# ACADEMIA

Accelerating the world's research.

# Converting local spectral and spatial information from a priori classifiers into contextual knowledge for imperv...

Giorgos Mountrakis

*ISPRS Journal of Photogrammetry and Remote …*

**Cite this paper**

Get the citation in MLA, APA, or Chicago styles

**Related papers**

A Comprehensive Review on Pixel Oriented and Object Oriented Methods for Information Extr…
Shridhar Jawak

Optimal Band Configuration for the Roof Surface Characterization Using Hyperspectral and LiDAR Ima…
Anna Jarocińska, Bogdan Zagajewski, Prakash Nimbalkar

Assessing the performance of two unsupervised dimensionality reduction techniques on hyperspect…
F. Canters
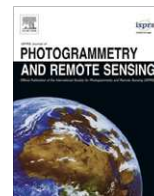
# Converting local spectral and spatial information from a priori classifiers into contextual knowledge for impervious surface classification

Li Luo, Giorgos Mountrakis *

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

## ABSTRACT

A classification model was demonstrated that explored spectral and spatial contextual information from previously classified neighbors to improve classification of remaining unclassified pixels. The classification was composed by two major steps, the a priori and the a posteriori classifications. The a priori algorithm classified the less difficult image portion. The a posteriori classifier operated on the more challenging image parts and strived to enhance accuracy by converting classified information from the a priori process into specific knowledge. The novelty of this work relies on the substitution of image-wide information with local spectral representations and spatial correlations, in essence classifying each pixel using exclusively neighboring behavior. Furthermore, the a posteriori classifier is a simple and intuitive algorithm, adjusted to perform in a localized setting for the task requirements. A 2001 and a 2006 Landsat scene from Central New York were used to assess the performance on an impervious classification task. The proposed method was compared with a back propagation neural network. Kappa statistic values in the corresponding applicable datasets increased from 18.67 to 24.05 for the 2006 scene, and from 22.92 to 35.76 for the 2001 scene classification, mostly correcting misclassifications between impervious and soil pixels. This finding suggests that simple classifiers have the ability to surpass complex classifiers through incorporation of partial results and an elegant multi-process framework.

© 2011 International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS). Published by Elsevier B.V. All rights reserved.

## 1. Introduction

Impervious surface detection is an important topic in remote sensing applications. A variety of research has been conducted on classification methods for imperviousness estimation based on remotely sensed images (Weng, 2007). Examples of these methods include multivariate regression (Bauer et al., 2005; Yang, 2006), spectral mixture (Wu and Murray, 2003; Lu and Weng, 2006; Powell et al., 2007; Franke et al., 2009) and machine learning models (Herold, 2003; Yang et al., 2003; Dougherty et al., 2004; Lee and Lathrop, 2006; Mohapatra and Wu, 2008; Esch et al., 2009; Hu and Weng, 2009; Mountrakis et al., 2011).

However, the majority of the imperviousness classification methods are pixel-based and do not consider spatial and contextual information from neighboring pixels that may improve classification accuracy. Methods that take into account labeling of neighbors when seeking to determine the most appropriate class for a pixel are said to be context sensitive, or simply context clas-

sifiers (Richards and Jia, 2006). Existing context classifiers can be usually summarized in four categories (Richards and Jia, 2006): (1) *Preprocessing:* In this method, the image is preprocessed via spatial filters or more advanced texture analysis algorithms before classification takes place in order to extract spatial features (e.g. Gong and Howarth, 1990, 1992; Binaghi et al., 2003). Typically, local spectral and spatial information is used to divide the image into a number of homogeneous objects composed of adjacent pixels with the similar characteristics (i.e. employ an image segmentation). Early on, Kettig and Landgrebe (1976) presented the geo-object based classification method (OBCM) through extraction of homogeneous objects before classification. A series of studies followed to explore further OBCM and have been proved successful in a number of recent applications (e.g. Blaschke and Hay, 2001; Benz et al., 2004; Hay and Castilla, 2008; Johansen et al., 2010; Lizarazo and Barros, 2010). (2) *Postprocessing:* Instead of processing an image before classification, a post-classification filtering method performs spatial context analysis on the classification results. Local spectral and spatial information is integrated by examining the labeling of neighboring pixels in the intermediate classification map using a spatial mask. The label of the center pixels within the spatial mask might be changed to the label most represented in the spatial mask (Townsend, 1986; Barnsley and Barr, 1996; Kim, 1996; De Voorde et al., 2007; Chormanski et al., 2008; Mas et al.,

* Corresponding author. Address: Intelligent Geocomputing Laboratory, Department 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).

2010). (3) *Probabilistic label relaxation:* In this method, local spectral and spatial information is integrated in the form of probabilities within a spatial mask for each pixel. The probabilities of each pixel are adjusted according to the labeling of neighboring pixels within a spatial mask and are used in the classification (Gong and Howarth, 1989; Richards and Jia, 2007; Yi et al., 2007; Reigber et al., 2010). (4) *Markov random fields* (Dubes and Jain, 1989; Khedama and Belhadj-Aissa, 2004; Tso and Olsen, 2005; Tolpekin and Stein, 2009): Similar to the probabilistic label relaxation method, the Markov random fields method integrates local spectral and spatial information by looking into the probabilities of adjacent pixels. However, this method measures the prior probability using Bayes' theorem to maximize the global posterior probability in order to incorporate the spatial context information.

The classification model adopted in this paper also incorporated contextual information. The process was divided into a priori and a posteriori classifications, where classified pixels from the a priori method assisted in the classification of the leftover pixels handled by the a posteriori method. Distinct from previous context classification methods, the pixels classified through the a posteriori classifier were neither previously labeled nor had a probability assigned to them, instead they were classified on the fly. Furthermore, a significant difference between this paper and our prior work (Luo and Mountrakis, 2010) is that this model integrated spectral and spatial information exclusively from a local neighborhood to identify unclassified pixels, while all other surrounding pixels at larger scale were ignored. The working hypothesis is that already classified neighboring pixels will contain enough class information of spectral reflectance (similar materials) and spatial structure. Also, by ignoring information from the entire scene we will limit misclassifications. Part of our investigation focused on neighborhood type and size identification.

In order to derive partially classified results, a hybrid classification structure comprised of a series of steps was proposed. Hybrid classifiers have demonstrated potential for higher classification accuracy over single classifiers (Steele, 2000; Liu et al., 2002, 2004; Coe et al., 2005; Mountrakis et al., 2009; Franke et al., 2009). The objective of this research was to improve impervious surface classification accuracy by integrating contextual information in a hybrid classification model. To assess the performance of this novel classification model, two Landsat images from 2001 and 2006 covering central New York were used, respectively.

## 2. Methodology

A hybrid multi-process classification model that integrated multiple classifiers was used as the classification model for this paper (Fig. 1). It was a progressive process comprised of multiple steps. In each step, parts of the dataset were classified while the remaining portions of the dataset were forwarded to subsequent steps. Initially, an a priori classifier was used to derive partially classified results. After the partially classified results were processed by a majority filter, an a posteriori classifier was implemented to identify the remaining unclassified pixels. This a posteriori classifier integrated spectral and spatial information of the partial classification results from the previous two steps (a priori and majority filter) as contextual information. This paper focuses on the a posteriori classifier and assesses potential benefits and tradeoffs of the method.

### 2.1. A priori classifier

The purpose of the a priori classifier was to produce a partially classified image that would act as the basis for subsequent classification steps. Any classification algorithm could be used as the a
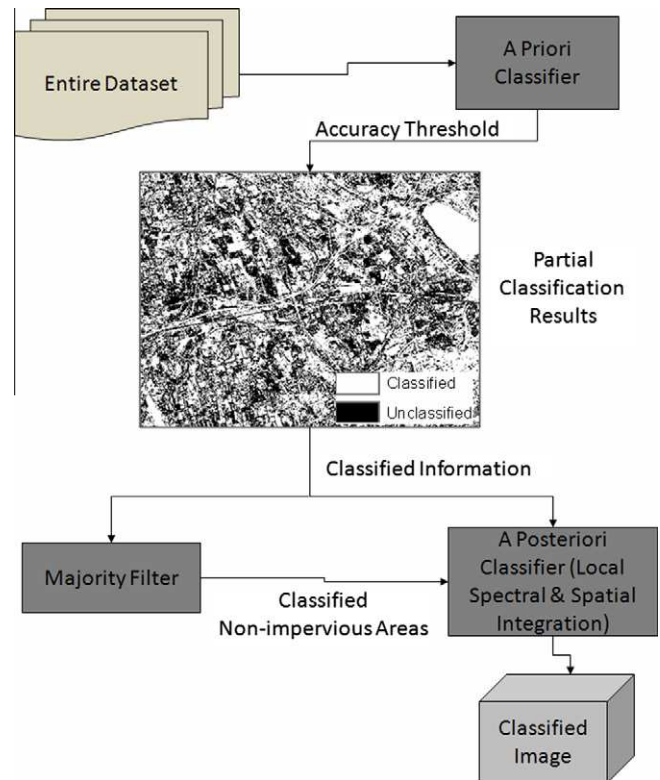


**Fig. 1.** The framework of the hybrid multi-process classification.

priori classifier, the only constraint was that their classification results should be able to provide a continuous range of accuracies. To determine the portion of partial classification results, a threshold was applied on the a priori classifier to ensure that the extracted pixels were classified with certain accuracy. The accuracy threshold was set up based on the calibration dataset and balanced the tradeoff of sufficient yet accurate partial results. Higher amount of already classified pixels (partial results) could reveal additional contextual information for later steps. However, as the classified pixel amount increased the classification accuracy decreased which could lead to additional but erroneous contextual information for the subsequent classifiers. Multiple thresholds were tested to identify the optimal for a given image problem, a process investigated in prior work (Luo and Mountrakis, 2010; Mountrakis and Luo, 2011).

In this experiment, the multi-layer perceptron feed forward neural network structure trained with a back-propagation learning algorithm was adopted as the a priori classifier. One thousand different neural network architectures were trained using the Levenberg–Marquardt backpropagation learning algorithm and the one with the best overall accuracy on the calibration dataset was identified (more on training datasets in Section 2.4). The input layer contained six nodes corresponding to the six Landsat ETM+ image bands (blue, green, red, near IR and two mid IR bands). Each node at the output layer represented one class (therefore a total of two nodes for impervious and non-impervious class, respectively) and was comprised of a logistic function. The range of each output node was continuous between 0 and 1. The node number in hidden layers was randomly selected during the training process of the 1000 architectures: between 6 and 15 for the first hidden layer and from 0 to 9 for the second hidden layer. The activation functions for the hidden layers were tangent sigmoidal functions. In order to translate a predetermined accuracy threshold to an output node

threshold of the selected best neural network, each node response was progressively lowered. The calibration dataset was scanned and the process was stopped when the accuracy threshold was met. A pixel was assigned to a given class when its response satisfied the output node threshold. Pixels satisfying both thresholds were assigned to the node with the strongest response. Those pixels that did not satisfy the node threshold of either class were forwarded to the next classifier.

## 2.2. Majority filter

After the implementation of the a priori classifier at a certain accuracy threshold, all pixels were categorized in three classes: impervious, non-impervious and unclassified. It was noticed that most non-impervious surface pixels in the reference dataset were grouped into homogeneous areas (e.g. forest or agricultural patches). Since adjacent pixels within these non-impervious areas exhibited spectral similarities, the non-impervious pixels identified from partially classified results provided contextual information for the classification of their neighboring pixels. In order to increase the contextual information for the leftover unclassified pixels, a majority filter was adopted targeting solely the non-impervious pixels. In the majority filter application, if all eight neighbors of an unclassified pixel were occupied by non-impervious pixels, this pixel was switched from the unclassified category to the non-impervious class. The process was repeated until no additional pixels were extracted. The majority filter was not applied on the opposite class of impervious pixels because the size of homogeneous regions of impervious class was variable compared to non-impervious regions. We should note that the majority filter was a simple processing step, the novelty of the proposed method lied on the a posteriori classifier and its localized application, presented below.

## 2.3. A posteriori classifier

At this stage, each pixel (i) was classified as impervious/non-impervious from the a priori classifier, (ii) was identified as non-impervious from the additional majority step or (iii) remained unclassified. The principle behind the a posteriori classifier was to use exclusively local information in order to label the remaining unclassified pixels. Two significant distinctions from our prior work (Luo and Mountrakis, 2010) and this work are that this paper investigates the incorporation of only local information and also adapts a much simpler classifier that supports a more wide-range adoption of the methodology. The adjusted minimum distance to means classification algorithm (De Jong et al., 2001) was incorporated as the a posteriori classifier to integrate local spectral and spatial information. Traditional minimum distance to mean classification calculates the Euclidean distance between each unknown pixel and mean vector for each class of the training data (Jahne, 2005). The unknown pixel was assigned to the class with the shortest distance. De Jong et al. (2001) extended the minimum distance to mean classification algorithm by introducing spatial distances which described the spatial proximity of an unclassified pixel to a homogeneous classified region. A weighting scheme was adopted to describe relative impact of spatial and spectral proximity.

In our classification model, distances were calculated as:

$$Dist^* = \min\left(\alpha \times Dist_k^{Spectral} + (1-\alpha) \times Dist_k^{Spatial}\right) \quad (1)$$

where $Dist^*$ is the minimum integrated distance among all the $k$ classes. $Dist_k^{Spectral}$ and $Dist_k^{Spatial}$ are the local spectral and spatial distances to $k$th class and $\alpha$ is the ratio factor used to describe the relative contribution between spectral and spatial information.

To calculate local spectral and spatial distances, a spatial proximity rule was applied to determine the neighboring pixels around an unknown pixel (e.g. a rectangular neighborhood, or a predetermined number of closest neighbors). For each class, we extracted spatial and spectral information of the already labeled pixels. Using the above equation, a distance $Dist^*$ was calculated for each class and the unclassified pixel under consideration was assigned to the class providing the smallest distance value.

More specifically, the spectral distance was calculated by the Euclidean distance between an unclassified pixel and the spectral signature of each already labeled pixel of a given class. By doing so, we spectrally compared class signatures only from the local area hypothesizing similar ground feature materials and therefore spectral responses. The local spatial distance for each class was the average distance between the unclassified pixel and all classified pixels from the same class. It should be noted that the local spectral and spatial distances were normalized using the maximum spectral and the spatial values to ensure no bias toward either spectral or spatial distances.

In order to test the effects of different neighborhood sizes (masks) of the unclassified pixels, two strategies were adopted: fixed and adaptive masks. In the case of a fixed mask, a predetermined square neighborhood was searched (e.g. a $5 \times 5$ area around an unclassified pixel). If there was no classified pixel within the neighborhood, the unclassified pixel was randomly assigned to impervious or non-impervious class. Since the number of such unclassified pixels was limited, the classification accuracies were not significantly affected. For the adaptive mask, the algorithm would start from the unclassified pixel and continue expanding the search area and progressively look into further away pixels until a predetermined number of classified pixels was incorporated in the calculations.

## 2.4. Training dataset

A Landsat 7 Enhanced Thematic Mapper Plus (ETM+) scene acquired on 18 April, 2006 was used to tackle the problem of binary impervious surface classification. This Landsat ETM+ image covered central New York and was comprised of $5884 \times 5661$ pixels. The spatial resolution was 30 m and six bands were used as inputs, including blue, green, red, near IR and two mid IR bands.

The reference data were produced using aerial digital orthophoto quarter quads (DOQQ) imagery acquired in 2006 with a spatial resolution of 0.6 m. On-screen digitizing was performed on a sampling subregion. If any portion of the 30 m pixel occupied any constructed impervious surface (e.g. roads, houses, parking lots), that pixel was assigned to the impervious class. The zero-tolerance binary classification was adopted as the purpose of this work is not to provide direct imperviousness estimates, but rather act as an efficient and accurate filtering step. By doing so, any subsequently applied subpixel algorithms would have the opportunity to focus on pixels with known imperviousness presence. This is expected to offer significant advantages as imperviousness overestimation is an issue for large scale subpixel classifiers, especially within rural areas (Yang et al., 2003; Homer et al., 2004). The aforementioned subpixel classifiers frequently assign imperviousness percentages larger than 0% where no impervious surfaces exist. This is not a significant statistical error (e.g. assigning 7% instead of 0%), however it has been shown to be quite significant for environmental and ecological applications where minor imperviousness changes in the low imperviousness range have considerable impacts (Klein, 1979; Forman and Alexander, 1998; Paul and Meyer, 2001).

The reference dataset was collected using a stratified sampling method on the digitized region in order to represent various types of impervious and non-impervious surfaces. A significantly high

number of 140,000 pixels was used as the overall reference dataset, including 70,000 impervious and 70,000 non-impervious pixels. The reference dataset was then divided into calibration dataset and validation dataset with a ratio of 3:7. The calibration dataset was used to train the classification model and the validation dataset was used to assess the classification performance.

The proposed multi-process classifier in this paper was comprised of three sequential algorithms (see Fig. 1): a priori classifier, majority filter and a posteriori classifier. As the classification progressed from one algorithm to the next, pixels were gradually classified. Therefore the entire reference dataset (and therefore the corresponding calibration and validation datasets) was classified sequentially by the three classifiers. The calibration dataset was only utilized in the training process of the single neural network (also used as the a priori classifier). The majority filter does not require any training, while the adjusted minimum distance to mean classifier (the a posteriori classifier) has only one parameter which was investigated further through a sensitivity analysis. Sections 3.1–3.3 present classification results on the validation dataset. Each section is associated with one of the sequential classifiers, namely the a priori classifier, the majority filter and the a posteriori classifier. Each section also presents results from the corresponding portions of the validation dataset as the dataset progressively moves through the three sequential classifiers. An assessment on the entire validation dataset is provided in Section 3.4 combining results from all three classifiers. The classification results using a single neural network model were used as benchmark to compare with the multi-process three classifier model. For consistency, the single neural network was exactly the same one adopted as the a priori classifier, therefore any performance differences are attributed to the majority filter and the a posteriori classifier.

The hybrid multi-process classification model was also applied on a second image to further investigate the applicability of the method. A Landsat 7 Enhanced Thematic Mapper Plus (ETM+) scene acquired on April 8th, 2001 was used for binary impervious surface classification. The 2001 Landsat ETM+ imagery with 30 m spatial resolution and six bands (blue, green, red, near IR and two mid IR bands) covered almost the same area as the 2006 imagery and comprised of 5991 × 5915 pixels. Reference data was collected from 19 sites throughout the entire scene. The detailed description of the reference dataset is available in Luo and Mountrakis (2010). The reference data contained 101,919 pixels in impervious class and 72,850 pixels in non-impervious class. These pixels were randomly divided into calibration dataset and validation dataset with a ratio of 3:7.

## 3. Experimental results

The proposed method was evaluated using a 2001 and a 2006 Landsat scene with corresponding reference data. For the 2006 scene a detailed presentation follows to allow in depth evaluation and replication of the approach (Sections 3.1–3.4). The 2001 scene was used to further evaluate the novelty of the work, namely the a posteriori classifier, in Section 3.5.

### 3.1. Classification results of the a priori classifier

The neural network with best overall accuracy on the entire calibration dataset from 1000 candidate neural networks with different architectures was adopted as (i) the a priori classifier in the multi-process hybrid algorithm, and (ii) the single neural network benchmark algorithm. In the a priori classifier case, the derived classification results were utilized as the partially classified results. The selected network contained 1 input layer, 2 hidden layers and 1 output layer, with 11 and 5 nodes for the two hidden layers,

respectively. The overall accuracy on the entire validation dataset using this 'best' neural network was 81.80% and the Kappa statistic was 63.60 (for detailed accuracy results see Table 3). Various accuracy thresholds were applied on the selected neural network to derive different partially classified results, with thresholds starting from 84% to 94%, with an increment of 2%. The selection of optimal accuracy threshold has been the focus of prior work (Luo and Mountrakis, 2010), here the focus is on the a posteriori classifier. The accuracy threshold was set larger than the overall accuracy acquired by the single neural network (81.80%) to ensure there were enough leftover pixels. In this experiment we selected the accuracy threshold of 92% since it produced the highest overall accuracy among all the accuracy thresholds when all three classifiers were merged on the entire validation dataset. The accuracy threshold was then translated into the threshold value of each output node resulting to 0.79 for the output node corresponding to impervious

**Table 1**
Classification accuracy for the a priori classifier in the corresponding validation subset.

| Class | Validation data | | | |
|---|---|---|---|---|
| | Impervious | Non-impervious | Total | User's (%) |
| *Classification* | | | | |
| Impervious | 28,672 | 2619 | 31,291 | 91.63 |
| Non-impervious | 2220 | 25,522 | 27,742 | 92.00 |
| Total | 30,892 | 28,141 | 59,033 | |
| Producer's (%) | 92.81 | 90.69 | | |
| Overall Accuracy (%) | 91.80 | | | |

**Table 2**
Comparison between a posteriori classifier and single neural network in the corresponding validation subset.

| Accuracies (%) | Single NN | A posteriori classifier | |
|---|---|---|---|
| | | Fixed mask | Adaptive mask |
| *Impervious* | | | |
| Producer's | 61.30 | 87.00 | 87.56 |
| User's | 64.87 | 71.20 | 71.16 |
| *Non-impervious* | | | |
| Producer's | 54.66 | 34.66 | 34.11 |
| User's | 45.59 | 58.96 | 59.63 |
| Overall accuracy | 61.30 | 68.68 | 68.85 |
| Kappa statistic | 18.67 | 23.95 | 24.05 |
| *Z-Score (P-value)* | | | |
| Single NN vs. fixed mask | 5.00 (<0.0001) | | |
| Single NN vs. adaptive mask | 4.98 (<0.0001) | | |

**Table 3**
Comparison between multi-process model and single neural network in the entire validation dataset.

| Accuracies (%) | Single NN | A posteriori classifier | |
|---|---|---|---|
| | | Fixed mask | Adaptive mask |
| *Impervious* | | | |
| Producer's | 80.38 | 83.71 | 83.97 |
| User's | 82.73 | 84.35 | 84.30 |
| *Non-impervious* | | | |
| Producer's | 83.22 | 84.47 | 84.38 |
| User's | 80.92 | 83.83 | 83.95 |
| Overall accuracy | 81.80 | 84.09 | 84.13 |
| Kappa statistic | 63.60 | 68.18 | 68.25 |
| *Z-Score (P-value)* | | | |
| Single NN vs. fixed mask | 10.13 (<0.0001) | | |
| Single NN vs. adaptive mask | 10.29 (<0.0001) | | |

(a) Overall accuracy using fixed mask method



(c) Overall accuracy using adaptive mask method



(b) Kappa statistics using fixed mask method
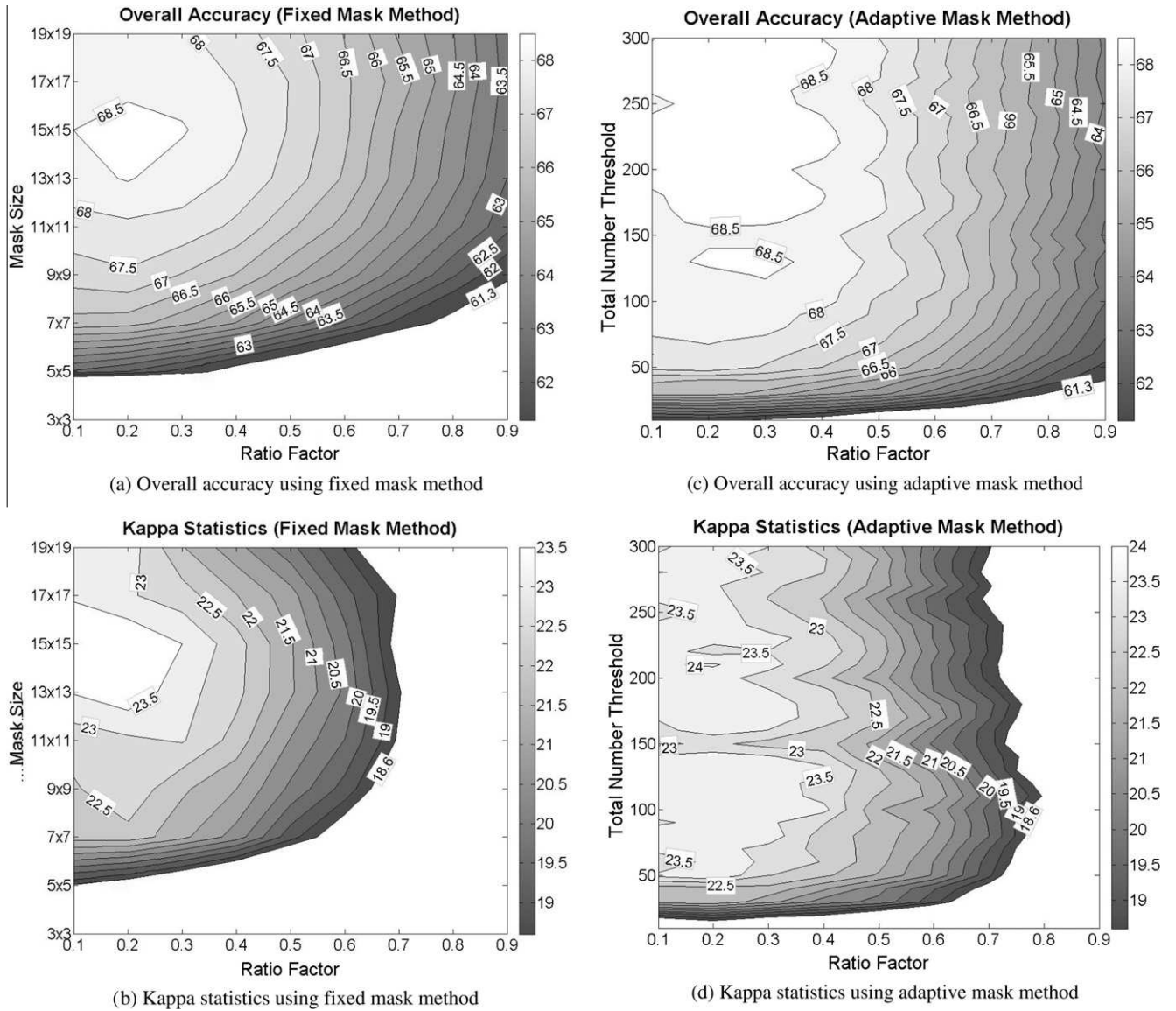


(d) Kappa statistics using adaptive mask method

**Fig. 2.** Contours of overall accuracies and Kappa statistics of classification results using a posteriori classifier.

class and 0.82 for the corresponding output node of non-impervious class. Therefore, pixels with response in the impervious output node larger than 0.79 were assigned as impervious class and the pixels with response larger than 0.82 in the non-impervious output node were directed to the non-impervious class. Pixels satisfying both thresholds were assigned to the class with the strongest response. The error matrix from the validation dataset is listed in Table 1. The validation dataset contained only pixels satisfying the pre-identified thresholds, an entire dataset comparison is provided in Section 3.4. Approximately 63.04% of all impervious pixels and 57.43% of pixels from non-impervious class in the entire validation dataset (98,000 pixels) were extracted.

### 3.2. Classification results of the majority filter

A 3 × 3 window was applied solely on unclassified pixels using the partial classification results (pixels classified in the above process using a priori classifier). All extracted pixels were assigned to the non-impervious class. The overall accuracy of the extracted pixels in this process was 77.13% using the multi-process classification model compared to 73.46% using the single neural network.

### 3.3. Classification results of the a posteriori classifier

Spectral and spatial information within local neighborhoods was integrated using the adjusted minimum distance to mean classification algorithm to identify all leftover pixels. Both the fixed and adaptive mask methods were tested for different neighborhoods sizes. A sensitivity analysis took place where the ratio factor $\alpha$ (see Eq. (1)) was assigned values from 0.1 to 0.9 with a step of 0.1 to assess the relative importance between spectral and spatial information. Fig. 2 illustrates the contours of overall accuracies and Kappa statistics obtained using fixed and adaptive masks with different spectral/spatial ratio factors. Numerous mask sizes were used for the fixed mask and multiple total point thresholds were incorporated for the adaptive mask. The masks used in fixed mask method were 3 × 3, 5 × 5, 7 × 7 up to 19 × 19, while the thresholds for total classified pixel number in the adaptive mask method was assigned between 10 and 300 with a step of 10. The accuracies of the benchmark neural network for the corresponding portion of the dataset were demonstrated as the lowest contour levels in order to emphasize improvements. The following conclusions can be extracted:

(i) Limited but necessary contribution of the spectral information as ratio values close to 0.2–0.3 provides the optimal accuracy. Higher ratio values did not perform as well suggesting spectral confusion that is resolved through the incorporation of spatial proximity information.

(ii) Incorporation of higher number of points in both the fixed and adaptive methods initially increased accuracy. However, there seems to be a saturation level beyond which improvements are not clear.

(iii) As hypothesized, both the fixed and adaptive methods provided significant improvements over the single neural network benchmark. Therefore, the incorporation of partially classified results is a plausible one.

Table 2 illustrates in further detail a comparison of classification results between the best fixed mask (mask size: 15 × 15; ratio factor: 0.2), the best adaptive mask (pixel total number threshold: 210; ratio factor: 0.2) and the single neural network (NN) benchmark. The improvements were substantial compared to the benchmark. The Z-test between Kappa statistics (Cohen, 1960) of the two models further indicated that improvements of the a posteriori classifiers were statistically significant compared to the single neural network classifier with a confidence level of 0.01. However, the producer's accuracy of non-impervious class was decreased using the a posteriori classifiers. It suggested that there was an overestimation of impervious pixels. Since the majority classifier identified most of the leftover non-impervious pixels from the a priori classification, only limited number of non-impervious pixels classified by the a posteriori classifier. Considering the improvement of the overall accuracies and accuracies of impervious class, the overestimation was acceptable. Note that all results reposted in this section apply on the portion of the validation dataset forwarded to the a posteriori classifier.

### 3.4. Classification results of the entire hybrid method

The overall accuracy and Kappa statistic of the entire dataset classification was contrasted for the multi-process models using fixed (mask size: 15 × 15; ratio factor: 0.2) and adaptive (pixel total number threshold: 210; ratio factor: 0.2) mask methods with the results using a single neural network (Table 3). The single neural network used here was the exact neural network adopted as the a priori classifier in order to stay consistent. The comparison illustrated that the multi-process model with incorporation of methods that integrated local spectral and spatial information based on partially classified results, improved impervious surface classification. The Z-test between Kappa statistics also indicated that improvements were statistically significant at the confidence level of 0.01. This is of particular importance because portions of the dataset were classified by the same algorithm (single neural network) in all three methods.

A visual assessment of algorithmic performance was also conducted. The thematic map of impervious surface within the entire Landsat ETM+ scene is shown in Fig. 3. This map was derived using the multi-process model with the adaptive mask method with a pixel total number threshold of 210 and a ratio factor of 0.2. A representative area was selected in order to further assess the performance of the multi-process model (rectangle in Fig. 3). Fig. 4 illustrates the classified image of this representative area using the multi-process model compared with the single neural network. Fig. 4d identifies the spatial footprints of each of the three processes of the multi-process models. Referring to the reference data from this area (Fig. 4a), the classified image derived from the single neural network contained a rather large quantity of misclassifications between roads (Fig. 4b). Most of the noise was caused by misclassification of soil pixels into the impervious class, a known
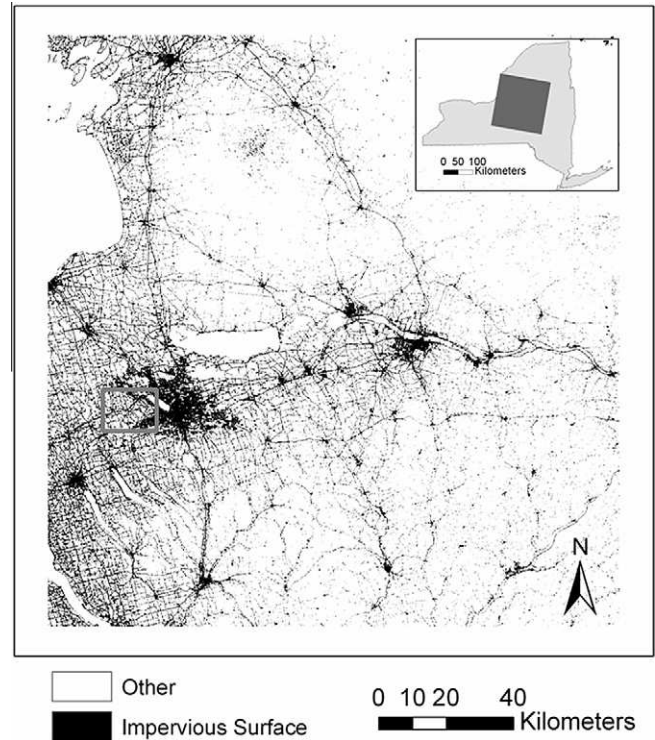


Fig. 3. Classification results of the entire study area.

problem for impervious extraction algorithms. The misclassification was caused by the similar spectral characteristics between soil and impervious pixels. The multi-process model removed most of the misclassified pixels with the additional spatial information from neighboring pixels (Fig. 4c). As a trade-off, there were a few road pixels missing; also, in some locations an overestimation of impervious surface pixels at the boundary between impervious areas and non-impervious areas was observed. This misclassification may be caused by the similar spectral features of pixels at the boundary and the lack of road pixels in the partially classified results. Despite of these misclassifications, the overall proposed approach of integrating local spectral and spatial information significantly improved the overall classification product.

### 3.5. Further evaluation of the a posteriori classifier on a 2001 scene

The classification process of the 2001 scene was similar to the process followed on the 2006 scene; it was a progressive process comprised of the a priori classifier, majority filter and the a posteriori classifier. The multi-layer perceptron feed forward neural network structure trained with a back-propagation learning algorithm was adopted as the a priori classifier and the adjusted minimum distance to mean classification algorithm was used as the a posteriori classifier. All the training processes of the a priori classifier and the a posteriori classifier were the same as the processes of the 2006 scene.

The accuracy thresholds of 92%, 94% and 96% were applied on the selected neural network (as the a priori classifier) to derive different partially classified results. Various spectral/spatial contribution ratio factors were tested during the a posteriori classification, starting from 0.1 to 0.9. Numerous mask sizes were also tested: masks used in the fixed mask process were 3 × 3, 5 × 5, 7 × 7 up to 19 × 19, while the thresholds for total classified pixel number in the adaptive mask method were assigned between 10 and 300 with a step of 10. This ensured consistency with the 2006 image.
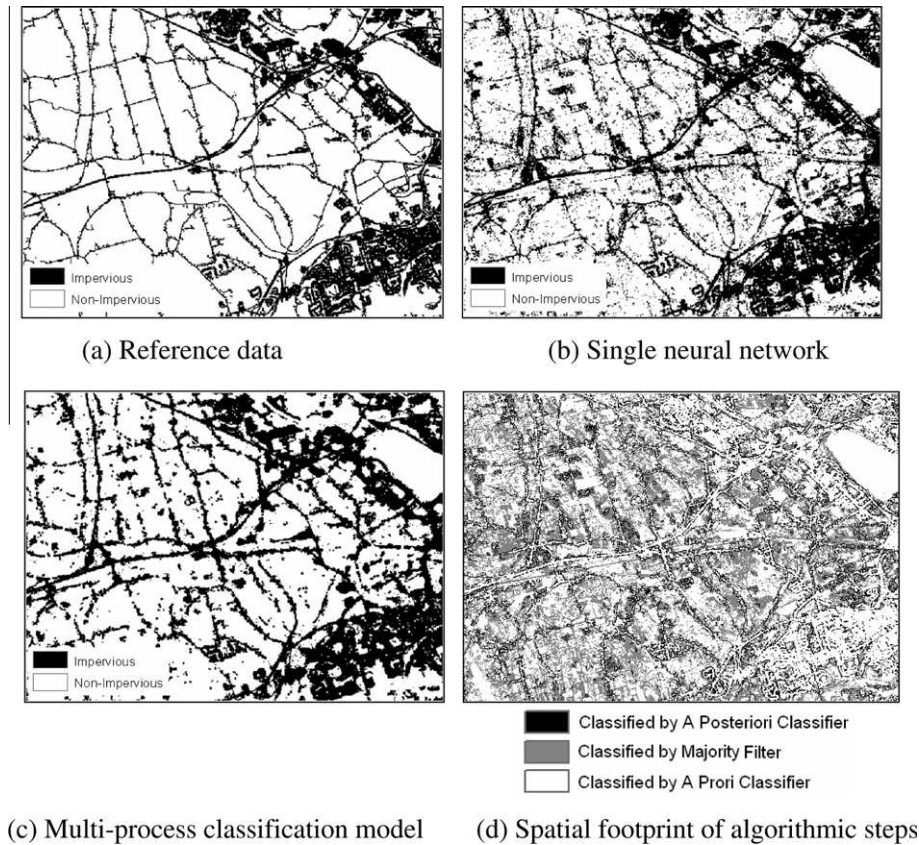
(a) Reference data

(b) Single neural network

(c) Multi-process classification model

(d) Spatial footprint of algorithmic steps

**Fig. 4.** Classification results of the selected representative area.

Among the three accuracy thresholds applied on the a priori classifier, the classification results with 96% achieved the best overall accuracy. Table 4 compares results between the best fixed mask (mask size: 19 × 19; ratio factor: 0.9), the best adaptive mask (pixel total number threshold: 290; ratio factor: 0.6) and the single neural network (NN) benchmark.

The Z-test between Kappa statistics indicated that improvements between our classification model and the single neural network were statistically significant at the confidence level of 0.01. It should be noted that different from the classification results of the 2006 imagery, the optimal spectral/spatial ratio and mask sizes were achieved at high values. This is partially attributed to the different composition of the reference data. The reference data of the 2006 imagery were selected from a single area while the reference data of the 2001 imagery were from 19 reference sites. The homogeneity of the multiple reference sites in 2001 suggested a higher dependency on spectral vs. spatial local information (captured by higher ratio factors).

In addition to the best result shown in Table 4, we investigated further results obtained with equal spectral and spatial contribution (ratio set to 0.5) along with smaller window sizes (11 × 11 for the fixed mask method and 150 points for the adjusted mask method). Smaller window sizes result in lower computational load, while equal spectral and spatial contribution would allow further evaluation of local information. Table 5 shows the corresponding classification accuracies, where Kappa values are still significantly higher than the single neural network in Table 4. The proposed method offers a more balanced classification accuracy distribution

**Table 4**
Comparison between a posteriori classifier and single neural network in the corresponding validation subset.

| Accuracies | Single NN | A posteriori classifier | |
|---|---|---|---|
| | | Fixed mask (19, 0.9) | Adaptive mask (290, 0.6) |
| *Impervious* | | | |
| Producer's | 28.67 | 54.87 | 46.40 |
| User's | 66.57 | 61.76 | 62.08 |
| *Non-impervious* | | | |
| Producer's | 91.55 | 80.06 | 83.36 |
| User's | 68.62 | 75.14 | 72.61 |
| Overall accuracy | 68.30 | 70.74 | 69.70 |
| Kappa statistic | 22.92 | 35.76 | 31.41 |
| *Z-Score (P-value)* | | | |
| Single NN vs. fixed mask | 7.17 (<0.0001) | | |
| Single NN vs. adaptive mask | 11.53 (<0.0001) | | |

**Table 5**
The a posteriori classification using 0.5 ratio factor and smaller mask sizes in the corresponding validation subset.

| Accuracies | Fixed mask (11, 0.5) | Adaptive mask (150, 0.5) |
|---|---|---|
| *Impervious* | | |
| Producer's | 48.41 | 42.49 |
| User's | 61.91 | 61.70 |
| *Non-impervious* | | |
| Producer's | 82.52 | 84.52 |
| User's | 73.16 | 71.46 |
| Overall accuracy | 69.91 | 68.98 |
| Kappa statistic | 32.39 | 28.87 |

across classes (compare Producer's accuracy for the impervious class in Tables 4 and 5).

## 4. Discussion

The multi-process classification model employed in this paper can be regarded as a context classifier since it explored the spatial and spectral information from neighbors to determine the most appropriate class for a pixel. However, distinct from the way that the previous context classifiers utilize local spectral and spatial information via extraction of homogeneous objects or the probabilities of neighbors within a spatial mask, the multi-process classification model incorporates contextual information derived from already classified results. In other words, we make use of local spectral and spatial information of labeled pixels hypothesizing that this information would also propagate to neighboring unclassified pixels. The classification process of the multi-process model is similar to a post-classification context classifier which processes the intermediate classification results derived from an a priori classifier. A significant difference however is that the multi-process model only classifies part of the entire dataset instead of the entire dataset and attempts to exploit the spectral/spatial relationships between the leftover pixels and classified pixels from the a priori classifier. Compared with using the global spectral and spatial information, the integration of local spectral and spatial information proved to be more useful knowledge for the identification of unclassified pixels. Adjacent pixels were dependent and correlated with each other while exhibiting similar spectral characteristics especially in homogenous regions. Global spectral and spatial information averaged information from all pixels thus reducing the effects of neighboring pixels during the classification. This is an indication that Tobler's (1970) first law of geography suggesting "everything is related to everything else, but near things are more related than distant things" is applicable in this experiment.

Of particular interest is the comparison between the proposed pixel-based method and geo-object based classification methods (OBCM). Several existing context classification methods exploit the local spectral and spatial relationships among neighboring pixel. The OBCMs group adjacent pixels with similar spectral/spatial properties into homogeneous segmentations (called geo-objects), however the class of each pixel is not directly determined. The local spectral and spatial information is utilized during the geo-objects derivation process, but any subsequent operations are restricted by this segmentation performance. In cases of ground features with very similar spectral responses (e.g. soil and impervious classes), it is common to obtain segmentation errors through overaggressive grouping. Our approach does not suffer from similar drawbacks as it operates at the pixel level.

Looking further into the classification results using the adjusted minimum distance to mean method, there was a clear effect of the spatial neighbor assignment process and the spectral/spatial ratio factor. While we cannot extrapolate general conclusions from a single implementation it is highly encouraging that improvements over the benchmark were noticeable for almost all neighborhood and ratio combinations. Also, distinct improvement patterns may also suggest a future computational optimization of the methodology by progressively limiting the parameter search space.

The adjusted minimum distance to mean classification algorithm used as the a posteriori classifier was based on the traditional simple minimum distance to mean classifier. It is simple to understand and computationally efficient. There is only one parameter to calibrate in a straightforward manner. Despite its simplicity, it has received promising results when the *local* spectral and spatial information were integrated as inputs. The classification accuracies were improved compared to using a single neural network, which is admittedly a complex classifier and difficult to master. Our finding suggests that simple classifiers have the ability to surpass complex classifiers through elegant processing of partial results.

## 5. Conclusions

The proposed classification model explored integration of local spectral and spatial information from partially classified results to assist classification of leftover unclassified pixels. The experimental results from the 2006 scene suggested that the overall accuracy and the Kappa statistic of the entire dataset were improved significantly by using our model compared to the single neural network used as benchmark. It is further encouraging that the focus of this paper, the a posteriori classifier, showed the largest improvements in the corresponding sub dataset, the most difficult to classify sub dataset. Kappa values increased from 18.67 to 24.05 in the 2006 scene, and from 22.92 to 35.76 in the 2001 scene classification.

The multi-process classification model described in this paper is flexible and effective. Various classification algorithms can be implemented as the a priori classifier to produce partially classified results at a certain accuracy level. This proposed model can be easily extended to other land cover classification applications and the model applicability is expected to increase as spatial resolution increases, since spectral and spatial local dependences of adjacent pixels may become prominent. Future work may also focus on optimization techniques for the ratio parameter identification (e.g. using genetic algorithms). In addition, more complex but still transparent algorithms could be implemented (e.g. decision trees).

## References

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

Bauer, M., Loeffelholz, B., Wilson, B., 2005. Estimation, mapping and change analysis of impervious surface area by Landsat remote sensing. In: Proceedings Pecora 16 Conference, ASPRS Annual Conference, Sioux Falls, SD, 23–27 October, 9 p (on CDROM).

Benz, U.C., Hofmann, P., Willhauck, G., Lingenfelder, I., Heynen, M., 2004. Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. ISPRS Journal of Photogrammetry and Remote Sensing 58 (3–4), 239–258.

Binaghi, E., Gallo, I., Pepe, M., 2003. A cognitive pyramid for contextual classification of remote sensing images. IEEE Transactions on Geoscience and Remote Sensing 41 (12), 2906–2922.

Blaschke, T., Hay, G.J., 2001. Object-oriented image analysis and scale-space. Theory and methods for modeling and evaluating multiscale landscape structure. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 34 (Part 4/W5), 22–29.

Chormanski, J., De Voorde, T.V., De Roeck, T., Batelaan, O., Canters, F., 2008. Improving distributed runoff prediction in urbanized catchments with remote sensing based estimates of impervious surface cover. Sensors 8 (2), 910–932.

Coe, S.E., Alberti, M., Hepinstall, J.A., Coburn, R., 2005. A hybrid approach to detecting impervious surface at multiple-scales. In: Proceedings ISPRS Joint Conference, 3rd International Symposium Remote Sensing and Data Fusion Over Urban Areas and International Symposium Remote Sensing of Urban Areas, Tempe, AZ, 14–16 March.

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

De Voorde, T.V., De Genst, W., Canters, F., 2007. Improving pixel-based VHR land-cover classifications of urban areas with post-classification techniques. Photogrammetric Engineering and Remote Sensing 73 (9), 1017–1027.

De Jong, S.M., Hornstra, T., Maas, H.G., 2001. An integrated spatial and spectral approach to the classification of Mediterranean land cover types: the SSC method. International Journal of Applied Geoscience 3 (2), 176–183.

Dougherty, M., Dymond, R.L., Goetz, S.J., Jzntz, C.A., Goulet, N., 2004. Evaluation of impervious surface estimates in a rapidly urbanizing watershed. Photogrammetric Engineering and Remote Sensing 27 (11), 1275–1284.

Dubes, R.C., Jain, A.K., 1989. Random field models in image analysis. Journal of Applied Statistics 16 (2), 131–164.

Esch, T., Himmler, V., Schorcht, G., Thiel, M., Wehrmann, T., Bachofer, F., Conrad, C., Schmidt, M., Dech, S., 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 (8), 1678–1690.

Forman, R.T.T., Alexander, L.E., 1998. Roads and their major ecological effects. Annual Review of Ecology and Systematics 29 (1), 207–231.

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 (8), 1712–1723.

Gong, P., Howarth, P.J., 1989. Performance analyses of probabilistic relaxation methods for land-cover classification. Remote Sensing of Environment 30 (1), 33–42.

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 (1), 67–73.

Gong, P., Howarth, P.J., 1992. Frequency-based contextual classification and grey-level vector reduction for land-use identification. Photogrammetric Engineering and Remote Sensing 58 (4), 423–437.

Hay, G.J., Castilla, G., 2008. Geographic object-based image analysis (GEOBIA): a new name for a new discipline. In: Blaschke, T., Lang, S., Hay, G.J. (Eds.), Object-based Image Analysis – Spatial Concepts for Knowledge-driven Remote Sensing Applications. Springer, New York, NY, pp. 81–92.

Herold, N., 2003. Mapping impervious surfaces and forest canopy using classification and regression tree (CART) analysis. In: Proceedings 2003 ASPRS Annual Convention, Anchorage, Alaska, 5–9 May (on CDROM).

Homer, C., Huang, C., Yang, L., Wylie, B., Coan, M., 2004. Development of a 2001 national land-cover database for the United States. Photogrammetric Engineering and Remote Sensing 70 (7), 829–836.

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 (10), 2089–2102.

Jahne, B., 2005. Digital Image Processing, sixth ed. Springer, New York.

Johansen, K., Arroyo, L.A., Phinn, S., Witte, C., 2010. Comparison of geo-object based and pixel-based change detection of riparian environments using high spatial resolution multi-spectral imagery. Photogrammetric Engineering and Remote Sensing 76 (2), 123–136.

Kettig, R.L., Landgrebe, D.A., 1976. Classification of multispectral image data by extraction and classification of homogeneous objects. IEEE Transactions on Geoscience and Remote Sensing 14 (1), 19–26.

Khedama, R., Belhadj-Aissa, A., 2004. Contextual classification of remotely sensed data using map approach and MRF. ISPRS Journal of Photogrammetry and Remote Sensing 35 (7), 11–16.

Kim, K.E., 1996. Adaptive majority filtering for contextual classification of remote sensing data. International Journal of Remote Sensing 17 (5), 1083–1087.

Klein, R., 1979. Urbanization and stream quality impairment. Water Resources Bulletin 15 (4), 948–963.

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 44 (6), 1642–1654.

Liu, X., Skidmore, A.K., Van Oosten, H., 2002. Integration of classification methods for improvement of land-cover map accuracy. ISPRS Journal of Photogrammetry and Remote Sensing 56 (4), 257–268.

Liu, W., Gopal, S., Woodcock, C.E., 2004. Uncertainty and confidence in land cover classification using a hybrid classifier approach. Photogrammetric Engineering and Remote Sensing 70 (8), 963–971.

Lizarazo, I., Barros, J., 2010. Fuzzy image segmentation for urban land-cover classification. Photogrammetric Engineering and Remote Sensing 76 (2), 151–162.

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 (2), 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 (6), 1220–1229.

Mas, J., Gao, Y., Pacheco, J.A.N., 2010. Sensitivity of landscape pattern metrics to classification approaches. Forest Ecology and Management 259 (7), 1215–1224.

Mohapatra, R.P., Wu, C., 2008. Subpixel imperviousness estimation with IKONOS imagery: an artificial neural network approach. In: Weng, Q. (Eds.), Remote Sensing of Impervious Surfaces, Boca Raton, FL, pp. 21–37.

Mountrakis, G., Luo, L., 2011. Enhancing and replacing spectral information with intermediate structural inputs: a case study on impervious surface detection. Remote Sensing of Environment 115 (5), 1162–1170.

Mountrakis, G., Im, J., Ogole, C., 2011. Support vector machines in remote sensing: a review. ISPRS Journal of Photogrammetry and Remote Sensing 66 (3), 247–259.

Mountrakis, G., Watts, R., Luo, L., Wang, J., 2009. Developing collaborative classifiers using an expert-based model. Photogrammetric Engineering and Remote Sensing 75 (7), 831–844.

Paul, M.J., Meyer, J.L., 2001. Streams in the urban landscape. Annual Review of Ecology and Systematics 32, 333–365.

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 (2), 253–267.

Reigber, A., Jager, M., Neumann, M., Ferro-Famil, L., 2010. Classifying polarimetric SAR data by combining expectation methods with spatial context. International Journal of Remote Sensing 31 (3), 727–744.

Richards, J.A., Jia, X., 2006. Remote Sensing Digital Image Analysis, fourth ed. Springer, Heidelberg, Germany.

Richards, J.A., Jia, X., 2007. A Dempster–Shafer relaxation approach to context classification. IEEE Transactions on Geoscience and Remote Sensing 45 (2), 1422–2431.

Steele, B.M., 2000. Combining multiple classifiers: an application using spatial and remotely sensed information for land cover type mapping. Remote Sensing of Environment 74 (3), 545–556.

Tobler, W., 1970. A computer movie simulating urban growth in the Detroit region. Economic Geography 46 (2), 234–240.

Tolpekin, V.A., Stein, A., 2009. Quantification of the effects of land-cover-class spectral separability on the accuracy of Markov-random-field-based superresolution mapping. IEEE Transactions on Geoscience and Remote Sensing 47 (9), 3283–3297.

Townsend, F.E., 1986. The enhancement of computer classifications by logical smoothing. Photogrammetric Engineering and Remote Sensing 52 (2), 213–221.

Tso, B., Olsen, R.C., 2005. A contextual classification scheme based on MRF model with improved parameter estimation and multiscale fuzzy line process. Remote Sensing of Environment 97 (1), 127–136.

Weng, Q., 2007. Remote Sensing of Impervious Surfaces. CRC Press, Boca Raton, FL.

Wu, C., Murray, A.T., 2003. Estimating impervious surface distribution by spectral mixture analysis. Remote Sensing of Environment 84 (4), 493–505.

Yang, X., 2006. Estimating landscape imperviousness index from satellite imagery. IEEE Geoscience and Remote Sensing Letters 3 (1), 6–9.

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 (2), 230–240.

Yi, S., Huang, B., Wang, C., 2007. Pattern matching for heterogeneous geodata sources using attributed relational graph and probabilistic relaxation. Photogrammetric Engineering and Remote Sensing 73 (6), 663–670.