Multiscale differential Geometry and Dense Hough transforms: a unified framework for line, circle and object detection in videos

Antoine Manzanera
ENSTA-ParisTech

Congreso de Informática y Sistemas, Villahermosa, México 2012

Context and Objectives

Computer Vision

- Semantic Gap
- Variability of aspects and conditions
- Prior knowledge, context,...

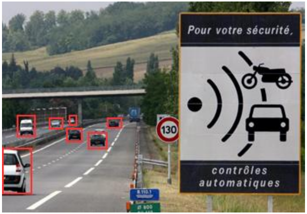
Context and Objectives

Analytical Shapes Recognition

- Lines, circles, conic curves,...
- Defined by an equation

Object Recognition

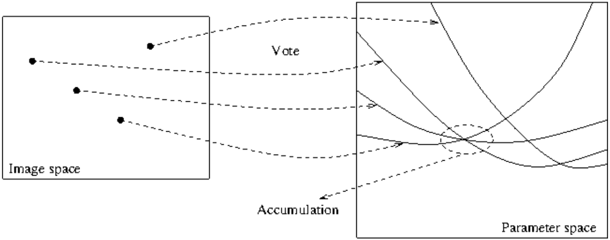
- Car, face, chair,...
- User defined parameters
- Example / Learning based



Presentation outline

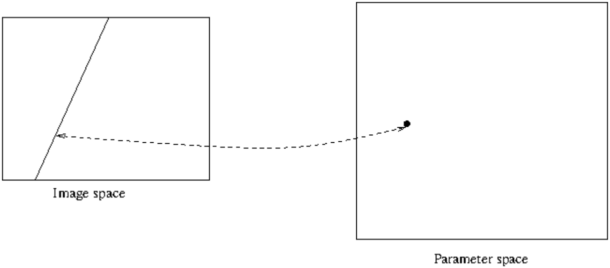
Hough Transform: global view

- One of the oldest applications of Computer Vision (End 50's, bubble chamber images)
- Adapted to both analytical (curves) or non analytical (objects) shapes
- Based on accumulation (vote) mechanism from image space (pixels) to multidimensional parameter space



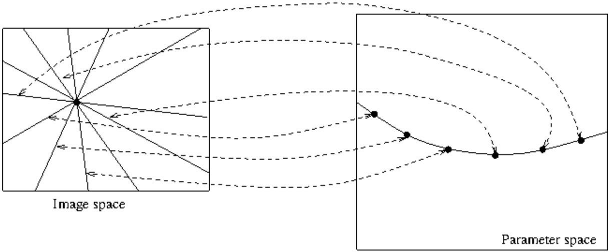
Hough Transform: details (1)

Every point of the parameter space (dimension: number of parameters) corresponds to one shape in the image space.
Example : One (θ, ρ) polar coordinates point correspond to one line.



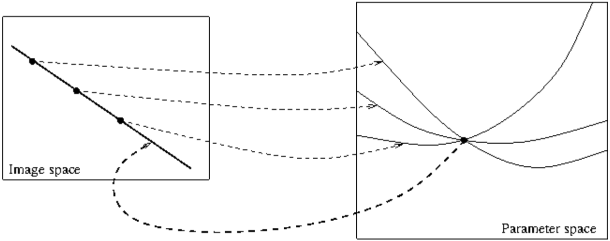
Hough Transform: details (2)

Every single curve of the parameter space corresponds to one point, or equivalently to one beam of shapes in the image space.
Example: One sine curve corresponds to one beam of line, i.e. one point.



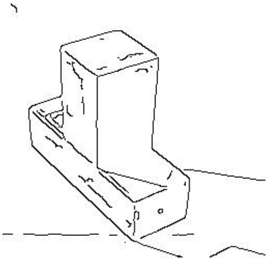
Hough Transform: details (3)

Reciprocally, different points from the same shape in the image space form a beam of curves in the parameter space, converging to one point defining the right shape.



Hough Transform: practice

So classically, the Hough transform (i.e. the result of the projection of all image points in the parameter space) is calculated from a limited set of points: the contours.
The best candidat shapes are then detected by computing the local maxima of the Hough transform.



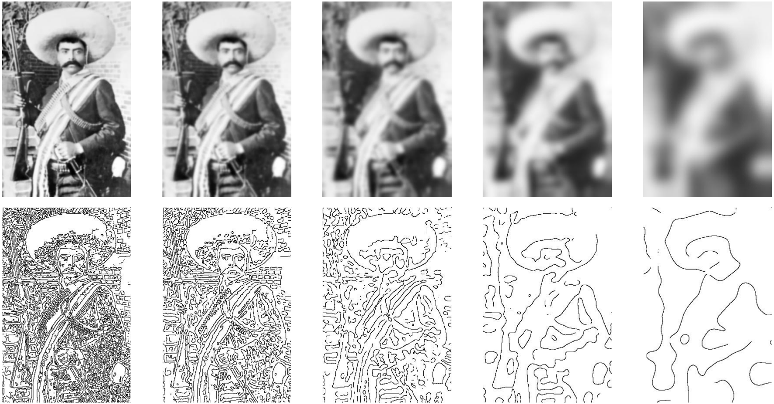
Contour image

Classical Hough transform: different accumulation points are visible

Scale space theory

- Mainly developed in the 80's for Computer Vision applications
- Key idea: Every measure is relative to scale.
- Principle: A multiscale version, i.e a stack of blurred images $(I^\sigma)_{\sigma \in \{\sigma_1, \dots, \sigma_n\}}$ is produced from one single image I, with σ the scale parameter.
- Causality principle: No new structure may appear in I^σ that is not already present in finer scales $I^{\sigma'}$, $\sigma' < \sigma$

Scale space causality illustrated on contours



$\sigma = 1.0$ $\sigma = 2.0$ $\sigma = 4.0$ $\sigma = 8.0$ $\sigma = 16.0$

Presentation outline

The differential model

- In Image Processing and Computer Vision, many algorithms are based on local features calculated using partial derivatives: Contrast, contours, segmentation...
- In the differential model the image is assimilated to a continuous and differentiable function.
- The local behaviour of the image near every point can be predicted thanks to the partial derivatives (Taylor expansion).

Notations

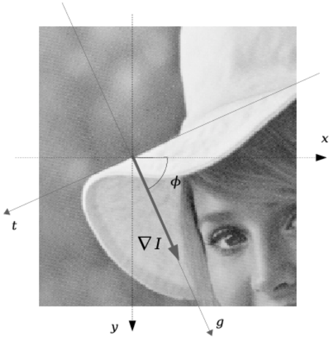
$$I_x = \frac{\partial I}{\partial x},\ I_y = \frac{\partial I}{\partial y},\ I_{xx} = \frac{\partial^2 I}{\partial x^2},\ I_{xy} = \frac{\partial^2 I}{\partial x \partial y},\ \text{etc.}$$

Order 1: Gradient and Isophote

At order 1, the main measure is the gradient vector:

$$\nabla I = (Ix, Iy)^T$$

- Its argument $\Phi = \arg \nabla I$ correspond to the direction of steepest ascent.
- Its modulus $||\nabla I||$ is a measure of contrast.

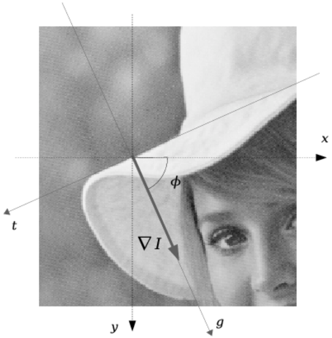


Order 1: Gradient and Isophote

Let **v** be a unit vector. The value of the first derivative along **v** is calculated as:

$$I_v = \mathbf{v}^T . \nabla I$$

In particular, the derivative along the direction orthogonal to the gradient is zero (isophote direction **t**).

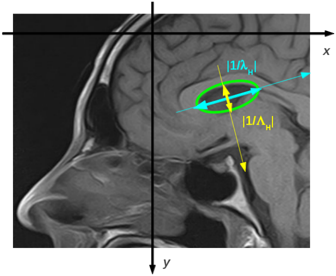


Order 2: Hessian and Curvature

At order 2, the main measure is the Hessian matrix:

$$H_I = \begin{pmatrix} I_{xx} & I_{xy} \\ I_{xy} & I_{yy} \end{pmatrix}$$

- Its eigen vectors (resp. eigen values Λ_H and λ_H) correspond to the directions (resp. intensities) of main curvatures.
- Its Frobenius norm $||H_I||_F = \sqrt{\Lambda_H^2 + \lambda_H^2}$ is a measure of global curvature intensity.



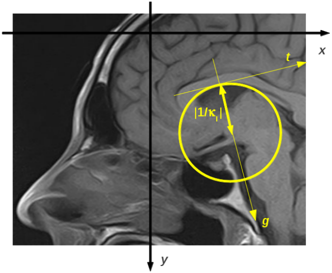
Order 2: Hessian and Curvature

Let **u** and **v** be two unit vectors. The value of the second derivative along **u** and **v** is calculated as:

$$I_{uv} = \mathbf{u}^T H_I \mathbf{v}$$

In particular, the second derivative along the isophote **t** is the curvature of the isophote, i.e. the inverse radius of the osculating circle:

$$\kappa_I = I_{tt} = \frac{I_{xx} I_y^2 - 2 I_{xy} I_x I_y + I_{yy} I_x^2}{||\nabla I||^3}$$



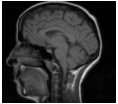
Multiscale derivatives

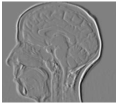
According to the Scale Space framework, a derivative measure on our discrete images only makes sense up to a scale parameter. The derivative is then estimated at scale σ using the convolution by the corresponding derivative of Gaussian function:

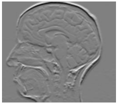
$$I^{\sigma}_{x^i y^j} = I \star \frac{\partial^{i+j} G_\sigma}{\partial x^i \partial y^j}$$

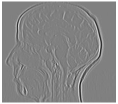
- G_σ is the 2d Gaussian function with variance σ^2.
- σ is the scale of estimation.
- $i + j$ is the order of derivation.

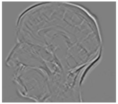
Multiscale Gaussian Derivatives

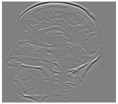

$I^{1.0}$

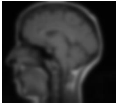

$I_x^{1.0}$

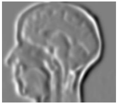

$I_y^{1.0}$

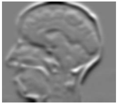

$I_{xx}^{1.0}$

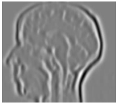

$I_{xy}^{1.0}$


$I_{yy}^{1.0}$

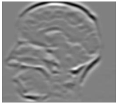

$I^{4.0}$

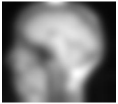

$I_x^{4.0}$


$I_y^{4.0}$

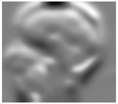

$I_{xx}^{4.0}$


$I_{xy}^{4.0}$


$I_{yy}^{4.0}$

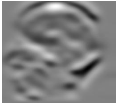

$I^{10.0}$


$I_x^{10.0}$


$I_y^{10.0}$


$I_{xx}^{10.0}$


$I_{xy}^{10.0}$


$I_{yy}^{10.0}$

Presentation outline

Partial derivatives and 1-to-1 Hough transforms

Classical approaches

- sparse: Only a few points (contours, key points) are voting.
- 1-to-many: Every point from the image space is voting uniformly on a n dimensional surface in the parameter space.
- many-to-1: Every n-tuple of points from the image space is voting for one unique point in the parameter space.

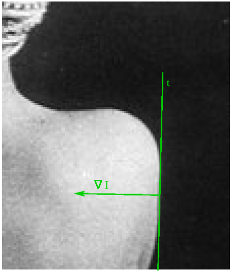
Partial derivatives and 1-to-1 Hough transforms

Hough transfoms based on partial derivatives

- dense: All the points are voting...
- inegalitarian: ...but their votes don't have the same weight!
- 1-to-1: Every point from the image space is voting for one unique point in the parameter space.

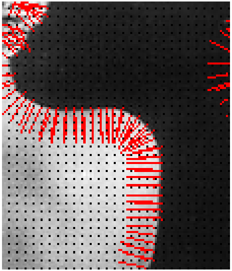
1-to-1 transform: order 1

At order 1, the gradient defines the isophote direction, and then the direction of the candidate line. The weight of the vote is the norm of the gradient.



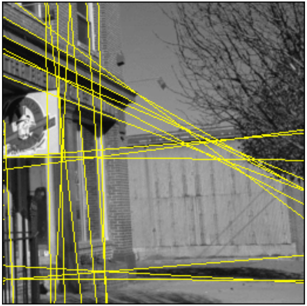
Gradient and line

Weight of the vote

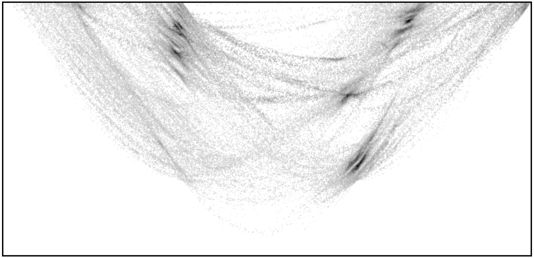


Main votes

1-to-1 transform: order 1



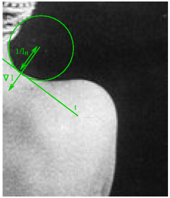
20 best lines



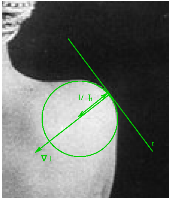
(ρ, θ) 1-to-1 transform

1-to-1 transform: order 2

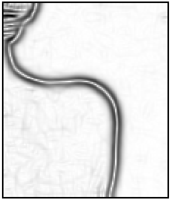
At order 2, the gradient direction and the isophote curvature define the radius and the centre of the osculating circle to the isophote curve, and then the equation of the candidate circle. The weight of the vote is the Frobenius norm of the Hessian matrix.



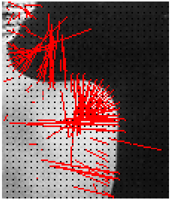
Positive curvature



Negative curvature



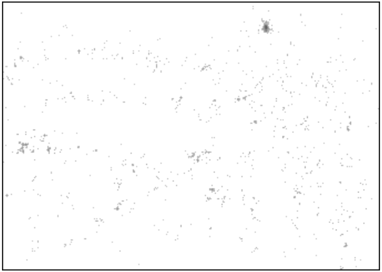
Weight of the vote



Main votes

1-to-1 transform: order 2

10 best circles



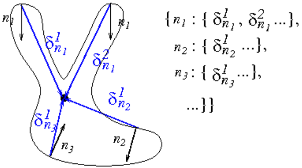
(ρ, x, y) 1-to-1 transform (level $\rho = 19$)

Presentation outline

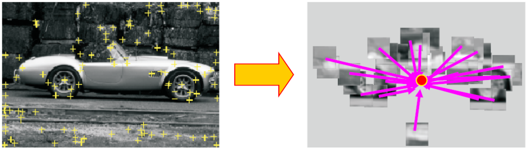
Object representation by R-Tables

The classical generalised Hough transforms are *sparse*: they are calculated from a reduced set of feature points: contour [Ballard 81], or salient points [Leibe 04].

$$\text{R-Table}: \{i, \{\vec{\delta_i^j}\}_j\}_i$$



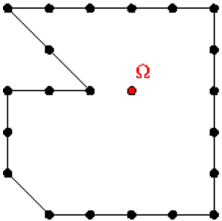
Contour

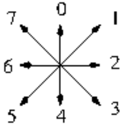


Salient points

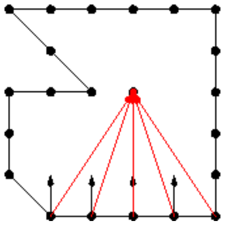
Construction of the R-Table

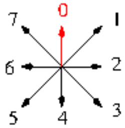
The R-Table is a shape model, constructed from a prototype. Let Ω be an arbitrary centre of the prototype. Every point M of the prototype is indexed by a geometrical feature i, corresponding to the row indices of the R-table. The R-table is constructed by adding the displacement vector $\overrightarrow{M\Omega}$ in the line of index i.
For example consider the following contour points as a prototype, indexed by the normal direction to the contour, quantised to 8 values:





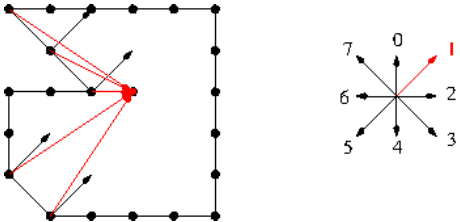
Construction of the R-Table





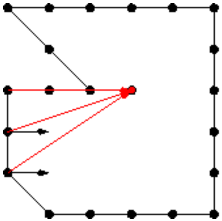
Index	List of vectors					
0	$\begin{pmatrix}-2\\-3\end{pmatrix}$	$\begin{pmatrix}-1\\-3\end{pmatrix}$	$\begin{pmatrix}0\\-3\end{pmatrix}$	$\begin{pmatrix}1\\-3\end{pmatrix}$	$\begin{pmatrix}2\\-3\end{pmatrix}$	end

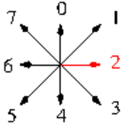
Construction of the R-Table



Index	List of vectors					
0	$\binom{-2}{-3}$	$\binom{-1}{-3}$	$\binom{0}{-3}$	$\binom{1}{-3}$	$\binom{2}{-3}$	end
1	$\binom{2}{-3}$	$\binom{3}{-2}$	$\binom{1}{0}$	$\binom{2}{1}$	$\binom{3}{2}$	end

Construction of the R-Table



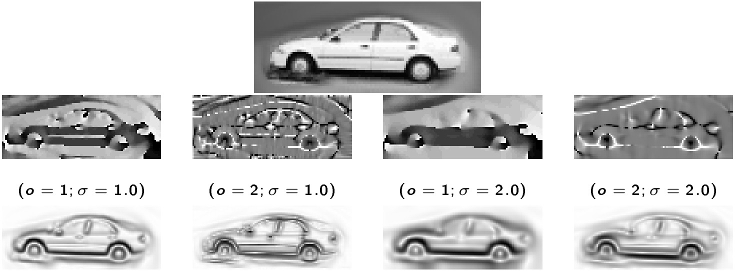


Index	List of vectors					
0	$\begin{pmatrix}-2\\-3\end{pmatrix}$	$\begin{pmatrix}-1\\-3\end{pmatrix}$	$\begin{pmatrix}0\\-3\end{pmatrix}$	$\begin{pmatrix}1\\-3\end{pmatrix}$	$\begin{pmatrix}2\\-3\end{pmatrix}$	end
1	$\begin{pmatrix}2\\-3\end{pmatrix}$	$\begin{pmatrix}3\\-2\end{pmatrix}$	$\begin{pmatrix}1\\0\end{pmatrix}$	$\begin{pmatrix}2\\1\end{pmatrix}$	$\begin{pmatrix}3\\2\end{pmatrix}$	end
2	$\begin{pmatrix}3\\-2\end{pmatrix}$	$\begin{pmatrix}3\\-1\end{pmatrix}$	$\begin{pmatrix}3\\0\end{pmatrix}$	end		

and so on...

Dense R-Tables indexed by derivatives

In the *dense* approach, the indices i of the R-table are the quantised mutiscale derivatives, available everywhere.



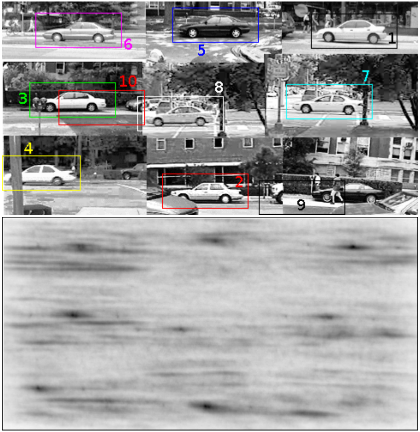
Weighted R-Table: $\{i, \{\vec{\delta}_i^j, \omega_i^j\}_j\}_i$

Generalised Hough Transform: Object Detection

Initial: $H(\mathbf{x}) = 0$ everywhere.
For all image point $\mathbf{x}$,
let $\lambda(\mathbf{x})$ the quantised derivative.
  For all occurence j of the R-Table associated to $\lambda(\mathbf{x})$, do:

$$H(\mathbf{x} + \delta^{j}_{\lambda(\mathbf{x})}) \mathrel{+}= \omega^{j}_{\lambda(\mathbf{x})}$$

The best object candidates are then located at the maxima of H (Right: Hough transform and the 10 best candidate cars).



Presentation outline

Contribution outline

- **Combination** of cumulative Hough approaches and Scale space derivatives.
- **Faster** computation (no more feature extraction).
- **More reliable** voting process: more voting points, more precise voting locations.

Bibliography (For further details)

[Manzanera 11] A. MANZANERA
Local Jet Feature Space Framework for Image Processing and Representation
Int. Conf. on Signal Image Technology and Internet Based Systems 261-268 (2011)

[Manzanera 12] A. MANZANERA
Dense Hough transforms on gray level images using multi-scale derivatives
Working document, ENSTA-ParisTech (2012)

Bibliography (Scale space)

- **[Lindeberg 98]** T. LINDEBERG
  Feature detection with automatic scale selection
  International Journal of Computer Vision 30(2), 77-116. (1998)
- **[Koenderink 87]** J.J. KOENDERINK and A.J. VAN DOORN
  Representation of Local Geometry in the Visual System
  Biological Cybernetics 55, 367-375. (1987)
- **[Van Vliet 98]** L.J. VAN VLIET, I.T. YOUNG and P.W. VERBEEK
  Recursive Gaussian derivative filters
  Proc. Int. Conf. on Pattern Recognition vol. 1, 509-514. (1998)

Bibliography (Hough transforms 1)

[Hough 59] P. HOUGH
Machine analysis of bubble chamber pictures.
In: Int. Conf. High Energy Accelerators and Instrumentation. (1959)

[Ballard 81] D.H. BALLARD
Generalizing the Hough transform to detect arbitrary shapes.
Pattern Recognition **13** (1981) 111–122

Bibliography (Hough transforms 2)

- **[O'Gorman 76]** F. O'GORMAN AND B. CLOWES
  Finding picture edges through collinearity of feature points
  IEEE Trans. on Computers C-25 449-456 (1976)
- **[Leibe 04]** B. LEIBE, A. LEONARDIS and B. SCHIELE
  Combined object categorization and segmentation with an implicit shape model
  ECCV Workshop on Statistical Learning in Computer Vision (2004)
- **[Valenti 08]** R. VALENTI and T. GEVERS
  Accurate eye center location and tracking using isophote curvature
  Int. Conf. on Computer Vision and Pattern Recognition (2008)