Proceedings of the 2013 International Conference on Machine Learning and Cybernetics, Tianjin, 14-17 July, 2013

Augmented Reality Technology Overview for Tourism App Development

Feng TIAN[1], Feifei XU[2], Jiacai FU[3]

[1]School of Design, Engineering & Computing, Bournemouth University, UK
[2]School of Tourism, Bournemouth University, UK
[3]School of Information Technology, Heilongjiang Institute of Technology, Harbin, China

Abstract:

With massive penetration of mobile phones the tourism industry is adopting augmented technology (AR) and apps supported by smartphones to create and provide a memorable and exciting experience to tourists with fun, interaction and learning. In this paper we first introduce the basic AR technology and then review several typical and well-known AR engines/platforms developed by companies and research institutes, focusing on their technical aspects. The challenges and recent development trends and directions are analysed to provide an insight for future research and development of AR technology for tourism and beyond.

1. Introduction

Augmented Reality (AR) is a technology which superimposes virtual 2D images, 3D models, videos, sounds, haptic feedback, etc. to the natural world [1], as illustrated in Fig. 1. It's a spectrum between the real world and the virtual reality (VR), which simulates and displays immersive, computer-generated environments.

Figure 1 A virtual bear on the palm

AR, as a technology, is not new but with a history stretching back decades of years. For years it's been an academic research area and later found applications in various areas. In 1992, a Boeing researcher Thomas Caudell has developed an AR-enabled maintenance-guide application which inserted relevant instructions into a real-time image viewed on a head-mounted display [2]. The application could help workers on a factory floor as they navigated the maze of electrical wiring in a giant airplane. AR has also been widely used for training and teaching purposes where annotations of components' name, location and operational steps, are placed in the real scene as guides to new users. Neumann and Majoros [3] suggested that AR can be used to support concurrent training and performance which greatly reduces the need for training as a distinct process. And in the medical filed AR can be applied to image guided surgery so that a surgeon can view and examine the CT or MRI data correctly registered on the patient [4]. AR's application is also seen in ultrasound imaging [5]. Comparative studies have shown that a guided process with AR-assisted instructions may yield better performance [6].

In addition, AR technology has been the subject of much interest and applied into a variety of applications such as advertisements, promotions, games, and information searches. With AR, a large number of videos have flooded the Internet with games, virtual shopping and search engines that add digital information on the top of live images or photos. The Swedish flat-pack furniture company Ikea is to augment the reality of shopping in its stores with its new catalogue, which will feature a smartphone icon on certain pages [7]. Holding a smartphone over the page will unlock a host of extra features provided by AR: with videos coming to life and photo galleries suggesting how a customer may decorate his houses and showing what is inside cabinets and wardrobes.

More recently, another area has seen AR's application, tourism industry. Searching under

978-1-4799-0260-6/13/$31.00 ©2013 IEEE

"Augmented Reality" in the Android Market or the iTunes App Store shows that at least half the results come under travel and tourism. This trend, helped by the massive popularity of smartphones and their constantly improving processors and sensors, along with the growth of high-speed wireless data network, will profoundly change the way we travel and interact with our surroundings. AR is an obvious travelling companion to help tourists find information about nearby ATMs and restaurants, the closest subway stop, and other points of interest in a new place. With AR you might aim a phone's camera at a restaurant, and on the screen he'll see not just the venue but also a review hovering above it. British digital agency Appshaker has worked on an augmented reality project for National Geographic Channel in a mall, allowing passersby to watch themselves "interacting" with various animals and dinosaurs on a big screen [8]. Similarly at Toronto's Royal Ontario Museum, visitors can use iPads or smartphones at a dinosaur exhibit to see how the beasts would have looked in real life [9]. Unlike virtual reality which is mainly used before or after the trip, AR can enhance on-trip experiences of tourists.

From technical point of view, the AR apps for tourism above-mentioned are based on two types of technologies. One is to use a smartphone's built-in GPS, accelerometer and compass to determine a traveler's position and line of site. Most of early apps fall into this category, including Metro Paris Subway [10], Etips Travel Guide [11], etc. The other is not dependent on either location detection or a network connection, but computer vision techniques to recognize and match objects (captured by camera) with those from a database/library stored in a server, such as Word Lens [12]. The object recognition technology itself can be further categorized into two types: marker based and markerless. In this paper we will review these technologies used by some well-known AR engines (for both smartphone users and app developers), analyze challenges and difficulties facing the current development, and conclude the paper by shedding a light to future research and development direction.

2. Overview of AR technology on smartphone

AR is to register and superimpose the computer-generated graphical content, including 2D/3D images, videos, sounds, etc., over the real world. Technically AR involves various technologies including image processing, computer vision, computer graphics and display, as illustrated below in Figure 2.

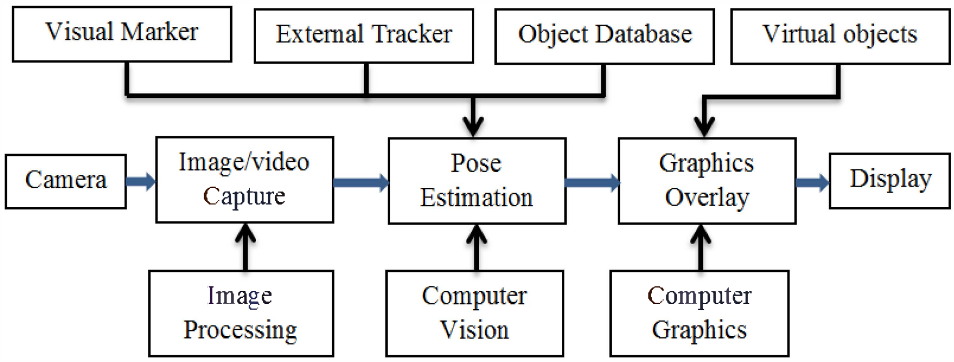

Figure 2 Outline of Augmented Reality Technology

It can be seen that the most sophisticated and yet important part is the pose estimation, which registers virtual objects with the real world (captured by smartphone camera) [13], involving image or object recognition and 3D object tracking. This is because the virtual objects or computer-generated imagery has to look realistic and be properly aligned with the real environment in order to create an authentic impression.

2.1 Marker-based and markerless tracking in AR

Traditionally the visual markers have been widely used to assist tracking and recognition [14][15]. The design of the markers can be barcode, symbol or a company/product logo. The use of these markers may limit the interactivity and is constrained to a range of photos or objects encapsulated within a border to create the marker. Therefore, in order to use this approach, these visual marks have to be printed previously and also be kept for future uses.

To overcome the limitation or constraints of the marker-based AR applications, much research has been done on markerless AR technology in which nature features or any part of the real environment may be used as a target that can be tracked in order to place virtual objects [16][17]. This mainly involves computer vision and image processing technology for pose estimation. The geometric primitives to be considered for the estimation include points, segments, lines, contours or a combination of these different features [18][19][20][21].

During recent years the mobile technologies have been advanced significantly, both in hardware and software. The smartphone CPU has reached one gigahertz, close to that of many little laptops (whose CPU speeds are around 1.5 gigahertz). In fact Samsung Galaxy S3 has a quad-core 1.4Ghz Exynos processor while iPhone 4 comes with a processor of dual-core 1.2Ghz [22]. Top smartphones also has come with graphics processing units (GPU), ideal for gaming and video streaming. The cameras in the smartphones are very advanced to feed raw data of the environment into computer vision algorithms. The camera in Samsung Galaxy S3 is 8 megapixel and supports 1080 video recording. The combination of these advances has enabled new markerless approaches like the ones based on natural features to break into the AR world [23]. Another advantage of the markerless systems is the possibility of extracting from the environment characteristics and information that may later be used by them.

2.2 Location-based tracking in AR

For the location-based AR, the position of objects on the screen of smartphone is computed using the user's position and the direction in which the user is facing. The AR apps will know which way you're facing (by the compass) and, the orientation at which you're holding your phone (by accelerometer) and where all the points of interest are relative to your location ((by GPS or Wi-fi). Compared to object-based (marker or markerless) approach, the location-based AR applications do not demand much computing from smartphone processor on image processing and computer vision algorithms. As a result most of AR apps for tourism are location based, such as [10] [11], though we see more object recognition based apps being developed and released.

3. Current AR platforms and engines for tourism game app development

Using a web browser like IE, FireFox or Chrome we are able to view rich contents online such as text, image, video. Based on WWW standard a web developer can use various programming languages (like HTML, XML, Javascript, PhP, etc.), web authoring tools (like Dreamweaver) or content management service (like Wordpress, Drupal, etc.) to develop a webpage which can be opened in browser. Similarly, to support and allow people to develop their own AR mobile applications, quite a few AR developing engines or platforms, so called browsers have been developed. In this section we will review and compare these platforms, providing an insight from technical point of view. Typical apps based on the platforms will also be covered.

3.1 Wikitude

Developed by the Austrian company Wikitude GmbH, Wikitude was the first publicly available application that used a location-based approach to AR [24].

Wikitude is a cross-platform AR SDK, which supports all mobile phone systems: Android, Blackberry, IoS and Windows phone. The SDK, coming with a library which unifies all AR technologies, is kept as lean and lightweight to fit nicely in apps. A user may create AR content using well-known web standards, getting most of HTML, JavaScript and CSS. With thousands of content providers, a user may find stunning AR contents around him/her. Though it's first developed using location-based tracking approach, the Wikitude ARchitect engine now provides support for recognizing and tracking markers as well as arbitrary natural images in real time. This way it's very easy to combine geo based and vision based algorithms and content to create one immersive AR experience.

Both native Londoners and the millions of who visit London each year are now presented with a unique opportunity to climb one of the world's greatest arenas, thanks to the "Up at The O2" attraction. Once at the summit, visitors can then take advantage of a stunning 360º view of London, and use the Wikitude-powered augmented reality element to further explore their surroundings.

3.2 Layar

Layar, an AR platform or browser, is developed by a Dutch-based company. It works by using a combination of the smartphone's camera, compass and GPS data to identify the user's location and field of view, retrieve data based on those geographical coordinates, and overlay that data over the camera view.

Data in the Layar browser comes in the form of *layers*. Layers are REST web services serving geo-located points of interest in the vicinity of the user. Layers are developed and maintained by user or third-party developers using a free API provided by the company. Various types of engaging AR experiences can be created, such as recognizing real world objects and displaying digital experiences on top of them with Layar Vision, or creating interactive features like 3D objects and animation. Also, the location based layers help users find nearby locations, including restaurants, shops and other businesses, as well as historical locations and monuments. Figure 3 below gives an outline of the Layar platform.

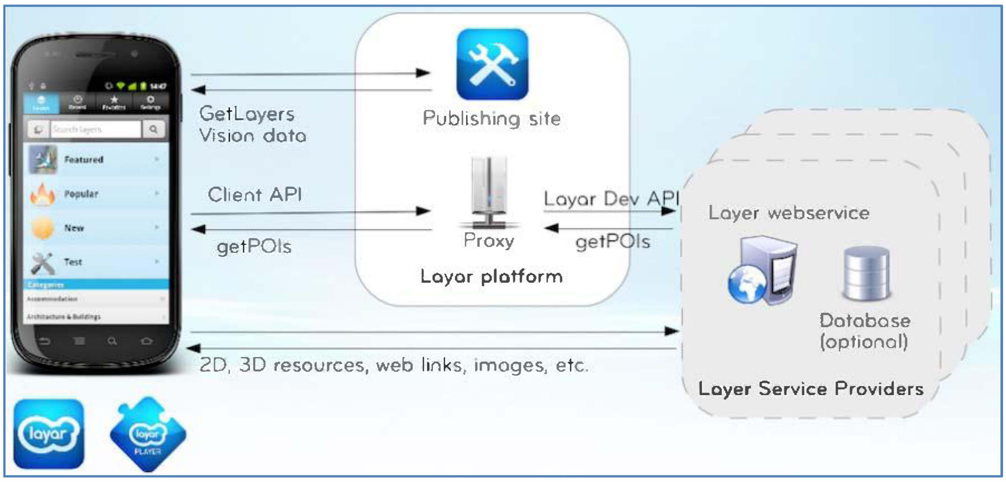

Figure 3 Outline of Layar platform (courtesy to Layar)

The Layar Server is the heart of the Layar service, which provides the interfaces to the Layar Reality Browser, the Layar Publishing site and the external Layar Service Providers which are to be created by third party developers. The key to make a layer work is to set up a web service. This web service needs to communicate with the layar server and generate appropriate getPOIs response based on Layar Developer API.

Lonely Planet is experimenting with just that in partnership with Layar [25]. They've enhanced all 42 Lonely Planet European city guidebooks (city guides, pocket city guides and Discover city guides) to work with the Layar app. Simply by scanning the cover through the app, travellers can enjoy live weather feeds, exclusive videos and photos, transport planners, activity ideas and hundreds of bookable hotel and tour listings – all for free.

3.3 String

String is a vision-based AR SDK for IOS, developed and owned by String Labs Ltd, a company in London.

According to the company [23], String recognizes framed images and understands where they are in 3D space, and then displays rich 3D graphics on top. The framed images are actually markers which are of PNG format. String is able to Track up to 10 markers simultaneously, very robust as long as the markers are of high contrast outline. To create a String based AR apps, the developers may use the Unity, a well-known 3D game engine, getting started without a single line of code.

3.4 KML/HTML Augmented Reality Mobile Architecture (KHARMA)

Three AR engines above mentioned are all proprietary technologies developed and owned by individual companies. Currently, there is no standard way to create or render AR applications so that AR apps may be able to talk to each other or share data. The KHARMA architecture, being proposed and developed by a Georgia Tech team, is a new open platform for AR that lets users create content using HTML and JavaScript, web development tools already in widespread use day [26]. KHARMA combines the keyhole markup language (KML) used by the Google earth mapping program with HTML. The team hopes that its open standard, an enhancement of existing Web protocols, will yield a common way for every Web browser to store, transmit, and manipulate data for augmented reality services. Argon is an AR browser which is being developed by the team. The basic framework may be able to incorporate image or object recognition algorithms.

4. Challenges and Solutions of Future AR Tourism Apps

Despite that more and more AR apps are being pushed to the market to provide tourists rich and exciting multimedia contents timely on site, there are challenges which both users and developers of AR apps are currently facing. In this section we will analyse two major issues and potential solutions to them.

The first major issue for tourism industry to fully adopt AR apps is that currently most of apps are location-based, relying purely on GPS positions and therefore what databases such as Google Maps say is supposed to be there, rather than what actually is. If the maps or databases are not updated to the latest incomplete or inaccurate, the apps won't work or at least are not able to overlay the graphics correctly. At this moment the accuracy of GPS in the smartphone is around 10 meters, depending on various factors such as number of GPS satellites available. This accuracy may not be enough for certain historical sites, such as Roman Forum, where many interesting objects or ruins are tight together. Even worse, the roof of car, tall buildings, mountains, and other obstructions can block line of sight to GPS satellites. The problem of losing GPS signal may be compensated by Wi-Fi supported by the smartphone. However nowadays it's still quite expensive or even unavailable for data roaming internationally.

To deal with this issue one solution is not to use GPS to figure out where to overlay a virtual object, but object recognition instead. Sony Computer Entertainment is developing a new AR system called Smart AR using the object recognition technology [27]. With Smart AR, certain real-world objects could become part of a game when viewed through a device such as the PlayStation Portable. This could allow game characters to appear on a tabletop, perhaps, or to respond to the movement of real objects. And Avideh Zakhor, a professor of electrical engineering at the University of California, Berkeley, focuses on creating truly believable augmented reality experiences. She, together with her team, uses image-recognition algorithms to chase the ultimate goal: pixel-perfect imposition of objects, surfaces, images, and textures onto the reality viewed through a smart phone or headset. Another solution obviously is to increase the accuracy of GPS, which may be realized in very near future with the fast pace of smartphone development.

The second major problem which the current tourism AR apps face is that using the smartphone for the apps runs down the battery very quick and poses potential for accidents when a traveller holds the phone to view graphic contents and in the meantime walking on a street in a new town.

The solution to this issue is wearable AR glasses. Many researchers and companies are working on developing the glasses for consumers. A company called Vuzix has developed the glasses which look like wraparound sunglasses, except you can't see directly through the lenses [28], as shown in Figure 4. Instead, small cameras centered on the outside of each lens feed continuous video through a smartphone to an LCD screen mounted inside each lens, creating a single stereoscopic field of view on the LCD, where computer graphics merge with the real world. Paul Travers, president of Vuzix, says that in the near future video glasses will deliver spectacular AR effects. "Instead of a little cell phone display, you'll have an image on the LCD that looks like an IMAX theater filling your field of view."

Figure 4 Vuzix Eyewear (Courtesy to Vuzix)

Babak Parviz, an associate professor of bio-nanotechnology at Seattle's University of Washington. Parviz has made a contact lens etched with a tiny, transparent electronic circuit that contains a single LED, allowing it to display text and images that would appear to hover in space at a readable distance in front of the eye.

5. Conclusion

Augmented Reality (AR) technology and system enriches the real world with the virtual objects and scenes, being images (2D/3D), videos, sounds, etc. After decades of research, the technology which used to be only available in the laboratory is now mature enough for mass adoption in the consumer market, including tourism industry. To enhance tourists' experience and provide them with more information either for learning or fun, many researchers and companies have developed AR engines, browsers and apps. In this paper we have looked into typical AR engines, focusing on their technologies which are either location-based, marker/markerless object recognition based or both. Some well-known tourism apps based on these engines are reviewed as well. With further advancement of smartphone and other hardware technology, tourists may interact seamlessly with virtual objects in the real world.

Reference

[1] T. Jebara, C. Eyster, J. Weaver, T. Starner, and A. Pentland. Stochasticks: Augmenting the billiards experience with probabilistic vision and wearable computers. Int'l Symp. On Wearable Computers, pp. 138–145, Cambridge, MA, USA, Oct.13-14 1997. IEEE CS Press. ISBN 0-8186-8192-6.

[2] D. Sims. New realities in aircraft design and manufacture. IEEE Computer Graphics and Applications, 14 (2): 91, 1994.

[3] Neumann, U., Majoros, A., "Cognitive, Performance and System Issues for Augmented Reality Applications in Manufacturing and Maintenance", Proceedings of IEEE Virtual Reality Annual International Symposium 1998, pp. 4-11

[4] W. Grimson, G. Ettinger, T. Kapur, M. Leventon, W. Wells, and R. Kikinis. Utilizing segmented MRI data in imageguided surgery. International Journal of Pattern Recognition and Artificial Intelligence, 11(8):1367-97, 1998.

[5] S. Andrei, D. Chen, C. Tector, A. Brandt, H. Chen, R. Ohbuchi, M. Bajura, and H. Fuchs. Case study: Observing a volume rendered fetus within a pregnant patient. Proceedings of IEEE Visualization, 17-21, 1993

[6] Tang, A Owen, C., Biocca, F., Mou, W., "Comparative Effectiveness of Augmented Reality on Object Assembly", Proceeding of the conference on Human Factor in Computing Systems, CHI 2003, April 5-10, 2003, Ft. Lauderdale, Florida, USA, pp. 73-80

[7] Ikea AR catalogue, http://crave.cnet.co.uk/software/ikeas-augmented-reality-catalogue-takes-you-inside-cabinets-50008677/

[8] National Geographic Lets You Pet Dinosaurs, http://www.adweek.com/adfreak/national-geographic-lets-you-pet-dinosaurs-mall-136591

[9] Royal Ontario Museum, http://www.rom.on.ca/dinos/exhibition/

[10] Metro Paris Subway, http://www.metroparisiphone.com/index_en.html

[11] Etips Travel Guide, http://www.etips.com/en/travel-en/augmented-reality.html

[12] Word Lens, http://questvisual.com/us/

[13] R. Azuma. Tracking requirements for augmented reality. Communications of the ACM, 36(7):50-51, 1993.

[14] Daniel Wagner, Dieter Schmalstieg. First Steps Towards Handheld Augmented Reality, Proceedings

of the 7th International Conference on Wearable Computers, IEEE Computer Society Press, 2003, ISBN: 0-7695-2034-0, 127 – 135.

[15] Hirokazu Kato, Mark Billinghurst. Marker Tracking and HMD Calibration for a Video-based Augmented Reality Conferencing System, Proceedings of the 2nd IEEE and ACM International Workshop on Augmented Reality '99, pp. 85-94, October 20-21, 1999 San Francisco.

[16] Andrew I. Comport, Éric Marchand, François Chaumette, A real-time tracker for markerless augmented reality, ACM/IEEE Int. Symp. on Mixed and Augmented Reality, ISMAR'03, pp. 36-45, Tokyo, Japan, October 2003.

[17] G. Simon and M.-O. Berger. Reconstructing while registering: A novel approach for markerless augmented reality. In IEEE Int. Symp. on Mixed and Augmented Reality (ISMAR'02), pages 285–294, Darmstadt, Germany, Sept 2002

[18] D. Dementhon and L. Davis. Model-based object pose in 25 lines of codes. Int. J. of Computer Vision, 15:123–141, 1995

[19] M. Dhome, M. Richetin, J.-T. Lapresté, and G. Rives. Determination of the attitude of 3-d objects from a single perspective view. IEEE Trans. on Pattern Analysis and Machine Intelligence, 11(12):1265–1278, December 1989.

[20] R. Kumar and A. Hanson. Robust methods for estimating pose and a sensitivity analysis. CVGIP: Image Understanding, 60(3):313–342, Novembre 1994.

[21] E. Marchand and F. Chaumette. Virtual visual servoing: a framework for real-time augmented reality. In EUROGRAPHICS'02 Conference Proceeding, volume 21(3) of Computer Graphics Forum, pages 289–298, Saarebrücken, Germany, September 2002.

[22] Smartphone CPU, http://smartphones.techcrunch.com/compare/245-355/Samsung-Galaxy-S3-vs-Apple-iPhone-5

[23] String, http://www.poweredbystring.com/

[24] Wikitude, http://en.wikipedia.org/wiki/Wikitude

[25] Lonely Planet meets Layar: http://www.lonelyplanet.com/blog/2012/06/25/lonely-planet-meets-layar/

[26] Georgia Tech AR project, https://research.cc.gatech.edu/kharma/content/kharma-framework

[27] Sony's Smart AR project, http://www.technologyreview.com/news/424147/sony-sets-its-sights-on-augmented-reality/

[28] Wearable AR glasses, http://www.vuzix.com/augmented-reality/index.htm