

Bayesian Classification for the Statistical Hough Transform

Rozenn Dahyot

School of Computer Science and Statistics
Trinity College Dublin, Ireland
Rozenn.Dahyot@cs.tcd.ie

Abstract

We have introduced the Statistical Hough transform [2] that extends the standard Hough transform by using a kernel mixture as a robust alternative to the 2 dimensional accumulator histogram. This work develops further this framework by proposing a Bayesian classification scheme to associate the spatial coordinates (x, y) to one particular class defined in the Hough space (θ, ρ) . In a first step, we segment the Hough space into meaningful classes. Then using the inverse Radon transform, we backproject the different classes into the image space. We illustrate our approach on a synthetic image and on real images.

1 Introduction

To overcome histogram limitations in the computation of the Standard Hough transform, we have proposed a new kernel based modelling of the density $p_{\theta\rho}(\theta, \rho)$ [2]. All the observations available are used in this modelling and no pre-segmentation of the edges is necessary. In this new formalism, the maxima in the Hough space cannot be associated directly with their contributing pixels (as with the standard Hough transform) since all the pixels have contributed to them at different degrees. We propose here to perform an unsupervised segmentation of our kernel estimate of the Hough transform, and using the inverse Radon transform, find the associated pixels in the spatial domain. We start in paragraph 2 by a short review, and we introduce our unsupervised dual segmentation scheme (in both the Hough and spatial domains) in section 3.

2 Statistical Hough Transform

Many works have been published on the Hough Transform since its first publication [4]. It is a very

robust estimator for estimating lines in a set of points [3]. As a consequence, it has been used extensively to recover aligned segments amongst the contours of images. Recent works [5, 1] have proposed to consider the angle of the gradient as an additional information for computing the Hough transform. In addition the variance of the angle is also computed and we have proposed in [2], to use it as variable bandwidths when estimating the Hough transform with a mixture of kernels. This new Statistical Hough Transform is replacing the discrete 2D histogram of the Standard Hough Transform, by a smoother continuous kernel estimate. We have shown [2] that such modelling was encoding better the alignment content of images.

First we define the sets of observations $\mathcal{S}_{\theta xy} = \{(\theta_i, x_i, y_i)\}_{i \in [1 \dots N]}$ and $\mathcal{S}_{xy} = \{(x_i, y_i)\}_{i \in [1 \dots N]}$, with θ_i is the angle direction of the gradient, and (x_i, y_i) is the spatial position on the pixel i of an image. Using the Bayes formula, the statistical Hough transform proposes to model the joined probability density function of the hough variables (θ, ρ) and the spatial variables (x, y) by:

$$p_{\theta\rho xy}(\theta, \rho, x, y) = p_{\rho|\theta xy}(\rho|\theta, x, y) \cdot p_{\theta xy}(\theta, x, y) \quad (1)$$

When x, y, θ are known, the variable ρ is deterministic since we have the relation:

$$\rho = x \cos \theta + y \sin \theta \quad (2)$$

Therefore we propose to model the conditional probability:

$$p_{\rho|\theta xy}(\rho|\theta, x, y, \mathcal{S}_{\theta xy}) = \delta(\rho - x \cos \theta - y \sin \theta) \quad (3)$$

where $\delta(\cdot)$ is the dirac distribution. As a consequence, only $p_{\theta xy}(\theta, x, y)$ is to estimate using kernels on the set of observations $\mathcal{S}_{\theta xy} = \{(x_i, y_i, \theta_i)\}_{i=1, \dots, N}$:

$$\hat{p}_{\theta xy}(\theta, x, y|\mathcal{S}_{\theta xy}) = \sum_{i=1}^N \frac{1}{h_{x_i}} k_x \left(\frac{x - x_i}{h_{x_i}} \right) \frac{1}{h_{y_i}} k_y \left(\frac{y - y_i}{h_{y_i}} \right) \frac{1}{h_{\theta_i}} k_{\theta} \left(\frac{\theta - \theta_i}{h_{\theta_i}} \right) p_i \quad (4)$$

where p_i is the prior on the observation (x_i, y_i, θ_i) . In this paper, all the pixels are selected with equiprobable prior $p_i = \frac{1}{N}$, $\forall i = 1, \dots, N$. By integration with respect to the spatial variables (x, y) , an estimate of the Hough transform $p_{\theta\rho}(\theta, \rho)$ can be computed:

$$\hat{p}_{\theta\rho}(\theta, \rho | \mathcal{S}_{\theta xy}) = \sum_{i=1}^N \frac{1}{h_{\theta_i}} k_{\theta} \left(\frac{\theta - \theta_i}{h_{\theta_i}} \right) R_i(\theta, \rho) p_i \quad (5)$$

where $R_i(\theta, \rho)$ is the Radon transform of the spatial kernels:

$$R_i(\theta, \rho) = \int \int \delta(\rho - x \cos \theta - y \sin \theta) \frac{1}{h_{x_i}} k_x \left(\frac{x - x_i}{h_{x_i}} \right) \frac{1}{h_{y_i}} k_y \left(\frac{y - y_i}{h_{y_i}} \right) dx dy \quad (6)$$

All kernels k_x , k_y and k_{θ} are chosen Gaussian. The spatial bandwidths are naturally set to the resolution grid of the image $h_{x_i} = h_{y_i} = 1$, $\forall i$ and the variable bandwidths of the angle are estimated from the image data [2]. If no observation is available for θ (i.e. the set of observations is \mathcal{S}_{xy}), then we simplify the kernel for the angle by a uniform distribution such that:

$$\hat{p}_{\theta\rho}(\theta, \rho | \mathcal{S}_{xy}) = \frac{1}{\pi} \sum_{i=1}^N R_i(\theta, \rho) p_i \quad (7)$$

3 Bayesian classification

We want to extract and recover all the alignment content from an image. This is performed in 4 steps. We illustrate them using the image *diamond* (see figure 4) with additional centered Gaussian noise ($\sigma_n = 20$).

1. **Estimates of the pdfs.** For an image I , compute both the pdfs $\hat{p}_{\theta\rho}(\theta, \rho | \mathcal{S}_{\theta xy})$ and $\hat{p}_{\theta\rho}(\theta, \rho | \mathcal{S}_{xy})$ (see figure 1). Since all the pixels are taken, the pdf $\hat{p}_{\theta\rho}(\theta, \rho | \mathcal{S}_{xy})$ gives an estimate of the distribution if no aligned content was occurring in the image I .
2. **Threshold of the pdf.** We extract the relevant regions in $\hat{p}_{\theta\rho}(\theta, \rho | \mathcal{S}_{\theta xy})$ that are significant of aligned content in the image. This is done by fitting robustly the surface $\hat{p}_{\theta\rho}(\theta, \rho | \mathcal{S}_{xy})$ to $\hat{p}_{\theta\rho}(\theta, \rho | \mathcal{S}_{\theta xy})$. In particular, the following values are estimated:

$$\hat{\mu} = \text{median} \left\{ \begin{array}{l} \hat{p}_{\theta\rho}(\theta, \rho | \mathcal{S}_{\theta xy}) \\ \hat{p}_{\theta\rho}(\theta, \rho | \mathcal{S}_{xy}) \end{array} \right\} \quad (8)$$

and the robust Median Absolute Deviation:

$$\hat{\sigma} = 1.48 \text{ median} \left| \frac{\hat{p}_{\theta\rho}(\theta, \rho | \mathcal{S}_{\theta xy})}{\hat{p}_{\theta\rho}(\theta, \rho | \mathcal{S}_{xy})} - \hat{\mu} \right| \quad (9)$$

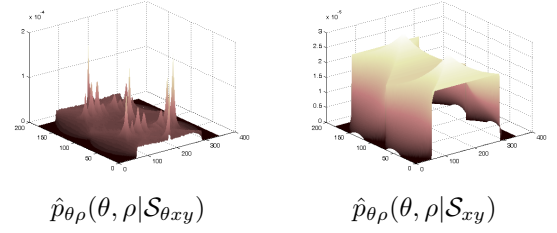


Figure 1. Density estimates in the Hough space.

Then we extract the alignment content by thresholding the pdf $\hat{p}_{\theta\rho}(\theta, \rho | \mathcal{S}_{\theta xy})$ such that:

$$\hat{p}_{\theta\rho A}(\theta, \rho, A) = \begin{cases} \hat{p}_{\theta\rho}(\theta, \rho | \mathcal{S}_{\theta xy}) & \text{if } \hat{p}_{\theta\rho}(\theta, \rho | \mathcal{S}_{\theta xy}) > (\hat{\mu} + 3\hat{\sigma}) \hat{p}_{\theta\rho}(\theta, \rho | \mathcal{S}_{xy}) \\ 0 & \text{otherwise} \end{cases} \quad (10)$$

A indicates the class *alignment* content and \bar{A} is its complement defined as:

$$\hat{p}_{\theta\rho A}(\theta, \rho, \bar{A}) = \begin{cases} \hat{p}_{\theta\rho}(\theta, \rho | \mathcal{S}_{\theta xy}) & \text{if } \hat{p}_{\theta\rho}(\theta, \rho | \mathcal{S}_{\theta xy}) \leq (\hat{\mu} + 3\hat{\sigma}) \hat{p}_{\theta\rho}(\theta, \rho | \mathcal{S}_{xy}) \\ 0 & \text{otherwise} \end{cases} \quad (11)$$

Hence we have $\hat{p}_{\theta\rho}(\theta, \rho | \mathcal{S}_{\theta xy}) = \hat{p}_{\theta\rho A}(\theta, \rho, A) + \hat{p}_{\theta\rho A}(\theta, \rho, \bar{A})$. Figure 2 shows the results of the thresholding: in (a) the thresholding map (view from above) in the Hough Space, and the corresponding estimate $\hat{p}_{\theta\rho A}(\theta, \rho, A)$.

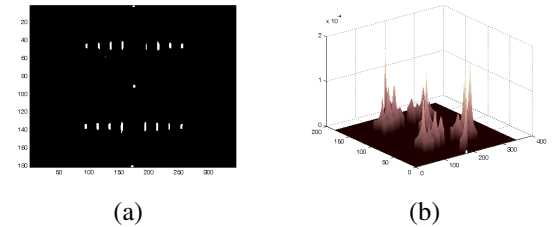


Figure 2. Segmentation in the $\theta\rho$ -space.

3. **Segmentation of $\hat{p}_{\theta\rho A}(\theta, \rho, A)$.** We first detect all maxima of $\hat{p}_{\theta\rho A}(\theta, \rho, A)$ and sort them in descending order. For each mode, we aim at segmenting the corresponding bump [6]. The class in the Hough space is computed by thresholding at

half height of the mode. In figure 3(a), the results of the segmentation in the Hough space is shown: the segmentation is unsupervised and gives here 20 classes $\{\mathcal{C}_j^{\theta\rho}\}_{j=1,\dots,20}$, each associated with a particular color.

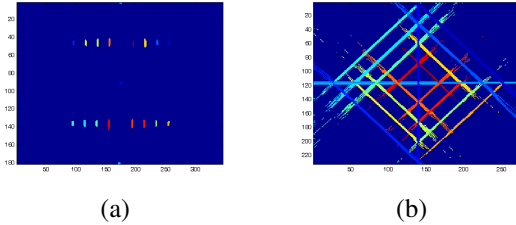


Figure 3. Classification in the $\theta\rho$ -space (a) and xy -space (b).

4. **Backprojection in the xy -space using Bayesian Classification.** The class $\mathcal{C}_j^{\theta\rho}$ in the Hough space is associated to the set of pixels \mathcal{C}_j^{xy} . Each pixel (x, y) is labelled as follow:

$$(x, y) \rightarrow \mathcal{C}_j^{xy} = \arg \max_{\mathcal{C}_k^{xy}} \hat{p}_{xy}(x, y, \mathcal{C}_k^{xy}) \quad (12)$$

where $\hat{p}_{xy}(x, y, \mathcal{C}_j^{xy})$ is computed by inverse Radon transform of $\hat{p}_{\theta\rho}(\theta, \rho, \mathcal{C}_j^{\theta\rho} | \mathcal{S}_{\theta xy})$. In figure 3(b), the result of the classification in the spatial domain is shown using the same color code as in 3(a). We note that all 19 linear contours are well detected.

4 Experimental Results

We illustrate our approach on several images in figure 4. The first is the image *diamond* with a very important Gaussian noise (not shown on the image). We can see that some relevant regions have been segmented in the Hough space and their backprojection in the spatial domain allows to recover the corresponding straight edges. The edges that are missed have low gradient magnitude. Therefore, because of the important noise level, the corresponding bandwidths for the angle in the Statistical Hough transform $\hat{p}_{\theta\rho}(\theta, \rho | \mathcal{S}_{\theta xy})$ are then very large, and no peaks can then be detected in the Hough space [2]. The second and the third images, *road* and *aerial*, present less well defined straight edges. For the *road* image, most of the aligned content is recovered apart from one side of the traffic sign. The aligned content in the *aerial* image is more difficult to extract but we note that our segmentation scheme allows the extraction of the main roads. The number of classes that have

been found in the Hough space is also reported. Some classes $\mathcal{C}_j^{\theta\rho}$, segmented in the Hough space, do not have any associated data in the spatial domain. The number of non-empty classes in the xy -space (also reported in fig. 4) can therefore be lower.

We have re-ordered the classes in the spatial domain in descending order of their number of pixels. In figure 5, the four main classes found in spatial domain are shown for the real images from figure 4. Main borders of the road are easily detected. The *aerial* image is more complex (i.e. more cluttered), however the main roads are also well detected.

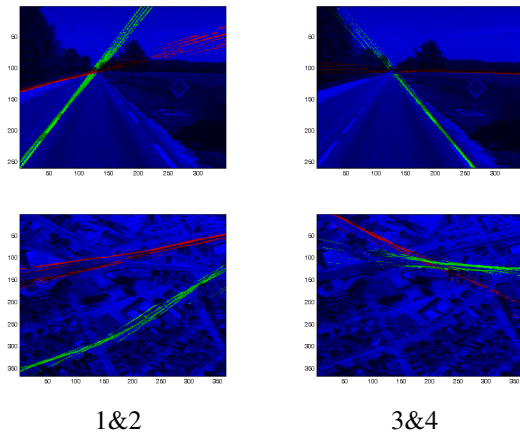


Figure 5. Main 4 classes \mathcal{C}_j^{xy} in xy -space.

Comparison with the Standard Hough transform.

When using the Standard Hough transform, edges are first segmented. This preliminary step can be understood in our statistical framework as choosing binary priors $\{p_i\}$. Edge segmentation is a sensitive operation especially in noisy images and some edges may be completely lost. Here, we have chosen equiprobable priors which remove the need of segmenting the edges. The Statistical Hough Transform is a more generic framework and it allows for smoother estimates of $p_{\theta\rho}(\theta, \rho)$ when choosing Gaussian kernels for instance. Its computation is however more time consuming as each pixel from the original images is used (not only the edges), and the infinite tail of the Gaussian kernels has also to be taken into account. As an alternative, other kernels that have a finite support (e.g. Epanechnikov) could be chosen to speed up the process.

The classification scheme presented in this paper considers the bumps of the peaks of the estimate of $p_{\theta\rho}(\theta, \rho)$. This is a different approach from using only the maxima as it is done in the Standard Hough transform for recovering the lines. The results presented here

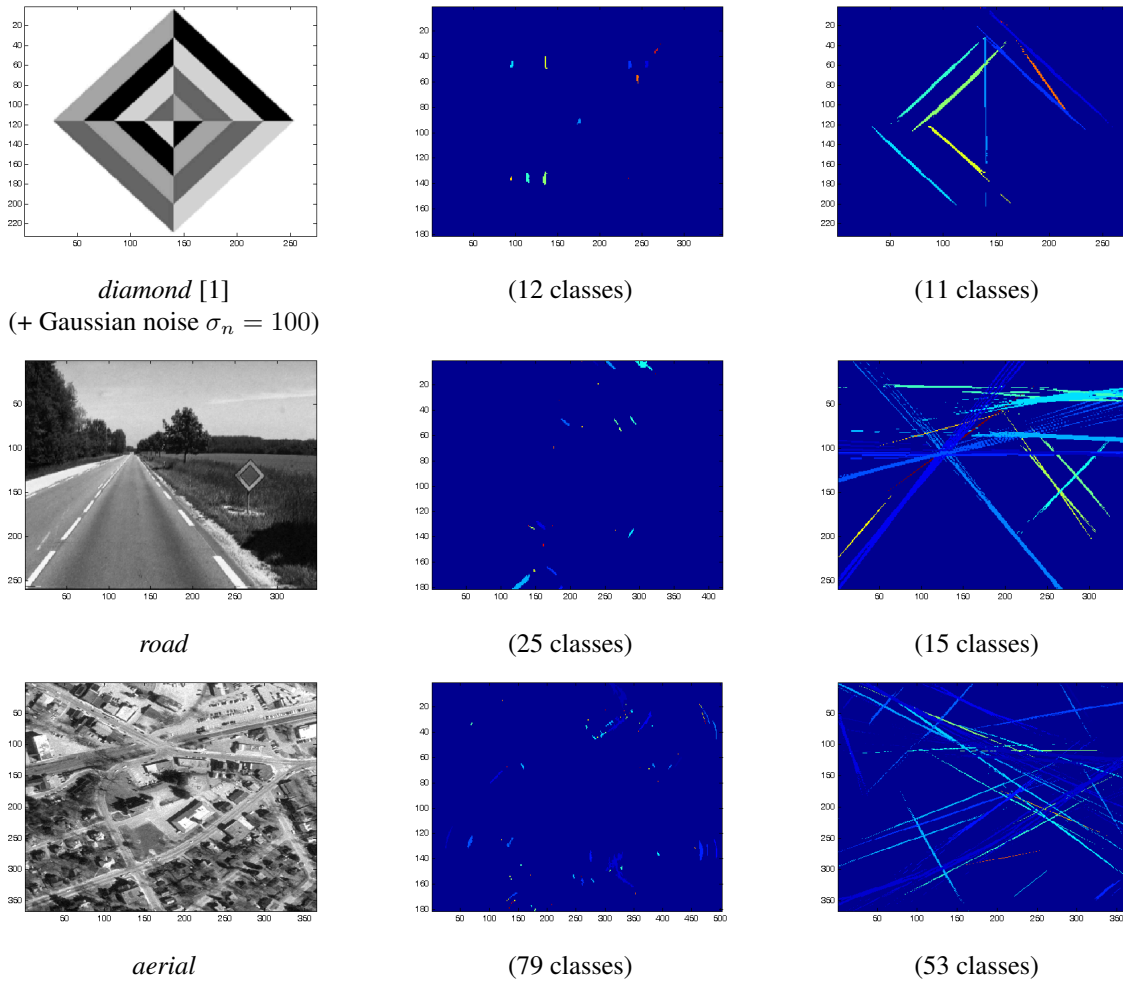


Figure 4. Classification results: in the Hough space (middle column) and spatial domain (right column).

shows that the shape of the peaks itself contains meaningful information for recovering *almost* aligned set of points such as the curvy roads in figure 5.

5 Conclusion

We have presented an unsupervised classification applied to dual spaces: the spatial and Hough spaces. The framework allows to recover the positions that are associated to a particular bump in the Hough space.

Acknowledgment

This work has been supported by the Enterprise Ireland Innovation Partnership IP-2006-412 and a Google Research Award.

References

- [1] A. Bonci, T. Leo, and S. Longhi. A bayesian approach to the hough transform for line detection. *IEEE Transactions on Systems, Man, and Cybernetics*, 35(6), November 2005.
- [2] R. Dahyot. Statistical hough transform. Technical Report TCD-CS-2007-37, Computer Science Trinity College Dublin, July 2007.
- [3] A. Goldenshluger and A. Zeevi. The hough transform estimator. *The Annals of Statistics*, 32(5), October 2004.
- [4] P. Hough. Methods of means for recognising complex patterns, 1962.
- [5] Q. Ji and R. M. Haralick. Error propagation for the hough transform. *Pattern Recognition Letters*, 22:813–823, 2001.
- [6] B. W. Silverman. *Density Estimation for Statistics and Data Analysis*. London: Chapman and Hall, 1986.