Image and Vision Computing 28 (2010) 1049–1062

Contents lists available at ScienceDirect

Image and Vision Computing

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

Delving deeper into the whorl of flower segmentation ☆

Maria-Elena Nilsback *, Andrew Zisserman

Visual Geometry Group, Department of Engineering Science, University of Oxford, 3 Parks Road, Oxford OX1 3PJ, UK

ARTICLE INFO

Article history:
Received 18 June 2008
Received in revised form 20 May 2009
Accepted 5 October 2009

Keywords:
Image segmentation
Spatial model

ABSTRACT

We describe an algorithm for automatically segmenting flowers in colour photographs. This is a challenging problem because of the sheer variety of flower classes, the variability within a class and within a particular flower, and the variability of the imaging conditions – lighting, pose, foreshortening, etc.

The method couples two models – a colour model for foreground and background, and a light generic shape model for the petal structure. This shape model is tolerant to viewpoint changes and petal deformations, and applicable across many different flower classes. The segmentations are produced using a MRF cost function optimized using graph cuts.

We show how the components of the algorithm can be tuned to overcome common segmentation errors, and how performance can be optimized by learning parameters on a training set.

The algorithm is evaluated on 13 flower classes and more than 750 examples. Performance is assessed against ground truth trimap segmentations. The algorithms is also compared to several previous approaches for flower segmentation.

1. Introduction

There is an interesting research theme in computer vision of using a general class model to initialize a segmentation, and then improving the segmentation using image-specific features. The LOCUS approach of Winn and Jojic [1] is a good example, where a shape matte (e.g. for a car or horse) is used to propose a foreground/background segmentation, colour distributions are then gathered in each region, and the final segmentation achieved using a colour based binary MRF optimized with graph cuts. The `ObjCut` algorithm of Kumar et al. [2] similarly proposes a position and configuration for a class instance (e.g. a cow or horse) using a pictorial structure with boundary shapes and texture features, and then, again, incorporates colour measured from the proposed regions to carry out a MRF segmentation with graph cuts. Similar ideas are present in several other recent class based segmentation methods [3–6]. Even though these methods are applied to a variety of classes, e.g. cars, cows, faces, horses, for the most part a different model is used for each class, and a different model is used for each view of the class (e.g. cars rear, cars side).

In this paper we introduce two variations on this theme: first, we reverse the order in which the features are used – we start with colour to propose a foreground/background segmentation and use this to initialize image-specific shape measurements; second, we use a generic shape model which is applicable across a number of classes and viewpoints.

Our goal is to automatically segment out the flower given only that the image is known to contain a flower, but no other information on the class or pose. This capability can be used to "power assist" interactive image segmentation – the flower is segmented without any manual interaction. In a graphics application the flower can then be cut-and-pasted into another image. However, our target application is automated flower classification from photographs, and segmentation forms the first step of this process [7–9]. Fig. 1 illustrates the challenge of the segmentation task, and shows the fitted generic flower shape model for several flower classes.

The method is evaluated on the Oxford 17 Flower Dataset available at [10]. The dataset has 17 flower classes (e.g. buttercup, daffodil, iris, pansy), with photographs exhibiting typical (large) variations in viewpoint, scale, illumination and background. Segmenting such photographs is challenging due to both the variety of colours and the variety of shapes. If we knew that we were looking for a daffodil or a bluebell, we could build one foreground and one background colour model for each of these classes – though this would still give problems with the sky being segmented as foreground in a bluebell photograph for example. But here we do not wish even to specify the flower class in advance. Shape also poses a challenge because of the many different types and whorls of petals. Even on the same flower there are local deformations of the petal shape.

Previous work on segmenting flowers using colour by Das et al. [11] was aimed not at extracting exact foreground regions, but

☆ Expanded version of a paper presented at the 18th British Machine Vision Conference (Warwick, September 2007).

* Corresponding author. Tel.: +44 1865283045.
E-mail addresses: men@robots.ox.ac.uk (M.-E. Nilsback), az@robots.ox.ac.uk (A. Zisserman).

0262-8856/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.imavis.2009.10.001

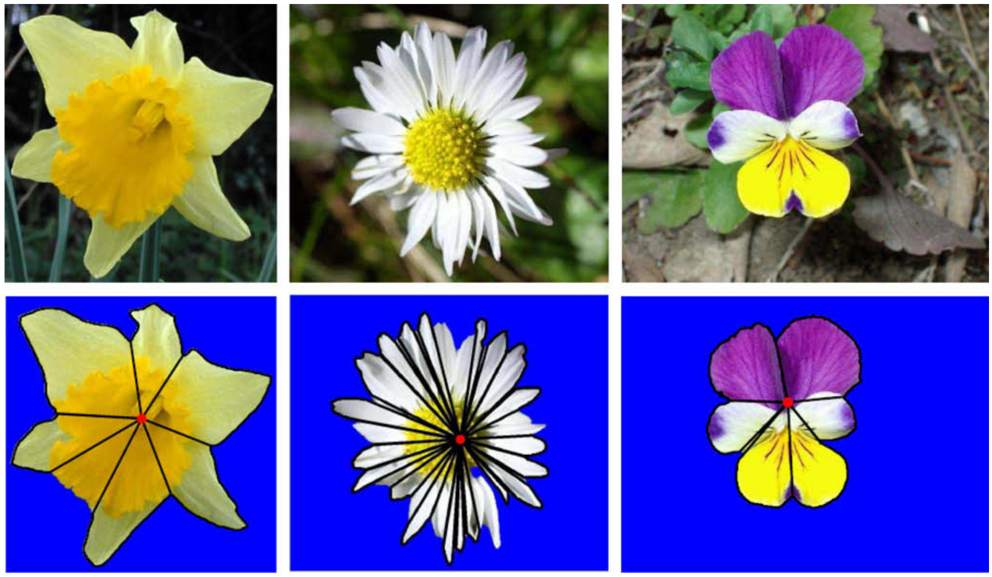

Fig. 1. Top row: typical flower photographs, note the variety of imaging conditions. The classes are (left to right) daffodil, daisy and pansy. Bottom row: the flower shape model fitted automatically (note the lines to the flower "centre") and the resulting segmentation. The same shape model is used in all cases despite differing numbers of whorls, the large variation in the number of petals (daffodil and pansy vs. daisy), and variation in the shape of the petals.

instead at isolating regions in the image that accurately describe the flower colour. Their method learns an image-specific background colour model from dominant colours in the periphery of the image. These colours are removed from a pool of possible foreground colours, and the remaining colours used to obtain a foreground region. The region is accepted provided it satisfies certain spatial conditions. In their implementation Das et al. [11] also always remove green, brown, gray and black from the pool of foreground colours. This is not suited to our database as the choice would fail to segment the centre of the sunflower, the spots on the tigerlily, the stripes on the pansy, etc. However, the strategies for hypothesizing and checking the image-specific background colours proposed in [11] are complementary to the use of a spatial model proposed here, and we return to this point later.

Saitoh et al. [9] proposed a method for extracting flowers regions. It is based on "Intelligent Scissors" [12], which find the path between two points that minimizes a cost function dependent on image gradients. The method works under the assumption that the flower is in focus and in the centre of the photograph and that the background is out of focus. Under this assumption the cost between any two points on the flower is smaller than the cost between a point in the background and a point in the foreground. By fixing the midpoint of the image as part of the flower this can be used as a starting point for finding the flower region. This method requires no prior colour information. We compare with this method to determine the impact of these assumptions.

This paper is organized as follows: In Section 2 we describe our segmentation algorithm, and introduce the flower shape model. The method is evaluated on a dataset of 13 different flower classes including daffodil, crocus, iris, tigerlily, wild tulip, fritillary, sunflower, daisy, colt's foot, windflower and pansy, with over 40 instances of each class. Segmentation performance is measured against ground truth trimap segmentations. The database and evaluation protocol are described in Section 3. Section 4 describes how the parameters of the algorithm are optimized over a training set, and Section 5 assesses the quality of the flower segmentation algorithm on a test set, and compares it to previous algorithms.

This paper is an extended version of [13]. The extensions include schedules for updating the foreground and background colour distributions, optimizing parameters on a training set, and comparisons with previous algorithms for flower segmentation. These generalizations and optimizations result in a 10% performance boost over the method of [13].

2. The segmentation algorithm

2.1. Overview

We first obtain an initial flower segmentation using *general* (non-class specific) foreground and background colour distributions. These distributions are learnt by labelling pixels in four training images for each class in the dataset as foreground (i.e. part of the flower), or background (i.e. part of the greenery), and then averaging the distributions across all classes. Given these general foreground and background distributions, a binary segmentation is obtained using the contrast dependent prior MRF cost function of [14], optimized with graph cuts. This is the method used in [7]. This segmentation may not be perfect, but is often sufficient to extract at least part of the external boundary of the flower.

The generic flower shape model is then fitted to this initial segmentation in order to detect petals. The model selects petals which have a loose geometric consistency using an affine invariant Hough like procedure. The image regions for the petals deemed to be consistent are used to obtain a new *image-specific* foreground colour model. The foreground colour model is then updated by blending the image-specific foreground model with the general foreground model. Similarly the *background* colour model is updated by blending an image-specific model with the general background distribution. The MRF segmentation is repeated using the new colour models. In cases where the initial segmentation was not perfect, the use of the image-specific foreground and background often harvests more of the flower. The steps of shape model fitting and image-specific foreground and background learning can then be iterated until convergence.

The algorithm is illustrated in Fig. 2. We first describe these stages in more detail and then give implementation details. Variations on the schedule for combining the image-specific and general distributions are discussed in Section 4.