Real-Time Regular Polygonal Sign Detection

Nick Barnes[1] and Gareth Loy[2]

[1] National ICT Australia, Locked Bag 8001, Canberra, ACT 2601,
Department Of Information Engineering, The Australian National University
Nick.Barnes@nicta.com.au

[2] Computer Vision and Active Perception Laboratory, Royal Institute of Technology (KTH), Stockholm, Sweeden
gareth@nada.kth.se

Summary. In this paper, we present a new adaptation of the regular polygon detection algorithm for real-time road sign detection for autonomous vehicles. The method is robust to partial occlusion and fading, and insensitive to lighting conditions. We experimentally demonstrate its application to the detection of various signs, particularly evaluating it on a sequence of roundabout signs taken from the ANU/NICTA vehicle. The algorithm runs faster than 20 frames per second on a standard PC, detecting signs of the size that appears in road scenes, as observed from a camera mounted on the rear-vision mirror. The algorithm uses the symmetric nature of regular polygonal shapes, we also use the constrained appearance of such shapes in the road scene to the car in order to facilitate their fast, robust detection.

1 Introduction

Improving safety is a key goal of road vehicle technology. Driver support systems aim to improve safety by helping drivers react to changing road conditions. Although full automation of road vehicles may be achievable in the future, our research focusses on systems that can assist drivers immediately. Rather than replacing the driver, we aim to keep the driver in the loop, while supporting them in controlling the car.

Road signs present information to alert drivers to changes in driving conditions. Critical information signs, such as speed, give-way, roundabout, and stop give information that a driver must react to, as opposed to informational and directional signs. These signs appear clearly in the road scene, and are well distinguished. However, drivers may sometimes miss such signs due to distractions or a lack of concentration. This makes detecting critical information signs and making the driver aware of any they may have missed a key target for improving driver safety. The lack of driver awareness of a sign may be detected through a lack of response.

We have previously demonstrated the application of the radial symmetry operator [1] to detecting speed signs, demonstrating real-time performance [2].

P. Corke and S. Sukkarieh (Eds.): Field and Service Robotics, STAR 25, pp. 55–66, 2006.

Here detection took advantage of the circle that must appear on Australian speed signs. The radial symmetry algorithm is a shape detector, based on image gradient, and so is robust to varying lighting conditions and occlusions, or the incomplete appearance of edges due to sign degradation under weather conditions. This detection algorithm was coupled with recognition to allow a full system that reported the current speed limit to the driver in real-time.

The regular polygon detector is a more general algorithm than radial symmetry that is able to detect all regular polygons including triangular, square, and octagonal shapes [3]. General regular polygons may be detected in an image up to a similarity transform. All the properties of the radial symmetry algorithm are maintained in terms of robustness to illumination changes, occlusion, and poor edges. In this paper, we adapt this algorithm specifically to real-time road sign detection. By only examining gradients that could be part of the road sign being sought, we are both able to simplify the detection process without significantly impacting robustness. This allows the detection to run at less than 50 ms per image on a standard PC installed in the vehicle. We demonstrate this performance on a real road sequence, showing robust and effective detection.

This class of regular polygon detectors has complexity $\mathcal{O}(Nkl)$, where l is the maximum width of the segments of the shape, k is the number of scales that are being considered, and N is the number of image pixels. Note that for small shapes k and l are small numbers. The generalised Hough transform [4] is a general algorithm that can perform the same function, however, even modern hardware-based implementations take multiple seconds to recognise a single shape [5]. This improved computational performance is the benefit of specialised algorithms for detection of known features.

Other work in perceptual grouping [6] takes a similar approach in terms of finding local support for shapes. However, this work does not use pixel-based gradient information, and works at the level of edge segments for gradient. A complexity of k^2 was reported where k is the number of edge segments, however for a cluttered image, if the segment size is one pixel, k may be of order of the size of the image.

The regular polygon shape detector class of algorithms is specialised for real-time performance by exploiting the nature of regular polygon-like shapes that have a defined centre point, using it as the voting centre. This makes the algorithm robust to pixels that are missed due to occlusion, poor gradient direction estimates or broken edges, as the vote at the centre point will still be high. The regular polygon detectors are parametric in the formulation, and can be easily and efficiently applied in situations where constraints are available from the embodiment of the vision system within, such as the appearance of a road sign to a car. In this paper, we focus specifically on the aspect of adapting the detector for road sign detection.

This detector improves on previous sign detection results by facilitating fast, robust detection that can reduce the region of interest for sign recognition to only a few pixels, returning the position, centroid, scale and shape of candidate signs. Thus, few pixels require examination for recognition, and the sign size and shape parameters are known. This allows computationally expensive recognition as processing is well-targetted, so comparatively little computation is needed to assess a candidate. The enhancements detailed in this paper reduce the number of false positive sign detections by ignoring all gradient edges that are not at an angle that could be part of a valid road sign.

2 Background

Road sign recognition research has been around since the mid 1980's. A direct approach is to apply normalised cross correlation to the raw traffic scene image. This brute force method is computationally prohibitive, but can be eased somewhat by approaches such as simulated annealing [7], or applying templates to an edge image of the road scene [8]. However, these methods are still computationally intensive and so unsuitable for a real-time system.

Many approaches have introduced separate stages for sign detection and classification of sign type (e.g., [9, 10, 11]). We argue that this is an effective means of managing computation for even a small number of sign types if the detection algorithm has low computational cost, facilitating real-time operation. In this manner, computationally intensive classification is only performed on a small fraction of the input stream.

Colour segmentation is the most common method for detection. Typically, this is based on the assumption that the wavelength that arrives at the camera from a traffic sign is invariant to the intensity of incident light. This assumption usually manifests in the statement that HSV (or HSI) space is invariant to lighting conditions [11]. A great deal of the research in this area exploits a detection stage based on this assumption (e.g., [11, 12, 9, 13, 14, 15, 16]), either finding the signs, or eliminating much of the image from further processing. However, the camera image is *not* invariant to changes of incident light chromaticity. Such variance occurs under different conditions, such as direct sunlight, heavy cloud, smog, and under headlights at night. Further, the colour changes as signs fade over time.

Another approach to detection is *a priori* assumptions about image formation. At its simplest, one can assume that the road is approximately straight, so much of the image can be ignored. Combining such assumptions with colour segmentation, Hsu and Huang [16] look for signs in only a restricted part of the image. However, such assumptions can break down on curved roads, or if the road is not planar. A more sophisticated approach is to use some form of detection to facilitate scene understanding, and thus eliminate a large region of the image. For example, Piccoli *et. al* [12] suggest large uniform regions of the image correspond to the road and sky, and thus signs are only likely to appear in the region alongside the road and below the sky. However, this is inadequate in more difficult road scenes, for example a scene with trees overhead casting shadows across the road. They also suggest ignoring one side of the image as signs will only come up on one side, however, this is not always the case, for example see Figure 3, where the sign is in the centre of the road.

In this paper we present an adaptation of the regular polygon detection algorithm to real-time detection of road signs. The detector uses gradient elements to detect signs based on shapes, detecting triangle, square, diamond, octagonal, and circular signs. The adaptation of the regular polygon detector that we present can detect all signs in an image in 50 ms at all the sizes that are practical for our driver assistance system vehicle platform.

3 The Structure of the Road Environment

There is much possible variation in sign appearance. Throughout the day, and at night time, lighting conditions can vary enormously. A sign may be well lit by direct sunlight, or headlights, or on the other hand it may be completely in shadow on a bright day. Further, heavy rain may blur the image of the sign. Ideally, signs have clear colour contrast, but over time they can become quite faded, but are still quite clear to drivers. Although signs appear by the road edge, this may be far from the car on a multi-lane highway — to the left or right, or very close on a single lane exit ramp — and whilst signs are generally a standard distance above the ground, they can also appear overhead or on temporary roadworks signs at ground level. With this nature, it is not simple to restrict the possible viewing positions of a sign within the image. By modelling the road [17], it may be possible to dictate parts of the image where a sign cannot appear, but road modelling has its own computational expense, and, as discussed previously, colour-based methods are not robust. However, the roadway is well structured. Critical signs, including give way, stop and roundabout signs have a highly regulated appearance. Unless the sign has been tampered with, signs will appear upright and approximately orthogonally to the road direction. Finally, signs are always placed to be easily visible, so the driver can easily see them without having to look away from the road. In this paper we assume that the signs are correctly oriented, however, with minor accidents, signs can occasionally appear tilted.

The regular polygon algorithm searches for near regular polygonal features. As a legal critical information signs must always appear orthonormal to the direction of the road, provided the system camera points in the direction of vehicle motion, the surface of signs will be approximately parallel to the image plane and so the polygonal shape of these signs will appear undistorted. On a rapidly curving road it may be that the sign only appears parallel to the image plane briefly, but this will be when the vehicle is close to the sign, so it will appear large in the image. If we are processing images at 20 Hz, and we are able to recognise a sign reliably from only a small number of frames then generally we are safe to assume that the sign will be approximately parallel to the image plane for many images.

4 Adapting Regular Polygon Detection to Real-Time Sign Detection

Detecting regular polygons has been shown to be a powerful means of locating triangular, square and octagonal road signs in images [3]. The method used was an extension of the radial symmetry transform [1], and utilised the intrinsic symmetry of equiangular regular polygons to detect these shapes independently of orientation. However, in the case of road sign detection the orientation of the target is known *a priori*, and thus orientation independence is not required. This section demonstrates how the introduction of an orientation constraint can be applied to the algorithm to: reduce the complexity; and increase the speed to real-time levels.

For the sake of completeness this section summarises the base part of the algorithm described in [3]. This sets the basis for the adaptation to road scenes and the real-time implementation described in the remainder of the section.

4.1 Review of Regular Polygon Detection

The existing algorithm detects the centroids of n-sided regular polygons in greyscale images as follows. Firstly the image gradient vector field is computed, all elements with magnitudes under a predefined threshold are set to zero, and the resulting vector field is denoted $\mathbf{g}$. For each radius under consideration, a vote image O_r is computed of the same size as the input image, and initialised to zero. Each non-zero element of $\mathbf{g}$ is considered in scan-line order, and the shape centroid locations voted for and against by this gradient element are computed. For a single gradient element $\mathbf{g}(\mathbf{p})$ these are given:

$$L_+ = \{L(\mathbf{p}, m) | m \in [-w, w]\}, \tag{1}$$

$$L_- = \{L(\mathbf{p}, m) | m \in [-2w, -w) \cup (w, 2w]\}, \tag{2}$$

respectively. Here $L(\mathbf{p}, m)$ describes a line of pixels in front of and behind the gradient element a distance r away, as shown in Figure 1, and is given by:

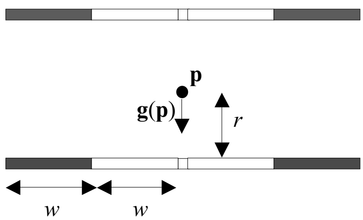


Fig. 1. Lines of pixels voted for by gradient element $\mathbf{g}(\mathbf{p})$, light lines indicate positive votes, dark lines negative votes.

$$L(\mathbf{p}, m) = \mathbf{p} + \text{round}\left(m\bar{\mathbf{g}}(\mathbf{p}) \pm r\mathbf{g}(\mathbf{p})\right), \tag{3}$$

where $\bar{\mathbf{g}}(\mathbf{p})$ is a unit vector perpendicular to $\mathbf{g}(\mathbf{p})$, m is an integer and

$$w = r \tan \frac{\pi}{n}. \tag{4}$$

Figure 2 shows the different voting lines for the same gradient element when searching for different shapes, the white lines show positive votes L_+ and the dark lines are the negative votes L_-. The negative votes are included to attenuate the response of straight lines too long to belong to the target shape. As each significant (non-zero) element of $\mathbf{g}$ is considered the total votes for each pixel are accumulated in the vote image $\mathbf{O}_r$.

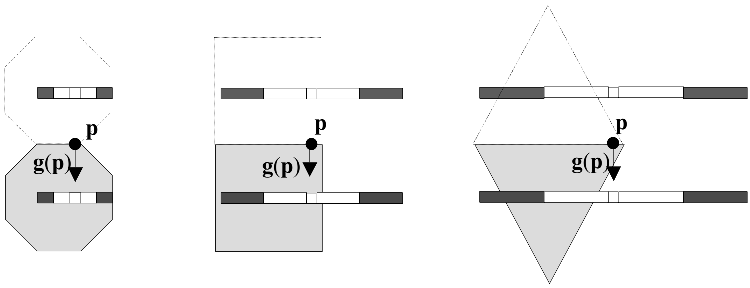


Fig. 2. Voting lines associated with a gradient element for different shapes.

In [3] a second vote image is constructed called the n-angle image. This votes for angles to ensure that the shapes detected have the required number, and does this in such a way as to facilitate the detection of these shapes at any orientation simultaneously. The n-angle image is a key component and accounts for a substantial portion of the computational time required. However, it is not required if the orientation of the target is known *a priori*.

4.2 Detecting Regular Polygons of Known Orientation

If the orientation of the targeted shape is known, the search can be targeted much more specifically, and a number of speedups and simplifications can be applied to both reduce the complexity and increase the performance of the method originally proposed in [3].

The orientation together with the number of sides of a regular polygon provides a constraint on the angles of gradient elements that are likely to occur at the edges of the shape. Specifically, for an n-sided figure with orientation θ, the gradient orientations likely to occur around this shape are given by

Fig. 3. Images from the start (52), middle (75), and end (99) of the sequence showing the roundabout sign used in this paper.

$$\angle \mathbf{g}(\mathbf{p}) = \left(\theta + \frac{2k\pi}{n} \pm \epsilon\right) \bmod \pi, \text{ for } k \text{ an integer,} \tag{5}$$

where ϵ is the angular tolerance on detected shape orientation with respect to that expected. This allows for some variation in the angle of real signs, and for some error in gradient values. Note that ϵ was actually a few degrees. Note that this method is equivalent to the previous algorithm, with *a priori* information replacing the second vote image to apply joint orientation selectivity. It will also remove any triangular candidates that are not oriented as signs, and so remove this noise from the image.

Only gradient elements satisfying the orientation constraint in Equation 5 need be considered when computing O_r. This both reduces the computation required to compute O_r and implicitly performs the same task that the n-angle image did in the unconstrained case, namely attenuating responses for votes whose angular spacing is not consistent with that of the target shape.

4.3 Embodied Vision Issues

In practice, we process a continous image stream, not single images. Between two images at 20 fps, only slight motion will occur. Accidental features are common, where two image features align coincidentally, such as alignment of multiple edges to form a regular polygon. We may partly deal with this issue by requiring that a detected regular polygon is present for t images, does not move far between frames, and that its radius does not decrease in size over time, and only increases by an amount appropriate for the vehicle speed. Further, as there will be artefacts at multiple radii for a single sign, we apply non-maximal suppression.

Further, the size that the sign will appear in the camera image mounted in the car is constrained by the closest distance that it will be from the car, and the camera focal length. There is also no point in detecting signs that are smaller than the minimum size required to classify the sign.

Fig. 4. Inside the ANU/NICTA intelligent vehicle. Cameras to monitor the road scene appear in place of the rear-vision mirror.

5 Real-Time In-Vehicle Implementation

We integrated regular polygonal sign detection into the ANU/NICTA Intelligent vehicle, see Figure 4. The algorithm was implemented in C++. The maximum radii of the largest circle that could be inscribed on the triangle was 12 pixels, and four radii was sufficient for the visible range of give way signs as they appeared from the vehicle. For a 320 x 240 image, the detection of four radii algorithm was able to run at 20 fps, running in less than 50 ms per frame, on a 3Mhz dual processor PC. This is a vast improvement over the approximately 1 fps performance of the original algorithm [3]. In previous work we found that classification using normalised cross correlation was able to run in less than 1ms per candidate, and so there is sufficient time left for each frame to classify several candidates. As discussed in [18], the threshold of gradients was set to only take the top 20% of the gradients in the image, improving the signal to noise ratio and enhancing computational speed.

6 Results

Previously, we ran the detector on a series of signs taken from the internet with the general algorithm. Images from this set were rerun with the new adaptation of the algorithm to demonstrate that it has not lost any generality for road scenes. The images and their regular polygon transform for the correct radius are shown in Figure 5. The results are quite comparable with the full algorithm, and are less noisy due to the exclusion of gradients that are irrelevant to the constrained orientation case of road signs.

We took a real image sequence from the cameras mounted in the vehicle. The sequence consists of 100 images. The sign is large enough to be classified starting from frame 52, until it falls outside the image in frame 99. Over this period the vehicle was travelling at approximately 60 kmph. This sequence was run through the complete system without any specific adjustments to the sequence itself. The first, last, and mid images in the sequence are shown

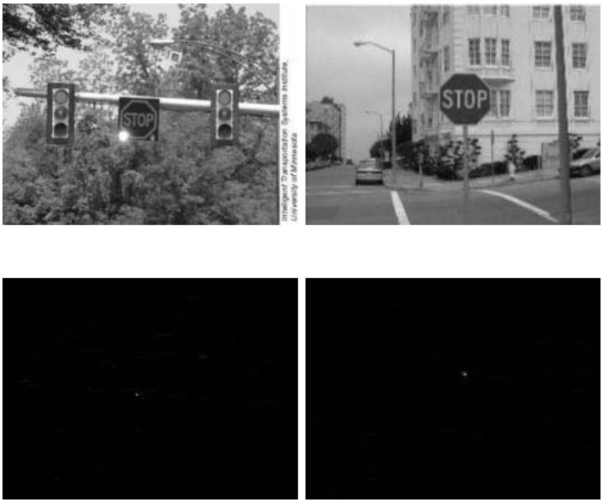
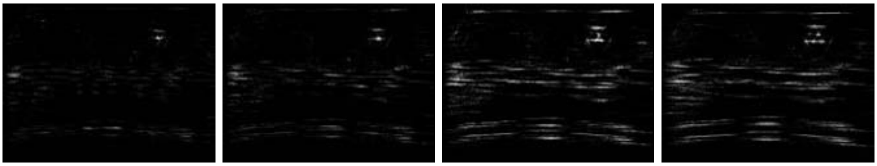


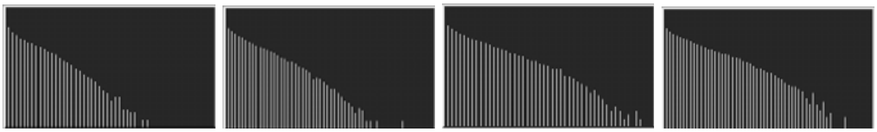
Fig. 5. Images of two stop signs from the internet, and corresponding regular polygon transform image using the algorithm presented for radius 20, and 28 respectively. Although not completely clear to eye in this reproduction, in both cases the sign is detected as the highest peak in the image.

in Figure 3. Figure 6 shows the vote images that were produced for the mid image in the sequence. At the position of the sign, the increasing radius can be seen as the lines produced by voting move towards, and then past the centre of the sign, the peak is across (b) and (c). For this image, the best radius was at (232,52) from top left. For this sequence, we took the best candidate in each frame for each radius, and checked if it was in the best ten candidates at each radius for the next t images. For the sequence, the frame by frame results for different t are shown in Figure 7. Figure 8 shows a summary of the number of times the sign was correctly detected during the sequence, and the number of false positives, given different values for t. Regardless of the detection rates over the sequence, it can be clearly seen that the sign is easily detected many times over the sequence. Also, with this small number of false positives, the recognition part of the system will be required to investigate a maximum of four image locations in any frame. Figure 9 shows the false positive rates for the images where the sign was to small to be detected. We expect that this can be improved by computationally simple post-processing of detected candidates. In initial trials with normalised cross correlation-based recognition, the sign was recognised 17 times in the sequence. Although this shows the sign being recognised many times while it is visible, the literature would suggest better recognition rates can be achieved with tuning.

In previous work round signs [2], we found that false positive repeat signs were rare for circular features. As can be seen from Figure 8 this was less the



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



(e) (f) (g) (h)

Fig. 6. The regular polygon detection image for image 76 shown above, for radii of (a) 6, (b) 8, (c) 10, and (d) 12. The images underneath show vote histograms. (b) and (f) show the correct radius with a single clear peak of votes at the sign, with few other candidates with a close number of votes.

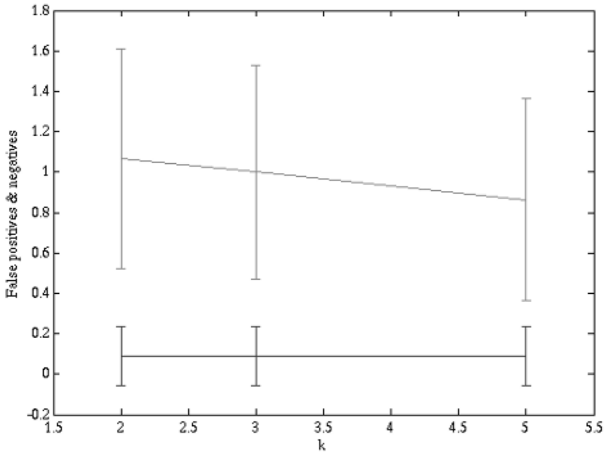


Fig. 7. Average and half standard deviation of false positives (top) and false negatives (bottom), given number of frames t that the sign is required to be present. Note that the error bars are set to one half of the standard deviation for clarity.

case for triangular signs. This is to be expected, as it is not uncommon for a multiple edges to align in an urban road-scene. However, as the number of expected edges increases towards the limit of a circle they will become less common as this requires more coincidentally placed edges.

numframes	times correct	false positives
2	44	47
3	44	44
4	44	38
5	44	37

Fig. 8. For the number of frames the sign is required to be present, the number of correct classification and the number of false positives across the whole 48 frames where the sign was detectable.

t number of frames	average false positives
2	3.01
3	2.80
4	2.46
5	2.26

Fig. 9. Number of false positives per frame versus t for the 52 frames where the sign was not large enough to be detected.

7 Conclusion

We have adapted the regular polygon algorithm for road sign detection. By constraining the gradient image voting, we were able to simplify the algorithm in terms of computation, and process images in less than 50 ms. The detection results over sample images show no significant loss of detection signal with respect to the original regular polygon detection algorithm. New image sequence results demonstrate stable performance at real time, and reliable detection, with sufficiently few false positives to support real-time classification.

Acknowledgements

National ICT Australia is funded by the Australian Department of Communications,Information Technology and the Arts and the Australian Research Council through *Backing Australia's ability* and the ICT Centre of Excellence Program. The support of the STINT foundation through the KTH-ANU grant IG2001-2011-03 is gratefully acknowledged.

References

1. G. Loy and A. Zelinsky, "Fast radial symmetry for detecting points of interest," *IEEE Trans Pattern Analysis and Machine Intelligence*, vol. 25, no. 8, pp. 959–973, Aug. 2003.
2. N. Barnes and A. Zelinsky, "Real-time radial symmetry for speed sign detection," in *Proc IEEE Intelligent Vehicles Symposium*, Parma, Italy, 2004.

3. G. Loy and N. Barnes, "Fast shape-based road sign detection for a driver assistance system," in *Proceedings of the 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems.IROS2004*, 2004, in press.
4. D. H. Ballard, "Generalizing the hough transform to detect arbitrary shapes," *Pattern Recognition*, vol. 13, no. 2, pp. 111–122, 1981.
5. R. Strzodka, I. Ihrke, and M. Magnor, "A graphics hardware implementation of the generalized hough transform for fast object recognition, scale and 3d pose detection," in *Proc 12th International Conference on Image Analysis and Processing*, Mantova, Italy, 2003, pp. 188–193.
6. G. Guy and G. Medioni, "Inferring global perceptual contours from local features," *International Journal of Computer Vision*, vol. 20, no. 1-2, pp. 113–33, Oct. 1996.
7. M. Betke and N. Makris, "Fast object recognition in noisy images using simulated annealing," A.I. Lab, M.I.T., Cambridge, Mass, USA, Tech. Rep. AIM-1510, 1994. [Online]. Available: citeseer.nj.nec.com/betke94fast.html
8. D. M. Gavrila, "A road sign recognition system based on dynamic visual model," in *Proc 14th Int. Conf. on Pattern Recognition*, vol. 1, Aug 1998, pp. 16–20.
9. P. Paclik, J. Novovicova, P. Somol, and P. Pudil, "Road sign classification using the laplace kernel classifier," *Pattern Recognition Letters*, vol. 21, pp. 1165–1173, 2000.
10. J. Miura, T. Kanda, and Y. Shirai, "An active vision system for real-time traffic sign recogntition," in *Proc 2000 IEEE Int Vehicles Symposium*, Oct 2002, pp. 52–57.
11. L. Priese, J. Klieber, R. Lakmann, V. Rehrmann, and R. Schian, "New results on traffic sign recognition," in *Proc. Intelligent Vehicles Symposium*. Paris: IEEE Press, Aug. 1994, pp. 249–254. [Online]. Available: citeseer.nj.nec.com/priese94new.html
12. G. Piccioli, E. D. Micheli, P. Parodi, and M. Campani, "Robust method for road sign detection and recognition," *Image and Vision Computing*, vol. 14, no. 3, pp. 209–223, 1996.
13. B. Johansson, "Road sign recognition from a moving vehicle," Master's thesis, Centre for Image Analysis, Sweedish University of Agricultural Sciences, 2002.
14. C. Y. Fang, C. S. Fuh, S. W. Chen, and P. S. Yen, "A road sign recognition system based on dynamic visual model," in *Proc IEEE Conf. on Computer Vision and Pattern Recognition*, vol. 1, 2003, pp. 750–755.
15. D. G. Shaposhnikov, L. N. Podladchikova, A. V. Golovan, and N. A. Shevtsova, "A road sign recognition system based on dynamic visual model," in *Proc 15th Int Conf on Vision Interface*, Calgary, Canada, 2002.
16. S.-H. Hsu and C.-L. Huang, "Road sign detection and recognition using matching pursuit method," *Image and Vision Computing*, vol. 19, pp. 119–129, 2001.
17. R. Labayrade, D. Aubert, and J.-P. Tarel, "Real time obstacle detection in stereovision on non flat road geometry through v-disparity representation," in *Proc IEEE Int Vehicles Symposium*, June 2002.
18. N. Barnes, "Improved signal to noise ratio and computational speed for gradient-based detection algorithms," in *Proc. IEEE Int. Conf. on Robotics and Automation (ICRA '96)*, Barcelona, Spain, April 2005.