Traffic Measuring and Authentication on Digital Video

Guorong Xuan, Yang Xiao, Xiaoguang Yang, Jidong Chen, Chengyun Yang, Ruhua Zhang

Abstract—**Several schemes for traffic monitor on video are developed. The first is the dynamic gap technique by means of digital video processing instead of subsurface electromagnetic detectors. The dynamic background refreshing and threshold selection are obtained by statistical methods. The second is density and queue length measurement in a traffic detection region. The third is interpretation and identification of traffic activities by Hidden Markov Models and entropy minimization to discover the relationships between hidden causes and observed scenes. The learning system learns a concise model of scene behavior directly from optical flow. The fourth is to authenticate traffic images with robust digital watermarking technique and hiding data in image losslessly.**

Index Terms—**dynamic gap, entropy minimization, Hidden Markov Models, watermarking, data hiding**

I. INTRODUCTION

THE ITS (Intelligent Transport Systems) which are now promoted as the national project, targeting the improvement of safety, transport capacity, and comfortableness as well as the environment protection. As one of the fundamental parts of ITS, vehicle flow measuring system is very important. The normally used subsurface electromagnetic detectors have some disadvantages, like the limitation of placing positions, being subject to damage, and so on. Therefore, video is chosen as a solution to measure vehicle flow because it's easy to handle with [1] [2] [3].

Our traffic flow measuring system contributes to the proper monitoring and control of the traffic flow, measuring velocity, vehicle length, and the queue length as well as the level of congestion using the images from the camera installed on poles or other tall structures [4]. Video is captured, digitized, and then measured by system. In this paper, we proposed a new scheme for vehicle flow measuring which recognizes the vehicle real-time by the gray scale variation on the video measuring lines. Adopting the dynamic gap technique in training samples, we can get the feature curves from the dynamic gap drawing, which makes the determination of the gray scale difference threshold illustratively and reliably. Consequently, the measuring system has demonstrated good performance in the practical application. As shown in Fig. 1, when the vehicle pass the video measuring line, system can recognize the vehicle, then count the number and signal the controlling system. Besides, queue length can be provided in meters or number of vehicles between the stop bar and the end of region [5].

Another scheme is to interpret and identify traffic activities by HMM and entropy minimization. We introduce a novel generalization of HMM for multiple observations per frame [6], enabling us to model scenes with arbitrary and changing numbers of participants, and solve for its entropy based MAP (Maximum A Posteriori) estimation formula. In minimizing the sum of entropies, the iterative MAP estimation minimizes uncertainty in the sufficient statistics and cross-entropy between the model and the data. Iterative estimation with trimming operator extinguishes weakly supported parameters, compressing and simplifying the model, and guaranteeing an increase in posteriori probability.

By using watermarking, the electronic-police system monitoring vehicles at intersection of cities has more and more applications. Color rectification and watermarking technology help to solve color saturation and authentication problem of traffic camera images. A lossless high-capacity data hiding for image watermarking based on integer wavelet and histogram adjustment is also proposed for both self-authentication and reversible data embedding [7].

II. DYNAMIC GAP TECHNIQUE

We are developing a traffic flow measuring system which can produce individual vehicle data (e.g., flow, velocity, headway, vehicle length), leading to better traffic flow modeling and improved understanding of driver behavior. Our strategy includes dynamic gap, threshold selection, dynamic background refreshing.

A. Dynamic Gap

Actually, video signal (3D) is the combination of the 2D video image (Fig. 1) and the time signal. It will be very hard to get features from these 3D signals. Herein, we are only interested in the gray scale of the image part on the video measuring line. Thus the recognition will be simplified a lot.

Arranging the images on one video measuring line in the order of time, "dynamic gap drawing" of this measuring line will be got.

Manuscript received March 15, 2003. This research is supported by the China 973 project (No.G1998030419).

Guorong Xuan (e-mail:grxuan@public1.sta.net.cn, webpage: http://grxuan.xiloo.com/), Yang Xiao, Jidong Chen, and Chengyun Yang are with the Dept of Computer Science, Tongji University, Shanghai 200092, China; Xiaoguang Yang and Ruhua Zhang, are with the Dept of Traffic and Transportation Engineering Tongji University, Shanghai 200092, China

As in Fig. 2, the continuing gray scale variation of the images on one measuring line is shown as three segments. The horizontal direction represents time and every vertical segment is the image's gray scale on this measuring line at that time.

Fig. 1. System frame

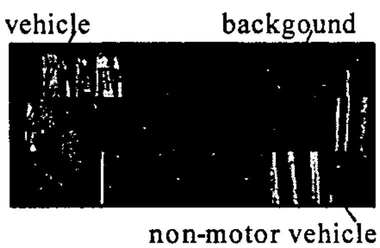
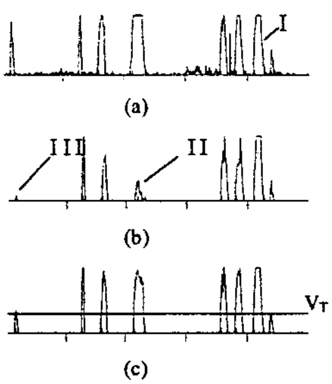


Fig. 2. Dynamic gap drawing

From this dynamic gap drawing, the road background and the images' gray scale variation during the vehicle passing can be clearly discerned. And vehicles and non-motor vehicles (e.g., bikes and pedestrians) can be easily differentiated from each other. So the gap drawing contains enough information to get vehicles' features and to complete the recognition as follows.

B. Threshold Selection



(a) RGB_T=15 (b) RGB_T=35 (c) RGB_T=25

Fig. 3. V value of different RGB_T

Feature V is the number of the points on the video measuring line, whose variation range of real-time signals (the R, G, B gray scales) with background is bigger than gray scale difference threshold RGB_T. The feature V gives the difference between real-time signals and background signals. The bigger the V is, the bigger the difference will be. If V is greater than threshold V_T, we can determine there are vehicles passing, otherwise, there is only background of the road. In this paper, V_T is called vehicle threshold. The value of threshold V_T can be determined by some methods. For example, to deal with the gap drawing of Fig. 2, segment 3, the process is as following (Fig. 3). V value at each moment is calculated and Fig. 3 shows V value of different RGB_T for a period of time. Based on Fig. 3, if RGB_T is small (RGB_T=15), some changes on background might be overstated (as I in (a)). If RGB_T is large (RGB_T=35), the gray scale variation caused by passing vehicles will be reduced (as II in (b)), or even omitted (as III in (b)). In this system, the RGB_T value is chosen as 25 by careful evaluation based on a lot of experiments. Comparing with the dynamic projection drawing, he V value in (c) basically coincides with the situation changes of background and vehicles. After RGB_T is chosen, V value graph can be drawn, on which the vehicle threshold V_T can be illustratively determined. In Fig. 3 (c) the horizontal straight line shows the V_T value. In practice, both RGB_T and V_T vary in small range (about ± 10). It will be possible to combine all variation of RGB_T and V_T of all the training samples to get the best recognition.

C. Dynamic Background Refreshing

As mentioned above, by comparing real-time signals with background signals, the degree of the gray scale variation on the video measuring line will be used to judge the vehicle threshold. Therefore, the background value will be very important. Due to the background change, we use dynamic background instead of fixed background for measurement.

In this paper, the dynamic background is that the latest gray scale on measuring line, which is continually updated and approved as background by monitoring program, will be used as the new background. For graduate and abrupt changes of background, the special strategy to refresh the dynamic background is necessary.

D. Results of the Measurements

Table I shows the traffic flow results of our system measuring on the video provided by the headquarters of Shanghai's traffic police, which records the traffic situation at the intersection of Yan'an Road and Zhong'shan Road N on August 21st, 2000. We select a lane from 11:00am-12:00am

As in Table I, we confirm that the measuring system has measured the traffic flow precisely. Using the method, the dynamic background is always the true background, and so it is correctly to detect the vehicles' passing. But a few vehicles are missed, which mostly caused by the vehicles shielding each other. Shielding phenomenon is a common problem of using video methods to measure vehicle flow. The solution is to install cameras vertically. Both the results with or without shielding are shown in Table I. The recognition rate of our system is above 90%, and if the camera is installed vertically, the recognition rate could be above 95%. According to these data, the results of this system are satisfying.

Table I
Results of measuring vehicles

A	B	C	D	E	F	G
11:00~11:10	78	1	7	4	89.7	94.6
11:10~11:20	91	0	7	7	92.3	100.0
11:20~11:30	93	0	6	5	93.5	98.9
11:30~11:40	76	1	9	5	86.8	93.0
11:40~11:50	81	1	3	1	95.1	96.3
11:50~12:00	65	3	4	3	89.2	93.5
11:00~12:00	484	6	36	25	91.3	96.3

A: Time period

B: Number of vehicles passing by

C: Extra counting number of vehicles

D: Total missing number of vehicles

E: Missing number of vehicles by shielding

F: Overall recognition rate (%)

G: Recognition rate without shielding (%)

Adopting the strategy of dynamic gap and dynamic background refreshing, we can also set one measuring line close to another on the video to measure the speed and length of the vehicle. We will get the pixel number vehicle passed on video at each time unit by means of dividing the distance between the two measuring lines by the recorded time from vehicle entering first line to its exiting the second line. Then our system computes a projective transform, or homography, between the image distance and real world distance, so the real vehicle parameter is obtained.

III. DENSITY AND QUEUE LENGTH MEASUREMENT

Our current system can process images received from the camera installed above the approach lane at the traffic signal intersection, and can directly measure queue and delay length as well as number of vehicles, vehicle type. It mainly has two processes, image preprocess and queue length measurement process.

Fig. 4. Density and queue length measurement

Fig. 4 shows segmentation of the scene into individual vehicles. In the region designated by users freely, every vehicle is found by system and marked by a rectangle. Queue length can be provided in meters or number of vehicles between the stop bar and the end of region. And information (such as location, scale etc.) for each vehicle is obtained. Density (or occupancy) can also be calculated by means of dividing the number of vehicles (or the sum of the vehicles' area) by the region's area. The experimental results show that our system has achieved accurate measurement within 10% error of maximum queue length under various conditions. These are the sort of information a traffic engineer needs to design an intersection and its controls.

IV. INTERPRETATION AND IDENTIFICATION BY HMM AND ENTROPY MINIMIZATION

A HMM as traditionally formulated has limited applicability because it is fundamentally a model of a single hidden process, observing a single fixed-length observation vector in each time step. In fact, pedestrian plazas and vehicle roadways have varying numbers of participants that constantly enter and exit the scene. Here, we introduce a novel graphical model that generalizes HMM to take a varying number of observations per time step and solve for its MAP reestimation formula. This model will learn holistic modes of activity in the scene, which the sorts of traffic modes that a traffic engineer would need to know when designing controls for an intersection. Rather than attempt simultaneous tracking of tens of variable-sized objects, with all the atendant sources of error, we will learn a distribution over low-level image processes. The patterns of moving traffic are choreographed by the (invisible) traffic lights into phases of horizontal and vertical traffic, as well as implicit sub-phases with different frequencies of right and left turns. Our model will have to "learn" the typical locations and directions of the moving pixels, as well as the dynamic changes of these patterns through entropy minimization. The multi-modal distribution of the moving pixels can be captured with multivariate Gaussian mixture models. HMMs can be extended to handle multiple observations per time step by treating each frame's flow-list as an observation sequence for the mixture model at that time step. We call the resulting model a multi-observation mixture counter HMM [6].

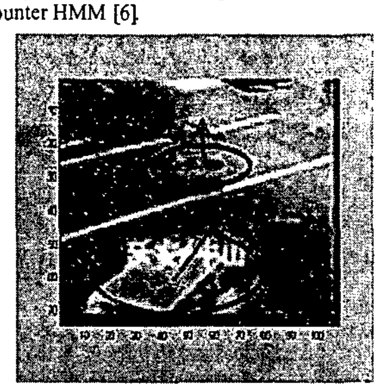


Fig. 5. (a) Minimum-entropy estimation

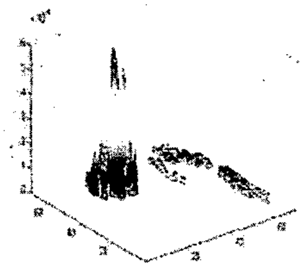


Fig. 5. (b) Minimum-entropy estimation

Maximize the posterior probability, equivalently, minimizing the sum of entropies, will reduce uncertainty in all respects. In this light, entropic EM is a search-less and highly efficient form of structure learning under a minimum description length constraint.

The entropic posterior defines a distribution over all possible model structures and parameterizations within a class; small, accurate models having minimal ambiguity in their joint distribution are the most probable. To find these models, we simply replace the M-step of EM (Expectation Maximization) with the entropic MAP estimator. This drives weakly supported parameters toward zero and concentrates evidence on surviving parameters, causing their estimates to approach the ML (Maximum-Likelihood) estimate. Structurally irrelevant parts of the model gradually expire, leaving a skeletal model whose surviving parameters become increasingly well-supported and accurate.

Experiment Results

We pointed the camera out the window at a busy way. The optical flow was spatially sub-sampled by a factor of 10. The rhythm of traffic is quite clear (Fig. 5 (a)). Arrows indicate the mean optical flow. Ellipses indicate standard deviation spatial iso-probability contours of each mixture component. The curve (Fig. 5 (b)) shows the mixture Gaussian distribution of an optical flow list between two frames. It is also possible to make short-term predictions of traffic dynamics by projecting the current hidden state probabilities through the transition matrix. If we combine this with the current relative activations of the mixture components, we get an estimate of where the model is expecting to see motion in the next frame.

V. WATERMARKING FOR AUTHENTICATION

With the development of recognition and network techniques, robust digital watermarking and image data hiding have drawn lots of interest recently. For example, electronic police system monitoring vehicles at intersection of cities has more and more applications. Do color rectification firstly to the image captured by the digital camera, which recover the red light effect in HVS, and then watermarking it. The watermarking technique is used in this paper for authentication of the images, showing the traffic rule violation of the driver.

Experimental pictures for robust digital watermarking are as follows: original image (Fig. 6), original watermark (Fig. 7), watermarked image (Fig. 8), watermarked image and then be cut out a part (Fig. 9) and its extracted watermark (Fig. 10), watermarked image and then be compressed 32 times by JPEG (Fig. 12) and its extracted watermark (Fig. 11).

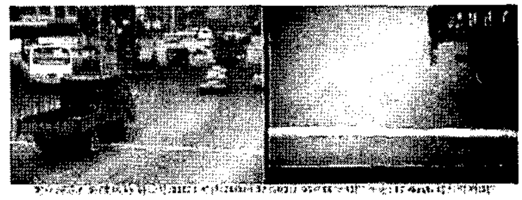


Fig. 6. Original image

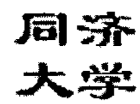


Fig. 7. Original watermark

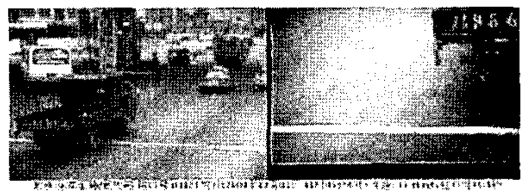


Fig. 8. Watermarked image

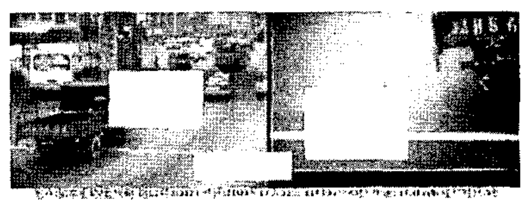


Fig. 9. Watermaked image by cutting out

Fig. 10. Watermark of Fig.9.

Fig. 11. Watermark of Fig.12.

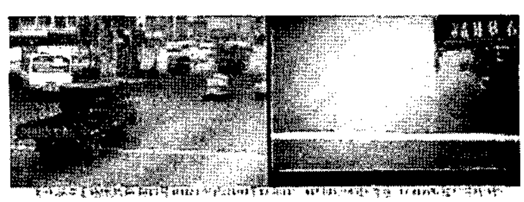


Fig. 12. Watermarked image compressed 32 times by JPEG

In special traffic images, even small modifications are not allowed for legal reasons. It serves for the purposes of both self-authentication and reversible data embedding. As being reversible, the original digital content can be completely restored after authentication. A lossless high-capacity data hiding for image watermarking based on integer wavelet and histogram adjustment is proposed [7]. After extracting data embedded, the original image should be reversible from watermarked image. The key problem for integer wavelet using in image watermarking is that the gray level should be held in given range. The dynamic range histogram adjustment on gray level is used to avoid overflow after returning space domain from wavelet domain.

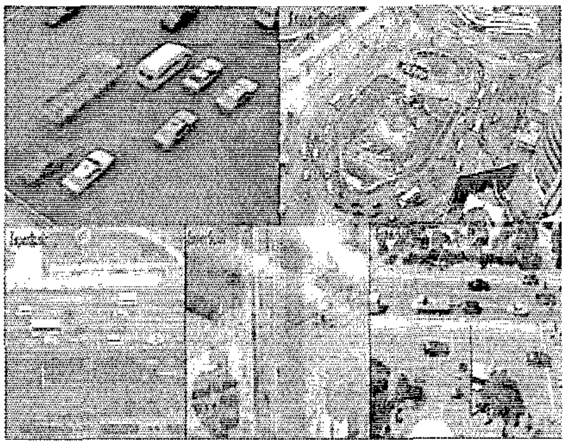
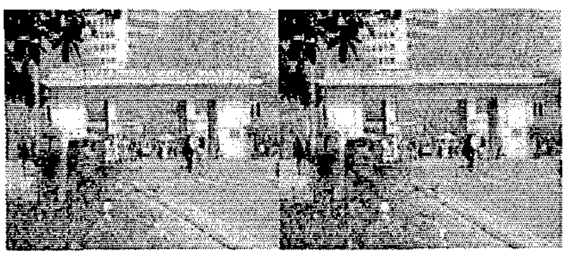


Fig. 13. Five traffic images



(a) (b)

Fig. 14. Original & watermarked images

Experimental pictures for lossless data hiding are as Fig. 13 and Fig. 14. The embedded watermark (Fig. 13 include five color traffic images, each size 352 × 288), original color image (Fig. 14 (a) size 1024 × 768, also named stego-image), and color watermarked image (Fig. 14 (b) size 1024 × 768) are shown. It is very hard to tell the difference between Original and watermarked images by human eyes. All these original image and five traffic images can be recovered losslessly.

VI. CONCLUSIONS

The measuring systems using concept of dynamic gap is developed. After the approach for obtaining the features is determined, the best results for training samples can be got by tuning the thresholds. The method for determining feature is got from experiments. Dynamic background technique makes sure the true background.

We process images received from the camera installed above the approach lane at the traffic signal intersection, and can directly measure queue length in meters as well as number of vehicles, vehicle location, vehicle scale and density in region that designated freely.

We have also developed an efficient method for finding compact hidden-variable probability models via entropy minimization. These models are highly predictive and often interpretable as theories that identify relationships between hidden causes and observed effects. The system learns a concise model of scene behavior directly from optical flow. We introduce a novel generalization of HMM for multiple observations per frame, enabling us to model scenes with arbitrary and changing numbers of participants. In the maximum structure case, the MAP estimator smoothly extinguishes excess parameters, thereby simplifying the model and preventing over-fitting.

The robust digital watermarking technique provides a good way to the authentication of transport images. A lossless high-capacity data hiding for image watermarking based on integer wavelet and histogram adjustment is also proposed for both self-authentication and reversible data embedding.

REFERENCES

[1] J.Bemsen, "Dynamic Thresholding of Grey-level Images",Proc. of Int. Conf. On Pattern Recognition, 1986.

[2] Y.J. Kim,Y.S. Soh, "Improvement of Background Update Method for Image Detector", Proc. of the 5th World Congress on Intelligent Transport Systems, *Bd.1*, 1998

[3] D.P.Panda, *et al.*, "Image Sensing System Technology Advancement For Adaptive Traffic Management", ITS World Congress, Orlando, Florida, October 1996.

[4] Guorong Xuan, Jianguo Jiang ,Peiqi Chai, *et al.*, "Real -Time Traffic Flow Measuring System Based on Gap Video Image Processing",Proc. of 7th Symposium on Transportation System Theory and Application, 1994.

[5] D.Beymer, B.Coifman, P. McLauchlan, J. Malik, "A Real-Time Computer Vision System for Area-Wide Traffic Surveillance and Vehicle Tracking", Transportation Research: Part C, 1998.10.

[6] M.Brand, V.Kettnaker, "Discovery and Segmentation of Activities in Video", IEEE Transactions on Pattern Analysis and Machine Intelligence, 22 (8), August 2000.

[7] Guorong Xuan, Jiang Zhu, Jidong Chen, Yun Q. Shi, *et al.*, "Distortionless Data Hiding Based on Integer Wavelet Transform", IEE journal, ELECTRONICS LETTERS, Volume 38, No 25, Dec.2002.