Computer Vision

Bogdan Alexe

bogdan.alexe@fmi.unibuc.ro

University of Bucharest, 2nd semester, 2020-2021

Course structure

1. **Features and filters: low-level vision**

   Linear filters, color, texture, edge detection

2. **Grouping and fitting: mid-level vision**

   Fitting curves and lines, robust fitting, RANSAC, Hough transform, segmentation

3. **Multiple views**

   Local invariant feature and description, epipolar geometry and stereo, object instance recognition

4. **Object Recognition: high – level vision**

   Object classification, object detection, part based models, bovw models

5. **Video understanding**

   Object tracking, background subtraction, motion descriptors, optical flow

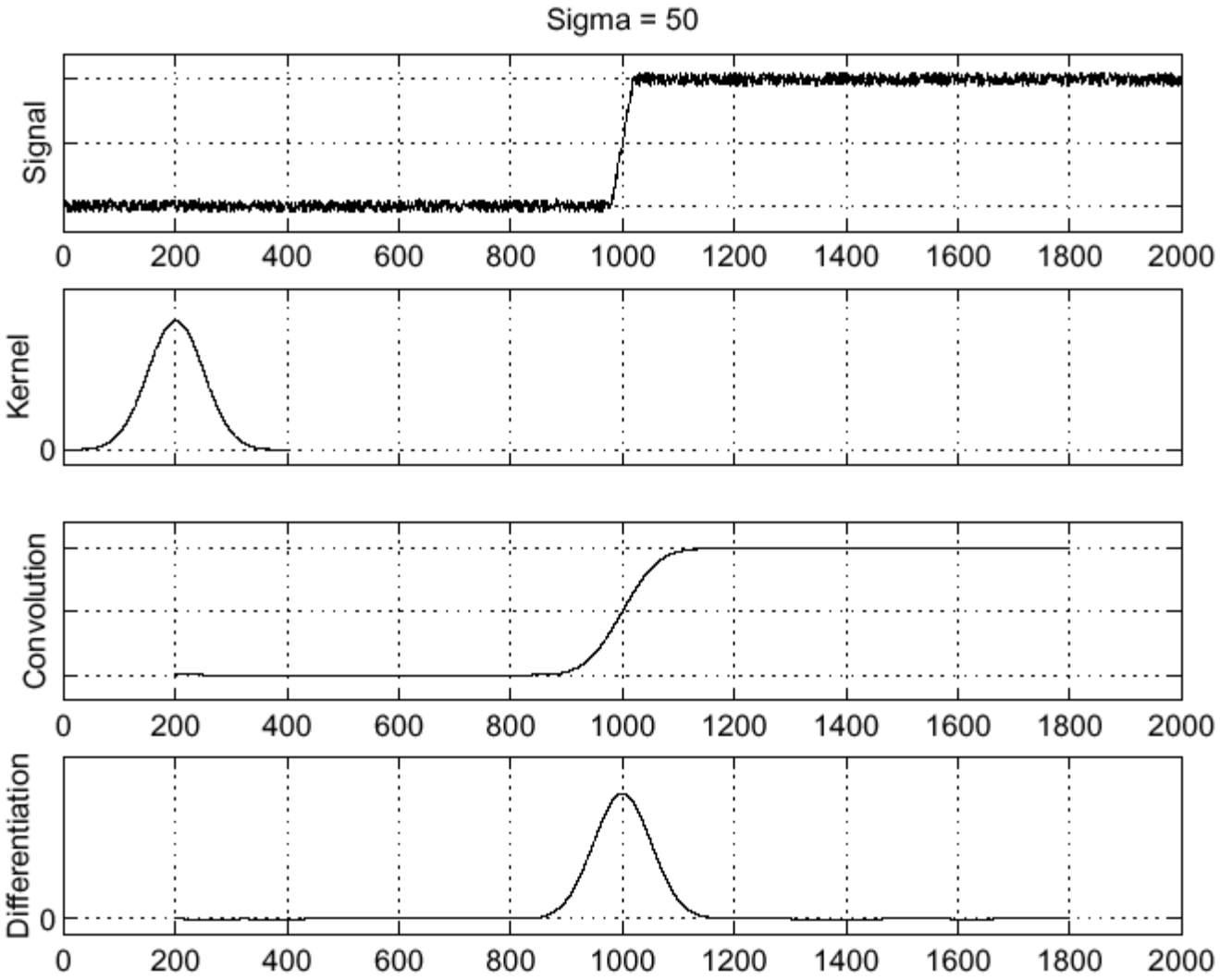
Edge detection using "Zero-crossings" – 1D

f

h

$h \star f$

$\frac{\partial}{\partial x}(h \star f)$



Where is the edge?

Look for peaks in $\frac{\partial}{\partial x}(h \star f)$ = zero's of the function $\frac{\partial^2}{\partial x^2}(h \star f)$

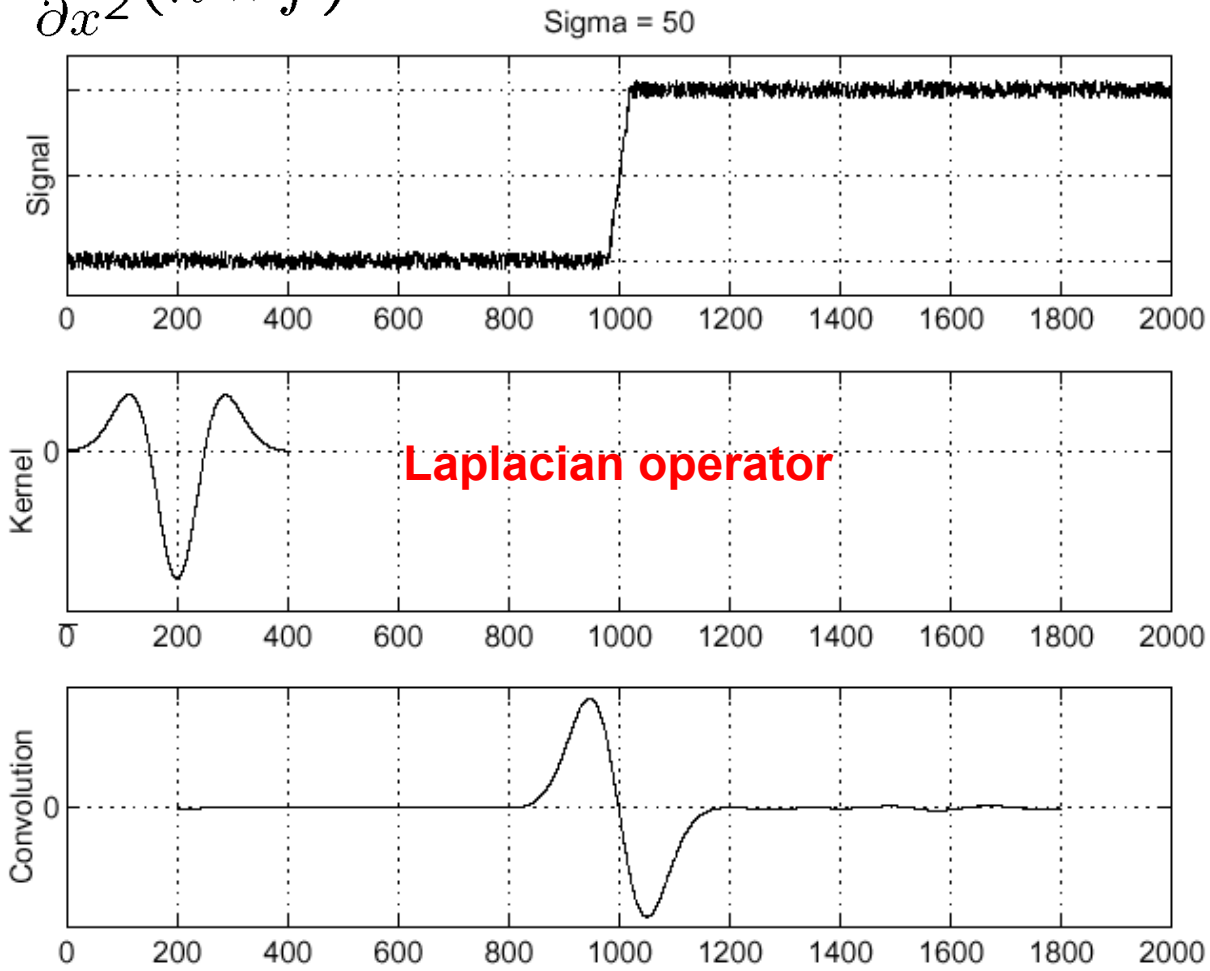
Edge detection using "Zero-crossings" – 1D

Consider $\frac{\partial^2}{\partial x^2}(h \star f)$

f

$\frac{\partial^2}{\partial x^2}h$

$(\frac{\partial^2}{\partial x^2}h) \star f$



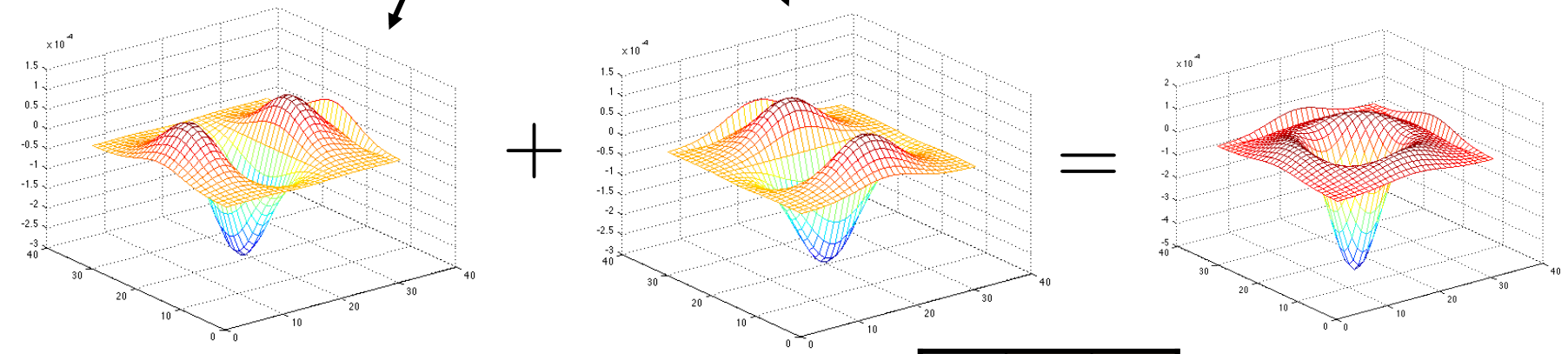
Where is the edge?

Zero-crossing of the function $(\frac{\partial^2}{\partial x^2}h) \star f$

Edge detection using "Zero-crossings" – 2D

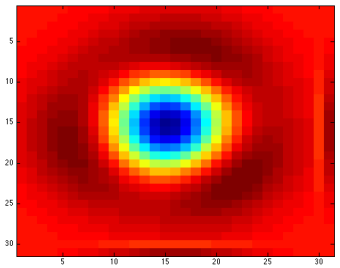
$$\nabla^2 h = \frac{\partial^2 h}{\partial x^2} + \frac{\partial^2 h}{\partial y^2}$$

Laplacian of Gaussian



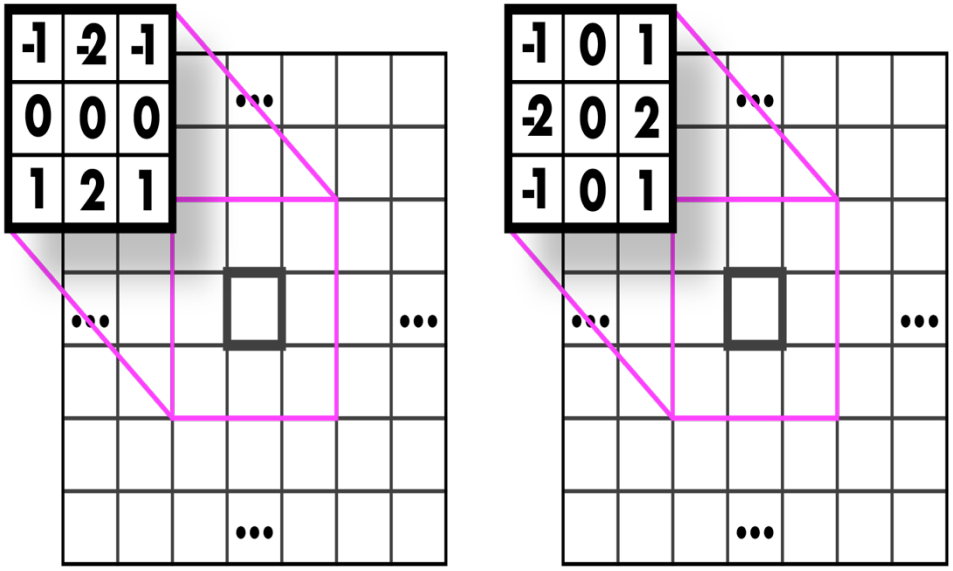
0	1	0
1	-4	1
0	1	0

Example of a Laplacian filter

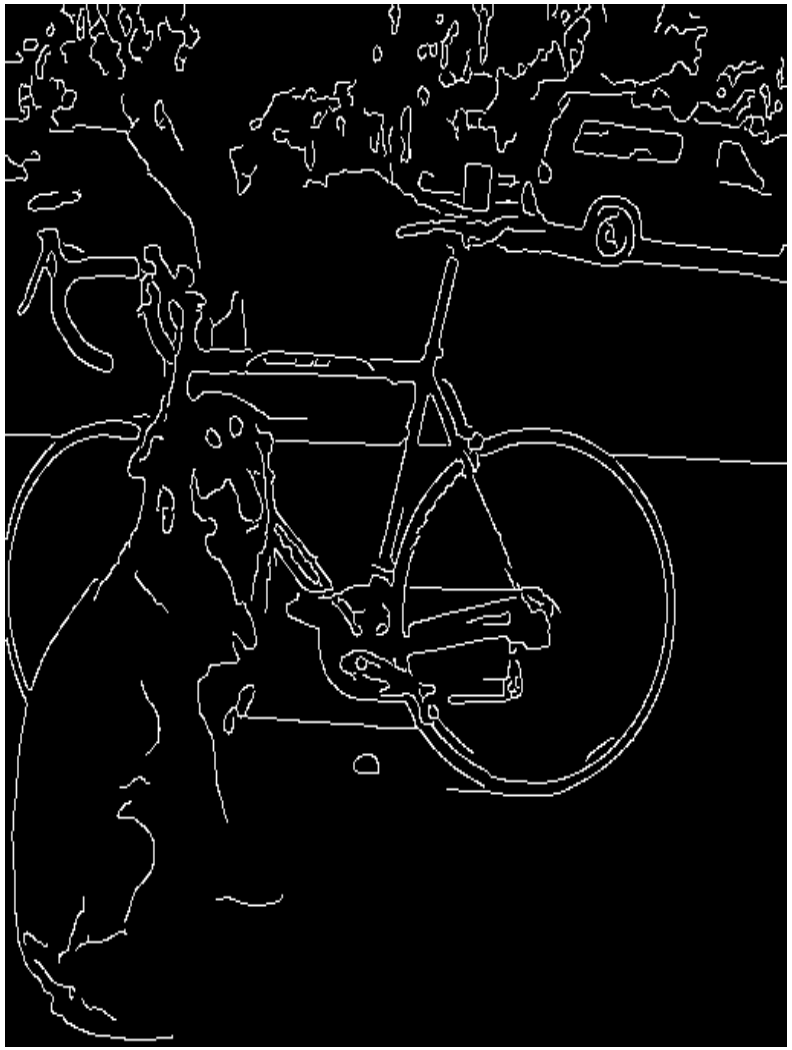


Another approach: gradient magnitude

- Don’t need 2nd derivatives
- Just use magnitude of gradient
- Are we done? No!



What we really want: line drawing



Canny Edge Detection

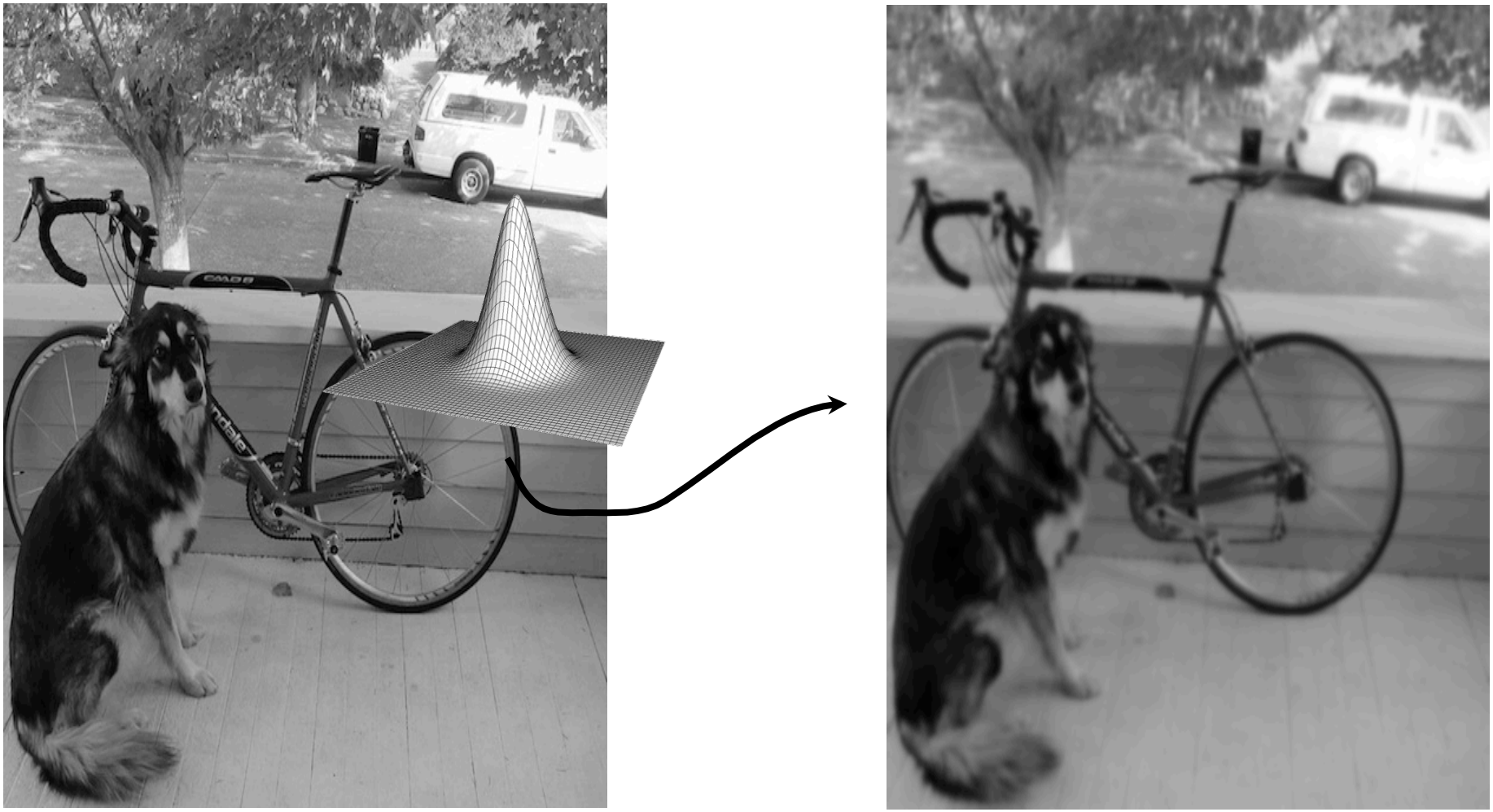
Algorithm:

- Smooth image (only want “real” edges, not noise)
- Calculate gradient direction and magnitude
- Non-maximum suppression perpendicular to edge
- Threshold into strong, weak, no edge
- Connect together components

http://bigwww.epfl.ch/demo/ip/demos/edgeDetector/

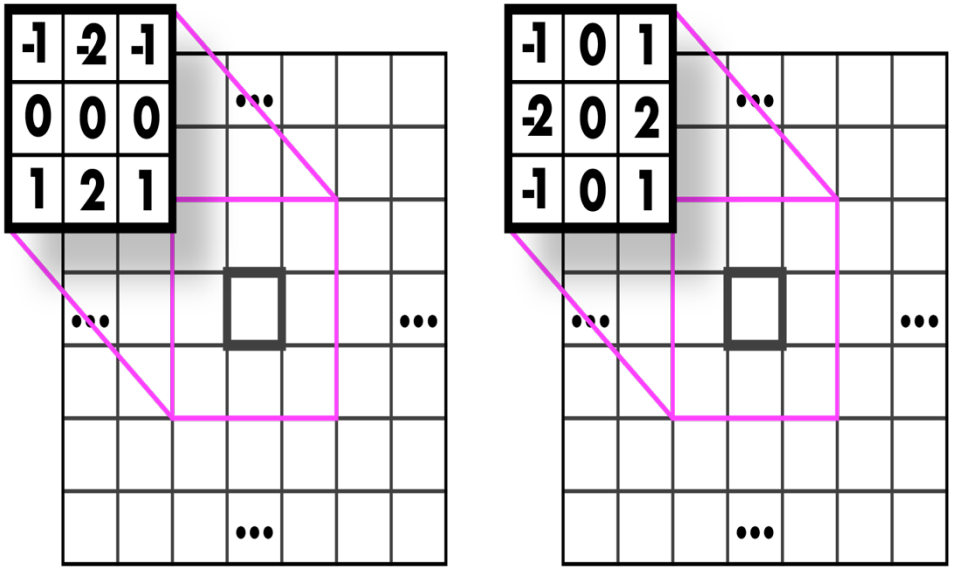
Smooth image

- Apply a Gaussian filter



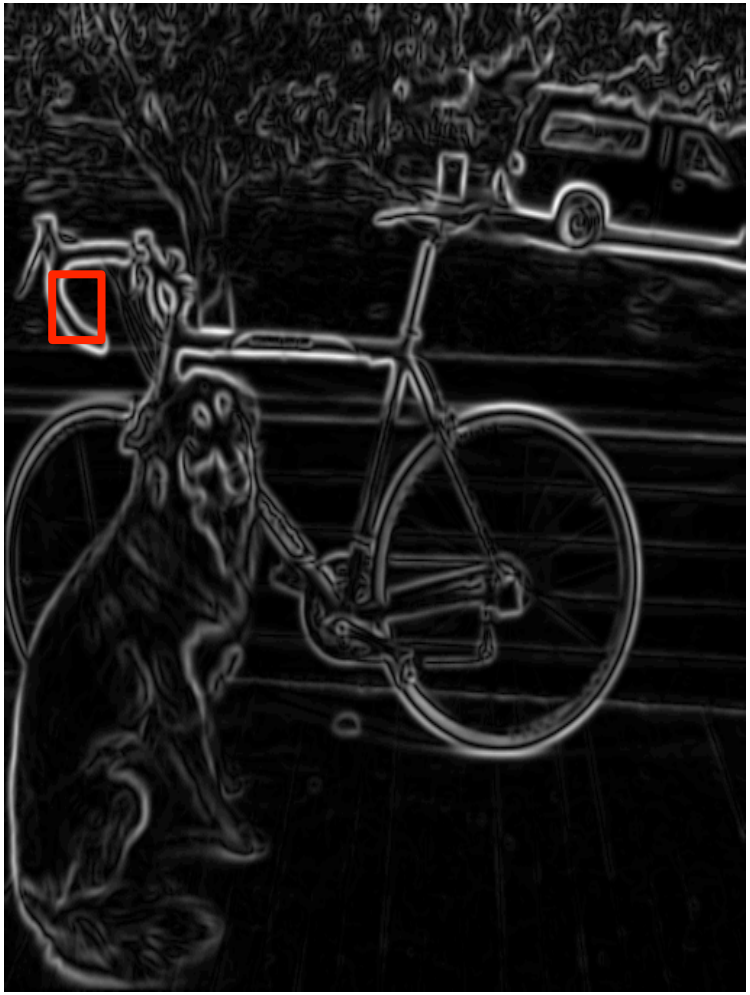
Gradient magnitude and direction

- Sobel filter

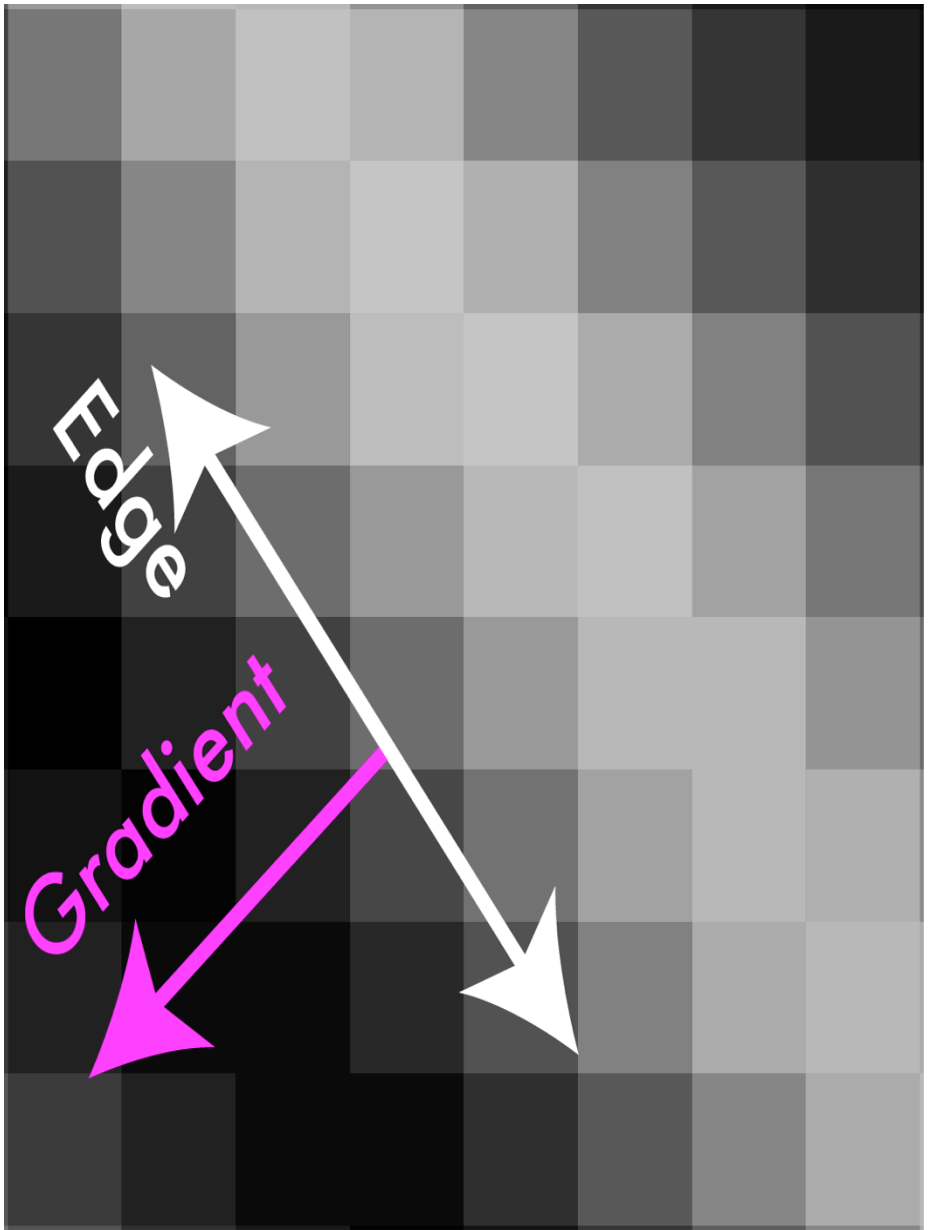


Non-maximum suppression

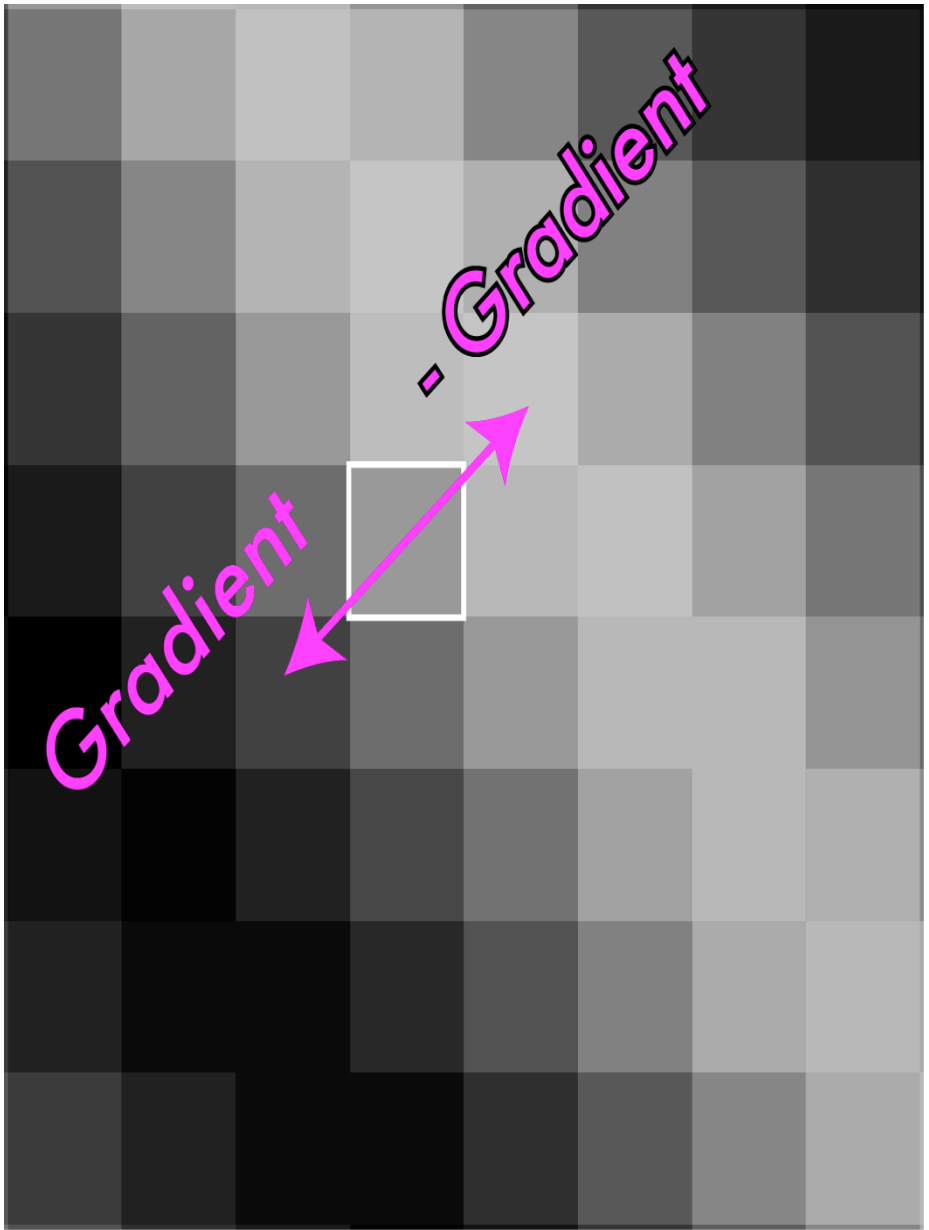
- Want single pixel edges, not thick blurry lines
- Need to check nearby pixels
- See if response is highest



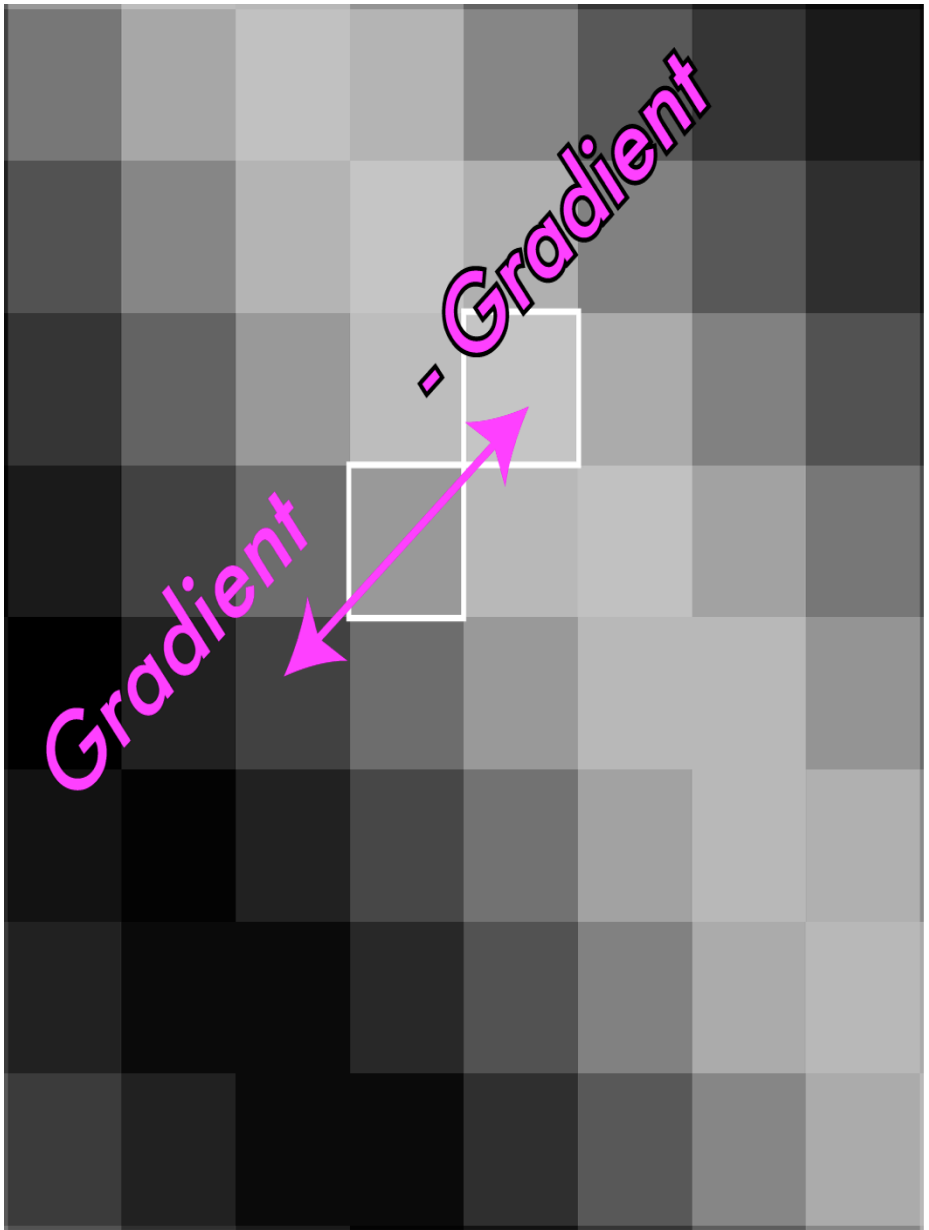
Non-maximum suppression



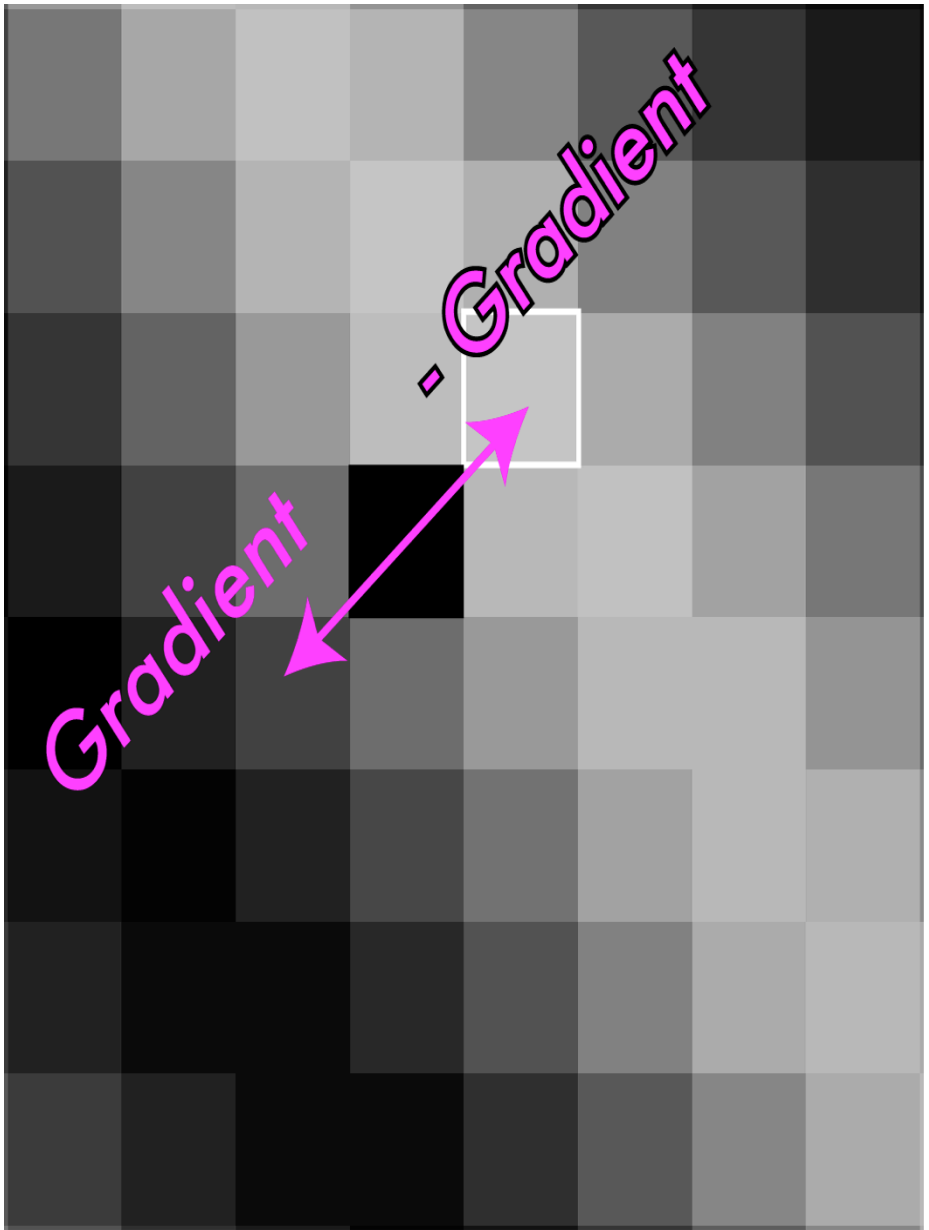
Non-maximum suppression



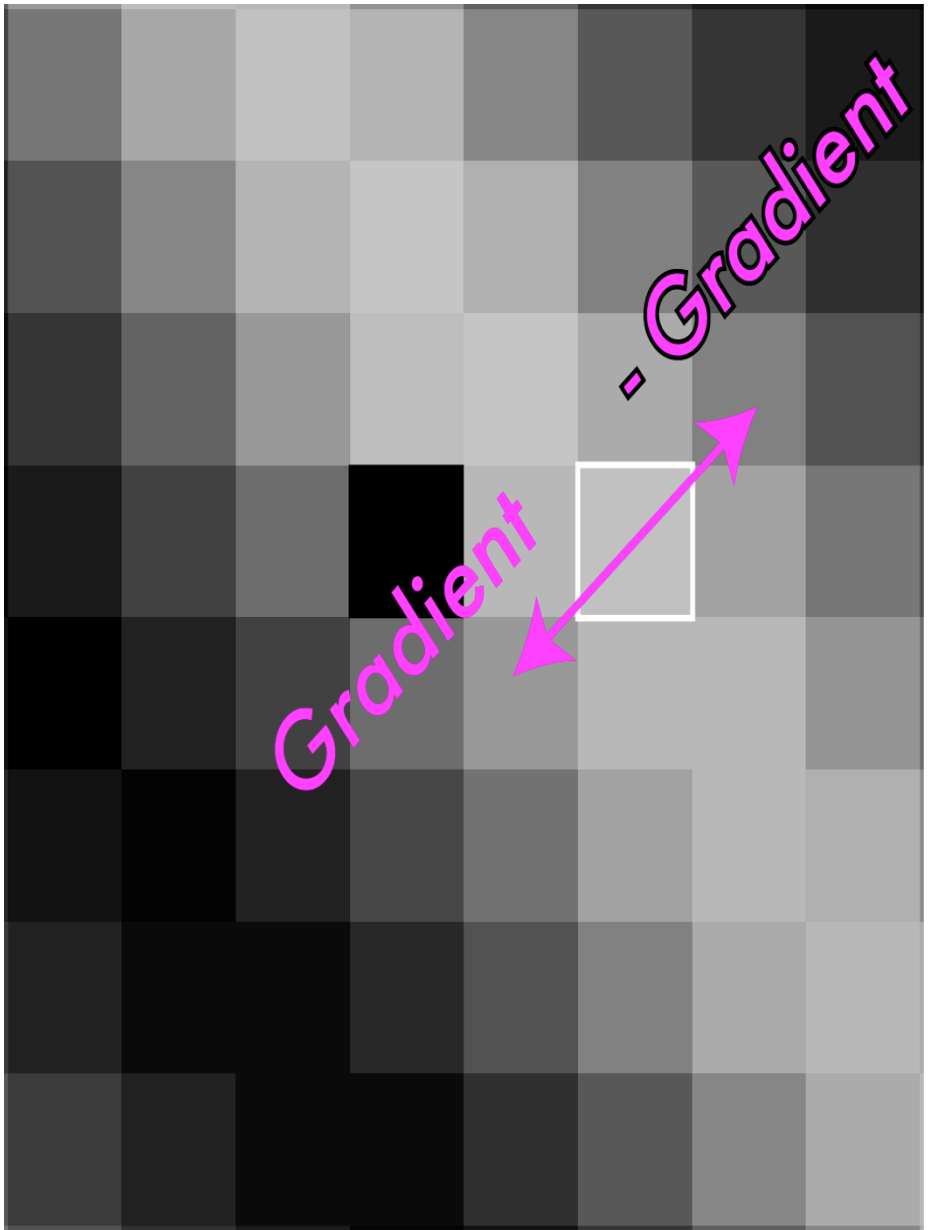
Non-maximum suppression



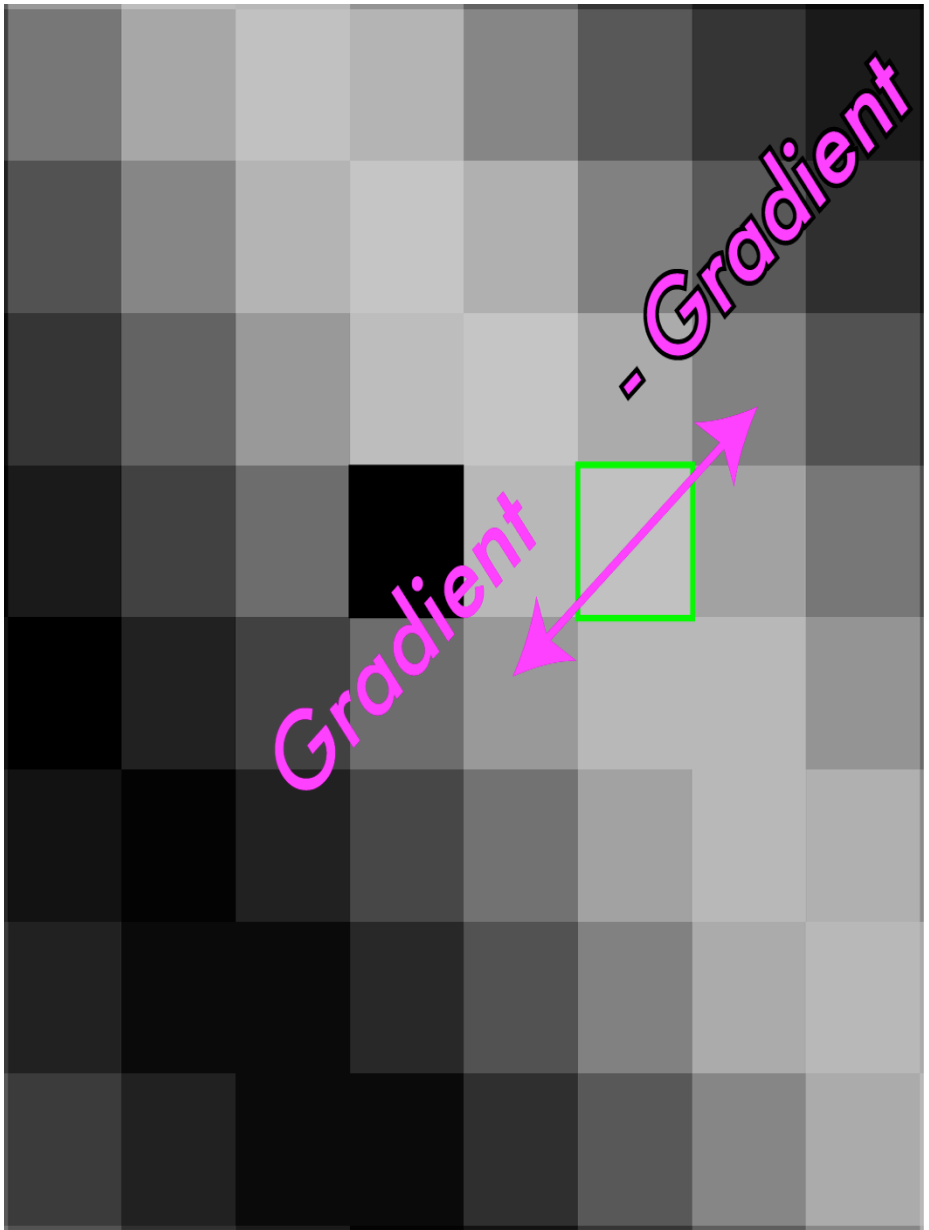
Non-maximum suppression



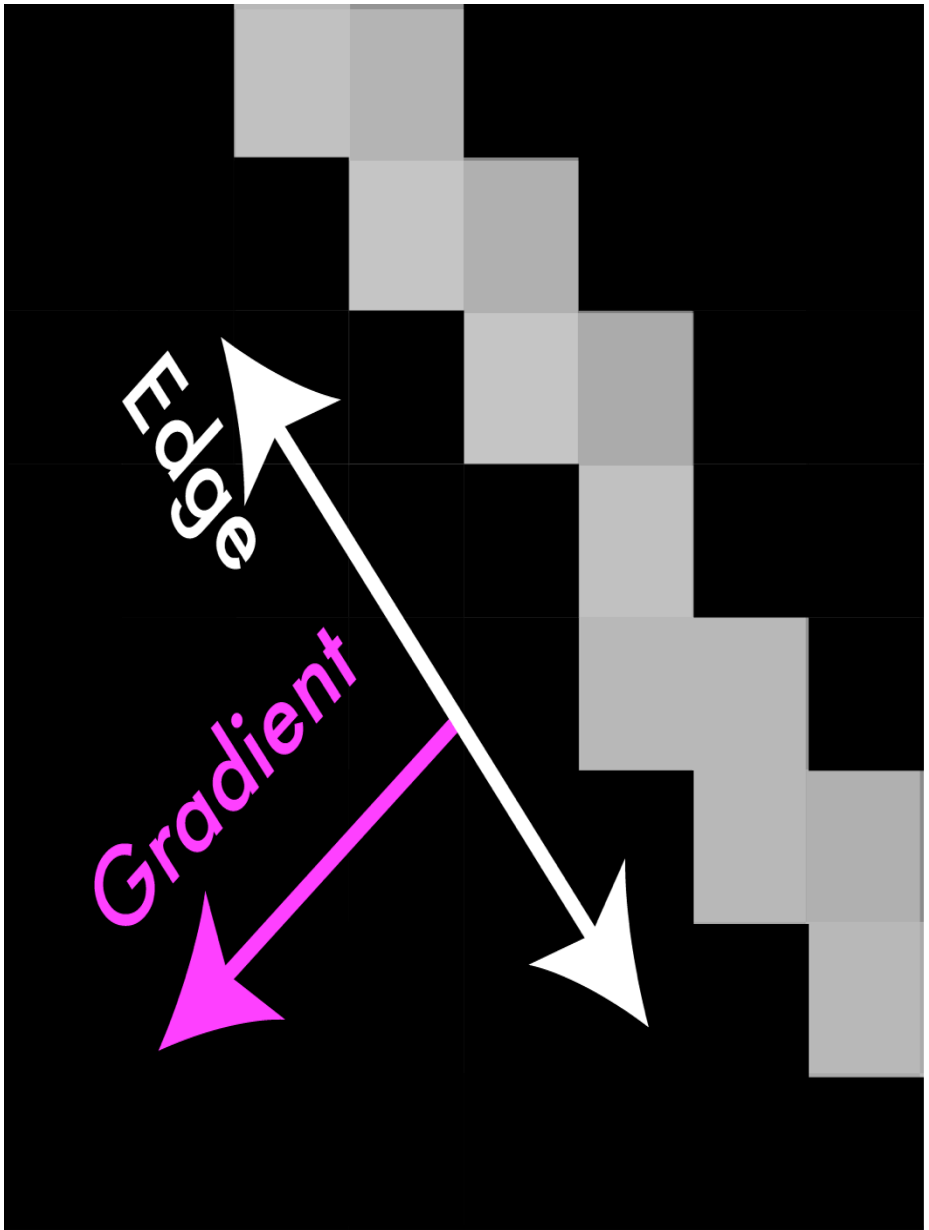
Non-maximum suppression



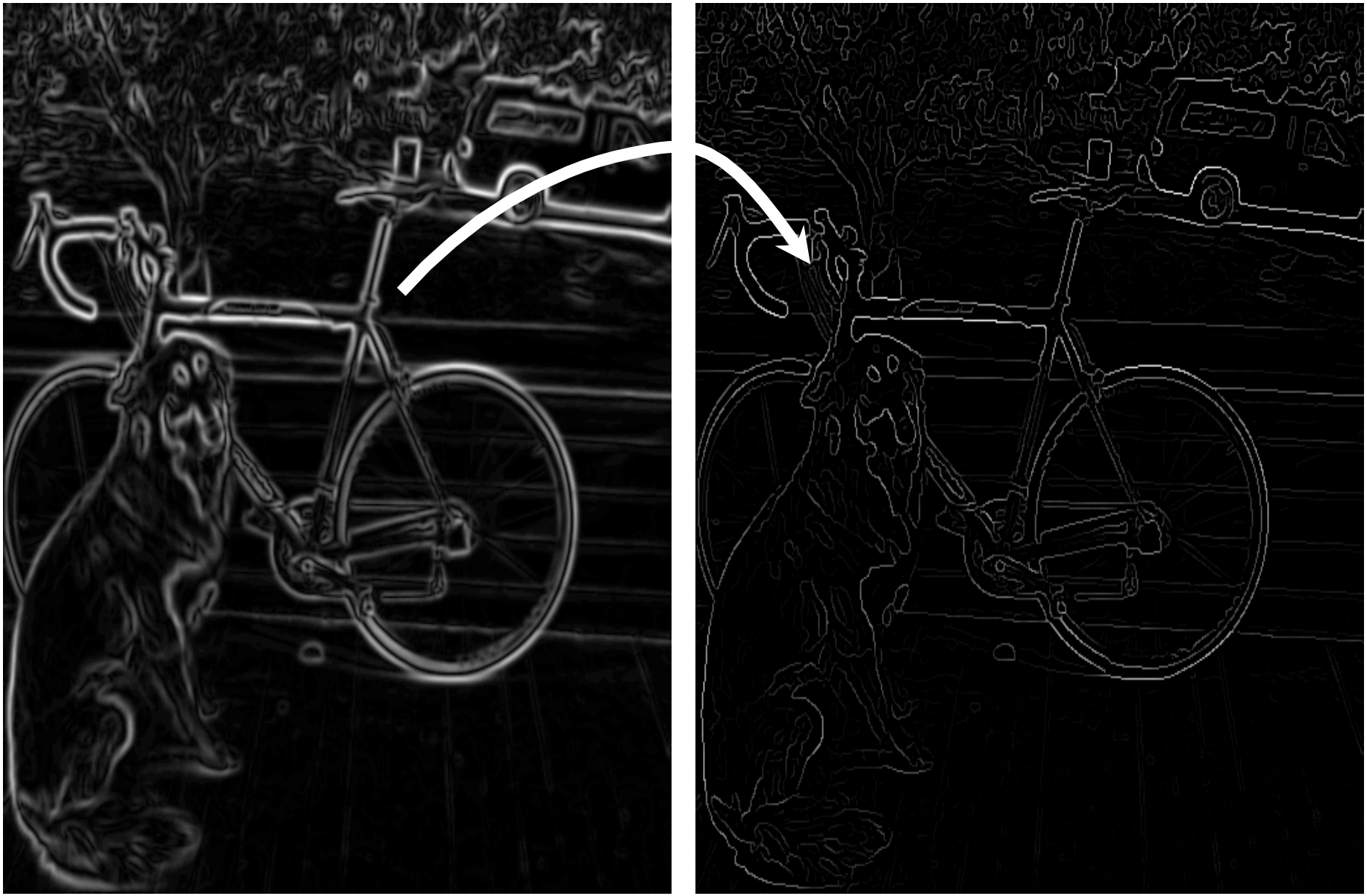
Non-maximum suppression



Non-maximum suppression

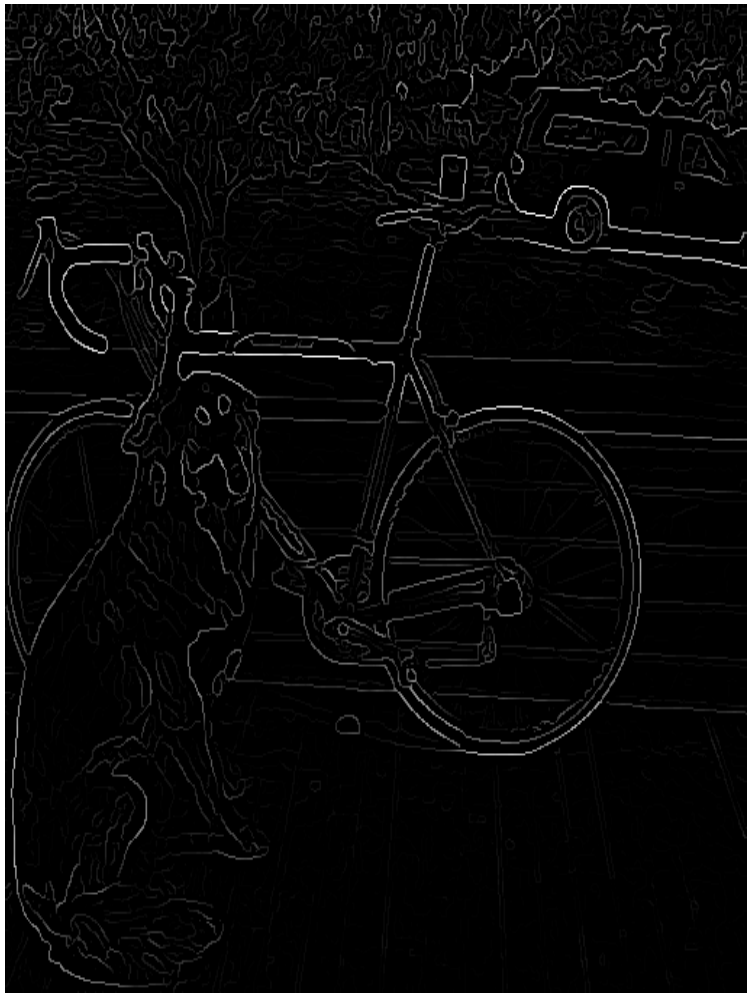


Non-maximum suppression



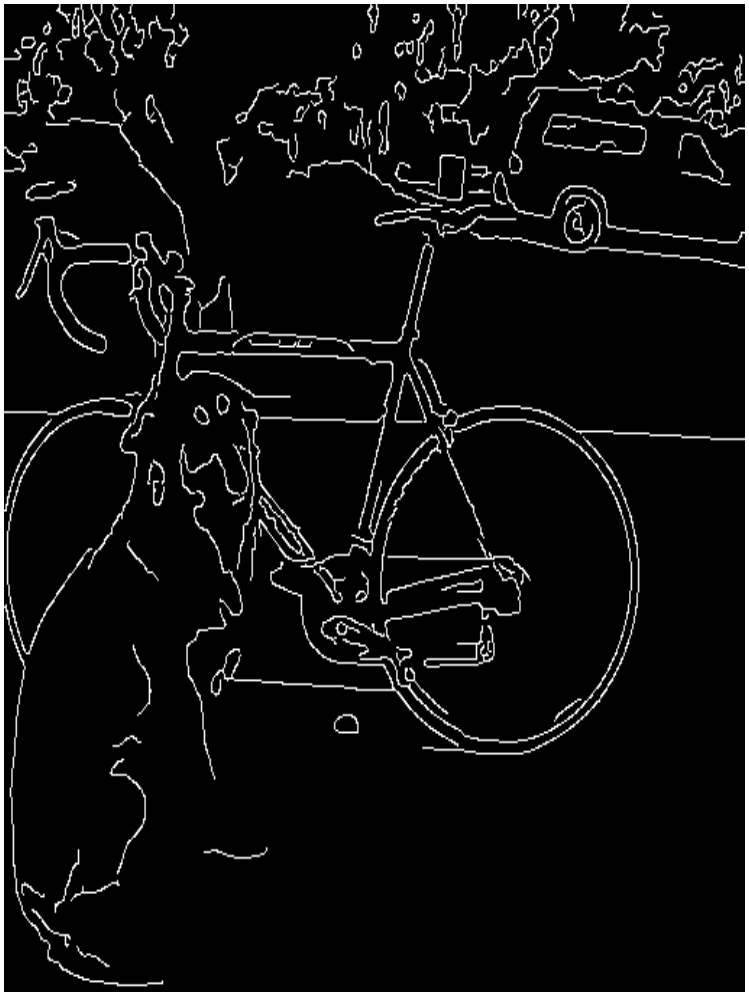
Threshold edges

- Still some noise
- Only want strong edges
- 2 thresholds, 3 cases
  - R > T: strong edge
  - R < T but R > t: weak edge
  - R < t: no edge
- Why two thresholds?



Connect them up!

- Strong edges are edges!
- Weak edges are edges
  iff they connect to strong
- Look in some neighborhood
  (usually 8 closest)

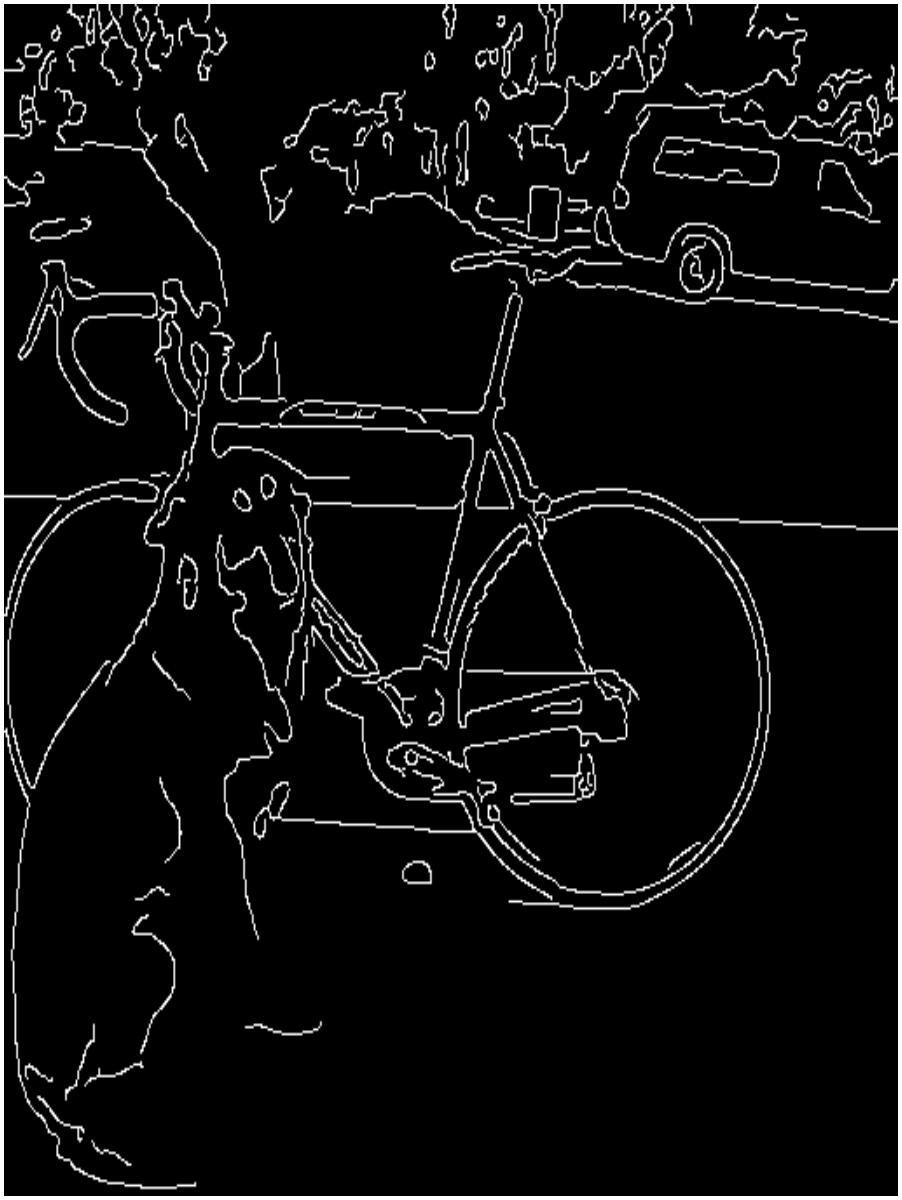


Canny Edge Detection

Algorithm:

- Smooth image (only want “real” edges, not noise)
- Calculate gradient direction and magnitude
- Non-maximum suppression perpendicular to edge
- Threshold into strong, weak, no edge
- Connect together components
- Tunable: Sigma, thresholds

Canny Edge Detection



Canny Edge Detection - demo

http://bigwww.epfl.ch/demo/ip/demos/edgeDetector/

Course structure

1. **Features and filters: low-level vision**

   Linear filters, color, texture, edge detection

2. **Grouping and fitting: mid-level vision**

   Fitting curves and lines, robust fitting, RANSAC, Hough transform, segmentation

3. **Multiple views**

   Local invariant feature and description, epipolar geometry and stereo, object instance recognition

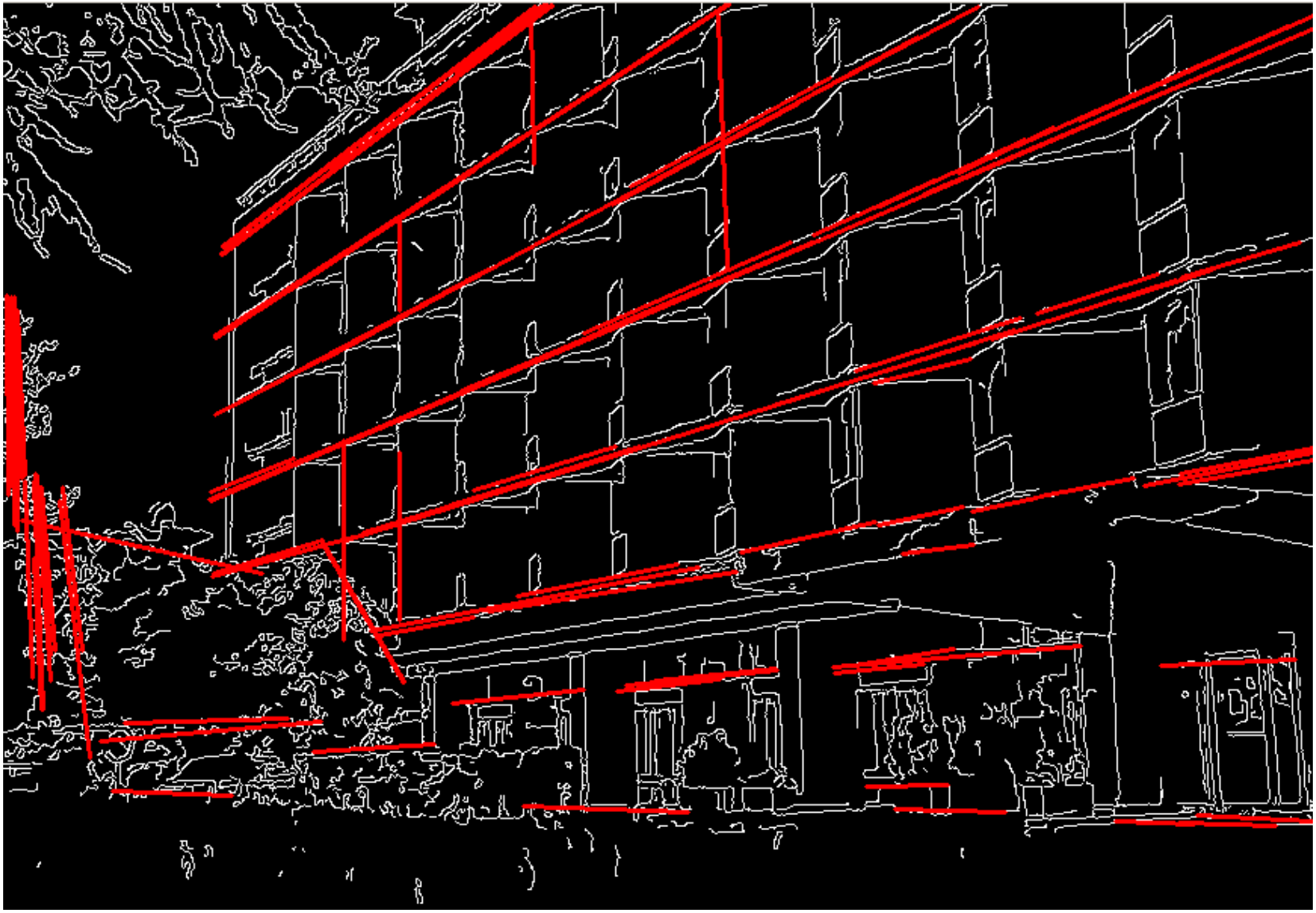
4. **Object Recognition: high – level vision**

   Object classification, object detection, part based models, bovw models

5. **Video understanding**

   Object tracking, background subtraction, motion descriptors, optical flow

Fitting

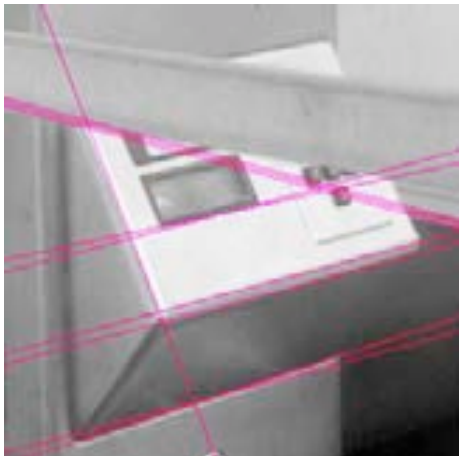


Fitting

- We have learned how to detect edges.
- We would like to form a higher-level, more compact representation of the features in the image by grouping multiple features according to a simple model

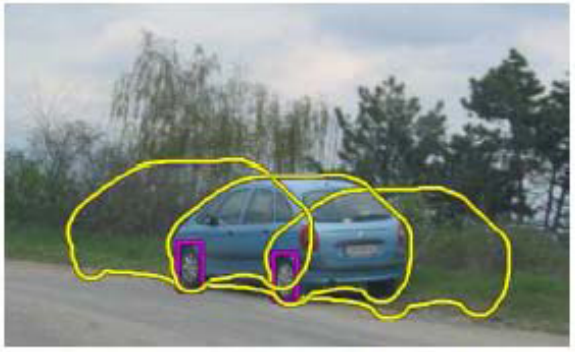
Fitting

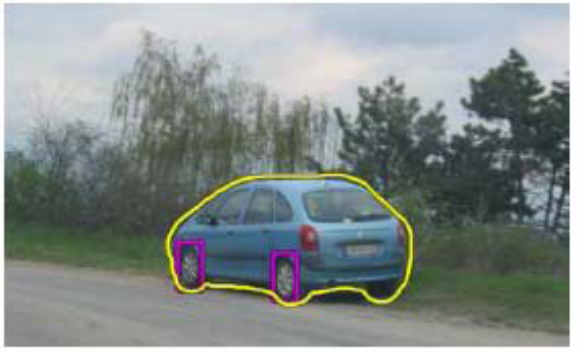
- Choose a *parametric model* to represent a set of features



simple model: lines

simple model: circles

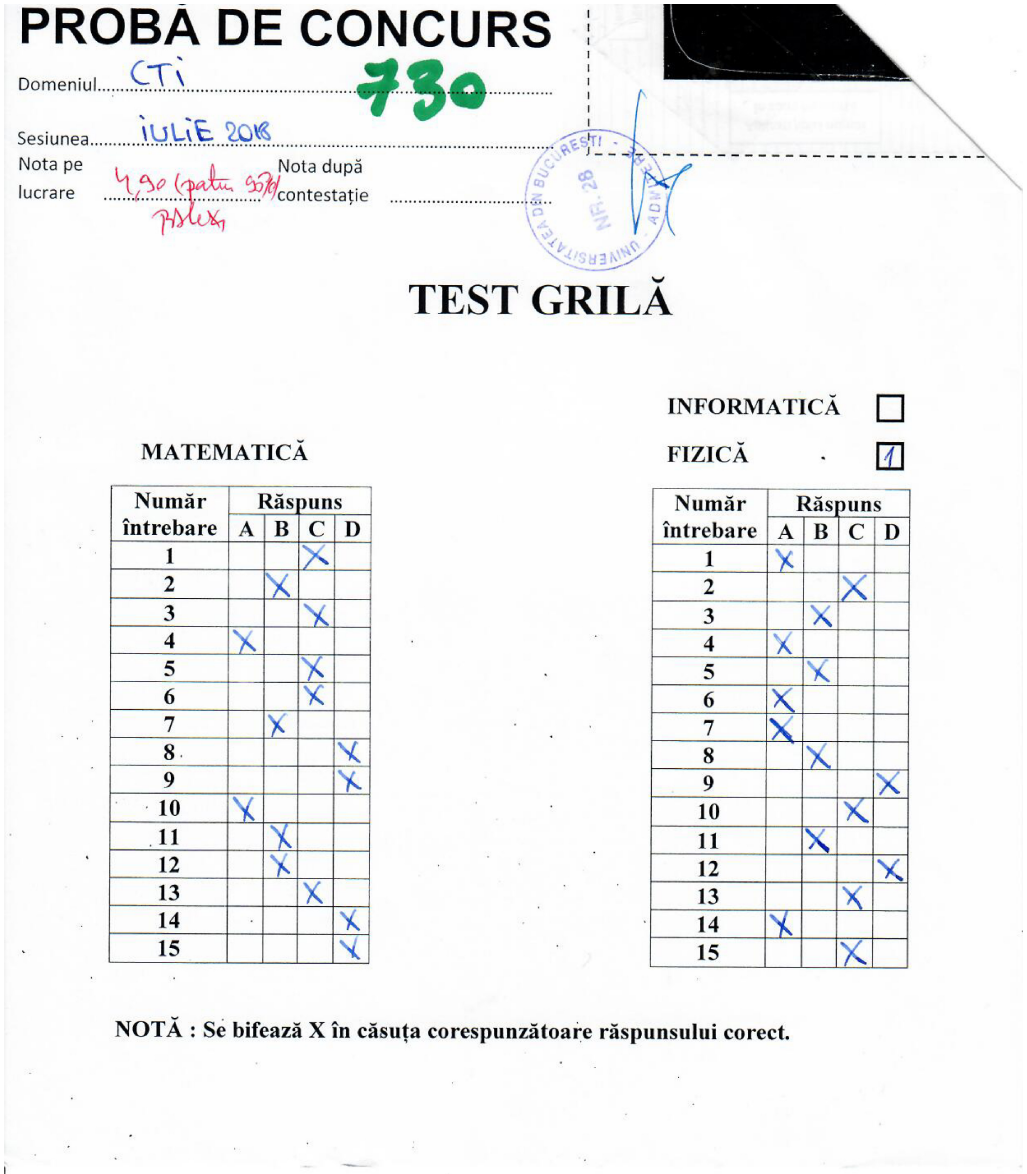




complicated model: car

Slide credit : K. Grauman

Lab 2: Line detection - application



PROBĂ DE CONCURS

Domeniul CTI 730

Sesiunea IULIE 2018

Nota pe lucrare 4,90 (patru 90) Nota după contestație .....

TEST GRILĂ

INFORMATICĂ ☐

FIZICĂ ☑

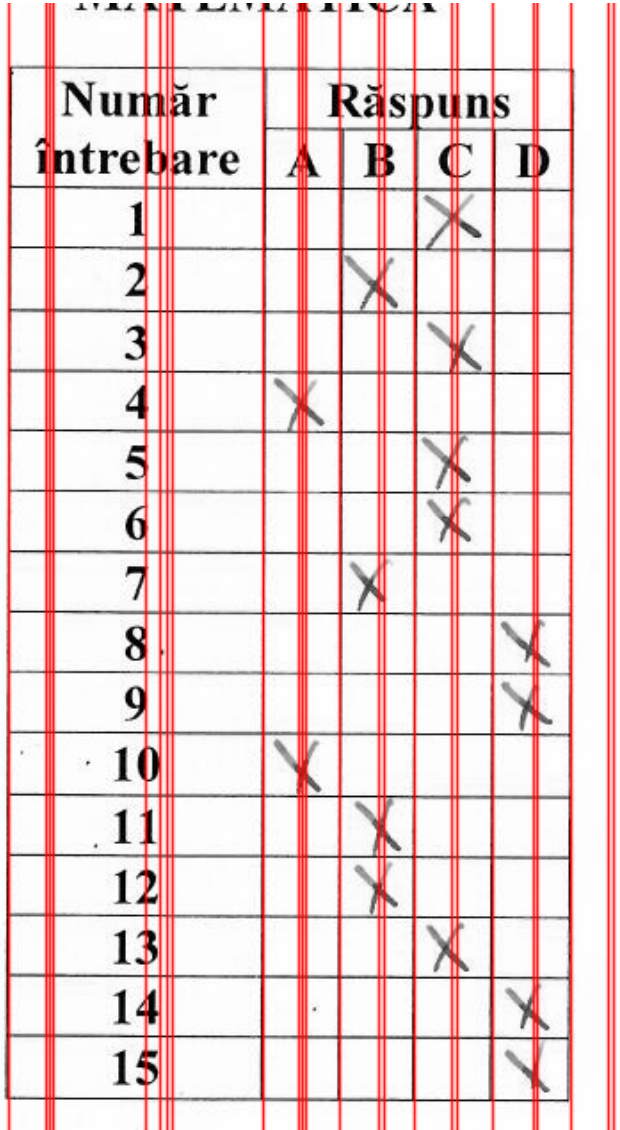
MATEMATICĂ

Număr întrebare	Răspuns A	B	C	D
1			X	
2		X		
3			X	
4	X			
5			X	
6			X	
7		X		
8				X
9				X
10	X			
11		X		
12		X		
13			X	
14				X
15				X

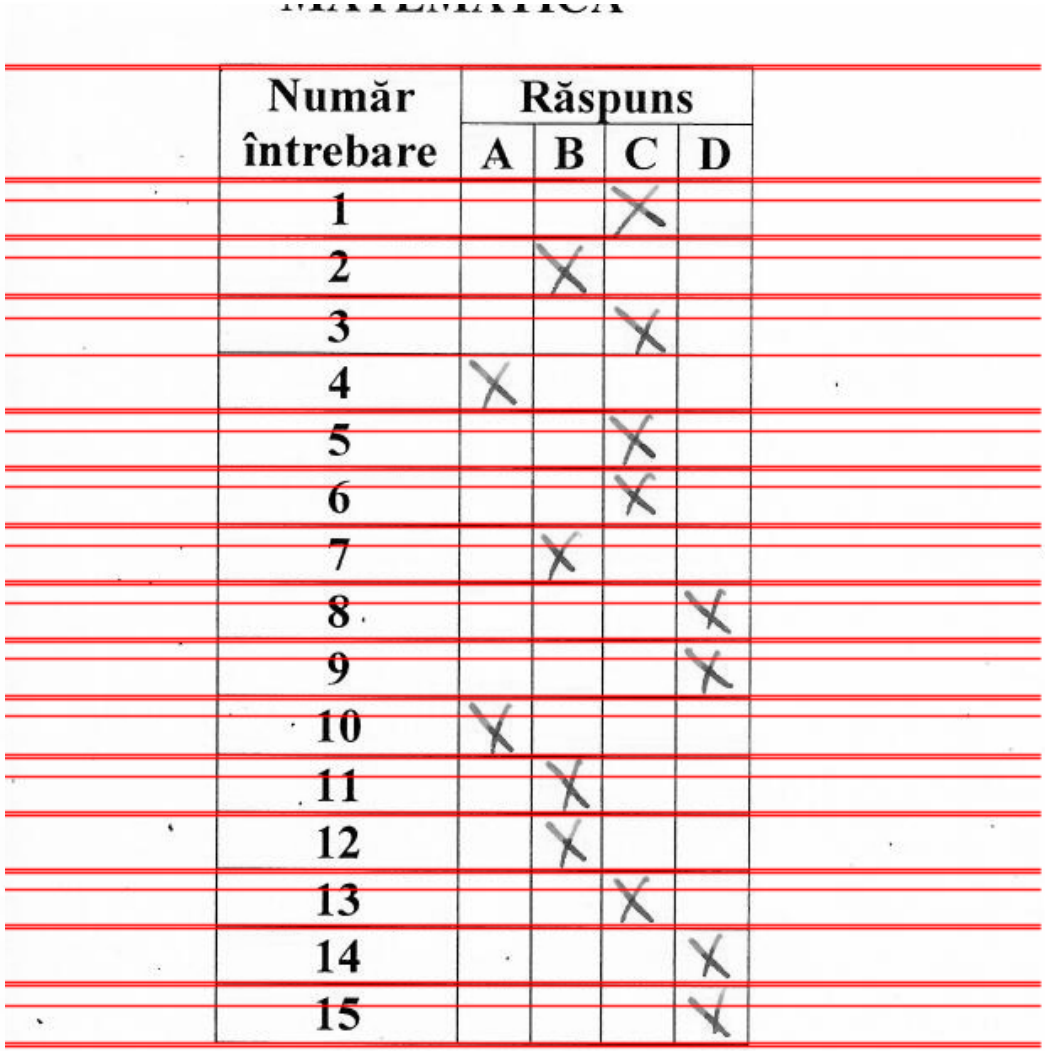
Număr întrebare	Răspuns A	B	C	D
1	X			
2			X	
3		X		
4	X			
5		X		
6	X			
7	X			
8		X		
9				X
10			X	
11		X		
12				X
13			X	
14	X			
15			X	

NOTĂ : Se bifează X în căsuța corespunzătoare răspunsului corect.

Lab 2: Line detection - application

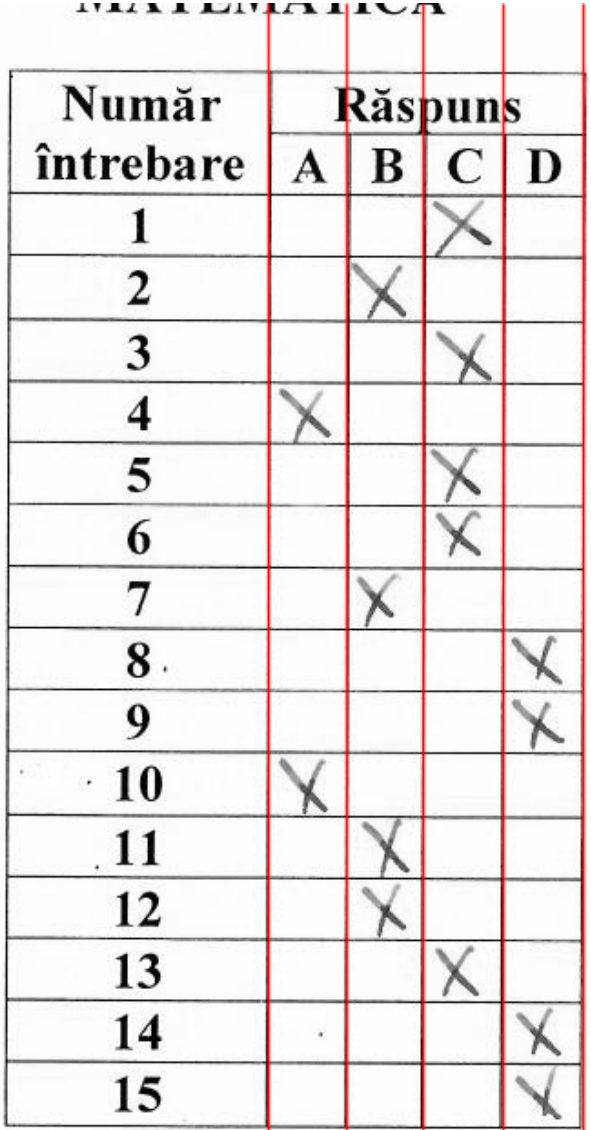


Număr întrebare	Răspuns A	B	C	D
1			X	
2		X		
3			X	
4	X			
5			X	
6			X	
7		X		
8				X
9				X
10	X			
11		X		
12		X		
13			X	
14				X
15				X

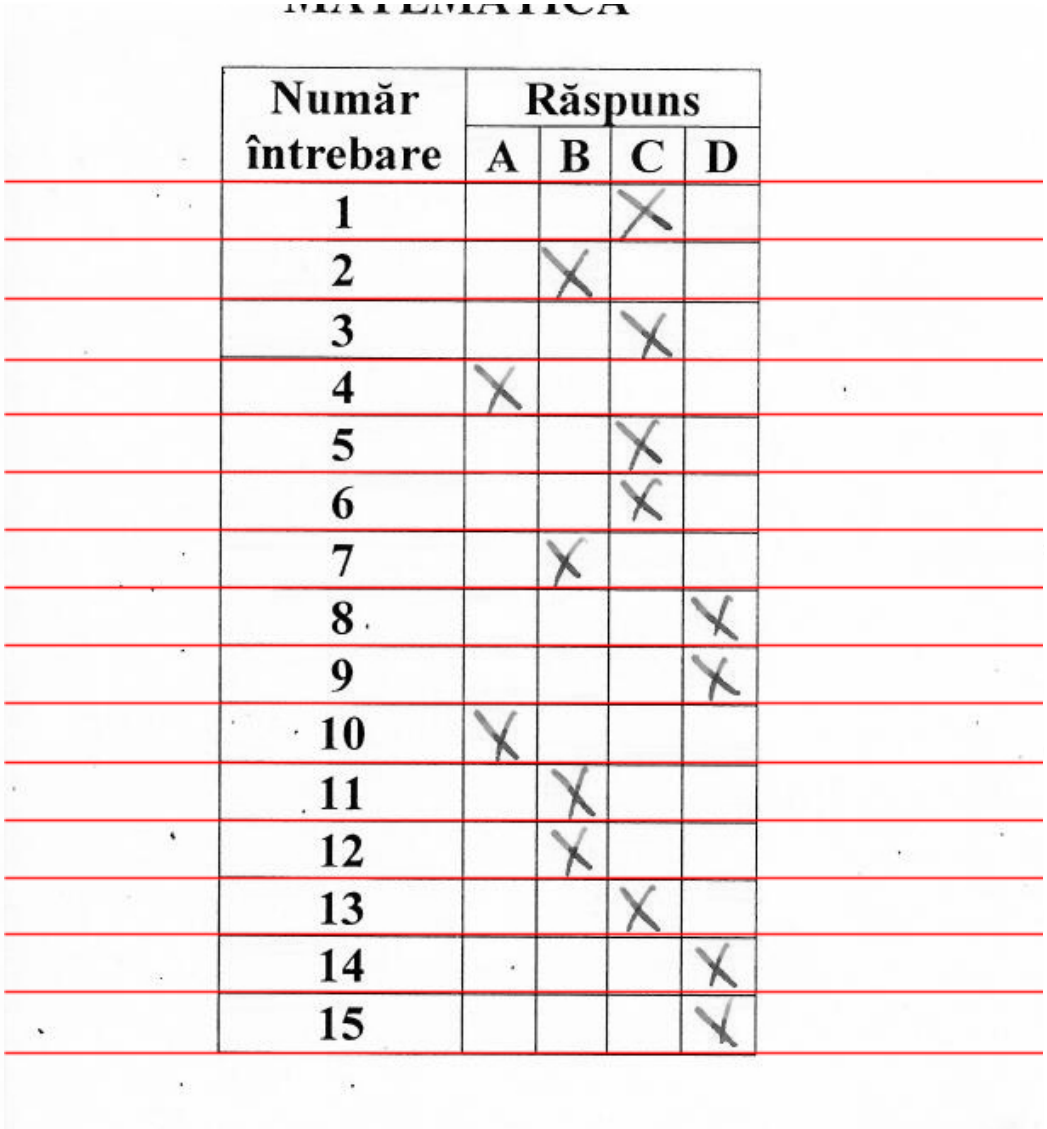


Număr întrebare	Răspuns A	B	C	D
1			X	
2		X		
3			X	
4	X			
5			X	
6			X	
7		X		
8				X
9				X
10	X			
11		X		
12		X		
13			X	
14				X
15				X

Lab 2: Line detection - application



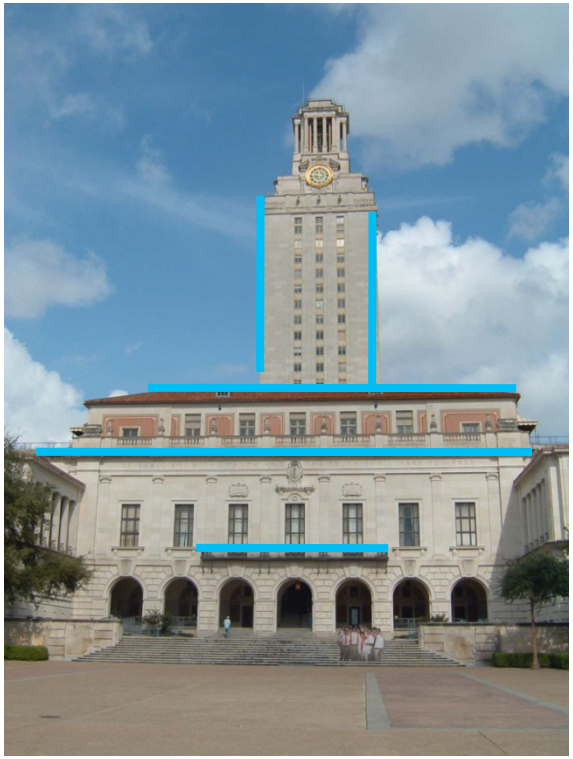
Număr întrebare	Răspuns			
	A	B	C	D
1			X	
2		X		
3			X	
4	X			
5			X	
6			X	
7		X		
8				X
9				X
10	X			
11		X		
12		X		
13			X	
14				X
15				X

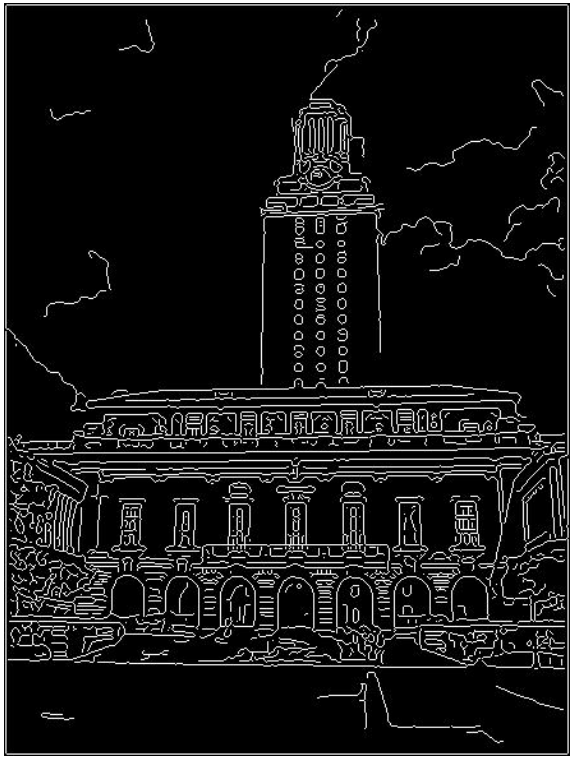


Număr întrebare	Răspuns			
	A	B	C	D
1			X	
2		X		
3			X	
4	X			
5			X	
6			X	
7		X		
8				X
9				X
10	X			
11		X		
12		X		
13			X	
14				X
15				X

Fitting: challenges

Case study: Line detection





- **Noise** in the measured feature locations
- **Extraneous data:** clutter (outliers), multiple lines
- **Missing data:** occlusions

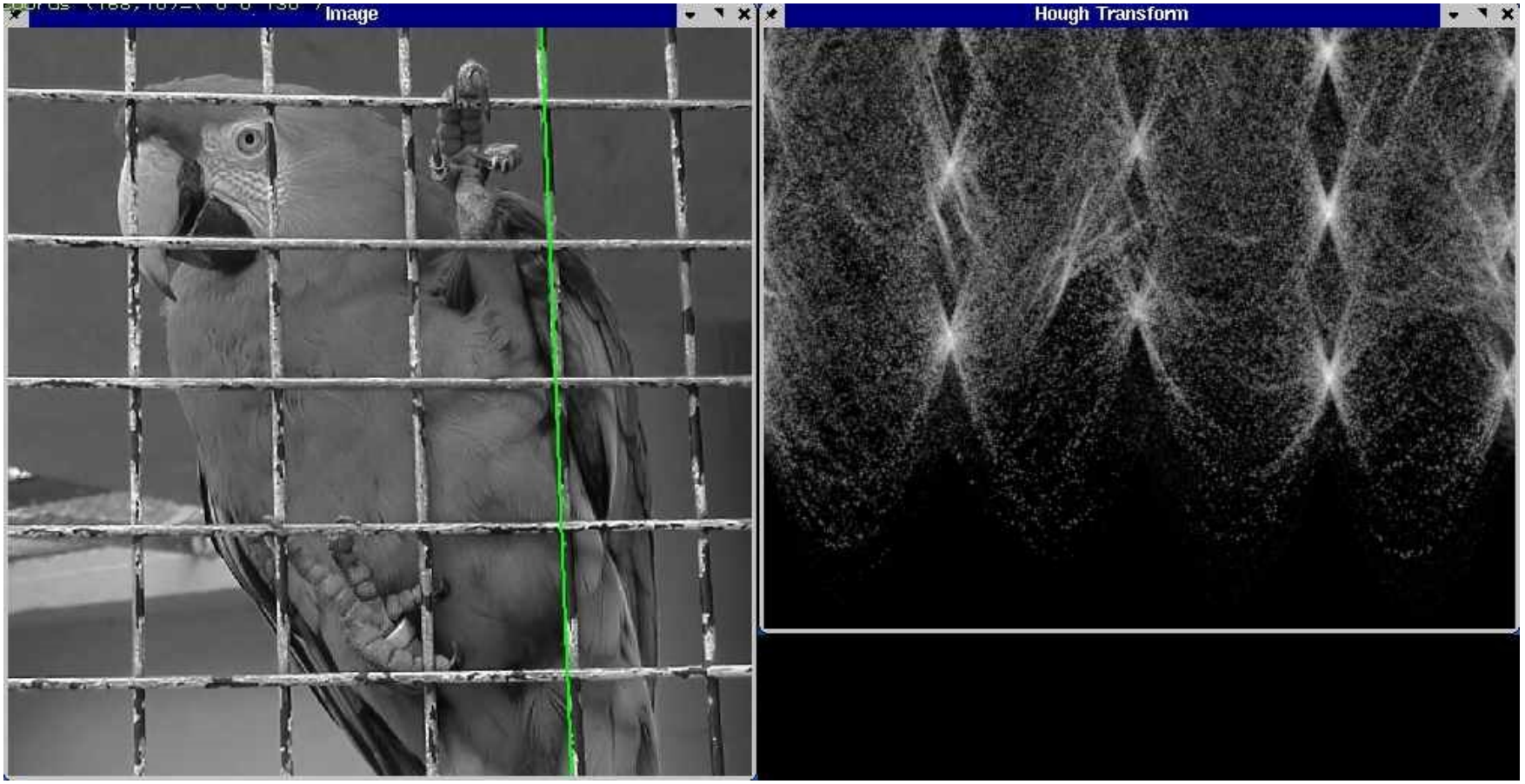
Slide credit : S. Lazebnik

Fitting: challenges

- If we know which points belong to the line, how do we find the "optimal" line parameters?
  - Least squares
- What if there are outliers?
  - Robust fitting, RANSAC
- What if there are many lines?
  - Voting methods: RANSAC, Hough transform

Slide credit : S. Lazebnik

Fitting: The Hough transform



Slide credit : S. Lazebnik

Voting schemes

- Let each feature vote for all the models that are compatible with it
  - Hopefully the noise features will not vote consistently for any single model
  - Missing data doesn't matter as long as there are enough features remaining to agree on a good model

Slide credit : S. Lazebnik

Hough transform

- An early type of voting scheme
  - Discretize *parameter space* into bins
  - For each feature point in the image, put a vote in every bin in the parameter space that could have generated this point
  - Find bins that have the most votes

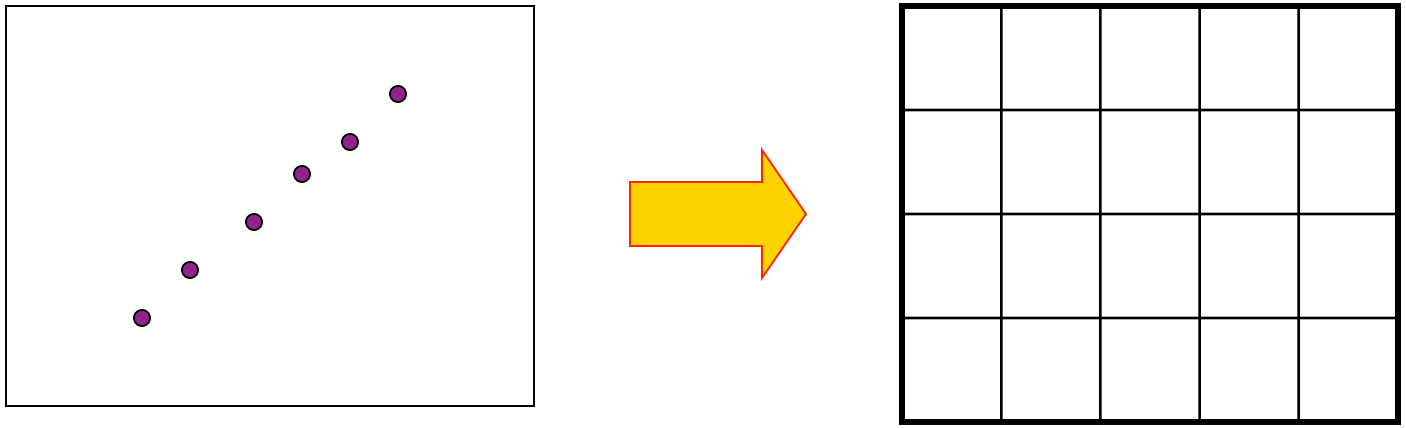


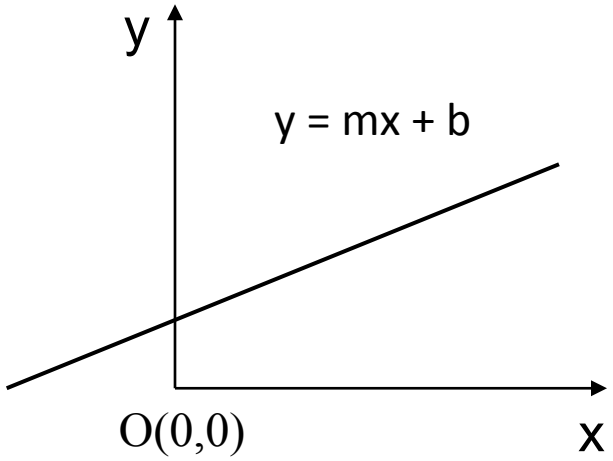
Image space

Hough parameter space

P.V.C. Hough, *Machine Analysis of Bubble Chamber Pictures*,
Proc. Int. Conf. High Energy Accelerators and Instrumentation, 1959

Slide credit : S. Lazebnik

Parameterization of a line



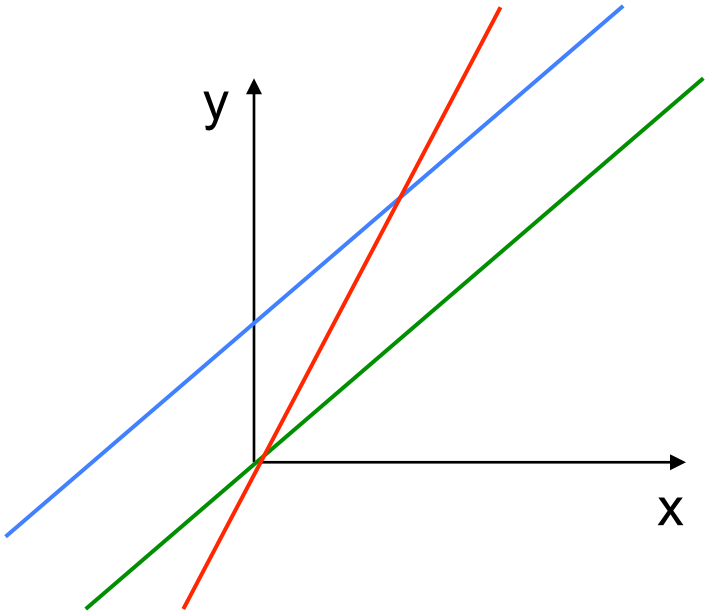
Parameterization of a line in image plane: $y = m * x + b$,

m – slope (defines the angle between the line and Ox axis),

b – intercept (bias to origin O(0,0))

(m,b) are line parameters

Parameterization of a line - examples



d_1: y = x, so m=1, b = 0

d_2: y = x + 2, so m=1, b = 2

d_3: y = 2x, so m=2, b = 0

Parameterization of a line in image plane: $y = m * x + b$,

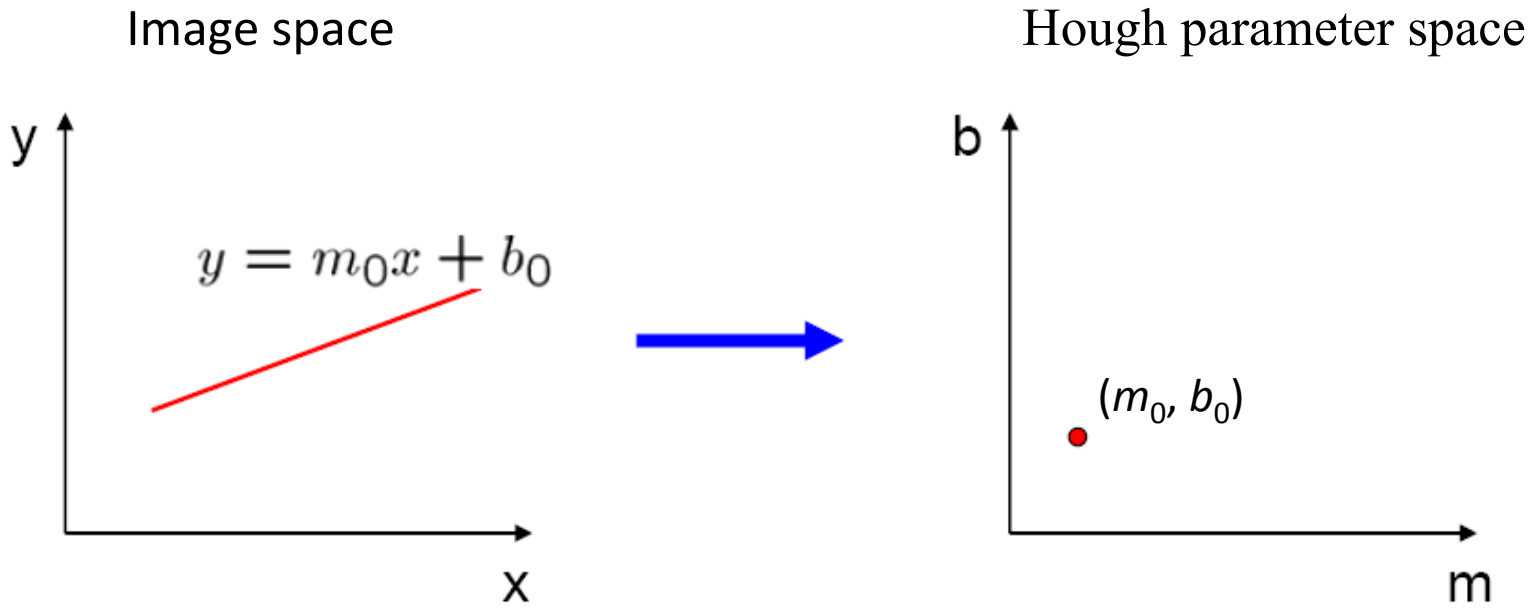
m – slope (defines the angle between the line and Ox axis),

b – intercept (bias to origin O(0,0))

(m,b) are line parameters

Parameter space representation

- A **line** in the image corresponds to a **point** in Hough space



Slide credit : S. Seitz

Hough parameter space

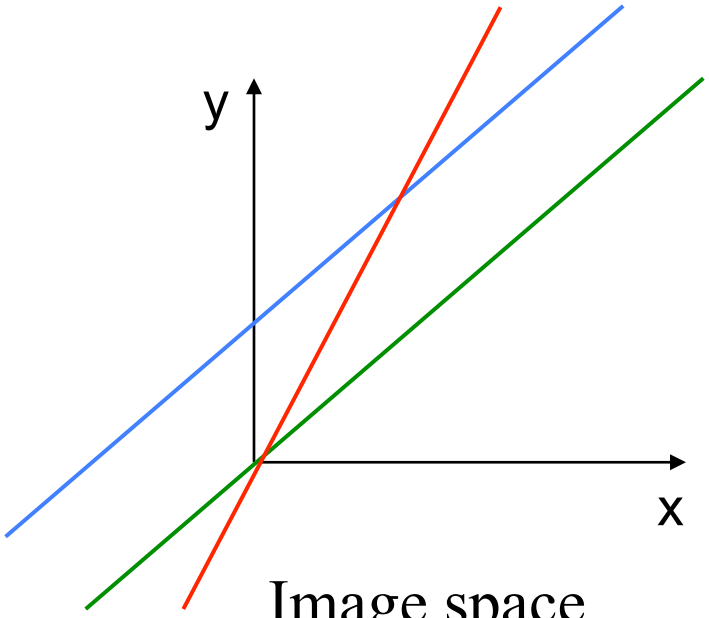
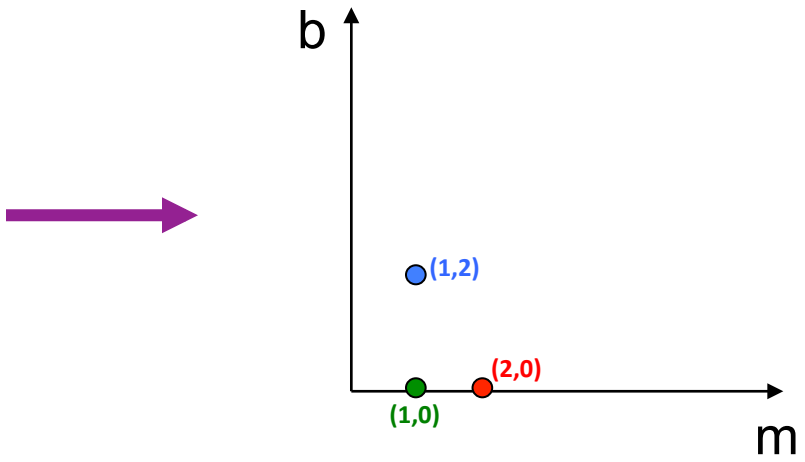


Image space



Hough space

d_1: y = x, so m=1, b = 0

d_2: y = x + 2, so m=1, b = 2

d_3: y = 2x, so m=2, b = 0

d_1: m=1, b = 0 -> (1,0)

d_2: m=1, b = 2 -> (1,2)

d_3: m=2, b = 0 -> (2,0)

A line in the image space correspond to a point in the Hough space.

Correspondence between image and Hough spaces

A line in the image space correspond to a point in the Hough space.

What does a **point** (x_0, y_0) in the image space map to in the Hough space?

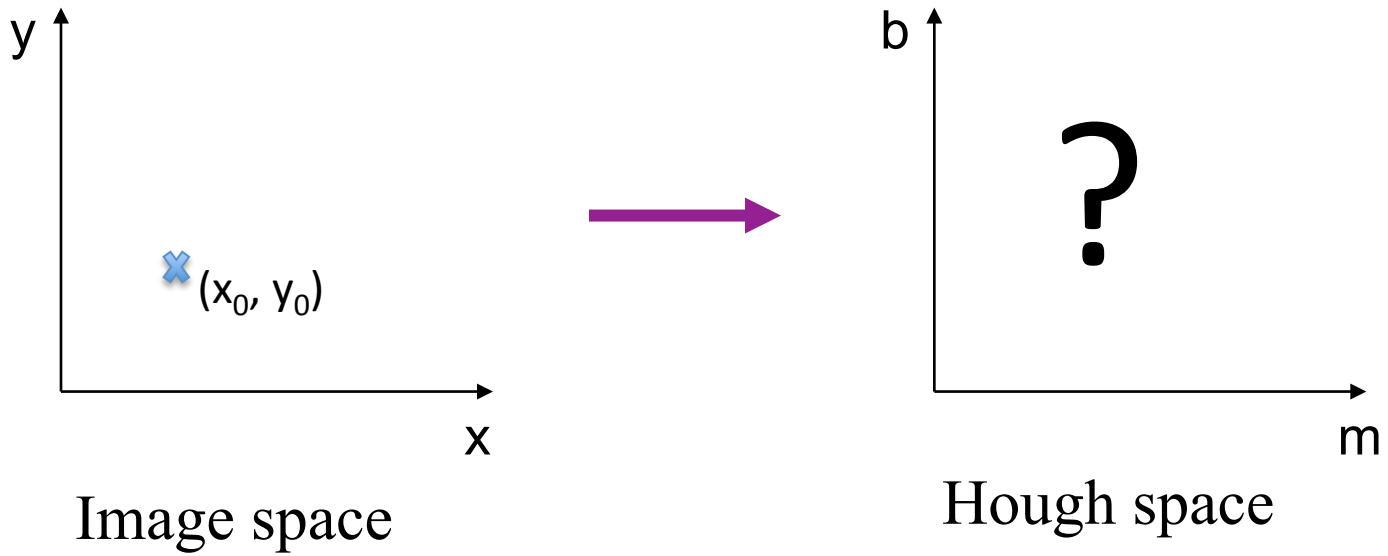


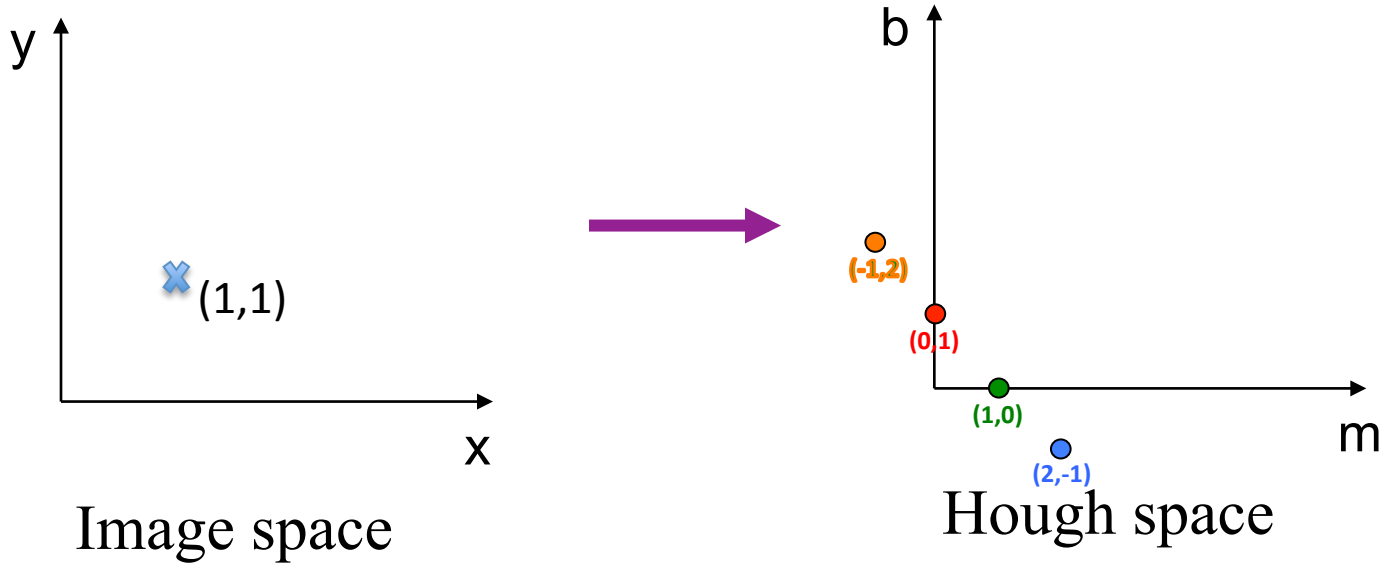
Image space

Hough space

Correspondence between image and Hough spaces

A line in the image space correspond to a point in the Hough space.

What does a **point** (x_0, y_0) in the image space map to in the Hough space?



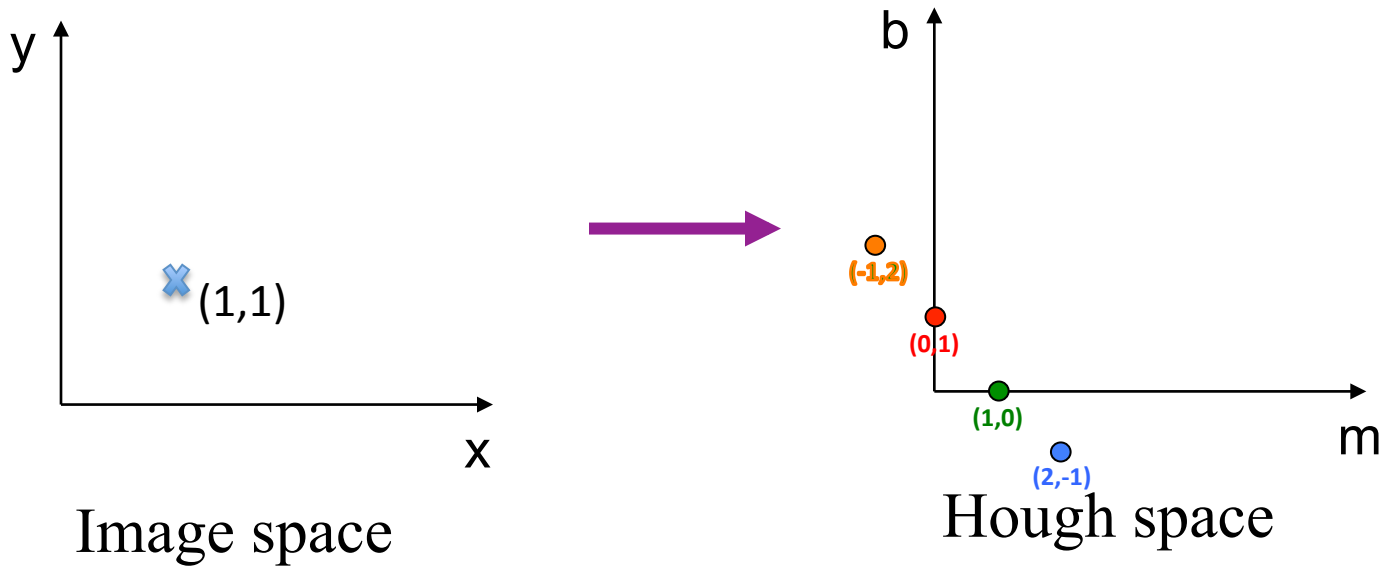
There is an infinity of lines of the form y = mx+b that contain point (1,1):

y = x (m=1, b=0); y = 2x-1 (m=2,b=-1), y = 1 (m=0, b = -1), y = -x + 2 (m=-1, b =2),

Correspondence between image and Hough spaces

A line in the image space correspond to a point in the Hough space.

A **point** (x_0, y_0) in the image space maps to a line in the Hough space



There is an infinity of lines of the form y = mx+b that contain point (1,1):

y = x (m=1, b=0); y = 2x-1 (m=2,b=-1), y = 1 (m=0, b = 1), y = -x + 2 (m=-1, b =2),

All 4 points (1,0), (2,-1), (0,-1), (-1,2) are on the line b = -m+1

Correspondence between image and Hough spaces

A line in the image space correspond to a point in the Hough space.

A **point** (x_0, y_0) in the image space maps to a line in the Hough space

Parameterization of a line in image plane: $y = m * x + b$,

All lines that contain point (x_0,y_0) can have any slope m and any intercept b, the only condition that should be satisfied is

$y_0 = m * x_0 + b$.

$y_0 = m * x_0 + b \Leftrightarrow b = -x_0 * m + y_0$ (parameterization of the line in the Hough space with slope $-x_0$ and intercept y_0)

$(x_0,y_0) = (1,1) \Rightarrow b = -m+1$

Parameter space representation

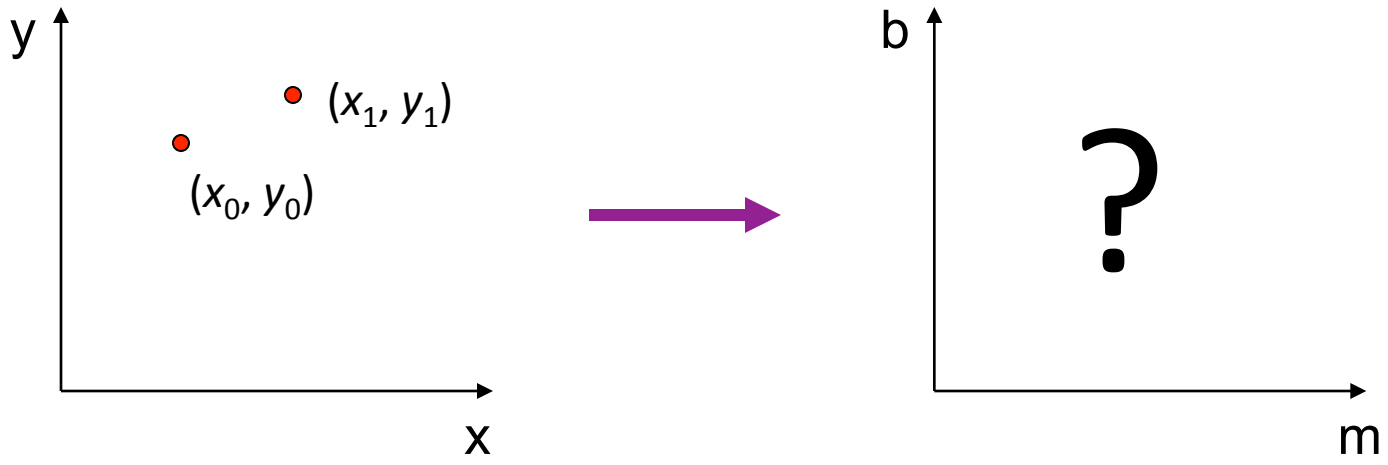


Image space

Hough space

Where is the line that contains both (x_0, y_0) and (x_1, y_1)?

Parameter space representation

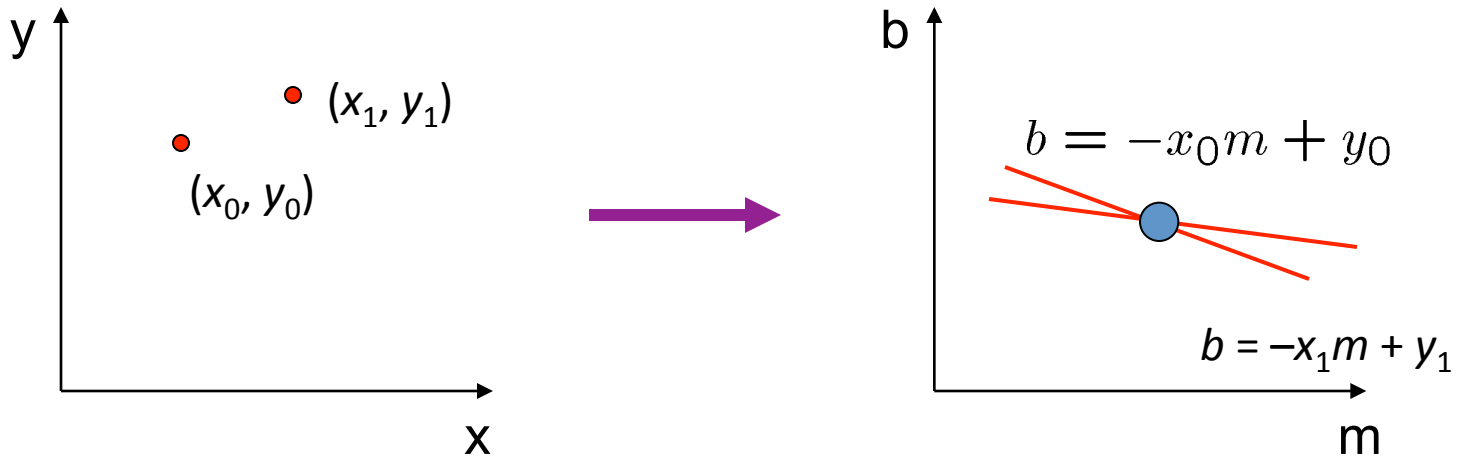


Image space

Hough space

Where is the line that contains both (x_0, y_0) and (x_1, y_1)?

Point (x_0,y_0) maps to the line in the Hough space: $b = -x_0m + y_0$

Point (x_1,y_1) maps to the line in the Hough space: $b = -x_1m + y_1$

The line that contains both (x_0, y_0) and (x_1, y_1) maps to a point in the Hough space. This point is the intersection of lines $b = -x_0m + y_0$ și $b = -x_1m + y_1$

Finding lines in an image with Hough

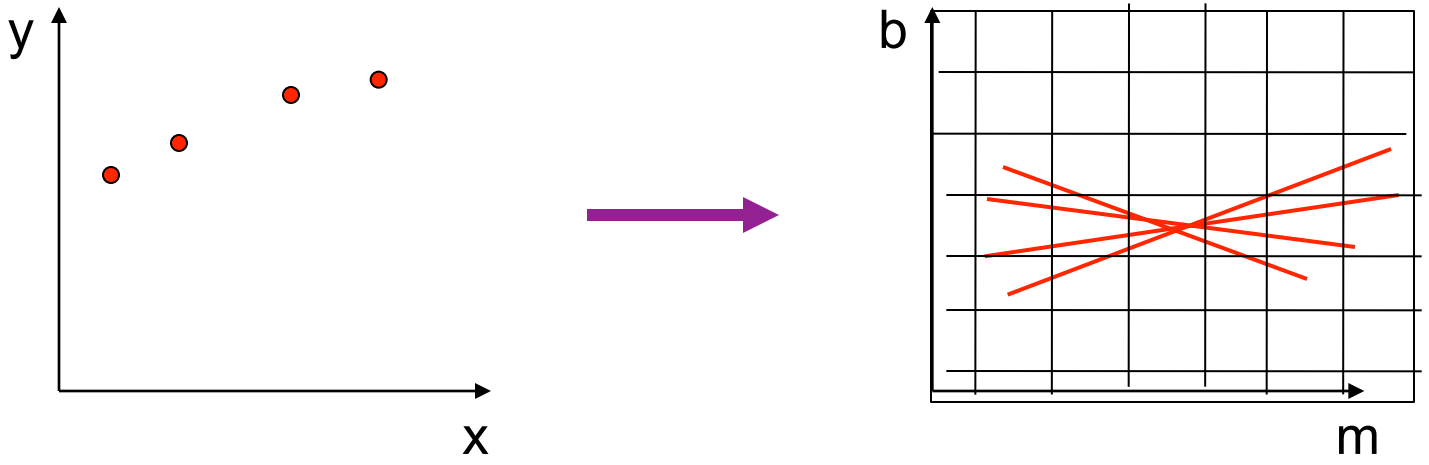


Image space

Hough space

How we can use the previous observations to find parameters (m,b) that define the most probable line in the image space?

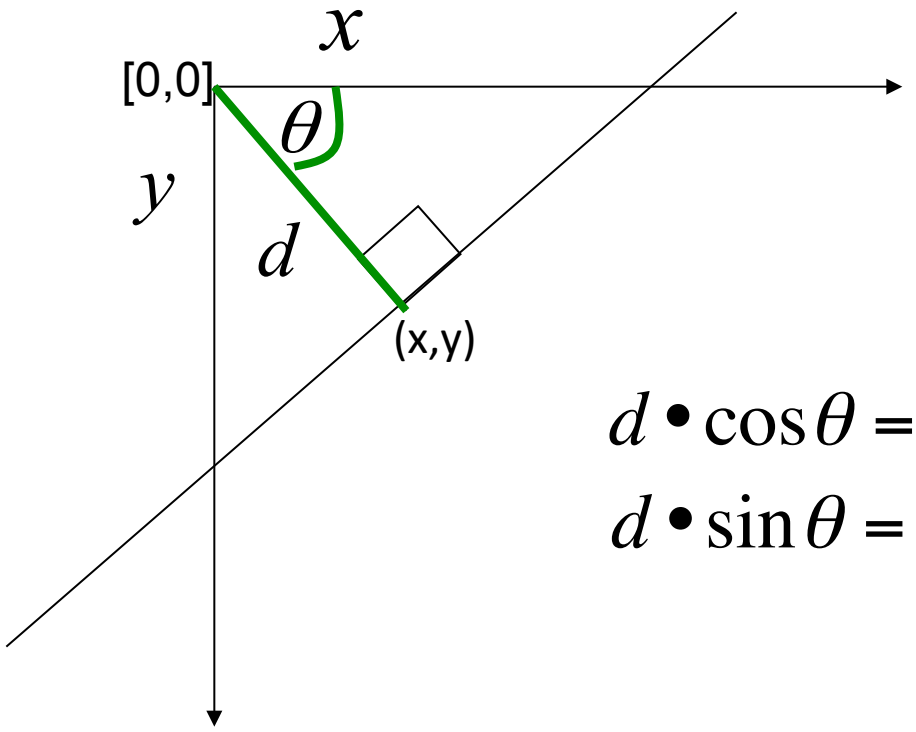
- each edgel point (part of an edge) from image space will vote (for an infinity of lines) in the Hough space (here a line is a point)
- discretize the Hough space in discrete intervals/cells; the cell with largest number of votes determines the parameters of the line

Polar representation

Problems with the (m,b) space:

- Unbounded parameter domains
- Vertical lines require infinite m

Alternative: polar representation



d: distance from origin to line (length of the perpendicular)

θ: angle between the perpendicular and Ox axis

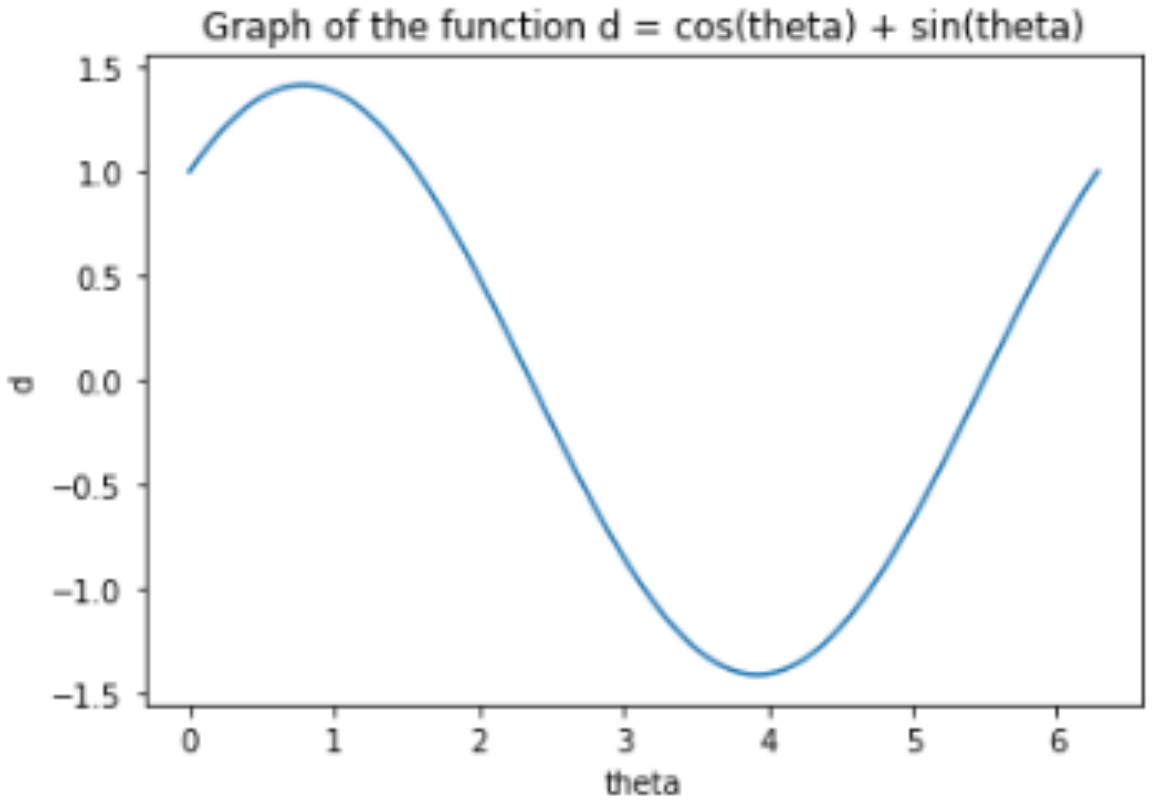
$$d \bullet \cos\theta = x \qquad d \bullet (\cos\theta)^2 = x \bullet \cos\theta$$

$$d \bullet \sin\theta = y \qquad d \bullet (\sin\theta)^2 = y \bullet \sin\theta$$

$$d = x \bullet \cos\theta + y \bullet \sin\theta$$

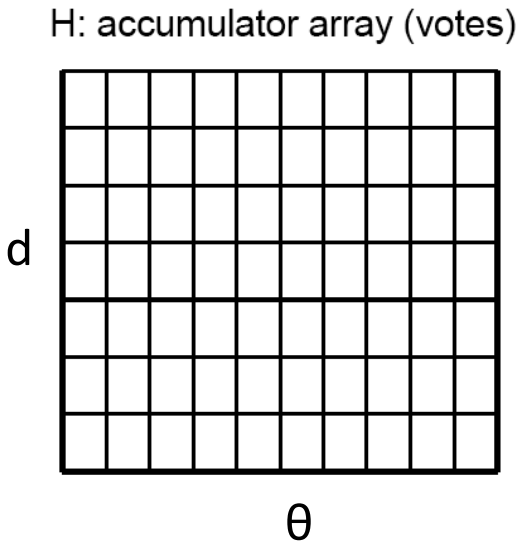
Point in the image space → sinusoid curve in the Hough space

```python
import numpy as np
import matplotlib.pyplot as plt
x, y = 1,1
theta = np.linspace(0,2*np.pi,100)
d = x * np.cos(theta) + y * np.sin(theta)
plt.figure,
plt.plot(theta, d)
plt.xlabel('theta')
plt.ylabel('d')
plt.title('Graph of the function d = cos(theta) + sin(theta)')
plt.show()
```



Hough transform-algorithm outline

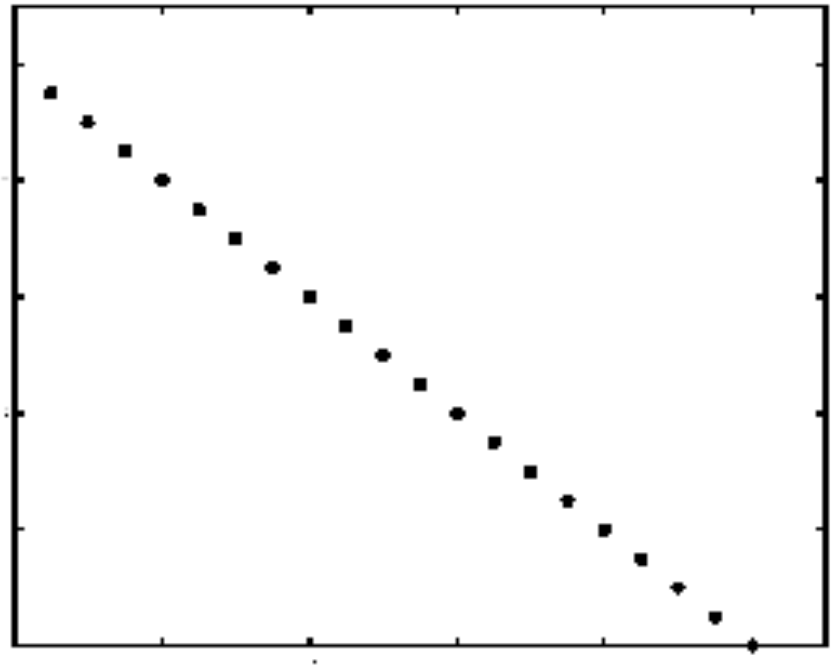
- Initialize accumulator H to all zeros
- For each feature point (x,y) in the image
  For θ = 0 to 180
    $d = x \cos \theta + y \sin \theta$
    $H(\theta, d) = H(\theta, d) + 1$
  end
end
- Find the value(s) of (θ, d) where H(θ, d) is a local maximum
  - The detected line in the image is given by
    $d = x \cos \theta + y \sin \theta$



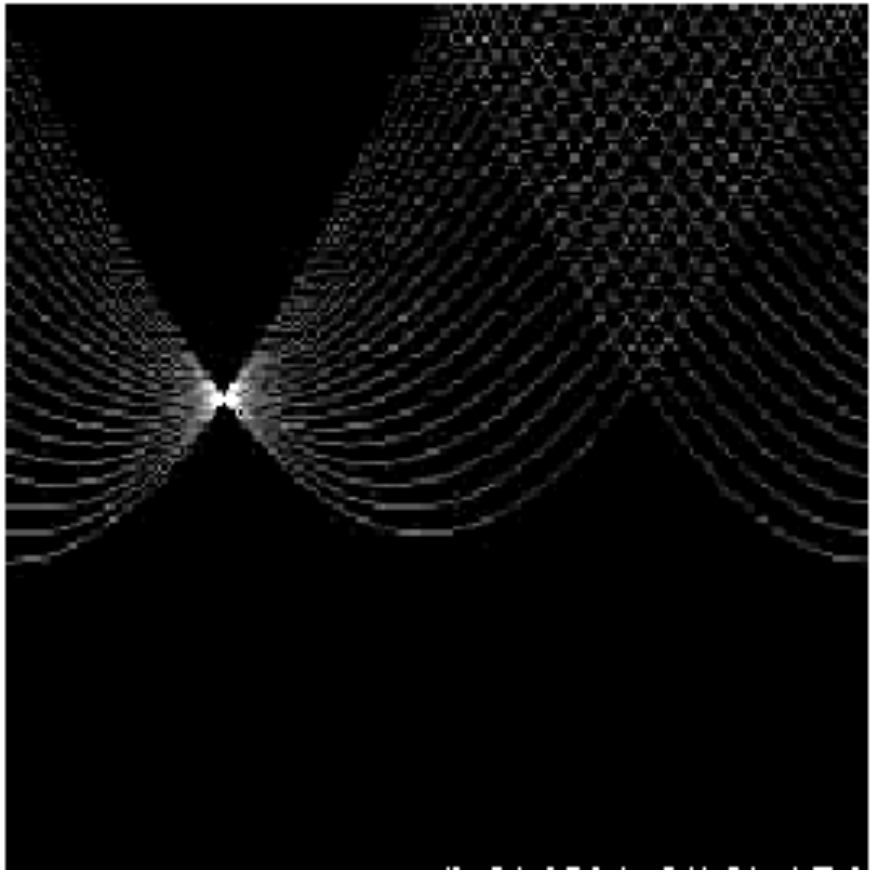
Hough transform - demo

https://www.aber.ac.uk/~dcswww/Dept/Teaching/CourseNotes/current/CS34110/hough.html

Basic illustration



features



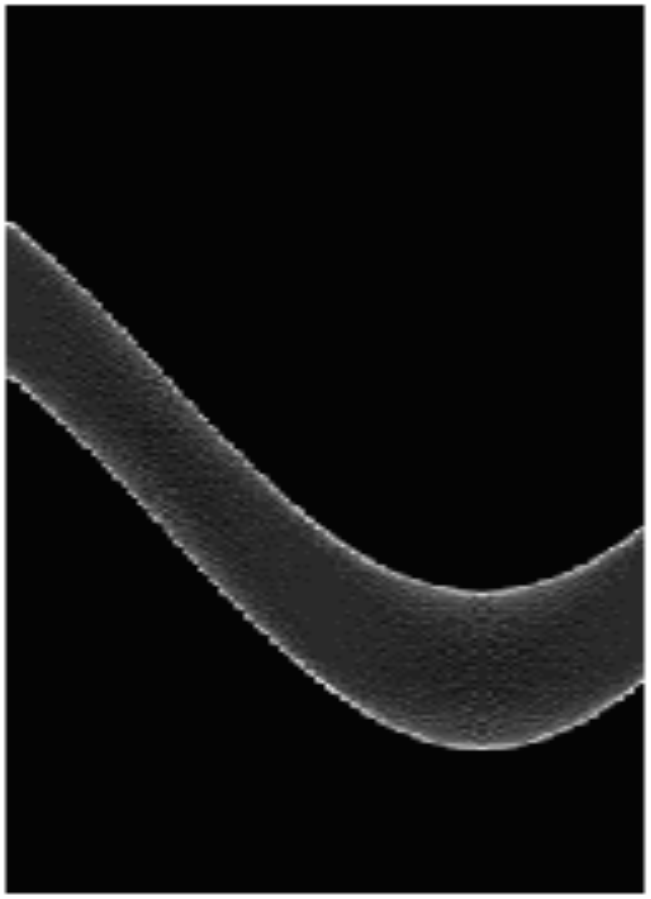
votes

Hough transform demo

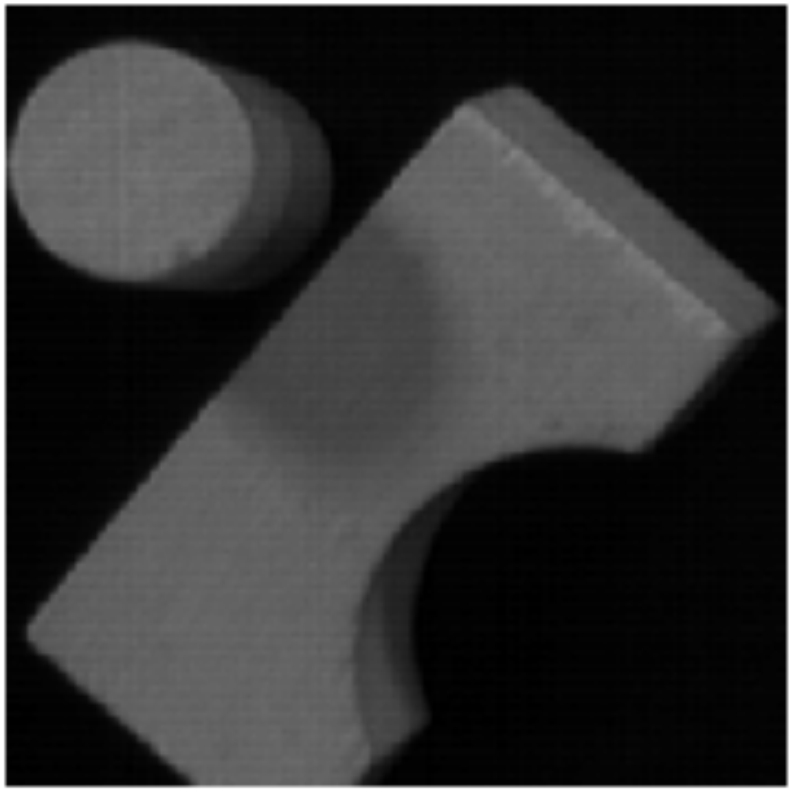
Other shapes

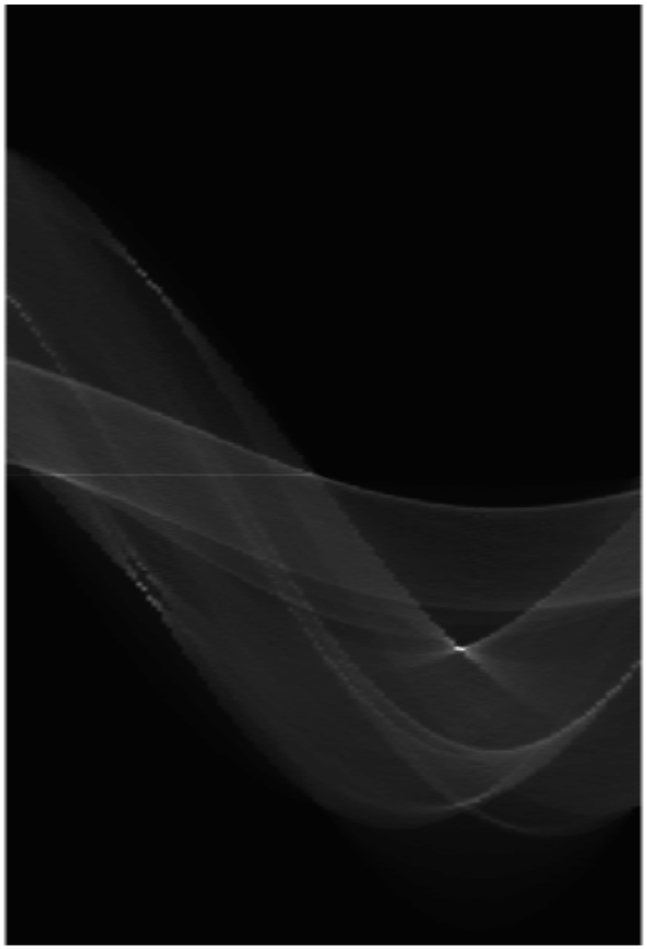
Square

Circle

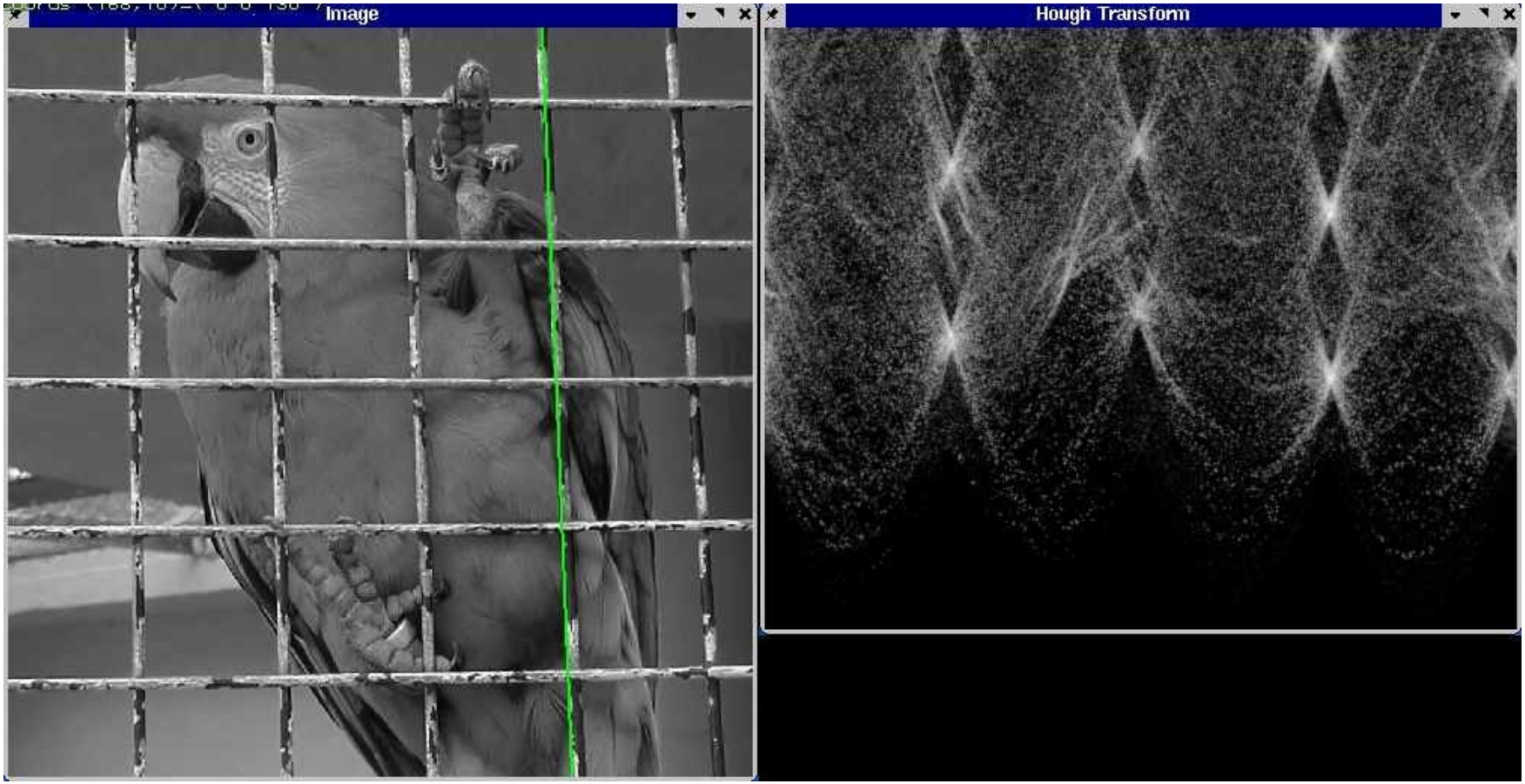


Several lines

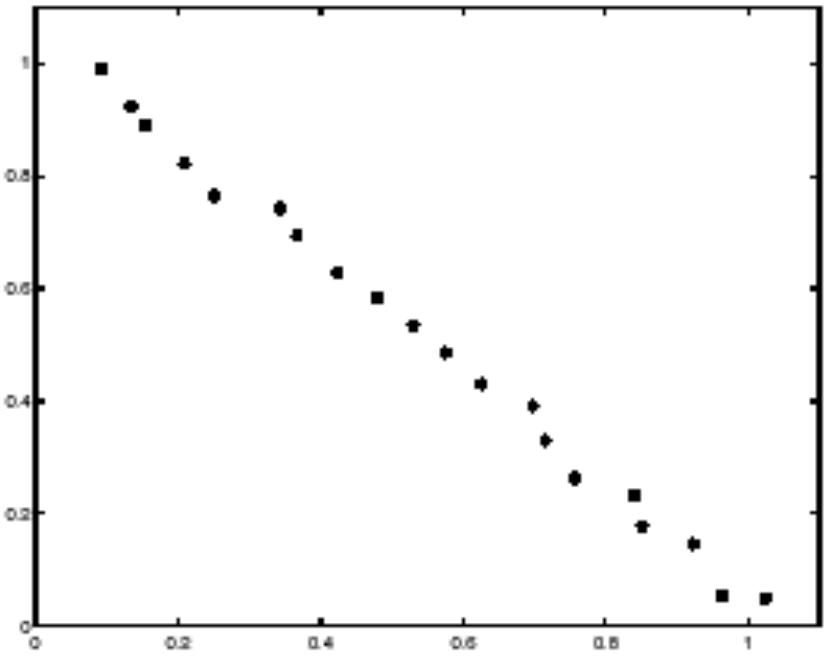




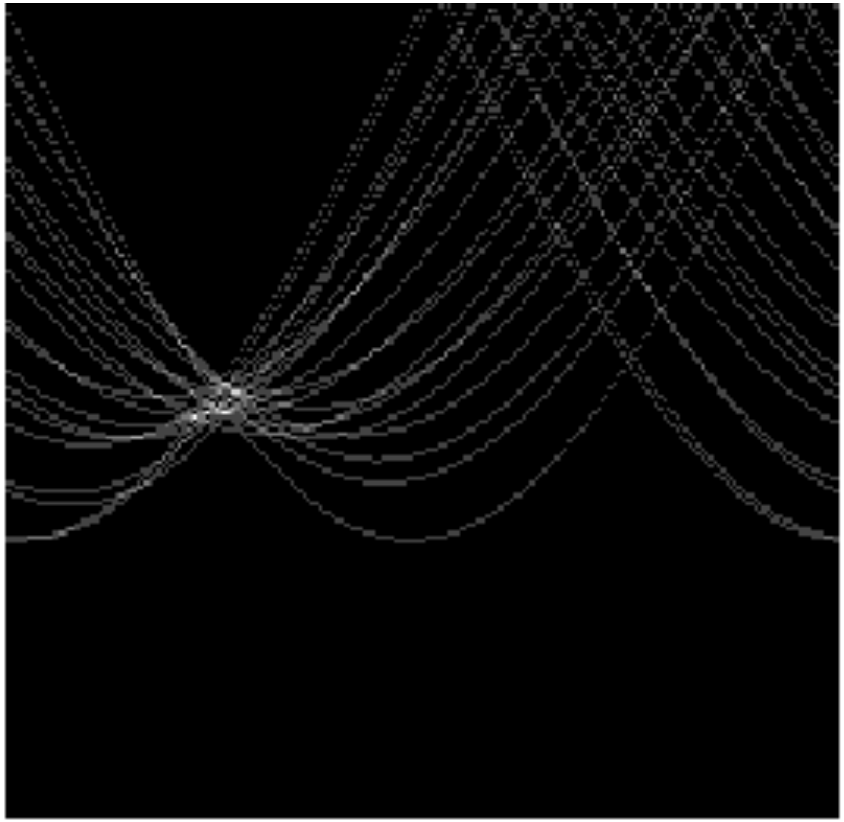
A more complicated image



Effect of noise



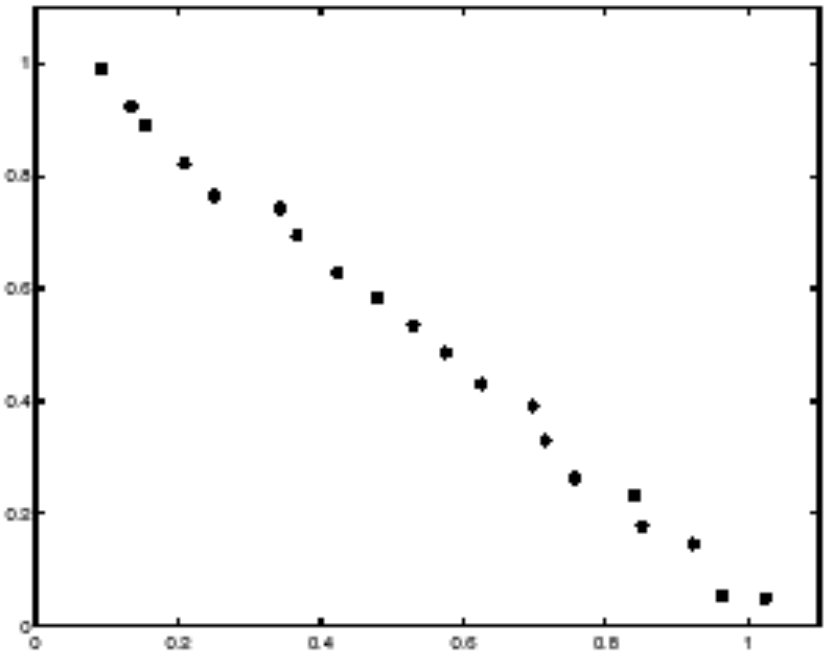
features



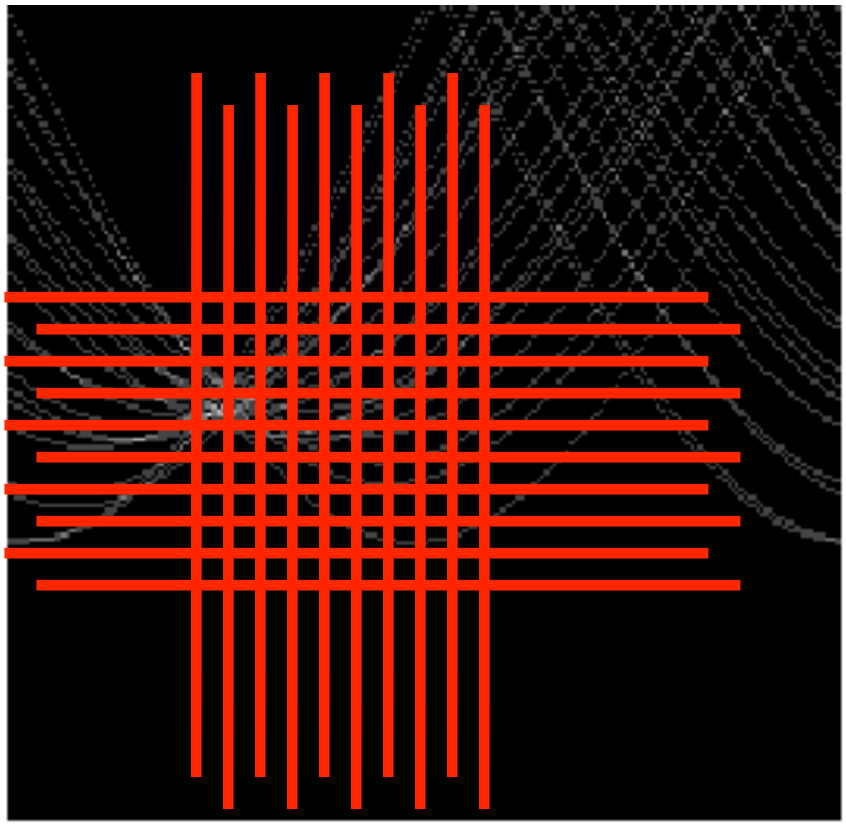
votes

- Peak gets fuzzy and hard to locate

Effect of noise



features

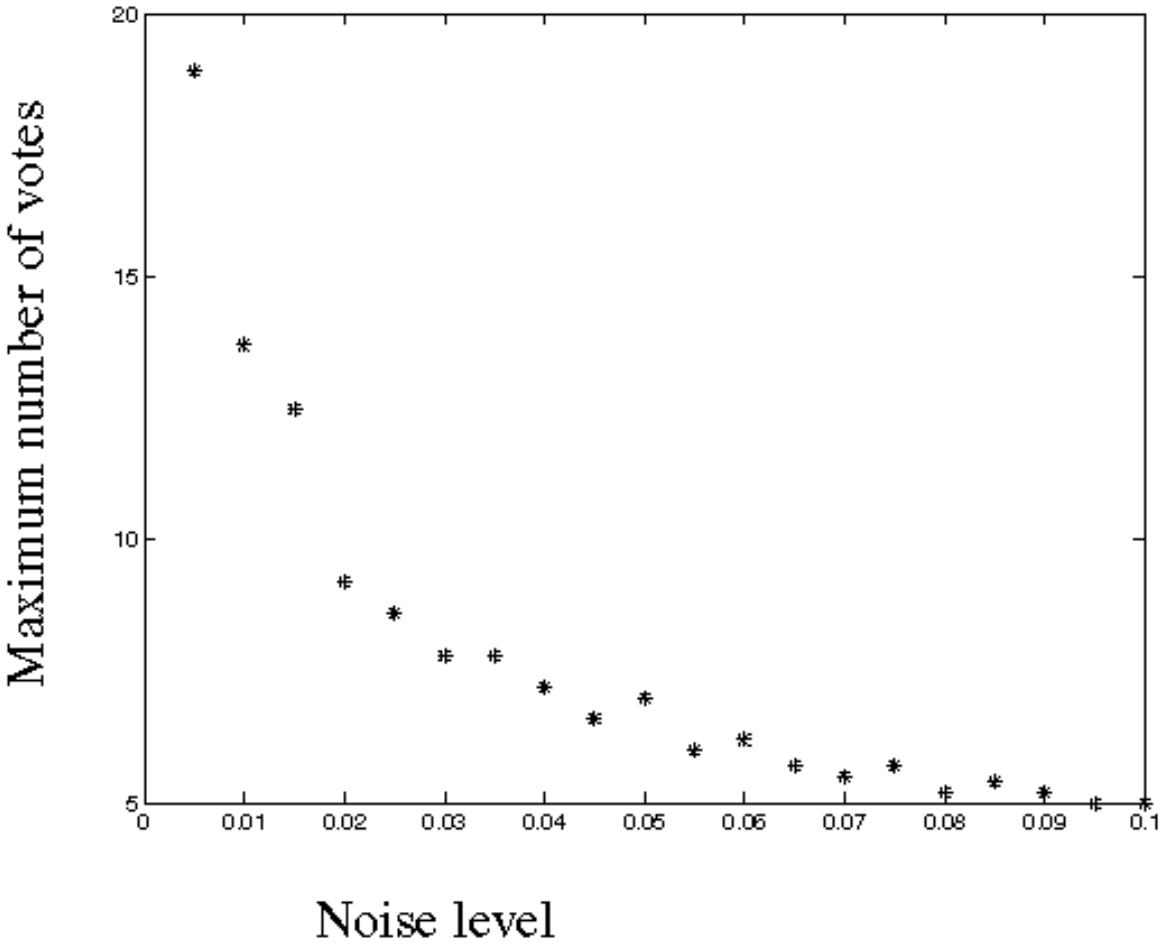


votes

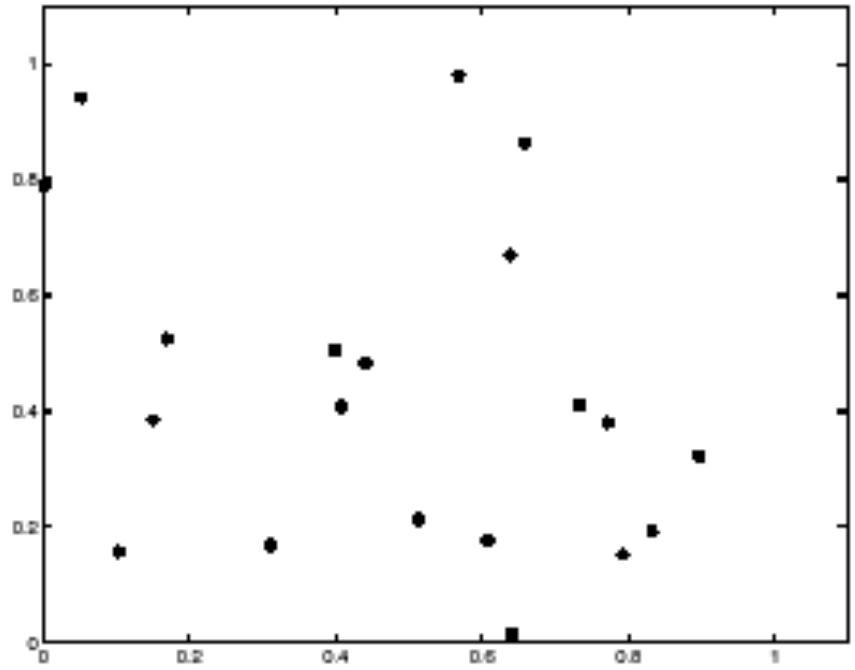
- Peak gets fuzzy and hard to locate

Effect of noise

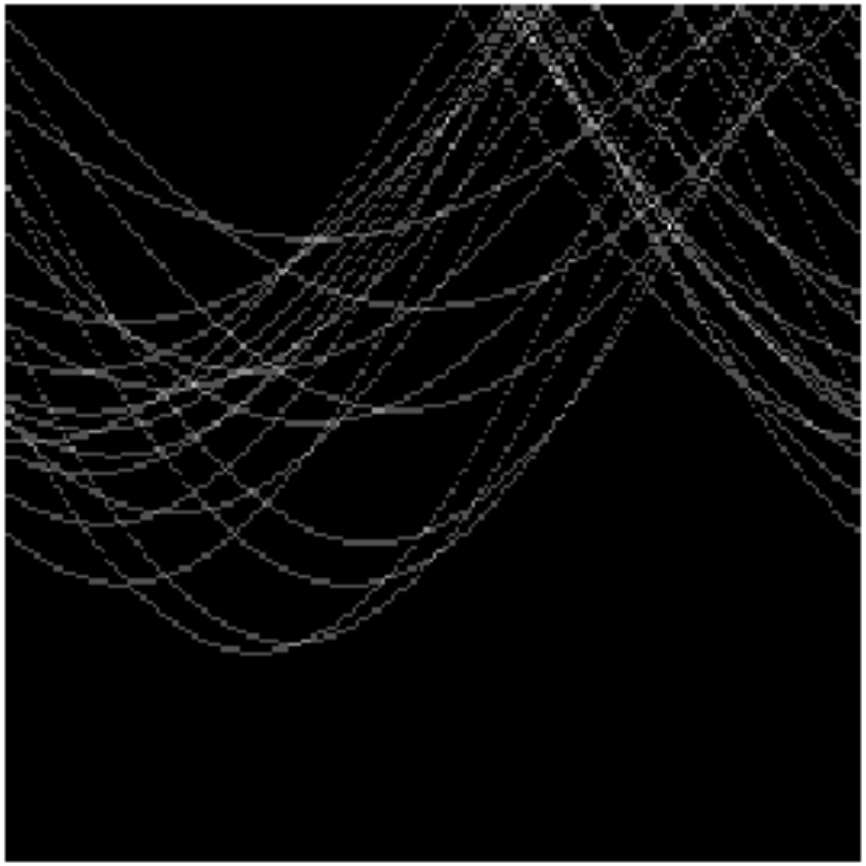
- Number of votes for a line of 20 points with increasing noise:



Random points



features

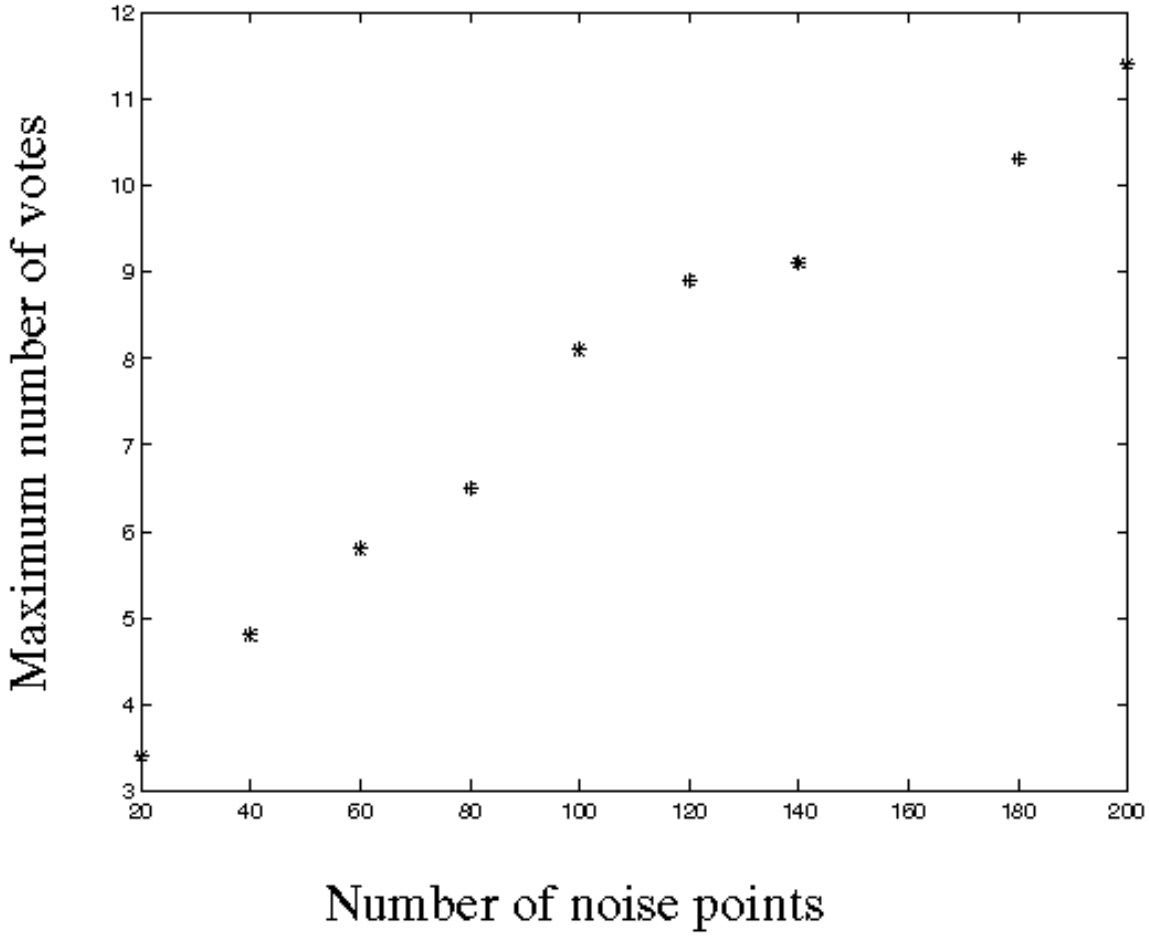


votes

- Uniform noise can lead to spurious peaks in the array

Random points

- As the level of uniform noise increases, the maximum number of votes increases too:

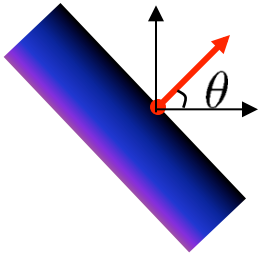


Dealing with noise

- Choose a good grid / discretization
  - **Too coarse:** large votes obtained when too many different lines correspond to a single bucket
  - **Too fine:** miss lines because some points that are not exactly collinear cast votes for different buckets
- Increment neighboring bins (smoothing in accumulator array)
- Try to get rid of irrelevant features
  - E.g., take only edge points with significant gradient magnitude

Incorporating image gradients

- When we detect an edge point, we also know its gradient orientation
- How does this constrain possible lines passing through the point?



$$\nabla f = \left[\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right]$$

$$\theta = \tan^{-1}\left(\frac{\partial f}{\partial y} / \frac{\partial f}{\partial x}\right)$$

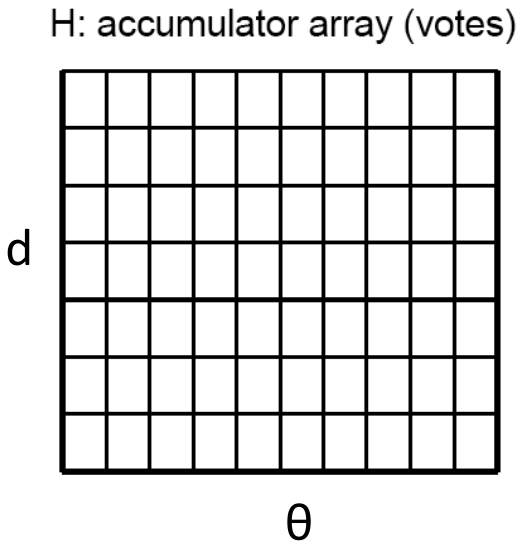
- Modified Hough transform:

- For each edge point (x,y)
  θ = gradient orientation at (x,y)
  d = x cos θ + y sin θ
  H(θ, d) = H(θ, d) + 1
  end

Algorithm outline with image gradients

- Initialize accumulator H to all zeros
- For each feature point (x,y) in the image
  θ = gradient orientation at (x,y)
  d = x cos θ + y sin θ
  H(θ, d) = H(θ, d) + 1

  end



- Find the value(s) of (θ, d) where H(θ, d) is a local maximum
  - The detected line in the image is given by
    d = x cos θ + y sin θ

Algorithm outline with image gradients

- Initialize accumulator H to all zeros
- For each feature point (x,y) in the image

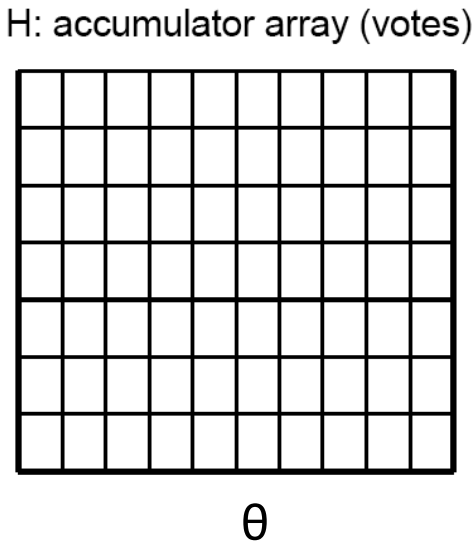
  θ = gradient orientation at (x,y)

  mag = gradient magnitude at (x,y)

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

  $H(\theta, d) = H(\theta, d) + mag$

  end



- Find the value(s) of (θ, d) where $H(\theta, d)$ is a local maximum
  - The detected line in the image is given by

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

Hough transform for circles

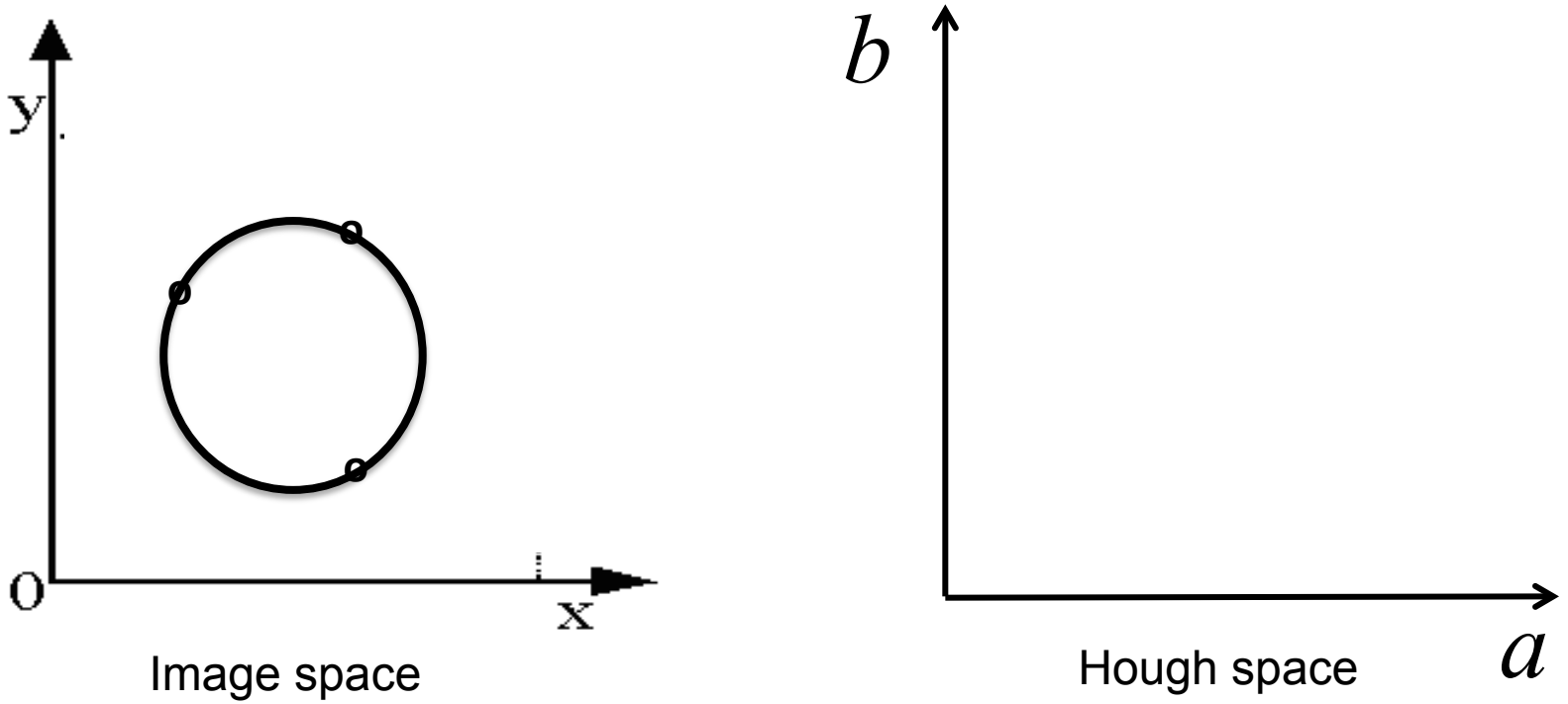
- How many dimensions will the parameter space have?
- Given an unoriented edge point, what are all possible bins that it can vote for?
- What about an *oriented* edge point?

Hough transform for circles

- Circle: center (a,b) and radius r. Equation?

$$(x_i - a)^2 + (y_i - b)^2 = r^2$$

- For fixed radius r

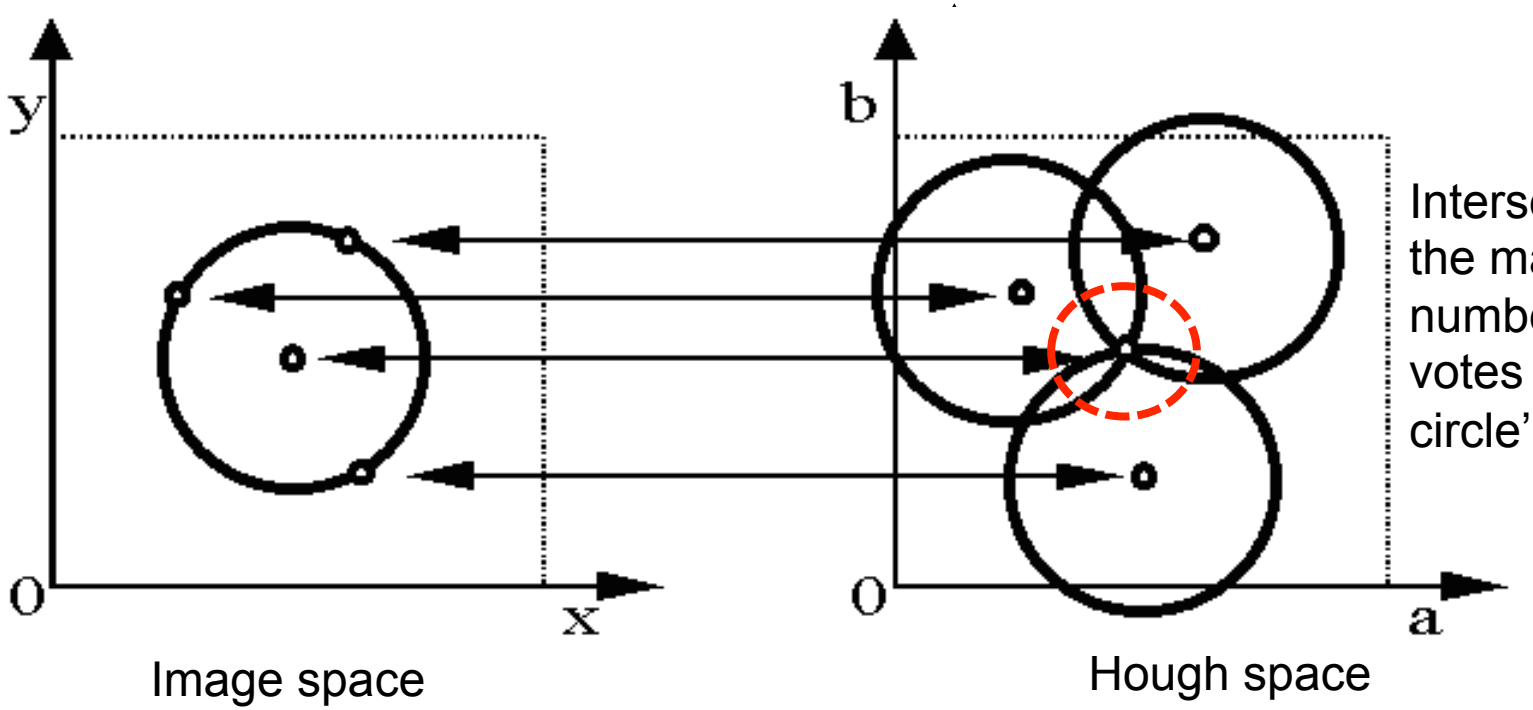


Hough transform for circles

- Circle: center (a,b) and radius r. Equation?

$$(x_i - a)^2 + (y_i - b)^2 = r^2$$

- For fixed radius r



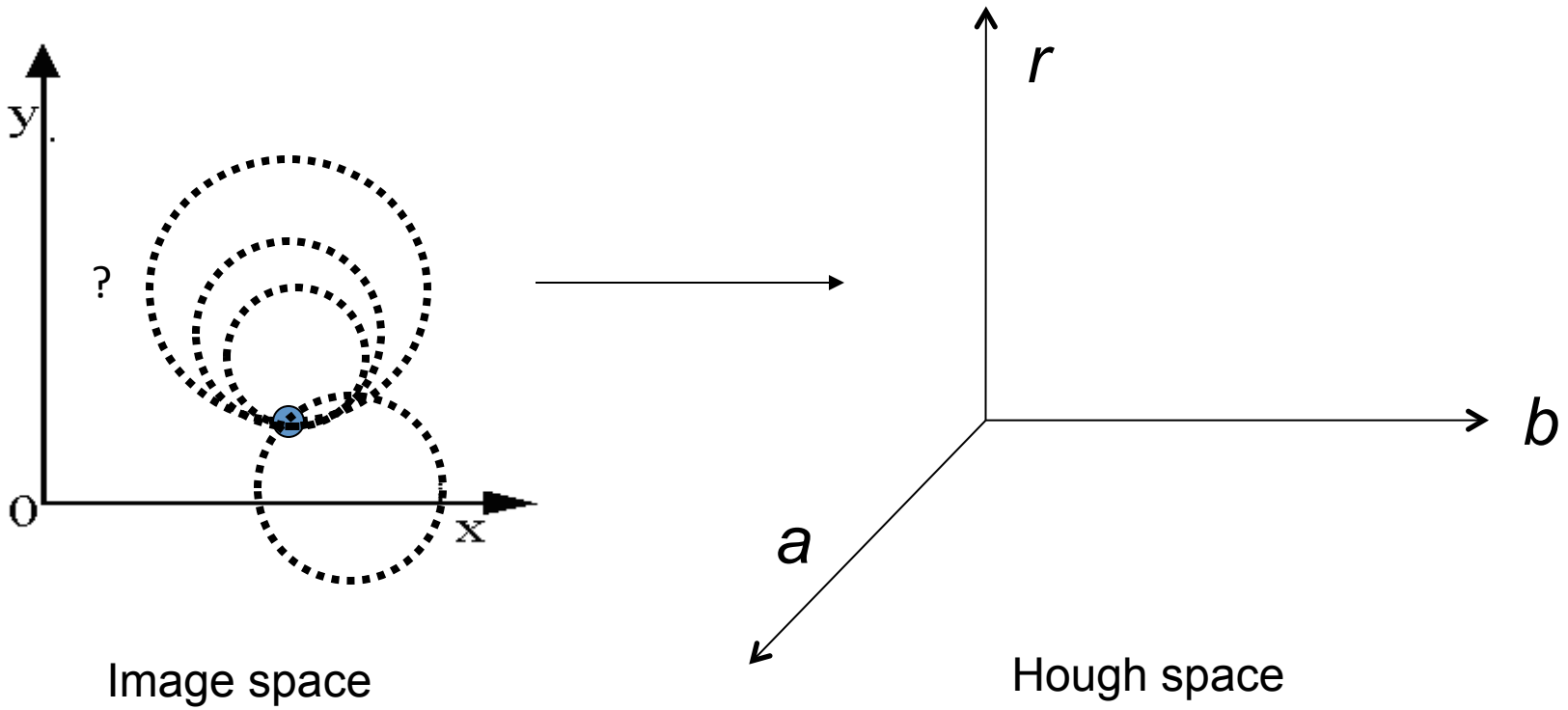
Slide credit: Kristen Grauman

Hough transform for circles

- Circle: center (a,b) and radius r. Equation?

$$(x_i - a)^2 + (y_i - b)^2 = r^2$$

- For unknown radius r



Slide credit: Kristen Grauman

Hough transform for circles

- Circle: center (a,b) and radius r. Equation?

$$(x_i - a)^2 + (y_i - b)^2 = r^2$$

- For unknown radius r

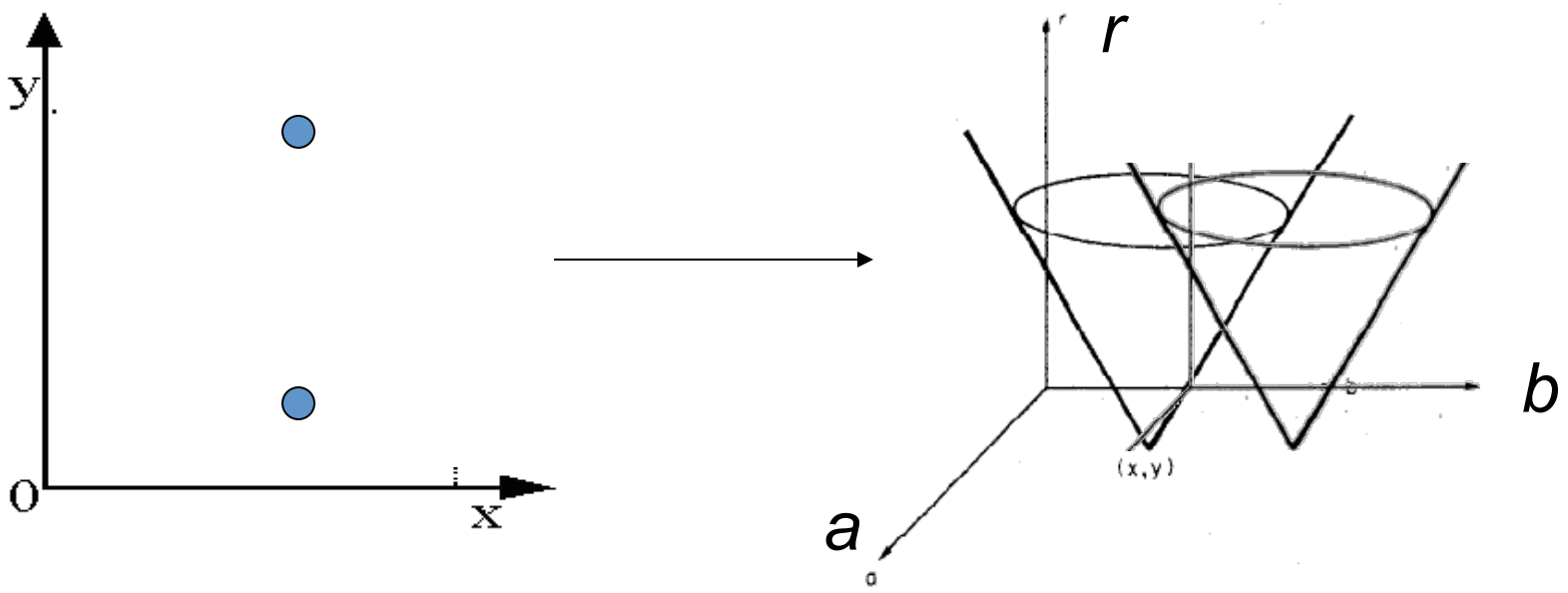


Image space

Hough space

Slide credit: Kristen Grauman

Hough transform for circles

- Circle: center (a,b) and radius r. Equation?

$$(x_i - a)^2 + (y_i - b)^2 = r^2$$

- For unknown radius r, with known gradient direction

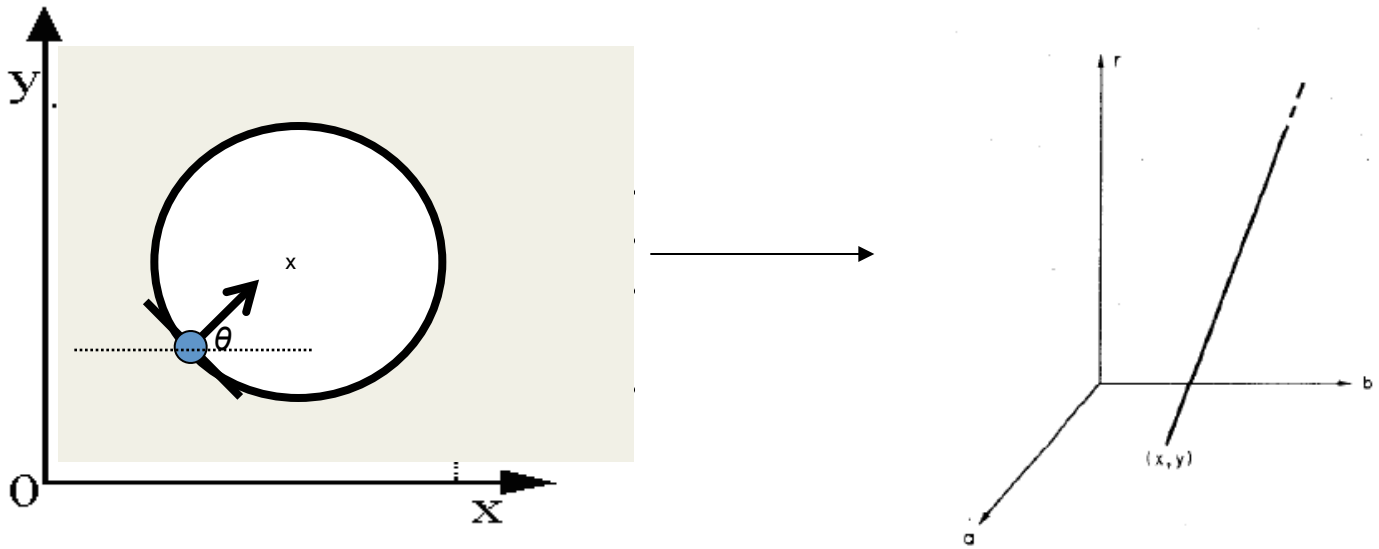


Image space

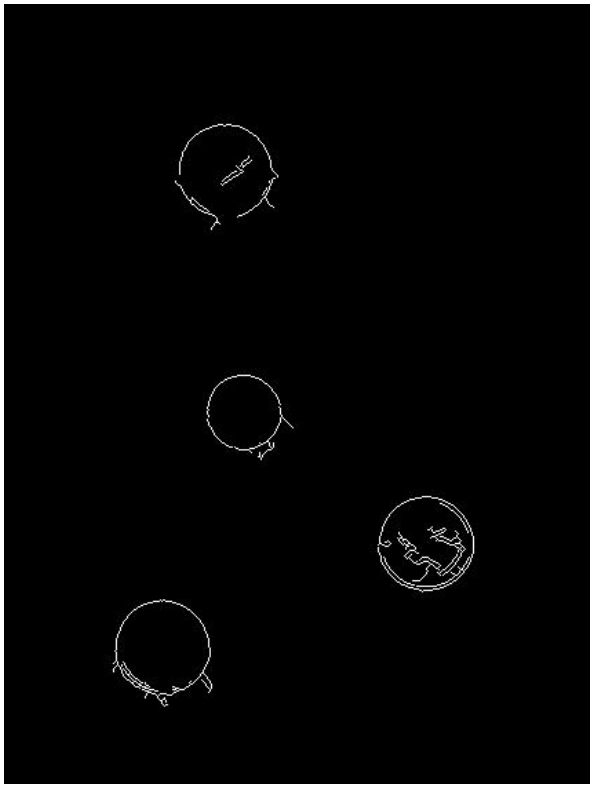
Hough space

Slide credit: Kristen Grauman

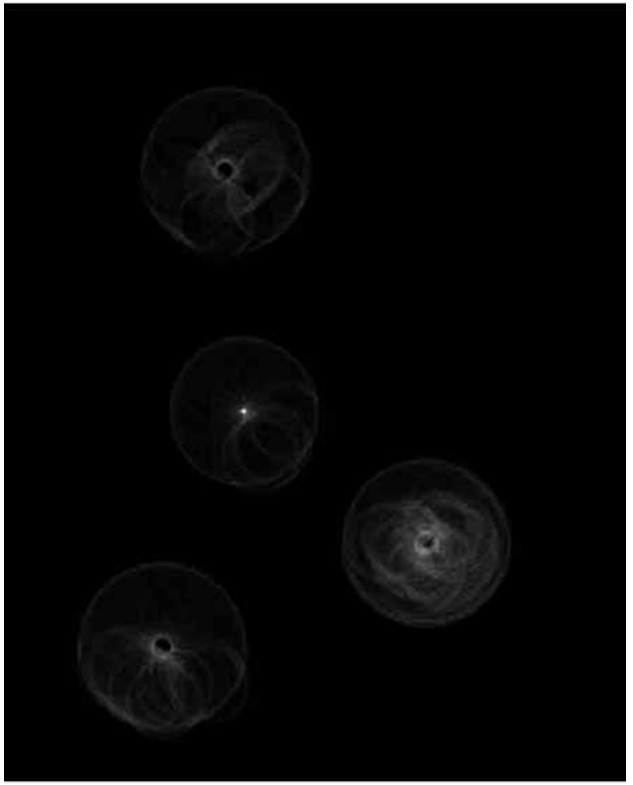
Circle detection using Hough transform

Original image

Edges



Votes ray 1



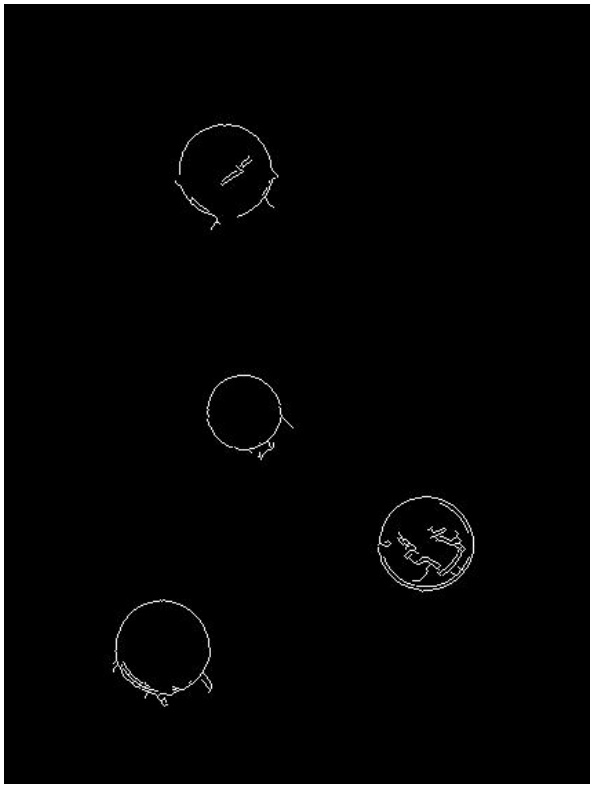
Use two Hough transforms, one for each ray (of known length)

Slide credit: Kristen Grauman

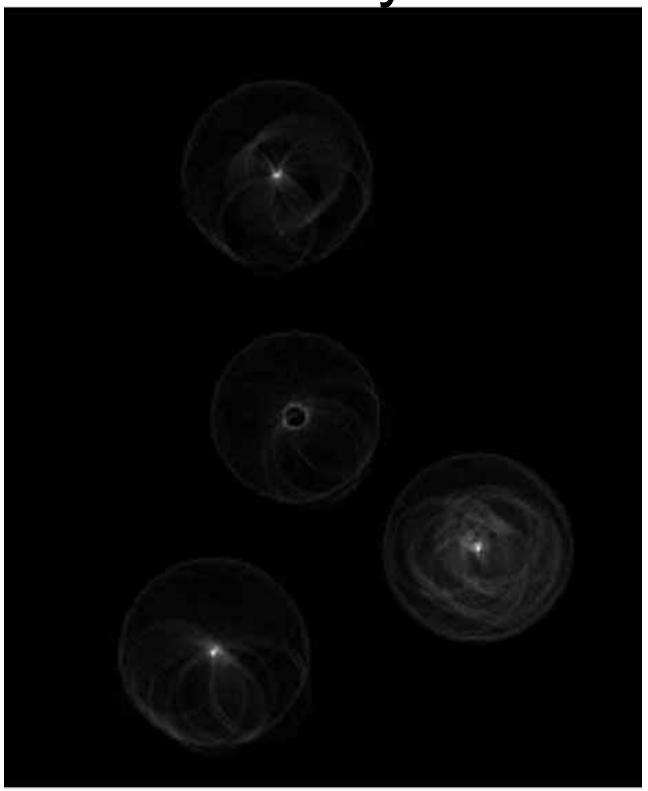
Circle detection using Hough transform

Original image

Edges



Votes ray 2

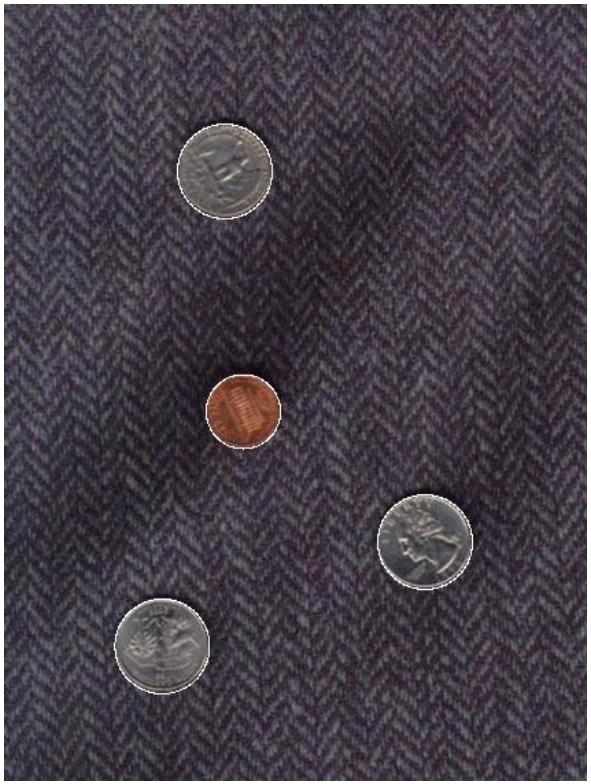


Use two Hough transforms, one for each ray (of known length)

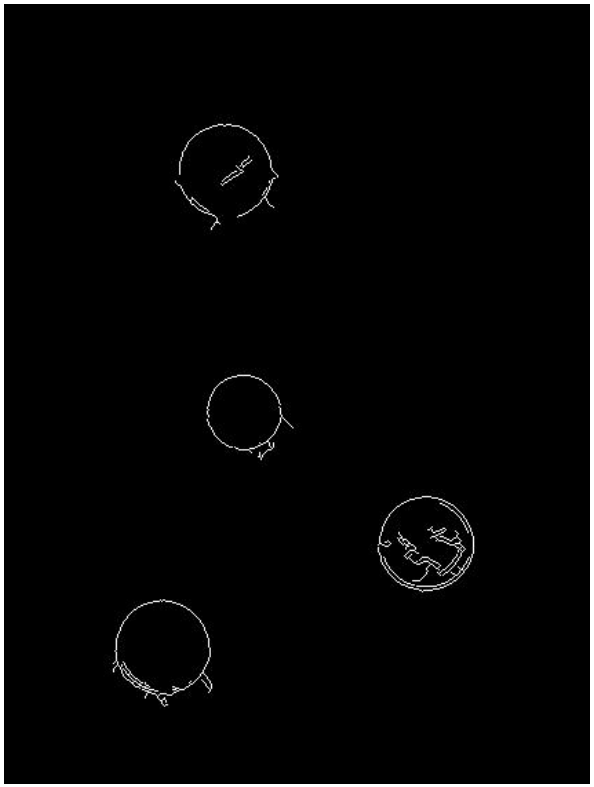
Slide credit: Kristen Grauman

Circle detection using Hough transform

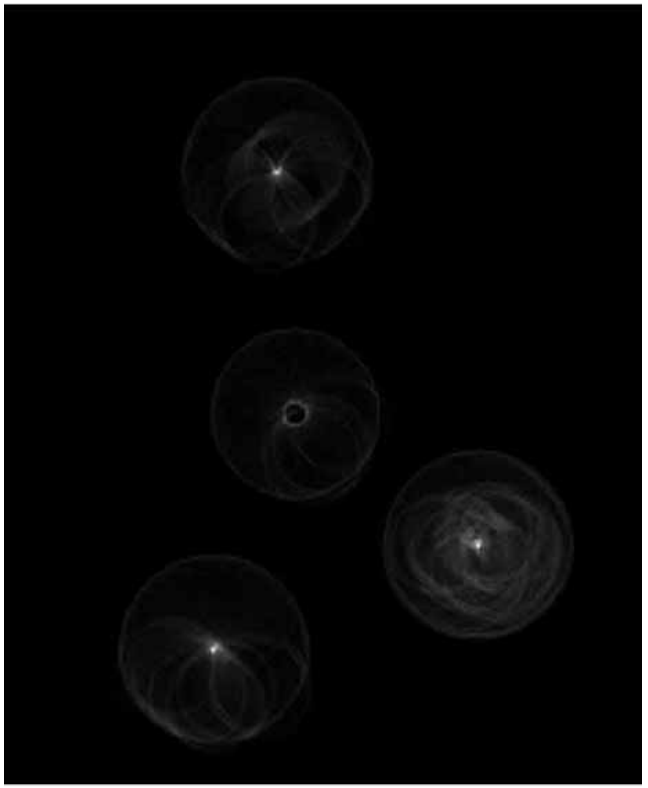
Combined detections



Edges



All votes



Slide credit: Kristen Grauman

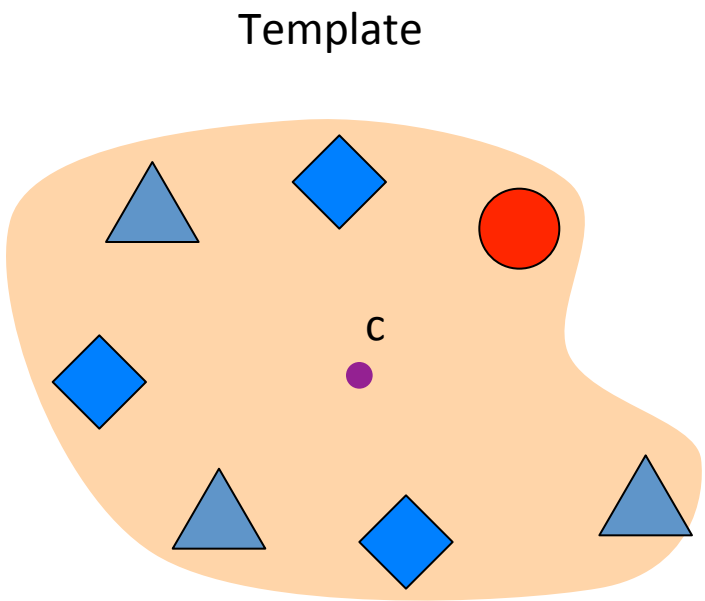
Hough transform: Pros and cons

- Pros
  - Can deal with non-locality and occlusion
  - Can detect multiple instances of a model
  - Some robustness to noise: noise points unlikely to contribute consistently to any single bin
- Cons
  - Complexity of search time increases exponentially with the number of model parameters
  - Non-target shapes can produce spurious peaks in parameter space
  - It's hard to pick a good grid size

Slide credit: Kristen Grauman

Generalized Hough transform

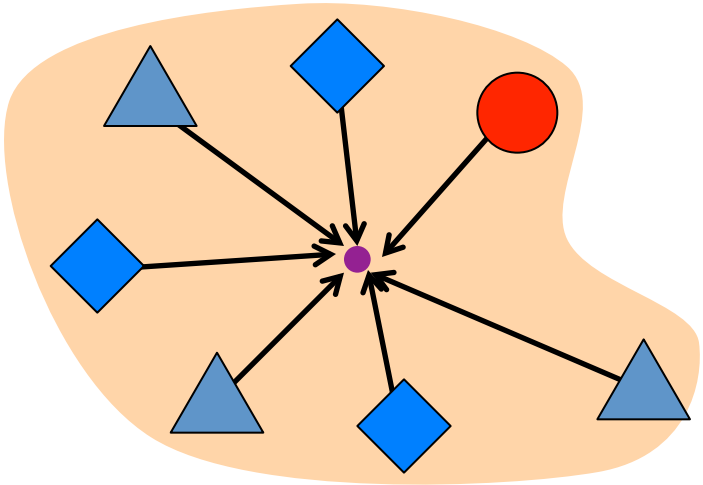
- We want to find a template defined by its reference point (center) and several distinct types of landmark points in stable spatial configuration



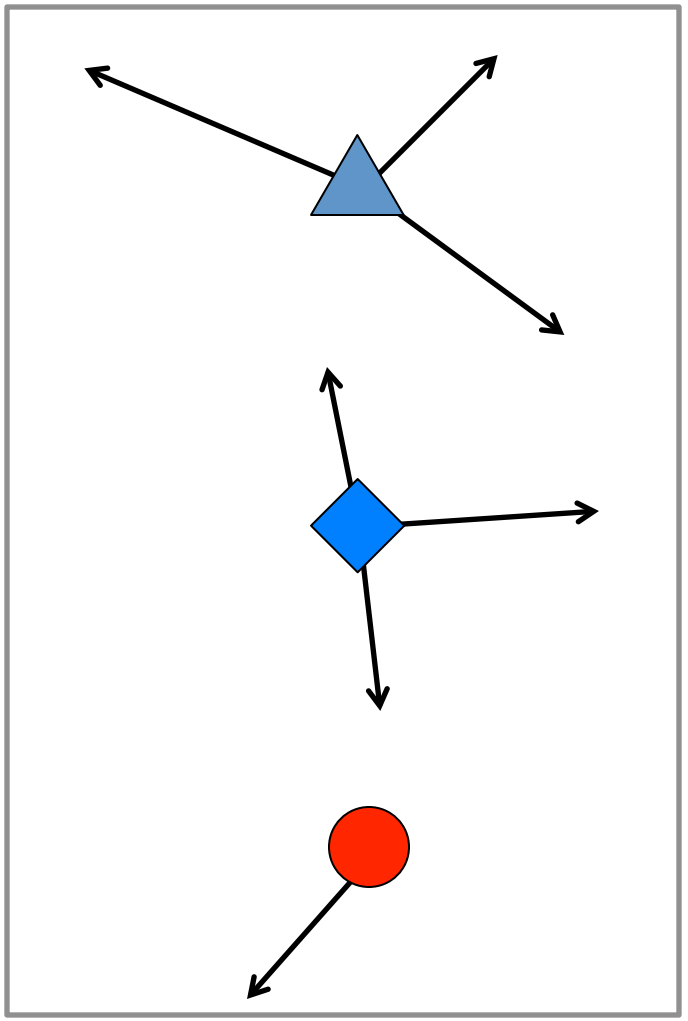
Generalized Hough transform

- Template representation: for each type of landmark point, store all possible displacement vectors towards the center

Template

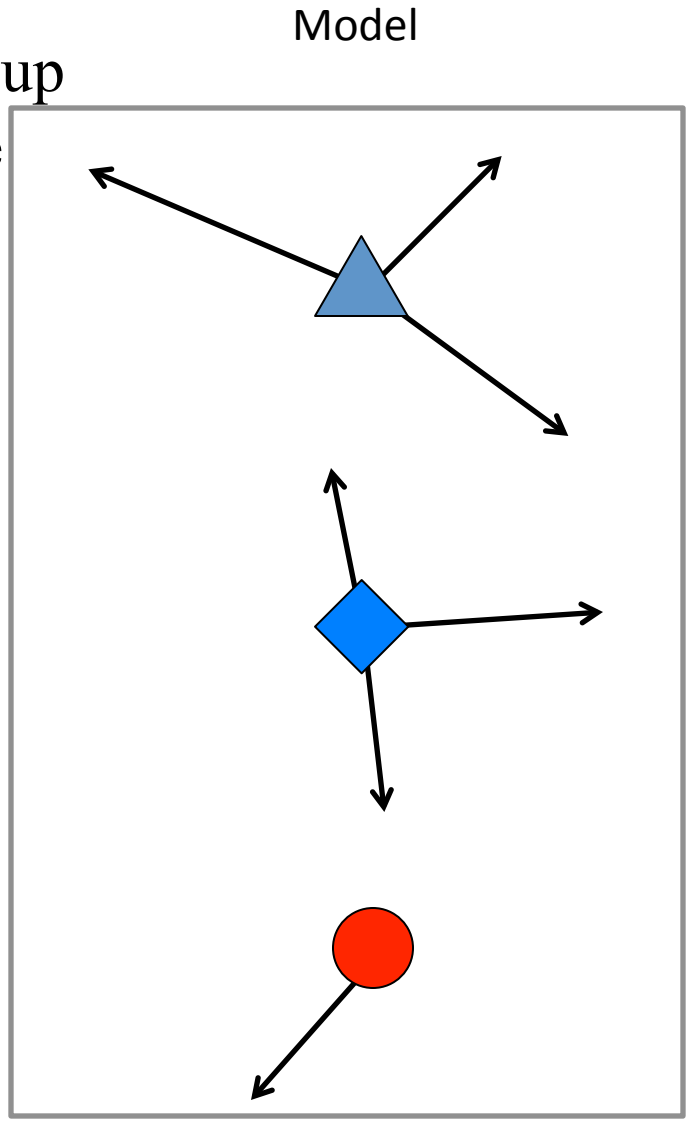


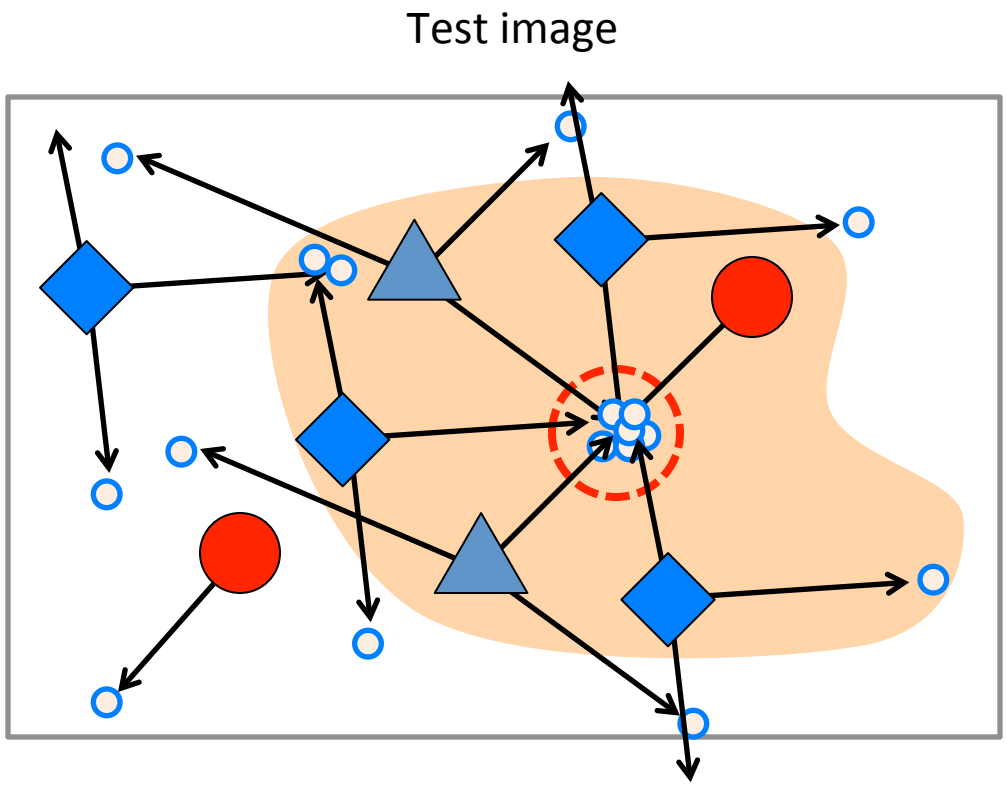
Model



Generalized Hough transform

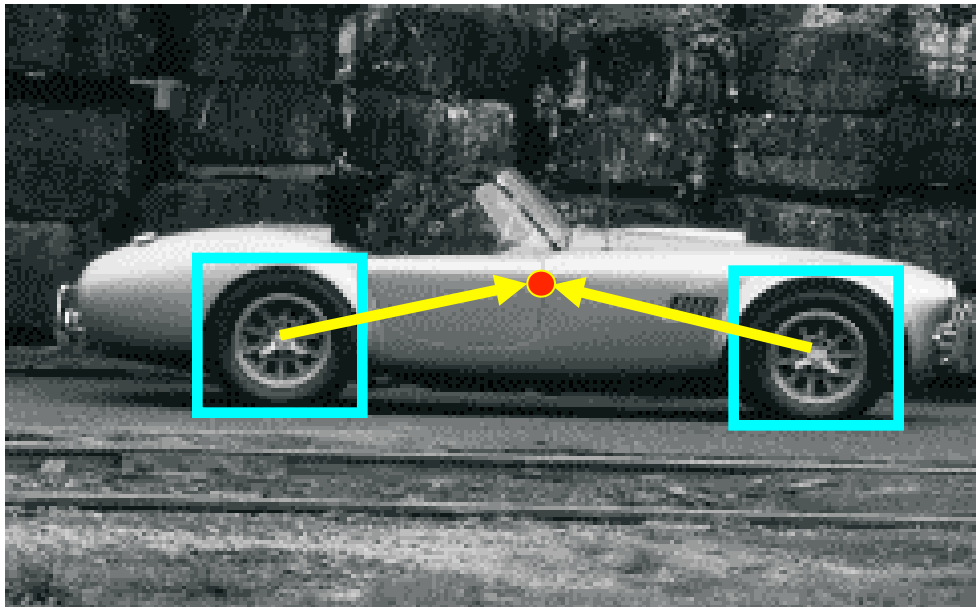
- Detecting the template:
  - For each feature in a new image, look up that feature type in the model and vote for the possible center locations associated with that type in the model





Application in recognition

- Index displacements by “visual codeword”



training image

visual codeword with displacement vectors

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

Application in recognition

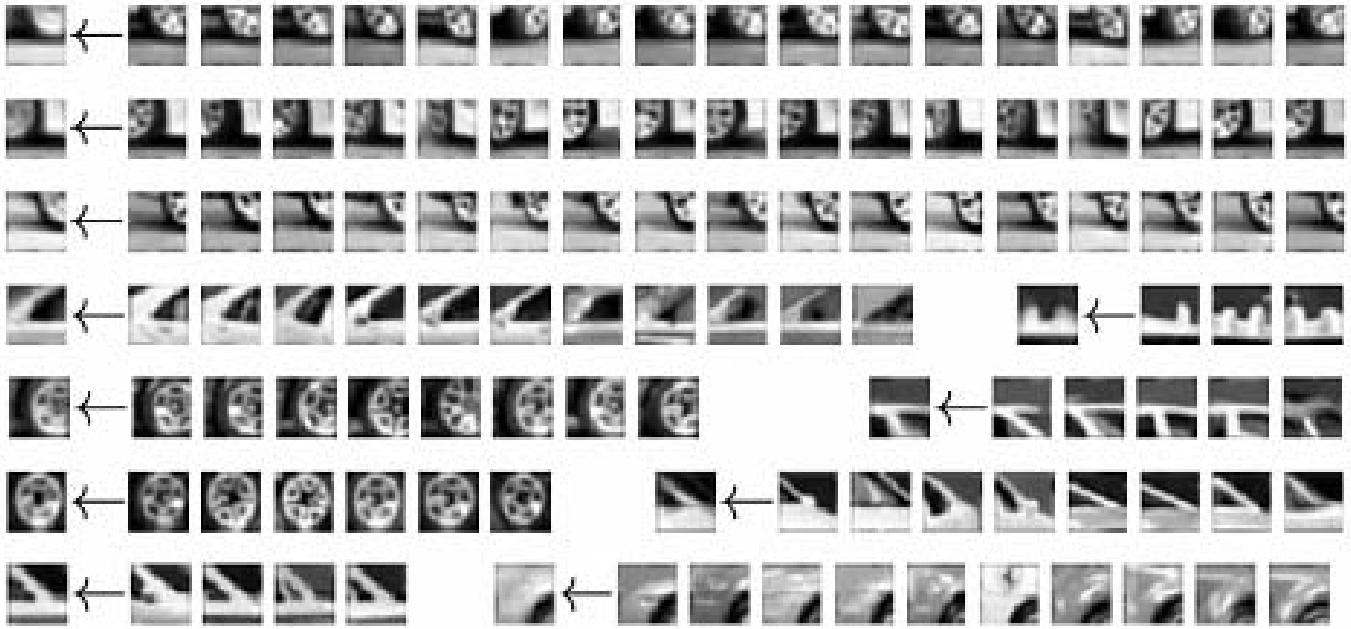
- Index displacements by “visual codeword”

test image

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

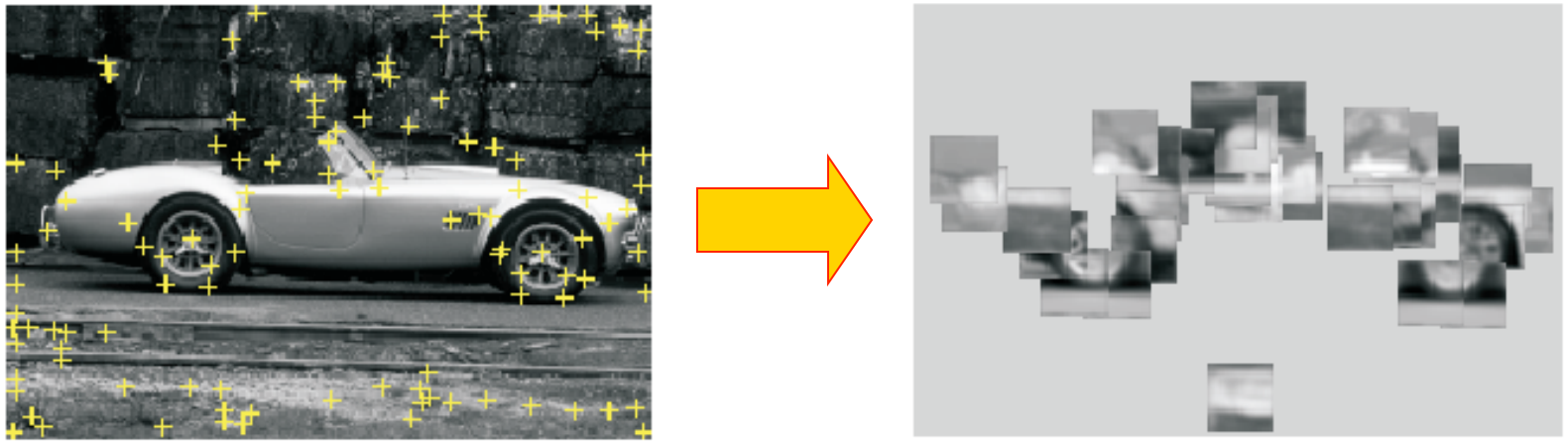
Implicit shape models: Training

1. Build *codebook* of patches around extracted interest points using clustering



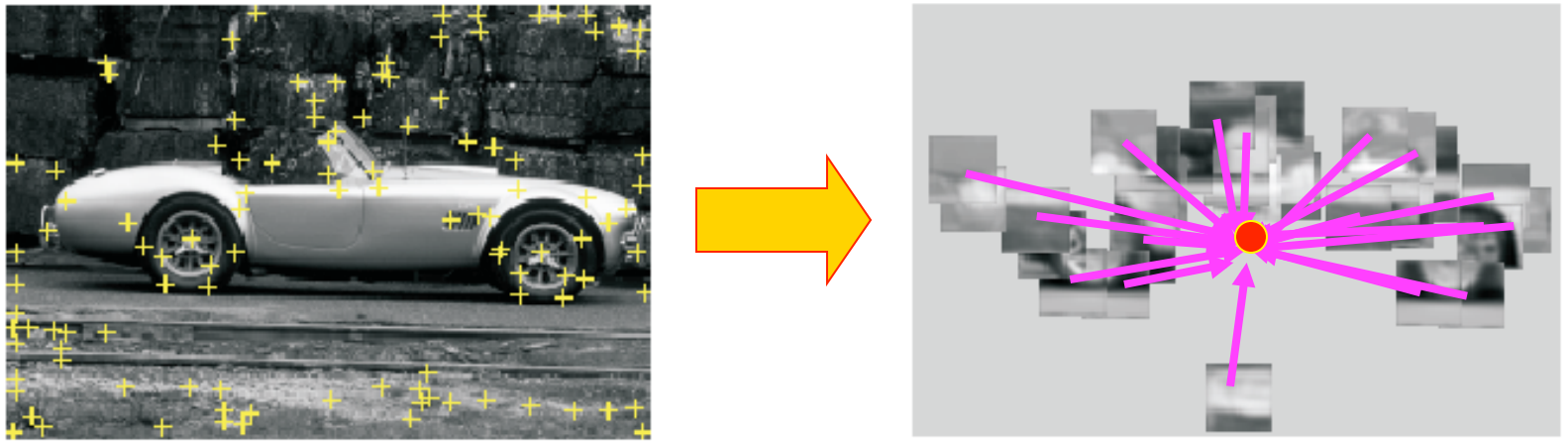
Implicit shape models: Training

1. Build *codebook* of patches around extracted interest points using clustering
2. Map the patch around each interest point to closest codebook entry



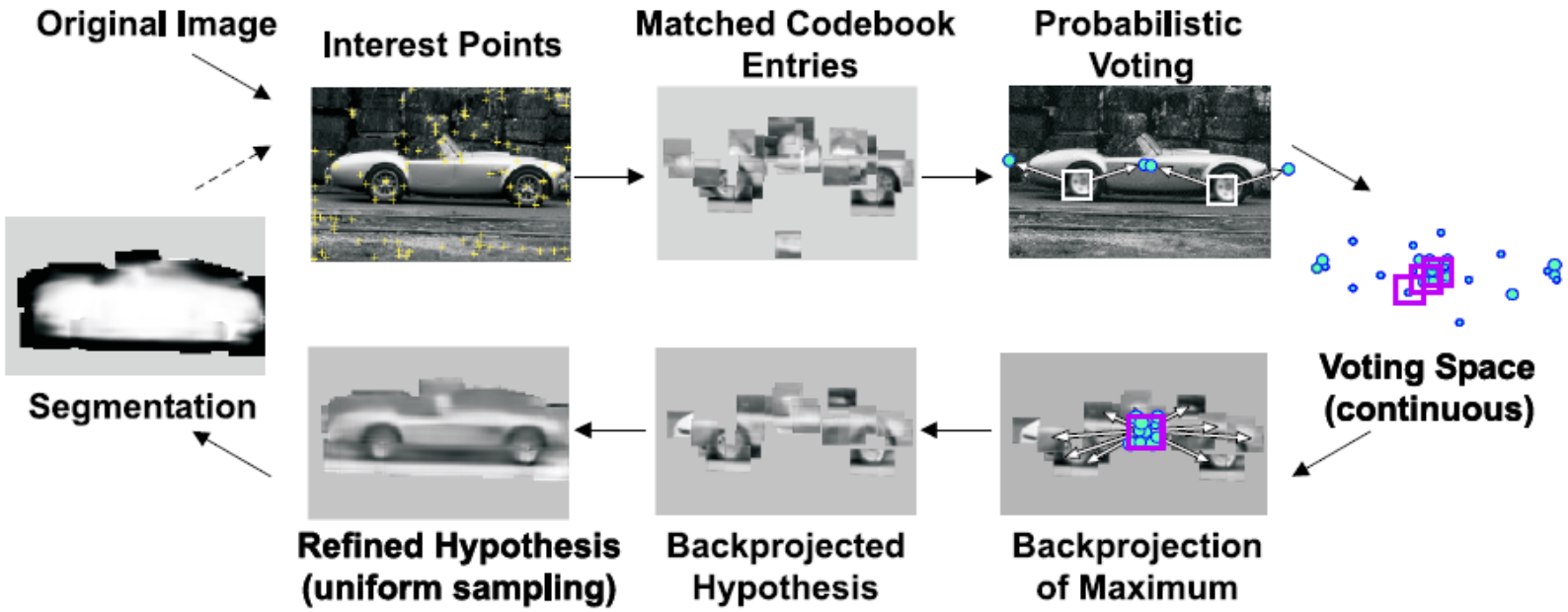
Implicit shape models: Training

1. Build *codebook* of patches around extracted interest points using clustering
2. Map the patch around each interest point to closest codebook entry
3. For each codebook entry, store all positions it was found, relative to object center

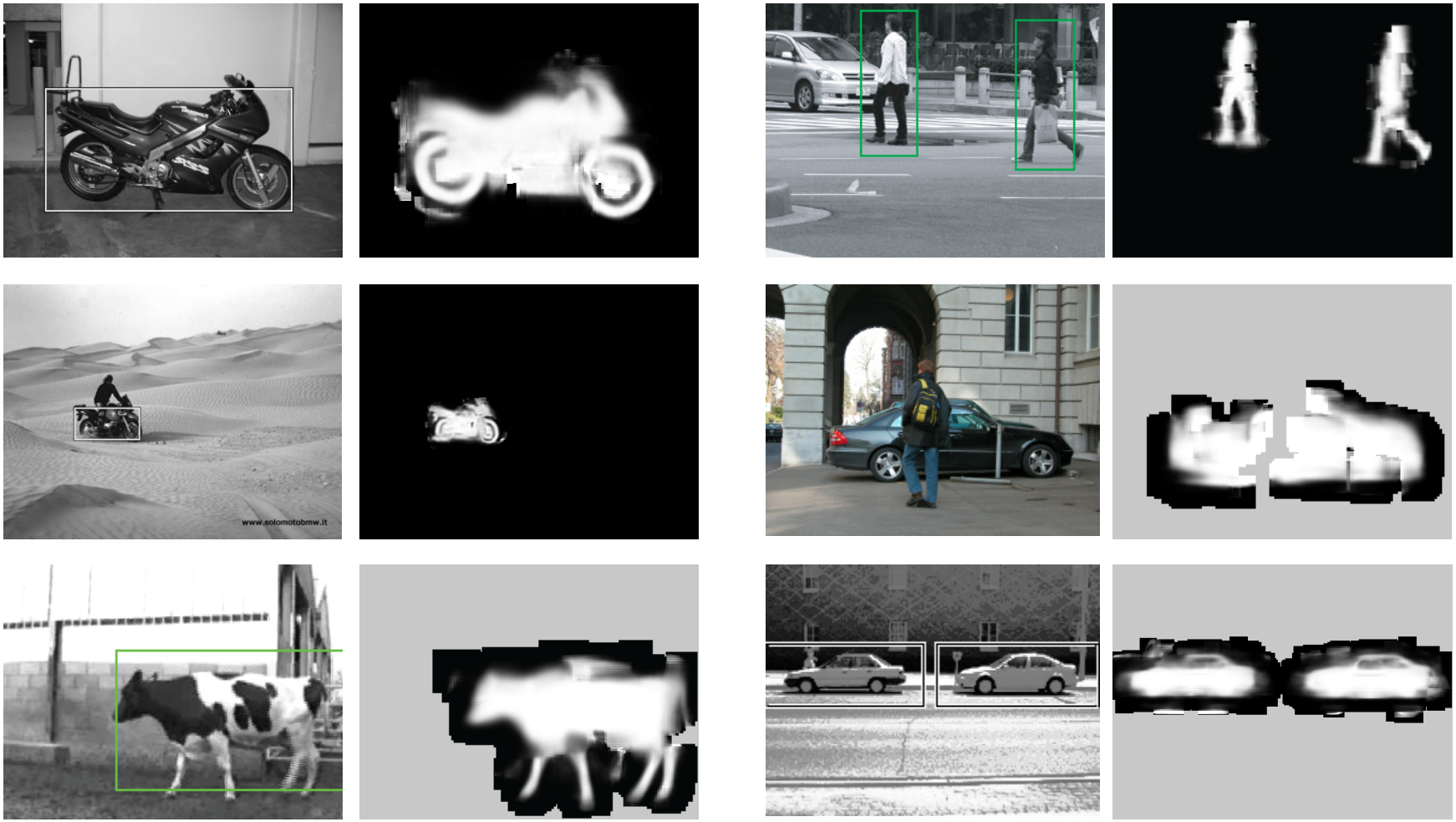


Implicit shape models: Testing

1. Given test image, extract patches, match to codebook entry
2. Cast votes for possible positions of object center
3. Search for maxima in voting space
4. Extract weighted segmentation mask based on stored masks for the codebook occurrences

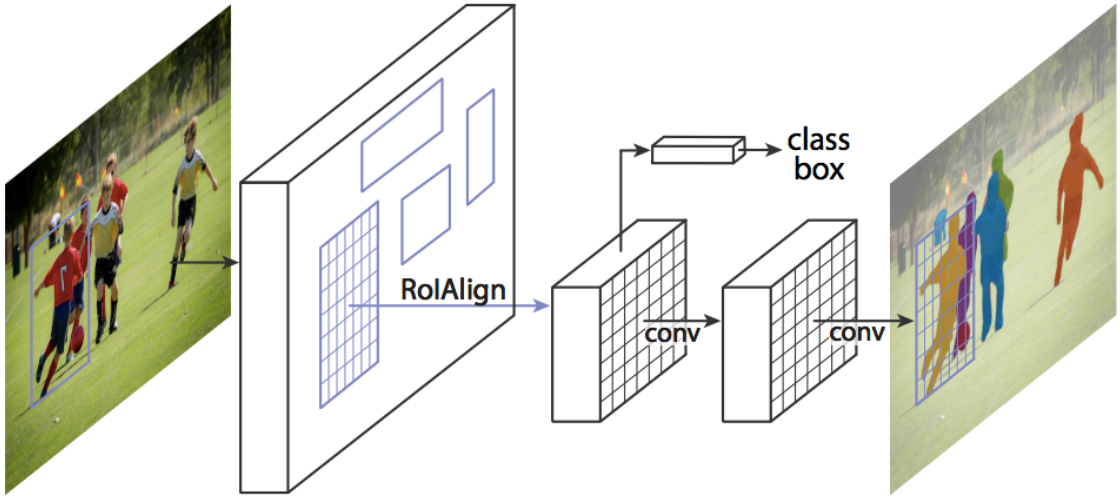


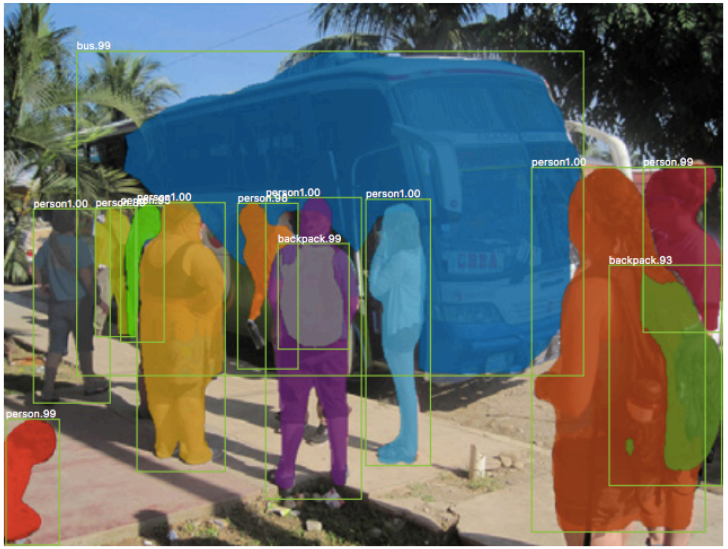
Additional examples

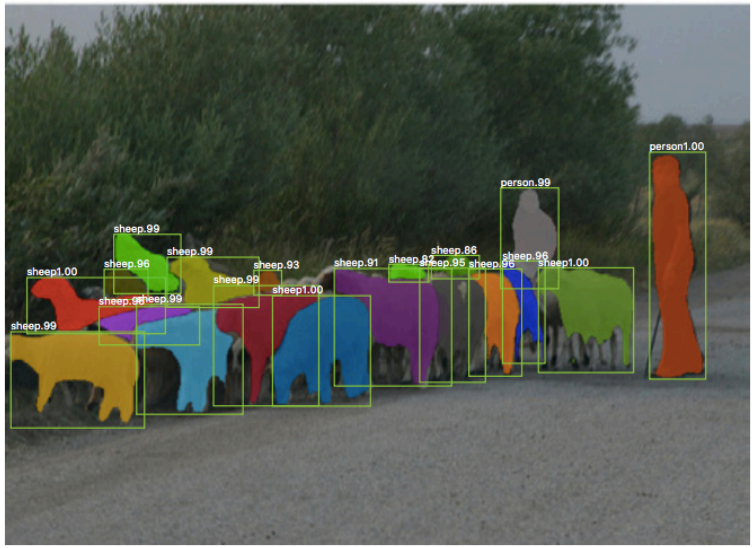


B. Leibe, A. Leonardis, and B. Schiele,
Robust Object Detection with Interleaved Categorization and Segmentation,
IJCV 77 (1-3), pp. 259-289, 2008.

These days: Mask R-CNN







K. He, G. Gkioxari, P. Dollar, and R. Girshick, Mask R-CNN, ICCV 2017