Demisting the Hough Transform for 3D Shape Recognition and Registration

Oliver J. Woodford
Minh-Tri Pham
Atsuto Maki
Frank Perbet
Björn Stenger

Cambridge Research Laboratory,
Toshiba Research Europe Ltd.
Cambridge, UK
firstname.lastname@crl.toshiba.co.uk
www.toshiba-europe.com/research/crl/cvg/

1 Overview

The Hough transform [2], named after Hough's 1962 patent describing a method for detecting lines in images, has since been generalized to detecting, as well as recognizing, many other objects: parameterized curves , arbitrary 2D shapes, cars [4], pedestrians [1], hands and 3D shapes [3, 5, 6], to name but a few. This popularity stems from the simplicity and generality of the first step of the Hough transform—the conversion of *features*, found in the data space, into sets of *votes* in a Hough space, parameterized by the pose of the object(s) to be found. The second stage of the Hough transform then simply sums the likelihoods of the votes at each location in Hough space, and selects the modes of the resulting distribution.

A problem with this latter step is that the summation can create modes where there are only a few outlier votes. Indeed, as stated in [1], no probabilistic interpretation that fully explains this approach has yet been provided. A second problem is that, given a required accuracy, the size of the Hough space is exponential in its dimensionality. The application we are concerned with, object recognition and registration from 3D geometry (here, point clouds), suffers significantly from both these problems. The Hough space, at 8D (one dimension for class, three for rotation, three for translation and one for scale), is to our knowledge the largest to which the Hough transform has been applied, and the feature-to-vote conversion generates a high proportion of incorrect votes, creating a "mist" of object likelihood throughout that space, as shown in figure 1(a).

In the face of this adversity, we have developed two important contributions which enable inference on this task, and potentially many others, using the Hough transform to be both feasible and accurate:

- We introduce the *intrinsic Hough transform*, which substantially reduces memory and computational requirements in applications with a high dimensional Hough space.
- We introduce the *minimum-entropy Hough transform*, which greatly improves the precision and robustness of the Hough transform.

2 Methods

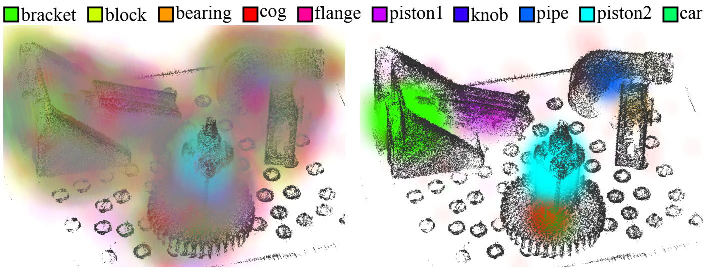
We note that while the Hough space increases exponentially with its dimensionality, the number of votes generated in applications using the Hough transform generally do not, implying that higher dimensional Hough spaces are often sparser. We exploit this sparsity by sampling the Hough space only at locations where the probability is likely to be non-zero—at the locations of the votes themselves. Since the votes are intrinsic to the distribution, we call this the intrinsic Hough transform.

Recently Barinova *et al*. [1] introduced an alternative vote-based inference framework which exploits the assumption that *only one vote cast by each feature is correct*, with the result that correct votes are able to explain away incorrect votes from the same feature for the first time. We exploit this same assumption within the Hough transform, by optimizing the weights of votes w.r.t. the information entropy of the vote distribution. A lower entropy distribution contains less information, making it more peaky and hence having more votes in agreement. Since information in Hough space is the location of objects, minimizing entropy constrains features to be generated by as few objects as possible, enforcing Occam's razor. We call this approach the minimum-entropy Hough transform.

Our paper presents a local method for minimizing entropy using iterated conditional modes, optimizing weights for each feature in series. An initialization strategy is proposed which helps reach a good minimum.

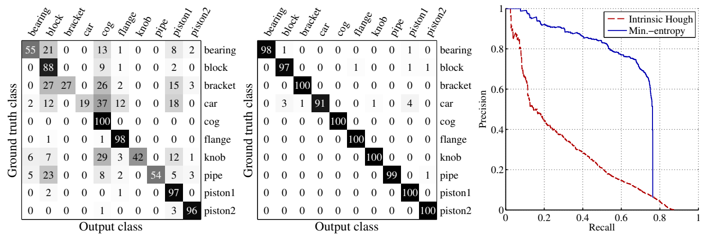
3 Results

Our paper presents experiments on a dataset of 1000 test point clouds of 10 different objects, captured using a multi-view stereo method, and evaluates both recognition and registration performance of our methods.



(a) Standard Hough (b) Min.-entropy Hough

Figure 1: **Demisting the Hough transform.** Posterior distributions over translation and ten object classes (six of which are present in the scene), with scale and rotation marginalized out.



(a) Intrinsic Hough (b) Min.-entropy Hough (c) Precision-recall

Figure 2: **Quantitative results.** (a,b) Recognition confusion matrices and (c) precision-recall curves for correct recognition and registration.

While intrinsic Hough allows the Hough transform to be applied to the full 8D Hough space of this problem, the minimum-entropy Hough transform dramatically improves recognition rates (figure 2(b)), as well registration performance, as indicated by the improvement in precision vs. recall shown in figure 2(b). It can be seen qualitatively in figure 1 that this improvement comes from the latter method "explaining away" incorrect votes.

4 Conclusion

We present two extensions of the Hough transform, which are not task specific; they can be applied, either together or independently, to any application that does or is able to use the standard Hough transform. We demonstrate, through applying these extensions to the task of 3D shape recognition and registration, that the assumption that only one vote generated by each feature is correct is a powerful constraint in vote-based frameworks, which can dramatically improve inference by "clearing the mist" of incorrect votes.

[1] O. Barinova, V. Lempitsky, and P. Kohli. On detection of multiple object instances using Hough transforms. In *Proceedings of CVPR*, 2010.

[2] R. O. Duda and P. E. Hart. Use of the Hough transformation to detect lines and curves in pictures. *Commun. ACM*, 15:11–15, 1972.

[3] J. Knopp, M. Prasad, G. Willems, R. Timofte, and L. Van Gool. Hough transform and 3D SURF for robust three dimensional classification. In *Proceedings of ECCV*, pages 589–602, 2010.

[4] B. Leibe, A. Leonardis, and B. Schiele. Robust object detection with interleaved categorization and segmentation. *International Journal of Computer Vision*, 77(1-3):259–289, 2008.

[5] M.-T. Pham, O. J. Woodford, F. Perbet, A. Maki, B. Stenger, and R. Cipolla. A new distance for scale-invariant 3D shape recognition and registration. In *Proceedings of ICCV*, 2011.

[6] F. Tombari and L. Di Stefano. Object recognition in 3D scenes with occlusions and clutter by Hough voting. In *Proceedings of PSIVT*, pages 349–355, 2010.