Markov Nets

Dhruv Batra,
10-708 Recitation

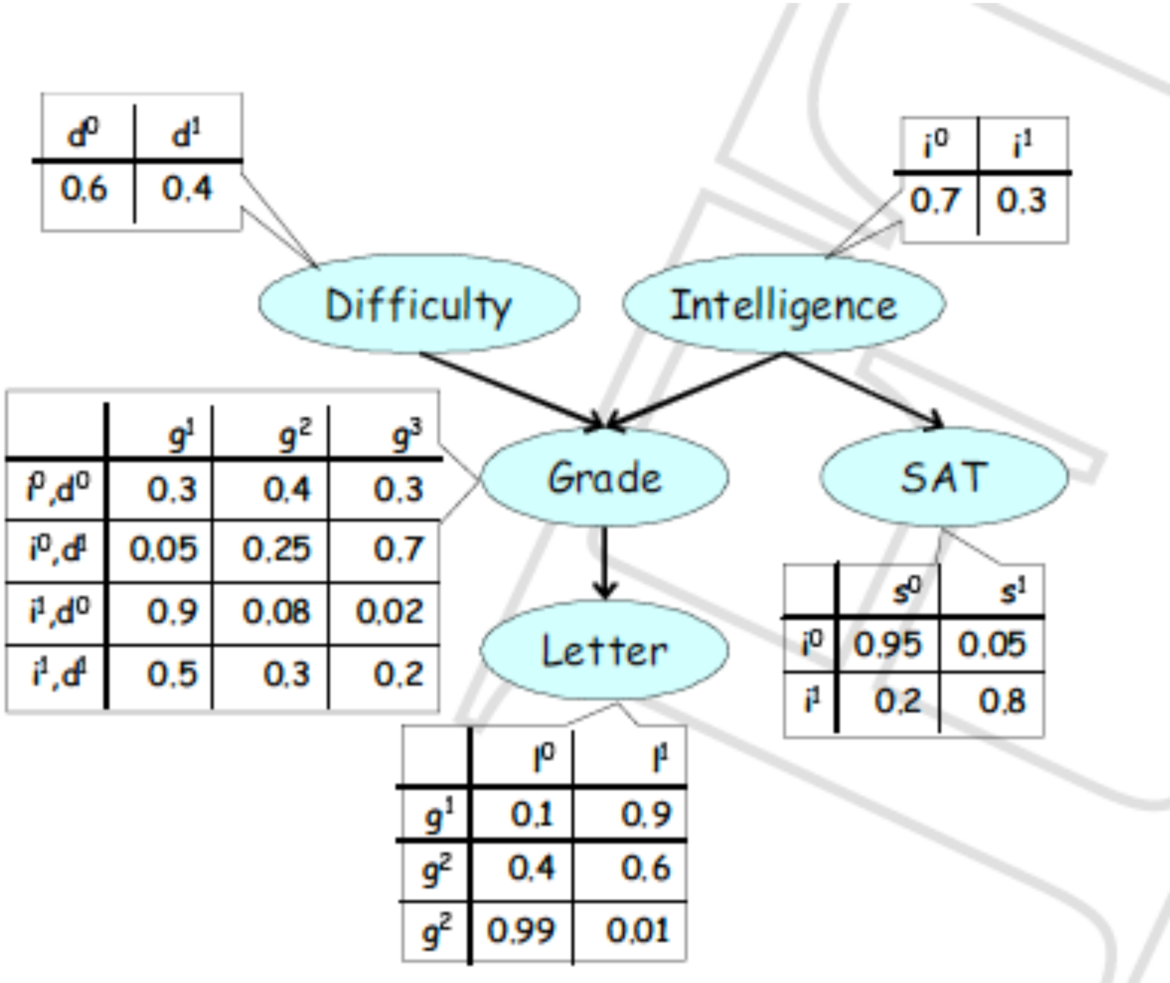
10/30/2008

Contents

- MRFs
  - Semantics / Comparisons with BNs
  - Applications to vision
  - HW4 implementation

Semantics

- Bayes Nets



d^0	d^1
0.6	0.4

i^0	i^1
0.7	0.3

	g^1	g^2	g^3
i^0,d^0	0.3	0.4	0.3
i^0,d^1	0.05	0.25	0.7
i^1,d^0	0.9	0.08	0.02
i^1,d^1	0.5	0.3	0.2

	s^0	s^1
i^0	0.95	0.05
i^1	0.2	0.8

	l^0	l^1
g^1	0.1	0.9
g^2	0.4	0.6
g^2	0.99	0.01

Semantics

- Markov Nets

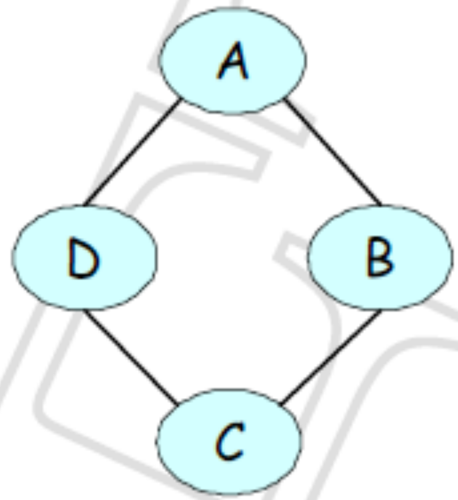


Figure 5.1 The undirected graph for the Misconception example.

$\pi_1[A, B]$

a^0	b^0	30
a^0	b^1	5
a^1	b^0	1
a^1	b^1	10

$\pi_2[B, C]$

b^0	c^0	100
b^0	c^1	1
b^1	c^0	1
b^1	c^1	100

$\pi_3[C, D]$

c^0	d^0	1
c^0	d^1	100
c^1	d^0	100
c^1	d^1	1

$\pi_4[D, A]$

d^0	a^0	100
d^0	a^1	1
d^1	a^0	1
d^1	a^1	100

Semantics

- Decomposition
  - Bayes Nets

$$P(I, D, G, S, L) = P(I)P(D)P(G \mid I, D)P(S \mid I)P(L \mid G).$$

  - Markov Nets

$$P(a, b, c, d) = \frac{1}{Z}\pi_1[a, b] \cdot \pi_2[b, c] \cdot \pi_3[c, d] \cdot \pi_4[d, a]$$

where

$$Z = \sum_{a,b,c,d} \pi_1[a, b] \cdot \pi_2[b, c] \cdot \pi_3[c, d] \cdot \pi_4[d, a]$$

Semantics

- Factorization

Definition 5.3.2: Let $\mathcal{H}$ be a Markov network. A distribution P *factorizes* over $\mathcal{H}$ if it is associated with:

- a set of subsets $\boldsymbol{D}_1, \ldots, \boldsymbol{D}_m$, where each $\boldsymbol{D}_i$ is a complete subgraph of $\mathcal{H}$,
- factors $\pi_1[\boldsymbol{D}_1], \ldots, \pi_m[\boldsymbol{D}_m]$,

such that

$$P(X_1, \ldots, X_n) = \frac{1}{Z} P'(X_1, \ldots, X_n)$$

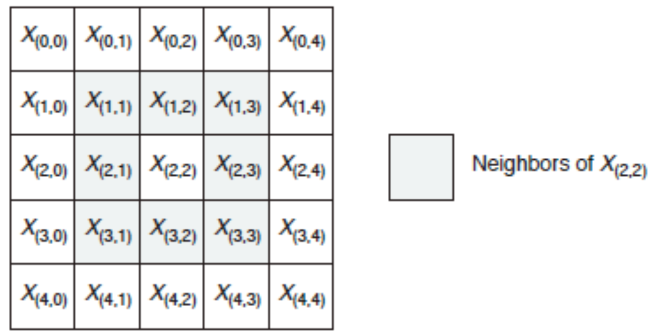
where

$$P'(X_1, \ldots, X_n) = \pi_i[\boldsymbol{D}_1] \times \pi_2[\boldsymbol{D}_2] \times \cdots \times \pi_m[\boldsymbol{D}_m]$$

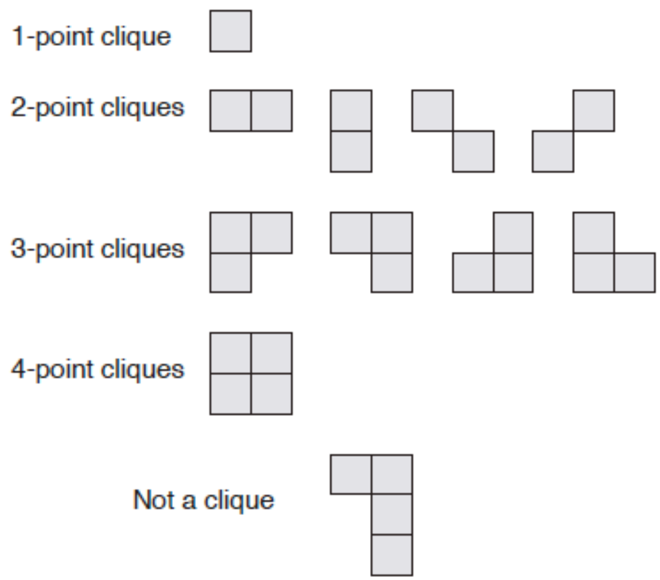
- What happens in BNs?

Example of Neighborhood System and Clique

- Example of 8 point neighborhood



- Example of cliques for 8 point neighborhood



Semantics

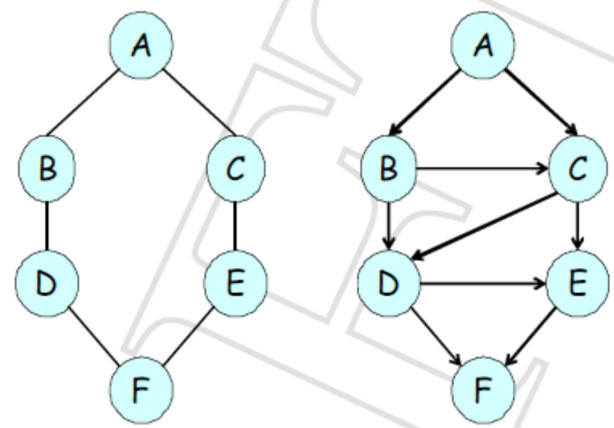
- Active Trails in MNs

Definition 5.2.2: Let $\mathcal{H}$ be a Markov network structure, and $X_1 — \ldots — X_k$ be a path in $\mathcal{H}$. Let $\boldsymbol{E} \subseteq \mathcal{X}$ be a set of *observed variables*. The path $X_1 — \ldots — X_k$ is *active* given $\boldsymbol{E}$ if none of the X_i's, $i = 1, \ldots, k$, is in $\boldsymbol{E}$. ∎

- What happens in BNs?

Semantics

- Independence Assumptions



	Markov Nets	Bayes Nets
Global Ind Assumption, d-sep	$\mathcal{I}(\mathcal{H}) = \{(\boldsymbol{X} \perp \boldsymbol{Y} \mid \boldsymbol{Z}) : sep_{\mathcal{H}}(\boldsymbol{X};\boldsymbol{Y} \mid \boldsymbol{Z})\}.$	
Local Ind Assumptions	$\mathcal{I}_p(\mathcal{H}) = \{(X \perp Y \mid \mathcal{X} - \{X,Y\}) : X—Y \notin \mathcal{H}\}.$	
Markov Blanket	$(X \perp \mathcal{X} - \{X\} - \boldsymbol{U} \mid \boldsymbol{U}) \in \mathcal{I}$	

Representation Theorem for Markov Networks

If joint probability distribution *P*:

$$P(X_1, \ldots, X_n) = \frac{1}{Z} \prod_{i=1}^{m} \phi_i(\mathbf{D}_i)$$

Then

H is an I-map for *P*

If *H* is an I-map for *P*

and

P is a positive distribution

Then

joint probability distribution *P*:

$$P(X_1, \ldots, X_n) = \frac{1}{Z} \prod_{i=1}^{m} \phi_i(\mathbf{D}_i)$$

Semantics

- Factorization

$$P(X_1, \ldots, X_n) = \frac{1}{Z} P'(X_1, \ldots, X_n)$$

where

$$P'(X_1, \ldots, X_n) = \pi_i[\boldsymbol{D}_1] \times \pi_2[\boldsymbol{D}_2] \times \cdots \times \pi_m[\boldsymbol{D}_m]$$

- Energy functions

$$\pi[\boldsymbol{D}] = \exp(-\epsilon[\boldsymbol{D}]),$$

- Equivalent representation

$$P(X_1, \ldots, X_n) \propto \exp\left[-\sum_{i=1}^{m} \epsilon_i[\boldsymbol{D}_i]\right].$$

Semantics

- Log Linear Models

Definition 5.6.9: A distribution P is a *log-linear model* over a Markov network $\mathcal{H}$ if it is associated with:

- a set of features $\phi_1[\boldsymbol{D}_1], \ldots, \phi_k[\boldsymbol{D}_m]$, where each $\boldsymbol{D}_i$ is a subclique in $\mathcal{H}$,
- a set of weights $w_1, \ldots, w_k$,

such that

$$P(X_1, \ldots, X_n) = \frac{1}{Z} \exp\left[-\sum_{i=1}^{k} w_i \phi_i[\boldsymbol{D}_i]\right].$$

Semantics

- Energies in pairwise MRFs

$$E(f) = \sum_{\{p,q\} \in \mathcal{N}} V_{p,q}(f_p, f_q) + \sum_{p \in \mathcal{P}} D_p(f_p),$$

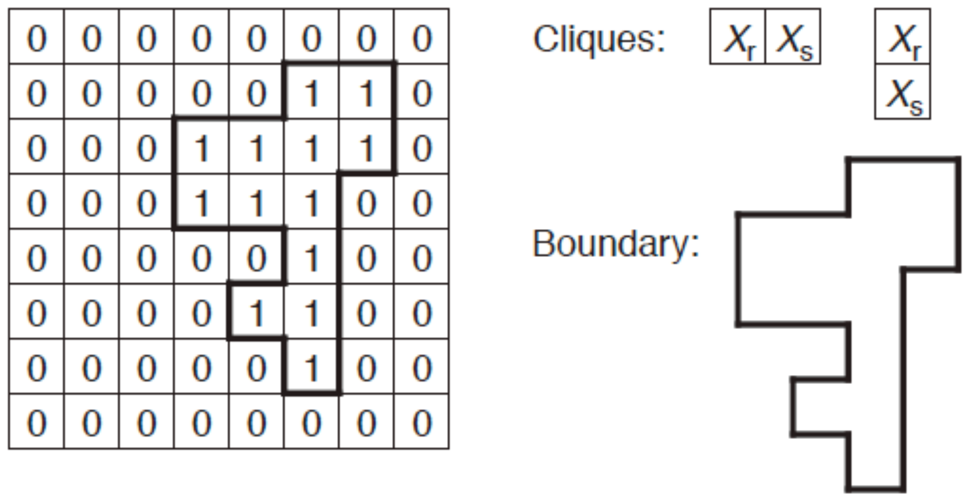
- What is encoded on nodes and edges?

$$E(f) = E_{smooth}(f) \quad + \quad E_{data}(f).$$

Semantics

- Priors on edges
  - Ising Prior / Potts Model
  - Metric MRFs

The Ising Model: A 2-D MRF[100]



- Potential functions are given by

$$V(x_r, x_s) = \beta\delta(x_r \neq x_s)$$

where β is a model parameter.

- Energy function is given by

$$\sum_{c \in \mathcal{C}} V_c(x_c) = \beta(\text{Boundary length})$$

- Longer boundaries $\Rightarrow$ less probable

Metric MRFs

- Energies in pairwise MRFs

$$E(f) = \sum_{\{p,q\} \in \mathcal{N}} V_{p,q}(f_p, f_q) + \sum_{p \in \mathcal{P}} D_p(f_p),$$

$$\begin{aligned} V(\alpha, \beta) = 0 \quad &\Leftrightarrow \quad \alpha = \beta, \\ V(\alpha, \beta) \quad &= \quad V(\beta, \alpha) \geq 0, \\ V(\alpha, \beta) \quad &\leq \quad V(\alpha, \gamma) + V(\gamma, \beta), \end{aligned}$$

Applications in Vision

- Image Labelling tasks
  - Denoising
  - Stereo
  - Segmentation, Object Recognition
  - Geometry Estimation

MRF nodes as pixels

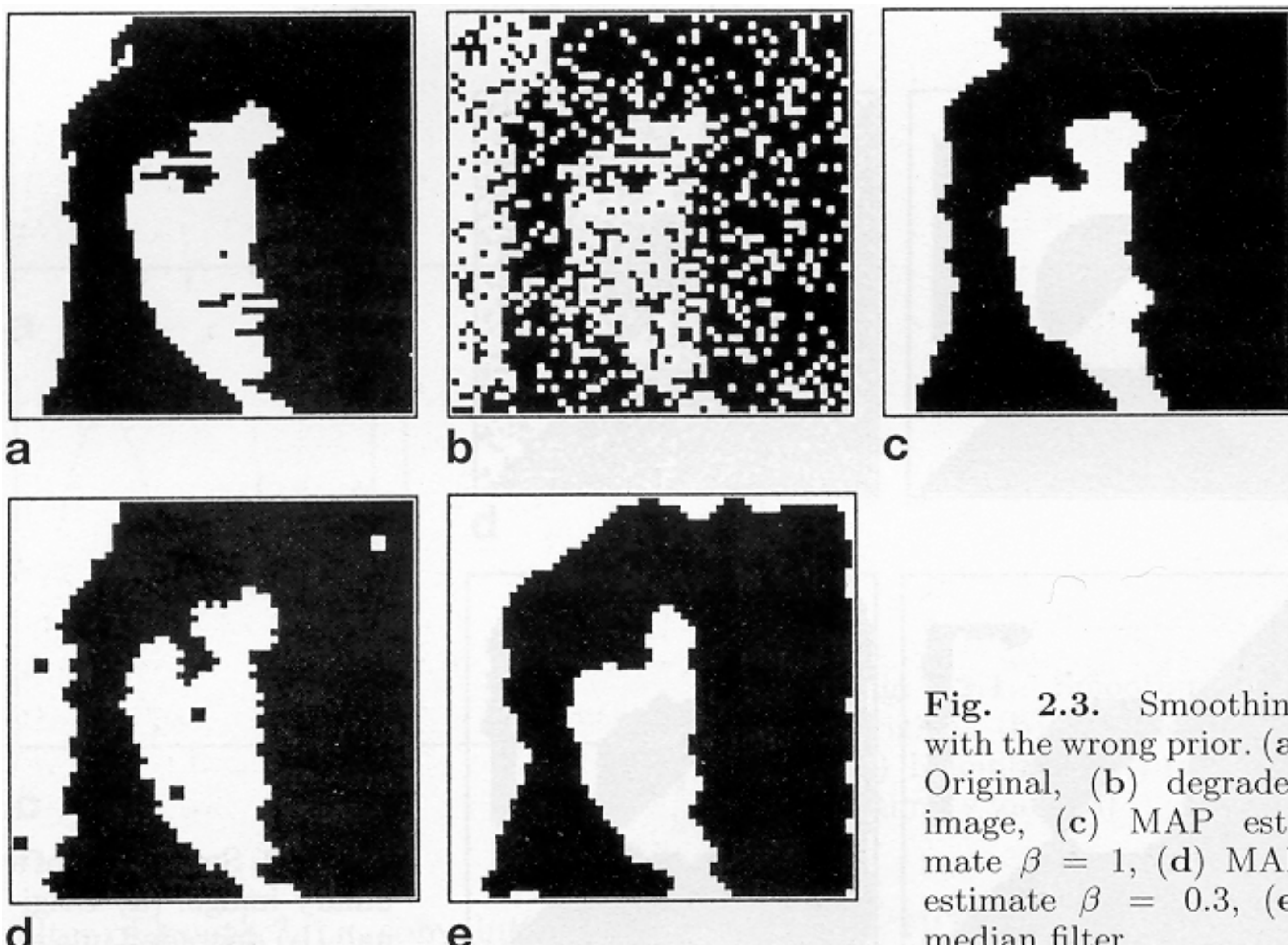
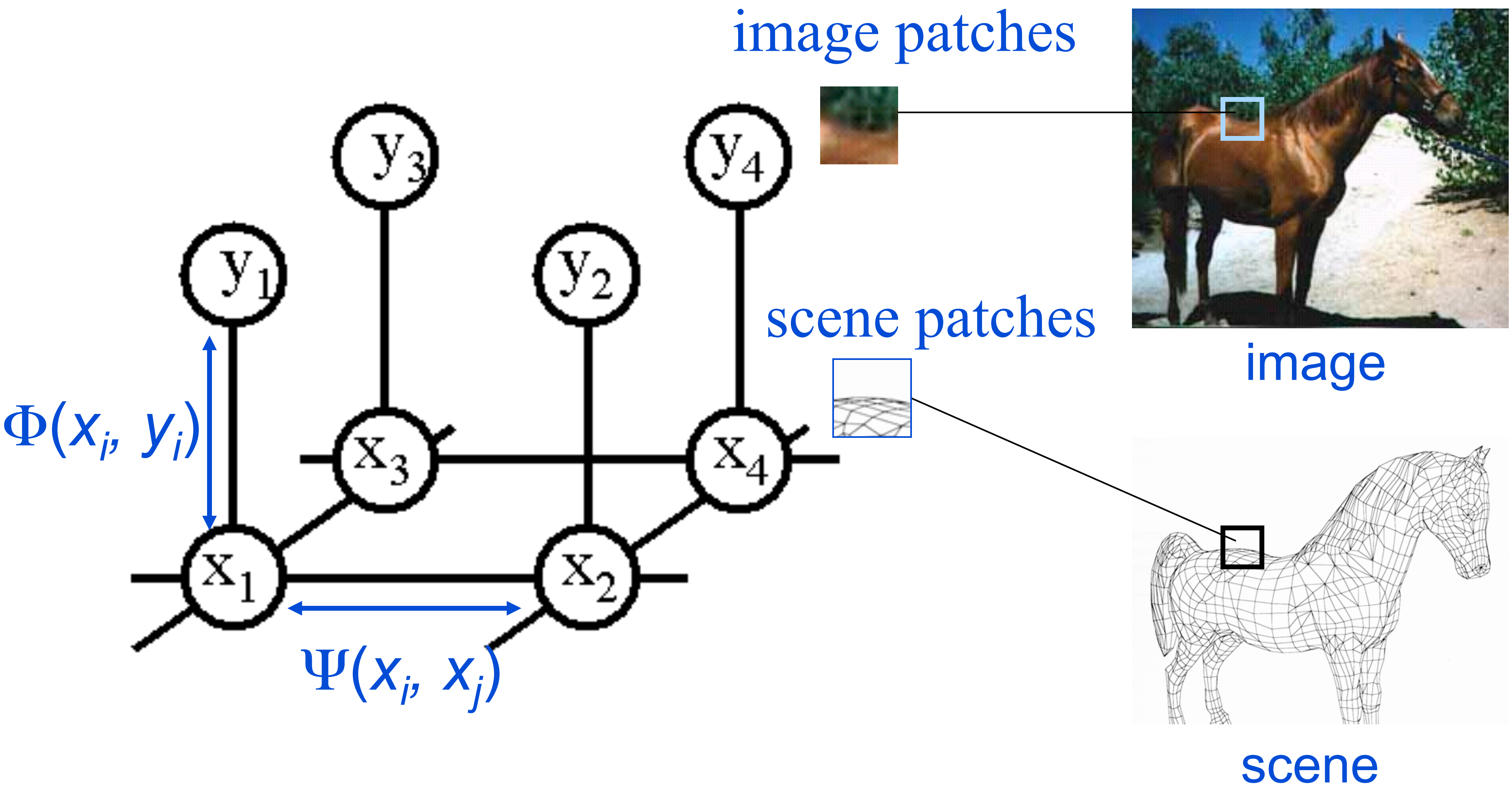


Fig. 2.3. Smoothing with the wrong prior. (**a**) Original, (**b**) degraded image, (**c**) MAP estimate $\beta = 1$, (**d**) MAP estimate $\beta = 0.3$, (**e**) median filter

Winkler, 1995, p. 32

MRF nodes as patches



Network joint probability

$$P(x,y) = \frac{1}{Z} \prod_{i,j} \Psi(\underbrace{x_i, x_j}_{\text{neighboring scene nodes}}) \prod_i \Phi(x_i, y_i)$$

scene

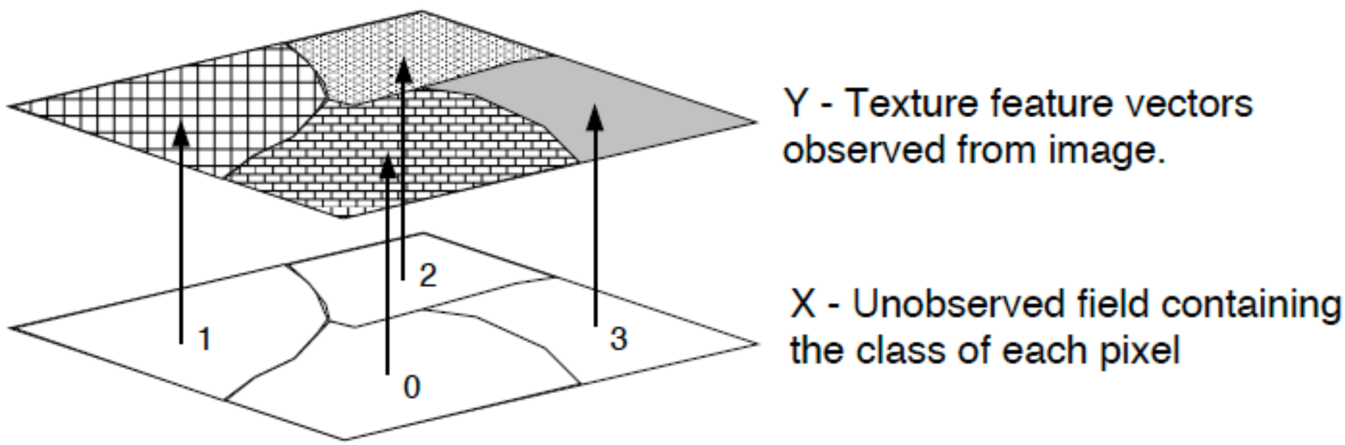
image

Scene-scene compatibility function

Image-scene compatibility function

local observations

Bayesian Segmentation Model



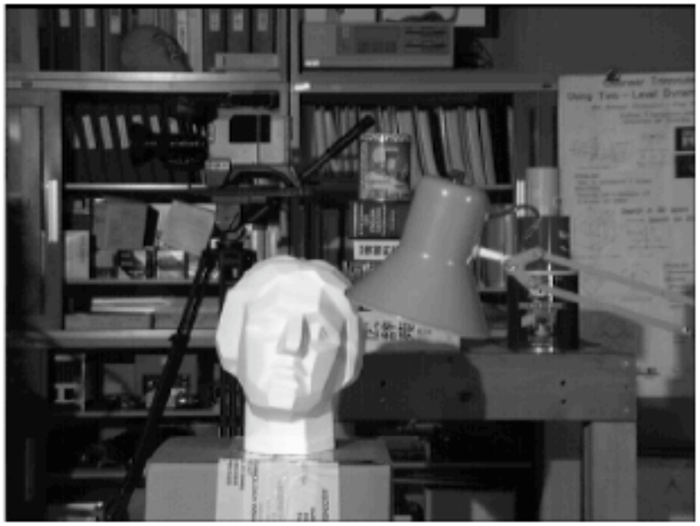
- Discrete MRF is used to model the segmentation field.
- Each class is represented by a value $X_s \in \{0, \cdots, M-1\}$
- The joint probability of the data and segmentation is

$$P\{Y \in dy, X = x\} = p(y|x)p(x)$$

where

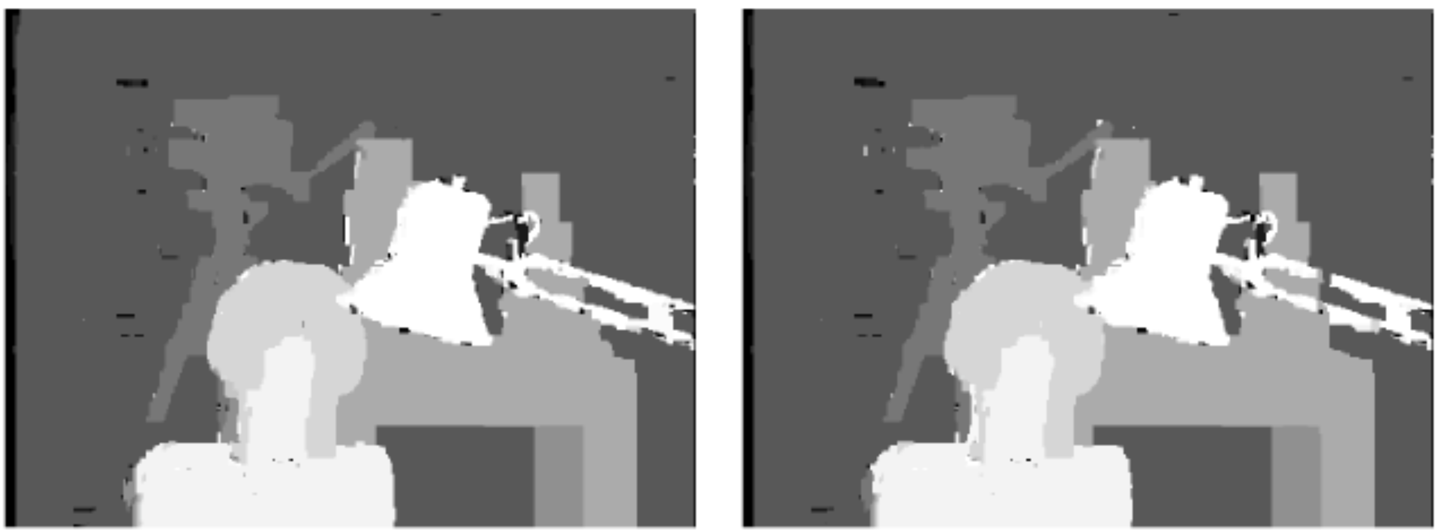
- $p(y|x)$ is the data model
- $p(x)$ is the segmentation model

Stereo

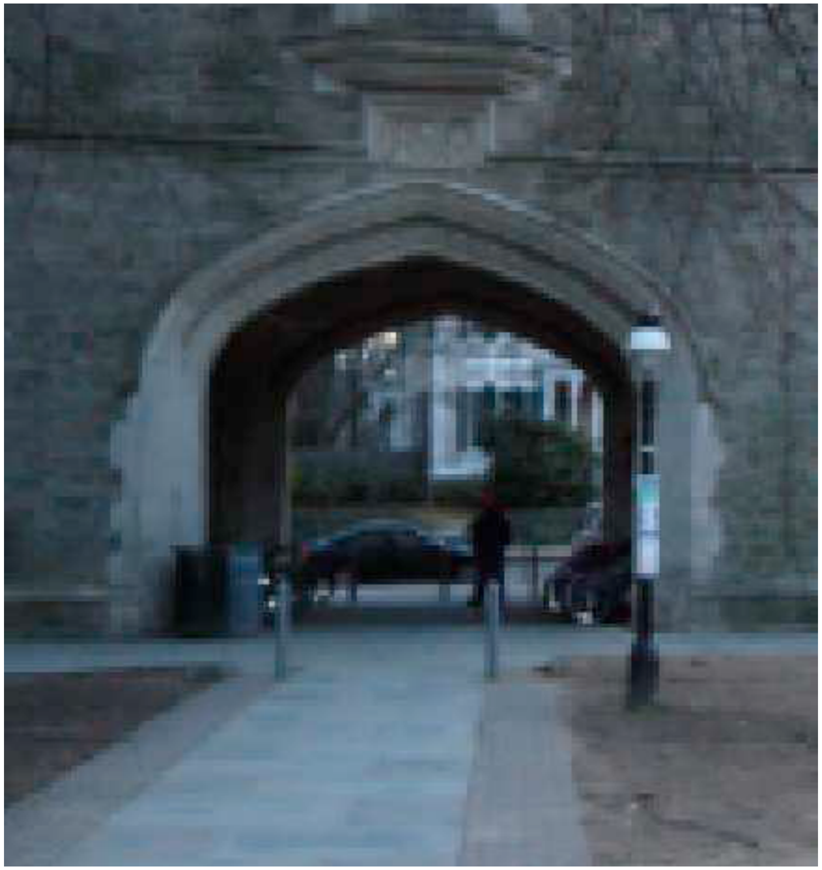


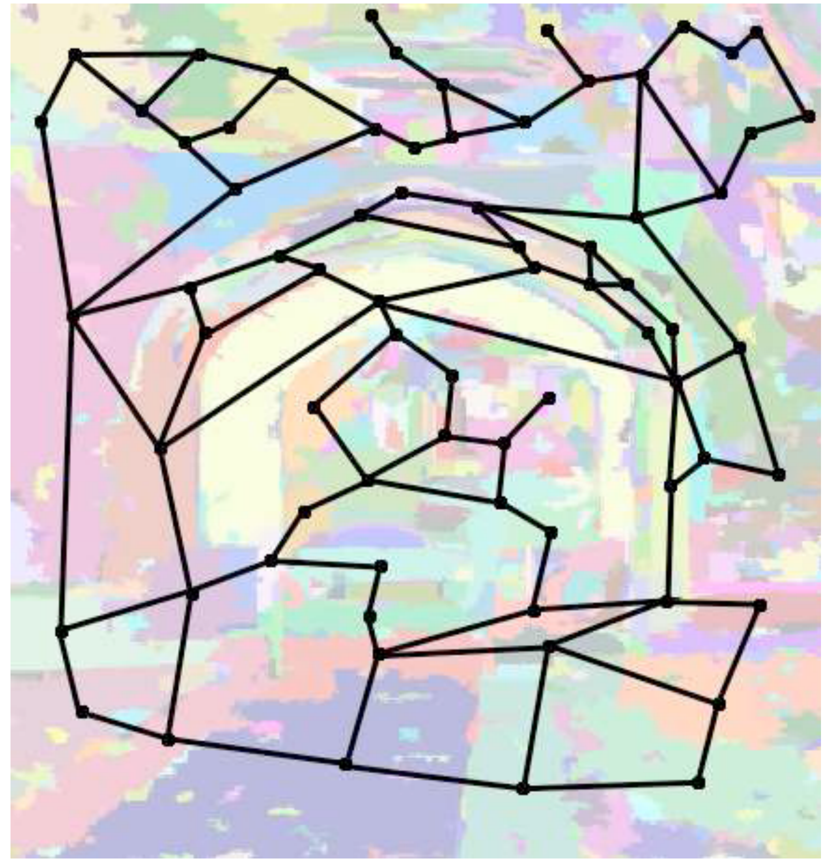
(a)

(b)

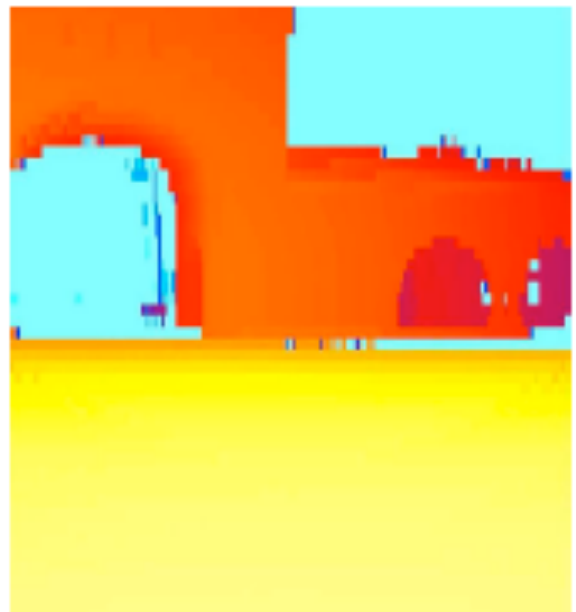


Geometry Estimation



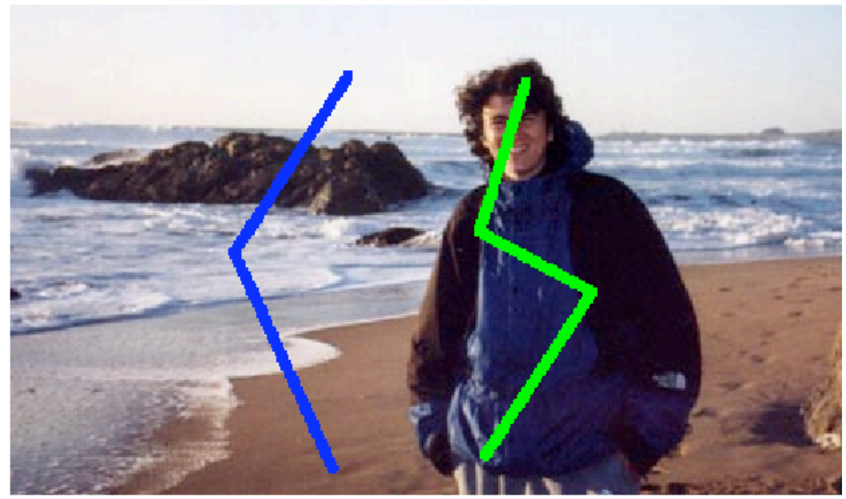


Geometry Estimation



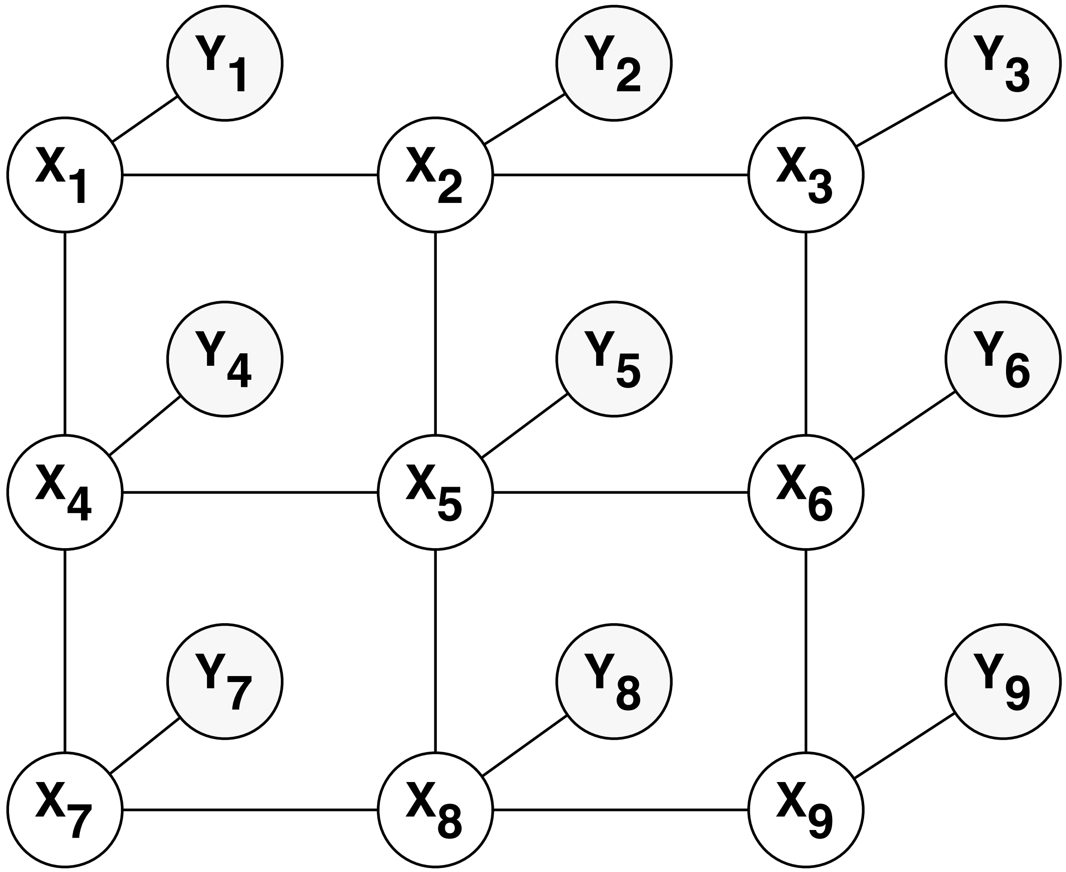
HW4

- Interactive Image Segmentation

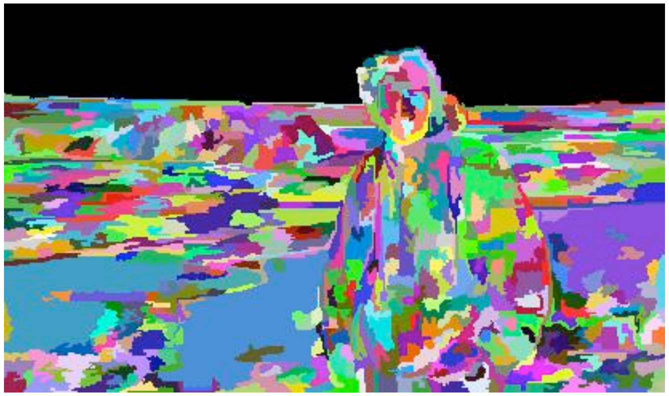
 

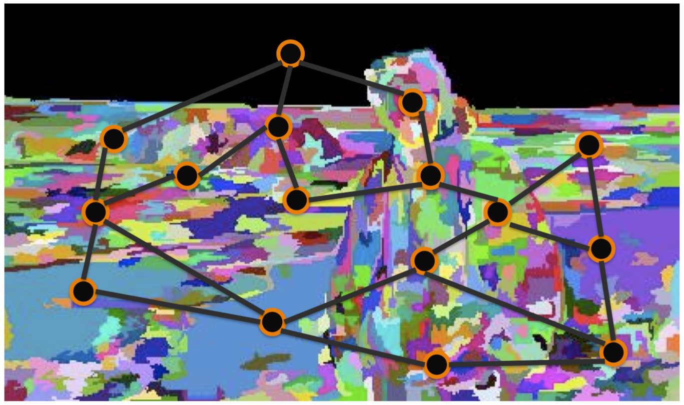
HW4

- Pairwise MRF



HW4





HW4

- Step 1: GMMs
- Step 2: Adjacency matrix / MRF Structure
- Step 3: MRF parameters

$$\Phi(X_i = \mathrm{fg}, y_i) = P(y_i|\mathrm{GMM}_{\mathrm{fg}})$$
$$\Phi(X_i = \mathrm{bg}, y_i) = P(y_i|\mathrm{GMM}_{\mathrm{bg}})$$

$$\Psi(x_i, x_j) = \exp\left\{-\beta \times I\left(x_i \neq x_j\right)\right\},$$

- Step 4: Loopy BP
- Step 5: Segmentation Masks

Slide Credits

- Bill Freeman, Fredo Durand, Lecture at MIT
  - http://groups.csail.mit.edu/graphics/classes/CompPhoto06/html/lecturenotes/2006March21MRF.ppt

- Charles A. Bouman
  - https://engineering.purdue.edu/~bouman/ece641/mrf_tutorial/view.pdf

- Fast Approximate Energy Minimization via Graph Cuts, Yuri Boykov, Olga Veksler and Ramin Zabih. IEEE PAMI 23(11), November 2001.

- Make3D: Learning 3-D Scene Structure from a Single Still Image, Ashutosh Saxena, Min Sun, Andrew Y. Ng, To appear in IEEE PAMI 2008

Questions?