Biological Vision and Applications
Module 05-04: Context-based attention models

Hiranmay Ghosh

Context-based Model

A comprehensive model for top-down + bottom-up attention

$P(O \mid v_l, v_c) = \frac{1}{P(v_l \mid v_c)}.(v_l \mid O, v_c).\underline{P(O \mid v_c)}$

Substituting $O = (o, x, \sigma)$ and expanding:

$P(O \mid v_c) = P(\sigma \mid x, o, v_c).P(x \mid o, v_c).P(o \mid v_c)$

Conditional independence: $P(v_l \mid O, v_c) = P(v_l \mid O)$

Substituting and rearranging:

$$P(O \mid v_l, v_c) = \underbrace{\frac{1}{P(v_l \mid v_c)}}_{1}. \underbrace{P(v_l \mid O).P(\sigma \mid x, o, v_c).}_{3} \underbrace{P(x \mid o, v_c).P(o \mid v_c)}_{2}$$

1. Bottom-up saliency (task-independent)

What catches spontaneous attention

- $1/P(v_l \mid v_c)$:
  - How unlikely is v_l given a context v_c
    - Independent of object hypothesis O
    - surprise factor
  - Represents bottom-up saliency
  - Bayesian model brings in experiential factor (temporal)

2. Top-down saliency (task specific)

Where to look for

- $P(x \mid o, v_c).P(o \mid v_c)$
  - $P(o \mid v_c)$: Prob of an object class to appear in a context
  - $P(x \mid o, v_c)$: Prob of an object class to appear at a certain location in a context
    - ... given that it appears
- $P(x \mid o, v_c).P(o \mid v_c) = P(o, x \mid v_c)$:
  - Represents the task-specific context-driven saliency of location

3. Appearance model (task specific)

What to look for

- $P(v_l \mid O).P(\sigma \mid x, o, v_c)$
  - $P(\sigma \mid x, o, v_c)$: appearance (scale, orientation) of the object
    - ... when it appears at a certain location
  - $P(v_l \mid O) = P(v_l \mid o, x, \sigma)$: The expected visual features
    - ... when the object appears at a certain location with a certain appearance

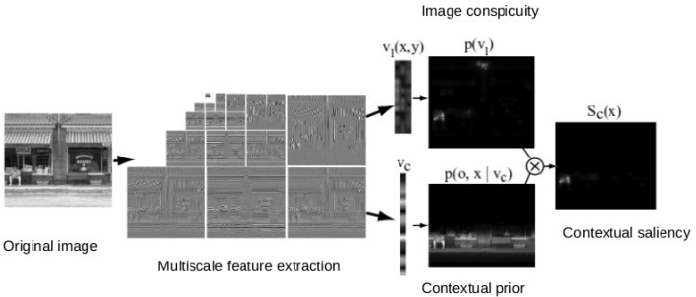
In Summary

$$P(O \mid v_l, v_c) = \underbrace{\frac{1}{P(v_l|v_c)}}_{1}. \underbrace{P(v_l \mid O).P(\sigma \mid x, o, v_c).}_{3} \underbrace{P(x \mid o, v_c).P(o \mid v_c)}_{2}$$

1. Bottom-up saliency (s_b): $\frac{1}{P(v_l|v_c)}$
   - What spontaneously draws attention
2. Top-down saliency (s_t): $P(o, x \mid v_c) = P(x \mid o, v_c).P(o \mid v_c)$
   - Where to look for for an object class
3. Visual features (v): $P(v_l \mid o, x, \sigma).P(\sigma \mid x, o, v_c)$
   - What to look for at a certain location to detect an object

- Overall saliency (where): $s_c = s_b \times s_t$
- Feature to look for (what): v

Local Computations



Global Computations

Quiz

Quiz 05-04

End of Module 05-04