Spatio-Temporal Event Selection in Basic Surveillance Tasks using Eye Tracking and EEG

Jutta Hild[1], Felix Putze[2], David Kaufman[2], Christian Kühnle[1], Tanja Schultz[2], Jürgen Beyerer[1,3]

[1]Fraunhofer Institute for Optronics, System Technologies and Image Exploitation (IOSB)
76131 Karlsruhe, Germany
{jutta.hild, juergen.beyerer}@iosb.fraunhofer.de

[2]Cognitive Systems Lab
Karlsruhe Institute of Technology
76131 Karlsruhe, Germany
{felix.putze, tanja.schultz}@kit.edu

[3]Vision and Fusion Laboratory
Karlsruhe Institute of Technology
76131 Karlsruhe, Germany

ABSTRACT

In safety- and security-critical applications like video surveillance it is crucial that human operators detect task-relevant events in the continuous video streams and select them for report or dissemination to other authorities. Usually, the selection operation is performed using a manual input device like a mouse or a joystick. Due to the visually rich and dynamic input, the required high attention, the long working time, and the challenging manual selection of moving objects, it occurs that relevant events are missed. To alleviate this problem we propose adding another event selection process, using eye-brain input. Our approach is based on eye tracking and EEG, providing spatio-temporal event selection without any manual intervention. We report ongoing research, building on prior work where we showed the general feasibility of the approach. In this contribution, we extend our work testing the feasibility of the approach using more advanced and less artificial experimental paradigms simulating frequently occurring, basic types of real surveillance tasks. The paradigms are much closer to a real surveillance task in terms of the used visual stimuli, the more subtle cues for event indication, and the required viewing behavior. As a methodology we perform an experiment (N=10) with non-experts. The results confirm the feasibility of the approach for event selection in the advanced tasks. We achieve spatio-temporal event selection accuracy scores of up to 77% and 60% for different stages of event indication.

Categories and Subject Descriptors

H.5.2. [**Information interfaces and presentation (e.g., HCI)]:** User Interfaces – *Input devices and strategies*

General Terms

Experimentation, Human Factors

Keywords

video surveillance task, eye tracking, EEG, event selection, event localization, experiment

GazeIn'14, November 16, 2014, Istanbul, Turkey.

http://dx.doi.org/10.1145/2666642.2666645

1. INTRODUCTION

In applications like video surveillance, (air) traffic control systems, but also computer games, or simulations, it occurs that a human operator has to monitor one or more screens presenting visually rich dynamic output. One important task for the operator is to scan the output for certain relevant events which might occur in the form of objects showing certain characteristics or activities, or creating certain situations of interest. In many cases, the operator has to report or disseminate the detection of an event.

Particularly, in safety- or security-critical real-time applications fast and reliable event dissemination is essential to make an appropriate response possible. For this purpose, the human operator is typically assisted by an interactive system tailored closely to the specific application. For example, current video surveillance systems started to provide automated algorithms for some typical subtasks like event detection or object tracking [2, 3]. However, in safety- and security-critical applications, the human-in-the-loop will always be an imperative part, and therefore it is usually the human operator who performs the initial dissemination step by applying a selection operation to the event on screen. The system reaction might be sending an alert to the staff responsible for performing the response, or the event might just be marked in the video image for further analysis by other operators [2, 3, 4].

Event selection requires the localization of the event both in time and space. State-of-the-art systems typically provide manual input techniques using a mouse or a joystick to perform the selection operation. Using such devices, event localization in space is determined by the mouse pointer position, and event localization in time is determined by the point in time the mouse or joystick button is clicked. Manual, pointer-based selection is an efficient technique for the selection of stationary objects on computer screens. However, in the case of the visually rich and dynamic input, it occurs that the human operator misses to manually select events. For example, this might happen just because the object carrying the event moved out of the camera image before the human operator could mark it. Making selection operations faster, easier and more comfortable in order to reduce the load of the human operator is therefore an important issue.

Recently, gaze-based interaction, which has been investigated and optimized over several years for the selection of stationary objects on a desktop [8], appeared to prove advantages over mouse input for moving object selection in video images [4]. As the eyes of a human operator typically lie on what he or she is working on [9], eye tracking appears a natural way for the spatial localization of

an event in a video image. The results of Hild et al. (2013) [4, 5] show that a gaze interaction technique combining gaze pointing for spatial and a key press on a keyboard for temporal event localization ("gaze+key") provides faster and equally accurate moving object selection than mouse input; moreover, the great majority of participants preferred the gaze-based interaction technique. As the reduction of physical effort by replacing mouse input by gaze+key proved to be successful, we propose to continue this process by developing a selection paradigm which does not involve any physical interaction.

Besides selection speed, ease and comfort, in safety-critical environments it is crucial that the selection procedure encompasses a way to prevent or at least minimize selection misses which may result from physical overload (caused by multiple simultaneously occurring events, physical or auditory distraction, or fatigue) or because of hesitation in unclear situations. In such cases, the operator may scan the screen as usual. But when detecting an event, the operator may not be able to perform the manual or even the gaze+key selection within the period of time the event (or at least the corresponding object) is visible on the screen, as also the gaze-based selection still requires an actively performed key press.

To cope with this challenge, this paper investigates a novel input method eluding any manual action or even any voluntary intervention for event selection. Our goal with this method is not to completely replace manual selection, but to augment it to cover situations in which manual event indication is not possible. The proposed input method uses eye tracking for the spatial localization of an event, and EEG measurement for temporal localization. In doing so, spatial localization is performed using gaze pointing like used by the above mentioned gaze+key technique. EEG is a striking way for temporal event location. With EEG, we can measure characteristic brain activity related to perception and attention of task-relevant events. This characteristic activity can be measured even when the operator has not (yet) generated any voluntary reaction to the event. We can therefore try to detect those brain activity patterns from an EEG stream recorded from the operator to localize events on the screen in time. EEG-based Brain-Computer Interfaces (BCIs) have a long history of detecting events which catch the user's attention; a classical BCI paradigm which tracks the user's attention is the P300 speller paradigm for letter spelling without manual intervention [10]. If such an eye-brain interface would be able to detect events, a system could, for example, send the detected event to an external video analyst or decision maker notifying that an event might have occurred which the human operator did not handle actively.

The combination of eye tracking and EEG has been investigated before by other authors. Yong et al. (2011) describe a hybrid point-and-click interface allowing self-paced operation of a BCI speller for text entry at arbitrary locations determined from eye gaze [17]. Zander et al. (2010) implement a visual-neural item selection process without manual intervention by combining eye tracking and EEG [18]. They use eye tracking to localize the user's attention in the spatial domain; the EEG signal replaces the traditional dwell time-based object selection. For evaluation of their approach, they use a search-and-select task on a simple graphical user interface. Other than our proposed approach, their paradigm requires the user to actively issue a mental command to trigger selection, which according to the authors should result in higher accuracy compared to implicit item selection. Their results support this finding. However, they also notice significantly higher reaction times for the active BCI compared to implicit methods. Huang and Lo (2013) integrate EEG information to enhance the controllability of a gaze-based cursor in a continuous cursor control task [6] reporting a significantly better precision of the control compared to pure gaze-control. Putze et al. (2013) [11] showed that the use of eye tracking and EEG allows spatio-temporal event localization: A number of circles were used as stakeholders for objects to which an event might occur. An event was represented by a red dot marking the circle for 2s. Spatio-temporal localization was possible for 78.5% of all cases if the circles were stationary, and for 64.5% of all cases if the circles were moving. While they showed the general feasibility of spatio-temporal event location with eye tracking and EEG, the achieved results require further validation to show their applicability to more realistic scenarios.

In this paper, we provide this validation with results from an updated experimental paradigm. With the newly designed experiments, we show the feasibility of the approach for less artificial experimental paradigms which are much closer to a real surveillance task in terms of the required viewing behavior. We do this by 1) reducing the size of the displayed objects, 2) replacing color-based indication of events by much more subtle cues, 3) adding visual distractors, and 4) avoiding unnatural viewing behavior.

The experimental task for the validation of the proposed event selection approach needs to provide a minimum number of events which are balanced over time. This requirement is close to impossible to fulfill with a continuous stream of real video images. Therefore, we use simulated video streams instead. While the new experiments are still highly abstracted from real world video analysis, they now correspond to typical real-world surveillance tasks, like monitoring a stationary or moving object for suspicious behavior. If we can reproduce the results from Putze et al. (2013) [11] on this much more challenging data set, we show that spatio-temporal event localization can be performed robustly for those use cases. Once this result is confirmed, we will be able to compare the proposed approach to mouse-based event selection or to the gaze+key technique in future work.

2. METHODOLOGY

Two experimental tasks were implemented. For each, a framework simulates a frequently occurring basic surveillance task demanding the observation of a single object: the observation of a stationary object, e.g., a house, and the observation of a moving object, e.g., a car, which are both tasks human operators are confronted with in real video surveillance.

2.1 Experimental Tasks

Figure 1 and Figure 2 show screenshots of the visual stimuli. The presentation of the visual stimuli is designed close to the visual parameters of overflight video images in terms of choice of colors, object size, and the presence of distractor objects. For authentic overflight images, have a look at Heinze et al. (2010) who provide several types of images. As the colors of video images are also often not very bright, the design uses rather pale colors and low contrast between background color and object colors.

The object of interest is represented by a light gray circle with a diameter of 0.72 degrees of visual angle (30 pixels on the screen in the used setting). In experimental task 1, the object of interest

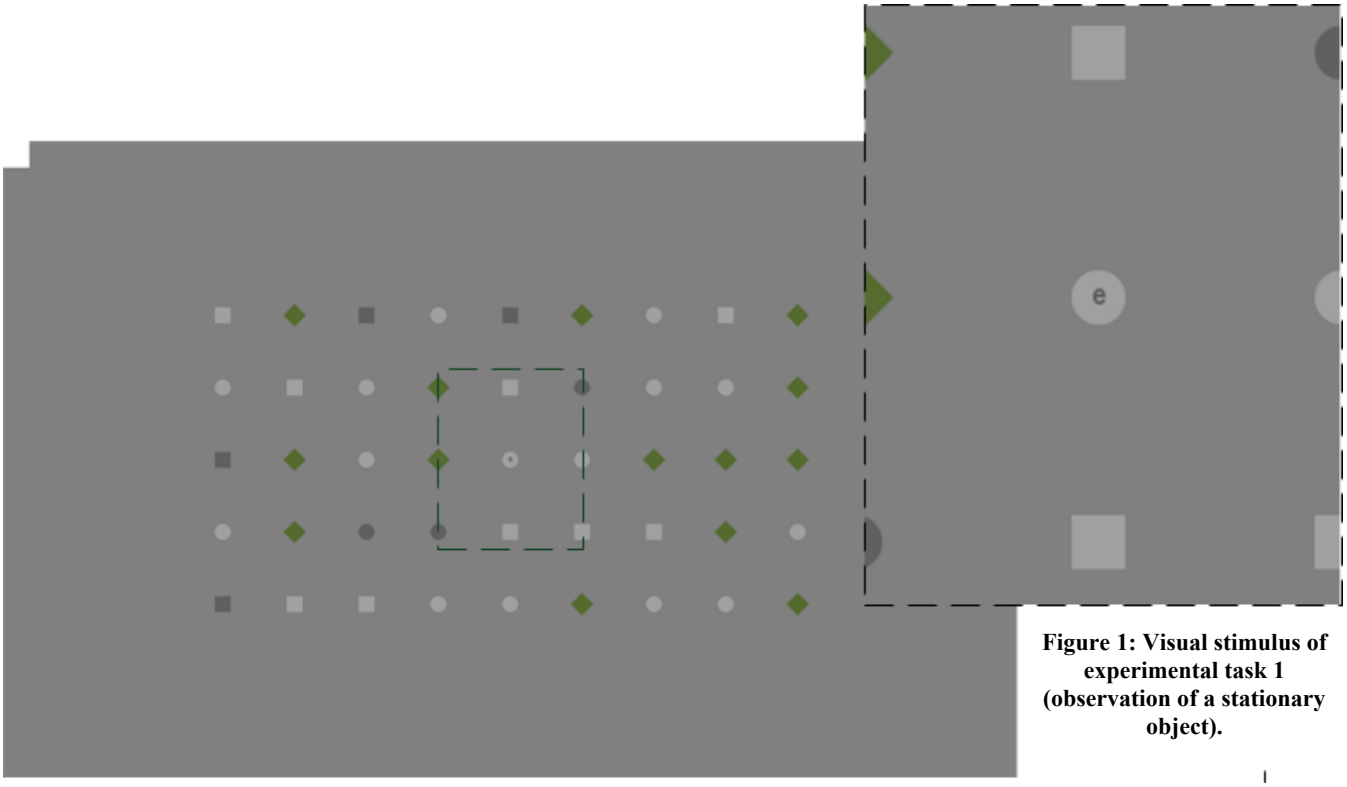

Figure 1: Visual stimulus of experimental task 1 (observation of a stationary object).

is located stationary in the center of a 9x5 grid of distractor objects, all having the same size. Distances between centers of distractor objects correspond to 3.16° of visual angle; colors and shapes vary, some look like the object of interest, but most differ in color and/or shape. The visual stimulus of experimental task 2 is similar to experimental task 1 but uses a different distractor object arrangement; the object of interest now moves straight-line through the grid (covering a range of 30.9°x16.2°) with a speed of 70 pixel/s or 100 pixel/s, speed randomly changing if the object of interest changes its moving direction (possibly at every starting point of the yellow arrows).

Another important issue is the design of the event indication. In fact, we use two event indication stages. The first stage is the repeated display of different letters on the object of interest; letters show up randomly for 2s, every two letter displays interrupted by at least 2s of no letter display. This first event indication stage corresponds to the event indication used by Putze et al. (2013) [11] using a red dot displayed for 2s on the objects. The letter display uses Arial bold, dark gray, and font size 16 to ensure good legibility, but yet implements a much more subtle cue as the clearly popping-out red dot. The second event indication stage implements an oddball paradigm. We distinguish between the target letter ‘e’, and distractor letters ‘h’, ‘p’, ‘n’ (all easily distinguishable from target shape ‘e’), and ‘c’ (close to target shape ‘e’ to ensure that the subjects look closely if a letter is displayed). Hence, the human observer has to use cognition to distinguish whether a shown letter is the target letter ‘e’ or not. We expect this to be a far more challenging temporal location compared to the color change (red dot) used in [11] which could be perceived by humans pre-attentively, not requiring any cognitive work.

Linking the two stages to aspects of the corresponding real surveillance tasks can be done as follows. For the stationary object of interest of experimental task 1 imagine a house being observed; the letter display might indicate a person appearing at a window of the house, while the type of letter might indicate a certain person appearing. For the moving object of interest of experimental task 2 imagine a car being observed; the letter display might indicate a person appearing at a window of the car, while the type of letter might indicate a certain person. As the focus of the investigation is on selecting events without any manual intervention by the user, we chose experimental tasks taking less than 15 minutes to complete. This avoids a vigilance decrement, which would typically occur after observation times of 15 minutes or more [15]. This way, we can disentangle event detection in situations of missing manual intervention from event selection in situations of missed events due to oversight. In order to check the attention of the subjects, we let them count the number of occurrences of the target letter ‘e’.

The duration of experimental task 1 was 8:30 min including 72 letters (24 of them target letters); the duration of experimental task 2 was 9:30 min including 63 letters (23 of them target letters). In both experimental tasks, the subjects were told to monitor the object of interest constantly, and to count the number of displayed letters ‘e’. To keep the number of errors in counting small, a dialog box popped up every 90 seconds asking to enter the number of ‘e’ since the last request.

2.2 Apparatus

Eye gaze measurement was done using a Tobii X60 remote eye tracker, placed beneath a 24’’ monitor (1920x1200 pixel) using a Tobii Monitor Mount. Data sampling rate was 60 Hz. The manufacturer reports an accuracy of 0.5 degrees of visual angle. Subjects were sitting with a distance of about 65cm to the monitor, so 0.5° correspond to a region of gaze estimation uncertainty of about 1.13 cm or 43 pixels on the screen. For two reasons, no chinrest was used in the experiment: first, the manufacturer reports a freedom of head movement of 44x22x30 cm (W x H x D) for the Tobii X60 not reporting any accuracy

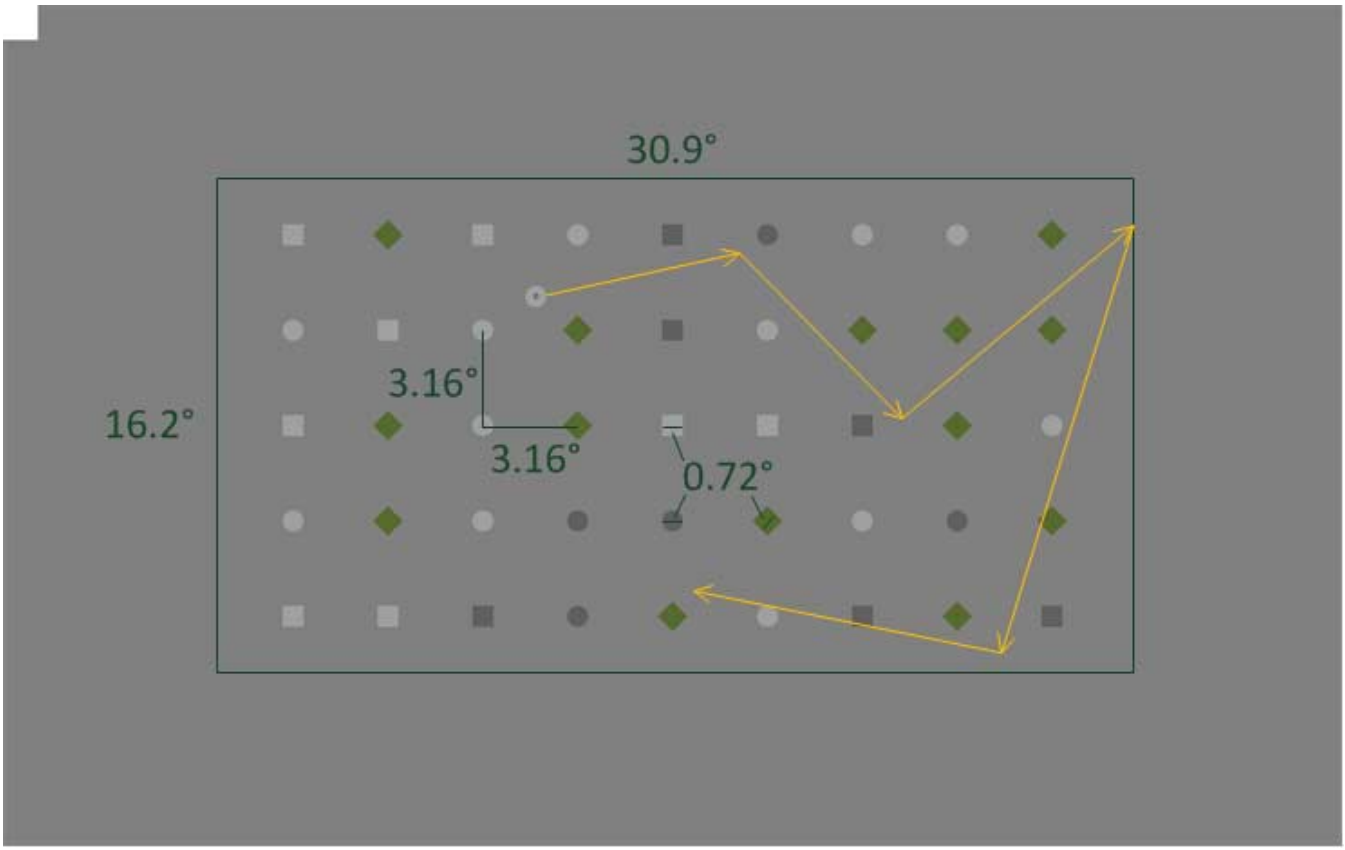

Figure 2: Visual stimulus of experimental task 2 (observation of a moving object).

decrement; second, in expert video surveillance using a chinrest would not be acceptable for the human operators. Eye tracker calibration was performed once at the beginning of the experiment.

EEG was recorded using a BrainProducts actiCHamp recorder with 28 active electrodes arranged according to the international 10-10 system referenced at Pz (see [11] for the electrode setup) and three monopolar Electrooculography (EOG) electrodes placed around the eyes according to [12]. Impedance was kept below 16kOhm for all electrodes. Synchronization of EEG data, eye tracking data and tasks was performed using a photo diode attached to the screen above a marker box which switched color during every letter vs. no-letter display (corresponding to the first event indication stage). This guarantees synchrony of data streams at a frame level.

2.3 Procedure

10 subjects (9 male, 1 female) between 21 and 42 years (average age = 25.9 years) volunteered in the study. All were students or colleagues. All had normal or corrected to normal sight. Five wore glasses. Three had used a remote eye tracker before; none had used EEG equipment before. The data of one additional recorded subject was excluded from analysis as technical problems with the EEG electrodes rendered several channels invalid.

After being provided with the EEG equipment, the subjects performed a short 9-point calibration routine. Then a short introduction into the experiment followed including an instruction of the tasks (fixate object of interest and count number of target letters ‘e’) and one training task for each experimental task (duration about 1 minute, 4 distractor letters, 3 target letters). Then, the investigator left the room, and the subjects independently performed first experimental task 1, and then experimental task 2.

3. RESULTS

3.1 EEG for Temporal Event Localization

For the analysis of EEG data, two different classification tasks are analyzed separately, corresponding to the two event indication stages: The first classification task distinguishes between ’letter’ versus ‘no letter’. The second classification task distinguishes between the target letter ‘e’ versus any of the distractor letters. For each classification task, EEG data is segmented according to the respective stimulus and labeled accordingly. Each trial comprises a window of 500ms of EEG data and is baseline corrected with 100ms of data prior to this window. ’No letter’ sequences are extracted from the EEG signal starting 1000ms prior to each event, such that they do not overlap with either the previous or the following ’letter’ sequence or the corresponding baseline normalization data.

Prior to extracting each window, channels that appeared to be erroneous during the recording session (high noise, loose electrodes) were excluded from further analysis. Data is re-referenced to a common average reference. Afterwards, each trial is corrected individually by detrending each channel to compensate for signal drifts. Additionally, each trial is preprocessed using a low pass filter with a cut off frequency of 15 Hz and a line noise filter.

A challenge of the data set is the frequent eye movement which contaminates the EEG recordings. Spatial filters provide a framework for dealing with non-neural influences on the signal. In this work, we employ Stationary Subspace Analysis (SSA) [14]. SSA is an unsupervised linear transformation which decomposes a multichannel signal in stationary and non-stationary components. Removing the non-stationary subspace and evaluating the system only on the stationary subspace can potentially increase the generalization ability of the classifier.

Features for every trial are extracted by down-sampling the signal to windows of 83ms for every electrode. Afterwards all features are z-normalized. A Support Vector Machine (SVM) with Radial Basis Function kernel is utilized as the classification model using LIBSVM introduced by Chang and Lin (2011) [1] with default parameters. We use the accuracy of this model as a quality metric for the recognition performance of our setup. For evaluation, we perform person-dependent leave-one-out cross-validation, for each experiment and classification problems separately.

Table 1: Recognition accuracy for both experiments and both classification tasks. We report results for classifiers with and without SSA as preprocessing step.

	Exp. 1	Exp. 2
Letter vs. No letter	76.0%	73.8%
Letter vs. No letter (SSA)	83.0%	80.7%
Target letter vs. Distractor letter	63.2%	59.1%
Target let. vs. Distractor let. (SSA)	63.6%	62.1%

Table 1 shows the resulting recognition accuracy for both classification problems. We see that the classifier provides a very robust detection of 'letter' vs. 'no-letter'. Despite a much more challenging task setup, those results come close to the results reported in [11]. This is the case for both experimental conditions, which indicates that both sporadic relaxing eye movements and eye blinks (experimental task 1) as well as continuous object tracing (experimental task 2) can be handled. SSA substantially increases recognition accuracy, with a much stronger impact on the 'letter' vs. 'no-letter' task with a relative improvement of up to 8.8%. As expected, discriminating target stimuli from distractors is a much more difficult task, resulting in lower recognition accuracy. Still, we see that the classifier is able to extract usable information from the provided data. Overall, those results indicate that the system is already able to reliably detect events which are perceived as potentially relevant to the task by the participant. For further analysis of the task-relevance of an event, the EEG signal can provide a reasonable estimate but are not yet reliable enough for fully automatic detection of critical events.

3.2 Eye Tracking for Spatial Event Localization

For spatial event localization, the fixation position on the object of interest had to be determined. For fixation detection, the fast and robust I-DT-algorithm was used, with a dispersion of 0.5 degree of visual angle; minimum fixation duration was set to 100ms [13].

Fixations have to fulfill two requirements. First, considering the distance of 3.16° between the object of interest and the neighboring distractor objects in experimental task 1, unambiguous spatial event localization requires a fixation distance (DIST) of 1.58° or less to center of the object of interest. Second, in order to complete the spatial localization within the time temporal localization requires (see subsection above), fixations must be completed 500ms after the letter was initially displayed; hence, as fixation duration is set to 100ms, fixation onset times (FOT) are restricted to fall into an interval between 0 and 400ms after the letter display started. In order to be able to compare the results of experimental tasks 1 and 2, the same distance condition of 1.58° was applied to the evaluation of experimental task 2; again, fixations had to be complete 400ms after the letter display started.

For experimental task 1 (stationary object of interest), a total of 72 letters (24 target letters) had to be localized for each subject. For experimental task 2 (moving object of interest), a total of 63 letters (23 target letters) had to be localized for each subject.

Table 2: Results for spatial localization for both experiments. We report results for hit rates, Distance of Fixations to the object of interest, and Fixation Onset Times.

	Exp.1			Exp.2		
	Hit Rate (%)	Mean DIST (SD)	Mean FOT (SD)	Hit Rate (%)	Mean DIST (SD)	Mean FOT (SD)
Letter	92.6	0.82° (0.32)	203 (129)	81.9	0.78° (0.40)	217 (132)
Target letter	94.2	0.81° (0.32)	195 (127)	84.3	0.82° (0.38)	221 (132)

Table 2 shows the results for both experimental tasks. Hit rates are clearly better for the stationary object of interest compared to the moving object, for both letters and target letters. While the distance of the fixations to the object of interest (DIST) as well as the fixation onset times (FOT) are practically the same for letters and target letters in both experiments, it appears that at least the lower hit rate might indicate that the subjects looked less thoroughly at the distractor letters.

Between letter displays, the subjects relaxed their attention a little, mean DIST being 1.13° (SD 0.56) for the stationary condition, and 1.59° (SD 1.13) for the moving condition. Thus, during the observation the object of interest was lying mostly within the area captured by parafoveal vision (5° of visual angle [7]).

3.3 Counting Task

The subjects were quite good in counting the number of target letters. Of 430 target letter occurrences, the subjects missed only 3 for the stationary condition, and 6 for the moving condition. It appears that as we expected no vigilance decrement occurred.

4. CONCLUSION

The results show that for basic surveillance tasks spatio-temporal event localization using eye tracking and EEG can be performed in both time and space in a robust fashion. For the 'letter' vs. 'no-letter' localization which would be a similar task as reported in [11], we achieve a temporal accuracy of 83.0% and a spatial accuracy of 92.6% for the stationary object paradigm; for the moving object paradigm, we achieve a temporal accuracy of 80.7% and a spatial accuracy of 81.9%. Assuming independence, we can expect to perform successful combined spatio-temporal localization for 76.9% (83.0% * 92.6%) resp. 66.1% (80.7% *81.9%) of all cases, fitting the results reported in [11] quite well.

The target letter/distractor letter localization represents a far more challenging task. Assuming independence, we can expect to perform successful combined spatio-temporal localization for 60.0% (63.6% * 94.2%) of all cases for the stationary object paradigm, and 52.4% (62.1%* 84.3%) of all cases for the moving object paradigm.

Hence, enhancing a given user interface for event selection by the proposed method can provide a valuable fallback if the human operator fails to select an event manually, as the proposed system provides a reasonable estimate of the task-relevance of an event as well as its spatial location. Given the lower accuracy for the second classification task, the EEG-based temporal event selection is not yet reliable enough for fully automatic detection of critical events.

Considering completion time, the eye-brain input achieves event selection within 500ms with the chosen parameters for spatio-temporal localization, which is very fast. Examples of reported selection completion times using the gaze+key technique are, e.g., 600ms or more for stationary object selection [16], or 768ms for moving object selection [5].

As the classifiers used are not fine-tuned, it can be expected that even better performance could be achieved, especially, if EEG and eye tracking data are combined in the classifier. Further improvement might result if event detection would additionally be extracted from gaze behavior.

Next steps will continue the analysis with the design of experimental tasks which add more of the complexity of surveillance tasks. For example, such tasks would comprise simultaneous events at the same time, distractors both far away as well as closer the event, or tasks which require more complex decision making and reasoning than the now used simple counting task. Furthermore, this might include tasks lasting for more than 15 min to investigate vigilance decrement, as well as comparing event selection by the eye-brain approach to mouse input or the gaze+key technique.

Another interesting research direction is the comparison of our approach with existing computer vision based techniques for event detection which are, as already mentioned in the introduction, increasingly getting integrated into interactive surveillance systems. It would be valuable to find out, how the automated components and the multimodal user interface should work together for best supporting a human operator in a surveillance task.

5. REFERENCES

[1] Chang, C.-C., and Lin, C.-J. 2011. LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology* 2, 3, 27:1-27:27.

[2] Fischer, Y., Krempel, E., Birnstill, P., Unmüßig, G., Monari, E., Moßgraber, J., Schenk, M., and Beyerer, J. 2014 (accepted for publication). Privacy-Aware Smart Video Surveillance Revisited. In *Proceedings of the 9th Security Research Conference (Future Security)*.

[3] Heinze, N., Esswein, M., Krüger, W., and Saur, G. 2010. Image exploitation algorithms for reconnaissance and surveillance with UAV. In *Proceedings of SPIE Defense, Security, and Sensing*, International Society for Optics and Photonics (2010), 76680U--76680U.

[4] Hild, J., Brüstle, S., Heinze, N., and Peinsipp-Byma, E. 2013. Gaze interaction in UAS video exploitation. In *Proceedings of SPIE Defense, Security, and Sensing*, International Society for Optics and Photonics (2013), 87400H--87400H.

[5] Hild, J., Müller, E., Klaus, E., Peinsipp-Byma, E., and Beyerer, J. 2013. Evaluating Multi-Modal Eye Gaze Interaction for Moving Object Selection. In *Proceedings of the Sixth International Conference on Advances in Computer-Human Interactions*. ACHI 2013. IARIA, 454-459.

[6] Huang, B. and Lo, A. 2013. Integrating EEG information improves performance of gaze based cursor control. In *Proceedings of Neural Engineering*. IEEE, 415-418.

[7] Irwin, D. 1992. Visual memory within and across fixations. In *Eye movements and visual cognition,* Springer, 146-165.

[8] Jacob, R.J.K. and Karn, K.S. 2003. Eye tracking in human-computer interaction and usability research: Ready to deliver the promises. In *Mind 2*, 3 (2003), 573-605.

[9] Just, M.A. and Carpenter, P.A. 1976. Eye fixations and cognitive processes. *Cognitive psychology* 8, 4 (1976), 441-480.

[10] Krusienski, D.J., Sellers, E.W., McFarland, D.J., Vaughan, T.M., and Wolpaw, J.R. 2008. Toward Enhanced P300 Speller Performance. *Journal of Neuroscience Methods* 167, 1 (2008), 15–21.

[11] Putze, F., Hild, J., Kärgel, R., Herff, Ch. Redmann, A., Beyerer, J. and Schulz, T. 2013. Locating User Attention using Eye tracking and EEG for Spatio-temporal Event Selection. In *Proceedings of the 2013 international conference on intelligent user interfaces.* IUI 2013. ACM, 129-136.

[12] Schlögl, A., Keinrath, C., Zimmermann, D., Scherer, R., Leeb, R., and Pfurtscheller, G. A. 2007. Fully Automated Correction Method of EOG Artifacts in EEG Recordings. *Clinical Neurophysiology* 118, 1 (Jan. 2007), 98-204.

[13] Salvucci, D.D. and Goldberg, J.H. Identifying fixations and saccades in eye-tracking protocols. 2000. In *Proceedings of the 2000 symposium on Eye tracking research & applications.* ETRA 2000. ACM, 71-78.

[14] von Bünau, P., Meinecke, F., Kiraly, F., and Müller, K.-R. 2009. Finding stationary subspaces in multivariate time series. *Physical Review Letters* 103, 21. APS.

[15] Teichner, W.H. 1974. The detection of a simple visual signal as a function of time on watch. *Human Factors* 16, 4, 339-352, Sage Publications.

[16] Ware, C., and Mikaelian, H. 1987. An evaluation of an eye tracker as a device for computer input. *ACM SIGCHI Bulletin* 17, SI, ACM, 183-188.

[17] Yong, X., Fatourechi, M., Ward, R.K., and Birch, G.E. 2011. The Design of a Point-and-Click System by Integrating a Self-Paced Brain Computer Interface with an Eye-Tracker. *IEEE Journal on Emerging and Selected Topics in Circuits and Systems* 1, 4 (December 2011), 590 –602.

[18] Zander, T.O., Gaertner, M., Kothe, C., and Vilimek, R. 2010. Combining Eye Gaze Input with a Brain-Computer Interface for Touchless Human-Computer Interaction. *International Journal of Human-Computer Interaction* 27, 1 (2010), 38–51.