Research on Operation Training of Large Vehicle Based on Augmented Reality Technology

Ming Wu, Gang An, Wenxuan Sun, Wei Li, and Luoguo Wang(✉)

Institute of NBC Defence PLA Army, Beijing, China
15810996827@163.com

Abstract. With the continuous development of new technologies such as virtual reality, augmented reality and mixed reality, its application in the field of large vehicle operation training is becoming more and more extensive. Based on the practice of smoking vehicle operation and training, this paper designs an intelligent interactive training system based on the integration of artificial intelligence technology and augmented reality technology, and researches and demonstrates the key technologies of its implementation. It realizes the superposition and interaction of virtual equipment, such as virtual animation and model, and real installation and real scene, provides a good virtual interactive experience of augmented reality, meets the training needs under complex conditions in the field, and provides technical support for the improvement of training quality and benefit.

Keywords: Augmented Reality · Training of Large Vehicle · Operation

1 The Introduction

Augmented reality, originating from virtual reality technology, is a technology that enhances the user's perception in the real environment by superimposing virtual information on the computer. It can superimpose virtual objects, objects, scenes or information generated by the computer into the real scene, so as to realize the reality enhancement of the physical world [2].

The application of augmented reality technology in equipment training has many advantages, such as the combination of virtuality and reality, economic security and good man-machine interaction, which are not available in traditional training [4]. It is the key point of the combination of simulation training and practical training, and plays a very important role in the improvement of training quality [1]. So far, the application of augmented reality technology in the field of equipment training in China is generally implemented under the constraints of indoor environment, good lighting conditions, QR code recognition assisted positioning and other technical means, there is no mature practical application of equipment training under complex conditions in the field [3].

X. Yuan et al. (Eds.): ICEKIM 2023, AHCS 13, pp. 254–264, 2023.
https://doi.org/10.2991/978-94-6463-172-2_28

2 Analysis of the Current Situation of Operation Training of Smoking Vehicles

2.1 Training Characteristics of Smoking Vehicle

Smoke vehicles [9] usually have different smoke generation system, the engine is mainly used to generate high temperature and high speed gas to atomize the smoke material at high temperature and high pressure, which is discharged into the atmosphere and condensed to form a shielding smoke screen, or the smoke material is dispersed at high pressure to form a solid smoke screen. Different smoke systems can be used cooperatively according to the characteristics and requirements of smoke screens, and can emit different smoke screens. Different smoke generation systems have different functions, performance parameters and smoke generation principles., the construction principle, operation flow and use method are different so much from each other. Coupled with the overall structure and production process limitations of smoking vehicles, most of the devices are exposed, and there are more parts, in particular, there are many pipelines and switches. It is easy to make novice in the process of implementation of operation training confused, and even fear of difficulties. The initial training and teaching of smoking vehicles is relatively difficult, and the teachers need to teach the trainees hand to hand in order to prevent the loss of personnel or equipment caused by misoperation, and the trainees need to spend a lot of time to fully master the actual operation of the smoke vehicle.

2.2 Training Method of Smoking Vehicles

2.2.1 Theoretical explanation

The theory explanation of smoking vehicles is usually organized by teachers and carried out in the form of unified learning. Its methods are also diverse. It can be taught theoretically with the help of indoor multimedia, or it can be explained intuitively with real equipment [8]. Its main content is to explain the construction principle and the smoke formation principle of different smoke screens. The theoretical content is strong, is the basis of smoke vehicle operation training.

2.2.2 Simulating Operations

The simulation operation is the training method before the physical operation, mainly under the organization of the teacher, the use of auxiliary training equipment, in accordance with the corresponding operation requirements, to simulate the operation. The main content that can be simulated is the operation flow and method of different kinds of smoke screen. Its purpose is to enable the trainees to master the basic operation process of smoking vehicles, and lay a foundation for the implementation of safe practical training.

2.2.3 Physical operation

Physical operation is the most critical step of smoke vehicle training and the ultimate goal of smoke vehicle training. Physical operation is usually implemented in a fixed

site, organized by the teacher, according to the number of physical objects, the scale of the training group and other factors to carry out group training. In the process of group training, the teacher should make clear the matters needing attention and relevant requirements. In case of any misoperation of the trainee in the process of operation, the teacher shall organize correction in time.

2.3 Training Effect of Smoking Vehicle

The training effect of smoking vehicle will be slightly different due to the different scenes where it is located. In one case, under the condition of a small number of physical objects and a small training scale, the method of hand-holding teaching is adopted for the training and teaching of smoking vehicle. The advantage of this training and teaching method is that it can enable new scholars to master the operation process of smoke vehicle relatively quickly, but the disadvantage is that the trainees lack of theoretical knowledge. Most of them only know the general process of smoke vehicle operation, but do not know the operation process of smoke vehicle system. They cannot flexibly deal with problems when they encounter them. The other is centralized training for a large number of trainees. The teaching methods are more diverse, the hardware facilities are more perfect, and the process is more systematic and comprehensive. But the disadvantage is that the audience is large, and the teachers are relatively limited, can not do hand-to-hand teaching, trainees can not be able to master the operation of smoking vehicles in the limited learning time.

3 Overall Design of the System

3.1 Requirement Analysis

According to the smoke vehicle training characteristics, augmented reality equipment is used to realize intelligent recognition of physical components and personnel actions. The virtual scene is superimposed on the real scene in timely. Through the human-computer interaction and operation guidance under the immersive virtual scene, the practical guided training and operation action correction are realized. It provides intelligent means for trainees to carry out autonomous training. The overall design should meet the needs of equipment training [10].

3.1.1 Better Ergonomic Design

In order not to affect the action of the physical operation, combined with the basic requirements of ergonomics during long-term training, augmented reality equipment needs to meet the basic requirements of lightweight, wearable split-piece structure design and power supply of long-endurance terminals, so as to free the hands of trainees and not affect their operational experience of the real installation.

3.1.2 Requirements for localization

The overall system hardware and software design is carried out under the constraints of localization. Augmented reality devices need to adopt non-cloud technology architecture

to achieve key functions such as intelligent identification and 3D registration. Portable terminals and augmented reality devices should have bidirectional control capabilities, while combining the means assisted by QR code recognition technology to maintain the effective operation of intelligent identification and training systems.

3.1.3 Good Environmental Adaptability

The system operation should fully consider factors such as the change of illumination conditions and the restriction of different operating angles under the complex conditions of the external field, so as to ensure the rapid and accurate input information of intelligent interaction.

3.2 Overall Hardware Architecture of the System

Wearable augmented reality devices that meet the requirements of complex external field conditions include augmented reality glasses, portable computing terminals and data exchange accessories [5]. The underlying hardware architecture of augmented reality smart glasses needs to include various sensors such as IMU, RGB, fisheye and voice, which are required for system recognition. It can complete video and picture information collection, interactive gesture capture, virtual and real superposition display, voice input and output and other functions. The computing terminal should have powerful computing and storage capabilities, provide storage space for the data needed for system operation, and realize bidirectional control system computing with augmented reality glasses. The data exchange accessory connects the augmented reality smart glasses and the portable computing terminal, providing audio, video and gesture interactive data transmission and stable power supply for the augmented reality glasses. To realize the recognition and processing capability of augmented reality, in addition to basic sensors, the core is a neural network processor based on edge computing technology, which can run deep learning models [6]. The overall hardware architecture is shown in Fig. 1.

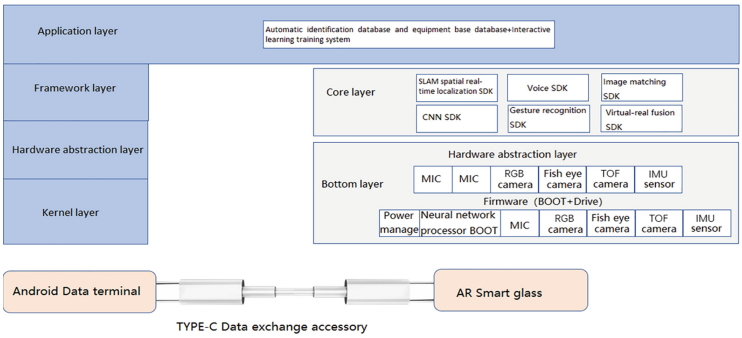

Fig. 1. Overall design framework of the system

3.3 Overall Software Architecture

In order to meet the needs of equipment training and provide a good interactive training experience, support on software design of the software provide development kit SDK, real-time position tracking module, spatial real-time positioning submodule, gesture recognition submodule, virtual-real fusion submodule, image comparison submodule, recognition resource library, information opening module and other software function modules. During the implementation process, video and image data from different scenes, lighting and angles of equipment parts or assemblies should be collected, and then data labels should be manually operated. Input TensorFlow neural network for model training, and get accurate training model. The data model adapted to the edge neural network processor and OpenVINO format was obtained through model optimization and transformation. The data model was deployed to the edge neural network processor system, and the local real-time intelligent object recognition was realized through OpenVINO inference acceleration. The intelligent recognition software architecture is shown in Fig. 2.

At present, the common toolkits for AR system development mainly include ARToolKit, which uses manual identification to track and locate, Vuforia and EasyAR, which uses natural features to track and locate. Taking Vuforia, a widely used development tool, as an example, it is necessary to register the identity information in the Vuforia website during the development process, and then edit and publish the virtual information. At the same time, there are certain requirements on the registered identity information to facilitate identification. However, based on the requirements of training localization operation, through the application practice and field test of the toolkit developed by the common augmented reality system, the technical process architecture based on deep learning is more in line with the equipment training requirements under complex conditions in the field, as shown in Fig. 3.

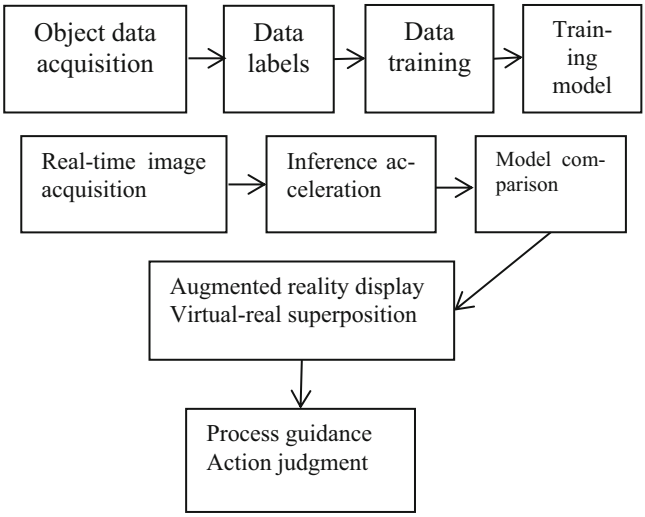

Fig. 2. Architecture diagram of custom training software

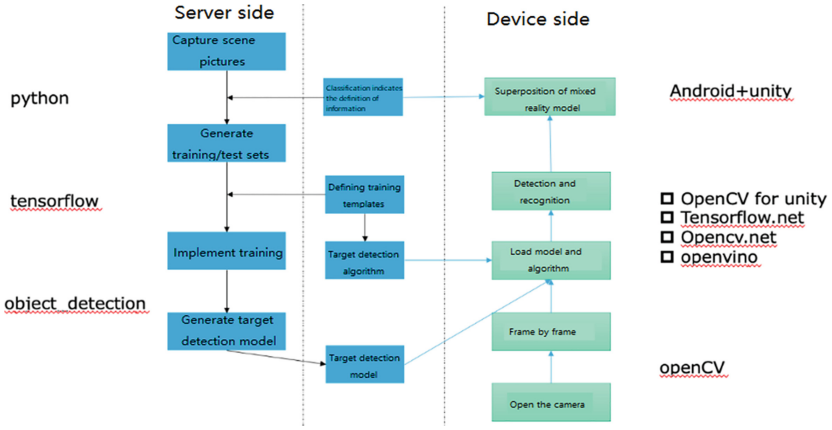

Fig. 3. Technical flow diagram of augmented reality application practice

4 Key Technology Implementation Approaches

4.1 Deep Learning Automatic Recognition Technology

At present, the mainstream application mode of deep learning mostly adopts independent image acquisition equipment combined with powerful computing server, and adopts wireless network or wired network. Wearable AR glasses are responsible for collecting the image or video data of the field equipment, and sending the image or video data to the server for recognition through the network. The results of recognition are fed back to the AR glasses. According to the recognition results, the AR glasses carry out corresponding operations, such as displaying 3D models, playing animation, video, text and other actions.

According to the overall design of the system, combined with the constraints of the actual application scenario. Remove the cloud architecture or dedicated server architecture identified by the server backend [11]. The recognition computing power is realized through the edge computing neural network processor of the terminal, that is, heterogeneous computing to complete the intelligent recognition computing power. Firmware development of edge neural network processor includes initialization and driver interface of processor BOOT, RGB camera, dual fisheye camera, TOF camera, IMU attitude sensing sensor, CODEC, power management, etc. and the hardware abstraction layer software design of the above hardware module. The hardware independent interface description is provided for the virtual-real fusion module, intelligent image recognition module, position tracking module, image comparison module, gesture recognition module, and spatial real-time positioning module.

CNN convolutional neural network is a key SDK to complete the edge neural network processor for deep learning image comparison, and the SDK relies on the OpenVINO toolkit. TensorFlow completes the model training and obtains the PB general data model. Config. json is used to complete the configuration of BLOB and JSON files to complete the deployment of heterogeneous computing.

Lightweight portable intelligent computing terminal, the transplantation of deep learning model to optimize deployment is a necessary condition to maintain real-time performance. In terms of technical implementation, the MobileNet feature extraction network is finally selected and the SSD algorithm is used to perform the classification task, which is the key technology to complete the performance guarantee.

4.2 Research on Optimization Technology of Deep Learning Results

Convolutional neural network is used to complete the task of logo identification. Logo identification not only has strong generalization ability, so that the recognition object is no longer limited to a specific identification, but also the recognition accuracy is high. However, in deep learning methods, convolutional neural networks have long training time and high training cost, and weak ability to adapt to new application scenarios.

In view of the problems such as the overall deep learning convolutional neural network training method of unit module, the number of recognition methods is large, inflexible, time-consuming for repeated modification, and poor scalability, etc. The equipment element model, which divides the identified unit into smaller modules, is used for classification learning and training. Through the intelligent scheduling algorithm of multi-element and multi-category recognition learning model results, the agile development of deep learning application is realized, and the rapid transplantation of equipment units of the same series and different models can be realized [12].

4.3 Research on Integration of QR Code Recognition and Other Technologies

The key technology of intelligent recognition adopts deep learning model. The characteristics of deep learning model are huge amount of data and time-consuming training, but the successfully trained model has superior recognition performance. In order to solve the problem of too long learning and training time, the classification learning technology of modular subdivision into assembly elements and the intelligent scheduling algorithm for multiple training models are adopted. As a result, the logo identification not only has strong generalization ability, so that the recognition object is no longer limited to specific identification, but also the recognition accuracy is high.

Intelligent identification based entirely on deep learning requires high load operation of edge computing units, which poses a severe challenge to the overall endurance of the system and the heat dissipation performance of augmented reality [13]. Therefore put forward backward compatibility training is usually not easy to be interference light signal as the medium, operating status of identifying objects by means of QR code the transmission of information, to ensure that the interference of augmented reality system under the normal use, at the same time, the application of the work force of portable computing terminal on the efficient computation of core business object. Therefore, in the technical research and practice, the secondary training objects, which are extremely difficult to identify or have large interference factors, can be identified by the necessary QR code assisted recognition, and the realization of a stable and reliable training system can be further promoted through the computing power balance mechanism.

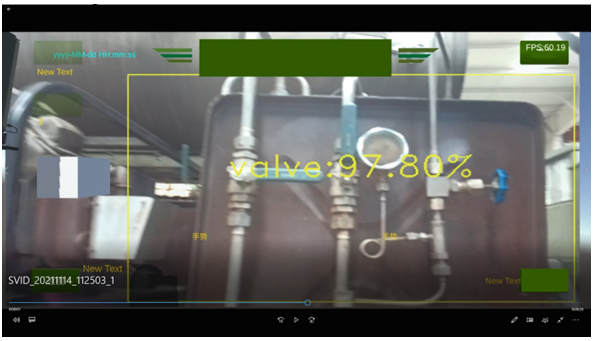

Fig. 4. Single test recognition accuracy of the box under indoor conditions

Fig. 5. Single test recognition accuracy of generator under indoor conditions

5 Verification of Core Functions

5.1 Image Comparison Module

The virtual and real fusion display presented by augmented reality technology is based on the accurate recognition of the target object. By taking multiple pictures of target objects from multiple angles, constructing feature points and developing a database matching feature points in a 3D engine. After a large number of deep learning, the recognition results are optimized to achieve target recognition. After testing, the recognition accuracy is more than 93% under good illumination condition. The identification results are shown in Fig. 4 and 5.

5.2 Virtual-Real Fusion Display Module

The identified virtual model and physical overlay reality are the presentation forms of augmented reality technology. After solving the overlapping display coincidence and contrast problems. SLAM (Simultaneous localization and mapping) is implemented with hardware based on TOF camera and IMU attitude sensing sensor. The display results are shown in Figs. 6 and 7.

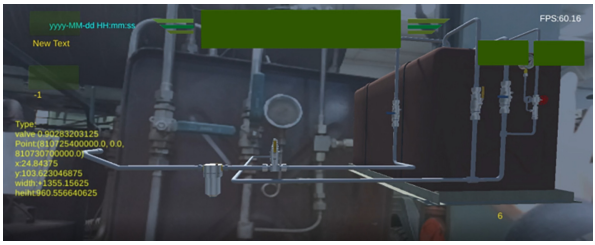

Fig. 6. Display effect of virtual and real superposition after box recognition

Fig. 7. Display effect of virtual and real superposition after generator identification

5.3 Gesture Recognition Module

Through gathering gesture images and autonomous learning machine, test device can realize automatic recognition of 6 static gestures and 2 dynamic gestures. Each gesture corresponds to an operation mode. The identification results are shown in Fig. 8 and 9.

Fig. 8. Recognizes the static OK gesture

Fig. 9. Recognize a dynamic gesture with five fingers open

6 Future Development Trend

After research, the domestic augmented reality recognition system has made progress in the field equipment training part of the key technology research, and verified and tried the multi-person cooperative mode training technology approach. However, due to the limitations of many factors such as time and data, especially for the research of multi-position cooperative military equipment training platform, there are still some problems worth further optimization and improvement.

First, it is necessary to overcome the convergence of video streaming media technology, deeply study the network security and performance optimization of local AD hoc network, and realize the effective operation of the system in collaborative mode.

Second, it is necessary to adapt the intelligent server for the recording, evaluation and other functions of training business. It is necessary to further study the back-end intelligent guidance technology of the server and the edge computing unit allocation and scheduling technology of the front-end smart glasses, so as to provide a technical basis for intelligent training.

Third, intelligent terminals can further integrate sensing devices such as physiological monitoring systems to provide metadata for military training evaluation platforms supported by data or models.

7 Concluding Remarks

The key technology of the field training of smoke vehicle based on augmented reality is the inheritance and development of the adaptation of augmented reality technology from indoor scene to complex environment [7]. At present, the research on augmented reality in China is still in the initial development stage. However, with the continuous emergence of corresponding hardware equipment, software platform and the establishment and development of theoretical system, the extensive application of augmented reality technology will become an inevitable trend. It can significantly improve the quality and effectiveness of training, which is of great significance to improve the level of information technology.

References

1. Chen Yuwen, Research on Augmented Reality Technology and Its Application in Military equipment and Simulation Training [J]. Journal of System Simulation, 2013, V. 25, No. 260, 264–268
2. HU Xiaoqiang. Virtual Reality Technology [M]. Beijing: Beijing University of Posts and Telecommunications Press, 2005.7
3. Yang Zhen, GAO Junxiong, Research on Application of Virtual Simulation Training for Equipment [J]. Computer and Digital Engineering, 2013, 99–102
4. Jon Peddie. Augmented Reality – Everywhere [M]. Deng Baosong. Beijing: Publishing House of Electronics Industry, 2020, 24–37
5. WU Yueming. Interactive Operation Method of Augmented Reality Workshop Layout Design [J]. Computer Integrated Manufacturing Systems, 2015(5)

6. Chen Jing, Wang Cong. Research progress and Core Technology of Augmented Reality: Top [J]. Standardization Research, 2012, 6:35–37
7. LI Daguang. Military Simulation Technology [J]. Encyclopedia Knowledge, 2006, (8): 63–64
8. HU Limin, ZHAI Wenjiang, HE Donghui, Theory of Military Training Energy [M]. Beijing: National Defense Industry Press, 2009
9. Academy of Military Sciences, Military Language of the Chinese People's Liberation Army (complete edition) [S]. Military Science Press, 2001
10. Hou Ying, Xu Weiwei. Survey of Augmented Reality technology [J]. Computer Measurement and Control, 2017(25):1–2
11. Qi Yue, Ma Hongmei. Augmented Reality: Features, Key Technologies and Applications [J]. Small Microcomputer Systems, 2004(5):900–902
12. Xia Defang, Liu Chuancai. Mobile augmented reality registration method based on manual identification [J]. Modern electronic technique, 2015, 38 (8): 26–30
13. James C M, Charles K B, John P. Improving robotic operator performance using augmented reality[C]. In: Proceedings of the human factors and ergonomics society 51st annual meeting, 1635–1639