Enhancing and replacing spectral information with intermediate structural inputs: A case study on impervious surface detection

Giorgos Mountrakis *, Li Luo

Department of Environmental Resources Engineering, State University of New York College of Environmental Science and Forestry, 1 Forestry Dr, Syracuse, NY 13210, USA

A B S T R A C T

This paper assessed the incorporation of road structural information in the classification process of impervious surface areas. A multi-process classification model was adopted and it consisted of an a priori classifier and an a posteriori classifier. The role of the a priori classifier was to classify the relatively simple portions of the image. This partial classification acted as the basis for the production of linear features using an iterative Radon transform. Spatial statistics derived from the linear features led to road structural intermediate inputs (RSIIs) (for example, distance to the closest segment endpoint). RSIIs were integrated with spectral information on the remaining unclassified pixels and an assessment was done to evaluate whether they would improve a binary impervious classification task. The experimental results on a 2006 Landsat ETM+ image suggested that classification accuracy improved by 8.4% for the portion of the dataset classified with the a posteriori classifier and led to an improvement of 3.2% over the entire dataset. In addition, a more challenging and wide-reaching hypothesis was tested, namely whether RSIIs could completely replace spectral information in portions of the image instead of complementing it. Exclusive use of RSIIs matched or improved classification accuracy obtained solely from spectral information, even when more than half of the validation dataset was forwarded to the a posteriori classifier. This finding offers an important contribution to the remote sensing community, since the proposed methodology handles the missing spectral information problem through exclusive analysis of the given degraded image; no external information, such as spectral information from other times and/or vector data, is needed.

© 2011 Elsevier Inc. All rights reserved.

1. Introduction

The human footprint on the earth's landscape is expressed through several land cover and land use changes. Constructed impervious surfaces such as roads, buildings and parking lots are a prominent component of these changes, with significant environmental impacts on humans, biodiversity, and earth resources. Impervious surface estimation has become a key indicator in environmental studies (Arnold & Gibbons, 1996), motivating remote sensing studies on accurate estimation of imperviousness. In recent research, remote sensing techniques have been successfully applied to imperviousness detection and mapping (Carlson, 2003; Elmore & Guinn, 2010; Esch et al., 2009; Forster, 1980; Gillies et al., 2003; Weng, 2007; Xian & Crane, 2005). Increasing interest has led to the investigation of various classification methods of imperviousness, ranging from spectral mixture models (Franke et al., 2009; Lu & Weng, 2006; Powell et al., 2007; Wu & Murray, 2003) and multivariate regression models (Bauer et al., 2005; Yang & Artigas, 2008) to machine learning algorithms such as decision tree

(Herold, 2003; Xian et al., 2008; Yang et al., 2003) and neural networks (Hu & Weng, 2009; Lee & Lathrop, 2006; Luo & Mountrakis, in press).

Among the classification methods, several studies have incorporated road structural information to improve classification accuracy. In early research of this field, Gong and Howarth (1990) generated an edge-density image using a Laplacian operator from SPOT multispectral data. The structural information derived was then incorporated into conventional classification procedures, improving the overall accuracy from 76.6% to 86.1% by using structural information. Wang (1995) applied the linear feature detection and analysis system (called LINDA) to extract linear feature networks in order to obtain a road network density map for urban–rural area segmentation using a Landsat TM image of the Kitchener–Waterloo area in Ontario, Canada. The results demonstrated the LINDA system as a useful method for remote sensing image segmentation. Zhang et al. (2002) proposed a new structural method based on road density combined with spectral bands for change detection. The road density maps were produced using a gradient directional profile analysis algorithm. This method reduced spectral confusion and improved classification accuracy. Guindon et al. (2004) extracted line-like features and computed line density to assist with segment-based land-cover classification using Landsat TM and ETM imagery. In addition to image derived road structural information, external road data have been utilized to assist the classification process. Thomas and Endreny (2008) converted vector road files to raster data and created spatially explicit

* Corresponding author at: Dept. of Environmental Resources Engineering, College of Environmental Science and Forestry, State University of New York, 419 Baker Hall, 1 Forestry Dr., Syracuse, NY 13210, USA. Tel.: +1 315 470 4824; fax: +1 315 470 6958.

E-mail address: gmountrakis@esf.edu (G. Mountrakis).

representations of impervious area associated with roadways to improve national land cover database estimates of road network impervious cover. Esch et al. (2009) presented a flexible and large-area mapping approach combining Landsat images and road network vector data using support vector machines to estimate impervious surface.

In contrast to previous research, in this study road structural information was extracted from a partially classified image. A multi-process classification model consisting of an a priori classifier and an a posteriori classifier was developed. The a priori classifier labeled a portion of the entire image. This portion was used to generate road structural information to assist the classification of leftover pixels using the a posteriori classifier. The concept of "intermediate inputs"—structural information created by partially classified images—was first presented by Luo and Mountrakis (2010). In this paper, we make use of a much more elaborate road identification process for the purpose of creating road structural intermediate inputs for a significantly different evaluation process. We examine not only whether structural information can complement spectral information, but also whether it can entirely replace it. This examination is an important scientific but also practical investigation, because the advanced age of existing satellite platforms makes this equipment prone to errors; for example, in 2003 the Scan Line Corrector that compensates for the forward motion of Landsat 7 failed, leading to gaps in spectral returns towards the edges of a scene. We tested our method using imagery from central New York, with some surprising results.

2. Study area and dataset

A region of 15 km × 12 km located in central New York was selected for evaluation of the proposed classification approach. This region covered the rural area of Syracuse, New York. The major land cover types include water, grass, forest, bare soil, and impervious surface. The predominant impervious feature is the road network, which has a limited number of developed urban blocks. A Landsat 7 Enhanced Thematic Mapper Plus (ETM+) scene taken on 18 April, 2006 was used in this study. The ground sample distance (GSD) is 30 m and 6 bands (blue, green, red, near IR, and two mid IR bands) were utilized as the spectral information for the classification (Fig. 1).

The research presented in this paper aimed at estimating impervious surface at the binary level instead of impervious percentage distribution; therefore, the goal was to identify whether a pixel contained any amount of impervious cover. In our dataset creation, if any portion of the 30 m pixel occupied constructed impervious then that pixel was assigned to the impervious class. This rather strict binary threshold was used to ensure that the spatial structure of impervious surfaces remained intact. In addition, from the practical perspective we expect our analysis to provide an accurate filtering step for subpixel imperviousness classification models, allowing them to focus only on pixels with known imperviousness. This would limit overestimation issues, a typical problem in rural areas due to spectral confusion, especially with some soil types.

For binary classification purposes, Digital Orthophotography (DO) from 2006 with a GSD of 0.6 m was overlaid on the 30 m Landsat imagery. The reference dataset was produced with on-screen digitizing. The entire reference dataset was then divided equally into two parts: west and east (Fig. 1). From each of the two dataset parts 18,000 pixels were selected, adopting a stratified sampling technique (9000 impervious pixels and 9000 non-impervious pixels) that led to the creation of the calibration and validation datasets. The calibration dataset was used to train and test the classification model, while the validation dataset was utilized to assess the performance. All experimental results in Section 4 refer to the validation dataset.

3. Methodology

3.1. The multi-process classification framework

Luo and Mountrakis (2010) presented the concept of "intermediate inputs": statistical calculations based on partially classified images. The classification process was organized in two general processes: the a priori classifier that is utilized to create a partial classification, and the a posteriori classifier that converts the partial results into contextual information (intermediate inputs) under the hypothesis that they will assist with the classification of the unclassified leftover pixels. This paper follows the same conceptual framework; however, the a posteriori classifier is completely different
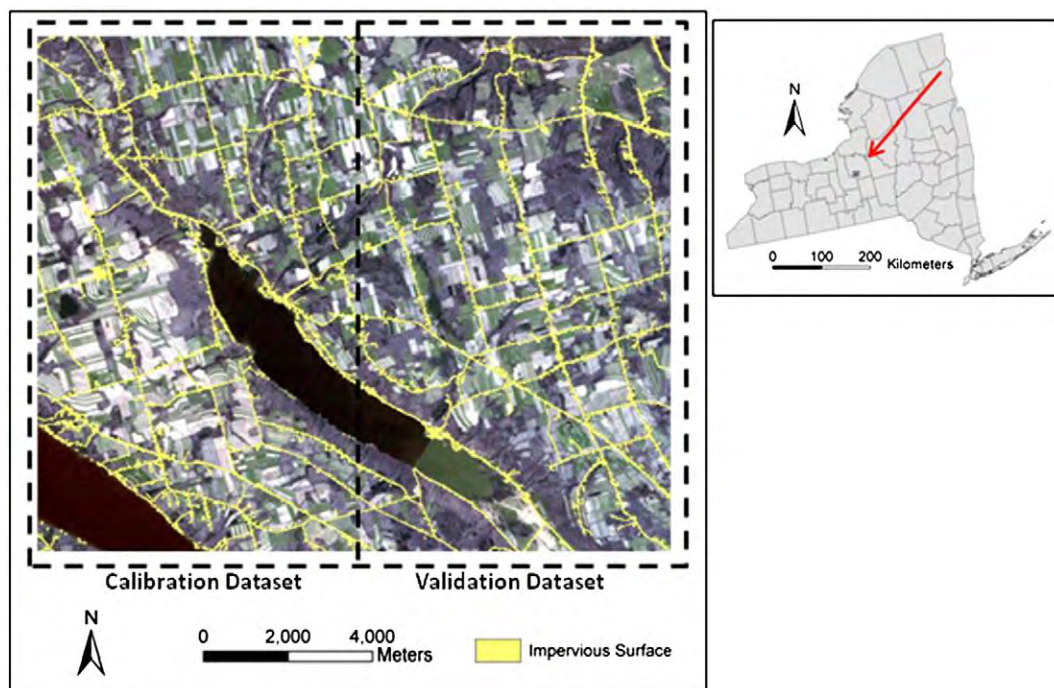

Fig. 1. The Landsat 7 ETM+ imagery used in this research displayed using RGB bands. Impervious surface reference data are overlaid in yellow. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

from that used in previous work. Furthermore, our assessment sheds light on an even more important question than the enhancement of spectral information with intermediate inputs: it investigates the replacement of spectral information with contextual intermediate inputs for certain image portions.

The general framework of the multi-process classification model is illustrated in Fig. 2. A partially classified image was initially produced using a backpropagation neural network classifier (a priori classifier) and an accuracy threshold, leading to three possible outcomes: pixels falling into either the impervious or the non-impervious class, or pixels that remained unclassified. Already-labeled pixels were further analyzed to derive road structural information. This information was used in two subsequent classifiers. The first one was a selective dilation morphological operation, a process that assigned only impervious pixels by connecting road segments. The second classifier, named the a posteriori classifier and the focus of this paper concentrated on classifying the leftover pixels from the a priori and the dilation processes. The final classification product is a simple mosaic of the three classifiers (dotted rectangles in Fig. 2) because there was no spatial overlap in the operation of these two classifiers.

3.2. A priori classifier and partial classification derivation

The multi-layer perceptron feedforward neural network trained using the backpropagation learning algorithm was adopted as the a priori classifier to produce the partially classified results. The six bands of the Landsat ETM+ imagery from the calibration dataset served as the input nodes, and the two output nodes represented impervious and non-impervious classes, respectively. Different hidden layer architectures were tested (1000) in order to determine the optimal hidden layer structure and weights. The architecture of each of the 1000 candidate neural networks was randomly selected from 5 to 10 nodes for the first hidden layer and between 0 and 5 nodes for the second hidden layer. The range of the hidden layer nodes was kept low to avoid overfitting. The calibration dataset was divided into training and testing datasets, with a size ratio of 7:3. Each neural network candidate was trained using the training dataset, and the corresponding overall classification accuracy was calculated on the testing dataset. The network with the highest testing accuracy was adopted as the a priori classifier.

The number of classified pixels extracted by the a priori classifier was determined using an accuracy threshold. The accuracy threshold was translated into an output node threshold for the neural network, using the testing dataset as a guide; this process involved gradually decreasing output node response until the given accuracy was achieved. As the accuracy threshold increased, the number of classified pixels would become smaller but with higher accuracy. To the contrary, implementation of a lower accuracy threshold would classify more pixels with the a priori classifier leaving a fewer but more challenging number of pixels for the a posteriori classifier. In order to study the effects of road structural information on different a priori contribution levels, six accuracy thresholds from 84% to 94% with an increment of 2% were implemented, generating different partial classifications.

3.3. Incorporation of a priori classifier information

Road structural information was generated from the partially classified image. The motivation was to complement or replace the spectral information of the unclassified pixels with contextual relationships as expressed through road structural information to increase classification accuracy. The fact that our method examined the replacement (instead of complement) of spectral with contextual information in image portions is a major motivation for this research. The lack of quality spectral information in remotely sensed images is a common problem due to atmospheric effects, topography, and sensor malfunctions.

Existing research has reported road extraction using numerous approaches, such as dynamic programming (Amini, 2009; Bentabet et al., 2003; Gruen & Li, 1997; Laptev et al., 2000), Hough transform (Lee & Moon, 2002) and Radon transform (Li et al., 2010; Zhang & Culoigner, 2007), mathematic morphology (Letitia, 2010; Valero et al., 2010; Zhu et al., 2005), and template matching (Kim et al., 2004; Shi & Zhu, 2002). In this paper, an iterative localized Radon transform approach was applied to detect the road linear feature. We should note that the focus of this paper is not to present a novel road
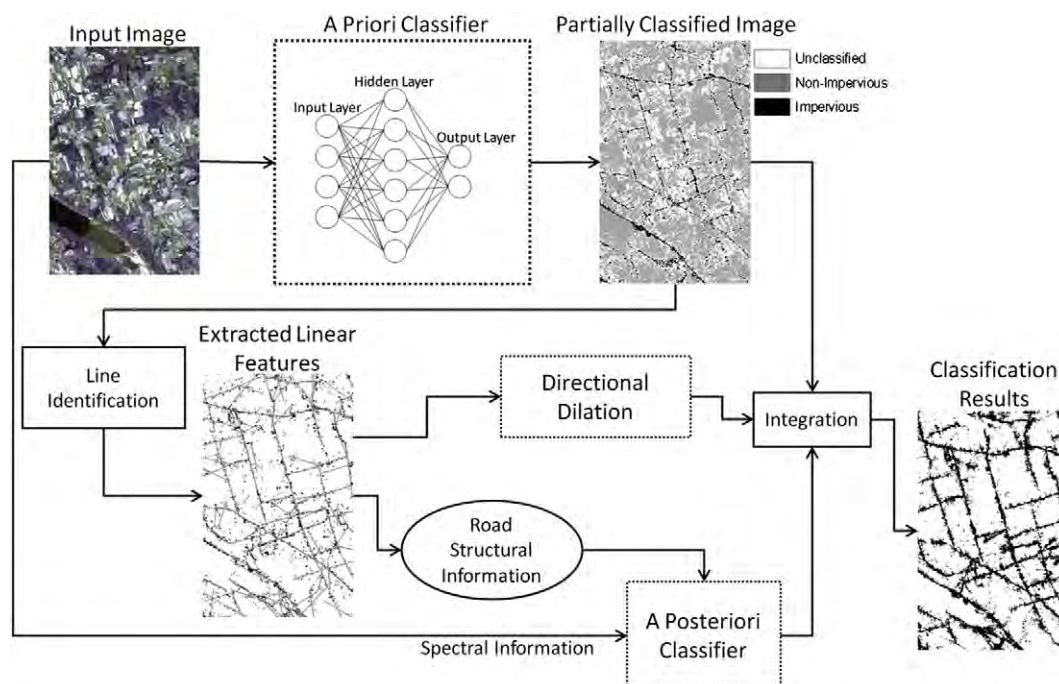

Fig. 2. The framework of the multi-process classification model.

detection mechanism, instead facilitate further usage of detected road structure.

Within the a posteriori classifier, several processes were initiated (Fig. 2). Firstly, preprocessing was applied on the partial classification image from the a priori classifier. Pixel assignment changed from three categories (impervious, non-impervious, and unclassified) into two, by merging the non-impervious and unclassified classes. This new binary map was processed using the mean of Angular Texture Signature and a closing morphological operation, in order to remove non-road clusters and strengthen road linear features. The iterative localized Radon transform approach was then applied to extract road linear features. Based on the detected road linear features, a directional morphological operation was applied and five road structural intermediated inputs were generated. The five inputs were utilized in two ways during the classification process of leftover pixels: first by complementing with the spectral information, and later by replacing it. Details on all these process are presented in the following subsections.

3.3.1. Linear feature detection

3.3.1.1. Road preprocessing.
The impervious pixels in the derived binary partial classification contained both road and non-road pixels (e.g. buildings and parking lots) under the impervious class. The existence of non-road clusters interfered in the detection of road linear features. To compensate for this, the mean of Angular Texture Signature (ATS) (Gibson, 2003; Zhang & Culoigner, 2007) was adopted as a shape descriptor to remove non-road clusters. The higher the ATS mean value, the larger its probability of belonging to a non-road cluster. The impervious pixels with corresponding high ATS mean values larger than a threshold identified by visual inspection (e.g. 0.41 for the accuracy threshold of 92% applied on a priori classification) were regarded as non-road clusters and removed from the binary image. After the non-road clusters were removed using the ATS mean values, a closing morphological operation was then applied on the updated binary partially classified image (impervious pixels and others) using a 3×3 rectangle structure element. The purpose of the closing operation was to smooth sections of road segments and fill gaps in them.

3.3.1.2. Iterative localized Radon transform algorithm.
The Radon transform introduced by Johann Radon (Radon, 1917), is an integral transform which computes projections of an image matrix along specific directions. Although the basic theory of Radon transform is simple, several problems arise when applied globally to the entire image for linear feature detection. Since the integral transformation of the straight lines is performed globally, it may have difficulty detecting shorter line segments (Murphy, 1986). Long line segments displaying some curvature may also not produce suitable peaks or troughs during the transform (Copeland et al., 1995). Since the extraction of linear features is crucial for the generation of road structural information in this research, additional attention was given to this process. The final winning strategy was an iterative localized Radon transform algorithm (Fig. 3). Based on the idea of iterative Radon transform proposed by Zhang and Culoigner (2007), the derived binary image (after the ATS and closing preprocessing) was partitioned into smaller subimages (50×50 pixels), most of which contained a few road lines. The Radon transform was performed within each subimage using an iterative process. At each iteration, only one road line was detected from the peak of the Radon transform intensity image. Pixels within a one pixel buffer area from the detected road centerline were then eliminated from the input subimage, and the Radon transform was performed again on the new subimage. The iterative process would stop when either the intensity peak reached a threshold or no more pixels were left. The intensity threshold was identified through visual inspection on
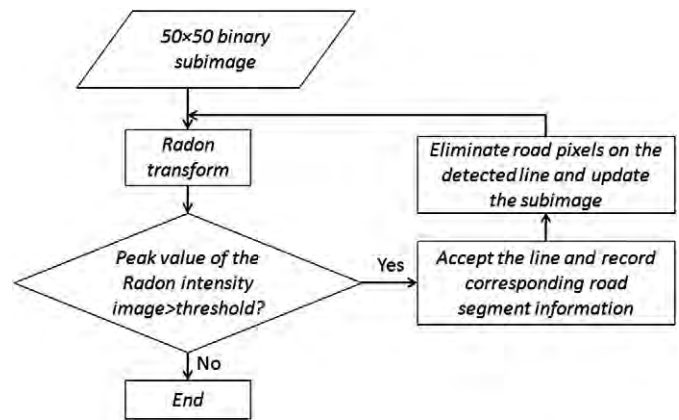

Fig. 3. The flowchart of the iterative localized Radon transform.

the calibration dataset and was set to 30% of the highest peak value identified at the first iteration. Each detected road line was then transformed to several line segments consisted of connected pixels within a one pixel buffer area from the detected road line. Angle, position, and endpoint information of all line segments were recorded and prepared for further processing.

3.3.2. Directional dilation

In order to refine the detected road network further, a directional dilation morphological operation was selectively applied. The size of the structure element was 5×1, where the length of five pixels was deemed appropriate to fill in gaps while the width of one pixel reflected the typical road width of a binary classification at the 30 m pixel size. The directional dilation was only applied on the line endpoints, as identified by the Radon process above, with the goal to elongate the discrete road segments. The direction of the linear structure element was dynamically adapted to reflect the angle of the road segment where the endpoint pixel was located. After directional dilation, endpoint information was updated using the new road segments.

3.3.3. Intermediate inputs derivation based on road structure

A refined partially classified image was produced after the application of the directional dilation. Through dilation the number of impervious pixels increased as more pixels were identified as road pixels; this process in reverse reduced the number of non-impervious pixels and unclassified pixels. The endpoint information was updated as well, since each road segment had been elongated. Five road structural intermediate inputs (RSIIs) were generated exclusively on the unclassified pixels through analysis of the road segment properties (segment length and direction and endpoint identification). The five RSIIs were:

RSII 1: The distance from an unclassified pixel to the closest road segment. The underlying assumption is that impervious surfaces are typically close to road pixels.

RSII 2: The distance from an unclassified pixel to the closest road segment endpoint. This input emphasized further proximity to road segments under the assumption that if an unclassified pixel is close to a road endpoint it has a higher chance of being part of the road network as well.

RSII 3: The angle between the line connecting an unclassified pixel and its closest endpoint, and the line describing the segment owning that endpoint (Fig. 4). The expectation was that angles close to 0° or 180° would be road pixels as they would continue along the same direction.

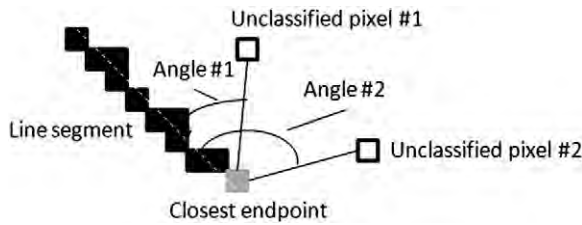RSII 4: The ratio between the sum of the distances from the unclassified pixel to the two closest endpoints and the

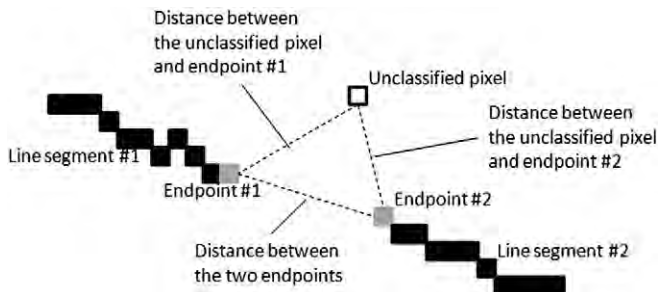**Fig. 4.** Calculation of the third road structural intermediate input.

Fig. 5. Calculation of the fourth road structural intermediate input.

Fig. 6. Kappa statistics using each RSII at different accuracy thresholds.

distance between the two endpoints (Fig. 5). The two endpoints were forced to belong to different line segments. Ratio values close to one would indicate possible merger of the two line segments.

RSII 5: The length of the line segment on which the closest endpoint of the unclassified pixel was located. This input was designed as supplementary to the other RSIIs under the assumption that longer segments should incorporate higher flexibility in the other RSII metrics because there is more clear evidence of road presence.

4. Experimental results

4.1. Focusing on the a posteriori classifier

The a posteriori classifier's purpose was to identify the unclassified pixels remaining from the application of the a priori and dilation classifiers. The experimental results illustrated in this section are based on the validation dataset pixels remaining after the application of the two aforementioned methods.

4.1.1. RSII sensitivity analysis

A sensitivity analysis was conducted on the five road structural intermediate inputs (RSIIs) to assess their performance during the classification process of the leftover pixels after directional dilation. A simple threshold-based calibration approach was adopted and each road structural intermediate input was applied separately. For each RSII, 100 gradually increasing thresholds from the minimum values to the maximum values were generated, using a fixed small step size. These thresholds were applied in the calibration dataset and the optimal threshold with the best overall accuracy was selected. The optimal thresholds were then applied on the validation dataset. Fig. 6 shows the kappa statistics of using each RSII separately. The process was replicated for 6 different accuracy thresholds for the a priori classifier. The RSII 2, which describes the distance from an unclassified pixel to its closest endpoint, and the RSII 3, which describes the angles between the line connecting the unclassified pixel with its closest endpoint pixel and the line in which the endpoint pixel was located, revealed the best kappa statistics at all accuracy thresholds. The kappa statistic curves of RSII 2 and RSII 3 show a roughly increasing tendency with the increasing of accuracy thresholds, which indicates that the
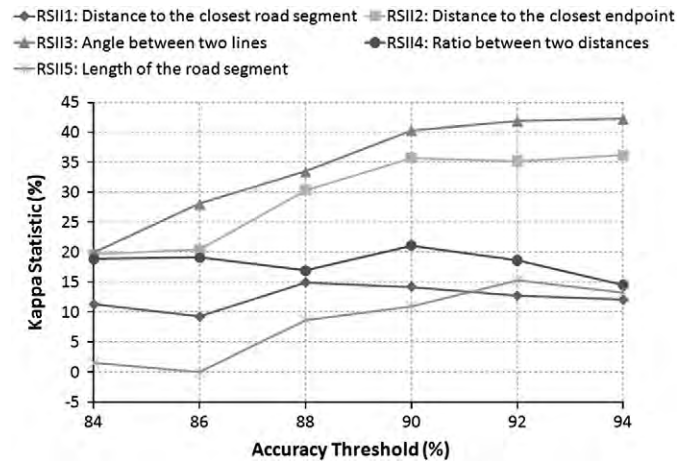
performances of the three inputs could have improved because more pixels were identified in the partial results. By contrast, the performances of RSII 1, RSII 4 and RSII 5 were not affected significantly by the number of the classified pixels within the partial results, since the curves of the two inputs exhibit only a small fluctuation at all accuracy thresholds. Being supplementary information to other RSIIs, RSII 5, which calculates the length of the line segment on which the closest endpoint of the unclassified pixel is located, illustrated the least efficiency within the five inputs, especially for low accuracy thresholds where the low kappa statistics indicated that the classification process was almost random.

4.1.2. Evaluation of RSII integration and replacement of spectral information

After applying different a priori classifier accuracy thresholds and corresponding directional dilation processes, three different scenarios were tested in classifying the leftover pixels using the identical training/testing dataset. The first scenario used exclusively spectral information from the six bands and acted as the benchmark case. The second scenario used exclusively the five road structural intermediate inputs (RSIIs) and examined the hypothesis that contextual information could completely replace spectral information. The third scenario assessed whether the integration of the six bands with the five RSIIs would further improve classification performance. All three scenarios incorporated the same multi-layer perceptron feedforward neural network training approach described in the a priori classifier (Section 3.2). The fourth scenario was created to assess exclusive use of a single RSII (RSII 3) using a simple threshold-based method. The purpose of this last scenario was to investigate whether additional RSIIs together with a neural network complex classifier offered substantial benefits.

An assessment took place using multiple accuracy thresholds, therefore leaving unclassified pixels of variable number and classification complexity. The Y axis on the right in Fig. 7 shows the portion of the leftover pixels within the entire validation dataset which were forwarded to all these algorithms. As the accuracy threshold increases, more pixels need to be identified in the following classification steps. The overall accuracies from the four scenarios are illustrated in Fig. 7 (Y axis on the left). As the accuracy threshold increased, the overall accuracies of leftover pixels increased as well. Along the same lines, the corresponding classification complexity decreased due to the fact that more pixels remained to be identified with higher accuracy.

Three important findings can be extracted from Fig. 7:

   i) Exclusive usage of RSIIs clearly outperformed classifiers using only spectral inputs, at all accuracy threshold levels. This is the
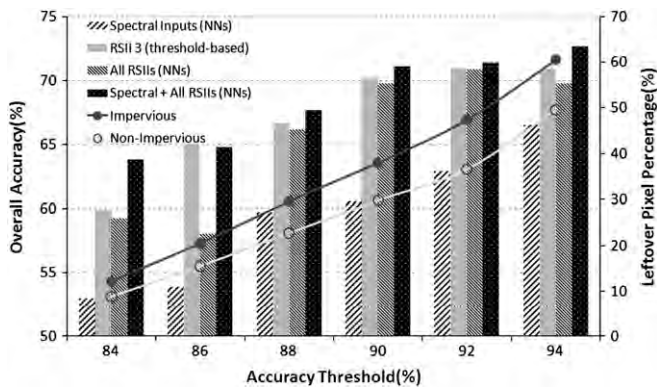
Fig. 7. The classification results of four scenarios for the a posteriori classifier.

major finding of this paper suggesting that missing spectral information can be overcome for the impervious classification task. To put this in further perspective, Fig. 8 demonstrates the portion of the validation image (identified as "classified by a posteriori") that corresponds to a 94% accuracy threshold for the a priori classifier. Using exclusively RSIIs would not only match the 66.5% accuracy offered by a spectral only classifier, but would even surpass it (69.76% for the RSII-based neural network and 70.93% for the RSII-based threshold classifier). Considering the distribution of the white pixels in Fig. 8, it can be inferred that the quality of spectral information is low due to impervious and soil confusion.

ii) Integration of spectral and RSII-derived information did provide significant improvements when compared to spectral-only classifiers. The improvements were not as pronounced when contrasted with RSII-only classifiers, which

Fig. 8. Contribution of each classifier for a 94% accuracy threshold.

leads to the same conclusion as above, namely that spectral confusion prevails in almost all accuracy thresholds.

iii) Looking specifically into the two RSII-only classifiers, the simple one dimensional classifier slightly surpassed the five dimensional neural network classifier. This could be attributed to training overfitting of the neural network and lack of further expressiveness of the additional four RSIIs. It may also be caused by the different road network representation in the calibration and validation datasets. However, further testing is necessary to derive more concrete conclusions.

4.2. Evaluation of the multi-process framework on the entire validation dataset

In this section, the classification results of the a priori, the dilation and the a posteriori classifiers were integrated to further study the performance of the RSIIs. The statistical evaluation was based on the entire validation dataset and it was followed by a visual inspection.

4.2.1. Statistical assessment

The statistical assessment compared five methods. The first one, named single neural network (NN), used the selected a priori classifier directly on the entire validation dataset (i.e. no dilation or a posteriori classifiers were implemented). The other four options used the same a priori algorithm and directional dilation operation with the same accuracy threshold (92%), but with different a posteriori classifiers as defined in Section 4.1.2.

The classification accuracies using the five methods are summarized in Table 1. Comparing with the single neural network, better overall accuracies and kappa statistics were achieved by adopting the multi-process classifier. Within the classification results employing multi-process classifiers, the majority of pixels were identified in a priori classification (57.43%) and a posteriori classification (37.61%), with only about 5% of all the pixels were classified using the directional dilation operation. The dataset portion corresponding to the a priori classifier and the directional dilation process was classified at relatively higher accuracy levels (88.07% and 87.67% respectively) compared to all a posteriori classifiers performance (app. 63–71%). This is attributed to the higher spectral confusion present in the leftover dataset forwarded to the a posteriori classifier. The three classification methods involving RSIIs as the input information provided similar overall accuracies and kappa statistics, and all surpassed the classification results using only spectral inputs. This was consistent with the classification results of the dataset portion corresponding to the a posteriori classifier (Section 4.1.2). As it can be expected the improvements in overall accuracy metrics depend not only on the algorithmic performance of the a posteriori classifier but also on the portion on the dataset forwarded to it; higher overall gains are expected as the classification difficulty increases, a significant benefit of the proposed approach.

4.2.2. Visual assessment

Fig. 9 demonstrates the final binary classification maps using the five classification methods from Table 1. Spatial distribution of different classification results was inspected and compared. The visual maps illustrate two major benefits using RSIIs:

i) The utilization of RSIIs successfully solved the common misclassification problem between soil and impervious pixels, a result of their similar spectral characteristics. This finding is in accordance with prior work (Luo & Mountrakis, 2010). As shown in the triangle in Fig. 9, there is a misclassification of non-impervious pixels into the impervious class that occurred between roads using the single neural network model (Fig. 9 (a)) and using only spectral inputs in the posteriori classifier (Fig. 9 (b)). Most of the misclassified pixels have been corrected in RSII-based classification maps (Fig. 9 (c), (d) and (e)).

Table 1
The classification accuracies of different classification models on validation dataset.

	Single-process	Multi-process classifier					
	Single NN	A priori	Dilation	A posteriori (select one)			
				Spectral inputs (NN)	RSII 3 (threshold-based)	All RSIIs (NN)	Spectral + all RSIIs (NN)
Dataset percentage (%)	100	57.43	4.96	37.61			
Accuracy within corresponding dataset	N/A	88.07	87.67	62.94	70.90	70.86	71.39
Overall accuracy (%)	77.77	–	–	78.91	81.91	81.89	82.09
Kappa statistic (%)	55.53	–	–	57.82	63.81	63.78	64.18

Note: Dataset percentage reflects the portion of the entire dataset classified by a given classifier. Overall accuracy and Kappa statistic reflect performance on the entire dataset.

ii) By employing the RSIIs, the general structure of the road networks has been approximately captured (Fig. 9 (c), (d) and (e)). The road breaks were connected and the gaps between road segments were filled. The road classification was particularly accurate when the road lines were straight and the road network structure appeared simple (i.e., the roads were regularly distributed). Even in difficult situations where roads would meet in unusual angles some of them were successfully detected (two elongated rectangles in Fig. 9). Further studies are necessary though to expand this method in curved roads.

Fig. 9. Binary impervious maps using different classification models.

However, the visualized maps also show several problems using RSIIs. Firstly, the impervious pixel classification relied closely on the detected road linear features. As shown in the circles in Fig. 9, it is clear that the wrong detection of road segments in RSII-based classification models (Fig. 9 (c), (d) and (e)) aggravated the misclassification of impervious pixels that already existed in the impervious maps derived by non RSII-based classification models. Second, the roads in the impervious maps using exclusively RSIIs (Fig. 9 (c) and (d)) appeared with high thickness due to the overestimation of road pixels across the road direction. The overestimation was introduced mostly because the classification was derived based on road structural information from limited partially classified results. This overestimation was reduced when the classification integrated with additional spectral information (Fig. 9 (e)).

5. Discussion and conclusions

The initial purpose of this paper was to examine whether road-based structural inputs from an intermediate classification can supplement spectral information and improve classification results. For the most efficient a priori classifier (using the 92% accuracy threshold) the neural network a posteriori classifier increased accuracy from 63% to 71.4% for the associated portion of the dataset. This also led to an improvement over the entire dataset from 78.9% to 82.1% (+6% in terms of Kappa). Visual inspection indicated their applicability in typical cases of spectral confusion, especially their ability to address the common misclassification between soil and impervious pixels.

A more far-reaching hypothesis was the investigation as to whether contextual information as expressed through the RSIIs could completely replace spectral information in portions of the image using a posteriori classifiers. Atmospheric effects, topography, and sensor malfunctions often degrade considerably the quality of spectral information. The issue of absence of spectral information for classification purposes has not been adequately examined, mostly due to the associated difficulty. In this paper, two different scenarios were studied. Firstly, an approach similar to that discussed above was incorporated where a neural network-based algorithm was applied, but this time using exclusively all five RSIIs (thus ignoring spectral information). The superiority of an RSII-supported neural network was clear over a spectral-supported neural network for all accuracy thresholds applied on the a priori classifier (i.e. for different proportions of the dataset forwarded to the a posteriori classifier). This could be expected in small accuracy thresholds (e.g. 84% to 88%), considering that the leftover pixels classified here are some of the most challenging cases; it made sense that spectral confusion would be high, therefore rewarding an RSII-supported method. The RSII-supported superiority continued though in higher accuracy thresholds; for example, for a 94% threshold a 3–4% advantage was present. This is especially encouraging considering that more than half of the entire validation dataset was forwarded to the a posteriori classifier. In other words, completely eliminating spectral information and replacing it with contextual information for half of the dataset would not only provide equivalent results, it would actually improve them.

The aforementioned evaluation implemented a five RSII input using a complex neural network approach. The extreme alternative was tested, namely a one RSII input using a simple binary threshold approach. The underlying motivation was to investigate the power of individual RSIIs along with any over-generalization issues that may be associated with complex algorithms such as neural networks. The RSIIs sensitivity analysis suggested that RSII 2 (the distance from an unclassified pixel to its closest endpoint) and RSII 3 (the angle between the line connecting the unclassified pixel with its closest endpoint pixel and the line segment direction on which the endpoint pixel was located) were more prominent inputs. RSII 3 was tested using a simple binary threshold rule to classify the remaining pixels

after the a priori classification. The obtained results were equivalent (on occasion better) than those of the five RSII input neural network. This could be partially attributed to the higher grid-type road network on the validation dataset (see Fig. 1) allowing a single RSII input to capture that variability. Nonetheless, the results were encouraging to the degree that for practical implementation purposes and in the absence of spectral information, the binary threshold would be the preferred method, mostly due to its simplicity in RSII input creation and algorithmic training.

There are some limitations of the proposed methodology. The derivation of road structural information was strongly dependent on the road identification accuracy. Mistakenly detected line segments brought in overestimation of impervious pixels, while missed line segments had the opposite effect. These errors resulted from two sources: misclassification of the partial results from the a priori classifier, and errors in the road identification process using the iterative localized Radon transform algorithm. Both algorithms could be further optimized in the future; for example, the backpropagation neural network could be replaced by a more transparent algorithm such as decision tree, or a potentially stronger classifier such as a support vector classifier, especially in the presence of limited samples. The iterative localized Radon transform could also be improved through peak value profile analysis and dynamic identification of optimal density threshold. Furthermore, other road detection processes could be implemented depending on the site specifics, type of imagery and user expertise. In addition, the proposed methodology is designed for rural or semi urban environment; it is unknown how it would perform in highly developed urban areas. Therefore, a moving filtering window is suggested for an approximate estimation of impervious density before application of this method. This can easily be incorporated within the local subimage iterative Radon analysis: for example, through the selective application of the method where the a priori classifier indicates an impervious presence of less than a given threshold.

This paper acted as the proof of concept for replacement of spectral with contextual information, in portions of the image. In future work, the concept of RSIIs could be implemented in remotely sensed images of higher than the examined 30 m resolution, because additional RSII dependencies may come to the surface which are currently masked at our working scale. Higher resolution imagery could facilitate more accurate road detection, especially for roads not wide enough to register a strong spectral response. Furthermore, the methodology could expand to other feature identification problems, assuming an appropriately adjusted feature detection method; one example is river classification in the presence of dense vegetation. Lastly, the presented work involved backpropagation neural networks as the core classification methodology. The approach is flexible and it easily supports replacement with other classification methods (e.g. decision trees and support vector machines), therefore accommodating user expertise and preference.

Appendix A. Supplementary data

Supplementary data to this article can be found online at doi:10.1016/j.rse.2010.12.018.

References

Amini, J. (2009). Road extraction from satellite image using a fuzzy-snake model. *The Cartographic Journal*, *46*, 164−172.

Arnold, C. L., & Gibbons, C. J. (1996). Impervious surface: The emergence of a key urban environmental indicator. *Journal of the American Planning Association, 62*, 243−258.

Bauer, M., Loeffelholz, B., & Wilson, B. (2005). Estimation, mapping and change analysis of impervious surface area by Landsat remote sensing. *Proceedings of Pecora 16 conference, ASPRS annual conference, Oct. 23–27, Sioux Falls, SD.*

Bentabet, L., Jodouin, S., Ziou, D., & Vaillancourt, J. (2003). Road vectors update using SAR imagery: a snake-based method. *IEEE Transactions on Geoscience and Remote Sensing, 41*, 1785−1803.

Carlson, T. N. (2003). Applications of remote sensing to urban problems. *Remote Sensing of Environment, 86*, 273−274.

Copeland, A. C., Ravichandran, G., & Trivedi, M. M. (1995). Localized Radon transform-based detection of ship wakes in SAR images. *IEEE Transactions on Geoscience and Remote Sensing, 33*, 35−45.

Elmore, A. J., & Guinn, S. M. (2010). Synergistic use of Landsat Multispectral Scanner with GIRAS land-cover data to retrieve impervious surface area for the Potomac River Basin in 1975. *Remote Sensing of Environment, 114*, 2384−2391.

Esch, T., Himmler, V., Schorcht, G., Thiel, M., Wehrmann, T., Bachofer, F., et al. (2009). Large-area assessment of impervious surface based on integrated analysis of single-date Landsat-7 images and geospatial vector data. *Remote Sensing of Environment, 113*, 1678−1690.

Forster, B. C. (1980). Urban residential ground cover using Landsat digital data. *Photogrammetric Engineering and Remote Sensing, 46*, 547−558.

Franke, J., Roberts, D. A., Halligan, K., & Menz, G. (2009). Hierarchical multiple endmember spectral mixture analysis (MESMA) of hyperspectral imagery for urban environments. *Remote Sensing of Environment, 113*, 1712−1723.

Gibson, L. (2003). Finding road networks in Ikonos satellite imagery. *Proceedings of ASPRS 2003 Conference, May 5–9, Anchorage, Alaska (CD-ROM).*

Gillies, R. P., Symanzik, J. B., & Rodemaker, E. J. (2003). Effects of urbanization on the aquatic fauna of the Line Creek watershed, Atlanta−A satellite perspective. *Remote Sensing of Environment, 86*, 411−422.

Gong, P., & Howarth, P. J. (1990). The use of structural information for improving land-cover classification accuracies at the rural–urban fringe. *Photogrammetric Engineering and Remote Sensing, 56*, 67−73.

Gruen, A., & Li, H. (1997). Semiautomatic linear feature extraction by dynamic programming and LSB-snakes. *Photogrammetric Engineering and Remote Sensing, 63*, 985−995.

Guindon, B., Zhang, Y., & Dillabaugh, C. (2004). Landsat urban mapping based on a combined spectral–spatial methodology. *Remote Sensing of Environment, 92*, 218−232.

Herold, N. (2003). Mapping impervious surfaces and forest canopy using classification and regression tree (CART) analysis. *Proceedings of 2003 ASPRS annual convention, May 5–9, Anchorage, Alaska.*

Hu, X., & Weng, Q. (2009). Estimating impervious surfaces from medium spatial resolution imagery using the self-organizing map and multi-layer perceptron neural networks. *Remote Sensing of Environment, 113*, 2089−2102.

Kim, T., Park, S., Kim, M., Jeong, S., & Kim, K. (2004). Tracking road centerlines from high resolution remote sensing images by least squares correlation matching. *Photogrammetric Engineering and Remote Sensing, 70*, 1417−1422.

Laptev, I., Mayer, H., Lindeberg, T., Eckstein, W., Steger, C., & Baumgartner, A. (2000). Automatic extraction of roads from aerial images based on scale space and snakes. *Machine Vision and Applications, 12*, 23−31.

Lee, S., & Lathrop, R. G. (2006). Subpixel analysis of Landsat ETM+ using self-organizing map (SOM) neural networks for urban land cover characterization. *IEEE Transactions on Geoscience and Remote Sensing, 70*, 963−971.

Lee, T. H., & Moon, W. M. (2002). Lineament extraction from Landsat TM, JERS-1, and EDM for geological applications. *Proceedings of International Geoscience and Remote Sensing Symposium, Jun. 24–28, Toronto, Ontario, Canada* (pp. 3276−3278).

Letitia, S. (2010). Adaptive structuring element in fuzzy morphology for automatic extraction of urban road network from high resolution aerial images. *International Journal of Engineering Science and Technology, 2*, 4182−4191.

Li, X., Zhang, S., Pan, X., Dale, P., & Cropp, R. (2010). Straight road edge detection from high-resolution remote sensing images based on the ridgelet transform with the revised parallel-beam Radon transform. *International Journal of Remote Sensing, 31*, 5041−5059.

Lu, D., & Weng, Q. (2006). Spectral mixture analysis of ASTER images for examining the relationship between urban thermal features and biophysical descriptors in Indianapolis, Indiana, USA. *Remote Sensing of Environment, 104*, 157−167.

Luo, L., & Mountrakis, G. (2010). Integrating intermediate inputs from partially classified images within a hybrid classification framework: An impervious surface estimation example. *Remote Sensing of Environment, 114*, 1220−1229.

Luo, L., & Mountrakis, G. (in press). A multi-process model of adaptable complexity for impervious surface estimation. International Journal of Remote Sensing.

Murphy, L. M. (1986). Linear feature detection and enhancement in noisy images via the Radon transform. *Pattern Recognition Letters, 4*, 279−284.

Powell, R. L., Roberts, D. A., Dennison, P. E., & Hess, L. L. (2007). Sub-pixel mapping of urban land cover using multiple endmember spectral mixture analysis: Manaus, Brazil. *Remote Sensing of Environment, 106*, 253−267.

Radon, J. (1917). Über die Bestimmung von Funktionen durch ihre Integralwerte längs gewisser Mannigfaltigkeiten. *Berichte über die Verhandlungen der Sächsische Akademie der Wissenschaften (Reports on the proceedings of the Saxony Academy of Science), 69* (pp. 262−277).

Shi, W., & Zhu, C. (2002). The line segment match method for extracting road network from high-resolution satellite images. *IEEE Transactions on Geoscience and Remote Sensing, 40*, 511−514.

Thomas, K. E., & Endreny, T. A. (2008). Improving national land cover database estimates of road network impervious cover using vector road networks in GIS. *Surveying and Land Information Science, 68*, 21−27.

Valero, S., Chanussot, J., Benediktsson, J. A., Talbot, H., & Waske, B. (2010). Advanced directional mathematical morphology for the detection of the road network in very high resolution remote sensing images. *Pattern Recognition Letters, 31*, 1120−1127.

Wang, J. (1995). Application of the linear feature detection system−LINDA to image segmentation from remotely sensed data. *Remote Sensing Reviews, 13*, 49−66.

Weng, Q. (2007). *Remote sensing of impervious surfaces.* Boca Raton, Florida: CRC Press.

Wu, C., & Murray, A. T. (2003). Estimating impervious surface distribution by spectral mixture analysis. *Remote Sensing of Environment, 84*, 493−505.

Xian, G., & Crane, M. (2005). Assessments of urban growth in the Tampa Bay watershed using remote sensing data. *Remote Sensing of Environment, 97*, 203−215.

Xian, G., Crane, M., & McMahon, C. (2008). Quantifying multi-temporal urban development characteristics in Las Vegas from Landsat and ASTER data. *Photogrammetric Engineering and Remote Sensing, 74*, 473−481.

Yang, J. S., & Artigas, F. J. (2008). Estimating impervious surfaces area of urban watershed using ASTER data. *Journal of Environmental Informatics, 12*, 1−8.

Yang, L., Huang, C., Homer, C. G., Wylie, B. K., & Coan, M. J. (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*, 230−240.

Zhang, Q., & Couloigner, I. (2007). Accurate centerline detection and line width estimation of thick lines using the Radon transform. *IEEE Transactions on Image Processing, 16*, 310−316.

Zhang, Q., Wang, J., Peng, X., Gong, P., & Shi, P. (2002). Urban built-up land change detection with road density and spectral information from multi-temporal Landsat TM data. *International Journal of Remote Sensing, 23*, 3057−3078.

Zhu, C., Shi, W., Pesaresi, M., Liu, L., Chen, X., & King, B. (2005). The recognition of road network from high-resolution satellite remotely sensed data using image morphological characteristics. *International Journal of Remote Sensing, 26*, 5493−5508.