Semantic Analysis of Traffic Video Using Image Understanding

Jian Wu[1,2]
1.The Institute of Intelligent Information Processing and Application, Soochow University, Suzhou 215006, China;
2.Jiangsu Yihe Traffic Engineering Co., Ltd., Suzhou 215002, China
Email: szjianwu@163.com

Zhi-ming Cui[1,2], Heng-jun Yue[1,2], Guang-ming Zhang[1,2]
1.The Institute of Intelligent Information Processing and Application, Soochow University, Suzhou 215006, China;
2.Jiangsu Yihe Traffic Engineering Co., Ltd., Suzhou 215002, China
Email: szzmcui@suda.edu.cn

***Abstract*—Video retrieval technology has always been the important application domain of image engineering. Aiming at the problem that low-level features cannot describe high-level semantic completely and accurately in the content-based video retrieval technology, this paper introduces the methods of image understanding and proposes a video semantic analysis framework using scene analysis for recognizing some semantic events in the traffic video. The experimental results show that the feasibility of the framework is good and the event recognition method based on scene analysis has good recognition effect for vehicle status in traffic video.**

***Index Terms*—video retrieval; semantic analysis; image understanding; scene analysis**

I. INTRODUCTION

With the development of society economy and the increasement of living level, vehicle turns into the necessary means of transportation travel. With the increasing number of vehicles, traffic problems are also becoming more serious, traditional television monitoring technology can not meet the requirements of the current traffic management. In recent years, intelligent transportation system (ITS) becomes a new generation of traffic management systems which is based on many new technologies, such as computer network, video transmission, image processing, video processing and computer vision. ITS is applied in many scenarios, including airports, stations, passenger flow, highway intelligent scheduling, operations management, vehicle scheduling. To a certain extent, it improves transportation efficiency, relieves traffic congestion and improves road capabilities. Effective control of the vehicle can reduce traffic accidents, so as to save the social cost. How to extract sensitive data from mass data timely and efficiently has become a research focus.

Most of the existing image retrieval systems describe the image content directly using the traditional low-level image features, such as color, texture, shape. Because of considerable difference between these characteristics and people's understanding, image retrieval based on low-level features is always unsatisfactory [1~4]. In academic work and application systems, it is not described as an image understanding process. To improve the efficiency as much as possible, most retrieval system process retrieval based on simple image features rather than scene analysis. While browsing the video, we just want to know how much content the video has and watch clips of interest. How to integrate semantic features of images to search retrieval is the key to improve system performance. With the plenty rich content from video analysis and different search requirements, the problem of scene analysis is unable to avoid, and the core problem can be solved only by the image understanding method.

Video image understanding focused on the interpretation of video sequences, both related to the spatial characteristics of images and the time characteristics of the video sequences. This paper focuses on the application of image understanding methods with hierarchical structure in the traffic video semantic analysis. As an important component of traffic incident detection part, the results of vehicle status analysis directly impact on the application level of traffic video analysis technology. According to the real-time and accuracy of video analysis, this paper proposed a new video semantic analysis framework using image understanding. Firstly, obtain the moving vehicles by background subtraction method, then track the detected vehicles with fast normalized cross-correlation method based on prediction to get accurate vehicle trajectory, and extract the interested video information, such as vehicle I/O time, vehicle type, vehicle color, etc. Lastly, based on the above structured data, the useful information for decision can be gotten by video querying and statistical analysis.

II. FRAMEWORK FOR TRAFFIC VIDEO SEMANTIC ANALYSIS

A. Video Image Understanding

Based on the image object, image understanding is a science which takes knowledge for core, researches what the image has, the relationship between them, what scene the image is and how to use the scenarios. Image

© 2012 ACADEMY PUBLISHER
doi:10.4304/jmm.7.1.41-48

understanding, computer vision and artificial intelligence are closely linked. Visual information is the low-level data of image understanding, theoretical starting point is the machine vision; knowledge information is the high-level object of image understanding which based on artificial intelligence [5]. By analyzing and processing of video sequences, the image understanding obtained the understanding scene or object behavior from video. With the progress of technology, more and more applications are partial for the understanding of image sequence and have a wide applications range, including intelligent video surveillance, human-computer interaction, video compression, motion extraction and analysis, robot vision, etc.

Actually, video image understanding is to simulate people to understand visual scene, "visual" means that the human eyes obtain information from the objective world, "sense" means that processing visual information input to the brain. Similar to the human visual process, the purpose of the video image understanding is through the computer to complete the automatic scene understanding [6]. As shown in figure 1, the understanding course of video image understanding has hierarchy.

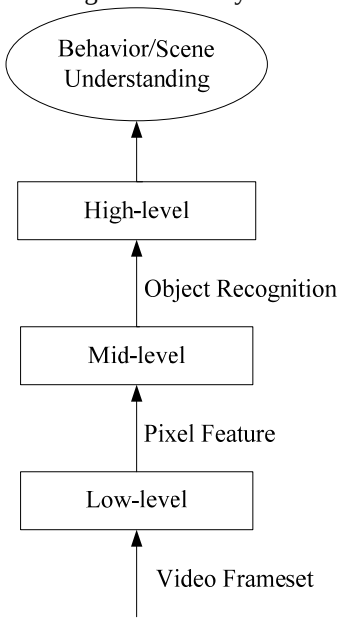


Figure 1. Video image understanding machanism

The video image segmentation in the lowest level is responsible for separating foreground from background of video sequence, the foreground is concerned content of upper understanding; object recognition and classification in the interlayer is responsible for identifying object in the foreground areas and object classification; the goal of behavior/scene analysis in uppermost layer is to understand the behavior of object and explain the scene state. The specific content at all levels includes:

(1) Video image segmentation.

The processing of input original video frame belongs to pixel-level operations, in fact it uses variety of image processing technology to transform the image to the image. Although the input image sequence data is great, actually the concern data of the upper understand is very small, always only the moving foreground constantly changing in the sequence, so the video image segmentation must be able to segment the moving target from foreground.

(2) Object recognition.

Object recognition means identification the separated foreground target, understanding their respective categories, such as judge the foreground region is human or pieces of image sequences belonged to the same goal. Similar to the video image segmentation, object recognition and image classification have differences in time-space operations and spatial operations.

(3) Behavior/scene understanding.

Behavior and scene understanding in uppermost layer is responsible for understanding the meaning of action and explaining the state of the scene. Understanding of individual behavior comes from the explanation of motion state and action; in the process of understanding, some prior knowledge is needed inevitably, which can be obtained through prior training. A large number of individual behavior constitute the state of scene, the understanding of scene means the clustering of individual behavior.

B. Application Framework

In the 21st century, intelligent traffic system will be the main trend and it's irreversible. In strengthening the road infrastructure, cities must consider how to improve transport efficiency and safety, reduce resources waste, ensure science and technology for the sustainable development of entire city. With the extensive use of traffic surveillance systems, the amount of traffic data is also growing at an alarming rate. It is unrealistic to manage the video artificially. Although the machine control is superior than people, the primary level of video information intelligent analysis directly leads low utilization efficiency of video information.

Intelligent Transportation System is an integrated transport and management systems, which uses advanced information technology, communication technology, sensor technology, control technology and computer technology, etc. It works in real-time, accuratly and efficiently. By close cooperation of people, vehicles and road, to improve transport efficiency, reduce traffic congestion and accidents, alleviate energy consumption and environmental pollution. Some well-known domestic and foreign research institutions and companies are focusing on intelligent transportation technology research and development, and the specific directions are Internet technology, intelligent video analysis, intelligent transportation platforms, etc.

This paper researches traffic video semantic analysis, and its intelligent transportation system framework is shown in figure 2.

© 2012 ACADEMY PUBLISHER

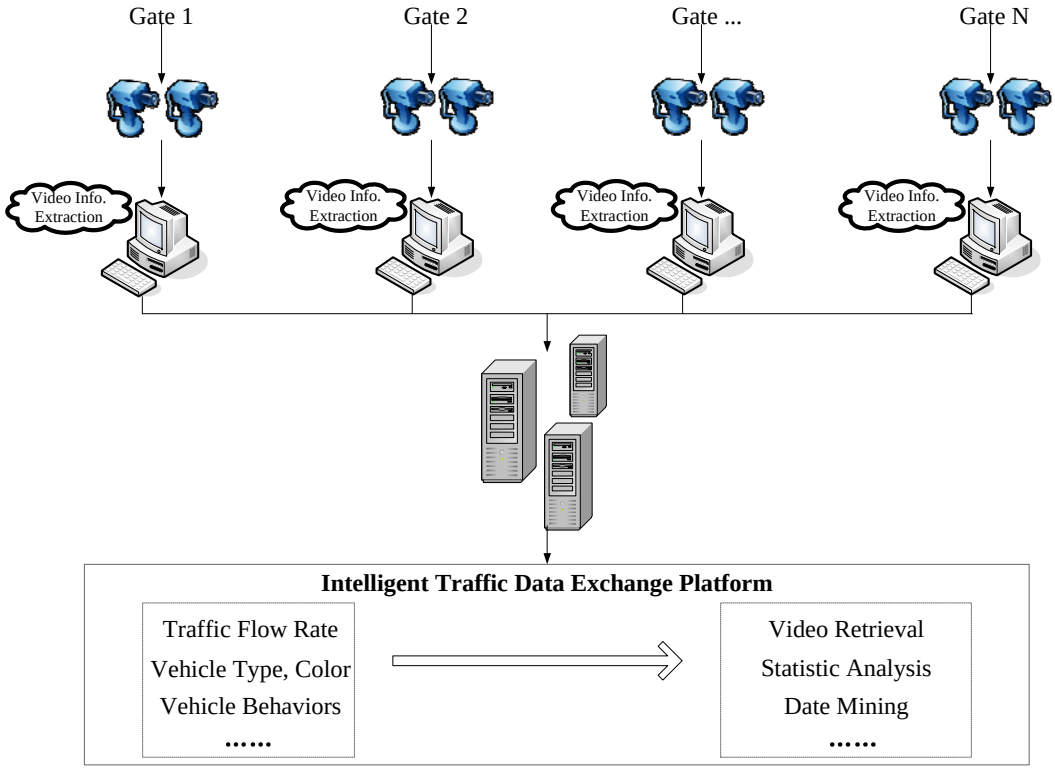


Figure 2. Framework of video analysis technology

Intelligent video analysis technology can be deployed on gate of road security system. After collecting traffic video data in real-time, extract sensitive data traffic from video by the video semantic analysis technology, and send to the server in symbolic form. Then, based on the extracted semantic information, the variety of application systems, using artificial intelligence theory, tools and techniques for statistical analysis and data mining, can serve the urban transportation and management.

III. VEHICLE TRACKING

A. Fast Normalized Cross-correlation

Cross-correlation has some advantages, simple, strong ability of anti-noise, etc. It is a statistical approximation method, commonly used in template matching and pattern recognition. Get the image f and the template t, to measure the similarity c:

$$c(u,v) = \sum_{x,y} f(x,y)t(x-u,y-v) \tag{1}$$

But there is some shortages in the process of matching using the formula (1) . For example, when the grayscale of the image varies with the position, the correlation of the matching area of the template and image may smaller than the correlation of some bright spot in the template and image, resulting in failure of image matching. Besides, the value of $c(u,v)$ is dependent on the size of the template. To solve this problem, it is necessary to normalize the cross-correlation. The coefficient of the so-called normalized cross correlation (NCC) is shown as formula (2).

$$\delta(x,y) = \frac{\sum_{i=-m/2}^{m/2}\sum_{j=-n/2}^{n/2} f(x+i,y+j).t(i,j) - m.n.\mu_f.\mu_t}{\left\{\left(\sum_{i=-m/2}^{m/2}\sum_{j=-n/2}^{n/2} f^2(x+i,y+j) - m.n.\mu_f{}^2\right).\left(\sum_{i=-m/2}^{m/2}\sum_{j=-n/2}^{n/2} t^2(i,j) - m.n.\mu_f{}^2\right)\right\}^{1/2}} \tag{2}$$

In the formula (2), $\delta(x,y)$ is the NCC coefficient of the image f and template t, while m and n is the size of the template t, while μ_f and μ_t are respectively calculate by the formula (3) and formula (4).

$$\mu_f(x,y) = \frac{1}{m.n}\sum_{i=-m/2}^{m/2}\sum_{j=-n/2}^{n/2} f(x+i,y+j) \tag{3}$$

$$\mu_t(x,y) = \frac{1}{m.n}\sum_{i=-m/2}^{m/2}\sum_{j=-n/2}^{n/2} t(i,j) \tag{4}$$

It can be seen from the formula of NCC, the more resemble the template and the area to be matched are, the more approach to the value of $\delta(x,y)$ is , and it can effectively solve the existing problem in matching of cross-correlation. But it can be seen in the same time that the process of cross-correlation normalizing involve many operation of square root and involution, resulting in a lot

© 2012 ACADEMY PUBLISHER

of calculation, and the complexity of the algorithm depends on the size of template t. So in this article we use fast normalized cross-correlation algorithm (FNCC) [7][8] , reducing the complexity of the traditional normalized cross-correlation algorithm.

FNCC make use of the concept of sum-table proposed by J.P.Lewis, improving the formula of NCC using sum-table [9]. When x=0,1,2,⋯⋯,M-1, and y=0,1,2,⋯⋯,N-1, the sum-table of a two-dimensional discrete variable $f(x,y)$ is expressed as:

$$S_\mu(x,y)=f(x,y)+S_\mu(x-1,y)+S_\mu(x,y-1)-S_\mu(x-1,y-1) \quad (5)$$

$$S_\sigma(x,y)=f^2(x,y)+S_\sigma(x-1,y)+S_\sigma(x,y-1)-S_\sigma(x-1,y-1) \quad (6)$$

$$S_c(x,y)=f(x,y).t(x,y)+S_c(x-1,y)+S_c(x,y-1)-S_c(x-1,y-1) \quad (7)$$

In the formula (5), (6) and (7), if x<0 or y<0, $S_\mu(x,y)$, $S_\sigma(x,y)$ and $S_c(x,y)$ are all equal to 0. We can reform the NCC formula by the three sum-table. In the formula (2), $\sum_{i=-m/2}^{m/2}\sum_{j=-n/2}^{n/2} f(x+i,y+j)$、$\sum_{i=-m/2}^{m/2}\sum_{j=-n/2}^{n/2} f^2(x+i,y+j)$、$\sum_{i=-m/2}^{m/2}\sum_{j=-n/2}^{n/2} f(x+i,y+j).t(x+i,y+j)$ are shown as formula (8), formula (9) and formula (10):

$$\sum_{i=-m/2}^{m/2}\sum_{j=-n/2}^{n/2} f(x+i,y+j)=S_\mu(x+m/2,y+n/2)-S_\mu(x-m/2-1,y+n/2)-S_\mu(x+m/2,y-n/2-1)+S_\mu(x-m/2-1,y-n/2-1) \quad (8)$$

$$\sum_{i=-m/2}^{m/2}\sum_{j=-n/2}^{n/2} f^2(x+i,y+j)=S_\sigma(x+m/2,y+n/2)-S_\sigma(x-m/2-1,y+n/2)-S_\sigma(x+m/2,y-n/2-1)+S_\sigma(x-m/2-1,y-n/2-1) \quad (9)$$

$$\sum_{i=-m/2}^{m/2}\sum_{j=-n/2}^{n/2} f(x+i,y+j).t(i,j)=S_c(x+m/2,y+n/2)-S_c(x-m/2-1,y+n/2)-S_c(x+m/2,y-n/2-1)+S_c(x-m/2-1,y-n/2-1) \quad (10)$$

We substitute the above three formula in the process of cross-correlation normalizing, namely FNCC. Then we can introduce the FNCC method to the object tracking , and update the moving object basing on forecasting.

B. Vehicle Tracking Based on Prediction

In this paper, we update the tracking template in real time [10][11], making use of the moving vehicle tracked in the current frame to update the tracking template. There is a feature that the moving vehicle is stiff object, so the vehicles in two adjacent frames changes little. We can track the vehicles make use of it. To some extent, it meets the requirement of high matching probability of normalized cross-correlation. So it improves the efficiency of tracking as well as the accuracy of tracking.

We can get the moving trajectory of tracked object in the tracking process, and predict the position of the moving object in the next frame [12]. Accordingly, we should normalize the moving templates and the predicted regions fast, and it can reduce the time consumed in the fast normalized cross-correlation. In the tracking process, at first we should track the moving object and then get the trajectory of it. Suppose the position of the moving vehicle in the current frame f_n is P_n while the position in the former frame f_{n-1} is P_{n-1}. Then the size of the area to be matched refers to the area position of the moving object in the current frame. We can calculate it in accordance with the formula (11) and formula (12) as follows:

$$SR.x = t.x + 2\times(P_n.x - P_{n-1}.x) \quad (11)$$

$$SR.y = t.y + 2\times(P_n.y - P_{n-1}.y) \quad (12)$$

In the formula (11) and formula (12), $SR.x$ and $SR.y$ is respectively the length and width of the predicted area, and $t.x$ and $t.y$ is respectively the length and width of the template, while $P_n.x$ and $P_n.y$ are the adjacent points in the trajectory, as well as $P_{n-1}.x$ and $P_{n-1}.y$.

IV. VIDEO CONTENT EXTRACTION

A. Vehicle Type Recognition

Vehicle type identification is an important component of the Intelligent Transportation system. It is now the hot spot domestic and foreign. This paper presents a method of vehicle type identification based on drive line detection and divided the vehicles into four types, they are car, truck, van and bus.

Basic steps of vehicle type identification are as follows: (1) First get the background image of the current video frame.

Background subtraction method is faster than the inter-frame subtraction method and optical flow method, and can detect the moving object accurately. So we adopt the method to get the moving regions. Before the operation of subtracting background, we should first build the background model. It is reasonable to assume that for some time there is only small gray scale change in the background while the gray scale of the prospect varies a lot to different vehicles, even the different parts of the same car is not the same. The schematic diagram of background extraction course is shown in figure 3.

© 2012 ACADEMY PUBLISHER



(a) Video Frame

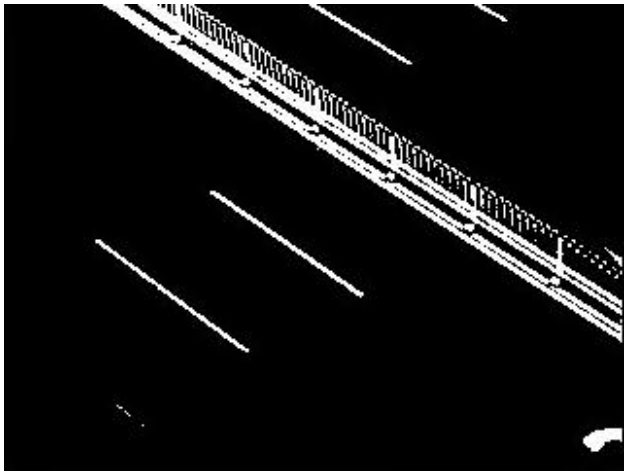


(b) Extracted background frame

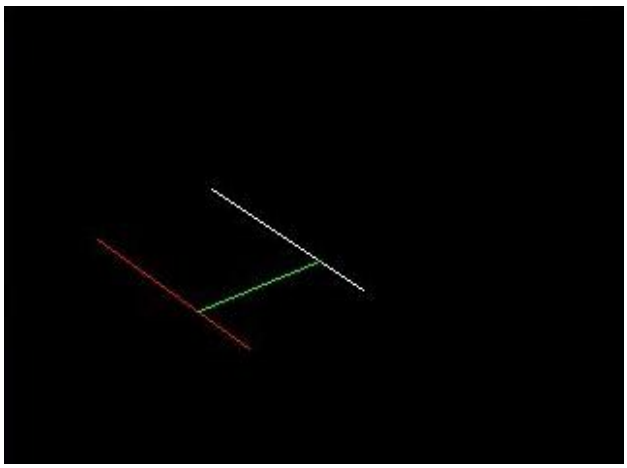
Figure 3. Background extraction

(2) Identify two drive lines using Hough Tranform.

It is often necessary in computer vision to identify or locate some geometry, such as lines, circles, ellipses, and other graphics. Hough transform provides a method to detect the lines and is simple. Which is developed to detect circles, ellipses, and other general graphics, Hough transform's theory is to map a set of points of some graphic to one point. The point record the number of the points in the set and the number recorded assist the program in finding the point. The point is the parameter of the graphic and the range of the parameter is called parameter space. Hough transform using straight lines with polar coordinates to represent. There is two parameters and they are associated by polar-style, so the voting process only need to traverse one of them and search the peak in the two-dimensional parameter space. The schematic diagram of drive line width determination is shown in figure 4.



(a) Binarization of background frame



(b) Calibration of drive line

Figure 4. Determination of drive line width

(3) Using the ratio of measured drive line width and actual width to calculate actual area of vehicle S_{Real}.

$$S_{Real} = (\frac{d}{l})^2 S_{Oval} \tag{13}$$

$$S_{Oval} = \pi ab \tag{14}$$

Where d is the actual lane width, l the measured lane width, a and b are the lengths of the major semi-axis and the minor semi-axis of the external ellipse.

(4) Classify the vehicles with the actual area value of the vehicle S_{Real}.

B. Color Recognition

First we should extract the color information and classify it to the specified color category. There are several key questions. First, we should select the appropriate color space. Second is to define the color characteristics and quantified. At last, we must measure the similarity and match, in other words, how to define the similarity of the color features and match fast. As the RGB color space is used universally in the images and it is easy to access the RGB color of the vehicles. But there is a problem that we can't commonly use Euclidean

© 2012 ACADEMY PUBLISHER

distance to properly describe the distance of two different colors.

In addition, when the color changes the three value of R ,G and B will have non-linear change, it is not conducive to the color quantization and classification. Therefore, we choose HLS space in this paper and we will first transform the RGB color space into the HSL color space.

Make $\max$ the maximal value of R,G and B, while $\min$ the minimal. Then the fomula used to transform the RGB color space into the HSL color space is as follow :

$$H=\begin{cases}0^\circ, & if\ \ \max=\min\\ 60^\circ\times\dfrac{G-B}{\max-\min}+0^\circ, & if\ \ \max=R\ and\ G\ge B\\ 60^\circ\times\dfrac{G-B}{\max-\min}+360^\circ, & if\ \ \max=R\ and\ G<B\\ 60^\circ\times\dfrac{B-R}{\max-\min}+120^\circ, & if\ \ \max=G\\ 60^\circ\times\dfrac{R-G}{\max-\min}+240^\circ, & if\ \ \max=B\end{cases} \tag{15}$$

$$L=\frac{1}{2}(\max+\min) \tag{16}$$

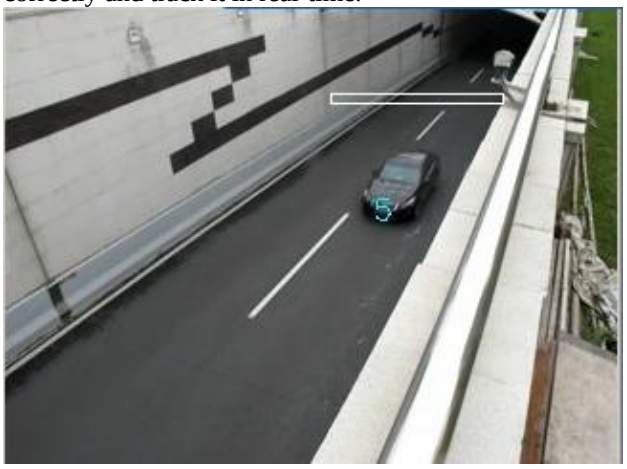
$$S=\begin{cases}0 & if\ \ L=0\ or\ \max=\min\\ \dfrac{\max-\min}{\max+\min}=\dfrac{\max-\min}{2L}, & if\ \ 0<L\le\dfrac{1}{2}\\ \dfrac{\max-\min}{2-(\max+\min)}=\dfrac{\max-\min}{2-2L}, & if\ \ L>\dfrac{1}{2}\end{cases} \tag{17}$$

The analysis found that L component to a large extent dictate the colors of black, white and gray, in other words, when the L component is in a fixed range the color will be one of the three colors no matter what the value of H component and S component is. So the method to quantify and classify the color is to distinguish the colors of black, white and gray according to the value of L component and then determine the other colors according to the value of H component. With this method we can divide the HSL color space into seven colors of black, white, gray, red, yellow, green and blue, which comprise almost all the vehicles of pure color in the video.

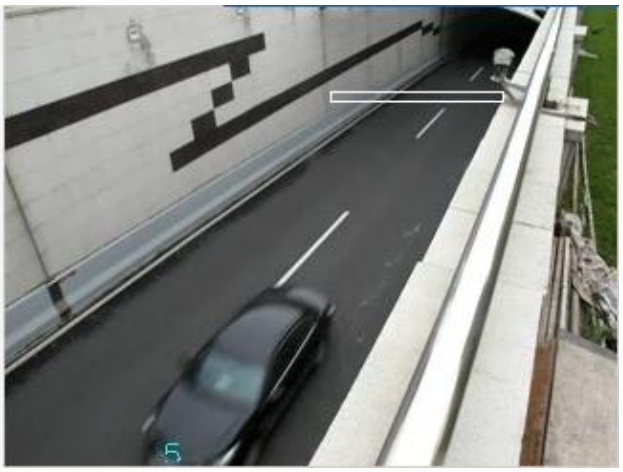
V. Experimental Results and Analysis

A. The Tracking Effect

The screenshots of this article is from the video which is shot in the tunnel of DuShuHu in Suzhou and processed with the tracking algorithm of normalized cross-correlation. The tracking result is shown in figure 5 and figure 6. The fifth vehicle drive in the monitoring area at 624th frame and out at the 657th frame. The fifth figure intercepts the 630th frame and 650th frame, from which we can see the currently tracking although the closer the vehicle away from, the bigger it is. Similarly, the two frames intercepted show that the van drive in at 931st frame and out at 960th frame. Although the size and shape of the van changes, we can calibrate the van correctly and track it in real-time.



(a) One moment of vehicle 5# tracking



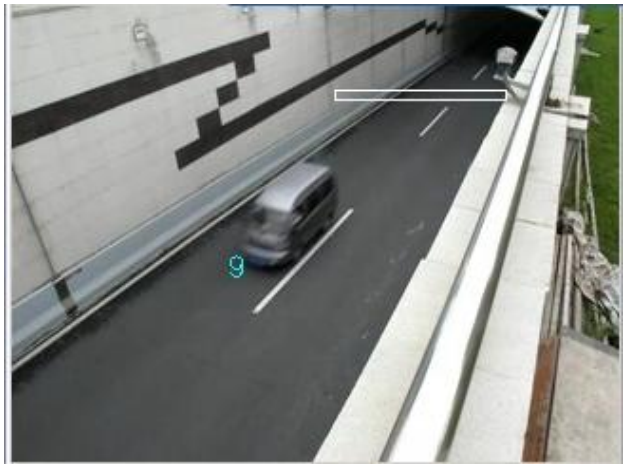
(b) Another moment of vehicle 5# tracking

Figure 5. The tracking cource of vehicle 5#



(a) One moment of vehicle 9# tracking

© 2012 ACADEMY PUBLISHER



(b) Another moment of vehicle 9# tracking

Figure 6. The tracking course of vehicle 9#

B. Video Content Extraction Result

Based on accurately tracking, we can extract the video information we are interested in with the method of vehicle style identification and color classification designed above. In this paper, we primarily judge and find the numbers of the vehicles, the frame number the vehicles drive in and out, the vehicle category and the vehicle color, and eventually write in the database for the further experiments. It is shown in Figure 7 that the algorithm involved in this article can identify the data of the vehicle mentioned above and finally write them to the database.

cameraID	videoID	vehicleID	frameIn	frameOut	vehicleType	vehicleColor	memo1
0101	2009072806	1	167	184	Van	White	
0101	2009072806	2	661	703	Car	Black	
0101	2009072806	3	719	726	Car	White	
0101	2009072806	4	924	957	Car	Black	
0101	2009072806	5	1042	1042	Car	Black	
0101	2009072806	6	1066	1066	Car	Silver	
0101	2009072806	7	1231	1260	Van	Blue	
0101	2009072806	8	1745	1772	Van	Silver	
0101	2009072806	9	2094	2103	Car	Blue	
0101	2009072806	10	2339	2377	Car	Blue	
0101	2009072806	11	2419	2448	Car	Black	
0101	2009072806	12	2614	2657	Car	Black	
0101	2009072806	13	2623	2623	Car	White	
0101	2009072806	14	2673	2708	Car	Black	
0101	2009072806	15	2940	2940	Car	White	
0101	2009072806	16	3144	3144	Car	Yellow	
0101	2009072806	17	3172	3342	Van	Black	
0101	2009072806	18	3352	3352	Car	Black	
0101	2009072806	19	3470	3497	Car	Yellow	
0101	2009072806	20	3539	3566	Car	Yellow	
0101	2009072806	21	3741	3748	Van	White	
0101	2009072806	22	3857	3889	Car	Black	
0101	2009072806	23	3892	3892	Car	White	
0101	2009072806	24	4642	4665	Van	White	
0101	2009072806	25	4929	4929	Car	White	
0101	2009072806	26	5009	5037	Car	Black	
0101	2009072806	27	5013	5036	Car	Black	

Figure 7. Extraction result of video content

C. Video Querying

After the above steps, we transform the video data into text and can query them for the specific needs. As is shown in the figure 8, the users set the search criteria, including the date, the time, vehicle style, vehicle color and so on. According information written in the previous steps, the system will return back the frame set and extract the corresponding video frames in real-time. So the users can quickly check the video clip they want.

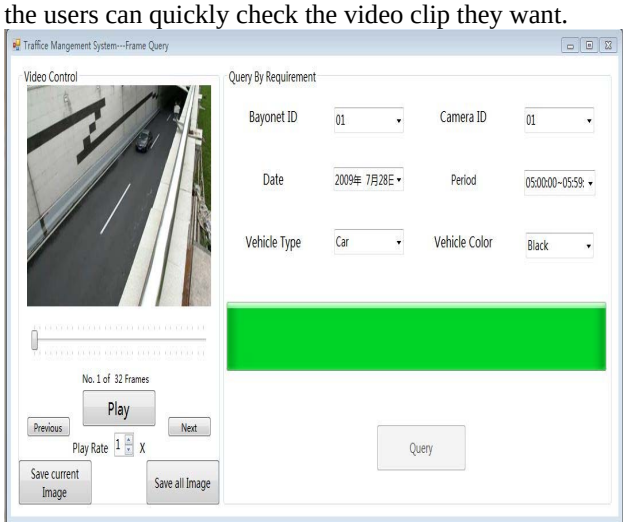
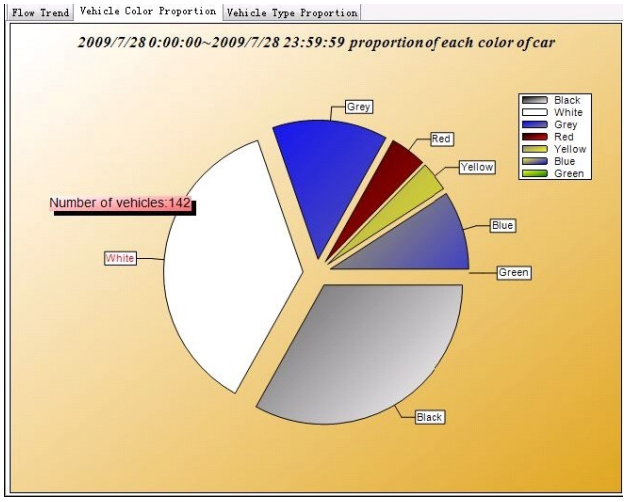


Figure 8. Video querying window

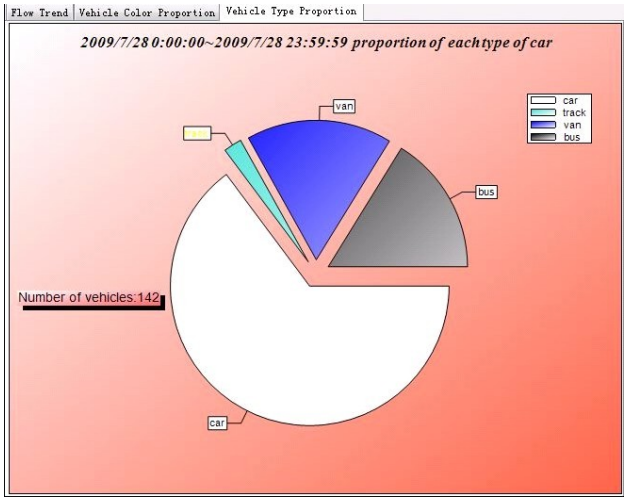
D. Statistical Analysis

With the methods involved in the paper, users are allowed to set a specific date and time ,and statistic the vehicle style and color, finally get the percents of all the categories and colors. The statistic of the vehicle category can show the current road situation, while the statistic of the vehicle color can show the preference vehicle color of the public and assist the manufacture and trade company in making a strategic decision. These statistics can also assist the traffic control department with the corresponding analysis and decision. As is shown, the figure 9(a) is the ratio diagram of the vehicle colors of the video got in one day of July in 2009, while the figure 9(b) is the ratio diagram of the vehicle types.



(a) According to vehicle color

© 2012 ACADEMY PUBLISHER



(b) According to vehicle type

Figure 9. Vehicle statistical analysis

VI. CONCLUSION

Regarding to the problem that traditional video retrieval methods cann't make full use of semantic information of video flow, this paper introduces the image understanding method to traffic video analysis, and proposes a video semantic analysis framework based on scene analysis. Firstly, obtain the moving vehicles by background subtraction method, then track the detected vehicles with fast normalized cross-correlation method based on prediction to get accurate vehicle trajectory, and extract the interested video information, such as vehicle I/O time, vehicle type, vehicle color, etc. Lastly, based on the above structured data, the useful information for decision can be gotten by video querying and statistical analysis. The experimental results verified the feasibility of the framework.

ACKNOWLEDGEMENT

This research was partially supported by the Natural Science Foundation of China under grant No. 60970015, the 2009 Special Guiding Fund Project of Jiangsu Modern Service Industry (Software Industry) under grant No. [2009]332-64, the 2010 Program for Postgraduates Research Innovation in University of Jiangsu Province under grant No. CX10B_041Z, the Applied Basic Research Project (Industry) of Suzhou City under grant No. SYJG0927 and No. SYG201032, and the Beforehand Research Foundation of Soochow University.

REFERENCES

[1] Lee D, Barber R, NiBlack W, et al. Indexing for complex queries on a query-by-content image database[C]. In Proceedings of ICPR 1994, pp.142-146.

[2] Pentland A, Picard R W, Sclaroff S. Photobook: content-based manipulation of image database[J]. International Journal of Computer Vision, 1996, 18(3):233-254.

[3] Simith J R, Chang S F. Tools and techniques for color image retrieval[C]. In Proceedings of IS&T/SPIE 1996, vol.2670, pp.1-12.

[4] Wu J K. Content-based indexing of multimedia database[J]. IEEE Transations on Knowledge and Data Engineering, 1997, 9(6):978-989.

[5] XIE Zhao. Research for key issues and methods in image understanding[D]. Hefei Unversity of Technoly, 2007.

[6] LIANG Ying-hong, WANG Zhi-yan, CAO Xiao-ye, XU Xiao-wei. Application of video image understanding to measurement of pedestrain flow[J]. Computer Engineering and Design, 2008, 29(5):1203-1206.

[7] D. M. Tsai, C. T. Lin. Fast normalized cross-correlation for defect detection[J]. Pattern Recognition Letters, 2003, 24, 2625-2631.

[8] GUO Wei, ZHAO Yi-Gong, XIE Zhen-Hua. An Improved Normalized Cross-correlation for Template Matching of Infrared Image[J]. Acta Photonica Sinica, 2009, 38(1) : 189-193.

[9] Lewis J P. Fast normalized cross-correlation [A]. Proceeding of Vision Interface[C]. Canada Quebec, 1995, 120-123.

[10] HOU Zhi-Qiang, HAN Chong-Zhao. A Survey of Visual Tracking[J]. Acta Automatica Sinica, 2006, 32(4), 603-617.

[11] XUE Chen, ZHU Ming, LIU Chun-xiang. Review of tracking algorithms under occlusions[J]. Chinese Journal of Optics and Applied Optics, 2009, 2(5), 388-393.

[12] Jian Wu, Heng-jun Yue, Zhi-ming Cui, Jian-ming Chen. Moving Object Tracking Method Based on Adaptive On-line Clustering and Prediction-based Cross-correlation[C]. In Proceedings of IEEE CIT 2010, Bradford, UK, pp.487-494.



Jian Wu was born in Nantong on the 29th April, 1979, and got master degree in the field of computer application technology from Soochow university, Suzhou city, China in 2004. The main research direction is computer vision, image processing and pattern recognition. He works as a teacher in the same college after his master graduation. Now he is pursuing the doctoral degree. He was awarded the Third Prize of 2007 Suzhou City Science and Technology Progress and the 2008-2009 Soochow University Graduate Scholarship Model.

Zhi-ming Cui was born in Shanghai on the 4th July, 1961. Professor, PhD Candidate Supervisor. The main research direction is deep web and video mining.

Heng-jun Yue was born in Henan on the 11th July, 1984. Master, his main research direction is video processing.

Guang-ming Zhang was born in Suzhou on the 10th Februry, 1981. PhD Candidate, his main research direction is image processing and video retrieving.

© 2012 ACADEMY PUBLISHER