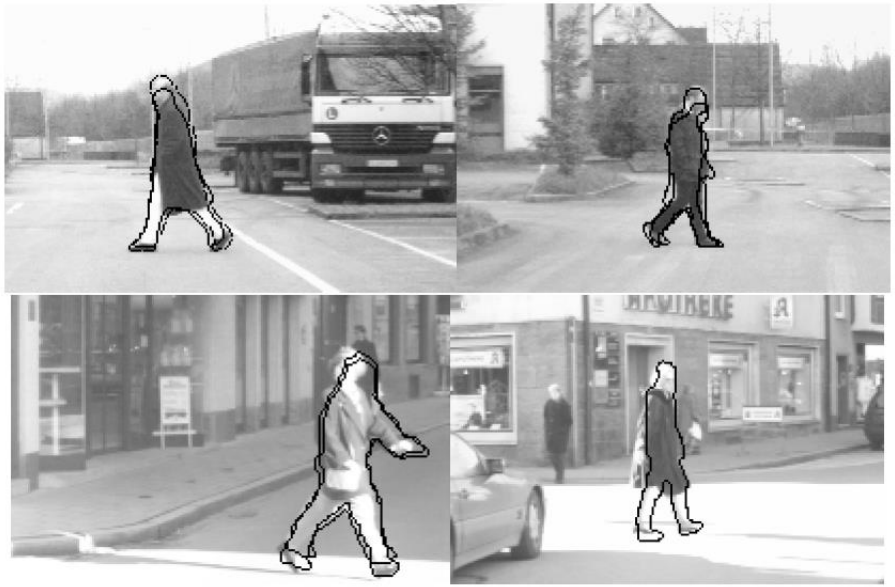
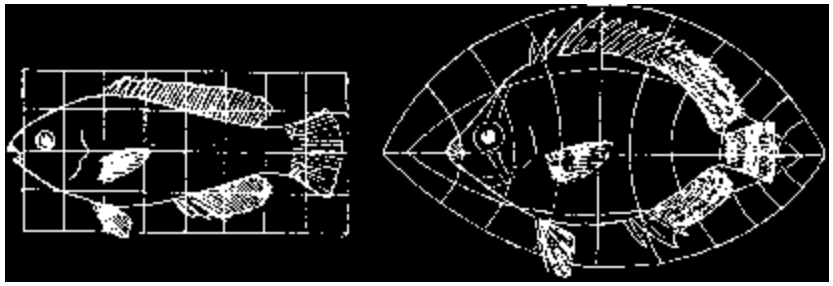
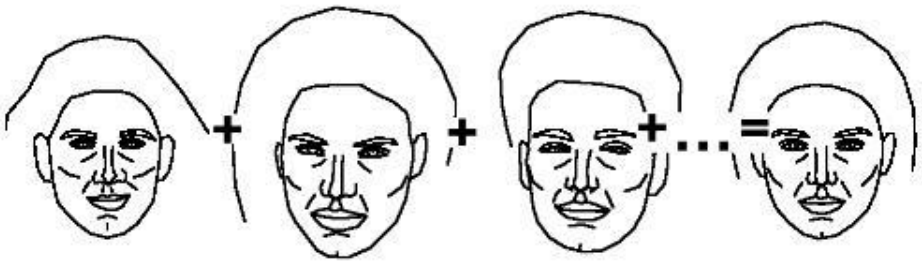
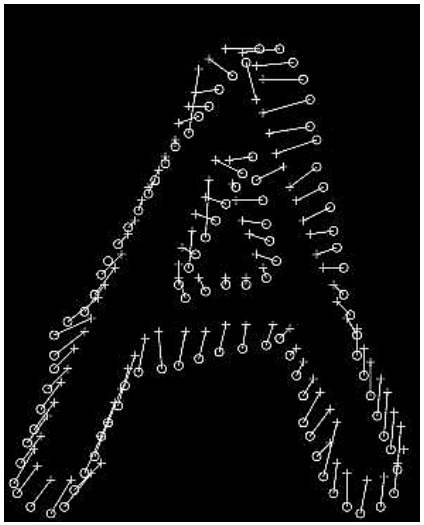




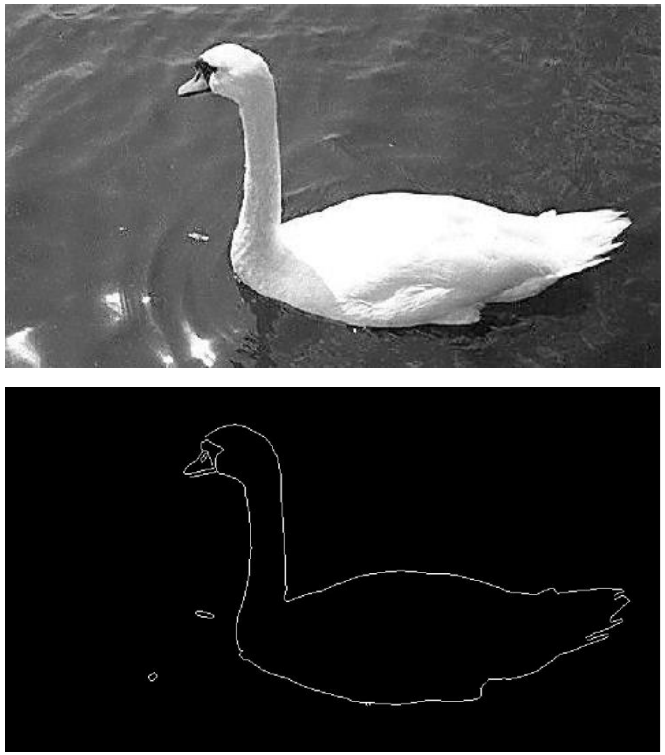
Shape Matching



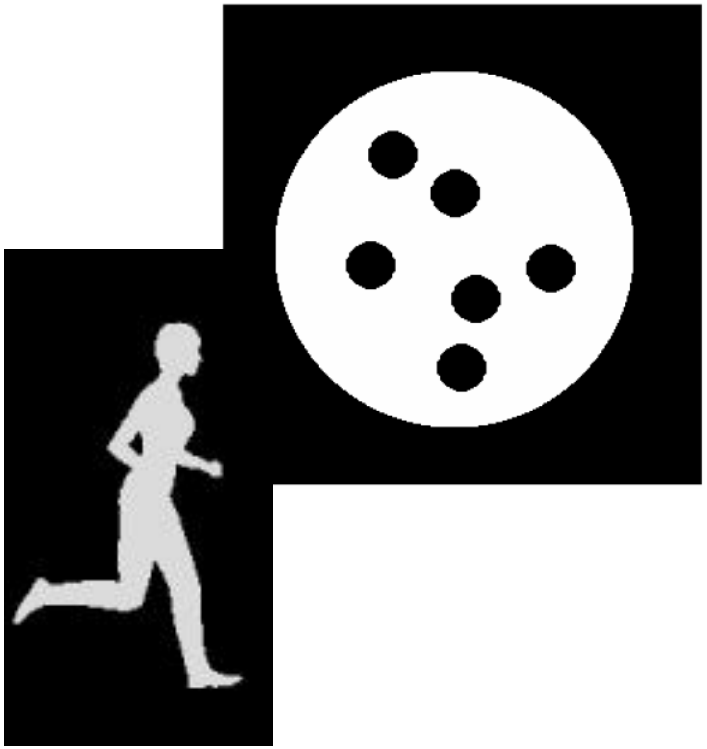


- Where have we encountered shape before?

Low-level features



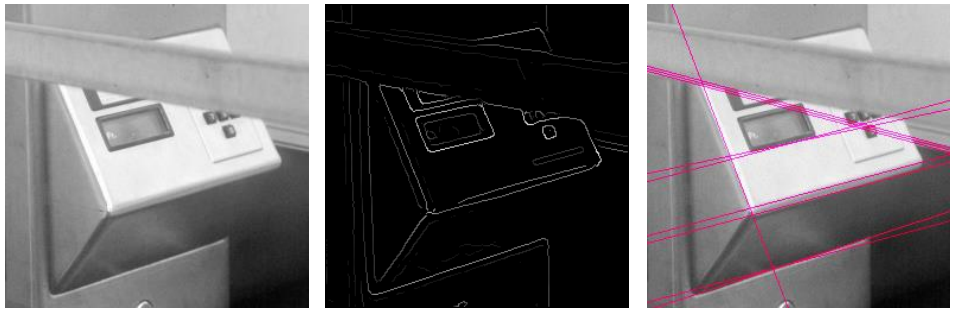
Edges



Silhouettes

Fitting

- Want to associate a model with observed features



[Fig from Marszalek & Schmid, 2007]

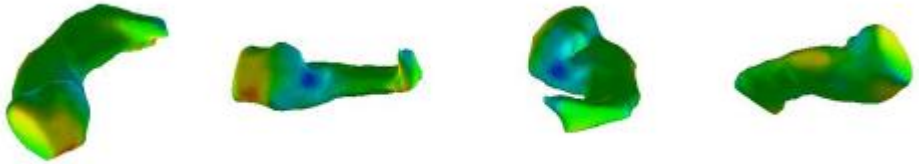
For example, the model could be a line, a circle, or an arbitrary shape.

Deformable contours

Visual Dynamics Group, Dept. Engineering Science, University of Oxford.

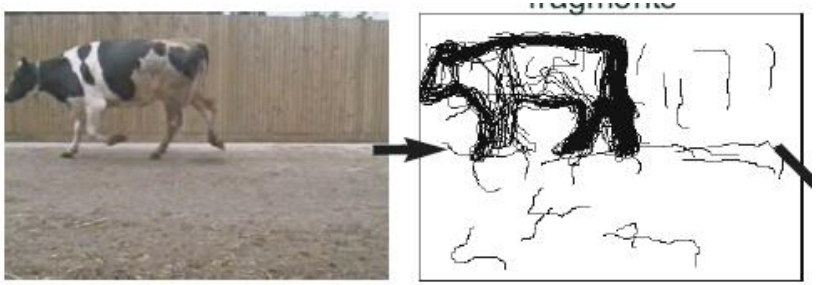
Applications:
- Traffic monitoring
- Human-computer interaction
- Animation
- Surveillance
- Computer Assisted Diagnosis in medical imaging

Role of shape



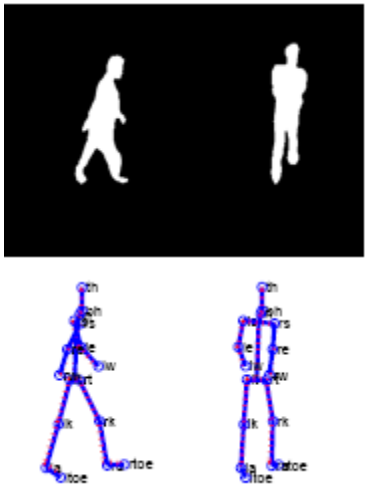
Analysis of anatomical structures

Figure from Grimson & Golland



Recognition, detection

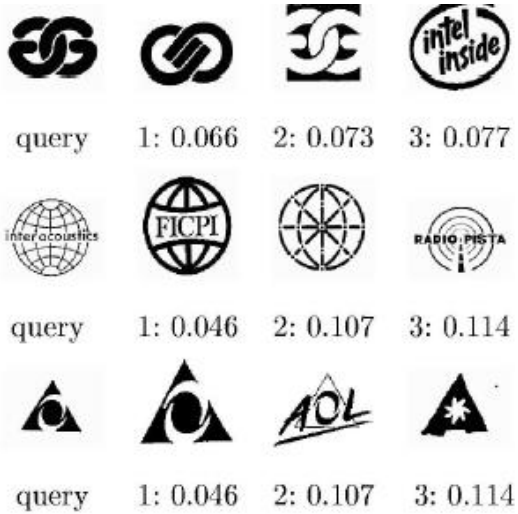
Fig from Opelt et al.



Pose

Morphology

http://usuarios.lycos.es/lawebdelosfosiles/i



Characteristic feature

Fig from Belongie et al.

Shape in recognition

Questions

- What features?
- How to compare shapes?

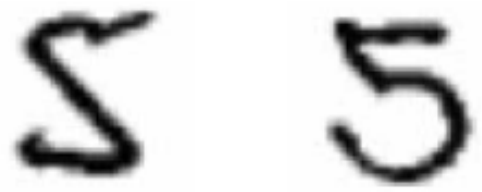


Fig. 1. Examples of two handwritten digits. In terms of pixel-to-pixel comparisons, these two images are quite different, but to the human observer, the shapes appear to be similar.

Chamfer distance

- Average distance to nearest feature

$$D_{chamfer}(T, I) \equiv \frac{1}{|T|} \sum_{t \in T} d_I(t)$$

- T: template shape→ a set of points
- I: image to search→ a set of points
- $d_I(t)$: min distance for point t to some point in I

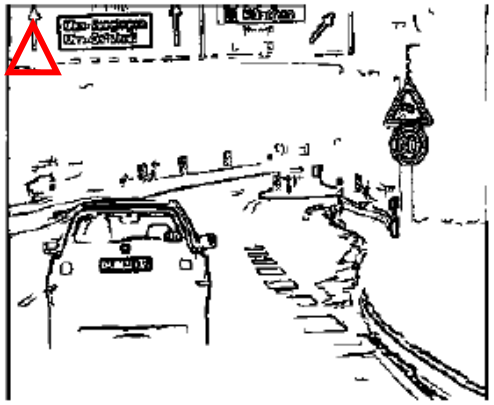
Chamfer distance

- Average distance to nearest feature

$$D_{chamfer}(T, I) \equiv \frac{1}{|T|} \sum_{t \in T} d_I(t)$$

How is the measure different than just filtering with a mask having the shape points?

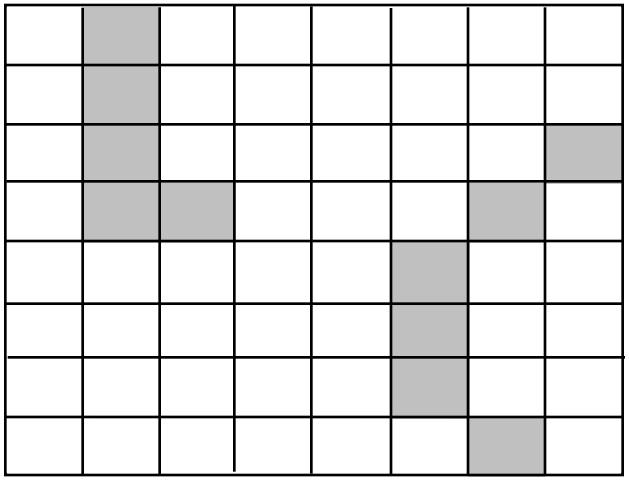
How expensive is a naïve implementation?



Edge image

Distance transform

Image features (2D)



Distance Transform

1	0	1	2	3	4	3	2
1	0	1	2	3	3	2	1
1	0	1	2	3	2	1	0
1	0	0	1	2	1	0	1
2	1	1	2	1	0	1	2
3	2	2	2	1	0	1	2
4	3	3	2	1	0	1	2
5	4	4	3	2	1	0	1

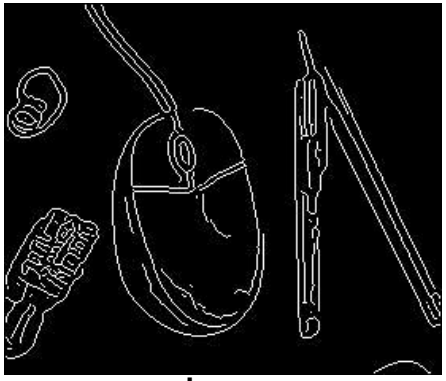
Distance Transform is a function $D(\cdot)$ that for each image pixel p assigns a non-negative number $D(p)$ corresponding to distance from p to the nearest feature in the image I

Features could be edge points, foreground points,…

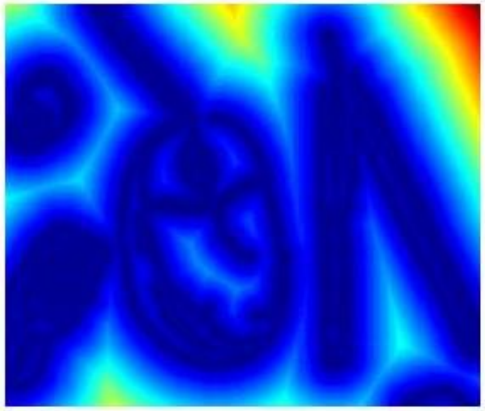
Source: Yuri Boykov

Distance transform

original



edges



distance transform

Value at (x,y) tells how far that position is from the nearest edge point (or other binary mage structure)

```
>> help bwdist
```

Distance transform (1D)

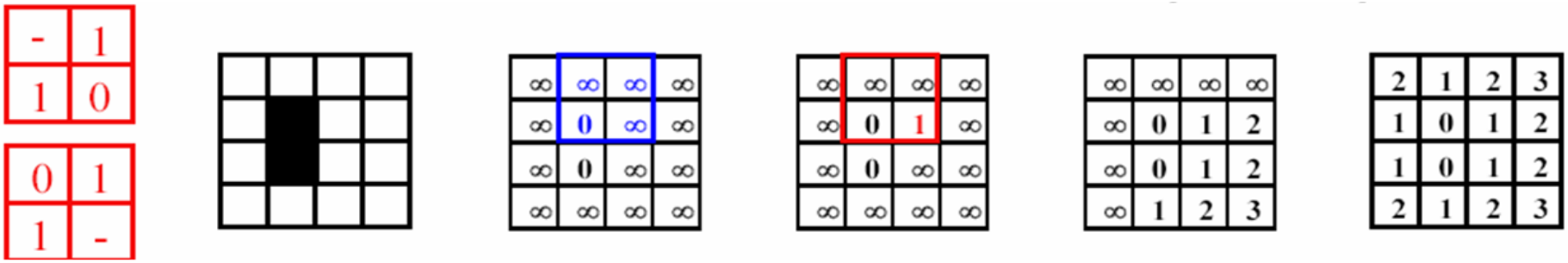
Two pass O(n) algorithm for 1D L_1 norm

1. Initialize: For all j
   $D[j] \leftarrow 1_{\mathbf{P}}[j]$ // 0 if j is in **P**, infinity otherwise

Adapted from D. Huttenlocher

Distance Transform (2D)

- 2D case analogous to 1D
  - Initialization
  - Forward and backward pass
    - Fwd pass finds closest above and to left
    - Bwd pass finds closest below and to right

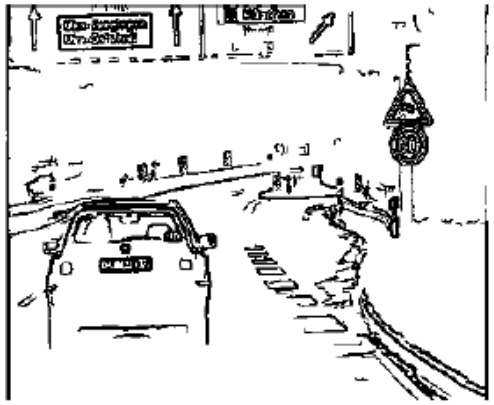


Adapted from D. Huttenlocher

Chamfer distance

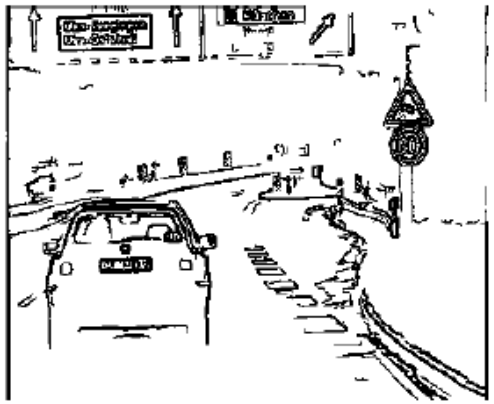
- Average distance to nearest feature

$$D_{chamfer}(T, I) \equiv \frac{1}{|T|} \sum_{t \in T} d_I(t)$$



Edge image **Distance transform image**

Chamfer distance



Edge image **Distance transform image**

Fig from D. Gavrila, DAGM 1999

A limitation of active contours

- External energy: snake does not really “see” object boundaries in the image unless it gets very close to it.

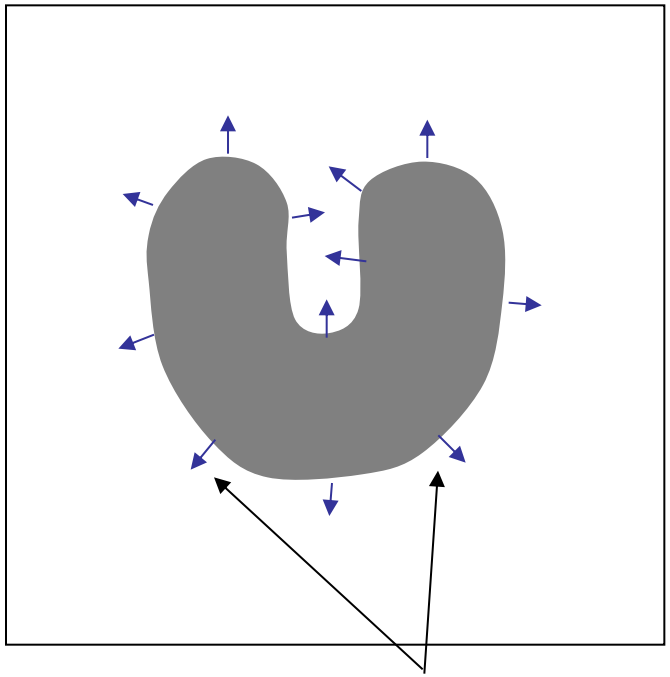
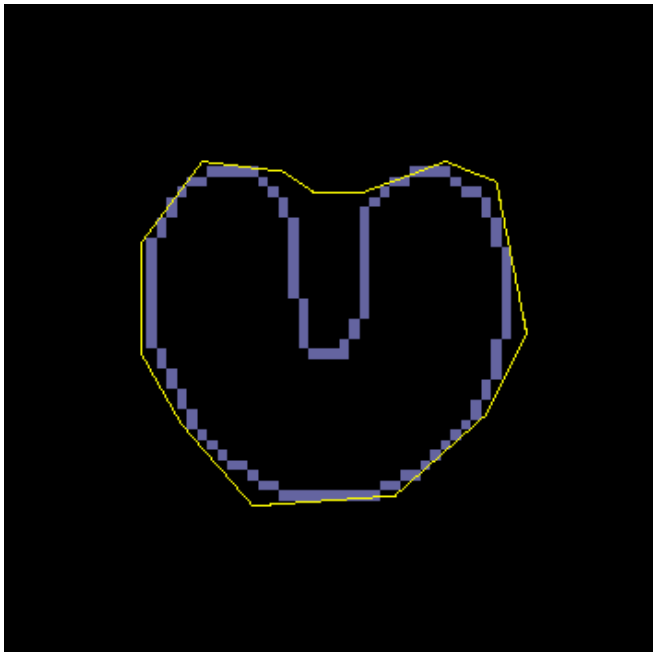


image gradients ∇I
are large only directly on the boundary

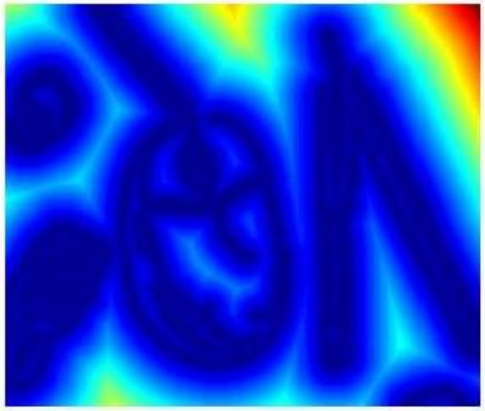


Distance transform can help

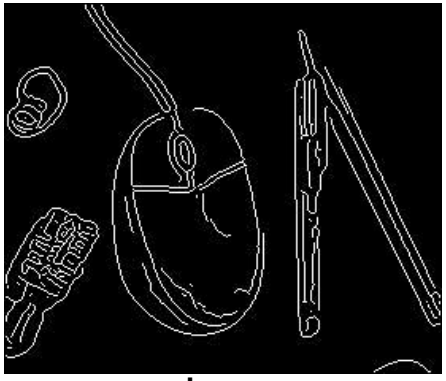
- **External image cost** can also be taken from the **distance transform** of the edge image.

original

-gradient

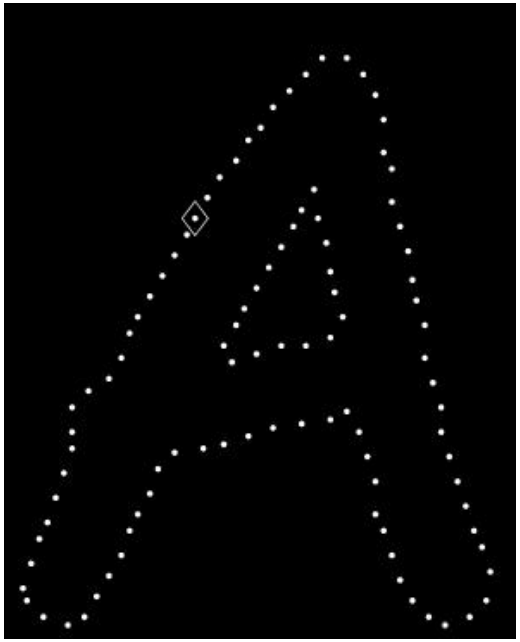


distance transform

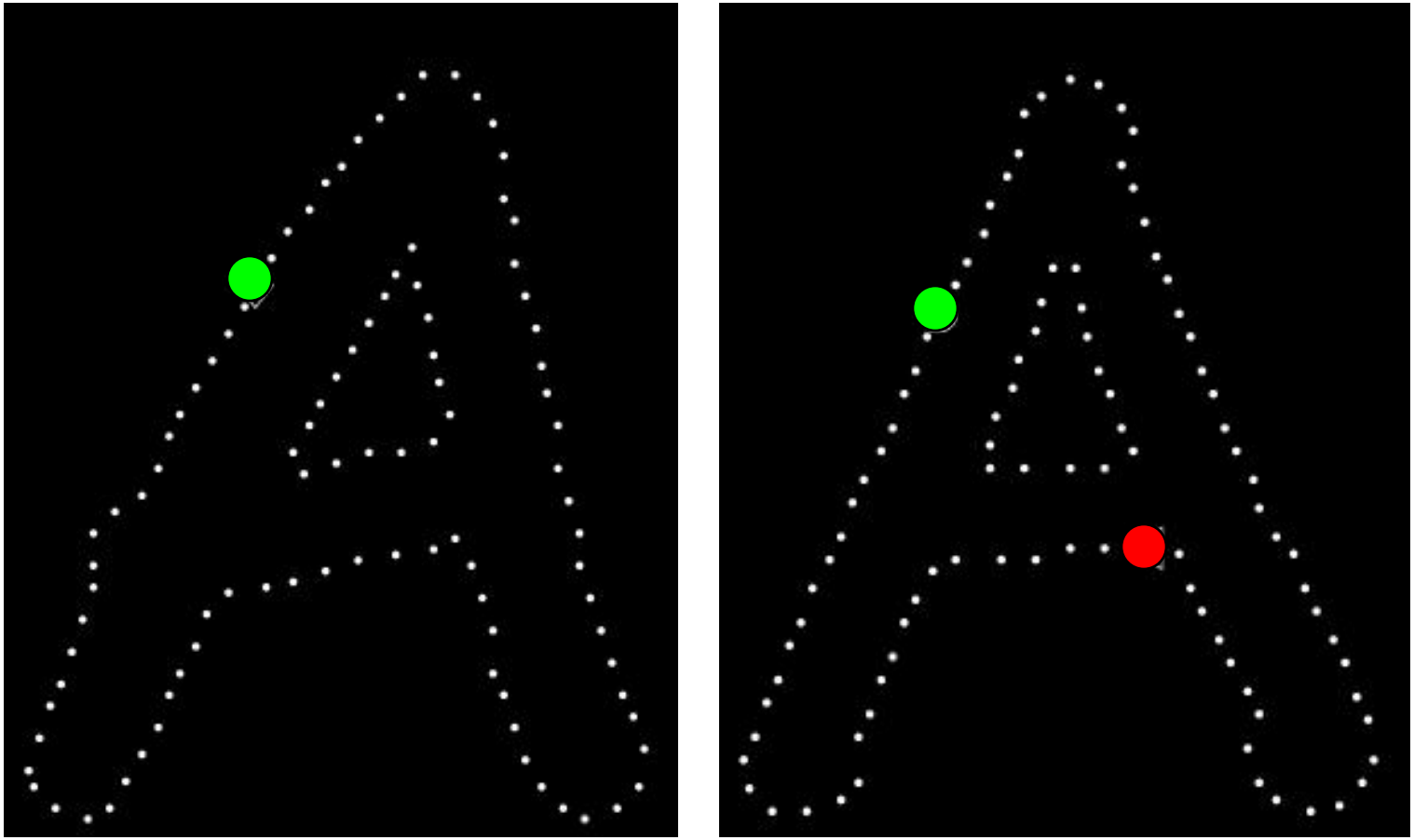


edges

- What limitations might we have using only edge points to represent a shape?
- How descriptive is a point?

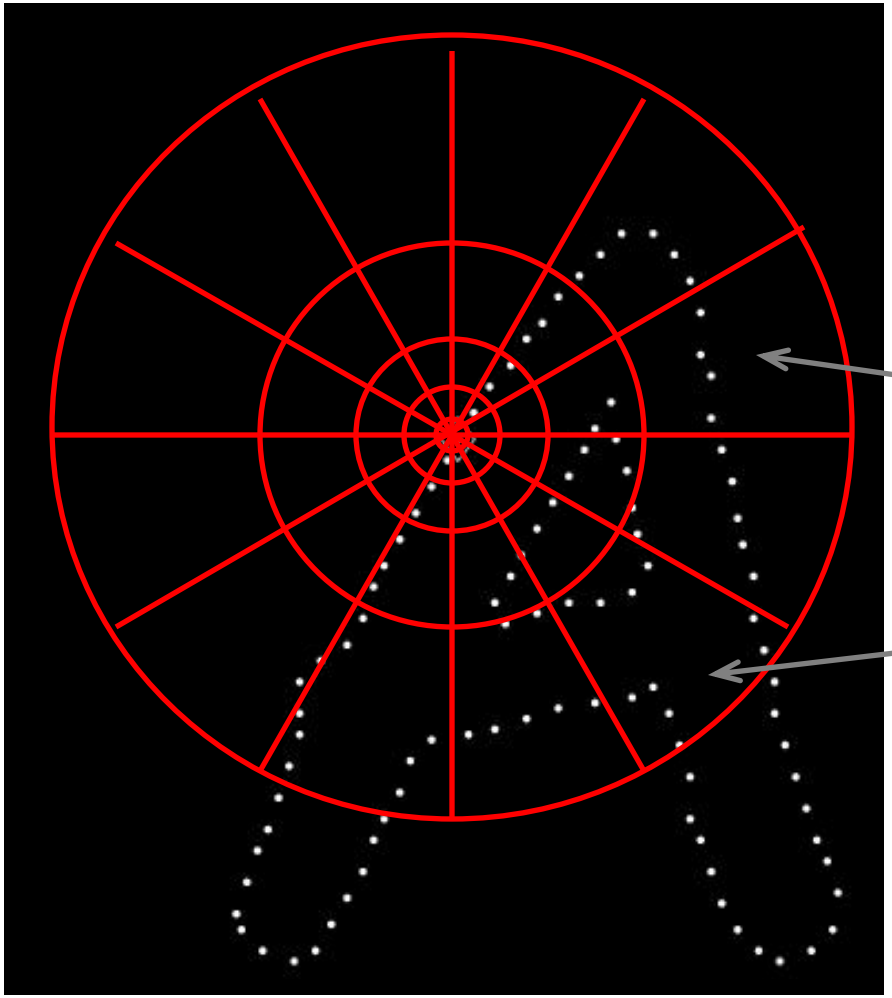


Comparing shapes



What points on these two sampled contours are most similar? How do you know?

Shape context descriptor



Count the number of points inside each bin, e.g.:

Count = 4

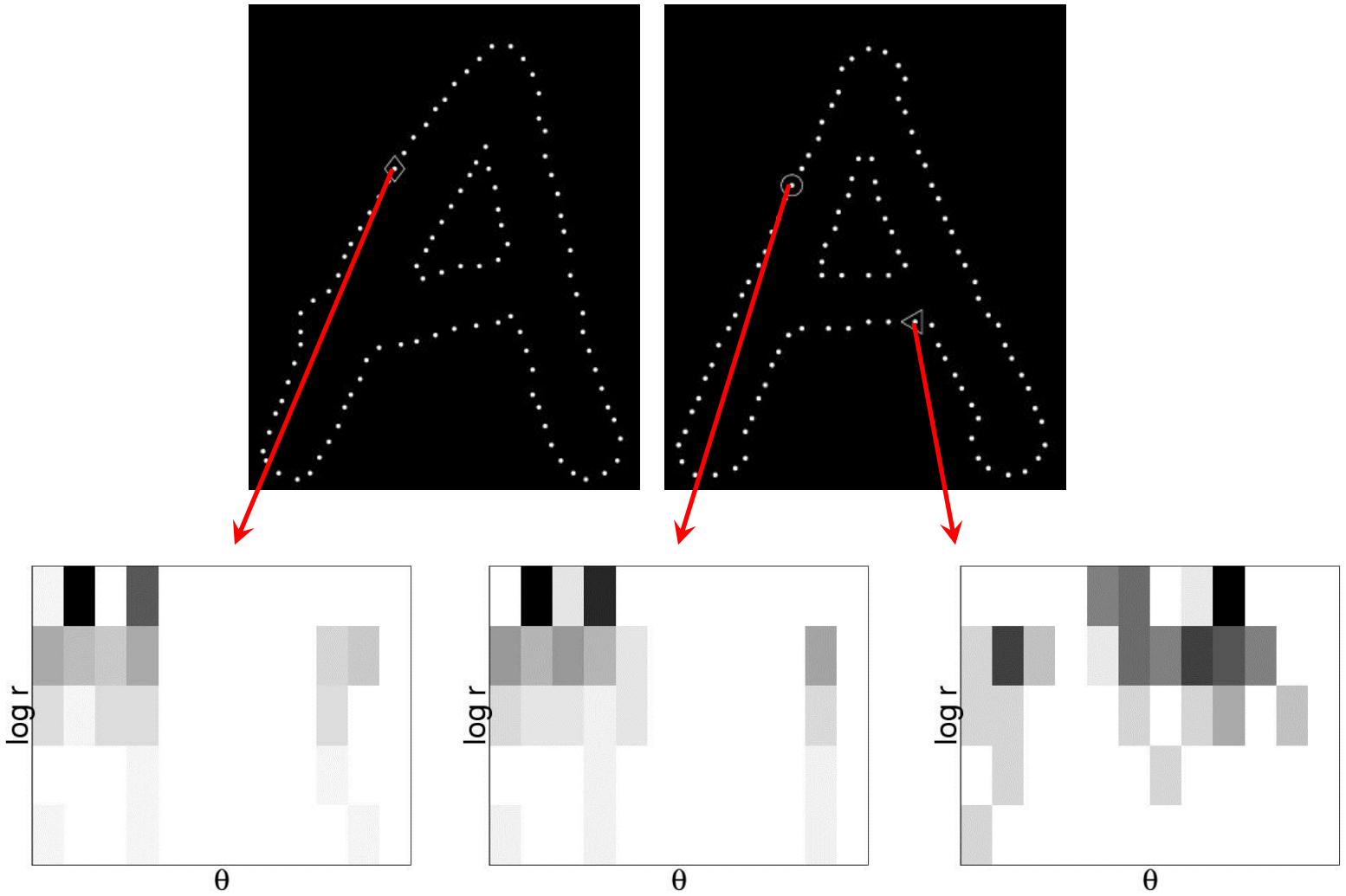
⋮

Count = 10

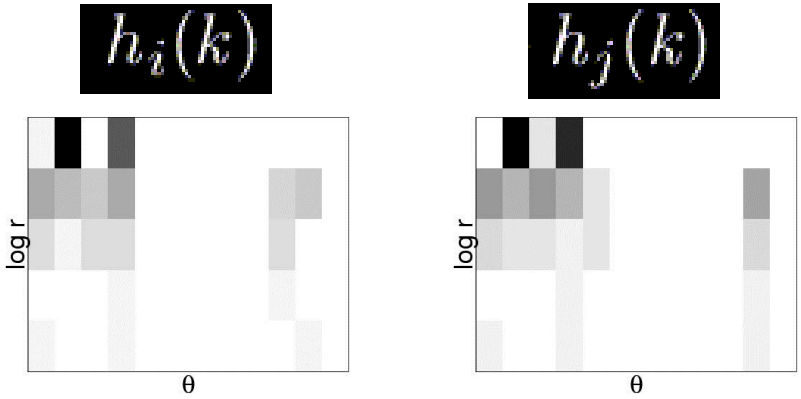
Compact representation of distribution of points relative to each point

Shape context slides from Belongie et al.

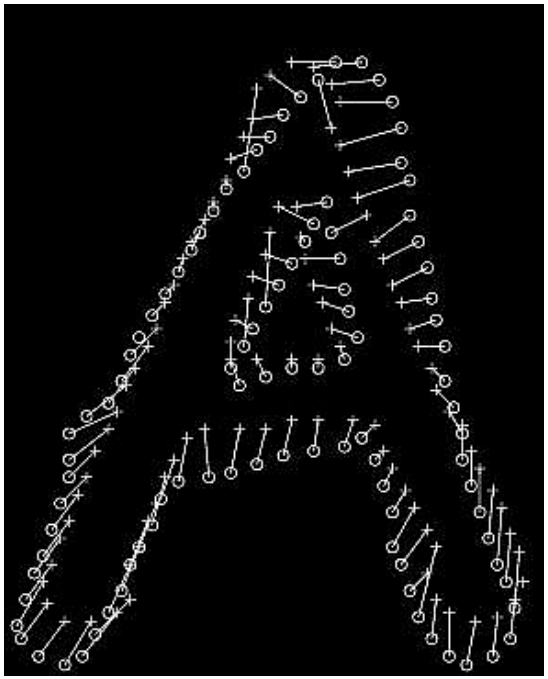
Shape context descriptor



Comparing shape contexts



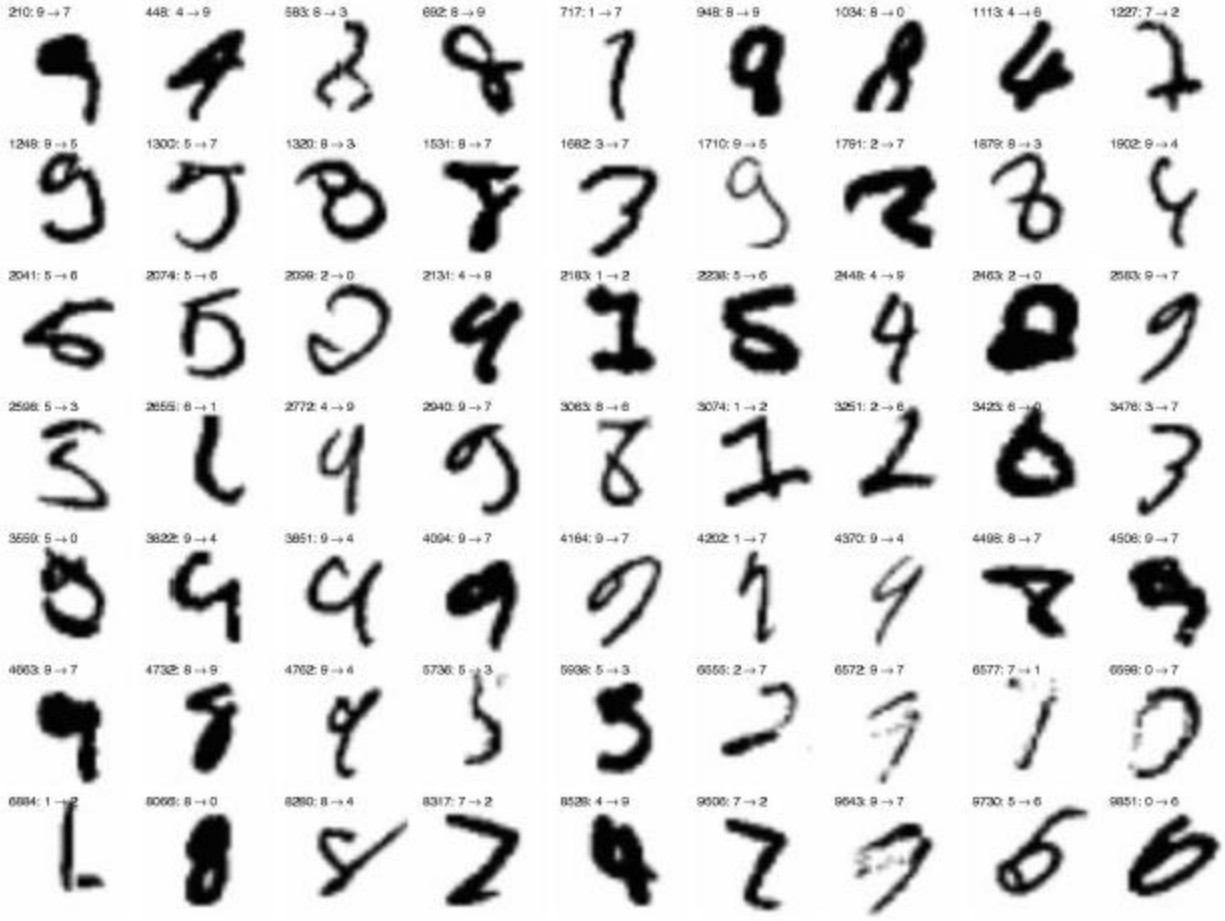
$$C_{ij} = \frac{1}{2}\sum_{k=1}^{K}\frac{[h_i(k) - h_j(k)]^2}{h_i(k) + h_j(k)}$$



Recover correspondences by solving for least cost assignment, using costs C_{ij}

(Then use a deformable template match, given the correspondences.)

Shape context matching with handwritten digits

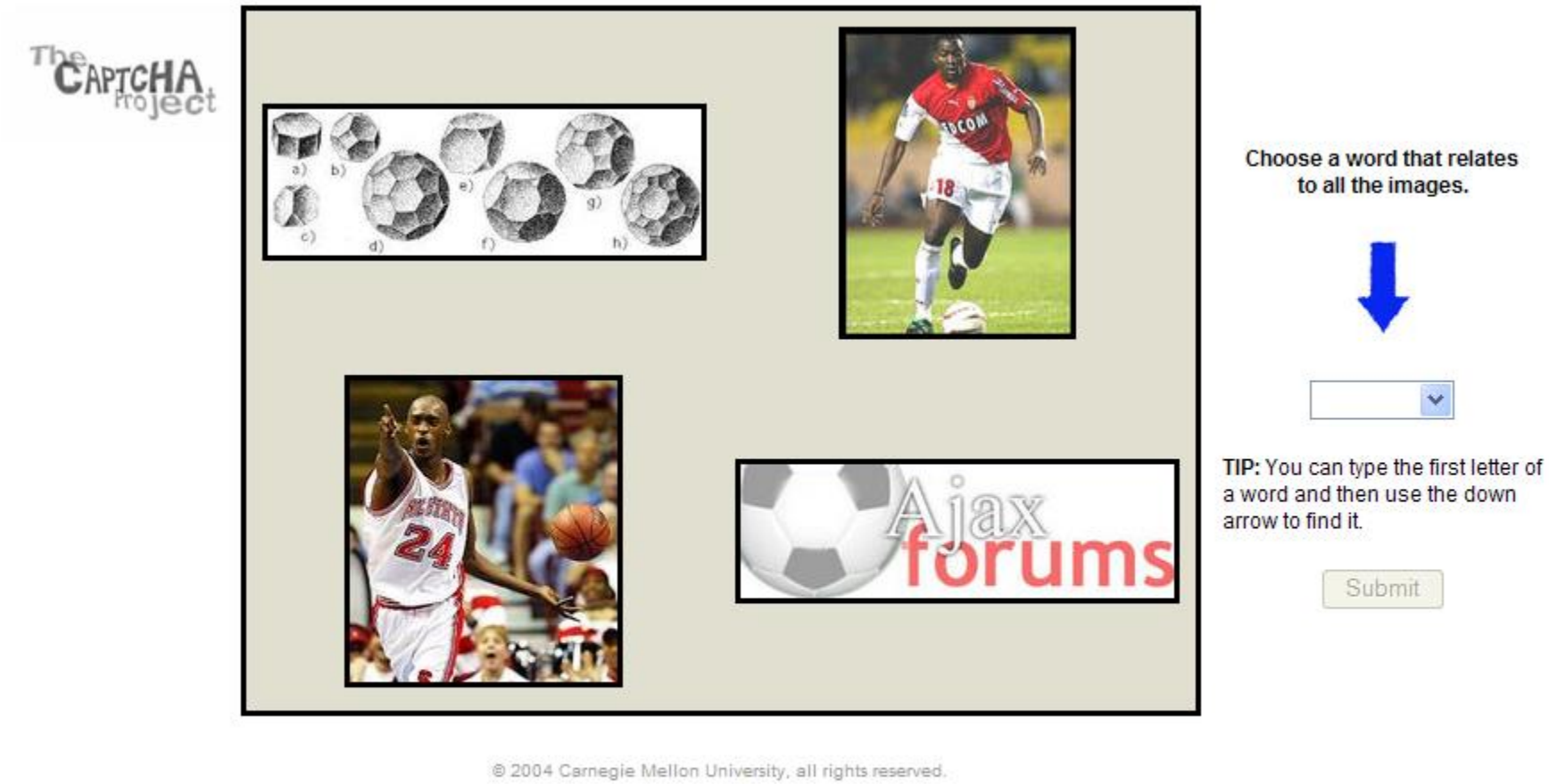


Only errors made out of 10,000 test examples

CAPTCHA’s

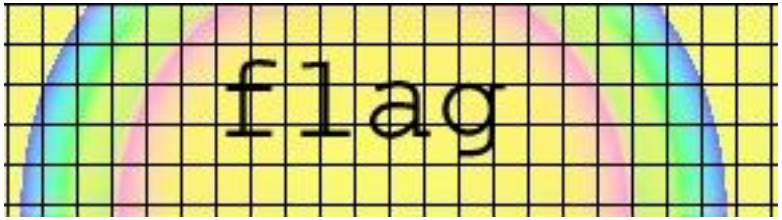
- CAPTCHA: Completely Automated Turing Test To Tell Computers and Humans Apart
- Luis von Ahn, Manuel Blum, Nicholas Hopper and John Langford, CMU, 2000.
- www.captcha.net

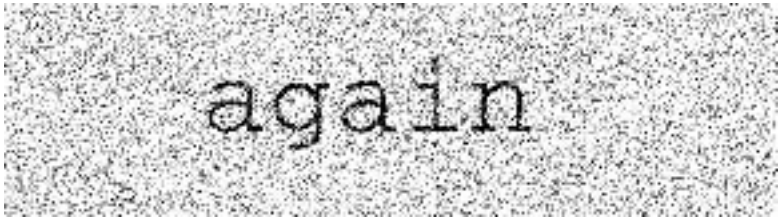
Image-based CAPTCHA



Shape matching application: breaking a visual CAPTCHA

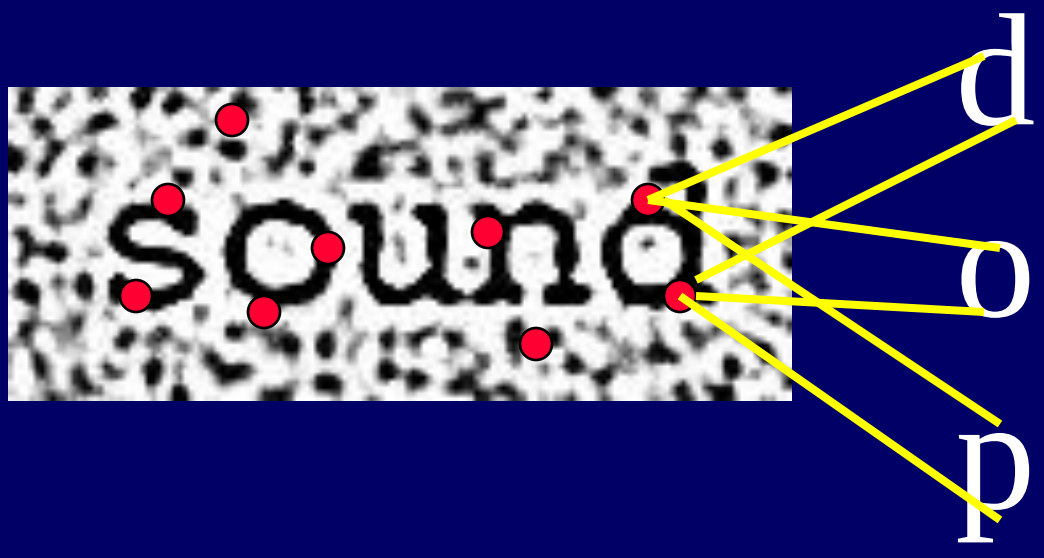
- Use shape matching to recognize characters, words in spite of clutter, warping, etc.





Recognizing Objects in Adversarial Clutter: Breaking a Visual CAPTCHA, by G. Mori and J. Malik, CVPR 2003

Fast Pruning: Representative Shape Contexts



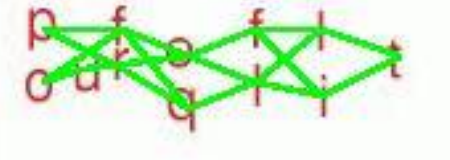
- Pick k points in the image at random
  - Compare to all shape contexts for all known letters
  - Vote for closely matching letters
- Keep all letters with scores under threshold

Algorithm A: bottom-up

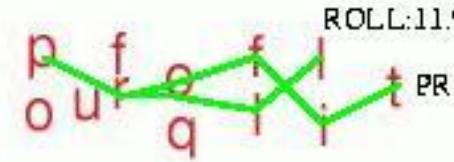
- Look for letters
  - Representative Shape Contexts
- Find pairs of letters that are “consistent”
  - Letters nearby in space
- Search for valid words
- Give scores to the words

Input

Locations of possible letters



Possible strings of letters



Matching words

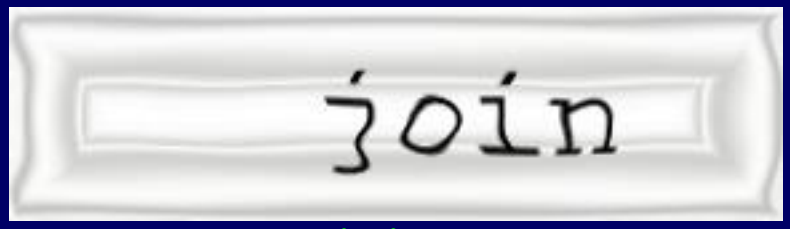
EZ-Gimpy Results with Algorithm A

- 158 of 191 images correctly identified: 83%
  - Running time: ~10 sec. per image (MATLAB, 1 Ghz P3)

horse

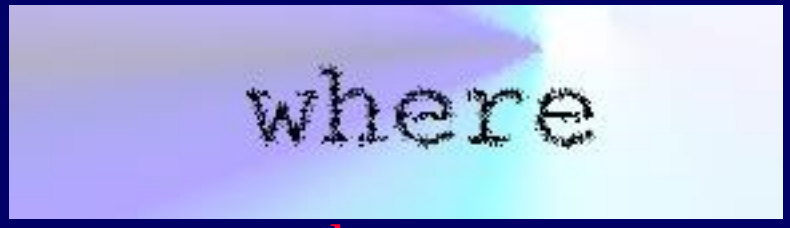
spade

smile



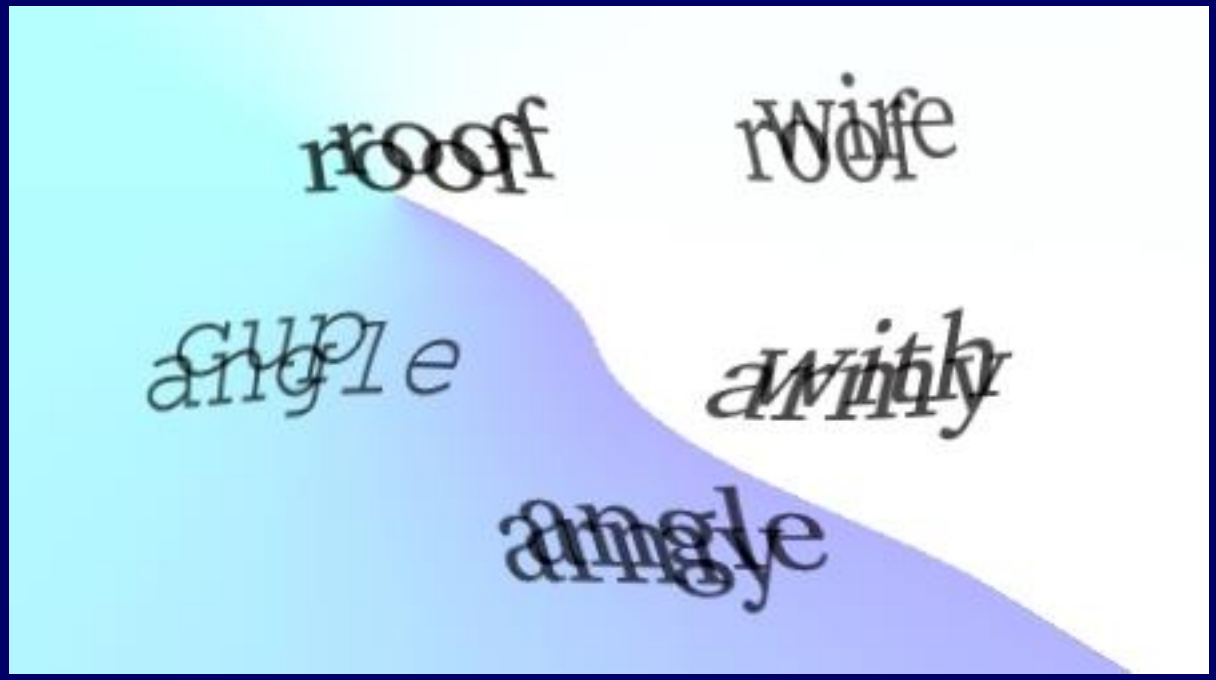
join

canvas



here

Gimpy



- Multiple words, task is to find 3 words in the image
- Clutter is other objects, not texture

Algorithm B: Letters are not enough

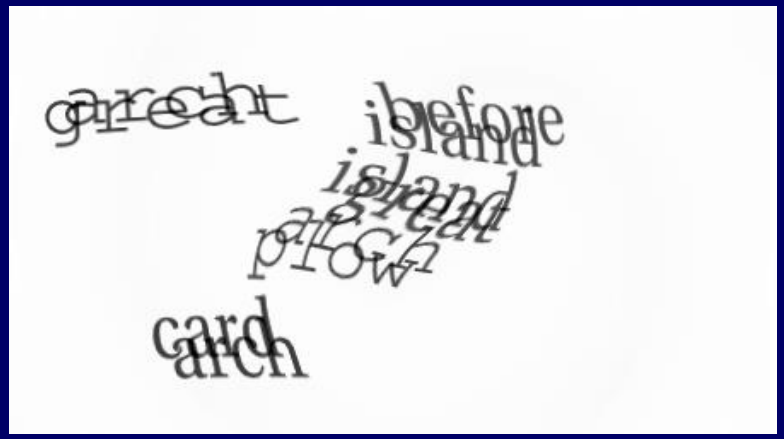



- Hard to distinguish single letters with so much clutter
- Find words instead of letters
  - Use long range info over entire word
  - Stretch shape contexts into ellipses

- Search problem becomes huge
  - # of words 600 vs. # of letters 26
  - Prune set of words using opening/closing bigrams

Results with Algorithm B

dry clear medical

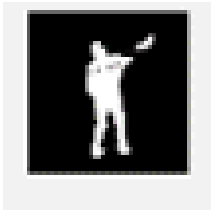


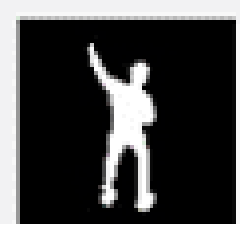
card arch plate

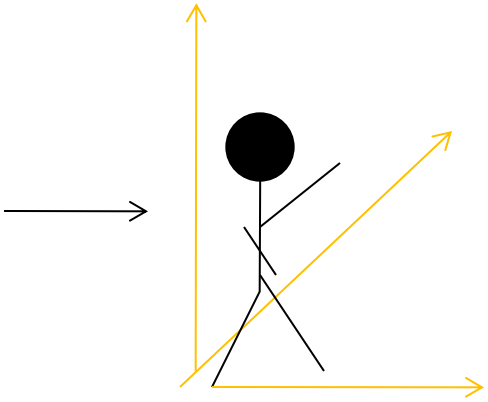
# Correct words	% tests (of 24)
1 or more	92%
2 or more	75%
3	33%
EZ-Gimpy	92%

door farm important

Shape matching application II: silhouettes and body pose







What kind of assumptions do we need?

Fun with silhouettes

- Liu Ren, Gregory Shakhnarovich, Jessica Hodgins, Hanspeter Pfister and Paul Viola, Learning Silhouette Features for Control of Human Motion
- http://graphics.cs.cmu.edu/projects/swing/