

## NAVIGATION AND GRAVITATION

S.S. BEAUCHEMIN

*The University of Western Ontario  
London, Canada*

M.T. KOTB

*The University of Western Ontario  
London, Canada*

H.O. HAMSHARI

*The University of Western Ontario  
London, Canada*

**Abstract.** We propose a mathematical model for vision-based autonomous navigation on general terrains. Our model, a generalization of Mallot's inverse perspective, assumes the direction of gravity and the speed of the mobile agent to be known, thus combining visual, inertial, and navigational information into a coherent scheme. Experiments showing the viability of this approach are presented and a sensitivity analysis with random, zero-mean Gaussian noise is provided.

**Key words:** autonomous navigation, optical flow, perspective mapping, inverse perspective mapping

### Introduction

Needless to say, there is a growing interest in the field of vision-based autonomous navigation, partly due to its important applications in natural and man-made environments (Batavia et al., 2002; Baten et al., 1998; Choi et al., 1999; Desouza and Kak, 2002; Tang and Yuta, 2001; Tang and Yuta, 2002; Wijesoma et al., 2002).

The complexity of the navigation problem increases with that of the terrain and the environment in general. Navigation over rough terrains requires a vision system to react correctly with respect to the conditions posed by navigational surfaces with significant irregularities.

In general, perception systems rely on sensors such as sonar, lasers, or range finders. In addition, their outputs may be fused to increase the

reliability of the perception process. Environmental data captured and fused in this way may then be used for essential navigation tasks such as relative position and egomotion estimation (Jin et al., 2003; Kim and Kim, 2003), obstacle detection and avoidance (Ku and Tsai, 1999), and path planning (Desouza and Kak, 2002).

Relative position may be estimated through spatial relations to external objects such as land-marks or through incremental movement estimation using odometry and gyroscopes. Path planning methods depend on many factors, such as the complexity of the navigational tasks, the level of knowledge about the environment, and the dimension of the navigational problem. For instance, in the 1-dimensional case, the navigation is kept with a fixed distance to a reference line. In 2 and 3-dimensional navigation, landmarks are commonly used for estimating the current position and performing local path planning. Topological and geometrical relations between environmental elements and features are represented by various spatial maps that are established prior to the navigational task. For instance, landmark maps hold the information about the position of landmarks on the terrain, whereas passability maps represent the traversable paths on the terrain, or the location of obstacles in the environment. With the knowledge of the relative position of the moving sensor and the information held in the landmark and passability maps, navigating becomes a relatively trivial task. However, solving this problem with potentially unreliable information about the environment or the location of obstacles is very challenging.

The use of motion information for navigational purposes such as optical flow poses significant problems in general. The difficulty of obtaining numerically accurate and meaningful optical flow measurements has been known for some time (Barron et al., 1994). For this and other reasons, if one could impose additional constraints onto the spatiotemporal structure of optical flow, one could most probably obtain better flow estimation. For instance, Mallot's inverse perspective mapping model eliminates optical flow divergence, provided that the navigational path on the terrain remains perfectly flat. As a result, when the sensing agent moves in a straight line, the optical flow estimates are then isotropically parallel and their magnitudes describe the corresponding terrain heights.

In this contribution, we propose a generalization of this model for uneven terrains, modeled as triangulations of randomly generated height points. As we demonstrate, it is possible to maintain a correct optical flow pattern in spite of the motion experienced by the visual sensor while navigating on an uneven terrain. We also provide noise analysis to the reconstructed 3d world model. Our proposed model will be used without being provided with landmark or passability maps. Ultimately, the mobile agent is required to make real-time navigational decisions, using the perceived information

from the scene. Incremental movement estimation using odometers and gyroscopes will be used for relative position estimation and the determination of the direction of the gravity field.

This contribution is organized as follows: section 1 defines the coordinate systems involved, section 2 is a synopsis of Mallot's perspective and inverse perspective mapping model, section 4 outlines the problems encountered while applying this model on uneven terrains, section 5 is a description of our proposed perspective and inverse perspective mathematical models, and section 6 presents a noise sensitivity analysis for our proposed model.

## 1. Coordinate Systems

The projection of a 3d world point onto the image plane involves three coordinate systems and two transformations. The world coordinate system  $W$  is described with 3 primary axes,  $X$ ,  $Y$ , and  $Z$ . A point in the 3d world is denoted by  $P_w$  and the coordinates of this point are  $(P_{w_x}, P_{w_y}, P_{w_z})$ . The point  $P_w$  is transformed from the world coordinate system  $W$  into the camera coordinate system  $C$  giving a point  $P_c = (P_{c_x}, P_{c_y}, P_{c_z})$ , where  $c_x$ ,  $c_y$ , and  $c_z$  are the sensor axes defined in the world coordinate system. The point  $P_c$  is projected into the image plane giving a corresponding point  $P_i$ . This point is described in an image plane coordinate system  $I(a, b)$ .

## 2. Mallot's Model

Mallot's model presents an inverse perspective scheme for navigation on flat terrains. It is a bird's eye model where the imagery recorded by the visual sensor undergoes a mathematical transformation such that the sensor's gaze axis becomes perpendicular to the navigational surface. This transformation effectively nulls the perspective effects within the resulting optical flow and allows for a simple procedure to estimate obstacle locations.

### 2.1. PERSPECTIVE MAPPING

Perspective mapping or projection may be written in the following way:

$$\begin{pmatrix} P_{I_a} \\ P_{I_b} \\ -f \end{pmatrix} = \frac{-f}{P_{c_z}} \begin{pmatrix} P_{c_x} \\ P_{c_y} \\ P_{c_z} \end{pmatrix} \quad (1)$$

where  $f$  is the focal length. Figure 1 shows the world map of a triangulated flat terrain captured by a perspective visual sensor.

Figure 2a shows the perspective mapping image from the visual sensor specified in Equation (1), moving along the diagonal of the terrain. The

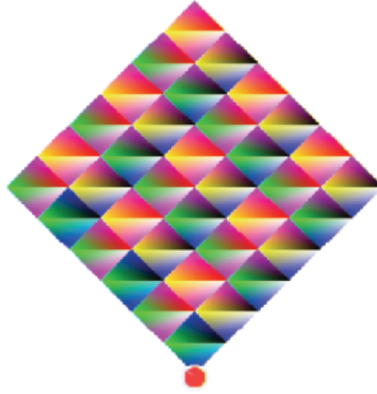


Figure 1. The map of a flat terrain, where the circle represents the position of the mobile agent.

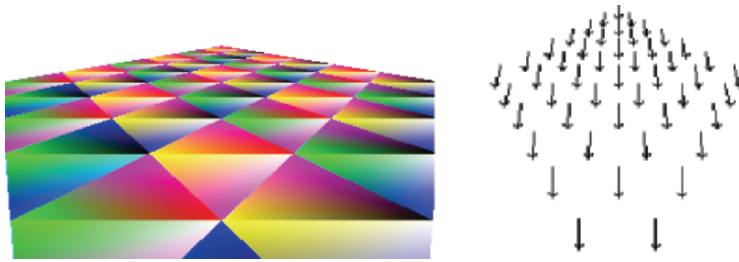


Figure 2. **a) Left:** Perspective mapping from the visual sensor, moving along the diagonal of the terrain. **b) Right:** The optical flow of the perspective mapping in **a)**.

terrain in the image is a square surface, as shown in Figure 1. The obvious perspective effects resulting from projection are noted in Figure 2a. Figure 2b shows the optical flow of the perspective mapping from Figure 2a as the visual sensor moves along a straight line on the terrain. It can be easily seen from Figure 2b that the Focus Of Expansion (FOE) is located at the horizon. In addition, from Equation (1) and Figure 2b, it can be understood that perspective effects are in direct relation with, among other things, the relative height of the visual sensor from the navigational surface.

Applying the transform  $T$  from Mallot's model results in the correct, perspective free, optical flow vector field. The transformation  $T$  is shown in Figure 3, where  $P$  is the point which the camera looks at. Figures 4a and 4b show the perspective mapping and optical flow respectively, after applying the transformation  $T$  onto the sensor imagery.

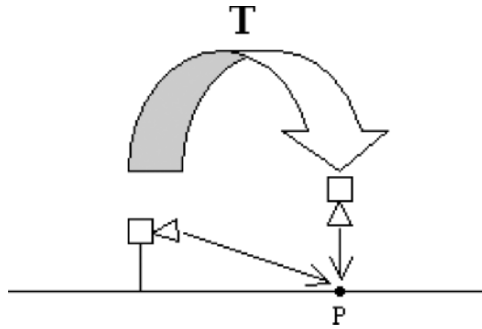


Figure 3. Camera transformation  $T$ .

Since inverse-perspective optical flow is a function of depth, then different terrain heights have different optical flow vector magnitudes. Figure 5a shows a global map that has some spikes in the middle of the terrain. Figures 5b and 5c show the perspective mapping and the optical flow respectively. As shown in Figure 5c, the optical flow vectors that represent the motion of the spikes with respect to the camera are longer than those which represent the flat part of the terrain.

## 2.2. INVERSE PERSPECTIVE MAPPING

Equation (2) presents the inverse perspective mapping as per Mallot's model. This mapping gives a point  $P_w$  which corresponds to a point  $P_i$  in the image plane. The inverse perspective mapping involves two transformations, one from the image plane coordinate system to the camera coordinate system, and a second transform from the camera coordinate system to the world coordinate system:

$$\begin{pmatrix} W_x \\ W_y \end{pmatrix} = \beta \cdot \gamma \tag{2}$$

where

$$\beta = \frac{-h}{N_x P_{I_a} + N_y P_{I_b} - N_z f}$$

and

$$\gamma = \begin{pmatrix} U_x P_{i_a} + U_y P_{i_b} - U_z f \\ V_x P_{i_a} + V_y P_{i_b} - V_z f \end{pmatrix}$$



Figure 4. **a) Left:** Perspective mapping taken by the visual sensor, moving along the diagonal of the terrain after applying the transformation  $T$ . **b) Right:** The optical flow of the perspective mapping in a).

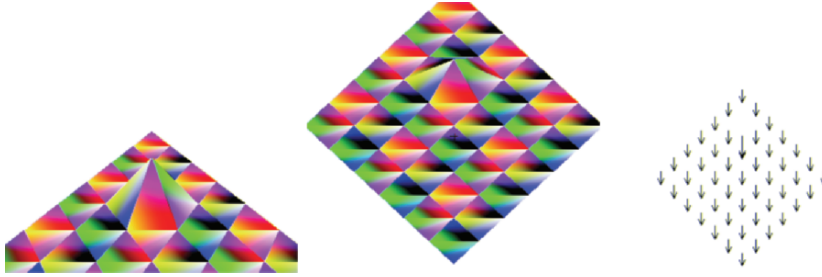


Figure 5. **a) Left:** Camera view from an arbitrary point for a terrain that contains a spike. **b) Center:** Perspective mapping of the image perceived by the visual sensor, pictured at the start of the simulation. **c) Right:** Optical flow obtained from the perspective mapping in b).

Here,  $(N_x, N_y, N_z)$ ,  $(U_x, U_y, U_z)$ , and  $(V_x, V_y, V_z)$  are the sensor axial components described in the world coordinate system  $W$ ,  $(P_{i_a}, P_{i_b})$  is a point in the image plane described by the image plane coordinate system  $I(a, b)$ , and  $h$  is the height of the visual sensor from the ground. Figure 6 shows the inverse perspective mappings corresponding to the images in Figure 5.

The result of this transformation on the optical flow field displayed in Figure 3 is shown in Figure 5c, where the totality of optical flow vectors are parallel to each other, as expected from the application of the inverse perspective mapping.

### 3. Mallot's Model and Uneven Terrain

Generally, applying Mallot's model from a mobile agent moving on an uneven terrain yields optical flow fields in which vectors may not be exhibiting parallelism among their constituent vectors. This is exemplified by the



Figure 6. a) **Left:** Inverse perspective mapping for the image in Figure 5a. b) **Right:** Inverse perspective mapping for the image in Figure 5b.

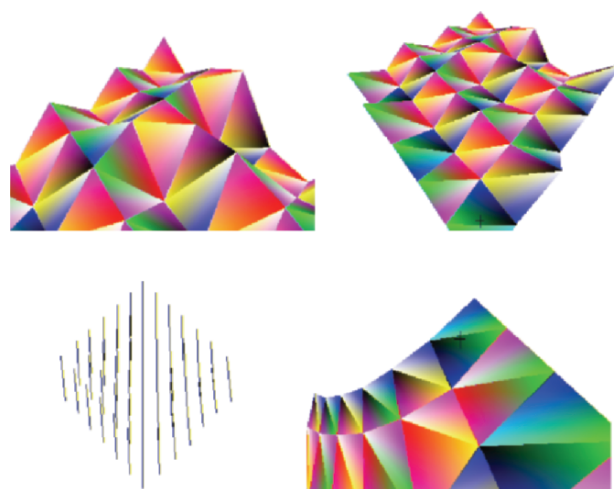


Figure 7. a) **Top-left:** Camera view from an arbitrary point for a terrain. b) **Top-Right:** Perspective mapping of the image perceived by the visual sensor, at the start of the simulation. c) **Bottom-Left:** Optical flow obtained from the perspective mapping in b). d) **Bottom-Right:** Inverse Perspective mapping for the image in b).

following case, where Figure 7a shows a 3d surface of irregular terrain and Figures 7b, 7c, and 7d display the perspective mapping, resulting optical flow, and the inverse perspective mapping respectively.

Figure 8 shows the reason behind the incorrect optical flow of Figure 7c. The point  $P$  in Figure 8 is on the terrain and the dashed line represents the path the agent must follow to keep the angle between its sensor and the horizon constant as it travels on the surface. Because the surface is uneven,

this angle varies and the optical flow vectors deviate from the parallelism they should exhibit. This demonstrates the inadequacy of Mallot's model on irregular navigational terrains.

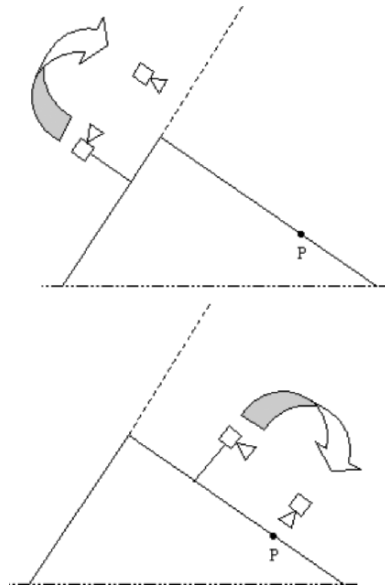


Figure 8. Mallot's inverse perspective on rough terrain.

#### 4. The Proposed Model

As previously stated and under the conditions created by an uneven navigational surface, Equation (2) yields an incorrect optical flow, and a further transformation  $T$  is needed to null its effects. Hence, Equation (2) may be rewritten as:

$$\begin{pmatrix} Q_{i_a} \\ Q_{i_b} \\ 1 \end{pmatrix} = \frac{-f}{P_{c_z}} \begin{pmatrix} P_{c_x} \\ P_{c_y} \\ P_{c_z} \end{pmatrix} T(\theta) \quad (3)$$

where  $T(\theta)$  is a rotation matrix, and  $\theta$  is the angle between the optical axis of the camera and the perpendicular to the absolute horizon<sup>1</sup>. For Mallot's inverse perspective to be valid under the hypothesis of an uneven

<sup>1</sup> We define the absolute horizon as the plane perpendicular to the vector describing the direction of the gravitational field.



navigational terrain, one must find the transformation  $T$  which allows the visual sensor's angle relative to the absolute horizon to remain constant regardless of the slope of the terrain over which the agent moves.

As the agent navigates, the transformation  $T(\theta)$  evolves in relation to the angle that the sensor makes with the direction of the gravitational field. Provided that the agent is fitted with adequate gyroscopic equipment, then the vector describing the direction of gravity is available and the plane to which this vector is perpendicular represents the flat navigational surface which Mallot's model requires to perform adequately.

Assuming that the agent is so equipped as to instantaneously measure the pitch and roll angles it makes with respect to the aforementioned plane, then the model can be generalized in the following fashion:

$$\begin{pmatrix} P_{i_a} \\ P_{i_b} \\ 1 \end{pmatrix} = \begin{pmatrix} Q_{i_a} \\ Q_{i_b} \\ 1 \end{pmatrix} \cdot P(\alpha) \cdot R(\phi) \tag{4}$$

where  $\alpha$  and  $\phi$  are the respective pitch and roll angles:

$$P(\alpha) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha & \sin \alpha \\ 0 & -\sin \alpha & \cos \alpha \end{pmatrix}$$

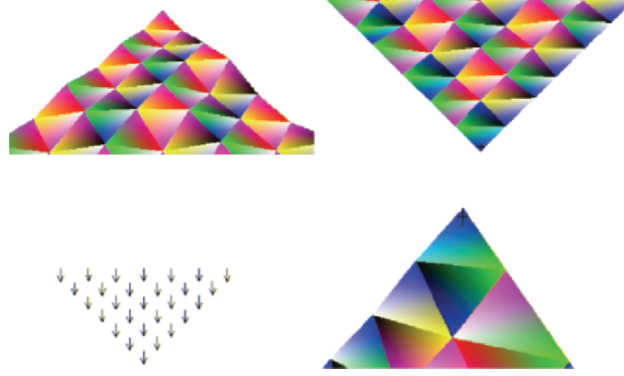
$$R(\phi) = \begin{pmatrix} \cos \phi & \sin \phi & 0 \\ -\sin \phi & \cos \phi & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

As it navigates on an uneven terrain, the mobile agent experiences height variations with respect to any arbitrarily determined reference point on the terrain. This, of course, introduces unwanted perspective effects, even while pitch and roll are being corrected in the imagery acquired by the sensor. Therefore, a third transformation, this time requiring both the gravimeter and the speed of the sensory agent as inputs, needs to be formulated.

Figure 11 shows the agent moving on such a rough terrain. As the camera moves further down, the height of the camera with respect to a terrain point  $P$  decreases, thus creating a perspective effect. The following Equation shows the transformation  $T_{h'}$  which compensates for the perspective:

$$T_{h'} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -h' & 1 \end{pmatrix} \tag{5}$$

where  $h'$  is the difference in camera height with respect to a virtual plane, normal to the direction of gravity. It is obtained in the following way:



*Figure 9.* **a) Top-left:** Camera view from an arbitrary point about a moderately rough terrain. **b) Top-Right:** Perspective mapping of the image perceived by the visual sensor for the camera view in a). **c) Bottom-Left:** Optical flow obtained from the perspective mapping in b). **d) Bottom-Right:** Inverse Perspective mapping for the image in b).

assuming that the robot is moving with velocity  $V$ , then the distance per time interval traversed by the robot is equal to:

$$\delta S = Vt \quad (6)$$

Given that the angle of the terrain surface is known to be  $\rho$  by way of a gravimeter, then the change in camera height  $h'$  with respect to the virtual plane is obtained as follows:

$$h' = \delta S \sin \rho \quad (7)$$

The next Equation shows how this last transformation is combined with the two previous ones:

$$\begin{pmatrix} P_{i_a} \\ P_{i_b} \\ 1 \end{pmatrix} = \begin{pmatrix} Q_{i_a} \\ Q_{i_b} \\ 1 \end{pmatrix} \cdot T_{h'} \cdot P(\alpha) \cdot R(\phi) \quad (8)$$

Figures 9a and 10a show the camera view for different terrains with different roughness. Figures 9b and 10b are the perspective mapping; Figures 9c and 10c are the optical flow; and Figures 9d and 10d are the inverse perspective mapping, respectively.

## 5. Noise Analysis and Sensitivity

The orientation and magnitude of ground-truth optical flow fields were corrupted by two independent, zero-mean Gaussian distributions. Consider

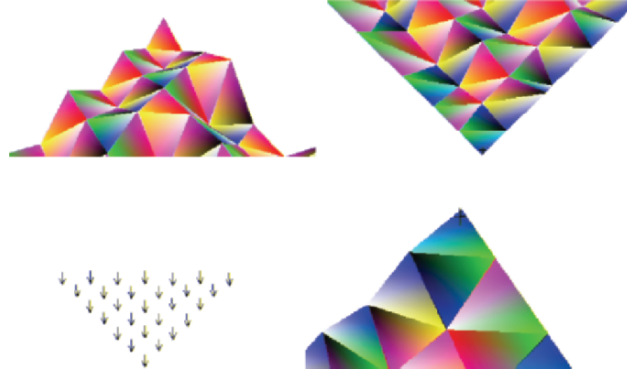


Figure 10. **a) Top-Left:** Camera view from an arbitrary point about a very rough terrain. **b) Top-Right:** Perspective mapping of the image perceived by the visual sensor for the camera view in a). **c) Bottom-Left:** Optical flow obtained from the perspective mapping in b). **d) Bottom-Right:** Inverse Perspective mapping for the image in b).

$\epsilon_{angle}$ , a randomly generated number from a zero-mean Gaussian distribution with standard deviation  $\sigma_{angle}$ . We formed the disturbance angle  $\theta_d$  as:

$$\theta_d = \epsilon_{angle} 2\pi. \quad (9)$$

Consider  $\epsilon_{mag}$ , a randomly generated number from a zero-mean Gaussian distribution with standard deviation  $\sigma_{mag}$ . We formed the disturbance value to be added to the magnitudes of optical flow vectors as:

$$\ell_{noisy} = \epsilon_{mag} \times \ell_{orig}. \quad (10)$$

The output noise in the terrain reconstruction process is represented by the Sum of Squared Errors (SSE) between a noise-free inverse perspective mapping and the noisy one, reconstructed with the corrupted optical flow vectors. The following Equation represents our noise metric:

$$\text{SSE} = \sum_{i=1}^n \sqrt{(x_i - \bar{x}_i)^2 + (y_i - \bar{y}_i)^2} \quad (11)$$

where  $x_i$  and  $y_i$  are the reconstructed coordinates of a point  $P_i$  in the image, and  $\bar{x}_i$  and  $\bar{y}_i$  are the corresponding noisy ones.

Figure 12 shows the relation between the two standard deviations  $\sigma_{angle}$ , and  $\sigma_{mag}$ , within the range 0.0001 and 0.05 with step 0.01 and the SSE metric. We observe that the error increases non-linearly with the progression of the standard deviation that corrupts the magnitude of the optical

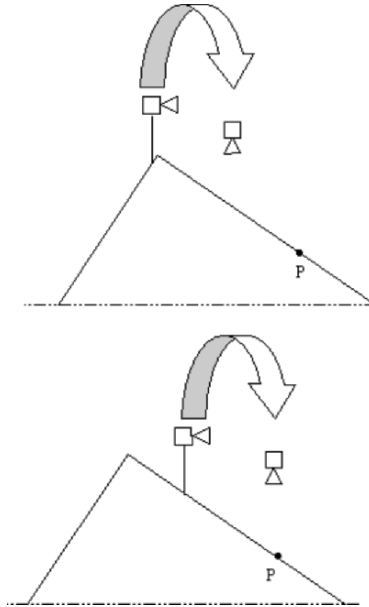


Figure 11. The camera, for the proposed model.

flow vectors. However, the output error behavior for the input optical flow directional error appears to be linear.

It is apparent from this analysis that linear input noise generates non-linear output noise in the terrain reconstruction process. We believe this effect to be mainly due from expected sources, including the behavior of perspective projection equations and the relationship between optical flow from a bird's eye perspective and the depth of environmental surfaces.

## 6. Conclusion

We proposed a mathematical model for optical flow-based autonomous navigation on uneven terrains. We provided a detailed explanation on the inadequacy of Mallot's inverse perspective scheme for uneven navigational surfaces. The model was extended to include these types of surfaces.

Our generalization of Mallot's model relies on the knowledge of the direction of the gravitational field and the speed of the mobile agent. We believe that visual information must be fused with other sources of information, such as one's position with respect to the direction of gravity, odometry, and inertial information. In addition, our model can be further

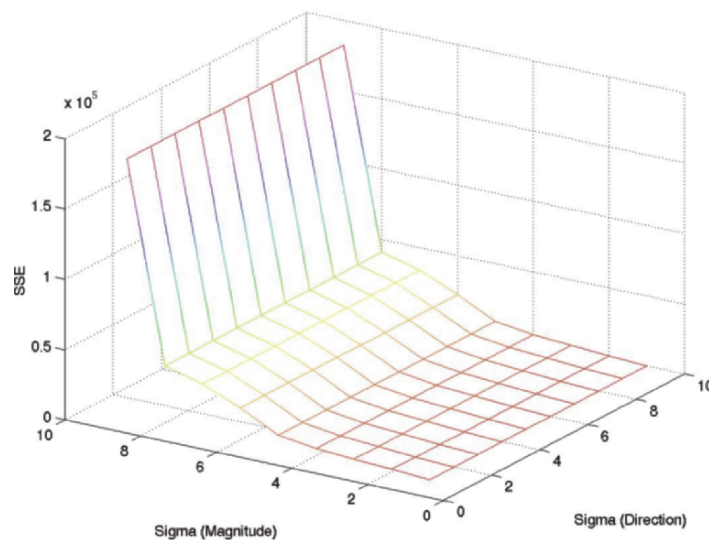


Figure 12. SSE versus standard deviation, representing the noise in optical flow vector magnitudes and directions. Each unit in the graph represents 0.01 standard deviation. The experiment displays a standard deviation range from 0.0001 to 0.05.

extended to compensate for acceleration, as long as this information is made available to the vision system through odometry.

We are currently working towards generalizing our approach to stereo vision systems, so as to obtain multiple channels of visual information, onto which cue selection and integration could be performed, thus enhancing the robustness of the approach.

## References

- Barron, J. L., D. J. Fleet, and S. S. Beauchemin: Performance of optical flow techniques. *IJCV*, **12**: 43–77, 1994.
- Batavia, P. H., S. A. Roth, and S. Singh: Autonomous coverage operations in semi-structured outdoor environments. In *IEEE RSJ Int. Conf. Intelligent Robots and Systems*, October 2002.
- Baten, S., M. Lutzeler, E. D. Dickmanns, R. Mandelbaum, and P. J. Burt: Techniques for autonomous, off-road navigation. *IEEE Intelligent Systems*, **13**, 1998.
- Choi, W., C. Ryu, and H. Kim: Navigation of a mobile robot using mono-vision and mono-audition. In *IEEE Int. Conf. Systems, Man and Cybernetics*, 1999.
- Desouza, G. N. and A. C. Kak: Vision for mobile robot navigation: a survey. *IEEE Trans. Pattern Analysis Machine Intelligence*, **24**, 2002.
- Jin, T., S. Park, and J. Lee: A study on position determination for mobile robot navigation

- in an indoor environment. In Proc. *IEEE Int. Symp. Computational Intelligence in Robotics and Automation*, 2003.
- Kim, Y. and H. Kim: Dense 3d map building for autonomous mobile robots. In Proc. *IEEE Int. Symp. Computational Intelligence in Robotics and Automation*, 2003.
- Ku, C. and W. Tsai: Obstacle avoidance for autonomous land vehicle navigation in indoor environments by quadratic classifiers. *IEEE Trans. Systems, Man and Cybernetics*, **29**, 1999.
- Tang, L. and S. Yuta: Vision-based navigation for mobile robots in indoor environments by teaching and playing-back scheme. In Proc. *IEEE Int. Conf. Robotics and Automation*, 2001.
- Tang, L. and S. Yuta: Indoor navigation for mobile robots using memorized omnidirectional images and robot's motion. In Proc. *IEEE RSJ Int. Conf. Intelligent Robots and Systems*, 2002.
- Wijesoma, W. S., K. R. S. Kodagoda, and A. P. Balasuryia: A laser and a camera for mobile robot navigation. In Proc. *Int. Conf. Control, Automation, Robotics and Vision*, 2002.