Signal Processing: Image Communication 44 (2016) 29–43

Contents lists available at ScienceDirect

Signal Processing: *Image Communication*

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

Fast two-cycle curve evolution with narrow perception of background for object tracking and contour refinement

Yaochen Li [a], Yuanqi Su [b,*], Yuehu Liu [a]

[a] *Institute of Artificial Intelligence and Robotics, Xi'an Jiaotong University, Xi'an 710049, China*
[b] *Department of Computer Science and Technology, Xi'an Jiaotong University, Xi'an 710049, China*

ARTICLE INFO

Article history:
Received 26 May 2015
Received in revised form
16 January 2016
Accepted 1 March 2016
Available online 4 March 2016

Keywords:
Object tracking
Contour refinement
Curve evolution
Region competition
Maximum likelihood

ABSTRACT

The problem of object contour tracking in image sequences remains challenging, especially those with cluttered backgrounds. In this paper, the fast two-cycle level set method with narrow perception of background (FTCNB) is proposed to extract the foreground objects, *e.g.* vehicles from road image sequences. The curve evolution of the level set method is implemented by computing the signs of region competition terms on two linked lists of contour pixels rather than by solving partial differential equations (PDEs). The curve evolution process mainly consists of two cycles: one cycle for contour pixel evolution and a second cycle for contour pixel smoothness. Based on the curve evolution process, we introduce two tracking stages for the FTCNB method. For coarse tracking stage, the speed function is defined by region competition term combining color and texture features. For contour refinement stage which requires higher tracking accuracy, the likelihood models of the Maximum a posterior (MAP) expressions are incorporated for the speed function. Both the tracking and refinement stages utilize the fast two-cycle curve evolution process with the narrow perception of background regions. With these definitions, we conduct extensive experiments and comparisons for the proposed method. The comparisons with other baseline methods well demonstrate the effectiveness of our work.

1. Introduction

Road image sequences contain lots of useful information of the world, which can be applied to traffic scenes' reconstruction [1]. These road image sequences are usually captured by cameras mounted on the roofs of moving cars. Numerous traffic elements exist in these sequences, such as the moving vehicles. The problem of locating and extracting these traffic elements is important in the communities of signal processing and computer vision.

Matching an object contour to the given template is a core part in many visual tasks [2,3]. In addition to single-image tasks, the problem of object contour tracking in image sequences has been studied extensively by researchers. Fan et al. [4] propose a model adaptation framework that combines matting and tracking. The scribbles for matting are based on coarse tracking results. Accurate boundaries are then obtained based on short-term features and long-term appearances. However, the results obtained with this method are poor in the conditions of cluttered backgrounds. In [5], Zhang et al. use a novel layered directed acyclic graph (DAG) based framework for detection and segmentation of the primary object in video. This framework depends on the generation of object proposal. The dynamic programming solution is comparatively time consuming. Bai et al. [6] present the Video SnapCut, which is based on the collaboration of local classifiers. The tracking paradigm naturally supports local user editing. However, this method involves too many user interactions and obtains poor results in cluttered backgrounds. The SeamSeg method presented in [7] utilizes the patch seams across frames. Their energy function takes into account the similarity of patches along the seams, temporal consistency of motion and spatial coherency of seams. Although the SeamSeg method is computationally efficient, it fails to handle complete occlusions for only two adjacent frames are considered at one propagation. For example, it cannot track multiple cars well in vehicle image sequences. Varas et al. [8] propose a video object segmentation approach that extends a particle filter to a region-based image representation. The prediction step applies a co-clustering between the previous image object partition and a partition of the current one, which allows us to tackle the evolution of non-rigid structures. Nevertheless, the experimental results are poor when the background is close in appearance features with the object region. Moreover, it cannot handle the occlusion

* Corresponding author.

http://dx.doi.org/10.1016/j.image.2016.03.001
0923-5965/

cases. More literatures for object detection and tracking in images and videos are described in [9–11].

Level set method is a powerful tool [12–15] for object tracking in image sequences due to its flexibility with respect to topological changes of the contours. The basic concept of this method assumes the boundary curve to be the zero level set of a function defined over the entire image domain. The boundary curve is then evolved by the predefined speed function. Conventionally, the level set implementation of the foreground curve evolution is based on the solution of certain partial differential equations (PDEs). However, the computational complexity of this PDE-based solution has limited the use of the level set method.

Two different approaches for the curve evolution of level set method have been examined by researchers. In the first approach, the update of the level set function is implemented globally over the entire image domain. The Chan–Vese (CV) model [16] is a typical example. The CV model applies the Mumford–Shah function [18] into the level set framework, which gives a piecewise constant description of an image. The curve evolution is driven by an energy function and incorporates a "fitting" term. Originally, the CV model is applied to gray scale images. Chan et al. [17] extend the traditional scalar-field CV model into a vector-field representation. Nevertheless, this model still ignores the consistency of spatial structure of the multichannel images. Gibou et al. [20] separate the CV model into two different cycles and connected the level set algorithm with K-means, as well as nonlinear diffusion preprocessing. The solving of the CV model using this method updates the level set function over the whole image grid.

The second approach focuses on the local evolution of the curve, and only solves the level set equations around the neighborhood of the zero level set. Peng et al. [19] propose a method to reinitialize the Hamilton–Jacobi equation for a few steps for each iteration of level set curve evolution. The level set curve evolution is performed on a localized region around the contours. Shi et al. [21] propose an approximation method of the level-set-based curve evolution without solving PDEs. The curve evolution is implemented by a switching mechanism between two linked lists of boundary pixels. The curve evolution process consisted of two cycles: one for data dependent terms and another for smoothness regularization. Although the curve evolution in Shi's method is defined locally around the zero level set, the extraction of the foreground and background features is based on the global image domain. As a result, the foreground and background regions are confined to possess unique color and texture information in the experimental datasets, *e.g.* gray-scale image sequences with static backgrounds. This has limited its usability. The method we propose in this paper falls into the second category.

In the present paper, we propose the fast two-cycle level set method with narrow perception of background (FTCNB) for road image sequences with dynamic background. This method consists of two tracking stages: (1) the object tracking stage and (2) the contour refinement stage. In the first stage, our algorithm is implemented based on the speed function that combines mid-level color and texture features. The object tracking is fast in speed and has adequate tracking accuracy. In the second stage, a maximum a posterior (MAP) expression is defined to infer the speed function. This is necessary for the generation of high-accuracy tracking contours. Both of these stages utilize a fast two-cycle curve evolution process with the narrow perception of background regions. The curve evolution process mainly consists of two cycles. The first cycle is designed for contour pixel evolution, while the second cycle is for contour pixel smoothness. The flow diagram of the proposed method is shown in Fig. 1. The main contributions of this paper are summarized as follows:

- *The computations of the level set speed functions*: In the object tracking stage, the speed function is defined by the region competition term that combines color and texture descriptors. The mid-level visual cues such as superpixels are applied to generate the color descriptors. The edge histogram descriptor is utilized as the texture descriptor. To obtain more accurate contours, the speed function is defined by the likelihood models of MAP expressions in the contour refinement stage.
- *FTCNB level set method for outdoor image sequences*: In both the object tracking and contour refinement stages, the proposed algorithms confine the perception of background to the close regions near the foreground contour. The two-cycle curve evolution algorithms are then implemented to achieve the resulted contours.
- *Combination of both spatial and temporal information from the image sequences*: The mid and low-level spatial cues are utilized in the object tracking and contour refinement stages, respectively. The temporal information between frames is considered for the contour refinement stage, *i.e.* the optical flows.

This paper builds upon and extends our previous work in [22], with a more detailed description of the algorithms and additional evaluation results. The remaining parts of this paper are organized as follows: Section 2 shows the object tracking stage of FTCNB method for coarse tracking. The data structure of the FTCNB level set method is introduced firstly. Motivated by the energy function, the speed function is defined and computed. The curve evolution process is then implemented. Section 3 shows

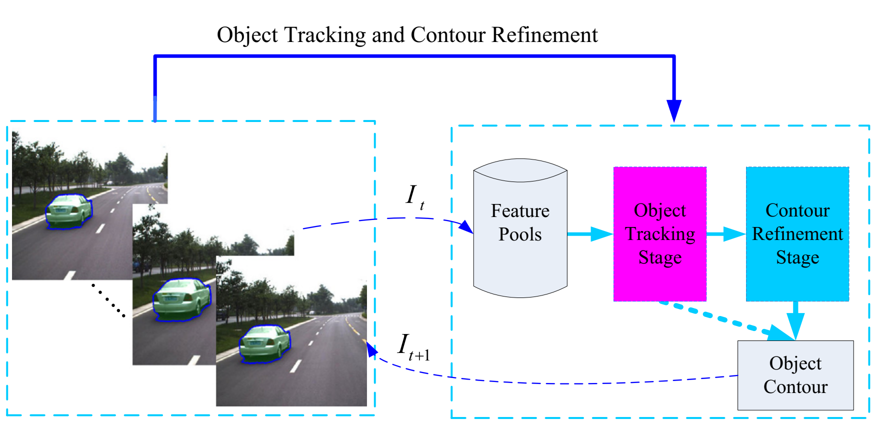

Fig. 1. Flow diagram of FTCNB level set method for object tracking and contour refinement.