A Gaze-Controlled Interface to Virtual Reality Applications for Motor- and Speech-Impaired Users

Wei Ding[1], Ping Chen[2], Hisham Al-Mubaid[3], and Marc Pomplun[1]

[1] University of Massachusetts Boston
[2] University of Houston-Downtown
[3] University of Houston-Clear Lake
{wei.ding,marc.pomplun}@umb.edu, chenp@uhd.edu, hisham@uhcl.edu

Abstract. This project aims to overcome the access barriers to virtual worlds for motor- and speech-impaired users by building a gaze-controlled interface for Second Life that will enable them to interact with the virtual world by just moving their eyes. We have conducted a study to assess (1) the facilitation of gaze-controlled text input using word prediction technique to speed up chat-style text input in virtual worlds, (2) the influence of screen layout on the efficiency of text input, (3) the effect of the maximum number of suggested words on typing efficiency, and (4) the performance of non-disabled vs. motor-impaired users. Non-disabled subjects and Amyotrophic Lateral Sclerosis (ALS) patients have participated in our experiment. Experimental results show that on average the patients took less time and fewer corrections per letter than did the non-disabled subjects. This finding suggests that our interface design is suitable for motor-impaired users.

Keywords: Eyegaze-controlled input, word prediction, interface agent, virtual world, accessibility.

1 Introduction

A virtual world is a computer-simulated online environment, and its users create avatars to inhabit there and interact with each other. Current virtual worlds (e.g., Second Life, World of Warcraft) often have thousands of concurrent online users and depict a world very similar to the real world. Because the multifaceted nature of virtual worlds offers so much more vivid and richer perceptual stimuli than traditional Web sites, e-mails, instant messages, and chat rooms, the experience can feel astonishingly real. Emerging from online multiplayer games and online chat rooms in the 1990's, virtual worlds have expanded from entertainment-themed games to various fundamental areas of human society including economy, education and training, healthcare, research, and our social life. Virtual worlds allow users to interact without revealing their real identity including their race, skin color, gender, social class, or disability, so users are able to create rich virtual lives—identities they can tailor to their desires: old people become young, infirm people become vibrant, paralyzed people become agile [11]. As our society progressively invents and integrates more advanced technologies including virtual worlds, overcoming technology access barriers for different user groups and promoting equalization across the whole society become even more critical. Research shows that virtual worlds can be especially beneficial to physically challenged users, not only to enhance their independence and mental health, but also to increase career and education opportunities for them [2].

To provide concrete assistance to physically challenged users and evaluate our work, we will build an efficient client system to access Second Life. Second Life is a massive general-purpose 3D virtual world with great potentials for promoting online community and collaboration [3]. According to the latest statistics, its 16.4 million registered users spent 36.8 million hours on Second Life in November 2008 [6]. Users can walk, run, fly, and "teleport" around vast realms offering shopping malls, bars, homes, universities, parks, and even embassies. To socialize, users (through their avatars) can schmooze, flirt, and comfort one another using lifelike shrugs, slouches, nods and other gestures while typing instant messages or talking directly. Users can also create and trade virtual properties and services with a huge collection of flexible tools. A recent study showed that many people with disability start using Second Life to escape from their disability and appear as able-bodied users [10].

Our research aims to overcome the access barriers to virtual worlds for motor- and speech-impaired users by building a gaze-controlled interface for Second Life that will enable them to interact with the virtual world by just moving their eyes. Motor-impaired and speech-impaired users, such as people with ALS, cerebral palsy, or muscular dystrophy, are a fairly large user group. For example, cerebral palsy occurs in approximately 1.4 to 2.4 of every 1,000 people. Currently, there are more than 500,000 people with cerebral palsy in the United States [4]. As a starting point, we have conducted a preliminary study to assess

(1) the facilitation of gaze-controlled text input using word prediction techniques to speed up chat-style text input in virtual worlds
(2) the influence of screen layout on the efficiency of text input
(3) the effect of the maximum number of suggested words on typing efficiency
(4) the performance of non-disabled vs. motor-impaired users

Our method offers, after each keystroke, a small number of suggested words or phrases to minimize the extra cognitive load imposed by the process of browsing these suggestions. Both non-disabled subjects and Amyotrophic Lateral Sclerosis (ALS) patients have participated in our experiment.

In our experiment, we analyzed two variables that are indicative of the subjects' task performance: the average time taken to type a word and the average number of triggering the delete key. Experimental results show that on average the patients took less time and fewer corrections per letter than did the non-disabled subjects. This finding suggests that our interface design is suitable for motor-impaired users, at least for those whose impairment is at the level of the patients in our study. Furthermore, word prediction has improved performance for both groups of subjects. The number and display position of the suggested words, however, do not seem to be important for the usability of the system.

2 Gaze-Controlled Text Input

One important input type employed in assistive technology utilizes an individual's eye gaze. Rather than using a mouse to select items on the computer screen, the user selects targets by gazing at the icons [7]. The user's current gaze position is usually determined through an infrared camera that is directed at the user's eye. This camera can be attached to a headset that the user needs to wear (e.g., the SR Research EyeLink-II system), or the camera may observe the user from a remote position (e.g., the LC Technology EyeGaze system). Several different kinds of gaze-controlled interfaces have been developed for physically challenged persons [5][8][9]. The most widely employed paradigm in this field of research is "typing by eye", which enables the user to type text by fixating and thereby "pressing" keys on a virtual on-screen keyboard [9].

Sixteen non-disabled subjects (three of them with gaze-control experience) and three ALS patients (two of them with gaze-control experience) participated in the experiment after giving informed consent. The study was approved by the UMass Boston Institutional Review Board. The ALS patients were unable to move any of their limbs and were tested while sitting in their wheelchairs. None of them showed severe eye-movement abnormalities. All subjects received the same task—using a preliminary version of the typing interface to copy texts that were shown to them on top of the screen. Two different layouts were used, which differed in the location of the words suggested by the word prediction algorithm (see Figure 1). The actual word prediction algorithm [1] was not yet incorporated in this version of the interface, but the algorithm was used off-line to generate predictions for the prefixes of all words used in the experiment. While this approach could not utilize the full capabilities of the algorithm, it gave at least an indication of whether word prediction would increase typing performance. In different experimental conditions, the interface provided a maximum of 0, 1, 2, or 3 different suggestions at a time.

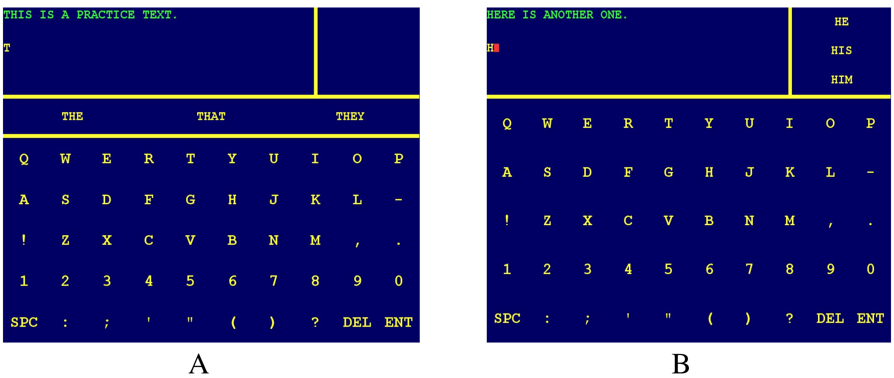

Fig. 1. The two interface layouts used in the preliminary study. The subjects' task was to type the text shown at the top of the screen. (A) Word suggestions shown below the text field; (B) word predictions shown to the right of the text field.

The combination of layout (two levels) and number of word suggestions (four levels) led to a total of eight different experimental settings. Each subject performed ten trials, starting with two practice trials to become familiar with the interface (one trial for each layout), followed by eight experimental trials, in which the eight conditions were tested in randomized and counterbalanced order. Each subject was given the same eight texts, whose order was also randomized and counterbalanced. Each text consisted of approximately 80 characters.

3 Results and Discussion

We analyzed two variables that are indicative of the subjects' task performance: The average time taken to type a text, and the average number of pressing the *delete* key. In order to account for differences in the length of the actually typed texts, both variables were divided by the number of characters that subjects typed. Trials with more than ten mistakes, i.e., differences between model text and typed text (Levenshtein distance), were excluded from analysis. For the non-disabled subjects, varying the number of suggested words between one and three did not influence either of the two performance variables, all ps > 0.5. However, offering no word prediction led to longer time taken per letter than for any of the three word prediction conditions, all ts(15) > 3.23, ps < 0.01 (see Figure 2A). Similarly, without word prediction, subjects pressed the delete key more often than when one, two, or three words were suggested, all ts(15) > 3.51, ps < 0.01 (see Figure 2B). The comparison of layouts A and B for those trials with word prediction did not reveal any significant effects on either the time taken per letter, t(15) = 0.53, p > 0.5, or the number of delete operations per letter, t(15) = 0.48, p > 0.5.

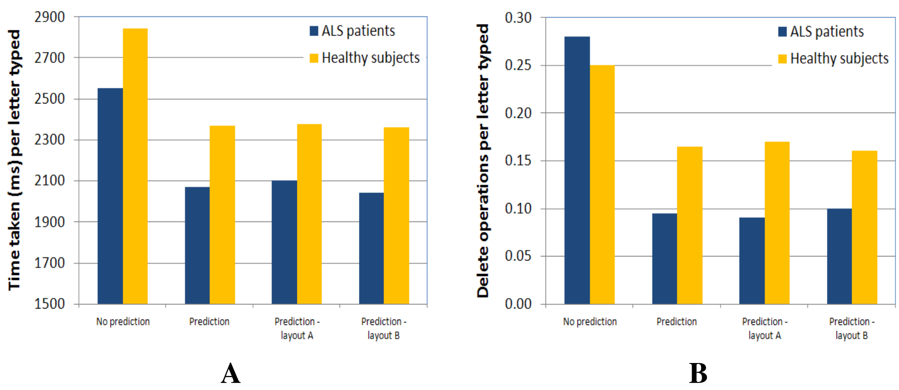

Fig. 2. (A) Time taken and (B) delete operations per letter typed for the ALS patients and the non-disabled subjects. Each chart compares the "no word prediction" and "word prediction" (1, 2, or 3 words) conditions (two leftmost groups of columns). Among the "word prediction" trials, layouts A and B are compared (two rightmost groups of columns).

When comparing the data of the non-disabled subjects with those of the ALS patients, we need to be aware of the fact that only three ALS patients were available for the study, which is insufficient for meaningful statistical analysis. Moreover, as mentioned above, a greater proportion of patients than non-disabled subjects were familiar with gaze control. Despite these limitations of the preliminary study, it is interesting to see that on average the patients took less time and fewer corrections per letter than did the non-disabled subjects (see Figure 2). This finding suggests that our interface design is suitable for motor-impaired users, at least for those whose impairment is at the level of the participating ALS patients. Furthermore, word prediction seems promising to improve performance for both groups of subjects. The number and display position of the suggested words, however, do not seem to be important for the usability of the system. These findings will be considered for the design of the future systems.

References

1. H. Al-Mubaid, P. Chen, Application of word prediction and disambiguation to improve text entry for people with physical disabilities, International Journal of Social and Humanistic Computing, vol.1, no.1, 2008.
2. M.G. Brodwin, E. Cardoso, T. Star, Computer assistive technology for people who have disabilities: computer adaptations and modifications, Journal of Rehabilitation, 70, 28–33, 2004.
3. Linden Lab, Second Life, www.secondlife.com
4. L. Pellegrino L, Cerebral palsy, In M Batshaw, ed., Children with disabilities, 5th ed., pp. 443–466. Baltimore: Paul H. Brooks Publishing, 2002
5. M. Pomplun, B. M. Velichkovsky, H. Ritter, An artificial neural network for high precision eye movement tracking, In Nebel, B. & Dreschler-Fischer, L. (Eds.), Lecture notes in artificial intelligence: AI-94 Proceedings, 63-69. Berlin: Springer Verlag, 1994.
6. Second life key metrics, http://s3.amazonaws.com/static-secondlife-com/economy/stats_200811.xls
7. L. E. Sibert, R. J. K. Jacob, Evaluation of eye gaze interaction, In Proceedings of the SIGCHI conference on Human factors in computing systems. The Hague, The Netherlands, pp. 281–288 2000.
8. J. Spaepen, M. Wouters, Using an eye-mark recorder for alternative communication, In A. M. Tjoa, H. Reiterer and R. Wagner. (Eds.), Computers for Handicapped Persons, Vienna: R. Oldenbourg, 1989.
9. D. M. Stampe, E. M. Reingold, Selection by looking: A novel computer interface and its application to psychological research, In J.M. Findlay, R. Walker & R.W. Kentridge (Eds.), Eye Movement Research: Mechanisms, Processes, and Applications. Amsterdam: Elsevier, 1995.
10. S. Stevens, www.disabilitynow.org.uk/living/other/secondlife_may_2007.htm, 2007
11. Washington Post, Real hope in a virtual world, http://www.washingtonpost.com/wp-dyn/content/story/2007/10/05/ST2007100502446.html?hpid=topnews