International Journal of Civil Engineering and Machinery Manufacture
ISSN: 2518-0215, Volume 4, Issue 4, December, 2019

Research on Spatial Location and Interaction Method of Object in VR Application

Boxiong Yang[1, 2], Heng Xiao[1,2*], Ping Yin[1,2], Chun Xiong[1,2], Bo Zhou[1,2]
[1]School of Information & Intelligence Engineering, University of Sanya, Sanya, 572022, China
[2]Academician Workstation of Guoliang Chen, University of Sanya, Sanya, 572022, China

Abstract: With the development and progress of modern sensor technology and embedded processing technology as well as various location method and algorithm, human computer interaction capability in virtual reality technology is greatly enhanced. The experience feelings of users will be more real in the virtual three-dimensional environment. The various location methods of virtual reality in recent years have been researched and discussed in the paper. The relative sensors and interaction devices have been studied and compared such as PS, Eye+PS Move HTC, Vive Light House Kinect, etc. Meanwhile, the advantages and disadvantages of those methods and corresponding equipment have been analyzed and compared. The respective location accuracy, scope of application and applicable scenarios have been deeply analyzed and discussed, which provide reference for the application of space location and human computer interaction in virtual reality.

Keywords: Virtual reality; Spatial location; Interaction method; User experience

1. Introduction

In all kinds of interactive applications of VR(Virtual Reality), how to determine the positions in space of the targets within the range of activity space is the key factor which effects the user experience[1]. Some virtual reality applications with the interactive function do not have an ideal using effect and vivid experience, such as body feeling games, roaming, simulative training and reproducible scenes[2]. The applications of virtual reality technology are able to accurately locate and quickly obtain body movements, gestures and fine behavior of fingers of users, and it greatly enhances the immersion and scene of reality of users[3]. In this paper, we research different interactive devices and methods of target location in current virtual reality technology applications. We also compare the advantages with the disadvantages of these devices and methods. The research results could provide referable value for space location and man-computer interaction of virtual reality technology applications.

2. Infrared Optical Location

The location principle of infrared optical method mainly uses multiple infrared transceivers to cover the range of activity space of users in interactive virtual reality. We place infrared light spots upon game players or experiencers, and we make sure of position information in space of them by capturing the signal of infrared receiver which is reflected back by reflective spot[4]. In order to reduce the blocked line-of-sight and achieve all-round tracing, location method of infrared optical requires at least three high-definition cameras, generally it needs more than six, as shown in Fig.1.

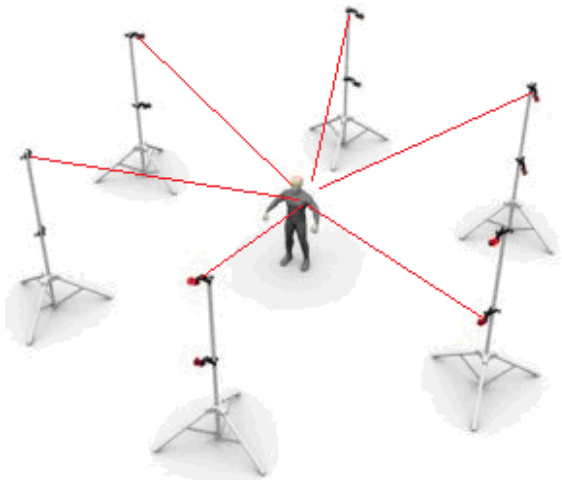


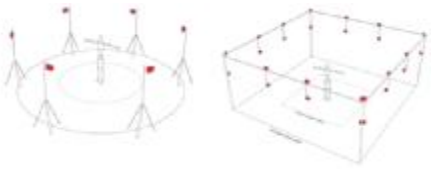
Figure 1. Principle diagram of infrared optical location method

The most representative product of spatial location method of infrared optical is the OptiTrack motor tracking system. OptiTrack captures the human body, HDM(Helmet Mounted Display), fixed and tailor-made marked points on interactive devices at the shooting rate of hundreds of frames per second. It could build up the information of body's bone or three-dimensional position

of marked points real-time and precisely. OptiTrack provide a set of location data of and body action data for users of virtual reality application field.

OptiTrack motor system capturing system is able to capture all-body movements of 2 to 4 people in space of 3m x 3m to 25m×20m with the data delay less than 5 ms, and it could achieve the multiple people and things of space location at the accuracy as high as 0.1mm[5]. At the same time it connects seamlessly with main engine such as Unity/UE4, and it supports customized developments of real-time data stream interface of six-degree freedom.

According to the traced effect and captured range, there are two ways to installing when using OptiTrack. The two ways are respectively tripod installation and truss installation, as shown in Fig.2.(a) and Fig.2.(b).



(a) Tripod installation (b) Truss installation

Figure 2. Installation method of OptiTrack

3. Location Method of Binocular Imaging

Location method of infrared optical needs to put several infrared reflector on users, and it will increase unnecessary inconvenience. Principle of binocular distance measurement in use of two cameras to take pictures of users in a non-contact way. It captures the user's spatial position and posture through the binocular distance measurement algorithm. The method is able to make the users get rid of the bondage and discomfort of the sensors and make it more convenient for users.

location method of binocular imaging is a kind of typical space location method which is based on image. It collects color images or infrared images with the help of at least two cameras. The camera feeds back into two infrared cameras through infrared fractional laser dot matrix and it makes 3D scans for x, y and z coordinates of users. When the scan is over, it transmits data to a computer[6]. Then the computer will calculate the transmitted data, and locate the space through binocular stereo space location algorithm, the basic principle of binocular stereo location is shown in Fig.3. The ligature of the center points of the two camera lens is called baseline of the system. We ensure the coordinate point (x_l, y_l) and (x_r, y_r) on image plane and the coordinate point (X, Y, Z) of point P on world coordinate system in advance. When the camera coordinate system and world coordinate system coincides, the x, y planes of image plane coordinate system are paralleled with the world coordinate system [7].

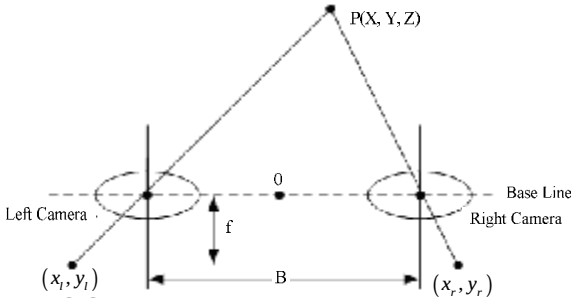


Figure 3. Sketch map of binocular imaging

When the two cameras are the same, and the correspondent coordinate axises are paralleled, but the original points are different, as shown in Fig.4. We will make the coordinate system of the first camera coordinate coincide with world coordinate system, and then calculate x coordinate of point P.

$$X_1 = x_1(f - Z_1)/f \quad (1)$$

In the same way, we enable the second coordinate system of the camera to coincide with the world coordinate system, and get:

$$X_2 = x_2(f - Z_2)/f \quad (2)$$

The length of the baseline is B, and the z coordinate of point P gets the same value of the coordinate system of the two cameras.

$$X_2 = X_1 + B \quad (3)$$

$$Z_1 = Z_2 = Z \quad (4)$$

Plug Eq. (3) and Eq.(4) into Eq.(1) and Eq.(3), and we can get:

$$X_1 = x_1(f - Z)/f \quad (5)$$

$$X_1 + B = x_2(f - Z)/f \quad (6)$$

Solve the equation and get:

$$Z = f - fB/(x_2 - x_1) \quad (7)$$

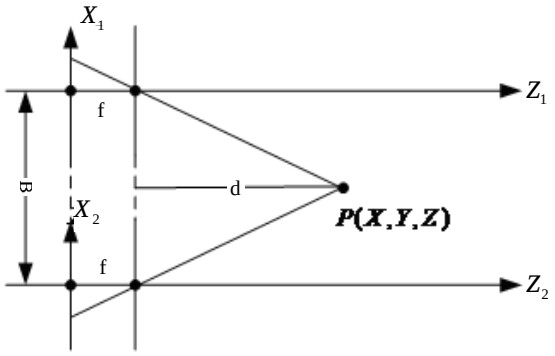


Figure 4. Parallax in parallel binocular imaging

Eq.(7) connects the distance of the object and image plane with the parallax ($|x_2 - x_1|$). The size of the parallax is associated with depth, so the parallax includes the spatial information of the 3D object. When the parallax and the distance of baseline and the focal length is confirmed, we can calculate the z coordinate of point P and get the depth information of the point[8].

Actually, the most famous equipment which uses the method to locate the space of the users who play the virtual reality games is Kinect body equipment. There are three cameras in Kinect. The center camera of Kinect is the RGB color camera, and the left and right sides are respectively the infrared transmitter and infrared CMOS camera, as shown in Fig.5. Kinect is equipped with base motor to achieve chase focal technology, it will realize movement of the object and capture the posture when it turns with the movement of the target.

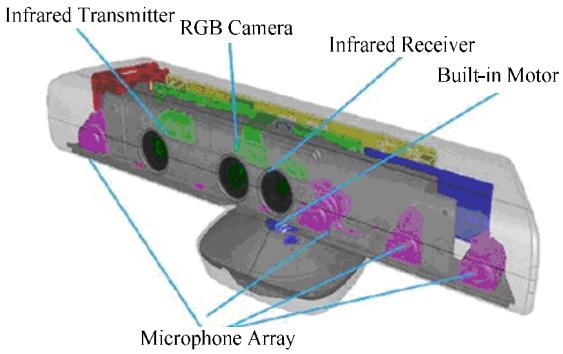


Figure 5. Main components of Kinect body equipment

Kinect uses a technique which is called optical encoding. The technology uses simple continuous lighting instead of the pulse signal. Kinect do not need special light-sensitive chips, so ordinary CMOS light-sensitive chips could solve the image calculating , it will reduce the cost greatly[9]. Kinect projects some random dot matrix by IR (Infrared Radiation), and only uses an ordinary CMOS sensor to capture this dot matrix. When the depth of the scene changes, the dot matrix which is captured by camera will also change, and we can infer depth information through the change. At the same time, Kinect can also make pixel level assessment of depth of field images. Segmentation strategy is used to distinguish human body from the background, and to get the depth of field images which the background objects are removed off. In addition, in the whole working process of the image capturing, Kinect could converse the images of 3D range to skeleton tracking system and capture overall motion of users. Then it will quickly generate a complete movement model and output it according to the movement of skeleton, as shown in Fig.6.

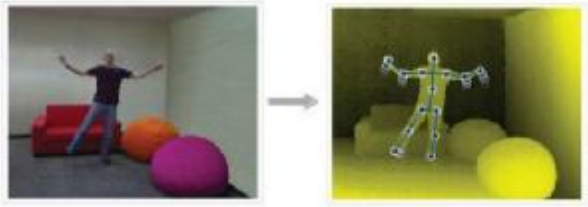


Figure 6. Skeleton tracking of Kinect image

Kinect captures the gestures of users with the help of Prime Sense software and cameras detection, and then compare the captured images with human body models preserved in system. Every human body model which accords with existing internal object will be created into related bone model. System will transform the model into a virtual character, the character will be triggered by identifying the key parts of the human body skeleton model. The location accuracy of Kinect is centimeter level. The best perspective is 57°, and the best working distance is 1.2m to 3.5m. In this area, we can detect location information of 6 individual. And it contains movement track of 2 people, and detailed record of 20 groups, including the tracing range of trunk, limbs and fingers. So it will complete a full body feeling[10].

Kinect body feeling equipment is generally used for optional peripherals for the XBox One game machine, the current refreshing rate is about 30fps. Although the delay makes less effect on people compared with delay of helmet, But the users still feel the location result of delay obviously. The inadequacy makes people lack of more fluent user experience.

4. Location Method of Spatial Triangulation

Triangulation method is used to realize location and direction of three-dimensional space, and it is the common method of location in geodesy and geology. We can also use this method to obtain the user's information of location and movements in virtual reality technology. It is used to use two or more detectors to detect the direction of objects in different positions in activity space of the target objects. Then we make use of triangle geometry principle to ensure the location distance of the objects. As shown in Fig.7, we design the reference coordinate system according to the plan in advance. We install three receivers Ro, Rx and Ry respectively on point of origin, the X axis and the Y axis which are set before. At the same time, there are three emitters Fa, Fb and Fc on device T(holder or helmet) which are worn by objects. We use some methods(such as optical method, pulse method, etc) to calculate the distance of three transmitters or reflectors and between Ro, Rx and Ry. So we can calculate space coordinate information of the objects within the range of space which could be detected by space triangulation method.

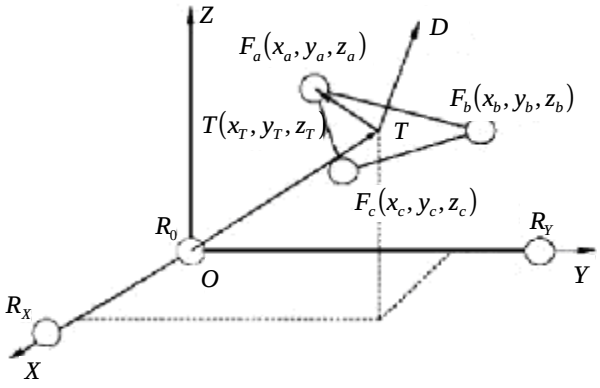


Figure 7. Measuring method of spatial triangulation

The typical representatives which make use of space triangulation method to locate the virtual reality applications are PS Eye + PS Move, VirTouch, etc. PS Eye uses wide angle camera with high resolution. It can not only take photos and videos of a variety of specifications, but also accurately measure depth of field and dolly moves of the targets in camera. PS Move includes a moving handle to tag the ball, a three-axis gyroscope, a three-axis acceleration and a magnetic sensor. It can provide six shaft sensor, vibration, and operation function of body feeling. There are different colors lighting balls equipping with handles of PS Move. The lighting balls show images significantly different from background in camera, so it is convenient extracted it by computer vision algorithm. The three-dimensional space location of PS Move depends on the realization of the top lighting ball. More far the lighting ball is from the camera, the area it projects into the camera is more smaller. We can calculate the changes of distance accurately according to the size of images, and ensure the accurate coordinates of it in three-dimensional space[11].

There are two problems need to be solved when using the method of camera and lighting ball to locate. One is how to accurately and steadily judge and distinguish different marked points from the frame of camera. Another one is how to know the position and posture in world space about camera itself, in order to calculate the position and posture of the experiencers correctly. The general treatment is to use a professional camera whose frame rate is greater than 100 hz and the shutter system is global, so phenomenon of the motion blur can be avoided when the high speed moving objects appear on the screen. Another is using infrared LED to fill all around camera with light, and adopts the material of high reflectivity to make marked points which users wear. Because infrared camera itself has blocked most of the information of visible lighting, so marked points could be greatly apparent in the picture[12].

VirTouch transmits infrared light by infrared LED and reflect it by a finger ring (glass inlays outside the ring, infrared thermal reflective coating inside the ring). It uses a visible light filter camera to filter the visible light, and it captures inputted signal generated by infrared controller through the infrared light. There are two functions of the ring: (1) reflect the infrared light; (2) accurately identify the input signal of single point when its center point is the identification point. What is the most important is that the solution could form a three-dimensio al input and identify the degree of depth of field all around. Because this method has high accuracy and stability, and it could achieve low delay by adjusting camera parameter. In theory, this method can be extended to infinite space, so triangulation space location method has become the first choice of many museum builders of virtual reality currently. Of course, there is still significant limitations identifying a number of experiencer. Because there is no possible that the marked points combine indefinitely, if the two groups of marked points get too close, it is easy to be false-positive or unable to identify when two experiencers are back to back. In addition, the user of complex scenes are more likely to be sheltered by obstacles and side leakage happens. The virtual reality technology which adopts the interactive method is using in an open or regular shapes(such as square or rectangular) room.

5. Location Method of Laser Scanning

The location method of laser scanning works through two laser sensors worn by users or location tracing sensors handled by users. It gets information of position and direction in the two ways. The location method of laser scanning technology has advantages of low cost, high accuracy and distributed processing. The technology almost has no delay and is not afraid of keeping out, it could capture the objects even if the handle of it is on the back or under leg. The location method of laser scanning technology has avoided the disadvantages of high complexity, high cost of equipment and slow speed of calculating and easy effect of natural light. It has realized high precision, high response speed of indoor location.

The typical product of location method of laser scanning technology is HTC Vive. It uses a different location method from optics. It is full of helmets, handles of infrared sensors and light house used for location itself. The users put the light houses in the two corners of an empty room, and the two light houses is equivalent to two fixed laser launch base stations, as shown in Fig.8.

The advantage of HTC Vive include:

1. Synchronization: the overall plate of led lights a time, the sensors of handle and helmet lit up together as synchronous signal.

2. X axis scanning: transverse laser lights up the photosensitive sensors of handles and helmets.
3. Y axis scanning: vertical laser lights up the photosensitive sensors of handles and helmets.

Figure 8. Sketch map of HTC Vive light house

There are many photosensitive sensors installing on the displayers and comptrollers of helmets. When LED of station flashes, it will synchronize the time of all equipment, then it starts to scan. We can measure the time of laser of X axis and Y axis arriving at the sensors. When the laser sweeps across the sensor has the sequence, the time several sensors which perceive the signal also has sequence. So the angel of X axis and Y axis of each sensor which is relative to the base station is known to us. Then position of sensors which installed on handle and helmet has calibrated in advance and the position is fixed. So we can calculate the position shift of helmets and handles and motion locus according to position shift. There are more than 70 photosensitive sensors on helmets and handles on HTC Vive. When laser scans across, the helmet begins to count. We can calculate accurate location relative to the laser emitter by use of the relationship of the position and time of receipting laser. If the photosensitive sensors are enough at the same time, it will generate a 3D model. The model can not only detect the position of helmet, but also capture the direction of helmet. The calculating process generally follows the steps as shown in Fig.9.

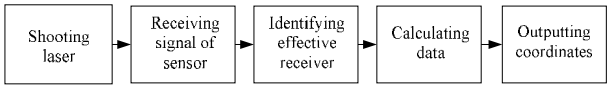


Figure 9. Calculation process of HTC Vive

The biggest advantage of HTC Vive light house is that it requires a very small amount of calculation, because it do not have to image first and then identify the features from the images like the location system based on image. HTC Vive uses only simply photosensitive, and need not image. So it will not involve a large number of computation, image processing and avoid possibility of performance loss and instability. Comparing with other interactive devices, delay of HTC Vive is less, and it could response faster, so it is better for user experience.
Although HTC Vive light house is the virtual reality interactive location devices with the best experience so far, but apparently there is a problem. The limitation of the power and the safety of irradiating human body, it covers a very limited distance and range. Generally, a space region of 5m×5m. In addition, installation of this kind of equipment is little cumbersome, and it may be very difficult for the first operation of ordinary users.

6. Experiment and Comparison

This article selects several typical applications of interactive virtual technology and compare the performance of them. We compare performance of different aspects of location precision, delay time and recognition scope. The performance index as shown in Table 1.

Table 1. Comparison of all kinds of location methods

Typical representative	Location Principle	Location Accuracy	Delay time	Identification Scope
Soptitrack	Infrared Optics	Submilli-Meter	10ms	9-900 m2
Kinect	Binocular Imaging	Centimeter	30ms	4╳3m2
Ps eye/ Ps move	Triangulation	Centimeter	30ms	30cm Expansible
Htc vive	Laser scanning	Millimeter	20ms	5╳5 m2

We could improve immersion of virtual reality by extending field of view and more accurate method of location and tracing, and it will the key development of virtual reality technology in the future. Each of the above methods has its advantages and is suitable for different application, and develops and improves constantly. Of course, now there are some other location technology currently, such as wifi technology, RFID radio frequency identification technology, UWB technology, low-power Bluetooth BLE4.0 technology, ZigBee technology, etc. However, because of the limited precision of location, they are rarely applied in the field of virtual reality technology.

7. Acknowledgements

The paper is supported by Industry University Research Innovation Foundation of Ministry of Education of the People's Republic of China (No. 2018A01003).

References

[1] Zhao Qinping. A survey on virtual reality. Science in China(Series F:Information Sciences). 2009, (1), 2-46.

[2] FP Brooks. What's real about virtual reality. Computer Graphics & Applications IEEE, 2010, 19(6), 16-27.

[3] Xiao Bao, Yurong Mao, Qiang Lin. Mechanism of Kinect-based virtual reality training for motor functional recovery of upper limbs after subacute stroke. Neural Regeneration Research. 2013, 13, 127-138.

[4] Multi angle optical location and tracking system. Integration Technology. 2013, 1(2), 34-40.

[5] Ma Le, Song Rong, Liu Yan. Design and experimental verification of rehabilitation training system based on Virtual Reality Technology. Computer Application and Software.

[6] Yang J.C., Chen C.H., Ming C.J. Integrating video-capture virtual reality technology into a physically interactive learning environment for English learning. Computers & Education. 2010, 55(3), 1346-1356.

[7] Yang Boxiong, Hu Xuhua, Huang Jing. Fingertip location through the Gesture Space Points Sets Constructed by Kinect Hand Tracking. Journal of Beijing Normal University (Natural Science Edition). 2015, 51(6), 571-575.

[8] Zhao Hanli, Meng Jiaru, Han Lizhen. Design and Implementation of Customized Virtual Body Building System. Chinese Journal of image and graphics. 2015, 20(7), 953-962.

[9] Kong Liwen, Xue Zhaojun. Research progress of brain computer interface technology based on virtual reality environment. Journal of electronic measurement and instrument. 2015, (3), 317-327.

[10] Lin Yi, Chen Jing, Liu Yue. Virtual Reality Mental Model and Augmented Reality Hybrid Mobile Navigation System Design based on User Experience. Journal of Computer Science. 2015, 38(2), 408-422.

[11] Chang Yang, Lv Yuanyuan, Han Yingjiang. Real estate display system based on Virtual Reality Technology. Electronic Technology and Software Engineering. 2014, (7), 80-81.

[12] Chen Wei. HTC virtual reality helmet is about to start scheduled. Computer and Network. 2016, 1, 34.