Optical and Acoustical Methods in Science and Technology 2013

Research and Parameter Optimization of the Pattern Recognition Algorithm for the Eye Tracking Infrared Sensor

K. Murawski[a,*], K. Różanowski[b] and M. Krej[b]

[a]Military University of Technology, Institute of Teleinformatics and Automatics
S. Kaliskiego 2, 00-908 Warsaw, Poland

[b]Military Institute of Aviation Medicine, Aviation Bioengineering Department
Z. Krasińskiego 54, 01-755 Warsaw, Poland

The paper presents the process and results of tests of the pattern recognition algorithm. The algorithm has been developed for the sensor to track the activity of the eyes. The study was conducted on a group of ten people. The group was selected, so that it was possible to determine the influence of gender, age, eye color, contact lenses and glasses on the result of the algorithm. The object of the study was also to optimize the pattern recognition algorithm in terms of CPU utilization. For this purpose the duration of the algorithm steps were measured as a function of the allocation of tasks to multiple processor cores.

DOI: 10.12693/APhysPolA.124.513

PACS: 07.05.Pj, 02.60.Pn

1. Introduction

This paper presents the research and process optimization of the pattern recognition algorithm for analyzing images taken in infrared. In the presented case, in view of the mode (real time) sensor operation the image processing time is essential. Processing time determines the frequency of loading of frames, and thus affects the rate of detection of the observed phenomena. This area of applying sensors is not the only one, in which the detection speed is important. In the literature many examples are presented [1–11], in which detection times and making a decision are key due to information importance. The reduction is the result of the signal processing optimization algorithm (I) or the use of hardware implementation (II). In the first case signal characteristics are looked for that limit the number of calculations [12–20] or algorithm modifications are made to make better use of the processing power. Variant (II) concerns increasing the detection speed by using specialized hardware solutions [21, 22].

The conducted optimization process was carried out in accordance with (I). The pattern recognition algorithm has been tested to determine the average time to process a frame. The influence of user features were tested on the average frame processing time, such as age, glasses, contact lenses and eye color. The calculated indexes were compared with the values obtained by the algorithm model.

The infrared sensor for eye tracking was built on the basis of a monochrome video camera, infrared emitter, emitter controller, lens and filter. The sensor in detail is presented in [23] and patern recognition algorithm in [2].

*corresponding author; e-mail: kmurawski7@gmail.com

2. Pattern recognition algorithm optimization

In order to optimize the algorithm the analysis time for a single video frame was reduced (image). The optimization was preceded by examining the algorithm properties. This has been done through experimentation. The experiment was conducted on a truck simulator. Its aim was to identify the algorithm operations that are characterized by the longest execution time. A total of ten people aged from 26 to 45 yrs (26, 27, 28, 29, 30, 32, 3×40, 45) participated. The study group consisted of five women and five men. Six people in the group had blue eyes, three people had green eyes and one person gray eyes. In the group of people with blue eyes four had no vision deficiencies. One person wore glasses and the other contact lenses.

All participants were subjected to the same test procedure. During the experiment, the same driving conditions were guaranteed. These included the level of illumination, the weight of the vehicle, the adhesion to the road, the maximum vehicle speed, weather conditions, obstacles, including vehicles and people, the type of terrain — suburban and urbanized. The study consisted of driving the route. The participant had to drive three sections. The first and third were marked outside the urban area.

The middle section ran through urban areas. During driving raw image from the sensor was stored in digital form. At the beginning of recording the program indicated the first point of the planned route. This point was out of town and was known only to the person controlling the test. The end of recording followed after storing 9000 frames (three two-minute segments, each with 3000 frames). The middle section was selected to test the algorithm. It required the driver to focus more on the road ahead. So the person driving was less able to control their impulses.

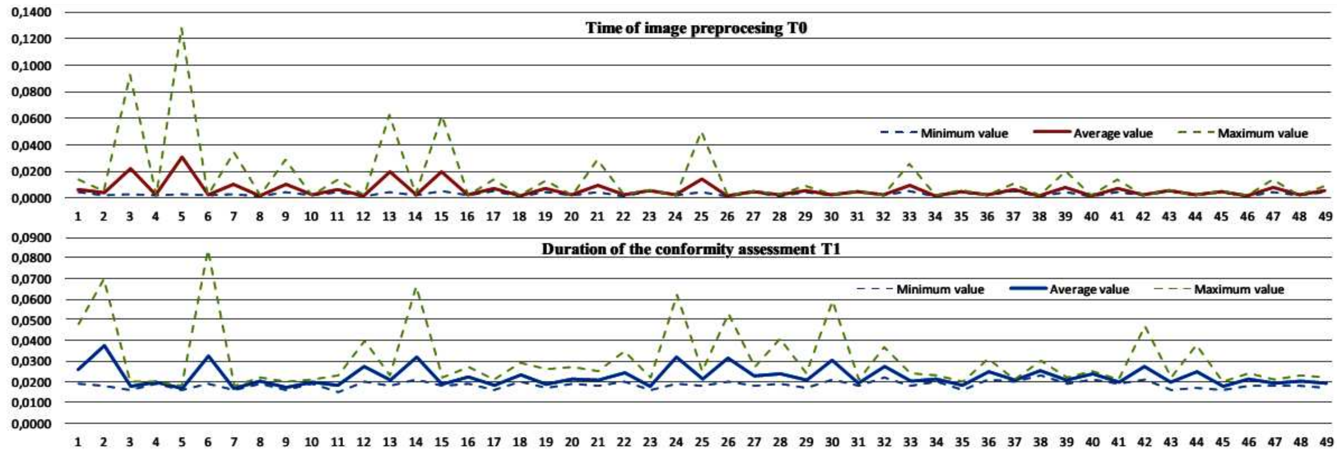

Fig. 1. Time series imaging the variable time of processing frames in a segment.

Each of segments was processed five times. During the processing the following were determined: time of image preprocessing T0, duration of the conformity assessment T1. The multiple processing segment was used to reduce the impact of changes in CPU load (caused by processes starting by the operating system) on the obtained values. The average values were calculated for all analyzed segments based on the obtained times. They were the basis for the determining the average time to process a frame in a segment of video. Sample time series showing the change in frame processing time in a segment is illustrated in Fig. 1. Changes in times T0 and T1 was studied as a function of age, eye colour, glasses, and contact lenses.

4.1. Studying the impact of a patient's age on the duration of the algorithm

The results show that the age of a tested person has practically no impact on the duration of output data processing. The impact of the age of the tested person is even less visible after averaging the results, Fig. 2.

On the basis of Fig. 2 it was found that the influence of age of the subject on the frame processing duration is negligible. The resulting frame processing time changed in the range from 0.026 s to 0.032 s and processing was done at a rate of 31 fps up to a maximum of 38 fps. The average time was 0.029 s. and the standard deviation equaled 0.002. It means that all the results obtained are clustered around the average and are in the range of 0.029 s $\pm$ 0.003 s. This results that the algorithm without modifications can be used in a wide range of user ages.

4.2. Studying the effect of eye colour on the duration of the algorithm

The impact of eye color on the duration of the algorithm was determined by analyzing the data obtained from persons. 60% of the people in the study declared to have blue eye colour. 30% declared to have green eyes and 10% indicated gray.

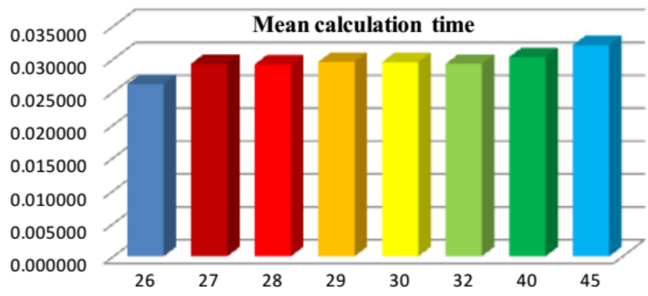

Fig. 2. The impact of a patient's age on the duration of processing a single frame.

The data show that the algorithm determines the position of the eyes fast regardless of eye colour. By analyzing the average values it is noticeable that the decision of determining is fastest for gray eyes and then green, and then blue. This difference, however, is symbolic. The average processing time was 0.030 s $\pm$ 0.0003 s. It is almost identical to time 0.029 s $\pm$ 0.003 s obtained in Sect. 4.1. According to the authors the obtained results are under the effect of not only eye colour, but also other factors such as facial features. This belief is the result of the fact that the algorithm processes a frame recorded in infrared light by a monochrome camera. Thus the impact of color should be almost unnoticeable. Therefore, the only differences in the time of calculation must result from the number of determined potential positions of the pupil. This, however directly depends on the number of detected contours and vertices on the face image plane.

4.3. The impact of glasses and contact lenses on the duration of the algorithm

Due to the composition of the tested group, calculations were performed on the subjects, who declared blue eye colour. For the rest of the eye colours there were not enough test subjects. In the composition of the study group only one person wore glasses (photochromic glass with UV protection and antiglare), one person was wearing contact lenses. The rest did not require vision correction.

By analyzing the data it was ascertained that the shortest times of processing a frame, an average of 0.029 s, was achieved by persons not needing vision correction. The person with contact lenses did slightly worse. The average processing time was 0.030 s. The longest frame processing time was achieved by the person wearing glasses. The average processing time then was about 0.033 s and was 0.004 s longer than the best result.

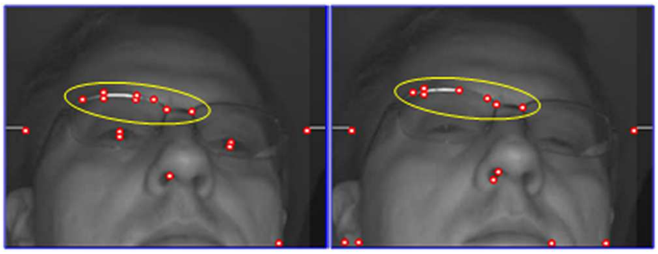

Fig. 3. Interference in the image caused by glasses.

The increases in frame processing time caused by the glasses are due to the introduction of an image element strongly different from the natural facial features. This results in the occurrence of additional contours and vertices, Fig. 3, which randomly increases the number of points of potential occurrences of the pupil and extends the evaluation process.

4.4. Algorithm elements designed to optimize — proposed changes

It can be noticed that the average value T1 (0.0254 s) is almost six times higher than the average value T0 (0.0043 s). For this reason, the main focus is on improving the process of determining the similarity of patterns of image slices. A small change was also introduced in the process of determining expected points of the pupil position. This change was based on a single and not, as was the case in the original algorithm, double determination of the image edge (edge detection). The image edge in the improved algorithm is transmitted simultaneously to two computation paths. They simultaneously determine the characteristic points resulting from edge analysis and the determination of coordinates of vertices on the image. In the proposed form the algorithm for each of the points appoints a separate computational process. This process is responsible for the calculation of the conformance of the tested fragment of the image with each of the patterns. As a result, the number of appointed processes for each image frame is a dynamic variable. In the adopted implementation the operation system is responsible for process management. The rest of the algorithm remains unchanged. The determined characteristic points are sorted by the value of compliance index in descending order. Afterwards a point, distant from the first in a series of at least 30 pixels, is searched for fulfilling the condition of minimal match, e.g. more than 70%.

The modified algorithm (Fig. 4) was compared with the original version. For this purpose, it was subjected to the same tests.

In the modified algorithm, the patient's age effect was also small. Frame processing time ranged from 0.013 to 0.015 s. The average processing time was 0.014 s and the standard deviation equaled 0.0006. The average processing time was less than twice the time to get without the optimization algorithm.

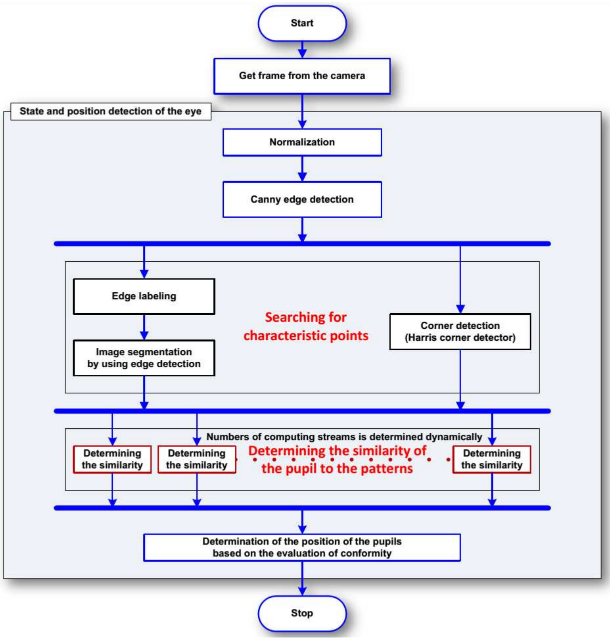

Fig. 4. Modified pattern recognition algorithm for the eye tracking sensor.

In the study determining the impact of eye color on the duration of frame processing, the shortest time was achieved by the person with grey eyes. Next in order were people with green eyes. The longest time was achieved by people with blue eyes. The determined times fit in the range of 0.0143 to 0.0146 s. The average processing time was 0.014 s and was more than half as shorter than the same time achieved for the algorithm before optimization. Equally positive results were obtained by studying the effects of glasses and contact lenses. After modifying the algorithm the impact of studied factors remained unchanged. The shortest times were obtained by those that do not require vision correction. The next person was the person wearing contact lenses. The determined length of time was in the range from 0.014 to 0.015 s. with a mean value equal to 0.0148 s and a standard deviation of 0.0006. As before, the resulting average is twice less than the same time determined before optimization. The quoted results are also confirmed by comparing of mean values T0 and T1 obtained before and after optimization. The average time T0 equalled 0.0043 s and T1 was equal to 0.0254 s. After optimization these times were accordingly 0.0023 s and 0.0123 s giving an average frame processing time equal to 0.0146 s.

Thanks to optimization the frame processing time was reduced by 51%. As a result, the increase in average frequency of acquiring and analysis of frames taken from the sensor was provided with a maximum of 33 fps to 68 fps.

5. Conclusions

The conducted experiments show that eye colour, age, contact lenses and glasses do not affect substantially on the duration of frame processing. In the first implementation of the algorithm these times were respectively: 30.0 ms ± 0.3 ms; 29.4 ms ± 3.2 ms; 30.8 ms ± 2.5 ms. The starting point for the algorithm optimization was to analyze the execution times of its steps. Finally, the two stages of the algorithm were selected for optimization: searching for characteristic points and determining the similarity of the image slices to the patterns. As a result of algorithm parameter optimization the total time of video frame processing was reduced to 14.57 ms, it is about 51% less than before optimization. As a result of optimization the average frequency of obtaining frames taken from the sensor increased from a maximum of 38 fps to 68 fps, which in turn enables the observation of fast-changing phenomena [24].

Acknowledgments

The project was co-financed by the European Union through the European Regional Development Fund under the Innovative Economy Program. Development Project contract no. POIG.01.03.01-10-085/09.

References

[1] Z. Bielecki, J. Janucki, A. Kawalec, J. Mikolajczyk, N. Palka, M. Pasternak, T. Pustelny, T. Stacewicz, J. Wojtas, *Metrol. Meas. Syst.* **19**, 3 (2012).

[2] E. Lacot, W. Glastre, O. Jacquin, O. Hugon, H. de Chatellus, *J. Opt. Soc. Am. A* **30**, 60 (2013).

[3] S. Shahbeig, H. Pourghassem, *J. Opt. Soc. Am. A* **30**, 13 (2013).

[4] Z. Hu, K. Uchimura, *Int. J. ITS Res.* **4**, no. 1 (2006).

[5] Z. Piotrowski, K. Różanowski, *Acta Phys. Pol. A* **118**, 131 (2010).

[6] J. Keinert, H. Dutta, F. Hannig, C. Haubelt, J. Teich, in: *Proc. Design, Automation & Test in Europe*, 2009, p. 135.

[7] W. Żorski, in: *Proc. 8th Int. Conf. Lecture Notes in Computer Science*, Vol. 4179, Springer-Verlag, 2006, p. 1071.

[8] M. Piszczek, M. Kowalski, M. Karol, K. Rutyna, M. Zarzycki, M. Szustakowski, K. Ludwikowski, *Acta Phys. Pol. A* **124**, 550 (2013).

[9] M. Życzkowski, *Acta Phys. Pol. A* **124**, 428 (2013).

[10] K. Murawski, *Electrotechn. Rev.* **9**, 184 (2010).

[11] K. Murawski, in: *Proc. XV Int. Conf. on Methods and Models in Automation and Robotics (MMAR)*, Międzyzdroje 2010, p. 356.

[12] S. Belongie, J. Malik, J. Puzicha, *IEEE Trans. PAMI* **24**, 509 (2002).

[13] J. Yang, A. Waibel, in: *Proc. WACV, Sarasota (FL)*, 1996, p. 142.

[14] D. Comaniciu, V. Ramesh, in: *Proc. Third IEEE Int. Workshop on Visual Surveillance*, 2000, p. 11.

[15] G. Bradski, in: *Proc. IEEE Workshop Applications of Computer Vision*, 1998, p. 214.

[16] J. Clemons, H. Zhu, S. Savarese, T. Austin, in: *Proc. 2011 IEEE Int. Symp. on Workload Characterization*, 2011, p. 91.

[17] E. Rublee, V. Rabaud, K. Konolige, G. Bradski, in: *Proc. 2011 Int. Conf. on Computer Vision*, 2011, p. 2564.

[18] C. Harris, M. Stephens, in: *Proc. Fourth Alvey Vision Conf.*, 1988, p. 147.

[19] I. Gordon, D.G. Lowe, in: *Proc. ISMAR*, 2004.

[20] E. Rosten, T. Drummond, in: *9th European Conf. on Computer Vision*, Vol. 1, 2006, p. 430443.

[21] W. Żorski, in: *Proc. 7th Int. Conf. Lecture Notes in Computer Science*, Springer-Verlag, Vol. 3708, 2005, p. 460.

[22] B. Draper, W. Najjar, W. Bohm, J. Hammes, B. Rinker, C. Ross, M. Chawathe, J. Bins, in: *Proc. Fifth IEEE Int. Workshop on Computer Architectures for Machine Perception (CAMP'00)*, 2000, p. 222.

[23] K. Różanowski. K. Murawski, *Acta Phys. Pol. A* **122**, 874 (2012).

[24] K. Różanowski, K. Murawski, *Acta Phys. Pol. A* **124**, 558 (2013).