MVA2002 IAPR Workshop on Machine Vision Applications, Dec. 11-13, 2002, Nara-ken New Public Hall, Nara, Japan

8—4 Application to the Traffic Occlusion with Multiple Views

Tameharu Hasegawa[1]
Department of Computer Science
Chiba Institute of Technology

Heitoh Zen[2]
Institute of Media and Information Technology
Chiba University

Abstract

Recently, since the traffic is increasing every year, the problems of traffic jam and accidents, *and so on*, are increasing, it is required that the ITS (Intelligent Transport Systems) provides the drivers with the information for smooth and safety drive. As one of ITS (Intelligent Transport Systems) techniques, the studies on the traffic flow measurement by image processing have been proposed widely. In the conventional methods, since the traffic flow is almost observed from one viewpoint, thus one vehicle is occluded by another one. It is difficult to extract an occluded vehicle by the image processing. In our proposal method, multiple views are used in order to solve the occlusion in the image taken from one viewpoint. On the other hand, the locations of cameras are not always placed at the desirable position for the traffic flow observation system. We propose a method to extract the movement vehicles from the multiple viewpoints which are placed as freely as possible. Even if the occluded vehicle exists in the image taken from one viewpoint, our method solves the occlusion by using the image from another viewpoint. As the pre-processing of the solution of occlusion problems, the movement regions extracted from the same vehicle in each view image is matched by the geometric transformation considered the height of the vehicle based on the obtained camera parameters.

1 Instruction

Recently, since the vehicle is often used as the movement means, the traffic is increasing every year. For this reason, the problems of traffic jam and accidents, *and so on*, are increasing.

In order to solve these problems, it is required that the ITS (Intelligent Transport Systems) provides the drivers with the information for smooth and safety drive. As one of ITS techniques, the studies of the traffic flow measurement by image processing have been proposed widely.

In the conventional methods, since they are almost observed the traffic flows from one viewpoint with one camera or stereo-camera, one vehicle is occluded by another vehicle. Thus, it is difficult to detect the parts of one vehicle occluded by another vehicle with a single viewpoint.

The method [1][2] to solve the occlusion problems has been proposed with a single view by S. Kamijo, *et al*. This method, however, can not be adapted to the case that vehicles are always overlapped each other in the image sequences. It is, needless to say, impossible to detect the vehicle occluded completely by another one.

Our proposal method uses multiple views, because the method extracts the occluded part in the image from another viewpoint when the occlusion is captured in the image from one viewpoint.

On the other hand, the locations of cameras are not always placed at the desirable place to extract the vehicles. The reasons are the problems of the road construction and the landscape, for example, cameras should not be set on the wall where the drivers can look at them. Our method recognizes the relation of the multiple viewpoints without knowing the location of cameras before, like the conventional method.

Therefore, we have proposed a method to extract the moving vehicles by using multiple views even if the occluded part of the vehicle exists in the image taken from one viewpoint [3][4]. Our method solves the occlusion on condition that the multiple viewpoints may be placed as freely as by the road managers, *and so on*.

This method does not use the features which depend on the road such as the lane markers of white lines but two pieces of the common information detected by the image processing. The common information is the movement vector and the front line of vehicle. The brightness and the shape of vehicle in each image are very different.

We define the model of the movement of vehicles in the three-dimensional coordinates, whereas the movements are combined at the multiple viewpoints. By using the geometrical relation, the occluded part of a vehicle is detected in the image from each viewpoint.

This time, we develop our method, then these parameters of the model are obtained more stably, and the matching procedure considered the height of vehicle is proposed as the pre-processing of the solution of occlusion problems.

2 Estimation of Parameters

In this section, the coordinates of three-dimensions and the camera parameters are defined in order to recognize the correspondence between images.

Generally, if multiple cameras are used, then the positions are fixed and known, or the calibration of each camera is done by using the intensity of brightness and the shape of object, for example, lane markers, *and so on*.

In our study, cameras may be placed by road managers, *and so on*, as freely as possible. Since the intensity of brightness and the shape of the vehicle between viewpoints placed with long distance are very different, it is difficult to match between these images like stereo cameras.

We aim at the movement direction of vehicle observed at each viewpoint, which is the same direction in three-dimensional space. We consider the direction parallel to the road.

[1] Address: 2-17-1 Tsudanuma, Narashino-shi, Chiba 275-0016 Japan. E-mail: **hasegawa@cs.it-nara.ac.jp**

[2] Address: 1-33 Yaoi-cho, Inage-ku, Chiba-shi, Chiba, 263-8522 Japan. E-mail: **zen@imit.chiba.ac.jp**

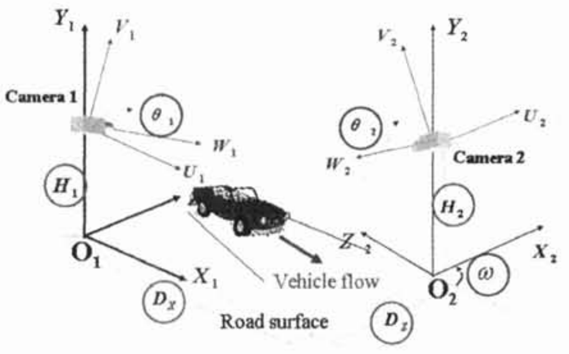


Figure 1: Coordinate systems and Camera parameters

At first, our proposed method tracks a vehicle to extract its flows from images if the condition for image processing is very good.

Secondarily, we obtain a rectangle around its region and five loci of a center of gravity and four center points of the rectangle. Every locus is approximated to a straight line. Then, the camera parameters are estimated by using the obtained straight lines.

Since the camera parameter are not estimated frequently because of using fixed cameras, we assume that the estimation is executed with good condition for the image processing, when the traffic quantity is little and the vehicle in the image is only one.

In Figure 1, the coordinate systems and the camera parameters to be obtained in our study are shown. These parameters are the external camera parameters, and the internal camera parameters have been already calibrated before our surveillance system has been set in the field.

We define the road coordinate system and the camera coordinate system in three-dimensional space as shown in Figure 1, where the marks enclosed by circle are the parameters to be obtained.

The road coordinate system for camera n is defined as (X_n, Y_n, Z_n), where the subscript n is the camera number, 1 or 2 in this paper. The plane X_nZ_n, which corresponds to the road plane, is assumed horizontal.

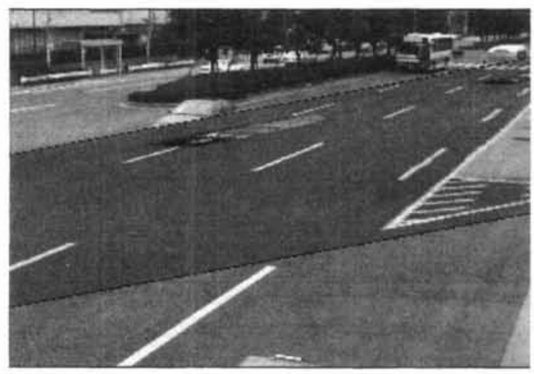
Camera 1 overlooks the road with the angle θ of depression and the height H_n. Meanwhile, the camera coordinate system is defined as (U_n, V_n, W_n), where the plane U_nV_n corresponds to the plane Y_nZ_n in the road coordinate system. Then, the plane X_1Z_1 and X_2Z_2 are set on the same road plane. The angle between planes Y_1Z_1 and Y_2Z_2 is defined as ω and the origin of the coordinate system (X_2, Z_2, Z_2) is $(D_X, 0, D_Z)$ in the coordinate system (X_1, Y_1, Z_1).

The image coordinate systems (x_n, y_n) for camera n are defined as follows: the axis x_n is parallel and the same direction to the axis U_n, and the axis y_n is parallel and the same direction to the axis V_n, and the origin of (x_n, y_n) is the center of image.

Then, it supposes that only the focus length f_n of the Camera n has been calibrated and obtained before, then the roll angle of every camera is ignored. Especially, the coordinate system (X_1, Y_1, Z_1) is defined as the basic coordinate system, that is, the world coordinate system. We call the image for camera n image n in the following explanation.

If the genetic transformation from image n onto image n' (n≠n') is considered, it is well that the parameters for the length is not the absolute length but the relative length, and so, we obtain the length as the ratio to H_1. Therefore, the parameters to be obtained in our study are θ_n, ω, H_n/H_1, D_X/H_1, and D_Z/H_1.

(a) Image 1



(b) Image 2 overlapped the transformed image 1

Figure 2: Conventional projective transformation

These parameters are obtained in the above order by using the information of the moving vehicle. How to obtain them is referred to [4].

3 Matching between views

The matching process, which is often used in the case of using multiple cameras, between each image taken from camera has determined feature points on a flat plane and obtained the projective transformation matrix by these correspondences. Then, the process has been done by the transformation. In this case, if the texture on the flat plane, it is coincided between images with views. Otherwise, it is not, for example, a texture is one on a solid body.

Figure 2 shows an example of a vehicle's texture. Then, feature points used to obtain the projective transformation matrix are the ends of lane markers, and are extracted by person in order to obtain the matrix as precisely as possible. Figure 2(a) is image 1, taken from viewpoint 1, and Figure 2(b) is the image transformed from that by the matrix and overlapped onto image 2 at the same time as the moment taken image 1.

The moving vehicle in Image 1 is pasted on the road in image 2, and then, its region is not almost overlapped to the region of the same vehicle in image 2, and its shape is much distorted. Considering that some vehicles approach each other at the congested time, it is very difficult to match the regions of the same vehicle between images. This difficulty is caused by the projective transformation which is not considered that a vehicle has a solid body with height.

Then, a geometric transformation considered that vehicle has a solid body with height is proposed, which is explained in the next paragraph.

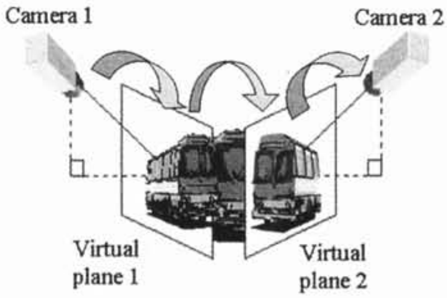
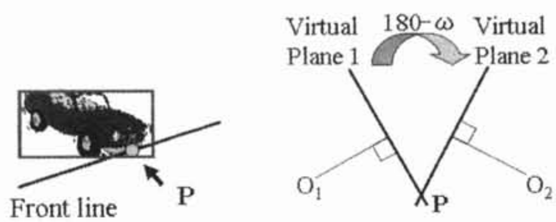


Figure 3: Virtual planes, and geometric transformation



(a) Front line and point p (b) Rotation of virtual plane

Figure 4: Transformation between virtual planes

3.1 Introduction of a virtual plane

In Figure 3, a perpendicular plane to the plane X_nZ_n and the straight line projected the axis of every camera's light onto it is created virtually. We call this plane a virtual plane. Then, since there is a virtual plane for a camera, the number of planes is equal to the number of cameras, that is, viewpoints.

Our proposal geometric transformation is done from the image 1 to image 2 through these virtual planes as shown in Figure 3. Then, it is necessary to obtain the location of the virtual plane, that is, the depth from the camera. In next paragraph, how to obtain the depth is explained in detail.

Similarly, the geometric transformation from image 2 to image 1 through these virtual planes can be done.

3.2 Proposal geometric transformation

In this paragraph, the process of our proposal geometric transformation is explained concretely.

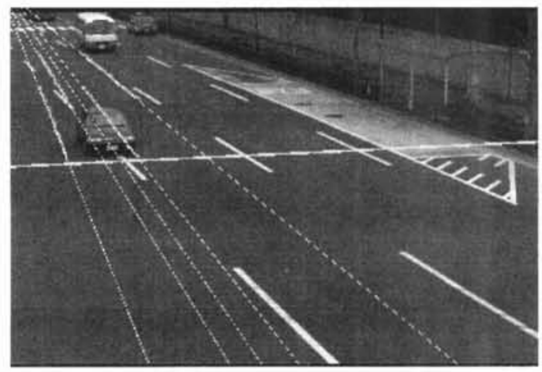
The geometric transformation from the image for camera 1 to the image for camera 2 through these virtual planes is done as follows:

1) Detection of a moving region in image 1,
2) Projective transformation of it from image 1 onto virtual plane 1,
3) Geometric transformation of the rotation for the above output from virtual plane 1 onto virtual plane 2, and
4) Projective transformation for the above output from virtual plane 2 onto image 2.

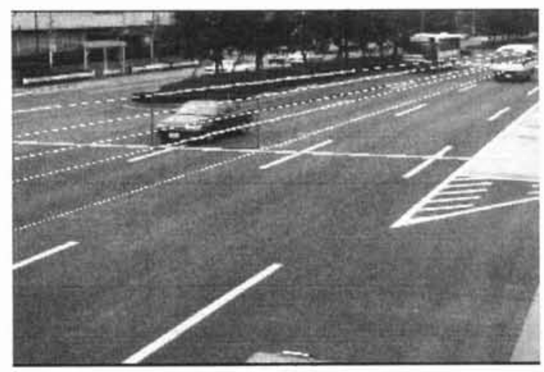
In fact, the above three transformations can be calculated once by the arrangement of them.

Next, how to determine the depth from the camera to the virtual plane is explained.

In this work, the circumscribed rectangle around the moving region for the vehicle and the straight line obtained under the front of it, we call a front line, are obtained through the process that estimates the camera parameters as shown in Figure 4(a).



(a) Middle results by image processing in image 1



(b) Middle results by image processing in image 2

Figure 5: Middle result by image processing

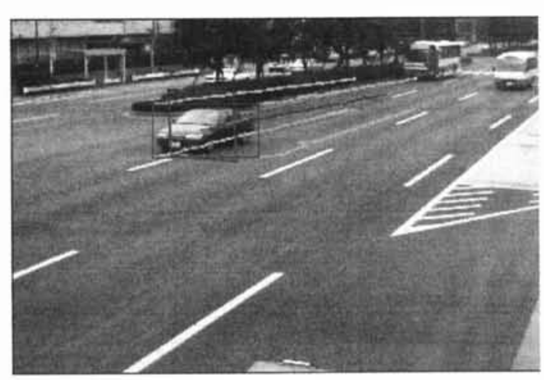


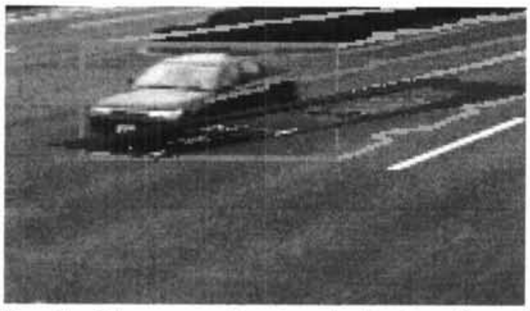
Figure6:

The point p in Figure 4(a) is defined as the middle point between two intersection points of the front line and the circumscribed rectangle.

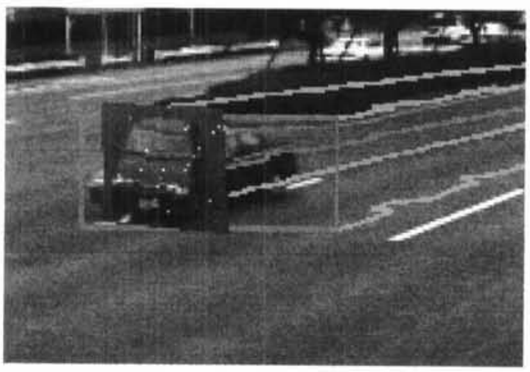
Then, since this can be obtained on image 1, point P in Figure 4(b) can be calculated by the projective transformation onto the plane X_nZ_n, of which axes are described as X_n/H_l and Z_n/H_l respectively, because camera parameters for length are obtained as the ratio to H_l.

Therefore, the coordinate of point P are obtained as the ratio to H_l. Then, the depth of virtual plane 1 and virtual plane 2 are created, and so, the transformation the above process 2, 3, and 4 is made possible.

The axis of the rotation in the above process 3 is the intersection of virtual plane 1 and virtual plane 2, where the intersection is determined through point P as shown in Figure 4(b). The angle becomes (180-ω) degree.



(a) Result of the conventional projective transformation



(b) Result of our proposed geometric transformation

Figure 7: Comparison of the conventional projective transformation and our proposal geometric transformation

Note that point P and every virtual plane for the geometric transformation from image 1 onto image 2 do not always agree with that point P and every virtual plane for the geometric transformation from image 2 onto image 1, because the point p obtained on one image do not always agree with the point p obtained on another image though we assume that they are correspondence in the process of the estimation of camera parameters.

Furthermore, the number of the point p becomes the number of the moving vehicles, though the number of the object, that is the moving vehicle, is one in the process of the estimation of camera parameters.

4 Experiments

In the experiments, the images taken from two viewpoints on the walkway bridge are used as the input, which are taken at video rate (1/30 *sec*) in 256 levels, monochrome and 720x486 pixels.

By processing these images sequences under good condition, our defined parameters were estimated.

Figure 5 shows the middle results obtained by image processing through the process of the requisition of the camera parameters. A black rectangle around a vehicle was described as the circumscribed rectangle of the vehicle region detected by image processing. The broken straight lines were used by obtaining the parameters θ_m, ω, and H_2/H_1. A white straight line in front of the vehicle is described as the front line, which is used by obtaining the parameters D_X/H_1 and D_Z/H_1.

Figure 6 shows the result of our proposal geometric transformation from image 1 to image2, where this transformation is done for the region inside the circumscribed rectangle of the moving vehicle, and the transformed region is overlapped onto image 2.

We compare our proposal geometric transformation with the conventional projective transformation as shown in Figure 7. The overlapped area of the region for the vehicle in image 1 and the transformed region for the vehicle in image 2 is large and the overlapped area, for example the front part of vehicle, is matched well, though the area is very small and the transformed region for the vehicle in image 2 is much distorted and stretched.

Therefore, our proposal geometric transformation is proved effectively. Then, we are developing the traffic surveillance system by solving the occlusion problems based on the information, for example, texture, the size and location, *and so on*, of the overlapped area by using our geometric transformation and the obtained camera parameters in the states of the hard congestion.

5 Conclusion

We proposed the method to extract vehicles by the observation of the traffic flow from the multiple viewpoints to solve the occlusion problems under the condition that cameras might be placed as freely as possible.

In our future works, we are developing the traffic surveillance system by solving the occlusion problems based on the information of the overlapped area by using our geometric transformation and the obtained camera parameters in the states of the hard congestion.

References

[1] S. Kamijo, and M. Sakauchi: "Vehicle tracking in low-angle and front view images based on spatio-temporal markov random field model," 8th World Congress on ITS, ITS00381(CD-ROM), 2001.

[2] T. Nishida, S. Kamijo, and M. Sakauchi: "Automated statistics acquisition from traffic images based on occlusion robust vehicle tracking method," 8th World Congress on ITS, ITS00642(CD-ROM), 2001

[3] T. Hasegawa, and H. Zen: "Road traffic surveillance with multiple views," Proc. of MVA2000, IAPR Workshop on Machine Vision Applications, pp.15-18, Nov. 2000.

[4] T. Hasegawa, T. Mahara, and H. Zen: "Road traffic surveillance with multiple views," 8th World Congress on ITS, ITS00397(CD-ROM), 2001

[5] T. Hasegawa, K. Nohsoh, and S. Ozawa: "Counting Cars by Tracking Moving Objects in the Outdoor Parking Lot," The Second World Congress on Intelligent Transport Systems, Vol.I, pp.230-235, Nov. 1995.