Context-Aware Based Mobile Augmented Reality Browser and its Optimization Design

Yi Lin, Yue Liu[(✉)], and Yong-tian Wang

Beijing Engineering Research Center of Mixed Reality and Advanced Display,
Beijing Institute of Technology, Beijing China
linyi_bit@163.com, {liuyue,wyt}@bit.edu.cn

Abstract. The latest research shows that despite user's increasing interests in the Augmented Reality service, its usage frequency is low and the usage time is short. The intrinsic reason lies in its ignorance of human factors, which causes the design defect in the design stage of Augmented Reality browsers. In order to improve the quality of user experience, this paper introduces the design and optimization of a mobile Augmented Reality browser system by adopting the concept of context-aware service pushing based on scene classification, and proposes a novel interactive mode which combines Virtual Reality and Augmented Reality (AR/VR) on the basis of mental model theory. The whole design flow is verified by the design of the AR/VR hybrid interface. Experimental result shows that the proposed system of adopting hybrid interactive mode based on context-aware frameworks can not only increase the user acceptance of Augmented Reality browser, but also significantly improve quality of experiences of human-computer interaction.

Keywords: Context-aware · Mobile augmented reality browser · Mental model · User experience

1 Introduction

Grubert et al. [1] at the University of Graz, Austria, carries on analysis and research on the status of the use of Augmented Reality service based on the data collected in LimeSurvey [2] in 2011 and the survey results are shown in Fig. 1. It can be seen from Fig. 1(a) and (b) that people have more understandings and interests in the Augmented Reality technology. However, it can be seen from Fig. 1(c) that about a third of the participants (34 %) tried out the browsers only a few times and on the other hand 42 % used the browsers at most once a week. In addition, it can be seen from Fig. 1(d) that the average session time with an Augmented Reality browser was between 1–5 min. Such results indicate that despite user's increasing interests in the Augmented Reality services, there is a contradictory phenomenon of high installation rate and low usage

---

Y. Lin—No academic titles or descriptions of academic positions should be included in the addresses. The affiliations should consist of the author's institution, town, and country.

rate. Inherent reason for such discrepancy lies in the browser's own design. How to improve the Augmented Reality browser design, and how to mobilize the enthusiasm of users are all worthy of study.
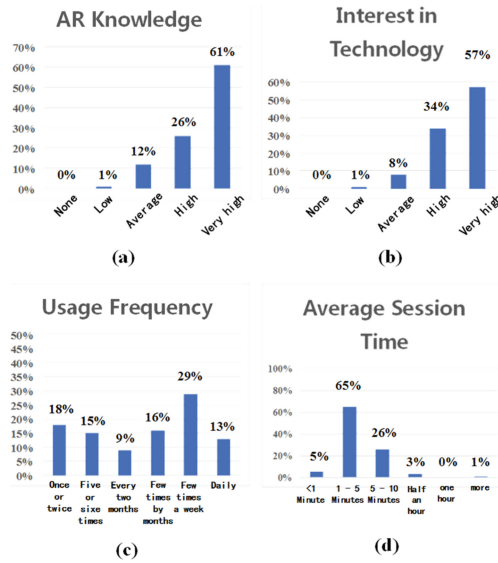

Fig. 1. User survey data of using augmented reality services (Source: Augmented reality browser survey, pp. 1–11)

At present, the following problems restrict the development of Augmented Reality browser, which includes poor cross-platform operation, hardware limitation of work mode and insufficient attention to the cognitive research. In order to improve the user experiences of mobile Augmented Reality browser, by analyzing the original requirements of users and finding the user's cognitive rules, this paper aims to create a new Augmented Reality browser system and optimize the design of new systems. With the development of software and optimization of interface design, the user experience of Augmented Reality browser can be significantly improved.

Lee et al. [3] of Chonnam National University in South Korea proposed a new method to deploy context-aware framework for augmented reality browser in 2008, the virtual models is embedded into physical environment by using augmented reality. It provides an immersive visual and interactive experience, realizes the bidirectional enhancement between the physical and virtual space. However, augmented reality visualization and interaction is established on the basis of tracking and registration based on markers. This system without markers will stop working in complex outdoor environment. Woensel et al. [4] proposed a new method which describes the context data and the relationships between each other by adopting semantic network technology in 2008. The context-aware framework constructed by employing their new method improves the efficiency of pushing personalized services. However, visualization of augmented reality required by application layer still is luck of in-depth research in this paper.

Coppola et al. [5] redefined the composition of context-aware browser in 2010. The search engine, Webpage content and automatically download module are integrated into the MoBe system they proposed. MoBe can provide detailed retrieval for web content according to the changes in the surrounding environment, however the user interface is still relatively drab, only images and text are used to organize the user interface, user's interactive experience still stay in two-dimensional space.

The mobile augmented reality browser and its optimization design proposed by this dissertation has two contributions. A contribution is that the classification accuracy of the system constructed by using the proposed scene classification method is improved, furthermore it deepen user's cognition of the operation object in an interactive environment. Another contribution is that AR/VR hybrid interactive mode based on Mental Model broadens the interactive space using AR technology. Not only can the consumer access virtual information in real world, they also can enter a fully virtualized environment for more information.

2   System Framework

Due to the hardware limitations of current mobile phone, such processing as image recognition which requires real-time and high-speed computing must be processed on PC side. The data exchange between mobile devices and Augmented Reality server is performed via a wireless network. In order to compensate the limited data processing ability of mobile devices, distributed architecture is applied in the proposed mobile Augmented Reality browser system. Different computing tasks are respectively performed in client and server. Both offline scene learning and online recognition are processed on the back-end cloud AR servers. Besides target tracking, user positioning and annotation rendering are also carried on the Mobile phone. With the exchange and processing of data across two side, AR/VR hybrid human-computer interaction experience can be provided to the customers on the mobile phone platform. First the camera is launched to capture image sequences of real world and feature points are extracted from the captured images. Then the data of feature points and user's position are sent to cloud server through wireless network. The target information is quickly recognized in cloud server and sent to mobile phone after the sample training is completed. Finally the virtual items are rendered around the real objects for augmented display. With the help of the attitude data from compass sensor, mobile client can achieve precise tracking, which means that the virtual tags always move with the real objects. Figure 2 shows the frameworks of MARB system.

2.1   Large Scale Scene Recognition

The application of Augmented Reality technology in mobile devices means that the mobile devices should have real-time image acquisition, 3D scene display and orientation tracking. Higher requirements for the calculation of equipment capacity and speed is made. Automatic target recognition which has the largest amount of calculation is processed by high-performance computing server.
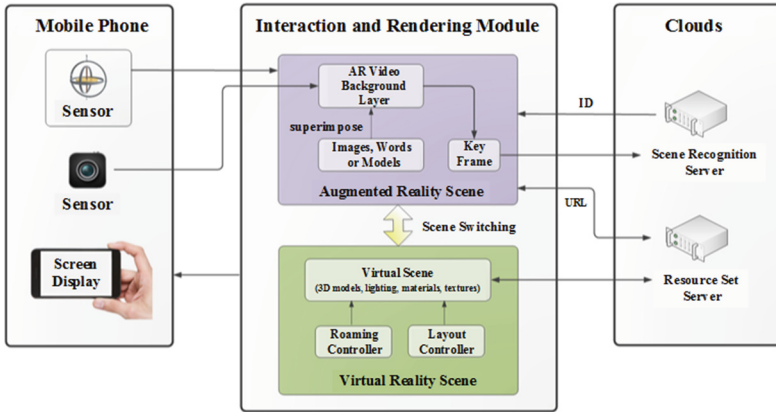
Fig. 2. Framework of MARB system

There are two stages for the large scale scene recognition by Mobile Augmented Reality Browser (MARB) proposed in this paper. Rough positioning of the point of interest (POI) is first achieved through the location data collected by GPS, then the sensing data and information of the target are transmitted to the cloud after the feature information is captured by a mobile phone camera. Finally, using the visual information retrieval method carries on building accurate recognition by the server.

Speeded up robust features [6] (SURF) has been widely used in image recognition and description. In order to achieve the accurate identification of large scale outdoor scene, local features (SURF key points) is adopted as the feature of object description [7] by MARB. The image information of the object is mapped into a set of keywords by using statistical analysis algorithm based on dictionary [8] proposed by Nister in 2006. By comparing the set of keywords constructed by training sample and set of keywords for image querying in online stage, the image that is the most close to the training sample is chosen as the final recognition result.

2.2   Tracking and Registration

The MARB system uses Keyframe-based real-time camera tracking [9]. In order to realize parallel tracking and matching for natural feature [10], parallel registration strategy based on double thread was firstly adopted in the system design. Then KLT (Kanade-Lucas-Tomasi) algorithm [11] is employed to tracking and registration in one thread and BRISK feature matching algorithm is used for wide baseline correction and the corresponding relations between 2D and 3D in another thread in another thread. Finally, by fusing the matching information based on the model and information of front and rear frame, the Objective optimization function is constructed. Finally, the iterative optimization algorithm is utilized to improve the tracking accuracy of the positioning.

3   Context-Aware Service Pushing Based on Scene Classification

3.1   The Classification of the Scene

As the best form of expression for providing relative information of physical environment in Augmented Reality technology, annotation has become a major component to improve the understanding of the real world. Such indicators as amount, preference, layout and view of combination are used to evaluate the representation of service information. By comparing current popular mobile Augmented Reality browser's annotation both in China and abroad based on four test indicators above mentioned such as Junaio, Layar, Wikitude, City Lens, Senscape and City One Click, survey results show that the current mobile Augmented Reality browser has the following problems. Firstly, it will bring the user's cognitive experience great inconvenience when the number of annotation is more than 20. Secondly, on the representation of annotation, too many hue will increase the cost of user's perception to the POI and reduce the efficiency of target retrieval. Thirdly, there are still such unsolved difficulties with the marker less recognition based on computer vision in the outdoor environment as lighting change, occlusion and transmission efficiency.

In order to solve the above-mentioned problems, this paper proposed a new concept of scene. The user must perform multiple steps of the task to achieve a goal. The environment (context) will change during the time of performing such steps of task. Multiple sets of tasks to be achieved by user in a context is defined as a scene. User will experience the process of performing multiple sets of tasks when planning to achieve a goal. Therefore, the system can provide services to assist the user to complete different tasks based on the perception and reasoning to the current context. Grouping tasks by using scene can not only easy to push perceived service for the user by the system, but also deepen the cognition of the operation object in an interactive environment. Take real estate information query as an example, a number of tasks are divided into three scenario of POI's positioning and selection, path finding navigation and POI's recognition and browsing in this paper.

3.2   The Structure of the Module and Context-Aware Services

(1) POI's Positioning and Selection Scenario. The structure of the module in POI's positioning and selection scenario is shown in Fig. 3. There are location and tracking module, the image capture module, context adapter and rendering module in the mobile client. Location service database and context reasoning in context service layer are distributed in cloud.

User firstly enters the POI's positioning and selection scenario after MARB is launched. User's goal is to identify and select the POI in this scenario. Therefore the first task to Figure out their location. Then, built on location search center, all POI are scanned within the specified radius. Finally, user should make the selection decision after the profile information of POI is checked through. In this scenario, according to the tasks the user needs to perform, MARB provide context-aware
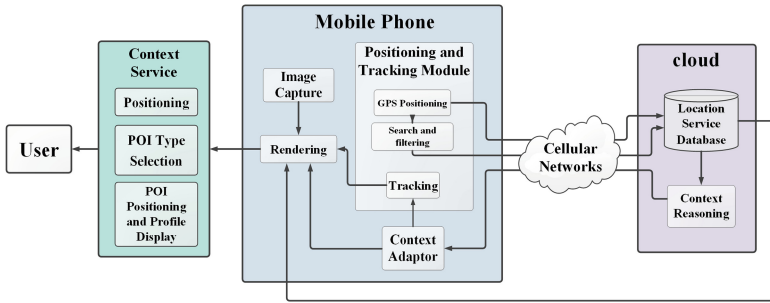
Fig. 3. Module structure and context-aware service in the scenario of POI's positioning and selection

services such as user positioning, POI type selection and POI positioning and profile display.

(2) Path Finding Navigation Scenario. The structure of the module in the path finding navigation scenario is shown in Fig. 4. There is context adapter, map calls, navigation mode selection, path generation and rendering in mobile client. Location service database and context reasoning is distributed in cloud.
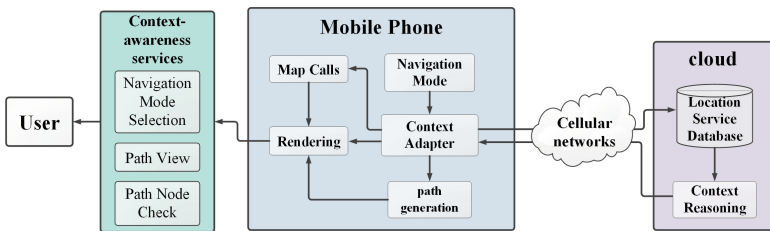

Fig. 4. Module structure and context-aware service in the scenario of path finding navigation

The navigation request is sent after the label of POI is selected in the first scenario, then MARB change into the path finding navigation scenario. By drawing a shortest path from the current position to the appointed place on a flat map, path navigation service is provided by MARB for users. In order to taking convenience for users to get the information of arriving at destination location by using different modes of transport, there are not only many navigation mode such as walking, by bus, by car, etc. available to select, but also draw a path for every navigation mode in this scenario. User can check node information such as bus stations, subway station or intersections, etc. through the path.

(3) POI's Recognition and Browsing Scenario. The structure of the module in POI's recognition and browsing scenario is shown in Fig. 5. There is image capture, listener, context reasoning, context adapter and rendering module in the mobile client. Location service database and resource services database is distributed in cloud.
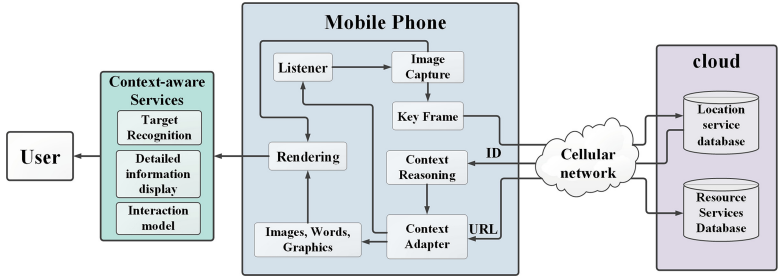
Fig. 5. Module structure and context-aware service in the scenario of POI's recognition and browsing

The user is firstly find a building wanted to know after reaching the real estate scene of POI. Then look at building details and deepen understanding through the interaction with the 3D model. Details of building such as building structure, type area, sales price and property information, etc. is present in two-dimensional images or text labels. In order to bring a more intuitive browsing experience, user can rotate or scaling a three-dimensional model of the building by touch control during the interaction. MARB provide context-aware services such as navigation mode selection, path view and path node check.

4   Improvement of Interaction Model

Compared with human-computer interaction in real environments, Augmented Reality can create a more immersing interactive experience for the user. However, the current interactive mode of Augmented Reality browsers only improves interactive experience in the horizontal dimension of user perception. It still lacks a novel interactive mode to guide the user obtain an in-depth understanding to POI from vertical aspect. Therefore, being aimed at the goal of improving interactive mode, this paper proposed an Augmented Reality and Virtual Reality (AR-VR) hybrid interactive mode based on the theory of mental mode. This type of interactive mode based on user's requirement has both types of Augmented Reality and Virtual Reality technology features.

In the outdoor part of the proposed system, augmented information which usually is concerned about is actively pushed to customer when preview housing such as the sales prices, residential area, surrounding environment, traffic conditions, and way of contact. The customers always have the demand of taking a closer look of inside structure of housing when some set of apartments is chosen. Clicking on the corresponding housing units preview image, interactive space can be smoothly switched to the VR environment. In pure virtual space, the user can not only roam in the house, but also replace the furniture or adjust the layout of items according to their own wishes. Touching AR button at the lower right corner, it can be returned again from VR indoor space back to outdoor AR scene. It is shown in Fig. 6.
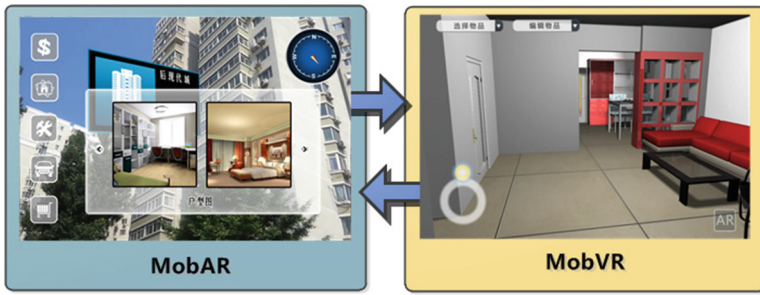
Fig. 6. AR and VR scene switch

Mental Model concept was first proposed by the Scottish psychologist Kenneth Craik in 1943, Johnson-Laird and others [12] give a clearer representation of this conception in the book <Mental Model>, He thinks Mental Model is the structural analogy in physical world, it generated after the information is offered and filtered by perception (visual, auditory, tactile, olfactory). Young suggested that Mental Models consist of several components, each part is divided into several groups, and the entire model can be described by series of affinity diagram of behavior [13].

Take the example of AR-VR hybrid mobile browsing system, the affinity diagram of behaviors is shown in Fig. 7. Such parts as check structure preview image and switch to indoor separated by the longitudinal axis above the horizontal axis are called mental space. Each mental space is cut into several parts. Below horizontal axis, interaction module is corresponding to each mental module. By contrast, weak point can be found from the interactive function. By interactive operation of click on image, AR space in the right side can be switched into VR environment from left side.

The Metal Model can be classified as two types of models. One type is Macro Mental Model. Its recognition must satisfy the user's original psychological needs of the product
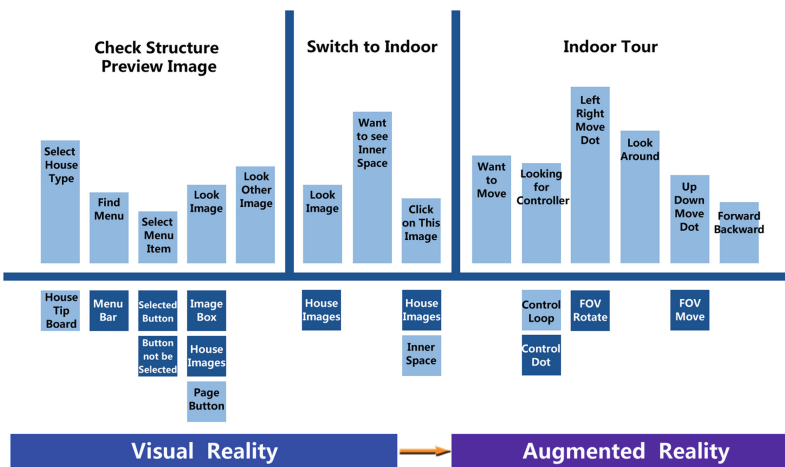

Fig. 7. Affinity diagram of behaviors

or service during the time of using the product or service by user. Another type is Micro Mental Model. Micro Mental Model refers to the Mental Model when a specific operation through the whole system is completed by user. After iterative usability testing, cognitive conflict point between the aspects of visual, interactive, copywriting and user can be solved.

The user experience interface should be designed to match the user's Mental Model. It means the interactive interface must correspond to user's habits and life cognition. In view of this, humanized design principles should reflects the ease of use and ease of identification of the design plan. The transform and match between Design Process Model and Cognitive Process Model is shown in Fig. 8.
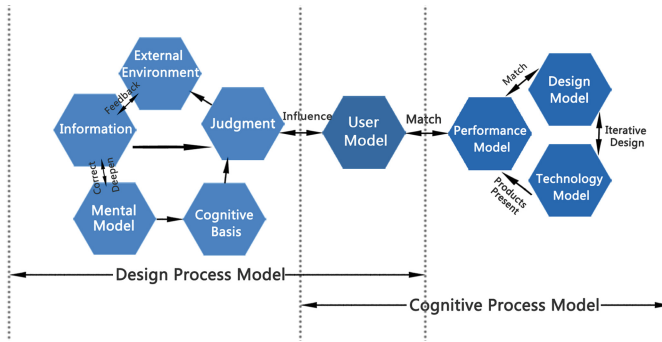

Fig. 8. Structural model of the user cognitive processes and product design

Macro Mental Model is a description of user's needs and demand structure. The purpose is to help designers to construct the experience process from the overall perspective. Micro Mental Model describe the interactive process between the user and interface just from the operation level. It can validate whether this interface is available, easy to use and easy to learn. During the process of housing information exchange, AR-VR hybrid mobile browsing system makes the use of macro Mental Model to build up the entire structure which interface process required based on user's psychological cognitive structure. Then, the macro Mental Model which has already horizontal built is taken into a specific vertical task to make up micro Mental Model. A specific operation workflow of a task in the entire system is shown in Fig. 9.

5   Design Example of a VR/AR Hybrid Interface Based on Context-Aware Service

5.1   Interactive Design in VR Scene

After the transverse and longitudinal design by mental model, interface of VR Indoor browsing and editing is produced. Scene roaming is conducted in first-person perspective view. As shown in Fig. 10, the direction controller is placed on the lower left side of the VR interface. Clicking-hold round point and sliding toward any orientation by
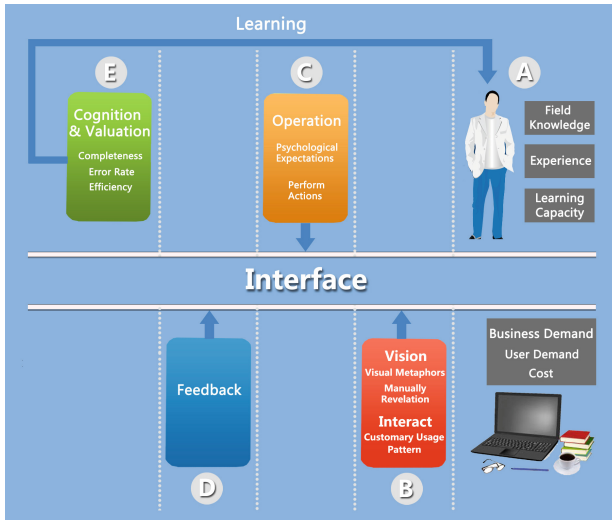
Fig. 9. Single-tasking operating process of micro mental model

Fig. 10. Indoor selecting and editing effects

touch input, user can move in relative direction in virtual space according to the controller. Two buttons are placed on upper left side of interface. The left button is item selecting, while item editing button at the right. Items in this circumstance are highly customizable. It means that objects can be replaced and rearranged to satisfy the needs of users. After the item is selected, clicking on item editing button, user can operate commands such as move, rotation or scaling which is appear in the pop-up menu. Only requires simple touches, by the editing operation to the objects, user can have immersive experience using different decoration scheme. A richer interactive experience for mobile housing preview is realized in this way. Scene can be switched back to AR scene when clicking on the AR button on the lower right corner of the VR interface.

5.2   Interactive Design in AR Scene

For the UI design in AR scene, considering the use environment is moving, limited to the space of interactive operation is narrow caused of mobile phone screen size, user cannot keep watching the screen for a long time compared to desktop computer screen. Therefore, the principles of interface design should reflect the concept of simple, easy to understand efficiency and clear at the first sight. UI design should be based on the following principles. First, due to the interference of video background, UI elements should be simple, intuitive, metaphorical and easy to identify. Second, gesture commands must be supported for slide, scale or other gestures. Last, switching design between scenes should be easy to understand and operate.

According to the preliminary analysis of user needs and later usability testing, visual design of UI is consist of icon, menu, description box and navigation compass. The navigation compass provides users with a sense of direction. The function menu offers information classified by different option item. There are two categories of description box, respectively is text boxes and picture boxes. This "drawer type" hierarchical menu can not only simplify the interface, but also consistent with the user's mental mode in the operation logic. According to the classification of type of housing by 3D Tips Board, there are different menu items corresponding to different types of housing. It simplifies the user's operation process and avoids interactive barriers caused by the confusion of concepts. Visual metaphors are used by two kind of next page button to guild the user to query for more hidden information. User interface design is shown in Fig. 11.

5.3   Transformation Between Virtual Reality Space
and Augmented Reality Space

By calling the checkpoints script of Unity engine, MARB system can switch scenes between Virtual Reality and Augmented Reality. Every scene obtains an identity number of corresponding checkpoints after the Virtual Reality and Augmented Reality scene

Fig. 11.   Indoor selecting and editing effects

files which have been completely designed are added into "building settings". Scene switching command "Application.LoadLevel(index)" is triggered by button events "GUI.Button" in a script when users are required to switch scenes. Index parameters in this command corresponding to the previous identity number of checkpoints in "building settings". In accordance with the above steps, scene switch can be made.

6   Evaluation Experiment and Conclusion

By randomly selecting 30 participants, the preliminary evaluation of AR interface as shown in Fig. 11 is conducted. The evaluation team consists of 18 males aged 20–30 and 12 females aged 22–31. Two male participants have been involved in the testing of Augmented Reality application. Questionnaire survey is designed to measure four aspects which are most important in human-computer interaction design: Aesthetic, Navigation, Recognition and Efficiency. Aesthetic is the measuring element of slinky of UI. Rationalization of interact can be inspected through Navigation aspect. Recognition is an important indicator of the visual design. Efficiency factor show the logic of interaction design.

The results of evaluation experiment are shown in Fig. 12 by radar chart, aesthetic is the measuring element of slinky of UI. Rationalization of interact can be inspected through Navigation aspect. Recognition is an important indicator of the visual design. Efficiency factor show the logic of interaction design. Predilection is the comprehensive evaluation of the scheme. According to the statistical report of feedback from evaluators, three kinds of software are basically the same in aesthetic and recognition aspects. It means that user is able to correctly identify the operating indicator during operating period. The aesthetic design of the interface can be separated from the complex video background. In navigation aspect, MARB and Layar are both better than Wikitude. Due to the hierarchical menu design, the former is lightly better than the latter. This is attributed to the iterative design based on micro Mental Model. Affected by the navigation factors, MARB and Layar have better efficiency than Wikitude. During the period of single task, the average response time for button interaction is less than 0.3 s.

Taking the logic order of operation for retrieving and browsing POI, numerous operating behaviors are divided into three scenarios which include "POI's positioning and selection", "Path finding navigation" and "POI's recognition and browsing". The system module of MARB providing context-aware services is built in each scene based on context-aware services framework. The efficiency of pushing Augmented Reality services is improved for each module in the scene. In addition, Virtual Reality and Augmented Reality hybrid interaction mode widen the interactive way for human-computer interaction, increase the user acceptance of browser and significantly improve human-computer interactive experience.
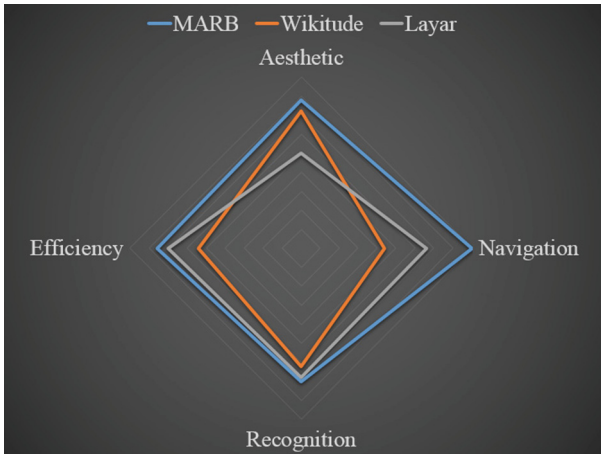
Fig. 12. Radar chart of evaluation experiment

References

1. Grubert, J., Langlotz, T., Grasset, R.: Augmented reality browser survey, Institute for Computer Graphics and Vision, University of Technology Graz, Technical report, 1101, pp. 1–11 (2011)
2. May, P.: Limesurvey. www.limesurvey.org
3. Lee, J.Y., Seo, D.W., Rhee, G.: Visualization and interaction of pervasive services using context-aware augmented reality. Expert Syst. Appl. **35**(4), 1873–1882 (2008)
4. Van Woensel, W., Casteleyn, S., De Troyer, O.: A framework for decentralized, context-aware mobile applications using semantic web technology. In: Meersman, R., Herrero, P., Dillon, T. (eds.) OTM 2009 Workshops. LNCS, vol. 5872, pp. 88–97. Springer, Heidelberg (2009)
5. Coppola, M., Della, M.V., Gaspero, D.L.: The context-aware browser. Intell. Syst. **25**(1), 38–47 (2010)
6. Bay, H., Ess, A., Tuytelaars, T., Van Gool, L.: Speeded-up robust features (SURF). Comput. Vis. Image Underst. **110**(3), 346–359 (2008)
7. Bay, H., Tuytelaars, T., Van Gool, L.: SURF: speeded up robust features. In: Leonardis, A., Bischof, H., Pinz, A. (eds.) ECCV 2006, Part I. LNCS, vol. 3951, pp. 404–417. Springer, Heidelberg (2006)
8. Nister, D., Stewenius, H.: Scalable recognition with a vocabulary tree, Scalable recognition with a vocabulary tree, pp. 2161–2168 (2006)
9. Dong, Z., Zhang, G., Jia, J., Bao, H.: Keyframe-based real-time camera tracking, pp. 1538–1545 (2009)
10. Yang, L., Chao, S., Mingmin, Z., Wenjie, Z., Zhigeng, P.: Parallel tracking and matching for natural feature based AR registration. J. Image Graph. **16**(4), 680–685 (2011)
11. Birchfield, S.: Derivation of Kanade-Lucas-Tomasi Tracking Equation, Department of Electrical and Computer Engineering, Clemson University (1997)

12. Johnson-Laird, P.N.: Mental Models: Towards a Cognitive Science of Language, Inference, and Consciousness. Harvard University Press, Cambridge (1983)
13. Young, I.: Mental Models: Aligning Design Strategy with Human Behavior. Rosenfeld Media, Brooklyn (2008)