Seam the Real with the Virtual:a Review of Augmented Reality

Rui Zheng[1, a *], Dan Zhang[1,b] and GuoSheng Yang [1,c]

[1] Faculty of Information Engineering, MinZu University of China, Beijing, 100081, China

[a]rzhengbj@163.com, [b]zhangdan_being@163.com, [c] gsyangmuc@163.com

Keywords: Aaugmented reality; 3D interaction; AR application; AR display

Abstract. In this paper, we reviewed several outstanding contributions over the past decade in the field of augmented reality. The papers are selected according to their citation rate in the academic resources of Google scholar. In addition, considering that augmented reality covers several topics, and it has intersection with pervasive computing and wearable computing in concept, only the promising technologies and some prominent cases are concerned. In our opinion, a typical AR application is a solution for special application scenarios, and in most cases a comprehensive system is needed. For simplicity, we divided the technologies adopted in AR application into three categories: acquiring the data of environment, registering the virtual object in real environment and presenting the composite data. At the end of the paper, we discussed the remained problems and future development trend of augmented reality.

Introduction

Augmented reality (AR) is always a very active field of research in past decades. Unlike virtual reality (VR), where the user is completely immersed in a virtual environment, AR allows the user to interact with the virtual images using real objects in a seamless way. Therefore, it has wider application prospects and more potential economic value. Since Steven Feiner et al. proposed their prototype of tour machine which is a mobile augmented reality systems in 1997 [1], there have emerged lots of AR applications. Ronald Azuma et al. summarized some promising application in 2001[2], except for the traditional areas of assembly, inspection and medical applications, the collaborative and commercial applications captured the attention of the public. AR addresses two major issues with collaboration: seamless integration with existing tools and practices and enhancing practice by supporting remote and collocated activities that would otherwise be impossible. In the Disney theme park, AR is embedded in a large number of applications, most of them are operated in collaborative manner [3]. The rise of digital museum produce opportunities of applying AR technologies, T. Miyashita et al. introduce their Louvre - DNP Museum Lab (LDML) project comprising six presentations, with the goal of gaining experience in innovative multimedia approaches to bring together visitors and artworks in [4]. For commercial applications, AR has been used for real-time augmentation of broadcast video, primarily to enhance sporting events and to insert or replace advertisements in a scene, an example is the FoxTrax system which highlights the location of a hard-to-see hockey puck as it moves rapidly across the ice. In recent years, augmented reality application is more popular, A.Y.C. Nee et al. summarized the applications in design and manufacturing [5] and Kangdon Lee evaluated the application prospects of AR in Education and Training [6].

Although the AR application has many kinds of types, mobile AR systems should receive more attention as it may have millions of potential users. The prototype that Steven Feiner et al. proposed is composed by a backpack, a head worn display, a handheld display and its stylus, which is cumbersome and uncomfortable for long time use. However, the capability that it demonstrated is fascinating. Thanks to technological advances of mobile phone, the function of tour machine can be accomplished in most major smartphone today, which integrated with miniaturized GPS, magnetometer, inclinometer and camera. Thereby a well-designed application in app store can achieve the AR function. Layar is an example, which is an open platform for developer to build their own AR and interactive print campaigns on mobile device [7].

The progress is remarkable. However, there are some difficulties remained thereby no killer apps about AR appear, Fig.1 shows compilation diagram of AR technologies. In which the technologies is classified according to the step of processing: acquiring data, registering virtual object and presenting composite data. The detail is expounded in the following sections.
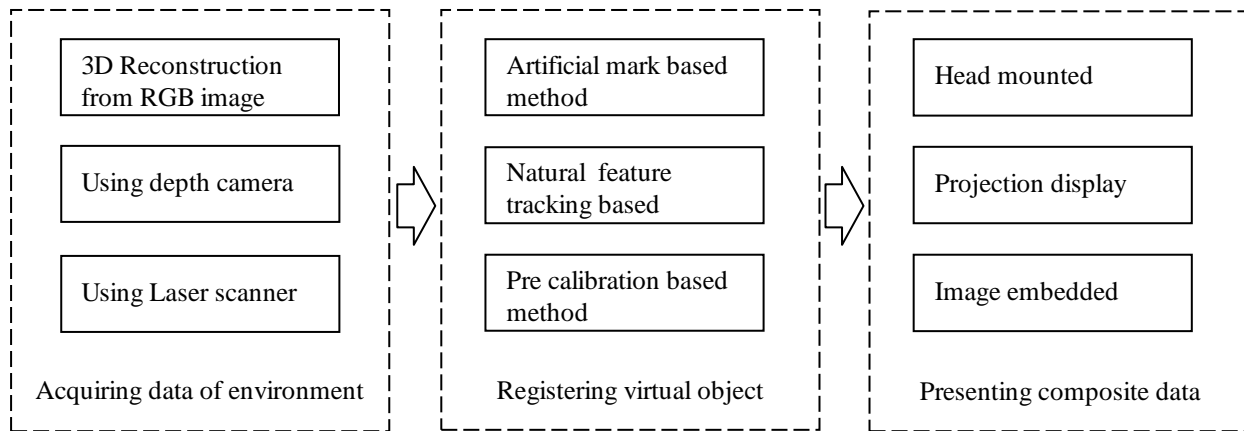
Acquiring data of environment	Registering virtual object	Presenting composite data
3D Reconstruction from RGB image	Artificial mark based method	Head mounted
Using depth camera	Natural feature tracking based	Projection display
Using Laser scanner	Pre calibration based method	Image embedded

Fig.1 Compilation Diagram of AR technologies

Acquiring data of environment

Accredited data of the real world is needed in AR application. There are some methods to achieve the goal. The traditional choice is using laser scanner, which is time consuming and the equipment that used for acquiring data is normally cumbersome, therefore it is not suitable for real time application. If the application uses an artificial mark to register the virtual object, only the image of the real background is needed, the 3D data of the environment may not be used. However, to get a dramatic effect, we expect the virtual object can interactive with the environment. For example, a virtual ball can bounce on the real surface until it stops. In this case, we must acquire 3D data of acceptable accuracy.

In computer vision, research on structure from motion (SFM) and multi-view stereo (MVS) has produced many compelling results, in particular accurate camera tracking and sparse reconstructions and increasingly reconstruction of dense surfaces. For robot vision, the similar problem is simultaneous localization and mapping (SLAM) which also needs to estimate the position of camera and the structure of environment. Such research focus more on real-time markerless tracking and live scene reconstruction based on the input of a single commodity sensor—a monocular RGB camera. The 'monocular SLAM' systems (MonoSLAM) [8] and the more accurate Parallel Tracking and Mapping (PTAM) system [9] allow researchers to investigate flexible infrastructure- and marker-free AR applications. Richard A. Newcombe et al. have been able to make live dense reconstruction with a single moving camera [10], although this method may be affected by light condition.

But while algorithms for estimating camera pose and extracting geometry from images have been evolving at pace, so have the camera technologies themselves. New depth cameras based either on time-of-flight (ToF) or structured light sensing offer dense measurements of depth in an integrated device. With the arrival of Microsoft's Kinect, such sensing has suddenly reached wide consumer-level accessibility. The opportunities for SLAM and AR with such sensors are obvious, but algorithms to date have not fully leveraged the fidelity and speed of sensing that such devices offer. As the data got by Kinect is sparse, an idea at hand is fusing multiple devices to form dense volumetric reconstruction. By using highly parallel general purpose GPU (GPGPU) techniques, Richard A. Newcombe et al. developed an algorithm [11], so that users can simply pick up and move a Kinect device to generate a continuously updating, smooth, fully fused 3D surface reconstruction. Using only depth data, the system continuously tracks the 6 degrees-of-freedom (6DOF) pose of the sensor using all of the live data available from the Kinect sensor rather than an abstracted feature subset, and integrates depth measurements into a global dense volumetric model.

Registering virtual object

After having acquired the data, the virtual objects need to be registered in the sight. As we mentioned before, for the method using artificial marker, the registering work is simpler as Yuji Ayatsuka et al. [12] does. The maker developed called Active CyberCode. Only by traditional computer user interfaces like keyboards or stick-pointers, once a character or information has appeared on or around a 2D code in a camera image, users are allowed to control them. The marker they proposed has a variable part for direct manipulations by fingers or other objects. A user can control a character or information invoked by a 2D code by "pushing" printed buttons. The variable part is recognized as an extended part of a code so that it costs almost the same as a normal CyberCode and is robust. Hirokazu Kato et al. [13] also registered the video pictures ID cards which has special patterns to build a AR conferencing system.

For those markerless methods based natural features, whether the problem can be solved mainly depends on the tracking effect of the features. In early stage, researcher preferred to use simple feature such as lines or points [14]. The result is apparently not reliable. Since the advent of SIFT, the situation has been greatly improved. However the extraction and matching of SIFT is time consuming, so some researcher draw support from special hardware such as GPU which has been mentioned before. In the meantime, some researcher tends to develop variation of the classic SIFT[15].

Presenting composite data

Adopting see-through HMD is the ordinary way of presenting composite data. In recent years, the focus shifted to projection-based display, as it is a good option for applications that do not require several users to wear anything, providing minimal intrusiveness. There are a variety of projection-based display techniques proposed for displaying graphical information directly on real objects or even daily surfaces in everyday life. Hrvoje Benko et al. presented an AR system called MirageTable. It is a curved projection-based augmented reality system, which digitizes any object on the surface, presenting correct perspective views accounting for real objects and supporting freehand physics-based interactions. The traditional HMD technology also has improved. Takemura, K. et al.[17] reported their work on catching the person's gaze which can effectively improve the user experience.

Conclusion

In the paper, we summarized recent trends of AR technologies and some of them may cause revolutions in the field. Besides some newly developed display technologies for cooperation based on projectors, the mobile application got the most attention as it may have millions of potential users. Following Moore's law, the computing power of mobile device is more and more strong in the next few years, and we believe the large-scale deployment of AR application will become reality.

Acknowledgement

References

[1] Steven Feiner; Blair MacIntyre; Tobias Hollerer; Anthony Webster, A touring machine: Prototyping 3D Mobile Augmented Reality Systems for Exploring the Urban Environment, Proceedings of First IEEE International Symposium on Wearable Computers (ISWC '97), 1997, pp 74–81. Cambridge, MA.

[2] Ronald Azuma; Baillot Yohan; Behringer Reinhold; Feiner Steven; Julier Simon; MacIntyre Blair, "Recent advances in augmented reality," IEEE Computer Graphics and Applications, vol.21, no.6, pp.34-47, Nov/Dec 2001.

[3] Information on http://www.souvr.com/event/201312/64016.html.

[4] T. Miyashita; P. Meier; T. Tachikawa; S. Orlic; T. Eble; V. Scholz; A. Gapel; O. Gerl; S. Arnaudov; S. Lieberknecht, An Augmented Reality Museum Guide, IEEE International Symposium on Mixed and Augmented Reality, 2008, pp.103-106, Cambridge, UK.

[5] A.Y.C. Nee ; S.K. Ong; G. Chryssolouris; D. Mourtzis, Augmented reality applications in design and manufacturing, CIRP Annals - Manufacturing Technology 61,2012, pp.657–679, Paris, France.

[6] Kangdon Lee, Augmented Reality in Education and Training, CIRP Annals-Manufacturing Technology, Tech Trends, March 2012, Vol. 56(2), pp 13-21, Feb 2012.

[7] Information on https://www.layar.com/.

[8] Davison, A.J.; Reid, I.D.; Molton, N.D.; Stasse, O., "MonoSLAM: Real-Time Single Camera SLAM," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.29, no.6, pp.1052-1067, June 2007.

[9] Klein, G.; Murray, D., "Parallel Tracking and Mapping for Small AR Workspaces," 6th IEEE and ACM International Symposium on Mixed and Augmented Reality, 2007, pp.225-234, Nov. 2007.

[10] Newcombe, Richard A.; Davison, A.J., "Live dense reconstruction with a single moving camera,", 2010 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp.1498,1505, 13-18 June 2010.

[11] Newcombe, Richard A.; Izadi, Shahram; Hilliges, Otmar; Molyneaux, David; Kim, David; Davison, Andrew J.; Kohi, Pushmeet; Shotton, Jamie; Hodges, Steve; Fitzgibbon, Andrew, "KinectFusion: Real-time dense surface mapping and tracking," 10th IEEE International Symposium on Mixed and Augmented Reality (ISMAR), 2011,pp.127,136, 26-29 Oct. 2011.

[12] Ayatsuka, Y., Rekimoto, J.: Active cybercode: a directly controllable 2d code. In: ACM Special Interest Group on Computer-Human Interaction (2006).

[13] Hirokazu Kato and 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, Oct.1999.

[14] Ansar, A.; Daniilidis, K., "Linear pose estimation from points or lines", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.25, no.5, pp.578,589, May 2003.

[15] Wagner, Daniel; Reitmayr, Gerhard; Mulloni, Alessandro; Drummond, Tom; Schmalstieg, D., "Real-Time Detection and Tracking for Augmented Reality on Mobile Phones," IEEE Transactions on Visualization and Computer Graphics, vol.16, no.3, pp.355,368, May-June 2010.

[16] Hrvoje Benko; Ricardo Jota; Andrew D. Wilson, MirageTable: freehand interaction on a projected augmented reality tabletop, Proceedings of ACM CHI 2012 Conference on Human Factors in Computing Systems, pp.199-208. May 2012.

[17] Takemura, K.; Takahashi, K.; Takamatsu, J.; Ogasawara, T., "Estimating 3-D Point-of-Regard in a Real Environment Using a Head-Mounted Eye-Tracking System,", IEEE Transactions on Human-Machine Systems, vol.44, no.4, pp.531-536, Aug. 2014.