roar @UEL
research open access repository

University of East London Institutional Repository: http://roar.uel.ac.uk

This paper is made available online in accordance with publisher policies. Please scroll down to view the document itself. Please refer to the repository record for this item and our policy information available from the repository home page for further information.

Author(s): Snailum, Nic P.
Title: A new perspective on the Roberts filter
Year of publication: 2010
Citation: Snailum, N.P. (2010) 'A new perspective on the Roberts filter.', Proceedings of Advances in Computing and Technology, (AC&T) The School of Computing and Technology 5th Annual Conference, University of East London, pp.151-156.

A NEW PERSPECTIVE ON THE ROBERTS FILTER

Nic P. Snailum
University of East London
n.snailum@uel.ac.uk

Abstract: When autonomous mobile robots need to navigate by vision in indoor environments, powerful direction-finding and steering clues are provided by the lines of perspective in the image. These are neither vertical nor horizontal. In a real-time navigation task it is important to obtain this information as efficiently as possible, and is the motivation for this research. The Roberts and Sobel operators are amongst the most efficient and well known for locating edge segments which provide the information required. In this paper the outputs of the Roberts operators are compared with that of Sobel, and a subtle variant of the Roberts is introduced and compared with these.

1. Introduction:

Many operations once carried out by human operators are now performed by robots. However, the progress toward automated applications has been far slower than was predicted some decades ago, and this is largely due to the unforeseen difficulties in giving inanimate objects sufficient 'intelligence' to perform the task required. Even the simplest tasks performed by human beings become quite a challenge when programming or training a robot to perform a task robustly. The research related here is part of such a case in which the robot is required to navigate autonomously from one location to another using only vision guidance (Lebegue et al, 1993, Schuster et al, 1993, Snailum et al, 1999). There are now a number of 'robot assisted' mobile-robot products available commercially where the ultimate responsibility for safety remains with the operator, but there remain few exploited applications which can be called truly 'autonomous' mobile robots (AMR). The 'Helpmate' delivery system was perhaps the first (Evans, 1995), but many years later and still only simple applications such as vacuum cleaners and lawnmowers have followed. These must be responsible for both the decision to move and also deciding the direction in which to move. In this research it is assumed that a higher intelligence has determined that the robot must move. Also the problem of localisation is not addressed, so the task here is reduced to recognising sufficient features from the environment to *allow* movement to occur accurately and consistently. It has been observed that much of the office environment may be viewed as open spaces connected by networks of corridors so that much of the navigation problem can be solved by corridor following behaviour (Horswill, 1993). These man-made environments contain visual clues which are generally seen with three main alignments: the horizontal (x), vertical (y) and depth (z). Although the x and y dimensions may be used for short-distance manoeuvring, it is the depth dimension that is particularly of interest for navigation, human or robot, as this provides long-distance 'steering' information. Visual clues for depth were extracted from a single image (Saxena et al, 2009), and from a video sequence (Zhang et al 2009).

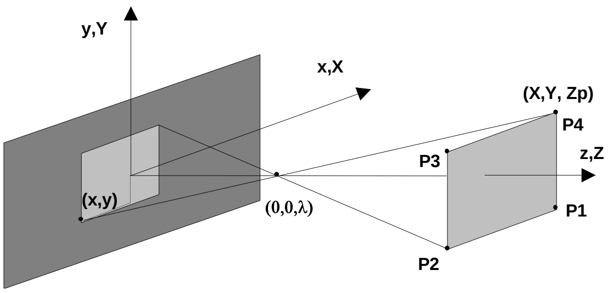


Figure 1(a) Pinhole model

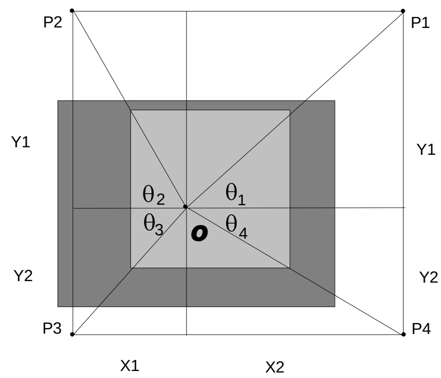


Figure 1(b) Projection of lines of interest into the image

2. Perspective:

A 'pinhole' perspective model is used, 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. Perspective was employed in locating vehicles on the road ahead (Sun et al 2006), and was used to successfully navigate an AMR in a corridor environment (Snailum, 2001). Using such a model, features of interest may be projected into the image, as shown in Figures 1(a and b).
The features of interest are the lines of perspective, $P_1 0$, $P_2 0$, $P_3 0$ and $P_4 0$, which represent depth in the image and 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, a distance 'Zp' from the camera.

3. Image Filters:

A glance through most image processing publications reveals a plethora of line or edge detection filters. The choice in many applications is limited by the context. If greater accuracy is required and time is not a constraint, as in off-line processing, the Canny detector may be usefully employed. For real-time analysis of images in an autonomous vehicle it is usually necessary to minimise the processing time of each image to allow a higher throughput of images, and hence more accurate vehicle control. With speed of processing in mind, the choice is limited to filters of small mask size which require fewer mathematical operations (particularly multiplications) per frame; and of these the most frequently used masks are typically 3x3 pixels, as in Prewitt, and Sobel. The mask coefficients for these varieties are shown in Figure 2. These masks approximate the first derivative of the image, similar to $\frac{\partial f}{\partial x}$ defined in equation (1).

$$\frac{\partial f}{\partial x} = f(x+1) - f(x) \quad (1)$$

3x3 masks in general have the requirement for 9 multiply accumulate (mac) operations per image location, and are thus more efficient than the equivalent 5x5 mask which require 25.

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix} \qquad \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$$

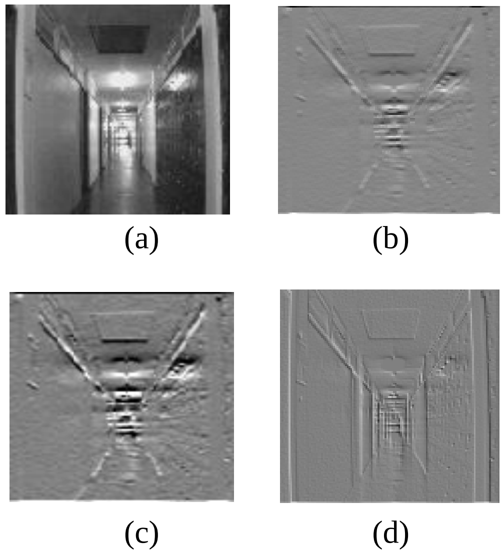
(a) (b)

$$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \qquad \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

(c) (d)

Figure 2 (a) Prewitt, (b) Sobel,
Roberts masks (c) 45˚, (d) 135˚

The zero and unity coefficients in the illustrated masks reduce the processing further. Each of the varieties shown must be orientated to detect the edge direction of choice. The pair of orthogonal Roberts masks (Roberts, 1965) shown in Figures 2(c) and (d) are amongst the first edge operators devised and have only 2x2 coefficients. These are therefore more efficient and correspondingly quicker to execute; again the mask orientation must be chosen with the edge direction in mind. However, the penalty of the smaller mask, is that of poor noise immunity, as illustrated in Figure 3(d). It is noted that although the Sobel output may at first appear noisier, after a suitable thresholding operation to locate meaningful edges, the output using this mask will be less noisy due to the strong response to the lines of interest. It is further noted that lines of one-pixel width in the original (upper diagonals) produce a double response from all three mask types, whereas true edges (ceiling panel) produce only the single response. This is due to the phenomenon of a line appearing as two edges, one positive to negative and one negative to positive. After suitable thresholding to produce a binary image, it would be usual to 'thin' the output yielding more accurate edges, but which is often costly in processing time. Subsequent operations are dependent on the application.



(a) (b) (c) (d)

Figure 3 (a) Original, (b) Prewitt (horizontal)
(c) Sobel (horizontal) (d) Roberts 45˚

4. A Roberts Variant:

It was noted that the 'useful' lines for navigation are those representing depth, which appear as diagonals (Straforini et al, 1993). Each of the masks illustrated so-far is able to detect these lines. Both Sobel and Prewitt also detect horizontal lines (which is their main purpose), and to isolate the lines of interest requires further processing. The Roberts locates only one diagonal, requiring application of the orthogonal pair to detect both.

A variant on the Roberts (RV) mask has been proposed which simplifies and improves the efficiency of the processing required to locate the diagonal lines. In one operation it locates both diagonal lines and has only a small response to horizontal or vertical lines, thus eliminating the need for further processing to remove them. The mask is shown in Figure 4.

4.1 Response to Test Images:

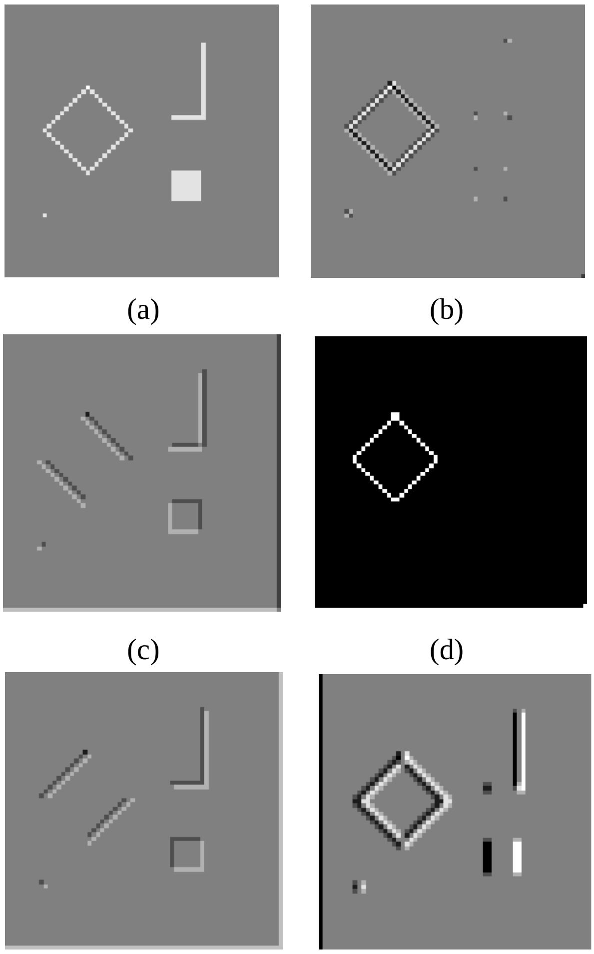
The originals (a) of Figures 5 and 6 were used to assess the response (b) of the Roberts Variant (RV) mask. Figure 5(a) is a clean image containing no added noise, whereas Figure 6(a) includes Gaussian noise of mean zero and standard deviation 0.001. In each case the filtered image has been thresholded and is provided in (d). For comparison, the response of the Roberts orthogonal pair is provided in (c) and (e), and a Sobel response in (f). The input images have a mid-grey background, with foreground objects either lighter or darker. Some of the output images also have a mid-grey background in order to be able to show both positive and negative responses of the masks.

$$\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$$

Figure 4 The Roberts Variant mask

The output images with a black background have been thresholded. It is clear that the RV mask passes diagonal lines and rejects both horizontal and vertical lines and edges. The response in Figure 5(b) is instructive, showing: the response of the RV mask to a single input pixel having four output pixels (two positive and two negative) as expected from the mask coefficients in Figure 4; the rejection of the single-pixel width vertical and horizontal lines, but exhibiting a response at the ends of the lines and at the corner; the response to the single-pixel width diamond shows the positive and negative responses on alternate sides. Only the larger values have passed the threshold in (d). Finally the Sobel response is provided for comparison in (f).

Much of the detail seen in Figure 5 has been lost in Figure 6 with the addition of noise to the original, but the passing of diagonal lines and rejection of horizontal and vertical lines is apparent. The original also has 'missing' foreground pixels in the diagonals to model a poor quality input image more realistically. Again, in Figure 6(f) the Sobel response is shown for comparison, after thresholding.



(a) (b) (c) (d) (e) (f)

Figure 5 Test image one
(a) Original (b) Roberts Variant response
(c) Roberts 45˚ (d) Thresholded (b)
(e) Roberts 135˚ (f) Sobel

4.2 Response to Real Images:

The images presented in Figures 7 and 8 show a typical corridor environment encountered by an AMR. The lines of perspective between floor and wall are easily discernible, but those between ceiling and wall are less so.

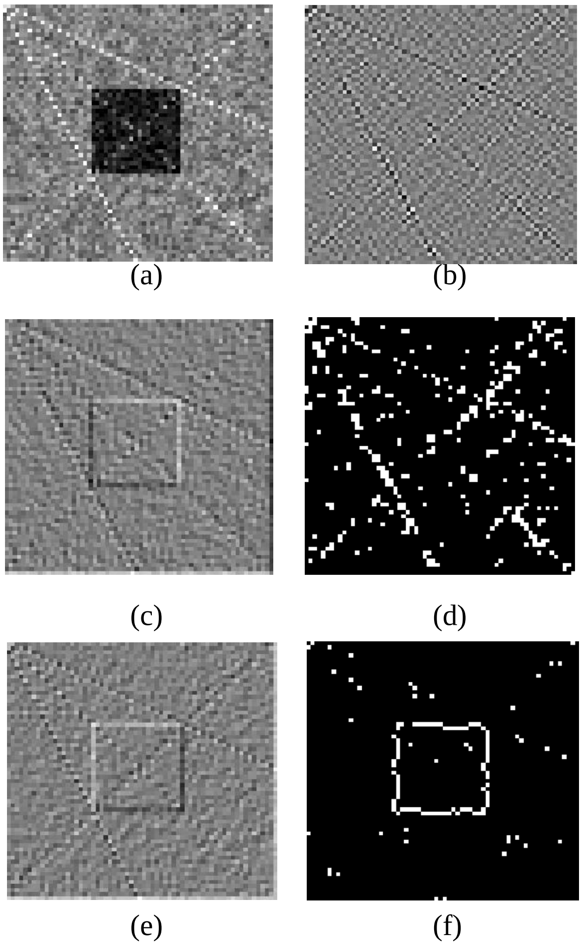


Figure 6 Test image two (a) Original (b) Roberts Variant response(c) Roberts 45˚ (d) Thresholded (b)(e) Roberts 135˚ (f) Sobel thresholded

Reflections from shiny surfaces and 'flaring' due to high contrast in low-light conditions illustrate some of the problems well-known to the image processing community. The effects of added noise to such a real image are illustrated. In Figure 8, the original includes Gaussian noise of mean zero and standard deviation 0.0015. A comparison is provided between the Roberts and the Roberts Variant responses. Note that Figures 8(b) and (c) have enhanced contrast for ease of perception. Even at modest noise levels it is seen that the response of the Roberts Variant degrades, and renders the mask unusable in poorer lighting conditions which would produce such noise. The original Roberts also degrades with noise (a well known phenomenon), but not so quickly as the variant due to the stronger response to uni-directional lines. The variant has a response to both diagonal directions but with less strength for each due to its ability to detect 'corners' as shown in Figure 5(b). In a noisy image, noise appears as a series of corners each of which has a response, thus reducing the apparent response to the diagonals.

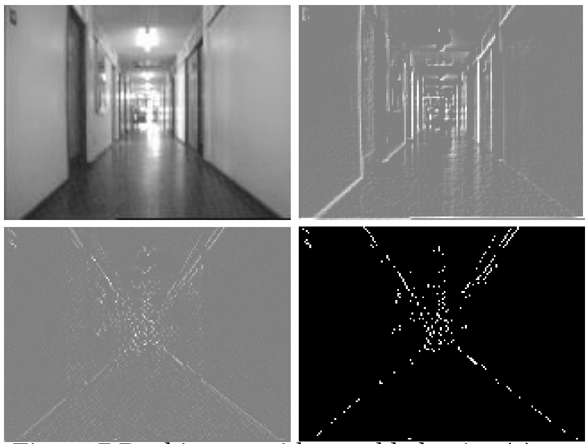


Figure 7 Real image with no added noise (a) Original (b) Roberts 135˚(c) Roberts Variant (d) Thresholded (c)

5. Conclusion:

The lines of perspective which are useful for steering autonomous mobile robots may be extracted from images of rooms and corridors by applying a variety of well-known masks. The Roberts and Sobel are used as references, whilst testing a variant of the Roberts mask which appears to give some efficiency advantages. The Roberts mask is a 2x2 element mask and is inherently efficient, and as its coefficients are unity, its application only involves addition and subtraction and no multiplication. Instead of applying the pair of Roberts orthogonal masks in two separate operations, it is found that by applying the proposed variant in a single operation, a similar result is obtained more efficiently. The response of the variant may be regarded as a 'corner detector' thus detecting diagonal but not vertical or horizontal lines, and this makes it particularly useful when attempting to recover lines of perspective from an image. It is noted that the resultant output is more noisy than the outputs obtained after application of the two Roberts masks, and that it degrades quickly with additive noise as does the original Roberts. It is often difficult to quantify the results of such operations, and visual outputs only have been provided. Further work will attempt this quantification so that a better understanding of relative performances may be obtained.

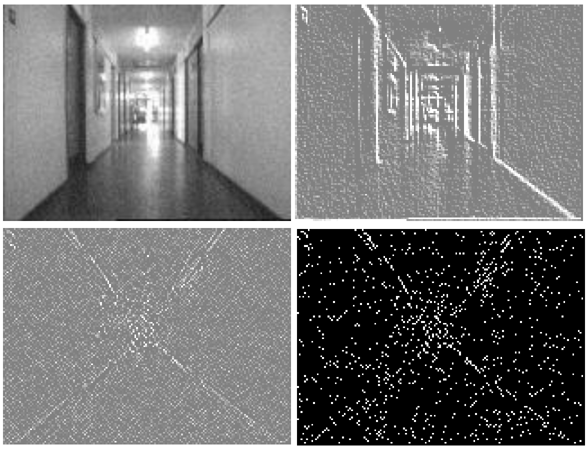


Figure 8 Real image with added noise
(a) Original (b) Roberts 45˚
(c) Roberts Variant (d) Thresholded (c)

6. References:

Evans J., "Helpmate: A Service Robot Success Story," *Service Robot,* Vol. 1, no. 1, 1995, pp 19-21.

Horswill I., "Polly: A Vision-Based Artificial Agent," *Proc. of the 11th Nat. Conf. on A.I,,* California, 1993.

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

Roberts L.G., "Machine perception of 3-D Solids", *in J. T. Tippett (Ed.) Optical and Electro-Optical Information Processing,* MIT Press, Cambs., Massachusetts, 1965.

Saxena A., Sun M., Ng A. Y., "Make3D: Learning 3D Structure from a Single Still Image", *IEEE Trans. Pattern Analysis & Machine Intelligence,* Vol 31, no.5, May 2009.

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

Snailum N., Smith M.C.B., Dodds S., "Vision Based Mobile Robot Navigation in a Visually Cluttered Environment", *Proc. Towards Intelligent Mobile Robots, Bristol,* 1999.

Snailum N., "Mobile Robot Navigation using Single Camera Vision", *Thesis,* 2001.

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

Sun Z., Bebis G., Miller R., "On-Road Vehicle Detection: A Review", *IEEE Trans. Pattern Analysis & Machine Intelligence,* Vol 28, no.5, May 2006, pp 694-711.

Zhang G., Wong T., "Constructed Depth Maps Recovery from a Video Sequence", *IEEE Trans. Pattern Analysis & Machine Intelligence,* Vol 31, no.6, Jun 2009