Gesture based natural user interfaces

Daniel Smith
University of Auckland
Dsmi151@aucklanduni.ac.nz

ABSTRACT
Natural User Interfaces have received a lot of attention in recent years. A multitude of touch based devices such as modern phones and tablets have all but become the norm and there is ongoing research into more natural and immersive forms of human computer interaction. The idea of a natural user interface is very broad, covering concepts ranging from voice and touch controlled devices to seemingly futuristic brain wave or thought controlled devices. The standard keyboard and mouse paradigm, while powerful, is quickly becoming replaced or at least supplemented by alternative "natural" modes of interaction. There can be a variety of reasons why a more natural method of interaction is desired. Natural interactions have been credited as being more intuitive and thus easier to learn, more realistic and thus more fun in gaming settings, or simply more practical or safe such as a hands free (voice controlled) phone for use while driving. This report focuses primarily on the use of gestures as a means of interaction with computing devices. The advantages and disadvantages of gestural interaction will explored and real world applications will be looked into. Although other forms of natural interaction will be looked at, deeper exploration of them is outside the scope of this report.

Author Keywords
Natural User Interface; Gestures;

ACM Classification Keywords
H.5.m.

General Terms
Design; Standardization; Performance;

INTRODUCTION
There is currently a large amount of research being conducted in the field of natural user interaction and interfaces. As we gather more and more computing devices around ourselves we look for new ways to interact with these devices. Gone are the days when it is acceptable to spend hours learning a how to interact with a system. The keyboard and mouse are a powerful interaction paradigm but they are not ideal in every situation, consider for example a mobile phone or tablet requiring a keyboard and mouse to function. With the advance of technology and the emergence of affordable alternatives such as motion sensing devices there has been great advancement in what has become known as natural user interaction. References [1][5] state that although these forms of interaction are natural in that they may rely on body actions mimicking real life, they still require a pre defined language and thus are not truly natural interactions. Nonetheless, natural interaction is the term generally accepted for this sort of interaction. The sections in this will discuss the emerging gestural interaction paradigm, possible uses for this paradigm, challenges and obstacles as well as looking into the realistic performance and efficiency of this type of interaction.

CHI'12, May 5–10, 2012, Austin, Texas, USA.

REAL WORLD APPLICATIONS
There are an endless number of real world applications that would potentially benefit from gestural interaction. This report will briefly mention three of these that were covered in the literary review. Problems and solutions presented later in the report will directly relate to these applications.

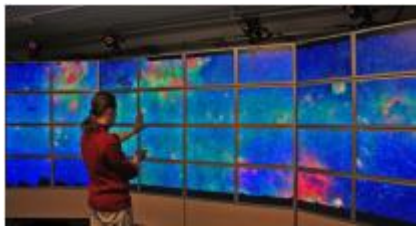

Figure 1: Wall sized display.

Wall sized displays
Reference [2] explores navigating within very large datasets such as very large maps on wall-sized displays such as the one shown in figure 1. This kind of visualization draws its strength from the user's ability to move freely about in front of the display and view large amounts of data at once. This kind of interaction makes it impractical for a keyboard and mouse to be used for navigation as they typically require the user to be seated and stationary. While the keyboard and mouse are usually very good tools for navigating these kinds of datasets on a desktop, an alternative interaction needs to be designed for the same task on these larger displays.

CAD type programs

Reference [3] claims that conventional interaction devices (the keyboard and mouse) are too restrictive for creative work, and that creating 3D designs using these tools can take a long time. [3] Proposes a system based on multi-hand gestures to allow designers to work in a more natural and unrestricted way. It is claimed that this kind of interaction will allow designers to be more creative and expressive in their designs. Using a multi-hand gestural system also provides a simple means of collaboration between designers and there has been some research indicating collaborative tasks benefit from natural user interactions [6].

Figure 2. Multimodal game control.

Games and gaming

Reference [4] proposes and prototypes a game control system which combines a gestural control element with a handheld controller shown in figure 2. Natural mapping between the real world gestures and the in game characters actions make the prototype games easy to learn, and fun to play. It is also worth noting that fitness games make great use of gestural interaction systems and can even use the usually negative side effect of fatigue in a positive way [1].

CURRENT PROBLEMS WITH GESTURAL INTERACTION

Standardization

Any form of human computer interaction requires standards and conventions to be effective. Imagine if a different keyboard with a different layout was required for every application on a desktop, it would be simply impossible for the user to get anything done. Similarly the emerging concept of gestural interaction requires a standard set of commands and gestures, or at least conventions governing how they should be created. Gestural command libraries have many requirements for them to be effective and some of these requirements will be discussed further later in the article. Each gesture should be different enough that the computer is able to correctly interpret what action the user is trying to make, and easy enough to learn and remember that the user is actually able to correctly make the action. Gestures need to be able to provide a high level of control while causing as little fatigue as possible. Reference [3] proposes a framework for gesture based CAD system and identifies the need for a suitable library. Reference [1] claims that standardization is more important even than optimization for the advancement of gestural interaction and likens this to the keyboard – although the layout is far from optimal it is fairly standard.

Fatigue

Consider navigating on a wall sized display such as in reference [2] by pointing at various locations on the screen with a fully extended arm. Although this would be a highly natural and intuitive interaction, as well as being easy for the computer to detect and interpret, holding an arm up in this manner for any extended period of time (even a few minutes) will be highly fatiguing to most people. The mouse and keyboard generally just uses small muscle groups, with the arms in a relaxed and supported position. Small motions of the fingers, especially at a distance, are quite difficult for a camera based system to detect. If the system is intended to be used for extended periods of time then fatigue can be a serious consideration that needs to be taken into account when selecting appropriate gestures for use in the system. Conversely, the aspect of fatigue can be seen as a positive allowing such systems to be used in the context of physical exercise [1].

Intuitivism and learnability

A big part of the drive for natural interaction is the desire for more intuitive systems. A user should be able to act with virtual objects in the same way they interact with real objects [3] [5], without having to learn an arbitrary system of controls. In reality this is never the case and users are forced to learn an artificial language of gestures. References [1] [5] point out that in real life there are real cultural differences in the meaning of gestures from culture to culture, and because of this it is very difficult or potentially impossible to create a fixed gestural language that is truly intuitive. Reference [4] demonstrates a gaming application where gestural control is used and claims that due to the natural mapping between the gestural controls and the actions in game that users were able to instantly understand how to play the game with little or no instruction. Many tasks do not map naturally to gestural interaction, for example in [2] it is observed that the act of panning by pointing is very natural and intuitive but there is no obvious or intuitive solution for zooming.

Tactile feedback and guidance

Purely gestural systems have no form of tactile feedback. This denies the user of any sort of feedback through their sense of touch. In a keyboard or a mouse it is clear when a key press or click has been achieved, but this is not inherently true with a gestural interaction. The lack of tactile feedback also results in a lack of guidance for user actions and this can lead to accuracy issues or difficulties using the system [2]. Consider scrolling the wheel of a mouse compared to making circular motions with the hand in mid air – clearly with the mouse wheel it is much harder

to go wrong. The issue of tactile feedback can be addressed by combining a gestural interface with a handheld device, current game controllers such as the Nintendo Wii already do this and references [2] [4] both report success combining these techniques.

APPROACHES

In their study on panning and zooming on a wall sized display [2] investigates a number of gestural interaction techniques and compares their performance. Three key factors dimensions were identified in design of interaction techniques in this study. Handedness (one handed verses two handed) is a key decision. It is hypothesized that two handed techniques will perform better, but this is not the only factor to consider. If two hands are used for the panning and zooming operations, is there much room to add other operations? Linear verses Circular gestures was the next dimension considered. Circular gestures naturally return back to their start point, while linear gestures can require "clutching", repositioning the hand or finger back to the starting position. It was hypothesized that although linear gestures map better to the task of zooming users would prefer and eventually become faster with circular gestures because of the absence of clutching. The last factor identified was the level of guidance given to the user. This was in the form of an additional hand held device (or absence of one) with varying level of control. The user would hold the device to supplement their gestures in order to control the system. A total of twelve interaction techniques were developed from the combination of these dimensions.

[3] Proposes a framework for a gesture based CAD system and identifies there are currently three main techniques for capturing motion data; gesture recognition systems, glove based devices, and full body systems. The beginnings of a gesture library are suggested and some of this is implemented into a prototype system. The gestures selected should be varied enough for the computer to recognize. Fairly easy to learn or intuitive, and the set of gestures should be able to cover the full range of design activities found in a CAD program.

Similar to [2], reference [4] experiments with the fusion of a physical controller with gesture controlled interactions. The controller itself is tracked (position, orientation) as well as having buttons or controls which the user is able to operate. The user can also make gestures with free hands resulting in actions within the game. Actions that are logical to be based on the motion sensing component of the interaction (e.g. swinging a sword) are mapped to the movement of the controller. Actions that require a higher degree of control, such as movement, are mapped to buttons on the controller. Additional actions can be performed with the free hand to provide a gaming experience that fully engages both hands.

METHODOLOGIES

References [2] and [4] do some analysis of their proposed interaction systems. [3] Identifies the need for heavy usability testing in creation of a full gesture library for their proposed system but does no experimentation or evaluation at this stage. Two prototype games were developed for the interaction concept designed in [4] and these were demonstrated to real users at a trade-show. One game involved wielding the controller like a sword – which mapped to the game character swinging an in game sword with a one to one mapping. The free hand in this game was used to "force push" enemies away from the in game character, this was done by making a pushing motion with the free hand. They also demonstrated a first person based game where the controller was used for movement and shooting while the offhand was used to control a magic shield in game. Observations were made about the usability and learnability of the system. The interaction techniques designed in [2] were compared empirically in a study of 12 participants. Performance time and error rate were recorded as well as observations made about usability and fatigue factors. The techniques were trialed by participants on a display wall consisting of 32 high-resolution LCD screens.

Reference [6] notes the lack of data and analysis into whether natural interface options are actually faster or more efficient than their traditional desktop counterparts. An experiment is devised comparing paradigms in a traditional desktop task: pair-programming. Debugging tasks are devised to be completed by pairs on both desktops and also on multi-touch tables. Participants only do half of the test (they do not use both devices) and the performance between the two groups is measured. Reference [6] also investigates whether the amount of gestures made between the participants is an indication of how "natural" the system is. They also investigate whether an increase in non-verbal communication between participants results in better performance, and whether one interaction mode supports more of this non-verbal communication.

RESULTS

It is found that guidance has a significant effect on performance [2] and that some tasks map more naturally to a handheld device rather than actions in mid air [2][4]. Two handed techniques generally perform better than one handed techniques [2] and in terms of entertainment users find using both hands actively in gaming to be highly enjoyable [4]. Natural mapping has a massive effect on performance with [4] stating that users were able to understand naturally mapped controls without any prior instruction and [2] noting that although linear gestures require "clutching" which should slow them down they still outperformed the clutch-free circular gestures proposed. This is attributed to the linear gesture mapping to the problem more naturally although some of this is explained by the difficulty of performing a circular gesture with the thumb. Overall the best techniques found by [2] involved

two handed control with a high level of guidance as well as linear gestures. One handed gestures with little guidance performed the worst. [4] Found that their second game, which was mapped less naturally and was more complex, took longer for participants to learn, but once they did they found it to be entertaining. [2] Also found that device-free interactions caused the most fatigue.

Figure 3. Gesture controlled prototype game

Reference [6] found that participants using the multi-touch table were able to complete the tasks much faster than those using the desktop. They also found that this corresponded with a greater amount of gestures and non verbal communication between within the pairs. They noted however that pointing with the mouse was not counted as non-verbal communication although this is a very common practice for people sharing a computer. The results obtained in [6] show a significant improvement in terms time to complete the task. No results were gathered regarding the correctness of the answers submitted.

SUMMARY AND FUTURE WORK

There is huge potential for natural user interfaces in a range of real world applications. This report has briefly looked at a few examples focusing on gesture based interactions and gone into some of the challenges associated with creating an effective gesture based system. Though there is still skepticism that these systems can be considered truly natural [1][5] other studies such as [4] have shown that a high level of correspondence between the control system and the systems response allows for a very "natural" interaction that people can understand with little or no training.

Although some concrete results and working prototypes have been discussed in this report [2], [4] they are still very basic systems. Real world systems are much more complex and have more than a couple of controls and it has yet to be seen whether these systems would perform as well with a much larger array of interactions required. [4] Notes that they are just scratching the system with what a gesture based interaction could do for games in general.

Natural interface paradigms may slowly replace or at least supplement the traditional mouse and keyboard but this does not mean the need for GUIs will be removed. [3] Suggests completely moving away from the WIMP paradigm but as noted in [1] without access to visual cues a user is required to remember everything (recognition vs. Recall). Instead interfaces and GUIs will need to be redesigned to take full advantage of the possibilities offered by gestural interaction.

REFERENCES

1. D.A. Norman. 2010. Natural user interfaces are not natural. *Interactions* 17. 2010. http://doi.acm.org/10.1145/1744161.1744163
2. M. Nancel, J. Wagner, E. Pietriga, O. Chapuis, and W. Mackay. Mid-air pan-and-zoom on wall-sized displays. In *Proceedings of the 2011 annual conference on Human factors in computing systems.* CHI '11. http://doi.acm.org/10.1145/1978942.1978969
3. M.F. Shiratuddin, Kok Wai Wong , Non-contact multi-hand gestures interaction techniques for architectural design in a virtual environment, *Information Technology and Multimedia (ICIM), 2011 International Conference on* , vol., no., pp.1-6. 2011. http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6122761&isnumber=6122715
4. D. Ionescu, B. Ionescu, C. Gadea, S. Islam, Multimodal control of virtual game environments through gestures and physical controllers, *Virtual Environments Human-Computer Interfaces and Measurement Systems (VECIMS), 2011 IEEE International Conference on* , vol., no., pp.1-6. 2011. http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6053846&isnumber=6052192
5. A. Malizia, A. Bellucci. The artificiality of natural user interfaces. *Commun. ACM* 55.March 2012. http://doi.acm.org.ezproxy.auckland.ac.nz/10.1145/2093548.2093563
6. A. Soro, S.A. Iacolina, R. Scateni, S. Uras. Evaluation of user gestures in multi-touch interaction: a case study in pair-programming. In *Proceedings of the 13th international conference on multimodal interfaces* (ICMI '11). http://doi.acm.org.ezproxy.auckland.ac.nz/10.1145/2070481.20705