Shape Matching and Classification: Algorithms and Performance Evaluation

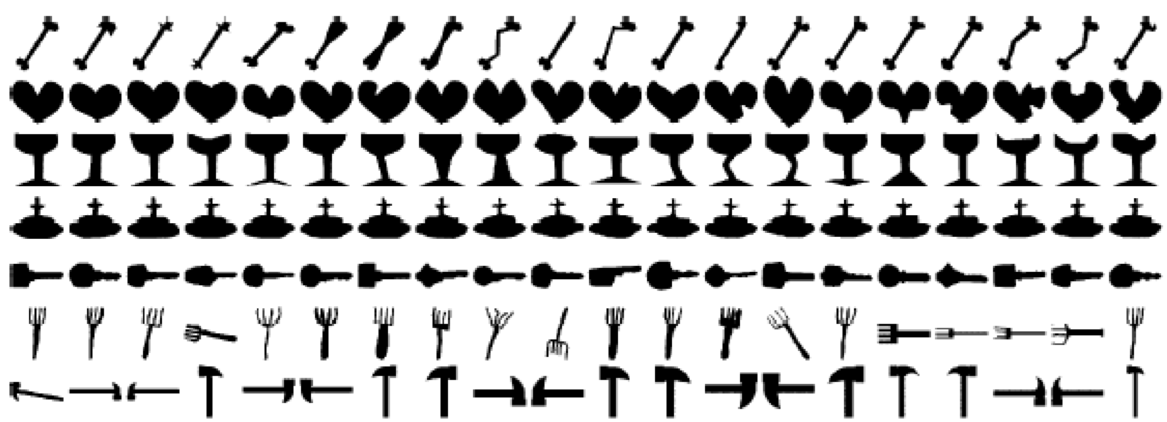
Song Wang

Department of Computer Science and Engineering

University of South Carolina

January 29, 2010

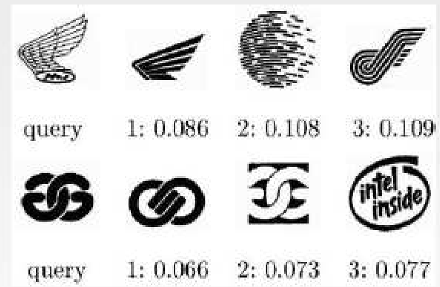
Shape matching for classification and retrieval



Shape matching examples

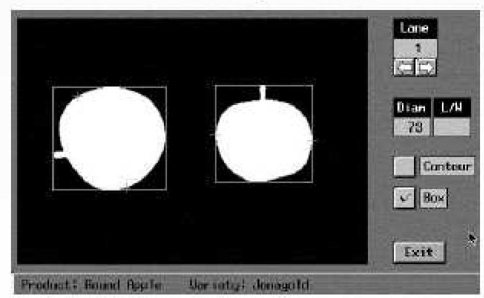
Hieroglyph Lookup

Trademark Lookup

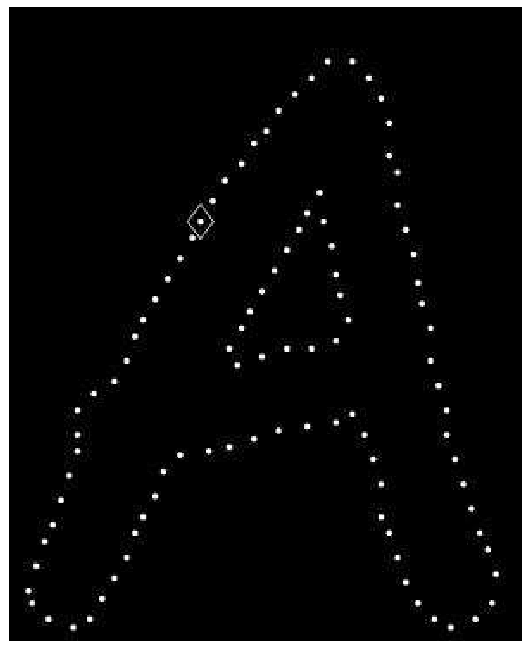


Fingerprint Matching

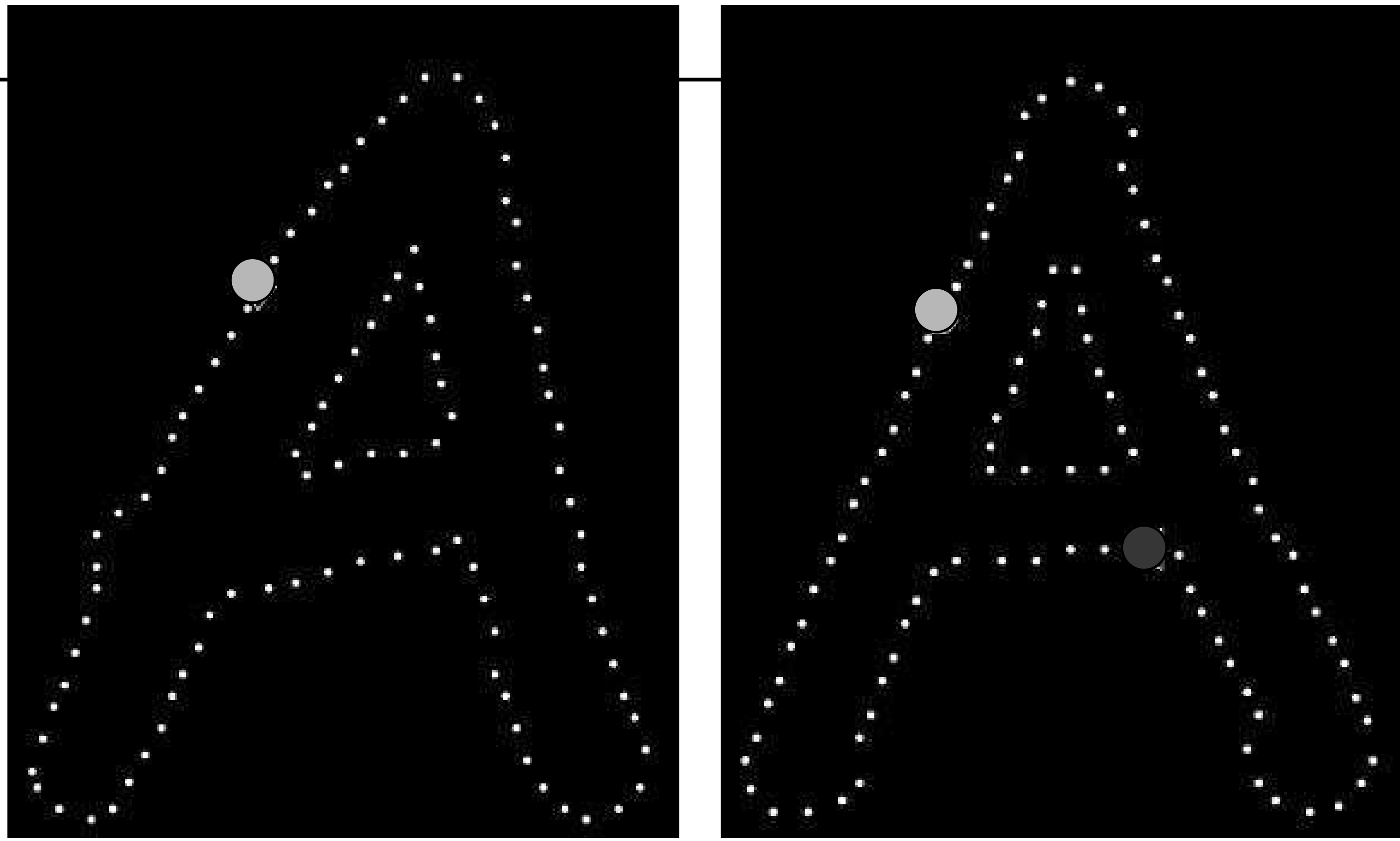
Fruit Inspection



Shape: Edge Points

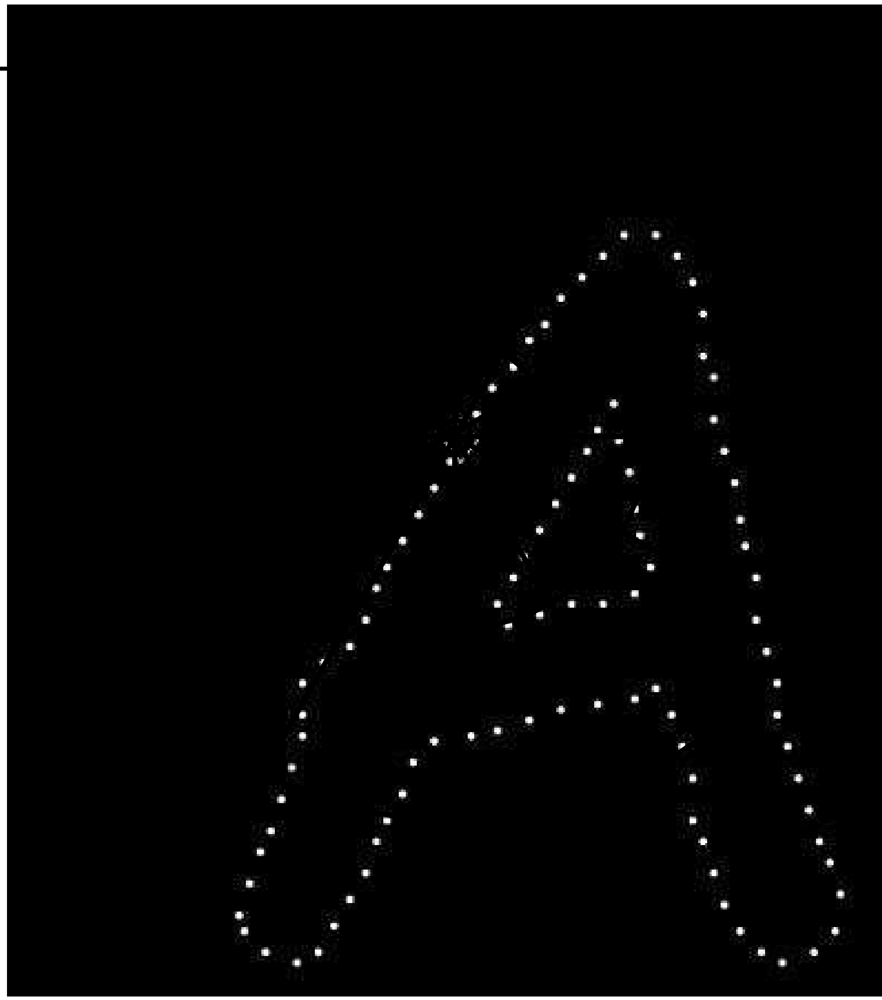


Comparing shapes



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

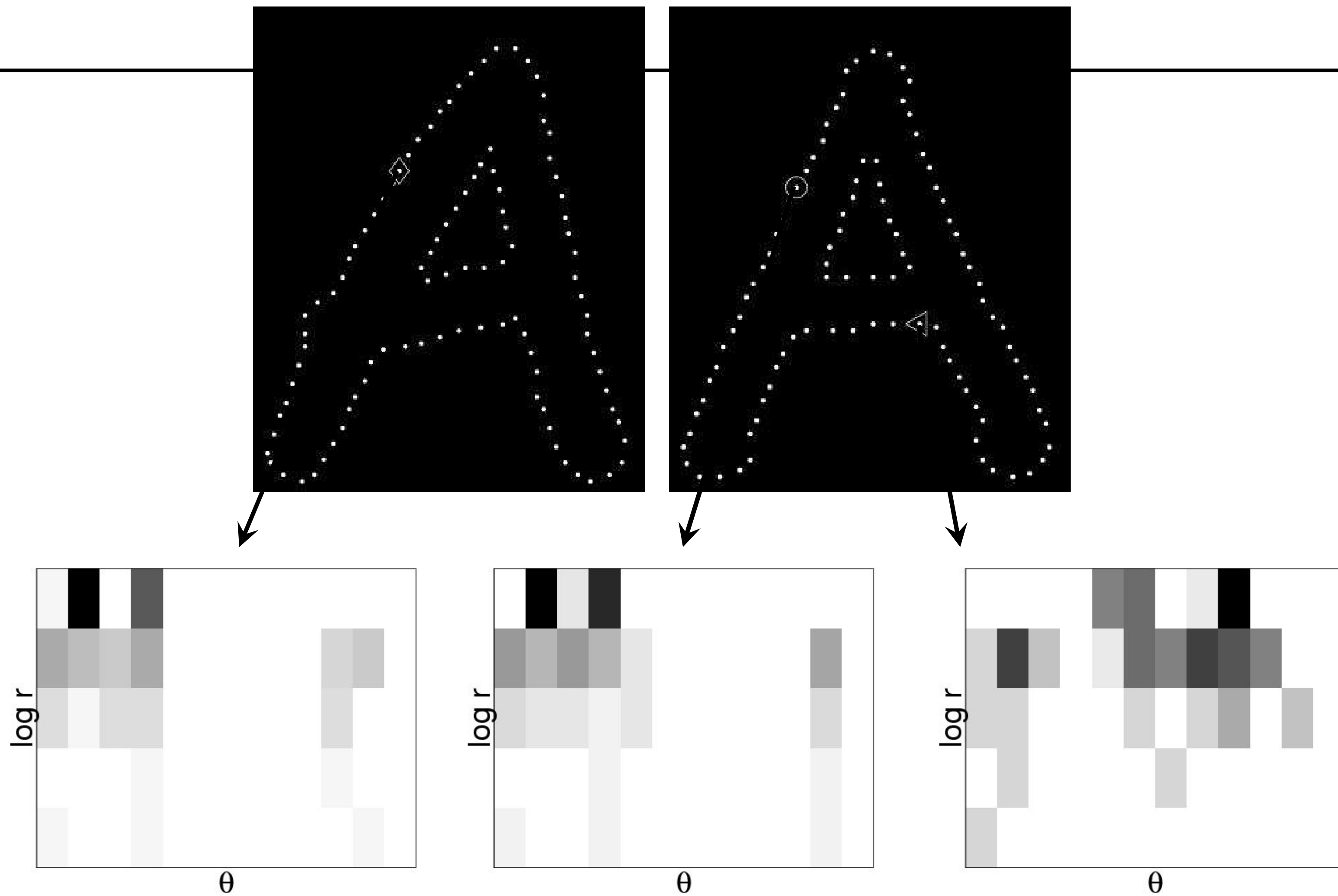
⋮

Count = 10

Compact representation of distribution of points relative to each point

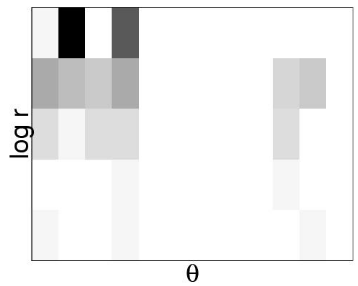
Shape context slides from Belongie et al.

Shape context descriptor

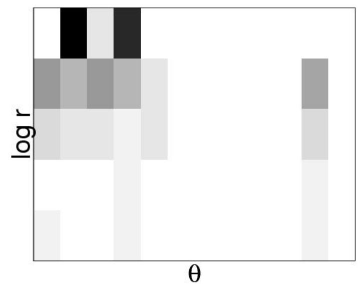


Comparing shape contexts

$h_i(k)$

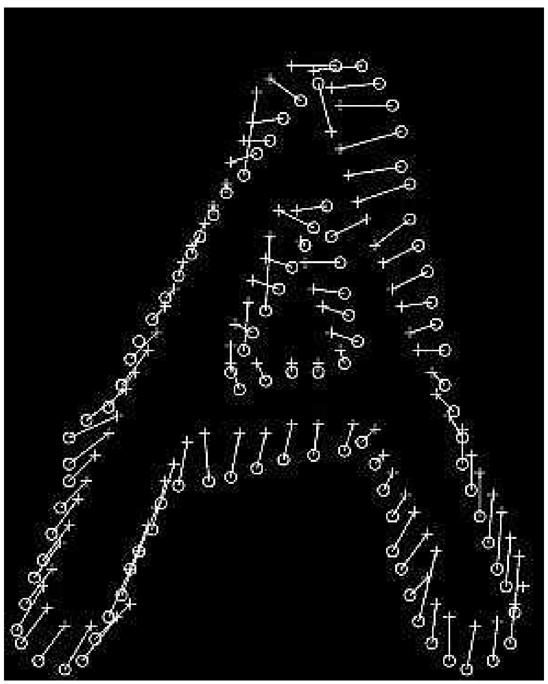


$h_j(k)$



Compute matching costs using Chi Squared distance:

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



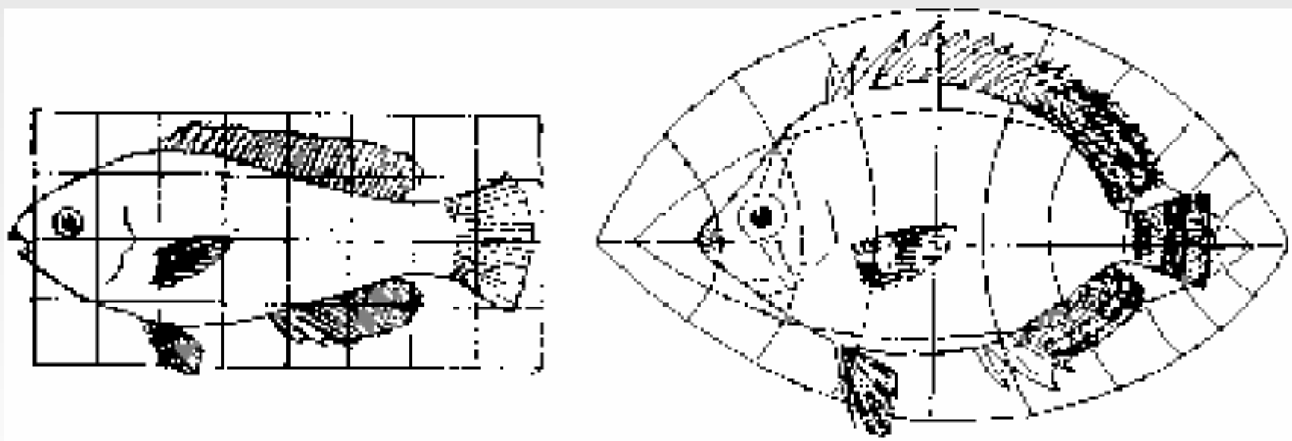
Recover correspondences by solving for least cost assignment, using costs C_{ij}

Bipartite Graph Matching

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

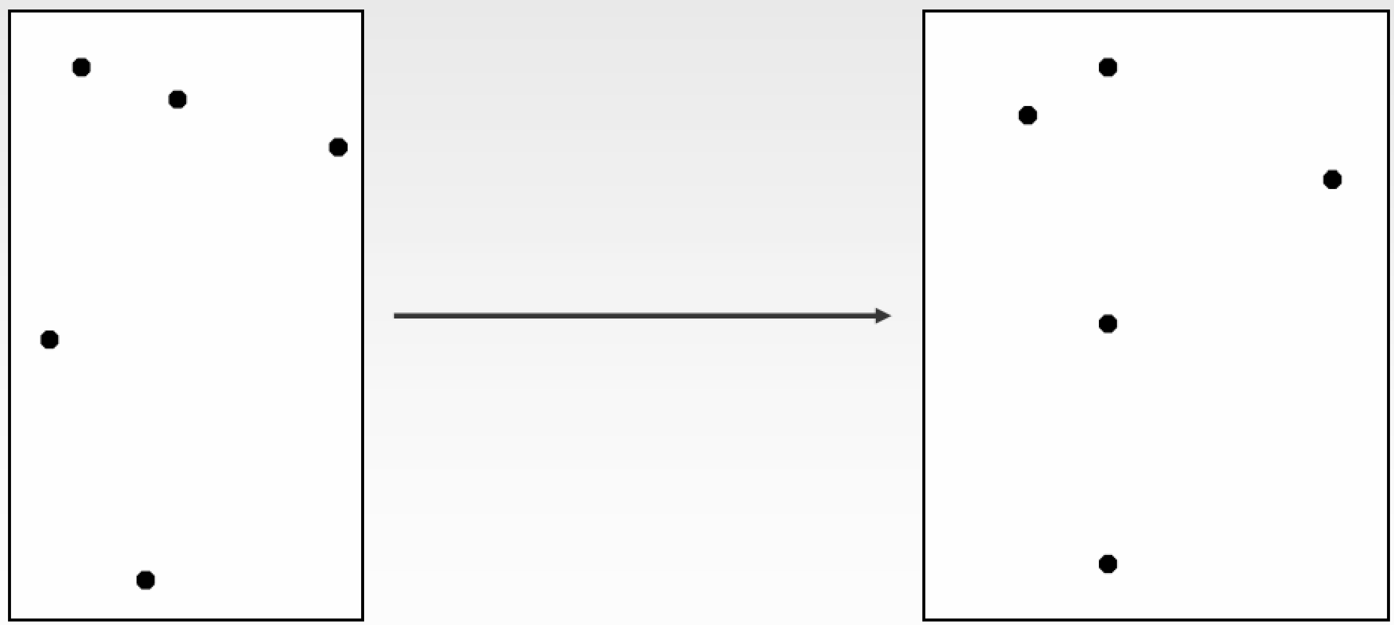
Matching with Shape Contexts

Modeling Transformations



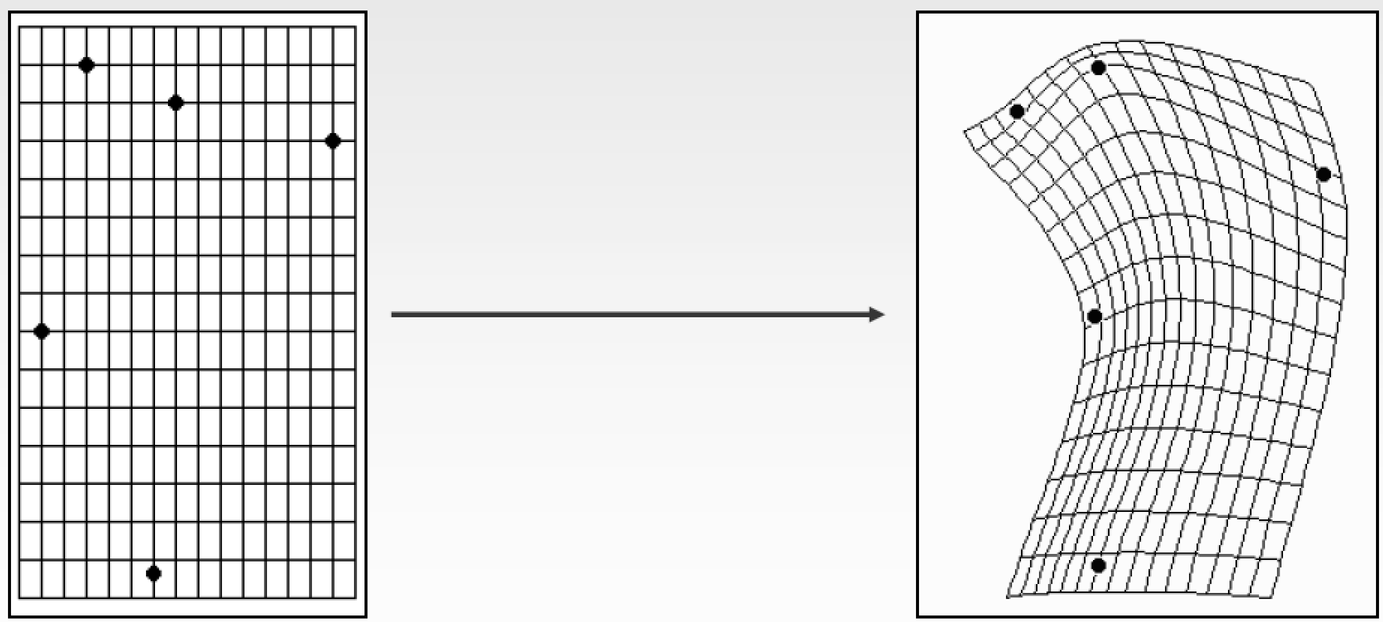
Matching with Shape Contexts

Thin Plane Spline (TPS) Model
(2D Generalization of Cubic Spline)

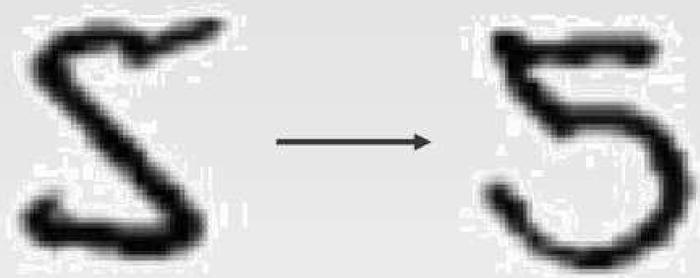


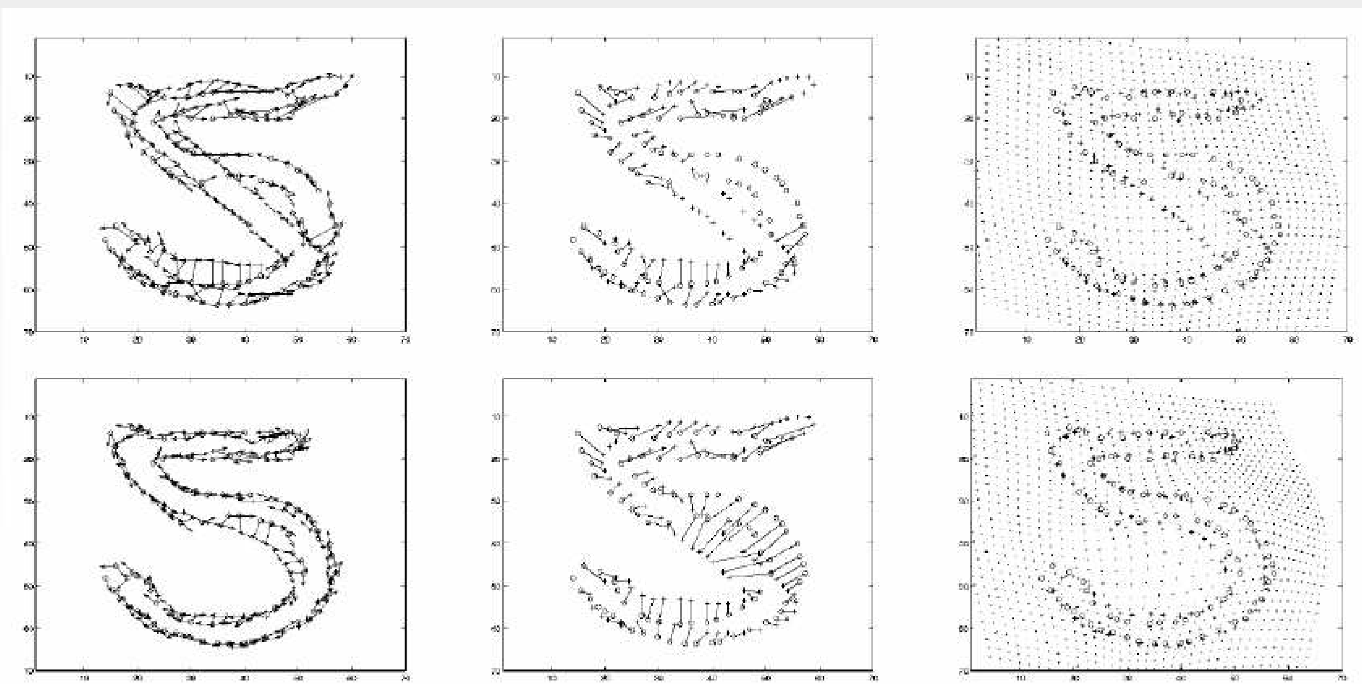
Matching with Shape Contexts

Thin Plane Spline (TPS) Model
(2D Generalization of Cubic Spline)



Matching with Shape Contexts





Matching with Shape Contexts

Invariance and Robustness

- Invariant under translation and scaling
- Insensitive to small affine distortion
- Can be made invariant to rotation

Shape Distance Measures

Shape Context Distance:

$$D_{sc}(P,Q) = \frac{1}{n}\sum_{p \in P} \arg\min_{q \in Q} C(p,T(q)) + \frac{1}{m}\sum_{q \in Q} \arg\min_{p \in P} C(p,T(q))$$

Appearance Distance

$$D_{ac}(P,Q) = \frac{1}{n}\sum_{i=1}^{n} \sum_{\Delta \in Z^2} G(\Delta)\left[I_P(p_i+\Delta) - I_Q(T(q_{\pi(i)})+\Delta)\right]^2$$

Thin-Plate Bending Energy

$$E = \iint \left[\left(\frac{\partial^2 f}{\partial x^2}\right)^2 + 2\left(\frac{\partial^2 f}{\partial xy}\right)^2 + \left(\frac{\partial^2 f}{\partial y^2}\right)^2\right] \mathrm{d}x\,\mathrm{d}y$$

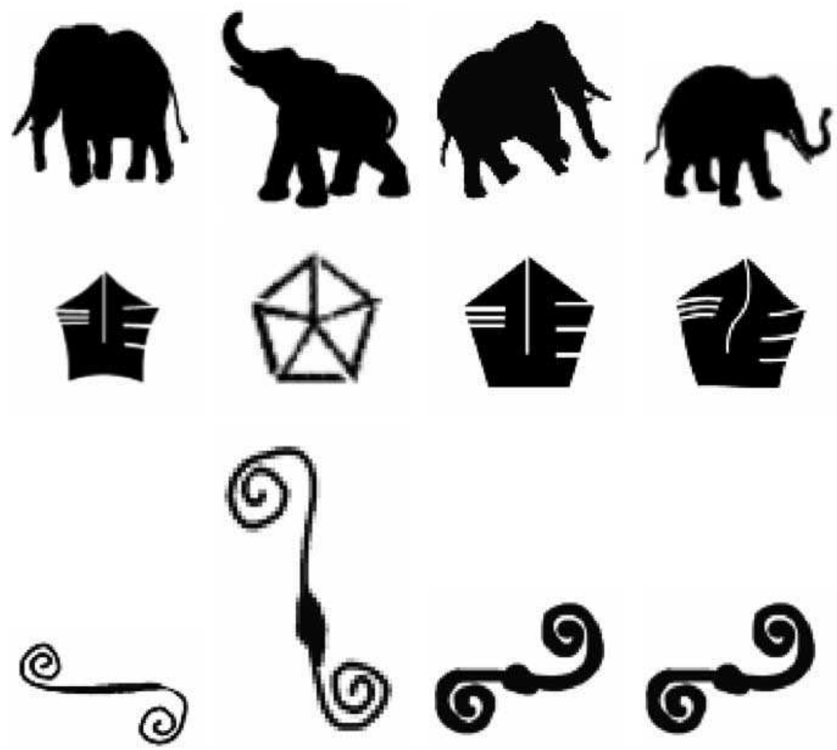
Shape context matching with handwritten digits

Only errors made out of 10,000 test examples

MPEG 7 Data Set

70 shape categories

20 images in each category



Bulleye test:

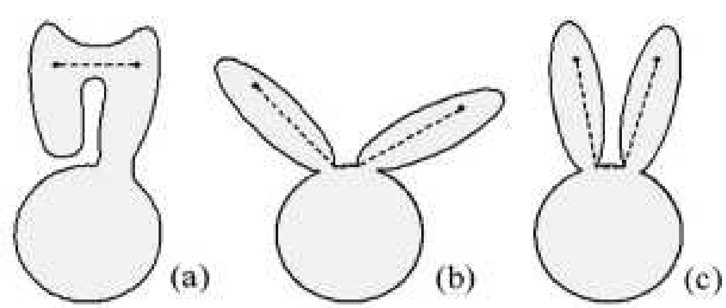
- Treat each image as a query
- Count the top 40 matches
- Recognition rate = the percentage of the 20 images in the same category is included in the top 40 matches

Rate = 76.51%

Inner Distance Shape Context

We use the inner-distance to build shape descriptors that are robust to articulation and capture part structure.

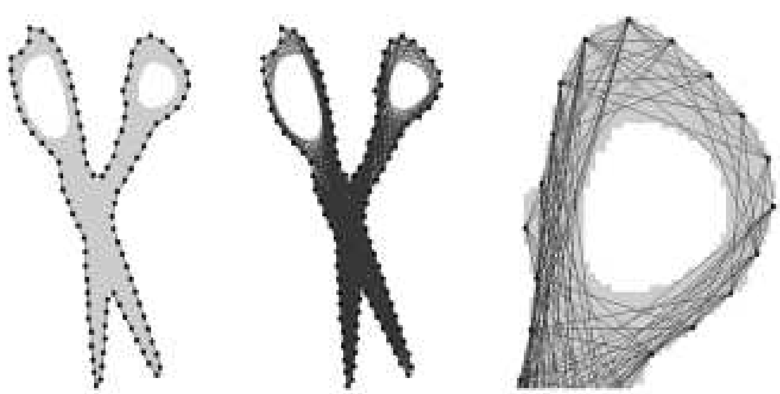
Inner-distance is defined as the length of the shortest path between landmark points within the shape boundary.



The inner-distance: Computation

Shortest path algorithms:

- (1) Build a graph with the sample points. For each pair of sample points p_1 and p_2, if the line segment connecting p_1 and p_2 fall entirely within the object, Let an edge between p_1 and p_2 is added to the graph with its weight equal to the Euclidean distance $|| p_1 - p_2 ||$.
- find the inner-distance between all pairs of points according to the graph.

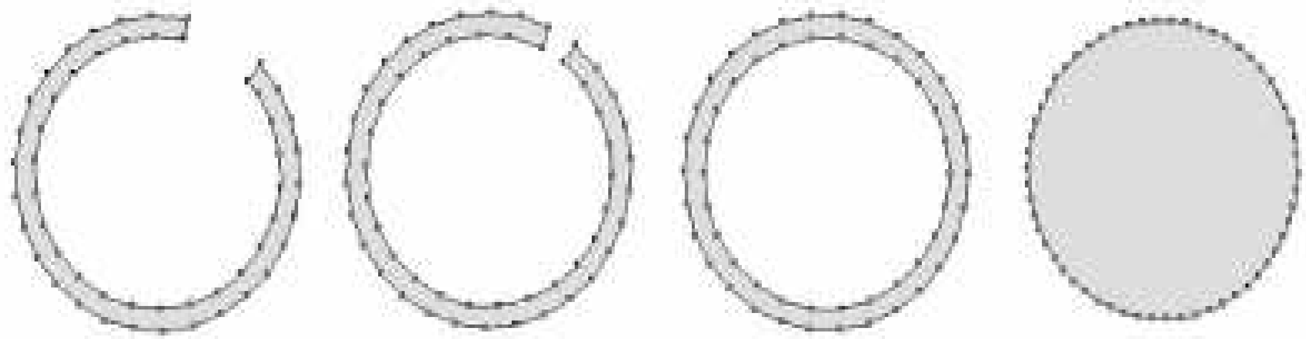


Note:

(1) Neighboring boundary points are always connected.

(2) (2) The inner-distance reflects the existence of holes without using samples points from hole boundary.

The inner-distance: ability to capture structures



With about the same number of sample points, the four shapes are virtually indistinguishable using distribution of Euclidean distance. However, their distributions of the inner-distance are quite different except for the first two shapes.

Note: more sample points will not affect the above statement.

Inner-Distance Shape Context (IDSC)

To extend the shape context, Euclidean distance is directly replaced by the inner-distance.

Inner-Distance Shape Context (IDSC)

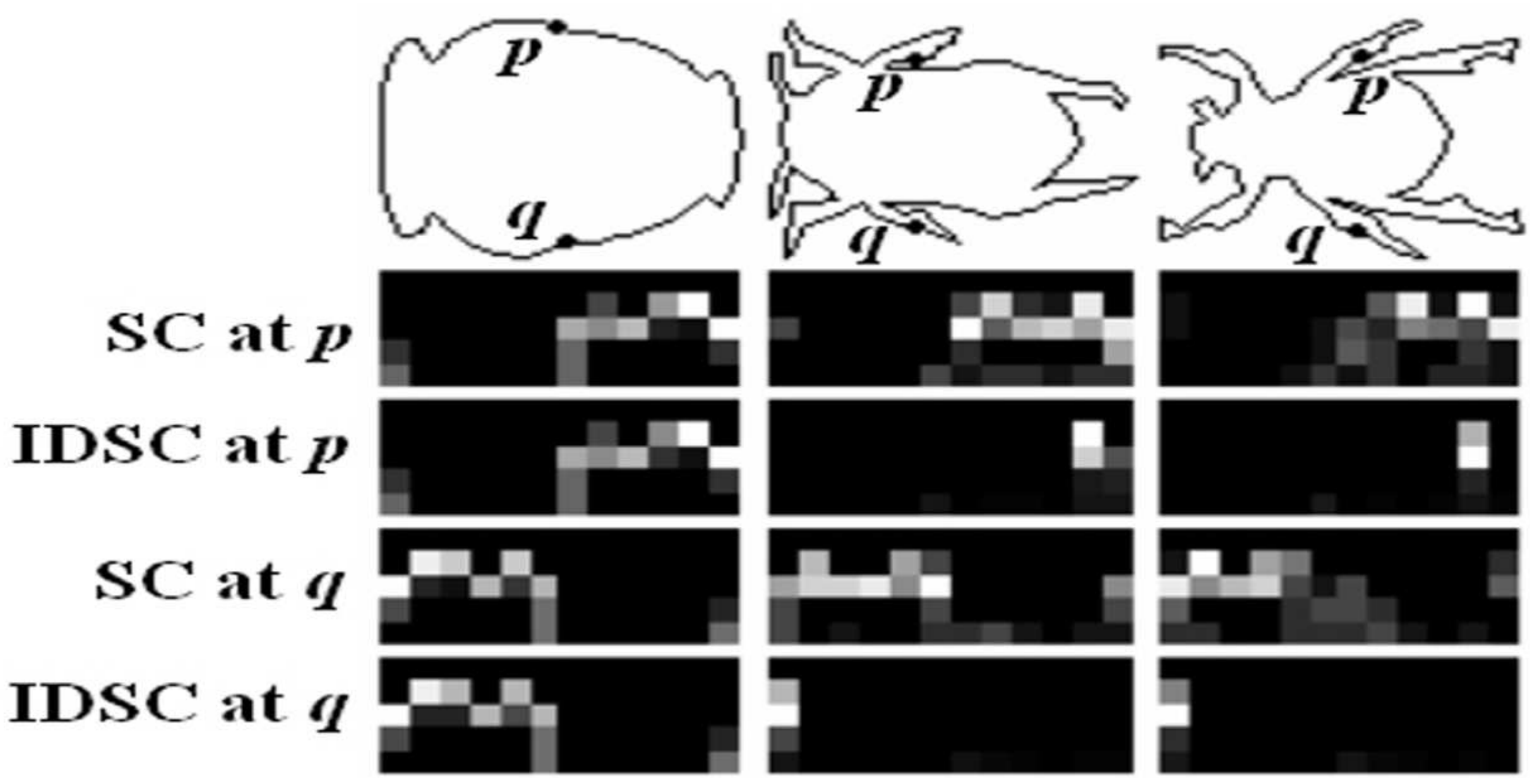
Euclidean Distance is replaced by inner distance

Angle is replaced by *inner-angle*

Inner-angle is used for the orientation bins.

Noise may reduces the stability of the inner-angle, smoothing contour before computing it.

Inner-Distance Shape Context (IDSC)



In the histogram, the x axis denotes the orientation bins and the y axis denotes log distance bins.

Shape matching through Dynamic programming

DP is used to solve the matching problem since it uses the ordering information provided by shape contours.

By default, assumes the two contours are already aligned at their start and end points.

Without this assumption, one simple solution is to try different alignments at all points on the first contour and choose the best one.

$$\text{the complexity}: O(n^2) \Rightarrow O(n^3)$$

Shape distance

The matching cost $C(\pi)$ is used to measured the similarity between shapes.

IDSC+DP is better than SC+DP

- Better performance
- Only two parameters to tune
  - The penalty τ for a point with no matching, usually set 0.3.
  - The number of start points *k* for different alignments, usually set 4-8.
- Easy to implement since it does not require the appearance and transformation model.

Experiment

MPEG7 CE-Shape-1 shape database is widely tested, which consists of 1400 silhouette images from 70 classes. Each class has 20 different shapes.

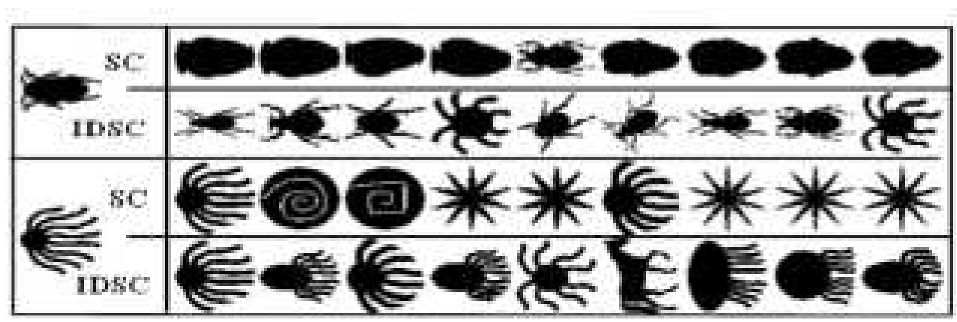
- Bullseye test: for every image in the database, it is matched with all other images and the top 40 most similar candidates are counted.

Experiment

- The score of the test is the ratio of the number of correct hits of all images to the highest possible number of hits (which is 20x1400).

Alg.	CSS [32]	Vis. Parts[24]	SC+TPS[5]	Curve Edit[38]	Dis. Set[18]
Score	75.44%	76.45%	76.51%	78.17%	78.38%

Alg.	MCSS[22]	Gen. Mod.[44]	**MDS+SC+DP**	**IDSC+DP**	
Score	78.8%	80.03%	**84.35%**	**85.40%**	



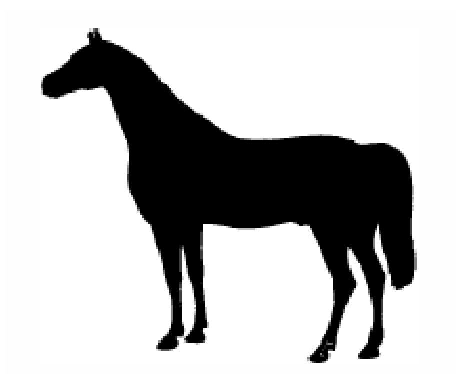
Locally Constrained Diffusion Process (LCDP) Method for Shape Classification (CVPR'09)

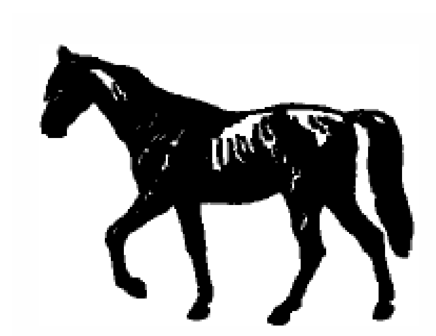
Basic Idea:

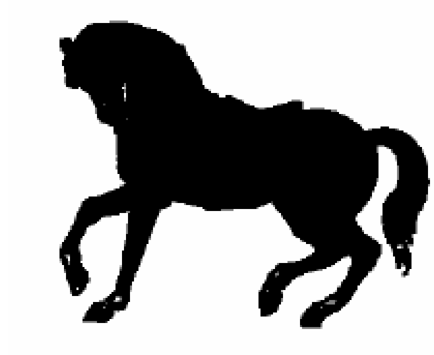
Assume we have a pair-wise shape distance measure (IDSC or SC). If *a* is similar to *b*, and *b* is similar to *c* → This provide the hint that *a* might be similar to *c*.

Shape distance *d(a,c)* may be better refined by considering both *d(a,b)* and *d(b,c).*

This need a set of shape instances and is completely unsupervised







Markov Chain

K is a Gaussian to shape distance

$$D(x_i) = \sum_{j=1}^{n} k(x_i, x_j)$$

Probability

$$P(x_i, x_j) = \frac{k(x_i, x_j)}{D(x_i)}.$$

Stable Probability

$$P* = P^t, \quad t \rightarrow \infty$$

Problem when Using to Shape Instances

Sensitive to noise / outliers

- Solution: Build a K nearest neighboring graph instead of the fully connected graph
- For each shape instance, we only consider K nearest neighboring shape instances for calculating probability, which is used for Markov chain iteration

The shape space might be sparse

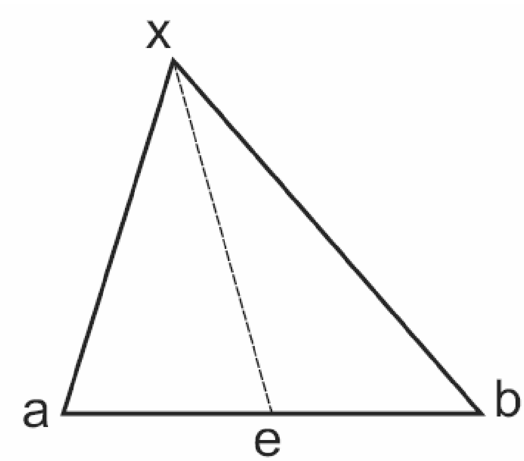
- Shape instances in the same category might still have large deformation
- We want to interpolate them by constructing more shape contours to make the shape space denser
- But construct new shape instances explicitly is complex
- In fact, we only need the shape distance to the new shape instances and do not need them explicitly

Ghost Point

Construct a new shape instance by averaging *a* and *b*.

$$||h(x) - \frac{h(a)+h(b)}{2}||^2 = \frac{||h(x)-h(a)||^2}{2} + \frac{||h(x)-h(b)||^2}{2} - \frac{||h(a)-h(b)||^2}{4}$$

$$\rho(x,\mu(a,b))^2 = \frac{1}{2}\rho(x,a)^2 + \frac{1}{2}\rho(x,b)^2 - \frac{1}{4}\rho(a,b)^2$$



In this paper, for each shape instance *a*, find its nearest neighbor *b* and construct a new shape instance that averaging *a* and *b*.

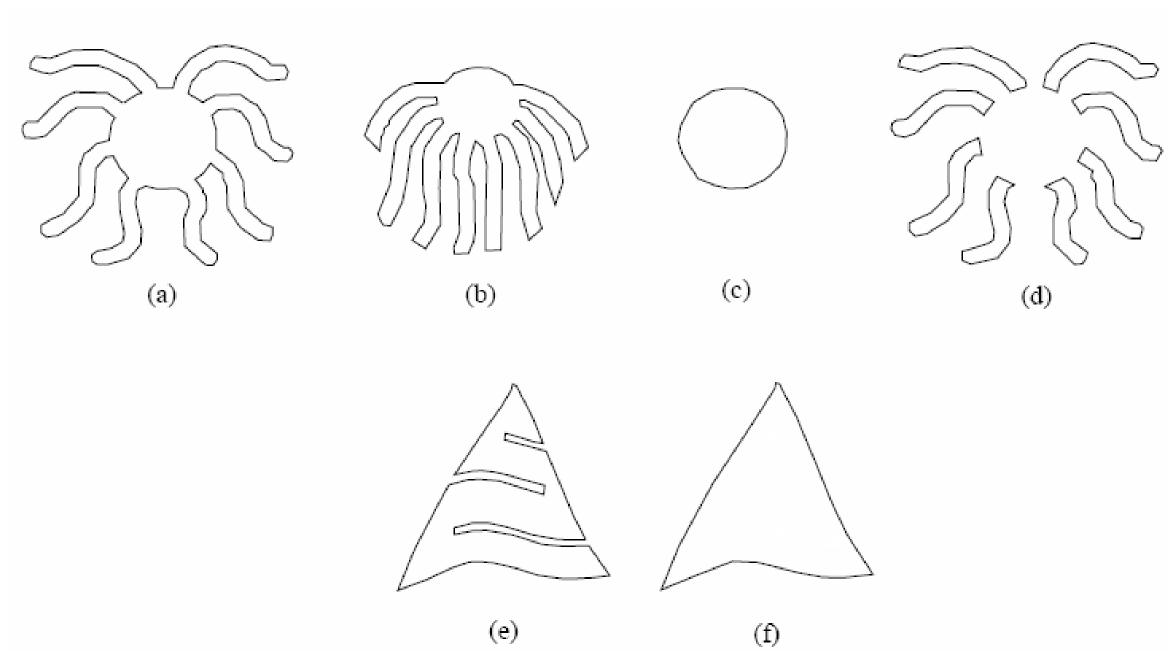
Results on MPEG7 Dataset

Alg.	IDSC [12]	IDSC + LAM	IDSC +DM	IDSC +LP[25]	IDSC +LCDP	IDSC +LCDP +unsupervised GP
Score	85.40%	89.00%	78.56%	91.00%	**92.36%**	**93.32%**

Our Recent Work: Under Review

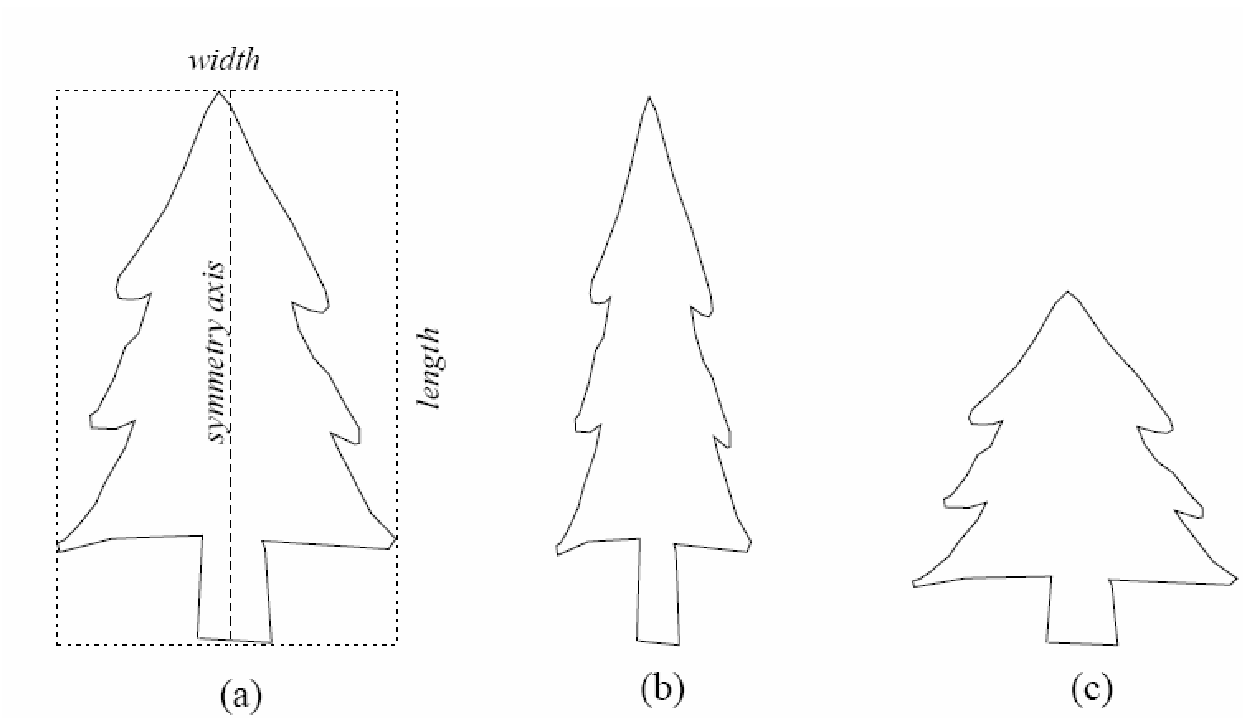
Two Perceptually Motivated Strategies for Shape Classification

First, shape decomposition → base structure and strand structures

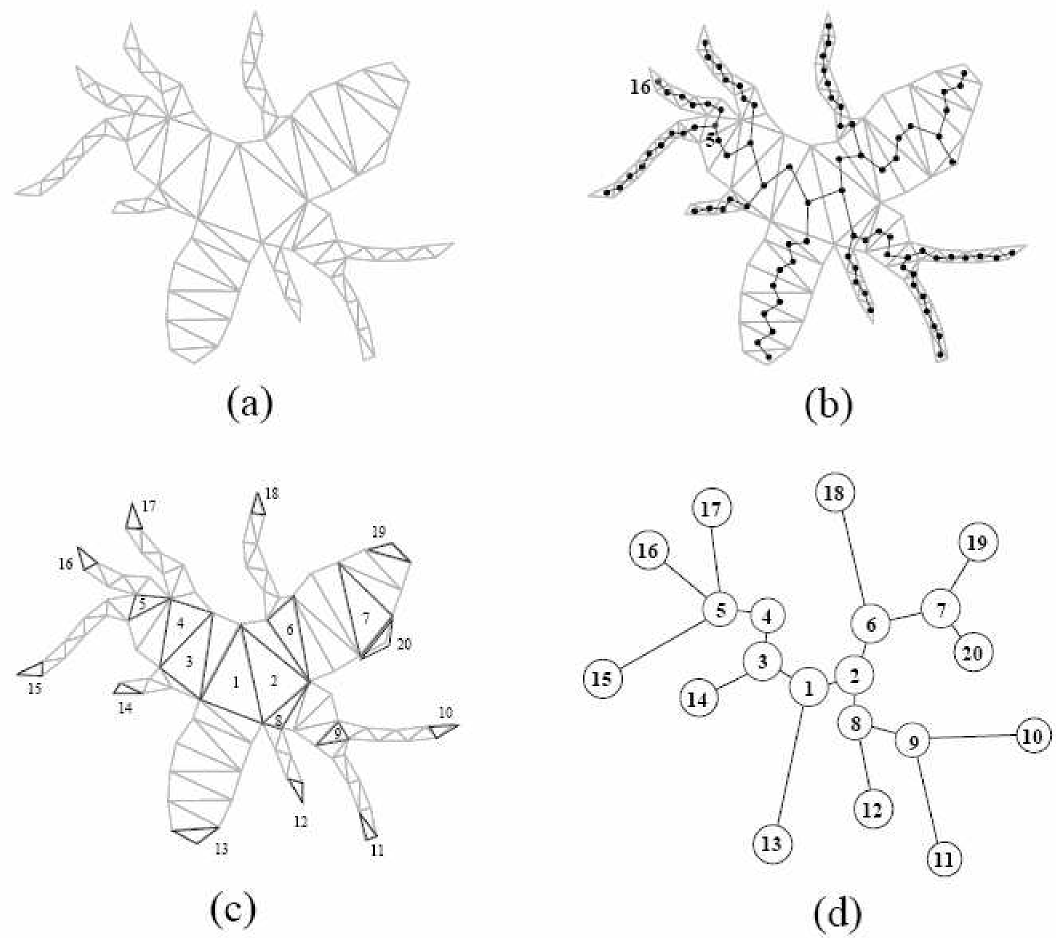


Cont.

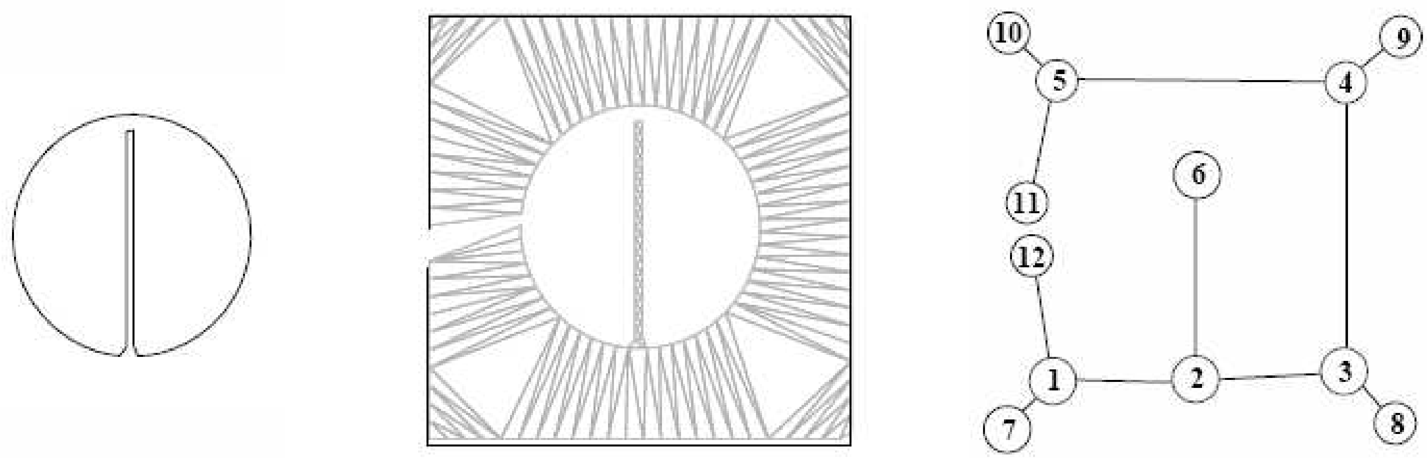
Second, invariance to the aspect ratio if the shape is symmetric



Shape Decomposition for Outward Strands

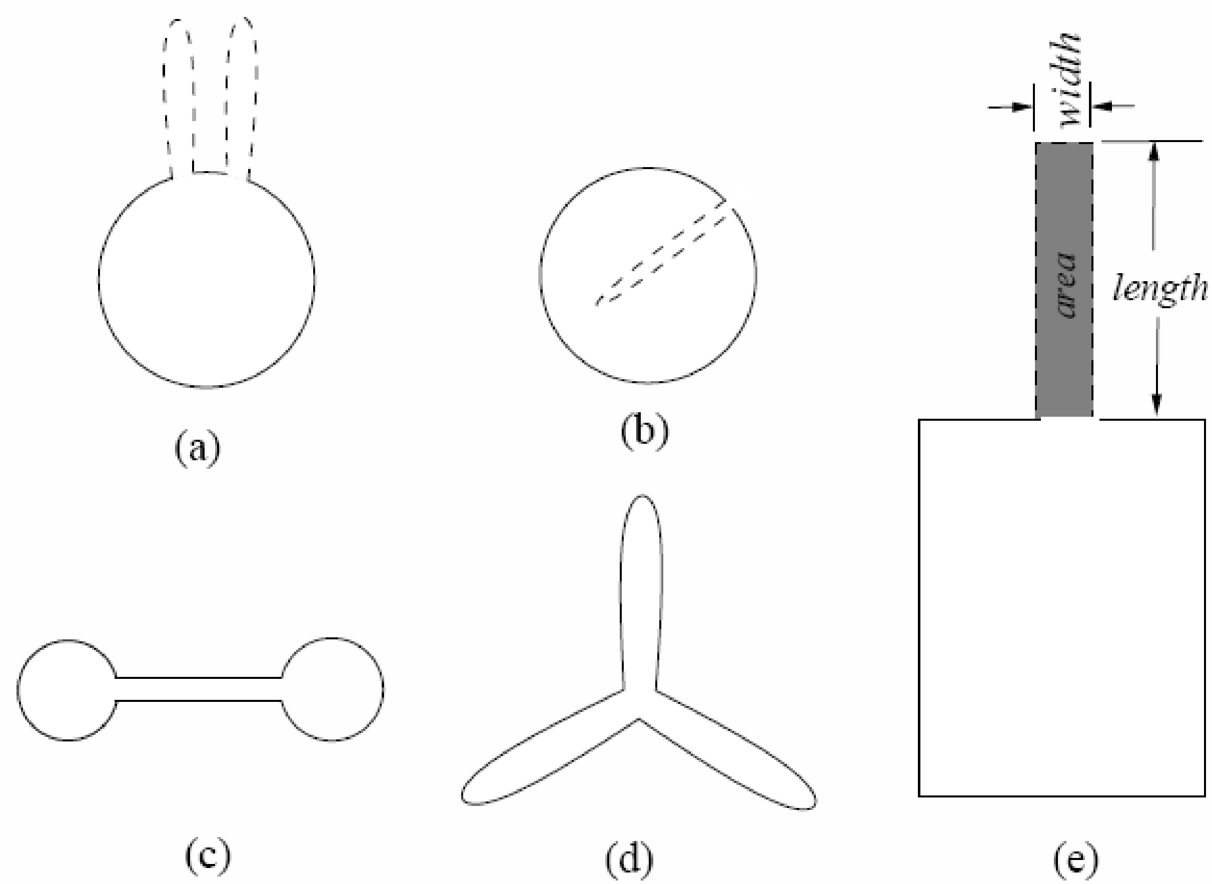


Shape Decomposition for Inward Strands



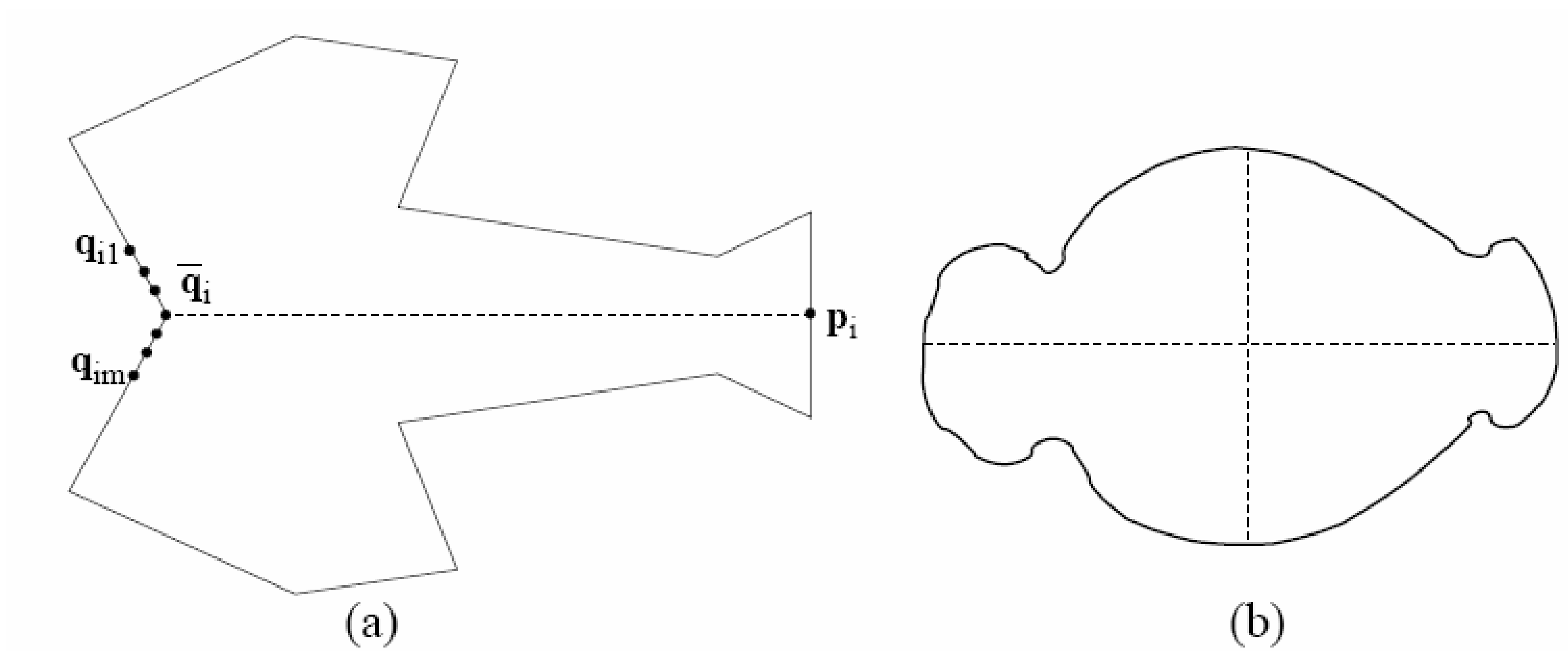
(a) (b) (c)

Strands Determination



(a) (b) (c) (d) (e)

Symmetry Axis Detection



Algorithm

E **and** ***F*** **are the base structures of S after outward and inward strand removal,** ***n*** **is the number of outward strands**

$$\phi_1(S_1, S_2) = \min\{C(S_1, S_2), C(E_1, E_2) \cdot c(n_1, n_2), C(F_1, F_2)\},$$

We then scale the target shape contour S_2 along its symmetry axis to a shape contour S_2' so that it has the same aspect ratio as the template,

$$\phi_2(S_1, S_2) = \min\{C(S_1, S_2), C(S_1, S_2')\}.$$

Combining two Strategies

$$\phi(S_1, S_2) = \min\{\phi_1(S_1, S_2), \phi_2(S_1, S_2)\}.$$

Results on MPEG7

Method	Rate
Proposed method + IDSC(χ^2) + LCDP	**95.60** %
IDSC(χ^2) + LCDP + unsupervised GP [32]	93.32 %
IDSC(χ^2) + LCDP [32]	92.36 %
IDSC(χ^2) + LP [5]	91.61 %
Contour Flexibility [29]	89.31 %
Proposed method + IDSC(χ^2)	**88.39** %
Shape-tree [13]	87.70 %
Triangle Area [2]	87.23 %
IDSC(EMD) [17]	86.56 %
Hierarchical Procrustes [18]	86.35 %
Symbolic Representation [11]	85.92 %
IDSC(χ^2) [16]	85.40 %
Shape L'Âne Rouge [20]	85.25 %
Multiscale Representation [1]	84.93 %
Polygonal Multiresolution [4]	84.33 %
Fixed Correspondence [27]	84.05 %
Chance Probability Function [26]	82.69 %
Curvature Scale Space [19]	81.12 %
Generative Model [28]	80.03 %

What is the next?