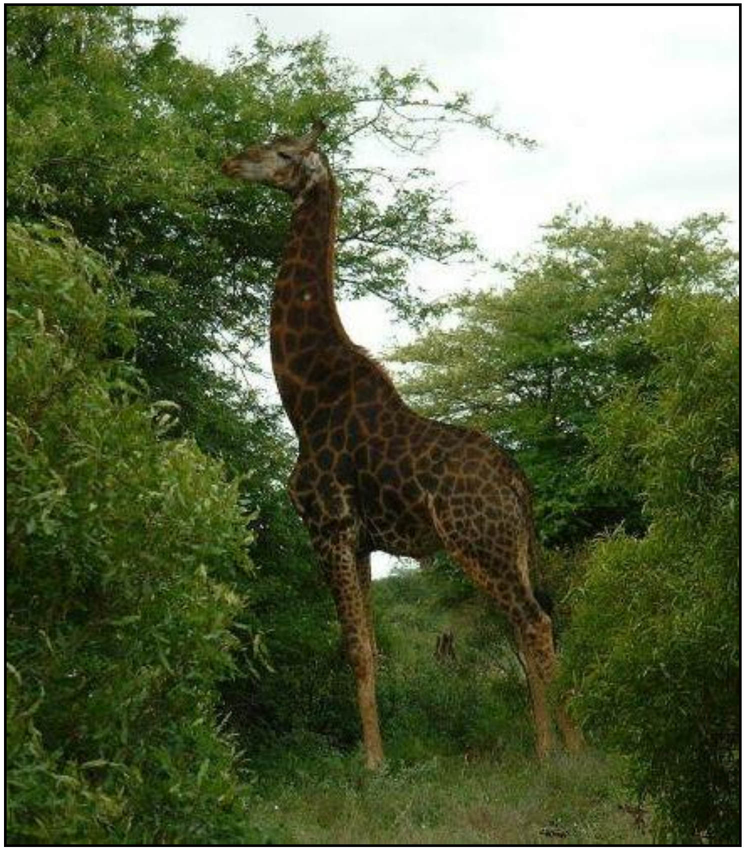
Skeleton-search: Category-specific object segmentation/recognition using a skeletal shape model

Nhon Trinh & Benjamin Kimia
Brown University

British Machine Vision Conference
Sep 08, 2009

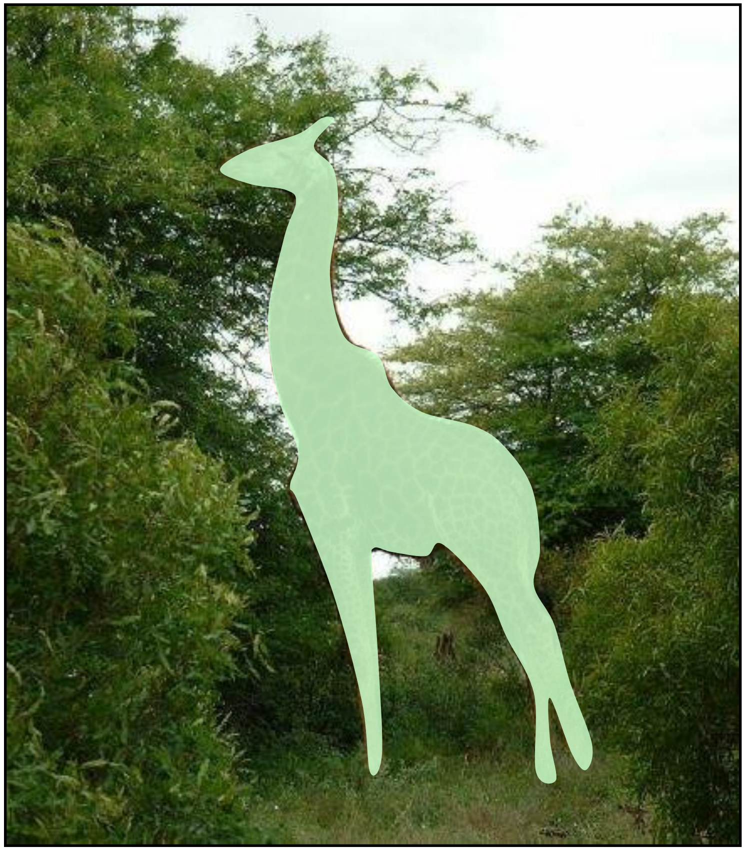
Category-specific object recognition



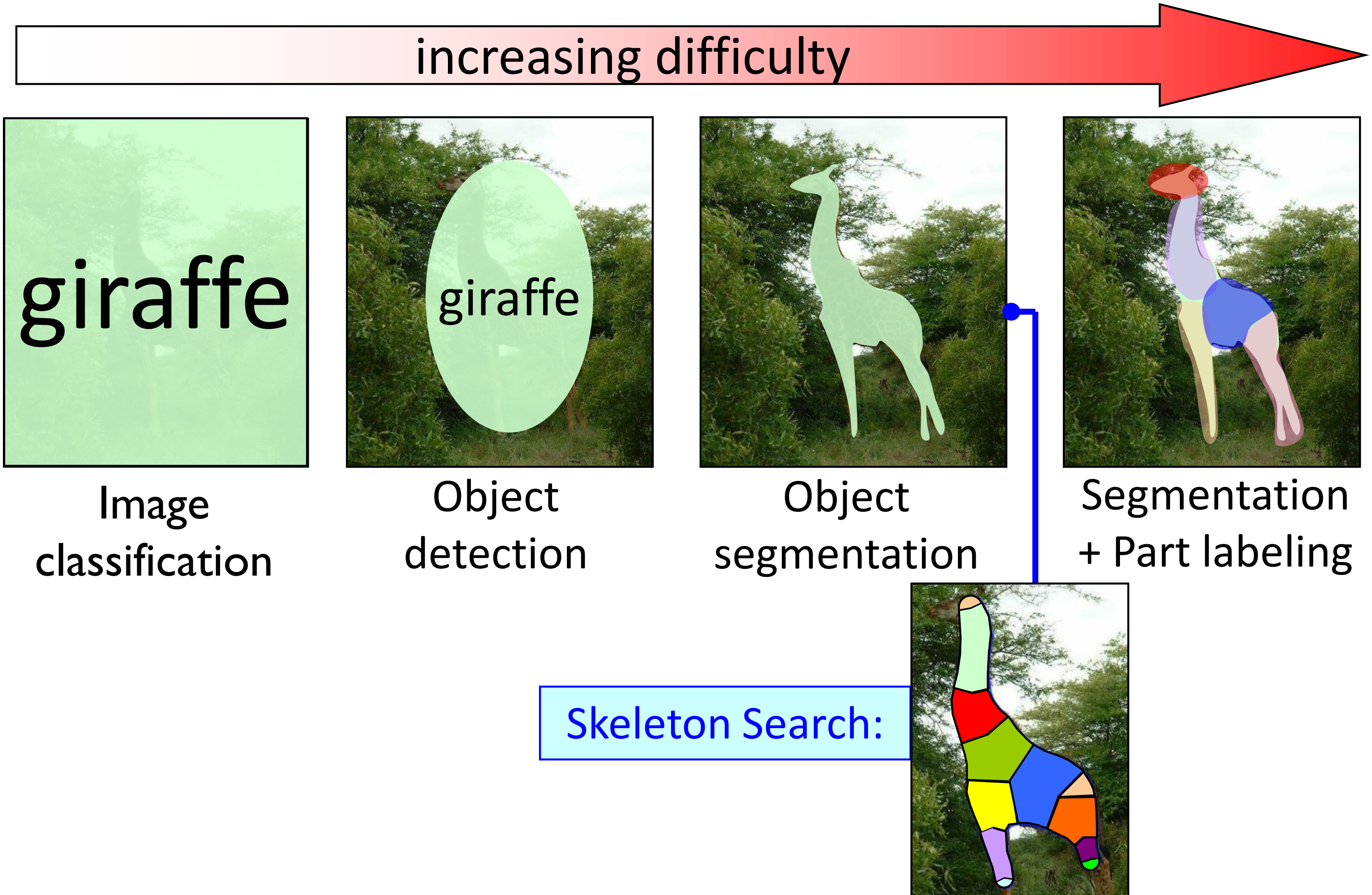
- Q: Is there a giraffe in this image?
- A: Yes.
- Q: Really? Can you delineate it?


giraffe

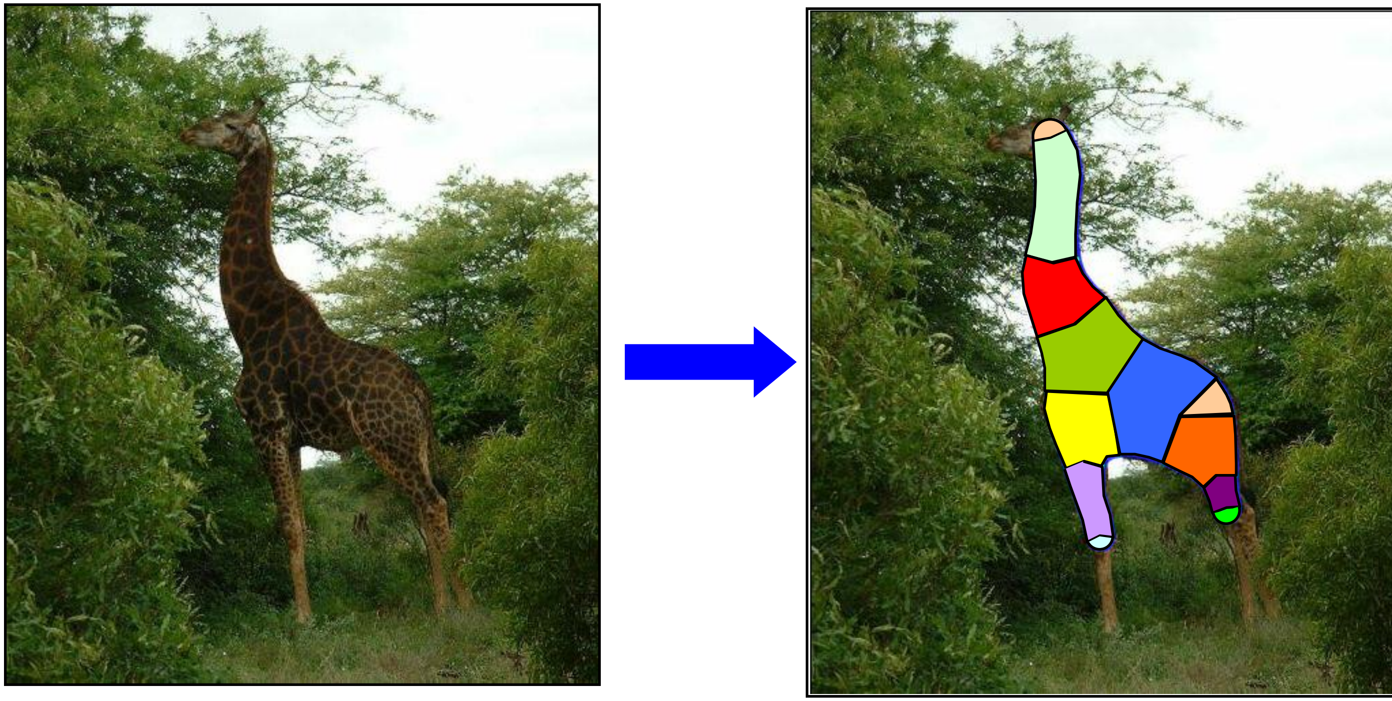

giraffe



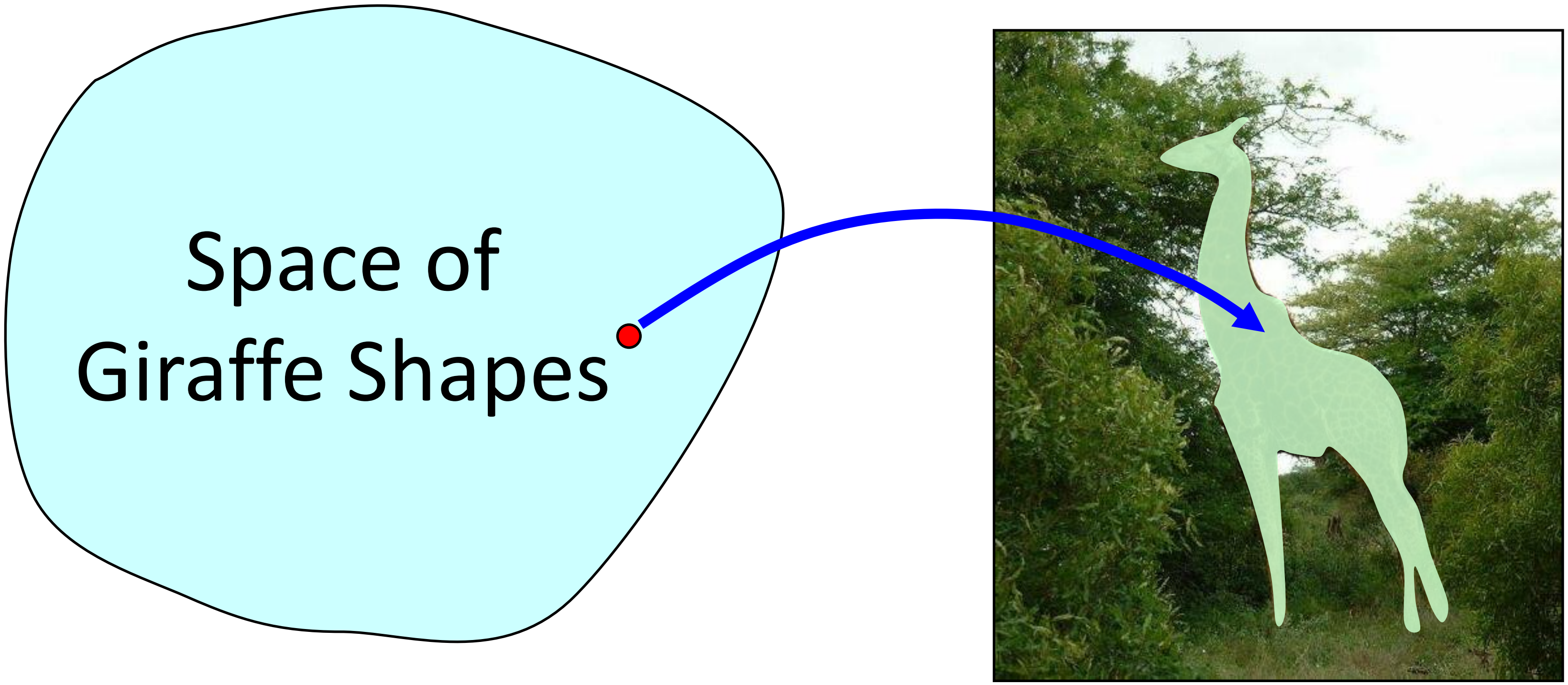
Faces of Object Recognition



Our goal



Top-down approach



1. How to represent shapes of an object category?
2. How to measure support for a shape in an image?
3. How to search for the best supported shapes?

Contributions

1. How to represent shapes of an object category?

   → Fragment-Based Generative Model for Shape

2. How to measure support for a shape in an image?

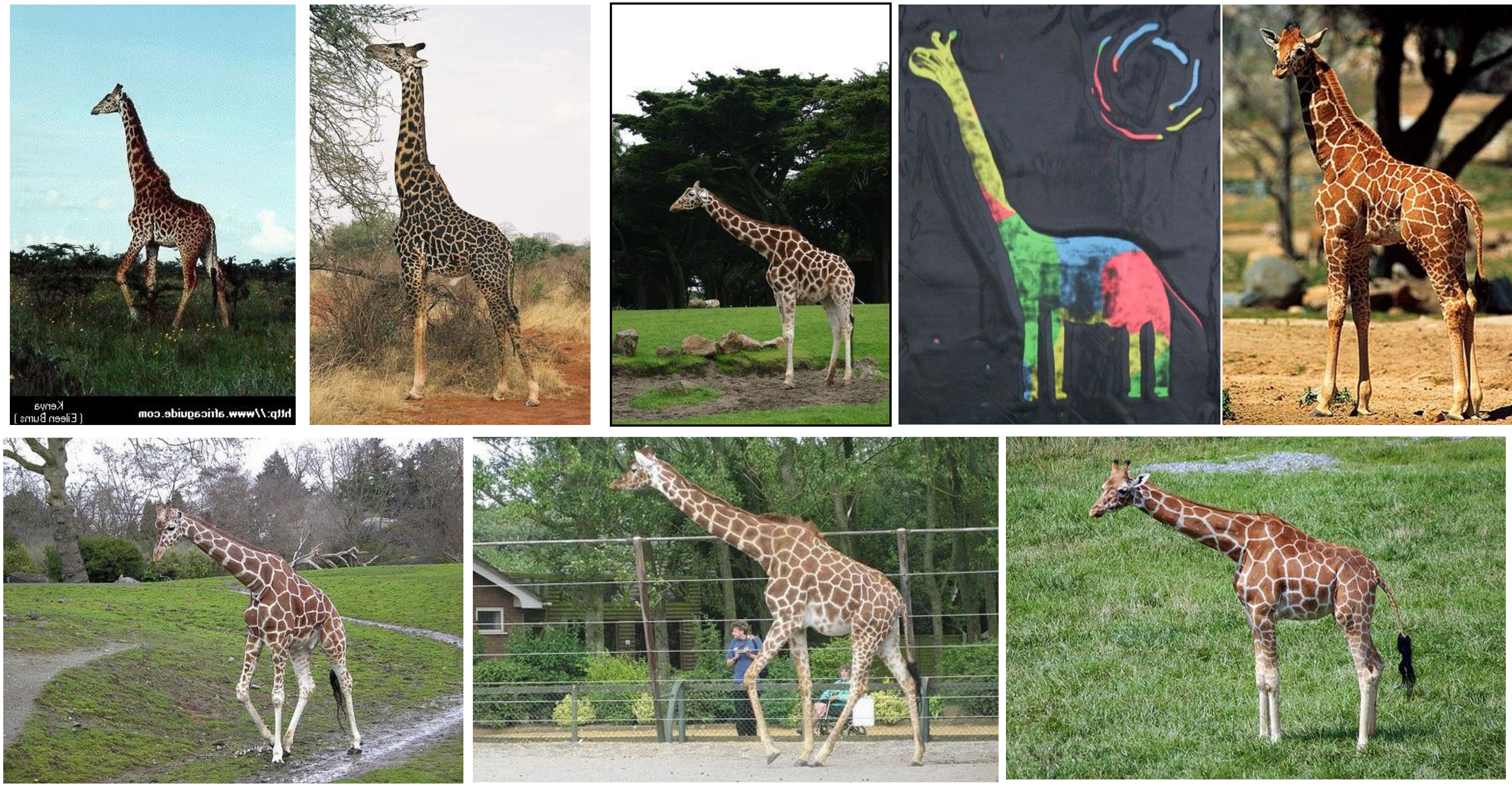
   → Improvement to Oriented Chamfer Matching

3. How to search for the best supported shapes?

   → Extension to the Viterbi algorithm to compute multiple solutions

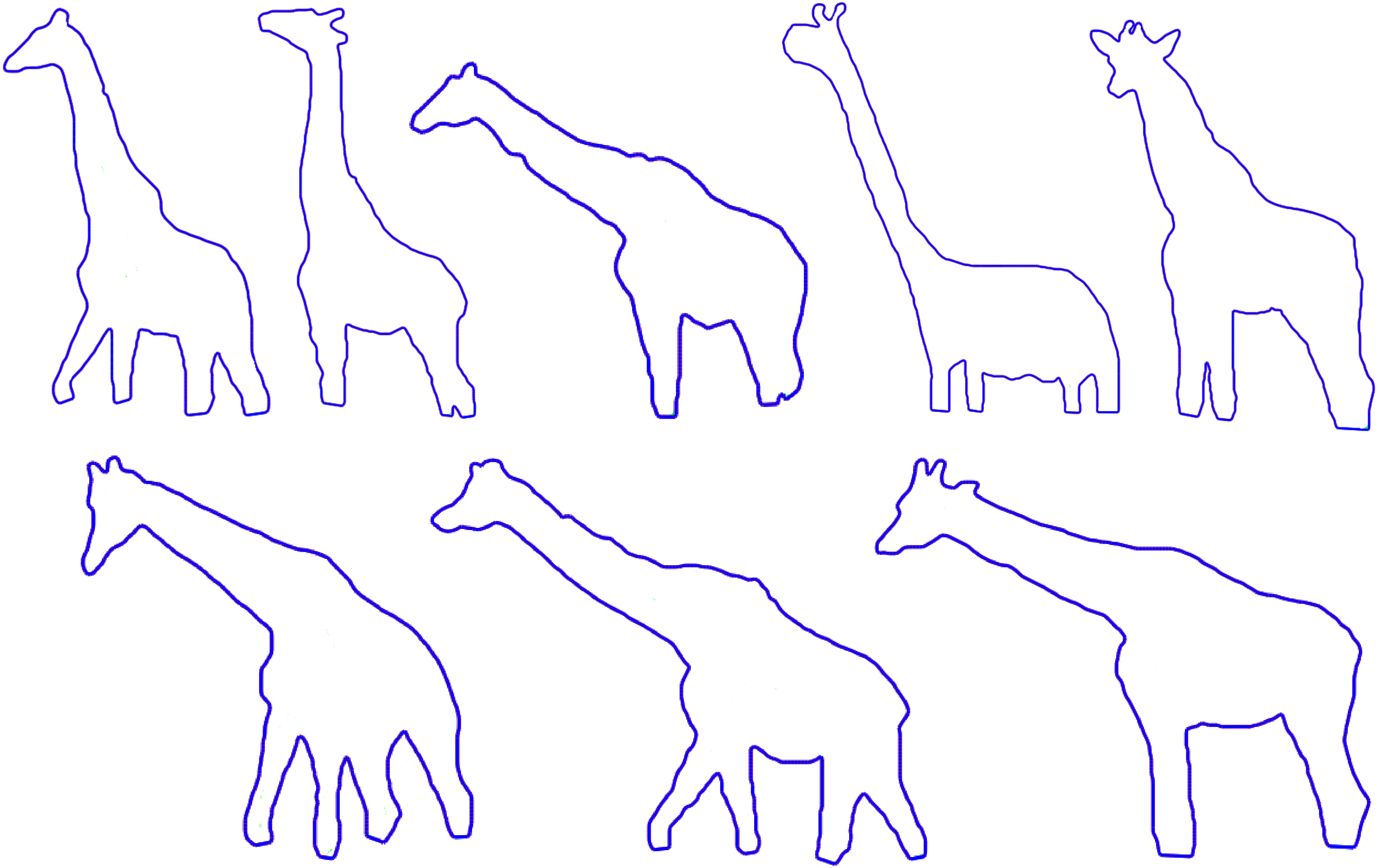
1. How to represent shapes of an object category?

Giraffes in Images

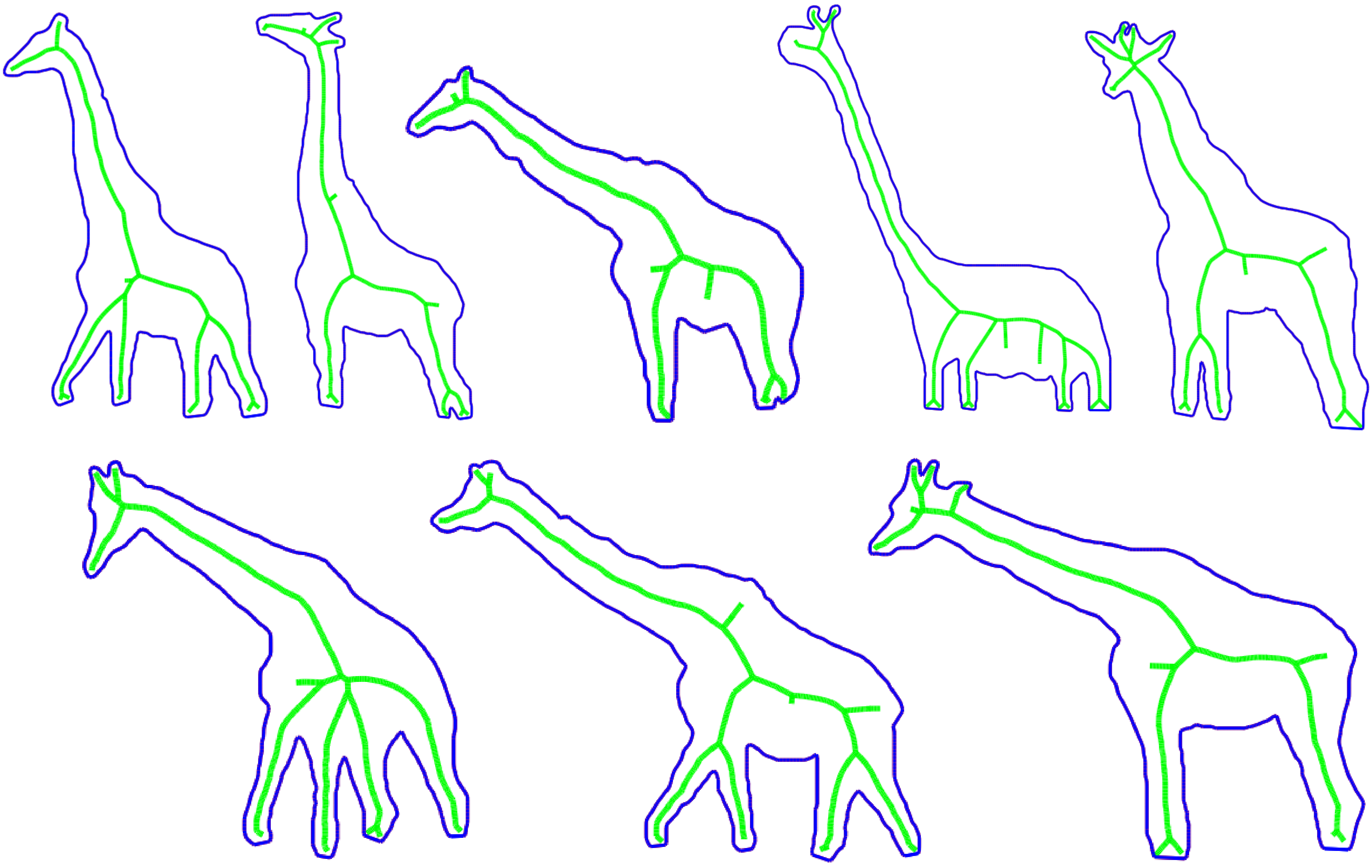


Courtesy of Vittorio Ferrari

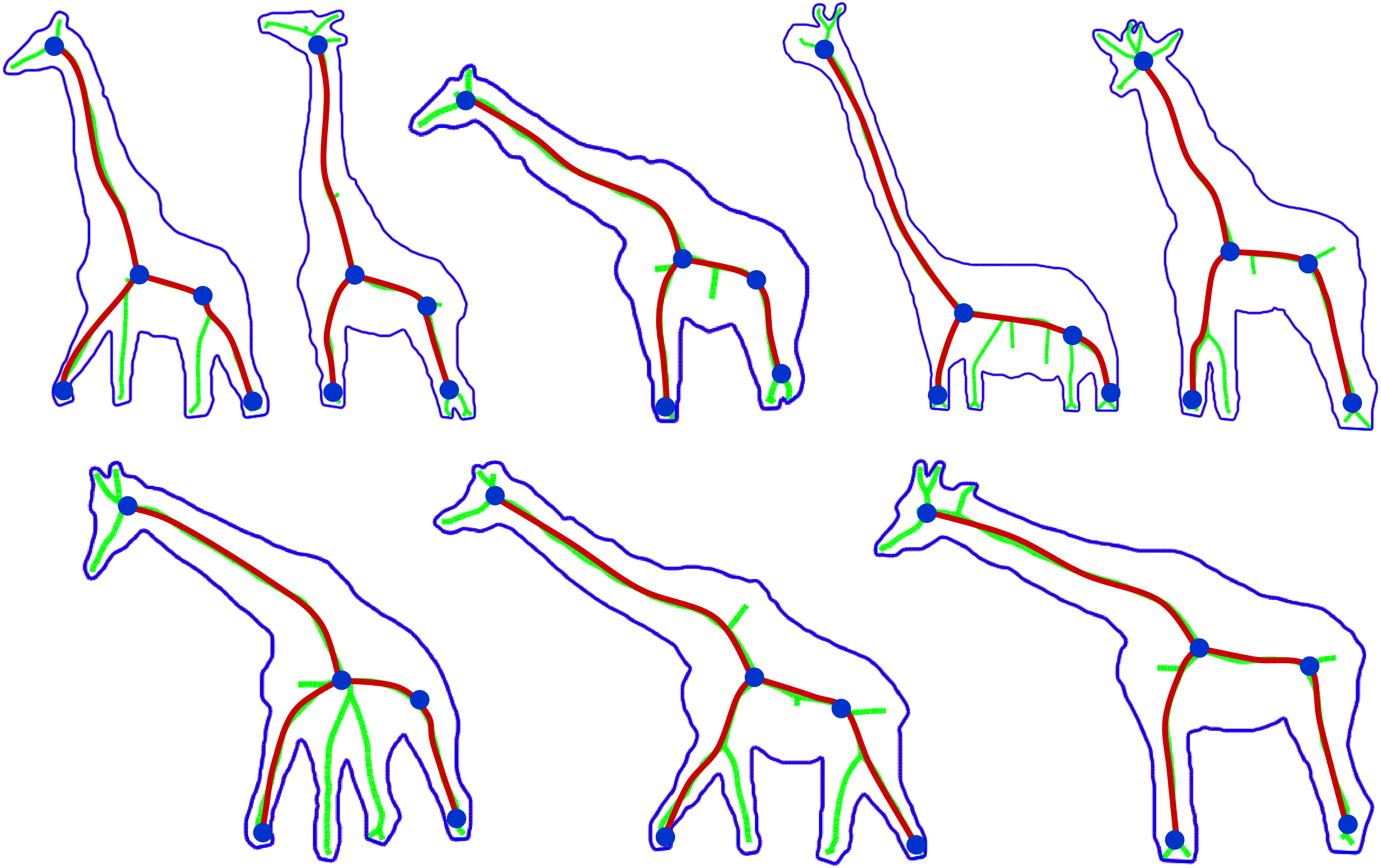
Giraffe Shapes



Giraffe Skeleton



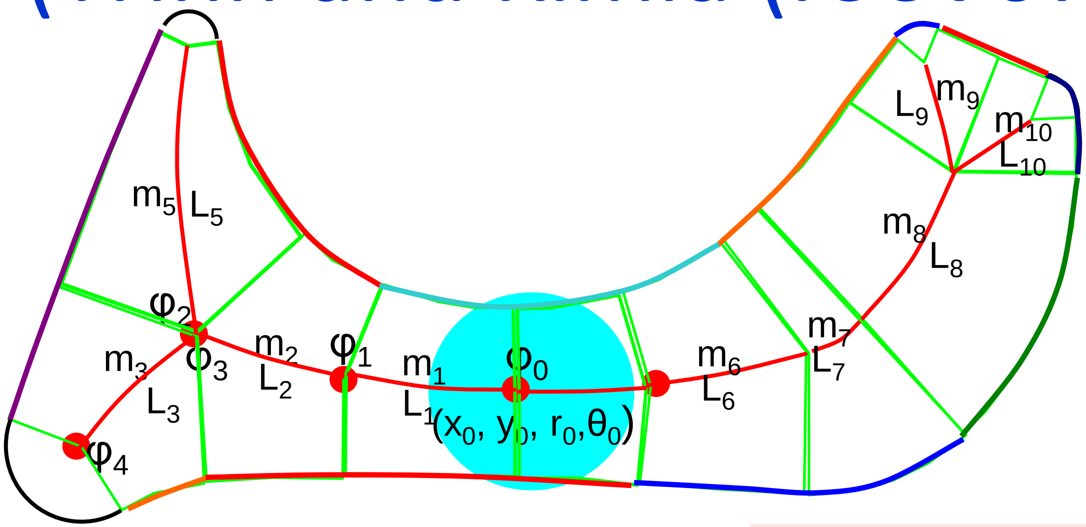
Shared Skeletal Topology

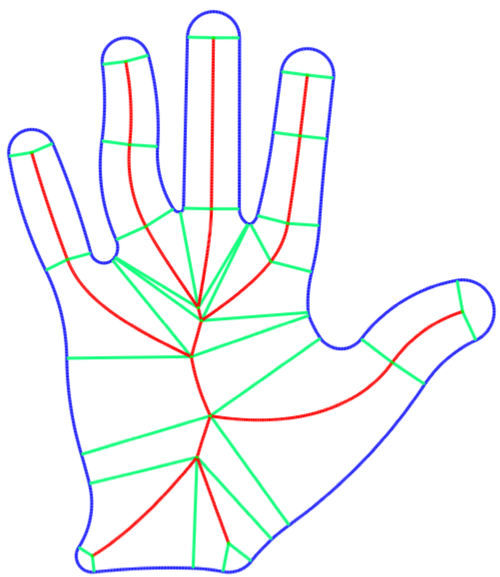


Idea:

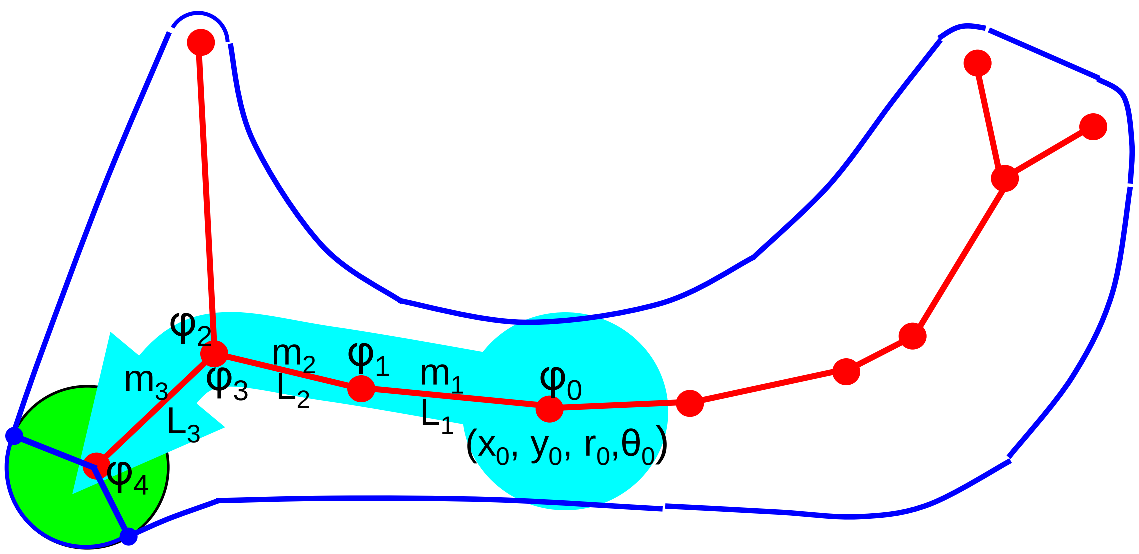
Represent a shape using its skeleton

Intrinsic Symmetry-based Shape Model (Trinh and Kimia (ICCV07)



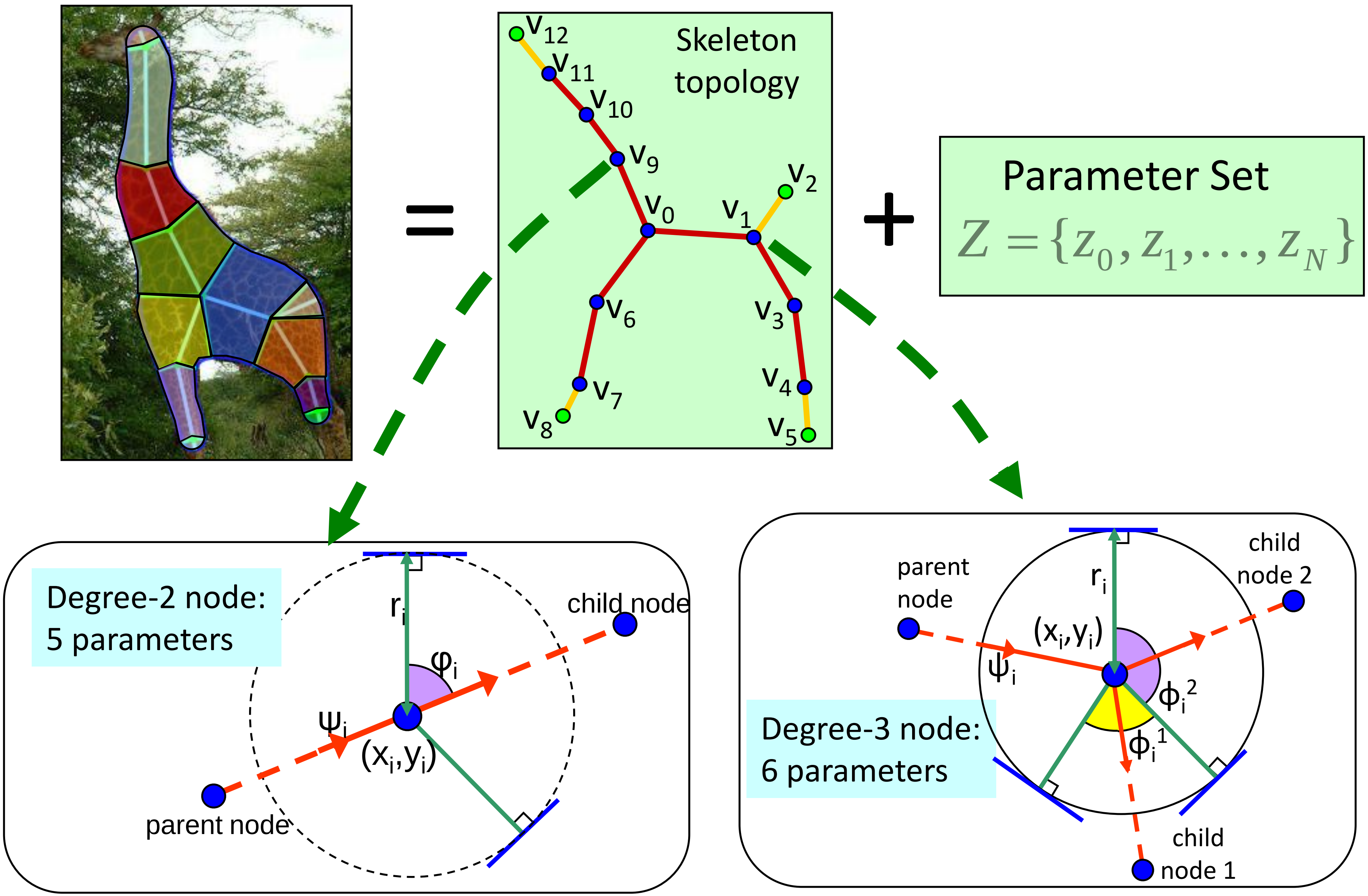


Intrinsic Shape Model for Segmentation

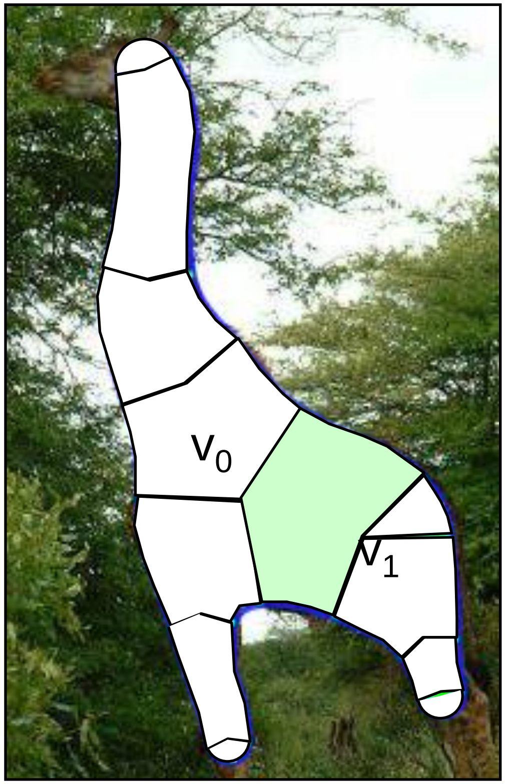


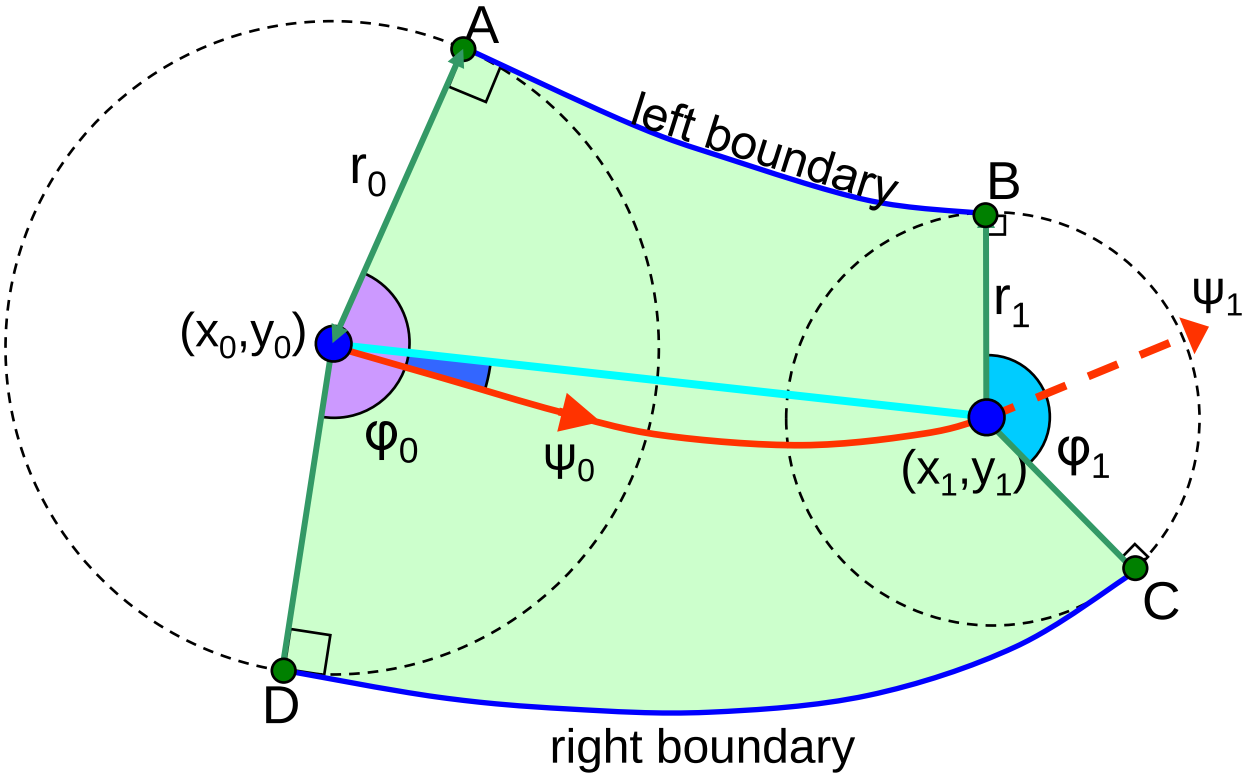
- Drawback: global dependency of each fragment's boundary on other fragments.
- New model: able to reconstruct each fragment LOCALLY from its adjacent nodes.

Fragment-Based Generative Model for Shape



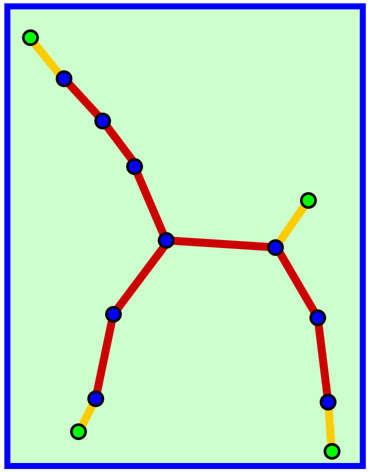
Reconstructing a Shape Fragment's Boundary



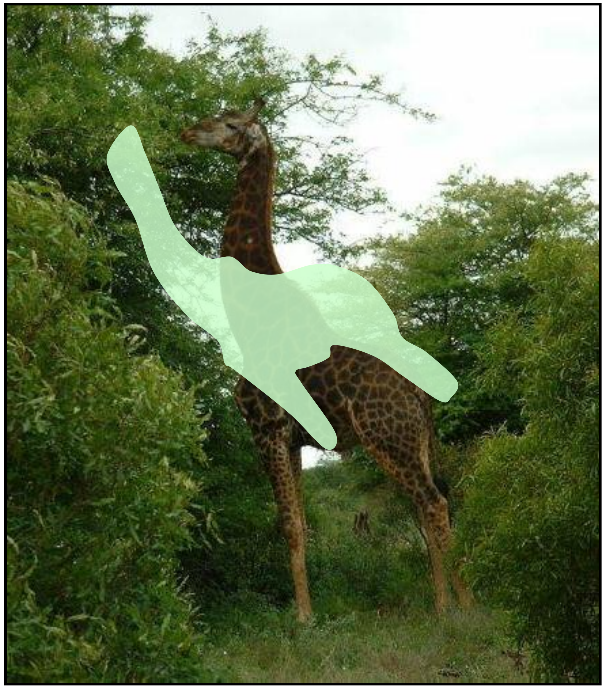


- Interpolate A→B and D→C contours using smooth bi-arcs (Kimia *et al.*, IJCV 2003).

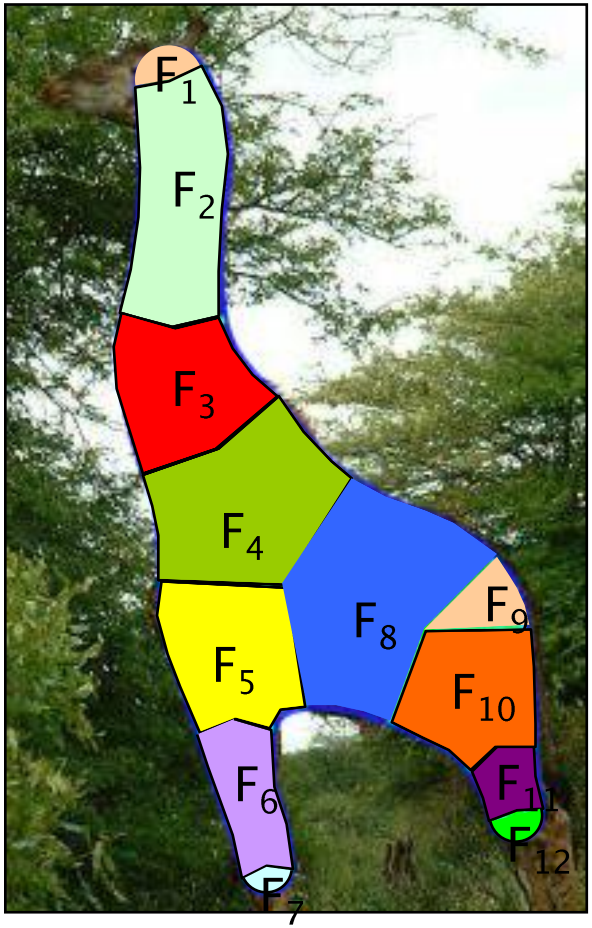
Generative Model



2. How to measure support for a shape in an image?



Cost function



- Cost of a shape = sum of its fragments' costs.

$$f(Z) = \sum_{i=1}^{N} f_i(F_i)$$
$$= \sum_{i=1}^{N} f_i(z_{\hat{i}}, z_i)$$

Cost of a shape fragment

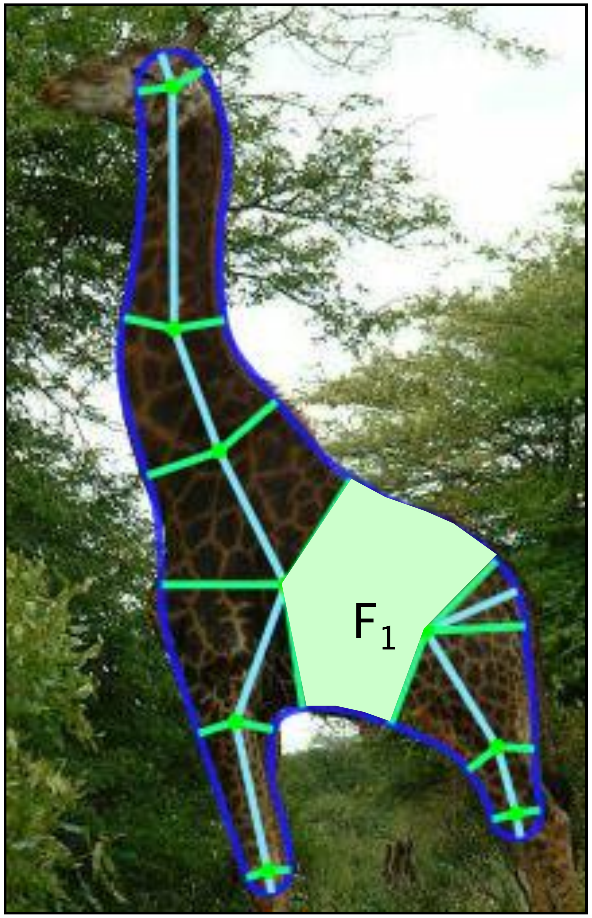
$$f_i(F_i) = g_i(F_i) + d_i(F_i)$$

cost of fragment

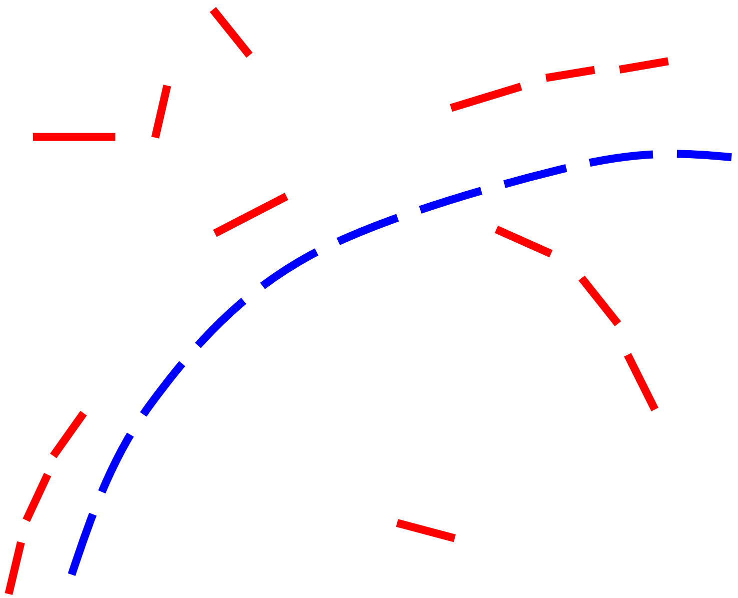
shape prior

image support

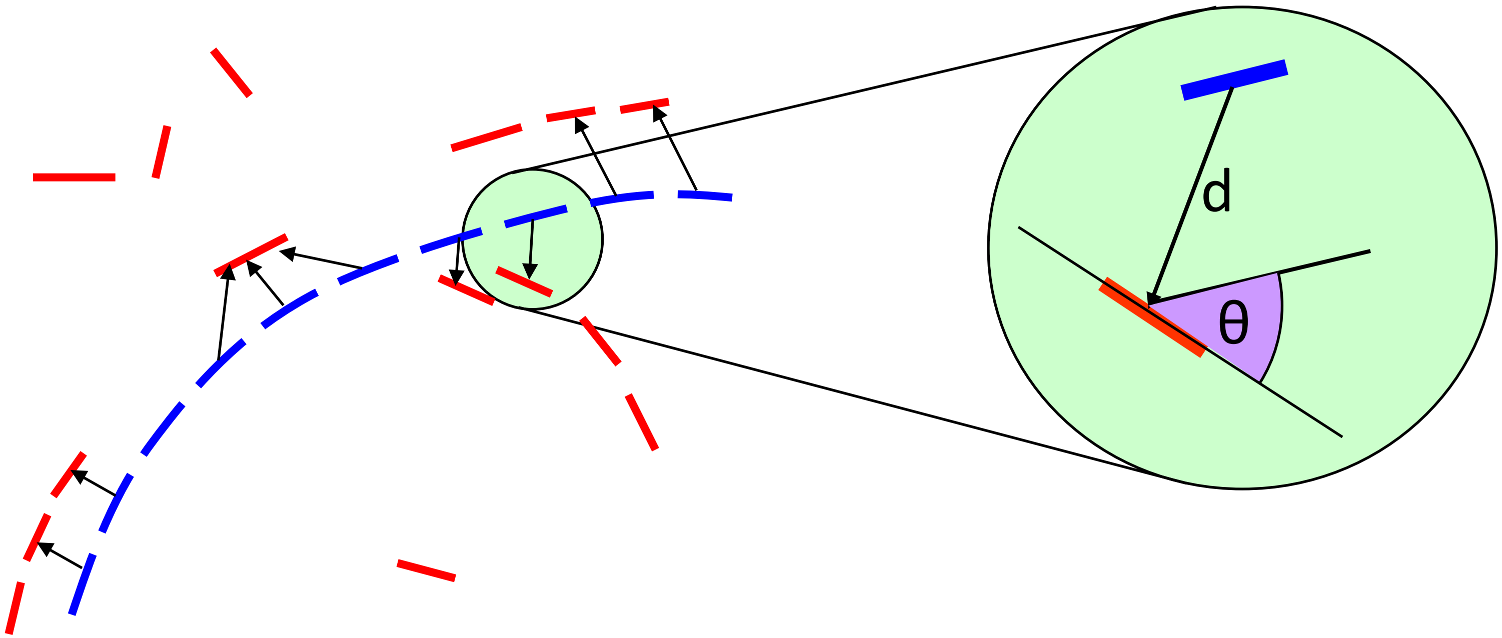
- Shape prior: uniform distribution on the fragment's intrinsic parameters.
- Image support:
  - Region appearance
  - Edge support for pair of boundary contours



Oriented Chamfer Matching (OCM)
(Shotton *et al*, PAMI'08 and Jain *et al*, CVIU'07)



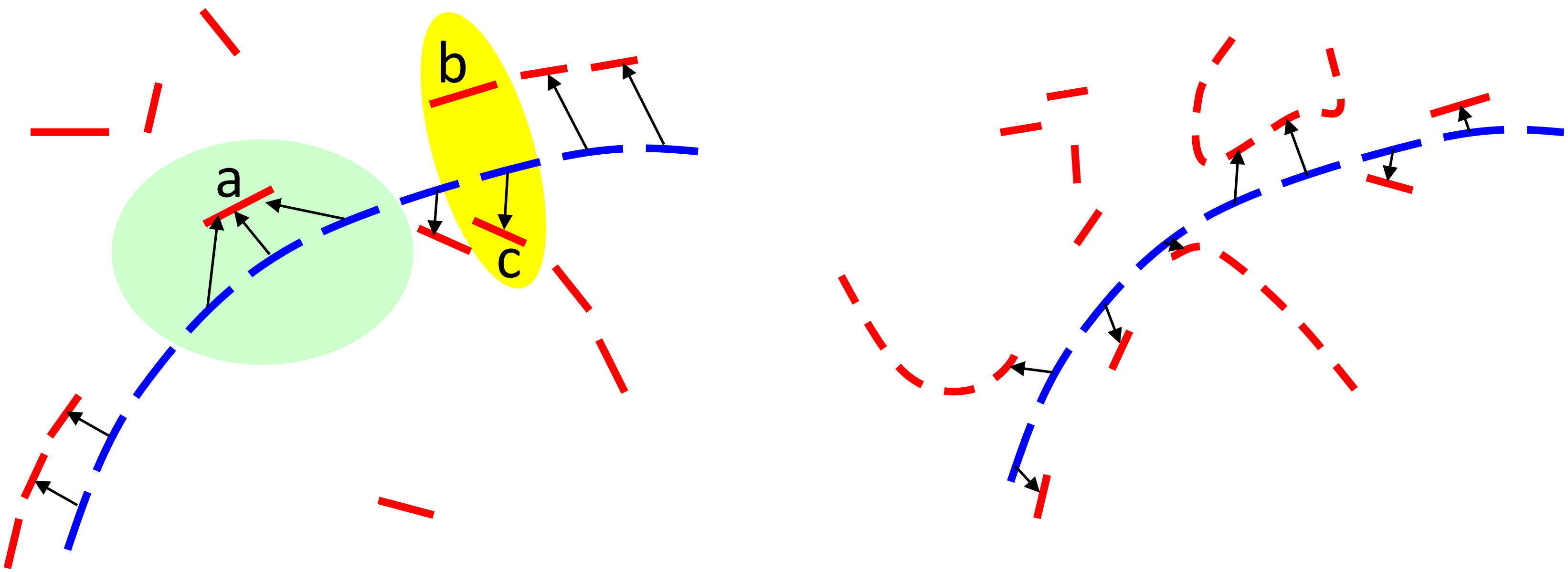
Oriented Chamfer Matching (OCM)
(Shotton *et al*, PAMI'08 and Jain *et al*, CVIU'07)



- Match each contour point to its closest edge
- OCM cost:

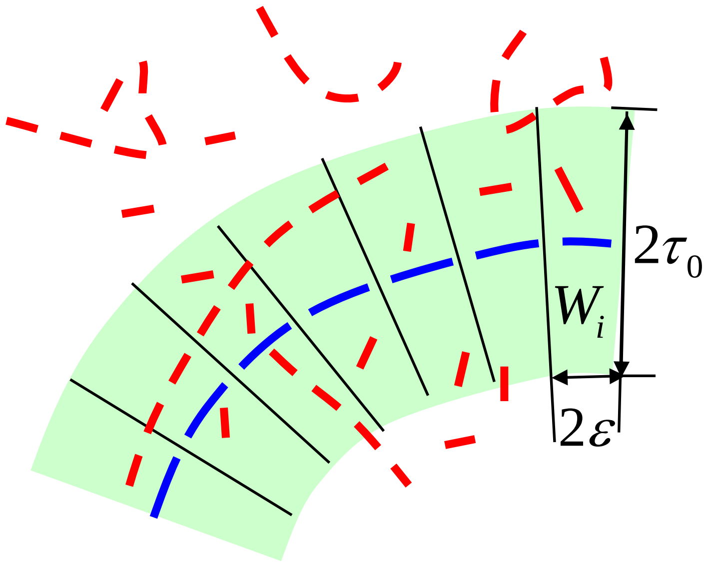
$$d_{OCM} = \frac{1}{N}\sum_{1}^{N}\left[\lambda_1 \underbrace{\min\left(\frac{d_i}{\tau_1},1\right)}_{\text{normalized distance}} + \lambda_2 \underbrace{\min\left(\frac{\theta_i}{\tau_2},1\right)}_{\text{normalized orientation difference}}\right]$$

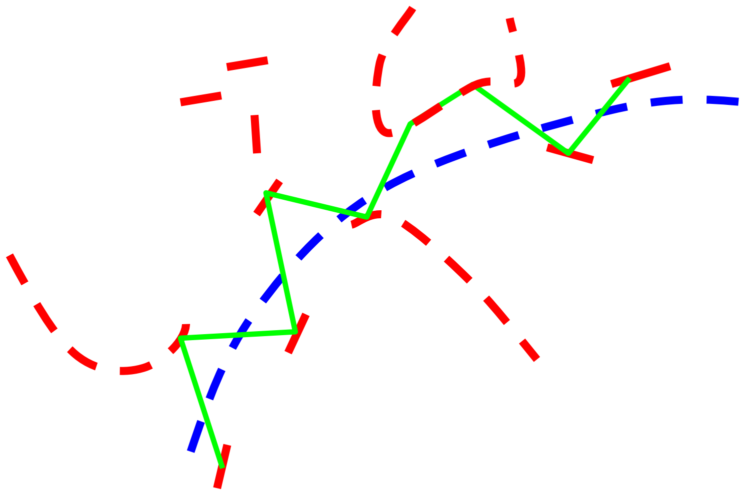
Drawbacks of OCM



- Over-counting support when edges missing.
- Under-counting support when many spurious edges present.
- Awarding accidental alignment.

Improvement: Contour Chamfer Matching (CCM)

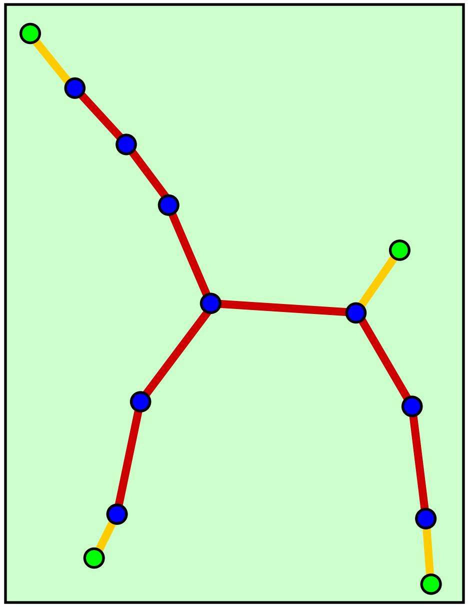




- Partition edges into thin stripes.
- Match contour points to image edges using OCM cost.
- Penalize orientation discrepancies between query contour and the contour connecting image edges.

How to search for the best supported shapes?

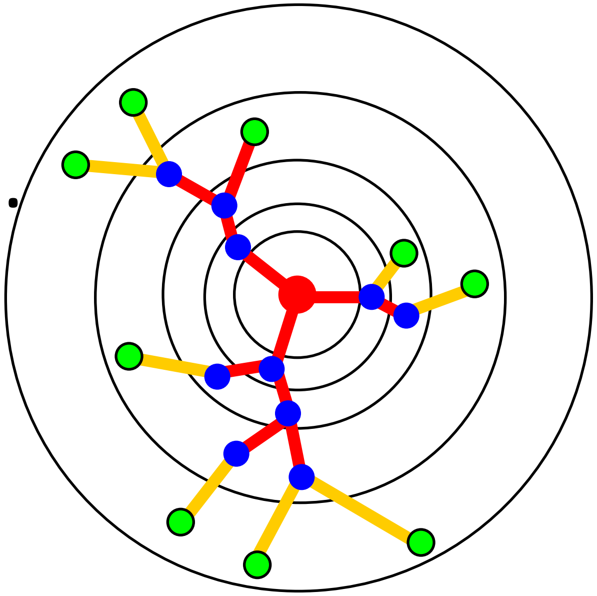
Single Global Solution



$$f(Z) = \sum_{i=1}^{N} f_i(z_{\hat{i}}, z_i)$$

$$Z^* = \arg\min_{Z} f(Z)$$

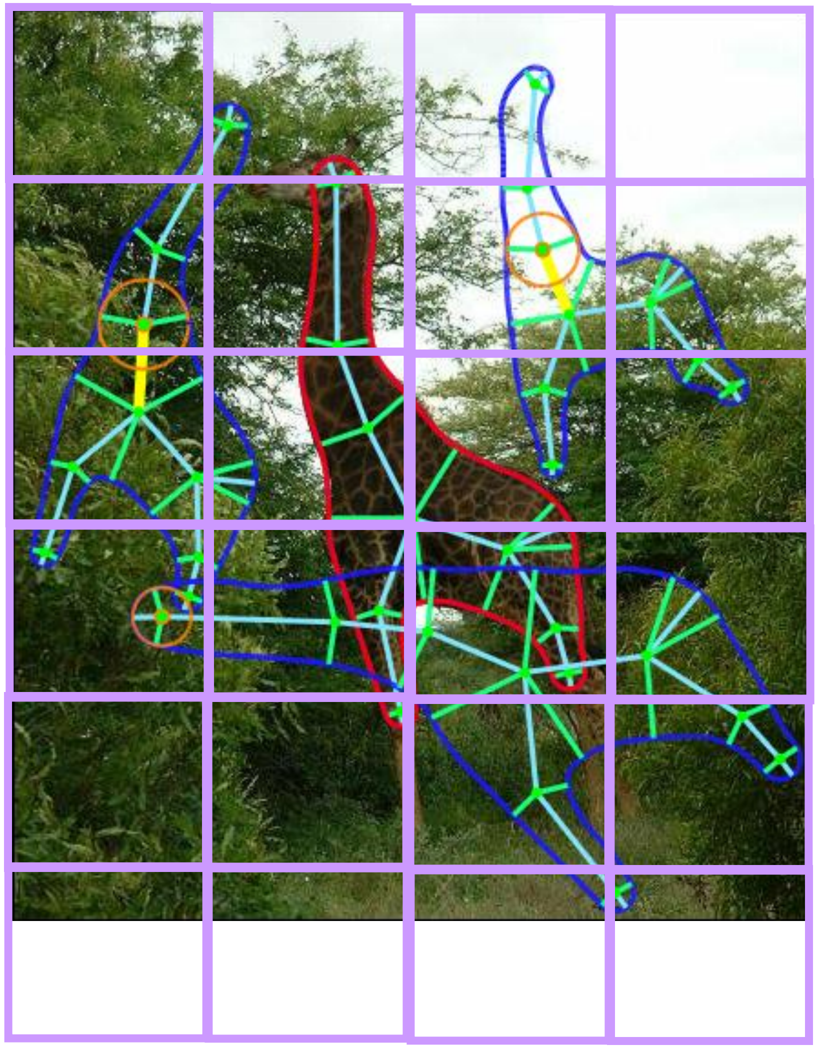
- Use Viterbi algorithm on a tree.



The need for multiple solutions

Single-Pass Multiple Solution Using DP

- Candidate pool: optimal solutions for each position of root node.
- Differential Exclusion Principle
- Trimming: discarding non-max solutions the candidate pool.



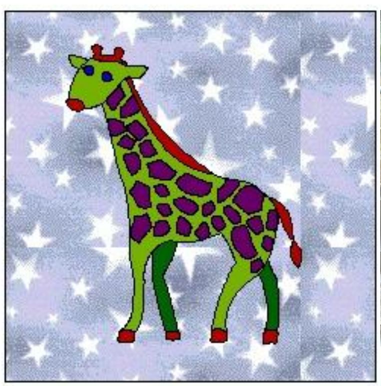
Experiments

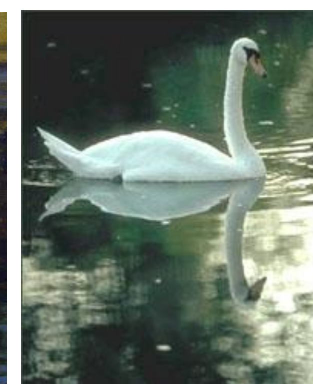
Dataset: ETHZ Shape Classes

- 255 images
- 5 categories: giraffes, bottles, applelogos, swans, mugs.



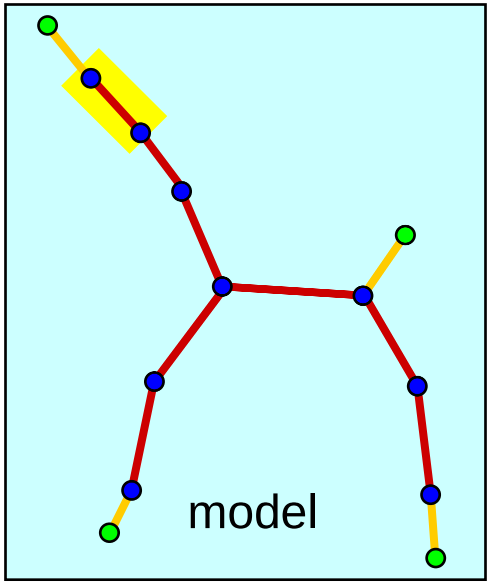


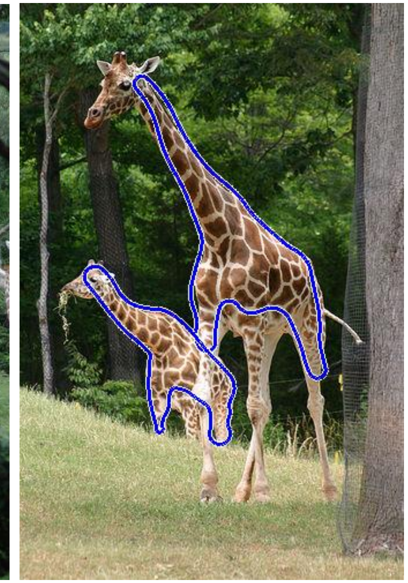


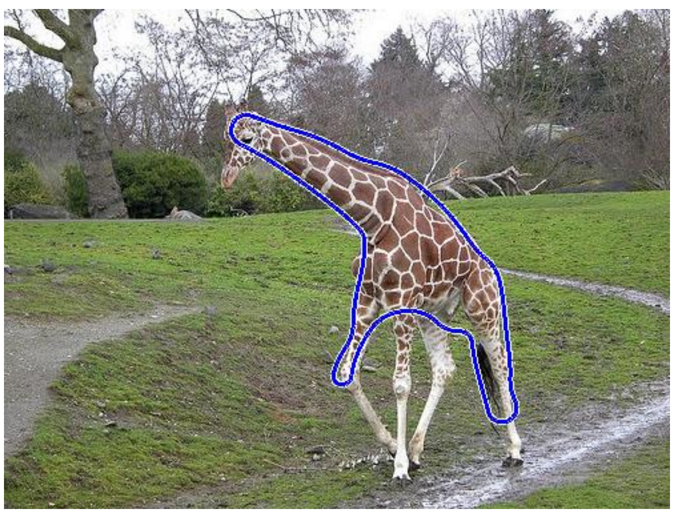



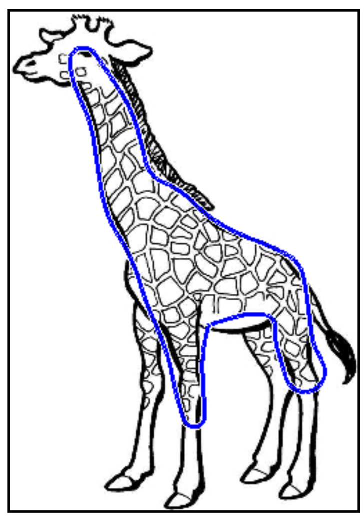
Courtesy of Vittorio Ferrari

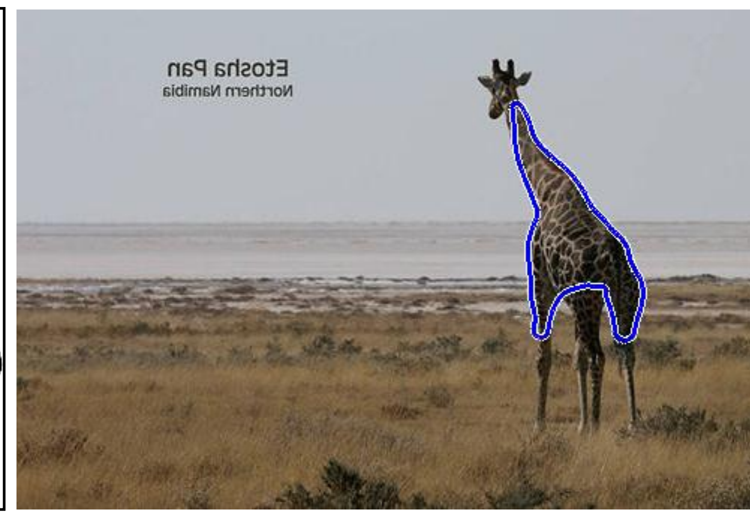
Detection / Segmentation - Giraffes



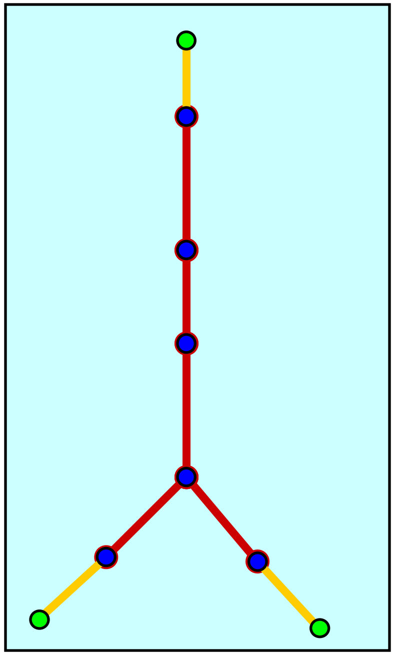




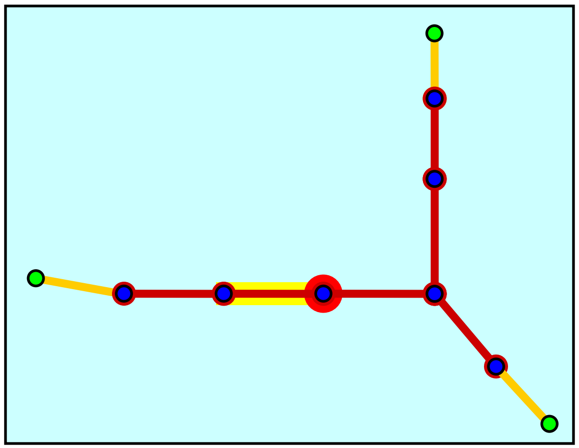


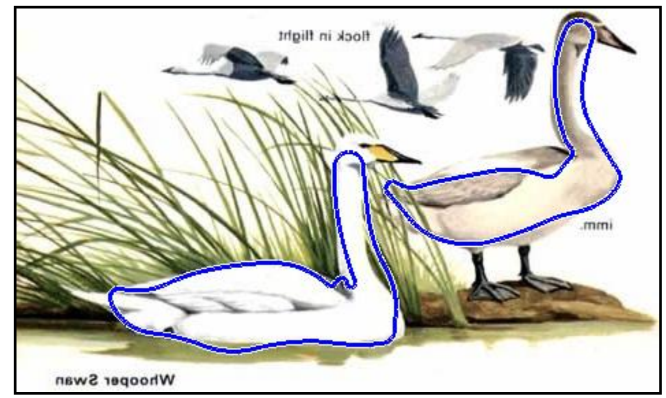


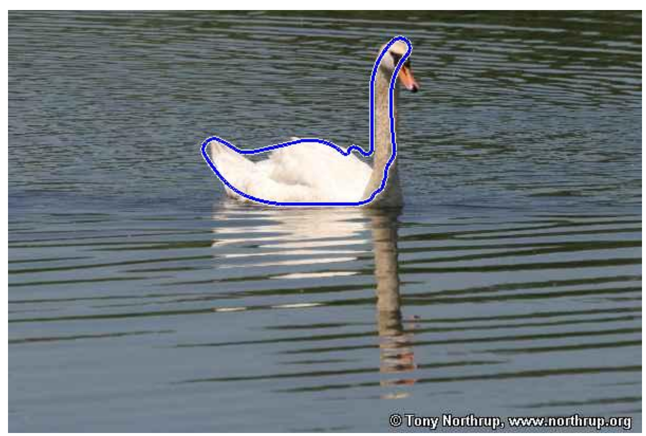
Detection/Segmentation - Bottles



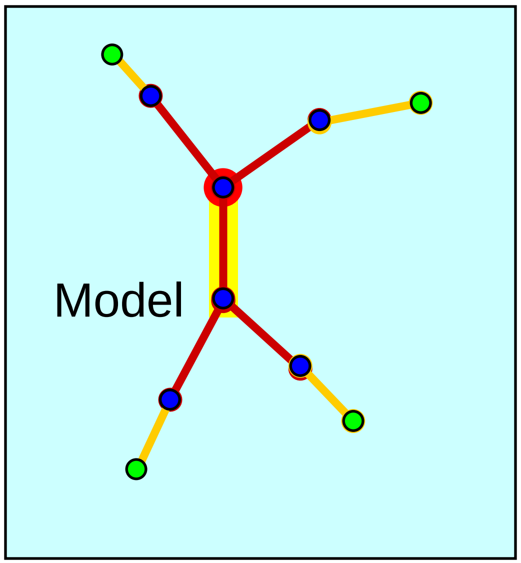
Detection/Segmentation - Swans

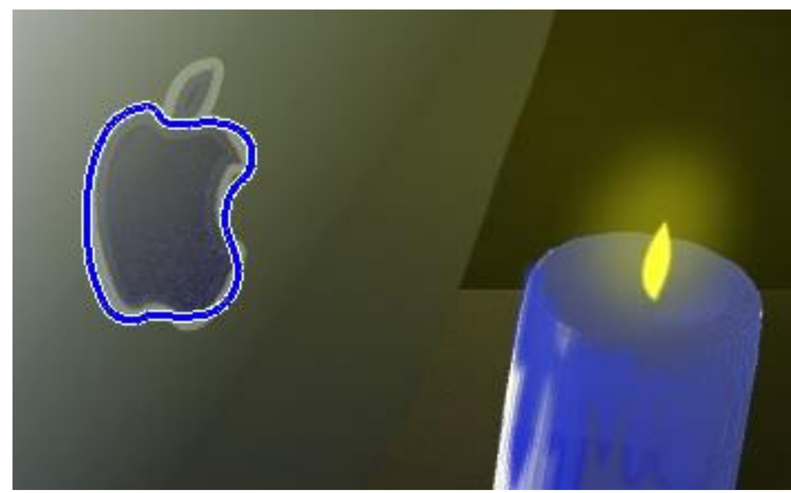




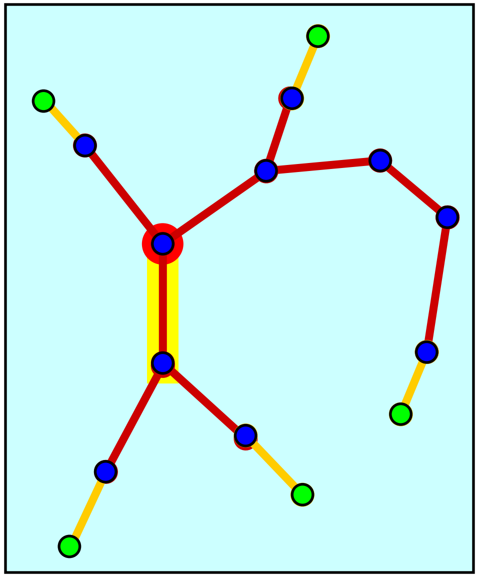


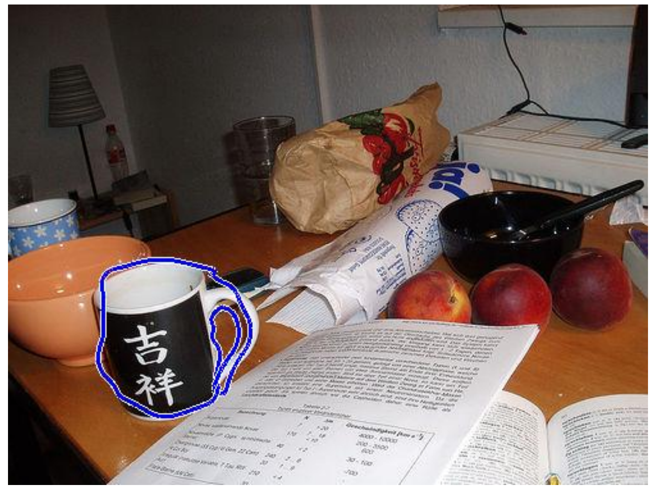
Detection/Segmentation - Applelogos

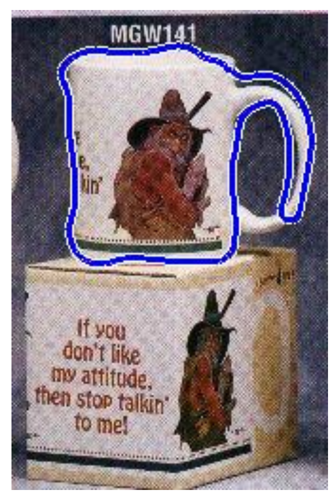




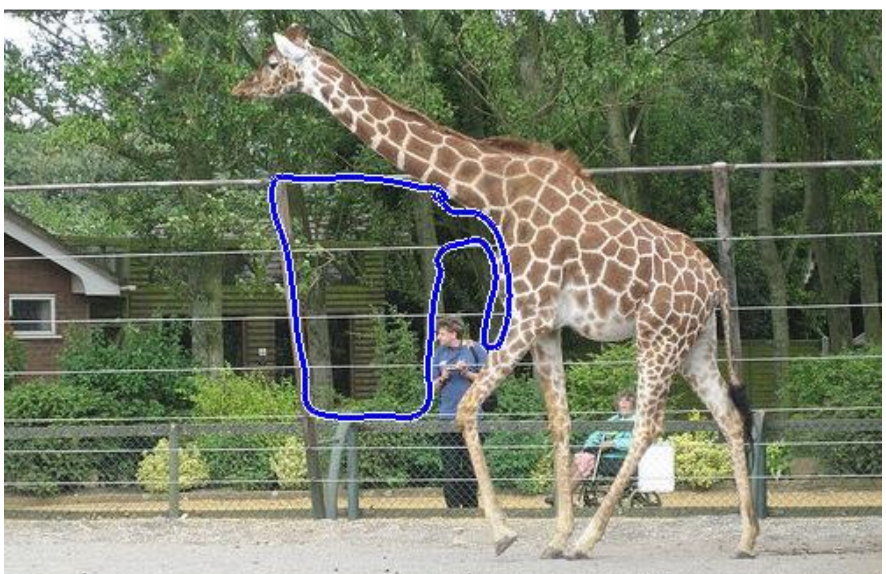
Detection/Segmentation - Mugs





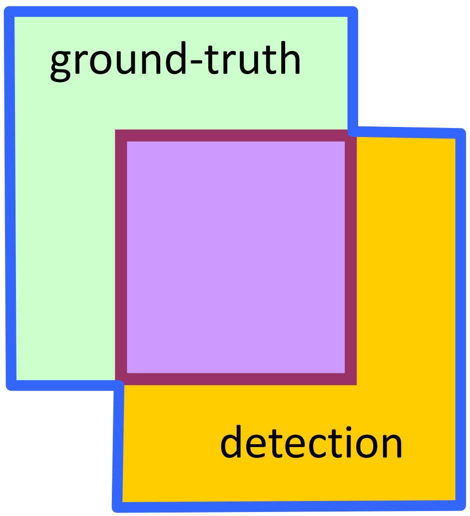


False Positives



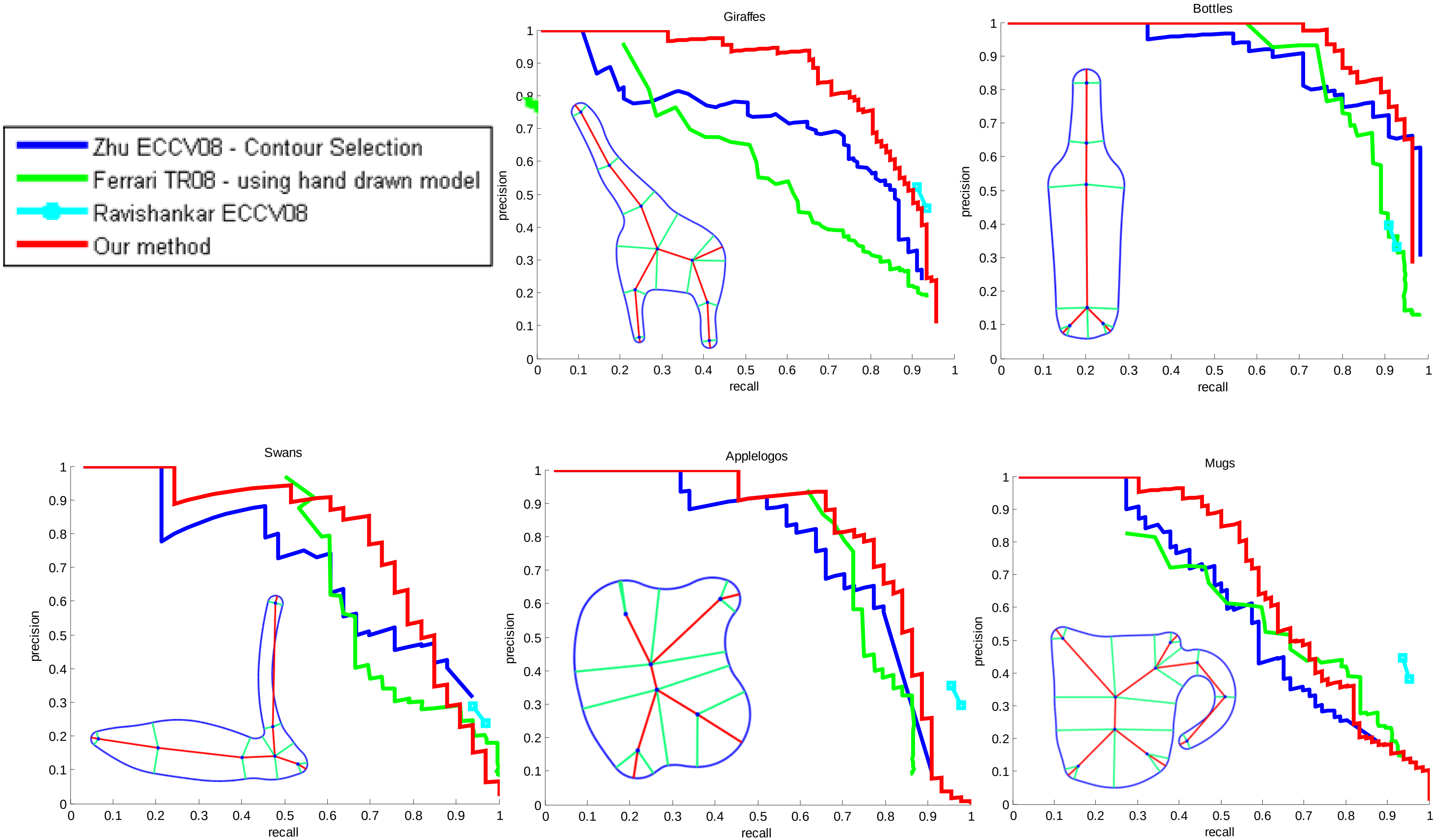
Object Detection Evaluation

- PASCAL criterion:



$$\frac{area(\det \cap gt)}{area(\det \cap gt)} \geq 0.5$$

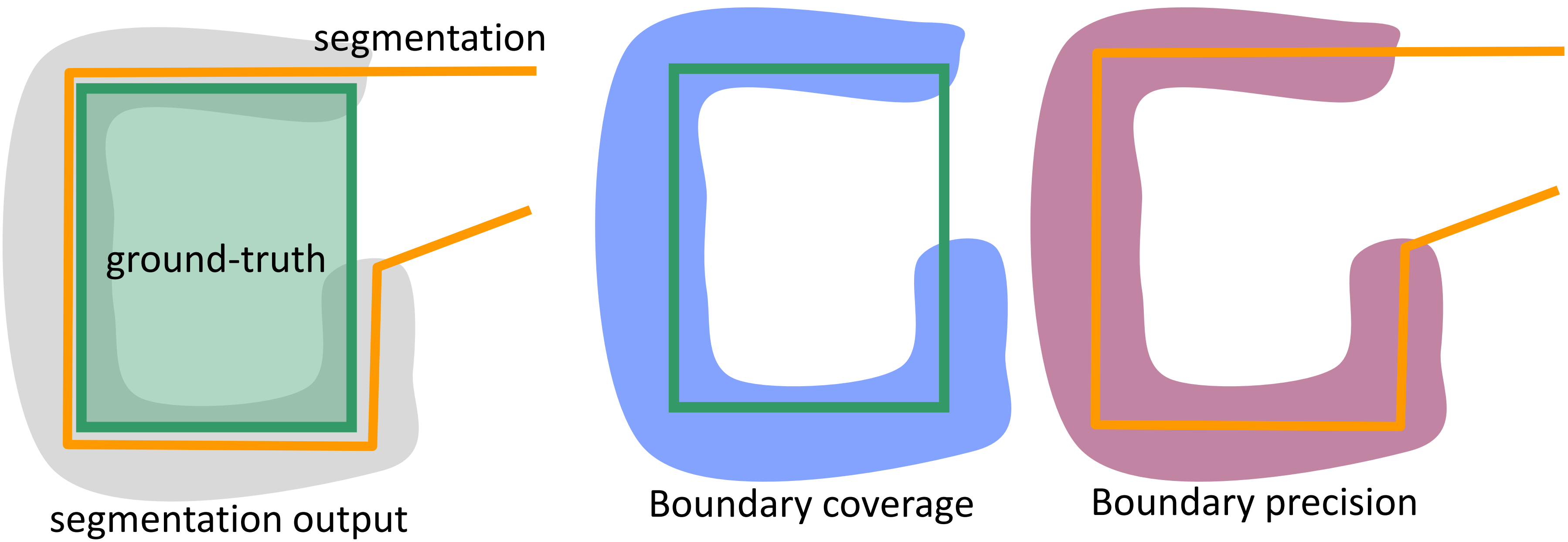
Object Detection Performance



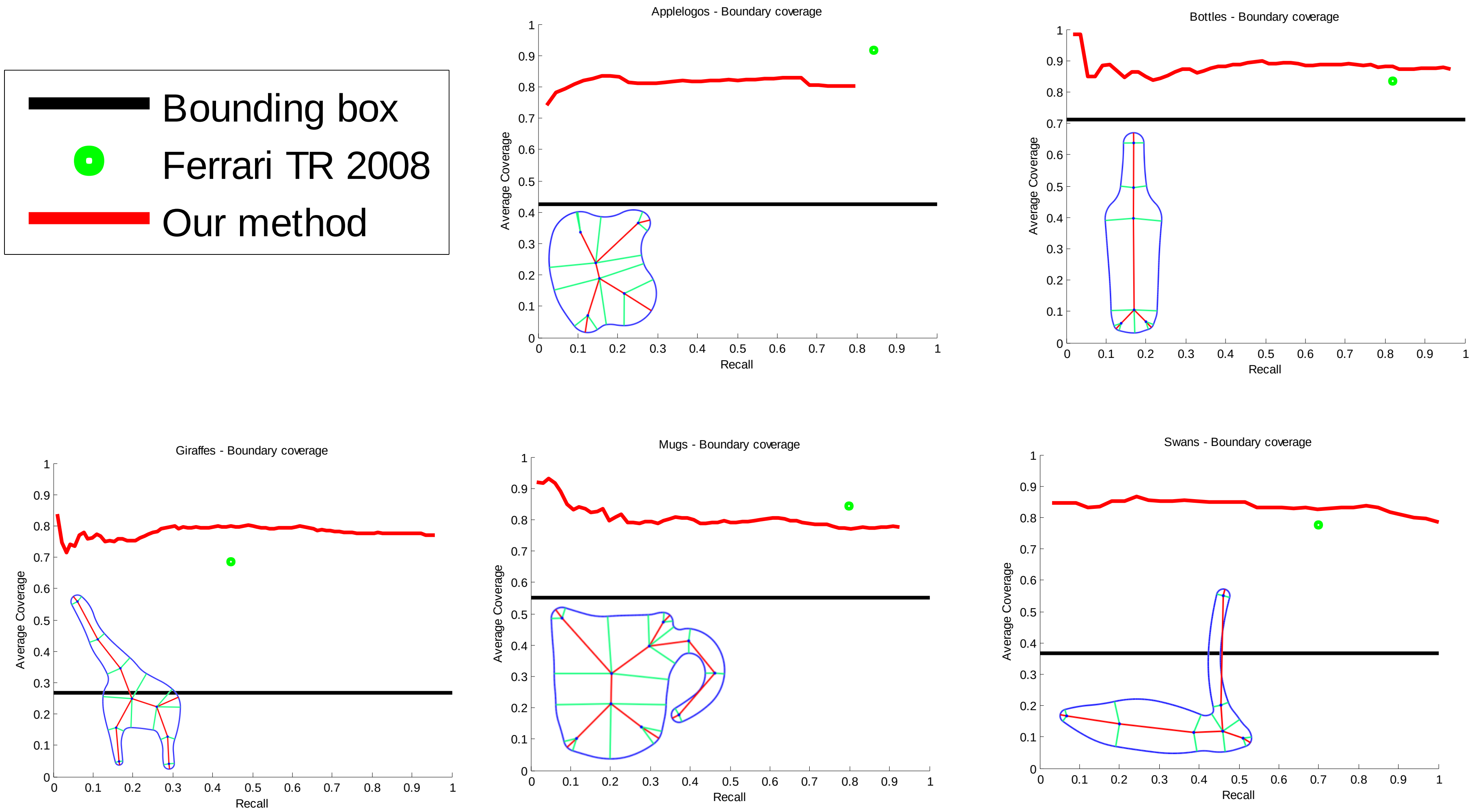
Evaluation: Segmentation Performance

(Ferrari *et al*, INRIA Tech Report 2008)

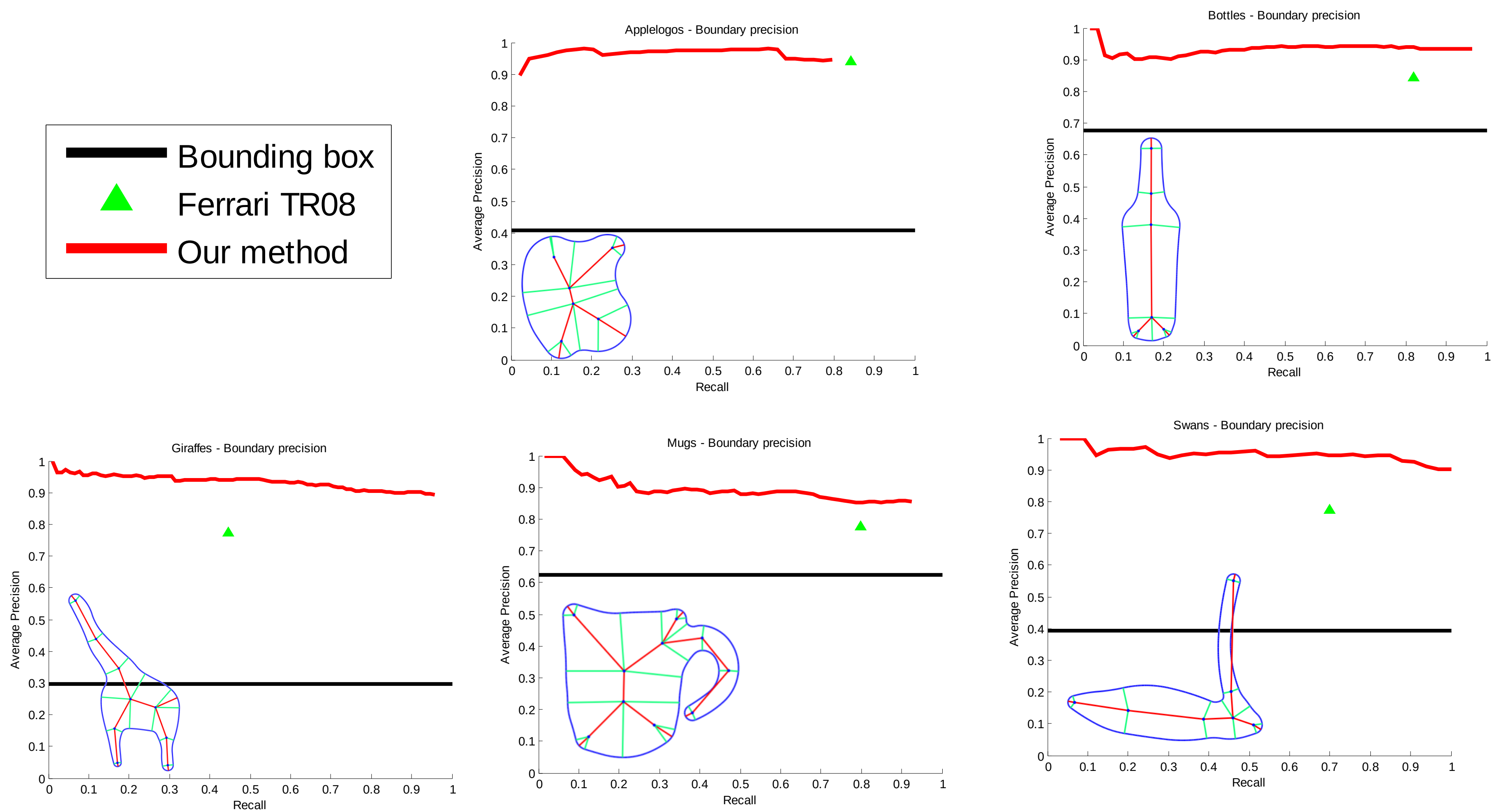
- Boundary Coverage: proportion of the ground-truth that is close to the segmented shape.
- Boundary Precision: proportion of the segmented shape that is close to the ground-truth.



Performance – boundary coverage



Performance – Boundary Precision



Summary

- A skeleton-based generative model for shape where each fragment can be reconstructed locally.
- Improvement to Oriented Chamfer Matching cost.
- Extension to Viterbi algorithm to compute multiple solutions in a single pass.

Thank you

Questions?

Email: ntrinh@lems.brown.edu