ATLANTIS PRESS

Atlantis Highlights in Engineering, volume 5

International Conference on Precision Machining, Non-Traditional Machining and Intelligent Manufacturing (PNTIM 2019)

Application of Augmented Reality in Campus Navigation

Li Zhigang*, Qi Guanglei, Hu Wenkai, Ma Xiangyu, Guo Qinsheng
Department of Computer Science and Technology
Century College
Beijing University of Posts and Telecommunications
Beijing, China
* E-mail: owent@sina.com

Abstract—**Augmented reality(AR) has the unique quality of providing a direct link between the physical reality and virtual information about that reality. In the application of a visitor visiting the campus with a smartphone, the location and pose information of the phone are acquired and calculated in real time. Then registration in 3D is realized by tracking the visitor's viewpoint and the phone post based on the result data. Also the location and orientation of the campus buildings relative to the visitor are identified. When the visitor is watching campus buildings, additional digital information appears to become part of the real world in the visitor's perception utilized by AR technology. As a result, the visitor not only appreciates the appearances and styles of the buildings directly, but also learn their history from the additional information which they can't intuitively understand.**

Keywords-Augmented Reality; Location Based Service; Pose; Navigation

I. INTRODUCTION

With the vigorous development of China's education, many new campus buildings have risen to the ground, and the campus has become more beautiful and full of academic atmosphere. During the holidays, parents take their children to visit the campus to become a tourist attraction. The annual College entrance examination counseling meeting also brings a lot of parents and middle school students to the campus. Additionally, with the frequent inter-school academic exchanges and mutual visits between teachers and students, the requirements for campus navigation assistant (or software) have gradually increased. In the traditional way, visitors often use navigation software or are accompanied by acquaintances directly to the campus. Imagining a kind of software with which you could see more than others see, hear more than others hear, and perhaps even touch, smell and taste things that others cannot, it is wonderful for visitors. The software implemented with Augmented Reality (AR) technology is not only as assistant for campus navigation, but also a commentator which can tell the history and cultural stories behind the campus.

AR is the technology to create a next generation, reality-based interface[1][2][3][4] and is moving from laboratories around the world into various industries and consumer markets. AR supplements the real world with virtual (computer-generated) objects that appear to coexist in the same space as the real world. It was recognized as an emerging technology, and with the rapid development of mobile internet technology, today's smartphones are powerful and small enough to support the computing and graphical overlay, we are able to embrace this very new and exciting kind of human-computer interaction in the palm-sized smartphones[5][6][7].

The traditional navigational software often uses the geographical location information to present only the campus building as a floor plane or a simple three-dimensional map in the phone. However, in the same location, due to the visitor surrounded by buildings in different orientations, he often needs to identify the orientation to confirm which building is in front of him. Based on the geographical location information, the paper also crucially exploits the orientation and post information of smartphones, automatically identifying the scenes (location and orientation) that the visitor is watching, this is called "Registered in 3D", and then integrates real scenes and virtual objects in real time to achieve a stronger sense of immersion.

II. AR TECHNOLOGY SUMMARY

A. AR basic theory and process

What is AR is an important technical problem. Some people associate the visual combination of virtual and real elements with the special effects in movies such as Avatar. The most widely accepted definition of AR was proposed by Azuma in his 1997 survey paper. According to Azuma [1], AR must have the following three characteristics i.e. combines real and virtual, Interactive in real time, and registered in 3D. Two decades have passed, and AR has made great progress. The focus and difficulty of the system has also changed, but these three elements are basically indispensable in the AR system.

It can be seen that a complete AR system requires at least three components: a tracking component, a registration component, and a visualization component. A fourth component—a spatial model (i.e., a database)—stores information about the real world and about the virtual world[2]. The real-world model is required to serve as a reference for the tracking component, which must determine the visitor's location in the real world. The virtual-world model consists of the content used for the augmentation. Both parts of the spatial model must be registered in the same coordinate system.

As showed in Fig. 1, AR uses a feedback loop between human visitor and computer system. The visitor observes the AR display and controls the viewpoint. The system tracks the visitor's viewpoint, registers the pose in the real world with the virtual content, and presents situated visualizations.

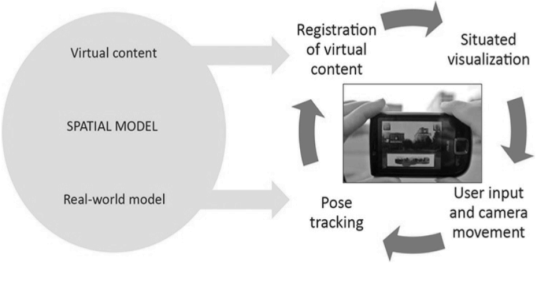

Figure 1. The Conceptual flow of a typical AR system

B. Spatial registration

For the visitor utilizing an application in a smartphone for navigation in the campus, since the information of location and pose of the smartphone capturing the real scene may be constantly changing dynamically, it is necessary to analyze the information in real space in real time. Three technical approaches are generally used in the processing of spatial location tracking and registration, which are sensor-based tracking technology, vision-based tracking technology and hybrid tracking technology[8][9][10][11]. The core of the approaches is the scene recognition and semantic understanding. The physical world we are in is not only composed of various three-dimensional structures, but also composed of transparent windows, brick walls, tables and so on. For AR, we can only overlay the virtual curtain onto the plane with geometric information, however, we can more accurately overlay the virtual curtain onto the windows with semantic understanding. In the campus tour, the visitor may take pictures of different campus buildings in the same location. In this case, the geographical location is necessary but not sufficient. For example, there is a south-facing teaching building, when the visitor takes a photo with his smartphone in front of it, the orientation of his smartphone's camera is "north direction", whereas, the visitor turns around and takes photos of the experimental building opposite the teaching building, the orientation of his smatphone's camera becomes "south direction". As shown in Fig. 2, only by using the difference of the orientation information of the smartphone, it can be truly understand whether the scene in front of the visitor is the teaching building or the experimental building. By this way, AR can achieve a combination of virtual and real more accurately.

Figure 2. Scene understanding requirements in the campus

In the campus tour, for large 3D objects such as campus buildings, the use of image recognition technology to achieve spatial registration is clearly problematic, and the recognition rate is very low. While, multi-modal data fusion combing non-visual data of multi-sensor measurements in the phone is a feasible way to achieve semantic understanding. In the implementation of the campus tour, the geographical location information and the smartphone posture information representing the orientation of the phone and the visitor will be first acquired and calculated in real time, and then the multi-modal fusion which combined the two kind of information will be used to understand the scene and realize the spatial registration and tracking.

C. Virtual and real approaches

In case of correctly semantic understanding of the scene and clarifying the real-world objects in the visitor's smartphone lens, the additional virtual information can be accurately superimposed on the real world. Generally, there are two kind of virtual superimposed information, video and model. They may appear or disappear as the scene changes. Moreover, the superimposed videos and 3D models should be selected and made elaborately so as to attract the visitor and give him vivid and valuable enjoyment.

D. Real-time interaction approaches

Gesture-based interaction is a commonly used in smartphones such as in the games of Racing Cars. Also multiple interaction modes may be superimposed on a phone by current popular ways such as touch screen and voice etc. Suitable interaction can bring more changes and fun to AR experience.

III. INERTIAL MEASUREMENT UNIT

A. Principles of inertial measurement unit

In most high-performance smartphones, the pose and orientation of the phone are generally determined by an inertial measurement unit (IMU), which is a combination of a three-axis gyroscope and a three-axis accelerometer, and is often fabricated using a micro electro mechanical (MEMS) principle. Among them, the MEMS gyroscope usually has a movable capacitive plate in two directions, the radial capacitive plate plus the oscillating voltage forcing the object to move rapidly, and the lateral capacitive plate to measure the capacitance change due to the lateral Coriolis motion. Because the Coriolis force is proportional to the angular velocity, so the angular velocity can be calculated from the change in capacitance, while the accelerometer measures the acceleration in a certain direction. The corresponding application programming interface (API) can be invoked by the smartphone's development kit (SDK) to measure in three orientations (North, North, East and West). As shown in Fig. 3, the phone coordinate system is defined, the X axis is horizontal and points to the right, the Y axis is vertical and points to the front, and the Z axis points directly above the front of the screen. When the phone swings left and right (rotating around the y-axis), the changing roll angle Φ(roll) is obtained, the range is (-90 to 90), and when the phone swings back and forth (rotating around the x-axis), the varying pitch angle θ is obtained (Pitch), ranging from (-180 to 180), when the phone's landscape is converted to portrait or vertical to horizontal (rotating around the z-axis), a varying yaw angle is obtained.

The visitor usually uses the camera in the smartphone to take pictures of the real scene. Since the position of the camera on the phone is fixed, the pose and orientation of the camera can be determined by the pose of the phone.

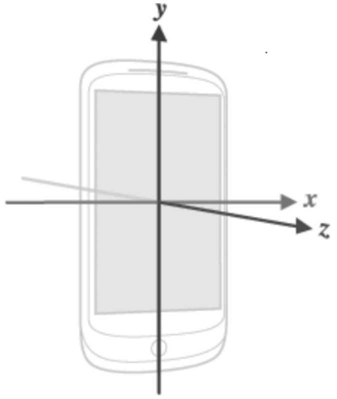

Figure 3. Coordinate system in mobile phone

B. Location and pose calibration

In the example of my campus navigation, there are three buildings, i.e. Mingde Building, Library building and Xingzhi Building, with a square in the centre. In order to realize the campus navigation, it must first calibrate the three buildings. The geographical locations are shown in Fig. 4. The direction of the "upward" on the map is northward. The central dot on the map represents the current position of the visitor. In the current position, if the visitor (that is the phone) faces south, west and east, his viewpoint will correspond to the three buildings respectively. Before scheme implementation, it is necessary to calibrate and record the values of location coordinates (eg. what is the east longitude and the north latitude), and the IMU data (x, y, z) of the phone facing viewing the three buildings.

Figure 4. Map of the example campus

C. Error estimate of smartphone pose

Due to the use of low-cost IMU sensors in general smartphones, there is a large error in IMU measurement data[12]. In actual applications, since the post of the mobile phone cannot be kept constant when taking pictures at the same position, there is usually a certain error between the actual value of the IMU and the stored value of the calibration. When the smartphone is not facing directly the designated building, but tilted at an angle, what is the AR appearance due to the error?

The optical schematic of the smartphone's camera is simplified as shown in Fig. 5. Where the line OF represents the optical axis, the line segment of AB is imaging sensors (usually Complementary Metal-Oxide-Semiconductor or CMOS used on smartphones), the double arc of CD is the lens, the COD is the plane of lens, the point F is the focus, and the length of the line segment of OF represents the focal length.

Suppose that the angle α between the ray CW and DZ is the angle of view i.e. angle between OX and OY, which is parallel to CD and DZ respectively.

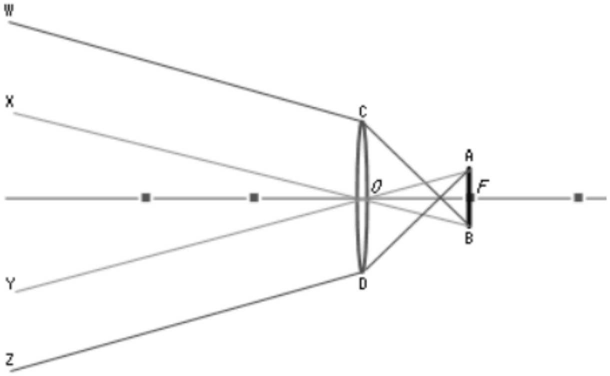

Figure 5. The simplified optical schematic of camera

In the following experiment, three Huawei brand Android smartphones were used as samples, of which cameras support Exchangeable Image File format (Exif). Some parameters of the camera obtained from the Exif list in Table 1, where the column of Name means the model of the phones, the unit of Focus Length and Equivalent focus Length relative to 35mm lens are mm, and unit of the angle of view α is° .

TABLE I. SAMPLE PARAMETERS

Name	**Focus Length**	**Equivalent focus Length**	**Angle of view**
Honor6	3.8	27	77.4
ChangX8	3.4	25	81.7
Honor10	4.0	27	77.4

Now we calculate the approximate imaging field (in width and height) from a distance based on angle of view. Assume the distance from the lens to the scene (for example, a building) is d , imaging width is w, then we have $w = 2d\tan\frac{\alpha}{2}$.

A group values calculated in Table 2. If the smartphone deflect an angle $\Delta 2\alpha$ around the direction perpendicular to the plane COD of Figure 5, then the field of view moves out of the above scene by a distance ΔW in width, which can be calculated by $\Delta w = 2d(\tan\frac{\alpha+\Delta\alpha}{2} - \tan\frac{\alpha}{2})$

TABLE II. FIELD OF VIEW AND DEVIATIONS IN WIDTH(UNIT IS M).

$\Delta 2\alpha$	***0***	***5***	***10***	***15***	***20***
w=20	ΔW=3.8	3.10	6.74	11.2	16.88
w=30	ΔW=3.4	5.60	10.11	16.8	25.3
w=40	ΔW=4.0	6.20	13.5	22.4	33.8

For a building, the width is generally more than 40m, If these angle $\Delta 2\alpha$ sensing from IMU (drift error usually small than 10° per hour[12]) in front of the building small than 40meters, then it is still assured that the building is within the scope of the camera, and deviations of angle have no effect on the AR presentation.

IV. IMPLEMENTATION AND EXPERIMENTS

The hardware device for development is the three Android phone above with their camera as viewfinder sensors, while, the software for development includes Unity3d, vuforia SDK, C# and Android SDK. Augmented reality development is implemented according to the MVC model architecture. Among them, C# is used for the scripts of data acquisition and process control. The vuforia SDK provides support for augmented reality development. Unity3 provides APIs of smartphone sensors, and rendering tools for computational graphical objects. After the augmented reality software developed has been debugged , then it will be packaged into an app that can be run directly on an Android phone.

When augmented reality application software (an app) runs on a smartphone, it constantly recognizes the real scene through its camera, while continuously capturing the location and pose information of the phone using its GPS and IMU sensors. In case that the data of location and pose meet the given conditions, the spatial registration is completed, and then the virtual reality is superimposed on the real sense. For example, in the case that the visitor is already near the central dot of Fig. 4, he runs the app on the phone and points the phone to Mingde Building. The UI on the phone will display the real scene of the building. Meanwhile, a video window will be superimposed on it, which will introduce the building age, the European style features and administrative office in a multimedia way, and then present the internal structure model which can be touched for interaction. Later, when the visitor turns himself facing to the west side, the UI on the phone will be scene of the library, then a virtual model of its internal floor is superimposed and presented, and the library's book volume, opening hours and other information are played by voice.

V. CONCLUSION

Visiting the college campus has become a tourist hotspot and the demand for campus navigational assistance is growing. Obtaining the location and pose of the smartphone by its GPS and IMU data, thus the viewpoint of the phone or the visitor is accurately grasped. Then, the multi-sensor fusion approaches are used for the spatial registration, and a combination of the real and virtual world appears in the eyes of the visitor, which helps the visitor not only understand the school style, but also deeply understands the history, tradition and culture of the school. The implementation of the application also provides a good idea and technical route for the design of Smart Campus Projects. It can be combined with the Smart Campus Projects or the Push Technology of the campus information system to realize hot news push and campus culture service further.

REFERENCES

[1] AZUMA R T. A survey of augmented reality [J]. Presence: Teleoperators and Virtual Environments, 1997, 6(4): 355-385.

[2] Dieter Schmalstieg,Tobaia Höllerer. Augmented Reality_ Principles [M]. Addison-Wesley Boston • Columbus • Indianapolis • New York • San Francisco • Amsterdam• Cape Town，2014

[3] BILLINGHURST M, KATO H. Collaborative augmented reality [J]. Communications of the ACM, 2002, 45(7): 64-70.9

[4] R. J. K. Jacob. What is the next generation of human-computer in-teraction? In R. E. Grinter, T. Rodden, P. M. Aoki, E. Cutrell, R. Jeffries, and G. M. Olson, editors, CHI'06: Proc. Int'l Conf. on Human Factors in Computing Systems, pp. 1707–1710, Montréal, Québec, Canada, April 2006. ACM Press. ISBN 1-59593-298-4.

[5] Daniel Wagner，Gerhard Reitmayr，Alessandro Mulloni, et a1. Real—Time Detection and Tracking for Augmented Real@on Mobile Phones[J]．IEEE Transactions on Visualization and Computer Graphics, 2010，16(3):56—80.

[6] E. Jonietz. TR10: Augmented reality. http://www.techreview.com/special/emerging/, Mar. 12, 2007.

[7] WANG P, OUYANG S, ZHANG X, ZU R, CAI S. A Study on Campus Cultural Learning and Navigation. System Using Mobile Augmented Reality and Location Based Services[C]. Proceedings of World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education. 2014: 2011-2016.

[8] M. Kumar and T. Winograd. GUIDe: Gaze-enhanced UI design. In M. B. Rosson and D. J. Gilmore, editors, CHI'07: Proc. Int'l Conf. on Human Factors in Computing Systems, pp. 1977–1982, San Jose, CA, USA, April/May 2007. ISBN 978-1-59593-593-9.

[9] [9]E. Foxlin and M. Harrington. Weartrack: A self-referenced head and hand tracker for wearable computers and portable VR. In [4], pp. 155–162.

[10] T. H. Höllerer and S. K. Feiner. Mobile Augmented Reality. In H. Karimi and A. Hammad, editors, Telegeoinformatics: Loca-tion-Based Computing and Services. CRC Press, Mar. 2004. ISBN 0-4153-6976-2.

[11] D.W.F. van Krevelen and R. Poelman: A Survey of Augmented Reality Technologies, Applications and Limitations, The International Journal of Virtual Reality, 2010, 9(2):1-20

[12] MEMS motion sensor: 3-axis digital output gyroscope ST Microelectronics.