Optik 123 (2012) 769–774

Contents lists available at ScienceDirect

Optik

journal homepage: www.elsevier.de/ijleo

Display systems and registration methods for augmented reality applications☆

Dongdong Weng, Dewen Cheng, Yongtian Wang*, Yue Liu

School of Optics and Electronics, Beijing Institute of Technology, Beijing 100081, China

ARTICLE INFO

Article history:
Received 23 December 2010
Accepted 26 May 2011

Keywords:
Augmented reality
Display systems
Interactive projection
Head-mounted display
Volumetric 3D display

ABSTRACT

Introductions are given to various display systems and corresponding registration methods developed and realized at Beijing Institute of Technology for augmented reality applications. An interactive projection system is used in a theatrical performance, and the registration between the virtual scene (projected image) and the real objects (dancers) is realized through a video camera working in the near infrared waveband. An ultra-light wide-angle head-mounted display system is designed for an interactive exhibition item in a scientific museum, and a multi-user six-degree-of-freedom tracking system is developed for the indoor registration, which uses infrared markers projected on the ceiling and a camera fixed on top of each head-mounted display. A fixed-position viewing system is developed for the on-site digital reconstruction of a historical site, which provides visitors with the visualization of the original magnificent buildings on the natural field of the ruins. The azimuth and pitch of the viewing system are detected by two rotary encoders. A volumetric 3D display based on multi-angle projection is applied in a navigation system for endoscopic sinus surgery. A 3D digital scanner is employed to scan the skin on the patient's head for the alignment between the patient and the 3D model from his CT or MRI images. The registration of the surgical tool is also achieved by optical means through tracking cameras.

1. Introduction

Augmented reality (AR) technology is developed from virtual reality (VR), and it is sometimes referred to as mixed reality (MR) in existing literatures. Whereas a VR system tries to immerse its user inside a computer-generated environment, an AR system integrates virtual objects into the real world in real time and tries to let its user believe that the virtual and real objects coexist in the same space [1]. Display and registration are the two key issues for AR. A suitable display device presents the merged environment vividly and effectively to the user, and a good registration method solves the problem of real-time alignment among the objects in the real and virtual worlds.

Although there have been successful applications of AR technology in the fields of medicine [2], industry [3], culture [4], education [5], entertainment [6] and defense [7], the number of cases is not as many as expected. We believe that AR technology can be more widely used if its display device and registration method are carefully chosen and specifically tailored for the requirement of each application. At Beijing Institute of Technology, we have developed various AR applications in areas such as artistic performance, exhibition showcase, historical site reconstruction and surgery guidance, using display devices ranging from projectors to head-mounted displays to volumetric displays, and registration methods based on digital cameras or rotary encoders or scanners. This paper describes a few examples.

2. Interactive projection system

Used in theatrical performances or exhibitions, a simple AR system composed of a digital projector and a video camera can produce spectacular effects. Fig. 1 shows a schematic drawing of the interactive projection system. With the projector fixed on the ceiling and its light beam bent by a reflective mirror, the image is projected on the ground. The registration between the virtual scene (projected image) and the real objects (people moving on the image) is realized through a downward-looking video camera. The video stream collected by the camera is sent to the computer. A background frame is obtained by accumulating previous frames, which is subtracted from the current frame. The motion of the people can thus be detected in real time, which is used to trigger the interaction. In order to reduce the impact of stray light on tracking accuracy, infrared light sources are placed near the camera and a near-infrared band-pass filter is attached to the camera lens.

Fig. 2 shows an actual application of such a setup in an artistic performance. The projected image is a pool of water. According to

☆ This paper is based on an invited presentation at the Conference on Current Developments in Lens Design and Optical Engineering X, Optics + Photonics, 2009.

* Corresponding author. Tel.: +86 10 68911272.
E-mail address: wyt@bit.edu.cn (Y. Wang).

0030-4026/$ – see front matter © 2011 Elsevier GmbH. All rights reserved.
doi:10.1016/j.ijleo.2011.05.033

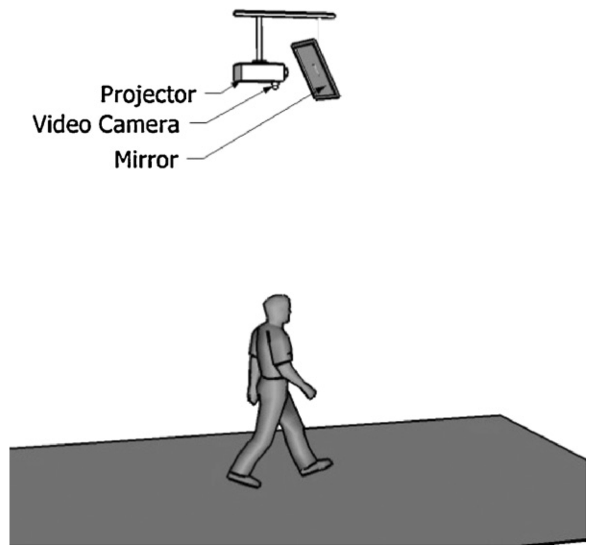

Fig. 1. Schematic drawing of interactive projection system.

Fig. 3. Eyepiece design layout and HMD prototype assembly.

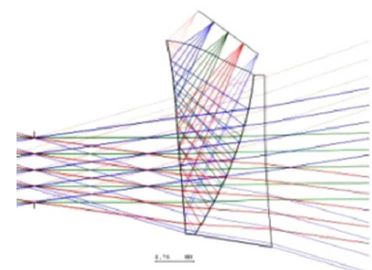

Fig. 4. Setup for optical see-through mode.

the real-time tracking result of the locations of the dancers, virtual ripples are projected onto the stage to create a beautiful scenery of dancing on water.

3. Head-mounted display system

Just like VR systems, many AR systems adopt head-mounted displays (HMDs) as the display devices. The difference here is that the HMDs in AR systems have the see-through feature which is needed for the combination of real scene and virtual objects. This see-through characteristic can be realized by optical or video means [1]. In the video see-through mode, two video cameras are fixed on the HMD to capture the video streams of the real world for left and right eyes respectively. The video signals are sent to the computer, which renders the virtual objects at the right positions in each frame of the real scene images in real time according to the result of registration. The composite video streams are then sent back to the HMD, displayed on its miniature display units and imaged by the eyepieces to the right position for the user to observe. In the optical see-through mode, the user can see the real world directly through two optical image combiners, usually half mirrors with a semi-transmittance coating, which are placed in the light paths of the two eyepieces. The virtual objects are also generated by the computer according to the result of registration and displayed on the miniature display units in the HMD, which are imaged by the eyepieces to the correct positions in the real world, to provide the user with an augmented view. In any case, an eyepiece is needed for each eye, and its design determines many important specifications of the HMD.

For many applications, the HMD system needs to have a large view field and the display devices used in it are usually very small. The exit pupil diameter of the eyepiece should be large enough to fit for different interpupillary distance of different users, and its exit pupil distance should also be sufficiently large to accommodate users wearing glasses. Used in a head-worn device, the eyepiece should be light in weight and small in size. All these contradictory requirements add difficulties to the design of the eyepiece.

An ultra-light wide-angle eyepiece [8] is designed on the model of Takahashi's free-form prism structure [9]. To achieve a wide field of view of $45^\circ \times 34^\circ$ and an exit pupil diameter of 8 mm with a 0.61″ OLED, three XY polynomial surfaces are employed and optimized. The design layout, the HMD prototype assembly and the setup for the optical see-through mode with a compensating optical component are shown in Figs. 3 and 4 respectively.

Video see-through HMDs are applied in an AR exhibit item in Guangdong Science and Technology Center, which is composed of three different scenes including Undersea World, Pearl River Tour and Bird's Eye View of Guangzhou City from a Fire Balloon. It was initially designed to allow the visitors to move around freely, but later changed to sit the visitors on a trailer for safety reasons. The exhibit item has been in successful operation for more than a year.

Fig. 2. Interactive projection applied in artistic performance.

Fig. 5. Augmented view from the scene "Undersea World".

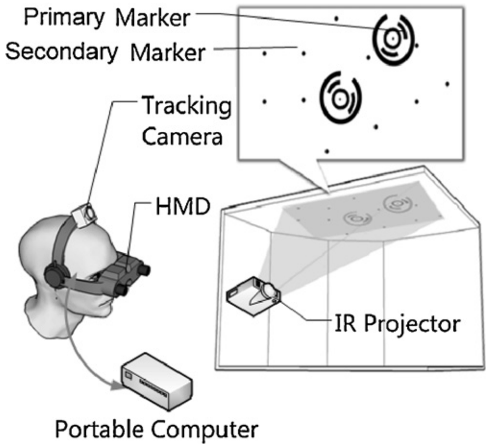

Fig. 6. System configuration.

Fig. 5 demonstrates an example of the augmented view from the scene "Undersea World", where virtual fishes are vividly to the real background of coral and seaweed.

A multi-user tracking system based on computer vision is developed for the real-time registration of the users of this indoor application. Visible fiducials such as ARToolkit [10] could be used, but they would produce an annoying side effect to the environment and interfere with the user's observation of the real world to a great extent. Thus, infrared markers are deployed instead, which also make the tracking system more stable and less susceptible to the influence of environmental light.

The infrared markers could be generated from flashing or patterned LEDs, but the installation of such markers would be complicated as LEDs need power supply. In order to remove the power requirement, Yusuke et al. [11] proposed reflective fiducials created by a special infrared reflective coating. During the system operation, the fiducials are illuminated by an infrared light, tracked and identified through an infrared camera to complete the task of pose estimation for the users. However, the installation of these coated fiducials in the working environment is still not easy.

In our system, the fiducials are projected on the specific location of the ceiling through a projector. According to the size of the tracking area, one or more projectors can be employed. The self-made simple projector(s) can be installed on the ground or the wall. Near-infrared light sources are used, thus the fiducials generated are invisible. In order to improve system stability and enlarge tracking area, the system adopts a primary–secondary marker deployment solution, as shown in Fig. 6. The primary markers use the circular two-dimensional barcodes which have the self-checkout ability, and the secondary markers are just simple light-spots, which are distributed around the primary markers. Using the identification results of a few primary markers, the system can further identify a larger number of secondary markers which have very simple structure. Therefore, this solution greatly simplifies the deployment and design process of the system.

4. Fixed-position viewing system

For outdoor AR applications, it is always difficult and sometimes forbidden to place artificial fudicials. Although intensive studies [12–14] have been carried out on markerless tracking using natural features, few of the current algorithms can be applied reliably outdoor due to the complexity of the environment and the drastic changes in lighting and the environment itself along the time of the day and the seasons of the year. Therefore, another kind of display devices, a fixed-position viewing system [15], is developed for certain outdoor AR applications. This display system can be operated

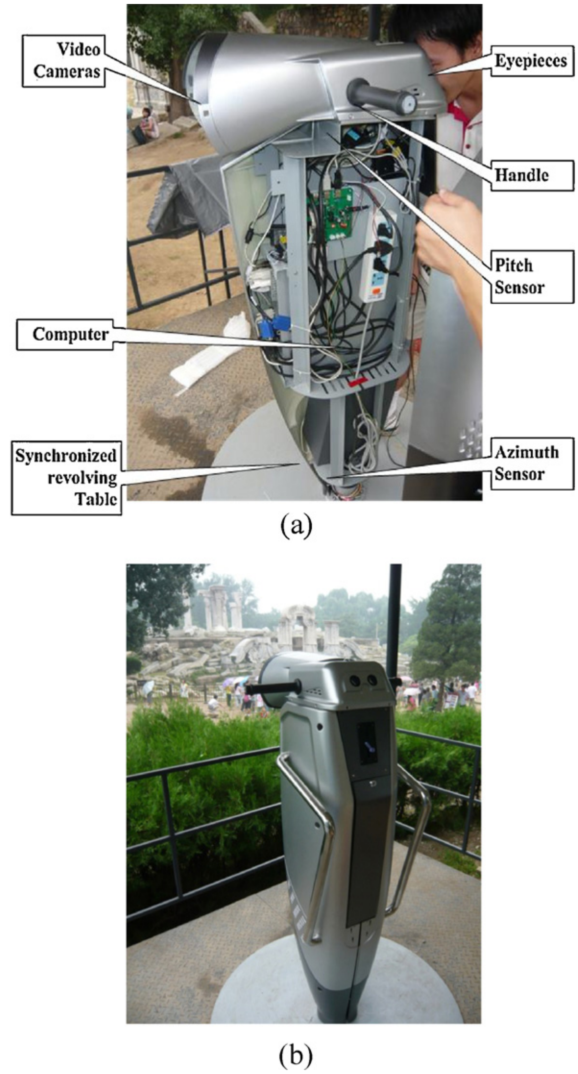

Fig. 7. Fixed-position viewing system.

in the same way as a scenery-viewing telescope often seen on the commanding point of a park or on the visitor's platform of a television tower, but its main function is augmenting the real scene with virtual objects rather than magnifying it. With its position fixed, the viewing system has only two degrees of freedom from angular movements, namely azimuth and pitch, which can be easily and accurately detected by two rotary encoders attached to the rotating axes.

Fig. 7 shows the configuration of the fixed-position viewing system. Its working principle is similar to a video see-through HMD. Two video cameras are installed in the front of the system, and the composite or augmented video streams are displayed on the LCD displays inside the system and shown to the user through the eyepieces.

The project of Digital Yuanmingyuan is an application using such a viewing system. Yuanmingyuan was a vast and beautiful royal garden built continuously by six emperors of the Qing Dynasty, which was once referred to as "the Garden of All Gardens". Unfortunately it was destroyed by the Anglo-French allied force in 1840s, and it is nowadays a relics park. With the help of our viewing system, the original magnificent buildings in this garden can be reconstructed on the natural field of the ruins for visitors to observe. The Western-Style Quarter of Yuanmingyuan was chosen as the first experimental site for this project, because relatively detailed

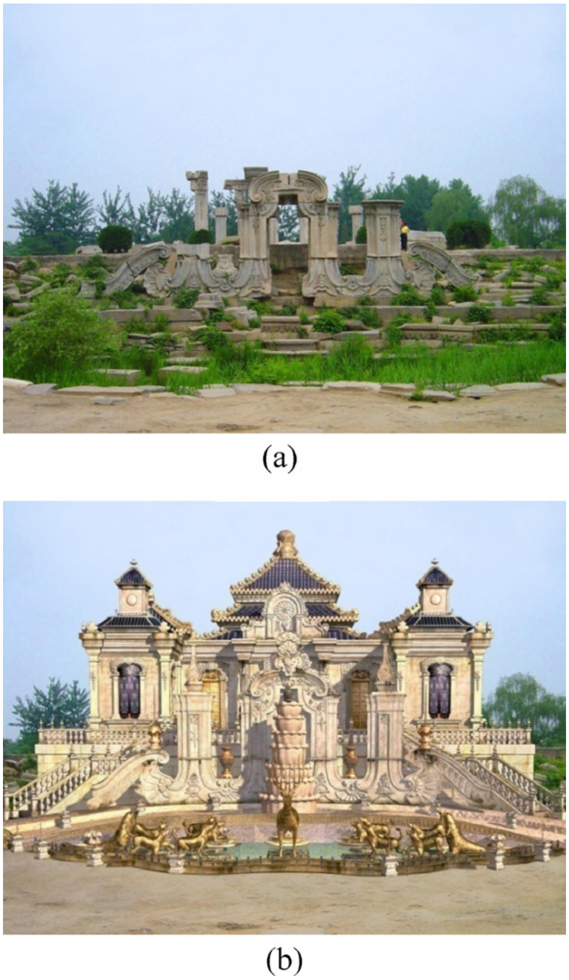

Fig. 8. Relic and augmented view of Dashuifa.

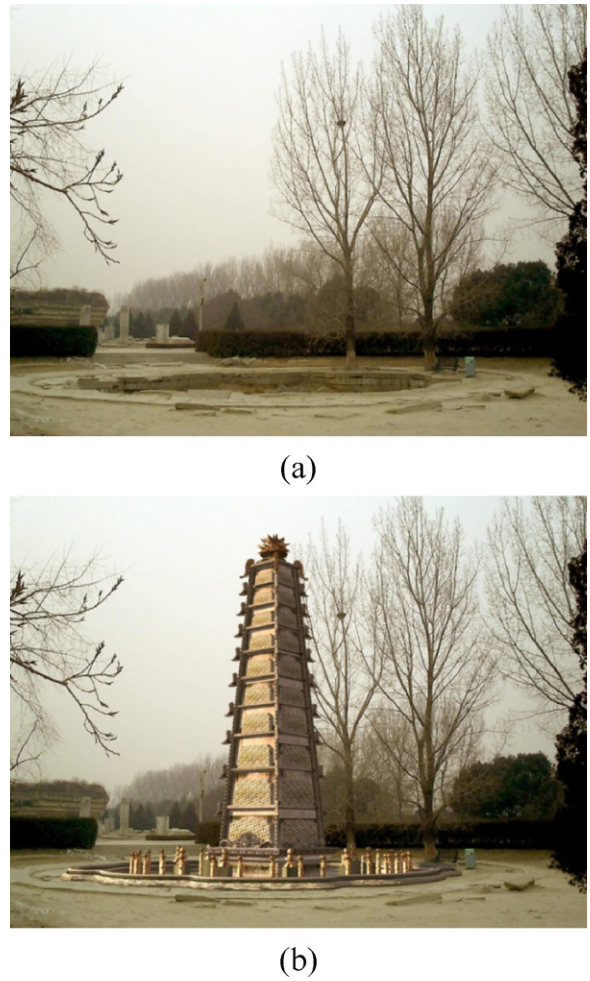

Fig. 9. Relic and augmented view of Xishuita.

information of the original scenery can be found from the huge stone ruins that still exist nowadays, and from the drypoint paintings by the missionaries who designed these building.

Figs. 8 and 9 give respectively the relic and augmented views of Dashuifa (Grand Waterworks) and Xishuita (West Water Tower) in the Western-Style Quarter of Yuanmingyuan. Fig. 10 shows the first generation of the fixed position viewing system operating on the site in our experiment. The visitors of the park queued up to try the system and almost all of them praised the result enthusiastically. Fig. 11 demonstrates the second generation of the system in a successful trial commercial operation. The final design as installed in Yuanmingyuan can be seen in Fig. 7.

5. Volumetric 3D display system

Volumetric 3D display is a new technology which in recent years underwent rapid development and started to be applied in different aspects in the field of computer vision. A prototype of the Perspecta Spatial 3D Platform was acquired from Actuality Systems, Inc. [16], and is used as the display device in an AR-based navigation system for endoscopic sinus surgery, which is developed jointly by our group and surgeons in the General Hospital of the People's Liberation Army.

The endoscopic surgery is a minimal invasive technique, which can effectively reduce the trauma and pain that a patient suffers during the operation to remove a polyp or tumor, for example, as well as the duration of hospitalization. However, because the endoscope offers only a limited keyhole view of the operated area, and because the sinus is very close to some vital parts such as the carotid artery, optic nerve, orbit and dura, incorrect movement of the surgical tool during an operation could lead to serious problems such as puncturing the cranial cavity, rupturing the carotid artery or injuring the optic nerve. Cerebrospinal fluid leakage and loss of sight are the two major complications of endoscopic sinus surgery [17]. A surgical navigation system tracks the surgical tool

Fig. 10. First generation system in experiment.

Fig. 11. Second generation system in trial operation.

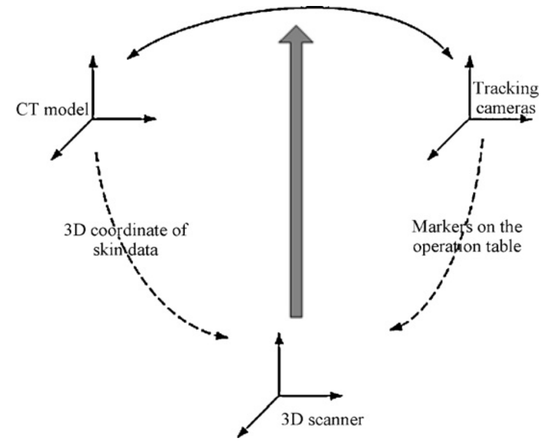

Fig. 13. Coordinate systems and registration.

in real time and shows the surgeon the relative position of the tool to the vital organs mentioned above so as to ensure that they are not harmed during the operation. For this purpose, CT (computed tomography) and/or MRI (magnetic resonance imaging) scans need to be performed on the patient to obtain parallel medical images of his (or her) head. Different segmentation and reconstruction algorithms are applied to these images to acquire a 3D model of the patient's skin, skull, brain matter, eyeballs, blood vessels and other parts of interest. During the endoscopic surgery, this 3D model is aligned with the patient's head by certain means of registration, and the position of the surgical tool is tracked, so that the relative position of the tool to the 3D model of the vital organs can be visualized. In existing surgical navigation systems, this relative position is shown to the surgeon in three plane views perpendicular to one another. The usage of volumetric 3D display technology represents a significant improvement welcomed by surgeons.

Fig. 12 gives a schematic drawing of our surgical navigation system, in which both the registration of the 3D model to the patient's head and the tracking of the surgical tool are realized by optical means. As indicated in Fig. 13, a common reference coordinate system is attached to a 3D digital scanner, which is employed to scan the skin on the patient's head as well as the reflective markers attached on the operation table. The point clouds from the patient's face are used to align the CT/MRI 3D model to the patient under this reference coordinates, and those from the reflective markers are used to calibrate the coordinate system of the tracker for the surgery tool. The tracker is composed of two cameras, which captures the infrared images of the reflective markers attached on the surgery tool and on the operation table during the operation. The computer calculates the position and orientation of the surgery tool in the reference coordinates in real time from the stereo images obtained by the cameras. With advanced algorithms, the tracking and registration can reach sub-millimeter accuracy and satisfy the requirement of the surgical navigation system.

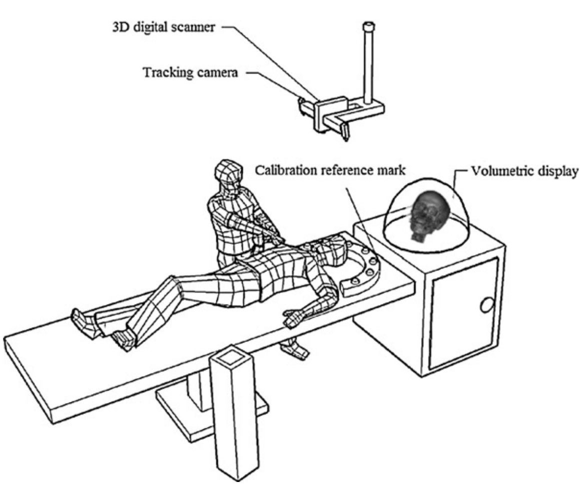

Fig. 12. Schematic drawing of surgical navigation system.

The Perspecta Spatial 3D Platform adopted a single high-frame-rate MEMS-based projector. Based on a rotating relay mirror mechanism, it projects images from different angles of one 3D model onto a rotating screen. The device is compatible with CT and MRI images in the DICOM format, and can display the 3D model of a patient's internal organs from a set of parallel CT or MRI images. However, the limit of transmission bandwidth between the display device and computer is the bottleneck for real-time display, which is essential for a surgical navigation system. There are 100 million voxels in each frame of volumetric 3D display, thus the data flow is huge. The problem can be partially solved by writing a specific application program with the Perspecta SDK, which directs the computer to transmit only the moving part of the 3D graphics, which, in the case of our navigation system, is the surgical tool. But in some cases when the surgeon wants to zoom in or zoom out the volumetric 3D display, for example, he will still have to wait, because now the whole display content is changed, and the amount of computation and data transmission is tremendous. We are currently working on a total solution by developing a new volumetric 3D display system that adopts multiple mini-projectors and a rotating screen [18]. Since the mini-projectors can be divided into groups and controlled by a number of synchronized computers, the data transmission problem can thus be eliminated.

6. Conclusions

Augmented reality technologies can be applied in many aspects of our life as seen in the examples described in this paper. To design a successful AR system, it is important to select an appropriate display device and corresponding registration method according to the specific requirements of the application.

Acknowledgment

The research activities described in this paper are supported by the National Natural Science Foundation of China (60827003), National Basic Research Program of China (2010CB732505), the Hi-Tech Research and Development Program of China (2007AA01Z325, 2008AA01Z303, 2009AA01Z308) and the

Innovation Team Development Program of the Ministry of Education of China (IRT0606). The authors are affiliated with the Key Laboratory of Photoelectronic Imaging Technology and System (Beijing Institute of Technology), Ministry of Education of China.

References

[1] R.T. Azuma, A survey of augmented reality, Presence: Teleop. Virt. Environ. 6 (4) (1999) 355–385.
[2] A. State, M.A. Livingston, et al., Techniques for augmented-reality systems: realizing ultrasound-guided needle biopsies, in: SIGGRAPH 96, New Orleans, USA, 1996, pp. 439–446.
[3] T.P. Caudell, D.W. Mizell, Augmented reality: an application of heads-up display technology to manual manufacturing processes, in: International Conference on System Sciences, Hawaii, USA, 1992, pp. 659–669.
[4] V. Vlahakis, M. Ioannidis, et al., Archeoguide: an augmented reality guide for archaeological sites, IEEE Comput. Graph. Appl. 22 (5) (2002) 52–60.
[5] M. Fjeld, B.M. Voegtli, Augmented chemistry: an interactive educational workbench, in: International Symposium on and Augmented Reality, Darmstadt, Germany, 2002, pp. 259–260.
[6] C. Scherrer, J. Pilet, et al., The haunted book, in: International Symposium on Mixed and Augmented Reality, Cambridge, UK, 2008, pp. 163–164.
[7] J. Simon, et al., The software architecture of a real-time battlefield visualization virtual environment, in: IEEE Virtual Reality 99, Los Alamitos, CA, USA, 1999, pp. 29–36.
[8] D. Cheng, Y. Wang, et al., Design of an optical see-through head-mounted display with a low F-number and large field of view using a free-form prism, Appl. Opt. 48 (14) (2009) 2655–2668.
[9] K. Takahashi, "Image display apparatus comprising an internally reflecting ocular optical system", (1996) United States Patent 5699194.
[10] ARToolkit library, http://www.hitl.washington.edu/research/shared_space/download/.
[11] N. Yusuke, K. Masayuki, Y. Naokazu, Localization system for large indoor environments using invisible markers, in: Proceedings of the 2008 ACM Symposium on Virtual Reality Software and Technology, Bordeaux, France, 2008.
[12] M. Ozuysal, M. Calonder, V. Lepetit, P. Fua, Fast keypoint recognition using random ferns, IEEE Trans. Pattern Anal. Machine Intell. 32 (3) (2010) 448–461.
[13] G.L. David, Distinctive image features from scale-invariant keypoints, Int. J. Comput. Vision 60 (2004) 91–110 (Kluwer Academic Publishers, Hingham, MA, USA).
[14] V. Lepetit, P. Lagger, P. Fua, Randomized trees for real-time keypoint recognition, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Washington, DC, USA, 2, 2005, pp. 775–781.
[15] Y. Wang, Y. Liu, et al., "Fixed-Position Augmented Reality Viewing System for On-Site 3D Digital Reconstruction", Chinese Patent ZL 2005 1 0105577.X.
[16] G.E. Favalora, et al., 100 million-voxel volumetric display, Proc. SPIE 4712 (2002) 300–312.
[17] D. Mehta, Atlas of Endoscopic Sinonasal Surgery, Lea & Febiger, Philadelphia, 1993.
[18] Q. Zhu, Y. Wang, et al., "Volumetric 3D display system based on multiple projectors and rotating screen", Chinese Patent ZL 2007 1 01078874.