社団法人 情報処理学会　研究報告
IPSJ SIG Technical Report

2004－CVIM－145　(7)
2004／9／10

オプティカルフローによるドミナントプレーン抽出

大西 直哉 † 　井宮 淳 ††

† 千葉大学　自然科学研究科 〒263-8522 千葉市稲毛区弥生町 1-33　Tel +81-43-290-3257.
†† 千葉大学総合メディア基盤センター　〒263-8522 千葉市稲毛区弥生町 1-33　Tel +81-3-4212-2000,
†,†† {ohnishi,imiya}@media.imit.chiba-u.ac.jp

あらまし　オプティカルフローを用いた移動ロボットのためのドミナントプレーン推定法を提案する。ドミナントプレーンとはカメラの搭載されたロボットから見て最も広い領域を占める平面である。これは実世界の床面などに対応し、ロボットの移動可能な領域となる。したがって、視覚情報からのドミナントプレーンを推定することは、ロボットの自律行動のために重要な仕事である。本論文では、ドミナントプレーンとそれ以外の領域でのオプティカルフローの差異によって、ドミナントプレーンを抽出する。そこで、ドミナントプレーンのオプティカルフローがアフィン変換で近似できることを示し、ドミナントプレーンのフローを推定する。実験では、人工画像と実画像を用いて、ドミナントプレーン推定の精度を調べた。

キーワード　ロボットビジョン，オプティカルフロー，ドミナントプレーン，ホモグラフィー.

Dominant plane detection from optical flow for robot navigation

Naoya OHNISHI†　and　Atsushi IMIYA††

†School of Science and Technology, Chiba University　Yayoicho 1-33, Inage-ku, 263-8522, Chiba, Japan　Tel +81-43-290-3257.
††IMIT　Yayoi-cho 1-33, Inage-ku, 263-8522, Chiba, Japan　+81-43-290-3257.
†,†† {ohnishi,imiya}@media.imit.chiba-u.ac.jp

Abstract　In this paper, we develop an algorithm for dominant plane detection using the optical flow observed by means of a vision system mounted on an autonomous mobile robot. The dominant plane estimation is an essential task for the autonomous navigation and the path planning of the mobile robot, since the dominant plane occupies the largest domain in the image and the robot moves on the dominant plane. We show that the points on the dominant plane in a pair of two successive images are combined with an affine transformation if the mobile robot obtains successive images for optical flow computation. Therefore, our algorithm detects the dominant plane from images observed by an uncalibrated camera without any assumptions of camera displacement. We experiment with the detection of the dominant plane using a real image sequence and evaluate an error with a test image sequence.

Key Words　Robot vision, Optical flow, Dominant plane, Homography.

1 Introduction

In this work, we aim to develop an algorithm for dominant plane detection using the optical flow observed by means of a vision system mounted on an autonomous mobile robot. The dominant plane is the planar area which occupies the largest domain in the image observed by a camera. Assuming that the robot moves on a planar area, dominant plane estimation is an essential task for the autonomous navigation and the path planning of the mobile robot, since the robot moves on the dominant plane (e.g., floors and ground areas). We show that the points on the dominant plane in a pair of two successive images are combined with an affine transformation if the mobile robot obtains successive images for optical flow computation. Therefore, our algorithm detects the dominant plane without calibration of the camera or any assumptions of camera displacement.

For the autonomous navigation of mobile robots, vision sensors and sonar or laser sensors are generally used. Sonar and laser sensors [1] provide simple methods of obstacle detection. These sensors are effective in obstacle detection for collision avoidance, since these methods can obtain range information to an object. On the other hand, stationary vision sensors have difficulty in obtaining range information. However, vision sensors mounted on a mobile robot can obtain an image sequence from the camera motion. The image sequence provides the motion and structure from correspondences of points on successive images [2]. Additionally, vision sensors are fundamental devices for the understanding of the environment, since the robot collaborates with a human being. Moreover, visual information is valid for the path planning of the mobile robot in a long sequence, because the vision system can capture environmental information quickly for a large area compared to present sonar- and laser-based systems.

There are many methods for the detection of obstacles or planar areas using vision systems [3]. For example, the edge detection of omni and monocular camera systems [4] and the observation of landmarks [5] are the classical ones. However, since these methods are dependent on the environment around a robot, they are difficult to apply in general environments. If a robot captures an image sequence of moving objects, the optical flow [6] [7] [8], which is the motion of the scene, is obtained for the fundamental features in order to construct environment information around the mobile robot. Additionally, the optical flow is considered as fundamental information for the obstacle detection in the context of biological data processing [9]. Therefore, the use of optical flow is an appropriate method from the viewpoint of the affinity between the robot and human being.

Haag and Nagel[10] extracted moving cars from a sequence of images observed by a fixed camera using optical flow and the model fitting technique. Furthermore, they applied the same methodology to human motion tracking [11]. We develop the detection of stationary obstacles from an image sequence observed from a mobile robot. This problem is the dual problem to the detection of moving objects from an image sequence observed from a stationary camera, since the motion of objects and the camera have geometrically dual properties. We address the problem of detecting obstacles and the dominant plane using optical flow from an image sequence observed by a moving robot for the purpose of navigation.

The obstacle detection using optical flow is proposed in [12] [13]. Enkelmann [12] proposed an obstacle-detection method using the model vectors from motion parameters. Santos-Victor and Sandini [13] also proposed an obstacle-detection algorithm for a mobile robot using the inverse projection of optical flow to ground floor, assuming that the motion of the camera system mounted on a robot is pure translation with a uniform velocity. However, even if a camera is mounted on a wheel-driven robot, the vision system does not move with uniform velocity due to mechanical errors of the robot and unevenness of the floor.

In this work, we develop an algorithm for the estimation of the optical flow field of the dominant plane from a moving uncalibrated camera. For the estimation of the optical flow field of the dominant plane, we show that the homography of the image sequence of the dominant plane is calculated approximately by affine transformation in Section 2. Furthermore, our algorithm detects obstacles by applying pattern matching to the computed optical flow field and estimated flow field. Our method dose not require any physical assumptions for the robot motion. Moreover, our algorithm detects the dominant plane without camera calibration. We show experiments for the detection of the dominant plane using an image sequence for the evaluation of optical flow computation. Moreover, using a real image sequence observed by the camera mounted on a mobile robot, we evaluate the validity of own method for long image sequences

2 Approximation of the dominant plane motion by affine transformation

In this section, we show that the corresponding points in a pair of successive images which are a projection of the dominant plane in a space are connected by an affine transformation. Therefore, the corresponding points (u, v) and (u', v') on the dominant plane are expressed by

$$\begin{pmatrix} u' \\ v' \end{pmatrix} = \begin{pmatrix} a & b \\ d & e \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} + \begin{pmatrix} c \\ f \end{pmatrix}, \tag{1}$$

if the camera displacement is small. That means, a homography between the two images of a planar surface is approximated by an affine transformation if the camera displacement is small.

Setting $\mathbf{H}$ to be the 3×3 matrix [14], the homography between the two images of a planar surface is expressed as

$$\mathbf{p} = \mathbf{H}\mathbf{p}', \tag{2}$$

where $\mathbf{p} = (u, v, 1)^\top$ and $\mathbf{p}' = (u', v', 1)^\top$ are the corresponding points of two images. The matrix $\mathbf{H}$ is expressed as

$$\mathbf{H} = \mathbf{K}(\mathbf{R} + \mathbf{t}\mathbf{n}^\top)\mathbf{K}^{-1}, \tag{3}$$

where $\mathbf{K}$, $\mathbf{R}$, $\mathbf{t}$, and $\mathbf{n}$ are the projection matrix, the rotation matrix, the translation vector, and the plane normal of the planar surface, respectively. The matrixes $\mathbf{K}$ and $\mathbf{K}^{-1}$ are affine transformations since those matrixes are the projection matrix of the pinhole camera. Assuming that the camera displacement is small, the matrix $\mathbf{R}$ and the matrix $\mathbf{t}\mathbf{n}^\top$ are approximated by an affine transformation. These geometrical and mathematical assumptions are valid when the camera mounted on the robot moves on the dominant plane. These assumptions allow us to describe the relationship between (u', v') and (u, v) as the affine transformation. Therefore, setting *a*,*b*,*c*,*d*,*e*,*f* to be affine coefficients, Eq.(2) is expressed as Eq.(1). In the next section, we develop an algorithm for the estimation of these six parameters from a sequence of images.

3 Dominant plane detection from image sequence

In this section, we develop an algorithm for the detection of the dominant plane from the image sequence observed by a camera mounted on a mobile robot. When the camera mounted on the mobile robot moves on the ground plane, we obtain successive images which include a dominant plane area and obstacles. Therefore, the computed optical flow from the successive images describes the motion of the dominant plane and obstacles on the basis of the camera displacement. Since the corresponding points in the dominant plane motion are combined with an affine transformation, we can compute the affine coefficients in Eq.(1) from the optical flow on the dominant plane. Once the affine coefficients are computed, we can estimate the dominant plane motion in the image from the affine coefficients. The dominant plane motion is a *planar flow field* as shown in Fig.1. The difference between the estimated planar flow field and the computed optical flow field enables us to detect the dominant plane area on the image by using the matching of flow vectors on the image planes.

3.1 Planar flow estimation

First, using a pair of successive images from a sequence of images obtained during the camera motion, we compute the optical flow field $(\dot{u}, \dot{v})$. Since optical flow is the dense correspondence of points between an image pair, Eq.(1) can be applied to each point on the dominant plane. Let (u, v) and (u', v') be the points of the corresponding image coordinates in the two successive images. Point (u', v') can be calculated by adding the flow vector $(\dot{u}, \dot{v})$ to (u, v), if (u, v) and (u', v') are a corresponding points pair. Thus, by setting $(\dot{u}, \dot{v})$ to be the optical flow at the point (u, v) in the image coordinate system, we have the relation

$$\begin{pmatrix} u' \\ v' \end{pmatrix} = \begin{pmatrix} u \\ v \end{pmatrix} + \begin{pmatrix} \dot{u} \\ \dot{v} \end{pmatrix}. \tag{4}$$

If we obtain (u, v) and (u', v'), we can compute the affine coefficients in Eq.(1). If three points are non-collinear, Eq.(1) has a unique solution. Since the dominant plane occupies the largest domain in the image, we select three points randomly for computing the affine coefficients in the successive pair of images. After we compute the affine coefficients, using Eq.(1) again, we can compute the motion of the image sequence of the dominant plane, that is, the collection of the corresponding points

$$\begin{pmatrix} \hat{u} \\ \hat{v} \end{pmatrix} = \begin{pmatrix} u' \\ v' \end{pmatrix} - \begin{pmatrix} u \\ v \end{pmatrix} \tag{5}$$

in the image sequence can be regarded as the motion flow of the dominant plane on the basis of the camera motion. We call this flow *planar flow* $(\hat{u}, \hat{v})$.

3.2 Dominant plane detection

Next, we compute the dominant plane area using the estimated planar flow and the computed optical flow. Setting ε to be the tolerance of the difference between the optical flow vector and the planar flow vector, if

$$\left| \begin{pmatrix} \dot{u}_t \\ \dot{u}_t \end{pmatrix} - \begin{pmatrix} \hat{u}_t \\ \hat{u}_t \end{pmatrix} \right| < \varepsilon \tag{6}$$

is satisfied, we accept point (u_t, v_t) as the point in the dominant plane.

In the case that we select at least one point on the obstacle area in the image, the estimated planar flow is no longer the dominant plane motion. Therefore, the detected dominant plane area is very small. Since the dominant plane occupies the largest domain in the image, in such cases, it becomes evident that the selection of points is incorrect. In those cases, consequently, we select three points randomly again.

Once we have detected the dominant plane at a certain frame in the image sequence, the planar flow of subsequent images can be estimated robustly using the least-squares method, because dense optical flows are used for the estimation of affine coefficients. Assuming that the robot displacement is small, the dominant plane of the successive images changes negligibly. Therefore, using the optical flow on the estimated dominant plane in the previous image, we estimate the affine coefficients using the least-squares method as shown in Fig.4. Setting (u_i, v_i) and (u_i', v_i') $(0 \leq i \leq n)$ to be corresponding points, the mean-squared errors E_u and E_v associated with Eq.(1) are

$$E_u = \sum_{i=1}^{n} \{u_i' - (au_i + bv_i + c)\}^2, \tag{7}$$

$$E_v = \sum_{i=1}^{n} \{v_i' - (du_i + ev_i + f)\}^2, \tag{8}$$

where n is the number of points for estimation. Therefore, we can compute affine coefficients which minimize the errors E_u and E_v.

3.3 Procedure for the dominant plane detection

Our algorithm is summarized as follows.

1. Compute optical flow $(\dot{u}, \dot{v})$ from two successive images.

2. Compute affine coefficients in Eq.(1) by random selection of three points.

3. Estimate planar flow $(\hat{u}, \hat{v})$ from affine coefficients.

4. Matching the computed optical flow $(\dot{u}, \dot{v})$ and estimated planar flow $(\hat{u}, \hat{v})$ by using Eq.(6).

5. Detect the dominant plane. If the dominant plane occupies less than half of the image, then go to step(2).

Fig.5 shows the procedure of the dominant plane detection from the image sequence.

4 Experiment

In this section, we evaluate the performance of our method for steps(2) and (5), since the estimation of affine coefficients and the dominant plane detection are essential to the resolution of our problem. Step(2) estimates the parameters from the image sequence. Step(5) evaluates the area of the dominant plane in

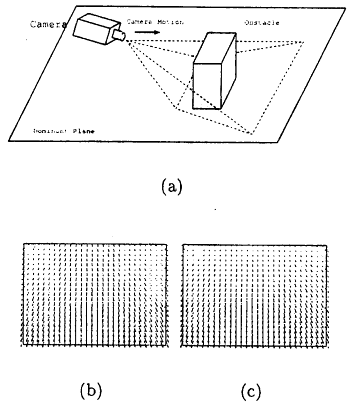

Figure 1: Planar flow of the image sequence. (a)Example of the camera displacement and the environment, (b)Computed optical flow, (c)Estimated planar flow.

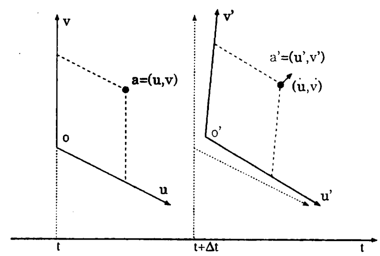

Figure 2: Optical flow generated from affine transformation. Optical flow $(\dot{u}, \dot{v})$ on the dominant plane is described as the affine transformation.

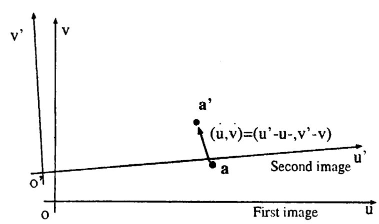

Figure 3: Optical flow computed from corresponding points in two successive images. Optical flow is described as the correspondence of the points of the two successive images. Therefore, the point (u', v') is $(u, v) - (\dot{u}, \dot{v})$ and planar flow $(\hat{u}, \hat{v})$ is $(u', v') - (u, v)$.

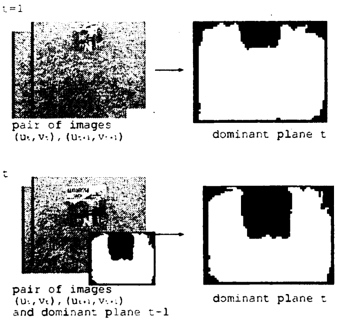

Figure 4: For the first frame, our algorithm uses a pair of images for the detection of the dominant plane. For subsequence frames, the dominant plane in the previous frame is used for the application of the least-squares method.

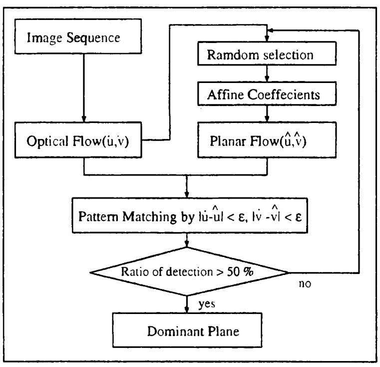

Figure 5: Procedure for dominant plane detection. (1)Compute optical flow $(\dot{u}, \dot{v})$ from two successive images. (2)Compute affine coefficients in Eq.(1) by random selection of three points. (3)Estimate planar flow $(\hat{u}, \hat{v})$ from affine coefficients. (4)Matching the computed optical flow $(\dot{u}, \dot{v})$ and estimated planar flow $(\hat{u}, \hat{v})$ by Eq.(6). (5)Detect the dominant plane. If the dominant plane occupies less than half of the image, then go to step(2).

the images. These experiments are presented in subsections 4.1 and 4.2, respectively. For the computation of optical flow, we use the Lucas-Kanade method with pyramids [15]. The tolerance for the matching of flow vectors in Eq.(6) is set to be $\varepsilon = 0.2$, which is fixed experimentally.

To evaluate of the accuracy of the dominant plane detection, we define the error ratio $E(t)$. Setting $D(u, v, t)$ and $M(u, v, t)$ to be the dominant plane on the image coordinate system (u, v) at the frame t detected from optical flow and detected manually, respectively, and setting

$$D(u, v, t) = \begin{cases} 0 & \text{for dominant plane} \\ 1 & \text{for the other area} \end{cases}, \tag{9}$$

$$M(u, v, t) = \begin{cases} 0 & \text{for dominant plane} \\ 1 & \text{for the other area} \end{cases} \tag{10}$$

we evaluate the ratio

$$E(t) = 100\frac{\sum_{u,v} |M(u, v, t) - D(u, v, t)|}{\sum_{u,v} 1} \tag{11}$$

as a function of time t.

4.1 The error rate of affine coefficients with uniform motion

First, we show the validity of the affine transformation in two successive images. From Eq.(1), we define the matrix $\mathbf{A}_t$ as

$$\mathbf{A}_t = \begin{pmatrix} a & b & c \\ d & e & f \\ 0 & 0 & 1 \end{pmatrix}, \tag{12}$$

where t is the frame number in the image sequence. Assuming that the motion of the camera is constant in two successive sequences of images, the relation $\mathbf{A}_{t+1} \simeq \mathbf{A}_t$ is satisfied. Therefore, we define the error of affine coefficients as

$$e(t) = |\mathbf{A}_{t+1} - \mathbf{A}_t|^2. \tag{13}$$

In this experiment, we use the Marbled-Block image sequence [16], since the velocity of the camera displacement should be constant. Figure 6 shows the error of affine coefficients. Figure 7 and Fig.8 show the image sequence of detected dominant plane and the error ratio of the dominant plane, respectively.

Figure 6 shows that $|\mathbf{A}_{t+1} - \mathbf{A}_t|^2$ approaches zero with time. Therefore, the affine coefficients are constant when the camera displacement is constant. We show the validity of the affine transformation in two successive images.

4.2 Detection of the dominant plane

Second, we show the validity of the random selection of three points for the computation of affine coefficients.

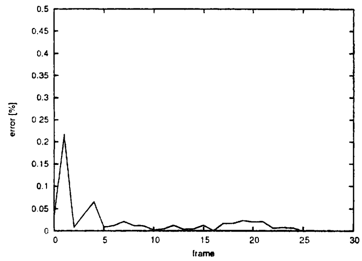

Figure 6: Error of affine coefficients. The vertical axis is the error of affine coefficients e(t) as given in Eq.(13). The horizontal axis is the frame number of the image sequence.

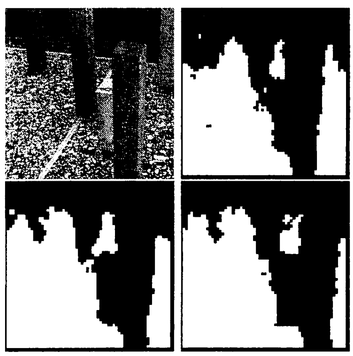

Figure 7: Detected dominant plane in the 0, 10 and 20 frames. From left to right, the original image, and detected dominant plane in the 0, 10 and 20 frames. In the images of the dominant plane, the white areas are the dominant planes and the black areas are the obstacle areas.

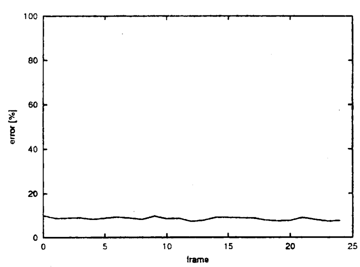

Figure 8: Error ratio of the dominant plane detection using the image sequence shown in Fig.7. The vertical axis is the error ratio E(t) in Eq.(11). The horizontal axis is the frame number of the image sequence.

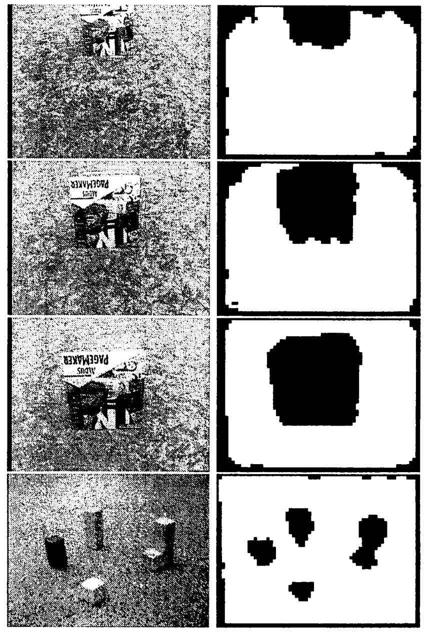

Figure 9: Detected dominant plane. The first colomn is the original image, and the second colomn is shown for the dominant plane.

In this experiment, we use image sequences of forward motion and rotational motion, because the motion of a mobile robot mainly consists of these motion. For the image sequence of forward motion, we prepared image sequences in which the percentages of the dominant plane are 90%, 80% and 70%, since we assume that the dominant plane occupies more than half of an image. These percentages of the dominant plane are computed using $M(u, v, t)$ in Eq.(10). For the image sequence of rotational motion, we use the Rotating Blocks image sequence [17].

The result of dominant plane detection is shown in Fig.9. The first colomn is the original image, and the second colomn is shown for the dominant plane. In the images of the dominant plane, the white areas are the dominant planes and the black areas are the obstacle areas. Table 1 lists error ratios evaluated with Eq.(11).

Figure 9 and Tab.1 show that the correctness of regions detected as the dominant plane increases with time.

4.3 Dominant plane detection in the long sequence

Finally, we present the dominant plane detection in a long image sequence for the navigation and path plan-

Table 1: Error ratios of the dominant plane detection in Fig.9. For the second row, F90, F80 and F70 are the image sequences of forward motion in which the percentages of the dominant plane are 90%, 80% and 70%, respectively. Rotation gives the image sequence of rotational motion. This image sequence is shown in the first row of Fig.9.

	error ratio E(t) [%]			
frame	F90	F80	F70	Rotation
0	23.22	24.01	20.33	19.37
1	14.40	12.99	12.29	14.50
2	10.91	11.83	10.45	11.09
3	10.54	11.08	10.55	11.29

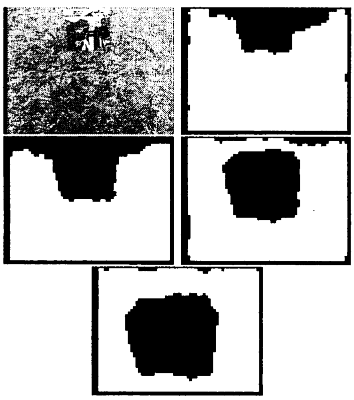

Figure 10: Detected dominant plane at the 50, 100, 150 and 200 frames.

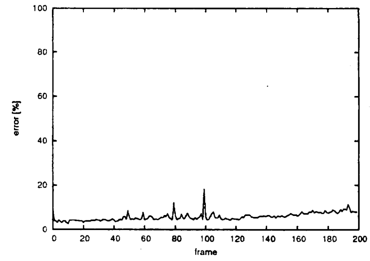

Figure 11: Error ratio E(t). The vertical axis is the error ratio E(t) in Eq.(11). The horizontal axis is the frame number of the image sequence.

ning of the mobile robot. Figure 10 and Fig.11 show the results for the dominant plane detection and the error ratio, respectively.

As evident from Fig.11, even if the error ratio increases in a sequence, the error ratio becomes small with a subsequent frame. Therefore, our method detects the dominant plane without accumulation of the error ratio. Our algorithm runs on a 2.0 GHz Pentium4 for 320×240 images and takes 0.25 seconds per frame for the detection of the dominant plane.

5 Conclusion

We developed an algorithm for the estimation of the optical flow field of the dominant plane detection from a sequence of images observed through a moving uncalibrated camera. Furthermore, we showed that corresponding points on dominant planes in a pair of successive images are combined with an affine transformation. If we compute the affine coefficients, we obtain the dense planar flow easily from the predetected images. Using this principle, if we compute the affine coefficients which relate the corresponding points in two successive images, we can obtain the dense planar flow easily from the predetected images. This property of the points in a dominant plane allows us to design an algorithm which detects the dominant plane by the simple pattern matching of the flow vectors in a series of dominant planes.

Although there is a model-based approach in dominant plane detection proposed in [12], our method is a non-model-based approach which uses matching with the predictive image. In addition, our algorithm detects the dominant plane without requiring camera calibration, since our algorithm uses the affine approximation on the correspondences between points in a pair of images. Results of experiments using real and artificial image sequences show that the dominant plane can be detected robustly. These experimental results support the application of our method to the navigation and path planning of a mobile robot with a vision system.

References

[1] Hahnel, D., Triebel, R., Burgard, W., and Thrun, S., Map building with mobile robots in dynamic environments, In Proc. of the IEEE ICRA03, (2003).

[2] Huang, T. S. and Netravali, A. N., Motion and structure from feature correspondences: A review, Proc. of the IEEE, **82**, 252-268, (1994.

[3] Guilherme, N. D. and Avinash, C. K., Vision for mobile robot navigation: A survey IEEE Trans. on PAMI, **24**, 237-267, (2002).

[4] Kang, S.B. and Szeliski, R., 3D environment modeling from multiple cylindrical panoramic images, *Panoramic Vision: Sensors, Theory, Applica-*

tions, 329-358, Ryad Benosman and Sing Bing Kang, ed., Springer-Verlag, (2001).

[5] Fraundorfer, F., A map for mobile robots consisting of a 3D model with augmented salient image features, 26th Workshop of the Austrian Association for Pattern Recognition, 249-256, (2002).

[6] Barron, J.L., Fleet, D.J., and Beauchemin, S.S., Performance of optical flow techniques, International Journal of Computer Vision, **12**, 43-77, (1994).

[7] Horn, B. K. P. and Schunck, B.G., Determining optical flow, Artificial Intelligence, **17**, 185-203, (1981).

[8] Lucas, B. and Kanade, T., An iterative image registration technique with an application to stereo vision, Proc. of 7th IJCAI, 674-679, (1981).

[9] Mallot, H. A., Bulthoff, H. H., Little, J. J., and Bohrer, S., Inverse perspective mapping simplifies optical flow computation and obstacle detection, Biological Cybernetics, **64**, 177-185, (1991).

[10] Haag, M. and Nagel, H.-H., Beginning a transition from a local to a more global point of view in model-based vehicle tracking, ECCV98, **1**, 812-827, (1998).

[11] Wachter, S. and Nagel, H.-H., Tracking persons in monocular image sequences, Computer Vision and Image Understanding, **74**, 174-192, (1999).

[12] Enkelmann, W., Obstacle detection by evaluation of optical flow fields from image sequences, Image and Vision Computing, **9**, 160-168, (1991).

[13] Santos-Victor, J. and Sandini, G., Uncalibrated obstacle detection using normal flow, Machine Vision and Applications, **9**, 130-137, (1996).

[14] Hartley, A. and Zisserman, A., *Multiple View Geometry in Computer Vision*, Cambridge University Press, (2000).

[15] Bouguet J.-Y., Pyramidal implementation of the Lucas Kanade feature tracker description of the algorithm, Intel Corporation, Microprocessor Research Labs, OpenCV Documents, (1999).

[16] Marbled-Block sequence: recorded and first evaluated by Otte and Nagel (KOGS/IAKS, University of Karlsruhe),
website: http://i21www.ira.uka.de/image_sequences/

[17] Rotating Blocks: Otago Optical Flow Evaluation Sequences
website: http://www.cs.otago.ac.nz/research/vision/Research/