Available online at www.sciencedirect.com
ScienceDirect
Pattern Recognition 40 (2007) 2341–2355
PATTERN RECOGNITION
THE JOURNAL OF THE PATTERN RECOGNITION SOCIETY
www.elsevier.com/locate/pr

A visual approach for driver inattention detection

T. D'Orazio[a,*], M. Leo[a], C. Guaragnella[b], A. Distante[a]

[a]Institute of Intelligent Systems for Automation, C.N.R. Via Amendola 122/D-I, 70126 Bari, Italy
[b]DEE-Electrics and Electronics Department, Politecnico di Bari, via E. Orabona 4, Bari 70125, Italy

Received 25 July 2006; received in revised form 6 December 2006; accepted 16 January 2007

Abstract

Monitoring driver fatigue, inattention, and lack of sleep is very important in preventing motor vehicles accidents. A visual system for automatic driver vigilance has to address two fundamental problems. First of all, it has to analyze the sequence of images and detect if the driver has his eyes open or closed, and then it has to evaluate the temporal occurrence of eyes open to estimate the driver's visual attention level. In this paper we propose a visual approach that solves both problems. A neural classifier is applied to recognize the eyes in the image, selecting two candidate regions that might contain the eyes by using iris geometrical information and symmetry. The novelty of this work is that the algorithm works on complex images without constraints on the background, skin color segmentation and so on. Several experiments were carried out on images of subjects with different eye colors, some of them wearing glasses, in different light conditions. Tests show robustness with respect to situations such as eyes partially occluded, head rotation and so on. In particular, when applied to images where people have eyes closed the proposed algorithm correctly reveals the absence of eyes. Next, the analysis of the eye occurrence in image sequences is carried out with a probabilistic model to recognize anomalous behaviors such as driver inattention or sleepiness. Image sequences acquired in the laboratory and while people were driving a car were used to test the driver behavior analysis and demonstrate the effectiveness of the whole approach.

Keywords: Iris detection; Behavior analysis; Neural network; Mixture Gaussian model

1. Introduction

The detection of driver visual attention is very important for developing automatic systems that monitor driver inattention, driver fatigue, and sleepiness. A great number of fatalities occurring in motor vehicles could be avoided if these behaviors were detected and alarm signals were provided to the driver. The literature reports many attempts to develop safety systems for reducing the number of automobile accidents: these systems detect both the "driving" behavior by monitoring lane keeping, steering movements, acceleration, braking and gear changing [1], and also the "driver" behavior by such means as tracking the driver's head and the eye movements, monitoring the heart and breathing rates, the brain activity [2], and recognizing the torso and arm/leg motion [3]. Repeated experiments have shown that among all driver performance and bio-behavioral measures tested, the percentage of eyelid closure over time (Perclos) reliably predicts the most widely recognized psychophysiological index of loss of alertness. In this work we address the problem, crucial for automotive applications, of developing a robust eye recognition algorithm that can be used for detecting the Perclos measure but also for modelling different driver behaviors.

1.1. Related works

Many works presented in the literature propose real time methods for eye tracking based on active infrared (IR) illumination approaches. Some commercial products are coming to the market such as SmartEye and EyeAlert Fatigue Warning System [4,7] and also many Institutions are actively involved in automotive research projects. Different types of IR light sources have been devised to emit non-coherent energy synchronized with the camera frame rate, which generates bright and dark pupil images. Pupils can be detected by a simple thresholding

* Corresponding author. Tel.: + 39 080 5929442; fax: +39 080 5929460.
E-mail address: dorazio@ba.issia.cnr.it (T. D'Orazio).

0031-3203/$30.00
doi:10.1016/j.patcog.2007.01.018

of the difference between the dark and the bright pupil images [9–15]. The success of these active approaches depends on several factors: the brightness and size of the pupils, which are often functions of face orientation; external illumination interference; the distance of the subject from the camera; the need for stable lighting conditions (not strong sun light). In addition, glasses tend to disturb the IR light so much that the red eye effect may appear very weak. In recent years large improvements have been made in miniaturizing the cameras with the compact IR illuminator, in designing configurations that produce diffuse lighting, and in selecting allowable levels of IR irradiation. However, in many cases an initial calibration phase is still required during which the intensity of the active IR illuminators has to be tuned in order to operate in different natural light conditions, multiple reflections of glasses, and variable gaze directions.

Alternative approaches that use standard cameras with classical algorithms for eye detection in cluttered images have also been investigated [8]. SeeingMachine [5] proposes a couple of stereo cameras to determine the 3D position of matching features on the driver face. Starting from an initial calibration of the subject in a green room, the extracted 3D features are then used to capture the 3D pose of the person's face as well as the eye gaze direction, blink rates and eye closure [6]. Instead, in our work we investigate the subject of eye detection with monocular images within the visual spectrum and normal illumination. Generally the eye detection algorithms in cluttered images require two steps: locating face to extract eye regions and then eye detection from eye windows. The face detection problem has been addressed with different approaches: neural networks, principal components, independent components, skin color based methods, face models [16–18]. Each imposes some constraints: frontal view, expressionless images, limited variations of light conditions, hairstyle dependence, uniform background, and so on. An exhaustive review has been presented in Ref. [19].

Many works on eye or iris detections, assume either that eye windows have been extracted or rough face regions have already been located [20–22,24–27]. In Ref. [21], the eye detection method is performed within the possible eye region of the candidate face field. In this case it can be applied only after a face detection system has extracted a small number of candidates for eye regions. Left and right eye templates were used to detect eyes by a method that is unaffected by slight rotation, scaling, and translation (up to 10%). The algorithm proposed in Ref. [20] requires the detection of face regions in order to extract intensity valleys. Only at this point do the authors apply a template matching technique to extract iris candidates. The authors deal with the difficult problem of face region extraction both in intensity images and in color images.

Skin color models are strictly related to the considered images and cannot be so general as to be applied in every light condition and with different colors of skin. Region-growing methods or head contour methods on intensity images require strong constraints such as plain background. In Refs. [22,24] the first step is also a face detection algorithm, based on skin color segmentation of the input image with the constraints of there being only one face and a simple background; then the facial feature segmentation is based on gray value reliefs in Ref. [22] or on template matching of edge and color features in Ref. [24]. A probabilistic framework is used in Ref. [23] for locating precisely eyes in face areas extracted formerly by using a face detector. In Ref. [25] linear and non-linear filters have been used for eye detection: oriented Gabor wavelets form an approximation of the eye in gray level images; non-linear filters are applied to color images to determine the color distribution of the sclera region. In both cases a face detection step is applied, which assumes the face as the most prominent flash tone region in the image. The same algorithm has been used in Ref. [26] for tracking iris and eyelids in video.

In Ref. [27], lip and skin color predicates are used as a first step to segment lip region and skin regions in the image: the two holes above the lip region that satisfy some fixed size criteria are selected as the candidate for the eyes region. A hierarchical strategy is applied to track the eyes in a video sequence and evaluate the driver visual attention by finite state automata. In Ref. [28] the authors make use of multicues extracted from a gray level image to detect the eye windows within an a priori detected face region. The precise iris and eye corner locations are then detected by a variance projection function and an eye variance filter. In Ref. [29], face regions are also initially determined by using rules derived from quadratic polynomial models, then eye components are extracted after the segmentation of skin pixels and lips. In Ref. [30] the distribution of discriminant features that characterize eye patterns are learnt statistically and provided to probabilistic classifiers to separate eyes and non-eyes. Also in this case the eye localization method requires a face detection step based on hierarchical classifiers [31].

The use of eye detection algorithms in the visible spectrum for automotive applications is not straightforward for several reasons: the problem of face segmentation (distinguishing faces from a cluttered background) should not be avoided, as has been done in many papers by imaging faces against a uniform background; the common use of skin color information to segment the face region is based on computationally expensive initializations and is not so general as to be applicable in different light conditions and with different skin colors; finally, the more precise the location of the eye regions in an initial step, the more reliable the results of the subsequent eye detection algorithms.

The second problem that a visual system for automotive applications has to solve is to model the eye occurrence on image sequences to detect driver status. Substantial amount of data have been collected to study driver behavior from a suite of vehicle sensors and unobtrusively placed video cameras [32]. Human error is known to be a causal factor in many accidents. Inattention and fatigue play an important role in human errors since they affect cognitive aspects and impair perception and the ability to make a decision to react, and also degrade the actual performance of actions. Detecting the driver's state and the driver's fatigue level in particular is actually a difficult task. Research on this subject is still incomplete. At this time it is not possible to provide an exact quantitative assessment of fatigue.