

Robust Recognition of Scaled Shapes using Pairwise Geometric Histograms

A.P.Ashbrook, N.A.Thacker, P.I.Rockett and C.I.Brown
Electronic Systems Group
University of Sheffield
United Kingdom
A.P.Ashbrook@sheffield.ac.uk

Abstract

The recognition of shapes in images using Pairwise Geometric Histograms has previously been confined to fixed scale shape. Although the geometric representation used in this algorithm is not scale invariant, the stable behaviour of the similarity metric as shapes are scaled enables the method to be extended to the recognition of shapes over a range of scale. In this paper the necessary additions to the existing algorithm are described and the technique is demonstrated on real image data. Hypotheses generated by matching scene shape data to models have previously been resolved using the generalised Hough transform. The robustness of this method can be attributed to its approximation of maximum likelihood statistics. To further improve the robustness of the recognition algorithm and to improve the accuracy to which an objects location, orientation and scale can be determined the generalised Hough transform has been replaced by the probabilistic Hough transform.

1 Introduction

The existing Pairwise Geometric Histogram (PGH) algorithm provides a robust solution for the recognition of arbitrary, 2D shape in a manner which is invariant to position and orientation but which is not invariant to changes in shape size or affine transformations of the shape data [1]. Current work in this area has focused on establishing this technique as a generic solution to the problem of rigid object recognition by tackling the scale and view constraints. A recent analysis has shown that the algorithm is highly scalable and may be applied to recognition tasks involving very large numbers of models [2, 3] – this is a necessary requirement if 3-dimensional objects are represented by large numbers of 2-dimensional views. The technique has also been shown to be complete in the sense that a model shape may be reconstructed from the set of PGHs which describe it [4] confirming that the complete representation for a shape is unambiguous.

Recognition is performed by finding edge data in scene images which is consistent with edge data from stored model shapes. The use of edge information

ensures robust performance as these features are reliably extracted from images. The complete algorithm comprises the following steps:

1. Model image data (during training) and scene image data (during recognition) are processed to extract edges which are then approximated by line segments.
2. Model histograms (during training) and scene histograms (during recognition) are constructed for each line segment (reference line) by comparing this line to all other lines and making histogram entries according to the measured relative angle and perpendicular distances. This representation encodes local shape geometry in a manner which is invariant to rotation and translation.
3. Scene line labelling is performed by finding good matches between scene histograms and model histograms using the Bhattacharyya metric. This statistical metric is appropriate as PGH's are joint probability distributions of local shape geometry. Each scene line label is a hypothesis of the scene content.
4. Object classifications are confirmed by finding consistent labelling (hypotheses) within a scene using a hough transform.

In this paper we explain how the algorithm can be extended to the recognition of scaled shapes with reasonable computational overhead. Although the PGH representation is not scale invariant an analysis of the effect of scale on the similarity metric shows that it is stable across ranges of scale. This property may be utilised to recognise shapes at *any* scale.

When the constraint of fixed scale shape is removed the process of resolving hypotheses becomes more demanding. Previously this was done using generalised Hough transforms. The robustness of this method can be attributed to its approximation of maximum likelihood statistics. To further improve the robustness of this stage of recognition it has now been replaced by an implementation of the probabilistic Hough transform [5].

2 The Effect of Scale on the Similarity Metric

Good correspondences between scene lines and model lines are found by matching scene PGHs to model PGHs using the Bhattacharyya similarity metric. The effect that scaling shape data has on the representation and, consequently, on the similarity metric is described in this section.

Because distances are encoded in PGHs the representation is not scale invariant so matching histograms constructed from the same shape data but at different scales does not result in perfect match scores. The effect on a histogram as the shape data it represents is continuously increased in scale is a continuous stretching of the entries along the perpendicular distance axis. This is demonstrated in figure 1 (a) for histograms constructed for the line data in figure 1 (b) as it is scaled.

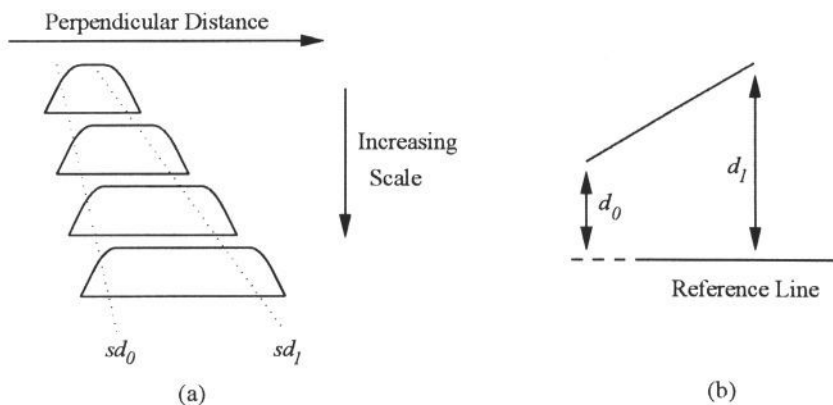


Figure 1: (a) Cross sections through PGH entries for the line data at scale s . (b) The line data used in this example

Because PGHs change smoothly as the shape data is scaled the similarity metric degrades smoothly as a function of scale. This is shown in figure 2 for a geometric histogram constructed from real shape data (The line data used in this example was taken from the shape model used in section 5). Because good match scores are attained over a reasonable range of scale individual PGHs, although not scale invariant, can effectively represent a range of scale. Although the change in the similarity metric is always smooth the rate at which it degrades, and therefore the range of scale that can be reasonably represented, depends upon the actual line data encoded in the histograms.

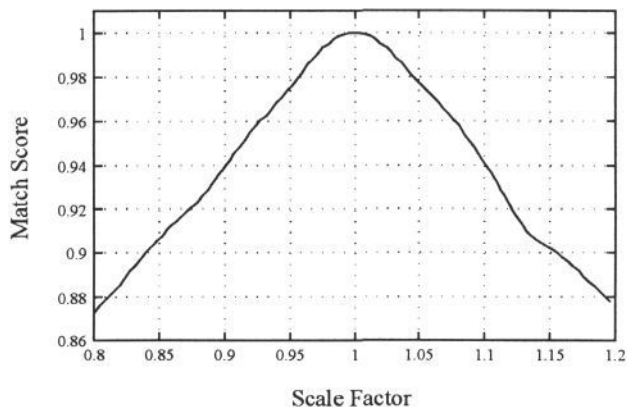


Figure 2: The effect of scale on the similarity metric

To improve the performance of recognition in the presence of scene clutter and to provide some robustness to shape deformation, PGHs are constructed from lines which lie in a locally defined region – typically a circular window centered on the reference line is used. To maintain the stability of the similarity metric it is important that lines entering or leaving the region of interest as the shape data

is scaled do not introduce sudden changes in the representation. By weighting each histogram entry by the length of line within the region of interest, each entry degrades smoothly to nothing as the line data responsible for it is scaled beyond the region of interest.

3 Scene Line Labelling

Each scene line is labelled with a hypothesis of the scene content according to the best correspondences found between it and the set of model lines. The positions and orientations of models suggested by these labels are implied by the positions and orientations of the scene lines themselves. When the constraint of fixed scale shape is removed a strategy has to be employed which ensures that good correspondences are still identified. Labels for variable scale data must be extended to include a hypothesis of the scale of the scene line.

As previously shown, although the PGH representation is not scale invariant a single PGH is capable of representing a range of scale because of the stability of the similarity metric. This property can be utilised to recognise shapes at *any* scale.

3.1 The Training Algorithm

A model line may be represented across a specified range of scale by storing a number of histograms for the line at suitable scale intervals. The actual number and scale of stored histograms depends upon both the scale range specified and upon how closely all scales within the range need to be represented.

The training algorithm we have used takes a range of scale defined from S_{min} to S_{max} and a threshold on the similarity metric, D_{th} , and stores model histograms in such a way that any known shape which is later encountered within the scale range will result in match scores of at least D_{th} , in the absence of noise.

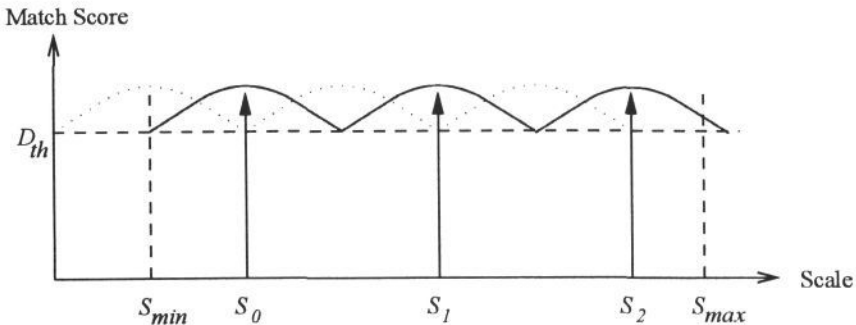


Figure 3: Adding histograms to represent a new model line across a range of scale

The algorithm is depicted in figure 3. Initially a temporary histogram is constructed at scale S_{min} and this is compared to temporary histograms constructed at larger scales until the scale S_0 , where the match score equals D_{th} , is found. The most efficient way to do this is to use a bisections search across the scale range. By

symmetry, a histogram stored at scale S_0 is guaranteed to result in match scores above D_{th} when compared to histograms constructed at scales between S_{min} and S_0 . The stored histogram is then compared to histograms constructed at scales greater than S_0 until the scale where *this* match score equals D_{th} is found. A temporary histogram is then constructed at this scale and then in a similar manner this used to find the scale S_1 where the second histogram is stored. This process is continued until the complete range of scale is covered.

After matching scene PGHs to model PGHs which are stored in this way, the scales associated with the best matches provide estimates of the scale of the scene data. If models can appear in a scene at any scale within the specified range then each estimate will have a corresponding uniform error which extends either side of the scale estimate to the points where the temporary histograms were constructed during training. Figure 4 shows the best match scores obtained when matching a PGH constructed from a model line at different scales to model histograms which were stored using this algorithm (D_{th} was specified as 0.9). For this scale range and similarity metric threshold an average of 6 histograms were stored per model line, so labelling shapes over this scale range results in a six-fold increase in computation.

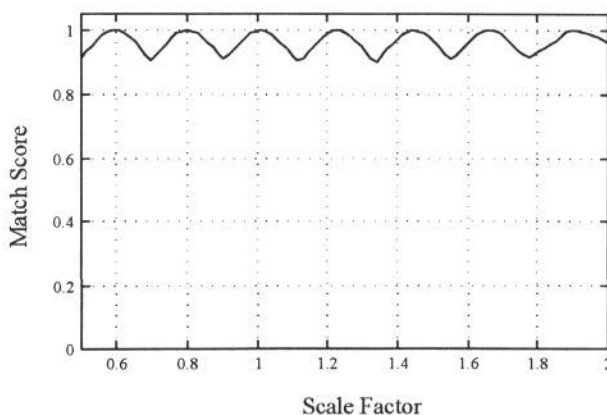


Figure 4: The best match scores as a function of scale

4 The Probabilistic Hough Transform

The recognition process is completed by finding scene line labelling which is consistent with stored models. For fixed scale data each label hypothesises the presence of a shape at some position and orientation. Consistency amongst these parameters for a set of scene lines was previously found using generalised Hough transforms. For variable scale shape each scene line label also provides a scale hypothesis and this additional information must also be considered.

Although the generalised Hough transform used in the existing PGH algorithm performs reliably for fixed scale data an improved method has to be adopted to account for the effect of scale error in a coherent fashion when the fixed scale

constraint is removed. To achieve this we have implemented a probabilistic Hough transform (PHT) which takes proper account of errors, resulting in improved robustness and more accurate determination of model position, orientation and scale. In fact, because the PHT is derived from robust, maximum likelihood statistics the results it gives are the most probable in a maximum likelihood sense. To constrain the entries made into the PHT a single entry is made for each *pair* of scene lines and only when the pair of lines are in reasonable agreement about the position, orientation and scale of the model.

The probabilistic Hough transform is in general constructed from a set of N input features. In this application each input feature is a pair of scene line labels which hypothesises the position, orientation and scale of a model in the scene. The probabilistic Hough transform $H(m)$ used to find the position m of a model in a scene is given by the expression:

$$H(m) = \sum_{i=0}^{N-1} \log [P(h_i|m)] + K \quad (1)$$

Where h_i is the position of the model hypothesised by the i_{th} pair of scene line labels and $P(h_i|m)$ is conditional probability density function that a pair of scene lines will hypothesise a position h_i given that the model is actually positioned at m . In other words, the error on the model location given a pair of scene lines.

4.1 Modelling the Error on the Model Location Hypotheses

The hypothesised location of a model is subject to error because of the variability of the line segmentation process and because of the uniform error on the scale estimates. By modelling these errors the entry which should be made in the PHT for a pair of scene lines can be determined.

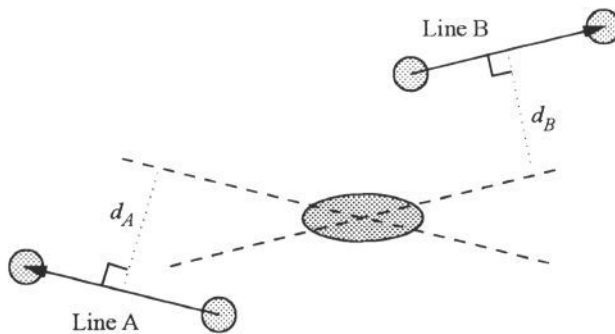


Figure 5: The line segmentation error model

The line segmentation process can be approximately modelled by assuming gaussian errors on the position of the line end points, as shown in figure 5. During training the perpendicular distance, d , from each model line to an arbitrary reference point on the model shape is recorded. Consequently, for each pair of scene

lines, extended lines at the appropriate perpendicular distance will intersect at the hypothesised position of the model. Applying standard error propagation [6] the error on the point of intersection ΔI is given by:

$$\Delta I = \nabla f^T C_{end} \nabla f \quad (2)$$

Where f describes the point of intersection in terms of the two pairs of endpoints and C_{end} is endpoint error covariance matrix.

The validity of this model is checked by measuring the error on the position of a shape predicted by pairs of scene lines. Figure 6 shows a frequency distribution of the normalised residual errors for a typical scene. The approximately gaussian shape of this plot with an extended tail implies that the error model is sufficiently accurate.

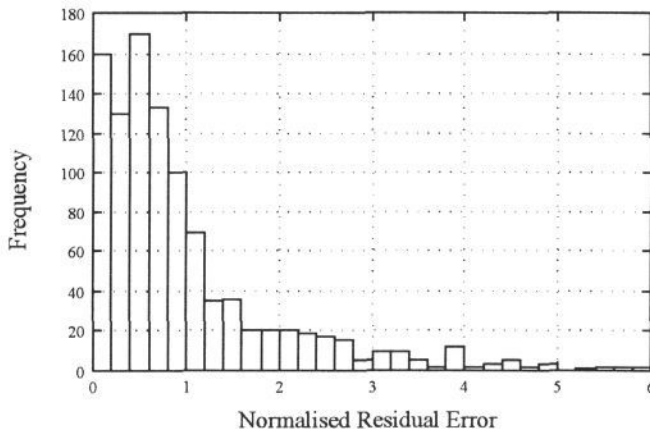


Figure 6: Distribution of normalised residual errors on predicted shape position

The effect of the uniform scale error, as described in section 3, on the hypothesised model position is demonstrated in figure 7. The scales and scale errors associated with Lines A and B constrain the position of the model to lie within the bands defined between the dotted lines. However, if both of these lines belong to the same shape then the model position must satisfy both constraints and lie in the shaded region. Furthermore, if the scene lines belong to the same shape then they must also be at the same scale which gives the constraint represented by the dashed line. The section of the dashed line which intersects the shaded region satisfies all of the constraints so if both lines A and B are from the same model its position must lie on this section.

Because the scale and segmentation errors are independent $P(h_i|m)$ may be determined by convolving the segmentation and scale error functions. This is then used to make entries in the PHT.

4.2 The Orientation and Scale Hough Transforms

The PHT described is used to locate models using the positions, orientations and scales hypothesised by scene line labels, but it does not determine the orientations

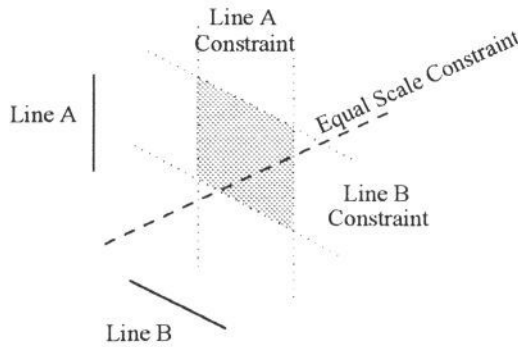


Figure 7: The possible positions of a shapes origin constrained by a pair of lines

or scales of models explicitly. This is done separately using 1-parameter Hough transforms.

For each model position determined a 1-parameter orientation Hough transform and a 1-parameter scale Hough transform is constructed then scene lines which are consistent with the model position are used to make orientation and scale entries. The orientation entry is determined from the difference in orientation between the scene line and the model line to which it matched. The scale entry is determined by comparing the perpendicular distance from the scene line to the model position to this same distance in the model itself. Peaks in these Hough transforms give the orientation and scale of the model at this position in the scene.

5 Demonstration

The modifications to the algorithm are demonstrated here on cluttered scenes containing views of real 3-dimensional objects at different scales. To allow control of the scale of the data the scene was imaged at one scale then scaled appropriately before being segmented into lines. In these examples the models were stored over a range of scale from 0.5 to 2.0, requiring on average six histograms to be stored per model line. Figure 8 shows that the person model has been successfully located in scenes of scale 0.75 and 1.5. The probabilistic Hough transforms constructed for the person model for each of these scene are shown in figure 9 (the intensity of the background entries have been artificially increased so that they can be seen). Finally, the scale Hough transforms for the located models are shown in figure 10 – note that positions of the peaks correspond to the scales of the models.

6 Conclusions

Object recognition using PGHs has previously been confined to shapes of fixed scale because the geometric representation used in the algorithm is not scale invariant. The similarity metric used to match scene and model line data, however, has been shown to behave in a stable manner over ranges of scale. This property enables the recognition of shape over scale ranges with reasonable computational

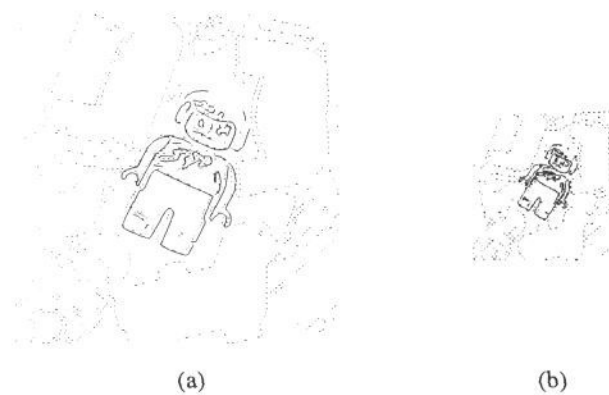


Figure 8: Scenes containing views of model shapes at (a) a scale of 1.5, (b) a scale of 0.75

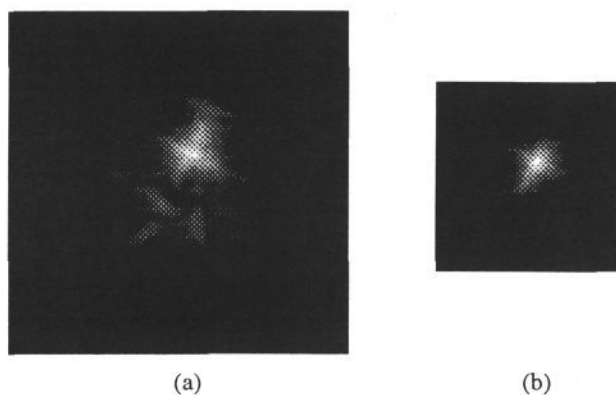


Figure 9: The probabilistic Hough transforms used to locate the models in the scenes above

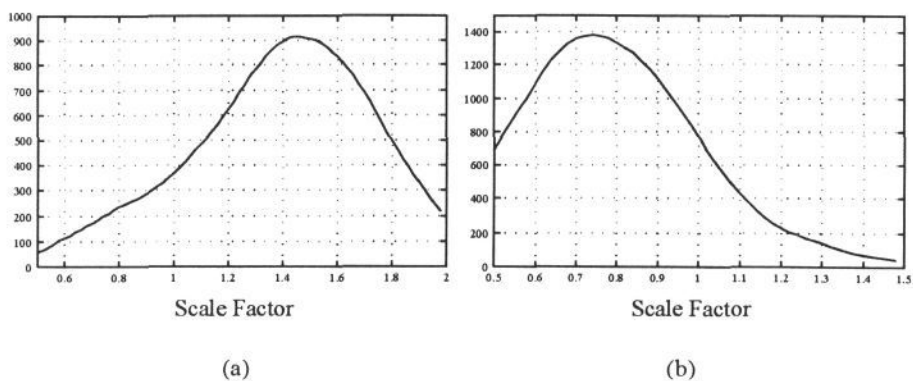


Figure 10: The scale Hough transforms for the models located at (a) a scale of 1.5, (b) a scale of 0.75

overhead. In fact the lack of scale invariance proves to be useful as this provides a constraint on the position of a model in a scene which is used when combining individual model hypotheses to complete the recognition process.

To take correct account of the constraints imposed by individual model hypotheses the generalised Hough transform (used previously) has been replaced by a probabilistic Hough transform. This not only takes account of hypothesis parameters such as position and scale but also takes account of the errors on these quantities. This both improves robustness of the algorithm and allows the position, orientation and scale of encountered models to be determined more accurately (the most probable parameters are determined in a maximum likelihood sense).

Finally, the improved algorithm has been successfully demonstrated on views of real objects at different scales in highly cluttered scenes.

Acknowledgements

This work was partly supported by SERC/ESPRC contract No. GR/J104464 and the University of Sheffield.

References

- [1] A.Evans, N.A.Thacker, J.E.W.Mayhew. "The use of Geometric Histograms for Model-Based Object Recognition", Proc 4th BMVC, Guildford, pp429-438, Sept 1993.
- [2] A.P.Ashbrook, N.A.Thacker and P.I.Rockett. "Pairwise Geometric Histograms: A Scaleable Solution for the Recognition of 2D Rigid Shape", Proc SCIA95, June 1995.
- [3] A.P.Ashbrook, N.A.Thacker and P.I.Rockett. "Multiple Shape Recognition using Pairwise Geometric Histogram based Algorithms", Proc IPA95, July 1995.
- [4] P.A.Riocreux, N.A.Thacker and R.B.Yates. "An Analysis of Pairwise Geometric Histograms for View-Based Object Recognition". Proc 5th BMVC, York, pp75-84, Sept 1994.
- [5] R.S.Stephens. "A Probabilistic Approach to the Hough Transform". Proc BMVC, 1990.
- [6] N.A.Thacker "Statistics and Estimation in Algorithmic Vision". Notes from 1995 ESPRC Machine Vision Summer School.