



Image-based plant phenotyping with incremental learning and active contours



Massimo Minervini, Mohammed M. Abdelsamea, Sotirios A. Tsafaris*

IMT Institute for Advanced Studies, Piazza S. Ponziano, 6, Lucca, Italy

ARTICLE INFO

Article history:

Received 31 January 2013

Received in revised form 8 July 2013

Accepted 22 July 2013

Available online 6 August 2014

Keywords:

Agriculture

Phenotyping

Plant segmentation

Machine learning

Gaussian mixture model

Active contour model

ABSTRACT

Plant phenotyping investigates how a plant's genome, interacting with the environment, affects the observable traits of a plant (phenome). It is becoming increasingly important in our quest towards efficient and sustainable agriculture. While sequencing the genome is becoming increasingly efficient, acquiring phenotype information has remained largely of low throughput. Current solutions for automated image-based plant phenotyping, rely either on semi-automated or manual analysis of the imaging data, or on expensive and proprietary software which accompanies costly hardware infrastructure. While some attempts have been made to create software applications that enable the analysis of such images in an automated fashion, most solutions are tailored to particular acquisition scenarios and restrictions on experimental design. In this paper we propose and test, a method for the segmentation and the automated analysis of time-lapse plant images from phenotyping experiments in a general laboratory setting, that can adapt to scene variability. The method involves minimal user interaction, necessary to establish the statistical experiments that may follow. At every time instance (i.e., a digital photograph), it segments the plants in images that contain many specimens of the same species. For accurate plant segmentation we propose a vector valued level set formulation that incorporates features of color intensity, local texture, and prior knowledge. Prior knowledge is incorporated using a plant appearance model implemented with Gaussian mixture models, which utilizes incrementally information from previously segmented instances. The proposed approach is tested on *Arabidopsis* plant images acquired with a static camera capturing many subjects at the same time. Our validation with ground truth segmentations and comparisons with state-of-the-art methods in the literature shows that the proposed method is able to handle images with complicated and changing background in an automated fashion. An accuracy of 96.7% (dice similarity coefficient) was observed, which was higher than other methods used for comparison. While here it was tested on a single plant species, the fact that we do not employ shape driven models and we do not rely on fully supervised classification (trained on a large dataset) increases the ease of deployment of the proposed solution for the study of different plant species in a variety of laboratory settings. Our solution will be accompanied by an easy to use graphical user interface and, to facilitate adoption, we will make the software available to the scientific community.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

Plants have always been a crucial source of food, feed, fiber, and fuel. Thus, striving for a more sustained agriculture (Peleman and van der Voort, 2003), together with the breeding industry, researchers try to identify, improve, and breed key traits to satisfy growing demand, increase resistance to parasites and diseases, and minimize environmental impact (less water, less fertilizer).

Understanding biological function and the complex processes involved in the development of plants relies on understanding the interaction between genetic information and the environment, and how they affect the phenotype (the appearance or behavior) of the organism and

consequently desirable traits. Model plant systems, such as *Arabidopsis thaliana*, have been used extensively for this purpose (Licausi et al., 2011). Unlike humans and most other mammals, they are considered ideal models to study natural variation and decipher the “genotype to phenotype” link (O'Malley and Ecker, 2010) – arguably a problem that concerns all life science fields. It is expected, that the integration of these findings with bioinformatics and systems biology will lead to the construction of a “virtual plant analog”, which will permit researchers to investigate gene activity at every stage of plant development (Benfey and Mitchell-Olds, 2008; Holtorf et al., 2002).

The rate of throughput for acquiring genetic information (with sequencing and microarrays) has achieved game-changing levels (Mardis, 2008). However, as of today, inexpensive and automated phenotyping (phenomics) remains a bottleneck (Furbank and Tester, 2011; Spalding and Miller, 2013). Until recently most phenotypes (e.g., related to plant growth) were acquired in destructive ways (e.g., weigh the plant, or cut out and measure a leaf) or involved

* Corresponding author.

E-mail addresses: massimo.minervini@imtlucca.it (M. Minervini), mohammed.abdelsamea@imtlucca.it (M.M. Abdelsamea), s.tsafaris@imtlucca.it (S.A. Tsafaris).

human survey (e.g., measuring leaf size or plant radius) in situ without destructing the plant. Naturally these methods are faced with low throughput and high productivity cost. Consequently, there has been a growing interest towards developing solutions for the automated analysis of visually observable traits of the plants. Several consortia such as the International and European Plant Phenotyping Networks (IPPN and EPPN), and the iPlant Collaborative Project (Goff et al., na) have been established to promote and accelerate phenotype discovery and analysis, and increase our understanding of biology (Benfey and Mitchell-Olds, 2008; Finkel, 2009; Micol, 2009). Several approaches based on imaging techniques and computer vision have been proposed to increase the throughput of non-destructive phenotyping (without penalizing accuracy) with solutions that can be affordable and easy to deploy (Golzarian et al., 2011; Spalding and Miller, 2013); however, such systems usually require sophisticated analysis algorithms to segment the plant from the background. As of now the majority of solutions pose strict experimental conditions to ease the complexity of the analysis task that follows.

In this paper we propose and test, an algorithm and a software system for the automated segmentation and analysis of time-lapse top-view plant images from phenotyping experiments of *Arabidopsis* rosettes. Example images of *Arabidopsis* rosettes (referring to the circular and radial cluster of leaves they form when growing) are shown in Fig. 1. We use data which we acquired in a general laboratory setting with a static camera that captures many plants at the same time (Tsaftaris and Noutsos, 2009), with the purpose of collecting test data (which are lacking in the public domain) and demonstrating the challenging aspects of the problem of plant segmentation. While our paper focuses on the algorithm and the software solution, as an example that relates it to phenotyping experiments we measure plant growth, estimated through projected rosette area of 19 *Arabidopsis* Columbia (Col-0) wild-type plants. The system involves minimal user interaction (necessary to establish the statistical experiments that follow) and at every time instance (in our context a digital photograph) segments the plants in images that contain many instances of the same species (an example input image is shown in the bottom of Fig. 1). We rely on a combination of level set and learning based segmentation to incrementally incorporate information from previous time instances, allowing us to adapt to changes in the scene. We learn an appearance model of the plant relying on Gaussian mixture models (GMM) of color and texture features collected from previously segmented instances. For each unseen instance we classify each pixel in the image to obtain a probabilistic map of pixels most likely belonging to a plant.

The probabilistic output assists the localization of a plant within an image (many plants exist in an image). Each individual plant is then segmented using a new active contour model that incorporates probabilistically weighted (using the model output) features of pixel intensity and texture. Once each plant has been segmented, several measurements relevant to the study of plant growth are extracted. Finally, the plant appearance model is updated by re-training the GMM to include the newly processed data in an online and incremental fashion.

Our inclusion of the iterative learning aspect allows us to handle images with complicated and changing background in an automated fashion, which challenge currently available solutions. Overall the proposed approach obtains an accuracy in segmentation higher than 96.4% (dice similarity coefficient) and is not affected by challenges in the scene. While here it was tested on one mutant of *Arabidopsis*, the fact that we do not employ shape driven models and we do not rely on fully supervised classification (trained on a large dataset) favors the deployment of the proposed solution for the study of different plant species in a variety of laboratory settings. The automated phenotyping solution proposed in this paper improves upon the accuracy results obtained by the state of the art in plant phenotyping, even in an environment that is not strictly controlled, thus accommodating a broader range of experimental scenarios — handling multiple plants simultaneously without an explicit scene arrangement and with minimal user intervention. Our solution will be accompanied by an easy to use graphical user

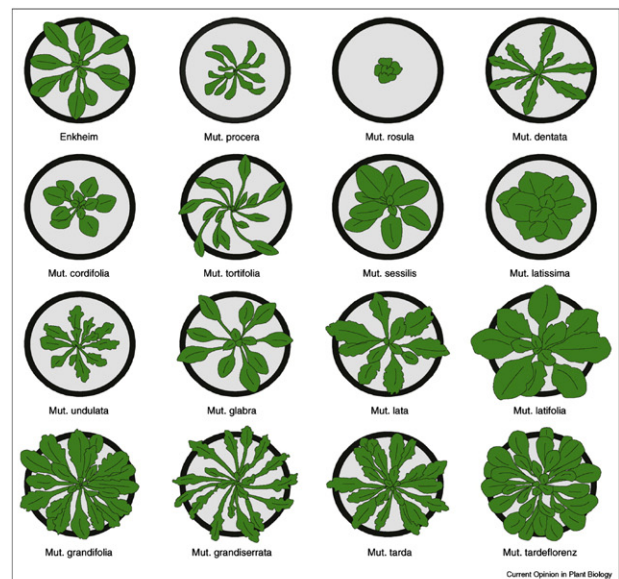


Fig. 1. *Arabidopsis* images showing shape variability. Sketch representations of *Arabidopsis* mutants adapted from Micol (2009) (top); two plants imaged at different growth stages, earlier shown left-most (middle); and a top-view example many individual plants (bottom).

Download English Version:

<https://daneshyari.com/en/article/4374911>

Download Persian Version:

<https://daneshyari.com/article/4374911>

[Daneshyari.com](https://daneshyari.com)