Int J Comput Vis
DOI 10.1007/s11263-013-0623-2

Demisting the Hough Transform for 3D Shape Recognition and Registration

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

Received: 19 September 2012 / Accepted: 1 April 2013

Abstract In applying the Hough transform to the problem of 3D shape recognition and registration, we develop two new and powerful improvements to this popular inference method. The first, *intrinsic Hough*, solves the problem of exponential memory requirements of the standard Hough transform by exploiting the sparsity of the Hough space. The second, *minimum-entropy Hough*, explains away incorrect votes, substantially reducing the number of modes in the posterior distribution of class and pose, and improving precision. Our experiments demonstrate that these contributions make the Hough transform not only tractable but also highly accurate for our example application. Both contributions can be applied to other tasks that already use the standard Hough transform.

Keywords Hough transform · Object recognition · 3d shape · Registration

1 Introduction

The Hough transform Duda and Hart (1972), named after Hough (1962) patent describing a method for detecting lines in images, has since been generalized to detecting, as well as recognizing, many other objects or instances: parameterized curves (Duda and Hart 1972), arbitrary 2D shapes (Ballard 1981), object motions (Bober and Kittler 1993), cars (Gall and Lempitsky 2009; Leibe et al. 2008) pedestrians (Barinova et al. 2010; Gall and Lempitsky 2009), hands Okada (2009) and 3D shapes (Knopp et al. 2010; Pham et al. 2011; Tombari and Di Stefano 2010), 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. Various different approaches to learning this feature-to-vote conversion function have been proposed, including the *implicit shape model* Leibe et al. (2008) and *Hough forests* (Gall and Lempitsky 2009; Okada 2009).

The second stage of the Hough transform simply sums the likelihoods of the votes at each location in Hough space, then selects the modes. One problem with this step is that the summation can create modes where there are only a few outlier votes. 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 (R&R) 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 Fig. 1a.

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.

O. J. Woodford (✉) · M-T. Pham · A. Maki · F. Perbet · B. Stenger
Toshiba Research Europe Ltd., 208 Cambridge Science Park, Milton Road, Cambridge CB4 0GZ, UK
e-mail: oliver.woodford@crl.toshiba.co.uk