Computer Vision
Grouping and Segmentation

Professor Hager
http://www.cs.jhu.edu/~hager

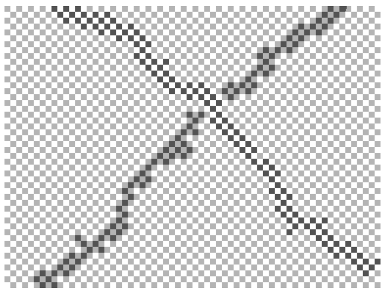
Grouping and Segmentation

- G&S appear to be one of the early processes in human vision
- They are a way of *organizing* image content into "semantically related" groups
- In some applications, segmentation is the crucial step (e.g. some types of aerial image interpretation).
- We'll just touch on a few simple approaches of image grouping and segmentation
- We'll also present a couple of useful geometric grouping methods.

CS 461, Copyright G.D. Hager
with slides shamelessly stolen from D. Forsyth

Grouping

- **Grouping** is the process of associating similar image features together

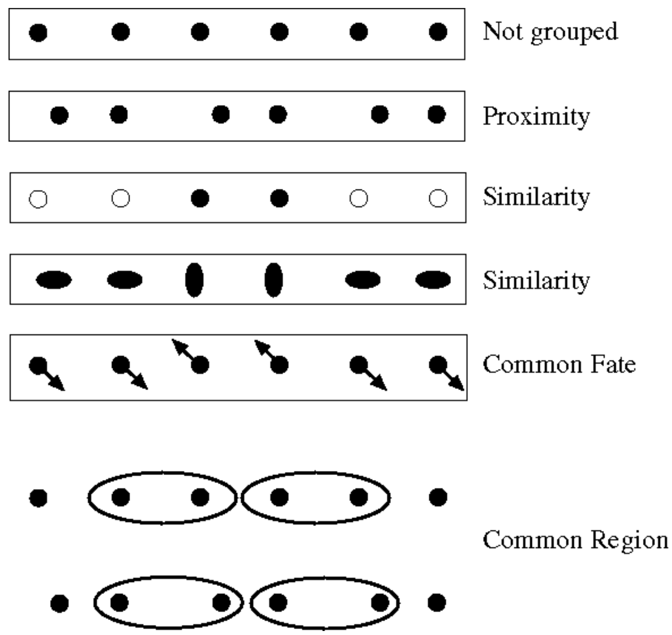


Grouping

- **Grouping** is the process of associating similar image features together

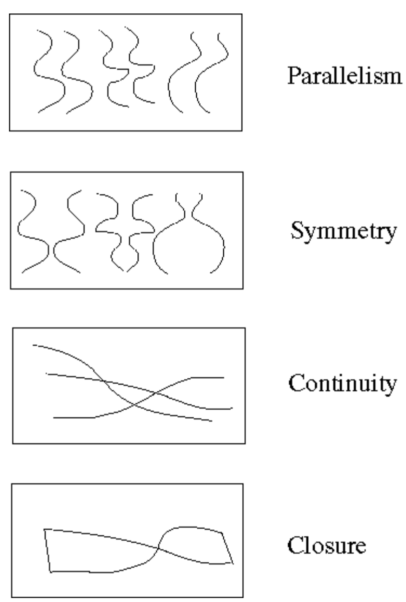
- The Gestalt School:

  – **Proximity**:tokens that are nearby tend to be grouped.
  – **Similarity:**similar tokens tend to be grouped together.
  – **Common fate**:tokens that have coherent motion tend to be grouped together.
  – **Common region:**tokens that lie inside the same closed region tend to be grouped together.
  – **Parallelism:**parallel curves or tokens tend to be grouped together.
  – **Closure:**tokens or curves that tend to lead to closed curves tend to be grouped together.
  – **Symmetry:**curves that lead to symmetric groups are grouped together.
  – **Continuity:**tokens that lead to “continuous ” (as in “joining up nicely ”, rather than in the formal sense): curves tend to be grouped.
  – **Familiar Conguration:**tokens that, when grouped, lead to a familiar object,tend to be grouped together



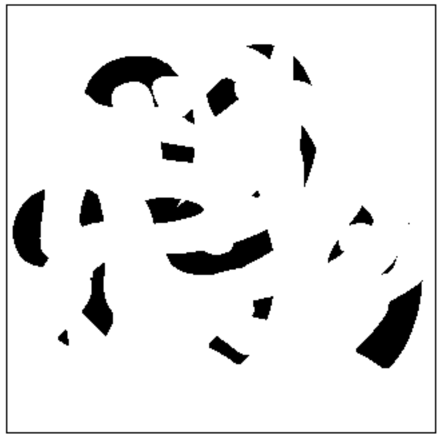
10/2/2003

CS 461, Copyright G.D. Hager
with slides shamelessly stolen from D. Forsyth



10/2/2003

CS 461, Copyright G.D. Hager
with slides shamelessly stolen from D. Forsyth



10/2/2003

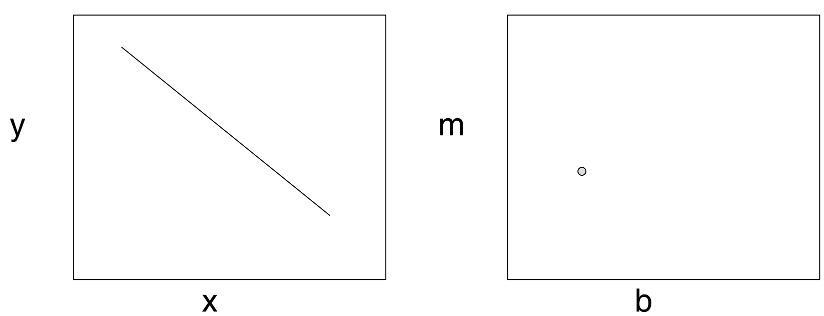
CS 461, Copyright G.D. Hager
with slides shamelessly stolen from D.
Forsyth

10/2/2003

CS 461, Copyright G.D. Hager
with slides shamelessly stolen from D.
Forsyth

The Hough Transform

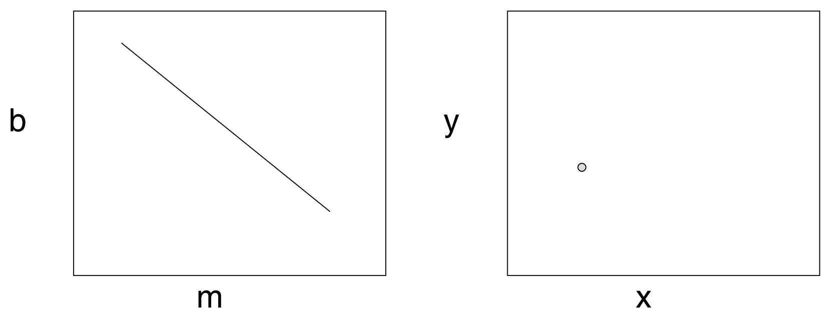
- How can we detect (group) extended line structures in images?



$$y = m x + b$$

The Hough Transform

- How can we detect extended line structures in images?



$$b = - x\, m + y$$

Hough Transform Idea

- Each edge point in an image is a constraint on line parameters:
  - constraint is a line
  - each unique point adds another constraint
- Algorithm:
  - Initialize a 2-D array of counters to zero.
  - For each edge point (x,y), increment any counter which contains a parameter point (b,m) satisfying $b = -x\, m + y$
  - Threshold counters
  - Group edgels that belong to above threshold counters into contours
  - Optional: Refit lines to get higher precision.

10/2/2003
CS 461, Copyright G.D. Hager
with slides shamelessly stolen from D. Forsyth

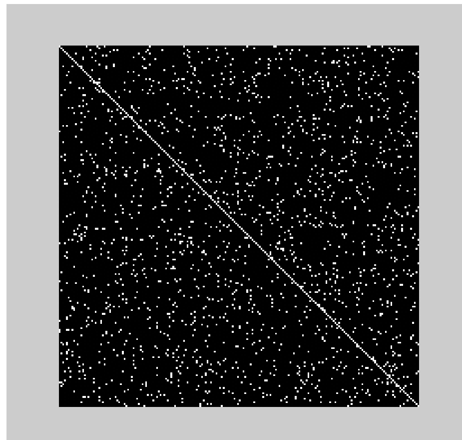
Hough Transform Variation

- Problem:
  - m and b are, in principle, unbounded
    - rewrite as sin(t) x + cos(t) y = d
    - For each discrete value of t from 0 to pi, increment counter with nominal d minimal distance from exact value
  - coarse grid leads to grouping of distinct lines
    - in post-fitting stage, re-hough at higher resolution
- Generalizations
  - Any linear in parameters model: e.g $a x^2 + b y^2 = 1$ can use the same algorithm
  - For f(x,a) = 0, choose any cell a_c s.t. $f(x,a_c) < t$ for some threshold t.
- Limitations:
  - the curse of dimensionality

RANSAC: A More General Method

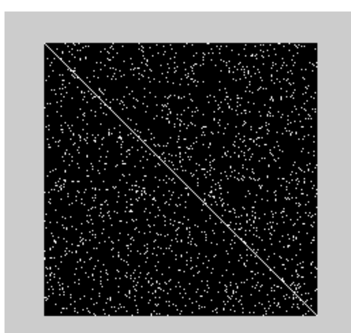
- The Hough transform has the nice feature that it provides a method for detecting geometric structure in clutter.
- The HT suffers from the curse of dimensionality.
- RANSAC (Random Sample Consensus) is a robust estimation technique that avoids (to some degree) the curse of dimensionality

RANSAC: General Idea



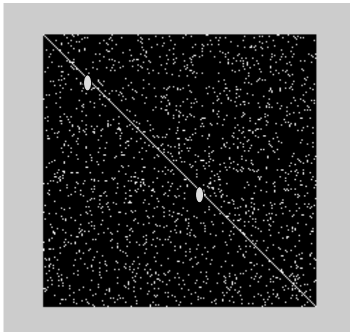
How to find the line in this?

RANSAC: General Idea



Basic observation: if we were to choose the "right" pair of points

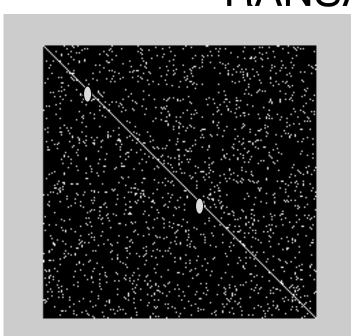
RANSAC: General Idea



Basic observation: if we were to choose the “right” pair of points then it would be fairly obvious that many other points would agree

Algorithm:
For some number of tries N
pick a pair of points
fit a line to these points
count the number of other pts that agree
Return: line parameters with best agreement

RANSAC: General Idea



Algorithm:
For some number of tries N
pick a pair of points
fit a line to these points
count the number of other pts that agree
Return: line parameters with best agreement

Questions:
1) How many points
2) What do we mean “agree”

RANSAC

- # of points is $N = \log(1-p)/\log(1-(1-\varepsilon)^s)$
  - assumes sampling is independent
  - p is the probability one of the samples is all inliers
  - s is sample size (2 for us)
  - ε is percentage of outliers

- Consensus:
  - Assume the points (inliners) are contaminated with Gaussian noise
  - Distance to the line is a χ^2_m model where (in our case) m =1
    - Choose a probability threshold from χ^2_1 giving a distance $t= 3.82\ \sigma^2$ for $\alpha = .95$
    - inliner if $\text{distance}^2 < t^2$
    - outlier if $\text{distance}^2 >= t^2$

Image Segmentation: A Brief Overview

- Segmentation
  - criteria
  - region group and counting

Segmentation: Definitions

- An affinity measure $d(R_1,R_2) \rightarrow \Re$ or a homogeneity measure m(R)
  - note possibly $d(R_1,R_2) = |m(R_1) - M(R_2)|$
- An threshold τ (could operate on distance or homogeneity)
- A region definition (e.g. square tiles)
- A neighborhood definition
  - 4 neighbors

```
  x
x 0 x
  x
```

  - 8 neighbors

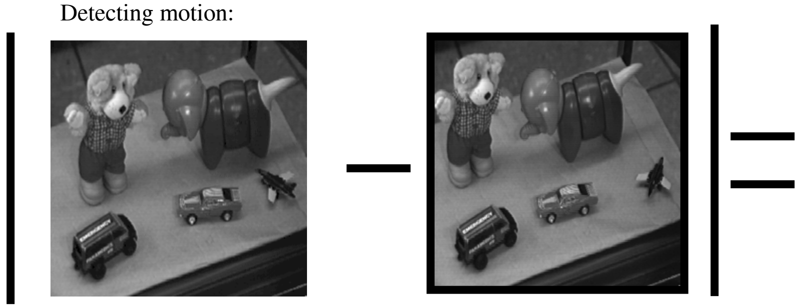
```
x x x
x 0 x
x x x
```


Simple Thresholding

- Choose an image criterion c
- Compute a binary image by b(i,j) = 1 if c(I(i,j)) > t; 0 otherwise
- Perform "cleanup operations" (image morphology)
- Perform grouping
  - Compute connected components and/or statistics thereof

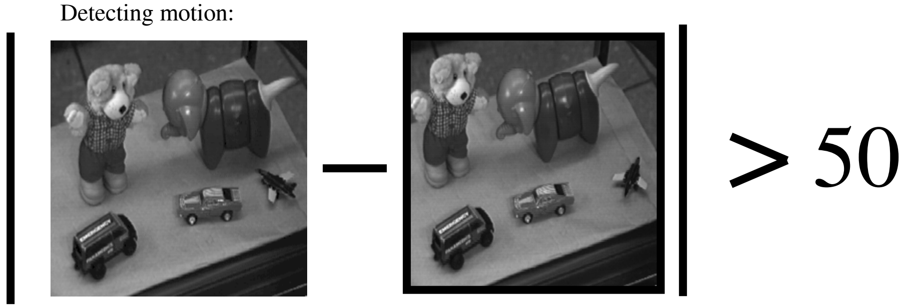
An Example: Motion

Detecting motion:

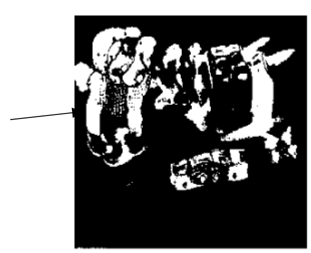


Thresholded Motion

Detecting motion:



Candidate areas for motion



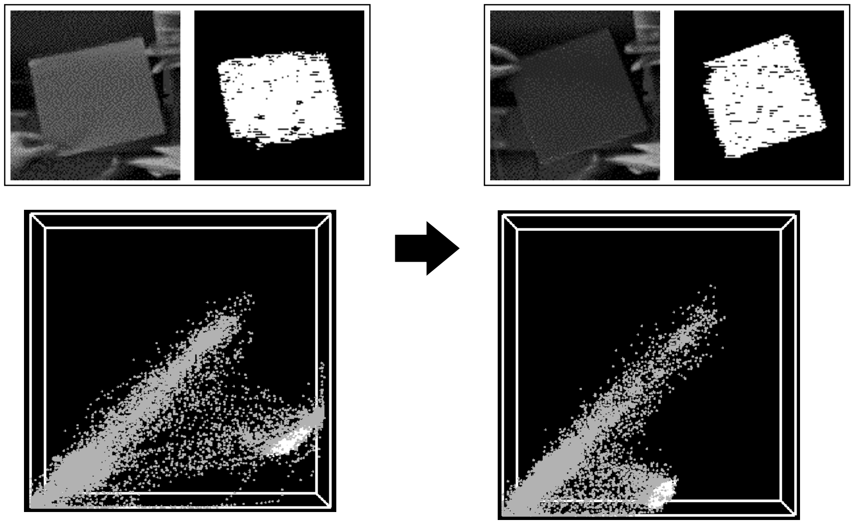
A Closer Look

Color: A Second Example

- Color representation
  - DRM [Klinker et al., 1990]: if P is Lambertian, has *matte* line and *highlight* line
  - User selects *matte* pixels in *R*
  - Compute first and second order statistics of cluster
  - Decompose ellipsoid $(\mathbf{S}, \mathbf{R}^T, \mathbf{T})$ of variance of matte cluster
  - Color similarity $\gamma(\mathbf{I}(x, y))$ is defined by Mahalanobis distance

$$|\mathbf{S}^{-1}\mathbf{R}^T(\mathbf{I}(x, y) - \mathbf{T})|$$

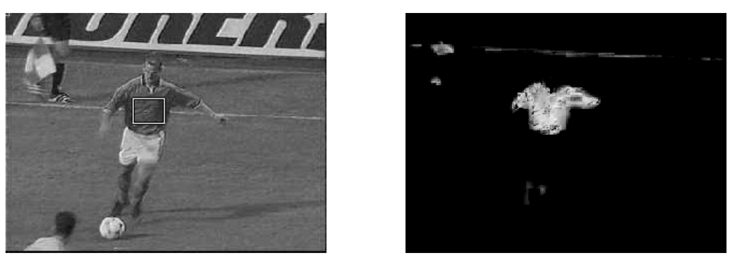
Homogeneous Color Region: Photometry



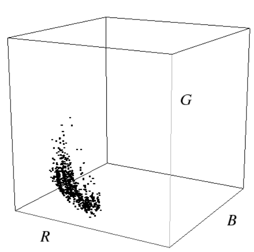
10/2/2003

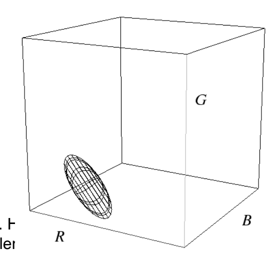
CS 461, Copyright G.D. Hager
with slides shamelessly stolen from D. Forsyth

Homogeneous Region: Photometry



Sample





PCA-fitted ellipsoid

10/2/2003

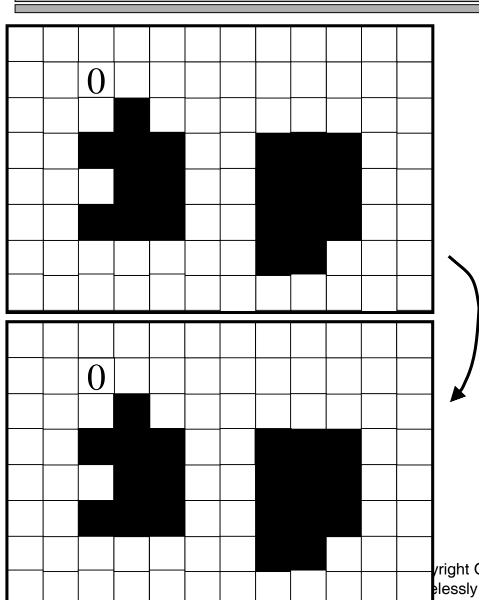
D. H
toler
Forsyth

Binary Image Processing

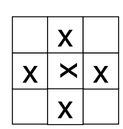
After thresholding an image, we want to know something about the regions found ...

- [x] How many objects are in the image?
- [ ] Where are the distinct "object" components?
- [ ] "Cleaning up" a binary image?

Connected Component Labeling

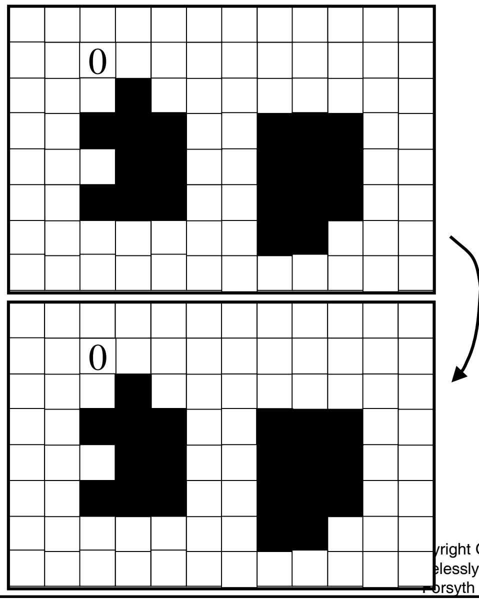


Goal: Label contiguous areas of a segmented image with unique labels



4 neighbors vs. 8 neighbors

Connected Component Labeling



Algorithm

1. Image is A. Let A = -A;
2. Start in upper left and work L to R, Top to Bottom, looking for an unprocessed (-1) pixel.
3. When one is found, change its label to the next unused integer. Relabel all of that pixel's unprocessed neighbors and their neighbors recursively.
4. When there are no more unprocessed neighbors, resume searching at step 2 -- but do so where you left off the last time.

Morphological Operators summarized

Let S_t be the *translation* of a set of pixels S by t.

$$S_t = \{ x + t \mid x \in S \}$$

The *dilation* of a binary image A by a mask S is then

$$A \oplus S = \bigcup_{b \in S} A_b$$

The *erosion* of a binary image A by a mask S is

$$A \ominus S = \{ x \mid x + b \in A, \forall b \in S \}$$

The *closing* of a binary image A by S is

$$A \bullet S = (A \oplus S) \ominus S$$

The *opening* of a binary image A by S is

$$A \circ S = (A \ominus S) \oplus S$$

The “Poor Man’s” Closing

- Note that median (or more generally any order statistic) filtering is one way of achieving similar effects. On binary images this can be also implemented using the averaging filter

Limitations of Thresholding

- A uniform threshold may not apply across the image
- It measures the uniformity of regions (in some sense), but doesn’t examine the inter-relationship between regions.
- Local “disturbances” can break up nominally consistent regions
  - note hysteresis thresholding is one solution to this!!
- Later in the course we’ll return to talk about methods around this