Vineyard detection from unmanned aerial systems images

(Article begins on next page)

19 April 2024

Vineyard detection from unmanned aerial systems images

(or

Vineyard plot and row detection using centimetre-resolution unmanned aerial systems images)

Lorenzo Comba[a], Paolo Gay[a,b,*], Jacopo Primicerio[a,c], Davide Ricauda Aimonino[a]

[a] DI.S.A.F.A. – Università degli Studi di Torino, 44 Via Leonardo da Vinci, 10095 Grugliasco (TO) – Italy

lorenzo.comba@unito.it, paolo.gay@unito.it, davide.ricauda@unito.it

[b] CNR-IEIIT, 24 Corso Duca degli Abruzzi, 10129 Torino – Italy

[c] CNR – IBIMET, Via G.Caproni, 8, 50145 – Italy

j.primicerio@ibimet.cnr.it

* Corresponding author: Paolo Gay

Email: paolo.gay@unito.it Ph: +39 011 6708620 Fax: 011 6708591

Abstract

In viticulture, the adoption of precision agriculture techniques is nowadays increasingly essential to reach required high product quality standards. New reliable tools for mapping crop variability indexes in a vineyard or a single parcel are necessary to deploy site-specific management practices. In this paper, a new method to automatically detect vine rows in grey-scale aerial images is presented. The developed image processing algorithm is constituted by three main steps based on dynamic segmentation, Hough Space Clustering and Total Least Squares techniques. The procedure's reliability has also been proven in the presence of disturbance elements, like dense inter-row grassing, bushes and trees shadows, by properly detecting vine rows in the

vineyard images. Moreover, its adaptive features allow it to obtain optimal results in the presence of uneven image illumination due, for example, to the presence of clouds or steep terrain slopes. The extracted row and inter-row information, besides being the basis for vineyard characterization maps computation, like vine plants vigour maps, could also be used as a reference for other precision viticulture tasks such as, for example, path planning of unmanned ground vehicles.

Keywords: Vineyard detection, unmanned aerial vehicles, remote sensing, precision viticulture

1. Introduction

Nowadays, required high quality standards impose to many farms and growers the adoption of precision agriculture techniques and site-specific management practices (Kropff et al., 1997, Zhang et al., 2002; Stafford, 2006). This is especially true in modern viticulture, where the possibility to distinguish between zones of different grape quality within the same parcel and to execute appropriate management according to inherent crop variability could provide for an increase in economic benefits and a reduction of the environmental impact (Arnó et al., 2009). The implementation of a precision viticulture (PV) approach to vineyard management is a process that begins with the observation of vineyard performance and associated vineyard attributes, followed by the interpretation and evaluation of the collected data (Proffit et al., 2006; Tisseyre et al., 2007; Bramley and Reynolds, 2010).

Remote sensing is one of the most powerful tools in PV, being able to rapidly provide a description of grapevine shape, size and vigour over entire vineyards, and to find relationships

between these canopy descriptors and grape quality and yield (Lamb et al., 2004; Hall et al., 2008a; Hall et al., 2008b; Hall et al., 2011). The applications of satellite or airborne imaging in PV have nevertheless been limited by poor revisiting times, coarse spatial resolutions, high operational costs and complexity, and lengthy delivery of products (Zhang and Kovacs, 2012). With recent technological advances in aerospace engineering, the acquisition of Earth surface images at a low altitude (Low Altitude Remote Sensing system - LARS) by using Unmanned Aerial Systems (UAS) is being promoted as a suitable alternative for this purpose. This technology allows the acquisition of ultra-high spatial resolution aerial maps with low operational costs and near real-time image acquisition (Zhang and Kovacs, 2012).

To calculate vegetation indexes from aerial images, detection of vineyard plot is first needed, in order to obtain vineyard maps "cleaned" from all the features that are usually present in field aerial photography, such as roads, trees and bushes. Nonetheless, the extraction of pure vines pixels, discarding inter-rows areas, is crucial for enhancing the final vineyard maps quality and effectiveness for site-specific management (Hall et al., 2003; Kazmierski et al., 1011; Puletti et al., 2014). This task is usually complex and user-intensive, since the reflectance response of all these features is often similar to the canopy to be extracted. Many remote sensing systems and algorithms have been devised to study canopy reflectance response (Zarco-Tejada et al., 2005), but the possibility of exploiting the information potential of the images depends on the development of image processing methods for analysing textured images (Delenne et al., 2010). Starting from simple value thresholding techniques (see e.g. Hall et al., 2003), many methods have been developed to face the problems of vine rows and vineyard detection, either with texture or frequency analysis. Vine rows and parcels detection are strictly related since in most cases vine plots are identified using row information. Ranchin et al. (2001) proposed a method based on wavelets and multi-resolution analysis that obtained, on average, a score of 78% of correctly

identified plots, although it is complex and requires significant user intervention. Da Costa et al. (2007) proposed a method which takes into account the textural properties of vine images, where a thresholding operation on textural attributes allows discriminating between vine field and non-vine field pixels. With a different approach, using the typical periodic patterns of this crop, Delenne et al. (2006) and then Rabatel et al. (2008) introduced a recursive process based on the Fast Fourier Transform (FFT) and the Gabor filtering of the aerial images to show the delineation of vineyard plots and the evaluation of row orientation and inter-row width. Some factors, such as alternate treatments between inter-rows and high numbers of vines which are missing or too young, had the consequence of compromising the periodicity of the vine row patterns and then, consequently, performance degradation. More recently, Delenne et al. (2010) improved the performance of their vine plot detection method, based on the FFT and the frequency domain Gabor filtering, by enhancing the precision of the boundary locations with a precise adjustment of each row (still, the method presents an over-estimation of the vine parcels in case of missing vines). An active contour model developed by Bobillet et al. (2003) aims at fitting a line to each vine row through a global convergence process. However, the limit of this method lies in the initialization process, essential to reach the optimum of the convergence process, which requires segments to be placed as close to the underlying vine rows as possible. This problem has been solved by defining a model of roughly parallel and equidistant lines, although this approach limits the application area of the method, excluding less regular vineyards.

With a different approach, Smit et al. (2010) proposed a segmentation method that combined a thresholding technique with a subsequent graph-based analysis. This method provided good results in the case of a flawless segmentation step, which requires a manual setting of image pre-processing parameters and threshold values, but it had difficulties in managing the presence of non-vine vegetation near the parcels and within the inter-rows.

The objective of the method proposed in this paper is to use information collected in high-resolution aerial digital images to determine:

1- a mask which identifies all the pixels that represent vines foliage and discriminates the vineyards from the background comprising of inter-row soil and grass, bushes, trees and other elements of the rural area

2- straight lines representing the optimal alignments of vine plants in the vineyards.

The adoption of the proposed method should not be limited to vineyards with parallel and equidistant vine rows, so that no spatial periodicity is required by the image elaboration process. The algorithm must also be able to obtain reliable results even in the presence of disturbance elements in the image, such as roads, shadows, bushes, trees and dense inter-row grassing, and to deal with the occurrence of non-uniform illumination during image acquisition, due to the presence of clouds or to the slope angle of the terrain. Providing robustness in heterogeneous scenarios, these features, joined with a limited number of required parameters, must made the method able to operate autonomously, without user intervention.

This paper is structured as follows. Section 2 describes the algorithm, which can be divided in three main processing steps. Results of the method application to four aerial images are presented in Section 3. Moreover, for the sake of clarity, additional detailed images, relative to each processing step, are shown in the auxiliary material for all the sample elaborated images. An animation of the processing of the example image, used for the explanation of the algorithm in Section 2, is also available as additional material on the Journal website. Conclusions and future developments are finally discussed in Section 4.

2. Materials, methods and algorithm description

The image elaboration process presented in this study allows locating vineyard areas within grey-scale aerial images, obtained by a sensor camera installed on a UAV. Depending on the adopted airborne sensor, the input grey-scale digital images can consist of a single plane of an RGB image, a linear combination of R, G and B values, a NIR (Near InfraRed) image or they can be obtained as the result of an index computation such as, for example, the well known NDVI (Normalized Difference Vegetation Index), calculated as

$$NDVI = \frac{(NIR - R)}{(NIR + R)}$$
(1)

as proposed by Rouse (1974). In the presence of marked geometric distortion in imagery, due to the optical equipment, it is necessary to apply proper algorithm correction to obtain a not altered aerial image.

The main difficulty of vineyard and vine rows detection lies in the fact that the sensor response to grass, trees and any other kind of native vegetation are typically very similar, and in many cases indistinguishable, from vineyards. For this reason, spatial information like vine row location and orientation, inter-row width, etc., which is the object of this paper, cannot be derived from vineyard images by using threshold processing techniques and algorithms. On the other hand, the use of frequency transform-based methods presents low discriminating performances in the case in which vine row periodicity is not very evident (different row angles, few rows aligned with the same angle, etc.) and/or in the presence of other (disturbance) elements located in an image showing any kind of spatial periodicity.

The method proposed here is based on a sequence of operations, sketched in the flow chart in Figure 1, that, step by step, refine the image by masking non-vineyard vegetation components, finally leading to the extraction of the desired features.

The description of the method is divided in three main steps concerning (1) the preliminary detection of pixels (potentially) representing vine plants in order to obtain a clear binary image with clusters of interconnected pixels; (2) the determination of line parameters that best fit each group of pixels by using clustering techniques in the Hough Parameters Space (HPS) and (3) the Total Least Squares (TLS) identification of straight lines representing the optimal estimation of vine row alignment. Based on the results of this approximation, a final rejection of regions (groups of pixels) that do not represent vine rows is performed.

The image elaboration process has been implemented in a software framework (Matlab$^{©}$ r2013) that automatically collects the results, such as the spatial coordinates of each vine row, in tables and provides additional graphical outputs to facilitate the interpretation of the elaboration. The entire process is performed on a 3.2 GHz Quad-Core Intel Xeon server with 8 GB of RAM memory (1066 MHz DDR3 ECC).

2.1 Data

The proposed method is here presented by showing the effect of each step on a near-infrared image (see the raw image in Figure 2) of a vineyard located in Montalcino (Siena), in the southern area of Tuscany, Italy. The image was acquired with a Tetracam ADC-lite camera (Tetracam, Inc., Gainesville, FL, USA) installed on a modified Mikrokopter Hexa-II (HiSystems GmbH, Moomerland, Germany). The 3.2-megapixel CMOS sensor, fitted with an 8.5 mm lens, has a field of view of 43° and is able to achieve a 0.056 m/pixel ground resolution at a flight altitude of 150 m. The resolution of the acquired image is 2048 × 1536 pixels, whose values – Digital Numbers (DN) – use an 8 bit representation.

Furthermore, the effectiveness of the proposed method has been demonstrated by reporting the results of the elaborations applied to four sample vineyard images . The interested reader can find the results of each processing phase, applied to each image, in the additional electronic material.

2.2 Segmentation and first clustering

The first step of the developed procedure aimed at making a preliminary (rough) detection, in the aerial image, of pixels representing vineyard vegetation, obtaining a binary image where nonzero-value pixels represent potential vine plants. As already mentioned, the similar response to different kinds of vegetation and the presence of non-uniform (and uncontrollable) illumination of the targets, which locally affect pixels DN, prevent the classification of pixels by directly applying a threshold to the 255-values grey-scale image. Here, the problem of segmentation is handled firstly by processing the image with a dynamic local neighbouring window segmentation procedure, which provides a (segmented) image with a sharpened contrast between regions of interest (ROIs), i.e. vine rows and background pixels representing soil, grass and other vegetation. This technique consists in classifying each pixel based on the DN of its neighbours. For each image pixel, this procedure is performed n_w times by varying the specific set of neighbouring pixels and by recording all results in a proper counter. In more detail, subsets of pixels $\Omega_{x,y}$ of the entire image are selected by using a squared sliding window of size l_w, centered on the pixel at position (x, y). Best results have been obtained with l_w equal to 2 inter-row width, but good results can been obtained with l_w ranging from 1 up to 6 inter-row width, avoiding a fine calibration of the parameter for the profitable adoption of the method. The window slides with a moving step of $s_w = l_w/\sqrt{n_w}$ pixels, both in the horizontal and vertical directions, with the parameter n_w that can be chosen arbitrary between 1 and l_w^2. Obviously, the procedure effectiveness increase with higher n_w values requiring, contextually, a longer processing time.

For instance, in Figure 2 two subsets Ω_{x_a,y_a} and Ω_{x_b,y_b} of pixels are marked with squares centred on pixels $a = (1150,250)$ and $b = (950,850)$ respectively, and their enlargement is shown in Figures 3.a and 3.b. The size l_w of the windows, as previously discussed, has to be chosen large enough to include,

whenever is the case, at least two vineyard rows. The possible presence of a portion of vineyard in $\Omega_{x,y}$ is evaluated considering the intensity distribution histogram of DNs belonging to $\Omega_{x,y}$. The histogram of a region containing a vine row typically presents a bimodal behaviour. This condition is detected by introducing the normalized moment $J(\Omega_{x,y})$ defined as

$$J(\Omega_{x,y}) = \frac{\sum_{v=1}^{255} n(v) \cdot (v - \mu_{x,y})^2}{N} \qquad (2)$$

where $n(v)$ is the number of pixels characterized by DN intensity value v, $\mu_{x,y}$ is the mean of the local intensity distribution histogram and $N = l_w^2$ is the number of pixels belonging to $\Omega_{x,y}$. The index $J(\Omega_{x,y})$ is therefore independent from window size l_w and the standardization with respect to $\mu_{x,y}$ allows the segmentation process to be robust against non-uniform image illumination. If $J(\Omega_{x,y})$ is greater than fixed threshold $T_J = 5 \cdot 10^2$, then pixels that belong to $\Omega_{x,y}$ and have a DN value greater than $\mu_{x,y}$ are classified as part of a vine row and their counters are increased. Otherwise, if $J(\Omega_{x,y})$ is lower than T_J (e.g. in the case of bell-shaped intensity distribution histogram), no segmentation is applied, therefore counters will not be updated. Since $J(\Omega_{x,y})$ is normalized with respect to N, T_J threshold value is invariant to the mobile windows size l_w. Figures 3.c and 3.d show the intensity distribution histogram of subsets Ω_{x_a,y_a} and Ω_{x_b,y_b} (Figure 3.a and 3.b) respectively, while the segmentation of subset Ω_{x_b,y_b} is shown in Figure 4: counters of pixels marked in white colour will be increased by one unit. Each pixel, analysed n_w times with a different set of neighbouring pixels, therefore results in a counter ranging from 0 to n_w. An example of the achieved counter value for each pixel in the image is shown in Figure 5.a, where can be seen that the proposed procedure provides a segmented image with an enhanced separation between vines and the background. A threshold operation is then applied to the image derived from counters, obtaining a binary image where pixels are considered part of vine rows, or small vegetation, if their DN is greater than $0.8 \cdot n_w$, as displayed in Figure 5.b. Since the high contrast of the

segmented image, the threshold operation is very robust and the best threshold value settled to $0.8 \cdot n_w$ results to be invariant to the aerial image characteristics. As can be seen, vine rows are now identifiable, even if a number of other non-vine objects are still present in the image and need to be detected and eliminated with a different technique. A sequence of morphological operations is then adopted in order to clean the binary image, remove small objects and favouring the detection of sets of interconnected pixels $A \in \mathcal{A}$, where $\mathcal{A}$ is the set of all detected clusters in the processed image (Gonzales and Woods, 2002). For instance, the image resulting from the sequential application of Open, Close and Remove Small Objects operations is reported in Figure 6. Note that, due to the presence of dense grassing within vine inter-rows and/or small vegetation near the parcel, several vine rows could be classified as belonging to the same cluster of pixels that, therefore, will assume a complex shape, like the one highlighted in Figure 7. Isolation of each vine row in a separate cluster is then achieved by processing the entire set of clusters $\mathcal{A}$ with the second module of the procedure, which is based on a clustering procedure in the Hough parameter space.

2.3 Cluster isolation

Even if segmentation of the original image with the dynamic local neighbouring window handles the problem of non-uniform illumination of the image well and allows detection and deletion of large vegetation areas out the vineyard parcels, it could generate clusters of interconnected pixels, with complex shapes, which include two or more adjacent vine rows. As already mentioned, this could occur in the case of too dense grassing and/or other small crops detected between rows. It is necessary to distinguish and separate each vine row, deleting inter-row vegetation. This goal is reached firstly by projecting each cluster of pixels A in the digital image on the HPS (see e.g. Duda & Hart, 1972), and then to proceed with the extraction of common geometric features. In particular, this method is adopted here

to identify a set of rectilinear lines $\mathfrak{L}_A$ in a cluster A: each line l, with $l \in \mathfrak{L}_A$, is parameterized in terms of distance ρ from the origin of the image and the clockwise inclination angle θ relative to the vertical axis, which fits a single vine row in A well.

2.3.1 *Hough transform-based line detection*

To improve performance, the Hough Transform (HT) has been applied singularly to each cluster A, and not to the entire binary image directly (as, for instance, in Recio et al., 2013). Each pixel $p = (x, y)$ in the image space, $p \epsilon A$, has been transformed to the curve

$$s(p) = \{(\rho, \theta), \rho \in [-d, +d], \theta \in (-90°, +90°]: \rho - x \cos \theta - y \sin \theta = 0\} \qquad (3)$$

in the HPS. Then, the function

$$S\big((\rho, \theta), A\big) = \mathrm{card}\{s(p): s(p) \cap Z\big((\rho, \theta)\big) \neq \emptyset, \forall p \in A\} \qquad (4)$$

has been defined as the number of curves $s(p)$ that pass through a neighbourhood $Z(q)$ of $q = (\rho, \theta) \in HPS$, where

$$Z(q) = \{q' = (\rho, \theta): \|q - q'\|_{\infty, w} \leq 1, w = [w_\rho \quad w_\theta]^T \}, \qquad (5)$$

$\mathrm{Card}(X)$ is the cardinality of set X and $\|x\|_{\infty, w}$ is the weighted ℓ_∞ norm of x. Since the value of $S\big((\rho, \theta), A\big)$ represents the number of collinear pixels aligned to line (ρ, θ), up to an error of $\varepsilon_\rho = w_\rho^{-1}$ and $\varepsilon_\theta = w_\theta^{-1}$, the absolute and local maxima of this function define the lines that best fit the cluster of pixels A. A rectangular tessellation with a resolution of one pixel along ρ and 0.1° for θ has been applied on the HPS. Then, function $S\big((\rho, \theta), A\big)$ has been computed on each generated region, centred in (ρ, θ), $\rho \in P = \{-d, -d + 1, \dots, +d\}$, $\theta \in \Theta = \{-89.9°, -89.8°, \dots, +90°\}$, assuming $w = [0.5^{-1} \quad 0.05^{-1}]^T$.

However, the selection based on the sole cardinality of aligned pixels is not sufficient to avoid false classification, also because $S\big((\rho,\theta),A\big)$, $\rho \in \mathrm{P}$ and $\theta \in \Theta$ often provide clouds of local maxima which are hardly distinguishable. Considering that, in vineyards, vine rows are usually arranged in parallel, their detection has been performed by investigating the most recurring values of θ in the HSP. This task has first been achieved by defining the function

$$S_{\mathrm{H}}\big((\rho,\theta),A\big) = \begin{cases} S\big((\rho,\theta),A\big) & \text{if } S\big((\rho,\theta),A\big) > \dfrac{\mathrm{H}}{2} \\ 0 & \text{ow} \end{cases} \tag{5}$$

where $H = \max_{\substack{\theta \in \Theta \\ \rho \in P}} S\big((\rho,\theta),A\big)$, and then by projecting the $S_H(\cdot)$ values on the θ axis as

$$I(\theta,A) = \sum_{\rho \in \mathrm{P}} S_{\mathrm{H}}\big((\rho,\theta),A\big), \text{ for } \theta \in \Theta. \tag{6}$$

Indeed, Eq.(5) has been introduced to eliminate numerous low-values of $S(\cdot)$ that may affect the accuracy of the projection process in determining mainly recurrent θ orientations.

Function $I(\theta,A)$ has then been smoothed by using a moving average filter on a window of 21 samples and its maxima $\mathcal{O}_A = \{\theta_1, \cdots, \theta_m\}$ have been determined. In a typical case, Eq.(6) presents only one maximum θ_1, which corresponds to have only one main vine row orientation in cluster A. However, in general, a number of m possible different orientations could be detected.

Once the most recurring values of θ have been localized, the next step consists in finding, for each $\theta \in \mathcal{O}_A$, the local maximum values of $S\big((\rho,\theta),A\big)$, $\rho \in P$, which are located in $\mathcal{P}_{\hat{\theta}} = \{\hat{\rho}_1, \cdots, \hat{\rho}_n\}$. For this operation it is also convenient to smooth out $S\big((\rho,\theta),A\big)$ with a moving average filter, before maxima detection. $\mathcal{P}_{\hat{\theta}}$ values therefore allow to locate, for each $\hat{\theta}$, the lines best fitting the vine rows in cluster A.

As an explicative example, in Figure 6 a cluster of pixels (A_{64}) has been selected and highlighted. Its enlargement and the graphical representation of function $S\big((\rho,\theta),A_{64}\big)$ are shown in Figure 7 and 8 respectively. The smoothed out $I(\theta,A_{64})$ function is shown in Figure 9.a: in this case only one maximum value has been detected, located in $\theta_1 = 76.7°$. Then function $S\big((\rho,\theta_1),A_{64}\big)$, shown in Figure 9.b, has been computed to detect $\mathcal{P}_{\hat{\theta}_1} = \{\hat{\rho}_1,\cdots,\hat{\rho}_4\}$. Four detected lines, which fit the vine rows in cluster A_{64}, are shown in red in Figure 7.

2.3.2 *Separate vine rows*

If more than one line l has been detected in a group of interconnected pixels A, it is necessary to crop the cluster into simpler elongated sub-groups representing a single vine row. Cutting lines are defined as bisectors of the acute alternate interior angles of the intersection of two adjacent detected lines l_i and l_{i+1}, like those displayed with black dotted lines in Figure 7. In the dividing operation the cluster shape and the cutting segments, defined as the intersection between the cutting line and the cluster area, must respect two geometrical criteria to avoid undesired separation: (1) the ratio between cutting segment length and cluster surface area must not exceed a fixed threshold, allowing only short cutting segments, and (2) the intersection of two adjacent lines l_i and l_{i+1} must not be located within the cluster surface. In Figure 7 small black regions, selected using a five-pixel band width, represent pixels to be removed in order to disconnect portions of the cluster. These criteria prevent the use of the dividing operator to pixel regions that seem to be attributable to non-vine vegetation. After this operation has been performed on the set of clusters $\mathcal{A}$, a number of sub-clusters emerge, each one representing a single vine row (or still a non-crop vegetation region).

In the case of clusters without a tapered shape, the determination of a few main θ values is not possible, and a cloud of any-oriented line is therefore generated. However, these lines will be discarded by the next step of the procedure.

2.4 Total Least Squares vine rows identification

Finally, the last step of the proposed method is designed to detect the geometric features of each cluster by rejecting residual non-vine vegetation areas. This is performed by using the Total Least Square (TLS) technique (Golub & Van Loan, 1980; Markovsky & Huffel, 2007) to identify the best-fitting line of the pixels belonging to the cluster and then, using a measure of the approximation error (Balsamo et al., 2005), to determine whether to accept (or not) the cluster as a vine row. In this latter case, the TLS line is chosen as the descriptive row. Contrary to ordinary least-squares, the TLS solution is invariant to rotations and independent of reference-frame orientation.

Given a cluster A, the TLS identification of a line

$$r = \{(x, y) \in \mathbb{R}^2 \,|\, ax + by + c = 0\}$$

approximating pixels $(x_i, y_i) \in A, i = 1, \ldots M$, is achieved by minimizing

$$f(r) := \sum_{i=1}^{M} \text{dist}\big((x_i, y_i), r\big)^2 \tag{8}$$

where $\text{dist}\big((x_i, y_i), r\big)$ is the Euclidean distance between line r and pixel (x_i, y_i).

As can be evinced in the example in Figure 10, where the two detected lines obtained from the same cluster of pixels are shown, the row detected by the TSL technique is more reliable than the one achievable using the HT method.

Moreover, by using the TSL solution $\hat{r}$ it is possible to discriminate if the obtained line represents a vine row or an interpolation of a non-vine region. This can be checked by evaluating the root mean square total error

$$Q = \sqrt{\frac{1}{M}f(\hat{r})} \qquad\qquad (9)$$

and by discarding a cluster of interconnected pixels when value Q is greater than a threshold T_Q, that depends on the typical width of the vine. In our case, T_Q has been fixed at 7 pixels (i.e. 0.35 meters).

Even if a set of lines l has already been determined with the HT technique in the previous step, TLS line detection, by approximating in a ℓ_2-norm sense the cloud of pixels belonging to a cluster, is more reliable.

2.4.1 *Vine rows reconstruction*

To complete vineyards detection, vine rows reconstruction is needed to deal with the problem of missing vine plants or those too small for detection. Indeed, at the end of the previous processing steps, several shorter clusters of interconnected pixels can be obtained even if they should be part of a single, more extended, vine row. The vine rows reconstruction process consists in recursively evaluating all possible combinations of two lines detected by the TLS technique, merging them if (1) the difference between their inclination angles θ is lower than a threshold T_θ and (2) the distance between the nearest two ends of the segments is less than a threshold T_d, expressed in pixels. When these two criteria are fulfilled, the merging process is performed by computing a new approximating line with the TLS technique, using the pixels belonging to the union of the original clusters. The combination and checking process is restarted after each merge event, while the whole process is repeated three times with threshold values of

$T_\theta = 1°, 2°$ and $3°$ and $T_d = 20$, 50 and 100 pixels. Vine rows reconstruction process turn out to be very robust to T_θ and T_d threshold values modification: only T_d value need a rough calibration in the case of processing of images with very different spatial resolution.

3. Results

The performance of the developed procedure has been tested on four aerial NIR vineyard images (Figure 11) by comparing the results obtained by applying automatic and manual vine rows detection according to their overlapping rate.

In Figures11.a and 11.b, one complete vineyard and two partial ones, acquired at two different times and conditions of illumination, are represented. A vineyard composed by several parcels, located on a steep hill slope, is shown in Figure 11.c where non-crop vegetation disturbance is particularly present. The aerial image in Figure 11.d shows a vineyard characterized by a high number of missing plants and the nearby presence of urbanization, such as a swimming pool and a tennis court. In all four images, streets or small rural paths are present.

According to the seven different criteria described in Table 1, results have been quantified as the percentage of vine row length (measured in pixels), automatically obtained by the proposed method, compared to the overall manually detected vine row length. Single image and overall mean results are then detailed in Table 2. The *good detection* index was always greater than 90%, with an average value on the four images of 95.13%. Most undesired misclassifications, *missed detection* and *extra detection* errors, are extremely limited with mean indexes of 1,68% and 0,98% respectively. In image "D" (Figure 11.d), *missed detections* reach the highest value of 4,67% due to the presence of significant foliar masses that overlay one vine row. In the other images, this criticality, ascribed to dense shadows projected by some isolated trees, is almost negligible. *Extra detections*, mainly occurring in image "C" (Figure 11.c) and

image "D" (Figure 11.d), are mostly caused by misclassification of hedge and non-vine rows along streets and infrastructures. It should be noted that these rare eventualities produce short and isolated detections, which can be excluded with further verifications.

The proposed method gives positive results also in the case of vineyards with several missing vine plants, as shown on the right side of image "A" (Figure 11.a) or image "D" (Figure 11.d), where spaced out groups of vine plants are correctly recognized as part of the same vine row. Only a few *over-detection* errors occur in the case of not strictly rectilinear, very extended, vine rows, like in the case of the vineyard in image "D" (Figure 11.d).

Even if the percentage of correctly detected vine rows is higher than 95%, performance can be further increased by considering only the main focused vineyard of each image. Indeed, faulty detections (categories 2-7 in table 1) are mainly localized at the border of the image because of partially included vineyards, as shown in Figures 11.a and 11.b, where *over-detection* and *smaller detection* in images "A" and "B" can be completely addressed to the image boundary. The entire image elaboration process takes less than 3 minutes for each image.

4. Conclusions

In this paper, a new method to identify vine rows in aerial images, by determining straight lines that represent vine plants alignments, has been presented. The proposed method is able to process the original acquired image as it is, without any intervention or feedback from the user and requiring a very limited number of parameters for the calibration. The results obtained from the processing of a set of sample images demonstrate the effectiveness of the proposed method, with a very limited error rate, even in the case of images with disturbances (numerous missing plants, non-crop vegetation, dense inter-row grassing, etc.) and different vine row orientations in the same image. Indeed, the elaboration process does not require any spatial periodicity of the crop, and it is robust against irregular illumination,

like in the presence of clouds or steep terrain slopes. The computational effort is affordable and the results can be obtained in a few minutes of elaboration for complex images.

The extracted row and inter-row information could be used for many different tasks in precision agriculture as, e.g., a reference for the autonomous guidance systems of ground vehicles, allowing automatic path computation in the corridors between vineyards. The vineyard detection algorithm will be the basis for further developments of vigour maps, a powerful tool for the creation of field partitioning, allowing differentiated crop management and treatment.

Future work will concern the development of faster algorithms able to operate in real-time conditions. This configuration will allow the use of the detected row line for UAV path planning and flight control to obtain optimal conditions in image acquisition.

References

Arnó, J., Martínez Casasnovas, J.A., Ribes Dasi, M., Rosell, J.R., 2009. Review. Precision viticulture. Research topics, challenges and opportunities in site-specific vineyard management. Spanish Journal of Agricultural Research, 7(4), 779-790. doi:10.5424/sjar/2009074-1092

Balsamo, A., Mana, G., Pennecchi, F. 2005. On the best fit of a line to uncertain observation pairs. Metrologia, 42, 376-382. doi:10.1088/0026-1394/42/5/006

Bobillet, W., da Costa, J. P., Germain, C., Lavialle, O., Grenier, G., 2003. Row detection in high resolution remote sensing images of vine fields. Precision agriculture: Papers from the 4th European Conference on Precision Agriculture, Berlin, Germany, 81-87.

Bramley, R.G.V., Reynolds, A.G., 2010. Precision Viticulture: Managing vineyard variability for improved quality outcomes. Managing Wine Quality, 1, 445–480. doi:10.1533/9781845699284.3.445

Corcoles, J.I., Ortega, J.F., Hernandez, D., Moreno, M.A. 2013. Estimation of leaf area index in onion (Allium cepa L.) using an unmanned aerial vehicle. Biosystems Engineering, 115(1), 31-42. doi:10.1016/j.biosystemseng.2013.02.002

Da Costa, J.P., Michelet, F., Germain, C., Lavialle, O., Grenier, G. 2007. Delineation of vine parcels by segmentation of high resolution remote sensed images. Precision Agriculture, 8, 95-100. doi:10.1007/s11119-007-9031-3

Delenne, C., Rabatel, G., Agurto, V., Deshayes, M., 2006. Vine plot detection in aerial image using Fourier analysis. Proceedings of International Conference on Object-Based Image Analysis (OBIA), 4-5 july 2006, Salzburg (Austria), 6.

Delenne, C., Durrieu, S., Rabatel, G., Deshayes, M., 2010. From pixel to vine parcel: A complete methodology for vineyard delineation and characterization using remote-sensing data. Computers and Electronics in Agriculture, 70(1), 78–83. doi:10.1016/j.compag.2009.09.012

Duda, R., Hart, P., 1972. Use of Hough transformation to detect lines and curves in pictures. Graphics Image Process, 15, 11–15. doi:10.1145/361237.361242

Golub, G.H., Van Loan, C.F., 1980. An analysis of the total least squares. Siam Journal on Numerical Analisys, 17(6), 883-893.

Gonzalez, F., Castro, M.P.G., Narayan, P., Walker, R., Zeller, L., 2011. Development of an Autonomous Unmanned Aerial System to Collect Time-Stamped Samples from the atmosphere and Localize Potential Pathogen Sources. Journal of Field Robotics, 28(6), 961-976. doi:10.1002/rob.20417

Gonzalez, R.C., Woods, R.E., 2002. Digital Image Processing, second edition. Prentice Hall, 520-528.

Hall, A., Louis, J., Lamb, D., 2003. Characterising and mapping vineyard canopy using high-spatial-resolution aerial multispectral images. Computers & Geosciences, 29(7), 813–822. doi:10.1016/S0098-3004(03)00082-7

Hall, A., Lamb, D.W., Holzapfel, B., Louis, J., 2008a. Optical remote sensing applications in viticulture-a review. Australian Journal of Grape and Wine Research, 8(1), 36–47. doi:10.1111/j.1755-0238.2002.tb00209.x

Hall, A., Louis, J.P. and Lamb, D.W., 2008b. Low resolution remotely sensed images of winegrape vineyards map spatial variability in planimetric canopy area instead of LAI. Australian Journal of Grape & Wine Research, 14(1), 9-1. doi:10.1111/j.1755-0238.2008.00002.x

Hall, A., Lamb, D.W., Holzapfel, B. and Louis, J., 2011. Within-season temporal variation in correlations between vineyard canopy and winegrape composition and yield. Precision Agriculture, 12, 103-117. doi:10.1007/s11119-010-9159-4

Kazmiersky, M., Glemas, P., Rousseau, J., Tisseyre, B., 2011. Temporal stability of within-field patterns of NDVI in non irrigated Mediterranean vineyards. Journal international des sciences de la vigne et du vin, 45(2), 61-73.

Kendoul, F., 2012. Survey of advances in guidance, navigation, and control of unmanned rotorcraft systems. Journal of Field Robotics, 29(2), 315–378. doi:10.1002/rob.20414

Kropff, M.J., Wallinga, J., Lotz, L.A.P., 1997. Modelling for precision weed management. Precision Agriculture: Spatial and Temporal Variability of Environmental Quality. Wiley, Chichester (1997), 182–204.

Lamb, D.W., Weedon, M.M. and Bramley, R.G.V., 2004. Using remote sensing to predict grape phenolics and colour at harvest in a Cabernet Sauvignon vineyard: timing observations against vine phenology and optimising image resolution. Australian Journal of Grape & Wine Research, 10, 46-54. doi:10.1111/j.1755-0238.2004.tb00007.x

Markovsky, I., Huffel, S.V., 2007. Overview of total least-squares methods. Signal Processing, 87(10), 2283–2302. doi:10.1016/j.sigpro.2007.04.004

Proffit T., Bramley R., Lamb D., Winter E. (2006) - Precision viticulture - A new era in vineyard management and wine production. Winetitles, Australia, 6-48.

Puletti, N., Perria, R., Storchi, P., 2014. Unsupervised classification of very high remotely sensed images for grapevine rows detection. European Journal of Remote Sensing, 47, 45 – 54. doi: 10.5721/EuJRS20144704

Rabatel, G., Delenne, C., Deshayesa, M., 2008. A non-supervised approach using Gabor filters for vine-plot detection in aerial images. Computers and Electronics in Agriculture, 62(2), 159–168. doi:10.1016/j.compag.2007.12.010

Ranchin, T., Naert, B., Albuisson, M., Boyer, G. and Astrand, P., 2001. An automatic method for vine detection in airborne imagery using wavelet transform and multiresolution analysis. Photogrammetric Engineering and Remote Sensing, 67(1), 91–98.

Recio, J.A., Hermosilla, T., Ruiz, L.A., Palomar, J., 2013. Automated extraction of tree and plot-based parameters in citrus orchards from aerial images. Computers and Electronics in Agriculture, 90, 24–34. doi:10.1016/j.compag.2012.10.005

Rouse, J.W., Haas, R.H., Schell, J.A., Deering, D.W., 1974. Monitoring Vegetation Systems in the Great Plains with ERTS. Third Earth Resources Technology Satellite Symposium, 1, 309-317.

Smit, J.L., Sithole, G., Strever, A.E., 2010.Vine Signal Extraction – an Application of Remote Sensing in Precision Viticulture. South African Journal of Enology & Viticulture, 31(2) 65-74.

Stafford, J.V., 2006. The role of the technology in the emergence and current status of precision agriculture. In: Srinivasan A. (Ed.), Handbook of Precision Agriculture : Principles and Applications, Food Products Press, New York, 19–56

Tisseyre, B., Ojeda, H., Taylor, J., 2007. New technologies and methodologies for site-specific viticulture. Journal International des Sciences de la Vigne et du Vin, 41(2), 63-76.

Zhang, N., Wang, M., Wang, N., 2002. Precision agriculture: a worldwide overview. Computer and electronics in agriculture, 36, 113–132. doi:10.1016/S0168-1699(02)00096-0

Zarco-Tejada, P.J., Berjón, A., López-Lozano, R., Miller, J.R., Martín, P., Cachorro, V., González, M.R., De Frutos, A., 2005. Assessing vineyard condition with hyperspectral indices: Leaf and canopy reflectance simulation in a row-structured discontinuous canopy. Remote Sensing of Environment. 99(3), 271–287. doi:10.1016/j.rse.2005.09.002

Zhang, C., Kovacs, J.M., 2012. The application of small unmanned aerial systems for precision agriculture: a review. Precision Agriculture, 13(6), 693–712. doi:10.1007/s11119-012-9274-5

Notation table

DN	Digital Number
FFT	Fast Fourier Transform
HPS	Hough Parameters Space
HT	Hough Transform
LARS	Low Altitude Remote Sensing
NDVI	Normalized Difference Vegetation Index
NIR	Near InfraRed
PV	Precision Viticulture
ROI	Region of Interest
TLS	Total Least Squares
UAV	Unmanned Aerial Vehicle
UAS	Unmanned Aerial System
UGV	Unmanned Ground Vehicle

A	Group of interconnected pixels (Cluster) after first clustering
$\mathcal{A}$	Set of all detected clusters A in the processed image after first clustering
B	Group of interconnected pixels (Cluster) after vine rows detaching
$\mathcal{B}$	Set of all detected clusters B in the processed image after vine rows detaching
d	Diagonal of the image (in pixel)
$f(r)$	Sum of the Euclidean distances between the line r and the pixels $(x_i, y_i) \in A, i = 1, \dots M$
H	Maximum value of $S(\cdot)$
$I(\cdot)$	Projection of $S_H(\cdot)$ values on the θ axis
$J(\cdot)$	Normalized inertia moment of the intensity distribution histogram of $\Omega_{x,y}$
l_w	Sliding window size, used to select $\Omega_{x,y}$
l	Rectilinear line detected with HT
$\mathfrak{L}_A$	Set of all rectilinear lines detected in a cluster A
N	Number of pixels in $\Omega_{x,y}$
$n(v)$	Number of pixels characterized by DN intensity value v
n_w	Number of times that the neighbouring window segmentation process is applied to a

	pixel (x, y)
m	Number of θ values in $\mathcal{O}_A$
M	Number of pixels in A
$\mathcal{O}_A$	Set of most recurrent values of θ in cluster A
p	Generic pixel ϵA
$\mathcal{P}_{\hat{\theta}}$	Set of most recurrent values of ρ, for a fixed $\hat{\theta}$ value, in cluster A
q	Point (ρ, θ) in HPS
Q	Root mean square total error of $\hat{r}$
r	Line detected with TLS
$s(p)$	Curve in HPS, derived from p
$S(\cdot)$	Function defined as the number of curves $s(p)$ that pass through the neighbourhood $Z(\cdot)$
s_i	Projection of pixel i in the HPS
s_w	Sliding window moving step, used to select $\Omega_{x,y}$
T_Q	Threshold of Q
T_J	Threshold of $J_{x,y}$
T_θ	Threshold of difference between inclination angles θ between two lines

T_d	Threshold of distance between the nearest two ends of the segments
w_ρ	Reciprocal of ε_ρ
w_θ	Reciprocal of ε_θ
$Z(\cdot)$	Neighbourhood in the HPS
ε_ρ	Width of Z on the ρ axis
ε_θ	Width of Z on the θ axis
θ	Clockwise inclination angle with respect to the vertical axis
$\hat{\theta}$	Most recurrent values of θ
Θ	Set of θ values used for HPS tasselation
$\mu_{x,y}$	Mean of the intensity distribution histogram of $\Omega_{x,y}$
ρ	Distance from the origin of the image
P	Set of ρ values used for HPS tasselation
$\Omega_{x,y}$	Subset of pixels centred in pixel (x,y)
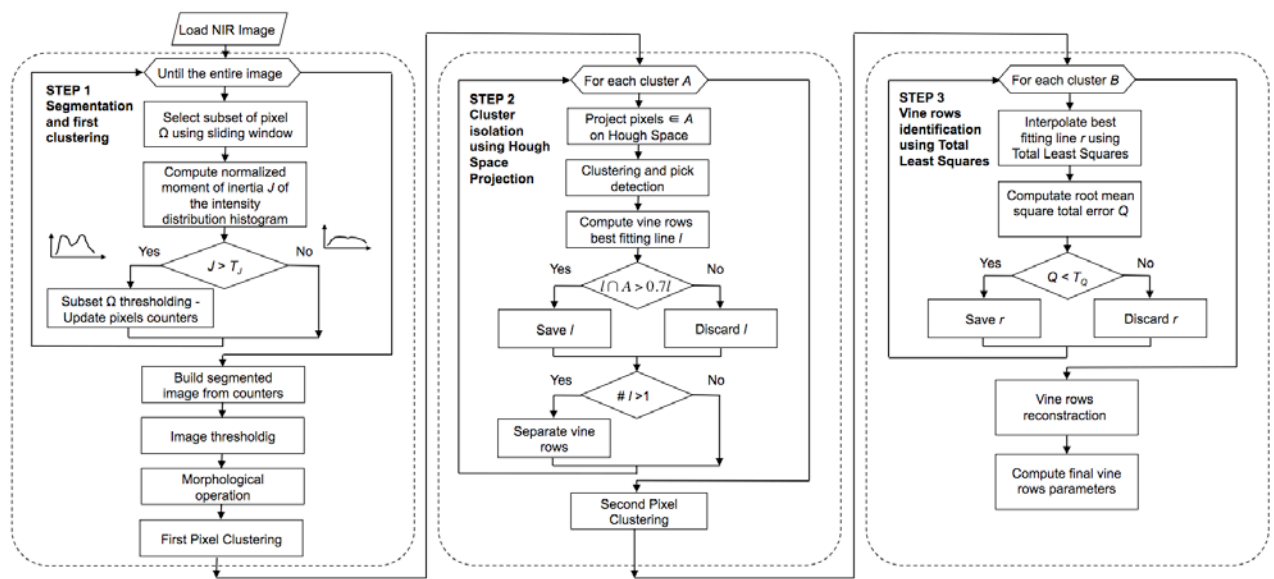
Figure 1. Flow chart of the developed image processing settlement
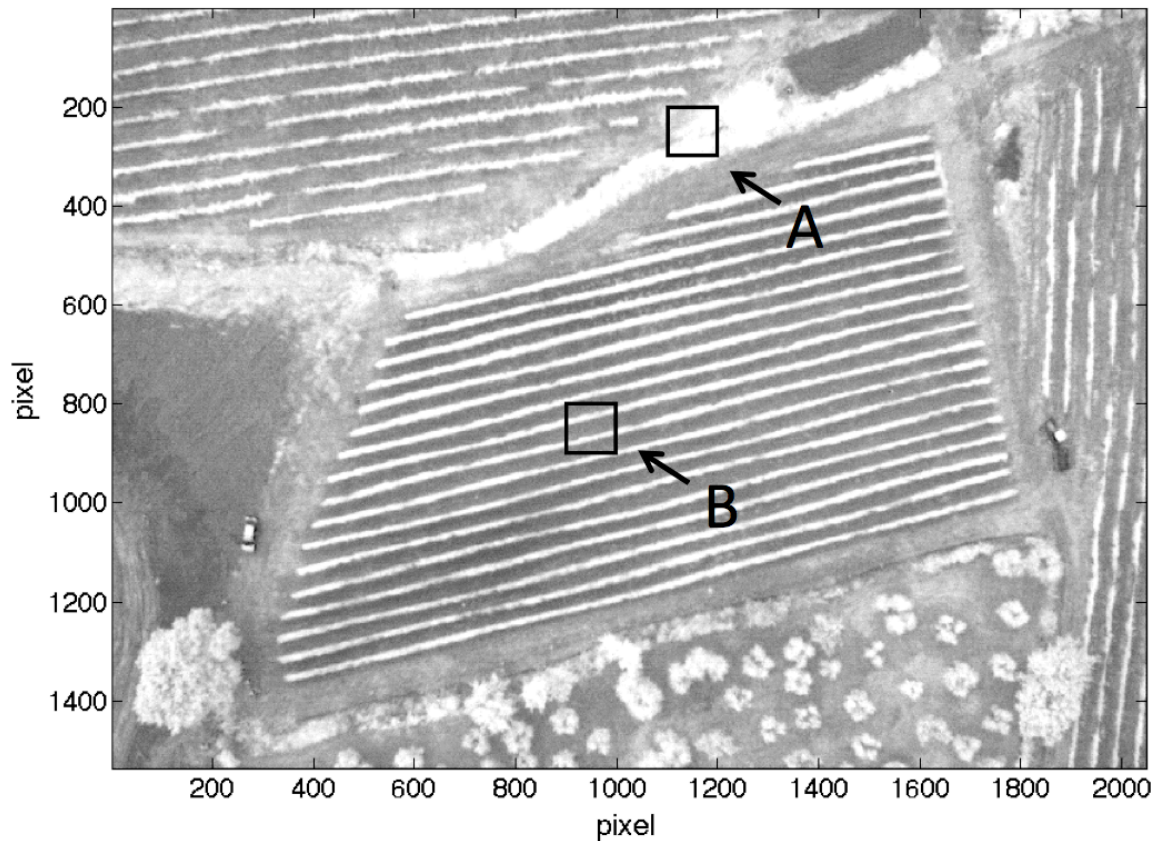
Figure 2. Near Infrared Gray-scale aerial image of a vineyard (Montalcino-Siena, Italy). Squares delimit subsets of pixel Ω_{x_a,y_a} and Ω_{x_b,y_b}, centered on pixels $a = (1150,250)$ and $b = (950,850)$ respectively, during the dynamic local neighboring window segmentation process
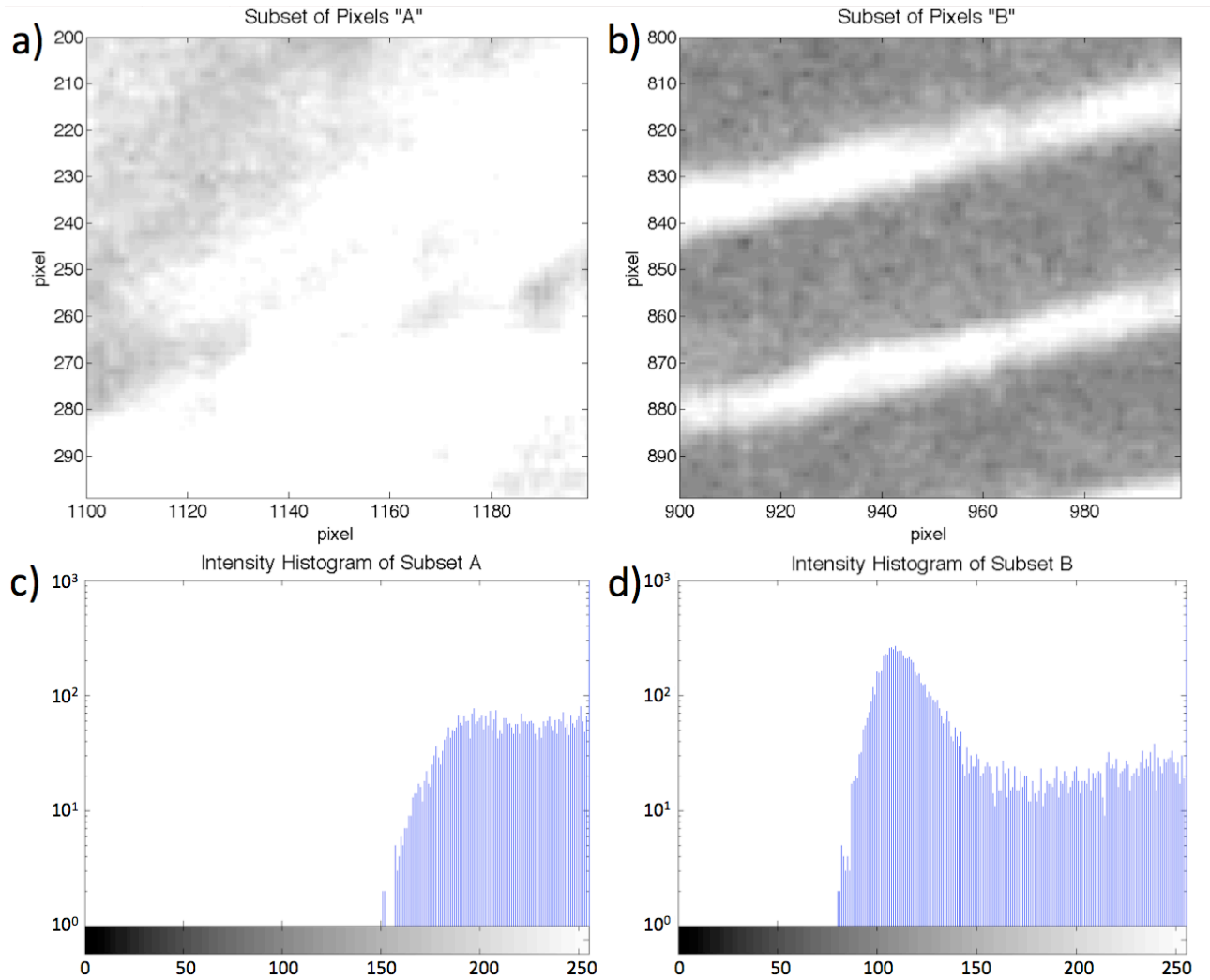
Figure 3. A detail of subsets of pixel Ω_{x_a,y_a} (a) and Ω_{x_b,y_b} (b) extracted from the near infrared image in Figure 2. Intensity distribution histogram of subsets of pixel Ω_{x_a,y_a} (c) and Ω_{x_b,y_b} (d). Normalized moment $J(\Omega_{x_a,y_a})$ of subset Ω_{x_a,y_a} is $4.01 \cdot 10^2$, lower than $T_J = 5 \cdot 10^2$, while $J(\Omega_{x_b,y_b})$ of subset Ω_{x_b,y_b} is $3.01 \cdot 10^3$.
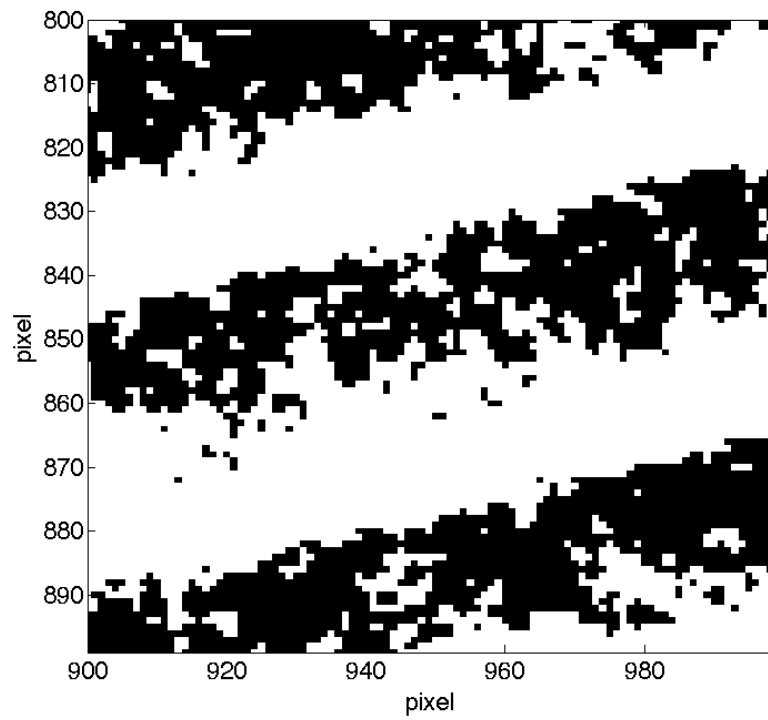
Figure 4. Threshold applied to subset of pixel Ω_{x_b,y_b}: pixel with values higher than $\mu_{x_a,y_a} = 138.6$ are represented in white color
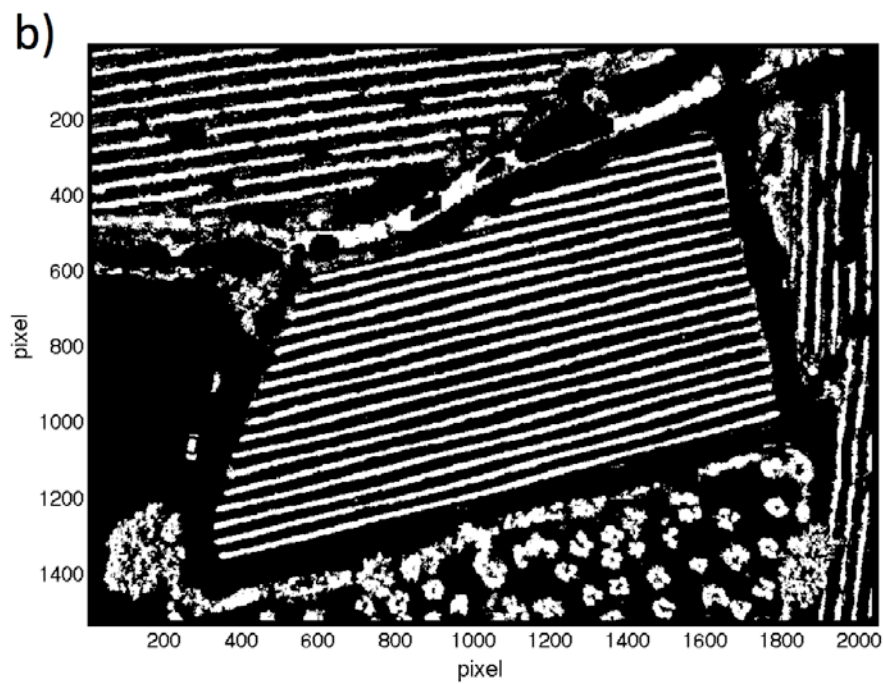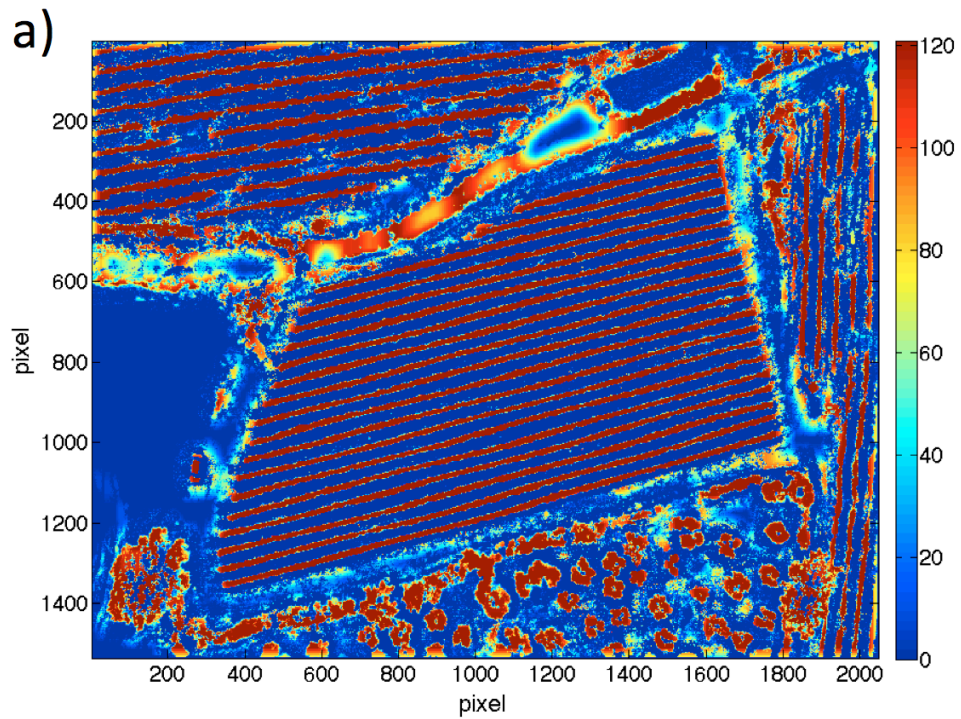
Figure 5. Resulting image (a) results of the dynamic local neighboring window segmentation. Pixel values range from 0 to 121: higher values characterize pixels that more likely belong to vine rows. Image thresholding (b): pixels with values higher than 96.8 are white colored
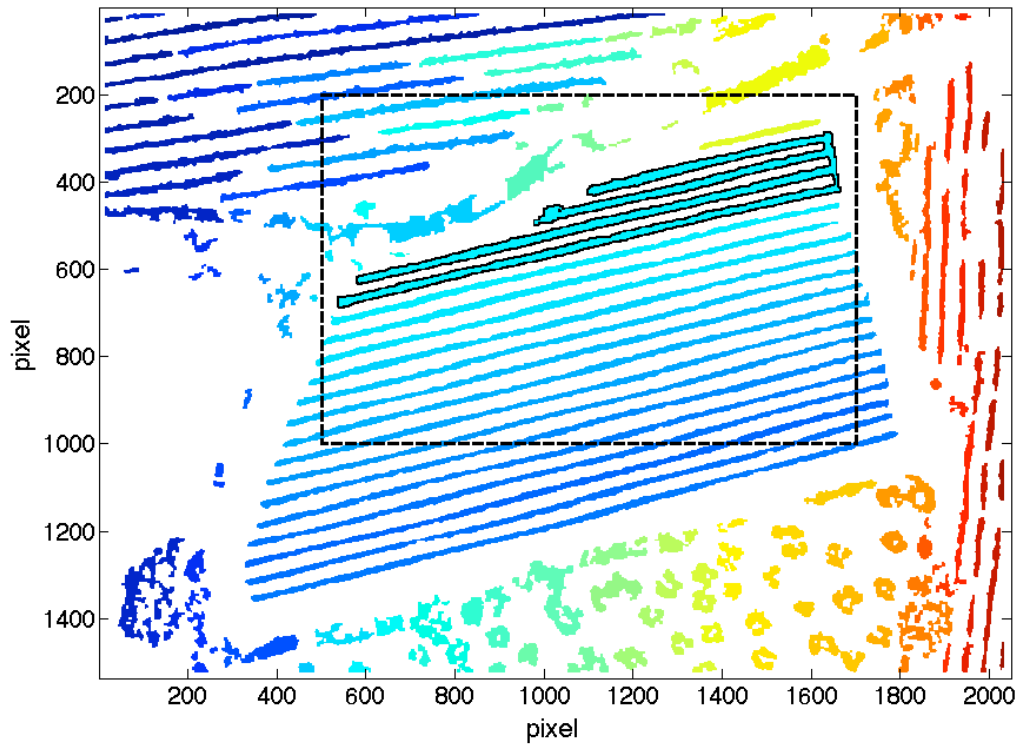
Figure 6. Resulting image after the sequentially application of Open, Close and Remove Small Objects operations. Each group of interconnected pixels is represented with a different color. Dashed rectangle defines a zoomed area reported in Figure 7, to further analyze the highlighted bordered cluster $\mathcal{C}_{64}$.
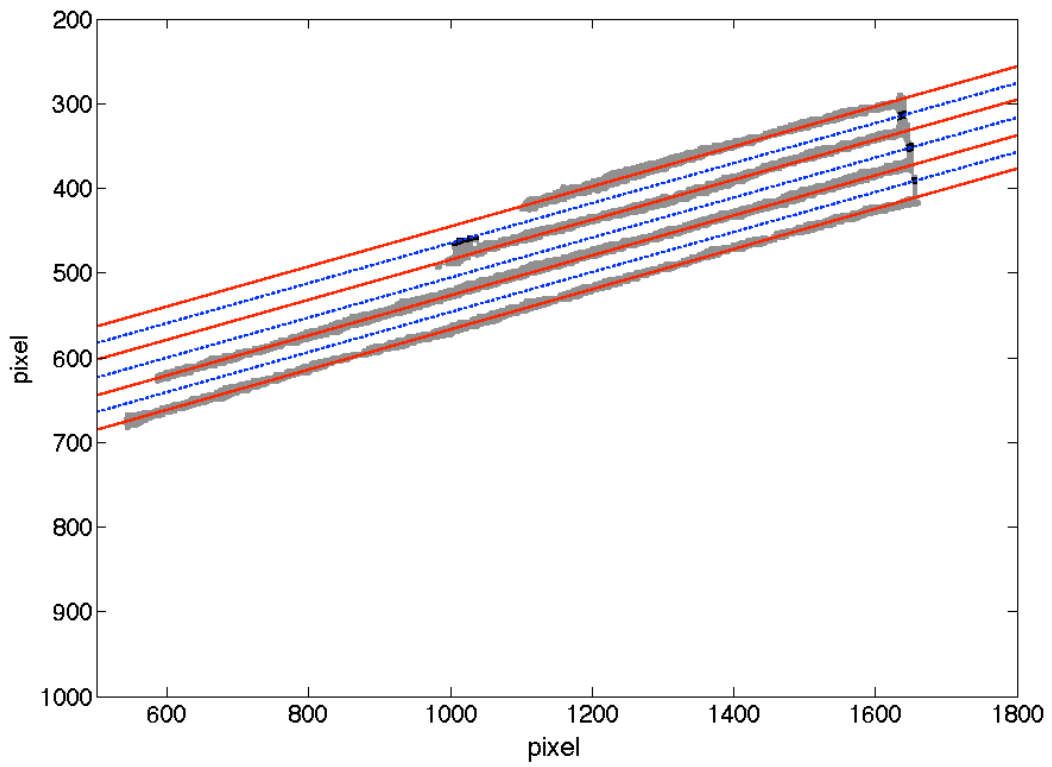
Figure 7. Detail of interconnected pixels highlighted in Figure 6, constituted by four vine rows wrongly interconnected (gray pixels). Best fitting lines (solid red) are detected with Hough transform technique. Cutting lines (dotted blue) have been used to select pixels to be removed (black pixels) in order to disconnect each single vine rows.
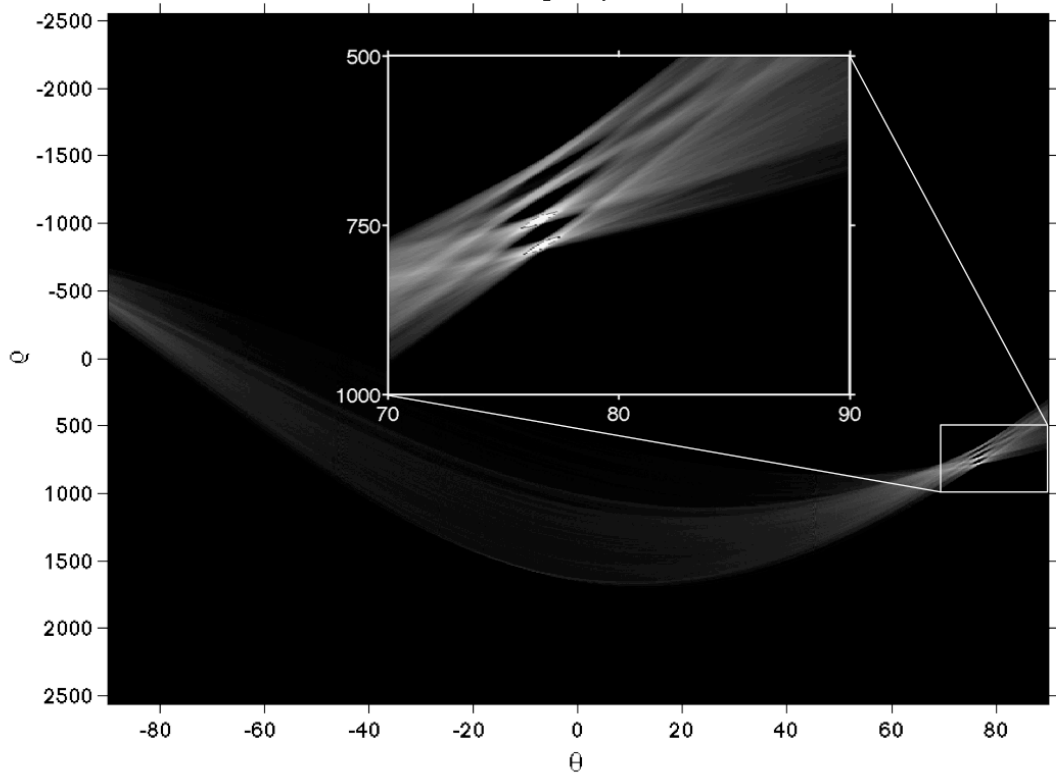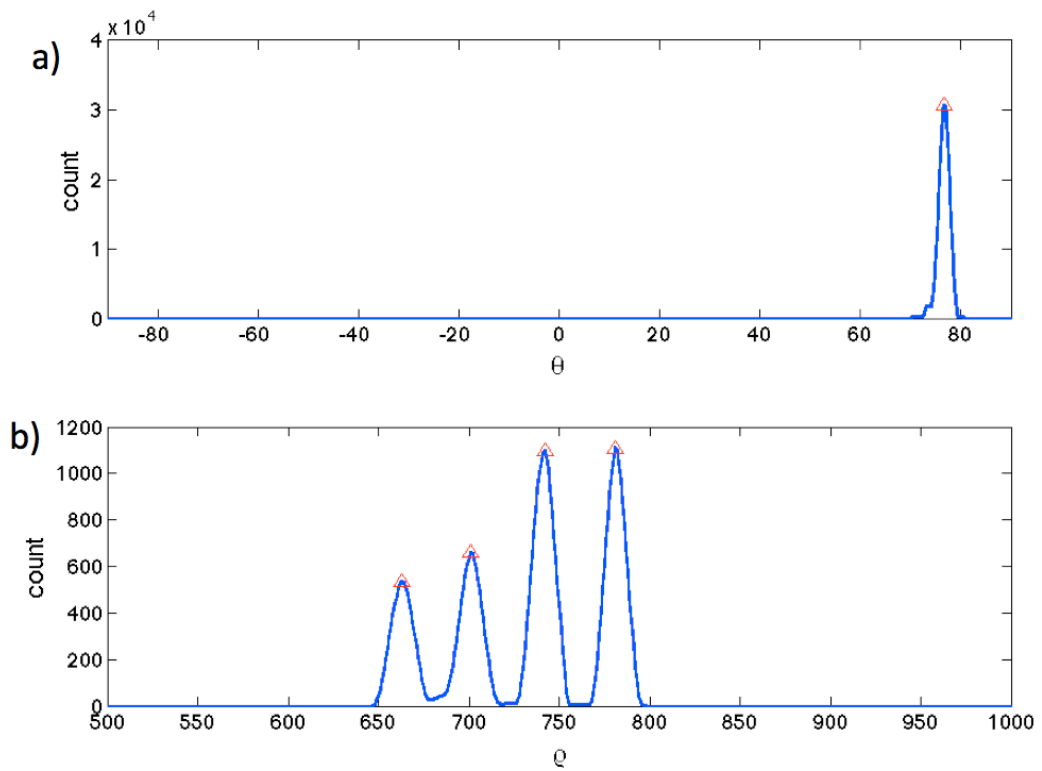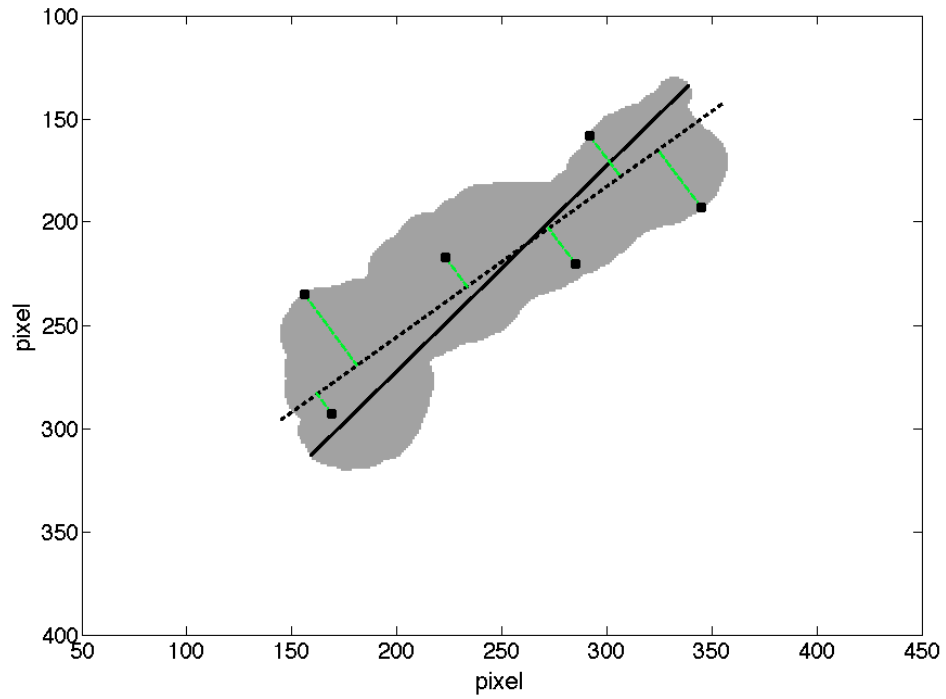
Figure 8. Function $S\big((\rho, \theta), \mathcal{C}_{64}\big)$ in HPS. Brighter pixels represent higher values.

Figure 9. (a) Smoothed function $I(\theta, A_{64})$, with the detected global maximum (Δ) in $\theta_1 = 76.7°$, and (b) function $S\big((\rho, \theta_1), A_{64}\big)$. Detected maxima are marked with Δ.

Figure 10. Lines detected using HT (black solid) and TLS (black dotted) have been obtained from the same cluster of pixels A. Some true (Euclidean) distances between black dotted line and the pixels $(x_i, y_i) \in A$ (black •), used in TLS method, are represented with green dashed lines as explicative example.
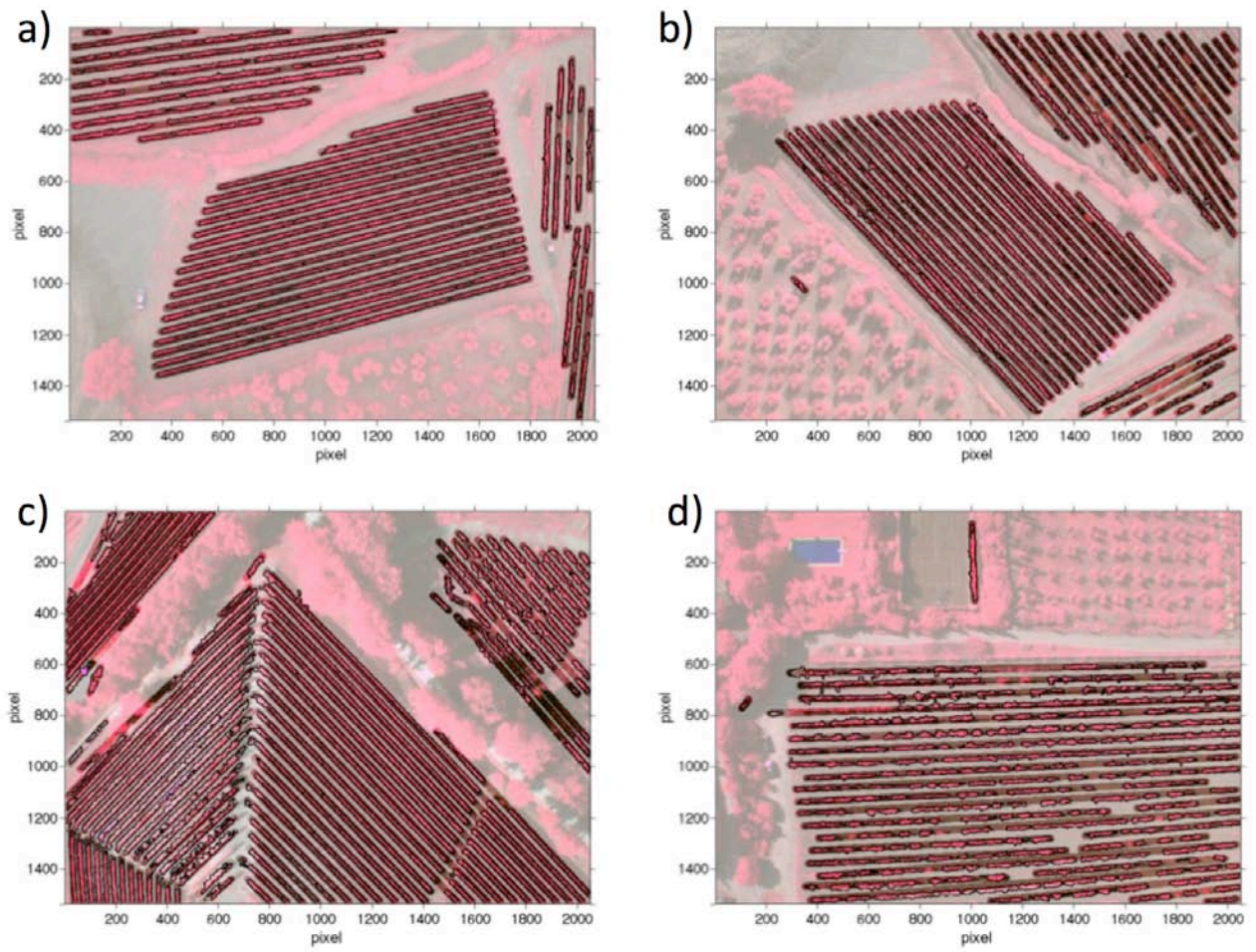
Figure 11. Final vine rows detection applied to four aerial NIR image. Regions of the image classified as vine rows are highlighted, fading background pixels.

Table 1

Criteria for vine rows detection results evaluation

Case	Meaning
1. *Good detection*	detection of the vine rows is correct, matching the manual evaluation
2. *Missed detection*	vine row totally not detected
3. *Extra detection*	non-vine row region wrongly classified as vine row
4. *Over-detection*	unique real vine row classified in several automatically detected vine rows
4. *Under-detection*	several real vine rows classified in a unique automatically detected vine row
5. *Smaller detection*	boundaries of automatically detected vine row smaller than real one
7. *Larger detection*	boundaries of automatically detected vine row greater than real one

Table 2.

Results of vine rows detection (expressed as percentage of vine row length with respect to the overall manually detected vine rows length) applied to four airborne NIR images. Overall mean performances of the developed image processing settlement are reported in the fifth column.

	Image A	Image B	Image C	Image D	Mean
1. Good detection	96,06%	96,30%	97,85%	90,32%	95,13%
2. Missed detection	0,00%	1,54%	0,50%	4,67%	1,68%
3. Extra detection	0,00%	0,15%	2,75%	1,01%	0,98%
4. Over-detection	0,46%	0,21%	0,68%	1,03%	0,59%
4. Under-detection	0,00%	0,00%	0,85%	0,00%	0,21%
5. Smaller detection	3,48%	1,96%	0,97%	3,98%	2,59%
7. Larger detection	0,00%	0,05%	5,59%	0,46%	1,53%