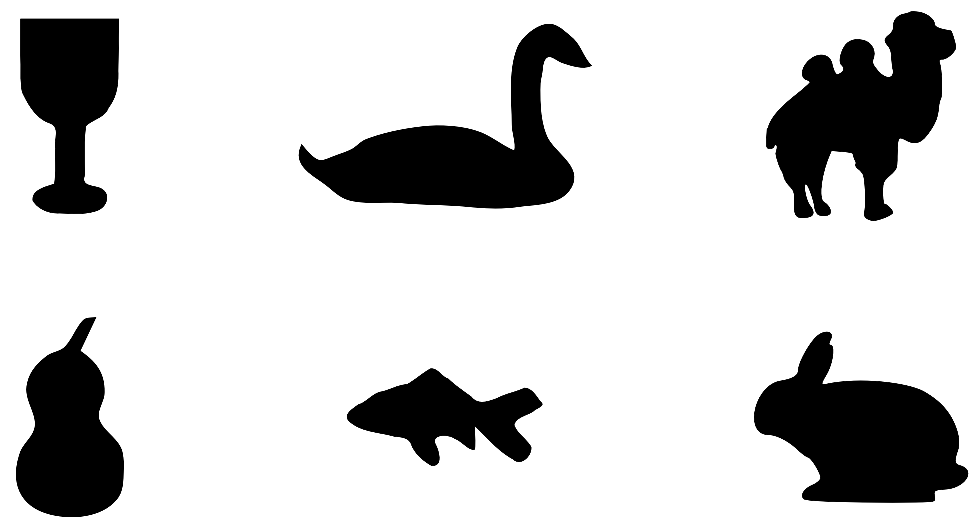
Shape Matching

Intro

Shape-Based Recognition

- Humans can recognize many objects based on shape alone
- Fundamental cue for many object categories
- Invariant to photometric variation.



Slide from Pedro Felzenszwalb

Intro

Shapes vs. Intensity Values

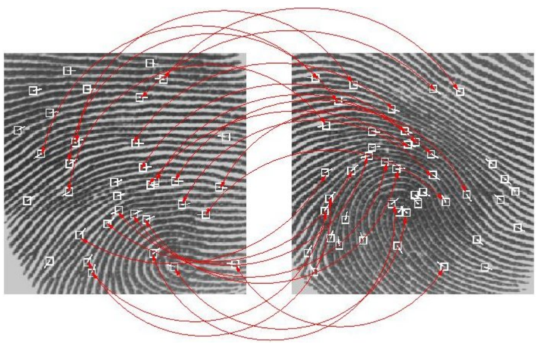
Similar to a human in terms of shape, but very different in terms of pixel values.

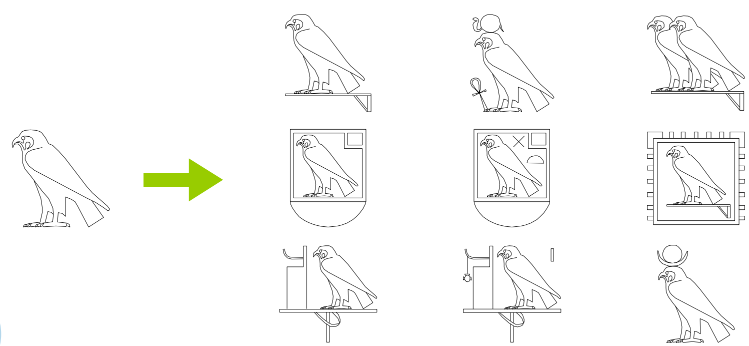
Images from Belongie et al.

Intro

Applications

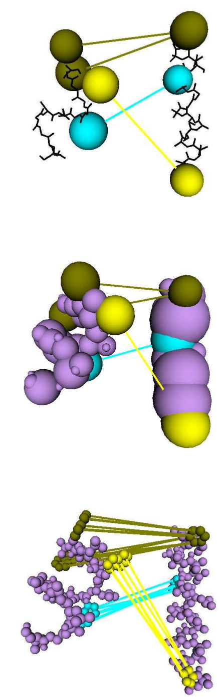
- Shape retrieval
- Recognizing object categories
- Fingerprint identification
- Optical Character Recognition (OCR)
- Molecular-biology





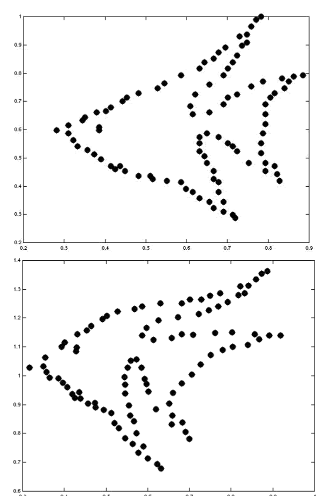
Western

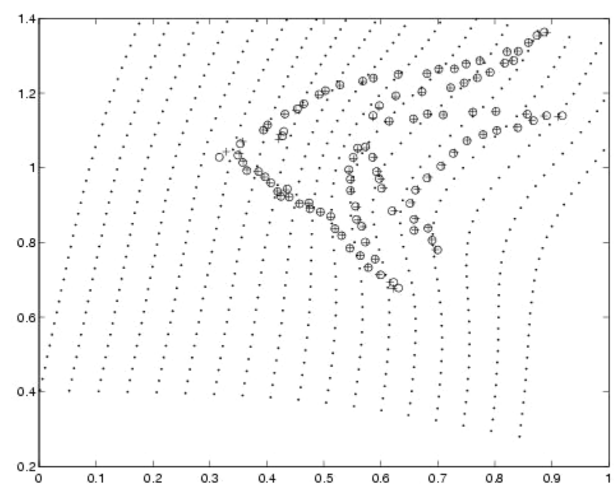
1909



Geometric Transformations

- Often in matching images are allowed to undergo some geometric transformation
- Related but not identical shapes can be deformed into alignment using simple coordinate transformations
- Find the transformations of one image that produce good matches to the other image





Images from Belongie et al.

Intro

Biological Shape

- D’Arcy Thompson: *On Growth and Form*, 1917
  - studied transformations between shapes of organisms

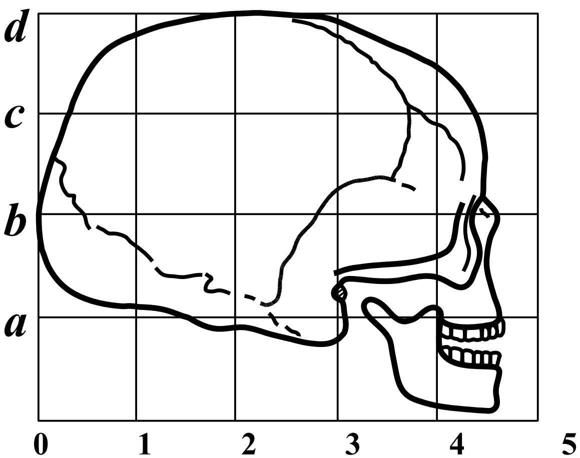


Fig. 177. Human skull

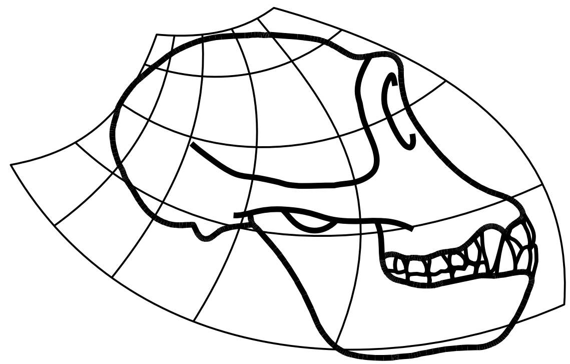


Fig. 179. Skull of chimpanzee.

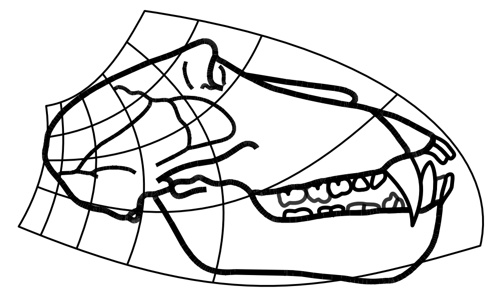
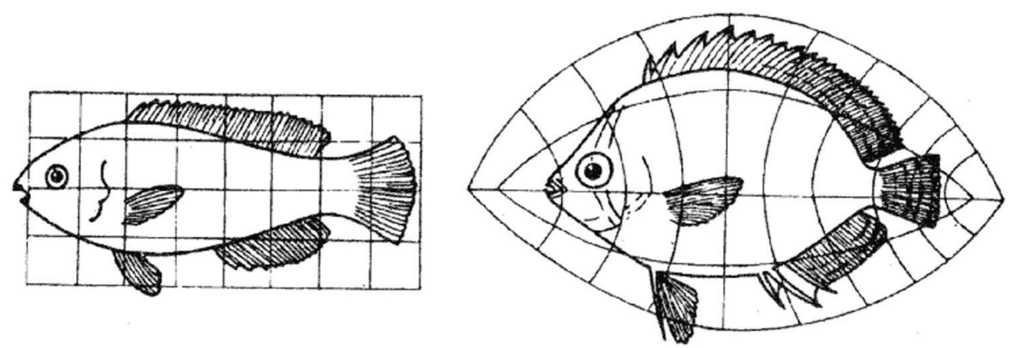


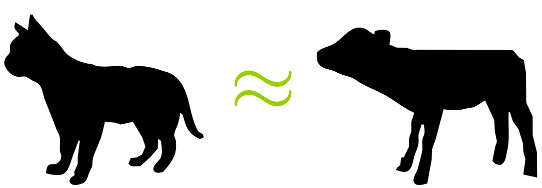
Fig. 180. Skull of baboon.

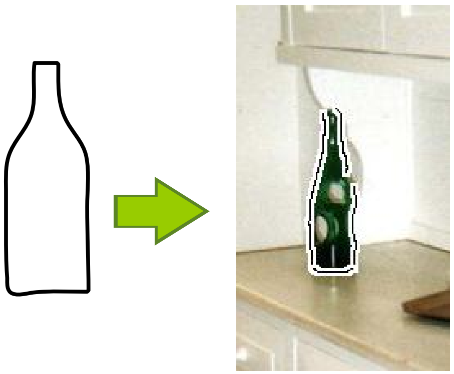


Slide from Belongie et al.

Related Problems

- Shape representation and decomposition
- Finding a set of correspondences between shapes
- Transforming one shape into another
- Measuring the similarity between shapes
- Shape localization and model alignment
- Finding a shape similar to a model in a cluttered image





Slide from Pedro Felzenszwalb

Intro

References

- **Shape Matching and Object Recognition Using Shape Contexts**, by S. Belongie, J. Malik, and J. Puzicha. Transactions on Pattern Analysis and Machine Intelligence (PAMI), 2002.
- **Recognizing Objects in Adversarial Clutter: Breaking a Visual CAPTCHA**, by G. Mori and J. Malik, in Proceedings IEEE Computer Vision and Pattern Recognition (CVPR), 2003.
- **Using the Inner-Distance for Classification of Articulated Shapes**, by H. Ling and D. Jacobs, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2005.
- **Comparing Images Using the Hausdorff Distance**, by D. Huttenlocher, G. Klanderman, and W. Rucklidge, Transactions on Pattern Analysis and Machine Intelligence (PAMI), 1993.
- **A Boundary-Fragment-Model for Object Detection**, by A. Opelt, A. Pinz, and A. Zisserman, Proceedings of the European Conference on Computer Vision (ECCV), 2006.
- **Hierarchical Matching of Deformable Shapes**, by P. Felzenszwalb and J. Schwartz, in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2007

Hausdorff Distance

Outline

- Shape Distance and Correspondence
  - ***Hausdorff Distance***
  - Shape Context
  - Inner Distance
- Hierarchical Approach
  - Hierarchical Matching
- Machine Learning Approach
  - Boundary Fragment Model

Hausdorff Distance

Comparing Images Using the Hausdorff Distance

1993

D. Huttenlocher, G. Klanderman, and W. Rucklidge

Hausdorff Distance

Overview

- Use Hausdorff distance to compare images to a model
- Fast and simple approach
- Tolerant of small position errors
- Model is only allowed to translate with respect to the image
- Can be extended to allow rotation and scale

Hausdorff Distance

Hausdorff Distance

- A means of determining the resemblance of one point set to another
- Examines the fraction of points in one set that lie near points in the other set

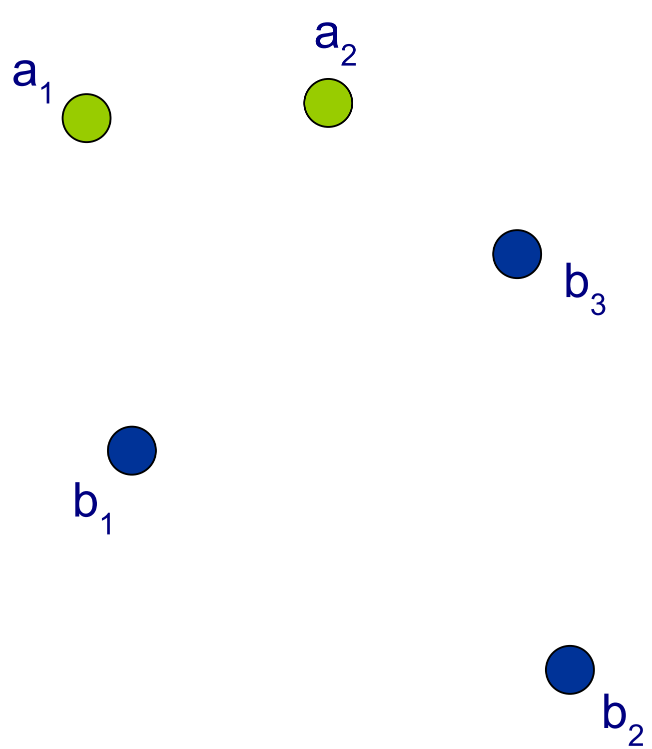
$$H(A,B) = \max\{h(A,B), h(B,A)\}$$

$$h(A,B) = \max_{a \in A}\left\{\min_{b \in B}\{d(a,b)\}\right\}$$

Hausdorff Distance

Example

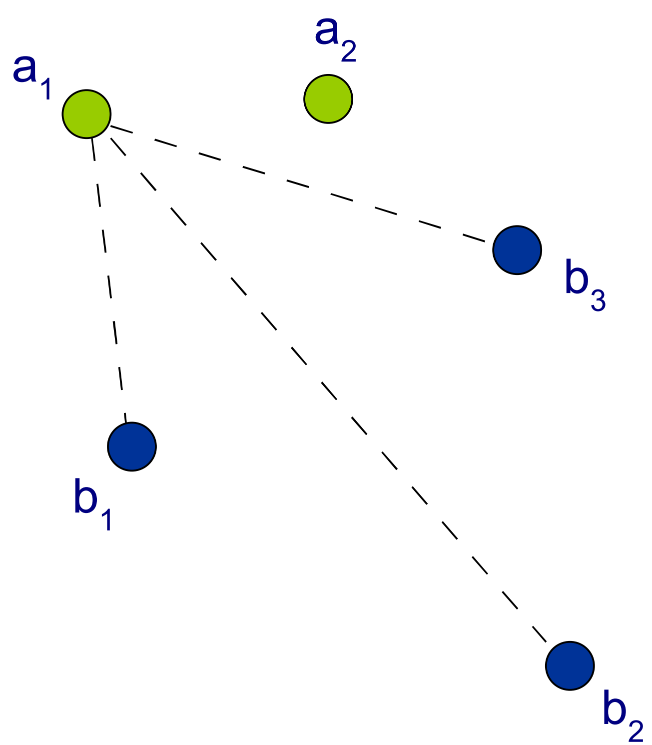
Given two sets of points A and B, find h(A,B)



Hausdorff Distance

Example

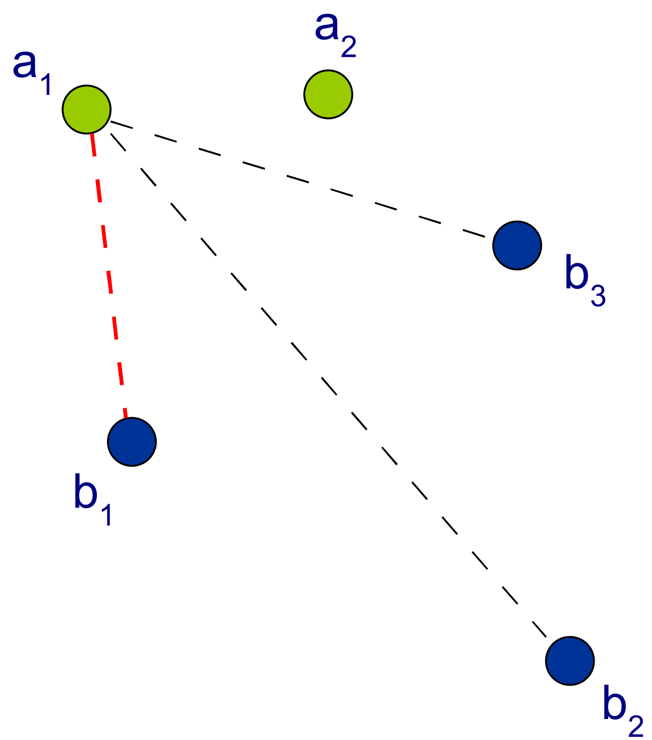
Compute the distance between a_1 and each b_j



Hausdorff Distance

Example

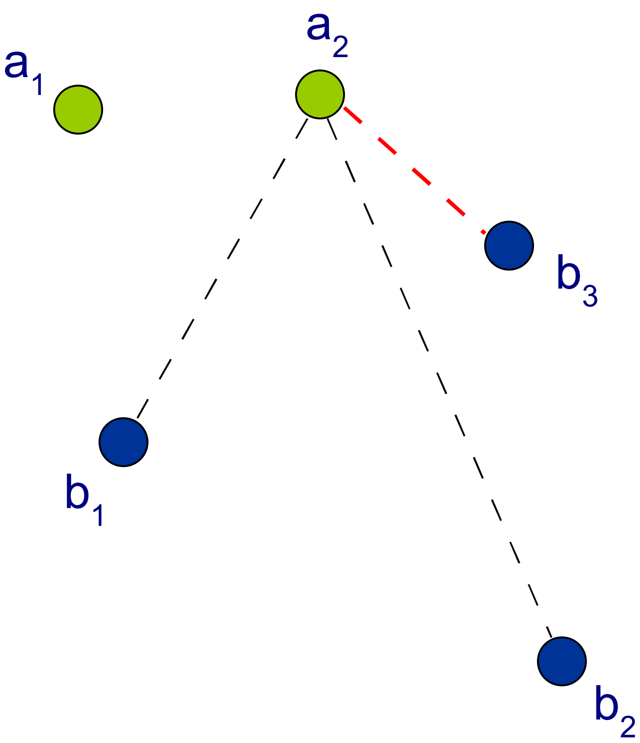
Keep the shortest



Hausdorff Distance

Example

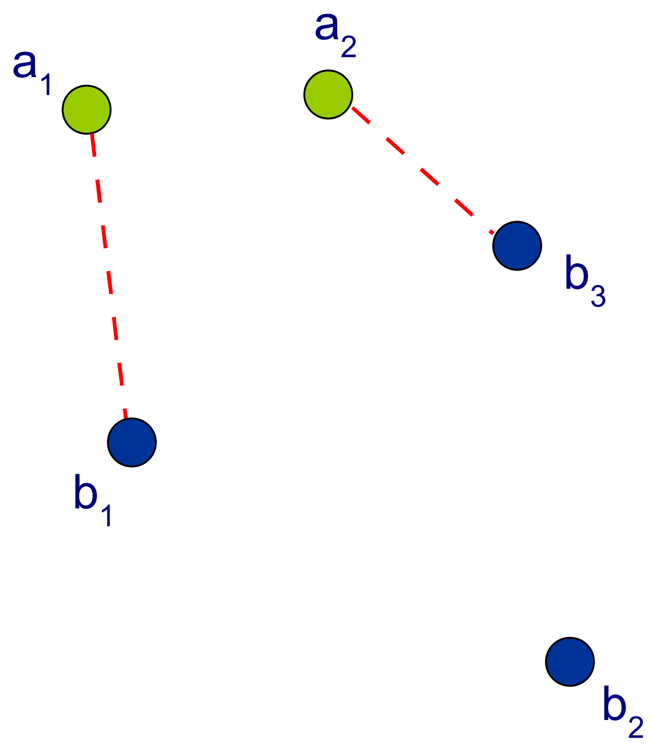
Do the same for a_2



Hausdorff Distance

Example

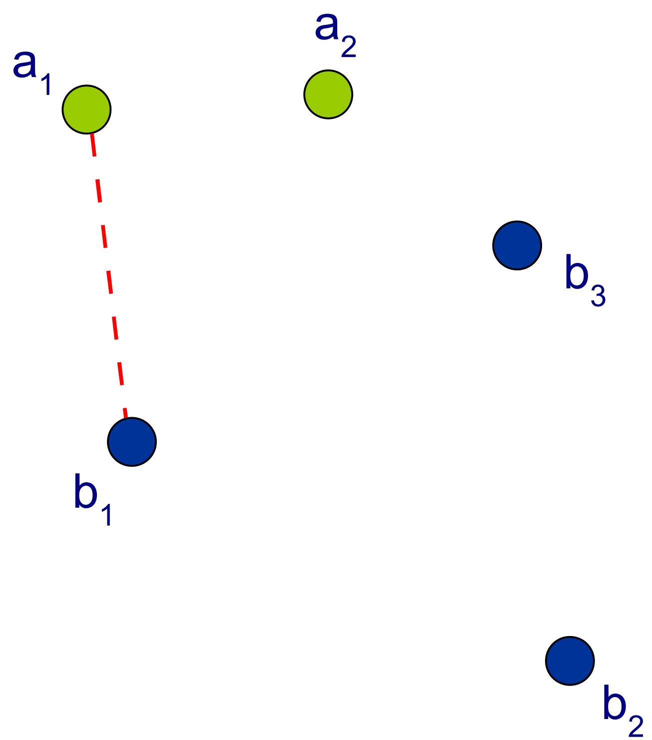
Find the largest of these two distances



Hausdorff Distance

Example

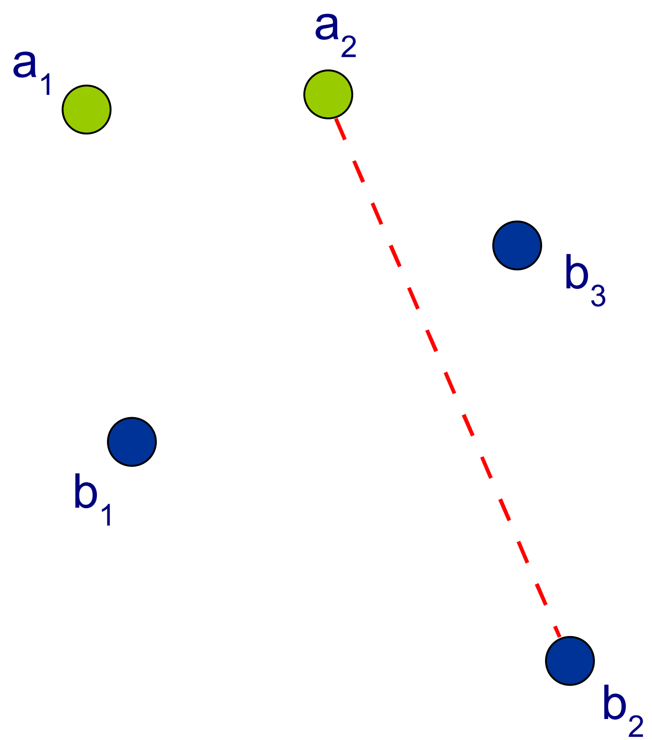
This is h(A,B)



Hausdorff Distance

Example

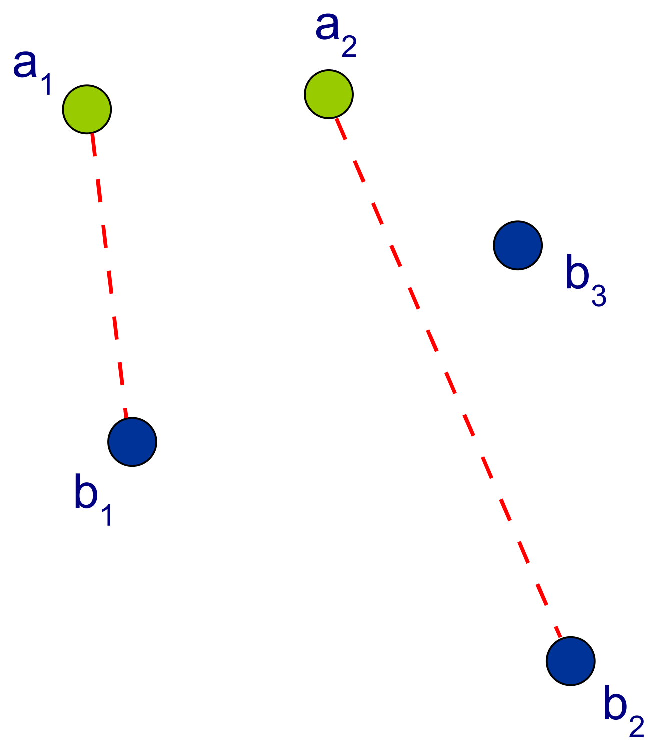
This is h(B,A)



Hausdorff Distance

Example

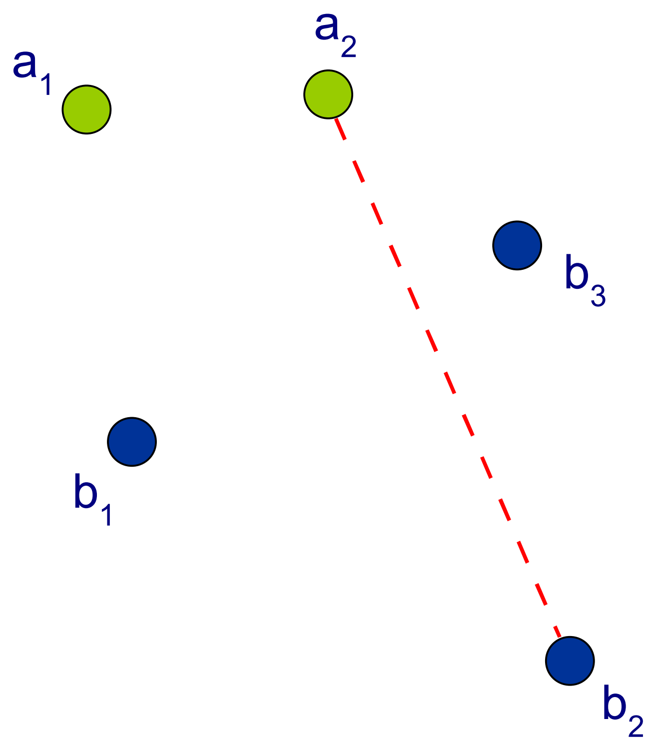
$H(A,B) = \max(h(A,B),h(B,A))$



Hausdorff Distance

Example

This is H(A,B)



Hausdorff Distance

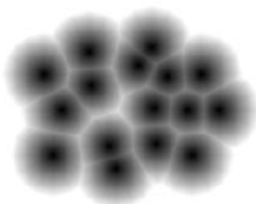
Generalization

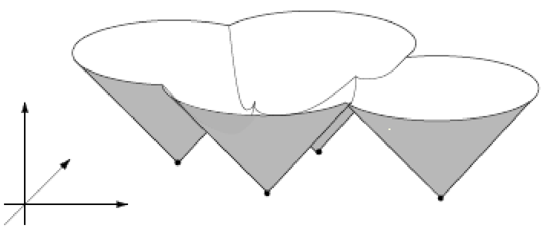
- Hausdorff distance is very sensitive to even one outlier in A or B
- Use kth ranked distance instead of the maximal distance
- Match if $h_k(A, B) < \delta$
  - k is how many points of the model need to be near points of the image
  - δ is how near these points need to be

$$h_k(A, B) = \underset{a \in A}{\mathrm{k}^{\mathrm{th}}} \left\{ \min_{b \in B} \{ d(a, b) \} \right\}$$

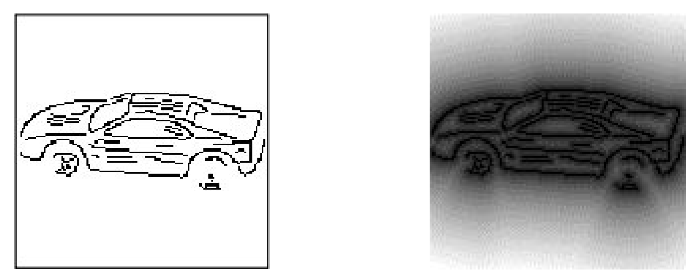
Hausdorff Distance

Distance Transforms



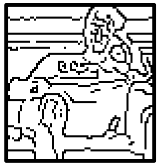


- Processing can be sped up by probing a precomputed Voronoi surface
- A Voronoi surface defines the distance from any location in B to the nearest point
- Can be efficiently computed using dynamic programming in linear time

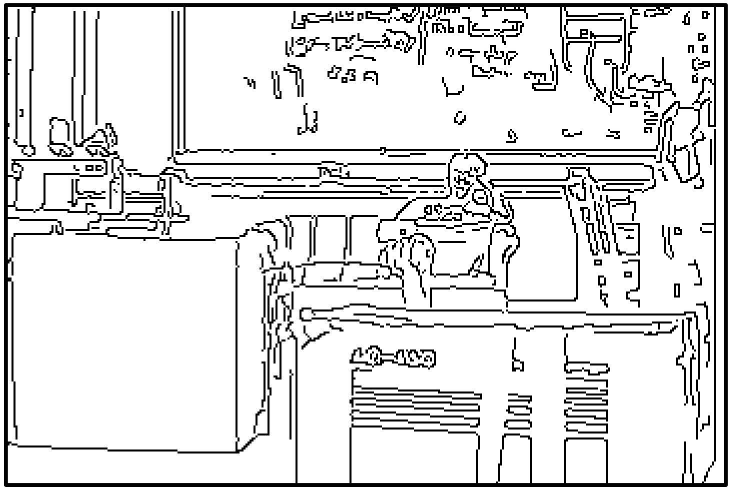


Hausdorff Distance

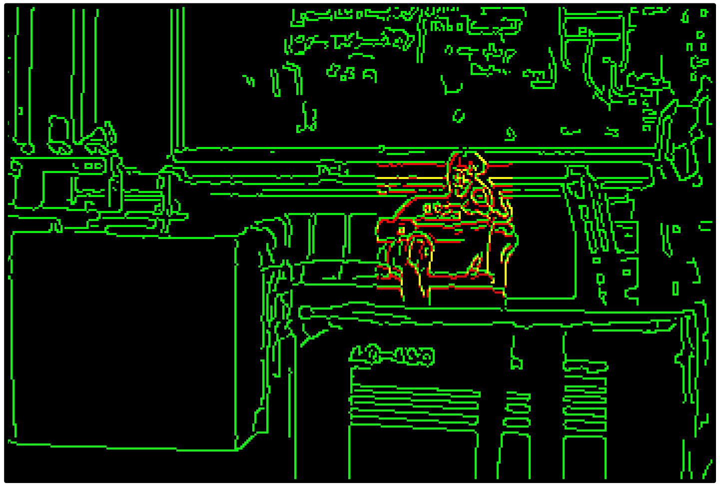
Example: Matching



Model



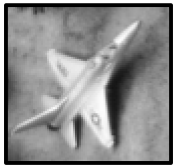
Edges



Match

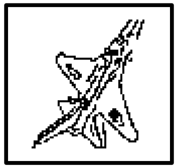
Hausdorff Distance

Example: Matching

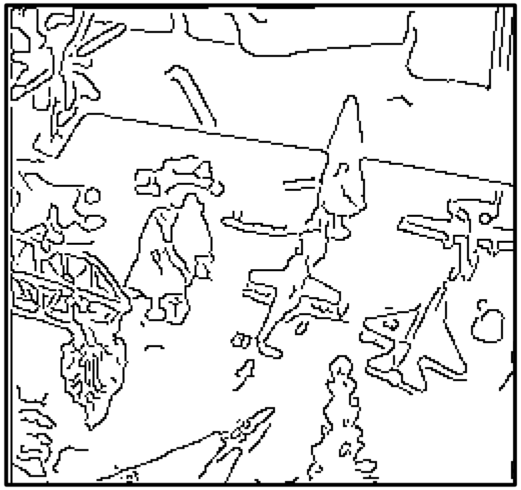


Model

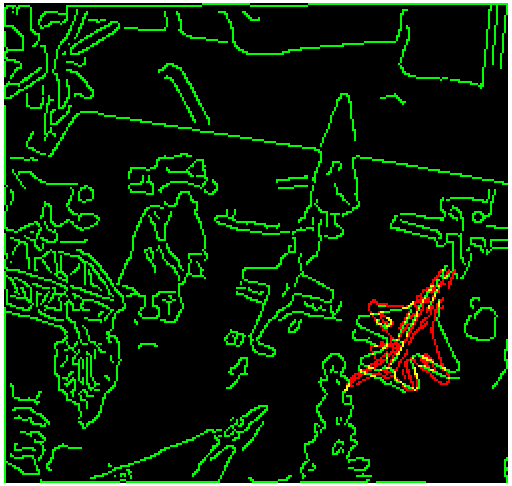
Image



Model



Edges



Match

Outline

- Shape Distance and Correspondence
  - Hausdorff Distance
  - ***Shape Context***
  - Inner Distance
- Hierarchical Approach
  - Hierarchical Matching
- Machine Learning Approach
  - Boundary Fragment Model

Shape Context

Shape Matching and Object Recognition Using Shape Contexts

2002

S. Belongie, J. Malik, and J. Puzicha

Shape Context

Overview

1) Solve for correspondences between points on the two shapes
   - Using shape contexts
2) Use the correspondences to estimate an aligning transform
   - Using regularized thin-plate splines
3) Compute the distance between the two shapes

Shape Context

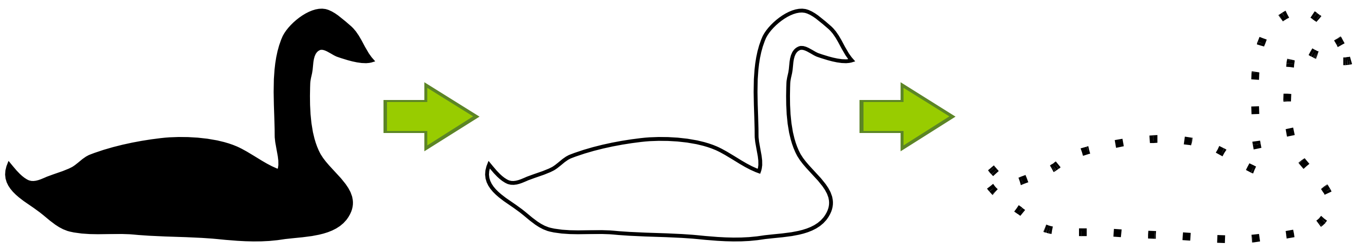
Related Work: Deformable Templates

- **The Representation and Matching of Pictorial Structures**, by Fischler & Elschlager (1973)
- **Structural image restoration through deformable templates**, by Grenander et al. (1991)
- **Deformable Templates for Face Recognition**, by Yuille (1991)
- **Distortion invariant object recognition in the dynamic linkarchitecture**, by von der Malsburg (1993)

Shape Context

Sampling Points

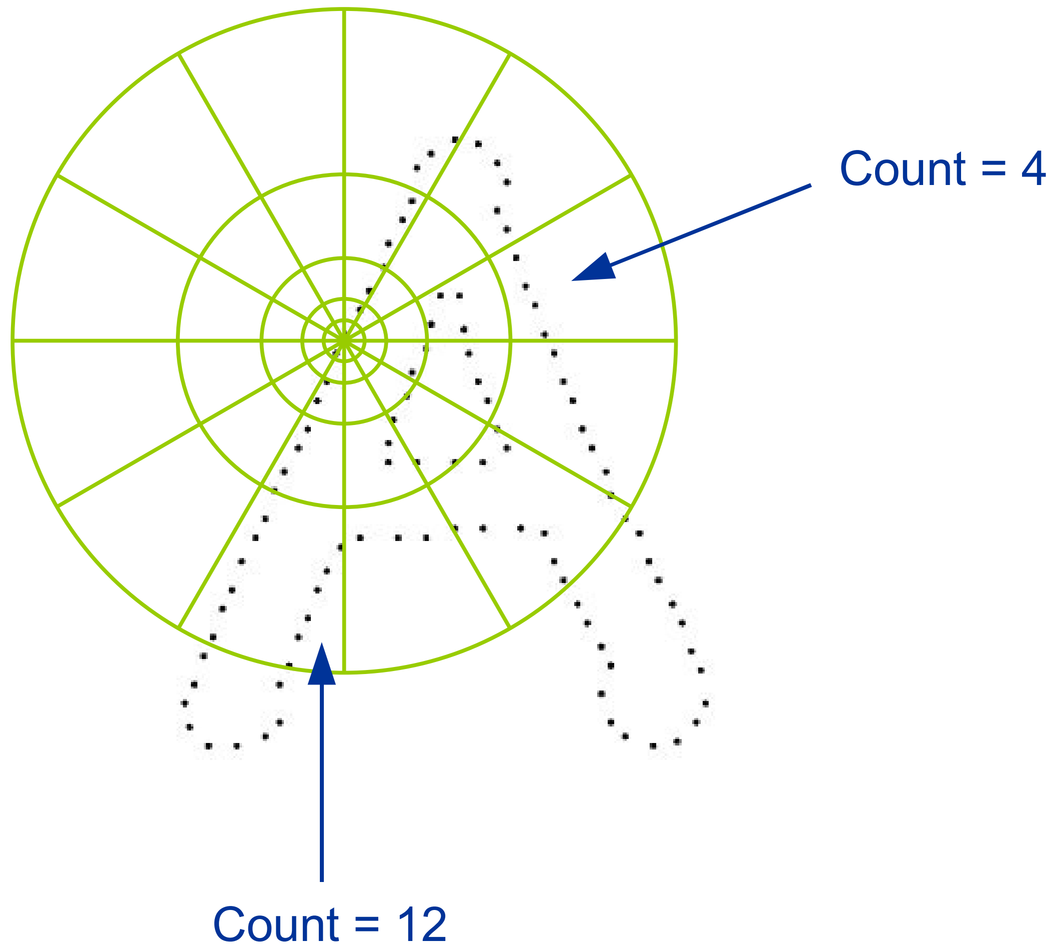
- A shape is represented by a set of points sampled from the edges of the object



Shape Context

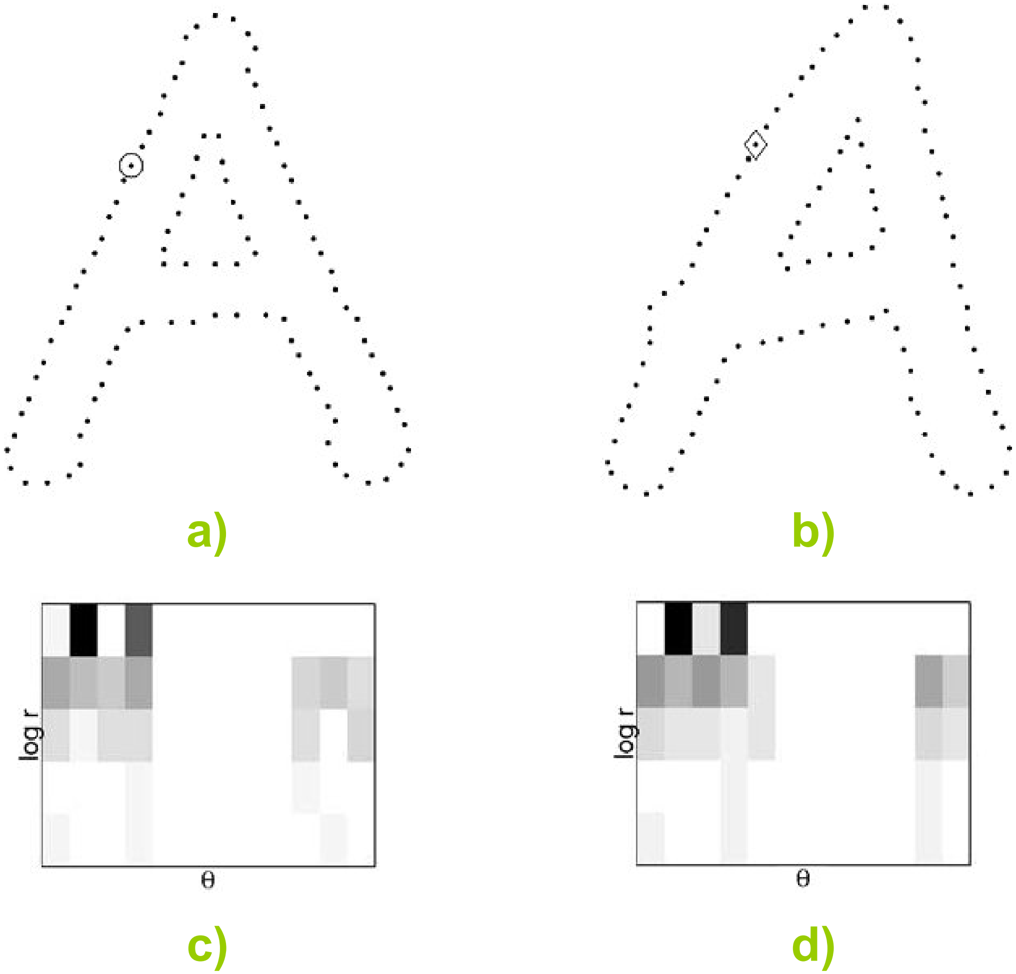
Shape Context: Log-Polar Histograms

Count the number of points inside each bin.



Slide from Belongie et al.

Example: Shape Contexts



Images from Belongie et al.

Shape Context

Point Correspondences

- Compute matching costs $C(p_i, p_j)$ using Chi Squared distance:

$$C(p_i, p_j) = \frac{1}{2} \sum_{k=1}^{K} \frac{[h_i(k) - h_j(k)]^2}{h_i(k) + h_j(k)}$$

- Minimize the total cost of matching, such that matching is 1-to-1

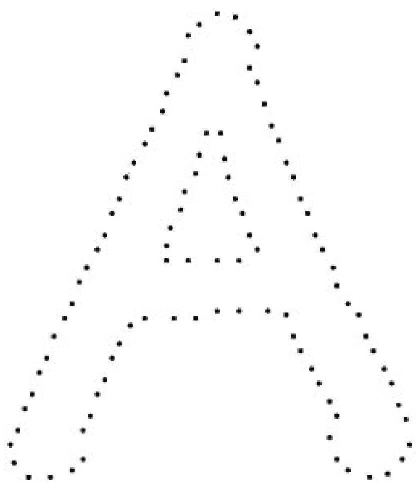
$$H(\pi) = \sum_i C\left(p_i, q_{\pi(i)}\right)$$

[Jonker & Volgenant, 1987]

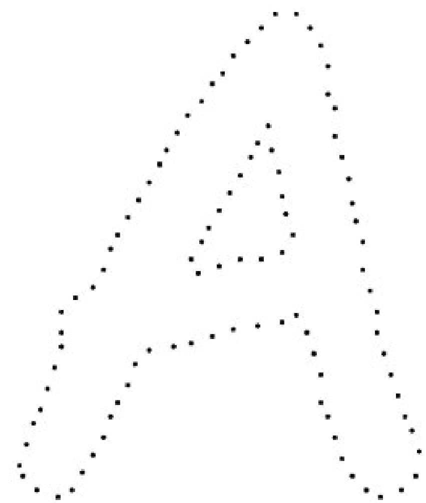
Slide from Belongie et al.

Shape Context

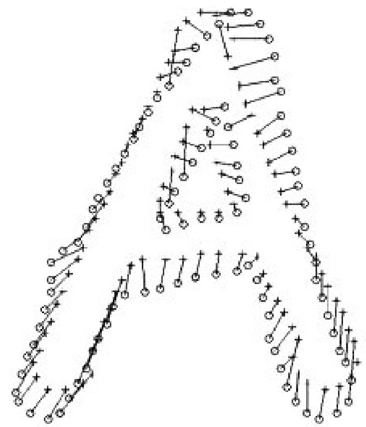
Example: Point Correspondences



a)



b)

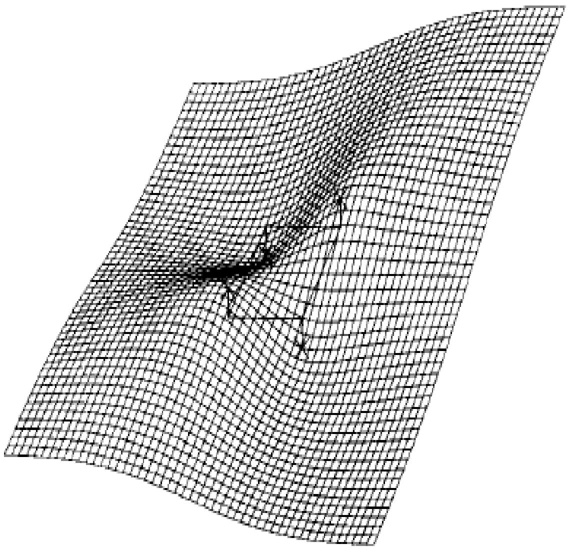


c)

Shape Context

Thin Plate Spline Model

- The name “thin plate spline” refers to a physical analogy involving the bending of a thin sheet of metal
- The 2D generalization of the 1D cubic spline
- Contains the affine model as a special case



Shape Context

Minimizing Bend Energy

- The Thin Plate Spline interpolation has the form:

$$f(x,y) = \underbrace{a_1 + a_x x + a_y y}_{\text{global affine transform}} + \underbrace{\sum_{i=1}^{n} w_i U\left(\left|\left|(x_i, y_i) - (x,y)\right|\right|\right)}_{\text{local non-linear transformations}}$$

where, $U(r) = r^2 \log r^2$

- Select a and w to minimize the bend energy:

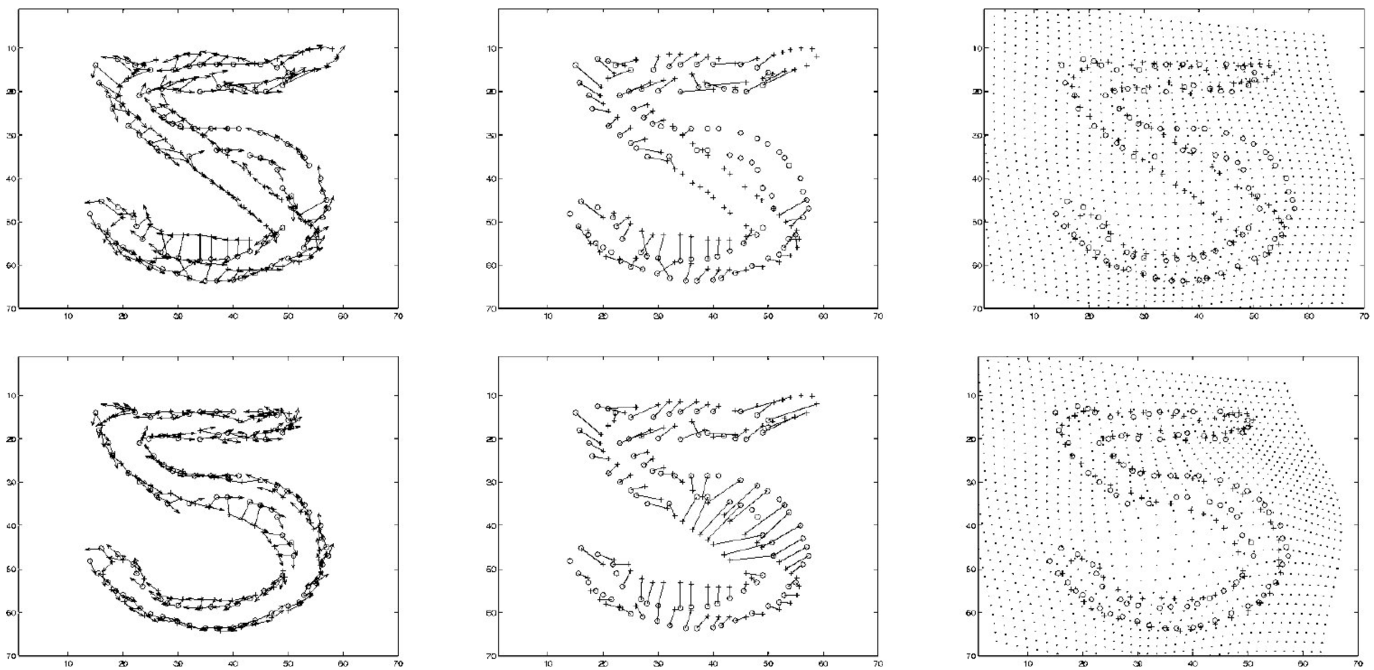
$$I(f) = \int\int_{\mathbb{R}^2} \left(\frac{\partial^2 f}{\partial x^2}\right)^2 + 2\left(\frac{\partial^2 f}{\partial x \partial y}\right)^2 + 2\left(\frac{\partial^2 f}{\partial y^2}\right)^2 dxdy$$

Shape Context

Example: Matching and Transformation

a)

b)



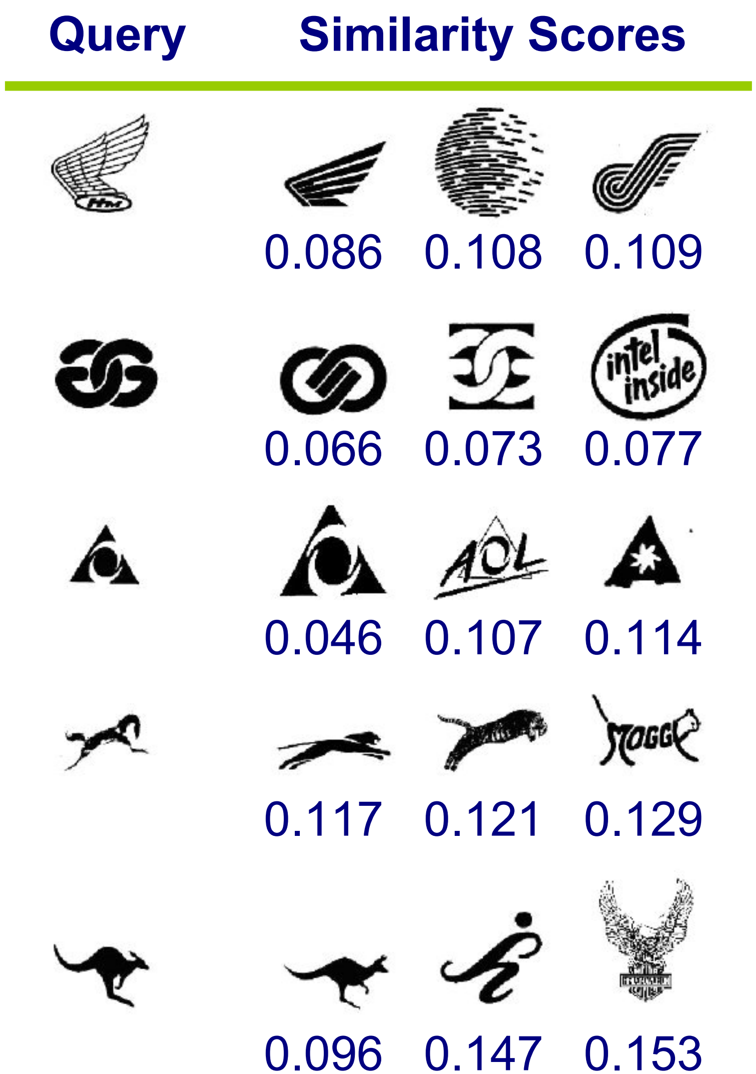
Images from Belongie et al.

Terms in Similarity Score

- Shape Context difference, $\mathcal{D}_{sc}$
- Local Image appearance difference, $\mathcal{D}_{ac}$
  - Orientation
  - Gray-level correlation in Gaussian window
  - … (many more possible)
- Bending energy, $\mathcal{D}_{be}$

$$\mathcal{D}_{sc} + 1.6 * \mathcal{D}_{ac} + 0.3 * \mathcal{D}_{be}$$

Shape Context Results



Images from Belongie et al.

Outline

- Shape Distance and Correspondence
  - Hausdorff Distance
  - Shape Context
  - ***Inner Distance***
- Hierarchical Approach
  - Hierarchical Matching
- Machine Learning Approach
  - Boundary Fragment Model

Inner Distance

Using the Inner-Distance for Classification of Articulated Shapes

2005

H. Ling and D. Jacobs

Inner Distance

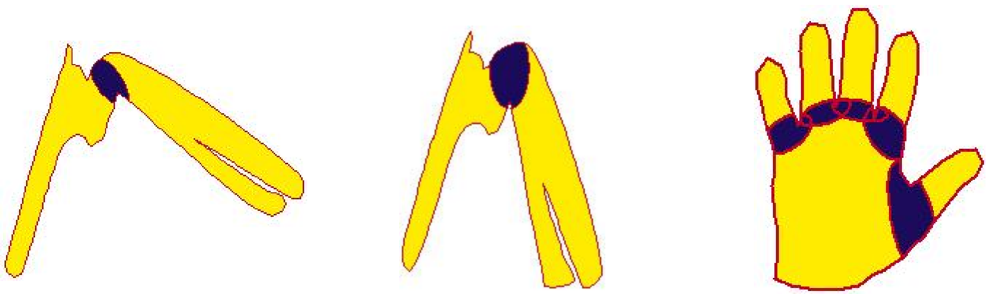
Overview

- Its difficult to capture the part structure of complex shapes with existing shape matching methods
- Replace euclidean distance with the inner-distance
  - Insensitive to shape articulations
  - Often more discriminative for complex shapes
- An extension to shape contexts

Inner Distance

Model of Articulated Objects

1) An object can be decomposed into a number of parts
2) Junctions between parts are relatively small with respect to the parts they connect
3) Articulation on the object is rigid with respect to any part, but can be non-rigid on the junctions
4) An object that has been articulated can be articulated back to its original form

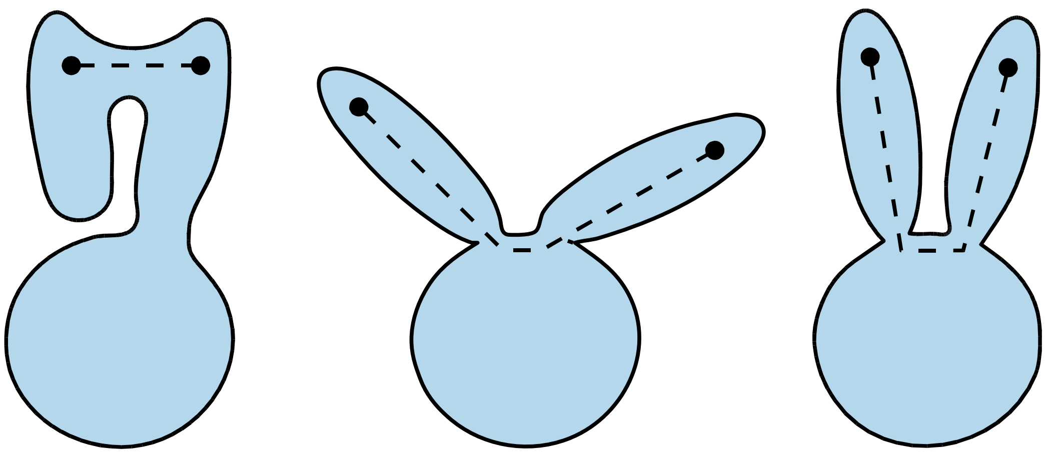


Images from Ling and Jacobs

Inner Distance

The Inner-Distance

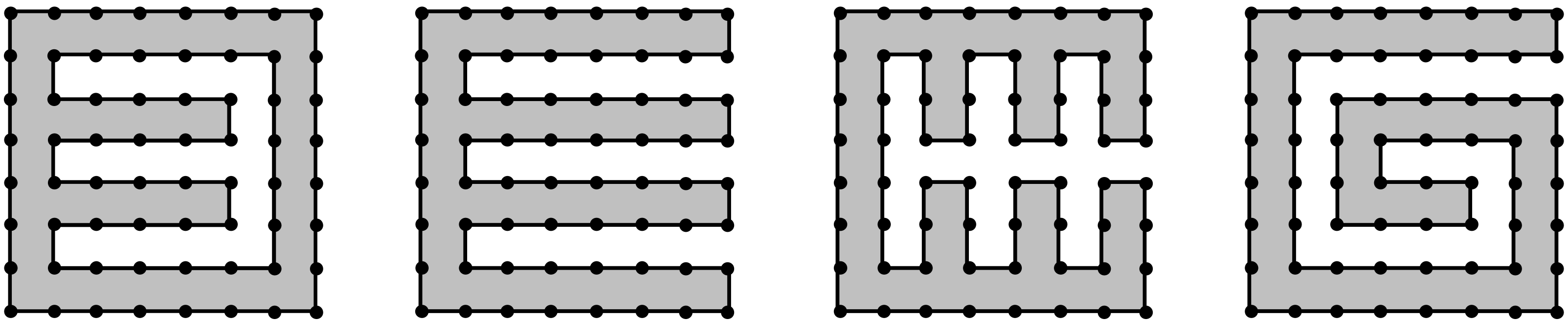
- The length of the shortest path between landmark points within the shape silhouette
- For convex shapes, the inner-distance reduces to the Euclidean distance
- Inner-Distance changes only due do deformations of the junctions



Images from Ling and Jacobs

Inner Distance

Inner-Distance vs Euclidean Distance



Images from Ling and Jacobs

Inner Distance

Computing the Inner-Distance

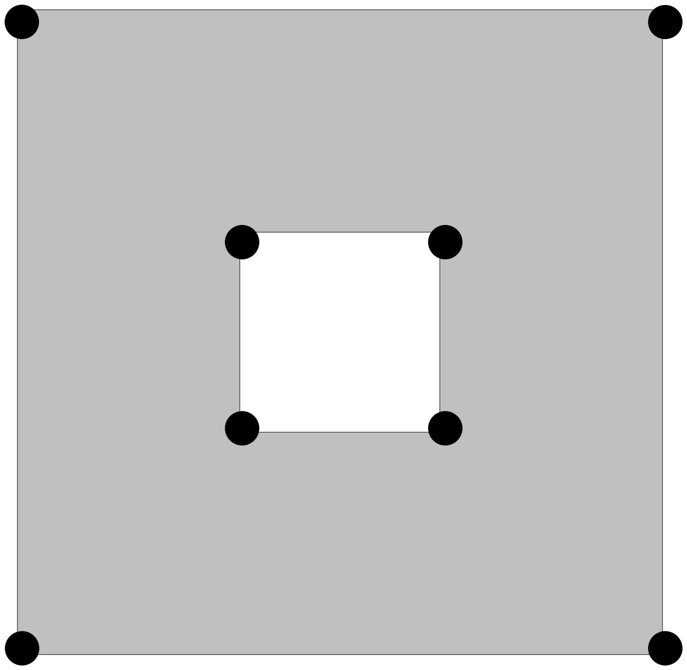
1) Build a graph on the sampled points
   - For each pair of points x,y.
     1. If line segment between them existed entirely within the object
     2. Build an edge connecting x and y with weight

$$w = ||x - y||$$

2) Apply a shortest path algorithm on the graph

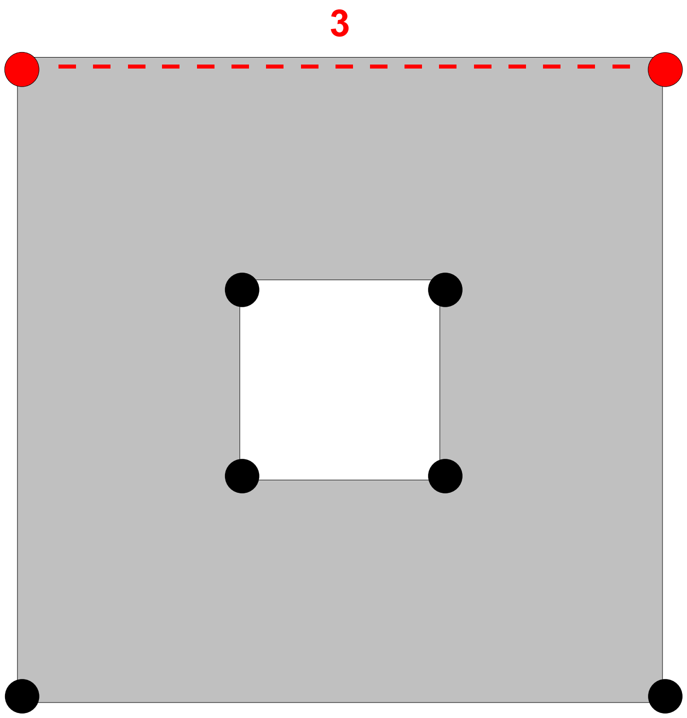
Inner Distance

Example: Inner Distance



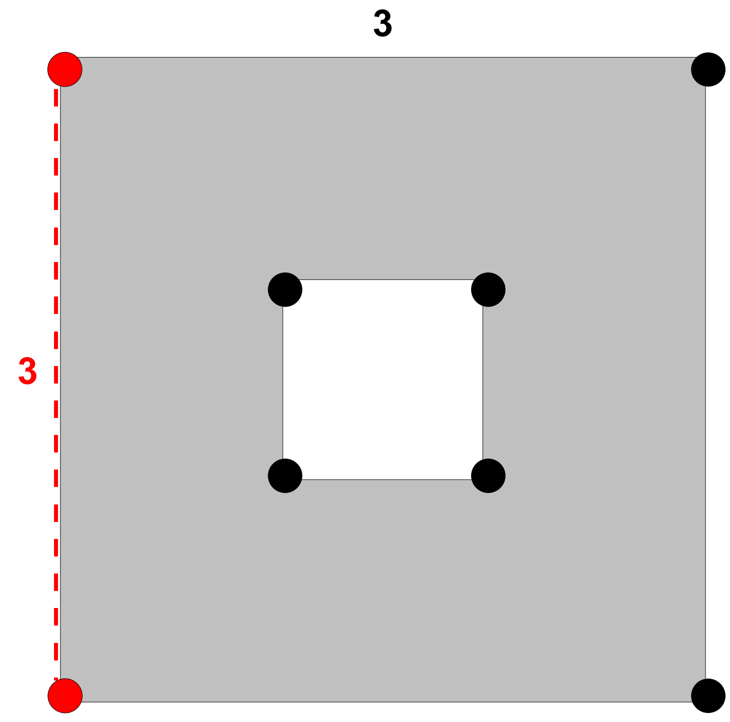
Inner Distance

Example: Inner Distance



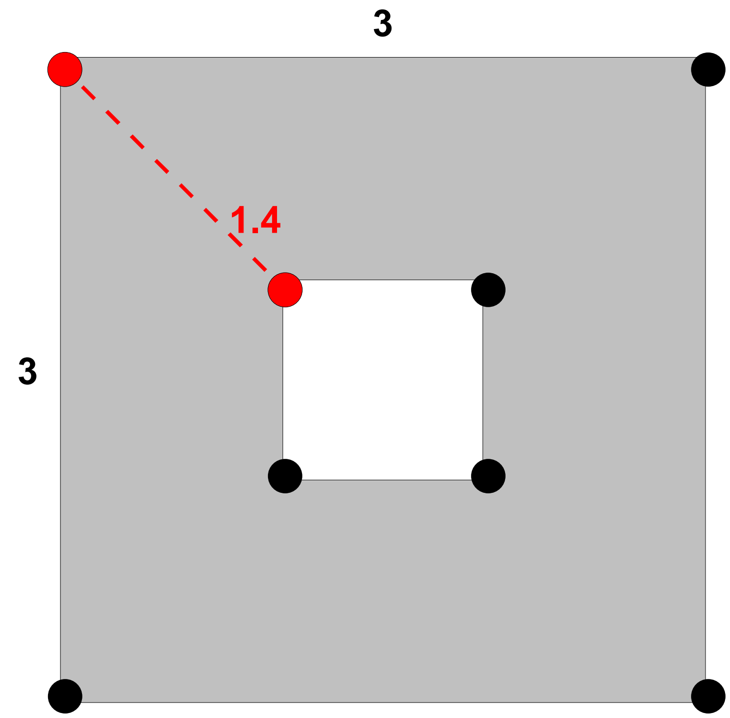
Inner Distance

Example: Inner Distance



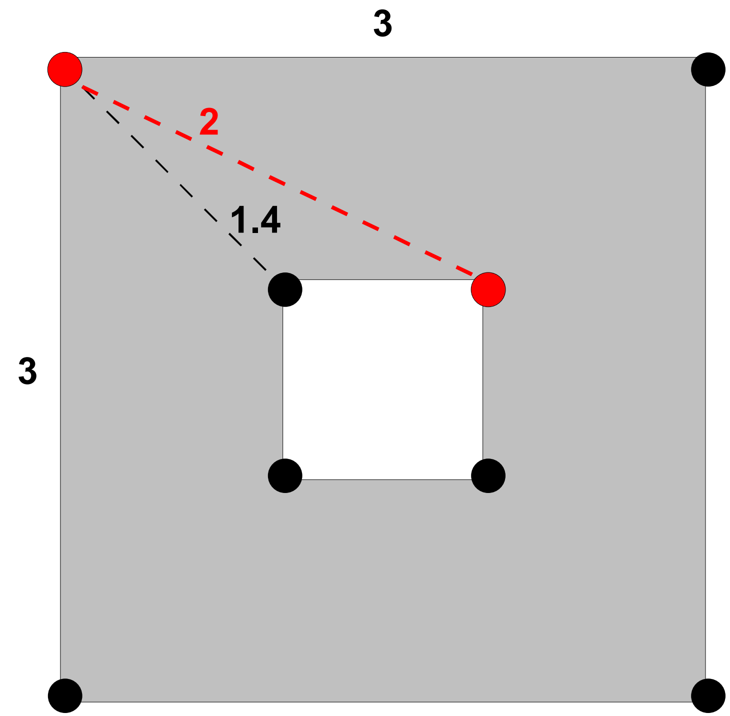
Inner Distance

Example: Inner Distance



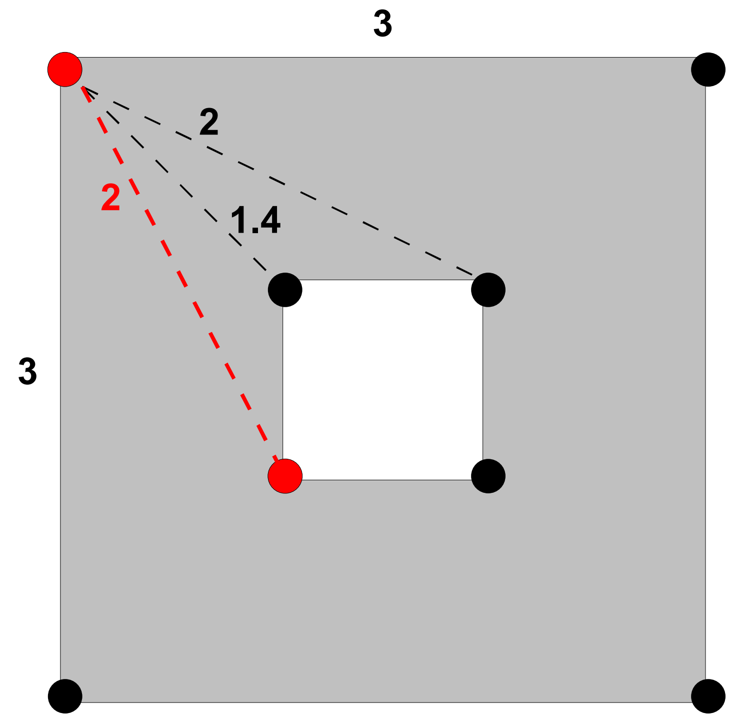
Inner Distance

Example: Inner Distance



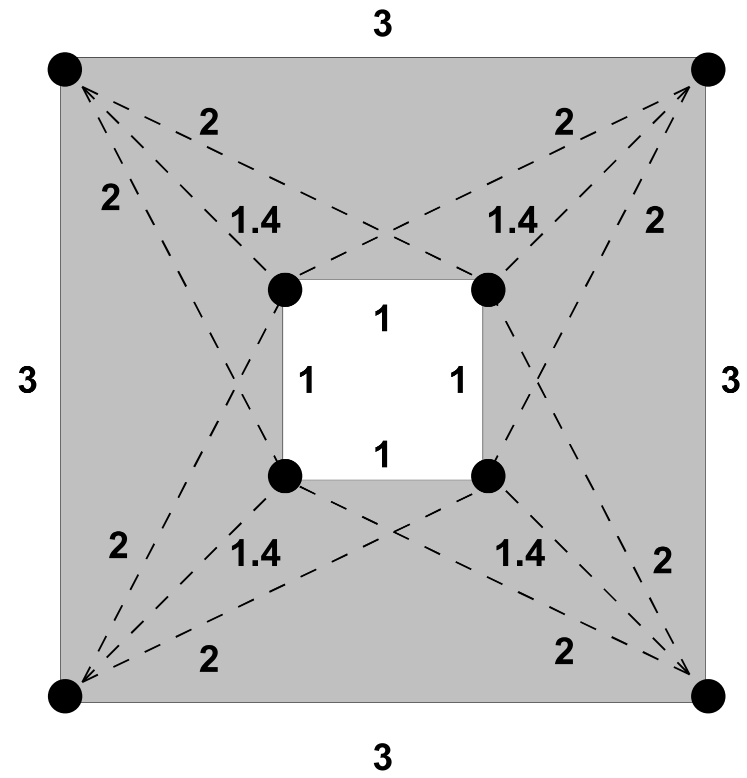
Inner Distance

Example: Inner Distance



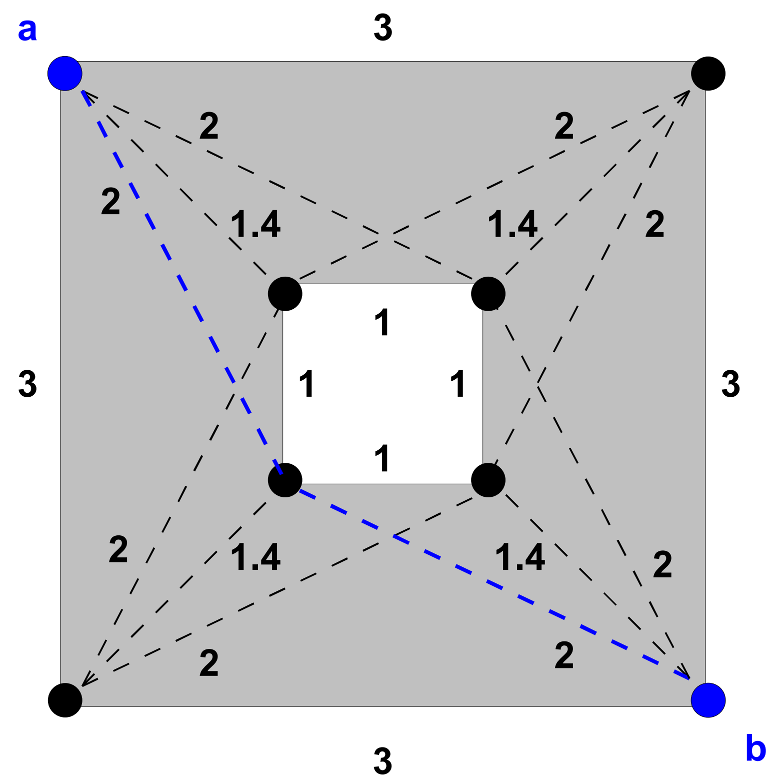
Inner Distance

Example: Inner Distance



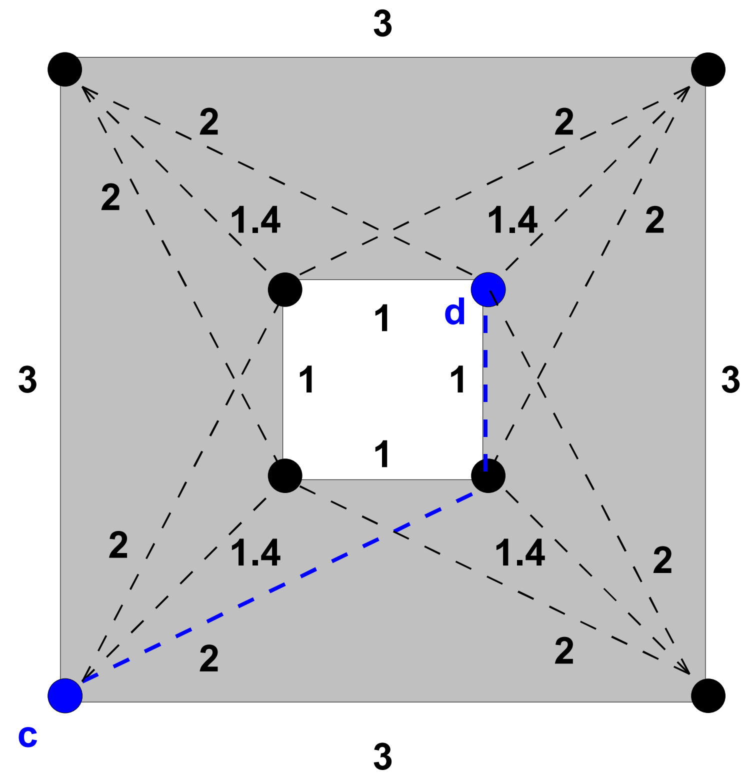
Inner Distance

Example: Inner Distance



$$d\,(a, b) = 4$$

Example: Inner Distance

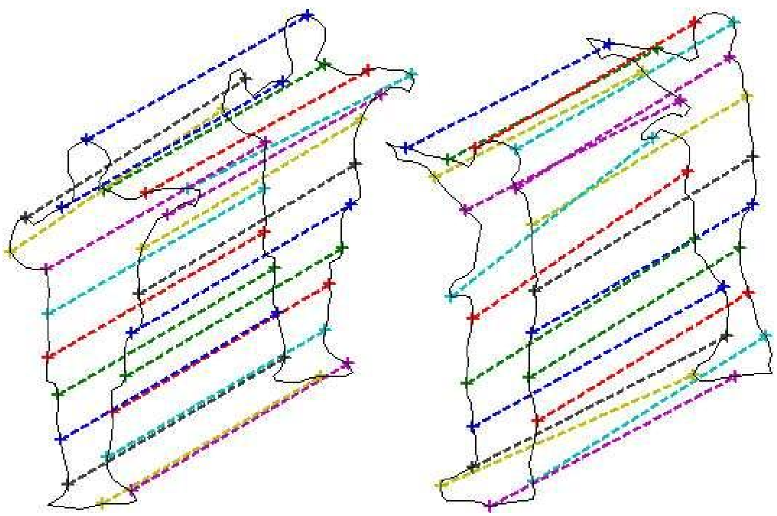


$$d\,(c, d) = 3$$

Inner Distance

An Extension to Shape Contexts

- Redefine the bins with inner-distance
  - Euclidean distance is replaced directly with the inner-distance

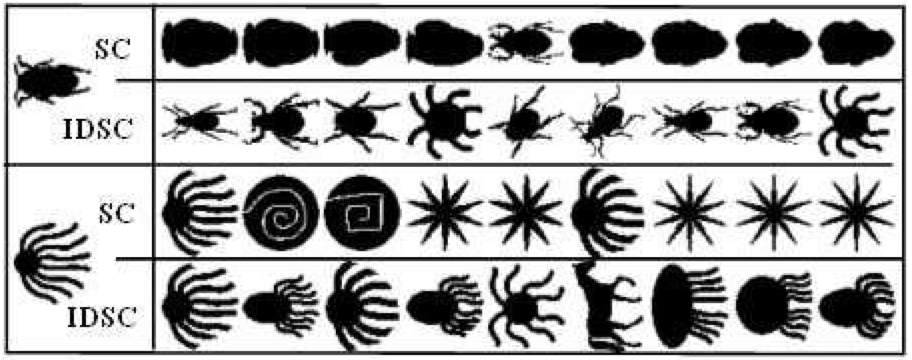


Images from Ling and Jacobs

Results (MPEG7 dataset)

Algorithm	CSS	Visual Parts	SC
Score	75.44%	76.45%	76.51%

Algorithm	Curve Edit	Gen. Model	IDSC
Score	78.17%	80.03%	85.40%



Shape Tree

Outline

- Shape Distance and Correspondence
  - Hausdorff Distance
  - Shape Context
  - Inner Distance
- Hierarchical Approach
  - ***Hierarchical Matching***
- Machine Learning Approach
  - Boundary Fragment Model

Shape Tree

Hierarchical Matching of Deformable Shapes

2007

P. Felzenszwalb and J. Schwartz

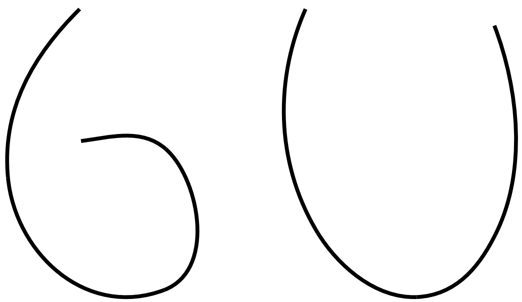
Shape Tree

Overview

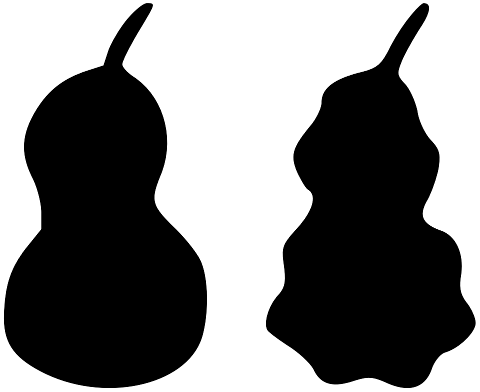
- Use hierarchical representation to capture shape information at multiple levels of resolution
- Capture global properties by compositing adjacent curve matches

Shape Tree

Local vs. Coarse Features



a)

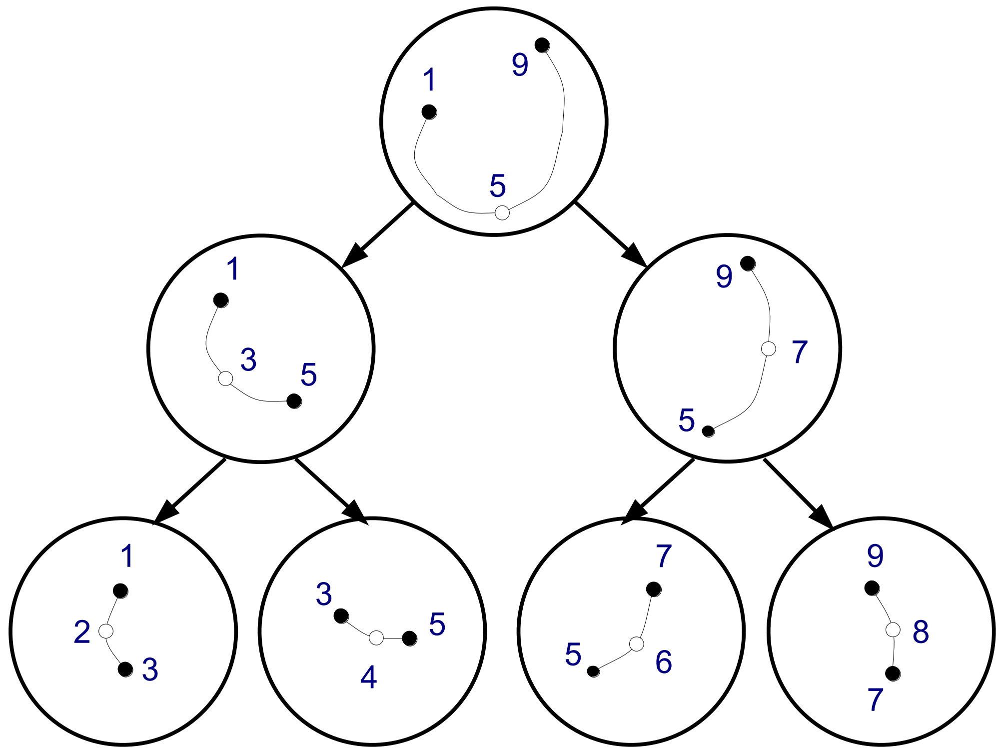


b)

Images from Felzenszwalb and Schwartz

Shape Tree

The Shape-Tree



Images from Felzenszwalb and Schwartz

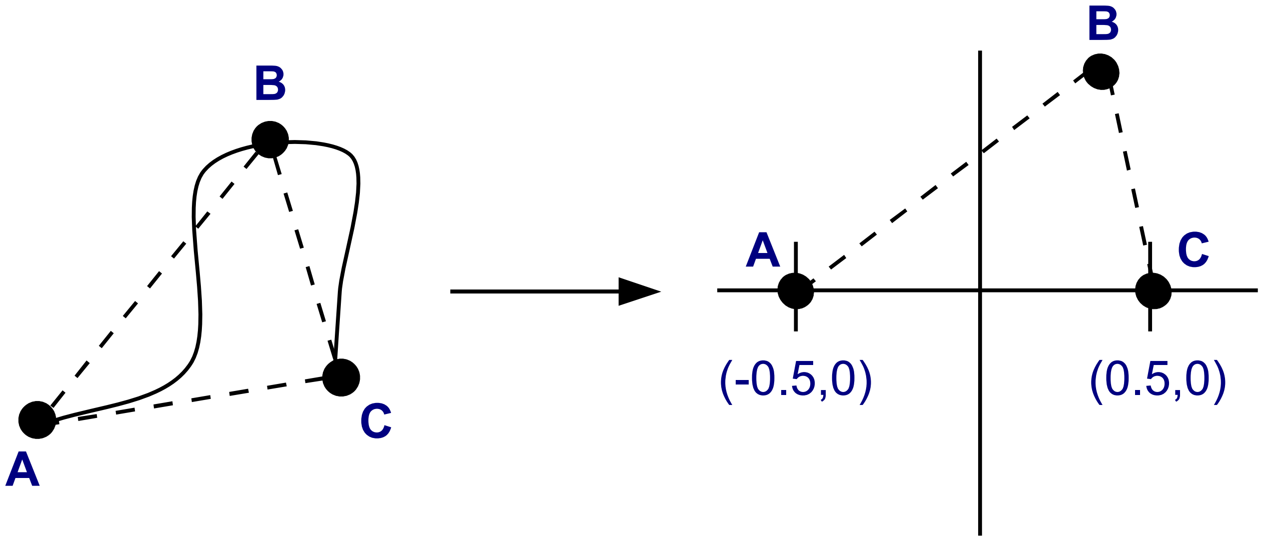
Shape Tree

Bookstein Coordinates

- Encode the relative positions of 3 points as a point in the plane
- A simple way to represent the relative location of a midpoint in the shape tree
- Given 3 points there exists a unique similarity transformation which maps:
  - P_1 to (-0.5, 0)
  - P_2 to (0.5, 0)
  - P_3 to the Bookstein coordinate

Shape Tree

Relative Locations



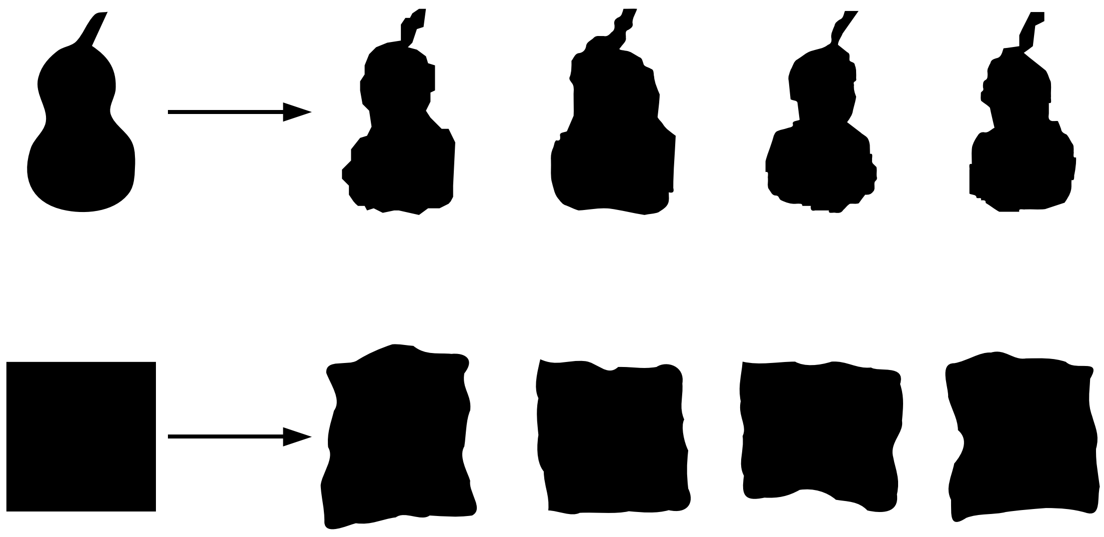
- Bookstein coordinates for representing B | A,C
- There exists a unique similarity transformation T taking:
  - A to (-0.5 , 0)
  - C to (0.5 , 0)
- We are interested in T(B)

Slide from Felzenszwalb and Schwartz

Shape Tree

Deformation model

- Independently perturb relative locations stored in a shape-tree
  - Reconstructed curve is perceptually similar to original
  - Local and global properties are preserved



Images from Felzenszwalb and Schwartz

Shape Tree

Distance Between Curves

- Given curves A and B
- Can't compare shape-trees for A and B built separately
- Fix shape-tree for A and look for map from points in A to points in B that doesn't deform the shape-tree much
- Efficient $\mathcal{O}\left(nm^3\right)$ DP algorithm, where

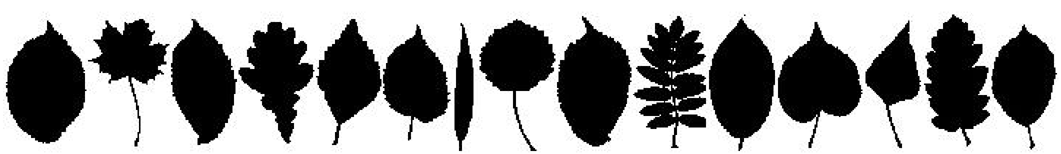
$$\left(n = |A|, m = |B|\right)$$

Shape Tree

Recognition Results

Swedish Leaf Dataset (15 species with 75 examples each)

Nearest Neighbor Classification			
Algorithm	**Shape-Tree**	Inner-Distance	Shape Context
Score	**96.28%**	94.13%	88.12%



MPEG7 Dataset

Bullseye Score				
Algorithm	**Shape-Tree**	Inner-Distance	Curve Edit	Shape Context
Score	**87.70%**	85.40%	78.14%	76.51%

Shape Tree

Matching in Cluttered Images

- Given M the model curve and C the set of curves in the image
- Use DP to match each curve in C to every subcurve of M
  - Running time is linear on total length of image contours and cubic in the length of the model
- Stitch partial matchings together to form longer matchings
  - Use compositional rule

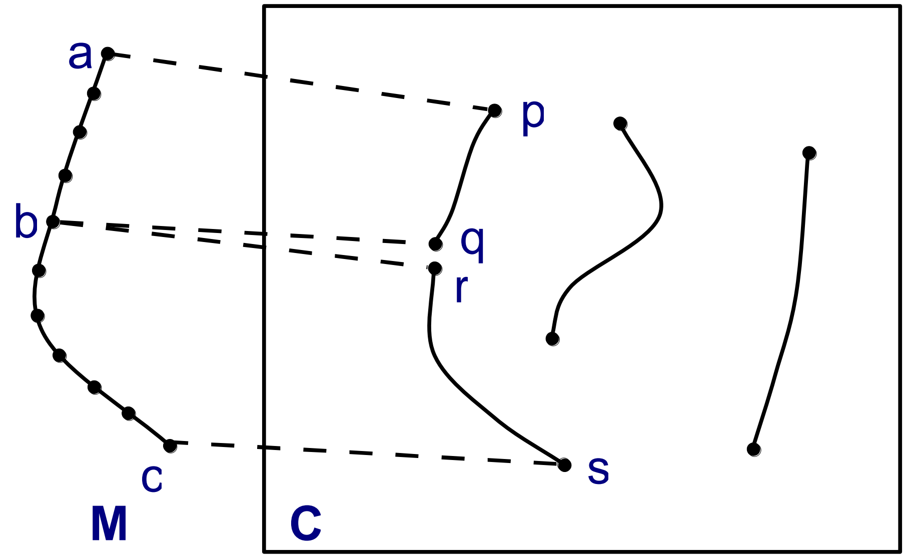
Shape Tree

Compositional Rule

If $||q-r|| < \tau$ compose Match([a,b], [p,q]) and Match([b,c], [r,s])

Match([a,b],[p,q]) = w_1

Match([b,c],[r,s]) = w_2



$$m = \frac{q+r}{2}$$

Match([a,c],[p,s]) =

$$w_1 + w_2 + dif\left((b|a,c), (m|p,s)\right)$$

Slide from Felzenszwalb and Schwartz

Example: Detection

Input Image

Edge Map

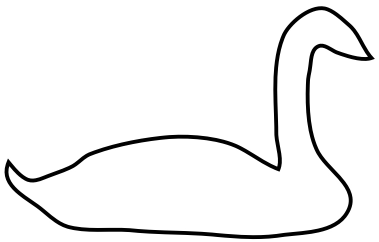
Contours

Detection

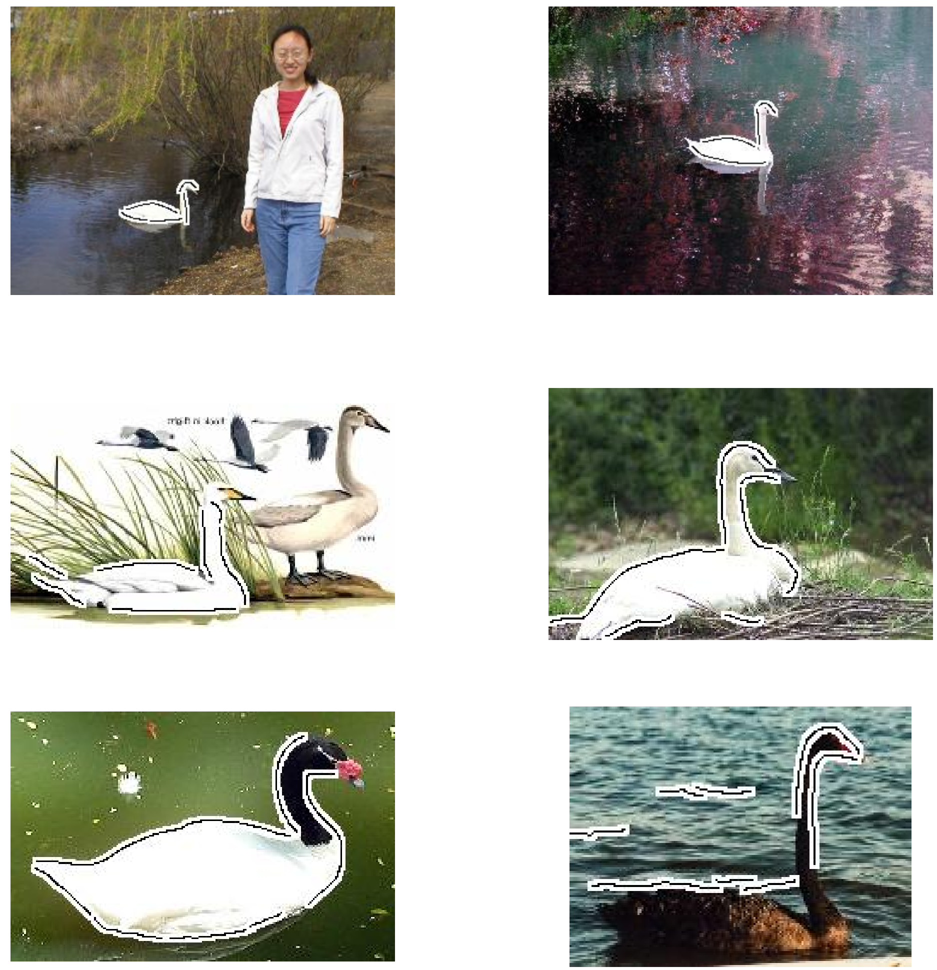
Images from Felzenszwalb and Schwartz

Shape Tree

Results



Model



Images from Felzenszwalb and Schwartz

Outline

- Shape Distance and Correspondence
  - Hausdorff Distance
  - Shape Context
  - Inner Distance
- Hierarchical Approach
  - Hierarchical Matching
- Machine Learning Approach
  - ***Boundary Fragment Model***

Boundary Fragment Model

A Boundary-Fragment-Model for Object Detection

2006

A. Opelt, A. Pinz, and A. Zisserman

Boundary Fragment Model

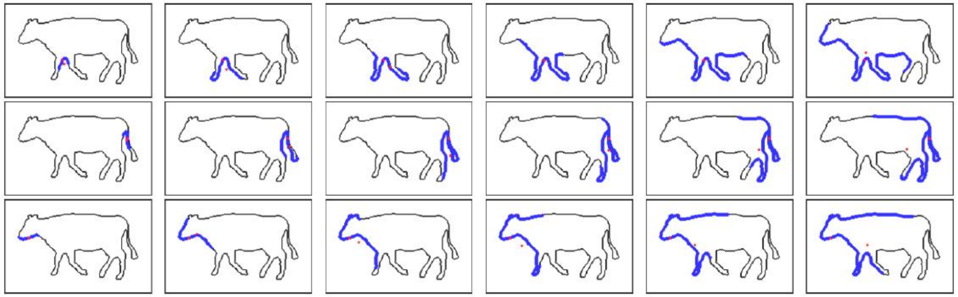
Overview

- Object class detection using object boundaries instead of salient image features
- A learning technique to extract discriminating boundary fragments
- Use boosting to select discriminative combinations of boundary fragments (weak detectors) to form a strong detector

Boundary Fragment Model

Learning Boundary Fragments

- Given
  - A training image set with the object delineated by a bounding box
  - A validation image set labeled with whether the object is absent or present, and the object's centroid
- From the edges of the training images identify fragments that:
  - Discriminate objects from the target category from other objects
  - Give a precise estimate of the object centroid

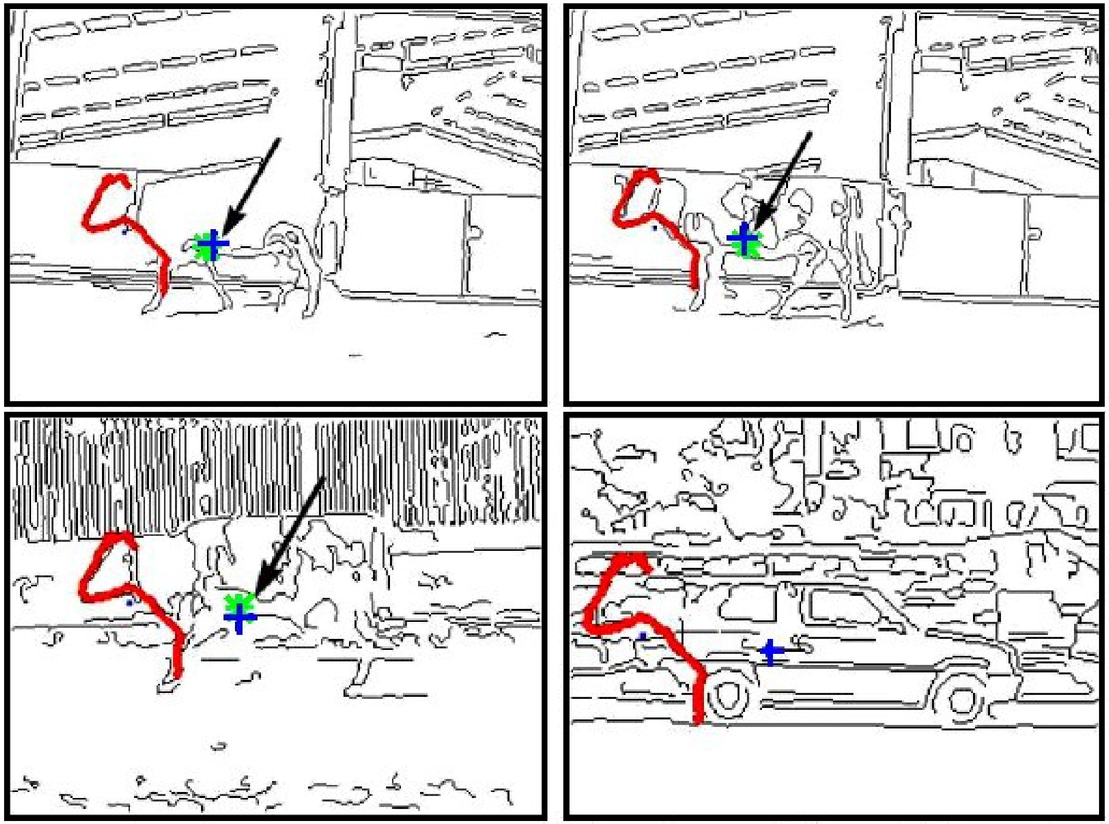


Boundary Fragment Model

Example: Good Boundary Fragment

+ = Estimated Centroid

* = Correct Centroid



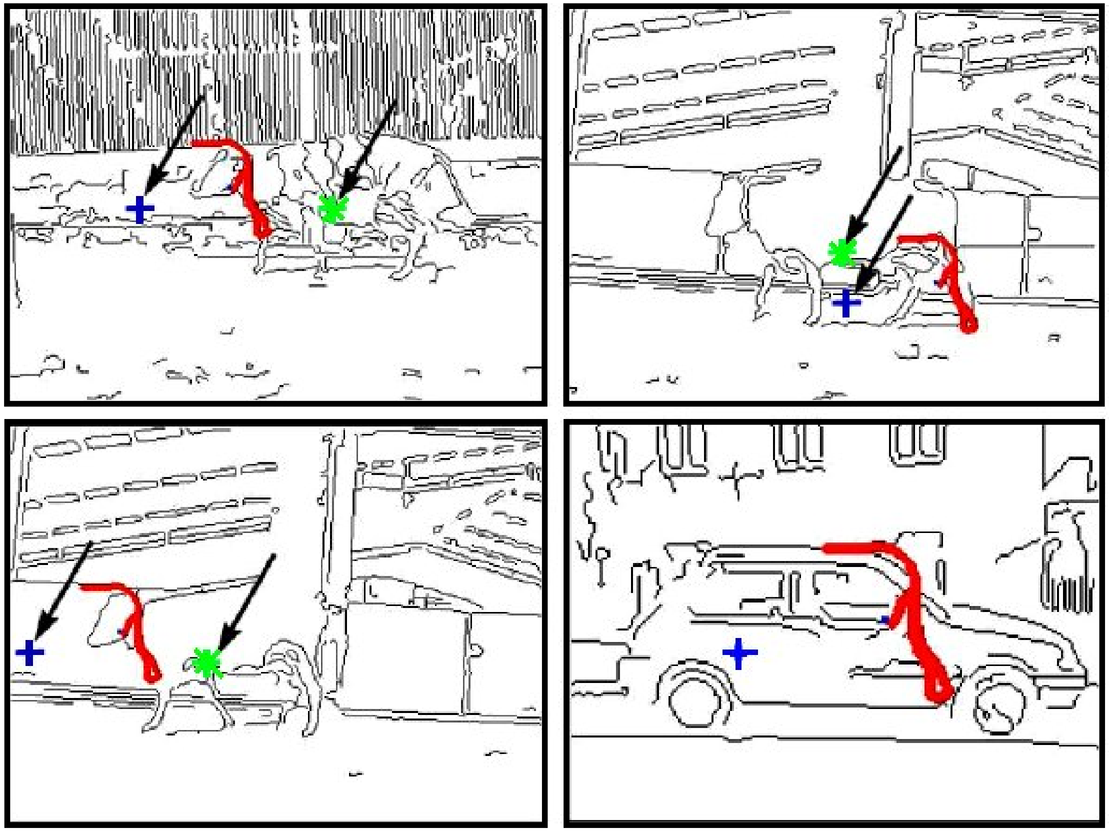
Images from Opelt et al

Boundary Fragment Model

Example: Poor Boundary Fragment

+ = Estimated Centroid

* = Correct Centroid

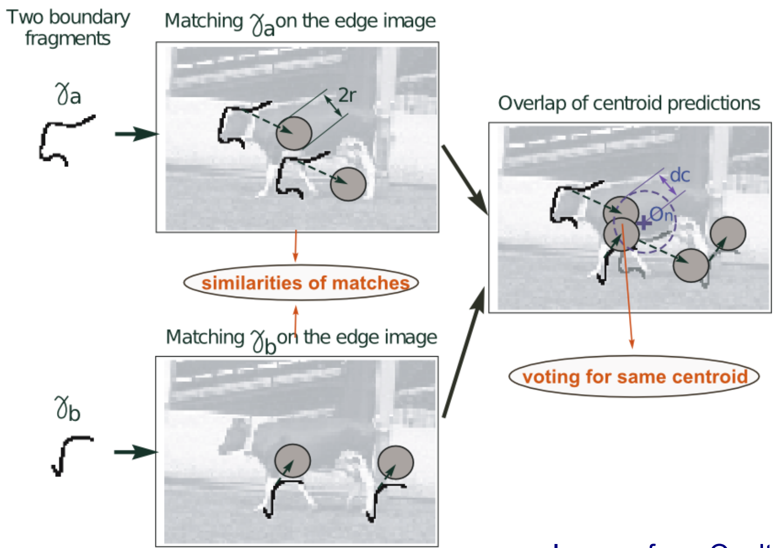


Images from Opelt et al

Boundary Fragment Model

Weak Detectors

- A weak detector is composed of k (typically 2 or 3) boundary fragments
- Detection should occur when
  - The k fragments match the image edges
  - The centroids concur
  - For positive images the centroid estimate agrees with the true object centroid



Images from Opelt et al

Boundary Fragment Model

Strong Detector

- Given weak detectors h_i
- Using AdaBoost
  - In each round find the weak detector that obtains the best detection results on the current weighting

$$H(I) = sign\left(\sum_{i=1}^{T} h_i(I)\, w_{h_i}\right)$$

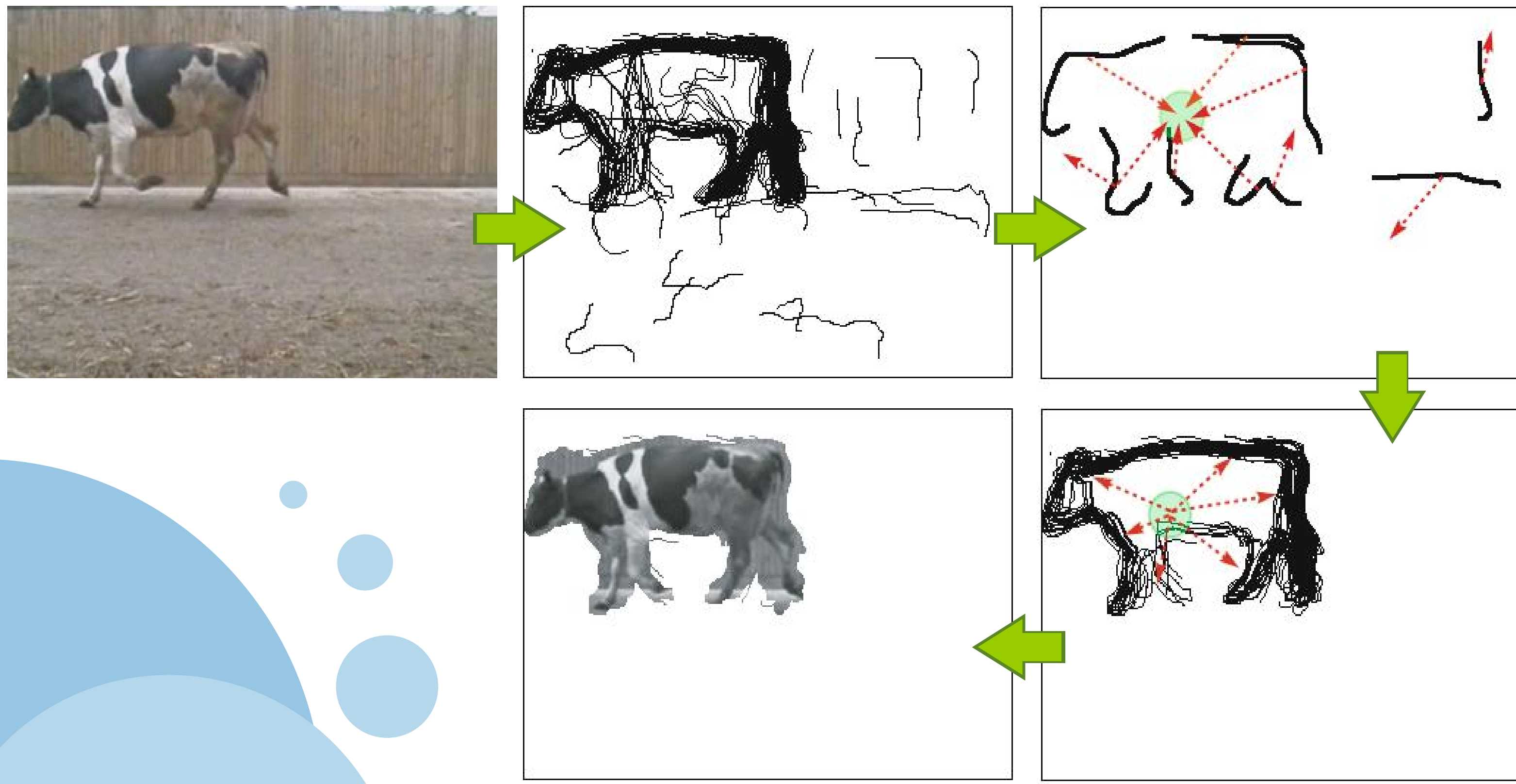
Boundary Fragment Model

Example: Detection and Segmentation

Original Image

All Matched Boundary Fragments

Centroid Voting on Subset of Fragments



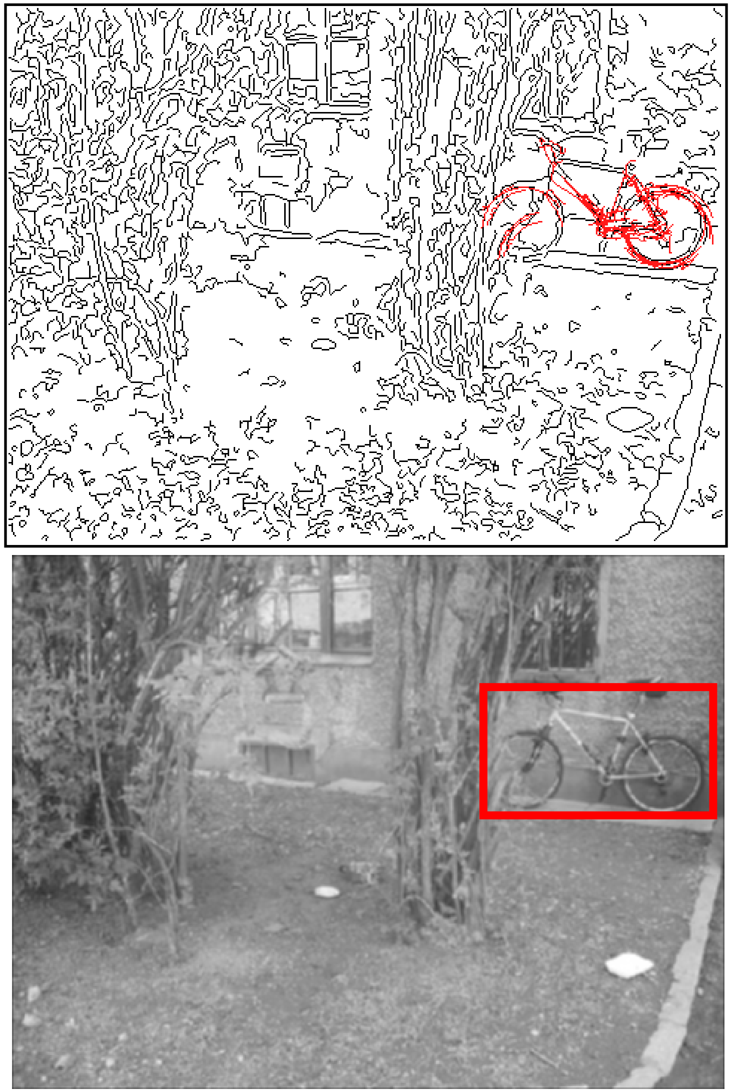
Detection and Segmentation

Backprojected Maximum

Images from Opelt et al

Boundary Fragment Model

Example: Detection and Localization



Images from Opelt et al

Boundary Fragment Model

Results

ROC Error Rate									
Algorithm	**BFM**	[12]	[22]	[25]	[2]	[3]	[14]	[26]	[28]
cars-rear	**0.50%**	8.80%	8.90%	21.40%	3.10%	2.30%	1.80%	9.80%	-
airplanes	**2.60%**	6.30%	11.10%	3.40%	4.50%	10.30%	-	17.10%	5.60%

Detection Error		
Algorithm	**BFM**	[18]
cars-rear	**2.25%**	6.10%

Breaking CAPTCHA

Recognizing Objects in Adversarial Clutter: Breaking a Visual CAPTCHA

2003

G. Mori and J. Malik

Breaking CAPTCHA

What is a CAPTCHA?

- **Definition:** Completely Automated Public Turing test to tell Computers and Humans Apart.
- Used to prevent automated SPAM.
- Also to read books!

Breaking CAPTCHA

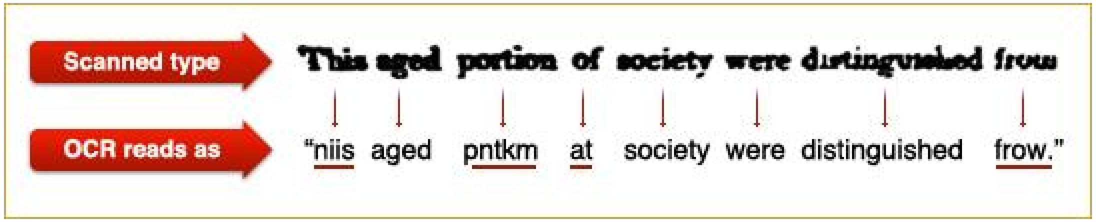
Applications of CAPTCHAs

- Preventing blog SPAM
- Protecting web site registration
- Protecting email addresses from scrapers
- Preventing dictionary attacks
- Online polling
- Blocking search engines
- Blocking email SPAM

Breaking CAPTCHA

Human Assisted OCR

- Roughly 60 million CAPTCHAs are solved by humans every day.
- Equivalent to about 150,000 hours of work.
- Why not use these CAPTCHAs for hard OCR tasks?



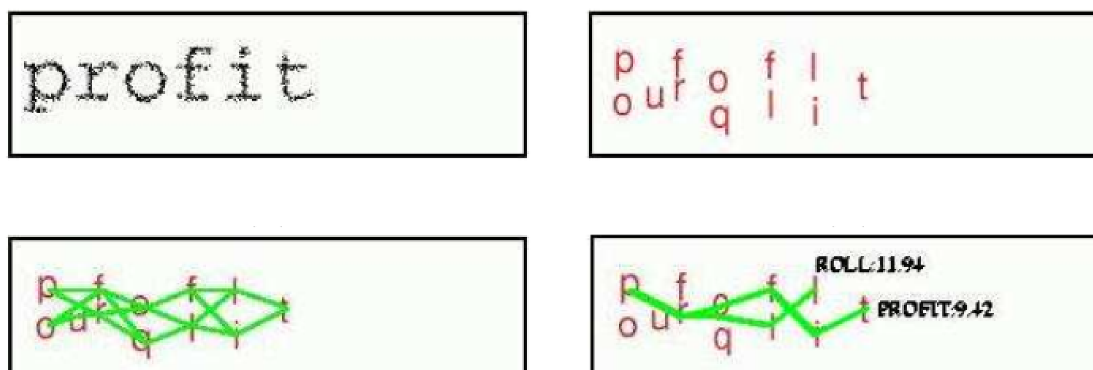
Why Break a CAPTCHA?

- CAPTCHAs help prevent SPAM
- They also offer challenges to the AI community
- A win-win situation:
  - If the CAPTCHA is not broken then SPAM is blocked
  - If it is broken then an AI problem has been solved

Breaking CAPTCHA

Approach 1

- Detect letters using the Shape Context approach
  - Extended so that the SC includes the dominant tangential direction of the edges in each bin
- Form a directed acyclic graph of the letters to find candidate words
- Choose the most likely word based on the average deformable match cost of the individual letters

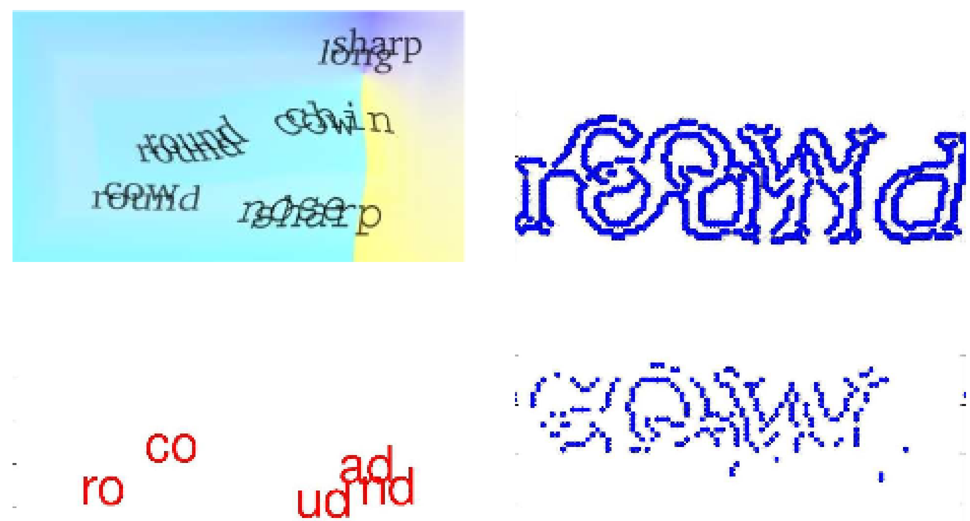


Images from Mori and Malik

Approach 2

- For harder CAPTCHAs matching on letter sized regions is to difficult

- Match on groups of letters instead

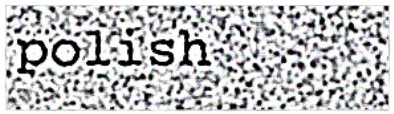


Images from Mori and Malik

Breaking CAPTCHA

Example: EZ-Gimpy

polish

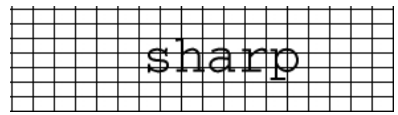


store

sound

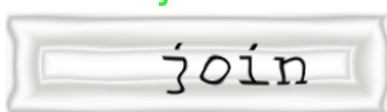
rice

east



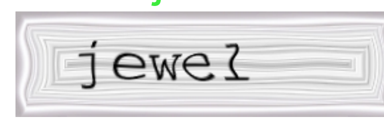
weight

join



sock

jewel



horse

space

mine

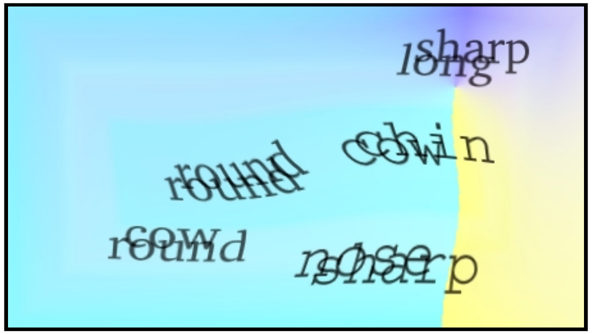
canvas

Images from Mori and Malik

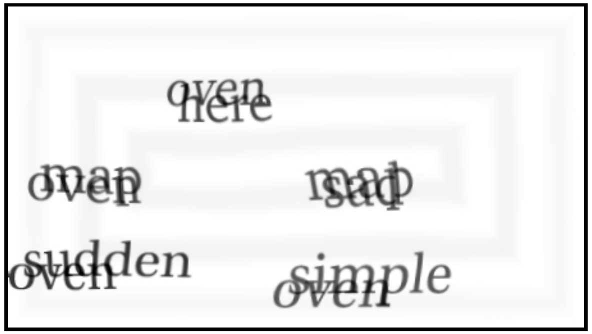
Breaking CAPTCHA

Example: 3 Word CAPTCHA

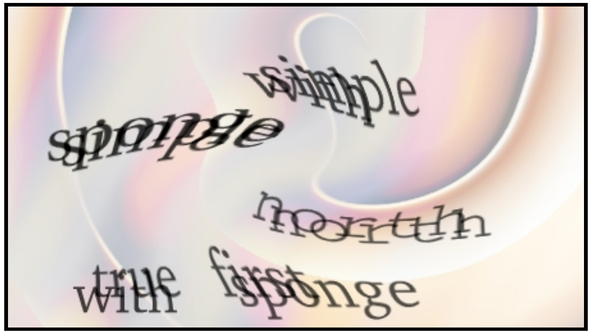
future key have



sharp round long



sudden apple oven



with true sponge

Images from Mori and Malik

Conclusion

Discussion Points

- How can shape matching be made more robust to clutter?
- What applications are not suitable for shape matching? Which are?
- How can methods like Shape Context take advantage of available training data?
- How can appearance and shape features be best combined?
- What other hard AI problems can be used as CAPTCHAs?