World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:5, No:6, 2011

A Vehicular Visual Tracking System Incorporating Global Positioning System

Hsien-Chou Liao and Yu-Shiang Wang

Abstract**—**Surveillance system is widely used in the traffic monitoring. The deployment of cameras is moving toward a ubiquitous camera (UbiCam) environment. In our previous study, a novel service, called GPS-VT, was firstly proposed by incorporating global positioning system (GPS) and visual tracking techniques for the UbiCam environment. The first prototype is called GODTA (GPS-based Moving Object Detection and Tracking Approach). For a moving person carried GPS-enabled mobile device, he can be tracking when he enters the field-of-view (FOV) of a camera according to his real-time GPS coordinate. In this paper, GPS-VT service is applied to the tracking of vehicles. The moving speed of a vehicle is much faster than a person. It means that the time passing through the FOV is much shorter than that of a person. Besides, the update interval of GPS coordinate is once per second, it is asynchronous with the frame rate of the real-time image. The above asynchronous is worsen by the network transmission delay. These factors are the main challenging to fulfill GPS-VT service on a vehicle.In order to overcome the influence of the above factors, a back-propagation neural network (BPNN) is used to predict the possible lane before the vehicle enters the FOV of a camera. Then, a template matching technique is used for the visual tracking of a target vehicle. The experimental result shows that the target vehicle can be located and tracking successfully. The success location rate of the implemented prototype is higher than that of the previous GODTA.

Keywords**—**visual surveillance, visual tracking, global positioning system, intelligent transportation system

I. Introduction

THE evolution of hardware and information communication technology (ICT) causes the development of Internet-based surveillance system. The installation of an IP-based camera is based on the current Internet infrastructure and thus increases the accessibility and maintainability. Cameras are widely installed at streets or communities. That is, the deployment is moving toward a ubiquitous camera (UbiCam) environment. For the Kaohsiung City in Taiwan as an example, 74.5 cameras are installed on average per square kilometers and totally over 15,000 cameras are installed at the streets.

On the other hand, the global positioning system (GPS) is the most popular system for the outdoor environment. Currently, a commercial differential GPS (DGPS) receiver can provide accuracy to within less than a few meters. GPS is widely used in various applications, such as car navigation, fleet management system. However, current GPS-related tracking system mainly displays the locations of vehicles on the digital map. It is difficult to realize the actual situation without the visual information.Therefore, the GPS is firstly incorporated with the visual information of cameras to provide a novel GPS visual tracking (GPS-VT) service in our previous studies. The first prototype is called GODTA (GPS-based Moving Object Detection and Tracking Approach) [1-2]. Its screen shot is shown in Fig. 1. The GPS location of the moving object is used to locate the target moving object on the real-time image of a camera. Such visual information shown on the right-hand side of Fig. 1 is useful to realize the actual situation of the target object.

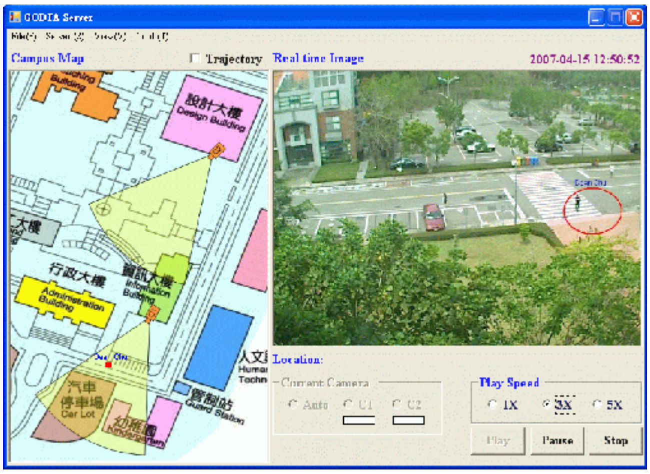


Fig. 1 The screen shot of GODTA prototype

In this study, the GPS-VT service is attempted to fulfill on the tracking of a vehicle. The moving speed of a vehicle is over ten times faster than a moving person. For the GODTA, the vehicle is easily outside the possible area marked with a red circle. It is the main challenge on fulfilling such a service. Three factors influence the synchronization of the GPS location and the vehicle in the real-time image of a camera.

1. Moving speed: When the vehicle is moving fast, the real-time GPS location is definitely asynchronous with the location in the real-time image of a camera.
2. Update interval: Currently, the update interval of a GPS receiver is once per second. However, the frame rate of the real-time image is 30 FPS (frames per second). That is, the update interval causes the asynchronous of GPS location and the vehicle in the real-time image.
3. Transmission delay: When a vehicle transmits its GPS location to a GPS-VT server. The transmission delay also causes the asynchronous of the GPS location and the vehicle in the real-time image. For the general GPRS

H. C. Liao is with the Department of Computer Science and Information Engineering, Chaoyang University of Technology, Wufong District, Taichung, Taiwan (R.O.C.) (phone: 886-4-23323000~4211; fax: 886-4-23742375; e-mail: hcliao@cyut.edu.tw).

Y. S. Wang is with the Department of Computer Science and Information Engineering, Chaoyang University of Technology, Wufong District, Taichung, Taiwan (R.O.C.) (e-mail: s9727604@cyut.edu.tw).

(General Packet Radio Service), its delay is longer than that of a Wi-Fi connection as shown in Fig. 2. The average transmission delay of a GPRS connection is 0.485 second.

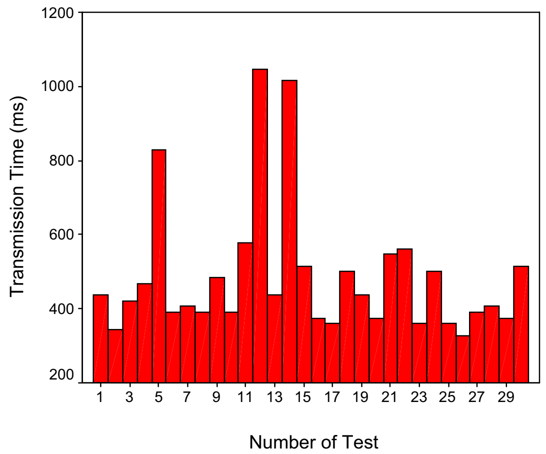


Fig. 2 The transmission delay of GPRS connection

In order to overcome the influence of the above factors, a back-propagation neural network (BPNN) is used to predict the possible lane of the target vehicle to appear. Then, the target vehicle can be located once it enters the field-of-view (FOV) of the camera. Then, a template matching technique is used for the tracking of the target vehicle.The remainder of this paper is organized as follows: Section 2 presents the related works. Section 3 presents the operation process of the GPS-VT for vehicles. Section 4 presents the implemented prototype and experimental study, and Section 5 gives the conclusion and suggestions for future works.

II. Related works

Object tracking is an important task of surveillance system. Many studies proposed methods for object tracking under various types of objects, conditions, or cameras. These studies can be classified into the following three categories:

A. GPS Tracking

In this category, object is tracking simply based on their GPS locations. For example, L. Kane et al. proposed a vehicle tracking system in public transport domain [3]. A passenger can access the real-time location of a vehicle, e.g., bus, by using public display, mobile device, notebook, and so on. It is expected to increase the usability and improve air pollution of a public transport system. G. Derekenaris et al. proposed a path planning approach for ambulances by integrating the GIS (Geographic Information System), GSM (Global System for Mobile communication), and GPS technologies [4]. Besides, W. H. Lee et al. proposed a system, called VPS (Vehicle Positioning System) for electronic toll collection (ETC) [5]. The deployment of VPS is simple and cheap than the current Infrared or microwave system.

B. Visual Tracking

The visual tracking is based on the real-time image of a camera. For the visual tracking of vehicles, the installation height of a camera influences the design and performance of the proposed approach. If the height is too small, the vehicles are easily to collide to each others. It increases the difficulty on locating and tracking of vehicles. So, the camera is assumed installed at a suitable height. For example, N. K. Kanhere and S. T. Birchfield proposed a method based on a set of feature points to segment vehicles when the collision occurs [6]. In additional to the collision condition, the change of illumination or shadow also influences the tracking performance. Therefore, J. Zhou et al. proposed a method by splitting the vehicle image into a set of blocks [7]. Then, the block matching based on their histograms are used to keep the tracking to be successful. Besides, the vehicle counting function based on the tracking technology is also studies in recent years for the intelligent transportation system (ITS) [9-10]. Such information can be used to analyze the traffic condition.

C. GPS Visual Tracking (GPS-VT)

In our previous studies, GPS is first incorporated with the camera to provide a GPS visual tracking services [1]. A moving person can be tracking visually based on his real-time GPS coordinate. Such a method is based on the transformation from GPS coordinate to the image coordinate. A camera must be calibrated to obtain the intrinsic and extrinsic parameters for coordinate transformation, including the focal length (f), translation (T) and rotation (R) matrices. An automatic calibration method is also proposed to overcome the troublesome on manual calibration [2]. The GPS-VT service is efficient since the GPS coordinate restricts the image area to be processed. It is still working when the target object is totally occluded or when the illumination is very low.According to the above description, GPS-VT is suitable for the vehicle tracking. Especially many vehicles has installed GPS receiver for car navigation or fleet management. Several advantages are expected for fulfilling GPS-VT on vehicle tracking: Firstly, GPS can be the common coordinate system of all the cameras. It simplifies the step on establishing the spatial relationship among cameras. Secondly, the design of some mechanisms becomes simple, such as the handover mechanism when a vehicle moves through cameras or the labeling of the same vehicle on the FOV of different cameras. Lastly, the plate number of a vehicle can be transmitted with its coordinate. It means that the recorded video can be tagged with the number and useful for video retrieval.

III. GPS-VT for vehicles

In this section, the process of the GPS-VT service for vehicles is presented. The target vehicle is basically tracking based on its GPS coordinate before it enters the camera's FOV. Once it enters the FOV, a template matching technique is used to locate and keep tracking the target vehicle. Therefore, the process is divided into three phases: initialization, GPS tracking, and visual tracking as shown in Fig. 3.

A. Initialization Phase

The moving speed of a vehicle is fast. Therefore, the prediction of the possible lane to appear is very important in this study. In this phase, several steps must be done for the lane prediction and the operation of the later phases, including the road lane detection and the learning of back-propagation neural network (BPNN). They are presented below.

1. Road lane detection

The lane detection is done by using the following steps. Firstly, the original image is processed by using color filtering technique to preserve the bright colors and converted to black and white image. The white pixel is searched from the bottom of the image. If a while pixel can be found, another white pixel with a specific distance in the range of ±30 degrees is searched again. If a white pixel cannot be found, the range ±30 degrees is increased to ±35 degrees. Then, the desired lane can be found by connecting these found while pixels. An example of road lane detection is shown in Fig. 4. Fig. 4(a) is the original image. Fig. 4(b) is the result of color filtering and binarization. Fig. 4(c) is the result of detected lane.

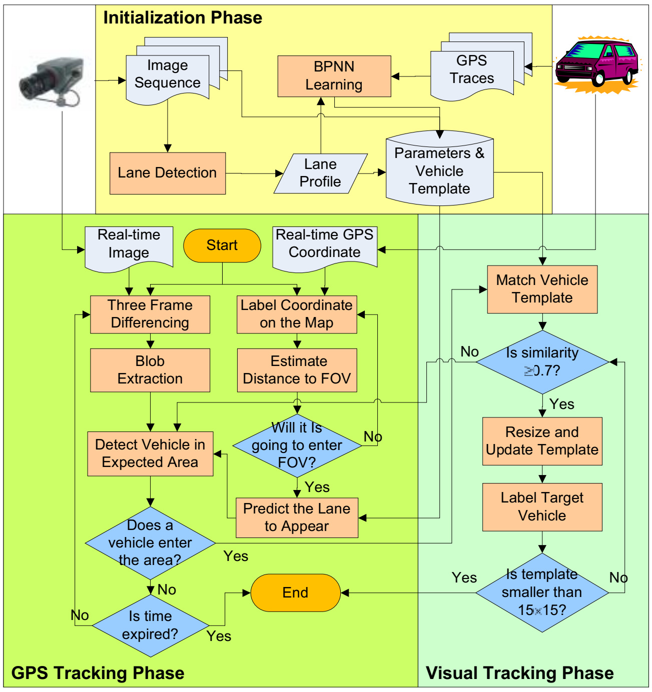
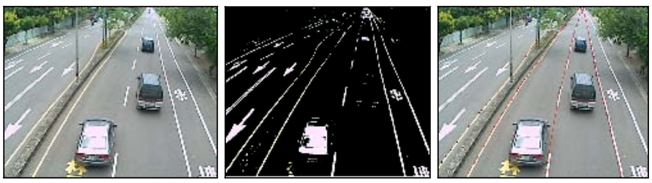


Fig. 3 The operation process of GPS-VT for vehicles



(a) (b) (c)

Fig. 4 The demonstration of road lane detection (a) original image (b) color filtering and binarization (c) the detected lane

2. Back-propagation neural network (BPNN)

In order to overcome the asynchronous problem caused by the high moving speed, a BPNN is used to predict the possible lane of a vehicle. The BPNN consists of three layers. The possible lane is predicted according to the last GPS coordinate before entering the FOV of a camera. Therefore, the input layer is the last GPS coordinate and the output layer is the corrected lane number. The nodes of the hidden layer are operated in constructive algorithm [9]. The number of hidden neurons in the hidden layer is increased until the mean error square (MES) is stable. Function of hidden layer and output layer is Sigmoid function as shown in Eq. (1). The BPNN model is depicted in Fig. 5.

$$f(x) = \frac{1}{1+e^{-x}} \quad \text{...} (1)$$

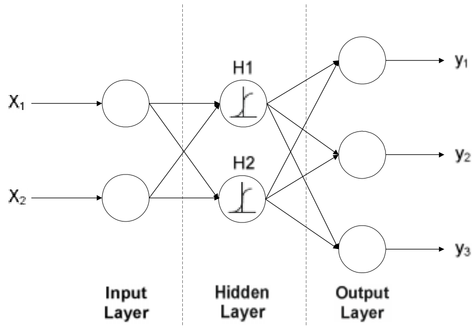


Fig. 5 The model of three-layered back-propagation neural network

B. GPS tracking phase

In this phase, the system track the target vehicle according to received GPS coordinate. The GPS coordinate is used to mark the location of the vehicle on the digital map, e.g. Google map. The distance of the received GPS coordinate to the FOV is computed. And, the expected distance to FOV is computed according to the moving speed of the vehicle. The possible lane of the vehicle to appear is also predicted by using the BPNN in this phase.When the vehicle is expected to enter the FOV in the next second, the corresponding area of the predicted lane is checked whether a vehicle appears. When a vehicle is found in the area, a pre-stored template of the target vehicle is matched with the detected object by using conventional library [11]. If the similarity of template matching is larger or equal to 0.7, it means the target vehicle is located. The matched object is deemed as the new template and the process moves to the next phase.

C. Visual Tracking Phase

In this phase, the closest object in front of the moving direction of the located target is performed the template matching. The template is updated when the similarity is large or equal to 0.7. However, two factors influence the result of template matching. One is that the size of the target vehicle is reduced quickly; the other is that the vehicle may occlude with the other vehicles. If the size of the template does not update simultaneously, the visual tracking is easily failed in the middle of tracking. The following steps are used to solve the above problem.

(1) If the width or height of the object to be matched is less or equal to 1.2 times of the template, the template is updated as this object if the similarity of template matching is large or equal to 0.7. The size of the template will be reduced as the matched object.

(2) If the width or height of the object is larger than 1.2 times of the template, it means that the collision condition

occurs. The current template is reduced and matched with the object continuously until the similarity of template mathcing is not higher than the matching of the previous size.

The above steps are performed until the template is smaller than 15×15.

IV. Experimental Study

A prototype system of GPS-VT for vehicles was implemented using Microsoft Visual C# 2008. The camera is AXIS 207W and installed at a pedestrian bridge with a height about 4.6 meters. The distance of the road in the FOV of the camera is about 155 meters. The screen shot of the prototype is shown in Fig. 6. The left-hand side is the real-time image. The blue rectange above each lane is the area to activate template matching function. When the target vehicle is located, it is marked with a red rectangle. Its plate number is also shown above the upper left corner of the rectangle. The digital map, the prediction result, and templates are shown on the right-hand side from top to bottom.

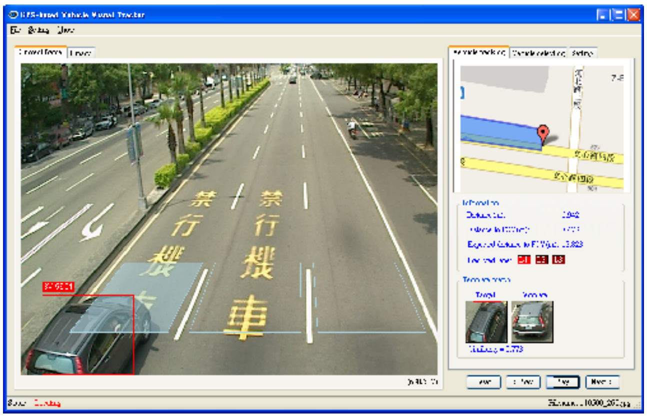


Fig. 6 The prototype's screen shot of GPS-VT for vehicle

The GPS-VT service for a vehicle is demonstrated by the prototype in Fig. 7. In Fig. 7(a), the GPS coordinate of a target vehicle is received and marked on the digital map. The process is in the GPS tracking phase. The corresponding distance to the FOV of the camera is also calculated, e.g., 13.394 meters. Then, an expected distance to FOV is calculated according the speed of the vehicle. In Fig. 7(b), the vehicle is going to enter the FOV. The lane prediction result of BPNN is shown on the screen, e.g., L1. From now on, the process enters the visual tracking phase. For every vehicle enters the blue rectangle, it is matched with the pre-stored template of the target vehicle. When the result of the template matching is higher than a predefined threshold, e.g. 0.7, the vehicle is located as the target. Then, the target vehicle is keep tracking as shown in Fig. 7(c). The tracking is ended until its size is smaller than 15×15 pixels. The moving trajectory is also displayed on the image as shown in Fig. 7(d).

(a) (b) (c) (d)

Fig. 7 The demonstration of GPS-VT for a vehicle (a) GPS tracking phase (b) the last second before entering FOV (c) target vehicle is located (d) visual tracking result

In advance, an experiment was designed to evaluate the performance of this prototype and the previous prototype, GODTA. The target vehicle driving through the pedestrian bridge was repeated and recorded for 30 times. There are three lanes in the road. 10 times were recorded for every lane. The network transmission delays from 0 to 800 milliseconds are added to the transmission of GPS coordinates. Then, the locating success rate of the target vehicle is computed and listed in Table I. The success rate of GODTA can achieve 97 percent under no delay condition. However, the success rate is decreased significantly as the increase of the delay. For the delay, 400 milliseconds, in the general case depicted in Fig. 2, the success rate of GODTA is only 63 percent. Oppositely, the success rate of the GPS-VT for vehicle is kept 100 percent without the influence of the transmission delay. It shows that the proposed approach can overcome the asynchronous problem caused by the high moving speed of vehicles.

TABLE I
EXPERIMENTAL RESULTS

Methods	GODTA		GPS-VT for Vehicles	
Delay(ms)	Locating time (ms)	Success Rate (%)	Locating time (ms)	Success Rate (%)
0	466	97%	30.1	100
200	604	90%	32.2	100
400	822	63%	32.2	100
600	1071	46%	32.2	100
800	1243	16%	42.6	100

V. CONCLUSION AND FUTURE WORKS

In our previous study, the GPS-VT service is designed for tracking a moving person. In advance, the service is designed for tracking the vehicle in this paper. In order to overcome the asynchronous problem between the GPS coordinate and real-time camera image, a BPNN is used to predict the possible lane of the target vehicle to appear. The experimental results also show that the proposed GPS-VT for vehicles service can locate and track the target vehicle successfully and precisely. However, the service relies on the training of BPNN. The training step will be automated for realizing GPS-VT service in UbiCam environment with thousands of camera in the future.Besides, the GPS-VT service for vehicles will be fulfilled on a pan-tilt-zoom (PTZ) camera. Unlike the fixed camera in this study, the size of the target vehicle becomes more and more small or large as the vehicle leaving or moving toward the camera. The zoom-in capability of a PTZ camera can keep the size of the vehicle in an almost constant size and increase the tracking distance. It is helpful for improving the quality of GPS-VT service.

REFERENCES

[1] H. C. Liao and P. T. Chu, “*A Novel Visual Tracking Approach Incorporating Global Positioning System in a Ubiquitous Camera Environment*,” Information Technology Journal, Vol. 8, No. 4, 2009, pp. 465-475.

[2] H. C. Liao and H. J. Wu, “*Automatic Camera Calibration and Rectification Methods*,” Measurement + Control Journal, Vol. 43, No. 8, Oct. 2010, pp. 251-254.

[3] L. Kane, B. Verma, and S. Jain, “*Vehicle Tracking in Public Transport Domain and Associated Spatio-temporal Query Processing*,” Computer Communications, Vol. 31, Issue 12, July 2008, pp. 2862-2869 .

[4] G. Derekenaris, J. Garofalakis, C. Makris, J. Prentzas, S. Sioutas, and A. Tsakalidis, “*Integrating GIS, GPS and GSM technologies for the Effective Management of Ambulances*,” Computers, Environment and Urban Systems, Vol. 25, Issue 3, May 2001, pp. 267-278.

[5] W. H. Lee, S. S. Tseng, and C. H. Wang, “*Design and Implementation of Electronic Toll Collection System based on Vehicle Positioning System Techniques*,” Computer Communications, Vol. 31, Issue 12, July 2008, pp. 2925-2933.

[6] N. K. Kanhere and S. T. Birchfield, “*Real-Time Incremental Segmentation and Tracking of Vehicles at Low Camera Angles Using Stable Features*,” IEEE Transactions on Intelligent Transportation Systems, Vol. 9, Issue 1, 2008, pp. 148-160.

[7] J. Zhou, D. Gao, and D. Zhang, “*Moving Vehicle Detection for Automatic Traffic Monitoring*,” IEEE Transactions on Vehicular Technology, Vol. 56, Issue 1, 2007, pp. 51-59.

[8] B. T. Morris and M. M. Trivedi, "*Learning, Modeling, and Classification of Vehicle Track Patterns from Live Video*", IEEE Transactions on Intelligent Transportation Systems, Vol. 9, Issue 3, 2008, pp. 425-437.

[9] J. W. Hsieh, S. H. Yu, Y. S. Chen, and W. F. Hu, “*Automatic Traffic Surveillance System for Vehicle Tracking and Classification*,” IEEE Transactions on Intelligent Transportation Systems, Vol. 7, Issue 2, 2006, pp. 175-187.

[10] T. Y. Kwok and D. Y. Yeung, “*Constructive Algorithms for Structure Learning in Feedforward Neural Networks for Regression Problems*,” IEEE Transactions on Neural Networks, Vol. 8, Issue 3, 1997, pp. 630-645.

[11] AForge.Net Framework: http://code.google.com/p/aforge