

# Improving voluntary pupil size changes for HCI

Jan Ehlers  
University of Ulm  
Dept. General Psychology  
89069 Ulm  
+49(0)731-5031154  
jan.ehlers@uni-ulm.de

Juliane Georgi  
University of Ulm  
Dept. General Psychology  
89069 Ulm  
+49(0)731-5031154  
juliane.georgi@uni-ulm.de

Anke Huckauf  
University of Ulm  
Dept. General Psychology  
89069 Ulm  
+49(0)731-5031150  
anke.huckauf@uni-ulm.de

## ABSTRACT

Previous research [11] refers to pupil size as a passive information channel that provides insight into the affective state of the viewer but defies any voluntary control. However, since physiological arousal is influenced by various cognitive processes, we assume that pupil behavior can be brought under control by strategies of emotional regulation and cognitive processing. In the present paper we provide a methodological approach for examining the potentials and limits of active control of pupil dilation. Based on [3], we developed methods applying graphical feedback on systematic pupil diameter changes to utilize mechanisms of operant conditioning to gradually enable voluntary control over pupil size. Calculation models are introduced to carefully disentangle task relevant and irrelevant pupil dynamics. Based on mean values, single measuring and interpolation, we conceived computational rules to validate pupil data in real-time and determine criteria for artefact rejection. Extensive research based on the depicted methodology may shed further light on learning achievements related to emotional control and will reveal the potential of pupil-based input channels for the future development of affective Human-Computer Interfaces.

## Categories and Subject Descriptors

H.5.2 [Information Interfaces and presentation]: User Interfaces – Input devices and strategies.

## General Terms

Human Factors, Algorithms, Performance, Experimentation, Design

## Keywords

Affective Human-Computer Interface, Pupil size, Biofeedback, Emotions, Voluntary control

## 1. INTRODUCTION

Emotions are an inherent part of our lives; modern research [e.g. 9] unanimously stresses their considerable influence on human cognition and behavior. Accordingly, utilization of emotional responses becomes more prevalent even in the traditionally rational concept of man-machine interaction. Several studies

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

*Conference '10*, Month 1–2, 2010, City, State, Country.  
Copyright 2010 ACM 1-58113-000-0/00/0010 ...\$15.00.

raised various physiological measures to provide the computer system with information of the user's affective state, including galvanic skin responses [1] or heart rate [4]. Size and responsiveness of the human pupil may constitute an additional input channel for affective HCI; previous studies [e.g. 6; 11] suggest that the associated dynamics are related to both, cognitive and affective information processing.

Albeit variations in pupil size have been considered early to make up a possible input channel in HCI [7], extensive research efforts did not happen. The vast majority of studies [e.g. 11] refer to pupil dynamics as a passive information channel providing insight into affective experiences but defying any voluntary control. [5] reported pupil dilations up to 20% due to cognitive load, such as solving arithmetic problems; the magnitude of dilation appears to be proportional to task difficulty [8]. [11] find larger expansions during both, emotionally negative and positive stimuli compared to neutral conditions. In particular they report an onset of pupil dilation at about 400 ms and a gradual increase up to 2 or 3 seconds after stimulus onset.

[3] indicate that pupil behavior can be influenced intentionally by strategies of emotional regulation or cognitive processing. They find pupil size changes of about 20 % due to self-induced positive or negative emotions. However, they also report large variability between subjects in the general ability as well as huge variations in magnitude of the effect. We assume the true potential of pupil size-based mechanisms in HCI to be considered only when certain aspects of iteration and training are taken into account. Therefore signal processing has to provide learners with feedback that distinguishes precisely between intended and uncorrelated fluctuations and enables individual improvement in performance.

The present paper depicts methodical approaches to evaluate the potential and limits of pupil dilation in HCI. The aim is to train users while providing feedback on pupil size changes. For this purpose three algorithms compensating in real-time for task-unrelated samples and carefully smoothing the validated data for optimized usability are presented and compared. The proposed feedback scheme is conceived to display the functional separation between ordinary and strategy-based variations in pupil size and thus should facilitate training to control pupil dilation. First empirical results are exemplarily depicted. Succeeding investigations have to examine the question of how usable pupil control may become.

## 2. REAL-TIME BIOFEEDBACK ON PUPIL SIZE CHANGES

### 2.1 Aims of Real-Time Expansion Smoothing

According to [14] pupil size variations remain relatively stable in alert subjects; however, even under constant lighting conditions

reflexive constrictions, spontaneous fluctuations and especially eye blinks still pose a problem in pupil recordings; specifically, when considering real-time feedback. During blinks, eyelid impacts cause the eye tracker signal to be lost and zero-values in the pupil diameter are recorded. Furthermore, the lid gradually covers the pupil, resulting in distorted values and thus preventing reliable measurement during affective computing. Hence, pupil size information which is fed back to a user should contain intended size changes but be cleaned from physiological fluctuations. [10] conducted post-hoc analysis after auditory emotional stimulation and judged sudden increases or decreases of 0.75 mm within a 20 ms interval (recorded with a 50 Hz eye-tracking system) as artifacts. At a later stage they applied stricter criteria and rejected samples that include changes of at least 0.375 mm during the same period [11].

Subsections 2.2 to 2.4 depict three different algorithms to validate the transferred data. The search for an appropriate criterion that covers task-relevant variations and distinguishes them from invalid values constitutes the primary focus of the procedures. Of course, an elaboration of the best possible method is an iterative process and still ongoing.

such a case the new value is kept in the data and part of the upcoming averaging procedure whereas the oldest measurement in the mean drops out. If the current value exceeds the predefined boundary it is rejected and the twelfth is used as a replacement. The columns in Figure 1 illustrate the course of pupil size variations during a 15 second baseline computation (extract of 12 seconds depicted). The average-based computation (left) rejects blink-related data; however, due to little up-to-date information within the averaging process (illustrated in the current mean curve), signal changes at about 0.5 and 8.3 seconds are lost and valid data is rejected.

### 2.3 Single Value Approach

Considering the difficulties outlined above, we applied stricter criteria for the valid range of values in a second filtering concept and moved away from the averaging procedure by trying a sample-to-sample approach. [2] studied the correlation between amplitude and peak velocity of pupil constriction to the light reflex in normal subjects. The results showed a mean amplitude of 1.92 mm (SD: 0.39) and an average peak velocity of 5.65 mm/sec (SD: 1.17).

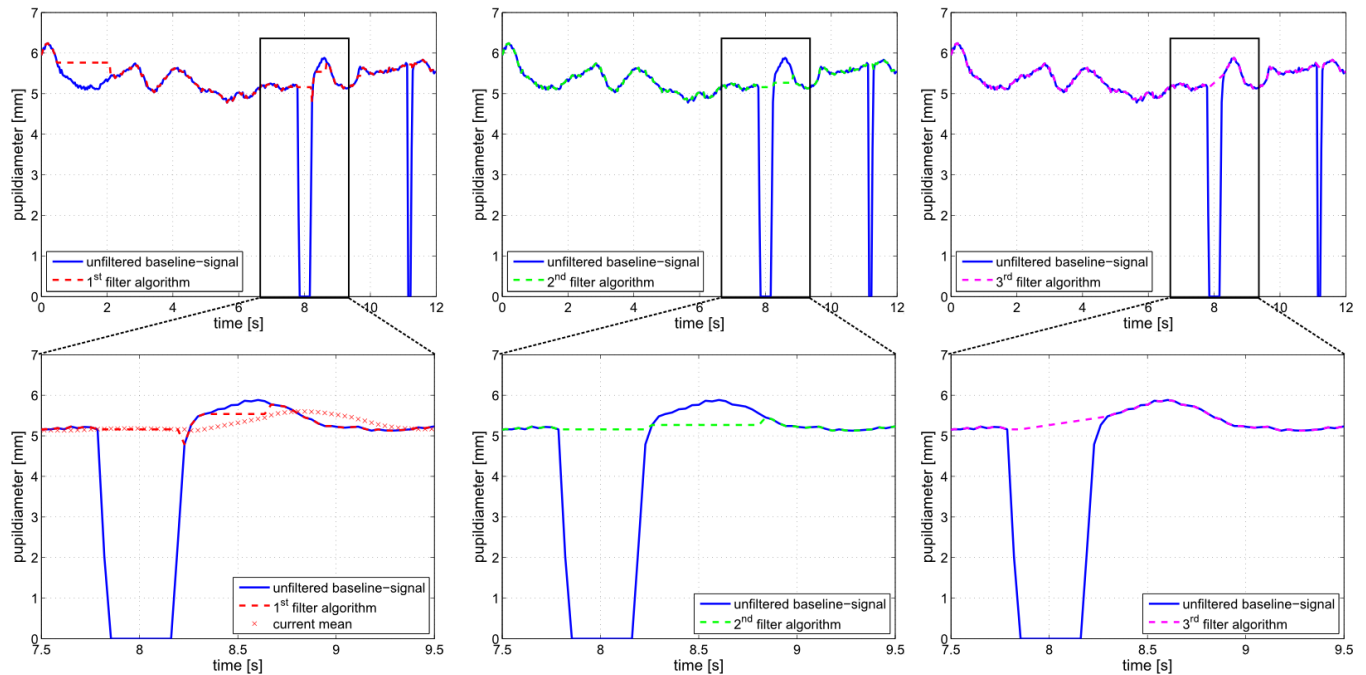


Figure 1. Results of different filter testing.

#### 2.1.1 Apparatus

To evaluate the algorithms we recorded pupil sizes with the SMI iView X™ Hi-Speed 1250 [13] featuring a sampling rate of 500 Hz. The experimental monitor had a refresh rate of approximately 30 Hz. This latter fact defines both, the basic conditions and the restrictions of our data filtering methods.

### 2.2 Mean Value Approach

To compensate unrelated contractions and re-dilations we continuously averaged over a window of twelve measurements adding up to a 400 ms interval at a 30 Hz sampling rate. Any value is compared to the preceding window and classified as valid while it meets the predetermined criterion of 0.375 mm [11]. In

For the current implementation the reported peak velocity is used as a limit for detecting (un-)valid measurements. Converting the specifications to our sampling rate of 30 Hz leads to an allowable sample-to-sample change of 0.1883 mm. If the distance between two values exceeds this range then the latest measurement is substituted by the last valid value. The middle column of Figure 1 shows the application of the single value procedure; blink artifacts are rejected and, in contrast to the first method, a valid pupil contraction at the outset is detected. Still, the dynamics subsequent to the first blink are rejected as well. Due to the stricter criterion, variations are sorted out at an earlier stage compared to the mean value computation.

## 2.4 Interpolation Approach

The interpolation method builds upon the criterion of the single value implementation and combines it with the idea of linear interpolation. It is activated when eyelid movements (zero-values) are detected. The algorithm marks the starting point of a blink for the interpolation ( $pd_s(t_s)$ ). For interpolation, three prior values are considered; size-differences are calculated and compared to the criterion of 0.19 mm as the highest permitted value. The number of comparisons is adjustable; after studying various raw data-signals, we decided to have five comparisons executed. The latest valid value is consulted. Simultaneously, the incoming data to the right of the zero-value is monitored. Once a measurement greater than zero appears in the dataset, the value is taken as indication of the endpoint of interpolation. Again, the following three data points are pairwise compared according to the second criterion until the earliest valid value is found. This defines the end-point ( $pd_e(t_e)$ ). Interpolation is then performed according to the following equation:

$$pd(t) = pd_s(t_s) + (t - t_s) \frac{(pd_e(t_e) - pd_s(t_s))}{(t_e - t_s)}.$$

Execution of this algorithm leads to an eligible delay. If a predefined time lag is exceeded, the last valid value (interpolation starting point ( $pd_s(t_s)$ )) will be repeated. In contrast to the aforementioned approaches, the interpolation algorithm detects the complete valid dynamics in the exemplary dataset (Figure 1, right column).

## 2.5 Discussion of Filtering Methods

Problems using the mean value approach become particularly apparent during rapid pupil size variations: due to little up-to-date information within the averaging procedure, changes of mean values do not sufficiently meet the real dynamics and further variations on a new expansion-level remain unobserved until another decline in pupil diameter. Consulting fewer values in the averaging process or applying weighted means may be used to counteract these tendencies. It is debatable, whether the criterion of 0.375 mm that originally relied on a sample-to-sample comparison in a 20 ms interval, is suitable for a time period of 400 ms which allows much more variance between the included values. It would therefore be reasonable to reduce the averaging time interval or augment the boundary of 0.375 mm. Additional criteria is also needed to prevent the algorithm from including extreme measurement during eye blinks; distorted mean values due to such information lower the average and repeatedly led to the auto removal of complete datasets.

Calculation procedures on the basis of the single value algorithm may lead to a similar loss of data, given that computation starts with a zero-value – however, a quite unlikely case. Nevertheless, the interval right after the blink is cut off. This behavior can be considered a fundamental weakness of such a single value approach, since it reacts not only to sudden changes, indicating a blink, but also to other probably valid increases or decreases in the dataset. In case of a steep rise, of two to three seconds as outlined by [11], there is a risk that once the algorithm has detected two values exceeding the criterion, it may lose complete track of the rising movement and continuously repeat the last valid value, thereby distorting the reported feedback of pupil size. The rejection of valid dynamics within the exemplary data record demands further examinations of the current selection criterion. The interpolation approach bypasses this issue. But in contrast to

the previous algorithms, the computation entails a short delay of the output data stream. However, with regard to the temporal dynamics of pupil size variations during affective processing [11], this might be negotiable.

## 3. TRAINING VOLUNTARY PUPIL SIZE CHANGES

### 3.1 Conception of Training Sessions

A 15-seconds calibration procedure while looking on a light grey background precedes every task. The acquired pupil sizes are averaged and depicted as a black circle (Figure 2). A grey colored ring depicts the standard deviation of the baseline. These variations should primarily depict task-unrelated physiological fluctuations and serve as a reference to evaluate own efforts in voluntary pupil expansion/contraction during the specific tasks. Real-time feedback according to the single value approach (section 2.3) is provided in shape of a red circle. Pre-tests showed that constantly received feedback on the basis of untreated values leads to shaky expansion movements that are difficult to handle. As a consequence the last two forwarded values are averaged to smooth the feedback dynamics and ensure increased usability.

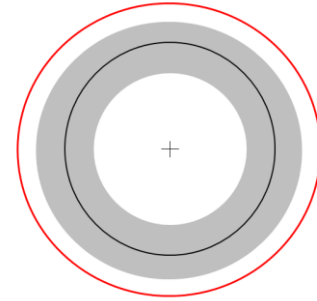


Figure 2. Feedback scheme.

### 3.2 Initial Results

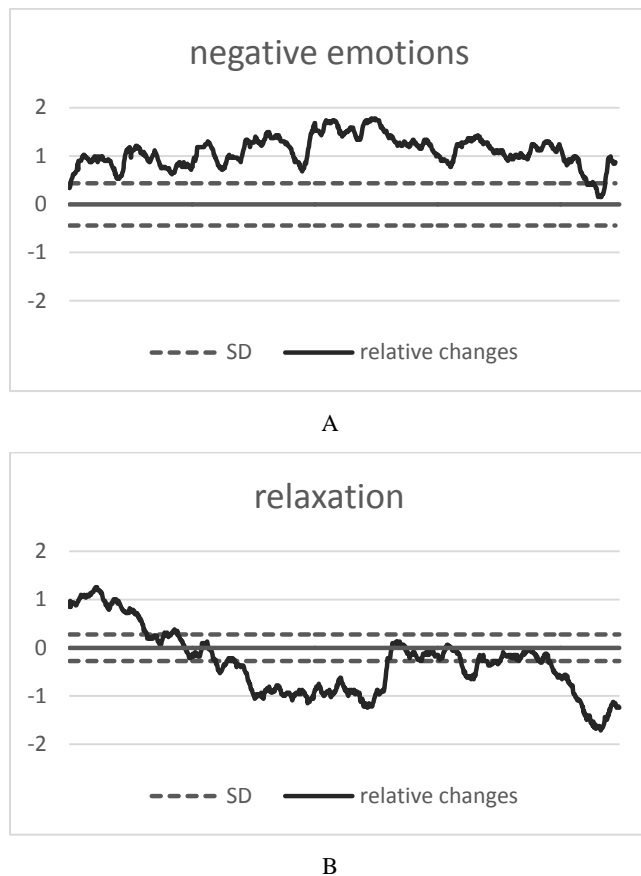
Figure 3 depicts first efforts of a subject to voluntary control variations in pupil size during a 30 seconds processing time. Pupil dynamics are pictured relative to individually determined baseline calculations (15 seconds) preceding every task. Continuous feedback was provided according to the descriptions in section 3.1. Baseline mean is set to zero on the y-axis; the dotted lines illustrate limits of the standard deviation. Calculations are based on the single value algorithm (section 2.3).

Compared to a neutral baseline, utilizing emotional associations (here: the imagination of negative autobiographical memories) leads to a gradual increase of pupil size variations (3 A) during the first 15 seconds, similar to findings in [11]. By contrast, self-induced relaxation is associated with a linear 15-second decline below variations of the neutral condition (3 B).

## 4. CONCLUSION

The depicted empirical findings tend in the same direction as reported by [3]. However, they should be considered carefully, since suitable measures for relevant pupil dilations in HCI (e.g. latency, maximum, mean) still need to be specified and evaluated. Longer-term studies may then reveal the true potential of voluntary controlled pupil size variations in HCI. Likewise, the proposed filter algorithms need further testing and elaboration. They are yet to be compared and further examined, particularly under extreme conditions, e.g. when prolonged eyes-closed

periods contaminate the data sample or subjects blink rapidly and several times in succession.



**Figure 3. Relative changes in pupil size due to experimental instructions. Explanations see text. The abscissa depicts time, the ordinate corresponds to relative changes in pupil size.**

## 5. ACKNOWLEDGMENTS

We like to thank Daniel Staiger for valuable programming assistance, Nikola Bubalo for critical discussions and David Kowalski for continuous support. This research was supported by the Collaborative Research Center (sfb transregio 62) by the Deutsche Forschungsgemeinschaft (DFG).

## 6. REFERENCES

- [1] Ark, W., Dryer, D. and Lu, D. 1999. The Emotion Mouse. In *Human-Computer Interaction: Ergonomics and User Interfaces*, H. J. Bullinger and J. Ziegler, Eds. Lawrence Erlbaum, 818-823.
- [2] Bremner, F. D. 2012. Pupillometric Evaluation of the Dynamics of the Pupillary Response to a Brief Light Stimulus in Healthy Subjects. *Investigative Ophthalmology & Visual Science* 53 (October 2012), 7343-7347. DOI= [10.1167/iovs.12-10881](https://doi.org/10.1167/iovs.12-10881).
- [3] Ekman, I., Poikola, A., Mäkäräinen, M., Takal, T. and Hämäläinen, P. 2008. Voluntary Pupil Size Change as Control in Eyes Only Interaction. In *Proceedings of the 2008 symposium on Eye tracking research & applications*. ETRA '08. ACM, New York, NY, 115-118. DOI= [10.1145/1344471.1344501](https://doi.org/10.1145/1344471.1344501).
- [4] Healey, J. and Picard, R.W. 2005. Detecting stress during real-world driving tasks using physiological sensors. *IEEE Transactions on Intelligent Transportation Systems* 6 (June 2005), 156-166. DOI= [10.1109/TITS.2005.848368](https://doi.org/10.1109/TITS.2005.848368).
- [5] Hess, E. H. 1972. Pupillometrics. In *Handbook of Psychology*, N.S. Greenfield and R. A. Sternbach, Eds. Holt, Rinehart & Winston, 491-531.
- [6] Hyönä, J., Tommola, J. and Alaja, A. M. 1995. Pupil dilation as a measure of processing load in simultaneous interpretation and other language tasks. *Q. J. Exp. Psychol. A* 48 (August 1995), 598-612.
- [7] Jacobs, R. J. K. 1996. The future of input devices. *ACM Comput. Surv.* 28, 4es, Article 138 (December 1996). DOI= <http://doi.acm.org/10.1145/242224.242400>.
- [8] Kahneman, D. and Beatty, J. 1966. Pupil diameter and load on memory. *Science*, 154 (December 1966), 1583-1585. DOI= [10.1126/science.154.3756.1583](https://doi.org/10.1126/science.154.3756.1583).
- [9] Loewenstein, G. and Lerner, J. S. 2003. The role of affect in decision making. In *Handbook of Affective Science*, R. Davidson, H. Goldsmith and K. Scherer, Eds. Oxford University Press, 619-642.
- [10] Partala, T., Jokiniemi, M. and Surakka, V. 2000. In *Proceedings of the 2000 symposium on Eye tracking research & applications*. ETRA '00. ACM, New York, NY, 123-129. DOI= <http://doi.acm.org/10.1145/355017.355042>.
- [11] Partala, T. and Surakka, V. 2003. Pupil size variation as an indication of affective processing. *International Journal of Human-Computer Studies* 59 (July 2003), 185-198. DOI= [http://dx.doi.org/10.1016/S1071-5819\(03\)00017-X](http://dx.doi.org/10.1016/S1071-5819(03)00017-X).
- [12] Picard, R. W. 1997. *Affective Computing*. MIT Press, Cambridge, MA.
- [13] SensoMotoric Instruments, *iView XTM Hi-Speed 1250*, <http://www.smivision.com/en/gaze-and-eye-tracking-systems/>.
- [14] Wilhelm, B., Giedke, H., Luëdtke, H., Bittner, E., Hofmann, A. and Wilhelm, H. 2001. Daytime variations in central nervous system activation measured by pupillographic sleepiness test. *Journal of Sleep Research* 10, 1-7.