ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/278965906

Using Non-Calibrated Eye Movement Data To Enhance Human Computer Interfaces

CONFERENCE PAPER · JULY 2015

DOI: 10.1007/978-3-319-19857-6_31

READS

25

2 AUTHORS:

Pawel Kasprowski

Silesian University of Technology

25 PUBLICATIONS **84** CITATIONS

SEE PROFILE

Katarzyna Harezlak

Silesian University of Technology

16 PUBLICATIONS **9** CITATIONS

SEE PROFILE

Available from: Pawel Kasprowski

Retrieved on: 16 February 2016

This is a pre-print. The final version of the paper was published in Springer, Smart Innovation, Systems and Technologies, Vol. 39, 2015 as part of Proceedings of the 7th KES International Conference on Intelligent Decision Technologies (KES-IDT 2015) and is available in Springer Link library via http://dx.doi.org/10.1007/978-3-319-19857-6_31

Using Non-Calibrated Eye Movement Data To Enhance Human Computer Interfaces

Pawel Kasprowski, Katarzyna Harezlak

Institute of Informatics
Silesian University of Technology
Gliwice, Poland
pawel.kasprowski@polsl.pl, katarzyna.harezlak@polsl.pl

Abstract. Eye movement may be regarded as a new promising modality for human computer interfaces. With the growing popularity of cheap and easy to use eye trackers, gaze data may become a popular way to enter information and to control computer interfaces. However, properly working gaze contingent interface requires intelligent methods for processing data obtained from an eye tracker. They should reflect users' intentions regardless of a quality of the signal obtained from an eye tracker. The paper presents the results of an experiment during which algorithms processing eye movement data while 4-digits PIN was entered with eyes were checked for both calibrated and non-calibrated users.

1 Introduction

The usefulness of the eye movement analysis was confirmed in research in many areas of interests. It may be used for example in advertisements developing, sociology, medicine and cognitive studies [3][6]. Recently, a lot of attention has been focused on possibilities to use eye movement for enhancing human-computer interfaces. Using gaze information as a new input device in a way similar to mouse seems to be the promising technique, making cooperation with computers even easier for unexperienced users. Nevertheless, the eye movement processing still faces a lot of usability problems so a lot of effort must be done to make this technique really user friendly. One of the examples is a so called Midas touch problem [7], which addresses the difficulty to decide when user looks at a button if he wants to click it or just to read its caption.

One of the most important obstacles in making eye movement based interfaces robust and convenient is the necessity to calibrate an eye tracker for each user before any usage [11]. The aim of the studies presented in this paper was to check whether it is possible to utilize information obtained from an eye tracker without prior calibration done by the user that is being measured. One of the simplest tasks - the PIN entering - was taken into consideration and these studies are based on the research discussed in [9].

The main contribution of the research presented in the paper is the introduction of the idea to shorten eye tracking sessions by carrying out the same calibration for various users. Thus, the novel so called *regression based* algorithm was implemented and

adfa, p. 1, 2011.
© Springer-Verlag Berlin Heidelberg 2011

This is a pre-print. The final version of the paper was published in Springer, Smart Innovation, Systems and Technologies, Vol. 39, 2015 as part of Proceedings of the 7th KES International Conference on Intelligent Decision Technologies (KES-IDT 2015) and is available in Springer Link library via http://dx.doi.org/10.1007/978-3-319-19857-6_31

compared to an intuitive *distance based* algorithm. The correctness of the analysed task realization for both cases and factors influencing the results were analysed.

All classic experiments based on eye tracking methods are conducted in accordance with a commonly used schema. Each trial starts with a calibration process. The aim of that step is to find the correlation between coordinates of user's gaze point and coordinates represented in an eye tracker system. During a calibration users are asked to move their eyes over a screen as a reaction to a presented stimulus. Dependent on an experiment type there may be various stimuli used, yet the most popular is a point jumping over a screen. Each change of the point position triggers eyes movement. Recorded eye movement samples correspond to a given point on a screen and a user is expected to keep the focus in the same point for a while, long enough to collect such set of samples, which ensure good adjustment of a screen and an eye tracking system coordinates. The time of a single point presentation is usually set within scope of 2 – 3 seconds [4]. A number of point's locations and their dispersion on the screen are other issues [8][1]. It is obvious that the higher quality may be achieved for more points, yet taking user's convenience into consideration the lowest possible number is better solution. The reason is time required to perform an eye tracker calibration – too long can be wearisome for participants, discouraging them for the involvement in core experiments.

The main problem of the calibration is that it must be repeated before every trial as it depends on an environment used in experiments and on characteristic features of a particular user. As the calibration process is rather cumbersome for a user, the idea of the paper was to check if, for some human-system interactions, which do not require highly accurate eye tracker's adjustment, it is possible to omit calibration step and still achieve satisfactory results. Tasks regarding entering a PIN, in which focusing eyes on a specific area is sufficient to determine a digit, can be taken as an example. Such task may be used to lock and unlock computer screen with eyes or to enter PIN at ATM [10]. According to [2] it prevents shoulder surfing attack and is generally more difficult to forge.

2 The experiment

The eye movements were registered using the Eye Tribe - an eye tracking system working with sampling rates 60 Hz. The accuracy and spatial resolution declared by manufacturer equals 0.5° – 1° and 0.1° respectively. The eye tracker was placed below the screen. A PC computer was used to control the experiment (show stimulus, control the eye tracker and save recordings). The users were sitting centrally at a distance of 60 cm.

The experiment consisted of two parts. During the first part one participant was calibrated using a classic scenario with 9 points evenly distributed on the screen. It lasted for about 20 seconds. Then a screen with evenly distributed circles with digits from 0 to 9 was displayed and the participant was asked to look at a digit and press a trigger button, then look at subsequent digit and press the trigger and so on. The PIN

This is a pre-print. The final version of the paper was published in Springer, Smart Innovation, Systems and Technologies, Vol. 39, 2015 as part of Proceedings of the 7th KES International Conference on Intelligent Decision Technologies (KES-IDT 2015) and is available in Springer Link library via http://dx.doi.org/10.1007/978-3-319-19857-6_31

was a four digits sequence for which every two subsequent digits were always different. After four trigger clicks the attempt (a trial) was saved.

The second (and more interesting) part of the experiment started when some number of different participants were asked to enter their PINs with eyes in the same manner, but this time without any calibration. These participants' eye movements were registered using a calibration function built for the first user. Eye movement data for every PIN entering was saved (later referred to as a *trial*).

There were 802 trials collected including 204 with own calibration (called 'calibrated') and 598 without own calibration (called 'non-calibrated'). 41 participants took part in the experiment.

To examine samples gathered during the experiments two own developed methods were used. Both of them are based on sets of fixations extracted from eye movement signal and are described in details in the subsequent section.

3 Method

The purpose of the PIN extraction algorithm described in this section was to obtain information about a sequence of digits pointed with eyes from the recorded eye movement data. In the algorithm two phases may be distinguished :

- Extraction of fixations,
- Assigning fixations to digits on the screen.

3.1 Extraction of fixations

Typical eye movement signal consists of two events: fixations - when eye is relatively still and the brain acquires information from the scene; and saccades - a rapid movement when an eye position changes to another fixation. The extraction of fixations from a raw eye movement signal may be done using different algorithms [12][13]. It was our own implementation of one of the most popular dispersion-threshold algorithm (I-DT) used in this work.

At first the algorithm classifies each eye movement sample according to a simple rule: if the distance among this sample and five previous samples is less than a specified threshold (Th) the sample is classified as a potential part of a fixation (F) and it is classified as a potential part of a saccade (S) otherwise. In the next step all neighboring F-points are gathered together as potential fixations. Every fixation has four attributes: its start time, x and y coordinates of its center and the fixation duration. The subsequent steps convert this preliminary list of fixations into the final list using different techniques for fixation merging and removing. All details of the algorithm are presented in [5].

The value of threshold parameter (Th) started from 0.2 deg. and was increased by 0.2 deg. until one of following conditions was met: I-DT algorithm returned exactly four fixations or threshold value reached 8 deg.

In the latter case the trial was rejected and no further analyses were done.

This is a pre-print. The final version of the paper was published in Springer, Smart Innovation, Systems and Technologies, Vol. 39, 2015 as part of Proceedings of the 7th KES International Conference on Intelligent Decision Technologies (KES-IDT 2015) and is available in Springer Link library via http://dx.doi.org/10.1007/978-3-319-19857-6_31

3.2 Discovering chosen digits

After extracting the four most dominant fixations, it was assumed that these fixations occurred while a person was looking at specific digits. The next task was to discover which digits were pointed with eyes. There were two different methods used in the research to pair fixations with proper digits.

Distance based approach.

The first – and the most obvious - method divides a screen into regions of interest (ROIs) using digits locations as points in Voronoi diagram. As a result every fixation is classified as a digit which is the closest one for this fixation. It may happen that a fixation location is almost in the middle between two digits (near a boundary of a Voronoi cell). As a result one of the digits must always be chosen but we may expect that the choice is somehow random in such case. Therefore, an additional step and an additional parameter: proximity coefficient (PCF) were introduced. After finding the closest digit for a fixation, it is checked whether a distance between that digit and the fixation multiplied by PCF is still lower than distances between the fixation and all other digits:

$$\forall_{0 \le i \le 9} \|F - D_i\| > min_{0 \le i \le 9}(\|F - D_i\|) * PCF \quad (1)$$

where F is fixation's location and D_i is a location of digit i.

If this condition does not hold for any fixation in a sequence, the whole trial is rejected. Obviously, for PCF equal to one there are no rejections (there is always one minimal distance) and as its value increases the number of rejections increases as well. For instance PCF equal to 2 means that the distance between the closest digit and the fixation must be twice lower than the distances between that fixation and all other digits.

To check whether a simple, one point calibration is enough to improve results, the additional assumption was added that the first digit of PIN is known. Therefore, the extended version of the algorithm introduces the additional step: before any classification all fixations are shifted in space so that the first fixation is positioned exactly in a location of the first digit. The confirmation of usefulness of such activity may be subsequently used for one point calibration displayed for example in the middle of the screen.

Regression based approach.

The next method was using slightly different approach. The basic assumption was that the regression model for a correctly adjusted PIN should provide the lowest error. Thus, the algorithm starts with building regression models that map 4 fixations into 4 digits for every possible combination of PIN digits. There are 7290 possible combinations when assumption that subsequent digits are always different holds. For every such model new fixations' positions are calculated and Mean Square Error between these new and correct positions of PIN's digits are calculated. At the end there is a list of possible PIN numbers with MSE for every PIN available. The PIN with the lowest error is chosen as a correct one.

This is a pre-print. The final version of the paper was published in Springer, Smart Innovation, Systems and Technologies, Vol. 39, 2015 as part of Proceedings of the 7th KES International Conference on Intelligent Decision Technologies (KES-IDT 2015) and is available in Springer Link library via http://dx.doi.org/10.1007/978-3-319-19857-6_31

The only problem in the algorithm described above is which regression function should be chosen to obtain reliable results. Usually, second degree polynomial function is used for eye trackers calibration [1] but such model is too precise. For 4 points it is able to build a function that maps given fixations to any sequence of digits with almost no errors. Therefore, it was a first degree polynomial function used to evaluate new values for X and Y independently:

$$X_{new} = A_x * X + B_x \tag{2}$$

$$Y_{new} = A_y * Y + B_y \tag{3}$$

The Levenberg Marquardt algorithm was used to calculate coefficients for each 4 fixations – 4 digits pair. Because we were not interested in mirror mapping of PIN numbers, an additional assumption that coefficients A_x and A_y must be positive numbers was made.

Similarly to the proximity coefficient in the distance based method, it was necessary to add possibility to allow for rejection of trials for which values found are not reliable. Therefore, an additional *min_error* (MER) parameter was introduced. If the lowest value of MSE for the trial is higher than MER, the trial is rejected as unreliable.

To make a fair comparison to the distance based approach that uses information about the first digit, there was also a version of the algorithm evaluated that calculates models only for PINs starting with a known digit.

4 Results

The most obvious results that may be taken into consideration is the absolute accuracy (ABS), which is measured as a ratio between the number of trials with PIN found correctly and the number of all trials. However, there are two more detailed factors that may be used when analyzing results of the algorithms described in the previous section. At first, each algorithm rejects some number of trials for which it assumes that recognition is impossible. So, the first factor to be analyzed is an acceptance rate (ACR). This factor is influenced by PCF and MER parameters accordingly to the algorithm used. Then, for all remaining trials, PIN is evaluated. The number of PINs found correctly to the number of all evaluated trials is defined as a correctness rate (CRR).

Ideally, both ACR and CRR should be 100%. However, it can be expected that both factors are dependent on each other – when the acceptance rate decreases, the remaining samples are of better quality and the correctness increases. And when the acceptance rate increases, more low quality samples are taken into account during the next step, which may result in lower correctness rate. Tuning of this two factors depends on the purpose of the trial (see **Conclusion** for examples). Obviously, the absolute accuracy (ABS) is the result of multiplication of the two above factors (ABS = ACR*CRR).

This is a pre-print. The final version of the paper was published in Springer, Smart Innovation, Systems and Technologies, Vol. 39, 2015 as part of Proceedings of the 7th KES International Conference on Intelligent Decision Technologies (KES-IDT 2015) and is available in Springer Link library via http://dx.doi.org/10.1007/978-3-319-19857-6_31

As it was presented in the previous section, the first rejection takes place after the 'extraction of fixations' step. All trials, for which it was impossible to find exactly four fixations are rejected. The next step when trials may be rejected depends on the algorithm used. For the distance based algorithm the acceptance rate depends on the proximity coefficient (PCF). As it was described in the previous section, increasing PCF decreases the number of accepted samples. For the regression based algorithm the *min_error* coefficient (MER) may be tuned to reject dubious trials. If MER is high, all samples are accepted and as it decreases, the acceptance rate (ACR) decreases as well.

To illustrate described dependency, the ACR and CRR values for trials when the regression based algorithm was used with different values of min_error (MER) parameter was presented in Fig 1.

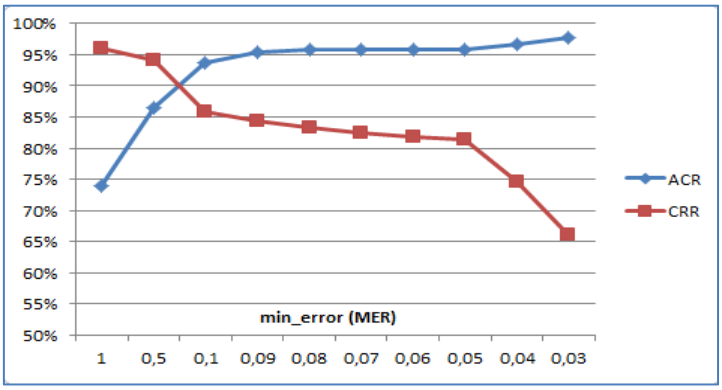

Fig. 1. ACR and CRR values for different min_error (MER) in regression based algorithm

The results obtained for both algorithms and both types of samples are shown in following tables. Table 1 presents values of the acceptance rate (ACR) and the correctness rate (CRR) for calibrated and non-calibrated trials, for the distance based (DIST) algorithm with different values of the proximity coefficient (PCF).

Table 1. CRR and ACR for different PCF for distance based algorithm

	Calibrated		Non-calibrated	
PCF	CRR	ACR	CRR	ACR
1	95%	94%	58%	68%
1.1	95%	94%	66%	57%
1.2	95%	93%	75%	46%

As it can be seen, the results for calibrated trials are quite good and stable for different values of PCF and the results are significantly better than for non-calibrated trials. As it could be expected, higher value of PCF increases the correctness (CRR) but in the same time decreases the acceptance (ACR). For PCF>=1.2 more than a half

This is a pre-print. The final version of the paper was published in Springer, Smart Innovation, Systems and Technologies, Vol. 39, 2015 as part of Proceedings of the 7th KES International Conference on Intelligent Decision Technologies (KES-IDT 2015) and is available in Springer Link library via http://dx.doi.org/10.1007/978-3-319-19857-6_31

of non-calibrated trials is rejected. The best value of ABS for non-calibrated trials is only 39% while it is about 90% for all PCF values, when only calibrated trials are taken into consideration.

The results for the regression based algorithm (REGR) are presented in Table 2. They were calculated for different values of *min_error* (MER) parameter.

Table 2. CRR and ACR for different MER values for regression based algorithm

	Calibrated		Non-calibrated	
MER	CRR	ACR	CRR	ACR
1	74%	96%	56%	94%
0.5	86%	94%	66%	91%
0.1	94%	86%	74%	80%
0.08	96%	83%	81%	73%
0.06	96%	82%	85%	61%
0.04	97%	75%	90%	44%

It is visible that the results for calibrated trials are worse than for DIST algorithm with ABS about 80%. However, the results for non-calibrated trials for the regression based algorithm are significantly better than for the distance based one, with ABS reaching 60% for MER=0.5. The algorithm is especially efficient in rejecting low quality trials. For instance, 74% correctness (CRR) was achieved for the acceptance rate (ACR) 80%, while for DIST algorithm the same correctness rate was achieved for ACR amounting only to 46%.

The next research question was how the simplest possible, one point calibration can improve the results. Because there were only trials with four points available the only way to check it was to assume that the first digit of PIN is known. For DIST algorithm it resulted in shifting fixations so that the first fixation overlapped the first (known) digit (see **Method** section for details). For REGR algorithm only PIN numbers starting with the known digit were considered as candidates (see **Method** section as well).

Table 3. Results achieved for DIST algorithm with the first fixation shift.

	Calibrated		Non-calibrated	
PC	CRR	ACR	CRR	ACR
1	96%	95%	76%	83%
1.1	96%	94%	81%	77%
1.2	96%	94%	85%	70%
1.3	98%	93%	89%	62%
1.4	98%	91%	91%	56%

This is a pre-print. The final version of the paper was published in Springer, Smart Innovation, Systems and Technologies, Vol. 39, 2015 as part of Proceedings of the 7th KES International Conference on Intelligent Decision Technologies (KES-IDT 2015) and is available in Springer Link library via http://dx.doi.org/10.1007/978-3-319-19857-6_31

When considering DIST algorithm (Table 3) the results for the calibrated trials are better for the case of fixation shifting but the difference is not significant (ABS is equal about 91% in most cases). However, the results for non-calibrated samples are significantly better with 63% for ABS, in the best case comparing to 39% for tests without shifting.

Table 4. Results for REGR algorithm with the first digit known

	Calibrated		Non-calibrated	
MER	CRR	ACR	CRR	ACR
0.5	87%	94%	69%	90%
0.1	95%	85%	83%	74%
0.08	96%	83%	88%	69%
0.06	98%	80%	91%	58%
0.04	98%	74%	92%	43%

The results for the regression based algorithm (Table 4) did not improve outcomes significantly when only PINs with a correct first digit were taken into account. Such condition reduced the number of PINs for which models were calculated ten times (from 7290 to 729) but it did not affect algorithms performance.

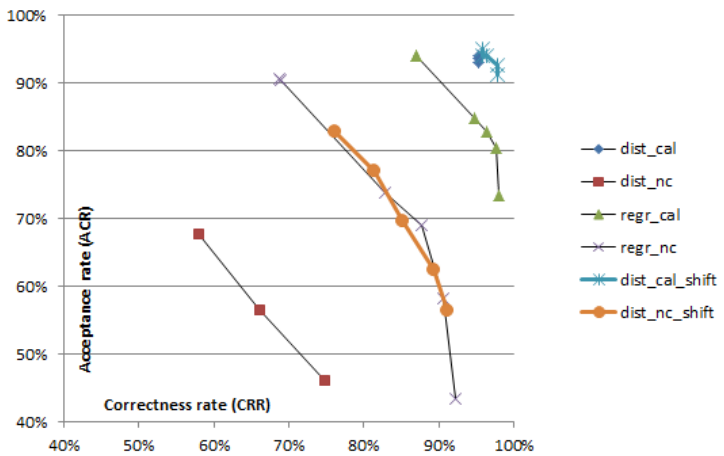

Fig. 2. ACR and CRR for different algorithms and calibrated (cal) and non-calibrated (nc) trials

The comparison of the algorithms and the sets was presented in Fig 2. The distance based algorithm performed very well for calibrated data (*dist_cal* and *dist_cal_shift*) while its results were very unsatisfactory for non-calibrated data (*dist_nc*). The regression algorithm was not as good as the distance based one for calibrated data (*regr_cal*) but it outperformed it for non-calibrated one (*regr_nc*). Adding information about the first digit improved outcomes of the distance based algorithm

This is a pre-print. The final version of the paper was published in Springer, Smart Innovation, Systems and Technologies, Vol. 39, 2015 as part of Proceedings of the 7th KES International Conference on Intelligent Decision Technologies (KES-IDT 2015) and is available in Springer Link library via http://dx.doi.org/10.1007/978-3-319-19857-6_31

(*dist_nc_shift*) but even with this information it is not better than the normal regression based outcome (*regr_nc*). As it was shown in Table 4, adding information about the first digit of PIN did not improve significantly the results for the regression based algorithm so it was not included in Fig 2.

5 Conclusions

The findings of the research can be divided into two groups. First of them regards trials proceeded by the per-user calibration. The results obtained for such recordings confirmed the possibility of using eyes for providing information of PIN type. Another conclusions may be drawn from outcomes obtained for various scopes of the introduced parameters - *proximity coefficient* (PCF) and *min_error* (MER). They show to what extent the size of the area of interests can be reduced not to decrease the efficiency of the method.

The second group of the findings concerns the problem of omitting a calibration process. The experiments presented in this paper showed that it is possible to use eye tracker as a pointer for simple and well defined tasks even without a prior per-user calibration. It is possible if such task does not require point to point gaze mapping, yet point to area of interests adjustment is acceptable. The studies of using eyes for PIN providing fulfil this requirement. The results obtained for the non-calibrated trials are worse than for the calibrated ones, however values of the analyzed factors indicated that in most cases proper values could be obtained.

A novel regression based algorithm was introduced and it was shown that it outperforms the distance based one for the non-calibrated samples. Additionally, it may be tuned for various types of interfaces using a *min_error* parameter. When the correctness of recognition is important, the *min_error* value may be increased and it was shown that results become more reliable (in sake of higher rejection rate). Such scenario may be useful for instance for gaze pointing of PIN at ATM when we want to be sure that PIN entered is correct even if the user is forced to enter it several times due to rejections. On the other hand there are applications in which an approximated gaze position is enough and rejections are rather undesirable for an interface to be fluent. That is the case of for instance interactive games. For such applications *min_error* value may be low, resulting in lower rejection but also with lower overall accuracy.

Providing opportunity for removing trials with the low quality before any analyses starts is the important contribution in improving the efficiency of data processing.

Additionally, it was shown that one point calibration enhances results for the distance based algorithm. However, the improvement for non-calibrated samples does not make this algorithm better than the regression based one. It shows that further studies on more complicated regression based algorithms for using the eye movement signal for human computer interaction may provide results improvement.

This is a pre-print. The final version of the paper was published in Springer, Smart Innovation, Systems and Technologies, Vol. 39, 2015 as part of Proceedings of the 7th KES International Conference on Intelligent Decision Technologies (KES-IDT 2015) and is available in Springer Link library via http://dx.doi.org/10.1007/978-3-319-19857-6_31

6 Bibliography

1. Blignaut, Pieter; Wium, Daniël. The effect of mapping function on the accuracy of a video-based eye tracker. In: Proceedings of the 2013 Con-ference on Eye Tracking South Africa. ACM, 2013. p. 39-46
2. De Luca, A.; Denzel, M.; Hussmann, H.: Look into my eyes!: Can you guess my password?. In: Proceedings of the 5th Symposium on Usable Privacy and Security. ACM, 2009. p. 7.
3. Duchowski, Andrew. Eye tracking methodology: Theory and practice. Vol. 373. Springer, 2007.
4. Harężlak K., Kasprowski P., Stasch M.: Towards accurate eye tracker calibration – methods and procedures. Proceedings of the 17th International Conference in Knowledge Based and Intelligent Information and Engineering Systems - KES2014. Procedia Computer Science, Vol. 35, pp. 1073 – 1081, Elsevier, 2014 (ScienceDirect)
5. Harężlak K., Kasprowski P.: Evaluating quality of dispersion based fixation detection algorithm. 29th International Symposium on Computer and Information Sciences. Information Sciences and Systems 2014. Springer Lecture Notes on Electrical Engineering, pp 97-104, 2014.
6. Holmqvist, Kenneth, et al. Eye tracking: A comprehensive guide to methods and measures. Oxford University Press, 2011.
7. Jacob, Robert JK. "The use of eye movements in human-computer interaction techniques: what you look at is what you get." ACM Transactions on Information Systems (TOIS) 9.2 (1991): 152-169.
8. Kasprowski P., Harężlak K., Stasch M.: Guidelines for the Eye Tracker Calibration Using Points of Regard. Information Technologies in Biomedicine. Volume 4. Advances in Intelligent Systems and Computing. Springer International Publishing, Vol. 284, pp. 225-236, 2014.
9. Kasprowski P., Harężlak K.,: Cheap and easy PIN entering using eye gaze. 3rd Conference on Cryptography and Security Systems. Annales UMCS Informatica, Volume 14, 1, pp 75-84, Lublin, 2014.
10. Kumar, M, et al.: Reducing shoulder-surfing by using gaze-based password entry. In: Proceedings of the 3rd symposium on Usable privacy and security. ACM, 2007. p. 13-19.
11. Ramanauskas, N.: Calibration of video-oculographical eye-tracking system. Electronics and Electrical Engineering, 2006, 8.72: 65-68.
12. Salvucci, D. D., Goldberg, J. H.: Identifying fixations and saccades in eye-tracking protocols. In Proceedings of the 2000 Symposiumon Eye Tracking Research and Applications, 71–78, NY: ACM Press (2000)
13. Shic, F., Chawarska, K., Scassellati, B.: The incomplete fixation measure. In Proceedings of the 2008 Symposium on Eye Tracking Research and Applications, 111–114, NY: ACM Press (2008)