

HUMAN EYES LOCALIZATION FOR DRIVER INATTENTION MONITORING SYSTEM

Karel Horak, Miloslav Richter, Ilona Kalova

Brno University of Technology
Department of Control and Instrumentation
Kolejni 4, Brno
Czech Republic
{horakk, richter, kalova}@feec.vutbr.cz

Abstract: Paper presents a machine vision framework for human eyes detection in non-trivial scenes, especially for detection of a driver's pupils while a car is moving on a road. The article describes a technique of image acquisition, which takes advantage of an electronic camera with a specific reflective illuminator. Besides, an image processing method for pupils detection based on convolution is presented too and in the end of the paper there is demonstrated function of a tracking algorithm with estimation of the eye blinking. All the exploited designs and methods were proposed with respect to further realization of a real-time application determining a level of the driver's fatigue.

Key-Words: image processing, driver, fatigue, pupil, detection, tracking.

1 Introduction

Currently the most features in new cars are primarily intended to reduce aftereffects of an accident. On the opposite side there are methods intended to prevent from accidents e.g. by means of monitoring of the driver's level of an attention. Increasing number of accidents caused by low vigilance of drivers is a very serious problem, that's why visual systems for face detection [6] and monitoring of the driver's fatigue become part of some new cars. Inattention of the driver can be generally detected by means of a great number of various symptoms [9]. In this paper a simple and efficient method based on visual detection of human's pupils [3] is introduced.

In the flow chart below (Fig. 1) is shown the whole process of monitoring of driver's fatigue. First step is acquisition of a video sequence of the driver by means of an electronic camera (chapter 2). Follows image preprocessing and detection of human's pupils in acquired images separately (chapter 3). Waveforms of these detected coordinates of pupils are then filtered due to eventual errors in the previous stage of pupil detection. After it a frequency and duration of the driver's blinking is estimated and a fatigue level of the driver is determined in the end (chapter 4).

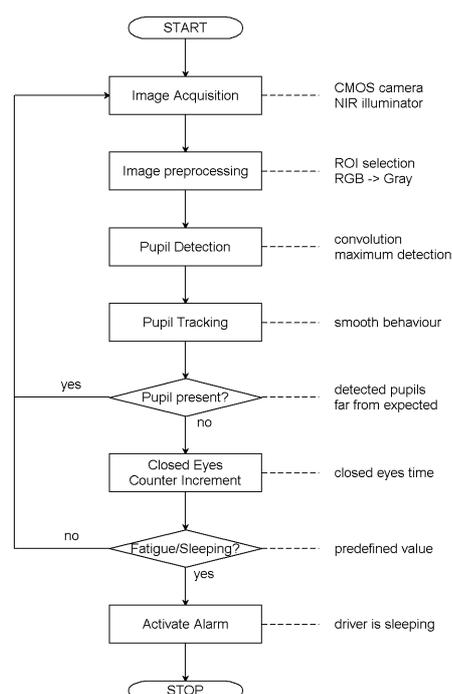


Fig. 1. The flow chart of the driver's fatigue monitoring system

2 Image Acquisition System

2.1 Parameters of the camera

In visual systems, a driver's fatigue is usually estimated from human features like blinking, yawning, moving of a head etc. [9]. On this account it is necessary to acquire images especially from the front of the driver. Images have to contain at least regions of all possible face positions. Accordingly a selected camera is emplaced on a windscreen inside a car in distance approximately 600 mm (considering 6 mm focal length and 1/2.5" CMOS sensor) from driver's face. Images with a standard resolution 640x480 pixels are acquired since the resolution of the image as well as a framerate of the camera are not critical aspects for our solution. Framerate of the camera don't exceed value of 15 fps, because a driver generally falls asleep very slowly.

2.2 Infra-red illuminator

In images acquired by the standalone camera without any additional illuminator occur reflections of outside sources of light as headlamps of others vehicles, street illuminations, sunlight etc. [8]. These ambient light sources together with relatively high complexity of the car interior and the driver's head and eyes moving make the pupil detection and their tracking very difficult. Therefore a special illuminator with a set of infra-red diodes was suggested and then constructed. The designed illumination device is very simple and it is formed by total of eight diodes distributed evenly along the circumference of the camera lens as you can see on the next illustration (Fig. 2).

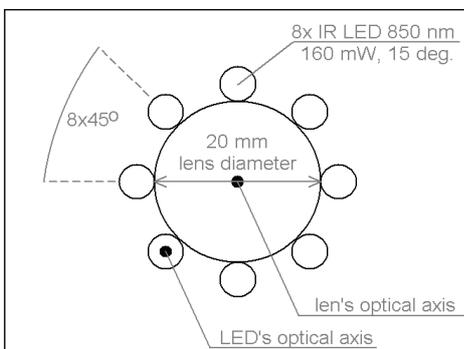


Fig. 2. LEDs distribution around the camera lens

All the eight diodes emit IR light in a continuous mode on wavelength 850 nm with 15° aperture and 50 mA of direct current (160 mW). Such powerful illumination is sufficient to suppress of a usual light sources. However, in case of very strong external sources of light (e.g. direct sunlight at noon) a flash mode instead of the continual lighting can be used. In this case the electric current through circuit is switched to approximately 500 mA, but only for a short time. Naturally, an exposure time of the CMOS sensor has to be decreased properly.

2.3 Acquired images

Due to the so-called bright pupil effect [1], diodes have to be placed necessarily close to the lens optical axis as possible. Then the human pupils reflect practically all the NIR rays back to the camera and produce bright pupil effect mentioned above. Exactly, the bright pupil effect is caused by a retina, which reflects almost all infra-red wavelengths back outwards from the eye. On the contrary, when eyes are illuminated off the lens optical axis, they seem dark since NIR rays do not reflect back into the camera lens. Examples of both images acquired without and with the NIR illuminator you can see on the next image (Fig 3).

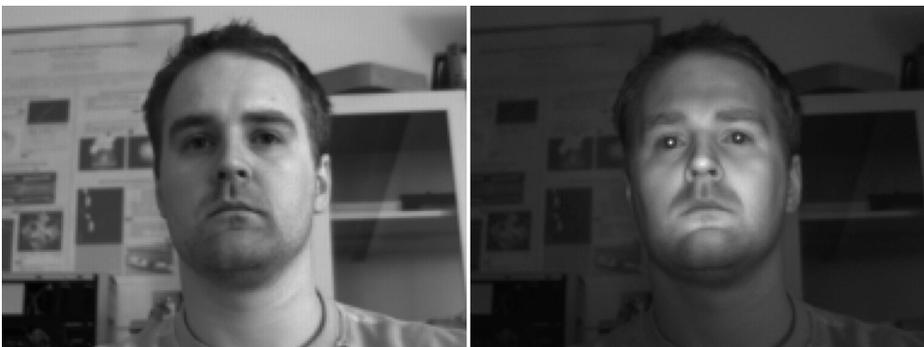


Fig. 3. Comparison of images acquired without (on the left) and with (on the right) the NIR illuminator

In addition to the illuminator it is possible to use the filter to suppress all wavelengths shorter than the dominant wavelength of infra-red LEDs (approximately 850 nm). Through an experiment we took advantage of a glass filter, which is able to filter out all rays with wavelengths below 780 nm (it means all visible wavelengths). Such image acquisition procedure leads to a simplified video sequence without surrounding items (Fig. 4). The human pupils are then detected more easily and with higher reliability than on images like on the previous figure.



Fig. 4. Influence of using the filter – original image (left) and image acquired with the filter (right)

Besides advantage described above the glass filter causes that only an illuminated face is visible on the video sequence. This drawback unfortunately makes impossible the accurate detection of a whole face region. Considering the complete face region may be needed in the future, a more general technique of the image acquisition without mentioned filter was chosen. It follows that we consider more complex images like on the right image of the figure 3.

3 Detection of Pupils

It is obvious from the figure 3 that pupils create two bright spots inside the small dark area caused by the iris and the sclera. If images acquired with and without the infra-red illuminator are available at the same time (Fig. 3), their subtraction can be used for simplifying of the pupil detection [2]. Unfortunately the subtracted images not only contain bright spots that belong to eyes, but often also contain other adjacent regions because of the different illumination level of the face (Fig. 5). Furthermore this method requires a couple of images for only one step in an image processing sequence.

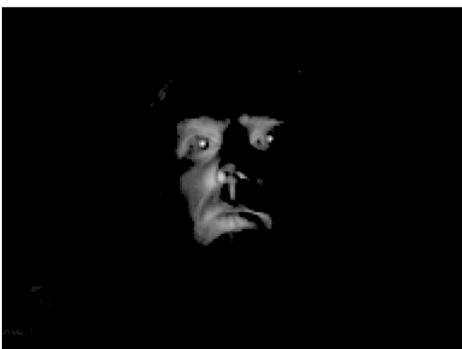


Fig. 5. The differential image mostly contain just two peaks

Considering above mentioned disadvantages of differential images we decided to only use the simple images obtained via the camera equipped with continuously running illuminator (on the left in Fig. 4).

Selection of the region of the interest (Fig. 6 on the left) is a first step of an image preprocessing leading to the detection of both pupils. In our case images are reduced to 60 percent of original dimensions. It was found out through an experiment that a range of the head movements are surprisingly small (Fig. 6 on the right). Moreover the range of head movements of a belted body in a car will be more limited than the range of our practically unbounded movements. Second and also last image preprocessing procedure is a color conversion from RGB model to ordinary grayscale image. Conversion is performed by means of the standard color transformation [7].

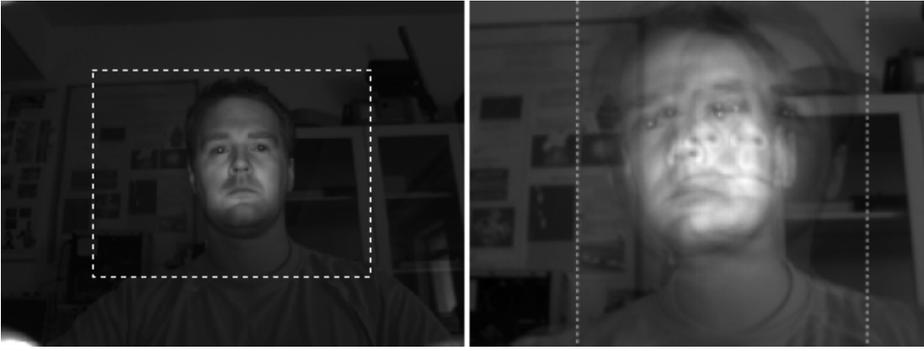


Fig. 6. The ROI selection (on the left) and two boundary lines of expected driver's head positions (on the right)

When the cropping and color conversion are done, the grayscale image is available and both pupils have to be detected in each image of the acquired video sequence. As denoted before, human eyes form two bright spots on the dark neighborhood caused by the iris and the sclera (see detail of the eye in Fig. 7). With respect to a rotational independence of human eyes, pupils can be localized by means of a standard convolution filter [5] with an appropriate kernel given by the equation (1). Vertical and horizontal dimensions of the kernel are denoted by letters R and C respectively.

$$g(x, y) = f(x, y) * k(u, v) = \sum_{v=1}^R \sum_{u=1}^C f(x-u, y-v) \cdot k(u, v) \quad \forall x, y \in \langle 1; \dim(f) \rangle \quad (1)$$

In our case kernel denoted $k(u, v)$ should imitate a brightness distribution near the center of pupils in the acquired image. Shape of the kernel and corresponding values are both shown on the right of the next figure (Fig. 7). The kernel is designed symmetrically and contains only few coefficients due to a low computational cost. Calculation of a convolution value is then executed by means of only five multiplications and one summation.



Fig. 7. The detail of the eye (on the left) and corresponding kernel values (on the right)

In correctly acquired images the convolution results in a new image with at least two peaks. Two biggest peaks generally belong to the searched pupils and so a simple detection of the two maxima results in correct coordinates of pupils (Fig. 8). In case that the driver turns his head too much, pupils not have to be detected properly since some of the two peaks belong to pupils get lost.



Fig. 8. Driver's eyes are open (on the left) and closed due to blinking (on the right)

The described detection of pupils is applied on the each image of the sequence separately. Potential errors (e.g. closed eyes or detection of some wrong peak) are rare and handled by the pupil tracking algorithm, which is presented in the next chapter.

Experiments with kernels with bigger dimensions and kernels containing different values in various placements led to results comparable to the result just depicted and no significant difference was observed. Nevertheless, if dimensions

of the kernel exceed dimensions of the relevant eye region, an expected result of the convolution (Fig. 8) naturally faded out very rapidly and a useless result is obtained. It follows the described technique for detection of pupils is invariant to a scale of the driver's face only in very limited range. The changes of the scale might be easily caused by the driver leaning back or forward, though such little changes are evaluated correctly.

4 Pupils Tracking and Blinking Estimation

4.1 Tracking algorithm

From the previous chapter we know positions of the two biggest peaks (Fig. 9) in the filtered image. However with respect to the used detection method it is necessary to point out that a correct result is not always guaranteed. From executed experiments we modestly extrapolate that almost all images are processed correctly. It follows majority of computed couples of coordinates (x,y) of the pupils is correct and the rest of couples has to be filtered properly.



Fig. 9. Pupils detected correctly (on the left) and incorrectly during blink (on the right)

For the filtering of an incorrectly detected coordinates we assume that they are isolated values (different from their vicinity). It means detection errors appear more or less randomly. On the next figure (Fig. 10) you can see two graphs. The first of them is the graph containing original values of x-coordinates of both pupils. An upper curve corresponds to the left eye (right cross in the figure Fig. 9) and bottom curve to the right eye (left cross). In case of closed eyes or another situation when one or both pupils are not visible, the graph contains obvious errors.

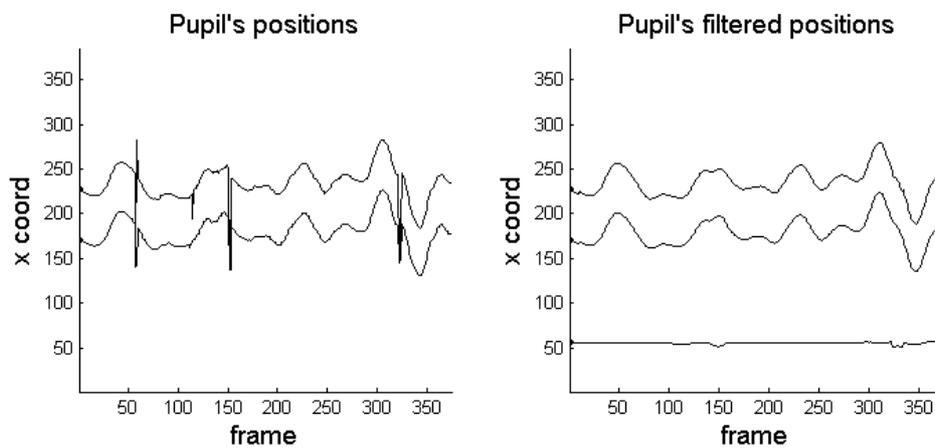


Fig. 10. Graphs of the original (on the left) and filtered (on the right) values of x-coordinates of pupils

The second graph contains curves of the data from the left graph after passing through a low frequency filter. Resulting smooth curves also correspond to a limited human's acceleration of the head movement. Third shape of the curved line depicted below both filtered curves is distance between x-coordinates of the left and right eye (strictly between x-coordinates of the both pupils). Because a distance between the human eyes/pupils is generally constant, the shape of this curve has to be straight line in the ideal case.

4.2 Estimation of blink moments

The each blink of eyes is a moment, when both left and right eye is closed. For determination moments of blinking it is necessary to find positions in the graph, where both upper and bottom curves are discontinuous at the same time. This determination is simply done by the subtraction and subsequent thresholding of the original and filtered graphs (Fig. 10). In next illustration (Fig. 11) are highlighted computed moments of blinking and one error through the detection of the pupils.

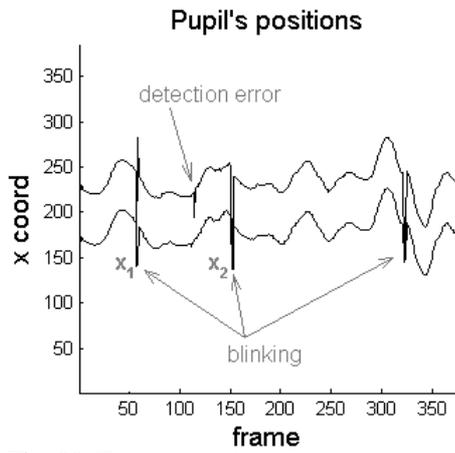


Fig. 11. The recognition between errors and moments with closed eyes

From detected moments can be easily computed both durations and frequency of the blinking. The duration of the each blink is simply proportional to duration of the corresponding error in the graph and frequency of the blinking is given by the following equation (2).

$$f_b = \frac{C}{N-1} \cdot \sum_{i=1}^{N-1} \frac{1}{x_{i+1} - x_i} \quad [s^{-1}] \quad (2)$$

Symbol f_b denotes the average number of blinks per second, symbol C represents the number of frames per second of the camera and finally N is the number of detected moments of the blinking. A mentioned frequency and durations of blinking are very good indices to estimation of the driver's fatigue [4]. Increases of the frequency or duration of the blinking indicate oncoming of the driver's fatigue and thus his lower attention. An acoustic and/or visual alarm has to be activated in such cases to prevent an accident.

5 Conclusion

In the presented paper was introduced a framework for the image acquisition and consequential detection of driver's fatigue. Image acquisition hardware was assembled and exploited in order to following localization of driver's pupils. On the basis of known positions of pupils in the image, both frequency and duration of the blinking were calculated because of indicating of the driver's fatigue. The next research will be aimed to improve reliability and accuracy of the detection of pupils and also to analyze regions of eyes in order to determine some parameters of eyelids e.g. an absolute position or closing speed.

Acknowledgement: The introduced paper was supported by the Czech Science Foundation under the project GA102/09/1987 and by the Ministry of Education of the Czech Republic under the project MSM0021630529. Both of them are very gratefully acknowledged.

References:

- [1] Batista, J.P., A Real-time Driver Visual Attention Monitoring System, *Pattern Recognition and Image Analysis*, Vol. 3522, 2005, pp. 200-208.
- [2] Bergasa, L.M. et al., Visual Monitoring of Driver Inattention, *Studies in Computational Intelligence*, Vol. 132, 2008, pp. 19-37.
- [3] Morimoto, C.H., et al., Pupil detection and tracking using multiple light sources, *Image and Vision Computing*, Vol. 18, 2000, pp. 331-335.
- [4] Rashmi, P., Preeti, B., Intelligent Monitoring System for Driver's Alertness, *Knowledge-Based Intelligent Information and Engineering Systems*, Vol. 4692, 2008, pp. 471-477.
- [5] Russ, J.C., *The Image Processing Handbook*, CRC Press Inc., 1994.
- [6] Saleh, A.A., A Simple and Novel Method for Skin Detection and Face Locating and Tracking, *Computer Human Interaction*, Vol. 3101, 2004, pp. 1-8.
- [7] Sonka, M., Hlavac, V., Boyle, R., *Image Processing, Analysis and Machine Vision*, Thomson, 2008.
- [8] Vernon, D., *Machine Vision*, Prentice Hall, 1991.
- [9] Vural, E. et al., Drowsy Driver Detection through Facial Movement Analysis, *Human-Computer Interaction*, Vol. 4796, 2007, pp. 6-18.