Integrated Somatosensory Interaction System and VR System Based on WLAN

Sheng-Dian JIANG$^{1,2}$ and Zong-Cai RUAN$^{1,2,a,*}$

$^1$Key Laboratory of Child Development and Learning Science, Ministry of Education, China
$^2$School of Biological Sciences & Medical Engineering, Southeast University
Nanjing, China

$^a$rzc.rcls@seu.edu.cn
$^*$Corresponding author

Keywords: Kinect, VR, WLAN, Somatosensory interaction.

Abstract. Special devices are still needed to provide the interaction of the Virtual Reality (VR) system. Many researches use the somatosensory interaction technology to improve the experience of users. However, building data interaction between VR system and somatosensory system remains a prominent problem. In this paper, we use technology of Wireless Local Area Networks (WLAN) to establish a real-time platform of data communication between WorldViz system (a VR system) and somatosensory system which consists of two Kinect data acquisition systems. As the data was collected by two Kinect sensors, we encoded the data and sent it to WorldViz system by constructing WLAN in the whole platform. According to the experiment results, the constructed WLAN played well in data delay and communication transmission. Besides, real-time interaction can be obtained in our research with the similar accuracy of handheld devices.

Introduction

Virtual Reality (VR) has drawn more attention in recent years. The realistic and ingenious digital environment created by VR enables users to interact with the virtual environment with the necessary equipment. As one of the three typical features of VR (Immersion, Interaction, and Imagination), interaction has close relationship to the immersion of VR environment. However, the interaction throughout the scene remains a problem which users have to touch specialized devices to control the movements of the avatar and the transform of main view. This brings users a very bad experience.

To change the terrible interaction experience, many ideas are proposed to give a better solution [1]. Somatosensory interaction technology collects body data from the special sensor like Kinect and then sends the data to VR devices to accomplish the interaction [2]. As a low-cost sensor, the Microsoft Kinect integrates many powerful components. It is composed of a traditional RGB camera, an infra-red (IR) camera, a microphone and a built-in motor [3]. Therefore, we explore integrate somatosensory interaction system (two Kinect sensors) and the WorldViz system [4] together, in wish to strengthen the use’s interaction experience without use of handheld devices.

For the sake of setting up the connection between somatosensory system and WorldViz virtual reality system, an appropriate network topology must be designed to accomplish the data communication. For one thing real-time operation is necessary...
for interaction which low latency must be guaranteed in the network. For another, the devised network should have the ability to handle large-scale data eruption with a high accuracy. Considering the above factors, we select WLAN as the basic network. WLAN technology is a high-speed wireless IP network technology developed in the late twentieth Century [5] and the simplicity and convenience make it popular in many fields. In this paper, we implemented the data communication between VR system and somatosensory interaction system via WLAN and we designed a network topological structure used socket programming based upon UDP which performed well in the communication of Client/Server (C/S).

In our research, we firstly acquired head rotation data, hand state data, hand position data and body position data from the Kinect sensors. Secondly, we built a WLAN which contains somatosensory client system and WorldViz server system. Finally, the data received from the client system was used to control WorldViz system. In the Fig.1, Data processing was followed by Data encoding in the client system and Data parsing was needed after Data reception in the server system.

In this paper, we proposed a Kinect-based interaction technique in VR environment. Considering the ability of the second-generation Kinect, The Kinect sensor can detect and calculate three head rotation angles (pitch, yaw, and roll). Meantime, Kinect sensors can describe human body with 25 skeleton joints in three-dimension space. Somatosensory interaction means that users can interact with the virtual environment like what we do in the real environment. For this purpose, we separate the natural interaction into two steps: navigation and interaction. In the first step, navigation refers to the head rotation to control the virtual environment main view rotation and links the user body movement to the virtual environment main view movement or the movement of avatar in virtual environment. In the other step, interaction represents that hand position data together with the hand state (open or close) data can be used to interact with the virtual object.

In fact, the head rotation data is calculated after the Kinect sensor ‘see’ the user’s face which means three-dimension space detected by a Kinect sensor has a limited angle of view (Fig.2). In view of this, we used two Kinect sensors to enhance the ability of data acquisition. Each Kinect sensor was put on one of the top of WorldViz screen so that the head rotation data would be caught when the user’s face turn around. As is showed in Fig.3, we assume that the user would face the WorldViz screen 1 or the WorldViz screen 2 in all the experiment time.

![Figure 1. Whole platform flowchart.](image-url)
C/S client is mainly responsible for the acquisition, processing, and encoding of the somatosensory data. As we can see from the Fig.4, Kalman filter [6] is used after the raw data received from the Kinect sensor in the client data processing. The measured data affected by many complicated environmental factors (slight shaking of the body, body occlusion) often contain uncertainty noise. Since the general frequency filter is not suitable for filtering the real-time data, Kalman filter is a recursive solution to the discrete-data linear filtering problem and its status as the optimal estimator for one-dimensional linear system with Gaussian error statistics. We chose Kalman filter to smooth the noisy data and provide optimal estimates of the raw data.

The last step in the client system is data encoding. In our research, five types of data (Table 1) were encoded and sent to the server system. Each sending data was encoded to a string type with four parts (Kinect index, data types, data, remarks for the main Kinect). To minimize the network transmission pressure, the data with same type will be compared with the last sending data. Then, sending the data only when the compared results exceed the threshold which is decided by the experience of the interaction of the virtual environment.

<table>
<thead>
<tr>
<th>Data</th>
<th>DataTypes</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>HeadRotation</td>
<td>1</td>
<td>pitch</td>
<td>yaw</td>
<td>roll</td>
</tr>
<tr>
<td>HandState</td>
<td>2</td>
<td>RightHandState</td>
<td>LeftHandState</td>
<td>0</td>
</tr>
<tr>
<td>Bodyposition</td>
<td>3</td>
<td>x</td>
<td>y</td>
<td>z</td>
</tr>
</tbody>
</table>
WLAN Construction

The whole system is mainly composed of client system part and server system part. Building a compatible network between the two parts is significant for the data transmission. Beyond many network topology structures, tree network topology structure is expandable, low-cost and suitable for sub primary or sub level hierarchical management system. In our research, we choose tree network structure as the basic structure to build the whole system structure (Fig.5).

Communication protocol refers to the rules and conventions that must be followed when communication is completed by both systems. Transmission Control protocol (TCP) and User Data Packet Protocol (UDP) are widely used in communication protocol [7]. TCP is a connection-based communication protocol that can request retransmission of the lost data. When two computer systems need reliable data transmission, TCP can establish a stable and reliable connection through the network. While the UDP does not need to establish a connection before the data transmission and UDP does not have retransmission request function. UDP just transfer data packets. Not only can’t guarantee that the receiver receive the packet, but does not guarantee that the receiver would receive the data correct in content or with the right order. Although UDP cannot provide reliable data transmission like TCP, UDP is much more efficient than TCP in terms of transmission efficiency. In our system, reliable data transmission is not quite essential to the performance of the server system and TCP would affect the speed of real-time data transmission. So UDP is used as the network protocol of our WLAN.

Socket programming [8] is highly applied in building UDP-based WLAN. UDP protocol provides communication identification marks for transmission data. The Socket monitors the target port in real time. When the transmitted data arrives, the Socket can transmit the data to the application layer, thus realizing the network communication between the systems. For the implementation of C/S communication program, both on the client and server must first create a Socket object which is used to indicate the endpoint datagram. The client and server application send or receive data at Socket.

A good network topology should meet the requirements of low latency. In view of the problem of clock synchronization in two systems, we choose the client system
clock to measure the data delay time. The whole measurement process consists of two parts: the client sends a frame of data to the server, and then the server sends the data frame back immediately. Suppose that the network consumption time is equal for the two times, we measure twenty times and the result is shown in Fig.6. Contrasting with the wired network delay time, in spite of less stabilization and more delay time, the distinction between the two ways is small to feel for a real-time data communication platform. Moreover, the simplicity and convenience of WLAN makes it a better plan.

![Network Delay Time](image)

**Figure 6. The delay time of the network communication.**

**Server System**

The server system is primarily responsible for data reception, data parsing and sending data to WorldViz system. As is shown in Fig.7, the flow diagram of server system describes that received data will be processed differently depending on the data types (Table 1). Because of two Kinect sensor data acquisition systems, the first problem is how to handle two Kinect sensor data to obtain the better result. According to the principle that Kinect sensor faces the side of the face which is less susceptible to occlusion and other factors, we firstly choose the main Kinect sensor’s data as the data sent to the WorldViz system. Then, updating the main Kinect sensor every frame based on the ‘Remarks’ in the string data. The ‘Remark’ state alter while the smaller angle of roll of the head rotation change.

![Flow Diagram of Server System](image)

**Figure 7. The flow diagram of server system.**

WorldViz virtual reality system includes hardware system and software system. The hardware platform consists of 3D VR helmet, infrared optical tracking and 3D fusion active stereo projector, projection screen and so on. The software platform can complete the modeling and control of virtual avatar in virtual environment. In this
paper, we focus on collect the somatosensory data to complete the following functions:

1. **View Rotation Navigation**: using the rotation angle of the head to control the change of the view angle of the virtual scene.

2. **View Position Navigation**: using the body position data which will convert the coordinate system before to control the virtual scene perspective to move forward or backward.

3. **Hand Control**: using hand position information combined with hand status to complete the virtual hand movement in the scene. When the hand state is open, only the movement of the virtual hand can be controlled. While the hand is in the close state, the virtual hand can complete the grasping of the same position object.

**Conclusions**

In this paper, we researched data interaction between somatosensory interaction system and WorldViz VR system in wish to control the virtual environment without handheld devices. Kinect sensors were used to obtain the information of the users and WLAN was constructed to build the connection of systems. In this work, we substitute the expensive handheld devices with taking advantage of low-cost devices which provide a better immersion in VR environment. Besides, this work will provide a new thought for the improvement of nowadays VR application experience. At the meantime, the whole structure would ameliorate the effect of some experiments in education and psychology. However, the current platform still has some problems. Large-scale data processing ability in the network is poor and data packet loss problem caused by unreliable communication still will make unsmooth control of virtual environment. The future work would focus on improving operational smoothness experience and the capability of network data transmission.

**Acknowledgement**

This work was supported by the Fundamental Research Funds for the Central Universities (2242016K40040, 2242017K40200).

**References**


