IPAM
February 2013

Western
University

Segmentation with non-linear constraints on appearance, complexity, and geometry

Yuri Boykov

Andrew Delong | Lena Gorelick | Hossam Isack | Anton Osokin | Frank Schmidt | Olga Veksler






Overview

Standard linear constraints on segments

- intensity log-likelihoods, volumetric ballooning, etc.

1. Basic non-linear regional constraints

- enforcing intensity distribution (KL, Bhattacharia, L_2)
- constraints on volume and shape

2. Complexity constraints (label costs)

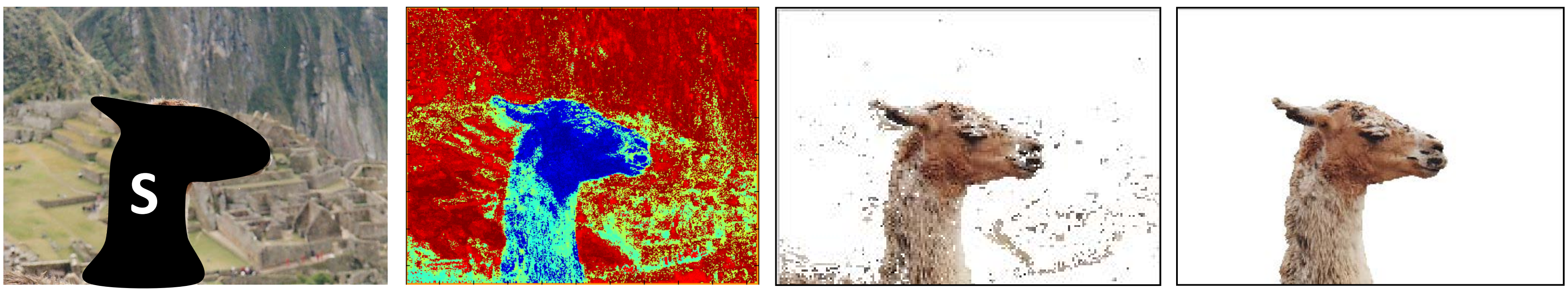
- unsupervised and supervised image segmentation, compression
- geometric model fitting

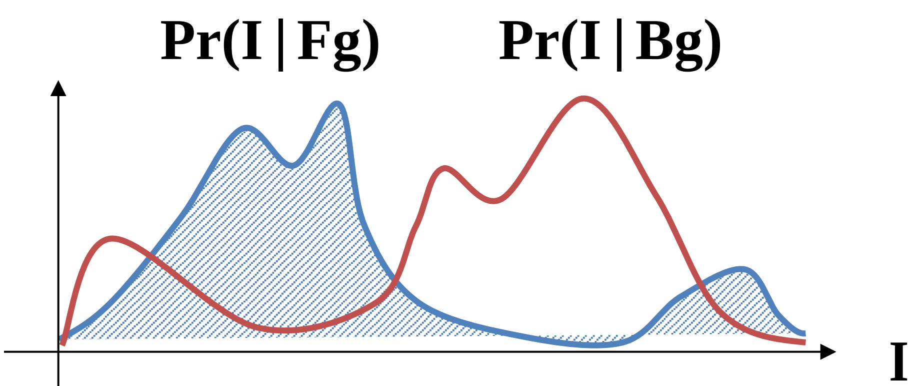
3. Geometric constraints

- unsupervised and supervised image segmentation, compression
- geometric model fitting (lines, circles, planes, homographies, motion,…)

Image segmentation
Basics

$$E(\mathbf{S}) = \sum_p \langle \mathbf{f}_p, \mathbf{S}_p \rangle + \mathbf{B}(\mathbf{S})$$





$$f_p = -\ln\left(\frac{\Pr(I_p \mid fg)}{\Pr(I_p \mid bg)}\right)$$

Linear appearance of region S

$$R(S) = \langle f, S \rangle$$

Examples of potential functions $\boldsymbol{f}$

- Log-likelihoods $\quad f_p = -\ln Pr(I_p)$
- Chan-Vese $\quad f_p = (I_p - c)^2$
- Ballooning $\quad f_p = -1$

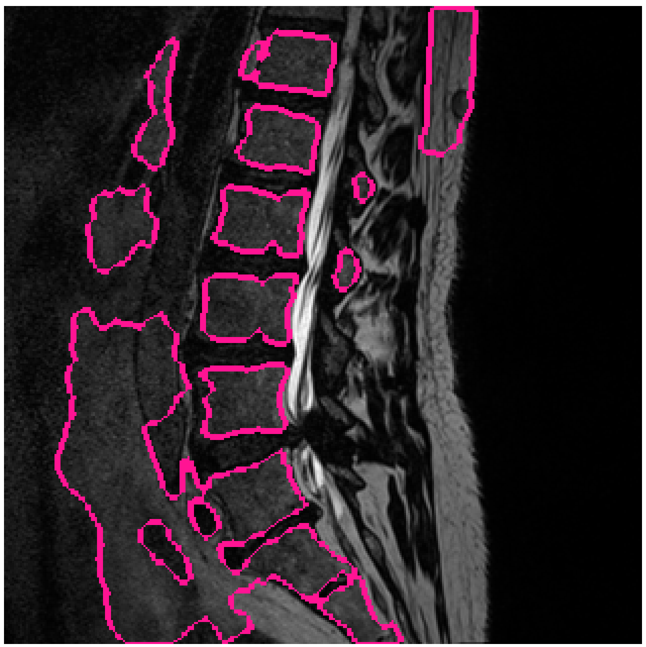
Part 1

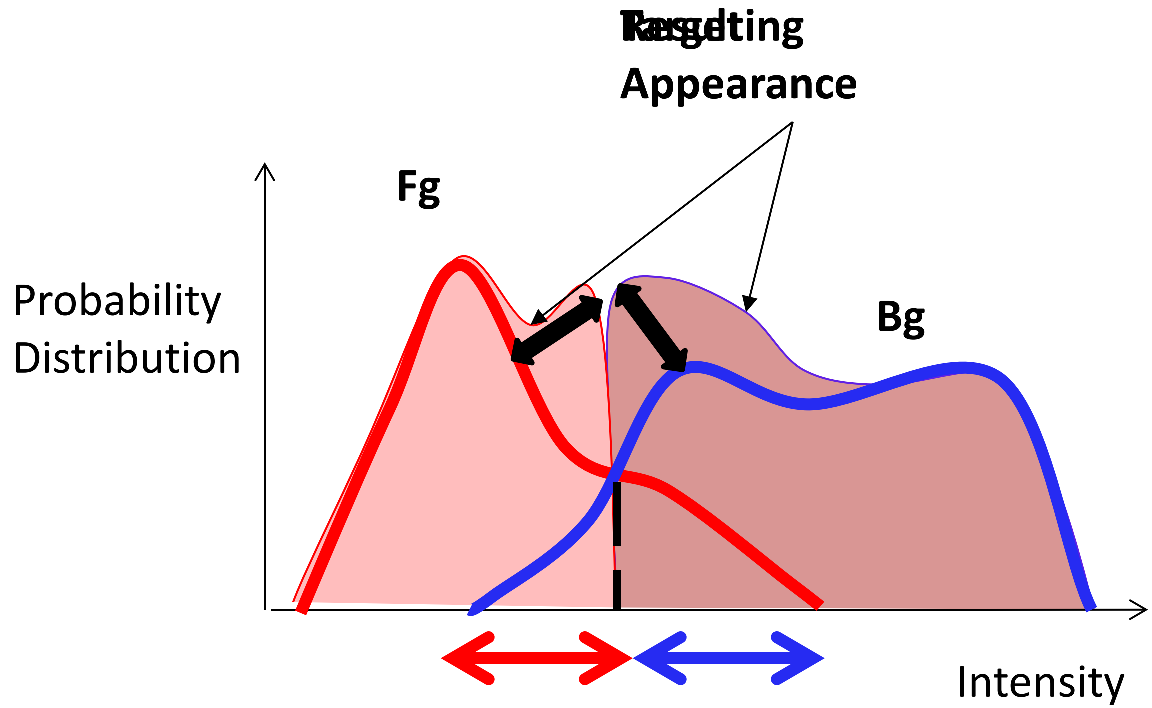
Basic non-linear regional functionals

$$\sum_{p \in S} -1 \quad \Rightarrow \quad (vol(S) - vol_0)^2$$

$$\sum_{p \in S} -\ln Pr(I_p \mid S) \quad \Rightarrow \quad || hist(S) - hist_0 ||$$

Standard Segmentation Energy





Minimize Distance to Target Appearance Model

$$\sum_{p \in S} -\ln\left(\frac{\Pr(I_p \mid fg)}{\Pr(I_p \mid bg)}\right)$$

$$R(S) = \| S - T \|_{L^2}$$

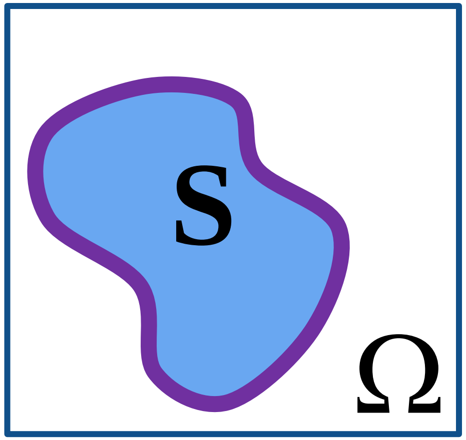
$$R(S) = KL(S \| T)$$

$$R(S) = Bha(S, T)$$

Non-linear harder to optimize regional term

$$\mathbf{E(S)} = \mathbf{R(S)} + \mathbf{B(S)}$$

non-linear
regional term



– **appearance** models
– **shape**

Related Work

- Can be optimized with gradient descent
  - first order approximation

Ben Ayed et al. Image Processing 2008,
Foulonneau et al., PAMI 2006
Foulonneau et al., IJCV 2009

We use higher-order approximation based on **trust region** approach a general class of

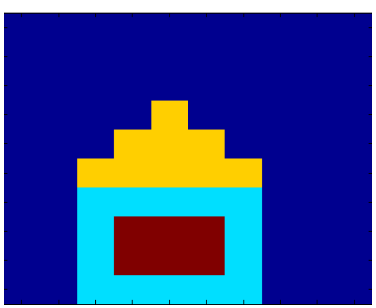
non-linear regional functionals

$$\mathbf{R}(\mathbf{S}) = \mathbf{F}(\langle \mathbf{f}_1, \mathbf{S} \rangle, \cdots, \langle \mathbf{f}_k, \mathbf{S} \rangle)$$

Regional Functional Examples

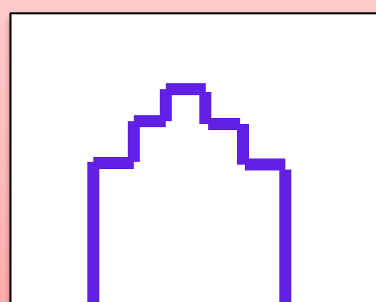
Volume Constraint

$$\mathbf{R(S) = (\langle 1, S \rangle - V_0)^2}$$



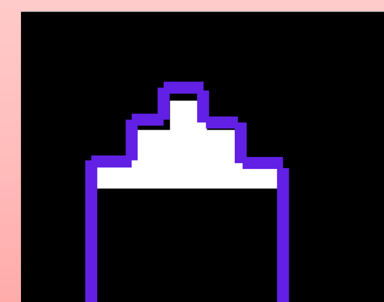
$$\mathbf{|S| = \langle 1, S \rangle}$$

$$\mathbf{f(x) = 1}$$



$$\mathbf{\langle f_i, S \rangle}$$

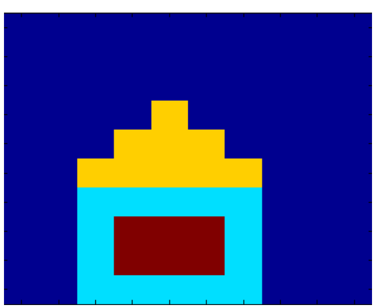
$\mathbf{f_i(x)}$ for bin i



Regional Functional Examples

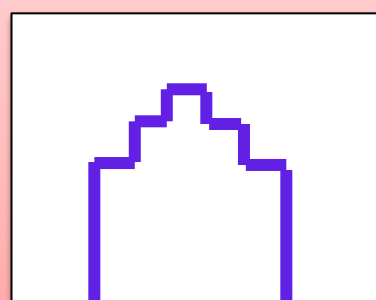
Bin Count Constraint

$$\mathbf{R(S)} = \sum_{i=1}^{k} \left(\langle \mathbf{f_i}, \mathbf{S} \rangle - \mathbf{V_i} \right)^2$$



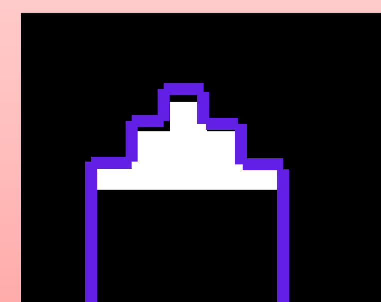
$$|\mathbf{S}| = \langle \mathbf{1}, \mathbf{S} \rangle$$

$$\mathbf{f(x)} = \mathbf{1}$$



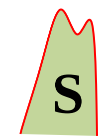
$$\langle \mathbf{f_i}, \mathbf{S} \rangle$$

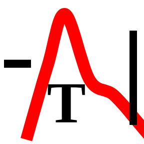
$\mathbf{f_i(x)}$ for bin i



Regional Functional Examples

- Histogram Constraint

$R(S) = \|$ 

$-$ 

$\|_{L^2}$

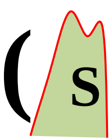
$$R(S) = \sum_{i=1}^{k} \left(P_i(S) - V_i \right)^2$$

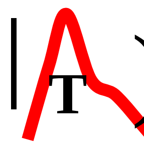
$$P_i(S) = \frac{\langle f_i, S \rangle}{\langle 1, S \rangle}$$

Regional Functional Examples

- Histogram Constraint

$$R(S) = KL(\;S \;\|\; T\;)$$



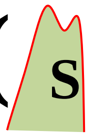


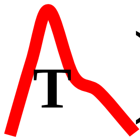
$$R(S) = \sum_{i=1}^{k} P_i(S) \log \frac{P_i(S)}{V_i}$$

$$P_i(S) = \frac{\langle f_i, S \rangle}{\langle 1, S \rangle}$$

Regional Functional Examples

- Histogram Constraint

$\mathbf{R(S) = Bha(}$ 

$\mathbf{,}$ 

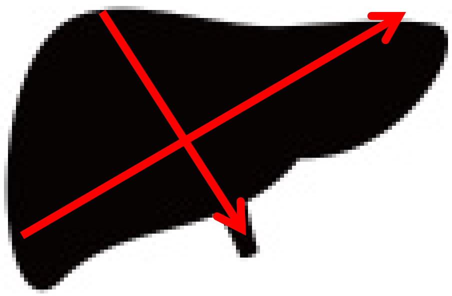
$\mathbf{)}$

$$\mathbf{R(S) = -\log\left(\sum_{i=1}^{k}\sqrt{P_i(S)\cdot V_i}\right)}$$

$$\mathbf{P_i(S) = \frac{\langle f_i, S\rangle}{\langle 1, S\rangle}}$$

Shape Prior

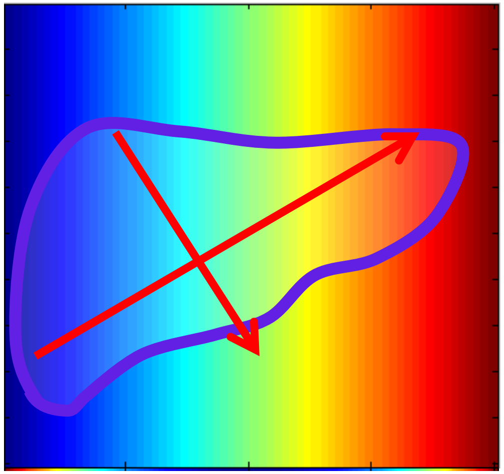
- Volume Constraint is a very crude shape prior



- Can be generalized to constraints for a set of shape moments $\mathbf{m}_{\mathbf{pq}}$

Shape Prior

- Volume Constraint is a very crude shape prior



$$f_{pq}(x,y) = x^{2}y^{3}$$

$$m_{pq}(S) = \langle f_{pq}, S \rangle$$

$$f_{pq}(x,y) = x^{p}y^{q}$$

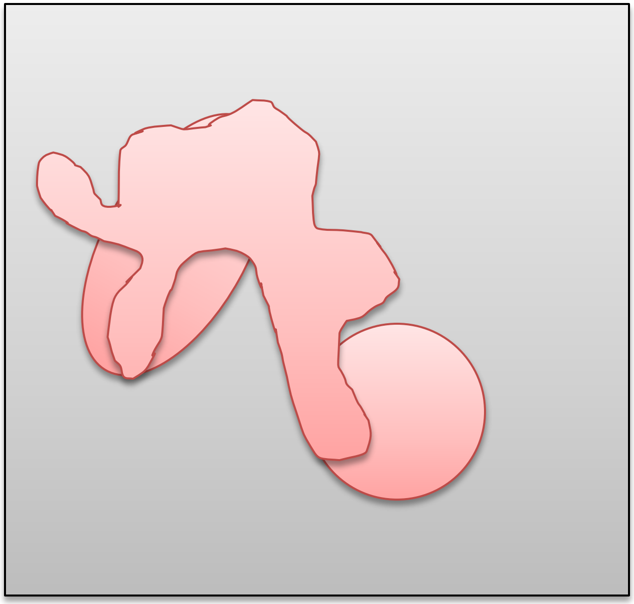
Shape Prior using Shape Moments m_{pq}

$$m_{00} = \text{Volume}$$

$$(m_{10}, m_{01}) = \text{Center Of Mass}$$

$$\begin{pmatrix} m_{20} & m_{11} \\ m_{11} & m_{02} \end{pmatrix} = \begin{matrix} \text{Principal Orientation} \\ \text{Aspect Ratio} \end{matrix}$$

…



Shape Prior using Shape moments

- Shape Prior Constraint $\mathbf{R(S)} = \text{Dist}(\mathbf{S}, \mathbf{T})$

$$\mathbf{R(S)} = \sum_{\mathbf{p+q} \leq k} (\mathbf{m_{pq}(S)} - \mathbf{m_{pq}(T)})^2$$

$$\mathbf{m_{pq}(S)} = \langle \mathbf{f_{pq}}, S \rangle$$

$$\mathbf{f_{pq}(x,y)} = \mathbf{x^p y^q}$$

Optimization of Energies with Higher-order Regional Functionals

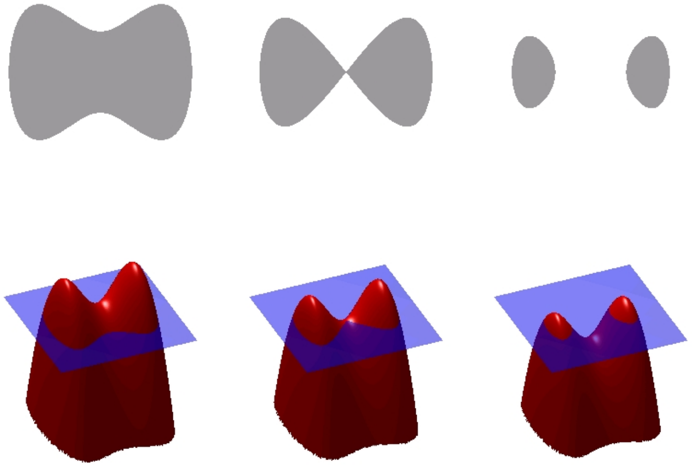
$$\mathbf{E}(\mathbf{S}) = \mathbf{R}(\mathbf{S}) + \mathbf{B}(\mathbf{S})$$

Gradient Descent (e.g. level sets)

$$\mathbf{E(S)} = \mathbf{R(S)} + \mathbf{B(S)}$$

- Gradient Descent
- First Order Taylor Approximation for R(S)
- First Order approximation for B(S) (“curvature flow”)
- Robust with tiny steps
  - Slow
  - Sensitive to initialization

Ben Ayed et al. CVPR 2010,
Freedman et al. tPAMI 2004



http://en.wikipedia.org/wiki/File:Level_set_method.jpg

Energy Specific vs. General

- Speedup via energy- specific methods
  - Bhattacharyya Distance
  - Volume Constraint

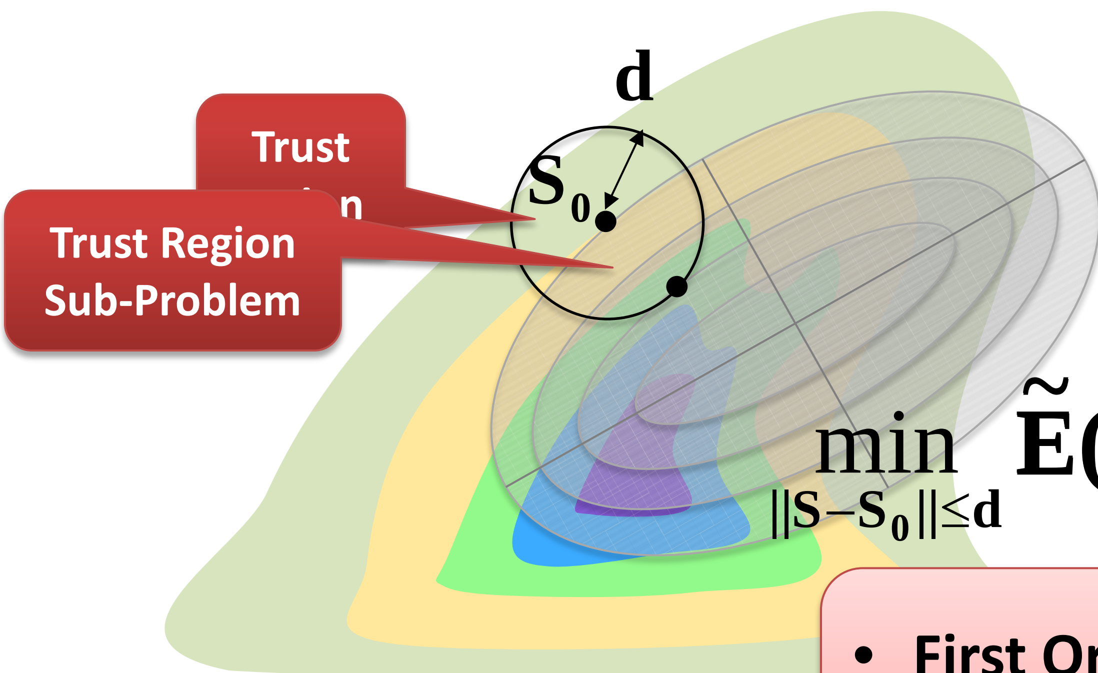
Ben Ayed et al. CVPR 2010,
Werner, CVPR2008
Woodford, ICCV2009

- We propose
  - trust region optimization algorithm for **general** high-order energies
  - **higher-order** (non-linear) approximation

General Trust Region Approach
An overview

- The goal is to optimize

$$\mathbf{E(S) = R(S) + B(S)}$$



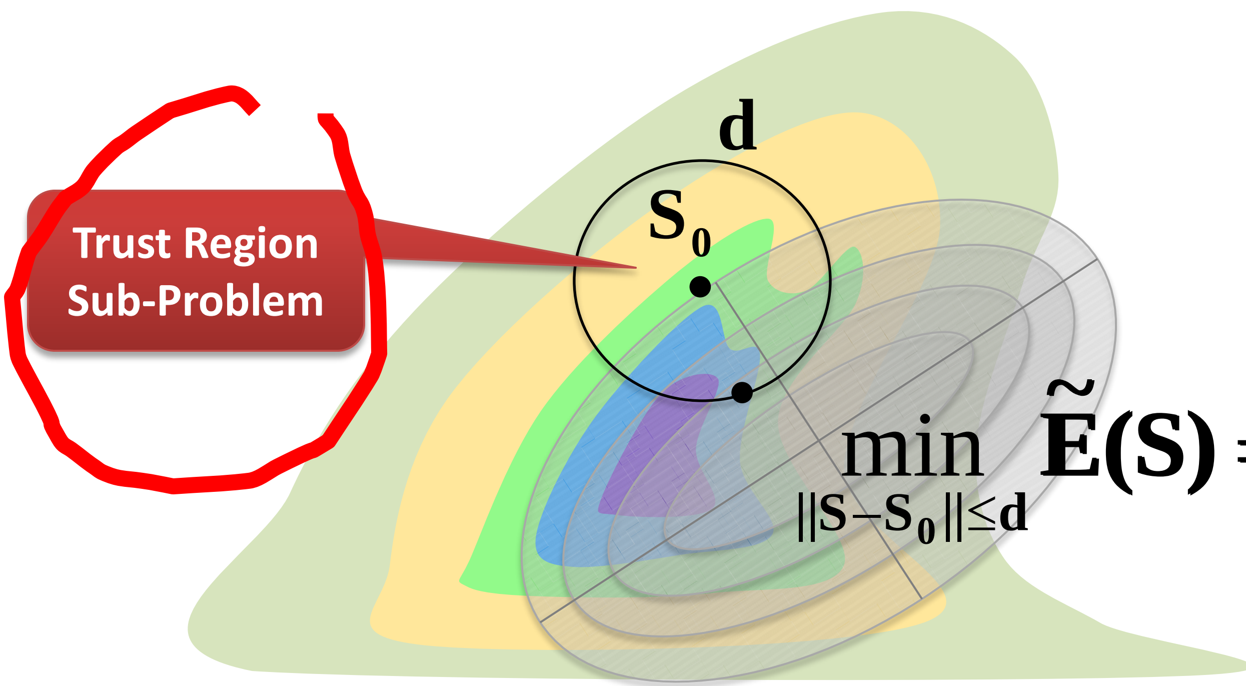
$$\min_{\|S - S_0\| \le d} \tilde{\mathbf{E}}(\mathbf{S}) = \mathbf{U_0(S) + B(S)}$$

- **First Order Taylor for R(S)**
- **Keep quadratic B(S)**

General Trust Region Approach
An overview

- The goal is to optimize

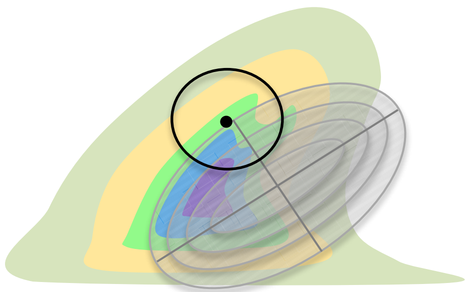
$$\mathbf{E(S) = R(S) + B(S)}$$



$$\min_{\|\mathbf{S}-\mathbf{S}_0\| \le \mathbf{d}} \ \tilde{\mathbf{E}}(\mathbf{S}) = \mathbf{U}_0(\mathbf{S}) + \mathbf{B}(\mathbf{S})$$

Solving Trust Region Sub-Problem

- Constrained optimization



$$\text{minimize} \quad \tilde{\mathbf{E}}(\mathbf{S}) = \mathbf{U_0}(\mathbf{S}) + \mathbf{B}(\mathbf{S})$$
$$\mathbf{s.t.} \; \| \mathbf{S} - \mathbf{S_0} \| \leq \mathbf{d}$$

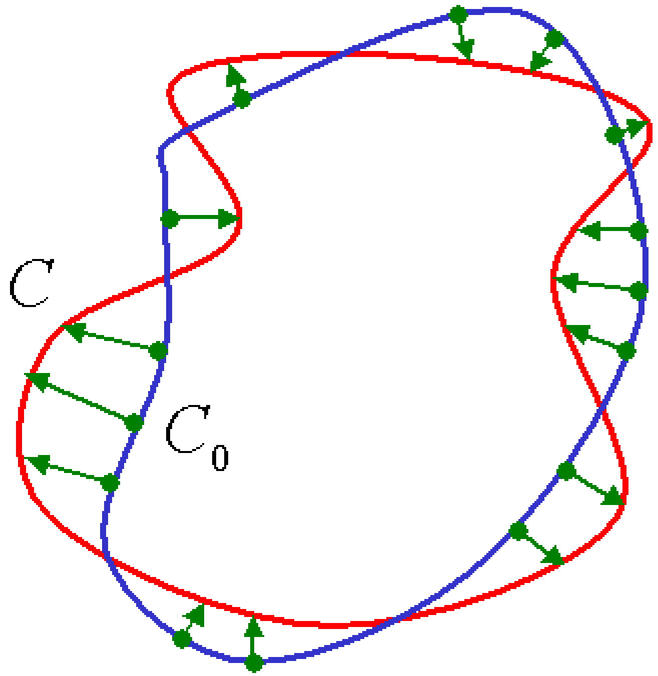
- Unconstrained Lagrangian Formulation

$$\text{minimize} \quad \mathbf{L}_\lambda(\mathbf{S}) = \mathbf{U_0}(\mathbf{S}) + \mathbf{B}(\mathbf{S}) + \lambda \| \mathbf{S} - \mathbf{S_0} \|$$

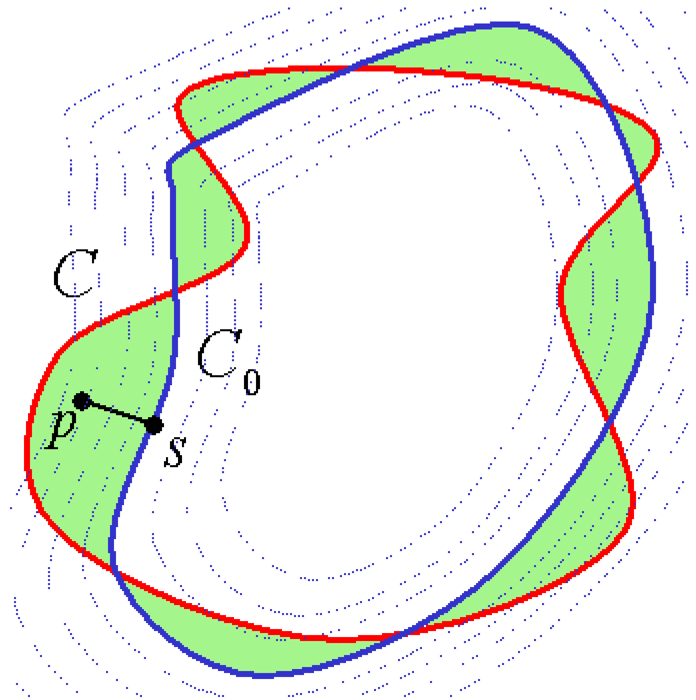
L_2 distance can be approximated with unary terms [Boykov, Kolmogorov, Cremers, Delong, ECCV'06]

- Can be optimized **globally** using graph-cut or convex continuous formulation

Approximating distance $||S - S_0||$



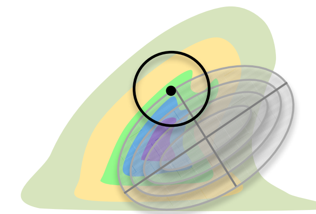
$$\langle dC, dC \rangle = \int_{C_0} dC_s^2 \cdot ds$$



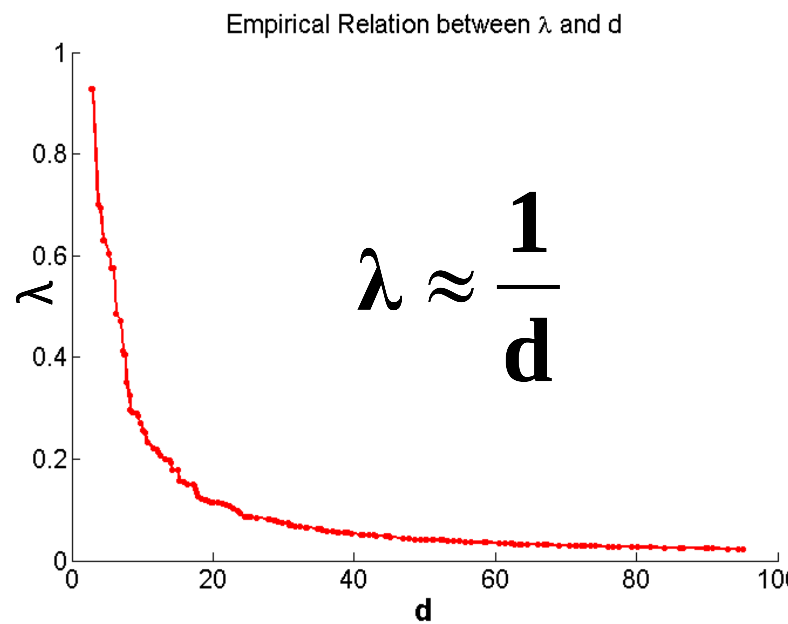
$$dist(C, C_0) = 2 \int_{\Delta C} d_0(p) \cdot dp$$

[BKCD – ECCV 2006]

Trust Region

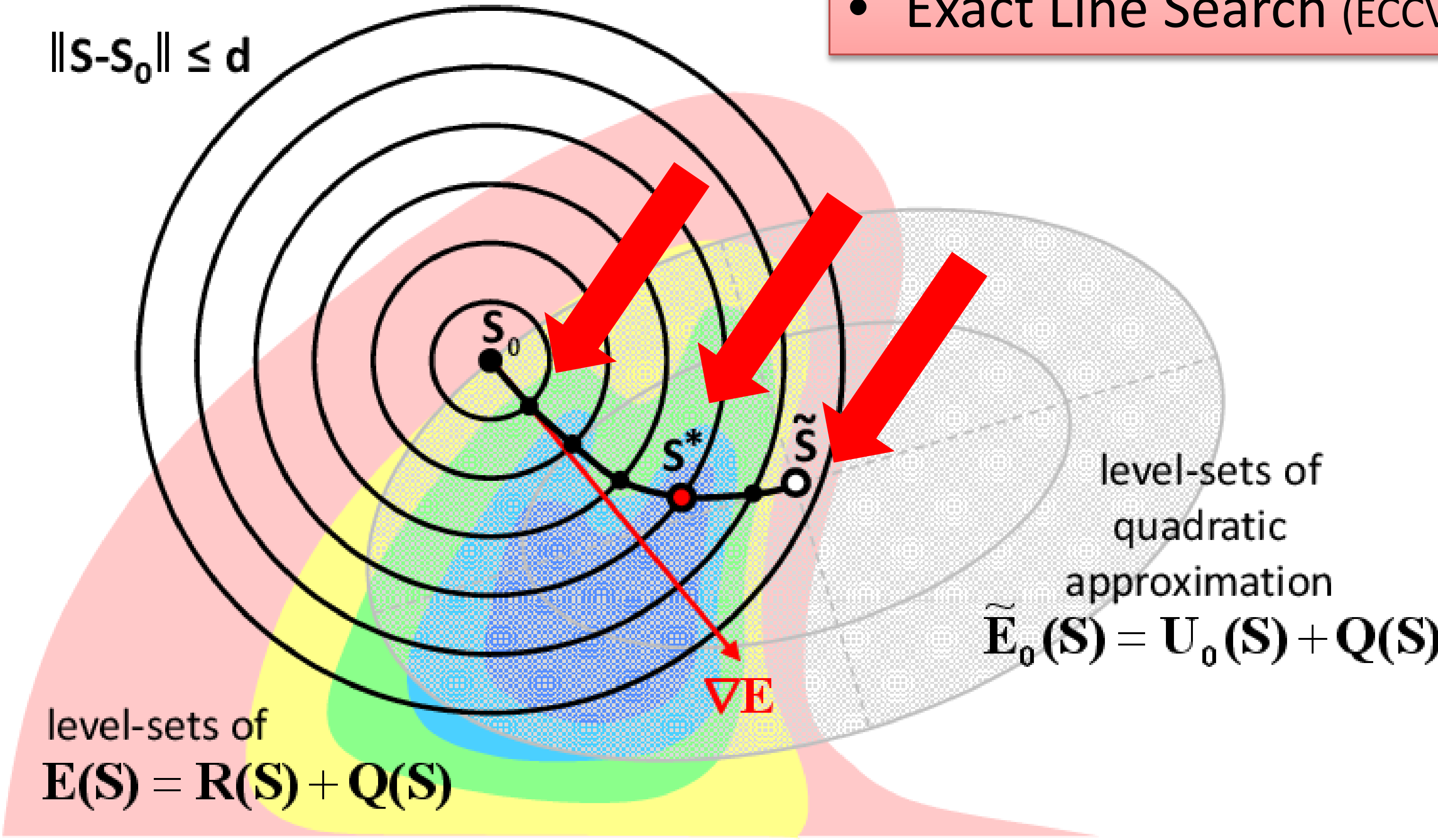


- Standard (adaptive) Trust Region
  - Control of step size **d**
- Lagrangian Formulation
  - Control of the Lagrange multiplier **λ**



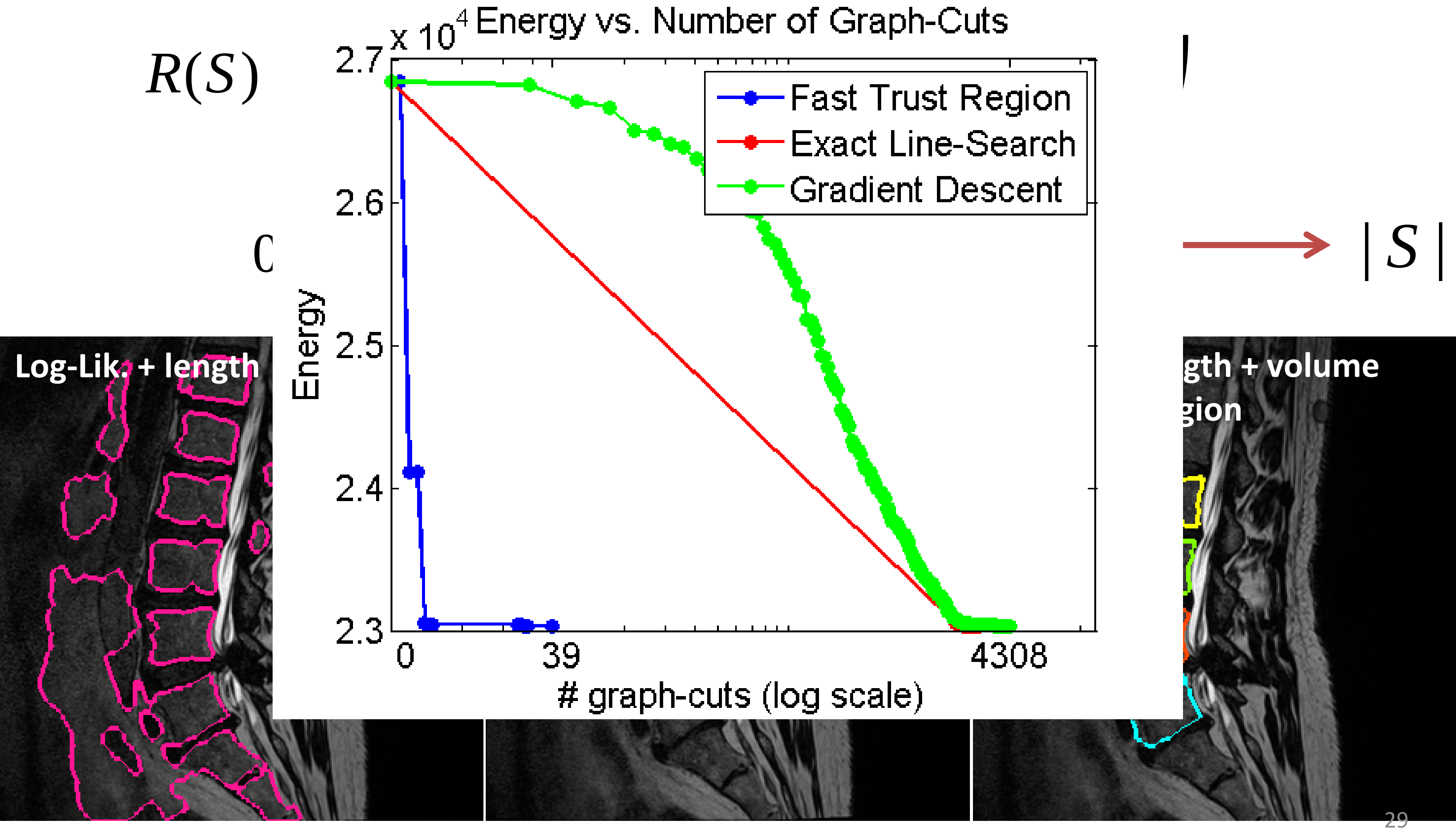
Spectrum of Solutions for different λ or d

- Newton step
- "Gradient Descent"
- Exact Line Search (ECCV12)



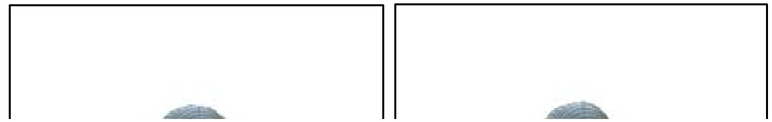
Volume Constraint for Vertebrae segmentation

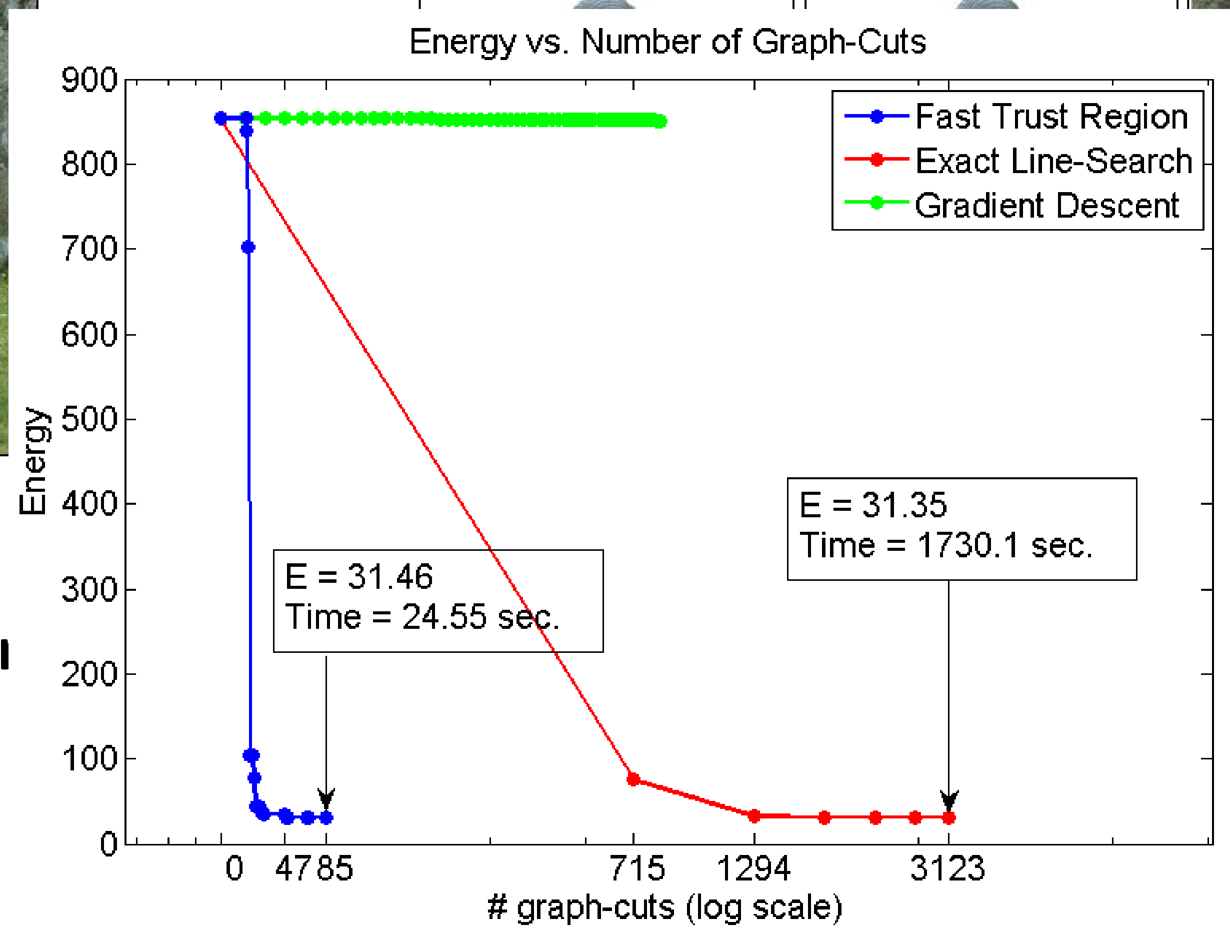
$R(S)$



$|S|$

Appearance model with KL Divergence Constraint

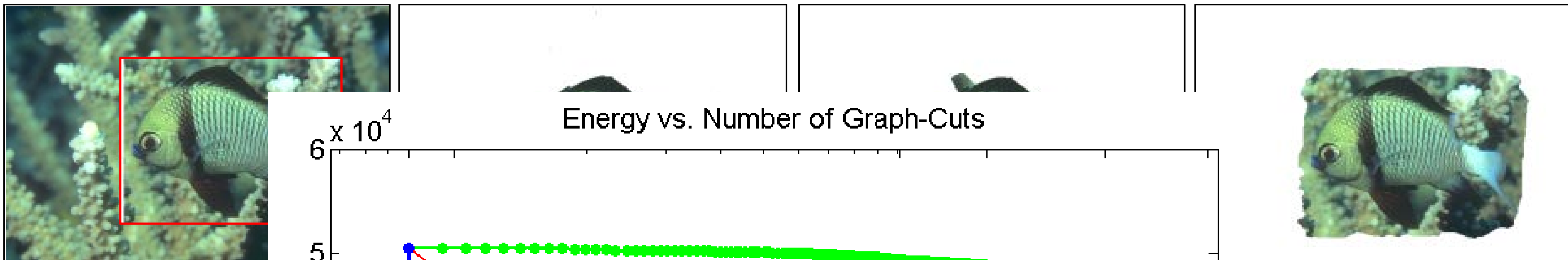




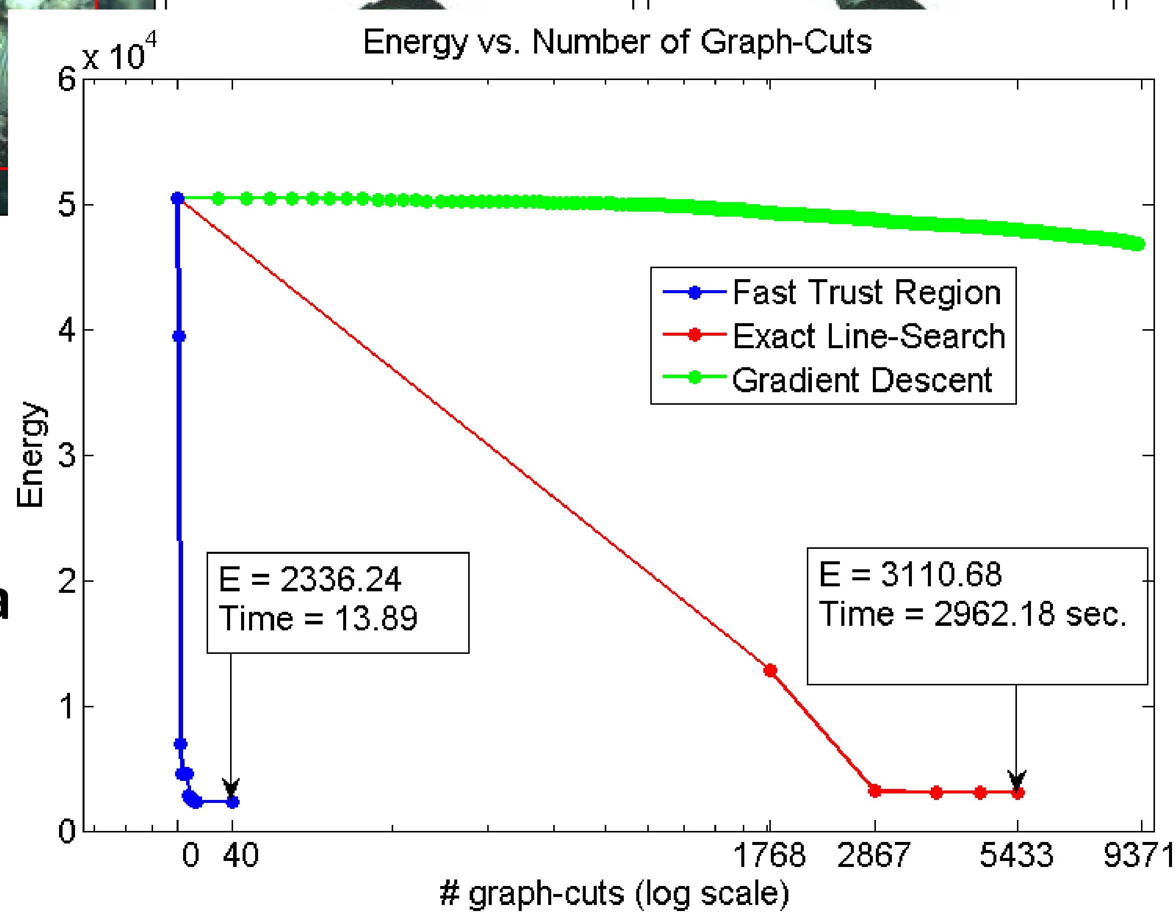
adient Descent”

Appearai

Appearance Model with Bhattacharyya Distance Constraint

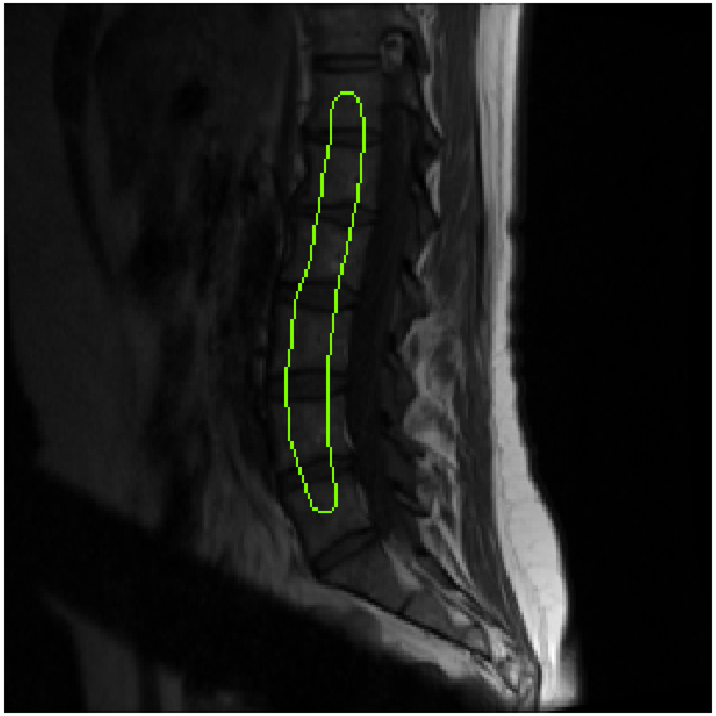


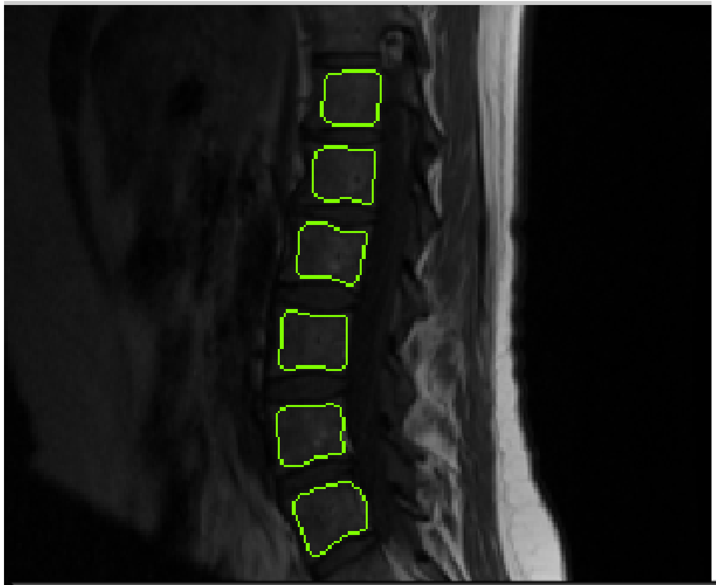
“Gradient Descent”



Appeara **h**

Shape prior with Tchebyshev moments for spine segmentation





Log-Lik. + length + Shape Prior
Fast Trust Region

Second order Tchebyshev moments computed for the user scribble

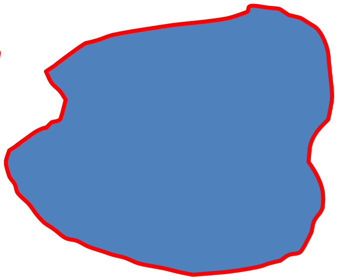
Part 2

Complexity constraints on appearance

Segment appearance ?

- sum of log-likelihoods (linear appearance)
  - ~~histograms~~
  - ~~mixture models~~

assumes i.i.d. pixels



- now:

allow sub-regions (n-labels)

with <u>constraints</u>

  - based on information theory (MDL complexity)
  - based on geometry (anatomy, scene layout)

Natural Images: GMM or MRF?

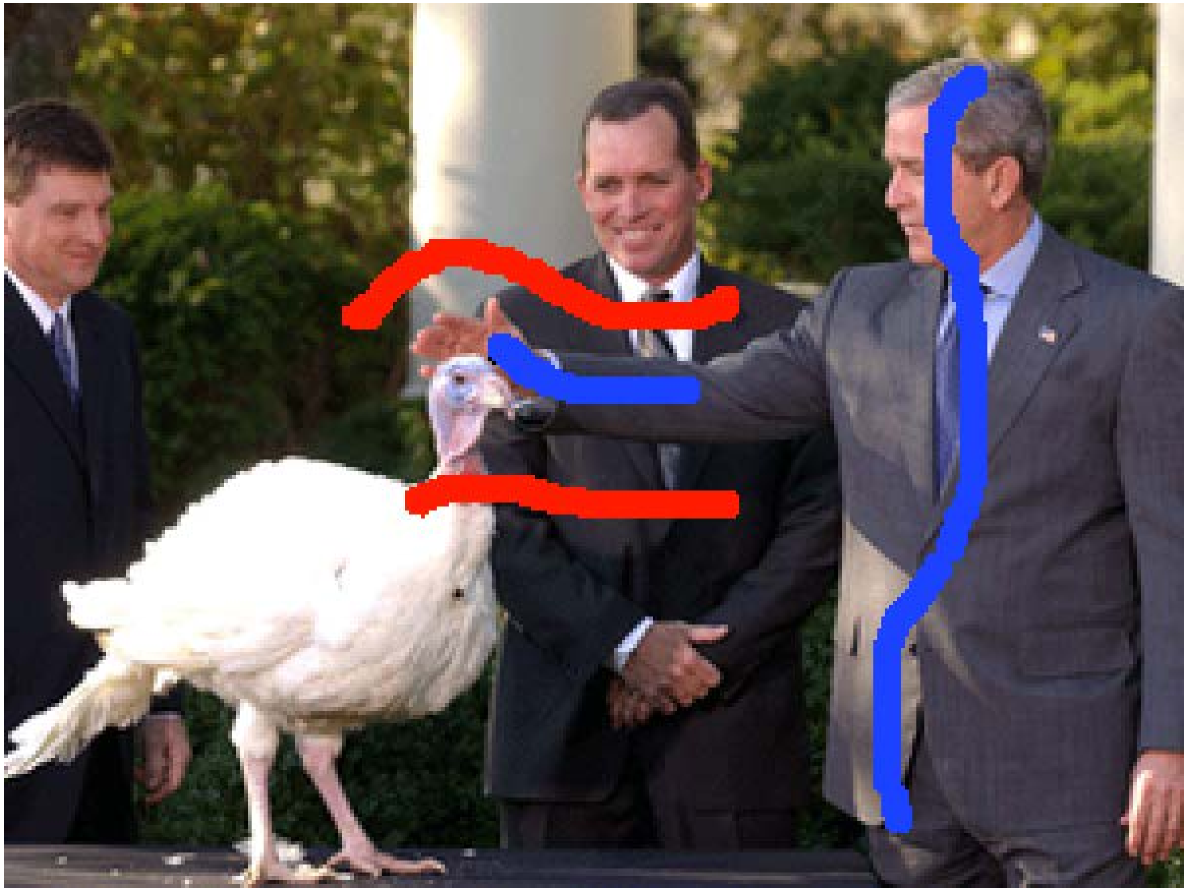
are pixels in this image i.i.d.? NO!

Natural Images: GMM or MRF?

Natural Images: GMM or MRF?

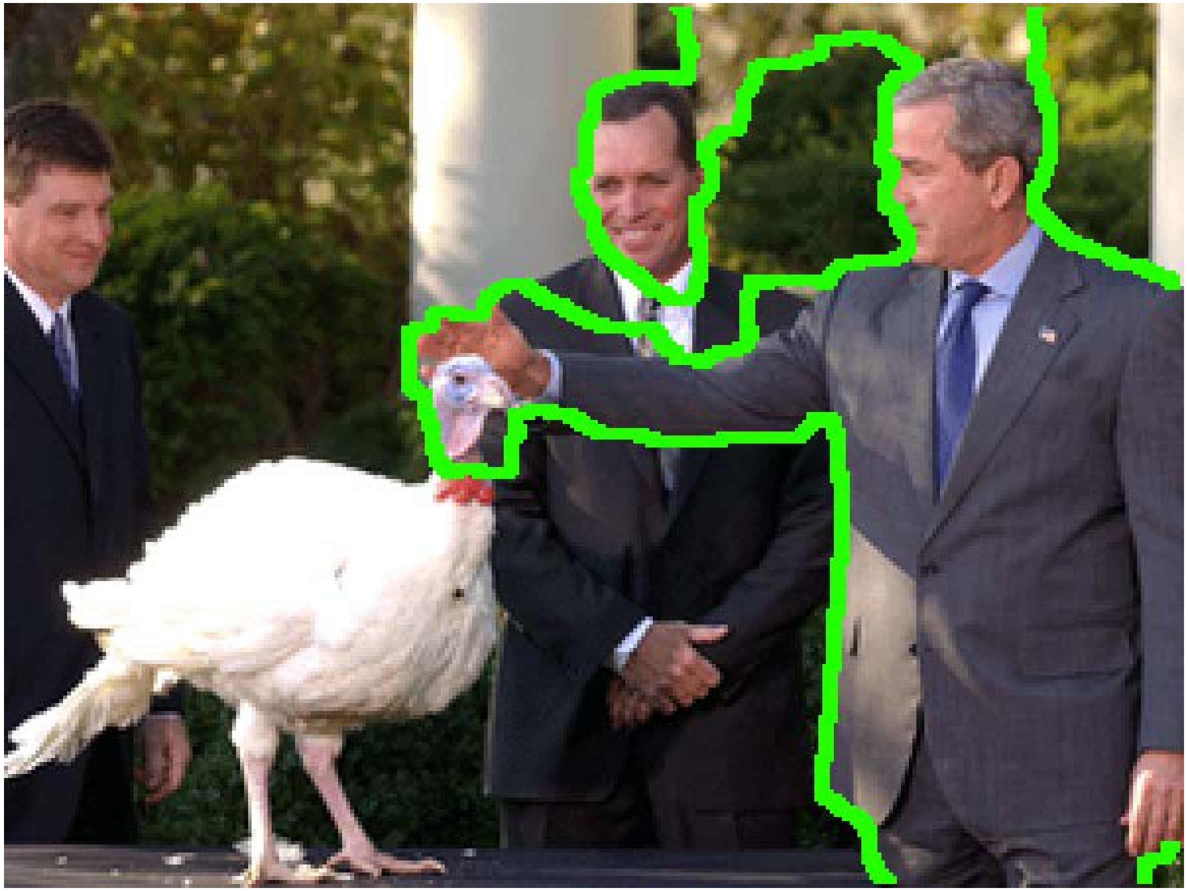
Natural Images: GMM or MRF?

Binary graph cuts



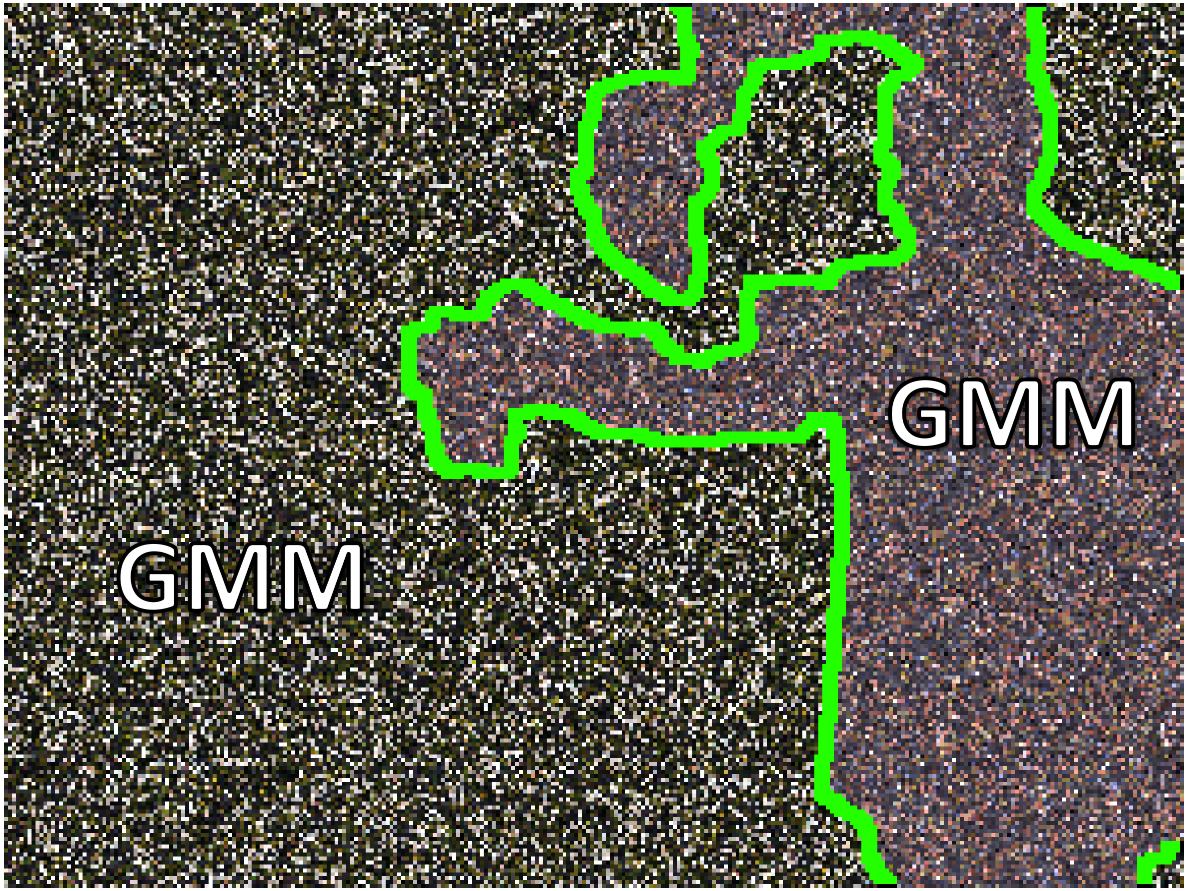
[Boykov & Jolly, *ICCV* 2001] [Rother, Kolmogorov, Blake, *SIGGRAPH* 2004]

Binary graph cuts



[Boykov & Jolly, *ICCV* 2001] [Rother, Kolmogorov, Blake, *SIGGRAPH* 2004]

Binary graph cuts

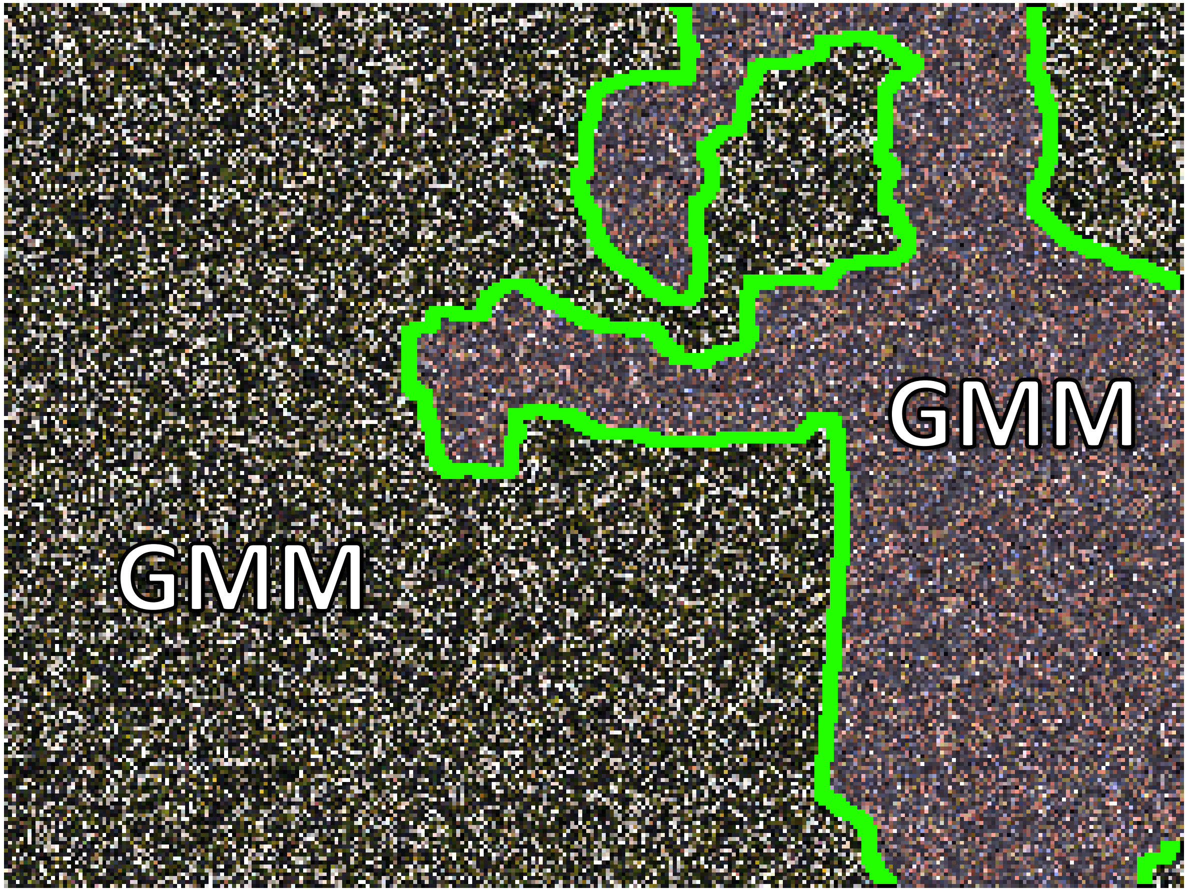


[Boykov & Jolly, *ICCV* 2001] [Rother, Kolmogorov, Blake, *SIGGRAPH* 2004]

A Spectrum of Complexity

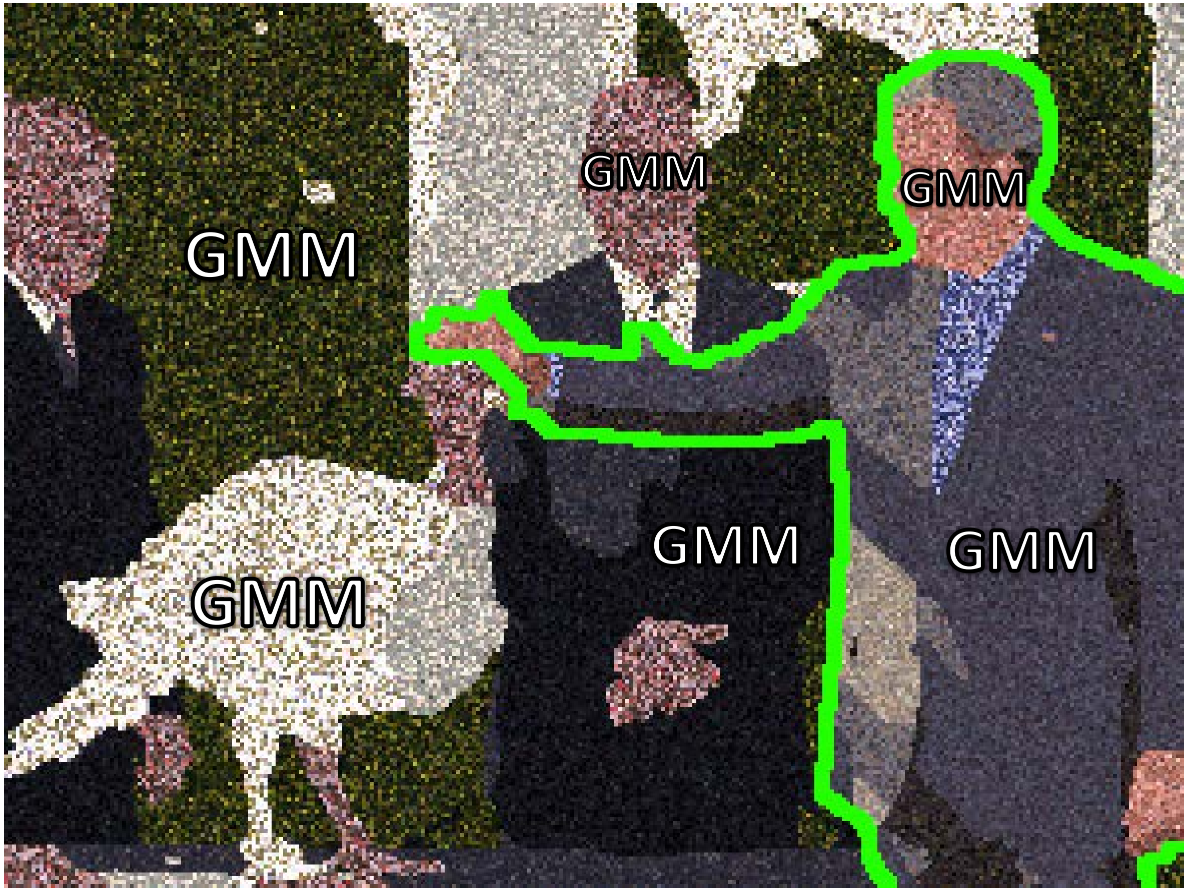
- Objects *within* image can be as complex as image itself
- Where do we draw the line?

Single Model Per Class Label



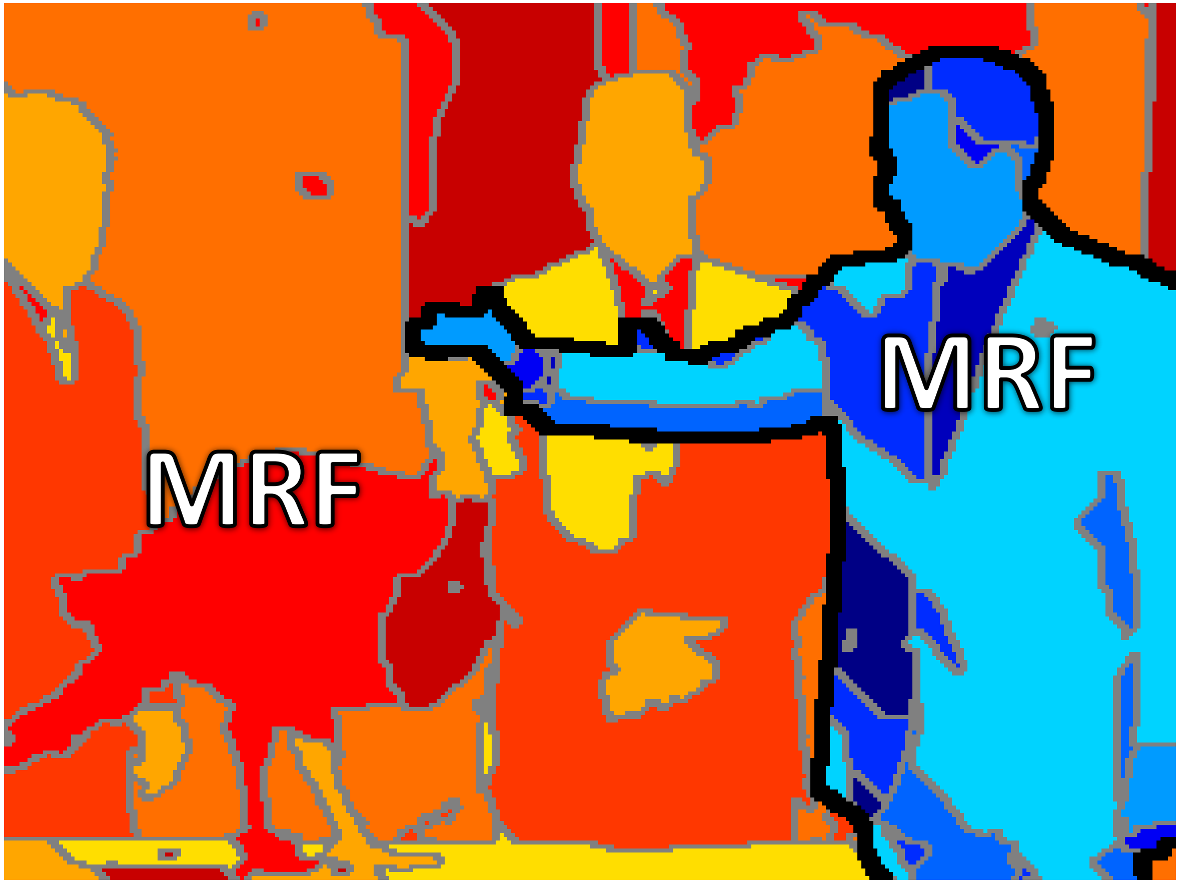
Pixels are **identically distributed** inside each segment

Multiple Models *Per Class Label*



Now pixels are **not identically distributed** inside each segment

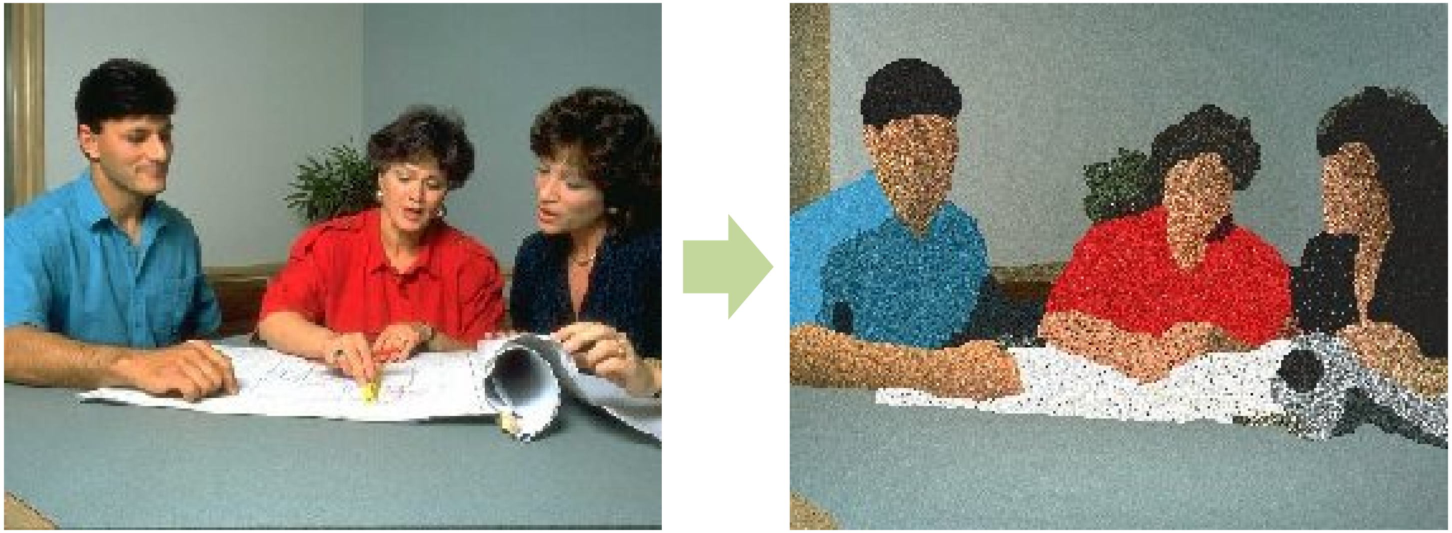
Multiple Models Per Class Label



Our Energy ≈ Supervised Zhu & Yuille!

- Leclerc, *PAMI'92*
- Zhu & Yuille. *PAMI*'96; Tu & Zhu. *PAMI*'02
- Unsupervised clustering of pixels

$$\min_{labeling} \left\{ \text{color similarity} + \text{boundary length} + \text{MDL regularizer} \right\}$$

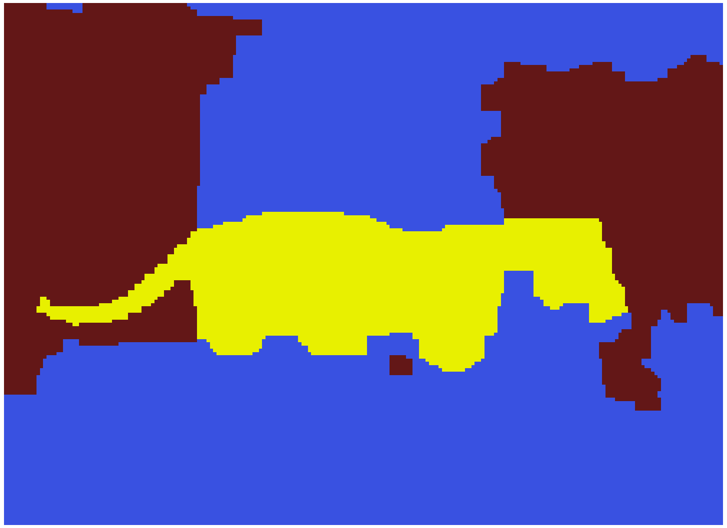


α-expansion (graph cuts) can handle such energies with some optimality guarantees [IJCV'12,CVPR'10]

(unsupervised image segmentation)

Fitting color models

each label *L* represents some distribution $Pr(I|L)$



$$-\ln \Pr(I_p \mid L_p)$$

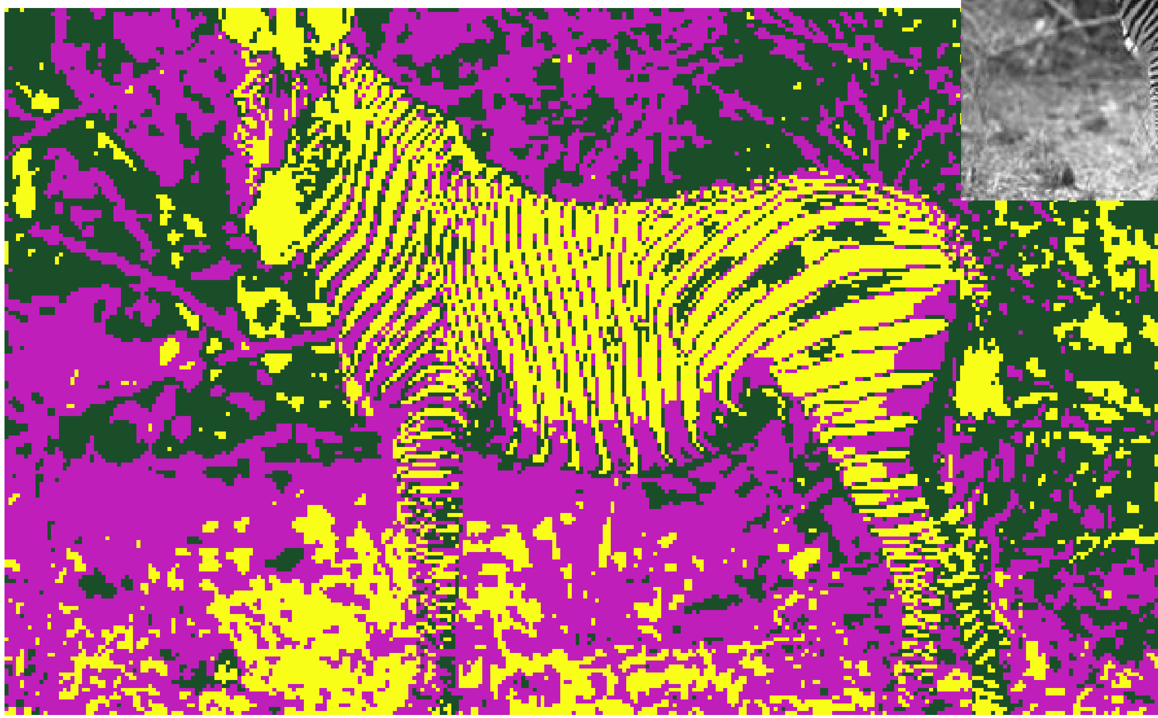
$$E_I(\mathbf{L}) = \sum_p \| p - L_p \| + \sum_{(p,q)\in N} w \cdot \delta(L_p \neq L_q) \;\; + \sum_{L\in\Lambda} h_L \cdot \delta_L(\mathbf{L})$$

information theory (MDL) interpretation:

= number of bits to compress image *I* **losslessly**

(unsupervised image segmentation)

Fitting color models



Label costs only

Delong, Osokin, Isack, Boykov, IJCV 12   (UFL-approach)

(unsupervised image segmentation)

Fitting color models

Spatial smoothness only [Zabih & Kolmogorov, CVPR 04]

(unsupervised image segmentation)

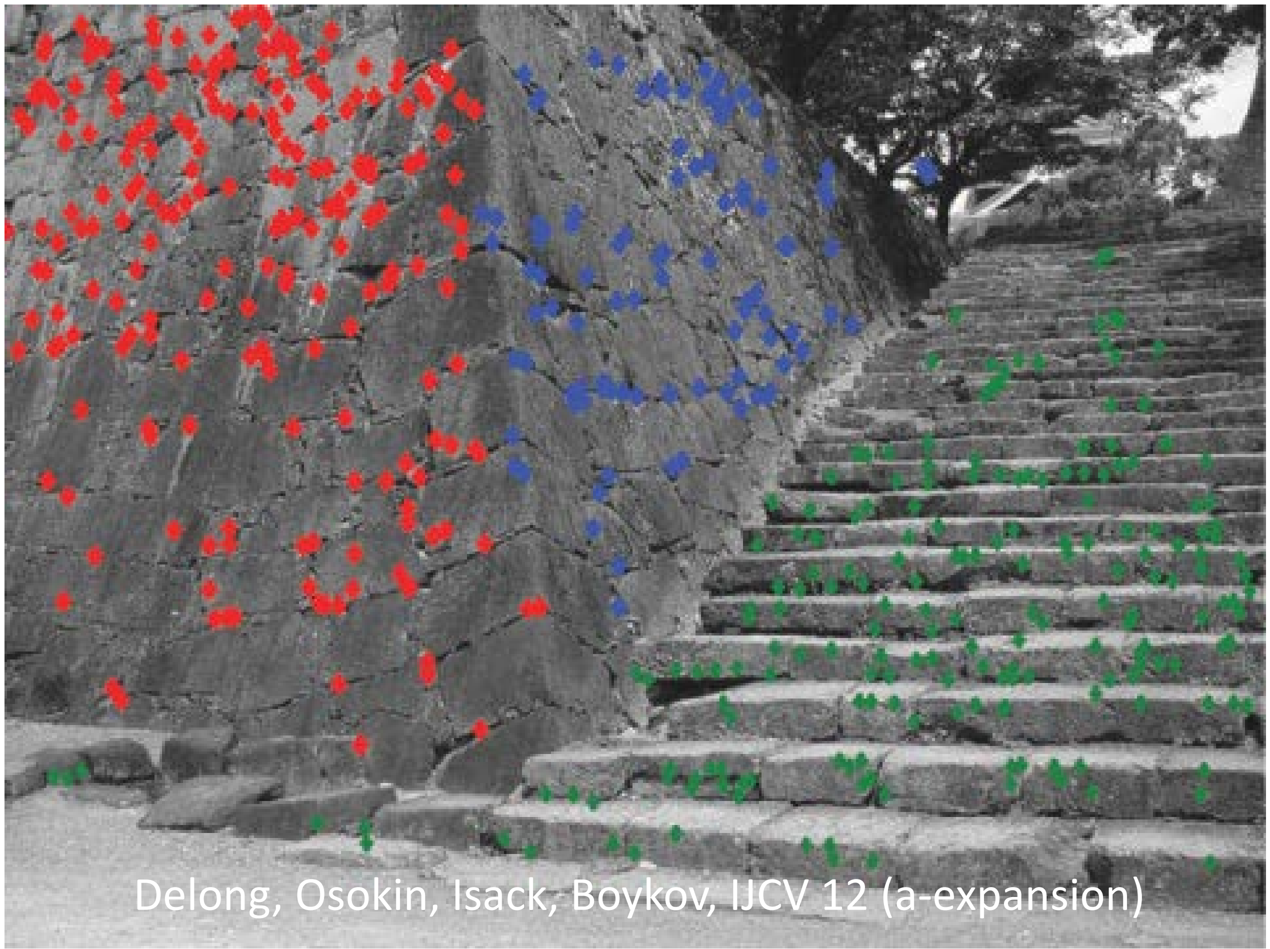
Fitting color models

Spatial smoothness + label costs

Zhu & Yuille, PAMI 1996 (gradient descent)
Delong, Osokin, Isack, Boykov, IJCV 12 (a-expansion)

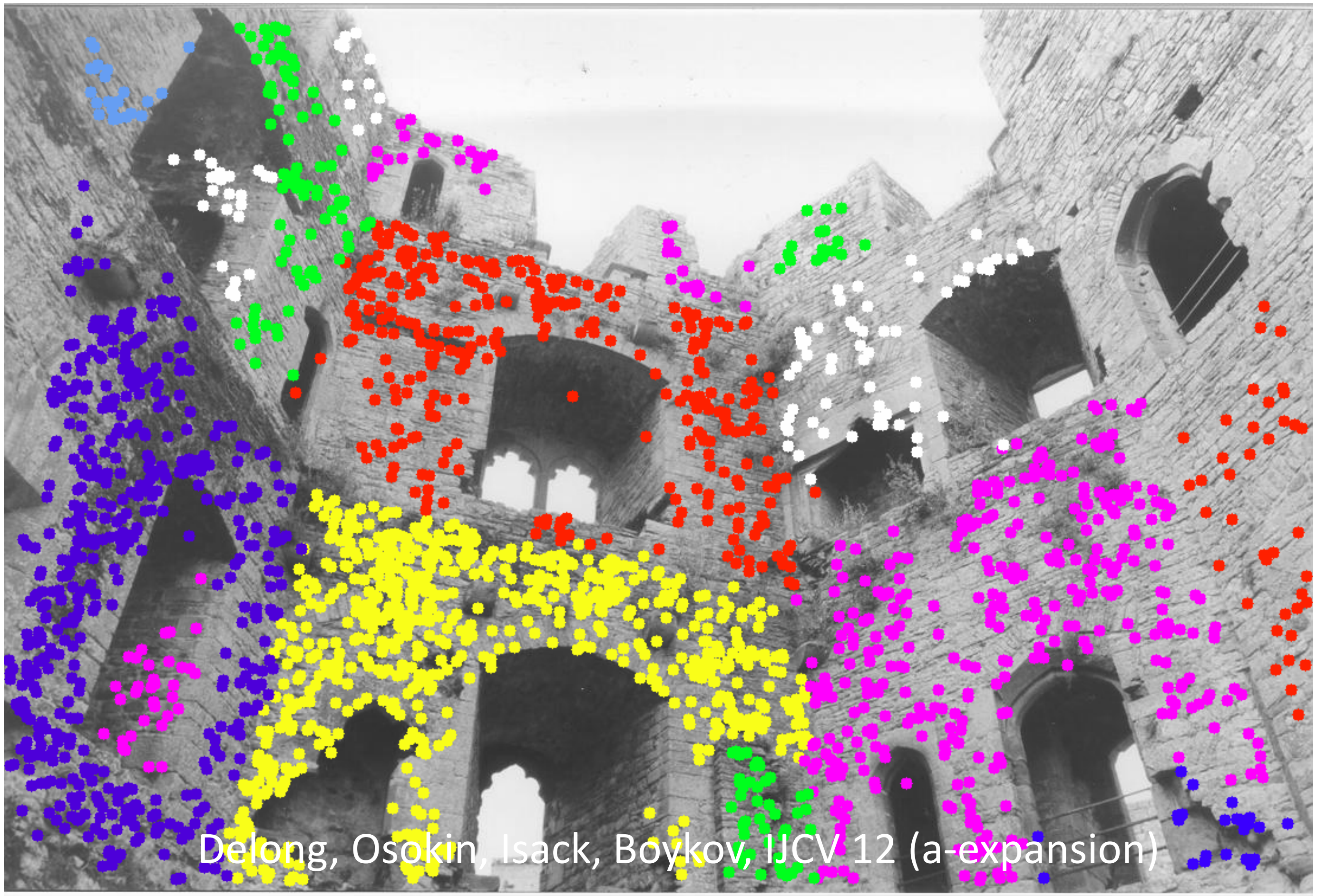
Fitting planes (homographies)

$$E(\boldsymbol{L}) = \sum_{p} \| p - L_p \| + \sum_{(p,q)\in N} w \cdot T(L_p \neq L_q) \quad + \sum_{L\in\Lambda} h_L \cdot \delta_L(\boldsymbol{L})$$

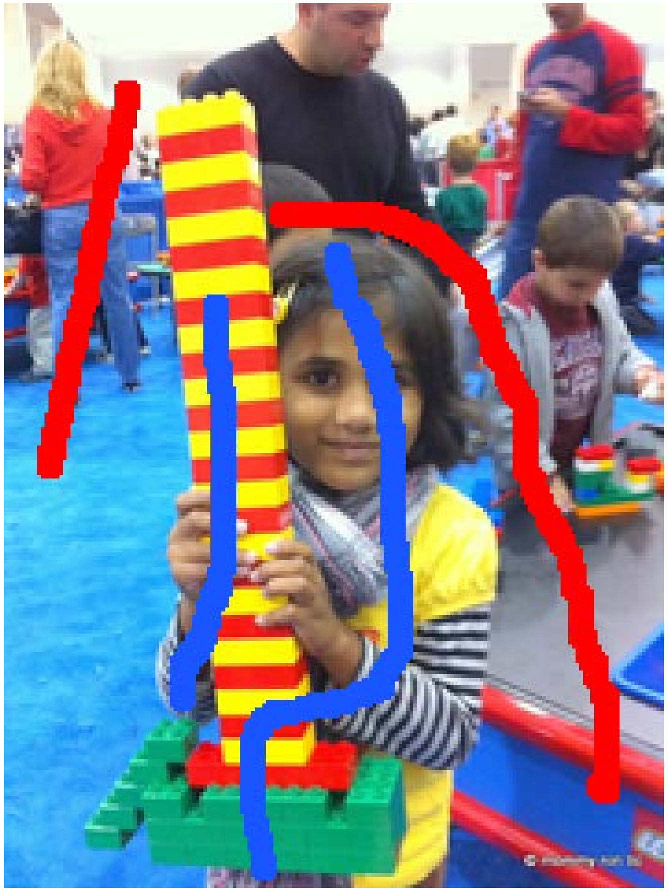


Fitting planes (homographies)

$$E(\mathbf{L}) = \sum_{p} \| p - L_p \| + \sum_{(p,q)\in N} w \cdot T(L_p \neq L_q) \quad + \sum_{L\in\Lambda} h_L \cdot \delta_L(\mathbf{L})$$



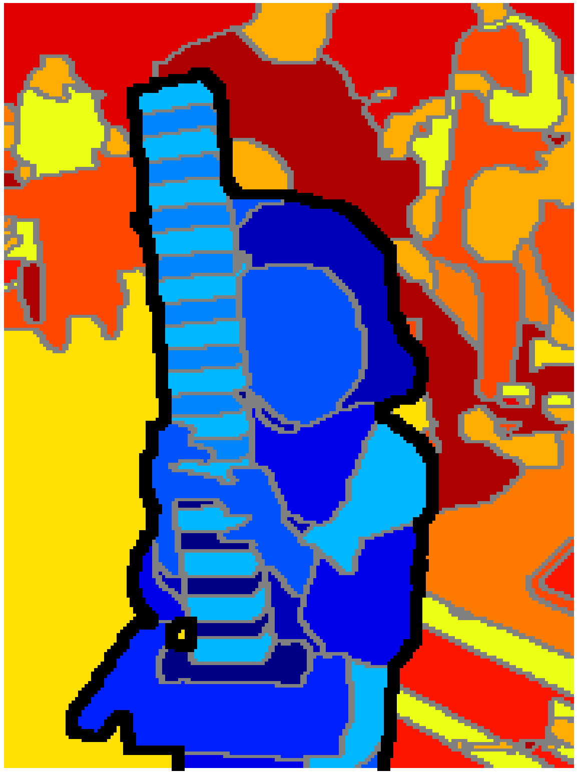
Back to interactive segmentation



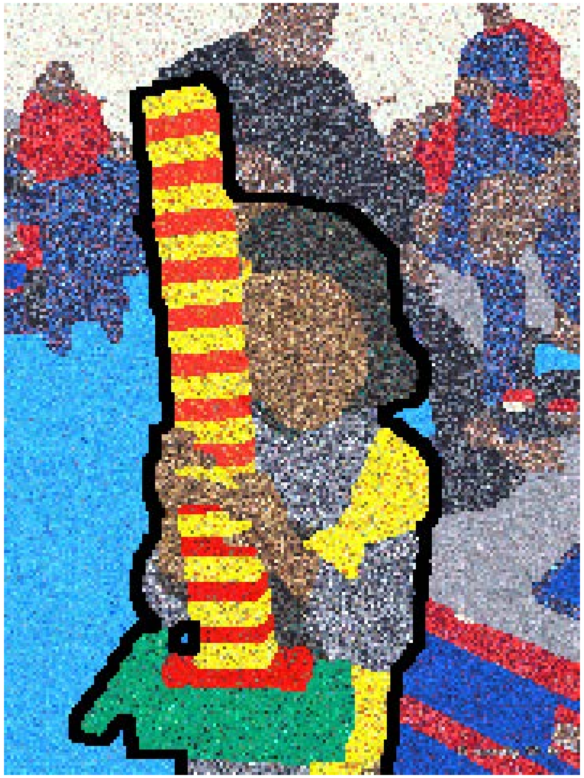
EMMCVPR 2011

segmentation

"sub-labeling"

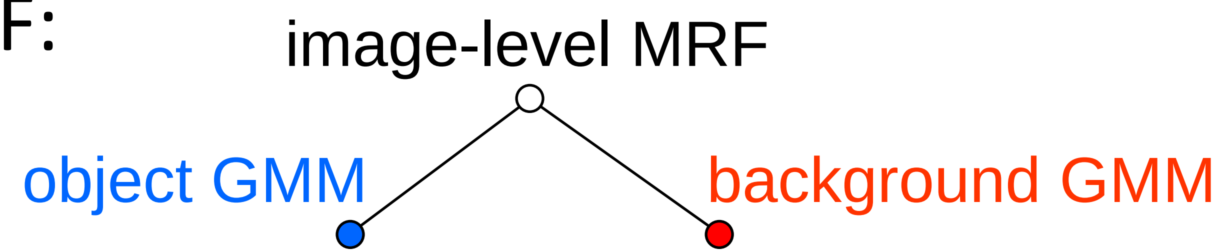


colour models

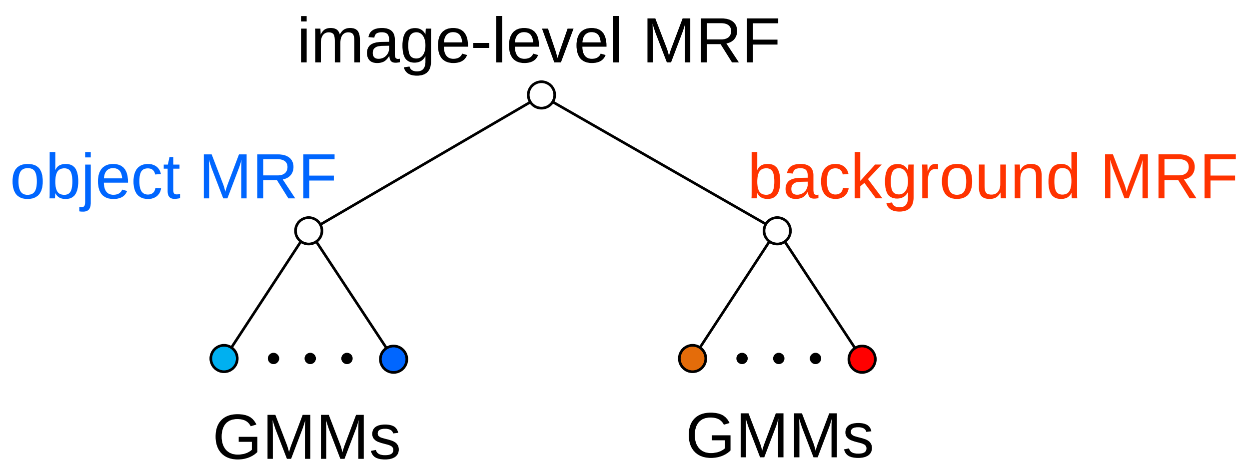


Main Idea

- Standard MRF:



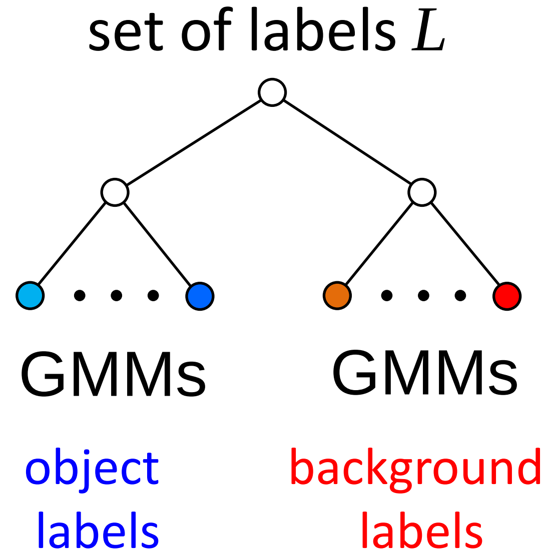
- Two-level MRF:



unknown number of labels in each group!

Our multi-label energy functional

$$E = \underbrace{\sum_{p\in P} D_p(f_p)}_{\text{data costs}} + \underbrace{\sum_{pq\in N} w_{pq}\cdot c(f_p, f_q)}_{\text{smooth costs}} + \underbrace{\sum_{l\in L} h_l \cdot I(\exists p : f_p = l)}_{\text{label costs}}$$



- Penalizes number of GMMs (labels)
  - prefer fewer, simpler models
  - MDL / information criterion regularize "unsupervised" aspect
- Discontinuity cost c is higher between labels of different categories

two categories of labels (respecting hard-constraints)

More Examples

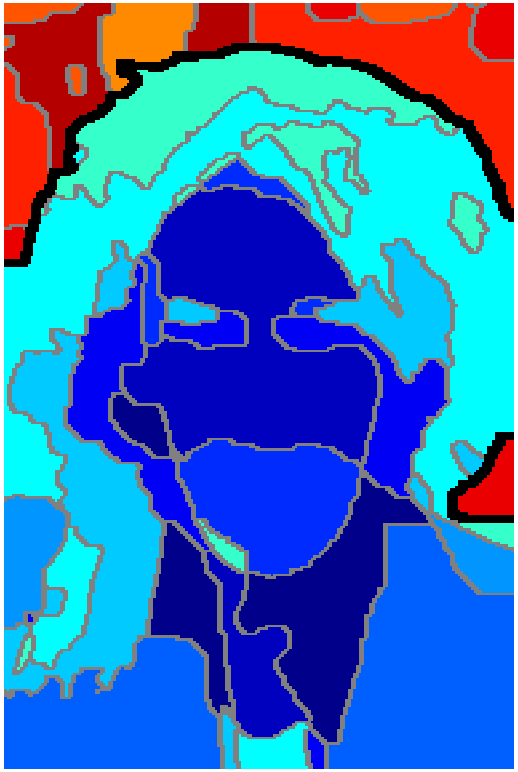
standard 1-level MRF

2-level MRF

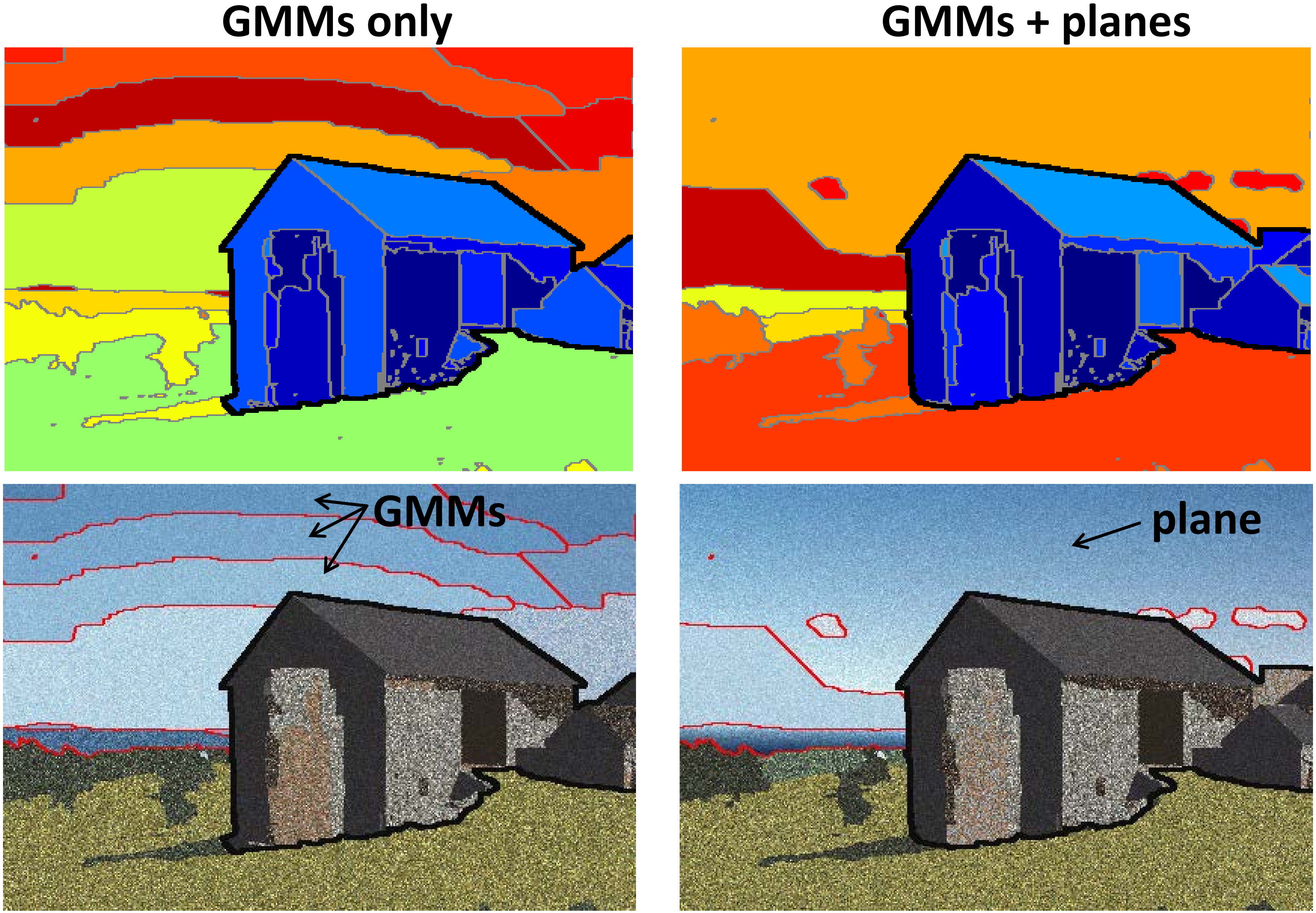
More Examples

standard 1-level MRF

2-level MRF



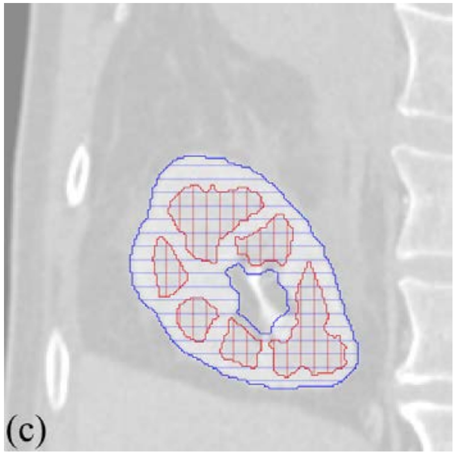
Beyond GMMs



Part 3

Geometric constraints on sub-segments

Towards biomedical image segmentation…



- Sub-labels maybe known apriori
  - known parts of organs or cells
  - interactivity becomes optional

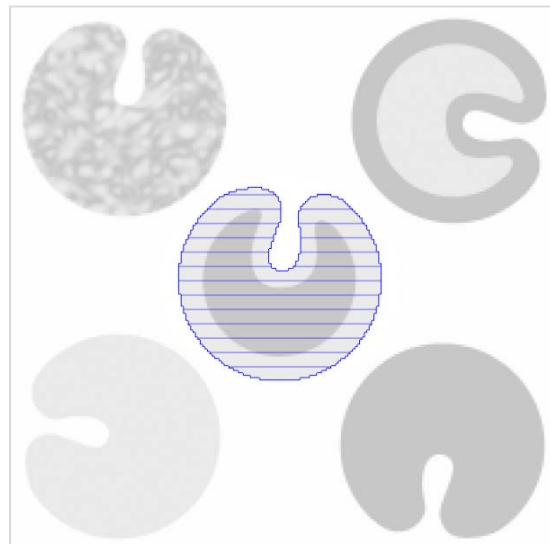
- Geometric constraints should be added
  - human anatomy (medical imaging)
  - or known scene layout (computer vision)

Illustrative Example

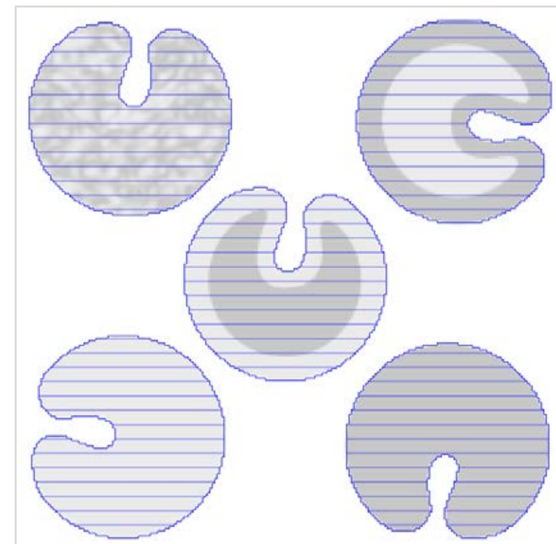
- We want to distinguish between these objects

input

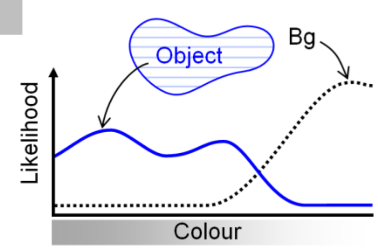
what we want



what mixed model gives



mixed appearance model:

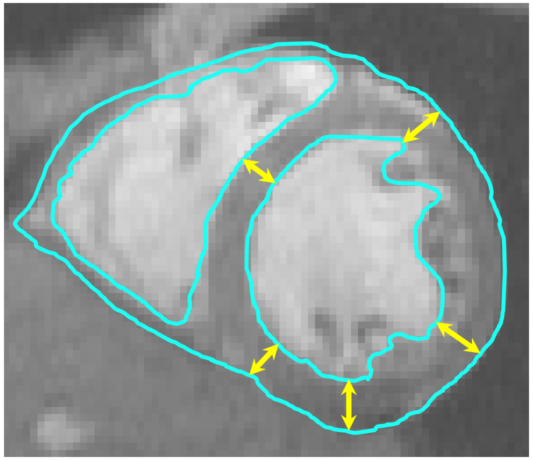


main ideas

– sub-labels with distinct appearance models (as earlier)

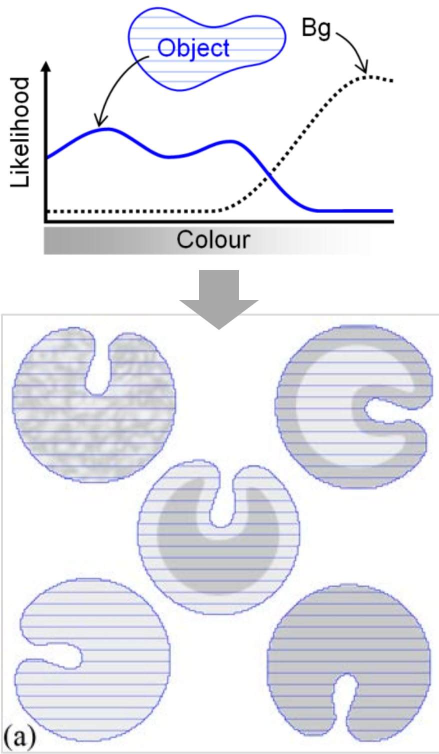
– basic geometric constraints between parts
  - inclusion/exclusion
  - expected distances and margins

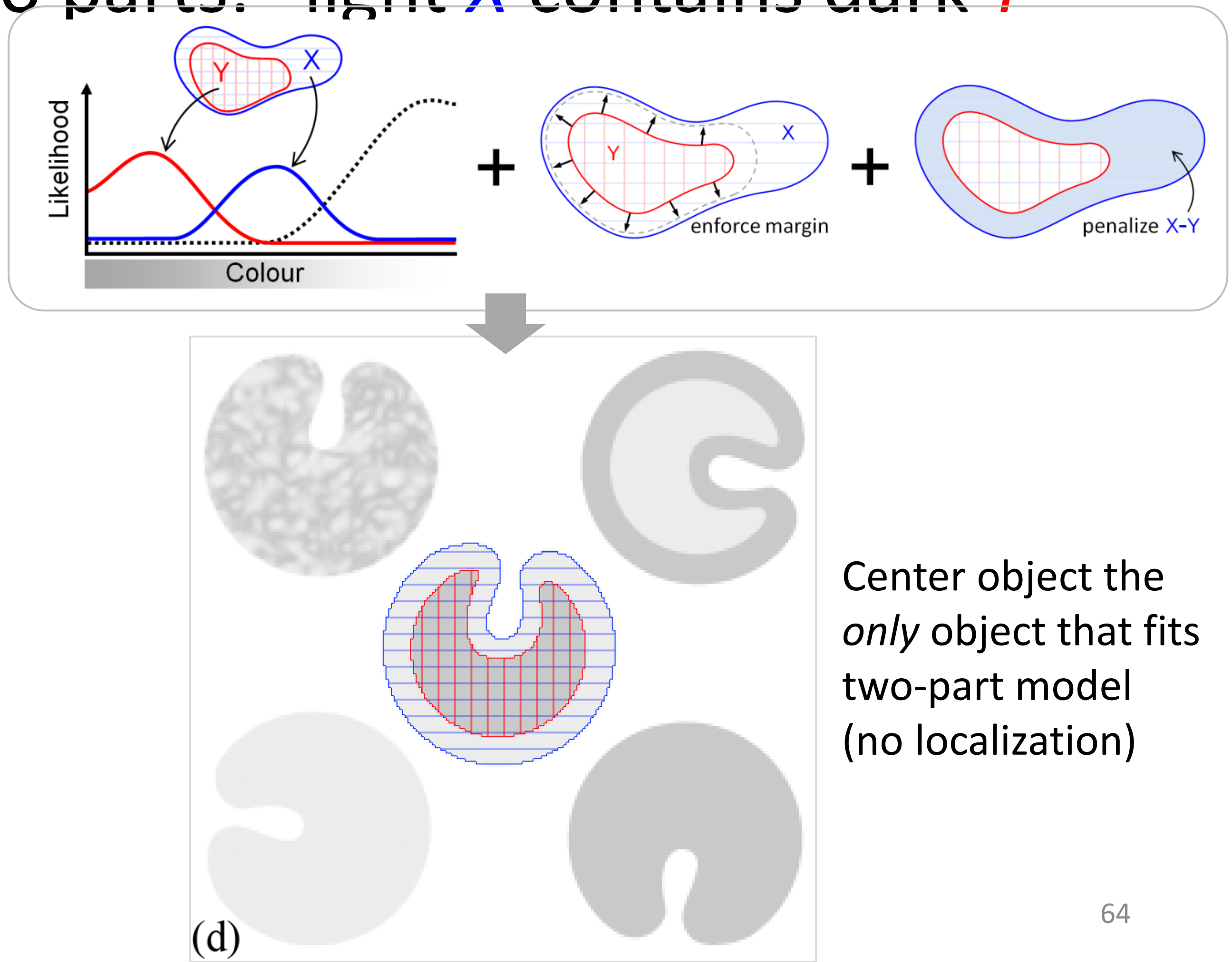
For some constraints globally optimal segmentation can be computed in polynomial time



Illustrative Example

- Model as two parts: "light *X* contains dark *Y*"





Center object the *only* object that fits two-part model (no localization)

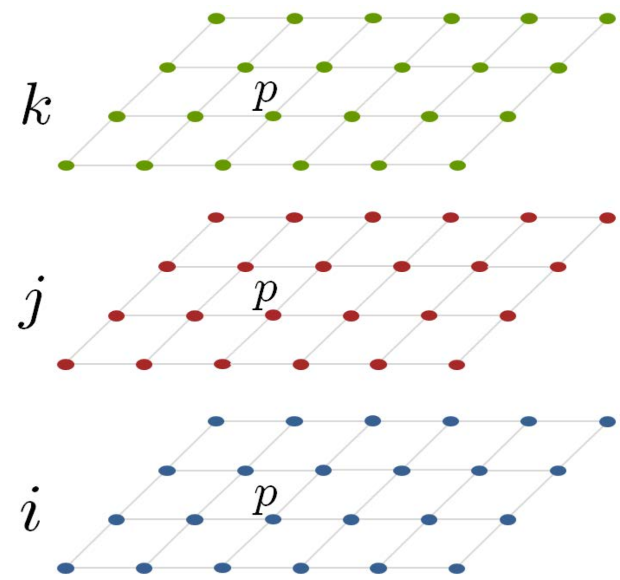
Our Energy

Let $\mathbf{x} \in \mathbb{B}^{\mathcal{L}\times\mathcal{P}}$ over objects L and pixels P

$$E_{mult}(\mathbf{x}) = \sum_{p\in\mathcal{P}} D_p(\mathbf{x}_p) + \sum_{i\in\mathcal{L}} V^i(\mathbf{x}^i) + \overbrace{\sum_{\substack{i,j\in\mathcal{L}\\ i\neq j}} W^{ij}(\mathbf{x})}^{\text{interaction terms}}$$

'layer cake'

a-la Ishikawa'03



variables x_p^i, x_p^j, x_p^k

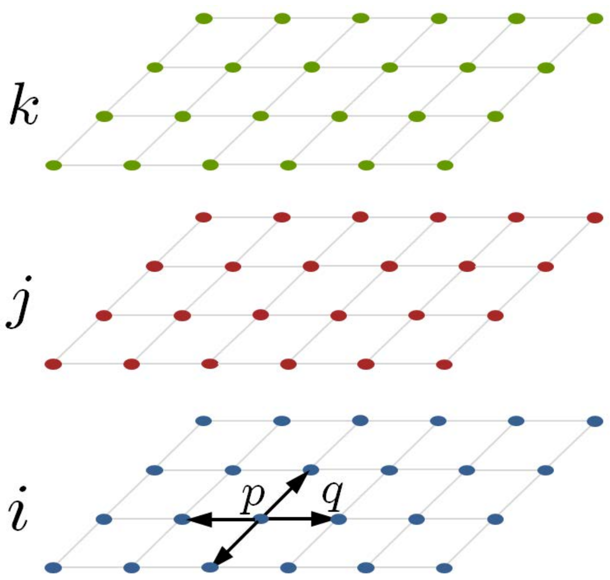
Surface Regularization Terms

Let $\mathbf{x} \in \mathbb{B}^{\mathcal{L}\times\mathcal{P}}$ over objects L and pixels P

$$E_{mult}(\mathbf{x}) = \sum_{p\in\mathcal{P}} D_p(\mathbf{x}_p) + \sum_{i\in\mathcal{L}} V^i(\mathbf{x}^i) + \overbrace{\sum_{\substack{i,j\in\mathcal{L}\\ i\neq j}} W^{ij}(\mathbf{x})}^{\text{interaction terms}}$$

- Standard regularization of each independent surface

$$V^i(\mathbf{x}^i) = \sum_{pq\in\mathcal{N}^i} V^i_{pq}(\mathbf{x}^i_p, \mathbf{x}^i_q)$$



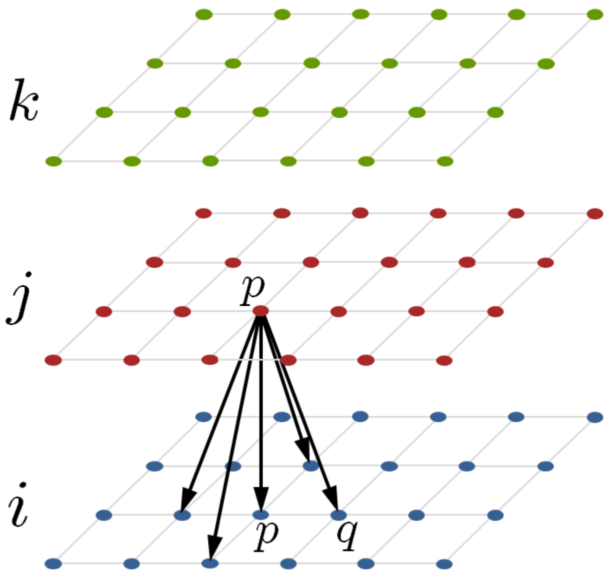
Geometric Interaction Terms

Let $\mathbf{x} \in \mathbb{B}^{\mathcal{L}\times\mathcal{P}}$ over objects L and pixels P

$$E_{mult}(\mathbf{x}) = \sum_{p\in\mathcal{P}} D_p(\mathbf{x}_p) + \sum_{i\in\mathcal{L}} V^i(\mathbf{x}^i) + \overbrace{\sum_{\substack{i,j\in\mathcal{L}\\ i\neq j}} W^{ij}(\mathbf{x})}^{\text{interaction terms}}$$

- Inter-surface interaction

$$W^{ij}(\mathbf{x}) = \sum_{pq\in\mathcal{N}^{ij}} W^{ij}_{pq}(\mathbf{x}^i_p, \mathbf{x}^j_q)$$



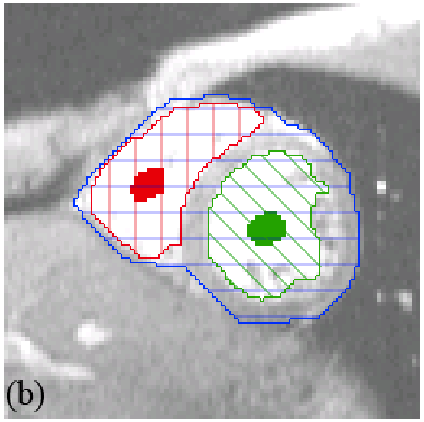
So What *Can* We Do?

- Nestedness/inclusion of sub-segments
  ICCV 2009 (submodular, exact solution)
- Spring-like repulsion of surfaces, minimum distance
  ICCV 2009 (submodular, exact solution)
- Spring-like attraction of surfaces, Hausdorf distance
  ECCV 2012 (approximation)

- Extends *Li, Wu, Chen & Sonka* [PAMI'06]
  - no pre-computed medial axes
  - no topology constraints

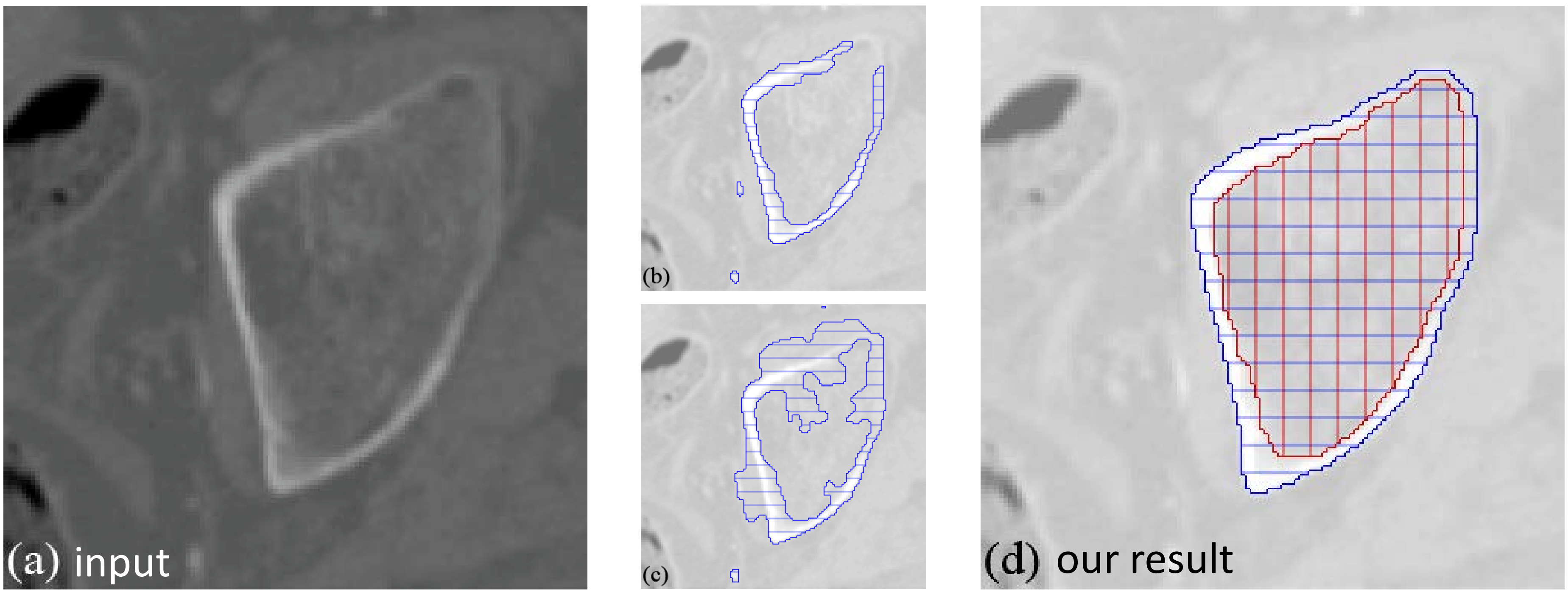
Applications

- Medical Segmentation
  - Lots of complex shapes with priors between boundaries
  - Better domain-specific models

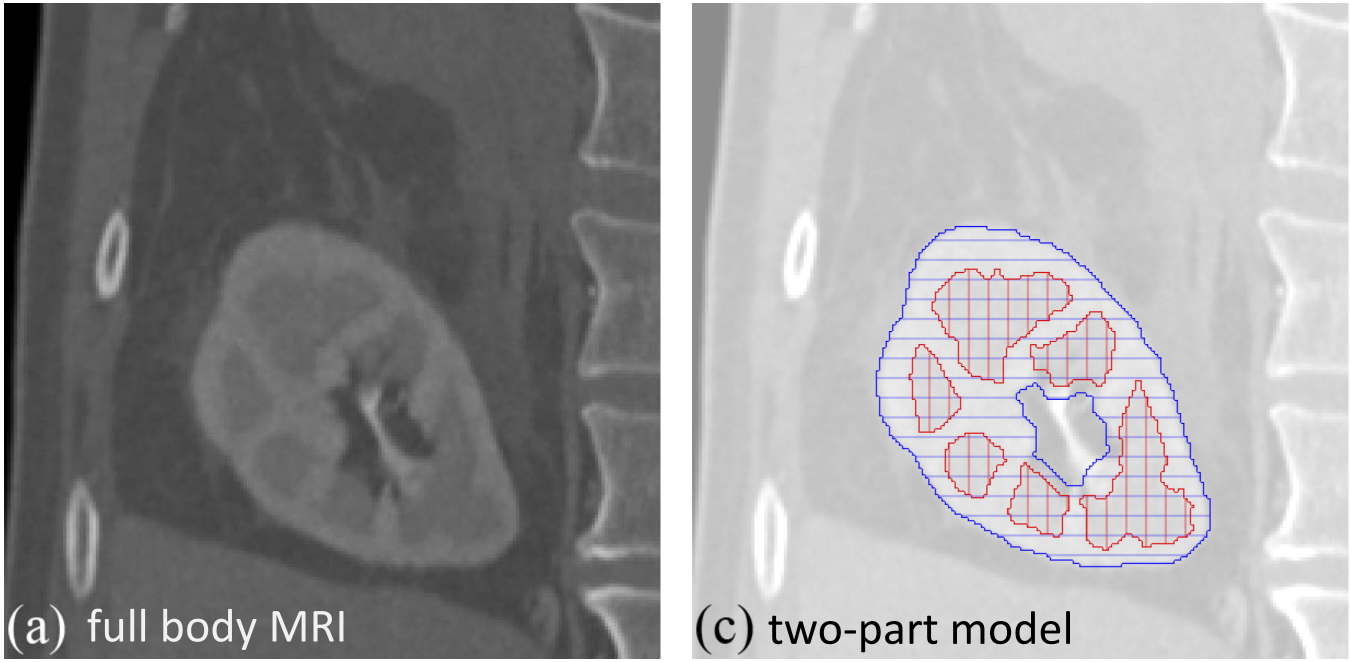


- Scene Layout Estimation
  - Basically just regularize Hoiem-style data terms [4]

Application: Medical



Application: Medical



Application: Medical

full body MRI

Conclusions

- linear functionals are very limited
  - be careful with i.i.d. (histograms or mixture models)

- higher-order regional functionals
- use sub-segments regularized by

  - complexity or sparcity (MDL, information theory)
  - geometry (based on anatomy)

literature: ICCV'09, EMMCVPR'11, ICCV'11, , IJCV12, ECCV'12