Wands are Magic: A comparison of 3D pointing interfaces

5 authors, including:

Martin Henschke
Australian National University
8 PUBLICATIONS **22** CITATIONS

SEE PROFILE

Richard Jones
Australian National University
23 PUBLICATIONS **142** CITATIONS

SEE PROFILE

Tom Gedeon
Australian National University
430 PUBLICATIONS **5,943** CITATIONS

SEE PROFILE

Sabrina Caldwell
Australian National University
27 PUBLICATIONS **36** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Model identification and rule base extraction View project

Emotion recognition in the wild View project

Wands Are Magic: A Comparison of Devices Used in 3D Pointing Interfaces

Martin Henschke, Tom Gedeon, Richard Jones, Sabrina Caldwell, and Dingyun Zhu

College of Engineering and Computer Science,
Australian National University, Acton ACT, Australia
{martin.henschke,tom.gedeon,richard.jones,sabrina.caldwell,
dingyun.zhu}@anu.edu.au

Abstract. In our pilot study with 12 participants, we compared three interfaces, 3D mouse, glove and wand in a 3D naturalistic environment. The latter two were controlled by the same absolute pointing method and so are essentially identical except for the selection mechanism, grasp action versus button. We found that the mouse performed worst in terms of both time and errors which is reasonable for a relative pointing device in an absolute pointing setting, with the wand both outperforming and favored by users to the glove. We conclude that the presence of a held object in a pointing interface changes the user's perception of the system and magically leads to a different experience.

Keywords: "magic wand", "3D mouse", "hand gesture", "fatigue", "user satisfaction".

1   Introduction

Tasks that are typically accomplished by human beings can be separated into two broad categories: tasks achieved through physical manipulation of objects, such as chopping wood with an axe or writing with a pen and paper, and communication with other people, commonly done both using words and a series of communicative gestures. When not operating physical tools, gestures are usually a form of communication and often intertwined with spoken language[3]. Although operating computers can be considered a form of communication, typical day-to-day interactions with a standard desktop or laptop computer fall into the physical object category. Interfaces that use motion or body capture allow users to perform these sorts of tasks without requiring a device to be held, with interfaces implementing gestures or methods of interaction found in both categories. The challenge of making these interfaces sufficiently ubiquitous remains formidable, and in the meantime it remains unclear if such interfaces are superior or preferred by users to an object-based interface, in a comparable setting. The ergonomic factors relating to these devices, namely how much physical strain or fatigue they cause with continued use, remain an issue in interface design[9], and merit further study.

Our pilot study used a series of pointing interaction methods in a 3D interface, comparing operating an interface while holding a physical object to performing interactions in the absence of one. We wish to determine if there are any changes in the mode of operation, the severity and locality of discomfort caused by their operation and the preferences of the users.

2   Background

The MIT Media Room[1] is a seminal example of an object-free interface, in which users would use a combination of spoken language and pointing gestures at a large projected display to create and modify the state of various shapes in a scene. A similar system, using a 3D spatial environment and gesture interface captured interactions with a specially designed glove and camera system[8]. Kjeldsen and Kender (1996) presented a camera-based pointing system applied directly to performing mouse tasks on a standard desktop computer[7]. Another similar system provided a means of absolute pointing through capture of the user's entire arm and projecting a line from the user's shoulder to their fingertip[5]. The system used multiple cameras to make any form of device such as a glove or sensor unnecessary.

Research into the ergonomics or feasibility of such systems remains limited, though study into developing ergonomic gestures has been conducted[9]. An assessment of performance of gesture interfaces in VR environments reported the interfaces were almost four times slower than using a traditional mouse system, and fatigue was a common complaint[2]. A study comparing Wii remote and Kinect interfaces found a hands free interface performed tasks faster and was preferred by users to the Wii interface, which used in-built accelerometers rather than arm movements for tracking[4]. Another study, looking specifically at the ergonomic considerations of interfaces with and without physical objects indicated that performing a task with a 'virtual object' was more difficult and fatigue-inducing than an equivalent one performed with a real object. The trial found users extended their fingers earlier and farther with the virtual object, making it the more fatiguing of the two interactions[6].

3   The System

We have developed a 3D pointing interface, in which users are able to select, grab and drag objects in a virtual space by pointing and performing interactions with the given device. The interface makes use of an absolute pointing system, in which both the user's position in the work area and the screen's position and dimensions are used to determine where user is pointing. The system works by having the user physically point at the display surface with their hand. From here, capturing user input involves detecting the position of the user's forearm and hand to construct a 3D vector, which is then projected forward to the 2D plane on which both the Kinect and the display surface rest. This provides a 2D point on the plane where the user is pointing at any given point in time. The system allows for a relatively flexible definition of pointing, with different arm orientations and holding the device in different ways possible.

Our pilot makes use of a total of three different interaction methods. The first is a wired glove interface as an approximation of operating a user interface without physically manipulating a device. It operates using an Essential Realities P5 Wired Glove, weighing 82g for tracing finger movements and a Microsoft Kinect that tracks general body movements. The second is a Wand interface, operating with Kinect in the same way as the Wired Glove interface, but with a hand-held controller in place of a glove. The controller is from a Nintendo Wii, and is identical to a standard Wii Remote controller, except that it has a slimmer, longer and more cylindrical profile, weighing 78g. The controller is only used for button presses as position information is acquired from the Kinect. An OnmiMotion Air Mouse, weighing 65g, is used for our third device, using a relative pointing system in contrast to the other two devices. This gyroscopic mouse is similar to a standard desktop mouse in appearance, but moves the cursor through rotating the body of the mouse rather than laser tracking.

4   Experiment

The general research question of this paper is *does the absence of a physical object of manipulation impact how the user interacts with an interface?* We define a physical object in this context as one that is held rather than worn, and is manipulated in-hand. We are looking at three particular aspects of this question: user performance, induced fatigue and preference.

A 3D visualization was conducted for the experimental trials, designed to have the appearance of being performed in a Gothic château, with the tasks themselves resembling ones being accomplished by casting magic spells. This allowed users to treat each device as operating according to their own expectations, and adopt an approach that felt contextually natural in operating the devices. A total of three separate kinds of tasks were performed:

- A selection task, in which the user was required to select a ghost on the screen by pointing at it and performing a selection. For the glove, this was achieved by quickly tapping the index finger down and back up again, and for the wand this was done by pressing the 'A' button on the shaft of the device. For the mouse, the left mouse button was used. The position and size of the targets was randomized. Nine selections were performed per trial.
- A select-drag task, in which the user was required to select a key, randomly placed near the bottom of the display, then 'grab' the object and drag it over a lock. In the case of the glove, the grab was performed by forming a fist, and in the case of the wand it was by pulling a trigger beneath the device. Six drag tasks were performed per trial.
- A select-drag-drop task, in which the user was required to select a firefly and drag it to a cage, then release it. The target area remained the same for each task, but the positions of the selectable objects were randomized. Six of these tasks were performed per trial.

Fig. 1. Selecting a firefly in the select-drag-drop trial using the P5 Wired Glove

The trials presented the tasks consecutively with a brief break between each task explaining what needed to be done before it was begun. At the end of each sequence of tasks, the participant was given a 5-10 minute break to rest their arm before continuing with the next interface device.

During the task completion phase of the experiments, the mouse cursor is hidden from participants, as our trial is designed for a naturalistic environment without relative location cues, and could be used in a natural environment. Selection gestures are accompanied with a small burst of stars at the area pointed to, so users know where their gesture was directed.

Over the course of the trials, the system keeps track of every button press and gesture, the position of the pointing location and a time signature. Video from the Kinect was captured, to observe the movements each user makes when operating each interface. At the end of each trial, the user was also queried on the discomfort felt in their arms, and asked where specifically the discomfort, if any, was located and how severe it was on a scale of 0 to 10. At the end of each trial, the user was also asked to fill out a short questionnaire about the device they had just used, using a combination of Likert-scale questions and short written answers.

4.1 Results

The trial was run with 12 participants, 8 male, and 4 female, between the ages of 21 and 36, with a mean age of 25. All participants were computer science students, and reported regular computer use. Of those, 9 reported familiarity with the Kinect. Other gesture capture technologies such as the Nintendo Wii were also familiar to participants, with 1 participant having never used any motion capturing devices.

The relationship between the number of incorrect selections made and the time taken to complete the trial demonstrates a moderate to strong trend for most users, so these values are used as a measure of how accurately and easily each participant was able to use each device.

Table 1. Mean and Standard Deviations for time and errors in completing selection tasks for each device (time in seconds)

	Glove		Mouse		Wand	
	Time	Errors	Time	Errors	Time	Errors
μ	136.8s	50.3	149.0s	71.5	79.7s	45.8
σ	62.1	18.4	81.3	33.8	33.8	20.7

Performance between the glove interface and wand were roughly equivalent in terms of accuracy, but users on average took substantially longer in performing operations with the glove. The mouse was on average the poorest performing device, but also the device with the greatest level of variance. The wand can be seen here as having generally the best and most consistent performance, though all devices are demonstrated as being able to perform quickly and with relatively few errors.

During the trial, a variety of arm positions were adopted. The method of control and orientation of the arm was consistent enough between participants to be separated into three categories, ordered from lowest induced fatigue to highest:

- The shoulder at rest, with the elbow kept at the side and the forearm extended pointing at the display. Holding the device in this manner, users typically made wrist and small forearm movements to control the interface.
- The shoulder partially at rest, the elbow out from the side and bent, with the forearm raised and the wrist bent to face the device to the display. Forearm rotation and wrist movements were the primary method of interaction.
- The arm fully extended with the elbow locked. The user moves the entire arm at the shoulder in this position.

While users typically kept a consistent arm orientation for the trials, it was not uncommon for them to change positions. This was most commonly moving from an outstretched arm to a relaxed or partially relaxed arm, likely to combat fatigue, or from a relaxed or partially relaxed arm to an outstretched arm, usually due to issues with accuracy. Figure 3 shows the prevalence of various orientations with each device, by counting the number of times and the duration of the trial during which the stance appeared for all participants.

On average, the glove interface was reported to be the most fatiguing by participants ($\mu = 4$, $\sigma = 1.9$), followed by the wand ($\mu = 2.4$, $\sigma = 2.2$) with the mouse being the least stressful to use ($\mu = 1.3$, $\sigma = 2.5$).

Fig. 2. From left to right: shoulder at rest, partially extended forearm, fully extended arm
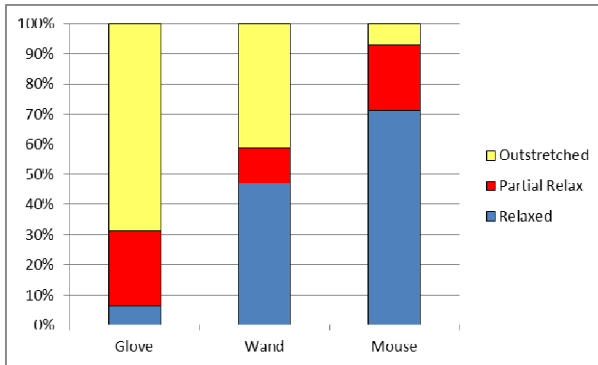
Fig. 3. Prevalence of different arm positions for all participants using the interface

Immediate fatigue induced from use, and the extent of continued fatigue from use for each device was also reported in the questionnaire. The results are shown in Table 2, with no significant difference between the wand and mouse ($p > 0.05$), but more pronounced variation between the glove and all other devices. This confirms the glove as being the most fatigue-inducing device to operate.

Table 2. Distribution-free analysis of questionnaire results pertaining to immediate and continually induced fatige from each device

	Immediate Fatigue	*Continuous Fatigue*
Non-parametric test	$\chi^2 = 7.023$, p=0.03	$\chi^2 = 11.286$, p=0.004
Pair-wise Glove vs. Wand	p = 0.008	p = 0.015
Pair-wise Glove vs. Mouse	p = 0.04	p = 0.005
Pair-wise Wand vs. Mouse	p = 0.55	p = 0.232

In the questionnaire, users were asked to grade each interface system on a seven-level Likert scale, how natural, intuitive, learnable, reactive, accurate and generally easy to use the interface was perceived to be. Of those, only two yielded a statistically significant result in non-parametric analysis. In both instances, pair-wise analysis revealed no significant difference between the glove and mouse. The wand however was found to be more intuitive than the glove ($p=0.016$) and the mouse ($p=0.018$). The wand was also found to be easier and quicker to learn than the glove ($p=0.024$). The fact that all other results lacked a clear correlation suggested personal preference played a large part in these reports.

Users were also asked to rank each interface in terms of preference, and provide comments as to those rankings. The results of this are shown in Figure 3, indicating a general preference for the wand interface. Comments justifying this preference refer-enced the wands appearance or comfortable grip or being easier to press buttons than on the mouse or with finger gestures with the glove. When the mouse was preferred over the wand, it was typically due to being the less strenuous to operate, with

complaints of the mouse being its unpredictable behavior and difficulty in reaching buttons. Preferences for the glove were reported as being the more natural of the interfaces, while complaints included a poor responsiveness and fatiguing to use.
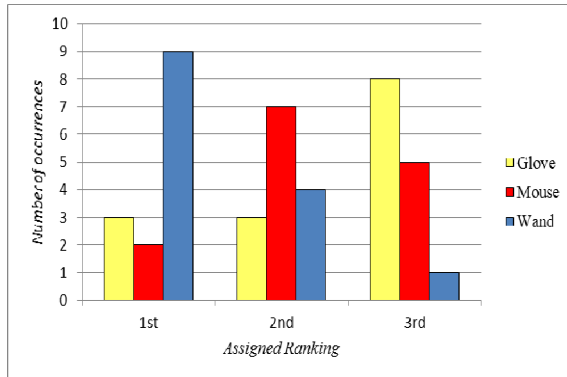

Fig. 4. Sum of rankings given by each user to the three interface devices

5   Discussion and Future Work

Our expectations were that users would find differences in induced fatigue and preference between the devices, but pronounced differences in performance and speed were not expected. The results suggest the presence of an object held in the user's hand has a profound impact on how users view the system, and the amount of effort users expended.

The results from observations suggest that interfaces with a physical object held in hand tended to encourage users to use a less direct form of pointing, performing pointing with the upper arm fully or partially at rest, and the elbow at the side, while having no object encouraged users to keep their arms fully extended. This fits the natural analogue of attempting to point to an object with your hand, in which the user raises it and looks down their arm towards their index finger, where the wand was interpreted much more generally and had more varied usage patterns. A generally larger number of changes in posture were seen with wand users; they would find configurations they liked rather than what they felt was necessary. The wand was observed to also be subject to in-hand manipulation; users were observed turning or rotating the device within the palm of their hand rather than performing the equivalent rotation with their wrist or forearm. This rather counter-intuitively seems to be accompanied with a better performance index for the wand. Less physical movement by the user is required overall to move the pointer to the desired area, but as the region of selection is far smaller, inaccuracy was expected to be higher. This may be attributed to the fact that selection and grabbing was reported to be much easier with the wand and mouse when compared to the glove. Users on average seemed to exert more energy attempting to perform selections and grabs, as the glove required a strong, deliberate motion to select an object on screen compared to the other two. That a

strong correlation can be seen between increasingly extended arm, trial length and fatigue matches expectations, and remains an important consideration in interface design.

The poor results for the mouse can largely be attributed to the absence of a visible cursor and the use of rotation as opposed to lateral motion as with a desktop mouse. Overshooting targets and quickly clicking to inform the user where the cursor was occurred regularly in the trials.

The variation in user preference may reflect the background of participants. Comments such as 'familiarity with the Wii may have been useful' appeared in the questionnaire, despite the fact the system functions quite differently to the Wii, but the method in which users operated the interface may have reflected that experience. Preferences for any given device appears to be heavily influenced by those users found to perform most effectively, though there were some instances where this was supplanted by familiarity with that device, particularly with the mouse.

This trial has indicated substantial nuances in the problem that merit further investigation. Controlling and examining each of the findings in this paper in more detail will be the focus of later experiments. Running trials for extended periods of time and encompassing more complex interactions will be performed in future work.

References

1. Bolt, R.A.: "Put-that-there": Voice and gesture at the graphics interface. SIGGRAPH Comput. Graph. 14, 262–270 (1980)
2. Cabral, M.C., Morimoto, C.H., Zuffo, M.K.: On the usability of gesture interfaces in virtual reality environments. In: Proc. of the 2005 Latin American Conference on Human-Computer Interaction, pp. 100–108. ACM, Cuernavaca (2005)
3. Cassell, J.: A Framework For Gesture Generation and Interpretation. Computer Vision and Machine Interaction, pp. 191–125 (1998)
4. Francese, R., Passero, I., Tortora, G.: Wiimote and Kinect: gestural user interfaces add a natural third dimension to HCI. In: Proc. of the International Working Conference on Advanced Visual Interfaces, pp. 116–123. ACM, Capri (2012)
5. Fukumoto, M., Suenaga, Y., Mase, K.: "Finger-Pointer": Pointing interface by image processing. Computers & Graphics 18, 633–642 (1994)
6. Kim, Y., Lee, G.A., Jo, D., Yang, U., Kim, G., Park, J.: Analysis on virtual interaction-induced fatigue and difficulty in manipulation for interactive 3D gaming console. In: Consumer Electronics (ICCE), pp. 269–270 (2011)
7. Kjeldsen, R., Kender, J.: Toward the use of gesture in traditional user interfaces. In: Automatic Face and Gesture Recognition, pp. 151–156 (1996)
8. Maggioni, C.: A novel gestural input device for virtual reality. In: Virtual Reality Annual International Symposium, pp. 118–124 (1993)
9. Nielsen, M., Störring, M., Moeslund, T.B., Granum, E.: A Procedure for Developing Intuitive and Ergonomic Gesture Interfaces for HCI. In: Camurri, A., Volpe, G. (eds.) GW 2003. LNCS (LNAI), vol. 2915, pp. 409–420. Springer, Heidelberg (2004)