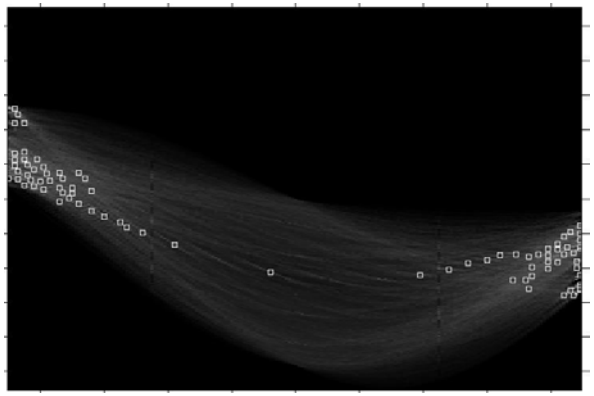
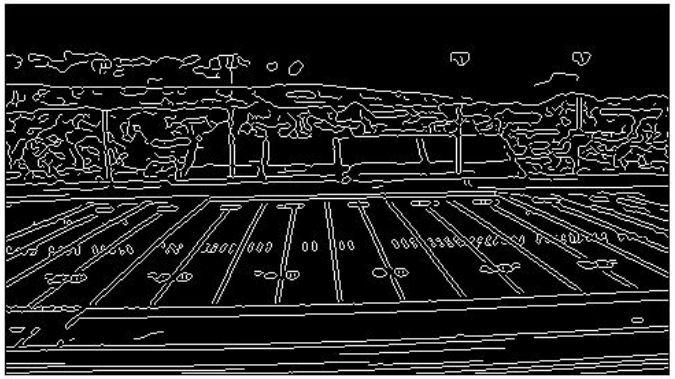
CS 4495 Computer Vision

Finding 2D Shapes and the Hough Transform

Aaron Bobick

School of Interactive Computing





Administrivia

- Today: Modeling Lines and Finding them
- CS4495: Problem set 1 is still posted.
  - Please read the directions carefully.
  - You ***can*** use Matlab (or OpenCV or Octave) edge operators
  - You ***cannot*** use Matlab (or OpenCV or Octave) Hough methods.
  - Due Sunday, Sept 7th 11:55pm. (T-Sq does not allow 11:59pm)
    - No – there is no grace period. Really. Really.
- Apparently Piazza put my email address on a "bounced-back" list so I was not getting Piazza email. Should be better now.

Now some "real" vision…

- So far, we applied operators/masks/kernels to images to produce new image

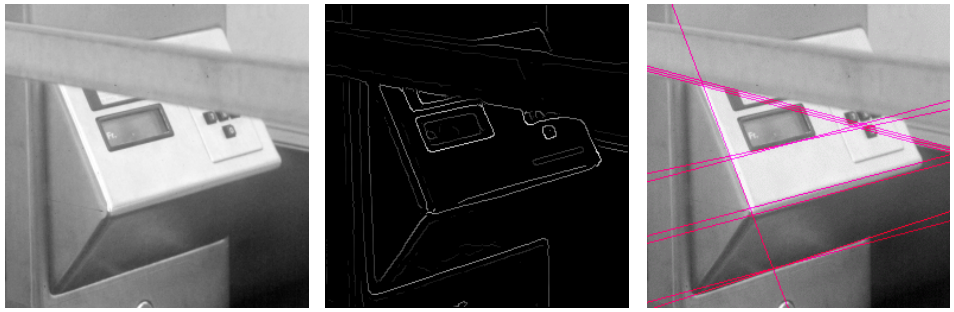
Image processing: $F : I(x, y) \longrightarrow I'(x, y)$

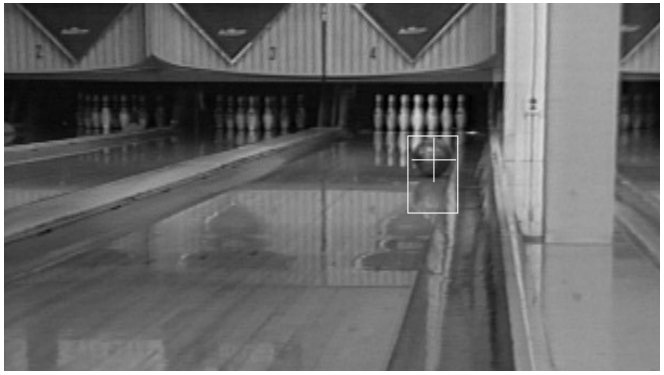
- Now real vision:

$$F : I(x, y) \longrightarrow \text{good stuff}$$

Fitting a model

- Want to associate a model with observed features





[Fig from Marszalek & Schmid, 2007]

For example, the model could be a line, a circle, or an arbitrary shape.

Fitting

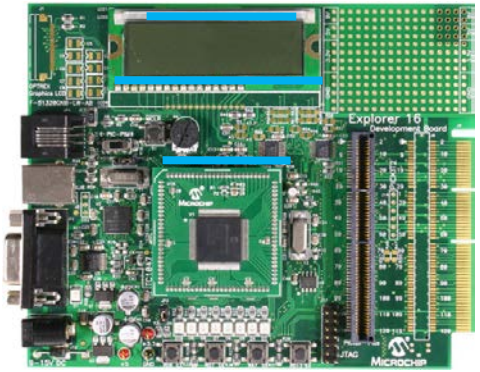
- Choose a ***parametric model*** to represent a set of features – e.g. line, circle, or even scaled templates.
- Membership criterion is not local
  - Can't tell whether a point in the image belongs to a given model just by looking at that point
- Three main questions:
  1. What model represents this set of features best?
  2. Which of several model instances gets which feature?
  3. How many model instances are there?
- Computational complexity is important
  - It is infeasible to examine every possible set of parameters and every possible combination of features

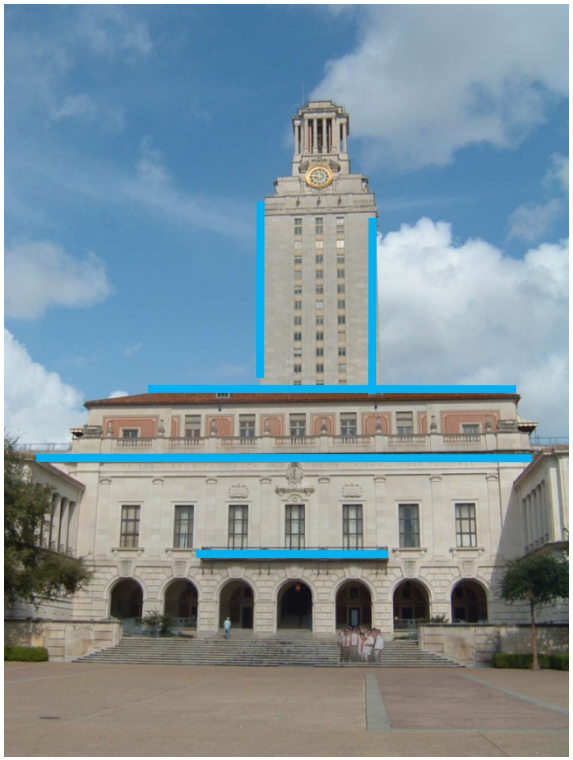
Source: L. Lazebnik

Example: Line fitting

- Why fit lines?

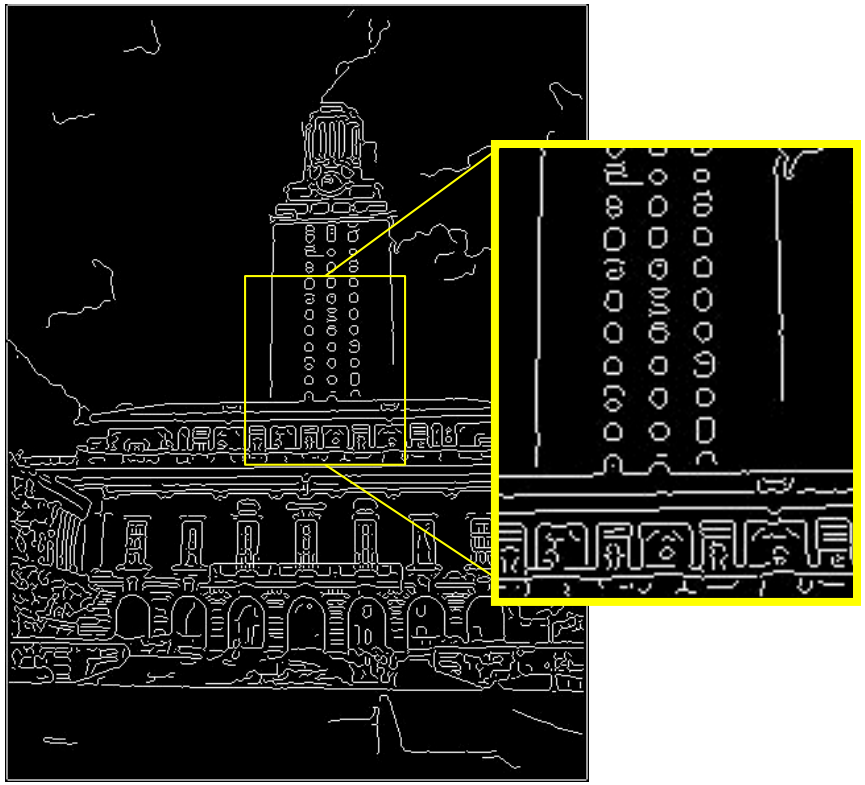
Many objects characterized by presence of straight lines





- Wait, why aren't we done just by running edge detection?

Difficulty of line fitting



- **Extra** edge points (clutter), multiple models:
  - which points go with which line, if any?
- Only some parts of each line detected, and some parts are **missing:**
  - how to find a line that bridges missing evidence?
- **Noise** in measured edge points, orientations:
  - how to detect true underlying parameters?

Voting

It's not feasible to check all possible models or all combinations of features (e.g. edge pixels) by fitting a model to each possible subset.

Voting is a general technique where we let the features vote for all models that are compatible with it.

1. Cycle through features, each casting votes for model parameters.
2. Look for model parameters that receive a lot of votes.

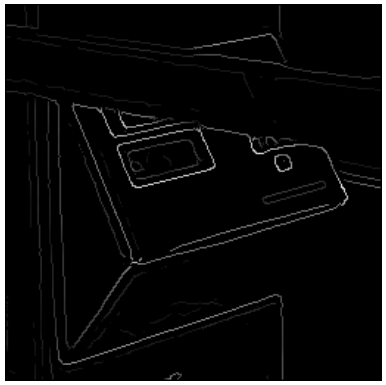
Voting – why it works

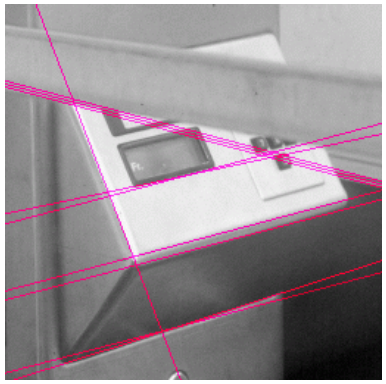
- Noise & clutter features will cast votes too, but typically their votes should be inconsistent with the majority of "good" features.
- Ok if some features not observed, as model can span multiple fragments.

Fitting lines

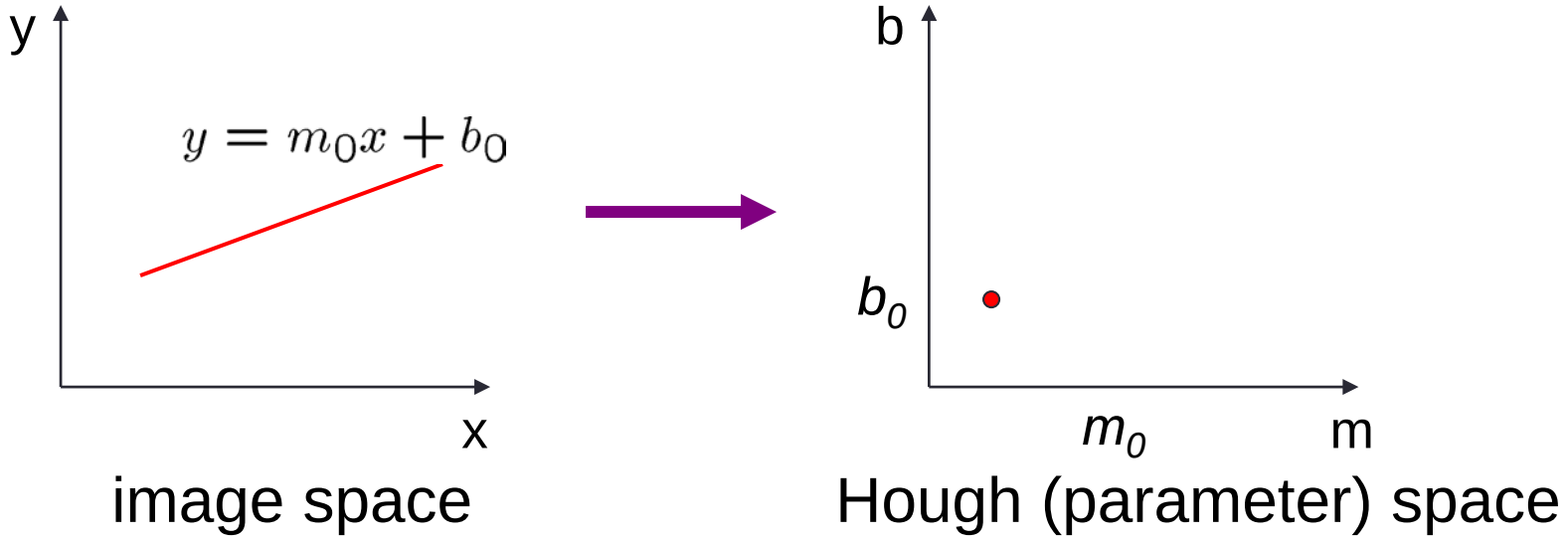
- Given points that belong to a line, what is the line?
- How many lines are there?
- Which points belong to which lines?

- **Hough Transform** is a voting technique that can be used to answer all of these
  - Main idea:
  - 1. Record all possible lines on which each edge point lies.
  - 2. Look for lines that get many votes.





Finding lines in an image: Hough space

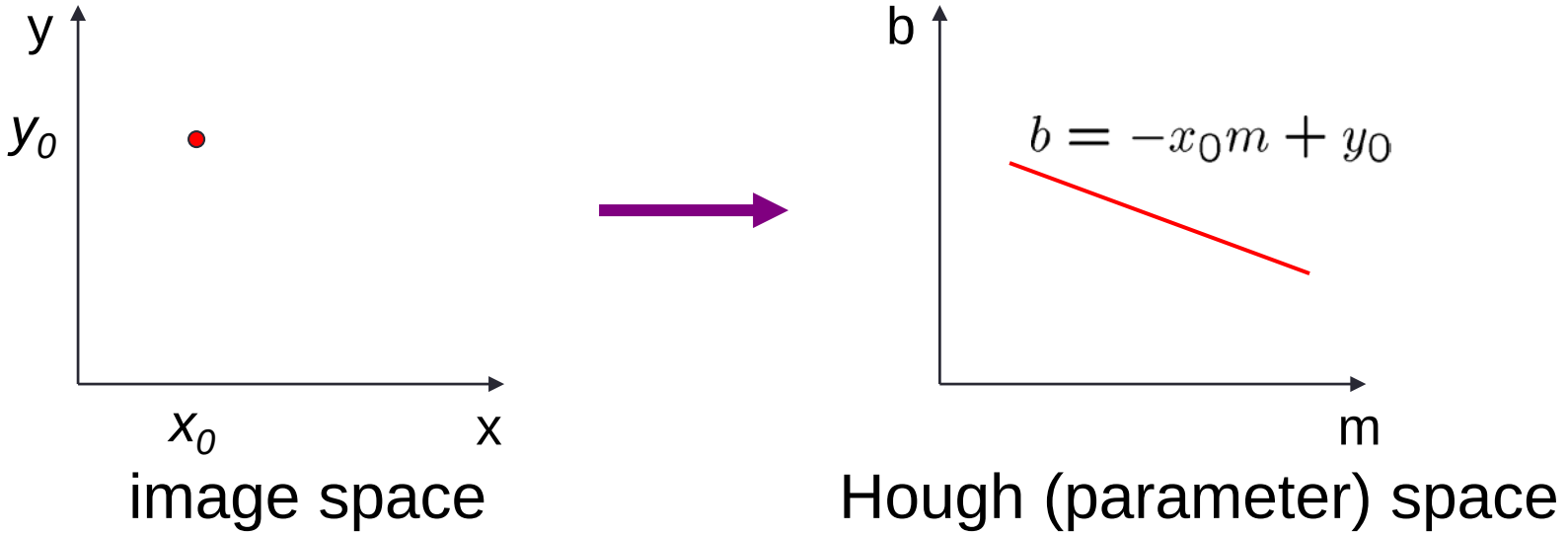


Connection between image (x,y) and Hough (m,b) spaces

- A line in the image corresponds to a point in Hough space
- To go from image space to Hough space:
  - given a set of points (x,y), find all (m,b) such that y = mx + b

Slide credit: Steve Seitz

Finding lines in an image: Hough space



Connection between image (x,y) and Hough (m,b) spaces

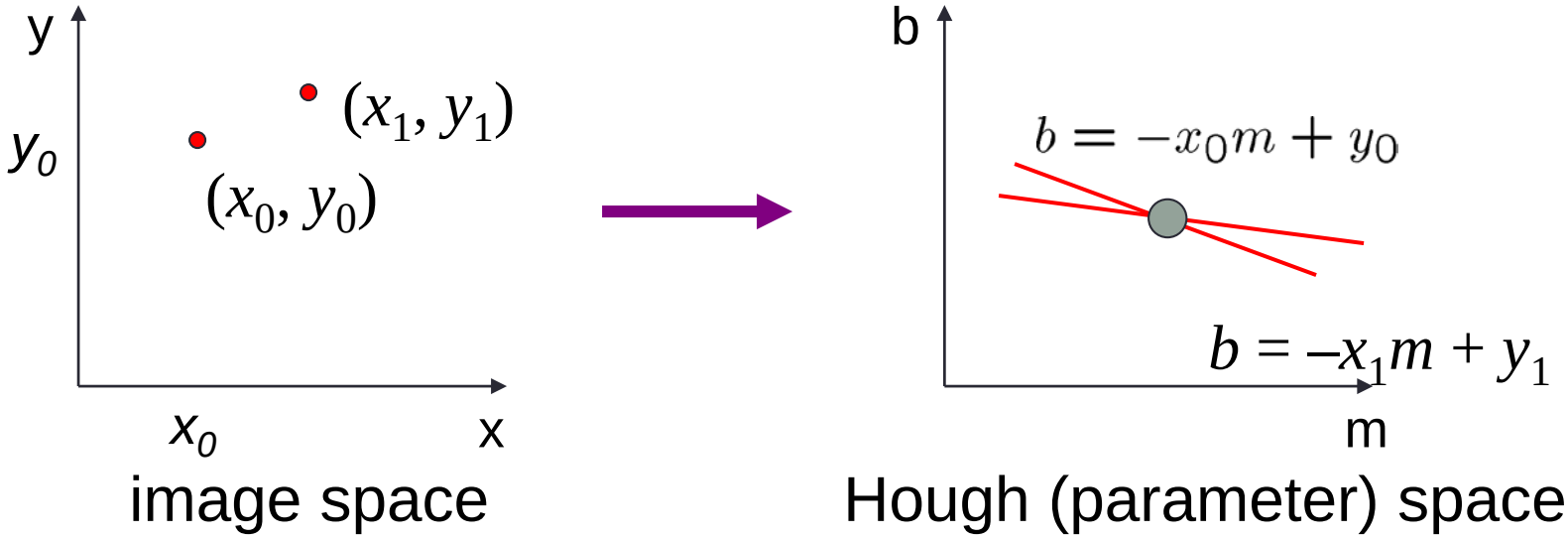
What does a point (x_0, y_0) in the image space map to?

Answer: the solutions of $\boldsymbol{b = -x_0m + y_0}$

this is a line in Hough space

Slide credit: Steve Seitz

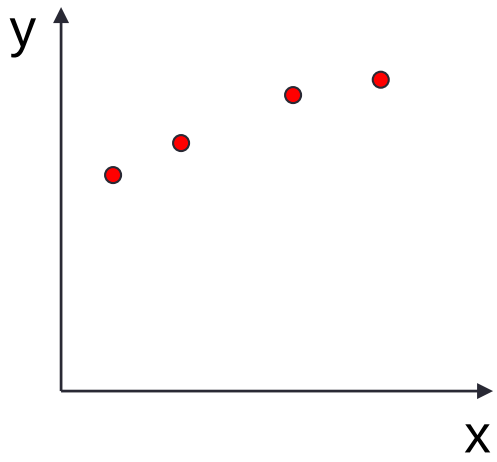
Finding lines in an image: Hough *transform*



What are the line parameters for the line that contains both (x_0, y_0) and (x_1, y_1)?

- It is the intersection of the lines $\boldsymbol{b = -x_0m + y_0}$ and $\boldsymbol{b = -x_1m + y_1}$

Finding lines: Hough algorithm



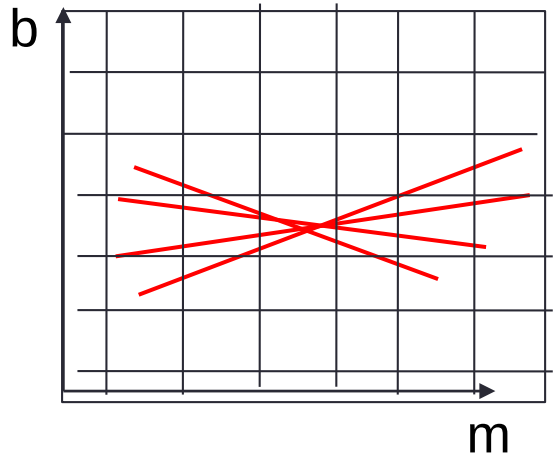


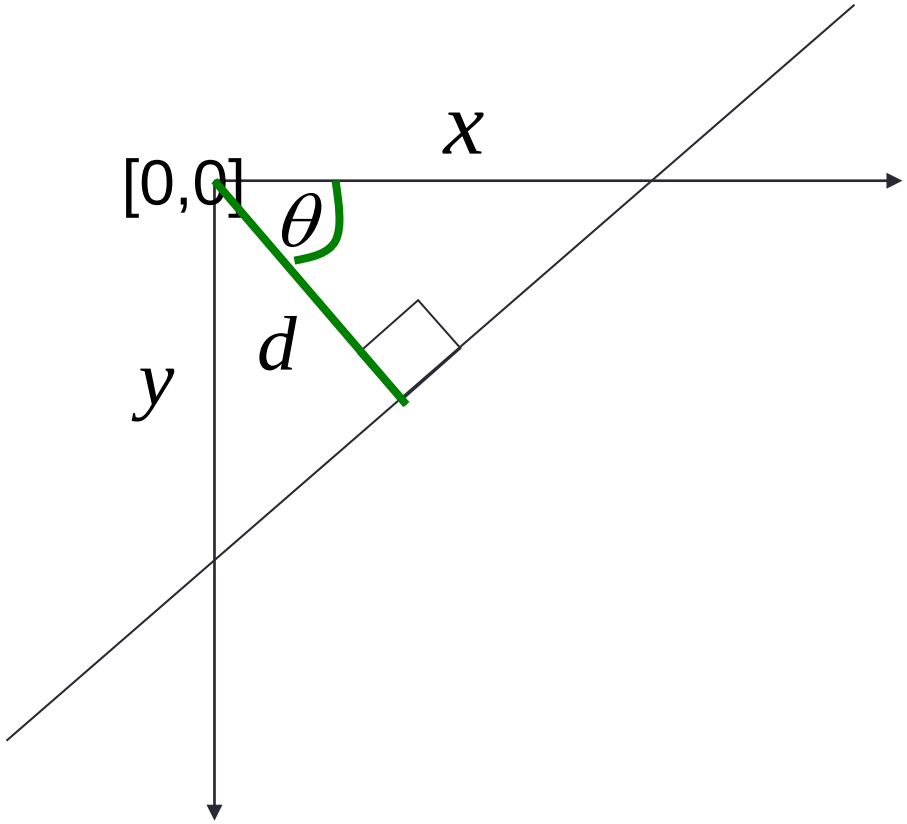
image space

Hough (parameter) space

- How can we use this to find the most likely parameters (m,b) for the most prominent line in the image space?
- Let each edge point in image space vote for a set of possible parameters in Hough space
- Accumulate votes in discrete set of bins; parameters with the most votes indicate line in image space.

Polar representation for lines

Issues with usual (*m,b*) parameter space: can take on infinite values, undefined for vertical lines.



d : perpendicular distance from line to origin

θ : angle the perpendicular makes with the x-axis

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

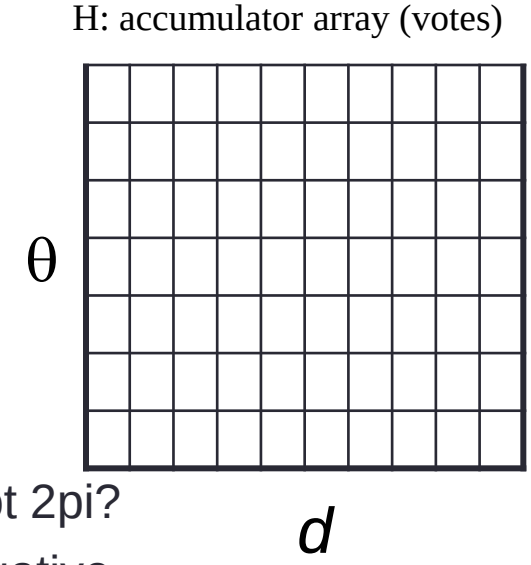
Point in image space → sinusoid segment in Hough space

Hough transform algorithm

Using the polar parameterization:

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

H: accumulator array (votes)



Basic Hough transform algorithm

1. Initialize H[d, θ]=0
2. for each edge point I[x,y] in the image

   for θ = 0 to 180 // some quantization; not 2pi?

   $d = x\cos\theta - y\sin\theta$ // maybe negative

   H[d, θ] += 1
3. Find the value(s) of (d, θ) where H[d, θ] is maximum
4. The detected line in the image is given by $d = x\cos\theta - y\sin\theta$

Space complexity? k^n (n dimensions, k bins each)

Time complexity (in terms of number of voting elements)?

Example: Hough transform for straight lines

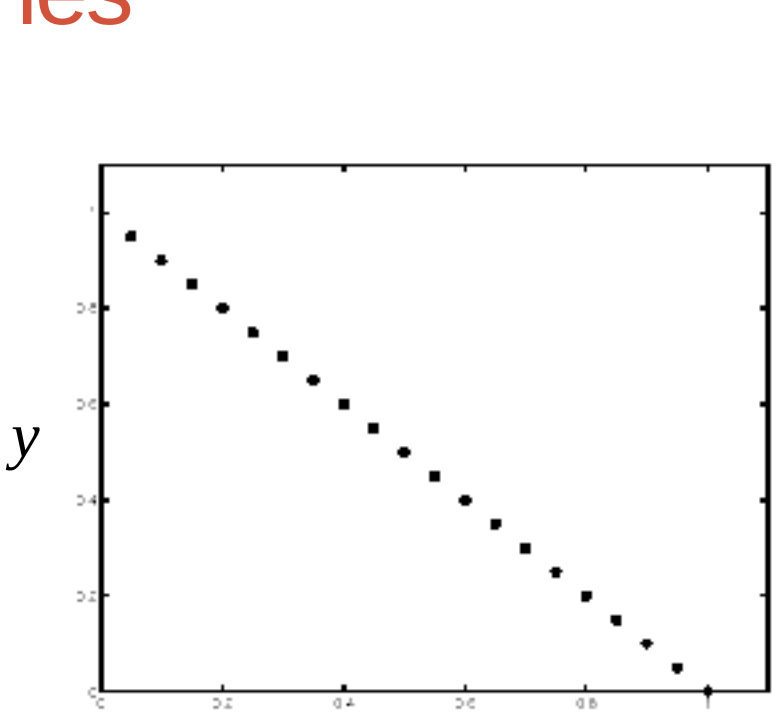
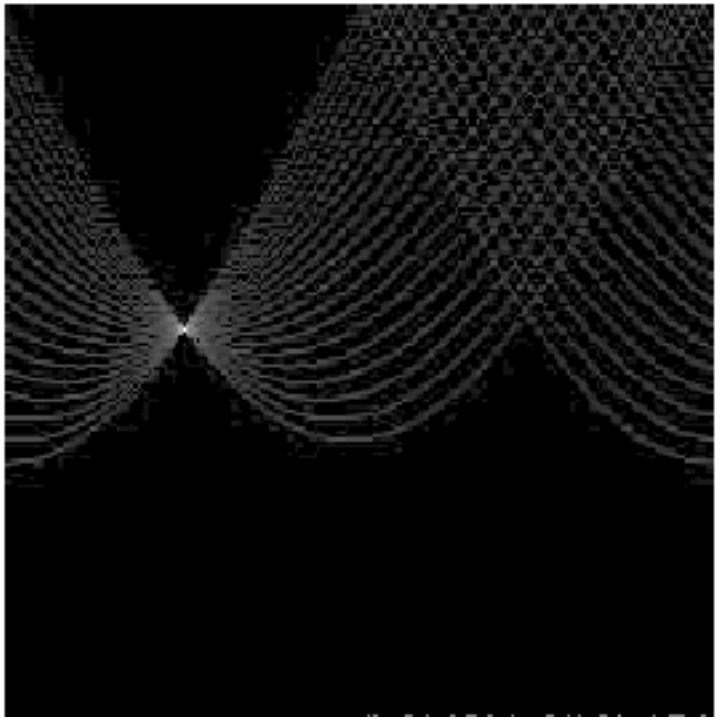


Image space
edge coordinates

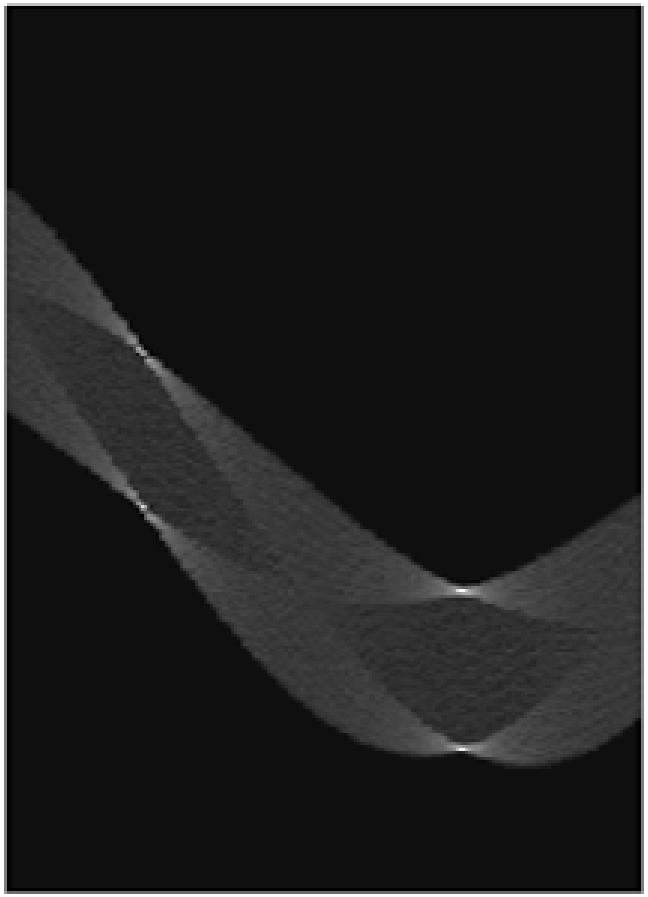


Votes

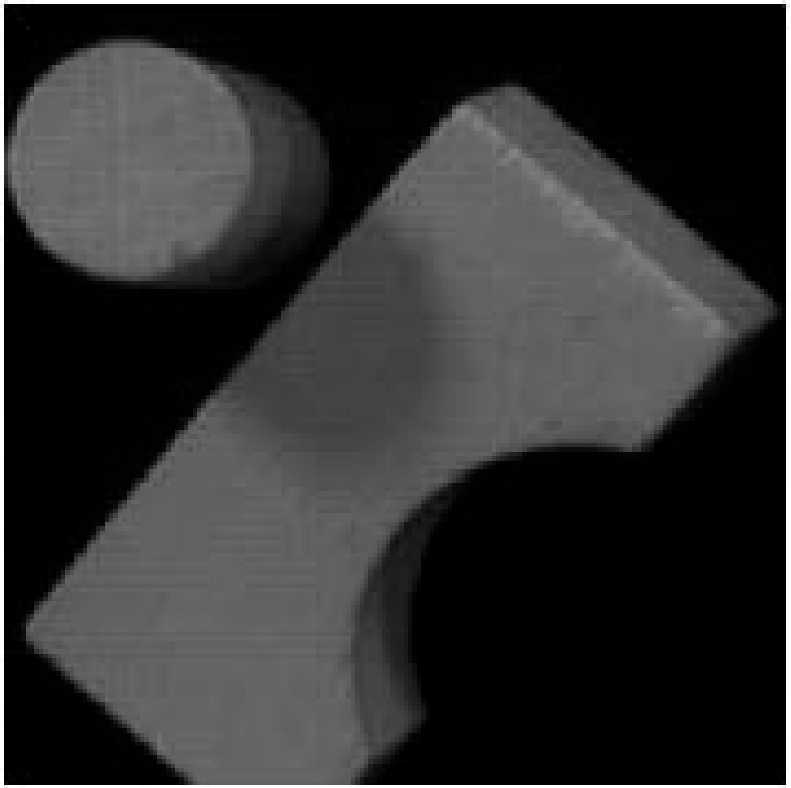
Bright value = high vote count
Black = no votes

Example: Hough transform for straight lines

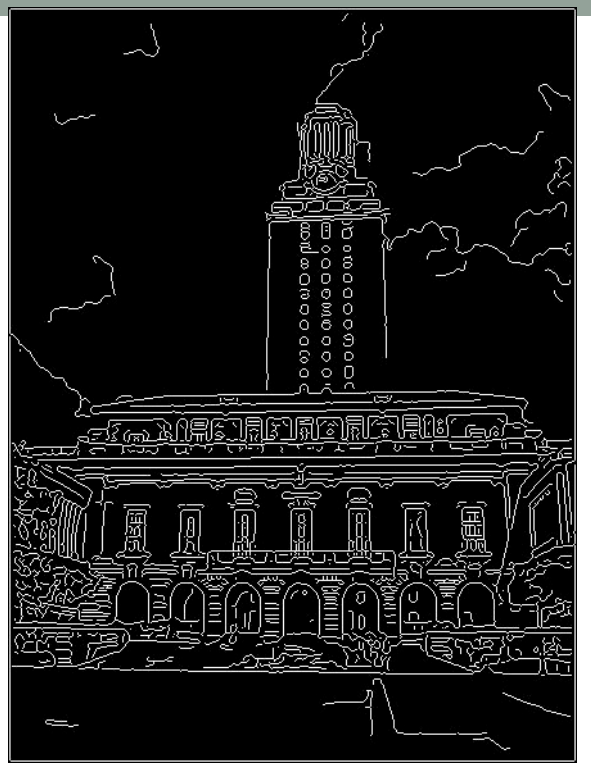
Square :

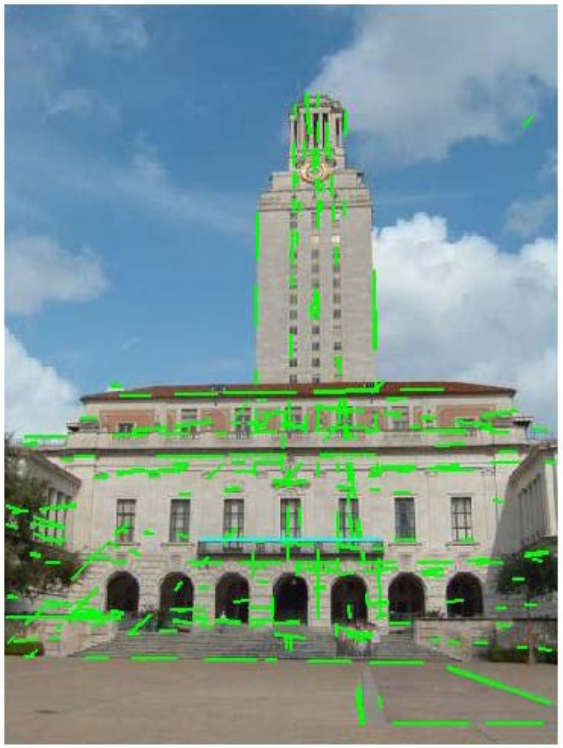


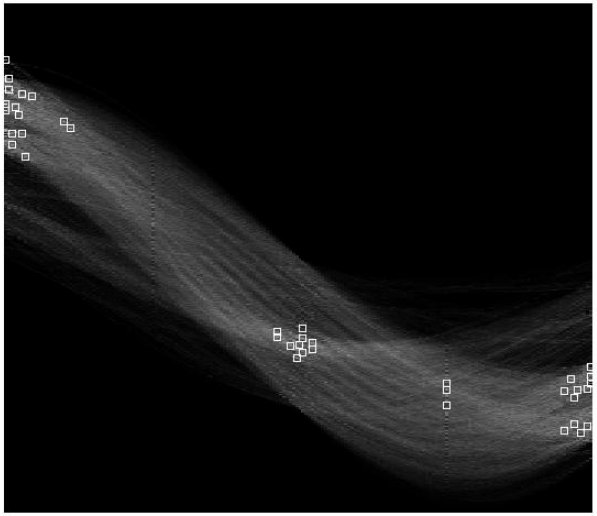
Example: Hough transform for straight lines

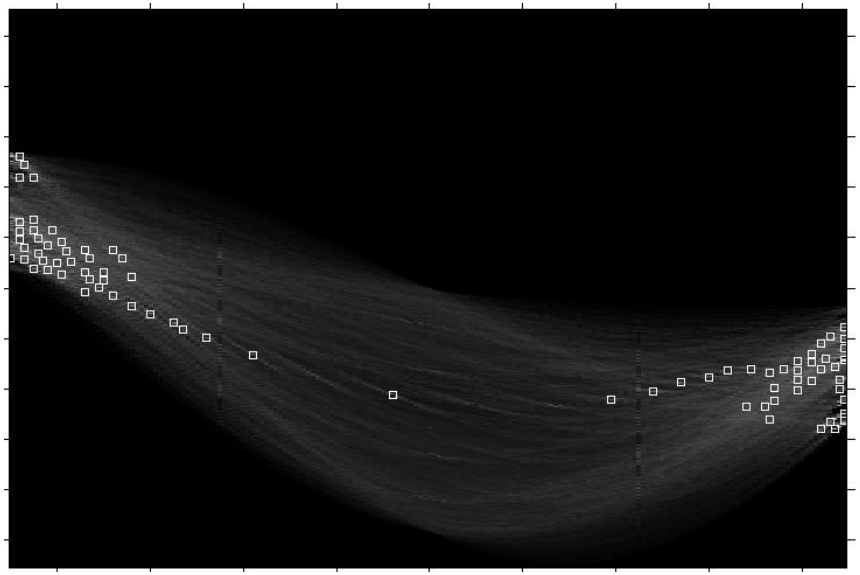


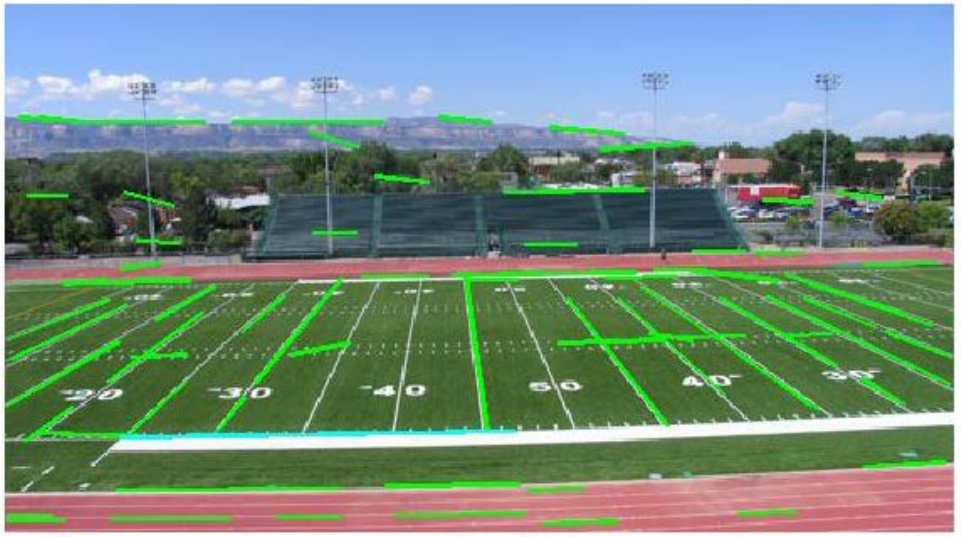
Hough demo..






ρ
θ





Showing longest segments found

Impact of noise on Hough

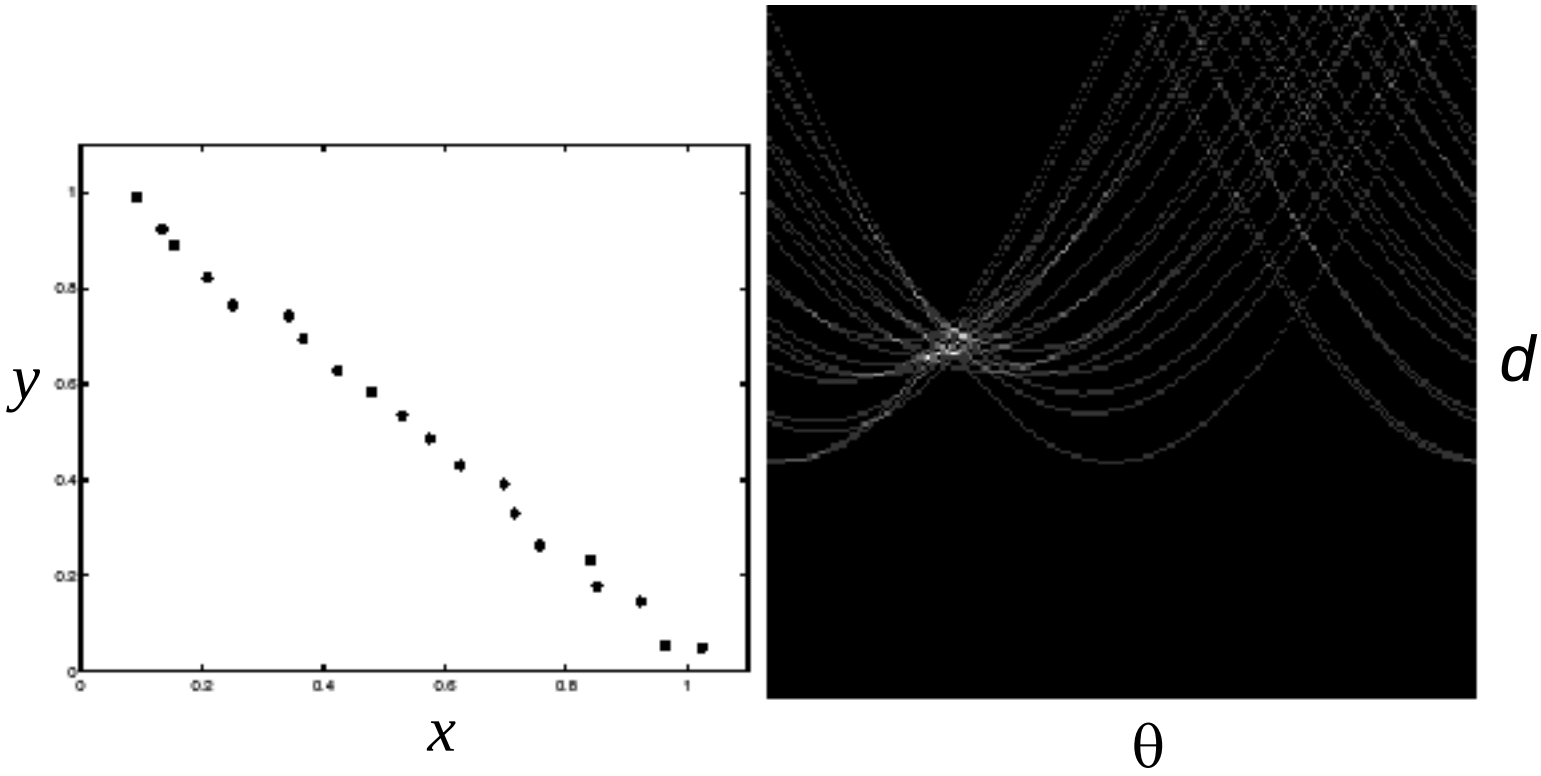
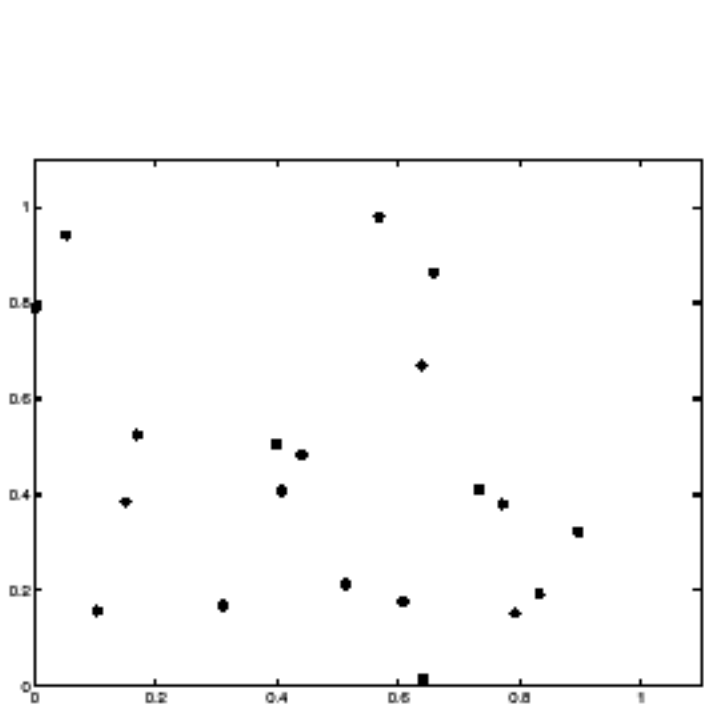


Image space edge coordinates

Votes

What difficulty does this present for an implementation?

Impact of noise on Hough



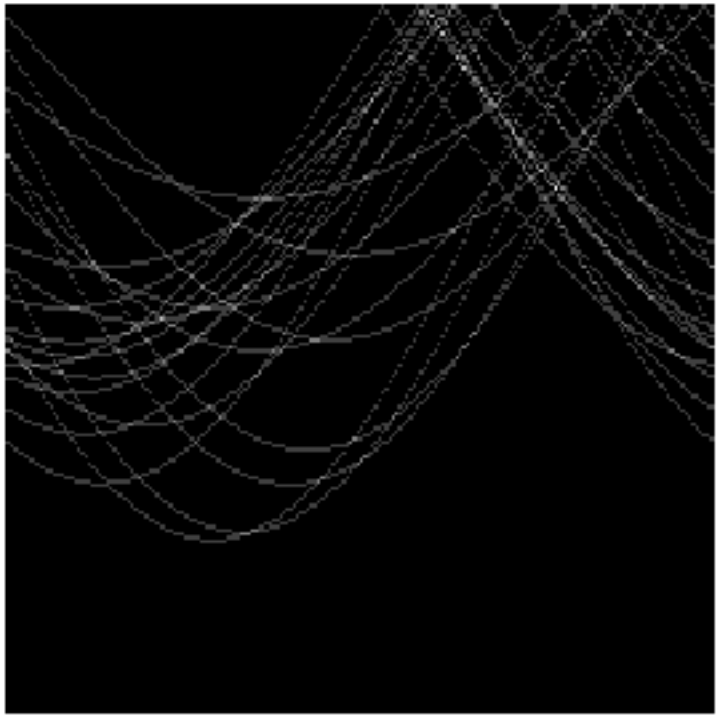
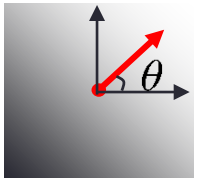


Image space edge coordinates

Votes

Here, everything appears to be “noise”, or random edge points, but we still see peaks in the vote space.

Extensions



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

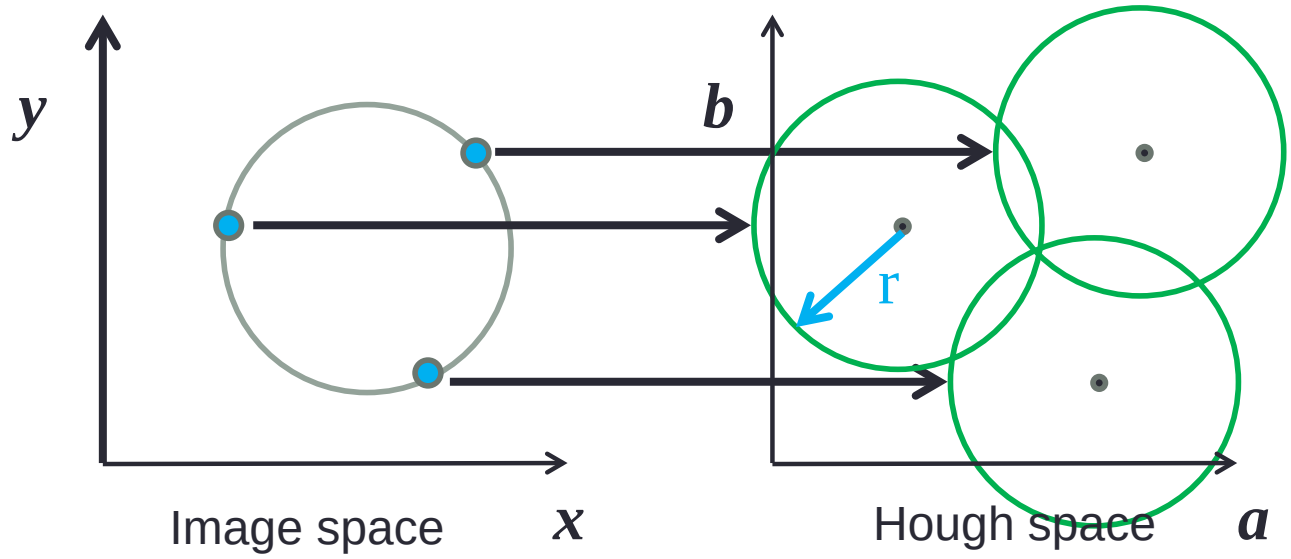
$$\theta = \tan^{-1}\left(\frac{\partial f}{\partial y} / \frac{\partial f}{\partial x}\right)$$

1. Initialize H[d, θ]=0
2. For each ***edge*** point in $E(x, y)$ in the image
   θ = gradient at (x,y)
   $d = x\cos\theta - y\sin\theta$
   $H[d, \theta] \mathrel{+}= 1$
3. Find the value(s) of (d, θ) where H[d, θ] is maximum
4. The detected line in the image is given by $d = x\cos\theta - y\sin\theta$

- Extension 2
  - give more votes for stronger edges
- Extension 3
  - change the sampling of (d, θ) to give more/less resolution
- Extension 4
  - The same procedure can be used with circles, squares, or any other shape

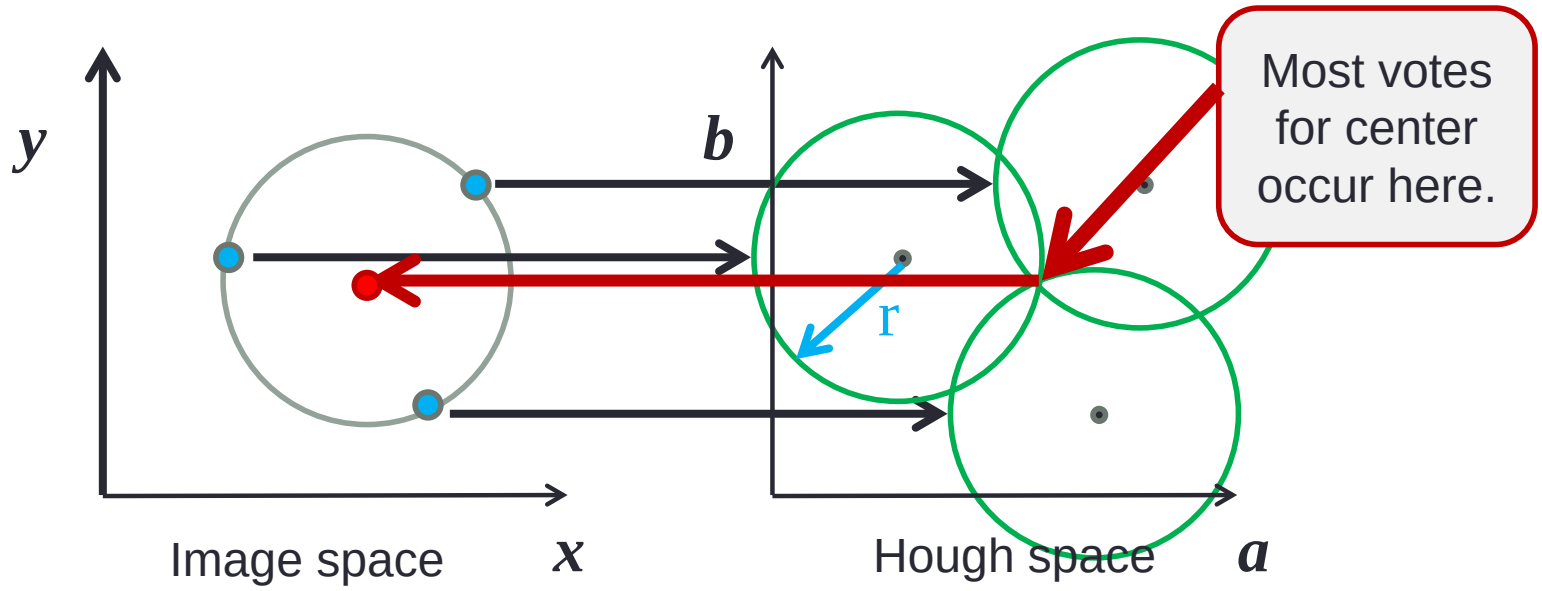
Hough transform for circles

- Circle: center (a,b) and radius r $(x_i - a)^2 + (y_i - b)^2 = r^2$
- For a fixed radius r, unknown gradient direction:



Hough transform for circles

- Circle: center (a,b) and radius r $(x_i - a)^2 + (y_i - b)^2 = r^2$
- For a fixed radius r, unknown gradient direction:



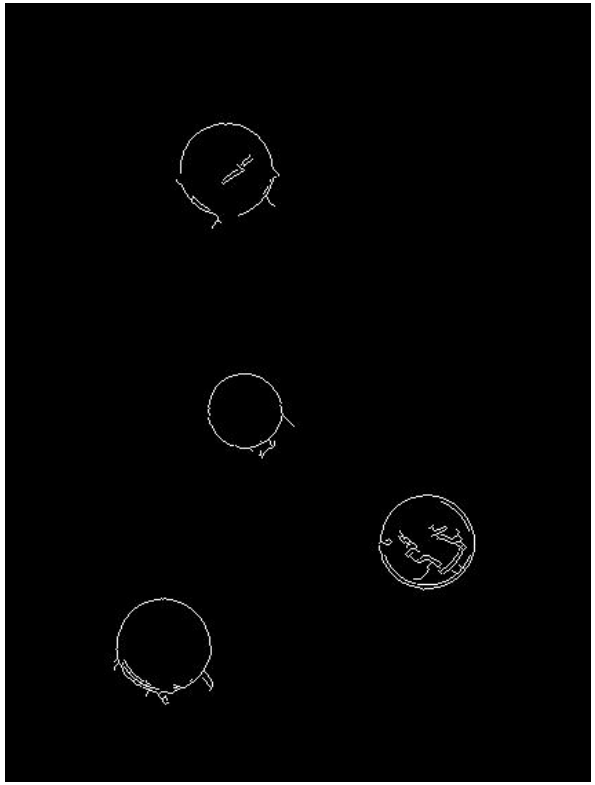
Example: detecting circles with Hough

Crosshair indicates results of Hough transform, bounding box found via motion differencing.

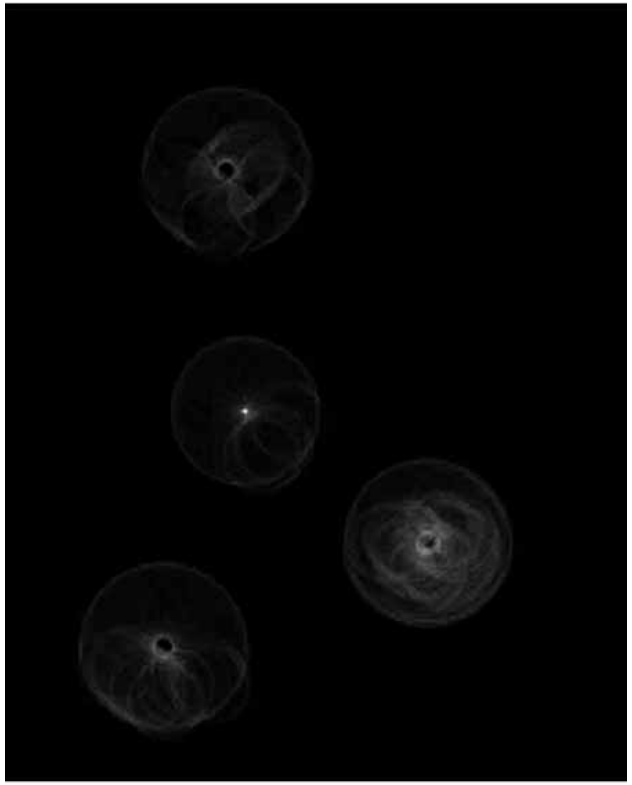
Example: detecting circles with Hough

Original

Edges



Votes: Penny



Note: a different Hough transform (with separate accumulators) was used for each circle radius (quarters vs. penny).

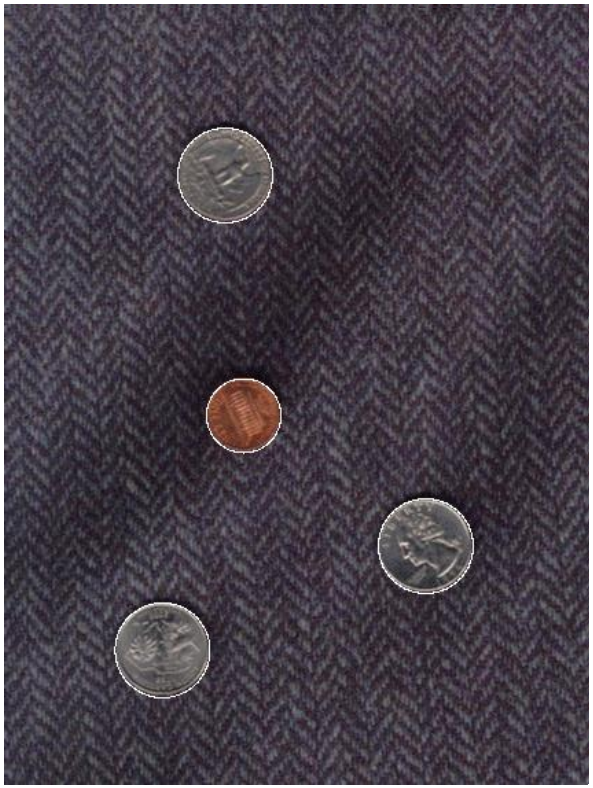
Example: detecting circles with Hough

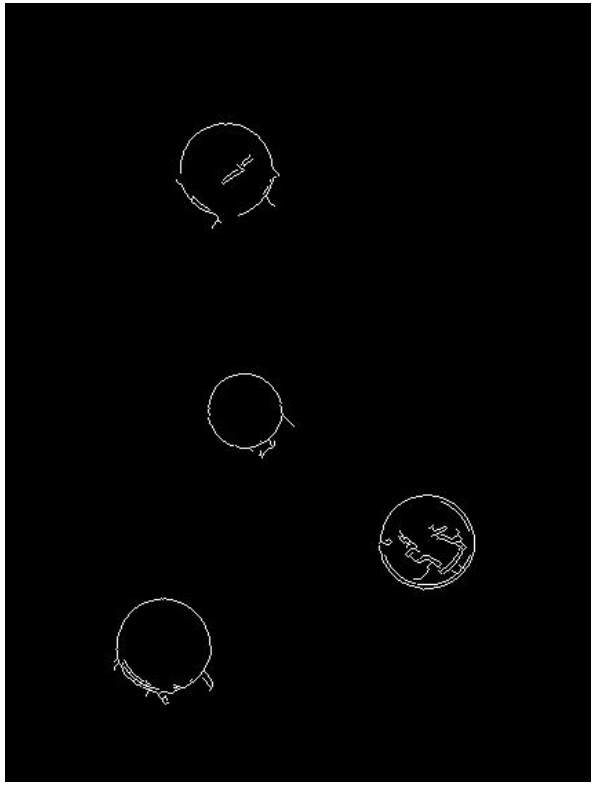
Original

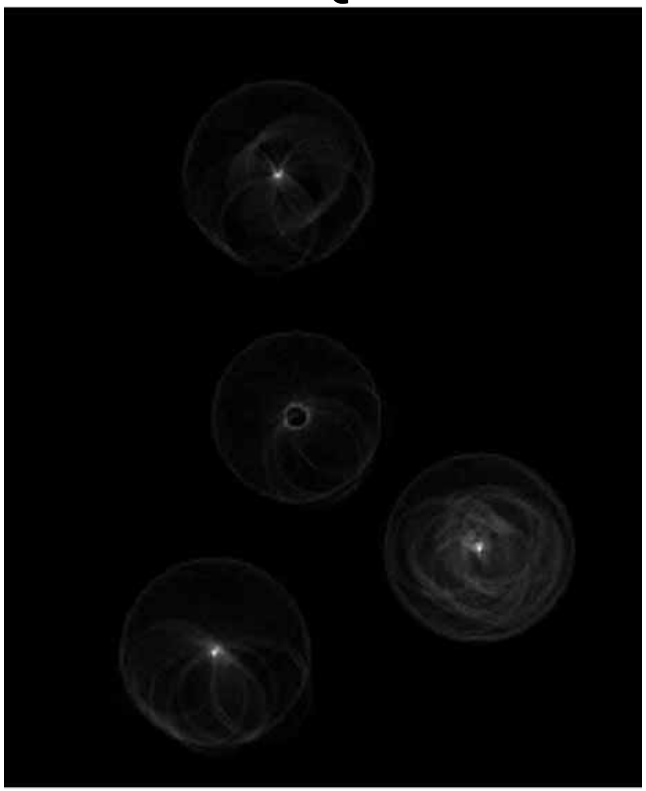
Edges

Votes: Quarter

Combined detections







Coin finding sample images from: Vivek Kwatra

Hough transform for circles

- Circle: center (a,b) and radius r $(x_i - a)^2 + (y_i - b)^2 = r^2$
- For ***unknown*** radius r, no gradient:

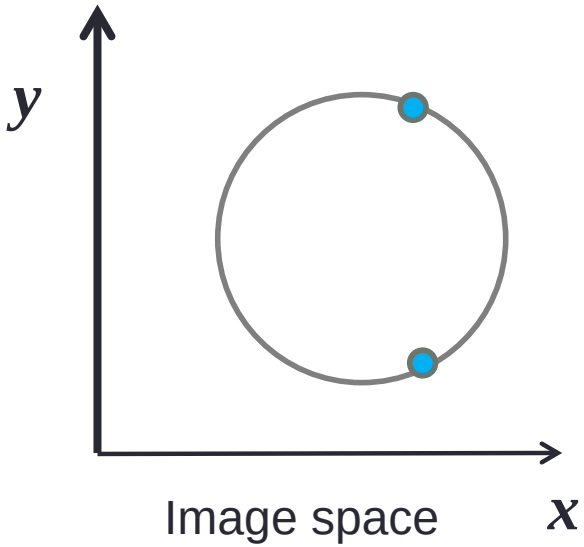
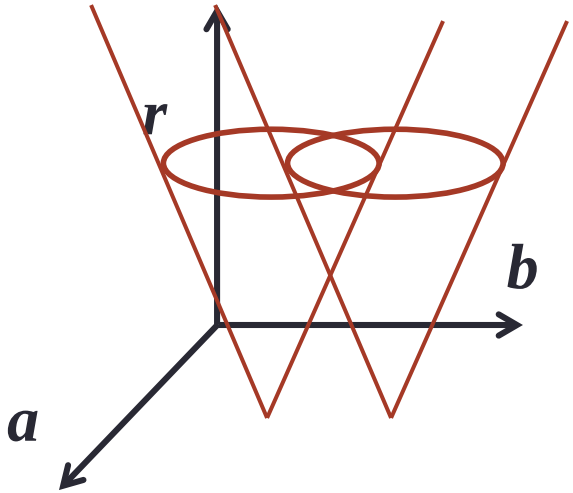


Image space



Hough space

Hough transform for circles

- Circle: center (a,b) and radius r $(x_i - a)^2 + (y_i - b)^2 = r^2$
- For ***unknown*** radius r, with ***gradient***:

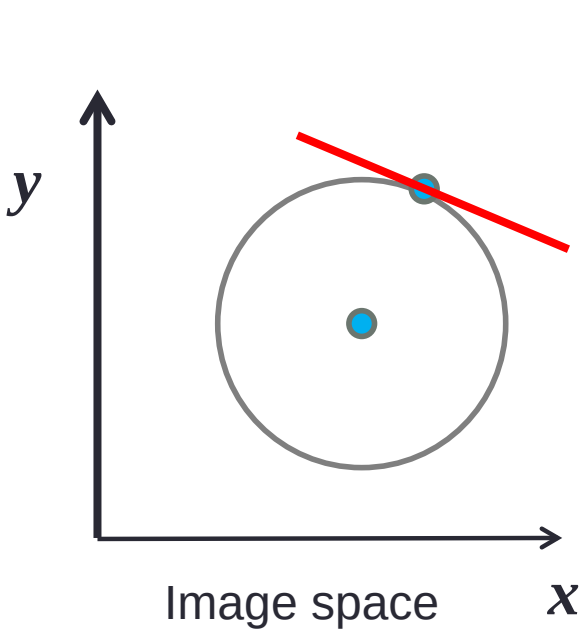
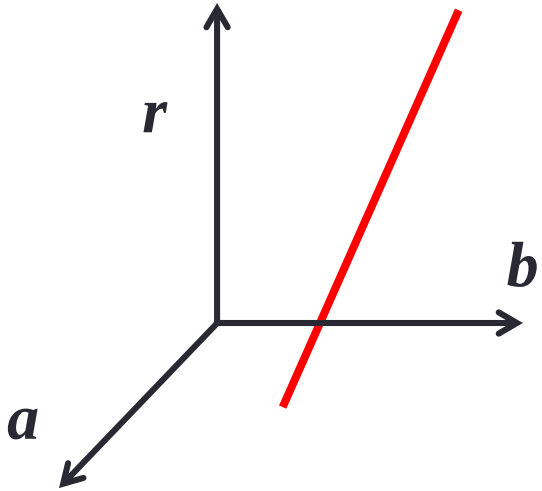


Image space



Hough space

Hough transform for circles

1. For every edge pixel (x,y) :
2. For each possible radius value r:
3. For each possible gradient direction θ:
   %% or use estimated gradient
4. a = x – r cos(θ)
5. b = y + r sin(θ)
6. H[a,b,r] += 1
7. end
8. end
9. end

Hough transform for circles

For every edge pixel (x,y) :

  For each possible radius value r:

    For each possible gradient direction θ:

      %% or use estimated gradient

    $a = x - r\cos(\theta)$

    $b = y + r\sin(\theta)$

    H[a,b,r] += 1

  end

end

Voting: practical tips

- Minimize irrelevant tokens first (take edge points with significant gradient magnitude)
- 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
- Vote for neighbors, also (smoothing in accumulator array)
- Utilize direction of edge to reduce free parameters by 1
- To read back which points voted for “winning” peaks, keep tags on the votes.

Hough transform: pros and cons

- Pros
  - All points are processed independently, so can cope with occlusion
  - Some robustness to noise: noise points unlikely to contribute consistently to any single bin
  - Can detect multiple instances of a model in a single pass

- Cons
  - Complexity of search time increases exponentially with the number of model parameters
  - Non-target shapes can produce spurious peaks in parameter space
  - Quantization: hard to pick a good grid size

Generalized Hough Transform

- **Non-analytic** models
  - *Parameters express variation in pose or scale of fixed but arbitrary shape (that was then)*
- **Visual code-word** based features
  - *Not edges but detected templates learned from models (this is "now")*

Generalized Hough transform

- What if want to detect arbitrary shapes defined by boundary points and a reference point?

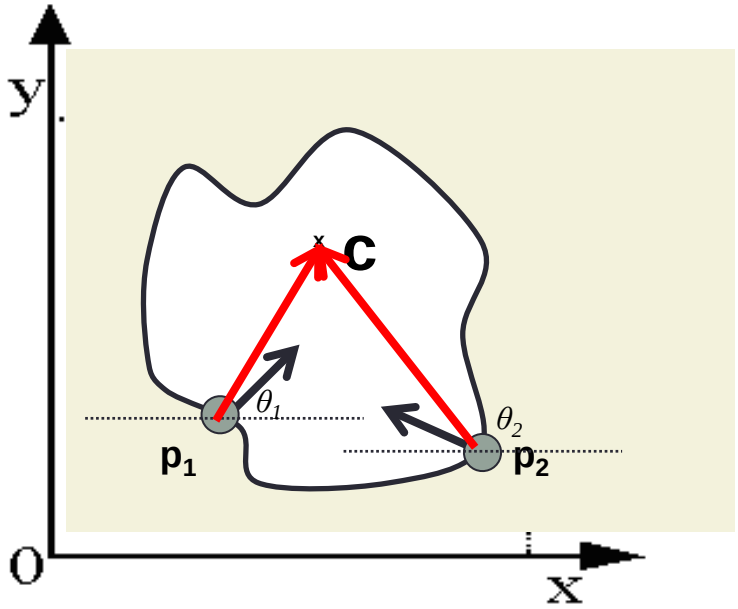


Image space

At each boundary point, compute displacement vector: $\mathbf{r = c - p_i}$.

For a given model shape: store these vectors in a table indexed by gradient orientation θ.

[Dana H. Ballard, Generalizing the Hough Transform to Detect Arbitrary Shapes, 1980]

Generalized Hough transform

- Recognition:

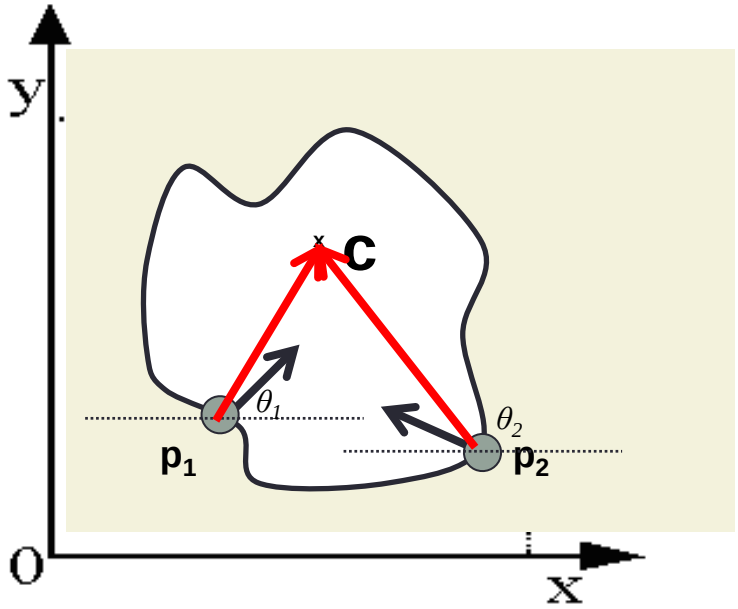
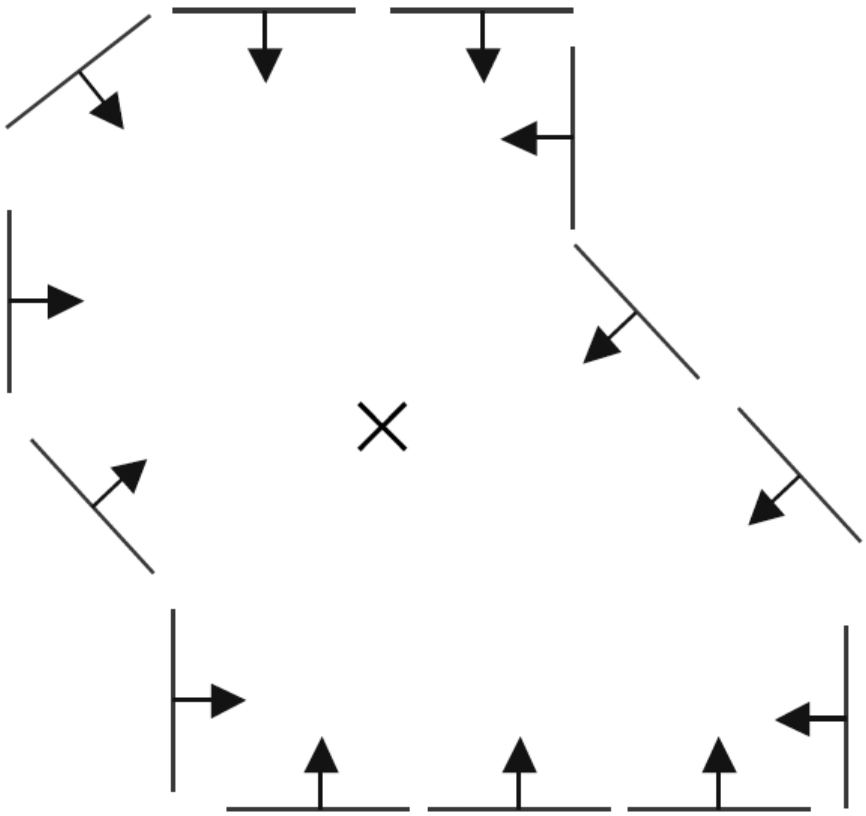


Image space

1. At each boundary point, measure the gradient angle θ
2. Look up all displacements in θ displacement table.
3. Vote for a center at each displacement.

[Dana H. Ballard, Generalizing the Hough Transform to Detect Arbitrary Shapes, 1980]

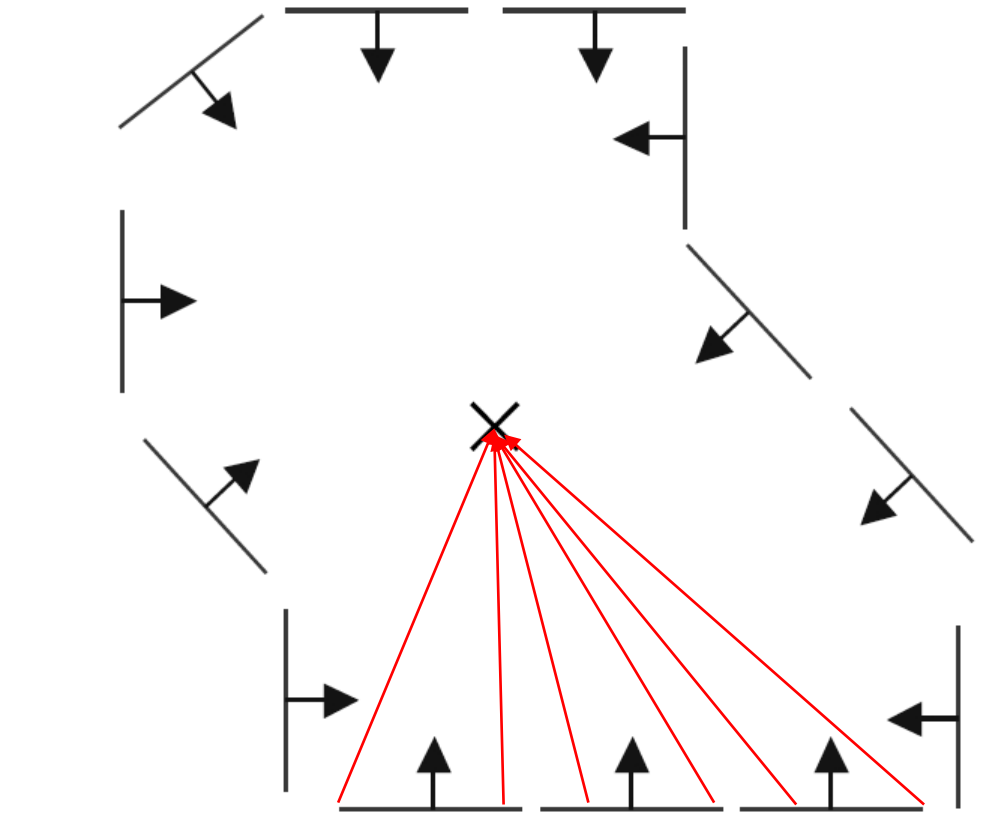
Example (known scale and orientation)



Source: L. Lazebnik

Example

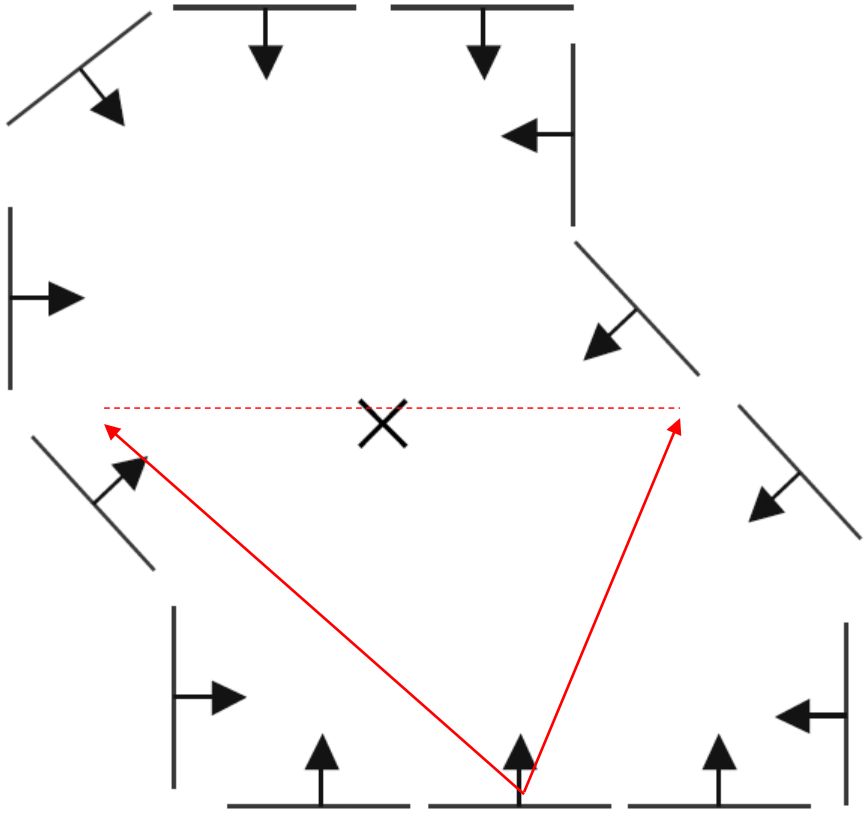
Looking at the bottom horizontal boundary points (all the same θ), the set of displacements ranges over all the red vectors.



displacement vectors for model points

Example

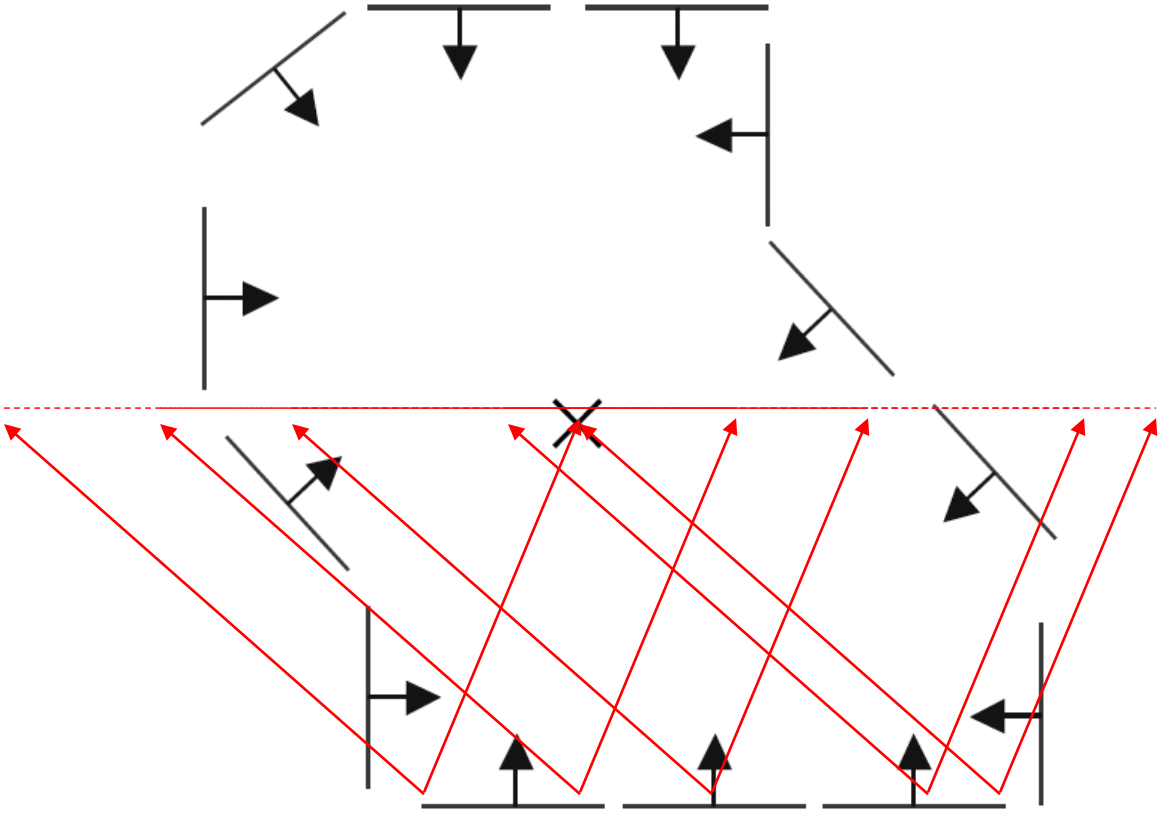
At recognition, each bottom horizontal element votes for all those displacements.



range of voting locations for test point

Example

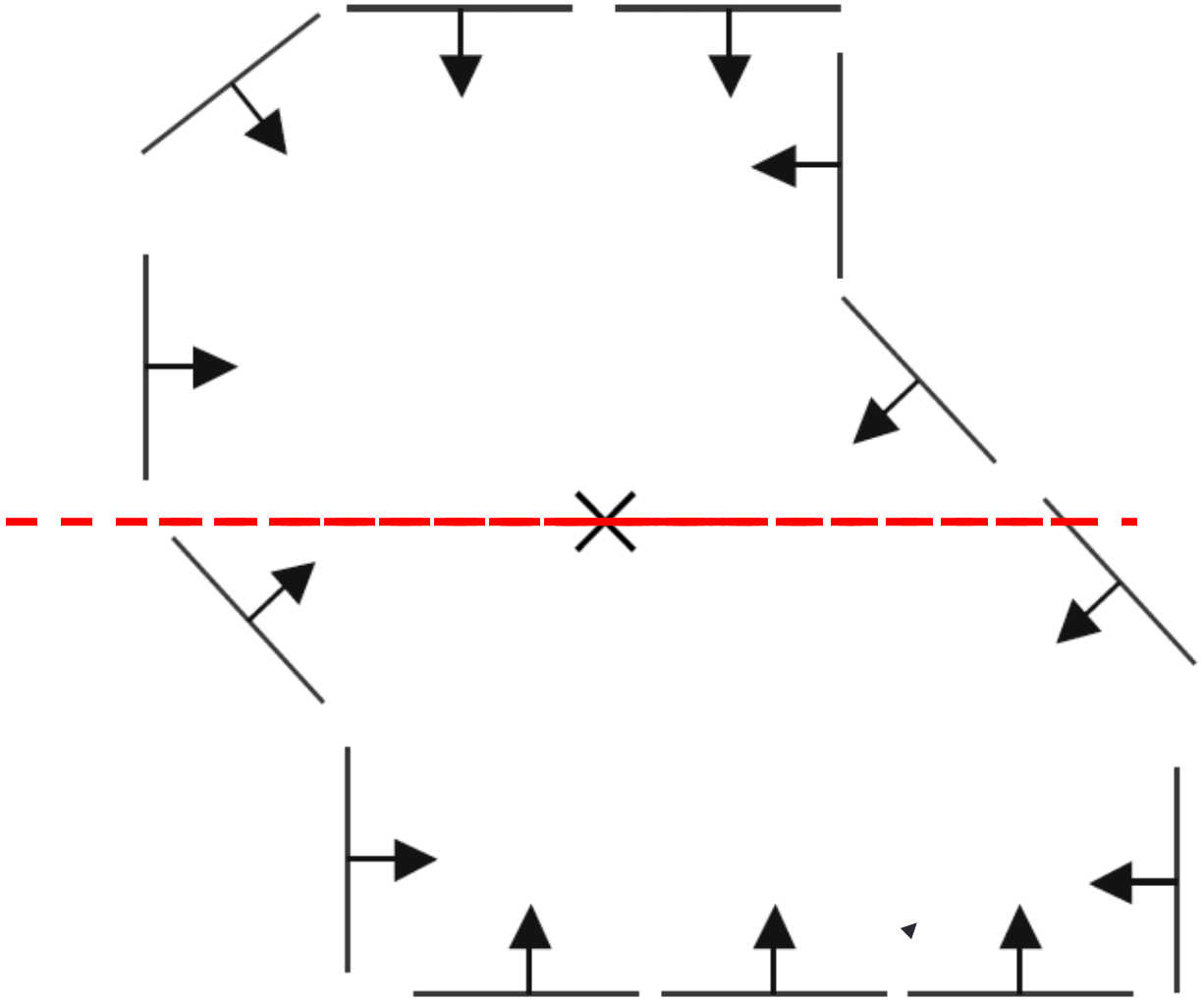
At recognition, each bottom horizontal element votes for all those displacements.



range of voting locations for test point

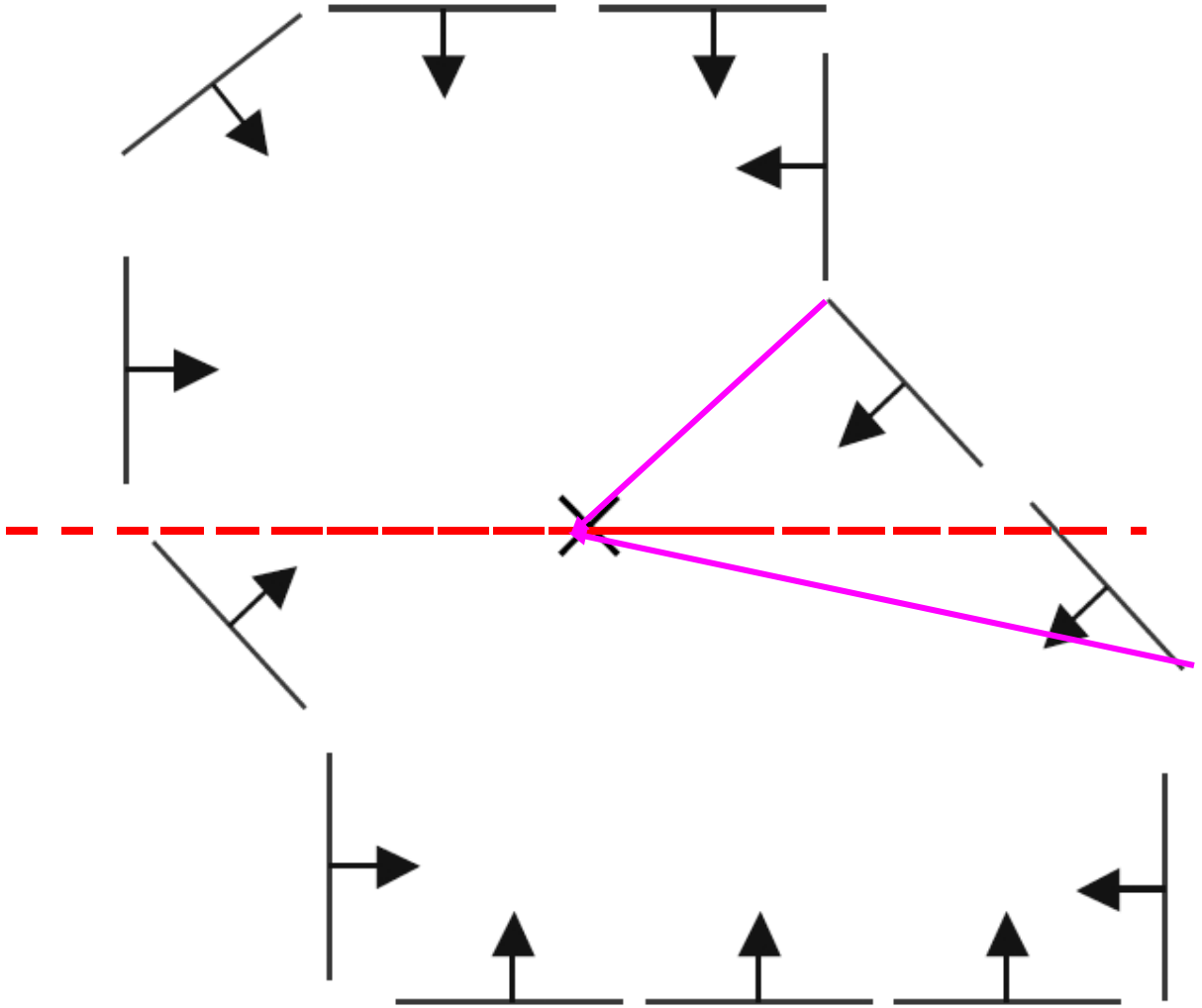
Example

At recognition, each bottom horizontal element votes for all those displacements.



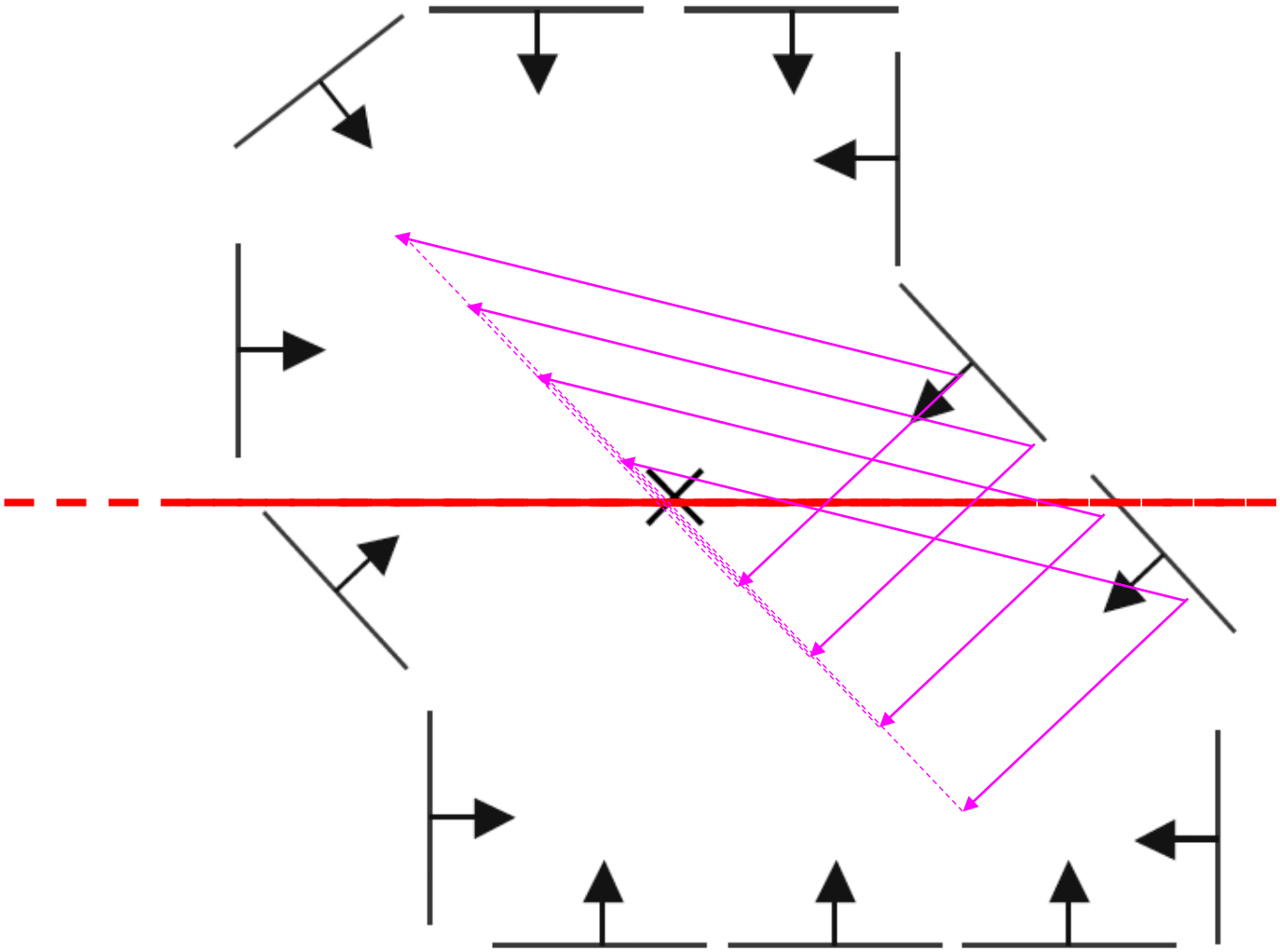
Example

Now do for the leftward pointing diagonals.



Example

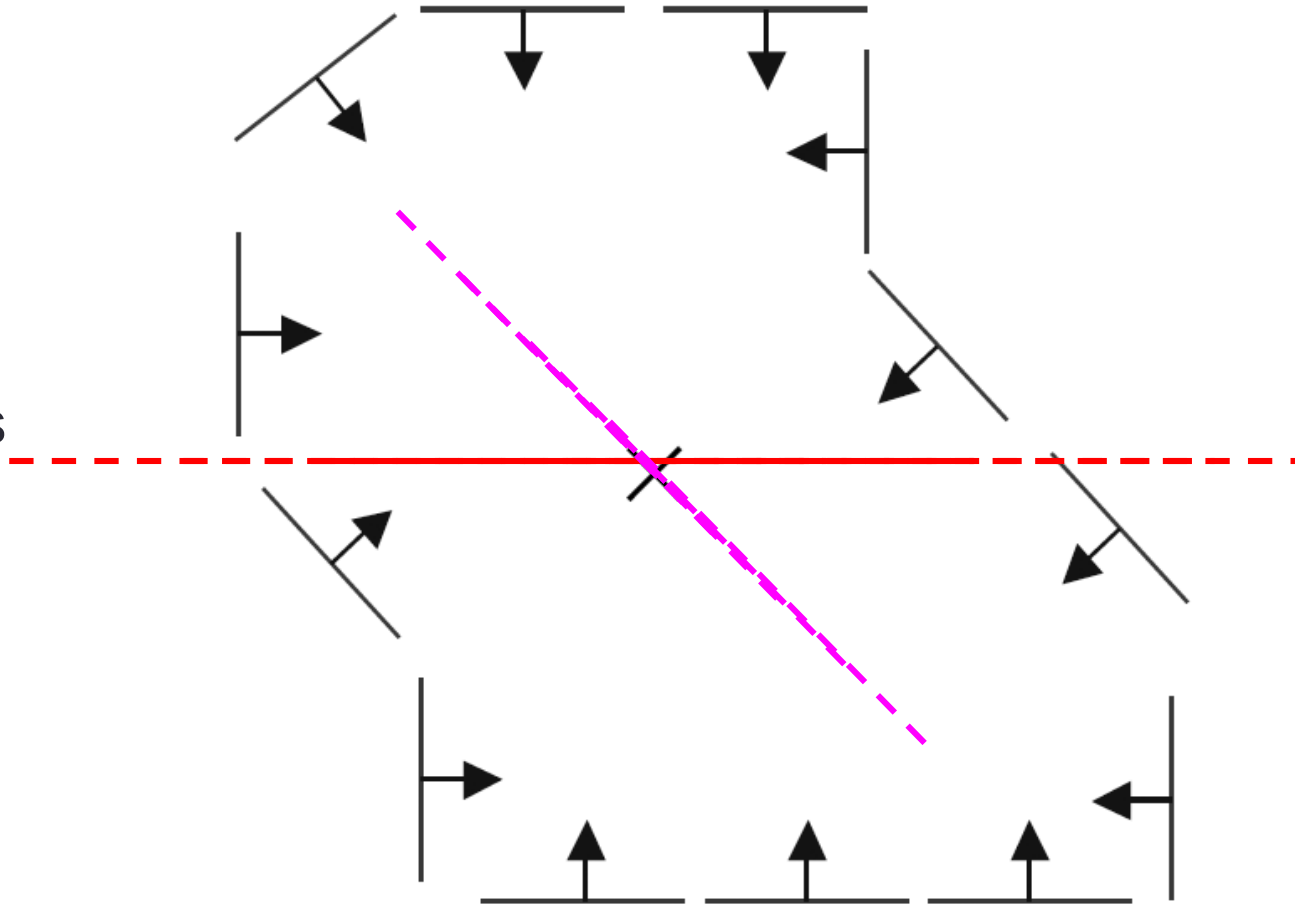
Now do for the leftward pointing diagonals.



Example

Now do for the leftward pointing diagonals.

And the center is found.



Generalized Hough transform

If orientation is known:

1. For each edge point

   Compute gradient direction θ

   Retrieve displacement vectors r to vote for reference point.

2. Peak in this Hough space (X,Y) is reference point with most supporting edges

Generalized Hough transform

If orientation is unknown:

For each edge point

For each possible master θ*

Compute gradient direction θ

New θ’ = θ - θ*

For θ’ retrieve displacement vectors r to vote for reference point.

Peak in this Hough space (now X,Y, θ*) is reference point with most supporting edges

[Dana H. Ballard, Generalizing the Hough Transform to Detect Arbitrary Shapes, 1980]

Generalized Hough transform

If scale S is unknown:

For each edge point

For each possible master scale S:

Compute gradient direction θ

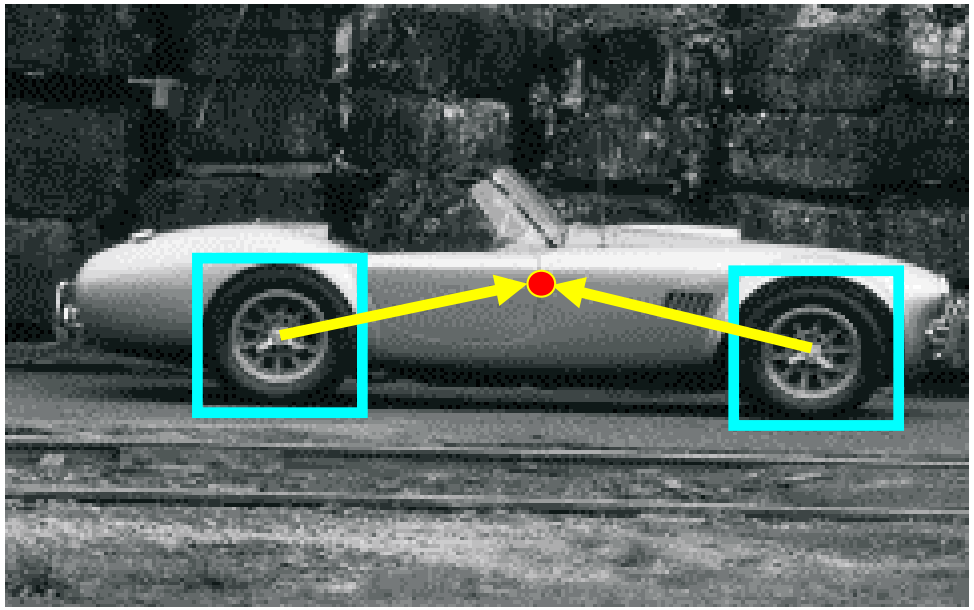
For θ’ retrieve displacement vectors r

Vote r scaled by S for reference point.

Peak in this Hough space (now X,Y, S) is reference point with most supporting edges

Application in recognition

- Instead of indexing displacements by gradient orientation, index by “visual codeword”



training image

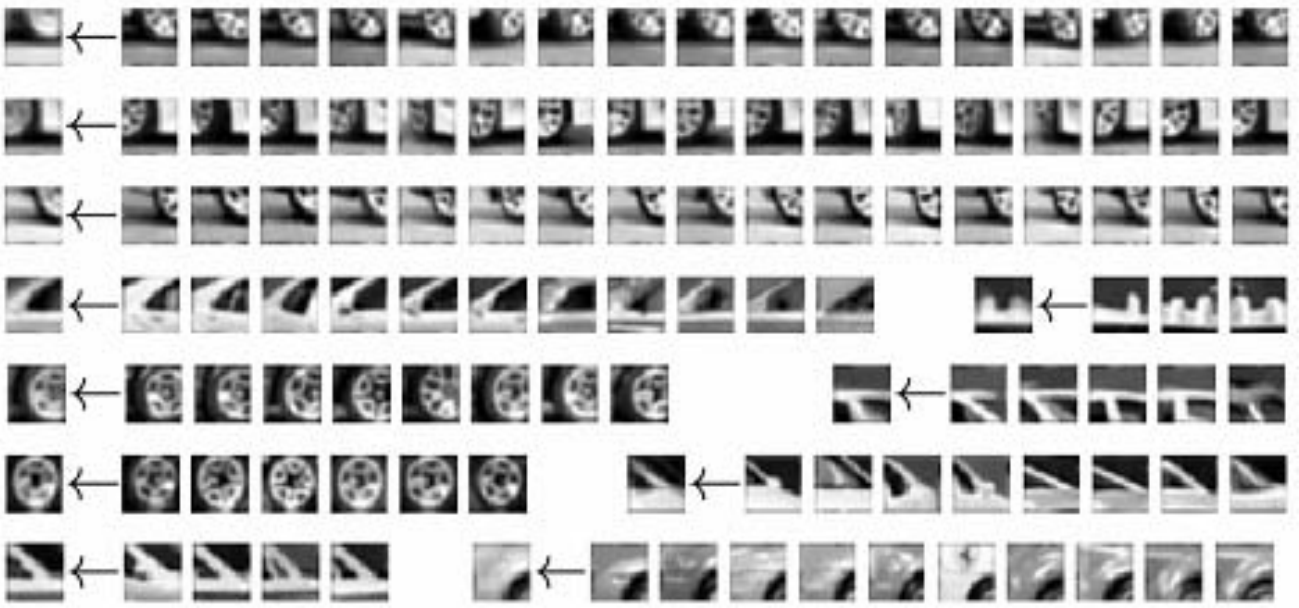
visual codeword with displacement vectors

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

Source: S. Lazebnik

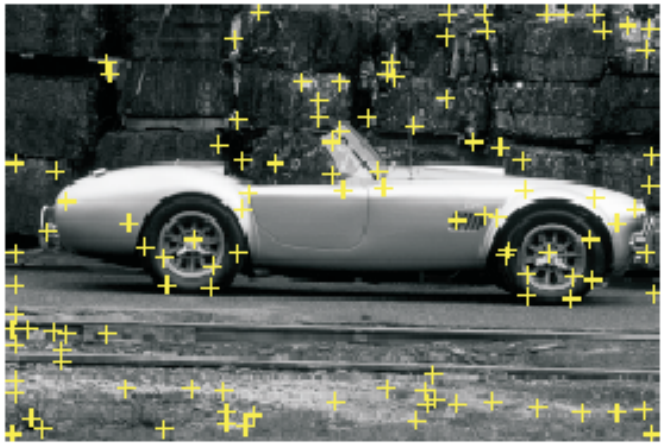
Training: Visual code-words

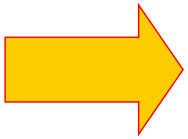
1. Build codebook of patches around extracted interest points using clustering (more on this later in the course)

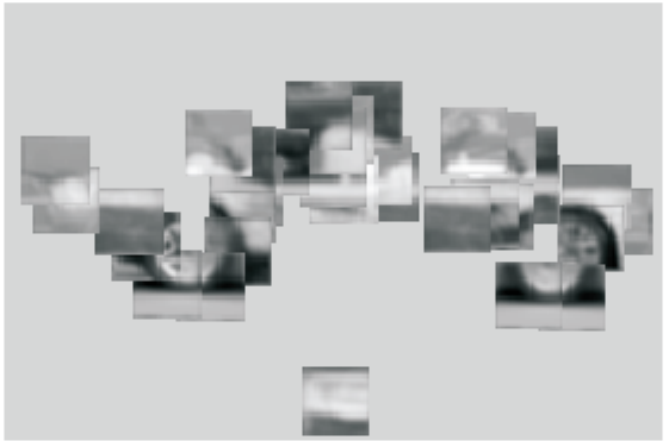


Training: Interest points

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

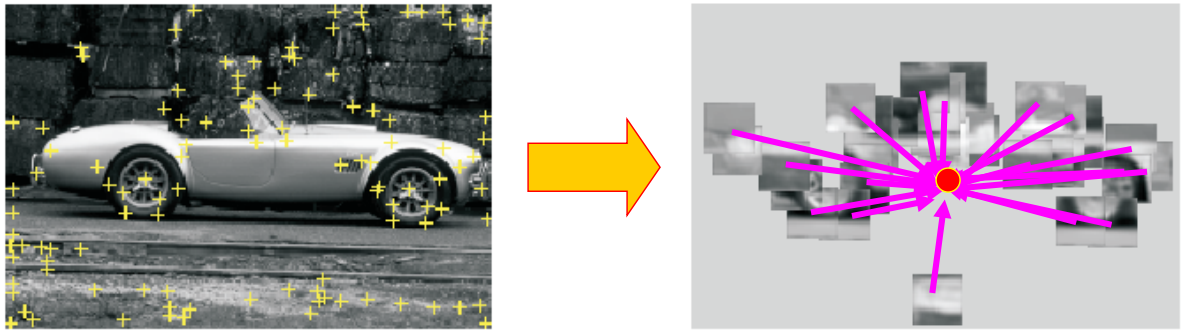






Training: Displacements

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 displacements relative to object center



Application in recognition

- Instead of indexing displacements by gradient orientation, index by “visual codeword”

test image

Summary

- Fitting problems require finding any supporting evidence for a model, even within clutter and missing features.
  - associate features with an explicit model
- Voting approaches, such as the Hough transform, make it possible to find likely model parameters without searching all combinations of features.
  - Hough transform approach for lines, circles, …, arbitrary shapes defined by a set of boundary points, recognition from patches.
- Labor Day Weekend is over. Time for the problem set!