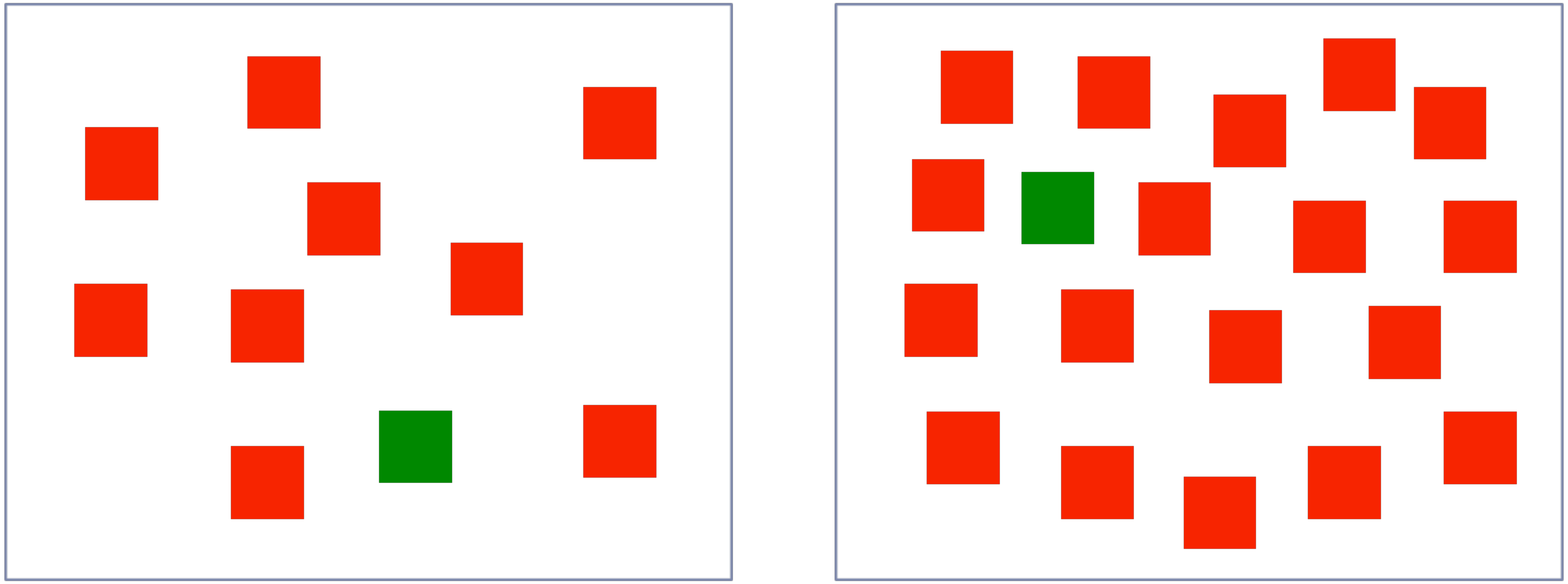
Gestalt Principles for Attention and Segmentation in Natural and Artificial Vision Systems

Gert Kootstra, Center for Autonomous Systems, KTH

Visual search

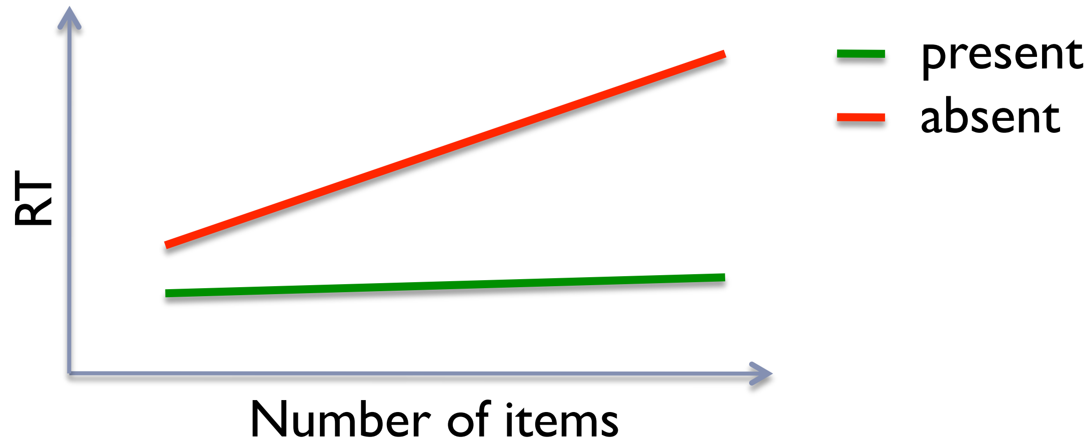
- Pop-out experiments: basic features



(Treisman & Gelade '80)

The pop-out effect

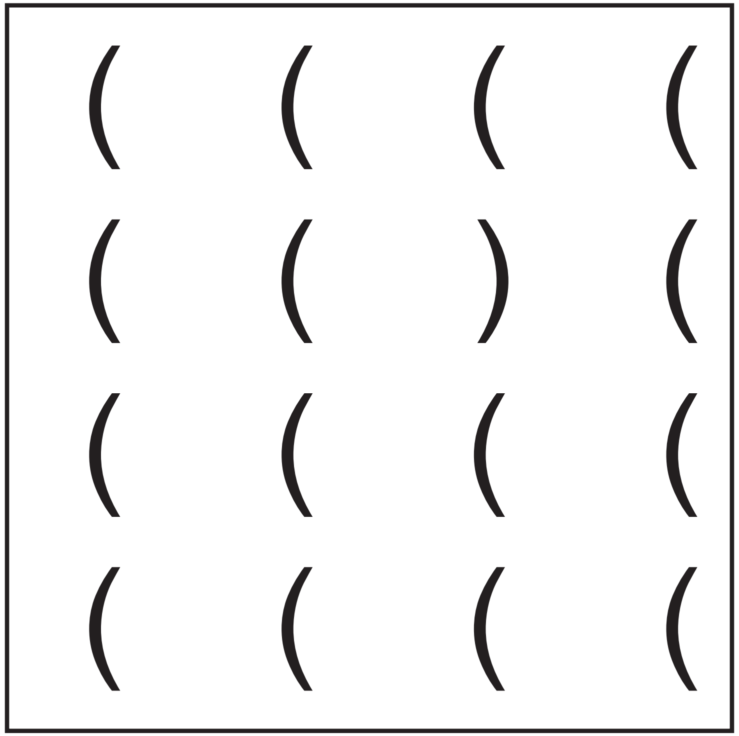
- This stimulus results in a pop-out effect



- Efficient search for the target

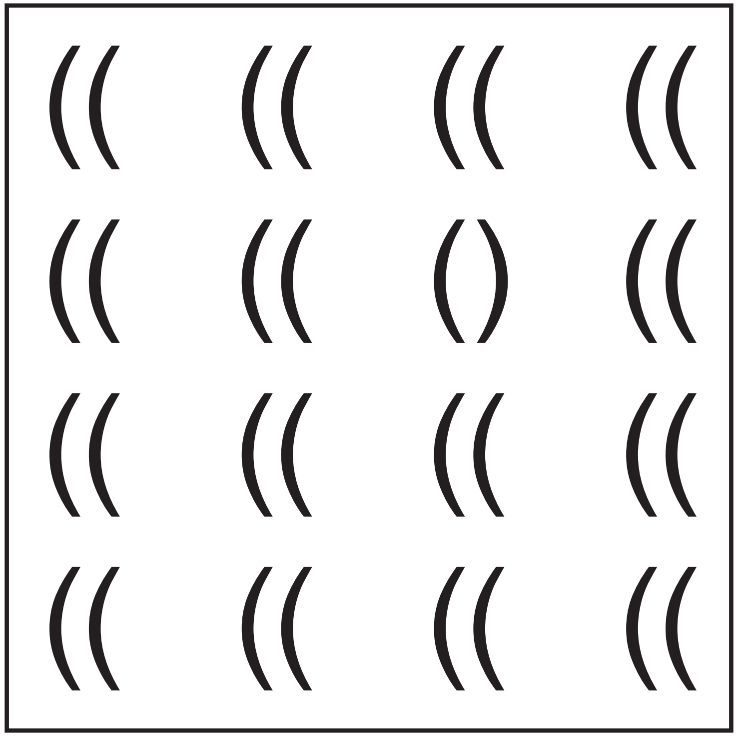
Less efficient search

- Find the) among (



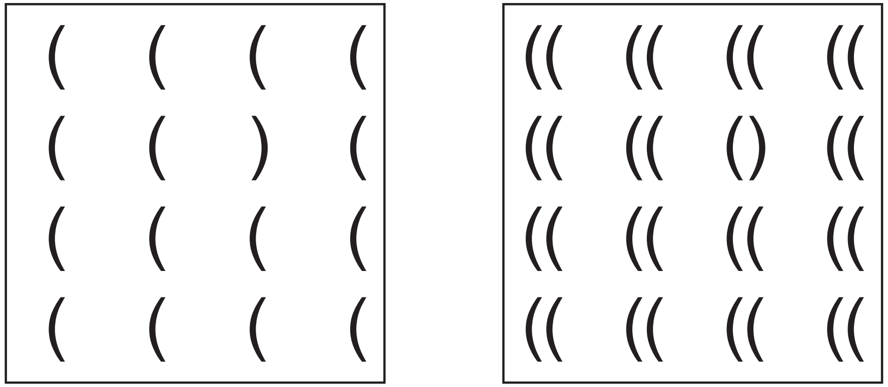
Configural superiority

- Now make it harder by adding more items



Configural superiority

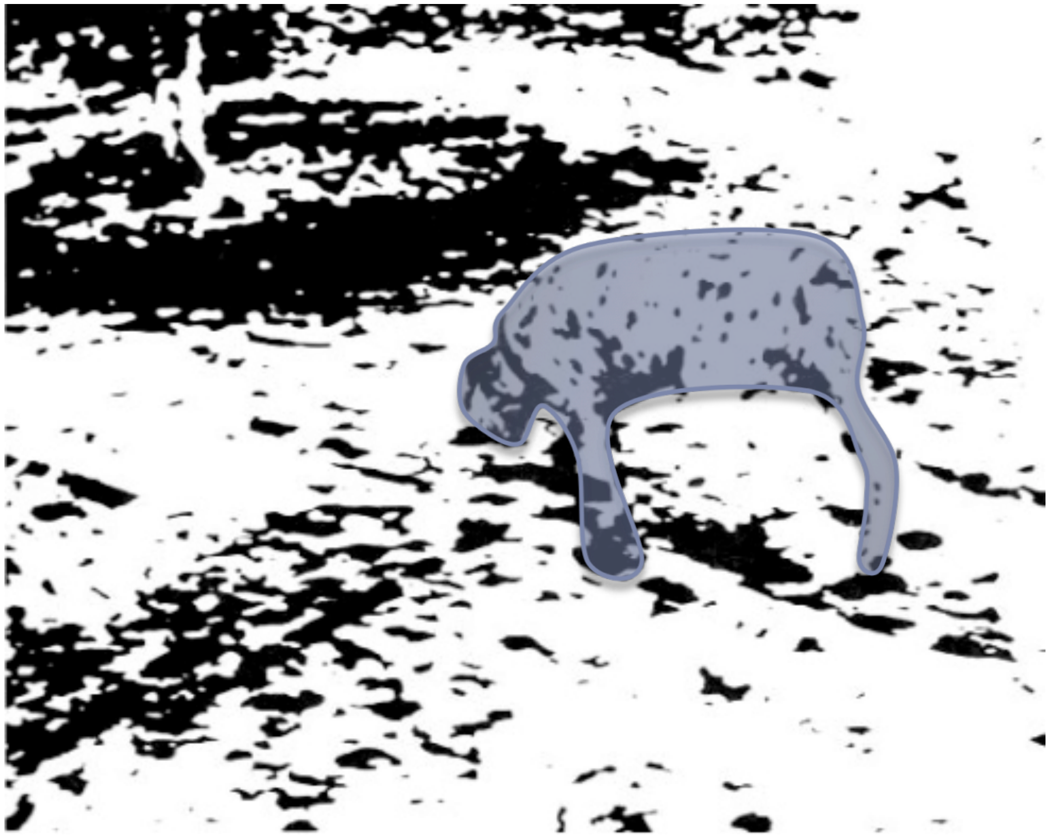
- Search becomes more efficient
- Configural superiority
  - Not perception of 32 items, but 16 figures, () or ((
- Features



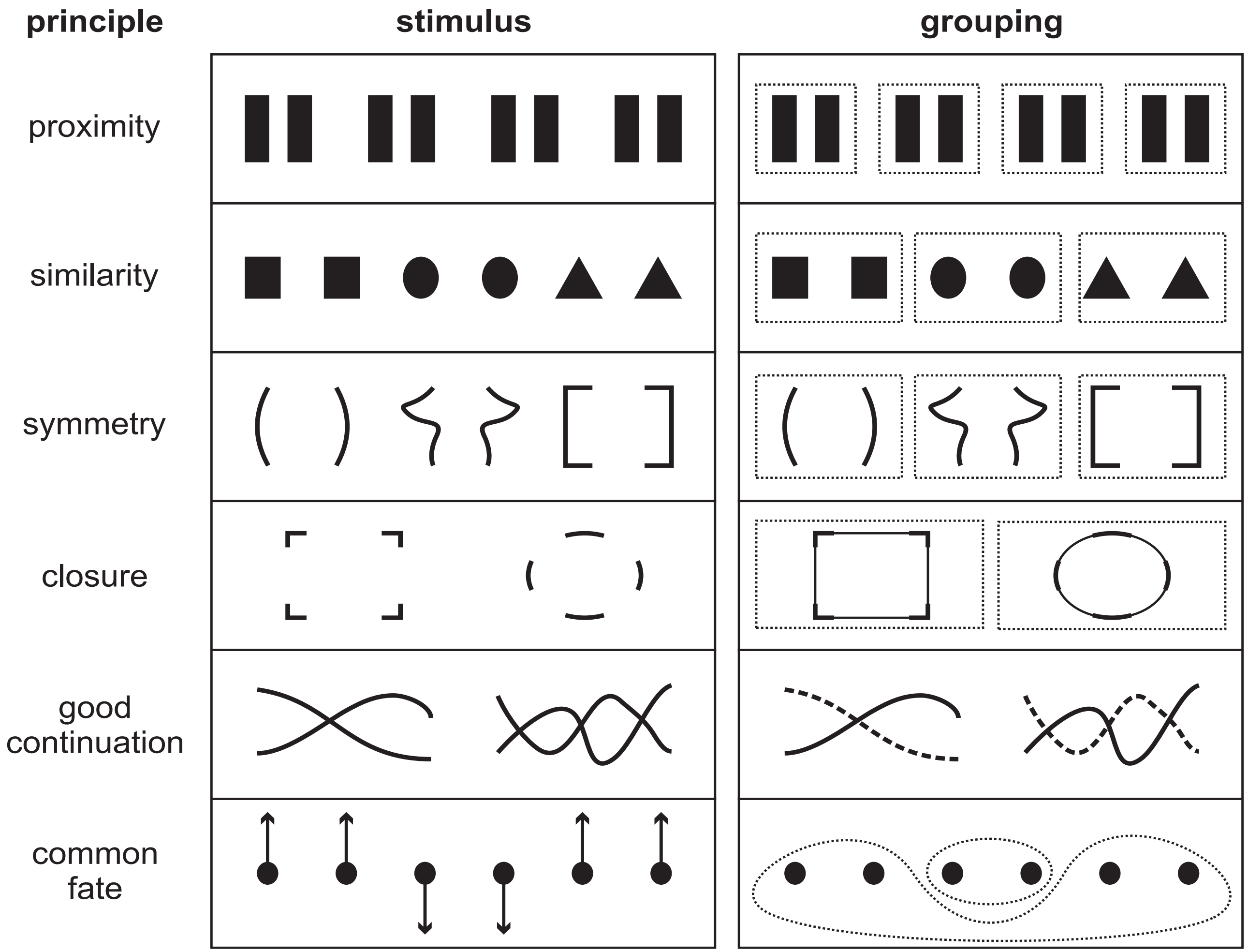
  - Curvature
  - Emerging features: symmetry and closure
- Higher-level features subsume basic features

Gestalt: From parts to wholes

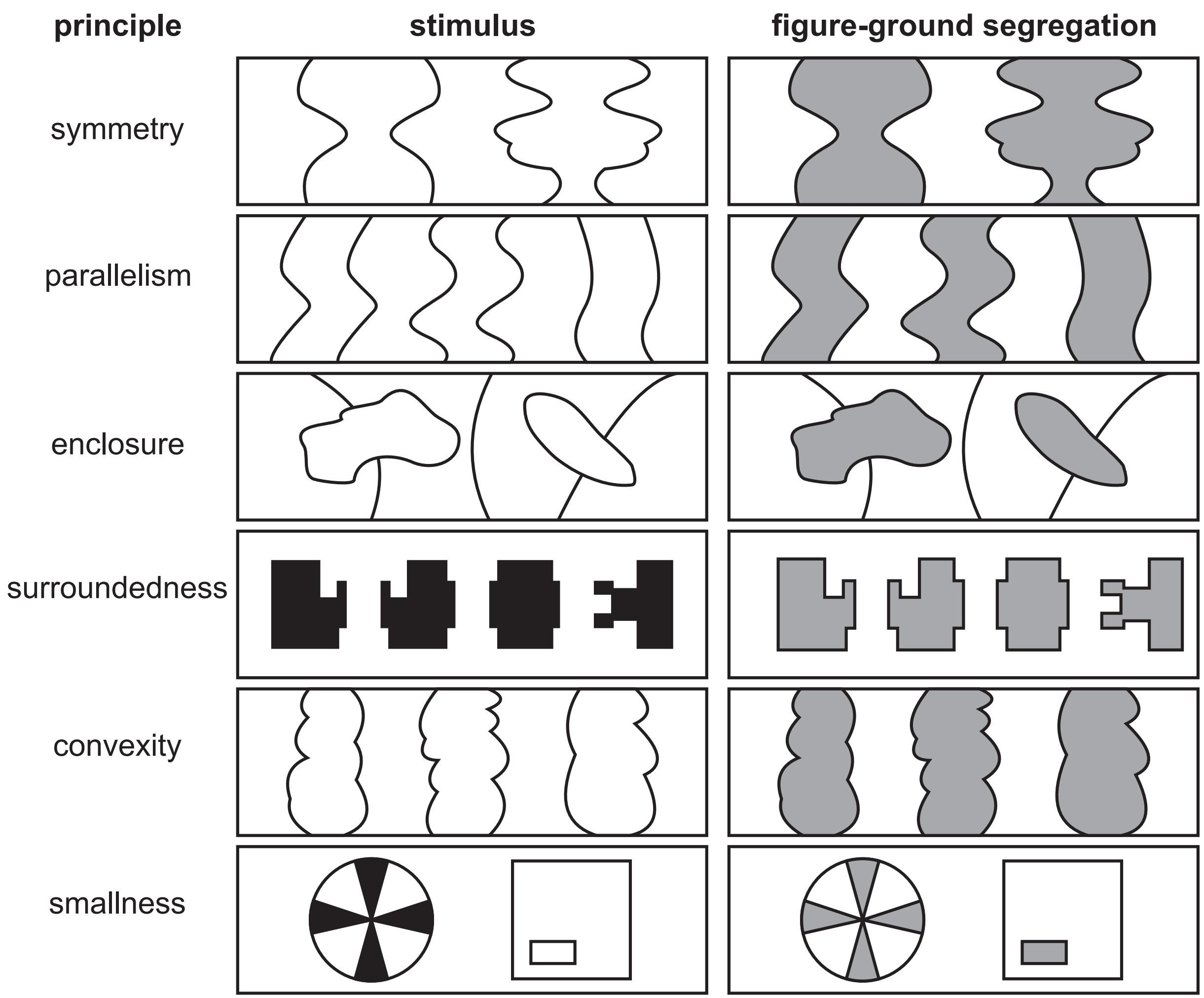
- The visual system has the tendency to group the parts into larger whole (objects/scenes).



Gestalt: Grouping



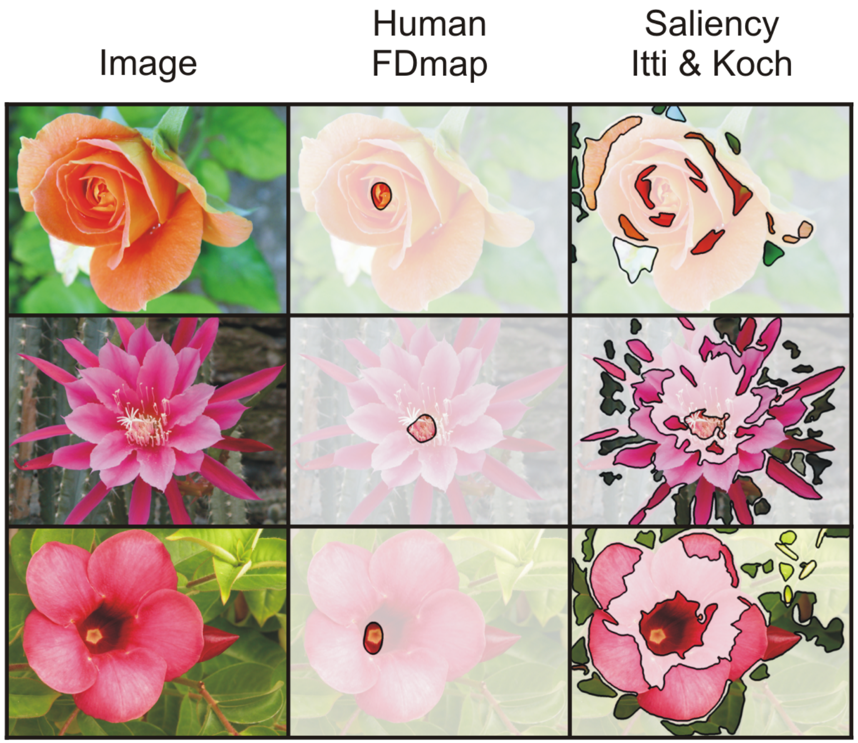
Gestalt: Figure-ground segregation



Visual Attention

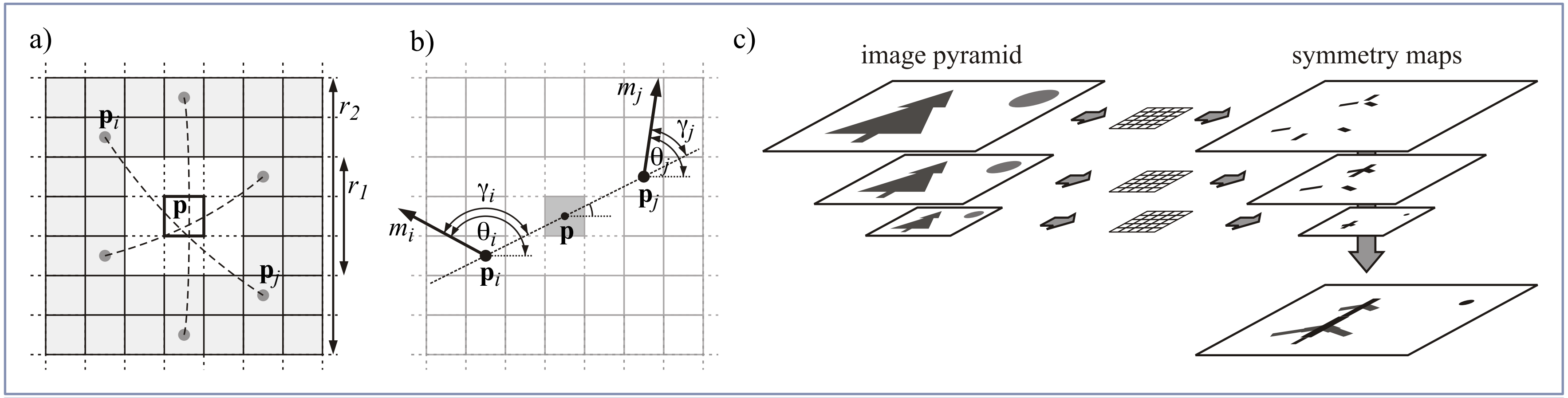
Prediction of gaze: basic features

- Predictions of human eye fixations based on contrasts of basic features



Multi-scale symmetry model

- Calculation of local symmetry at different scales to obtain a symmetry-saliency map



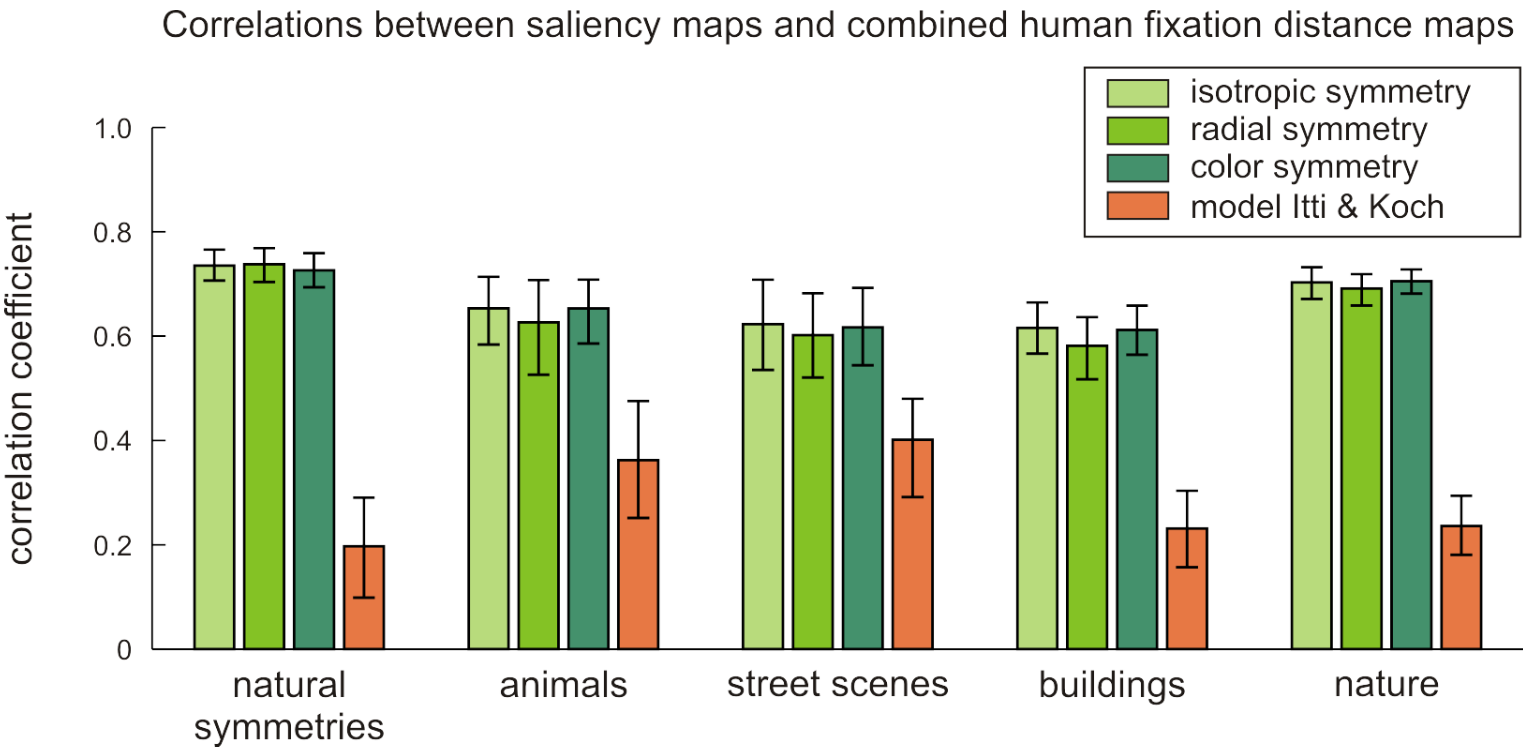
(Kootstra *et al*, Cogn. Comp. 2011)

Based on (Reisfeld *et al*, 1995)

Symmetry-saliency model: Results

Symmetry-saliency model: Results

- Human eye fixations are better predicted using symmetry than using center-surround contrast



(Kootstra *et al*, Cogn. Comp. 2011)

Automatic Object Detection and Segmentation

Gestalt in Machine Vision

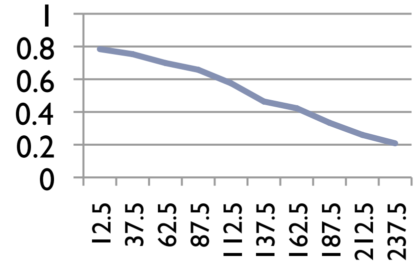
- Machine vision also faces the problem of how to go from parts to wholes
- Gestalt principles can be used as bottom-up features for
  - Attention
  - Segmentation

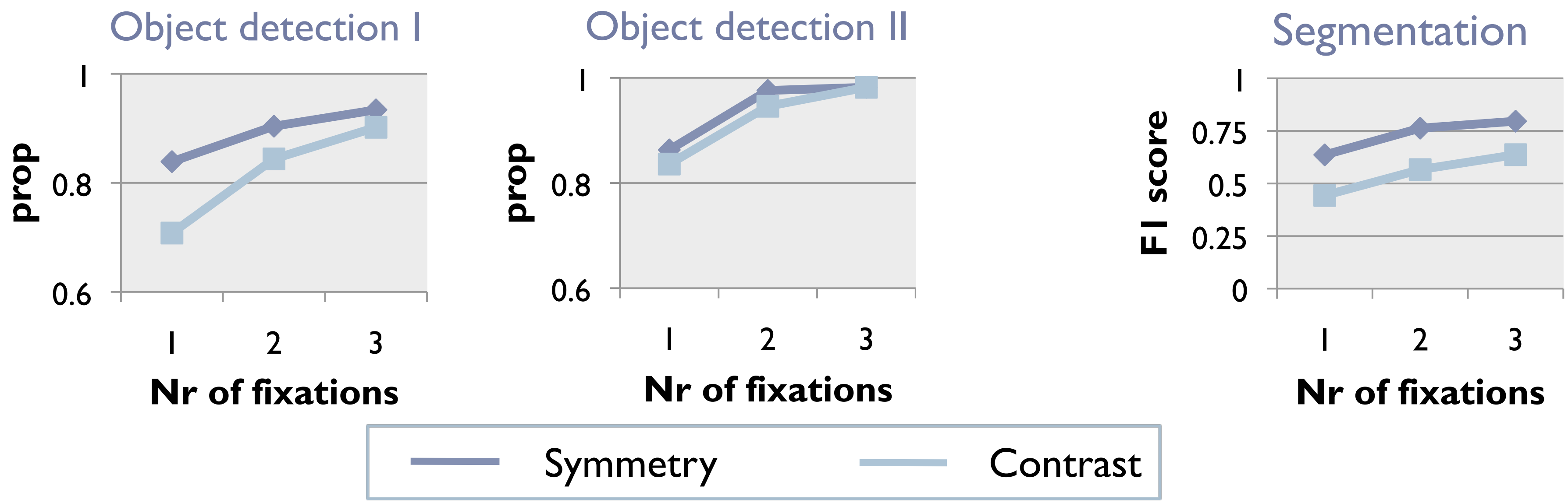
General Objectives

- Detection and segmentation of unknown objects
- No prior knowledge, so bottom-up methods
- Purpose
  - To be able to interact with the scene
  - To be able to learn new objects in the scene

Object Detection

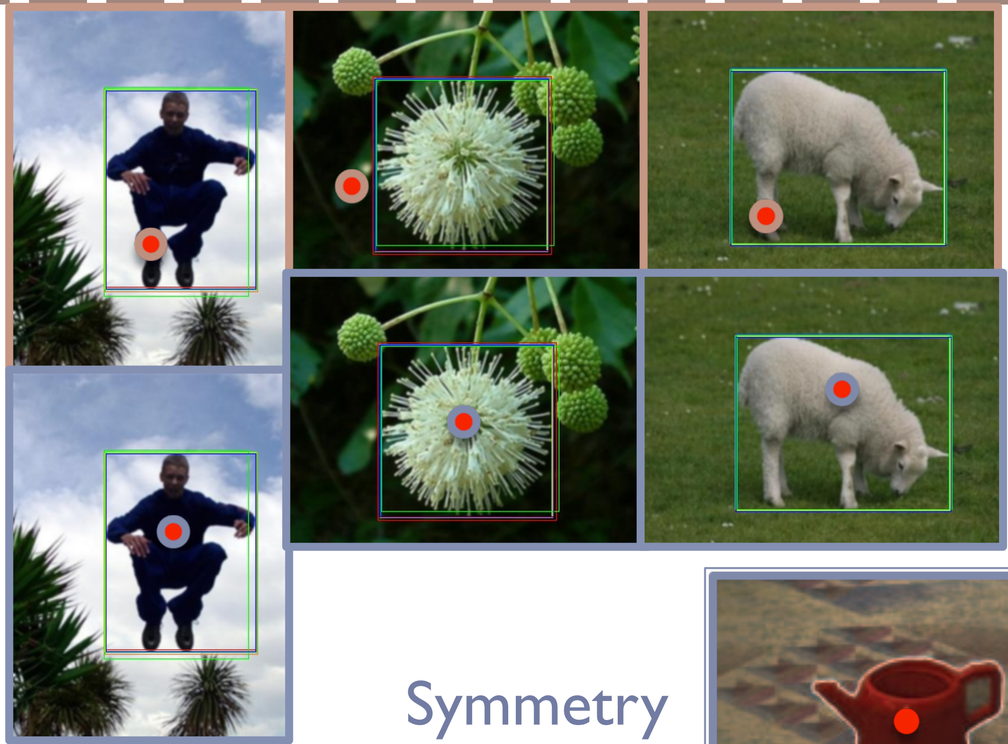
- The symmetry-saliency model selects fixation points that are:
  - Often on the salient objects in a scene
  - Close to the center of the object





(Kootstra *et al*, ICPR 2010)

Object Detection Results



Contrast

Symmetry

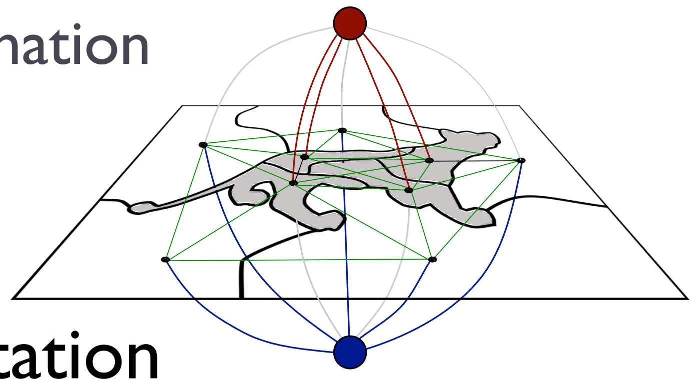
Symmetry

Contrast

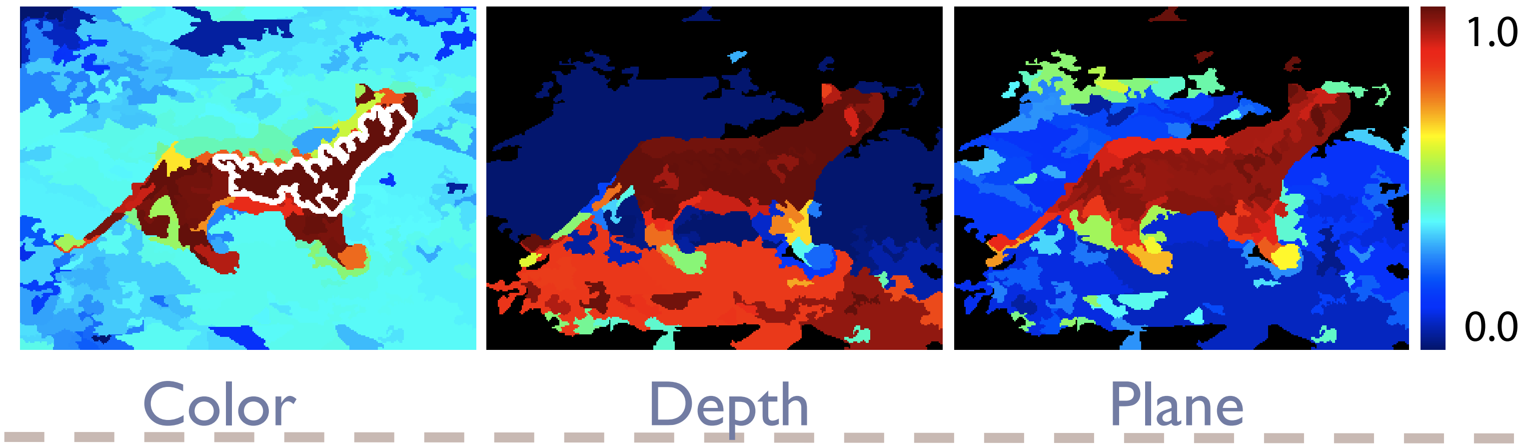
Object Segmentation

Super Pixels and Similarities

- We use a super-pixel representation
  - From 300.200 pixels to hundreds of super pixels
  - More reliable depth and plane information



- Markov Random Field
- Three similarity cues for segmentation

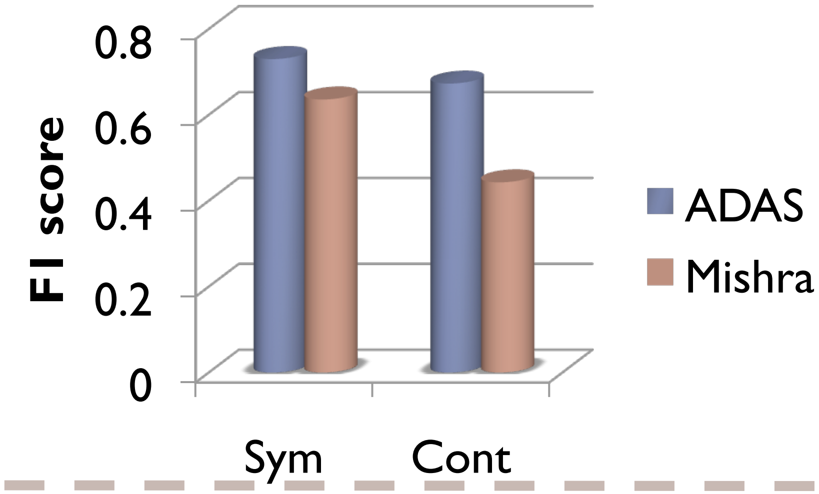


(Kootstra *et al*, Humanoids 2010)

Object Segmentation

- Iterative: from fixation point to segment

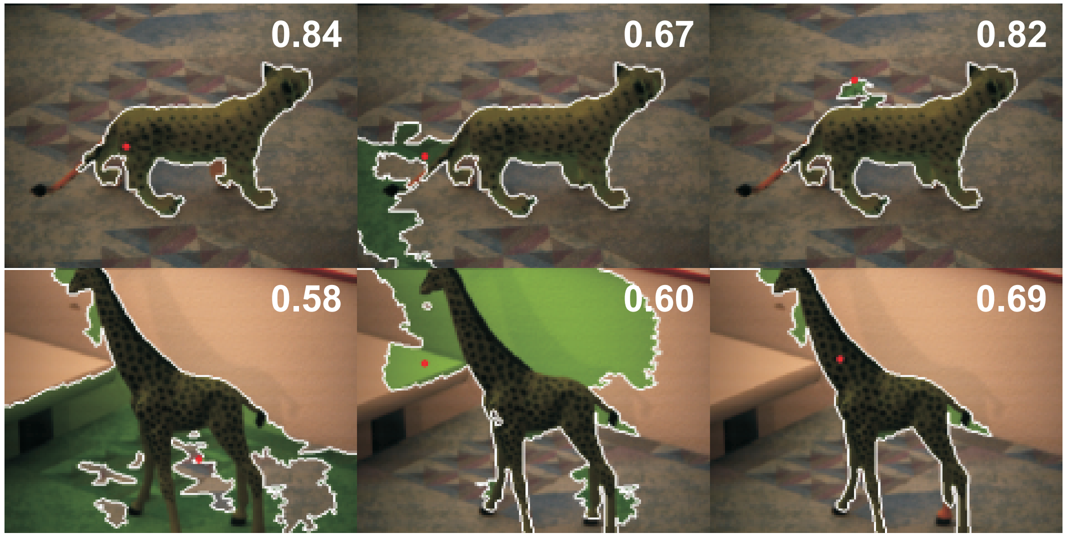
- Results, compared to (Mishra et al 2009)
  - Better segmentation
  - Much faster (50-100ms)



Segment Evaluation

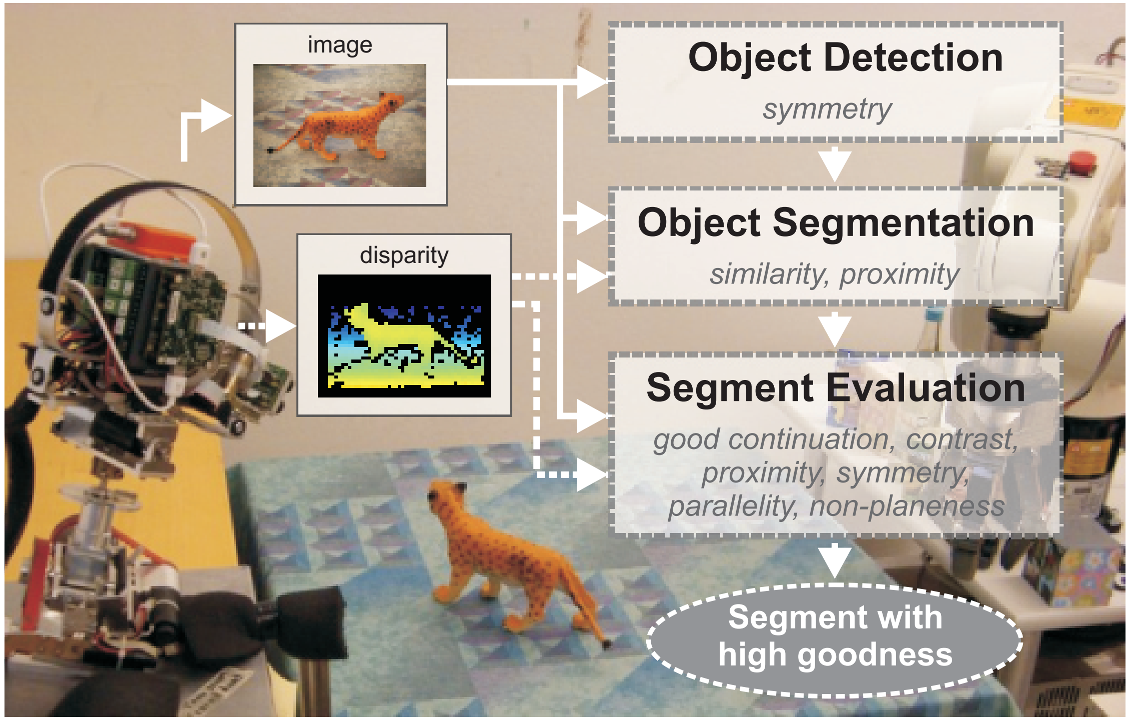
Segment Evaluation

- The bottom-up methods can fail
  - Failure of object-detection method to find object
  - Failure of segmentation method to correctly find the object borders.
- Goal of evaluation
  - Find best segment among candidates



(Kootstra *et al*, ICRA 2011)

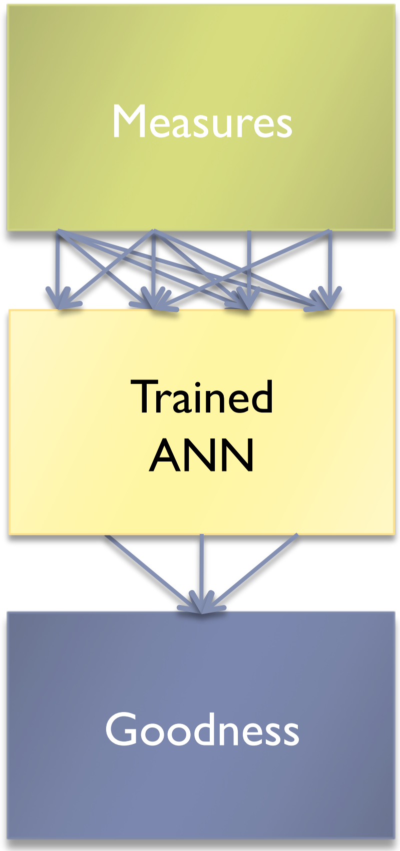
Segment Evaluation



- Use Gestalt principle of *figural goodness*
  - Nice, ordered, simple forms have high goodness

Gestalt measures for goodness

- Segment goodness
  1. Good continuation
  2. Color contrast
  3. Plane contrast
  4. Symmetry
  5. Parallelism
  6. Color uniqueness
  7. Out-of-planeness

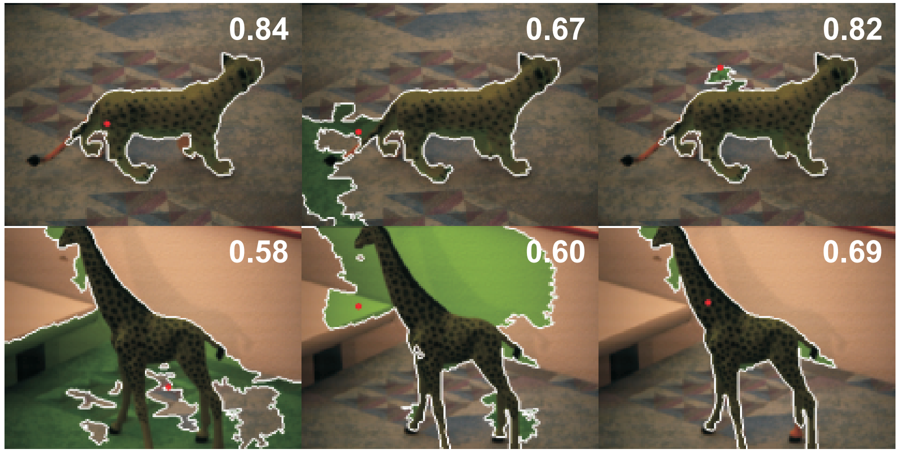


Results

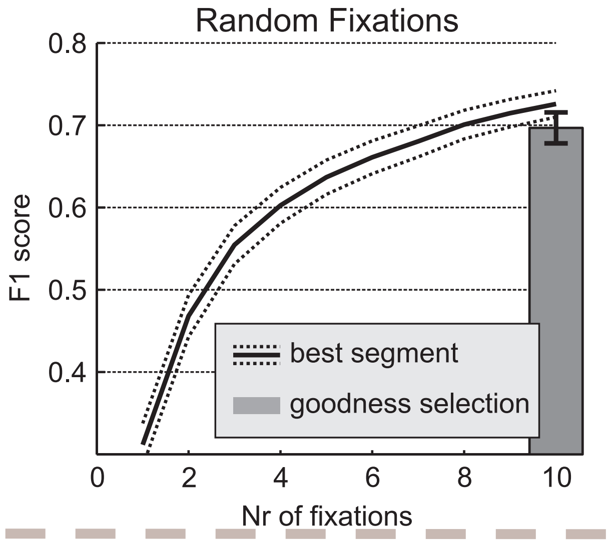
Measure	Correlation	R^2 measure
good continuation	0.56	0.31
color contrast	0.58	0.34
plane contrast	0.64	0.41
symmetry	0.63	0.39
parallelity	0.61	0.37
color uniqueness	0.71	0.51
out-of-planeness	0.77	0.59
Linear combination	–	*0.80*
Neural network	0.93	*0.87*

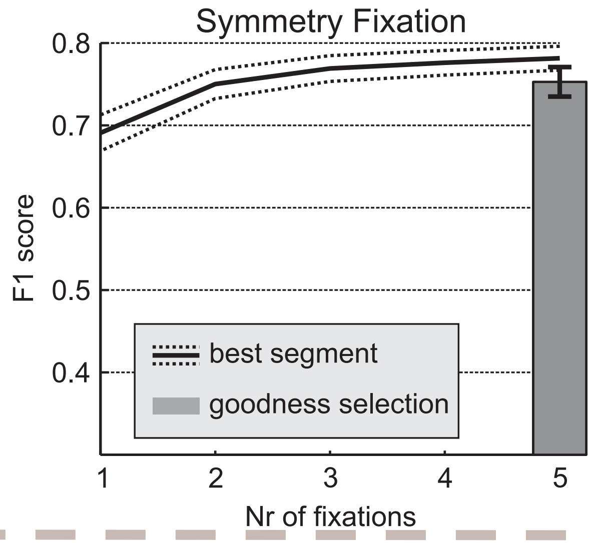
- R^2 measure of individual measures reasonable
- Linear combination: improvement
  - Shows that measures are complimentary
- Trained neural network outperforms linear comb.

Results


0.84
0.67
0.82
0.58
0.60
0.69


0.97
0.96
0.22
0.96
0.94
0.28
0.09
0.97
0.85


Random Fixations
F1 score
0.8
0.7
0.6
0.5
0.4
0
2
4
6
8
10
Nr of fixations
best segment
goodness selection


Symmetry Fixation
F1 score
0.8
0.7
0.6
0.5
0.4
1
2
3
4
5
Nr of fixations
best segment
goodness selection

Discussion

Discussion

- Gestalt principles for
  - Prediction of eye movements (symmetry)
  - Object detection (symmetry)
  - Object segmentation (proximity, similarity)
  - Segment evaluation (7 principles)
- Gestalt principles provide good features for bottom-up and autonomous object detection and segmentation

Speed

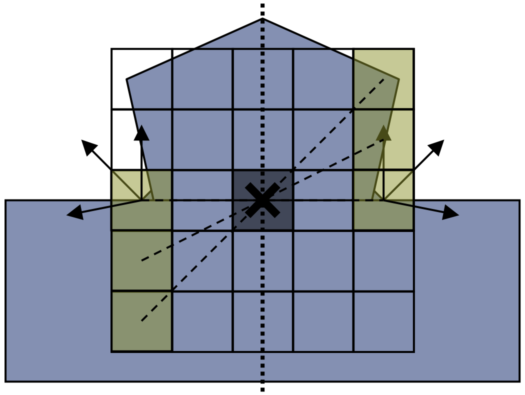
- Symmetry detection
  - CPU: 50 ms for 460x480 images
  - GPU: 5-10 ms
- Super-pixel segmentation
  - CPU:
    - Super pixels, Lab, stats: 100 ms
    - Graph-cut segmentation: 4-8 ms
- Evaluation
  - CPU: 5-10 ms

Thank you for you attention

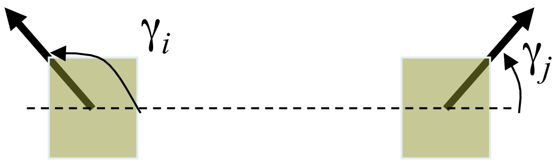
kootstra@kth.se

Symmetry-saliency model

- Symmetry as a salient feature



Do this for all pixel pairs in kernel

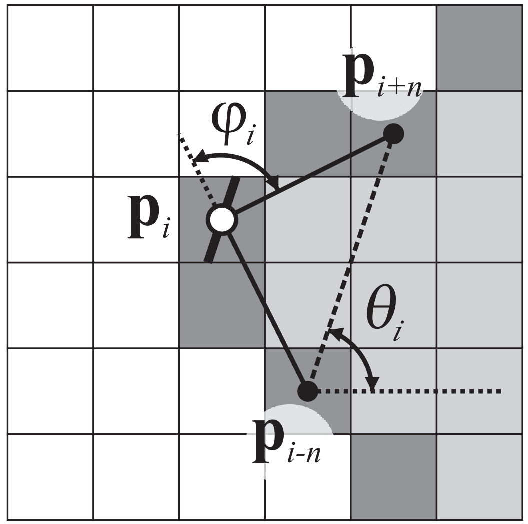


$$s_{ij} = \left(1 - \cos\left(\gamma_i + \gamma_j\right)\right) \cdot \left(1 - \cos\left(\gamma_i - \gamma_j\right)\right)$$

$$S(p) = \sum_{ij \in \Gamma} s_{ij} \cdot w_{ij}$$

Gestalt measures

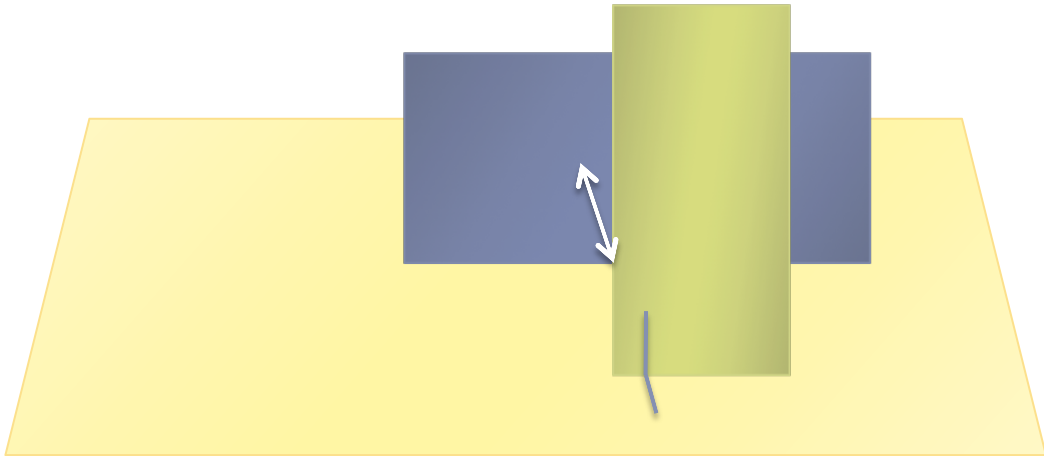
- Good continuation
  - True object boundaries are generally smooth
  - Based on the curvature φ_i



- Color contrast
  - Contrast of color at left and right side of contour usually high at true object boundaries
  - $c_i = \sqrt{0.3(r_i - r_b)^2 + 0.59(g_i - g_b)^2 + 0.11(b_i - b_b)^2}$

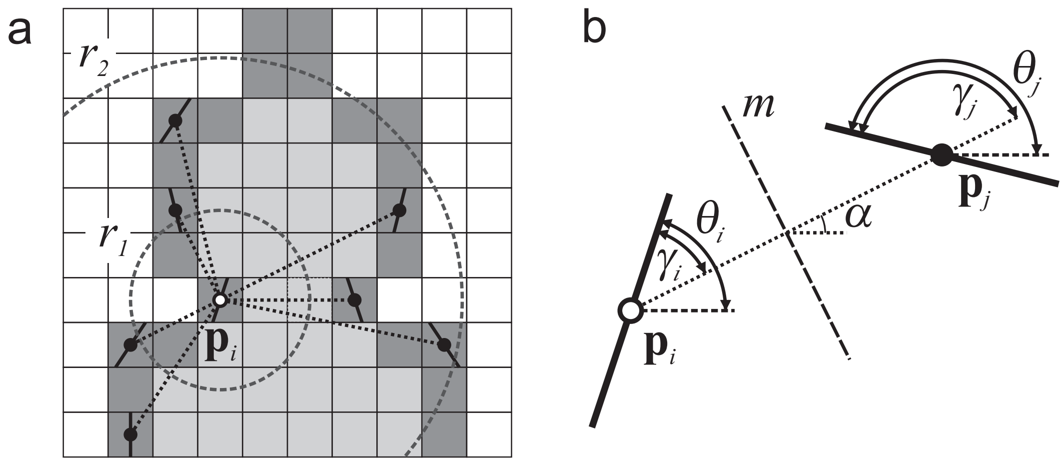
Gestalt measures

- Plane contrast
  - Left and right side of true object boundaries are usually in different plane
  - Contrast based on difference in depth and orientation of the super-pixel planes



Gestalt measures

- Symmetry
  - Total amount of symmetry of the segment



  - Comparing all contour elements ($r_1 < d < r_2$) with each other
  - $$\begin{aligned} \xi_{i,j} &= \cos^2(\gamma_i + \gamma_j) \cdot \left(\sin^2(\gamma_i) \cdot \sin^2(\gamma_j)\right) \\ \lambda_{i,j} &= \log(1 + c_i) \cdot \log(1 + c_j) \\ s_{i,j} &= \xi_{i,j} \cdot \lambda_{i,j} \end{aligned}$$

Gestalt measures

- Parallelism
  - Similar to symmetry, but now for parallel contour elements
- Color uniqueness
  - The object usually has a distinct color from the background
  - Comparing the L*ab* color histograms of the segment with the complete image

Gestalt measures

- Out-of-planeness
  - Assumption: objects are placed on a supporting surface, which can be detected as dominant plane
  - Comparison of planes of the super pixels in the segment and the dominant plane