Expert Systems with Applications 41 (2014) 1139–1152

Contents lists available at ScienceDirect

Expert Systems with Applications

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

Detecting driver drowsiness using feature-level fusion and user-specific classification

Jaeik Jo [a], Sung Joo Lee [a], Kang Ryoung Park [b], Ig-Jae Kim [c], Jaihie Kim [a,d,*]

[a] School of Electrical and Electronic Engineering, Yonsei University, Seoul 120-749, Republic of Korea
[b] Division of Electronics and Electrical Engineering, Dongguk University, Seoul 100-715, Republic of Korea
[c] Imaging Media Research Center, Korea Institute of Science and Technology, Seoul 136-130, Republic of Korea
[d] Sunway University, 46150 Petaling Jaya, Selangor, Malaysia

ARTICLE INFO

Keywords:
Drowsiness detection system
Blink detection
Eye state classification
Feature-level fusion
User-specific classification

ABSTRACT

Accurate classification of eye state is a prerequisite for preventing automobile accidents due to driver drowsiness. Previous methods of classification, based on features extracted for a single eye, are vulnerable to eye localization errors and visual obstructions, and most use a fixed threshold for classification, irrespective of variations in the driver's eye shape and texture. To address these deficiencies, we propose a new method for eye state classification that combines three innovations: (1) extraction and fusion of features from both eyes, (2) initialization of driver-specific thresholds to account for differences in eye shape and texture, and (3) modeling of driver-specific blinking patterns for normal (non-drowsy) driving. Experimental results show that the proposed method achieves significant improvements in detection accuracy.

1. Introduction

According to the American National Highway Traffic Safety Administration (NHTSA) Liu & Subramanian, 2009, approximately 100,000 accidents per year occur because of driver drowsiness. In response to this mounting problem, methods for detecting driver drowsiness have been intensively studied in the automotive field (Adachi et al., 2008; Bergasa, Nuevo, Sotelo, Barea, & Lopez, 2006; Bhowmick & Chidanand Kumar, 2009; Damousis & Tzovaras, 2008; Dong, Hu, Uchimura, & Murayama, 2011; Eriksson & Papanikotopoulos, 1997; Ersal, Fuller, Tsimhoni, Stein, & Fathy, 2010; Flores & Armingol, 2010; Ince & Yang, 2009; Jimenez-Pinto & Torres-Torriti, 2009; Jo, Lee, Jung, Park, & Kim, 2011; Jo et al., 2010; Kawato & Ohya, 2000; Kircher, Uddman, & Sandin, 2002; Kurylyak, Lamonaca, & Mirabelli, 2012; Li, 2008; Liu, Hosking, & Lenn, 2009; Minkov, Zafeiriou, & Pantic, 2012; Murphy-Chutorian, Doshi, & Trivedi, 2007; Noguchi, Shimada, Ohsuga, Kamakura, & Inoue, 2009; Orazio, Leo, Guaragnella, & Distante, 2007; Panning, Al-Hamadi, & Michaelis, 2011; Parmar, 2002; Patel, Lal, Kavanagh, & Rossiter, 2011; Rongben, Lie, Bingliang, & Lisheng, 2004; Saradadevi & Bajaj, 2008; Shuyan & Gangtie, 2009; Sukno, Pavani, Butakoffand, & Frangi, 2009; Torkkola, Massey, & Wood, 2004; Tran, Craig, Wijesuriya, & Nguyen, 2010; Tsuchida, Bhuiyan, & Oguri, 2010; Uliyar & Ukil, 2012; Vural et al., 2007; Wang, Ding, Fang, & Liu, 2009; Wu & Chen, 2008; Wu & Trivedi, 2007; Yang et al., 2009; Yunq, Meiling, Xiaobing, Xiuxia, & Jiangfan, 2009). As with the detection of driver inattention (Dong et al., 2011; Jo et al., 2011), the methods used for drowsiness detection can be divided into three categories: those based on driving behavior (Ersal et al., 2010; Kircher et al., 2002; Liu et al., 2009; Torkkola et al., 2004; Yang et al., 2009), those based on physiological features (Damousis & Tzovaras, 2008; Patel et al., 2011; Shuyan & Gangtie, 2009; Tran et al., 2010), and those based on visual features (Adachi et al., 2008; Bergasa et al., 2006; Bhowmick & Chidanand Kumar, 2009; Eriksson & Papanikotopoulos, 1997; Flores & Armingol, 2010; Ince & Yang, 2009; Jimenez-Pinto & Torres-Torriti, 2009; Jo et al., 2010, 2011; Kawato & Ohya, 2000; Kurylyak et al., 2012; Li, 2008; Minkov et al., 2012; Murphy-Chutorian et al., 2007; Noguchi et al., 2009; Orazio et al., 2007; Panning et al., 2011; Parmar, 2002; Rongben et al., 2004; Saradadevi & Bajaj, 2008; Sukno et al., 2009; Tsuchida et al., 2010; Uliyar & Ukil, 2012; Vural

Abbreviations: ASM, active shape model; CCA, canonical correlation analysis; ECD, eye closure duration; EER, equal error rate; ESD-Value, eyelid state detection value; FEC, frequency of eye closure; GLCM, grey-level co-occurrence matrix; GRBF, Gaussian radial basis function; LBP, local binary pattern; LDA, linear discriminant analysis; NHTSA, national highway traffic safety administration; PCA, principal component analysis; PERCLOS, percentage of eye closure; PHOGs, pyramid histogram of oriented gradients; ROI, region of interest; SVM, support vector machine.

* Corresponding author. Address: Sunway University, 46150 Petaling Jaya, Selangor, Malaysia. Tel.: +82 2 2123 2869; fax: +82 2 312 4584.

E-mail addresses: jaeik@yonsei.ac.kr (J. Jo), sungjoo@yonsei.ac.kr (S.J. Lee), parkgr@dongguk.edu (K.R. Park), kij@imrc.kist.re.kr (I.-J. Kim), jhkim@yonsei.ac.kr (J. Kim).

0957-4174/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.eswa.2013.07.108

et al., 2007; Wang et al., 2009; Wu & Chen, 2008; Wu & Trivedi, 2007; Yunq et al., 2009). Methods based on driving behavior detect drowsiness by monitoring vehicle speed, lane observation, steering, acceleration, braking, and gear changes. The main drawback of these methods is that their accuracy depends on the individual characteristics of the vehicle and its driver. In contrast, methods based on physiological features detect drowsiness by measuring heart rate and brain activity. Although these methods show good detection accuracy, they also depend on peripheral measuring equipment that must be attached to the driver's body. Finally, methods based on visual features detect drowsiness using information obtained from a camera, and thus neither depend upon vehicle or driver characteristics nor require intrusive measuring equipment. As such, visual feature-based methods have emerged as the preferred avenue for research.

A great number of "visual feature-based methods" for drowsiness detection have been proposed and studied. Among the visual features used by these methods are eye state information (Adachi et al., 2008; Bergasa et al., 2006; Bhowmick & Chidanand Kumar, 2009; Eriksson & Papanikotopoulos, 1997; Flores & Armingol, 2010; Ince & Yang, 2009; Jo et al., 2010, 2011; Kurylyak et al., 2012; Li, 2008; Minkov et al., 2012; Noguchi et al., 2009; Orazio et al., 2007; Panning et al., 2011; Parmar, 2002; Sukno et al., 2009; Tsuchida et al., 2010; Uliyar & Ukil, 2012; Wang et al., 2009; Wu & Trivedi, 2007; Yunq et al., 2009), head movement (Kawato & Ohya, 2000; Murphy-Chutorian et al., 2007), yawning (Rongben et al., 2004; Saradadevi & Bajaj, 2008) and facial expression (Jimenez-Pinto & Torres-Torriti, 2009; Vural et al., 2007). Methods using eye state to measure driver drowsiness have generally done so by calculating values such as the percentage of eye closure (PERCLOS) Wierwille, Ellsworth, Wreggit, Fairbanks, & Kim, 1994, eye closure duration (ECD), and the frequency of eye closure (FEC) Orazio et al., 2007. Methods using head movement measure drowsiness by estimating head posture. Methods based on yawning locate the driver's mouth and then train the system with images of normal and yawning mouths. Methods that use facial expressions to detect drowsiness generally combine several facial cues, such as yawning, blinking, and eyebrow rising. It should be noted that some of the above features suffer deficiencies in timing: yawning generally occurs well before drowsiness sets in, and head nodding generally occurs after the driver falls asleep. Thus, methods based on these features cannot detect the onset of drowsiness precisely, and are therefore unsuitable for reliable detection systems. On the other hand, eye status information is well-suited for such systems, since the closing of eyes and the appearance of unusual patterns of blinking have been shown to directly indicate the onset of drowsiness. Indeed, methods based on eye status information have already shown superior accuracy in detecting drowsiness (Vural et al., 2007).

The core technology used by these methods is an algorithm for classifying eye state (i.e., as open or closed). Prior work on such algorithms can be categorized into three methods: texture-based methods (Bergasa et al., 2006; Eriksson & Papanikotopoulos, 1997; Flores & Armingol, 2010; Jo et al., 2010, 2011; Kurylyak et al., 2012; Li, 2008; Minkov et al., 2012; Orazio et al., 2007; Panning et al., 2011; Parmar, 2002; Uliyar & Ukil, 2012; Wu & Trivedi, 2007), shape-based methods (Adachi et al., 2008; Ince & Yang, 2009; Noguchi et al., 2009; Sukno et al., 2009; Tsuchida et al., 2010; Wang et al., 2009), and combined texture–shape methods (Bhowmick & Chidanand Kumar, 2009; Yunq et al., 2009). Texture-based methods extract texture features for eye state classification using various feature extraction methods. Minkov et al. (2012) described a blinking detection method using the following features: raw-image intensities, the magnitude of the responses of Gabor filters, the pyramid histogram of oriented gradients (PHOGs), and optical flow. Classification was achieved through a support vector machine (SVM) using Gaussian radial basis function (GRBF) kernels. Jo et al. (2011) introduced an eye state classification method that uses the combination of appearance and statistical features. The appearance features are extracted using principal component analysis (PCA) and linear discriminant analysis (LDA), and the statistical features are acquired using the sparseness and kurtosis of the histogram from the eye-edge image. Uliyar and Ukil (2012) proposed a method based on canonical correlation analysis (CCA) coupled with local binary pattern (LBP) histogram features calculated from the input eye image. Their experiments show that the coupling of these features results in 10–12% improvement in eye state classification accuracy, compared to methods using normalized intensity-based features. Panning et al. (2011) introduced an algorithm for eye state classification using the eyelid state detection value (ESD-Value) calculated by comparing the pixel values in the region of interest (ROI) with an experientially selected threshold value. In addition to the above, features such as Tensor PCA (Wu & Trivedi, 2007), Gabor response waves (Flores & Armingol, 2010; Li, 2008), frame differencing (Bergasa et al., 2006; Kurylyak et al., 2012), and histogram (Eriksson & Papanikotopoulos, 1997; Parmar, 2002) and edge (Jo et al., 2010; Orazio et al., 2007) of the local eye image, have all received attention in the growing body of research on texture-based methods of eye state classification.

As for shape-based methods, eye state has generally been classified using a measurement of the distance between the upper and lower eyelids. Wang et al. (2009), Sukno et al. (2009), and Noguchi et al. (2009) extracted eye contours using an active shape model (ASM) and measured eyelid distance using landmarks in both eyes. Adachi et al. (2008) carried out a similar measurement, but used two search windows to detect the eyelids.

In the combined methods of classification, both texture and shape features are extracted from the input eye image to increase robustness. Bhowmick and Chidanand Kumar (2009) proposed a method in which texture features such as histogram energy and contrast on grey-level co-occurrence matrix (GLCM) are fused with several different shape features, including Hu's seventh moment, compactness, and top-hat and bottom-hat area ratio and eccentricity. The combined features are then used to train the nonlinear SVM in eye state classification.

Unfortunately, all of the above classification methods have two distinct problems. First, they suffer direct and unavoidable performance degradation when eye localization errors occur, or when the eye region is obstructed (e.g., by eyeglasses). Second, these methods generally apply the same classification thresholds to all drivers, regardless of individual differences in eye shape, scale, and blinking frequency.

To address the first of these problems, we propose a method of eye state classification that uses feature-level fusion of both eyes, instead of only one. This approach eliminates a number of errors attributable to single-eye obstructions. To address problems caused by the application of fixed thresholds, we examine an initial period of driving to calculate a baseline probability that the driver will have open eyes at any given moment, and then classify eye state as open or closed using a maximum a posteriori (MAP) classifier and a user-specific threshold. Further, during the initial period of driving, we obtain the driver's blinking pattern, based on which normal (non-drowsy) driving behavior is learned using a two-dimensional (2D) Gaussian model.

The remainder of this paper is organized as follows: in Section 2, we describe the proposed method, comprising eye detection and tracking, eye state classification, and drowsiness detection; in Section 3, we present experimental results with an image database collected from a vehicle under various conditions; and in Section 4, we provide some conclusions.