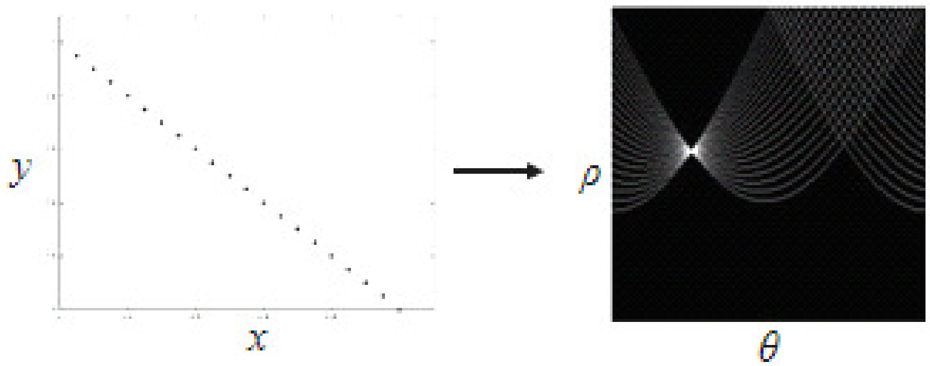
Hough Transform



lecturer: Jiří Matas, matas@cmp.felk.cvut.cz

authors: Jiří Matas, Ondřej Drbohlav

Czech Technical University, Faculty of Electrical Engineering

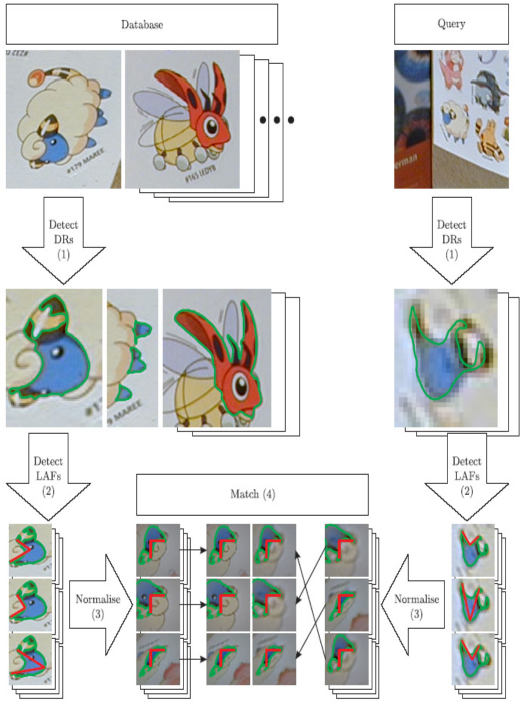
Department of Cybernetics, Center for Machine Perception

Many slides thanks to Kristen Grauman and Bastian Leibe

Why HT and not Recognition with Local Features?

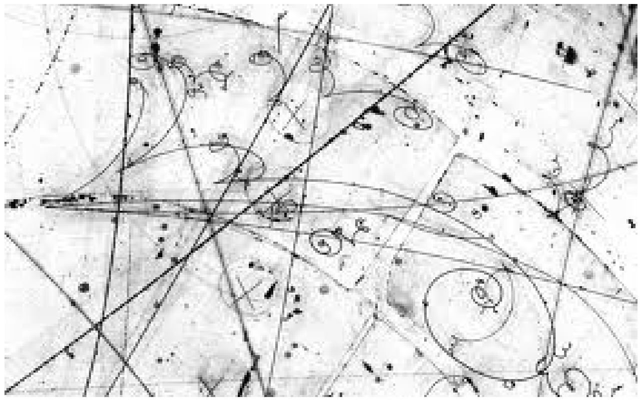
Strengths:

- applicable to many objects (e.g. in image stitching)
- is real-time
- scales well to very large problems (retrieval of millions of images)
- handles occlusion well
- insensitive to a broad class of image transformations



Weaknesses:

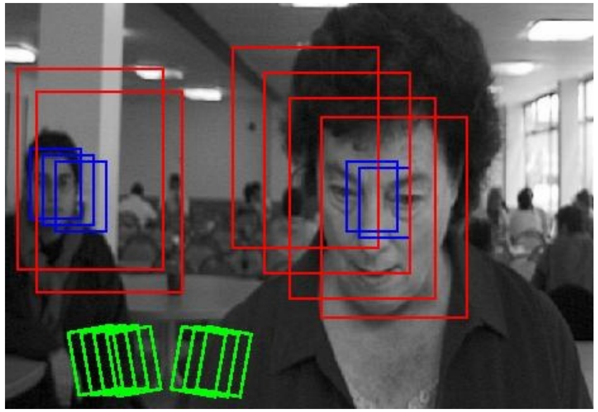
- applicable to recognition of specific objects (no categorization)
- applicable only to objects with distinguished local features



Why HT and not the Scanning Window (Viola-Jones) Method ?

Strengths:

- applicable to many <u>classes</u> of objects
- not restricted to specific objects
- often real-time



Weaknesses:

- extension to a large number of classes not straightforward (standard implementation: linear complexity in the number of classes)
- occlusion handling not easy
- full 3D recognition requires too many windows to be checked
- training time is potentially very long

Hough Transform

CENTER FOR MACHINE PERCEPTION

- A method for detecting geometric primitives based on evaluation of an objective function:

$$J(\Omega_c) = \sum_{i=1}^{M} p(\mathbf{x}_i, \Omega_c)$$

$\Omega_c \in \mathcal{R}^N$ is the parameter space, $\mathbf{x}_i$ are *tokens* (image points of interest)

- Origin: Detection of straight lines

- Examples of Ω_c for different geometric primitives:
  - Straight line: $\Omega_c = (a, b) \in \mathcal{R}^2$ $\quad y - ax - b = 0$
  - Circle: $\Omega_c = (x_c, y_x, r) \in \mathcal{R}^3$ $\quad (y - y_c)^2 + (x - x_c)^2 - r^2 = 0$

- Parameters evaluated on a grid
  - Discretization of Ω_c: $\quad \Omega = N_1 \times N_2 \times N_3 \times ...$

Comparison: Template Matching and HT

- Template Matching:

```
for all ω ∈ Ω
    J(ω) = 0
    for all x = (x, y) ∈ Image // for all x_i ~ tokens
        if ω ∈ Ω(x//x_i)
            J(ω) = J(ω) + p(x//x_i)
        else
            /* nothing */
```

$$\texttt{for all}\ \omega \in \Omega$$
$$J(\omega) = 0$$
$$\texttt{for all}\ \mathbf{x} = (x, y) \in \text{Image} \ // \text{ for all } \mathbf{x}_i \sim \text{tokens}$$
$$\texttt{if}\ \omega \in \Omega(\mathbf{x}//\mathbf{x}_i)$$
$$J(\omega) = J(\omega) + p(\mathbf{x}//\mathbf{x_i})$$

  - Complexity: $O(|\Omega| \times |P|)$

- HT: (basic idea: each "token" votes for all primitives it is consistent with)

$$\texttt{for all}\ \mathbf{x}_i$$
$$\texttt{find}\ \Omega(\mathbf{x}_i)$$
$$J(\omega) = J(\omega) + p(\mathbf{x}_i)$$

  - Complexity: $O(|\Omega(\mathbf{x}_i)| \times |P|); \ \ |\Omega(\mathbf{x}_i)| \ll |\Omega|$

HT for Straight Lines: Parametrization (1)

- Line parametrization:

$$ax + by + c = 0, \qquad (a \neq 0 \vee b \neq 0) \qquad (1)$$

$$(x, y) : \text{ point coordinates} \qquad (2)$$

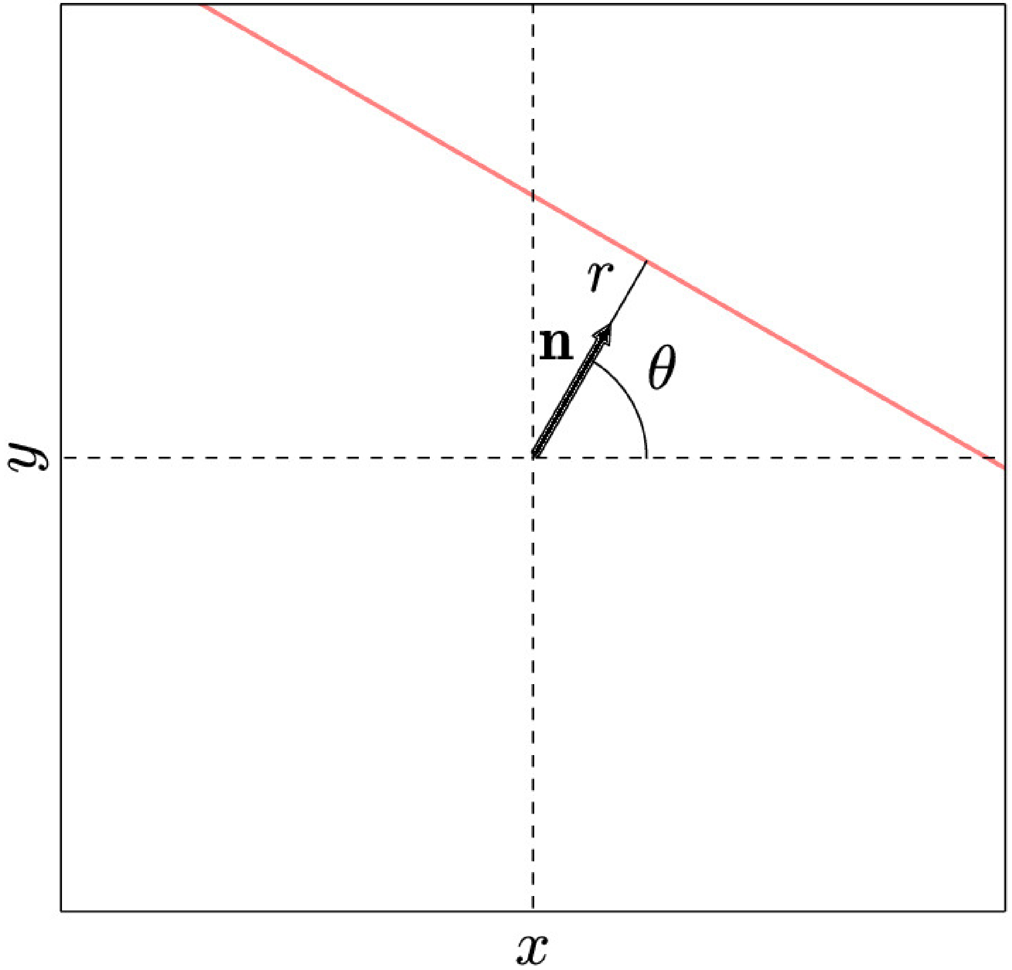
$$(a, b, c) : \text{ line parameters} \qquad (3)$$

- There are 3 line parameters (a, b, c) in this equation.
- The equation is homogeneous. Parameters (a, b, c) and (ka, kb, kc) $(k \neq 0)$ represent the same line. Thus, there are only 2 degrees of freedom (2 DOFs) as expected (orientation and shift)
- A 2-DOF representation:

$$x \cos\theta + y \sin\theta - r = 0, \qquad (\theta \in [0, 2\pi[,\ r \geq 0),\ \text{or} \qquad (4)$$

that's what we'll use $\rightarrow$ $\quad \theta \in [0, \pi[,\ r \in \mathbb{R})$ (5)

HT for Straight Lines: Parametrization (2)



$$x \cos\theta + y \sin\theta = r, \qquad (1)$$

$$(\theta \in [0, \pi[,\ r \in \mathbb{R}) \qquad (2)$$

or,

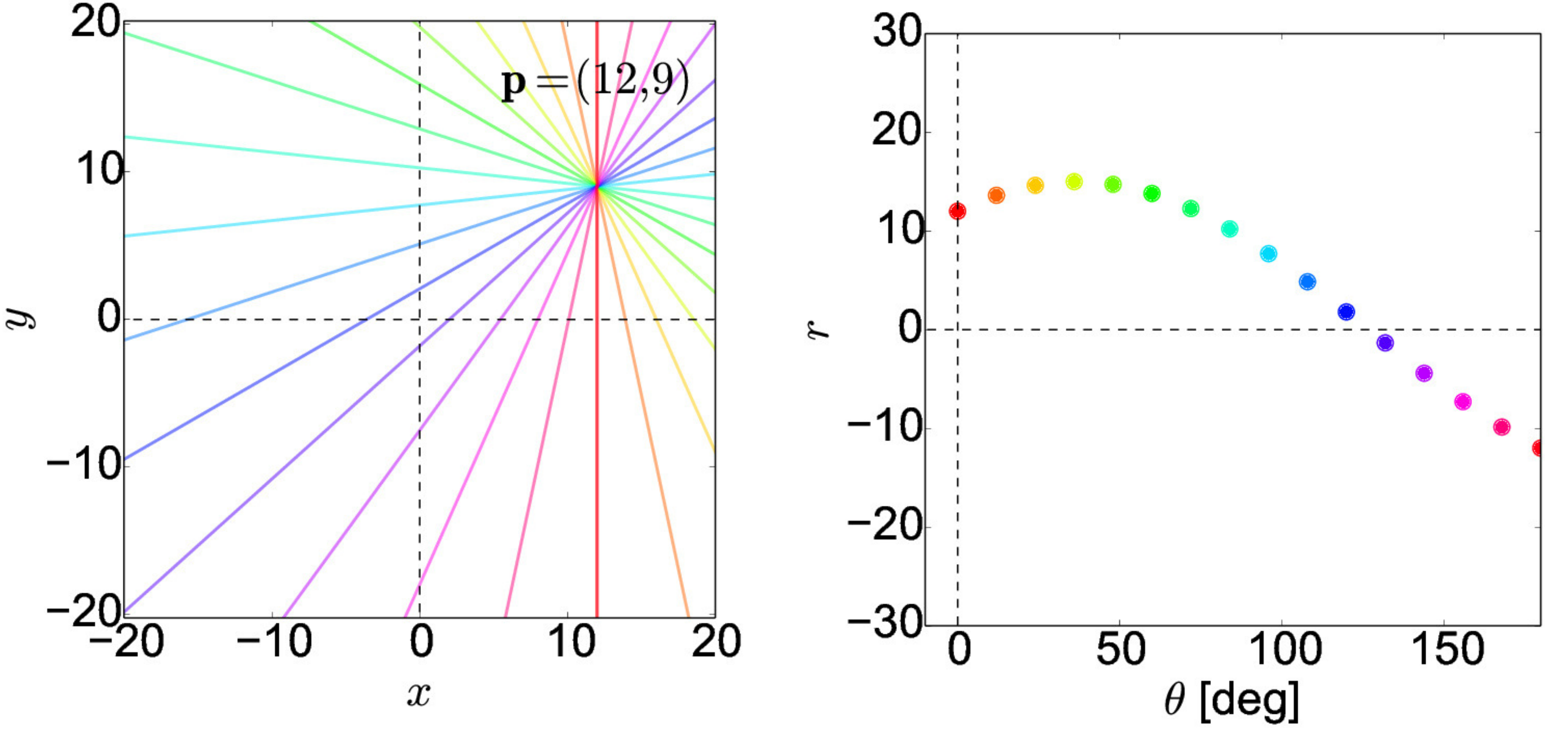
$$(x, y) \cdot (\cos\theta, \sin\theta) = r \qquad (3)$$

$$\theta \in [0, \pi[,\ r \in \mathbb{R}) \qquad (4)$$

Note: $\mathbf{n} = (\cos\theta, \sin\theta)$ (thus $\|\mathbf{n}\| = 1$)

HT for Straight Lines (3)

A point p votes for all lines it can be incident with.



Subset of lines incident with p

Corresponding line parameters

HT for Straight Lines (4)

A point p votes for all lines it can be incident with.

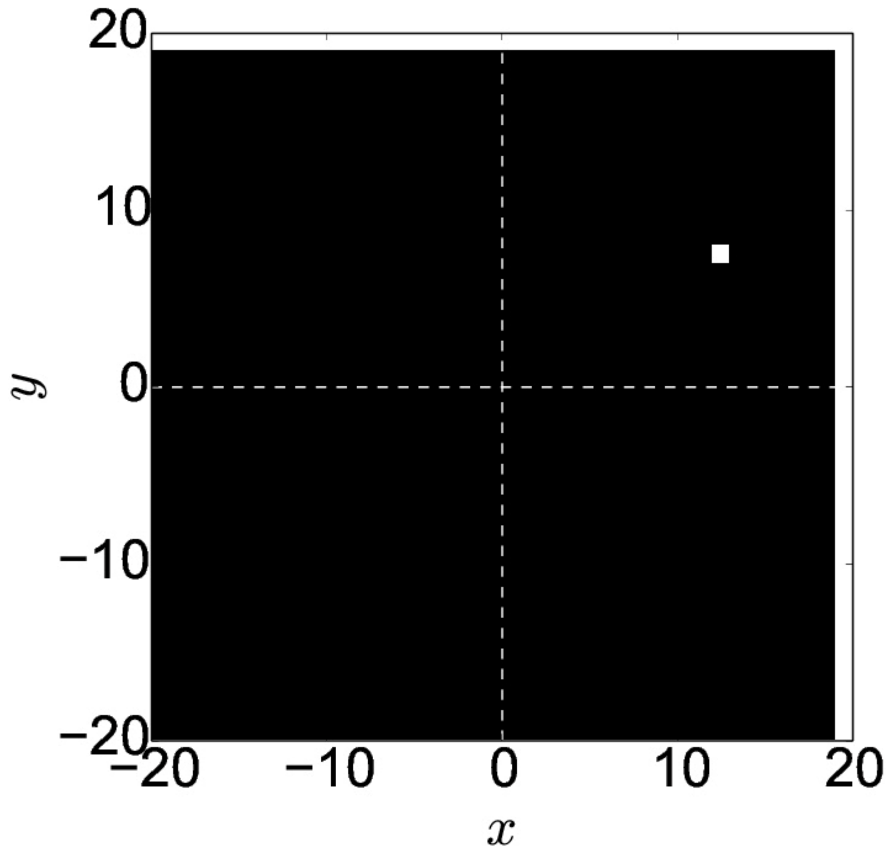
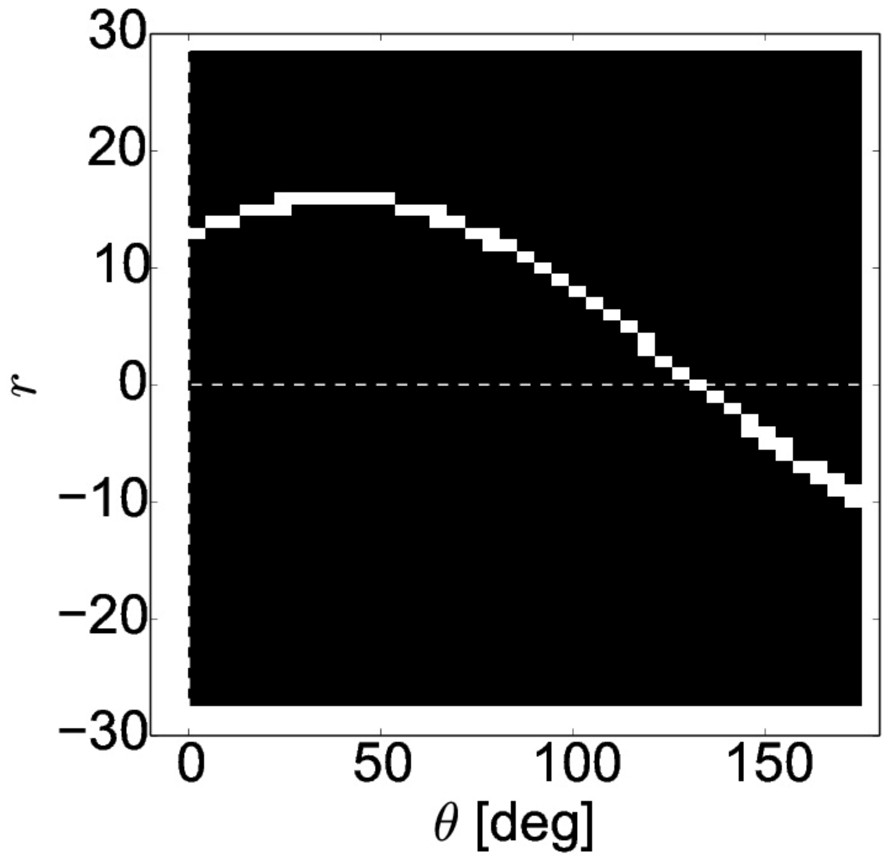


Image with a single point



Accumulator storing votes

HT for Straight Lines (5)

Multiple points; accumulating votes

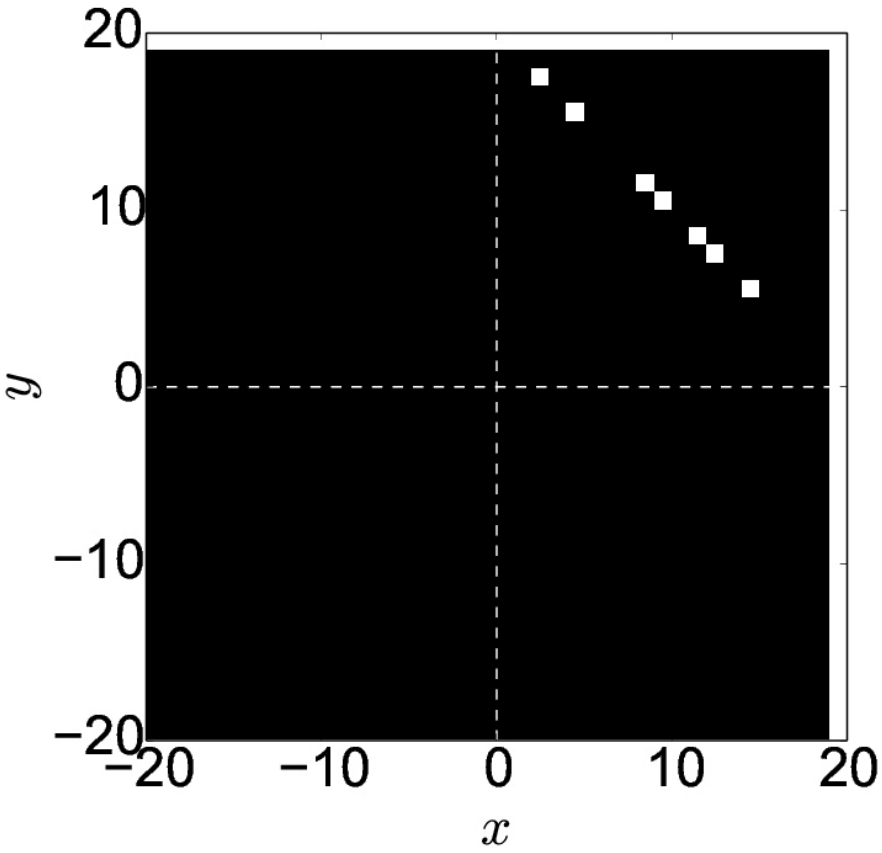
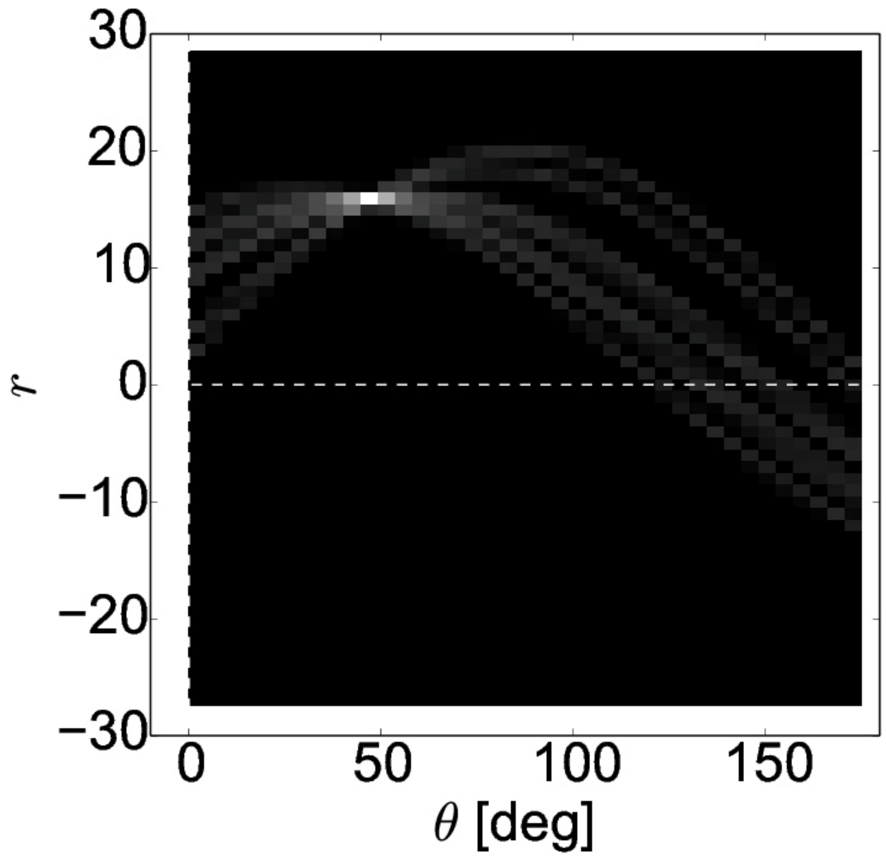


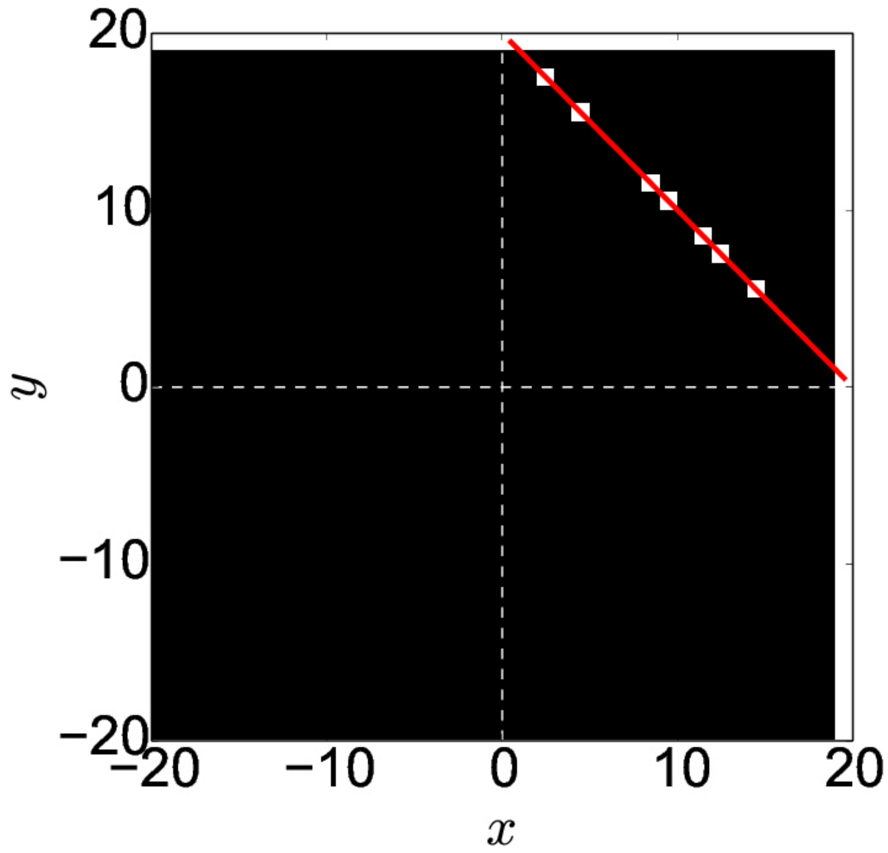
Image with multiple points

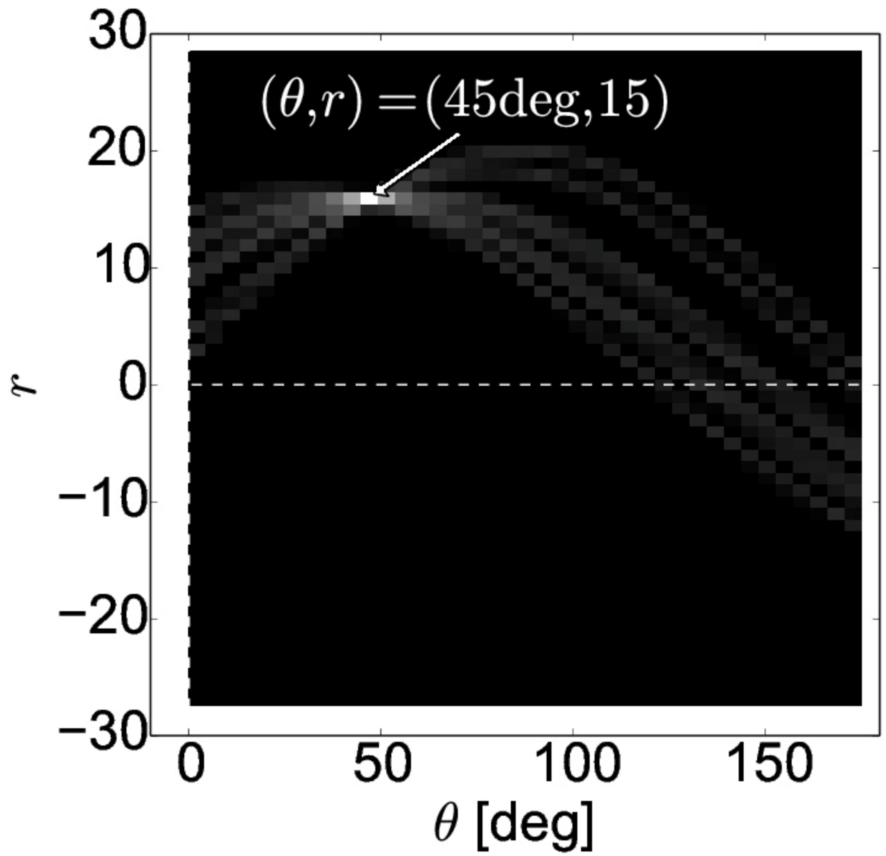


Accumulated votes from all points

HT for Straight Lines (6)

Multiple points; accumulating votes





Line with maximum number of votes

Accumulator maximum

HT for Straight Lines (7)

CENTER FOR MACHINE PERCEPTION

1. Define the *minimal* parametrization (p, q) of the space of lines:
   - Most common: angle – distance from origin (θ,r)
   - Other options: tangent of angle – intercept (a,b) , nearest point to center, …
2. Quantize the Hough space:
   - Identify the maximum and minimum values of *a* and *b*, and the number of cells,
3. Create an accumulator array A(p,q); set all values to zero
4. (if gradient available) : For all edge points (x_i,y_i) in the image
   - Use gradient direction
   - Compute a from the equation
   - Increment A(p,q) by one

   (if gradient not available): For all edge points (x_i,y_i) in the image
   - Increment A(p,q) by one for all lines incident on x,y
5. For all cells in A(p,q)
   - Search for the maximum value of A(p,q)
   - Calculate the equation of the line
6. To reduce the effect of noise more than one element (elements in a neighborhood) in the accumulator array are increased

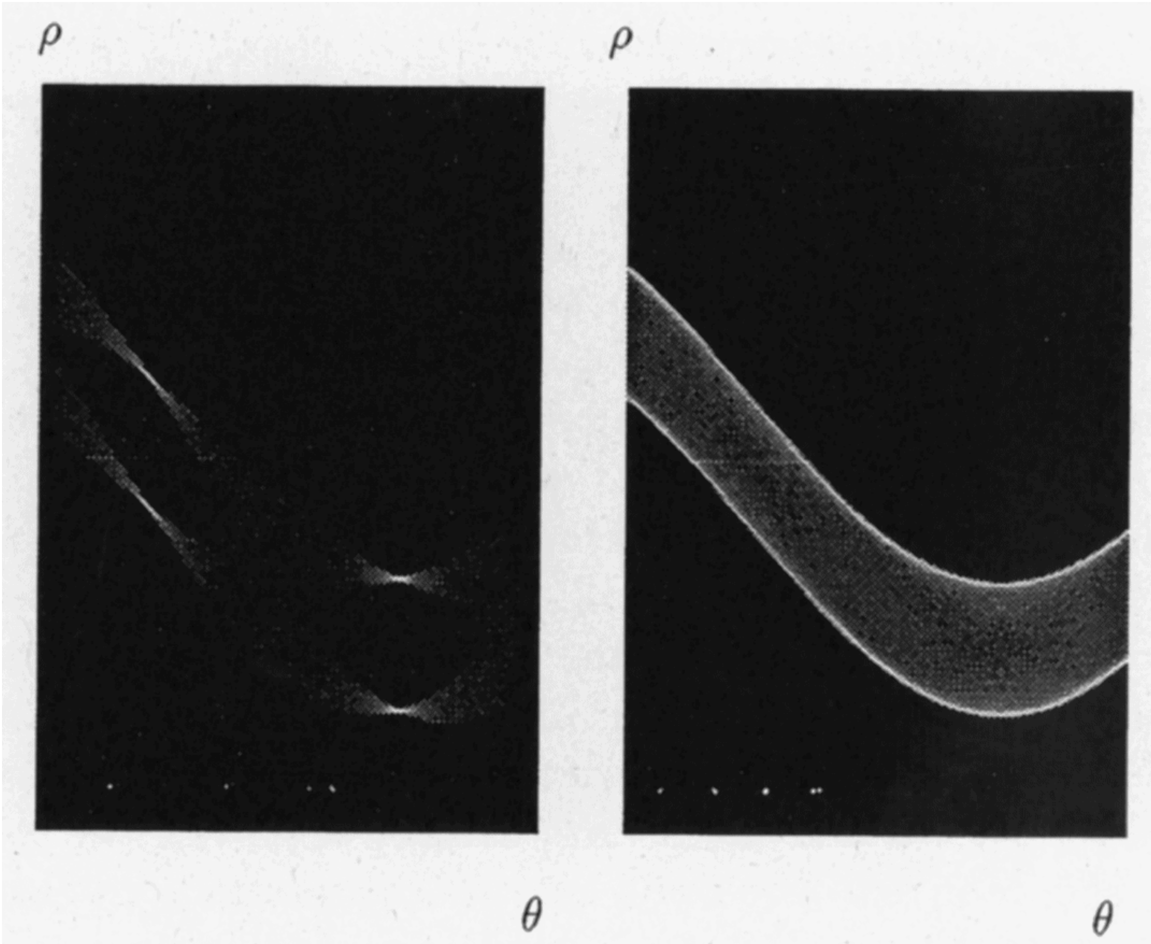
HT for Straight Lines: Variations

- Besides the discussed representation:
  - The form $y = a\,x + b$ has a singularity around 90°.
    Can be overcome by considering two cases, $y = a\,x + b$ and $x = a\,y + b$

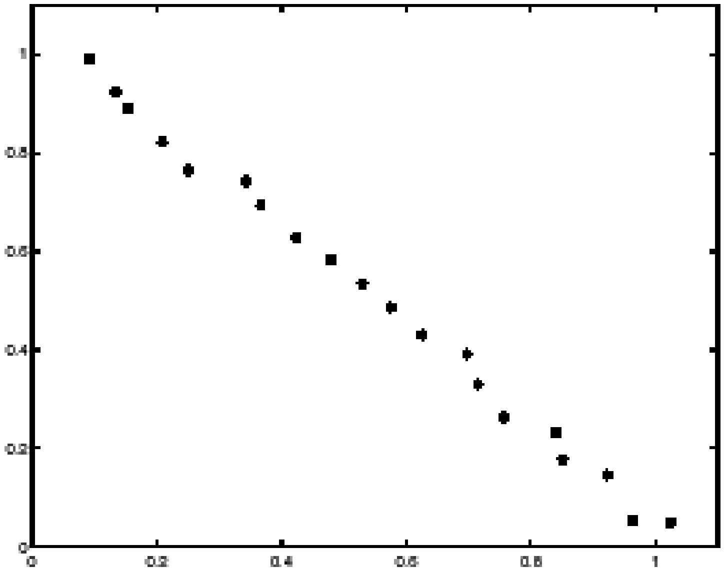
- Using gradient orientation
  - Uses not only point but also orientation consistent with the edge orientation
  - Variables: $P, \Omega, \phi : P \to \langle 0, \pi)$
  - In HT: for $\Omega(\mathbf{x}_i, \phi(\mathbf{x}_i))$
  - Can be used by weighting the strength of the vote by: $|\phi - \psi|$
    $\psi \ldots$ line orientation, $\phi \ldots$ gradient orientation

Examples

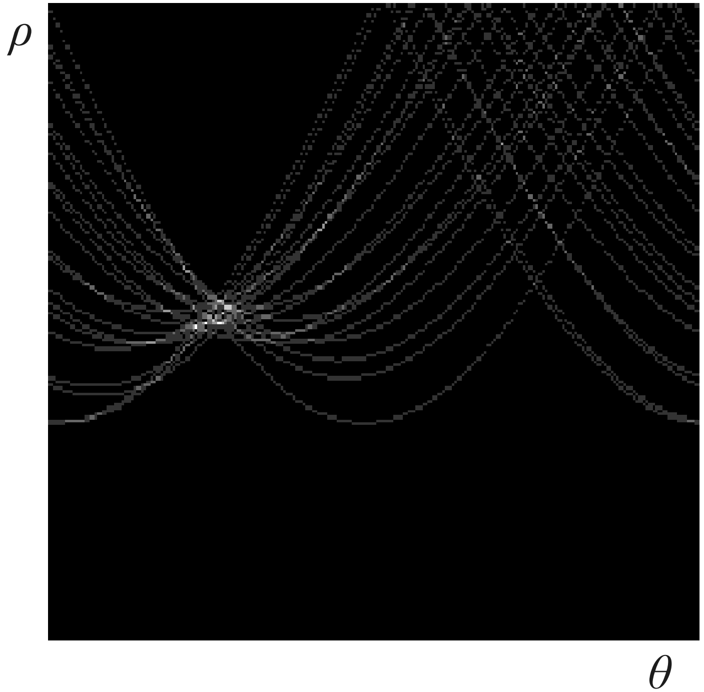
- Hough transform for a square (left) and a circle (right)



Hough Transform: Noisy Line



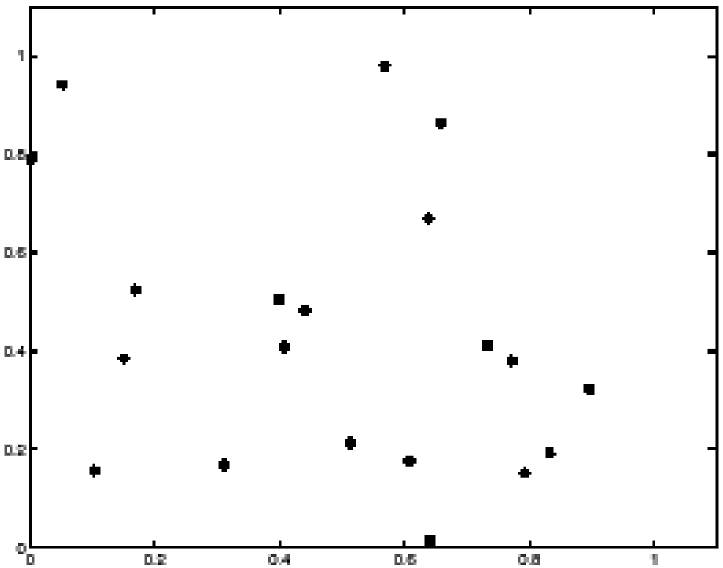
Tokens



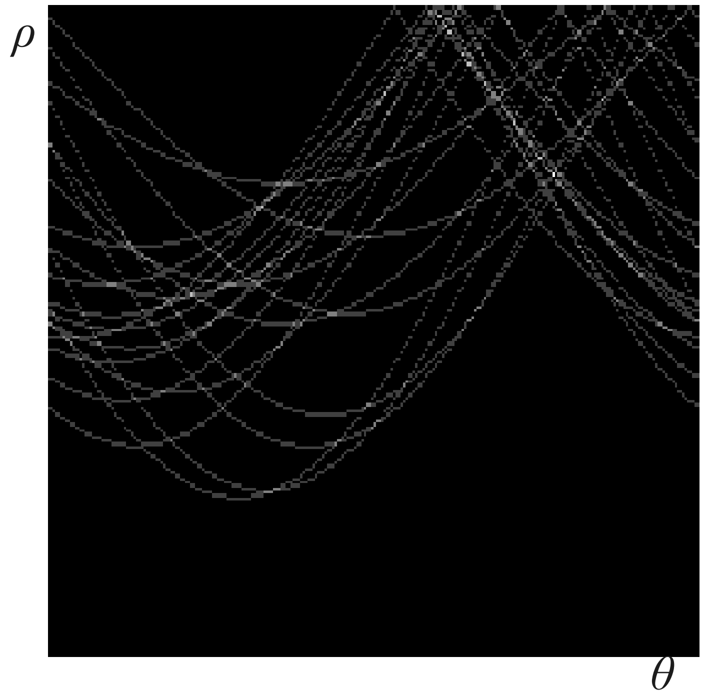
Votes

- Problem: Finding the true maximum

Hough Transform: Noisy Input



Tokens



Votes

- Problem: Lots of spurious maxima

HT for different primitives (1)

Circles with known, fixed radius

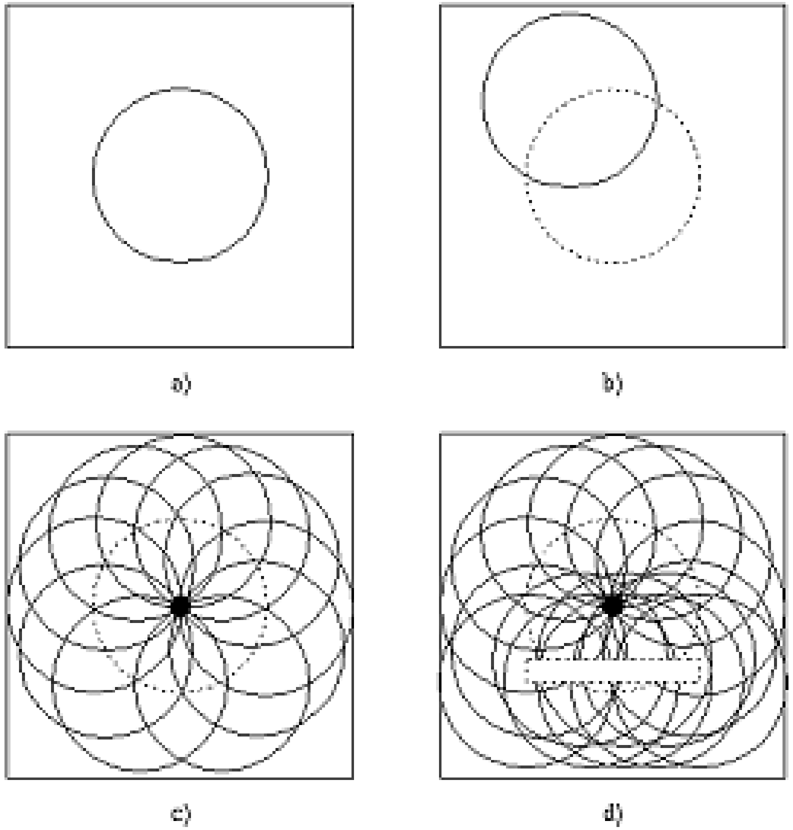
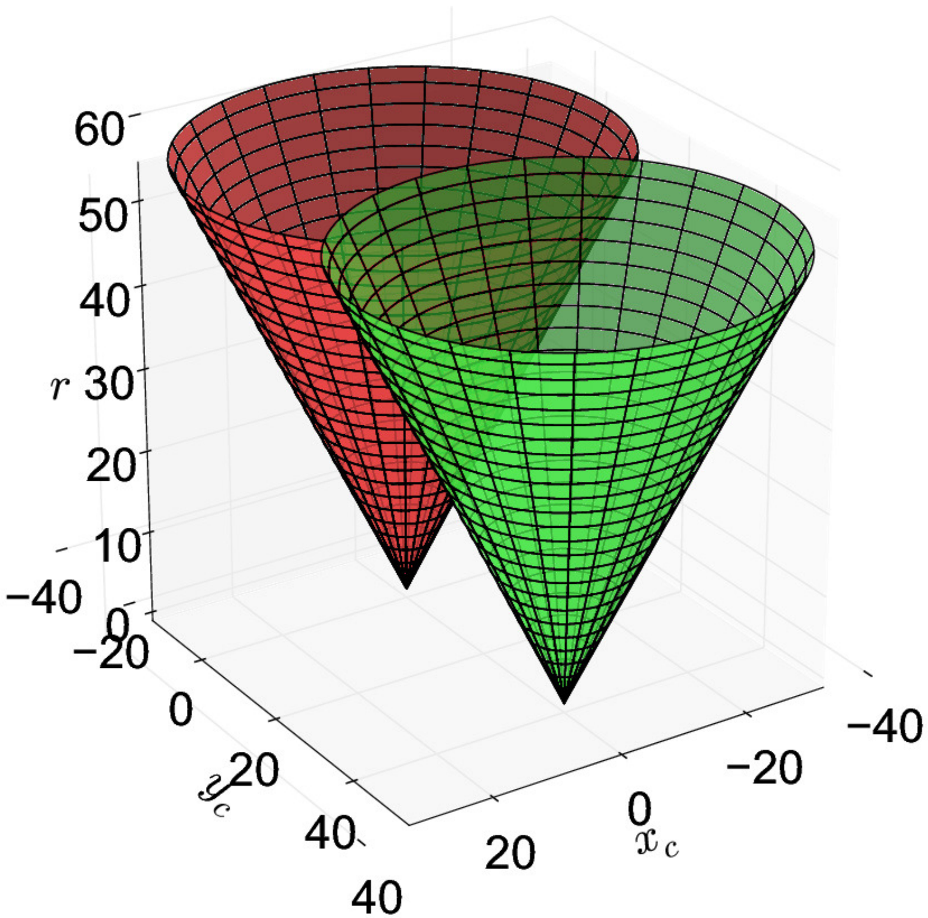


Figure 5.29 *Hough transform - example of circle detection. (a) Original image of a dark circle (known radius r) on a bright background, (b) for each dark pixel, a potential circle-center locus is defined by a circle with radius r and center at that pixel, (c) the frequency with which image pixels occur in the circle-center loci is determined; the highest-frequency pixel represents the center of the circle (marked by •), (d) the Hough transform correctly detects the circle (marked by •) in the presence of incomplete circle information and overlapping structures (see Figure 5.34 for a real-life example).*

HT for different primitives (2)



Voting surface for a point at (0,0) and at (0, 40)

Circles with unknown radius

$$(x - x_c)^2 + (y - y_c)^2 = r^2$$

(x,y) : point coordinates
(x_c, y_c) : circle centre
r : circle radius

The accumulator is 3-dimensional

HT for multiple instances

1. $p_1 = HT(P, \Omega)$: strongest result of HT
2. Set $P_1 = P \setminus p_1$
3. Unvote p_1
4. $p_2 = HT(P_1, \Omega)$
5. Cont. to get as many instances as required

- Greedy
- Sequential

Hough Transform Problems

CENTER FOR MACHINE PERCEPTION

1. Search space (accumulator size) gets prohibitively large easily
   - Line segments: θ, ρ, t_1, t_2
   - Circular arc: r, c_x, c_y, t_1, t_2
2. Cost function must be additive.
3. Greedy assignment rule of a token to primitive
4. No global objective function for multiple primitives (global optimization for one primitive only)

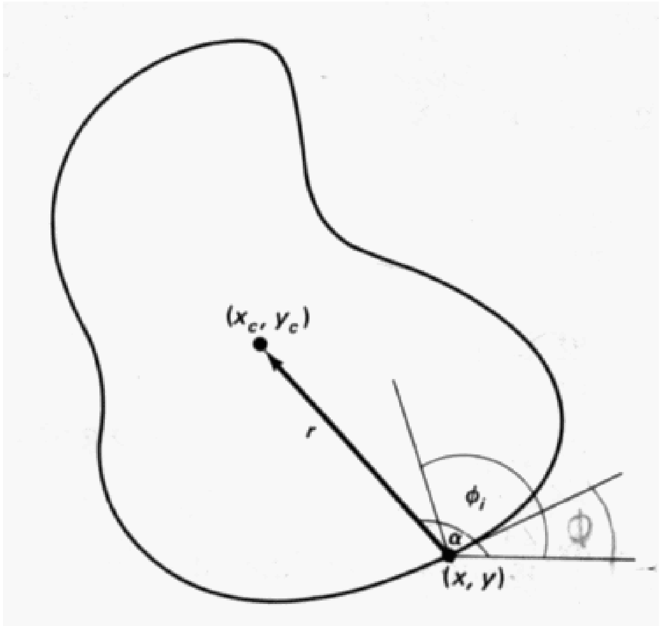
When is the Hough transform useful?

CENTER FOR MACHINE PERCEPTION

- Textbooks often imply that it is useful mostly for finding lines
  - In fact, it can be very effective for recognizing arbitrary shapes or objects (Generalized HT)
- The key to efficiency is to have each feature (token) determine as many parameters as possible
  - For example, lines can be detected much more efficiently from small edge elements (or points with local gradients) than from just points
  - For object recognition, each token should predict location, scale, and orientation (4D array)
- Bottom line: The Hough transform can extract feature groupings from clutter in linear time!

Generalized Hough Transform [Ballard81]

- Generalization for an arbitrary contour or shape
  - Choose reference point for the contour (e.g. center)
  - For each point on the contour remember where it is located w.r.t. to the reference point
  - Remember radius r and angle ϕ relative to the contour tangent
  - Recognition: whenever you find a contour point, calculate the tangent angle and 'vote' for all possible reference points

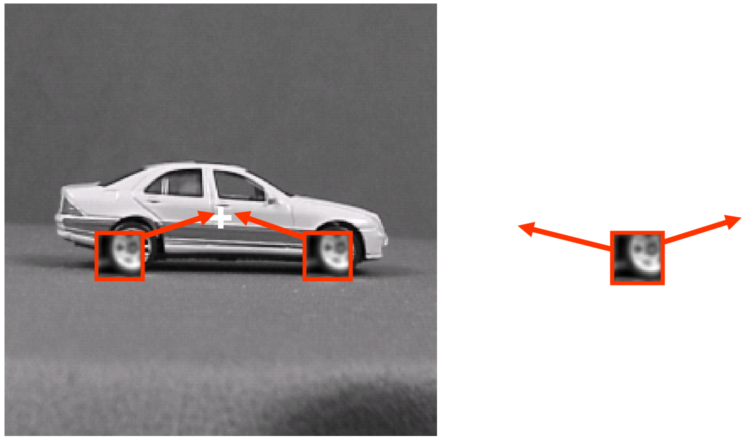


  - Instead of reference point, can also vote for transformation

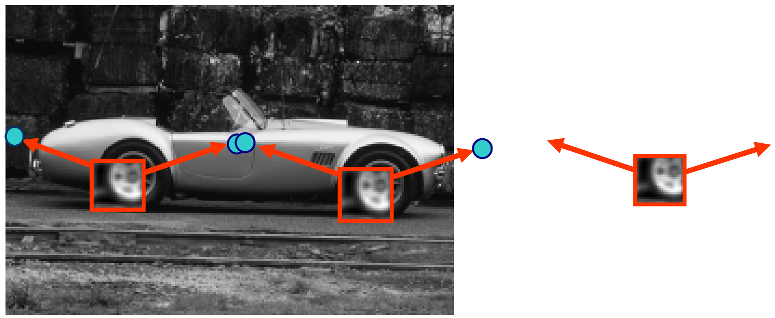
⇒ The same idea can be used with local features!

Gen. Hough Transform with Local Features

- For every feature, store possible “occurrences”



For new image, let the matched features vote for possible object positions

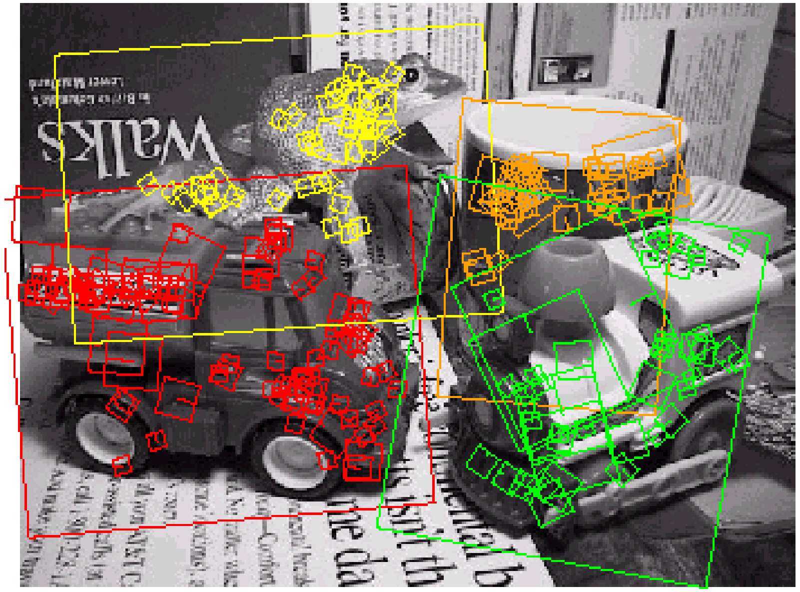


– Object identity
– Pose
– Relative position

Finding Consistent Configurations

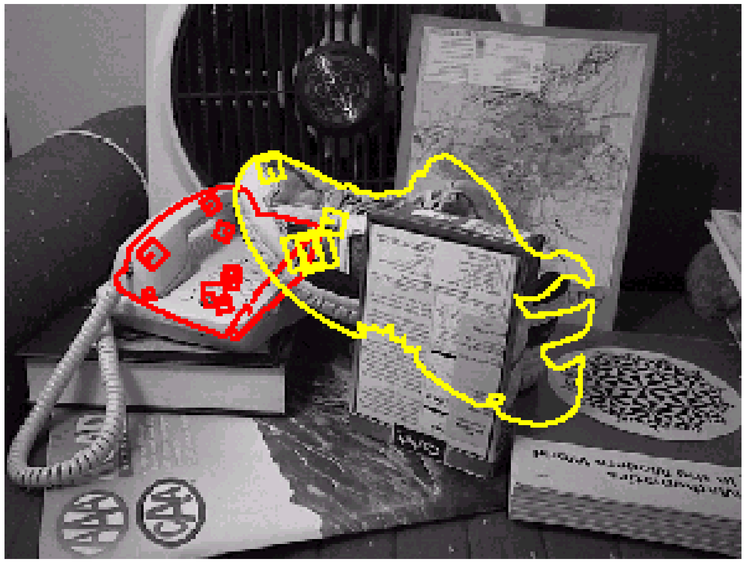
- Global spatial models
  - Generalized Hough Transform [Lowe99]
  - RANSAC [Obdrzalek02, Chum05, Nister06]
  - Basic assumption: object is planar

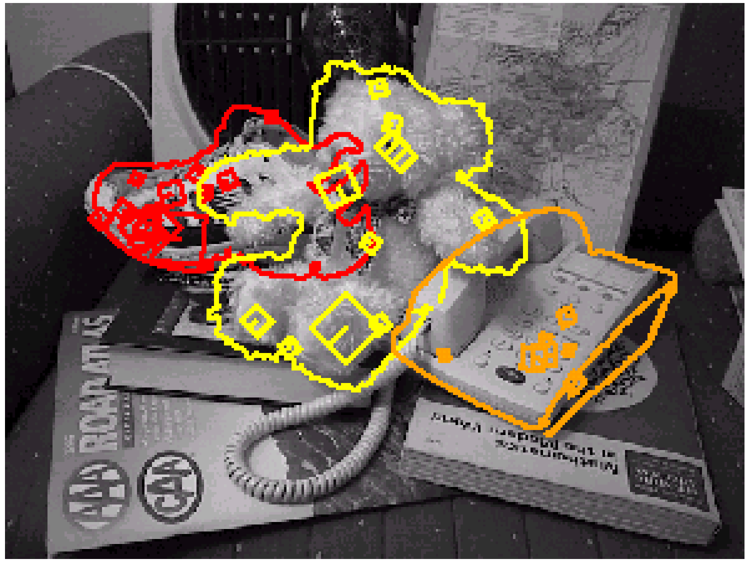
- Assumption is often justified in practice
  - Valid for many structures on buildings
  - Sufficient for small viewpoint variations on 3D objects



3D Object Recognition

- Gen. HT for Recognition
  - Typically only 3 feature matches needed for recognition
  - Extra matches provide robustness
  - Affine model can be used for planar objects





Comparison

Gen. Hough Transform

- Advantages
  - Very effective for recognizing arbitrary shapes or objects
  - Extracts groupings from clutter in linear time
- Disadvantages
  - Quantization issues
  - Only practical for small number of dimensions (up to 4)
- Improvements available
  - Probabilistic Extensions
  - Continuous Voting Space

  } **[Leibe08]**

RANSAC

- Advantages
  - General method suited to large range of problems
  - Easy to implement
  - Independent of number of dimensions
- Disadvantages
  - Standard implementation finds a single model at a time (cf. Hough Transform)
- Many variants available, e.g.
  - PROSAC: Progressive RANSAC [Chum05]
  - Preemptive RANSAC [Nister05]

RHT = Randomized Hough Transform [Xu93]

CENTER FOR MACHINE PERCEPTION

In: $E = \{e_i\}, m(\Omega, e) = 0$

Out: $\Omega_{S_1}, \Omega_{S_2}, ..., \Omega_{S_N}$

Repeat:

I. Hypothesis

1. Select random M feature points $e_{k_1}, ..., e_{k_M}$
2. Compute $\Omega_k : m(\Omega_k, e_{k_j}) = 0, j = 1, ..., M$

II. Pre-Verification

3. Add 1 to accumulator Ω_k
4. If (accumulator$(\Omega_k) > T_1$) goto *III.*

   Else goto *I.*

III. Verification

5. Find support for Ω_k
6. If (support $(\Omega_k) > T_2$) output Ω_k
7. Reset accumulator

Proc v Out Omega-S-N, a v algoritmu Omega-N

Probabilistic Hough Transform [Kiryati et al. 91]

CENTER FOR MACHINE PERCEPTION

Idea: Evaluate $\sum_{i=1}^{N} p(x_i, \Omega)$ using only a fraction $f = \frac{k_{MAX}}{N}$ of N points x_i

Algorithm:

1. Select k_{MAX} points at random
2. Perform standard HT

Analysis:

- Selection of k_{MAX} is incorrect

  $\Rightarrow$ the number *L* of selected points from L_N points of a line in a random subset of k_{MAX} points is governed by hypergeometric, not binomial distance

$$P(L_N) = \frac{\binom{L}{L_N}\binom{N-L}{k_{MAX}-L_N}}{\binom{N}{k_{MAX}}}$$

$$\mu = \overline{N} \qquad \sigma^2 = k_{MAX}\frac{L_N(N-L_N)}{N^2}\left(1 - \frac{k_{MAX}-1}{N-1}\right)$$

PHT = Monte Carlo Evaluation of $\sum_{i=1}^{N} p(x_i, \Omega)$

CENTER FOR MACHINE PERCEPTION

Idea:

1. Evaluate $\sum_{i=1}^{N} p(x_i, \Omega)$ using only a fraction $f = \frac{k_{MAX}}{N}$ of N points x_i
2. Apply standard MC analysis to find k_{MAX} in PHT to guarantee $P\{\text{false\_positive}\}$ and $P\{\text{false\_negative}\} < \epsilon$

Algorithm:

1. Select a random point
2. Vote and return it
3. Finish if k_{MAX} reached

CHT = Cascaded Hough Transform [Tuytelaars et al. 97]

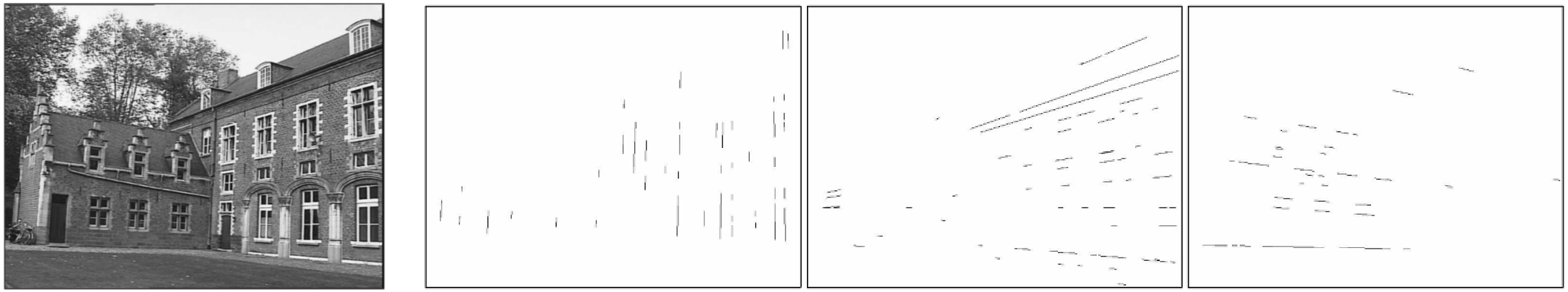
CENTER FOR MACHINE PERCEPTION

- Finds structures at different hierarchical levels by iterating one kind of HT (fixed points, fixed lines, lines of fixed points, pencils of fixed lines)
- Uses duality of lines and points in image and parameter spaces
- Algorithm:

1. First HT: detects lines in the image and keeps dominant peaks in the parameter space
2. Second HT: detects lines of collinear peaks in parameter space and keeps vertices where several straight lines in the original image intersect (*vanishing points*)
3. Third HT: applied to the peaks of thto detect collinear vertices (*vanishing lines*)

layer	meaning of detected features
layer 0	(the original image)
Hough 1 ↓	
layer 1	points ~ lines lines ~ convergent lines
Hough 2 ↓	
layer 2	points ~ intersection points lines ~ collinear intersection points
Hough 3 ↓	
layer 3	points ~ lines of intersection points

CHT: Experiments



Lines belonging to one of the three detected vanishing points

Aerial image of buildings and streets (left), the corresponding edges (right)

macros.tex
sfmath.sty
cmpitemize.tex

Thank you for your attention.