

Towards Mobile Augmented Reality (MAR)

Xingdong Ma

School of International Education

Canvard College, Beijing Technology And Business University

davidxm99@126.com

Keywords: Augmented Reality (AR); Mobile Augmented Reality (HAR); Cloud Computing

Abstract. This paper first introduces the concept and the characteristics of Augmented Reality. The overall system structure and the functionalities of the components are outlined, followed by the discussion of the architectures of HAR. The cloud computing platform is proposed.

Introduction

The "reality-virtuality continuum" relates to any particular situation where a mixture of scenes or objects, real or virtual, is displayed, as illustrated in Figure 1. The *real environment*, at the left end of the continuum, is the environment that consist only real objects, such as the display of a real-world scene. The *virtual environment*, at the right end of the continuum, defines the environment that consists only virtual objects, such as a computer graphic simulation. As indicated in Figure 1, a Mixed Reality environment is one in that real world and virtual world objects are presented together within a single display.

Augmented Reality, AR, is a typical interdisciplinary developed on the basis of virtual reality technology, VR, based on very wide range of researches and applications, involving many technical fields, such as computer graphics and image processing, HCI design, psychology, mobile computing, computer network technology, signal processing technology, and information visualization, and design of new display and sensors, etc.[1]

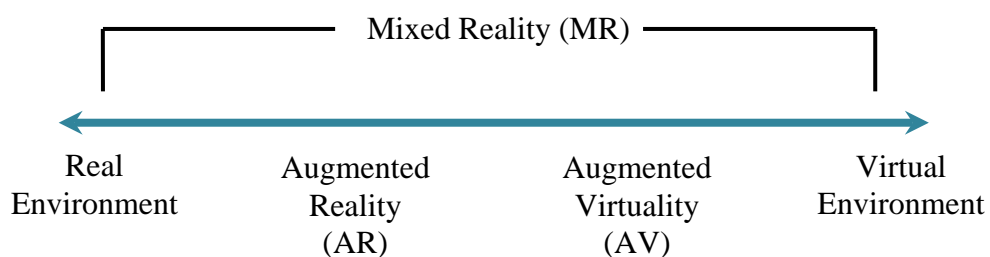


Figure.1. Reality-Virtuality Continuum

Unlike traditional virtual reality technology, AR superimposes computer generated virtual images information accurately on the real objects images in real time, allowing users view the combined images in real time, and to interact with AR devices seamlessly. AR technology needs to have the characteristics in three areas: (1) a combination of real and virtual images, (2) real-time interaction, (3) alignment of the virtual image according to the real physical objects [2].

With the advances in mobile technologies, the mobile devices, smart phones and tablets, are equipped with equal or more processing power than the desktop computers [3]. Augmented reality applications, with media-rich contents that combine the realistic and the virtual environments, will be the most promising trend for future mobile devices.

Despite the improved processing power, the mobile devices are still incapable of providing real-time AR applications which include complex real scene analysis and realistic virtual scene rendering [4]. In order to reduce the workload of the mobile devices, wireless AR structure is proposed in this paper.

The Structure and Algorithms of HAR System

The Structure of AR System. The structure of the functional modules of HAR is very much the same as the general AR, as shown in Fig. 2.

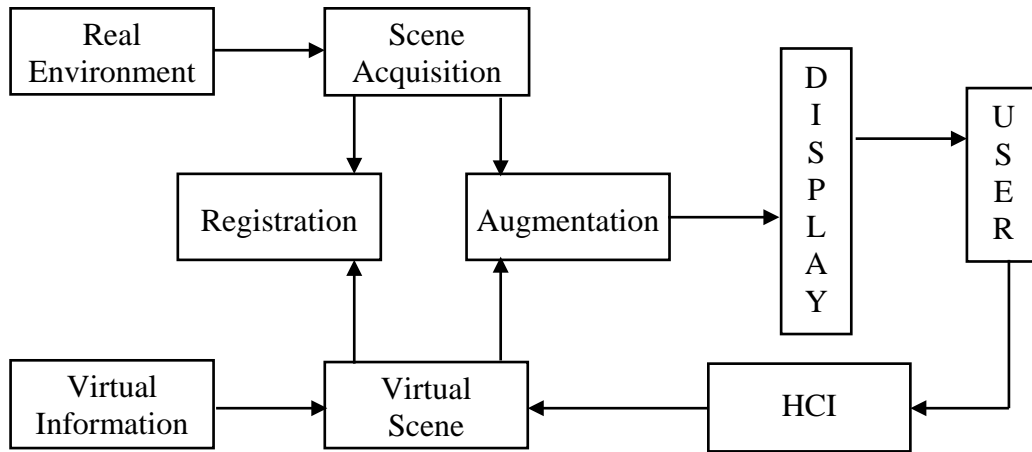


Figure.2. Structure of Function Modules of AR System

A typical augmented reality system is usually composed of several subsystems: the scene acquisition system, tracking and registration system, virtual scene generator, and augmentation (virtual real synthesis) system, display system and human-computer interaction interface. The scene acquisition system (camera) is responsible for accessing to information in the real environment such as the environment image or video sequence, the tracking and registration system is responsible for tracking the camera orientation and view direction, the virtual scene is responsible for rendering the virtual information, the augmentation system is responsible for the synthesis of the virtual scene and the real scene. Augmented reality system establishes the real space and the virtual space according to the geometric consistency of embedded real space, and forms a fusion of virtual and real environment, and then transmits the result to the display system that presents to the user with the equipment and environment for interaction. The accurate registration, together with the final display of the output, determines the user's final perceived environmental effect.

Real world space can be established by processing the real environment image. According to the tracking registration technology, the camera pose and the virtual image of the spatial location can be determined; the augmentation is performed through the arrangement of the virtual image and real image registration.

Scene Identification Algorithms. The scene identification module is a key unit for augmented reality. The image recognition includes two stages: target feature points extraction (feature point detection, feature point description), and target feature points matching. The feature extraction and matching is generally to search for the feature points in the image, and with a strong feature description. The feature points to environmental may change according to some transformations, such as rotation, scaling, and lighting, etc.

Visual tags are usually used as markers to be put in the real scene of interest. They can be easily recognized and identified by the AR system. They can be used both indoors and outdoors.

Based on the identification of the tracking registration technology, the image of the target is placed in the scene, and the camera acquired images must include this identification image. By the feature test of the scene image acquired by the camera, the identification marks and extracts the reference points (such as the corner points, the center point, etc.). The logo is placed artificial, so the sign the coordinates of the reference point in the 3D scene knowledge known, also camera intrinsic

parameters are known, which according to perspective projection algorithm in computer vision, you can get a camera with respect to mark the transformation matrix, to obtain registration information.

Current registration systems based on tracking are: ARTag [5], ARToolKit [6], etc. ARToolKit is suitable for small scale application, while ARTag is faster when dealing with large scale application.

The computational complexity for marker based tracking, registration and identification is relatively low, with good accuracy. However, there are also some shortcomings, such as when the logo is partially occluded, the system will not be able to complete tracking and registration.

HARs that do not use artificial markers usually facilitate image recognition technologies. The markerless based identification, tracking and registration technology, also known as natural feature points registration and tracking, does not require any artificial marker being placed in the scene. It directly uses the natural features of the scene and the recognizable features of the real objects, such as inflection points, surface material, etc., as the identified reference point for feature extraction.

Because the real scene are often complex, under different light intensity, different perspectives, different relative motion, etc., the difficulty of markerless based tracking registration is much more difficult than that of marker based approach. The classical algorithm are Harris [7], SIFT [8], SURF [9], etc. The comparison of the algorithms can be found in [10].

Table. 1. Comparison of the Algorithms

| Algorithm | Accuracy | Localization | Scale Invariance | Rotation Invariance | Real Time |
|------------------|-----------------|---------------------|-------------------------|----------------------------|------------------|
| Harris | Average | Average | None | Average | Average |
| SIFT | Very good | Best | Good | Average | Bad |
| SURF | Best | Very good | Average | Best | Average |

Current techniques in image retrieval of computer vision can be based on two types: global features and local features. The global features mainly refer to the color features and contour features, whereas the feature is mainly the steering characteristics. The color feature recognition rate is generally not high but the time consuming is shorter. PHOG (Histogram of Oriented Gradients) is a commonly used for contour feature. The PHOG features have low recognition accuracy under complex scenes, and it cannot meet the requirements of high recognition rate under the massive target with the increase of number of targets. Local feature is the steering characteristics, feature points in preserving the important features of image and graphics at the same time, can effectively reduce the quantity of the data, thus improving the content of information and speed up the calculation. The global features is conducive to the image of the reliable matching, the real-time processing possible.

There are two kinds of feature extraction algorithms, the edge based feature point detection, and the gray scale feature point detection.

The edge detection algorithm is based on the detection of the edge of the image, and the quality of the edge detection is the basis of the detection algorithm. This method has too much dependence on the image segmentation and edge detection. Any problem during the process of edge detection will have a greater impact on the extraction of feature points, or even get the erroneous feature points and image segmentation. Thus it is not suitable for real-time processing. And based on gray feature point detection algorithm only based on image gray information of the image of feature point detection without relying on other local feature, with a fast algorithm, more in line with the requirements of the real-time, at present commonly used feature point detection algorithm based on gray with Harris and Susan, fast, sift and surf algorithm, each have its unique advantages, in the practical application need to choose according to the specific circumstances.

The calculation of the registration of the unlabeled tracking is large, and the application of the equipment is limited because of the limited equipment resources, so the application of the mobile device is affected.

Wireless Processing Architecture

Although the processing power has been greatly improved, the state of the art mobile devices, smart phones and tablets, have the capabilities of speech recognition and video communication; AR applications demand even more resources for media-rich contents. There are two approaches to develop HAR applications, to develop more advanced hardware, and to facilitate the new processing architecture. There are two major architectures in ARs, autonomous architecture, and distributed architecture [11].

Autonomous architecture, Figure 3, install all the AR modules within one single device, such as a smart phone or a tablet. The processing is performed on the single device as well. The autonomous system requires much greater processing abilities of the mobile device.



Figure. 3. Autonomous Architecture of HAR

Distributed architecture, Figure 4, on the other hand, delegate the processing to a remote server. The server thus performs the scene analysis, registration, and augmentation, while the mobile device is only responsible for scene capturing and visualization.

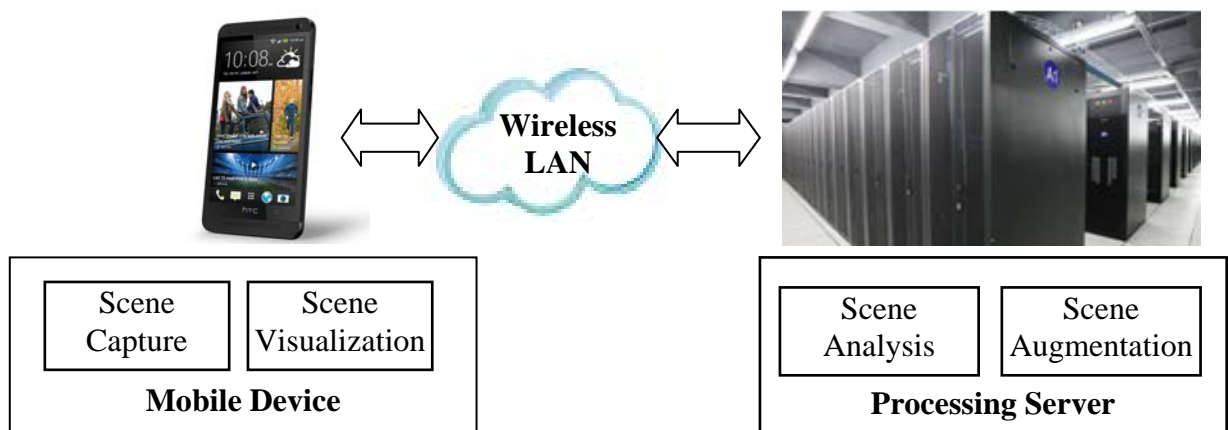


Figure. 4. Distributed Architecture of HAR

Augmented reality show the effect of the need for a large number of complex operations, including of user location tracking, rendering, rendering, these complex computing tasks to in server of the mobile terminal and a system of coordination, the part of the work in the mobile terminal by the

mobile phone to realize and another part of the complex operation put in the server implementation. High performance multi-core graphics workstation computing and processing ability is far superior to the mobile device. At the same time, 4G network has made data transmission makes real time network communication of the large amount of data possible, mobile devices and high performance workstations; collaborative integrated work environment is the key of system design. The advantage of the client/server (C/S) architecture for the implementation of the complementary between mobile devices and servers. With the computation of initialization and calibration, three-dimensional modeling will by a multi-core server completed construction of a cloud based computing technology support large-scale training images and fast retrieval and highly concurrent enhanced content access service integrated service platform, to achieve the system with high concurrency, high availability, scalability characteristics; and the image acquisition, feature extraction, the actual rendering, real-time registration will by the mobile device is completed.

A typical computer vision C/S architecture based on mobile augmented reality system consists of image acquisition, image preprocessing, feature extraction, wireless transmission, image recognition, tracking, rendering module, the system through the camera to obtain images, and pretreatment, reduce network traffic for image compression, through the communication module to upload the server, the server for feature extraction and image recognition, the server will return the target mark according to the recognition result, enhance the content of the corresponding query, target identification, enhance the content returned to the client through the communication module, the client can use the tracking registration module get the virtual information added to the dynamic position of the real scene, the rendering module to overlay virtual the enhanced content information to the client real-time camera interface.

To enhance the target recognition using C/S architecture, make full use of massive cloud storage and computing power, image recognition to generate samples and training, so as to realize the fast object recognition. The key algorithm of target recognition is used as the application of the algorithm, and can be used to control the instance and parameters of the algorithm.

To enhance the real information and training sample data storage needs of the massive augmented reality respectively research and develop the reality storage subsystem and enhance the practical interest point information management subsystem. The former building on top of the non structured storage mechanism, using based on bucket hash scheme, both to avoid the common consistent hashing the scalability problem, but also to avoid the locating in DHT network delay problem. The information management subsystem of interest points is built on the structure of the structure, and is the interest point information, such as all kinds of metadata information, and index to speed up the retrieval speed.

For large-scale image processing, the use of map/reduce based parallel computing framework, improve the efficiency of sample training and target recognition.

In order to share the enhanced real business request load, cache hotspot data, enhance user experience, will deploy a number of users access cloud near the user's place. The service access cloud will receive all kinds of requests from users, enhance the actual information of the hot spot, and reduce the load pressure of the real service processor.

For the AR directly affect the user experience of positioning technology, the technical scheme of integrated positioning technology to provide more efficient, the base station location, WLAN positioning, a combination of GPS positioning technology, make full use of the server information storage and computing power, improve the speed and accuracy of positioning.

Proposed Cloud Computing Platform for AR

Cloud based service architecture provides the nearest access to AR services, cloud storage and image parallel computing service, as depicted in Figure 5. Cloud servers uses parallel computing to solve large-scale off-line fast image training and online identification problem; designed to enhance the real image parallel processing framework. The framework adopts distributed architecture, based on map/reduces calculation method, and the parallel implementation of image retrieval, training

samples such as business of augmented reality, to achieve the system with high concurrency, high availability, and scalability characteristics. In the image feature extraction: firstly, the feature points of pre trained fusion of multiple effects on image retrieval factors screening model for feature selection, in the image retrieval using relatively simple linear search method for image retrieval. The method more benefit in distributed processing and memory occupancy is low and support incremental training, to join the new cloud reference image can be quickly added to the index, can be quickly used in cloud object recognition system.



Figure. 5. Cloud Computing Platform of AR Services

Using cloud storage, unstructured and structured storage problem of AR information is settled, and the metadata management of massive AR information and fast index are solved. In augmented reality business application environment, it will store the massive augmented reality information data, and carry on the quick retrieval to these data. The distributed structured storage mechanism to save images and other information, using structured storage mechanism to preserve various augmented reality metadata information; by establishing feature indexing and query mechanism in structured storage mechanism, in order to achieve the purpose of efficient data query.

Summary

Mobile devices, also can be categorized as "see-through" devices in AR, are most commonly used scene capture and visualization devices. Although there is no predominant scene identification algorithm, marker based approach is more common for indoor environments, while markerless approach can be more widely used for outdoor environments.

Taking into account of the limited computing resources of the mobile devices, distributed architecture for HAR is more promising. It transfers the heavy work load of AR to the processing server that facilitates the complex computing technology. Thus the hand held devices are used only for scene capturing and visualization tasks. In the distributed architecture, the real-time requirement of the HAR can be improved.

References

- [1] S. Feiner, B. Macintyre, and D. Seligmann, "Knowledge-based augmented reality", *Communications of the ACE*. vol 36, pp23-62, 1993
- [2] P. Milgram, F. Kishino, "A taxonomy of mixed reality visual displays," *IEICE Transactions on Information Systems*, Vol E77-D, No.12, 1994, pp. 1321-1329

- [3] Lee, "Towards wireless augmented reality - a review of its enabling technologies", ISUVR'09. International Symposium on Ubiquitous Virtual Reality, pp25-28, 2009
- [4] S. Irshad, D. Rohaya, and A. Rambli, "User Experience of mobile augmented reality: a review of studies", 3rd International Conference on User Science and Engineering (i-USEr), pp 125-130, 2014
- [5] M. Fiala. "ARTag, an improved marker system based on artoolkit[J]. National Research Council Canada, 2004, 6: 1-36
- [6] H. Kato, "ARToolKit: library for Vision-Based augmented reality[J]. IEICE, PRMU, 2002: 79-86
- [7] Ozuysal M., Fua P., Lepetit V. "Fast keypoint recognition in ten lines of code". IEEE Conference on Computer Vision and Pattern Recognition, pp.1-8, 2007
- [8] Lowe D.G. Distinctive image features from scale-invariant keypoints[J]. International journal of computer vision, 2004, 60(2): 91-110
- [9] Bay H., Tuytelaars T., Van Gool L. "Surf: Speeded up robust feature", Computer Vision–ECCV 2006. pp.404-427, Springer, 2006
- [10] Azuma R.T., Hoff B.R., Neely H.E., et al. Making augmented reality work outdoors requires hybrid tracking[C]. 1998, 1: 219-224
- [11] H. Lopez, A. Navarro, and J. Relano, "An analysis of augmented reality systems", Fifth International Multi-conference on computing in the global information technology, pp245-250, 2012