Original papers

Automatic green fruit counting in orange trees using digital images

Walter Maldonado Jr. [a,*], José Carlos Barbosa [b]

[a] Department of Exact Sciences, College of Agricultural and Veterinarian Sciences, Sao Paulo State University (FCAV/UNESP), Brazil's Citrus Defense Fund (FUNDECITRUS), Via de Acesso Prof. Paulo Donato Castellane, s/n, Jaboticabal, SP 14884-900, Brazil
[b] Department of Exact Sciences, College of Agricultural and Veterinarian Sciences, Sao Paulo State University (FCAV/UNESP), Via de Acesso Prof. Paulo Donato Castellane, s/n, Jaboticabal, SP 14884-900, Brazil

ABSTRACT

Yield estimation is an important factor in a production process planning. In the case of citrus crops, can be useful in industrial management and as guidance for farmers, showing a decisive role in the product market strategies and cultivation practices. Several techniques are being studied for estimating citrus crop yield. On the basis of the known correlation between the number of visible fruits in a digital image and the total of fruits present in an orange tree, we developed a method for green fruit feature extraction with a combination of the techniques of color model conversion, thresholding, histogram equalization, spatial filtering with Laplace and Sobel operators and Gaussian blur. In addition, we built and tested an algorithm to recognize and count them, with detection rates of false-positives of 3% in images acquired in good conditions. It is possible to estimate the mean number of visible fruits in the trees within a tolerated error of 5% with up to 46 images and taking approximately 8 min without any human interaction.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Among the technology resources that are part of the modern production processes aiming to improve management strategies and crop yield, stand out the Global Positioning System (GPS), Geographic Information Systems (GIS), Remote Sensing (RS), Variable Rate Technology (VRT), yield mapping and advances in sensors and information technology. These improvements have enabled farmers to see the field in a wide and detailed manner, helping resource planning and diminishing its waste, taking into account the variability observed in the field (Annamalai et al., 2004).

Citrus crops are known to show an alternated behavior regarding its yield, as an outcome of plant's intrinsic energy storage mechanisms. These mechanisms generate a large temporal variation with respect to fruit yield and fruit quality (Isagi et al., 1997; Noguchi et al., 2003). Consequently, the known prediction methods do not show satisfactory results, becoming more inefficient as the prediction period increases (Sakai et al., 2008). For that reason, an efficient and practical sampling method is important, in order to be applied regularly for a better follow up on the plants.

Knowledge and follow up on each plant individually have becoming each time more important, since this alternated behavior dynamic has a proved endogenous and individual source

(Sakai et al., 2008). This makes it essential to a reasonable sampling plan not to be destructive. Moreover, traditionally, procedures as liming and fertilizing take into account soil and climate conditions, leaving aside plants individual characteristics, treating them uniformly. With the rise of precision agriculture, more specific knowledge on the plants individually is gaining practical advantages, being possible through detailed maps and equipment capable of vary the application rates with spatial precision of centimeters.

Important attempts on citrus yield prediction have been made. Sakai et al. (2008) demonstrated that prediction was successful for one-year forecasting in a season based on a very short ecological time series, with relative root-mean-square error (RRMSE) of 0.678 and correlation coefficient (CC) of 0.734. Greater prediction time increased the RRMSE to 0.993 and decreased the CC to 0.116, indication long-term unpredictability. A prediction system in real time showed positive results (Annamalai et al., 2004) but only when the fruits are in a more advanced maturation state because they are segmented in the images by their color. In this stage, the application of techniques to improve yield is not possible and there is not enough time for market and industrial planning. The system purposed by Triboni and Barbosa (2004) has the same disadvantage, counting only the ripe fruits in the trees. There is also an *Android* application that showed efficient for fruit counting in the same conditions, with estimations up to 90% between the real yield values and the ones estimated for individual trees (Gong et al., 2013).

* Corresponding author.
  E-mail address: walter@agroestat.com.br (W. Maldonado Jr.).

Leaves and fruits were analyzed for differences in the reflected electromagnetic spectrum through hyperspectral images (Annamalai et al., 2004). Aerial images acquired with help of an aerial model remotely piloted were used for fruit recognition (MacArthur et al., 2006) and embedded sensors were used for image acquisition and posterior processing for fruit recognition and counting (Annamalai, 2004). Digital image processing methods as the *watershed* transform were used for the segmentation of clustered fruits (Chinchuluun et al., 2006), neural networks were used in aerial hyperspectral images (Ye et al., 2006) and the number of new leaves and floral flushes were studied, with help of vegetation indices, for relation with the number of fruits through a least-squares method for yield estimation, without positive results (Ye et al., 2007). More recently, the performance of the TBVI (*Two-Band Vegetation Index*) was reassessed (Ye et al., 2008), as the green fruit detection through hyperspectral images (Okamoto and Lee, 2009). The use of digital images for yield prediction in citrus is promising, given that there is a functional relationship between the number of visible fruits in the images and the total number of fruits in the plants (Triboni and Barbosa, 2004).

However, those methods have several limitations, requiring highly specialized labor or expensive equipment and generally a low efficiency (Gong et al., 2013). In addition, some are not affordable by small farmers. All the cited methods need improvements, considering that even in hyperspectral images the number of false-positives is large, precluding its application in estimation processes. Further studies are needed (Okamoto and Lee, 2009) mainly because even with good results, low cost and easy application for prediction with ripe fruits (Gong et al., 2013), it is still necessary to enable the yield estimation with more time before harvesting.

Texture analysis through digital image processing techniques enables the generation of images that simulate a bas-relief effect, as in bronze sculptures of famous artists as Donatello. The bas-relief representation gives us a more enhanced notion of depths and texture. Computerized techniques of bas-relief generation are been studied and evaluated (Belhumeur et al., 1999; Cignoni et al., 1997; Ji et al., 2014). Regardless the limitation of the digital image to be a visual representation in two dimensions, the main concept of these images generation is the simulation of the effect present in the sculpture of a solid material.

Obtaining those bas-relief representations from common digital images involves several digital image processing procedures, as border detection, image enhancement, spatial filtering and fusion, among others (Durand and Dorsey, 2002; Paris and Durand, 2006; Wang, 2011; Weiss, 2006). Border detection is one of the fundamental problems of digital image processing and computer vision, being obtained by operators well known like the Laplacian, Sobel, Robert's cross operator, Prewitt's and Canny's (Wang, 2011). A bas-relief representation of a common digital image would be result of a combination from the above methods, for example.

Therefore, bas-relief representations are important for the extraction of texture features of the objects in the image, what becomes clear regarding the spherical shapes, that shows a region of high brightness and other with low brightness, considering a single source of illumination. The concept of texture is of extreme importance in images, enabling the definition of measures of properties like smoothness, rugosity and regularity. Hence, texture can be defined as the difference on the intensity pattern between pixels in a neighborhood (Gonzalez and Woods, 2000).

After the texture analysis for feature extraction the next step is the recognition and classification of those features. In this case, we are interested in the classification of a region of the image as a fruit or not. For that task, there are several reports of the use of machine learning techniques as, for example, neural networks, also in the recognition of orange fruits (Regunathan and Lee, 2005). Another important method of machine learning is the Support Vector Machines (SVMs). In a similar way of the neural networks, the SVMs require a learning step and, after its training they become a pattern classifier. This type of method has a wide application in image processing, with positive results in the recognition and classification of apples, applied in images treated with Otsu's thresholding (Otsu, 1975). The authors verified segmentation errors ranging from 3 to 25% using a linear SVM and errors lower than 2% using an adjustable method for the classifier.

SVMs are maximum margin classifiers. Instead of the association of probability distribution models to the training vectors, SVMs try to split the different classes finding the adequate immediate borders between them. For this purpose, they build separation hyperplanes in the separation region between those classes (Keuchel et al., 2003).

This recognition method has large application in image processing applied to agriculture. It is reported as the best method of recognition and classification of oil palm areas in aerial images, showing a good performance even with a small training set (Li et al., 2015). It is also reported as one of the best methods in the recognition and classification of different varieties of rapeseed (Kurtulmus and Ünal, 2015), and presented 100% of correct classifications of mango fruits quality, overcoming all other tested methods (Sa'ad et al., 2015). There is also reports of its use in flower cultivation, showing good results in the detection of diseased rose leaves (Nagasai and Rani, 2015).

Early yield estimation of orange crops could enable the modification of cultivation practices in order to increase fruit size and quality, provide better planning for market and industrial process and make possible the generation of precise yield maps to be used in VRT. The objective of this work was to build and test an algorithm for automatic counting of green fruits in digital images of orange trees, on the basis of texture analysis and SVMs.

2. Materials and methods

The relationship between the number of visible fruits in the images and the number of fruits counted by the algorithm was evaluated through linear regression analysis. The 1328 images were obtained with collaboration of Citrosuco Company in the 2011 crop year. All images were taken at 2 m of distance between the end of the tree's canopy and the camera, in field conditions, in different hours of the day and several climate and lightning conditions. Regular digital cameras were used (Sony DSC-W530) and images were taken in a 2592×1944 pixels spatial resolution (approximately 5 Megapixels), with flash always on and without optical or digital zoom.

At the moment the image was taken, it was identified regarding the plants' variety and age. Tests were performed in orange trees from four varieties (Hamlin, Natal, Pêra, Valência) in three groups of ages (from 3 to 5 years old, from 6 to 10 years old or older than 11 years). Each plant was photographed twice (one image from each accessible side of the canopy).

The automation of manual tasks is always very attractive because brings up the potential of time saving and precision at its execution. Consequently, the performance of a proposed method should be tested against a well known and reliable method. In this case, the automatic method was compared to the visual manual counting, where all the images were counted by human inspectors in digital monitors.

The manual counting on the images was helped by the development of a software, in which the user just needed to click over the fruits to increase an automatic counter. The results of this step enabled the study of the relationship between the number of visible fruits in the images and the outcome of the algorithm

developed. The automatic method was applied to all of the collected images from the 2011 crop, with exception only for those cases where the image quality was very low, impairing the algorithm performance significantly or even forbidding the manual counting process.

All the procedures used in this work were developed on the basis of the implementation of C++ methods, making use of the Open Source Computer Vision library (OpenCV) (Bradski, 2015). Feature extraction was constructed on the basis of texture analysis, through the generation of bas-relief representations from common two-dimension digital images.

2.1. Bas-relief representations

Bas-relief representations were obtained by the combination of the techniques of color model conversion, thresholding, histogram equalization, spatial filtering with Laplace and Sobel operators and Gaussian blur (Fig. 1).

Considering the facts that the images were obtained in open areas, susceptible to variations on the illumination conditions, and the plant canopies are irregular, it is indispensable the adjustment of brightness. Among the methods used for this purpose, the most common are the log transformation and gray levels histogram equalization (Gonzalez and Woods, 2000; Kurtulmus et al., 2011; Savvides and Kumar, 2003), that have been applied to the value layer of the HSV color model.

2.2. The Sobel operators

Sobel operators are gradient operators responsible for edges detection in the horizontal and vertical directions, with the advantage of providing, for once, the differentiation and smoothing effects. Once derivation increases the noise of the image, the smoothing effect is one attractive feature of these operators

Fig. 1. Flowchart of the purposed algorithm for the generation of bas-relief representations.

(Gonzalez and Woods, 2000). In a region of an image with size of 3×3 given by

$$\begin{bmatrix} z_1 & z_2 & z_3 \\ z_4 & z_5 & z_6 \\ z_7 & z_8 & z_9 \end{bmatrix}$$

when Sobel operators are applied through spatial filtering by the masks

$$\begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$$

and

$$\begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$$

it yields the respective derivatives based on those masks:

$$G_x = (z_7 + 2z_8 + z_9) - (z_1 + 2z_2 + z_3)$$

and

$$G_y = (z_3 + 2z_6 + z_9) - (z_1 + 2z_4 + z_7)$$

in which values of z are the levels of gray of the overlapping pixels by the masks in any position of the image. The calculations yield a single value in the center of the mask being applied, and a new value should be calculated for each pixel of the image, considering its edge limitations together with the size of the mask. The result after the calculations for all the pixels is a new image, highlighting the edges found.

2.3. The Laplacian operator

The Laplacian of a two-dimensional $f(x, y)$ is a second order derivative given by:

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}.$$

As in the case of the Sobel operators, this equation can be implemented in the digital form in several ways. For the case of a region of size 3×3, the most frequently used masks are:

$$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix}.$$

and

$$\begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix},$$

where it defines the respective derivative equations from those masks:

$$\nabla^2 f = 4z_5 - (z_2 + z_4 + z_6 + z_8)$$

and

$$\nabla^2 f = 8z_5 - (z_1 + z_2 + z_3 + z_4 + z_6 + z_7 + z_8 + z_9),$$

where again the values of z are the gray levels of the gray pixels overlapped by the mask. The definition of the digital laplacian requires that the associated coefficient with the central pixel is positive and the other external coefficients are negative. Besides that, once this operator is a derivative, the summation of its coefficients should be null. The result of the application of the mask is null in
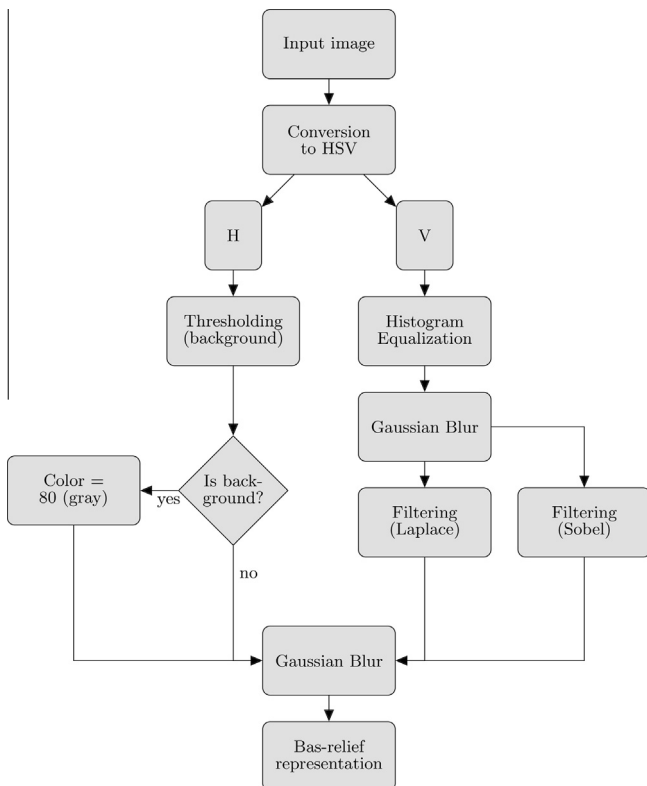
every case in which the not null neighboring coefficient presents the same value as the central pixel.

However, even responding to transitions on pixels' intensity, the laplacian has little use in this form for edges detection. It happens because, for being a second order derivative filter, the operator is extremely sensible to noise in an unacceptable way. It is also important to bring up that it produces double borders, incapacitating it to detect its directions. Therefore, the laplacian shows a secondary role as a detector to determine if a pixel is in the light or dark side of a border (Gonzalez and Woods, 2000).

This operator shows a more general use of localization of the edges through its zero cross property, based on the convolution of the image by the laplacian of a two-dimensional Gaussian function given by:

$$\nabla^2(x,y) = -\frac{1}{\pi\sigma^4}\left[1 - \frac{x^2+y^2}{2\sigma^2}\right]e^{-\frac{x^2+y^2}{2\sigma^2}}$$

where σ is the standard deviation (Marr and Hildreth, 1980).

2.4. Fruit feature extraction

The most complex and challenging step of digital image processing is the segmentation of the desired characteristic. This gets even more difficult when the image acquisition occurs in non standardized conditions (Zhao et al., 2015). It was possible to highlight the fruits in the images of orange trees canopies through a combination of the techniques of color model conversion, thresholding, Gaussian blur and spatial filtering with Laplace and Sobel masks. The images took in the field showed a large amount of strange elements that needed to be removed before the procedures of recognition and counting of the fruits, as the soil, the sky and the weeds. In addition, the fruits are easily confused with the leaves of the tree.

The prevailing color of green fruits in orange trees is very similar to the leaves, what ruled out this parameter of the segmentation attempts. For consequence, the H (Hue) and S (Saturation) components of the color model were not used in the pre-processing for feature extraction. On the other hand, the V (Value) component brings very important information about brightness, being for that reason the selected one for fruit feature extraction.

Histogram equalization is largely used in digital image processing applied to agriculture. Its properties of contrast stretching are used in the detection of diseases in cotton leaves (Warne and Ganorkar, 2015), for example. In the case of the citrus images, it showed good results in the standardization of the illumination from the images taken in very different conditions.

The next step is when texture analysis takes place, by means of the use of the Laplace and Sobel filters. Those techniques consist basically on the convolution of the image by predefined kernels (or masks). Hence, a pixel or a group of pixels with neighbors presenting similar levels of gray represent a smooth texture region (low frequency) and pixels with neighbors presenting levels of gray very different represent a harsh texture region (high frequency). In the case of the citrus canopies, the fruits can be classified as smoother regions (low frequencies) than the leaves, what enabled its segmentation. The Sobel high-pass filter was used in the horizontal direction because of the illumination pattern of the fruits (which have an approximately spherical shape), characterized by a high value of brightness in the upper half and a low value in the lower visible half.

The soil was removed through thresholding applied to the Hue component of the HSV color model. The numerical scale of this component makes it possible to identify pixels by its colors. In the case of the soil, it has a very different color value from the rest of the image, being safely removed based on its reddish coloring.

Part of the sky was also removed based on very high values of the brightness component in the images.

The bas-relief representation is now built as a combination of the above techniques. As noticeable in Fig. 2, the fruits are highlighted as in bas-relief sculptures, with its upper half brighter and its lower half darker. The automatic generation of these representations enables the use of method of recognition and interpretation for the number of visible fruits estimation, defining the first step of the development of the automatic system for citrus yield estimation.

2.5. Fruit counting

2.5.1. Spatial masks

The bas-relief representation highlights the spherical shapes in a particular manner. When there is a single illumination source from a superior position, it determines a well defined brightness pattern: the visible part of the upper semi-sphere showing a high brightness value and the visible part of the lower semi-sphere a low value. This enabled the definition of two spatial masks, capable of detecting the fruits in the pre-processed images, from several dimensions. The main principle is the ratio between the upper and the lower parts which, in the case of the fruits, outcomes higher values than in the rest of the image in most of cases, acting as a preliminary classifier.

2.5.2. Support Vector Machines (SVMs)

After the first step of identification, the use of more robust methods for fruit recognition becomes necessary, considering that
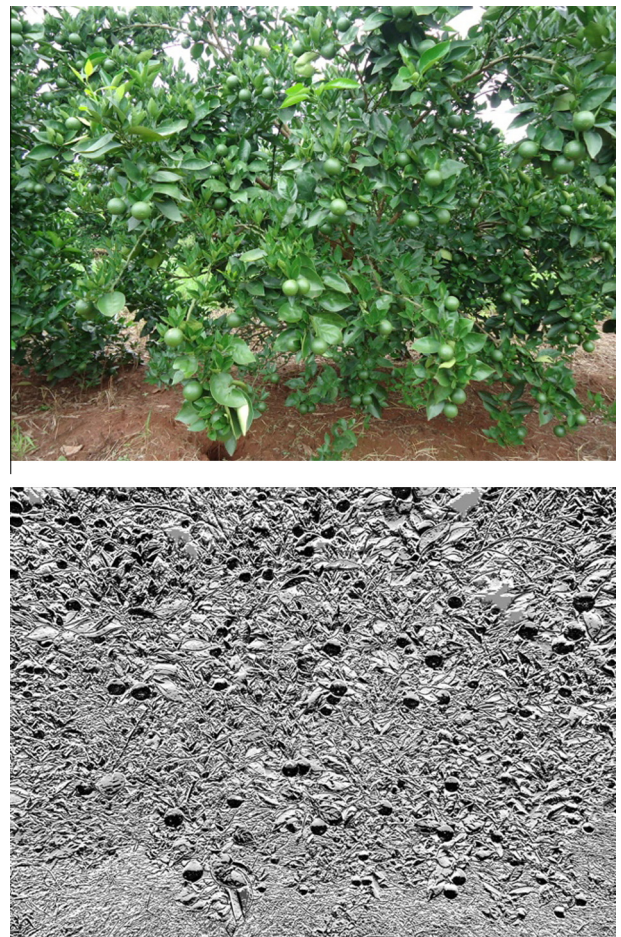

Fig. 2. Bas-relief representation of an orange tree highlighting the green fruits on the basis of texture and brightness analysis.

the orange tree images present countless shapes, positions and textures of leaves, that can be confounded with fruits, even in the bas-relief representations. Therefore, SVMs where chosen for the next step, the fruits recognition.

Hard margin SVMs define linear limits from linearly separable data. Let T be a training set with n observations, $\mathbf{x}_i \in X$ its respective labels $y_i \in Y$ in which X is the data space and $Y = \{-1, +1\}$. T is linearly separable if it is possible to split the data into classes $+1$ and -1 through a hyperplane given by:

$$f(x) = \mathbf{w} \cdot \mathbf{x} + b = 0$$

where $\mathbf{w} \cdot \mathbf{x}$ is the scalar product of the vectors, $\mathbf{w} \in X$ is the vector normal to the described hyperplane, $\frac{b}{\|\mathbf{w}\|}$ is the distance from the hyperplane to the origin, with $b \in \mathbb{R}$. This enables the calculation of a signal function $g(\mathbf{x}) = sgn(f(\mathbf{x}))$ to be used for obtaining the classifications (Lorena and Carvalho, 2007):

$$g(\mathbf{x}) = sgn(f(\mathbf{x})) = \begin{cases} +1 \text{ if } \mathbf{w} \cdot \mathbf{x} + b > 0 \\ -1 \text{ if } \mathbf{w} \cdot \mathbf{x} + b < 0. \end{cases}$$

The limits of the separation margins of the data should be calculated given some restrictions which assure that there is no training data between the margins, what is the origin of the name of this SVM as a hard margin SVM. The final step for the definition of these limits for classification falls in a quadratic optimization problem, which has a wide and established mathematical theory (Schölkopf and Smola, 2002). Since the objective function being minimized is convex and the points that satisfy the restrictions form a convex set, the problem has a global minimum. This type of problem can be solved with the introduction of a Lagrangian function, considering the restrictions to the objective function associated to parameters named Lagrange multipliers α_i, as exposed by:

$$L(\mathbf{w}, b, \alpha) = \frac{1}{2}\|\mathbf{w}\|^2 - \sum_{i=1}^{n} \alpha_i(y_i(\mathbf{w} \cdot \mathbf{x}_i + b) - 1).$$

2.5.3. The Hough transform

The Hough transform (Hough, 1962) brings up an efficient approach to the recognition of lines in images. Considering the point (x_i, y_i) and a line $y_i = ax_i + b$, infinite lines pass through (x_i, y_i) satisfying the line's equation, for different values of a and b. However, it is possible to rewrite this line as $b = -x_i a + y$ and, as the values of x_i and y_i are known, its new representation in the parameter space (axis x and y replaced by a and b) is given that each point (x_i, y_i) describes a corresponding line. When three or more lines intercept in a single point (a', b'), the corresponding points (x_i, y_i) belong to a same line $y_i = ax_i + b$.

Now, considering the equation of a circle

$$(x - c_1)^2 + (y - c_2)^2 = c_3^2,$$

the Hough transform application occurs in an analogous way. The basic difference is the presence of three parameters (c_1, c_2 e c_3), that will produce a three-dimensional parameter space. The procedure then is to increment c_1 and c_2 and find the c_3 that solves the circle equation, followed by the update of a corresponding accumulator (Gonzalez and Woods, 2000).

2.5.4. Algorithm strategy for fruit recognition

Five different classifiers were defined for fruit detection because of the several conditions in which images were taken, as for example different canopy sizes and different fruit sizes. Either the preliminary classifiers (based on spatial masks) or the SVMs were subdivided (Fig. 3). This enabled a more efficient training of the
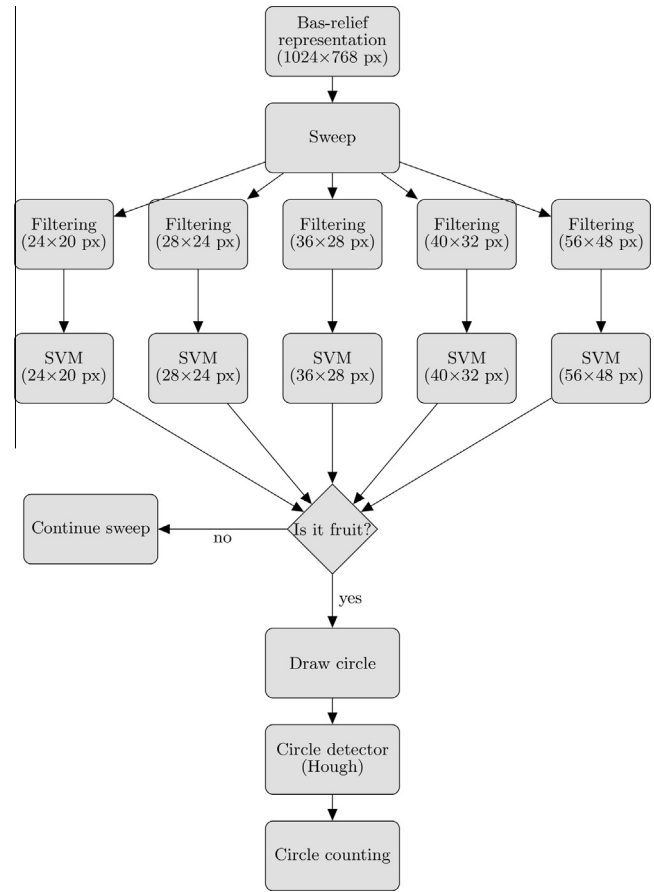

Fig. 3. Flowchart of the fruit recognition algorithm in bas-relief representation images of orange trees.

SVMs, increasing significantly its detection power without increasing the number of false-positives.

The algorithm testing showed that the lower the distance between the camera and the plant canopy, the greater the number
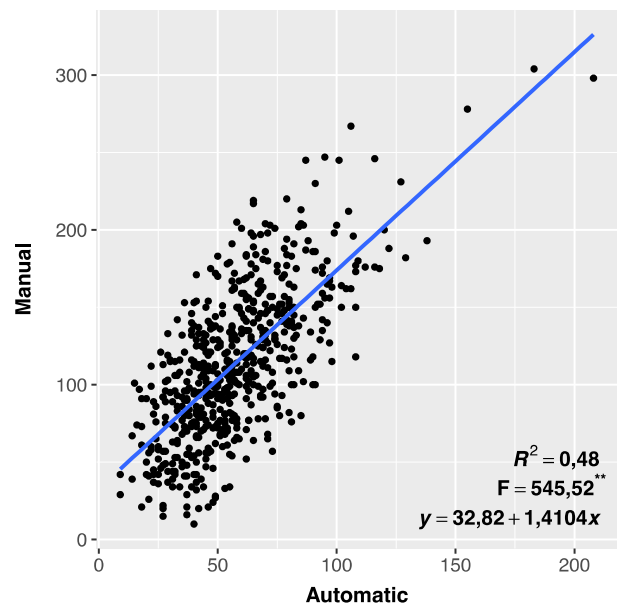

Fig. 4. Linear regression for the relationship between the number of fruits obtained by the algorithm and the number of fruits found by the inspectors in the manual counts of images from 591 orange trees (1182 images). Data from the 2011 crop year.

of false-positives detected. Those cases occurred mainly in the detection patterns of lower dimensions, confusing leaves (larger in these cases) with fruits. For that reason, a rule was implemented to disregard the counting results of those patterns depending on the amount of fruits detected by the patterns of higher sizes.

Lastly, all the detected fruits were stored in a binary image as circles with radius proportional to the size of the pattern of the detection. Consequently, the fruits recognized by more than one pattern size resulted in overlapping white circles in that image, avoiding multiple counting of a single fruit. This binary image was the subjected to the Hough circles detector (Yuen et al., 1989) and the final number of fruits is equal to the number of detected circles. The method was tested on a computer with Unix-based operating system, 2.4 GHz Intel Core i5 processor and 8 GB 1600 MHz DDR3 RAM.

2.5.5. Number of required images for visual fruit number estimation

After the model fitting, the proportional deviations from the estimates were calculated, in order to find out the needed number of images to estimate the mean number of visible fruits per tree in a population, given a tolerated error rate D in percentage of the mean. The calculations were made based on sample size determination and given by Thompson (1992):

$$n = \frac{t^2 s^2}{D^2}$$

where t is the value of Student's t with $n-1$ degrees of freedom at the $\alpha = 5\%$ level of probability, s^2 is the variance of the proportional deviations and D is the tolerated error for the mean estimation.
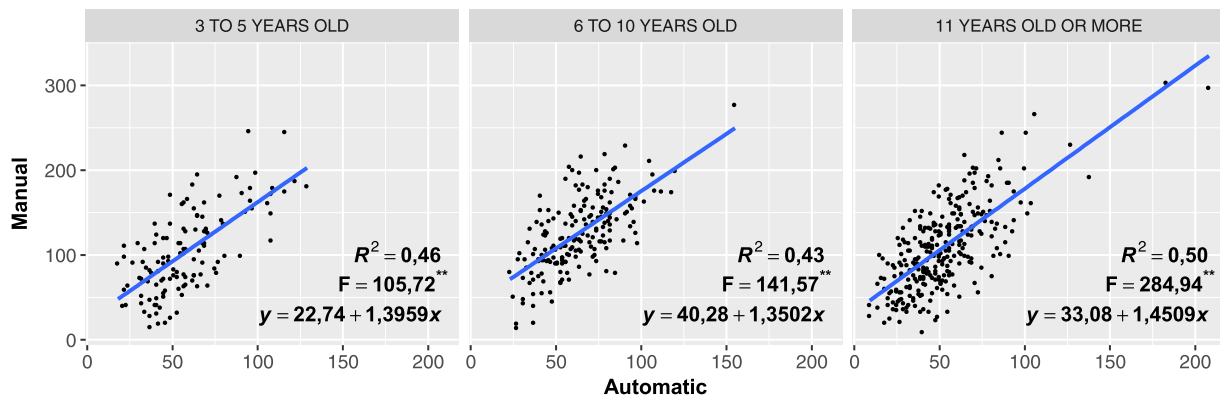

Fig. 5. Linear regression for the relationship between the number of fruits counted by the algorithm and those visually counted by the inspectors in images of 591 plants (1182 images) stratified by age groups. Data from the 2011 crop year.
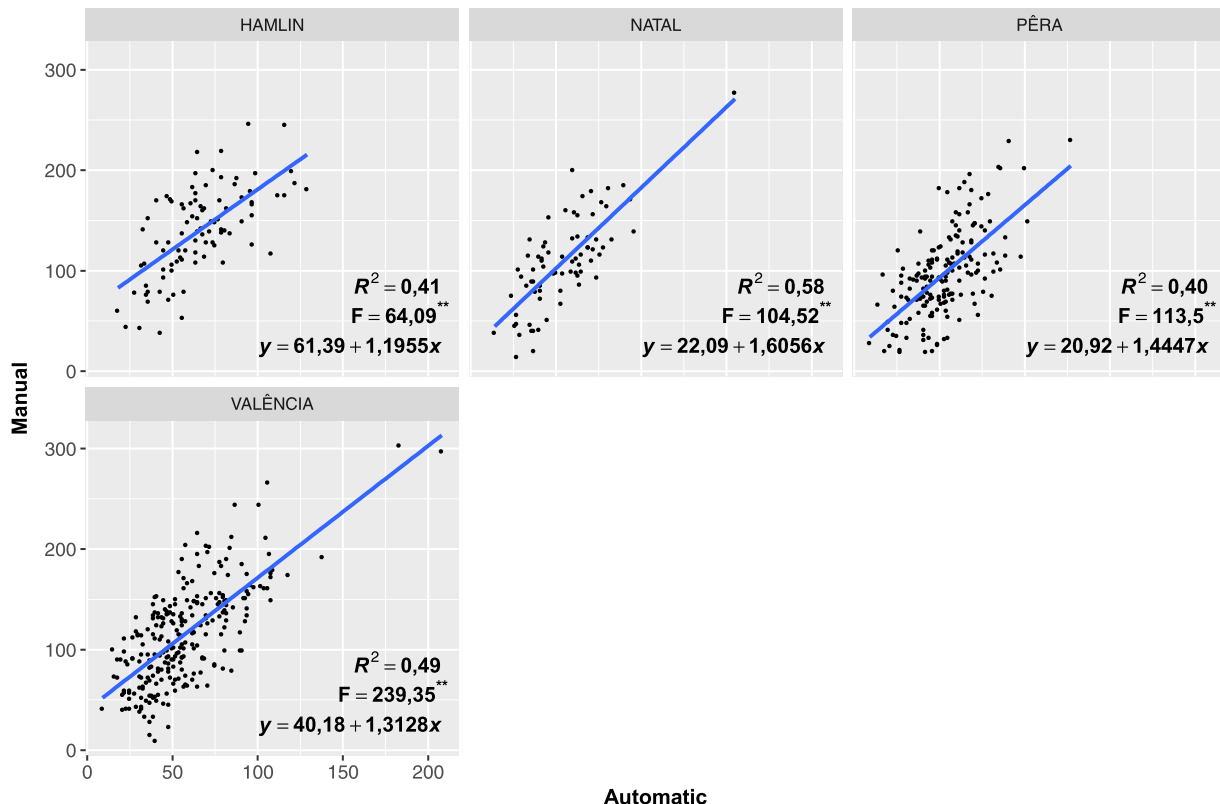

Fig. 6. Linear regression for the relationship between the number of fruits counted by the algorithm and those visually counted by the inspectors in images of 591 plants (1182 images) stratified by varieties. Data from the 2011 crop year.
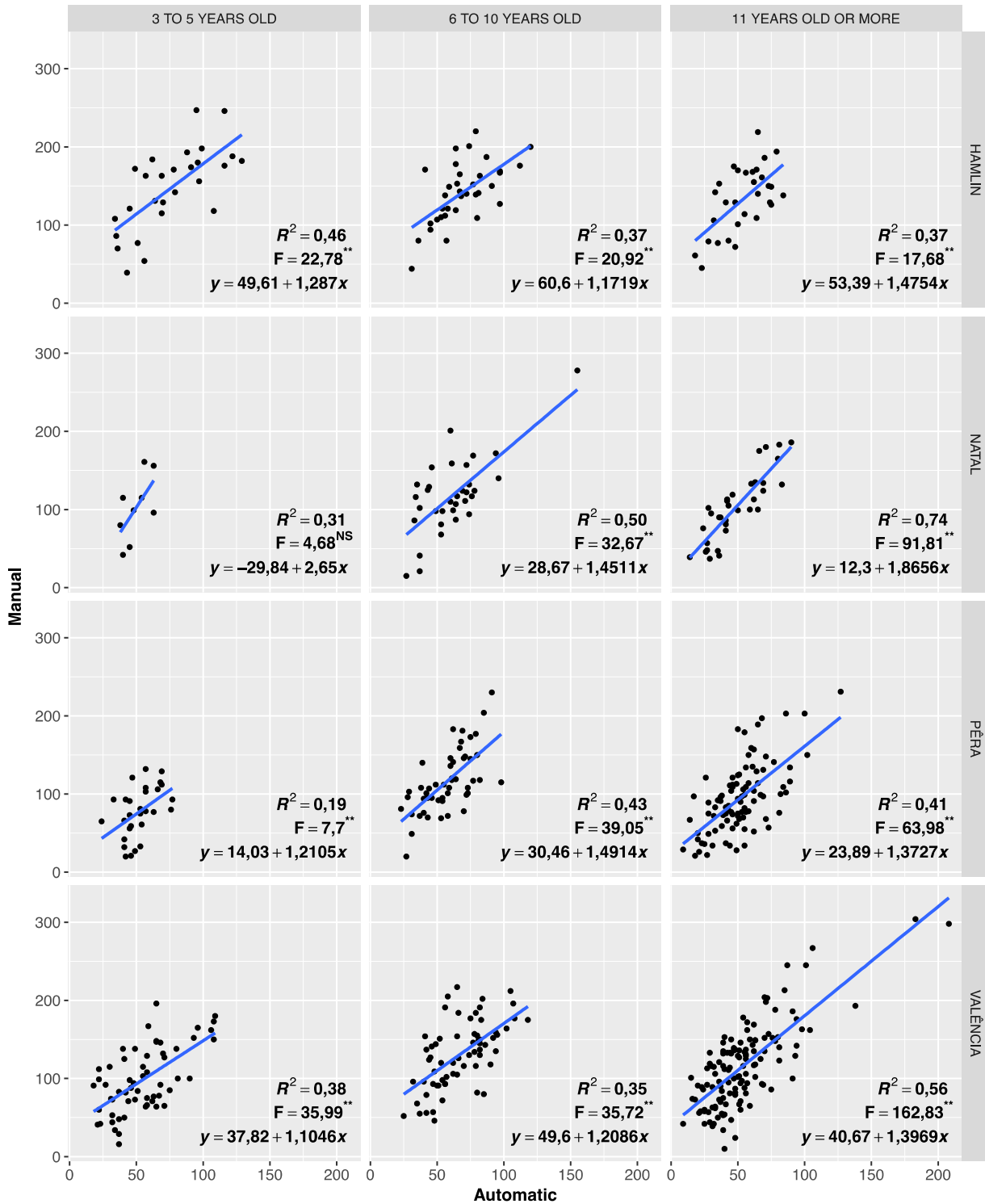
Fig. 7. Linear regression for the relationship between the number of fruits counted by the algorithm and those visually counted by the inspectors in images of 591 plants (1182 images) stratified by the combinations between varieties and age groups. Data from the 2011 crop year.

3. Results and discussion

3.1. Comparison between automatic and manual counting methods

Fig. 4 shows the overall performance of the method presenting the results of the linear regression. This model fit was obtained from 1182 images of 591 trees. Two hundred images of 100 plants were removed from the initial set for being in bad conditions for the automatic counting and even for the manual evaluations too. Each dot in the chart represents the summation of the counts of two images taken from the same tree found by the algorithm in the x axis and the summation of the visual manual counts found by the inspectors in the y axis.

Among the factors that could interfere with the algorithm performance, major importance was given to the age and the variety of the plants. An age effect was expected due to the size of the

plant, in which the distinctness of the fruits inside the canopy, the number of fruits occluded by the leaves and the fruit size could be affected, even with a fixed distance from the plant canopy. Regarding varieties, some would present more fruits only recognizable by human inspectors inside the canopy than others. Those age and variety effects can be observed in Fig. 5 and in Fig. 6.

Therefore, still regarding Fig. 5, it becomes possible to identify a difference for age group 1 (3–5 years old). Basically, the intercept of this model was less than the other age groups due to the number of fruits not recognized by the algorithm but counted by the inspectors manually. These fruits would be the partially occluded, that are not evident enough to be accounted by the algorithm but are recognizable by the human capabilities.

With respect to the varieties, the stratification is also interesting mainly because of the structural difference between the plants. For that reason, all the combinations between varieties and groups of age were studied individually seeking more precise models, as shown in Fig. 7. These results show that the linear regression is capable of representing the relationship between automatic and manual methods in 11 of the 12 combinations. The exception is for Natal with age between 3 and 5 years, probably because of the low number of points available for model fitting and testing. Values of the coefficient of determination for the fit (R^2) ranged from 0.19 (Pêra from 3 to 5 years old) to 0.74 (Natal with age equal or older than 11 years) (Fig. 7).

3.2. Methods benchmarking

Both methods were benchmarked for comparison regarding the time spent on its execution. In the manual approach, the most of the time was spent in the recognition and identification step by the inspectors, making the effect of computational resources negligible (Table 1).

The automatic method provided a reduction on time spent of approximately 85% in comparison with the manual method and a reduction of about 90% in comparison with the method of Kurtulmus et al. (2011). The time savings here presented would be extremely important for practical purposes, enabling the use of a greater number of images for the estimation of the total number of fruits per plant.

3.3. Algorithm performance

The algorithm was not able to detect all fruits in the images. Moreover, there were detections of parts of the image that were not fruits, being called false-positives. These errors are mostly a consequence of the uneven light conditions, image quality, absent or not enough flash light in the moment of image taking, distance between the camera and the plant canopy, irregular plant canopies and variable fruits sizes. The measured algorithm performance is available in Table 2.

The results of the performance evaluation of the algorithm points to the fact that the percentage of correct identifications were constant and the amount of false-positives were low. False-positive rates greater than the observed here were reported by Kurtulmus et al. (2011) in their method, reaching up to 40.3%, which confirms a better performance of the here proposed method. It also is worth mentioning that our results take into account a final set of 1182 images, after the removal of only the low quality images, i.e. the ones with indistinguishable fruits or blurred by camera motion. This contrasts with the cited work that reports results only in two images taken in different conditions: sunny and shadow side.

As a concern with the effect of the quality of the image (dark image, blurred image, direct light on the lens, etc.), they were grouped into four distinct groups: very good, good, regular and bad. The detection rates kept around 50% when the image quality

Table 1
Descriptive statistics of the time spent in seconds by the manual and automatic fruit counting methods in 1182 orange trees images.

Statistic	Manual	Automatic
Mean	64.9	9.79
Median	63.2	9.80
Minimum	49.8	8.50
Maximum	94.7	11.8
Standard deviation	13.1	0.31
Coefficient of variation	20%	3%

Table 2
Algorithm performance evaluation for the detection of green fruits in images of orange trees through the percentage of correct identifications ± the standard error of the mean grouped by varieties and age groups in 1182 individual images. Data from the 2011 crop year.

Stratum	Images	Correct identifications (%)	False-positives (%)
General	1182	49.03 ± 1.42	14.68 ± 0.52
Variety			
Hamlin	187	44.27 ± 1.29	11.24 ± 0.89
Natal	154	46.48 ± 1.78	14.22 ± 1.29
Pêra	355	52.21 ± 1.94	17.75 ± 0.91
Valência	486	49.32 ± 2.99	13.96 ± 0.90
Age group			
3–5 years old	248	53.53 ± 2.03	17.27 ± 1.02
6–10 years old	374	47.30 ± 1.02	12.66 ± 0.78
11 years old or more	560	48.22 ± 2.73	14.86 ± 0.84

Table 3
Algorithm performance evaluation for the detection of green fruits in images of orange trees through the percentage of correct identifications ± the standard error of the mean grouped by image quality in 1182 individual images. Data from the 2011 crop year.

Stratum	Images	Correct identifications (%)	False-positives (%)
General	1182	49,03 ± 1,42	14,68 ± 0,52
Image quality			
Very good	281	46,63 ± 0,74	3,20 ± 0,21
Good	349	44,74 ± 0,93	8,34 ± 0,41
Regular	464	51,06 ± 1,54	20,90 ± 0,71
Bad	88	62,91 ± 16,45	42,40 ± 3,70

was not bad, being greater in the last case. The false-positive rates increased considerably in the direction of the decrease of image quality, together with the standard deviation of the mean (Table 3), clarifying the effect of performance decrease as lower the image quality gets. It is also worth bringing attention to the 'very good' group, made by 281 images and with a false-positive detection rate of approximately 3%, the lowest ever reported for green fruit counting in orange trees to the best of our knowledge. The algorithm was capable of detecting fruits with diameter ranging from 20 to 45 pixels in the images, which were resized to 1024 × 768 pixels.

An example of the algorithm application in a 'very good' image is available in Fig. 8. It becomes clear the capacity of recognition of the fruits in the plant's canopy. Nevertheless, improvements are still needed aiming to increase the correct identifications rate, what would be necessary to achieve a better precision in the estimation of total fruits per plant.

Finally, we calculated the number of required images to perform an estimation of the visible fruits per plant using the algorithm. Since the use of specific models for each combination of varieties with age groups would outcome more precise results, the obtained values for a precision of 5, 10, 15 and 20% are available in Table 4.

Fig. 8. Green fruit detection in an image of a Valência orange tree of age group 1 (from 3 to 5 years old).

Table 4
Number of required images for mean visible fruits estimation in orange trees for the tolerated error of 5, 10, 15 e 20% for each combination between varieties and age groups.

Stratum	Images	Sample units			
		5%	10%	15%	20%
General	1182	278	70	30	18
Varieties × age groups					
Hamlin × 3–5 years old	54	140	36	16	8
Hamlin × 6–10 years old	70	46	12	6	2
Hamlin × 11 or more years old	60	48	12	6	2
Natal × 3–5 years old	18	208	52	24	14
Natal × 6–10 years old	66	386	96	42	24
Natal × 11 or more years old	66	62	16	6	4
Pêra × 3–5 years old	58	824	206	92	52
Pêra × 6–10 years old	102	112	28	12	6
Pêra × 11 or more years old	186	254	64	28	16
Valência × 3–5 years old	116	338	84	38	22
Valência × 6–10 years old	130	88	22	10	6
Valência × 11 or more years old	256	458	114	50	28

In the case of Hamlin from 6 to 10 years old and Hamlin with 11 years or mode are presented the best results of the algorithm. In these cases, it is possible to estimate the mean number of visible fruits with a precision of 5% with only 48 images. It means that, if 100 samples are taken with 48 images each from an orange trees population of the Hamlin variety with 6 years or more, in 95 of them $t_{\alpha=5\%}$ the value of the mean number of visible fruits would be within the interval of the population mean ± 5%. It is important

to point out that this procedure would take approximately 470 s (7.8 min) to complete, without any human interaction.

4. Conclusions

We developed a method for green fruit feature extraction in digital images of orange trees. In addition, we built and tested an algorithm to recognize and count them, with detection rates of false-positives of 3% in images acquired in good conditions. It is possible to estimate the mean number of visible fruits in the trees within a tolerated error of 5% with up to 46 images and taking approximately 8 min.

Acknowledgments

References

Annamalai, P., 2004. Citrus Yield Mapping System Using Machine Vision. University of Florida.

Annamalai, P., Lee, W.S., Burks, T.F., 2004. Color vision system for estimating citrus yield in real-time. In: ASAE/CSAE Annual International Meeting. ASAE/CSAE, Ontario.

Belhumeur, P.N., Kriegman, D.J., Yuille, A.L., 1999. The bas-relief ambiguity. Int. J. Comput. Vis. 35, 33–44. http://dx.doi.org/10.1023/A:1008154927611.

Bradski, G., 2015. Open Source Computer Vision Library [WWW Document]. GitHub. <https://github.com/Itseez/opencv> (accessed 5.31.15).

Chinchuluun, R., Lee, W.S., Burks, T.F., 2006. Machine vision-based Citrus yield mapping system. In: Proc. Fla. State Hort. Soc., pp. 142–147.

Cignoni, P., Montani, C., Scopigno, R., 1997. Computer-assisted generation of bas- and high-reliefs. J. Graph. Tools 2, 15–28. http://dx.doi.org/10.1080/10867651.1997.10487476.

Durand, F., Dorsey, J., 2002. Fast bilateral filtering for the display of high-dynamic-range images. In: Proceedings of the 29th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH '02. ACM, New York, NY, USA, pp. 257–266. http://dx.doi.org/10.1145/566570.566574.

Gong, A., Yu, J., He, Y., Qiu, Z., 2013. Citrus yield estimation based on images processed by an Android mobile phone. Biosyst. Eng. 115, 162–170. http://dx.doi.org/10.1016/j.biosystemseng.2013.03.009.

Gonzalez, R.C., Woods, R.E., 2000. Processamento De Imagens Digitais. Edgard Blucher, São Paulo.

Hough, 1962. Method and means for recognizing complex patterns.

Isagi, Y., Sugimura, K., Sumida, A., Ito, H., 1997. How does masting happen and synchronize? J. Theor. Biol. 187, 231–239. http://dx.doi.org/10.1006/jtbi.1997.0442.

Ji, Z., Ma, W., Sun, X., 2014. Bas-relief modeling from normal images with intuitive styles. Vis. Comput. Graph. IEEE Trans. On 20, 675–685.

Keuchel, J., Naumann, S., Heiler, M., Siegmund, A., 2003. Automatic land cover analysis for Tenerife by supervised classification using remotely sensed data. Rem. Sens. Environ. 86, 530–541. http://dx.doi.org/10.1016/S0034-4257(03)00130-5.

Kurtulmus, F., Lee, W.S., Vardar, A., 2011. Green citrus detection using "eigenfruit", color and circular Gabor texture features under natural outdoor conditions. Comput. Electron. Agric. 78, 140–149. http://dx.doi.org/10.1016/j.compag.2011.07.001.

Kurtulmus, F., Ünal, H., 2015. Discriminating rapeseed varieties using computer vision and machine learning. Expert Syst. Appl. 42, 1880–1891. http://dx.doi.org/10.1016/j.eswa.2014.10.003.

Li, L., Dong, J., Njeudeng Tenku, S., Xiao, X., 2015. Mapping oil palm plantations in Cameroon using PALSAR 50-m orthorectified mosaic images. Rem. Sens. 7, 1206–1224. http://dx.doi.org/10.3390/rs70201206.

Lorena, A.C., Carvalho, A.C., 2007. Uma introdução às support vector machines. Rev. Informática Teórica E Apl. 14, 43–67.

MacArthur, D.K., Schueller, J.K., Lee, W.S., Crane, C.D., MacArthur, E.Z., Parsons, L.R., 2006. Remotely-piloted helicopter citrus yield map estimation. Am. Soc. Agric. Biol. Eng. http://dx.doi.org/10.13031/2013.20892.

Marr, D., Hildreth, E., 1980. Theory of edge detection. Proc. R. Soc. Lond. B Biol. Sci. 207, 187–217.

Nagasai, S., Rani, S.J., 2015. Plant disease identification using segmentation techniques. Plant Dis. 4.

Noguchi, Y., Sakai, K., Asada, S., Garciano, L., Sasao, A., 2003. Modeling of alternate bearing in Satsuma Mandarin. J. Jpn. Soc. Agric. Mach. 65, 55–61. http://dx.doi.org/10.11357/jsam1937.65.6_55.

Okamoto, H., Lee, W.S., 2009. Green citrus detection using hyperspectral imaging. Comput. Electron. Agric. 66, 201–208. http://dx.doi.org/10.1016/j.compag.2009.02.004.

Otsu, N., 1975. A threshold selection method from gray-level histograms. Automatica 11, 23–27.

Paris, S., Durand, F., 2006. A fast approximation of the bilateral filter using a signal processing approach. In: Computer Vision–ECCV 2006. Springer, pp. 568–580.

Regunathan, M., Lee, W.S., 2005. Citrus fruit identification and size determination using machine vision and ultrasonic sensors. Am. Soc. Agric. Biol. Eng. http://dx.doi.org/10.13031/2013.19821.

Sa'ad, F.S.A., Ibrahim, M.F., Shakaff, A.Y.M., Zakaria, A., Abdullah, M.Z., 2015. Shape and weight grading of mangoes using visible imaging. Comput. Electron. Agric. 115, 51–56. http://dx.doi.org/10.1016/j.compag.2015.05.006.

Sakai, K., Noguchi, Y., Asada, S.-I., 2008. Detecting chaos in a citrus orchard: reconstruction of nonlinear dynamics from very short ecological time series. Chaos, Solitons Fractals 38, 1274–1282. http://dx.doi.org/10.1016/j.chaos.2007.01.144.

Savvides, M., Kumar, B.V., 2003. Illumination normalization using logarithm transforms for face authentication. In: Audio-and Video-Based Biometric Person Authentication. Springer, pp. 549–556.

Schölkopf, B., Smola, A.J., 2002. Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond. MIT Press.

Thompson, S.K., 1992. Sampling, Wiley Series in Probability and Mathematical Statistics. Wiley, New York.

Triboni, H.de R., Barbosa, J.C., 2004. Estimativa do número de frutos por amostragem de parte da copa em laranjeiras. Rev. Bras. Frutic. 26, 454–458.

Wang, M., 2011. 3D Digital Relief Generation. Bournemouth University.

Warne, P.P., Ganorkar, S.R., 2015. Detection of diseases on cotton leaves using K-mean clustering method. Int. Res. J. Eng. Tech. 2, 425–431.

Weiss, B., 2006. Fast median and bilateral filtering. In: ACM SIGGRAPH 2006 Papers, SIGGRAPH '06. ACM, New York, NY, USA, pp. 519–526. http://dx.doi.org/10.1145/1179352.1141918.

Ye, X., Sakai, K., Asada, S., Sasao, A., 2008. Application of narrow-band TBVI in estimating fruit yield in citrus. Biosyst. Eng. 99, 179–189. http://dx.doi.org/10.1016/j.biosystemseng.2007.09.016.

Ye, X., Sakai, K., Garciano, L.O., Asada, S.-I., Sasao, A., 2006. Estimation of citrus yield from airborne hyperspectral images using a neural network model. Ecol. Model. 198, 426–432. http://dx.doi.org/10.1016/j.ecolmodel.2006.06.001.

Ye, X., Sakai, K., Manago, M., Asada, S., Sasao, A., 2007. Prediction of citrus yield from airborne hyperspectral imagery. Precis. Agric. 8, 111–125. http://dx.doi.org/10.1007/s11119-007-9032-2.

Yuen, H.K., Princen, J., Dlingworth, J., Kittler, J., 1989. A Comparative Study of Hough Transform Methods for Circle Finding. Alvey Vision Club, pp. 29.1–29.6. http://dx.doi.org/10.5244/C.3.29.

Zhao, M., Qin, J., Li, S., Liu, Z., Cao, J., Yao, X., Ye, S., Li, L., 2015. An automatic counting method of maize ear grain based on image processing. In: Li, D., Chen, Y. (Eds.), Computer and Computing Technologies in Agriculture VIII, IFIP Advances in Information and Communication Technology. Springer International Publishing, pp. 521–533.