Vine field monitoring using high resolution remote sensing images: segmentation and characterization of rows of vines

Suivi des vignes à l'aide des images de télédétection à haute résolution : segmentation et caractérisation des rangs de vignes

Jean-Pierre DA COSTA*, Christian GERMAIN, Olivier LAVIALLE,
Saeid HOMAYOUNI and Gilbert GRENIER

LAPS CNRS – ENITAB – ENSEIRB, Université Bordeaux 1
351 cours de La Libération, 3305 Talence cedex, France
Tél. +33 5 4000 2634, Fax +33 5 4000 8406
Corresponding author: jean-pierredacosta@laps.u-bordeaux1.fr

Abstract: A new framework for the segmentation and characterization of row crops on remote sensing images has been developed and validated for vineyard monitoring. This framework operates on any high-resolution remote sensing images since it is mainly based on geometric information. It aims at obtaining maps describing the variation of a vegetation index such as NDVI along each row of a parcel.
The framework consists in several steps. First, the segmentation step allows the delineation of the parcel under consideration. A region-growing algorithm, based on the textural properties of row crops, was developed for this purpose. Once the parcel under consideration is delineated, a boundary smoothing process is applied and the row detection process begins. Row detection operates by means of an active contour model based on a network of parallel lines. The last step is the design of vegetative vigor maps. Row vigor is computed using pixels neighboring the lines of the network. Once row vigor is obtained on the rows, 2D vigor-maps are constructed. The values measured on the row are propagated to the inter-row pixels, producing «continuous» vigor maps ready to be exported to a GIS software.
We successfully exercised our framework on vineyard images. The resulting parcel segmentations and row detections were accurate and the overall computational time was acceptable.

Keywords: remote sensing, image processing, row crop, vine

Introduction

High-resolution remote sensing imagery is becoming a reality in the context of Precision Agriculture. In the future, it will be an essential tool for site-specific crop management and will help in the global knowledge of land use and quality control. Precision agriculture technologies and more particularly remote sensing imagery are expected to provide detailed maps of the crops. Such maps, which can be frequently updated, will enable the determination of management zones within the fields [1][2].
Applied to row crops such as vines, precision agriculture concepts imply that the field can no longer be considered as a whole. Each row should be processed separately and inter-rows, which often consist of bare soil or grass, should be considered separately. Very high-resolution images are then necessary to attain accurate row discrimination and characterization.
Considering the particular spatial arrangement of row crops, the design of management maps requires the following steps: image acquisition, field delineation, row detection and row processing. The results can then be integrated into a Geographic Information System (GIS). In this paper, we describe a complete framework, from image acquisition to row processing, and we apply it to the computation of vigor maps for vine parcels.

Material and Methods

Acquisition methods and images

Our study is carried out on very high-resolution remote sensing images taken from a plane. Pixel resolution is about 0.10 m. The available data (figure 1) are either color or multi-spectral images of several vineyards in the Bordeaux area, taken during the 2002 to 2003 and 2004 growing seasons i.e. from April through August. The methods described in this paper apply to grey level images, obtained from a single spectral band or from some combination of several spectral bands. For instance, the Normalized Difference Vegetation Index (NDVI) or the Ratio Vegetation Index (RVI), widely used in agricultural engineering, are meaningful measures of plant vigor [3][4][5]. The choice of the appropriate combination depends on the context and is not addressed here.

Figure 1 – Typical high-resolution vineyard remote sensing image

Parcel delineation

The segmentation algorithm for parcel delineation is derived from region-growing methods [6][7]. It relies on the specificity of the observed texture. Indeed, the spatial layout of row crops shows a strong anisotropy.
The segmentation is carried out by a region-growing algorithm based on anisotropy features. Region growing methods is suited to the delineation of a unique region but requires some supervision from the operator. We consider this basic task to be acceptable and entrust it to the end-user.
The aggregation proceeds in the four steps hereafter.

Reference region selection and feature computation – In order to characterize the row crop texture, we choose a textural feature based on local orientation. The orientation is associated with a confidence index related to texture anisotropy. The computation of the textural feature is carried out as follows :

- *Computation of the local orientation and the anisotropy index* – These attributes are computed for each pixel of the image. For this purpose, we use the « adaptive framework for unbiased orientation estimation » described in [8].
- *Manual selection of a rectangular area inside the parcel under consideration* – The reference rectangle constitutes the seed of the region-growing algorithm.
- *Computation of the Directional Mean Vector V_{ref} in the reference region* – This vector [9] gives us both the orientation of the plot and its anisotropy.

Coarse aggregation stage – Pixels surrounding the reference region are iteratively processed and aggregated if they belong to the parcel. The decision of aggregating a pixel p to the reference region is taken on the basis of the inner product of its feature vector $V(p)$ by the reference feature vector V_{ref}. The pixel p is aggregated if the inner product is greater than a given threshold d_{th}.
The objective of the coarse aggregation stage is to aggregate many pixels as fast as possible. To do so, small windows W around the reference region W_{ref} are analyzed by computing the window feature vector and by comparing it to the reference vector V_{ref}. If the inner product is greater than the threshold d_{th} then all the pixels inside W are aggregated simultaneously.

Fine aggregation stage – After coarse aggregation, pixels at the border of the parcel are examined one by one. We use the same threshold value d_{th}. If the inner product is lower than the threshold, the pixel is added to the parcel. This process is carried out for all neighboring pixels until aggregation is no longer possible.

Boundary smoothing – When a pixel p is at the border between the parcel under consideration and a path, the observation window $W(p)$ combines both rows and noisy structure-less paths. This local layout may produce an irregular feature vector $V(p)$ and the inner product between $V(p)$ and V_{ref} may be below the threshold. This leads to an irregular border. To get smoother borders and to fill the holes that may appear inside the parcel, a morphological closing is processed.

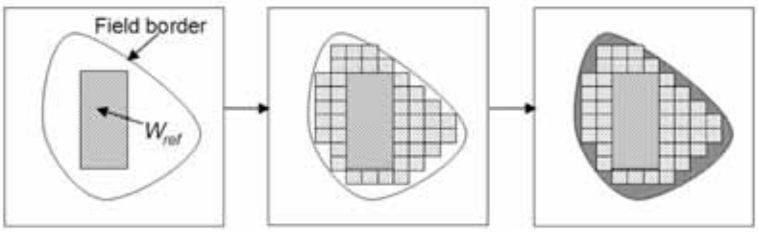

Figure 2 - The three basic segmentation steps: choice of a reference, coarse aggregation and fine aggregation.

Row detection

Most computer vision methods for detecting crop rows are implemented in real-time systems for automatic guidance of agricultural implements [10][11] where, for computational reasons, segmentation of crop/weed pixels is usually avoided. Some authors carry out the row detection using a simple thresholding [1]. Nevertheless, the choice of a threshold is awkward, especially when grass is present between rows.
Contrary to the real-time context, computational time is not a critical issue for precise location of rows in crop monitoring. In contrast, row position and angle estimations have to be very precise. The specific arrangement of individual crops suggests the use of the Hough transform (e.g. [10][12]). However, previous attempts, not reported in this paper, have shown that some limitations of the Hough transform because of the large number of linear structures and the presence of undesirable alignments.
In order to retrieve the rows with high accuracy, we implement an active contour network, which aims at fitting a line to each row through a global convergence process. Such a network has already been used in [14] where it was applied to an image completely covered by vines. We propose here to generalize this approach to any image showing a vineyard segmented from its neighborhood.

Definition of an active contour model or snake – A snake is a geometric object whose features can evolve over time [15]. Rigid models composed of predefined shapes are more suited to our application. Only their orientation and position can move to a stable state, while respecting a given set of constraints. The stable state corresponds to the minimum of a pre-defined energy E: $E = \mu E_{\text{int}} + (1-\mu)E_{ext}$

The internal energy E_{int} reflects the internal constraints that the snake undergoes. The external energy E_{ext} takes into account the effect of the image on the snake. $\mu \in [0,1]$ is a weighting parameter. The evolution of the snake is an iterative process. The snake stops moving when the minimum of the energy is reached.

Choice of the model – The most appropriate model is a network of quasi-parallel segments. Each segment S_i is completely defined by the two parameters θ_i and $P_i=(x_i,y_i)$ (fig. 3). P_i is the pivot around which S_i rotates. During the evolution of the snake, the positions and the orientations change until the energy of the snake reaches a minimum.

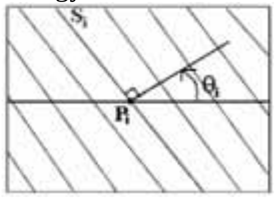

Figure 3 - Pivot P_i and orientation θ_i

Initialization of the snake – The segments have to be located as close to the underlying rows as possible.
Good approximations of row orientation and spacing can be obtained by spectral methods. Since we already know the global orientation of the rows, thanks to the delineation step, the row spacing is all that is left to estimate. In order to do so, the image is rotated to the horizontal. Then the grey levels of the pixels are projected horizontally to produce a 1-D signal. On this signal, we compute a 1-D Fast Fourier Transform (FFT). The maximum of the power spectrum provides with the mean row spacing.
Before starting the iterative process, the rough orientation, previously obtained thanks to the Directional Mean Vector, and the mean spacing measured on the power spectrum are used to fix the location of the segments on the rotated image. Pivots are placed on a vertical line that intersects the largest number of rows. Hereafter, the method consists in finding all the minima of the luminance along that line.

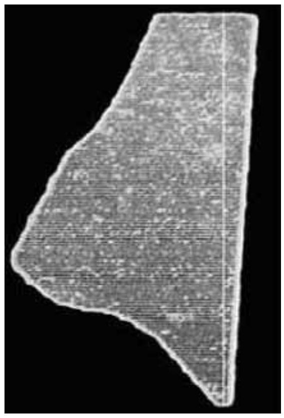

Figure 4 - Vertical line orthogonal to a maximum number of rows

Convergence of the snake – The iterative convergence algorithm, described in [14] carries out the minimization of the energy, E. The external energy E_{ext} models the attraction of the segments by the rows. It must be minimal when the segments are close to the rows. As the rows are characterized by a weak luminance, we define E_{ext} as the integral of the luminance along the segments. The internal energy E_{int} enables the control of the snake shape. It is composed of two parts E_x and E_θ. E_x maintains a regular spacing between segments while E_θ ensures a regular angle between segments.

Results

We present here the results obtained at each step of the framework for the image in (fig. 1). Hereafter, we will propose some approaches to the use of these results for cartography.

Segmentation step and row detection step

The window aggregation step provides us with a rough estimation of the field shape. It also shows some difficulties concerning regions of the field that differ from the reference region: for instance, the upper left area, which suffers from low vegetative vigor, is left out after the coarse aggregation step. The pixel aggregation step refines the boundaries, and the post segmentation processing fills the holes and smoothes the sharpest variations along the edges. The segmentation result is accurate as shown in (fig. 5).
As described previously, the image is rotated before the row detection step so that the line network appears horizontally. The initialization stage performs well. Pivots are placed correctly in spite of the absence of foliage in some locus. After convergence of the active network, the lines fit almost perfectly the central axis of the vine rows (fig. 6) overcoming the irregularity of foliage and the absence of vine plants.

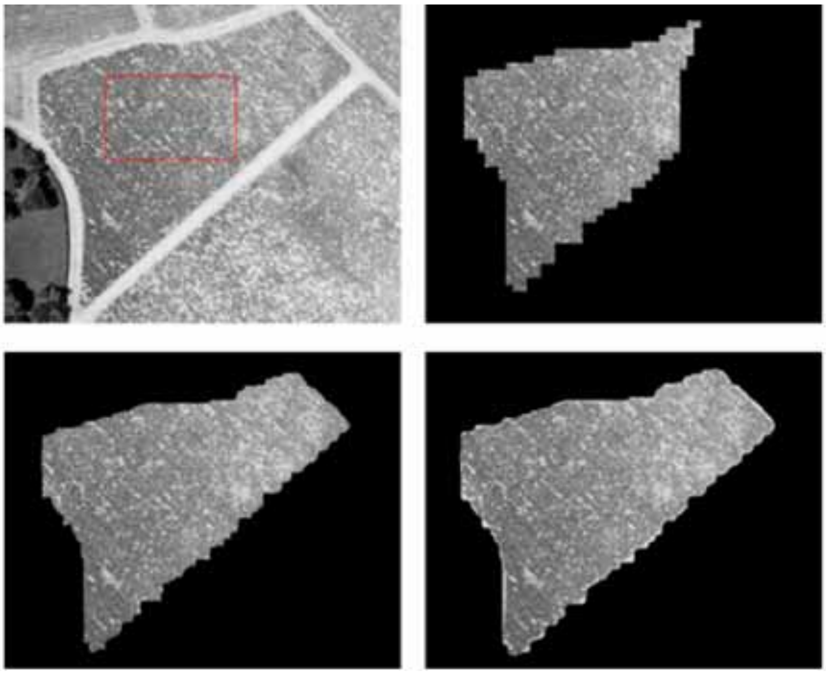

Figure 5 - The four-steps segmentation process

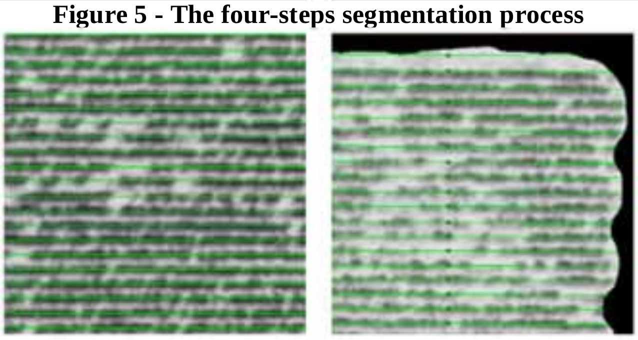

Figure 6 - Details of the row detection step.

Row feature computation

The active contour model makes it possible to take measurements along crop rows and to assign the measured feature to a precise locus on the row. The feature to be measured will depend on the type of crop and on the type of information needed by the end-user. In the case of vineyards, it may be useful to characterize the vine canopy by measuring its width (or any vegetative index) along vine rows. In all cases, the measured index must be assigned to a particular vine plant on the row.
The grey level profile along a line orthogonal to the vine rows shows an alternation between dark pixels corresponding to foliage and light pixels corresponding to soil. In order to produce a vegetative index $V(i,j)$

for a given locus *(i,j)* on the vine row, we propose to simply invert the data, to sum the inverted pixel values vertically around the considered locus and to assign the final value to the pixel on the row.

Note that *V(i,j)* may take into account both vine and non-vine pixels in the event of missing plants. Non-vine pixels may thus have an influence on *V(i,j)*. When working with the NDVI index, this influence is reduced since soil pixels have very low NDVI values. Figure 7 shows an extract of the vegetative index map, with values computed only on row pixels.

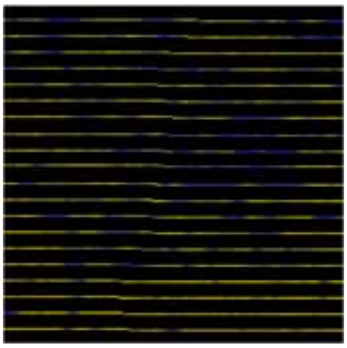

Figure 7 - Discontinuous vigor map

Continuous map design

At this stage, the computed vegetative index is assigned only to the pixels located on the lines of the network. Despite the fact that it accurately depicts vine behavior, the discontinuity of this map makes it inappropriate for visualization purposes (see fig. 7). The conversion of this discontinuous vegetative index map into a continuous 2-D map can be carried out either by using a 2-D interpolation or by the duplicating the vegetative index values from the lines of the network to the inter-rows. In the case of 2-D interpolations, the krieging function of a GIS could fulfill this task.
However, we propose here to produce a continuous vegetative index map. For this purpose, we carry out the propagation of the vegetative index between the lines of the network using a 2-D Deriche isotropic filter. The recursive implementation of this filter requires a low computational time [16]. Finally, the continuous vigor map is rotated back to the initial orientation of the vine parcel. The result is shown on figure 8. The palette used shows missing plants in blue, the average vigor areas in yellow, and the most vigorous areas in orange.

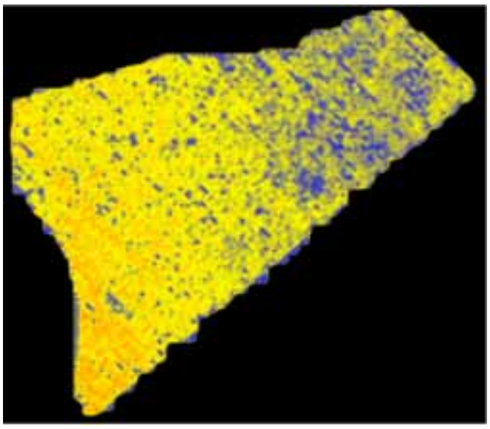

Figure 8 - Continuous vigor map.

Discussion

Using our framework with high-resolution remote sensing images, discriminating row crops from non crop areas is possible. This framework should allows us to draw several maps revealing row crop characteristics, among which density, spacing, NDVI or biomass. It should be noted that our framework can be adapted to multi-spectral or hyper-spectral images, applying each step either on a single original wavelength or on a combination of several wavelengths such as the well known NDVI or RVI indices.
In some cases, the automatic segmentation step might not be needed, especially when digital orthophotos are used together with field boundaries maps registered in a GIS. Otherwise, the segmentation step requires very low input from the end-user who just needs to draw a small rectangle inside the field he wants to survey rather than clicking on each vertex of the polygon surrounding the field.
Considering computational time, Table 1 provides detailed results obtained on Fig. 6a (size 1829×1605). Algorithms were processed on an Intel Pentium M 2.0 Ghz with 1 Gbytes of memory.

Orientation computation	02 s
Automatic segmentation	31 s
Rotations (direct and inverse)	03 s
Row detection	04 s
Computation of the vigor on row pixels	< 1 s
Inter-row filling	01 s
Other	01 s
Total	43 s

Table 1 - Computational time for fig. 1 processed on an Intel Pentium M 2.0 Ghz with 1 Gbytes of RAM.

Rotations, row detection and vigor map computation take only a few seconds. The automatic segmentation is the most time-consuming step. The cost of the segmentation step goes from 20 to 40 seconds, depending on the surface to be segmented. This duration may seem too long and an improvement of the segmentation cost should be feasible. However, the automatic segmentations turns out to be often worthwhile compared to a manual delineation. The row detection and map drawing steps, take about 10 seconds altogether which is considerably shorter than clicking and identifying each row.

Conclusion

We have proposed a new framework for the segmentation and characterization of row crops using high-resolution remote sensing images. More specifically, we have introduced new algorithms for parcel delineation and for row detection in the case of row crops.
The segmentation relies on a region-growing algorithm based on a textural description of row crops. The coarse to fine scheme allows to detect parcel boundaries accurately. The row detection algorithm implements an active contour network which aims at placing lines on the very middle of the rows. The process consists in minimizing an appropriate energy function in order to converge quickly towards the optimal line network.
The resulting lines are the basis for a robust estimation of a vegetative index. This index, computed along each line of the network, avoids taking into account the soil and grass pixels in the inter-rows.
The last step consists in the propagation of values measured on lines, to inter-row pixels in order to obtain "continuous" maps of the row vigor.
Our framework has been successfully applied to draw vigor maps of vine parcels. Such maps may help to split the parcel into differentiated management zones, and can even be compared with in-field measurements taken on soil or vines, e.g. yield, electro-resistivity, etc.
Prospective work concerns the reduction of the processing time for the segmentation step and the implementation of a new network model taking into account row curvature in order to strengthen slope variation robustness.

Acknowledgments: This work has been done under the patronage of the *Institut des Sciences de la Vigne et du Vin* and of the *Centre Inter-professionnel des Vins de Bordeaux*, with the financial support of the FEDER. We are grateful to the vine growers Château Palmer, Château Grand Baril and Château Luchey-Halde, for providing image data. Finally, we thank the "Œnologie-Ampélologie" laboratory, ENITAB-INRA, for its technical support.

References

[1] A. Hall, J. Louis, D.W Lamb, 2003. Characterizing and mapping vineyard canopy using high spatial resolution aerial multispectral images. *Computers & Geosciences*, **29**, 7, 813-822.
[2] B. Tisseyre, N. Ardoin, F. Sevila, 1999. Precision viticulture: precise location and vigour mapping aspects. *Proc. of the ECPA*, Odense, Denmark, 319–330.
[3] A. Hall, D.W. Lamb, B. Holzapfel, J. Louis, 2002. Optical remote sensing applications in viticulture – a review. *Australian Journal of Grape and Wine Research*, **8**, 36-47.
[4] R.M. Pearson, L.D. Miller. 1972. Remote mapping of standing crops biomass for estimation of the productivity of the short grass prairie. *Int. Symp. on remote sensing of the environment*, Ann Arbor, MI.
[5] J.W. Rouse, R.H. Haas, D.W. Deering, J.A. Schell, J.C. Harlan, 1974. Monitoring the vernal advancement and retrogradation (green wave effect) of natural vegetation; NASA/GSFC Type III Final Report, Greenbelt, MD., 371.
[6] C.R. Brice, C.L. Fennema. 1970. Scene Analysis Using Regions. Artificial Intelligence, 1-3/4, 205-226.
[7] Y. Chang, X. Li,1994. Adaptive image region growing. *IEEE Trans. on Image Processing*, **3**, 868-872.
[8] F. Pouliquen, J.P. Da Costa, C. Germain, P. Baylou. 2005. A new adaptive framework for unbiased orientation estimation in textured images. *Pattern Recognition*, **38**, 11, 2032-2046.
[9] C. Germain, J.P. Da Costa, O. Lavialle, P. Baylou. 2003. Multiscale estimation of vector field anisotropy. Application to texture characterization. *Signal Processing*, **83**, 7, 1487-1503.
[10] R. Keicher, H. Seufert, 2000. Automatic guidance for agricultural vehicles in Europe. *Computers and Electronics in Agriculture*, **25**, 169-194.
[11] T. Hague, N.D. Tillet, 2001. A bandpass filter-based approach to crop row location and tracking. *Mechatronics*, **11**, 1-12.
[12] J.A. Marchant, 1996. Tracking of row structure in three crops using image analysis. *Computer and Electronics in Agriculture*, **15**, 161-179.
[13] H.T. Søgaard, H.J. Olsen, 2003. Determination of crop rows by image analysis without segmentation. *Computers and Electronics in Agriculture*, **38**, 141-158.
[14] W. Bobillet, J.P. Da Costa, C. Germain, O. Lavialle, G. Grenier, 2003. Row detection in high resolution remote sensing images of vine fields. *Proc. of ECPA*, Berlin, Germany.

[15] M. Kass, A. Witkin, D. Terzopoulos, 1988. Snakes: Active Contour Models. *International Journal of Computer Vision*, **4**, 321-331.
[16] R. Deriche, 1990. Fast algorithms for low-level vision. *IEEE Transactions on PAMI*, **12**, n° 1, 78-87.