IIT Jodhpur

Biological Vision and Applications
Module 05-04: Probabilistic models

Hiranmay Ghosh

Information theoretic model

- Challenges à-prior defined features for attention in Cognitive models
- Based on Shannon's information theory:
  - The event that is least likely to occur has the maximum information value
  - Self-information of an event x: $-\log P(x)$
- Image region that is least likely to occur in an image is the most salient one
- How to decide what is least likely to occur?

Shannon's Information theory

A generative model of an image

Independent Component Analysis

- An image region (signal) is a manifestation of some underlying hidden processes
- It is a weighted sum of contributions from each process
  - The processes are independent of each other
  - The processes are hidden
  - The weights are unknown
- We resort to Natural Scene Statistics
  - Observe a large number of images
  - Learn the processes and the weights
- Given a new image, we would expect each region to be a weighted sum of contributions from some of these features

What is unexpected ?

- Model a new image with the learned features
  - Select features for best fit over all image regions
- There will be some outlier regions, which do not fit
  - Has least probability to occur
  - Has most information value
- These image regions are the salient ones
  - Lower is the probability, the higher is the saliency
- Bruce. Features that draw visual attention: ...
- Bruce & Tsotsos. Saliency Based on Information Maximization (2006)

Bayesian model

Based on "surprise" – brings in experential factor

- Image regions (observed data) are caused by some hidden states (features)
- Bayesian definition of surprise
  - M: a continuous range of states (features that manifests as image regions)
  - $p(m)$: the prior pdf for states (characterizes alternate models)
  - $p(m \mid D)$: the posterior pdf for the states, after experiencing some data D
- The surprise factor of the data D is defined as
  - The difference between the posterior and the prior pdfs for M
    - $S(D) = KLD(p(m), p(m \mid D)) = \int_m p(m).\log \frac{p(m)}{p(m|D)}.dm$

Surprise
... contd.

- The surprise factor of the data D: $S(D) = \int_m p(m).\log\frac{p(m)}{p(m|D)}.dm$
- Baye's Theorem: $p(m \mid D) = \frac{p(m)P(D|m)}{P(D)}$
- Using Baye's Theorem and simplifying $[\int_m p(m) = 1]$
  - $S(D) = \log P(D) - \int_m \log P(D \mid m).dm$
- The model of the environment builds incrementally as the agent observes
  - For repeated observations, posterior of one observation is the prior for the next
  - Leads to change awareness

Itti & Baldi. Bayesian Surprise Attracts Human Attention (2005)

Outlier and surprise

Criticism to information theoretic model

- Outlier D: $P(D \mid M_{\text{best}})$ to have low value
- Suppose $P(D \mid M')$ is low for all possible alternate models M'
- D has little informative value for discriminating the models for best explanation of the observation
- Cannot be treated as "surprise" and cannot be considered to be a guiding factor for attention

Eye movement ?

- Fixations and saccades are guided by WTA and RI policies as in Cognitive models

Quiz

Quiz 05-04

End of Module 05-04