MVA2000 IAPR Workshop on Machine Vision Applications, Nov. 28–30, 2000, The University of Tokyo, Japan

1—5 Road Traffic Surveillance with Multiple Views

Tameharu Hasegawa and Heitou Zen *
Tokyo University of Mercantile Marine

Abstract

Recently, since the car is often used as the movement means, the traffic is increasing every year. ITS (Intelligent Transport Systems) which provides drivers with plural information for smooth and safety drive has been required. As one of them, several studies have been proposed for the road traffic surveillance in the image processing. The traditional methods have almost observed the traffic flow from one viewpoint with one camera or stereo-cameras. In this case, it is difficult to detect the part of one vehicle occluded by another one with single view. Our proposed method for the road traffic surveillance extracts vehicles with multiple views. Extracted motions in the different viewpoint are complementary to each other, and compared by using the estimated camera parameters of which cameras are placed freely by road managers, *and so on.* As a result, the occlusion of vehicles is detected. Our implemented method has been tried to real traffic scenes and successfully detected the individual vehicle.

1 Introduction

Recently, since the vehicle is often used as the movement means, the traffic is increasing every year, and the problems of the traffic jam and accident and so on are increasing. ITS (Intelligent Transport Systems) which provides the drivers with plural information for smooth and safety drive has been required to solve these problems.

As one of them, several studies have been proposed for the road traffic surveillance in the image processing. In the traditional methods, since they have almost observed the traffic flow from one viewpoint with one camera or stereo-camera, one vehicle has been occluded by another vehicle. It is difficult to extract it from one viewpoint.

Our proposed method uses multiple views to solve the occlusion in the traffic flow observation. When the occlusion is captured with one view, the occluded vehicle is extracted with another view.

On the other hand, the placement of cameras are not always placed at the desirable position for the vehicle extraction because of problems from the road construction and the landscape, for example, cameras are set on the wall where the drivers cannot view them. Therefore, we propose a method to extract the vehicles with multiple views on condition that the cameras are placed freely by the road managers, and so on.

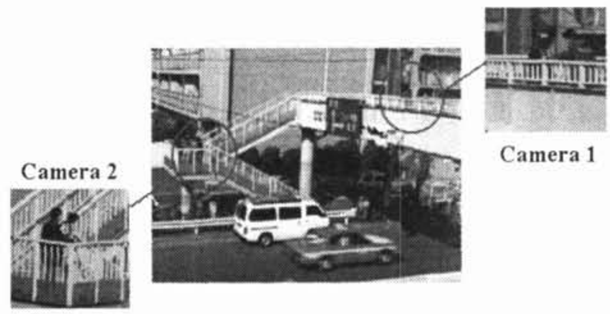


Figure 1: Example of placement with multiple views.

Without using the features depending on the road such as the lane markers of white lines, our method estimates the geometrical relation of viewpoints. Our useful features are the movement vector, the front line, and the edge of vehicle.

We define the model of the movement of vehicles in the three dimensional coordinates, whereas the motions with the multiple views are combined. As a result, the model parameters are estimated geometrically, and the occlusion of vehicle is extracted by using the model parameters.

Our implemented method has been tried to real traffic scenes and successfully detected the individual vehicle.

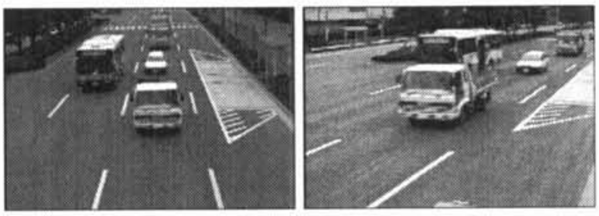
2 Estimation of the camera parameters

In our work, cameras are placed as "freely" as possible. Figure.1 shows an example of the scene of taking pictures. Figure.2 shows the occluded state to be recognized finally.

2.1 Definition of the coordinate system

We aim at the movement direction of vehicle observed at each viewpoint, which is the same direction

*Address: 2-1-6 Etchujima Koto-ku, Tokyo 135-8533 Japan. E-mail: {hasegawa, zen}@ipc.tosho-u.ac.jp



(a) Camera 1 (b) Camera 2

Figure 2: Example of images at the far separated viewpoints.

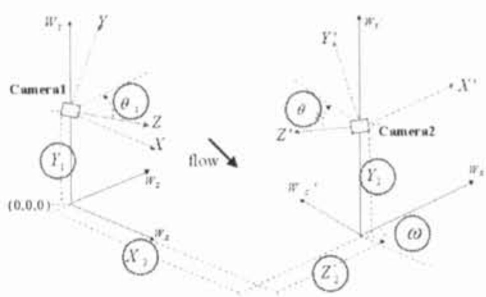


Figure 3: Road coordinate system and camera coordinate system.

in three dimensional space. We consider the direction parallel with the road.

Note that we may estimate thier parameters with good condition for the image processing, when the traffic quantity is little and the vehicle in the image is only one.

We define the road coordinate system and the camera coordinate system as shown in Figure 3, where the marks enclosed by circle is the parameters obtained by our method.

The road coordinate system for camera 1 is defined as (W_X, W_Y, W_Z). The plane $W_X W_Z$, which corresponds to the road plane, is assumed horizontal.

Then, it supposes that only the focus length f_1 of camera1 and f_2 of camera 2 have been acquired beforehand.

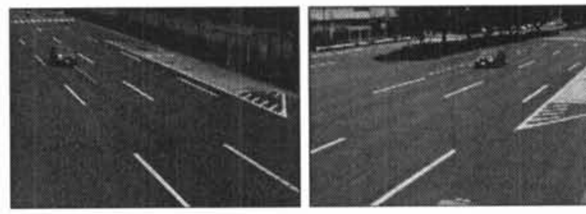
2.2 Estimation of parameters

2.2.1 Outline of the estimation

The representative flow vector in Figure 4 is obtained after the error flows are rejected from the optical flows. We referred to [1][2] for the detection of the optical flows.

The error flows are removed by using the derection and length of flows, and the result of the edge detection.

The process flow of the estimation is shown in Figure 5.



(a) Camera 1 (b) Camera 2

Figure 4: Representative flow of which length is ten times of original length.

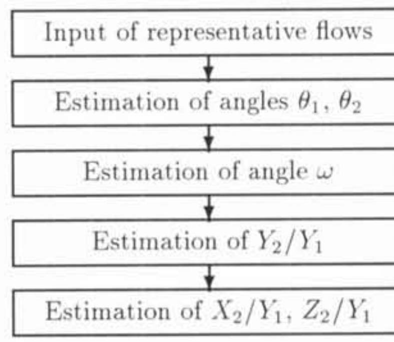


Figure 5: Outline of the estimation.

2.2.2 Estimation of angles θ_1 and θ_2 of depression

The angles θ_1 and θ_2 are estimated by using every representative flow obtained from every camera image respectively.

We explain how to estimate of angle θ_1. Since the movement direction of vehicle is unchangeable in very small time, we suppose that the present representative flow and the previous one are parallel vectors each other. On this assumption, the inclination of these flows are compared in the road coordinate system (W_X, W_Y, W_Z) for camera 1.

The relation between the road coordinate system (W_X, W_Y, W_Z) and the camera coordinate system (X, Y, Z) is transformed by:

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_1 & -\sin\theta_1 \\ 0 & \sin\theta_1 & \cos\theta_1 \end{pmatrix} \begin{pmatrix} W_X \\ W_Y - Y_1 \\ W_Z \end{pmatrix}. \quad (1)$$

The relation between the image coordinate system (x, y) and the camera coordinate system (X, Y, Z) is transformed by:

$$x = f_1 \frac{X}{Z} \quad and \quad y = f_1 \frac{Y}{Z}, \quad (2)$$

where focus length f_1 is the known parameter.

Moreover, the equation of the representative flow is set as:

$$y = a_1 x + b_1, \quad (3)$$

where a_1 and b_1 are computed from the representative flow vector. Thus, the representative flow is represented by the equation of plane as follows in the road coordinate systems.

$$f_1 a_1 W_X + (b_1 \sin\theta_1 - f_1 \cos\theta_1)(W_Y - Y_1) + (f_1 \sin\theta_1 + b_1 \sin\theta_1) W_Z = 0. \quad (4)$$

Since the vehicle run on the road, the flow is parallel with the plane $W_X W_Z$. Therefore, it is considered that W_Y is constant. Then, instead of the equation of plane, equation (4) becomes a straight line equation.

On the assumption that the representative flows are parallel in the sequential frames, which a_1, b_1 in the i th frame are denoted as $a_1^{(i)}$, $b_1^{(i)}$ respectively, the parallel condition is represented as follows:

$$\frac{f_1 a_1^{(i)}}{f_1 \sin\theta_1 + b_1^{(i)} \cos\theta_1} = \frac{f_1 a_1^{(i-1)}}{f_1 \sin\theta_1 + b_1^{(i-1)} \cos\theta_1}. \quad (5)$$

In the same way, the angle θ_2 of depression for camera 2 can be obtained.

2.2.3 Estimation of angle ω

In the same way as previous paragraph, the parallel condition of the representative flow is used to estimate the angle ω as shown in Figure 3.

When a, b in equation (3) for camera 1 are denoted as a_1, b_1 respectively and a, b for camera 2 are denoted as a_2, b_2 respectively, the parallel condition is represented as follows:

$$\frac{f_1 a_1}{f_1 \sin\theta_1 + b_1 \cos\theta_1} = \frac{f_2 \sin\theta_2 \sin\omega + f_2 a_2 \cos\omega + b_2 \cos\theta_2 \sin\omega}{f_2 \sin\theta_2 \cos\omega - f_2 a_2 \sin\omega + b_2 \cos\theta_2 \cos\omega}. \quad (6)$$

The angle ω is estimated by solving equation (6).

2.2.4 Estimation of the height Y1, Y2 of the camera

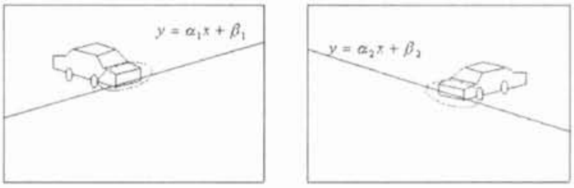
In the this paragraph, based on the assumption that the length of the flow transformed on the road plane XZ is equal, the height Y_1 and Y_2 of the camera are estimated. However, in spite of obtaining the values of Y_1 and Y_2, the value obtained in fact is the ratio between Y_1 and Y_2. We assume that it is enough to relate one camera image with another.

We define the start point $(W_{X_{1s}}, 0, W_{Z_{1s}})$ and $(W_{X_{2s}}, 0, W_{Z_{2s}})$, and the end point $(W_{X_{1e}}, 0, W_{Z_{1e}})$ and $(W_{X_{2e}}, 0, W_{Z_{2e}})$ of the representative flow for camera 1 and camera 2 respectively. The start point and the end point of the representative flow for camera 1 are defined.

This condition is described as:

$$\sqrt{(W_{X_{1e}} - W_{X_{1s}})^2} = \sqrt{(W_{X_{2e}} - W_{X_{2s}})^2}, \quad (7)$$

where these points are computed by equations (1) and (2).



(a) Camera 1 (b) Camera 2

Figure 6: Straight line in the front of vehicle.

2.2.5 Estimation of X_2 and Z_2

In the same as the previous paragraph, in spite of obtaining the values of X_2 and Z_2, the values obtained in fact are the ratios X_2/Y_1 and Z_2/Y_2.

In this session, we assume that the front line of vehicle in every image of camera 1 and camera 2, is equivalent in three dimensional space as shown in Figure 6. Then, we estimate the rations X_2/Y_1 and Z_2/Y_2.

In Figure 6, every equation of the straight line in the images is respectively described as:

$$y = \alpha_1 x + \beta_1 \quad and \quad y = \alpha_2 x + \beta_2. \quad (8)$$

The inclinations α_1 and α_2 are computed from the inclination of the representative flow in three dimension space. Because the front line becomes perpendicular to the representative flow in the road coordinate system $W_X W_Y W_Z$ where W_Y is fixed. In the road plane $W_X W_Z$, the inclination α' of the front line computed as perpendicular inclination to the representative flow.

The front region extracted from the image is the front part of the regions of which the optical flows exist.

3 Detection of the moving vehicle

The region of the moving vehicle is the sum of the regions where the optical flows exist in the image.

The regions between both images are matched by the camera parameter obtained in the previous section. Though it is not enough accurate to use these parameters for matching point to point, it is enough for matching region to region.

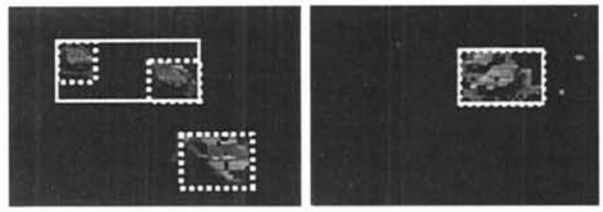
Thus, by comparing the number of the regions which is matched to each other, the state of the occlusion is recognized. For example, when two region in the image for camera 2 are matched to one region in the image for camera 1, the state for camera 1 becomes occlusion, and the number of vehicles becomes two.

4 Experiments

In the experiments, the images taken at two points on the walkway bridge were used as the in-

(a) Camera 1 (b) Camera 2

Figure 7: Example of the occlusion case.



(a) Camera 1 (b) Camera 2

Figure 8: Result of matching vehicle regions by using the camera parameters.

put, which were taken at video rate (1/30 *sec*) in 256 levels, monochrome and the size 720×486 pixels.

The result of the estimation of camera parameters from Figure 2 showed in Table 1. The result of our proposed method were almost similar to the measurement values.

In the case of occlusion in Figure 7, the viewpoints were both sides on the pedestrian bridge, and three vehicles existed in the image for camera 1, and two vehicles for camera 2. The first vehicle for camera 1 was not for camera 2. The second and third vehicles for camera 1 became the occlusion for camera 2.

The result of the extraction of vehicles was shown in Figure 8, where the region in the image for camera 2 was extracted as one region though two vehicles existed in the image for reasons of occlusion.

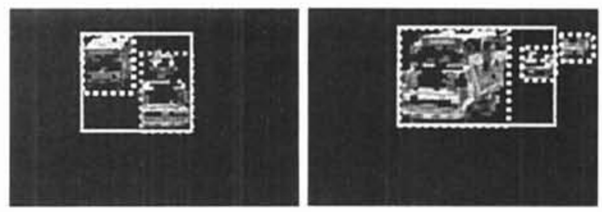
By using the camera parameters, the extracted regions were matched between cameras as shown in Figure 8, where the regions of vehicles were surrounded with break rectangle and the regions matched between both images were surrounded with bold rectangle.

In the result of the experiment, three vehicle were recognized in the image for camera 1, and the first vehicle was recognized which it run out of the image for camera 2. The second and the third vehicles were recognized that they were overlapped in the image for camera 2.

In the another case of worse occlusion in Figure 2, the viewpoints were a side and a middle place on the pedestrian bridge. The result was shown in Figure 9.

Table 1: RESULT OF ESTIMATION

	Measurement	Proposed method
θ_1	16°	16.7°
θ_2	14°	13.9°
ω	− 75°	− 77.4°
Y_2/Y_1	0.64 (4.5/7.0)	0.67
X_2/Y_1	0.85 (6.0/7.0)	0.86
Z_2/Y_1	2.57 (18/7.0)	2.52



(a) Camera 1 (b) Camera 2

Figure 9: Result of matching vehicle regions by using the camera parameters.

5 Conclusion

As one of ITS technique, we proposed a method to extract vehicles by image processing in order to observe the traffic flow. We proposed the method to extract vehicles for the road traffic surveillance of the traffic flow with the multiple views even if the occluded part of vehicle exists with one view. and cameras were placed as freely as possible by the road manager and so on.

Our method mainly used the movement vectors, that is the optical flows, by estimating the camera parameters, which are defined in the three dimension space. The time of obtaining them is when the number of moving vehicles is one.

Thus, the regions of vehicle were matched between images with multiple views by using the obtained parameters.

As the result of the region matching, the region of occlusion was recognized. Our implemented method was tried to real traffic scenes and successfully detected the individual vehicle.

References

[1] N. Ohta: "Image movement detection with reliability indices", *IEICE Trans.*, vol.E74, no.10, pp.3379-3388, Oct. 1991.

[2] *http://www.ail.cs.gunma-u.ac.jp/Labo/program.html.*