Research on the Multiplayer Online Virtual Experiment System Based on Kinect Somatosensory Interaction

Wenlan Guo[1, a], Jialiang Zhang[2], Yafei Cai[3] and Bowen Sun[4]

[1, 2, 3, 4]Department of Computer Science and Technology, Harbin University of Science and Technology, Harbin, 150080, China

[a]email: guowenlan@sohu.com

Keywords: multiplayer; virtual experiment; Kinect; somatosensory interaction; Unity 3D

Abstract. According to the needs for many people conducting a certain virtual experiment at the same time in different places, a multiplayer online virtual experiment system built by Kinect and Unity 3D engine is proposed. In this system, the Kinect is implemented to complete the motion operation, which can improve users' immersion. To achieve the function of multiplayer online, the Remote Process Call technique is introduced. The results show that the system is steady, and it can and meet the need of multiplayer online. The system can be applied in experiment education of all types of schools.

Introduction

Virtual experiment system is supported by virtual reality technology and other technology, which can accomplish experiment data collecting and analysis by using a computer. Virtual experiment allows users to achieve interaction with virtual world by taking advantage of visual sense and other sense. So it is important that the experiment is based on virtual reality technique in experiment education. However, some inventions of somatosensory equipment made the traditional interaction of processing with mouse and keyboard of in trouble. The virtual experiment technology could meet the requirements for no restraints, but great expense and special purpose limited their popularity among regular schools. Kinect from Microsoft was cheaper than the equipment mentioned, and it can collect depth and color images, recognize people gesture and track body position. So Kinect is a good choice to develop virtual experiment.

At present, many applications based on Kinect somatosensory interaction have been developed. For example, Kinect-based Ergonomic Assessment System [1] designed by H.Haggag can aid in ergonomic analysis. Virtual assembly technology based on Kinect [2] investigated by Y.Y. Chen allow human simulating before the real operations. But they did not focus on the Virtual Experiment System based on Kinect. Virtual Assembly System based on Kinect [3] and Virtual Experiment System for Electrician Training built by Hongjian Liao and Zhe Qu make use of the Kinect in teaching activities, but they did not solve the problem of multiplayer online.

The ability of team work is important in experiment operation. Therefore, the function of multiplayer online is significant. Remote Call Process technique is employed to achieve the multiplayer control. This paper studies the implementation of Kinect in the multiplayer online virtual experiment system in which Kinect is used to collect body information and a LAN was built to realize multiplayer cooperation. This solution is of certain innovation and generalization.

Kinect Pose Recognition Technology

Kinect can capture depth images and track user's skeleton. It has a set of infrared cameras. The infrared ray launched by launching camera can be caught by receiving camera after hitting something. Kinect calculates the difference between launching and receiving time, and generates an image of the very time where each pixel represents a value of short type. The last three bits of the value index the user, while the others give the distance between user's body and Kinect, which is shown in Fig 1 (a). Then Kinect can separate human from background.

Kinect can recognize human body parts in a depth image based on random decision tree, while the tree was trained by many series of body images with body parts marked in advance [4]. The training result was that the feature value of pixel x belongs to body part c in image I was computing by Eq. 1. In Eq. 1 $d_I(x)$ is the depth value at pixel x in image I, u and v are offset parameters.

$$f_\theta(I,x) = d_I\left(x + \frac{u}{d_I(x)}\right) - d_I\left(x + \frac{v}{d_I(x)}\right) \tag{1}$$

In actual use, Kinect generates a body image I_0 with the infrared cameras, computes the feature value of each pixel in the image, and compares the feature f_θ with threshold τ repeatedly to reach the probability distribution of the pixel x belonging to body part c in I_0, that is $P_t(c \mid I, x)$ [5]. And final result is the average probability distribution of the random decision forest as Eq. 2:

$$P(c \mid I, x) = \frac{1}{T}\sum_{t=1}^{T} P_t(c \mid I, x) \tag{2}$$

So Kinect can determines how many pixels belongs to body part c and other body parts, then marks the body parts in the body image generated by itself which is shown in Fig 1 (b).

(a) The meaning of pixel value (b) Body parts in depth images (c) Depth image and skeleton

Fig.1. Pos recognition technology

Per-pixel information is inferred based on the body part recognition described above. Reliable proposals for the positions of 3D skeletal joints must be generated by pooling across the per-pixel information. In case of the outlying pixels degrade the quality of global estimate severely. Kinect employs a local mode-finding approach based on mean shift with a weighted Gaussian kernel [6].

$$f_c\left(\hat{x}\right) \propto \sum_{i=1}^{N} \omega_{ic} \exp\left(-\left\| \frac{\hat{x} - \hat{x}_i}{b_c} \right\|^2\right) \tag{3}$$

Eq.3 is a density estimator. In it $\hat{x}$ is world space coordinate, N is the image pixel number, ω_{ic} is a pixel weighting, $\hat{x}_i$ is the secondary projection of image pixel x_i into world space according to the depth $d_I(x_i)$, and b_c is a learned per-part bandwidth. The pixel weighting ω_{ic} considers both the inferred body part probability at the pixel and the surface area of the pixel. With the density per body part computed in Eq.3, Kinect can roughly locate the positions of every joint in the depth images.

Above works contribute to the Kinect pose recognition technology. Kinect separates the body images from the background, generates a user's skeleton with joints positions, as is shown in Fig 1 (c).

The Kinect Software Develop Kit provides all the information above to developers for gesture recognition. Using the joints data, we can define gestures to operate in the virtual experiment. And it also provides the results of fist detection. We can realize the grabbing effect by programming.

Remote Process Call Technique and Multiplayer Online Function

RPC is a kind of protocol to request service from remote computer program, and the developers do not need to understand the underlying network technology. It can help deliver information data between communication programs. The mode RPC implemented is Client/Server mode. The program requester is the client while the program provider is the server.

Unity 3D is a multi-platform and fully-integrated dedicated game engine. Developers can easily edit and render virtual scene, manage game logic function based on C# computer languages. There is a useful component named Network View in Unity 3D which can realize the RPC function. To achieve the goal of building a local area network (LAN), particular references need to be added at the top of the scripts. The terminal type can be checked by using the class of *NetworkPeerType*. With a LAN established, system creates a server with the programming statements below:

MasterServer.RegisterHost(HUSTVRlabServer, name);

After the connecting between server and client is established, RPC protocol will find and call the functions marked by [RPC] across the LAN. And any computer can send RPC call to others including its own to achieve communication among all the players of the experiment system.

Virtual Experiment System Frame Structure

The system was consisted of three modules: experiment operating module, gesture collecting and processing module and network communication module.

Experiment operation module includes several parts: virtual scenes, user interface, experiment and operation logic control and grading mechanism. Experiment logic control was programming according to real experiment, and it was added hints in the certain steps which may be dangerous in real experiment. Operation logic control takes charge in processing human gesture data and getting rid of useless information to ensure the running efficiency. While grading mechanism calculates the accuracy of user's operation by adding test variables at key steps.

Gesture collecting and processing module processes the human gesture data. By analyzing real experiment, we defined several poses, such as selecting, grabbing, releasing and walking around.

Network communication module has two main functions: One is for sharing operation data, which is transferring operating information to other terminals while the user is operating. The other is sharing his operation intention, which is transferring user's hands positions when he does not operate.

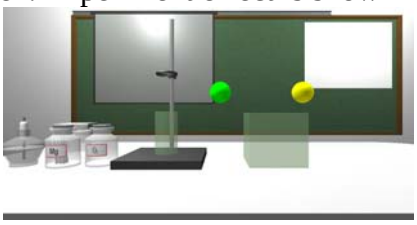
Case and Application

Some chemistry experiment projects may be dangerous without fully understanding. And chemistry drug usually cost too much. So virtual chemistry experiment would benefit both in safety and cost control. So this system focused on chemistry, it allowed user conducting experiment repeatedly to avoid human mistakes in real experiment. This paper introduces the solution of multiplayer virtual experiment for chemistry, taking the "Burning magnesium sticks in oxygen and air" as an example.

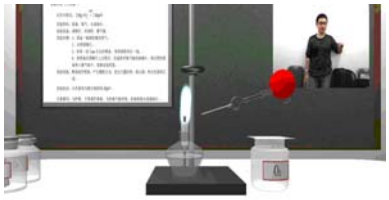
When the system runs, user can choose the multiplayer or single player mode according to his demand. If he chooses the former mode, he will come to a menu to create or join a server. If one user is unaware of the existence of server and clicks the creating server button, there will be a hint says that "Server exists, please join it". Then he should click the joining button to join in the LAN with system searching and finding the server based on the particular key code such as " HUSTVRlabServer ".

After succeed in connecting to LAN, the system can distinguish user's hands model from his companies' by the *NetworkView ID* which is distributed by the Unity 3D Network component, and the colors of models are different, as is shown in Fig 2 (a). When a user waves his hand, the corresponding model will move in the same way according to the data provided by Kinect SDK. While one grabs any object and moves it in the virtual scene, the control script will call the RPC functions within the LAN. So anyone's operation will be shown in the screens of connected terminals.

In the experiment of "Burning the magnesium sticks both in oxygen and air ", user should grab a stick of magnesium, approach it to the flame of alcohol lamp to ignite it and stretch it into a bottle of oxygen. Experiment effect is shown in Fig 2 (b).



(a) Different colors for different user



(b) Grabbing a stick and igniting it

Fig.2. Case and Application

Test Results

There are several gestures needed in the virtual experiment including: object-grabbing, object-moving, object-releasing and character-walking. The test results are shown in Table 1.

Table 1. Test for gestures.

Gestures	Test times	Correct times	Precision
Grab & move	1000	982	98.2%
Release	1000	993	99.3%
Walking	500	494	98.8%

It was by fixed key code that the server and client found and connected to each other. We got no situation like "connection failed" in a stable state of LAN. All the test results showed that the multiplayer online virtual experiment system based on Kinect was in good stabilization and accuracy.

Conclusions

This paper discusses the implementation of Kinect somatosensory technology in virtual experiment, and innovates the interaction of virtual experiment, and greatly improves user's immersion.

The introduction of RPC technique changes the experiment mode from players taking turns to multiplayer cooperating at the same time. It also increases teaching efficiency and enhances teaching effect. And it fills the gaps that no virtual experiment allows multi-user operation up till now.

The proposal is a good reference and guidance for the establishment of virtual experiment base on somatosensory technology, and is of great popularization also.

References

[1] H. Haggag, M. Hossny, S. Nahavandi, D. Creighton, “Real Time Ergonomic Assessment for Assembly Operations Using Kinect,” Proc. Computer Modelling and Simulation (UKSim), 495-500 (2013).

[2] Y.Y. Chen, ”Gesture recognition based on Kinect and application in the virtual assembly technology,” Electronic Design Engineering, 4-7 (2013).

[3] Hongjian Liao， XiaoLi Long. Study on Virtual Assembly System Based on Kinect Somatosensory Interaction[C]. International Conference on Information Science and Cloud Computing，2013:55-60.

[4] Jamie Shotton，Andrew Fitzgibbon. Real-Time Human Pose Recognition in Parts from Single Depth Images[DB/OL]. Microsoft Research Cambridge & Xbox Incubation, June 2011.

[5] V. Lepetit, P. Lagger, and P. Fua. Randomized Trees for Real-time Key point Recognition[C]. San Diego: CVPR. 2005: 775–781.

[6] D. Comaniciu and P. Meer. Mean shift: A robust approach toward feature space analysis[C]. IEEE Trans. PAMI. 2002: 603-619.