Vision Based Mobile Robot Navigation in a Visually Cluttered Environment

Nic Snailum, Martin Smith, and Steven Dodds
Mobile Robots Research Unit
University of East London
Dagenham, Essex, RM8 2AS
United Kingdom
n.snailum@uel.ac.uk, m.c.b.smith@uel.ac.uk, s.j.dodds@uel.ac.uk

26.3.99

Keywords: polar transform, vanishing point, robot navigation, lateral position

Abstract

A novel method for determining the vanishing point in the television image of a corridor using a polar histogram is presented. A method is also proposed for using the angles of lines that converge on the vanishing point to generate steering signals to enable control of the lateral displacement of a robot travelling down a corridor. Results are presented which show that the method is robust with real images in the presence of noise, visually cluttered images and where some of the expected features are missing.

1 Introduction

1.1 Motivation

A vision-based corridor roaming robot needs to retrieve information from the environment to enable it to travel parallel to a wall, which may be achieved by steering towards 'the vanishing point' (see below) in the camera image. This does not allow the determination of the robot's lateral position in the corridor, and the information for this has typically been derived using another sensor. In this paper we present a method of obtaining the required lateral position information from the same camera image.

When a 3-D real-world scene is projected onto a 2-D image, parallel lines in 3-space transform into converging lines in the image plane due to perspective. These lines of perspective (LOPs) converge at one of a number of vanishing points (VPs). An image may in general contain three vanishing points, corresponding to the three principal scene axes, but under certain conditions one or more of these VPs may not exist in the image plane. For example, if the image plane has an axis parallel to one of the 3-space principal axes, the corresponding VP is at infinity. In the situation under consideration, with the camera mounted horizontally, the vertical axis in the image plane is parallel to the vertical axis in 3-space, and that VP is thus at infinity. Also, when the vehicle mounted camera is pointing along the axis of the nominal direction of travel, the VP representing lateral image features is at infinity, leaving only the VP formed by features parallel to the axis of travel (the z-axis in Figure 1(a)). It is the location of this VP and the properties of the image features producing it which are considered in this paper. It was noted by Nair [1], that in indoor environments, the camera elevation angle may be regarded as constant due to smooth floors, so that camera rotation is restricted to a single degree of freedom about the vertical axis, i.e. camera panning. Under these conditions, as the camera pans, the VP locus traces a straight horizontal line in the image, known as the vanishing line (VL).

Proc. TIMR. 99 "Towards Intelligent Mobile Robots", Bristol 1999.
Technical Report Series, Department of Computer Science, Manchester University, ISSN 1361 – 6161.
Report number UMCS-99-3-1. http://www.cs.man.ac.uk/cstechrep/titles99.html

1.2 Previous work

There are many environments in which the location of VPs conveys useful information, and has been used in applications for calibrating camera parameters [2], inferring object orientation [3], and for navigation purposes [4],[5],[6]. Lebegue extracted line segments in specific a priori 3D orientations and used the calculated vanishing points to correct heading inaccuracies arising from odometry errors [5].

Most methods used for extracting vanishing points generate an accumulation of line-pair intersections. A summary of these methods is provided by Tai [7]. These methods fall into two main groups, Gaussian sphere and Hough methods. All of these involve a two-dimensional accumulator grid, the size of which determines both the accuracy and computational workload. Tai presents a probabalistic method which provides a measure of VP validity [7]. Palmer used a Hough transform to locate line segments (as is usual) and then operated on the accumulator bins with a smoothing operator to achieve high accuracy [3].

The Hough transform maps the (x,y) plane to one of several parameter spaces. For example (m,c), where m is the slope and c is the intercept of the straight lines, $y = mx + c$, joining pairs of points in the x-y plane. Singularities due to vertical lines may be avoided by using another parameter set based on the equation:

$$\rho = x\cos(\theta) + y\sin(\theta) \tag{1}$$

where ρ and θ are the *polar co-ordinates* of a general point on the straight line.

This polar transform maps straight lines in the x-y plane to *curves* in the θ-ρ plane. In this case, the parameters, θ_H and ρ_H, together precisely define a straight line and all pairs of points on a particular line map onto the same pair, (θ_H, ρ_H). This is a common form of Hough transform used to identify straight lines. The polar transform allows the location of the points of intersection of pairs of straight lines by identifying the overlap of sinusoids in (θ, ρ) space. A similar technique solves equation (1) for the overlaps using pairs of (x, y) co-ordinates believed to be co-linear, and then plots the points (θ, ρ) directly without need of the sinusoids. In each case, the origin for the transform is taken to be the x-axis, so that

$0 \leq \theta < \pi$ and $0 \leq \rho < D$ where D is the length of the image diagonal in number of pixels.

Environments where mobile robots are required to operate often produce images containing a vanishing line (VL), and it is useful to distinguish between lines of similar orientation situated above and below the VL in order to avoid confusion of image features. The Hough transform, however, does not enable this but a modification to overcome this problem is presented in the following section.

Polar mapping of images seems to have been largely ignored, though it has received attention as a technique for determining object motion, particularly in connection with optical flow [1],[8]. Here, the range of polar angle is effectively doubled by choosing an origin within the image. Nair [1] uses the focus of expansion of a moving image as the polar origin to determine motion toward the observer. The merits of a polar transform using the VP as origin do not appear to have been documented. Polar transforms are, for example, able to distinguish between features above and below the VL, and may be used to detect and locate the image lines required for navigation. In section 2, a modified polar transform for detecting the VP is presented that reduces the memory overhead required for 2-D grid methods by using a one-dimensional accumulator. In section 3 the results are presented using the transform on some real images to locate the VP and indicate the lateral position of the camera for guidance purposes. The experiments which follow were carried out on a particularly visually cluttered corridor chosen to develop a robust method

2 The Proposed System

2.1 Image model

Many paths through which a robot must travel have features similar to a 'corridor'. In this study, a corridor model is used in which at least one of the wall/ceiling and wall/floor interface lines is extracted. A 'pinhole' perspective model is used, as shown in Figure 1(a). Thus, a point (X, Y, Z), in the real three-dimensional world is mapped onto a point (x, y) in the image x-y plane. Using such a model, features of interest may be projected into the image, as shown in Figure 1(a and b).

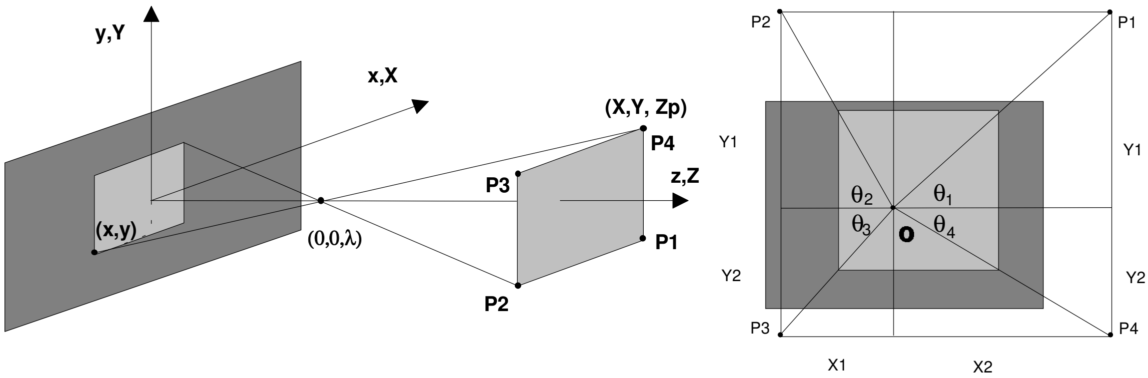


Figure 1: (a) Perspective model (b) Projection of features of interest onto the image plane (dark grey rectangle)

The features of interest are the lines of perspective, $P_1 0$, $P_2 0$, $P_3 0$ and $P_4 0$, which meet at the vanishing point 0. The points P_1, P_2, P_3 and P_4, lie on a cross section of a corridor perpendicular to the nominal direction of motion of the vehicle, some distance 'Zp' from the camera.

2.2 The automatic steering control signal

The polar plot based system directly yields estimates of the image angles, θ_1, θ_2, θ_3 and θ_4, therefore an expression for the required steering signal to actuate an automatic steering controller can be expressed in terms of these angles. As will be seen, for control of only one degree of freedom of motion, these angles contain redundant information and advantage may be taken of this in optimising the accuracy of the steering signal and hence the control. It is evident that:

$$\left.\begin{array}{ll} \tan(\theta_1) = y_1/x_2, & \tan(\theta_3) = y_2/x_1 \\ \tan(\theta_2) = y_1/x_1, & \tan(\theta_4) = y_2/x_2 \end{array}\right\} \tag{2}$$

where y_2 is the view point height and is usually known. It follows that:

$$\tan(\theta_1)/\tan(\theta_2) = \tan(\theta_4)/\tan(\theta_3) = x_1/x_2 \tag{3}$$

$$\tan(\theta_2)/\tan(\theta_3) = \tan(\theta_1)/\tan(\theta_4) = y_1/y_2 \tag{4}$$

The relationship between the angles subtended by the lines of perspective is thus dependent on the geometry of the corridor and the lateral position of the camera in the corridor. These angles and their relationships therefore contain information from which a steering signal for automatic control can be derived. If, for example, the robot is required to navigate at a central lateral position in the corridor, then $x_1 = x_2$ and therefore from equation (3), $\tan(\theta_1) = \tan(\theta_2)$ from which $\theta_1 = \theta_2$ as expected. Similarly, $\tan(\theta_3) = \tan(\theta_4)$ from which $\theta_3 = \theta_4$. In this case, the steering signal could be simply obtained from $\theta_2 - \theta_1$ or $\theta_4 - \theta_3$. If, on the other hand, the steering control system is required to respond to an arbitrary lateral position reference input, x_r, then the steering signal would be:

$$e = x_r - (x_1 - x_2) \tag{5}$$

Assuming a corridor with a constant ceiling height, y_1 and y_2 are nominally constant and therefore equations (2) may be used to express x_1 and x_2 in terms of θ_1, θ_2, θ_3 and θ_4:

$$\left.\begin{array}{l} x_1 = y_1/\tan(\theta_2) \text{ or } x_1 = y_2/\tan(\theta_3) \\ x_2 = y_1/\tan(\theta_1) \text{ or } x_2 = y_2/\tan(\theta_4) \end{array}\right\} \tag{6}$$

This allows *four* possible error signals to be formed. Thus, substituting the four possible combinations of expressions for x_1 and x_2 into equation (5) yields:

$$\left.\begin{array}{ll} e_1 = x_r - (y_1/\tan(\theta_2) - y_1/\tan(\theta_1)) & e_2 = x_r - (y_1/\tan(\theta_2) - y_2/\tan(\theta_4)) \\ e_3 = x_r - (y_2/\tan(\theta_3) - y_1/\tan(\theta_1)) & e_4 = x_r - (y_2/\tan(\theta_3) - y_2/\tan(\theta_4)) \end{array}\right\} \tag{7}$$

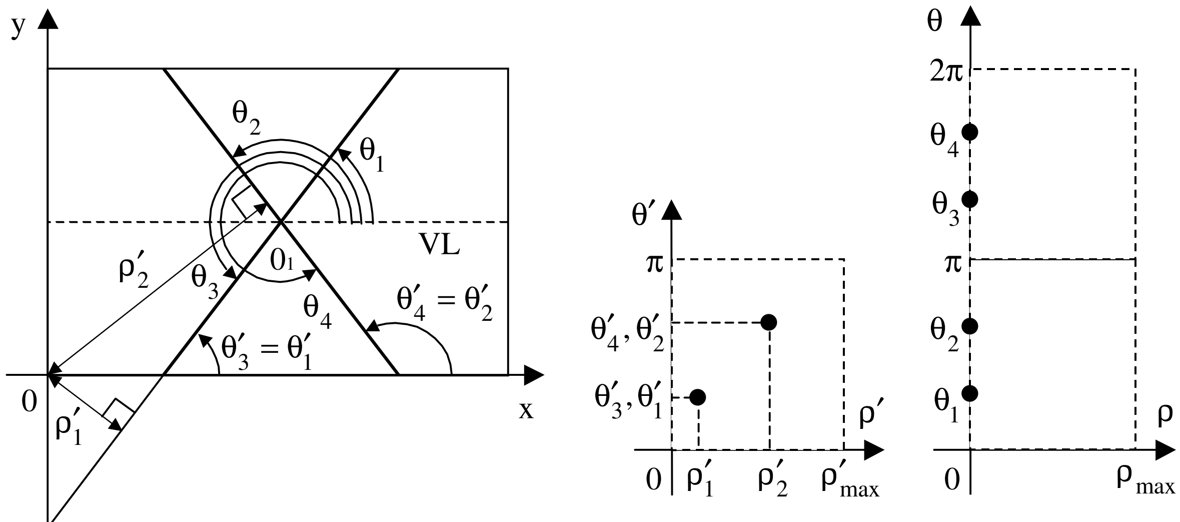
The control accuracy would then be optimised by selecting the error from $(e_1(\theta_1,\theta_2), e_2(\theta_1,\theta_3), e_3(\theta_2,\theta_4), e_4(\theta_3,\theta_4))$ expressed in terms of the pair of angles (θ_i,θ_j) deemed to be the most reliable by selecting the two highest peaks of the histograms presented in section 3.1.

2.3 Proposed polar transform and the corresponding Hough transform.

The Hough transform is modified by changing the origin for the angular transformation to the VL instead of the x-axis. This doubles the bounds of the image angles. Thus, if instead of taking the Cartesian origin for the polar distance transformation, the VP is made the origin, the range of θ is extended from $0 \le \theta < \pi$ to:

$$0 \le \theta < 2\pi \tag{8}$$

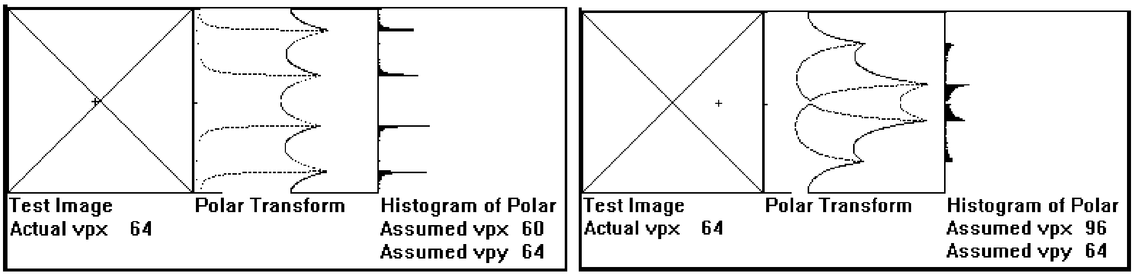
and the polar distance range is $0 \le \rho < D_1$ where D_1 is the diagonal length in pixels of the part image between the y co-ordinate extremity and the vanishing line, whichever is the largest. The image angle range extension allows lines of similar orientation above and below the VL to be distinguished from one another, as illustrated in Figure 2.



a) Ideal Cartesian image model b) Standard Hough accumulation c) The proposed accumulation

Figure 2: Standard and modified Hough transforms of symmetrical floor-wall-ceiling interface lines

In this case, the 'range', ρ, is no longer defined as the perpendicular distance of the line from the origin, but is the actual distance of the pixel from the origin. This has the effect of elongating the pixel clusters in Figure 2(c), the minimum values of ρ coinciding with that of the Hough transform, i.e., zero, for the lines of the ideal corridor model which pass through the origin, O_1. If the origin is taken as a first approximation as some other point on the vanishing line, as illustrated in Figure 3, the pixel clusters are further dispersed in the direction of the angle axis in Figure 2(c). This is due to the continuous variation in the angle, θ, subtended by the line pixels to the new origin, O_1, with position along the line. Figure 3 presents a test image (left) and its polar transform (centre) using two (incorrect) assumed vanishing points (AVPs). The correct vanishing point is located at (x,y) = (64,64). Figure 3(a) shows the effect of the AVP = (60,64), and in Figure 3(b) the AVP is (96,64). The AVP location in each instance is marked.



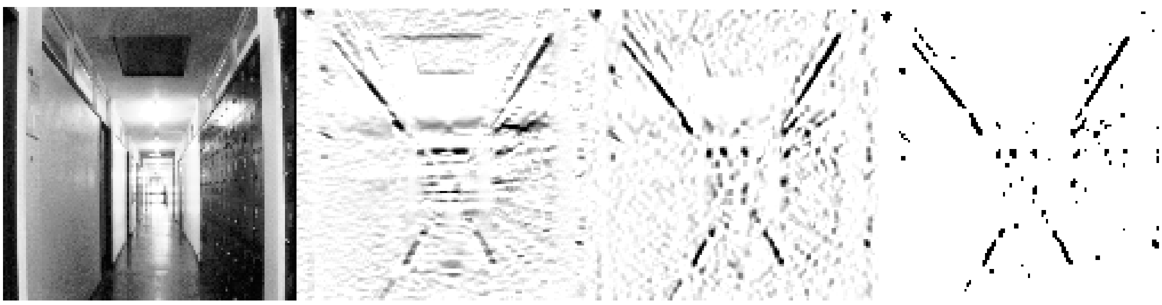
a) AVP at 60 b) AVP at 96

Figure 3: Test image and polar transform with erroneous AVP locations

As proven in section 2.2, the lateral position of the camera and hence the robot can be determined from the angular data of the lines of perspective alone. Thus the ‘range’ data (ρ) in the transform is not needed. All pixels subtending a similar angle to the origin are thus counted identical, and the transform may be reduced once again to a one dimensional plot of the number of feature pixels vs. polar angle, that is, a histogram, as shown in Figure 3 (a and b) (right). It is observed that as the AVP becomes less accurate, the polar histogram peaks reduce and there is a greater overlap of the image features in the polar transform, blurring the peaks. However, for distinct image lines the peaks remain distinct with moderate AVP inaccuracy as will be shown in section 3. Thus, the angle estimates required for the steering error algorithm of section 2.2 may be taken as the medians of the histogram peaks.

2.4 Image pre-processing

Pre-processing is required for the reduction of the data in real-world images to images suitable for use by the system presented above. Several data reduction schemes have been tried on different real, cluttered images, and the result of one computationally efficient sequence is shown in Figure 4.



a) Original image b) Prewitt horizontal filter c) Prewitt vertical filter d) threshold of 50%

Figure 4: Results of an image pre-processing sequence.

As shown, filters are first applied to the original image, one after the other. Correlation filters using only integer kernel values, are more efficient in operation than those requiring floating point calculations, especially when further restricted to 2^n, which may be implemented using binary shifts. The Prewitt filter for example requires only addition and subtraction of image pixel values, as its kernel comprises only +1, 0, or –1 values. A threshold function is then applied to yield a binary image defining the pixels for which values of θ are calculated. It was found that histogram equalisation of the original image improved the extraction of the perspective lines marginally. (Images were reduced from 768x576 pixels to 128x128 pixels to reduce the processing time.)

3 Results

3.1 Generation of polar plots and histograms

After pre-processing various images in a similar fashion to that described in section 2.4, the polar transform and polar histogram were obtained. The effect of using the 'wrong' origin was investigated and found to confirm the theory. The effect of additive noise was also investigated to further assess the method's robustness. Figure 5 presents the polar plot and polar histogram of the image in Figure 4(a) after pre-processing.

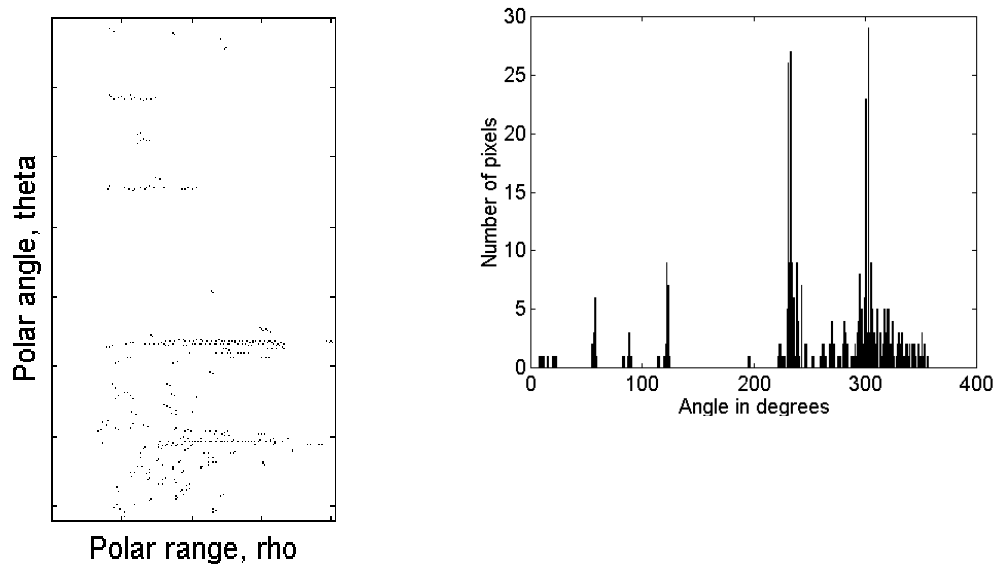


Figure 5: Polar plot and polar histogram from the image in Figure 4(a)

3.2 Determination of the vanishing point

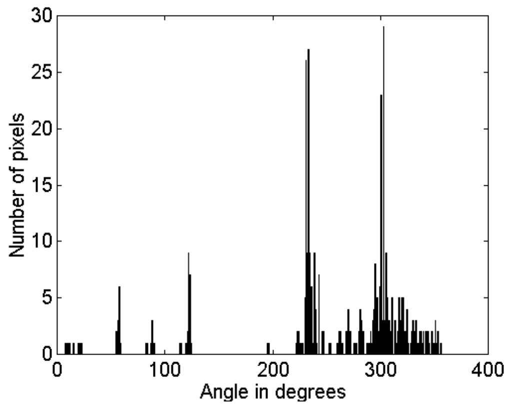


Fig 6 a) with correct AVPx = 68

Figure 6 presents polar histograms of the same image calculated from different AVP 'x' positions. The four peaks are noticeable when the AVP is correct, and still distinguishable with realistic AVP errors. The two main peaks maintain a good "signal to noise" ratio despite the error, which leads to the expectation that the method will be robust in the presence of inaccurate steering. Apparently there are no significant trends except the differences in peak values

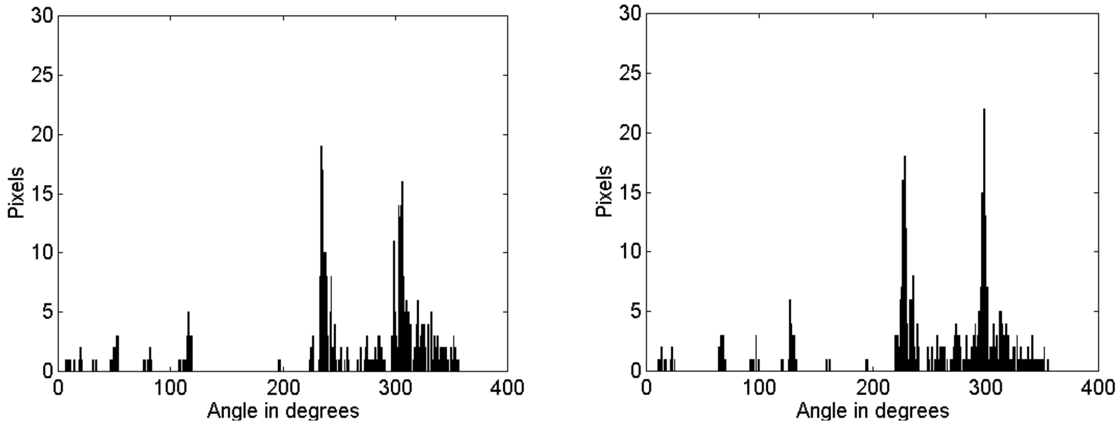


Fig 6 b) with VP-5, AVPx = 63

Fig 6 c) with VP+5, AVPx = 73

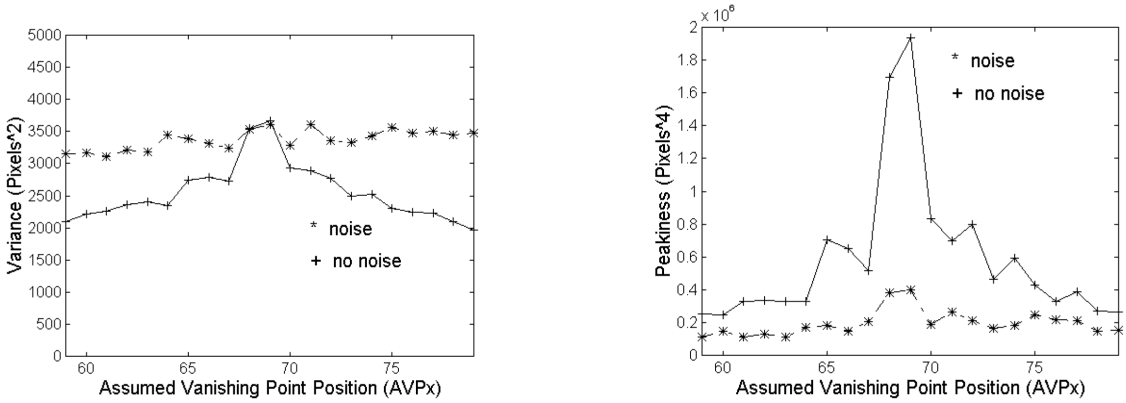
Figure 6: Polar histograms

Two measures have been tested to optimise the AVP position:

i) the polar histogram variance $var = \Sigma\ (ph-\mu)^2$ (9)

ii) the polar histogram 'peakiness' $pk = \Sigma\ (ph-\mu)^4$ (10)

The values of these two measures yielded by differing AVPs are plotted under two conditions: 1) using the raw image, and 2) with additive Gaussian noise (0,0.005), as shown in Figure 7.



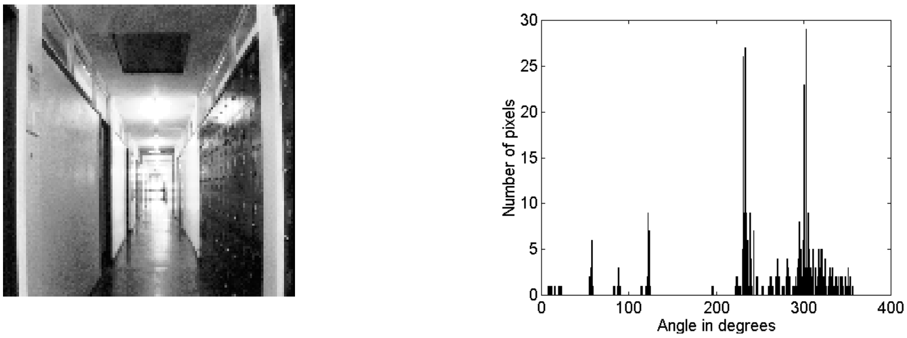
a) variance

b) peakiness

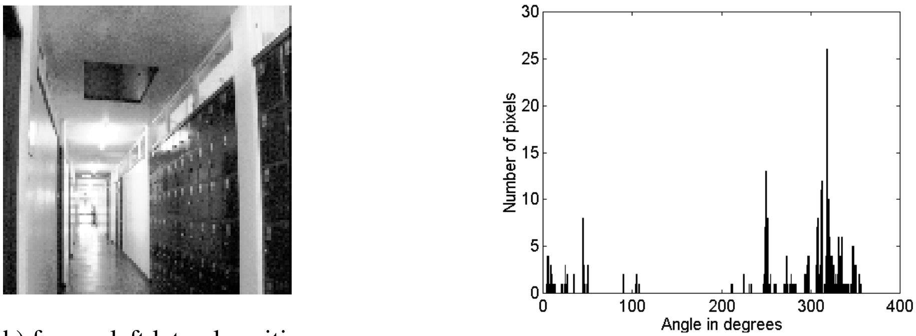
Figure 7: Values of the optimising measures for differing AVPs

3.3 Determination of lateral position

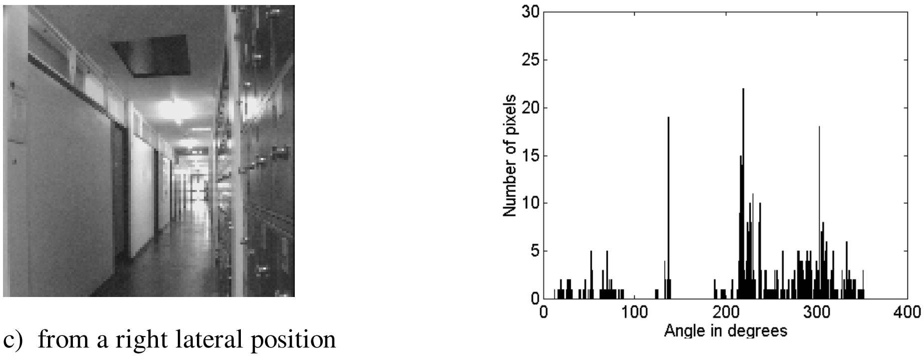
Figure 8 consists of three corridor images and the corresponding polar histograms of their lines of perspective after pre-processing as described in 2.4. It is noted that not all the expected features appear in all images due to occlusion, lack of contrast or poor definition. (a) is a copy from figure 4 where the camera is in a 'central' position in the corridor. (b) is from the extreme left position, and (c) from the extreme right position. The same pre-processing has been performed on all three images, (although it is only optimal for the 'centre' image), prior to polar transformation and polar histogram processing. The angles associated with the peaks in each histogram are tabulated (using the model in figure 2) for comparison in Table1. Those angles that are not certain due to noise etc. without further processing are greyed. The trend is seen most easily in the top-left peak which is observable in all three images, which varies from 70° with the vehicle touching the left wall, to 38° with the vehicle touching the right wall. Note: The top-right peak found at 57° is deemed unsafe despite a good signal to noise ratio. The image feature so-indicated is a line belonging to a row of lockers at that angle, and has been extracted in this instance because we have applied non-optimal pre-processing for this image, for comparison. However, the two left-hand peaks being consistent with a right lateral viewpoint are sufficient to enable us to classify that peak as invalid for lateral control, though the peak is still useful in locating the vanishing point in 3.2.



a) from a central lateral position



b) from a left lateral position



c) from a right lateral position

Figure 8: Images and polar histograms

	Angular position (θ°) relative to vanishing line			
Image	θ_4 (Bottom right LOP)	θ_3 (Bottom left LOP)	θ_2 (Top left LOP)	θ_1 (Top right LOP)
Left	46	-	70	41
Centre	57	58	52	58
Right	52/53 & 71	43	38	57

Table 1. Comparison of angles of calculated lines of perspective

4 Conclusions

In order to enable high speed, real time motion control using inexpensive sensor and sensor signal processing hardware and software, a new technique of transforming a Cartesian television image into polar co-ordinates using a vanishing point as the polar origin has been proposed. The polar histogram so-produced has been shown to be suitable for deriving a robot steering signal. Two optimising functions were considered which are accurate under low noise conditions. Under noisy conditions the 'peakiness' measure was found to outperform the variance. These measures enable the x position of the vanishing point to be determined. In addition, the method has been shown to be robust with respect to moderate errors in the assumed vanishing point position and noise, and produces an output that can be utilised to form an error signal for automatic steering control of the robot.

References

[1] D. Nair and J. Aggarwal, "Moving Obstacle Detection from a Navigating Robot", *IEEE Trans. Robot and Auto.,* Vol 14, No. 3, June 1998, pp 404-416.

[2] B. Caprile and V. Torre, "Using Vanishing Points for Camera Calibration", *Int. J. Comput. Vis.,* Vol 4, Mar 1990, pp 127-139.

[3] P. Palmer and A. Tai, "An Optimised Vanishing Point Detector", *British Mach. Vis. Conf.,* 1993, pp 529 - 538.

[4] M. Straforini, C. Coelho and M. Campani, "Extraction of Vanishing Points from Images of Indoor and Outdoor scenes", *Image Vis. Comput. J., Vol 11, No. 2, pp 91-99, 1993.*

[5] X. Lebegue and J. Aggarwal, "Significant Line Segments for an Indoor Mobile Robot", *IEEE Trans. Robot and Auto.,* Vol 9, No. 6, Dec 1993, pp 801-815.

[6] R. Schuster, N. Ansari, and A. Bani-Hashemi, "Steering a Robot with Vanishing Points", *IEEE Trans. Robot and Auto.,* Vol 9, No. 4, Aug 1993, pp 491-498.

[7] A. Tai, J. Kittler, M. Petrou and T. Windeatt, "Vanishing Point Detection", *British Mach. Vis. Conf.,* 1992, pp 109 - 118.

[8] M. Tistarelli and G. Sandini, "On the Advantage of Polar and Log-polar Mapping for Direct Estimation of Time-to-impact from Optical Flow," *Pattern Anal. and Machine Intell.,* vol. 15, no. 4, pp 401-410, 1993.