Automatic Detection and Segmentation of Orchards Using Very High Resolution Imagery

Selim Aksoy, *Senior Member, IEEE*, Ismet Zeki Yalniz, and Kadim Taşdemir, *Member, IEEE*

Abstract—Spectral information alone is often not sufficient to distinguish certain terrain classes such as permanent crops like orchards, vineyards, and olive groves from other types of vegetation. However, instances of these classes possess distinctive spatial structures that can be observable in detail in very high spatial resolution images. This paper proposes a novel unsupervised algorithm for the detection and segmentation of orchards. The detection step uses a texture model that is based on the idea that textures are made up of primitives (trees) appearing in a near-regular repetitive arrangement (planting patterns). The algorithm starts with the enhancement of potential tree locations by using multi-granularity isotropic filters. Then, the regularity of the planting patterns is quantified using projection profiles of the filter responses at multiple orientations. The result is a regularity score at each pixel for each granularity and orientation. Finally, the segmentation step iteratively merges neighboring pixels and regions belonging to similar planting patterns according to the similarities of their regularity scores and obtains the boundaries of individual orchards along with estimates of their granularities and orientations. Extensive experiments using Ikonos and QuickBird imagery as well as images taken from Google Earth show that the proposed algorithm provides good localization of the target objects even when no sharp boundaries exist in the image data.

Index Terms—Orientation estimation, periodic signal analysis, regularity detection, texture analysis, texture segmentation.

I. INTRODUCTION

REMOTE sensing has been a very valuable tool for agricultural studies. In particular, remotely sensed imagery can be used as an effective way for locating, delineating and classifying agricultural sites, monitoring their change in time over large areas, and identifying potential plantation areas for decision makers. For example, the Control with Remote Sensing campaign within the frame of the European Union's (EU) Common Agricultural Policy (CAP) includes detailed guidelines about how very high spatial resolution (VHR) imagery can be used for parcel identification and categorization.

S. Aksoy is with the Department of Computer Engineering, Bilkent University, 06800 Ankara, Turkey (e-mail: saksoy@cs.bilkent.edu.tr).

I. Z. Yalniz was with the Department of Computer Engineering, Bilkent University, Ankara, 06800, Turkey. He is now with the Department of Computer Science, University of Massachusetts, Amherst, MA 01003-9264 USA (e-mail: zeki@cs.umass.edu.).

K. Taşdemir is with the European Commission, Joint Research Centre, Institute for Environment and Sustainability, Monitoring Agricultural Resources Unit, 21027 Ispra, Italy (e-mail: kadim.tasdemir@jrc.ec.europa.eu).

However, realization of these guidelines in a wide scale is an extremely laborious task because they are often implemented using manual photo-interpretation. Consequently, development of automatic and robust methods has become an important research problem when the analysis goes beyond local sites to cover a wide range of landscapes in national and even international levels.

Availability and wide coverage of VHR imagery has enabled detailed analysis of agricultural sites in the scale of individual plants. However, even though it may be possible to perform a dichotomous vegetation versus nonvegetation classification [1], the traditional approach for land cover classification using pixel-based spectral information has been unsatisfactory because it is often not possible to discriminate between certain terrain classes using only spectral information in VHR images with limited spectral resolution. A popular alternative to pixel-based spectral classification has been object-based image analysis that relies on image segmentation before feature extraction and classification. However, segmentation algorithms aim to find image regions that satisfy some form of homogeneity criteria, but defining such criteria and setting the corresponding parameters for accurate segmentation of VHR images is still a very hard task with robustness, generalizability, and repeatability problems in different landscapes.

Another approach is to design methods for automatic detection of specific target landscape features by incorporating different types of information exploiting their peculiarities [2]. For example, we developed an algorithm that combined spectral, textural, and shape information for the detection of linear strips of woody vegetation such as hedgerows and riparian vegetation that are important elements of the landscape ecology and biodiversity [3]. Other target objects of particular interest include permanent crops like orchards comprising fruit or nut-producing trees, vineyards, and olive groves. For example, permanent crops are of great importance economically as well as in terms of spatial coverage in Europe. Therefore, the EU CAP regulations (regulation EC 73/2009, previously EC 1782/2003) include several support and payment schemes for permanent crops, including nuts (hazelnuts, almonds, walnuts, pistachios, locust beans) among some other crop production. Due to different aid schemes concerning this sector, many member states dedicate much effort to the control of their subsidies. For example, the existing subsidies for nuts, regulated in EC 73/2009, are only applied if eligibility conditions for the application of support schemes for nuts, regulated by EC 1121/2009, are satisfied. Two main conditions are: 1) homogeneous and geographically continuous orchards are eligible whereas isolated trees and single rows of trees alongside roads or parcels are not eligible; 2) the land should have a minimum density not less than 125 trees/ha for hazelnut orchards, 50 trees/ha for walnuts,
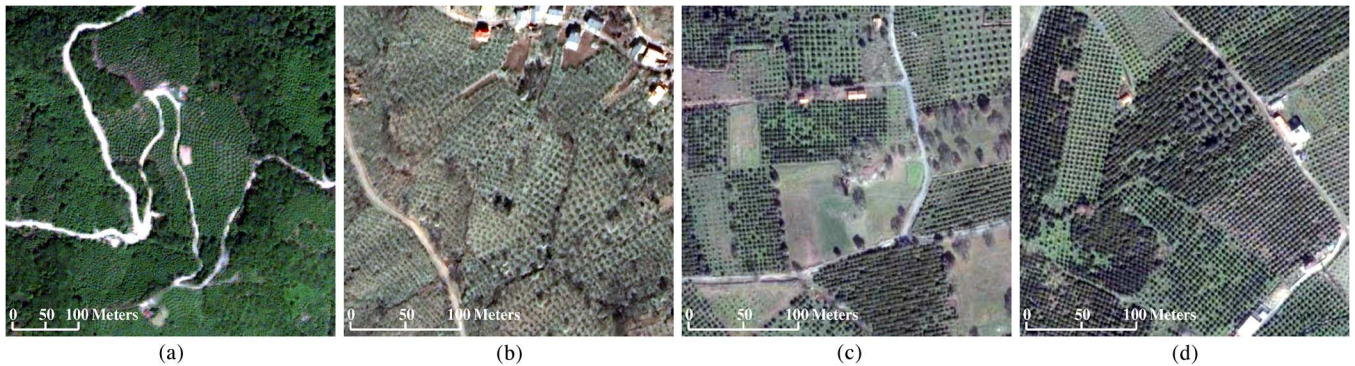
Fig. 1.   Example 500 × 500 pixel images containing orchards. The data sets are described in Section II. Pansharpened multispectral data are shown, but only the panchromatic information was used in the study. Contrast enhancement was performed on all raster images in this paper for better visualization. (a) Giresun—Ikonos; (b) Giresun—QuickBird; (c) Izmir—Google Earth; (d) Izmir—Google Earth.

almonds and pistachios, and 30 trees/ha for locust bean groves. In order to improve the quality and efficiency of the control tasks for permanent crops, there is a considerable potential through a better use of the increasing amount of VHR imagery.

A common property of these objects is that they all exhibit a wide range of spectral characteristics and a large amount of within-class variability in their appearance that make their detection a very hard problem using traditional land cover analysis techniques. Yet, they also possess distinctive spatial structures that can be observable in detail in VHR images with less than 1-m spatial resolution. Therefore, modeling of these spatial structures is of great interest for the development of flexible and effective solutions for their detection in diverse landscapes. At the same time, it is important that these models take into consideration the scale variations and irregularities in the planting patterns for the robustness of the detection process.

This paper describes our work on the automatic detection and segmentation of orchards in VHR imagery. As shown in Fig. 1, orchards consist of trees or shrubs planted in a regular grid structure with a grass or bare soil base. Texture analysis has been the common choice for the recognition of such structures. Even though forests and grasslands can be detected using microtexture methods such as co-occurrence analysis [4] or morphological transformations [5], structures such as orchards and other similar structures such as vineyards and olive groves require macrotexture methods that can model the spatial arrangements in the planting patterns of individual trees. For example, Warner and Steinmaus [6] used the variance of the distance between the peaks of an autocorrelogram computed at a particular orientation within a fixed window as a feature for the identification of pixels belonging to orchards. The autocorrelograms were computed at four principal orientations and a threshold on variance was used for the final detection. Trias-Sanz [7] analyzed the variograms of image patches to estimate primary orientations and used a set of rules based on the periodicities along these orientations for the classification of the patches as forests, orchards, vineyards, tilled fields, and untilled fields. Amoruso et al. [8] applied this method to the characterization of olive groves. Ruiz et al. [9] produced an image where each tree was represented by a pixel by using local thresholding of the input data and used features obtained from semivariogram analysis, Hough transform, histogram of minimum distances, and Fourier transform on this image for classification of parcels as regular versus nonregular. Tasdemir [10]

used a self-organizing map with the spectral values of the pixels within a fixed window as the features of the center pixel to classify that pixel as hazelnut, forest, urban, bare area, and agriculture. Helmholz and Rottensteiner [11] used a rule-based classifier combining texture analysis using Markov random fields, image segmentation, structural analysis using Hough transform for the detection of parallel lines, and radiometric analysis using normalized difference vegetation index (NDVI) for the verification of cropland and grassland objects in a GIS. Ranchin et al. [12] used a texture index derived from the density of edges computed using wavelet analysis for the differentiation of reference parcels as vine or nonvine. Wassenaar et al. [13] analyzed the peaks of the Fourier transform of an image parcel to estimate the spacing and orientation of the vineyard structure in that parcel. Chanussot et al. [14] used the Radon transform of the Fourier spectrum to estimate the orientations of vineyards in image patches. Delenne et al. [15] compared the contrast feature computed from co-occurrence matrices with the features computed from the peaks of the Fourier spectrum where the vine versus nonvine classification was obtained by thresholding these features at each pixel. Rabatel et al. [16] also used frequency analysis for vine plot detection. After computing the Fourier spectrum for an image window, they found the highest peak, designed a Gabor filter with scale and orientation parameters extracted from this peak, thresholded the Gabor filter output to obtain the plot corresponding to these parameters, and repeated this procedure with the next highest peak in the spectrum. Delenne et al. [17] improved the delineation of vineyard boundaries by fitting lines to the vine rows within the detected vineyards, and adjusting these lines based on the local minima of image values and interrow widths. The initial vineyard detection in [17] chose the method in [15] over the one in [16] as the former was much simpler to implement while providing equivalent results.

A common property of most of these methods [9], [11]–[13] is that they aim to classify existing parcels that are assumed to contain single orchards or vineyards that are uniform in texture. However, this assumption cannot always be satisfied in a very large scale detection task because the parcel boundaries may not be available for many areas or an available agricultural parcel may consist of different units [11] that violate the homogeneity assumption. On the other hand, an image-wide prior segmentation cannot be used to obtain the target parcels because the structures of interest do not exhibit uniform spectral, textural, or

morphological characteristics expected by the state-of-the-art segmentation algorithms. Similarly, the methods that classify individual pre-extracted patches [7], [14] are limited in their capacity of modeling different scales and locally varying orientations because a fixed set of parameters are often considered within each patch. Furthermore, the methods that make decisions on individual pixels centered within fixed size windows [6], [10] may result in a lot of false positives (commission) and false negatives (omission) when there is a large amount of local within-class variation such as different plant spacings and curved rows.

In this paper, we describe a new unsupervised method for simultaneous detection and segmentation of orchards in complex scenes. While detection is a very hard problem when images contain multiple orchards with significantly varying scales and orientations, segmentation is even harder when such orchards neighbor each other and other textured areas. Our algorithm, first, localizes regular spatial structures as candidate regions and then employs an iterative region growing and merging process to segment the individual orchards. The localization step adapts a model that we recently proposed for structural texture analysis [18]. The model uses the idea that textures are made up of primitives appearing in a near-regular repetitive arrangement. The texture model for the orchards involves individual trees that can appear at different sizes with regular spatial patterns at gradually changing orientations. The tree size is related to the granularity of the texture primitives, and the planting pattern corresponds to the structural properties of the texture. The detection algorithm uses periodic signal analysis to compute a regularity score at each pixel for a given range of granularities and orientations. Then, the segmentation algorithm iteratively merges neighboring pixels and regions belonging to similar planting patterns according to the similarities of their regularity scores, and obtains the boundaries of individual orchards.

The contributions of this paper include the extension of the texture model for orchard localization to handle computational issues for processing large amounts of remotely sensed data, and the segmentation algorithm for the delineation of individual orchards. Our approach differs from other related work in that it can simultaneously localize multiple orchards along with estimates of their granularities and orientations in complex images containing different kinds of textures as well as nontextured areas even when no sharp boundaries exist in the image data. The unsupervised localization and delineation ability eliminates the requirement of existing parcel information and enables large-scale studies. The ability to estimate granularities and orientations eliminates the need for prior information about the planting patterns such as crown sizes and placement periods where the rows of trees have to be approximately equally spaced. Furthermore, the algorithm is robust to variations in granularity and irregularities in the spatial patterns because the local processing can handle missing or overlapping crowns and gradually changing orientations that are typical problems in areas with a varying surface structure and a strong relief.

The rest of the paper is organized as follows. The study sites and the corresponding data are described in Section II. Periodic signal analysis and its application for multi-granularity and multi-orientation regularity detection are discussed in Section III. The detection and segmentation algorithm for the delineation of individual orchards is described in Section IV. Quantitative performance evaluation using Ikonos and Quick-Bird data as well as images captured from Google Earth is presented in Section V. Finally, conclusions are given in Section VI.

II. STUDY SITES AND DATA

A. Giresun Data Set

The first data set used in this paper is from the Giresun province in the Black Sea region of Turkey. Being an accession country to the EU, Turkish authorities work on establishing a national Land Parcel Identification System and on improving the information on specific sectors to be reformed, through assessment of the potential of VHR imagery to map crops and to discriminate among them. Among many crops, permanent crops (mainly hazelnut, olive, and citrus fruits) cover a considerable part of the arable and more marginal land in Turkey. In particular, about 70% of the world's hazelnut production is produced along the Black Sea coast in northern Turkey. A specific property of the region is the strong relief, which makes hazelnut production the main cultivation there. While the assessment of hazel orchards in the area is of great importance for a national reform policy, it also poses an interesting research challenge in terms of permanent crop detection: the hazel orchards in the region are often small and have a high planting density relative to orchards in other countries, and the natural vegetation in the area can also be spatially discontinuous.

The data set used for evaluating the performance of the detection of hazel orchards in Turkey includes a panchromatic Ikonos image (1-m spatial resolution, covering an area of 147 km2, acquired in 2007) of the Merkez county (referred to as *Merkez07*), a panchromatic QuickBird image (0.6-m spatial resolution, covering an area of 80 km2, acquired in 2006) of the Yaglidere county (referred to as *Yaglidere06*), and another panchromatic QuickBird image (0.6-m spatial resolution, covering an area of 145 km2, acquired in 2008) of the Merkez county (referred to as *Merkez08*) in the province of Giresun. 15 subscenes, each with size 1000 × 1000 pixels, were cut from these images (five subscenes were used from each image) as a diverse sample of orchards with different characteristics.

B. Izmir Data Set

The second data set is from the Seferihisar county of the Izmir province in the Aegean coast of Turkey. Seven images, each with size 1680 × 1031 pixels, that were saved from Google Earth are used as another set of test cases including orchards with citrus trees such as tangerine and satsuma. The zoom level for these images was adjusted so that the spatial resolution was as close as possible to that of the QuickBird images. The color images were converted to gray scale by adding together 30% of the red value, 59% of the green value, and 11% of the blue value [19]. These images are referred to as *Izmir* in the rest of the paper.

III. REGULARITY DETECTION

Regularity detection aims to quantify the local structure in images so that areas consisting of near-regular repetitive arrangements of trees achieve higher regularity scores. These

scores can then be used to localize the orchards. However, since an image may contain multiple orchards involving trees with different sizes planted along different dominant orientations, accurate delineation of orchards also requires the estimation of these granularities and orientations. The Fourier spectrum or variogram-based approaches that are often found in related work are natural ways of modeling regularity, but they are limited in their capacity of simultaneous localization and estimation. In particular, the algorithm in [7] can produce an orientation estimate for the periodic structure but does this for whole image windows where localization is not straightforward. That algorithm also has a lot of parameters that may be hard to adjust for different windows as observed in [8]. The algorithm in [16] can handle multiple scales and orientations by thresholding the peaks in the Fourier spectrum and can localize the corresponding structures using carefully designed Gabor filters, but can do this only in small windows (e.g., 500×500) with only a few structures so that the Fourier spectrum has clear peaks corresponding to rows of trees that are approximately equally spaced.

The regularity detection algorithm in this paper adapts the structural texture model proposed in [18]. The algorithm starts with the enhancement of potential tree locations by using an isotropic spot filter. A pyramid scheme is used where a fixed filter enhances tree-like objects in each level that also reduces the image size for fast processing of large VHR images. Then, the local extrema in the filter responses are assumed to correspond to potential tree locations without any strict requirement for their exact detection, and the structure of these locations is analyzed by converting the image data into 1-D signals by using projection profiles within oriented sliding windows. The regularity of the texture primitives, i.e., trees, is quantified in terms of the periodicity of these projection profiles at multiple orientations. The result is a regularity score at each pixel for each granularity and each orientation. The details of these steps are described below.

A. Preprocessing

The first step in the modeling of orchards using the arrangements of individual trees is the enhancement of the tree-like objects in the image. Using the assumption that individual trees appear approximately as circular blobs with intensities generally darker than their surroundings in visible wavelengths, a suitable filter for such enhancement is the spot filter. The isotropic Laplacian of Gaussian (LoG) filter shown in Fig. 2(a) is a well-known spot detector [20]. Given the Laplacian operator

$$\nabla^2 = \frac{\partial^2}{\partial r^2} + \frac{\partial^2}{\partial c^2} \tag{1}$$

and the 2-D Gaussian function

$$G(r, c) = e^{-\frac{r^2+c^2}{2\sigma^2}} \tag{2}$$

with the scale parameter σ, the expression for the LoG filter expressed in row (r) and column (c) coordinates can be obtained as

$$\nabla^2 G(r, c) = \frac{\partial^2 G(r, c)}{\partial r^2} + \frac{\partial^2 G(r, c)}{\partial c^2}$$
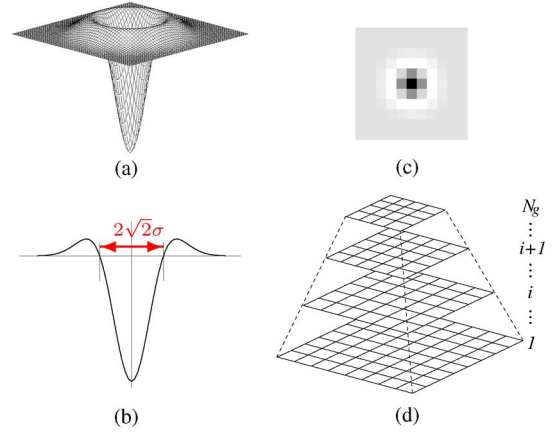

Fig. 2. (a) Three-dimensional plot of the Laplacian of Gaussian filter. (b) The filter cross-section showing zero crossings. (c) The filter with granule size of 3 pixels. (d) Illustration of the image pyramid used for multi-granularity analysis.

$$= \frac{\partial}{\partial r}\left(\frac{-r}{\sigma^2}e^{-\frac{r^2+c^2}{2\sigma^2}}\right) + \frac{\partial}{\partial c}\left(\frac{-c}{\sigma^2}e^{-\frac{r^2+c^2}{2\sigma^2}}\right)$$

$$= \left(\frac{r^2+c^2-2\sigma^2}{\sigma^4}\right)e^{-\frac{r^2+c^2}{2\sigma^2}}. \tag{3}$$

The selection of σ depends on the scales of the spots, i.e., trees, of interest. Since the length of the cross section between the zero crossings of the LoG filter is $2\sqrt{2}\sigma$ as shown in Fig. 2(b), the σ parameter can be selected according to the granularities of the trees in the target orchards. Different granularities can be modeled using a multiscale approach where a separate spot filter is designed for each granularity [18]. Alternatively, in this paper, we use a fixed spot filter in a pyramid scheme to handle multiple granularities in a given image. Given the minimum and maximum granularities of interest, g_{min} and g_{max}, respectively, in pixels, a set of granularities $\mathcal{G} = \{g_1, \ldots, g_{N_g}\}$ is computed as

$$g_i = \begin{cases} g_{min}, & i=1 \\ \sqrt{2}g_{i-1}, & i=2,\ldots,N_g \end{cases} \tag{4}$$

where $N_g = \lfloor 2\log_2(g_{max}/g_{min}) + 1 \rfloor$. The scale factor of $\sqrt{2}$ between two consecutive granularities is chosen for scale invariance [21]. Next, given a fixed spot filter with parameter σ_0, the i'th level in the pyramid shown in Fig. 2(d) is constructed by reducing the size of the original image by a scale factor of σ_0/σ_i using bilinear interpolation where $\sigma_i = g_i/(2\sqrt{2})$, $i=1,\ldots,N_g$ as shown in Fig. 2(b). Then, the resulting images in each level of the pyramid are filtered using the LoG filter with parameter σ_0. We use a spot filter with granule size (cross section) of 3 pixels, corresponding to $\sigma_0 = 3/(2\sqrt{2})$ and selected as the simplest possible spot filter, as shown in Fig. 2(c). In addition to being a well-founded way of handling multiple granularities, the pyramid scheme also enables faster processing at increasing levels due to the reduced image size.

B. Projection Profiles

The pixels having high responses (local maxima) in the image that is enhanced using a particular spot filter indicate

possible locations of tree-like objects at that particular granularity. Furthermore, in a neighborhood with a regular planting pattern, the locations of local maxima along a line with an orientation that matches the dominant direction of this pattern will also follow a regular repetitive structure. We measure the existence of the regularity of the local extrema along a particular orientation in a particular spot filter output using the projection profile along that orientation in an image window. Given a window constructed symmetrically on both sides of a line representing a particular orientation, the vertical projection profile is computed as the summation of the values in individual columns in perpendicular direction to the line.

Fig. 3(b) shows the vertical projection profile of a window cropped from the spot filter response of a QuickBird image shown in Fig. 3(a). It can be observed that the profile contains a signal with a periodic structure consisting of successive peaks with similar shapes corresponding to the trees if the orientation of the window matches the orientation of the regular planting pattern. Furthermore, regularity along multiple parallel image rows enhances these peaks as well. On the other hand, if there is no significant regular pattern or if the orientation of the window does not match that of the orchard, the peaks have arbitrary shapes. In practice, it is quite unlikely to observe perfectly periodic signals in the projection profiles computed from real images. This is mostly due to various factors such as variations in the granularities of the trees, their imperfect arrangements, missing or overlapping crowns, and the restrictions of the terrain on the planting pattern. Hence, the analysis of the projection profile for periodicity should support a relaxed definition of regularity and should take such complications into account.

C. Periodic Signal Analysis

We assume that the regularity of the planting pattern along a particular orientation at a particular granularity can be quantified in terms of the periodicity of the corresponding projection profile. The goal of this step is to measure the amount of periodicity and to locate the periodic part within the larger profile signal. The proposed periodic signal analysis is based on the alternating pattern of peaks and valleys in the profile signal where the peaks correspond to the trees and the valleys correspond to the spacing between the trees. In addition to providing an abstraction level compared to the absolute signal values, the representation based on peaks and valleys makes the analysis invariant to intensity and granularity variations as described below.

The projection profile $x[n]$, $n = 1, \ldots, N_p$ where N_p is the window width in pixels, is segmented into peaks and valleys by finding the zero crossings, the local minima in the positive plane, and the local maxima in the negative plane as shown in Fig. 3(c). The analysis is based on three constraints. The first constraint uses a new sequence of width values $x_w[i]$, $i = 1, \ldots, N_s$ where N_s is the total number of peaks and valleys in the segmented projection signal. The values in this width signal, which correspond to the crown sizes and the amount of spacing between the trees, are expected to be similar to each other in a periodic signal corresponding to a regular pattern. Since the intensity level in the original image affects only the heights of the peaks and valleys, the use of width values also
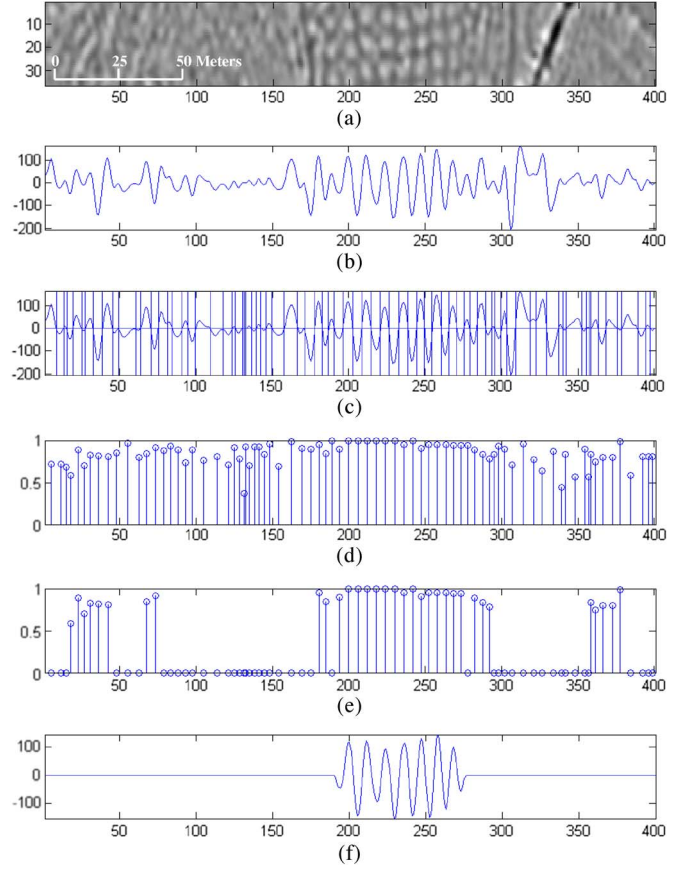

Fig. 3. Periodic signal analysis for the projection profile of an example image window. (a) A window of 400×35 pixels cropped from the spot filter response of a QuickBird image. (b) Vertical projection profile of the window. (c) Segmentation of the profile into its peaks and valleys. (d) Regularity scores computed using (5). (e) Regularity scores after elimination using peak-valley alternation and peak width constraints. (f) Periodic intervals of the profile located by thresholding the values in 3(e) by 0.9 and eliminating the isolated points.

enables invariance to intensity variations. Our initial model in [18] used a two-level wavelet decomposition of the width signal to check for the existence of high-frequency components indicating irregular instances of peaks and valleys. The energies of the detail coefficients of the two-level wavelet transform using the Haar filter were used to compute an irregularity score that quantified the amount of variations in the width values where the scores close to zero were candidates to be part of a regular periodic signal. One limitation of that model for the localization of orchards in VHR imagery is the downsampling by four during the two-level wavelet transform. An upsampling of the results by four is needed to reconstruct a value for each peak and valley, but this process may blur the local details.

In this paper, we use a direct computation of a regularity score using sliding windows without any need for upsampling where

$$x_{reg}[i] = 1 - \frac{1}{2} \left| \frac{x_w[i-1] - x_w[i]}{x_w[i-1] + x_w[i]} - \frac{x_w[i+1] - x_w[i+2]}{x_w[i+1] + x_w[i+2]} \right| \tag{5}$$

for $i = 2, \ldots, N_s - 2$ is the regularity score that is in the $[0, 1]$ range. The values that are close to 1 indicate high regularity, and increasing degrees of deviations from the

regular pattern due to various possible local distortions discussed above make the score approach to 0. The second part in (5) corresponds to the wavelet energies. The numerators inside the absolute value compute the detail coefficients in the first level of the wavelet transform, the denominators provide normalization using the average coefficients of the wavelet transform to enable invariance of the resulting values to different granularities, and the subtraction corresponds to the second level of the wavelet transform. The direct computation in (5) results in a separate value for each peak and valley with only a total of three undefined terms at the image boundaries.

The second constraint selects the parts of the signal where there are alternating peaks and valleys corresponding to a regular planting pattern of trees and the spacing between the trees where trees and ground must follow each other. This corresponds to the elimination of the consecutive peak-peak and valley-valley pairs from the sequence x_{reg} by setting the corresponding values to 0.

Finally, the third constraint eliminates the peaks whose width values are too small (< 2 pixels) or too large (> 5 pixels) with respect to the expected sizes of the trees (3 pixels) in a particular granularity. Fig. 3(d) shows the regularity scores computed using (5), Fig. 3(e) shows the resulting values after elimination based on the second and the third constraints, and Fig. 3(f) shows the parts of the projection profile detected to correspond to a regular pattern by thresholding the small values in Fig. 3(e). In addition, isolated points remaining after thresholding are also removed because multiple peaks and valleys are needed to form a regular pattern.

D. Multi-Granularity and Multi-Orientation Analysis

An image may contain orchards that have different dominant orientations and are composed of trees at different granularities. Therefore, multiple granularities are approximated using the pyramid scheme as described in Section III-A, and the projection profiles for multiple orientations are analyzed by sliding image-wide oriented windows over each spot filter output as described in Sections III-B and C.

Each window is characterized by a line corresponding to the symmetry axis of the window and a height parameter defining the extent of the window on both sides of this line. The symmetry axis is parametrized by a distance parameter d and an orientation parameter θ that are measured with respect to the center pixel of the image. Given an image with N_r rows and N_c columns at a particular granularity, the window is defined using the inequality

$$|r\cos(\theta) - c\sin(\theta) - d| < \frac{\delta}{2} \qquad (6)$$

where r and c are the row and column coordinates, respectively, δ is the window height, and θ is measured relative to the horizontal axis in clockwise direction. For each pixel, all combinations of $d \in [-\sqrt{(N_r/2)^2 + (N_c/2)^2}, \sqrt{(N_r/2)^2 + (N_c/2)^2}]$ and $\theta \in [-90°, 90°)$ values produce a set of windows with symmetry axes passing through that pixel at 180 different orientations as shown in Fig. 4. The θ values are restricted to $[-90°, 90°)$ due to the symmetries.

The projection profile corresponding to each window is computed using summation along $\theta + 90°$, and the regularity scores
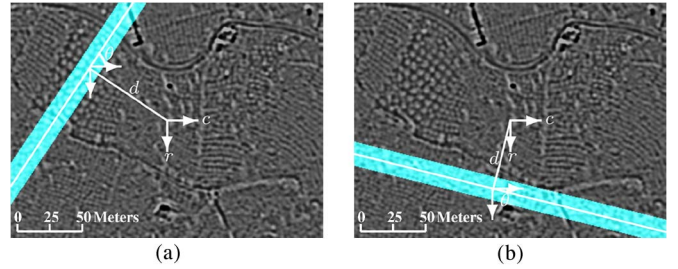

Fig. 4. Example windows for computing the projection profiles. Each window is marked as green together with the symmetry axis that is marked as white. Both images are 400×300 pixels. (a) $d = -120$, $\theta = -55°$, $\delta = 31$. (b) $d = 90$, $\theta = 15°$, $\delta = 31$.

calculated for each peak and valley as in (5) are recorded back to the corresponding pixels on the symmetry axis defining that window. The resulting regularity scores for all granularities $g \in \mathcal{G} = \{g_1, \ldots, g_{N_g}\}$ and all orientations $\theta \in \Theta = [-90°, 90°)$ for each pixel (r, c) are stored in a 4-D matrix denoted as $\rho(r, c; g, \theta)$ where the values for granularities larger than g_1 are upsampled to the original resolution using nearest neighbor interpolation. We denote the scores for a particular pixel for all granularities and all orientations as the regularity spectrum of that pixel. Example spectra are shown in Fig. 5.

IV. ORCHARD DETECTION AND SEGMENTATION

The detection and segmentation algorithm makes use of the regularity of the planting pattern in the neighborhood of each pixel for all granularities and orientations of interest, which is quantified by the matrix $\rho(r, c; g, \theta)$ calculated in Section III. In [18], we computed the maximum regularity score for each pixel from all granularities and orientations, and thresholded the resulting scores to obtain a binary map that separates the regular areas from the rest of the image. In this paper, we propose a segmentation algorithm based on the whole set of granularities and orientations, to obtain the boundaries of individual orchards that may neighbor other orchards or other textured areas.

The algorithm starts with filtering the regularity scores $\rho(r, c; g, \theta)$ for each granularity and orientation with a Gaussian smoothing filter to suppress the noisy values. Then, a region growing and merging process is run as follows.

1) Compute the maximum regularity score for each pixel as

$$\rho^*(r, c) = \max_{g, \theta} \rho(r, c; g, \theta). \qquad (7)$$

2) Identify all pixels whose regularity scores $\rho^*(r, c)$ are greater than a threshold τ_h, and construct a list $\mathcal{L}_0$ that contains these pixels in descending order of their scores. These pixels are used as seeds for region growing.

3) For each pixel (r_k, c_k) in $\mathcal{L}_0$, flood fill its surrounding region as follow.

   a) Initialize the region $\mathcal{R}_k$ with the pixel (r_k, c_k), and set the regularity score for $\mathcal{R}_k$ as $\rho(\mathcal{R}_k; g, \theta) = \rho(r_k, c_k; g, \theta)$.

   b) Construct a new list $\mathcal{L}_k$ that contains the previously unprocessed neighbors of (r_k, c_k) that have a maximum regularity score greater than a threshold τ_l where $\tau_l < \tau_h$. This list contains the candidates for growing.
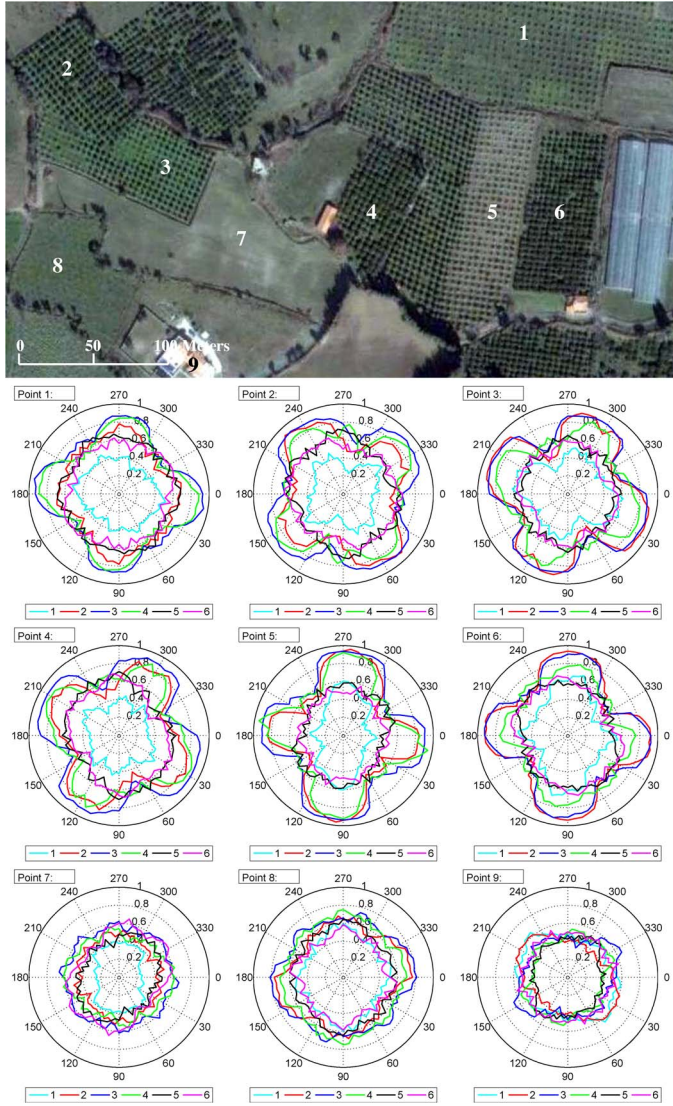
Fig. 5. Example regularity spectra for pixels belonging to different structures in an image with 750×420 pixels. The pixels of interest are marked on the color image with numbers. The pixels from 1 to 6 belong to orchards, and the ones from 7 to 9 belong to other kinds of objects. The spectra are shown as polar plots with the radii in $[0, 1]$ representing the regularity scores as a function of the angles representing the orientations increasing in clockwise direction. The orientations in $[-90°, 90°)$ are mapped to $[0°, 360°)$ for a complete spectrum. The spectra for six different granularities are shown with different colors. The peaks of the spectra for the pixels belonging to regular structures correspond to the dominant orientation and granularity of the planting pattern whereas the spectra for irregular neighborhoods do not indicate any significant peaks.

c) Randomly select a pixel (r'_k, c'_k) from $\mathcal{L}_k$, and add it to the region $\mathcal{R}_k$ if the distance

$$\frac{1}{|\mathcal{G}||\Theta|} \sum_{g \in \mathcal{G}} \sum_{\theta \in \Theta} |\rho(\mathcal{R}_k; g, \theta) - \rho(r'_k, c'_k; g, \theta)| \quad (8)$$

is less than a threshold τ_d. The $|\cdot|$ operation on sets denotes cardinality, whereas it corresponds to absolute value for real numbers inside the summation.

d) When a new pixel (r'_k, c'_k) is added to $\mathcal{R}_k$, update the regularity score for $\mathcal{R}_k$ as

$$\rho(\mathcal{R}_k; g, \theta) = \frac{|\mathcal{R}_k|\rho(\mathcal{R}_k; g, \theta) + \rho(r'_k, c'_k; g, \theta)}{|\mathcal{R}_k| + 1} \quad (9)$$

and extend $\mathcal{L}_k$ by adding the previously unprocessed neighbors of (r'_k, c'_k) that have a maximum regularity score greater than τ_l.

e) Repeat from 3c until $\mathcal{L}_k$ becomes empty.

4) Repeat from 3 until $\mathcal{L}_0$ becomes empty.

5) Construct a graph where the nodes are the resulting regions, $\mathcal{R}_k$, and neighboring regions are connected with undirected edges. Assign a weight to each edge using the distance

$$\frac{1}{|\mathcal{G}||\Theta|} \sum_{g \in \mathcal{G}} \sum_{\theta \in \Theta} |\rho(\mathcal{R}_k; g, \theta) - \rho(\mathcal{R}_t; g, \theta)| \quad (10)$$

where $\mathcal{R}_k$ and $\mathcal{R}_t$ are the regions connected by that edge.

6) Find the edge with the smallest weight in the graph. If that weight is less than the threshold τ_d, merge the corresponding regions $\mathcal{R}_k$ and $\mathcal{R}_t$ as new $\mathcal{R}_k$, update the regularity score for $\mathcal{R}_k$ as

$$\rho(\mathcal{R}_k; g, \theta) = \frac{|\mathcal{R}_k|\rho(\mathcal{R}_k; g, \theta) + |\mathcal{R}_t|\rho(\mathcal{R}_t; g, \theta)}{|\mathcal{R}_k| + |\mathcal{R}_t|} \quad (11)$$

and update the edges that are incident to these regions.

7) Repeat from 6 until no edge with a weight smaller than τ_d remains.

8) Eliminate small regions and relabel all remaining regions consecutively.

Steps 2)–4) correspond to a growing process involving merging pixels to existing regions, and steps 5)–7) correspond to a second growing process involving merging regions to obtain the final segmentation. The growing and merging steps are shown in Fig. 6. Note that, since $\rho(r, c; g, \theta) \in [0, 1]$, all thresholds τ_h, τ_l, and τ_d are in the [0, 1] range. Guidelines for selecting these thresholds are given in Section V.

The pixels are randomly selected from the list of candidates for growing in step 3c) to avoid any direction bias in the selection order while significantly reducing the computational load when compared to the alternative where all candidates are considered at every iteration. Once the iterations complete when all candidate lists become empty and no further merging becomes possible, a granularity and orientation estimate for each region can be selected as

$$\{g^*(\mathcal{R}_k), \theta^*(\mathcal{R}_k)\} = \arg \max_{g, \theta} \rho(\mathcal{R}_k; g, \theta). \quad (12)$$

The estimates $g^*(\mathcal{R}_k)$ and $\theta^*(\mathcal{R}_k)$ indicate the dominant granularity and orientation, respectively, of the planting pattern inside each region whose boundary is delineated during the region growing process. These estimates can be used for further selection and processing of the regions of interest in planning, mapping, classification, and monitoring applications.

V. PERFORMANCE EVALUATION

The performances of the proposed orchard detection and segmentation algorithms were evaluated with respect to different parameter settings using the data sets described in Section II. Since the algorithms are fully unsupervised, i.e., no training is required, the detection and segmentation results can be computed once the parameters are set. All parameters except the
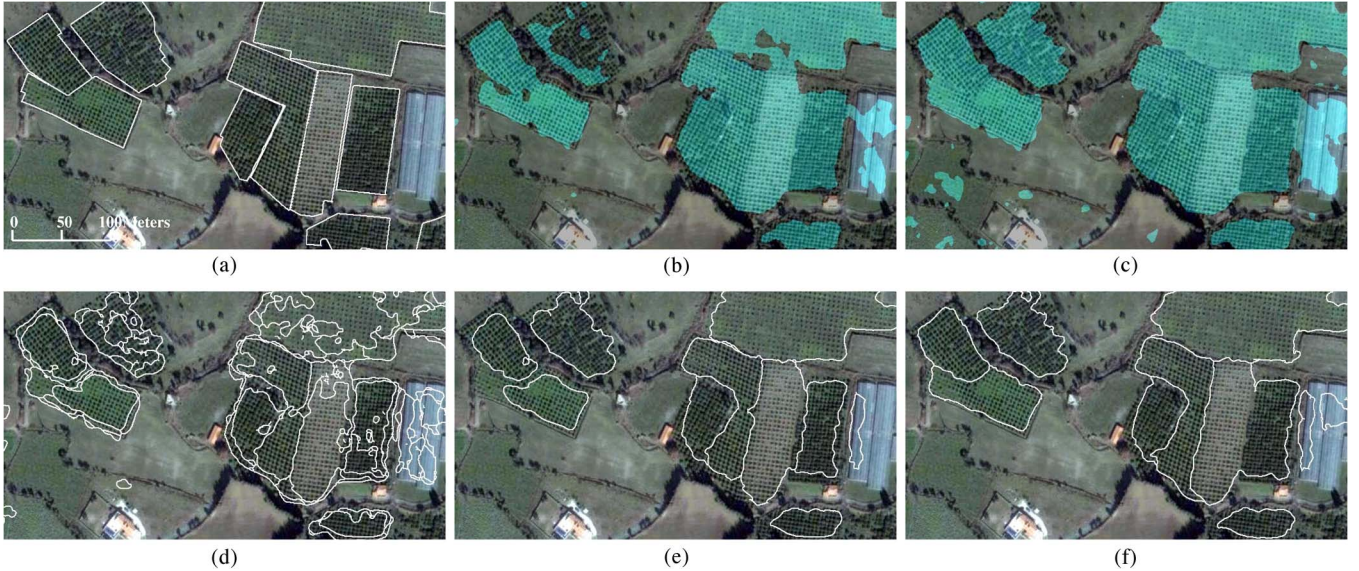
Fig. 6. Illustration of the segmentation process for the example image in Fig. 5. Boundaries of the regions in the reference data are shown in (a). Map of pixels that have regularity scores above $\tau_h = 0.85$ and are used as seeds ($\mathcal{L}_0$) for region growing are shown as green in (b). Map of pixels that have regularity scores above $\tau_l = 0.80$ and are considered as candidates ($\mathcal{L}_k$) for region growing are shown as green in (c). Boundaries of the regions at the end of the growing process (step 4) for $\tau_d = 0.05$ are shown in (d). Boundaries of the regions at the end of the merging process (step 8) for $\tau_d = 0.05$ are shown in (e). For comparison, results of merging for $\tau_d = 0.06$ are shown in (f).

segmentation thresholds can easily be assigned intuitive values according to the orchards of interest. We used pixel-based performance measures to evaluate the accuracy of orchard detection, and object-based performance measures to evaluate the accuracy of orchard segmentation. The reference data for each image were produced using manual photo-interpretation. Objective performance criteria were computed to evaluate both site-specific and cross-landscape performance.

A. Evaluation of Orchard Detection

We used the same set of values for all parameters for all data sets to evaluate the effectiveness of the proposed algorithms for sites with different characteristics. In particular, the set of granularities $\mathcal{G} = \{g_1, \ldots, g_{N_g}\}$, corresponding to the expected tree sizes in pixels in the multiscale pyramid, was fixed as $\{2, 2\sqrt{2}, 4, 4\sqrt{2}, 8, 8\sqrt{2}\}$ using (4) with parameters $g_{\min} = 2$, $g_{\max} = 12$, and $N_g = 6$. The window height δ in (6) for computing the projection profiles was fixed at 7 pixels. If the window height is further increased, only the orchards occupying larger areas can be found. However, it is also not desirable to decrease the window height too much because the projection is no longer effective when it includes partial rows of trees. Therefore, as a tradeoff, the window height of 7 pixels was used as approximately twice the expected size of the trees corresponding to the granule size (cross section) of 3 pixels for the spot filter that was used to process all granularities. This fixed window height was in fact implicitly adaptive to the granularities of interest because of the pyramid scheme used.

Orchard detection was evaluated at the second step of the algorithm in Section IV. In addition to the fixed parameters described above, three different Gaussian smoothing filters with sizes 11×11, 21×21, and 31×31 pixels were used for suppressing the noisy values in $\rho(r, c; g, \theta)$. The standard deviation of the Gaussian used for each filter was set to one fourth of the filter width. The threshold τ_h was varied from

TABLE I
PARAMETER SETTINGS USED FOR ORCHARD DETECTION EVALUATION

	Values
Set of granularities $\mathcal{G}$ (pixels)	$\{2, 2\sqrt{2}, 4, 4\sqrt{2}, 8, 8\sqrt{2}\}$
Window height δ (pixels)	7
Smoothing filter size (pixels)	11×11, 21×21, 31×31
Regularity score threshold τ_h	0.60 to 0.95 with increments of 0.01
Angle increment θ_{inc} (degrees)	$1°$, $5°$, $10°$

0.60 to 0.95 with increments of 0.01 to convert the regularity scores $\rho^*(r, c)$ to a binary map of orchards. We also designed an optional parameter to decrease the number of orientations used in the multi-orientation analysis step to reduce the complexity of the regularity score computation. As an alternative to using all 180 orientations from $-90°$ to $89°$, this range was sampled using an angle increment θ_{inc} of $1°$, $5°$, and $10°$, corresponding to 180, 36, and 18 orientations, respectively. The parameter settings used for orchard detection evaluation are shown in Table I. These settings corresponded to 324 different parameter combinations for each data set.

A particular choice of smoothing filter size, angle increment, and regularity score threshold produces a binary map of regular areas detected as orchards. Precision and recall were used as the quantitative performance criteria as in [3]. Given the reference data that were obtained by manual labeling of the orchard areas as positive and the rest of the image as negative, pixel-based precision and recall were computed as

$$\text{precision} = \frac{\text{\# of correctly detected pixels}}{\text{\# of all detected pixels}} \qquad (13)$$

$$\text{recall} = \frac{\text{\# of correctly detected pixels}}{\text{\# of all pixels in the reference data}}. \qquad (14)$$

Recall (producer's accuracy) can be interpreted as the number of true positives detected by the algorithm, while precision (user's accuracy) evaluates the tendency of the algorithm for false positives. Finally, the F_β measure [22] that provides a
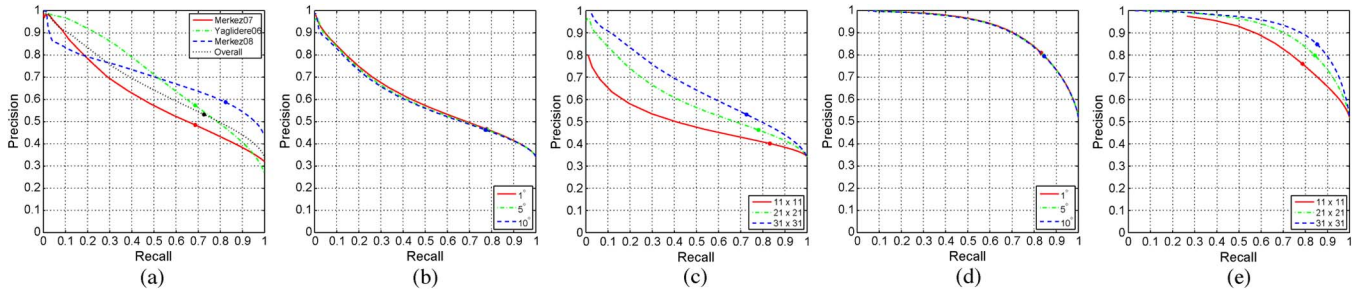
Fig. 7. Precision versus recall curves for different settings used for orchard detection evaluation. (a) shows the site-specific performance (by combining the images specific to each site) and the cross-landscape performance (by combining all images) for the Giresun data set by varying the regularity score threshold τ_h for fixed smoothing filter size of 31×31 and angle increment $\theta_{inc} = 5°$. (b) and (d) show the performances for the whole Giresun and Izmir data sets, respectively, for different values of θ_{inc} at fixed smoothing filter size of 21×21. (c) and (e) show the performances for the whole Giresun and Izmir data sets, respectively, for different values of smoothing filter size at fixed $\theta_{inc} = 5°$. The best F_1 value is marked on each curve.

way of combining precision and recall into a single measure that falls between the two was computed as

$$F_\beta = \frac{(\beta^2 + 1) \times \text{precision} \times \text{recall}}{\beta^2 \times \text{precision} + \text{recall}}. \quad (15)$$

The F_β measure attaches β times as much importance to recall as precision. The F_1 measure $(\beta = 1)$ was used in the experiments below to rank the performances of different parameter settings.

Fig. 7 shows precision versus recall curves for different settings. Table II summarizes the parameter settings that obtained the best detection performance among all combinations. When all parameter combinations were considered, the following conclusions can be derived.

- We observed that more accurate detections were obtained for QuickBird images (*Yaglidere06* and *Merkez08*) compared to the Ikonos image (*Merkez07*) because the individual trees that made up the texture were more visible in the increased spatial resolution. We also observed that the time of the image acquisition affected the results as higher accuracy was obtained when the individual trees were more apparent in the panchromatic image. The results for Izmir data taken from Google Earth were more accurate than those for the Giresun data because the texture patterns were more distinctive in the former. The proposed algorithm was successful in detecting these patterns even under the lossy compression of the JPEG format of the input images. The lower accuracy for the Giresun data was mainly due to the irregularities in the planting patterns, mixed appearances of other types of trees within the orchards, and the deformations in the visual appearance of the patterns due to the strong relief in the region.
- The best parameter settings were very similar for individual sites as well as for the whole data set. The best performing Gaussian smoothing filter of 31×31 pixels realized better noise suppression compared to the smaller filters and achieved higher precision and recall for all data sets. The best performing angle increment was obtained at $5°$, but the performances of $1°$ and $10°$ were also very similar. This left the regularity score threshold as the only significant parameter in the algorithm. The minor differences in the best threshold values among different data sets were due to differences in the spatial resolution, time of image acquisition, and texture content. Site-specific adjustment

TABLE II
PARAMETER SETTINGS THAT OBTAINED THE BEST PERFORMANCE IN ORCHARD DETECTION. THE PARAMETERS AND THE PERFORMANCE MEASURES ARE DEFINED IN THE TEXT. THE DATA COLUMN SHOWS THE DATA SET USED FOR COMPUTING THE AVERAGE PERFORMANCE MEASURES IN EACH ROW

Data	Filter	τ_h	θ_{inc}	Precision	Recall	F_1
Merkez07	31×31	0.73	$10°$	0.4763	0.7069	0.5691
Yaglidere06	31×31	0.75	$5°$	0.5732	0.6864	0.6247
Merkez08	31×31	0.70	$5°$	0.5875	0.8243	0.6860
Overall	31×31	0.73	$5°$	0.5322	0.7262	0.6142
Izmir	31×31	0.77	$5°$	0.8483	0.8531	0.8507

of the threshold leads to higher precision and recall for individual images compared to the average values in Fig. 7 and Table II. Such thresholds can be set by using automatic thresholding techniques for local analysis. Overall, the similar performances for different parameter values were possible because the proposed algorithm exploited the regularity in the structure in the projection profiles using the periodicity analysis in a way that was invariant to contrast, scale, and orientation differences in the raw image data.

- When the overall detections were considered, the following sources of error were identified. Most of the false positives were observed along roads where there was a repetitive contrast difference on both sides, and around some greenhouses where a similar regular contrast difference was observed due to parallel edges. These false positives may be eliminated by using a threshold on vegetation using multispectral information (e.g., thresholding of NDVI). False negatives mostly occurred at small vegetation patches that were marked in the reference data due to a few rows of regularly planted trees but were not large enough for the algorithm in order to observe a periodic signal with the selected parameter settings. The parameters can be locally adjusted if a minimum mapping unit is required for a particular site.
- We also compared the results to those of the algorithm in [18] on the Giresun data set. The average F_1 measure was obtained as 0.5186, 0.6099, 0.6835, and 0.5874 for *Merkez07*, *Yaglidere06*, *Merkez08*, and the whole data, respectively, when the algorithm in [18] was applied. The improvements achieved in this paper were mainly due to pyramid-based handling of the granularities and the direct computation of the regularity scores without any need for upsampling.
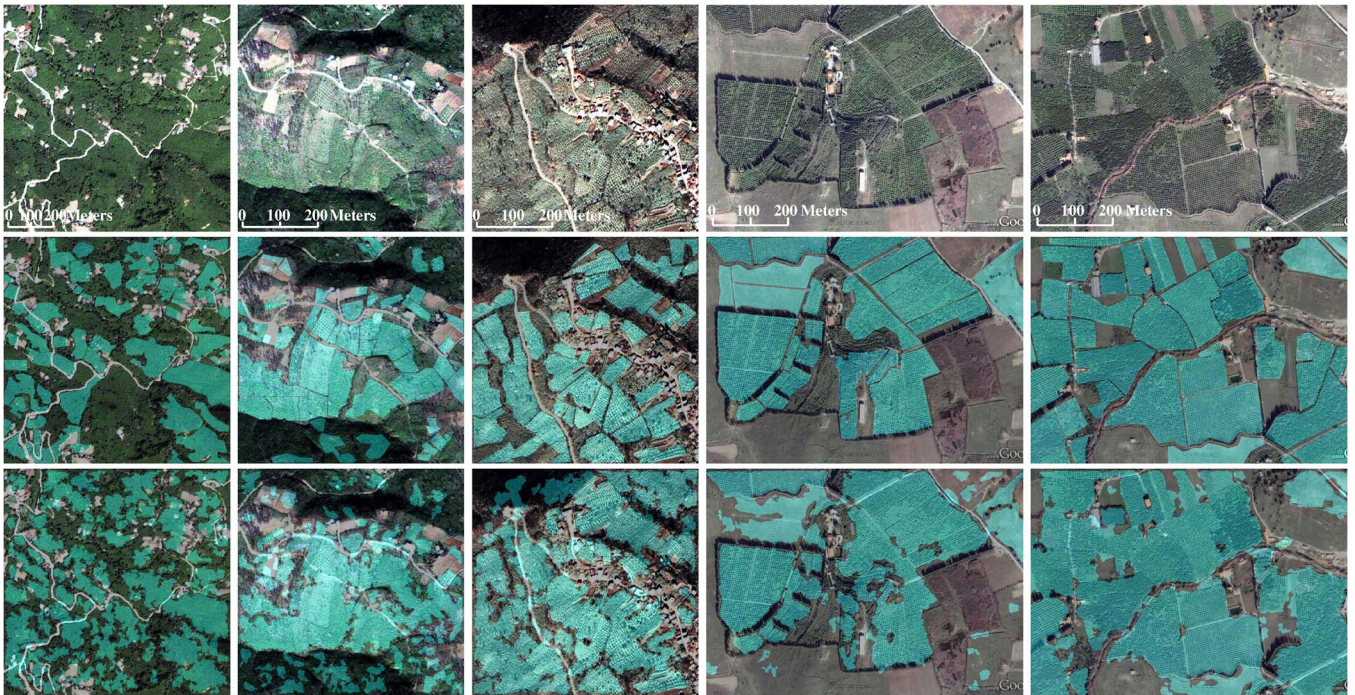
Fig. 8. Example results for orchard detection. The first three columns show examples from the Giresun data set (*Merkez07*, *Yaglidere06*, and *Merkez08*, respectively, each with 1000 × 1000 pixels), and the last two columns show examples from the Izmir data set (each with 1400 × 1000 pixels). The first row shows the original images. The second row shows the reference data where the orchard areas are overlayed as green. The third row shows the areas detected by thresholding the regularity scores using the τ_h values given in Table II as green.
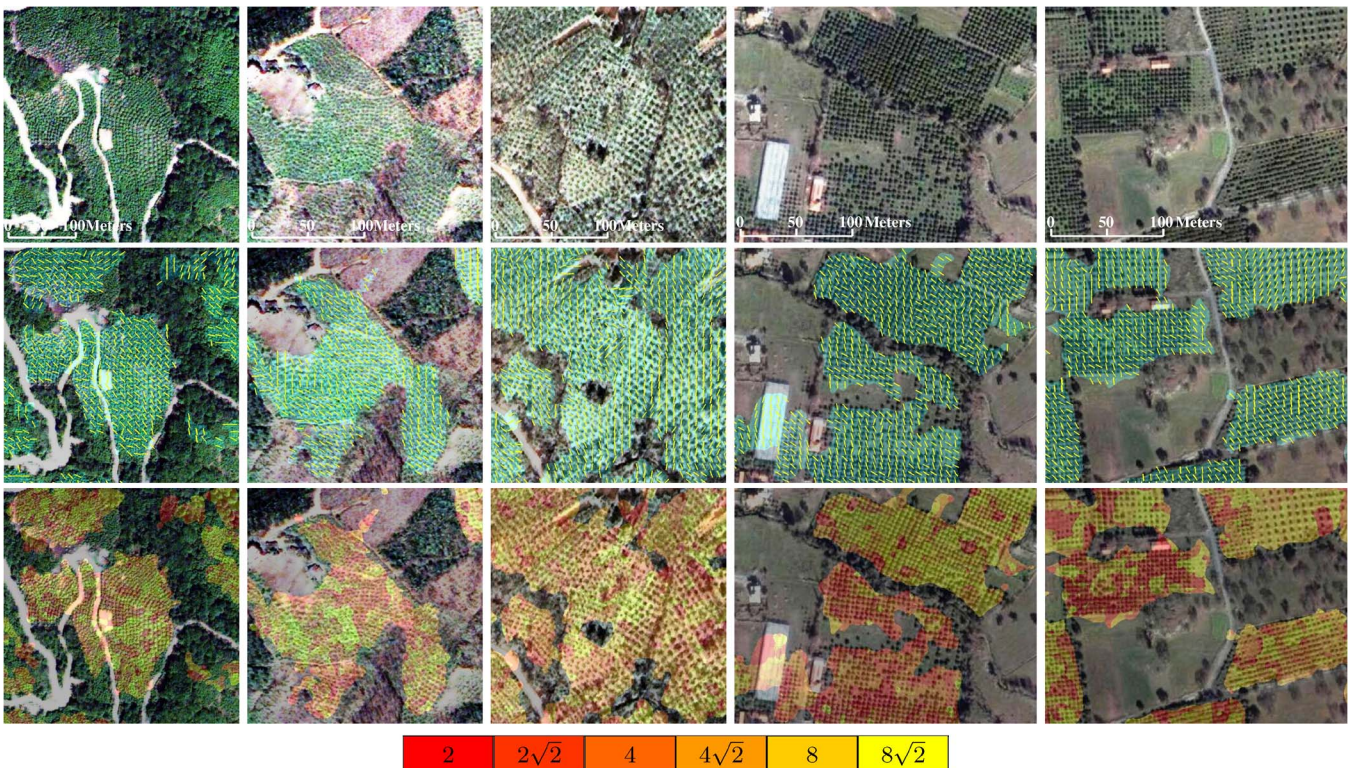

Fig. 9. Examples of local details of orchard detection. The first three columns show examples from the Giresun data set (*Merkez07*, *Yaglidere06*, and *Merkez08*, respectively, each with 350 × 350 pixels), and the last two columns show examples from the Izmir data set (each with 450 × 350 pixels). The first row shows the original images. The second row shows the orientation estimates as yellow line segments overlayed on the areas detected by thresholding the regularity scores. The third row shows the granularity estimates using the color map shown on the fourth row. The orientation and granularity estimates at each pixel are computed as the θ and g values, respectively, for which the regularity score $\rho(r, c; g, \theta)$ is maximized in (7). Note that, in highly structured areas, similarly high regularity scores can be obtained at 90° or even 45° rotated projects. The orientation estimates for pixels in such areas may alternate between these principal orientations and may yield a noisy picture when visualized. However, the segmentation algorithm uses the whole regularity spectrum, so this is only a visualization issue.

TABLE III
PARAMETER SETTINGS USED FOR ORCHARD
SEGMENTATION EVALUATION

	Values
Segmentation threshold τ_h	0.81 to 0.90 with increments of 0.01
Segmentation threshold τ_l	0.70 to 0.77 with increments of 0.01
Segmentation threshold τ_d	0.05 to 0.07 with increments of 0.005

Example detections are shown in Fig. 8. The quantitative evaluation does not currently reflect the quality of the results very precisely because the reference data remain approximate. The set of parameters that were selected by maximizing the average F_1 measure for a group of images with such approximate reference data may not always correspond to the best results. Therefore, we examined the local details to evaluate the effectiveness of the proposed algorithm in identifying regular plantation areas as well as the accuracy of the granularity and orientation estimates. Fig. 9 shows the details of these estimates. We observed that orientation estimation was more accurate than granularity estimation. These examples showed that even the gradually changing orientations could be estimated smoothly, and the localization of the regular plantation areas was very accurate even when no sharp boundaries existed in the image data. Furthermore, the individual trees that were located inside the orchards but had types different from those belonging to the orchards and the locations of missing crowns were successfully isolated in the detection results.

B. Evaluation of Orchard Segmentation

Orchard segmentation was evaluated using the whole algorithm in Section IV. Different values were considered for the segmentation thresholds. In particular, the threshold τ_h was varied from 0.81 to 0.90 with increments of 0.01, the threshold τ_l was varied from 0.70 to 0.77 with increments of 0.01, and the threshold τ_d was varied from 0.05 to 0.07 with increments of 0.005. The parameter settings used for orchard segmentation evaluation are shown in Table III. These settings corresponded to 400 different parameter combinations for each data set.

Precision and recall can also be used to evaluate the segmentation performance. Most of the segmentation evaluation measures in the literature are based on matches between two complete partitionings of the whole image, and are not directly applicable to the problem studied in this paper where the goal is to delineate particular objects, not to partition the whole land cover. We used an object-based evaluation procedure similar to the ones in [3] and [23] that was adapted from the work of [24] on range image segmentation evaluation. This procedure used the individual reference orchards in the reference data and the output orchards in the produced segmentation map, and classified every pair of reference and output objects as correct detections, overdetections, underdetections, missed detections, or false alarms with respect to a threshold T on the amount of overlap between these objects. The overlap was computed in terms of number of pixels. A pair of reference and output objects was classified as an instance of correct detection if at least T percent of each object overlapped with the other. A reference object and a set of output objects were classified as an instance of overdetection if at least T percent of each output object overlapped with the reference object and at least T percent of the reference object overlapped with the union of the output objects. An output object and a set of reference objects were

TABLE IV
PARAMETER SETTINGS THAT OBTAINED THE BEST PERFORMANCE IN ORCHARD SEGMENTATION. THE PARAMETERS AND THE PERFORMANCE MEASURES ARE DEFINED IN THE TEXT. THE DATA COLUMN SHOWS THE DATA SET USED FOR COMPUTING THE AVERAGE PERFORMANCE MEASURES IN EACH ROW

Data	τ_h	τ_l	τ_d	Precision	Recall	F_1
Merkez07	0.86	0.72	0.065	0.5867	0.4226	0.4913
Yaglidere06	0.88	0.72	0.070	0.6967	0.5443	0.6111
Merkez08	0.85	0.70	0.070	0.6533	0.5791	0.6140
Overall	0.87	0.72	0.065	0.5255	0.3570	0.4252
Izmir	0.89	0.75	0.070	0.5835	0.8132	0.6795

classified as an instance of underdetection if at least T percent of each reference object overlapped with the output object and at least T percent of the output object overlapped with the union of the reference objects. A reference object that was not in any instance of correct detection, overdetection, and underdetection was classified as missed detection. An output object that was not in any instance of correct detection, overdetection, and underdetection was classified as false alarm. For a T value between 0.5 and 1.0, any object can contribute to at most three classifications (at most one correct detection, one overdetection and one underdetection) [23], [24]. If an object was included only in a single classification instance, that instance was used as its unique classification. When an object participated in two or three classification instances, the instance with the largest overlap was selected for that object. An overlap threshold of $T = 0.6$ was used in the experiments in this paper.

Given the reference data that were obtained by manual delineation of individual orchards and once all reference and output objects were classified into instances of correct detections, overdetections, underdetections, missed detections, or false alarms, precision and recall were computed as

$$
\begin{aligned}
\text{precision} &= \frac{\text{\# of correctly detected objects}}{\text{\# of all detected objects}} \\
&= \frac{N - FA}{N}
\end{aligned} \tag{16}
$$

$$
\begin{aligned}
\text{recall} &= \frac{\text{\# of correctly detected objects}}{\text{\# of all objects in the reference data}} \\
&= \frac{M - MD}{M}
\end{aligned} \tag{17}
$$

where FA and MD were the number of false alarms and missed detections, respectively, and N and M were the number of objects in the output and reference maps, respectively. Then, F_β was computed as in (15) to combine precision and recall into a single measure.

Table IV summarizes the parameter settings that obtained the best segmentation performance among all combinations. When all parameter combinations were considered, the following conclusions can be derived.

- The pattern in the relative accuracies of different data sets was similar to the pattern in the detection results. In particular, more accurate segmentations were obtained for QuickBird images compared to the Ikonos image, and the most accurate segmentations were obtained for the Google images due to similar reasons discussed in detection evaluation.
- The effects of changing the parameter values were similar for all data sets. In general, increasing τ_h increased
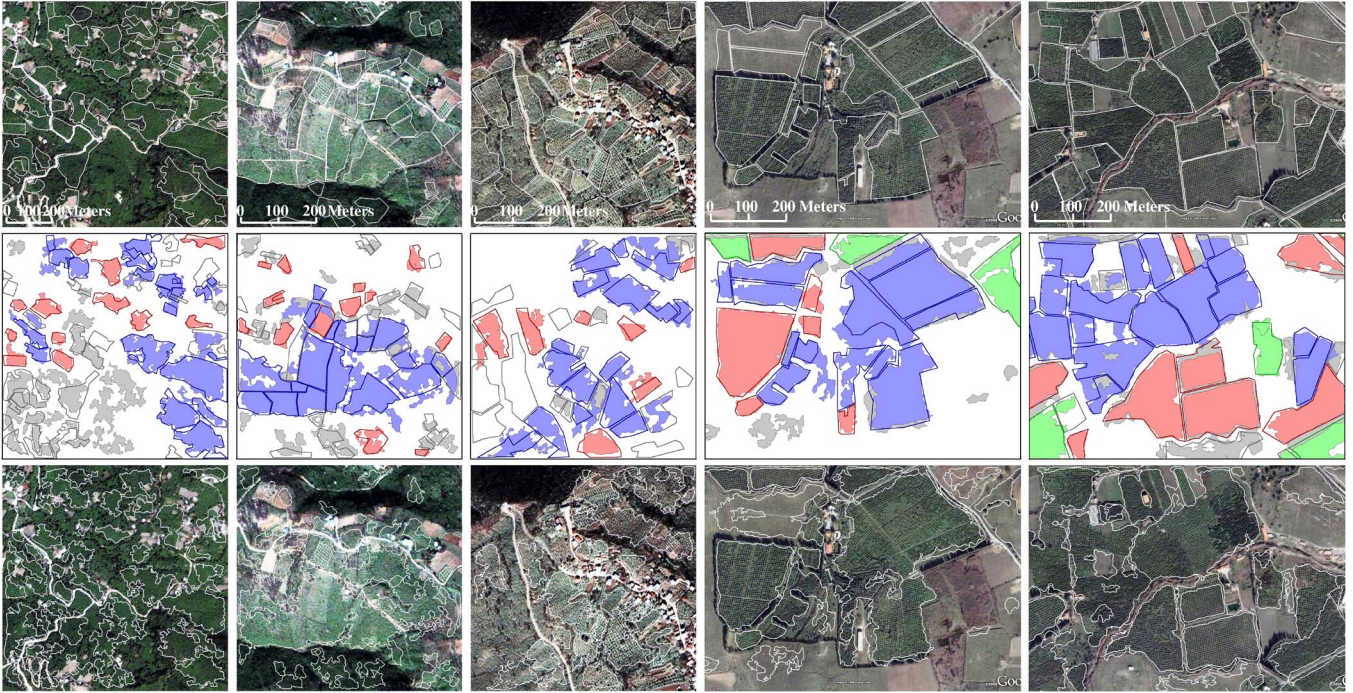
Fig. 10. Example results for orchard segmentation. The first three columns show examples from the Giresun data set (*Merkez07*, *Yaglidere06*, and *Merkez08*, respectively, each with 1000×1000 pixels), and the last two columns show examples from the Izmir data set (each with 1400×1000 pixels). The first row shows the reference boundaries. The second row shows the segment classification according to the object-based performance measures. The algorithm outputs that correspond to correct detections (red), overdetections (green), underdetections (blue), and false alarms (gray) are shown as regions. The reference objects that are correctly detected (red), overdetected (green), underdetected (blue), and missed (gray) are shown as overlayed boundaries. The third row shows the detected objects overlayed on the visible bands.

precision while decreasing recall. For small values of τ_d, an increase also resulted in increased precision and recall, but after some point, increasing it further started decreasing both precision and recall because too many regions started merging. The effect of τ_l depended on τ_d. In general, increasing τ_l decreased both precision and recall, but for larger τ_d, increasing τ_l led to an increase in both precision and recall.

- The best performing τ_h value was significantly higher than the best τ_h for detection. This was expected because the seed points in the growing process were selected within the most regular areas.
- The best performing τ_l value was close to but slightly lower than the best τ_h for detection. This was also expected because the algorithm stopped growing at the boundaries of regular areas.
- τ_d was the most sensitive parameter. Even though a relatively large range of τ_d values worked similarly well in finding the boundaries of neighboring orchards planted at different granularities and orientations, small variations in τ_d led to overdetections or underdetections for neighboring orchards with very similar granularity and orientation patterns. This effect was the most apparent in *Merkez08* images as the accuracy decreased by approximately 20% when the parameters with the best average performance in the whole Giresun data were used (fourth row in Table IV) instead of the parameters with the best average performance in *Merkez08* images (third row in Table IV). This was the main reason in the decrease in overall accuracy when the average performance for the whole Giresun data was considered in parameter selection compared to the

site-specific performances. Local adjustment of τ_d could lead to more accurate boundaries, particularly when the granularity and orientation differences between neighboring orchards were subtle. However, the values presented in Table IV were selected based on average performance for all images in a particular data set.

Example segmentations are shown in Fig. 10. The approximations in the manually created reference data affected the results more negatively compared to the detection evaluation. For example, some of the orchard polygons in the reference data for Giresun were extensions to the point data collected by a team from the Ministry of Agriculture and Rural Affairs of Turkey during a field trip. However, some parcels contained multiple orchards having different characteristics whereas some neighboring parcels contained orchards having very similar granularity and orientation patterns. Since the orchard segmentation algorithm in this paper was based on local differences in the planting patterns, the segmentation boundaries obtained did not always correspond to such parcel boundaries. Similarly, gradual changes in the orientation of the planting pattern were often merged into the same output region because the growing process also gradually adjusted the orientation estimates. Furthermore, some reference polygons included several isolated trees that did not belong to the orchard, but the algorithm often successfully left out those trees outside the segmentations. Consequently, such approximations in the reference data resulted in noisy labeling of some of the produced segmentations as overdetection, underdetection, missed detection, and false alarm. Therefore, we also examined the local details to evaluate the effectiveness of the proposed algorithm in delineating the orchard boundaries. Fig. 11 show some example

Fig. 11. Examples of local details of orchard segmentation. The first row shows examples from the Giresun data set (each with 350×350 pixels). The first and second images are examples from *Merkez07*, the third and fourth are examples from *Yaglidere06*, and the fifth is an example from *Merkez08*. The second row shows examples from the Izmir data set (each with 450×350 pixels). The detected objects are overlayed in pseudocolor on the visible bands.

details. These examples showed that the proposed fully unsupervised algorithm that used only texture information computed from panchromatic data could produce accurate segmentation boundaries.

C. Computational Complexity

The proposed detection and segmentation algorithms were implemented in Matlab with the only exception that the periodic signal analysis step described in Section III-C was implemented in C. The overall processing of one 1000×1000 pixel Giresun image took 19 min on the average using the unoptimized Matlab/C code on a personal computer with a 2-GHz Intel Xeon processor. The running times were computed when the smoothing filter size was set to 31×31 pixels, the angle increment θ_{inc} was set to $1°$, and the rest of the parameters were fixed at the values given in Tables I and IV.

We performed a code profile analysis to investigate the time spent in different steps. Of the overall 19 min, on the average, preprocessing using multi-granularity spot filters in Section III-A took 0.2% of the time (2 s), image rotations for multi-orientation analysis in Section III-D took 3.6% of the time (0.7 min), periodic signal analysis to compute the regularity scores for all pixels for all granularities and orientations in Section III-C took 5% of the time (1 min), upsampling the resulting scores back to the original resolution to compute $\rho(r, c; g, \theta)$ in Section III-D took 9% of the time (1.7 min), smoothing the regularity scores in $\rho(r, c; g, \theta)$ before segmentation in Section IV took 76% of the time (14.5 min), and segmentation in Section IV took 6.2% of the time (1.2 min).

We also analyzed how different parameters affect the running times. The time complexity is linear in the angle increment θ_{inc}. Increasing θ_{inc} from $1°$ to $5°$ decreased the running time from 19 min to 3.9 min, and increasing it to $10°$ decreased the running time to 2 min. These significant savings in running time did not lead to any noticeable decrease in accuracy as shown in Fig. 7. All main steps, namely multi-granularity and multi-orientation computation of the regularity scores, smoothing of the scores, and segmentation benefited from the increase in θ_{inc} because of the reduction in the number of orientations to be

processed. Similarly, reducing the smoothing filter size from 31×31 pixels to 21×21 pixels decreased the running time from 19 min to 14.5 min, from 3.9 min to 3 min, and from 2 min to 1.6 min for θ_{inc} at $1°$, $5°$, and $10°$, respectively. The main cause of the decrease in running time was the shorter convolution time during the smoothing of the regularity scores in $\rho(r, c; g, \theta)$ before segmentation. The other steps were not affected because the number of granularities and orientations remained the same. However, decreasing the smoothing filter size led to a loss of accuracy as shown in Fig. 7 because more smoothing resulted in more stable regularity scores.

Computing the maximum regularity score for each pixel at step 1 was the dominant operation with 30% of the time during segmentation. After these scores were computed, the running time of segmentation depended on the segmentation parameters τ_h, τ_l, and τ_d. In general, decreasing τ_h increased the running time as the number of seed pixels increased. Similarly, decreasing τ_l increased the running time because the number of candidate pixels for growing increased. The running time analysis showed that the proposed algorithm provided flexibility for possible adjustment of the parameters by the users for different tradeoffs between computation time, detection, and segmentation accuracy.

VI. CONCLUSION

Development of flexible automatic methods for object detection in agricultural landscapes continues to be an important research problem when the analysis goes beyond local sites to cover a wide range of landscapes. In this paper, we presented a new unsupervised method for the detection and segmentation of orchards. The method used a structural texture model that was based on a near-regular repetitive arrangement of texture primitives. The model applied to orchard detection used trees at different granularities as the texture primitives and the planting patterns at different orientations as their arrangement. Multi-granularity isotropic filters were used for the enhancement of potential tree locations in panchromatic data without any strict requirement for the detection of individual trees. The method quantified the regularity of the planting patterns in terms of the

periodicity of the projection profiles computed from the filter responses within oriented sliding windows. This resulted in a regularity score at each pixel for each granularity and each orientation. Then, a segmentation process iteratively merged neighboring pixels and regions belonging to similar planting patterns according to the similarities of their regularity scores, and obtained the boundaries of individual orchards along with estimates of their granularities and orientations.

Extensive experiments using Ikonos and QuickBird imagery as well as images taken from Google Earth showed that the proposed algorithm could accurately localize and delineate orchards in complex images with different spatial resolutions and different characteristics. Examination of local details showed that the boundaries of neighboring orchards were accurately detected even when the granularity and orientation differences between the orchards were subtle and no sharp boundaries existed in the image data. Furthermore, variations in the granularities of the trees within the same orchard, their imperfect arrangements, missing crowns, and the restrictions of the terrain on the planting pattern were successfully handled. The results can be further improved by incorporating a postprocessing step that involves the utilization of multispectral data for eliminating some false positives caused by regular structures that are not orchards and narrow paths within groups of orchards having similar orientations.

REFERENCES

[1] A. Baraldi, T. Wassenaar, and S. Kay, "Operational performance of an automatic preliminary spectral rule-based decision-tree classifier of space-borne very high resolution optical images," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 9, pp. 3482–3502, Sep. 2010.

[2] G. G. Wilkinson, "Results and implications of a study of fifteen years of satellite image classification experiments," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 3, pp. 433–440, Mar. 2005.

[3] S. Aksoy, H. G. Akcay, and T. Wassenaar, "Automatic mapping of linear woody vegetation features in agricultural landscapes using very high-resolution imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 1, pp. 511–522, Jan. 2010.

[4] F. Kayitakire, C. Hamel, and P. Defourny, "Retrieving forest structure variables based on image texture analysis and IKONOS-2 imagery," *Remote Sens. Environ.*, vol. 102, no. 3/4, pp. 390–401, Jun. 2006.

[5] I. Epifanio and P. Soille, "Morphological texture features for unsupervised and supervised segmentations of natural landscapes," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 4, pp. 1074–1083, Apr. 2007.

[6] T. A. Warner and K. Steinmaus, "Spatial classification of orchards and vineyards with high spatial resolution panchromatic imagery," *Photogramm. Eng. Remote Sens.*, vol. 71, no. 2, pp. 179–187, Feb. 2005.

[7] R. Trias-Sanz, "Texture orientation and period estimator for discriminating between forests, orchards, vineyards, and tilled fields," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 10, pp. 2755–2760, Oct. 2006.

[8] N. Amoruso, A. Baraldi, C. Tarantino, and P. Blonda, "Spectral rules and geostatic features for characterizing olive groves in quickbird images," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Cape Town, South Africa, Jul. 13–17, 2009, vol. 4, pp. IV-228–IV-231.

[9] L. A. Ruiz, J. A. Recio, and T. Hermosilla, "Methods for automatic extraction of regularity patterns and its application to object-oriented image classification," in *Proc. Photogramm. Image Anal.*, Munich, Germany, Sep. 19–21, 2007, pp. 117–121.

[10] K. Tasdemir, "Classification of hazelnut orchards by self-organizing maps," in *Proc. 6th IAPR Workshop Pattern Recognit. Remote Sens.*, Istanbul, Turkey, Aug. 22, 2010, pp. 1–4.

[11] P. Helmholz and F. Rottensteiner, "Automatic verification of agricultural areas using ikonos satellite images," in *Proc. ISPRS Workshop High-Resolution Earth Imag. Geospatial Inform.*, Hannover, Germany, Jun. 2–5, 2009.

[12] T. Ranchin, B. Naert, M. Albuisson, G. Boyer, and P. Astrand, "An automatic method for vine detection in airborne imagery using wavelet transform and multiresolution analysis," *Photogramm. Eng. Remote Sens.*, vol. 67, no. 1, pp. 91–98, Jan. 2001.

[13] T. Wassenaar, J.-M. Robbez-Masson, P. Andrieux, and F. Baret, "Vineyard identification and description of spatial crop structure by per-field frequency analysis," *Int. J. Remote Sens.*, vol. 23, no. 17, pp. 3311–3325, Sep. 2002.

[14] J. Chanussot, P. Bas, and L. Bombrun, "Airborne remote sensing of vineyards for the detection of dead vine trees," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Seoul, Korea, Jul. 25–29, 2005, vol. 5, pp. 3090–3093.

[15] C. Delenne, S. Durrieu, G. Rabatel, M. Deshayes, J. S. Bailly, C. Lelong, and P. Coureton, "Textural approaches for vineyard detection and characterization using very high spatial resolution remote sensing data," *Int. J. Remote Sens.*, vol. 29, no. 4, pp. 1153–1167, Feb. 2008.

[16] G. Rabatel, C. Delenne, and M. Deshayes, "A non-supervised approach using gabor filters for vine-plot detection in aerial images," *Comput. Electron. Agric.*, vol. 62, no. 2, pp. 159–168, Jul. 2008.

[17] C. Delenne, S. Durrieu, G. Rabatel, and M. Deshayes, "From pixel to vine parcel: A complete methodology for vineyard delineation and characterization using remote-sensing data," *Comput. Electron. Agric.*, vol. 70, no. 1, pp. 78–83, Jan. 2010.

[18] I. Z. Yalniz and S. Aksoy, "Unsupervised detection and localization of structural textures using projection profiles," *Pattern Recognit.*, vol. 43, no. 10, pp. 3324–3337, Oct. 2010.

[19] R. C. Gonzales and R. E. Woods, *Digital Image Processing*, 3rd ed. Englewood Cliffs, NJ: Prentice-Hall, 2008.

[20] J. Malik and P. Perona, "Preattentive texture discrimination with early vision mechanisms," *J. Opt. Soc. Amer. A*, vol. 7, no. 5, pp. 923–932, May 1990.

[21] J. L. Crowley and A. C. Parker, "A representation for shape based on peaks and ridges in the difference of low-pass transform," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-6, no. 2, pp. 156–170, Mar. 1984.

[22] R. A. Baeza-Yates and B. Ribeiro-Neto, *Modern Information Retrieval*. Reading, MA: Addison-Wesley, 1999.

[23] B. Ozdemir, S. Aksoy, S. Eckert, M. Pesaresi, and D. Ehrlich, "Performance measures for object detection evaluation," *Pattern Recognit. Lett.*, vol. 31, no. 10, pp. 1128–1137, Jul. 15, 2010.

[24] A. Hoover, G. Jean-Baptiste, X. Jiang, P. J. Flynn, H. Bunke, D. B. Goldgof, K. Bowyer, D. W. Eggert, A. Fitzgibbon, and R. B. Fisher, "An experimental comparison of range image segmentation algorithms," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 18, no. 7, pp. 673–689, Jul. 1996.

Selim Aksoy (S'96–M'01–SM'11) received the B.S. degree from the Middle East Technical University, Ankara, Turkey, in 1996, and the M.S. and Ph.D. degrees from the University of Washington, Seattle, in 1998 and 2001, respectively.

He has been an Assistant Professor at the Department of Computer Engineering, Bilkent University, Ankara, since 2004, where he is also the Co-Director of the RETINA Vision and Learning Group. Before joining Bilkent, he was a Research Scientist at Insightful Corporation, Seattle, where he was involved in image understanding and data mining research sponsored by the National Aeronautics and Space Administration, the U.S. Army, and the National Institutes of Health. During 1996–2001, he was a Research Assistant at the University of Washington, where he developed algorithms for content-based image retrieval, statistical pattern recognition, object recognition, graph-theoretic clustering, user relevance feedback, and mathematical morphology. During the summers of 1998 and 1999, he was a Visiting Researcher at the Tampere International Center for Signal Processing, Tampere, Finland, collaborating in a content-based multimedia retrieval project. His research interests include computer vision, statistical and structural pattern recognition, machine learning, and data mining with applications to remote sensing, medical imaging, and multimedia data analysis.

Dr. Aksoy is a member of the IEEE Geoscience and Remote Sensing Society, the IEEE Computer Society, and the International Association for Pattern Recognition (IAPR). He was one of the Guest Editors of the special issues on Pattern Recognition in Remote Sensing of IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, *Pattern Recognition Letters*, and IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING in 2007, 2009, and 2012, respectively. He served as the Vice Chair of the IAPR Technical Committee 7 on Remote Sensing during 2004–2006, and as the Chair of the same committee during 2006–2010. He is an Associate Editor of Pattern Recognition Letters. He received the CAREER Award from the Scientific and Technological Research Council of Turkey (TUBITAK) in 2004.

Ismet Zeki Yalniz received the B.S. and M.S. degrees in computer engineering from Bilkent University, Ankara, Turkey, in 2008 and 2006, respectively. He is currently working toward the Ph.D. degree in the Department of Computer Science, University of Massachusetts at Amherst, Amherst.

He is broadly interested in combining computer vision and information retrieval concepts to offer practical solutions for data and/or computation intensive problems. On the computer vision side, he worked on texture analysis and segmentation, video event detection, document image analysis, and recognition. On the information retrieval side, he worked on the retrieval of noisy (OCRed) text documents focusing on OCR evaluation and error correction, duplicate document detection, and translation identification. His current research focus is on the development of effective and fast image search techniques.

Kadim Taşdemir (M'09) received the B.S. degree in electrical and electronics engineering from Bogazici University, Istanbul, Turkey, in 2001, the M.S. degree in computer science from Istanbul Technical University, Istanbul, Turkey, in 2003, and the Ph.D. degree in electrical and computer engineering from Rice University, Houston, TX, in 2008.

After his Ph.D., he worked as an Assistant Professor at the Department of Computer Engineering, Yaşar University, Izmir, Turkey. He is currently a Researcher with the European Commission Joint Research Centre, the Institute for Environment and Sustainability, Ispra, Italy. His research interests include detailed knowledge discovery from high-dimensional and large data sets, particularly multi- and hyperspectral imagery, self-organized learning, manifold learning, data mining, and pattern recognition. He is currently working on developing advanced control methods for monitoring agricultural resources using remote sensing imagery.