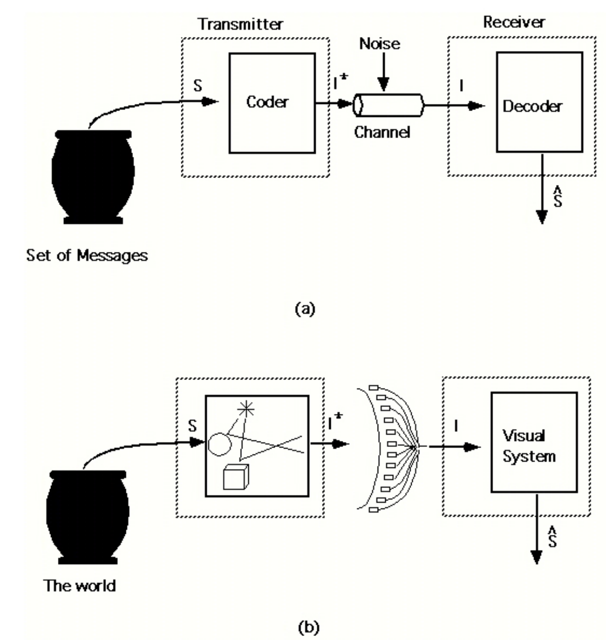
Modeling Vision as Bayesian Inference: Is it Worth the Effort?

Alan L. Yuille. UCLA.
Co-Director: Centre for Image and Vision Sciences.
Dept. Statistics
Joint Appointments:
Computer Science, Psychology

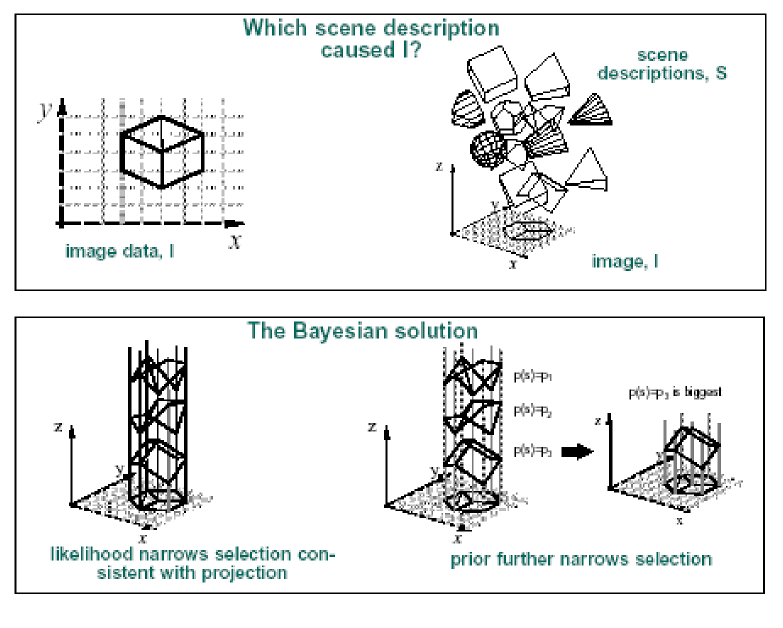
What is the purpose of Vision?

- To extract information about the world from the input images.
- This is a decoding problem: inverse inference.



What is Bayes? Inverse Inference.

- Inverse Inference/Image decoding:
- Posterior P(S|I) = P(I|S) P(S)/P(I)
- Likelihood P(I|S), Prior P(S).



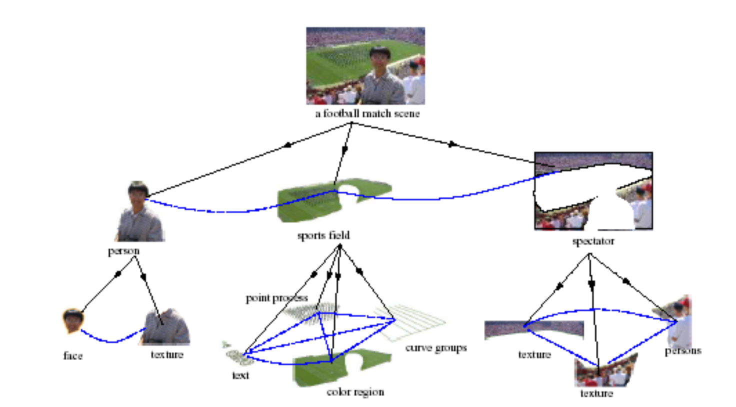
Cf. Pavan Sinha

How does this relate to the Brain?

- Marr’s computation level. Information Processing.
- Bayes derives from Decision Theory – same roots as signal detection theory and ideal observers.
- Plausible neuronal implementations.

The Challenge of Bayes

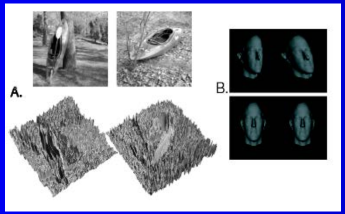
- What are P(I|S) and P(S) for realistic images?

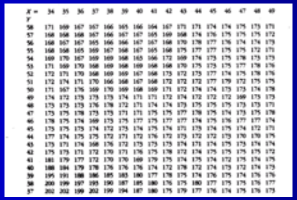


- Probability Distributions on Structured Representations (Graphs/Grammars).
- Considerable progress – in many communities; NIPS,Machine Vision, Machine Learning,Artificial Intelligence. Natural Language processing…

Why is Bayes Complicated?

- It is complicated because of the difficulty of the vision problem.
- In information processing terms – images are extremely complex and ambiguous.





- Vision is much more complicated – by many orders of magnitude – than any “solved” inverse inference problem.

Bayes Research Program

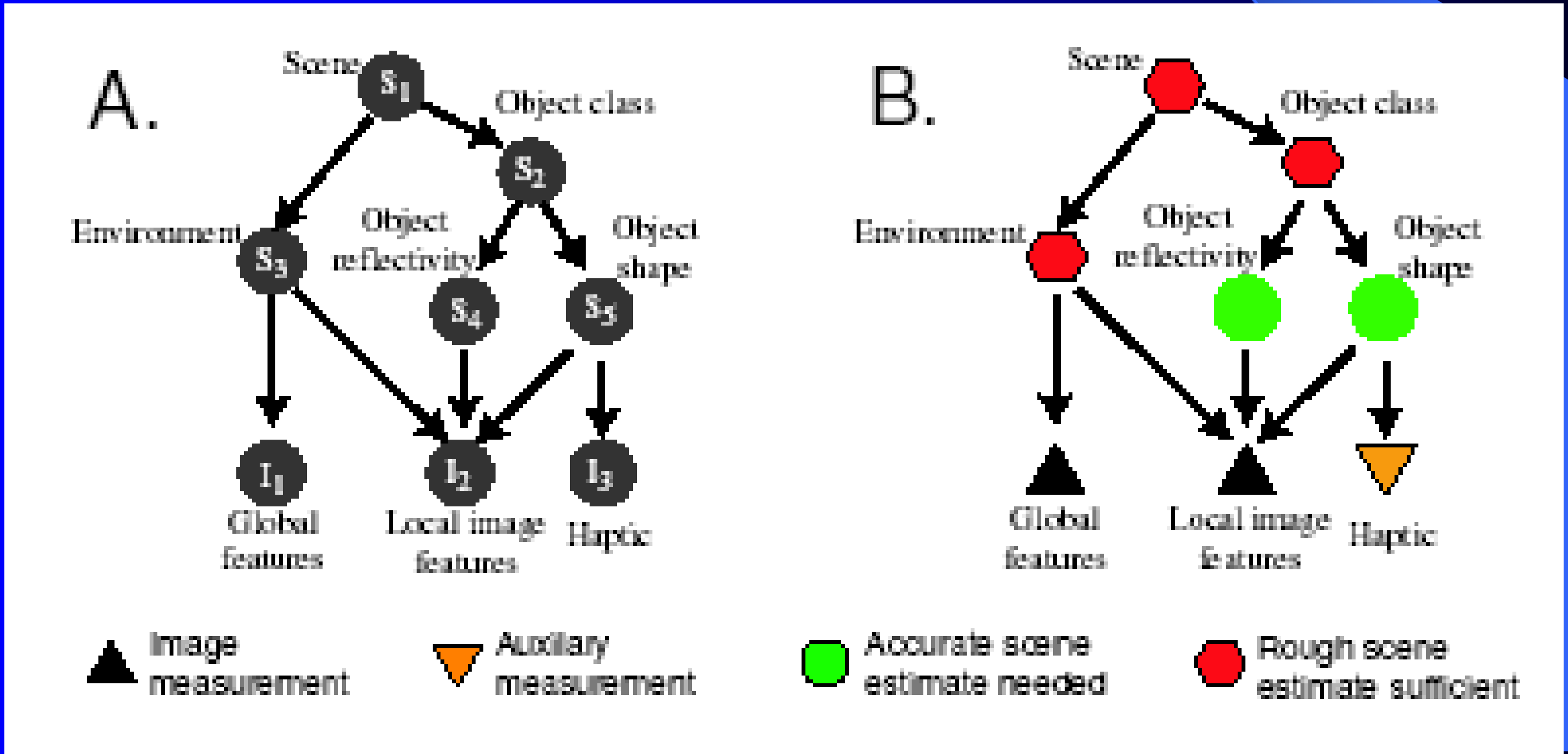
- This program is to model the world with increasingly realistic distributions $P(I|S)$ and $P(S)$.
- Relate performance on these models to Psychophysics experiments.
- Speculate (test?) neuronally plausible implementations of these models.

Brief History: 1980's.

- Convergent strands:
- Signal Detection Theory, Ideal Observers.
- Pattern Theory and Bayesian Inference.
- Energy Function formulations.
- Computer Graphics Psychophysics.
- Axis of Bayes: Brown, Harvard, MIT?

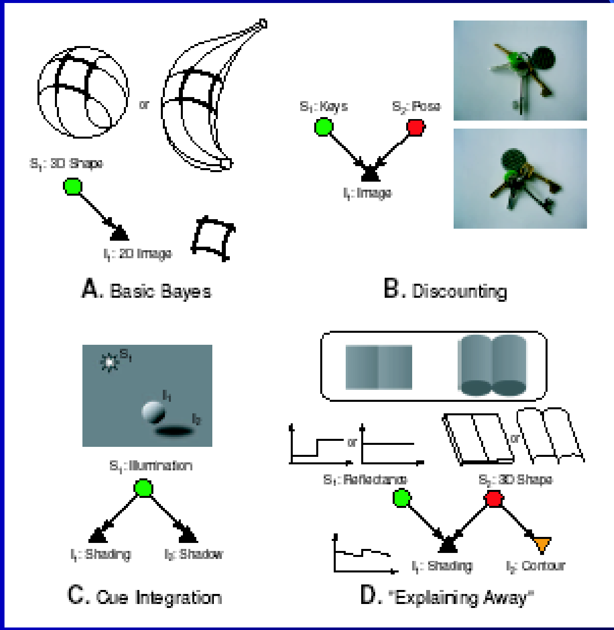
Probabilities defined on Graphs

- Graphs represent the causal generation of images. (Pearl 1988).



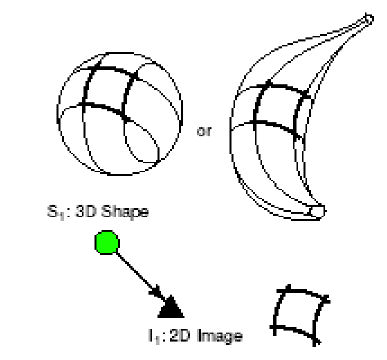
Taxonomy of Vision Tasks.

- Symbolic taxonomy.
- Kersten, Mamassian, Yuille (2004). 150 references…



Basic Bayes

- Bulthoff and Mallot (1988).
- The perception of shape and depth from visual cues.

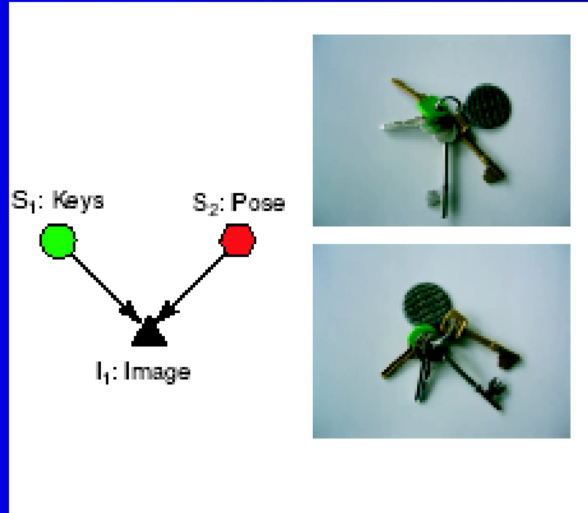


A. Basic Bayes

- Trade-off between data-driven and prior-driven (e.g. Weiss et al 1997).

Discounting and Task Dependence

- What do you care about? The identity of the keys? Their location? Their materials?



B. Discounting

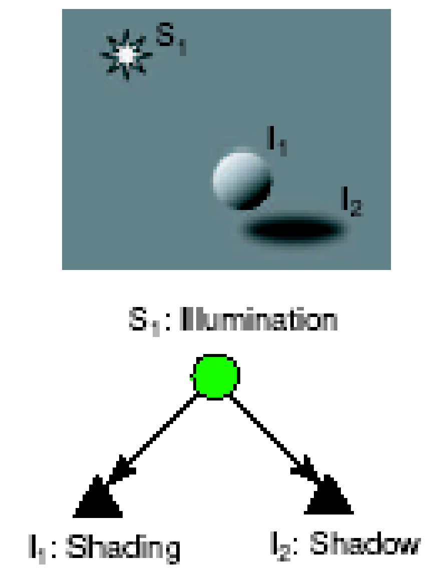
Cue Integration

- How to combine cues? Clark and Yuille 1990 (Amazon says "1 copy still available!").

  Bulthoff and Yuille 1991.
- Weak coupling

  or strong coupling.

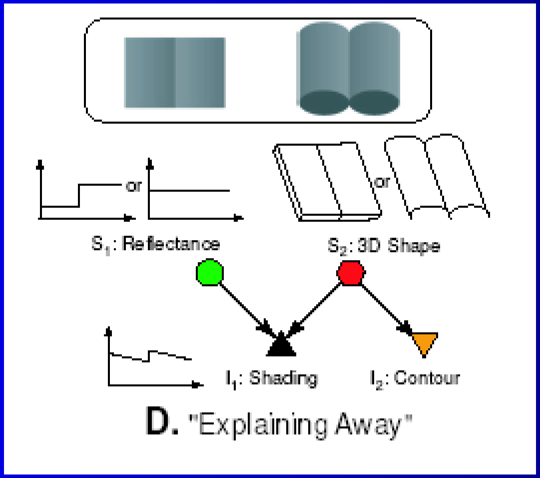
- Knill's talk.



C. Cue Integration

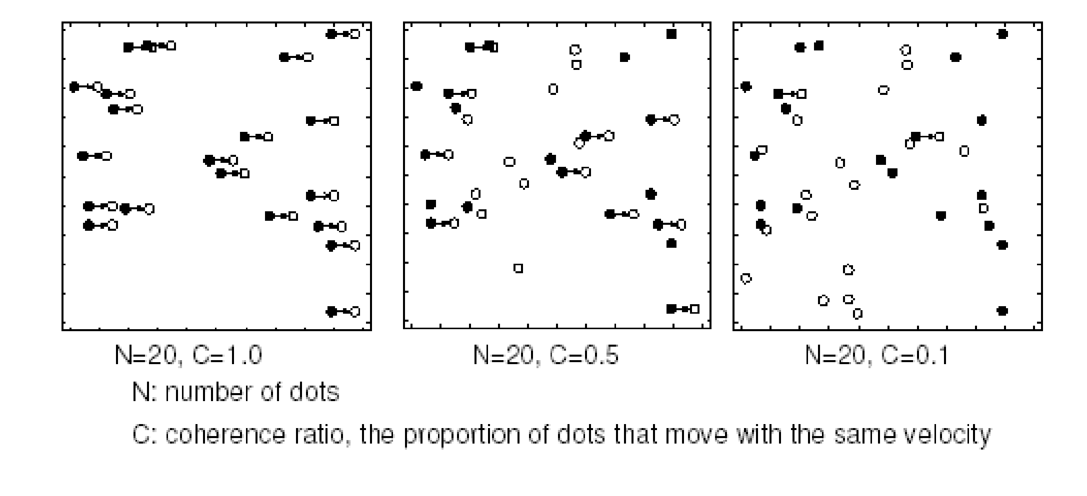
Perceptual “Explaining away”

- Knill and Kersten 1991.
- Small changes of the boundary shape can explain away the intensity changes as being due to geometry and not to materials.



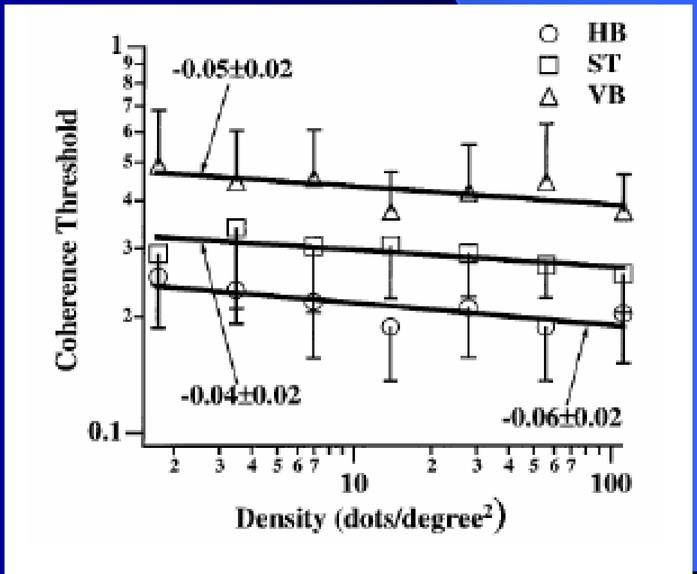
Ideal versus Bayes Ideal

- Ideal for the experiment? Or Ideal for real images?
- Random Dot Kinematograms (RDK's).



Ideal for the experiment?

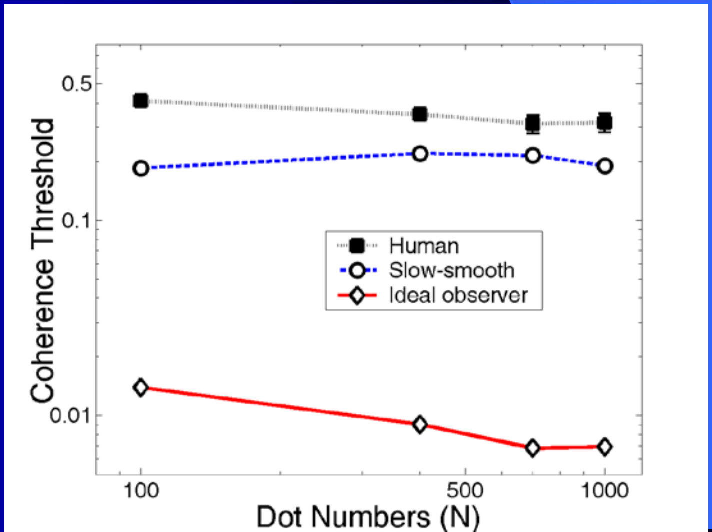
- Barlow&Tripathy (1997), Lu&Yuille (2006) derived an Ideal Observer model for these stimuli.
- This model predicts trends, but humans are much worse than the ideal.



Ideal for the Real World?

- But human performance is well matched to a Bayesian model that uses a slow&smooth prior (Lu&Yuille 2006).
- Slow&Smooth (Yuille&Grzywaxz 1987, Weiss et al 19970.

Slow&smooth – consistent with knowledge of motion sequence statistics.

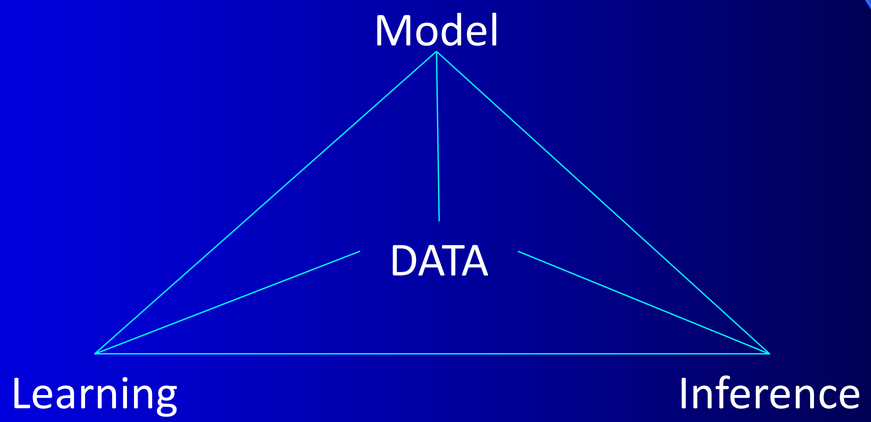


Neural Hypotheses

- The graphical nature of the models makes it straightforward to propose neuronally plausible implementations – nodes of graph as neurons. Koch, Marroquin, Yuille (1987), Lee (1995).
- Population models. Heroic efforts to test these models by Lee and his collaborators.
- Analysis by Synthesis: Feedforward and feedback connections (Mumford 1991). Some experimental support from fMRI (Murray et al.).

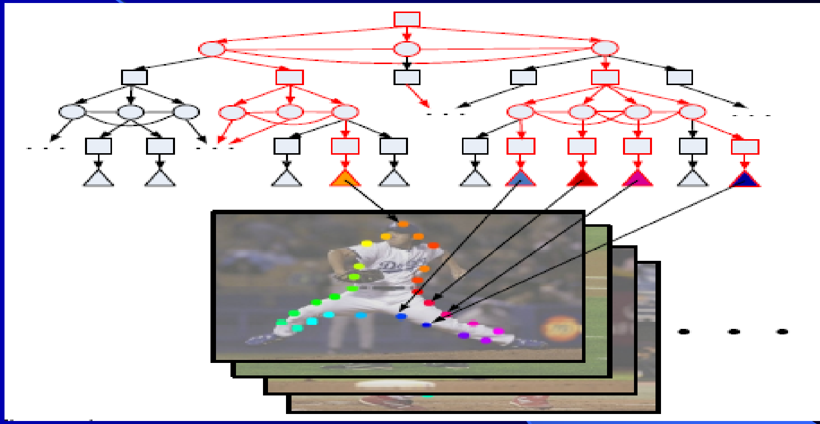
Bayesian Research Program.

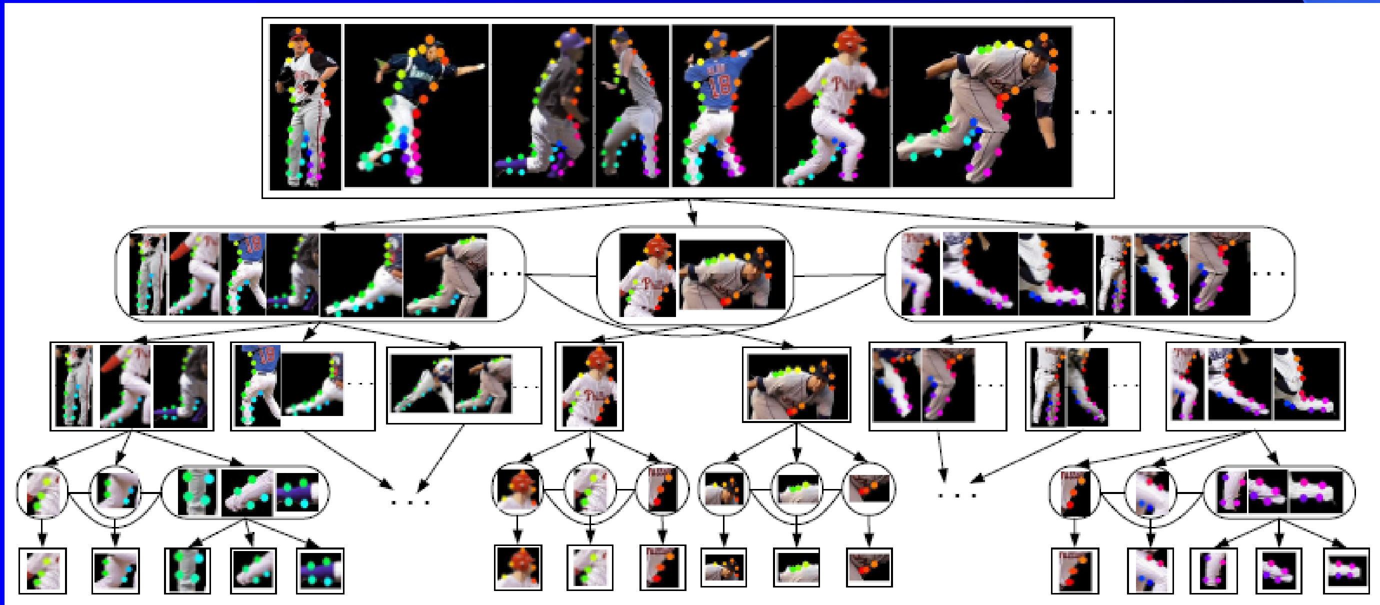
- Increasing sophisticated models, inference algorithms, learning.



And/Or Probabilistic Grammar:

P(I|S) & P(S) for human Poses.



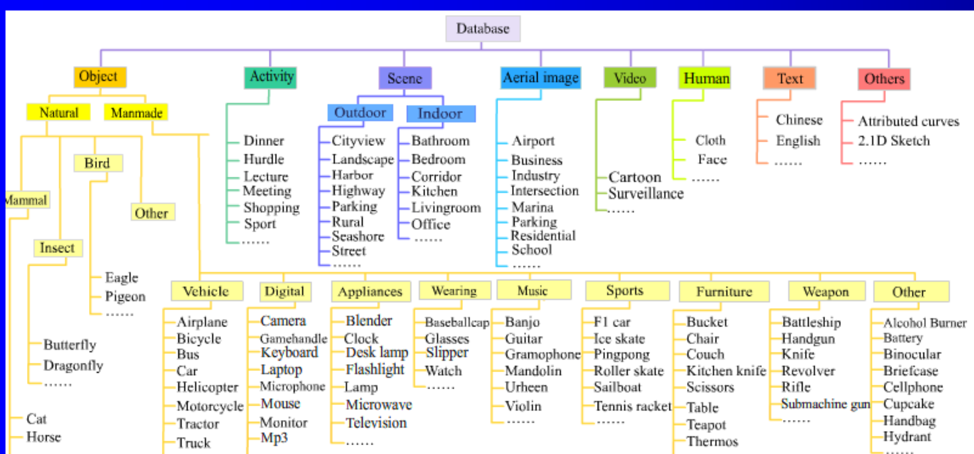


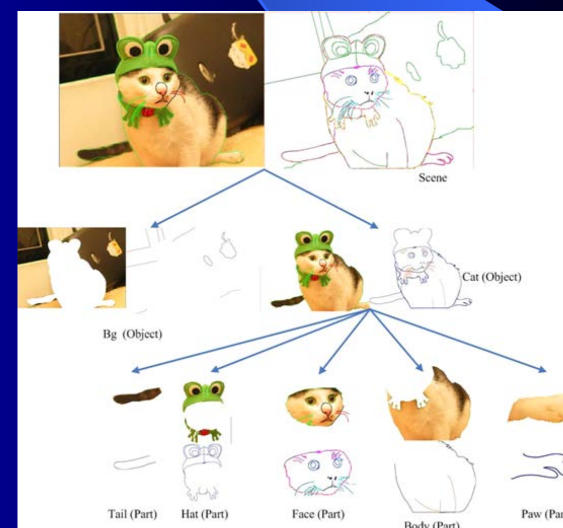
Bayes Research Program

- The Lotus Hill "Genome Project". Hand parse millions of images.
- Benchmarks for evaluating algorithms.
- Facility for learning generative models of images.
- Directed by Prof. S-C. Zhu.

Lotus Hill

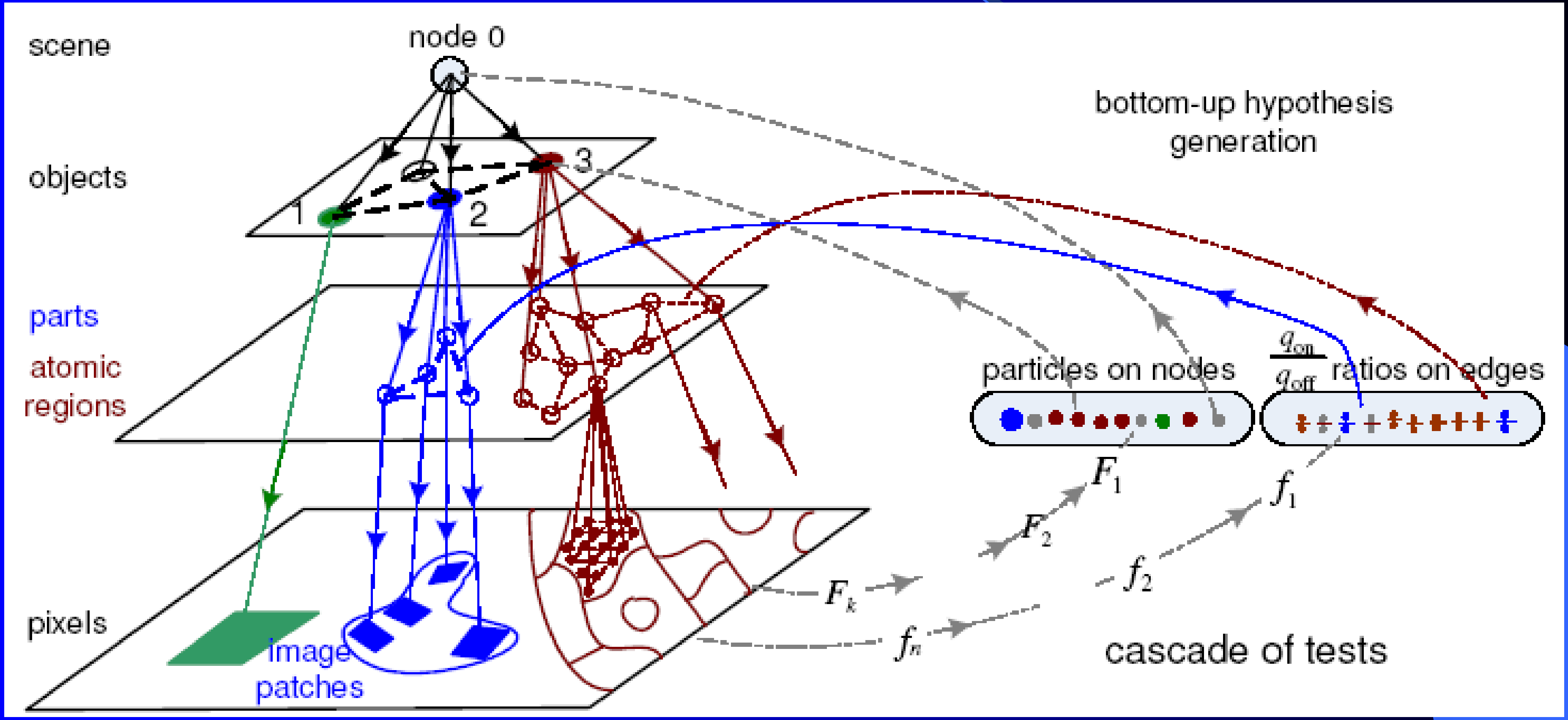
- Hand parsed images (dataset released).





Inference Algorithms: Bottom-Up/Top-Down

Integrating generative and discriminative methods

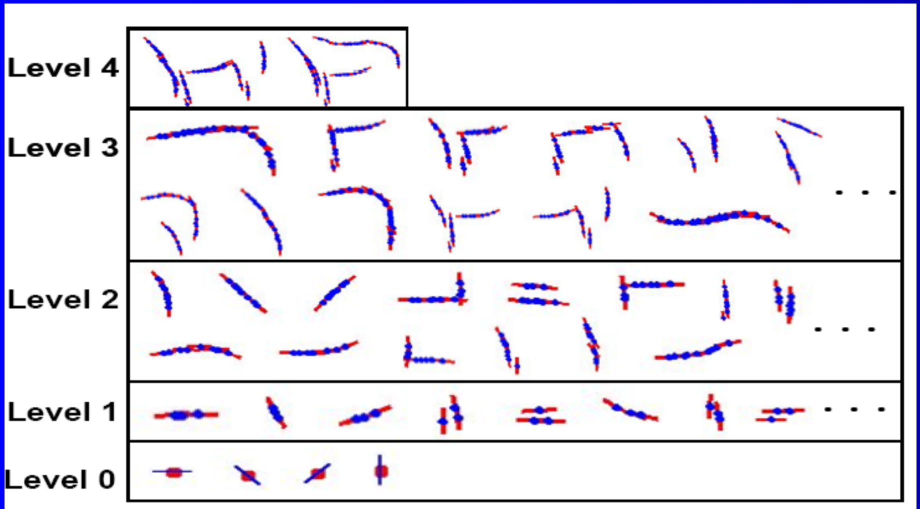


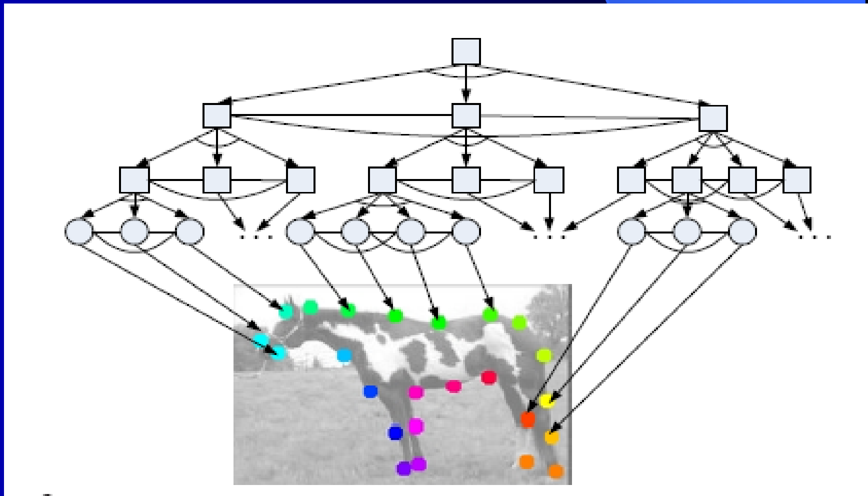
Unsupervised Learning

- Learn a model of a horse by hierarchical composition. (Leo Zhu & Yuille).
- Strategy: (i) suspicious coincidences,
- (ii) competitive exclusion,
- (iii) build hierarchy by composition.

Unsupervised Learning.

- Learn hierarchical schema by composition.
- Input 10 images, unknown pose/osition.
- Output; hierarchical model of a horse.
- Can detect and parse horses >300 real images.





Vision and Cognition

- The same probabilistic modeling techniques are being successfully applied to model other aspects of cognition.
- Vision, language, reasoning, motor control, and so on.
- This offers a theoretical framework for all of cognition.

Is Bayes complicated enough?

- **"To myself I am only a child playing with pebbles on the beach, while vast oceans of truth lie undiscovered before me."** – Isaac Newton.
- You can't explore the ocean without the right techniques and tools.

Conclusion

- “Bayes” – probabilistic models on structured representations – is extremely promising as a model of vision/cognition.
- There is good progress at scaling Bayes up to deal with the complexities of realistic images and visual tasks.
- There is encouraging progress at using Bayes to model psychophysics.

Facilities for Learning

- How to learn Bayes?
- Tutorial Programs – e.g. the UCLA IPAM website. Videos and Pdf's of lectures by world experts.