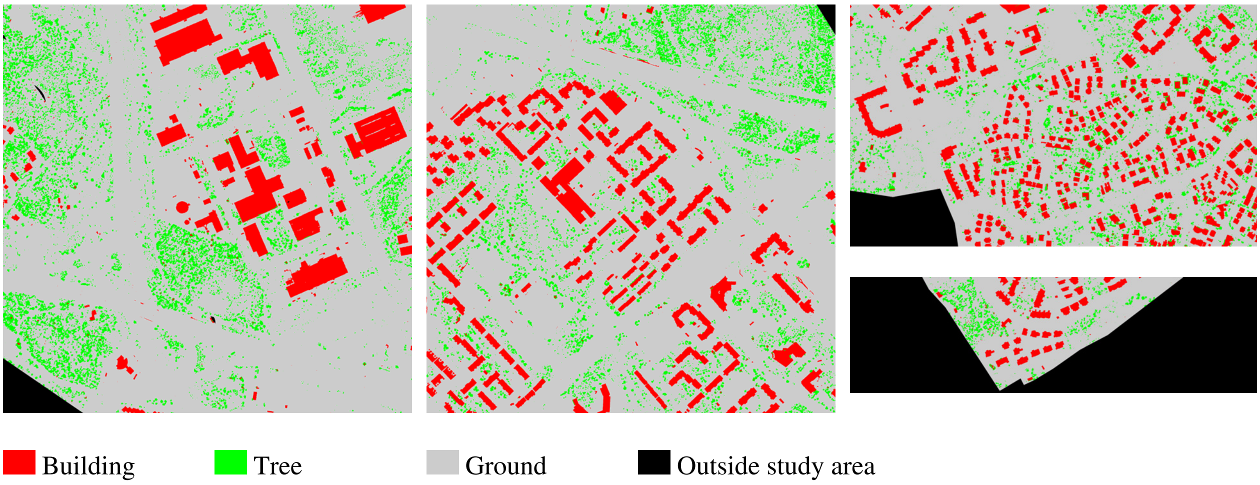

Figure 2. Building detection results for the industrial area (left), high-rise residential area (middle) and low-rise residential area (right). The width of each area is 900 m.

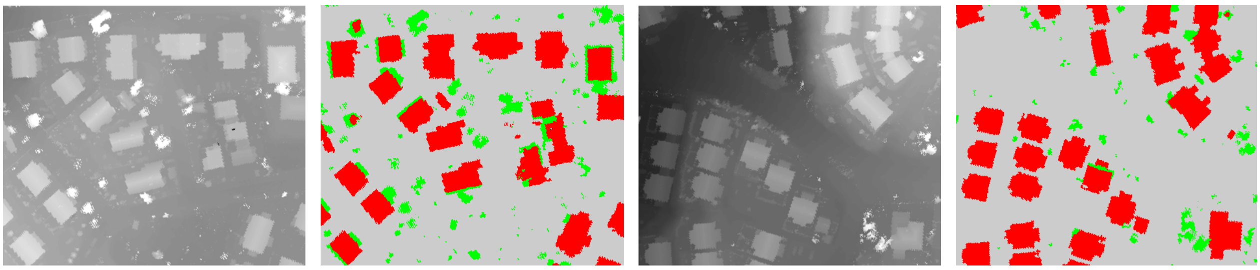

Figure 3. DSM and building detection results for two subareas of the low-rise residential area. The legend for the building detection results is presented in Figure 2.

Table 1. Accuracy of the building detection results estimated pixel by pixel. The accuracy of previous results obtained using manually created classification rules is shown in brackets. The classification of the ground was based on prior classification of laser points.

	Area			
	Industrial area	**High-rise res. area**	**Low-rise res. area**	**All**
Completeness	**98%** (97%)	**94%** (92%)	**94%** (91%)	**95%** (93%)
Correctness	**89%** (93%)	**87%** (93%)	**78%** (85%)	**84%** (90%)
Mean accuracy	**93%** (95%)	**90%** (93%)	**86%** (88%)	**90%** (92%)
Buildings classified as trees	**0%** (1%)	**0%** (2%)	**2%** (5%)	**1%** (3%)
Buildings classified as ground	**2%** (2%)	**5%** (5%)	**4%** (4%)	**4%** (4%)

3.2 Land-use classification using E-SAR data

3.2.1 Data

In the land-use classification study, the classification of the main land-use classes from high-resolution airborne E-SAR data was studied. The E-SAR data were acquired by DLR (German Aerospace Center). The goal was to recognize water, forest, open areas and built-up areas. Open areas included roads, and built-up refers to areas covered with buildings as interpreted using the SAR images (due to the side-looking geometry of the SAR sensor, buildings shifted slightly from their real position). The size of the study area was about 5.2 km^2. The SAR data included L band images with four polarizations (HH, HV, VV, VH) and X band images with two polarizations (HH, VV). The pixel size of the images on the ground was 1 m × 1 m. Simple texture images were produced by calculating the variance in a moving window (25 m × 25 m). The segmentation was carried out by using all the image channels except LVH, which is similar to LHV. In addition to the segments used for classification, a lower segmentation level with very small segments was created for the purpose of texture calculation. A total of 38 attributes were exported for each segment to be used in the classification tree classification (for a description and formulas of the attributes, see Definiens Imaging, 2003):

- mean value of the segment in channels LHH, LHV, LVV, XHH and XVV,
- standard deviation of the segment in channels LHH, LHV, LVV, XHH and XVV,
- a texture measure ‘average mean difference to neighbours of sub-objects’, calculated for the segment from channels LHH, LHV, LVV, XHH and XVV,
- mean value of the segment in the texture image calculated with the moving window from channel XHH,
- shape attributes for the segment: size, length, width, length/width, compactness, elliptic fit, rectangular fit, border length, shape index, density, main direction, asymmetry,
- texture measures based on the shape of the sub-objects (sub-segments) of the segment: mean size of sub-objects, standard deviation of the size of sub-objects, mean density of sub-objects, standard deviation of the density of sub-objects, mean asymmetry of sub-objects, standard deviation of the asymmetry of sub-objects, mean direction of sub-objects, standard deviation of the direction of sub-objects,
- mean value of the segment in LHH divided by the mean value in LHV,
- mean value of the segment in LHH divided by the mean value in LVV.

A set of 87 reference points located inside homogeneous areas was used to determine the training segments. These points were selected and classified manually on the basis of the SAR imagery, an aerial ortho image and map data. If a segment contained a point, it became a training segment of the corresponding class. A set of 519 reference points was used to estimate the accuracy of the results. The points used for training and accuracy estimation were separate sets of points, although some points were located in almost the same places in both sets. The points used for accuracy estimation were collected from aerial imagery and not specifically for the SAR image study. Some problematic points may therefore exist (e.g. buildings in the reference data but not in the SAR images). Some more details concerning the imagery, reference points and segmentation can be found in Matikainen et al. (2004a).

3.2.2 Classification and results

The classification tree obtained for the land-use classification study is shown in Figure 4. For this tree, the best pruning level was determined to be 0, i.e. the full tree without pruning. The structure of the tree was again simple, and only four attributes were used for classification. The classification results are shown in Figure 5. The confusion matrices and accuracy estimates for the classification can be found in Tables 2 and 3. In addition to the completeness, correctness and mean accuracy of the individual classes, the overall accuracy and the result of the Kappa analysis (KHAT statistic; Congalton and Green, 1999) are presented. The same information for previous classification results obtained using manually created rules (Matikainen et al., 2004a) is shown in brackets.

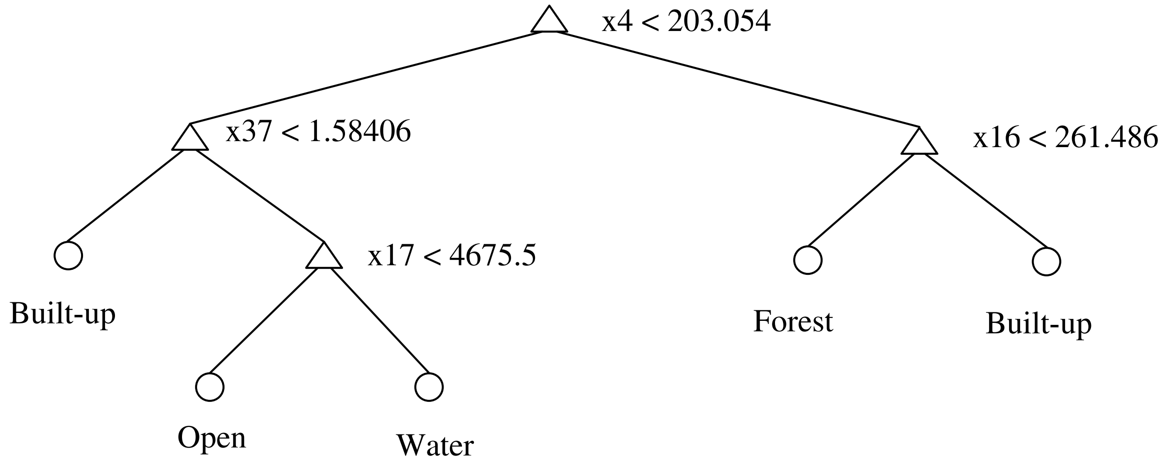

Figure 4. Classification tree for land-use classification. Attribute x4 is mean value in channel LHV, attribute x37 is mean value in LHH divided by mean value in LHV, attribute x16 is mean value in the texture image calculated with the moving window from channel XHH, and attribute x17 is the size of the segment.

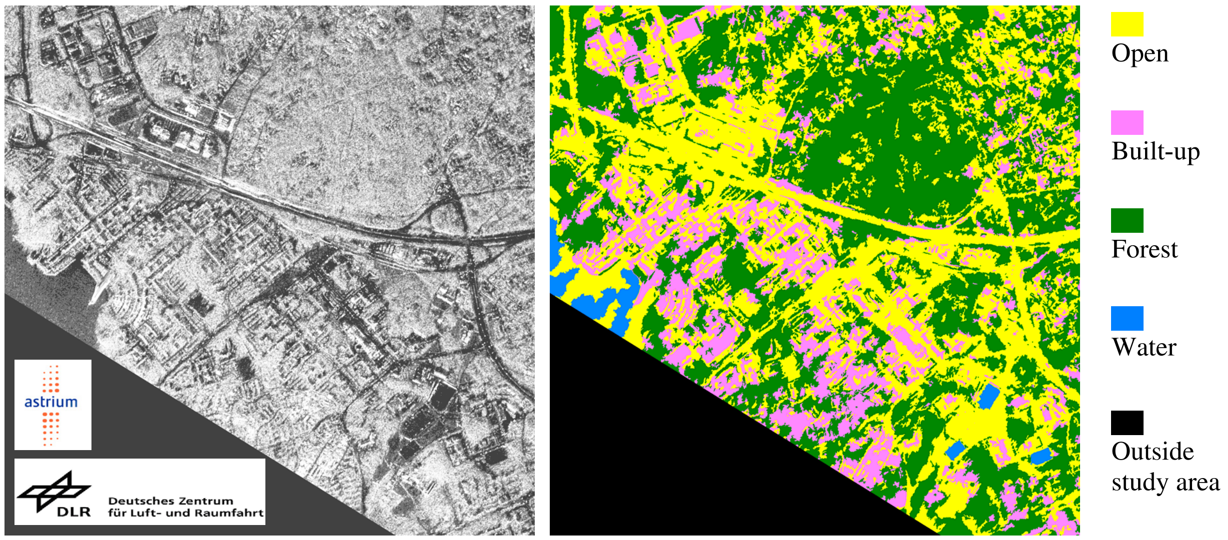

Figure 5. E-SAR image (L band with polarization HH) (left) and land-use classification results (right). The size of the area is 2.5 km × 2.5 km. Image data © DLR and Astrium GmbH.

Table 2. Confusion matrix for the land-use classification results. Corresponding figures for earlier results obtained using manually created classification rules are shown in brackets.

Classif. result	**Reference points**									
	Water	**Forest**	**Open**	**Building**	**Garden/ grass**	**Garden/ trees**	**Car park**	**Multi-storey car park**	**Road**	**All**
Water	**7** (11)	**0** (0)	**5** (0)	**0** (0)	**0** (0)	**0** (0)	**0** (0)	**0** (0)	**0** (0)	12 (11)
Forest	**0** (0)	**104** (104)	**7** (15)	**29** (21)	**1** (0)	**12** (12)	**3** (3)	**2** (1)	**8** (8)	166 (164)
Open	**4** (0)	**3** (3)	**94** (93)	**28** (28)	**8** (8)	**4** (4)	**27** (24)	**4** (2)	**88** (82)	260 (244)
Built-up	**1** (1)	**2** (2)	**5** (3)	**59** (67)	**0** (1)	**5** (5)	**5** (8)	**0** (3)	**4** (10)	81 (100)
All	12	109	111	116	9	21	35	6	100	519

Table 3. Confusion matrix and accuracy estimates for the land-use classification results. The reference points were divided into four classes corresponding to those used for classification. Corresponding figures for earlier results obtained using manually created classification rules are shown in brackets.

Classification result	**Reference points**				
	Water	**Forest** (forest, garden/trees)	**Open** (open, garden/grass, car park, road)	**Built-up** (building, multi-storey car park)	**All**
Water	**7** (11)	**0** (0)	**5** (0)	**0** (0)	12 (11)
Forest	**0** (0)	**116** (116)	**19** (26)	**31** (22)	166 (164)
Open	**4** (0)	**7** (7)	**217** (207)	**32** (30)	260 (244)
Built-up	**1** (1)	**7** (7)	**14** (22)	**59** (70)	81 (100)
All	12	130	255	122	519
Completeness	**58%** (92%)	**89%** (89%)	**85%** (81%)	**48%** (57%)	
Correctness	**58%** (100%)	**70%** (71%)	**83%** (85%)	**73%** (70%)	
Mean accuracy	**58%** (96%)	**78%** (79%)	**84%** (83%)	**58%** (63%)	
Overall accuracy					**77%** (78%)
Kappa					**0.64** (0.66)

4. DISCUSSION

4.1 Classification experiments

The quality of the building detection results was satisfactory and comparable to that obtained using manually created classification rules (see Figure 2, Figure 3 and Table 1). The completeness of the results was 95%, which means that 95% of the map's building pixels were detected as buildings. The correctness was 84%, i.e. 84% of the building pixels in the classification results were labelled as buildings on the map. The mean accuracy, which is a combined measure of completeness and correctness, was 90%. The test areas were separate from the training area and contained a large variety of buildings of different sizes, and different types, materials and colours of roofs. Despite their simplicity, the rules thus proved to work well with the test data. It should also be noted that some of the errors shown by the accuracy measures are due to the different appearance of the buildings on the map and in the remotely sensed data (see Section 3.1.1). For example, buildings with wide eaves are larger in the classification results than on the map, which reduces correctness. The mean accuracy of the classification tree results was 2 percentage units lower than the mean accuracy achieved using the manually created rules. The manually created rules, however, were originally defined for first pulse laser scanner data and it is possible that a slight improvement might be achieved in these results by tuning the rules for last pulse data. The classification process applied with the manually created rules in eCognition was more complex than the classification tree classification. The attributes used in this classification included the GLCM homogeneity calculated from the DSM, the mean value in the red channel of the aerial image and the shape of the segment (standard deviation of length of edges). Additionally, the size of the segment and information on the classes of neighbouring segments were used to correct some small, misclassified segments. The rules were selected after investigating histograms of various attributes in the classes building and tree.

The overall quality of the land-use classification results was also satisfactory and near to that obtained using manually created rules (see Figure 5, Table 2 and Table 3). The overall accuracy of the classification was 77%, which is only 1 percentage unit lower than the overall accuracy achieved earlier using the manually created rules. The results of the kappa analysis for the classification tree classification and the previous classification were 0.64 and 0.66, respectively. In the classification tree classification, however, some problems occurred with the class water. A large part of the water area was classified as open and some open areas were also classified as water. The distinction between water and open was based on the size of the segment (see the classification tree in Figure 4), and this proved to be an inadequate criterion in practice. The number of training segments for water was only 3, which is one likely reason for the problem. The mean accuracy of the class built-up area was also slightly lower in the classification tree classification. When the manually created rules were used, the number of rules was considerably higher and the classification process was more complex. In particular, many rules had to be created for the classes built-up and water to achieve satisfactory results. Neighbourhood information was used in addition to the basic attributes of the segments.

It must be noted that the comparisons between the classification tree results and earlier results give information on the quality of these particular classifications. It is possible that different results would be obtained in other studies. It is also possible that improvements might be achieved with both the classification tree method and manually created rules by further testing and developing the methods. In the case of the classification tree method, several methods and improvements are available for the construction of the tree (see, for example, Breiman et al., 1984; Safavian and Landgrebe, 1991; Pal and Mather, 2003; Lawrence et al., 2004; Zambon et al., 2006). For example, the classification trees that can be created with the Matlab software and that were used in this study are univariate trees, where each split is based on one attribute. The splits are thus perpendicular to the axes of the attribute space, and many such splits may be needed to separate the classes (Breiman et al., 1984). Methods for creating multivariate decision trees have also been developed. These may be better suited to classification problems in which the separation between classes is based on combinations of attributes (see Breiman et al., 1984; Friedl and Brodley, 1997). Several alternative splitting rules and pruning methods also exist, and advanced methods of using training data have been developed (see, for example, Pal and Mather, 2003; Lawrence et al., 2004; Zambon et al., 2006). It is also known that factors such as the total number of training objects and the number of training objects in different classes affect the results (see, for example, Friedl and Brodley, 1997; Pal and Mather, 2003).

The benefits of the classification tree method have been discussed by several authors (see, for example, Breiman et al., 1984; Safavian and Landgrebe, 1991; Hansen et al., 1996; Friedl and Brodley, 1997; Lawrence and Wright, 2001; Thomas et al., 2003; Huang and Lees, 2004; Lawrence et al., 2004). From our point of view, the method proved to be well suited to the type of region-based and rule-based classifications we were interested in. A large number of different attributes can be presented for the method as input data, and it automatically selects the most useful ones. This differs from most automatic classification approaches, where the user must select suitable attributes before classification. The method can also use different attributes for distinguishing different classes. The structure of the classification tree is easy to understand and it gives information on the usefulness of different attributes in the classification task, although caution is needed in the interpretation of this information (see Breiman et al., 1984). The main benefit of the classification tree classification in our study was its high level of automation. Once the scripts had been created and the input data were in the correct format, the tree could be created and the classification carried out in seconds. Compared with the process of manually defining the rules and stages of classification, this achieves considerable time-saving. The manual work, including different analyses and experiments to find a good classification process and rules for a new application, may well take weeks or at least days.

4.2 Classification trees and permanent reference data in new classification applications

Combined with suitable training data, the classification tree method could be a useful tool in the development of new classification applications and in testing the feasibility of new datasets. As discussed above, region-based and rule-based classifications exploiting several different types of attributes can be carried out with a high level of automation using the method. It might thus be possible to achieve satisfactory classification results relatively easily and automatically. The automatically created classification trees could also be used as a starting point for the further manual improvement of the rules. This is possible because the classification rules in the tree are easy to understand and could also be programmed in some programming language or implemented in rule-based classification packages (see, for example, Huang and Jensen, 1997; Lawrence and Wright, 2001). It would also be possible to carry out additional classifications after the classification tree classification to correct some obvious misclassifications. For example, information on the classes of neighbouring segments could be exploited at this stage.

For the purpose of developing new applications on the basis of new datasets, a permanent reference dataset would be useful. The dataset could consist of reference points for which detailed information on the land cover and land use is available and is kept up-to-date. The points should be located inside homogeneous regions, so that transformation from training points to training segments is possible. Different point sets could also be collected, taking into account the characteristics of different spatial resolutions and application areas. Additionally, the reference data could include map data if suitable, up-to-date data are available. When new, remotely sensed datasets become available, the permanent reference data, together with segments and attributes derived from the remotely sensed data, could be used as a basis for the rapid construction of classification trees. Depending on the characteristics of the data and application, the land-cover/land-use information of the reference data could be generalized into the desired classes. If the reference data are representative enough and the characteristics of the remotely sensed data are stable, the classification rules could then be applied to classify new areas. It is possible, however, that variations in intensity and other characteristics between different datasets and areas require changes in the rules. In this case, the rules could be further improved and modified.

5. CONCLUSIONS

The classification tree method is an interesting method that can be used to create a tree-structured classification hierarchy and rules automatically from training data. A large number of different attributes can be presented for the method as input data, from which it automatically selects those most useful for classification. The method is thus well suited to dealing with the diverse datasets and attributes that are typical in the region-based and knowledge-based interpretation of remotely sensed data. Tests with the classification tree method in two applications – building detection using laser scanner and aerial image data and land-use classification using E-SAR data – were carried out using the classification tree tools available in the Statistics Toolbox of the Matlab software. The segments to be classified and attributes for the segments were obtained from the eCognition software. The results were satisfactory and the classification accuracy was near to that obtained in earlier studies using manually created classification rules. The structure of the classification trees in both applications was very simple. The most important benefit of the classification tree method was its high level of automation. The creation of the tree was very fast compared with the process of manually defining rules for classification. A combination of the classification tree method and permanent, up-to-date reference data could be a useful tool in the development of new classification applications and in testing the feasibility of new remotely sensed datasets. Together with segments and attributes derived from remotely sensed data, it could be used for the rapid construction of classification trees, which could then be directly used in classification or as a starting point for further development of the rules. With this method, it might be possible to achieve satisfactory classification results relatively easily and automatically.

6. REFERENCES

Breiman, L., Friedman, J. H., Olshen, R. A., Stone, C. J., 1984. Classification and regression trees, Wadsworth, Inc., Belmont, California, 358 p. (The Wadsworth Statistics/Probability Series.)

Chan, J. C.-W., Chan, K.-P., Yeh, A. G.-O., 2001. Detecting the nature of change in an urban environment: a comparison of machine learning algorithms. Photogrammetric Engineering & Remote Sensing, Vol. 67, No. 2, pp. 213-225.

Congalton, R. G., Green, K., 1999. Assessing the Accuracy of Remotely Sensed Data: Principles and Practices, Lewis Publishers, Boca Raton, 137 p.

Definiens, 2006. http://www.definiens.com/, (accessed 29 Nov. 2006).

Definiens Imaging, 2003. eCognition, Object Oriented Image Analysis, User Guide 3, Definiens Imaging GmbH, München, Germany.

Ducic, V., Hollaus, M., Ullrich, A., Wagner, W., Melzer, T., 2006. 3D vegetation mapping and classification using full-waveform laser scanning, Proceedings, International Workshop, 3D Remote Sensing in Forestry, Vienna, 14th –15th Feb. 2006, pp. 211-217, http://www.rali.boku.ac.at/fileadmin/_/H85/H857/workshops/3drsforestry/Proceedings_3D_Remote_Sensing_2006.pdf, (accessed 29 Nov. 2006).

Friedl, M. A., Brodley, C. E., 1997. Decision tree classification of land cover from remotely sensed data. Remote Sensing of Environment, Vol. 61, pp. 399-409.

Hansen, M., Dubayah, R., Defries, R., 1996. Classification trees: an alternative to traditional land cover classifiers. International Journal of Remote Sensing, Vol. 17, No. 5, pp. 1075-1081.

Haralick, R. M., Shanmugam, K., Dinstein, I., 1973. Textural features for image classification. IEEE Transactions on Systems, Man, and Cybernetics, Vol. SMC-3, No. 6, pp. 610-621.

Helldén, U., 1980. A test of Landsat-2 imagery and digital data for thematic mapping, illustrated by an environmental study in northern Kenya, Lunds Universitets Naturgeografiska Institution, Rapporter och Notiser 47, Lund, Sweden, 63 p.

Hodgson, M. E., Jensen, J. R., Tullis, J. A., Riordan, K. D., Archer, C. M., 2003. Synergistic use of lidar and color aerial photography for mapping urban parcel imperviousness. Photogrammetric Engineering & Remote Sensing, Vol. 69, No. 9, pp. 973-980.

Huang, X., Jensen, J. R., 1997. A machine-learning approach to automated knowledge-base building for remote sensing image analysis with GIS data. Photogrammetric Engineering & Remote Sensing, Vol. 63, No. 10, pp. 1185-1194.

Huang, Z., Lees, B. G., 2004. Combining non-parametric models for multisource predictive forest mapping. Photogrammetric Engineering & Remote Sensing, Vol. 70, No. 4, pp. 415-425.

Im, J., Jensen, J. R., 2005. A change detection model based on neighborhood correlation image analysis and decision tree classification. Remote Sensing of Environment, Vol. 99, pp. 326-340.

Jensen, J. R., 2005. Introductory Digital Image Processing, A Remote Sensing Perspective, Third edition, Pearson Prentice Hall, Upper Saddle River, 526 p. (Prentice Hall Series in Geographic Information Science.)

Jung, F., 2004. Detecting building changes from multitemporal aerial stereopairs. ISPRS Journal of Photogrammetry & Remote Sensing, Vol. 58, pp. 187-201.

Lawrence, R. L., Wright, A., 2001. Rule-based classification systems using classification and regression tree (CART) analysis. Photogrammetric Engineering & Remote Sensing, Vol. 67, No. 10, pp. 1137-1142.

Lawrence, R., Bunn, A., Powell, S., Zambon, M., 2004. Classification of remotely sensed imagery using stochastic gradient boosting as a refinement of classification tree analysis. Remote Sensing of Environment, Vol. 90, pp. 331-336.

Li, D., Di, K., Li, D., 2000. Land use classification of remote sensing image with GIS data based on spatial data mining techniques, International Archives of Photogrammetry and Remote Sensing, Amsterdam, The Netherlands, Vol. XXXIII, Part B3, pp. 238-245.

The MathWorks, 2006. http://www.mathworks.com/, (accessed 29 Nov. 2006).

The MathWorks, 2003. Online documentation for Statistics Toolbox, Version 4.0.

Matikainen, L., Karjalainen, M., Kaartinen, H., Hyyppä, J., 2004a. Rule-based interpretation of high-resolution SAR images for map updating. The Photogrammetric Journal of Finland, Vol. 19, No. 1, pp. 47-62.

Matikainen, L., Hyyppä, J., Kaartinen, H., 2004b. Automatic detection of changes from laser scanner and aerial image data for updating building maps, The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Istanbul, Turkey, Vol. XXXV, Part B2, pp. 434-439.

McCauley, S., Goetz, S. J., 2004. Mapping residential density patterns using multi-temporal Landsat data and a decision-tree classifier. International Journal of Remote Sensing, Vol. 25, No. 6, pp. 1077-1094.

Pal, M., Mather, P. M., 2003. An assessment of the effectiveness of decision tree methods for land cover classification. Remote Sensing of Environment, Vol. 86, pp. 554-565.

Quinlan, J. R., 1993. C4.5: Programs for Machine Learning, Morgan Kaufmann, San Mateo, California, 302 p.

Richards, J. A., Jia, X., 1999. Remote Sensing Digital Image Analysis, An Introduction, Third, revised and enlarged edition, Springer-Verlag, Berlin, 363 p.

Rogan, J., Miller, J., Stow, D., Franklin, J., Levien, L., Fischer, C., 2003. Land-cover change monitoring with classification trees using Landsat TM and ancillary data. Photogrammetric Engineering & Remote Sensing, Vol. 69, No. 7, pp. 793-804.

RuleQuest Research, 2006. RuleQuest Research, data mining tools, http://www.rulequest.com/, (accessed 29 Nov. 2006).

Safavian, S. R., Landgrebe, D., 1991. A survey of decision tree classifier methodology. IEEE Transactions on Systems, Man, and Cybernetics, Vol. 21, No. 3, pp. 660-674.

Simard, M., De Grandi, G., Saatchi, S., Mayaux, P., 2002. Mapping tropical coastal vegetation using JERS-1 and ERS-1 radar data with a decision tree classifier. International Journal of Remote Sensing, Vol. 23, No. 7, pp. 1461-1474.

Soininen, A., 2005. TerraScan User's Guide (03.10.2005), Terrasolid, 169 p.

Terrasolid, 2006. http://www.terrasolid.fi/, (accessed 29 Nov. 2006).

Thomas, N., Hendrix, C., Congalton, R. G., 2003. A comparison of urban mapping methods using high-resolution digital imagery. Photogrammetric Engineering & Remote Sensing, Vol. 69, No. 9, pp. 963-972.

Townsend, P. A., 2001. Mapping seasonal flooding in forested wetlands using multi-temporal Radarsat SAR. Photogrammetric Engineering & Remote Sensing, Vol. 67, No. 7, pp. 857-864.

Tso, B., Mather, P. M., 2001. Classification Methods for Remotely Sensed Data, Taylor & Francis, London and New York, 332 p.

Tullis, J. A., Jensen, J. R., 2003. Expert system house detection in high spatial resolution imagery using size, shape, and context. Geocarto International, Vol. 18, No. 1, pp. 5-15.

Zambon, M., Lawrence, R., Bunn, A., Powell, S., 2006. Effect of alternative splitting rules on image processing using classification tree analysis. Photogrammetric Engineering & Remote Sensing, Vol. 72, No. 1, pp. 25-30.

ACKNOWLEDGEMENTS

The author wishes to thank Professor Juha Hyyppä for his valuable comments on the study and the manuscript, and Harri Kaartinen and Mika Karjalainen for their contribution in creating the datasets. Harri Kaartinen originally collected and preprocessed a major part of the reference data and orthorectified the aerial imagery. Mika Karjalainen rectified the E-SAR images to the local coordinate system and helped in the preprocessing of the reference data. The author also wishes to thank TopoSys GmbH for laser scanner data, DLR and Astrium GmbH for E-SAR data (ProSmart II), FM-Kartta Oy for aerial image and map data, the city of Espoo for map data, and the National Land Survey of Finland for map data (used, for example, when collecting the reference points). The financial support from the Academy of Finland (project 'Novel map updating with remote sensing imagery') is gratefully acknowledged.