

WATARIDORI: Multiple Ghost Cursors for Speech-Based Cursor Movement

Yoshiyuki Mihara
Dept. of Mathematical and
Computing Sciences
Tokyo Institute of Technology
mihara1@is.titech.ac.jp

Shin Takahashi
Dept. of Computer Science
University of Tsukuba
shin.takahashi@acm.org

Etsuya Shibayama
Dept. of Mathematical and
Computing Sciences
Tokyo Institute of Technology
etsuya@is.titech.ac.jp

ABSTRACT

We propose a speech-operated cursor that is easy to move to any desired position. Speech-recognition systems are useful for people with physical disabilities, because such systems allow them to manipulate a computer using voice only. However, using traditional speech-recognition systems it is difficult to move the cursor to a desired position. To make it easier to specify the desired position, our system displays multiple aligned ghost cursors, which are user-friendly. Users can easily specify a one-dimensional coordinate by looking at the aligned ghost cursor. We examined the practicality of this technique.

Categories and Subject Descriptors: K.4.2 [Social Issues]: Assistive Technologies for Persons with Disabilities; H.5.2 [User Interfaces]: Graphical User Interfaces (GUI), Voice I/O;

Additional Keywords and Phrases: mouse cursor

INTRODUCTION

By using speech-recognition systems such as IBM ViaVoice [1] and Dragon NaturallySpeaking [2], a user can manipulate a computer application without requiring a keyboard or mouse. Speech-recognition systems are useful tools for people with physical disabilities. However, cursor operations using the voice alone are rarely simple, as the techniques used result in complex operations.

- In a text editor, there are two ways to reposition the cursor in a document: the user either states the word that he or she wants to select, or specifies the direction and distance to move (*e.g.*, lines, words, and letters). For example, “Move right four words”.
- When a user wants to reposition the cursor to somewhere else on the screen, there are two options: either to name the desired target, *e.g.*, an icon, so that the cursor navigates there directly, or to specify the absolute coordinates in a subdivided screen, or relative ones, by using centimeters, pixels, or units (a unit is a region in the subdivided screen); for example, “Move right twenty pixels”.

Of these techniques, the last is particularly complicated, but is needed when the user wants to select objects that do not

have a name, or to draw new objects in a particular position using a drawing tool. To specify relative coordinates using centimeters etc., requires a ruler, but none of the available speech-recognition systems provides a ruler for the whole screen. Even when there are rulers at the top and to the left of the screen, the user needs to move his or her eyes frequently to relate the desired position to the values on the rulers.

DISPLAY MULTIPLE GHOST CURSORS

We propose an interface in which multiple ghost cursors are displayed, aligned either vertically or horizontally with the actual cursor when the user specifies the direction to move from the initial state (Fig. 1(a)). Fig. 1(b) shows the multiple ghost cursors that appear in the horizontal direction when the user issues the command, “Move right”.

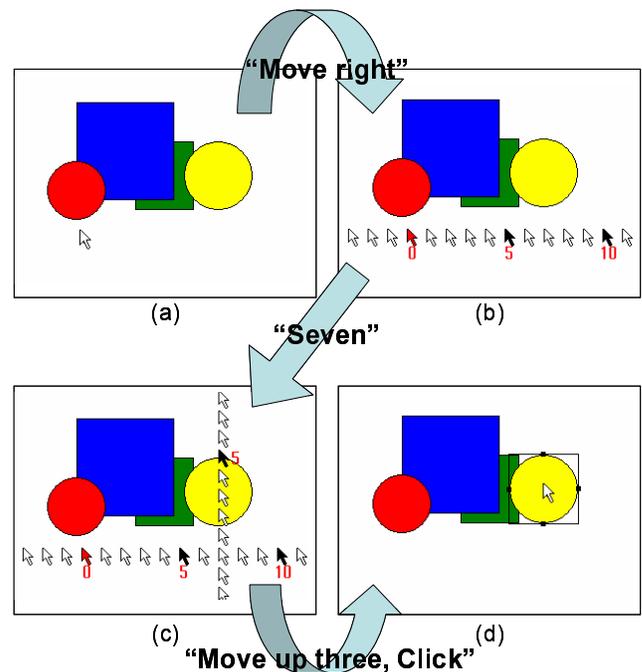


Figure 1: An example of commands and cursor movement

When the multiple ghost cursors are displayed, the user specifies the number of the ghost cursor that is in the desired position. If the user wants to select the right most yellow circle

shown in Fig. 1 when the current position is the red cursor, the user should issue the command “Seven” followed by “Move up three, Click” (Fig. 1(c),(d)).

Although repositioning the cursor involves vertical and horizontal steps, users can easily specify one-dimensional coordinates by looking at the aligned ghost cursors.

Easily Viewable Expression

As multiple ghost cursors might make it difficult for a user to specify a specific cursor, every fifth cursor is a different color and is numbered (Fig. 1(b)), allowing the user to identify the desired cursor immediately. The user does not need to relate the cursor to a ruler, and the system decreases the number of eye movements required.

The system also provides feedback as to where the focus is currently located (Fig. 1(c)). When the user issues the command, “Seven”, the cursor does not move immediately. New aligned cursors appear at the identified point, oriented vertically to the ghost cursors that were originally displayed (Fig. 1(c)). The user can then see whether any one cursor is in the desired position. If no cursor is in the desired position, the user reissues the number of the identified cursor, *e.g.*, “Six” or “Eight”. When the user reissues the number of the cursor, new aligned ghost cursors move and the user can see again whether any one cursor is in the desired position. If the user wants to select an object, he or she issues the command, “Click”, or if the user wants to move the cursor, he or she issues the command “Move”.

PRELIMINARY EXPERIMENT

We tested our technique by an experiment in which the subjects specified the position of an object eight times. For comparison, we tested another technique in which the subjects specified absolute coordinates using a subdivided screen with rulers above and to the left of the window. In both the subdivided screen and our system, the interval was 20 pixels. When the user issued the command “ready”, a cursor and the object to be specified appeared on the screen. Six individuals were recruited for the test. We recorded the times required to select the object using the two techniques and then interviewed each subject.

The times required to select the object differed by 10% at most between the two techniques. Although three subjects preferred our technique because it required less concentration to select the object, one subject found the two-step approach of our technique tedious.

RELATED WORK

Manaris *et al.* proposed the system SUITEKeys, which is a speech interface for controlling a mouse [3]; the user identifies the desired direction and then the cursor moves in the direction at a certain rate of pixels every second until the user issues a stop command. This technique requires the user to specify the value of the time-axis. Specifying one dimension is easier than specifying two dimensions, such as absolute coordinates.

However, a moving cursor system requires time to process the stop command; consequently, the cursor does not stop immediately at the desired position and moves further ahead

during the delay. Karimullah and Sears [4] therefore positioned a ghost cursor ahead of the actual cursor to indicate where the cursor will stop after the stop command is issued.

In these systems, the user must wait until the moving cursor is at the desired position. Instead, we used the concept of ghost cursors, which requires the user to specify one visual dimension.

FUTURE WORK

Zooming in on a Specified Area

In the current implementation, the user can specify a position only within 16 pixels, because we implemented the cursor as a 16x16-pixel icon. To specify a position more accurately, the system has to zoom in on the area where the cursor is located. The user can then specify the cursor position in the zoomed image.

Non-Verbal Voice Input

We also want to introduce a function that allows the user to specify the cursor by a non-verbal voice command, using the technique proposed by Igarashi *et al.* [5]. For example, if the user wants to move right to position 18, the user first states, “Move right”, then specifies a number near the desired position, such as “Fifteen”, and then shifts the position by using a non-verbal input “ta ta ta”, until the focus arrives at cursor 18.

CONCLUSION

We proposed an interface that displays multiple ghost cursors for individuals with physical disabilities that hinder their ability to use a keyboard and mouse. Multiple ghost cursors provide users with an easy way to identify any position by specifying a one-dimensional value with our user-friendly view. However, a test showed that some subjects felt that specifying two steps was tedious. We have enhanced our visual effects, and in a future study we will improve the commands issued to select objects.

REFERENCES

1. IBM ViaVoice:
<http://www-306.ibm.com/software/voice/viavoice>
2. Dragon Naturally Speaking:
<http://www.scansoft.com/naturallyspeaking/overview>
3. Bill Manaris, Renée McCauley and Valanne MacGyvers: “An Intelligent Interface for Keyboard and Mouse Control - Providing Full Access to PC Functionality via Speech”, *Proc. of International Florida AI Research Symposium (FLAIRS-01)*, pp.182-188, 2001.
4. Azfar S. Karimullah and Andrew Sears: “Speech-Based Cursor Control”, *Proc. of the International ACM SIG-CAPH Conference on Assistive Technologies (ASSETS)*, pp.178-185, 2002.
5. Takeo Igarashi and John F. Hughes: “Voice as Sound: Using Non-verbal Voice Input for Interactive Control”, *Proc. of Symposium on User Interface Software Technology (UIST)*, pp.155-156, 2001.