J. Vis. Commun. Image R. 38 (2016) 540–549

Contents lists available at ScienceDirect

J. Vis. Commun. Image R.

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

Constellational contour parsing for deformable object detection ☆

Chen Huang [a], Tony X. Han [a], Zhihai He [a], Wenming Cao [b,*]

[a] Department of Electrical and Computer Engineering, University of Missouri, Columbia, MO 65211, USA
[b] Department of Computer Science, Shenzhen University, Shenzhen, Guangdong Providence 518060, China

ARTICLE INFO

Article history:
Received 18 July 2015
Revised 6 February 2016
Accepted 30 March 2016
Available online 30 March 2016

Keywords:
Object detection
Shape matching
Optimization
Additive similarity measure
Dynamic programming
Constellational contour parsing

ABSTRACT

In this paper we propose a novel framework for contour-based object detection from cluttered environments. Given a contour model for a class of object, it is first decomposed into fragments, then in the test image we simultaneously perform selection of relevant contour fragments in edge images, grouping of the selected contour fragments, and finding best geometry-preserving matching to model contours. Finding the best matching is inherently a computationally expensive problem. To address this challenge, we developed local shape descriptors and an additive similarity metric function which can be computed locally while preserving the capability of matching deformable shapes globally. This allows us to establish a constellational shape parsing framework using low-complexity dynamic programming to find optimal configuration of contour segments in test images to match the model contour. To effectively detect objects with large deformation, we augmented the metric function with a local motion search, modeled the relationship between different shape parts using multiple concurrent dynamic programming shape parsers. Our experimental results show that the proposed method outperforms the state-of-the-art contour-based object detection algorithms on two benchmark datasets in terms of average precision.

1. Introduction

Object detection is an important and challenging task in computer vision. It allows localization of previous unseen objects in new images. In general, two main paradigms can be distinguished: *appearance-based* and *contour-based*. Appearance-based approaches [1–4] form the dominant paradigm using local image patch features and constructing rules and models with powerful classification and learning methods for object detection. These types of methods often require a large image dataset for training. Recently, several contour-based methods have demonstrated proficiency at the task of object detection, such as [5–8]. Compared to other image cues, the outline contour provides a set of powerful and robust features for object characterization which are largely invariant to changes in illumination, object colors and texture [9,10]. More importantly, the outline contour can efficiently represent image structures with large spatial extents [11]. Recent work on contour-based object detection has shown that, using a single hand-drawn example (e.g., a bottle sketch) instead of a large set of training data, enables detection of objects with a wide variety of textures, poses, and sizes from highly cluttered images with significant illumination variations and shape deformations [12–16].

Contour-based object detection is a challenging task. It operates on the edge or contour map extracted from the image obtained by an edge detector, such as Canny [17] or Pb [18]. The edge pixels are grouped into edge fragments in a bottom up process using an edge-linking algorithm, e.g., [19]. Given the contour of target object as a model, the goal of contour-based object detection is to select a small subset of edge fragments that minimize the dissimilarity to the model contour. as shown in Fig. 1.

Shape contours obtained from object segmentation often exhibit a large degree of intra-class of variations and inter-class ambiguities [20]. For example, animals of similar species are often very similar to each other, except for some small distinguishable features.

During contour-based object detection, we need to identify the subset of contour segments from the cluttered image to best match the example shape model. The shape description, representation, and matching scheme should be able to accommodate large shape deformation of the same object class. Furthermore, the detection process also needs to deal with a so-called hallucination problem, where a subset of edge segments in the background form a shape similar to the target object, resulting in a false positive during detection [14].

☆ This paper has been recommended for acceptance by Zicheng Liu.

* Corresponding author.
E-mail addresses: chenhuang@mail.missouri.edu (C. Huang), hantx@missouri.edu (T.X. Han), HeZhi@missouri.edu (Z. He), wmcao@szu.edu.cn (W. Cao).

http://dx.doi.org/10.1016/j.jvcir.2016.03.028
1047-3203/

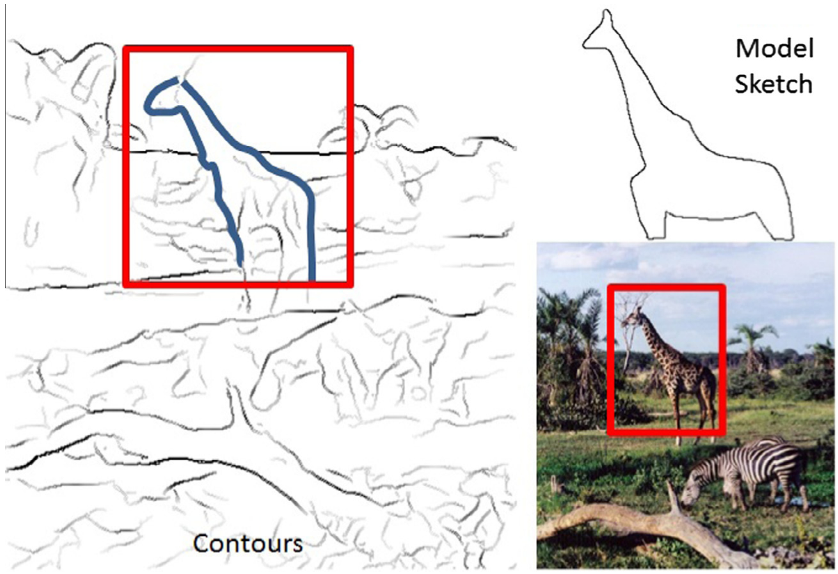

Fig. 1. An example of contour-based object detection results generated by the proposed method.

This work presents a new method, called *constellation contour parsing* (CCP), for contour-based object detection. Given a shape prototype of the object, such as a hand-drawn sketch or a model image, we aim to search through the clouds of contour segments in the test image to identify the optical configuration of contour segments which best matches the shape prototype, as shown in Fig. 1. We use the term *constellation* because this problem is similar to finding a specific constellation of stars in the sky.

Finding the best constellation is combinatorial optimization problem. After developing local shape descriptors and an additive similarity metric function which can be computed locally while preserving the capability of matching deformable shapes globally, we establish a constellational shape parsing framework using dynamic programming. To effectively detect objects with large deformation, we perform implicit shape parts analysis, model the relationship between different shape parts using multiple concurrent dynamic programming shape parsers, and finalize the detection result using voting scheme. Our experimental results show that the proposed method outperforms the state-of-the-art contour-based object detection algorithms on two benchmark datasets.

Grouping edge pixels into contours can increase the discriminative power compared to considering an unstructured set of edge pixels. However edge and contour detection are sensitive to local image changes and noise. The contour of one object may be broken into several fragments. Some fragments of the target object could be missing. The object contour could wrongly connect with contours in the background clutter [21]. These types of errors can significantly limit the applications of the image contour-based object detection method. CPP does not require a complete edge contour of objects, because what CPP are trying to search is only a set of contour segments rather than complete edge contour. So the proposed detection method is robust to broken edges and can obtain very impressive detection results even when the edge quality is bad. This fact is illustrated in Fig. 2.

In this paper, we choose to explore an object detection system that exploits only contour-based information, because we believe shape contour provides a powerful and generic feature since it is invariant to extreme lighting conditions and large variations in texture or color. Clearly the eventual goal of any detection algorithm is to combine sensibly many different useful types of features, but for the purpose of this paper, we deliberately ignore these to show just how powerful the shape contour is.

The major contribution of this work lies in the following three aspects. (1) We introduced the scheme of constellational contour parsing (CCP) with anchor points, observation rays, and additive shape metric function to convert a combinatorial contour grouping problem into a dynamic programming optimization problem with low complexity. (2) We developed an implicit shape parts analysis scheme using multiple observation points to handle object deformation. (3) The proposed method is able to detect objects with different scales, rotations, perspective changes, and deformations at low computational complexity and achieves very encouraging and highly competitive results.

The remainder of the paper is organized as follows. Section 2 reviews the related work on contour-based object detection. Section 3 provides an overview of the proposed system. The constellational contour parsing is explained in Section 4. Experimental results are presented in Section 5.

Concluding remarks and discussions are given in Section 6.

2. Related work

Shape is widely considered as one of the best features in computer vision [22]. In contrast to gradient and texture-based representations, shape is more descriptive at a larger scale, ideally capturing the object of interest as a whole [23–25]. A successful shape matching scheme consists of two components: a compact and accurate shape representation, and an efficient as well as robust distance measurement. There are many ways to represent shapes. Examples include axial representation [26,27], primitive-based representation [28], reference points and projection based representation [29], histograms of oriented gradients [2], hierarchical Fourier descriptor [30], etc.

Measuring the similarity between two shapes often can be done in two ways: (1) by computing the direct difference in features extracted from shape contours, which are invariant to the choice of starting points and robust to a certain degree of deformation; and (2) by performing matching to find the detailed point-wise correspondences to compute the differences, such as the famous Shape Context [31] and Self-Similarity [32]. The latter has recently become dominant due to its ability to capture intrinsic properties, thus leading to more accurate similarity measures. Inner Distance Shape Context (IDSC) improves the distance of the Shape Context