Honours Project Proposal

The Future of Augmented Reality in Computer Games

Dean Lynch

Abstract

Augmented reality has been used to display novelty graphics, 3d models and useful information in applications for a while but it has never been the focal point of a mainstream game. As hardware is becoming more and more portable this project will examine what input methods are available to a device that is using augmented reality as its solo graphical output and will assess those inputs on their suitability to be used as an input method for computer games. The methods that will be assessed are QR cards, hand tracking and a peripheral input. They will be assessed on how easy they are to learn to use and on how accurate they allow the user's inputs to be. This project will develop one or more applications which will allow a group of testers to use easy input method and then feedback from the test group will be used to assess the methods.

1. Introduction

Augmented reality is when a scene or view of the physical world has been augmented by the overlaying of graphics or a graphic effect in real time. Augmented reality should not be confused with virtual reality where a whole scene is replaced with computer graphics or it should not be mistaken for CGI in movies and television shows as though these are a combination of real life scenes and computer graphics they are applied after filming and not in real time.

Currently, numerous augmented reality applications exist for handheld devices such as smart phones or handheld gaming devices. In these, the world is seen through the device's camera but with computer-generated models superimposed onto it. For example the app 'TwittARound' for the iPhone allows you to look through the phones camera and as you scan the world around you it superimposes icons over the camera image which indicate that someone in that direction is 'tweeting' on Twitter and tells you how far away they are and allows the user to press on the icon if they wish to view that persons Twitter feed. More advanced AR applications exist for HMDs and HUDs, Head Mounted Displays (HMDs) and Heads-Up Displays (HUDs) are devices which are worn by the user and are usually built into helmets, they contain a number of displays on the inside which the user views a feed of the outside world captured by cameras on the outside of the device. These devices fully immerse the user allowing them to only see the world through the displays and offer a great deal of detail due to the top-of-the range displays used in them, however, these devices are very expensive and because of this they are restricted to use in the military (Livingston et al. 2011) and medical training.

Recently HMDs have become more available to the public due to several companies starting to develop smaller, more affordable devices such as the Sony 'Personal 3D Viewer Head Mounted Display' along with other devices from Siemens and Sentics.

The question is where does augmented reality fit into the future of computer games? Is it merely a fad, a feature used to display menus and character select screens, or does its future lie with it becoming a core graphical mechanic? Where it is the main and maybe only way graphics would be represented in a game.

One of the challenges of AR is the user input. There are a number of input methods available like QR cards, hand tracking and an input device (e.g. a smart phone's inputs), and with each of these inputs there comes the problems of accuracy and usability. QR cards are cards which have a QR code on them, this code is a symbol like a barcode which is used to represent the position of a 3D model which will be displayed through the AR device, and the virtual model will be moved through manipulation of the QR card. Hand tracking is when the user's hands are followed and their gestures interoperated by the application to interact with virtual objects. An input device such as a smart phone can be used by using the phones sensors (tilt, touch etc) to give the application feedback as to what the user would like to do.

This research will show the strengths and weaknesses of the different input methods by assessing them on their usability and accuracy, then a verdict will be drawn as to whether they are suitable for use as computer game inputs.

2. Research Question

2.1 Specific Research Question

What input methods are available for augmented reality and how appropriate are they for use in computer games?

This question can be split up into two main parts; the methods available and their suitability for computer games. The methods that this project will look at are the use of QR cards, hand and finger tracking and an input peripheral. Each of these input methods will need to be assessed for their suitability for use in computer games, such assessments will have to answer the questions 'How accurate is this method for inputting the user's data and intentions?' and 'How easy is this method for the user to learn to use? Does it take a long time to get used to?'.

2.2 Project Objectives

1. Research the implementation of the three different input methods and learn to use the device environment they will be implemented on.
2. Build three applications which are identical in their task but use different input methods or one application which can switch between the different inputs.
3. Using the applications test each inputs accuracy and usability and assess their suitability for computer games.

3. Literature Review

The project will look at three methods of input Hand tracking, QR cards and use of an input peripheral.

The first method that will be assessed will be hand tracking. The hand tracking method uses the users hand and finger movements to recognise specific gestures that are interoperated into inputs to the application. Hand tracking is well suited to augmented reality (AR) as it requires no extra hardware to be implemented, this is because AR already requires there to be a camera to provide the live feed. Hand tracking can be broken down into two different methods, one that tracks the user's fingers to calculate their inputs by following where they point or the position of the fingers. The other is one that follows the user's whole hand and looks for specific gestures and movements of the whole hand; this method traditionally detects the shape of the hand and uses specific shapes as inputs as well e.g. a 'thumbs up' gesture. Crowley, Berard and Coutaz (1995) use finger tracking as an input in their application to draw and move objects in an application, in their application the finger tracking method looks for a cross section of the viewed area which resembles a reference pattern. This method works well in the application which is run using a computer with a camera and a projector suspended above a table, but it works well because the finger being tracked will always be a set distance from the camera therefore the reference pattern method has little trouble finding the user's finger. With an application such as the one proposed in this project this method would be unwise as with the device being portable the need for the finger being tracked to be a specific distance away finding a surface to input on would be troublesome and even impossible if the user is on the move.

Full hand tracking allows a lot more freedom as it doesn't need the user to be positioned in a specific place or distance from the camera, though it is more computationally expensive, this is because it is not matching a hand to a reference pattern but must find the hand and separate it from the input image then process it to see if there has been any gestures made. The hand tracking method has many steps which make it work, most of which are used to isolate the hand itself so that its gestures can be tracked. The main steps in the hand tracking method are background subtraction, noise removal, separating the skin colour region and selecting the hand gesture region (F.-S. Chen et al 2003). The background subtraction step separates the unneeded background from the foreground where the hand will be, the noise removal step clear up the image, the next step extracts the areas which are skin coloured which is where the hand will be. The last step is to take the image which is left and isolating the hand from that, the hand is found by using motion detection and edge detection. Now that the hand has been isolated, its gesture needs to be interoperated. In the experiment by F.-S. Chen et al (2003) hidden Markov models are used for pattern recognition to identify the hand gesture and their evidence suggests that this is a very accurate way to identify gestures with the possibility of lower error rates if a larger training set for the Markov model.

Another method for identifying gestures is to use feature extraction, this technique is used by Shahzad Malik(2003) and works but finding the peaks and valleys along the perimeter of the skin coloured area which identifies the fingers. Once the peaks and valleys are found the technique can be used to identify different gestures from the number of peaks which appear, Malik's application identifies two different gestures (one pointing and one pinching) but the method could be used to identify more gestures. The advantage this method has over the Markov model method is that it allows the tracking of individual fingers and also it does not require the vast memory needed to contain the Markov training set and doesn't require the program to be 'trained' at all.

The second method that will be assessed will be the use of QR cards as inputs. A QR card is a card with a quick response (QR) code on it, QR codes were developed by Denso Wave a division of Denso Corporation and released in 1994. QR codes are a scan able symbol which contains information in both the horizontal and vertical directions and hold a considerably larger amount of data than a traditional barcode. QR codes come in a range of versions from version 1 to version 40, each version being able to hold a different amount of data (Table 1), QR code data can also be restored if the code is dirty or damaged and can be viewed from any direction(Denso Wave, 2011).

Symbol Size	21 * 21 – 177 *177 modules(size grows by 4 modules/side)

Type and Amount of Data (mixed use is possible)	Numeric	Max. 7,089 characters
	Alphanumeric	Max. 4,296 characters
	8-bit bytes (binary)	Max. 1,817 characters

Table 1

These features make the codes very versatile and due to the large amount of data they can hold it makes them a great choice as an input for AR on a portable devise as data can be held on the code instead of taking up space on the device. The steps involved in receiving input from a QR card are first to identify and tract the card, then to decode the cards code and finally to draw the virtual object over the cards position (Jian-tung Wang et al. 2010). The first step identifies the card by finding the three reference patterns, which look like square bulls-eye patterns, and using them to build a transformation matrix to the position of the card. The bulls-eye patterns are found by identifying all square-like objects in the scene captured by the devices camera and checking them against criteria such as:

"

1. The polygons are not allowed to intersect themselves.
2. The number of vertices of each polygon must be exactly 4
3. The angles of each polygon must be close to 90 degree "

(Jian-tung Wang et al. 2010, p. 416).

The second step is to decode the QR code, this is done by putting the code into a QR decoder, the decoder translates the QR code into whatever data it contains be that a number or letter string which will correspond to a model in the program. The final step is to rotate and transform this model in 3d space and then overlay it over the captured feed, this will use the transformation matrix generated in the first step and should place the model directly on top of the QR card. Another feature of QR codes is that they allow for error correction to correct occlusion or if the code is damaged, this makes QR codes good for a portable device as the cards may get covered up or damaged if being used on the move. QR cards are good for use with portable devices as they aren't required to be places specifically and they can be moved around easily, they do not require any additional hardware to work, only the devices camera which is used for the live feed. P. Rojtberg and A. Qlwal (2010) use QR cards along with a portable device to mark the positions of 3D models in their application, this application shows QR cards working well to mark the positions of virtual objects and also allows the user to mark the position of the light source which directly affects the model. Though this is a limited application it shows the QR cards potential and is open to further development such as a simple cannonball game.

The final method that will be assessed will be the use of a peripheral device, such as a smart phone, as an input device. This method works by taking the data gathered by the smart phone's accelerometer or touch screen and applying that to the application, P. Rojtberg and A. Qlwal (2010) made an application which allows users to use a smart phone's touch screen to add annotations onto a AR scene, there is also an overlay with buttons that the user can use to toggle the annotation mode and to change the position of the lighting which is affecting the virtual models. This input method is good for a portable application as the smart phone has its own processor and can perform some of the calculations, allowing the main device to focus on the graphics while the peripheral deals with all the input calculations. This method would require the smart phone to have a separate application designed to send the data to the main device, this could be wirelessly or the phone could be wired to the main device, the phone outputs the accelerometer data and the touch screen data in a useable numerical way so there would be no excess processing of the data once it reaches the main application. This method can work well with QR codes with the smart phones screen displaying the QR code while the input is received from the accelerometer as the user tilts or turns the device.

4. Research Method

4.1 Research Design

Different input methods will be for implemented in an application; the application will have users perform simple tasks to test the accuracy of the input methods. A group of people will use each method and then they will be questioned to find out their opinions on the usability of the methods.

4.2 Strategy and Framework

The projects final product will be three applications which are practically identical though have different input methods or will be one application which allows the input method to be changed. The applications will be implemented for use on the Sensics SmartGoggles Natalia device, this device is a head mounted display with inbuilt processor, head and hand trackers. The application will ask the user to complete a number of tasks; each will be used to measure the input method's accuracy and usability by timing how long it takes for the user to complete it. The tasks are:

Button Pressing test:

This test will display a row of three buttons labelled A to C (fig. A) and will give the user a sequence they are to enter by pressing the buttons in order.

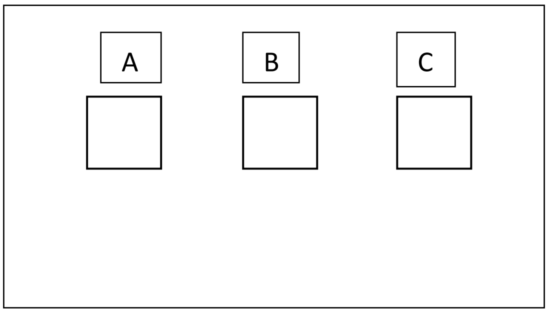

Fig. A

For this test the inputs will be used as follows:

The hand tracking method will have the user hover their hand over the box to select the button.

The QR method will have the user hold the card in the box area to select the button.

The input peripheral method will use a tilt sensor to move a cursor over the box to select.

The Maze Test:

This test will display a maze over a portion of the screen and the user must guide a virtual object through the maze.

For this test the inputs will be used as follows:

The hand tracking method will move the virtual object to co-inside with the movement of the users hand when a specific gesture is made e.g. a thumbs up.

The QR method will have the virtual object follow the movement of the QR card, this method will need the user to position the card at the start of the maze to begin.

The input peripheral method will have the virtual object move in the direction the peripheral's tilt sensor says it is being tipped.

4.3 Data Collection

Data will be collected on the accuracy and usability of the input method by having a test group of volunteers play through the applications tests using each input method and timing how long it takes them to complete each test, this will assess how usable the input method is, and its accuracy. The volunteers will then be asked to fill out a questionnaire asking them to mark each method out of ten in the fields of how easy it was to learn, how accurate they found it, which was the hardest to learn and which one they'd prefer to use. The combination of the timings for the tasks and the questionnaires will allow a conclusion to be drawn as to which was the easiest to learn, which was the easiest to use and which was the most accurate and getting the users intended inputs into the application.

4.4 Analysis of data

The data collected will have to be analysed thoroughly as just the timings of the tasks and the questionnaires alone will not be sufficient at reaching a conclusion, this is due to different users abilities and speeds of getting used to the input methods will be different, also some users may have a preference already which will make them biased. One way to ensure a fair assessment of each method will to split the volunteers into three groups having each one begin on a different input method so that they are not favourable to the one they started on. Another way to avoid spurious results tainting the data is to have a large group of volunteers be assessed, this will help create a better indication of what the average usability rating for each method will be.

5. Significance of Study

This study will help assess AR as a medium for computer games by assessing the suitability of the input methods available for computer games. The study will show people the options available when making AR applications and help choose the right input method for the task. Furthermore the study will show if the input methods available are accurate and usable enough to allow AR to play different types of games as games such as first person shooters and real time strategy games require a lot of accuracy and need the user to be able to respond quickly so easy to use input is a must.

6. Bibliography

Barandiaran, I., Paloc, C. & Graña, M. 2010, "Real-time optical markerless tracking for augmented reality applications",*Journal of Real-Time Image Processing*, vol. 5, no. 2, pp. 129-138.

Chen, F., Fu, C. & Huang, C. 2003, "Hand gesture recognition using a real-time tracking method and hidden Markov models",*IMAGE AND VISION COMPUTING*, vol. 21, no. 8, pp. 745-758.

Crowley, J.L., Berard, F. & Coutaz, J. 1995, "Finger tracking as an input device for augmented reality", *The International Workshop on Face and Gesture Recognition*,Switzerland, Zurich.

Jian-tung Wang, Chia-Nian Shyi, Hou, T.-. & Fong, C.P. 2010, "Design and implementation of augmented reality system collaborating with QR code", IEEE, , pp. 414.
Lee, H., Billinghurst, M. & Woo, W. 2011, "Two-handed tangible interaction techniques for composing augmented blocks",*Virtual Reality*, vol. 15, no. 2, pp. 133-146.

Livingston, M.A., Ai, Z., Karsch, K. & Gibson, G.O. 2011, "User interface design for military AR applications", *Virtual Reality*,vol. 15, no. 2, pp. 175-184.

Malik, S. 2003,"Real-time hand tracking and finger tracking for interaction". [online].cs.toronto.edu. Available from: http://www.cs.toronto.edu/~smalik/downloads/2503_project_report.pdf [Accessed 09 December 2011].

Piekarski, W. & Thomas, B.H. 2003, "Interactive augmented reality techniques for construction at a distance of 3D geometry", *Eurographics Virtual Environments Conference*, Switzerland, Zurich, 2003.

Rojitburg, P. & Olwal, A. 2010, "Tangible Interfaces using Handheld Augmented Reality", *Proceedings of SIGRAD 2010(Swedish Chapter of Eurographics Conference)*,Vasteras, Sweden, 2010, pp.17-26.

Sangyoon Lee & Hong Hua 2011, "Effects of Viewing Conditions and Rotation Methods in a Collaborative Tabletop AR Environment", IEEE, LOS ALAMITOS, pp. 1245.

Uchiyama, H., Saito, H., Servières, M. & Moreau, G. 2011, "Camera tracking by online learning of keypoint arrangements using LLAH in augmented reality applications", *Virtual Reality*,vol. 15, no. 2, pp. 109-117.