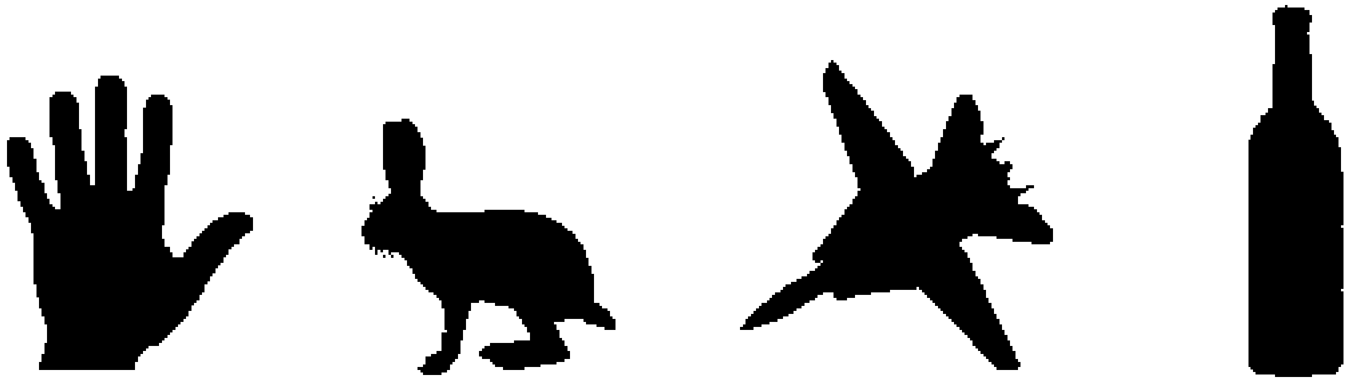
A Hierarchical Representation for Matching Deformable Shapes

Pedro Felzenszwalb
Joshua Schwartz

University of Chicago

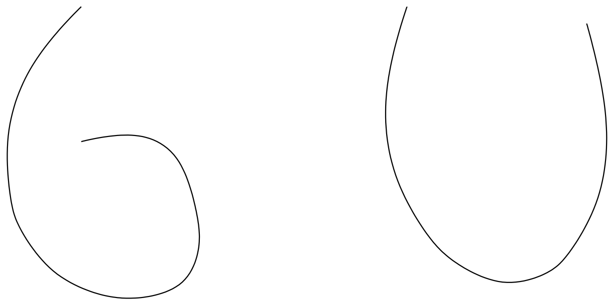
Shape-based recognition



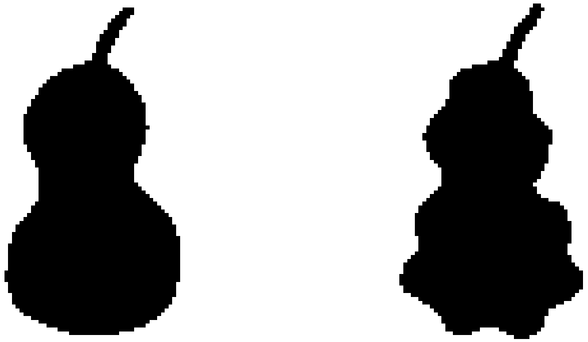
- Humans can easily recognize objects using shape information
- Classical approach for recognizing rigid object
- Important for many object categories
  - Fairly abstract representation

Local deformation models

- Measure amount of bending and stretching necessary to turn one curve into another --- only captures local information [Basri, et al 95], [Younes, 98]



can turn these into each other without much bending at any point



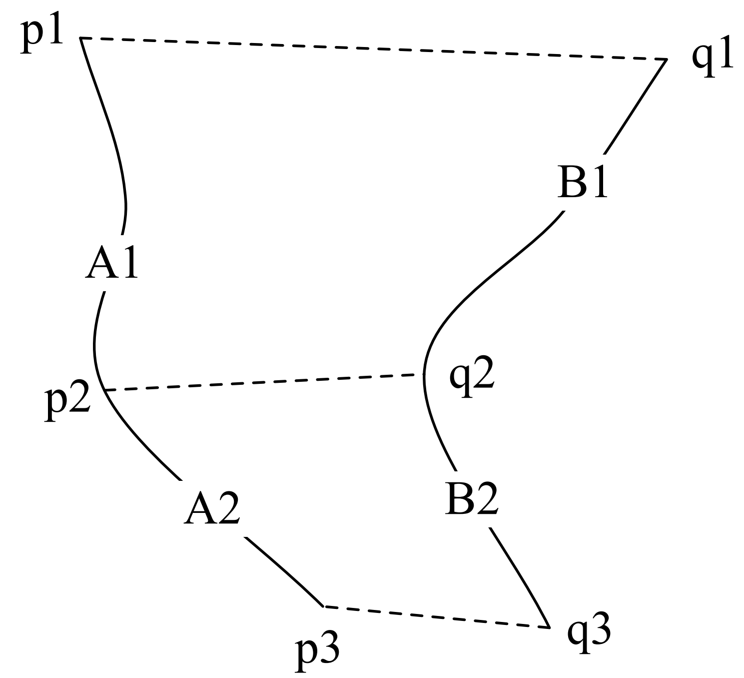
Similar objects with completely different local boundary properties

Compositional model

- Consider arrangement of points far from each other
- Combine matchings between subcurves to form longer matchings

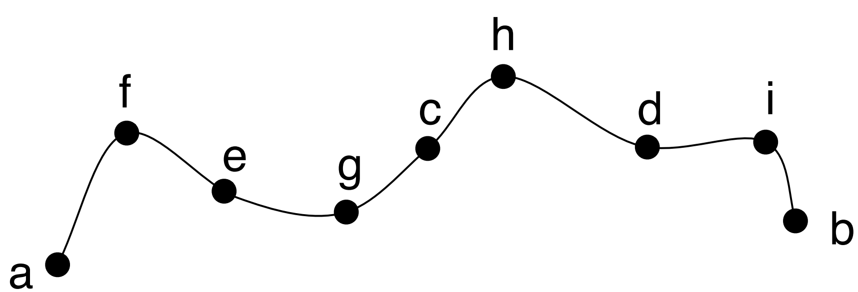
Combine matching from A_1 to B_1 with matching from A_2 to B_2

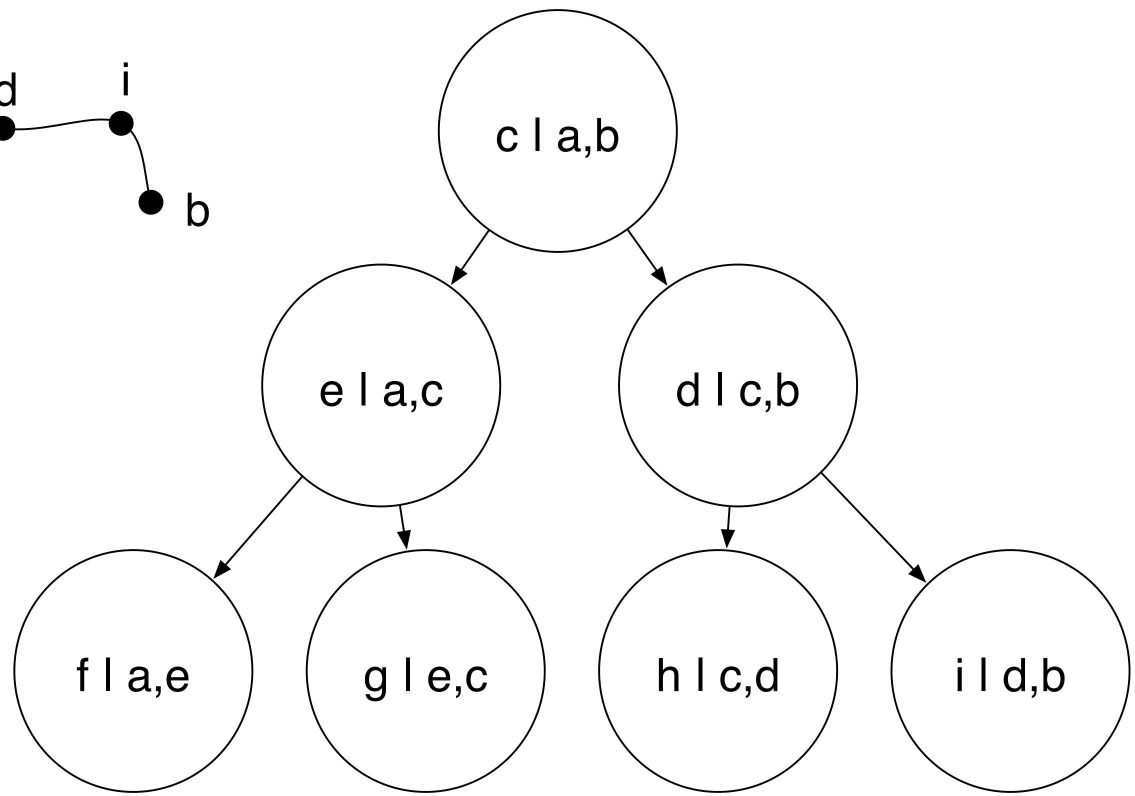
quality depends on quality of matchings being combined and arrangement of (p_1,p_2,p_3), (q_1,q_2,q_3)



Shape-tree

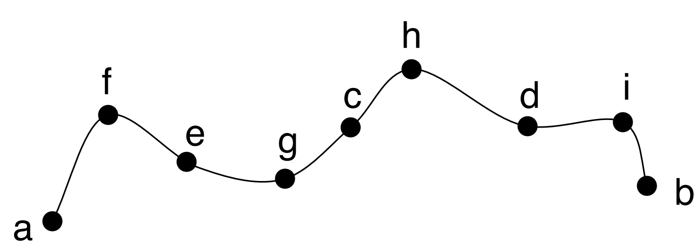
- Shape-tree of curve from a to b:
  - Select midpoint c, store location w.r.t. a,b frame
  - Left child is a shape-tree of sub-curve from a to c
  - Right child is a shape-tree of sub-curve from c to b

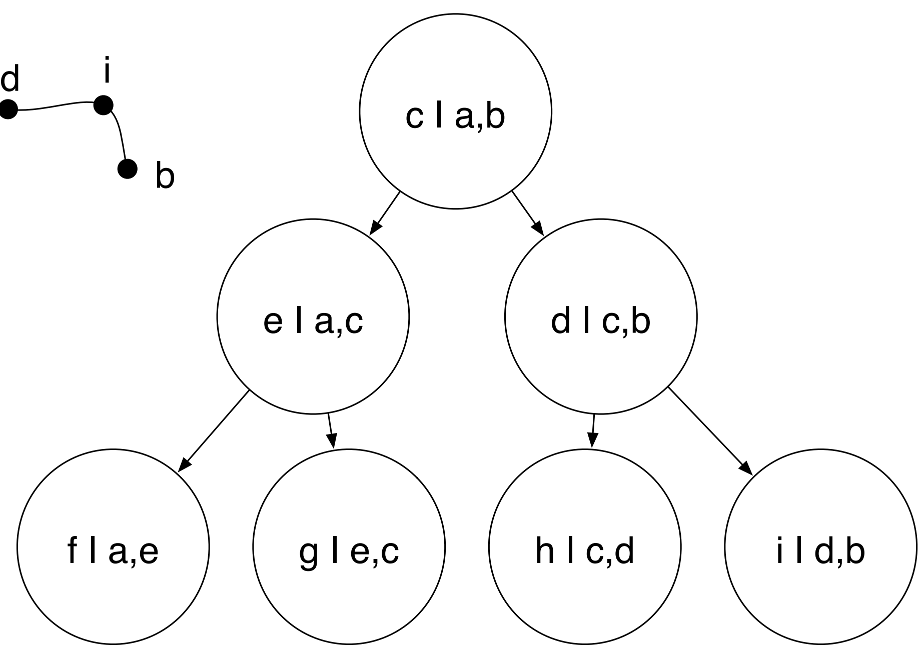




Shape-tree

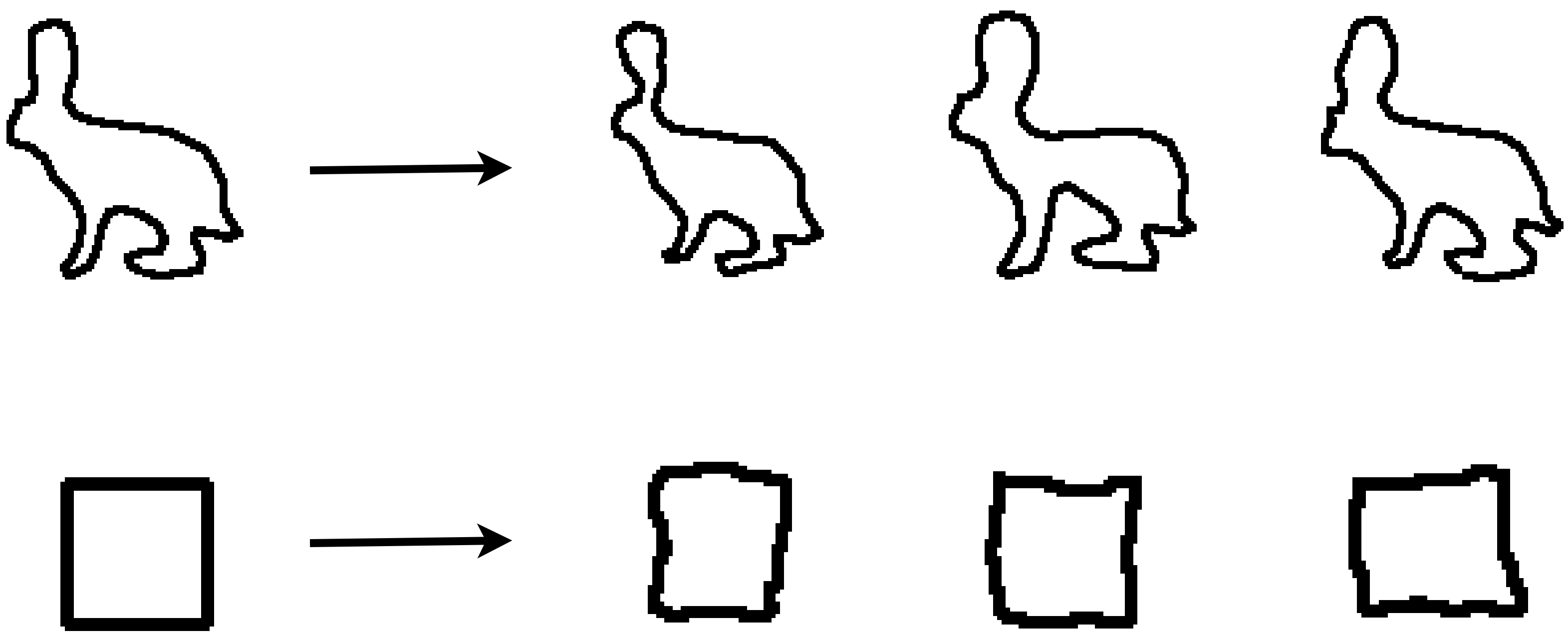
- Invariant to similarity transformation
- Subtree is shape-tree of sub-curve
- Given placement for a,b we can reconstruct the curve
- Bottom nodes captures local curvature
- Top nodes capture curvature of sub-sampled curve





Deformations

- Perturb relative locations stored in a shape-tree
  - Reconstructed curve is perceptually similar to original
  - Global properties are preserved



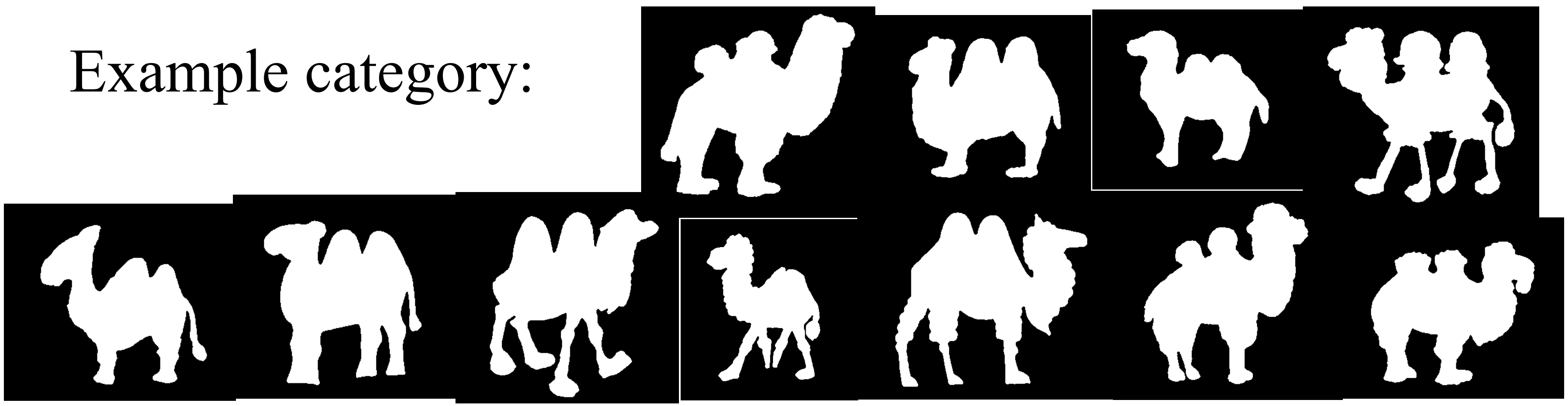
Distance between curves

- Define distance between shape-trees in terms of deformations applied to each node
- But a curve can be represented with multiple shape-trees!
  - We need to search over possibilities
- Given curves A,B
  - Fix shape-tree for A, search over shape-trees for B: $O(n^4)$
  - Jointly find correspondences and common tree: $O(n^4 \log n)$
- Reason about missing parts by cutting off sub-trees

Recognition results

MPEG7		
Shape-tree	Inner distance	G. M.
85.30	85.40	80.03

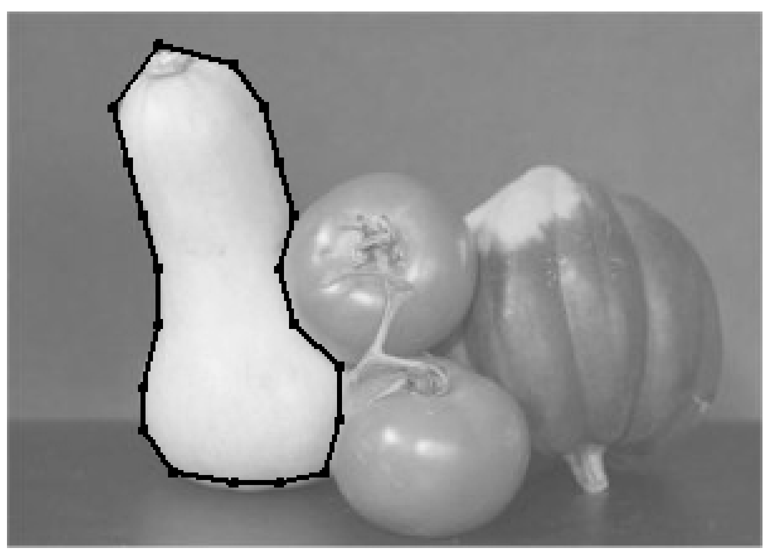
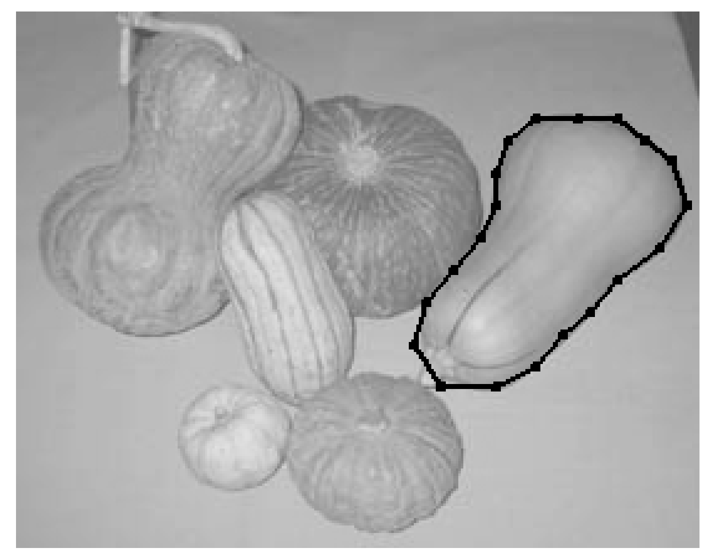
Example category:



Swedish leaves		
Shape-tree	Inner distance	Shape context
94.31 (mean)	94.13	88.12

Cluttered images

- Consider embedding deformed curve in images
  - Cost depends on deformation + image evidence
  - No edge detection
- Combine partial embeddings with bottom-up algorithm
  - Generalization of Dijkstra's shortest path (with D. Mcallester)
  - Find best match without considering bad ones

Problems

- Current local evidence measurement too weak
- Often place object at strange location
  - Close inspection shows that gradient is high along boundary
- What is going on?
  - We may need NMS
  - We may need to capture internal coherence
  - Could try finding multiple good solutions

Parts

- Sub-trees represent generic curves
- We can share sub-trees among different models
  - Useful for bottom-up matching
- Look for a context-free grammar for compactly representing all shape-trees of a big data set
  - Terminals l(a,b) are line segments from a to b
  - Sentences are curves
  - Non-terminals N(a,b) represent curve fragments

Examples

- L(a,b) generates an "almost straight curve" from a to b
- Productions
  - $L(a,b) \rightarrow L(a,c)\ L(c,b)$ where $c \sim (a+b)/2$
  - $L(a,b) \rightarrow l(a,b)$ if a near b

- Can also define B(a,b) to generate an elongated branch anchored at a and b
- etc.
- Learning is a challenge