Successive Observations of Trajectories of Vehicles with Plural Video Cameras

Hirokazu AKAHANE [*1] Satoshi HATAKENAKA[*2]
Professor, Faculty of Engineering, Chiba Institute of Technology [*1]
(2-17-1 Tsudanuma, Narashino-shi, Chiba 275-8588, JAPAN, TEL: +81-47-478-0444,
E-mail: akahane(at)ce.it-chiba.ac.jp)
Narita International Airport Corporation[*2]

In this study, a system was developed for successive observations of trajectories of vehicles with images of plural video cameras installed adjacently along roads through the use of techniques such as image processing and geometry transformation. This system can more highly equilibrate the observation precision and the observation range than the use of images of a single video camera. In addition, it can exclude the influence of occlusion where a vehicle is shadowed by another in video images. Furthermore, an observation equation was derived in combination of equations of triangles that approximate road surfaces with equations of a 3D projective transform. The measurement accuracy of this method was evaluated in comparison with results of a precision survey of trajectories of a 1/22 scale vehicle model.

Keywords: *vehicle tracking, vehicle behavior, image processing, Kalman smoother*

1. Introduction

In ITS, safe driving support systems that provide drivers with safety information, warning or automatic control that overrides manual driving maneuvers in order to avoid danger are required so as not to confuse the drivers as well as others who are driving in their vicinity and not to disturb traffic flow as a whole.

These requirements for part-time automatic vehicle control systems are totally different from requirements for full–time automatic control systems because the former have to override manual control in case of emergency avoiding panicking either the drivers or the passengers in the automatically controlled vehicles as

Table 1. Comparison of techniques for observing behavior of running vehicles

Observation techniques / Comparative items	Traffic detector	Measurement car with sensors	Video camera with image processing
Vehicles observed	Unspecified and many	Measurement car and its neighboring cars	Unspecified and many
Successive observations of trajectories of vehicles	Not possible	Possible	Possible but limited to coverage of video image
Method for positioning running vehicles	Detection of passage of vehicles at detector locations	GPS, etc.	Transform of image coordinates
Method for measuring running speeds of vehicles	Calculations based on time differences of passage of vehicles sensed by pairs of detectors and intervals of locations of them	Transform of rotation speeds of wheels, use of spatial filter, etc.	Not possible (Calculation based on differences in positions)
Method for measuring running accelerations of vehicles	Not possible	Acceleration sensors of distortion gauge type, gyroscope type, etc	Not possible (Calculation based on differences in positions)
Difficulty in measurement	Easy (Installation of additional detectors is difficult)	Easy	Difficulty in data acquisition for transform of coordinates and in installation of additional cameras

well as people in the neighboring vehicles that might still be manually driven. It is indispensable to grasp manual driving maneuvers, running behavior of the vehicles as their results, and interactions among vehicles with high precision in order to satisfy such requirements.

In this study, a system for successive observations of trajectories of vehicles with images of plural video cameras installed adjacently along roads was developed in order to construct a database of serial numerical data of positions, velocity, acceleration and so on that represent microscopic behavior of individual vehicles.

Table 1 shows characteristics of the observation by video images in comparison with other observation methods. The observation system developed in this study can more highly equilibrate the observation precision and the observation range than using images of a single video camera. In addition, it can exclude the influence of occlusion by framing a vehicle with plural video cameras at different angles at the same moment.

2. Basic framework of observation system

The basic composition of this system is the following:

2.1. Synchronous acquisition of video images

Black burst signals and SMPTE time code signals are supplied from a main camera to sub-cameras via daisy chain connections in order to synchronize image acquisition in each camera.

As for observations from open-air sites, GPS-based synchronizers [1] that generate black burst signals and SMPTE time code signals from GPS signals are used in order to synchronize acquisition of images by video cameras installed at separated sites with no cable-connection, for instance, from one rooftop to another.

2.2. Acquisition of video image coordinates

The TrackPoint [2] was used for the acquisition of the video image coordinates of vehicles in each frame. This application automatically tracks objects designated by an operator when they come in a video image. It has an interface by which an operator who monitors video images can manually modify automatic tracking when it strays away from an actual trajectory.

So-called features such as corner parts of license plates and bodies have to be chosen as tracking targets because of requirements for image processing. Points where tires contact road surfaces have the advantage of transform of image coordinates because they are on the road surfaces. However, those points are not necessarily practical because they are frequently shadowed by other vehicles in congested traffic flow when video images are acquired from the front and the rear of vehicles.

2.3. Transform of image coordinates

Video image coordinates of vehicles acquired by the method mentioned in section 2.2 were transformed into world coordinates in images of each video camera independently. In our past study, a 2D projective transform was applied supposing that vehicles moved on a single plane [3]. The parameters of the transform were identified based on both image coordinates and world coordinates of the required number of Ground Control Points (GCPs) on a road surface.

The errors in transform of coordinates depend on the positions of vehicles in video images because of the influence of characteristics, locations, and angles of cameras. The evaluations of the variances of the errors were applied to the smoothing algorithm mentioned below.

2.4. Matching of vehicles between images of adjacent cameras

Each vehicle has to match between images of adjacent cameras in order to connect and smooth its trajectories that consist of world coordinates transformed through the use of the method mentioned in section **2.3** in images of each camera.

When a vehicle is framed by both of a pair of adjacent cameras at the same moment, it can be manually matched, based on the acquired video images. When a vehicle is framed by a pair of adjacent cameras successively, except for a very short spatial gap, it can be manually matched, based on the video images acquired at intervals that are roughly proportional to the speed of the vehicle and the length the spatial gap.

This study applied the manual matching method mentioned above because automatic and/or real-time matching was not necessary for our present laboratory use such as analysis of occurrence processes of traffic accidents.

2.5. Estimation of trajectories of vehicles

The Kalman smoothing algorithm was applied to the least square estimations of unified trajectories based on world coordinates transformed independently in images of each camera through the use of the method mentioned in section 2.3 [3]. The estimation algorithm consists of the two processes as follows:

1) Management of errors in transform of coordinates

The errors evaluated in section **2.3** are processed. When a vehicle is framed by both of a pair of adjacent cameras at the same moment, world coordinates can be rationally unified using the error variance of coordinate transform in each camera image.

2) Management of errors caused by heights of tracking targets from road surfaces

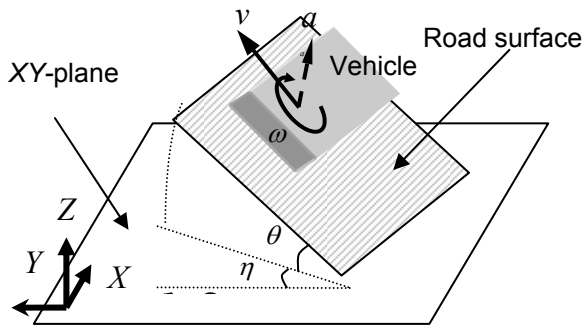

Figure 1. Vehicle motion model

The 2D projective transform in section **2.3** assumes that tracking targets are on the road surfaces. Therefore, the heights of tracking targets from road surfaces result in errors in the 2D projective transform. The errors can be compensated by values of heights.

When a vehicle is framed by both of a pair of adjacent cameras at the same moment, the height of the tracking target of the vehicle can be estimated by a principle similar to stereo photogrammetry. The Kalman smoothing algorithm can compensate the errors caused by the heights that are simultaneously estimated by the algorithm according to the least square criteria as well.

3. Models of vehicle motion and observation

In this study, the basic framework of the observation system mentioned in section **2** was revised as follows:

3.1. The state equation

Figure 1 illustrates the concept of the vehicle motion model newly introduced in this study [4]. Equation (1) indicates the state equation derived from the vehicle motion model where a height *h(t)* of a tracking target from road surfaces was embedded in the Kalman smoothing algorithm as a constant state variable.

$$\mathbf{x}(t+\Delta t) = \mathbf{f}(\mathbf{x}(t)) + \mathbf{q}(t) \qquad (1)$$

where,

$$\mathbf{x}(t) = \begin{pmatrix} X(t) \\ Y(t) \\ \eta(t) \\ v(t) \\ \omega(t) \\ a_l(t) \\ a_r(t) \\ \dot{\omega}(t) \\ \dot{a}_l(t) \\ \dot{a}_r(t) \\ h(t) \end{pmatrix}, \quad \mathbf{q}(t) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \kappa(t) \\ \lambda(t) \\ \mu(t) \\ 0 \end{pmatrix},$$

$$\mathbf{f}(\mathbf{x}(t)) = \begin{pmatrix} X(t) + v(t)\cdot\cos\theta(t)\sin\eta(t)\Delta t \\ \quad + a_l(t)\cdot\frac{1}{2}\cos\theta(t)\sin\eta(t)\Delta t^2 \\ \quad + a_r(t)\cdot\frac{1}{2}\cos\eta(t)\Delta t^2 \\ Y(t) + v(t)\cdot\cos\theta(t)\cos\eta(t)\Delta t \\ \quad + a_l(t)\cdot\frac{1}{2}\cos\theta(t)\cos\eta(t)\Delta t^2 \\ \quad + a_r(t)\cdot\frac{1}{2}\sin\eta(t)\Delta t^2 \\ \eta(t) + \omega(t)\cdot\Delta t + \dot{\omega}(t)\cdot\frac{1}{2}\Delta t^2 \\ v(t) + a_l(t)\cdot\Delta t + \dot{a}_l(t)\cdot\frac{1}{2}\Delta t^2 \\ \omega(t) + \dot{\omega}(t)\cdot\Delta t \\ a_l(t) + \dot{a}_l(t)\cdot\Delta t \\ a_r(t) + \dot{a}_r(t)\cdot\Delta t \\ \dot{\omega}(t) \\ \dot{a}_l(t) \\ \dot{a}_r(t) \\ h(t) \end{pmatrix}$$

t: time

Δt: scanning time

X, Y, Z: world coordinates

v: longitudinal velocity

θ: a gradient of a trajectory

η: a running direction of a vehicle

a_l: longitudinal acceleration

a_r: lateral acceleration

ω: angular velocity

h: a height of a tracking target from road surfaces

κ: random fluctuation in angular accelertion

λ: random fluctuation in longitudinal jerk

μ: random fluctuation in lateral jerk

$\dot{}$: notation for time derivative

3.2. 3D projective transform and identification of its parameters

Equation (2) indicates the 3D projective transform that was introduced instead of the 2D projective transform in order to consider gradients of road surfaces:

$$\left.\begin{aligned} x^j &= \frac{b_1^j X + b_2^j Y + b_3^j Z + b_4^j}{b_9^j X + b_{10}^j Y + b_{11}^j Z + 1} \\ y^j &= \frac{b_5^j X + b_6^j Y + b_7^j Z + b_8^j}{b_9^j X + b_{10}^j Y + b_{11}^j Z + 1} \end{aligned}\right\} \qquad (2)$$

where,

x, y: image coordinates

b_i: parameters

j: a suffix denotes a camera

The parameters of the 3D projective transform were identified by the DLT (Direct Linear Transformation) [5] using world and image coordinates of GCPs that had to be scattered not only on road surfaces, but in 3D space.

3.3. Approximation of road surfaces by triangles and formation of observation equation

Road surfaces that consisted of vertical and horizontal alignments were approximated by triangles whose vertices were GCPs on the surfaces as follows:

$$\alpha^k (X - X_1^k) + \beta^k (Y - Y_1^k) + \gamma^k (Z - Z_1^k) = 0 \qquad (3)$$

where,

$\alpha^k, \beta^k, \gamma^k$: elements of a normal vector of an approximation triangle

X_1^k, Y_1^k, Z_1^k: coordinates of a vertex of an approximation triangle

k: a suffix denotes an approximation triangle

The observation equation (4) was approximately derived when Equation (3) was substituted for Equation (2) eliminating coordinate Z and when h denotes a height of a tracking target from road surfaces:

$$\left.\begin{aligned} x^j(t) &= \frac{A_j^k X(t) + B_j^k Y(t) + C_j^k h(t) + D_j^k}{I_j^k X(t) + J_j^k Y(t) + K_j^k h(t) + L_j^k} + u^j(t) \\ y^j(t) &= \frac{E_j^k X(t) + F_j^k Y(t) + G_j^k h(t) + H_j^k}{I_j^k X(t) + J_j^k Y(t) + K_j^k h(t) + L_j^k} + v^j(t) \end{aligned}\right\} \qquad (4)$$

where,

$$A_j^k = \gamma^k b_1^j - \alpha^k b_3^j$$
$$B_j^k = \gamma^k b_2^j - \beta^k b_3^j$$
$$C_j^k = \gamma^k b_3^j$$
$$D_j^k = \alpha^k b_3^j X_1^k + \beta^k b_3^j Y_1^k + \gamma^k b_3^j Z_1^k + \gamma^k b_4^j$$
$$E_j^k = \gamma^k b_5^j - \alpha^k b_7^j$$
$$F_j^k = \gamma^k b_6^j - \beta^k b_7^j$$
$$G_j^k = \gamma^k b_7^j$$
$$H_j^k = \alpha^k b_7^j X_1^k + \beta^k b_7^j Y_1^k + \gamma^k b_7^j Z_1^k + \gamma^k$$
$$I_j^k = \gamma^k b_9^j - \alpha^k b_{11}^j$$
$$J_j^k = \gamma^k b_{10}^j - \beta^k b_{11}^j$$
$$K_j^k = \gamma^k b_{11}^j$$
$$L_j^k = \alpha^k b_7^j X_1^k + \beta^k b_7^j Y_1^k + \gamma^k b_7^j Z_1^k + \gamma^k b_8^j$$

u^j: measurement error in x imge coordinate

v^j: measurement error in y imge coordinate

The same pairs of Equation (4) as video cameras that simultaneously frame the identical vehicle from different positions and angles are obtained and collectively expressed by the following vectors:

$$\mathbf{y}(t) = \mathbf{g}(\mathbf{x}(t)) + \mathbf{e}(t) \qquad (5)$$

where,

$$\mathbf{y}(t) = \begin{pmatrix} \vdots \\ x^j(t) \\ y^j(t) \\ \vdots \end{pmatrix}, \quad \mathbf{e}(t) = \begin{pmatrix} \vdots \\ u^j(t) \\ v^j(t) \\ \vdots \end{pmatrix},$$

$$\mathbf{g}(\mathbf{x}(t)) = \begin{pmatrix} \vdots \\ \dfrac{A_j^k X(t) + B_j^k Y(t) + C_j^k h(t) + D_j^k}{I_j^k X(t) + J_j^k Y(t) + K_j^k h(t) + L_j^k} \\ \dfrac{E_j^k X(t) + F_j^k Y(t) + G_j^k h(t) + H_j^k}{I_j^k X(t) + J_j^k Y(t) + K_j^k h(t) + L_j^k} \\ \vdots \end{pmatrix}$$

3.4. Application of extended Kalman smoothing algorithm

The extended Kalman smoothing algorithm was applied to the state equation (1) and the observation equation (5) in order to obtain the least square estimator of the state vector $\mathbf{x}(t)$ that includes the height h as a constant and the world coordinates $X(t)$ and $Y(t)$ of a vehicle at every moment.

In the smoothing process, random fluctuation $\kappa(t), \lambda(t), \mu(t)$ and measurement errors $u^j(t), v^j(t)$ were assumed to be white noise. In addition, very large values were set to the variances of measurement errors $u^j(t)$ and $v^j(t)$ when the camera j did not frame a vehicle tracked at time t in order to practically process the information acquired just by other cameras that framed it.

The information of image coordinates of a tracking target of a vehicle at every moment by plural video cameras could be effectively and rationally integrated by the smoothing process considering the errors in transform of coordinates and errors caused by a height of a tracking target from road surfaces.

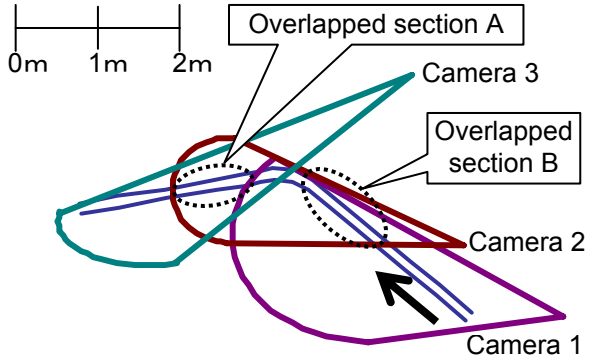

Figure 2. Horizontal alignment of the course and angles of video cameras

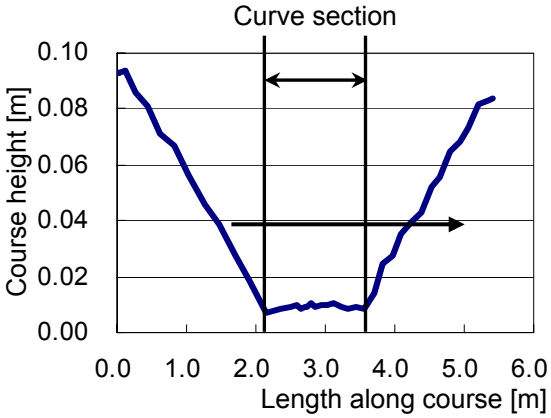

Figure 3. Vertical alignment of course

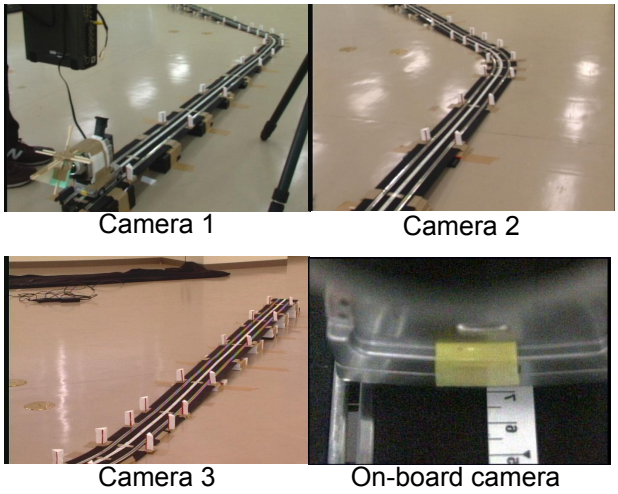

Figure 4. Sample images

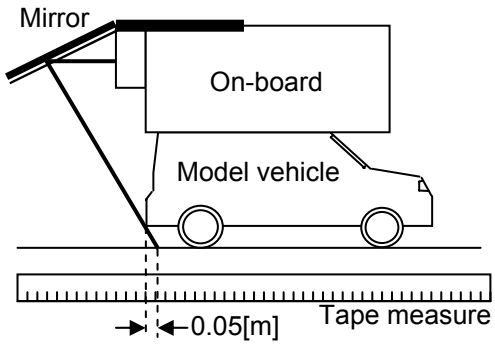

Figure 5. On-board camera and measure

4. Precision verification by a model experiment

In this experiment, video images of a 1/22 scale vehicle model that ran along a model course were acquired using plural video cameras synchronized by the method mentioned in section 2.1. A series of image coordinates of tracking targets was obtained by the method mentioned in section 2.2. Trajectories of the model of a vehicle were estimated by the Kalman smoothing algorithm where the 3D observation equation and the revised state equation mentioned in section 3 were embedded. The results of the estimation were verified in comparison with the precise trajectories measured by surveying instruments.

4.1. Settings in experiment

Figure 2 shows the horizontal alignment of the experiment course as well as the locations and the angles of the video cameras. Figure 3 illustrates the vertical alignment of the course where a flat curve section was located between a descent gradient and an ascent gradient of about 4 percent. The total length of the course was about 100 meters in full-scale. Figure 4 is sample images of the video cameras installed along the course. The model vehicle, only the speed of which was radio-controlled, ran along the guidance groove installed in the center of the course.

Figure 5 illustrates the video camera and the mirror carried by the model vehicle. This on-board camera that was syncronized with the other three cameras installed along the course aquired images of the divisions of the tape scale pasted on the course. The positions of the vehicle model at every 6 video frames that was equivalent to every two-tenth of a second were specified by reading values of the divisions of the tape scale from the video images of the on-board camera. The specified positions were precisely surveyed by a total station afterwards.

The positions of the GCPs higher than the course surfase as well as the GCPs located on the course were precisely surveyed in order to improve identification of parameters of the 3D projective transform.

4.2. Results of precision verification

Figure 6 illustrates triangles that approximate the surface of the course as well as the trajectory estimated

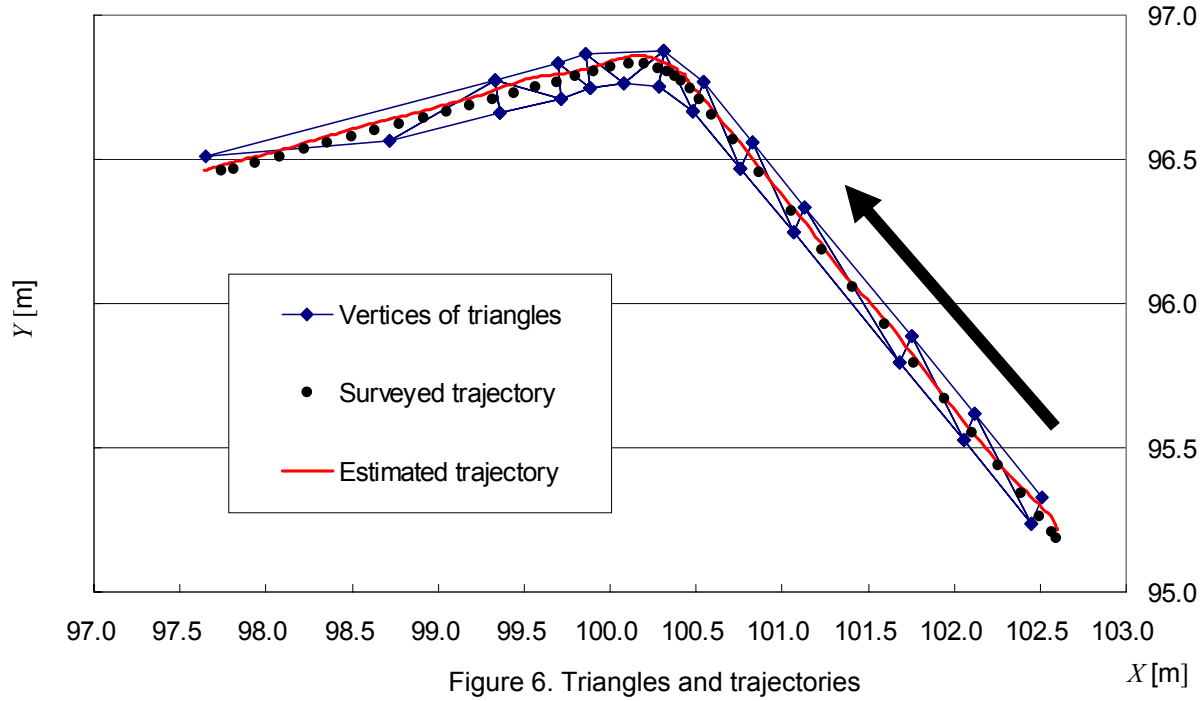

Figure 6. Triangles and trajectories

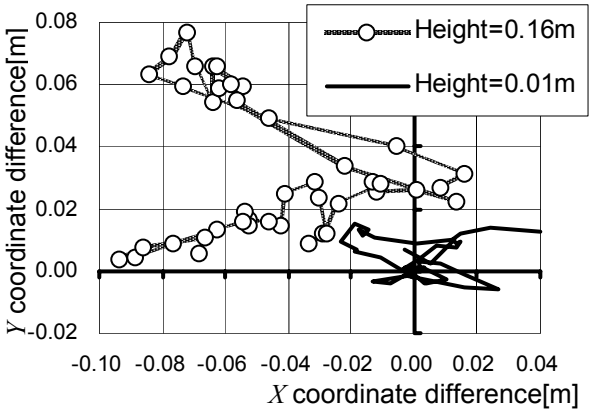

Figure 7. Coordinate differences between estimations and precise surveys

based on image coordinate values of the tracking target that was 0.16 meters in height from the course surface, together with the precise trajectory measured.

Figure 7 indicates differences in X and Y coordinates between the estimation and the precise surveys according to the heights of the tracking targets from the course surface. The differences tended to increase in proportion as the heights of the tracking targets increased especially in the gradients at both ends of the course. However, the absolute values of the differences were less than about 0.1 meters which was equivalent to about 2 meters in full-scale.

Figure 8 indicates speeds of the model vehicle calculated from differences of estimated coordinate values as well as the speeds calculated from differences of precisely surveyed values. The right vertical axis of

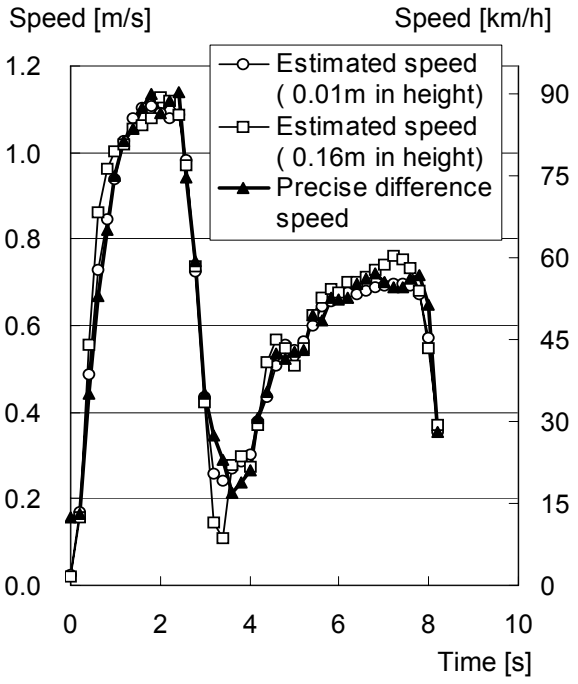

Figure 8. Speeds of model vehicle

Figure 8 shows the converted speeds of the model vehicle in kilometers per hour into full-scale.

The considerable change in speeds during a short period of time could be well estimated. No remarkable bias occurred, despite the defferences in positions. However, the accuracy of the estimation of the speeds tended to decrease in proportion as the heights of the tracking targets increased, as did the accuracy of the estimated positions.

Figure 9 illusrates the differences in X and Y

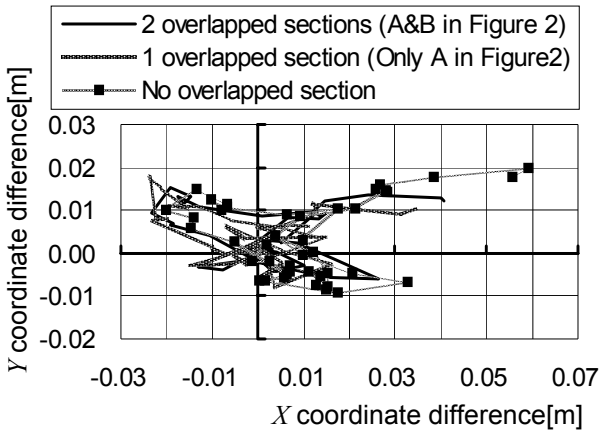

Figure 9. Coordinate differences according to the number of overlapped sections

coordinates between the precise surveys and the estimation where the tracking target was 0.01 meters in height from the course surface according to the number of overlapped sections shown in Figure 2. In the estimation with two overlapped sections, all of the observations, including the overlapped observations by the two cameras individually in sections A and B, were used. In the estimation with one overlapped section, the overlapped observations only in section A were used. In the estimation with no overlapped section, the observations by all cameras, except for one of two cameras in the two overlapped sections, were used. There was no remarkable increase in the differences of the estimation with one overlapped section in the positions in comparison with the estimation with no overlapped section. That meets the principle that the least square estimator of the height of the tracking targets from the course surface can be obtained based on observations in even one overlapped section.

5. Conclusions

The system was developed for successive observations of trajectories of vehicles with images of plural video cameras installed adjacently along roads through the use of techniques such as image processing and projective transform.

The equations of triangles that approximated road surfaces were combined with the equations of the projective transform in order to form the observation equation. The Kalman smoothing algorithm where the observation equation was embedded was applied to the observed data of the 1/22 scale model of a vehicle that ran on the experiment course. As a result, this system can more highly equilibrate the observation precision and the observation range than the use of images of a single video camera.

For further study, this system will be validated on real roads and will be applied as well to microscopic and dynamic analysis of vehicle behavior in traffic accidents, in diverging, in weaving, and so on.

6. Acknowledgments

This study was a part of the research project, “Basic and up-to-date research on ITS,” that the National Institute for Land and Infrastructure Management, the Ministry of Land, Infrastructure and Transport, entrusted to the Institute of Industrial Science, the University of Tokyo. The authors would like to express appreciation to both of the institutes, especailly to the members of the reseach project.

7. References

[1] http://www.totsu-eng.co.jp/
[2] http://www.gen.co.jp/product/index.html
[3] H. Akahane: “Vehicle Tracking with multiple Video Cameras,” *Proceedings of the 37th Symposium on Infrastructure Planning and Management*, JSCE, pp. 89-96, 1998.
[4] R. Horiguchi, H. Akahane and K. Funabashi: “Estimation of 3D trajectories of probe cars and driving behavior by the extended Kalman smoothing,” *Proceedings of the 57th Annual Meeting of JSCE*, CD-ROM, 2003.
[5] Japan Society of Photogrammetory and Remote Sensing: *Analytical Photogrammetory (Revised Edition)*, JSPRS, 1997.

Prof. Hirokazu Akahane Received a Doctor of Engineering degree from the University of Tokyo in 1986. Professor of Chiba Institute of Technology from 1995. Major research fields are road traffic control and management. Member of JSCE, JSTE and IATSS.

Satoshi Hatakenaka Received a Master of Engineering degree from Chiba Institute of Technology in 2004. Member of Narita International Airport Corporation from 2004.

- *Received date*
- *Received in revised form (if any)*
- *Accepted date*

- *Editor*