IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 4, April 2016

Augmented Reality: Analysis and Challenges in Interaction Techniques

R. Kugan[1], M. Balachandar[2], R.S. Ram Priya[3]

B.E., C.S.E, Anna University- BIT Campus, Trichy, India [1, 2]

Teaching Fellow, Department of C.S.E, Anna University- BIT Campus, Trichy, India [3]

Abstract: Augmented Reality is a trending technology which has applications in many areas such as education, entertainment, advertisement, designing, military etc., Due to the advances in Image Processing and Mobile Computing, the application and access of augmented reality is increasing seamlessly. This paper analyses various Interaction techniques being used in the augmented reality on different platforms and discusses their advantages and limitations. It also includes future possibilities that can be expected in Interaction techniques in augmented reality.

Keywords: Augmented reality, Interactive techniques, Dimension gap, Virtual object.

I. INTRODUCTION

The prime goal of AR is to incorporate virtual content with reality, such that the user gets the effect of reality has also got a lot of barriers, the major one being the interaction. AUGMENTED REALITY is a technology of enhancing while perceiving it. Hence, the interface we choose should the reality by superimposing virtual contents to it. The real-world is augmented with virtual contents like digital 3D models of real-world objects, menus, buttons and other graphical contents using AR [10]. By merging digital data with a live vision of the real world, we can better integrate them and make it feel more convincingly realistic for the user.

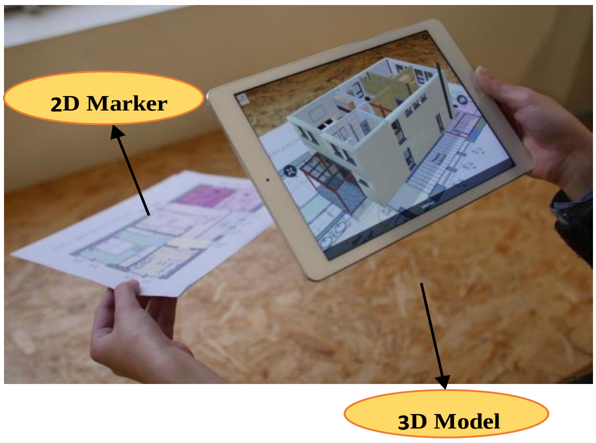

Fig1 AR app showing 3D model of building over a D marker

Fig 1 illustrates about augmented reality application that shows 3D model of building over 2D model.

A wide range of devices, from smartphones to mainframe systems already show support for implementing AR applications. The scope of AR applications in mobile devices is one of the important reasons for the growth of AR. Still, it overcomes the challenges in retaining the feel of reality. This paper will discuss the several methods currently used in interfacing of AR applications and complications in their implementation.

This paper classifies the interaction methods based on various criteria such as the spatial data, devices used and discusses the pros and cons of each method.

II. CLASSIFICATION

A. Based on Spatial Data

The current Human Computer Interaction techniques used in computing devices mostly requires only two dimensional data. Use 2D input data in computing, even while processing 3D content like modeling a 3D object in CAD software. But, in the case of AR which deals with three dimensional space and objects, definitely it requires 3D spatial data. So, the interaction methods can be classified into two wide categories, 3D input devices and 2D input devices.

B. Based on Tangibility

Tangibility of an interface is the attribute by which the user is able to feel its physical existence. Mouse, Pen, Joysticks, Touch Screen are all tangible input devices. They could be touched and felt by` the user. They can also provide haptic feedbacks. The Non-Tangible interfaces are Eye Tracking, Aural UI, Gesture recognition, and GPS data.

C. Based on number of simultaneous users

Based on the number of users simultaneously interacting with AR [2], it can be classified into Single User AR and Collaborative AR. The Interaction method for Single User AR is simple and comparatively easy to implement. But, when it comes to Collaborative AR, there are some difficulties in integrating the inputs of multiple users. Collaborative AR is achieved by either Cloud based collaboration or Co-located collaboration.

D. Based on AR displays

Since, the choice of interaction methods depends upon the type of the display device used in AR, this classification

IJARCCE
ISSN (Online) 2278-1021
ISSN (Print) 2319 5940
International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 4, April 2016

will help to easily comprehend the interaction requirements of each types of device. There are three types of display devices used in AR namely Hand held displays (Fig. 1), Head mounted displays (HMD) and Desktop or Projected display devices. The Hand held displays are mostly found in mobile computing devices [2] [10]. The interaction methods used by them are usually Touchscreens, Gesture recognition, Marker recognition, Voice recognition and GPS data.

The HMD devices either can be displayed mounted at the focal distance of eye or retinal projection devices. Among the other types, this is one of the most powerful user interface but also one of the most difficult one to implement. It gives an immense AR effect and the most suitable interaction method here is gaze tracking which requires to track the pupil dilations and eyeball movements. Hence this method has some complexities in successful implementation [2]. HMD can also be voice controlled. The Desktop or Projected display device use the same interaction methods used by hand held displays except that Touchscreens are rarely used.

III. INPUTS IN AUGMENTED REALITY

A.2D Input

Typically user interaction with graphical computations involving 3D content is achieved by 2D interaction methods like Touchscreen and various pointing devices. The same methods are followed in AR so far. But, these methods are proved to be inadequate in providing the immense effect of reality.

1) 2D Pointing Device

These devices are input interfaces which allow the user to control the movement of a pointer on the screen by providing

2D spatial data from the movement of user's hand or fingers.

Commonly used pointing devices are 2D mouse, Touchpads, TrackPoint and Joystick. Among these devices, the Touchpad can simultaneously track more than one point in the 2D plane.

2) Touch Screen

Touchscreen is an input interface which is a sensor panel coated over the electronic display. The user's touch is detected by change in the capacitance in the panel. Touch Screens with capacity of simultaneously tracking ten points are available in mobile devices now. Multipoint touch helps to create various combinations of input which does various actions like panning, rotating and scaling. Touchscreens are vastly used in hand held devices. Hence, In AR applications touchscreens can be used to manipulate virtual buttons.

3) Eye Tracking

Eye tracking is the technique of mapping the gaze of eyeballs relative to head. Eye tracking is done by three methods namely Eye attached tracking, Optical tracking and Electric potential measurement. In Eye attached tracking, a tracker like a contact lens is attached to the eyeballs and their position is traced [9]. As a tracker needs to be attached to the eye, this method might give an uncomfortable feel to the user. Optical tracking the movements of eyeball is recorded by a video recorder or any other optical sensor. More sensitive optical tracking uses more intricate data such as corneal reflections, retinal veins, eye lens position and pupil center position. Infrared imaging devices are used instead of usual optical sensors in case of low light or dark environment applications. Electric potential measurement method tracks the changes in direction of electric potential between retina and cornea. This Electrooculogram readings can be used to obtain the direction of electric field and hence the eyeball's gaze [9]. This method can also be used in complete darkness. Eye tracking is an excellent choice of interaction method in the case of HMD AR devices.

4) Gaze Tracking

Gaze tracking might seem similar to eye tracking, but this method has a substantial difference that the Gaze tracking simultaneously track both the gaze of the eyeball and that of the head whereas eye tracking only traces eyeball movements [9]. So, this can be considered as an advanced form of eye tracking with the advantage of data about the direction towards where the head gazes.

The data is collected by either magnetic or video based head trackers.

B. 3D Input

1)3D Mouse / Pen

A typical 2D mouse only records two dimensional movements done by the user. But, a 3D Mouse provides data about a third dimension too, thus making it suitable while dealing with real-world 3D content. An example of 3D mouse consists of a typical 2D mouse with a tracker ball in it which gives additional spatial data.

2) Motion capture gloves

Motion capture technology is a widely used technique in Animation Industry. This tech helps to capture the motion of a number of points in 3D space. Motion tracking techniques can be simply categorized into Optical and Non-Optical methods. Optical methods use camera sensors to track the movement of visually distinct points. The Non-Optical methods use inertial sensory systems, mechanical motion measurements and magnetic systems.

Motion capture gloves are used by two means, by using an external camera and making visible points in the gloves, the other one is using gyroscope, magnetometer and accelerometer to measure the movement of points marked in hands [11]. The data acquired from these sensors are integrated to useful 3D coordinates. The accuracy of this system depends upon the quantity and quality of the sensors used. Hence, more number of sensors give more accuracy and hence the cost will rise with it. This method also needs specific software to convert and integrate the data received from the sensors usefully. So, this method offers greater accuracy at high cost.

3)3D Touch

3D Touch or Force Touch was designed by an Apple Inc. This feature adds an extra input parameter to conventional

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 4, April 2016

Touchscreen that it measures the magnitude of pressure applied. This tech can be used in AR applications and this will offer some advantages over usual touchscreens [12].

4) The MATRIX
The MATRIX (Multipurpose Array of Tactile Rods for Interactive eXpression) is a special interactive device to capture complete movement of the human palm and fingers over a bed of spring forced push rod.


Fig 3. MATRIX

As the palm applies sufficient force on the rods, it moves down against the spring force. The hand position was tracked by displacement of each rod in the matrix [3]. The accuracy can be increased with increasing number of rods. This technology is not so cost effective and requires specific software too.

5) Kinect
Kinect is a series of motion sensing devices released by Microsoft for their gaming console XBOX 360. It captures the 3D gesture of the human hand using a camera, a depth sensor and a microphone arranged in an array. Infrared light and CMOS sensor are used for depth measurement. Many attempts are being taken to use this Kinect devices in interacting with AR applications.

IV.CAMERA BASEDSYSTEM

The camera, being a mandatory device in implementing AR, can also be used as an interaction medium. Using camera as an interaction device eliminates the need of any extra devices thus makes the system cheaper and simpler. Image processing plays the vital role in these domains of devices. They can be either Marker based or Marker less. Marker based systems use an easily distinguishable image or pattern as a target and frames different actions based on their movement. Markers less systems use Gesture recognition for interaction. The new Marker less systems developed by Wikitude and Vuforia uses real world objects as Markers.

A. Marker Based
Markers are Image Targets based on which the AR systems work. We can manipulate them to be used as an interaction medium. We can get 3D motion data from Markers and thus overcome Dimension gap which is beneficial over 2D interaction devices.
The Pranav Mistry's "Sixth Sense" uses simple colored bands in fingers for 2D pointing. As depth calculation cannot be done using these simple colored bands, it is limited to 2D interaction. This method is simple and easy to be processed by the device in real-time.
Virtual interaction menus and buttons can generated by masking the main marker. The subordinate Markers can be used as pointing and selecting tool.
3D pointing can be achieved using a 2D image as marker. But this requires so much processing and calculations which needs high end processors to do it in real-time. Even this method gives a good user experience, it is slow particularly when it comes to mobile devices.

B. Hand Gesture
Hand Gesture has been an important area of research for decades for many. Hand Gesture is the most suitable interaction method to use in AR in terms of user experience and cost effectiveness. Hand gesture can be recognized with an ordinary camera available in the common computing devices or the more cutting-edge optical sensors [7]. This section deals with hand gesture recognition with a basic RGB camera. The minimum requirement is that the camera needs capacity of recording at 26 fps. The hand gesture recognition involves various continuous processes running in a loop. 1) Hand detection 2) Hand Shape Analysis 3) Hand trajectory Analysis 4) Gesture recognition.

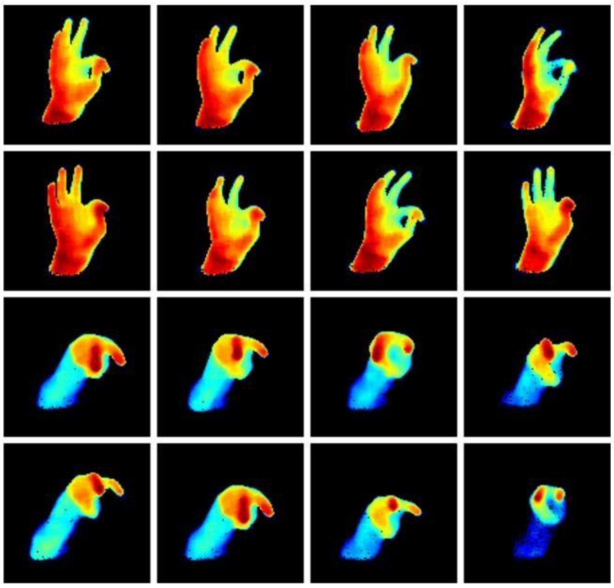
Fig. 4.Sample Hand gesture detected

Hand detection comprises of steps such as edge detection and skin color filtration. Common problems in these procedures are time complexity of algorithms and requirement of ambient light conditions. Plain background without any visual noises will yield the best outcomes which is not the typical case.
Shape Analysis is used to find the pose of the hand after the hand was detected. The pattern of hand pose is matched with the database contains various pattern and in advance machine learning process like neural network can be used for this purpose. Current hand gesture techniques which mostly work on simple gestures like waving and

DOI 10.17148/IJARCCE.2016.54176

showing different postures of fingers which give only planar data. So, the Dimension gap issue strikes here again.

V. AURAL USER INTERFACE

The ability of a machine or program to understand and execute the spoken command is known as speech recognition. In order to help the disabled people speech recognition is combined with AR. In proposed system AR engine is used to display text in AR environment [13]. The features of the system for these disabled person is that

- The replacement of markers with the face detection techniques for narrator's detection.
- To capture the narrator's voice AR engine is used which is back upped with the Audio Visual Speech Recognition (AVSR) techniques.
- Disabled people uses Text To Speech (TTS) engine to talk to the narrator.
- The brief description of the architecture and the working of the system are illustrated.

VI. CONCLUSION

Even though Augmented Reality is not a well-established domain yet, it definitely exhibits great potentials for future applications. So, AR will continue to be an area of unceasing research and this process can be accelerated with more practicable technologies like inexpensive and compact devices with impressive performance. When it comes to Interaction technique for AR, it needs to be compact, cheaper and able to give inordinate user experience. In addition to these, it should be able to process all the data in real-time. Further researches needs to be focused more on mobile devices. Because, the market of mobile devices is gigantic and these are the most appropriate devices to experience AR in day-to-day life. So, researches can be focused on Camera based interactions like gesture recognition using basic mobile cameras.

REFERENCES

[1] Mehdi Mekni and Andre Lemieux "Augmented Reality: Application, Challenges and Future Trends" ISBN: 978-960-747-368-1, pp: 205-214.
[2] Feng Zhou,Henry Been-Lirn Duh and Mark Billinghurst "Trends in Augmented Reality Tracking, Interaction and Display: A Review of Ten Years of ISMAR"
[3] D. Overholt. "The MATRIX: A Novel Controller for Musical Expression" In Proceedings of the International Conference on New Interfaces for musical Expression, pages 38–41, 2001.
[4] Pierre Ducher : "Interaction with augmented reality" in Human computer interaction with augmented reality technical reposrt volume-2, Issue 4, ISSN:2198-9494, 2014.
[5] Stuart Goose, Sandra Sudarsky, Xiang Zhang, and Nassir Navab "Augmented Reality Supporting Mobile Industrial Maintenance" Published by the IEEE CS and IEEE Communications Society 2003 IEEE.
[6] P. Milgram and F. Kishino." A taxonomy of mixed reality visual displays." IEICE TRANSACTIONS on Information and Systems, 77(12):1321–1329,1
[7] Andrew Wilson and Aaron Bobick, "Learning visual behavior for gesture analysis," In Proceedings of the IEEE Symposium on Computer Vision, Coral Gables, Florida, pp. 19-21, November 1995.
[8] Y. Genc, S. Riedel, F. Souvannavong, C. Akinlar, and N. Navab. "Marker-less tracking for ar: A learning-based approach. In Mixed and Augmented Reality", ISMAR 2002 .
[9] T. Ohno, N. Mukawa, and A. Yoshikawa. "FreeGaze: a gaze tracking system for everyday gaze interaction" In Proceedings of the 2002 symposium on Eye tracking research & applications. ACM, 2002.
[10] M. Anabuki and H. Ishii. "AR-Jig: A Handheld Tangible User Interface for Modification of 3D Digital Form via 2D Physical Curve." In ISMAR'07, pp. 55-66, 2007.
[11] Pengfei Lu,Matt Huenerfauth"Accessible motion-capture glove calibration protocol for recording sign language data from deaf subjects"Proceedings of the 11th international ACM SIGACCESS conference on Computers and accessibility Pages 83-90,2009.
[12] J. David Smith,T.C. Nicholas Graham,David Holman,Jan Borchers "Low Cost Malleable Surfaces with Multi-Touch Pressure Sensitivity",Horizontal Interactive Human-Computer Systems, International Workshop 2007, ISBN: 0-7695-3013-3, pp: 205-208.