New method for on-the-go yield assessment using machine vision

Javier Tardaguila*, Arturo Aquino, Borja Millan
Instituto de Ciencias de la Vid y del Vino (University of La Rioja, CSIC, Gobierno de La Rioja), Logroño, La Rioja, Spain
* Corresponding author. E-mail: javier.tardaguila@unirioja.es

Abstract

New non-destructive methods for yield components assessment in the vineyards are needed. The goal of this work was to develop a new methodology for on-the-go assessment of yield in vineyards using machine vision. Vine images were captured on-the-go at night time using a RGB camera embedded on an all-terrain vehicle (ATV). The ATV was equipped with artificial illumination and automatic triggering of the digital camera, and was driven through the vineyard at 7 km/h. Images from three different grapevine varieties (Cabernet Sauvignon, Syrah and Tempranillo) were taken in a VSP commercial vineyard located in Navarra (Spain) on September 2015, prior to harvest. Later on, yield per vine was manually weighted for validation of the new method. The images were processed with a new analysis algorithm which yielded overall average Recall and Precision values of 0.57 and 0.82, respectively. The yield estimation showed a determination coefficient (R^2) of 0.74 and 607 grams of mean error per segment (composed by three vines). The new non-destructive and efficient method can be applied in VSP commercial vineyards for on-the-go yield assessment prior to harvest.

Keywords: yield components, yield forecast, sensing technologies, non-invasive sensors

Resumen

Se necesitan nuevos métodos no-destructivos de estimación de los componentes de la de la producción. El objetivo de este trabajo es el desarrollo de una nueva metodología para la estimación de la producción en el viñedo mediante el uso de una plataforma móvil y visión por computador. Se capturaron imágenes de las vides utilizando un *quad* modificado equipado con una cámara digital RGB, un sistema de iluminación artificial y de control y disparo de las cámaras. Este vehículo se utilizó para capturar las imágenes en el viñedo durante la noche a una velocidad de 7 km/h. Se analizaron imágenes correspondientes a cuatro variedades (Cabernet Sauvignon, Syrah and Tempranillo) capturadas en un viñedo comercial con conducción VSP en septiembre de 2015, antes de la cosecha. Tras la toma de imágenes, se procedió a la cosecha manual para la validación del método. Las imágenes se procesaron con un nuevo algoritmo de análisis obteniendo una media de precisión y exhaustividad de 0,57 y 0,82 respectivamente. La estimación de la cosecha se realizó con un coeficiente de determinación (R^2) de 0,74 y 607 gramos de error medio por segmento (compuesto por tres cepas). El método presentado permite la estimación rápida y no destructiva de la cosecha utilizando la captura de imágenes automatizada desde una plataforma móvil antes de la cosecha.

Palabras clave: componentes del rendimiento, estimación del rendimiento, tecnologías de detección, sensores no invasivos

Introduction

Grapevine yield estimation is encouraged by its economical relevance (Dunn, 2010), and can help to optimize plant growth and fruit quality. Traditional yield predictions are performed using historical yield records, weather patterns and manual sampling on the field. This methodology is labour intensive, expensive and imprecise because the manual sampling is usually biased. This is consequence of the natural tendency for operators to pick healthier and bigger clusters when asked to randomly select them (Dunn, 2010). Moreover, due to the high labour demand, the sample size is often not enough to represent the spatial variability across the vineyard, resulting in an inaccurate and spatially coarse prediction (Nuske et al., 2011).

Yield estimation can be done using image analysis and manual acquisition of images on the field (Diago et al., 2012; Dunn and Martin, 2004; Herrero-Huerta et al., 2015), but a more automatized approach is desirable for commercial application, and to represent vineyard variability. Modified agricultural vehicles can be used to automate the image capture of large datasets (Font et al., 2015; Nuske et al., 2011), but this approach must face the limitation introduced by the lack of supervision during the capturing process, which greatly affects image quality. The segmentation process of images captured on the field is challenging, because of the uncontrolled scenario characteristics and the lack of uniformity in the colouration of the berry surface caused by the pruine (Diago et al., 2015). Also, it must be noted that not all the berries in a cluster are visible due to occlusions from other berries or vegetal material from the vine. To improve the results of yield estimation, a method that is less affected by these problems (occlusions and segmentation errors) must be used.

In this work, the use of Boolean model (Matheron, 1975; Serra, 1980) is proposed. From an image processing viewpoint, the practical advantage of this model relies in its capabilities to estimate the number of particles present in an image, even when the segmentation present errors or occlusions between particles.

The aim of this work is to offer a novel, precise and fast yield prediction in vineyards using images captured on-the-go and Boolean model.

Materials and methods

The experiments were conducted in 2015 in a commercial vineyard located in Navarra (Spain) of three different grapevine (V*itis vinifera* L.) varieties (Cabernet Sauvignon, Syrah and Tempranillo). The vines were growing in VSP, with north-south row orientation at 2 x 1 m disposition. The six first basal leaves of the selected vines were manually removed after berry-set. A plastic ribbon was used to allow the differentiation in the images of 8 segments per variety, each segment consisting in three vines. After the image capture, the vines were manually harvested and the clusters corresponding to each segment were weighted together to obtain ground truth data of the yield estimation.

The images were captured using a modified ATV (Trail Boss 330, Polaris Industries, Minnesota, USA) at a speed around 7 km/h (Figure 1). The vehicle was equipped with a Sony alpha 7-II digital mirror-less camera (Sony Corp., Tokyo, Japan) with a Vario-Tessar FE 24-70 mm lens. RGB images were saved at a resolution of 24 Mpx (6000 x 3376 pixels), 8 bits per channel and manually combined to obtain 24 sections. A 900 LED Bestlight panel was used for scene illumination. The ATV was fitted with an adjustable mechanical structure that allowed for different height and depth fixation to adapt to the vines configuration (Figure 1). The structure also provided protection against branch impact and allowed the attachment of the illumination equipment. The camera triggering was synchronized with the ATV movement using an inductive sensor attached to the rear axle. This sensor produced 3 pulses per rear-axle revolution, thus allowing to obtain images with an approximate 40% of superposition rate. To adapt the signal of the sensor, a custom-built controller based on Arduino MEGA (Arduino LLC, Italy) was developed. The controller conditioned the signal, stored the geopositioning data from a Leica Zeno 10 (Leica Heerbrug, St. Gallen, Switzerland) GPS receiver in a SD card and showed status information through a tft screen.

Results and discussion

The variability in the colouration of the clusters and the heterogeneous lighting originated errors in the segmentation of the images, with cross interference between the cluster and the cable class (representing the metal wire used in VSP). A filter based on morpholine image analysis (Soille, 2004) was used to improve the cluster segmentation. The segmentation was quantified using manually classified images as ground truth. The differences in the results between no-filter and filter application are not remarkable in terms of Recall (0.58 and 0.56 respectively), but are notable for the Precision (0.71 and 0.79), probing that false positives were correctly eliminated during the filtering, with little loss of true positives. The relative low values of Recall can be explained by the difficulty in pixel discrimination because of the lack of uniformity in the illumination. Figure 2 shows the regions manually segmented as clusters. As it can be observed, these regions were hardly distinguishable even by manual evaluation. An illumination improvement might enhance the segmentation process and thus Recall.

For comparison purposes, the area of the segmented clusters was used to estimate yield. The problems during the segmentation clearly affected the performance of this approach (Table 1), whose root-mean-square error (RMSE)=2.3 kg resulted in a lack of its practical application, even when the coefficient of determination was acceptable (R^2=0.67), as this estimator did not compensate for the occlusions and errors in the segmentation. Its slope (0.31) is far from 1, which is the expected slope for perfect estimator (Figure 3). In the other hand, the Boolean model was capable to correctly estimate yield, offering RMSE=0.6 kg with a slope of 0.81. It must be noted that the estimation refers to segments composed by three vines, so this value corresponds to an error around 200 g per vine.

In a similar study, Nuske et al. (2014) also used an ATV with artificial lighting for image capturing of grapevines. The collected images were analysed to identify visible berries to estimate yield. This setup allowed to asses yield with a R^2=0.73 for the best dataset. They also tried to boost the yield estimation through an evaluation of the self-occlusion of berries using 3D models of berries (ellipsoid 3D model) and clusters (convex hull 3D model). The results showed that the proposed correction models did not improved the overall estimation. In contrast to this, the Boolean estimator, that also compensates for partially occluded berries, generated better results (R^2=0.74).

Conclusions

This work presented a new method for accurate, non-destructive and in field grapevine yield estimation by using computer vision and Boolean models. The images were captured at a commercial speed, comparable to other agricultural equipment used in vineyard management using a modified ATV at night time.

The use of Boolean models allowed overcoming two of the major difficulties in visual yield estimation: this technique is robust against segmentation errors and partial occlusions. The capacity to estimate the visible berry number and the partially hidden ones was confirmed by the comparison between the results obtained with the Boolean model and an area based estimator.

The simplicity and precision of the Boolean model formulation makes it ideal for yield map generation. These maps will represent the spatial variability of the vineyards, thus allowing for grapevine zoning site specific management.

Acknowledgements.

The work leading to these results has received funding project of ReSolVe funded by the European Fund FP7 ERA-net Project, CORE Organic Plus. Borja Millán would like to acknowledge the research founding FPI grant by Universidad de La Rioja, Gobierno de La Rioja.

References cited

DIAGO, M., CORREA, C., MILLÁN, B., BARREIRO, P., VALERO, C., TARDAGUILA, J., 2012. Grapevine yield and leaf area estimation using supervised classification methodology on RGB images taken under field conditions. Sensors 12, 16988–17006. doi:10.3390/s121216988

DIAGO, M.P., TARDAGUILA, J., ALEIXOS, N., MILLAN, B., PRATS-MONTALBAN, J.M., CUBERO, S., BLASCO, J., 2015. Assessment of cluster yield components by image analysis. J. Sci. Food Agric. 95, 1274–1282. doi:10.1002/jsfa.6819

DUNN, G.., 2010. Yield Forecasting, Technical Booklet.

DUNN, G.M., MARTIN, S.R., 2004. Yield prediction from digital image analysis: A technique with potential for vineyard assessments prior to harvest. Aust. J. Grape Wine Res. 10, 196–198.

FONT, D., TRESANCHEZ, M., MARTÍNEZ, D., MORENO, J., CLOTET, E., PALACÍN, J., 2015. Vineyard yield estimation based on the analysis of high resolution images obtained with artificial illumination at night. Sensors 15, 8284–8301. doi:10.3390/s150408284

HERRERO-HUERTA, M., GONZÁLEZ-AGUILERA, D., RODRIGUEZ-GONZALVEZ, P., HERNÁNDEZ-LÓPEZ, D., 2015. Vineyard yield estimation by automatic 3D bunch modelling in field conditions. Comput. Electron. Agric. 110, 17–26. doi:10.1016/j.compag.2014.10.003

MATHERON, G., 1975. Random sets and integral geometry. Wiley, New York.

NUSKE, S., ACHAR, S., GUPTA, K., 2011. Visual yield estimation in vineyards: experiments with different varietals and calibration procedures. Technical Report CMU-RI-TR-11-39. Pittsburgh.

NUSKE, S., WILSHUSEN, K., ACHAR, S., YODER, L., NARASIMHAN, S., SINGH, S., 2014. Automated visual yield estimation in vineyards. J. F. Robot. 31, 837–860. doi:10.1002/rob.21541

SERRA, J., 1980. The Boolean model and random sets. Comput. Graph. Image Process. 12, 99–126. doi:10.1016/0146-664X(80)90006-4

SOILLE, P., 2004. Morphological image analysis: principles and applications, 2nd ed. Springer-Verlag, Berlin, Heidelberg. Doi:10.1007/978-3-662-05088-0

Figure 1: Modified ATV for automatic image capture on-the-go. Equipped with a GPS antenna LED Light Panel, RGB camera and inductive sensor for camera triggering (3 pulses per revolution).

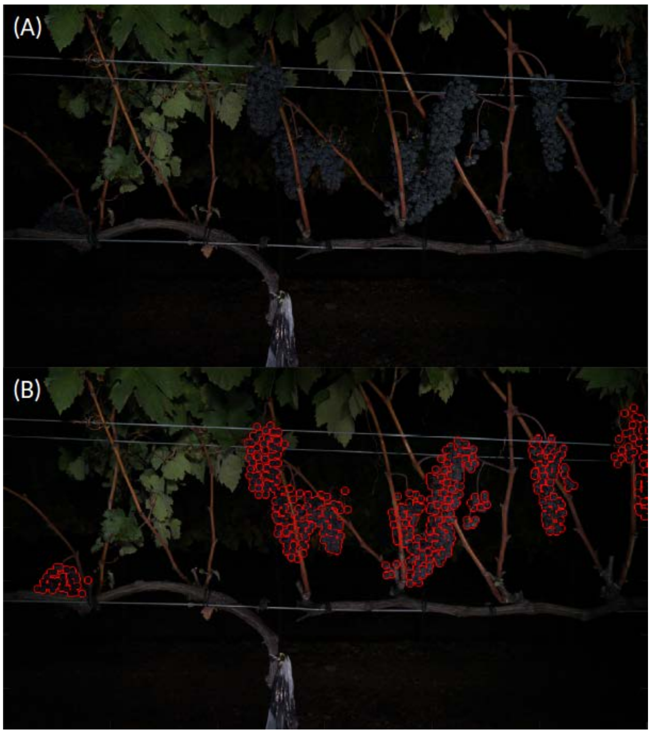

Figure 2: Example image of a vine captured on-the-go in Tempranillo VSP commercial vineyard (A). The berries were manually selected to establish a ground truth for the segmentation. The segmentation ground truth is showed in B.

Table 1: Results obtained for the estimation of yield per segment (composed of three vines) on images captured with an "on-the-go" platform. These measures were obtained when using an area based estimator and the Boolean model on images.

Grapevine variety	Manual harvest		Area based estimation			Boolean model		
	Mean Yield (g)	Number of segments	Mean Yield (g)	R^2	*RMSE* (g)	Mean Yield (g)	R^2	*RMSE* (g)
Cabernet Sauvignon	3406.25	8	1640.40	0.64	1863.7	3234.94	0.50	574.5
Syrah	3126.25	8	1274.50	0.85	2031.9	2948.21	0.84	399.9
Tempranillo	4782.5	8	1875.19	0.54	3194.1	4520.52	0.55	785.6
Global	3771.67	24	1596.70	0.67	2319.2	3567.89	0.74	607.5

RMSE: Root-mean-square error

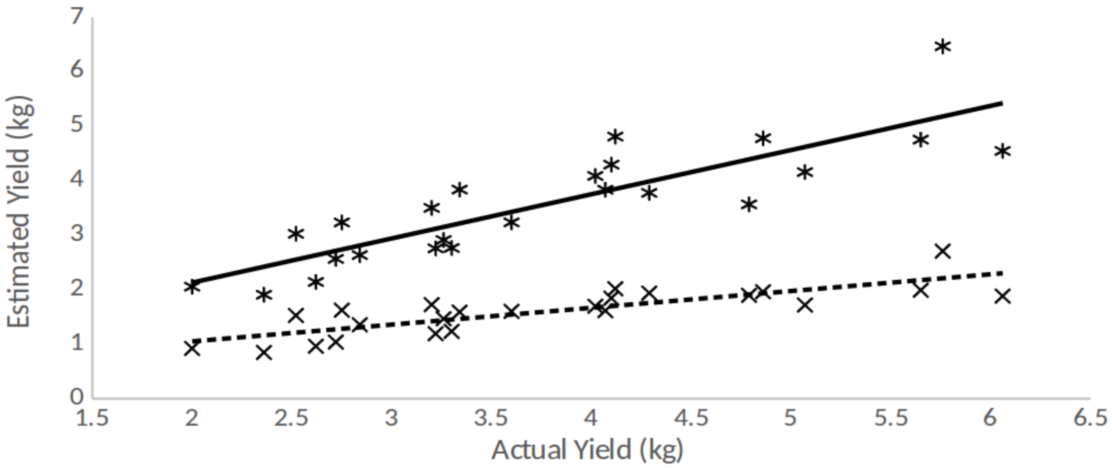

Figure 3: Yield estimation using the area based estimator (crosses and dashed line) $f(x)=0.31x+433.98$; $R^2=0.67$ and the Boolean model (stars and solid line) $f(x)=0.81x+506.87$; $R^2=0.74$ using images captured on-the-go.