Inferring shapes using Mean-shift
Seminar INRIA Sophia Antipolis

Rozenn Dahyot

School of Computer Science and Statistics
Trinity College Dublin, Ireland
Graphics Vision and Visualisation (GV2)
http://gv2.cs.tcd.ie/
https://www.cs.tcd.ie/Rozenn.Dahyot/
Rozenn.Dahyot@tcd.ie

July 6th, 2009
Introduction I

1. Reminder on the Standard Hough transform
2. Statistical Hough Transform: a Radon Transform based approach
3. Stochastic exploration using Mean-shift

Applications:
- Image processing
- Physics
- (geo)Statistics
- etc.
Proposed in 1962, the Hough transform aims at estimating lines from a set of 2D points $\mathcal{S} = \{(x_i, y_i)\}_{i=1}^N$.

As an alternative to the parametric relation $y = ax + b$, lines are defined with their normal equation:

$$\rho = x \cos \theta + y \sin \theta$$

$(\rho, \theta)$ are the latent variables to estimate.
Standard Hough Transform II
Standard Hough Transform III

Space \((x, y)\)

\((x_1, y_1)\)

Space \((\theta, \rho)\)
Standard Hough Transform IV

Space \((x,y)\)

\((x_2, y_2)\)

Space \((\theta, \rho)\)

\(\theta\)

\(\rho\)
Standard Hough Transform V

Space \((x,y)\)

\((x_3,y_3)\)

Space \((\theta,\rho)\)
Standard Hough Transform VI

- $S_{xy} = \{(x_i, y_i)\}_{i=1,\ldots,N}$
- Prior attached to each pixel:

$$p_i = \begin{cases} \frac{1}{c} & \text{if pixel } i \text{ belong to an edge} \\ 0 & \text{otherwise} \end{cases}$$

with the normalising constant $c$ defined such that $\sum_{i=1}^{N} p_i = 1$
Standard Hough Transform VII


Statistical Hough Transform

- $\mathcal{L}_{xy} = \{(x_i, y_i)\}_{i=1,\ldots,N}$, with bandwidths $\{(h_{x_i}, h_{y_i})\}_{i=1,\ldots,N}$

- $\mathcal{L}_{xy\theta} = \{(x_i, y_i, \theta_i)\}_{i=1,\ldots,N}$, with bandwidths $\{(h_{x_i}, h_{y_i}, h_{\theta_i})\}_{i=1,\ldots,N}$ with
  $\theta_i = \arctan \frac{I_y(x_i, y_i)}{I_x(x_i, y_i)}$ computed from the image $I(x, y)$.

- Prior attached to each pixel/data point:
  \[ p_i = \frac{1}{N} \]
Four random variables linked by the relation:

\[ \rho = x \cos \theta + y \sin \theta \]

The joint probability density function can be estimated by:

\[ p_{\theta \rho xy}(\theta, \rho, x, y) = p_{\rho | \theta xy}(\rho | \theta, x, y) \cdot p_{\theta xy}(\theta, x, y) \quad \text{(Bayes)} \]

\[ = \delta(\rho - x \cos \theta - y \sin \theta) \]
The function $p_{\theta xy}(\theta, x, y)$ can be modelled by:

- using observations $\mathcal{I}_{xy} = \{(x_i, y_i)\}_{i=1,\ldots,N}$

\[
\hat{p}_{\theta xy}(\theta, x, y|\mathcal{I}_{xy}) = p_{\theta}(\theta) \sum_{i=1}^{N} \frac{1}{h_x} k_x \left( \frac{x - x_i}{h_x} \right) \frac{1}{h_y} k_y \left( \frac{y - y_i}{h_y} \right) p_i
\]

- using observations $\mathcal{I}_{\theta,xy} = \{ (\theta_i, x_i, y_i)\}_{i=1,\ldots,N}$

\[
\hat{p}_{\theta xy}(\theta, x, y|\mathcal{I}_{xy}) = \sum_{i=1}^{N} \frac{1}{h_\theta} k_\theta \left( \frac{\theta - \theta_i}{h_\theta} \right) \frac{1}{h_x} k_x \left( \frac{x - x_i}{h_x} \right) \frac{1}{h_y} k_y \left( \frac{y - y_i}{h_y} \right) p_i
\]
Statistical Hough Transform IV

The p.d.f. of \((\theta, \rho)\) is computed by integration w.r.t. \(x\) and \(y\):

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

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

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

Taking \(\mathbf{x}_i = (x_i, y_i)\) and \(h_{xi} = h_{yi} = h_i\), with Gaussian kernels, then:

\[
R_i(\rho, \theta) = \frac{1}{\sqrt{2\pi} h_i} \ exp \left( - \frac{(\rho - x_i \cos \theta - y_i \sin \theta)^2}{2 h_i^2} \right)
\]
Statistical Hough Transform V
Statistical Hough Transform VI

\[ \hat{p}_{\theta \rho}(\theta, \rho | \mathcal{L}_{xy}) \]

\[ h_{x_i} = h_{y_i} = \min_{j \in [1; N]} (j \neq i) \{ \| \mathbf{x}_i - \mathbf{x}_j \| \} \]

\[ \hat{p}_{\theta \rho}(\theta, \rho | \mathcal{L}_{xy}) \text{ FB} \]

\[ h_{x_i} = h_{y_i} = h_i = 1, \ \forall i \]
\hat{p}_{\theta \rho}(\theta, \rho | S_{\theta xy})
Mean-shift algorithm I

**Definition (Mean-shift)**

Mean-shift is a gradient ascent method that defines a 'deterministic' Markov chain that converges toward a nearby local maximum.

For a random variable $\Theta = (\rho, \theta)$ and its estimated density function $\hat{p}_\Theta(\Theta)$ using kernel modelling, we can find a relation $f$ such that:

$$\left( \Theta^{(m+1)} = f(\Theta^{(m)}) \right) \land \left( \hat{p}_\Theta(\Theta^{(m)}) \leq \hat{p}_\Theta(\Theta^{(m+1)}) \right)$$

Starting with an initial position $\Theta^{(0)}$, by applying several time $f$, then $f^{(\infty)}(\Theta^{(0)}) = \Theta^{(\infty)}$ is a local maximum.
Mean-shift for Statistical Hough Transform I

- Our Kernel distribution $\hat{p}_{\rho\theta}(\rho, \theta|\mathcal{S}_x)$ is defined on $\mathbb{R} \times [\frac{-\pi}{2}; \frac{\pi}{2}]$ and is not linear in $\theta$.
- We change variables:

$$n = \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix}$$

and the equation of the line becomes:

$$\rho = x^T n$$

- Hypothesis: $n$ (or $\theta$) as a uniform distribution

$$\hat{p}_{\rho n}(\rho, n|\mathcal{S}_x) = \left(\frac{1}{\pi}\right) \sum_{i=1}^{N} R_i(\rho, n) p_i$$

$\hat{p}_{\rho n}(\rho, n|\mathcal{S}_x)$ is defined on $\mathbb{R}^1 \times \mathbb{S}^{d-1}$ where $\mathbb{S}^{d-1}$ is a unit hypersphere:

$$\mathbb{S}^{d-1} = \{n \in \mathbb{R}^d : \|n\| = 1\}$$
Mean-shift for Statistical Hough Transform II

1. **Mean shift step in** $\mathbb{R}^1$

$$\rho^{(m+1)} = \frac{\sum_{i=1}^{N} \left( \frac{x_i^T n_i^{(m)}}{h_i^2} \right) R_i(\rho^{(m)}, n^{(m)}) \ p_i}{\sum_{i=1}^{N} \frac{1}{h_i^2} R_i(\rho^{(m)}, n^{(m)}) \ p_i}$$

2. **Mean shift step in** $\mathbb{S}^{d-1}$

$$n^{(m+1)} = \frac{(A^{(m)})^{-1} b^{(m)}}{\|(A^{(m)})^{-1} b^{(m)}\|}$$
Mean-shift for Statistical Hough Transform III

with \((d = 2)\)

- \(A^{(m)} = [a_{k,j}]\) is a \(d \times d\) square symmetric matrix defined by:

  \[
  a_{k,j} = \sum_{i=1}^{N} \frac{x_{ik} x_{ij}}{h_i^2} R_i(\rho^{(m)}, n^{(m)}) p_i
  \]  

  \(∀k = 1, \ldots, d\)  

- the vector \(b\) is:

  \[
  b_k = \rho^{(m)} \sum_{i=1}^{N} \frac{x_{ik}}{h_i^2} R_i(\rho^{(m)}, n^{(m)}) p_i
  \]  

  \(∀k = 1, \ldots, d\)
Link with PCA I

\[ (x_i, y_i) \]

\[ \varepsilon_i = \rho - x_i \cos(\theta) - y_i \sin(\theta) \]
Link with PCA II

With

\[ \epsilon_i = \rho - x_i^T \mathbf{n}, \quad \forall i = 1, \cdots, N \]

- **PCA:**
  \[
  (\rho_{pca}, \mathbf{n}_{pca}) = \arg\max_{\rho, \mathbf{n}} \left\{ p(\epsilon_1, \cdots, \epsilon_N) = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi h}} \exp \left( -\frac{\epsilon_i^2}{2h^2} \right) \right\}
  \]

- **SHT:**
  \[
  (\rho_{sht}, \mathbf{n}_{sht}) = \arg\max_{\rho, \mathbf{n}} \left\{ \hat{p}_{\rho \mathbf{n}}(\rho, \mathbf{n}) = \left( \frac{1}{\pi} \right) \sum_{i=1}^{N} \frac{1}{\sqrt{2\pi h}} \exp \left( -\frac{\epsilon_i^2}{2h^2} \right) \right\}
  \]
Finding the global maximum with simulated annealing

L1: (ρ=10; θ=π/4)
L2: (ρ=-10; θ=π/4)
The bandwidth acts like the temperature in the simulated annealing approach. To search for the **global maximum**, an iterative scheme can be used starting with \( h_0 \) decreasing to \( h_{min} \):

\[
\begin{align*}
\text{Initial guess at } h_0: \quad (\hat{\rho}_0, \hat{\theta}_0) &= (\hat{\rho}_{pca}, \hat{\theta}_{pca}) \\
(\hat{\rho}_k, \hat{\theta}_k) &= f_k^\infty(\hat{\rho}_{k-1}, \hat{\theta}_{k-1}) \text{ computed with } h_k = (0.99)^k h_0 \\
\text{Until } h_k &= h_{min}
\end{align*}
\]
Finding the global maximum with simulated annealing III
Finding the global maximum with simulated annealing IV

Graph 1: 
- x-axis: $N_1$
- y-axis: $s_e(\rho)$
- Graph shows the behavior of $s_e(\rho)$ as $N_1$ varies from 0 to 100.

Graph 2: 
- x-axis: $N_1$
- y-axis: $s_e(\theta)$
- Graph shows the behavior of $s_e(\theta)$ as $N_1$ varies from 0 to 100.
Finding (local) maxima with Meanshift I

The bandwidth are fixed.

1. Generate guesses in the $\rho n$-space.
2. Iterate mean-shift algorithm till convergence.
3. Keep the highest maxima as estimates.

Illustrations

- Lines (d=2)
- Circles (d=3)
Finding (local) maxima with Meanshift II

Finding lines \((d = 2)\)

image

priors \(\{p_i\}\)
Finding (local) maxima with Meanshift III

Finding lines ($d = 2$)

\[ \hat{p}_{\rho \theta}(\rho, \theta) \]
Finding (local) maxima with Meanshift IV

Finding circles \((d = 3)\)

Equation of a circle:

\[ \mathbf{z}^T \mathbf{n} = \rho \quad \text{subject to } \| \mathbf{n} \| = 1 \]

with

\[
\mathbf{z} = \begin{pmatrix}
  z_1 = x^2 + y^2 \\
  z_2 = x \\
  z_3 = y
\end{pmatrix}
\]

Using the observations \(\mathcal{S}_x = \{ \mathbf{x}_i = (x_i, y_i) \}_{i=1,\ldots,N}\), we compute the observations \(\mathcal{S}_z = \{ \mathbf{z}_i = (x_i^2 + y_i^2, x_i, y_i) \}_{i=1,\ldots,N}\).
Finding (local) maxima with Meanshift V

Finding circles ($d = 3$)
Finding (local) maxima with Meanshift VI

Results obtained with 100 random guesses.
Finding (local) maxima with Meanshift VII

Results obtained with 100 random guesses.
Conclusion

We have proposed:

- a way to model robustly the pdf of latent variables that describe a shape (SHT).

- an algorithm to explore this pdf and infer the global maximum and/or the local ones.
Some Future challenges with MS-SHT

- Choose better priors for $n$ (or $\theta$) and define the corresponding MS algorithm to perform stochastic exploration.

- Improve the generation of initial guesses to start the MS algorithm.

- The current algorithm can be applied to many shapes: lines, hyperplanes, circles, ellipse, quartic etc. Can this approach be generalised to other (non-linear) shapes? 
Current/Future investigation: 3D shape inference I

- Using a turning table, an image of the object is taken every 5 deg.
- The silhouette of the object is segmented in each image.
- Each slice is computed by inverse Radon transform and stacked to give an estimate of the 3D shape.
Current/Future investigation: 3D shape inference II

- Project: STI Cell Centre of Competence hosted in GV2 group (Directors: Dr. Michael Manzke & Dr. Steven Collins).
- We are looking at 3D Shape inference from multiple camera views, inferred in Real-time.

Table: Amsterdam Library of Object Images (ALOI).
Current/Future investigation: 3D shape inference III

Table: Mean-shift 3D inference (©Jonathan Ruttle, PhD student GV2.)
Any question?

References:

