Efficient Expression and Deep Analysis Platform of Massive Traffic Video Data

Xisong Dong, *Member, IEEE*, Gang Xiong, *Senior Member, IEEE*, Bin Hu, Fenghua Zhu, *Member, IEEE,* and Zhen Shen, *Member, IEEE*

***Abstract*—Intelligent video surveillance technology has changed the traditional passive reception mode. It can real-time, automatically and intelligently analyze video data to improve its efficiency and reliability. This paper presents a set of efficient expression and deep analysis platform for massive traffic video, which can connect the surveillance cameras from urban transport systems, and perceive the state of people, vehicles, roads and other elements using intelligent visual perception technology, and extract relevant semantic structure information from the non-structured video content and display it in structured text language. Then, its main contents and key technologies are analyzed. Finally, an application case in Guangzhou, China is given. The system can be used for intelligent expression, deep analysis, detection, identification, archiving and management of massive traffic video data, to provide a complete information services for the managers and users.**

I. Introduction

As one of the hottest vocabularies, big data has been applied in many areas [1, 2]. In the field of city security and traffic management, many video surveillance equipment is deployed, which can produce many video data, for example, the number of traffic bayonet data maybe reach one billion or more, the number of data of a face database would be 10 million or more [3-5]. For such a huge amount of data, there are many limitation, so data retrieval speed, data statistics, analysis efficiency, and so on, are needed to be solved by big data technology [6-9].

Massive video presentation and analysis technology can change the traditional manual monitoring and passive query mode in video proceeding field, and can initiatively take real-time, automatic intelligent analysis of video image to improve its efficiency and reliability [1]. Massive video presentation and analysis includes such functions as detection, tracking, classification, identification, behavior analysis and abnormal alarm of moving objects. It can identity vehicles, pedestrians or other objects and label them, and then use the network characteristics to correlate the target tags, effectively analyze the real-time status of the target objects, and perceive all kinds of abnormal events, finally to make a warning to the relevant receptors on the occurrence of abnormal events [10-19].

This paper proposes the main research contents and key technologies of efficient expression and deep analysis platform of massive traffic video data, and gives a case study. The organization of the paper is as follows: Section II gives the main research contents of efficient expression and deep analysis platform for massive traffic video; Section III presents its key technologies; Section IV presents an application case; Finally, the conclusions are drawn out in Section V.

II. Main Research Contents

The system framework of Efficient Expression and Deep Analysis Platform of Massive Traffic Video Data proposed in this paper is consistent with that of Internet of Things (IoT), is shown in Fig. 1. And, its architecture is shown in Fig. 2. It includes three layers:

- Perception layer: including all kinds of video cameras, to acquire traffic video big data.
- Transmission layer: to transmit the collected data from perception layer to the application layer (service center) via fiber, Internet, wireless network, etc.
- Application layer: it mainly involves the expression, analysis, archiving and management of massive video data, and provides visualization of human-computer interaction, to display the vehicle, pedestrian information of video analysis as video clips, pictures, graphics, etc. And, such relevant information as license plate, motorcycle type, color, movement direction, vehicle images, can be recorded in cloud storage platform to facilitate the latter retrieval process.

A. Cloud Computing Platform

Massive video data cannot be analyzed by a small number of decentralized computers. So, cloud computing platform of massive video data expression and analysis is needed (Seen in Fig. 3), to achieve interoperability and value-added services. Integrating massive video data acquired by perception layer into cloud computing platform, to summarize, archive, analyze and mine the valuable information and to display the analysis results of vehicles, pedestrians and other objects in a graphical, the security management, illegal punishment and other applications can be achieved timely and accurately. Cloud computing platform can also achieve network

This work is supported in part by Natural Science Foundation of China (61233001, 61773381, 61533019, 61773382, and 91520301).

X. Dong is with the State Key Laboratory for Management and Control of Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing, 100190, China. (e-mail: xisong.dong@ia.ac.cn).

G. Xiong is with the Cloud Computing Center, Chinese Academy of Sciences, Dongguan, 523808, China. (e-mail: xionggang@casc.ac.cn).

B. Hu is with the with the State Key Laboratory for Management and Control of Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing, 100190, China. (corresponding author to provide e-mail: binhu@ia.ac.cn).

F. Zhu is with the Beijing Engineering Research Center of Intelligent Systems and Technology, Institute of Automation, Chinese Academy of Sciences, Beijing, 100190, China. (e-mail: zhufh@casc.ac.cn).

Z. Shen is with the Qingdao Academy of Intelligent Industries, Qingdao, 266109, China. (e-mail: shenzhen@qaii.ac.cn).

978-1-5386-1526-3/17/$31.00 ©2017 IEEE

management of the device, user rights, equipment expansion, increasing new business or changing business type, and other functions.

B. *Management Platform of Video Detection Data*

The management platform of video detection data integrates data display, data storage, data analysis and forecasting, and other functions. The platform integrates all data perceived by detection layer, stores it in the database, and aggregates, archives and analyzes it, then carries on the centralized, unified and effective management and deep analysis, to provide users with more effective information support.

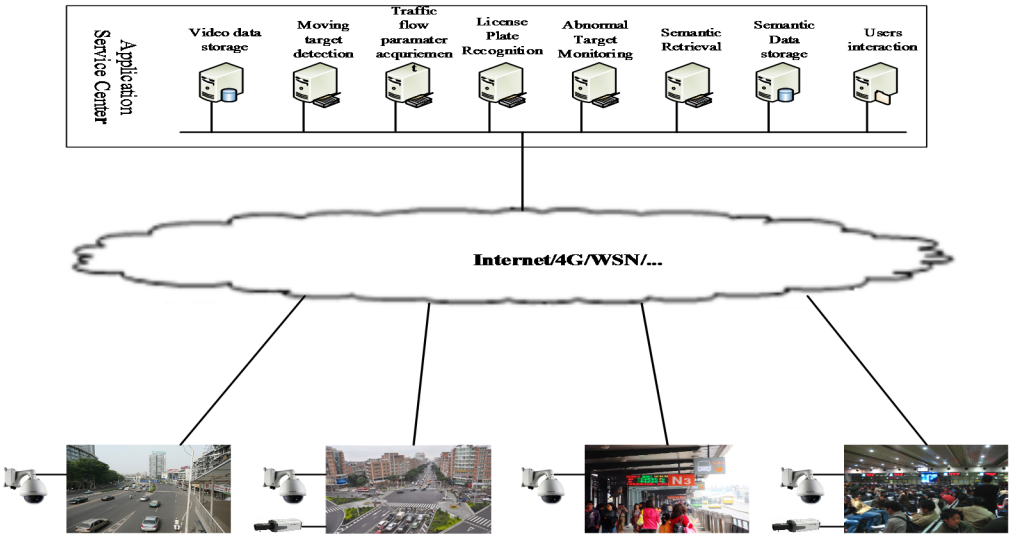

Fig. 1 The system framework of Big Data Platform for Urban Public Transportation

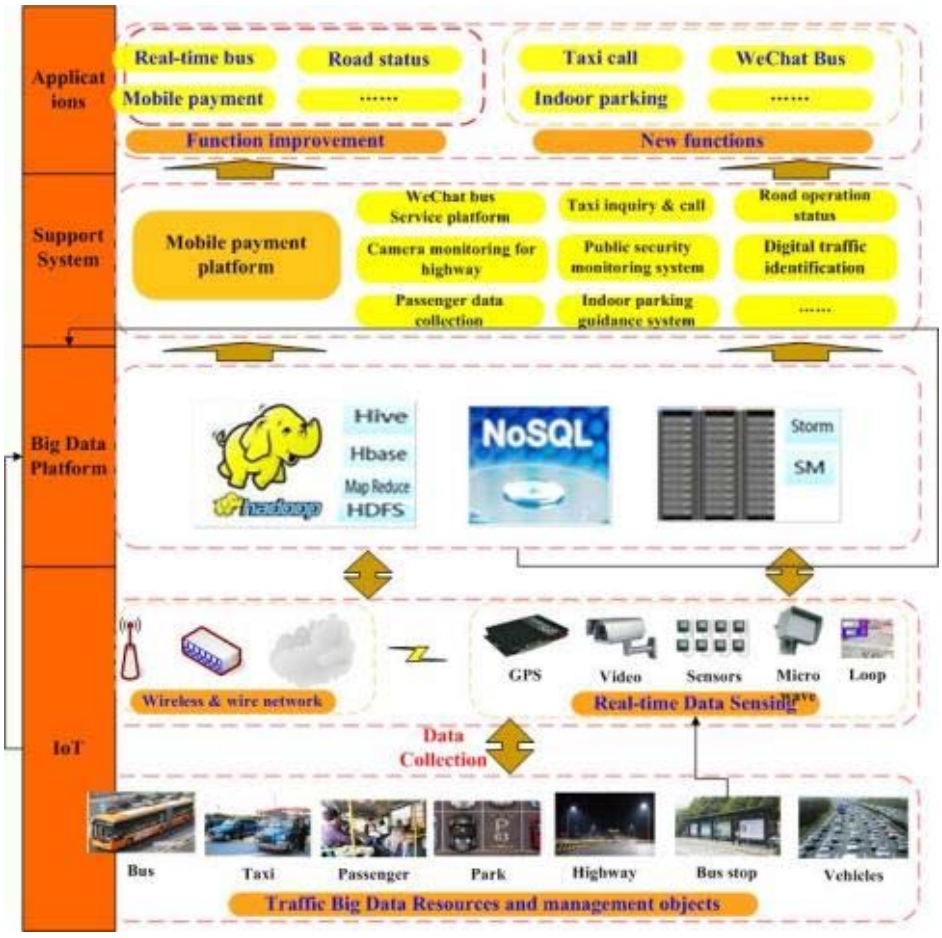

Fig. 2 The system Architecture of Big Data Platform for Urban Public Transportation

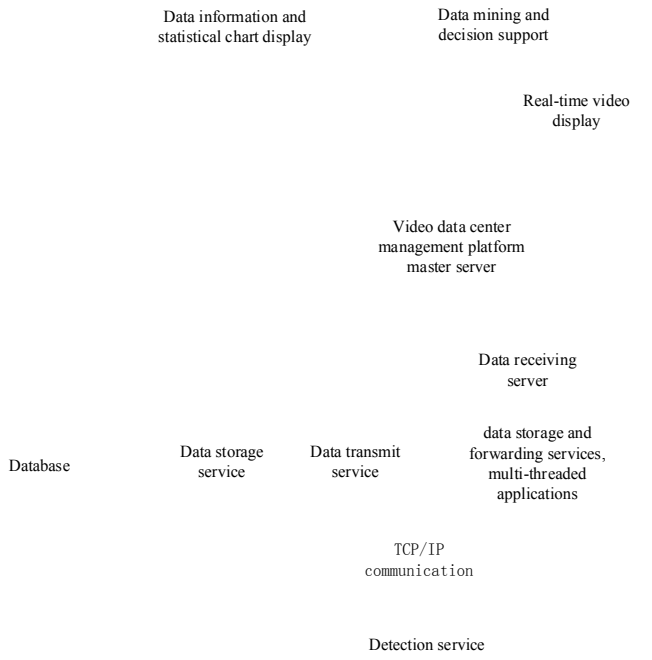

Fig. 3 Cloud Computing Platform

The platform implements a variety of data applications and management functions, including: real-time displaying the vehicle queue length at intersections, license plate number, and so on; showing the statistical analysis via curve, form or graphic; real-time automatic alarming when emergency; generating management programs for illegal processing and abnormal behaviors; providing timely, accurate and complete video detection information in order to carry out scientific and effective traffic management. The platform can also realize such functions as equipment management, users' rights management, video stream management, video storage management, equipment expansion.

C. *Intelligent Video Analysis System*

Intelligent video analysis system can analyze and process historical video and real-time video data using intelligent video processing technology, which includes two aspects: target analysis and behavior analysis. Target analysis includes: real-time detection and tracking of multi-motion targets; human or vehicle target classification based on online scene learning, which can automatically select and annotate similar samples, and so on. Behavior analysis mainly refers to detection and recognition of abnormal behaviors using classification methods.

D. *Moving Target Detection*

Moving target detection is to divide the specific pixel of the moving object from the video sequence, which is the prerequisite for realizing other video analysis (shown in Fig. 4). The actual environment is very complex, such as dynamic background, camera shaking, illumination changing, motion shadows, noise, camouflage, bad weather, which directly determines the accuracy of the target tracking, classification, etc.

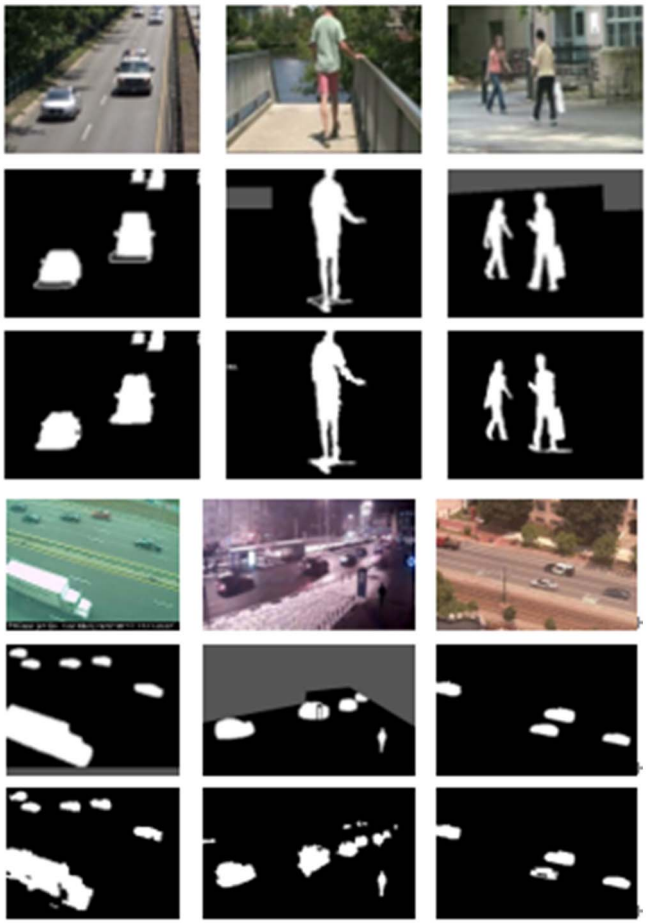

Fig. 4 Moving Target Detection

E. *Traffic Flow Parameter Acquisition*

Massive video resources can be collected by many surveillance cameras installed at roads and intersections to carry out traffic flow parameter acquisition, which can be the complementary or alternative means of coil, microwave radar and other traditional equipment. Through the video analysis, many traffic flow parameters can be collected, including: traffic volume, vehicle speed, occupancy, traffic density, vehicle type, queue length, and so on. The accurate quantification of road traffic flow conditions can be used for urban traffic induction to ensure road safety and smoothness. By analyzing historical laws and trends of traffic flow parameters, traffic management and decision planning for traffic managers can also be provided with reliable data.

F. *License Plate Recognition*

License plate recognition technology is a kind of computer vision technology, which can automatically capture, analyze, recognize, and record the license plate in real time. It has become the core technology for traffic penalty, vehicle retrieval, etc. By cameras installed at roads, intersections, stations and other locations, the vehicles' pictures can be collected automatically and their license plates can be detected and recorded. Based on the license plate number and the vehicle image, the vehicle information can be gotten for the application of parking fee, security and illegal handling.

G. Abnormal Target Monitoring

Using ubiquitous surveillance cameras in ITS, the abnormal targets can be monitored, which is the especial function that such other sensors as coil, laser radar and other traditional detectors do not have. Through video analysis, many abnormal events can be identified, including traffic accidents (vehicle collision, etc.), traffic violations (illegal parking, overspeed, etc.), emergencies (spilled goods, fireworks, etc.), perimeter invasion, overcrowding, and so on, can be identified. And, other types of exception events can be defined according to local regulations. In the application service center, once the system detects the target exception event, the linkage alarm can be immediately triggered and the timely and reliable reference information can be provided to managers. At the same time, the event can be classified and kept. This function is important for shortening the response time of the rescue, to ensure personnel safety and reducing the loss of life and property. In addition, the recorded abnormal event data can provide objective evidence for the event tracing, the fault confirmation of the parties and taking relevant penalties.

H. Semantic Retrieval

Through mass video analysis, the state of human, the vehicle and road can be perceived from mass video data, and the semantic structure information can be extracted and stored as massive semantic data. A user can directly enter the keyword text, such as "query the current location of the vehicle 'xxxxxx" or "query the red car at X-time at X-monitoring point to the east", to search the relevant information. In this cloud computing platform, the system can automatically search and match the user instructions from the semantic database, and send the relevant pictures, video clips to the user. This function will change the traditional passive monitoring and manual search query, to collect automatically text information and intelligence from massive video data, finally to achieve leapfrog development of video surveillance.

III. KEY TECHNOLOGIES

A. High-Accuracy Pedestrian and Vehicle Detection Algorithm

High-quality and high-accuracy pedestrian and vehicle detection is necessary to obtain high accuracy of intelligent video analysis methods, for examples: Deformable Part Model (DPM) is used to locally model the pedestrian and vehicle; the background modeling of video image is carried out by using robust texture description operator; the robust detectors is obtained by training of structural SVM; multiple detectors can be structed for different gestures of pedestrians and vehicles by unsupervised clustering method; using multi-task learning method to enhance the low-resolution pedestrian detection performance; using overall optimization method to improve the pedestrian detection accuracy under the shelter; and so on.

B. Storage and Analysis of Massive Video Data

Large amount of data and real-time communication are key technology that need to be addressed to ensure the correctness and availability of data. The choice of communication strategy is particularly important because of high frequency of data (multiple monitoring sites, possibly hundreds or even thousands of test data per second) and persistence (24-hours uninterrupted operation). Different formats, different coding, and even different transmission of massive video data need different coding techniques to provide algorithm support.

C. Efficient Parallel Computing Algorithm

Intelligent video surveillance system involves larger physical space and information calculation and interaction of vehicles, places, crowds, facilities and other entities, then efficient parallel computing of mass data is the key to achieve efficient interworking of sensor networks and mass video data processing and analysis. The calculated content contains not only complex logical operations, but also a large number of simple mathematical operations.

D. Intelligent Video Analysis Technology under Special Conditions

- Night detection

In the night, the video detector can automatically distinguish vehicles and their light reflection through the analysis of the texture of the detection area. When the detector determines that the sky is dark, it will automatically adjust the texture feature parameters, which can eliminate the impact of vehicle light on the detection accuracy.

- Background disturbances

The video detector has such function as background learning and continuous automatic update. It can adapt to any background disturbances, such as surface water reflection, throwing objects, trees and guardrail shadows. If such interference background exists in the detection area, the system can automatically define them as background objects through background learning.

- Vehicle movement shadows

During daytime, the video detector can automatically distinguish the vehicle and its shadows by analyzing the texture and chroma of the detection area. When the detector is determined as daytime, the texture feature parameters are automatically adjusted. This algorithm will effectively distinguish the motion shadow and texture-rich body of the missing texture, to eliminate the influence of the shadow on the detection accuracy.

E. Quick Video Search and Excavation Technology

Massive video data's quick browsing, retrieval and target trajectory mining system integrates video preview, massive video retrieval and cross-video target trajectory mining technology.

Massive video retrieval requires the following key technologies:

- The static and dynamic attribute extraction and analysis techniques based on static and dynamic features.
- The sample, color, texture retrieval matching technology based on static attribute, and the trajectory search matching technology based on dynamic behavior.

Target trajectory mining requires the following key technologies:

- To form useful metadata information based on video analysis technology.
- To mine and analyze metadata, to extract the target track which the user is interested in.

F. High-speed Robust Background Modeling Algorithm

The environment in surveillance scene is more complex, including lighting change (including moving target shadows), dynamic backgrounds, and so on, which provide challenges for background modeling. The traditional background modeling algorithm based on color information is not robust to the illumination change.

The pixel-based background modeling method has higher time and space complexity, and the background modeling algorithm based on texture cannot distinguish smooth background and prospect. Therefore, a background modeling framework combining color feature and texture feature is needed, which can solve the problem of illumination change and smooth background and foreground at the same time, aiming at the problems of color feature and texture feature.

G. Robust Target Tracking Algorithm

The target representation always has dramatic and complex in background in the monitoring scene, which poses a great challenge to traditional tracking method. A robust target tracking algorithm based on graph matching is needed to be proposed to solve complex target representation, such as Superpixel algorithm, spatial structure diagram model, the context of the target and the background.

H. Video Enrichment Algorithm

Video enrichment is to furthest compress original video length in space and time on the basis of not losing information. The concentration of time aspect means that the video enrichment technology can intelligently discard all frames without motion information. The concentration of space aspect refers to the original objects at different times can be displayed in the same frame of concentrated video. So, high-quality online video enrichment methods, as well as strobe problems of moving objects due to mutual occlusion between the concentrated video, are needed, including background modeling, target detection and tracking, target rearrangement and seamless suture, main background selection, online moving target filling optimization, image mosaic editing, and so on.

I. Target Recognition Algorithm

In the actual monitoring scene, the same target always has great differences in different cameras due to light, target gesture, camera perspective, environment and other factors, which causes a lot of difficulties for the target "ID". The traditional target reidentification algorithm based on metric learning is difficult to achieve high recognition rate in practical application. Therefore, it is necessary to study the method of multi-feature fusion to extract and integrate a variety of local features with strong discriminating ability, and to strengthen the structural information at the pixel-level, blob-level, part-level, to improve the recognition rate, including: to increase the weight of important parts of the matching process using Hierarchical Weighted Histograms (HWH); using bidirectional matching and graph matching method to match the key points (Maximized Stable) Color Region, MSCR; using Gabor Ternary Pattern HSV, GTP-HSV) to improve the accuracy of pedestrian re-identification under the disadvantages of scale, resolution and occlusion, and so on.

IV. Application Case

Based on the traffic big data platform, Guangzhou, China provides a comprehensive traffic information service mobile APP - "XingXunTong (XXT)", which mainly provides such functions as real-time bus service, taxi inquiries and taxi call service, traffic information service, mobile payment service, parking, travel planning, subway, etc. It has integrated more than 700 bus routes and more than 10,000 buses (coverage above 90%). As a result, with XXT, the traffic manager can conveniently grasp the real-time traffic conditions of Guangzhou.

A. Real-time Bus Services

It provides bus lines, transfer information, site location and other information query services.

- Line diagram, shows real-time vehicle position of the current line, electronic bus stops show real-time arrival information of the current site.
- Waiting reminder or arrival reminder: plays a reminder when the next car meets the conditions including departure/arrival site, specific route, reminder condition, and so on.
- Displaying real-time crowded degree of all buses.

B. Traffic Information Services

Real-time monitoring of the city's major traffic intersection can be obtained on the real-time video image of the mobile phone software, and the whole traffic condition of the city and real-time congestion status of main roads can be queried, and it has good enough accuracy rate. Sub functions mainly include road condition inquiries, traffic diagram, traffic ranking, traffic information and micro-blog sharing and so on. The advantages of the service are as follows.

- Covering all the features of the Google map road query, and getting more comprehensive and accurate query information, and more intuitive results.
- Achieving query function of traffic diagram, and showing road map of the trunk roads, which does not have similar function in the existing software.
- Road congestion ranking, not only shows the ranking of congested road sections, the starting point and the end point of each section, but also can be synchronized to the road map to view real-time traffic information.
- Traffic information shows traffic alert in the day.

The acquisition of traffic video information usually uses high definition video camera to monitor motor vehicles, non-motorized vehicles of the road sections in real time and recording the relevant image data with advanced technology of photo electricity, computer, image processing, pattern recognition, remote data access and so on, and it can help to

avoid congested roads, making traffic guidance from passive to active.

Fig. 5 Traffic video display screen

V. CONCLUSION

This article combines video big data, cloud computing and artificial intelligence technology, the content and key technologies of efficient expression and deep analysis platform of massive video data are proposed and researched.

- At the system level, such advanced technology as big data, cloud computing, and artificial intelligence, are used to construct the support platform of high quality expression and deep analysis and comprehensive utilization of massive video data, and provide information support for ITS construction.
- At the technical level, using advanced intelligent technology, the state of the human, vehicle, road and other elements in video surveillance scene can be detected, and the semantic structure information can be extracted from massive video data, and then unstructured video content can be described by structured text language.
- At the service level, the system contains the default query instructions, provides personalized information query service and active information push. The information and intelligence, which users are interested in, can be automatically collected from massive video data, to change the traditional passive monitoring and artificial query.
- At the algorithm level, based on multi-angle learning, probability graph model, depth learning and other advanced intelligent methods to achieve the complex environment can be adapted to the movement target detection, license plate recognition and other video analysis algorithms to ensure that in bad weather, low light, Jitter and other adverse environmental conditions, the system can generate the correct semantic structure information.
- At the application level, based on multi-angle learning, probability graph model, depth learning and other intelligent methods, the video analysis, such as moving object detection and license plate recognition under complex environment, can be realizes, to ensure generate the correct semantic structure information under adverse environmental conditions such as bad weather, low illumination and camera shake. The system not only has such basic functions as traffic flow parameter acquisition, license plate recognition and moving target detection, but also such high-end function as target anomaly monitoring, semantic retrieval, human-computer interaction visualization, information display, and so on.

At present, the platform is undergoing research and development, and the initial results have been applied in video big data processing of public transport in Guangzhou China.

REFERENCES

[1] T. Huang. Surveillance video: The biggest big data. Computing Now, 2014, 7(2): 82-91.

[2] X. Liu, X. Wang, W. Zhang, et al. Parallel Data: From Big Data to Data Intelligence. Pattern Recognition and Artificial Intelligence, 2017, 30(8): 673-682.

[3] Y. Lv, Y. Duan, W. Kang, et al. Traffic Flow Prediction With Big Data: A Deep Learning Approach. IEEE Transactions on Intelligent Transportation Systems, 2015, 16(2): 865-873.

[4] K. Wang, C. Gou, F. Wang. Parallel vision: an ACP-based approach to intelligent vision computing. Acta Automatica Sinica. 2016, 42(10): 1490−1500.

[5] K. Wang, C. Gou, N. Zheng, et al. Parallel vision for perception and understanding of complex scenes: methods, framework, and perspectives. Artificial Intelligence Review, 2017, 48(3): 298−328.

[6] G. Xiong, F. Zhu, X. Dong, et al. A Kind of Novel ITS based on Space-Air-Ground Big-data, IEEE Intelligent Transportation Systems Magazine, 2016, 8(1): 10-22

[7] K. Wang, Y. Lu, Y. Wang, et al. Parallel Imaging: A New Theoretical Framework for Image Generation. Pattern Recognition and Artificial Intelligence, 2017, 30(7): 577−587.

[8] X. Dong, G. Xiong, W. Kang, et al. A Kind of Adaptive Dynamic Transit Signal Priority Control Method, 12th World Congress on Intelligent Control and Automation, 2016: 999-1004.

[9] F. Wang, X. Wang, L. Li, et al. Steps toward parallel intelligence, IEEE/CAA Journal of Automatica Sinica, 2016, 3(4): 345-348.

[10] G. Xiong, F. Zhu, X. Liu et al. Cyber-physical-social System in Intelligent Transportation. IEEE/CAA Journal of Automatica Sinica, July 1, 2015, 2(3): 320-333.

[11] X. Dong, J. Zhou, B. Hu et al. A Framework of Future Innovative Urban Transport, 19th International IEEE Conference on Intelligent Transportation Systems, 2016: 19-23.

[12] G. Xiong, D. Shen, X. Dong, et al. Parallel Transportation Management and Control System for Subways, IEEE Transactions on Intelligent Transportation Systems, 2017, 18(7): 1974-1979.

[13] X. Dong, Y. Lin, D. Shen, et al. A Parallel Transportation Management and Control System for Bus Rapid Transit using the ACP Approach, IEEE Transactions on Intelligent Transportation Systems, 2017, 19(8): 2569 – 2574.

[14] L. Li, Y. Lv, F. Wang. Traffic Signal Timing via Deep Reinforcement Learning. IEEE/CAA Journal of Automatica Sinica, 2016, 3(3): 247 -254.

[15] X. Dong, G. Xiong, D. Fan, et al. Research on Bus Rapid Transit (BRT) and its Real-Time Scheduling, 6th IEEE International Conference on Service Operations and Logistics, and Informatics, 2011: 342-346.

[16] X. Wang, X. Zheng, Q. Zhang, et al. Crowdsourcing in ITS: The State of the Work and the Networking, IEEE Transactions on Intelligent Transportation System, 2016, 17(6): 1596-1605.

[17] L. Li, Y. Li, D. Cao, et al. Parallel Learning - A New Framework for Machine Learning. Acta Automatica Sinica, 2017, 43(1): 1-8.

[18] F. Wang. Parallel Control and Management for Intelligent Transportation Systems: Concepts, Architectures, and Applications, IEEE Transactions on Intelligent Transportation Systems, 2010, 11(3): 630-638.

[19] N. Zhang, F. Wang, F. Zhu, et al. DynaCAS: Computational Experiments and Decision Support for ITS, IEEE Intelligent Systems, 2008, 23(6): 19-23.