Neurocomputing 168 (2015) 336–347

Contents lists available at ScienceDirect

Neurocomputing

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

Video-based road detection via online structural learning

Yuan Yuan [a], Zhiyu Jiang [a,b], Qi Wang [c,*]

[a] Center for OPTical IMagery Analysis and Learning (OPTIMAL), State Key Laboratory of Transient Optics and Photonics, Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences, Xi'an 710119, Shaanxi, PR China
[b] University of the Chinese Academy of Sciences, 19A Yuquanlu, Beijing 100049, PR China
[c] School of Computer Science and Center for OPTical IMagery Analysis and Learning (OPTIMAL), Northwestern Polytechnical University, Xi'an 710072, Shaanxi, PR China

ARTICLE INFO

Article history:
Received 5 February 2015
Received in revised form
27 April 2015
Accepted 25 May 2015
Communicated by Xiaofei He
Available online 5 June 2015

Keywords:
Computer vision
Machine learning
Road detection
Structural SVM
Online updating
Road boundary

ABSTRACT

Video-based road detection is a crucial enabler for the successful development of driver assistant and robot navigation systems. But reliable detection is still on its infancy and deserves further research. In order to adapt to the situation consisting of environmental varieties, an online framework is proposed focusing on exploring the structure cue of the feature vectors. Through the structural support vector machine, the road boundary and non-boundary instances are firstly discriminated. Then they are utilized to fit a complete road boundary. After that, the road region is accordingly inferred and the obtained results are treated as ground truth to update the learned model. Three contributions are claimed in this work: online-learning updating, structural information consideration, and targeted sampling selection. The proposed method is finally evaluated on several challenging videos captured by ourselves. Qualitative and quantitative results show that it outperforms the other competitors.

1. Introduction

According to one recent report [1], road traffic injury remains an important health problem for the public. The total number of road traffic deaths keeps unacceptably high at 1.24 million per year, while the primary cause is unacceptably due to the driver's inattention and tiredness. To alleviate this situation, *Driver Assistance System* (DAS) [2–4] is developed and equipped, with the hope that it can serve as an autonomous reminder and guidance for the drivers. Among the various techniques enabling the DAS, road detection is the fundamental one, because it is the first step for a vehicle to become moveable and many other intelligent maneuvers are based on it. For example, *Lane Departure Warning* (LDW) [5,6], Lane Centering [7], and even full autonomous driving [8] rely on the results of road detection. Moreover, it can provide a significant contextual cue for target detection (e.g, vehicle or pedestrian) [9–12] and act as the prerequisite for robot navigation in an outdoor environment, which is widely researched in artificial intelligence and computer vision.

Because of its practical and theoretical importances, road detection has been thoroughly investigated in recent years. According to the types of sensing modalities used for this purpose, existing methods can be categorized into active sensor based and passive sensor based. For the *active sensor based* methods, the sensors project certain kinds of radiative lights and measure the reflection from its projection. Typical examples include *Light Detection And Ranging* (LIDAR) and *Radio Detection And Ranging* (RADAR). Several active sensors have been widely used for road understanding and great progress has been made since the DARPA Grand Challenge and Urban Challenge [13].

However, due to the restriction of limited perceptual range by the active sensors, and the risk of inter-vehicle inference or pollution to the environment, the *passive sensor based* methods have a tendency of dominating the trend because of their noninvasive characteristic. To be specific, the passive sensors obtain useful information from the environment by capturing the reflection of sun light or other artificial lights. This kind of method can provide intuitive understanding of all the surrounding environment and deliver more meaningful cues, which are crucial for the development of future intelligent transportation systems in mixed traffic conditions [14]. As for the sensor, video cameras that provide the visual data are the most frequent choice. Therefore, the term "video-based" is interchangeably used with "passive sensor based" for simplicity in the following sections.

In this paper, we address the problem of video-based road detection utilizing an online strategy. The major focus is on

* Corresponding author.
E-mail address: crabwq@nwpu.edu.cn (Q. Wang).

http://dx.doi.org/10.1016/j.neucom.2015.05.092
0925-2312/

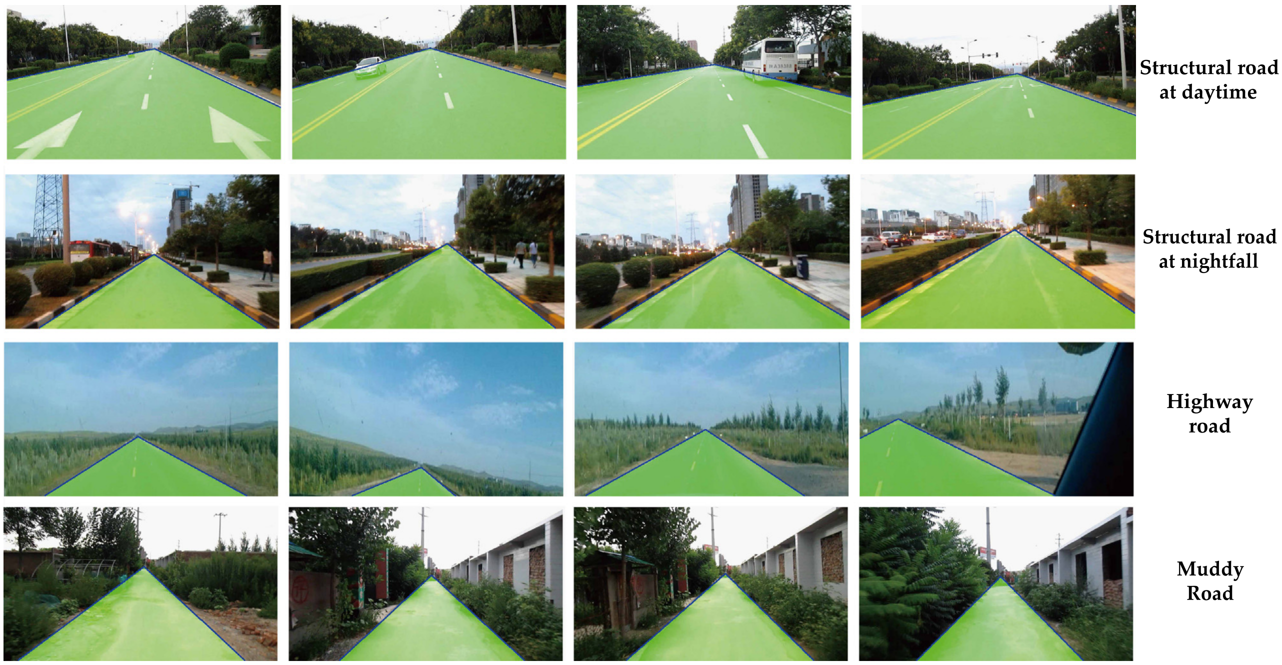

Fig. 1. Typical road detection results of the proposed method. From top to bottom, each row represents the detection results of a specific kind of scenery.

exploring the structural information of the input data through the structural SVM (SSVM). At the same time, the learned model is online updated to adapt to the changing environment. Fig. 1 shows some typical detection examples using the proposed method.

1.1. Related works

Since the presented work belongs to the passive sensor based type, we only review the video-based methods. With respect to the different emphasis on the prior knowledge, road detection can be divided into three groups: *model-based*, *feature-based*, and *learning-based*.

(1) *Model-based method* tends to have an assumption of road shape, which is actually treated as road model. Then the aim is to find the fittest parameters under the model assumption. Several strategies of model fitting [15–19] are used to get the road model. Oniga et al. [15] fitted a quadratic road model by RANSAC approach and the fitting result was refined by a region growing-like process so as to determine the road surface. Sappa et al. [16] proposed Least Square Estimation (LSE) based approach to fit a model for the road surface. Fardi et al. [17] utilized Hough domain to determine the road borders after using the Gaussian pyramid technique to model the scale information. Borkar et al. [18] employed RANSAC to eliminate outliers caused by noise and artifacts in the road and Kalman filter was finally used to smooth the road boundaries. Sawano and Okada [19] utilized an internal energy based on the tendency of a control point resisting changes in its state of motion in an image space, to represent the road model. Although *Model-based methods* can accurately determine the road region given a proper road model, it may be invalid to face the situations where road shapes change as the vehicle is moving. Therefore, it is difficult to find an appropriate model for unstructured roads with inconstant conditions.

(2) *Feature-based method* relies upon the extraction of image features to detect road boundaries and road region. The features such as color, gradient and texture are commonly used to measure local neighborhoods and a likelihood function is formulated by feature clustering [20], threshold segmentation [21] or region growing approach [22] to obtain the road region. For example, He et al. [23] assumed that the color components of road surfaces obey the Gaussian distribution and the road areas were detected based on the full color features. Sotelo et al. [24] utilized the Hue-Saturation-Intensity (HSI) color features for segmentation to model the road pattern. Alvares et al. [25] employed an illuminant-invariant, which was converted from the RGB space, as the feature space to accomplish the road detection task. The main advantages of the *feature-based method* are that it is insensitive to the shape of roads and little previous knowledge is needed. But it is sensitive to shadows and other illumination changes.

(3) *Learning-based method* [25–27] generally makes use of a trained neural network or classifier to distinguish between the road region and non-road region. Such methods are independent of special road markings and are capable of dealing with non-homogeneous road appearance, if the characteristics of road or non-road regions are properly represented by the feature space. Alvarez and Lopez [25] introduced a shadow-invariant feature space and it was used along with a likelihood-based classifier which was online learned to achieve road segmentation. Son et al. [27] constructed a probabilistic road model by supervised training and a posteriori probability based on visual information was then utilized to extract the road region. For *learning-based method*, although less prior knowledge is needed, it heavily relies on the training sets and training strategies. But unfortunately, most of the classifier and neural network are trained once, unable to adapt to the varieties of the environment.

Apart from the three types, most road detection problems can be successfully interpreted using a variant of the three above approaches or a combination of them. The proposed method in this work belongs to the learning based prototype, while taking advantages of advanced features and road boundary fitting.

1.2. Proposed framework

Though many works have been proposed, most of them are based on the assumption that the road area is consistent in intensity or color. However, in real-life environments, this assumption might fail because the intensity often varies a lot as the vehicle or robot is moving. Moreover, the shadows and occlusions