The Design and Construction of College English Multi-modal Assisted Teaching System Based on Virtual Classroom

Haiqi Liu(✉)

Dalian Naval Academy, Dalian 116018, China
carollikes@126.com

Abstract. In view of the problems of single teaching form, weak students' sense of classroom participation and insufficient educational resources in colleges and universities, the author of this paper uses the characteristics of VR technology with strong interactivity and multi-sensory experience, and combines the multi-modal teaching concept to develop a multi-modal auxiliary teaching system for college English based on virtual classroom. The system is modeled by 3DSMAX, and the function is designed by using unity3D in C# language to complete the system development. It gives full play to the characteristics of virtual reality technology, and designs virtual scenes from the perspectives of classroom, roles and scenes, thus constructing an English classroom teaching system with virtual perception. The virtual space can be created by combining VR hardware with students and teachers, and the human-computer interaction experience can be realized. During the development process, the Control Rotation. cs class is used to realize the function of rotating the viewing angle around the object by 360 degrees, and the Input Tracking class technology is used to obtain the user's eye position. The development of the system has promoted the current educational reform in China to a certain extent and contributed to the educational cause in China.

Keywords: VR technology · English teaching · Multi-modal teaching · system design

1 Introduction

Education is the foundation of our country, the future of our country depends on the growth of education, and education is also one of the important standards to test the comprehensive national strength. However, English is an indispensable and important subject in education. Because of its cross-cultural characteristics, English education can help our country cultivate multi-disciplinary international talents with communicative and speculative abilities. In the theory of foreign language education, there is a famous situational learning theory, which holds that students should integrate the interaction between people and the environment in the process of learning a foreign language, so that students can practically apply their foreign language learning knowledge to real life practice. But, under the background of the current era, the mainstream teaching mode

V. Balakrishnan et al. (Eds.): ICMEIM 2022, AHSSEH 6, pp. 377–386, 2022.
https://doi.org/10.2991/978-94-6463-044-2_49

is still the traditional classroom teaching, which uses courseware and books as a single interactive teaching method. This kind of teaching method lacks practical factors, the teaching environment is not intuitive, and students' interest in autonomous learning is not high because it is too boring and rigid, which directly affects the quality of English teaching. Moreover, among the English knowledge that students have learned, the knowledge of cultivating reading and writing ability is relatively large, while the knowledge of listening and speaking is relatively small, so that students can't fully apply English to their lives, which leads to the fact that the current teaching objectives can't meet the requirements of English quality education. To solve the above problems, China urgently needs to optimize the teaching mode and reform the English teaching system. [1] We found that multimodal teaching theory is suitable for optimizing teaching mode. Multimodal teaching can enable students to learn the language in person, and in this process, it can achieve the best effect of strengthening the understanding of English context and memorizing English vocabulary and sentence patterns. The focus of multimodal teaching is to comprehensively cultivate students' four basic foreign language learning skills: listening, speaking, reading and writing. The Vr technology is virtual reality technology, which uses the computer to build three-dimensional objects and combine them. It can break the limitation of time and space to build a virtual world, which can also provide users with sensory experience about vision, hearing and touch. VR technology combined with VR glasses, handles and other hardware devices creates a real-time interactive way to achieve accurate simulation. VR technology is just in line with the current development needs of English multimodal teaching display and interaction. [2] According to the above analysis, the author believes that a multi-modal English teaching system based on VR technology should be developed. This paper uses B/S architecture, 3DSMAX modeling, C# language and unity3D to design functions, and combines sqlsever database technology to complete the system development. This paper systematically explores the combination of vr technology and English learning, which is a new attempt of English autonomous learning mode. It can stimulate students' learning interest and improve their learning experience of speaking and listening through technical means outside the classroom.

2 The Key Technologies

2.1 VR Technology

VR technology refers to virtual reality technology. By using computers to build three-dimensional objects and combine them, it can break the limitation of time and space to build a virtual world. It integrates many other new high-end technologies, including multimedia perception technology and artificial intelligence concept technology. The VR technology has three main advantages: interactive Interaction, imaginative Imagination and Immersion. It is these three main advantages that make the VR technology develop rapidly. The development of this system also takes advantage of 3I features.

Interactivity means that users send dynamic instructions to the characters in the virtual world according to their needs, so that the characters can make timely action feedback according to their operations, so that the virtual characters can interact with the users. The realization of this feature requires all kinds of external equipment tools,

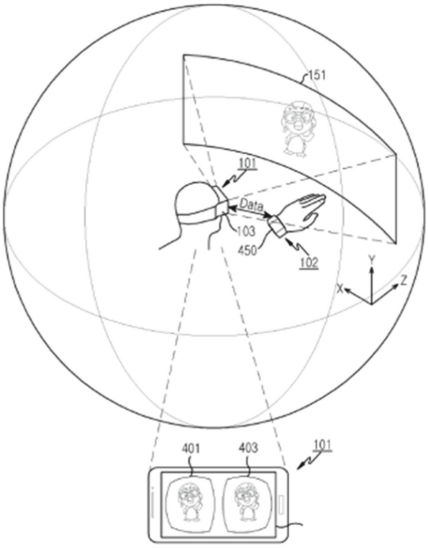

Fig. 1. Demonstration diagram of influencing factors of VR technology immersion

such as VR glasses, handles, VR head display, etc. Sense of immersion refers to the sense of reality that users participate in virtual activities as roles in the virtual environment. The factors affecting immersion are shown in Fig. 2. The main factors are the distance between the presented picture and the user, the adaptability of external hardware devices, the angle of virtual picture presentation, the depth of three-dimensional image, etc. Imagination means that users need to fully integrate with the virtual environment [3].

This system applies VR technology to realize English multi-modal teaching, and provides students with multi-sensory immersion experience. VR technology components that need help include auditory system, visual system, communication equipment, etc. The communication system relies on the access and transmission of information between the client and the server. The realization of vision takes VR glasses as the core. The auditory system can perform the function of speech recognition (Fig. 1).

2.2 3Ds Max

The 3DS MAX is a three-dimensional design software, which is developed by a subsidiary of Autodesk company named discreet. The first advantage of this system is the diversity of modeling methods. Among them, polygon modeling method is the most widely used modeling method by developers. Another advantage of 3DS is that its rendering function is more powerful than other softwares, because it has a built-in scanline renderer, which reduces the workload of manual rendering by developers. This renderer is called Menal Ray renderer, which is used in 3ds Max version 6.0 used in this system. Besides, it can also be connected with renderers such as Viay Fnatender Brezil and Lgtsepp for auxiliary use. Because of its excellent characteristics, 3DS MAX is widely used in various fields and industries, and the most used industries include game development, film animation, architectural design, medical visualization and so on [4].

2.3 Unity 3D

UNITY 3D is also simply called unity, which is a VR visualization engine tool developed by Unity Technologies. The VR engine is the core tool for building a VR system. It can coordinate the division of labor and cooperation among various components of the management system, and it can be developed only by using a unified programming language. The programming language of unity3D is mainly C#. This system can run in the mainstream systems on the PC side, and can also be published to the mobile side or even the web version in the form of WEBGL. In the process of using unity 3D, developers can only focus on the development of specific functions without considering the details of the underlying logic, so that the user's experience will be greatly improved. So, unity3D is widely used in three-dimensional animation, video games, architectural design and other industries. The Unity 3D is also a comprehensive tool that can gather multiple platforms to jointly develop the system [5].

2.4 Development Process

As shown in Fig. 2, the design flow of the overall architecture of the multi-modal English teaching system. The development of the system first uses 3DSMAX for preliminary modeling, and then uses Photoshop to beautify the maps of various objects, and outputs FBX format files with animation effects. When the model is finished, it is imported into Unity3D platform to build the environment scene. And then complete the functions that the system needs to achieve by writing scripts. The development of the module is basically completed after the realization of the function, and the system is established by exporting executable exe files through unity3d. Since VR hardware is built on different system platforms, it is necessary to select the corresponding target platform in the publishing settings. For example, HTC VIVE application is published on PC platform, and users can enter the designated webpage for learning. We can combine this system with HTC VIVE equipment for VR man-machine interaction, and enhance the user's sense of operation.

Turn on Unity and create a new empty project. On the menu of Unity, select Fiel-Build Settings, and select PC, Mac& Linux Standalone. And then, the FBX format modeling of 3dsmax is imported to build and render the scene. When some scenes are built, you can select Edit-Project Settings-Player in the menu, check "Virtual Reality Supported" in the "Other Settings" section, and click the Play button on the Unity interface for a preliminary demonstration. In the development of the system, it is necessary to use the Input Tracking class to obtain the position of each node visually. The implementation method is to use the sample script code shown in Fig. 3 and attach it to the camera [7].

The hardware environment required for the development of English multi-modal assisted instruction system is a Dell computer with NVIDIA P600 graphics card and intel CORE i7 memory 8G, and the operating system is Windows 10 64-bit. The 3D rendering of 3DMAX takes time and needs to be accelerated, which can be accelerated by real-time rendering of GPU hardware. This system adopts B/S mode for development. The database system adopts SQL server 2019 and the server is IIS6.0. The main choice of programming language of the system is C#. The VR modeling uses 3DSMAX 2020 software. Firstly, conceive, model in 3DsMax, beautify the maps of various objects by

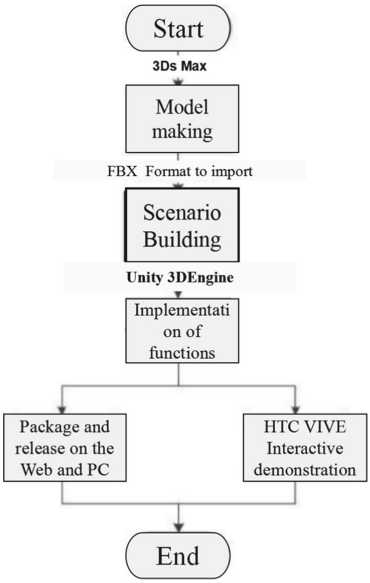

Fig. 2. The overall design framework of the system

```
using UnityEngine ;
    using UnityEngine.VR ;
        public class UpdateEyeAnchors :
MonoBehaviour
    {
    GameObject[] eyes = new GameObject[2] ;
    string[] eyeAnchorNames ={ "LeftEyeAnchor" ,
"RightEyeAnchor" } ; void Update ( )
    {
    for ( int i = 0 ; i < 2 ; ++i )
    {
    // If the eye anchor is no longer a child of us ,
don't use it
      if ( eyes[i] != null && eyes[i] 。
transform.parent != transform )
    {
    eyes[i] = null ;
    }
    // If we don't have an eye anchor , try to find one
or create one
if ( eyes[i] == null )
    {
        Transform t = transform.Find
( eyeAnchorNames[i] ) ;
    if ( t )
    eyes[i] = t.gameObject ;
    if ( eyes[i] == null )
    {
    eyes[i] = new GameObject( eyeAnchorNames[i] ) ;
        eyes[i] 。 transform.parent =
gameObject.transform ;
    }
    }
    // Update the eye transform
        eyes[i] 。 transform.localPosition =
InputTracking.GetLocalPosition ( ( VRNode ) i ) ;
        eyes[i] 。 transform.localRotation =
InputTracking.GetLocalRotation ( ( VRNode ) i ) ;
    }
    }
    }
```

Fig. 3. Realization of the code of obtaining the eye position by VR device

Photoshop, and output FBX format files with animation effect. Secondly, use Unity3D 2020.3.32 to design the scene and develop its functions. You need to install JDK to configure environment variables, download and install JDK on steamVR platform, and import FBX format files, sound effects, lights, etc. into Unity3D for integrated development. When the development is completed, the software is tested on the steamVR platform with HTC VIVE somatosensory handle, headphones and headband display.

After the test is completed, release the PC-side system for users to use. Through the introduction of the above key technical theories, the overall environment of the system development, the running process of related software and tools are determined, and the technical feasibility of the overall project of English multimodal assisted instruction system is also clarified [6].

3 Demand Analysis

The original intention of the development of English multi-modal assisted instruction system based on VR technology is to transform traditional English classroom with single teaching mode into interactive English classroom in virtual world, so that students can improve their understanding of words, contexts and sentence patterns in this process. This system is designed for the learning environment in both after-class and in-class situations. The user groups faced by the system are teachers and students. The client ports of the two types of users are the same, and no special development is made. The main functional modules of the system are divided into two parts, namely, off-line teaching and classroom-assisted teaching. Teachers mainly use classroom-assisted teaching modules, and students mainly use extracurricular online learning modules. In the online learning module, students train their listening and speaking ability after class through dialogue with NPC. The system requires teachers to design and set up virtual scenes related to topics according to their English teaching hours, and teachers should immerse themselves in explaining the application of different grammar words in each scene, which will improve students' English level in an all-round way by putting them in an English environment.

4 Function Realization

4.1 Extracurricular Online Learning

Students can click to enter the extracurricular online learning module, and they can see the list of scenes, such as VR scenes such as ordering food in restaurants, changing airport, and saving money in banks. When you select the VR scene category you are interested in, enter the VR scene, and the system will first demonstrate the English dialogue process between two NPC characters. During this period, students will watch the dialogue content from the perspective of the third person, and observe and learn in all directions in the virtual situation. You can also choose one of the NPC roles to exchange, have one-on-one communication and imitate the role dialogue. This function module is mainly realized by combining collision detection with speech synthesis module in Unity. First of all, it works on the Unity editor, adding the Box Collider component to each character, and selecting is Trigger in the trigger attribute, so that the collision body acts as a trigger. The Collider Ctrl script is used to control the sound effects and dialog display when students encounter different NPCs. The system will grade the sentences that students imitate, so as to improve their speaking ability. Speech evaluation is an important module for scoring spoken English standards. Speech evaluation is divided into speech recognition and evaluation, and then the user's scores on stress, plosive sound, fluency and synthesis are

returned. After the students read the sentences, the students' reading pronunciation and standard sentence pronunciation are sent to the scoring interface of speech evaluation, and the score of each word and the overall score will be obtained. And the overall comprehensive score, fluency score, completeness score and pronunciation score will be obtained. Through the functional sub-module of oral practice, students and users can be effectively helped to find and improve the shortcomings of oral pronunciation. During the learning process, students play it repeatedly according to their actual oral English mastery, and imitate and repeat it until they master it. You can also check the word translation of the physical objects in the virtual situation, so that students can exercise their spoken English while training their word memory and English language ability. During the learning of the dialogue content, you can choose to reread or repeat. After the demonstration of the dialogue, you can also exchange with the roles in the situation. [8] The realization of the function of the evaluation and numeration module needs the related technologies of data preprocessing, feature extraction algorithm and pattern matching. In the way of collecting Sound data, audio nodes such as Audio Clip and Movie Texture are used to add audio nodes to create the access mode of sound source, and then Sound is used to specify the playing mode of sound. The audio format after voice input is WAV or MPEG-1. During the data preprocessing stage, the system will pre-process the audio data input by users through external hardware devices by framing and windowing. And after the preprocessing, the sound file is processed by feature extraction. In the process of feature extraction, mathematical algorithms should be used to study parameters, including linear prediction method and perceptual weighting method [9].

4.2 Auxiliary Classroom Teaching

The classroom assistant teaching module of the system is equivalent to the VR electronic courseware of teachers, and students learn English through the multimedia classroom on campus. The functional module includes two sub-modules, namely, VR courseware resource library and VR classroom. When you click on the VR courseware resource library, you can see three-dimensional objects and scenes created and uploaded by other users. Teachers and users can directly quote these components to finished courseware with different themes to further improve VR courseware. By clicking on the VR classroom module, you can see the VR courseware composed of virtual scenes related to themes designed and built by the system according to the teacher's English teaching hours. In the classroom of multimedia teachers, teachers guide students to watch virtual scenes together. When a teacher talks about a key word, teachers and students can observe the related objects described by the word through VR scenes. Each object component in each scene will be described by English words and example sentences. When the user uses the handle to stay on an object, the object will be highlighted. The method of highlighting is to import a Highlighting System plug-in into the system, which is used to load scripts to the specified object, so that the object can achieve the intuitive effect of luminous reminder. The code statement to control the object to stay on is on (Color.blue), and the prompt to turn off the light is ConstantOff (Color c). There are many key codes besides the first time, which are saved in Plugins. This article uses Unity's own GUILayout to realize UI interface. Besides using GUILayout to build the

```
using                                   Gameobject m_ goRotateobject;       &&!UIManager.m.m_bIsInTextur
UnityEngine;                            public                              eWindowRect)
using                                   static                              {
system. collections;                    bool                                 m_fDeltaX = Input.GetAxis(
public                                  m_ bIsRotateobject;                 "Mouse x"
class                                   // use this for initialization      )*Time.deltaTime*m_fSpeed;
Contr1Rotation : MonoBehaviour          void                                m_fDeltaY = Input.GetAxis(
{                                       start () {                          "Mouse y"
public                                  m_ fSpeed = 80;                     ) * Time.deltaTime * m_fSpeed;
float                                   m_ bIsRotateobject =false;          m_goRotateObject.transform.
m_ fDe1tax;                             }                                   Rotate(
public                                  // Update is called once per frame  new
float                                   void                                vector3(0,-m_fDeltax,m_fDeltaY),
m_ fDe1tay;                             update () {                         space .wor1d);}
public                                  if                                  }
float                                   (input.GetMouseButton(0)){          }
m_ fspeed;                              if                                  }
public                                  (ClickAndDeal.m_bIsFreeze
```

Fig. 4. Control object rotation operation code

graphic panel, in the ContrlRotation.cs class, the method of dragging the object with the handle is also implemented to make the angle of view rotate 360 degrees around the object. The implementation code of this part is shown in Fig. 4. This system can also judge the coordinate position of the object by the static variables of the class, and view the object in all directions by rotating the coordinates of the handle, which also enables students to get a more real and three-dimensional feeling and improve the visualization of the learning content [10].

4.3 The System Testing

Before the system goes online, select a class of students to test the learning effect of vr virtual classroom teaching system. This system has 21 class hours per semester, and every 7 class hours has three stages of testing: early, middle and late. The control group uses traditional classroom teaching methods, while the experimental group uses vr virtual classroom English teaching system. The results of the three exams of the two groups of students are shown in the Fig. 5. The Fig. 5 shows that there is little difference in the initial stage. The experimental group has a low score due to maladjustment, and the mid-term result is close, but the experimental group starts to overtake ($P = .660 > .01$). The difference of the late experimental group is as high as 6.5 points compared with the control group ($P = .008 < .01$). According to the independent sample t test, there is no significant difference between the two groups in the average score of the middle test. The test results show that the system can effectively help students improve their English learning ability.

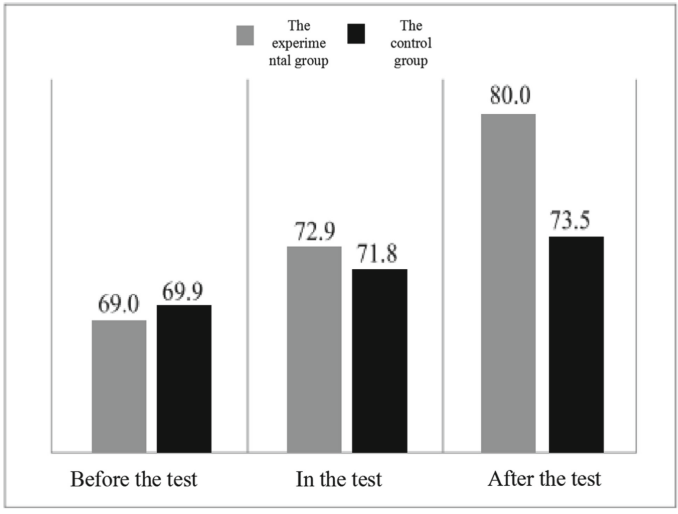

Fig. 5. Test result graph

5 Conclusion

The application system of computer English teaching is developed by combining VR technology with English multimodal teaching based on the theoretical research of multimodal teaching. This system can make the teaching process free from the constraints of time and environment, and maximize the use of learning resources. But my ability is limited, and the system still has many shortcomings.

Firstly, due to the limited energy, there are relatively few VR curriculum resources that can be developed. Secondly, the hardware equipment used in the virtual classroom on campus is not advanced enough, and the details of many VR models cannot be displayed completely. But China's science and technology are constantly developing, and the talent pool is abundant. I believe that these problems will be solved in the near future, and the classroom of VR technology will be gradually popularized, realizing the beautiful vision of VR teaching for all.

References

1. Xie Lin Lin. A Survey of Multi-modal Teaching in Senior High School English Class. Yan'an University. 2019.06
2. Liu Jie, Wu jianping, Liu Jiapei. Interactive Situational English Teaching Design Based on VR Technology. China Educational Technology Equipment. 2020.11
3. Tang Zhuowei. Research on College English Teaching Theory and AI and VR Technology Assisted English Autonomous Learning Method. Technology Enhanced Foreign Language Education. 2020.12
4. Nie Hui, Huang Baiping. Research on Innovation of College English Teaching Mode from the Perspective of VR/AR Technology. Journal of University of Shanghai for Science and Technology (Social Science Edition). 2020.06

5. Wang Jingjie, Hu Yunxuan, Zhao Jin. Research on the Application of VR Immersion Teaching in English Learning in Primary and Secondary Schools. Journal of Hubei Normal University (Philosophy and Social Sciences Edition). 2020.02
6. Fan Dongmei. Research on College English Teaching Mode Based on VR Technology. Overseas English. 2022.04
7. He xiaorui. On the Application of VR Technology in English Classroom Teaching in Senior High School. Teaching Management of College Entrance Examination. 2022.2
8. Sun Hongxing. Research on English Teaching in Secondary Vocational Schools Based on VR Technology. Henan Institute of Science and Technology. 2020.05
9. Xu Zhan. Design of classroom teaching assistant application based on VR equipment. The electronic world. 2018.09
10. Liu Zeyu, Zhang Siwei, Lu Jin. Design of Virtual Reality (VR) CAI System Based on "internet plus". Science Tribune. 2019.06