

Looking around with your brain in a virtual world

Plass-Oude Bos, D.; Duvinage, M.; Oktay, O.; Guruler, H.; Istanbulu, A.; Laar, B. van de; Poel, M.; Roijendijk, L.M.M.; Reunderink, B.; Tonin, L.; Bahramisharif, A.; Vliet, M. van
2010, Article in monograph or in proceedings (Salah, A.A.G.; Gevers, T. (ed.)), Proceedings
eINTERFACE'10 : 6th International Summer Workshop on Multimodal Interfaces July 12 - August 6,
2010 University of Amsterdam, The Netherlands, pp. 12-23)

Version of the following full text: Publisher's version

Downloaded from: <https://hdl.handle.net/2066/83539>

Download date: 2026-05-05

Note:

To cite this publication please use the final published version (if applicable).

Looking Around with Your Brain in a Virtual World

Danny Plass-Oude Bos, Matthieu Duvinage, Oytun Oktay, Jaime Delgado Saa, Huseyin Gürüler, Ayhan Istanbulu, Marijn van Vliet, Bram van de Laar, Mannes Poel, Linsey Roijendijk, Luca Tonin, Ali Bahramisharif, Boris Reuderink

Abstract—Offline analysis pipelines have been developed and evaluated for the detection of covert attention from electroencephalography recordings, and the detection of overt attention in terms of eye movement based on electrooculographic measurements. Some additional analysis were done in order to prepare the pipelines for use in a real-time system. This real-time system and a game application in which these pipelines are to be used were implemented. The game is set in a virtual environment where player is a wildlife photographer on an uninhabited island. Overt attention is used to adjust the angle of the first person camera, when the player is tracking animals. When making a photograph, the animal will flee when it notices it is looked at directly, so covert attention is required to get a good shot. Future work will entail user tests with this system to evaluate usability, user experience, and characteristics of the signals related to overt and covert attention when used in such an immersive environment.

Index Terms—Multimodal interaction, brain-computer interfacing, covert attention, eye tracking, electroencephalography, electrooculography, virtual environment, usability, user experience.

I. INTRODUCTION

So far, most brain-computer interfaces seek to replace traditional input modalities, like mouse or keyboard. However, current electroencephalography-based brain-computer interfaces (EEG-based BCIs) have considerable problems: low speed, low detection accuracies which varies highly between users, low bandwidth, sensitivity to noise and movement, often requiring training, and expensive and cumbersome hardware [1]. These make it difficult to make such BCIs an interesting input method for able-bodied users.

Allison et al. mention a number of considerations for BCI applications for this healthy user group [1]. In this report we touch upon some of them (extending the term BCI to interface using neurophysiological signals):

- **Hybrid BCI:** using BCI in combination with other input signals, either as independent command signal or as a modifier of commands from other inputs.
- **Induced disability:** in circumstances where conventional interfaces are not usable, BCI could function as a replacement, or when they provide not enough bandwidth, BCI could function as an extra input channel.
- **Mapping between cognition and output:** make systems natural in their use by letting the system respond in a way that

D. Plass-Oude Bos, W.M. van Vliet, B.L.A. van de Laar, M. Poel, and B. Reuderink are with HMI group, EEMCS Faculty, University of Twente, The Netherlands. M. Duvinage is with TCTS Lab, Electrical Engineering University of Mons, Belgium. M.O. Oktay is with Electronics and Communications, Corlu Engineering, Namik Kemal University, Turkey, and was partially supported by TUBITAK Research Grant No. 109E202. J. Delgado Saa is with Faculty of Engineering and Natural Sciences, VPA Lab, Sabanci University, Turkey, and IEE Group of Robotics and Intelligent Systems, Universidad del Norte, Colombia. H. Gürüler is with Department of Electronics and Computer Education, Faculty of Technical Education, Mugla University, Turkey. A. Istanbulu is with Faculty of Computer Engineering, Balikesir University, Turkey. L. Roijendijk is with Radboud University Nijmegen, Donders Institute for Brain, Cognition and Behaviour, Nijmegen, The Netherlands. L. Tonin is with CNBI, École Polytechnique Fédérale de Lausanne, Switzerland. Ali Bahramisharif is with Radboud University Nijmegen, Donders Institute for Brain, Cognition and Behaviour, and Radboud University Nijmegen, Institute for Computing and Information Sciences, both in Nijmegen, The Netherlands.

corresponds to what the user would expect. The interaction does not only consist of the system response however, but also of the user action [2]. Therefore, we would like to propose to extend this definition to include: to use brain activity or mental tasks that come naturally given the situation. This ensures that the system is most intuitive in the interaction, requiring no user learning or memorization.

- **Accessing otherwise unavailable information:** some processes have no outside expression (whether it is just a mental process, or the user is purposefully trying to inhibit such expressions), but could be detected from brain signals.

We developed a system that makes use of naturally-occurring neurophysiological activity to augment the user interaction with a virtual environment, which already uses conventional mouse and keyboard controllers, in a natural way. The main mode of feedback from any computer system is visual, through the computer screen, thus when looking for natural interaction it makes sense to look into tasks that are related to vision: overt and covert attention. Jacob and Karn mention that it is quite difficult to have the system respond to eye gaze in a natural way, which also happens in the real world [2]. The only example they give is human beings: people respond to being looked at, or what other people are looking at. In our prototype, we use this natural response by letting an animal flee when looked at directly. This induces a situational disability (animals cannot be looked at directly), which is solved by using covert attention to get a good view of the creature. But we also show another option for the natural mapping of eye input: when we move our eyes, our view changes. This natural mapping can be translated to adjusting a first person camera in a virtual environment based on the user's eye movement.

Our report will first dive into covert and overt attention, providing background information, the design and evaluation of the pipelines for signal processing and classification, and answering issues related to the use of these pipelines in an online, real-time setting. After this, the whole system is described, with the game application in particular, followed by a description of the online user evaluation experiments we plan to do.

II. COVERT ATTENTION

Covert attention is the act of mentally focusing on a target without head or eye movements [3]. While overt attention is said to be an indication of place of focus, covert attention is a possible confound. By detecting both, all options for spatial attention are covered. There is also a theory that covert attention guides saccadic movement, and that it is possibly a mechanism to scan the visual field for points of interest [4].

Offline experiments have shown that when attention is directed to the left visual hemifield, alpha activity decreases in the right posterior hemisphere while simultaneously increasing in the left hemisphere (and vice versa) [5]–[8]. It is also shown in [9]–[11] that not only left-right but also other directions of covert attention are strongly correlated with the posterior alpha.

Covert attention was measured using EEG. EEG and fNIRS are the most suitable methods for healthy users at the moment, because no

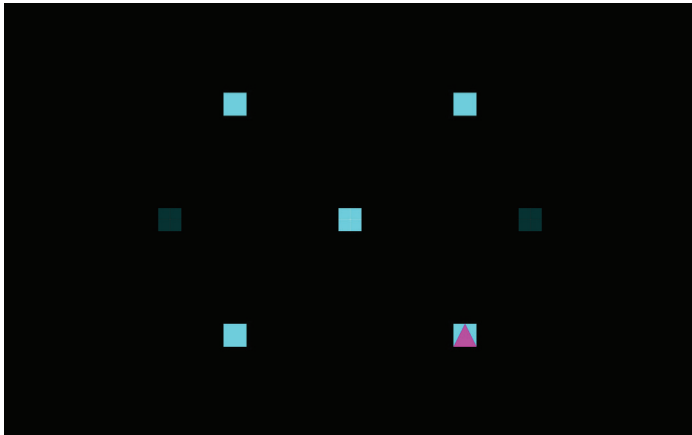


Fig. 1

COVERT ATTENTION SCREEN WITH FIXATION POINT IN THE CENTER, POTENTIAL TARGET SQUARES (DISTRACTORS), AND THE ARROW ON THE ACTUAL TARGET. THE DARKER SQUARES TO THE LEFT AND RIGHT WOULD NORMALLY NOT BE VISIBLE, BUT INDICATE THE ALTERNATIVE FIXATION POSITIONS.

surgery is required, the equipment can be used outside of a laboratory setup, and the equipment is relatively portable and affordable [1].

This section evaluates a number of potential pipelines, but also another important question: whether this correlation with posterior alpha depends on if a subject fixates centrally or if the same pattern will be observed irrespective of the location the subject's fixation point. While a central fixation point has been the norm in clinical laboratory experiments, in a practical application, this may only rarely be the case. Finally, some other research questions that are relevant for the online situation were looked into: what directions can we detect, how many trials are needed for training, and how long the trial window needs to be for classification?

A. Methods

The experiment is covert attention to the four directions of visual hemifields with three different fixation points. The task is to fixate at each fixation point in the screen which is 70 cm away from the eye of the subject and covertly attend to the direction of the pre-specified arrow. See Figure 1 for a screen shot of the situation. There are three fixation points: left, middle, and right, with six degrees of visual angle distance between them. The target focus can be one of 5 positions: either the fixation point itself (neutral), or one of the four diagonal directions. The focus targets were placed diagonally as earlier research indicated that this is best discriminable [11].

Fifty trials were recorded for each of these conditions consisting of a fixation position and target position. A trial starts with half a second showing the fixation cross, then for half a second the focus position for covert attention is indicated with a yellow circle inside one of the five potential positions. The other positions remain visible as distractors. After a period of 2 seconds plus a random duration of up to half a second, an up or down arrow is shown in the focus position. The participant then has a short period of time to press the corresponding arrow button (arrow up or down). This task ensures that the focus area is relevant to the participant, which may increase the effect on the brain activity for this paradigm. The trials were split up in five blocks, each containing ten repetitions for each condition in randomized order. The breaks in between blocks lasted until the participant pressed a key to continue.

Brain activity is measured during the task using the BioSemi

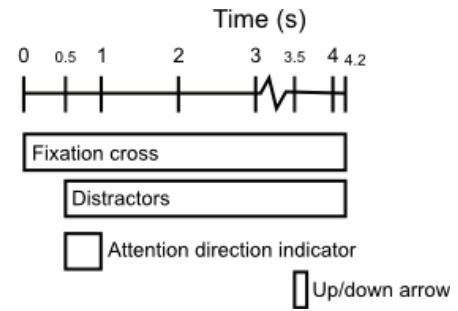


Fig. 2

DURING A TRIAL FIRST THE FIXATION CROSS IS SHOWN, THEN THE DIAGONAL POSITIONS APPEAR, AFTER WHICH THE FOCUS POSITION FOR COVERT ATTENTION IS INDICATED. AFTER A LITTLE WHILE AN UP OR DOWN ARROW IS SHOWN IN THE FOCUS POSITION. THE PARTICIPANT THEN PASSES THE CORRESPONDING BUTTON.

ActiveTwo EEG system, at 512 Hz sampling frequency, with 32 electrodes according to the montage shown in Figure 3. Electrooculogram (EOG) was also recorded to control for confounds in eye movements.

In total datasets were recorded for 8 participants, but for analysis the first two were left out because of marker issues. The last two sets were recorded at a late stage in the project, and thus were not used for every analysis that was conducted.

B. Results

a) Which pipeline performs best?: The four pipelines that were tested were:

Pipeline CA1:

- channels: occipito-parietal
- window: 0.5-2.0sec relative to focus indication stimulus
- feature extraction: CAR, bandpower 9-11Hz STFT, z-score normalization
- classifier: SVM (error cost: 0.1)

Pipeline CA2 (CA1 with whitening, and different SVM error cost parameter):

- channels: occipito-parietal
- window: 0.5-2.0sec relative to focus indication stimulus
- feature extraction: CAR, whitening, bandpower 9-11Hz STFT, z-score normalization
- classifier: SVM (error cost: 2.0)

Pipeline CA3:

- channels: occipito-parietal
- downsample to 256Hz
- window: 0.5-2.0sec relative to focus indication stimulus
- feature extraction: CAR, bandpass 8-14Hz, whitening, covariance
- classifier: logistic regression

Pipeline CA4 (CA1 with different SVM error cost parameter):

- channels: occipito-parietal
- window: 0.5-2.0sec relative to focus indication stimulus
- feature extraction: CAR, bandpower 9-11Hz STFT, z-score normalization
- classifier: SVM (error cost: 2.0)

Table I shows the performance accuracies per pipeline on average but also per subject. CA3 outperforms all others with 67% and 40% on average on the same datasets for two and four-class classification respectively. As pipeline CA3 was implemented in Matlab and not in Python it cannot be applied in the online situation. So for the game,

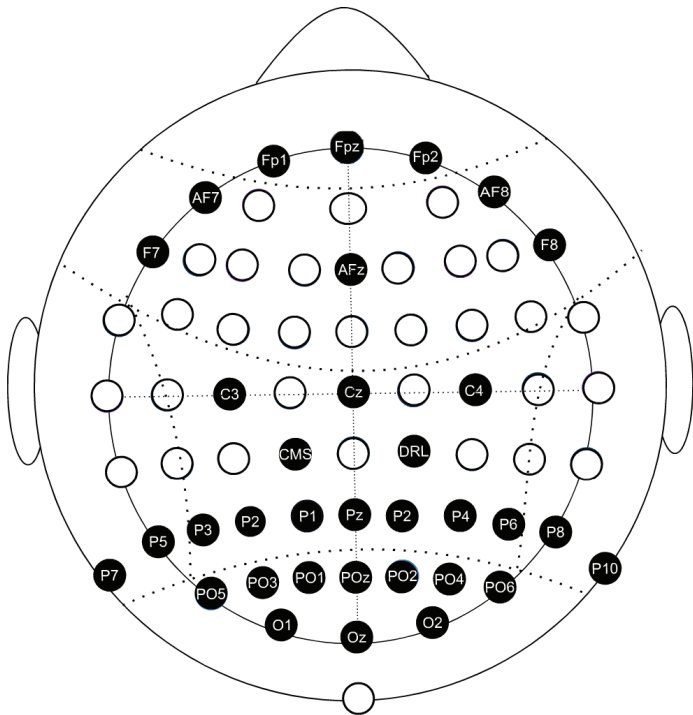


Fig. 3

ELECTRODE POSITIONING FOR EEG MEASUREMENT: 32 ELECTRODES POSITIONED MAINLY ON THE PARIETO-OCCIPITAL AREA AS THIS IS WHERE THE RELEVANT ALPHA MODULATIONS FOR SPATIAL COVERT ATTENTION ARE EXPECTED, AND SOME OTHERS TO LOOK AT ARTIFACTS AND TO OFFER THE POSSIBILITY TO APPLY CERTAIN SPATIAL FILTERS.

we will opt for the second-best pipeline in the two-class case, which is CA4.

b) *Does the position of the fixation point matter, with respect to the correlation of focus direction with parietal alpha, and with respect to detection accuracy?*: To answer this question, scalp plots were computed for each participant for each fixation position (left, middle, right), showing the relative difference in the alpha band (8–12 Hz) of each diagonal focus direction with the fixation point, see Figures 4-6. A time window from 0.5 to 2 seconds after the cue was used. The scalp plots were averaged over four subjects. The lateralization pattern is in line with what has been shown in literature [8], [9], [11]. As the eyes fixate on a different position, the excitation of the retina remains the same, and the mapping of the image to the occipital cortex is not expected to change. However, surprisingly, the patterns are a bit different for the different blocks, showing a migration of the alpha sources from one side to the other.

On average, there did not seem much of an accuracy difference between each of the fixation point conditions (28%, 30%, 32% and 30% for left, center, right, and pooled fixation points). When looking at our best participant however, we see an increase for the center fixation: 36% for left and right, 40% for pooled, but 45% for center fixation cross only.

c) *Which directions can be detected?*: Results based on datasets recorded from 4 different participants analyzed with pipeline CA3 indicate a performance above random. For a 4-class situation (each of the four directions) yields a 40% performance accuracy on average, and 52% on our best participant. The samples for the three different fixation points were pooled, so the classes indicate the covert attention direction relative to fixation. Random for four classes would have been 25%. For the two-class situation the bottom and top targets were

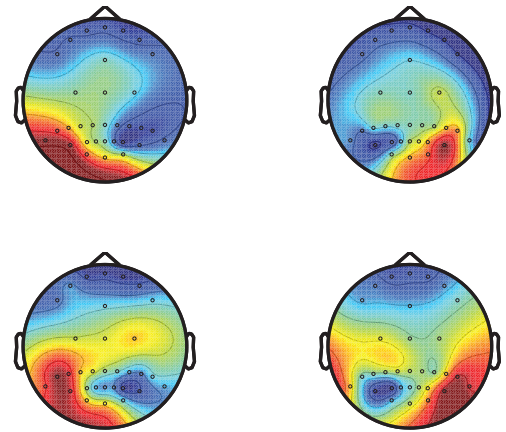


Fig. 4

RELATIVE DIFFERENCES OF EACH FOCUS DIRECTION WITH RESPECT TO THE FIXATION POSITION, WITH THE FIXATION ON THE LEFT.

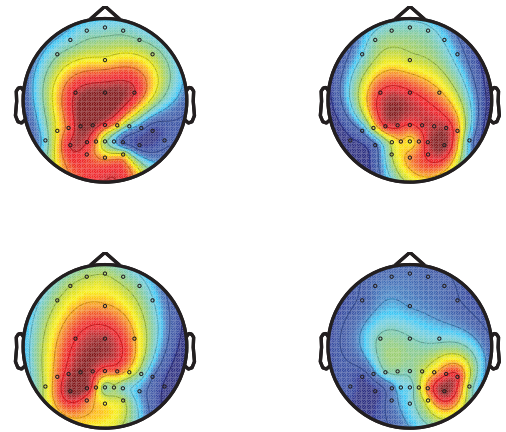


Fig. 5

RELATIVE DIFFERENCES OF EACH FOCUS DIRECTION WITH RESPECT TO THE FIXATION POSITION, WITH THE FIXATION ON THE CENTER.

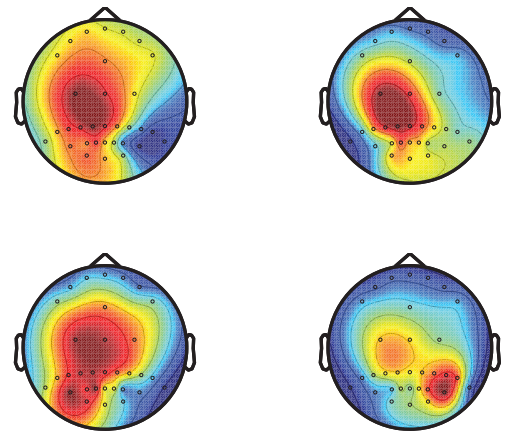


Fig. 6

RELATIVE DIFFERENCES OF EACH FOCUS DIRECTION WITH RESPECT TO THE FIXATION POSITION, WITH THE FIXATION ON THE RIGHT. ALL AVERAGED OVER FOUR SUBJECTS.

TABLE I

PERFORMANCE ACCURACIES OF THE COVERT ATTENTION PIPELINES PER PARTICIPANT. STANDARD DEVIATIONS OF THE PERFORMANCE SCORES ARE BETWEEN PARENTHESES.

4 classes	CA1	CA2	CA3	CA4
S3	32%	32%	33%	31%
S4	35%	30%	31%	31%
S5	44%	35%	52%	42%
S6	37%	34%	44%	35%
Avg	37% (4%)	33% (2%)	40% (2%)	35%
2 classes	CA1	CA2	CA3	CA4
S3	62%	60%	62%	60%
S4	61%	59%	57%	59%
S5	71%	71%	78%	85%
S6	66%	65%	72%	62%
Avg	65% (4%)	64% (5%)	67% (2%)	67% (11%)

merged to result in one class with samples to the left, and one class with samples to the right. For this, the average performance accuracy over 4 subjects with pipeline CA3 was 67%, with 78% for our best subject, against a random performance of 50% for two classes. The other pipelines show a similar pattern in performances, albeit slightly below the scores for CA3.

The classification performances for the different pairs of target directions (like top right vs. top left) were also analyzed. This confirms the information from literature that diagonally opposing targets (top left vs down right, and top right vs down left) are easier to distinguish than the other pairs.

d) How many trials are needed for training?: Figure 7 shows the two-class detection performance with pipeline CA3 for different training set sizes. The performance within the training set was evaluated using 10-fold cross validation. The plot shows no consistent increase in performance. After a peak at 120 trials, performance drops and flattens out.

e) What is the optimal window size?: For the online situation, preferably, the window size is minimal, because that way the data can be processed faster, which in turn could mean that updates can be computed more frequently. On the other hand, the classification accuracy is expected to be higher for longer window sizes (because you simply have more information).

Windows always start at 0.5 seconds after the stimulus, and then continue for the indicated window duration, except for the two-second window which starts at 0.0 seconds.

f) Does a blocked protocol yield a better performance?: In standard covert attention experiments there is only one fixation point, whereas in our experiment, this fixation point was randomized. To test whether this had unwanted side effects, we recorded one dataset which had the fixation points steady within each block, and one in which within a block this fixation point could jump around. The result was a 75% accuracy for both the blocked and not blocked condition of fixation points using pipeline CA3. Based on this one participant, there does not seem to be a difference between the two conditions.

C. Discussion and Conclusions

Pipeline CA3 (CAR, bandpass 8–14 Hz, whitening, covariance, logistic regression) performs best, but could not easily be translated from Matlab to Python. For the online situation we therefore decided to use the second best option: CA4, using CAR, bandpower 9–11 Hz STFT – adjusted to 8–14 Hz, z-score normalization, followed by an SVM classifier.

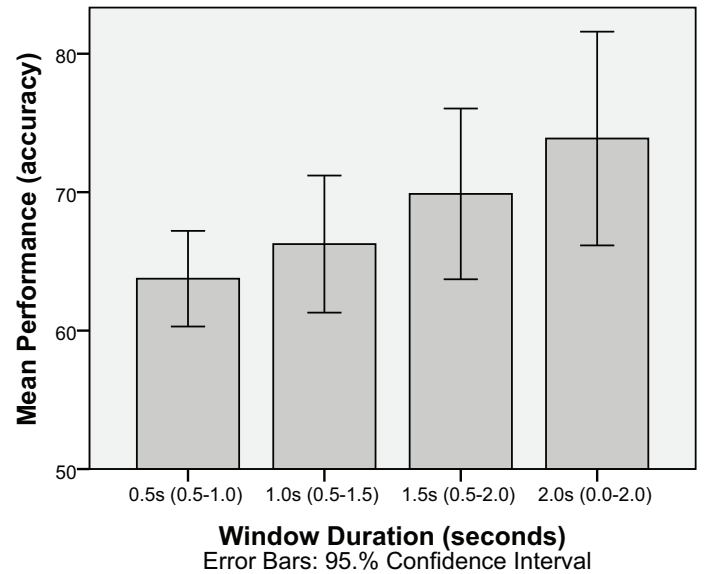


Fig. 8

TWO-CLASS COVERT ATTENTION PERFORMANCE FOR DIFFERENT WINDOW SIZES, AVERAGED OVER 6 SUBJECTS. IT SHOWS AN INCREMENTAL INCREASE FOR LONGER WINDOWS.

Different fixation points (left, middle, right) did not seem to have a significant impact on classification performance. When looking at the relative difference in parietal alpha between the focus direction and central fixation point, similar spatial patterns show which correspond to what is expected from literature. However, there also seems to be a migration of the alpha sources from one side to the other.

Although the four-class performance is above random, for an online game situation performance should be at a usable level. For this reason we decided to use two-class covert attention in the game.

The number of windows in the training dataset, strangely enough, does not seem to have a large impact on the classification performance. The performance does increase from 20 to trials samples, but after that it drops again, stabilizing around the same performance is shown at around 90 trials. As this is evaluated with 10-fold cross validation, about 80 trials would be enough if all trials are used.

The larger the trial window, the higher the performance. This is to be expected, but less fortunate for the online situation: the longer the window size, the longer it will take to get feedback on that particular window. However, we did not test beyond a size of two seconds, and the test for two seconds could not start at 0.5 seconds as the other windows did. This makes it possible that there are task-related eye movements in those 0.5 seconds that increase the performance.

However, most of these results are based on relatively little data.

III. EYE MOVEMENT

According to Jacob and Karn, using eye movement provides a number of features that make it an interesting input modality. Eye movements are not as intentional as mouse and keyboard input. This means that it can provide information on an intentional but also on a more subconscious level. A side effect is the Midas Touch problem: not every eye gaze has intentional meaning, so the system should somehow discern what to react to, and what not. Eye movement is faster than other input modalities, and already indicates the user's goal before any other action has been taken. Besides, no user training is required, as the relationship between the eye movement and the display is already established [2].

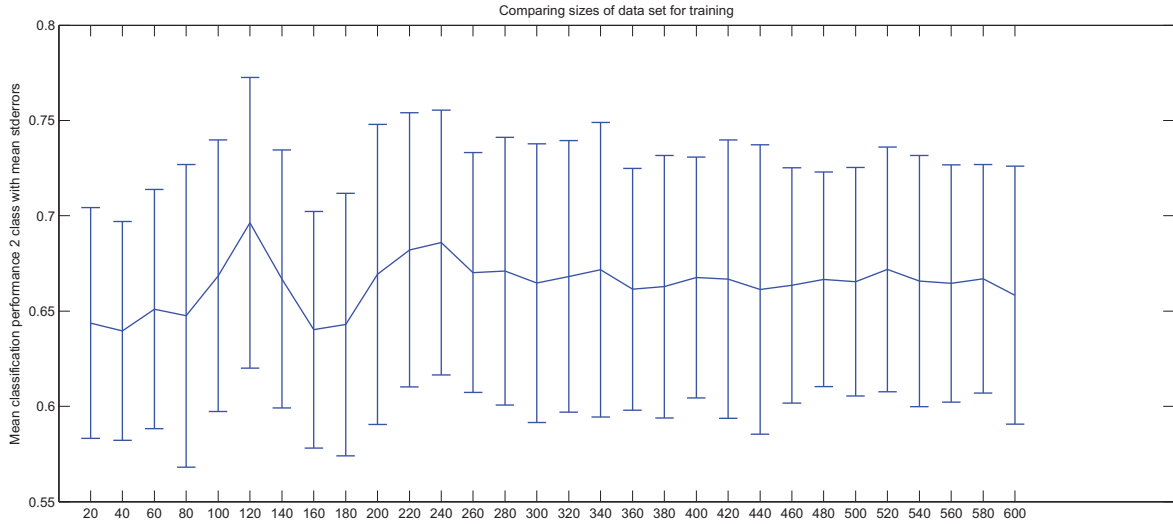


Fig. 7

TWO-CLASS COVERT ATTENTION PERFORMANCE FOR DIFFERENT TRAINING DATASET SIZES, FOR DATASETS 3-6.

Bulling et al. distinguish between the following types of eye movements. Fixations are the stationary states of the eyes during which gaze is focusing on a particular point on the screen, lasting between 100 ms and 200 ms. Saccades are very quick eye movements between two fixations points. The duration of a saccade depends on the angular distance the eyes travel during this movement. For a distance of 20 degrees, the duration is between 10 ms and 100 ms. Eye blinks cause a huge variation in the potential in the vertical electrodes around the eyes, and lasts between 100 ms and 400 ms [12]. For our application, saccades are the most relevant type of movement to detect.

There are a number of methods to determine eye movement or eye gaze, for example with special contact lenses, infrared light reflections measured with video cameras, or with electrodes around the eyes. The last example is also called electrooculography (EOG). The electrodes measure the resting potential that is generated by the positive cornea (front of the eye) and negative retina (back of the eye). When the eye rotates, the dipole rotates as well. By positioning the electrodes around the eyes as shown in Figure 9, one bipolar signal will be an indication of vertical eye rotation and the other for the horizontal axis.

For this system, we decided to use EOG for eye tracking. EOG signal analysis requires very little processing power, and can easily be done in real-time. Although this method is not that suitable for tracking slow eye movements (that occur when following a moving object), for fast saccades it is very robust. EOG can be used in bad lighting conditions (although it works better with good lighting), and in combination with glasses. The participant does not need to be restricted in the orientation to the screen (though for absolute eye gaze, then the position of the head would need to be tracked separately), nor do they have to wear an uncomfortable video camera system firmly mounted on the head [2]. Also, it is easy to incorporate in a wearable and unobtrusive setup [12].

This section explains the pipeline design, an eye blink detection and correction algorithm, the methods for the dataset recording and analysis, details the results of the evaluation, resulting in discussion and conclusions.

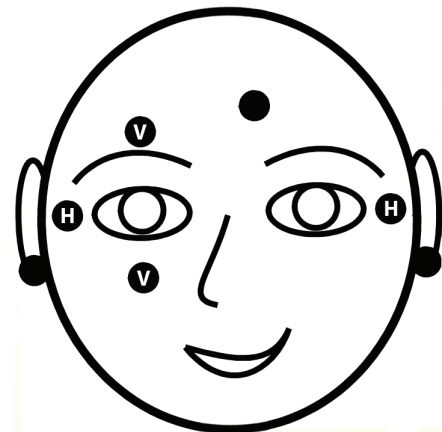


Fig. 9

ELECTRODE POSITIONING FOR EOG MEASUREMENT: BIPOLAR MEASUREMENTS OF TOP MINUS BOTTOM VERTICAL ELECTRODES AROUND THE RIGHT EYE AND RIGHT MINUS LEFT HORIZONTAL ELECTRODES NEAR THE CANTHI.

A. Pipeline

As described in [13], saccade detection can be used to construct an eye-tracker. The pipeline for eye movement is similar for both the vertical and horizontal EOG signals:

- 1) High pass filter (0.05 Hz) for drift correction which is very strong in the EOG signal.
- 2) Low pass filter (20 Hz) to reduce high frequency noise without affecting the eye movements.
- 3) Derivative in order to detect the rapid variations.
- 4) Thresholding to detect saccades and remove noise.
- 5) Integration in the saccade range which represents the features.
- 6) Linear regression between the angle and the integration result.
- 7) Conversion to x,y position.

The main steps are shown in Figures 10–13 and Figure 14.

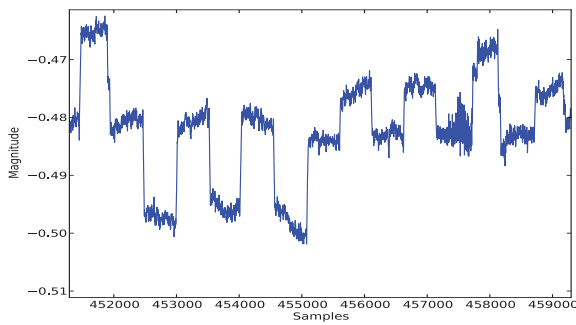


Fig. 10
EOG DATA IS NOISY AND DRIFTS OVER TIME.

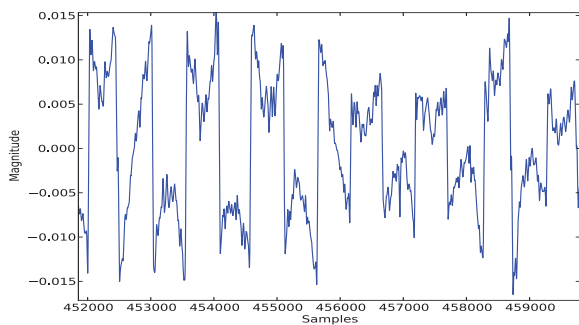


Fig. 11
FILTERED EOG DATA WITHOUT THE DRIFT AND HIGH FREQUENCY NOISE.

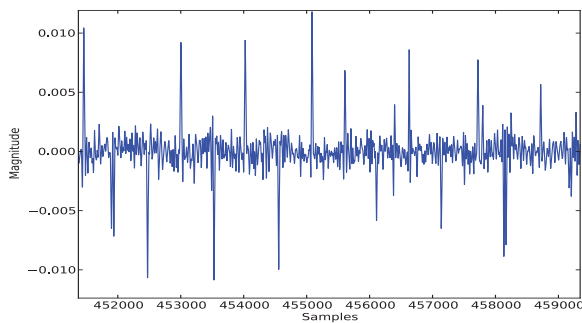


Fig. 12
THE HIGH VALUES OF THE DERIVATIVES INDICATE SACCADES.

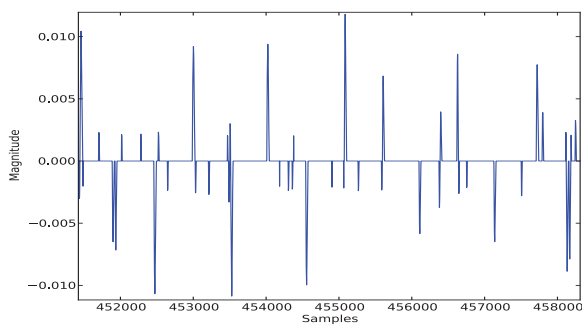


Fig. 13
INTEGRATION OF THE ABOVE-THRESHOLD SACCADE DERIVATIVE PROVIDES THE INPUT FOR THE LINEAR REGRESSION

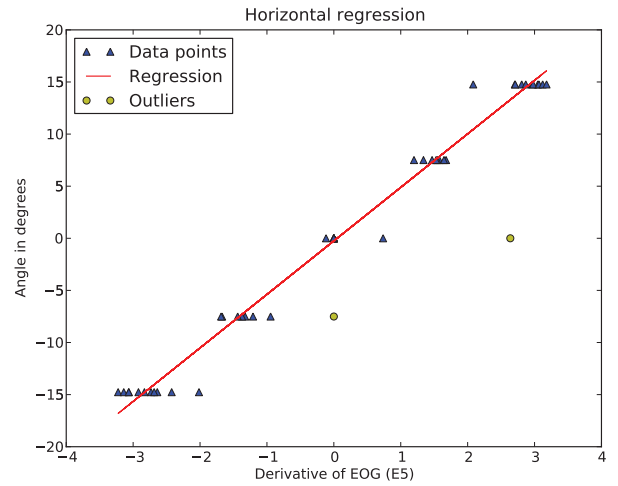


Fig. 14
THE REGRESSION SHOWS A HIGH CORRELATION BETWEEN THE PARAMETER OF EACH SACCADE AND THE JUMP IN ANGLE.

B. Eye Blink Detection

The previous pipeline can be enhanced by eye blink detection and correction. Eye blinks cause in large voltage changes in the vertical EOG signal, which result in a bad estimation of the current eye position 15.

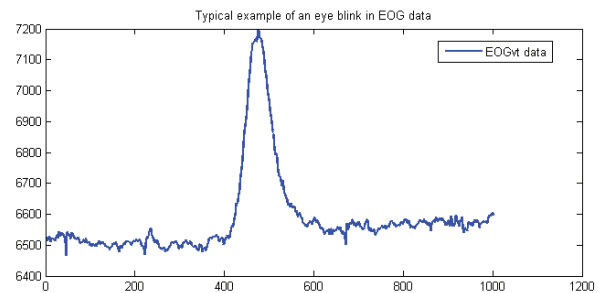


Fig. 15
TYPICAL OCCURRENCE OF AN EYE BLINK IN THE EOG SIGNAL, CHANNEL EOG VERTICAL TOP.

Inspired by [12], a template-based approach to eye blink detection was developed. There were EOG recordings with ten stimulus-based and hence unnatural eye blinks. Based on visual inspection, it was decided to use the EOG channel positioned right above the right eye for eye blink detection. The first five eye blinks were used to construct the initial template of 200 samples long (at a sample frequency of 512 Hz this is about 400 ms). The eye blink examples were aligned by taking a vertical offset such that the mean over time is zero.

This initial template was used to detect more natural eye blinks in the EOG data. Before determining the Euclidean distance between the template and the signal, the template was aligned using a vertical offset which minimizes the distance between the first and last ten samples of the template and the signal fragment under consideration. Figure 16 contains samples of EOG recordings from above the right eye of subject 6d and the corresponding Euclidean distance between the aligned template and the EOG data of the electrode above the right eye. Local minima in the distance below a threshold of 4000 mark the start of an eye blink.

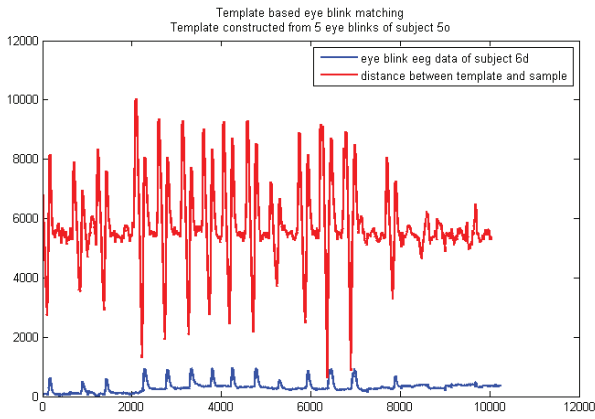


Fig. 16

TEMPLATE MATCHING ON EOG DATA OF SUBJECT 6D. LOCAL MINIMA IN THE DISTANCE BELOW 4000 INDICATE THE START OF AN EYE BLINK .

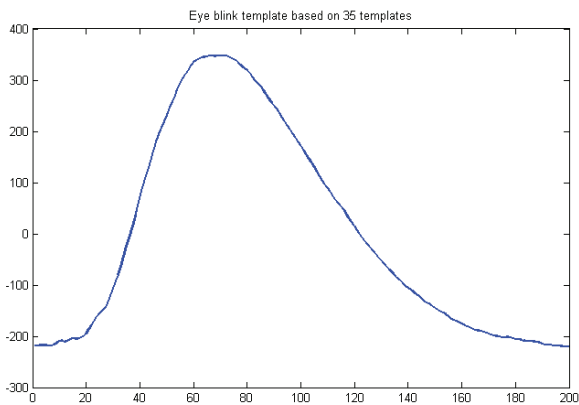


Fig. 17

THE FINAL EYE BLINK TEMPLATE CONSTRUCTED OUT OF 35 EXAMPLES.

These detected eye blinks were used to extend the eye blink examples to 35 and constructing a more realistic final eye blink template afterwards. The final eye blink template, see Figure 17, is used to construct an online eye blink detector.

For the online version, we only have a small sliding window to our disposal, which must be larger than the template size. Hence the template matching procedure has to be adapted, in particular the determination of a local minimum in the distance between the aligned template and the signal part. The start of an eye blink is now determined by a switch from decrease to increase (the first derivative changes sign) in the distance function and the constraint that the value of the local minimum is below the threshold. This threshold is dependent on the subject under consideration and can be determined by an online calibration.

Unfortunately, for this preliminary phase in the project, the eye blink detection and correction algorithm was not applied in the eye movement pipeline, because of time constraints.

C. Methods

The offline analysis protocol of the eye movement is twofold. In order to get enough data for training the linear regression, 25 trials were used. Each trial was composed of one target in the center of the

TABLE II

PERFORMANCE OF THE EYE MOVEMENT PIPELINE PER PARTICIPANT. HORIZONTAL AND VERTICAL ACCURACIES ARE FOR A PRECISION WITHIN 4CM. THE ERROR DISTANCE MEANS AND STANDARD DEVIATIONS ARE MEASURED FROM ACTUAL TARGET POSITION TO REGRESSION RESULT IN HORIZONTAL AND VERTICAL DIRECTIONS.

	Hacc	Herr avg	Herr std	Vacc	Verr avg	Verr std
S4	100.0%	1.0	0.8	94.9%	2.0	6.1
S5	90.9%	2.0	1.6	57.6%	3.9	3.7
S6	70.7%	3.2	3.2	34.3%	7.9	9.4
S8	77.8%	2.4	1.8	51.5%	5.3	4.5

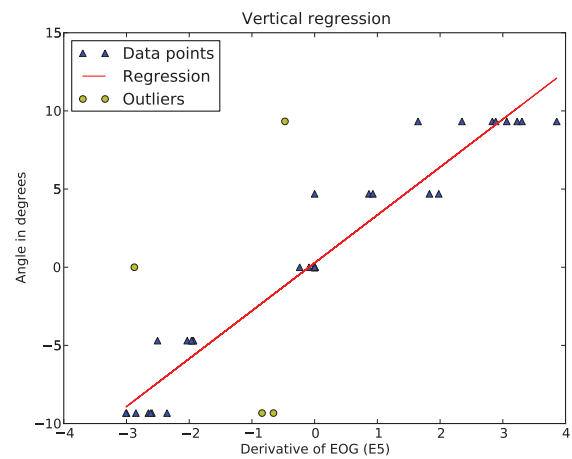


Fig. 18

DATA SHOWS LESS CORRELATION BETWEEN EOG FEATURES AND KNOWN ANGLE CHANGE FOR VERTICAL EYE MOVEMENT.

screen and one of five possibilities: extreme top, bottom, left, right and center targets. For horizontal and vertical eye movement there are separate pipelines, and the regression is also trained separately – for the pipeline details refer to the Pipeline section above.

For evaluation 100 trials were assessed. Because the system will be used as a kind of eye mouse, the performance evaluation was based on the accuracy of the system at N centimeters maximum deviation from the target. The screen was divided in a 5 by 5 grid, resulting in 25 potential target positions, which were selected randomly. The jump between the center and the target (Figure 19) of each trial is considered correct when the Euclidean distance between the EOG-based estimation point on the screen and the actual point is lower than N centimeters.

These trials were recorded using the BioSemi ActiveTwo hardware, with flat active electrodes positioned according to Figure 9. The distance between the user and the screen was 70 cm.

D. Results

Regarding the horizontal movements, the results are quite good as shown in Table II. Figure 21 shows the precision at N for two participants (S4 and S5). The curve is sharply increasing which shows the precision of this technique. However, regarding the vertical movements, the results are less good (see Table II, Figures 20, 22, and 18, the plots are again based on the data of the two participants S4 and S5).

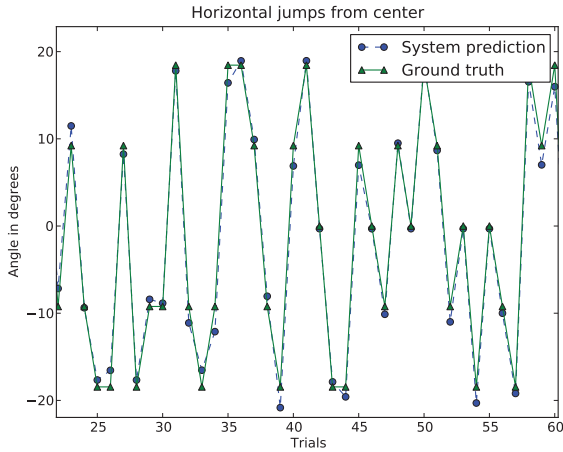


Fig. 19

THE JUMPS BETWEEN THE CENTER AND THE TARGET PROVIDED BY THE SYSTEM AND THE ACTUAL ONES ARE QUITE SIMILAR FOR THE HORIZONTAL AXIS.

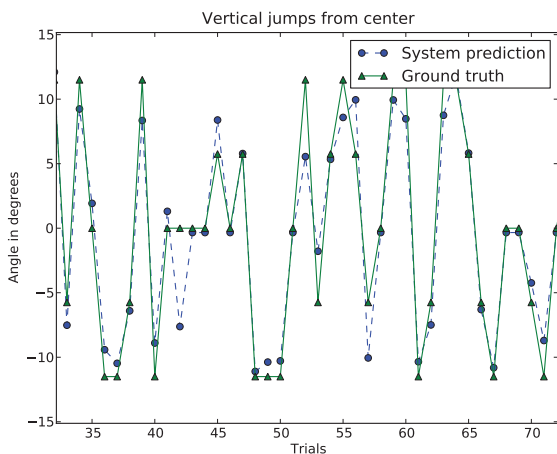


Fig. 20

THE PREDICTIONS OF THE SYSTEM ARE QUITE LESS GOOD THAN FOR THE HORIZONTAL JUMPS.

E. Discussion and Conclusions

Horizontal eye movement appears to be easily detectable: at a precision within 4 centimeters, the accuracy is about perfect for the best half of the participants. Within 2 centimeters it is about 90%. For vertical eye movement, the performance is a little less good: around 80% for a precision within 2 centimeters.

Visual inspection of the vertical EOG data shows that sometimes there is no sign of the vertical movement when there should be one. Maybe the sensors were not positioned optimally. Also, the vertical distance is smaller than the horizontal distance, meaning that the eyes will turn less degrees, resulting in a smaller potential chance. Moreover, the eye blink detection was not applied in the pipeline. This should also improve performance.

We still need to evaluate the optimal window length for eye movement detection. The window step should be quite short, to give the user the sense of continuous interaction.

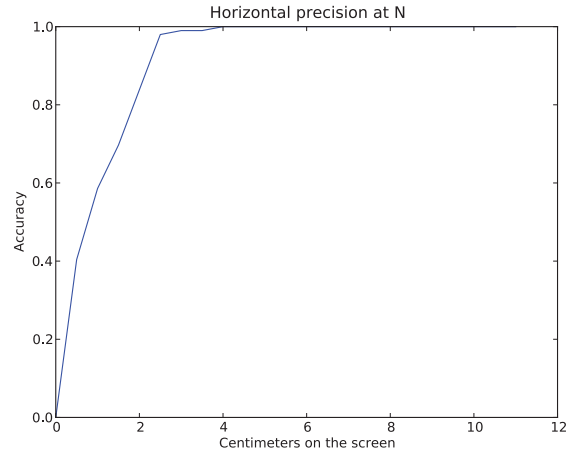


Fig. 21

HORIZONTAL PRECISION AT *N* CURVE IS SHARP AT THE BEGINNING WHICH IS GOOD.

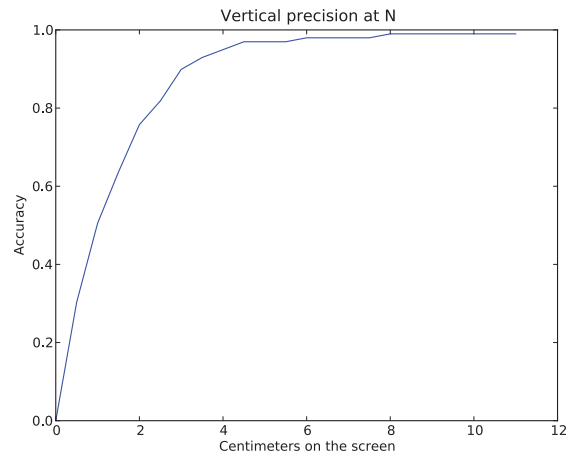


Fig. 22

VERTICAL PRECISION AT *N* CURVE IS LESS SHARP AT THE BEGINNING THAN THE CORRESPONDING HORIZONTAL CURVE.

IV. APPLICATION AND SYSTEM

The previous sections describe the development and evaluation of the pipelines for covert and overt attention. The goal is to use these covert and overt attention paradigms in a setting which requires real-time input. For this, the pipelines need to require minimum user and system training, the training needs to be integrated within the game, and the signal analysis and classification needs to be fast enough for such an application. As for the application itself: it had to be a game that requires the two paradigms to be used, in a way that is intuitive to the user, and reacts to them in a realistic manner. This way it also serves as proof that such signals can be a valuable addition to a game, and that they allow for new types of games. This section will first describe the game, and then the system bringing it all together.

A. Wild Photoshoot

In the game that was developed, you are a wildlife photographer. On an uninhabited island, you try to make pictures of rare wild animals. But of course wild animals are not that easy to make a good photograph of. First you have to follow animal tracks over the



Fig. 23

A SCREENSHOT OF WILD PHOTOSHOOT IN TRACKING MODE: TWO ANIMAL FOOTPRINTS ARE SHOWING IN THE CENTER-BOTTOM OF THE SCREEN.



Fig. 24

A SCREENSHOT OF WILD PHOTOSHOOT IN PHOTOSHOOT MODE: THE ANIMAL IS TO THE LEFT, COVERT ATTENTION (FOCUS SQUARE) IS STILL NEUTRAL, AND THE EYE GAZE (FIXATION CROSS) IS ALSO CENTERED.

island to find where the creature is hiding (see Figure 23). Then you go into photoshoot mode in which you try to get a good picture (see Figure 24). When you look directly at the animal, it will flee and you will have to track it again. Thus you have to covertly look at the animal to focus the camera to get your money shot.

Keyboard is used to walk and turn. The system detects eye movement, and adjusts the first person camera to reflect the view angle. When looking towards the left side of the screen, the camera turns towards the left as well, until the user is again looking at the center of the screen. It not only reacts to horizontal, but also to vertical eye movement. The first person camera angle can be adjusted manually by mouse. When in photoshoot mode, covert attention is needed to focus on the animal without looking at it directly. The

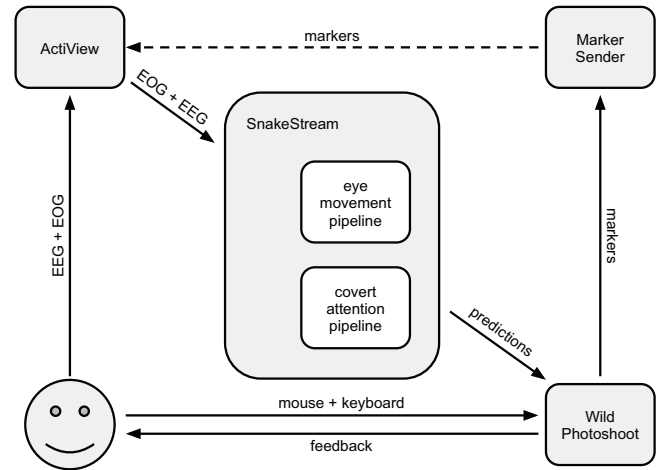


Fig. 25

THE DIFFERENT COMPONENTS AND COMMUNICATION BETWEEN THEM WITHIN THE WILD PHOTOSHOOT SYSTEM.

animal will appear to the left or right of the screen center. The classification of covert attention is divided into two classes: left and right. During the focus period, multiple covert attention classifications are performed. A simple majority vote determines the covert attention direction for the period. If the animal is on the same side as the covert attention, it will result in a nice wildlife photograph. If not, the user will have to try again. After five tries, the animal will notice and flee.

This is a game that uses multiple input modalities: mouse, keyboard, EOG-based overt attention, and EEG-based covert attention. It also creates situational disability for eye movement by letting the animal flee when looked at directly, introducing a natural need for covert attention. The mental tasks for both overt and covert attention come naturally given the situation, and the mapping to system response is based on real-world interaction as well. Finally through covert attention, we access information about the user that would not be available through other means.

B. System Design

Figure 25 shows how the different system components interact.

The user performs the user actions described in the previous section: looking by moving the eyes, and covertly attending without looking at the target directly. The user also interacts directly with the game through the keyboard to move around in the virtual environment.

EEG is measured in order to detect covert attention. EOG is measured to detect eye movement. This is done using Biosemi ActiveTwo hardware with 32 active electrodes, 7 additional flat-type electrodes on the skin, and separate CMS plus DRL. The raw data is sent over USB to the computer, where the Biosemi ActiView software sends the data over TCP/IP to the signal analysis software.

SnakeStream handles reading data from ActiView, passing the data in the appropriate formats to the signal analysis pipelines, and sending the prediction results from the pipelines on to the game environment of Wild Photoshoot. Snakestream works together well with the Golem and Psychic Python libraries, and supports the use of different markers and different sliding windows for each pipeline.

Within the game, keyboard input is used to move around the virtual world, eye movement to adjust the camera angle, and covert attention to shoot a great picture of the animal. The game can send markers to the EEG stream to give commands to the signal analysis software,

and to annotate the data for later offline analysis of the recordings. Because of limitations of the game engine software, it has to use the marker server to do this.

The marker server is a small application that receives marker values over TCP/IP and forwards them to the parallel port so it is added to the EEG stream. It also implements a simple queuing mechanism to ensure that markers do not get overwritten.

V. ONLINE EVALUATION

Due to time constraints, the online evaluation has not yet been executed. However, the preparations are ready. The online evaluation will look into two aspects. One is the influence of the online immersive situation on the signals measured and the classification performance for covert attention. The second is to evaluate the system usability and user experience when adding the eye movement camera adjustment to the interaction.

The goal is to run the main experiment for 10 participants. Each session will start with a recording of a clinical session, so we know for each participant what the theoretical performance is. This can then later be compared to the performance on the in-game training sessions. Next there will be two game sessions: one with eye movement and covert attention, and one with just covert attention where the camera is only adjusted by mouse. This way we can determine the effect of the eye movement based camera adjustment. After each game session there will be questionnaires to evaluate usability and user experience. The experimental design will be of a random crossover type, in which participants are randomly assigned to either the group where first only covert attention is used and secondly also eye movement or the group for which the order is inverted. This way possible learning or conditioning effects will be averaged out.

We will use two different questionnaires, administered to the participants in such a way that they only have to fill in one page with questions. The two different questionnaires are the SUS [14] and an adapted version of the presence questionnaire by Witmer et al. [15]. The SUS provides us with a standardized well validated scale that has been tested extensively in the field of HCI. For a more in-depth knowledge of the presence, immersion and control the user experiences within our game, we took the most interesting scales from the presence questionnaire and added some items particularly of interest to BCI research. Whereas normal input devices such as the mouse and keyboard provide the system with reliable input, using a BCI will not provide the user with perfect transmission of their intention to the system. This has its reflection on the user and we want to measure to what extent it alters the user experience.

VI. DISCUSSION AND CONCLUSIONS

The goal of this project was to develop a prototype that uses naturally occurring neurophysiological activity for natural user tasks, applying them in a way that supports intuitive interaction, with natural system responses. Pipelines for overt and covert attention have been developed and evaluated. A game that uses them in an intuitive manner has been designed and implemented, as well as a platform that provides the communication glue between each of these components.

Covert attention into four directions is detectable, but not well enough to be used as such in a game. The current game therefore only uses two classes: left and right. Detection accuracy did not decrease significantly for different fixation points. Around 80 trials will be enough for a training set for two classes. Larger trial windows result in higher performances, but this has not been tested beyond 1.5 seconds.

Horizontal eye movement seems to be detectable quite well. Vertical eye movement seems a little bit more problematic: sometimes

it does not show even though it is expected. This could be an inherent problem as the vertical distance between targets is smaller than the horizontal distance. Applying eye blink detection and correction could also improve performance. Optimal window length and training protocol still need to be determined.

However, most of these results are based on relatively little data. Analysis should be redone at a later stage, including all datasets. Possibly additional sets need to be recorded.

Other future work consists of performing the online experiments, perhaps implementing the CA3 pipeline that worked better than the others, using eyes closed data to determine a personalized alpha band, and evaluating eye movement detection based on EEG without the need for separate EOG electrodes.

REFERENCES

- [1] B. Allison, B. Graimann, and A. Gräser, "Why Use A BCI If You Are Healthy?" in *BRAINPLAY 07 Brain-Computer Interfaces and Games Workshop at ACE (Advances in Computer Entertainment) 2007*, 2007, p. 7.
- [2] R. Jacob and K. Karn, *Eye tracking in human-computer interaction and usability research: Ready to deliver the promises (Section Commentary)*. Elsevier Science, 2003.
- [3] M. Posner, "Orienting of attention," *The Quarterly Journal of Experimental Psychology*, vol. 32, no. 1, pp. 3–25, 1980.
- [4] R. Wright and L. Ward, *Orienting of attention*. Oxford University Press, USA, 2008.
- [5] M. I. Posner and S. E. Petersen, "The attention system of the human brain," vol. 13, pp. 25–42, 1990.
- [6] M. Worden, J. Foxe, N. Wang, and G. Simpson, "Anticipatory biasing of visuospatial attention indexed by retinotopically specific alpha-band electroencephalography increases over occipital cortex," *Journal of Neuroscience*, vol. 20, no. 6, p. 63, 2000.
- [7] S. P. Kelly, E. Lalor, R. B. Reilly, and J. J. Foxe, "Independent brain computer interface control using visual spatial attention-dependent modulations of parieto-occipital alpha," 2005, pp. 667–670.
- [8] M. A. J. van Gerven, A. Bahramisharif, T. Heskes, and O. Jensen, "Selecting features for BCI control based on a covert spatial attention paradigm," vol. doi:10.1016/j.neunet.2009.06.004, 2009.
- [9] T. A. Rihs, C. M. Michel, and G. Thut, "Mechanisms of selective inhibition in visual spatial attention are indexed by α -band EEG synchronization," vol. 25, no. 2, pp. 603–610, 2007.
- [10] M. A. J. van Gerven and O. Jensen, "Attention modulations of posterior alpha as a control signal for two-dimensional brain-computer interfaces," vol. 179, no. 1, pp. 78–84, 2009.
- [11] A. Bahramisharif, M. A. J. van Gerven, T. Heskes, and O. Jensen, "Covert attention allows for continuous control of brain-computer interfaces," vol. 31, no. 8, pp. 1501–1508, 2010.
- [12] A. Bulling, D. Roggen, and G. Tröster, "Wearable eog goggles: eye-based interaction in everyday environments," in *CHI '09: Proceedings of the 27th international conference extended abstracts on Human factors in computing systems*. New York, NY, USA: ACM, 2009, pp. 3259–3264.
- [13] R. Barea, L. Boquete, M. Mazo, and E. López, "Wheelchair guidance strategies using eog," *J. Intell. Robotics Syst.*, vol. 34, no. 3, pp. 279–299, 2002.
- [14] J. Brooke, *SUS: a "quick and dirty" usability scale*. London: Taylor and Francis, 1996.
- [15] B. G. Witmer and M. J. Singer, "Measuring presence in virtual environments: A presence questionnaire," *Presence: Teleoper. Virtual Environ.*, vol. 7, no. 3, pp. 225–240, 1998.



Danny Plass-Oude Bos did her internship at the University of Nijmegen, implementing physiological artifact detection in an online EEG-based BCI system. In 2008 she obtained her master in Human-Computer Interaction on BrainBasher, looking into the user experience of using BCI for games. At the moment she is working as a PhD student at the University of Twente, still attempting to merge BCI with HCI by researching how BCI can be made a more intuitive means of interaction.



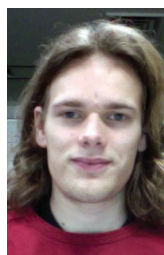
Dr. Ayhan Istanbulu received his undergraduate degree (1993) and his PhD from the Electronic and Computer Science Education Department of the Gazi University (2003), Turkey. He worked as an instructor at the University of Mugla, Turkey in the Department of Electronic and Computer Science Education (2001-2006). He currently works as assistant professor in the Computer Engineering Department of Balikesir University, Turkey. He has participated in the European Remote Radio Laboratory project (EU LdV). His current research interests include the

investigation of information technologies to support electronic and computer engineering education, mobile learning and intelligent tutoring.



Matthieu Duvinage, TIME student, holds Electrical Engineering degrees from the Faculty Polytechnique of Mons (UMons, Belgium, 2009), and Electrical Engineering degree from SUPELEC (France, 2009), a degree of fundamental and applied physics from Paris Sud XI Orsay (France, 2009) and a degree of management science from the School of Management at the University of Louvain (UCLouvain, 2011). His master thesis at Multitel (Mons, Belgium) dealt with robust low complexity speech recognition. He obtained an F.R.S-FNRS grant for pursuing a

PhD thesis about developing a lower limb prosthesis driven by a neural command in close partnership with the Free University of Brussels (ULB).



Marijn van Vliet has recently graduated from Twente University in the field of Human Media Interaction, where he devoted most of his time studying Brain-Computer Interfaces. He did an internship in Tokyo on virtual worlds and is recently accepted for a PhD position at the University of Leuven, where he will continue his BCI work.



Oytun Oktay obtained his bachelors degree in electronics engineering from Gebze Institute of Technology in 2008 with a insight of numerical electromagnetics. He started computational science and engineering master's program in Istanbul Technical University but left it after two years for a research assistant position at Namik Kemal University. Now he is working as a research assistant in Electronics and Communication Engineering Department on the subject of functional brain networks.



Bram van de Laar obtained his Bachelor degree in Computer Science (2006) and Master degree in Human Media Interaction (2009). With a broad interest in technology, such as: 3D, video, networking, music, sounds, haptics, brain-computer interfacing and physical exertion, Bram tries to combine different media to create a synergy by exploiting different modalities. User experience and 'added value' play an important role in this philosophy. As a PhD candidate Bram gets the space to explore the possibilities in this area.



Jaime Delgado Saa obtained his Bachelor and Master in Electronics Engineering in 2003 and 2008 respectively, at the Universidad del Norte (Colombia). From 2004 to 2009 he worked at the Universidad del Norte at the department of Electrical and Electronics Engineering. Currently he is a PhD student at Sabanci University at the Vision and Pattern Analysis group, working in the field of Brain Computer Interfaces. His work is focuses in the analysis temporal and spatial of EEG signals.



Mannes Poel is a senior researcher at the Human Media Interaction group of the department of Computer Science, University of Twente. He has a background in mathematics and theoretical computer science. After working on verification of concurrent processes his interest shifted towards Neural Networks, Machine Learning and Artificial Intelligence. At the moment his research focuses on Machine Learning in the context of Human Behavior Computing, including BCI. Currently he supervises several PhD students in these areas. He (co-)authored several

papers on Machine Learning in HCI and BCI, some of them also focusing on affective dialogue management.



Huseyin Güürler works as a lecturer at Mugla University in the department of Electronics and Computer Science. He has a bachelor degree in the field of Electronics and Computer from Marmara University, Istanbul, and master degree in the field of Statistics and Computer, Mugla University, Mugla, Turkey. His MSc thesis is about data mining and knowledge discovery in student databases. Now, he is a PhD student, researching RLS disease diagnosis from physiological signals.



Linsey Roijendijk obtained her Bachelor degree in Computer Science (2008) and her Master degree in Artificial Intelligence specialized in cognitive research (2009). Since November 2009, she is working as a PhD student at the Biophysics department of the Donders Institute for Brain, Cognition and Behaviour, Nijmegen, the Netherlands. Currently, her work focuses on investigating variabilities in EEG based Brain Computer Interfaces (BCI) and using covert spatial attention for BCI.



Luca Tonin obtained his Master degree in Electronic Engineering (2008) at University of Padova. Integration between Brain-Computer Interface and robotic devices was the main topic of his master thesis. Since January 2010, he is working as PhD student at EPFL in the CNBI (Chair on Brain-Machine Interface) laboratory. He is actively involved in TOBI (Tools for Brain-Computer Interface) an European Project started in November 2008. Currently, his research and PhD thesis are focused on covert attention for BCI purposes.



Ali Bahramisharif got his bachelor's and master's degree in the field of Electrical and Electronics Engineering from Sharif University of Technology and Tarbiat Modarres University, Tehran, Iran, respectively. Since 2008, he is working as a PhD student at the Intelligent Systems group of Radboud University Nijmegen, and as of February 2009 he has appointed as a part-time researcher at the Donders Institute for Brain, Cognition and Behaviour, Nijmegen, the Netherlands. Research interest includes brain-computer interfaces (BCI), machine learning

methods and source localization. His recent research focuses on different aspects of covert spatial attention.



Boris Reuderink obtained his Masters degree at the University of Twente in 2007, after working on different machine learning problems. He combines his big interests brains and intelligence with his PhD position at the University of Twente, in which he focuses on making brain-computer interfaces work in real-world settings for healthy users.