Displays as Advanced Human-Computer Interface

Yoshio Nakatani

Ritsumeikan University, 1-1-1 Noji-Higashi, Kusatsu, Shiga 525-8577, Japan

ABSTRACT

This paper briefly summarizes the historical background of a new human-computer interface used in the current tablet computers and smart phones. Its idea can go back to NLS and Dynabook, both proposed in late 1960s. Their concepts were inherited to the current new devices. A key concept of the next generation human-computer interface may be "user experience." Human motion (gesture) recognition should be introduced, but gestures have different meanings in different cultures. Is the next technology universal or culture-bound?

1. INTRODUCTION

The tablet computers and the smart phones have become very popular. A new human-computer interface called "multi touch" is preferred and a market of touch panel display is rapidly growing. Is the multi-touch completely a new concept? Does this suddenly emerge? The answer is "No." Then, the next question is which idea was its origin and what comes next. This paper overviews the history of human-computer interface technologies and shows that the multi-touch has a close connection to the flow of this history and that the next technology will follow the same stream.

2. HISTORY OF CURRENT STREAM

2.1 WYSIWYG

The important turning point of human-computer interface can far go back to the NLS (oN Line System) by D. Engelbart in 1968 [1]. He invented the mouse, the graphical user interface, and the working hypertext system. The NLS introduced the bit-map display which realized interaction between a user and the system by using the mouse. It enabled the user to edit hyper-linked concept images by using the mouse and to share them with other user on the network. Figure 1 shows the operation of the world first word processor by using the first mouse. Since the NLS, the display has become a "face" of the computer.

The basic concept of the NLS was inherited to the Alto system (Xerox) in 1973. The Alto was equipped with all input and output devices that are equipped on the current desktop computers, which was followed by the Xerox Star workstation. The important concept of the Alto is called the "direct manipulation" which allows users to operate objects on the screen as if they operate them directly in the metaphorical world of desks in an office. Results of inputs are correctly feedback on the screen in real time. This feedback system is called the WYSIWIG (What You See Is What You Get). The new display device was an input device for operation from the very beginning.

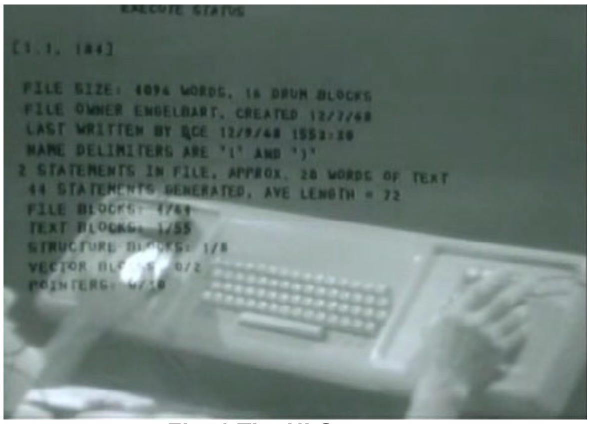

Fig. 1 The NLS system

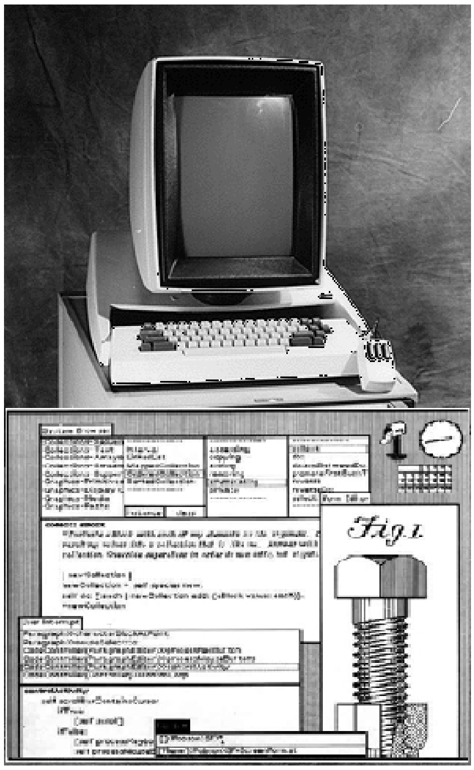

Fig. 2 The Alto system and its screen shot

The total concept of the Alto was then inherited to the Macintosh (Apple computer) in 1984. The Macintosh was the first commercial computer which realized the WYSIWYG interface and is highly evaluated in its easiness-to-use. The limitation, however, should be remembered that the key board and the mouse were the only devices for input. Scrolling was realized by the scroll bar operation, and the selection was realized by the mouse click. Thus, the Macintosh was the first computer which realized the direct manipulation, but it was not operated directly without any devices but indirectly by the mouse.

The Macintosh strongly pushed the human-computer interface forward in terms of software. The most important concept here is “Affordance [2].” The affordance was originally proposed by the perceptual psychologist J. J. Gibson to refer to the actionable properties between an actor (a person and an animal) and the world. The world always provides actors with meanings very essential to their existence and living. A flat stone, for example, affords “sitting” to people. A flat plate on a door affords “pushing.” Of course a stone affords many meanings at once: throwing, patting, kicking, and so on. Perception is a selection among the provided meanings based on the actor’s interests at that time. The same thing is true with artifacts. A well-designed object is self-explanatory and affords its function. Users can use it without deep reasoning. The Macintosh proved that the software, as well as the hardware, also affords meanings. Icons are the good example. Icons graphically represent functions provided on the computer. The icons of the Macintosh by Susan Kare successfully guided the users to the proper functions which they wanted to use [3]. Her icons were very simple compared with the current Windows icons, but directly represented the functions, although there were limitations to its expressiveness. (See Figure 3.)

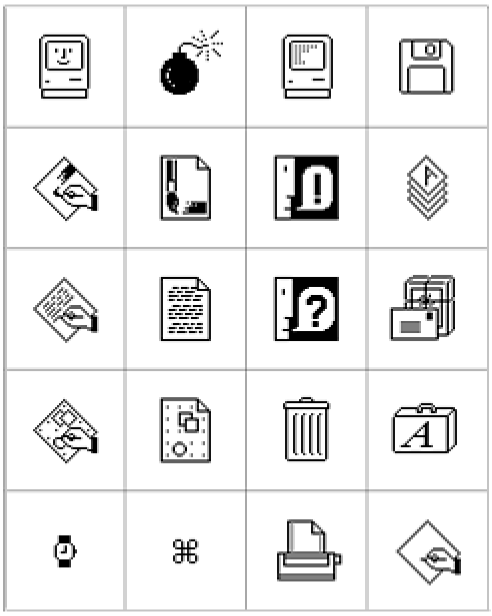

Fig. 3 Icons designed by S. Kare

Clicking an icon with the mouse means starting the corresponding software application. In this sense, the icons need to be recognized as the software switches. Is the software switch the same as the hardware switch? The software switch boxes of the late 1980s tried to imitate the hardware switches by having shadows around the boxes. But pushing the software switches and selecting them with the mouse are quite different. In fact, a mouse click action is not naturally triggered by the icons. Any icons do not afford “clicking.” Connection between an icon and mouse click requires training. This is not the “direct” manipulation. Touching realizes “direct” manipulation.

2.2 Touch-screen

Touch-screen refers to touching the display of the computer with a finger or hand. It dated back to the early 1970s. Before that, electronic instrument builders experimented touch-sensitive capacitance sensors to control the sounds. IBM built the first touch screens in the late 1960s. In 1972, Control Data released the first touch-screen computer PLATO IV, a terminal used for educational purposes that employed single-touch points as its user interface. After this, variety of systems have introduced touch-screen interface, such as ticket vending machines and bank cash dispensers.

Touching is a quite natural action. It does not require any devices. It only requires fingers. Some objects afford touching. The same is true with paintings. The paintings, especially representational paintings, translate a part of the world to canvas. When we are in front of a painting, some of us go closer to it and would like to touch an object which is drawn in it. Well-drawn objects, including good icons, afford touching. The icons symbolize functions. They are abstract, but at the same time, they are representational. If their level of abstraction is proper, that is, if they appropriately tell their meanings, they afford touching. Not all of the icons afford touching. On this point, Susan Kare is highly acclaimed.

The problem of touch-screen from the viewpoint of ergonomics is that significant pressure is required to detect “pushing” certainly while it is not easy because a finger is rather soft. A finger is a broad and ambiguous point of contact with the screen. So icons for touch-screen have to be big and the existing systems with touch-screen interface have rather large displays. Fingernails as stylus can be a solution. The fingernail's hard, curved surface contacts the touch-screen at one very small point, and much less finger pressure is required. This can make icons smaller.

2.3 Mobile Computer

At the same time as the NLS, A. Kay, a major member of the Alto project at Xerox PARC, proposed the concept of the mobile tablet computer, named the Dynabook. The Dynabook aimed at a portable learning environment for children [4]. “Pleasure to use” and “easiness to use” is its core concepts. (See Figure 4.)

The proposed Dynabook was a sketch book or a note rather than a mobile computer. It was planned to giving children access to digital media and to encourage their creativity. Kay supports the idea that a child is trying to acquire a model of his surrounding environment in order

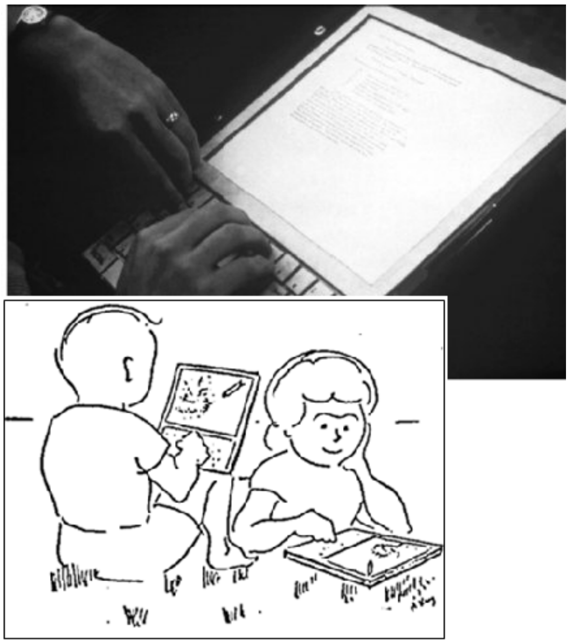

Fig. 4 An Image of the Dynabook

to deal with it and his/her theories are "practical" notions of how to get from idea A to idea B. So from the viewpoint of development support, the Dynabook is not a simple Web search engine but should be an experimental field where children try to operate and edit various kinds of concepts and ideas. The system itself should be operated by a simple and integrated small number of rules and objects, and the users can customize them.

The key concepts of the Dynabook were integrated into the Alto concept. A part of the concepts were inherited to the iPad and the iPhone of Apple computer. Especially, multi-touch realizes the “operation by a small number of rules.”

Considering this historical stream of concepts, the basic ideas behind the current new human-computer interface can be said to be born late in 1960s, although the multi-touch operation is a quite new technology.

3. DISPLAY AND FINGER MOTION

Currently the multi-touch is very popular in the mobile phones and mobile tablet computers. Before discussing on the multi-touch from the viewpoint of human-computer interface, one thing should be discussed. That is, “Does the display afford finger motion?”

The software switch boxes afford pushing, even if they are flat figures. This is because the software switch boxes are recognized analogous to the hardware switches. The existing touch panels use this.

After used to operate the software switch boxes, people could use the icons naturally. Here triggered action is “pushing” icons” at most. Of course, pushing is one of finger actions. How are actions, such as “glide,” “flip,” and “zoom in/out” triggered? Are they natural actions when we look at icons? Can we operate these actions without any explanations and prior information about the multi-touch? Moreover, touching the display with one or more fingers triggers different responses based on context. Can we easily understand what action should be made in a certain context? For example, to unlock the iPhone and start using it, a user has to slide a finger across its face, Can a “unlock” function and a “slide motion across the face” easily connected? Why is a “zoom out” motion unsuitable for unlock?

There are many tasks to be solved here.

4. NEXT GENERATION

4.1 User Experience

Then the next question is “Where does a new idea of the next generation come from?” This question is difficult to answer. A key concept should be “user experience.” User experience was proposed by D. Norman, who also proposed and strongly promotes “user-centered design” [2]. The ISO 9241-210 defines user experience as "a person's perceptions and responses that result from the use or anticipated use of a product, system or service." User experience is subjective in nature and is usually thought to be difficult to find any attractive new concept from it. But new concepts do NOT come from the seeds but from the needs. New ideas are buried in daily lives of people. New ideas can be found in analysis of people’s thoughts and motion. Multi-touch evokes people’s sense of anticipation and makes their user experience exciting. It is quite a new experience and fires their curiosity. People, however, will get bored if the incentive is curiosity alone, Naturalness in daily lives is required.

So we have to come back to “affordance” again. The device has to afford people’s natural daily actions. People use a device for a certain goal. The device is only its tool to realize the goal. The device must not obstruct the goal-oriented activity. If the device requires special operation to its user, it obstructs the goal-oriented activity and the goal will change to get used to the special operation. Device operation has to be natural. “Natural” means that the operation will be triggered by the device at one glance. The user has to use it without manuals and prior information. One of such natural operation will be daily gesture.

4.2 Gesture

People move their hands and bodies when they talk. Gesture is always with talks and is found across cultures and people. It is even found in blind people from birth.

Gesture is a natural action for the speakers to try to express their emotions and thoughts, although it is sometimes a simple action to give a rhythm to speech. Cognitive science finds that gesture accompanied by speech has the potential to display thoughts that are not conveyed in speech [6]. The speech-accompanying gesture serves two functions. First function is a tool for thinking. For example, gesture helps speakers retrieve words from memory. Gesturing while explaining a mathematic task improves performance on a simultaneously performed word recall task. Gesturing increases resources available to the speakers. Another function enables listeners to access to the unspoken thoughts of the speaker. A speaker sometimes sneaks in an idea in gesture unknowingly that does not cohere well with the ideas expressed in speech. Computers are very difficult to recognize such complicated communication, though.

Whilst all people and cultures share the same thoughts and emotions, the forms of communicating them differ from cultures. What is meant in one country can often mean another thing in another country. There are varieties to express the same thing. For example, vertical nodding means “yes” in many countries, but horizontal head shaking means “yes” in India. Moreover, preference to a certain motion differs from country to country. When Japanese people beckon someone, they move the palm down and the hand flapping up and down at the wrist. Many westerners think this rude. The bigness of gesture also differs from culture to culture. In Asian countries, gesture is rather smaller than in European and American countries. Thus while we can express our thoughts and emotions effectively with gesture, its effectiveness has a limitation (See Figure 5.).

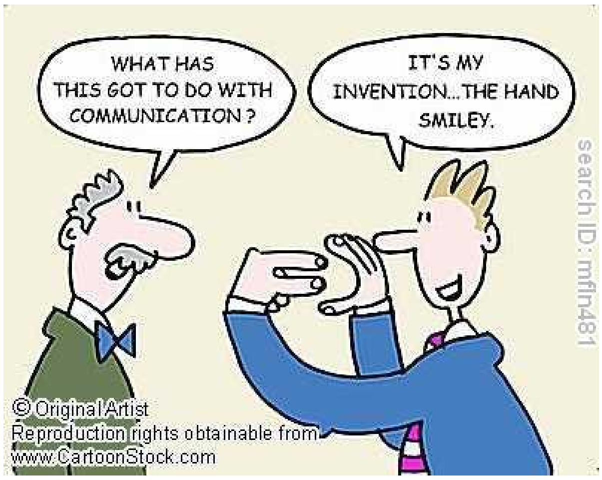

Fig. 5 A Limitation of Gesture

On the other hand, culture-depending gesture can be effectively used if the system has the capability of customizing its recognition of the meaning of gesture. It can widen its gesture-based interface to various culture, context, age, and so on. Good interface can afford natural gesture of each culture.

4.3 Interface of Next Generation

A three-dimensional display system is rapidly accepted. It is certainly innovative, even if it is not always necessary. But how about input methods? Human motion recognition will be rapidly developing as a new human-computer interface which supports people naturally operate computers and mobile phones in their daily lives. Multi-touch certainly opened new horizons for a touch-based input method. There are, however, many people who do not like surface touch of display from the viewpoint of sanitation and delay of reaction. How about possibilities of operation without surface touch? If motions can be sensed at close distance from the display surface, there are many ways to use it by combining with surface touch. Culture-depending gesture can be effective on such display. There has already been advance proposal of such interface. (For example, see [7][8].) The problem is that gesture is limited to small motion of fingers in case of mobile computers. Are there any differences in such small finger gesture from culture to culture? The answer may be “yes.” Differences in finger gesture can be found between boys and girls in Japan. Japanese girls prefer immature gesture, that is “Kawaii” motion. The same differences can be found between cultures. Expressiveness of finger gesture can be developed more richly based on the observation of people’s daily gesture. At that time, can the next-generation display trigger natural finger gesture from the viewpoint of affordance?

REFERENCES

[1] W. K. English, D. C. Engelbart, and M. L. Berman, “Display-Selection Techniques for Text Manipulation,” IEEE Transactions on Human Factors in Electronics, Vol. HFE-8, No. 1, pp. 5-15 (1967).

[2] D. A. Norman, "The Psychology of Everyday Things," Basic Books (1988).

[3] Soojung-Kim Pang, "Interview with Susan Kare," http://library.stanford.edu/mac/primary/interviews/kare/trans.html (2001).

[4] designboom, “susan kare is the queen of icons, ” http://www.designboom.com/portrait/kare.html (2011 Aug 21).

[5] A. C. kay, "A Personal Computer for Children of All Ages," Proceedings of the ACM National Conference, Boston (1972). http://www.mprove.de/diplom/gui/kay72.html (2011 Aug 21).

[6] S. Goldin-Meadow, “The Role of Gesture in Communication and Thinking,” Trends in Cognitive Sciences, Vol.3, No.11, pp.419-429 (1999).

[7] C. Brown et al., “A System LCD with Integrated 3-Dimensional Input Device,” SID Symposium Digest of Technical Papers, Vol.41, Issue 1, pp.453-456 (2010).

[8] T. F.Huang et al., “Tracking Hand-gestures with Depth Detection for a 3D Interactive Display,” SID Symposium Digest of Technical Papers, Vol.42, Issue 1, pp.1830-1833 (2011).