

Gesture Based User Interface Experience – Evolution and Challenges

Conor Raftery (G00274094), *Software Development (Honours), GMIT*

May 2019

Abstract—The purpose of this paper is to research the User Interface as it moves from purely physical (mouse, keyboard, touch screen) to include intuitive interaction through gestures. This paper delves into the impact of Gesture Based User Interface Experience and its evolution and challenges. The User Experience Evolution and gestures as a communication tool, along with similar topics, are presented within this paper. Other areas that are discussed include the challenges for design of applications and the challenges for implementation.

1 Introduction

Gesture has always been recognized and used as an intuitive and natural interaction. It has been demonstrated that young children can readily learn to communicate with gestures before they even learn to talk. Studies show that from the age of 10 months, babies start to produce simple gestures like pointing and giving [1]. Gesture Controlled User Interfaces (GCUI) use body and hand movements as input controls for an application or system, along with using speech recognition. Initially, gesture interaction has been reduced to a simple command interface. More recently, techniques such as recognizing and interpreting human actions, capturing body movements, and animating virtual humans, have given rise to a number of virtual reality applications with more natural interfaces.

In such applications, the produced gestures can be perceived as an ubiquitous and a natural interface. Gesture technologies have found its way into the everyday household with the introduction of Nintendo's Wii [2] system. Wii makes extensive use of the gesture interface in a wide variety of games, from tennis to bowling, allowing the user to interact in a more natural style than previously, when interaction was controlled via buttons and joysticks which were interfaced to the GUI. Microsoft's Kinect [3] also offers similar interaction without a remote, using the human skeleton for input into the system. Researchers are now

trying to explore even further into 3D virtual environments (VEs) in Human Computer Intelligent Interaction (HCII) [4]. This may allow it to take advantage of both styles of interaction, and so expand the capabilities of the interface.

According to ISO 9241-9 [5], usability is defined as "the extent to which a product can be used by specified end-users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use". The essence of usability engineering is to work iteratively in order to achieve the goal of better usability. However, there is not one single recipe whose application guarantees 100% success. Several authors ([6], [7] and [8]) have gone through these steps and summarized some of the most important issues of usability engineering. The first question of usability engineering is the question of what goals we actually want to achieve. Usability is the essence of a system or program, and Gesture UI needs to adhere to these functional needs in order to prove that its function is necessary.

2 User Experience Evolution

"Originally with just a keyboard and screen, computers have evolved to include a mouse, touch screen, voice control, virtual & augmented spaces and gesture recognition."

To meet the challenges of ubiquitous computing, ambient technologies and an increasingly older population, researchers have been trying to break away from traditional modes of interaction. A history of studies over the past 30 years reported by some researchers [9] suggests that Gesture Controlled User Interfaces (GCUI) now provides realistic and affordable opportunities, which may be appropriate for older and disabled people. Looking back through the brief history, traditional GUI was developed at a time when no one could even think about using Gesture based methods to control their technology, and that touch screen technology is the limit.

"Buttons & Controls can't change. They are there even if you don't need them."
- Steve Jobs [10]

Now it is hopelessly outdated and pulls behind together numerous manipulators. Through recent years, new dimensions have been added to these technologies, and have altered the way we now see and use GUI. Artificial Intelligence (A.I.), for one, has had a major impact in this area [11]. With the evolution of A.I., the emergence of Voice Controlled technologies have emerged. Examples of these technologies can be found in mobile devices and desktops world-wide [12]. The noteworthy voice assistants include 'Siri' and 'Alexa', but unlike the touchscreens, which completely killed the push-button devices market, voice interfaces became only an additional option for interaction with systems.

Along with voice control, other gesture based technology has become available to the everyday user. Take a look at today's current batch of hardware, which use more experimental ways to incorporate gesture control. For example, the Myo Armband [13] uses electrical muscle activity to recognize certain gestures, or the Leap motion controller [14] which reacts on certain gestures by tracking hand skeletal movements. There is no question that, with new and exciting innovative gesture technology, we will become more advanced than ever before.

With all of this on-coming change and strangeness, comes disgruntled users. This is expected as it has proven many times in history that users do not tend to accept change easily [15]. Each generation provides a different challenge to incorporate these new technologies into society. Older generations [16] will initially refuse to use it as they do not find the need for this technology, whereas the newer generations [17] have been brought up within the age of technology, and will want clarification as to how it will aid in their technology endeavour. Until Gesture UI becomes a better polished piece of technology, users will refuse to use it as it will not benefit or aid in their day to day lives.

3 Gestures as a Communication Tool

Our bodily actions are equally strong as our words we use in communication. Movements of hands, face, legs, and different pieces of the body to express something could be either voluntary or an automatic instinctive response, and such activities are known as gestures. Purely expressive display of gestures allow us to convey our feelings, sentiments and thoughts very explicitly while there are certain actions that are only attributed to proxemics. In the latter, gestures are governed by the nature of the physical space we share with people surrounding us. Use of gestures in non-verbal communication is extensive while in verbal communication they are just an addition to words in order to make the communication more effective and appealing. Study of gestures and body language is not only given utmost importance in psychology but also in neurology because the brain plays a key role in controlling our body movements [18]. Let's try to understand how gestures are significant in our everyday life.

A gesture is a form of non-verbal or non-vocal communication in which visible bodily actions convey specific messages, either instead of, or in conjunction with, speech. Gestures incorporate movement of the hands, face, or other parts of the body. [19] Gestures differ from physical non-verbal communication that does not communicate specific messages, such as purely expressive displays, proxemics, or displays of joint attention. Gestures enable people to convey an assortment of emotions and contemplations, from contempt and hostility to approval and affection, often together with body language in addition to words when they speak.

Non-verbal communication goes far in leaving an amazing impact on individuals

you run over in your everyday life. A significant piece of non-verbal communication is gestures, which incorporate body developments from jerking your eyebrows, to squirming with your fingers. Gestures become a powerful tool of communication in the hands of a specialist and now and then, pass on considerably more than words. Gestures establish non-verbal correspondence, which supplement verbal methods of correspondence [19]. They have a characterizing sway on how one gets words and can represent the deciding moment the effect of the verbally expressed word. Signals incorporate any deliberate or unexpected body development made throughout a discussion. In a formal situation, for example, interviews, signals have a major task to carry out, more along these lines, in light of the fact that the hopefuls, in their apprehension, don't focus on the manner in which their body responds. This opens them to according to the general population on the opposite side and empowers them to frame certain impressions about the competitors. For example, pointing your fingers in a formal set-up is viewed as discourteous.

Gestures prove to be useful, particularly to depict you as a certain individual, responsible for his/her self. The correct gestures utilized at the opportune time can upgrade the significance of the verbally expressed words and even add another layer to them. They, actually, become the second line of correspondence, notwithstanding the verbally expressed word. A confident handshake, for example, uncovers an individual's certainty level and self-conviction and tells the other individual that you mean business. A gesture of the head amid a discussion passes on affirmation and demonstrates that you are focussed on the discussion. On the other hand, unseemly gestures can diminish the import of the verbally expressed word and make a negative impression. For example, individuals who nibble their nails, gaze at the ground, sweat because of tension, uncover their absence of certainty just as anxiety, which neutralizes them and leaves a poor impression [18].

The essential thought behind correspondence is to pass on your considerations and have an important talk. Gestures have the ability to go about as viable instruments of correspondence as well. Not exclusively would they be able to be utilized to supplement and bolster what is being spoken, some of the time, they can go about as the essential device of imparting what you think, and do that more successfully than even words. [20] For example, while growing up, we recollect our educators putting a finger on their lips, a flag for the understudies to stay silent. Such is the intensity of gestures.

4 Challenges for Design of Applications

Gestures are an absolute necessity in portable applications, however, it's dependably a test to make them evident for users. Touch interfaces give numerous chances to utilize natural gestures like tap, swipe and pinch to complete things, yet dissimilar to graphical UI controls, gesture-based interactions are frequently

hidden from users. So unless users have prior knowledge that a gesture exists, they won't attempt to use it. Therefore design for discovery is critical. [21] You should make certain you give the correct prompts and visual signifiers that assist the users to how they can interact with an interface.

Instructional exercises, tutorials and walkthroughs are a significant well known practice for motion and gesture driven applications. Consolidating instructional exercises in your application much of the time implies demonstrating a few directions to the client to clarify the interface. In any case, a UI instructional exercise isn't the most exquisite approach to clarify the core functionality of an application. The serious issue with upfront instructional exercises and tutorials is that clients need to recall those actions once the client begins to use the application. [21] Too much information at once might prompt to more confusion and perplexity. For instance, some applications begin with an obligatory multi-page tutorial and clients need to calmly peruse all the data and commit it to their memory. That is awful design since it expects clients to work upfront even before they really attempt to use the application.

With regards to instructing clients to utilize a UI, It is recommended to do so mainly by educating in the context of the action (when a user actually needs it)[9]. So as to show individuals another gesture, you need to begin gradually. Given some iteration, directions can be transformed into an increasingly progressive revelation. Utilize tips to provide a spotlight on explaining a single interaction, as opposed to attempting to clarify every conceivable activity in the UI. Indicate gestures by giving self-evident, relevant pieces of information.

Gestures, usable as they seem to be, would be nothing without animation. Developers and designers could utilize animation to pass on information about available actions or gestures. For instance, in order to make users aware that they can interact with a certain component, a text command could be created right on the interactive component and animate the result of interaction. There are three well known techniques to help instruct users, based on the use of animation.[21] The first is a hint motion. Hint motion, or animated visual hint, shows a preview of how to interact with a component when performing the action. It intends to make a relationship between the gesture and the activity that it triggers. A second technique is content teases. Content teases are subtle visual hints that demonstrate what's conceivable. An example would be to display a 'card' to a user. The card then animates slightly to the left or to the right, showing a 'card' underneath. This indicates to the user that they can access the other card by either swiping left or right. The third and last technique is affordance. You can give a few components of your UI a high affordance to direct clients toward features in an interface, and use bounces or pulses as an indicator of an available gesture. An example of this method can be found in Apple iOS. When a user taps the camera icon, the lock screen bobs up, uncovering the camera application underneath.

5 Challenges for Implementation

Getting the technology to work is hard, yet the actually critical step is getting the human-framework communication right, making it simple for individuals to utilize the frameworks. Here are the issues. Looking at touch screen technology, Contact and detecting innovation is winding up increasingly prevalent, regardless of whether it is on cell phones and tablets, navigation systems, or even cooking appliances. These give extraordinary chances, and obviously, extraordinary chances present great difficulties.[22] Some are technical, however increasingly more they are design and interaction challenges - how to guarantee that the capacities of the innovation are all around coordinated to the requirements and abilities of the general population who use them.

As organizations, like UICO[23], make anticipated capacitance sensing robust under brutal ecological conditions (cold, heat, downpour, snow, through gloves, under substantial vibration, and so forth.) the scope of application domains will extend. These place extreme pressures on implementing decent interaction design. As the technology of touch sensitive applications progress, various opportunities for interaction development emerge. The discussion of the "interface" between individuals and products with the supposition that it is a physical presence, a panel or a general structure with which individuals interact. Indeed, considering something a "touch" or "multi-touch" interface infers a physical structure that is intended to be touched. Be that as it may, as we push ahead, the alternatives extend past insignificant physical touch. We may permit cooperation at any area on a gadget, or even without contacting. The new design considerations must apply to interface inputs that utilize contact or not (touchless) with outputs that can include any medium or sensory modality.[22]

How does an individual realize that the input has been recognized and comprehended by the system? Feedback. To be powerful, feedback must be prompt (under 100 msec., in a perfect world under 50), informative, and clear. The conventional beeps that connote that a touch has been gotten, are insufficient to signify that a more complex multi-touch or gestural input has been received and understood properly. [24] In design, too little consideration is paid to the quality and nature of feedback with complex association modes. Feedback may likewise fuse data about how to turn around the task if the individual trusts that the info was inaccurately translated or that it was wrong. Appropriate, rich feedback is fundamental to make it possible for individuals to distinguish and address errors, regardless of whether the error is with the system or the individual.

Gestures can be perplexing, utilizing an expansive vocabulary including numerous fingers, taps, and movement in a wide range of ways: up, down, left, right, round, tapping, with one, two, three, or four fingers. Include full body movement, speech, eye gaze, posture, and so forth to the scope of things being detected and it turns into an overwhelming test for individuals to gain proficiency with the required actions and after that to recollect them at the fitting occasions and execute them in a way that the system can translate appropriately. To aggravate the troubles, contending companies have built up their very

own language of gestures and interactions, now and again with the end goal of product differentiation, now and then so as to explore the mind boggling brush of copyrights and licenses. [24] The outcome is a perplexing trap of contending collaboration standards delivered by contending makers, causing extraordinary disarray among individuals who need to utilize the products from various sellers. In spite of the fact that the essential wrongdoers are the enormous three - Apple (iOS), Google (Android), and Microsoft (Windows and Phone 8), others are similarly blameworthy. The field is in grave need of standards.

6 Conclusion

As we have seen, Gesture Based UI has been researched and deployed since the late 90's [1]. The journey has started then and we can see the evolution of gesture based systems from the researches and developers over the last couple of years. At first it was with troublesome technologies like sensor, glove and so forth, now it becomes simpler with webcam, image processing software and gaming devices. Poor ease of use was an issue in the beginning period, however at this point it's instinctive, intuitive and natural.[19] In the early research, gesture control or gesture recognition was unpredictable and complex, however at this point it's simple vision technique using hand, head or even whole body gesture. In the beginning, PC application operations was the principle focus. However, at this point it is broadly acknowledged for ambient device and ubiquitous computing. In recent investigations, more focus has been given to controlling home appliances, to utilize cell phone, TV screen interactions, use in work environments, gaming, and even for activities in the household. Another very significant viewpoint is that this technology is now extremely reasonable in terms of cost, while it was quite expensive in the early years.

This paper is the achievement of studying where gesture controlled UI for old and disabled individuals has been researched, alongside the other gesture technologies. From this study, it has been recognized that the impaired and older individuals need more innovation in this area of technology to enhance their well being, [24] given into consideration the limitations of gesture based technology. We can utilize this technological innovation for every day exercises, but cannot rely on it for more advanced tasks.

References

- [1] K. De Bot and R. W. Schrauf, *Language development over the lifespan*. Routledge, 2010.
- [2] T. Schlömer, B. Poppinga, N. Henze, and S. Boll, "Gesture recognition with a wii controller," in *Proceedings of the 2nd international conference on Tangible and embedded interaction*. ACM, 2008, pp. 11–14.

- [3] Z. Zhang, "Microsoft kinect sensor and its effect," *IEEE multimedia*, vol. 19, no. 2, pp. 4–10, 2012.
- [4] T. Moeslund, M. Störring, W. Broll, F. Aish, and Y. Liu, "The arthur system: an augmented round table," *Computer*, vol. 1, no. 1, pp. 277–282, 2003.
- [5] D. Natapov, S. J. Castellucci, and I. S. MacKenzie, "Iso 9241-9 evaluation of video game controllers," in *Proceedings of Graphics Interface 2009*. Canadian Information Processing Society, 2009, pp. 223–230.
- [6] R. W. Pew, "Evolution of human-computer interaction: From memex to bluetooth and beyond," in *The human-computer interaction handbook*. L. Erlbaum Associates Inc., 2002, pp. 1–17.
- [7] J. Nielsen, *Usability engineering*. Elsevier, 1994.
- [8] B. Shneiderman and H. Hochheiser, "Universal usability as a stimulus to advanced interface design," *Behaviour & Information Technology*, vol. 20, no. 5, pp. 367–376, 2001.
- [9] M. Bhuiyan and R. Picking, "A gesture controlled user interface for inclusive design and evaluative study of its usability," *Journal of software engineering and applications*, vol. 4, no. 09, p. 513, 2011.
- [10] T. V. Wilson and W. Fenlon, "How the iphone works," *Retrieved from: on Apr*, vol. 24, no. 2009, p. 9, 2007.
- [11] T. M. Massaro and H. Norton, "Siri-ously? free speech rights and artificial intelligence," *Nw. UL Rev.*, vol. 110, p. 1169, 2015.
- [12] B. Tang, "The emergence of artificial intelligence in the home: Products, services, and broader developments of consumer oriented ai," 2017.
- [13] J. G. Abreu, J. M. Teixeira, L. S. Figueiredo, and V. Teichrieb, "Evaluating sign language recognition using the myo armband," in *2016 XVIII Symposium on Virtual and Augmented Reality (SVR)*. IEEE, 2016, pp. 64–70.
- [14] F. Weichert, D. Bachmann, B. Rudak, and D. Fisseler, "Analysis of the accuracy and robustness of the leap motion controller," *Sensors*, vol. 13, no. 5, pp. 6380–6393, 2013.
- [15] A. L. Russell, "Stages in learning new technology: Naive adult email users," *Computers & Education*, vol. 25, no. 4, pp. 173–178, 1995.
- [16] D. Oblinger, "Boomers gen-xers millennials," *EDUCAUSE review*, vol. 500, no. 4, pp. 37–47, 2003.
- [17] W. Strauss and N. Howe, "Generation z," 1991.

- [18] Y. C. Wu and S. Coulson, “How iconic gestures enhance communication: An erp study,” *Brain and language*, vol. 101, no. 3, pp. 234–245, 2007.
- [19] C. Müller, A. Cienki, E. Fricke, S. Ladewig, D. McNeill, and S. Tessendorf, *Body-Language-Communication*. Walter de Gruyter, 2013, vol. 1.
- [20] R. M. Krauss, Y. Chen, and P. Chawla, “Nonverbal behavior and nonverbal communication: What do conversational hand gestures tell us?” in *Advances in experimental social psychology*. Elsevier, 1996, vol. 28, pp. 389–450.
- [21] M. Brereton and J. Buur, “New challenges for design participation in the era of ubiquitous computing,” *Co-Design*, vol. 4, no. 2, pp. 101–113, 2008.
- [22] R. Z. Khan and N. A. Ibraheem, “Hand gesture recognition: a literature review,” *International journal of artificial Intelligence & Applications*, vol. 3, no. 4, p. 161, 2012.
- [23] A. S. Rath, D. Devaurs, and S. N. Lindstaedt, “Uico: an ontology-based user interaction context model for automatic task detection on the computer desktop,” in *Proceedings of the 1st Workshop on Context, Information and Ontologies*. ACM, 2009, p. 8.
- [24] S. Escalera, V. Athitsos, and I. Guyon, “Challenges in multi-modal gesture recognition,” in *Gesture Recognition*. Springer, 2017, pp. 1–60.