

Fast Plane Parameter Estimation From Stereo Images

Shigeki Sugimoto and Masatoshi Okutomi

Department of Mechanical and Control Engineering, Graduate School of Science and Engineering,
Tokyo Institute of Technology
2-11-2-S5-22 O-okayama, Meguro-ku, Tokyo, 152-8550 Japan
{shige, mxo}@ok.ctrl.titech.ac.jp

Abstract

For controlling a vehicle or a mobile robot which moves around on the ground, to estimate the plane parameters (i.e. the distance and plane normal) of the ground plane is an important issue. In this paper, we propose an efficient and robust method for plane parameter estimation from stereo images. Our approach is based on a “direct” method, which makes full use of pixel data in two image regions, for estimating the plane parameters. We formulate a fast direct method for the plane parameters by incorporating inverse compositional expression into the SSD function between two images. The proposed method is about eight times faster than the conventional method without loss of precision, and has the advantage of more robustness and efficiency to a method based on the ICIA algorithm of homography estimation. The validity of the proposed method is demonstrated through experiments using synthetic and real images.

1. Introduction

To obtain the 3D geometric parameters (i.e. the distance and the plane normal) of the ground plane is an important issue for controlling a vehicle or a mobile robot which moves around on the ground. These parameters are required for deciding the action of a vehicle facing a slope or a difference in the ground level. Especially for a biped robot which dramatically changes its pose and shape relative to the plane, the 3D parameters will be indispensable for inputs to joint control.

Because typical roads and floors have few feature points, *direct* methods [3], where every pixel directly contributes to a measurement, are more preferable than feature based methods [8]. There are mainly two approaches based on the direct methods for estimating the plane parameters.

Since the image deformation between two images with respect to the plane can be specified by a homography (eight-parameter projective deformation), we can estimate the eight parameters of a homography by an Inverse Compositional Image Alignment (ICIA) algorithm [1], and successively apply singular value decomposition (SVD) for obtaining plane parameters [6]. Although the eight parameters can be efficiently estimated by the ICIA algorithm, this approach tends to lack robustness because we have to solve

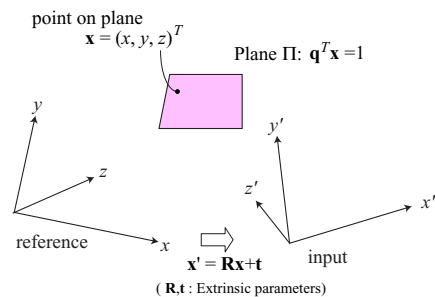


Figure 1. Stereo configuration and plane parameters

the problem of eight degrees of freedom (d.o.f).

On the other hand, we can formulate a direct method, similar to [4], for the plane parameters of three d.o.f which are equivalent to homography parameters under epipolar constraints. However, this method is more time-consuming than the method based on the ICIA algorithm for homography estimation.

In this paper, we propose a novel formulation of a direct method for efficiently estimating the plane parameters from stereo images. In our formulation, we incorporate inverse compositional expression into the SSD function between the reference image and an image warped from the other by the plane parameters. Our method is about eight times faster than the conventional method without loss of precision, and about two times faster than the full eight parameter estimation using the ICIA algorithm. The validity of the proposed method is shown by some experiments using synthetic and real images.

2. Efficient Direct Plane Parameter Estimation

2.1. Homography – Basic Notations

Let \mathbf{x} and \mathbf{x}' denote a scene point with respect to the two different camera views. We write: $\mathbf{x}' = \mathbf{R}\mathbf{x} + \mathbf{t}$, where \mathbf{R} and \mathbf{t} respectively denote the rotation matrix and the translation vector between the two camera coordinate frames (see Fig1).

Let $I[\mathbf{u}]$ and $I'[\mathbf{u}']$ be the pixel values of the reference image I and the other image I' , respectively, where $\mathbf{u} = (u, v)^T$ and $\mathbf{u}' = (u', v')^T$ respectively denote the corresponding points in I and I' . For avoiding complexity, let \mathbf{u} and \mathbf{u}' be in the *canonical* image configuration.

Let Π be a plane with a unit plane normal \mathbf{n} and a distance d in the 3D coordinate frame of the reference camera.

Thereby, the relationship between \mathbf{u} and \mathbf{u}' can be written as a 3×3 homography matrix \mathbf{P} as follows [2]:

$$\tilde{\mathbf{u}}' \sim \mathbf{P}\tilde{\mathbf{u}}, \quad (1)$$

$$\text{where } \mathbf{P} = \mathbf{R} + \mathbf{t}\mathbf{q}^T, \quad \mathbf{q} \equiv \mathbf{n}/d. \quad (2)$$

Note that \mathbf{q} determines a plane equation in the reference camera coordinate frame as $\mathbf{q}^T \mathbf{x} = 1$. In this section, we refer to \mathbf{q} as the plane parameter vector to be estimated.

2.2. Conventional Direct Method

Let $\mathbf{w}(\mathbf{u}; \mathbf{p})$ denote the homography warps derived from (1), where $\mathbf{p} = (p_1, p_2, \dots, p_9)^T$ is a homography parameter vector which is a function of \mathbf{q} , as indicated by (2). A conventional direct method for estimating \mathbf{q} minimizes the SSD function as follows:

$$\sum_{\mathbf{u} \in \text{ROI}} \{I[\mathbf{u}] - I'[\mathbf{w}(\mathbf{u}; \mathbf{p}(\bar{\mathbf{q}} + \Delta\mathbf{q}))]\}^2, \quad (3)$$

where $\bar{\mathbf{q}}$ and $\Delta\mathbf{q}$ respectively denote a known current estimate and unknown increments of \mathbf{q} . Herein, ROI denotes a region of interest in I . Equation (3) is essentially equivalent to an optical flow based representation in [4].

Applying Gauss-Newton optimization to (3) yields

$$\Delta\mathbf{q} = -\mathbf{H}^{-1}\mathbf{b}, \quad (4)$$

$$\text{where } \mathbf{H} \equiv \sum_{\mathbf{u} \in \text{ROI}} \left\{ \left[\frac{\partial I'}{\partial \mathbf{w}} \frac{\partial \mathbf{w}}{\partial \mathbf{p}} \frac{\partial \mathbf{p}}{\partial \mathbf{q}} \right]^T \left[\frac{\partial I'}{\partial \mathbf{w}} \frac{\partial \mathbf{w}}{\partial \mathbf{p}} \frac{\partial \mathbf{p}}{\partial \mathbf{q}} \right] \right\}, \quad (5)$$

$$\mathbf{b} \equiv \sum_{\mathbf{u} \in \text{ROI}} \left\{ e \left[\frac{\partial I'}{\partial \mathbf{w}} \frac{\partial \mathbf{w}}{\partial \mathbf{p}} \frac{\partial \mathbf{p}}{\partial \mathbf{q}} \right]^T \right\}, \quad (6)$$

$$e \equiv I[\mathbf{u}] - I'[\mathbf{w}(\mathbf{u}; \mathbf{p}(\bar{\mathbf{q}}))]. \quad (7)$$

The final estimate of \mathbf{q} is obtained after iterations in each of which $\bar{\mathbf{q}}$ is updated by $\bar{\mathbf{q}} \leftarrow \bar{\mathbf{q}} + \Delta\mathbf{q}$ after computing $\Delta\mathbf{q}$ by (4). In the conventional case, $\partial I'/\partial \mathbf{w}$ and $\partial \mathbf{w}/\partial \mathbf{p}$ at each pixel should be re-computed in each iteration because these differentials are evaluated at the current estimate of $\bar{\mathbf{q}}$. These per-pixel and per-iteration computations for obtaining \mathbf{H} impart large computational costs on the conventional method.

2.3. Our Proposed Method

A homography matrix can be written as

$$\mathbf{P} = \bar{\mathbf{P}}[\mathbf{I} + \Delta\mathbf{P}]^{-1}, \quad (8)$$

where $\bar{\mathbf{P}}$ and $\Delta\mathbf{P}$ respectively represent a current estimate of \mathbf{P} and a matrix with small elements. Let $\bar{\mathbf{p}}$ and $\Delta\mathbf{p}$ be parameter vectors composed respectively by the elements of $\bar{\mathbf{P}}$ and $\Delta\mathbf{P}$. Equation (8) is a fundamental in the ICIA algorithm [1] for fast homography estimation.

Assume that $\bar{\mathbf{p}}$ and $\Delta\mathbf{p}$ are functions of $\bar{\mathbf{q}}$ and $\Delta\mathbf{q}$, respectively, and $\Delta\mathbf{p} \rightarrow \mathbf{0}$ when $\Delta\mathbf{q} \rightarrow \mathbf{0}$, where $\mathbf{q} = \bar{\mathbf{q}} + \Delta\mathbf{q}$ (concrete expressions are presented later). Then we re-write (3) as

$$\sum_{\mathbf{u} \in \text{ROI}} \{I[\Delta\mathbf{w}(\mathbf{u}; \Delta\mathbf{p}(\Delta\mathbf{q}))] - I'[\mathbf{w}(\mathbf{u}; \bar{\mathbf{p}}(\bar{\mathbf{q}}))]\}^2 \quad (9)$$

where $\mathbf{w}(\mathbf{u}; \bar{\mathbf{p}})$ and $\Delta\mathbf{w}(\mathbf{u}; \Delta\mathbf{p})$ denote the homography warps derived respectively from $\tilde{\mathbf{u}}' \sim \bar{\mathbf{P}}\tilde{\mathbf{u}}$ and $\tilde{\mathbf{u}}' \sim [\mathbf{I} + \Delta\mathbf{P}]\tilde{\mathbf{u}}$.

Applying Gauss-Newton optimization to (9) yields

$$\Delta\mathbf{q} = -\mathbf{H}^{-1}\mathbf{b}, \quad (10)$$

where

$$\mathbf{H} \equiv \sum_{\mathbf{u} \in \text{ROI}} \left\{ \left[\frac{\partial I}{\partial \Delta\mathbf{w}} \frac{\partial \Delta\mathbf{w}}{\partial \Delta\mathbf{p}} \frac{\partial \Delta\mathbf{p}}{\partial \Delta\mathbf{q}} \right]^T \left[\frac{\partial I}{\partial \Delta\mathbf{w}} \frac{\partial \Delta\mathbf{w}}{\partial \Delta\mathbf{p}} \frac{\partial \Delta\mathbf{p}}{\partial \Delta\mathbf{q}} \right] \right\},$$

$$\mathbf{b} \equiv \sum_{\mathbf{u} \in \text{ROI}} \left\{ e \left[\frac{\partial I}{\partial \Delta\mathbf{w}} \frac{\partial \Delta\mathbf{w}}{\partial \Delta\mathbf{p}} \frac{\partial \Delta\mathbf{p}}{\partial \Delta\mathbf{q}} \right]^T \right\}, \quad (11)$$

$$e \equiv I[\mathbf{u}] - I'[\mathbf{w}(\mathbf{u}; \mathbf{p}(\bar{\mathbf{q}}))]. \quad (12)$$

As in the conventional method, $\bar{\mathbf{q}}$ is updated by $\bar{\mathbf{q}} \leftarrow \bar{\mathbf{q}} + \Delta\mathbf{q}$ in each iteration. However, in this case, $\partial I/\partial \Delta\mathbf{w}$ and $\partial \Delta\mathbf{w}/\partial \Delta\mathbf{p}$ are constant in each iteration because these differentials are evaluated at $\Delta\mathbf{q} = \mathbf{0}$ (i.e. $\Delta\mathbf{p} = \mathbf{0}$) [1]. $\partial I/\partial \Delta\mathbf{w}$ denotes the gradients of the reference image. In addition, $\partial \Delta\mathbf{w}/\partial \Delta\mathbf{p}$ can be written as

$$\frac{\partial \Delta\mathbf{w}}{\partial \Delta\mathbf{p}} = \begin{bmatrix} u & v & 1 & 0 & 0 & 0 & -u^2 & -uv & -u \\ 0 & 0 & 0 & u & v & 1 & -uv & -v^2 & -v \end{bmatrix}. \quad (13)$$

The remainder $\partial \Delta\mathbf{p}/\partial \Delta\mathbf{q}$ depends on the expression of $\Delta\mathbf{p}(\Delta\mathbf{q})$. We derive it next.

From (2), we can write

$$\mathbf{P} = \mathbf{R} + \mathbf{t}[\bar{\mathbf{q}} + \Delta\mathbf{q}]^T. \quad (14)$$

The relationship between $\Delta\mathbf{p}$ and $\Delta\mathbf{q}$ is defined by re-writing (14) into the form of (8).

Sherman-Morrison's formula [5] gives the inverse of (14):

$$\mathbf{P}^{-1} = \mathbf{R}' + \mathbf{t}'[\bar{\mathbf{q}}' + \Delta\mathbf{q}]^T, \quad (15)$$

$$\text{where } \mathbf{R}' \equiv \mathbf{R}^{-1}, \quad \mathbf{t}' \equiv -\mathbf{R}^{-1}\mathbf{t}, \quad (16)$$

$$\bar{\mathbf{q}}' + \Delta\mathbf{q}' \equiv \frac{\mathbf{R}[\bar{\mathbf{q}} + \Delta\mathbf{q}]}{1 + [\bar{\mathbf{q}} + \Delta\mathbf{q}]^T \mathbf{R}^{-1} \mathbf{t}} \quad (17)$$

Re-writing (15) yields

$$\mathbf{P}^{-1} = [\mathbf{I} + \mathbf{t}'\Delta\mathbf{q}'^T \bar{\mathbf{P}}] \bar{\mathbf{P}}^{-1}, \quad (18)$$

$$\text{where } \bar{\mathbf{P}} \equiv \mathbf{R} + \mathbf{t}\bar{\mathbf{q}}^T. \quad (19)$$

The inverse of (18) is written as

$$\mathbf{P} = \bar{\mathbf{P}}[\mathbf{I} + \mathbf{t}'\Delta\mathbf{q}'^T \bar{\mathbf{P}}]^{-1}. \quad (20)$$

Therefore, comparing (20) with (8) gives

$$\Delta\mathbf{P} = \mathbf{t}'\Delta\mathbf{q}'^T \bar{\mathbf{P}}. \quad (21)$$

Finally, substituting (17) and (16) for (21) obtains

$$\Delta\mathbf{P} = -\frac{1}{1 + \bar{\mathbf{q}}^T \mathbf{R}^T \mathbf{t} + \Delta\mathbf{q}^T \mathbf{R}^T \mathbf{t}} \mathbf{R}^T \mathbf{t} \Delta\mathbf{q}^T. \quad (22)$$

Eq. (22) is directed to the expression of $\partial \Delta\mathbf{p}/\partial \Delta\mathbf{q}$ evaluated at $\Delta\mathbf{q} = \mathbf{0}$. We derive

$$\frac{\partial \Delta\mathbf{p}}{\partial \Delta\mathbf{q}} = \frac{1}{\kappa} \begin{bmatrix} \alpha_1 & 0 & 0 & \alpha_2 & 0 & 0 & \alpha_3 & 0 & 0 \\ 0 & \alpha_1 & 0 & 0 & \alpha_2 & 0 & 0 & \alpha_3 & 0 \\ 0 & 0 & \alpha_1 & 0 & 0 & \alpha_2 & 0 & 0 & \alpha_3 \end{bmatrix}^T, \quad (23)$$

where

$$\kappa \equiv -(1 + \bar{\mathbf{q}}^T \mathbf{R}^T \mathbf{t}), \quad (24)$$

$$\alpha_1 \equiv t_1 r_1 + t_2 r_4 + t_3 r_7, \quad \alpha_2 \equiv t_1 r_2 + t_2 r_5 + t_3 r_8,$$

$$\alpha_3 \equiv t_1 r_3 + t_2 r_6 + t_3 r_9, \quad (25)$$

Therein, $t_i (i = 1, 2, 3)$ and $r_j (j = 1, \dots, 9)$ respectively denote the elements of \mathbf{t} and \mathbf{R} .

Unfortunately, $\partial \Delta \mathbf{p} / \partial \Delta \mathbf{q}$ can not be constant in each iteration because κ depends on the current estimate $\bar{\mathbf{q}}$, which varies from iteration to iteration. However, κ is a simple scalar independent of pixels. Therefore, the final Gauss-Newton optimization algorithm can be written as

$$\Delta \mathbf{q} = -\kappa \mathbf{H}'^{-1} \mathbf{b}', \quad (26)$$

where

$$\mathbf{H}' \equiv \sum_{\mathbf{u} \in \text{ROI}} \left\{ \left[\frac{\partial I}{\partial \Delta \mathbf{w}} \frac{\partial \Delta \mathbf{w}}{\partial \Delta \mathbf{p}} \left[\kappa \frac{\partial \Delta \mathbf{p}}{\partial \Delta \mathbf{q}} \right]^T \right] \times \left[\frac{\partial I}{\partial \Delta \mathbf{w}} \frac{\partial \Delta \mathbf{w}}{\partial \Delta \mathbf{p}} \left[\kappa \frac{\partial \Delta \mathbf{p}}{\partial \Delta \mathbf{q}} \right]^T \right] \right\}, \quad (27)$$

$$\mathbf{b}' \equiv \sum_{\mathbf{u} \in \text{ROI}} \left\{ e \left[\frac{\partial I}{\partial \Delta \mathbf{w}} \frac{\partial \Delta \mathbf{w}}{\partial \Delta \mathbf{p}} \left[\kappa \frac{\partial \Delta \mathbf{p}}{\partial \Delta \mathbf{q}} \right]^T \right] \right\}, \quad (28)$$

$$e \equiv I[\mathbf{u}] - I'[\mathbf{w}(\mathbf{u}; \mathbf{p}(\bar{\mathbf{q}}))]. \quad (29)$$

Compared with (10), we simply replace $\partial \Delta \mathbf{p} / \partial \Delta \mathbf{q}$ by $\kappa \partial \Delta \mathbf{p} / \partial \Delta \mathbf{q}$ which is a constant 9×3 matrix composed by the elements of \mathbf{R} and \mathbf{t} , as described in (23). This replacement renders the product (1×3 column vector) of $\partial I / \partial \Delta \mathbf{w}$, $\partial \Delta \mathbf{w} / \partial \Delta \mathbf{p}$, and $\kappa \partial \Delta \mathbf{p} / \partial \Delta \mathbf{q}$ at each pixel constant in each iteration. Therefore, \mathbf{H}' and its inverse can be pre-computed before the iteration process.

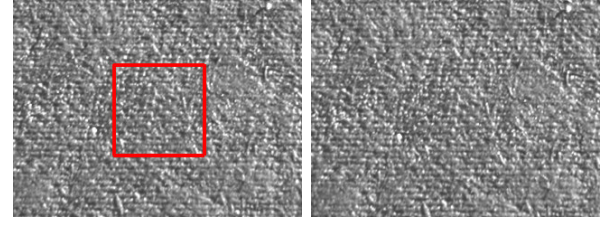
Note that $\partial \Delta \mathbf{w} / \partial \Delta \mathbf{p}$ and $\kappa \partial \Delta \mathbf{p} / \partial \Delta \mathbf{q}$ are independent of the scene. The product (2×3 matrix) of the two matrices at every pixel can be computed when the cameras are calibrated. Additionally, the computational costs of κ are negligible compared to \mathbf{b}' , which requires per-pixel computations in each iteration. Therefore, we can consider that the per-parameter computational costs of this algorithm are almost equivalent to those of the ICIA algorithm; the proposed method is faster than the ICIA algorithm of homography estimation.

3. Experimental Results

We demonstrate the capability of the proposed method by a comparative study in simulation using three methods that estimate plane parameters. We also show the experimental results of the proposed method for real environments.

3.1. Simulation

In our comparative study, the convergence stability and the computational time of three methods were evaluated. These methods are the conventional direct method described in Section 2.2, a method which combines the ICIA algorithm for estimating homography [1] with SVD for obtaining the plane parameters as presented in [7], and our proposed method presented in Section 2.3.



(a) Reference image (b) The other image

Figure 2. Stereo images used in simulation for plane parameter estimation: (a) is an example of reference images I_0 . The reference images are created by warping the other image (b) using randomly generated plane vectors. The rectangle denotes a 100×100 ROI whose position is fixed.

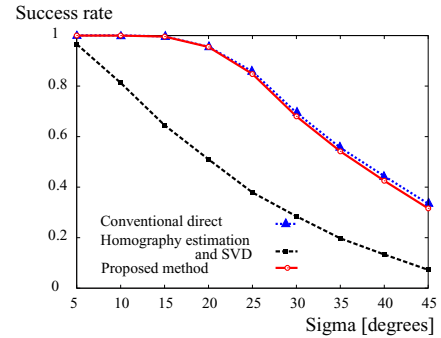


Figure 3. Success rate after five iterations. For each σ , 5000 sets of plane parameters were generated randomly.

We first set the plane parameters of a reference plane as $\mathbf{q}_0 = \mathbf{n}_0 / d_0$, where $\mathbf{n}_0 = (0, 0, 1)^T$ and $d_0 = 15.24$ (in meters). We referred to the fixed reference plane for setting a target plane, as the reference plane was rotated and shifted by a set of three values that were generated randomly by zero-mean Gaussian noise of a certain standard deviation σ .

Using each set of the three values, we computed the plane parameters \mathbf{q} of the target plane and warped the other image I' , as shown in Fig.2(b), by the target plane parameters for creating a reference image, as in Fig.2(a). We also added Gaussian noise of standard deviation 4 gray levels to both images. Then we run the algorithms of the three methods starting from \mathbf{q}_0 .

For each σ , we evaluated the frequency of convergence over 5000 trials. We judged that an algorithm succeeded after five iterations, if the angle between \mathbf{q} and the estimates was less than 0.5 degrees. The results are shown in Fig. 3.

Figure 3 shows that the two direct methods for plane parameter estimation give far higher success rates than the method for homography. Under the epipolar constraints of the stereo images, the image deformation is restricted to the epipolar lines. This restriction is remarkably profitable for acquiring high stability.

Table 1 shows the computational time of the three methods. Comparing our method with the others shows that the five-iteration time of our method is about one-eighth as long as the conventional method, and about one-twice as long as the homography estimation method.

3.2. Real Stereo Image Sequence

We mounted two cameras on a car, and sequentially estimated the 3D parameters of the ground plane by the pro-

Method	pre-computation	per iteration	total(5 iterations)	total(100 iterations)
Conventional direct	—	15.19	75.95	1519
Homography estimation and SVD	11.09	2.613	24.16	272.4
Proposed method	2.526	1.312	9.087	133.5

Table 1. Computational time (ms) in plane parameter estimation: Averaged over 100 trials for a 100×100 ROI. All algorithms were implemented using C-language and run on a Linux PC (Pentium-IV 2.8 GHz).

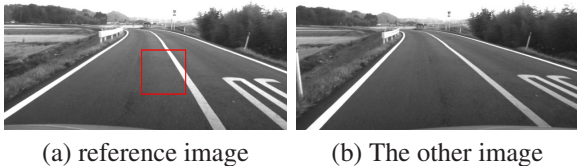


Figure 4. A reference image and the other image (640x480 pixels) from two calibrated cameras mounted on a car. A 100×100 ROI fixed on the reference image is used for sequentially estimating the plane parameters of the ground.

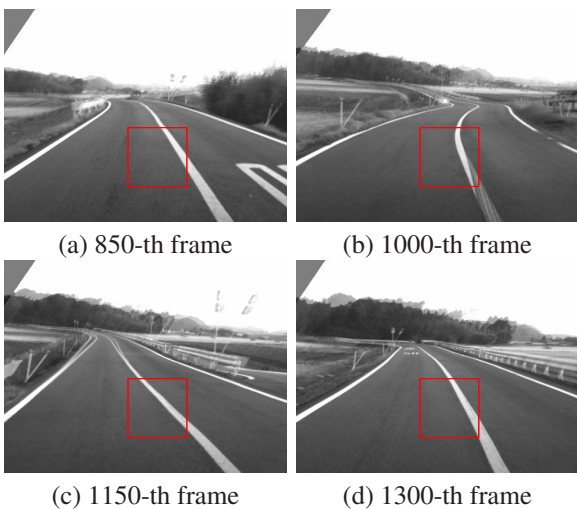


Figure 5. Estimation results for real environments: At each frame number, the reference image and the image warped from the other image by the estimated parameters are overlapped.

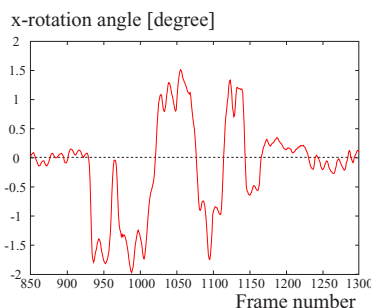


Figure 6. Rotation angle with respect to x-axis (parallel to the horizontal axis of the image plane) at each frame number. Zero indicates that the estimated normal has the same angle of the plane normal obtained at a static situation.

posed method. Fig.4 shows an image pair in the stereo image sequences. In this experiment, we first estimated initial plane parameters at a static situation. The initial param-

eters were used at the first estimation. Then the parameters estimated at each frame number were sequentially used as initial parameters at the next.

Fig.5 shows the results of the plane parameter estimation. Each image in this figure shows the overlapped image of the reference image and the image warped from the other image by the estimated plane parameters. We see that the plane parameters were satisfactorily estimated at each frame since the warped image corresponds to the reference image not only in the ROI but also in a large area on the road region.

Fig.6 shows the rotation angle between the initial parameters estimated at a static situation and the parameters estimated at each frame, with respect to x-axis (parallel to the horizontal axis of the image plane). In the actual situation, the car went up a uphill slope and went down an easy two-step slope. We can see the slope shapes from Fig.6.

4. Conclusions

We have proposed an efficient direct method for plane parameter estimation using stereo images. The proposed method is far faster than the conventional method, and has the advantage of more robustness and efficiency over the ICIA algorithm of homography estimation. We will extend the proposed method to multi-baseline stereo or planar region detection in the future research.

References

- [1] S. Baker and I. Matthews. Lucas-kanade 20 years on: A unifying framework. *IJCV*, 56(3):221–255, 2004.
- [2] O. Faugeras and F. Lustman. Motion and structure from motion in a piecewise planar environment. *Report de Recherche de l'INRIA*, 1988.
- [3] B. K. P. Horn and J. E. J. Weldon. Direct methods for recovering motion. *IJCV*, 2:51–76, 1988.
- [4] Q. Ke and T. Kanade. Transforming camera geometry to a virtual downward-looking camera: Robust ego-motion estimation and ground-layer detection. In *CVPR*, volume 1, pages 390–397, 2003.
- [5] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical Recipes in C*. Cambridge Univ Press, 1992.
- [6] A. Seki and M. Okutomi. Ego-motion estimation by matching dewarped road regions using stereo images. In *ICRA2006*, pages 901–907, 2006.
- [7] A. Seki and M. Okutomi. Robust obstacle detection in general road environment based on road extraction and pose estimation. In *IV2006*, pages 437–444, 2006.
- [8] G. Stein, O. Mano, and A. Shashua. A robust method for computing vehicle ego-motion. In *IV2000*, pages 362–368, 2000.