Syllabus Notes

- PDF's updated
- Today we will hopefully get to some of §16. It is worth reading this chapter.
- For those going on in vision, this chapter is a must!
- Book does segmentation, then lines. We will do lines than segmentation.
- Probability reference (may need login "me", pw, "read4fun"):

  http://www.cs.arizona.edu/people/kobus/teaching/reading/statistical_modeling/ua_cs_only/probability.pdf

Fitting

- Choose a parametric object/some objects to represent a set of tokens
- Most interesting case is when criterion is not local
  - can’t tell whether a set of points lies on a line by looking only at each point and the next.

- Three main questions:
  - what object represents this set of tokens best?
  - which of several objects gets which token?
  - how many objects are there?

(you could read line for object here, or circle, or ellipse or...)

Fitting and the Hough Transform

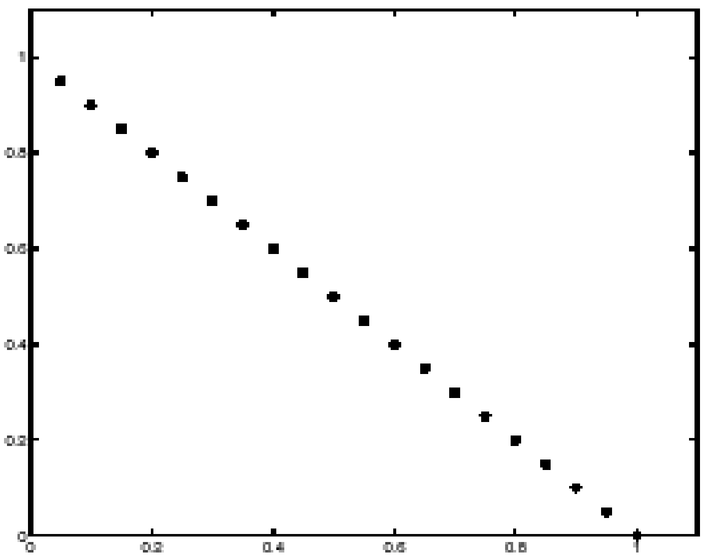
- Purports to answer all three questions
  - in practice, answer isn't usually all that much help
- We do for lines only
- A line is the set of points (x, y) such that

$$(\sin\theta)x + (\cos\theta)y + d = 0$$

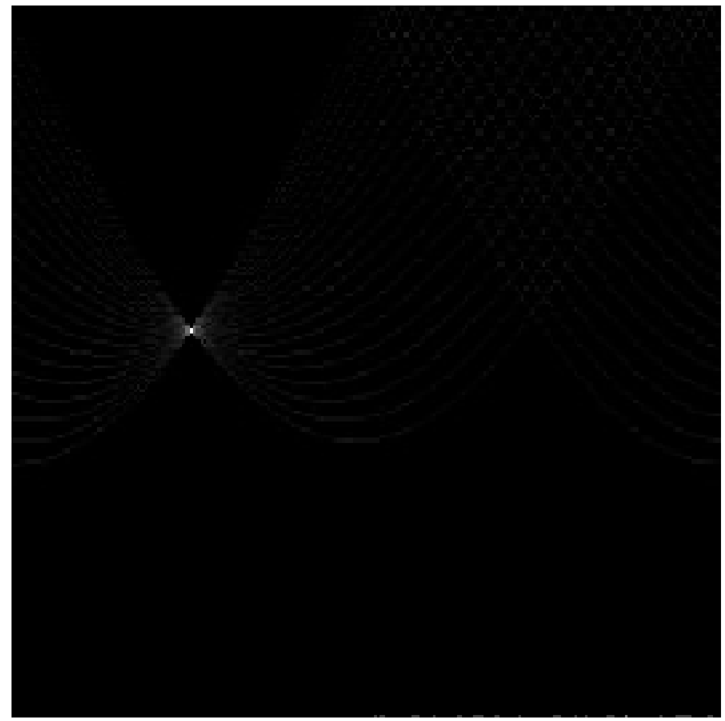
- Different choices of θ, $d>0$ give different lines
- For any (x, y) there is a one parameter family of lines through this point, given by

$$(\sin\theta)x + (\cos\theta)y + d = 0$$

- Each point gets to vote for each line in the family; if there is a line that has lots of votes, that should be the line passing through the points



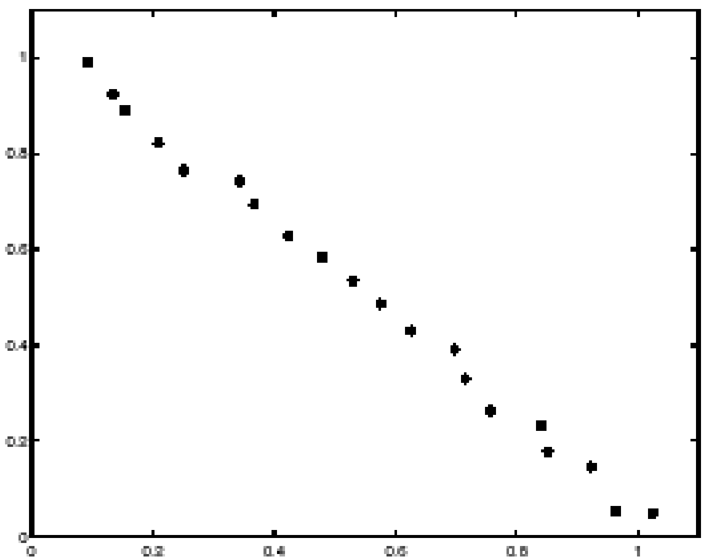
tokens



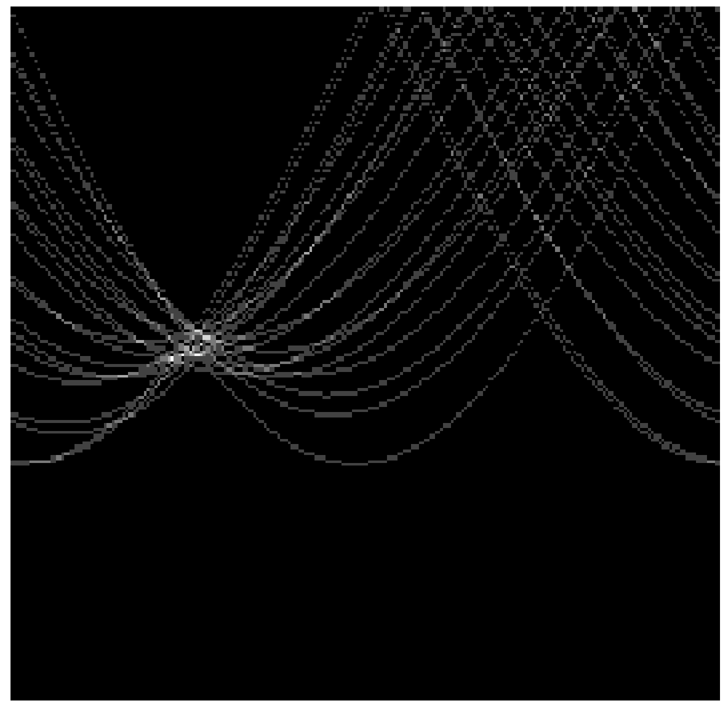
votes

Mechanics of the Hough transform

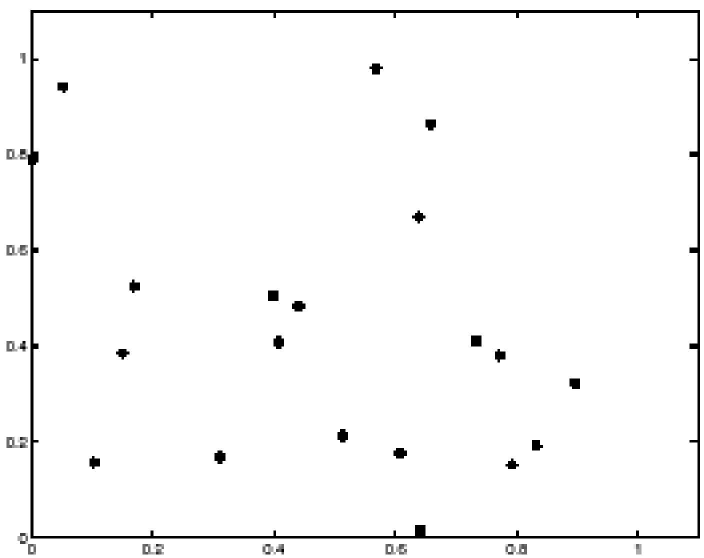
- Construct an array representing θ, d
- For each point, render the curve (θ, d) into this array, adding one at each cell
- Difficulties
  - how big should the cells be? (too big, and we cannot distinguish between quite different lines; too small, and noise causes lines to be missed)
- How many lines?
  - count the peaks in the Hough array
- Who belongs to which line?
  - tag the votes
- Hardly ever satisfactory in practice, because problems with noise and cell size defeat it



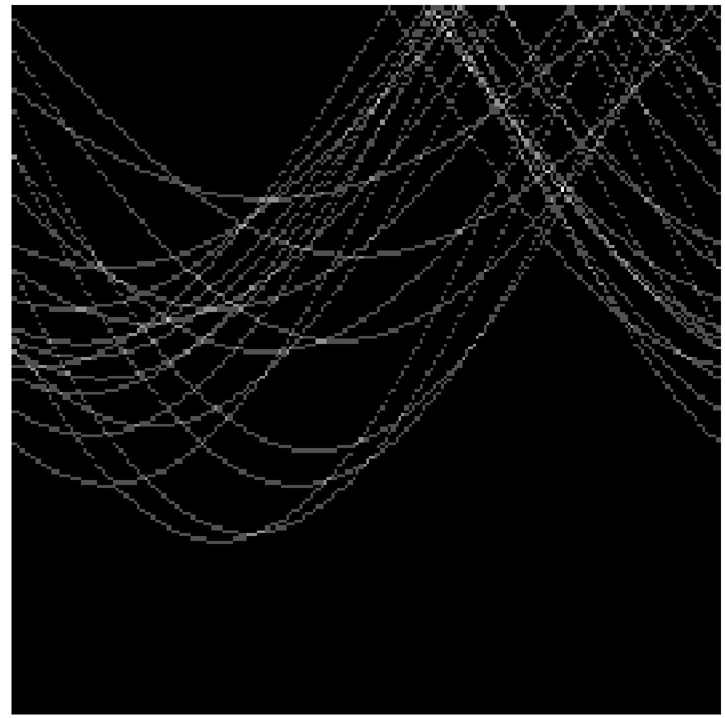
tokens



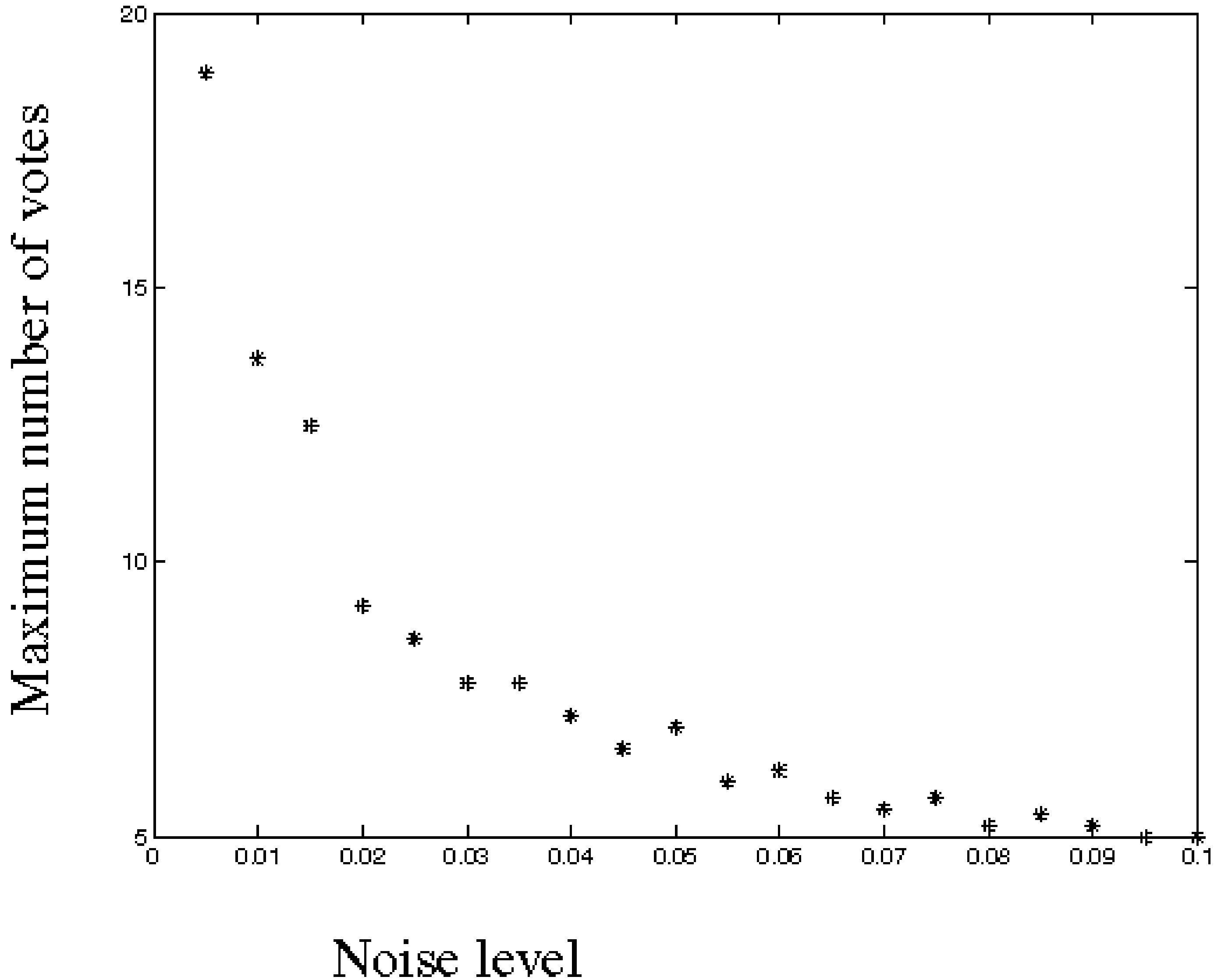
votes



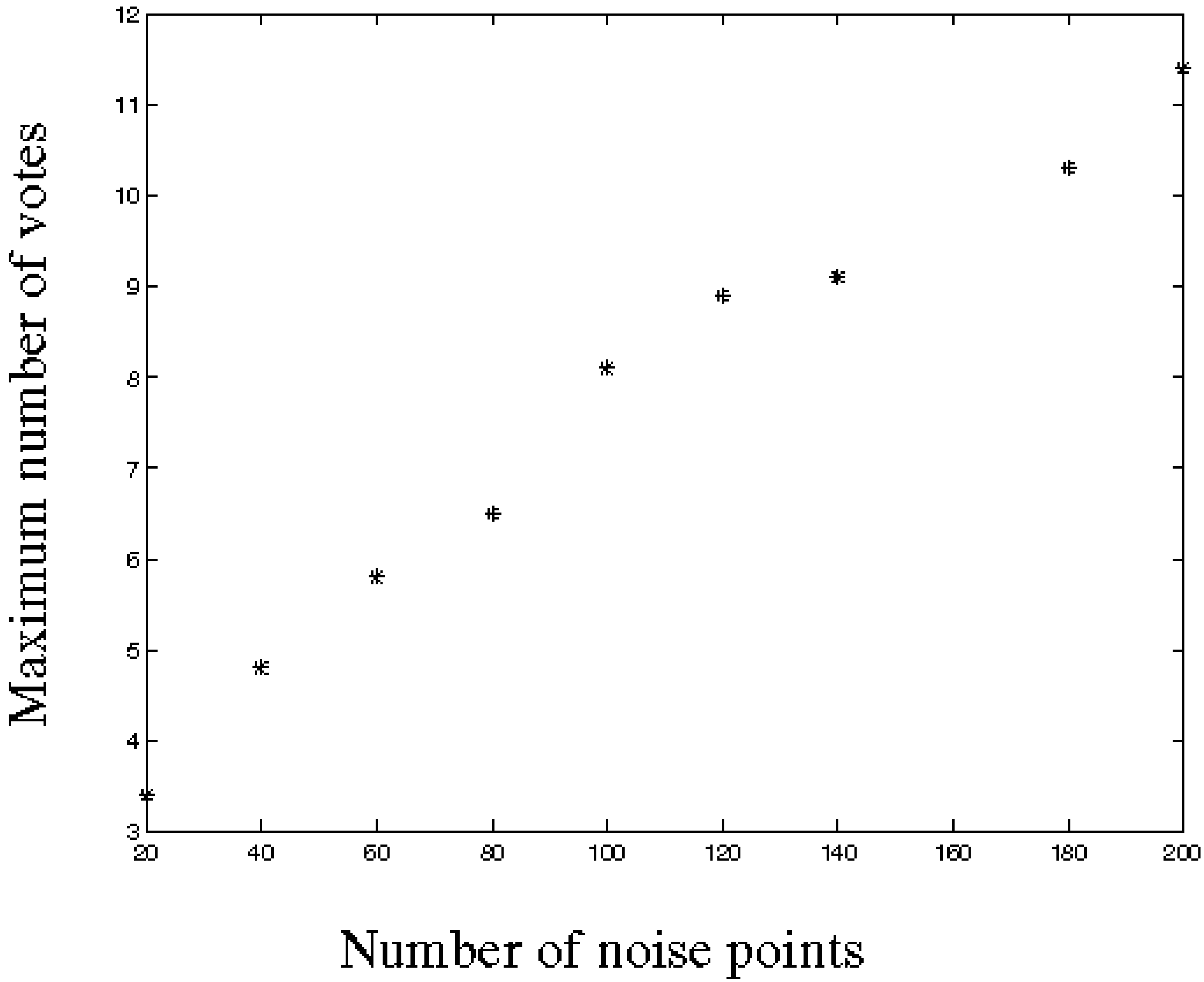
1
0.8
0.6
0.4
0.2
0
0
0.2
0.4
0.6
0.8
1



Votes for a real line of 20 points versus noise



Votes for a line in a picture that does not have one



Probabilistic Fitting

- Given a model with parameters θ
- Now consider some observations, **x**
- Suppose that the observations are independent
- So, given the model, the probability of observing the data is given by

$$P(\mathbf{x} \mid \theta) = \prod P(x_i \mid \theta)$$

- But what we really want is the probability of the model (parameters) given the data!

Probabilistic Fitting

- Bayes rule: $P(A \mid B) = P(B \mid A)P(A) / P(B)$

- So, $P(\theta \mid \mathbf{x}) = P(\mathbf{x} \mid \theta)P(\theta) / P(\mathbf{x})$

- $P(\theta)$ is the prior probability on the parameters (often taken to be uniform

- $P(\mathbf{x})$ is usually not of interest

- Often use $P(\theta \mid \mathbf{x}) \quad P(\mathbf{x} \mid \theta)$

Probabilistic Fitting

- Now the objective is to find the parameters θ such that this *likelihood* is maximum

- Note--this is the same as finding the parameters which minimize the **negative log likelihood** (very convenient if data is independent).

$$\underset{\theta}{\text{minimize}} \quad -\sum \log(P(x_i \mid \theta))$$

- Back to lines: au+bv+c=0 where a^2+b^2=1

- Algebraic fact: Distance squared from (x,y) to this line is (ax+by+c)^2

- **Generative model** for lines: Choose point on line, and then, with probability p(d), **normally distributed** (Gaussian) go a distance d from the line.

- Now the probability of an observed (x,y) is given by

$$P((x,y)\mid\theta) \propto \exp\left(-\frac{(ax+by+c)^2}{2\sigma^2}\right)$$

Now the probability of an observed (x,y) is given by

$$P((x,y)\mid\theta) \propto \exp\left(-\frac{(ax+by+c)^2}{2\sigma^2}\right)$$

The negative log is

$$\frac{(ax+by+c)^2}{2\sigma^2}$$

And the negative log likelihood of multiple observations is

$$\frac{1}{2\sigma^2}\sum_i (ax_i+by_i+c)^2$$

From the previous slide, we had that the negative log likelihood of multiple observations is given by

$$\frac{1}{2\sigma^2}\sum_i (ax_i + by_i + c)^2 \qquad \text{(where } \mathrm{a}^2 + \mathrm{b}^2 = 1\text{)}$$

We could solve this by considering derivatives, but this should be recognizable as homogeneous least squares

Thus we have shown that least squares is maximum likelihood estimation under normality (Gaussian) error statistics)!