Gesture Based User Interface Experience – Evolution and Challenges

"Gesture controls and the challenges of creating intuitive controls"

Gannon, Cian

g00337022@gmit.ie

May 1, 2019

Contents

Abstract   Gesture-based UI has come a long way in the last 10 years. Gesture technology continues to grow slowly year over year with new technologies like Myo and Alexa taking the user interaction with the machine to the next logical solution. In this research paper, I will look into the history of user interaction with computers. Review gestures in everyday life and how gestures are used, and the modern age of gesture-based UI and the challenges faced by developers.

1 User Experience Evolution

1.1 Early Electronic Computers

Computers and how we use them have changed dramatically as each generation is introduced to faster and better machines. How we interact with computers has changed drastically since the first electronic computers in the mid-1900s.

Early electronic computers such as the ENIAC (Electronic Numerical Integrator and Computer). [10] The ENIAC was the first general purpose computer, this meant it could solve many problems as it could be programmed. The ENIAC was the first real step in user interaction with a computer, although the user had to have knowledge of the machine first, they could still operate it using over 3,000 switches and wires which could change how the machine made calculations. [10] Although a primitive method of user interaction with a machine, it was the first 'modern' computer which the user had some control over its operation and change its operation by interaction.
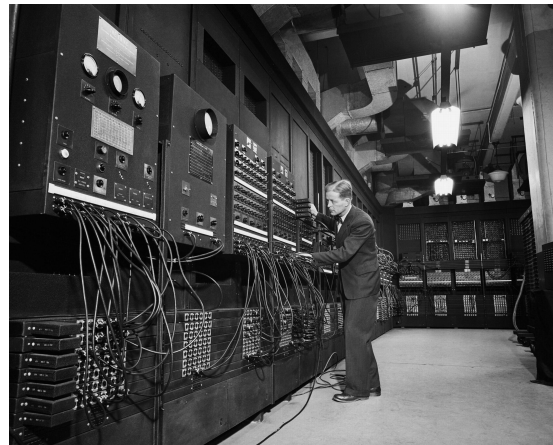

Figure 1.1: ENIAC

Moving on from the ENIAC new computers were built with observations on how to improve over the ENIAC. Machines like the Manchester Baby [4] allowed programs to be saved which was a huge leap in user intractable computers. Once a program was created by a programmer of the machine they could be saved and used again at a later date. This was a big step over the ENIAC where it took time to reprogram the machine each time they wanted to compute a specific task.

As computers advanced and went from large machines that took up an entire room to smaller and smaller devices. During the 1960s to early 1970s, computers started being more accessible by university's and businesses. These general purpose computers allowed the user to interact with the computer using punch cards. [9] Punch cards have been along since long before the computer emerging in the late 1800s. [9] These punch cards allowed computers to be more easily accessible by a large number of people. Students would be able to make a program with these punch cards and use them to give the instructions to the computer. Businesses would also use these punch cards to calculate business related queries.
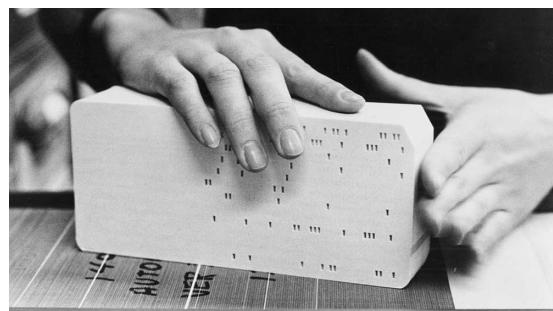

Figure 1.2: Punch card

1.2 Home Computers

Modern computers started when the computer entered the home. [5] These home computers started in the last 1970s with the Apple II, TRS-80 Model I and the Commodore PET. These three home computers created enough competition to keep prices low enough for many users to start getting into the first generation of home computers. These home computers were the first real leap for users interacting with computers. These computers used DOS. [11] DOS (Disk Operating System) was designed to make interaction with a computer easier for users who never interacted with computers before this point. These computers also had a Keyboard to make interaction with

Figure 1.3: Apple II

the computer easier. With the keyboard and a basic OS (Operating System), the home computer was a great success and turned the computer to something that only organizations used to a more personal item.

These new computers allowed users to buy software from vendors that would add extra functionality to their computer through the floppy disk drive. The same floppy disk drive could be used to save data to floppy disks. This allowed the user to easily transfer data between machines.

Later home computer improved in hardware, software and their operating system. Operating systems improved dramatically after hardware manufacturers moved away from DOS. Operating systems such as Windows 3/3.1 and System 1-7 (Classic Mac OS) made user interaction with the home computer easier and allowed multi-tasking by using "tabs" which the user could flick between. These tabs were controlled with new technology for the time, the computer mouse.

The computer mouse was a perfect addition to the home computer as it allowed more refined selection to the new OSs introduced by the newer models of home computers. The mouse paired with the keyboard allowed greater control over computers that couldn't be achieved by just a keyboard.

Over the years operating systems would improve and add more and more features to improve user interact-ability. A great example of the modern operating system is Microsoft. Microsoft has become the largest OS provider to hardware manufacturers. Microsoft has released an OS for nearly every device in every situation. OSs such as Windows 10 for home computers,

Figure 1.4: Apple II

Windows Phone for mobile phones, Windows Server 2019 for servers and legacy OSs such as Windows Embedded Compact for systems such as cash registers. Windows has stayed in the OS market by staying relevant and innovating.
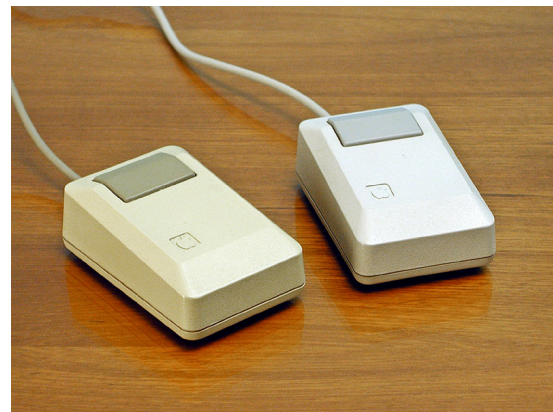
With the latest OSs such as Windows 10 users can now use voice controls via Cortana. This was a very important addition for people who are blind or visual impaired. Cortana allows the users to search the internet, launch applications and generally help the users.

1.3 Touch devices

As computers got smaller and cheaper it leads to companies like Samsung and Apple developing new mobile phones called smartphones. These phones moved away from the traditional full button interface. These smartphones had touch screens which made more room for the screen, which increased its usability.

Modern smartphones have used this touchscreen capability to the next level by adding fingerprint reader into them and adding gesture controls while the phone is off. The user can do gestures such as a two finger swipe down to pause music, or an arrow gesture to play the next song. Face unlocking as also become more popular moving further away from directly interacting with the device.

Figure 1.5: Modern Smart phone

2 Gestures as a Communication Tool

2.1 Origin of Gestures

Gestures have been used as a communication tool since before speech was invented. Gestures convey information rapidly in early human culture. [7] Gestures such as point and facial gestures such as fear would convey danger with a location. Although not very detailed, it quickly conveyed to the rest of the early humans that danger was close. [7]

Gorillas have been seen to use gestures and can be taught new ones. Koko the gorilla is a prime example of how gestures may have been used in early human development. She could use gestures to convey emotions and needs. [1]

Gestures were a vital part of early human communication and continue to be a communication tool in other primate troops.

Figure 2.1: Koko Gorilla Sign Language

2.2 Modern Gestures

Modern gestures differ from culture to culture, so I'll be talking about the gestures used in the western world.

Gestures used today are extremely commonplace. A single gesture might have multiple meaning given the context it's used in. An example is the stop gesture. The stop gesture involves facing your palm of your hand toward the user with fingers placed upwards and close together. This gesture is used to say stop, talk to the hand, giving way to a pedestrian while driving and thanking another driver who has given way to you. [3]

Modern gestures are used everywhere to convey a message quickly. A gesture becomes common place due to many reasons, mostly cultural. A thumbs up in the western world is a common gesture to say 'good job' or just to say 'OK'. This is a common gesture used in the western world but in the middle east and central Africa, it is used to convey frustration with another person by directing the gesture towards them as an insult. [3]

Figure 2.2: Hand Gestures

2.3 Gestures used in the Workplace

Gestures are used in high noise environments such as mining, construction, flying and military service.

The military is a great example of how gestures are used to communicate information where verbal communication is not possible or difficult. The military uses gestures in every branch in every nation to some extent. Military gestures have standard definitions which carry over generations. Military gestures don't evolve like gestures used by the general public for this reason. [2]

Military gestures are used on planes and helicopters to tell solders they are near the drop off point and to tell other solders without shouting out loud to do a given task. [2]

Figure 2.3: Military Gesture

3 Challenges for Design of Applications

The biggest design challenge for developers of gesture controls is market adoption, standardization and gesture choice. From getting users to use the technology to make the technology comfortable to use on a daily basis. Developers both in hardware and software face many challenges in developing with this technology.

3.1 Market Adoption

3.1.1 Myo Armband

One of the biggest challenges of gesture-based technology is market adoption. If the technology isn't adopted by a large segment of the population developers mostly ignore it in favour of a bigger market, such as mobile phones and the traditional desktop.

Devices like the Myo armband started out with great success raising 14.5 Million dollars in 2013 with 30,000 pre-orders for their product. The problem started with lacklustre support from developers who weren't as supportive of the technology as the customers were. [6]

Devices like Myo Armband suffered due to developer support which caused customers to lose interest with the technology.

Figure 3.1: Myo Armband

3.1.2 Kinect

Even industry giant Microsoft tried to break into the gesture controls space with the introduction of the Kinect back in 2010. [12] The Kinect was Microsoft's attempt of adding gesture controls to their gaming console the Xbox 360. With support from Microsoft, many developers made Kinect only games (Kinect Sports, Kinect Adventures!, Kinectimals) and other added Kinect support (Elder Scrolls: Skyrim, Forza Motorsports 4). With support from Microsoft, the Kinect looked to be on the rise as sales reached 10 Million. [8]

Kinect v1 **Kinect v2**

Figure 3.2: Kinect V1/2

From 2010 (Release of Kinect V1) to 2013 (Release of Kinect V2) customers opinion had changed on the Kinect. Many felt the Kinect games felt gimmicky' and games that added Kinect support didn't make using it over a traditional controller better. This dislike of gesture controls came to ahead with the release of the Xbox One, which originally was debuted with a new Kinect, which had to be plugged in at all times. With a price tag 100 dollars more than their competition, many blamed the Kinect for this price difference.

With the Xbox One release having a Kinect market adaption should have been even better than the previous iteration of the Kinect. Unfortunately, the announcement of the Xbox One was a disaster causing lots of controversies that was then directed towards the Kinect. The Kinect got lacklustre support and many had a concern of privacy with an always-on camera and speaker. The Kinect was dropped in later models of Xbox Ones with users having to buy an adapter to get one to work with the newer models. This adapter was then dropped leaving users who enjoyed the connect behind.

3.2 Standardization

A huge problem with gesture controls is standardizing the technology involved. Each company makes its own proprietary technology which leads to lacklustre developer support if the consumer-base isn't high enough. An example is the Xbox Kinect. The Kinect got support from Microsoft with games only for Kinect, which is fine as the game is developed by Microsoft. The problem with support comes from third party companies. These companies want the biggest profit they can get and by creating a Kinect only game it restricts their consumer-base to Kinect only users. Companies would rather release on both Xbox and PlayStation and add in Kinect features. The problem with this style of development is it leads to these Kinect features being thought of after the game has already been made, leading them to be pushed to the side as 'gimmicky' by consumers.

Microsoft is trying to create a standard using their Microsoft Launcher on Android Phones. By allowing the user to customize what gestures do for them, making the gestures more comfortable for the user and by putting this feature on the Android Play-Store giving users around the world access to this feature set.

3.3 Gesture Choice

During the development of gesture-based technology, the developers need to be wary of how the gestures are used. Gestures that mean one thing here means something else over there, so if developers aren't careful they can be met with lots of criticism. A huge part of gesture development is researching the gesture controls to see which gesture would fit best for a particular action.

With less and fewer buttons on phones, many developers have to keep gestures in mind when developing applications for phones. Luckily both Android and iPhone share similar designs allowing developers to carry these gestures across devices, hitting a larger market of users. Phones are a word wide item used by billions. Many people take gestures on phones for granted as we become accustomed to them. If an application doesn't support basic gestures such as swiping up and down to navigate it can make it feel quite old and awkward to use so many will just not use it.

An important part of developing gestures for users is to keep some sort of standard. Many expect to be able to swipe up and down to navigate with one finger, changing this to require two for your application will annoy user away from it. This is why hardware manufacturers like Apple will include how developers should handle certain gestures.

4 Challenges for Implementation

Tracking users with gesture-based technology is one thing, interpreting whether the intended function is meant in another thing. Intent and context is integral to human interaction and especially when using gestures. Intent and context is a very important part of gesture technology moving forward.

4.1 Voice Controls

4.1.1 Context and Intent

A big problem with voice controls is interpreting whether the user is trying to enact a function or the user is just talking to another person. Alexa, for example, will commonly mistake the two people talking about Alexa for them trying to interact with the device. The Alexa machine has difficulty understanding the context of the conversation and activates when hearing the keyword being spoken. This keyword is used to activate the device with the remaining speech being the function the user wishes to enact. Manufacturers are still trying to grasp when the user is talking about the device or trying to use it. Basic workaround such as Calling the device Google Assistant and making the wake word 'Hey Google' is

Figure 4.1: Alexa Active

a basic workaround which has worked for Google, but failed for Amazon. Amazon customers often complain their home assistant thinks they're trying to use it when they are just talking about it. This is a problem with how Amazon advertised their devices under Amazon Alexa and Alexa Compatible which has caused the user to think of the entire device as Alexa rather than an Amazon Echo with Alexa built in.

4.1.2 Interpreting Intent

Interpreting what the user wants is a hard task. There are many ways of phrasing sentences and figuring out what the user is trying to do is a challenging task.

Devices like Amazon Echo have three major methods of figuring out user intent. Alexa will separate a phrase into three parts.

- Intents

- Slots

- Slot Types

The three parts help break down a sentence and figure out what the user wants the device to do. 'Intents' are like functions in programming, 'Slots' are the variables and 'Slot Types' are the variable types. Using these three parts Alexa can break the sentence down

and figure out what has to be done. For example, if the user says 'What is one plus two', Alexa will use the sentence to find an Intent that best matches what the user has asked. Once found Alexa will get the slots in the sentence. The slots are 'one' and 'two'. Alexa will then use slot types to figure out what the slots are. Slot types are numerous ranging from numbers to German cities to American politicians. With the slot type, Alexa can figure out what is being said. Using the slot type in this example Alexa will figure out that 'one' and 'two' are numbers and return '1' and '2' instead of the string representation. Using this method Alexa can figure out most of the user requests.

4.2 Physical Controls

Tracking physical gestures is very tricky because unlike voice gestures there is no wake word to recognize tracking and cease after the function has been enacted.

Microsoft's Kinect for example, its tracking works be imposing a digital skeleton onto the users in the frame in order to track location and limb movement. This method makes for accurate tracking of the entire body and can slowly be expanded upon while also being compatible with previous games and applications.

Microsoft's Kinect is an extremely accurate gesture tracking technology which has been adapted to work in industries beyond its core gaming focus. But its accuracy may cause problems when the user wishes it to stop tracking. Microsoft has come up with multiple ways to fix the problem of gesture tracking when the user wishes to pause the software. The user can do a specific gesture

Figure 4.2: Kinect Tracking

to bring up the Xbox menu in order to pause the game and can use this gesture in order to return to the game. But the main method of interacting with menus and pausing or exiting an application is by using the built-in voice gestures on the Kinect. By using these voice gestures the user can still remain away from any physical input and control any application through gestures. The Kinect can turn on the Xbox console through voice gestures and then can be controlled through gestures both voice and moving your arm.

5 Conclusions

We are moving further and further away from traditional methods of interacting with computers. We've gone from having to use dials and wires, to using a keyboard and mouse. Between these two stages, there was a slow progression to make interacting with computers easier and require less and less knowledge of how a computer works. Gesture-based technology is the next step in human-machine interaction, the questions aren't if it'll happen, but when. In its current state gesture-based technology has faced issues with developer adoption, which ultimately leads to consumer dissatisfaction with the product.

Gestures as a communication tool is one of the oldest methods of communication and will remain as long as two people can communicate with one another. The limitations in gesture technology in the intent and context behind the gesture. Human language requires context and is full of innuendos, which carries over to gesture-based communication. These small innuendos are hard to teach machines, so in the current generation, workaround needs to be used in order to fix this small problem.

Ultimately gesture controls will replace interaction with a computer for most people. The challenge comes with adoption rate with physical gestures as voice gestures are being implemented into device users are already familiar with and have such as smartphones.

Bibliography

[1] Jane Meredith Adams. "Can Koko the Gorilla really talk?." In: *Biography* 3.5 (1999), p. 72. ISSN: 10927891. URL: `http://search.ebscohost.com/login.aspx?direct= true&db=ukh&AN=1851709&site=eds-live`.

[2] J Christopher Brill et al. "A wireless tactile communication system for conveying US Army hand-arm signals". In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. Vol. 50. 20. SAGE Publications Sage CA: Los Angeles, CA. 2006, pp. 2247–2249.

[3] R. Breckinridge Church, Martha Wagner Alibali, and Spencer D. Kelly. *Why Gesture? : How the Hands Function in Speaking, Thinking and Communicating.* Gesture Studies Vol. 7. John Benjamins Publishing Company, 2017. ISBN: 9789027228499. URL: `http://search.ebscohost.com/login.aspx?direct=true&db=e000xww&AN= 1503991&site=eds-live`.

[4] Mary Croarken. "The beginnings of the Manchester computer phenomenon: people and influences". In: *IEEE Annals of the History of Computing* 15.3 (1993), pp. 9–16.

[5] Mary Dee Dickerson and James W Gentry. "Characteristics of adopters and non-adopters of home computers". In: *Journal of Consumer research* 10.2 (1983), pp. 225–235.

[6] Darrell Etherington. "Thalmic Labs Raises 14.5M To Make The MYO Armband The Next Big Thing In Gesture Control". In: *TechCrunch* (). URL: `https://techcrunch. com/2013/06/05/thalmic-labs-raises-14-5m-to-make-the-myo-armband-the- next-big-thing-in-gesture-control/`.

[7] IOSIM IASMINA et al. "THE IMPORTANCE OF CLASSIFICATION AND INTERPRETATION OF MOST COMMON GESTURES IN NONVERBAL COMMUNICATION." In: *Agricultural Management / Lucrari Stiintifice Seria I, Management Agricol* 19.1 (2017), pp. 61–66. ISSN: 14531410. URL: `http://search.ebscohost. com/login.aspx?direct=true&db=bth&AN=123300134&site=eds-live`.

[8] Sarah Kessler. *Microsoft Kinect sales set new Guinness World Record.* Mar. 2011. URL: `http://edition.cnn.com/2011/TECH/gaming.gadgets/03/10/kinect. sales.record.mashable/index.html`.

[9] ENSMENGER NATHAN. "Punch-Card Systems and the Early Information Explosion, 1880—1945 Lars Heide." In: *Technology and Culture* 52.2 (2011), p. 420. ISSN: 0040165X. URL: `http://search.ebscohost.com/login.aspx?direct=true&db= edsjsr&AN=edsjsr.23020594&site=eds-live`.

[10] Abram Taylor. "Electronic Numerical Integrator And Computer (ENIAC)." In: *Salem Press Encyclopedia* (2018). URL: `http://search.ebscohost.com/login.aspx? direct=true&db=ers&AN=89116373&site=eds-live`.

[11] Oklahoma State Dept. of Vocational, Stillwater. Curriculum Technical Education, and Instructional Materials Center. "Disk Operating System–DOS. Teacher Packet. Learning Activity Packets." In: (1991). URL: `http://search.ebscohost.com/ login.aspx?direct=true&db=eric&AN=ED381154&site=eds-live`.

[12] Zhengyou Zhang. "Microsoft kinect sensor and its effect". In: *IEEE multimedia* 19.2 (2012), pp. 4–10.