Vision-based Detection of Kerbs and Steps

Stephen Se and Michael Brady
Department of Engineering Science
University of Oxford
Oxford OX1 3PJ, U.K.
{syss,jmb}@robots.ox.ac.uk

Abstract

A key component of a technological aid for the partially sighted (TAPS) is a system to detect kerbs and steps. In this paper, a vision-based kerb detection system is described which uses the Hough Transform to find clusters of parallel lines in the image as evidence for a kerb. This is combined with a stereo vision-based obstacle detection algorithm. Experiments show that kerb regions are identified correctly from the images. An error analysis of the obstacle detection algorithm enables the kerb height, and its uncertainty, to be determined.

1 Introduction

The system we describe here provides part of an obstacle avoidance capability for ASMONC (Autonomous System for Mobility, Orientation, Navigation and Communication), a project which aims to provide a full navigation and mobility capability for partially sighted people. A major requirement for the vision system in ASMONC is to detect small obstacles as well as kerbs/steps to help the user navigate safely along a path.

Ground Plane Obstacle Detection (GPOD) using stereo disparity was first reported by Sandini *et al.* [4] and subsequently refined by Mayhew *et al.* [7] and by Li [5, 6]. It has been adapted to detect small obstacles for the partially sighted [11, 8]. The Sobel edge detector used currently in GPOD detects vertical and near vertical edges, which are then matched by the PMF algorithm [9, 10]. This works well in practice for obstacle detection as they are often standing on the ground.

However, apart from obstacles, a partially sighted person also needs to be informed of the presence and locations of kerbs or steps. We have developed a separate module using a general Canny edge detector [2] and the Hough Transform [12] to detect such obstacles.

2 Ground Plane Obstacle Detection

GPOD parameterises the ground plane using measurements of disparity. It includes an initial calibration stage in which the ground plane parameters are extracted from images of the ground containing line features but no obstacles. It

then compares the disparity values in a new image pair with the expected ground plane disparity to detect differences (hence obstacles). The ground plane disparity d varies linearly with cyclopean image plane position [6], that is

$$d = au + bv + c \tag{1}$$

where (u, v) is the cyclopean image coordinates.

We obtain a least-squares fit for the ground plane parameters (a, b, c) by orthogonal regression. However, as the image coordinates and the measured disparities are not noise-free, we study the covariance of the (a, b, c) estimate because this tells us how much confidence we can have in our ground plane estimate.

Re-arranging Equation 1, we have $au + bv - d + c = 0$, and so we minimise

$$C = \sum_{i=1}^{n} \frac{(\mathbf{p}_i^\top \mathbf{l})^2}{l_a^2 + l_b^2 + l_c^2}$$

where $\mathbf{p}_i = (u_i, v_i, d_i, 1)$ and $\mathbf{l} = (l_a, l_b, l_c, l_d)$. The ground plane parameters (a, b, c) are obtained as $(-\frac{l_a}{l_c}, -\frac{l_b}{l_c}, -\frac{l_d}{l_c})$. This problem is equivalent to

$$\min_{\mathbf{l}} C = \mathbf{l}^\top \mathbf{Q} \mathbf{l}$$

subject to

$$l_a^2 + l_b^2 + l_c^2 = 1,$$

in which $\mathbf{Q} = \sum_{i=1}^{n} \mathbf{p}_i \mathbf{p}_i^\top$, i.e.

$$\mathbf{Q} = \begin{bmatrix} \sum u_i^2 & \sum u_i v_i & \sum u_i d_i & \sum u_i \\ \sum u_i v_i & \sum v_i^2 & \sum v_i d_i & \sum v_i \\ \sum u_i d_i & \sum v_i d_i & \sum d_i^2 & \sum d_i \\ \sum u_i & \sum v_i & \sum d_i & n \end{bmatrix}$$

The Lagrangian for this is given by

$$C' = \mathbf{l}^\top \mathbf{Q} \mathbf{l} + \mu(l_a^2 + l_b^2 + l_c^2 - 1)$$

Setting $\frac{\partial C'}{\partial \mathbf{l}} = 0$, we have

$$\mathbf{Q} \mathbf{l} + \mu \begin{bmatrix} l_a \\ l_b \\ l_c \\ 0 \end{bmatrix} = \mathbf{0}$$

The solution for (l_a, l_b, l_c) is the eigenvector corresponding to the least eigenvalue of the matrix

$$\begin{bmatrix} q_{11} - \frac{q_{14}^2}{q_{44}} & q_{12} - \frac{q_{14}q_{42}}{q_{44}} & q_{13} - \frac{q_{14}q_{43}}{q_{44}} \\ q_{12} - \frac{q_{14}q_{42}}{q_{44}} & q_{22} - \frac{q_{24}^2}{q_{44}} & q_{23} - \frac{q_{24}q_{43}}{q_{44}} \\ q_{13} - \frac{q_{14}q_{43}}{q_{44}} & q_{23} - \frac{q_{24}q_{43}}{q_{44}} & q_{33} - \frac{q_{34}^2}{q_{44}} \end{bmatrix}$$

and

$$l_d = -\frac{l_a q_{41} + l_b q_{42} + l_c q_{43}}{q_{44}}$$

where q_{ij} denotes the $(i,j)^{th}$ element of matrix $\mathtt{Q}$.

The estimate of $\mathbf{l}$ gives (a,b,c). Now we need to analyse the covariance of $\mathbf{l}$. To do this, we follow the technique outlined in Faugeras [3] (pages 151-158) for the constrained minimisation case. Assuming that $\mathbf{l}_0$ has been obtained by minimising the criterion function $C(\mathbf{p}_0, \mathbf{l})$ subject to the constraint above, this defines implicitly a function $\mathbf{f}$ such that $\mathbf{l} = \mathbf{f}(\mathbf{p})$ in a neighbourhood of $(\mathbf{p}_0, \mathbf{l}_0)$.

We define the vector $\Phi(\mathbf{p}, \mathbf{l})$ by

$$\Phi(\mathbf{p}, \mathbf{l}) = \begin{bmatrix} \mathtt{Q}^{[1]}\mathbf{l} - (l_a/l_c)\mathtt{Q}^{[3]}\mathbf{l} \\ \mathtt{Q}^{[2]}\mathbf{l} - (l_b/l_c)\mathtt{Q}^{[3]}\mathbf{l} \\ \mathtt{Q}^{[4]}\mathbf{l} \\ l_a^2 + l_b^2 + l_c^2 - 1 \end{bmatrix}$$

where $\mathtt{Q}^{[i]}$ denotes the ith row of matrix $\mathtt{Q}$. The Jacobian of $\mathbf{f}$ is given by

$$\boldsymbol{\nabla}\mathbf{f} = -(\frac{\partial \Phi}{\partial \mathbf{l}})^{-1}\frac{\partial \Phi}{\partial \mathbf{p}}$$

Assuming that the error at each point is independent and that the errors are isotropic, the covariance matrix $\Lambda_{\mathbf{p}}$ for the original data is block diagonal in form with $\mathbf{p}_i$'s covariance matrix as the ith block, assuming all points have the same diagonal covariance matrix,

$$\Lambda_{\mathbf{p}_i} = \begin{bmatrix} \sigma^2 & 0 & 0 & 0 \\ 0 & \sigma^2 & 0 & 0 \\ 0 & 0 & \sigma^2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

The covariance matrix for $\mathbf{l}$ is given by

$$\Lambda_{\mathbf{l}} = \boldsymbol{\nabla}\mathbf{f}\Lambda_{\mathbf{p}}\boldsymbol{\nabla}\mathbf{f}^{\top}$$

The ground plane parameters (a,b,c) are obtained as $(-\frac{l_a}{l_c}, -\frac{l_b}{l_c}, -\frac{l_d}{l_c})$, in which l_c is always close to unity and its variance is smaller than those of the others by a magnitude of at least 2. For this reason, we take the variances for a, b and c to be $\sigma_{l_a}^2$, $\sigma_{l_b}^2$ and $\sigma_{l_d}^2$ respectively [1].

From Equation 1, we can now determine the disparity variance, using the error propagation formula for the product of two uncorrelated distributions [1] given by $var(xy) = x^2 var(y) + y^2 var(x)$. We find that the expected disparity is

$$\sigma_d^2 = (a^2\sigma_u^2 + u^2\sigma_a^2) + (b^2\sigma_v^2 + v^2\sigma_b^2) + \sigma_c^2 \qquad (2)$$

3   Hough Transform

Typically, in Europe and the USA, a kerb is accompanied by a number of parallel lines close together. Weak perspective projection is a good approximation to perspective projection in this case as the variation in depth of the scene is small compared to the depth along the line of sight. As a result, parallel lines in the world project to parallel lines in the image, so we can use the Hough Transform [12]
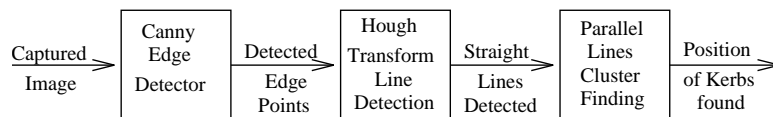
Figure 1: *Canny edge detection and Hough Transform kerb finding.*
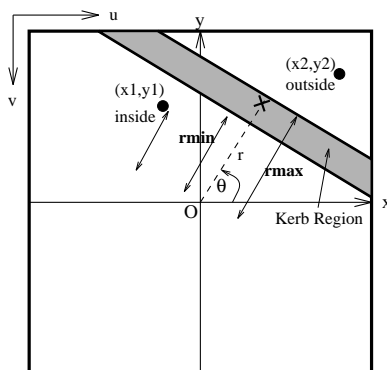
Figure 2: *Geometry for edge point classification.*

to detect clusters of parallel lines in the image as evidence for a kerb. As we do not need disparity information, we just consider a single image. The flow diagram is as shown in Figure 1.

Since kerbs are usually long, we model them as infinite lines. For point (u, v), the quantised Hough space is (r, θ) where θ is the angle rotated, and r is the distance from the origin of the x-y coordinate system as shown in Figure 2. We have

$$r = (u - \frac{W}{2}) \cos \theta + (\frac{W}{2} - v) \sin \theta$$

where W is the dimension of image.

We accumulate evidence for straight lines from the set of detected edge points obtained from the Canny detector. Then we extract the small number of (r_i, θ_i)s which receive most support. From these straight lines, we search for clusters of at least three parallel lines which are close together, and return $(\theta, r_{min}, r_{max})$ for the kerb region perpendicular to slope $\tan \theta$ extending from r_{min} to r_{max} as shown in Figure 2. This gives the estimated position of the kerb region.

Figure 3 shows some images of the same kerb viewed from various angles. Figure 4 shows the kerb regions found.
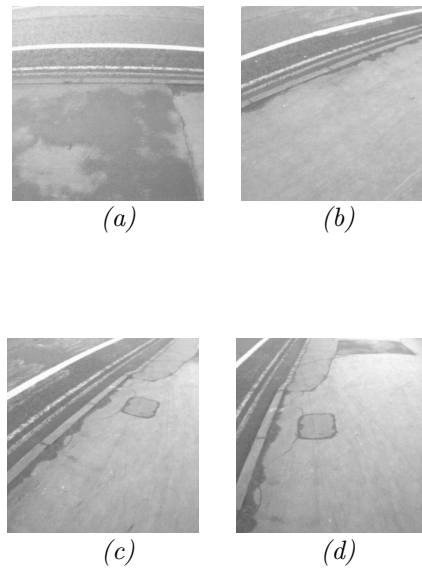
Figure 3: *Kerb images from various angles. (a) Horizontally. (b) 25 degrees to the horizon. (c) 40 degrees to the horizon. (d) 70 degrees to the horizon.*
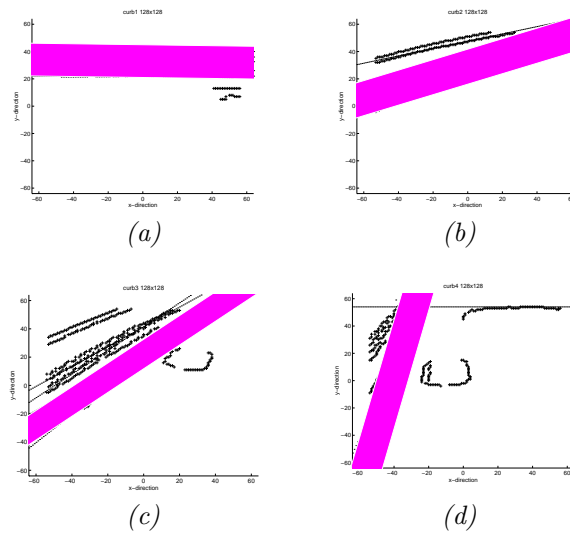

Figure 4: *Canny detected edge points with Hough Transform detected lines and expected kerb region for kerb images. (a) Horizontally. (b) 25 degrees to the horizon. (c) 40 degrees to the horizon. (d) 70 degrees to the horizon.*
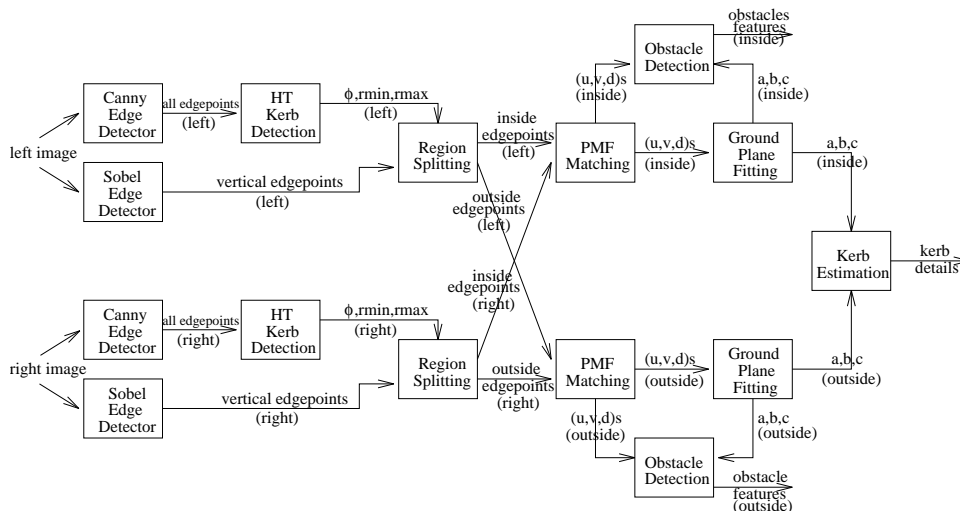
Figure 5: *Flow diagram for the integration of Ground Plane Obstacle Detection and Kerb Detection.*

4 Estimation and Uncertainty

After a candidate kerb has been localised, we can segment the image edge points into three regions: one for the kerb, one for pavement and one for road. Moreover, we would like to determine whether it is a step-up or step-down and estimate its size.

In the x-y coordinate system, for each edge point (x, y) found by the edge detector, we can compute the distance to the origin of the line passing through this point and perpendicular to slope $\tan\theta$:

$$r = x\cos\theta + y\sin\theta$$

We compare the r for each point with r_{min} and r_{max}:

- If $r > r_{max}$, then the point is classified as belonging to the outside region, for example, point (x_2, y_2) in Figure 2.

- If $r < r_{min}$, then the point is classified as belonging to the inside region, for example, point (x_1, y_1) in Figure 2.

- If $r_{min} < r < r_{max}$, the point is in the kerb region.

We do this for both images, provided that there are sufficiently many features on the ground. We then perform stereo matching and ground-plane fitting in the inside and outside regions individually, finally obstacle detection. The flow diagram in Figure 5 shows the integration of GPOD with kerb detection.

We have

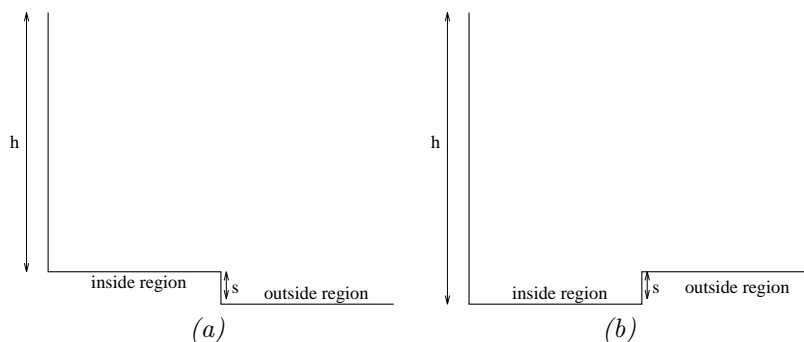$$u = \left(\frac{r_{min} + r_{max}}{2}\right)\cos\theta + \frac{W}{2}$$

Figure 6: *Kerb situations. (a) A step-down. (b) A step-up.*

$$v = \frac{W}{2} - (\frac{r_{min} + r_{max}}{2}) \sin \theta$$

for the estimated kerb position in each image. Averaging between the left and right images gives the cyclopean kerb position (u_k, v_k) which can be converted to give approximate 3-D range information with the roughly calibrated camera intrinsic parameters.

Let (a_i, b_i, c_i) and (a_o, b_o, c_o) be the ground plane parameters obtained for the inside and outside regions respectively. We have

$$d_{inside} = a_i u_k + b_i v_k + c_i$$

$$d_{outside} = a_o u_k + b_o v_k + c_o$$

If $d_{inside} > d_{outside}$, it is a step-down, else it is a step-up.

For a step-down of size s as shown in Figure 6(a), the effect on the ground plane parameters from inside region to outside region is like increasing the height h of the camera system by s. It can be shown [11] that

$$d_{outside} = d_{inside}(\frac{h}{h + s}) \Rightarrow s = (\frac{d_{inside}}{d_{outside}} - 1)h$$

Similarly for a step-up case as shown in Figure 6(b),

$$d_{outside} = d_{inside}(\frac{h}{h - s}) \Rightarrow s = (1 - \frac{d_{inside}}{d_{outside}})h$$

Using the error propagation formula for the quotient of two uncorrelated distributions [1], we have

$$var(s) = h^2 (\frac{var(d_{inside})}{d_{outside}^2} + \frac{d_{inside}^2 var(d_{outside})}{d_{outside}^4})$$

In this way, not only can we estimate the height of the kerb, but its uncertainty can be computed.

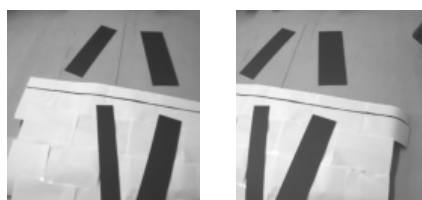Figure 7: *Stereo pair for a step-down.*

Figure 8: *Stereo pair for a step-down from another angle.*

Figure 7 shows an 128x128 stereo image pair for an artificial kerb scene. The kerb is approximately 8cm high and 100cm away from the cameras, which are 120cm above the ground. We apply the above approach to find the kerb region in order to segment the image: we find that the estimated kerb distance is 109cm. Assuming $\sigma_u^2 = \sigma_v^2 = \sigma^2 = 1$, the ground plane fitting as described in Section 2 gives:

- The inside region ground plane parameters are -0.0348, 0.2144 and 33.2667 with variances $3.838 * 10^{-5}$, $1.5275 * 10^{-5}$ and 0.4985 respectively.

- The outside region ground plane parameters are 0.1037, 0.2295 and 20.8608 with variances $9.8115 * 10^{-5}$, $7.1816 * 10^{-5}$ and 0.6240 respectively.

- The disparity for the inside kerb region is 44.0466 with variance 0.7388.

- The disparity for the outside kerb region is 40.8520 with variance 1.2992.

- Since the inside disparity is greater than the outside disparity, it is a step-down: its mean size and standard deviation are estimated as 9.38cm and 4.41cm respectively.

Figure 8 shows the step-down from another angle, Figure 9 shows the graph for the matched edge points together with the fitted ground plane using all points and the fitted ground planes using just points in the lower or upper region. This is a typical two-population problem and it can be seen that ground-plane fitting the points in each individual region will help in more accurate small obstacle detection.
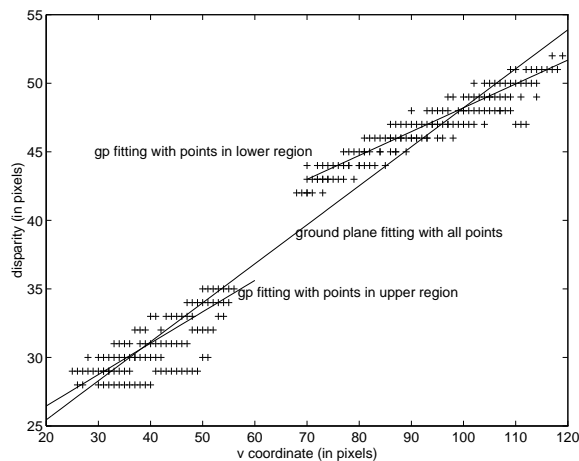
Figure 9: *Ground plane fitting using just points in the lower region, upper region and whole region.*

5 Conclusions

We have applied the Hough Transform to detect kerbs in images and demonstrated its feasibility on some real road scenes. The uncertainty analysis of the ground plane parameters fit enables us to investigate kerb estimation uncertainty. The kerb detection module developed has been integrated with the GPOD system, which can now provide kerb information to the partially sighted and also detect small obstacles better.

Currently, we are experimenting with the kerb detection algorithm in various scenes to detect different types of kerbs, and extending it for steps/stairs.

The current kerb detection implementation, however, takes 3 seconds on the average to process a pair of 128x128 images on an Ultra-Sparc machine. So before it can be actually used by the partially sighted, we need to achieve at least near real-time speed. From Figure 5, we can see that it can be optimised substantially by parallelisation: the left and right images can be processed individually, and the inside and outside regions can be processed separately.

Acknowledgements

References

[1] P.R. Bevington and D.K. Robinson. *Data Reduction and Error Analysis for the Physical Sciences*. McGraw-Hill, second edition, 1992.

[2] J. Canny. A computational approach to edge detection. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 8(6):679–698, 1986.

[3] O. Faugeras. *3 Dimensional Computer Vision - A Geometric Viewpoint*. MIT Press, 1993.

[4] F. Ferrari, E. Grosso, G. Sandini, and M. Magrassi. A stereo vision system for real time obstacle avoidance in unknown environment. In *Proceedings of IEEE International Workshop on Intelligent Robots and Systems IROS '90*, pages 703–708, 1990.

[5] F. Li. Visual control of AGV obstacle avoidance. DPhil First Year Report, Department of Engineering Science, University of Oxford, 1994.

[6] F. Li, J.M. Brady, I. Reid, and H. Hu. Parallel image processing for object tracking using disparity information. In *Second Asian Conference on Computer Vision ACCV '95*, pages 762–766, Singapore, December 1995.

[7] J.E.W. Mayhew, Y. Zheng, and S. Cornell. The adaptive control of a four-degrees-of-freedom stereo camera head. In H.B. Barlow, J.P. Frisby, A. Horridge, and M.A. Jeeves, editors, *Natural and Artificial Low-level Seeing Systems*, pages 63–74. The Royal Society, London, 1992.

[8] N. Molton, S. Se, J.M. Brady, D. Lee, and P. Probert. A stereo vision-based aid for the visually impaired. *Image and Vision Computing*, 1997. to appear.

[9] S.B. Pollard, J.E.W. Mayhew, and J.P. Frisby. Implementation details of the pmf stereo algorithm. In J.E.W. Mayhew and J.P. Frisby, editors, *3D Model Recognition From Stereoscopic Cues*, pages 33–39. MIT, 1991.

[10] S.B. Pollard, J. Porrill, J.E.W. Mayhew, and J.P. Frisby. Disparity gradient, lipschitz continuity, and computing binouclar correspondences. In J.E.W. Mayhew and J.P. Frisby, editors, *3D Model Recognition From Stereoscopic Cues*, pages 25–32. MIT, 1991.

[11] S. Se. Visual aids for the blind. DPhil First Year Report, Department of Engineering Science, University of Oxford, 1996.

[12] M. Sonka, V. Hlavac, and R. Boyle. *Image Processing, Analysis and Machine Vision*. Chapman and Hall Computing, 1993.