The Primal Sketch.

Images are regions and boundaries.

- Model the intensity within regions by Markov Random Fields.
- Model the intensity changes at the boundaries by "sparse coding" models. → sketch graph → structure unmixing.

Algorithm to compute the sketch graph and determine the HRF models.

Ingredients:

1. Sparse Coding models.

Represent an image by

\[ I(x,y) = \sum_{i=1}^{N} c_i B_i(x,y) + \epsilon(x,y) \]

\( B_i(x,y) \): \( i = 1 \ldots N \) are a set of basis functions. Gabor, Wavelets, etc.

Typically \( \{ B_i \} \) is an over-complete basis (i.e., \( N > \) image size).

\( \epsilon(x,y) \) is often assumed to be additive zero mean Gaussian noise.

Wavelet Theory & Harmonic Analysis (Mallat)
(2) The Primal Sketch

Overcomplete $\rightarrow$ Sparse Coding
require only a small number of coefficients
to be non-zero for each image.

Learn the set of basis functions that best
encode images sparsely $\rightarrow$ Gabor-like elements
(Olshausen et al.)

Gives a set of basis functions, can use
a matching pursuit algorithm (Mallat & Zhang) to
test encode the image.

Greedy algorithm: $\rightarrow$ Select the mth coefficient

$$c_m \rightarrow \min$$

$$\sum_{x,y} \left( I(x,y) - \sum_{i=1}^{m} c_i b_i(x,y) \right)^2.$$ 

Assume $c_1 \ldots c_m$ given

Results with small number of basis functions
are not impressive for reconstruction

$\rightarrow$ problems: (a) direct sensor texture requires
well (b) give uneven edges, which
are poorly aligned.
(3.) The Primal Sketch

Ingredients:

(2) Markov Random Field (MRF) model:

Set of filters \( \{ F_k : k = 1 \ldots K \} \)

Response at \((x, y)\) to filter \(k\) is

\[
[F_k + I](x, y) = F_k(x, y)
\]

Zhu, Wu, Mumford \rightarrow use the Minimax Entropy principle to get a distribution

\[
p(I) = \frac{1}{Z} e^{-\frac{1}{\alpha} \sum_{k} \psi_k(F_k + I(x, y))}
\]

(see previous lectures)

Feature statistics: summarized by histogram

\[
h_{k,bz} = \frac{1}{|\Lambda|} \sum_{(x,y) \in \Lambda} \delta(z : F_k(x, y))
\]

\(\delta(z : y) = 1\) if \(x \in \text{mem}\) \(z\)

\(\delta(z : x) = 0\), otherwise

\(p(I)\) is max entropy distribution based on

the histogram of the features. (as for weak membrane)

require

\[
\sum_{I} p(I) h_{k,z} = h_{\text{obs}}^z
\]
The Primal Sketch

General Result: Information Theory
Statistical Physics

For sufficiently large \( N \), the only images that are likely are those for which the observed histogram to equal to the true histogram (ergodically) and all of these are equally likely.

Image Ensemble \( \Omega(h) = \{ I : h(I) = h \} \).

So, you can either do the distribution by the MRF model \( p(I) \) or, by the uniform distribution over the image ensemble.

The distribution over the image ensemble is better if you have boundary conditions that you need to satisfy.

Note: how to select the factors \( k \)? This can be done by model selection, but requires many combinations of factors to be selected.

A greedy "filter pursuit" method is used: (Zhu, Lu, Mumford) (Lu, Zhu, Liu)
The Primal Sketch

Edges and Ridges $\mapsto$ Structure Domain
$\mapsto$ apply "sparse coding" only to edges and ridges.

In structure domain: $\Lambda_{nt}$

$$I(x,y) = \sum_{i=1}^{n} B(x,y|\theta_i) + \epsilon(x,y), \quad (x,y) \in \Lambda_{nt}$$

$\theta_i$ - geometric ad photometric parameters
(1 The $B(x,y|\theta)$ have minimal overlap)

Edge Second. $\mapsto$ construct along length of edge

$B_2(x,y | \theta) = f((x-u) \sin \alpha - (y-v) \cos \alpha)$

$f_0(x) = \begin{cases} 
-\frac{1}{2}, & x < 0 \\
0, & x = 0 \\
\frac{1}{2}, & x > 0 
\end{cases}$

$g_{\frac{x}{s}}(x)$ Gaussian with sd $s$.

Edge profile $+() = a + b f_0(x) + g_{\frac{x}{s}}(x)$

Ridge profile is a composition of two edge profile
Blob function
Corner & junction $\mapsto$ modeled as composition of edges and ridges

Altogether:
$$B(x,y | \theta) = \text{length with type } X$$

$\theta = (t,u,v,\alpha,\beta,\gamma,s,a,b)$

$X$: center orientation +3pronged + arcwise unclear.
16. The Primal Sketch

\[ S_{str} = (\Theta, i : i = 1 \ldots n) \] be the sketch graph

\[ V = \bigcup_{d=0}^2 V_d, \quad V_d \text{ set of nodes with degree } d \]

\[ d = 0 \quad \text{glob} \]
\[ d = 1 \quad \text{end point} \]
\[ d = 2 \quad \text{corner} \]

\[ p(S_{str}) \propto \exp \left\{ \frac{1}{2} \sum_{d=0}^2 V_d |V_d| \right\} \quad |V_d| = \text{no. of nodes with } d \text{ arms} \]

\[ \text{Integrated Model} \]

\[ \Rightarrow \text{combine structure model with texture model} \]

\[ S = (S_{str}, S_{tex}) \]

\[ p(I, S) = \frac{1}{Z} \exp \left\{ -\frac{1}{2} \sum_{i=1}^n \frac{1}{2} \sum_{(xy) \in S_{str} \cap \gamma} \right\} \]

\[ \left\{ \frac{1}{2} \sum_{i=1}^n \frac{1}{2} \sum_{(xy) \in S_{str} \cap \gamma} \right\} \sum_{n=1}^K \Phi_n(I'(xy)) \cdot \Phi_n(I'(xy)) \]

\[ \theta_{tex}(S_{tex}) = \text{pm} \quad \text{penalty, neg of vertex,} \]

Computational Algorithm

Greedy cy (6) based

1. Fix Edges

2. Model Edges (structure model)

3. Model Texture
(7) The Primal Sketch

First Part → Sketch Procedure

Objective Function for Computing $S_{tr}$

Hypothesis Test

$H_0: I(x,y) = \mu + N(0, \sigma^2)$

$H_1: I(x,y) = R(x,y, \theta) + N(0, \sigma^2)$ \space \forall \theta$

$L(\theta) = \prod_{(x,y) \in \Gamma} (I(x,y) - \mu)^2 - (I(x,y) - R(x,y, \theta))^2$

Add prior

$L(S_{tr}) = \frac{1}{2\sigma^2} \sum_{i=1}^{n} L(\theta_i) - S_{tr}(S_{tr})$

Note: The "default" model is Gaussian, but it really should be t-exact (HT)

Sketch Purified

Phase 0: Edge & ridge detectors. Based on localized linear filters gives an initial skeleton for sketch graph.

Battery of filters

Lanni, DoG's, Gates

Output edge-ridge map. Multiple scales + Threshold model & RMS

Phase 1: Sequentially add more filters (normalize input to have zero mean and variance 1)

Algorithm outputs frontal edge-ridge peaks of maximum strength

Connect true segments by spatial grouping
The Primal Sketch

Next: define profile by averaging the image intensity on the 20 x 20 region

Compute $\Delta l(B)$ to fit a linear function.

Repeat until you edge [edge]

Phase 2

Fix corners and junction by graph spencer

Greedy algorithm to decrease:

$$L(S_{set}) = \frac{1}{2} \sum_{i=1}^{n} \Delta L(B_i) - \frac{4}{2} \sum_{d=0}^{d} Val(d)$$

Set of graph operations $\rightarrow$ gives $A_{str}$

Texture clustering and synthesis

$A_{tex} = \Lambda / A_{str}$

Apply $k$-mean clustering to divide $A_{tex}$ into homogeneous texture regions.

Then can learn models and apply texture synthesis

(small no. of filters within 3-7)

Advantages:

(i) Derivative to model edges properly rather than guessing them. (Weak matching, requires curve)
(ii) Generic model for entire image.