Computers in Industry 98 (2018) 226–240

Contents lists available at ScienceDirect

Computers in Industry

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

Connected attribute morphology for unified vegetation segmentation and classification in precision agriculture

Petra Bosilj*, Tom Duckett, Grzegorz Cielniak

Lincoln Centre for Autonomous Systems, School of Computer Science, University of Lincoln, UK

ARTICLE INFO

Article history:
Received 22 September 2017
Received in revised form 20 December 2017
Accepted 22 February 2018

Keywords:
Precision agriculture
Attribute morphology
Crop/weed discrimination
Max-tree

ABSTRACT

Discriminating value crops from weeds is an important task in precision agriculture. In this paper, we propose a novel image processing pipeline based on attribute morphology for both the segmentation and classification tasks. The commonly used approaches for vegetation segmentation often rely on thresholding techniques which reach their decisions globally. By contrast, the proposed method works with connected components obtained by image threshold decomposition, which are naturally nested in a hierarchical structure called the max-tree, and various attributes calculated from these regions. Image segmentation is performed by attribute filtering, preserving or discarding the regions based on their attribute value and allowing for the decision to be reached locally. This segmentation method naturally selects a collection of foreground regions rather than pixels, and the same data structure used for segmentation can be further reused to provide the features for classification, which is realised in our experiments by a support vector machine (SVM). We apply our methods to normalised difference vegetation index (NDVI) images, and demonstrate the performance of the pipeline on a dataset collected by the authors in an onion field, as well as a publicly available dataset for sugar beets. The results show that the proposed segmentation approach can segment the fine details of plant regions locally, in contrast to the state-of-the-art thresholding methods, while providing discriminative features which enable efficient and competitive classification rates for crop/weed discrimination.

1. Introduction

Robust vision systems are a core technology for building autonomous robots in precision agriculture. Such systems automate time-consuming manual work in the field while increasing yield and reducing the reliance on herbicides and pesticides. To achieve this, the developed vision systems need to be able to monitor the crop and target only the specific plants that need treatment. Specifically, a number of approaches to discriminate value crops from weeds were developed [1–4] and employed in robotic systems, which use this information to perform tasks such as mechanical weeding and selective crop spraying.

Mathematical morphology [5], and specifically attribute morphology, offers a versatile framework to perform multi-scale spatial analysis of image content in various image domains. Efficient implementations rely on hierarchical image representations and enable fast processing of large amounts of image data.

The contributions of this paper are threefold. Firstly, we propose a novel and unified pipeline for crop/weed detection and classification relying fully on attribute morphology. Secondly, we evaluate the approach on a publicly available sugar beets classification dataset [6] as well as a newly collected dataset focused on onion crops, which exhibits a higher variation in lighting and registration errors, thus requiring a more robust solution. Finally, we demonstrate the locality of the proposed approach and its ability to segment the fine details of plants, in contrast to the state-of-the-art global thresholding methods, as well as the discriminative properties of the provided features by obtaining competitive classification rates for crop/weed discrimination.

In the following section, we give a brief overview of related work from both precision agriculture and image morphology. Then, in Section 3 we explain the basic principles of attribute morphology, highlighting its advantages compared to standard structuring element morphology and explaining the data structure which enables the efficient implementation of the proposed

* Corresponding author.
E-mail addresses: pbosilj@lincoln.ac.uk (P. Bosilj), tduckett@lincoln.ac.uk (T. Duckett), gcielniak@lincoln.ac.uk (G. Cielniak).

https://doi.org/10.1016/j.compind.2018.02.003
0166-3615/

pipeline. The core of the proposed approach is presented in Section 4. The data and experimental setup are explained in Section 5 followed by the results in Section 6. Finally, we conclude the paper and highlight future research directions in Section 7.

2. Related work

In order to apply per-plant treatments in precision agriculture, the vision system first performs segmentation, thus discarding all non-vegetation pixels, followed by classification of the remaining vegetation pixels to determine the correct treatment for plant regions of different types. We examine the related work through this two-step process.

Several choices for the segmentation step were explored in the literature, including colour-index based images calculated from *RGB* images (examples include *ExG* [7], *ExR* [8], *CIVE* [9], *VEG* [10]), normalised difference vegetation index (NDVI) images obtained from multi-spectral cameras as the difference-sum ratio of the near infra-red and visible red components [11,12], images in colour spaces such as LAB [13] and different hue-based colour spaces [14]. The choice of input image is then thresholded to separate soil from vegetation. The threshold decision is usually reached globally, e.g. using Otsu's threshold selection method [15], resulting in methods sensitive to varying lighting conditions and requiring post-processing to locally adjust the output of thresholding by removing noise. More robust segmentation approaches were developed using machine learning-based methods [16,14,2,13] but they come at an increased computational cost and are not well suited for real-time applications. For a recent overview of segmentation techniques applied to vegetation segmentation the reader is referred to [17].

Following the segmentation step, the foreground (vegetation) pixels are further classified into crops and weeds. Distinguishing between multiple weed classes is sometimes also of interest. Classifying only the vegetation pixels instead of all image pixels significantly reduces the computational load of classification. Colour information is often not enough to perform classification successfully, so additional information about texture and shape is often introduced. The two main approaches to classification are local pixel or grid-based approaches [3,18,4] and region-based approaches [1,19], which can also be used in conjunction [12] to benefit from the advantages of both approaches. While the region-based approaches are typically very fast, as they deal with several tens of regions per image, they cannot cope well with occlusions to reach a fine-grained decision on a vegetation patch with overlapping crops and weeds. An additional component labelling step on the segmented image is required to prepare the input for a region-based classifier. On the other hand, pixel-based approaches suffer from high computational cost. This is partially mitigated by classifying only certain pixels on a grid and interpolating the classification values of other pixels. However, due to their high classification accuracy and robustness to partial occlusion, the strength of these approaches lies in applying them to the limited amount of pixels for which the region-based approaches do not reach a certain decision. In this paper, we propose a novel pipeline for both segmentation and region-based classification of plants, while the development of a complementary pixel-based classified is left for future work.

Mathematical morphology, with the recent developments in hierarchical image representation and attribute morphology, offers a versatile and efficient framework to perform multi-scale spatial analysis of image content in various image domains. Historically applied to segmentation problems [20–22], various morphological techniques were recently successfully applied to a large number of image processing and computer vision problems including object detection [23,24], segmentation [25,26], image retrieval [27–29], scene classification for remote sensing [30–32] and more. Fast processing is achieved by using a hierarchical image decomposition such as the max-tree [22], relying on efficient construction algorithms, parallelisation and simultaneous calculation of attributes used throughout the processing pipeline, allowing attribute morphology approaches to be applied to images as large as several Gpx, with reported speeds of up to 370 Mpx/s when using parallelisation [33].

3. Attribute morphology and hierarchies

Classical approaches to mathematical morphology rely on the concept of a structuring element (*SE*) to define the basic operations of erosion and dilation, and then opening and closing. The erosion operation will erode or shrink the boundaries of foreground regions, thus making the foreground shrink in size and removing all small foreground components, with dilation being the complementary operation. Combining erosion and dilation sequentially produces the opening operator, which enables the removal of small foreground components without introducing big changes to other foreground elements. The complementary operator of closing is obtained by first applying dilation and then erosion. The SE is a (typically small) binary image with a defined origin, with which the input image is "probed" to calculate the output image. Thus, an erosion corresponds to placing the SE at all positions in the input image, and placing a foreground pixel at the SE origin in the output image if all the SE pixels fall onto foreground pixels of the input image. Similarly, with dilation a foreground pixel is placed at the SE origin in the output image if any of the SE pixels fall onto the foreground in the input image. Finally, an opening operator corresponds to placing an SE at all positions in the input image, and placing foreground pixels on all the SE pixels in the output only if the whole SE falls into the foreground.

However, relying on a structuring element to define an opening has several drawbacks: the boundaries are not faithfully preserved, the method is not rotationally invariant (i.e. designing a single SE to respond to elongated objects is not possible and thus multiple line-like SEs with different orientations need to be used), and shape and size are treated together, making it difficult to filter objects based on only one of these characteristics.

To address these problems, morphology has moved in the direction of connected filters [34,26]. The first such operators were binary opening and closing by reconstruction [35,36], which still rely on a SE to define which foreground regions should be removed from the image but fully reconstruct all the remaining components. The problems of rotational invariance and decoupling of shape and size are addressed in attribute morphology [37,22], in which the SE is omitted completely. Instead, in attribute morphology the decision to keep or discard is reached at region level, thus only keeping the regions where an attribute satisfies a certain criterion. This allows using criteria such as "area of the region is greater than 100" to process the input image. The difference between an opening with an SE, opening by reconstruction and an area opening on a binary image is shown in Fig. 1.

All the basic operators defined in SE morphology are *increasing*, meaning that they preserve the order of binary images such that $B_1 \subseteq B_2$ then $F(B_1) \subseteq F(B_2)$ where $F(\cdot)$ is the operation of erosion, dilation, opening or closing by an SE. This allows the extension of binary SE morphology to greyscale images relying on the principles of threshold decomposition [38] and stacking [39]. A greyscale image $f: E \to \mathbb{Z}, E \subseteq \mathbb{Z}^2$ is represented by its upper-level sets, defined as $\mathcal{L}^k = \{f \geq k\}$ with $k \in \mathbb{Z}$, i.e. the set of images obtained by thresholding an image at all possible values of their pixels (similarly one can work with lower-level sets $\mathcal{L}_t$). The result of applying a morphological operator to a greyscale image can then