Gesture Based User Interface Experience Evolution and Challenges

B.Sc.(Hons) in Software Development

Kamila Michel

May 2, 2019

Abstract

Recently, the control of devices using gestures has found wide application in various areas of science [2]. Following the changes in the way human interact with electronic devices, we can notice dynamic development in this field. An example of appliances using gestures can be modern televisions, computer games, automation systems, mobile robots, etc. The evolution of electronics and video systems has contributed to the development of a new ways of interaction between human and gadgets, which creates completely new possibilities in terms of design and interface. The integration of vision systems with electronic changes the current approach to the user experience with devices.

1 Introduction

Using the latest technology, we rarely think about how our communication with the machine works. Most often, it adopts an easily digestible graphic form for many so natural that the transition to the text interface, for some people would be an extraordinary difficulty. The interface is such an important element of human communication with the machine, with many identifying it with the computer itself. Changing the interface is associated with

a change in the perception of the computer itself. The interface defines the way in which a person communicates with a computer. Human communication processes run on five levels - intrapersonal, interpersonal, group, mass and extra personal. In the case of human-computer communication, we are dealing with extra personal communication, and therefore non-human communication, with entities other than humans, i.e. with animals or machines. The purpose of this research paper is to bring the reader closer to the evolution of the most popular gesture-based user interfaces. I will focus on those that are mostly used in everyday life by the average user of digital devices.

The term interface is derived from physics, where it means the surface separating the two phases of matter [9]. Adapted for the needs of computer science, this term means a device or program that allows the transfer of information between devices, programs, as well as between a computer and a human.

Throughout history, machine operation has been carried out in very different ways. The oldest machines were controlled using levers and cranks. The next stage was the use of punched cards. Joseph Jacquard was considered the creator of this patent. In 1805 [8] he used plates to control the threads during weaving, which allowed him to create a fabric with a repeatable pattern.

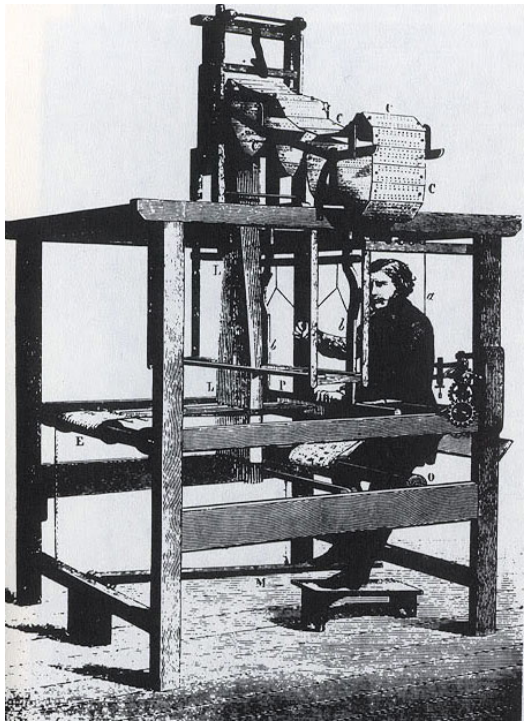

Figure 1: An old Jacquard loom.

Currently, the most common computer used to control the machine is an intermediate link in human-machine contact. Therefore, Human-Computer Interaction (HCI) takes on new meaning and is an indispensable element of our life.

Initially, up to the 1980s, punched cards were used to program computers. Well known to all of us computer mouse was invented by Douglas Engelbart in the early 1960s in his research lab at Stanford Research Institute. It's still used today, allow to move the cursor over the screen by moving it over a flat surface. Of course, an indispensable element is the keyboard, which we use to enter commands and text.

Currently, the touch interface is gaining popularity. Such an interface is associated with touch screens suitable for portable devices. There are several types of construction of such screens. Typically, capacitive and resistive overlays are used for their construction. In other types of touch screens, for example, the interruption of the infrared light emitted by the LEDs is used.

Another example is the multi-touch or multi-touch interface. This technology allows to operate graphical interfaces with more than one finger.

The most natural form of communication with the environment is speech. Hence the idea of using speech to control a computer. In 1952 [4], attempts were made to recognize speech, when K. H. Davis, R. Biddulph, and S. Balashek developed a system for recognizing spoken digits. Recognizing colloquial speech from any person is a very complex problem. Therefore, some simplifications are used, consisting in recognizing the speech of a particular person or narrowing down the recognized set of words.

2 User Experience Evolution

The field of user interface design has been developing very dynamically for years. Progress is driven primarily by the boom in the consumer electronics market, especially mobile, where intuitive handling is a feature that can encourage (or discourage, if not be assured) the use of the equipment on a par with other functions.

Communication with the first computers appeared in the 1960s and was carried out using punched cards and then perforated paper tape [5].

Figure 2: Programmer standing beside punched cards.

This stack of 62,500 punched cards 5 MB worth held the control program for the giant SAGE military computer network.

Along with miniaturization and above all with dissemination computers have changed the operating system. The Command Line Interface (CLI) [10] text mode has appeared. The interface became the keyboard on which short text commands were written. Everyone suddenly had to learn typing. The delay in CLI was reduced to seconds instead of days. The user interface was a series of transactions such as task-response. It allowed to change the user's views on the transaction in response to real-time data.

Figure 3: Teletype Model ASR-33 teleprinter.

The next key user interface progression was the introduction of video display terminals. Over the years, the operating systems has evolved towards the Graphical User Interface (GUI). The service has been enriched with a

manipulator called a computer mouse. The first GUI was developed by researchers at Xerox Palo [14] Research Center (parc) in the 1970s and was the start of a succession of computer graphic innovations to GUIs which has led us to where we are today.

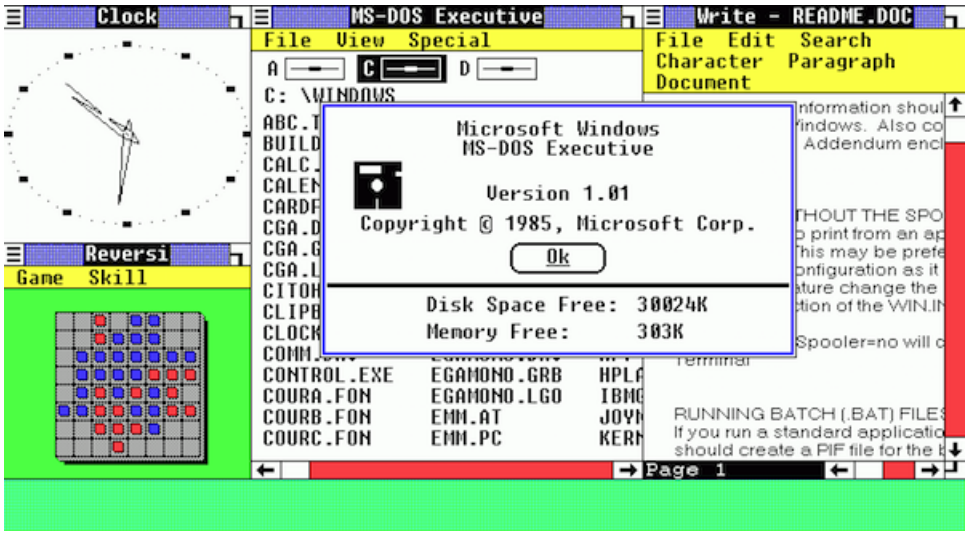

Figure 4: Windows 1.0

The first fully integrated desktop computer was released in 1981, named Xerox Star. Another machine that was released was Apple Lisa Office System 1 (1983), VisiCorp Visi On (1984) and Mac OS System 1 (1984).

Figure 5: Apple Lisa computer

Graphic operating systems have been constantly modified and improved. With the development of a new industry - computer games, type manipulators appeared joystick and gamepad. Both Nintendo [7] in Wii and Microsoft in Xbox 360 had motion detection functions. In order to operate graphical operating systems, we started using special resistive screens and late touch screens.

Figure 6: Nintendo Wii

However, along with the commercial success of touch screens, users have gained the ability to touch digital content without having to engage other input devices. Along with keyboards and mouse cursors, touch and voice inputs are common in both desktop and mobile devices now. The presence of physical buttons can be reduced to a minimum, the best example of which are smartphones and tablets - most of them have a minimal number of physical buttons. Personal computers have the functionality of a touch screen, similarly to portable computers. A process of shortening the distance between human action and its effect in the area of digital reality is revealed. Further hardware interfaces are abandoned for the most intuitive way of communicating with devices. Effective use of the touch screen is not possible without the use of appropriate software, and this is an enlarged user interface.

Figure 7: Touch screens

Another example of gesture based user interface are haptic gloves, VR or AR gloves. They duplicate a well-known pattern of interaction with digital machines and participation in a cyberculture-dominated visuality. Such technologies enable an almost complete and direct connection of a human with the machine by connecting the user's brain with the device or software they control. These technologies have been developed for several decades, but still unknown to the experience of the average user. Already in the 1990s, Derrick de Kerckhove described this phenomenon as the possibility of extending man's senses to artificial vision, hearing and touching, which would give people the ability to experience artificial consciousness, but as he himself noted, the computers at the time did not allow to generate high quality of images in real time. An additional problem was the price and inaccuracy and delays in determining the position of the head of a person who uses virtual reality goggles (Gaudiosi, 2012). Indeed, the VR concept came back after 20 years, when in 2012 a new generation head-mounted display (HMD) type display, Oculus Rifta, was presented. From this moment on the market appeared a dozen or so devices of this type. Both addressed to the owners of PCs (HTC Vive goggles), but also Sony PlayStation VR console games and many mobile systems with Samsung GearVR at the helm.

Figure 8: Haptic Glove for Virtual Reality

Modern, pioneering solutions generally have several disadvantages. The first basic is low reliability. The second, especially in the case of revolutionary solutions, is the necessity of first convincing the user and then the need to train and improve in the process of using new technology or a new interface.

3 Gestures as a communication tool

Experiments of an American scientist in the 1960s discovered the importance of both types of communication - verbal (content) and non-verbal (body language and voice). Based on the research, he developed the principle 7-38-55, which is known today as the rule of Mehrabian [6]. It says that the message consists of three components in certain proportions: 7 percent is words, 38 percent is the tone of voice, 55 percent is body language.

Figure 9: Albert Mehrabians 7-38-55 Rule of Personal Communication

Body language arouses the interest of every human being. It has been the subject of research by scientists since the 1960s. Sometimes the terms "body language" and "non-verbal communication" are used interchangeably, but non-verbal communication is a broader concept, it includes many elements. The body language itself has multi-channel transmission. The channels can be divided into two groups:

- body movements: facial expressions, eye contact, gestures, body position, touch;
- spatial dependencies: the distance that we hold with the interlocutor during the interaction.

The other elements are the duration of the communication process, a group of signals related to movement, touch and the sound of the voice. The division of non-verbal communication into groups is less important at the beginning of learning about the communication process. More important is the knowledge of those elements of body language that directly affect the correct coding and decoding of the message. Body language in human communication is extremely important. Maintaining consistency and proportion between the

importance of verbal and non-verbal communication supports the efficiency of communication and favors reaching an agreement.

3.1 Gestures

Gestures are any body movements made either consciously or unconsciously. Basic gestures are the same for all cultures, such as returning the body's position towards the interlocutor, saluting in the army around world. It happens that the same gesture may have different meanings depending on the culture and nationality of the person who performs it. Sometimes we talk about someone who has so-called Italian pronunciation, which means that it helps with gestures, at the same time we consider the hands, arms, feet, legs, the position of the head and the direction of the entire figure.

Figure 10: Thumbs Up

Thumb raised up, which is a symbol of social networking site Facebook and around the world associated with the words "I like it", means that something is okay. In the Mediterranean countries the importance of the thumb up may be slightly different. Giving the thumbs up to somebody in Greece is equivalent to giving the middle finger to an American. Similarly, in Australia, New Zealand, USA - the thumb is suddenly raised up, when accompanied by an aggressive face grimace, has analogous vulgar significance as in Greece.

Figure 11: Handshakes

In western countries, there is the habit of giving a hand to greet. The host always opens the official meeting. He gives his hand first. If the meeting is on neutral ground, the inviting person plays the role of the host. A firm grip, not very strong, eye contact and a smile give a partner atmosphere of equal rights at a business meeting.

Gesture control has many advantages. In addition to being intuitive and more convenient for the user, it is also worth emphasizing the practical aspect of touchless interfaces - gesture control does not consume the device to the extent that using a traditional keyboard or touch screen. Although properly designed can work without problems for years, regular use of them inevitably leads to their gradual destruction.

3.2 Gesture control - application overview

Gesture control will provide completely new possibilities in the field of human-device interaction, and in many areas where they have been rather limited. Trade is one of them. In this case, gesture control combined with face recognition technology will facilitate establishing relationships with clients. For example, the extension of the display functionality, which currently presents predefined promotional content, the possibility of selecting them by the buyers will make the passive screen so-far become a digital consultant for them. This will increase their interest in the store's offer, which will certainly translate into more sales.

It is also anticipated that gesture control will find many applications in cars, both in interaction with on-board devices as well as outside the car. In the latter case, people can use the security system cameras installed at the rear of the vehicle, providing reverse control and those in the side mirrors, which are part of the system that informs the driver about the appearance of objects in the blind spot. Cameras should automatically start when the proximity sensor detects that the driver is approaching the car. Then he will be able to open the door or the trunk lid of the car. Inside the vehicle, controlling on-board devices using gestures is safer, faster and more convenient than using buttons or knobs, for example to pick up a telephone call or change the settings of multimedia devices that occur in cars.

Examples of Voice User Interface (VUI) are: intelligent Apple Home Pod speakers with Siri's personal assistant, Google Home smart speakers and Amazon Echo speakers with Alex's intelligent personal assistant. This group also includes applications with a voice interface that complement traditional graphical user interfaces, for example in smartphones or smart TV.

4 Challenges for design of applications

The application should be intuitive and friendly. Developers often create software that requires the user to spend a lot of time to learn. The interface can be easy to use for programmers because they know the system well, but the user just wants to do the job without time-consuming preparation (and can quickly change the program to another one if it's too difficult to use). The programmer may think about the program and the user separately, but the user is part of the system and the programmer must take this into account. To understand how people, use digital devices, programmer need to take into account the psychological aspect, the way they perceive things, look for factors that make the user think that the system is a help to him, not an impediment. Thanks to the understanding of Human - Computer Interaction (HMI), the programmer is more likely to create an effective and popular application.

4.1 Voice commands

When designing application that uses speech command, developer cannot rely on similar principles as in GUI design. For human-machine communication to be effective, several conditions must be met. In the examples, illustrating the use of voice, it is worth suggesting to the user to communicate with him as accurately as possible, providing as much information as possible about his query. For people, when communicating, often use shortcuts, which VUI obviously will not understand. The solution is to present a list of functions for each option. It is important that one of them is the option to go out at any time and return to the main menu. It is also important to communicate in a simple visual form to the user that the device "listens" to his message. Otherwise, if VUI reacts only when a person ends speaking, he may not be sure if the device works.

4.2 Response Time

The response time at which the application reacts has a significant impact on usability. In addition, interface should be designed in such a way that the action triggers the response without unnecessary delays. Thanks to this, the user will have a sense of real interaction, such as when zooming in and zooming out the screen, the size of which changes smoothly depending on the spacing of his fingers. The reaction to the gesture should occur immediately, otherwise the user may think the application does not work.

4.3 Human Memory

Another important factor to keep in mind is human short-term memory, which is usually a matter of seconds. To remember something for a longer time, the user must practice (repeat) or record information, for example, save them. Adding many gestures [1] to the application that the user will not be able to remember can frustrated him. There should be an appropriate system provided that correctly recognizes these gestures. Developer must be confident that gestures will be quickly and correctly recognized, but also must provide instructions that allow easy and quick learning of these gestures.

4.4 Excessive amount of information

It is also important that in the case of dynamic transitions do not overwhelm the user with an excessive amount of information displayed on the screen. They must be limited only to the most important ones and do not extend the time of their presentation longer than necessary. For example, information about the current magnification is useful only until the user sets the desired one.

4.5 Range between the device and the user

A short range is usually required for portable devices, such as smartphones or tablets, from a greater distance of the hand movement is controlled by, among others, TV sets and decoders. In both cases the priority is user convenience and the effectiveness of gesture recognition, but the difference in distance has a significant impact on the details of implementation, just like the fact that navigating the smart TV menu, the user makes small gestures, while playing, moves the whole body.

5 Challenges for implementation

For the gesture interface to be used in applications should have a whole range of properties. Below are some categories and their requirements [11]:

5.1 Fast

System equipped with a gesture interface should recognize gestures in real time, without unnecessary delays (delays below 45ms are received by the human being as an immediate response, and over 300ms - as a reaction slowed down) and with a small number of errors.

Challenges: For a machine vision researcher, highly discriminating and computationally efficient functions are interesting, but more techniques must be developed.

5.2 User adaptability and feedback

Property of learning new gestures by memorizing them through the interface, as well as correcting user actions through the interface, i.e. teaching the user the right gestures. In the examples of games for the player, an additional experience may be learning new gestures.

Challenges: Hand gesture [13] systems have a classifier [12] that is offline [3] trained. In order to educate it online, it need a lot of different gestures. It is an impractical process. The challenge is to give the user information non-intrusive way and without increasing the users cognitive load is an additional problem.

5.3 Accuracy(detection, tracking and recognition)

This concept consists of three elements correctly implemented: hand detection, tracking [15] and gesture recognition with the hand or parts of it. The detection checks whether the hand gesture has appeared in the camera. The tracking checks how close it is to the basis of the distance method. Then the diagnosis checks the level of confusion between gestures and others.

Challenges: One of the challenges is detachment due to different shapes, size and color of the hand as well as the lighting in which it is located. Additional difficulty is fast moving and other uncontrolled conditions as, for example, more objects detected.

5.4 Low mental load

Gestures should be simple, short, natural and using the level of human processing habits.

Challenges: The user should use gestures that are simple, quick and natural. The user should remember only a few positions for a specific set of tasks. Icons could help to associate gestures.

5.5 Comfortable

Avoiding gestures that require strong and / or long-lasting muscle-straining. There are two types of stress: static stress and dynamic stress.

Challenges: For stress index measures, experiments vary from subjective questionnaires to electronic devices, such as Electromyograms (EMG), to measure muscle activity. Analytical approaches used to asses stress are based on mathematical models. While mathematical models are not accurate, the electronic devices attached to the users arm disobey the come as you are constraint.

6 Conclusions

In this paper, I described how user's interface evaluated and changed over the years. I focused mostly on gesture-based interface that currently play a dominant role in communication between human and machines. In retrospect, we can say that the first interfaces were quite primitive. They limited the scope of the recipient's influence on the digital environment to text codes, and in an extremely scant way affects the senses of the recipient (because only through luminous controls or simple characters). The analysis reveals the history of the transition from extremely hermetic and inaccessible forms of communication to much more available solutions. It is closer to the human senses and ways of experiencing the physical reality that surrounds us. The ubiquitous visual metaphor of the desk is slowly giving way, and new and constantly developing projects allow people to have almost tangible contact with cyberspace. The user interfaces are not only more user-friendly and user-oriented, but above all, they are getting closer and closer to us. Close to the point that perhaps in the future they will be inseparably connected with us - all due to the coming brain-computer interfaces.

References

[1] Opportunities and challenges for touch and gesture-based systems. https://jnd.org/opportunities_and_challenges_for_touch_and_gesture-based_systems/, 2 June 2013.

[2] Massimo Fascinar. Clemente Giorio. *Kinect in Motion - Audio and Visual Tracking by Example.* April 2013.

[3] Lode Hoste and Beat Signer. Criteria, challenges and opportunities for gesture programming languages.

[4] R. Biddulph K. H. Davis and S. Balashek. Automatic recognition of spoken digits. *The Journal of the Acoustical Society of America*, 24(6), November 1952.

[5] Clive Maxfield. How it was: Paper tapes and punched cards. https://www.britannica.com/science/interface-physics, October 2011.

[6] Albert Mehrabian. *Silent Messages*. Wadsworth, Belmont, Calfornia, 1 edition, 1971.

[7] Wii Nintendo. http://wii.com//, 2006.

[8] Jeremy Norman's. Joseph-marie jacquard's loom uses punched cards to store patterns. http://www.historyofinformation.com/detail.php?id=441.

[9] The Editors of Encyclopaedia Britannica. Interface. https://www.britannica.com/science/interface-physics.

[10] History of Science and Mathematics Beta. When was the first command-line interface developed? https://hsm.stackexchange.com/questions/2339/when-was-the-first-command-line-interface-developed.

[11] Juan P. Wachs, Mathias Klsch, Helman Stern, and Yael Edan. Vision-based hand-gesture applications. commun acm. *Commun. ACM*, 54:60–71, 01 2011.

[12] Noor Adnan Ibraheem Rafiqul Zaman Khan. Hand gesture recognition. *International Journal of Artificial Intelligence and Applications*, 3(4), July 2012.

[13] Anupam Agrawal Siddharth S. Rautaray. Real time hand gesture recognition system for dynamic applications. *International Journal of UbiComp (IJU)*, (1), January 2012.

[14] Usabilla. A short history of computer user interface design. https://medium.theuxblog.com/a-short-history-of-computer-user-interface-design-29a916e5c2f5, May 15, 2017.

[15] M. K. VIBLIS and K. J. KYRIAKOPOULOS. Gesture recognition: The gesture segmentationproblem. *Journal of Intelligent and Robotic Systems*, pages 151, 158, 2000.