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

2005－CVIM－148（17）
2005／3／4

ドミナントプレーンによる軌道制御

大西 直哉 † 井宮 淳 ††

† 千葉大学　自然科学研究科
†† 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

あらまし　本論文では、ドミナントプレーンを用いたロボットの軌道設計法を提案する。ドミナントプレーンとは画面の 50%以上を占める平面と定義し、運動するカメラによって得られるオプティカルフローから、ドミナントプレーンを検出することができる。ドミナントプレーンに属するピクセルの動きは、運動同相性を保っている。したがって、ドミナントプレーンの上では、オプティカルフロー抽出における、開口問題を避けることができる。さらに、本論文において提案する軌道制御アルゴリズムは、3 枚の連続する画像を利用するため数値的に安定であることを示す。

Dominant plane derives dominant motion

Naoya OHNISHI† and Atsushi IMIYA††

†School of Science and Technology, Chiba University
††Institute of Media and Information Technology, Chiba University
Yayoi-cho 1-33, Inage-ku, 263-8522, Chiba, Japan

Abstract　We propose a strategy for the detection of dominant motion in the scene. The dominant motion is the motion of pixels on the dominant plane, which occupies more than 50% in the scene. Therefore, the dominant motion is not affected by the aperture problem which is a critical problem in optical-flow-based motion detection from sequence of images. We apply dominant motion of scene to the ego-motion estimation of autonomous robot.

1 Introduction

In this work, we aim to develop an algorithm for robot motion planning using the dominant plane. 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 the dominant plane (for examples, floors and ground areas), the use of the dominant plane is an appropriate method for the autonomous navigation and the path planning of the mobile robot.

The autonomous navigation of mobile robots desires a simple mechanism and light devices, since mobile robots are restricted their payload (for examples, a power supply, a capacity of input devices and a computing power). 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 [1]. Additionally, vision sensors are fundamental devices for the understanding of the environment, since the robot collaborates with a human being. Furthermore, 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 [2][3][4]. For example, the edge detection of omni and monocular camera systems [5] and the observation of landmarks [6] are the classical ones. However, since these methods are dependent on the environment around a robot, these methods are difficult to apply in general environments. If a robot captures an image sequence of moving objects, the optical flow [7] [8] [9], 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 [10]. Therefore, the use of optical flow is an appropriate method from the viewpoint of the affinity for the collaboration between the robot and human being.

Haag and Nagel[11] 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 [12]. We develop the detection of static 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 static 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 [13] [14]. Enkelmann [13] proposed an obstacle-detection method using the model vectors from motion parameters. Santos-Victor and Sandini [14] 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.

Optical-flow-based motion segmentation has long history in computer vision [4]. Most fundamental segmentation is extraction of moving objects, since on the pixels of the moving objects and the background, flow vectors are not zero and zero, respectively. This method also allows us to separate object, whose depths from a camera are finite, from a scene, whose depth from a camera is infinite. Furthermore, optical flow in an image sequence observed from moving camera, for example a camera mounted on a mobile robot is used for the detection of obstacles in front of camera. In this paper, we extend these ideas on the segmentation of image sequence using optical flow to the detection of dominant plane in an image sequence. For the detection of dominant plane, we are not required to deal with the aperture problem for the optical flow on the boundary of moving object, as shown in Fig.1, since the dominant plane occupies the largest region in an image and the optical flow vector on the dominant plane can be computed stable.

2 Preliminary

2.1 Estimation of the optical flow

We detect dominant plane using optical flow, that is, for

$$\boldsymbol{u}^\top \nabla f(x,y,t) + f_t = 0, \quad \boldsymbol{u} = \left(\frac{dx}{dt}, \frac{dy}{dt}\right)^\top \tag{1}$$

$\boldsymbol{u}$ is constant at each pixels in a windowed region. Therefore, our optical flow vectors is the solution

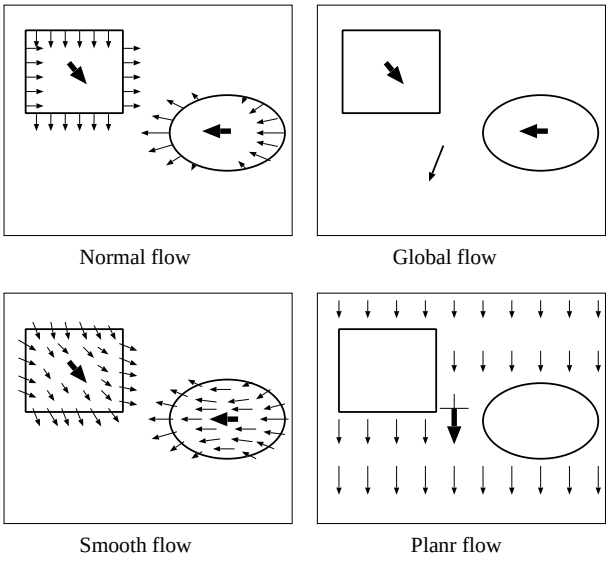

Figure 1: Normal flow has a aperture problem. Global flow dose not apply when many objects exist. Smooth flow dose not preserve motion discontinuities. Our planar flow on a dominant plane has no such problem since the dominant plane occupies the largest domain in the image observed by a moving camera.

which minimise the condition

$$J_{LK}(w) = \int\int_{\Omega(\boldsymbol{x})} w^2(\boldsymbol{x}) |\boldsymbol{u}^\top \nabla f(x,y,t) + f_t|^2 \mathrm{d}\boldsymbol{x} \tag{2}$$

where $\Omega(\boldsymbol{x})$ is the finite windowed area continued at $\boldsymbol{x}$, and w is the weight function. For the details of the derivation of the optical flow using the Lucas-Kanade method with pyramids, see Appendix A.

2.2 Definition of the dominant plane

As stated in Section 1, the dominant plane has the following properties

1. The plane occupies the largest domain in the image.
2. The distance from a camera to the plane is finite in a space.

The examples of a application of the definition of (1) and (2) are shown in Fig. 2 and Fig. 3, respectively. In Fig. 2, a obstacle is placed in front of a camera on the right and the camera moves toward a obstacle. Therefore, the top-right area of the image is not detected as the dominant plane.

In Fig. 3, the front of the camera is edge of the dominant plane. and the camera moves forward. Therefore, the top area of the image is not detected as the dominant plane.

Furthermore, our geometric assumptions are that

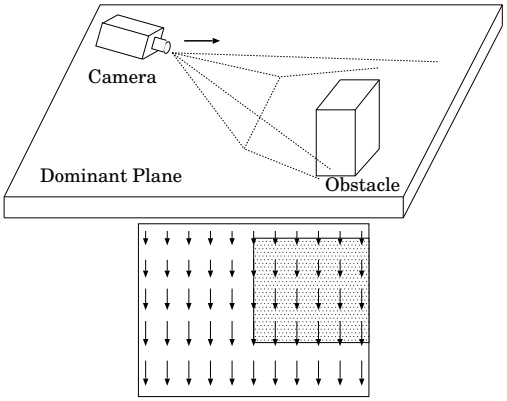

Figure 2: left: Camera moves toward a obstacle. right: Dominant plane of the captured image. The white area is the dominant plane. Since the obstacle is placed in front of the camera on the right, this area is not dominant plane.

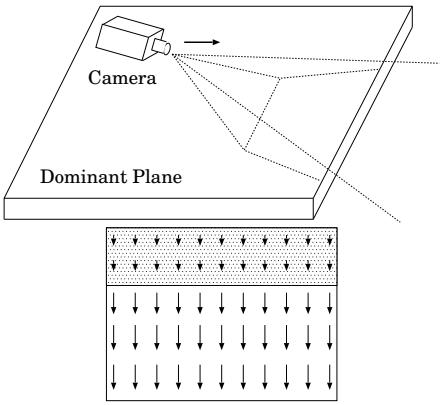

Figure 3: left: Camera moves on the edge of the dominant plane. right: Dominant plane of the captured image. Since the depth of the top area of the image is infinite, this area is not dominant plane.

1. Images are observed from a camera mounted on a mobile robot moving on a plane perpendicular to the axes of the gravity.

2. So, the plane in an image is horizontal.

3. The dominant plane is a finite plane in an environment on which a robot moves.

4. The dominant plane is a region more than 50% in an image in a sequence of images.

These definitions and the assumptions enable the robot to navigate autonomously in a environment with obstacles.

2.3 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

$$\begin{pmatrix} u' \\ v' \end{pmatrix} = \begin{pmatrix} a & b \\ d & e \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} + \begin{pmatrix} c \\ f \end{pmatrix}, \tag{3}$$

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 $\boldsymbol{H}$ to be the 3×3 matrix [15], the homography between the two images of a planar surface is expressed as

$$\boldsymbol{p} = \boldsymbol{H}\boldsymbol{p}', \tag{4}$$

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

$$\boldsymbol{H} = \boldsymbol{K}(\boldsymbol{R} + \boldsymbol{t}\boldsymbol{n}^\top)\boldsymbol{K}^{-1}, \tag{5}$$

where $\boldsymbol{K}$, $\boldsymbol{R}$, $\boldsymbol{t}$, and $\boldsymbol{n}$ are the projection matrix, the rotation matrix, the translation vector, and the plane normal of the planar surface, respectively. The matrices $\boldsymbol{K}$ and $\boldsymbol{K}^{-1}$ are affine transformations since those matrices are the projection matrix of the pinhole camera. Assuming that the camera displacement is small, the matrix $\boldsymbol{R}$ and the matrix $\boldsymbol{t}\boldsymbol{n}^\top$ 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. For the details of the derivation of the homography using our coordinate system, see Appendix B. 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.(4) is expressed as Eq.(3). In the next section, we develop an algorithm for the estimation of these six parameters from a sequence of images.

2.4 Planar flow estimation

Using a pair of successive images from a sequence of images obtained from the camera during robot motion, we compute the optical flow field $(\dot{u}, \dot{v})$. Since optical flow is the correspondence of dense points between an image pair, Eq.(3) 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

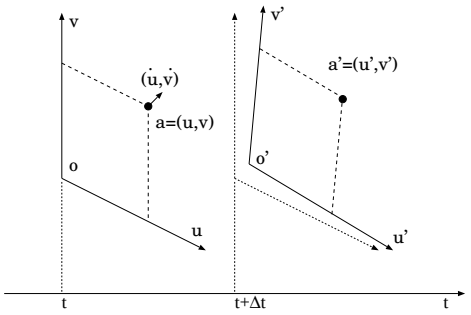

Figure 4: Optical flow generated by affine transformation. Optical flow $(\dot{u}, \dot{v})$ on the dominant plane is described as an affine transformation.

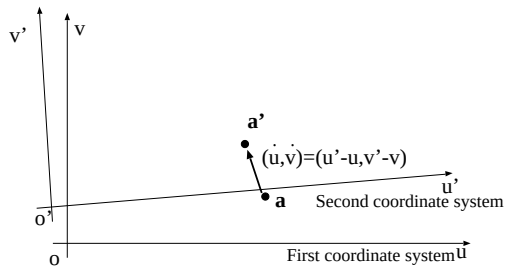

Figure 5: Optical flow computed from corresponding points in two successive images. Optical flow is described as the correspondence of the same points in the two successive images. Therefore, point (u', v') is $(u + \dot{u}, v + \dot{v})$ and planar flow $(\hat{u}, \hat{v})$ is $(u' - u, v' - v)$.

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 pair of corresponding points. Thus, by setting $(\dot{u}, \dot{v})$ to be the optical flow at point (u, v) in the image coordinate system, as shown in Figure 4, we have the relation

$$\begin{pmatrix} u' \\ v' \end{pmatrix} = \begin{pmatrix} u \\ v \end{pmatrix} + \begin{pmatrix} \dot{u} \\ \dot{v} \end{pmatrix}. \tag{6}$$

If we obtain (u, v) and (u', v'), we can compute the affine coefficients in Eq.(3). If three points are non-collinear, Eq.(3) has a unique solution. Since the dominant plane occupies the largest domain in the image, we select three points randomly to compute the affine coefficients in the successive pair of images. After we compute the affine coefficients, using Eq.(3) again, we can compute the motion of the image sequence of the dominant plane, that is, the collection of corresponding points

$$\begin{pmatrix} \hat{u} \\ \hat{v} \end{pmatrix} = \begin{pmatrix} u' \\ v' \end{pmatrix} - \begin{pmatrix} u \\ v \end{pmatrix} \tag{7}$$

in the image sequence can be regarded as motion flow of the dominant plane on the basis of camera displacement as shown in Figure 5. We call this flow *planar flow* $(\hat{u}, \hat{v})$. This planar flow is used as temporal-model for dominant plane detection.

2.5 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

$$\left| \begin{pmatrix} \dot{u}_t \\ \dot{u}_t \end{pmatrix} - \begin{pmatrix} \hat{u}_t \\ \hat{u}_t \end{pmatrix} \right| < \varepsilon \tag{8}$$

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

In the case that at least one point on the obstacle area in the image is selected, 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, we consequently select another three points randomly. Figure 6 shows examples of each case.

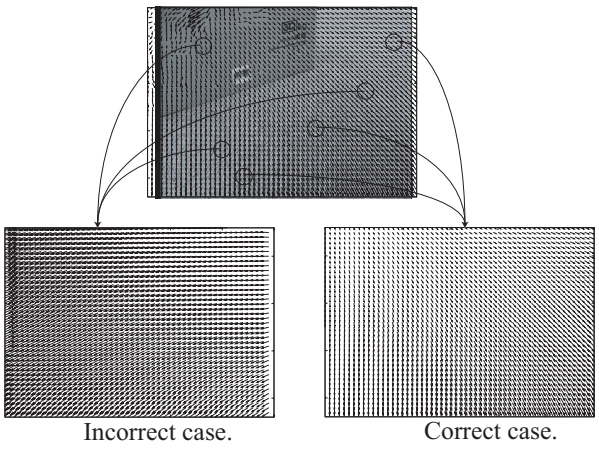

Figure 6: Examples of random sampling. Bottom-left is incorrect case, since the point is selected on the obstacle area. Select another points randomly. Bottom-right is correct case. This planar flow is equal to optical flow in 50% or more of area.

For the estimation of the affine parameters of the affine transformation. We are required to detect at least pair of three points in the regions which are corresponded by the affine transform. Therefore, if we select a correct pair of three points, we can compute correct affine parameters. However, if we select miss matched pairs, we can not compute the correct affine parameters. This geometrical property derives

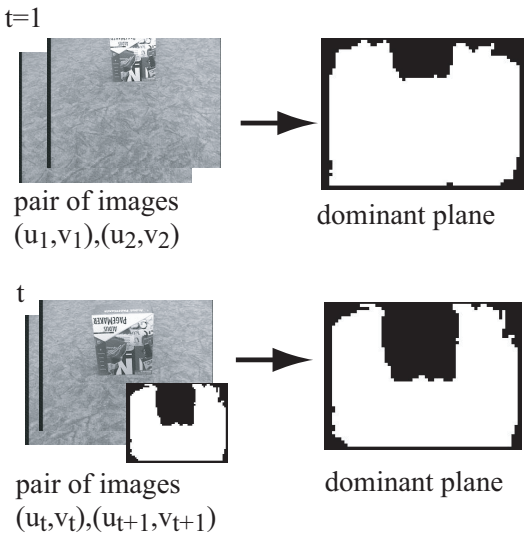

Figure 7: For the first frame, our algorithm uses a pair of images for the detection of the dominant plane. For subsequence frames, the dominant plane in the previous frame is used for the application of the least-squares method.

a voting-based algorithm [16][17] for the computation of the affine parameters.

Once we have detected the dominant plane in a certain frame of 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 Figure 7. 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.(3) are

$$E_u = \sum_{i=1}^{n} \{u'_i - (au_i + bv_i + c)\}^2, \tag{9}$$

$$E_v = \sum_{i=1}^{n} \{v'_i - (du_i + ev_i + f)\}^2, \tag{10}$$

where n is the number of points used for estimation. Therefore, we can compute affine coefficients which minimize errors E_u and E_v. This algorithm generates planar flow as temporal-model at subsequent frames

2.6 Procedure for dominant plane detection

Our algorithm is summarized as follows.

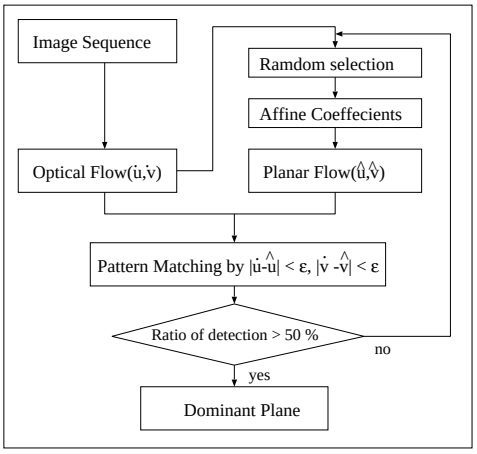

Figure 8: Procedure for dominant plane detection. (1) Compute optical flow $(\dot{u}, \dot{v})$ from two successive images. (2) Compute affine coefficients in Eq.(3) by random selection of three points. (3) Estimate planar flow $(\hat{u}, \hat{v})$ from affine coefficients. (4) Match the computed optical flow $(\dot{u}, \dot{v})$ and estimated planar flow $(\hat{u}, \hat{v})$ using Eq.(8). (5) Detect the dominant plane. If the dominant plane occupies less than half of the image area, then return to step(2).

1. Compute optical flow $(\dot{u}, \dot{v})$ from two successive images.
2. Compute affine coefficients in Eq.(3) by random selection of three points.
3. Estimate planar flow $(\hat{u}, \hat{v})$ from affine coefficients.
4. Match the computed optical flow $(\dot{u}, \dot{v})$ and estimated planar flow $(\hat{u}, \hat{v})$ using Eq.(8).
5. Detect the dominant plane. If the dominant plane occupies less than half of the image area, then return to step(2).

Figure 8 shows the procedure of dominant plane detection from the image sequence.

2.7 Navigation using the dominant plane

We describe an algorithm for navigation of a mobile robot using the dominant plane. The percentages of the dominant plane in the image determine for the robot a strategy of a robot motion. If the percentage of the dominant plane in the image is greater than 80% and the dominant plane of the left side of the image is greater than that of right side, the robot moves forward. If the percentage of the dominant plane in the image is less than 80% and the dominant plane of the left side of the image is greater than that of right side, the robot rotates counterclockwise to avoid collision of

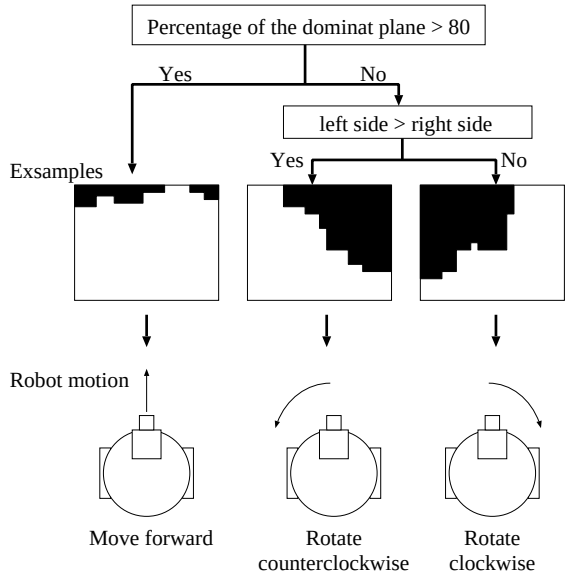

Figure 9: Strategy of a mobile robot. Left: Robot moves forward. Middle: Robot rotates counterclockwise. Right: Robot rotates clockwise.

obstacles. If the percentage of the dominant plane in the image is less than 80% and the dominant plane of the right side of the image is greater than that of left side, the robot rotates clockwise. These algorithm are summarized in Figure 9.

3 Experiment

3.1 Dominant plane detection

We experiment for the dominant plane detection from optical flow. The camera moves forward to the edge of the table and captures the image sequence, as shown in the first row of Fig. 10. The obstacle is on the right of the table. For the computation of optical flow, we use the Lucas-Kanade method with pyramids [18]. The result of the experiment shows Fig. 10. In Fig. 10, the right column shows the detected dominant plane. The white and black areas are shown the dominant plane and the obstacle area, respectively. The depths of pixels in background are infinite. Therefore, it is impossible to detect flow vectors on these pixels. Our algorithm detects optical flow vectors for pixels where depths are finite. The result shows that the dominant plane can be detected accuracy as the definition of the dominant plane in Section 2.2.

3.2 Moving obstacles detection

Next, we experiment for the detection of moving obstacles. The robot moves forward, and captures the image sequence. The obstacle is in front of the robot,

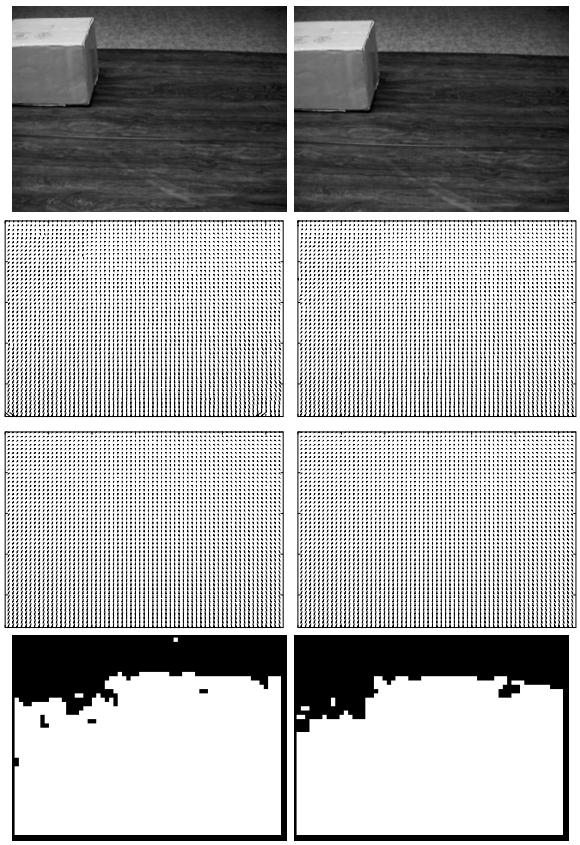

Figure 10: The first, second, third, and forth rows show the original image, the optical flow, planar flow, and the dominant plane, respectively. The left and right columns show the first and second image of the sequence, respectively.

and moves to left, right, back, and front of the robot. Figures 11, 13, 15 and 17 show the experiment result in each case. These results show that the algorithm detects moving objects in the environment around a robot.

3.3 Navigation of the mobile robot

Figure 18 shows the image sequence and the dominant plane observed form the robot. In these sequence, there is a reflective rail in the center of images. In a middle sequences, robot detects obstacles in the central region of images. However, since these parts are small, the robot moves forward considering these region as noise.

4 Conclusion

We developed an algorithm for planning of the robot motion using the dominant plane, which detected from a sequence of images observed through a moving uncalibrated camera. Since our algorithm uses simple

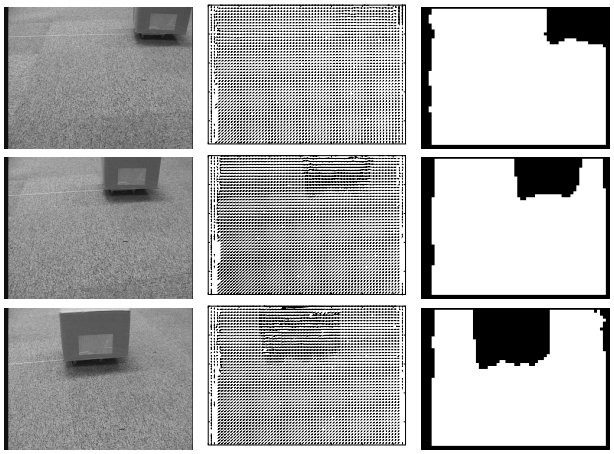

Figure 11:

Figure 12: The obstacle moves left. The first, second, and third columns show the original image, the optical flow, and the dominant plane, respectively. The sequences are ordered in time from the top row.

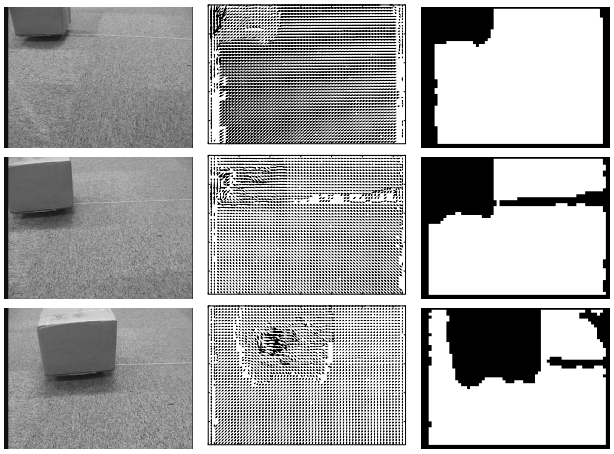

Figure 13:

Figure 14: The obstacle moves right. The first, second, and third columns show the original image, the optical flow, and the dominant plane, respectively. The sequences are ordered in time from the top row.

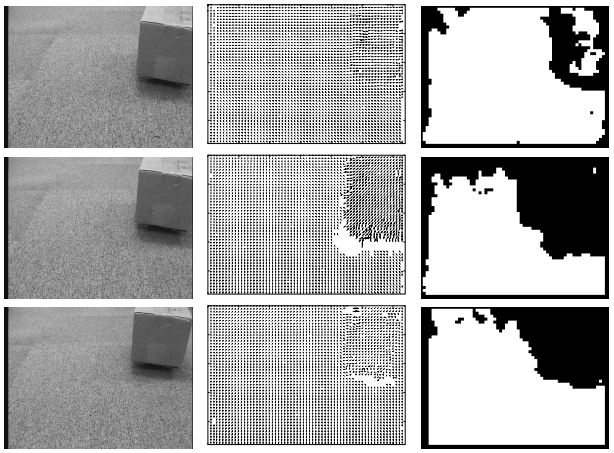

Figure 15:

Figure 16: The obstacle moves front. The first, second, and third columns show the original image, the optical flow, and the dominant plane, respectively. The sequences are ordered in time from the top row.

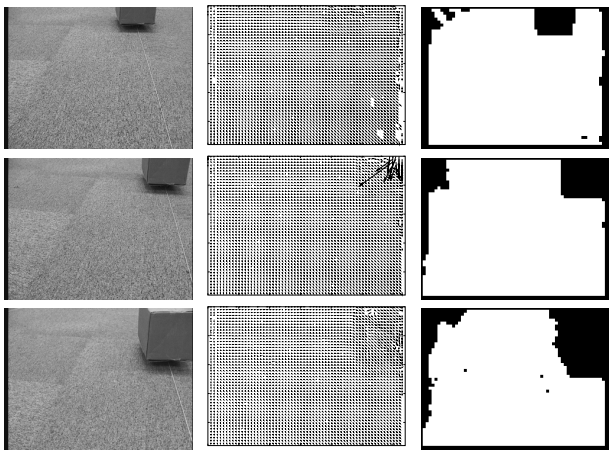

Figure 17: The obstacle moves back. The first, second, and third columns show the original image, the optical flow, and the dominant plane, respectively. The sequences are ordered in time from the top row.

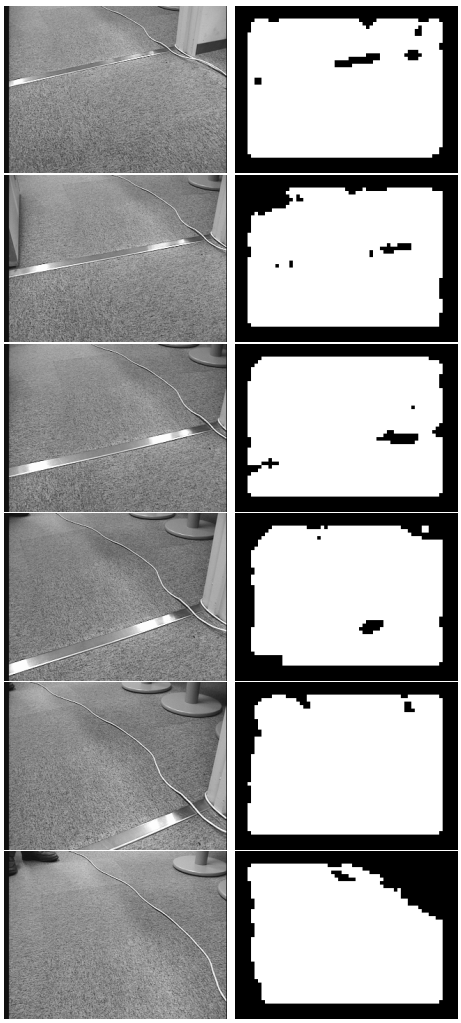

Figure 18: Left and right columns show the image sequence and the dominant plane, respectively

method and input devise, a robot dose not demand a large payload.

References

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

[2] Baker, S., and Matthews, I., Lucas-Kanade 20 years on: a unifying framework, International Journal of Computer Vision, **56**, 221-255, (2004).

[3] Nagel, H.-H., Image sequence evaluation: 30 years and still going strong, Proc. of the ICPR'00, 149-158, (2000).

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

[5] Kang, S.B. and Szeliski, R., 3D environment modeling from multiple cylindrical panoramic images, *Panoramic Vision: Sensors, Theory, Applications*, 329-358, Ryad Benosman and Sing Bing Kang, ed., Springer-Verlag, (2001).

[6] 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).

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

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

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

[10] 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).

[11] 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).

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

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

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

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

[16] Fermin, I. and Imiya, A., Planar motion detection by randomized triangle matching, Pattern Recognition Letters, **18**, 741-749, (1997).

[17] Fermin, I., Imiya, A., and Ichikawa, A., Randomized polygon search for planar motion detection, Pattern Recognition Letters, **17**, 1109-1115, (1996).

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