Intelligent Video Surveillance System Based on Cloud Network

Liwei Li, Haibin Xie and Peng Li
College of Intelligence Science and Technology, National University of Defence Technology, Changsha, Hunan, China
Email: 840822463@qq.com

Abstract. There are two shortcomings in traditional video surveillance systems: First, the amount of data is so huge that the processing, analysis, and viewing of massive data requires a lot of manpower, material resources and financial resources. Second, it is difficult to store and transmit video data. Mass video data often needs to be stored and passed to the backstage for analysis, and huge data makes data storage and transmission difficult. In view of the above trouble, this paper proposes a set of intelligent video surveillance system based on cloud network. It contains modules such as face detection, quick retrieval, etc., and uses the cloud network as a medium to transmit data and information. It can complete the automatic detection, capture, enhancement, and coding processing of the human face in the video, search and compare with the background database automatically and quickly.

Keywords. Intelligent video surveillance system; cloud network; face.

1. Introduction
At present, people’s requirements for security are becoming more and more urgent, especially the role of video surveillance system in people’s life. Traditional video surveillance is mostly used for video capture, storage and playback. Limited early warning capabilities for hazards, and it requires people to continuously monitor the video to achieve video surveillance, which consumes a lot of manpower and time. Furthermore, with the increase of data, the traditional monitoring mode can not meet the storage and transmission requirements of massive video data.

Therefore, the intelligent application of video surveillance is widely studied. Intelligent video surveillance adopts machine learning method to automatically understand and analyze video data. At present, there have been many research achievements in intelligent video surveillance technology. Bouwmans et al. [1, 2] introduced target detection techniques according to the perspective of background modeling. [3-5] studied the target tracking method from single-camera and multi-camera tracking. [6-8] introduced target classification and identification. [9-11] make an overview of the behavioral recognition algorithm.

In this paper, we propose a set of intelligent video surveillance system used cloud network.

The physical architecture of the system consists of front-end (camera integrated with smart modules) and back-end (database). The front and rear ends are connected through cloud network, which also serves as a cloud server for storage and computing for large amounts of data. The intelligent module of the system first processes and encodes the detected face, and then compares it with the face code contained in the database. The functional modules of the system are: the intelligent module has the functions of video receiving, intelligent coding, face detection and capture, image enhancement, etc. The database has the function of updating.
The working process of the system is that after the smart module receives the video, it detects, captures, enhances, and encodes the face in sequence. Then transmits the encoding to the background work machine through the network, and compares it with the face encoding. The encoding result is fed back to the camera to trigger the event switch.

Based on the network cloud, the system realizes the functions of automatic face detection, intelligent coding and fast retrieval, significantly enhances the intelligent level of video monitoring, saves the hardware resources of data storage, and is conducive to the upgrading of intelligent video monitoring system.

2. The Overall Architecture of the Intelligent Video Surveillance System

The execution flow of the intelligent video surveillance system is shown in figure 1. The physical structure of the architecture can be summarized as three parts and a set of systems, namely three parts of front-end, back-end, cloud and intelligent video surveillance network system.

The front end consists of cameras integrated with smart modules. The smart module is the core of the smart video surveillance architecture. It is responsible for the intelligent processing of the video information collected by the camera, including automatic face detection and capture, face quick retrieval and other processes, and has the function of positioning and time recording, thereby retaining the timestamp and location stamp of the detected target face.

The back end is composed of a face database, and the face coding database is obtained through intelligent coding. According to application scenarios of different scales, databases of different sizes can be constructed. In large-scale application scenarios of national networking, massive databases can be built. Unified and centralized storage and management of all databases will bring greater difficulties and challenges. This intelligent video surveillance architecture requires separate management of each database. For example, the suspect face database of a police station in Changsha, Hunan and the suspect face database of the Beijing Municipal Public Security Bureau can be managed and saved separately, but both are within the scope of the intelligent video surveillance network system. The smart module on any camera and any database can interact through the network.

The cloud refers to a cloud server used to store a large number of face codes, and quickly retrieve and compare face codes. With huge storage space and very powerful computing power, it can meet the storage, transportation and fast calculation of massive data.

The intelligent video surveillance network system uses the network as the medium to connect the front-end, back-end, and cloud through the network into a unified whole. In a certain application scenario, all cameras that comply with the system specifications, all face databases that comply with the system specifications, and cloud networks with big data processing and computing capabilities are used as front-end, back-end, and cloud, respectively, to transmit information through the network and data flow, which form an intelligent video surveillance networked system.
The information flow of the system is as follows: The front-end camera continuously receives the video information stream, and the smart module processes it in real time, performing face detection continuously. When the target face is detected, the smart module captures, enhances, and encodes it, and then uploads the encoding to the cloud.

The back-end database can be continuously updated. First, intelligent coding is performed to obtain the coding database, and then the coding database is uploaded to the cloud. After each database updates, the updated database is uploaded to the cloud.

With the powerful computing and storage capabilities, the cloud server in the cloud stores massive amounts of data such as each database code, and compares the target code with the code in the database.

3. The Design of Our System
The design of the intelligent video surveillance system is shown in figure 2.

![Diagram of the intelligent video surveillance system](image)

**Figure 2.** The design of the intelligent video surveillance system.

The intelligent video surveillance system consists of three components: front-end, back-end, and cloud, involving six functional modules, namely video receiving module, face detection and capture modules, image enhancement module, intelligent coding module, and quick search module.

The video receiving module refers to the module that receives and collects surveillance video. This module converts the behaviors and events that occur in the physical world into video information storage, realizing the transformation of event information to analog information. It is the source of surveillance video information, and its performance determines the clarity, real-time, accuracy, and availability of the collected video information. This module usually consists of a camera and an intelligent module integrated on it. The camera model can be freely selected, and the intelligent
module intelligently processes the collected information in real time. Face detection, face capture, image enhancement, smart coding, etc. are all functional modules integrated on the smart module.

The face detection module refers to the functional module that is integrated in the smart module to automatically detect the face in the video, and is mainly responsible for real-time face detection of surveillance video information. The face detection module is the basis of the following work, and the accuracy and speed determine the discriminative and practicality of face coding. This module is mainly composed of face detection algorithms, and the timeliness and accuracy of the algorithms are factors that need to be considered comprehensively.

The face capture module refers to a functional module integrated in the smart module to automatically capture the face detected in the video and standardize it into a uniform size. This module is the follow-up operation of the face detection module, and is an important step in capturing face images. This module is mainly composed of a face capture algorithm, and its main task is to quickly and accurately extract the face to form a face image.

The image enhancement module refers to the functional module integrated in the smart module to enhance the processing of the face image. This module performs enhanced denoising processing on unclear face images to obtain higher-resolution face images, which lays the foundation for subsequent intelligent face coding and fast face retrieval. This module is mainly composed of image enhancement algorithms, and its main task is to enhance face images to obtain clearer images.

The intelligent coding module refers to the functional module integrated in the intelligent module to encode the face image. This article uses Binary Gradient Mode (BGP) and its improvement methods. The intelligence is embodied in the following aspects: First, the image information is expressed by concise coding, which greatly improves the processing efficiency, and reduces the resource consumption in the storage and transmission process; Second, it encodes according to the relationship between adjacent pixels, which is very robust to light intensity; Third, it introduces prior knowledge for face recognition; Fourth, it uses cascaded encoding to extract face depth information. This module is responsible for encoding face images, which is a concentrated reflection of the intelligence of an intelligent video surveillance system. At the same time, it lays the foundation for the subsequent rapid retrieval of human faces.

The fast retrieval module refers to a software module that quickly retrieves a code matching the target face code from a face code database in the cloud. This module takes the target code as the code to be searched, and uses the database code as the search set to quickly retrieve the code to be searched from the search set, thereby determining the identity information of the target face to be searched. The speed and accuracy of this module are its two main considerations, which are mainly composed of fast retrieval algorithms.

4. The Workflow of the Intelligent Video Surveillance System

The workflow of the intelligent video surveillance system proposed in this article is shown in figure 3.

The workflow of the system includes three parts: front-end operation, background operation and cloud operation. Front end operation uses all directional cameras to detect video stream information in real time and transmit it to the intelligent module for real-time processing. Firstly, the face frame in the video information is extracted as the image to be queried; Then, the face is captured and normalized to obtain the image data that meets the code standard; Then, the face image is enhanced by interpolation and filtering; Finally, the output target image is intelligently encoded to obtain a robust target code with strong recognition ability, and transmitted to the cloud through the network.

In the background operation stage, the face database stores a large number of face images of objects of interest, and all images are processed in the same way as the above content for image enhancement and intelligent encoding to obtain the encoding database; then the encoding database is transferred to the cloud through the network.

Cloud operation: after receiving the target code and the code of the face database, the cloud quickly retrieves the code matching the target code from the code database, so as to judge whether the target
object is consistent with the existing face database. Finally, the comparison result is sent back to the front end to trigger the event switch.

![Diagram of intelligent video surveillance system workflow]

**Figure 3.** The workflow of the intelligent video surveillance system.

### 5. Conclusion

In this paper, we propose an intelligent video surveillance system using cloud network, which is described in detail from several perspectives such as overall architecture, system design, and workflow. The intelligent video surveillance system conforms to the trend of the era of big data processing and analysis, and applies innovative modes such as networking and intelligence, which enhances the level of intelligence of video surveillance, and solves the problem of excessive consumption of hardware resources in traditional models. The intelligent video surveillance network system and the three-part model of front-end, back-end and cloud proposed in this paper have broad application prospects.

### References

[1] Bouwmans T, F El Baf and Vachon B 2008 Background modeling using mixture of Gaussians for foreground detection - A survey Recent Patents on Computer Science 1 (3).
[2] Wojek C, Dollar P, Schiele B and Perona P 2012 Pedestrian detection: An evaluation of the state of the art IEEE Transactions on Pattern Analysis and Machine Intelligence 34 (4) 743-761.
[3] Yilmaz A, Javed O and Shah M 2006 Object tracking: A survey ACM Computing Surveys 38 (4) 13.
[4] Wang X 2013 Intelligent multi-camera video surveillance: A review Pattern Recognition Letters.
[5] Yi W, Lim J and Yang M H 2013 Online object tracking: A benchmark Proceedings of IEEE Conference on Computer Vision and Pattern Recognition.
[6] Huang K Q, Ren W Q and Tan T N 2014 A review on image object classification and detection Chinese Journal of Computers 37 (6) 1225-1240.
[7] Andreopoulos A and Tzotsos J K 2013 50 years of object recognition: Directions forward Computer Vision & Image Understanding 117 (8) 827-891.
[8] Zhang X, Yang Y, Han Z, Wang H and Gao C 2013 Object class detection: A survey ACM computing surveys 46 (1) 1311-1325.
[9] Hu Q, Qin L and Huang Q M 2013 A survey on visual human action recognition Chinese Journal of Computers 36 (12) 2512-2524.
[10] Morris B T and Trivedi M M 2008 A survey of vision-based trajectory learning and analysis for surveillance IEEE Transactions on Circuits and Systems for Video Technology 18 (8) 1114-1127.

[11] Aggarwal J K and Ryoo M S 2011 Human activity analysis: A review ACM Computing Surveys 43 (3) 1-43.