

Received August 3, 2020, accepted August 27, 2020, date of publication August 31, 2020, date of current version September 11, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3020637

# Intelligent Video Surveillance of Tourist Attractions Based on Virtual Reality Technology

JIE HUANG<sup>1</sup>, ANMIN HUANG<sup>1</sup>, AND LIMING WANG<sup>2</sup>

<sup>1</sup>College of Tourism, Huaqiao University, Quanzhou 362021, China

<sup>2</sup>College of Tourism, Xinjiang University of Finance and Economics, Urumqi 830012, China

Corresponding author: Liming Wang (wanglimingxjufe@163.com)

This work was supported by the Ministry of Education in China (MOE) Foundation of Humanities and Social Sciences (to Anmin Huang) under Grant 18YJA630039.

**ABSTRACT** This question combines virtual reality technology with video surveillance technology to realize an intelligent video surveillance system based on virtual reality technology in tourist attractions. It analyzes the technology and theory involved in the video surveillance system based on virtual reality technology and then designs the architecture and functions of the system. The system is divided into three parts, such as local monitoring client, streaming media server, and remote monitoring client. It detailed the design and implementation of each module. We use the OpenCV library and moving target detection algorithm to realize the intelligent analysis of surveillance video and use the OpenGL engine to render the surveillance video of VR panorama of tourist attractions to achieve the effect of virtual reality. The video surveillance system implemented is tested and analyzed in combination with the actual application of tourist attractions. The test results show that the system can operate stably and meet the design indicators. The remote monitoring delay is controlled within one second. The remote monitoring client has achieved the effect of virtual reality.

**INDEX TERMS** Virtual reality, tourist attractions, smart video, video surveillance.

## I. INTRODUCTION

With the maturity of sensor technology and the development of cloud computing, the amount of information and data in daily life has exploded, from weather forecasts to geographic information, from traffic flow to social media, and data continues to grow from the original two-dimensional space. The benchmark gradually evolves into three-dimensional space. This type of spatial data that has both time and space and changes with time is called Spatio-temporal data [1]. The Spatio-temporal big data has the characteristics of Spatio-temporal, multidimensional, multi-type, and dynamic correlation. For Spatio-temporal data, what needs to be stored is not only the position of the object space but also Spatio-temporal objects whose scope changes with time. It is used to find the position, scope, and scope of a moving object in the past, present, or even predict the future at a certain time or time interval behavior [2]. Travel navigation and tourism planning in daily life, in other words, Spatio-temporal big data involves all aspects of urban development

planning and personal life. The traces left by each of us in the Spatio-temporal environment can be aggregated into big data, through simple location attributes and all things are interconnected, allowing data to break industry and regional restrictions and create all-round value [3], [4]. At the same time, with the gradual maturity of virtual reality technology, a new video immersion interactive experience is becoming more and more popular. By combining virtual reality technology with video surveillance technology, it will give users a new way to view surveillance video. The interactive experience allows the user to more intuitively and accurately grasp the situation of the surveillance area while viewing the surveillance video, giving the user a sense of immersion [5]. In the process of the development of the security business in the direction of intelligence, technologies such as virtual reality will become new technological competition points in the future. Therefore, it is of great practical significance to develop an intelligent video surveillance system based on virtual reality technology.

Huawei series on the phone conference announced Huawei involved in the VR field and released a VR helmet HuaweiVR, which is Huawei's first VR helmet [6]. This was

The associate editor coordinating the review of this manuscript and approving it for publication was Zhihan Lv.

the most influential VR product among the many giant companies that entered the VR field across the industry at that time. Intel will be one of the first technology providers to implement live sports on multiple VR devices [7]. Besides, Intel also announced and computer vision company HypeVR cooperation. HypeVR is a focus on six degrees of freedom of ultra-high-definition photo virtual reality motion capture and playback technology developed by the company [8]. The two companies are preparing to incorporate HypeVR's stereoscopic video content into the alloy project. The project mainly develops integrated fusion real solutions to achieve the perfect integration of the real world and the virtual world. At the same time, Intel also announced that it will cooperate with some leading OEM manufacturers to commercialize this open hardware platform in the fourth quarter [9]. ICRRealTech demonstrated a panoramic VR virtual camera. Forte has cooperated with well-known security companies such as milestone, axis, Avigilon, BOSCH, and Genetec, and has successfully deployed some 3D/AR concept monitoring platforms. Samsung Electronics launched the first virtual reality camera Gear360, it can be 360 degrees shooting. Each fisheye lens is equipped with a 15-megapixel sensor, which can capture still images with a resolution of  $7776 \times 3888$  and a frame rate of 30 frames. Dahua's annual report mentions the need to strengthen the research and deployment of VR/AR and writes in the future development trend of the industry in the annual report [10], [11]. Video surveillance technology will be deeply integrated with new application fields such as VR/AR and redefine product properties. ST-Hadoop is the first completely open-source MapReduce framework that provides native support for Spatio-temporal data and is specifically designed for Spatio-temporal data [12]. Spatiotemporal data perception can be injected into each of its layers. MD-HBase is an extensible data management system for LBS, which mainly deals with multi-dimensional spatial data, but can also deal with Spatio-temporal data [13]. The DeepFace algorithm proposed by the Facebook company. It achieved an accuracy of 97.35% on the LFW data set, which is the first time that it approaches human-level performance. The DeepID algorithm proposed the Chinese University of Hong Kong has achieved 97.45% accuracy on the LFW dataset [14]–[16]. It builds a posture correction network model through the point distribution model and the corresponding posture parameters to generate a virtual face view and finally used Gabor transform to extract features to achieve facial posture correction [17]. The CASIA Huang put forward humans can be considered as a whole against the local information and generates a network structure for the best path [18]. The network has two paths, one is used to infer the overall structure of the face, the other is used to infer the local texture of the face, and finally, the feature maps generated by the two paths are synthesized into a frontal face view [19].

For the current monitoring system, the role of VR in the front-end is mainly reflected in two aspects. VR scenes can be spliced and synthesized using the front-end panoramic

camera, and then post-processed through supporting software, which not only satisfies users' traditional operations such as zooming and zooming but also can change users' viewing angles [20], [21]. The use of VR technology allows the front-end camera to perfectly integrate the reality of the video itself with enhanced information after digital calculation. We allow users who use video surveillance to quickly obtain information about the target object when viewing real-time video images. For VR technology, the back-end application of the video surveillance system is mainly based on the security control experience of VR itself. In the previous deployment and control of the monitoring system, the user looked at the layout of the entire system as the third vision, which was prone to deviation [22]. The VR technology can allow users to visually transition from third to first vision. The surrounding environment within a realistic target and monitoring scope of integration into the same screen, the user VR can direct screen and video equipment Interaction, especially some large-scale government projects, such as projects that use city electronic maps [23]. VR equipment can directly participate in the installation of cameras, anti-theft, access control or building control, fire protection, and other systems [24]. Users no longer look at the deployment of the system from the perspective of the layout but the entrant. The intelligent video surveillance system based on virtual reality technology in this question is based on the android system. The system can be divided into a local monitoring client, a streaming media server, and a remote monitoring client.

## II. VIRTUAL REALITY TECHNOLOGY VIDEO SURVEILLANCE DESIGN ANALYSIS

### A. OPTIMIZED VIRTUAL REALITY ALGORITHM DESIGN

The optical flow field is an instantaneous velocity field, which represents the instantaneous trend of changes in the gray value of pixels in the image. In actual situations, the change in the gray distribution of each pixel in the video stream is usually used to characterize the movement of the target. The basic principle of the optical flow method for detecting moving targets is calculating the motion vectors of the pixels in the entire image one by one, the optical flow field is established [25]. If there is no moving target in the monitoring area, the motion vectors of all pixels in the image are the change should be continuous. On the contrary, because there is relative motion between the target and the background image, the motion vector at the location of the moving target in the image is different from the motion vector of the neighborhood, namely the background, to achieve moving target detection [26]. When calculating the optical flow field, the motion vector for each pixel can be obtained because there is relative motion between the moving target and the camera. However, calculating the optical flow field is a very complicated process, and in the actual situation, there are factors such as light changes, and the brightness of the moving target surface may change, which violates the use conditions of the basic constraint equation of optical flow. Therefore,

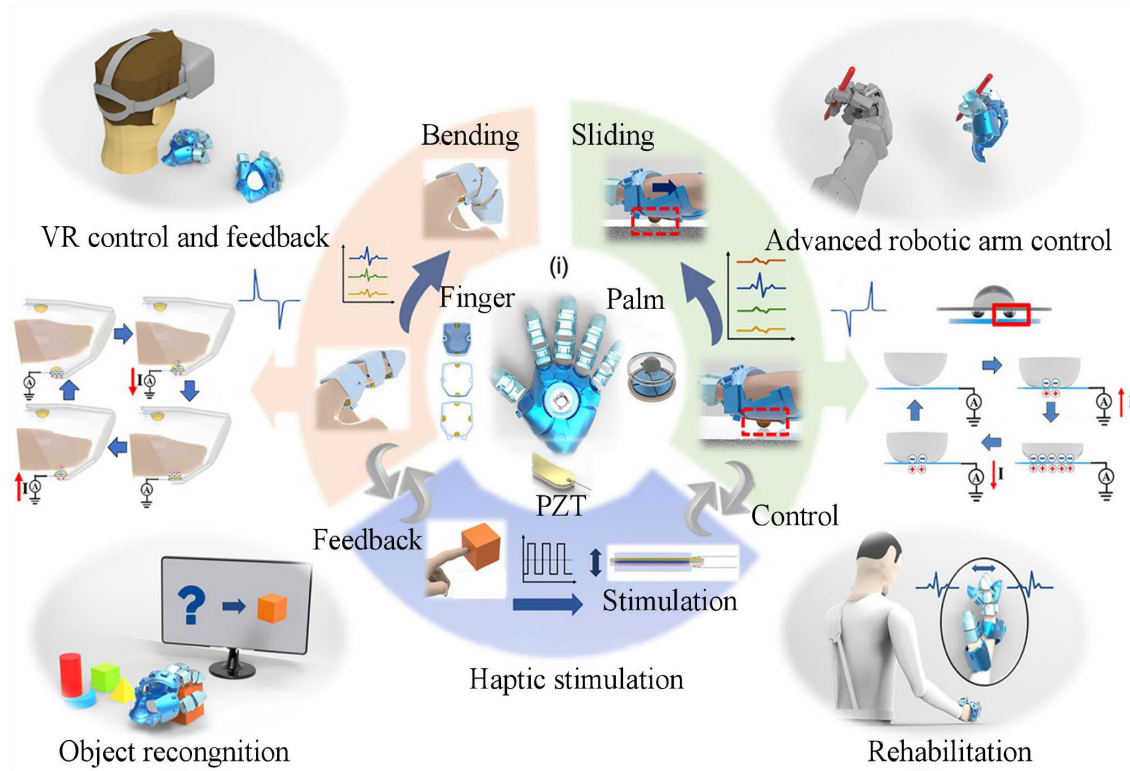


FIGURE 1. Optimized virtual reality algorithm model.

the optical flow field calculated when the optical flow method is used in practice will have a large deviation, as shown in Figure 1.

To perform edge detection on the image, it detects the image boundary, and then connect the contour line of the target boundary. Before boundary detection, filters are often used to reduce the influence of noise, and then edge detection is performed. The gradient corresponds to the first derivative, and a coordinate system is established for a two-dimensional image. For a continuous two-dimensional image, the function can be expressed as  $f(x,y)$ . The vector of the gradient is defined as:

$$\Delta f(x, y) = [K_x, K_y]^T = \left[ \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right]^T \quad (1)$$

The magnitude of the gradient:

$$|\nabla f(x, y) = \text{mag}(\nabla f(x, y))| \quad (2)$$

The direction of the gradient:

$$\phi(x, y) = \arctan(K_x/K_y) \quad (3)$$

Among them,  $j$  corresponds to the  $x$ -axis direction, and  $i$  corresponds to the  $y$ -negative axis direction, expressed as a simple convolution template is:

$$K_x = f(x, y + 1) - f(x, y) \quad (4)$$

Among them, corresponds to the  $x$ -axis direction, and  $i$  corresponds to the  $y$  negative axis direction, expressed as

a simple convolution template as:

$$K_y = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \quad (5)$$

S2 is mainly to reduce the dimensionality of three-dimensional spatial data to a number represented by one-dimensional UINT64. The main transformation process is the spherical latitude and longitude coordinates are transformed into spherical axes coordinates, then transformed into coordinates on the circumscribed cube projection surface, transformed into corrected coordinates, and finally transformed into 64bit values through the Hilbert curve.

The latitude and longitude are converted into radians, and the conversion between radians and angles is as formula (6):

$$1^\circ = \frac{\pi}{180} \text{rad} \quad (6)$$

The latitude and longitude radian are converted into a point on the coordinate system, as in formula (7):

$$H(\text{lat}, \text{lng}) = f(x, y, z) \quad (7)$$

The faces of the tangent cube are projected separately from the center of the sphere. S2 is to project all the points on the sphere onto the 6 faces of the circumscribed cube to make the sphere flat, as in formula (8):

$$f(x, y, z) = k(\text{face}, u, v) \quad (8)$$

In the previous step, the spherical rectangle on the spherical surface is projected on a certain surface of the square, and the

shape is similar to a rectangle. Due to the different angles on the spherical surface, the area of each rectangle will be different even if it is projected on the same surface. Therefore, a secondary transformation is needed to correct the projection of the spherical rectangle, as in formula (9):

$$k(\text{face}, u, v) = m(\text{face}, s, t) \quad (9)$$

$$f_d(x, y, t_1, t_2) = \begin{cases} 1 & x > 1 \\ 0 & x \leq 1 \end{cases} \quad (10)$$

It used the difference between two adjacent frames of video to determine whether there is a moving target. The simplest way to implement the frame difference method is to perform different operations on two adjacent frames of images to subtract the corresponding pixels in different frames. When the difference method operation result exceeds the set threshold, the difference image value is 1, and vice versa is 0, and formula (11) can be used to express the specific implementation principle.

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} \quad (11)$$

When the different result is greater than T, it is considered that the two frames of video have changed significantly and there is a moving target. Otherwise, it is considered that there is no moving target. Because the frame difference method selects adjacent video frames when performing the different calculations, the frame difference method is less sensitive to light changes in actual calculations, which greatly reduce the impact of light changes [27]. Besides, the algorithm is simple and easy to implement, and the execution process takes less time, which is convenient for implementation on mobile devices with relatively poor hardware performance. The background difference method to detect a moving target first need to perform a different operation between the current frame and the background image, and then compare the number of pixels obtained with the size of the threshold function. If the number of pixels obtained is greater than the threshold function, it is considered that there is a moving target [28]. The background difference method can effectively segment a complete moving target, but it is very sensitive to light changes, which will affect the accuracy of the recognition results, and how to select background frames is also worth studying. There are two common methods for this algorithm to select a background image. The background image is the first frame of the video, and then take each frame of the video and the first frame for different operations, and compare the different results with a threshold. The background image first selects the first frame of the video and then performs a different operation between the next frame and the background frame [29], [30]. If the different result is greater than the threshold function, it is determined that there is a moving target in the video. For subsequent detection, the background image needs to be updated and the current frame is taken as New background image. In the intelligent video surveillance system, real-time is the basic requirement for target detection,

and high real-time can ensure that the surveillance system promptly reminds users that suspicious targets have broken into the surveillance area. Although the frame difference method has general performance in the detection of motion characteristics, the frame difference method is simple and efficient, has no sense of light changes, and does not require high hardware requirements. It is suitable for the implementation of embedding devices. Therefore, the frame difference method is selected to achieve moving targets [31].

The operator is a second-order differential operator. It is a scalar, which is the anisotropic operation and is sensitive to gray-scale mutations. In digital images, the Laplacian is the two-dimensional equivalent of the second derivative.

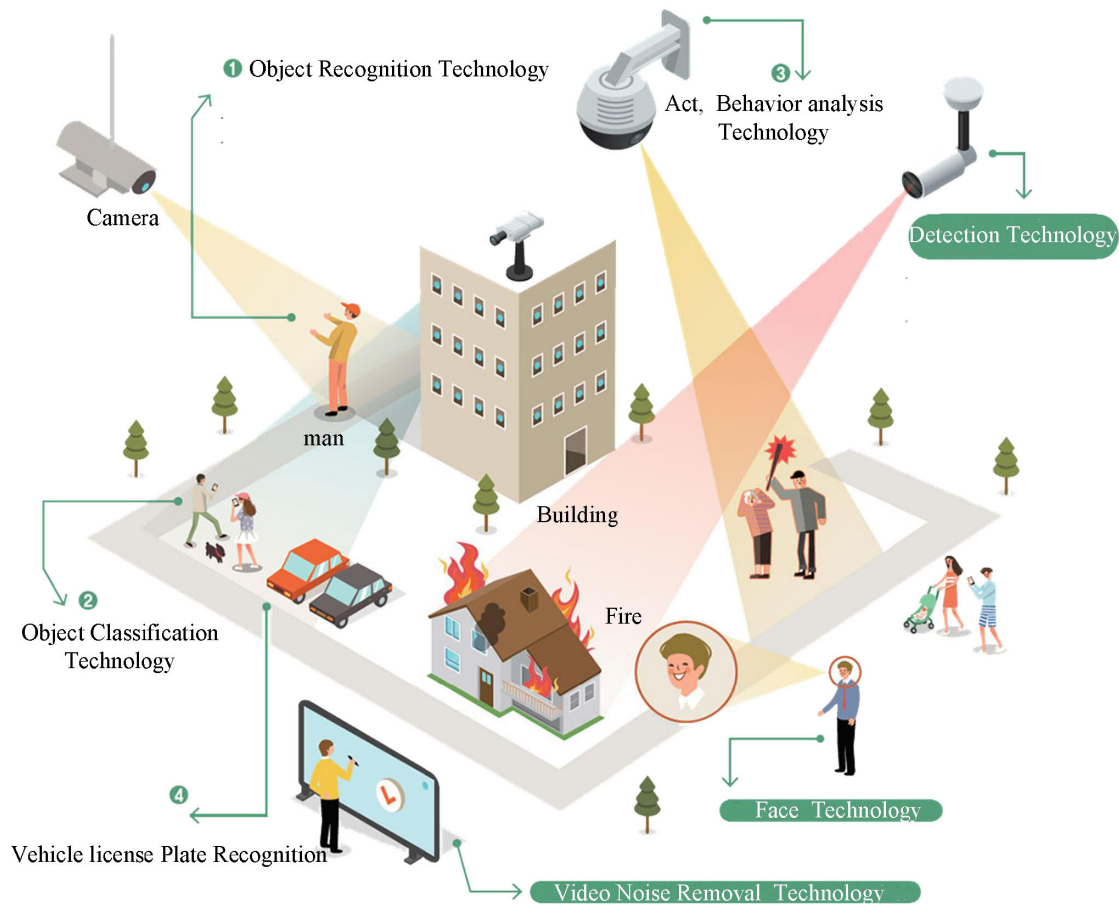
The color image or grayscale image obtained by the device contains a lot of information. In many cases, a lot of information in the image is useless to the algorithm. Before extracting useful image features, the image performs preprocessing to enhance the required information as much as possible and filter out the useless information as much as possible to highlight the key points.

## B. THE THEORETICAL BASIS OF DYNAMIC IMAGE RECOGNITION

The intelligent video surveillance system based on virtual reality technology designed in this subject is based on the android system and can be divided into local surveillance clients, streaming media servers, and remote surveillance clients. The camera of the system uses a VR panoramic camera. The local monitoring client is mainly responsible for the collection of VR panoramic surveillance video, surveillance video management, and intelligent monitoring, and pushes the collected surveillance video to the streaming media server. The streaming media server plays a relay role. On the one hand, it receives the VR panoramic monitoring video media stream packaged and transmitted by the local monitoring client according to the streaming media protocol (RTSP). On the other hand, it can respond to the request of the remote monitoring client by using the streaming media protocol. Stream VR panoramic surveillance video media to the remote surveillance client [32]. The remote monitoring client needs to obtain the VR panoramic monitoring video media stream from the streaming media server and render the monitoring video so that the user can view the monitoring video remotely with the effect of virtual reality. The transmission network of the system includes the wired connection from the VR panoramic camera to the local monitoring client, the local monitoring client and streaming media server, and the streaming server and remote monitoring client using a 3G/4G network or WiFi communication wireless connection. The overall system architecture is shown in Figure 2.

At present, there are various video surveillance systems in the market, but the survey found that these video surveillance systems have some disadvantages, mainly in the following aspects. The video surveillance systems are becoming more and more popular, but these Most video surveillance systems require staff to guard them 24 hours a day. In the future,





**FIGURE 2.** Intelligent video surveillance system framework.

video surveillance systems will be used on more occasions. If each surveillance system has a dedicated person on duty 24 hours a day, it will consume a lot of manpower, material, and financial resources. Most of the current video surveillance systems are unable to intelligently analyze and process surveillance video images, and cannot detect suspicious targets in time. Since the video surveillance system cannot provide intelligent early warning, it is prone to fail to detect abnormalities in time due to the negligence of the surveillance staff. At present, the surveillance video collected by the video surveillance system is 2D video, and there is no 3D video intuitive, which is not conducive to an intuitive and accurate understanding of the monitoring area.

To completely cover the surveillance area, a method of using multiple surveillance cameras to collect environmental images from different angles is often used. However, this method not only increases the cost of the system but also increases the complexity of system installation. In summary, the current video surveillance system still has some problems to be solved urgently. To better meet the user's functional requirements, an intelligent video surveillance system based on virtual reality technology designed in this topic improves the functions of the current video surveillance

system. By adopting the VR panoramic camera, the user can intuitively and accurately obtain the situation of the surveillance area without blind spots, and process and analyze the surveillance video. When a suspicious target enters the surveillance area, it can be intelligently warned without needing to be guarded 24 hours a day. At the same time, the system has the characteristics of real-time and stability. If the network allows, users can remotely view surveillance videos through Android devices. From the user's point of view, the video surveillance system in this topic is mainly divided into two parts: local monitoring and remote monitoring. Local monitoring provides users with real-time monitoring, monitoring video management, and intelligent monitoring functions. The monitoring allows users to view surveillance videos in real-time through Android mobile devices anytime, anywhere with the effect of virtual reality, and at the same time receives alarm emails sent by local monitoring clients.

The design of the surveillance video management module is divided into the UI layer and file operation layer. The UI layer is responsible for the application interface, including surveillance video file display controls and operational controls. To visually display surveillance video files, the display controller selects ListView. The file operation layer uses the

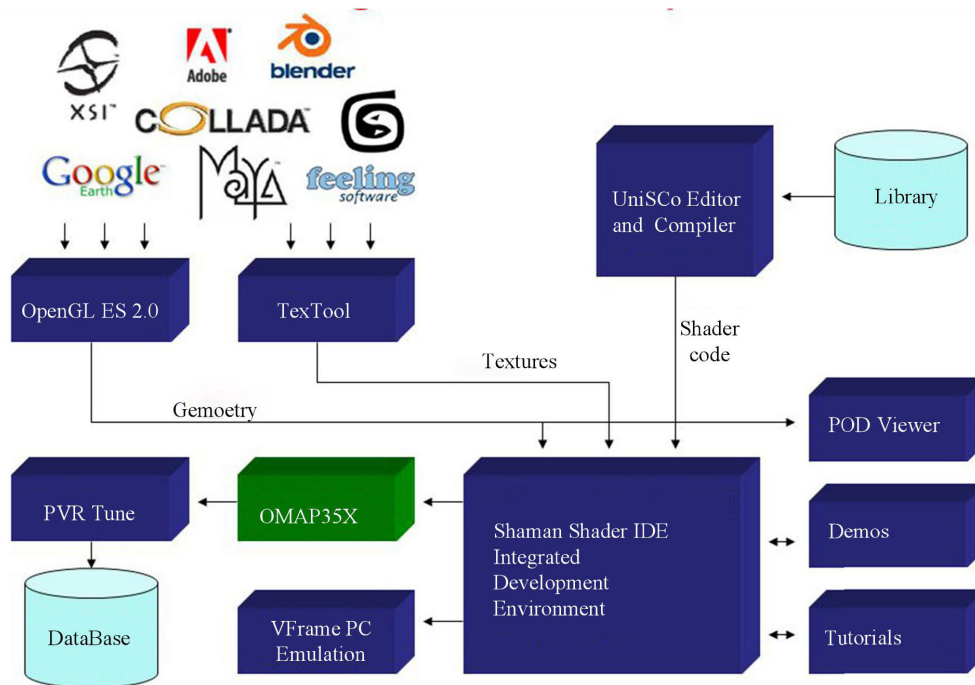


FIGURE 3. Function call flow chart.

FileAPI of the android system to obtain and delete recording files. The user opens the application software installed on the remote monitoring client, enters the URL corresponding to the local monitoring client's push, and then obtains the VR panoramic monitoring video from the streaming media server. The user can achieve the effect of virtual reality by selecting the interactive mode and playback mode. There are two interactive modes: Motion made and touch mode. In Motion mode, the monitoring video content displayed on the screen of the remote monitoring device is updated by capturing the movement of the android device or the movement data of the user's body. In Touch mode, users can touch the screen to interact, and control the content displayed on the screen through sliding and zooming operations. There are two playback modes, Normal mode, and cardboard mode. In normal mode, there is only one display area on the screen, and VR glasses cannot be worn. In cardboard mode, the screen is divided into two identical areas on the left and right for display [33]. Users can achieve a stronger virtual reality effect by wearing VR glasses.

According to the workflow design of the remote monitoring client, the functional structure design of the remote monitoring client is shown in Figure 3. It can be divided into three modules, namely the streaming media player module, the VR video rendering module, and the mode selection module.

The Java layer mainly consists of two class libraries, *ijkplayer-java*, and *ijkplayer-Exo*, and *ijkplayer-Exo* relies on *ijkplayer-java*. There are *IMediaPlayer* interface and

*AbstractMediaPlayer* abstract class in *ijkplayer-java*. The underlying JNI to *ijkplayer-java* is based on *replay*. *Fila* uses an event loop to control the process, but *ijkplayer-java* does not use *replay*'s event loop, but a message queue, but the essence is the same. The *ijkPlayer* function call process is shown in Figure 3.

The process of using OpenGL ES to achieve VR video rendering is as follows. The streaming media player module first obtains the VR panoramic monitoring video media stream data from the streaming media server and decodes it. The decoded data are transferred to the surface, and *SurfaceTexture* captures each frame of the image stream. One frame is used as the texture of OpenGL ES, that is, the texture, and then the texture is deformed and rendered, and after the rendering is completed, it is passed to *GLSurfaceView* for display. VR panoramic video has a higher degree of freedom, and it also makes it easier for users to interact with video content, which greatly enriches the viewing experience. When wearing VR glasses, the lens can increase the user's viewing area. The immersive viewing experience can lead users to enter the virtual reality world more quickly. The mode selection module provides users with different interaction modes and playback modes to realize the sense of virtual reality that VR panoramic video brings users. According to the characteristics of VR panoramic video, the mode selection module is divided into the UI layer, motion interaction layer, touch interaction layer, matrix transformation layer, normal playback layer, cardboard playback layer, and viewport control layer.

### III. ANALYSIS OF MONITORING DESIGN OF TOURIST ATTRACTIONS

#### A. DESIGN OF VIDEO CAPTURE MODULE

The video acquisition module is responsible for the collection of VR panoramic surveillance video data. The camera in the local surveillance client uses Huawei's TYPE-CVR panoramic camera. In the third chapter, the module is divided into the UI layer, camera function layer, and camera preview layer. The UI layer uses SurfaceView as the camera preview. Data collected by the VR panoramic camera are processed by the camera function layer and the camera preview layer and then encapsulated into a bitmap, and then drawn to the canvas owned by the SurfaceView and displayed on the screen. The camera function layer controls the camera-related hardware through the APICamera in the android system. The main methods and functions used are shown in Table 1. To use the camera, you need to declare permissions through the following code.

**TABLE 1.** Methods involved in camera functional layer.

Equipment name	Values	Access method
Camera	3	WIFI/LAN
Smoke/combustible gas/Co gas, etc.	3	LORA
Air conditioning controller	4	RS485
Thermometer	3	LORA
Door sensor detection	2	LORA
Lighting control	2	RS485
Curtain control	3	RS485
Access control equipment	4	RS485

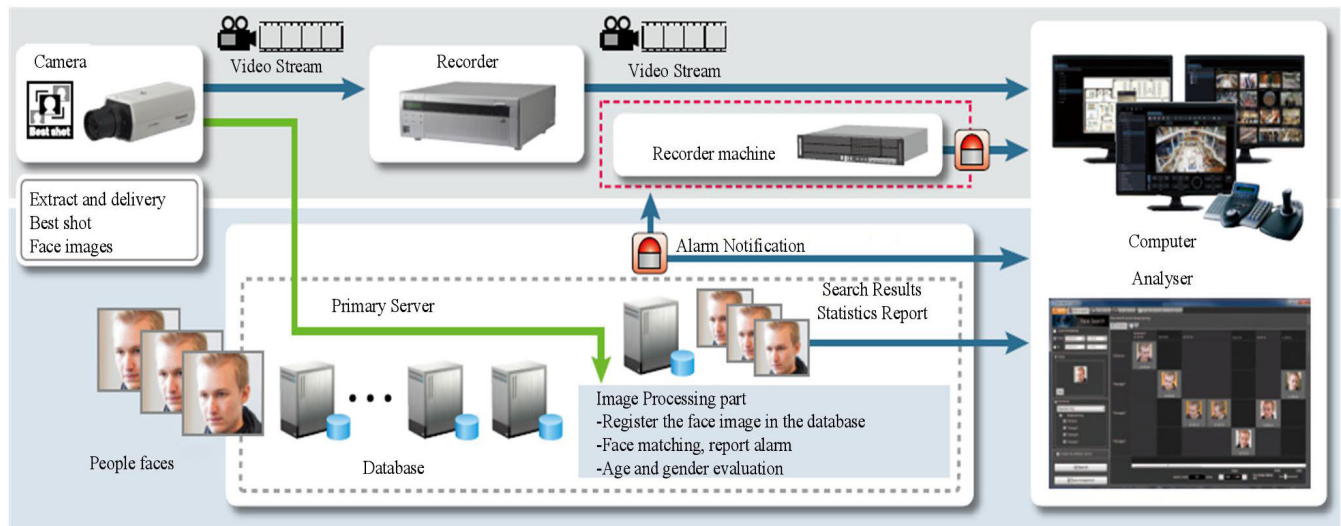
The intelligent monitoring module mainly processes and analyzes monitoring video data to determine whether there are moving targets in the monitoring area. OpenCV is a popular vision library in the development of computer vision applications. It consists of a large number of C functions and a small number of C++ classes and integrates a large number of general algorithms for digital image processing, which can quickly solve digital image processing problems. After writing the C/C++ code, we use VisualStudio to verify the code. It fixed the camera in a corner of the laboratory. The original color scene images collected in the laboratory are shown in the figure. The screen when the surveillance area is moving and an intrusion target moves in the surveillance area. Then the program will acquire the collected images frame by frame, and perform grayscale and matrix processing on the foreground image and background image, and extract the foreground image and background image to determine whether there is a target intrusion. The foreground map of the non-intrusion target is moving in the monitoring area and the intrusion target moving in the monitoring area obtained through the program processing.

The point with a value in the foreground image matrix format corresponds to the pixel points on the contour of the target that broke into the monitoring area and the points affected by factors such as light changes in the monitoring area. Therefore, traverse the foreground image matrix to count the number of non-zero pixels in the matrix, and then compare the number of pixels with the set threshold. If it is greater than the threshold, it means that a target has entered the monitoring area and alarm immediately. If it is less than the set threshold indicates that non-zero pixels are caused by environmental factors. When a suspicious target is detected to break into the monitoring area, the mail sending class is called to send an alarm email to the designated mailbox, and the current frame and the previous frame image are attached to the email, and the alarm email is sent using the package.

When the recording function is turned on, the method of the SDK will be called, and the local surveillance client will start recording and save the surveillance recording files in the reader folder. To browse the video files intuitively, the layout adopts the ListView control and then implements a custom adapter by inheriting BaseAdapter. When inheriting BaseAdapter and rewriting getView, you need to load the ListView layout file. The layout file includes a TextView control and a CheckBox control. TextView control is used to display the surveillance video file name, and the CheckBox control is used to select the file to be deleted. The control is invisible by default state, when the item of ListView is long pressed, the control changes from the invisible state to a visible state. To provide data for the layout file, construct an ArrayList object fileList to record the video files in the reader folder. Each element in the fileList is of map type, the key value is the video file ID, and the Value is the video file name.

The binarized image obtained by segmenting the moving target is shown in Figure 4. The background object refers to the static or very slowly moving object, and the foreground object corresponds to the moving object. Therefore, object detection can be seen as a classification problem, that is, to determine whether a pixel is a background point. Experiments show that our method has a much better detection effect and calculations. The amount for it is very small and the memory footprint is small, which makes the method can be used for embedded systems cameras.

Based on the single-stage detection network structure, RetinaFace adopts a multi-task learning strategy and executes four detection branches to simultaneously predict the face classification value, face candidate frame, five face key points coordinates, and the three-dimensional position of each face pixel. Among them, RetinaFace uses the existing face classification loss function and face candidate frame regression loss function to predict the face classification value and face candidate frame. It used additional supervision information and face key point regression loss function to achieve face alignment. On the dense regression branch, a self-supervised learning method is used to convert two-dimensional face information into three-dimensional face information [34]. Secondly, a grid decoder based on graph convolution



**FIGURE 4.** Four monitoring detection branches.

is used to map three-dimensional face information into two-dimensional face information. The dense regression loss function is used to compare the position differences of five facial feature points between the original image and the coded image.

### B. SYSTEM TEST DESIGN

When the database and collaborative filtering algorithm, we also need to show through the pages, so this paper describes the design meter push shows the main flow of the page. Web design can be divided into three layers for development: presentation layer, logic control layer, and the data layer. The data layer creates transmission objects that can be transmitted to the database. That creates a class corresponding to the form in the database. The logic control layer is mainly to edit the operation process of how to store, update, and delete data, and set the return value or threshold to make logical judgments. Finally, the presentation layer feeds back the processed data to the user by obtaining the data and parses the corresponding HTML. JavaScript and other front-end code to render the page, and finally presents it to the user. When the front end interacts with the user, a visual page is displayed, and the user operation page will send a request to the server. The background will call the service layer level by level according to the designed relevant business logic for processing, and call the corresponding method. When a series of operations such as adding, deleting, modifying, and searching data needs to be performed, such as saving or deleting collected data such as a user's interest degree parameter or character shape, the database is called to complete. When a response is returned after the background is completed, the browser gets the HTML code, parses the HTML code, requests the resources in the code, and renders the page and finally presents it to the user.

The best of the intelligent display system is to build an environment by simulating a small booth and perform functional tests according to the exhibitor's participation process. According to the design block diagram of the overall scheme of the display system, the tests are performed according to the functions divided by the modules. Test from the following three aspects. The first is the recognition and perception module based on the Internet of Things. The first step is to test the realization of the designed APP function. The second step is to test the reading and writing function of the exhibition card with UHF tags, which requires the required hardware equipment such as readers, displays, antennas, tags, coaxial cables, mouse, and keyboard. The third step is to test the performance of the sensor.

After the local monitoring client, the remote monitoring is connected to WiFi, and open the local monitoring client software and click the open remote monitoring button. The remote monitoring client monitoring screen is shown in Figure 5. The local monitoring client can choose to include pure video, audio, and video. The resolution of the surveillance video can be selected from low resolution, medium resolution, and high resolution. The local monitoring client will transmit the monitoring video to the streaming media server after enabling the remote monitoring. The remote monitoring client can obtain the monitoring video from the streaming media server in real-time by accessing the specified URL. It can be seen from Figure 5 that both the local monitoring client and the remote monitoring client of the video monitoring system implemented in this project can be monitored in real-time, reaching the standard of the video monitoring system and having certain practical application value.

Due to a large amount of interaction with memory and network transmission, the serialization process must be used. Because Java's native serialization mechanism will cause a lot of data redundancy and the serialization time is



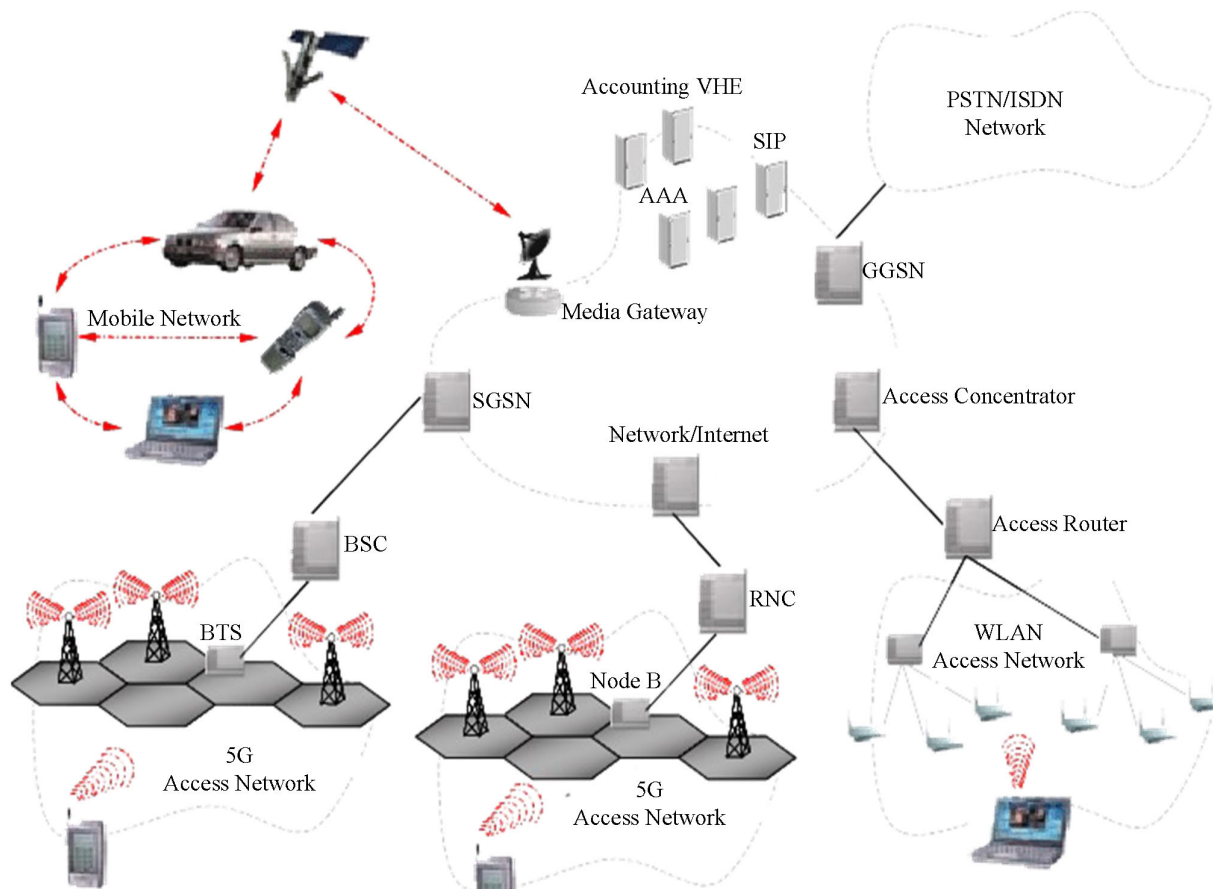


FIGURE 5. Test process design.

relatively long, so the serialization framework is used to solve this problem. This article uses two serialization frameworks, Avro and Prorostuff. Among them, the afro is a framework commonly used by many frameworks in the Hadoop ecosystem and has better compatibility when communicating with hides and base. Prorostuff is independent of language and platform, has a small footprint, is fast, and is suitable for data storage or RPC data exchange format. Serialize the data are divided into two parts: on the STDataBean serialization process and to STHBaseBean serialization process. When writing STDataBean, based on the serialization framework Prorostuff, the property structure information encapsulated in STDataBean is serialized using a serializer and stored in the local cache LocalCache. When writing or reading data, information related to the STDataBean in the local cache is destabilized to ensure that there is always a global schema to control the information of the table structure.

#### IV. RESULTS AND ANALYSIS

##### A. ANALYSIS OF SYSTEM PERFORMANCE TEST RESULTS

When performing a delay test on the system, you must first ensure that the network is in good condition and try to ensure exclusive bandwidth. We download time on Baidu on the web page. The current time will be displayed on the web page, and

then open the local monitoring client to start streaming. The content collected by the camera is the web page displaying the current time. It opens the remote monitoring client and plays the video collected by the local monitoring client. The test process is shown in Figure 6. It tested 100 times and then remotely monitor the delay. The unit is second, and then the weighted average of the measured delay data is calculated as the system delay, and the weighted average of the delay

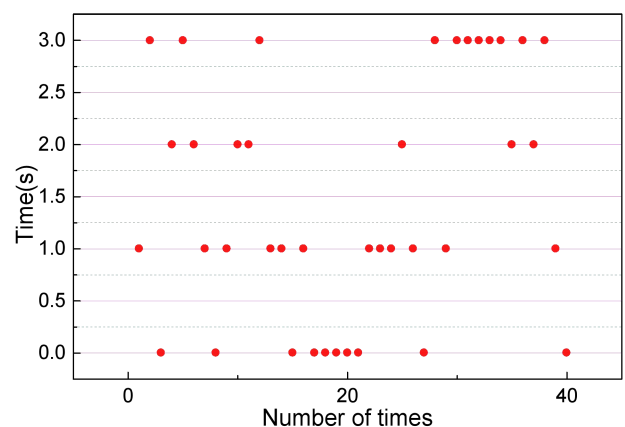


FIGURE 6. System delay result.

is one second, meeting the remote monitoring delay in the performance indicators does not exceed 2 seconds.

The mobile target detection test of the system is mainly to test whether the local monitoring client can accurately identify whether a mobile target breaks into the monitoring area. The test content putS the monitoring system in three different environments. The numbers of the three scenes are scene 1, scene 2, and scene 3. Then there are no people and people in the monitoring area at slow, medium, and fast speeds. When moving in the monitoring area, the local monitoring client collects 100 frames of monitoring video for each moving speed and detects moving targets on it. Finally, it is judged whether the detection result of the system is consistent with the actual situation. According to the log, you can get as shown in Figure 7. The test results show that the success rate of moving target detection meets the performance index, the success rate of moving target detection is not less than 90%.

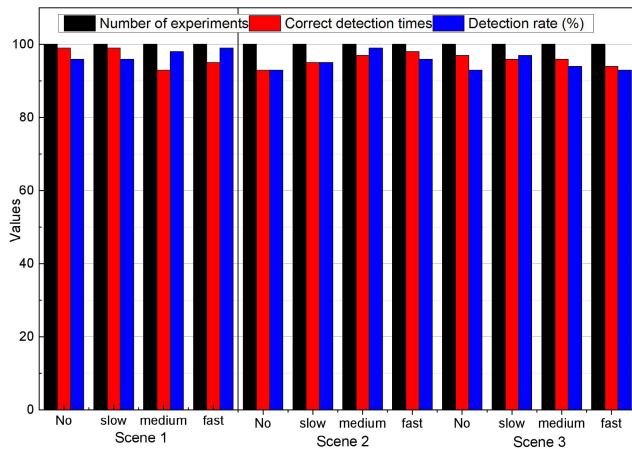


FIGURE 7. Moving target detection test.

The software system compatibility needs to be considered when designing the software. To verify that the VR panoramic player implemented by the remote monitoring client using ijkPlayer and OpenGL ES in this subject is compatible with mainstream Android versions, We run the remote monitoring client software on Huawei Mate7 and use the GT test tool to test the VR players VRMonitor. Among them, Storm Mirror VR and Orange VR are the top two VR players in the android application market by downloads. The GT test tool tests VRMonitor, Storm Mirror VR, and Orange VR respectively as shown in Figure 8. The test results of CPU and memory consumption are showed in Figure 8, where min, max and represent the minimum and maximum values respectively and average. The test results show that the VR player implemented in this topic is compatible with mainstream Android versions, and compared with mainstream VR players on the market, although it consumes slightly more memory, it greatly reduces CPU overhead.

Based on the above test results, a comparison table of design indicators and test results can be obtained. Figure 8

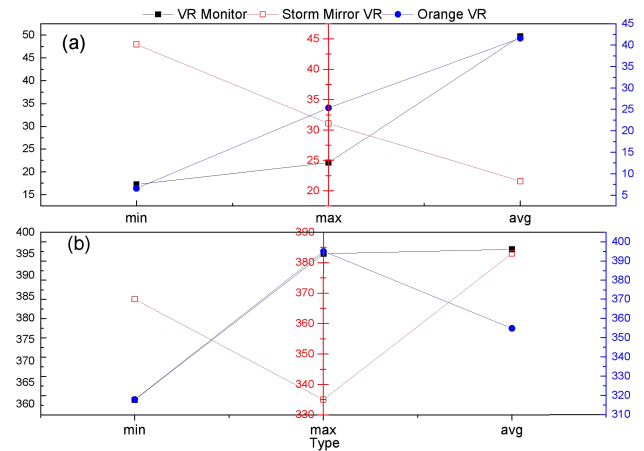
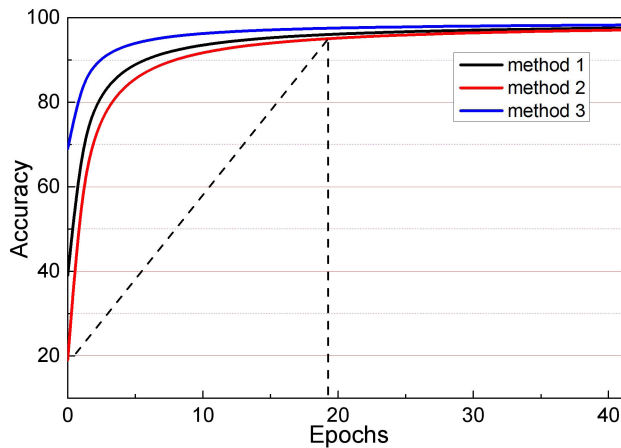


FIGURE 8. VR player performance comparison.

verifies that the local surveillance client can collect VR panoramic surveillance video in real-time and push the collected surveillance video to the streaming server. The remote surveillance client can obtain the VR panoramic surveillance video from the streaming server for real-time remote monitoring. It verifies that the video surveillance system can intelligently detect whether there are suspicious targets in the surveillance area and provide early warning, which meets the functional index. It can be seen from Figure 8 that the remote monitoring client can normally play VR panoramic monitoring video, support split-screen playback, user motion tracking, and touch-screen interaction, etc., and meet the functional indicators. The system has a high success rate of mobile target detection in different environments. It is lower than 90%, and the faster-moving target moves, the higher the accuracy rate. When there is no one moving in the monitoring area, there may be misjudgment due to the influence of the environment, and the test result meets the performance index.

## B. POOLING METHOD AND SIZE SELECTION ANALYSIS OF MONITORING ACCURACY RESULTS

The multi-pose face recognition comparative experiment includes three experimental methods. Method one is the experimental baseline method. RetinaFace is used for face detection. MobileFaceNets and ArcFace are used for facial feature extraction. The facial feature matching results of each video image frame is independent and is not affected by adjacent video image frames. Adding a face pose estimation method based on the multi-point perspective poses a solution method and a face pose correction algorithm fused with pose information. Method three is the multi-posture face recognition algorithm proposed in this paper. A face tracking method based on adjacent video image frame comparison is added, and a face feature matching method based on face tracking is used to determine people's face feature matching results. The experimental results of the face recognition comparison experiment are shown in Figure 9.

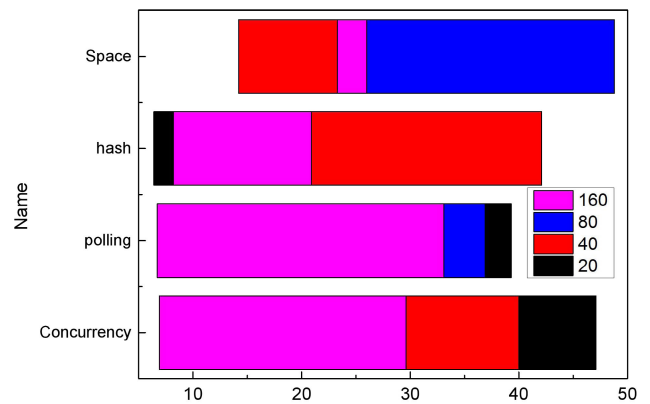


**FIGURE 9.** Experimental results of face recognition comparison experiment.

Comparative Method II and Method apparent from the results of three experiments, matching face based on facial feature tracking of people who can improve the recognition accuracy of a face recognition algorithm. Because of a larger range of motion scenes of video surveillance, face detection obtained through human Face images generally has low resolution, and there are also environmental interference factors such as occlusion and lighting, which cause the feature matching results of some face images to be incorrect. The face feature matching method based on face tracking first uses the face tracking method to obtain the same. The matching results of all facial features of an identifiable individual in adjacent video image frames are compared with the matching results to correct the wrong matching results, thereby improving the recognition accuracy of the multi-pose face recognition algorithm. From the experimental results of the above two comparative experiments, it can be seen that the multi-posture face recognition algorithm for video surveillance proposed in this paper can achieve considerable recognition accuracy in video surveillance scenarios and has high robustness to face pose changes. It can meet the application requirements of video surveillance systems.

We improve the speed of face detection by modifying the minimum side length of the anchor. The face tracking method based on the comparison of adjacent video image frames is introduced in detail to lay the foundation for the facial feature matching method. Subsequently, the face poses an estimation method based on the multi-point perspective poses solution method is introduced in detail to provide post information on the face poses correction algorithm. The facial feature extraction method is based on MobileFaceNet and ArcFace and the facial feature matching method based on face tracking is introduced in detail. Finally, through face recognition comparison experiments, it is verified that the multi-posture face recognition algorithm proposed in this chapter can achieve considerable recognition accuracy in video surveillance scenarios and has high robustness to face pose changes, which can meet the requirements of video surveillance systems.

This experiment verifies the constructed space heat tree algorithm by comparing with the load balancing algorithm using polling, analyzes the dynamic load balancing strategy based on the space heat tree, and whether it is better than the polling algorithm, hash in the spatiotemporal big data application service scenario. The algorithm is more effective. Among them, to get the results faster, the tested data set is taken from the tourist data table of the Bifengxia Scenic Area, containing about 1w pieces of data. It takes the average response time of all requests in each group as the reorganization result value. The experimental results are shown in Figure 10.



**FIGURE 10.** Comparison of load balancing algorithms experimental results.

According to the experimental results, for spatiotemporal big data scenarios, when the amount of concurrency is small, the performance of the two algorithms is comparable. When the amount of concurrency gradually increases, the load balancing algorithm based on space heat can effectively balance high concurrent requests. When the server-side load is 160 concurrent requests, the performance of the load balancing algorithm based on space heat is improved by about 10.18% compared with the polling method.

Since data is mainly stored on HDFS, for frequently accessed data, all or part of it is stored on a storage medium SSD with higher access performance to improve write performance. For data that is rarely accessed, it is stored on archive storage media to reduce its storage cost. Here, hot and cold data storage directories are divided in advance, the corresponding StoragePolicy is set, and the corresponding subsequent programs write data in the corresponding classification directories, and automatically inherit the storage policy of the parent directory. According to the creation time of the data, the hot and cold data are distinguished. The more historical data, the lower the probability of being queried and analyzed. To see the effect easily, the time is divided into one day, timed tasks are set, and data migration is performed using scripts. The experimental results are shown in Figure 11. The historical data with a data volume of 1G is dumped into the cold storage. Based on the billing method, the data only needs to be written to the hot storage. Through the data

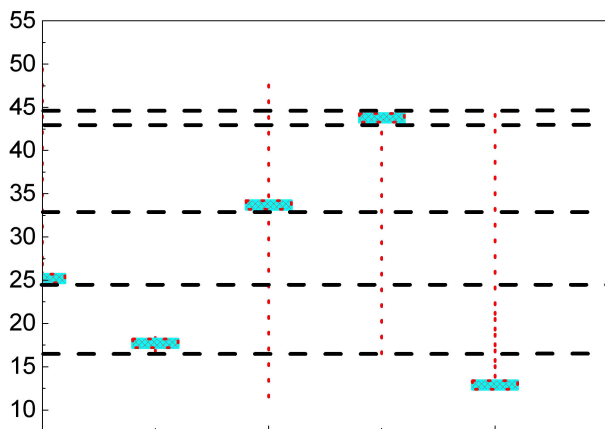


FIGURE 11. Results before and after migration.

subscription method, the data are incrementally written to cold storage, but the dialog is also generated when data is deleted. Using the mark removal method on a secondary basis and selecting data removal can effectively improve the performance of data separation. As shown in Figure 11, this scheme has a drawback that needs to increase called mark2Del field, before and after each data migration field is troublesome.

### C. OPTIMIZATION OF LAYER SELECTION

The item-based collaborative filtering algorithm mainly calculates the number of users interested in an item based on the user-item correlation matrix. When the number of users of a certain item is close, the similarity will also tend to be close. However, this formula does not combine the user's degree of interest. Maybe the user is interested in two items, but the degree will be different. There are too many items calculated recommended parameters for item 32, item 45, and item 150 are the same. Comparing the results of the unmodified code, it is found that the calculation results of the recommended parameters have changed due to the introduction of the degree of interest, and the ranking of the recommended items has also changed accordingly, so the recommended items are also different. The result data after running show that the item-based collaborative filtering algorithm plays a role in recommending items. Since 70% of the data is used as training in the testing phase of the recommendation algorithm, the result verification at this stage will use the remaining 30% of the data to complete the data verification. Top-N recommended classification usually uses two index parameters to measure, the recall rate and accuracy rate. First, the value of  $N$  needs to be determined, and different values of  $n$  are used to test the impact on the accuracy of the algorithm, to perform an initial range sorting the entire data set. As shown in Figure 12, when  $n$  takes different values, the improved algorithm is a simulation diagram of the accuracy of the improved algorithm without the number of users.

From the above results, it is found that when we take a relatively small value, as the number of users increases, the increase in accuracy tends to be stable. As the value

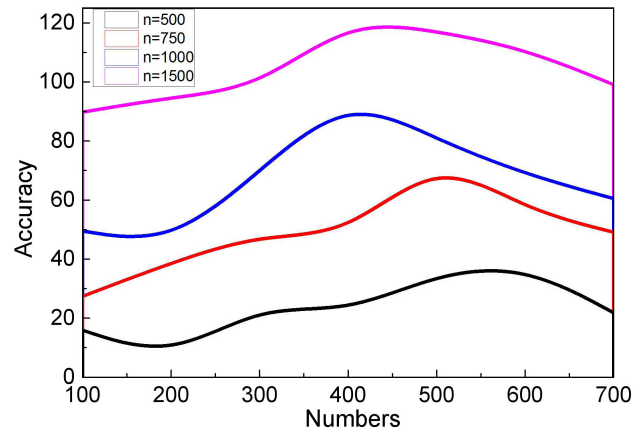


FIGURE 12. Video accurate simulation diagram under different values.

of  $N$  gradually increases, the accuracy rate also increases, but the volatility and randomness of the accuracy rate also increase. Therefore, in comprehensive consideration, when the value is 1000, compared with other values, not only the accuracy rate is relatively high, but also relatively stable as the number of users increases. Therefore, the value of  $N$  in this article is 1000. The recommendation result is achieved by the improved algorithm after the introduction of interest as the weight has been significantly improved in the two reference indicators of accuracy and recall. The accuracy rate is increased by about 5% and the recall rate is increased by about 3%. The abscissa of the figure is the number of users, which increases as the number of users in the test increases, and gradually tends to a stable value. To verify the rationality of the two indicators, the above-mentioned reference indicator is based on the top- $n$  recommendation strategy, that is, 1000 recommended values to compare with the test set. The scope of the two sets is very wide, and it is just an initial ranking stage. The recommended accuracy results obtained through simulation are shown in Figure 13.



FIGURE 13. Recommended accurate simulation results.

As shown in Figure 13, it can be observed that the accuracy rate of the improved algorithm is about 75%, that is, about 7.5 items out of the 10 items recommended to the visitors are recommended correctly. In the beginning, because the



number of users was 100, the accuracy rate was low. As the number of subsequent users gradually increased, the accuracy rate gradually increased, and finally tended to a stable trend. The accuracy of the improved recommendation algorithm is higher than that of the original algorithm.

The video capture device is installed in the air with a ground height of 5.1 meters, and its optical axis is perpendicular to the ground. The program runs without error, and the screen displays normally. The video acquisition device displays the movement status of all moving targets on the road under normal traffic conditions and draws the movement trajectory, detects the movement speed of the moving target, and submits it to the system for analysis.

Comparative study of the advantages and shortcomings of various passenger flow collection technologies, and determined the technical route of adopting image analysis technology. Although the algorithm is more complicated and has certain requirements on the angle and performance of the equipment, it has strong environmental adaptability, high accuracy, and promotion. According to the characteristics of passenger flow data, this paper implements and trains the algorithm from customer identification, area detection, and feature extraction.

## V. CONCLUSION

Based on optimized virtual reality technology to realize intelligent video surveillance of tourist attractions, this paper proposes implementing an intelligent video surveillance system based on virtual reality technology on the android platform. This system has the advantages of intuitive and accurate viewing of surveillance video, no surveillance blind spots, good user experience, and ability intelligent analysis of the characteristics of surveillance video has great practical application value in the field of security. We designed and implemented an intelligent video surveillance system based on virtual reality technology. The system is combined with virtual reality technology to provide users with a new experience of viewing surveillance video. It designs OpenCV and android to realize the moving target detection algorithm on the android platform to achieve the effect of intelligent monitoring. According to this solution, we designed and implemented the function of remote monitoring client to view VR panoramic monitoring video. The VR panoramic monitoring video player in the remote monitoring client can provide users with different interaction methods and playback modes. It performed a functional test and a performance test on the system, show the test results, and some test screenshots, the test results are consistent with the design indicators.

## REFERENCES

- [1] J. H. Park, C. Lee, C. Yoo, and Y. Nam, "An analysis of the utilization of Facebook by local Korean governments for tourism development and the network of smart tourism ecosystem," *Int. J. Inf. Manage.*, vol. 36, no. 6, pp. 1320–1327, Dec. 2016.
- [2] G. Tieri, G. Morone, S. Paolucci, and M. Iosa, "Virtual reality in cognitive and motor rehabilitation: Facts, fiction and fallacies," *Expert Rev. Med. Devices*, vol. 15, no. 2, pp. 107–117, Feb. 2018.
- [3] A. Huertas, A. Moreno, and T. H. My, "Which destination is smarter? Application of the (SA) 6 framework to establish a ranking of smart tourist destinations," *Int. J. Inf. Syst. Tourism (IJIST)*, vol. 4, no. 1, pp. 19–28, May 2019.
- [4] X. Yang, L. Lin, P.-Y. Cheng, X. Yang, Y. Ren, and Y.-M. Huang, "Examining creativity through a virtual reality support system," *Educ. Technol. Res. Develop.*, vol. 66, no. 5, pp. 1231–1254, Oct. 2018.
- [5] J. D. Hemanth, U. Kose, O. Deperlioglu, and V. H. C. de Albuquerque, "An augmented reality-supported mobile application for diagnosis of heart diseases," *J. Supercomput.*, vol. 76, no. 2, pp. 1242–1267, Feb. 2020.
- [6] C. E. Mora, J. Martín-Gutiérrez, B. Añorbe-Díaz, and A. González-Marrero, "Virtual technologies trends in education," *EURASIA J. Math., Sci. Technol. Edu.*, vol. 13, no. 2, pp. 469–486, Jan. 2017.
- [7] N. Chung, H. Lee, J.-Y. Kim, and C. Koo, "The role of augmented reality for experience-influenced environments: The case of cultural heritage tourism in Korea," *J. Travel Res.*, vol. 57, no. 5, pp. 627–643, May 2018.
- [8] A. Milosavljević, D. Rančić, A. Dimitrijević, B. Predić, and V. Mihajlović, "Integration of GIS and video surveillance," *Int. J. Geographical Inf. Sci.*, vol. 30, no. 10, pp. 2089–2107, Oct. 2016.
- [9] S. C. H. Li, P. Robinson, and A. Oriade, "Destination marketing: The use of technology since the millennium," *J. Destination Marketing Manage.*, vol. 6, no. 2, pp. 95–102, Jun. 2017.
- [10] N. Elmqaddeem, "Augmented reality and virtual reality in education. Myth or reality?" *Int. J. Emerg. Technol. Learn.*, vol. 14, no. 3, pp. 234–242, Feb. 2019.
- [11] G. Papanastasiou, A. Drigas, C. Skianis, M. Lytras, and E. Papanastasiou, "Virtual and augmented reality effects on K-12, higher and tertiary education students' twenty-first century skills," *Virtual Reality*, vol. 23, no. 4, pp. 425–436, Dec. 2019.
- [12] M. Brengman, K. Willems, and H. Van Kerrebroeck, "Can't touch this: The impact of augmented reality versus touch and non-touch interfaces on perceived ownership," *Virtual Reality*, vol. 23, no. 3, pp. 269–280, Sep. 2019.
- [13] M. G. Hanna, I. Ahmed, J. Nine, S. Prajapati, and L. Pantanowitz, "Augmented reality technology using microsoft HoloLens in anatomic pathology," *Arch. Pathol. Lab. Med.*, vol. 142, no. 5, pp. 638–644, May 2018.
- [14] M. C. T. Dieck and T. Jung, "A theoretical model of mobile augmented reality acceptance in urban heritage tourism," *Current Issues Tourism*, vol. 21, no. 2, pp. 154–174, Jan. 2018.
- [15] R. Eiris, M. Gheisari, and B. Esmaceli, "PARS: Using augmented 360-degree panoramas of reality for construction safety training," *Int. J. Environ. Res. Public Health*, vol. 15, no. 11, p. 2452, Nov. 2018.
- [16] J. Garzón, J. Pavón, and S. Baldiris, "Systematic review and meta-analysis of augmented reality in educational settings," *Virtual Reality*, vol. 23, no. 4, pp. 447–459, Dec. 2019.
- [17] S. Kavanagh, A. Luxton-Reilly, B. Wuensche, and B. Plimmer, "A systematic review of Virtual reality in education," *Themes Sci. Technol. Edu.*, vol. 10, no. 2, pp. 85–119, Dec. 2017.
- [18] Y. A. Sekhavat and H. Zarei, "Sense of immersion in computer games using single and stereoscopic augmented reality," *Int. J. Hum.-Comput. Interact.*, vol. 34, no. 2, pp. 187–194, Feb. 2018.
- [19] A. A. García, I. G. Bobadilla, G. A. Figueroa, M. P. Ramírez, and J. M. Román, "Virtual reality training system for maintenance and operation of high-voltage overhead power lines," *Virtual Reality*, vol. 20, no. 1, pp. 27–40, Mar. 2016.
- [20] B.-J. Sohn, D.-S. Park, and J. Choi, "Attitude confidence and user resistance for purchasing wearable devices on virtual reality: Based on virtual reality headgears," *J. Intell. Inf. Syst.*, vol. 22, no. 3, pp. 165–183, Sep. 2016.
- [21] E. A. O'Connor and J. Domingo, "A practical guide, with theoretical underpinnings, for creating effective virtual reality learning environments," *J. Educ. Technol. Syst.*, vol. 45, no. 3, pp. 343–364, Mar. 2017.
- [22] C. Gomez, S. Chessa, A. Fleury, G. Roussos, and D. Preuveneers, "Internet of Things for enabling smart environments: A technology-centric perspective," *J. Ambient Intell. Smart Environments*, vol. 11, no. 1, pp. 23–43, Jan. 2019.
- [23] A. Darwish, A. E. Hassanien, M. Elhoseny, A. K. Sangaiah, and K. Muhammad, "The impact of the hybrid platform of Internet of Things and cloud computing on healthcare systems: Opportunities, challenges, and open problems," *J. Ambient Intell. Humanized Comput.*, vol. 10, no. 10, pp. 4151–4166, Oct. 2019.

- [24] L. H. Son, S. Jha, R. Kumar, J. M. Chatterjee, and M. Khari, "Collaborative handshaking approaches between Internet of computing and Internet of Things towards a smart world: A review from 2009–2017," *Telecommun. Syst.*, vol. 70, no. 4, pp. 617–634, Apr. 2019.
- [25] A. T. Chatfield and C. G. Reddick, "A framework for Internet of Things-enabled smart government: A case of IoT cybersecurity policies and use cases in U.S. Federal government," *Government Inf. Quart.*, vol. 36, no. 2, pp. 346–357, Apr. 2019.
- [26] G. Misra, V. Kumar, A. Agarwal, and K. Agarwal, "The Internet of Things (not)-a technology analysis and survey on vision, concepts, challenges, innovation directions, technologies, and applications (an upcoming or future generation computer communication system technology)," *Amer. J. Electr. Electron. Eng.*, vol. 4, no. 1, pp. 23–32, Feb. 2016.
- [27] B. Li and Y. Li, "The Internet of Things drives supply chain innovation: A research framework," *Int. J. Organizational Innov.*, vol. 9, no. 3, pp. 71–92, Jan. 2017.
- [28] C. K. Ng, C. H. Wu, K. L. Yung, W. H. Ip, and T. Cheung, "A semantic similarity analysis of Internet of Things," *Enterprise Inf. Syst.*, vol. 12, no. 7, pp. 820–855, Aug. 2018.
- [29] C. T. Lai, P. R. Jackson, and W. Jiang, "Shifting paradigm to service-dominant logic via Internet-of-Things with applications in the elevators industry," *J. Manage. Analytics*, vol. 4, no. 1, pp. 35–54, Jan. 2017.
- [30] R. Basatneh, B. Najafi, and D. G. Armstrong, "Health sensors, smart home devices, and the Internet of medical things: An opportunity for dramatic improvement in care for the lower extremity complications of diabetes," *J. Diabetes Sci. Technol.*, vol. 12, no. 3, pp. 577–586, May 2018.
- [31] H. Lee, O. Na, Y. Kim, and H. Chang, "A study on designing public safety service for Internet of Things environment," *Wireless Pers. Commun.*, vol. 93, no. 2, pp. 447–459, Mar. 2017.
- [32] W. Serrano, "Digital systems in smart city and infrastructure: Digital as a service," *Smart cities*, vol. 1, no. 1, pp. 134–154, Dec. 2018.
- [33] S. N. V. Yuan and H. H. S. Ip, "Using virtual reality to train emotional and social skills in children with autism spectrum disorder," *London J. Primary Care*, vol. 10, no. 4, pp. 110–112, Jul. 2018.
- [34] H. Habibzadeh, K. Dinesh, O. R. Shishvan, A. Boggio-Dandry, G. Sharma, and T. Soyata, "A survey of healthcare Internet of Things (HIoT): A clinical perspective," *IEEE Internet Things J.*, vol. 7, no. 1, pp. 53–71, Jan. 2020.



**JIE HUANG** was born in Xinjiang, China, in 1994. He received the bachelor's and master's degrees from Xinjiang University, in 2016 and 2019, respectively. He is currently pursuing the Ph.D. degree in tourism management with the School of Tourism, Huaqiao University, Fujian, China. He has published seven articles, three of which have been indexed by CSSCI. His research interests include intelligent management of scenic spots and tourism big data.



**ANMIN HUANG** received the Ph.D. degree in management science and engineering from the School of Management, Tianjin University, China, in 2006. He is currently a Professor and a Ph.D. Supervisor with the School of Tourism Management, Huaqiao University, Fujian, China. He is also the Director of the Research Center of Tourism Planning and Scenic Area Development, Huaqiao University. He has published over 60 articles and his work has appeared in leading tourism and geography journals, such as the *International Journal of Contemporary Hospitality Management*, *Scientia Geographica Sinica*, and *Economic Geography*. His current research interests include intelligent management of scenic spots, tourism big data, and virtual reality technology.



**LIMING WANG** was born in Xinjiang, China, in 1994. He received the bachelor's and master's degrees from Xinjiang University, in 2016 and 2019, respectively. He currently works with the Tourism Management College, Xinjiang University of Finance and Economics, China. He has published a total of five articles. His research interests include tourism informatization and scenic spots management.

...