한국도로학회 논문집
第10券1號 2008年 6月
pp.159~156

원거리 차량 추적 감지 방법

Methodology for Vehicle Trajectory Detection Using Long Distance Image Tracking

오 주 택* Oh, Ju Taek
민 준 영** Min, Joon Young
허 병 도*** Heo, Byung do

요 지

최근 교통감시시스템은 실시간의 영상검지시스템(VIPS)을 가장 선호하고 있으며, 그 수요는 매년 증가하고 있는 추세이다. 일반적으로 영상검지시스템은 공간기반의 검지알고리즘을 사용하고 있으며, 교통량, 속도, 점유율 등의 교통정보를 제공하고 있다. 현재 전 세계적으로 이미 상용화되어 있는 대부분의 영상검지시스템들은 Tripwire기반의 검지영역 내 차량의 존재유무를 판단하여 교통정보를 수집하는 알고리즘으로 구성되어 있으나, 개별차량에 대한 검지는 불가능한 한계를 갖고 있다. 반면 개별차량의 추적시스템은 보다 구체적인 공간적 교통정보를 제공할 수 있어 사고검지, 급차선 변경 등 교통정보를 보다 다양화 할 수 있다는 장점이 있으나 추적길이가 불과 100미터이내이며, 그 이상 관측하기 위해서는 운영자가 카메라를 줌인을 하여 영상을 확대하여야 한다. 따라서 본 논문에서는 차량 추적의 효과를 높이기 위해서 기존의 100미터 이내 추적거리를 여러 대의 CCTV시스템을 이용하더라도 200미터이상으로 확대함으로써 사고 또는 비정상적 차량흐름을 검지할 수 있는 알고리즘을 제안한다.

핵심용어 : 지능형교통시스템, 원거리 차량추적, 비디오 영상처리시스템, 폐쇄회로 티비, 배경영상 차영상

Abstract

Video image processing systems (VIPS) offer numerous benefits to transportation models and applications, due to their ability to monitor traffic in real time. VIPS based on a wide-area detection algorithm provide traffic parameters such as flow and velocity as well as occupancy and density. However, most current commercial VIPS utilize a tripwire detection algorithm that examines image intensity changes in the detection regions to indicate vehicle presence and passage, i.e., they do not identify individual vehicles as unique targets.

If VIPS are developed to track individual vehicles and thus trace vehicle trajectories, many existing transportation models will benefit from more detailed information of individual vehicles. Furthermore, additional information obtained from the vehicle trajectories will improve incident detection by identifying lane change maneuvers and acceleration/deceleration patterns. However, unlike human vision, VIPS cameras have difficulty in recognizing vehicle movements over a detection zone longer than 100 meters. Over such a distance, the camera operators need to zoom in to recognize objects. As a result, vehicle tracking with a single camera is limited to detection zones under 100m.

This paper develops a methodology capable of monitoring individual vehicle trajectories based on image processing. To improve traffic flow surveillance, a long distance tracking algorithm for use over 200m is developed with multi-closed circuit television (CCTV) cameras. The algorithm is capable of recognizing individual vehicle maneuvers and increasing the effectiveness of incident detection.

Keywords : ITS, Tracking, Long Distance Tracking, VIPS, CCTV, Background Subtraction

* Associate Research Fellow, The Korea Transport Institute, South Koera, e-mail: jutaek@koti.re.kr
** Professor, Sangji Youngseo College, South Korea, e-mail: wmjy@paran.com (교신저자)
*** Visiting Professor, Sangji Youngseo College, South Korea, email: byungdo@paran.com

1. INTRODUCTION

ITS technology includes advanced traffic management systems, advanced traveler information systems, advanced public transportation systems, and advanced sensor systems for on-line surveillance, such as traffic video analysis (Chen et al., 2003).

The quest for better traffic information, and the consequent increasing reliance on traffic surveillance, has increased the need for better vehicle detection such as wide-area detectors. Meanwhile, the high costs and safety risks associated with lane closures have directed the search towards non-invasive detectors mounted beyond the edge of the pavement. One promising approach is vehicle tracking via video image processing, which can yield traditional traffic parameters such as flow and velocity. Recently, closed circuit television (CCTV), video image processing systems (VIPS) and probe cars have been used for traffic monitoring and traffic data collection. Especially, wide area CCTV surveillance systems have been extensively deployed to monitor freeways in urban areas. However, while CCTVs have proven to be very effective in monitoring traffic flows and supporting incident management, they simply provide images that must be interpreted by trained operators (Namkoong et al., 2004).

Spatial traffic information such as the vehicle tracks or trajectories can be more useful than spot information at a single point, because it is possible to measure true density instead of simply recording detector occupancy. In fact, by averaging trajectories over space and time, the traditional traffic parameters are more stable than corresponding measurements from point detectors, which can only be averaged over time. Additional information from the vehicle trajectories could lead to improved incident detection, both by detecting stopped vehicles within the camera's field of view and by identifying lane change maneuvers and acceleration/deceleration patterns that are indicative of incidents beyond the camera's field of view (Coifan et al., 1998).

However, unlike human vision, the VIPS cameras have difficulty in recognizing vehicle movements over a detection zone longer than 100 meters. Over such a distance, the camera operators need to zoom in the camera to recognize objects. As a result, vehicle tracking with a single camera is limited to detection zones under 100m. Therefore, most current VIPS operate with the tracking length of about 70m at most, which only allows 2~5 seconds for a vehicle to pass the detection zone. This is too short for incidents detection on the roads. For taking prompt or necessary action when an incident occurs, the tracking length has to be sufficiently long. Generally, a CCTV camera has a visible area of about 120~150m with operator's naked eye. At greater distances, the operators must zoom in to be able to see clearly. Therefore, vehicle tracking with a single camera is limited to under 70m.

This paper develops a methodology capable of monitoring individual vehicle trajectories based on image processing. To improve traffic flow surveillance, a long distance tracking algorithm, over 200m, with multi-CCTV cameras is developed. The algorithm is capable of recognizing individual vehicle maneuvers and increasing the effectiveness of incident detection. Experiments were conducted on image data captured with two video cameras installed on a downtown street in Wonjucity in South Korea.

2. PRIOR RESEARCHES AND STATE OF PRACTICE

The previous image processing and object tracking techniques have been mostly applied to traffic video analysis to address queue detection, vehicle classification and volume counting (Chen et al., 2003).

From the computer vision literature, the different tracking approaches for video data can be classified as 1) Model-based tracking, 2) Region-based tracking, 3) Active contour-based tracking, and 4) Feature-based tracking(Coifan et al., 1998).

Model-based tracking (Koller et al., 1993) is highly accurate (Ed- alternatively, "has high accuracy") for a small number of vehicles. The most serious weakness of this approach, however, is the reliance on detailed geometric object models. It is unrealistic to expect to be able to have detailed models for all vehicles on the roadway.

In region-based tracking, the process is typically initialized by the background subtraction technique. Most commercial VIPS, such as CMS Mobilizer, PEEK VideoTrack, and Nestor TrafficVision, track vehicles using region-based tracking, i.e., vehicles are segmented based on movement within a short, 70m-length (Coifman et al., 1998). This approach works fairly well in free-flowing traffic. However, these systems suffer the problem of one target occluding another when two targets become merged together by the tracking software. Thus, under congested traffic conditions, vehicles partially occlude one another instead of being spatially isolated, which increases the difficulty in segmenting individual vehicles.

Complementary to the region-based approach, active contour-based tracking is based on active contour models or snakes. The basic idea is to have a representation of the bounding contour of the object and keep updating it dynamically. The advantage of having a contour-based representation instead of a region-based one is reduced computational complexity. However, the inability to segment vehicles that are partially occluded remains. If a separate contour could be initialized for each vehicle, then each one could be tracked even in the presence of partial occlusion (Koller et al., 1994a).

An alternative approach to tracking abandons the idea of tracking objects as a whole and instead tracks sub-features such as distinguishable points or lines on the object. The advantage of the feature-based tracking approach is that even in the presence of partial occlusion, some of the features of the moving object remain visible. Furthermore, the same algorithm can be used for tracking in daylight, twilight or night-time conditions. It is self-regulating because it selects the most salient features under the given conditions, such as window corners, bumper edges during the day and tail lights at night.

3. DEVELOPMENT OF Long Distance Tracking ALGORITHM

This chapter explains the basic idea of the long distance algorithm developed from this research.

For long distance tracking, multiple cameras need to be used. However, the problem of multiple cameras is that when individual cameras present overlapping images, the occurrence of an intersection area between the overlapped images is unavoidable. For example, two CCTV cameras

A and B were used connectively in this research. The total surveillance area was around 260m, and the intersection area was around 20m as shown in Figure 1.

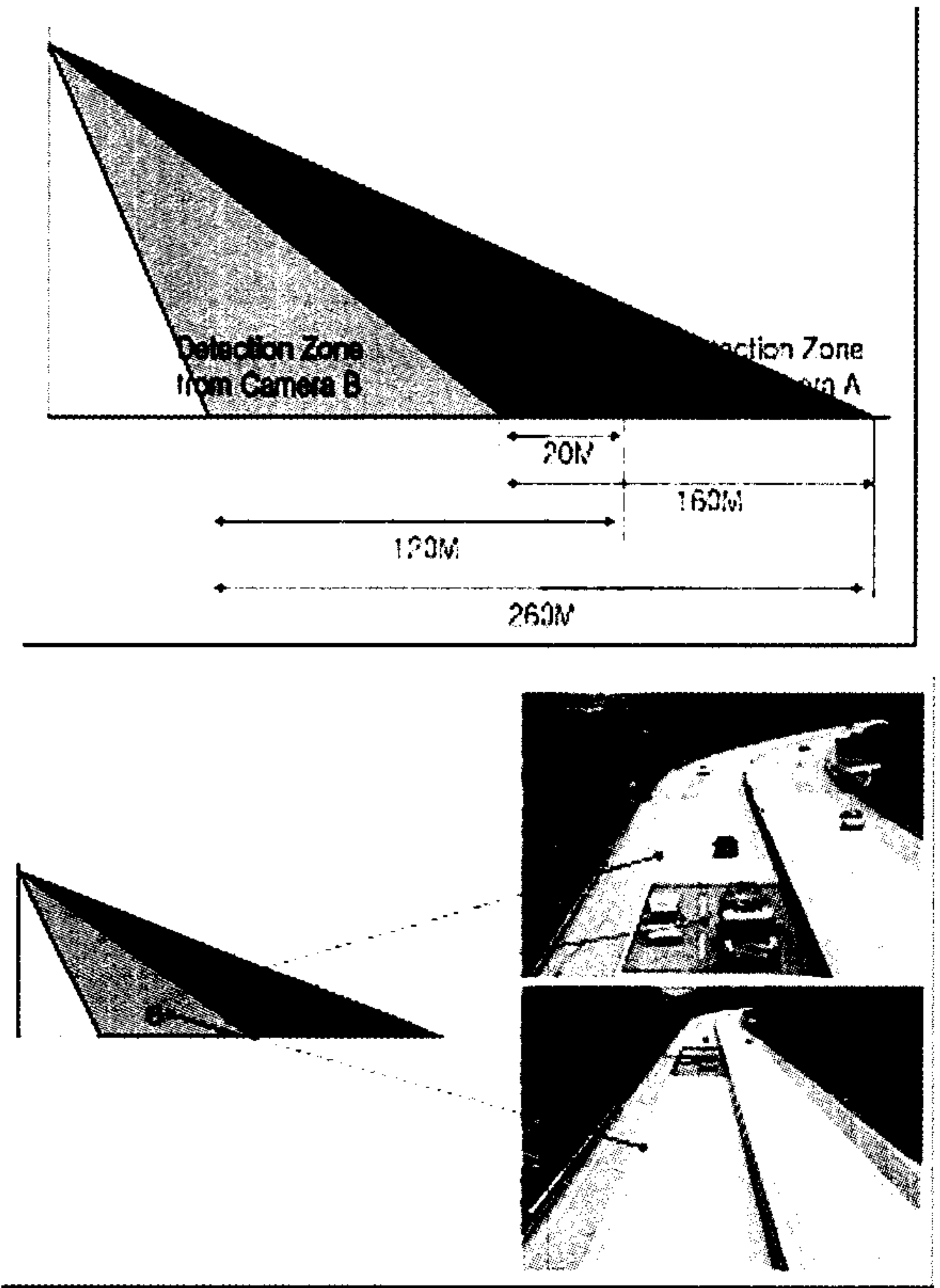

Figure 1. The feature based as means to improve surveillance performance, long distance tracking, total surveillance area is about 260m

For long distance tracking, we conducted the following four steps, as shown in Figure 2. First moving objects have to be extracted exactly using image processing, background subtraction, threshold and morphology, after which the tracking algorithm is applied.

The basic concept of long distance tracking involves processing the connectivity between multi-cameras of each individual vehicle on the intersection area.

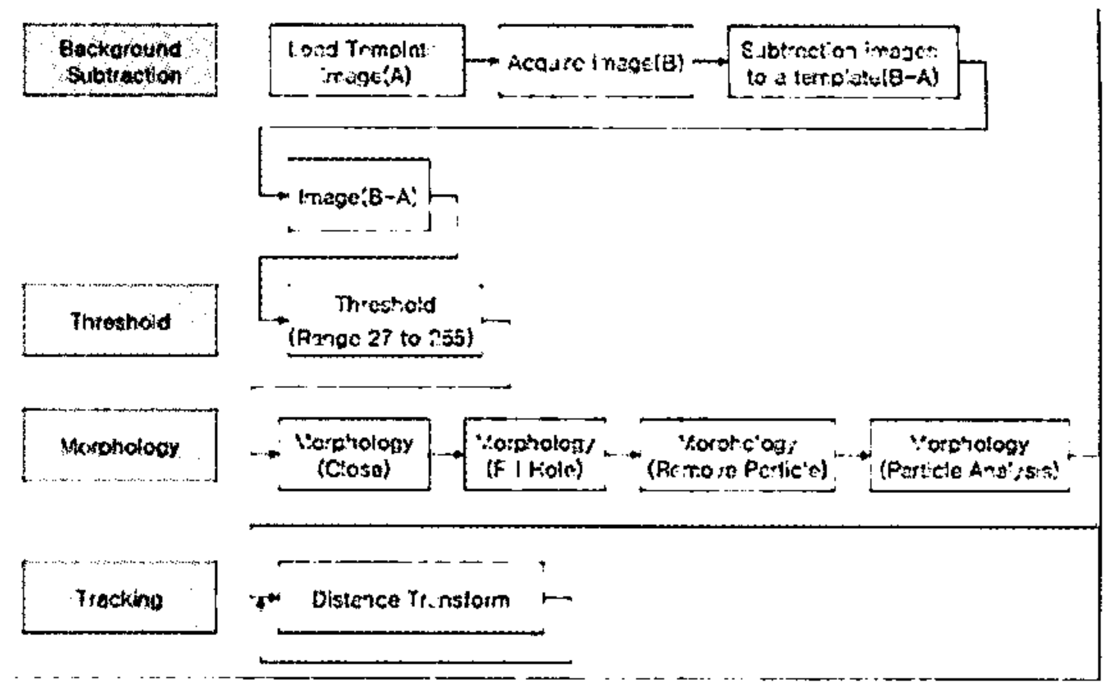

Figure 2. The processing connection flow for multi-cameras

Firstly, in moving object extraction the vehicles passing through the detection area are used by video background subtraction algorithm, the background template, $f(x, y, t_0)$ in the detection area is saved beforehand and the current frames $f(x, y, t_i)$are taken, and the differences of the two images are calculated pixel by pixel. A difference image between two images taken at time t_0 and t_i may be defined as Eq. (1) and an example is presented in Figure 3.

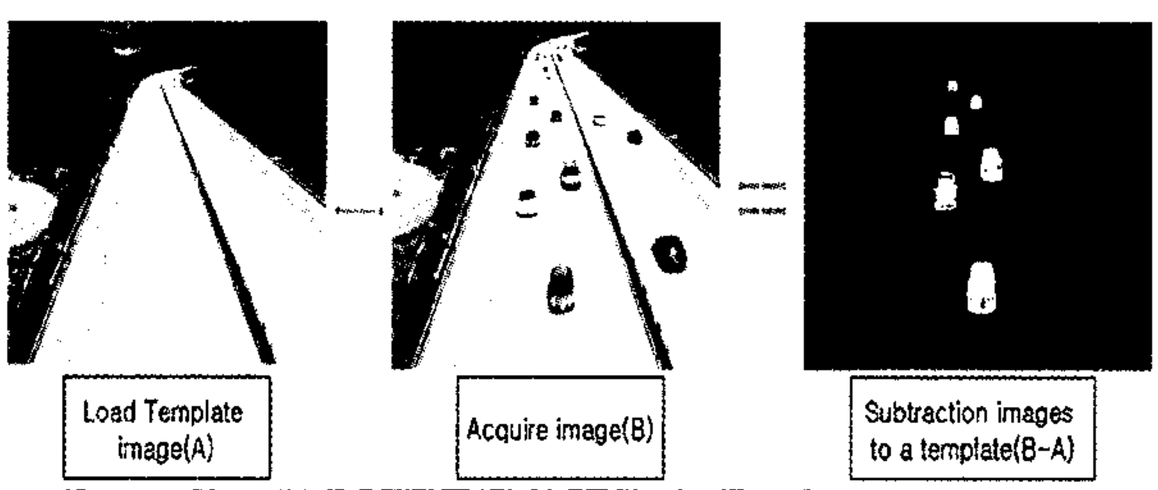

Figure 3. Background subtraction for extracting individual vehicle from road

Secondly, the threshold for binarization is defined (see Eq. (1)). In an ideal case, the histogram with a color or gray distribution has a deep and sharp valley between two peaks representing the objects and background. However, for most real images it is often difficult to detect the valley bottom precisely (Otsu, 1979).

We have chosen the threshold heuristically based on the experimental position, because the threshold exhibits wide variance depending on each position and time. Therefore, the optimal threshold can be obtained by running the experiments several times. In this experiment, the threshold was set up at a gray level of 27.

$$d_{0,i}(x,y)=\begin{cases}1 & \text{if} \mid f(x,y,t_0)-f(x,y,t_i) \mid > \Theta \\ 0 & otherwise\end{cases} \tag{1}$$

Thirdly, mathematical morphology is a tool for extracting image components that are useful in the representation and description of a region's shape, such as boundaries, skeletons, and the convex hull (Gonzalez et al., 1992).

As shown in Figure 4, the morphology process in this research consists of three methods: closing by using dilation of a 3×3 block, filling in the hole with vehicles, and removing very small objects from this frame when considered to be noise. After the moving objects in this frame are extracted, drawing the MOR(Moving Object Region), the center and (x,y), (top,left), (right, bottom) of MOR coordinates of each object can be obtained, and then save these information to reference table(RF).

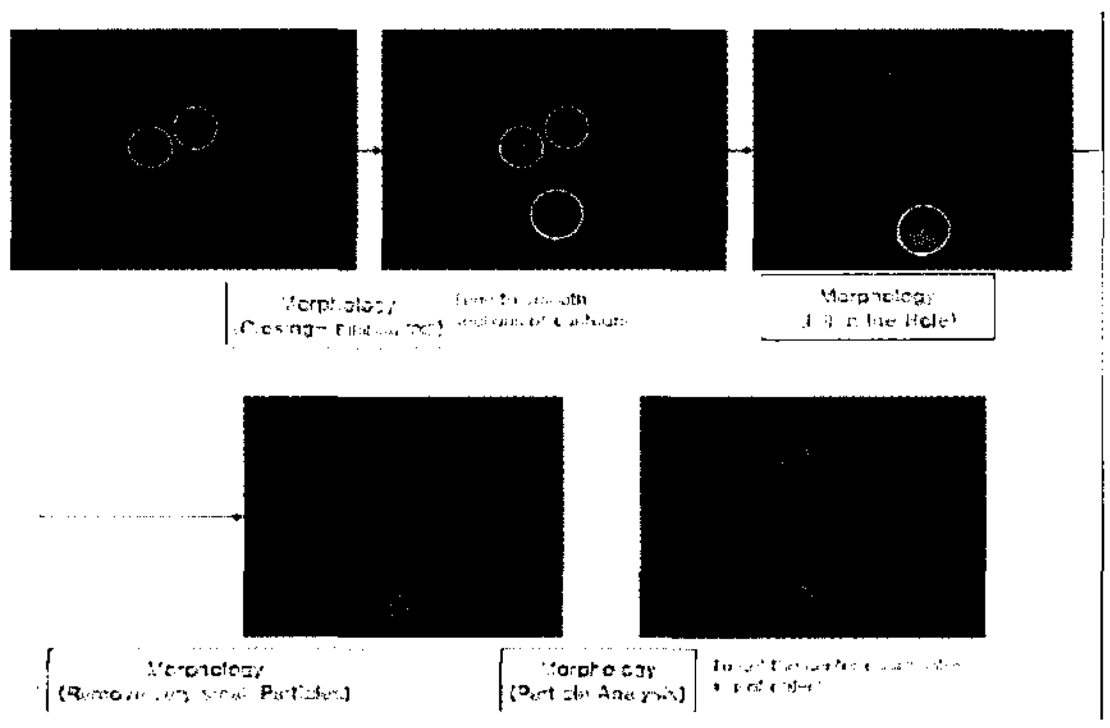

Figure 4. The morphology process for extracting moving objects

Finally, the most important work of long distance tracking is connectivity with each vehicle inter cameras. In long distance tracking process with multiple cameras, if a field test has been conducted with two cameras, A and B, the image is acquired from camera A, assigned an identity, referred to as ID throughout the rest of this paper, and the individual vehicle is tracked and entered into area by saving the ID and coordinates to the reference table. At this time, all indexes in the reference table(RT) are assigned a default value of 0. More detailed illustrate to grant the vehicle ID from camera A, vehicle ID can be decided to grant according to existing the ID in the RF.

For giving vehicle ID, comparing with frame i and frame i+1, if vehicle ID is not in the RT at frame i, the new ID and MOR coordinates have been generated, considering as new vehicle entering into detection zone A. And the other side, if vehicle ID is existed in the RT, the MOR coordinates (x,y), (top, left), (right, bottom) in the RT is updated, considering as passing vehicle within detection zone in frame i-1. The same ID have been sustained with frame i and frame i-1 by calculating minimum distance of MOR center.

In the next, monitoring with camera A and camera B within intersect area, the same ID have been detected from camera A and B simultaneously.

For camera B which focuses on the neighboring area of the camera, the vehicle ID and coordinates are extracted from RT, and the value of the index is changed to 9999 to enable camera B to sustain the tracking of this ID's vehicle.

And the other side, in the absence of any reference from camera B, this index is assigned 0, if vehicle exit the detection zone B from camera B.

The processing flow of vehicle tracking with multiple cameras is shown as Figure 5.

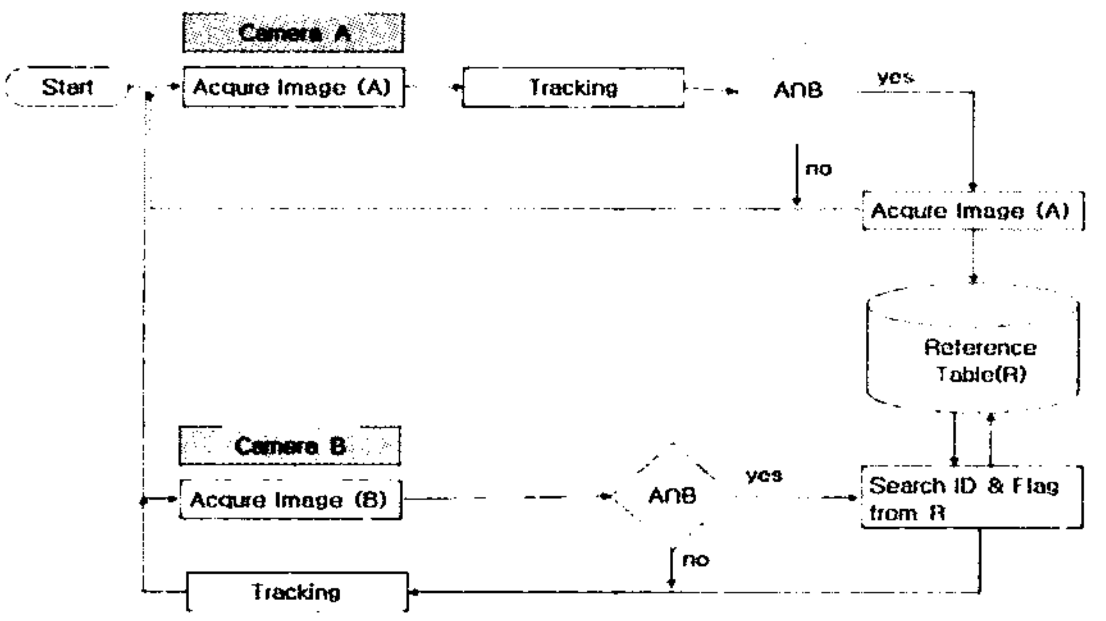

Figure 5. The process of long distance tracking process with multi cameras

4. EXPERIMENTAL RESULTS AND DISCUSSION

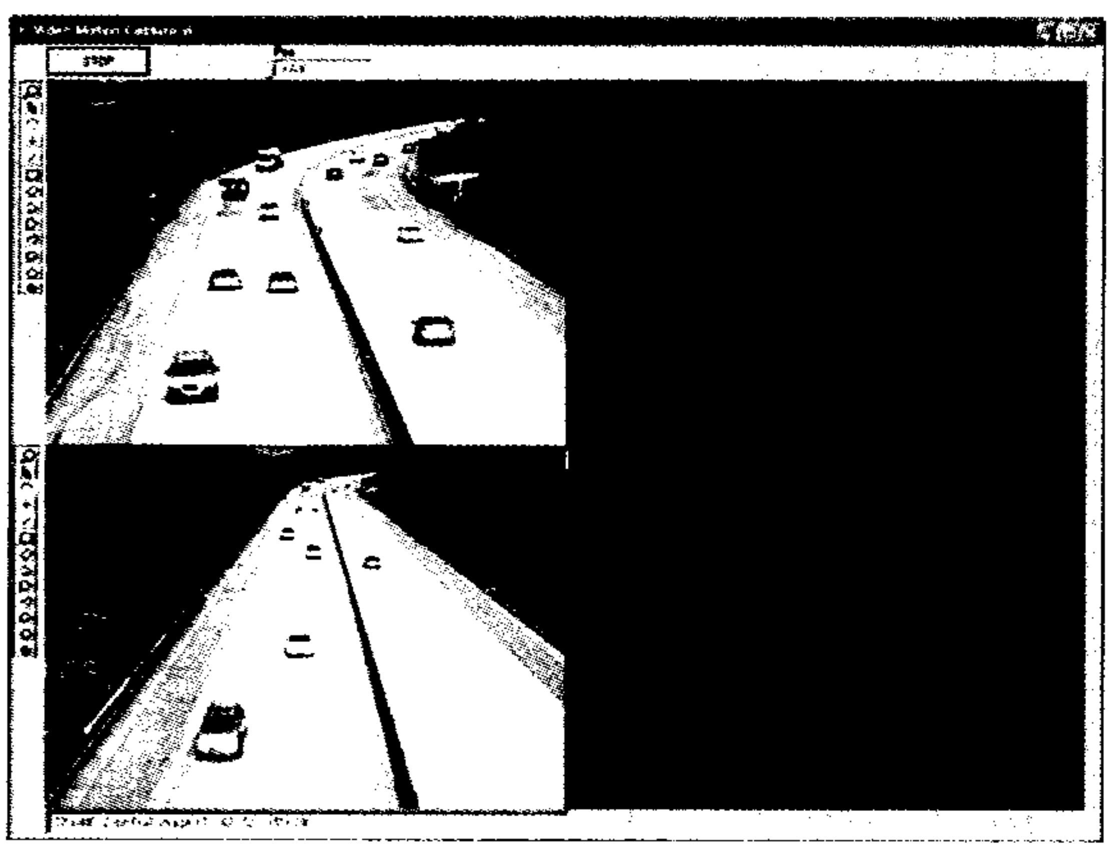

Figure 6. The example of vehicle tracking images with dual camera A and B

To run an experiment of long distance tracking to improve traffic flow surveillance, a multiple-camera installation is ideal. However, in the absence of any such suitable site, we installed multiple cameras at one position temporarily.

The field testing was conducted on Sep., 9, 2007, for 15 minutes from 4pm, with two cameras A and B, on Jungang highway in Wonju city in South Korea. Detection areas were divided into two zones: one of around 120m for camera A and the other of 160m for camera B. The intersection area of the two cameras was around 20m. Figure 6 shows the images in the detection zones A and B used for this research.

Figure 7 also shows the results of the experimental tests. The first vehicle, ID 10485, was referenced from camera B, and exited all detection zones. The second and third vehicles, ID 24522 and 89985 respectively, were tracked by camera B, referred from the reference table, and the index value was changed to 9999. On the other side of the road, the last vehicle to enter the detection zone, ID 38774, was not added into the reference table, and was assigned an index value of 0.

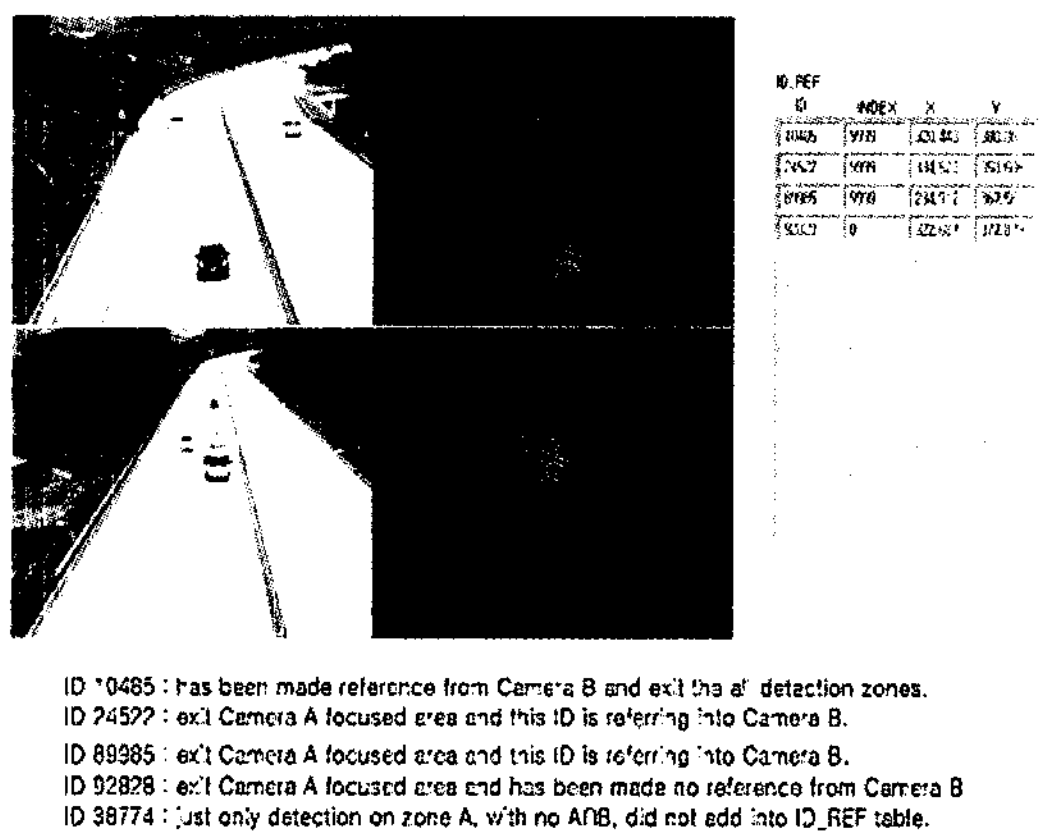

Figure 7. Experimental result, Jungang highway near by Wonju city in South Korea

To verify these test results, we also obtained the measuring items of traffic information such as volume, speed and occupancy rate for each one minute unit. Volume was accumulated by passing through each detection zone respectively for one minute, while the speed and occupancy rate were calculated from the average value for 1 minute.

Table 1. Measuring traffic information from field testing of long distance tracking in Zone A and B, about 260m

Time (min.)	Zone A			Zone B			Zone A+B			Error
	Vol. (a)	Speed (km/h)	Occ. (%)	Vol.	Speed (km/h)	Occ. (%)	Vol. (b)	Speed (km/h)	Occ. (%)	Vol. |a-b
1	11	98	8.6	11	101	5.9	11	103	7.4	0
2	21	101	16.3	21	105	10.9	21	106	14.5	0
3	12	94	11.0	13	95	8.9	13	103	7.4	1
4	12	103	9.9	13	102	8.2	13	103	9.2	1
5	21	92	21.2	21	97	14.2	21	98	19.2	0
6	13	95	10.7	13	101	5.0	13	97	7.0	0
7	13	102	11.0	13	101	7.1	13	99	9.8	0
8	18	105	14.3	18	109	8.7	18	104	12.0	0
9	15	99	12.5	15	101	8.2	15	103	10.4	0
10	10	94	8.7	12	97	5.7	12	95	9.0	2
11	20	93	18.3	20	100	6.9	20	96	11.8	0
12	11	93	10.2	11	101	5.4	11	101	8.6	0
13	13	91	12.8	13	91	8.8	13	101	7.8	0
14	25	102	21.0	25	110	11.6	25	99	18.1	0
15	21	91	20.0	22	93	11.1	22	104	13.1	1

The test results are shown in Table 1. A total of 236 and 241 vehicles passed through detection zones A and B, respectively. The difference was caused by occlusion within zone A when a large vehicle occluded a smaller vehicle as it passed, thereby removing it from the vehicle count, as described above. Zone A had more errors than zone B, because occlusion occurred more frequently in the more distance area.

In actuality, 243 vehicles passed through the two zones. The two missed vehicles equated to a total error rate of 0.8%. The error within the measured data was 2.07%: 236 vs. 241 vehicles.

5. CONCLUSIONS

Most problems of tracking techniques arise due to vehicle occlusions and inaccuracies in separating individual vehicles. Although traffic information such as volume, speed and occupancy rate can be obtained simply by these approaches, vehicle flow patterns cannot be estimated because of the very short detection length.

Most problems of commercial systems indicate that two targets may become merged together if one moving target, including its shadow, occludes another. However, tracking length problems have not been discussed in research papers or commercial systems.

The objective of this research was to relate traffic safety to VIPS tracking. The proposed long distance tracking algorithm will help to understand individual vehicle maneuvers and increase the effectiveness of incident detection for improved safety monitoring.

Generally, a CCTV camera has a visible area of about 120~150m with the operator's naked eye, whereas the operator must zoom in at greater distances. Therefore, vehicle tracking is limited to under 70m with a single camera and multi-CCTV camera systems are required for understanding the traffic flow efficiently in cases of accident or unforeseen disturbances in long distance tracking of over 200m.

Analyzing traffic data from long distance tracking will improve safety monitoring by providing a warning message of dangerous driving patterns to drivers, automatic traffic monitoring and incident detection in real time, and, especially, data on frequent incidents.

In future work, various analysis methods based on these data will be developed and applied to field tests, especially, we have conducted the analysis with other detectors, loop detectors, IR detectors, and other tripwire detectors several times for the verifying our system.

Furthermore, more than two cameras will be applied to an extended tracking distance of over 300~400m.

* Unfortunately, we could not obtain the observed value of speed and occupancy rate, because the loop or video detectors was not installed on the experimental points. Therefore, the errors were calculated just only from the measured volumes vs the observed ones.

ACKNOWLEDGEMENTS

This research was supported by a grant (06 Transportation Core C01) from Transportation System Efficiency Program funded by Ministry of Construction & Transportation of Korean government.

REFERENCES

B. Coifman, D. Beymer, P. McLauchlan, J. Malik, (1998) "A Real-Time Computer Vision System for Vehicle Tracking and Traffic Surveillance," *Transportation Research Part C, pp.271-288.*

E. J. Park, K. Y. Lee, J. H. Lee, M. S. Jang, *"Development of Vision-Based Vehicle Tracking for Extracting Microscopic Traffic Information,"* working paper.

Koller, D., Daniilidis, K., Nagel, H., (1993) "Model-based object tracking in monocular image sequences of road traffic scenes,". *International Journal of Computer Vision 10, pp.257-281.*

Koller, D., Weber, J., Huang, T., Malik, J., Ogasawara, G., Rao, B., Russell, S., (1994b) "Towards robust automatic traffic scene analysis in real-time," *Vol.1, ICPR, Israel, pp.126-131.*

N. Otsu, (1979), "A Threshold Selection Method from Gray level Histogram," *IEEE Transaction on Systems, Man, and Cybernetics Vol.SMC-9, No.1, pp.62-66.*

Rafael C. Gonzalez, Richard E. Wood, (1992), *"Digital Image Processing"*, Addison Wesley.

S. Kamijo, Y. Matsushita, K. Ikeuchi, (Jun., 2000) "Traffic monitoring and Accident Detection at intersections," *IEEE Trans. on Intelligent Transportation Systems, Vol. 1, pp.108-118.*

S.-C. Chen, M.-L. Shyu, S. Peeta, C. Zhang, (Sep., 2003) "Learning-Based Spatio-Temporal Vehicle Tracking and Indexing for Transportation Multimedia Database System," *IEEE Trans. on Intelligent Transportation Systems, Vol. 4, No. 3, pp.154-167*

S.-C. Chen, M.-L. Shyu, C. Zhang, (Aug., 2001) "An Intelligent Framework for Spatio-Temporal Vehicle Tracking," Proc. *4th Int. IEEE Conf. Intelligent Transportation Systems, Oakland, CA, pp.213-218.*

S. M. Smith, J. M. Brady, (1995) "Asset-2: Real-time motion segmentation and shaping Tracking," *IEEE Trans. on Pattern Anal. Machine Intell., Vol. 17, pp.814-820.*

S. Namkoong, H. Tanikella, B. Smith, (Jul., 2004) "Design and Field Evaluation of a System Integrating CCTV Surveillance with Video Image Vehicle Detection Systems (VIVDS)", *Paper submitted for presentation at the 2005 Annual Meeting of the Transportation Research Board and publication in the Transportation Research Record.*

접 수 일: 2008. 3. 17
심 사 일: 2008. 3. 21
심사완료일: 2008. 4. 25