International Conference on Artificial Reality and Telexistence
Eurographics Symposium on Virtual Environments (2016)
D. Reiners, D. Iwai, and F. Steinicke (Editors)

Blurry (Sticky) Finger: Proprioceptive Pointing and Selection of Distant Objects for Optical See-through based Augmented Reality

Ja Eun Yu and Gerard J. Kim

Digital Experience Laboratory, Korea University, Korea

Abstract

Most AR interaction techniques are focused on direct interaction with close objects within one's reach (e.g. using the hands). Interacting with distant objects, especially those that are real, has not received much attention. The most prevalent method is using a hand-held device to control the cursor to indirectly designate a target object on the AR display. This may not be a natural and efficient method when used with an optical see-through glass due to its multi-focus problem. In this paper, we propose the "Blurry (Sticky) Finger" in which one uses the finger to aim and point at a distant object, but focusing only on the target with both eyes open (thus without the multi-focus problem) and relying upon the proprioceptive sense. We demonstrate and validate our claim through an experiment comparing three distant pointing/selection methods: (1) indirect cursor based method using a 3D air mouse, (2) proprioceptive finger aiming (Blurry Finger) with a cursor, (3) proprioceptive finger aiming without a cursor. In the experiment, Blurry Finger showed superior performance for selecting relatively small objects and in fact showed low sensitivity to the target object size. It also clearly showed advantages in the initial object selection where the hand/finger starts from a rest position. The Blurry Finger was also evaluated to be the most intuitive and natural.

Categories and Subject Descriptors (according to ACM CCS): H.5.2 [INFORMATION INTERFACES AND PRESENTATION]: User Interfaces—Input devices and strategies

1. Introduction

The continued innovations and advances in computer vision, mobile/cloud computing and portable display devices have brought about a renewed interest in augmented reality (AR) as a prominent information visualization and interaction medium. In particular, optical see-through (OST) displays are becoming more compact and fashionably designed, and thereby getting accepted to the mass users. As for interaction, early OST-based AR systems have relied mostly on simple command or menu based methods e.g. via voice or button/touch input [Goo, Eps]. Recently, light and small mountable sensors (e.g. camera, depth sensors) have allowed for more natural interaction [Lea, Sof, OKA11].

However, most AR interaction techniques are focused on direct interaction with close virtual objects within one's reach (e.g. using hands) [LGB08, CDDGC13, PCB11, SL13, BVBC04, HBW11]. Interacting with distant objects, especially those that are real, has not received much attention. The most popular method is using a hand-held device to control a cursor on the AR display to indirectly designate a target object. This may not be the most natural and efficient method. Moreover, such a cursor based method can only be used on the overlaid display space in OST glasses. In most AR glasses, the augmented display is overlaid on a small part of the entire visual field. Figure 1 illustrates two possible cases of AR glasses, those with the augmentation space: (1) fully overlapping with that of the whole display (but rare in actual products) and (2) partially overlapping (more typical such as the Google Glass). In the latter typical case, the cursor can only move and cover the small rectangular region (and objects within) in the top left corner of the visual field.

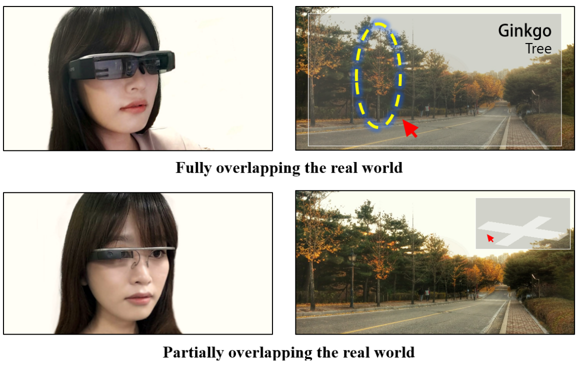


Figure 1: *Two possible coverage of the overlaid screen over the visual field in the AR glasses.*

Also, in OST based AR systems, one significant concern remains with regards to the multi-focus problem, i.e. the user having to frequently switch one's focus between an object of interest in the real

DOI: 10.2312/egve.20161434

world and the augmentation (e.g. a cursor) on the glass display.

In this paper, we propose, "Blurry (Sticky) Finger," in which one uses the unfocused blurred finger, the sense of proprioception and ocular dominance, to aim, point and directly select a distant object in the real world with both eyes open for the purpose of further interaction. The word "Sticky" is used to acknowledge the work of Pierce et al. [PFC*97] in which the original "Sticky Finger" concept was first applied to distant object selection in virtual environments. Figure 2 illustrates the basic concept and how such a technique could be effectively used e.g. for object query system using a see-through glass. In a typical scenario, the user focuses on and selects an object in the real world (by proprioceptive pointing and designation), which in turn is captured by the on-glass camera and then identified and recognized for final augmentation. Such an object selection method can be used regardless of the coverage of the augmentation screen because it does not require the use of the cursor. Note the "blurry" fingertip is used for aiming to (with the aimed position adjusted according to the offset between the eye and mounted camera in the camera image) and designating the target object (possibly with help of a technique such as lazy snapping [LSTS04]).

The hypothesis is that such an approach would reduce user fatigue since the user will switch one's focus less frequently [YK15], while still being able to point effectively through the proprioceptive sense (despite the blurry finger). In addition, we posit that, compared to the usual indirect cursor based pointing (e.g. using the hand/finger or hand-held tracking device to point and control a cursor for selection), the proposed technique will perform better and even be preferred for its directness.

In the rest of the paper, we demonstrate and validate our claim through an experiment comparing three distant pointing/selection methods: (1) indirect cursor based method using a 3D air mouse, (2) proprioceptive finger aiming with a cursor, (3) proprioceptive finger aiming without a cursor. We begin by discussing other research work related to ours.

Figure 2: *Blurry (Sticky) Finger: proprioceptive pointing and selection, using the finger to point at (and even segment) the target but without focusing on it (thus blurred), relying upon the proprioceptive sense (user view, left). Due to the difference in the positions of the user eye and camera, there is an offset that must be corrected to identify the intended target point in the camera view (right).*

2. Related Work

Remote object selection in the 3D virtual space has been studied extensively [BH97]. While a variety of methods exist, virtual ray casting [Bol80] is one of the most common and popular approaches, in which a 3D (virtual) ray emanates from the hand (whose position is tracked by a sensor), and the user "shoots" for the object of interest. A slight variant is the virtual flashlight in which a volumetric cone is formed with the apex at the hand toward the target object [LG94]. Despite seemingly extensible, these techniques (for 3D virtual space) have rarely been applied to the augmented reality situation, especially for interacting with the real objects [CKY*13, KBCW03]. Heidemann et al. did apply the ray casting technique to AR interaction with real objects, but using the video see-through display in which the whole image was in focus similarly to the case of virtual environments [HBB04].

In most AR applications, only a handful of pre-selected real world objects are recognized, rather than to purposely be newly selected to begin with. To apply the virtual ray casting or flashlight techniques, the user would need to either use a hand tracking device that would require a cumbersome process of calibrating its spatial coordinates with that of the view point, or recognize and track the hand from the head mounted camera or sensor which would restrict the hand position to stay within the sensing volume thereby making the ray/cone casting difficult [HBW11,SKMC99,BMR*12]. Pierce et al. suggested the "Sticky Finger" as one of their image plane interaction techniques similar to the Blurry Finger, for selecting and manipulating an object in virtual environments [PFC*97]. In the case of Blurry Finger, the selection is made over the real environment rather than over the rendered image space where everything including the virtual finger is in clear focus.

Recent availability and popularization of inexpensive environment capture sensors (e.g. depth sensors) have made it viable to model and include real world objects as potential interaction objects even for mobile AR systems [BH97, WKSE11, PCUB12, ZBM94, PR12, LBB12, BW01]. Li et al. [LJ09], Ha et al. [HW06], Mistry and Maes [MM09] and Colaço et al. [CKY*13] all have employed 3D hand gesture based interaction using small depth cameras mounted on the head mounted display (or other parts of the body), but only for interacting with objects within one's arm's reach.

The efficiency of hand-based selection is closely related to the human's hand-eye coordination capability. The proprioceptive sense of the hand/arm (at whose end the interaction object is situated) [SV99] and the visual system work together to resolve for a consistent spatial interpretation and manipulation of the interaction object [HESR00, HASM02, GVdHVG84, VDLD02]. In the case of mouse based cursor control or remote object selection with the cursor, the user would try to resolve the difference between the cursor (rather than the device or hand itself) and the target. Thus the hand-eye coordination is still applied but indirectly for relative movement of the cursor by the hand [SHAZ00, BCF*10]. Due to the offset between the target object/cursor and the hand, the proprioceptive sense becomes less contributing and effective, possibly resulting in degraded performance [BCF05].

Despite such a projected performance problem, hand based pointing of remote objects by the ray casting is still reasonably natural and familiar in the sense it adds on to the traditional remote control (e.g. for TV) very nicely. The air mouse such as the Wii-mote is such an example [BVLG11,LGE,Nin]. However, the use of the remote control can bring about significant fatigue on the wrist, also affecting the interaction performance. With the Blurry Finger, the finger/hand comes within the camera/sensor (usually equipped with an OST glass) view naturally because it is used for aiming with the user's eye (thus the finger can be detected, tracked and used without a hand-held sensing device). Even though the target object is far, it seems as if attached to the finger, which was the very idea of the aforementioned image plane based methods [PFC*97].

3. System Prototype

In order to support the proposed "proprioceptive" pointing, it is necessary to track the finger position in the camera image space and use it to ultimately identify the aimed target object in the line of sight (see Figure 2). Tracking of the finger itself can be accomplished relatively easily employing conventional finger tracking algorithms [LH07, OSK02, LB04]. The problem is that the camera view point is slightly different from that of the user's eye, i.e. what the user aims may not be seen as aimed in the camera view. The offset (*dist*) between the user eye and camera can be used to compute the position of the intended target assuming its depth is known. The formulation is illustrated in Figure 3 and Equations (1) and (2).

$$dist = x_0 + x_1 \quad (1)$$

$$x_0 = \frac{tan(\Theta_0)}{tan(\frac{\varphi_0}{2})} \times \frac{f_0}{2} \quad (2)$$

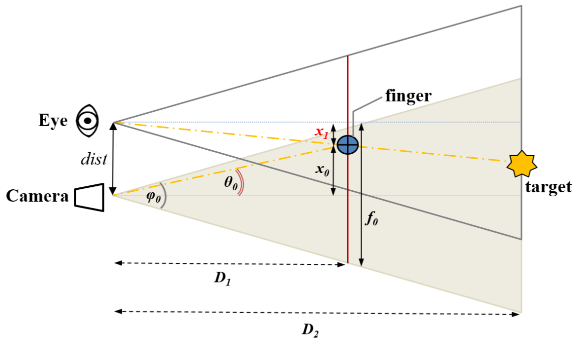


Figure 3: *Computing for the location of the intended target (yellow star) and its projected location (blue cross) with respect to the camera from the finger position and offset of the eye from the camera.*

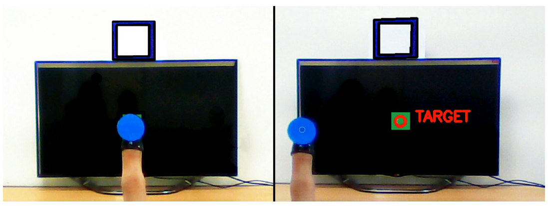
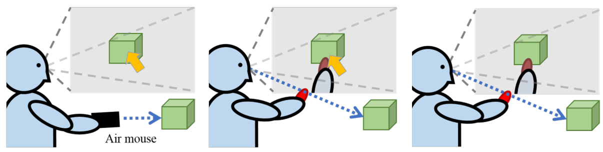


Figure 4: *Views from the user (left) where the finger is coincident with the target (green square). In the camera view (right), the finger is not coincident with the target, but the target is identified and marked with the red circle by the formulation in Figure 3.*

In Figure 3, the target is assumed to be at a fixed depth, D_2, and the finger location with respect to the camera, is found by the finger tracking algorithm, x_0, with its depth, D_1, also estimated from the size of the finger as well. The offset between the camera and the eye, *dist*, and the field of view of the camera, φ_0, can simply be measured or obtained. As for the eye position, we use that of the dominant eye (between the two) for a given user (for more details, see Section 4). The camera is positioned laterally (in one dimension) from the eye as shown in Figure 3. From the figure, it is possible to obtain, the location of the target as seen by the user eye, x_1, and location of the intended target object (yellow star). In turn the lateral location of the target object with respect to the camera is again easily obtained by subtracting this value from *dist*. Then the target object can be back-projected to the camera image space and its image coordinate value computed. Figure 4 shows the test implementation, namely, the views from two cameras and the adjusted proxy position extending into the correct target.

4. Experiment

The main purpose of the experiment was to comparatively demonstrate and validate the potential advantage of the Blurry Finger over the conventional remote cursor based interaction using a hand-held tracking device under different operational conditions. Figure 4 illustrates the difference between the two methods. With the cursor based pointing by the hand-held device ("Hand-Cursor"), the user makes a selection by controlling and moving a cursor (marked in yellow in Figure 5a) over the target object with the hand-held device. With the Blurry Finger ("BF"), the user aims at the object to be in the same line of sight with the finger (marked in red in Figure 5b and 5c). In its original form (Figure 5c), the Blurry Finger shows no cursor. Due to the possible concern that users might not be too familiar with a cursor-less method, we added the comparative case of the Blurry Finger with a cursor ("BF-Cursor") as shown in Figure 5b (the cursor is marked in yellow). Both Hand-Cursor and BF-Cursor use the cursor. Hand-Cursor requires and relies on focusing on the cursor entirely to make a selection, while BF-Cursor uses it only in a supplementary fashion (just to give the user an assurance of the current aim) and the user need not focus on it. Still, a fully overlapping AR glasses would be needed for any cursor based approach, whereas BF can be used for any type of AR glasses.



(a) Cursor based pointing with hand-held device (b) Blurry Finger with cursor (c) Blurry Finger without cursor

Figure 5: *The three compared methods of distant object selection.*

4.1. Experimental Design and Hypotheses

Factor	Levels
Interface Type	Hand-Cursor, Blurry Finger, Blurry Finger-Cursor
Target Object Size	120mm, 60mm, 30mm
Moving Distsance	600mm, 400mm, 200mm

Table 1: *Factors and their levels in the experiment.*

The comparison was carried out through a distant object (at a fixed depth) selection task varied in two dimensions: three object sizes and three inter-object moving distances (in making successive selections). Therefore, the experiment was designed as a 3 factor (3 interfaces × 3 object sizes × 3 moving distances) repeated measure within subject (see Table 1). The main dependent variables were the task completion time, error rate (i.e. total trials – the number of successful task completions), and responses to a general usability survey (including the user preference).

Our main hypotheses were as follows:

- H1-1: Overall, the Blurry Finger and Blurry Finger-Cursor will exhibit better interaction performance than the Hand-Cursor (due to the help of proprioceptive sense and visible finger despite being blurry).
- H1-2: The Blurry Finger and Blurry Finger-Cursor will particularly exhibit better performance as the target object size gets smaller and the moving distance gets longer (i.e. the indirect nature of Hand-Cursor will incur more cognitive cost with the more difficult task).
- H2: Blurry Finger-Cursor will show less error rate than the Blurry Finger (i.e. the help of the cursor will be significant).
- H3: Users will prefer and rate the usability to be high in the order of Blurry Finger-Cursor, Blurry Finger then Hand-Cursor (due to Blurry Finger's directness and less tiring nature).

4.2. Experimental Set-up

Figure 6: *The experimental set up: the air mouse used for Hand-Cursor (left) and the head mounted camera connected to a smart phone for the Blurry Finger / Blurry Finger-Cursor (right).*

The interfaces were tested for selecting objects at a fixed depth (240 cm, a nominal TV viewing distance) that appear successively on a large smart HD TV screen (48 inch, 106 cm × 60 cm). Ideally, the experiment should be conducted over real 3D objects (to reflect actual applications as proposed in Figure 2). Instead, selection of 2D virtual objects on a screen at a fixed depth was used as the experimental task to ensure the efficiency and controllability of the experimental conditions. It is reasonably expected that the experimental results (by the nature of the Blurry Finger method) can later be applied to the selection of real 3D objects over the 2D visual field. The Hand-Cursor used an air mouse called the Logitech MX Air which provided orientation tracking via an internal gyroscope for a screen cursor control at the resolution of 800 dpi [Log].

As for the Blurry Finger, a head mounted USB camera (with an FOV of $40^\circ \times 32^\circ$) connected to a smart phone (LG-F320S G2) running the Android operating system was used (see Figure 6). With the computational power of the Google glass platform it was not possible, with our implementation, to run the finger detection and error adjustment for real time interaction. However, the system was built as it was as a platform for a mobile AR system. The camera was mounted such that it has a lateral offset with the eyes sideways (see Figure 3). The OST glass was not used in the experiment because a fully overlapping type was not available (which would cripple the use of any cursor based approach), and to eliminate any bias from external factors such as its usability, wearability and fatigue. The cursors for the Hand-Cursor and BF-Cursor were displayed on the 2D screen instead.

For an easy detection of the finger, a color marker was worn on the fingertip. For the Blurry Finger-Cursor, a small cursor was displayed at a location where the target was believed to be according to the formulation explained in Section 3. The position of the cursor on the TV display was calculated by the user-end device, then sent to the display side through the Bluetooth.

4.3. Experimental Task and Process

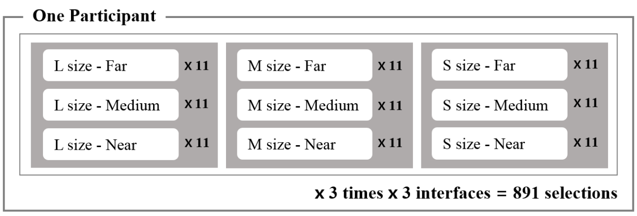
Q1: Ease of Use	Rate how easy you felt the interface to be in accomplishing the given task. (1: Very difficult – 7: Very easy)
Q2: Naturalness	Rate how natural and intuitive you felt the interface to be in accomplishing the given task. (1: Very contrived – 7: Very natural)
Q3: Confidence	Rate how much you were confident or assured that the intended target was being selected (1: Not confident – 7: Very confident)
Q4: Ease of Learning	Rate how easy it was for you to learn the interface in accomplishing the given task. (1: Very difficult to learn – 7: Very easy to learn)
Q5: Future use	Rate the willingness to use the interface again. (1: Not willing – 7: Very willing)
Q6: Fatigue	Rate how fatigued you were after using the interface to accomplish the given task. (1: Not fatigued – 7: Very fatigued)

Table 2: *The six usability-survey questions*

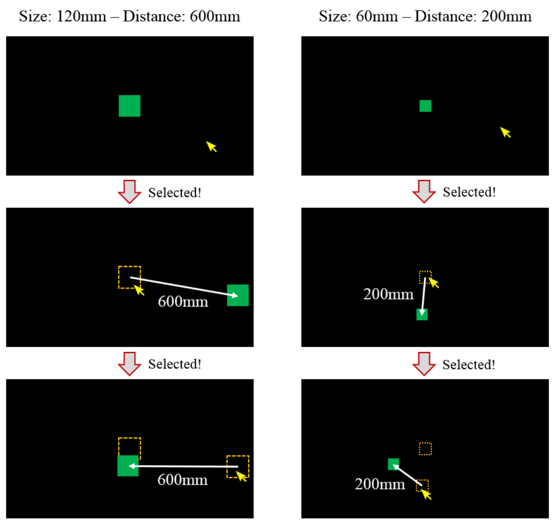
Sixteen paid subjects (12 men and 4 women between the ages of 23 and 34, mean = 26 / one left-handed) participated in the experiment ($10 compensation for their participation). To use the formulation explained in Section 3, it was necessary to measure the lateral offset between the camera and each subject's eye (*dist*). Between the two eyes, the dominant eye was identified using the simple but well-known Miles test [MOB03]. Fourteen subjects had the right eye as the dominant one. For each subject, subject specific parameters were used to compute the amount of error compensation between the eye and the camera.

After collecting their basic background information, the subjects were briefed about the purpose of the experiment and given instructions for the experimental tasks. A short training was given to allow the subjects to become familiarized with the experimental process and three interaction methods. Subjects were instructed not to make focus on the finger when using the BF and the BF-Cursor (which may have been somewhat unnatural, but after the short training, all users became sufficiently comfortable with the methods).

In a single experimental trial block, the subject was asked to use the given interface to make a series of selection (11) of square objects that appeared on the TV. The user was to either aim the target with the air mouse (Hand-Cursor) or with the finger (BF, BF-Cursor) and stay on the aimed object for 1 second to indicate selection. After a successful selection, the next target would appear at some other location on the screen. Three object sizes (120mm, 60mm, 30mm) and three moving distances to the next object (600mm, 400mm, 200mm) were used. Between successive targets, the user was given 3.5 seconds to make a successful selection, otherwise, the task was deemed as a failure and the next target appeared (see Figure 7b). Each block consisted of 11 selections with a particular object size and a moving distance, and three such series were carried out by a single subject for all 27 combinations ($3 \times 3 \times 3$). Thus a total of 891 object selections were made (see Figure 7a). For a subject, the presentation order of the interface type was counter-balanced and similarly for the selection task combination. The task was carried out in a sitting position and at the start of the block, the user was asked to put one's hand/device on one's knee. The time to complete the object selection, and the number of successful completions were measured. After the treatments, the user filled out a general usability survey (answered in 7 Likert scale) as shown in Table 2.



(a)



(b)

Figure 7: *(a) Task sets for one participant, (b) two examples of the task set (the first three of the 11 successive selections with a particular object size and inter-object moving distance).*

5. Results

Figure 8 shows the task completion times among the three interfaces according to the target object size. For the Large (120mm), the Hand-Cursor gave the fastest performance with a statistical significance (ANOVA - $F(2, 4292) = 303.68$, Scheffe test - p-value < 0.012) On the other hand, BF performed better for the Medium (60mm), and both the BF and BF-Cursor performed better for the Small (30mm), and surprisingly their performances across different object sizes stayed mostly unchanged. (BF: $F(2, 4250) = 0.5$, p-value < 0.607, BF-Cursor: $F(2, 4222) = 2.462$, p-value < 0.085).

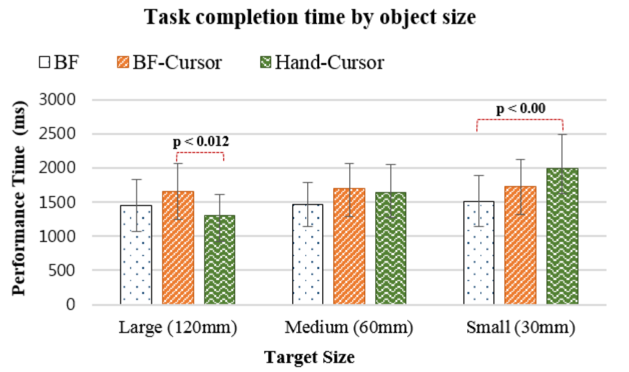


Figure 8: *The task completion times among the three interfaces according to the object size.*

Figure 9 shows the task completion times among the three interfaces according to the moving distance. Obviously, the performance time is generally longer for all interface types with longer moving distances. We had expected that the BF and BF-Cursor would perform better as the distance got larger, but no statistically significant differences were detected. However, we could see that the average amount of change when using the Hand-Cursor was the largest (not statistically significant).

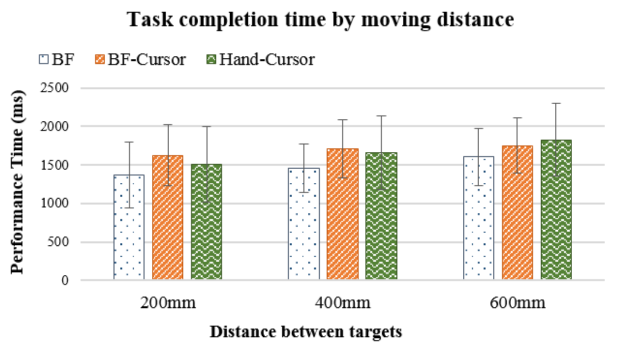


Figure 9: *The task completion times among the three interfaces according to the moving distance.*

Figure 10 shows the time taken for completing the initial object selection. Note that for the first object selection in the task block, the user started with one's hand from the knee position in a sitting posture. We considered that there could exist significant differences in the time spent for the initial object selection. For instance, for the Hand-Cursor, we expected the user to move the device-grasped hand from the knee position and consciously "find" the cursor on the screen, while for the BF (and even BF-cursor) such an overhead would not be present (with the finger acting as the cursor itself). Also note that the statistics in Figure 8 and 9 do not include the time taken for the initial object selection. In Figure 10, we see a similar trend to Figure 8, but for the Hand-Cursor, the performance (initial object selection time) was consistently the worst ($F(2, 141) = 41.097$, Tukey test - p-values < 0.00). In an actual application situation, it is quite possible that the object selection task occurs only infrequently with the users hand switching between the resting and aiming position incurring the observed performance hit. In summary, our hypotheses of H1-1, H1-2 were shown to be valid.

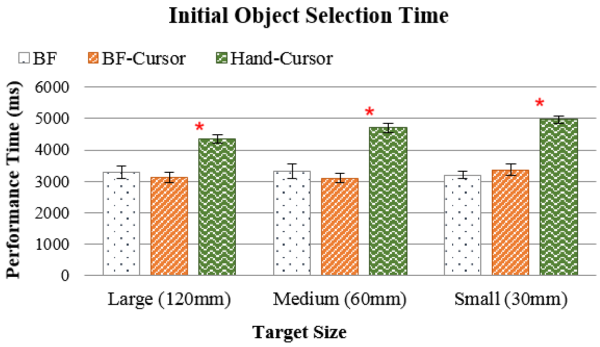


Figure 10: *The initial object selection performance among the three interfaces according to the object size.*

As for the error rate (i.e. total trials – number of successful selections), Table 3 shows that, as expected, more errors occurred with the increasing difficulty of the task, i.e. smaller target object and longer moving distance. However the occurrence of an error was generally very low, e.g. 1~2 times out of total 891 trials for each subject. Nevertheless, statistically significant differences in task completion times were found in the case of the Large (Kruskal-Wallis test, $p < 0.031$) and Medium object size (Kruskal-Wallis test, $p < 0.047$). Namely, the Hand-Cursor had lower error rate for the Large (Mann-Whitney, vs. BF: $p < 0.035$, vs. BF-Cursor: $p < 0.043$) and the Medium (vs. BF-Cursor, $p < 0.026$). No differences were found between the BF and BF-Cursor.

For a given interface type, the Hand-Cursor and BF were affected by the object size, the highest error rate for the Small (Hand-Cursor: with Medium size – $p < 0.008$ / with Large – $p < 0.002$, BF: with Medium size – $p < 0.019$ / with Large – $p < 0.039$). However, BF-Cursor was not affected as such possibly by the synergy between the proprioception and the existence of the cursor. No significant differences were found with regards to the moving distance. Thus, the hypothesis H2 is not rejected.

	Target Size		
	Large	Medium	Small
BF	1.063	1	2.312
BF-Cursor	0.875	1.5	1.688
Hand-Cursor	0.25	0.563	2.75
	Moving Distance		
	200mm	400mm	600mm
BF	1.313	0.938	2.125
BF-Cursor	0.75	1.5	1.8125
Hand-Cursor	1.0	1.25	1.313

Table 3: *Number of error occurrences (out of the total 891) according to the target object size and moving distance.*

Finally, Figure 11 shows the collective responses to the usability survey. To summarize the results:

- Ease of use and the level of fatigue were perceived at similar levels, unexpectedly.
- BF was evaluated to be the most natural, most attractive for future reuse (Kruskal-Wallis test, $p < 0.016$), and easy to learn.
- Hand-Cursor and BF-Cursor both gave high assurance to the user about their task completion (Kruskal-Wallis test, $p < 0.031$).

In the post briefing, subjects expressed their preference in the order of BF (13), BF-Cursor (2), and then the Hand-Cursor (1). Blurry Finger was generally considered the most intuitive and natural. Blurry Finger-Cursor was received generally negatively being distracting and even confusing sometimes. But at the same time, subjects expressed the lack of confidence in their task completion with (explicit) cursor-less method, even though with the Blurry Finger, the finger effectively served the purpose of the cursor. Hand-Cursor as was shown quantitatively was felt to be the most difficult method for interacting with small objects due to its indirectness and multi-focus efforts. Thus, we conclude that H-3 was accepted.

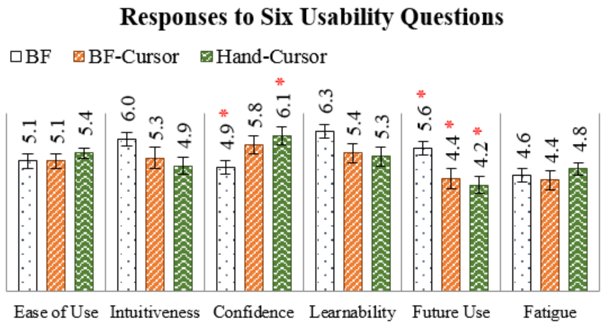


Figure 11: *Responses to the usability survey. Statistical differences by the Mann-Whitney test are marked with the asterisks.*

6. Application

Figure 12: *The Blurry Finger based AR based image searching system: (top) system set up with the OST display and (below) a scene from actual usage. The result on the OST display was also shown on the monitor*

In contrast to the limited (i.e. selecting of virtual objects at a fixed depth), to fully demonstrate the advantage of the Blurry Finger, we developed a prototype AR based image searching system similar to the one proposed in Figure 2. This prototype, shown in Figure 12, used the Blurry Finger to allow the user to encircle and designate a real object, segment out the image from the camera using the adjusted aims, match it against the database (Google Cloud Vision), and display the result on an OST display (Liteye LE-500).

Note that the formulation in Section 3 assumes that the distance to the target object is known (or obtained by a sensor). Likewise, the experiment was conducted at a known fixed depth without a depth sensor. In actual usage, even without a depth sensor, we can expect the user to designate an object with enough "room" to sufficiently compensate for the error from the camera-eye offset using a nominal default depth value. Figure 13 illustrates a user designating a can at different depth yet using a nominal depth value, and still being able to bring about correct matches.

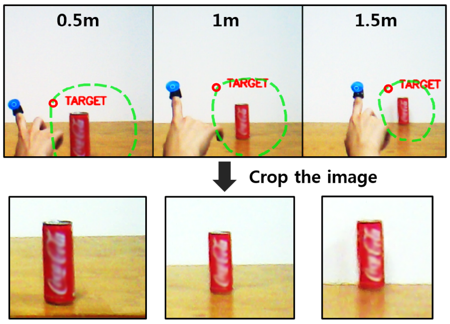


Figure 13: *The user designates the object tolerant enough to select the target that one's wanted to choose. The depth of the target object almost doesn't affect the accuracy of selection.*

7. Conclusion

AR applications are becoming more and more popular and interactive especially in the form of OST glasses due to its non-isolating nature. The Blurry Finger provides an intuitive and natural means to interact with the objects for both virtual and real environment in a seamless way. The suggested technique relies on one's innate proprioceptive sense and hand-eye coordination, and can reduce the eye stress due to the multi-focus problem in OST glasses. In addition, being a cursor-less approach, it can be used for any types of AR glasses (with a camera). Even though the aiming fingertip was perceived to be blurred due to binocularity, the correct object can still be selected without much error because our perceptual systems internally uses just one dominant eye. The evaluation experiment also showed promising results with the Blurry Finger exhibiting generally higher and stable interaction performance and usability compared to the conventional 3D mouse based cursor control method. Even though the prototype in the experiment was configured in a limited way (for selecting 2D objects at a fixed depth plane), its usability was demonstrated through another actual application, an AR image search system. We believe that the Blurry Finger approach is viable step toward realizing a more natural interaction for OST based AR applications in the future.

Mixed results were obtained with regards to the provision of a cursor. The cursor seemed to help the user make finer selection and give high confidence, but also perceived as distracting and obscuring when used with the Blurry Finger because the finger also shares the role of the cursor. More experiments will be needed to test the Blurry Finger under a wider range of operating conditions including with different interaction depths, varying the forms of the objects and cursor (e.g. closer coupling with the fingertip to alleviate the distraction problem) and other interactive tasks such as dragging and combining with (finger) gestures. Despite showing good performance for interacting with relatively small objects, the resolution up to which Blurry Finger can remain efficient (despite the finger being seen blurry) must be further investigated. We also plan to develop and test actual interactive OST based AR applications using the Blurry Finger, IoT (Internet of Things) appliance control systems, etc. For practicality, the Blurry Finger will also need to overcome the operational problems such as the cumbersome process of user customization (e.g. having to measure the eye to camera distance) and instability of the mounted camera.

Acknowledgements

This work was supported by Institute for Information & communications Technology Promotion(IITP) grant funded by the Korea government(MSIP) (No.R0190-16-2011,Development of Vulnerability Discovery Technologies for IoT Software Security) and the Forensic Research Program of the National Forensic Service (NFS), Ministry of Government Administration and Home Affairs, Korea. (NFS-2015-DIGITAL-04)

References

[BCF05] BERNIER P.-M., CHUA R., FRANKS I. M.: Is proprioception calibrated during visually guided movements? *Experimental Brain Research 167*, 2 (July 2005), 292–296. doi:10.1007/s00221-005-0063-5. 2

[BCF*10] BIEG H.-J., CHUANG L. L., FLEMING R. W., REITERER H., BÜLTHOFF H. H.: Eye and pointer coordination in search and selection tasks. In *Proc. the Symposium on Eye-Tracking Research & Applications* (2010), pp. 89–92. doi:10.1145/1743666.1743688. 2

[BH97] BOWMAN D. A., HODGES L. F.: An evaluation of techniques for grabbing and manipulating remote objects in immersive virtual environments. In *Proc. the 1997 symposium on Interactive 3D graphics* (1997), pp. 35–ff. doi:10.1145/253284.253301. 2

[BMR*12] BAILLY G., MÜLLER J., ROHS M., WIGDOR D., KRATZ S.: Shoesense: a new perspective on gestural interaction and wearable applications. In *Proc. SIGCHI Conference on Human Factors in Computing Systems* (2012), pp. 1239–1248. doi:10.1145/2207676.2208576. 2

[Bol80] BOLT R. A.: *"Put-that-there": Voice and gesture at the graphics interface*. ACM, 1980. 2

[BVBC04] BUCHMANN V., VIOLICH S., BILLINGHURST M., COCKBURN A.: Fingartips: gesture based direct manipulation in augmented reality. In *Proc. the 2nd international conference on Computer graphics and interactive techniques in Australasia and South East Asia* (2004), pp. 212–221. doi:10.1145/988834.988871. 1

[BVLG11] BAILLY G., VO D.-B., LECOLINET E., GUIARD Y.: Gesture-aware remote controls: guidelines and interaction technique. In *Proc. the 13th international conference on multimodal interfaces* (2011), pp. 263–270. doi:10.1145/2070481.2070530. 2

[BW01] BOWMAN D. A., WINGRAVE C. A.: Design and evaluation of menu systems for immersive virtual environments. In *Virtual Reality, 2001. Proceedings. IEEE* (2001), pp. 149–156. doi:10.1109/VR.2001.913781. 2

[CDDGC13] CORBETT-DAVIES S., DUNSER A., GREEN R., CLARK A.: An advanced interaction framework for augmented reality based exposure treatment. In *2013 IEEE Virtual Reality (VR)* (2013), pp. 19–22. doi:10.1109/VR.2013.6549351. 1

[CKY*13] COLAÇO A., KIRMANI A., YANG H. S., GONG N.-W., SCHMANDT C., GOYAL V. K.: Mime: compact, low power 3d gesture sensing for interaction with head mounted displays. In *Proc. the 26th annual ACM symposium on User interface software and technology* (2013), pp. 227–236. doi:10.1145/2501988.2502042. 2

[Eps] EPSON:. Epson moverio. [online]. URL: http://www.epson.com/moverio. 1

[Goo] GOOGLE:. Google glass [online]. URL: https://developers.google.com/glass/. 1

[GVdHVG84] GIELEN C., VAN DEN HEUVEL P., VAN GISBERGEN J.: Coordination of fast eye and arm movements in a tracking task. *Experimental Brain Research 56*, 1 (Apr. 1984), 154–161. doi:10.1007/BF00237452. 2

[HASM02] HAYHOE M., AIVAR P., SHRIVASTAVAH A., MRUCZEK R.: Visual short-term memory and motor planning. *Progress in brain research 140* (2002), 349–363. doi:10.1016/S0079-6123(02)40062-3. 2

[HBB04] HEIDEMANN G., BAX I., BEKEL H.: Multimodal interaction in an augmented reality scenario. In *Proc. the 6th international conference on Multimodal interfaces* (2004), pp. 53–60. doi:10.1145/1027933.1027944. 2

[HBW11] HARRISON C., BENKO H., WILSON A. D.: Omnitouch: wearable multitouch interaction everywhere. In *Proc. the 24th annual ACM symposium on User interface software and technology* (2011), pp. 441–450. doi:10.1145/2047196.2047255. 1, 2

[HESR00] HELSEN W. F., ELLIOTT D., STARKES J. L., RICKER K. L.: Coupling of eye, finger, elbow, and shoulder movements during manual aiming. *Journal of motor behavior 32*, 3 (June 2000), 241–248. doi:10.1080/00222890009601375. 2

[HW06] HA T., WOO W.: Bare hand interface for interaction in the video see-through hmd based wearable ar environment. In *International Conference on Entertainment Computing* (2006), pp. 354–357. doi:10.1007/11872320_48. 2

[KBCW03] KIYOKAWA K., BILLINGHURST M., CAMPBELL B., WOODS E.: An occlusion-capable optical see-through head mount display for supporting co-located collaboration. In *Proc. the 2nd IEEE/ACM International Symposium on Mixed and Augmented Reality* (2003), p. 133. doi:10.1109/ISMAR.2003.1240696. 2

[LB04] LETESSIER J., BÉRARD F.: Visual tracking of bare fingers for interactive surfaces. In *Proc. the 17th UIST* (2004), pp. 119–122. doi:10.1145/1029632.1029652. 3

[LBB12] LEE G. A., BAI H., BILLINGHURST M.: Automatic zooming interface for tangible augmented reality applications. In *Proc. the 11th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and its Applications in Industry* (2012), pp. 9–12. doi:10.1145/2407516.2407518. 2

[Lea] LEAPMOTION:. Leap motion controller [online]. URL: https://www.leapmotion.com. 1

[LG94] LIANG J., GREEN M.: Jdcad: A highly interactive 3d modeling system. *Computers & graphics 18*, 4 (July 1994), 499–506. doi:10.1016/0097-8493(94)90062-0. 2

[LGB08] LEE M., GREEN R., BILLINGHURST M.: 3d natural hand interaction for ar applications. In *2008 23rd International Conference Image and Vision Computing New Zealand* (2008), pp. 1–6. doi:10.1109/IVCNZ.2008.4762125. 1

[LGE] LGELECTRONICS:. Lg magic remote [online]. URL: http://www.lg.com/us/magic-remote. 2

[LH07] LEE T., HOLLERER T.: Handy ar: Markerless inspection of augmented reality objects using fingertip tracking. In *2007 11th IEEE International Symposium on Wearable Computers* (2007), pp. 83–90. doi:10.1109/ISWC.2007.4373785. 3

[LJ09] LI Z., JARVIS R.: Real time hand gesture recognition using a range camera. In *Australasian Conference on Robotics and Automation* (2009), pp. 21–27. 2

[Log] LOGITECH:. Mx air rechargeable cordless air mouse [online]. URL: http://support.logitech.com/en_us/product/mx-air-rechargeable-cordless-air-mouse. 4

[LSTS04] LI Y., SUN J., TANG C.-K., SHUM H.-Y.: Lazy snapping. In *ACM Transactions on Graphics (ToG)* (2004), vol. 23, pp. 303–308. doi:10.1145/1015706.1015719. 2

[MM09] MISTRY P., MAES P.: Sixthsense: a wearable gestural interface. In *ACM SIGGRAPH ASIA 2009 Sketches* (2009), p. 11. doi:10.1145/1667146.1667160. 2

[MOB03] MAPP A. P., ONO H., BARBEITO R.: What does the dominant eye dominate? a brief and somewhat contentious review. *Perception & Psychophysics 65*, 2 (Feb. 2003), 310–317. doi:10.3758/BF03194802. 4

[Nin] NINTENDO:. Wii remote [online]. URL: http://www.nintendo.com/wiiu/accessories. 2

[OKA11] OIKONOMIDIS I., KYRIAZIS N., ARGYROS A. A.: Efficient model-based 3d tracking of hand articulations using kinect. In *Proc. the British Machine Vision Conference* (2011), vol. 1, pp. 101.1–101.11. doi:10.5244/C.25.101. 1

[OSK02] OKA K., SATO Y., KOIKE H.: Real-time fingertip tracking and gesture recognition. *IEEE Computer graphics and Applications 22*, 6 (Dec. 2002), 64–71. doi:10.1109/MCG.2002.1046630. 3

[PCB11] PIUMSOMBOON T., CLARK A., BILLINGHURST M.: Physically-based interaction for tabletop augmented reality using a depth-sensing camera for environment mapping. In *Proc. the 26th International Conference on Image and Vision Computing New Zealand* (2011), pp. 161–166. 1

[PCUB12] PIUMSOMBOON T., CLARK A., UMAKATSU A., BILLINGHURST M.: Poster: Physically-based natural hand and tangible ar interaction for face-to-face collaboration on a tabletop. In *3DUI, IEEE Symposium on* (2012), pp. 155–156. doi:10.1109/3DUI.2012.6184208. 2

[PFC*97] PIERCE J. S., FORSBERG A. S., CONWAY M. J., HONG S., ZELEZNIK R. C., MINE M. R.: Image plane interaction techniques in 3d immersive environments. In *Proc. the symposium on Interactive 3D graphics* (1997), pp. 39–ff. doi:10.1145/253284.253303. 2, 3

[PR12] PRISACARIU V. A., REID I.: 3d hand tracking for human computer interaction. *Image and Vision Computing 30*, 3 (Mar. 2012), 236–250. doi:10.1016/j.imavis.2012.01.003. 2

[SHAZ00] SMITH B. A., HO J., ARK W., ZHAI S.: Hand eye coordination patterns in target selection. In *Proc. the 2000 symposium on Eye tracking research & applications* (2000), pp. 117–122. doi:10.1145/355017.355041. 2

[SKMC99] SASAKI H., KURODA T., MANABE Y., CHIHARA K.: Hit-wear: A menu system superimposing on a human hand for wearable computers. In *Proc. ICAT 99* (1999), pp. 146–153. 2

[SL13] SEO D. W., LEE J. Y.: Direct hand touchable interactions in augmented reality environments for natural and intuitive user experiences. *Expert Systems with Applications 40*, 9 (July 2013), 3784–3793. doi:10.1016/j.eswa.2012.12.091. 1

[Sof] SOFTKINETIC:. Softkinetic depthsense [online]. URL: http://www.softkinetic.com. 1

[SV99] SCARCHILLI K., VERCHER J.-L.: The oculomanual coordination control center takes into account the mechanical properties of the arm. *Experimental brain research 124*, 1 (Jan. 1999), 42–52. doi:10.1007/s002210050598. 2

[VDLD02] VAN DONKELAAR P., LEE J.-H., DREW A. S.: Cortical frames of reference for eye-hand coordination. *Progress in brain research 140* (2002), 301–310. doi:10.1016/S0079-6123(02)40058-1. 2

[WKSE11] WACHS J. P., KÖLSCH M., STERN H., EDAN Y.: Vision-based hand-gesture applications. *Communications of the ACM 54*, 2 (Feb. 2011), 60–71. doi:10.1145/1897816.1897838. 2

[YK15] YU J., KIM G. J.: Eye strain from switching focus in optical see-through displays. In *Human-Computer Interaction* (2015), pp. 550–554. doi:10.1007/978-3-319-22723-8_59. 2

[ZBM94] ZHAI S., BUXTON W., MILGRAM P.: The "silk cursor": investigating transparency for 3d target acquisition. In *Proc. the SIGCHI Conference on Human Factors in Computing Systems* (1994), pp. 459–464. doi:10.1145/191666.191822. 2