****

**Gesture Based User Interface Experience –**

**Evolution and Challenges.**

**Aaron Healy**

**G00333148**

**Contents**

Introduction 3

User Experience Evolution 4-5

Gestures as a communication tool 6

Challenges for design of applications 6

Challenges for Implementation 7-8

Conclusion 9

References 10

**Introduction**

The aim of this research paper is to analyse the changes being made in the User Interface as it changes from being a purely physical experience using a mouse, keyboard or controller into an interaction experience using gestures. As technology advances at lightning speed, users are always looking for the next thing to be invented to change their experience. In this paper I will investigate why gestures seem to be the next advancement in user interfaces. I will look at how the user interface has advanced in the past and how this could be done in the present day. I will look at the challenges that are faced to bring this gesture technology to a level where it can be implemented in everyday life.

Following my research, I will give my thoughts on where I think the future of the gesture-based user interface lies and how or if I believe it can be successful. I will discuss the challenges I believe this technology will face and how I believe they can be overcome.

**User Experience Evolution**

To discuss the evolution of the user interface experience I am going to look at the past and how human interaction with computers evolved over the last 70 or so years. The way humans interact with computers has changed so much since the humble beginnings of using the punch card to program early computers in the 1940s and 1950s. The keyboard then became the next big thing. It dominated human interaction with computers for most of the 20th century. However, something happened in 1979 that drastically changed the way humans interacted with computers, Steve Jobs visited Xerox Parc where he found inspiration in the form of a GUI and mouse. These two user interface inventions created a massive change in the way humans interacted with computing technology. Apple sold over a million Macintoshes by 1988 which pushed companies like IBM and Compaq to create their own computer mice.

The keyboard and mouse were then the main form of user interaction with computers. As the 1990s started, the laptop became a popular device and started to overtake the desktop. This brought changes in the design of the mouse and keyboard. Apple invented trackballs and trackpads into their designs. IBM created pointing sticks and incorporated them into their laptops. After this the next physical change to user interaction came with the introduction of a stylus for touchscreen devices. Palm released a stylus for use with their PalmPilot in 1997. This brought along handwriting recognition which was something new and advanced to users.

Voice recognition technology also started to become a form of interaction with the introduction of Dragon NaturallySpeaking in 1997. Dragon sold several million copies of its voice interaction software, although voice recognition did not take off as form of computing interaction until far later in the future when today’s generation of voice assistants became popular. After this user interaction devices seemed to just build on the current hardware available. Slight changes were made to already available technology such as multitouch mice.

The later years of the 2000s saw a huge move forward in user interface design which was led once again by Apple. Touch interfaces became the next big thing and with the introduction of the iPhone became the main way users interacted with their mobile devices. Multipoint capacity touch allowed users to interact with digital content in new ways. All phone manufacturers soon had their own versions of touch screen devices available and revolution of touchscreen technology began. Touch has now become a ubiquitous part of the way we interact with everyday devices such as kiosks, ATMs and house appliances. A technology that only gained mainstream popularity around ten years ago has now become an expectation to have as a minimum on all devices. This is a key example of how humans will always want the latest technology available to them. Once Apple released the iPhone in 2007 with its touchscreen design everyone had to have one. It became the newest fashion accessory as much as it was a mobile phone. The touchscreen technology allowed for new, slicker designs in mobile phones and allowed better user interaction.

While touchscreen technology has been advanced over the last decade to reach unbelievable levels of accuracy and precision, the next advancements in user interaction has already begun with other technologies being improved such as voice recognition and new technologies being created such as virtual and augmented reality spaces to record user gestures. In terms of voice recognition, the technology started to gain mainstream recognition in 2011 with Apple’s Siri voice assistant. Since then a number of devices have been released including the Google Now in 2012 and Amazon’s Alexa devices which were originally released in 2014. These devices are controlled by the user using voice technology. They have become very popular among the current generation as a home assistant. Users can ask them to play music, give them news updates and even control home appliances through them using the latest technology.

Looking at these devices it is easy to see why some consider them essential in modern living. The current generation demand everything to be available at an instant and these voice-controlled devices allow for that. They are marketed in a way that makes you believe they will make your life easier by giving you quicker and easier access to information and there is a viable argument for this. They can automate chores, including vacuuming, grocery shopping, and even cooking. The hands-free technology may also increase autonomy for the elderly and individuals with disabilities. This shows how technology like using a keyboard and mouse to control a computer in the 90s has evolved into an individual controlling multiple device in their home by just using their voice. This improvement and advancement in technology is truly fascinating and that’s coming from a twenty-two-year olds perspective. I can only imagine what some of the older generation think of this new voice-controlled technology. It may not be the main market for the companies making these devices but the older generation being able to control their homes via voice technology is a great positive from the creation of this technology.

Moving on from voice control, the newest user interface technology being experimented with is user gestures such as body movements being recorded, and this data being used to do something. This technology is being used a lot now in computer games and could be viewed as the future of computer games. The Nintendo Wii brought gesture recognition to gaming’s centre stage; the combination of the Wii remote and the Wii sensor bar meant the game console could detect a person’s movements and map them into a 3D computer generated world. Sony PlayStation soon followed by creating its own Move and Eye technology, which are quite like Nintendo’s own gesture recognition work. Not to be outdone, Microsoft’s Xbox 360 Kinect camera took gesture recognition one step farther. Its camera allowed the Kinect to recognise and track 20 individual joints of the human body. It also provided an experience where you could control the Xbox 360 menu with your hands instead of a traditional controller. These three innovations brought gesture recognition gaming into the mainstream and since then major funding has been pumped into the new gesture recognition technology research and development. The result of this investment has been devices such as the Oculus Rift to be created which is a virtual reality headset that allows the player to be transformed into any environment and use their gestures to control their movements in the game. This is the direction gaming is heading and gesture-based user interaction is at the centre of it.

**Gestures as a communication tool**

Gestures are the simplest form of communication between individuals. More than half of all human communication takes place nonverbally. You are constantly sending nonverbal messages. Even when you speak, your listeners judge you and your message based on what they see as well as on what they hear. The way you move and react can tell the person you’re communicating with whether you care, whether you’re being truthful or how well you’re listening to them. Gestures can also be used when verbal communication not used at all. A simple nod when walking past someone on the street is a form of communication through gesture. A wave towards someone when leaving is equivalent to saying goodbye to them.

If we want to look at where gestures came from, we can look at the thumbs up and thumbs down gestures. The thumbs up gesture is known universally as a “yes” or “good” gesture versus the thumbs down being known as a “no” or “bad” gesture. The Oxford English Dictionary records the first positive meaning of thumbs-up in Over the Top, a 1917 book by an American who served in the British army in the first World War, about his experiences at war. In this book it is described as a slang phrase in use by his British counterparts: “Thumbs up, Tommy’s expression which means ‘everything is fine with me.'”. This shows that the simple thumbs up gesture may have just started as a simple way of signalling that everything was okay when verbally speaking wasn’t an option. Over the decades it has made its way into modern culture more and more and now with the likes of Facebook using it as a “Like” button it is only becoming more known. This is the case for most gestures, they started as a way of communicating something between a certain group of people and over time worked their way into other groups eventually becoming a known gesture with a certain meeting. Some of these gestures become more well known than others, some become known all over the world, others just stay in certain countries or regions.

Gestures are used by all us all everyday when communicating with others, however the people we communicate all know what most of our gestures mean, however some gestures can have completely different meanings in different parts of the world. Gestures can be a saviour for travellers who are travelling in a foreign country where there is a language barrier. Gestures can allow them to communicate with locals without speaking the same language, however if they were use a certain gesture that means something positive to them could mean something highly offensive to the locals. For example, In the Philippines, using your hand to make a “come here” gesture is one of the most offensive things you can do. This gesture is deemed so bad that you can get arrested for using it. This proves that although gestures are a great tool for communication, they can also get you in trouble for no wrong doing on your part.

Writing this research paper has allowed me to think about gestures more than I ever would have planned to and what I have discovered in watching other people speak to me or interact with is that gestures are used by every person in every single conversation they have. We have come oblivious to them being used as we see them so often but once you look out for them, they are easy to spot. They are the first level of communication, they can even be used between animals and humans, think of when a trained dog is being told to sit, a gesture is usually being used. Gestures are a key aspect of human interaction which is why it easy to see why technology is being developed to take these gestures and incorporate them into user interfaces.

**Challenges for design of applications**

In terms of designing gesture applications and the correct gestures to use, it is always important to think about the target user. Gestures between different age groups can mean different things. Different regions of the world also have different meanings for different gestures as mentioned previously. Therefore, it essential when designing an application that uses gestures as users will not use an application that takes a lot of learning. The gestures you choose should be as universal as possible so that users will instantly understand why they are doing them. If your application is a game, the gestures should also in some way mimic the action being completed in the game. This way the user will be able to feel like they are gesturing what the player is doing in the game. A simple example of this would be the Wii sports games and the gestures used to make to play tennis or throw a baseball. All were like the action you would do in real life.

Designing gesture-based applications can be different to designing any other sort of application. The whole application must be designed around what gestures are available on the device the application is being designed for. The gestures must be the first thing thought about in the design process otherwise you could end up with an application that is not compatible with the gestures available.

**Challenges for implementation**

In terms of the way gestures are tracked this can depend on the certain device being used. For this research I will look at a few different devices and how they track gestures. I felt this was the best way to research how gestures are tracked. The devices I decided to look at were the Kinect V2, the leap motion controller and the Oculus Rift.

**Kinect V2**

The Kinect V2 uses two cameras, a depth camera and a colour camera. These cameras help track up to six skeletons at once. Each of these skeletons has 25 joints that are tracked. The Kinect skeleton returns joints and not bones. It’s an important difference that can have big implications when you think about how bodies move in space and how the gestures they make will be tracked. The joints are numbered from zero to twenty-four. There are eleven properties associated with each joint. Two for colour, two for depth, three for camera and four for orientation. The Kinect has an infrared sensor which is used with the camera co-ordinates to find 3D points of each joint in space. These are used for joint positioning in 3D projects.

The depth range of the Kinect is eight meters, but the skeleton tracking range is 0.5m to 4.5m. It doesn’t usually find a skeleton closer than 1.5m so the depth will usually fall somewhere between 1.5 and 4.5. The width range can go up to about six meters and the height range will depend on the distance from the camera but can usually find about five meters in height. All these features allow the Kinect v2 to be very good at picking up full body gestures such as the user doing dance moves. The main challenge the Kinect V2 faced in recognising user gestures was being able to recognise smaller gestures such as finger movements. They overcame this by using joints as the points to track. This allows them to track each thumb and tip of the hand to accurately see what gesture is being made.

**Leap Motion Controller**

The leap motion controller is a device that plugs into a computer and can track hand movement. It consists of two cameras and three infrared LEDs. These track infrared light which have a wavelength outside the visible light spectrum. The leap motion controller features wide angle lenses. This allows the device to have a large interaction space of 8ft. This space is in shape of an inverted pyramid. This allows accurate tracking of all hand movements inside this pyramid. Background objects are compensated for, the images are analysed to reconstruct a 3D representation of what the device sees. This is essential as otherwise the device could get confused between background objects moving and hand movements. One problem that the leap motion controller can offer is a difficulty to pick up all hand gestures correctly in harsh light conditions such as bright sunshine. They have not a solution for this problem just a recommendation to use it in locations where the lighting may be like that.

The leap motion can track hand and finger movements very accurately as this is its prime function. It focuses just on the user’s hand movement and no other part of the body. This allows it to be very accurate at getting slight movements as the camera is focused just on the hands.

**KAI**

KAI is a gesture controller designed to make your life easier and more productive. It is used to intuitively track your hand movements and translate them into digital interactions, it does what any pointing device does but just by moving your hand. It is wearable on the hand and doesn’t feature any cameras. It uses sensors to detect movement in the finger and hands to translate into gestures.

The KAI is designed for hand and finger gestures only. Since it has no cameras, range is not an issue, it means you do not need to keep your arms visible to a camera. It also means it is not affected by lighting conditions.

After looking at these three different devices I can see both the positives and negatives of each device and the way they track gestures. The devices that use gestures can track more body movements but are often not as accurate and can be obstructed by light. The sensor-based objects are highly accurate to small movements but can only be used on one part of the body and focus on gestures from one area.

**Conclusion**

Throughout doing this research paper I found out about the current technology being used in tracking user gestures. I learned about how far the technology has come in the years since the main form of communication with a computer was the keyboard and mouse. The latest technology can very accurately track users’ actions and complete actions on the screen based on the gestures. If the technology improves at the speed it has in the last ten years or so then the future of gesture-based user interfaces is looking very good. The more accurate gesture-based devices can get, the better. If a level of accuracy can be reached where every user gesture is read perfectly every time, then I can see a future where every computer is controlled totally by our gestures.

In terms of gaming, I see gestures being the next generation of gaming. Current designs can only be improved with better graphics. The big gaming improvement can only be to bring gesture-based gaming to a high level. Imagine a shooter style game where you are holding a virtual gun in your hands and using gestures to aim and shoot in a realistic style. It is the future of gaming and I predict huge investment in this area in the future. On a final note, I believe gesture-based technology is the next big technology, it would make computing easier, could transform gaming and could be highly usable in-home appliances.

**References**

<https://vicara.co/> - The KAI

<https://mygoodplanet.com/vicara-kai-tech/> - The KAI

<https://www.cmo.com/features/articles/2017/7/20/a-brief-history-of-ui-and-whats-coming.html#gs.3zdfcf> -The History of UI

<https://stevivor.com/news/the-past-present-and-future-of-gaming-gesture-recognition/> - The Past and Present of Gesture Gaming

<https://www.fastcompany.com/90202770/are-smart-assistants-good-for-users-weighing-the-pros-and-cons> - Smart Assistants and Voice Control

<https://socialmettle.com/hand-gestures-in-different-cultures> - Hand Gestures in Cultures

<https://www.jameco.com/jameco/workshop/howitworks/xboxkinect.html> - Kinect Technology

<https://skarredghost.com/2016/12/02/the-difference-between-kinect-v2-and-v1/> - Kinect v2

<https://www.helpguide.org/articles/relationships-communication/nonverbal-communication.htm/> - Non-Verbal Communication

<https://www.leapmotion.com/technology/> - Leap Motion