Journal of
Applied Remote Sensing

Cultivated land information extraction from high-resolution unmanned aerial vehicle imagery data

Lei Ma
Liang Cheng
Wenquan Han
Lishan Zhong
Manchun Li

SPIE

Cultivated land information extraction from high-resolution unmanned aerial vehicle imagery data

Lei Ma,[a,b] Liang Cheng,[a,b,*] Wenquan Han,[a,c] Lishan Zhong,[a,b] and Manchun Li[a,b,*]

[a]Nanjing University, Jiangsu Provincial Key Laboratory of Geographic Information Science and Technology, 22 Hankou Road, Nanjing 210097, China

[b]Nanjing University, Collaborative Innovation Center for the South China Sea Studies, 22 Hankou Road, Nanjing 210097, China

[c]Nanjing Institute of Surveying, Mapping & Geotechnical Investigation, Co. Ltd, 88 ChuangYi Road, Nanjing 210019, China

Abstract. The development of precision agriculture demands high accuracy and efficiency of cultivated land information extraction. Simultaneously, unmanned aerial vehicles (UAVs) have been increasingly used for natural resource applications in recent years as a result of their greater availability, the miniaturization of sensors, and the ability to deploy UAVs relatively quickly and repeatedly at low altitudes. We examine the potential of utilizing a small UAV for the characterization, assessment, and monitoring of cultivated land. Because most UAV images lack spectral information, we propose a novel cultivated land information extraction method based on a triangulation for cultivated land information extraction (TCLE) method. Thus, the information on more spatial properties of a region is incorporated into the classification process. The TCLE comprises three main steps: image segmentation, triangulation construction, and triangulation clustering using AUTOCLUST. Experiments were conducted on three UAV images in Deyang, China, using TCLE and eCognition for cultivated land information extraction (ECLE). Experimental results show that TCLE, which does not require training samples and has a much higher level of automation, can obtain accuracies equivalent to ECLE. Comparing with ECLE, TCLE also extracts coherent cultivated land with much less noise. As such, cultivated land information extraction based on high-resolution UAV images can be effectively and efficiently conducted using the proposed method. [DOI: 10.1117/1.JRS.8.083673]

Keywords: unmanned aerial vehicle; AUTOCLUST; cultivated land; object-based classification; Voronoi diagram.

Paper 13296 received Aug. 9, 2013; revised manuscript received Jan. 23, 2014; accepted for publication Jan. 27, 2014; published online Feb. 20, 2014.

1 Introduction

Precision agriculture is an emerging farm management strategy that is having a profound effect on the way people farm.[1] Maps of the distribution of cultivated land are basic requirements for any management strategy for cultivated land.[2,3] The economical and feasible use of remote sensing technology to quickly and accurately extract cultivated land information has become a focus for research, and it continues to present difficulties.[4] Cultivated land (of a minimum polygon area of 400 m^2) has been defined in a study monitoring the national geographic conditions of China, for its 12th Five-Year Plan. It is defined as the land on which crops have been planted after reclamation and which has been frequently cultivated and managed. In China, this mainly comprises paddy fields and dry farmed land.

*Address all correspondence to: Liang Cheng, lcheng@nju.edu.cn; Manchun Li, limanchun_nju@126.com

0091-3286/2014/$25.00

Currently, the methods employed for mapping of cultivated land mainly employ remote sensing technologies, and generally, the monitoring of cultivated land makes use of data derived from satellite remote sensing.[5,6] However, there are several limitations to satellite imagery. Due to the low spatial resolution, it is difficult to identify cultivated land of small areal extent in critical regions; this requires image data of higher spatial resolution, and the development of a supporting cultivated land information extraction method, for specific or general cases.[7] In addition, because precision agriculture requires small-scale and high-accuracy mapping of cultivated land, the low spatial and temporal resolutions of satellite images are unable to meet the high demands of this application. In this situation, the emergence of unmanned aerial vehicle (UAV) technology offers a new opportunity for the mapping of cultivated land.

UAVs are an important new remote sensing approach, and there is increasing interest in their use for obtaining very large-scale aerial imagery for assessment and monitoring of cultivated land.[8–10] UAVs are a flexible image acquisition tool for such a purpose. UAVs have several advantages over piloted aircraft: they can be deployed quickly and repeatedly, for example, to assess flooding after sudden rainfall events; there is no risk to an on-board pilot; and they can be used to obtain very high-resolution imagery (VHR) with lower image acquisition costs than with piloted aircraft. As such, UAVs are well suited for cultivated land remote sensing applications. In addition, in China, the current monitoring of the national geographic conditions for China's 12th Five-Year Plan demands a more efficient multiscale monitoring technology for use nationwide. For large-scale extraction of cultivated land information, UAVs have the potential to reduce the number of ground-based measurements required for obtaining a cultivated land inventory, as well as assessment and monitoring, by providing low-cost, subdecimeter-resolution digital imagery.

Image acquisition is limited by the payload of the current UAVs. At present, only lightweight commercial cameras can be carried, and consequently, the acquired images lack spectral information, and a large number of research goals common to remote sensing technology cannot be applied effectively and directly to this new data source. In particular, vegetation indices, such as the normalized difference vegetation index (NDVI), cannot be calculated, which would have great significance for vegetation extraction. In addition, it is still difficult to extract accurate cultivated land covers from UAV images based on the conventional pixel-based classification methods. Using only spectral information, these methods have only limited success for extraction of cultivated land information,[11] especially when the optical images are from UAVs, because their spectral information is not sufficient, and it is difficult to distinguish between dry lands and paddy fields by means of spectral information alone. An object-based image analysis (OBIA) is a more suitable method than pixel-based classification for high- and very-high-resolution imagery.[12–14] OBIA has been widely discussed with regard to processing VHR images and many methods have been developed to obtain appropriate objects.[14–20] Recently, OBIA has also been used successfully for analysis of UAV imagery over rangelands.[21–24]

For the extraction of cultivated land information from UAV images, object-oriented classification methods are recommended as an alternative solution to processing UAV images because they can incorporate a maximal amount of information on the spatial properties of a region into the classification process. Thus, a method for extraction of cultivated land information is proposed based on segmentation of high-spatial-resolution UAV images and detection of spatial features in a triangulation network. The framework for application of UAV data comprises data acquisition, data processing, and applications. The goal of this study is to test the feasibility of a UAV-based cultivated land assessment and monitoring approach, and to optimize this new method for extracting cultivated land information from VHR imagery in flat areas. This paper describes the details for acquisition and processing of UAV imagery, consisting of mission planning, image acquisition, orthorectification, and mosaicking and proposes a triangulation-network-based cultivated land information extraction method, as well as reporting and analyzing the results of the image analysis. In addition to communicating recent advances in the development and application of these technologies, this paper should also serve as a guide for individuals and organizations that are seeking to use UAV in the future.

2 Related Work

2.1 *Research Progress in Cultivated Land Information Extraction*

In the extraction and mapping of cultivated land information, studies in different areas have been carried out on different spatial scales, most often divided into local, regional, and global areas.[25] In order to achieve the best extraction of cultivated land information at different spatial scales, different sensor data and different mapping methods are required. Many researchers' results conclude that extraction of cultivated land information over a local area should make use of existing datasets, such as Landsat TM/ETM+, SPOT, LISS, ASTER, CBERS, and THEOS, to make analysis more cost-effective. Information extraction methods have typically involved photograph interpretation, digital image computation, image segmentation, classification, and image fusion. Landsat TM/ETM+, MODIS, MERIS, AVHRR, and SPOT VGT data are more typically applied to regional areas, and cultivated land information is identified through the time-series analysis method, the supervised or unsupervised classification method, and the mask method. Over a global area, MODIS, MERIS, AVHRR, and SPOT VGT data are most often recommended, and information extraction methods incorporate unsupervised clustering (ISODATA), machine learning algorithms for time-series data, and other auxiliary data (including statistical data and real-time ground data). Ozdogan et al.[7] have summarized the advantages and disadvantages of the sensors of different satellite image acquisition systems for the mapping of cultivated land information, including Landsat, SPOT, CBERS, AVHRR, MODIS, MERIS, and IRS. In addition, synthetic aperture radar (SAR) has been widely applied to research on rice crops because of its ability to observe in all weathers in both day and nighttime. Studies have already been carried in rice-growing areas, involving mapping and growth monitoring by means of multitemporal radar data. These studies show that multitemporal SAR data are very effective for distinguishing irrigated rice fields from other farmland types.[26–30] Currently, information extraction in combination with a variety of other data sources is also becoming more important.[31–33]

In addition, the implementation of a high-resolution Earth observation system—a special project under China's Eleventh Five-Year Plan—has contributed to the further development of VHR image data acquisition technology in China. This study into extraction technologies for high-resolution cultivated land information will promote the application of remote sensing technology in the agricultural sciences. Several researchers have already conducted studies in this field. For example, Liu et al.[34] make use of aerial imaging data (at 3-m resampling resolution) to extract information on high-quality prime farmland through segmentation (object segmentation through edge detection), feature analysis (calculating eigenvalues), and construction of rules (using the C4.5 to build decision trees). In contrast, for high-resolution satellite image data, an object-oriented classification method, which aims to improve the mean shift algorithm and also supports vector machine, has been proposed and applied to the extraction of cultivated land information.[35] Lu et al.[36] make use of hyperspectral data (with ground resolution of 2×2 m) to extract areas covered by complex agricultural land. Although these approaches have focused on extraction of cultivated land information by means of high-resolution data, they all have certain limitations to a greater or lesser extent: they either lack pertinence or apply the same method to different datasets and do not consider the unique features of specific cultivated land plots, ultimately yielding unsatisfactory results. Sun and Xu,[37] focusing on the Quickbird panchromatic band image, first segmented the image data and then extracted information from agricultural land plots of regular shape, in combination with the corner-feature points (following the cultivated land appearance rule, generally incorporating four or more corner-feature points). The approach of Sun and Xu[37] offers a new perspective for the extraction of cultivated land information from high-resolution panchromatic images by designing algorithms in a way that makes full use of geometric and textural features of cultivated land in high-resolution image data, as well as the unique properties of cultivated land, in order to develop a rapid and efficient method for extracting cultivated land plots from high-resolution image data.

2.2 *Progress in the Extraction of Typical Surface Features and Cultivated Land from UAV Data*

The number of potential applications for UAV has increased considerably in recent years as a result of their greater availability and the miniaturization of sensors, GPS, inertial measurement units, and other hardware.[22,38,39] The combination of UAVs and remote sensing technology has enabled the technologies to be used to acquire spatial data on land cover, resources, and environments for processing remote sensing data. Because of the high frequency and high resolution of data obtained from UAV, they have received more attention from researchers and manufacturers, which has further expanded the potential range of applications and users.

Currently, UAV are predominantly used as a photogrammetric data acquisition platform to map and monitor. The rectified images and their derivatives, such as image mosaics, maps, and drawings, can be used for interpretation and analysis. For example, research groups have focused on applications that make use of rotary and fixed-wing UAV for forest resource management and monitoring,[40] vegetation monitoring,[41] classification of hyperspectral UAV imagery,[42] road following,[43] river monitoring,[44] recording of large-scale urban and suburban sites,[45] mapping of archaeological sites,[46] and investigating the accuracy of digital surface map production.[47] In recent research, Dunford et al.[48] used UAV technology to quantify riparian terrain and vegetation units and identify standing dead wood and canopy mortality, for the benefit of riparian managers. To classify rangeland vegetation from a 5-cm-resolution UAV image mosaic, Laliberte and Rango[22,24] developed and evaluated an image processing workflow that included the integration of resolution-appropriate field sampling, feature selection, and object-based image analysis. UAV was developed as a flexible and powerful tool for site-specific vineyard management.[10] Lin et al.[49] also demonstrated the feasibility and effectiveness of using a fixed-wing UAV system for rapid observation and risk assessment of ice jam formation over the Yellow River.

In summary, UAVs have been widely applied to surface feature information extraction, obtaining comparatively good results. In particular, Laliberte et al.[22–24] have made use of UAV to carry out research on the extraction and assessment of rangeland information. However, there are few studies that focus on extraction of cultivated land information, and still fewer focusing on UAV as a source of high-resolution images. Therefore, it is necessary to study the extraction of cultivated land information from UAV image data.

3 Methodology

UAV images are generally of poor quality and have a narrow frame; thus they cannot be directly used for mapping and require orthorectification and fusion processing before being used for subsequent cultivated land information extraction. Here, we first introduce the current UAV data acquisition and processing method, and then extract cultivated land information from the processed digital orthophoto map (DOM). In essence, our method is still object-oriented. Segmentation is first performed to partition an image into a series of nonoverlapping, meaningful, and homogeneous objects. Cultivated land has visually obvious features and rectangular traits, as well as concentrated clustered distribution, so the optimal segmentation scale for cultivated land information can be determined by visual judgment. After successful UAV image segmentation, we find that the main factors interfering with cultivated land information extraction are information from roads and concentrated residential areas. The cultivated land information can be successfully identified if these surface features are eliminated. The clustered distribution of cultivated land means that the segmented cultivated land plots are larger because of the very high resolution of the UAV image. Compared with cultivated land areas, the segmented concentrated residential areas appear fragmented as a result of abundant internal surface features, including houses, woodlands, roads, and concrete surfaces, which form segmentation objects of various shapes [Fig. 8(c)]. We, therefore, consider using the cluster approach to eliminate such surface features, including residential land and woodland, in order to achieve the ultimate goal of cultivated land information extraction.

Using our proposed method, we first eliminate the obviously strip-shaped objects with a rectangular fit feature from the total segmented objects, which includes roads and ditches. Next, we extract the central point of the segmentation object and obtain the fragmented

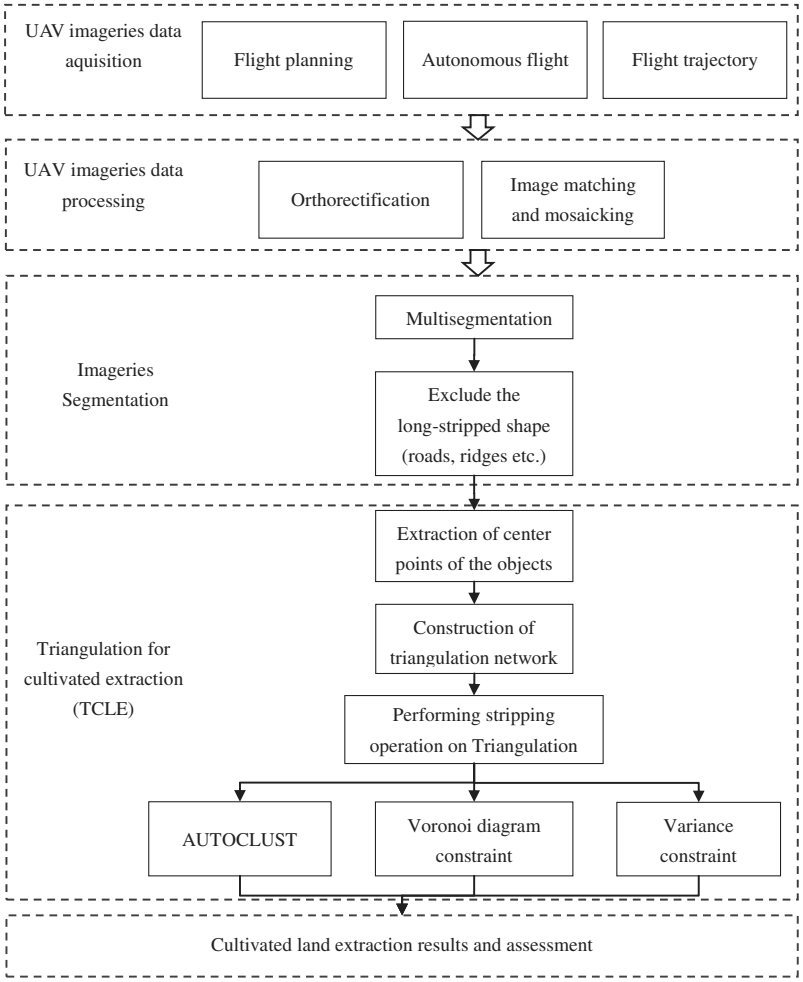


Fig. 1 Technical framework diagram of cultivated land information extraction using unmanned aerial vehicle (UAV) imagery.

distribution features of the segmentation objects within a residential area by constructing a triangulation network. The triangulation density in concentrated residential areas is high, while it is sparser in contiguous cultivated land areas. We then cluster residential areas by a triangulation network cluster approach, based on graph theory. Finally, we trim the cluster results with a Voronoi diagram and the maximum variance constraints, and complete the cultivated land information extraction. A detailed explanation of the above-mentioned procedures will be given in the following part of this article. The main technical framework diagram is shown in Fig. 1.

3.1 *Data Acquisition and Processing*

We used a fixed-wing UAV with Canon EOS 5D Mark II digital camera to acquire imagery at 80% forward lap and 60% sidelap. Imagery was acquired in August 2011 at a flying height of 750 m above the ground over the plain landscape of the Deyang Experimental Range in Sichuan, China. The camera's resolution was 5616×3744 pixels, resulting in an image footprint of 1123×748 m with a pixel resolution of 0.2 m. The camera focal length is 24.5988 mm, and the pixel size is 0.0064 mm. Figure 2 shows our fieldwork operations, collecting images by UAV and ground control points (GCPs) by real-time kinematic (RTK, which is a relatively new and commonly used method of GPS measurement).

GCPs play an important role on orienting data. From experience, GCPs for a UAV platform are suitable for setting one point per five images in single airstrip, setting one point on three adjacent air strips, and ensuring points are well distributed across a whole area. In addition,

(a) (b)

Fig. 2 Data acquisition. (a) UAV platform for acquiring imagery. (b) Collection of ground control points by real-time kinematic.

GCPs should be located in areas in which the surface features are distinct, easily distinguished, or where the color contrast is large with nearby surface features.

Image processing consists of (1) orthorectification and mosaicking of UAV imagery and (2) clipping of the mosaic. With traditional aerial photograph platforms, wind speed and wind direction exert a tremendous influence on the UAV flight attitude. There were multiple challenges associated with orthorectification of this imagery. These included the relatively small image footprints, considerable image distortion due to the use of a low-cost digital camera, difficulty in locating GCPs and generating tie points, and relatively large errors in exterior orientation parameters.[22] Additionally, the traditional digital photogrammetry systems (e.g., leica photogrammetry suite) have limitations when processing the UAV images. Therefore, we selected the digital photogrammetry grid (DPGRID) developed by Wuhan University to process the data, which is a professional software system that can produce a DOM from the UAV images. The Inpho software developed by Professor Fritz Ackermann in Germany could be another potential option for this step of the process.

At this stage, we first selected all of the images from air strips of the position orientation system data and removed images from the air strips that were taken when the UAV was climbing. Then we used the DPGRID to perform the data preprocessing, including distortion correction, strip arrangement, and image rotation. When the preprocessing was achieved, the software automatically completed the matching and extracts feature points. If the steps were successful, the external orientation of UAV scenes could be performed using GCP coordinates and tie points in order to improve the quality of the triangulation process.[50] Finally, we adjusted the color palate to be uniform, implemented an image mosaic, and created a DOM, ready for extraction of cultivated land.

3.2 *Image Segmentation*

Remote sensing image segmentation is the foundation and key to object-oriented information extraction.[51] The accuracy of segmentation results is directly related to the accuracy of subsequent information extraction results. On images, different objects typically occur on different scales; therefore, the chosen scales of interest will be guided by the purpose of the analysis. Here, we used multiscale segmentation in eCognition 8, an OBIA program, for image classification. In the multiscale segmentation process, the main parameters involved are scale and the composition of homogeneity. In this, the scale defines the maximum homogeneity criteria, with regard to the weighted image layers for resulting image objects; the higher the value, the larger the resulting image objects. The composition of homogeneity criterion is characterized by color and shape; each pair of criteria has a weighted percentage equalized to a value of one. Shape defines the textural homogeneity of the resulting image objects, including smoothness and compactness (shape = smoothness + compactness). In this, smoothness optimizes the resulting image objects, creating smooth borders within the shape criterion. Compactness optimizes the resulting image objects with regard to the overall compactness within the shape criterion.

In application of this step of the process, we usually follow two principles: (1) to segment the image at the largest possible segmentation scale, on the condition that necessary precision is met, considering the maximum segmentation scale at which cultivated land plots may be segmented and (2) to adopt a color standard on the premise that the necessary shape standard is met. The reason for the partitioning of these two principles is that the most important part of the image data is the spectral information, and high weighting of the shape criterion will reduce the quality of segmentation results.

The cultivated land is basically in a contiguous concentrated distribution, but is inevitably mixed with surface features, such as roads, ditches, and ridges, which can be segmented in very-high-resolution UAV images. Consequently, the center points of strip-shaped surface features will affect the accuracy of our method when extracting the center point of the segmentation object. On the other hand, these surface features are in a long-strip shape, contrasting greatly with rectangular isometric cultivated land. In order to improve the reliability of our method, these strip-shaped surface features need to be removed prior to construction of the triangulation network. Here, rectangular fit features are adopted to identify and remove these strip-shaped surface features.

The basic principle of rectangular fit features is to calculate the area of a rectangle of size equal to that of the object, and then to calculate the ratio of the area of the object outside the rectangle region and the area of the rectangle within the object region. In this, the first step is to create a rectangle of area equal to that of the image object. The length and width ratio of the rectangle is equal to that of the image object. Second, the area ratio of the object outside of the rectangle to the object within the rectangle is calculated. A ratio of 0 represents a complete mismatch, and 1 represents an exact match.

3.3 *Constructing the Triangulation Network*

A model of spatial proximity in a discrete point dataset should be properly modeled to elucidate sudden relative changes as a result of clustering. Voronoi diagrams and their accompanying dual Delaunay triangulations are used for precise and unique modeling of discrete point data with a succinct structure. Thus, they are solid candidates for analyzing spatial clustering. Delaunay triangulations are considered so powerful in succinctly representing proximity relationships that several clustering methods have been based on them.[52] Here, we propose two steps to construct a triangulation from segmentation objects, comprising calculation of the center points of objects and generation of Delaunay triangulations.

To extract the central point of a polygon, its area must first be calculated. At first, determining the area of a polygon seems complex, but the final formula is elegantly simple. Consider a polygon made up of line segments between N vertices (x_i, y_i), $0 \le i \le N-1$. The last vertex (x_N, y_N) is assumed to be the same as the first, and the polygon is closed (Fig. 3).[53]

The area is given by

$$A = \frac{1}{2}\sum_{i=0}^{N-1}(x_i y_{i+1} - x_{i+1} y_i). \tag{1}$$

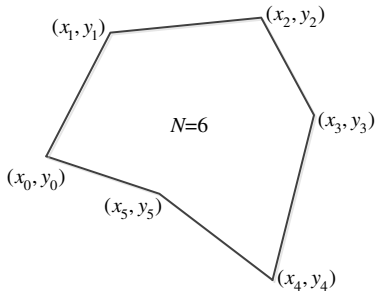


Fig. 3 Example of a typical polygon for calculating the centroid.

The center of the polygon, or the centroid, is also known as the center of gravity or the center of mass. As in the calculation of the area above, x_N is assumed to be equal to x_0; thus the polygon is closed. The centroid is calculated by the following equation:[50]

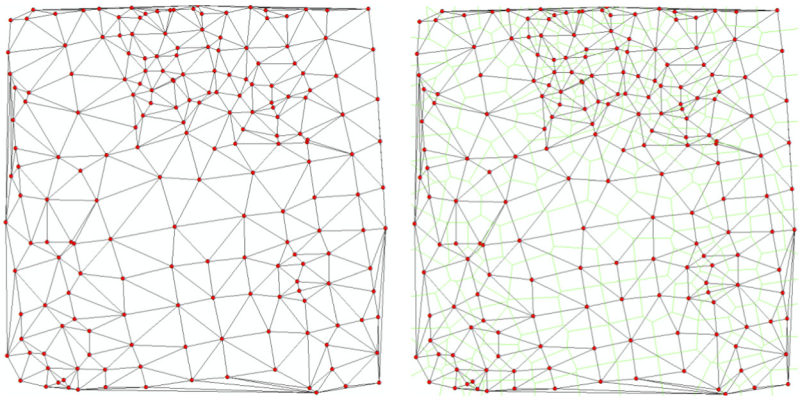
$$C_x = \frac{1}{6A}\sum_{i=0}^{N-1}(x_i + x_{i+1})(x_i y_{i+1} - x_{i+1} y_i), \tag{2}$$

$$C_y = \frac{1}{6A}\sum_{i=0}^{N-1}(y_i + y_{i+1})(x_i y_{i+1} - x_{i+1} y_i). \tag{3}$$

The Delaunay triangulation of a set of points on a plane is defined to be a triangulation such that the circumcircle of every triangle in the triangulation contains no point from the set in its interior. Such a triangulation exists for a given set of points, and it is the dual of the Voronoi tessellation.[54] In many engineering applications, and especially in geographic information system (GIS), the triangulation of a set of points in two dimensions is an ongoing problem. Many different triangulations can be constructed on a given set of points, and Delaunay triangulation is the most popular among them.[55] Each Delaunay edge is an explicit representation of a neighborhood relation between points. In addition, many algorithms also exist with an optimal expected time complexity. The most popular approaches are classified into six main categories: divide-and-conquer algorithms, sweep-line algorithms, incremental algorithms, fast incremental construction algorithms, gift-wrapping algorithms, and convex hull-based algorithms. The construction of Delaunay triangulation algorithms is not the research focus of this paper, so here we adopt the mature divide-and-conquer algorithms to construct the Delaunay triangulation. It was reported that Dwyer's divide-and-conquer algorithm[56] was strongest overall and was the most resistant to poor data distribution with a $O(n \log n)$ worst case. Figure 4 shows a Delaunay triangulation and Voronoi diagram constructed by extracting the central points of polygons.

3.4 *Method of Detection of Spatial Features in the Triangulation Network Model*

Automatic detection of spatial features in the triangulation network model is carried out with a computer in order to extract a region of interest. Here, we first perform a stripping operation on the constructed triangulation network and remove the long edge. Next, the triangulation network is clustered with the AUTOCLUST clustering method, and point groups are deleted in the dense regions. Then, area constraints are performed with a Voronoi diagram, and the average triangle area of the clustering triangulation network is calculated, allowing deletion of those triangulation



(a) (b)

Fig. 4 Voronoi diagrams and their dual Delaunay triangulations. (a) Delaunay triangulations. (b) Voronoi diagrams.

network clusterings with an average area that is too large. Last, forest lands with variance features are excluded.

Many studies have been carried out on the range of distribution of point groups. The spatial distribution range of a point group is still an uncertainty. In GIS, the convex hull is often used in place of the point group distribution range. The drawback of this method is that when the point group distribution is of convex and dent type, its convex hull may contain several nonpoint dent regions and cannot reflect the true distribution range of the point group. The extraction process based on the Delaunay triangulation is as follows: (1) construct a Delaunay triangulation in the point group [Fig. 4(a)]; (2) set a visual proximity distance d and perform a stripping operation on the triangle based on d, to traverse three sides of a triangle. If a side length is larger than d, then its connection is cut; (3) repeat process (2) until all the sides of all the triangles are less than d; (4) extract the side equal to 1, which is associated with the number of triangles, as the boundary of the point group distribution range. In Fig. 5, we find that by carrying out the stripping operation, triangles with larger side lengths are deleted, and the results after the stripping is closer to the distribution range of the point group and in accordance with the basic principles of human spatial cognition.

The threshold distance d of the stripping operation can be preset on the basis of experience. The larger the d value, the fewer the number of sequential strippings and the closer the obtained range of polygons is to a convex hull. The smaller the d value, the greater the number of sequential strippings and the deeper the obtained range of polygons is curved into the dent. Thus, a reasonable set of d values can obtain spatial clustering results of a group of points.

The triangulation network acquired by Delaunay triangulation segmentation can automatically adapt to changes in the density of points in different regions. In regions with dense points, the obtained triangle will be relatively small, and the side length is relatively short. The obtained triangle in regions with relatively sparse points is large, and correspondingly, the side length is relatively long. In regions with the same point density, both the size of the triangle and the side length are nearly identical. The side lengths of the connected triangles in different density regions vary greatly. AUTOCLUST belongs to the family of graph-based algorithms[52] that first construct a graph. It goes through short-long clustering to deliver P with cluster identifiers; the method utilizes both global and local variations.

The standard deviation and its square, the variance, are the most valuable and popular measures for describing dispersion. Thus, we may say that AUTOCLUST detects points whose incident edges exhibit lengths that have unusually large dispersion. In order to define this concept precisely, we introduce the following notation. Let $\mathrm{DD}(P)$ denote the Delaunay diagram of P. The graph $\mathrm{DD}(P)$ is a planar map that contains vertices and edges. Edges of $\mathrm{DD}(P)$ will be called Delaunay edges. For a point $p_i \in P$ (a vertex of $\mathrm{DD}(P)$), the neighborhood $N(p_i)$ is the set of Delaunay edges incident to point p_i. In the AUTOCLUST algorithm, we mainly use statistical variables $\mathrm{Mean\_St\_Dev}(P)$, $\mathrm{Local\_Mean}(p_i)$, and $\mathrm{Local\_St\_Dev}(p_i)$ to establish long side and short side sets. The statistical equations are defined as follows:[52,57]

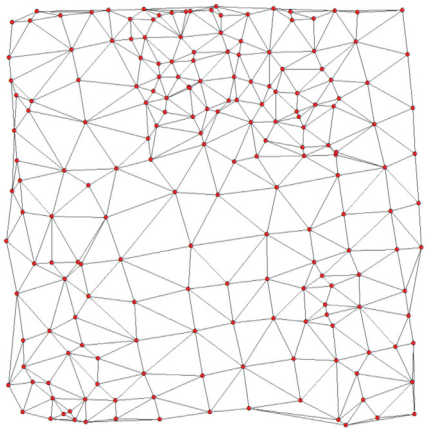


Fig. 5 Stripping outside the triangle in a Delaunay triangulation.

$$\text{Local\_Mean}(p_i) = \sum_{e \in N(p_i)} |e| / \|N(p_i)\| = \sum_{j=1}^{d(p_i)} |e_i| / d(p_i), \tag{4}$$

$$\text{Local\_St\_Dev}(p_i) = \sqrt{\sum_{j=1}^{d(p_i)} [\text{Local\_Mean}(p_i) - |e_j|]^2 / d(p_i)}, \tag{5}$$

$$\text{Mean\_St\_Dev}(P) = \sum_{i=1}^{n} \text{Local\_St\_Dev}(p_i)/n, \tag{6}$$

$$\text{Relative\_St\_Dev}(p_i) = \text{Local\_St\_Dev}(p_i)/\text{Mean\_St\_Dev}(P), \tag{7}$$

where $\text{Local\_Mean}(p_i)$ is the average value of sides connected to point p_i. $d(p_i)$ is the number of sides connected to point p_i. $\text{Local\_St\_Dev}(p_i)$ is the standard deviation of sides connected to point p_i. The average of the $\text{Local\_St\_Dev}(p_i)$ values as we let p_i be each point in P is denoted by $\text{Mean\_St\_Dev}(P)$. $\text{Relative\_St\_Dev}(p_i)$ denotes the ratio of $\text{Local\_St\_Dev}(p_i)$ and $\text{Mean\_St\_Dev}(P)$.

In the AUTOCLUST algorithm, using local and global statistics, short-long clustering classifies each edge adjacent to each p_i into one of three groups: short edge [denoted by $\text{Short\_Edges}(p_i)$], long edge [denoted by $\text{Long\_Edges}(p_i)$], and other edge [denoted by $\text{Other\_Edges}(p_i)$]. e_j belongs to $\text{Short\_Edges}(p_i)$ if and only if the length of e_j is less than $\text{Local\_Mean}(p_i) - \text{Mean\_St\_Dev}(P)$. Thus, a short edge could be expressed by the following equation:

$$\text{Short\_Edges}(p_i) = \{e_j | |e_j| < \text{Local\_Mean}(p_i) - \text{Mean\_St\_Dev}(P)\}, \tag{8}$$

where an edge $e_j \in N(p_i)$.

An edge $e_j \in N(p_i)$ belongs to $\text{Long\_Edges}(p_i)$ if and only if the length of e_j is less than $\text{Local\_Mean}(p_i) + \text{Mean\_St\_Dev}(P)$. It is expressed by the following equation:

$$\text{Long\_Edges}(p_i) = \{e_j | |e_j| > \text{Local\_Mean}(p_i) + \text{Mean\_St\_Dev}(P)\}. \tag{9}$$

$\text{Other\_Edges}(p_i)$ consists of the subset of $N(p_i)$ whose edges are greater than or equal to $\text{Local\_Mean}(p_i) - \text{Mean\_St\_Dev}(P)$ and less than or equal to $\text{Local\_Mean}(p_i) + \text{Mean\_St\_Dev}(P)$. It is expressed by the following equation:

$$\text{Other\_Edges}(p_i) = N(p_i) - [\text{Short\_Edges}(p_i) \cup \text{Long\_Edges}(p_i)]. \tag{10}$$

Delaunay edges in $\text{Long\_Edges}(p_i)$ are exceptionally long edges in the k-neighborhood of p_i, while Delaunay edges in $\text{Short\_Edges}(p_i)$ are exceptionally short edges. Short-long clustering removes exceptionally long Delaunay edges in the 1-neighborhood to extract approximate cluster boundaries. It then eliminates some short Delaunay edges in the 1-neighborhood to overcome the chain effect.[58] Short-long clustering extends the local neighborhood to the 2-neighborhood to remove inconsistently long edges, as a third step in the process. Further details are included in the original paper.[52] Three main phases describe the AUTOCLUST algorithm; the steps of the algorithm are shown in Fig. 6.

Step 1: Construct Delaunay triangulation network, compute statistical magnitudes of $\text{Local\_Mean}(p_i)$, $\text{Local\_St\_Dev}(p_i)$, and $\text{Mean\_St\_Dev}(P)$, and delete all the sides in both the short side set and the long side set, hence forming the rough boundary of each cluster.

Step 2: After short-long clustering, compute connected components and label each point p_i with its connected component $\text{CC}[p_i]$; then recuperate all intracluster edges by reconnecting every edge $e_{ij} = (p_i, p_j)$ satisfying $\text{CC}[p_i] = \text{CC}[p_j]$ (where $\text{CC}[p]$ is the connected component of p, which means that if edges $e_{ij} = (p_i, p_j)$ and $e_{ik} = (p_i, p_k)$ are in $\text{Other\_Edges}(p_i)$, then $\text{CC}[p_i] = \text{CC}[p_j]$). Figure 6(b) illustrates the result after short-long clustering, while Fig. 6(c)

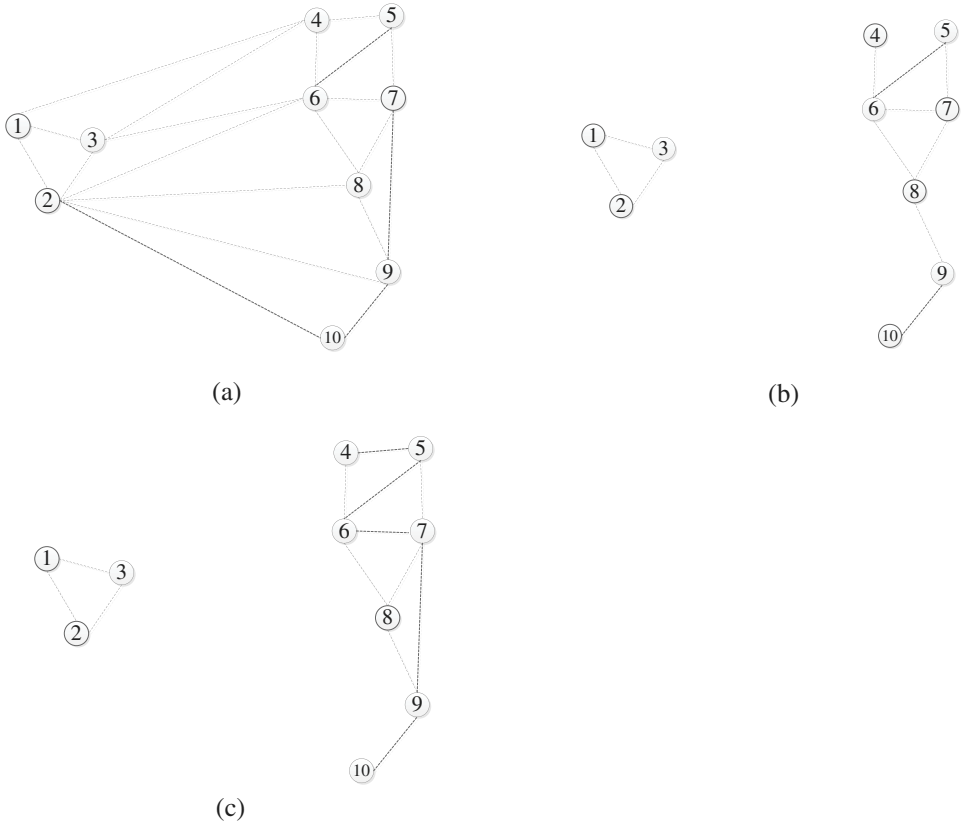


Fig. 6 (a) A set P of points with DD(P). (b) Short-long clustering results (two clusters). (c) Recuperating intracluster edges in DD(P).

depicts the result of recuperation. The recuperated subgraph approximates the shapes of the clusters.

Step 3: Expand the scope of neighboring points and expand to all the points of the path length to p_i not larger than 2. Then use Local_Mean_2, $\text{DD}(P)$ to carry out further adjustments, and look for connection sides between different regions. After the first two phases, the current subgraph of $\text{DD}(P)$ is the result of local analysis with regard to global effects. However, to complete the analysis, the immediate neighborhood is extended slightly. Thus, we extend the notion of the neighborhood of a vertex p_i in a graph G as follows. The k-neighborhood $N_{k,G}(p_i)$ of a vertex p_i in graph G is the set of edges of G that belong to a path of k or fewer edges starting at p_i. We denote the mean length of edges in $N_{k,G}(p_i)$ by $\text{Local\_Mean}_{k,G}(p_i)$. That is, $\text{Local\_Mean}_{k,G}(p_i) = \sum_{e \in N_{k,G}(p_i)} |e| / \|N_{k,G}(p_i)\|$. For each point p_i, this phase simply removes all edges in $N_{2,G}(p_i)$ that are long edges. That is, all $e \in N_{2,G}(p_i)$ such that $|e| > \text{Local\_Mean}_{2,G}(p_i) + \text{Mean\_St\_Dev}(P)$.

The AUTOCLUST algorithm only uses the variation characteristics of the sides of the triangulation network, but ignores the areas of the triangles in the triangulation network. The adjacent triangulation network connected by edges belonging to $\text{Other\_Edges}(p_i)$ does not easily cluster, but the triangulations within the range are small triangles with a highly concentrated point group. To analyze this feature, we further introduce the Voronoi polygon to determine the point density features, because the density features of point group can be understood as the number of points contained within a unit area. From the point of view of spatial competition, the greater the point density, the more the points contained within the unit area. The greater the number of points, the smaller the spatial scope influenced by each point, namely, the smaller the associated Voronoi polygon area. The Voronoi polygon is the dual of the Delaunay triangulation, so we calculate point distribution density on the basis of the Voronoi polygon area. This comprises the following two steps: (1) construct Delaunay triangulation on the point groups and (2) construct its dual

Voronoi diagram on the basis of Delaunay triangulation and calculate area A of each Voronoi polygon. At this point, the density of each point can be described as 1/A. The smaller 1/A, the smaller the density, and vice versa. Generally, when we perform an AUTOCLUST cluster, we find that there are phenomena of insufficient clustering in many regions; therefore, we constrain with a Voronoi polygon, by performing comprehensive constraint with upper and lower threshold values. In other words, we consider that when the Voronoi polygon is greater than the upper limit of the threshold value, the center point corresponding to the polygon is definitely cultivated land, and when the Voronoi polygon is lower than the limit of the threshold value, the center point corresponding to the polygon must belong to the fragmented area. Consequently, we try to set loose upper and lower limits of threshold values. We use the following formulae to filter point areas of high and low point density.

$$\text{Area} < \partial_1 \times \text{Voronoi\_mean}, \tag{11}$$

$$\text{Area} > \partial_2 \times \text{Voronoi\_mean}, \tag{12}$$

where Voronoi_mean is the average value of the generated Voronoi polygon areas, ∂_1 and ∂_2 are the coefficients, and ∂_1 is smaller than ∂_2. Obviously, the coefficient here determines the degree of clustering of points.

In the AUTOCLUST (graph-based) clustering process, we find some large triangulation networks, and those triangulation networks whose area is smaller than their surrounding triangulation networks are identified as clustering, too. The numbers of this kind of triangulation network are usually small, and the mean triangle area is much larger than that of triangulation networks in the fragmented region, which require manual identification. Therefore, we introduce the concept of clustering triangulation mean area to remove incorrect clustering from the AUTOCLUST (graph-based clustering) process.

$$\text{Local\_Mean}_{\Delta\text{area}}(C_j) = \sum_{i=1}^{n} T_i/n, \tag{13}$$

$$\text{Mean}_{\Delta_{\text{area}}}(C) = \sum_{j=1}^{m} \text{Local\_Mean}_{\Delta_{\text{area}}}(C_j)/m, \tag{14}$$

$$\text{Mean\_Dev}_{\Delta_{\text{area}}}(C_j) = \sqrt{\sum_{j=1}^{m} [\text{Local\_Mean}_{\Delta_{\text{area}}}(C_j) - \text{Mean}_{\Delta_{\text{area}}}(C)]^2/m}, \tag{15}$$

where $\text{Local\_Mean}_{\Delta_{\text{area}}}(C_j)$ is triangle mean of j clustering, $\text{Mean}_{\Delta_{\text{area}}}(C)$ is the average of m clustering mean, $\text{Mean\_Dev}_{\Delta_{\text{area}}}(C_j)$ is the variance of all the clustering means, j is the number of the cluster, and i is the serial number of the triangle in j cluster. If $\text{Local\_Mean}_{\Delta_{\text{area}}}(C_j) > \text{Mean}_{\Delta_{\text{area}}}(C) + \text{Mean\_Dev}_{\Delta_{\text{area}}}(C_j)$, then the cluster is excluded.

By observing a large number of experimental images, we find that our proposed method is able to cluster the concentrated residential areas and exclude them. However, several sporadic woodlands are present in the plains. Consequently, the extraction and segmentation objects are usually large due to the small difference in the texture of woodland to cultivated land, and it is difficult to exclude this in the method we propose. Therefore, we recommend a variance constraint method for trimming the cultivated land information extraction results. The rough texture of woodland has a large variance, while the relatively smoother cultivated land has a smaller variance. Here we use two times of sum of the maximum mean deviation and the standard deviation of the maximum deviation as the threshold for identification. Computation is as follows:

$$\text{Max.diff} > 2 \times \text{mean\_Max.diff} + \text{Standard Deviation\_Max.diff}, \tag{16}$$

where mean_Max.diff is the average value of the Max.diff feature for all of the objects. Standard Deviation_Max.diff is the standard deviation of the Max.diff feature.

Table 1 Example of confusion matrix.[59]

		Actual class				
		1	2	3	4	$\sum$
Predicted class	1	n_{11}	n_{12}	n_{13}	n_{14}	n_{1+}
	2	n_{21}	n_{22}	n_{23}	n_{24}	n_{2+}
	3	n_{31}	n_{32}	n_{33}	n_{32}	n_{3+}
	4	n_{41}	n_{42}	n_{43}	n_{44}	n_{4+}
	$\sum$	n_{+1}	n_{+2}	n_{+3}	n_{+4}	n

3.5 *Accuracy Assessment*

In order to evaluate the performance of both methods, classification accuracies (overall, user's, and producer's) and the kappa index[59] were calculated for three experiments. The measures of classification accuracy are derived from the confusion matrix (Table 1). In the confusion matrix, the main diagonal element of the matrix is showed as n_{ii}, which is the number of class i extracted correctly. The off-diagonal elements contain those cases where there is a disagreement in the labels, such as n_{ij} is the number of which class i is incorrectly extracted as class j ($i \neq j$). In the example shown in Table 1, q is the number of classes ($q = 4$). According to the confusion matrix, overall accuracy (OA) can be calculated by Eq. (16); user's accuracy (UA) can be calculated by Eq. (18); producer's accuracy (PA) can be calculated by Eq. (19); kappa index can be calculated by Eq. (20).[59]

$$\mathrm{OA} = \frac{\sum_{k=1}^{q} n_{kk}}{n} \times 100\%, \tag{17}$$

where n depicts all of the elements for the classification, q is the number of the classes, and n_{kk} is the number of class k extracted correctly.

$$\mathrm{UA}(k) = \frac{n_{kk}}{n_{k+}} \times 100\%, \tag{18}$$

where $\mathrm{UA}(k)$ is the user's accuracy of class k; $n_{k+} = \sum_{i=1}^{q} n_{ki}$, $k = 1, 2, 3, 4$.

$$\mathrm{PA}(k) = \frac{n_{kk}}{n_{+k}} \times 100\%, \tag{19}$$

where $\mathrm{PA}(k)$ is the producer's accuracy of class k; $n_{+k} = \sum_{i=1}^{q} n_{ik}$, $k = 1, 2, 3, 4$.

$$\mathrm{Kappa} = \frac{n \sum_{k=1}^{q} n_{kk} - \sum_{k=1}^{q} n_{k+} n_{+k}}{n^2 - \sum_{k=1}^{q} n_{k+} n_{+k}}, \tag{20}$$

where n depicts all of the elements for the classification.

4 Experimental Studies

The experiments were carried out over a cultivated land area interspersed with infrastructure, such as ditches and roads. The analysis results can demonstrate the applicability of our method to the extraction of cultivated land information from UAV imagery. The UAV images were acquired in August 2011 at a flying height of 750 m above the ground over a plain landscape at the Deyang Experimental Range in Sichuan, China (with 0.2-m resolution). First, we segment the images with multiscale segmentation method and a suitable segmentation scale. We then carry out the triangulation cluster for cultivated land information extraction (TCLE) method

(a) (b) (c)

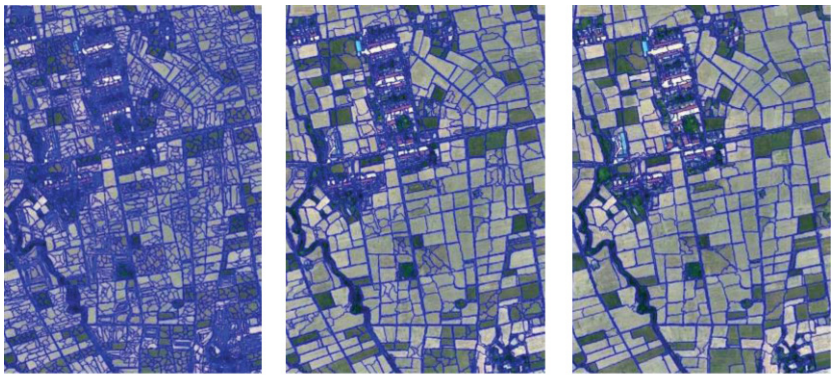
Fig. 7 Three UAV images for experiments 1, 2, and 3, respectively. (a) An image used in experiment 1, with a size of 3350×4400 pixels (only a few houses are contained in the noncultivated land). (b) An image used in experiment 2, with a size of 2400×1500 pixels (noncultivated land is more broken, covering houses, cement floor, and woodland). (c) An image used in experiment 3, with a size of 5000×5000 pixels (noncultivated land area distribution is more complex, comprising more land use types).

and eCognition for cultivated land information extraction (ECLE) on the segmentation results to extract cultivated land information. Last, we compare the extracted results acquired by these two methods. Further image classification experiments over three different areas are employed to verify the validity of the proposed TCLE approach (Fig. 7).

4.1 *Experiment 1*

4.1.1 *Multisegmentation*

This experiment covers a complicated zone with a spatial resolution of 0.2 m and size of 3350×4400 pixels, as shown in Fig. 7(a). In order to attest the reliability of the segmentation scale, we chose cultivated land information extraction; we utilize eCognition 8 to carry out multiscale segmentation (Fig. 8) by color/shape and compactness/smoothness parameters. Color/shape and compactness/smoothness were set to 0.6/0.4 and 0.5/0.5, respectively (Table 2). The image was segmented at three segmentation scales: 50, 100, and 180, respectively. The segmentation statistics are shown in Table 2. We easily find that with an increase of segmentation scale, the area of segmented objects also increases. Compared to segmentation scale parameters of 50 and 80, when the segmentation scale parameter is 180, the cultivated land plots with the same pixel feature in the 0.2-m-resolution UAV imagery data are found to be better clustered



(a) (b) (c)

Fig. 8 Image multisegmentations of UAV image mosaics (shape and compactness parameters: shape 0.6, compactness 0.5), showing (a) segmentation at scale parameter 50, (b) segmentation at scale parameter 100, and (c) segmentation at scale parameter 180.

Table 2 Parameter values used in multiresolution segmentation algorithm.

Image segmentation parameters				
Scale	Color/shape	Smoothness/compactness	Numbers of objects	Median area of objects (sq.m)
50	0.6/0.4	0.5/0.5	6749	78.5
100	0.6/0.4	0.5/0.5	1740	304.5
180	0.6/0.4	0.5/0.5	761	695.3

[Fig. 8(c)]. So the segmented objects can represent a true geographic object at a proper scale (e.g., cultivated land), which means a better one-to-one relationship between segmented objects and visually interpreted objects.

4.1.2 *Triangulation cluster for cultivated land information extraction*

Figure 9(a) is the constructed triangulation network, and the yellow points are center points of segmentation objects. We find that areas A, B, C, and D are obviously residential areas, that the triangulation network density here is greater than that of the surrounding cultivated land area, and that the side length of the triangulation network is much smaller than that of the triangulation network corresponding to the surrounding cultivated land area. The stripping operation is then performed on the triangulation network using AUTOCLUST clustering to automatically obtain high-density triangulation network area, in which the red dots are clustering points. The extraction effect is shown in Fig. 9(b). The red dots are the center points of the noncultivated land plots, which are identified by clustering, in this; regions B, C, D, and E can be better identified as a whole. For region A, there are some differences from our expected clustering results, mainly because of the uniform structure of the building in the central part of region A and the larger adjacent blocks of woodland, resulting in scattered segmentation objects in this region, with a corresponding triangulation network that is less dense, and an insufficient clustering phenomenon. However, this does not prevent us from giving priority to AUTOCLUST as clustering method to identify the fragmented rural residential areas because the total number of AUTOCLUST clustering points is 170 [Fig. 9(b)], compared to the correct number of clustering network points, which is 150. By visual discrimination, we know that 44 center points of the cultivated land are mistakenly identified as clustering points, resulting in excessive clustering with an error rate of 29.3%. On the other hand, in Fig. 9(b), 27 triangulation network points that are not clustered are identified as clutivated land, which mainly lie in rectangular and triangular areas, leading to insufficient clustering phenomena with a clustering error rate of 18.0%. Due to the excessive clustering and insufficient clustering phenomena, we introduce the Voronoi polygon area constraint method. For those insufficient clustering areas, we set the lower limit for the threshold value [Fig. 9(c)]. In addition, influenced by blocks of woodland, there is no obvious density change in the triangulation network between the residential area and the cultivated land border, which leads to incorrect clustering of that part of the cultivated land near to residential areas and to an excessive clustering phenomenon. We eliminate the error by setting the Voronoi polygon upper limit to the threshold value [Fig. 9(d)].

After segmentation, we carry out clustering analysis by the triangulation construction method. The experiment demonstrates that residential areas and road information in contiguous cultivated land can be better excluded by these methods. However, the ditches and isolated woodlands in the imagery are not identified or excluded [Fig. 11(a), elliptical area]. Here, we adopt the gray-value variance constraints approach to exclude ditches and woodlands. The mean and variance of each segmentation object are calculated by using a gray value. Due to the smooth surface of cultivated land, its variance is small. The result is shown in Fig. 11(a).

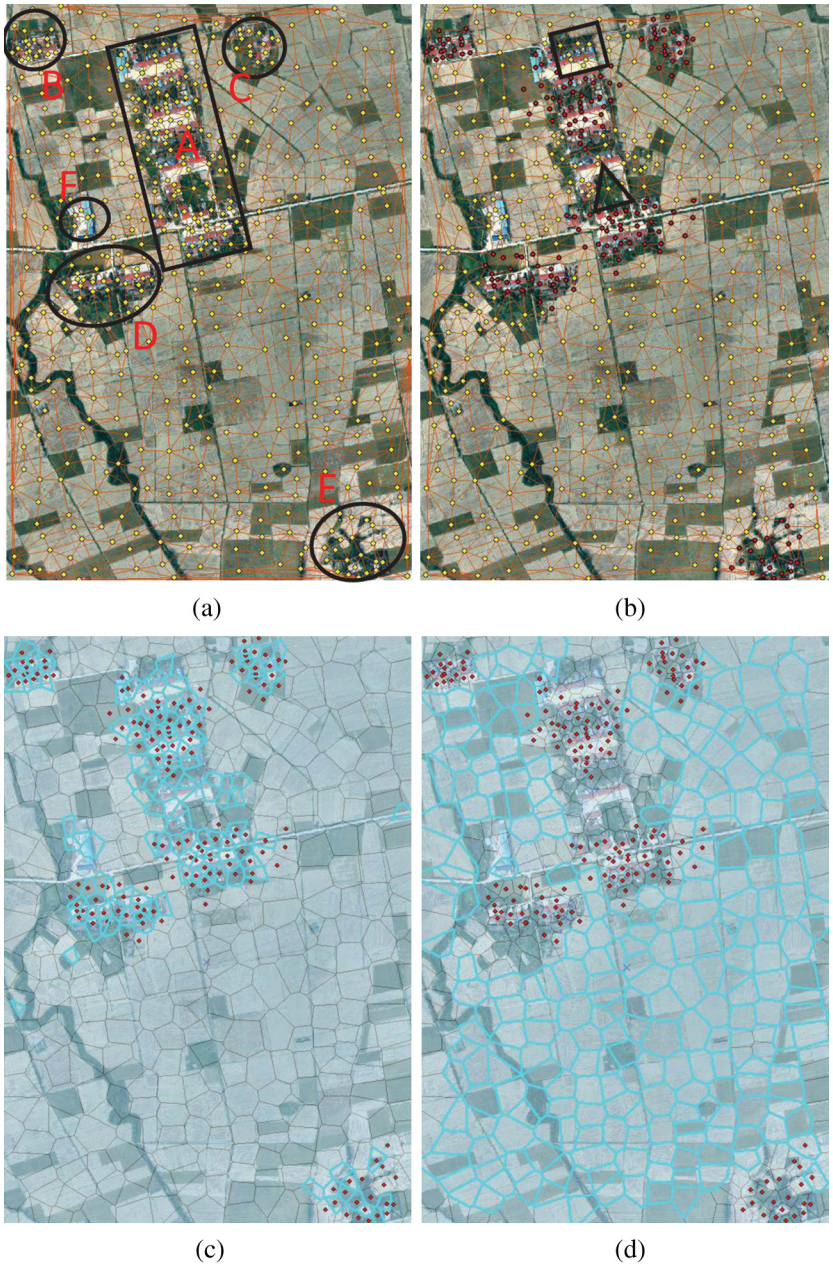


Fig. 9 Construction and feature detection of triangulation network. (a) The constructed triangulation network with the extracted center points (yellow dots are the center points of each segmentation surface). (b) After the stripping operation, we use AUTOCLUST clustering to construct a triangulation network, in which the red dots are clustered points. (c) Marker results of a Voronoi diagram with polygon area <472 (0.5 × Voronoi_mean). (d) Marker results of Voronoi diagram with polygon area >850 (0.9 × Voronoi_mean).

4.1.3 *eCognition for cultivated land information extraction*

For images analysis, we adopted the eCognition 8 object-oriented classification method to extract cultivated land information[60] and compared it with the method proposed. In the actual classification, the object-oriented method only requires two to three times the number of training samples in the image band and can achieve high and stable classification accuracy.[61] To separate cultivated land and noncultivated land, here we chose 14 cultivated land samples (e.g., dry land, paddy field, etc.) and 15 noncultivated land samples (e.g., roads, residential area, woodland,

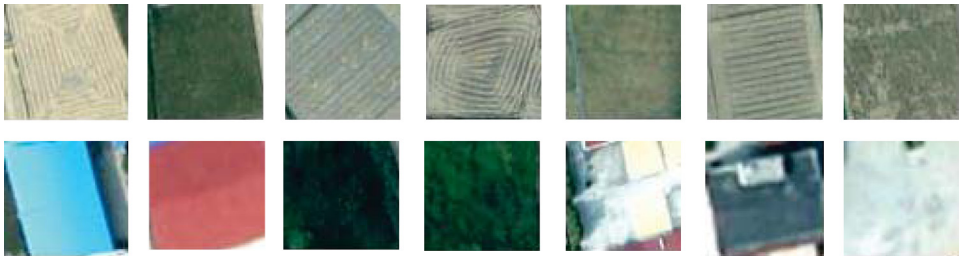


Fig. 10 The selected sample types: (upper) some cultivated land samples and (lower) some non-cultivated land samples.

concrete ground, etc.) (Fig. 10). Using the samples, we extracted segment-based information for 14 features as a base for the feature selection process based on the research by Laliberte and Rango,[24] including spectral, spatial, and contextual features. It is not always easy to predict which features or feature combinations will work best.[24] These samples served as inputs for feature space optimization function of nearest-neighbor classification in eCognition 8 to determine the optimal combination of optimal spectral, spatial, and contextual features. After analysis, the maximum classification distance was achieved with eight-dimensional feature space (best separation distance 6.42). For the classification, a nearest-neighbor classification algorithm was used with the following features by feature space optimization: area, max.diff, HIS transformation hue, standard deviation of the green, the means of the red, standard deviation of the red, the mean difference to neighbors of the red, and standard deviation of the blue.

4.1.4 *Comparing TCLE with ECLE*

Here, an assessment of accuracy was conducted for the study area visual interpretation method by determining overall, producers', and users' accuracies, and kappa index.[59,62] Due to the ultra-high resolution of UAV images, we use the visual interpretation method to verify accuracy. We overlay the extracted objects and reference objects and then calculate the area ratio of each extracted object to its corresponding reference object. If the ratio of an extracted object to a reference object is >60%, it is taken as a correctly extracted object; otherwise, it is taken as a wrongly extracted object. The number of correctly extracted cultivated land objects from the visual interpretation method is 362. The number of extracted cultivated land objects by the TCLE method is 343, and the number of correctly extracted cultivated land objects from the TCLE method is 325. Overall accuracy is 92.8%, with kappa index of 0.855 (Table 3). The number of extracted cultivated land objects with TCLE method is 319, in which the number of correctly extracted cultivated land objects is 308. Overall accuracy is 91.5%, and kappa index is 0.828 (Table 3). Experiments show that the accuracy of the two methods is roughly equivalent.

Table 3 Accuracy analysis of experiment 1.

	Proposed method		eCognition 8	
	Cultivated land	Noncultivated land	Cultivated land	Noncultivated land
Cultivated land	325	18	308	11
Noncultivated land	37	381	54	389
Producer's accuracy (%)	89.8		85.1	
User's accuracy (%)	94.8		96.6	
Overall accuracy (%)	92.8		91.5	
Kappa index	0.855		0.828	

Fig. 11 The finally extracted cultivated land map. (a) Extraction effect map from the triangulation for cultivated land information extraction (TCLE) method. Ditches in elliptical area and independent woodland in the rectangular area are correctly removed by variance constraints. (b) Extraction effect map from the eCognition for cultivated land information extraction method. Green represents cultivated land and red noncultivated land.

However, there are sporadically misclassified cultivated land plots in the cultivated land extracted by the ECLE method [Fig. 11(b)]. Our TCLE method can guarantee the concentrated and contiguous properties of cultivated land as a whole and, to some extent, avoid sporadic phenomena of misclassification. Thus, our proposed method can maintain the integrity of extracted cultivated land by maintaining the required accuracy.

Extracting cultivated land with the ECLE method requires collecting a large number of samples. The selection of samples and feature spaces is a heavy workload, but according to researchers' experiences, sample selection errors will be different. Using the nearest-neighbor classification method, the distance calculation depends only on the selected training samples; therefore, changes in the training sample directly affect the calculation results. Using the TCLE method proposed in this paper greatly improves the automatic processing of cultivated land information extraction. For the regional features of fragmented rural residential areas within contiguous cultivated land, we propose a specific cultivated land information extraction method, eliminating the process of sample selection after segmentation and of training the corresponding rules. The proposed AUTOCLUST algorithm makes full use of cultivated land distribution features, achieving automatic clustering of fragmented residential areas and simplifying the human-computer interaction in the process of extracting cultivated land information.

4.2 *Experiment 2*

Experiment 2 was carried out over a more complicated zone with a spatial resolution of 0.2 m and size of 2400×1500 pixels, as shown in Fig. 7(b). Noncultivated land is more broken up, covering houses, cement floors, and woodland, and the cultivated land has different phenological crops [Fig. 7(b)]. Due to the single small object on the noncultivated land, a smaller segmentation scale of 120 is set, with the same color and compactness as Experiment 1 [Fig. 12(a)]. The threshold values are set as $\partial_1 = 0.5$ and $\partial_2 = 0.9$ (Table 6), to modify the clustering results (Table 6). The classification result from TCLE is shown in Fig. 12(b), and the classification result from ECLE is shown in Fig. 12(c). Compared with TCLE, ECLE

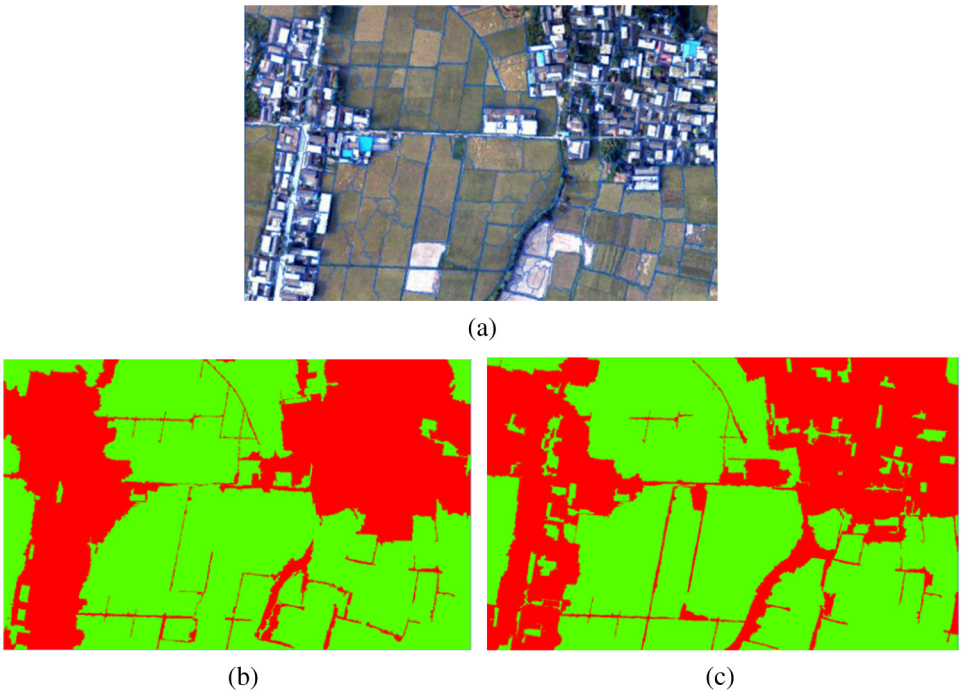


Fig. 12 Results for experiment 2. (a) Segmentations at scale parameter 180 (shape and compactness parameters: shape 0.6, compactness 0.5). (b) Extraction effect map from the TCLE method. (c) Extraction effect map from eCognition. Green represents cultivated land and red noncultivated land.

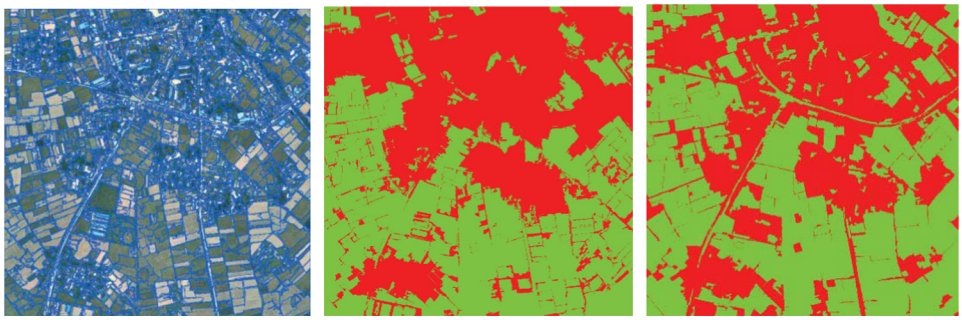
is unable to detect more noncultivated land. The overall accuracies are 90.3 and 87.6%, respectively (Table 4).

4.3 *Experiment 3*

The rural-urban fringe zones were surveyed with a spatial resolution of 0.2 m and size of 5000×5000 pixels, as shown in Fig. 7(c), containing a variety of land use types. This area was selected for Experiment 3 to verify the proposed method further. Compared with Experiments 1 and 2, the noncultivated land area distribution is more complex in Experiment 3. The segmentation parameters are set as scale 120, with the same color and compactness as Experiment 1 [Fig. 13(a)]. The threshold values are set as $\partial_1 = 0.5$ and $\partial_2 = 0.9$

Table 4 Accuracy analysis of experiment 2.

	Proposed method		eCognition 8	
	Cultivated land	Noncultivated land	Cultivated land	Noncultivated land
Cultivated land	144	17	143	20
Noncultivated land	42	403	55	388
Producer's accuracy (%)	77.4		72.2	
User's accuracy (%)	89.4		87.7	
Overall accuracy (%)	90.3		87.6	
Kappa index	0.762		0.705	



(a) (b) (c)

Fig. 13 Results for experiment 3. (a) Segmentations at scale parameter 180 (shape and compactness parameters: shape 0.6, compactness 0.5). (b) Extraction effect map from the TCLE method. (c) Extraction effect map from eCognition. Green represents cultivated land and red noncultivated land.

(Table 6), to modify the clustering results. Figure 13(b) shows the classification result from TCLE, and Fig 13(c) shows the classification result from ECLE. The overall accuracies are lower: 86.9 and 85.8%, respectively (Table 5).

This experiment also yields the following results. (1) The accuracy of the proposed method is basically the same as that found in Experiments 1 and 2. This further validates the effectiveness and reliability of the proposed method, due to the later experiment's complexity. (2) The overall detection accuracy of both methods in Experiments 2 and 3 is lower than that in Experiment 1, which is mainly a result of the more complicated surface in Experiments 2 and 3. The increase in surface complexity leads to an increase in the proportion of noncultivated land types that contain multiple objects in the scene. Due to the fragmented edge between noncultivated lands and cultivated lands, more small cultivated land objects are not extracted in Experiment 3. (3) Approximately 50 objects belonging to clustered small crop fields are wrongly removed. It leads to the number of correctly extracted land objects being relatively smaller. It also shows that the method without training samples faces more technical difficulties.

In this study, different thresholds are used to gain a general result in the three experiments; they are listed in Table 6. Due to the high homogeneity of objects, a better result is gained at scale 120 for Experiment 2 with $\partial_1 = 0.5$ and $\partial_2 = 0.9$. According to these experiments, we find that a scale of 100 to 200 would be most efficient for our method, but with some dependence on the segmentation results. All three experiments set parameters as $\partial_1 = 0.5$ and $\partial_2 = 0.9$ (Table 6). Thus, from these experiments, in order to avoid excessive clustering and insufficient clustering, we recommend $\partial_1 = 0.5$ and $\partial_2 = 0.9$ as general values so as to complement the AUTOCLUST algorithm.

Table 5 Accuracy analysis of experiment 3.

	Proposed method		eCognition 8	
	Cultivated land	Noncultivated land	Cultivated land	Noncultivated land
Cultivated land	572	133	670	125
Noncultivated land	117	1082	145	964
Producer's accuracy (%)	83.0		82.2	
User's accuracy (%)	81.1		84.3	
Overall accuracy (%)	86.9		85.8	
Kappa index	0.717		0.709	

Table 6 Parameters for the three experiments.

Parameters	Scale	Color	Compactness	∂_1	∂_2
Experiment 1	180	0.6	0.5	0.5	0.9
Experiment 2	120	0.6	0.5	0.5	0.9
Experiment 3	180	0.6	0.5	0.5	0.9

5 Conclusion

The results of this study indicate that UAV-based systems for land inventory and monitoring show great promise. There is increasing interest in the use of very-large-scale aerial imagery for cultivated land assessment and monitoring,[8–10] and UAVs are flexible image acquisition tools for such a purpose. However, limited by the carrying capacity of UAV, only consumer-oriented lightweight cameras can be installed on board. As a result, spectral information of the acquired image is not sufficient, and a large number of research questions cannot be directly and effectively investigated using this new type of data source, such as vegetation indices like NDVI, which have great significance for vegetation extraction.

Overcoming this difficulty, this paper proposes a UAV image-cultivated land information extraction method on the basis of triangulation space feature detection (TCLE). The method automatically extracts the center point of a segmentation object, automatically constructs a Delaunay triangulation and Voronoi diagram with a divide-and-conquer algorithm, and clusters the fragmented surface features, such as residential areas, with an AUTOCLUST algorithm. Last, the clustering results are further amended with a Voronoi diagram constraint, clustering mean area constraint, and Max.diff.

Compared with the classification software eCognition, which is a currently popular method for extracting cultivated land information from high-resolution imagery (ECLE), our method (TCLE) not only saves the trouble of selecting samples, but also automatically extracts cultivated land information from UAV high-resolution image data only on the basis of the distribution characteristics of cultivated land. Comparative experiments show that our method can not only guarantee equivalent accuracy to the ECLE method, but also ensure the integrity of cultivated land information extraction.

However, this method mainly considers the fragmentation and relatively compact features of residential areas, as well as contiguous characteristics of cultivated land. It is difficult to remove those plots that are similar to cultivated land after segmentation in the contiguous plots and scattered areas, such as uncultivated grasslands and large tracts of woodland in hilly areas. In addition, it is also easier to remove the clustered small crop fields of similar size to houses [Fig. 13(b)]. Moreover, technology for the extraction of cultivated land information from high-resolution imagery in mountainous regions is worthy of further detailed research. We believe that it is more important to develop an ultrahigh-resolution cultivated land information extraction method or algorithm to exploit the feature advantages of high-resolution remote sensing images. In particular, research into a cultivated land information extraction method for ultrahigh-resolution panchromatic remote sensing images (those images acquired by UAVs, and VHR satellite images, such as those from Quickbird) should put emphasis on the study of algorithms and on the study of the spatial distributions of cultivated land. Further work will extract cultivated land plots and perform research into algorithms that have a strong ability to adapt to different regions, by using the shape and texture information provided by high-resolution image data in combination with a variety of other data sources.

Acknowledgments

This work is supported by the National Natural Science Foundation of China (Grant Nos. 41371017 and 41001238), the National Key Technology R&D Program of China (Grant No. 2012BAH28B02). Sincere thanks are given for the comments and contributions of anonymous reviewers and members of the editorial team.

References

1. S. Liaghat and S. K. Balasundram, "A review: the role of remote sensing in precision agriculture," *Am. J. Agric. Biol. Sci.* **5**(1), 50–55 (2010), http://dx.doi.org/10.3844/ajabssp.2010.50.55.
2. Y. Q. Huang et al., "Research on development of agricultural geographic information ontology," *J. Integr. Agric.* **11**(5), 865–877 (2012), http://dx.doi.org/10.1016/S2095-3119(12)60077-X.
3. Y. X. Cheng et al., "Cold damage risk assessment of double cropping rice in Hunan, China," *J. Integr. Agric.* **12**(2), 352–363 (2013), http://dx.doi.org/10.1016/S2095-3119(13)60235-X.
4. S. W. Dong et al., "Extraction of cultivated land using ETM+ image based on multi scale fractal signature," *Trans. Chin. Soc. Agric. Eng.* **27**(2), 213–218 (2011).
5. M. C. Feng et al., "Monitoring winter wheat freeze injury using multi-temporal MODIS data," *Agric. Sci. China.* **8**(9), 1053–1062 (2009), http://dx.doi.org/10.1016/S1671-2927(08)60313-2.
6. J. L. Ding, M. C. Wu, and T. Tiyip, "Study on soil salinization information in arid region using remote sensing technique," *Agric. Sci. China* **10**(3), 404–411 (2011), http://dx.doi.org/10.1016/S1671-2927(11)60019-9.
7. M. Ozdogan et al., "Remote sensing of irrigated agriculture: opportunities and challenges," *Remote Sens.* **2**(9), 2274–2304 (2010), http://dx.doi.org/10.3390/rs2092274.
8. E. R. Hunt et al., "Acquisition of NIR-green-blue digital photographs from unmanned aircraft for crop monitoring," *Remote Sens.* **2**(1), 290–305 (2010), http://dx.doi.org/10.3390/rs2010290.
9. B. Li, R. Y. Liu, and S. H. Liu, "Monitoring vegetation coverage variation of winter wheat by low-altitude UAV remote sensing system," *Trans. Chin. Soc. Agric. Eng.* **28**(13), 160–165 (2012).
10. J. Primicerio et al., "A flexible unmanned aerial vehicle for precision agriculture," *Precision Agric.* **13**(4), 517–523 (2012), http://dx.doi.org/10.1007/s11119-012-9257-6.
11. J. Scheiewe, L. Tufte, and M. Ehlers, "Potential and problems of multi-scale segmentation methods in remote sensing," *GeoBIT/GIS* **6**, 34–39 (2001).
12. M. E. Hodgson et al., "Synergistic use of lidar and color aerial photography for mapping urban parcel imperviousness," *Photogramm. Eng. Remote Sens.* **69**(9), 973–980 (2003), http://dx.doi.org/10.14358/PERS.69.9.973.
13. A. S. Laliberte et al., "Object-oriented image analysis for mapping shrub encroachment from 1937–2003 in southern New Mexico," *Remote Sens. Environ.* **93**(1), 198–210 (2004), http://dx.doi.org/10.1016/j.rse.2004.07.011.
14. Q. Yu et al., "Object based detailed vegetation classification with airborne high spatial resolution remote sensing imagery," *Photogramm. Eng. Remote Sens.* **72**(7), 799–811 (2006), http://dx.doi.org/10.14358/PERS.72.7.799.
15. R. L. Kettig and D. A. Landgrebe, "Classification of multispectral image data by extraction and classification of homogeneous objects," *IEEE Trans. Geosci. Remote Sens.* **14**(1), 19–26 (1976), http://dx.doi.org/10.1109/TGE.1976.294460.
16. G. J. Hay, K. P. Niemann, and G. F. Mclean, "An object-specific image texture analysis of H-resolution forest imagery," *Remote Sens. Environ.* **55**(2), 108–122 (1996), http://dx.doi.org/10.1016/0034-4257(95)00189-1.
17. D. Geneletti and G. H. Gorte, "A method for object-oriented land cover classification combining Landsat TM data and aerial photographs," *Int. J. Remote Sens.* **24**(6), 1273–1286 (2003), http://dx.doi.org/10.1080/01431160210144499.
18. M. Mueller, K. Segl, and H. Kaufmann, "Edge- and region-based segmentation technique for the extraction of large, man-made objects in high-resolution satellite imagery," *Pattern Recognit.* **37**(8), 1619–1628 (2004), http://dx.doi.org/10.1016/j.patcog.2004.03.001.
19. V. Walter, "Object-based classification of remote sensing data for change detection," *ISPRS J. Photogramm. Remote Sens.* **58**(3–4), 225–238 (2004), http://dx.doi.org/10.1016/j.isprsjprs.2003.09.007.

20. R. Mathieu, C. Freeman, and J. Aryal, "Mapping private gardens in urban areas using object-oriented techniques and very high-resolution satellite imagery," *Landsc. Urban Plan.* **81**(3), 179–192 (2007), http://dx.doi.org/10.1016/j.landurbplan.2006.11.009.
21. A. Rango et al., "Unmanned aerial vehicle-based remote sensing for rangeland assessment, monitoring, and management," *J. Appl. Remote Sens.* **3**(1), 033542 (2009), http://dx.doi.org/10.1117/1.3216822.
22. A. S. Laliberte et al., "Acquisition, orthorectification, and object-based classification of unmanned aerial vehicle (UAV) imagery for rangeland monitoring," *Photogramm. Eng. Remote Sens.* **76**(6), 661–672 (2010), http://dx.doi.org/10.14358/PERS.76.6.661.
23. A. S. Laliberte, C. Winters, and A. Rango, "UAS remote sensing missions for rangeland applications," *Geocarto Int.* **26**(2), 141–156 (2011), http://dx.doi.org/10.1080/10106049.2010.534557.
24. A. S. Laliberte and A. Rango, "Image processing and classification procedures for analysis of sub-decimeter imagery acquired with an unmanned aircraft over arid rangelands," *GIsci. Remote Sens.* **48**(1), 4–23 (2011), http://dx.doi.org/10.2747/1548-1603.48.1.4.
25. P. S. Thenkabail et al., *A Global Irrigated Area Map (GIAM) Using Remote Sensing at the End of the Last Millennium*, International Water Management Institute, Colombo, Sri Lanka (2008).
26. T. Le Toan et al., "Rice crop mapping and monitoring using ERS-1 data based on experiment and modeling results," *IEEE Trans. Geosci. Remote Sens.* **35**(1), 41–56 (1997), http://dx.doi.org/10.1109/36.551933.
27. A. Rosenqvist, "Temporal and spatial characteristics of irrigated rice in JERS-1 L-band SAR data," *Int. J. Remote Sens.* **20**(8), 1567–1587 (1999), http://dx.doi.org/10.1080/014311699212614.
28. Y. Shao, J. Liao, and C. Wang, "Analysis of temporal radar backscatter of rice: a comparison of SAR observations with modeling results," *Can. J. Remote Sens.* **28**(2), 128–138 (2002), http://dx.doi.org/10.5589/m02-019.
29. K. S. Lee and S. I. Lee, "Assessment of post-flooding conditions of rice fields with multitemporal satellite SAR data," *Int. J. Remote Sens.* **24**(17), 3457–3465 (2003), http://dx.doi.org/10.1080/0143116021000021206.
30. A. Bouvet and T. Le Toan, "Use of ENVISAT/ASAR wide-swath data for timely rice fields mapping in the Mekong River Delta," *Remote Sens. Environ.* **115**(4), 1090–1101 (2011), http://dx.doi.org/10.1016/j.rse.2010.12.014.
31. L. Cheng et al., "Fusion of laser scanning data and optical high-resolution imagery for accurate building boundary derivation," *J. Appl. Remote Sens.* **7**(1), 073570 (2013), http://dx.doi.org/10.1117/1.JRS.7.073570.
32. L. H. Tong et al., "Integration of LiDAR data and orthophoto for automatic extraction of parking lot structure," *IEEE J. Sel. Topics Appl. Earth Obs. Remote Sens.* **7**(2), 503–514 (2013), http://dx.doi.org/10.1109/JSTARS.2013.2269193.
33. L. Cheng et al., "Integration of LiDAR data and optical multi-view images for 3D reconstruction of building roofs," *Opt. Lasers Eng.* **51**(4), 493–502 (2013), http://dx.doi.org/10.1016/j.optlaseng.2012.10.010.
34. Y. X. Liu et al., "High quality prime farmland extraction pattern based on object-oriented image analysis," in *Proc of Geoinformatics 2008 and Joint Conf. on GIS and Built Environment: Classification of Remote Sensing Images*, pp. 71470–71471, SPIE, Guangzhou, China (2008).
35. J. Shen et al., "Cropland extraction from very high spatial resolution satellite imagery by object-based classification using improved mean shift and one-class support vector machines," *Sens. Lett.* **9**(3), 997–1005 (2011), http://dx.doi.org/10.1166/sl.2011.1361.
36. S. Lu et al., "Comparison between several feature extraction/classification methods for mapping complicated agricultural land use patches using airborne hyperspectral data," *Int. J. Remote Sens.* **28**(5), 963–984 (2007), http://dx.doi.org/10.1080/01431160600771561.
37. X. D. Sun and H. Q. Xu, "Comer extraction algorithm for high-resolution imagery of agricultural land," *Trans. Chin. Soc. Agric. Eng.* **25**(10), 235–241 (2009).

38. A. Rango, A. Laliberte, and C. Winters, "Role of aerial photos in compiling a long-term remote sensing data set," *J. Appl. Remote Sens.* **2**(1), 023541 (2008), http://dx.doi.org/10.1117/1.3009225.
39. A. Rango and A. S. Laliberte, "Impact of flight regulations on effective use of unmanned aircraft systems for natural resources applications," *J. Appl. Remote Sens.* **4**(1), 043539 (2010), http://dx.doi.org/10.1117/1.3474649.
40. A. Horcher and R. J. M. Visser, "Unmanned aerial vehicles: applications for natural resource management and monitoring," presented at *Council on Forest Engineering Annual Meeting*, Hot Springs, USA (27–30 April 2004).
41. R. Sugiura, N. Noguchi, and K. Ishii, "Remote-sensing technology for vegetation monitoring using an unmanned helicopter," *Biosyst. Eng.* **90**(4), 369–379 (2005), http://dx.doi.org/10.1016/j.biosystemseng.2004.12.011.
42. A. S. Laliberte, C. Winters, and A. Rango, "Acquisition, orthorectification, and classification of hyperspatial UAV imagery," 2007, http://www.tucson.ars.ag.gov/rise/2007/Posters/19SlaughterPoster.pdf (29 January 2014).
43. J. Egbert and R.W. Beard, "Road following control constraints for low altitude miniature air vehicles," in *American Control Conf.*, pp. 353–358, IEEE, New York (2007).
44. N. Masahiko et al., "UAV borne mapping system for river environment," presented at *Proc. of 28th Asian Association of Remote Sensing Conf.*, ACRS, Kuala Lumpur, Malaysia (12–16 November 2007).
45. S. Spatalas, V. Tsioukas, and M. Daniil, "The use of remote controlled helicopter for the recording of large scale urban and suburban sites," presented at *Culture of Representation*, Xanthi, Greece (March 2006).
46. P. Patias et al., "Photogrammetric documentation and digital representation of the Macedonian palace in Vergina-Aegeae," presented at *Proc. of 2007 XXI Int. CIPA Symp.*, Camera and Imaging Products Association, Athens, Greece (1–6 October 2007).
47. R. B. Haarbrink and H. Eisenbeiss, "Accurate DSM production from unmanned helicopter systems," *Int. Arch. Photogramm., Remote Sens. Spatial Inf. Sci.* **XXXVII**(Pt. B1), 1259–1264 (2008).
48. R. Dunford et al., "Potential and constraints of unmanned aerial vehicle technology for the characterization of Mediterranean riparian forest," *Int. J. Remote Sens.* **30**(19), 4915–4935 (2009), http://dx.doi.org/10.1080/01431160903023025.
49. J. Y. Lin et al., "Experimental observation and assessment of ice conditions with a fixed-wing unmanned aerial vehicle over Yellow River, China," *J. Appl. Remote Sens.* **6**(1), 063586 (2012), http://dx.doi.org/10.1117/1.JRS.6.063586.
50. R. Salvini, S. Riccucci, and M. Francioni, "Topographic and geological mapping in the prehistoric area of Melka Kunture (Ethiopia)," *J. Maps* **8**(2), 169–175 (2012), http://dx.doi.org/10.1080/17445647.2012.680779.
51. N. Han et al., "NDVI-based lacunarity texture for improving identification of torreya using object-oriented method," *Agric. Sci. China* **10**(9), 1431–1444 (2011), http://dx.doi.org/10.1016/S1671-2927(11)60136-3.
52. V. Estivill-Castro and I. Lee, "Argument free clustering for large spatial point data sets via boundary extraction from Delaunay diagram," *Comput. Environ. Urban Syst.* **26**(4), 315–334 (2002), http://dx.doi.org/10.1016/S0198-9715(01)00044-8.
53. B. Paul, "Calculating the area and centroid of a polygon," 1988, http://www.seas.upenn.edu/~sys502/extra_materials/Polygon%20Area%20and%20Centroid.pdf (11 Febraury 2014).
54. S. H. Lo, "Delaunay triangulation of non-uniform point distributions by means of multi-grid insertion," *Finite Elem. Anal. Des.* **63**, 8–22 (2013), http://dx.doi.org/10.1016/j.finel.2012.08.005.
55. B. Zalik and I. Kolingerova, "An incremental construction algorithm for Delaunay triangulation using the nearest-point paradigm," *Int. J. Geogr. Inf. Sci.* **17**(2), 119–138 (2003), http://dx.doi.org/10.1080/713811749.
56. R. A. Dwyer, "A faster divide-and-conquer algorithm for constructing Delaunay triangulations," *Algorithmica* **2**, 137–151 (1987), http://dx.doi.org/10.1007/BF01840356.

57. I. Lee and V. Estivill-Castro, "Fast cluster polygonization and its applications in data-rich environments," *Geoinformatica* **10**(4), 399–422 (2006), http://dx.doi.org/10.1007/s10707-006-0340-x.
58. S. Guha, R. Rastogi, and K. Shim, "CURE: an efficient clustering algorithm for large databases," in *Proc. of the 1998 ACM SIGMOD Int. Conf. on Management of Data*, pp. 73–84, ACM, New York (1998).
59. G. M. Foody, "Status of land cover classification accuracy assessment," *Remote Sens. Environ.* **80**(1), 185–201 (2002), http://dx.doi.org/10.1016/S0034-4257(01)00295-4.
60. Definiens, *Definiens eCognition Developer 8.0 User Guide*," Definens AG, Munich, Germany (2009).
61. S. K. Bo and L. Ding, "The effect of the size of training sample on classification accuracy in object-oriented image analysis," *J. Image Graph.* **15**(7), 1106–1111 (2010).
62. R. G. Congalton and K. Green, "Assessing the accuracy of remotely sensed data: principles and practices," *Photogramm. Rec.* **25**(130), 204–205 (2010), http://dx.doi.org/10.1111/phor.2010.25.issue-130.

Biographies of the authors are not available.