Manuscript Version 6. Official copy "Mouse" in *Berkshire Encyclopedia of Human-Computer Interaction*, W.S. Bainbridge (ed).

The Computer Mouse and Related Input Devices

Shumin Zhai
IBM Almaden Research Center

Human input to computer systems is a critical and integral part of any human-computer interaction system. Input should also be designed as such, inseparable from the design of the output (display) components as well as the "interaction techniques" between input and output. This article focuses on the extreme end of input, but the reader is reminded that a good input device has to be compatible with the rest of the entire interactive system.

The most common input device – the computer mouse

The most common input device used today is the computer mouse. The invention of the mouse is commonly accredited to Douglas Engelbart and his colleagues who pioneered interactive computing with their online system NLS. In a legendary live public demonstration of NLS in 1968, at the Fall Joint Computer Conference held at the Convention Center in San Francisco, Engelbart and his colleagues at SRI demonstrated the mouse, along with hypertext, object addressing and dynamic file linking, shared-screen collaboration with audio and video interface and many other innovations.

There are various possible alternatives to the mouse, including touch screens, styli and tablets, joysticks, and trackballs. Various studies, such as (Card, English, & Burr, 1978) have compared the relative merits of these devices. The consensus of that literature is that for interacting with graphical user interfaces (GUI) on desktop computers, the mouse serves most users quite well. The mouse is more direct and more natural than a trackball or a rate controlled joystick, but less fatiguing than touch screens operated by a finger or a stylus. The mouse also enables well-coordinated actions between movement of the cursor and the selection of an object by button clicks, which is a challenge for many other devices.

The underlining technology for the mouse has evolved over many generations. Figure 1 shows Engelbart and his colleagues' mouse prototype with two wheels sensing the horizontal and vertical movements of the mouse respectively. One lasting successful design uses a rolling ball to drive two orthogonal optical encoders for two dimensional movements. Figure 2 shows the inside of a mouse of this type of sensing mechanism. Many design and manufacturing details, such as the size, weight, and location of the ball in the mouse body, affect the quality of use.

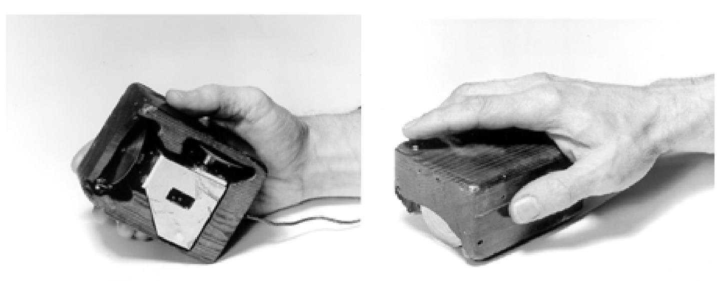

Figure 1. The first mouse prototype circa 1964

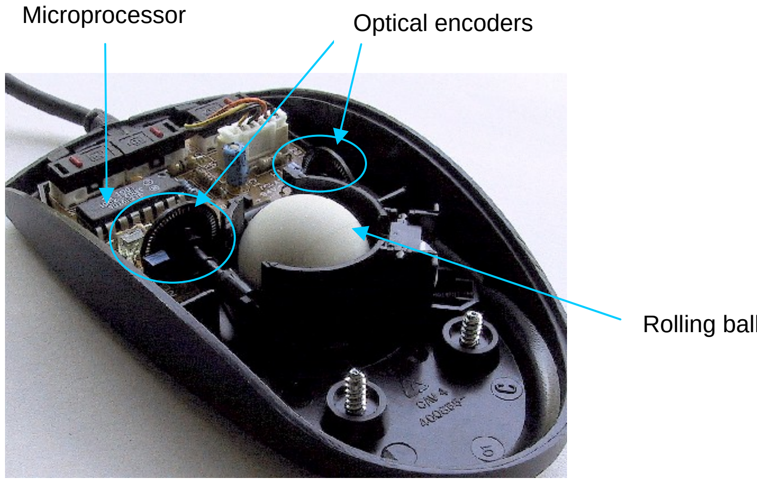

Figure 2. The inside of a typical mechanical mouse

A mouse in fact is a digital computer in itself, equipped with a processor and firmware program that compute the x y displacement in “mickies” based on the number of impulses measured by the optical encoders. The mickies are periodically sent to a host computer in packets based on such standards as PS/2 mouse/keyboard protocol or the USB (Universal Serial Bus) Human Interface Device (HID) protocol.

A more recent technological advancement is the use of a small, high frequency optical sensor (camera) embedded at the bottom of the mouse to measure the mouse’s movement speed by image correlation. Such a solid state design without moving parts is not only less susceptible to dust and debris that tend to be picked up by a rolling ball, but also simpler in assembly complexity. Another recent change is the “tail-less mice” that use wireless communication rather than a cable to connect the mouse with a computer.

The shape and size of a mouse are critical factors to its usability. There are differing views on the design principles that should guide the choice of these dimensions. One view emphasizes that the form should fit the shape of and provides good support to the hand. The products of this view are often called “ergonomically” designed mice. The alternative theory emphasizes that the user should be able to manipulate the mouse with different fingers and hand postures with dexterity and flexibility. It views a locked-in hand shape as a disadvantage. In practice, visually interesting mouse shapes that have a strong “shelf appeal” may dominate design and purchase decisions.

Multi-stream, Multi-hand, and Multi-degree-of-freedom Input

Multi-stream input

The mouse became the de facto standard input device in an era when the operation of the WIMP (windows, icons, menus, and pointer) interface was conceived as completely serial driven by a single stream of input. Today a computer user is no longer limited to a single input stream. In the past few years, document scrolling has been increasingly operated by a dedicated stream of input, such as a separate track wheel (wheel mouse) or a miniature force sensing joystick (e.g. the ScrollPoint™ mouse) embedded on the top of a mouse. See (Zhai & Smith, 1999) and (Hinckley, Cutrell, Bathiche, & Muss, 2002) for two studies on these multi-stream inputs.

Figure 3. Multi-stream mice

Multi-hand input

A more advanced form of multi-stream input should be operated by both hands which may enable a higher degree of efficiency both physically and cognitively (Bier, Stone, Pier, Buxton, & DeRose, 1993; Buxton, 1986; Leganchuk, Zhai, & Buxton, 1998). For example, with two hands each "grabbing" a corner of a graphical object, one would be able to move, rotate and resize it simultaneously, rather than switching between these actions by mode switching as is done in today's common interfaces. An influential theory for designing two-handed input is the bimanual manipulation theory of Yves Guiard (Guiard, 1987). Briefly, Guiard's theory views humans' two hands as cooperative but asymmetrical in function. The non-dominant hand tends to initiate action, operate at a macro scale, and set the frame of reference while the dominant hand tends to operate at a finer scale and complete the operation within the frame of references. A typical example is hammering: the non-dominant hand holds the nail and the dominant hand operates the hammer cooperatively with reference to the non-dominant hand.

Multiple degrees of freedom input

For three dimensional interfaces such as an immersive or a desktop virtual reality system, one would need at least 6 degrees of freedom to be able to fully manipulate an object (x, y, z, pitch, yaw, roll). Various 6 DOF devices have been developed, from free-moving (in the air) devices that is typically based on electromagnetic sensors relative to a base, to stationary force sensitive devices (Figure 4). See (Zhai, 1998) for a review of these devices.

Figure 4. A sample 6 DOF input device -- The Spaceball ™

Mobile input devices

As computing moves beyond the desktop paradigm, alternative input devices to the mouse have been sought. After a period of experimentation with the trackball and other alternatives, touchpads and force sensitive pointing sticks embedded in the keyboard (Figure 3) are the most widely used input devices in laptop computers. The most commonly used touchpads detect the electrical capacitance change under the finger tip. A force sensitive stick transfers finger pressure to two pairs of strain gauge at the base of the stick, which are used to measure the two dimensional force vector applied by the finger tip. The force vector drives the *speed* of a mouse cursor typically by a non-linear force to speed transfer function (hence called rate-controlled joystick) (Rutledge & Selker, 1990).

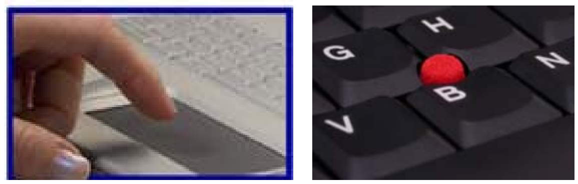

Figure 5. Sample laptop computer input devices: touchpad and Trackpoint™

A pen (a stylus) is a natural input device. In fact, light pens were used as the primary input device on graphic terminals before the mouse became popular. A pen offers greater dexterity than the mouse, enabling users to draw and write. One drawback with using a stylus on a vertical screen is fatigue. Unsupported arm movement becomes tiresome rather quickly. Another subtle but important weakness of the stylus has to do with device acquisition. A mouse (or a touchpad, or an embedded miniature joystick) stays where it is left and can be easily re-acquired by the hand. This is not true with a stylus. If the user has to frequently switch back and forth between typing on a keyboard and pointing with a stylus, device (re)acquisition becomes a hindrance. In fact, the pointing stick integrated between the keys (Figure 5) has a greater advantage than the mouse in this regard.

However, for handheld devices or tablet PCs, the stylus is used to drive most if not all operations. Little device switching is needed. This is one reason that the stylus as an input device is regaining popularity.

Another problem with a stylus is that it, together with the user's hand, may obscure the very object the user needs to look at. This reminds us of a point made earlier – effective input has to be designed together with interaction techniques and output displays. Many of today's usability issues with pen-based interfaces have resulted from "transplanting" existing desktop GUI interfaces to mobile forms.

There have been various successful designs of interaction techniques particularly suited for pen-based interactions. "Marking menus", which uses consistent pen gesture marks defined on nested pie menus that a novice can gesture by looking at the menu but an experienced users can gesture by recall, is one example (Kurtenbach, Sellen, & Buxton, 1993). The "Shorthand Aided Rapid Keyboard" (SHARK) system accommodates novice and experienced users in a similar way: one can tap or trace letters on a graphical keyboard by looking at the keyboard, but a more experienced user can recall the same pattern from memory and simply write the pattern of a word, fully taking advantage of a pen (Zhai, Kristensson, & Smith, 2004).

Sensing and Contextual Input

It is expected that computers will take more input from sensors of various kinds. Some of these sensors will be likely integrated into the mouse. For example, a mouse can be made "touch sensitive", which can display or hide crowding GUI widgets depending on whether a user's hand is on the mouse (Hinckley & Sinclair, 1999). With sensors, computers will be able to take various contextual inputs, such as user presence, posture, physiological variables (heart rate, galvanic skin conductance, EMG etc), or eye-gaze. A critical research challenge is to make computer systems take appropriate actions based on these contextual inputs that help users to achieve their goals.

For example, a user's eye-gaze has long been explored as a source of input (Jacob, 1991). Two fundamental limitations have to be overcome in using eye-gaze for interaction appropriately. First, eye-gaze accuracy is likely limited to one visual degree. Second, the eye, whose movement is driven both by the mind and by the scene, is not a natural control organ, in contrast with the hand. One approach to overcome these limitations is the MAGIC (Manual Acquisition with Gaze Initiated Cursor) system, which takes advantage of the eye-gaze information *implicitly*, without resolving to conscious and unnatural "eye-control". When an input device is touched the mouse cursor appears at the location where the eye is gazing on the screen, hence reducing the need to make large movements by hand to move the cursor to that location on the screen. See (Zhai, Morimoto, & Ihde, 1999) for more details.

Models and Theories

Work on computer input as a human-computer interaction research area has benefited from many basic scientific instruments such as theories, models, taxonomies, and controlled experimentations and observations.

The most frequently used model for input research is Fitts' law (Fitts, 1954), typically expressed as follows:

$$T = a + b \log_2\left(\frac{D + W}{W}\right) \qquad (1)$$

where T is the expected movement time; D and W are target distance and size respectively, and a and b are empirically determined constants that can serve as performance measurements of the input system used. a indicates the reaction time independent of target location and size; b indicates the time increase rate as the index of difficulty (*ID*) of the task increases. $ID = \log_2(D/W + 1)$ takes the unit of bit, analogous to information in communication systems. With the use of Fitts' law, research results (a and b) can be stated and generalized beyond the specific task parameters used in the experiments for testing input performance, enabling a more systematic and more objective understanding of a device's effectiveness.

Fitts' law, particularly suited for characterizing pointing tasks, can be viewed as one of the "laws of action". Recently similar laws of actions have been studied in other human computer interaction tasks such as pen stroking (law of crossing) and path steering on the computer screen. See (Accot & Zhai, 1997; Accot & Zhai, 2002) for more details.

To deal with the perplexing diversity of potential input devices, researchers have often used taxonomies to classify the design space of an input device. (Buxton, Billinghurst , Guiard, Sellen, & Zhai, 1994/2002) and (Card, Mackinlay, & Robertson, 1991) show two examples of such endeavors.

There are also various theoretical views on computer input. We have mentioned Guiard's theory of bimanual manipulation that has guided designs for two-handed input systems, and touched on Buxton's doctrine on facilitating novice to expert transition with common input patterns, as embodied in such methods as marking menus and the SHARK text input system. Jacob and Silbert's theory on "the perceptual structure of input" (Jacob & Sibert, 1992) states that the integration or separation of different degrees of freedom in input should match those of human perception. The "control order compatibility" principle (Zhai, 1995) emphasizes that the transfer function from sensing to display, varying from zero-order position control to first order rate control to higher order control functions, has to be compatible with the mechanical properties of the device. For example, only a self-centering (e.g. isometric device and elastic joystick) can function as an effective rate control device.

Input is a rich and fascinating area of HCI research. This article only touches on its very surface. See related topics in this volume and the further readings for more information.

By Shumin Zhai

Further Readings

Accot, J., & Zhai, S. (1997). Beyond Fitts' Law: Models for Trajectory-Based HCI Tasks. *Proceedings of CHI 1997: ACM Conference on Human Factors in Computing Systems*, 295-302.

Accot, J., & Zhai, S. (2002). More than dotting the i's - foundations for crossing-based interfaces. *Proceedings of CHI 2002: ACM Conference on Human Factors in Computing Systems, CHI Letters 4(1)*, 73 - 80.

Bier, E. A., Stone, M. C., Pier, K., Buxton, W., & DeRose, T. D. (1993). Toolglass and magic lenses: the see-through interface. *Proceedings of SIGGRAPH 93*, 73-80.

Buxton, W. (1986). There is more to interaction than meets the eye: Some issues in manual input. In D. A. Norman & S. W. Draper (Eds.), *User Centered System Design* (pp. 319-337): Lawrance Erlbaum Associates.

Buxton, W., Billinghurst , M., Guiard, Y., Sellen, A., & Zhai, S. (1994/2002). *Human Input to Computer Systems: Theories, Techniques and Technology (book manuscript)*: available at http://www.billbuxton.com/inputManuscript.html.

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.

Card, S. K., Mackinlay, J. D., & Robertson, G. G. (1991). A morphological analysis of the design space of input devices. *ACM Transactions on Information Systems, 9*(2), 99-122.

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.

Guiard, Y. (1987). Asymmetric division of labor in human skilled bimanual action: the kinematic chain as a model. *Journal of Motor Behavior, 19*(4), 486-517.

Hinckley, K., Cutrell, E., Bathiche, S., & Muss, T. (2002). Quantitative Analysis of Scrolling Techniques. *Proceedings of CHI 2002: ACM Conference on Human Factors in Computing Systems*, 65-72.

Hinckley, K., & Sinclair, M. (1999). Touch-sensitive input device. *Proceedings of CHI'99: ACM Conference in Human Factors in Computing Systems.*, 223-230.

Jacob, R. J. K. (1991). The Use of Eye Movements in Human-Computer Interaction Techniques: What You Look At is What You Get. *ACM Transactions on Information Systems,, vol. 9*(no. 3), 152-169.

Jacob, R. J. K., & Sibert, L. E. (1992). The perceptual structure of multidimensional input device selection. *Proceedings of CHI'92: ACM Conference on Human Factors in Computing Systems*, 211-218.

Kurtenbach, G., Sellen, A., & Buxton, W. (1993). An empirical evaluation of some articulatory and cognitive aspects of "marking menus". *Human Computer Interaction, 8*(1), 1-23.

Leganchuk, A., Zhai, S., & Buxton, W. (1998). Manual and Cognitive Benefits of Two-handed Input: An Experimental Study,. *ACM Transactions on Computer-Human Interaction, Vol 5*(No. 4), 326-359.

Rutledge, J., & Selker, T. (1990). Force-to-Motion Functions for pointing. *Proceedings of Human-Computer Interaction - INTERACT'90*, 701-705.

Zhai, S. (1995). *Human Performance in Six Degree of Freedom Input Control*. Unpublished Ph.D. Thesis, University of Toronto.

Zhai, S. (1998). User Performance in Relation to 3D Input Device Design, Computer Graphics. *Computer Graphics, 32*(4), 50-54.

Zhai, S., Kristensson, P.-O., & Smith, B. A. (2004). In Search of Effective Text Input Interfaces for Off the Desktop Computing. *Interacting with Computers, 16*(3), to appear.

Zhai, S., Morimoto, C., & Ihde, S. (1999). Manual And Gaze Input Cascaded (MAGIC) Pointing. *Proceedings of CHI'99: ACM Conference on Human Factors in Computing Systems*, 246-253.

Zhai, S., & Smith, B. A. (1999). An Experimental Study of Document Scrolling Methods. *IBM Systems Journal,, Vol 38*(4), 642-651.