# Spatial Contextual Noise Removal for Post Classification Smoothing of Remotely Sensed Images


Yu Qian
Dept. of Computer Science
The University of Texas at Dallas
Richardson, TX 75083-0688, USA
qianyu@student.utdallas.edu

Kang Zhang
Dept. of Computer Science
The University of Texas at Dallas
Richardson, TX 75083-0688, USA
kzhang@utdallas.edu

Fang Qiu
Prog. in Geographic Information Sciences
The University of Texas at Dallas
Richardson, TX 75083-0688, USA
ffqiu@utdallas.edu



## ABSTRACT

Extracting accurate land use and land cover information from remote sensing data is a challenging problem due to the gap between theoretically available information in remote sensing imagery and the limited classification ability based on spectral analysis. Traditional classification techniques based on spectral analysis of single pixel usually produce "noisy" results that contain many wrongly classified pixels. This paper presents a novel post classification method to detect the pixels that are wrongly classified and reassign them to correct fields in spatial context. The strategy is demonstrated through the classification of a benchmark digital aerial photograph. The experimental results show that the proposed approach can produce a more accurate classification than previous approaches.



## Keywords

Remote sensing, GIS, spatial data mining, post classification smoothing, noise removal


## 1. INTRODUCTION

Remote sensing imagery provides a huge amount of data about earth surface for global and detailed analysis. Mining useful information from remote sensing data would help GIS and earth science researchers significantly. Among the mining approaches commonly used in GIS and remote sensing area, automatic classification of the remotely sensed image is the most widespread analysis approach and in many cases a preliminary step for further analysis. While many digital image processing techniques have been applied to classify remotely sensed data, the extraction of land use and land cover information from the data is still a challenging problem [2, 18]. Traditional classification techniques classify land cover/use on a basis of spectral distribution of the pixels within an image, whereby each pixel is associated with the most similar spectral class. Since pixels of the same land use/cover class may not have similar spectral property, methods based on spectral analysis can produce results that are "noisy" due to the high spatial frequency of the land covers. Moreover, the popular classification algorithms are based on single pixel analysis, producing a geometric outline of land covers that does not correspond to real spatial entity representation such as fields, roads, and streams [4]. Such problems become more severe as modern remote sensing devices are increasing the spatial resolution to meet the requirements on accuracy and provide more detailed information for the study area. As the image resolution becomes finer and the number of pixels used to represent single object increases, more noise will be produced in the classified image. For example, images taken by IKONOS satellite can have a resolution of 1m, which is much smaller than the sizes of the objects under study such as urban, forest, and water. As a result, per-pixel-based classifiers using only spectral information can produce a large amount of "salt and pepper" noise pixels both inside and outside the objects in the classified image, which decreases the classification accuracy significantly.

To solve such problems, recent effort [10, 14] has been made to discover and utilize other implicit image information to allow a more accurate classification, among which spatial contextual information, characterized by the distribution of pixels of the objects being studied, is the most related information to the shape and outline of the objects. According to the way to integrate spatial contextual information, there are two categories of classification approaches [3]: per-field classification and object-based classification. Per-field classification requires priori information about boundaries of objects in the image, which limits its applications to many areas [18]. In contrast, object-based classification does not need GIS input. Object-based classification usually starts with an initial step of grouping neighboring pixels into meaningful areas/objects through advanced image segmentation techniques [2, 5]. Then the classification is performed on the generated objects instead of pixels. Therefore, results of object-based classification rely highly on the correctness of object generation step. If the object is generated incorrectly or inaccurately, the later classification becomes meaningless. Further, because an object usually consists of a large number of pixels, when it receives a wrong classification, the statistics of the area represented by that object-class can be seriously affected and a big area may be wrongly classified.

To avoid drawbacks of the aforementioned approaches, this paper presents a novel method to remove noise pixels produced by per-pixel-based classification. Through transforming each image pixel into a spatial data point, the problem becomes a spatial data cleaning problem and advanced spatial mining methods can be applied. The idea is based on the observation that although spectral analysis would produce noisy classification results, the majority of pixels of each field in the image are classified correctly, which provides desirable contextual information for later spatial analysis. Further, the pixels that are wrongly classified into a field appear inconsistent with other pixels of the same field on spatial distribution. Conducting spatial analysis on each classified field could distinguish from others the pixels that are wrongly classified. Then each wrongly classified pixel can be reassigned to its most related field in spatial context. Compared with previous approaches, such a strategy has three advantages:


Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.



*SAC'05*, March 13-17, 2005, Santa Fe, New Mexico, USA.


Copyright 2005 ACM 1-58113-964-0/05/0003…$5.00.

- The classification does not require prior GIS knowledge.
- There is no risk of classifying big areas wrongly because the basic unit to be classified is pixel instead of object.
- Noise pixels produced by per-pixel-based spectral analysis can be merged into correct fields, which facilitate a transition from pixel-based to object-based image interpretation.

Figure 1 illustrates the whole procedure of the approach, which consists of three steps: 1) image classification based on spectral analysis, 2) post classification smoothing of each classified field, and 3) combining the fields into the final result. Step 1 uses existing spectral analysis methods to classify the given image into several land use/cover fields. Step 2 discovers noise pixels from each classified field and reallocates them to correct fields. In Step 3, all fields are combined together to constitute the final classification result.

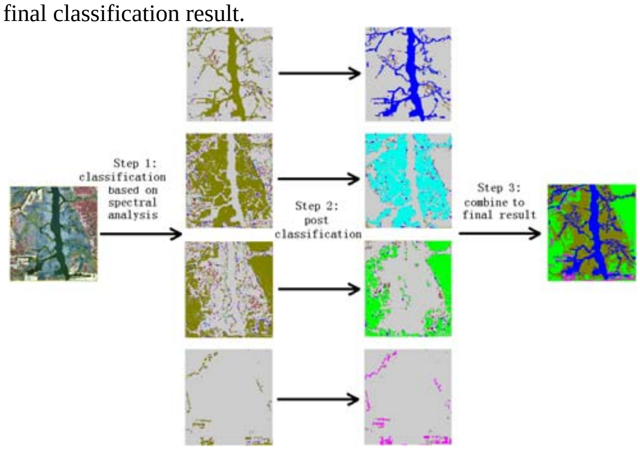



**Figure 1. The three steps of the proposed approach.**

Among the three steps, Step 2 is the key step of the post-classification procedure. It accepts a classified field as input and outputs a "clean" field. Step 2 can be divided into four stages:

- Mapping each classified field to a spatial data set, in which each data point corresponds to an image pixel at the same position.
- Modeling spatial dependency among data points with the connections of a *k*-mutual neighbor graph [8] constructed on the spatial data set.
- Separating the data points into noise and true data through partitioning the *k*-mutual neighbor graph with the *k*-core algorithm [13].
- Reassigning noise points to appropriate data sets. The class ID of the noise point is determined by the majority voting on the class IDs of its *k*-nearest neighbors. After all noise points have been reallocated, resulting data sets are mapped back to image fields.

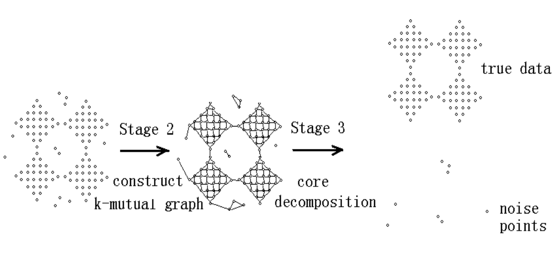



**Figure 2. The stages 2 and 3 of the second step to discover noise points.**

Figure 2 illustrates stages 2 and 3, the process to discover noise points through applying the *k*-core algorithm to partition the *k*-mutual neighborhood graph constructed on the data set.

In summary, the contributions of this paper include:

1) A novel post-classification method that can remove noise pixels produced by per-pixel-based classification more effectively so that more accurate classification of remote sensing images can be obtained.

2) A framework to classification of remote sensing data that connects per-pixel-based approaches and spatial noise removing methods.

The rest of the paper is organized as follows. Section 2 provides a brief review on previous post-classification methods and clarifies the motivation of our work. Section 3 explains the spatial noise removing method. Section 4 classifies a benchmark digital aerial photograph and presents a comparison between several representative post-classification methods. Discussion is conducted in Section 5. Section 6 concludes the paper.

## 2. RELATED WORK

The idea of using post-classification procedure to remove noise pixels is not new. Since 1980s there have been post-classification approaches [16, 17] to noise elimination through a moving window and majority logical filter. The basic idea of majority filter is straightforward. Given a moving window with a size $n$ x $n$, where $n$ is a user-defined threshold, the classified image is divided into many small squares with $n$ x $n$ pixels as the window moves. Then the class ID of each pixel in the moving window will be determined by majority voting on the class IDs of the $n$ x $n$ pixels so that the minority can be removed as noise. Although majority filter is easy to implement, its weaknesses are apparent. First, the size of neighborhood for the filter has to be very large for noise to be sufficiently removed, while a large size neighborhood may alter the boundaries between classes and create zigzag bounding polygons. Moreover, a loss of meaningful information in classified data is shown because of the geometric and dimensional non-correspondence of real objects with the moving window implementation matrix. Finally, such methods need extensive editing operations on classified images in order to be stored in GIS databases [4].

The latest progress on post-classification smoothing in literature is the size-based filter proposed by Jensen et al. [9]. All the regions that have an area less than a user-defined threshold are selected as noise. If a noise region has the same color as background, it is called "interior noise". Otherwise, it is "exterior noise". The task of size-based removal is to merge interior noise into corresponding fields and remove exterior noise. Size-based noise removal can preserve more details of the field boundary than majority filter. It, however, still has two drawbacks. The first drawback is due to the threshold on region size. Some small-size regions may not be noise while some big-size regions may be. Further, the region group operation puts pixels of the same spectral class into the same region only if they are neighbors, which is a too strict condition in some cases. For example, pixels may occupy a big area without neighboring to one another, given they are interleaved. In such a case size-based filter would regard all pixels as noise and remove the whole region.

Unlike the aforementioned approaches, the post-classification method introduced in this paper does not work on image pixels but map them to spatial data points to take advantage of spatial mining techniques. While significant research progress has been made in the field of spatial data mining, there is little work that can be applied to process remote sensing imagery. One of the major reasons is the lack of an appropriate mapping between images and spatial data sets. An image before classification contains various kinds of information such as texture, size, shape, color, and etc. It is very difficult, if not impossible, to transform all the information to a spatial data set that contains data points with two coordinates only. To process spectral and spatial information separately, the proposed approach first classifies image pixels based on spectral information. After classification, spatial properties of pixels of the same class can be mapped to spatial coordinates directly. Thus each class becomes a spatial data set and the advanced spatial analysis methods can be applied to discover patterns or outliers from this class. This provides us a chance for effective spatial contextual noise removal.

# 3. SPATIAL NOISE DISCOVERY AND REMOVAL

The key issue of post-classification is how to discover and remove noise pixels. This section will introduce the spatial noise removing method, which has been described as Step 2 in the whole procedure shown in Figure 1. As designed in Section 1, Step 2 is a stand-alone process that accepts classified fields from existing advanced classification methods and outputs the fields with noise removed. As depicted in Figure 2, Step 2 consists of four stages. The first stage is a direct mapping from pixels to data points through constructing a 2-D coordinate system for the given image. For rectangular images, the *x*-axis is along the bottom edge of the image while the *y*-axis along the left side edge, as illustrated in Figure 3 (a). For images of irregular shapes, the bottom edge and left side edge of its bounding rectangle is used as the axes, as illustrated in Figure 3 (b). Each image pixel becomes a 2-D data point in the coordinate system and the whole image becomes a spatial data set. It is important to note that although the bounding rectangle of an image is not unique and the definitions of bottom and side edge depend on how we put the image, the distances between the data points in the generated spatial data set is unchanged for the same image. Thus the later stage will not be affected because it uses only distance information to construct a *k*-mutual neighbor graph on the data set, which will be introduced in Section 3.1. Section 3.2 will apply the *k*-core algorithm [13] to partition the graph to detect outliers while Section 3.3 will reallocate the discovered noise points to correct fields.

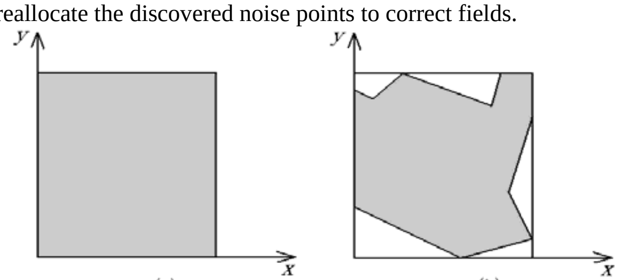
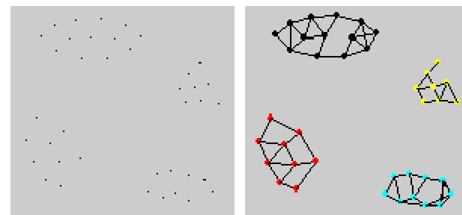



**Figure 3. Mapping from a (a) rectangular, (b) irregular image (shown in grey) to a spatial data set.**

## 3.1 *K*-mutual Neighbor Graph

*K*-mutual neighbor graph [8] is a commonly used graph structure to model the spatial dependency between data points. As illustrated in Figure 4 (b), each vertex of a *k*-mutual neighbor graph represents a data item. For each pair of data items, only if both data items are among the *k*-most similar data items of each other, can there be an edge between the two corresponding vertices. In this paper, we refer to data items as points on a 2-D space $\Re^2$ and the similarity of two data points is measured by the Euclidean distance between them. Such a similarity definition has a solid theoretical foundation [6] and follows Tobler's first law of geography: everything is related to everything else but nearby things are more related than distant things [15].



**(a) (b)**

**Figure 4. (a) A 2D data set; (b) its *4*-mutual neighbor graph.**

Generally, the advantages of representing data using a *k*-mutual neighbor graph include: first, data points that are far apart are completely disconnected from the graph. Second, the constructed graph is able to represent the natural density dynamically. In a dense region, the neighborhood radius of a data point is determined narrowly, and in a sparse region, the neighborhood radius becomes wide. Third, the number of graph edges is linear to the number of vertices. The first two advantages decide the graph structure can be used to distinguish noise and true data while the last one guarantees the efficiency of operations to be performed on the graph.

## 3.2 The *K*-core Algorithm

The task of noise discovery from a spatial data set can be accomplished through partitioning the constructed *k*-mutual graph into small sets of vertices called *cores*. The notion of a core was introduced by Seidman [13]. Let $G = (V, E)$ be a graph, $V$ is the set of vertices and $E$ is the set of edges. A sub-graph $H_k = (W, E \mid W)$ induced by the set $W$ is a *k-core* or a *core of order k* iff $\forall v$ in $W$: degree $(v) \geq k$ and $H_k$ is the maximum sub-graph with this property. The core of maximum order is also called the *main core*.

The algorithm for determining the core hierarchy is simple: from a given graph $G=(V, E)$, recursively delete all vertices of degrees less than $k$ and lines incident with them, the remaining graph is the *k*-core. Known as the *k*-core algorithm, it costs only $O(m)$ time, where $m$ is the number of edges for the given graph [13]. The *k*-core algorithm has previously been used to produce layouts for very big graphs [1] and generate spatial clusters [11, 12], in which its efficiency and effectiveness have been verified. To avoid confusion, we will hereafter use $k_c$ as the $k$ used in *k*-core algorithm and $k_m$ as the $k$ used in *k*-mutual graph. For a *k*-mutual graph with $n$ vertices and $m$ edges, we have $m \leq k_m\, n/2$ if $n$ is the number of vertices[1], so applying *k*-core algorithm to a *k*-mutual graph requires only linear time.

A sketch map of the core hierarchy is shown in Figure 5. Given a dataset $D$, suppose the corresponding *k*-mutual graph is $G=(V, E)$, $|V|=n$ and $|E|=m$. Let $k=0, 1, 2, \ldots, x$ and apply the *k*-core algorithm to $G$. The set of cores obtained are denoted by $H_0$, $H_1$, ..., $H_{x-1}$, $H_x$ for $k=0, 1, 2, \ldots, x$, respectively, where $H_i$ represents the core of order $i$. The higher the order is, the darker

[1] Given a data set, the corresponding *k*-nearest neighborhood graph $G=(V, E)$, and the *k*-mutual neighborhood graph $G'=(V', E')$, we have $|V|=|V'|$, $|E'|\leq|E|\leq k|V|$, and $|E|+|E'|=k|V|$, so $|E'|\leq k|V'|/2$.

the area is. To understand the concept of core-ID, let $S_x = H_x$, $S_{x-1} = H_{x-1} - H_x$ ,…, $S_1 = H_1 - H_2$, $S_0 = H_0 - H_1$, and $S_x = H_x$ is the main core, a core-ID of a vertex is $i$ if and only if the vertex belongs to $S_i$.

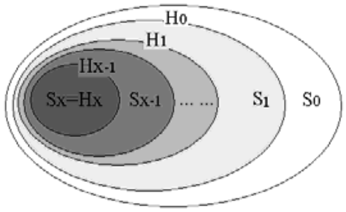



**Figure 5. An example sketch map of core hierarchy.**

It is not hard to justify why core decomposition suits spatial noise removal. According to the definition of the core, each vertex of a core of $k_c$ must have a degree greater than or equal to $k_c$ even after all other cores with orders smaller than $k_c$ have been removed, i.e., removing the cores of smaller orders will not decrease or affect the quality of the cores of bigger orders. While the connectivity of a $k$-mutual graph reveals the density distribution, removing cores of small orders matches the natural rule that removing noise should not decrease or affect the density of true data.

## 3.3 Reallocating Noise

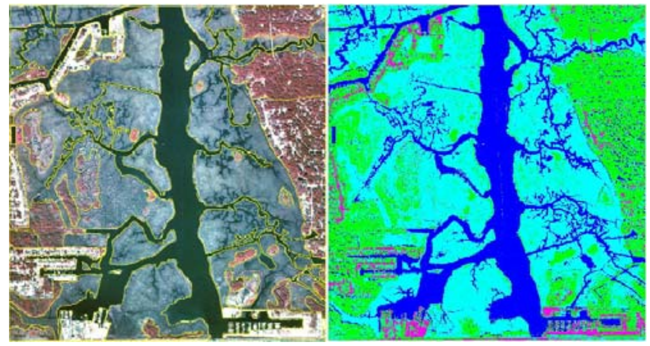
Post-classification of remote sensing images, however, cannot really remove noise pixels because all pixels need to appear in the final classification result. Noise pixels should be reallocated to the fields they belong to. The reallocation scheme is based on a central concept in statistical pattern recognition, called nearest neighbor consistency [7]: an element is consistent with its $k$-nearest neighbors. Ding and He [6] further extend this concept to data clustering: for any element in a cluster, its $k$-nearest neighbors should also be in the same cluster. Similarly, this paper extends the concept to image classification: if a pixel belongs to a field, its $k$-nearest neighbors should belong to the same field too. Based on the nearest-neighbor consistency, if the $k$-nearest neighbors of a pixel do not belong to a class, the pixel should not belong to the class either. The more nearest neighbors are in the field, the more possible the pixel belongs to the field. Thus the class ID of each noise point can be decided easily by majority voting on the class IDs of its $k$-nearest neighbors in all data sets. Such a strategy is supported by our experimental studies, as presented in the following section.

# 4. APPLICATION TO CLASSIFICATION

This section will evaluate the effectiveness of the three-step strategy through the classification of a benchmark aerial photograph used by Jensen et al. [9], as shown in Figure 6 (a). Section 4.1 will introduce the classification system we use while Section 4.2 will compare the final result with two other representing approaches.

## 4.1 The Classification System

Jensen et al. [9] introduce an image interpretation system to extract rural and urban land use and land cover information using a back-propagation neural network. The system successfully shows a well-trained image interpreter can produce better classification results than the approaches using traditional statistical methods. The system, however, still relies on the per-pixel-based classification. Since pixels of the same field may not have consistent spectral class and pixels of different objects may have similar spectral property, the produced classification result is "noisy", as shown in Figure 6 (b).



**(a) (b)**

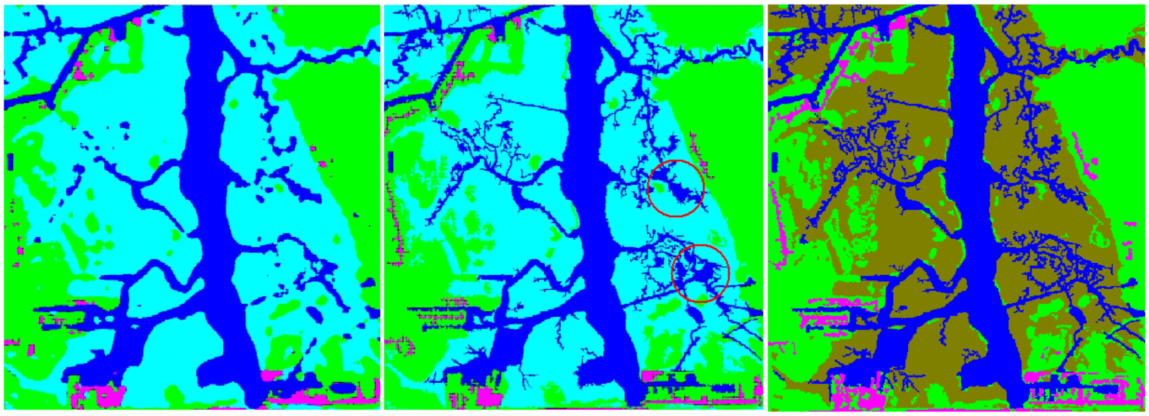
**Figure 6. (a) Digital National Aerial Photography Program (NAPP) image (1x1m) of Jacksonville Beach, FL.**
**(b) The classification result base on spectral analysis without post-classification.**

## 4.2 Experimental Results

After noise pixels of each field are discovered and reallocated, the "clean" fields are combined to the final results. Figure 7 provides a comparison between the results produced by (a) focal majority filter, (b) size-based filter, and (c) our approach. Among the three approaches, our approach preserves most details of boundaries of geographic fields. Figure 7 (a) shows a significant information loss produced by focal majority filter due to the fixed-size moving window. Figure 7 (b) circles several small regions of wetland that are removed wrongly by size-based filter due to the surrounding water. Section 2 has explained why they are removed: 1) the size threshold used in the size-based filter removes all small regions as noise while some of them are actually true data; 2) the pixels belong to the same region may not be connected. Many wetland pixels are separated from main body by surrounding rivers. Removing such pixels would cause inaccurate boundaries of both wetland and water.

# 5. DISCUSSION

Extracting land use and land cover information from high spatial resolution data with visual image interpretation and manual digitization into GIS database is a laborious procedure. The proposed automatic classification approach may help ease the burden. According to Jensen et al. [9], the potential time saving could be more than 50 percent using the automatic classification. For the post-classification procedure proposed in this paper, a good selection of $k_c$ can reduce the running time significantly. A too big $k_c$ would cause too many points to be reallocated while a too small $k_c$ would cause noise pixels remained in the field. While the core decomposition is a fast procedure that requires only linear time with regard to the number of pixels, the reallocation cost is much higher due to the distance computation performed between each noise pixel and all other pixels. The more noise pixels discovered, the slower the procedure would be. Current setting allows users to participate in the $k_c$ selection through a 3D visualization [11, 12] so that the whole post-classification process is optimized. Although such a strategy requires a little expertise to judge noise pixels from true data, the 3D visualization makes user intervention quite effective and efficient. To make the whole approach automatic, prior knowledge about noise distribution is needed, which could be a topic of our future research.



**(a)** **(b)** **(c)**

**Figure 7. The post-classification result of (a) focal majority filter, (b) size-based filter, and (c) our approach.**

# 6. CONCLUSION

Recent availability of multispectral imagery with very fine spatial resolution has been increasing the lag between the large amount of image/spatial data produced and the limited classification ability. Traditional classification methods based on spectral analysis cannot extract land use and land cover information accurately. A successful approach must take spatial contextual information into consideration. This paper presents a novel post-classification approach that can remove noise pixels produced in spectral analysis so that the overall accuracy of the classification is improved. Compared with previous approaches, the proposed method has significant advantages on the quality of classification results. Future work will focus on two directions: how to evaluate the classification accuracy automatically and how to integrate the proposed method with more classification systems.

# 7. REFERENCES


[1] Batagelj, V., Mrvar, A., and Zaversnik, M. (2000) Partitioning approaches to clustering in graphs, *Proc. Graph Drawing'1999,* LNCS, 2000, pp. 90-97.

[2] Benz, U. C., Hofmann, P., Willhauck, G., Lingenfelder, I. and 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 (2004), pp. 239-258.

[3] Blaschke, T., Lang, S., Lorup, E., Strobl, J., and Zeil, P. (2000) Object-oriented image processing in an integrated GIS/remote sensing environment and perspectives for environmental applications. Cremers, A. and Greve, K. (eds), *Environmental Information for Planning, Politics, and Public*, Vol. II, Metropolis-Verlag, pp. 555-570.

[4] Caprioli, M. and Tarantino, E. (2001) Accuracy assessment of per-field classification integrating very find spatial resolution satellite imagery with topographic data. *Journal of Geospatial Engineering*, Vol. 3, No. 2, pp. 127-134.

[5] De Kok, R., Schneider, T. and Ammer, U. (1999) Object-based classification and applications in the alpine forest environment. *International Archives of Photogrammetry and Remote Sensing*, Vol. 32, Part 7-4-3 W6.

[6] Ding, C. and He, X. F. (2004) K-nearest-neighbor consistency in data clustering: incorporating local information into global optimization. In *Proc. of the 19th ACM Annual Symposium on Applied Computing (SAC'04)*, pp. 584-589.

[7] Fix, E. and Hodges, J. L. (1951) Discriminatory analysis, nonparametric discrimination: consistency properties, *Technical Report 4*, USAF School of Aviation Medicine, Randolph Fields, TX.

[8] Jain, A. K. and Dubes, R. C. (1988) *Algorithms for Clustering Data.* Prentice-Hall advanced reference series. Prentice-Hall, Inc., Upper Saddle River, NJ.

[9] Jensen, J. R., Qiu, F. and Patterson, K. (2001) A neural network image interpretation system to extract rural and urban land use and land cover information from remote sensor data. *Geocarto International*, Vol. 16, No. 1, pp. 1-10.

[10] Li, J. and Narayanan, R. M. (2004) Integrated spectral and spatial information mining in remote sensing imagery. *IEEE Trans. on Geoscience and Remote Sensing*, Vol. 42, No. 3, pp. 673-685.

[11] Qian, Y., Zhang, G. and Zhang, K. (2004) FACADE: A Fast and Effective Approach to the Discovery of Dense Clusters in Noisy Spatial Data, In *Proc. ACM SIGMOD 2004 Conference*, Paris, France, 13-18 June 2004, ACM Press, pp. 921-922.

[12] Qian, Y. and Zhang, K. (2004) Discovering spatial patterns accurately with effective noise removal. In *Proc. 9th ACM SIGMOD Workshop on Research Issues in Data Mining and Knowledge Discovery (DMKD'04)*, Paris, France, 13 June 2004, ACM press, pp. 43-50.

[13] Seidman, S. B. (1983) Network structure and minimum degree. *Social Networks*, 5, pp. 269-287.

[14] Shekhar, S., Schrater, P. R., Vatsavai, R. R., Wu, W., and Chawla, S. (2002) Spatial contextual classification and prediction models for mining geospatial data, *IEEE Trans. on Multimedia*, Vol. 4, No. 2, pp. 174-188.

[15] Tobler, W. R. (1979) *Cellular Geography, Philosophy in Geography*, Gale, W. R. and Olsson, W. R. Eds, The Netherlands: Reidel.

[16] Tomas, I. L. (1980) Spatial postprocessing of spectrally classified Landsat data, *Photogrammetic Engineering and Remote Sensing*, 46, pp. 1201-1206.

[17] Townshend, J. R. G. (1986) The enhancement of computer classification by logical smoothing, *Photogrammetic Engineering and Remote Sensing*, 52, pp. 213-221.

[18] Walter, V. (2004) Object-based classification of remote sensing data for change detection, *ISPRS Journal of Photogrammetry and Remote Sensing,* 58 (2004), pp. 225-238.