Vehicle Lateral Position Estimation Method Based on Matching of Top-View Images

Tomoaki Teshima, Hideo Saito, Shinji Ozawa
Keio University
Department fo Information and Computer Science
3-14-1 Hiyoshi, Kohoku, Yokohama
Kanagawa, 223-8522, Japan
{tomoaki,hideo,ozawa}@ozawa.ics.keio.a.jp

Keiichi Yamamoto, Toru Ihara
Mitsubishi Fuso Truck & Bus Corporation
10 Okura, Nakahara, Kawasaki
Kanagawa, 211-0023, Japan
{keiichi.yamamoto, tohru.ihara}
@mitsubishi-fuso.com

Abstract

In this paper, a method to estimate the lateral position of the vehicle from a sequence of moving camera images is proposed. Proposed method relies on the plane projective transform (Homography) between the ground and the image plane of the present and the next frame. Homographies are obtained according to the information of the calibrated camera and the vehicles' speed. The movement of the camera is obtained by the registration of the 2 consecutive input images. The proposed method does not rely on extraction of features such as lines, flow vectors or lane markers, but based on matching of warped top-view images between two consecutive frames. Therefore, vehicle movement can be estimated even from images without such explicit image features or without special sensors or GPS. In this paper, experiments are done with both synthesized and real images and enough accuracy is shown even compared to conventional method.

1. Introduction

Understanding the motion or a movement of the camera is very important in many applications, such as autonomous driving [1, 2], vehicle driver's support[3, 4], vehicle navigation as NAVLAB [5] and CMU-VAMORS [6], etc. Many researchers have developed many methods to estimate the motion of the camera. Each method can categorize to FOE based method[4], optical flow based method and marker based method[3] which includes the use of GPS[7]. We propose another method which does not rely on any marker or GPS, in this paper. The accuracy of our method is higher than FOE based and optical flow based method.

The position of the camera is obtained by accumulating the movement between the each 2 consecutive frames. The movement is obtained by using the plane projective transform matrix(Homography), the vehicles' speed information and the calibrated information of the camera in our method. A single camera which is fixed on the vehicle is focused and such a condition that lane markers are not always seen is considered in our method. Since the camera is fixed to the vehicle, the movement of the camera is equal to the movement of the vehicle.

Our method is articulated in 6 steps including 1 preprocess. 1. Calibrate the camera and according to the calibrated information, homography between the ground and the image plane $\mathbf{H}_P$ is obtained. 2. compute the homography $\mathbf{H}_N$ between the ground and the image plane at the next frame by using the hypothesis of the angle of the vehicles' direction θ and speed of the vehicle l obtained from the sensor. 3. Convert the present input image to a top-view image using $\mathbf{H}_P$. 4. Convert the next input image to a top-view image using $\mathbf{H}_N$. 5. compute the SAD (Sum of Absolute Difference) value between the top-view image obtained from 3 and 4. 6. Repeat 2–5 for the entire hypothesis of the angle of the direction.

The movement which is obtained from the each frame is accumulated to obtain the lateral position of the camera. Note that the movement of the camera, which can be expressed by l and θ, is not the output of our method, but the input information. The repeating process of deciding the hypothesis and the evaluating is the particular feature of our method.

2. Proposed Method

A homography can be obtained from either the correspondence points and the geometrical relation The information of the geometrical relation is used in our method to obtain a homography.

Figure 1 shows the frame work of our method. $\mathbf{H}_P$ is a homography which converts the present frame to a top-view

The 18th International Conference on Pattern Recognition (ICPR'06)
0-7695-2521-0/06 $20.00 © 2006 IEEE

image according to the calibrated information of the camera. $\mathbf{H}_N$ is a homography which converts the next frame to a top-view image according to the calibrated information and the speed and the angle of the direction of the vehicle. The speed information can be obtained from the vehicles' sensor. The hypothesis of the angle of the direction is chosen to compute the $\mathbf{H}_N$.

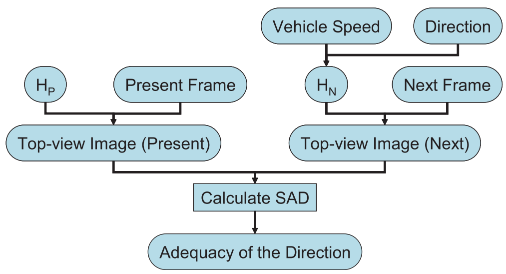
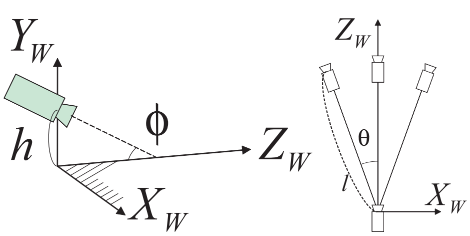


Figure 1. Frame work of the method

2.1. Camera model

We assume that the intrinsic parameters **A** and the height from the ground h and the angle ϕ of the camera against the horizontal level are already obtained in the pre-process.



(a) Geometrical position of the camera

(b) Relation of the camera

Figure 2. Camera model used in our method

Figure 2 (a) shows the relation between the world coordinate and the cameras' position. The plane which is expressed by $Y_W = 0$ is the ground. The origin of the world coordinate is fixed at the point on the ground directly under the camera. The axis of the camera is coplanar with the Y_W–Z_W plane.

Figure 2 (b) shows the relation of the camera positions in the two consecutive frames. The distance between the two cameras is expressed by the speed of the vehicle l. The angle of the direction θ becomes zero when the camera moves straight forward.

An assumption is made in our method that the angle of the camera equals the angle of the direction as shown in Fig.2 (b), since the camera is fixed to the vehicle and the vehicle heads to the direction where it moves.

2.2. Homography for the present frame

In the situation shown in section 2.1, the rotation matrix **R** and the translation vector $\boldsymbol{t}$ of the camera can be decided from 2 parameters which are already obtained, the height h and the angle ϕ of the camera against the horizontal level.

$$(\mathbf{R}|\boldsymbol{t}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \phi_c & \phi_s & h\phi_c \\ 0 & -\phi_s & \phi_c & h\phi_s \end{pmatrix} \quad (1)$$

where s_θ, c_θ, s_ϕ, c_ϕ denotes $\sin\theta$, $\cos\theta$, $\sin\phi$, $\cos\phi$. The planar projective transform matrix $\mathbf{H}_P$ is obtained by multiplying the intrinsic parameter matrix **A** and the matrix in Eq.(1). $\mathbf{H}_P$ is a matrix which converts the input image of the present frame to the top-view image. Since h and ϕ do not change, matrix $\mathbf{H}_P$ is computed only once.

2.3. Homography for the next frame

Another homography $\mathbf{H}_N$ should be computed in order to register the image plane at the next frame to the current frame. To obtain $\mathbf{H}_N$, computation of the rotation matrix **R** and the translation vector $\boldsymbol{t}$ is required again, according to the vehicles' speed l and the hypothesis of the angle of the vehicles' direction θ.

According to l and θ, the rotation matrix **R** and the translation vector $\boldsymbol{t}$ can be rewritten as following.

$$(\mathbf{R}|\boldsymbol{t}) = \begin{pmatrix} c_\theta & 0 & s_\theta & -ls_\theta \\ s_\phi s_\theta & s_\phi & s_\phi c_\theta & hc_\phi + lc_\theta s_\phi \\ c_\phi s_\theta & -s_\phi & c_\phi c_\theta & hs_\phi + lc_\theta c_\phi \end{pmatrix} \quad (2)$$

According to l, h, ϕ and θ, **R** and $\boldsymbol{t}$ can be obtained as shown in Eq.(2). Since h and ϕ is fixed, $\mathbf{H}_N$ can be obtained from l and θ.

2.4. The position and the movement of the vehicle

The 2 consecutive input frames are converted to the top-view images according to the homographies $\mathbf{H}_N$ and $\mathbf{H}_P$ as shown in Fig.3. The 2 top-view images must become completely identical if the hypothesis of the angle of the vehicle is correct. Thus, the adequacy of the hypothesis can be evaluated by computing the similarity between the 2 top-view images. Sum of Absolute Difference (SAD) value is used in our method to measure the similarity. SAD value

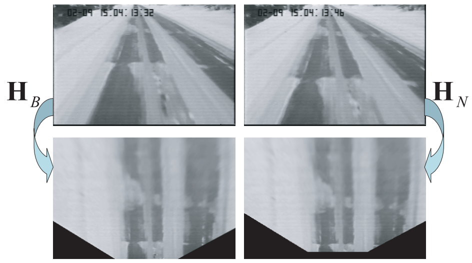


Figure 3. Trasforming the input images to the top-view images

is obtained by calculating the absolute difference of each intensity of the pixel which corresponds in the image at the current frame and the next frame.

The vehicles' direction at each frame is obtained by choosing the hypothesis of the direction which gives the minimum SAD value. Thus, the movement $\boldsymbol{v}_i$ and the position $\boldsymbol{p}_i$ of the vehicle at frame i can be expressed as

$$\boldsymbol{v}_i = l_i \begin{pmatrix} \sin\theta_i & \cos\theta_i \end{pmatrix}^\top \tag{3}$$

$$\boldsymbol{p}_i = \sum_{j=0}^{i} \boldsymbol{v}_i = \sum_{j=0}^{i} l_j \begin{pmatrix} \sin\theta_i & \cos\theta_i \end{pmatrix}^\top \tag{4}$$

where l_i expresses the speed and the θi expresses the angle of the direction of vehicle at frame i.

3. Experiments

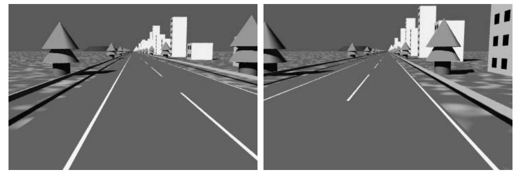
Two types of experiments have been done. The first experiment has been done for the evaluation of the correctness and the accuracy of our method by using the synthesized images. Second experiment has been done by using some real image sequences.

3.1. Experiment with synthesized image

The images which are used as the input images are synthesized by POV-Ray for Windows ver 3.6. Three types of images are synthesized. Each type of images is shown in Fig.4. The first type includes the lane marker and the guard rail. The second type includes only the guard rail. The third type has only the texture drawn on the road. The vertical level of the camera is fixed at 2 meters, and the lateral position changes as shown in Fig.5. The camera axis goes through the point 14 meter ahead on the ground, and the camera moves straightforward parallel to the ground.

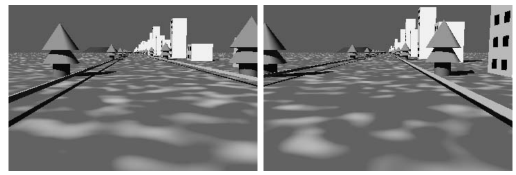
The estimated results are shown in Fig.5 compared to the correct position. The result obtained from each input image sequence has very small difference compared to the amplitude of the position. This shows that our method do not rely on any explicit marker.

We consider that the error compared to the correct position is small enough in our method, since our methods' target is not to obtain an accurate position, but to be able to apply in various condition.



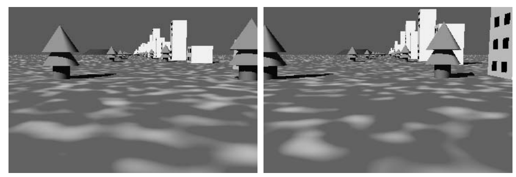
(a) With lane marker and rail (frame 1)

(b) With lane marker and rail (frame 300)



(c) With lane marker only (frame 1)

(d) With lane marker only (frame 300)



(e) Nothing on ground (frame 1)

(f) Nothing on ground (frame 300)

Figure 4. Synthesized input images

3.2. Experiment with real image

Two types of the real image sequences are used in this experiment, the vehicle on the normal road and on the

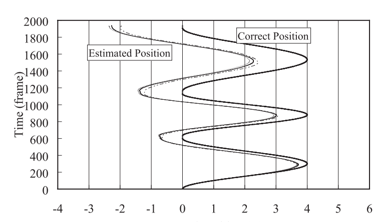


Figure 5. Estimated result of synthesized image

The 18th International Conference on Pattern Recognition (ICPR'06)
0-7695-2521-0/06 $20.00 © 2006 IEEE

snow road. The image sequence which is on the normal road has the lane marker drawn on the road. During this experiment, comparison between the results from our method, lane marker based method and the optical flow based method has been done. Lane marker based method estimates the vehicles' position according to the lateral position of the lane marker in the image. In the optical flow based method, vehicles' position is obtained from the movement of the camera. The movement of the camera is obtained from the optical flow by pattern matching. Figure 6 shows the results. In this section, the result obtained from the lane marker based method is considered as the correct value since the lane marker is available. In this experiment, the result from our method showed better result than the conventional method and enough accuracy compared to the correct movement as shown in Fig.6.

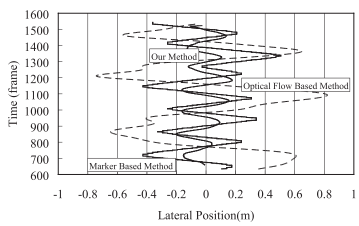


Figure 6. Estimated result on a normal road

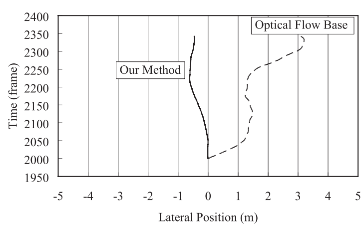


Figure 7. Estimated result on a snow road

Another experiment has been done to show the performance of our method on a road with snow. The estimated result of the experiment is shown in Fig.7. Because the lane marker is not availble, the results are compared between our method and the optical flow based method. The vehicle is running straightforward along the road in this experiment so the estimated position must become straight. In this experiment, correct position can not be obtained, but the result of our method shows more natural and reasonable movement along the road.

These experiments show that our method has enough accuracy to estimate the lateral position of the vehicle and can be applied in the situation which was impossible to estimate.

Computation time per frame in each situations are shown in table 1. Our methods requires improvement to use in real time.

Table 1. Computation time per frame in each situations

scene	frame	time	fps
road2	200	8:41	0.384
road3	443	19:57	0.370
snow	400	16:55	0.394

4. Conclusion

In this paper, a method which estimates the lateral position of the vehicle is proposed. Proposed method is based on the registration of the top-view images which are obtained from the homography. Therefore, our method requires none of image features such as line, lane marker, rail, or sensorsbut a calibrated camera. Experiments have been done and enough accuracy to estimate the vehicles' movement has been shown.

References

[1] P.Martinet and C.Thibaud: "Automatic Guided Vehicles: Robust Controller Design in Image Space", *Autonomous Robots*, Vol. 8, No. 1, pp. 25–42, January 2000.

[2] M. A. Sotelo, F. J. Rodriguez, L. Magdalena, L. M. Bergasa and L. Boquete: "A Color Vision-Based Lane Tracking System for Autonomous Driving on Unmarked Roads", *Autonomous Robots*, Vol. 16, No. 1, pp. 95–116, 2004.

[3] J. Manigel and W. Leonhard: "Vehicle Control by Computer Vision", *IEEE Transactions on Industrial Electronics*, Vol. 39, No. 3, pp. 181–188, June 1992.

[4] T. Teshima, H. Saito, S. Ozawa, K. Yamamoto, T. Ihara: "Estimation of FOE Without Optical Flow for Vehicle Lateral Position Detection", *Proceedings of IAPR Conference on Machine Vision Applications*, pp. 406–409, May, 2005.

[5] E. D. Dickmanns and B. D. Mysliewitz: "Recursive 3-d road and relative ego-state recognition", *IEEE Transactions on Pattern Analysis and Machine Intellgence*, Vol. 14, No. 2, pp. 199–213, 1992.

[6] T. Kanade, C. Thorpe, M. Herbert, and S. Shafer: "Vision and navigation for the carnegie-mellon navlab", *IEEE Transactions on Pattern Analysis and Machine Intellgence*, Vol. 10, No. 3, pp. 361–372, 1988.

[7] M. El B. El Najjar, P. Bonnifait: "Road-Matching Method for Precise Vehicle Localization", *Autonomous Robots*, Vol. 19, pp. 173–191, 2005.