A study on Two-Dimensional Scrolling with Head Motion

CLIPS-IMAG Technical Report
Ref: TR-IMAG-CLIPS-IIHM-199901
January 8, 1999
CLIPS laboratory
BP 53
38041 Grenoble Cedex 9
FRANCE

AUTHORS

François Bérard (http://iihm.imag.fr/fberard/)

ABSTRACT

This technical report presents an innovative interaction technique for simultaneously scrolling the content of a window in horizontal and vertical dimensions. Scrolling is controlled by head rotation via a non-intrusive computer vision system. Preliminary user studies show that this technique is easy to learn and provides substantial speed and accuracy advantages over scrollbar manipulation. Completion times for tasks including both scrolling and pointing are more than 30% faster for head-and-mouse than for mouse-and-scrollbar interactions.

KEYWORDS

2D scrolling, Input Device, Computer Vision, non Intrusiveness.

AVAILABILITY

This document is available in electronic format at the following addresses:

http://iihm.imag.fr/publs/1999/TR199901_PWindowRate.ps.gz

http://iihm.imag.fr/publs/1999/TR199901_PWindowRate.pdf

A study on Two-Dimensional Scrolling with Head Motion

François Bérard

Ref: TR-IMAG-CLIPS-IIHM_199901

HCI group Technical Report
CLIPS-IMAG
B.P. 53
38041 Grenoble Cedex 9
http://iihm.imag.fr/

Scrolling the content of a window can be troublesome, especially when both horizontal and vertical displacements must be interleaved with point-and-click actions. This sort of activity is very common when editing large documents such as modifying drawings at high resolution or working on a large spreadsheet document. Current graphical user interfaces offer horizontal and vertical scrollbars for navigating over document workspaces. As a result, shifting attention diagonally across a document display induces extraneous articulatory tasks that interrupt the central flow of the domain task. As shown in Figure 2, diagonal shifting involves moving the cursor from the editing area to a scrollbar, moving the elevator at an approximate position, manipulating the second scrollbar in a similar way, possibly switching multiple times between the scrollbars to adjust the region of interest, finally moving the cursor back to the editing area.

One way to address this problem is to allocate navigation and pointing tasks to independent devices. Browsing performance (i.e., navigating, pointing and clicking) can be more efficient when scrolling is controlled with a device distinct from the mouse and the vertical scrollbar [3]. However, two devices can be worse than just one if they do not match task requirements [2, 3]. We propose to use the face as a new input device for scrolling, thereby leaving the mouse to pointing and selection tasks. We demonstrate experimentally a significant improvement of this apparatus over the joint use of horizontal and vertical scrollbars.

The perceptual window

Our system, called "Perceptual Window", is based on a Computer Vision head tracker that measures the parameters used to control window scrolling. The system is "perceptual" in the sense that it reacts directly to user's actions in relation to the content of a graphical window, without any need for separate physical contact.

SYSTEM OPERATION

As the head is tilted downward from a *neutral* area, the window content is scrolled up. The more the head is tilted, the faster the scrolling. Upward scrolling is stopped by moving the head back to its neutral position. Similarly, tilting the head upward, left or right causes vertical or lateral displacements in the document view. Scrolling speed is governed by an exponential rather than a linear relationship to this movement, permitting fine adjustments or large translations depending on the degree of head rotation, this is represented on the Figure 1.

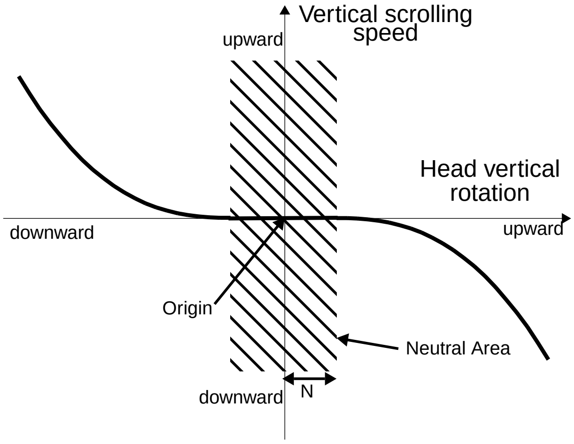

Figure 1: Transfer function for upward / downward head rotations.
N is the radius of the neutral area.

Head motion scrolling is controlled with the “tab” key in two ways. When the user first depresses “tab”, the neutral area is set to the user’s current head location. Secondly, the system is switched into head-scrolling mode: scrolling occurs only while the “tab” key is depressed; it stops immediately when the key is released. A movie of the perceptual window is available on the web to provide a better illustration of the system[1].

IMPLEMENTATION

The input to the Perceptual Window is provided by a video camera set in front of the user’s head (on top of their monitor). Head motion is detected by tracking a region of the face over time using correlation matching [1]. For this particular experiment, the region was chosen manually as the tip of an eyebrow or the frame of eyeglasses. For a real world application, the Perceptual Window would need to be autonomous, that is, independent of user intervention for its initialization. We have previously described how the cooperation of multiple vision techniques can satisfy the autonomy requirement [1]. The vision system tracks the target in real-time by focusing on the recovery of the translation of one facial feature only. The simplicity of the technique allows the system to process more than 60 frames per second with no dedicated hardware on a 350 Mhz Power Macintosh. We thus obtain tightly coupled interaction leading to a sense of direct interaction with the content of the graphical window rather than the window itself. A user study reported in the next section demonstrates the added-value of the Perceptual Window.

1. http://iihm.imag.fr/demos/pwindow/

An Experiment

The experimental task was designed to represent editing a large 2D document through a small window. The navigation component of the task reflects cues for where to go next given by the current content of the window. This occurs for example when following lines and columns of a spreadsheet, when following a mechanical part's shape on a technical sketch, or via a mental representation of a large picture when only a small part of it is visible.

In our task, subjects worked with a document displayed on a GUI. The notional area of the document was 2400x3000 pixels while the visible area was only 400x500 pixels (i.e., 2.8% could be viewed at any time). Subjects were presented with a succession of targets, clicking on each one causing it to disappear and the next to appear (targets are small crosses 12x12 pixels large, drawn with 4 pixels wide lines).

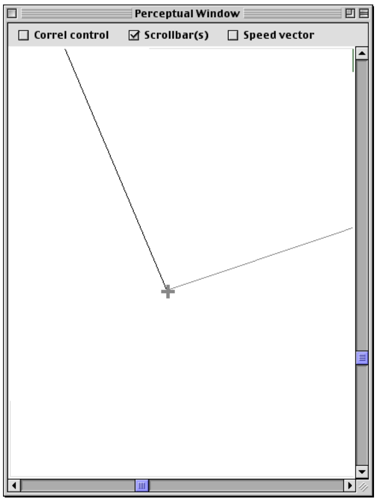

Figure 2: A target along the path

The target locations were distributed randomly throughout the document. Subjects knew where to scroll for the next target by following a line drawn between the last target location and the next. Another line is drawn between the new target and the next one, so that subjects can anticipate the next scrolling direction before clicking on a target (see Figure 2). Subjects can infer their current location from the scrollbar handles.

Subject had to locate and click on 50 targets, once while scrolling with the scrollbars and once with head motions (the sequence of target locations is the same for the two different conditions). Half of 8 volunteer subjects started with scrollbars, the other half with head motion. All of the subjects used scrollbars on a daily basis but none of them had any previous experience with head-motion scrolling thus they were trained with a practice run of 30 targets. The overall experiment time was about half an hour per subject.

Results

Data were analyzed with a paired samples t-test which revealed a highly significant difference between scrolling modalities ($t(7) = 7.04$; $p = .000204$). As shown in

Figure 3, head motion outperformed scrollbars by an average improvement of 32% (scrollbar average completion time was 290 s. versus 198 s. for head motion).

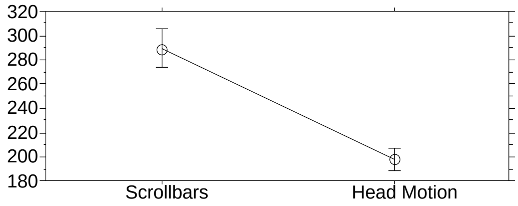

Figure 3: Completion time means and standard errors by scrolling modality.

Discussion

The most striking result from the experiment is the importance and the constancy of performance improvement. Performance improvement ranges from 15 to 42%, with 75% of the users improving more than 30%. Clearly, head motion better suits our 2D navigational task than scrollbars. This would be a big win if this transposes well to real-world tasks.

It also appears that subjects were able to use this very new modality surprisingly well: in less than a minute, they knew how to use it and could do so with great facility. With minimal training, all of the subjects were a lot faster with head motion than with the scrollbars. They all preferred head motion over the scrollbars. One of them commented that scrolling control with the head was very *natural*: he simply had to orient his head towards what he wanted to see, and it just appeared in the middle of the window. Perceptual Window has great potential for improving the transparency of interaction with content.

In addition to scrolling tasks, this work shows that perceptual interfaces based on computer vision can now be considered as a realistic modality in our everyday use of computers. Many workstations come with built-in image processing capabilities such as frame grabbers and digital cameras. These capabilities offer a whole range of new input possibilities (head, face, hands, body...) at no extra hardware cost, and allow potentially more natural interactions than physical devices. However, the lack of autonomy of currently available Computer Vision techniques are the main limitations. For example, our tracker must be manually initialized every time the user looks away from the monitor. We are working on it.

Acknowledgements

Many thanks to Leon Watts for his contributions to this work.

References

1. Coutaz, J. Bérard, F. and Crowley, J. Coordination of Perceptual Processes for Computer Mediated Communication. *Proceedings of Automatic Face and Ges-*

ture Recognition (AFGR'96). IEEE Computer Society Press 1996, 106-111.

2. Kabbash, P. Buxton, W. and Sellen, A. Two-Handed Input in a Compound Task, in *proceedings of CHI'94* (Boston MA, April 1994), 417-423.
3. Zhai, S. Smith, B. A. and Selker, T. Improving Browsing Performances: A study of four input devices for scrolling and pointing tasks. *Proceedings of INTERACT'97*. Chapman & Hall 1997, 286-293.