



Visualization-informed noise elimination and its application in processing high-spatial-resolution remote sensing imagery

Yu Qian^{a,*}, Fang Qiu^b, Jie Chang^b, Kang Zhang^a

^aDepartment of Computer Science, The University of Texas at Dallas, Richardson, TX 75083, USA

^bProgram in Geographic Information Sciences, The University of Texas at Dallas, Richardson, TX 75083, USA

Received 6 March 2006; received in revised form 30 January 2007; accepted 5 February 2007

Abstract

Noise removal is perhaps one of the most fundamental and challenging tasks for extracting useful information from a spatial data set. One of the challenges is that there is no general agreement on the definition of noise that can be universally applied to all different domains. This paper proposes a novel technique called Visualization-Informed Noise Elimination (VINE) to support a customized noise removal through incorporation of domain knowledge. The VINE technique consists of three steps of consecutive operations. First, a k -mutual neighbor graph is derived from a spatial data set to model the spatial proximity among data points. Next, a fast partitioning method is employed to reassemble graph nodes into groups. Last, a 3-dimensional (3D) visualization model is created to provide a layered view of the partitioned data, which allows an informed identification and elimination of noise by tailoring to the requirements of a specific domain. The flexibility and customizability provided by this novel technique ensures an effective differentiation of noise from valid data and demonstrates various advantages over traditional methods with improved results. When adapted in post-classification smoothing of high-spatial-resolution remotely sensed images, this approach was able to discover and reassign noise (such as shadows often seen in high-spatial-resolution images) to its proper target class. By incorporating domain knowledge and making use of spatial contextual information, the VINE technique could produce results significantly superior to existing approaches such as majority filter and size-based noise removal.

© 2007 Elsevier Ltd. All rights reserved.

Keywords: Spatial data set; Noise removal; Visualization; Remote sensing; High-spatial-resolution images

1. Introduction

Recent years have witnessed an increased availability of enormous amounts of spatial data, acquired by remote sensing satellites, medical equipments, various digital cameras, etc. The widespread accessibility to spatial data is leading

to growing research interest in extracting and discovering interesting and useful but implicit characteristics and patterns that may exist in the data (Han and Kamber, 2000). For applications that involve pattern discovery from spatial data, data cleaning is usually one of the most fundamental and challenging tasks. For example, data cleaning through noise removal is very critical to the achievement of an accurate feature classification from high-spatial-resolution remotely sensed images.

*Corresponding author. Tel.: +1 214 648 4191;
fax: +1 214 648 4080.

E-mail address: qianyu_utd@acm.org (Y. Qian).



Fig. 1. A benchmark data set containing ambiguous noise.

The motivation of this paper can be simply illustrated with a benchmark data set used by CHAMELEON (Karypis et al., 1999). As shown in Fig. 1, the benchmark contains six English letters crossed by a horizontal line. To accomplish effective data cleaning, a question must be answered first: is the horizontal line noise or valid data? For pure letter recognition, the line is usually regarded as noise, which may be caused by a fold in the paper. It, however, could also be considered as part of valid data, for example, representing a “strike out” style or a decoration, in a different application. Noise is therefore domain-specific and its definition varies in different applications. Many methods have been proposed to remove noise, but few of them were specifically designed to handle domain-specific noise. The fact that the horizontal line not only overlaps the six letters but also has a similar density to them poses a challenge for existing noise-removing approaches to segregate it from the letters. A comparable situation can be observed in remotely sensed images, where shadows of high-rise buildings and tall trees may have similar spectral characteristics to dark objects (such as water body), making them extremely difficult to be removed using traditional approaches of noise removal.

Noises like these are often referred to as being *ambiguous* (Qian and Zhang, 2004), and their effective segregation from valid data requires domain knowledge. In this paper, we proposed a core-based Visualization-Informed Noise Elimination (VINE) technique that is specifically designed to address this issue so that knowledge of domain experts can be incorporated into the data cleaning process with the support of a visualization tool. The development of such a technique allows knowledge-guided noise removal informed and supported by the visualization tool, which works

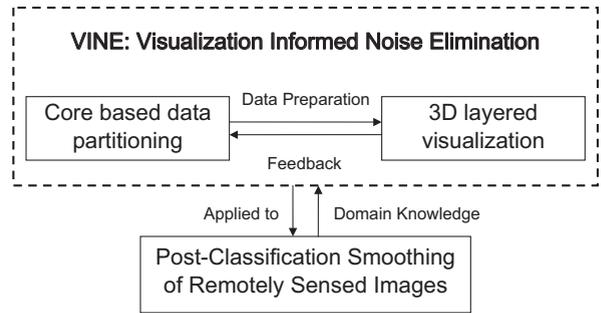


Fig. 2. Structure of proposed core-based VINE approach and its application to post-classification smoothing of remotely sensed images.

reciprocally with the noise differentiation and elimination process. Reciprocally in the sense that the data are first reorganized into different cores for visualization purposes, the visualization of the reorganized data then makes it possible to examine the characteristics of noise and valid data that function as feedback in reconfiguring the removing method with domain-specific parameters (Fig. 2). It is worth to mention that VINE is part of an effort that aimed to construct a spatial data clustering system named FAÇADE (Qian et al., 2004). To demonstrate the effectiveness and efficiency of the proposed technique, we tested the above benchmark spatial data set and two high-spatial-resolution remotely sensed images, and compared the results with those obtained using traditional noise-removing approaches.

2. Background and related work

The difficulty in removing ambiguous noise lies in the lack of a universally accepted definition of noise. To define noise, it is a common practice to first identify certain criteria that can be quantified with a threshold by which noise can be segregated from valid data. Frequently used criteria include distance between data points, size of data groups, and number of surrounding points (i.e., degrees of nodes in a constructed nearest-neighbor graph of the original data). Accordingly, existing noise-removing methods can also be classified into three categories: distance based, size based, and degree based. Popular distance-based methods include DBSCAN (Ester et al., 1996), OPTICS (Ankerst et al., 1999), and DENCLUE (Hinneburg and Keim, 1998). Distance-based approaches can discover points distributed sparsely as noise, but they cannot identify noise having density similar to valid data

points. Conversely, they can mistakenly mark sparsely distributed valid data points as noise. In addition, thresholds for distance between data points are usually difficult to set without prior knowledge of the data set. Size-based approaches, such as those employed in BIRCH (Zhang et al., 1996) and RandomWalk (Harel and Koren, 2001), must rely on a certain clustering method to first segment data points into groups, but the clustering method itself usually suffers from the presence of noise, thus resulting in a typical chicken and egg dilemma. Besides, the threshold for cluster size is similarly difficult to determine without numerous experiments. Existing degree-based noise-removing methods, such as the one used in SNN (Ertoz et al., 2003), are usually based on the construction of a nearest or mutual neighbor graph. They often also require the specification of a parameter to construct the graph and a threshold for degrees of corresponding vertices, falling short of effectively removing ambiguous noise.

Instead of relying on clustering methods to identify noise, wavelet transformation (Mallat, 1989) reduces noise in images through a hierarchical decomposition of original data into a nested subsurface of approximations at different resolution (Chang et al., 1998; Fan and Xia, 2000; Lang et al., 1996; Lebart and Boucher, 1996; Strela, 2000; Vidal-Pantaleoni and Marti, 2004). The main idea of wavelet-based noise removal is to represent the data using a basis function, where the large coefficients are usually corresponding to the valid information and the smaller ones to noise. By modifying the coefficients of the basis function properly, noise can be effectively removed from the data (Li et al., 2002). Recently, wavelet-based methods have also been used in finding clusters automatically in noisy environments (Sheikholeslami et al., 2000). However, it is not always a trivial work to select or define an appropriate mother wavelet and resolution for an unknown data set.

None of the aforementioned noise-removing methods, however, is purposely designed to encode domain-specific information. Since noise is often application dependent, requirements of a specific domain area should be taken into consideration before removing noise. In our proposed approach, noise is not defined until the prior domain knowledge is incorporated. To achieve this, the data points are first preprocessed to create a k -mutual neighbor graph, and then a fast graph partitioning method, the k -core algorithm (Seidman, 1983), is

applied to partition graph nodes, i.e., data points, into groups. The partitioning of the data points enables users to classify each group as either noise or valid data through the support of a visualization tool. This approach provides the flexibility to assign data of different densities into the same class according to custom specifications of a specific domain.

It is important to clarify that our approach is different from those aiming at discovering network intrusions, detecting fraud, etc. Their primary goal is to discover outliers and abnormal situations based on criteria such as distance (Aggarwal and Yu, 2001; Bay and Schwabacher, 2003; Knorr et al., 2000), density (Breunig et al., 2000), and graph-related features (Shekhar et al., 2001). Outliers are often explicit and well-defined patterns in nature, usually not comprising ambiguous noise that we are targeting here.

3. Data preprocessing and visualization

As mentioned earlier, the preprocessing of the data set for visualization is accomplished through constructing a k -mutual neighbor graph, and decomposing the resulting graph into small groups using the k -core algorithm (Seidman, 1983). With data groups generated, the visualization is then made possible to incorporate subjective domain knowledge into configuring the subsequent noise-removing scheme.

3.1. k -mutual neighbor graph

The k -mutual neighbor graph (Jain and Dubes, 1988) is an effective graph structure to model the spatial proximity between data points. It is one of the variants of the well-known k -nearest-neighbor graph that is widely used to model spatial dependency. As illustrated in Fig. 3, each vertex of a k -mutual neighbor graph represents a data point. For each pair of data points, an edge can be established only if both data points are among the k -most similar data points of each other. For spatial data sets, the similarity between two data points is usually measured by the Euclidean distance.

The use of a k -mutual neighbor graph exhibits the following advantages (Jain et al., 1999): data points that are far apart are completely disconnected; a higher degree implies a point belonging to a more uniformly distributed region; data points of uniformly dense and of uniformly sparse regions have

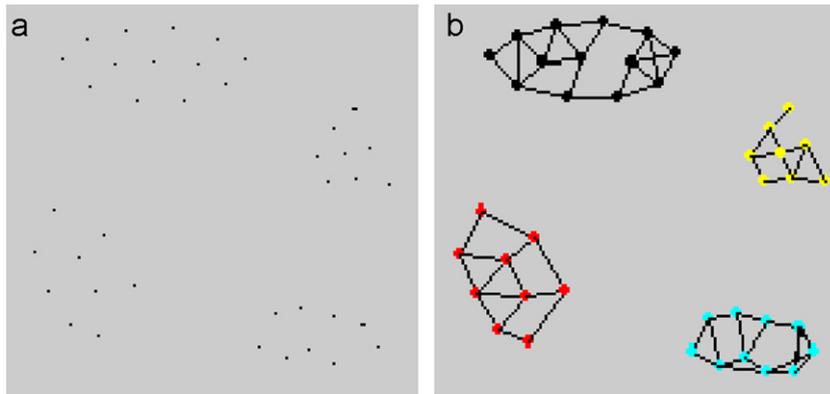


Fig. 3. (a) A 2D spatial data set and (b) its 4-mutual neighbor graph.

similar degrees; the number of graph edges is linear to the number of vertices. The first three advantages allow the effective distinguishing of noise from valid data, regardless of the underlying density of the spatial data distribution, while the last guarantees the efficiency of subsequent graph operations.

3.2. Decomposing data into cores

Once a k -mutual graph is constructed, a graph partitioning method can be applied. The notion of a core (or cluster) is introduced by Seidman (1983). Let $G = (V, E)$ be a graph. V is the set of vertices and E is the set of edges. A subgraph $H_k = (W, E|W)$ induced by the set W is a k -core or a core of order k iff $\forall v$ in W : $\text{degree}(v) \geq k$ and H_k is the maximum subgraph with this property. The core of the maximum order is also called the *main core*. The cores have the following properties:

- They are nested: $\forall i < j \rightarrow H_j \subseteq H_i$.
- There exists an efficient algorithm to determine the core hierarchy.
- A core is not necessarily a connected subgraph.

A core hierarchy is sketched in Fig. 4. Given a graph $G = (V, E)$, $|V| = n$ and $|E| = m$, G is partitioned into the set of cores, denoted by $H_0, H_1, \dots, H_{x-1}, H_x$, where H_i represents the core of order i . In Fig. 4, the higher the order, the darker the area. To explain the concept of *core-ID*, let $S_x = H_x$, $S_{x-1} = H_{x-1} - H_x, \dots, S_1 = H_1 - H_2, S_0 = H_0 - H_1$, 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 .

Determining the core hierarchy from a given graph $G = (V, E)$ (i.e., assigning a core-ID for each

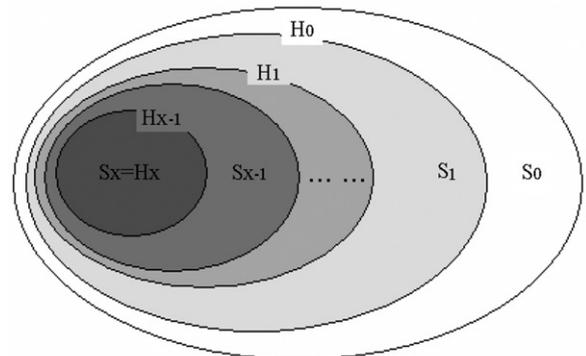


Fig. 4. A core hierarchy showing core decomposition result.

vertex) is called *core decomposition*, which costs only $O(m)$ time for the given graph with m edges (Seidman, 1983). Previously being used to produce layouts for very big graphs (Batagelj et al., 2000), the core decomposition algorithm is now utilized in this paper to partition the k -mutual graph constructed above. It is important to note that removing the vertices of small degrees is different from removing cores of small orders. Fig. 5 shows such a difference and the advantage of removing cores of small orders in preserving valid data. The use of core decomposition for spatial noise removal is justifiable as the removal of the cores with smaller orders will not affect the integrality of the cores of bigger orders. Therefore, cores with small orders, which are often assumed as noise, have no impact on the distribution of valid data if being removed. Additionally, vertices with slightly different degrees that are connected with each other would belong to the same core. This is consistent with the common recognition process of human beings.

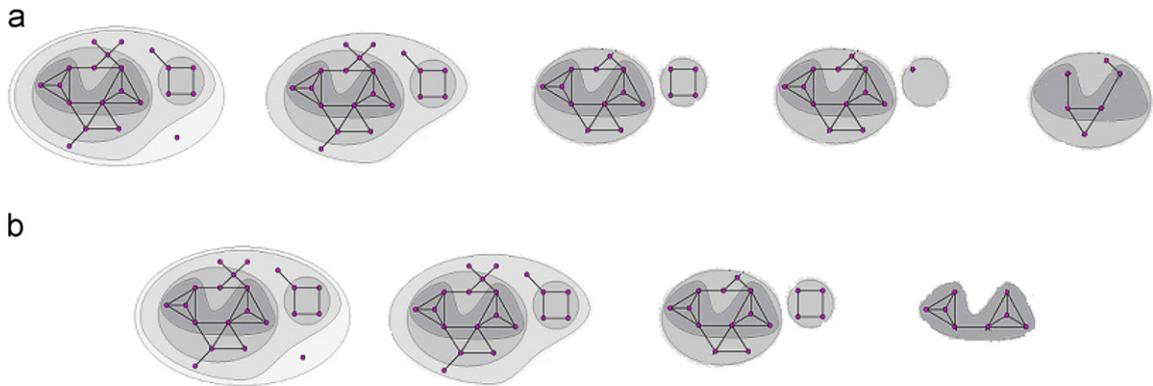


Fig. 5. A comparison between (a) degree-based noise removal: begin from original graph, remove vertices whose degree = 0, 1, 2, and 3 at each step—no true cluster discovered, and (b) core-based noise removal: remove 0, 1, 2—core step by step—true cluster discovered and noise removed.

The core decomposition method applied to a k -mutual neighborhood graph is thus referred to as *core-based noise removal* in this paper. The greatest advantage of the core-based noise removal is its flexibility. Unlike traditional approaches that have used hardwired thresholds to separate the data set into two parts, core-based removal allows users to assign cores with various and sometimes discontinuous core-IDs (i.e., the orders of the cores) to noise through the visualization process to be discussed below. This enables interactive participation of users in the noise-removing process and avoids troublesome parameter tuning on a trial and error basis.

3.3. Visualization

To incorporate domain knowledge, we build a visualization-supporting tool, which provides qualitative insights into the data. After being reorganized into groups of different cores, the data points can be displayed as layers in 3-dimensional (3D) space according to the hierarchy structure of the cores. The reorganized data are mapped to the three graphical dimensions of the data visualization tool in such a way that the original two dimensions of the data points are assigned to the x - and y -axis, respectively, and the core-IDs of data points to the z -axis. Additionally, different layers of data points are also rendered with different colors to better reveal the hierarchical structure.

Fig. 6 illustrates the 3D layered visualization of the core hierarchy of the benchmark, with points of different core groups displayed in different layers with different colors. The effectiveness of core-

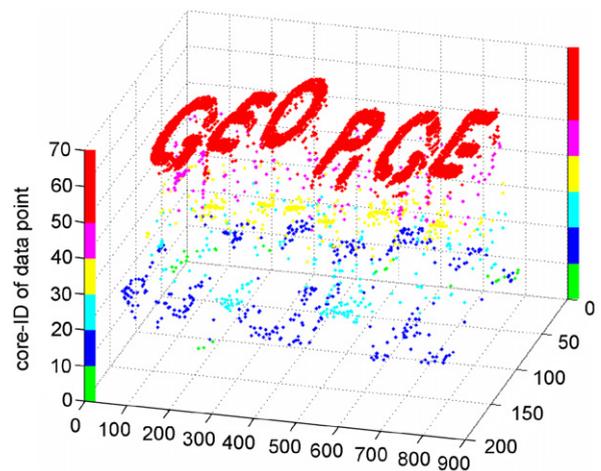


Fig. 6. A 3D visualization of core hierarchy.

based data decomposition can be manifested by the effective segregation of noise and valid data into different cores in the visualization tool. The horizontal line crossing “GEORGE” is segmented into a layer below the six letters, followed by the “salt and pepper” noise points formed as other layers under it. Through visual inspection of the layers, users can easily make a judgment as to which layer is noise or valid data. Intuitively, a user can set an appropriate threshold of core-ID defining the boundary, above which layers are retained as valid data while layers below are removed as noise. In addition, a set of ranges can also be specified to assign some of the layers as valid data and others noise. Unlike traditional noise-removing approaches that only allow a single fixed threshold to cut through data points, the proposed technique

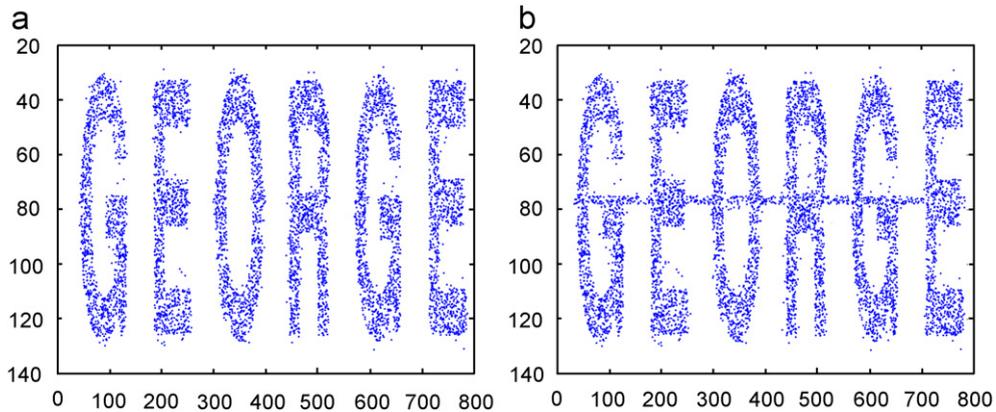


Fig. 7. Results of customizable noise removal with horizontal line treated as (a) noise and (b) true data.

provides the flexibility of defining noise at the level of individual layer. For example, the horizontal line can set either as noise to be removed or as valid data to be assembled with other desirable layers in a customizable fashion. A query can be formulated to achieve these using core-IDs. For example, a query to extract the horizontal line crossing “GEORGE” is ($coreID > 30$) and ($coreID < 36$) in the example. Fig. 7 presents two different results with the horizontal line being removed as noise or included as valid data as the two possible outcomes of the customizable process.

Fig. 8 compares the results of three different noise-removing methods applied to the CHAMELEON benchmark. Each row in Fig. 8 corresponds to one method and contains three results of this method produced with different parameters. Figs. 8(a–c) show the results of the distance-based method used by DBSCAN (Ester et al., 1996), (d–f) those of degree-based methods, and (g–i) display the results of the proposed core-based VINE approach. For each method, three different algorithm parameters are used to demonstrate the influence of parameters on the quality of corresponding outcomes. It is observed that the core-based noise-removal approach enables noise to be well segregated with all the three sets of testing parameters, while either distance-based or degree-based approaches cannot effectively differentiate noise from valid data. It is also observed that the outcomes of the core-based approach are only slightly different with different parameters, demonstrating the robustness of the approach compared with other methods. Moreover, Fig. 8 illustrates that the roles of data partitioning and the visualization tool are reciprocal. The data partitioning process preparing

cores of different core-IDs makes possible the visualization with layered structure and rendering with colors. Visualization on the other hand allows parameter selection for customizable noise removal through 3D graphic rotation and navigation of the cores, which makes the whole noise-removal process visualization-informed.

4. Post-classification smoothing of remotely sensed images

To further evaluate the effectiveness of the core-based VINE technique, we adapt it to a real-world application involving remote sensing image classification. In the context of remote sensing application, noise refers to sensor responses that are unrelated to target brightness. Noise can be originated internally by accumulated electronic errors from various components of the sensor, or introduced externally from environment, such as atmospheric attenuation and shadows from high buildings and tall trees. For a sensor system to be effectively utilized, the noise level of the acquired image should be kept small relative to signal (Campbell, 2002). This is often measured by using the signal-to-noise ratio (SNR). Various sensors may exhibit different SNR levels due to different configurations of spatial and spectral resolutions. For example, sensors such as Landsat TM of lower spatial resolution with a large instantaneous field of view permit a greater quantity of total energy to be focused on a sensor detector and allow a longer dwelling time of measurement, thus yield a signal that is much greater than background noise (i.e., higher SNR) compared with sensors of higher spatial resolution, such as IKONOS, QuickBird, and OrbView. Similarly,

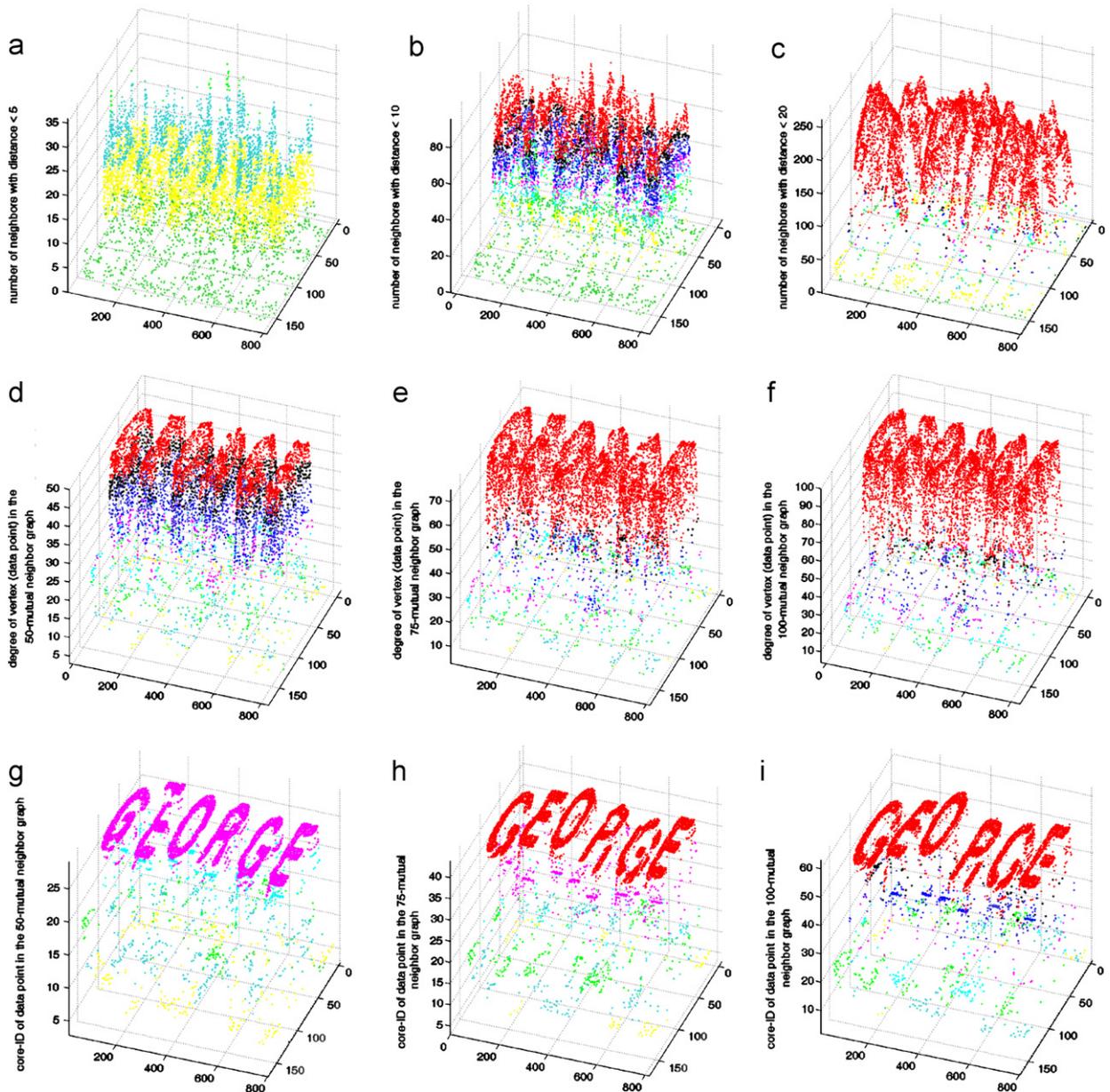


Fig. 8. Effects of distance-based (a–c), degree-based (d–f), and core-based (g–i) noise removal illustrated and compared using 3D visualization in VINE.

sensors of lower spectral resolution may have a higher SNR level with a broader wavelength band over which a given detector operates (Lillesand et al., 2004).

Traditional land use/cover classification techniques are based solely on the spectral information of single image pixels without considering their spatial context, i.e., the spatial distribution of the pixels and the spatial relationship of their corresponding geographic entities. This leads to classification

results that may be “noisy” and produces geometric outlines of land use/cover entities that may not correspond to those of their real-world counterparts, such as fields, roads, and streams (Shekhar et al., 2002). Such problems become more severe with the advent of higher-spatial-resolution remote sensing data such as those acquired by IKONOS, QuickBird, and OrbView satellites because of their lower SNR levels. The refinement of spatial resolution in these sensors that have similar wavelength

bands, in the visible-NIR spectrum, to Landsat TM also brings forth increased within-class spectral heterogeneity and between-class spectral confusion. A large amount of randomly distributed noise may be produced by per-pixel-based classifiers using only spectral information. Unlike noise seen in our benchmark images of textural and shape data set, these noise pixels may not be “erroneous” results but just unwelcome reminders of the complexity of the earth’s surface. They may be correct classifications, for example, of a single tree grown in the middle of large grassland or a small sandbar within a wide area of a water body, which may not be visible in lower-spatial-resolution images. The presence of many of these unwanted noise pixels may significantly degrade the quality of the final classification map and ultimately limit the effective use of high-spatial-resolution imagery. As a result, there will likely be a significant return to “on-screen” photo-interpretation and manual digitization of the high-resolution digital remote sensor data (Jensen et al., 2001).

To address such problems, efforts (Li and Narayanan, 2004; Shekhar et al., 2002; Zhao et al., 2003) have been made to utilize spatial contextual information to improve the accuracy of traditional per-pixel-based classifiers, among which per-field classification and object-based classification approaches (Blaschke et al., 2000) have been recently explored. Per-field classification requires prior information about an object (often with the object’s boundary predefined by a GIS), which has limited its wider application (Walter, 2004). In contrast, object-based classification does not need input from a GIS. It usually starts with grouping neighboring pixels into meaningful objects through advanced image segmentation techniques (Benz et al., 2004; De Kok et al., 1999). Classification is then performed on the grouped objects instead of individual pixels. The results of object-based classification rely heavily on the correctness of object generation. An object usually consists of a large number of pixels. If the pixels are not grouped correctly or accurately, the statistics of the object classes can be misleading, causing large areas to be misclassified.

We proposed to employ the core-based VINE method to process post-classified images produced by per-pixel-based approaches. The logic of conducting post-classification noise removal is based on three observations. Firstly, the existence of many pixels of unwanted but true small ground features

randomly distributed in high-spatial-resolution imagery, if considered during the image classification process, can be difficult to remove. The presence of unwelcome small features will not have any negative impact on the derivation of the decision rule of any image classification algorithm, because they are often ignored and could be separated and dealt with independently later after classification is completed. Secondly, the majority of pixels in an image are usually correctly classified with per-pixel-based approaches, which provide valuable contextual information for further spatial analysis. Finally, noise pixels are usually inconsistent with the pixels of their surrounding objects in spatial distribution (Qian et al., 2005). Conducting spatial analysis of the pixels can identify incorrectly classified noise and reassign them to their proper target class according to their spatial dependency with the class. Starting from the aforementioned observations, the proposed VINE approach has advantages over per-field- and object-based classification approaches. VINE technique does not require prior GIS knowledge about field boundaries, which often are difficult to obtain and thus limit per-field classification approaches from being widely applied. The risk of misclassifying a large area is reduced because the basic unit to be processed is a pixel, instead of an object as in object-based approaches.

There have been other post-classification image smoothing approaches that were utilized previously (Tomas, 1980; Townshend, 1986), primarily through a moving window-based focal majority filter where the class ID of the center pixel in the moving window is determined by majority voting of the class IDs of all the pixels in the window. The minority can then be removed as noise. The majority filter is easy to implement, but its weaknesses are also apparent. On the one hand, the size of neighborhood for the filter is fixed and must be very large for noise to be sufficiently removed; on the other hand, a neighborhood of large size may cause the boundaries between classes to be altered, creating zigzag bounding polygons (Caprioli and Tarantino, 2001).

Jensen et al. (2001) proposed a size-based filter for post-classification smoothing based on the region group operator, where regions having an area less than a user-defined threshold are regarded as noise and removed. Compared with the majority filter, size-based noise removal can better preserve details of boundaries between classes. This method requires the specification of a single threshold of region size

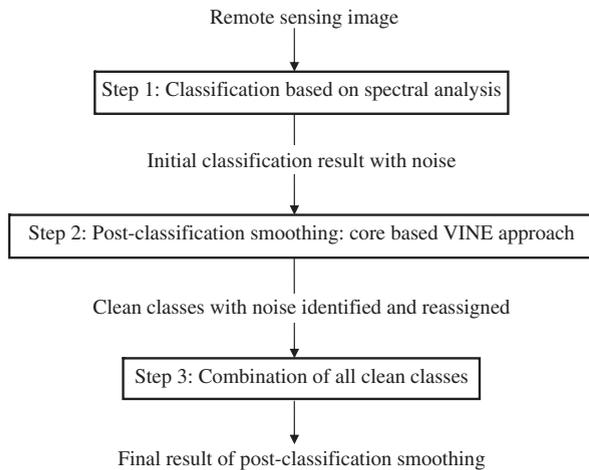


Fig. 9. Three steps of applying core-based VINE to post-classification smoothing of remote sensing image.

that applies to all classes. It, however, is possible that some small regions of a class are actually valid classes, while some larger regions of a different class can be noise. Another problem with this filter is that if the noise pixels are interlaced between two different adjacent object classes, they may remain as unclassified and not removed.

Our post-classification smoothing process is carried out in three steps as shown in Fig. 9. First, a per-pixel classifier is used to classify a remotely sensed image. Each individual class is then separated into different layers and processed with the proposed noise-removing method. Finally, the “cleaned” classes are combined to form a final map. In order for the core-based VINE technique to be applied, an adaptation of the original algorithm is needed. It is necessary to transform the classified pixels into vector data points with spatial coordinates first. The distances between data points are then calculated based on their spatial coordinates and a k -mutual neighbor graph is constructed. Unlike general noise removal that is often conducted in one layer, noise identified here in each layer is not actually removed but reallocated to another proper class. Noise reallocation is achieved through putting each noise pixel identified in one layer of a class back into other layers to recompute k -nearest neighbors. A measure based on the nearest-neighbor consistency (Fix and Hodges, 1951) is then derived to assess whether a noise pixel fits well with each class. The noise pixels will be assigned to the class with the best measure, i.e., the smallest average distance to its k -nearest neighbors among all other classes. The use of nearest-neighbor

consistency measure follows the Tobler’s (1979) First Law of Geography: everything is related to everything else, but nearby things are more related than distant things. The nearest neighbors of the noise pixels are now recomputed within each class, instead of on the whole map with a fixed moving window as in most filter-based approaches. The interactive nature of our visualization system also provides the alternative that allows users to force a noise pixel to be reallocated to any class, which is especially useful to accommodate users’ customized definition of noise.

To evaluate the effectiveness of the post-classification smoothing approach, two remotely sensed images are employed: an airborne NAPP digital orthophoto image (with 1 m spatial resolution) at Jacksonville Beach, FL as shown in Fig. 10a, and an IKONOS satellite multispectral image (with 4 m spatial resolution) at Singapore in Fig. 14(a). The classification of the two images is conducted using a neural network classifier proposed by Jensen et al. (2001) for extracting rural and urban land use and land cover information using the spectral information of each pixel. The results of initial classification of the two images are shown in Figs. 10b and 14b, respectively. Since the classification is primarily based on spectral analysis of the image only, the resulting maps are still very “noisy”.

The proposed VINE technique was first applied to the post-classification smoothing of the NAPP digital orthophoto. One noticeable type of noise resulting from the per-pixel classification map (Fig. 10b) is the sporadic shadows cast by tall trees. They are misclassified as water and are scattered inside the forest class and other classes. In addition to some newly developed residential subdivisions, small spots of new clearings can be observed within the forest regions, which are classified as urban based on their spectral characteristics. Theoretically, these pixels are not misclassified results but reminders of the complexity of the study area. In order to produce a useful and high-quality map, which is a simplified model of reality instead of the reality itself, the presence of these scattered urban pixels inside the forest class is considered unwelcome and should be removed as noise. Fig. 11 shows the four individual land use/cover classes (a) water, (b) wetland, (c) forest, and (d) urban, which are separated from the per-pixel-based classification of the NAPP digital orthophoto (Fig. 10a). K -mutual graph and core decomposition is then applied to each class by partitioning each class into many

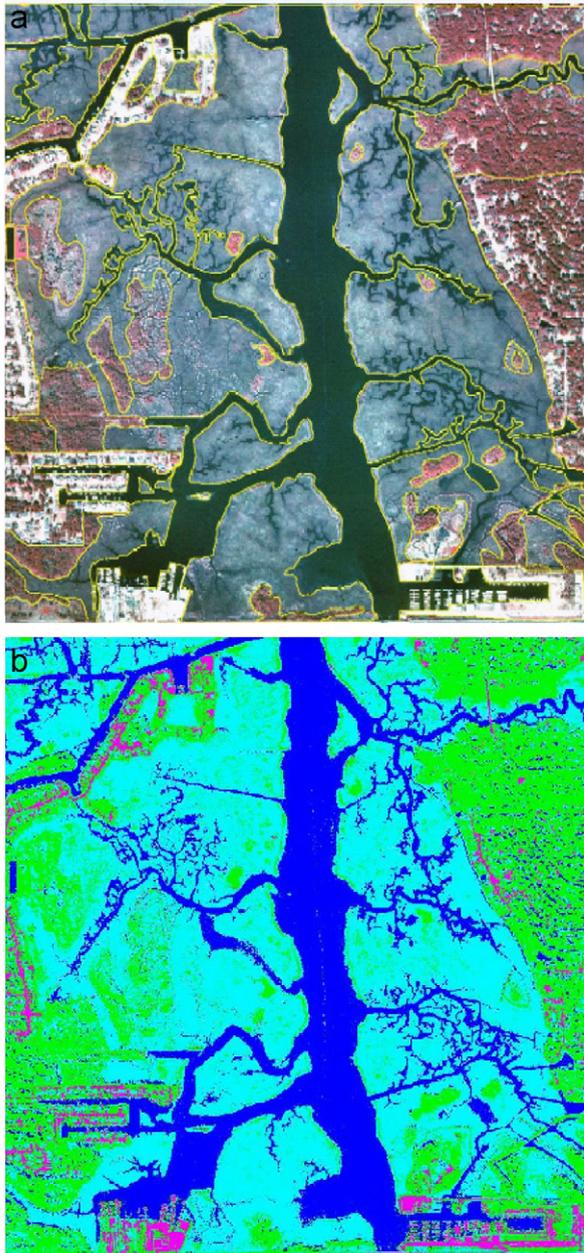


Fig. 10. (a) A digital National Aerial Photography Program (NAPP) image (with pixel size of 1×1 m) of Jacksonville Beach, FL. (b) Initial classification map (blue stands for water, green for forest, cyan for wetland and magenta for urban).

cores. The 3D visualization of the water class is shown in Fig. 12, which demonstrates that the core decomposition preserves the main body of water as layers of the higher core-IDs, which allows users to assign layers of lower core-IDs to other classes. For example, layers with core-IDs less than 10 are primarily regarded as noise for the water class and

most of them need to be reallocated to either forest or wetland class. The same process is repeated for other three classes.

After noise pixels for each class are detected and reallocated to their target classes, the individual maps of the “cleaned” classes are combined to form a final classification map. Fig. 13 provides a comparison between the results produced by (a) initial classification without post-classification smoothing, (b) focal majority filter, (c) size-based filter, and (d) core-based VINE approach. The result of focal majority filter (Fig. 13b) shows a significant boundary distortion due to the application of a fixed moving window of large size. Size-based filter demonstrates tremendous improvement, producing comparable results as that of our approach with most sporadic tree shadows and unwelcome urban spots removed. However, a close inspection of this result reveals that several small regions of wetland within the circles are incorrectly assigned to the surrounding water class (Fig. 13c) using size-based filter, which is quite obvious when compared with the initial classification map (Fig. 13a) and original image (Fig. 10a). This problem is attributed to the large threshold used in the size-based filter, which treats all small regions less than the threshold as noise, even though they might be disconnected regions belonging to the same target class. The comparison discloses an inherent weakness of the size-based approach due to the use of a fixed size threshold for all classes. This, however, is not a problem for core-based VINE smoothing, because the fixed size of the regions is not used as criteria to remove noise. The utilization of a k -mutual graph makes it possible to group small regions in heterogeneous areas into the cores of high orders, while larger regions in more homogeneous areas into the cores of a similar order. In addition to this sensitivity to patterns of spatial distribution, different and discontinuous threshold ranges are allowed for a different target class, and the spatial contextual information of the pixels (i.e., the spatial dependency in this case) is incorporated into the post-classification smoothing process. As a result, the core-based VINE smoothing method is able to produce the best result among the three approaches compared, with the details of the boundaries of the classes well preserved, sporadic shadows misclassified as water corrected, unwelcome urban spots removed, and legitimate large residential subdivisions retained, as shown in Fig. 13d.

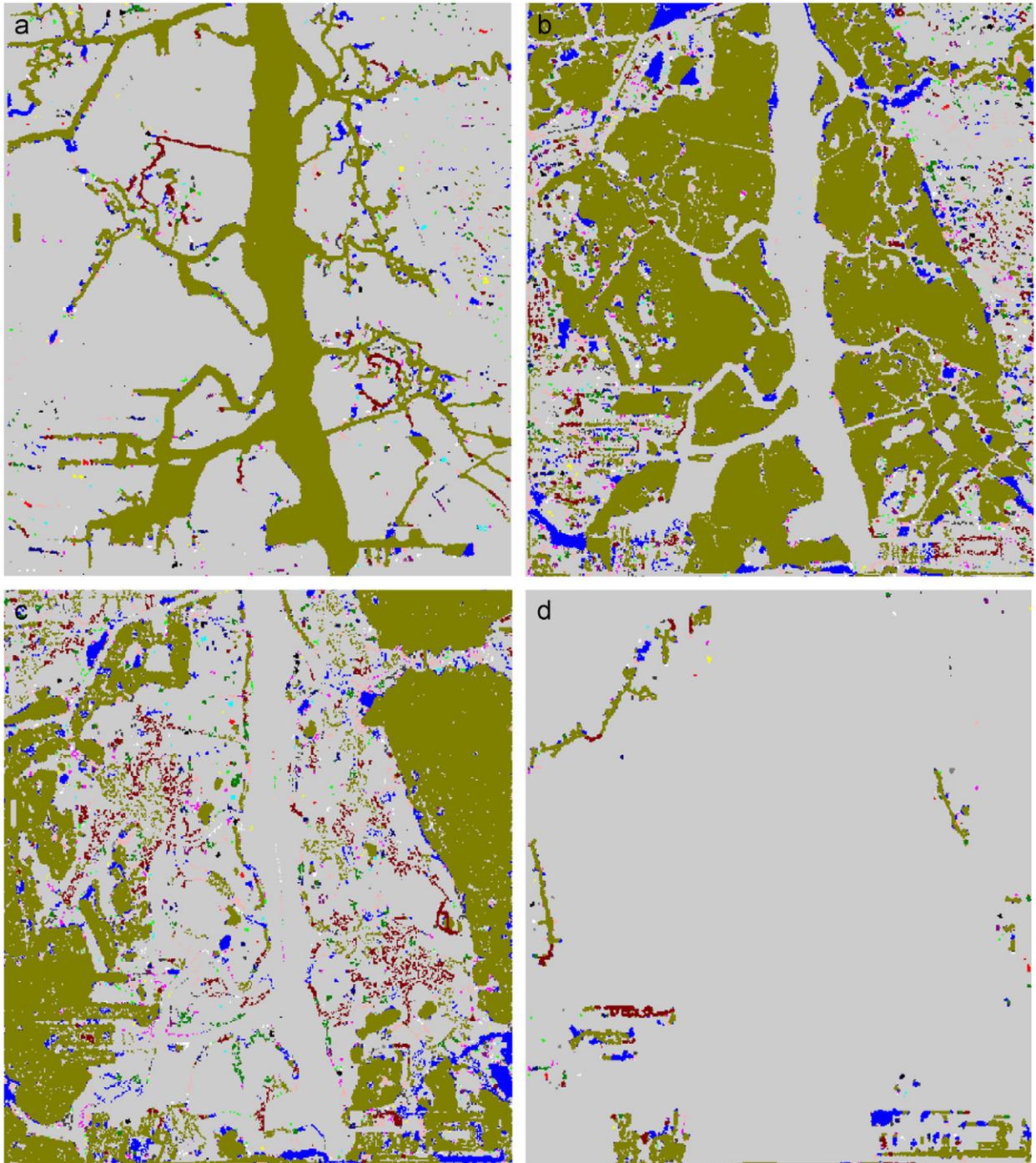


Fig. 11. Four individual classes extracted from initial classification with core decomposition applied. Different colors are used to differentiate different resulting cores. (a) Water, (b) wetland, (c) forest, and (d) urban.

We also applied the core-based VINE smoothing technique to the IKONOS high-spatial-resolution satellite image (Fig. 14a). Similar to the aforementioned NAPP image, shadows caused by tall trees

are also observed in the IKONOS image, a notorious problem for the classification of high-spatial-resolution satellite imagery. The presence of the shadows in the IKONOS image poses a bigger

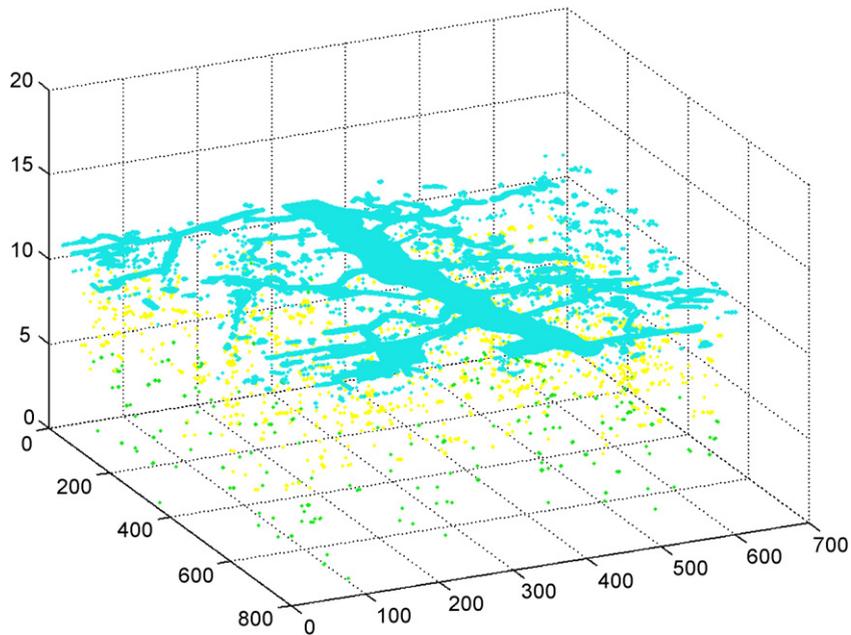


Fig. 12. An example 3D visualization of data partitions after applying core decomposition to water class.

challenge for noise removal. In addition to the sporadic shadows similar to those seen in the NAPP digital orthophoto, a new pattern of contiguous shadows extended from southwest to northeast was also detected along the highway, which is visible in the original image (Fig. 14a), but more obvious in the initial classification map (Fig. 14b) as misclassified water pixels. The highway itself, although perceptible from the IKONOS image (Fig. 14a), becomes less discernible on the classified image (Fig. 14b) due to the intrusion from surrounding trees and their shadows. Perceived also from the original image (Fig. 14a) are some newly planted individual trees, uniformly distributed inside the grass land and barren land in the center region of the image. On the initial classification map (Fig. 14b), they are classified as forest class. Spectrally, it is a correct classification but ontologically individual trees should not be regarded as forest, and should be treated as noise for the grass and barren land. Another interesting observation is over the bay area where the boats and the white wakes behind them were classified as urban because of their spectral similarity to the urban class. To produce a final classification map that is of practical use, any mobile objects (such as cars on the road, and the boats and their wakes in this case) should be removed as noise.

The result of size-based smoothing is displayed in Fig. 14c. The size threshold for a pixel region to be considered as noise was determined through multiple trial and error tests, with the one that produces the best outcome selected for the noise-removal operation. The post-classification smoothing result of the core-based VINE technique is shown in Fig. 14d. A visual comparison of the two easily reveals that the core-based VINE approach demonstrates some obvious superiorities to the size-based approach in many aspects. First, not only the sporadic shadows of individual trees that were misclassified as water are successfully identified and removed, but also the contiguous shadows extended along the highway were also effectively eliminated as shown in Fig. 14d. The 3D visualization of the water class cores obtained from initial classification is shown in Fig. 15a, where the third dimension is based on the core-IDs of the decomposed cores. The sporadic shadows (in green), contiguous shadow (in red), and the true water bodies (in blue) are well segregated from each other as different layers with various core-IDs. With the support of the visualization tool, the core-based decomposition makes it possible for an image analyst to bring his/her domain knowledge into the decision making process on whether a certain layer should be kept as water class or be regarded as

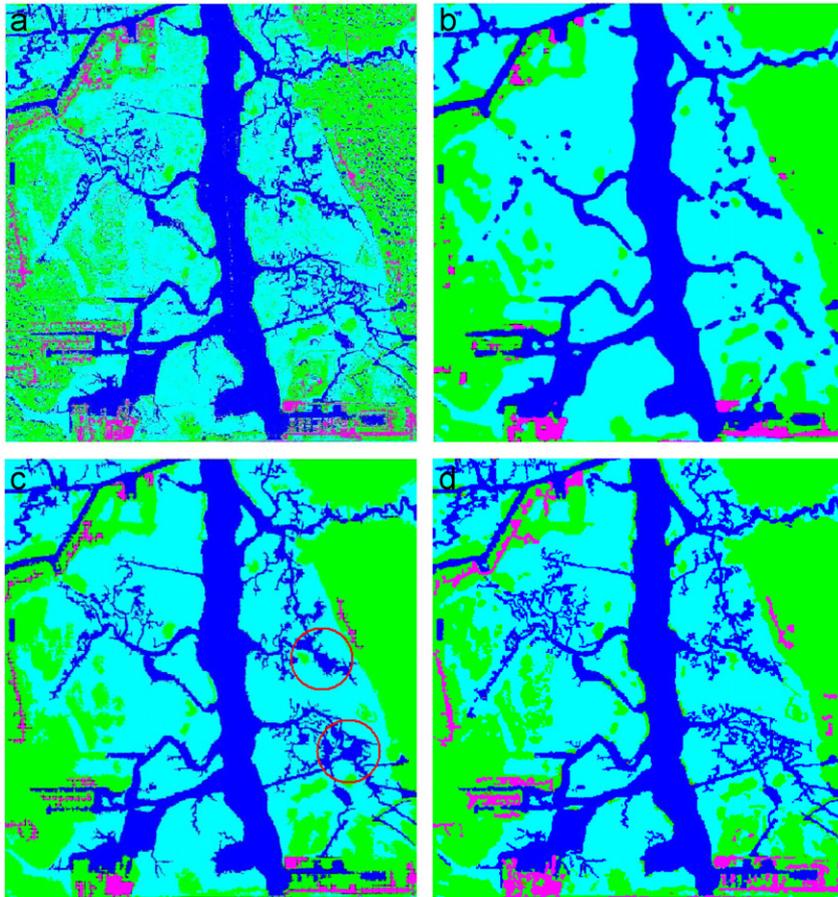


Fig. 13. A comparison of results obtained from applying (a) initial classification, (b) focal majority smoothing, (c) size-based smoothing, and (d) core-based VINE smoothing.

noise to be reassigned to another class. It is not difficult to understand that the mutual k -nearest-neighbor algorithm embedded in the VINE technique could easily separate the shadows from water, because the algorithm utilized spatial proximity to form graphs and most of the shadows are far away from the water bodies. However, just like the horizontal line crossing “GEORGE” in the benchmark data set shown in Fig. 1, the contiguous shadows intermingle with other features (such as the sporadic shadows) in the 2-dimensional (2D) space (Fig. 15b). In addition, sporadic and contiguous shadows are spectrally identical, making it very difficult to separate them. However, the core decomposition process could differentiate these two types of shadows by taking advantage of their difference in the patterns of spatial distribution, just like the algorithm could exploit the different spatial patterns between “GEORGE” and the crossing line

in the benchmark data set. The 3D visualization makes it possible for different types of noise to be grouped into dissimilar layers and treated differently if necessary, although they all would be reallocated to forest class in this case. While, in other cases such as an urbanized area, shadows cast by tall trees and high-rise buildings could be reassigned to urban class. In contrast, the size-based post-classification smoothing (Fig. 14c), although could remove most of the sporadic shadow with the exception of some at the northwest part of the image, failed to eliminate the major contiguous shadows along the highway and several other places. The size-based approach that uses a fixed size threshold for all noise types of all classes, could not handle sporadic shadows and contiguous shadows differently.

In order to completely eliminate the contiguous shadows, the size-based approach needs a much

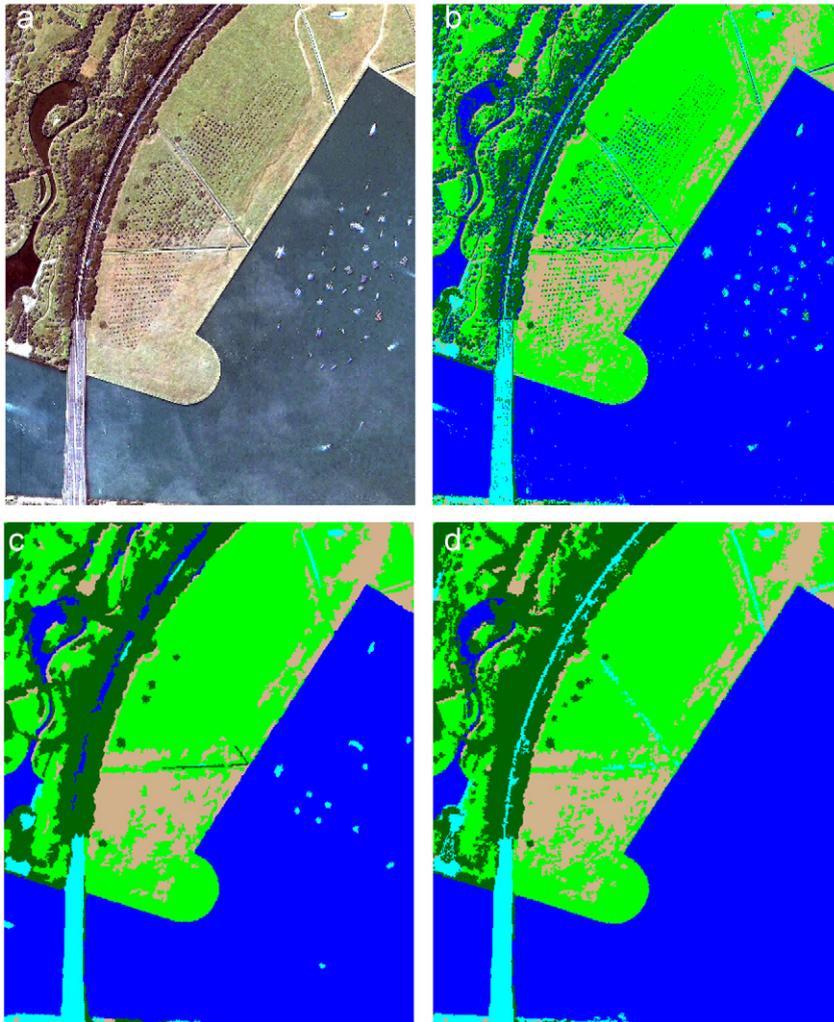


Fig. 14. (a) An IKONS high-spatial-resolution multispectral satellite image (with pixel size of 4×4 m) at Singapore. (b) Initial classification (blue stands for water, green for grass, dark green for tree, cyan for urban, and pink for barren land), (c) size-based smoothing result, and (d) core-based VINE smoothing result.

larger size threshold. As a consequence, regions of valid classification with an area less than the threshold could be eradicated unintentionally. This actually has already occurred with a small threshold that could not even remove the contiguous shadows. Most of the highway segments that pass through the forest region were inadvertently eliminated by the size-based approach (Fig. 14c). Due to the intrusion of surrounding trees and their shadows, the originally uninterrupted highway (Fig. 15a) is now broken into many disjointed pieces of urban pixel regions after classification (Fig. 15b). Most of these regions have an area less than the size threshold and therefore are discarded as noise by the size-based

approach (Fig. 14c). In the core-based VINE technique however, these disjointed pieces were able to constitute a big core of a continuous linear feature. This is evident when different core layers are rendered with different colors in the 3D visualization tool (Fig. 15c) and the corresponding 2D view. It is interesting to note from both Figs. 15c and d that the highway is in the same red layer as the bridge, indicating that they are grouped into similar cores in the urban class and are well separated from other urban features.

For both the size-based and core-based smoothing approaches, the removal of the individual trees from the forest class as unwanted noise was an easy

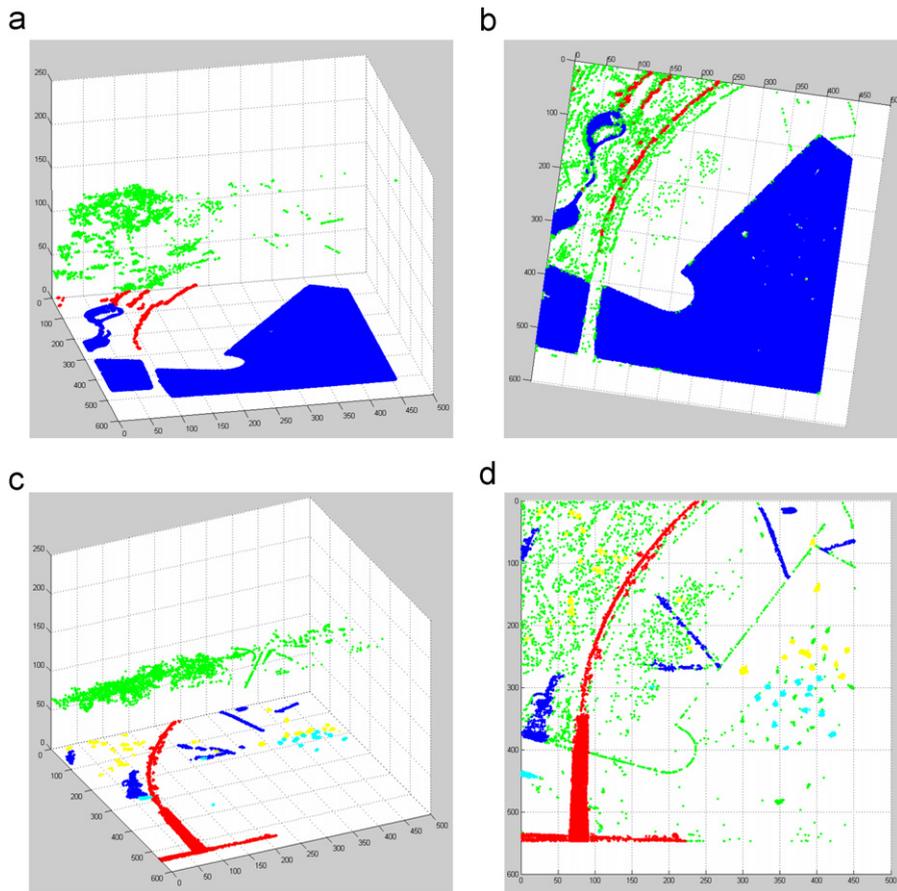


Fig. 15. (a) A 3D visualization of water class based on core-IDs showing that core decomposition separates shadows from true water with true water in blue and shadows in red and green; (b) corresponding 2D visualization of water class showing spatial distribution of shadows and true water bodies. (c) A 3D visualization of urban class with highway and bridge in red, country road and other urban area in blue, while boats and their wakes and other noise in cyan, yellow, and green; (d) corresponding 2D visualization of urban class.

task (Figs. 14c and d). However, the sized-based smoothing has the difficulty eliminating all the boats and their wakes from the urban class (Fig. 14c), because some of them have an area greater than the size threshold selected. This is not an issue for the core-based VINE technique as shown in Fig. 14d. An examination of their corresponding cores in 3D view of the urban class (Fig. 15d) explains why the boats and the associated wakes were completely removed by the core-based smoothing. They have different core-IDs (as well as rendering colors) with those of the highway, the bridge and other valid urban cores and therefore can be easily removed from urban class and reassigned to water class. Finally, it is worth mentioning that boundaries and linear features can be better preserved in core-based smoothing, which can be seen from Fig. 14d where the details of

shorelines of the two small lakes at the left part of the image, the three country roads crossing the grassland, and the coastline at the lower left corner of the picture are very well maintained when compared with those from size-based smoothing.

In addition to the above visual comparison, we also conducted statistical accuracy assessment in order to make a more objective evaluation. A total of 323 test points were selected for the accuracy assessment with 200 points directly generated for the whole image using stratified random sampling. However, only a few of these points fell into the areas where the noises were removed by the two post-classification smoothing approaches. In order to achieve a more effective accuracy assessment of these two noise-removal strategies, two masks that correspond to the difference between the original classified image and the two respective

Table 1a
Accuracy assessment of initial classification (Fig. 14b)

Classification error matrix						
	1	2	3	4	5	
1	72	6	17	0	2	97
2	1	59	4	14	4	82
3	7	12	30	1	5	55
4	2	12	2	24	4	44
5	10	2	4	0	29	45
Column total	92	91	57	39	44	323

Overall accuracy = 66.25%

	Producer's accuracy (%)	User's accuracy (%)	
1	78.26	74.23	
2	64.84	71.95	
3	52.63	54.55	
4	61.54	54.55	
5	65.91	64.44	
Conditional kappa is	0.64	0.61	0.450.480.59

Kappa = 0.5659, Variance = 0.001116

Table 1b
Accuracy assessment of size-based smoothing (Fig. 14c)

Classification error matrix						
	1	2	3	4	5	
1	83	1	15	0	7	106
2	0	73	8	8	3	92
3	3	6	34	0	5	48
4	1	9	0	31	3	44
5	5	2	0	0	26	33
Column total	92	91	57	39	44	323

Overall accuracy = 76.47%

	Producer's accuracy (%)	User's accuracy (%)	
1	90.22	78.30	
2	80.22	79.35	
3	59.65	70.83	
4	79.49	70.45	
5	59.09	78.79	
Conditional kappa is	0.70	0.71	0.65 0.66 0.75

Kappa = 0.6943, Variance = 0.000914

smoothed images were created, respectively. Then these two masks were combined for use to generate another stratified random sampling, from which 123 additional points were randomly created.

Table 1 displays overall accuracy and the Kappa coefficients (K_{hat} statistics) of the original classified image and two smoothed images. The overall accuracy of the initial classification, the size-based approach, and the core-based technique was 66.25%, 76.47%, and 91.95%, respectively. Kappa statistic was also computed, which are regarded as better representation of the general quality of a classification. The Kappa statistic for the core-based VINE smoothing is 0.8966. The Kappa statistic for the size-based noise removal is only 0.6943, while that for the initial classification was 0.5659. According to Fleiss (1981), Kappa coefficients >0.75 suggest strong agreement above chance. Landis and Koch (1977) suggest that Kappa coefficients >0.81 are almost perfect. To statistically compare Kappa coefficients, a significance Z-test was performed between the Kappa statistics of the two smoothing approach. A Z-score of 5.632 is achieved, suggesting the core-based smoothing method was significantly better than size-based approach at the 99% confidence level.

Table 1c
Accuracy assessment of core-based VINE smoothing (Fig. 14d)

Classification error matrix						
	1	2	3	4	5	
1	82	0	0	1	0	83
2	1	85	3	3	0	92
3	3	2	53	0	1	59
4	2	4	0	35	1	42
5	4	0	1	0	42	47
Column total	92	91	57	39	44	323

Overall accuracy = 91.95%

	Producer's accuracy (%)	User's accuracy (%)	
1	89.13	98.80	
2	93.41	92.39	
3	92.98	89.83	
4	89.74	83.33	
5	95.45	89.36	
Conditional kappa is	0.98	0.89	0.88 0.81 0.88

Kappa = 0.8966, Variance = 0.000376

5. Conclusion

The core-based VINE technique is proposed based on a three-step procedure consisting of graph

construction, core partitioning, and data visualization to inform a customized noise removal. For n data point, constructing a k -mutual graph of a 2D data set requires $O(n \log n)$ time (Xu et al., 1998) when using a k - d tree; core decomposition can be completed in $O(n)$ time based on the definition of k -mutual graph (Batagelj et al., 2000). The generation of 3D layered display is a simple point plotting based on the core hierarchy. In summary, all these operations can be completed within $O(n \log n)$ time. The actual running time of the program applied to the two remotely sensed images is both less than 2 h. In contrast, visual image interpretation and manual digitization for the same image require 2–3 person days to achieve a similar quality (Jensen et al., 2001). For space complexity, the implementation requires $O(kn)$ space to compute and store the list of adjacency links of the k -mutual graph of the whole data set with n data points.

The construction of the k -mutual graph allows the effective segregation of noise from valid data in areas of different size and densities. The data partitioning and visualization are functionally reciprocal. Data partitioning reorganized the data for visualization, and the visualization of the data in the resultant layered structure supports an informed noise removal and reassembling of the remaining data according to custom domain requirements. With such visualization support, the parameter determination for noise removal is much straightforward and no longer a trial and error process.

The proposed approach is adapted to improve post-classification smoothing of two remotely sensed images. Both cases demonstrate the advantages of the core-based VINE approach over previous methods through side-by-side visualization comparisons. Statistical accuracy assessment was also performed on the IKONOS image to evaluate the apparent improvement of the proposed technique over the size-based approach. Experimental results show that the core-based VINE technique can successfully discover and reassign noise pixels to their proper target classes and improve the overall classification accuracy of the high-spatial-resolution remote sensing images. Compared with other post-classification smoothing approaches, the core-based VINE relies on neither a fixed moving window, nor a fixed threshold of region size, and can effectively differentiate and remove shadows as noise through a flexible and customizable process informed by domain knowledge through visualization support. By capitalizing on the spatial contextual information

embedded in a k -mutual graph, the core-based VINE is also able to better preserve the details of the between-class boundary than the focal majority filter and the size-based filter. Compared with per-field and object-based classification approaches, core-based VINE does not require any prior GIS boundary and avoids the possible misclassification of large areas.

References

- Aggarwal, C., Yu, P., 2001. Outlier detection for high dimensional data. In: Proceedings of the 2001 International Conference on Management of Data, Santa Barbara, CA, USA, pp. 37–46.
- Ankerst, M., Breunig, M., Kriegel, H.P., Sander, J., 1999. Optics: ordering points to identify the clustering structure. In: Proceedings of the 1999 International Conference on Management of Data, Philadelphia, PA, USA, pp. 49–60.
- Batagelj, V., Mrvar, A., Zaversnik, M., 2000. Partitioning approaches to clustering in graphs. In: Proceedings of the Seventh International Symposium on Graph Drawing, Stirin Castle, Czech Republic, LNCS, pp. 90–97.
- Bay, S., Schwabacher, M., 2003. Mining distance based outliers in near linear time with randomization and a simple pruning rule. In: Proceedings of the Ninth International Conference on Knowledge Discovery and Data Mining, Washington, DC, USA, pp. 29–38.
- 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, 239–258.
- Blaschke, T., Lang, S., Lorup, E., Strobl, J., Zeil, P., 2000. Object-oriented image processing in an integrated GIS/remote sensing environment and perspectives for environmental applications. In: Cremers, A., Greve, K. (Eds.), *Environmental Information for Planning, Politics, and Public II*. Metropolis, Marburg, pp. 555–570.
- Breunig, M., Kriegel, H., Ng, R., Sander, J., 2000. LOF: identifying density based local outliers. In: Proceedings of the 2000 International Conference on Management of Data, Dallas, TX, USA, pp. 93–104.
- Campbell, J.B., 2002. *Introduction to Remote Sensing*, third ed. The Guilford Press, New York, London, 621pp.
- Caprioli, M., Tarantino, E., 2001. Accuracy assessment of per-field classification integrating very fine spatial resolution satellite imagery with topographic data. *Journal of Geospatial Engineering* 3, 127–134.
- Chang, S., Yu, B., Vetterli, M., 1998. Spatially adaptive wavelet thresholding with context modeling for image denoising. In: Proceedings of the 1998 International Conference on Image Processing, Chicago, IL, USA, vol. 1, pp. 535–539.
- De Kok, R., Schneider, T., Ammer, U., 1999. Object based classification and applications in the alpine forest environment. *International Archives of Photogrammetry and Remote Sensing* 32 (Part 7-4-3 W6).
- Ertoz, L., Steinbach, M., Kumar, V., 2003. Finding clusters of different sizes, shapes, and densities in noisy, high

- dimensional data. In: Proceedings of the 2007 SIAM International Conference on Data Mining, San Francisco, CA, USA, pp. 41–50.
- Ester, M., Kriegel, H.P., Sander, J., Xu, X., 1996. A density based algorithm for discovering clusters in large spatial databases with noise. In: Proceedings of the Second International Conference on Knowledge Discovery and Data Mining, Portland, OR, USA, pp. 226–231.
- Fan, G., Xia, X., 2000. Wavelet based statistical image processing using Hidden Markov Tree model. In: Proceedings of the 34th Annual Conference on Information Sciences and Systems, Princeton, NJ, USA, pp. 31–36.
- Fix, E., Hodges, J.L., 1951. Discriminatory analysis, nonparametric discrimination: consistency properties. Technical report 4, project number 21-49-004, USAF School of Aviation Medicine, Randolph Fields, TX.
- Fleiss, J.L., 1981. Statistical Methods for Rates and Proportions, second ed. Wiley, New York, 321pp.
- Han, J., Kamber, M., 2000. Data mining: Concepts and Techniques. Morgan Kaufmann/Elsevier, Los Altos, CA/Amsterdam, 500pp.
- Harel, D., Koren, Y., 2001. Clustering spatial data using random walks. In: Proceedings of the Seventh International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, pp. 281–286.
- Hinneburg, A., Keim, D.A., 1998. An efficient approach to clustering in large multimedia databases with noise. In: Proceedings of the Fourth International Conference on Knowledge Discovery and Data Mining, New York City, NY, USA, pp. 58–65.
- Jain, A.K., Dubes, R.C., 1988. Algorithms for Clustering Data. Prentice-Hall Advanced Reference Series. Prentice-Hall, Inc., Upper Saddle River, NJ, 320pp.
- Jain, A.K., Murty, M.N., Flynn, P.J., 1999. Data clustering: a review. ACM Computing Surveys 31, 264–323.
- Jensen, J.R., Qiu, F., 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 16, 1–10.
- Karypis, G., Han, E., Kumar, V., 1999. CHAMELEON, a hierarchical clustering algorithm using dynamic modeling. IEEE Computer 32, 68–75.
- Knorr, E., Ng, R., Tucakov, V., 2000. Distance based outliers: algorithms and applications. The International Journal on Very Large Data Bases 8, 237–253.
- Landis, J., Koch, G.G., 1977. The measurement of observer agreement for categorical data. Biometrics 33, 159–174.
- Lang, M., Guo, H., Odegard, J.E., Burrus, C.S., Wells, R.O., 1996. Noise reduction using an undecimated discrete wavelet transform. IEEE Signal Processing Letters 3, 10–12.
- Lebart, K., Boucher, J.M., 1996. Speckle filtering by wavelet analysis and synthesis. Wavelet Applications in Signal Image Process, IV, SPIE 2825, 644–651.
- Li, J., Narayanan, R.M., 2004. Integrated spectral and spatial information mining in remote sensing imagery. IEEE Transactions on Geoscience and Remote Sensing 42, 673–685.
- Li, T., Li, Q., Zhu, S., Ogihara, M., 2002. A survey on wavelet applications in data mining. SIGKDD Explorations 4, 49–66.
- Lillesand, T.M., Kiefer, R.W., Chipman, J.W., 2004. Remote Sensing and Image Interpretation, fifth ed. Wiley, New York, NY, 763pp.
- Mallat, S.G., 1989. A theory for multiresolution signal decomposition: the wavelet representation. IEEE Transactions on Pattern Analysis and Machine Intelligence 11, 674–693.
- Qian, Y., Zhang, K., 2004. Discovering spatial patterns accurately with effective noise removal. In: Proceedings of the Ninth ACM SIGMOD Workshop on Research Issues in Data Mining and Knowledge Discovery, Paris, France, pp. 43–50.
- Qian, Y., Zhang, G., Zhang, K., 2004. FACADE: a fast and effective approach to the discovery of dense clusters in noisy spatial data. In: Proceedings of the 2004 International Conference on Management of Data, Paris, France, pp. 921–922.
- Qian, Y., Zhang, K., Qiu, F., 2005. Spatial contextual noise removal for post-classification smoothing of GIS images. In: Proceedings of the 2005 ACM Symposium on Applied Computing, Santa Fe, NM, USA, pp. 524–528.
- Seidman, S.B., 1983. Network structure and minimum degree. Social Networks 5, 269–287.
- Sheikholeslami, G., Chatterjee, S., Zhang, A., 2000. WaveCluster: a wavelet based clustering approach for spatial data in very large databases. The International Journal on Very Large Databases 8, 289–304.
- Shekhar, S., Lu, C., Zhang, P., 2001. Detecting graph based spatial outliers: algorithms and applications (a summary of results). In: Proceedings of the Seventh International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, pp. 371–376.
- Shekhar, S., Schrater, P.R., Vatsavai, R.R., Wu, W., Chawla, S., 2002. Spatial contextual classification and prediction models for mining geospatial data. IEEE Transactions on Multimedia 4, 174–188.
- Strela, V., 2000. Denoising via block wiener filtering in wavelet domain. In: Proceedings of the Third European Congress of Mathematics, Barcelona, Birkhauser Verlag.
- Tobler, W.R., 1979. Cellular geography. In: Gale, W.R., Olsson, W.R. (Eds.), Philosophy in Geography. Reidel, The Netherlands, pp. 379–386.
- Tomas, I.L., 1980. Spatial postprocessing of spectrally classified landsat data. Photogrammetric Engineering and Remote Sensing 46, 1201–1206.
- Townshend, J.R.G., 1986. The enhancement of computer classification by logical smoothing. Photogrammetric Engineering and Remote Sensing 52, 213–221.
- Vidal-Pantaleoni, A., Marti, D., 2004. Comparison of different speckle-reduction techniques in SAR images using wavelet transform. International Journal of Remote Sensing 25, 4915–4932.
- Walter, V., 2004. Object based classification of remote sensing data for change detection. ISPRS Journal of Photogrammetry and Remote Sensing 58, 225–238.
- Xu, X., Ester, M., Kriegel, H., Sander, J., 1998. A distribution based clustering algorithm for mining in large spatial databases. In: Proceedings of the 14th International Conference on Data Engineering, Orlando, FL, USA, pp. 324–331.
- Zhao, J., Lu, C.T., Kou, Y., 2003. Detecting region outliers in meteorological data. In: Proceedings of the 11th International Symposium of ACM GIS, Atlanta, GA, USA, pp. 49–55.
- Zhang, T., Ramakrishnan, R., Linvy, M., 1996. BIRCH: an efficient data clustering method for very large databases. In: Proceedings of the 1996 International Conference on Management of Data, Montreal, Canada, pp. 103–114.