

Object Interpretation: Extending and Validating Object Recognition

Guy Ben-Yosef (1), Liav Assif (1), Daniel Harari (2), Ethan Fetaya (1), Shimon Ullman (1)

- 1) Department of Computer Science and Applied Mathematics, Weizmann Institute of Science
- 2) Department of Brain and Cognitive Science, Massachusetts Institute of Technology

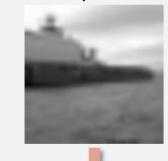
Motivation & Goal Motivation: To explore small local configurations Typical part model xplore 'out-of-part': Each part remains ambiguous Recognition is achieved by combinations of parts to compensate Contour 2 (mane contour) Full Interpretation explore 'in-part': Point 1 (the eye) Validate the part by full interpretation Recognition is achieved from single part Contour 4 (the lower-neck

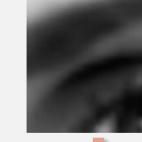
The 'full interpretation' problem:

To obtain the identity of the object, and also the identity and locations of a rich set of internal object features.

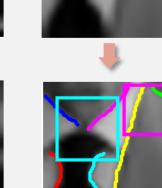
Goal: To produce full interpretation automatically

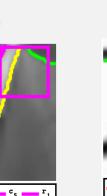
Examples for 'full interpretation':



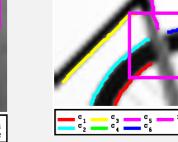












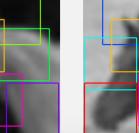
Currently there are no computational models for such full interpretation

Related models include:

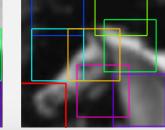
Deformable Part Model

[Felzenszwalb et al. (ver 2012) Internal features are image regions with their likelihood and relative position.

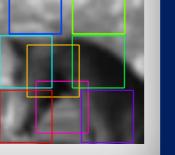












Problem Setting

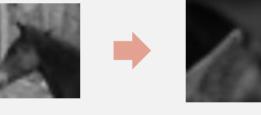
Input: small recognizable local configurations

- 1) Interpretation & recognition are possible on small configurations of the object in a fully local manner.
- 2) The number of internal features needed for interpretation is small. Hence a 'full interpretation' task can be computationally practical.
- 3) Components of such small local configurations are not likely to be recognized on their own. If so, recognition is done by the *relations* that hold between

Output: Interpretation & Recognition

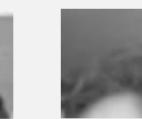
1) When humans recognize an object image, they also recognize its internal features. We assume that reaching human recognition performance must involve accurate full interpretation of the image.

Visually similar negative examples:









Approach

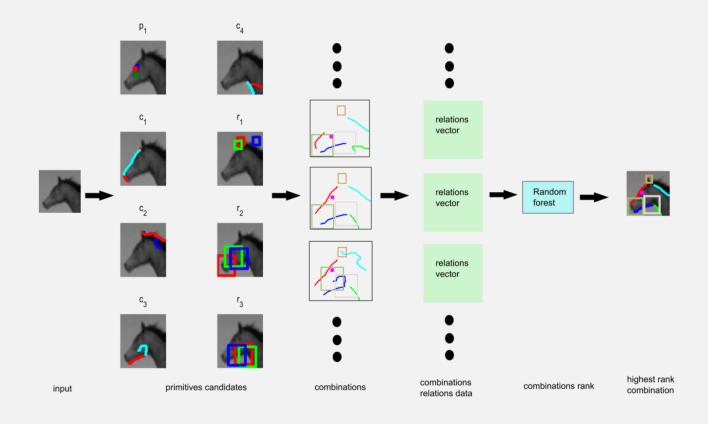
The interpretation scheme is based on a set of primitives and relations. The primitives are divided into three types, 2-D(regions), 1-D(contours), and 0-D(points).

The relations include unary (properties), binary, and more global relations between three or more primitives. Our set of relations go beyond relative location and relative distance. Examples are:

- local intensity extrema
- parallelism and continuity between two contours
- containment of point feature in region

The Interpretation process:

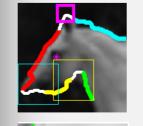
- · 'ends-in' relation between contour and point/region
- Cover of point feature by contour
- Segmentation and texture support along contours and between contours

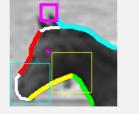


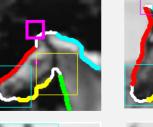
- Extract image measurements for candidate primitives of type points, contours, and regions
- 2) Score combinations of primitive candidates by their comparability with learned relations.
- Select the maximum-score combination as the final interpretation of the object structure.

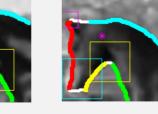
Model results & Discussion

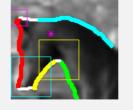
Positive examples

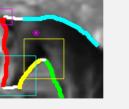


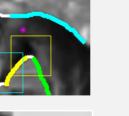


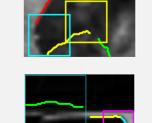


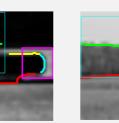


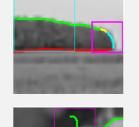


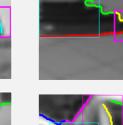




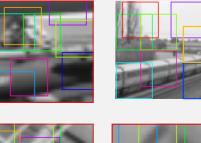












Top Negative examples DPM



Validation

