Structure of Vision Problems

Alan Yuille (UCLA).

Machine Learning

- Theory of Machine Learning is beautiful and deep.
- But, how useful is it for vision?
- Vision rarely has an obvious vector space structure.

Image Formation

- Images formation is complicated.
- E.g. the image of a face depends on viewpoint, lighting, facial expression.

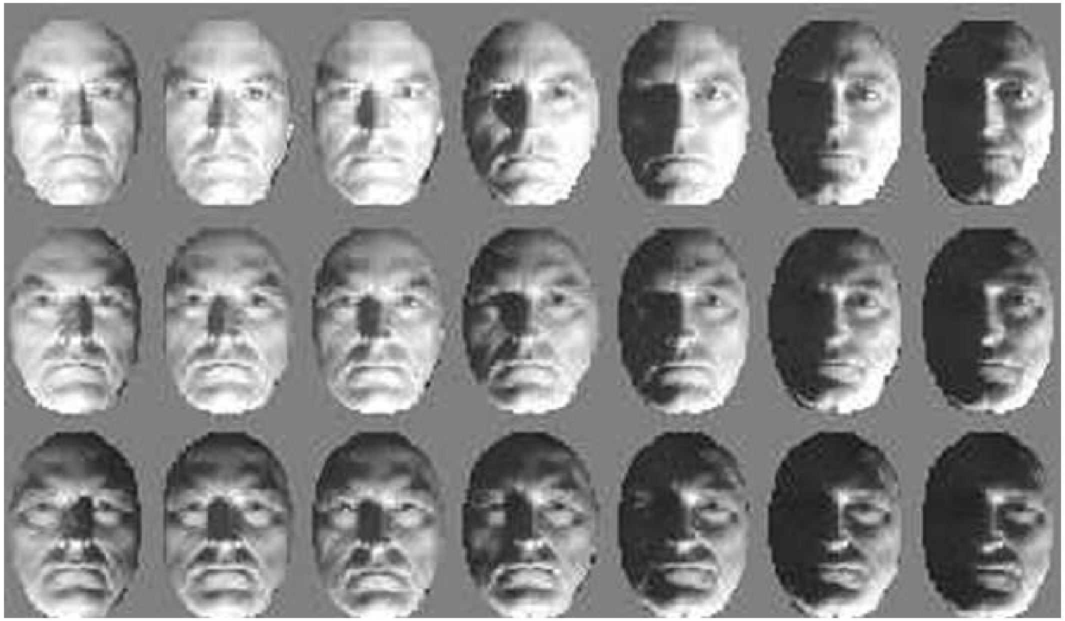


Image Formation.

- Parable of the Theatre, the Carpenter, the Painter, and the Lightman. (Adelson and Pentland).
- How many ways can you construct a scene so that the image looks the same when seen from the Royal Box?

Nonlinear Transformations

- Mumford suggested that images involve basic nonlinear transformations.
- (I) Image warping: $x \to W(x)$ (e.g. change of viewpoint, expression, etc.).
- (II) Occlusion: foreground objects occlude background objects.
- (III) Shadows, Multi-Reflectance.

Complexity of Images

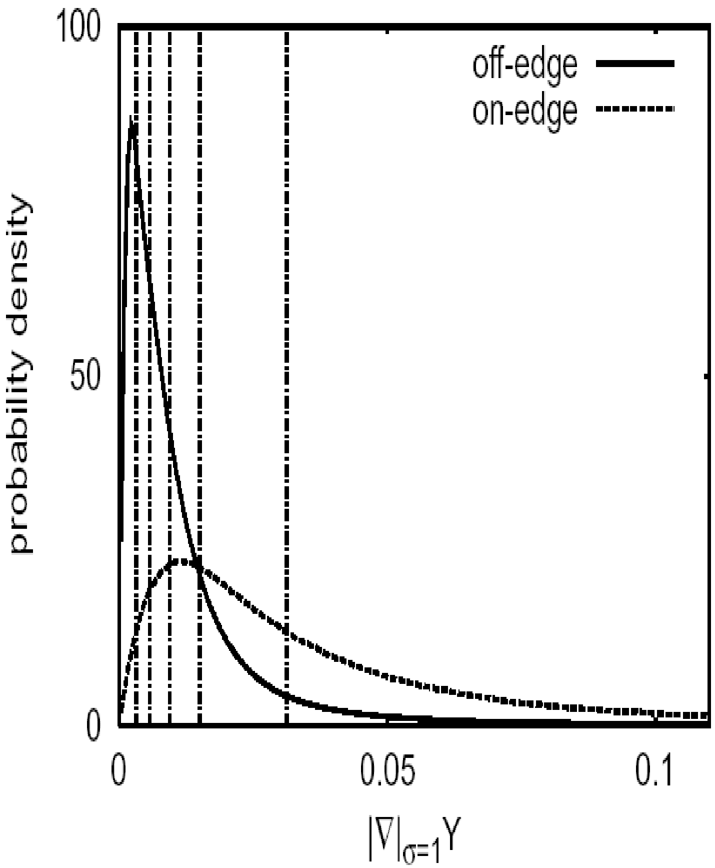
- Easy, Medium, and Hard Images.


Discrimination or Probabilities

- Statistical Edge Detection (Konishi,Yuille, Coughlan, Zhu).
- Use segmented image database to learn probability distributions of P(f|on) and P(f|off), where "f" is filter response.

P-on and P-off

- Let f(I(x)) = |grad I(x)|
- Calculate empirical histograms P(f=y|ON) and P(f=y|OFF).
- P(f=y|ON)/P(f=yOFF) is monotonic in y.
- So loglikelihood test is threshold on |grad (I(x)|.



P-on and P-off

- P-on and P-off become more powerful when combining multiple edge cues (by joint distributions).
- Results as good, or better than, standard edge detectors when evaluated on images with groundtruth.

P-on and P-off

- Why not do discrimination and avoid learning the distributions? (Malik et al).
- Learning the distributions and using log-likelihood is optimal provided there is sufficient data.
- But "Don't solve a harder problem than you have to".

Probabilities or Discrimination

- Two Reasons for Probabilities:
- (I) They can be used for other problems such as detecting contours by combining local edge cues.
- (II) They can be used to synthesize edges as a "reality check".

Combining Local Edge Cues

- Detect contours by edge cues with shape priors P_g (Geman & Jedynak).

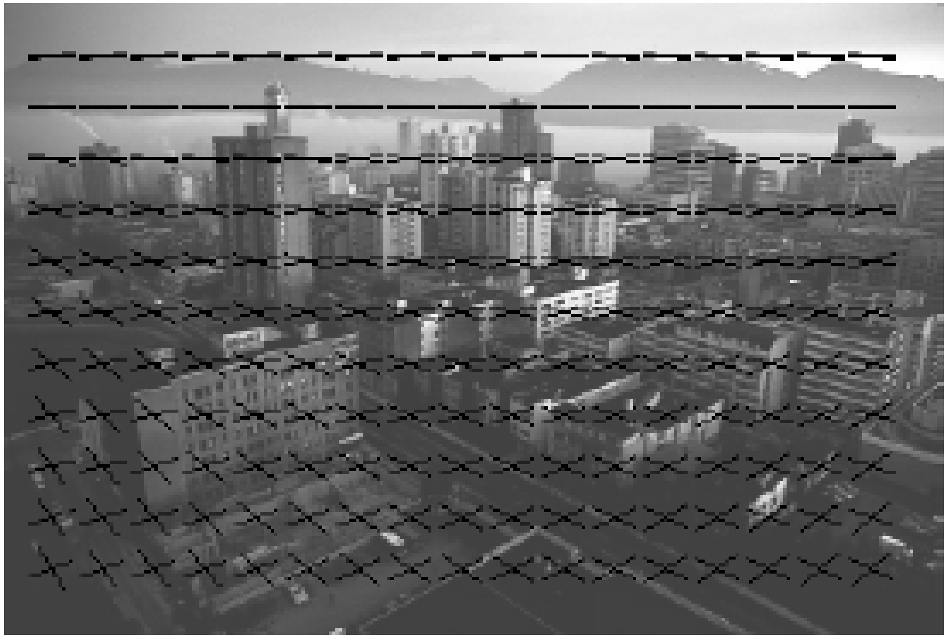
$$r(\{t_i\},\{y_i\}) = \frac{1}{N}\sum_{i=1}^{N}\log\frac{P_{on}(y_i)}{P_{off}(y_i)}$$

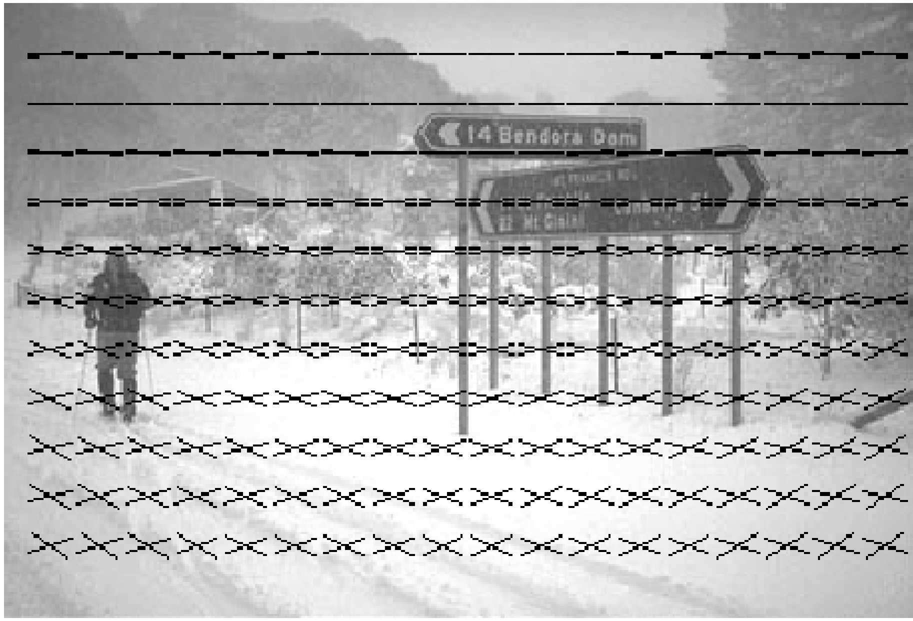
$$+\frac{1}{N}\sum_{i=1}^{N}\log\frac{P_g(t_i)}{U(t_i)},$$

$U(.)$ is uniform distribution.

Manhattan World

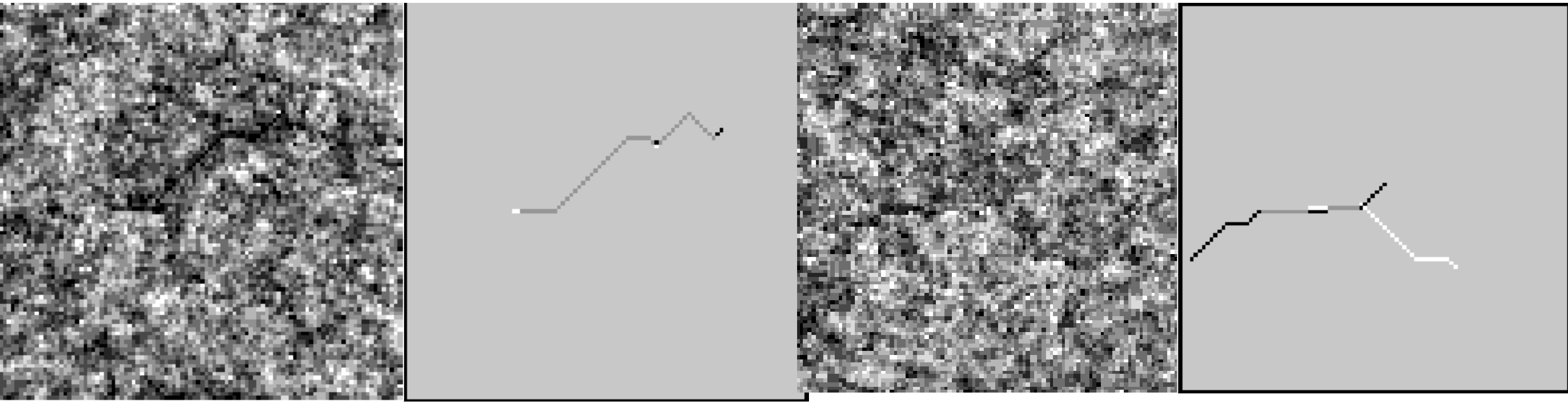
- Coughlan and Yuille use P-on, P-off to estimate scene orientation wrt viewer.





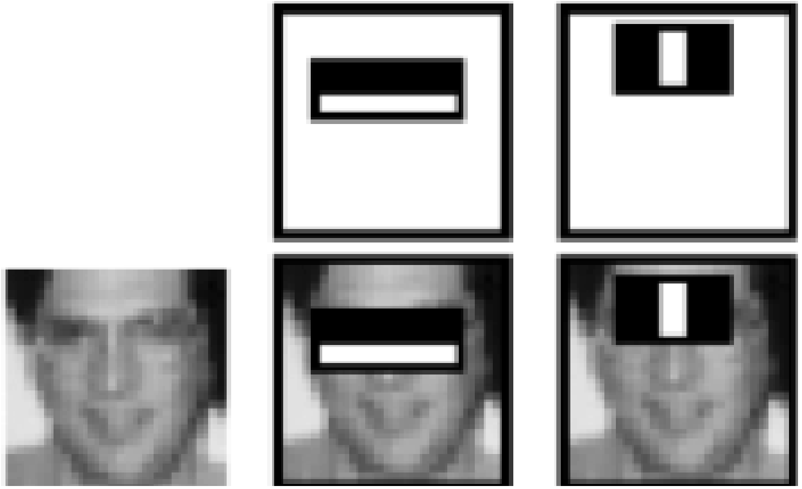
Synthesis as Reality Check

- Synthesis of Images using P-on, P-off distributions (Coughlan & Yuille).

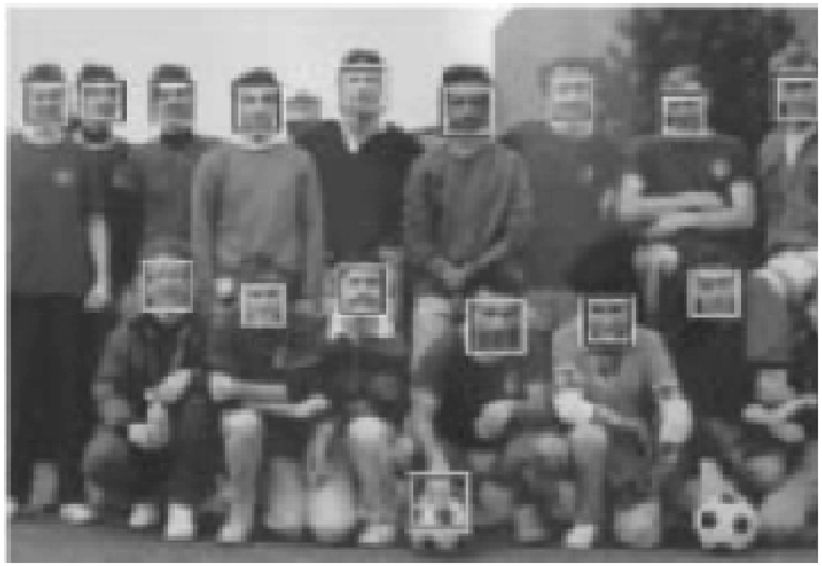


Machine Learning Success

- Fixed geometry, lighting, viewpoint.
- AdaBoost Learning: Viola and Jones.



a. the first two face features



b. an example of face detection

Machine Vision Success

- Other examples:
- Classification (Le Cun et al, Scholkopf et al, Caputo et al).
- Demonstrate the power of statistics – rather than the power of machine learning?

Bayesian Pattern Theory.

- This approach seeks to model the different types of image patterns.
- Vision as statistical inference – inverse computer graphics.
- Analysis by Synthesis (Bayes).
- Computationally expensive?

Example: Image Segmentation

- Standard computer vision task.
- Pattern Theory formulation (Zhu,Tu):

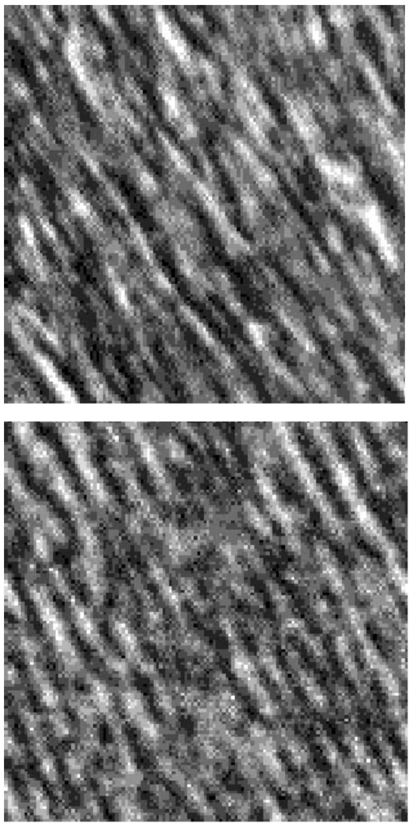
Decompose images into their underlying patterns.

- Requires a set of probability models which can describe image patterns.

Learnt from data.

Image Pattern Models

- Images (top) and Synthesized (bottom).



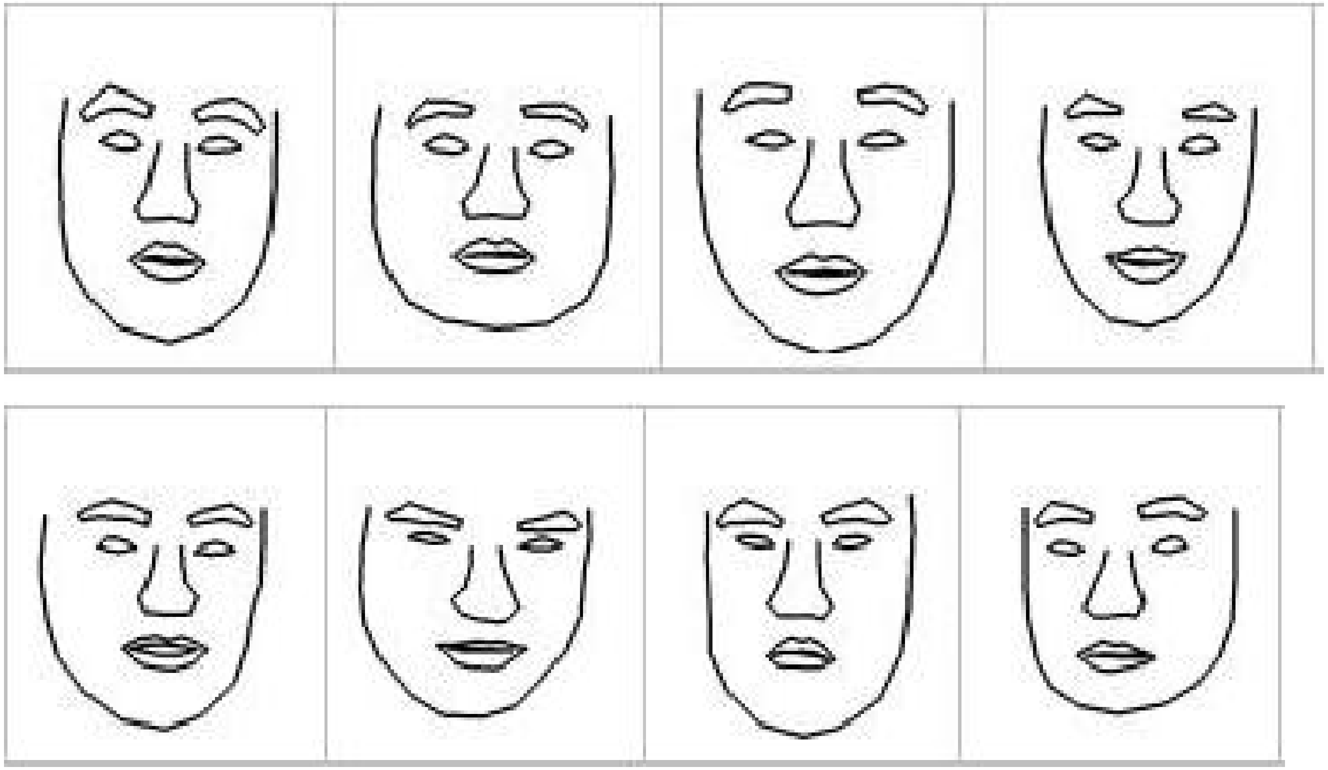


Image Parsing: Zhu & Tu

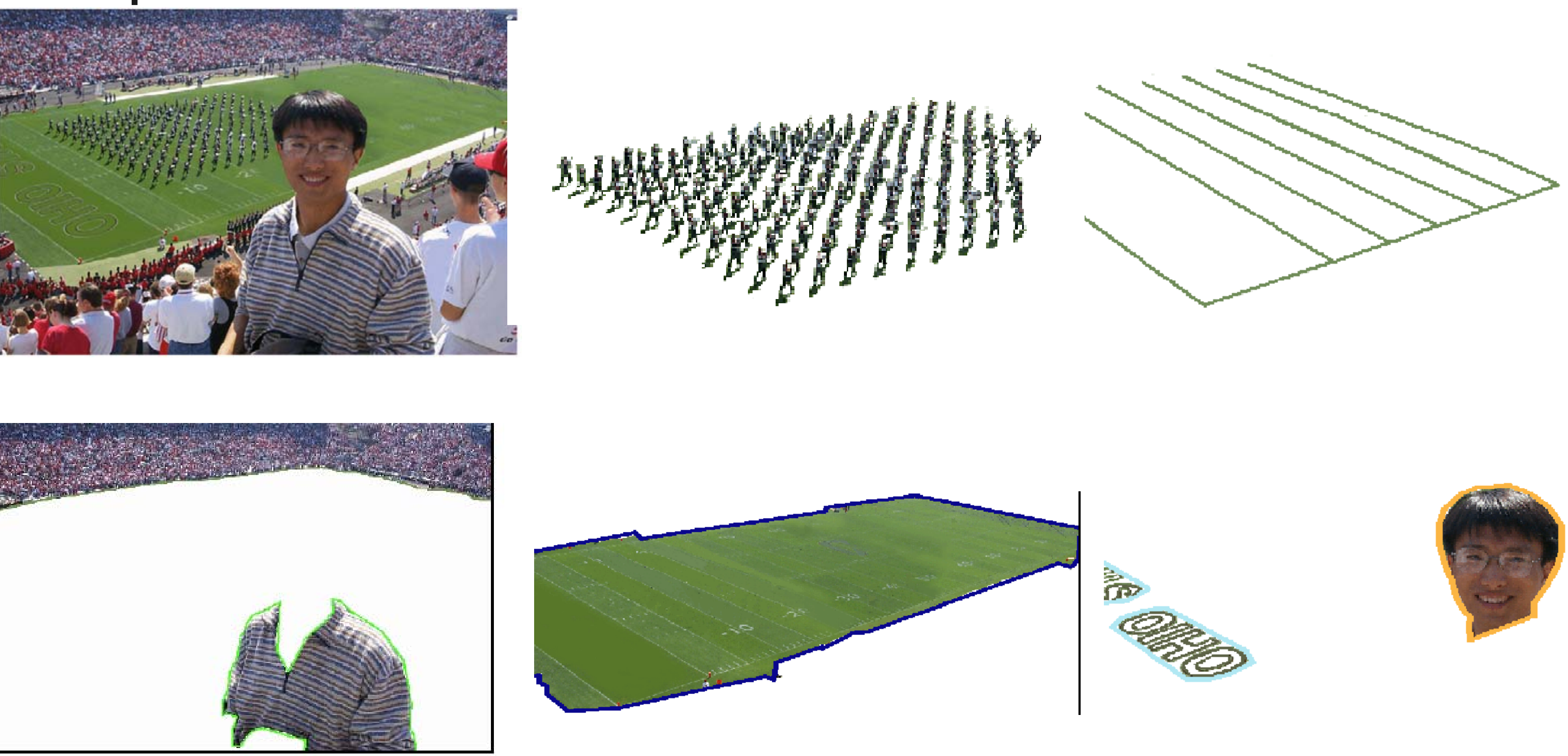
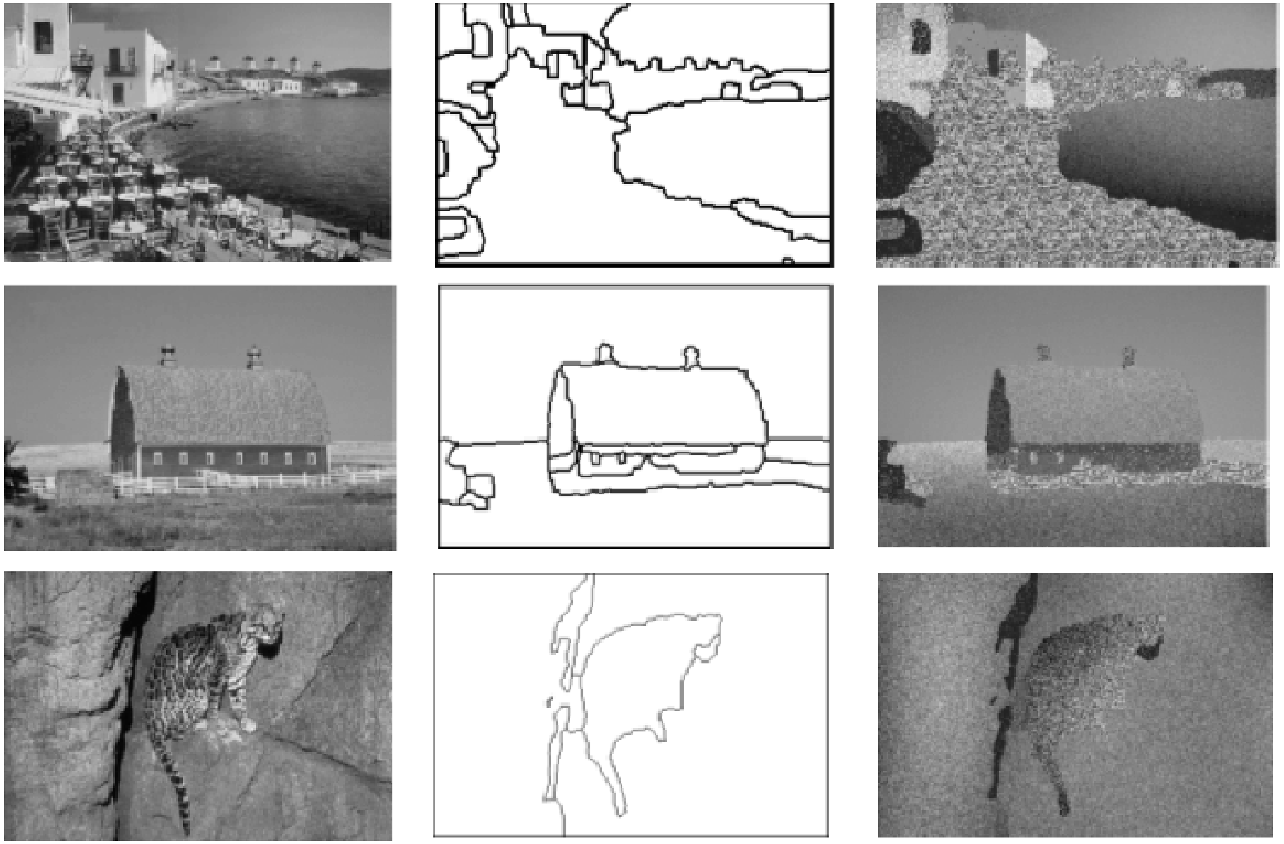


Image Parsing: Zhu & Tu.

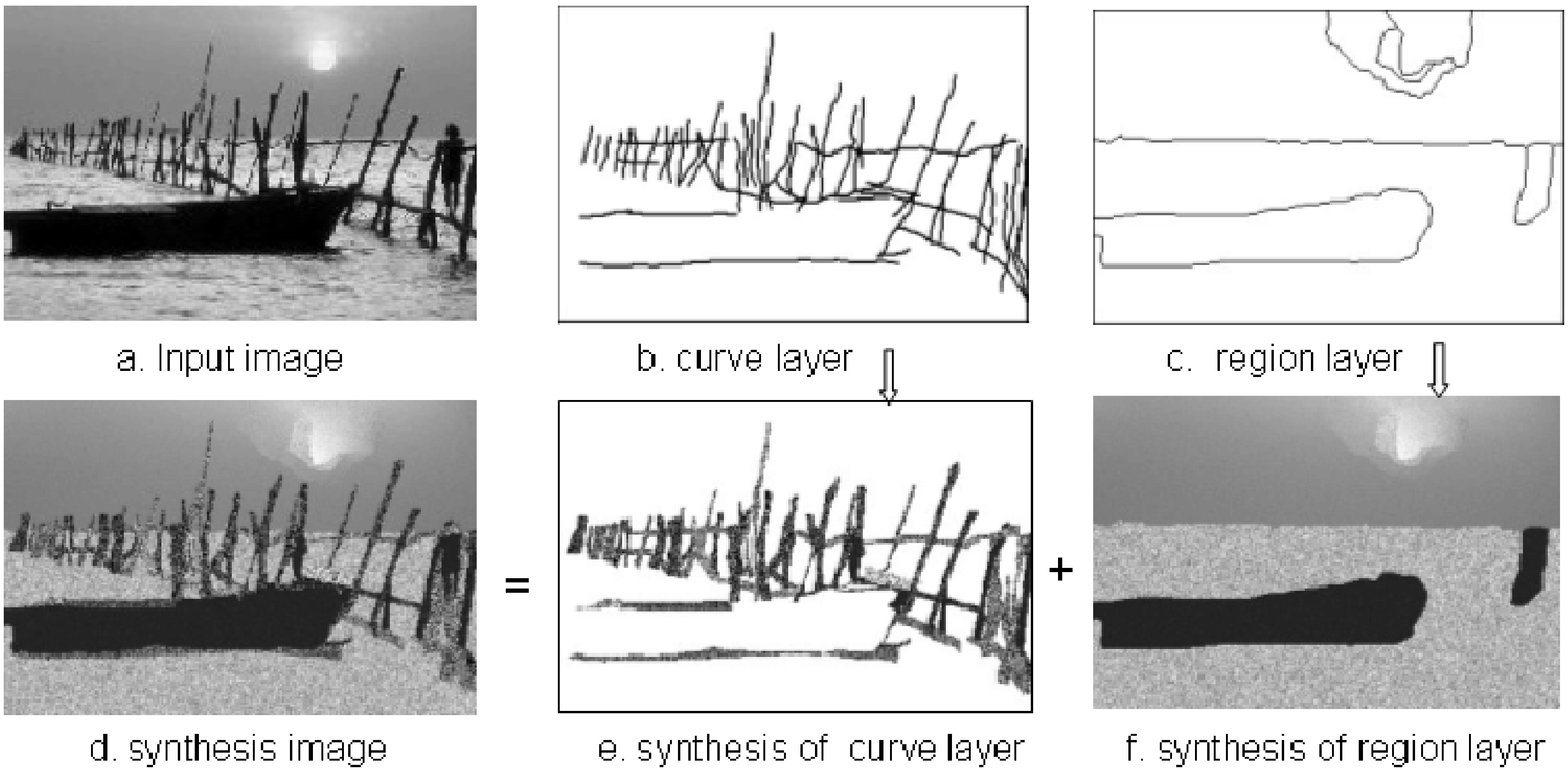
- Bayesian Formulation: model image as being composed of multiple regions.
- Boundaries of regions obey (probabilistic) constraints (e.g. smoothness)
- Intensity properties within regions are described by a set of models with unknown parameters (to be estimated).

Image Parsing Results:

Input, Segmentation, and Synthesis.



Regions, Curves, Occlusions.



Removing Foreground.

- "Denoising" images by removing foreground clutter.

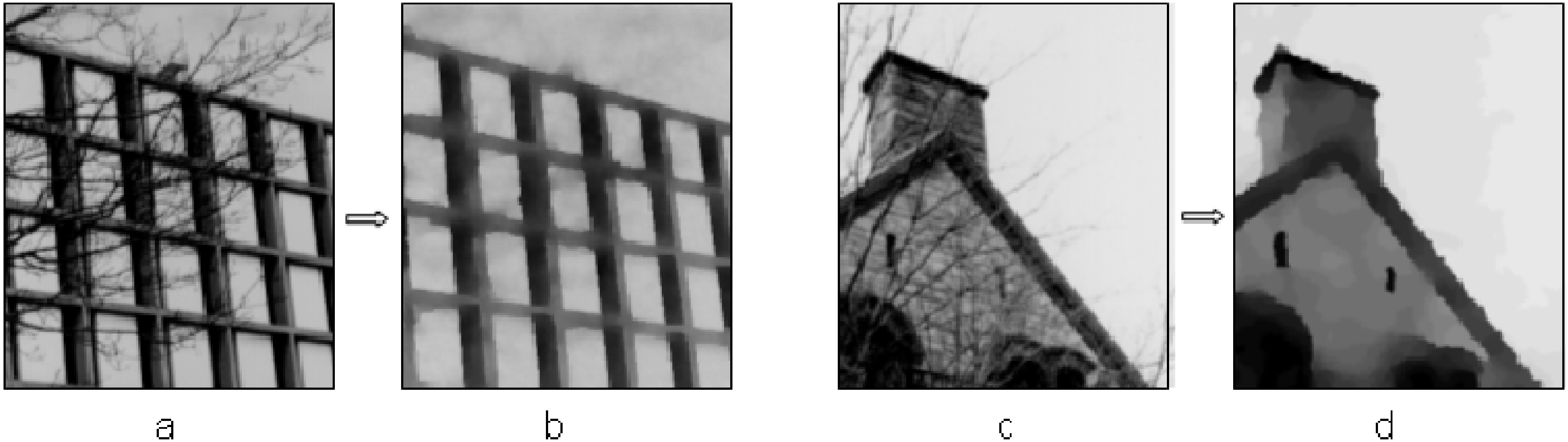
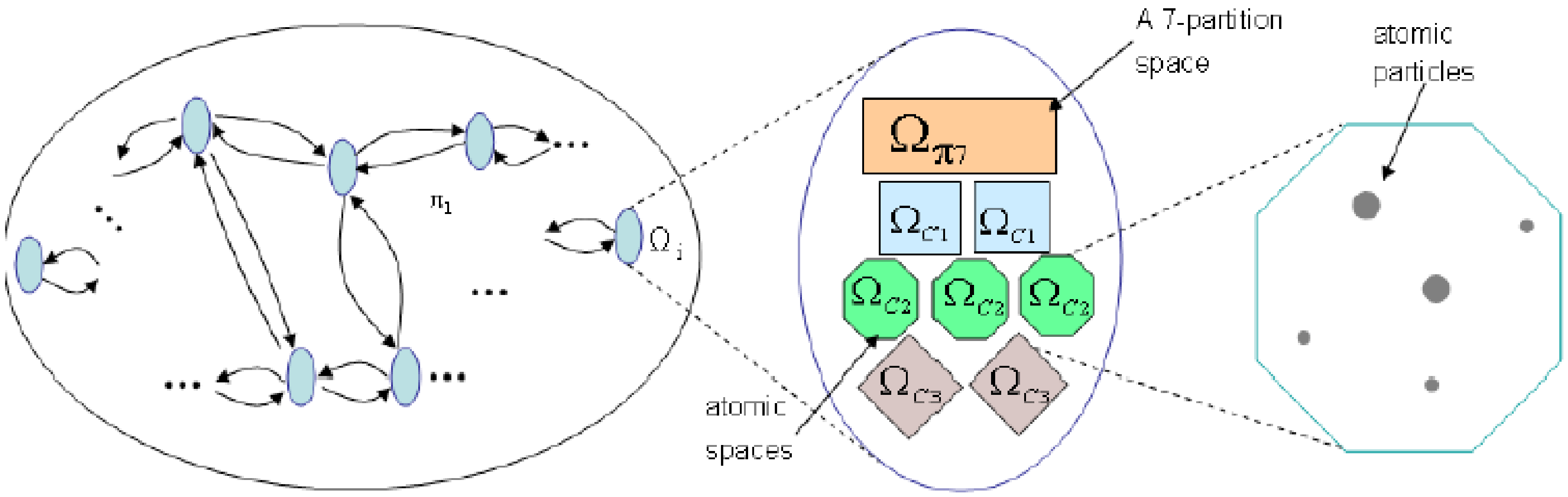


Image Parsing Solution Space

- No. regions, Types of regions, Properties of regions.



a). solution space b). a sub-space of 7 regions c). an atomic space

Machine Learning & Bayes.

- Zhu-Tu’s algorithm is called DDMCMC Data-Driven Markov Chain Monte Carlo.
- Discrimination methods (e.g. AdaBoost) can be used as *proposal probabilities*, which can be verified by Bayesian pattern models.

Machine Learning & Bayes

- Machine Learning seems to concentrate on discrimination problems.
- A whole range of other vision problems – image segmentation, image matching, viewpoint estimation, etc.
- Probability models for image patterns are learnable. These models give reality checks by synthesis.

Machine Learning & Bayes

- Machine Learning's big advantage over Bayes is speed (when applicable).
- AdaBoost may be particularly useful for combining local cues.
- Machine Learning for computational search to enable Bayesian estimation?