Transportation Research Part C 49 (2014) 32–42

Contents lists available at ScienceDirect

Transportation Research Part C

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

Real time detection of driver attention: Emerging solutions based on robust iconic classifiers and dictionary of poses

G.L. Masala *, E. Grosso

POLCOMING Department at University of Sassari, Italy

ARTICLE INFO

Article history:
Received 26 May 2014
Received in revised form 12 September 2014
Accepted 13 October 2014
Available online 5 November 2014

Keywords:
Automotive applications
Monitoring of driver attention
Driver assistance systems
Neural networks

ABSTRACT

Real time monitoring of driver attention by computer vision techniques is a key issue in the development of advanced driver assistance systems. While past work mostly focused on structured feature-based approaches, characterized by high computational requirements, emerging technologies based on iconic classifiers recently proved to be good candidates for the implementation of accurate and real-time solutions, characterized by simplicity and automatic fast training stages.

In this work the combined use of binary classifiers and iconic data reduction, based on Sanger neural networks, is proposed, detailing critical aspects related to the application of this approach to the specific problem of driving assistance. In particular it is investigated the possibility of a simplified learning stage, based on a small dictionary of poses, that makes the system almost independent from the actual user.

On-board experiments demonstrate the effectiveness of the approach, even in case of noise and adverse light conditions. Moreover the system proved unexpected robustness to various categories of users, including people with beard and eyeglasses. Temporal integration of classification results, together with a partial distinction among visual distraction and fatigue effects, make the proposed technology an excellent candidate for the exploration of adaptive and user-centered applications in the automotive field.

1. Introduction

Since late 90s, undesirable or unusual driver conditions have been clearly identified as a primary cause of car crashes and road deaths (Kircher et al., 2002; European Transport Safety Council, 2001). This problem attracted the interest of the scientific community, which has begun to study the development of intelligent and adaptive systems, namely Advanced Driver Assistance Systems (ADAS), suitable to monitor the diver's state of vigilance and give real-time support in accident avoidance (Liang et al., 2007; Batista, 2005).

As pointed out by Liang and Lee (2014), the nature of driver inattention can vary: fatigue and related symptoms like drowsiness and frequent nodding are very common in real cases but distraction from safe driving can also have a visual or cognitive cause.

Visual distraction has often to do with the on-board presence of electronic devices or tools like mobile phones, navigation and multimedia systems, requiring active control from the driver (for example pushing buttons or turning knobs); visual distraction can be also related to the presence of salient visual information away from the road, thus causing spontaneous

* Corresponding author at: POLCOMING: Department of Political Science, Communication, Engineering and Information Technologies, Computer Vision Laboratory, Viale Mancini, 5, 07100 Sassari, Italy. Tel.: +39 (79) 229486; fax: +39 (79) 229482.

E-mail address: gilmasala@uniss.it (G.L. Masala).

http://dx.doi.org/10.1016/j.trc.2014.10.005
0968-090X/© 2014 The Authors. Published by Elsevier Ltd.

off-road eye glances and momentary rotation of the head. Cognitive distraction happens whenever the mind of the driver is not sufficiently focused on the critical task of safe driving; symptoms of cognitive distraction are less apparent, and difficult to be detected or quantified by objective indicators. Most of times the analysis of cognitive distraction is therefore based on long behavioral patterns and sophisticated statistical techniques (Liang and Lee, 2014).

Focusing on fatigue and visual distraction, the paper investigates the design and the development of a fully automated driver assistance system based on advanced techniques coming from image analysis and related fields like pattern recognition and biometrics (Zhao et al., 2003).

In previous studies, computer vision techniques have been often proposed to detect driver attention (Batista, 2005; Singh and Papanikolopoulos, 1999) both by standard and day-night infrared cameras. In particular, these techniques have been adopted to detect signs of visual distraction, like off-road gaze direction and persistent rotation of the head, and changes in the facial features which characterize persons with reduced alertness due to fatigue: longer blink duration, slow eyelid movement, small degree of eye opening, nodding, yawns and drooping posture are among the most interesting conditions which has proved to be captured by vision-based approaches (Bergasa et al., 2008).

A common processing scheme, well discussed in Senaratne et al. (2011, 2007) includes the following steps:

- face localization;
- localization of facial features (e.g. eyes or mouth);
- estimation of specific cues related to fatigue or distraction;
- fusion of cues in order to determine the global attention level.

Concerning face localization, very robust techniques based on neural networks have been developed in late 90s (Rowley et al., 1998; Sung and Poggio, 1998). In 2004 Viola and Jones (Viola and Jones, 2004) proposed a new high performance algorithm based on integral images and robust classification; this algorithm is a de facto standard for real-time applications. Both the above approaches belong to the image-based subclass of the face detection techniques. More recently also feature-based approaches demonstrated a reasonable level of efficiency. In particular, Particle Swarm Optimization (Kennedy and Eberhart, 1995) has been proposed for locating and tracking a limited number of facial landmarks.

Research on facial features extraction mainly focused on eyes and mouth (Zhao et al., 2003); Gabor and SVM techniques have been successfully proposed to this aim (Senaratne et al., 2007). In order to work under low light conditions, researchers also proposed the use of infrared illuminators, exploiting high reflection of the pupils (Senaratne et al., 2011); as noted in Lenskiy and Lee (2012), however, IR based approaches show malfunctions during daytime and require the installation of additional hardware.

It is worth noting here that most of the literature defines the PERCLOS as the main cue for the estimation of driver's fatigue. PERCLOS is a measure of the time percentage during which eyes remain closed 80% or more; in order to compute this cue, every image frame is usually classified into two classes (closed eyes or open eyes): k-NN techniques, SVMs and Bayes approaches have been successfully applied to this purpose (Rowley et al., 1998). Other cues commonly used are head pose, eye blinking detection (Lenskiy and Lee, 2012), slouching frequency and postural adjustment. To the aim of this work the estimation of the head pose certainly represents the most interesting issue (Sung and Poggio, 1998); this information can be derived by applying both 2D and 3D approaches (Murphy-Chutorian and Trivedi, 2009).

Overall, previous studies show that the problem of detecting visual distraction and fatigue can be faced with fairly good results in driving simulators or constrained conditions. However, the application on a real moving vehicle presents new challenges like changing backgrounds and sudden variations of lighting. Moreover, a useful system should guarantee real time performance and quick adaptability to a variable set of users and to natural movements performed during driving.

In order to tackle the real problem and to reach a sufficient level of accuracy and performance, we propose here a driver assistance system based on robust iconic classifiers. Starting from a preliminary image data reduction step, and from a priori knowledge related to known head poses and known patterns (like, for instance, closed/open eyes), we show that iconic classifiers perform well with respect to changes in pose and facial features configuration, while ignoring unessential details like glasses, hairstyle and lighting conditions. As explained later in the text, the conceptual boundary between raw input data, feature extraction and classification can be somewhat arbitrary; moreover the proper classification of the input data can be heavily influenced by the collection of poses and patterns used in the learning phase. For this reason we propose a binary classification of poses and features, where the collection of possible configurations is simply categorized in "attentive" versus "inattentive" classes.

Following sections are organized as follows: Section 2 briefly introduces the adopted attention model and the fundamental methods applied for the various processing steps; Section 3 details the experimental setup, the data collection phase and experimental results. Finally Section 4 draws some conclusions and analyses possible outcomes of this research.

2. Approach and methods

Even though the adoption and the fusion of different cues usually shows some increase of performance, recent work (Masala and Grosso, 2013) demonstrates that this approach can be efficiently replaced by an alternative "fully iconic" approach, based on a generalized model of the "inattentive driver". This iconic generalization, derived by processing and classifying off-line a sequence or a selected set of images of a generic real user, is denoted here as "dictionary of poses"