Research on Network Architecture and Communication Protocol of Network Virtual Reality Based on Image Rendering

Guohui Yang1, Mengde Xu2

1Department of Electrical Engineering and Automation, Luoyang Institute of Science and Technology, Luoyang, 471023, China
2Department of Electrical Engineering and Automation, Luoyang Institute of Science and Technology, Luoyang, 471023, China

Abstract. With the rapid development of communication technology, mobile Internet technology and intelligent terminal technology, virtual reality (VR) system-related equipment has been highly integrated and internetized, forming a framework composed of terminals, content generation systems, application platforms and network transmission systems. Image-based rendering is a new method to realize virtual reality system. It overcomes the shortcomings of 3D graphics method. The key of virtual reality is the construction of virtual environment scene. In order to ensure the high interaction and real-time of VR system, this paper expounds the technical methods to meet the practical needs of VR wireless network transmission from the aspects of improving the transmission efficiency of VR wireless network, optimizing the network communication protocol and improving the transmission bandwidth of wireless network. The network structure and communication protocol of network virtual reality system are studied based on image rendering. Finally, a prototype of network virtual reality system for remote industrial control based on web is given.

1. Introduction

Virtual reality is a rapidly developing multidisciplinary integrated technology, which provides a new method for human-computer interaction and an effective tool for data visualization, design, testing and control. Network virtual reality is the combination of virtual reality and network technology [1]. It represents the future interaction interface between computers and people. This is a completely immersive interaction interface. Users seem to be really in the computer-generated world, whether they see, hear or feel it, they are like in the real world, and users can also issue commands to computers in a completely natural way [2]. Among various types of virtual reality technology, immersive virtual reality technology, because of its isolation of vision and hearing from the outside world, enables users to experience completely in the virtual world. In order to ensure the high interactivity and real-time of VR system, so that it can provide better visual experience for users, scholars at home and abroad have carried out a lot of research work for the needs of VR wireless network transmission technology [3]. Image-based rendering is a new method to realize virtual reality systems. It uses panorama sets to form a virtual environment. Roaming in the virtual environment is equivalent to selecting different panorama. Panoramas can be generated by computers, captured by panoramic cameras, or captured by ordinary cameras and then spliced. This paper focuses on the network structure and communication protocol of a typical network virtual reality system, and gives a
prototype implementation of a network virtual reality system.

2. Research on Network Architecture and Communication Protocol

2.1 Network Architecture

In virtual reality technology, the first problem to be solved is the modeling of virtual scenes, that is, the construction of the virtual world, and the quality of virtual three-dimensional space modeling is a prerequisite for creating immersion and realism. The scene is too simple, which will make users feel false, while complex and realistic scenes will inevitably increase the difficulty of interaction and affect the real-time performance [4]. In the peer-to-peer VE system model, each peer entity shares the hardware and software resources of other peer entities, and there is no difference between the client and the server. This peer-to-peer model (also called point-to-point or distributed network) enables each peer entity to send packets directly to any other peer entity. The classic VR system is mainly composed of input devices, output devices and computers running virtual reality applications. Now the related devices of VR system have been highly integrated and networked [5]. When the resource node transmits data, the transmitted data is packaged into a network transmission frame format through the resource network interface and output to the local communication node. According to the receiving destination node of the transmitted data. At each point, it takes a picture before and after, left and right respectively. This picture forms a movie segment of this point. Users can look in different directions at each point. When people move around in the environment, the system selects the corresponding viewpoint according to the location of people nearby. Generally, each peer-to-peer entity has a local copy of the database, and can also process calculation and rendering [6]. When a peer entity makes a change to the database, it must notify all other peers in the system of the change.

After analyzing and summarizing various image-based virtual reality systems, it is found that they all go through processes such as panorama mode selection. We have obtained the following image-based virtual reality system model, as shown in Figure 1.

Peer-to-peer model has the advantage of low delay because packets go directly from the sender to the receiver through the shortest path. But it also means that peer-to-peer communication will consume bandwidth. Moreover, when VE uses point-to-point communication over a wide area network (WAN), this model has scalability problems. It directly constructs a three-dimensional scene by using a limited
number of image samples of the three-dimensional virtual space to be built and on the basis of certain image processing algorithms and visual calculation algorithms. These two modeling methods have their own advantages and disadvantages. Among them, geometric modeling based on computer graphics has the advantage of facilitating the interaction between users and virtual objects in virtual scenes [7]. Using circuit switching or packet switching, the source communication node transmits the data to the destination communication node, and the destination communication node transmits the received data to the local resource network interface. Multiple server nodes are independent of each other, and can provide system services independently, so that when one server goes down, it will not affect the state of other servers, so it can provide services continuously, the whole system can operate normally, and ensure the high reliability of services [8]. Once the environment map is generated, as long as the viewpoint remains unchanged, the generation process of the scene in any line-of-sight direction is a projection process from the environment map to a new view plane. The source image it relies on can be either a real digital image taken, a virtual image generated by computer geometric modeling, or a mixture of the two, which requires less basic information and is easier to obtain.

2.2 Network Communication Protocol

The choice of communication protocol for a network virtual reality system depends on the application requirements, the network system being used, and the data to be transmitted. The communication protocols commonly used in network virtual reality systems are TCP and UDP. However, because the virtual objects in the scene are two-dimensional objects in the image. Therefore, it is very difficult and even impossible for users to interact with these two-dimensional objects [9]. BINIR method uses a group of images acquired in advance to encode the environment, and generates new views at different viewpoints through appropriate combination of these images, finally realizing complete roaming of the environment. It allows a virtual environment to send and receive data such as events, messages, files and streams. Virtual Reality Transfer Protocol (VRTP) is the application layer protocol of VR. It attempts to use a unified framework to support the transmission of all types of VR data, using the collection of protocol modules to provide customers with necessary connections. The exclusive bus of the main resource node is composed of clock line and data line, through which the main resource node can send data to each slave resource node at the same time; because it is only based on image analysis and processing, its operation complexity will not be Due to the change of scene complexity, it mainly depends on the working efficiency of the processor and does not need additional hardware acceleration devices. The pixels of the source image are moved to the interpolation position according to the corresponding relation to form the interpolation image. For different pixels moving to the same position, their depth value determines the occlusion relation between them.

One way bus structure is shown in Figure 2. In Figure 2, MR represents master resource; SR represents slave resource; MRNI represents master resource network interface; SRNI represents slave resource network interface.

![Figure 2 One Way Bus Structure](image)

Generally speaking, high real-time interaction and data consistency cannot be achieved at the same time [10]. At present, the consistency of shared virtual environment can be realized by different
methods, such as transaction management, event rotation, primary backup, cyclic token and lock mechanism. As for scalability. The new view can be composed of two adjacent reference images and their corresponding relations. The overall geometric model is not necessary. Image deformation is much faster than image rendering, and the deformation time is independent of the scene complexity. VR content has the characteristics of high resolution and high frame rate, and its massive data has a very high requirement for wireless network transmission bandwidth (real-time transmission rate). In order to avoid vertigo caused by its large delay, the wireless network transmission system needs to effectively control the end-to-end delay within 10ms. It establishes the form and content of individual protocol elements. The core of DIS is PDU. The DIS standard defines 27 different PDUs, of which 4 PDUs (entity state, fire, collision determination conflict resolution) are used to interact with the virtual environment. The bus is a one-way write-only bus, and the master resource node sends data to each slave resource node. When data is sent from the resource node to the primary resource node, requests are sent to the arbitration module through their own independent request signal lines. In this method, panorama drawing does not need real-time control, so it can draw more complex scenes and use more realistic lighting model. Because the source image obtained from any concentric circle can form a cylindrical panorama, and the relationship between these panoramas can be directly obtained through the concentric circle algorithm during sampling. The advantage of the method is that the user's position can be changed at will within the scope of concentric circle.

3. Prototype Implementation of a Network Virtual Reality System

Based on the above research on network virtual reality system, a prototype of network virtual reality system with client/server structure and using MMTP (Manufacturing Messaging Transfer Protocol) protocol to transfer VE data is developed and implemented. Like the view synthesis method, the view interpolation method also needs to establish the corresponding relation from the solid, so it still faces the problems caused by occlusion and depth discontinuity. It poses new challenges to the current network bearer, especially the wireless communication network. After receiving the response signal from the resource network interface, the slave resource sharing bus is occupied, the transmitted data is packaged by the slave resource network interface and transmitted to the master resource node, and after the data is transmitted, the bus possession right is released from the resource node. A variety of Internet devices connect to the server through the Internet to obtain the data provided by the service in the same way, and then present the data to the user in their own unique way in the client. Each server exists independently and does not affect the service provision. The advantage of this method is that it is easy to get panorama of any shape without any additional steps. The disadvantage is that modeling is quite time-consuming and cumbersome. Because the virtual scene generated by concentric circle splicing can only roam the horizontal plane, the field of view is limited. On the basis of the concentric circle mosaic technology, the field of view is extended, and the concentric sphere mosaic is constructed, and the cylindrical panorama is replaced by the spherical panorama. Once the parameters are estimated, the new views can be synthesized, and the three-dimensional corrected mosaic image can be constructed by using multiple views.

In this prototype system, the virtual environment is created by Super scape VRT, and the data communication and interaction between VE are completed by MMTP, in which the coding of MMS protocol is implemented by C++ and integrated into the virtual environment, and its flow is shown in Figure 3. Its decomposition makes full use of the coherence between frames, not only aiming at dynamic geometric objects, but also decomposing different lighting effects (such as shadows, highlights, reflections, etc.) in the illumination model. Each layer can control its refresh frequency, spatial resolution and rendering quality parameters separately. Because the arbitration module also receives the data frame sent from the resource node, the arbitration module can judge whether the data transmission is finished, so as to receive the next sending request from the resource node. When data server increases or decreases dynamically, CoServer can sense the change in real time and always provide data server information available to users. In the figure, path 1 represents the transmission path of MMSPDU request package, path 2 represents the transmission path of MMSPDU response.
package, and MMPM represents the MMS protocol machine. Although in theory a wider modeling space can be obtained by using the full seam method within the range of the lattice trajectory, since the determination of the lattice trajectory is based on a well-defined interaction route, once the interaction route is changed, it is possible that the image information of the original lattice trajectory will not well meet the needs of the new trajectory. When the viewpoint moves to a certain node, the node packet image is transferred into the memory at one time. When the viewpoint is fixed, all operations are performed in the memory, which better meets the delay requirement of virtual reality and enables the system to display in real time.

![Figure3 MMS Protocol](image)

The construction of virtual reality system based on image rendering is mainly to generate a panorama. Panoramas generally have three modes: cubic, cylindrical and spherical. The difficulty of obtaining and controlling different panorama modes are very different. Since the arbitration module also receives the data frame transmitted from the resource node, the arbitration module can judge whether the data transmission is finished, thereby receiving the next transmission request from the resource node. Node selection is actually to determine the shooting scheme, how often to shoot, how each node shoots, etc. Using the technology of separating the control plane and the data plane can not only give full play to the advantages of their respective frequency bands, but also meet the requirements of coverage, speed and traffic, etc., and can reduce the number of base stations and the cost of network distribution. Request and respond between client and server. When the client sends data requests, the server should understand the meaning of each request as clearly as possible. When the server goes down and restarts, the client will not be notified by message. In the process of camera shooting, tilt, tilt and other phenomena will occur. Therefore, one of Stitcher's tasks is to accurately locate the overlapping position between two adjacent photos. The vertical plan generated by the column panorama usually has a turning angle of 50 degrees. Because the cylindrical panorama is a surface image, to get the scene plan at a certain angle, we need to use a perspective deformation algorithm. If multiple resource nodes send request signals at the same time, they respond according to the priority defined in advance.
4. Conclusion
Based on image rendering, panorama, especially cylindrical panorama, is used to construct the scene. Panorama is an image independent of line of sight. Through real-time image processing, users can observe from any direction. Using image-based rendering technology, virtual reality can achieve deeper interaction and immersion than traditional geometry based rendering technology. The visual effect is more realistic and the rendering efficiency is higher. Users can view an object from different angles through a two-dimensional matrix. The application range of image-based rendering method includes virtual roaming, virtual tourism system, virtual real estate observation, building visualization, virtual museum and virtual reality game, especially for space and underwater robot teleoperation system in unknown environment. With the continuous improvement of network technology, computer hardware and software technology, the virtual environment based on image modeling and rendering technology is getting closer to the real environment. This paper expounds the technical methods to meet the real needs of VR wireless network transmission from the aspects of improving VR wireless network transmission efficiency, optimizing network communication protocol and improving wireless network transmission bandwidth. From this, it can be seen that a suitable network architecture and communication network protocol are important to design and implement the network virtual reality system, and there are still some problems to be solved and broken through in this aspect, such as how to better improve the real-time, scalability and heterogeneity of the network virtual environment.

References
[1] Elbamby M S, Perfecto C, Bennis M, et al. Toward Low-Latency and Ultra-Reliable Virtual Reality[J]. IEEE Network, 2018, 32(2):78-84.
[2] Alzubi A A. Bittorrent Protocol for Mobile Health Care System in Hybrid P2P Network Architecture[J]. Journal of Medical Imaging & Health Informatics, 2017, 7(2):508-513.
[3] Arya R, Sharma S C. Energy optimization of energy aware routing protocol and bandwidth assessment for wireless sensor network[J]. International Journal of System Assurance Engineering and Management, 2014:1-8.
[4] O’Connor, Eileen A, Domingo J. A Practical Guide, With Theoretical Underpinnings, for Creating Effective Virtual Reality Learning Environments[J]. Journal of Educational Technology Systems, 2017, 45(3):343-364.
[5] Lacasta Tintorer D, Manresa Domínguez, Josep Maria, Pujol-Rivera E, et al. Keys to success of a community of clinical practice in primary care: a qualitative evaluation of the ECOPIH project[J]. BMC Family Practice, 2018, 19(1):56.
[6] Fan R, Zhang Y, Li B, et al. Multidirectional parabolic prediction-based interpolation-free sub-pixel motion estimation[J]. Signal Processing: Image Communication, 2017, 53:123-134.
[7] Raj K B P, Oza P, Lahiri U. Gaze-sensitive Virtual Reality based Social Communication Platform for Individuals with Autism[J]. IEEE Transactions on Affective Computing, 2017:1-1.
[8] Yang J, Wang Y, Lv Z, et al. Interaction with Three-Dimensional Gesture and Character Input in Virtual Reality: Recognizing Gestures in Different Directions and Improving User Input[J]. IEEE Consumer Electronics Magazine, 2018, 7(2):64-72.
[9] Patti E, Mollame A, Erba D, et al. Combining Building Information Modelling and Ambient Data in Interactive Virtual and Augmented Reality Environments[J]. IT Professional, 2017:1-1.
[10] Mezzavilla M, Polese M, Zanella A, et al. Public Safety Communications above 6 GHz: Challenges and Opportunities[J]. IEEE Access, 2017, PP(99):1-1.