Evaluating Multi-Modal Eye Gaze Interaction for Moving Object Selection

Jutta Hild*, Elke Müller*, Edmund Klaus*, Elisabeth Peinsipp-Byma*, Jürgen Beyerer*†
*Fraunhofer Institute of Optronics, System Technologies and Image Exploitation (IOSB)
Karlsruhe, Germany
†Vision and Fusion Laboratory, Karlsruhe Institute of Technology (KIT)
Karlsruhe, Germany
e-mail: {jutta.hild; elke.mueller; edmund.klaus; elisabeth.peinsipp-byma; juergen.beyerer}@iosb.fraunhofer.de

***Abstract*—Moving object selection is a frequently occurring interaction task in expert video analysis. State-of-the-art, video analysts use the mouse as input device. As the selection of moving objects is more complex than the selection of static objects, particularly when objects move fast and unpredictable, using the mouse is less efficient than usual and induces more manual stress. In this contribution, multi-modal gaze-based interaction is proposed as an alternative interaction technique for moving object selection. In an experiment using an abstract moving circle scenario, the two gaze-based interaction techniques gaze + key press and liberal MAGIC pointing are compared to mouse interaction. Evaluation of both user performance and user satisfaction shows that at least the gaze + key press technique might be a promising interaction alternative for moving objects selection.**

Keywords-moving object selection; eye gaze interaction; multi-modal interaction; experiment; video analysis

I. INTRODUCTION

The selection operation is one of the basic interaction tasks in human-computer interaction for desktop computer systems with graphical user interfaces (GUI). To perform a selection task, most users use the mouse as an input device because it provides an intuitive and effective interaction technique for the selection of typical GUI elements of different sizes as, e.g., windows, icons, or buttons. However, there are applications that also require the selection of moving objects. E.g., in live-video analysis for security applications moving object selection is a frequently occurring interaction task.

A. Moving object selection in video analysis

Basically, video analysis is a visual search task aiming to detect and analyze objects and situations for conspicuous events, and to report them to the authorities concerned with the response actions. Therefore, video analysts are equipped with video exploitation systems providing a large representation of the scene currently by the video sensor. Figure 1 shows, as an example, the ABUL user interface [1]. Typically, analysts use the visualization in the large window in the center to perform their search task by continuously scanning it for conspicuous events. When detecting an event, the analysts have to mark it in the video stream and send the labeled video clip to the concerned authorities for further investigation/examination.

But, due to motion of the video sensor all objects including static ones like buildings move, too. Hence, object selection using mouse interaction can be rather challenging. The reason for this can be revealed easily considering the selection process for pointer-based input devices like mouse, consisting of the three parts:

1. Picking the object to be selected (by gaze)
2. Pointing at the object by moving the selection device onto the object (by motor action to position the mouse pointer)
3. Actuating the selection (by motor action to perform a mouse button click)

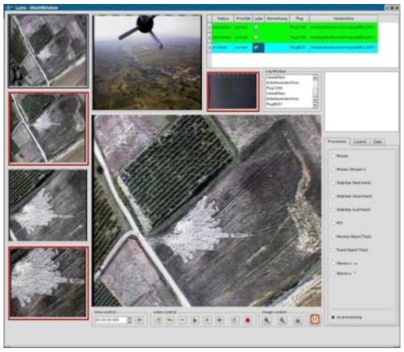

Figure 1. ABUL user interface tailored for moving object selection.

If a static object has to be selected the three parts are each performed once. If a moving object has to be selected it might be necessary to perform at least part two more than once, as the object might move away from the mouse pointer position before the user is able to perform the mouse click. There may be even a number of pick-point-click-loops necessary before the user finally is able to accomplish the moving object selection. This can happen, e.g., if due to high wind the video sensor, and accordingly the video image is shaking. Additionally, as well due to video sensor motion, objects are visible on the screen only for a limited time span, and accordingly selectable only for a short time. Thus, the analyst has to perform the moving object selection very fast, especially, if an object moves in addition to the sensor motion, as for example driving vehicles do.

However, as a missed object selection in security applications might result in severe consequences, it is important that the analyst is able to reliably select the moving objects. As a consequence, the load for the hand and arm motor systems is rather high. According to statements made by expert video analysts, their working time often lasts for several hours. Hence, mouse interaction often induces repetitive strain injury. For these reasons it is worth developing an alternative interaction concept which allows moving object selection, on the one hand, in a more efficient way, and, on the other hand, with reduced manual stress.

B. Proposing eye gaze interaction for moving object selection

The alternative which is proposed in this contribution is eye gaze interaction. Technically, eye gaze interaction is facilitated by an eye tracking device which provides continuously the current gaze position of the user on the screen. Over the years, a lot of different eye gaze interaction variants have been proposed [2,3,4,5]. They can be subdivided into unimodal and multimodal interaction techniques. While with the unimodal techniques the whole selection process described above is performed using eye gaze input, multi-modal eye gaze interaction techniques use at least one other input modality.

However, both unimodal and multimodal techniques have in common that they use eye gaze position for pointing. As a result, parts 1 and 2 of the mouse selection process described above are performed together, at the same time. This property is well suitable for moving object selection in video analysis. As introduced before, video analysis is basically a visual search task and displayed objects might be visible and selectable only for a short time. Therefore it is essential to ensure that the analysts can keep their visual focus of attention continuously on the object during the selection process. This is possible using eye gaze interaction. But it is not for mouse interaction where the visual attention has to be focused onto the mouse pointer when looking for it and repositioning it. Besides the advantage of not having to divide the visual focus of attention between object and pointing device, to perform parts 1 and 2 together should result in an overall shorter selection time. This could be beneficial if an object is visible in the video only for a short time or if its position changes quickly and unpredictably. Furthermore, as the user has no longer to move the pointing device by hand-motor action the user's manual stress is reduced, which could contribute to a decrease of hand motor complaints.

What makes the eye gaze interaction techniques different is the way part 3, the equivalent to the mouse click equivalent, is performed. Examples of uni-modal eye gaze interaction techniques use, e. g., a certain eye gaze dwell time on the object, or an eye blink [3]. Examples of multi-modal eye gaze interaction techniques use, e.g., a button press on a keyboard [2,6,7], manual pointing [4], or even speech [8]. All of them have their advantages, but as our goal is to provide an interaction concept for a, both mentally and physically, stressful working task, techniques that would add additional load are not appropriate.

For this reason, both dwell time and eye blink are not appropriate as both require unnatural eye movement. Furthermore, as stated, e.g., in [4], one should be careful to load gaze with a motor control task as it has been evolved for visual perception. Hence, performing a motor control task by eye gaze would result in additional mental load. Besides, to overload the eye normally used for information input with an information output task results in the so called "Midas Touch" Effect. As the eyes are an always-on device it is necessary to ensure by appropriate interaction design to make eye movements for perception distinguishable from eye movements for pointing or selection [3,5]. Therefore, for our concept eye gaze shall not be used for the final selection actuation.

As the multi-modal interaction techniques use an input modality for selection actuation different from eye gaze they do all avoid the Midas Touch effect [7]. Furthermore, both hand and speech are natural means for object selection. As state-of-the-art video exploitation systems not necessarily comprise a speech recognition system we did not consider speech as input modality at this time. The button press on a keyboard being referred to as "hardware button" [2], "Gaze and Keyboard" [5], "Eye tracking with manual click" [7], or "Eye + Spacebar" [6], has been reported by these (and other) authors to be a rather fast, easy-to-use, and effortless interaction technique. In addition to this, several times it was reported to be the favorite technique of subjects for selection tasks in experiments [5,6]. Because of these good assessments for static object selection, this technique was selected to be one of the appropriate candidates for performing moving object selection and was therefore evaluated in our experiment. From now on, we will refer to it as the "gaze + key press" technique (GK).

However, there is one major drawback of techniques using eye gaze for pointing. On the one hand, due to eye tracking inaccuracy and, on the other hand, due to anatomic and physiological properties of the eye, it is not possible to determine the exact gaze position on the screen. Eye tracker manufacturers often specify an accuracy of 0.5°. This means, whatever gaze position is estimated by the eye tracker, the true position might be located within a circle of 1° of visual angle around the estimated position. Furthermore, as the region of sharp vision, the fovea, covers 1° to 2° of visual angle, and as it jitters continuously to keep the visual stimulus on the retina alive it represents a "region of uncertainty" for the eye tracker [9]. To get higher accuracy, in [4] a gaze-based interaction technique called MAGI pointing, combining eye gaze and manual interaction, was proposed. The concept is to use eye gaze for coarse pointing, and a pointer-based interaction technique performed by hand for fine pointing and selection actuation. As one of the MAGIC pointing approaches, the so called liberal MAGIC pointing, performed faster than mouse interaction in a static object selection task, it was also considered to be an appropriate candidate for performing moving object selection and was therefore evaluated in our experiment. From now on we will refer to it as the "MAGIC liberal" technique (MAGIC-lib).

II. RELATED WORK

The challenge of moving object selection using mouse interaction has been addressed before by several authors. A comprehensive survey on the issue is provided by [10], including considerations on moving object selection in videos. Discussed techniques for adaption and improvement of mouse interaction can be categorized as techniques that extend the activation area of the mouse pointer according to object size [11] and techniques that extend the activation area of the object itself [10], both being reported to provide improved moving object selection using mouse.

Selection with eye gaze has been widely examined for static objects. As stated in the introduction, the GK technique was part of many evaluations [2,5,6,7] and found to be faster than mouse, intuitive, easy-to-use, and manually little stressful. However, selection accuracy was much lower than mouse, resulting in higher selection error rates. The MAGIC pointing paradigm was also evaluated in various variants by several authors and reported to be a good mouse alternative [4,12].

Contributions comprising gaze-based moving object selection are available for computer gaming applications. In [13] a survey is presented proving taxonomy of game genres and eye gaze interaction applicability. Two genres, shoot-them-up games and first-person shooter games, comprised moving object selection (shooting). When gaze was used as an input for first-person shooters (gaze being used for crosshair control, mouse click for triggering the weapon fire) studies showed that gaze input could not beat mouse input [13,14]. Targeting the objects with gaze provided much lower accuracy than the mouse. Another study used gaze as an input device for a chicken shoot game [13]. After a few practice trials, the participants were able to perform much faster with eye controlled shooting (combination of controlling the cross-hair by gaze position and triggering weapon fire by mouse click) than with mouse and keyboard.

III. EXPERIMENT

To examine whether, for moving object selection, the gaze-based interaction techniques gaze + key press (GK) or liberal MAGIC pointing (MAGIC-lib) would provide a promising alternative to mouse interaction an experiment was designed. To compare the three interaction techniques, a simple moving object selection task was designed simulating the real moving object selection in video analysis as described in the introduction. For every interaction technique, performance was measured by determining selection completion times and selection error rates. Satisfaction was evaluated using a questionnaire asking to rate selection accuracy, selection speed, task adequacy, and user friendliness. In addition, participants were asked for their favorite interaction technique for moving object selection.

A. *Interaction Techniques*

Object selection by mouse interaction used the typical point-and-click-procedure with a left mouse button click. With the GK interaction technique, object selection required to look at the object and to actuate the selection by pressing the ENTER-key while fixating the object. MAGIC pointing was implemented using the liberal approach. The pointer was continuously visualized at the currently measured fixation position using the I-DT fixation algorithm [15]. For selection, the mouse had to be moved over the object, and the left mouse button had to be clicked, just in the standard mouse interaction. The change of pointer control between eye tracker (gaze) and mouse (hand) worked as follows. By default, the control of the pointer lay with the eye tracker. In case of mouse movement, the pointer control was transferred immediately to the mouse. Pointer control was transferred back to the eye tracker only after the mouse had not been moved for 100 ms.

B. *Apparatus*

To capture the gaze data, a table-mounted Tobii 1750 remote eye tracker was used. It is a video-based eye tracker using pupil-center and corneal-reflection for gaze measurement. The eye-tracking hardware (video cameras and near infrared LEDs) is integrated with the display frame of a 17-inch TFT display with a 1280x1024 resolution. According to manufacturer information, the Tobii 1750 features an accuracy of 0.5°, and a spatial resolution of 0.25°. Gaze position is sampled at 50 Hz.

Participants sat at a distance of 60 cm from the monitor which according to the manufacturer allows a freedom of head-movement of 30 x 16 x 20 cm (W x H x D). Therefore, no chin-rest was used in order to let participants use the eye tracker with the same freedom of movement that would be required for video analysts at work.

Mouse input for mouse interaction and MAGIC-lib interaction was performed using a standard optical mouse. To perform the key press for the GK interaction, a standard keyboard's ENTER-key was used.

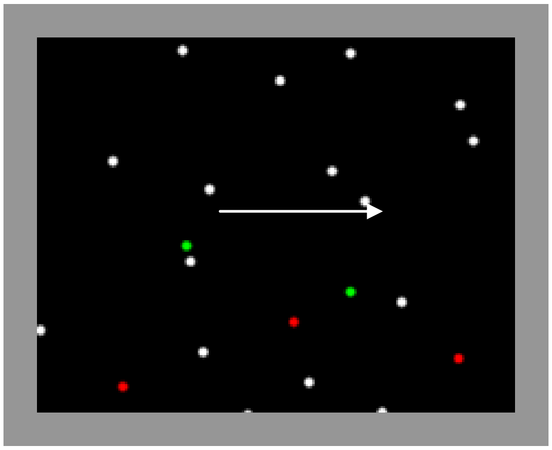

Figure 2. Screenshot of the experimental task. The arrow shows the moving direction.

C. *Experimental Task*

The moving object selection task was designed referring to movements of persons or vehicles in real world surveillance videos. Figure 2 shows a screen shot of the task. The objects were represented by white circles. All had the same size of a diameter of 24 pixels (corresponding to 0.63 cm on the used monitor, and 0.66° of visual angle). The objects moved over the screen from left to right on a black background. Each object moved with one of three constant speeds resulting in a visibility on the screen of about 7 to 9 seconds. While moving over the screen, some prior to the start of the experiment randomly chosen objects were highlighted in red for one second. In total, 294 objects moved over the screen, 91 of them were highlighted. To induce some of the time pressure and stress video analysts are faced with, the number of objects on the screen steadily grew during the task. In total, to perform the test task once lasted for about 180 seconds. The task was to select the highlighted objects. However, all objects were selectable. After being selected, an object changed its color into green as a visual feedback for the user.

As pointing based on gaze position can provide only limited accuracy, particularly, when allowing some head movement, a radius of acceptance of 80 pixels was assigned to the objects. Thus, the visible size of the circles was 24 pixels, while their selectable size had a diameter of 160 pixels (3.97° of visual angle). This approach is related to the object magnifications proposed, e. g, by [10] to improve mouse interaction. However, in contrast to their suggestions, our approach uses an invisible object magnification. This is more suitable for video analysis applications as they usually do not allow any part of the video to be covered by artificial visualizations. Regarding real applications, our radius of acceptance approach would require information about the position coordinates of the objects. This requirement is met as state-of-the-art video exploitation systems like ABUL [1] feature automatic target recognition algorithms which are able to preprocess the video material and to provide exploitation results like object position coordinates.

D. *Participants*

20 participants volunteered for this study. All of them were university students or members of our department and were therefore experienced users of desktop computer and mouse. Two of them could not complete the tasks due to calibration problems and were therefore excluded from the data analysis. The age of the remaining 18 participants (16 male, 2 female) ranged from 18 to 54 years (mean = 28.6 years). All participants had normal or corrected to normal vision, two of them wore glasses, three used contact lenses. Nine participants had already used the Tobii 1750 eye tracker once before.

E. *Procedure*

For the study, a within-subjects design was used. Each participant performed the selection task three times, each time using a different interaction technique. To control fatigue and training effects the participants were divided into 6 groups, each group performing the experiment with a different order of techniques, thus following a complete, counterbalanced design.

The experiment started with an explanation of the moving object selection task using a short practice task of about 50 moving circles. Participants were told to perform selections as fast as possible and with the least possible number of mistakes. In case of performing one of the two gaze-based interaction techniques, the participants completed the standard Tobii 1750 calibration using a grid of 9 calibration points before practicing. Participants were allowed to repeat the practice task until they felt to be familiar with the interaction technique. After that, in case of a gaze-based interaction technique a recalibration had to be completed, again using the 9-point-calibaration-procedure. Then, the test task of 294 circles was performed. Finally, the participants rated rate selection accuracy, selection speed, task adequacy, and user friendliness on a 5-point scale (1: best rating, 5: worst rating) and reported their favorite interaction technique for moving object selection.

IV. RESULTS

Figure 3 shows the results for the selection completion time (SCT). SCT was measured by the difference between the time of selection and the time of highlighting. A repeated-measures analysis of variance (ANOVA) shows that the participants' performance significantly varied with techniques ($F(2;50) = 510.7$; $p<0.001$). A post-hoc analysis with Bonferroni correction shows high significant differences in SCT between GK and both of the other techniques ($p<0.001$), and reveals a significant result between the SCT of mouse and MAGIC-lib ($p<0.01$). With an average SCT of 786 ms the GK technique was considerably faster as the mouse interaction technique with an average SCT of 1422 ms. Participants performed slowest with MAGIC-lib (average SCT = 1681 ms).

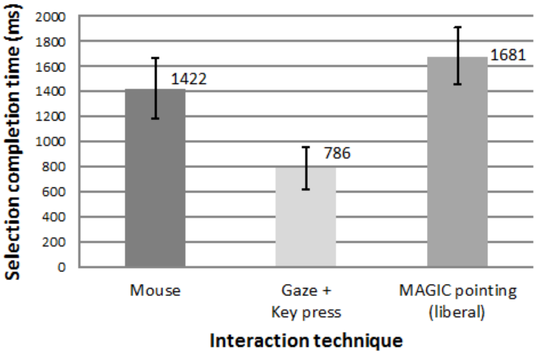

Figure 3: Selection completion time as a function of interaction technique.

Figure 4 shows the results for the selection error rate ((number of wrong selections / 294) * 100). It varies significantly with techniques as well ($F(2;50) = 17.41$; $p< 0.001$). Bonferroni-corrected comparisons show that the error rate for the MAGIC-lib technique is significantly higher than for GK and mouse ($p<0.0001$). The difference between GK and mouse is not significant. Participants performed with an average error rate of 3.3 % using mouse, with an average error rate of 3.9 % using GK, and with an average error rate of 8.3 % using MAGIC-lib. Error rates were similar for mouse and GK, whereas they were about double for MAGIC-lib.

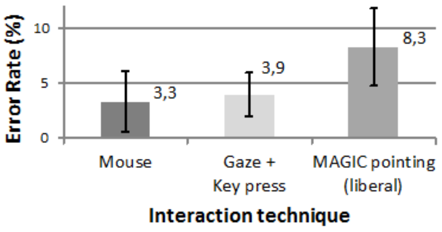

Figure 4. Selection error rate as a function of interaction technique.

Analyzing the errors in more detail, Figure 5 shows the percentage of missed objects (highlighted objects that could not be selected) and the percentage of false alarms (non-highlighted objects that were selected by mistake). An ANOVA of the number of missed objects shows a significant result (F(2;50) = 21.55; p<0.001). The Bonferroni post-hoc correction reveals significant differences between missed objects of MAGIC-lib and both GK and mouse (p<0.001). There is no significance between GK and mouse. In absolute numbers, participants missed with mouse on average 8.4 objects (out of 91), with GK 7.5 objects, and with MAGIC-lib 22.0 objects. Again, participants performed similar using mouse and GK and using MAGIC-lib considerably worse.

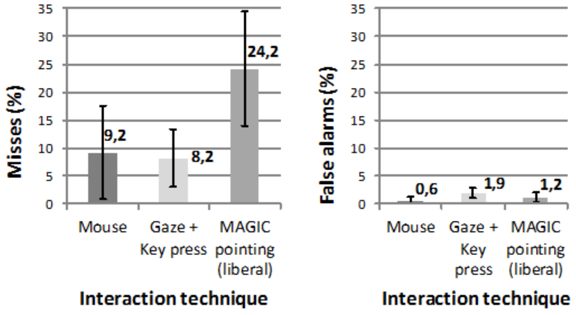

Figure 5: Misses and false alarms as a function of error rates.

An ANOVA of the number of false alarms shows a significant result as well (F(2;50) = 11.02; p < 0.001). Significant differences of false alarms between mouse and GK (p<0.001) are revealed by the post-hoc analysis with Bonferroni correction. At a p = 0.05 level a significant result is shown between mouse and MAGIC-lib. No significance is given between GK and MAGIC-lib (p=0.153). In absolute numbers, using mouse there were on average 1.3 (out of 203) objects selected by mistake, using GK it were 3.9 objects, and using MAGIC-lib it were 2.4 objects.

Figure 6 shows the results of the ratings of user satisfaction from the questionnaire. For all categories, MAGIC-lib got the worst ratings. For selection accuracy, the participants rated mouse and GK almost equally good. For selection speed, task adequacy, and user friendliness, GK was rated best, followed by mouse and MAGIC-lib.

Finally, asked for their favorite interaction technique for moving object selection, 16 out of the 18 participants voted for GK, the other two for mouse interaction.

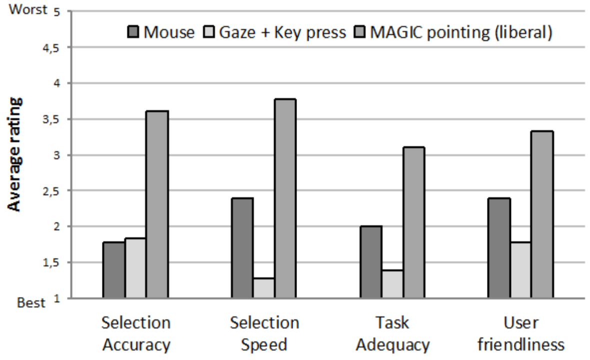

Figure 6. Subjective ratings for user satisfaction.

V. DISCUSSION

As shown for the selection of static objects [2,5,6,7], the GK technique also provides a very fast technique for the selection of moving objects. In the experiment presented here, participants were able to perform much faster than with mouse interaction. While the liberal MAGIC pointing approach provided a little faster object selection for static objects as mouse [4], this was not the case for moving object selection. Only one participant was able to select faster with MAGIC-lib than with the mouse. On the other hand, MAGIC pointing is a technique which is close to mouse interaction – usually at its end at least a small mouse interaction is completed. It is clear that it is rather challenging for the participants, both mentally and due to slightly different hand motor action, to use a technique that is almost like the one they are used to. Due to this fact, as already mentioned in the description of the procedure of the experiment, participants were allowed to repeat the practice task with each technique as often as they wanted. But, of course, even practicing for several minutes cannot compensate years of mouse experience. Moreover, some participants felt distracted by the permanent visibility of the mouse cursor. This could, of course, also have increased the selection completion time. The subjective estimations for selection speed in the questionnaire confirm the results of the performance measurement. All participants rated GK being the fastest selection technique. After all, 3 participants rated MAGIC pointing to be faster than mouse.

Considering the selection error rate, mouse and achieve an equally good error rate. Of course, this is due to the rather generous radius of acceptance which surrounded the circles. For static objects, it has been reported that for a size of 2 cm and larger [7], GK provides rather good selection accuracy. On the other hand, in video analysis it might in many cases be sufficient to select an object within a radius of uncertainty. Again, participants performed worst using MAGIC pointing. But again, this might be due to its complexity and its closeness to mouse interaction. After all, at least two participants were able to perform with the fewest number of selection errors using MAGIC pointing. As for selection speed, the ratings from the questionnaire for selection accuracy meet the results from the performance measurement. However, considering the error rate in more detail shows that there is at least one parameter that refers to the uncertainty of eye gaze pointing as use in the GK technique: the false alarm rate is, even if rather low for all the techniques, highest for GK. The result for the misses is in line with the results of the total error rate.

The ratings for task adequacy and user friendliness were also best for GK which make it the most comfortable technique to use. Regarding mouse interaction, some participants noted that even within the short time of the experiment the selection of moving objects was very fatiguing using mouse.

VI. CONCLUSION AND FUTURE WORK

The experiment showed that for moving object selection gaze-based interaction can provide advantages. Overall, the GK technique holds the largest number of positive characteristics. It is fast and, as a consequence, few objects are missed. It is easy-to-learn, requires only little manual effort – an important characteristic for video analysts working for several hours –, and easy-to-use which all in all made the participants to prefer it for moving object selection. Considering the work of a video analyst, particularly a short selection completion time is important. The less time the analyst needs to select an object the earlier they can draw their attention to the next object. However, the accuracy of GK was so well due to the acceptance radius around the circles. Hence, there is a need to confirm selection accuracy similar to mouse accuracy for moving object selection using real scenarios, including unpleasant selection situations like close objects, small objects, or shaking videos. In this case, using a large radius of acceptance could result in an overlap of selection regions of objects. This would prohibit to reliably select a certain object unambiguously. The performance of the standard input technique, the mouse interaction, was not as fast as GK interaction. But, at least mouse interaction was fast enough to catch almost as much highlighted objects as the very fast GK technique. However, only after the short experiment, participants noted that moving object selection using mouse interaction is very fatiguing. Regarding MAGIC pointing, the participants did not get along with this technique as good as with mouse or GK. However, there were persons performing fast and accurate using this technique. Bearing in mind the distraction of some participants due to the permanent display of the pointer using MAGIC pointing it could be interesting to investigate other variants like the conservative approach or MAGIC touch [12] for moving object selection.

ACKNOWLEDGMENTS

The underlying project to this article is funded by the WTD 81 of the German Federal Ministry of Defense under the promotional reference E/E810/BC013/AF144. The authors are responsible for the content of this article.

REFERENCES

[1] N. Heinze, M. Esswein, W. Krüger, G. Saur. “Image exploitation algorithms for reconnaissance and surveillance with UAV.” In Proceedings of SPIE 7668, Paper 76680U, Airborne Intelligence, Surveillance, Reconnaissance (ISR) Systems and Applications VII: 7 April 2010, Orlando, Florida, United States. Bellingham, WA: SPIE, 2010.

[2] C. Ware and H. Mikaelian, “An Evaluation of an Eye Tracker as a Device for Computer Input”, ACM SIGCHI Bulletin, vol. 18(4), pp. 183-188, 1987.

[3] R. J. K. Jacob, “The use of eye movements in human-computer interaction techniques: what you look at is what you get.” ACM Transactions on Information Systems (TOIS) 9, vol. 2, pp. 152-169, 1991.

[4] S. Zhai, C. Morimoto and S. Ihde, “Manual and gaze input cascaded (MAGIC) pointing.” In Proceedings of the SIGCHI conference on Human factors in computing systems: the CHI is the limit, pp. 246-253. ACM, 1999.

[5] M. Kumar, A. Paepcke and T. Winograd, “Eyepoint: practical pointing and selection using gaze and keyboard.” In Conference on Human Factors in Computing Systems: Proceedings of the SIGCHI conference on Human factors in computing systems, vol. 28, pp. 421-430, 2007.

[6] X. Zhang, and I. MacKenzie. “Evaluating eye tracking with ISO 9241-part 9.” Human-Computer Interaction. HCI Intelligent Multimodal Interaction Environments, pp. 779-788, 2007.

[7] R. Vertegaal, “A Fitts Law comparison of eye tracking and manual input in the selection of visual targets.” In Proceedings of the 10th international conference on Multimodal interfaces, pp. 241-248. ACM, 2008.

[8] D. Miniotas, O. Spakov, I. Tugoy, and I. S. MacKenzie. “Speech-augmented eye gaze interaction with small closely spaced targets.” In Proceedings of the 2006 symposium on Eye tracking research & applications, pp. 67-72, ACM, 2006.

[9] I. S. MacKenzie, “An eye on input: research challenges in using the eye for computer input control.” In Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications, pp. 11-12. ACM, 2010.

[10] H. Khalad, T. Grossman and I. Pourang, “Comet and target ghost: techniques for selecting moving targets.” Proceedings of the SIGCHI Conference on Human Factors in Computing Systems CHI, pp. 839-848, ACM, 2011

[11] T. Grossman and R. Balakishnan, “The bubble cursor: enhancing targe acquisition by dynamic resizing of the cursor’s activation area.” Proceedings of the SIGCHI conference on Human factors in computing systems, pp. 281-290, ACM, 2005.

[12] H. Drewes and A. Schmidt, “The MAGIC touch: Combining MAGIC-pointing with a touch-sensitive mouse.” Human-Computer Interaction–INTERACT 2009, pp. 415-428, 2009.

[13] P. Isokoski, M. Joos, O. Spakov, and M. Benoît, “Gaze controlled games.” Universal Access in the Information Society 8, vol. 4, pp. 323-337, 2009.

[14] J. D. Smith and T. C. Graham. “Use of eye movements for video game control.” In Proceedings of the 2006 ACM SIGCHI international conference on Advances in computer entertainment technology, ACM, 2006.

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