Shakespeare's Complete Works as a Benchmark for Evaluating Multiscale Document-Navigation Techniques

Yves Guiard, Yangzhou Du, Jean-Daniel Fekete, Michel Beaudouin-Lafon,
Caroline Appert, Olivier Chapuis

Shakespeare's Complete Works as a Benchmark for Evaluating Multiscale Document-Navigation Techniques

Yves Guiard[a]
Michel Beaudouin-Lafon[b]

Yangzhou Du[b]
Caroline Appert[b]

Jean-Daniel Fekete[c]
Olivier Chapuis[b]

[a]Mouvement et Perception
CNRS & Univ. de la Méditerranée,
Marseille, France
yves.guiard@univmed.fr

[b]LRI & INRIA Futurs
Bât 490, Univ. Paris-Sud,
91405 Orsay, France

[c]INRIA Futurs & LRI
Bât 490, Univ. Paris-Sud
91405 Orsay, France

{du,mbl,appert,chapuis}@lri.fr Jean-Daniel.Fekete@inria.fr

ABSTRACT

In this paper, we describe an experimental platform dedicated to the comparative evaluation of multiscale electronic-document navigation techniques. One noteworthy characteristics of our platform is that it allows the user not only to translate the document (for example, to pan and zoom) but also to tilt the virtual camera to obtain freely chosen perspective views of the document. Second, the platform makes it possible to explore, with semantic zooming, the 150,000 verses that comprise the complete works of William Shakespeare. We argue that reaching and selecting one specific verse in this very large text corpus amounts to a perfectly well defined Fitts task, leading to rigorous assessments of target acquisition performance. For lack of a standard, the various multiscale techniques that have been reported recently in the literature are difficult to compare. We recommend that Shakespeare's complete works, converted into a single document that can be zoomed both geometrically and semantically, be used as a benchmark to facilitate systematic experimental comparisons, using Fitts' target acquisition paradigm.

Categories and Subject Descriptors

H.5.2. [User Interfaces]: Interaction styles; I.3.6. [Methodology and Techniques]: Interaction techniques.

General Terms

Measurement, Experimentation, Human Factors, Standardization

Keywords

Fitts' law, Target acquisition, multiscale document navigation techniques, evaluation standard, evaluation benchmark.

1. INTRODUCTION

This paper is concerned with the establishment of scientific norms within the human-computer interaction (HCI) research community for the quantitative evaluation of multiscale electronic-document navigation techniques, a domain in which the innovation rate is currently fairly high [1][5][12][13][17][19]. Twenty seven years of research after Card et al.'s study of simple, single-scale, pointing tasks [4] have recently led to an ISO recommendation for the experimental evaluation of "non-keyboard input devices" [14], which apparently has already exerted an appreciable impact on practice in input research [20]. Research on target acquisition in multiscale documents, however, is of more recent appearance [9]. We believe that collectively adopting a standard multiscale document, held as a benchmark, would help researchers to draw comparisons among the variety of available techniques and thus constitute a useful first step towards standardization in this important field of HCI research.

Below, we start by reminding the basics of Fitts target acquisition task and Fitts' law, and recall why this paradigm is useful. Next, we turn to multiscale documents, showing that Fitts' paradigm is still perfectly appropriate to assess performance in the multiscale case. We then list the main conditions that we believe have to be met for performance data from different studies to be comparable, taking examples from the current literature. We end with the suggestion that Shakespeare's complete works, implemented in a single, highly flexible platform which we have recently developed, and which we are making available to the HCI community, might constitute a suitable evaluation benchmark.

2. MODELING USER PERFORMANCE WITH FITTS' PARADIGM

2.1 Fitts' Law in Simple Pointing Tasks

We designate as Fitts' paradigm [6] the conceptual ensemble formed by a simple experimental task, target acquisition or pointing, and a simple quantitative relationship, Fitts' law. For a one-dimensional (1D) pointing task in which target width is W and target distance is D, the index of pointing difficulty (ID) is defined as

$$ID = \log_2 (D/W+1) \qquad (1)$$

and movement time MT has been repeatedly shown to follow the empirical relationship

$$MT = a * ID + b \qquad (2)$$

where a and b are constants that can be determined by fitting a straight line to measured data. Fitts' law is frequently used as a tool to characterize or measure the performance of input devices [16]. Regardless of the specifics of the task, performance can be characterized with just two numbers provided that the measures are appropriate.

2.2 Fitts' Law in Multiscale Document Navigation

With the ever increasing capacity of our computers to store and process information, GUIs have allowed users to handle larger and larger electronic documents. However, adjustments have been necessary for the classical GUI to accommodate huge document sizes: the most important change has been that documents have been made scalable, thanks to techniques such as the zoom or the fisheye [7][8][9].

Not denying that there is still room for improving our single-scale pointing techniques in GUIs— obviously there is quite some research going on in this area of HCI (for a recent review, see [2])—, we believe the tailoring of new multi-scale target acquisition techniques to navigate very large documents will be a concern of great importance for HCI in the foreseeable future.

It is important to realize that Fitts' paradigm, in essence, remains applicable when the visualization scale of a document becomes a user-controlled variable. For example, zooming in and out is essentially neutral with regard to the target acquisition task because a zoom rescales both D and W, thus conserving the ratio D/W and hence the ID.

3. STANDARDIZING THE EVALUATION OF TARGET ACQUISITION TECHNIQUES

The whole value of Fitts' pointing paradigm rests in the fact that it generically captures the central property of target reaching tasks, ignoring all particulars of the tasks but the variables of interest, target distance D and target width W. Since the pioneering work of Card et al. [4], it has become routine practice in HCI research to resort to such measures as the throughput (dimensionally in bit/s), computed either as the ratio of mean ID to mean MT or the inverse of Fitts' law slope, to compare different pointing techniques or devices [20].

3.1 An ISO standard for the Evaluation of single-scale pointing devices

Because the intercept and the slope of Fitts' law, which quantify the performance of a given human-technique system, are sensitive to very many auxiliary parameters such as movement direction, any property of the interface, or seemingly unimportant choices made in the experimental design (e.g., number of trials per session), it is cautious for any comparison between or among pointing devices to implement the competing devices on the same platform and to test them in the same experimental design. The recent publication of an ISO standard [14][20] for pointing devices is an interesting step in this direction. The ISO 92419-9 standard, which concerns "non-keyboard input devices", defines a standardized multidirectional tapping task, thus controlling for the effect of movement direction. This standard also specifies a valid method of varying D and W, and of computing throughput (see [22], however, for a criticism of the recommended computation). Soukoreff and MacKenzie [20] surveyed nine recently reported Fitts' law studies that used the standard, showing that the homogeneity of the resulting data set has improved.

3.2 Towards a Standard for Multiscale Target Acquisition

It should be emphasized that the ISO 92419-9 standard only considers the case of traditional single scale pointing, that is, the case in which the target is visible within the view and can be acquired by means of a simple movement of the cursor across the view. Since the early nineteen nineties, graphical interfaces have become multi-scale, a change that has become necessary to allow users to handle extremely large documents [7]. As far as the pointing task is concerned, this means that users now need to *navigate* the document to reach targets possibly located a long way away from the current view. Importantly, Fitts' paradigm still applies to the case of multiscale target acquisition, regardless of the navigation technique [9][11].

Presumably the most familiar multiscale technique is that allowing users to both pan (scroll) their document and zoom it in and out. It is with this *pan-and-zoom* technique that the applicability of Fitts' paradigm was first demonstrated [11]. Recently, a variety of other multiscale navigation techniques have been described in the literature, a sample of which is briefly described below.

With *Speed-Dependent Automatic Zooming* (SDAZ) [13], the zoom level is automatically coupled with scrolling speed, with the consequence that the visualization scale adapts to the scale of the user's intentional motion.

With *Zliding* [19] the zoom level is modulated by the stylus pressure, while dragging on the input device performs panning at that zooming level.

OrthoZoom (OZ) *Scroller* [1] provides a similar control scheme for zooming and scrolling in very large one-dimensional space, based on standard mouse input. While panning is controlled by moving the mouse body in the direction of navigation, zooming is controlled by moving the mouse along the orthogonal dimension.

With *Perspective-view navigation* (PV) the camera can be freely tilted relative to the document plane [5][12]. With PV the visualization scale varies over space, with a whole range of scales displayed at the same time in the same view, that is, the user is able to look at the global context without losing the current detail [17] (as will be apparent in Figure 2 below).

It is noteworthy that for each of the above mentioned—or, for this matter, any other—multiscale document navigation technique, the efficiency of target reaching performance can be most conveniently quantified by the mean rate of change of the current ID defined as $ID_t = \log (D_t /W +1)$ over the duration of the movement (typically, several seconds), with D_t being defined as the current horizontal distance from the virtual camera to the target. As shown with an example in Figure 1 (borrowed from [10]), the ID declines at an approximately constant rate over time. In this example a simple linear regression yields a slope of 3.7 bits/s, and this amounts to a direct measurement of the throughput, one that is free of the estimation problem described in [22]. This performance quantification method can be used profitably to assess performance for the whole variety of multiscale navigation techniques.
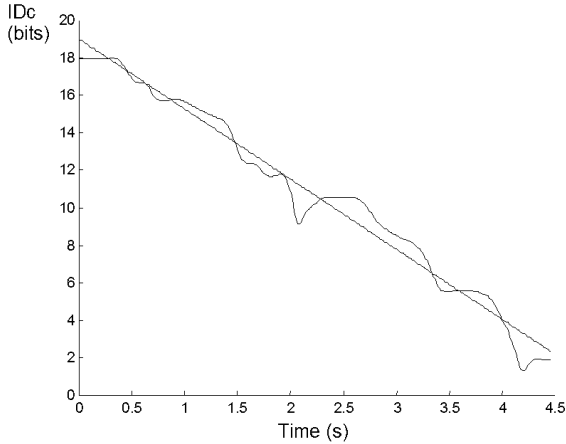
Figure 1. An illustration of the constant-rate decline of current *ID* over time during progression to the target, in a single reaching movement using the pan-and-zoom technique. The *r²* for the linear regression (straight line) equals .996.

Keeping in mind that the merits of a multiscale target-acquisition technique cannot be judged with the sole yardstick of Fitts' law performance, it remains that this yardstick is valuable as it captures an important component of electronic document navigation—namely, reaching a remotely located target. But obtaining a fair comparative evaluation of the throughput permitted by these various techniques requires some caution and a reasonably standardized approach. The central aspect of the standardization we recommend is recourse to Fitts' *target-reaching task paradigm*. That simply means defining a starting point and a target point in document space, with *D* and *W* measured in that space (so as to be independent of the level of scale, which typically will be subject to changes during navigation to the target). Specifically, *D* will measure the horizontal distance that separates the virtual camera from the target. In the simple 1D version of Fitts' paradigm that we use and recommend, the document will be linearly arranged along the *y* axis and lengths *D* and *W* will be measured along this single direction.

Recall that the case of interest here is that in which the *ID* is high enough (e.g., over 10 bits or so) that the target cannot initially be visible in the view, hence the necessity of view navigation [9][11]. Obviously acquisition of a single target in such a case will take a lot more time (e.g., 4.5 s for a 18-bit target in Figure 1) than traditional single scale pointing (where a typical movement takes less than a second). Such a time extension of the reaching movement has some interesting consequences from an experimental viewpoint, relative to traditional single-scale pointing. One is that there is no need to manipulate the *ID* as an explicit independent variable to evaluate throughput. A single, reasonably high level of *ID* will suffice. This can be easily understood with the help of Figure 1, where a single movement performed by the participant to reach an 18-bit target provides excellent data to estimate the rate of change of the *ID*: the steeper the slope, the most efficient the technique. Another consequence is that it does not really matter whether the evaluation experiment resorts to the so-called discrete or reciprocal version of Fitts' pointing paradigm (i.e., with a rest after each movement or with a concatenation of several to-and-fro movements). Finally, the problem of errors, so tricky in the case of single-scale pointing [16][20], vanishes: rather than providing one's participants with subtle speed-accuracy instructions, one will just ask them to maximize speed.

Another experimental component that needs to be standardized to make it possible to run fair comparisons among multiscale navigation techniques is the *document*, which provides the pointing environment. Whether participants are asked to navigate empty space, or a map, or a text document is not necessarily neutral with regard to performance. Hence the interest of a standard graphical material for multiscale navigation. In the next section, we present a prototype of such a material, a text document that we find attractive and that supports semantic zooming [3]. We will then turn to a description of the platform we have recently developed, a highly flexible tool aimed at accommodating the whole variety of known multiscale navigation techniques.

4. OUR EXPERIMENTAL PLATFORM

We have developed an experimental platform for evaluating multi-scale navigation techniques in which a number of techniques can be evaluated, including standard Pan-and-Zoom, SDAZ, OrthoZoom and Perspective View. Figure 2 shows a screen shot of the interface for navigating the document under perspective view. [1]
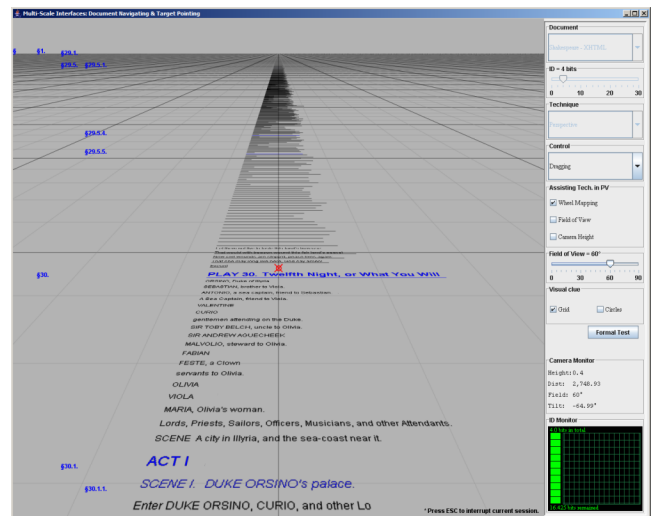

Figure 2. A screen shot of our experimental platform.

The program window displays two parts: the view area, where the document is visualized and interaction takes place, and the control area, on the right-hand side. The control area offers tools to set parameters like pointing task difficulty (*ID*) and to make a number of selections like navigation technique, mouse control mode, and type of visual cue. Also available at the bottom are a real-time monitor reporting camera state parameters and an *ID* monitor reporting the instantaneous level of *ID*.

---

[1] When visualizing a document in perspective one faces the problem that relatively few lines are actually legible (Figure 2). This is due to the limited level of detail of our state-of-the-art screens. As far as pixel density is concerned, the perspective visualization technique should benefit to an especially great extent from future improvements in screen technology.

4.1 The Document: Shakespeare's Works

The document we have implemented in the platform is a very large HTML file that comprises Shakespeare's complete work, which consists of 37 plays, about 150,000 lines overall. As all these lines are arranged in a single column, reaching and selecting one specific verse in this very large text corpus amounts to a perfectly well defined Fitts' task. By Fitts' measure, i.e. Eq. (2), here the *maximum* index of difficulty for reaching one line is

$$\log_2 (150\,000 / 1 + 1) \approx 17.2 \text{ bits,}$$

this *ID* corresponding to the limiting case where one must move from the first line to the last, or from the last line to the first. So this material is suitable for the study of very difficult target reaching. However, this does not exclude its use for pointing at lower levels of difficulty. For example, reaching one particular line at a distance of 10 lines (a task that can obviously be carried out in a single screen, at a single level of visualization scale) would correspond to a well defined Fitts task with an *ID* of 3.5 bits (i.e., $\log_2 11$).

After [1] we resorted to *Semantic zooming* [3] to visualize the Shakespeare document. Where the height of a verse is too small, it is simply represented as a graphical line (Figure 2). A multiscale table of contents (MSTOC) is permanently displayed on the left-hand side of the screen to show sections. There are three levels of heading in the MSTOC: PLAY, ACT and SCENE, entry §A.B.C. meaning PLAY A, ACT B, SCENE C. Note that these entries are displayed at a constant size but are aligned vertically with the position of the corresponding text. In the case of limited display area, only a subset of entries is clearly visible while the others become gray and fade with the background. The gray level and sampling choice is stable with the navigation so labels don't jump or quickly change while the user interacts with the system. They smoothly fade out or appear in a predictable fashion.

In recently published work on multi-scale pointing, navigation typically has taken place over a blank document—there was nothing to see but the geometric object standing for the target. The main shortcomings of such an option are that the task looks rather abstract and the participant does not have the experience of flying over some meaningful environment. Asking participants to navigate Shakespeare's complete works to reach some specific verse means offering them an environment that is not only visually structured but also, we feel, worth exploring.

4.2 Design Rationale

The platform offers several navigation techniques and is meant to evolve to accommodate new techniques as well as new input devices. We provide the platform with the source code to simplify the design of experiments and gather comparable results in term of multi-scale pointing. The platform is meant to be useful in the following situations:

- To implement a new navigation technique and compare it with existing ones already available in the platform.
- To test a new input device with existing navigation techniques.
- To test navigation techniques in different contexts, such as with smaller view ports or on new kinds of computers (PDA, Tablet PC etc.)

To support a wide range of navigation technique, the platform implements several features. Among the multiple multi-scale navigation techniques, some use a 2D surface while others rely on 3D. The platform implements several sophisticated 3D navigation techniques and offers support to implement new ones. Multi-scale techniques should also provide overviews when the view is zoomed in. These overviews can be used to help navigating so we provide them in the special context of Shakespeare's plays.

4.3 Allowing Perspective View

One noteworthy characteristics of our platform is that it allows the user not only to translate the document, that is, to pan and zoom, but also to tilt the virtual camera so to obtain freely chosen perspective views of the document. To explore the new navigation features provided with perspective view, we implemented three kinds of camera rotation in the platform, as illustrated in Figure 3.

Assuming the user knows the location of the target, an effective way to navigate is to simply tilt the camera until the area of interest enters the view, which we call a *panoramic* rotation (Figure 3A). But if the user wants to peek at a distant region without losing sight of the local selection, perhaps a better option is what we call the *lunar* rotation (Figure 3B), which consists in revolving the camera along a half-circle with the camera being constraining to remain oriented towards its current fixation point in document space. Another interesting option is what we call the *trans-rotation*, which consists in translating the camera with the constraints that it remains oriented to the same fixation point in the document and travels at the current altitude (Figure 3C).
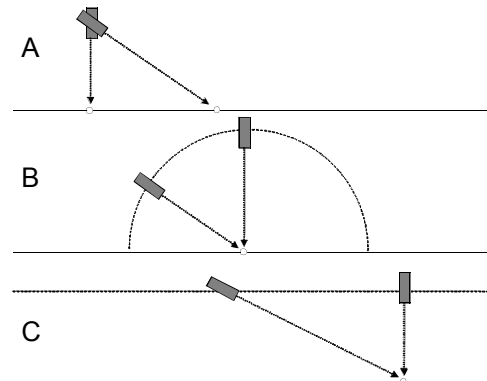

Figure 3. Three kinds of camera rotation, the panoramic rotation (A), the lunar rotation (B), and the trans-rotation (C). The camera's fixation point is shown as a small unfilled circle.

4.4 Semantic Grasping

The camera tilt facility we have implemented increases the versatility of our platform. However, one specific difficulty arises with regard to target grasping. Grasping one line of text near the horizon (see Figure 2) is problematic because in this part of the display one screen pixel maps to a considerable range in document plane. Overshooting by just one-pixel may result in missing thousands of lines. This problem is caused by the highly nonlinear variation of scale along the dimension of interest, the y axis in our implementation [5][12].

To deal with this problem, we used the snapping function of [1] for MSTOC entries. When the mouse cursor is moved over an entry, the cursor shape turns from a pointer into a hand. When the user drags the entry, (s)he does not suffer from grasping difficulty thanks to a underlying technique that helps the user to grasp the real target – as though the grasping had sub-pixel accuracy, and

the mouse seems to really drag the corresponding text line. We call this snapping function *semantic grasping*.

4.5 Input Parallelism

To enable our platform to be run on any hardware, the only input device required is a standard wheel mouse. In perspective view navigation, the mouse movement controls the camera translation parallel to document plane and does not change the visualization scale. When the mouse wheel is present, its rotation controls the *scaling factor*, that is to say, camera altitude in PZ and camera tilt in PV. Wheel-mouse scaling can also be used for navigating in a 2D zoomable configuration.

Other input devices can be used in our platform, thanks to the Java3D environment. Although managing multiple input devices is not standard in Java, it can be done with modest efforts and we provide several examples. This is required to support multi-scale navigation techniques described in [5][12].

4.6 Background Texturing

We are dealing with a generalized pointing task where the target may be located a long way away from the current view. Acquiring such a target requires first moving the view so as to make the target visible, then pointing at the target with the screen cursor. The former we call *view pointing*, or navigation, and the latter *cursor pointing* [10]. For the navigation of a blank document, background texture may guide navigation. For example, in the multi-scale pointing experiments of [10][11], where the camera remained orthogonal to the document plane, the target was surrounded by an infinite set of concentric circles: with the curvature information, participants always knew in what direction they had to travel and how far the target was.

This strategy remains suitable for reaching targets of middle level of pointing difficulty with a tilted camera (Figure 4 left). However, for perspective views with a target at a very large distance from the observation point, the curvature sometimes becomes hard to detect and the curvature varies with camera tilt. This affects the user's judgment so we designed a complementary visual cue: a mesh grid (Figure 4 right). This helps users perceive the camera position and orientation. Finally, when the target is out of view, an arrow appears to indicate the target direction. Because concentric circles and a mesh grid have their own advantages and disadvantages, both can be enabled in our platform.
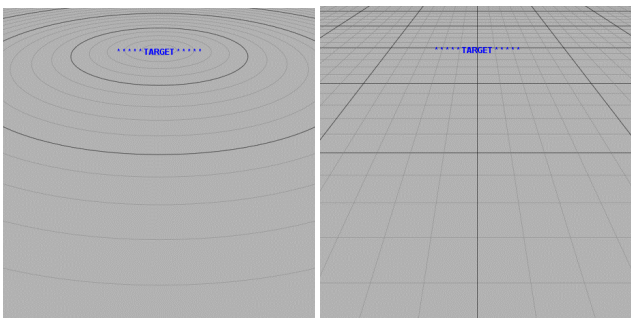

Figure 4. Two kinds of visual cues. Left: concentric circles. Right: mesh grid.

4.7 Deployment with "Java Web Start"

We implemented this experimental platform with Java SE and the Java3D library. Our program can be run on different operating system as long as there is a Java environment and supported 3D hardware acceleration. We use the Java Web Start Technology to simplify the distribution of our platform. We received our inspiration from the Web-based Test of Fitts' Law [24], which allows people to practice Fitts' law within a web browser [23]. Users can access our standard platform and play with the various navigation techniques at once.

Researchers or industrials interested in the test of new techniques or devices may also download the source code of the platform. The project is available at the following location:

https://gforge.inria.fr/projects/multiscalenav

5. REFERENCES

[1] Appert C., Fekete J.D. (in press), OrthoZoom Scroller: 1D Multi-Scale Navigation. *Proceedings of CHI 2006*. ACM Press.

[2] Balakrishnan, R. (2004). 'Beating' Fitts' law: Virtual enhancements for pointing facilitation. *International Journal of Human-Computer Studies, 61*, 857-874.

[3] Bederson, B.B., Hollan, J.D. (1994), PAD++: A Zooming graphical interface for exploring alternate interface physics. *Proceedings of UIST'94*, 17-26. ACM Press.

[4] Card, S. K., English, W. K., Burr, B. J. (1978). Evaluation of mouse, rate-controlled isometric joystick, step keys, and text keys for text selection on a CRT. *Ergonomics, 21*, 601-613.

[5] Du, Y., Guiard, Y., Chapuis, O., & Beaudouin-Lafon, M. (in press). Assisting Target Acquisition in Perspective View. Proceedings of the British Computer Society Conference on Human Computer Interaction (HCI'06). London (UK), Sept. 2006. ACM Press.

[6] Fitts, P. M. (1954), The information capacity of the human motor system in controlling the amplitude of movement, *Journal of Experimental Psychology, 47,* 381-391. (Reprinted in *Journal of Experimental Psychology: General, 121*, 262-269, 1992).

[7] Furnas, G. W. & Bederson, B. B. (1995). Space-scale diagrams: Understanding multiscale interfaces. *Proceedings of CHI'95*, 234-241. ACM Press.

[8] Furnas, G.W. (1986), Generalized fisheye views. *Proceedings of CHI'86*, 16-23. ACM Press.

[9] Guiard, Y., Beaudouin-Lafon, M. (2004), Target acquisition in multi-scale electronic worlds. *International Journal of Human-Computer Studies 61*, 875-905.

[10] Guiard, Y., Beaudouin-Lafon, M., Bastin, J., Pasveer, D., & Zhai, S. (2004). View size and pointing difficulty in multi-scale navigation. *Proceedings of AVI, Advanced Visual Interfaces* (Pp. 117-124). Gallipoli (Italie), May 25-28, 2004. ACM Press.

[11] Guiard, Y., Bourgeois, F., Mottet, D . & Beaudouin-Lafon, M. (2001), Beyond the 10-bit barrier : Fitts' law in multi-scale electronic worlds, *Proceedings of IHM-HCI 2001* (pp. 573-587). London : Springer.

[12] Guiard, Y., Chapuis, O., Du, Y., & Beaudouin-Lafon, M. (in press). Allowing camera tilts for document navigation in the standard GUI: A discussion and an experiment. *Proceedings of Advanced Visual Interfaces* (AVI 06). Venice (Italy), May 2006. ACM Press.

[13] Igarashi, T., Hinckley, K. (2000), Speed-dependent automatic zooming for browsing large documents. *Proceedings of UIST'00*, 139-148. ACM Press.

[14] ISO, 2002. Reference Number: ISO 9241-9:2000(E). Ergonomic requirements for office work with visual display terminals (VDTs)—Part 9—Requirements for non-keyboard input devices (ISO 9241-9) (Vol. February 15, 2002): *International Organization for Standardization*.

[15] Looser, J., Cockburn, A., and Savage, J. (2005), On the Validity of Using First-Person Shooters for Fitts' Law. *Studies, People and Computers XIX* (Volume 2): British Computer Society Conference on Human Computer Interaction. Edinburgh, Scotland, 33-36.

[16] MacKenzie, I. S. (1992), Fitts' law as a research and design tool in human-computer interaction, *Human-Computer Interaction, 7*, 91-139.

[17] Mackinlay, J.D., Robertson, G. G. and Card., S. K. (1991) The perspective wall: Detail and context smoothly integrated. *Proceedings of CHI'91*, 173-179. ACM Press.

[18] Perlin, K. and Fox, D. (1993), Pad: An Alternative Approach to the Computer Interface. *Proceedings of SigGraph'93*, Anaheim, CA, 57-64. ACM Press.

[19] Ramos, G., Balakrishnan, R. (2005). Zliding: Fluid zooming and sliding for high precision parameter manipulation. Proceedings of UIST 2005, 143-152. ACM Press.

[20] Soukoreff, R. W., & MacKenzie, I. S. (2004). Towards a standard for pointing device evaluation: Perspectives on 27 years of Fitts' law research in HCI. *International Journal of Human-Computer Studies, 61*, 751-789.

[21] Tzavaras, I., Spence, R. (1982), A bifocal display technique for data presentation. *Proceedings of Eurographics'82*, 27-43.

[22] Zhai, S. (2004). Characterizing computer input with Fitts' law parameters: The information and non-information aspects of pointing. *International Journal of Human-Computer Studies, 61*, 791-809

[23] http://java.sun.com/products/javawebstart/

[24] http://www.tele-actor.net/fitts/