

Regular polygon detection

Nick Barnes₁, Gareth Loy₂, David Shaw₁, Antonio Robles-Kelly₁

1. National ICT Australia, Locked Bag 8001, Canberra, ACT, 2601, Australia,

Dept. Information Engineering, The Australian National University.

2. Computer Vision & Active Perception Laboratory, Royal Institute of Technology (KTH), Stockholm

Abstract

This paper describes a new robust regular polygon detector. The regular polygon transform is posed as a mixture of regular polygons in a five dimensional space. Given the edge structure of an image, we derive the a posteriori probability for a mixture of regular polygons, and thus the probability density function for the appearance of a mixture of regular polygons. Likely regular polygons can be isolated quickly by discretising and collapsing the search space into three dimensions. The remaining dimensions may be efficiently recovered subsequently using maximum likelihood at the locations of the most likely polygons in the subspace. This leads to an efficient algorithm. Also, the a posteriori formulation facilitates inclusion of additional a priori information leading to real-time application to road sign detection. The use of edge orientation information also reduces noise compared to existing approaches such as the generalised Hough transform. Results are presented for images with noise to show stability. Also, the detector is applied to two separate applications: real-time road sign detection for on-line driver assistance; and, feature detection, recovering stable features in rectilinear environments.

1 Introduction

In scenes containing manufactured artefacts, features often appear as regular polygons. In road scenes, triangular, rectangular and octagonal signs display critical information. Polygonal shapes appear frequently in the structure of office environment. Corners are common partial polygons, but over a large range of scales regular polygons represent much more indoor structure. In outdoor scenes, brickwork, office buildings, windows, etc., all exhibit features that can be recognised repeatably by regular polygon detection.

The literature related to this topic is vast, covering line and analytical shape estimation and perceptual grouping, as well as road sign detection. There are two alternative schools of approach: model-driven approaches; and, perceptual grouping approaches that are frequently data driven. The model-driven approaches typically employ retinotopic mappings, whereby local features are detected purely locally, and pixel information is incorporated directly into

feature detection. Grouping approaches take low-level features, such as edges (but also alternatively raw edge pixels) and apply matching and clustering based methods.

In perceptual grouping, image elements are grouped according to some perceptual criteria. Several authors have investigated iterative relaxation style operators for grouping edgels. This began with Shashua and Ullman [19], which required many iterations. This was refined by Guy and Medioni [7] among others so as to be $\mathcal{O}(k^2)$ where k is the number of edgels. At a high level this problem can be composed as recovering a graph of relational arrangement. A common approach is spectral graph theory, e.g., [16, 18]. Probabilistic approaches in this area include Bayes nets [5] and combining evidence from raw edge attributes [3]. More recent approaches have used EM, e.g., [1, 4]. One of the basic algorithmic approaches here is examining pairwise relations. If the elements being examined is edge pixels then the complexity is at least $\mathcal{O}(N^2)$, where N is the image size.

Alternatively, we may fit functions to the image data. The Hough transform [8] and circular Hough transform [15] vote to an in-place representation of the shape to be detected. The circular Hough detects circle centres, and the Hough transform detects the closest line point to the origin (this mapping is in the same space in the log-Hough transform [21] in log-polar images). Each pixel ‘votes’ for each feature that it could possibly be a part of. This approach is inherently robust, as gaps, noise, and partial occlusion are ignored, but appear as a decreased strength of the feature. Such algorithms are local in their computation with each point only being registered for shapes that it can be part of. These methods have been generalised [2] to detect arbitrary shapes. The original Hough transform has a relatively simple mapping for each pixel, however, the generalisations vote into higher dimensions, and quickly become computationally intensive on serial architectures. Further, parallelisation is non-trivial due to the high dimensionality.

The Hough transform can be formulated as maximum likelihood estimation [20]. Kiryati and Bruckstein [10] formulated robust maximum likelihood line finding on a grid, assuming independent noise between points. Geyer *et al.* [6], applied a sampled likelihood function approach to estimation of the essential matrix. Makadia and Daniilidis [14] formulated a Hough-like algorithm in motion space for estimating robot position using a Radon transform formulation

with a distance-based soft characteristic function.

Variants of shape detector have also included edge orientation information. The radial symmetry detector [13] improves on the circular Hough giving $\mathcal{O}(Nl)$ performance, where l is the range of radii of the shapes being examined. We have taken a similar approach to regular polygons [12]. However, only one regular polygon type could be detected at a time and the angle vote applied to the whole image to resolve orientation was computationally expensive.

Road sign recognition research has been around since the mid 1980's. Good results may be achieved with a classification approaches such as normalised cross correlation [17]. However, this is computationally intensive per pixel. A standard approach is to first apply a low computational cost detection stage reducing classification to a fraction of the image stream. Typically this is through assumptions about scene structure, colour, or a combination of both (e.g., [17, 9]). Although these assumptions are sound for many road scenarios, the breakdown of a driver assistance system on hills and corners is not acceptable. Colour methods also break down under the enormous variation in lighting chrominance in outdoor situations, and segmenting out self similar regions is not robust in scenes with dense overhead branches that shadow the road. Instead, we propose a illumination invariant shape detection approach to find likely candidates, to be followed by recognition.

We propose a computationally efficient algorithm to recover regular polygons using an *a posteriori* probability approach taking advantage of locality and gradient information. Rather than summing support, we recover the likelihood of the probability density function for all regular polygons in an image. The algorithm breaks the task into two stages: 1) resolving likely regular polygons in a lower dimensional space; 2) taking the most likely regular polygons and resolving their final parameters in higher dimensional space. This allows us to recover all regular polygons in one computationally efficient algorithm. The algorithm is $\mathcal{O}(Nrlw)$, where N is the size of the image, l is the maximum radius length, r is the number of radii being considered, and w is the width of the operation of the distance function. w is close to one pixel for most implementations, and r and l tend to be small numbers leading to a more efficient algorithm. We demonstrate that in constrained cases this can lead to real-time implementations.

2 Regular polygons

Archimedes gave a scientific method for calculating π to arbitrary precision. The sum of the lengths of a regular polygon of n sides inscribed on a circle is smaller than the circumference, and the sum of the lengths of the sides circumscribed around a circle is greater than the perimeter of the circle. We take our regular polygon definition that circumscribed around a circle. For an n sided regular polygon, divide the circle into a set of $n \frac{n}{2\pi}$ angle isosceles triangles,

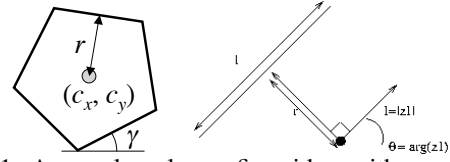


Figure 1: A sample polygon five sides, with parameters as defined above, and the line over which the centre may be for a given oriented edge pixels, as described in Equation 10.

where the centre point of the baseline falls on the circle. The length of the individual sides is $l = r2 \tan \frac{\pi}{n}$.

We may define a regular polygon as having a centroid in Euclidean 2-space (c_x, c_y) , the radius r of the circle it is circumscribed about, an orientation γ , and a number of sides n . Thus we may define a regular polygon transform as a function f_{rp} that maps the space of all possible regular polygons (regular polygon space) to 3D space. Regular polygon space is five dimensional, and the transform forms a mapping: $f_{rp} : \mathbb{R}^4, \mathbb{Z} \rightarrow \mathbb{R}^3$, where the integer dimension is the number of sides, and to form closed polygons this is three or greater, the space mapped to is image position plus orientation of edges. Specifically, we label regular polygon space: $\Phi = (c_x, c_y, r, \gamma, n)$. An example shape explaining the parameters is shown in Figure 1.

Mapping all regular polygons into the image from 5D regular polygon space is the integral across the space:

$$f_{rp} = \int_{\Phi} f_{rp}(\phi) d\phi, \quad (1)$$

The typical discrete image is a sampling over a regular Cartesian grid. The gradient of the image is a set of gradient pixels $\mathbf{x} = (\mathbf{x}_j, j = 1, \dots, s) \in \mathbf{I}$ corresponding to all the edge pixels with orientation of the regular polygons (as recovered by Sobel for example), plus any noise.

3 Likelihood formulation

A gradient image with a set of points \mathbf{x} may be regarded as the result of the transform from regular polygon space of some set of polygons $\phi_m \in \Phi$, where $m = 1 \dots M$. In order to recover ϕ_m from the \mathbf{x}_j , we may estimate the probability density function over regular polygon space given \mathbf{x}_j , considering the image as the result of a mixture of regular polygons, each of which is a Gaussian. That is:

$$f = \sum_{m=1}^M \alpha_m p(\phi_m | \mathbf{x}), \quad (2)$$

where α_m is a mixing parameter. If we assume a uniform mixture, then α_m is constant for all m . As $\sum_{m=1}^M \alpha_m = 1$, we may drop this term, and apply Bayes Law:

$$\sum_{m=1}^M p(\phi_m | \mathbf{x}) = \sum_{m=1}^m \frac{p(\mathbf{x} | \phi_m) p(\phi_m)}{p(\mathbf{x})} \quad (3)$$

Note that there is a distinction between a regular polygon in the scene and what appears in an image. An accidental view where several straight lines at different scene depths align to form a polygon is not a world structure, but is a polygon in the image. We may say that any apparent regular polygon is a regular polygon in the image. Further, with incomplete data, any edge pixel may be regarded as part a regular polygon. Thus, the edge image can be seen as being made up of regular polygons with noise. The result of this noise has two effects: edge displacement and gradient error. Due to standard image noise, or an imperfect scene edge (e.g., a faded road sign), the gradient edge in the image may be displaced slightly. There is also some residual uncertainty of position due to image sampling. All of these effects on the gradient edge may be reasonably approximated by a Gaussian in 3D over the parameters. The gradient error results from the effect of intensity noise in the image, and relation between orientation and intensity values is a sinusoid. For small angles the error in orientation will be linear and therefore may also be approximated by a Gaussian. Note that some authors choose to take gradient magnitude as a measure of edge certainty, however, it is also highly correlated with the intensity contrast of the underlying image. Thus, we threshold the gradient magnitude to basic noise, and set the remaining \mathbf{x}_j to unit magnitude. If required, the magnitude can also be included as a continuous measure of gradient certainty in this formulation.

We may take it that the \mathbf{x}_j are the result of a set of polygons with missing data. Also, that the noise in the appearance of the edge pixels in their position and estimated orientation is additive and independent between the points. Then we may take the probability of the Gaussian mixture of regular polygons in the image for the full set of points to be:

$$\sum_{m=1}^M p(\phi_m|\mathbf{x}) = \sum_{m=1}^M \prod_j \frac{p(\mathbf{x}_j|\phi_m)p(\phi_m)}{p(\mathbf{x}_j)} \quad (4)$$

Then, following Kiryati and Bruckstein [10], the probability density for an individual polygon ϕ_m and edge point \mathbf{x}_j can be modelled as a Gaussian in 3D $\mathbf{x}_j = (x_j, y_j, \theta_j)$ with zero mean and a point-specific covariance matrix:

$$P(\mathbf{x}_j|\phi_m) = \frac{\exp(-g(\mathbf{x}_j, f_{rp}(\Phi))\Sigma_j^{-1}g(\mathbf{x}_j, f_{rp}(\Phi)))}{2\pi|\Sigma_j|^{\frac{1}{2}}}, \quad (5)$$

where $g(\mathbf{x}_j, f_{rp}(\Phi))$ is a function giving the distance from the edge pixel to the closest point on the image projection of the regular polygon, Σ_j is the covariance matrix:

$$\Sigma_j = \begin{bmatrix} \sigma_{x_j}^2 & \sigma_{xy_j} & \sigma_{x\theta_j} \\ \sigma_{xy_j} & \sigma_{y_j}^2 & \sigma_{y\theta_j} \\ \sigma_{x\theta_j} & \sigma_{y\theta_j} & \sigma_{\theta_j}^2 \end{bmatrix} \quad (6)$$

and $|\Sigma_j|$ is the determinant of the covariance matrix. By the independence of noise between points, we may take the product over all edge points:

$$P(\mathbf{x}|\phi_m) = \prod_j \frac{\exp(-g(\mathbf{x}_j, f_{rp}(\Phi))\Sigma_j^{-1}g(\mathbf{x}_j, f_{rp}(\Phi)))}{2\pi|\Sigma_j|^{\frac{1}{2}}}, \quad (7)$$

Substituting back into Equation 3:

$$\sum_{m=1}^M p(\phi_m|\mathbf{x}) = \sum_{m=1}^M \prod_j \frac{\exp(-g(\mathbf{x}_j, f_{rp}(\Phi))\Sigma_j^{-1}g(\mathbf{x}_j, f_{rp}(\Phi)))p(\phi_m)}{2\pi|\Sigma_j|^{\frac{1}{2}}p(\mathbf{x}_j)} \quad (8)$$

Taking log-likelihoods and collecting constants:

$$\sum_{m=1}^M \log(\mathcal{L}(\phi_m|\mathbf{x})) = \sum_{m=1}^M \sum_j -g(\mathbf{x}_j, f_{rp}(\Phi))\Sigma_j^{-1}g(\mathbf{x}_j, f_{rp}(\Phi)) + const. \quad (9)$$

Matching to the regular polygon function: Two major issues are posed by the formulation of Equation 9. The term $g(\mathbf{x}_j, f_{rp}(\Phi))$ is the distance from \mathbf{x}_j to the nearest point on the regular polygon, i.e., this requires finding the nearest point. Also, Φ is a 5D space, and so optimising may be computationally expensive, while we would like to have a real time implementation.

As regular polygon space is high dimensional, the image is low dimensional, and edge points are sparse in general, we may pose the problem as a mapping from image edge pixels to possible regular polygons. We wish to form the inverse mapping f_{rp}^{-1} from \mathbf{x}_j to regular polygons that it may be a part of, i.e., Φ . f_{rp}^{-1} will allow us to create a function $g(\mathbf{x}_j, \Phi)$ that gives a distance metric from the point \mathbf{x}_j to the regular polygon defined by Φ .

Consider f_{rp}^{-1} for a given \mathbf{x}_j , for unknown r and γ . If we do not consider the edge gradient at \mathbf{x}_j , a regular polygon of any number of sides may have a centre anywhere, the subspace of Φ that \mathbf{x}_j may be an element of is large, and the function f_{rp}^{-1} is intractable. However, we may recover orientation using a standard image processing operator (e.g., Sobel) giving $\mathbf{x}_j = (x_j, y_j, \theta_j)$, where θ_j is the direction of the maximum intensity gradient. In this case the orientation of the regular polygon is constrained. Indeed, regardless of n , (c_x, c_y) is constrained to be on a line, orthogonal to the gradient, at a distance r from (x_j, y_j) . This relation is as shown in Figure 1. We maintain the standard of representing this in polar coordinates, but follow Weiman in representing a straight-line segment in the complex plane [22]. Consider that \mathbf{z}_1 is a complex number, with argument $\frac{\pi}{2} - \theta_j$, and modulus 1, we have:

$$h(\mathbf{x}_j) = \int_{t=-\frac{1}{2r}}^{t=\frac{1}{2r}} \mathbf{x}_j + r(\mathbf{z}_1 t + i\mathbf{z}_1)dt, \quad (10)$$

where \mathbf{x}_j is the (x_j, y_j) component of the point represented in the complex plane, l is the side length given n and r .

Using Equation 10, we may define a new regular polygon detection sub-space that is ambiguous in γ and n , but hence only 3D, and computationally simple. To form the detection subspace, we need only form the likelihood density function of Equation 9 over (c_x, c_y, r) . Let this new subspace be $\Psi = (c_x, c_y, r)$, and the function from this space to the image be f_{rps} (s for subspace).

In order to collapse the dimension of number of sides, we may assume the longest possible side length, that of a triangle. In the case of other shapes with shorter sides, this will lead to some edge pixels being considered that are not actually part of the true regular polygon. However, these ambiguities may be resolved subsequently. Thus, for any particular $\psi \in \Psi$, we may define g as the minimum distance from Ψ to the line defined in Equation 10.

Regular polygon detection: Using Equation 10 to find likelihood density across a Ψ , we may take the peaks of the distribution as the set of n regular polygons that are most likely in this subspace, by the assumption of zero mean, if we assume that the actual regular polygons are spatially separate. We will deal with the difficulties that arise from this assumption subsequently. Rather than recover the variance directly, we resolve the remaining dimensions in a second stage. This separation is computationally motivated. By finding most likely polygons in a reduced space we may focus computation on the most likely polygons, and ignoring parts of Ψ^1 that are not likely to have any regular polygons present. This will be sufficient to recover all polygons to some required likelihood, if we investigate regular polygons in the reduced space down to the required likelihood as the likelihood in Φ is an overestimate of that in Ψ .

In practice the variance constitutes: remaining dimensions of n and γ ; and, regular polygons (or partially visible regular polygons) that are overlaid. This may result in a figure that has a number edges of different length, for which likelihood is overestimated. For example, consider an octagon, where all sides are the length of sides for the corresponding triangle, although this is a correct octagon, its likelihood may be overestimated in Φ .

Resolving orientation and shape: Now we have resolved the mean for $\psi_h = (c_{xh}, c_{yh}, r_h) \in \Psi$ for the set of regular polygons that are most likely in Ψ . We may construct a likelihood function given this information to resolve the remaining dimensions, that is:

$$\log(\mathcal{L}(o, n | \mathbf{x}, c_x, c_y, r)) = \sum_j -g(\mathbf{x}_j, f_{rp}(\Phi)) \Sigma_j^{-1} g(\mathbf{x}_j, f_{rp}(\Phi)) + \text{const} \quad (11)$$

With a truncated distance function, for each point in detection subspace we may consider only the local \mathbf{x}_j , this

¹ and of the image if we truncate our distance function.

may be assembled by maintaining a list of the parameters for the \mathbf{x}_j that contributed to that point. This includes its basic parameters, plus its distance as defined by g in the 3D subspace, let this be d_j . We may iterate over n , and integrate over γ and compute the log-likelihood of the probability density function over the remaining parameters using Equation 12. This is now a reduced dimensional space search, however an efficient approximation to this algorithm is discussed in the implementation.

4 Implementation

Computation of the complete regular polygon likelihood distribution is expensive. Instead we discretely sample the distribution in a manner that is appropriate for the application. Note that a coarse-to-fine strategy may be employed or found peaks may be better located by subsequent EM. The likelihood distribution can subsampled by a simple scan line algorithm through x_j . If one considers the distance function to be the soft characteristic function of [14], then the log-likelihood estimation may be considered as a Hough-like algorithm for estimation of regular polygon parameters.

Using g^2 , the likelihood contribution of a regular polygon of x_j rapidly becomes small as distance increases. Thus, if $\mathcal{L}(\mathbf{x}_j | \Psi) < \epsilon$, we set it to zero. Given this formulation, the algorithm can operate by, for each edge pixel, step along the line of h for the width implied by ϵ , and adjust the log likelihood of each sample point of the distribution according to $\mathcal{L}(\mathbf{x}_j | \Psi)$. Alternatively, we may approximate by adjusting only the closest pixels to the line, and convolving the likelihood with a Gaussian. Thus for each \mathbf{x}_j , we update log-likelihoods for rlw points in Ψ , where w is the width implied by ϵ . In practice, small w is sufficient, for a fast implementation this may be one-two pixels. Also, for most specific applications, r may be over a small range.

The second stage of the algorithm can also be implemented more computationally efficiently. In practice, one is typically seeking either the set of k regular polygons with the maximum likelihood, or all regular polygons with a likelihood greater than a threshold. Let us assume initially that the \mathbf{x}_j in the list are all actually part of the polygon. In this case, γ will be a mixture of Gaussians corresponding to the orientations of each of the component edges. We may define the log-likelihood of the probability of a regular polygon as the sum over the edges, splitting Equation 12 over edges, and replacing $f_{rp}(\Phi)$ with each of the edges from Equation 10. In order for an \mathbf{x}_j to be an element of any particular edge, θ_j must fall within the expected variance of the mean for the orientation of the edge.

As edge length may be less with known n , an edge-based distance function (with orientation) is more constrained than $g(\mathbf{x}_j, f_{rps}(\Psi))$ adjusted to include orientation. Thus, if we approximate the distance function, by its distance d_j , multiplied by a distance function over θ_j , this is an upper bound on the actual edge distance function. If we take the

Given edge elements \mathbf{x}_j of an image, detect regular polygons with $n \in N$ sides. For each polygon radius r :

1. Estimate likelihood image:
 - (a) For each \mathbf{x}_j : Compute all triangle locations p_k that could generate \mathbf{x}_j , accumulate locations in a vote image, and record information: $p_k, \angle \mathbf{x}_j$ and distance m_k along line h .
 - (b) Convolve vote image with Gaussian to generate discrete approximation of likelihood image.
2. Evaluate each maxima q_i for each $n \in N$:
 - (a) Build a weighted angular histogram of all recorded $\angle \mathbf{x}_j$ values with $m_k < \frac{l(n,r)}{2}$, weighted by a Gaussian of $\|p_k - q_i\|_2$.
 - (b) Convolve histogram with string of delta functions $\delta(\gamma - \frac{2\pi}{n})$ corresponding with the p peak edges. The γ that maximises this convolution is the most likely orientation of an n sided polygon, and the values at $\gamma \pm \frac{2\pi}{n}u, u \in \mathbb{Z}$ indicate the support for each side.
 - (c) Total support is determined as a function of the support for each side of the shape.

Figure 2: Summary of the algorithm.

weighted \mathbf{x}_j over θ and convolve it with a Gaussian corresponding to our expected variance in θ , then this forms an approximate upper bound on the log-likelihood of the probability density function of individual edge orientations. We may use this to hypothesis test for γ and n , by taking each hypothesis in turn and summing over edges in the upper bound density function by direct look up at the expected angles, to form the upper bound density function across our remaining parameters. Rather than for all γ we may take the top p peaks as this corresponds to upper bounds likelihoods of the most likely edges. For each of these we may form the complete likelihood function of Equation 12. As the mean of the maximum likelihood orientation may not correspond to the mean orientation of the most likely edge. If required an EM search may be performed over orientation in the neighbourhood of these peaks.

For real applications, the number of radii being explored is not large, and the sampling can be coarse. Further if it is acceptable to overstate the likelihood, and precise position is not required, the algorithm may cease without removing edge pixels. Implementation is summarised in Figure 2.

With no *a priori* scene information, we may assume that the \mathbf{x}_j appear with an isotropic distribution, except for regular polygons. However, in a road scene, for example, the edges of the road and dividing lines often result in long piecewise straight edges. In this case, the probability of multiple edge points appearing together, with the same gradient, that are not part of something we wish to consider as a regular polygon, is greater than the probability of pixels appearing together randomly in the image.

In such a case, we may determine *a priori* distributions

from a set of images that is representative of the incoming images. From this we may adjust the likelihoods to reflect the number of supporting edges from the support for each of the edges. In such long clear line scenarios, we found it was effective to incorporate negative probability weighting at the ends of the line of influence of the \mathbf{x}_j . This prevented over emphasis of strongly contrasting long lines.

Finally, where other information is known *a priori* we may incorporate this information into our function, and simplify computation. This will be explained in more detail in the context of road sign detection in the results.

5 Experimental results

We evaluate performance in the presence of noise in artificial images, as well as on real images. By incorporating prior knowledge about the appearance of the road scene, we were able to adapt the algorithm to run at 16Hz for 320×240 images in which it was able to reliably detect a road sign in a test sequence taken from the vehicle in Figure 5(a). Finally, we demonstrate the application of a constrained version as a feature detector for wide-baseline matching for robot corridor navigation on the robot shown in Figure 5(b).

On still images: The detector was evaluated on a range of images containing regular polygons. Figures 3 and 4 show detection results for searching for 3 to 8 sided regular polygons over radii $r \in \{8, 11, \dots, 17\}$. Note that this non-continuous range of radii is adequate to detect shapes at neighbouring radii. Figure 3(b) shows the approximation of the likelihood image generated by the algorithm; note the peaks at the centres of the shapes detected in (c). The final output (d) illustrates both the impressive detection performance, and illustrates some of the artefacts of the method. Owing to the Gaussian modelling of edge and centroid locations the shapes detected do not always exactly overlay the edge locations of the original shapes. Also, polygons may be found with alignment of partial edges, and are awarded low (but not zero) strength accordingly.

The majority of incorrect hypotheses correspond to polygons of a small number of sides. There are two reasons for this: firstly, (for $n \leq 8$) as n increases it becomes less likely that n edges will be ‘accidentally’ aligned by noise; secondly the prevalence of right angles in built environments gives a strong prior for finding squares.

Figure 4 shows the detector operating on outdoor images, here the near-perfect regular polygon nature of the road signs ensure that they are strongly detected. The top right figure shows an example of the detected square shapes common around man-made structures, however, for sign detection the orientation of the diamond-shaped warning sign in the foreground sets it apart from the other squares.

Robustness to noise: The 180×240 test image $I : I(p) \in [0.25, 0.75]$ in Figure 3 (e) was corrupted by increasing amounts of additive Gaussian noise $\sigma = 0, 0.05, \dots, 1$ and performance was evaluated over 20 runs at each noise

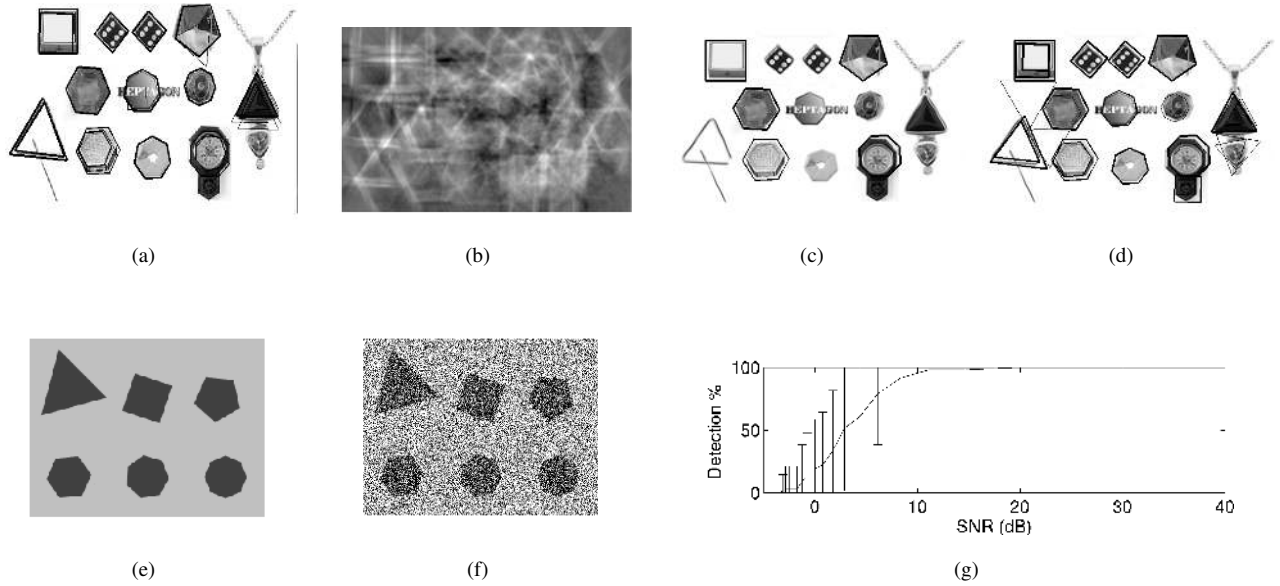


Figure 3: (a) Input image, (b) likelihood image for $r = 17$, (c) results for $r = 17$, and (d) results across all radii. Detection results faded from red (strong) to blue (weak). (e) Noise test image, (f) with SNR=3Db ($\sigma = 0.5$) (g) mean and standard deviation of detection performance with additive Gaussian noise.



Figure 4: Performance on outdoor images.

level at radius $r = 20$. Performance was quantified from the top six shape hypotheses which were taken to be correct if they located the shape with the correct number of sides within ± 4 pixels of its ground truth location. Figure 3 (f) shows the average detection performance versus noise, demonstrating stability in the presence of noise. At $\sigma = 0.5$ the average detection rate drops below 50%, but at this point it is difficult even for a person to differentiate between some of the shapes, Figure 3 (e).

Real-time sign detection: Critical information signs with information that requires driver action regularly appear on triangular, diamond, or octagonal backgrounds. In road scenes signs are designed to be easily visible; they appear at set orientation (approximately) unless damaged. Further, if the camera axis aligns with the vehicle's forward motion,



Figure 5: (a) Inside the intelligent vehicle. (b) The NOMAD robot for navigation in corridor-like environments with a camera pair on top.

signs will always be parallel to the image plane when they are close, even on curved roads and will be visible for many frames. Also, for a camera of approximately known focal length, their apparent size (radius) is constrained over a narrow range. Signs smaller than a few pixels cannot be recognised, so need not be detected. As signs have maximal world size, physical constraints mean they will never appear more closely than a set distance from the vehicle.

Detection aim to reduce each image to a small number of possible candidates. A sign need not be detected in every frame, but must be reliably detected while it is visible. Due to the appearance of regular polygons due to accidental features, it is efficacious to require a regular polygon to be detected over several images at a similar position and size as such accidental views will occur frequently in robot vision, but often will not be sustained.

In this trial, we constrained the detector to giveaway and roundabout signs, triangles, where the top edge is parallel to the ground plane, thus the number of sides is constrained, and orientation is constrained to a range *a priori*. For the standard configuration of the vehicle used in this sequence, four separate radii were sufficient of 6, 8, 10, and 12 pixels. The *a priori* constraint over number of sides and radii meant that gradient orientations outside the constraints could not be part of these oriented polygons, so could be disregarded *a priori*. These orientations were 0, 120 and 240 °, each $\pm 12^\circ$. Thus, only the first stage of the algorithm was required to recover all ambiguity. The constrained algorithm was able to run in less than 50ms per frame (320x240 image), based on an implementation in C++ on a standard PC, equivalent to that mounted in the vehicle. This is quite adequate for a sign to be visible for many frames in any reasonable situation for a road vehicle, and hence is real-time.

The road trial is on a sequence, see Figure 6. Of the sequence the sign is of a detectable size for 48 frames. A sample detection image is shown in Figure 6. For detection, we required the sign to be one of the two most likely triangles for any radius for an initial image, and then one of the top 10 for the next k images. False positive and false negative curves are shown in Figure 7. Although the sign was not detected as one of the top two candidates in every image, it was reliably detected many times over the image sequence. The processing of recognition using normalised cross correlation for each candidate is less than 1ms in our system, so the number of false positives was well within computation limits for real time processing.

Indoor scenes: In a project to develop vision-based robotic mapping, we applied the regular polygon detector to detect square-like features. Here n is set, but γ must be resolved, i.e., 1D ambiguity after initial detection. The detector was run at three different radii, and found all squares above a low likelihood (a few pixels on three edges). It detects stable edges that are the basis of squares, but is tolerant to incomplete edges and affine distortion. Figure 8 shows a sample image. The detector was applied to pair of images taken from the robot as it was navigating around indoor office environment, moving at approximately 30 cm/sec. With images three seconds apart the baseline was wide, approximately one metre. On corners the robot turned mostly in



Figure 6: An image from the middle of the roundabout sign sequence, with most likely triangles backprojected, and the likelihood function for radius 8.

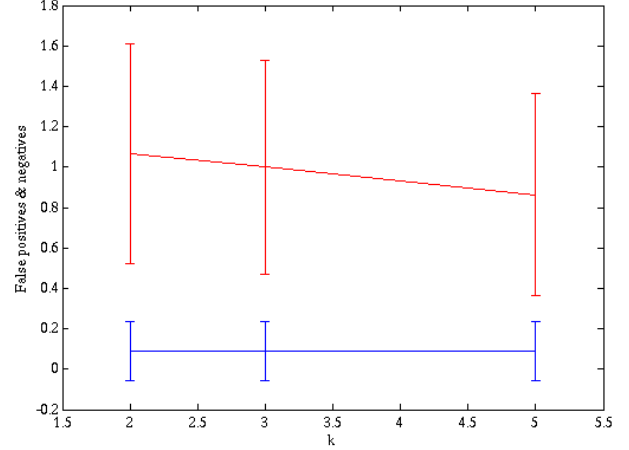


Figure 7: Average and standard deviation of number of false positives (top) and false negatives (bottom), given number of frames k that the sign is required to be present.

place, leading to shifts of more than half the image of features, see Figure 9. Features were matched according to size and orientation, the local greyscale environment and position, and used to calculate the fundamental matrix to recover motion. Results showed a high percentage of the found points matched (around 70% typically) correctly over frames moving down the corridor, but reduced over rotation as expected. Figure 9 shows a typical images taken from the sequence of 69 images as it turns around a corner. The matches show promising initial results. Figure 9 (c) shows the matches obtained by SIFT [11], using the reference implementation with its packaged parameters.² Note that the square feature detector will be most effective in this type of rectilinear environment, however, it is highly suitable to be part of a battery of detectors for matching. Its speed of operation makes it plausible for robotic mapping applications.

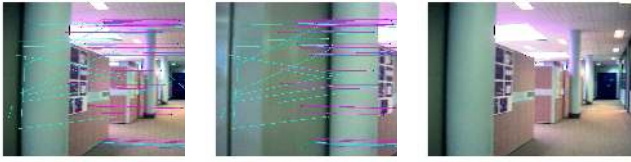
6 Conclusions

Regular polygon detection is an important problem with multiple applications in robotics and computer vision. Using an aposteriori probability approach, we presented a new

²<http://www.cs.ubc.ca/~lowe/keypoints/>



Figure 8: Square features: size and orientation are represented directly, likelihood is represented by colour, where red is most certain, and yellow is least certain.



(a) (b) (c)
Figure 9: Matches for turning a corner. The pink lines represent correct matches, while the others represent erroneous matches. (c) Matches found using SIFT, only one found.

algorithm for detecting regular polygons. We defined the continuous log-likelihood of the probability density function of regular polygons. In order to make the algorithm computationally efficient we find initial likely regular polygons in a lower dimensional space, and then resolve the remaining parameters for likely regular polygons. We presented an efficient algorithm based on this, and adaptations of this algorithm that can run at 16Hz detecting signs in an intelligent vehicle application. Experimental results show the efficacy and robustness of the algorithm, and its application to driver assistance, and as a feature detector for matching, applied to a robot sequence for mapping.

References

- [1] E R Hancock A Robles-Kelly. A probabilistic spectral framework for grouping and segmentation. *Pattern Recognition*, 37(7):1387–1405, 2004.
- [2] D H Ballard. Generalizing the hough transform to detect arbitrary shapes. *Pattern Recognition*, 13(2):111–122, 1981.
- [3] I J Cox, J M Rehg, and S L Hingorani. A bayesian multiple hypothesis approach to contour segmentation. *International Journal of Computer Vision*, 11:5–24, 1993.
- [4] D Crevier. A probabilistic method for extracting chains of collinear segments. *Image and Vision Computing*, 76(1):36–53, 1999.
- [5] W Dickson. Feature grouping in a hierarchical probabilistic network. *Image and Vision Computing*, 9(1):51–57, 1991.
- [6] C Geyer, S Sastry, and R Bajcsy. Euclid meets fourier: Applying harmonic analysis to essential matrix estimation in omnidirectional cameras. *OMNIVIS04: The fifth Workshop on Omnidirectional Vision, Camera Networks and Non-classical cameras*, 2004.
- [7] G Guy and G Medioni. Inferring global perceptual contours from local features. *International Journal of Computer Vision*, 20(1-2):113–33, Oct. 1996.
- [8] P V C Hough. Method and means for recognizing complex patterns. (3,069,654), December 1962. U.S. Patent, 3,069,654.
- [9] S-H Hsu and C-L Huang. Road sign detection and recognition using matching pursuit method. *Image and Vision Computing*, 19:119–129, 2001.
- [10] N Kiryati and A M Bruckstein. Heteroscedastic hough transform (htht): An efficient method for robust line fitting in the ‘errors in the variables’ problem. *Computer Vision and Image Understanding*, 78(1):69–83, Apr. 2000.
- [11] D G Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2):91–110, 2004.
- [12] G Loy and N Barnes. Fast shape-based road sign detection for a driver assistance system. In *Proc. 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems. IROS2004*, 2004.
- [13] G Loy and A Zelinsky. Fast radial symmetry for detecting points of interest. *IEEE Trans Pattern Analysis and Machine Intelligence*, 25(8):959–973, Aug. 2003.
- [14] A Makadia and K Daniilidis. Correspondenceless ego-motion estimation using an imu. In *Proc. IEEE Int. Conf. on Robotics and Automation (ICRA ’05)*, Barcelona, Spain, April 2005.
- [15] L G Minor and J Sklansky. Detection and segmentation of blobs in infrared images. *IEEE Trans. on Systems, Man, and Cybernetics*, 11(3):194–201, 1981.
- [16] P Perona and W T Freeman. Factorization approach to grouping. In *European Conference on Computer Vision 1998*, pages 655–70, 1998.
- [17] G Piccoli, E De Micheli, P Parodi, and M Campani. Robust method for road sign detection and recognition. *Image and Vision Computing*, 14(3):209–223, 1996.
- [18] S Sarkar and K L Boyer. Quantitative measures of change based on feature organisation: Eigenvalues and eigenvectors. *Computer Vision and Image Understanding*, 71(1):110–136, 1998.
- [19] A Sha’asua and S Ullman. Structural saliency: The detection of globally salient structures using a locally connected network. In *Proc. International Conference on Computer Vision*, pages 321–327, 1988.
- [20] R S Stephens. Probabilistic approach to the hough transform. *Image and Vision Computing*, 9(1):66–71, 1991.
- [21] C F R Weiman. Polar exponential sensor arrays unify iconic and hough space representation. In *Proceedings SPIE: Intelligent Robots and Computer Vision VIII: Algorithms and Techniques*, pages 832–842, 1989.
- [22] C F R Weiman and G Chaikin. Logarithmic spiral grids for image processing and display. *Computer Graphics and Image Processing*, 11:197–226, 1979.