

ROAD PLANE DETECTION USING DIFFERENTIAL HOMOGRAPHY ESTIMATED BY PAIR FEATURE MATCHING OF LOCAL REGIONS

Kenji Nishida

National Institute of Advanced Industrial
Science and Technology (AIST)
Central 2, 1-1-1 Umezono,
Tsukuba, Ibaraki 305-8568, JAPAN
kenji.nishida@aist.go.jp

Jun Fujiki

National Institute of Advanced Industrial
Science and Technology (AIST)
Central 2, 1-1-1 Umezono,
Tsukuba, Ibaraki 305-8568, JAPAN
jun-fujiki@aist.go.jp

Chikao Tsushiya

Nissan Motor Co., Ltd.
1-1, Morinosatoaoyama, Atsugi-shi,
Kanagawa 243-0123, JAPAN
ctsuchiya@mail.nissan.co.jp

Shinya Tanaka

Nissan Motor Co., Ltd.
1-1, Morinosatoaoyama, Atsugi-shi,
Kanagawa 243-0123, JAPAN
shinya_tanaka@mail.nissan.co.jp

Takio Kurita

Hiroshima University
1-7-1 Kagamiyama, Higashi-Hiroshima, Hiroshima 739-8521, JAPAN
tkurita@hiroshima-u.ac.jp

ABSTRACT

This paper presents a novel algorithm for road plane detection from an on-board camera. The algorithm employs the temporal difference of homography matrix, which is termed differential homography, caused by camera motion. Differential homography is estimated from optical flows of road plane regions, while using RANSAC algorithm to extract the majority optical flows. Since differential homography is estimated using the relationship between the image coordinate (location in an image) and the flows at the locations in an image. The proposed algorithm does not require the estimation of the homography matrix itself. Therefore, the proposed algorithm can be applied without calibration. The proposed algorithm effectively detect the optical flows from road region with using the pixel-pair feature matching. The algorithm is applied to the city traffic images distributed by UCL, and its average road detection ratio is found to be 75.6%. It is also applied to the previously collected suburban traffic images. A suitable detection result is obtained.

KEY WORDS

Road Plane Estimation, Optical Flow, Homography Matrix, ITS

1 Introduction

Road plane (drivable region) detection from visual information is one of the most important processes in autonomous vehicle navigation and driver assistance systems. The approaches to road plane detection can be roughly classified into two types – appearance-based detection and the

geometry-based detection (or segmentation).

Road plane detection on the basis of appearance is usually difficult, since the road plane does not have features significant for classification. Therefore, road markings are widely used to detect components of the road plane, such as driving lanes and cross walks [1, 2, 3].

The geometry-based approach involves the separation of the road plane and other objects through reconstruction of the 3D environment. Structure from motion (SfM)[4] is a typical geometry based method that uses motion cues of detected feature points. Although the road plane can easily be determined if the 3D environment is fully reconstructed, the actual road/obstacle segmentation is difficult; this is because feature points are rarely detected sufficiently for determining a complete 3D environment. Therefore, SfM is often employed in combination with other method. For example, Leibe et al. combined temporal integration and appearance-based detector with SfM in [5, 6], and Brostow et al. combined supervised learning with SfM [7].

A complete 3D reconstruction is not necessary for road plane detection; only 3D information of the road plane (homography) is required for road/obstacle segmentation. In this approach, a homography matrix is estimated from some visual cues, and the feature points that satisfy the homography are classified as road plane while the feature points that do not satisfy the homography are classified as obstacles. Nakai et al. employed stereo matching based on the homography of the road plane and the segmented road region using the two corresponding stereo images [8]. Simond employed super feature points mainly extracted from road markings for the stereo matching of the road plane [9]. Okada et al. extracted the projective invariant for horizon-

tal lines to estimate the road plane [10].

We propose a novel algorithm for road plane detection using the homography-based approach. In our algorithm, pixel-pair feature matching of local regions is used to estimate the temporal difference of homography matrix (differential homography) between two consecutive video frames; the regions that satisfy the estimated differential homography are classified as road plane.

The proposed algorithm is described in the following section, and the experimental results are shown in Section 3.

2 Road Plane Detection using Differential Homography

In this section, we propose a road plane detection method using differential homography. Differential homography refers to the temporal difference of homography matrices between the on-board camera and the road plane. We first describe the procedure of proposed method briefly and then discuss each step in detail.

2.1 Differential Homography

Let \mathbf{X} represent the coordinate on the road plane π ; and \mathbf{u} , the coordinate on the image I . The relation between \mathbf{X} and \mathbf{u} is expressed as follows:

$$\mathbf{u} = \mathbf{H}\mathbf{X}, \quad (1)$$

where \mathbf{H} denotes the homography matrix.

Figure 1 shows the relationship between \mathbf{X} and \mathbf{u} for moving camera. The homography matrix on time t is represented as \mathbf{H}_t ; therefore, the relationship between \mathbf{X}_t and \mathbf{u}_t is represented as follows

$$\mathbf{u}_t = \mathbf{H}_t\mathbf{X}. \quad (2)$$

Similarly, for time $t+\Delta t$, the relation of $\mathbf{X}_{t+\Delta t}$ and $\mathbf{u}_{t+\Delta t}$ is represented as follows:

$$\mathbf{u}_{t+\Delta t} = \mathbf{H}_{t+\Delta t}\mathbf{X}. \quad (3)$$

The temporal difference is represented as

$$\Delta\mathbf{u}_t = \mathbf{u}_{t+\Delta t} - \mathbf{u}_t \quad (4)$$

$$= (\mathbf{H}_{t+\Delta t} - \mathbf{H}_t)\mathbf{X} \quad (5)$$

$$= \Delta\mathbf{H}_t\mathbf{X}. \quad (6)$$

By substituting \mathbf{X} with $\mathbf{X} = \mathbf{H}^{-1}\mathbf{u}$, equation (6) is reformed as follows

$$\Delta\mathbf{u}_t = \Delta\mathbf{H}_t\mathbf{H}_t^{-1}\mathbf{u}. \quad (7)$$

When $\Delta\mathbf{u}_t$ (usually called optical flow) is caused only by the camera motion, the pixels projected from road plane π satisfy the equation (7). Therefore, we can segment the image I_t into road plane and non-road plane regions by estimating the differential homography $\Delta\mathbf{H}_t\mathbf{H}_t^{-1}$.

The road plane regions are segmented by the following procedure:

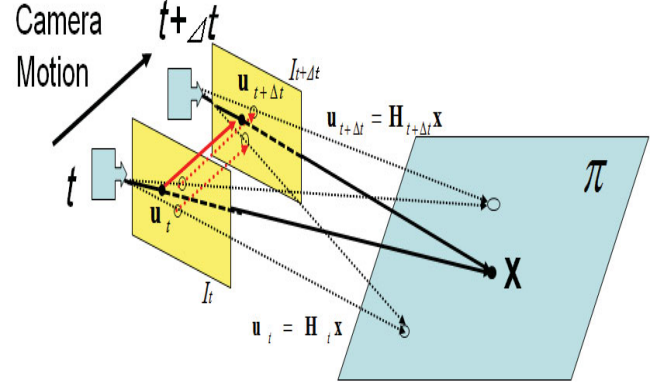


Figure 1. Homography Matrix for Moving Camera

1. Detection of optical flow from consecutive video frames.
2. Collection of the optical flows from the road plane region.
3. Estimation of differential homography $\Delta\mathbf{H}_t\mathbf{H}_t^{-1}$.
4. Estimation of flows for road plane region using $\Delta\mathbf{H}_t\mathbf{H}_t^{-1}$.
5. Comparison between detected flow and estimated flow.

The detailed procedure is described in the remainders of this section.

2.2 Flow Detection by Pair Feature Matching

Lucas-Kanade algorithm [11] for optical flow detection is not suitable for detecting the optical flow on the road plane, since the displacement of feature points are considerably too large for the texture of the road plane in many cases. Although Brox et al. proposed a large displacement optical flow detection algorithm in [12] for segmented sub-regions, it is also difficult to apply this algorithm to the road plane, which might contain various optical flows in the segmented region.

We employed feature matching of the image patch for detecting optical flow. Figure 2 shows the detection procedure as follows:

1. An image of video frame t is obtained.
2. Feature points, which have edge power and temporal difference, are extracted.
3. A patch P_t , which contains sufficient feature points, is selected in image I_t .
4. The most similar patch $P_{t+\Delta t}$ is found in image $I_{t+\Delta t}$.

5. Displacement between P_t and $P_{t+\Delta t}$ is considered as the optical flow.

We employed Pixel-Pair feature [13] for estimating the similarity of patches, since it is a powerful technique for estimating the similarity between one reference image (template) and many evidential images (this property is common in similar features described in [14, 15, 16]). The pixel-pair feature is an extension of the statistical reach feature (SRF) [14], in which the restriction on the distance between pixel pairs is removed. The definition of the pixel-pair feature and the similarity index $c(I, J)$ of a given pair of reference image I and evidential image J of the same size are described as follows (figure 3). Suppose the size of the input images is $W \times H$. Let grid Γ represent a set of pixel coordinates in the images I and J , for instance, $\Gamma := \{(i, j) | i = 1, \dots, W, j = 1, \dots, H\}$. We regard the image of size $W \times H$ as an intensity function defined on Γ . For an arbitrary pair (p, q) of grid points in Γ , we define the value $ppf(p \succ q; T_p)$ as follows:

$$ppf(p \succ q; T_p) := \begin{cases} 1 & \mathbf{I}(p) - \mathbf{I}(q) \geq T_p \\ -1 & \mathbf{I}(p) - \mathbf{I}(q) \leq -T_p \\ \phi(\text{empty}) & \text{otherwise} \end{cases} \quad (8)$$

Here, $T_p (> 0)$ is the threshold of the intensity difference. We adopt the grid-point pair (p, q) as a feature when $ppf(p \succ q; T_p) \neq \phi$ (ϕ stands for an empty set). Hereafter, we use the term $ppf(p \succ q)$ instead of $ppf(p \succ q; T_p)$, unless in the case of ambiguity.

We limit the number of pixel-pair features to N by selecting a set of pairs (p, q) with a selection policy s . We express the pixel-pair feature set RP_s as follows:

$$RP_s(\mathbf{p}, \mathbf{q}, \mathbf{I}, T_p, N) := \{ppf(\mathbf{p} \succ \mathbf{q}) \neq \phi\}, \quad (9)$$

where $\{\mathbf{p}, \mathbf{q} \in \Gamma \times \Gamma\}$, $\mathbf{p} = \{p_1, \dots, p_N\}$, and $\mathbf{q} = \{q_1, \dots, q_N\}$.

We define the incremental sign $b(\mathbf{p} \succ \mathbf{q})$ for the evidential image \mathbf{J} for computing the similarity between images \mathbf{I} and \mathbf{J} as follows:

$$b(\mathbf{p} \succ \mathbf{q}) := \begin{cases} 1 & \mathbf{J}(p) \geq \mathbf{J}(q) \\ -1 & \text{otherwise} \end{cases} \quad (10)$$

For a pixel pair $(p, q) \in RP_s$, a single-pair similarity $r(p, q, \mathbf{J})$ is defined as follows:

$$r(p, q, \mathbf{J}) = \{ppf(p \succ q) = b(p \succ q)\}. \quad (11)$$

$c_s(\mathbf{I}, \mathbf{J}, RP_s)$, the similarity index between image \mathbf{I} and \mathbf{J} , measured by using a pixel-pair feature set RP_s is defined as follows:

$$c_s(\mathbf{I}, \mathbf{J}, RP_s) = \frac{\sum_{(p,q) \in RP_s} r(p, q, \mathbf{J})}{|RP_s|} \quad (12)$$

The displacement between patch \mathbf{I} and the most similar patch in \mathbf{J} is considered as the optical flow for patch \mathbf{I} .

This procedure is repeated over the input image I_t ; however, optical flow cannot be detected for patches without sufficient feature points. Therefore, dense optical flow cannot be determined in such a case.

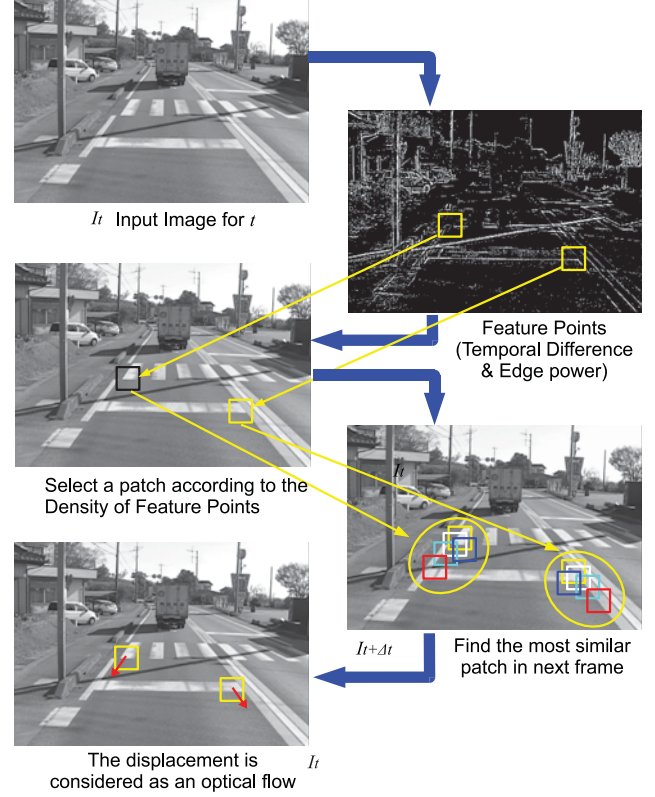


Figure 2. Detection Procedure for Optical Flow

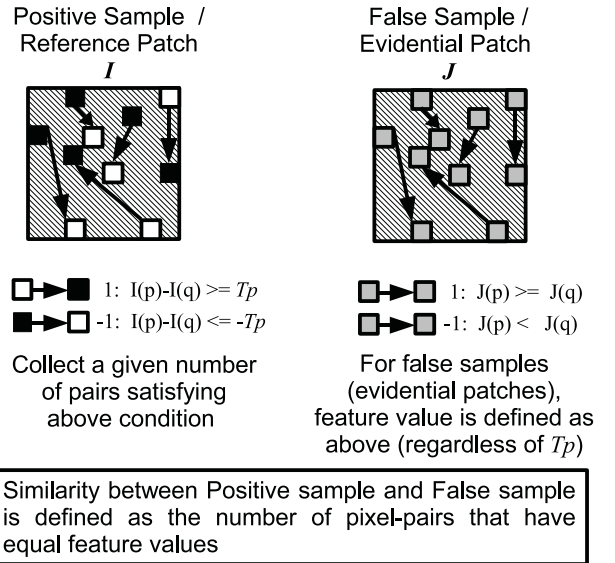


Figure 3. Pixel-pair Features

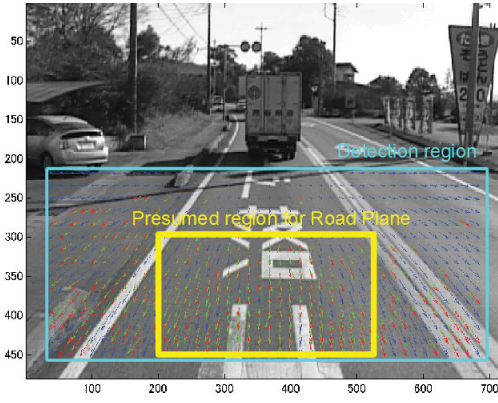


Figure 4. Detected Flow and Estimated Flow

2.3 Determining Differential Homography

Differential homography $\Delta H_t H_t^{-1}$ can be determined by solving equation (7), if reliable optical flows from a road plane region are collected. The reliability of the detected flows are estimated using the similarity index of P_t and $P_{t+\Delta t}$; and the majority of the lower-center part of the image (yellow rectangular area in figure 4) is presumed to be the road plane, although the road plane region is not determined yet (this is the ultimate goal of this paper).

Differential homography is determined by applying RANSAC [17] to the detected flows in the lower-center area (yellow rectangle) of the input image. The procedure is as follows:

1. A certain number of flows in the presumed region are collected.
2. The equation (7) is solved using flows in the yellow rectangular area (green arrows in the yellow area).
3. The estimated differential homography $\Delta H_t H_t^{-1}$ is obtained.
4. Flows are determined using estimated differential homography $\Delta H_t H_t^{-1}$ for comparing with the detected flows (red arrows in figure 4).
5. The flows (green arrows) are compared with the estimated flows (red arrows).
6. If one detected flow is identical to the estimated flow for the patch, the patch represents the road plane.

Once the differential homography $\Delta H_t H_t^{-1}$ is determined, the optical flow for the patches without sufficient feature points can be determined (blue arrows in figure 4).

2.4 Road Plane Detection

A patch is classified as the road plane if it satisfies one of the following conditions:



Figure 5. Example of Road Plane Detection

1. The detected flow and the estimated flow should be identical.
2. According to the estimated flow, the source patch and the destination patch should be identical.

Figure 5 shows an example of road plane detection. In this figure, the cyan area satisfies both condition 1 and 2, the blue area satisfies the condition 2, and the green area satisfies the condition 1.

3 Experimental Results

This section, the experimental results are presented. The proposed algorithm is first applied to the labeled dataset downloaded from “Motion-based Segmentation and Recognition Dataset” of UCL[18]. We examined the sequence of Seq16E5 (city traffic). Next, we applied our algorithm to previously collected suburban traffic dataset (unlabeled).

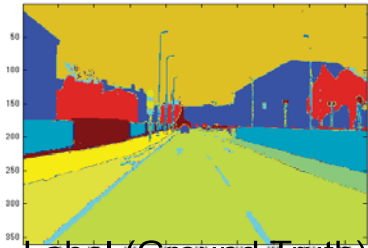
We adopted a 15×15 pixel patch for flow detection and randomly extracted 200 pixel-pairs from each patch. These 200 pairs contain both effective and ineffective pairs, which affect the matching score for patches. Therefore, patches with a low matching score are not employed for flow detection. Even though the proposed algorithm is fundamentally effective for any sort of camera motion, we examined only the straight running sequences for simplifying the optical flow detection. We limited the search range for flow detection to 30 pixels length \times 15 pixels width, which corresponds to approximately 50 Km/h of running speed with some pitching and yawing.

3.1 Result for City Traffic using UCL Dataset

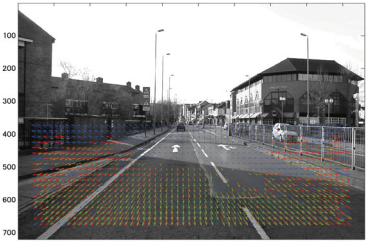
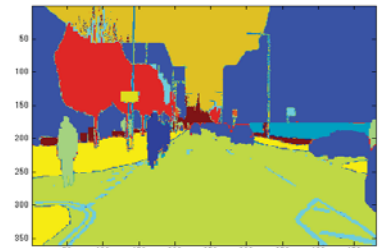
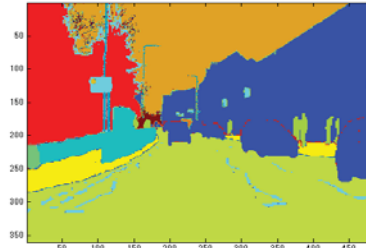
The top row of figure 6 shows the example images from UCL dataset sequence Seq16E5, and the second row shows the label for the image. The input images are obtained at 30 fps, and the labeled images are obtained at 1 fps; thus, detection ratio is computed at 1 fps. In the labeled images, the green region represents the road and the yellow region



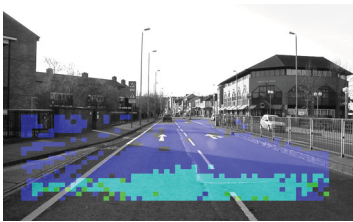
Input Image



Label (Ground Truth)



Detected Flows (Green), Estimated Flows (Blue and Red)

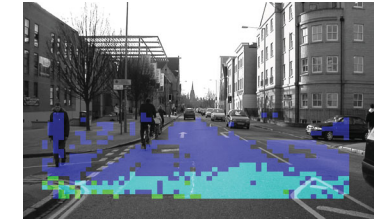


Road Plane Detection Result

(a)



(b)



(c)

Figure 6. Road Plane Estimation for City Traffic

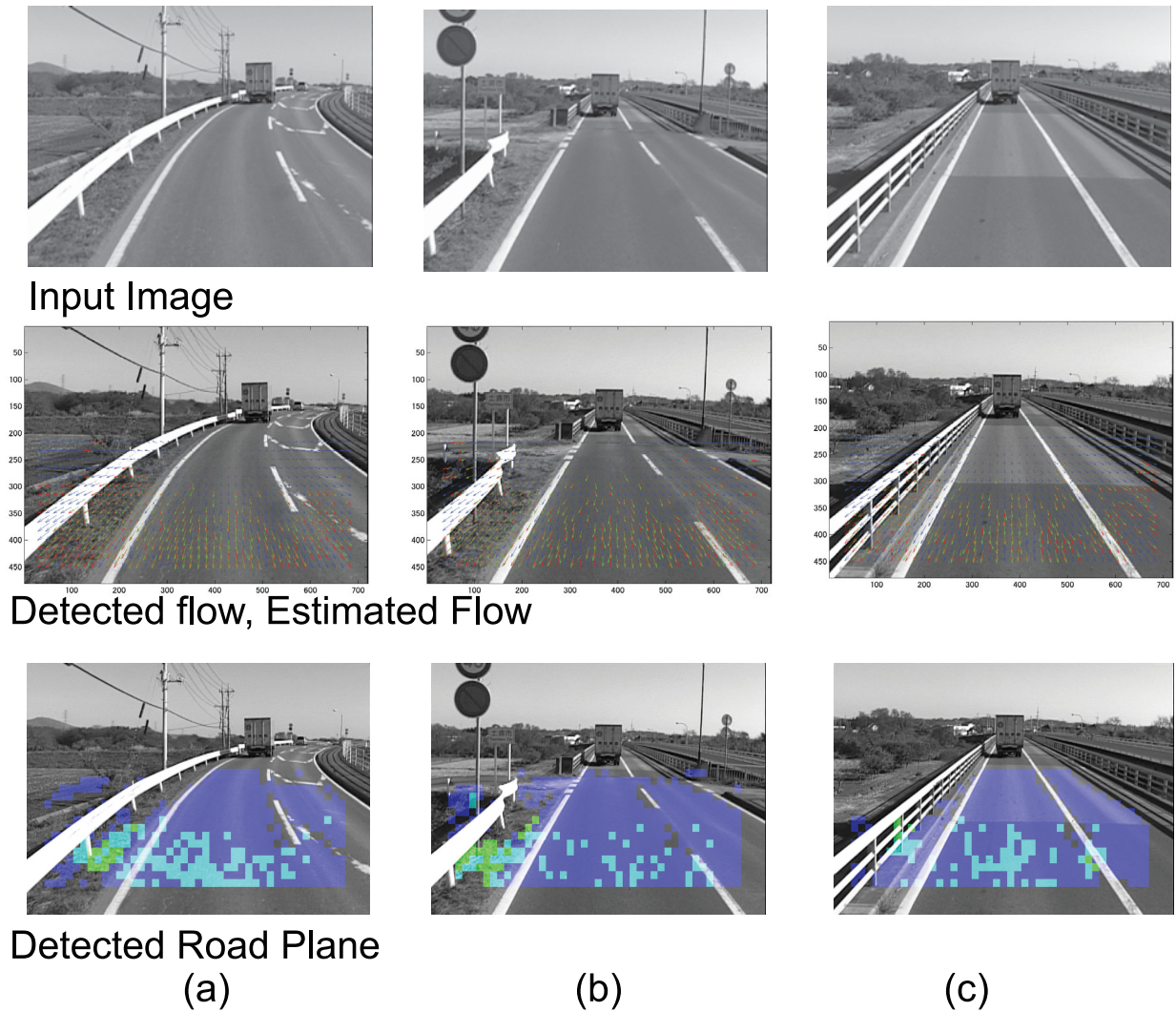


Figure 7. Road Plane Estimation for Suburban Traffic

represents the sidewalk; therefore, we extract these two regions as road plane.

The third row of figure 6 shows the optical flows (green arrow: detected flow; blue and red arrows: estimated flow), and the bottom row shows the detection result of the proposed algorithm. The average detection ratio for the road plane through the sequence was 75.6%.

3.2 Result for Suburban Traffic

Figure 7 shows the detection result for the suburban traffic in the vicinity of our laboratory. Since the ground truth for this dataset was not determined yet, the performance was not evaluated quantitatively. However, it showed a fairly good detection performance. In all examples, the guardrails are effectively omitted from the road plane. The result of figure 7 (a) shows that the detection performance is not affected by the slight curving of the road. In figure 7 (b), the cross road is accurately classified as road plane. In figure

7 (c), the parapet of the bridge, which is critical for driving safety, is effectively omitted.

4 Discussion

The proposed algorithm showed an efficient road plane detection performance, yet it is still imperfect in some ways.

The primary drawback of the proposed algorithm is that it cannot detect anything without camera motion, since the reliable optical flows cannot be detected. Figure 8 shows an example of the unreliable flow estimation for a low running speed.

The second drawback related to our implementation: we limited the search range for optical flow detection. Therefore, a large yawing results in unreliable optical flows, which affects the detection performance.

The third drawback is the faulty detection of patches without textures, such as the patches in the heavily shaded region. While determining the estimated flow using differ-

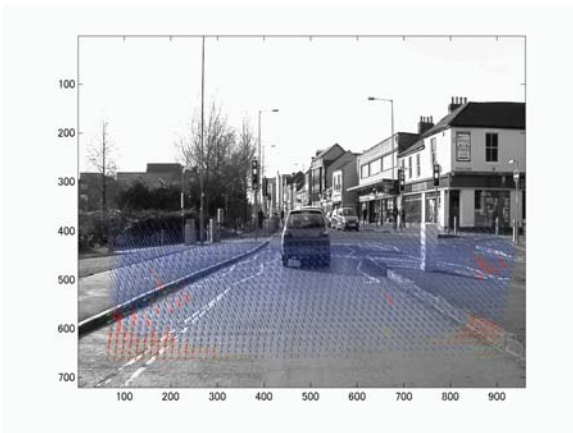


Figure 8. Unreliable Optical Flows in Low Speed

ential homography, no difference between source patch and destination patch can be observed in case of a textureless patches.

To resolve these drawbacks, we are plan to combine some of the following algorithms with our proposed algorithm:

1. Combination with appearance based detection.
2. Combination with rough estimation of camera (car) motion before the flow detection.
3. Considering temporal continuation.
4. Combination with segmentation algorithms.

5 Conclusion

The road plane detection algorithm for an on-board camera is proposed in this paper. The algorithm estimates differential homography from optical flows. Although detecting optical flows from road plane is usually considered difficult, local region matching using pixel-pair feature enabled to detect sufficient optical flows to estimate differential homography from road plane.

Our method showed a good road plane detection performance while it does not require camera calibration (this indicates the estimation of the homography matrix is not required). We are planning to combine our proposed algorithm with other detection algorithm for further improvement of road plane detection.

References

- [1] M.Bellino, YL. de Meneses, P.Ryser, J.Jact, "Lane detection algorithm for an on-board camera", *Photonics in the Automobile (Proceedings of SPIE)*, Thomas P. Pearsall, 2005.
- [2] S.Se, M.Brady,"Road feature detection and estimation", *Machine Vision and Applications*, val.14, pp.157-165, 2003.

- [3] Y.Zhou, R.xu, X.Hu, Q.Ye, "A robust lane detection and tracking method based on computer vision", *Meas. Sci. Technol.*, Vol.17, pp. 736-745, 2006.
- [4] P. Sturm, B. Triggs, "A factorization based algorithm for multi-image projective structure and motion," in *Proc. of European Conference on Computer Vision (ECCV)*, 1996, pp. 709-720.
- [5] B.Leibe, N.Cornelis, K.Cornelis, LV.Gool, "Dynamic 3D Scene Analysis from a Moving Vehicle", in *Proc.CVPR 2007*, 2007.
- [6] "3D Urban Scene Modeling Integrating Recognition and Reconstruction", *International Journal of Computer Vision*, 2007.
- [7] GJ.Brastow, J.Shotton, J.Fauqueur, R.Cipolla, "Segmentation and Recognition using Structure from Motion Point Clouds", in *Proc. ECCV08*, 2008.
- [8] H.Nakai, N.Takeda, H.Hattori, Y.Okamoto, K.Onoguchi, "A Practical Stereo Scheme for Obstacle Detection in Automotive Use", in *Proc. ICPR04*, 2004.
- [9] N.Simond, "Reconstruction of the road plane with an embedded stereo-rig in urban environments", in *Proc. IEEE Intelligent Vehicles Symposium*, pp. 70-75, June 2006.
- [10] R.Okada, Y.Taniguchi, K.Furukawa, K.Onoguchi, "Obstacle Detection Using Projective Invariant and Vanishing Lines", in *proc. ICCV 2003*, 2003.
- [11] B.D.Lucas, T.Kanade, "An iterative image registration technique with an application to stereo vision," in *Proc. of DARPA Image Understanding*, pp.121-130, 1981.
- [12] T.Brox, C.Bregler, J.Malik, "Large Displacement Optical Flow", in *proc. CVPR 2009*, pp.41-48, 2009.
- [13] K.Nishida, T.Kurita, Y.Ogiuchi, M.Higashikubo, "Visual Tracking Algorithm using Pixel-pair Feature", in *proc. ICPR 2010*, 2010.
- [14] R.Ozaki, et al., "Statistical Reach Feature Method and Its Application to Template Matching", in *Proc MVA 2009*, pp.174-177, 2009.
- [15] S. Kaneko, et al., "Robust image registration by increment sign correlation", *Pattern Recognition*, vol.35, no.10, pp.2223-2234, 2002.
- [16] M.Özuysal, et al., "Fast Keypoint Recognition in Ten Lines of Code", in *Proc. CVPR 2007*, 2007.
- [17] M. A. Fischler, R. C. Bolles. "Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography", *Comm. of the ACM*, Vol. 24, pp. 381-395, 1981.
- [18] <http://mi.eng.cam.ac.uk/research/projects/VideoRec/CamVid/>