Real-Time Feedback Towards Voluntary Pupil Control in Human-Computer Interaction: Enabling Continuous Pupillary Feedback

Juliane Georgi[1(✉)], David Kowalski[2], Jan Ehlers[1], and Anke Huckauf[1]

[1] Department of General Psychology, Ulm University,
Institute of Psychology and Education,
Albert-Einstein-Allee 47, 89069 Ulm, Germany
{juliane.georgi,jan.ehlers,anke.huckauf}@uni-ulm.de
http://www.uni-ulm.de/in/psy-paed/allgemeine-psychologie.html

[2] Ulm University, Institute of Microelectronics,
Albert-Einstein-Allee 43, 89081 Ulm, Germany
david.kowalski@uni-ulm.de
http://www.uni-ulm.de/in/mikro.html

Abstract. Since the late 90's pupil size variations have been considered a possible input channel in Human-Computer Interaction [7]. [4,5] showed that it is possible to manipulate pupil size via self-induced regulation strategies. A training based on graphical real-time pupillary feedback supported the learning process towards voluntary pupil size control. For successful learning the feedback has to be reliable, stable and on time. Taking this into account, spontaneous blinking poses one important problem during real-time feedback. This paper presents the process and elaboration of real-time data filtering methods. The final implementation consists of a two-state process. Blink replacement is achieved with a data-driven threshold. The filter was programed and tested in the framework of a study by [3]. The testing results were promising.

Keywords: Real-time pupillary feedback · Voluntary pupil size control · HCI · Data filtering · Spontaneous eye blinks

1 Introduction

The success of Human-Computer Interaction (HCI) grounds on the necessity of providing a suitable link between human physical input actions and technical devices. [7] stated that "the fundamental task in computer input is to move information from the brain of the user to the computer" (p. 1). He stressed the importance of finding faster, more natural and more convenient means for information transmission. [7] was one of the first to mention eye pupil diameter as a future possible input channel.

Since then, various studies have shown that pupil size changes are related to cognitive and affective information processing, for example [6,11]. Yet, so far, the

H.M. Fardoun et al. (Eds.): REHAB 2014, CCIS 515, pp. 104–114, 2015.
DOI: 10.1007/978-3-662-48645-0_10

majority of these studies have referred to pupil dynamics as merely passive, reactive information. Pupil size changes provide insight into affective experiences but seem to be defying any voluntary control, which is a crucial requirement in the case of feasible input channels. Ten years after [4,5,7] point to the possibility of intentional pupil size manipulations as a new way of transmitting information to a computer. They investigated the hypothesis that pupil variations can indirectly be controlled via several forms of physical and psychological self-regulation, for example positive thinking or changing the point of focus.

[4] provided real-time graphical feedback of the subjects' recorded pupillary behavior. This should allow the subjects to assess and learn to control their pupil variations, leading to voluntary pupil dilatations and constrictions. The authors see the quality and accuracy of the given feedback as essential to a successful learning process.

However, working with pupillary data raises two general problems. One is the sensitivity of the human pupil to changes in illumination, resulting in fluctuations of the recorded pupil size. This can be circumvented by providing constant lightning conditions, minimizing the possible influence of illumination on pupil size. The other one is the occurrence of spontaneous blinks. Blinking, characterized by a rapid closing and opening of the eyelid, causes the eye-tracker to record invalid values. During this process, the lid gradually occludes and uncovers the pupil leading to incorrect size measurements. The moment the eye is closed, the eye-tracker loses track of the pupil. For post hoc data analysis the method for removing the errors introduced by blinking does not have to meet any special requirements. Hence, with post hoc processing many methodical approaches are feasible.

However, in order to enable the control of a computer system by pupil size manipulations real-time eye tracking and feedback is required. Thus, handling blinks represents one important challenge. For providing real-time continuous feedback, the applied method has to be able to reliably detect and replace blinks by valid pupil size measurements without causing a too large delay and distortion in the data stream. Presentation of wrong or overly delayed values would disturb the feedback.

This paper presents the development and progress of an algorithm capable of fulfilling the above-mentioned specifications. In Sect. 2, the two filter approaches are presented, followed by their testing results and discussion. In Sect. 3 a conclusion is presented and implications for future work are elaborated.

2 Filtering Methods

The realization of real-time pupillary feedback needs to satisfy three premises, it has to be reliable, stable and on time. Reliability refers to the correctness and accuracy of the presented pupil size. Stability bears upon the ability to keep up reliable measurements when the pupil signal is disturbed which is particularly the case during blinking. The right timing assures that the depicted feedback values are related to the momentary pupil behavior at a specific time point.

Under these preconditions we developed and elaborated two filter algorithms for dealing with spontaneous eye blinks during real-time feedback.

Implementation and testing of the algorithms took place within the framework of a study by [3], evaluating the scopes and limits for training voluntary pupil control with real-time pupillary feedback. The subjects were instructed to try to control their pupil size as much as they can via self-induced imagination of positive or negative ideas, relaxation or by performing calculations. Figure 1 depicts the experimental setup and feedback schema presented to the participants. The basic feedback consisted of four circles. The use of graphical feedback instead of numeric values was considered to be more comprehensible. Likewise the presentation of circles was thought to be intuitive, since the shape of a circle is associated with the shape of the human pupil. The thicker black circle in the center represents the average pupil size of the subjects' baseline measurement. The two gray areas around the baseline circle depict the baseline average score plus/minus one standard deviation. These circles should serve the subjects as anchor points and remained constant during training. The dashed circle is the actual feedback. Its size varied according to the momentary pupil behavior. The fixation cross in the center should prevent the subjects' focus from drifting. For a more detailed description of the study refer to [3].

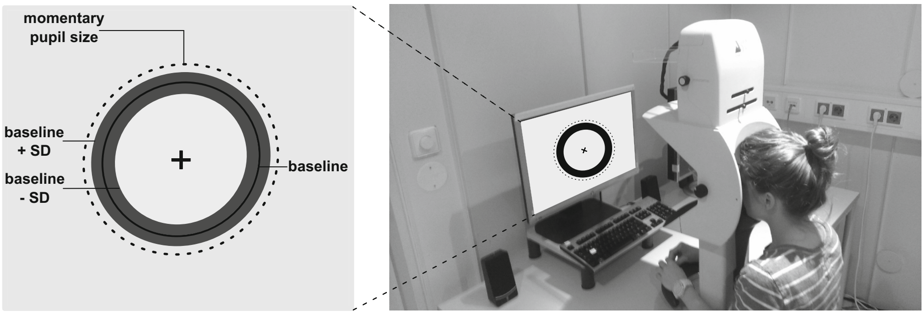

Fig. 1. Experimental setup and feedback schema as used in [3].

Pupil size was recorded with the SMI iView X™ Hi-Speed 1250 eye-tracker [13], featuring a sampling rate of 500 Hz. The experimental setup had a refresh rate of approximately 30 Hz, which implies one data point every 33.33 ms. The coding of the algorithms was done in PsychoPy v1.78 [12].

2.1 Value-to-Value Comparison Approach

Literature research of existing post hoc data analyzing methods indicates that one common rule to determine invalid data focuses on the comparison of consecutive measurement values. This approach is based on the assumption that the speed of pupil dilatation and constriction is physiologically limited. This proposes the definition of a criterion for the maximal allowed pupil diameter change during a fixed time interval. [9] have pointed out that applied criteria should best be derived from empirical data on actual physiological conditions. [10] investigated pupillary responses to emotionally provocative stimuli. They were recording pupil data with a 50 Hz eye-tracking system, hence the distance between two measurement points was 20 ms. They defined changes greater than 0.75 mm between two consecutive values as invalid measurements and removed them from the data set. In a subsequent study they halved their criterion to 0.375 mm [11]. Unfortunately, the authors did neither give a justification for their choice of criterion nor for its tightening in the second study.

Taken these considerations into account the first developed algorithm relies on a sample-to-sample comparison, as suggested by [10,11]. The filtering concept uses a criterion based on a study by [2]. This work provides a physiological foundation to the applied criterion. [2] investigated the amplitude and peak velocity of pupil constriction in the light reflex in 43 healthy subjects. The results yielded to an average peak velocity of 5.65 $\frac{\mathrm{mm}}{\mathrm{s}}$ ($SD = 1.17\,\frac{\mathrm{mm}}{\mathrm{s}}$). Adaptation of the given peak velocity by [2] to our sampling rate of 30 Hz leads to a maximum sample-to-sample change of approximately 0.19 mm, as can be seen in the following equation:

$$\begin{aligned} 5.65\,\frac{\mathrm{mm}}{\mathrm{s}} &= 0.0057\,\frac{\mathrm{mm}}{\mathrm{ms}} \\ 0.0057\,\frac{\mathrm{mm}}{\mathrm{ms}}\,33.33\,\mathrm{ms} &= 0.19\,\mathrm{mm}. \end{aligned} \tag{1}$$

The filtering process consists of the comparison of the absolute difference between each two measurement points to the criterion derived from equation (1). If a value is found to be exceeding this criterion, it is considered as invalid. The last valid measurement point is then taken as the comparison value for the following data points. Following the recommendation of [1], invalid values as well as zero values are replaced by the last valid value. Replacing ends once the last valid measurement and a following data point meet the criterion of 0.19 mm.

Testing Results and Discussion. Figure 2 depicts exemplary raw and filtered pupil diameter signals using the algorithm described in Sect. 2.1. The blinks are removed and replaced by the last valid measured value. However, the filtered data does not follow the raw data perfectly, especially between seconds 24 to 26. The lefthand side of Fig. 2 shows a close up of this time interval. A look at the raw data indicates a dilatation movement, consisting usually of two phases. The first one is a relatively steep and fast rising movement, whereas the second one is comparatively slower and continues until a certain base level is reached. This implies that the distance between two values, especially at the start of the upward movement could be somewhat bigger than 0.19 mm, as shown for the first two circled values. As a consequence the filter replaces the exceeding values by the last valid one and continues until one consecutive measurement meets the criterion, indicated for the third circled value.

These results point out one major drawback of this first filter implementation. They illustrate the trouble with keeping the comparison value fixed, once an

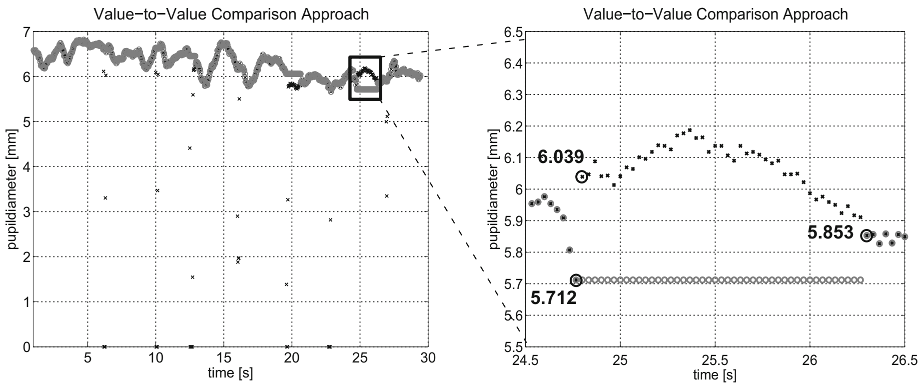

Fig. 2. Raw data signal (x) and filtered signal (o) after the Value-to-Value Comparison Approach. Close up of the time interval between 24 s to 26 s.

exceeding data point is found. Adherence of the comparison value leads to data replacements until a following measurement point matches the limit of 0.19 mm compared to the fixed value, potentially leading to massive data distortions.

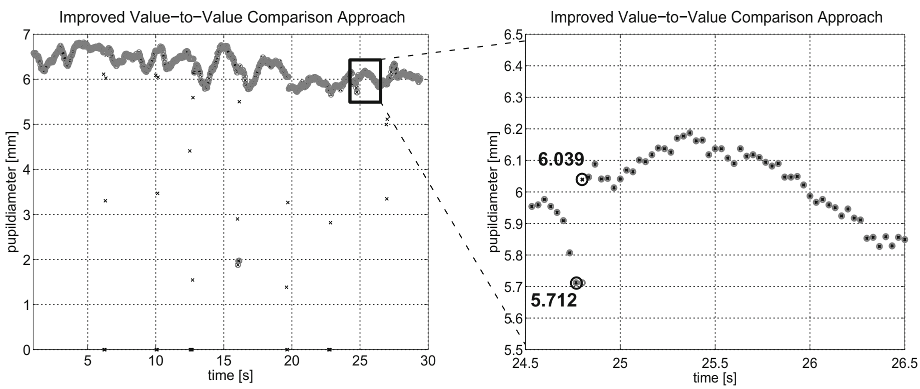

Fig. 3. Raw data signal (x) and filtered signal (o) after the Improved Value-to-Value Comparison Approach. Close up of the time interval between 24 s to 26 s.

In order to account for this issue, it is necessary to perform the value-to-value comparison independently of the last valid value. This requires buffering of two values and thereby introducing a delay. In the following explanation it is assumed that the first value is valid. The next incoming third measurement is then compared with the second. Two results are possible. If the comparison matches the criterion, then the second value becomes the next valid one and the following comparison is between the third and fourth value. If the criterion is not met, the first value is repeated, remains the valid value and the consecutive comparison is between the third and fourth value.

Figure 3 shows the same raw data and close up as in Fig. 2 with the modified filter. Studying the close up, it can be seen that the improvement leads to much less data distortions. However, correct data points are also replaced. This latter fact depicts two weak spots of a sample-to-sample change criterion.

The first problem is that the criterion is derived from a physiological characteristic. Therefore, it differs from person to person. This can be seen in the high standard deviation of the results by [2] (M=5.56 $\frac{\text{mm}}{\text{s}}$; SD=1.17 $\frac{\text{mm}}{\text{s}}$). These variations make it difficult to apply the same criterion to different persons. The second problem is that the applied criterion was based solely on average constriction speed and thus, disregarding the possibility that dilatation movements could be faster. A simple solution for both issues would be to enlarge the maximum allowed sample-to-sample change. However, the trade-off is that chances are increased to tolerate more invalid values.

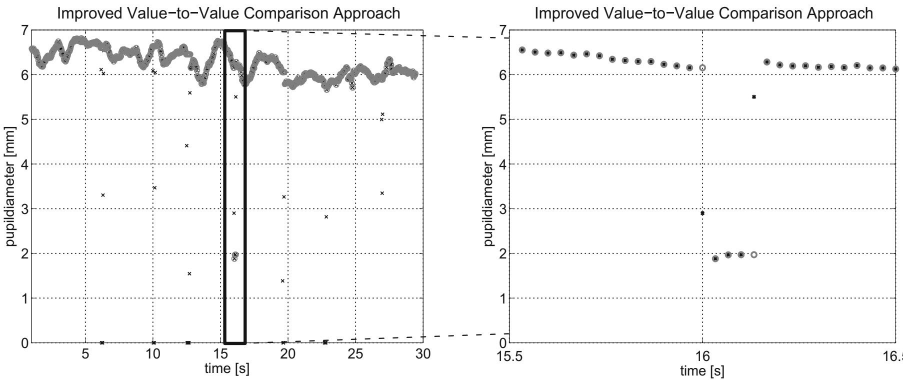

Fig. 4. Raw data signal (x) and filtered signal (o) after the Improved Value-to-Value Comparison Approach. Close up of the time interval between 15 s to 16 s.

Figure 4 illustrates another problem of the improved filter algorithm. Shortly after 16 s unusual small values are recorded. One reason for this might be, that the subject was blinking but did not close the eye completely. The close up on the left shows that the filter includes these values instead of replacing them, since the three adjacent values meet the criterion. This reveals a third weak spot of the applied criterion and filter. It does not check for unusual data.

As a consequence of these discussed problems of the first filtering concept an enhanced filtering algorithm was developed.

2.2 Enhanced Filter Approach (EFA)

The next proposed filter algorithm is built upon the improved value-to-value comparison approach and tries to overcome the implied problems. The implementation is based on a two-state-process and introduces the concept of a threshold. The threshold is a fixed value greater than zero, thereby indicating unusual small pupil diameter sizes, as well as blinks. This helps to detect blink movements at an earlier stage and to counteract the possibility of including unreasonable data. In this algorithm the value for the threshold is set to a maximal allowable pupil diameter of 3 mm. This choice is based on empirical experiences. Different than to the definition in the first algorithm, valid values are considered as three consecutive values which pairwise compared meet the allowable maximum deviation of 0.19 mm. This definition has proven to be more robust and reliable, than a simple value-to-value comparison. Figure 5 visualizes the basic working principle of the EFA algorithm, using a signal flowchart.

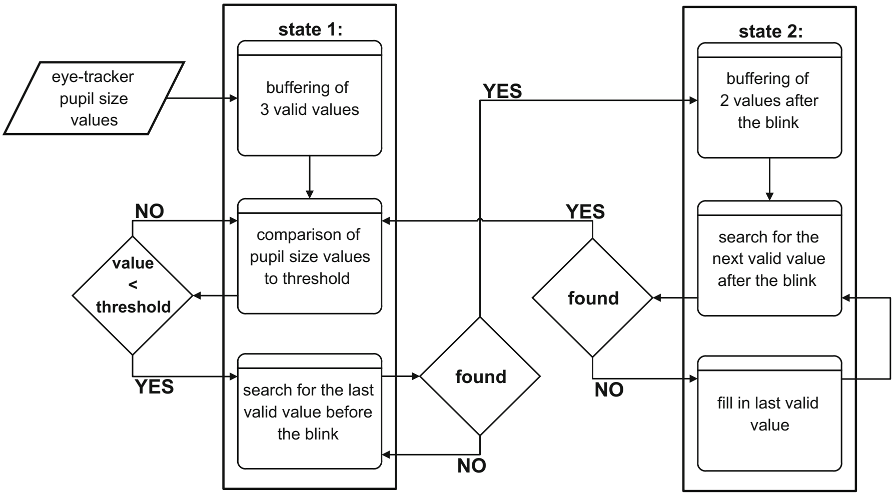

Fig. 5. Working principle of the EFA.

As shown in state one, each measurement starts with buffering three valid values. Three consecutive measurement points have to be greater than the threshold and meet the definition of valid values. This introduces a greater delay than in the previous filter where only two values need to be buffered. Nevertheless, these values constitute a save buffer, ensuring that a valid value is always given to replace invalid data. From then on, each incoming pupil size value is compared to the threshold. If it is greater, then it is considered valid and presented as feedback. If it is smaller, then a blink is detected and the actual filter process is activated. The EFA searches backwards in the preceding data for valid values, according to definition. The most recent valid value is taken as the replacement value and the filter switches to the second state. This state starts by buffering two values, thereby buying time to search for the next valid value in the prospective data. The two stored values are then again pairwise compared to the next incoming one, equivalent to the first process in state one. If they exceed the

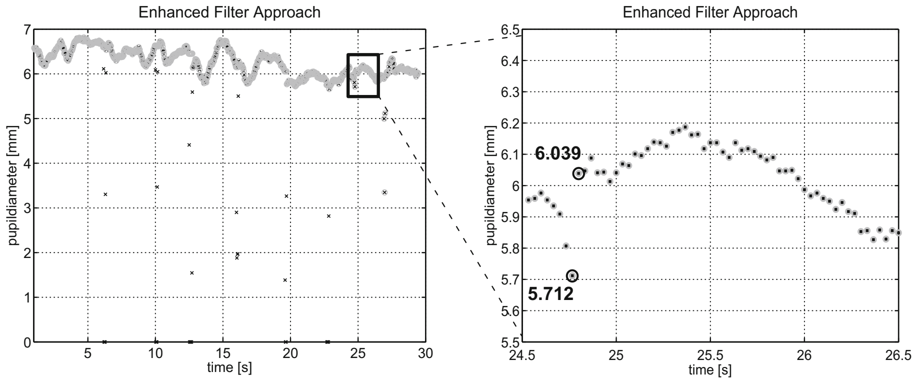

Fig. 6. Raw data signal (x) and filtered signal (o) after the EFA. Close up of the time interval between 24 s to 26 s.

threshold and are considered as valid, then the first one marks the endpoint of the blink and the filter switches back to state one. If they do not match both requirements, then the first value is replaced by the last valid value and the filter repeats the comparison process with the next incoming value.

Testing Results and Discussion. The EFA is compared to the previous filter concepts by using the same raw data signal and looking at the same time intervals.

Figure 6 depicts the time interval between 24 s and 26 s. The close up shows that the filter perfectly reproduces the raw signal course, in contrast to the previous approach (see Fig. 3).

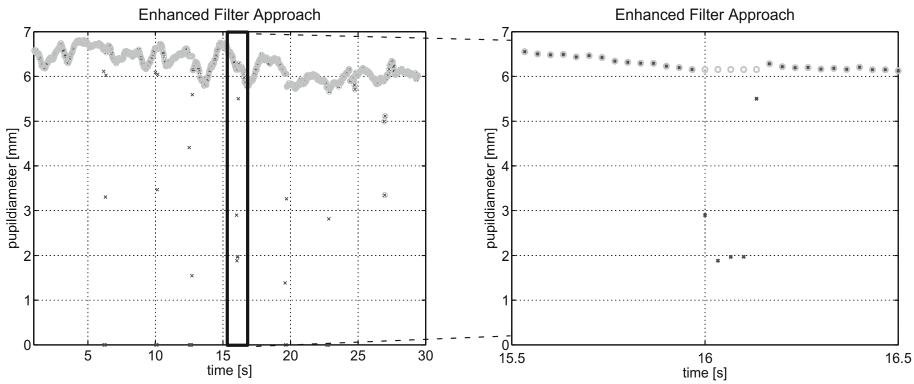

Fig. 7. Raw data signal (x) and filtered signal (o) after the EFA. Close up of the time interval between 15 s to 16 s.

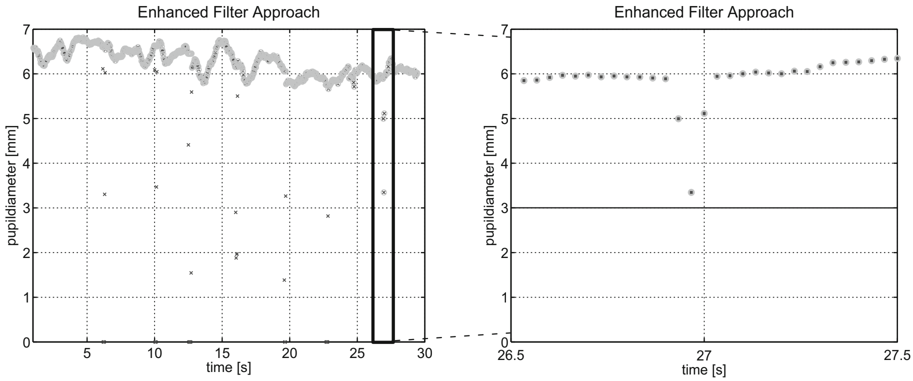

Fig. 8. Raw data signal (x) and filtered signal (o) after the EFA. Close up of the time interval between 26 s to 27 s.

The filter is also able to replace the unusual small pupil size recordings shortly around 16 s, as shown in Fig. 7. Both figures illustrate the benefits of the threshold. In contrast to a continuous comparison based on a fixed criterion, the threshold acts as a trigger to activate the search for valid values by pairwise comparison. This assures that valid values remain untouched by unusual large pupil size changes not associated to a blink.

Despite all of the mentioned aspects, introducing a constant threshold value does also create problems. The question to which value the threshold should be set is crucial. Figure 8 demonstrates this dilemma. The same data set but yet the invalid data around 26 s to 27 s, shown in the close up, are not intercepted. All three measurement points are greater than the threshold value.

3 Conclusion

This work presented the development and progress of filtering approaches in the context of live pupillary feedback. Feedback of pupillary behavior is thought to assist the learning process towards voluntary pupil control. [4,5] were the first to show that it is possible to indirectly manipulate pupil size variations with the help of self-regulation strategies. Real-time feedback of pupillary movements enhanced and facilitated this process. [4,5] and subsequently [3] paved the way for the possibility of pupil size control as a future input channel in HCI. A major problem of the realization of pupillary live feedback is the occurrence of blinks. Blinking leads to the recording of unusual small values. Presenting these values would interrupt the correct feedback of momentary pupil size. Therefore, such values need to be intercepted. While the removal of such values in post hoc data analysis does not need to meet any special requirements, the implementation of real-time filtering is a challenge. Real-time filters have to provide a reliable, stable and on time feedback. Feasible solutions for real-time filter concepts were elaborated. They were tested in the framework of a study by [3].

The first solution was a value-to-value comparison approach based on a maximal allowable sample-to-sample change criterion. The criterion was derived from average peak velocity of pupil constriction, given by [2]. This approach included a pairwise comparison of consecutive data points, with regard to the criterion. The implementation revealed the disadvantage of fixing the comparison value once two values were found to exceed the criterion. Potentially, this led to massive distortions in the filtered data, since the filter ceased to actually compare adjacent values.

In an improved version of this approach it was ensured that a continuous value-to-value comparison was maintained. However, the first results pointed out three major problems in applying a sample-to-sample change criterion. First, the derivation of the criterion from a physiological characteristic, introduces interindividual variations, thereby making it difficult to equally apply one fixed criterion. The second problem is the fact that the criterion based on average peak velocity of pupil constriction. Thus, this implementation did not respect that pupil dilatation movements could be faster. This led to replacement of correct data by the filter. A third weak spot is that this filter did not check for unusual data.

The EFA concept was based on the improved first approach and resolved its issues. The biggest advantage of this filter implementation, when compared with the first approach, is the use of a threshold. The threshold served as a trigger to activate the actual filtering, if there were values recorded which were smaller than the threshold. This helped to detect blinks and unusual data and, at the same time, ensuring that valid values remained untouched by unusual large pupil size changes not associated to a blink.

The results of the filter testings encouraged the implementation of the EFA. It could be proved that the EFA leads to better results as the first approach and overcomes its problems. Nevertheless, testing also revealed that the value of the threshold is crucial for successful filtering.

For a future development it is proposed to implement an adaptive threshold which is associated with actual pupil behavior. This should make it possible to identify unusual values dependent on the momentary signal trend, leading to more robust and reliable results. Furthermore, it is suggested to individualize the sample-to-sample change criterion, making it less susceptible to interindividual variations. Therefore, at the beginning of the data collection, the criterion could be adjusted by measuring a subjects peak velocity of constriction. Moreover, it might be possible to use linear interpolation to approach the actual signal course during the replacing of invalid values, as it was proposed by [8].

The current filter implementation and the proposed improvements still need to be investigated. It is important to further examine them in the context of real-time pupillary feedback, working towards a continuous, stable and direct feedback, which ensures a successful training on voluntary pupil size manipulation. This is a critical step towards the examination of how useful pupil size variations can be as one further input channel in HCI.

References

1. Bernhardt, P.C., Dabbs, J.M., Riad, J.K.: Pupillometry system for use in social psychology. Behav. Res. Methods. Instrum. Comput. **28**, 61–66 (1996)
2. Bremner, F.D.: Pupillometric evaluation of the dynamics of the pupillary response to a brief light stimulus in healthy subjects. Invest. Ophthalmol. Vis. Sci. **53**, 7343–7347 (2012)
3. Ehlers, J., Bubalo, N., Loose, M.C.A., Huckauf, A.: Towards voluntary pupil control - Training affective strategies?. Manuscript submitted for publication (2014)
4. Ekman, I., Poikola, A., Mäkäräinen, M., Takal, T., Hämäläinen, P.: Voluntary pupil size change as control in eyes only interaction. In: Proceedings of the 2008 Symposium on Eye Tracking Research & Applications, pp. 115–118. ACM, New York (2008)
5. Ekman, I., Poikola, A., Mäkäräinen, M.: Invisible eni: using gaze and pupil size to control a game. In: CHI 2008 Extended Abstracts on Human Factors in Computing Systems, pp. 3135–3140. ACM, New York (2008)
6. Hyönä, J., Tommola, J., Alaja, A.M.: Pupil dilation as a measure of processing load in simultaneous interpretation and other language tasks. Q. J. Exp. Psychol. A **48**, 598–612 (1995)
7. Jacob, R.J.K.: The future of input devices. ACM Comput. Surv. **28**, 177–179 (1996)
8. Marshall, S.P.: Method and Apparatus for Eye Tracking and Monitoring Pupil Dilation to Evaluate Cognitive Activity. US Patent No. 6,090,051 (2000)
9. Merritt, S.L., Keegan, A.P., Mercer, P.W.: Artifact management in pupillometry. Nurs. Res. **43**, 56–59 (1994)
10. Partala, T., Jokiniemi, M., Surakka, V.: Pupillary responses to emotionally provocative stimuli. In: Proceedings of the 2000 Symposium on Eye Tracking Research & Applications, pp. 123–129. ACM, New York (2000)
11. Partala, T., Surakka, V.: Pupil size variation as an indication of affective processing. Int. J. Hum-Comput. St. **59**, 185–198 (2003)
12. Peirce, J.W.: PsychoPy - psychophysics software in python. J. Neurosci. Methods. **162**, 8–13 (2007)
13. SensoMotoric Instruments, iView X™ Hi-Speed 1250. http://www.smivision.com/en/gaze-and-eye-tracking-systems/