Infrared Physics & Technology 76 (2016) 261–270

Contents lists available at ScienceDirect

Infrared Physics & Technology

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

Recognizing pedestrian's unsafe behaviors in far-infrared imagery at night

Eun Ju Lee, Byoung Chul Ko *, Jae-Yeal Nam

Department of Computer Engineering, Keimyung University, Sindang-dong, Dalseo-gu, Daegu 704-701, Republic of Korea

HIGHLIGHTS

- Recognizing pedestrian's behaviors using thermal image from moving vehicle at night.
- Designing the light convolutional neural networks with random forest classifier.
- Generating the pedestrian's unsafe behavior (PUB) dataset using thermal camera.
- Behavior recognition accuracy is higher than that of related algorithms.

ARTICLE INFO

Article history:
Received 11 February 2016
Accepted 8 March 2016
Available online 10 March 2016

Keywords:
Unsafe behavior recognition
Advance driver assistance system
Convolution neural network
Genetic algorithm
Spatial pyramid pooling
Boosted random forest

ABSTRACT

Pedestrian behavior recognition is important work for early accident prevention in advanced driver assistance system (ADAS). In particular, because most pedestrian-vehicle crashes are occurred from late of night to early of dawn, our study focus on recognizing unsafe behavior of pedestrians using thermal image captured from moving vehicle at night. For recognizing unsafe behavior, this study uses convolutional neural network (CNN) which shows high quality of recognition performance. However, because traditional CNN requires the very expensive training time and memory, we design the light CNN consisted of two convolutional layers and two subsampling layers for real-time processing of vehicle applications. In addition, we combine light CNN with boosted random forest (Boosted RF) classifier so that the output of CNN is not fully connected with the classifier but randomly connected with Boosted random forest. We named this CNN as randomly connected CNN (RC-CNN). The proposed method was successfully applied to the pedestrian unsafe behavior (PUB) dataset captured from far-infrared camera at night and its behavior recognition accuracy is confirmed to be higher than that of some algorithms related to CNNs, with a shorter processing time.

1. Introduction

According to the report of [1], if collision speed reduced from 50 to 31 km/h, fatality risk reduced from 33% to 5% when assuming a braking force 10 m before impact. For this reason, radar and light detection and ranging (LIDAR) and radar sensors have been using for pedestrian detection in advanced driver assistance system (ADAS) for avoiding pedestrian-vehicle collision for a long time. Although these sensors give the reliable sensing results, they are expensive and unable to distinguish what is the detected obstacle. Therefore, camera-based pedestrian has been a significant progress for the past decade. As the results of the study, Mobileye Company [2] has introduced artificial vision technologies related pedestrian's safety such as pedestrian collision warning system. However, current pedestrian detection systems are designed conservatively to give warning or control automatic braking in case pedestrian is detected [3] in the lane of driving. However, if a vehicle recognizes the pedestrian's unsafe behavior (PUB) in advance before pedestrian is approaching inside of driving lane, the driver can control the vehicle more effectively. For example, if a pedestrian is just standing at the curbside, the vehicle does not need to stop. In contrast, a pedestrian is positioning outside of the curbside, but if his behavior is running, the vehicle should reduce the speed. Therefore, recognizing pedestrian's behavior is a major emerging issue in ADAS to reduce the pedestrian fatality rate.

Most traditional methods of recognizing behavior of pedestrian use motion information in videos because it often provides discriminative cues for behavior classification [4]. However, motion information of pedestrian cannot be estimated exactly even if ego-motion is compensated owing to following reasons; the severe

* Corresponding author.
E-mail addresses: frigia@kmu.ac.kr (E.J. Lee), niceko@kmu.ac.kr (B.C. Ko), jynam@kmu.ac.kr (J.-Y. Nam).

http://dx.doi.org/10.1016/j.infrared.2016.03.006
1350-4495/

vibration of vehicles, different speed of vehicles, and pedestrian distortion caused by perspective camera view. In addition, several types of behaviors, such as 'standing', 'sitting', and 'lying-down', are static in nature road as shown in Fig. 1. Therefore, this paper focused on recognizing dynamic behaviors of pedestrian including safe and unsafe behaviors such as 'standing', 'walking', 'running', 'sitting', 'looking-back', and 'lying-down' from a single image.

The static behaviors of a pedestrian are valuable cues for inferring the pedestrian's intention when motion information is not available. However, the underlying premises need to be addressed before reliable behavior recognition using only a single image. Because the pedestrian detectors [3–5] are reliable enough to correctly localize pedestrian in a single image and pedestrian detection is beyond the scope of this research, we assume that pedestrian regions are detected correctly.

For many years, computer vision based ADAS technologies have been researched using a normal charge-coupled device (CCD) camera in daytime. However, because most pedestrian-vehicle accidents occur between 6 p.m. and 8 a.m., and the rate of pedestrian fatalities is highest between 4 a.m. and 6 a.m. due to inebriated drivers, inebriated pedestrians, drowsiness, and poor visibility [6], a CCD camera is ineffective in environments with poor illumination at night. To solve the illumination problem, near infrared (NIR) cameras are used in combination with an illuminator. However, it cannot distinguish pedestrian from the background when a pedestrian is located in front of a car's headlights as the similar limitation with CCD cameras.

Therefore, this study investigate recognize pedestrian's unsafe behaviors in a single image captured from far-IR (FIR) camera instead of CCD or NIR camera, particularly at night and in outdoor environments. In general, FIR cameras can detect relative differences in the amounts of thermal energy emitted or reflected from different body parts of a pedestrian in a scene, regardless of the illumination [7]. Therefore, it allows robust detection of pedestrian bodies in both indoor and outdoor environments.

1.1. Related works

Pedestrian's behavior recognition in static images is a very challenging problem because of the significant amount of pose, different viewpoint, pose deformation, and illumination variation [8]. In [4], authors categorized action recognition in still images into four main categories according body pose; model-based approaches, example-based approach, pictorial-structure-based approaches, and poselet-based approaches. Among four categories, state-of-the-art approaches typically use poselet-based methods [9–11] to infer human action. Poselets are represented as a configuration of body part locations such as head, leg, and shoulders, and it is expressed as a latent variable that is used for action recognition. Each poselet is trained separately using pattern classifier from specifically annotated 3D images and casts a vote for an object hypothesis for testing. Even though this method achieved good recognition performance, it has following limitations [4];

- It requires careful selection of color and shape descriptor for different databases.
- The optimal way to fuse of multiple features for consistent results remains an open problem.
- It requires a significant amount of computational time for poselet detection and recognition.

In addition, conventional researches related to action recognition have been using the PASCAL databases, which consist of annotated of human actions (e.g., Photographing, playing music, riding horsing, using computer) including the 2D locations of key points and kinds of pixel-level labels of image patches [12]. Therefore, PASCAL dataset and previous researched using this data is not appropriate to recognize pedestrian's unsafe behaviors.

To recognize pedestrian's behavior in a real road environment, some methods incorporate 2D pedestrian detections and motion and trajectory matching information. Jan et al. [13] introduced a data-based neural network to approximately partition the classification space nonlinearly in order to achieve an acceptable detection of abnormal activity of human behaviors while reducing computational complexity. However, this study focuses on classifying the human behavior in a car park as only suspicious or unsuspicious. Wang et al. [14] proposed recognizing of abnormal pedestrian's behaviors based on a rule induction classifier with optical-flow feature.

Keller and Gavrila [15] presented a study on pedestrian path prediction and action classification (walking vs. stopping) at short, sub-second time intervals. This study use both Gaussian process dynamical models based on lateral scene flow features and probabilistic hierarchical trajectory matching based on motion feature for path prediction. From the prediction results, predicting the position of pedestrians walking toward the road curbside or stopping before the road curbside is decided.

Although these methods showed reasonable results on a few safe and unsafe behaviors of pedestrian, they mainly depend on the motion information. However, motion estimation in moving vehicle is difficult due to speed and vibration of vehicle. In addition, these methods have the limitation for recognizing static pedestrian's behaviors such as 'standing', 'sitting', and 'lying-down'.

Fig. 1. Several types of static pedestrian's behaviors associated with specific behaviors: (a) standing on the road, (b) sitting on the road, and (c) lying-down on the road.