

Improving Pixel-based VHR Land-cover Classifications of Urban Areas with Post-classification Techniques

Tim Van de Voorde, William De Genst, and Frank Canters

Abstract

In this paper, three post-classification techniques are proposed to improve the information content, thematic accuracy, and spatial structure of pixel-based classifications of complex urban areas. A shadow-removal technique based on a neural network that was trained using the output of a soft classification is proposed to assign shadow pixels to meaningful land-cover classes. Knowledge-based rules are suggested to correct wrongly classified pixels and to improve the overall accuracy of the land-cover map. Finally, a region-based filter is applied to reduce high-frequency structural clutter. The three techniques were successfully applied to a pixel-based classification of a QuickBird image covering the city of Ghent, Belgium, improving the kappa index-of-agreement from 0.82 to 0.86 and transforming the shadow pixels into meaningful land-cover information.

Introduction

Nearly half of the Earth's population dwells in cities. In industrialized regions like Europe and the United States, the level of urbanization reaches nearly 80 percent (UNCHS, 2001). The continuing growth of cities since the late eighteenth century stretches the environment in and around urban areas to its limits. Local decision makers need effective and easy-to-use urban management tools to deal with problems such as disorganized growth, low quality of life, environmental degradation, and a deteriorating urban infrastructure. Lack of reliable and sufficiently detailed data is, however, a major obstacle for the problem analysis, planning, and monitoring phases of a sustainable urban management policy (E.C. Environment DG, 2004).

Very High Resolution (VHR) satellite images are potentially an attractive source of information for local and regional decision makers. Imagery with a resolution of 1 m and less not only allows a detailed mapping of land-cover, but the acquisition of images at multiple dates also facilitates monitoring dynamic areas where rapid changes in land-cover occur. This may substantially improve our understanding of the complex processes of change that take place within cities and at the urban fringe (e.g., the European Commission's MOLAND project: <http://moland.jrc.it>). Typical examples of the application of VHR data in an urban context are mapping urban green (Pillmann and Kellner, 2002; Ries *et al.*, 2002) and assessing impervious surface cover (Arnold

et al., 2000; Flanagan and Civco, 2001; Yang *et al.*, 2003). A comprehensive survey held among 19 large European cities showed that 13 of them already use some form of remotely sensed data, three of which use data from satellite platforms (Schmied and Pillmann, 2003).

Thematic mapping from remotely sensed data is typically based on image classification. Due to the spectral heterogeneity of urban surface types and the spatial complexity of urban scenes, VHR image classification of urban areas often results in land-cover maps with low thematic accuracy. Major problems that occur in classifying urban areas are: shadows hiding the underlying land-cover, spectral mixing due to the presence of transition zones between two classes or due to multiple class occurrence within single pixels, within-class heterogeneity due to varying lighting conditions or nearby solar reflection (e.g., on roof windows) or due to differences in the physical condition (e.g., aging of materials or dust cover) or chemical composition of materials, and finally (typical for pixel-based classifications) the salt-and-pepper effect, also referred to as structural clutter, which makes the resulting map look too complex and noisy. Specifically for pixel-based classifications, Jensen (2000) also points out that single-pixel classifiers have difficulties mapping from VHR data because texture information is, compared with lower resolution imagery, no longer present within a single pixel, but rather exists in the relationships between pixels.

Two main approaches can be taken to deal with the strong spectral-spatial variability in urban VHR imagery: image segmentation or post-classification processing. The former is applied before the classification process, and the latter after an image is classified.

For the segmentation approach, the image is divided into regions of similar pixels prior to classification. These so-called image segments do not necessarily have any cartographic meaning and can be considered as image primitives. Once they are created, they can be attributed to a land-cover class by any type of classifier. Many techniques of image segmentation have been developed (Pal and Pal, 1993). The most common methods to segment a full image are: global thresholding (a survey of these techniques is given by Sahoo *et al.*, 1988), region growing algorithms, watershed segmentation (Wegner *et al.*, 1997), and texture segmentation

Photogrammetric Engineering & Remote Sensing
Vol. 73, No. 9, September 2007, pp. 1017–1027.

0099-1112/07/7309-1017/\$3.00/0
© 2007 American Society for Photogrammetry
and Remote Sensing

Centre for Cartography and GIS, Geography Department
Vrije Universiteit Brussel, Pleinlaan 2, B 1050 Brussel
(tvdvoord@vub.ac.be).

algorithms. The latter can be based on spatial frequencies (Hofmann *et al.*, 1998), Markov Random Field models (Panjwani and Healy, 1995), co-occurrence matrices (Haralick and Shapiro, 1985), wavelet coefficients (Salari and Ling, 1995), or fractal indices (Chaudhuri and Sarkar, 1995). Recently, multi-resolution image segmentation (Baatz and Schäpe, 2000) has received quite some attention as it is a part of the object oriented image classification approach embedded in the commercial software Definions Professional®. Classifying image segments will produce a less cluttered image than a per-pixel classifier, but also leads to information loss due to generalization. One of the main difficulties in applying segment-based classification techniques is to obtain a set of segments that can be unambiguously linked to real-world objects without over-generalizing the structure of the scene.

As an alternative to segment-based classification, one may also apply a pixel-based classification approach and attempt to improve the structure of the obtained classification by applying dedicated post-classification techniques. In the post-classification phase of a satellite image classification, the classified image is manipulated to alter the class-labeling initially attained by the classifier. This usually entails the application of context-based rules or filters to the result of a pixel-based classification.

Contextual information, i.e., implicit information derived from the relationship between map elements, can be used for post-classification in different ways. Barr and Barnsley (1998), for instance, have proposed a reflexive mapping procedure that operates on individual regions of a classification, i.e., groups of adjacent pixels of the same land-cover class, to remove high-frequency structural clutter in a classified image. In their approach, all regions below a pre-specified area threshold are merged with the smallest neighboring region that exceeds the specified threshold. Simpler, though less efficient techniques for removing regions that are considered too small from a classification are: applying a standard majority filter (Gurney and Townshend, 1983) or a more sophisticated spatial reclassification technique (Barnsley and Barr, 1996; Gong and Howarth, 1992; Wharton, 1982) within a moving window of fixed size. The use of kernel-based approaches, however, as opposed to a region-based approach, has a number of disadvantages, including the difficulty of selecting an optimal kernel size and the fact that the kernel is an artificial construct that does not refer to the spatial units that occur in the land-cover scene.

Another way contextual knowledge can be used is by applying class-dependent rules that assign incorrectly labeled pixel regions to the correct land-cover class, based on context and geometry of the region, e.g., an image interpreter knows from their familiarity with the area that no cropland is present in the central business district. Small groups of cropland pixels surrounded by trees must therefore rather be part of an urban park and will be assigned to this land-cover type by applying a knowledge-based rule specifically designed to deal with these small erroneously labeled cropland regions. Context may also be introduced in the post-classification phase in the form of ancillary data sources that allow us to refine the outcome of the initial classification. Thomas *et al.* (2003), for instance, use land-use data and road centerlines in a raster-based spatial model to improve the results of a spectral classification. Elevation data from available digital surface models can also be used to distinguish paved surfaces from building roofs, even if both classes have similar spectral characteristics (Van de Voorde *et al.*, 2004).

Finally, a soft classification algorithm also provides information that can be useful for post-classification in the form of fuzzy membership values, class probabilities, or

end-node activations, depending on the type of soft classifier used. This information can be incorporated into knowledge-based rules or be used directly as input to another classifier to reclassify parts of a prior classification, as will be demonstrated for shadow removal in this research.

Both the post-classification and the segmentation approach appear to improve the quality of the end-product. Thomas *et al.* (2003), for instance, found that applying a knowledge-driven spatial model on the output of a pixel-by-pixel classification combined with ancillary data layers resulted in a higher mapping accuracy. Their raster-based model also performed better than an image segmentation approach, although the latter allowed for a more automated integration of spectral and contextual information and thus required less analyst input (Thomas *et al.*, 2003).

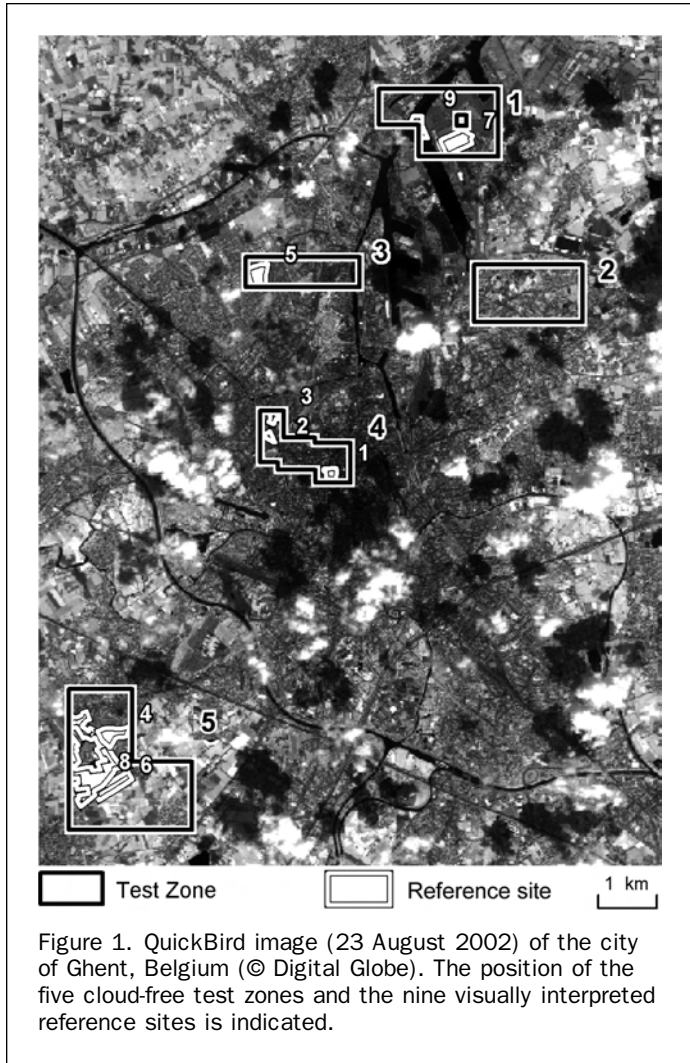
In this paper, we propose three straightforward post-classification techniques that can be applied to any soft pixel-by-pixel land-cover classification to improve its thematic accuracy and spatial structure: (a) post-classification shadow removal to improve the information content of the classification without conceding too much on accuracy; (b) rule-based spatial modeling to correct the classification errors that result from the pixel-by-pixel classification approach; and (c) structural filtering to reduce the complexity of the obtained land-cover map, and to improve its visual appearance as well as its thematic accuracy. The three post-classification enhancement techniques were tested on a pixel-based neural network classification of several test zones situated within a QuickBird image covering a spatially complex urban area.

Study Area and Data

Ghent, Belgium is the second most important city of the Flemish region. It covers an area of 158.18 square kilometers with a registered population of about 229,000 (2003). This corresponds to about 15 percent of the region's total urban population. About a third of the municipal area consists of built-up parcels. Between 1982 and 1999, the built-up area expanded with 22 percent, which corresponds to an average annual growth of 34 hectares (Baelus *et al.*, 2003).

The greater city has diverse land-cover and land-use types. The port area in the north contains mainly industrial buildings and transport infrastructure. The old city center consists of high density urban housing and is interwoven with canals. It is surrounded by low-density residential areas, agricultural land, and pasture.

Our study area is made up of five test zones with different types of land-use (Figure 1): the port area (zone 1), two zones with low density residential housing (zones 2 and 3), the old city center (zone 4), and a rural area (zone 5). Cloud-free image data were extracted for each zone from a QuickBird bundled image product acquired on 23 August 2002. A reference digital surface model (DSM) was created for each zone using a set of aerial photographs at a 1:12 000 scale. The DSMs were used in combination with GPS-measured ground control points to fully 3D ortho-rectify each image subset. The resulting ortho-rectified images have a resolution of 0.63 m in the panchromatic channel and 2.52 m in the multispectral bands. For the purpose of classification, the multispectral information was resampled to the 0.63 m resolution of the PAN-image. The multispectral bands were also fused with the PAN band (0.63 m resolution), using the Intensity Normalized Ratio band per band (INRBPB) algorithm proposed by Cornet (2003), but only to facilitate the visual sampling of training and validation data, not for defining the input bands for the classification.



Methods

Classification System

The classification system we adopted consists of seven land-cover classes: orange and red surfaces, grey surfaces, bare soil, water, grass, crops, shrub and trees. The class “orange and red surfaces” represents building materials of these colors such as red-tiled squares or parking spaces and roof tiles made of red clay. The “grey surfaces” class is an aggregation of several urban materials with a grey appearance such as concrete, asphalt, slate, etc. Preliminary tests taught us that these surface types could not be spectrally distinguished with QuickBird’s limited number of spectral bands. For this reason, we defined the super-class “grey surfaces.”

Shadow was also included as a separate class because shadows have a specific spectral response. As explained before, the problem of shadow removal is one of the issues we decided to deal with in the post-classification stage.

Training and Validation Data

Training data were obtained by selecting 20 training sites for each spectral class within each of the five test zones. Aerial photographs (1:12 000), the ortho-rectified satellite images, and the result of the image fusion were useful tools for this purpose. Each training site was field-visited to decide whether it was suitable or not, and changes were made if

necessary. Six actual training pixels belonging to the class considered were then selected within each of the verified training sites.

To reduce within-class spectral variability and reduce the risk of confusion with other classes, the super-class “grey surfaces” was divided into three spectral subclasses: light, medium, and dark grey. The purpose of splitting the “grey surfaces” class into three subclasses was solely for the benefit of the classification process itself. After the classification and before the validation, pixels assigned to one of the three grey subclasses were re-grouped into “grey surfaces.”

To be able to sample training data for each of the three grey classes (light, medium, and dark), we fell back on a grey-surface mask. This mask was developed by thresholding the red band and an NDVI image to exclude, respectively, red and vegetated surfaces. An unsupervised classifier was then applied on the image pixels within this mask to divide them into three subclasses of grey: light, medium, and dark. Training pixels for each subclass were defined by overlaying the corresponding part of the mask with a set of “grey surface” training sites that had been visually identified in the image and verified in the field, and then selecting training pixels within the overlapping area.

For validation, an exhaustive visual interpretation of nine small reference sites located within the five test zones of the study area (see Figure 1) was accomplished. Reference site selection was based on a proper representation of the different urban morphologies that characterize the study area. The edges of the polygons of the visual interpretation were eroded with a buffer operation to avoid edge effects during sampling, because while making the visual interpretation, it was often impossible to draw a strict line between two classes. The buffer was kept very small: only two pixels on each side of the polygons. A stratified random sample of validation points was then selected from this interpretation. We estimated the number of validation points to be sampled for each class from the class proportions obtained by applying a simple maximum likelihood classification on all the test zones. This method was chosen to ensure a proper representation of all classes in the validation set, according to their relative prevalence within the study area. Validation pixels for the shadow class were obtained by visually inspecting the selected validation pixels of all other classes on the fused satellite imagery. Validation pixels that fell inside shadowed areas were re-labeled as shadow. For validating the effects of the shadow removal on the classification, all shadow pixels in the validation set were again re-assigned to the classes to which these pixels actually belong, as determined from the exhaustive visual interpretation of the reference sites.

Table 1 lists the total number of training and validation pixels per class.

TABLE 1. NUMBER OF TRAINING AND VALIDATION PIXELS FOR EACH LAND-COVER CLASS

Class	Training	Validation
Grey surfaces	594 (light) 642 (medium) 594 (dark)	508
Orange and red surfaces	582	110
Bare soil	119	245
Water	279	120
Grass	698	347
Crops	211	244
Shrub and trees	639	289
Shadows	602	188

Neural Network Classification

To define and test our post-classification strategies, we developed a reference land-cover classification with an Artificial Neural Network (ANN). The network was created with Neuralware's NeuralWorks Predict® software, using the training data to build it. The accuracy of the resulting land-cover map was assessed with the validation data.

Several networks were trained and tested with two degrees of freedom: the selected input bands and the ratio between the amount of test and training cases. The combinations of input bands were grouped into three scenarios: only the multispectral bands, the multispectral bands with the PAN band, and the multispectral bands with the PAN and the local variance for a 3×3 moving window calculated on the PAN. Transformations of these input bands and a selection according to their relative contribution to the overall information content were accomplished with NeuralWorks Predict®. The transformed input variables that were retained in each scenario to actually perform the classification were chosen from a set of five mathematical transformations per original input band using a genetic-based variable selection algorithm embedded in the software.

The ratio of the amount of training samples used for learning (i.e., adjusting the network weights) to the amount of training samples used to determine when to halt the learning phase (the test data) proved to have a clear impact on the performance of the network, which was estimated on the independent validation data. The best result in terms of overall accuracy was obtained with the PAN, the multispectral channels and the local variance as input bands, and with a train/test ratio of 70/30 (PCC = 0.83, see Table 3). The architecture of the network that was trained with this input consists of 10 input variables, 16 hidden nodes, and 10 output nodes (the land-cover classes). The 10 input variables are mathematical transformations carried out on the input bands to scale them between 0 and 1. They represent the optimal set that was chosen by the genetic algorithm from a wide range of transformations (linear, n^{th} power, hyperbolic tangent, . . .) calculated on each input variable. The neural network architecture was created with the cascade-correlation learning algorithm, available in Predict®. This type of learning algorithm learns very quickly and the network determines its own size and topology (Fahlman and Lebiere, 1990).

This best classification result (further referred to as "original classification" or "initial classification") was used as a starting point for developing and testing the post-classification methods that are proposed in this paper.

The output of the neural network is a soft classification that provides us with an activation level for each land-cover class. The activation levels can be interpreted as indicators of the likelihood that a pixel belongs to a certain class (Foody, 1997) and were used in this study for post-classification shadow removal (see below).

Post-classification Shadow Removal

Especially in urban areas, shadows cast by trees, houses, and other buildings cause major difficulties to obtain useful land-cover maps. Large areas with useful information are lost because they are hidden by shadows. When they are not taken into account in the initial classification phase, shadowed areas are often classified as water or dark-grey areas. The amount of shadows that is present in an image of an urban area is related to the object height, the sensor collection azimuth, solar elevation angle, and solar azimuth at the time of acquisition. Specifying the sensor collection azimuth when placing the image order might be a first step to reduce the shadow cover. The sensor azimuth should be as close as possible to the solar azimuth (Sugumaran *et al.*, 2002).

However, even if the aforementioned parameters can be specified, shadows will still remain a major problem, particularly in urban areas. Techniques are therefore required to reduce the amount of shadows present in an image. In a recent paper, Dare (2005) presents a method for detecting and removing shadows from VHR images of urban areas. His shadow detection method is based on thresholding a density sliced, single band VHR panchromatic satellite image. Confusion between shadowed and non-shadowed areas such as water bodies is then filtered out by comparing the spectral variance of the regions identified as shadow. Because secondary illumination in shadowed regions makes many features visible, the spectral variance in those regions proved to be higher than within water bodies. After the shadows were identified, Dare (2005) reduced their severity with a radiometric enhancement for which the parameters were determined by comparing the histograms of the shadowed and non-shadowed regions.

To solve the shadow problem in this study, we applied a different technique: as was mentioned earlier, we dealt with shadow by considering it as an extra class in the classification process. After applying the ANN-classifier, we extracted the pixels identified as shadow from the initial classification. This allowed us to get a good idea of the approximate magnitude and location of the shadows within the five test zones (Table 2). We then re-assigned each pixel labeled as shadow to the land-cover class we assumed to be present at the pixel's location by applying two approaches that both use the class membership information produced by the soft classifier, i.e., the activations of the output nodes of the neural network. In a first attempt, shadow pixels were simply assigned to their second most likely class, i.e., the class corresponding to the second highest activated output node. In a second attempt, we trained a separate neural network to re-classify all shadow pixels into meaningful land-cover classes, using the per-class activation levels that are obtained for each pixel assigned to the shadow class by the initial network as an input for the second network. Training data for this second network were obtained by overlaying the original classification with the small visually interpreted reference sites. From this overlay, we could determine the actual class each shadow pixel belonged to according to the visual interpretation. The second network thus presents a mapping between the activation levels of the shadow pixels in the original classification and the target land-cover classes to which these shadow pixels belong.

Rule-based Classification Enhancement

Pixel regions (i.e., groups of adjacent pixels that have been assigned to the same class) may be misclassified due to spectral confusion between various urban surface types. Starting from the major problems we observed in our classification, we designed a set of knowledge-based rules that allowed us to solve many of the problems related to spectral confusion. Basically, what each rule does, is assign wrongly labeled regions to another class, depending on their size and on their spatial relationship with other regions

TABLE 2. AMOUNT OF SHADOW IN THE INITIAL CLASSIFICATION FOR EACH OF THE FIVE TEST ZONES

Test Zone	Type	% Shadow
1	Industrial	2.73%
2	Low-density residential	8.96%
3	Low-density residential	10.40%
4	High-density residential	13.58%
5	Rural	6.09%

(e.g., adjacency, inclusion). Some rules also make use of information on classification uncertainty produced by the neural network classifier, i.e., the identity of the second most likely class for the pixels that are part of a region. The actual definition of the rules is based on expert knowledge about the mislabeling of regions, which is gathered by observing systematic error patterns in the classified imagery. While defining the rules, special care was taken to ensure that the rule set could be successfully applied under various morphological conditions. In our study, this meant that the final set of rules had to be applicable to each of the five test zones, which were selected based on clear differences in urban morphology.

Structural Filtering

Structural image clutter, also referred to as the “salt-and-pepper” effect, is a common by-product of pixel-based classifications and is particularly conspicuous in high-resolution scene models, where image pixels are smaller than the elements in the scene (Strahler *et al.*, 1986). Land-cover classifications obtained from very-high-resolution satellite data are a good example. Because of the spectral variation observed within the boundaries of meaningful objects constituting the landscape, an individual pixel or a small group of adjacent pixels might be assigned to a different class than its neighbors, even though all pixels are part of the same object. This makes resulting land-cover maps look noisy or cluttered. In spatially complex environments such as urban areas, the within-class spectral heterogeneity of some surface types exacerbates this problem. Most of the unwanted clutter can be dealt with by applying rules based on contextual information. Barr and Barnsley (2000) make two assumptions about clutter areas: they are typically small in size, and they are likely to be adjacent to at least one land-cover parcel that does not represent structural clutter. Thus, if the clutter region is re-assigned to the land-cover class of that parcel, it will result in a more meaningful spatial pattern in that part of the image.

Following the work of Barr and Barnsley (2000), we applied a structural filter that deals with the most severely cluttered regions of our classification. First, we determine regions of adjacent same-class pixels in the classification. Then, the pixels of each region falling below a certain area threshold are re-assigned to the largest neighboring region.

In this study, a threshold value of 16 pixels was applied. We used this filter in combination with the rule-based procedure outlined above.

Results and Discussion

Classification Result

The quantitative accuracy of the original pixel-based supervised ANN classification is presented in Table 3. It was assessed, for all test zones together, with the independent validation data. The kappa index-of-agreement (Cohen, 1960; Rosenfield and Davis, 1979; Hudson, 1987) was calculated as a measure of the overall classification accuracy. It is shown at the bottom of the Table together with the percentage of correctly classified pixels (PCC). For the original classification, the kappa index is 0.82, without taking validation pixels of the shadow class and pixels labeled as shadow into account. When we did include shadow pixels in the validation process, we attained a kappa index of 0.80. The producer's accuracy (PA) of the shadow pixels is rather low (66 percent) because quite some shadow pixels from the validation sample were assigned to land-cover classes such as water, grey surfaces, and shrub and trees. We can also learn from the confusion matrix (Table 3) that shifts from one class to another caused by spectral confusion occur especially among the vegetation classes (shrub and trees, grass, crops), from grey surfaces to bare soil and from red surfaces to grey surfaces and bare soil.

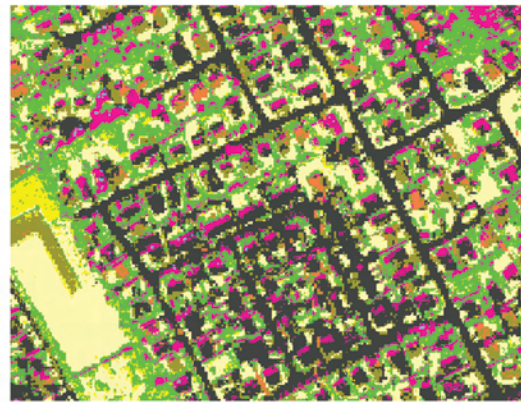
Despite the relatively high classification accuracy, a visual inspection of the resulting land-cover map points us at some problems. Plate 1b is an extract of the classification taken from test zone 3. The major problems can be summed up as follows: the classified image is excessively noisy (salt-and-pepper effect); parts of the image lack information due to the presence of shadow; some pixel regions are wrongly labeled due to spectral confusion between two or more surface types; the transition area between some adjacent land-cover patches consists of mixed pixels, which are misclassified, and finally, sun-glint effects alter the spectral response of certain pixels causing them to be wrongly classified as well. In Plate 1b, the small groups of bright yellow pixels in the central and northwestern part of the image are wrongly classified as crops because they form the

TABLE 3. ERROR MATRIX FOR THE ORIGINAL ANN-CLASSIFICATION, WITH AND WITHOUT VALIDATION OF SHADOW PIXELS

		Reference Data								Total	U.A. no Shadow	U.A. with Shadow
		Grey Surfaces	Red Surfaces	Bare Soil	Water	Grass	Crops	Shrub and Trees	Shadow			
Original ANN-classification	Grey surfaces	420	17	0	2	3	0	1	20	463	95%	91%
	Red surfaces	5	70	2	0	1	0	0	5	83	90%	84%
	Bare soil	68	16	239	0	8	0	1	2	334	72%	72%
	Water	1	0	0	118	0	0	0	14	133	99%	89%
	Grass	3	2	1	0	280	17	13	2	318	89%	88%
	Crops	0	1	2	0	37	205	27	0	272	75%	75%
	Shrub and trees	2	1	1	0	17	22	246	20	309	85%	80%
	Shadow	9	3	0	0	1	0	1	125	139		90%
	Total	508	110	245	120	347	244	289	188	2051		
	P.A. no shadow	84%	65%	98%	98%	81%	84%	85%			PCC: 85% kappa: 0.82	
P.A. with shadow	83%	64%	98%	98%	81%	84%	85%	66%		PCC: 83% kappa: 0.80		



(a)



(b)



(c)



(d)


	Grey surfaces		Grass
	Orange and red surfaces		Crops
	Bare soil		Shrub and trees
	Water		Shadow

Plate 1. (a) Extract of the PAN image © Digital Globe, (b) the original classification, (c) the result after shadow removal, and (d) the result after rule-based classification enhancement and structural filtering for part of test zone 3 (low density residential area).

transition zone between grass and trees. The same problem occurs for the brown bare soil pixels lying between roads and grass. The purple patches are shadows cast by houses or trees which hide the underlying land-cover and consequently reduce the information content. Table 2 illustrates this effect by showing for each test zone the percentages of shadow pixels in the classification. Especially in zone 4, the

shadows cause a significant loss of information because this zone lies in the city center, a dense urban area with relatively high buildings.

Shadow Removal

We attempted to solve the shadow problem with the post-classification shadow removal technique described above.

Within the area covered by the nine small reference sites, a random set of pixels that were classified as shadow was drawn from the original classification. The real class label obtained from the exhaustive visual interpretation of the reference sites as well as the corresponding end-node activations of the ANN that was applied to obtain the original classification were extracted for each of these pixels, in order to obtain training data to build a new network for the reclassification of shadows. A randomly selected validation set also consisting of pixels that were classified as shadow, but independent from the data used for the training, was used to assess the accuracy of the re-classification procedure.

The results of this validation are shown in Tables 4 and 5. We could not collect sufficient training and validation samples for shadow falling on "crops." No error measures could therefore be calculated for this class. When we simply allocated the shadow pixels to the class corresponding to the second highest activated end-node in the original classification, the accuracy calculated on the independent validation set was only 49 percent (Table 4). This is not surprising, because for all pixels classified as shadow, the highest activated end-nodes are mostly those corresponding to either "dark grey," "water," or "shrub and trees." Re-assigning shadow pixels to their second most likely class mostly results in pixels being attributed to one

of these three classes, while in reality shadow pixels may also belong to other classes. With the ANN-approach (Table 5), approximately 63 percent of the validation samples were correctly re-assigned to a new class. When an ANN is used to re-label the shadow pixels, the pattern of the activation levels of all classes is taken into account, which evidently includes implicit information on the actual land-cover beneath the shadow. This results in a substantial increase of producer accuracies for classes that are less spectrally confused with shadow, like "red surfaces," "bare soil," and "grass." For "water" and "shrub and trees" the producer accuracy decreases, yet the user accuracy is improved.

The network obtained for re-labeling shadow pixels was then applied to all pixels that had been assigned to the shadow class in the original classification. Plate 1c shows the classification result for the extract of test zone 3, after pixels that were classified as shadow have been re-assigned. Table 6 presents the confusion matrix for the classification without shadows. The kappa index drops from 0.82 to 0.79, yet one should keep in mind that the information content of the end-product is substantially increased. While the kappa of 0.82 only refers to the part of the area that is not assigned to the shadow class by the original classifier, the kappa of 0.79 refers to the entire area covered by the classification. It is obvious that the re-classification of pixels originally

TABLE 4. ERROR MATRIX FOR SHADOW RE-CLASSIFICATION WITH SECOND PROBABILITIES

		Reference Data							Total	U.A.
		Grey Surfaces	Red Surfaces	Bare Soil	Water	Grass	Crops	Shrub and Trees		
Second most likely identity of shadow pixels	Grey surfaces	393	394	0	128	118	0	70	1103	36%
	Red surfaces	3	72	0	0	3	0	0	78	92%
	Bare soil	1	52	2	3	15	0	11	84	2%
	Water	500	188	67	739	172	0	23	1689	44%
	Grass	0	0	0	0	0	0	0	0	0%
	Crops	1	1	0	10	11	0	10	33	N/A
	Shrub and trees	26	49	6	15	165	0	752	1013	74%
	Total	924	756	75	895	484	0	866	4000	
P.A.	43%	10%	3%	83%	0%	N/A	87%		PCC: 49% kappa: 0.35	

TABLE 5. ERROR MATRIX FOR ANN SHADOW RE-CLASSIFICATION

		Reference Data							Total	U.A.
		Grey Surfaces	Red Surfaces	Bare Soil	Water	Grass	Crops	Shrub and Trees		
Identity of shadow pixels after re-classification	Grey surfaces	428	129	0	76	44	0	21	698	61%
	Red surfaces	82	485	0	28	29	0	36	660	73%
	Bare soil	106	28	54	110	103	0	8	409	13%
	Water	206	65	10	635	4	0	11	931	68%
	Grass	77	24	8	27	204	0	92	432	47%
	Crops	0	0	0	0	0	0	0	0	N/A
	Shrub and trees	25	25	3	19	100	0	698	870	80%
	Total	924	756	75	895	484	0	866	4000	
P.A.	46%	64%	72%	71%	42%	N/A	81%		PCC: 63% kappa: 0.54	

TABLE 6. ERROR MATRIX FOR THE CLASSIFICATION AFTER SHADOW REMOVAL

		Reference Data							Total	U.A.
		Grey Surfaces	Red Surfaces	Bare Soil	Water	Grass	Crops	Shrub and Trees		
Land-cover map after shadow removal	Grey surfaces	471	18	0	3	3	0	3	498	95%
	Red surfaces	8	80	2	1	1	0	1	93	86%
	Bare soil	74	19	239	0	9	0	2	343	70%
	Water	31	0	0	120	0	0	0	151	79%
	Grass	16	3	1	0	285	17	18	340	84%
	Crops	0	2	2	0	37	205	29	275	75%
	Shrub and trees	3	4	1	0	18	22	303	351	86%
	Total	603	126	245	124	353	244	356	2051	
P.A.	78%	63%	98%	97%	81%	84%	85%		PCC: 83% kappa: 0.79	

labeled as shadow into their actual type of land-cover as identified in the field introduces an additional uncertainty, which explains the slight drop in the kappa value. However, taking into account that all shadow areas are removed from the classification and have been replaced by actual land-cover classes, the result may be considered as quite successful. It should be pointed out that for the water class there is a drop in user's accuracy from 99 percent to 79 percent. This can be explained by the still quite high proportion of shadow pixels on dark grey surfaces being reclassified as water by the neural network used for shadow reclassification. This also explains the drop in producer's accuracy for "grey surfaces" compared to the original classification. The erroneous assignment of shadow pixels to the water class can be clearly observed in Plate 1c. Correcting for these and other thematic errors was achieved by rule-based enhancement of the classification output (see below).

Knowledge-based Classification Enhancement and Structural Filtering

Table 7 shows the 10 knowledge-based rules in the order we have applied them on each of the five test zones. The rules can be divided into four groups:

- (a) Rules that put a threshold on a region's size to determine whether its label should be changed into an adjacent region's label (rules 4 and 7);
- (b) Rules that put a threshold on a region's size to determine whether its label should be changed into the label of the neighboring region with which it shares the largest border (rules 1 and 8);
- (c) Rules that change a region's label into the label of the region that completely surrounds it (rule 10);
- (d) Rules that use the second most probable class of a region's pixels, as indicated by the classifier, to
 - change the label of some of these pixels into the label of one of the neighboring regions (rules 2, 3, and 9);
 - change the label of *all* pixels that belong to the region into the label of one of the neighboring regions (rules 5 and 6).

Rule 1 was used to correct small patches that were misclassified as "water" mostly because of errors introduced by shadow-reclassification or confusion between "shadow" and "water" (see above).

A more severe problem in the classification of the Ghent area is the confusion between "bare soil" and "red surfaces." First of all, some parts of red-colored roofs were incorrectly identified as "bare soil" by the classifier. Because these erroneous "bare soil" patches are generally small in size, a rule of type (d), combining area thresholds and second

TABLE 7. OVERVIEW OF THE 10 KNOWLEDGE-BASED RULES IN THE ORDER THEY WERE APPLIED

- (1) <water> regions smaller than 250 pixels are assigned to the neighboring region with which they share the largest boundary
- (2) pixels belonging to <bare soil> regions smaller than 500 pixels are assigned to <grey surfaces> if their second probability is <grey surfaces> and their region is adjacent to a <grey surfaces> region which in turn is larger than 100 pixels
- (3) pixels belonging to <bare soil> regions smaller than 500 pixels are assigned to <red surfaces> if their second probability is <red surfaces> and their region is adjacent to a <red surface> region which in turn is larger than 100 pixels
- (4) <red surface> regions smaller than 150 pixels are assigned to <grey surfaces> if they are adjacent to a region of this class that is larger than 500 pixels
- (5) if the majority of the pixels inside a <red surfaces> region has <bare soil> as its second most probable class AND if this region is adjacent to <bare soil> regions with a total area of at least 12,000 pixels, then the <red surfaces> region is assigned to <bare soil>
- (6) if the majority of the pixels inside a <grey surfaces> region has <bare soil> as its second most probable class AND if this region is adjacent to <bare soil> regions with a total area of at least 12,000 pixels, then the <grey surfaces> region is assigned to <bare soil>
- (7) <red surfaces> regions larger than 27,000 pixels are changed into <bare soil>
- (8) <shrub and trees> regions smaller than 175 pixels that are adjacent to <crops> or <grass> are changed into one of these two classes, depending on the class with which they share the largest boundary
- (9) pixels belonging to <crops> regions that are adjacent to <shrub and trees> or <grass> and that are smaller than 4,500 pixels, are assigned to their second most probable class if this class is either <shrub and trees> or <grass>
- (10) <bare soil> regions that are completely surrounded by <water> are assigned to <water>

probabilities (rule 3), enabled us to correct wrongly classified parts of roofs while leaving correctly labeled "bare soil" pixels in other parts of the image unchanged. Plate 1d illustrates this rule's effectiveness. It clearly shows how parts of red roofs which were wrongfully classified as "bare soil" (Plate 1b and 1c) are re-assigned to "red surfaces" (Plate 1d). This rule introduces only a limited amount of new errors by incorrectly changing small bare soil patches into red surfaces.

Another part of the "bare soil" and "red surfaces" confusion is the presence of red surface patches inside larger bare soil regions (not shown on the Plate). Many of

these apparently small red patches are connected and form large regions with highly irregular boundaries. Rule 5 was applied to solve this problem. It looks at the region's neighborhood (regions to be transformed should be adjacent to large patches of bare soil) and the second most probable class of the region's pixels as identified by the classifier (should be bare soil). Sometimes entire fields of "bare soil" are wrongly assigned to the "red surfaces" class. Rule 7 deals with this type of confusion. It uses a threshold that exceeds the size of the largest red colored roof regions found in the area. As such, there is no risk of introducing new errors this way.

Rule 2 is similar to rule 3, but removes erroneous bare soil patches found within "grey" regions, e.g., on asphalted roads or on grey roofs (see Plate 1c). After the application of rule 2, the roads belong nearly exclusively to the "grey surfaces" class (Plate 1d). Similarly, rule 4 removes very small "red surfaces" clumps that occur in larger grey areas by applying an area threshold both on the region to be changed and the target region. These small red patches are often caused by cars present on asphalted or concrete roads. Rule 6 works in a similar fashion as rule 5, and solves confusion between grey surfaces and bare soil in non-residential areas.

Rules 8 and 9 deal with confusion between "crops," "grass," and the "shrub and trees" class. In the small agricultural field in the western part of Plate 1b, for instance, some pixels are classified as "shrub and trees" instead of crops, making the field look cluttered. After applying rule 8, this noise is removed and the field becomes more homogeneous (Plate 1d). Rule 9 deals with the problem of mixed pixels in the transition zones between "grass" and "shrub and trees" that are misclassified as "crops." Such misclassified patches are generally small, so again an area threshold is applied. A second condition for re-labeling is that the region needs to be adjacent to either "grass" or "shrub and trees," or to both. We then re-assign pixels in the region to the second most likely class, if this class is "grass" or "shrub and trees." If this last condition is not fulfilled, the pixel is left unchanged.

Finally, the type (c) rule in the set (rule 10) deals with structural clutter in large water bodies. Indeed, a very large number of small "bare soil" patches appeared within the waterways near the harbor (test zone 1). Because such patches are fully enclosed by regions of the "water" class, a simple rule could be defined to correct for this type of classification error.

The 10 post-classification rules were applied to the land-cover classification of the five test zones in the order specified in Table 7, after shadows had been removed. Before and after applying the rules, a structural filter assigning regions with a size smaller than 16 pixels (6.35 m²) to the largest neighboring region was applied to reduce image clutter. An extract of the result obtained is shown in Plate 1d. Using the same validation set as in the previous step, we obtained a kappa index of 0.86 (PCC of 88 percent), a substantial improvement of classification accuracy compared to the classification result obtained after shadow removal (kappa = 0.79, PCC = 83 percent). Table 8 shows the confusion matrix. After shadow removal, bare soil was the class with the lowest user's accuracy because of the confusion with red and grey surfaces (see Table 6). After post-classification, the user's accuracy for bare soil increases from 70 percent to 75 percent. For red surfaces there is also a slight improvement in the user's accuracy (from 86 percent to 89 percent). The confusion between crops and other vegetation classes is also greatly reduced. User's accuracy for crops increases from 75 percent to 91 percent, which proves that the knowledge-based rule used to reclassify the crop pixels in the transition zones between trees and grass performs well. We also notice a clear improvement for the "shrub and trees" class (+8 percent). Many of the shadow pixels that were erroneously assigned to the water class in the phase of shadow removal (see above) were successfully corrected by rule 1, increasing the user's accuracy for water from 79 percent to 89 percent. Producer's accuracies are improved for all classes.

Because we defined our rules starting from the actual errors, we observed in a particular classification, one should keep in mind that they are not unique in the sense that classification of other areas and/or the use of another classification key would require a different set of rules. Even for the same classification, a similar improvement might have been obtained with a different rule set. The order in which the rules are applied is also important because applying one rule to solve a particular problem might introduce new errors elsewhere in the image that need to be solved with yet another rule. The major advantage of rule-based post-classification, however, is its transparency and relative simplicity, which makes it possible for non-image specialists to use it for their own purposes. The fact that the rules we developed for our case study work well on five separate test zones, each characterized by different types of urban morphology, proves that the rules can, at least to a certain extent, be generally defined. On the other hand, it

TABLE 8. ERROR MATRIX FOR THE CLASSIFICATION AFTER SHADOW REMOVAL, RULE-BASED CLASSIFICATION ENHANCEMENT AND STRUCTURAL FILTERING

		Reference Data							Total	U.A.
		Grey Surfaces	Red Surfaces	Bare Soil	Water	Grass	Crops	Shrub and Trees		
Land-cover map after rule-based enhancement and structural filtering	Grey surfaces	509	24	0	2	5	0	2	542	94%
	Red surfaces	6	85	0	1	1	0	2	95	89%
	Bare soil	61	8	243	0	9	0	1	322	75%
	Water	15	0	0	121	0	0	0	136	89%
	Grass	8	4	0	0	308	25	23	368	84%
	Crops	0	1	2	0	16	219	2	240	91%
	Shrub and trees	4	4	0	0	14	0	326	348	94%
	Total	603	126	245	124	353	244	356	2051	
P.A.	84%	67%	99%	98%	87%	90%	92%		PCC: 88%	kappa: 0.86

should be mentioned that the definition of a proper set of rules requires a considerable effort from the operator. To reduce the amount of user intervention needed to define appropriate post-classification rules, one might think of developing an automated or semi-automated approach that is able to generate re-classification rules when provided with a certain classification and with additional knowledge on the spatial characteristics of classification error. This is clearly a topic for future research.

Conclusions

The prime objective of this work was to define a suitable strategy to produce an accurate land-cover classification for an urban area using Very High Resolution satellite data. Post-classification processing was applied to meet this goal by reducing some of the major problems of pixel-based classifications in a complex urban setting: the presence of shadows, misclassifications due to spectral confusion between classes, and structural clutter. An initial pixel-based classification of five test zones in and around the city of Ghent, Belgium was carried out with an artificial neural network classifier. The kappa index-of-agreement obtained with this approach was 0.82, with an individual class user's accuracy ranging from 0.72 to 0.99. The poorest performing class was bare soil, which was mostly confused with red and grey surface types. After applying an ANN-based shadow removal technique, the kappa index dropped slightly to 0.79. This appears to be only a slight trade-off, considering the fact that shadows hide between 3 percent and 14 percent of the area of the different test zones. The knowledge-based post-classification rules and the structural filter we applied to deal with misclassifications and structural clutter not only improved the visual appearance of the classification significantly, they also increased the overall kappa index to 0.86.

The methods described in this paper will be applied on two morphologically different Belgian cities in the near future, and will also be used to produce detailed reference classifications for sub-pixel classification of Landsat ETM+ data. Future research needs to consider developing automated procedures that are able to define rules directly from a given classification, using knowledge on image context and classification error.

Acknowledgments

The authors wish to express their gratitude to Marc Binard of SURFACES (Université de Liège) and to Nathalie Stephenne of IGEAT (Université Libre de Bruxelles) for their part in the collection of training and validation data for the land-cover classification. They also wish to thank Dennis Devriendt (Geography Department, Ghent University) for building the DSM's required for orthorectification. Belgian Science Policy is gratefully acknowledged for providing the funds for this research. The authors also would like to thank the anonymous reviewers for their comments.

References

Arnold, C.A., Jr., D.L. Civco, S. Prisloe, J.D. Hurd, and J. Stocker, 2000. Remote sensing-enhanced outreach education as a decision support system for local land use officials, *Photogrammetric Engineering & Remote Sensing*, 66(10):1251–1260.

Baatz, M., and A. Schäpe, 2000. Multiresolution segmentation – An optimization approach for high quality multi-scale image segmentation, *Angewandte Geographische Informationsverarbeitung XII, Beiträge zum AGIT-Symposium - Salzburg 2000* (J. Strobl, T. Blaschke, and G. Griesebner, editors), Herbert Wichmann Verlag, Karlsruhe, Germany, pp. 12–23.

Baelus J., G. Vloebergh, and J. De Greef, 2003. *Ruimtelijk structuurplan Gent – Informatief gedeelte*, Stad Gent, Ghent, Belgium, 210 p.

Barnsley, M.J., and S.L. Barr, 1996. Inferring urban land use from satellite sensor images using kernel-based spatial reclassification, *Photogrammetric Engineering & Remote Sensing*, 62(8): 949–958.

Barr, S.L., and M.J. Barnsley, 2000. Reducing structural clutter in land-cover classifications of high spatial resolution remotely-sensed images for urban land use mapping, *Computers and Geosciences*, 26(4):433–449.

Barr, S.L., and M.J. Barnsley, 1998. Inferring urban land use from very high spatial resolution remotely-sensed images using syntactic pattern recognition techniques, *Proceedings of ECO BP'98: International Symposium on Resource and Environmental Monitoring, International Archives of Photogrammetry and Remote Sensing, Commission VII*, 01–04 September, Budapest, Hungary, pp. 315–322.

Chaudhuri, B., and N. Sarkar, 1995. Texture segmentation using fractal dimension, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(1):72–77.

Cohen, J., 1960. A coefficient of agreement for nominal scales, *Educational and Psychological Measurement*, 20(1):37–46.

Cornet, Y., C. Schenke, S. de Bethune, M. Binard, and F. Muller, 2003. Stratégies de fusion d'images P/XS basées sur les principes colorimétriques et l'Égalisation de Statistiques Locales, *Bulletin SFPT, French Society for Photogrammetry and Remote Sensing*, 169 (2003–1):35–45.

Dare, P.M., 2005. Shadow analysis in high-resolution satellite imagery of urban areas, *Photogrammetric Engineering & Remote Sensing*, 71(2):169–177.

European Commission - Environment DG, 2004. *Towards a Thematic Strategy on the Urban Environment*, Communication from the Commission to the Council, the European Parliament, the European Social and Economic Committee and the Committee of the Regions, European Publication Office, Luxembourg, 56 p.

Fahlman, S.E., and C. Lebiere, 1990. The Cascade-Correlation Learning Architecture, *Advances in Neural Information Processing Systems 2* (D.S. Touretzky, editor), Morgan Kaufmann, San Mateo, California, pp. 524–532.

Flanagan, M., and D.L. Civco, 2001. Subpixel impervious surface mapping, *Proceedings of the ASPRS 2001 Annual Conference*, 23–27 April, St. Louis, Missouri, American Society for Photogrammetry and Remote Sensing, Bethesda, Maryland, unpaginated CD-ROM.

Foody, G.M., 1997. Fully fuzzy supervised classification of land-cover from remotely sensed imagery with an artificial neural network, *Neural Computing and Applications*, 5(4):238–247.

Gong, P., and P.J. Howarth, 1992. Land-use classification of SPOT HRV data using a cover-frequency method, *International Journal of Remote Sensing*, 13(8):1459–1471.

Gurney, M.C., and J.R.G. Townshend, 1983. The use of contextual information in the classification of remotely sensed data, *Photogrammetric Engineering & Remote Sensing*, 49(1):55–46.

Haralick, R.M., and L.G. Shapiro, 1985. Image segmentation techniques, *Computer Vision, Graphics, and Image Processing*, 29(1):100–132.

Hofmann, T., J. Puzicha, and J. Buhmann, 1998. Unsupervised texture segmentation in a deterministic annealing framework, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(8):803–818.

Hudson, W., 1987. Correct formulation of the Kappa coefficient, *Photogrammetric Engineering & Remote Sensing*, 53(4):421–422.

Jensen, J., 2000. *Remote Sensing of Environment: An Earth Resource Perspective*, Prentice Hall, Upper Saddle River, New Jersey, 544 p.

Pal, N.R., and S.K. Pal, 1993. A review on image segmentation techniques, *Pattern Recognition*, 26(9):1277–1294.

Panjwani, D., and G. Healey, 1995. Markov random field models for unsupervised segmentation of textured color images, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(10):939–954.

- Pillmann, W., and K. Kellner, 2002. Green space inventory in the City of Vienna: A BiotopMonitoring system based on remote sensing methods, *Urban Forests and Trees, Proceedings No. 1, COST Action E12*, European Communities, Office for Official Publications, Luxembourg.
- Ries, C., W. Pillmann., K. Kellner, and P. Stadler, 2002. Urban green space management information - Processing and use of remote sensing images and scanner data, *Proceedings of Environmental Informatics 2002: 16th International Conference on Informatics for Environmental Protection*, 25–27 September, Vienna, Austria, pp. 503–510.
- Rosenfield, A., and L.S. Davis, 1979. Image segmentation and image models, *Proceedings of the IEEE*, 67(5):764–772.
- Sahoo, P.K., S. Soltani, and A.K.C. Wang, 1988. A survey of thresholding techniques, *Computer Vision, Graphics, and Image Processing*, 41(2):233–260.
- Salari, E., and Z. Ling, 1995. Texture segmentation using hierarchical wavelet decomposition, *Pattern Recognition*, 28(12): 1819–1824.
- Schmied, A., and W. Pillmann, 2003. Tree protection legislation in European cities, *Urban Forestry & Urban Greening*, 2(2): 115–124.
- Strahler, A.H., C.E. Woodcock, and J.A. Smith, 1986. On the nature of models in remote sensing, *Remote Sensing of Environment*, 20:121–139.
- Sugumaran, R., D. Zerr, and T. Prato, 2002. Improved urban land-cover mapping using multi-temporal IKONOS images for local government planning, *Canadian Journal of Remote Sensing*, 28(1):90–95.
- Thomas, N., C. Hendrix, and R.G. Congalton, 2003. A comparison of urban mapping methods using high-resolution digital imagery, *Photogrammetric Engineering & Remote Sensing*, 69(9):963–972.
- UNCHS - United Nations Centre for Human Settlements – UNCHS, 2001. *The State of the World's Cities Report 2001*, Nairobi, Kenya, 125 p.
- Van de Voorde, T., W. De Genst, F. Canters, N. Stephenne, E. Wolff, and M. Binard, 2004. Extraction of land use/land-cover related information from very high resolution data in urban and suburban areas, *Remote Sensing in Transition* (R. Goossens, editor), *Proceedings of the 23rd Symposium of the European Association of Remote Sensing Laboratories*, 02–05 June 2003, Ghent, Belgium, Millpress Rotterdam, Netherlands, pp. 237–244.
- Wegner, S., H. Oswald, and E. Fleck, 1997. Segmentierung mit der Wasserscheidentransformation, *Spektrum der Wissenschaft*, 6:113–115.
- Wharton, S.W., 1982. A contextual classification method for recognizing land-use patterns in high-resolution remotely sensed data, *Pattern Recognition*, 15(4):317–324.
- Yang, L., C. Huang, C.G. Homer, B.K. Wylie, and M.J. Coan, 2003. An approach for mapping large-area impervious surfaces: Synergistic use of Landsat-7 ETM+ and high spatial resolution imagery, *Canadian Journal of Remote Sensing*, 29(2):230–240.

(Received 22 June 2005; accepted 03 October 2005; revised 10 February 2006)