

An Analysis of Pairwise Geometric Histograms for View-Based Object Recognition

P. A. Riocreux, N. A. Thacker and R. B. Yates

Department of Electronic and Electrical Engineering, University of Sheffield, PO Box 600, Mappin St., Sheffield, S1 3JD.

Abstract

A pairwise geometric histogram (PGH) encodes the probability of geometric co-occurrences between any line and the set of lines defining an object. An object therefore has a set of PGHs associated with it, one histogram for each line. We describe here the way in which the probability of geometric co-occurrence is calculated and entered in the histograms, the different ways these histograms can be defined and the completeness properties of the set of histograms in terms of arbitrary shape representation. We show that this representation provides unambiguous shape representation by demonstrating an inverse reconstruction algorithm. We conclude that the methods suggested in a previous paper for object recognition and location provide a complete solution to the problem of recognition of edge based descriptors for fixed 2D projected views of rigid objects.

1 Introduction

The process of PGH generation works on a line based approximation to the edge data in the scene. The set of lines comprising an object are obtained from an edge string segmentation algorithm which delivers a polygonal approximation to the edges present in an image. Each line is regarded in turn as a reference line and its geometric relationship with all other lines in the scene is computed. The geometric measures relating each pair of lines are stored in a histogram as a probability density distribution. In order to compute the probability of a particular geometric co-occurrence we need a statistical model of the geometric measurement process. The probabilistic encoding should therefore be related to the polygonal approximation algorithm.

The segmentation method that we use is a recursive algorithm which approximates an arbitrary string of individual edge elements (edgels). The algorithm proceeds as follows:

- A straight line is drawn between the endpoints of the string.
- The perpendicular distance of closest approach to the line is calculated for each point on the string, and the maximum identified.

- If the ratio of the maximum perpendicular distance to the length of the original line is greater than an approximation threshold, λ , then the string is split at this point, and both parts are reprocessed

This algorithm continues until all string fragments are represented to an accuracy specified by the threshold. The algorithm is similar to that proposed by Lowe [1] except that the segmentation criterion is based on a relative rather than absolute distance measure.

In order to properly encode the original data in the PGH we need to know the expected distribution of edgel orientations in the string. On the assumption that the most significant variation of the data away from purely linear behaviour will be second order, we can derive a relationship between the segmented line A-C and the parameters of the circular arc A-B-C (Figure 1).

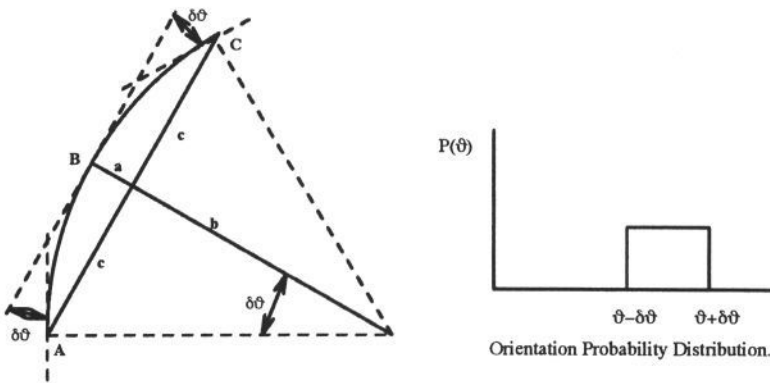


Figure 1: Relationship between the segmented line and the arc

If the approximated edge string represented an arc of a circle then the midpoint of the string would have the same orientation in the image, θ , as the approximated line with endpoints of orientation $\theta - \delta\theta$ and $\theta + \delta\theta$ where $\delta\theta$ can be related to the segmentation parameter λ as follows.

From Figure 1 it can be shown that $a/c = (1 - \cos(\delta\theta))/\sin(\delta\theta)$, and making the standard approximations to the trigonometric functions for small angles gives $a/c = \delta\theta/2$. The definition of the segmentation parameter states that $a/2c < \lambda$, and thus $\delta\theta < 4\lambda$.

Therefore, to second order, the edgel data will have a uniform probability of distribution of orientations between upper and lower bounds, with these bounds being fixed by the segmentation parameter λ .

On the basis of this model of shape approximation the appropriate method for making entries in the PGH for relative orientations between lines is in the form of a triangular distribution obtained by convolving the rectangular expected distributions of orientations found in the approximations to the reference line and the object line. This distribution needs to be integrated between the upper and lower limits of any bin in the PGH in order to compute the weighting vector, A_C , for each entry.

The source of correlation between small variations in object line rotation and the estimation of perpendicular distance is shown in Figure 2. Provided that the

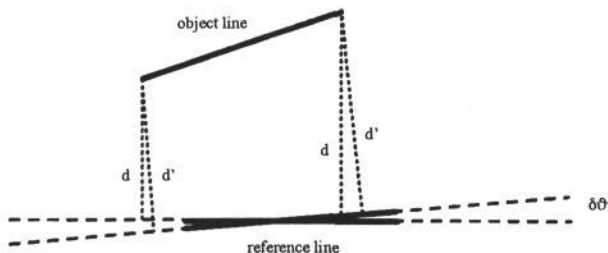


Figure 2: Effect of small angle variations on perpendicular distance

perpendicular distance is encoded on a scale which is not affected by small (typical) rotations in the reference line, correlated effects between the estimation of the perpendicular distance and the relative orientation are not be observable in the PGH and therefore do not need to be modelled. This is achieved when the entries into the PGH are convolved with a function which has a scale similar to the maximum expected correlation. We choose the form of this convolution to be a top hat, which has the property of invariance to line fragmentation. The total weighting function required to make entries into the PGH for a particular reference and object line pair W_{ro} is given by the outer product of the angle probability distribution, A_C , and the trapezoidal perpendicular distance probability distribution, L_C , multiplied by the length of both lines.

$$W_{ro} = |l_r||l_o|L_C \otimes A_C$$

Construction of the full histogram is then achieved by summing these components for all lines in the scene.

$$W_r = \sum_o W_{ro}$$

A typical histogram is shown in Figure 3.

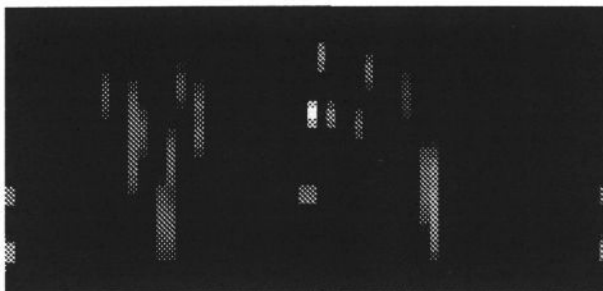


Figure 3: A typical pairwise geometric histogram

The segmentation algorithm adopted here is especially suited to a simple computation of the probability of geometric co-occurrences. Comparison of these histograms to determine the best match is then achieved using a finite approximation

to the Bhattacharya distance measure L_B , which is the appropriate way of comparing PDFs.

$$L_B = \sum_d \sum_\theta \sqrt{W_1(d, \theta)} \sqrt{W_2(d, \theta)}$$

This has been shown to have good discrimination properties on data from images containing occlusion clutter and line fragmentation [2].

2 Constructing Alternative Representations

Having established the method for generating the geometric co-occurrence probabilities it is now necessary to describe some of the alternative definitions of measurement axes for the whole PGH. The information we wish to encode is the relative angle and range of perpendicular distance between any two lines in the object. There are several ways of doing this depending on the information that we wish to use. The possible sources of information are:

- Angular handedness of relative orientation (clockwise/anti-clockwise)
- Direction to intersection of lines.
- Edge contrast orientation.

A subset of the resulting range of possible histograms is shown in Figure 4. In this, each small rectangle defines a region of area $0 - \pi \times 0 - d_{max}$. Multiple entries within a region implies summation of data with the orientations shown by the associated coordinate axes. Wrap around directions are shown at the edges of the plot.

Figure 4(i) shows the most general form of histogram, from which all the other types can be derived. The reference line is assigned an arbitrary direction, and the object line is assumed to be directed away from the intersection of the line extensions. From this information the relative angle and perpendicular distance are calculated. If the assumed direction of the object line agrees with the direction indicated by the contrast information, the entry is placed in the first or third columns of the histogram, otherwise in the second or fourth columns. If the intersection direction of the reference line agrees with the arbitrarily assigned direction the entry is placed in the first or second columns of the plot, and otherwise in the third or fourth columns. There are two alternatives for the constructed histogram depending on the initial choice of arbitrary direction.

If the assigned direction of the reference line is ignored then the histogram size is halved and we obtain Figure 4(ii). This still makes use of the edge contrast information, which is generally unreliable in practical situations. If the edge contrast information is ignored, we obtain Figures 4(iii) and (iv), which use and ignore the assigned direction of the reference line respectively. If the assigned direction rather than the intersection direction of the reference line is used to calculate the angle and distance measures, we obtain Figure 4(v). This type of histogram differs from Figure 4(iv) in the important fact that all object lines lying to one side of the reference line will be entered in the histogram on the same side of the distance axis.

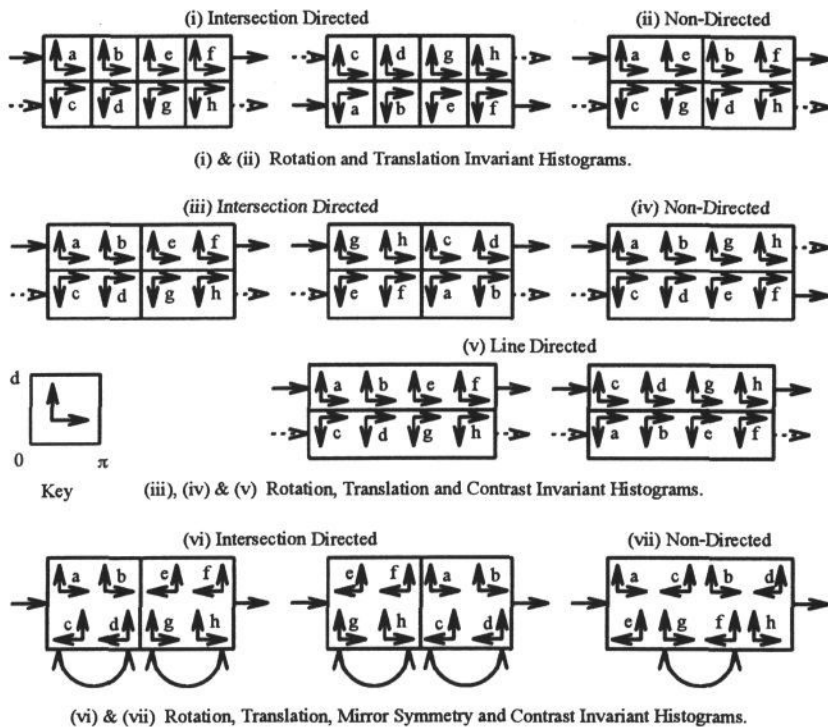


Figure 4: The family of pairwise histograms

Finally, if we ignore the handedness of the relative angle, we obtain Figures 4(vi) and (vii), using and ignoring the assigned direction respectively.

type	scale	rotation	translation	extension	mirror	contrast
(i)	No	Yes	Yes	Yes	No	No
(ii)	No	Yes	Yes	Yes	No	No
(iii)	No	Yes	Yes	Yes	No	Yes
(iv)	No	Yes	Yes	Yes	No	Yes
(v)	No	Yes	Yes	Yes	No	Yes
(vi)	No	Yes	Yes	Yes	Yes	Yes
(vii)	No	Yes	Yes	Yes	Yes	Yes

Table 1: Invariance properties of the various histogram types

Given a random population of data in these plots for arbitrary objects we would expect the disambiguational ability of each of these alternatives to increase with total histogram area. However, as mentioned above, absolute line orientation is not reliable in most practical applications, therefore the most useful representation must be considered to be that shown in Figure 4(iii). In the next section we will develop a theoretical argument regarding completeness of this representation.

3 Representational Completeness

Work has already been published regarding the smooth variation in PGH with continuous shape deformation and the resulting gradual change in similarity measure [3]. Work has also been presented demonstrating the robustness of the scheme to missing data and added noise [4]. However, for the recognition system to be fully quantified we need to assess its discriminability.

One approach would be to test the performance of the system on a set of standard test objects. This is important for system demonstration but is not really useful for quantitative analysis. A better approach would be to develop a prediction of performance based on a theoretical evaluation of the algorithm which may then be used to predict the expected performance of the recognition system for arbitrary data sets. The statistical basis of the PGH representation makes the possibility of such an evaluation much more likely than with other, more ad-hoc recognition systems.

Currently we can show that use of the Bhattacharya similarity measure will result in optimal recognition performance for individual histograms, but as these histograms are constructed with invariance properties there will inevitably be some cases where different objects could give rise to similar histograms, leading to ambiguity. However, full recognition schemes based on data combination in a neural network [5] or a Hough transform [2] make use of the complete set of pairwise histograms. If we could show that the set of histograms uniquely defines an object this would mean that the representation scheme was complete and it may then be possible to show that recognition is optimal. One way to do this is to show that an inverse transformation process exists which will reconstruct the original data, in this case the line based description of the object. This process is possible with some of the forms of PGH described above but in order to explain the process, some properties of the PGH need to be discussed.

Some of the types of PGH for an individual line can be considered as a projection along the direction of the line through the area of the object onto a projection axis (the perpendicular distance axis). This is only true for the histograms shown in Figure 4(i),(ii),(iii) and (iv) as the construction processes for the other histograms sometimes place data on the opposite side of the projection axis due to the lack of knowledge regarding absolute orientation of the reference line. Normalisation of entries from each line pair ensures that processes such as integration of contents for each angle bin onto a perpendicular projection axis will produce values which can be regarded as integrated edge densities. Thus, if the histogram includes entries for all the lines in the object, the mean of the distribution corresponds to the projection of the centroid of the original data onto the distance axis, as shown in Figure 5.

Projection of data in the histogram onto the angle orientation axis produces a 1D histogram which is completely scale invariant (the original form of PGH [5]). If the histogram again includes entries for all the lines in the object, this projection of the PGH must be identical for each line in the object but for a shift of origin. The relative orientation between lines in the object can thus be determined by determining the shift in the angle projection histograms.

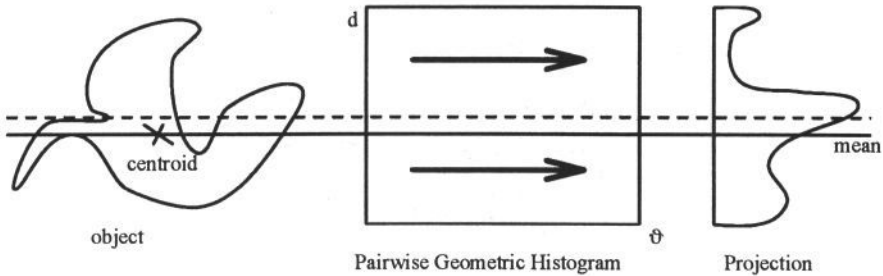


Figure 5: Calculation of the centroid from the projection onto the distance axis

4 Reconstruction Algorithm

We now have enough information with which to describe a reconstruction algorithm. Each histogram contains information regarding the integrated density of edges through the region of the object along a direction in the original image specified by the reference line. Each projection can be aligned with the centroid of the object using the distance projection histogram and oriented relative to others using the angle projection histogram. Thus the set of pairwise histograms provides a complete set of projections through the object analogous to a 2D image reconstruction process such as is commonly found in medical image processing applications. In the examples shown below we have used the exact values of orientation and centroid which are stored when the histograms are constructed. This has been done for speed, and does not invalidate the technique, as the matching and segmentation processes also have access to and use this data. Thus no data is being used which is not used by the recognition scheme. Similar but slightly less well defined results are obtained if the calculated values are used.

Many schemes have been described for reconstruction of images from projections, but most require the data to be of a particular nature, or are extremely computationally intensive [6]. The process requires the construction of a large set of linear equations for the known projections in terms of the unknown localised densities (pixels). Least-squares methods are clearly applicable and could be used to estimate the density of data in the image. For the set of n geometric histograms having m distance bins, this implies that we have nm constraint equations. However, the situation is slightly better than this as a zero projection value can specify a whole set of edge densities as will be described below.

As image plane pixels cannot contain a negative number of edgels, then if a projection total is zero, all image plane pixels that such a projection includes must have contained zero edgels. This can be used as the basis of a reconstruction algorithm when reconstructing the image with pixels whose side length is less than half the size of the projections, so that even in the worst case¹ a contiguous stream of pixels will remain, as shown in Figure 6.

The reconstruction is performed by eliminating those pixels in the reconstruction plane that could not have contained an edgel in the image plane. The data for lines lying within a small range of absolute angles is generally sparse, so recon-

¹A vertical or horizontal non-zero projection between two zero projections.

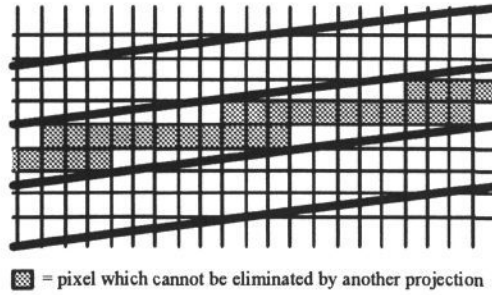


Figure 6: Elimination of half-size pixels

struction is performed with this data. Full reconstruction is performed by simple addition of the results of reconstructions for a complete set of small angle ranges. In this case it is natural to use the extents of the angle bins from a particular histogram as the angle ranges. The data for these processes is obtained by selecting an angle bin-distance bin entry from a type (v) histogram².

As the relative orientations of the reference lines of the histograms are already known, they may be used to linearly interpolate the corresponding angle bins in other histograms, so as to include all the angles in the range being considered, but possibly also including other angles. This allows projections for lines with the same range of absolute orientations, but with the projections parallel to each histogram's reference line. Thus the first histogram will give a band across the reconstruction plane, with subsequent histograms cropping this band. The limits parallel to each line will be well defined, as there will a projection from a histogram having the line as its reference.

Figure 7 shows the results of the reconstruction process. Figure 7(i) shows the original line data, (ii) shows the results of a projection from one histogram, (iii) shows the results of a reconstruction for a single angle range, and (iv) shows the complete reconstructed image³.

Figure 8 shows how the results change for coarser resolutions. In this case there are some pixels which have not been eliminated so that the image provides a superset of the correct pixels. This estimate of pixel location may be refined further by least squares solution of the remaining constraint equations but this has not been attempted here. Even in this current form the algorithm illustrates that even at relatively coarse histogram resolutions most of the shape information remains⁴.

5 Discussion and Conclusions

The process of entering probability density distributions in a histogram incorporates a model of both the segmentation process and the process error in the representational scheme. This makes the subsequent recognition scheme very robust

²This type of histogram is easily constructed from types (i), (ii) and (iii).

³Histogram size is 200×32 with 1 distance-bin blurring, and $\lambda = 0.03$ radians.

⁴Histogram size is 64×32 with 1 distance-bin blurring, and $\lambda = 0.03$ radians.

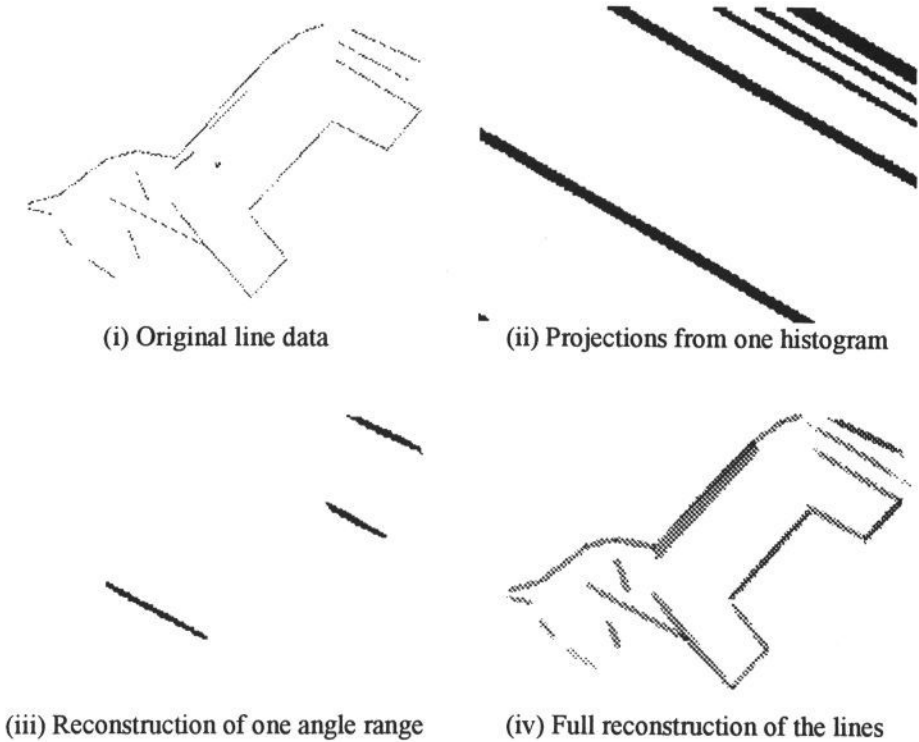


Figure 7: Fine resolution reconstruction process results

to both of these factors.

Reconstruction ability indicates that the set of histograms is a unique and complete representation of the image data. This must be a requirement of any representational scheme that is to discriminate between sets of arbitrary shapes. In practical use, the PGHs are typically 32×32 in size, resulting in a relaxation of the geometric constraints. Even so, the multiple redundancy of the data results in the retention of almost all salient edge information. If this histogram "blurring" is artificially increased, then this scheme is effectively able to represent deformable shapes as well as rigid ones. This appears in the reconstruction process as a thickening of the edges.

A recognition system which makes use of the set of geometric histograms in the appropriate manner can be considered as an algorithmic approximation to blurred edge template matching with edge orientation sensitivity. To date we have identified two separate ways of combining the recognition of a set of histograms into a single recognition process, the first is by set classification using a hierarchal neural network [5] which provided only scene content information and the second by use of the generalised hough transform [4] which provided object location. We are currently extending this later method to use a version of the probabilistic hough transform to improve robustness and allow the recognition of scaled objects. This work will be the subject of future papers.

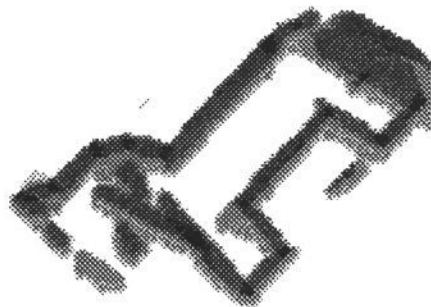


Figure 8: Coarse resolution reconstruction process results

We have shown that PGHs are a complete, unique and robust scheme for representing arbitrary shapes in terms of their edge information. Their behaviour under object deformation makes them ideal candidates for forming the basis of a view based recognition system.

Acknowledgments

The authors wish to acknowledge the work of Alun Evans in the early development of the ideas presented here. The work was partly supported by the SERC (grant number GR-J10464) and the University of Sheffield.

References

- [1] D.G.Lowe, Three-dimensional object recognition from two-dimensional images, *Artificial Intelligence*, vol. 31, no. 3, pp355-395, 1987.
- [2] A.C.Evans, Geometric Feature Distributions for Shape Representation and Recognition, Ph.D. Thesis, Sheffield, January 1994.
- [3] A.Evans, N.A Thacker and J.E.W.Mayhew, Pairwise Representations of Shape, Proc(I). International Conference on Computer Vision and Pattern Recognition (ICPR), The Hague, August 1992, pp133-136.
- [4] A.C.Evans, N.A.Thacker and J.E.W.Mayhew, The Use of Geometric Histograms for Model-Based Object Recognition, Proc. 4th BMVC93, Guildford, 21-23 Sept. 1993, pp429-438.
- [5] N.A.Thacker and J.E.W.Mayhew, Designing a Network for Context Sensitive Pattern Classification, *Neural Networks*, vol. 3, no. 3, pp291-300, 1989.
- [6] B.T.Kelley and V.K.Madisetti, The Fast Discrete Radon Transform - I: Theory, *IEEE Transactions on Image Processing*, vol. 2, no. 3, July 1993.