Computers and Electronics in Agriculture 151 (2018) 136–149

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Original papers

A pattern recognition strategy for visual grape bunch detection in vineyards

Rodrigo Pérez-Zavala[a], Miguel Torres-Torriti[a,*], Fernando Auat Cheein[b], Giancarlo Troni[c]

[a] Dept. of Electrical Engineering, Pontificia Universidad Católica de Chile, Chile
[b] Dept. of Electronic Engineering, Universidad Técnica Federico Santa María, Chile
[c] Dept. of Mechanical Engineering, Pontificia Universidad Católica de Chile, Chile

ARTICLE INFO

Keywords:
Grape bunch detection
Grape recognition
Precision viticulture
Histogram of oriented gradients
Local binary pattern
Support vector machine

ABSTRACT

Automating grapevine growth monitoring, spraying, leaf thinning and harvesting tasks, as well as improving yield estimation and plant phenotyping, requires reliable methods for detecting grape bunches across different vineyard environmental and plant variety conditions, in which illumination, occlusions, colors and contrast are the main challenges to computer vision techniques. This work presents a method that employs visible spectrum cameras for robust grape berries recognition and grape bunch detection that does not require artificial illumination nor is limited to red or purple grape varieties. The proposed approach relies on shape and texture information together with a strategy to separate regions of clustered pixels into grape bunches. The approach employs histograms of oriented gradients (HOG) as shape descriptor and local binary patterns (LBP) to obtain texture information. A review of the existing methods and comparative analysis of different feature vectors (DAISY, DSIFT, HOG, LBP) and support vector classifiers (SVM-RBF, SVDD) is also presented. Datasets from four countries containing 163 images of different grapevine varieties acquired under different vineyard illumination and occlusion levels were employed to assess the approach. Grapes bunches are detected with an average precision of 88.61% and average recall of 80.34%. Single berries are detected with precision rates above 99% and recall rates between 84.0% and 92.5% on average. The proposed approach should facilitate the estimation of yield, crop thinning measurements and the computation of leaf removal indicators, as well as the implementation guidance strategies for precise robotic harvesters.

1. Introduction

The industry of fresh table and wine grapes represent a major economic activity worldwide, being the fruit crop with highest value, with a market size of approximately 70 billion dollars; see Fig. 1, (FAO-OIV, 2016). Chile is worlds leading fresh table grape exporter (USDA, 2014), and one of the major wine producers, with 5% of the global production (OIV, 2017). Statistics show slightly increasing trends in surface and production during recent years. However, the viticulture industry is facing difficulties due to shortage of qualified fieldworkers and increasing labor costs (Canadian Agricultural Human Resource Council, 2016; Quackenbush, 2017), which affect productivity, quality, and timely harvesting. Additionally, agricultural tasks manually done are often time consuming, inaccurate and subjectively influenced or biased by workers (Nuske et al., 2014). These challenges motivate the development of new technologies that rely on novel sensors and robotics to ensure productivity, quality and economic competitiveness.

Robotic systems in agricultural applications involve three main components: mobility algorithms and hardware (guidance and mapping), perception subsystems, and end effector actuation mechanisms. This work focuses on the perception stage for detection and recognition of grape bunches in the field. For any robotic or automation process, the sensing stage is crucial for the correct performance of the unit. The detection and localization of grape bunches in vineyards has been an important challenge that still has not been completely solved.

The automatic recognition of grapes and grape bunches using machine vision could be employed to automate, manage and optimize current agricultural tasks, such as harvesting (Luo et al., 2016), spraying (Berenstein et al., 2010), grape counting for yield estimation models (Diago et al., 2015; Dunn and Martin, 2004; Liu and Whitty, 2015; Nuske et al., 2014), evaluating grape quality, size and grapevine phenotyping (Roscher et al., 2014; Yandun-Narvaez et al., 2017), detecting disease in clusters, predicting harvest time, quantifying and standardizing crop thinning and basal leaf removal tasks.

While several of the grape detectors developed recently have achieved good performance scores, improvements in grape bunch detection are still possible. The main contribution of this work is the development of a grape bunch detection approach, and not just a grape

* Corresponding author.
E-mail address: mtorrest@ing.puc.cl (M. Torres-Torriti).

https://doi.org/10.1016/j.compag.2018.05.019
Received 24 September 2017; Received in revised form 12 May 2018; Accepted 14 May 2018

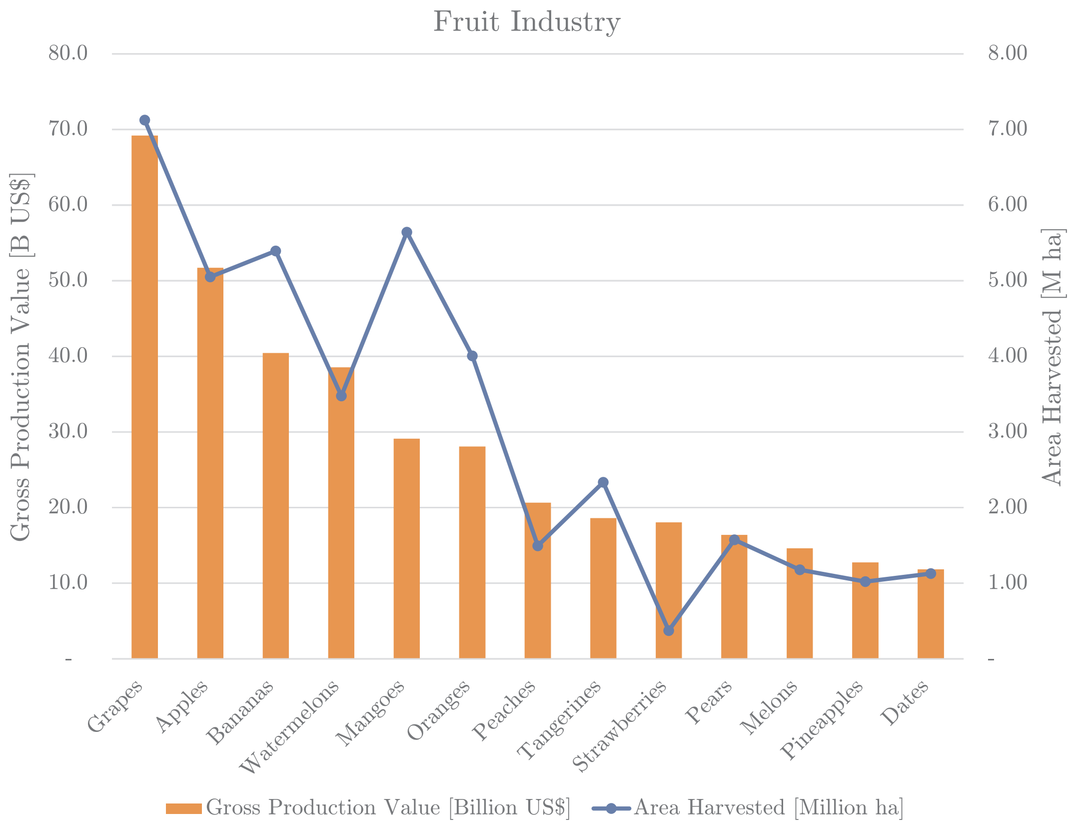

Fig. 1. Top fruit industry (FAO-OIV, 2016).

detector for single berries, which is the approach most frequently reported as will be discussed in more detail in the next section. The proposed method relies on a combination of shape and texture features to achieve higher accuracy and precision rates. Different features were tested in order to identify the best descriptor. Features considered in this study include local binary patterns (LBP), histogram of oriented gradients (HOG), Dense Scale Invariant Feature Transform (DSIFT) and DAISY. Superior results were obtained by combining HOG and LBP. A support vector machine (SVM) classifier was trained as a two class problem with positive and negative images. The performance of the two-class SVM was also compared to the single-class support vector data descriptor method (SVDD).

The proposed approach does not need special illumination and can operate under natural illumination. The results obtained with different datasets, having different levels of occlusion and shades, show the robustness of the approach. Since the method does not use color information, it can be applied to both purple/red and white grapes at different stages of development. We will refer to purple or red grape varieties simply as dark grapes to distinguish them from white grapes.

The paper is organized as follows. Section 2 discusses the existing methods, considering application scope, image processing techniques and the main aspects in which our contribution differs. Section 3 explains the proposed method, while Section 4 presents the experimental methodology and discusses the results. Finally, the conclusion and remarks about the ongoing research are mentioned in Section 5.

2. Related work

Approaches for the detection of grapes and the analysis of vines can be classified according to the scope of the application and the computer vision techniques that are employed. The approaches that are more relevant to the discussion of the strategy proposed here are summarized in Table 1 considering their scope and main underlying technique.

As far as the scope of application is concerned, most of the existing approaches have been conceived for yield estimation (Font et al., 2015; Nuske et al., 2014; Liu and Whitty, 2015; Dunn and Martin, 2004). Another group of approaches aim at identifying grape bunches, either for robotic harvesting (Luo et al., 2016; Reis et al., 2012; Chamelat et al., 2006) or to define regions that need to be analyzed for detailed growth monitoring, leave removal or pruning tasks. Vision algorithms that detect grape and foliage are presented in Berenstein et al. (2010) for selective spraying of hormones and pesticides. Methods for identifying vine parts are proposed in Fernández et al. (2013), Correa-Farias et al. (2012), Diago et al. (2015). Fernández et al. (2013) employ color and multispectral images of Cabernet Sauvignon vineyard scenes to segment stems, leaves, background and grapes. A similar work is presented by Diago et al. (2015) to classify the different parts of Tempranillo vines, including grapes, wood, background and leaves, using a white screen as background like Dunn and Martin (2004). Computer vision techniques have also been applied for measuring grape growth (Roscher et al., 2014), ripeness (Rodríguez-Pulido et al., 2012), and phenotyping (Roscher et al., 2014). Although the scope of application is not grape bunch identification, these works rely on similar image processing techniques employed in grape bunch identification and yield estimation. Recently, some authors have also explored the 3D reconstruction of grape bunches for compactness, volume and yield estimation (Herrero-Huerta et al., 2015; Ivorra et al., 2015; Liu and Whitty, 2015).

Considering computer vision techniques involved in the implementation of each approach, a considerable part of the existing methods rely on color information and some form of classification by color (Luo et al., 2016; Liu and Whitty, 2015; Font et al., 2015; Rahman and Hellicar, 2014; Reis et al., 2012; Diago et al., 2015). The use of descriptors that rely solely on color information is a common strategy that works reasonably well on red or dark grape varieties. However, many authors acknowledge the poor performance of the color-based approaches with white grape varieties due to reduced contrast of the grapes with respect to the leaves and background (Luo et al., 2016; Liu