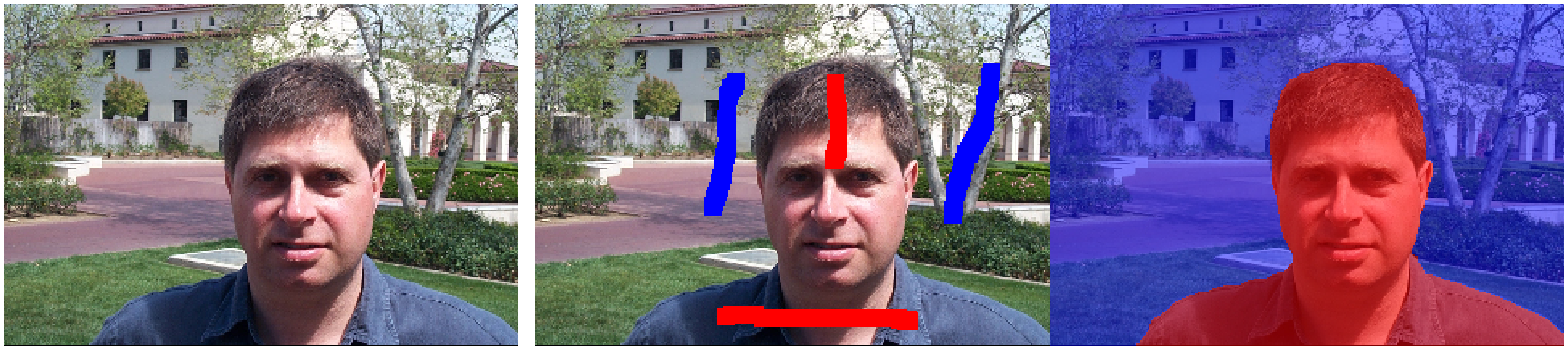
GraphCut segmentation

Tomáš Svoboda, svoboda@cmp.felk.cvut.cz
Czech Technical University in Prague
http://cmp.felk.cvut.cz/~svoboboda
and Jan Kybic

Segmentation, simplified intro

Label all pixels to a certain class.



Labeling for image analysis

Many image analysis and interpretation problems can be posed as *labeling problems*.

- Assign a *label* to image pixel (or to features in general).
- Image intensity can be considered as a label (think about pallete images).

Sites

$\mathcal{S}$ index a discrete set of m *sites*.

$$\mathcal{S} = \{1, \ldots, m\}$$

Site could be:

- individual pixel
- image region
- corner point, line segment, surface patch . . .

Labels and sites

A *label* is an event that may happen to a *site*.

Set of labels $\mathcal{L}$.

We will discuss *disrete labels*:

$$\mathcal{L} = \{\ell_1, \dots, \ell_M\}$$

shortly

$$\mathcal{L} = \{1, \dots, M\}$$

What can be a label?

- intensity value
- object label
- in edge detection binary flag $\mathcal{L} = \{\text{edge}, \text{nonedge}\}$
- in *image segmentation* flag $\mathcal{L} = \{\text{bckg}, \text{fg}\} = 0, 1$
- . . .

The Labeling Problem

Assigning a label from the label set $\mathcal{L}$ to each of the sites $\mathcal{S}$.

Example: image segmenation

Assign a label f_i from the set $\mathcal{L} = \{\mathrm{bckg}, \mathrm{fg}\}$ to site $i \in \mathcal{S}$ where the elements in $\mathcal{S}$ index the image pixels. The set

$$f = \{f_1, f_2\}$$

is called a *labeling*. In this particular case *segmentation*.

How many possible labelings?

Assuming all m sites have the same label set $\mathcal{L}$

$$\mathbb{F} = \underbrace{\mathcal{L} \times \mathcal{L} \times \ldots \times \mathcal{L}}_{m \text{ times}} = |\mathcal{L}|^m$$

Think about segmenting/labeling all pixels or image restoration problem.

The the m is the number of pixels in the image and $\mathcal{L}$ equals to number of intensity levels.

Many, many possible labelings. Usually only few are good.

Labeling problems in image analysis

- image restoration
- region segmentation
- edge detection
- object detection and recognition
- stereo
- . . .

Labeling with contextual analysis

In images, site neighborhood matters.

A probability $P(f_i)$ does not depend only on the site but also on the labeling around. Mathematically speaking we must consider conditional probability

$$P(f_i|\{f_{i'}\})$$

where $\{f_{i'}\}$ denotes the set of other labels.

no context:

$$P(f) = \prod_{i \in \mathcal{S}} P(f_i)$$

Remember what $\mathcal{S}$ denotes?

Markov Random Field – MRF

Let think about **images**:

- image intensities are the *random variables*
- the values depend only on their immediate spatial neighborhood which is the *Markov property*
- images are organized in a regular grid which can be seen as an *undirected graph*

$$P(f_i|f_{\mathcal{S}-\{i\}}) = P(f_i|f_{\mathcal{N}_i})$$

where $f_{\mathcal{N}_i}$ stands for the labels at the sites *neighboring* i.

MRF a Gibbs Random Fields

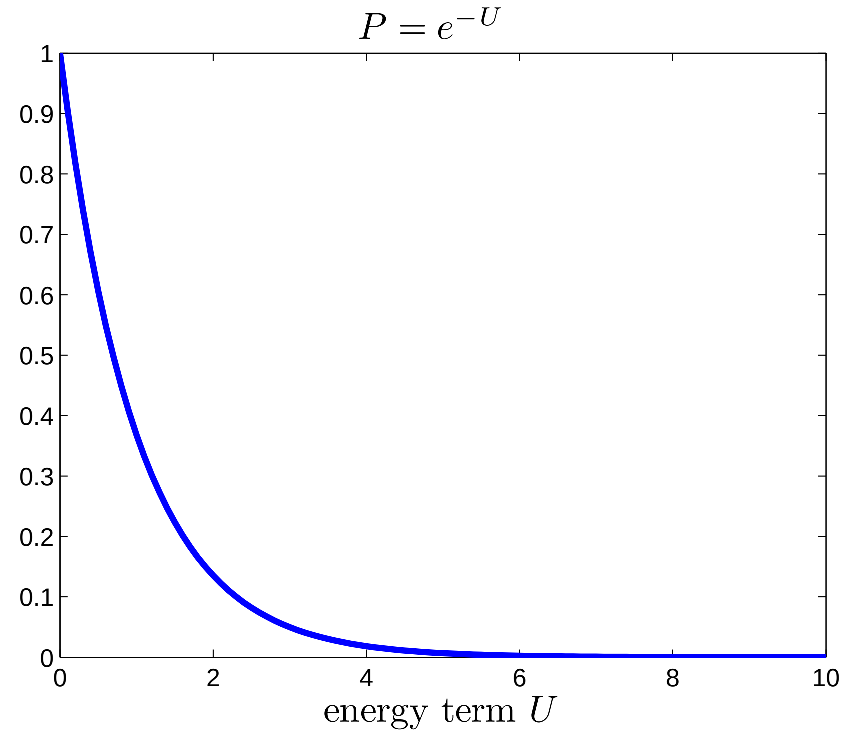
How to specify an MRF: Hammersley–Clifford theorem about equivalence between MRF and Gibbs distribution.

$$P(f) = Z^{-1} \cdot \mathrm{e}^{-\frac{1}{T}U(f)}$$

where

$$Z = \sum_{f \in \mathbb{F}} e^{-\frac{1}{T}U(f)}$$

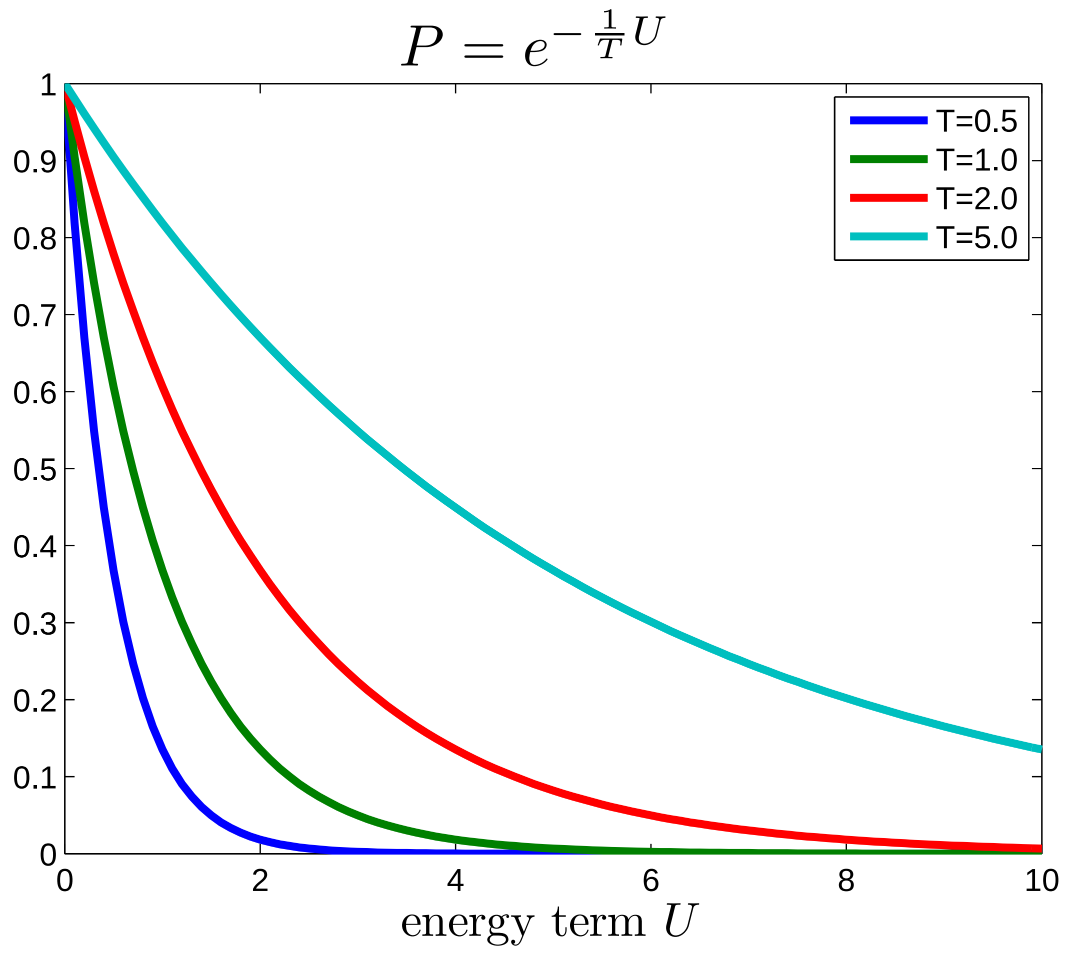
is a normalizing constant.



Gibbs distribution

$$P(f) = Z^{-1} \times e^{-\frac{1}{T}U(f)}$$

- T is *temperature*, $T = 1$ unless stated otherwise
- $U(f)$ is the *energy function*



Energy function

$$U(f) = \sum_{c \in \mathcal{C}} V_c(f)$$

is a sum of *clique potentials* $V_c(f)$ over all possible cliques $\mathcal{C}$.

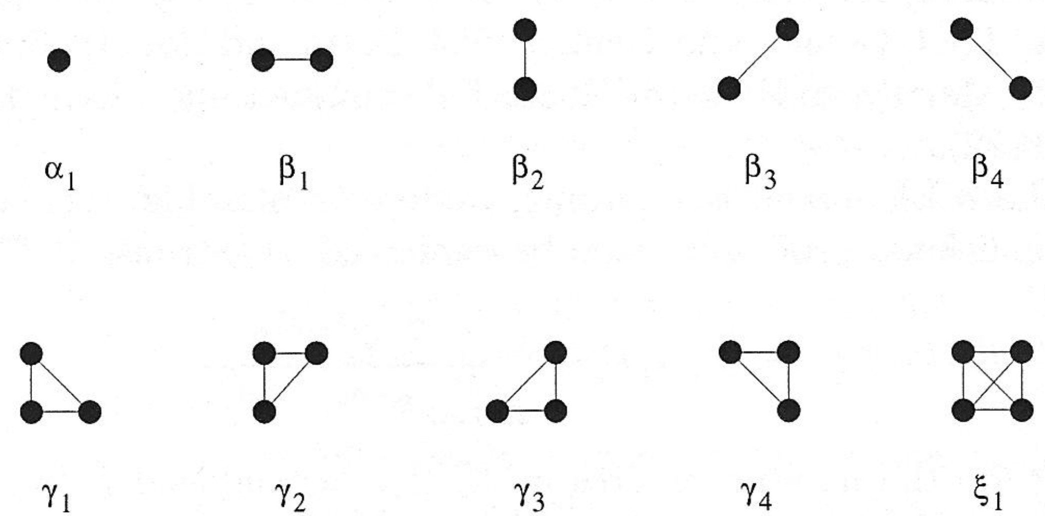


Figure 2.3: Clique types and associated potential parameters for the second-order neighborhood system. Sites are shown as dots and neighboring relationships as joining lines.

Simple cliques

Contextual constraints on two labels

$$U(f) = \sum_{i \in \mathcal{S}} V_i(f_i) + \sum_{i \in \mathcal{S}} \sum_{i' \in \mathcal{N}_i} W_{ii'}(f_i, f_{i'})$$

This can be interpreted as

$$U(f) = U_{data}(f) + U_{smooth}(f)$$

Neighborhood and cliques on a regular lattices

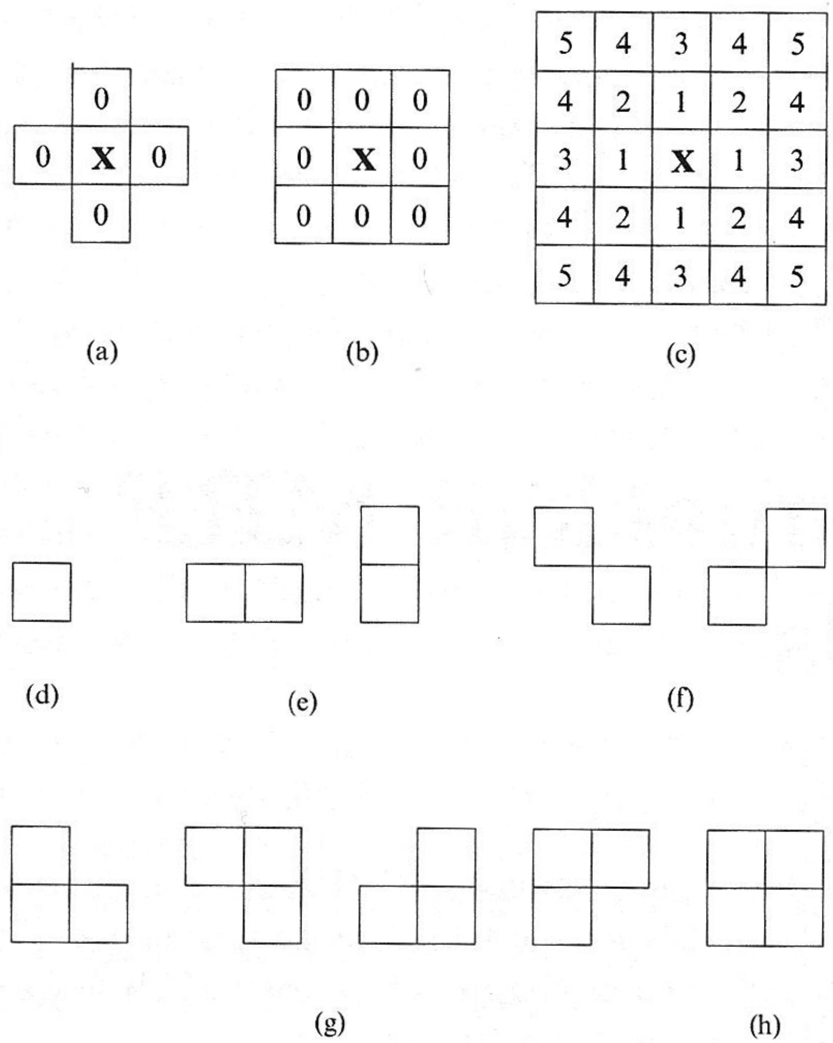


Figure 2.1: Neighborhood and cliques on a lattice of regular sites.

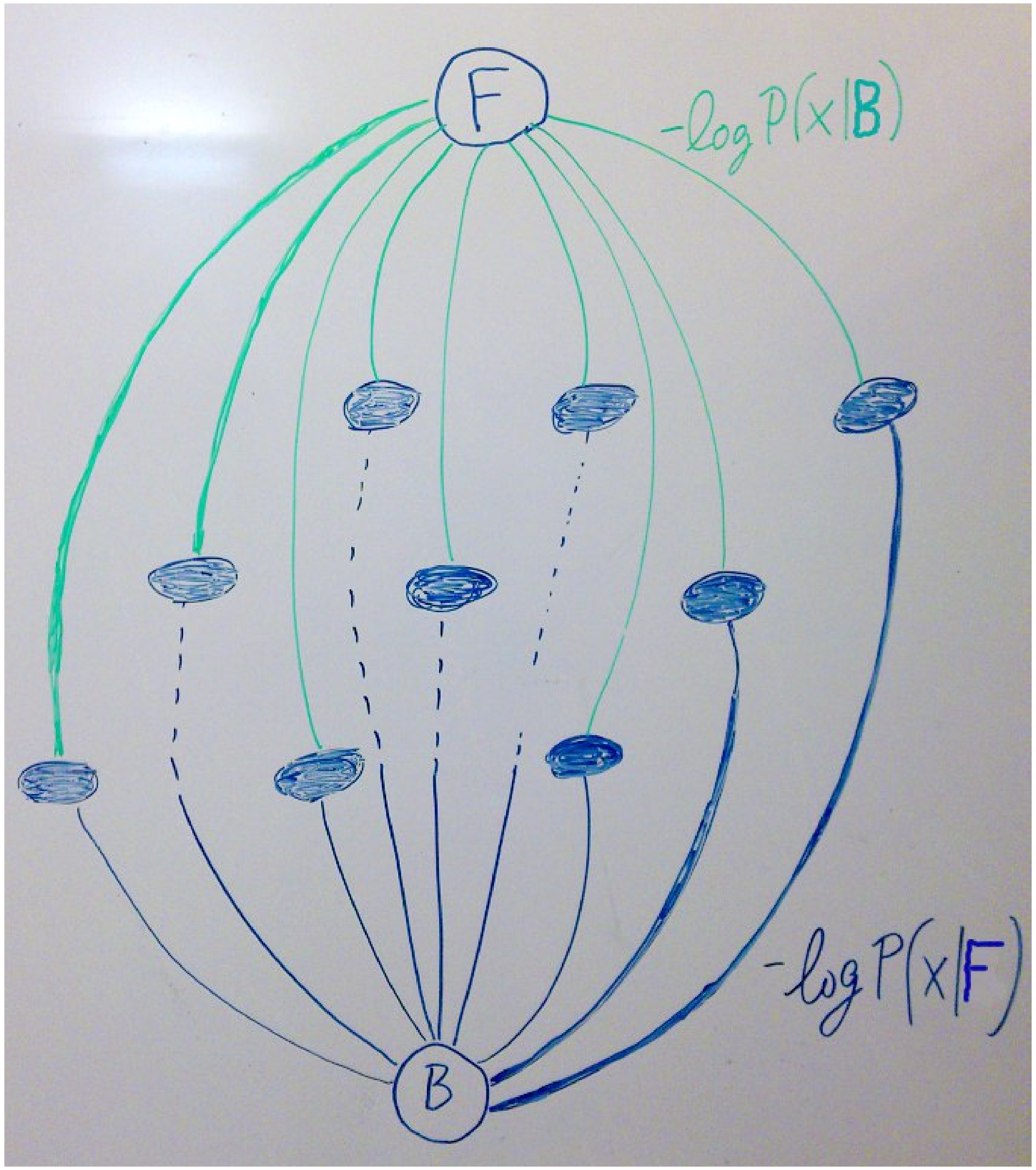
Finding best labeling (segmentation)

In MAP formulation we seek the most probable labeling $P(f)$.

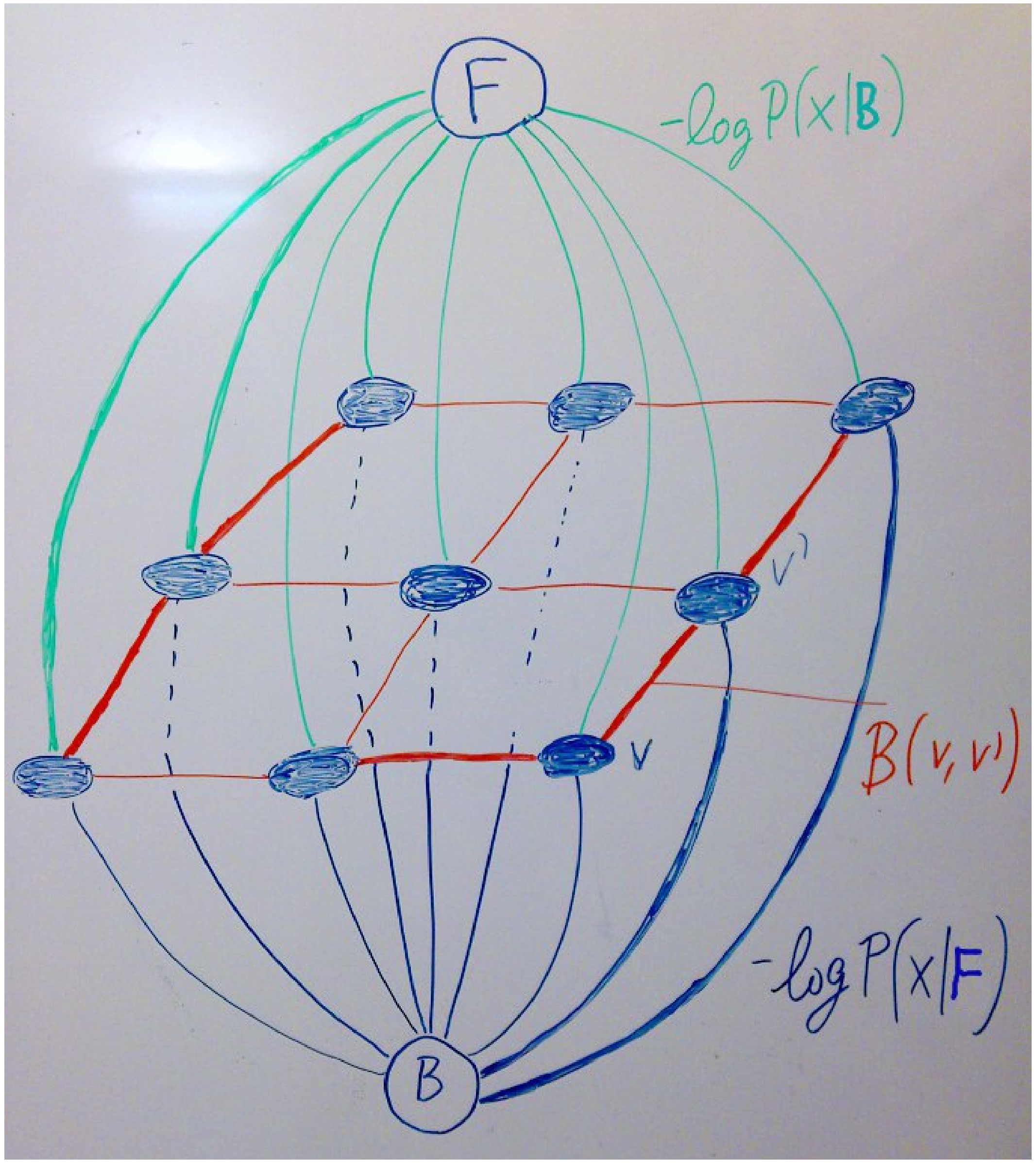
Minimize energy $U(f)$.

A combinatorial problem. Exact solution for binary problems. Many possible algorithms based on minimum cut, maximum flow (Ford-Fulkerson, Edmonds-Karp, Boykov-Kolmogorov, push-relabel, GridCut)

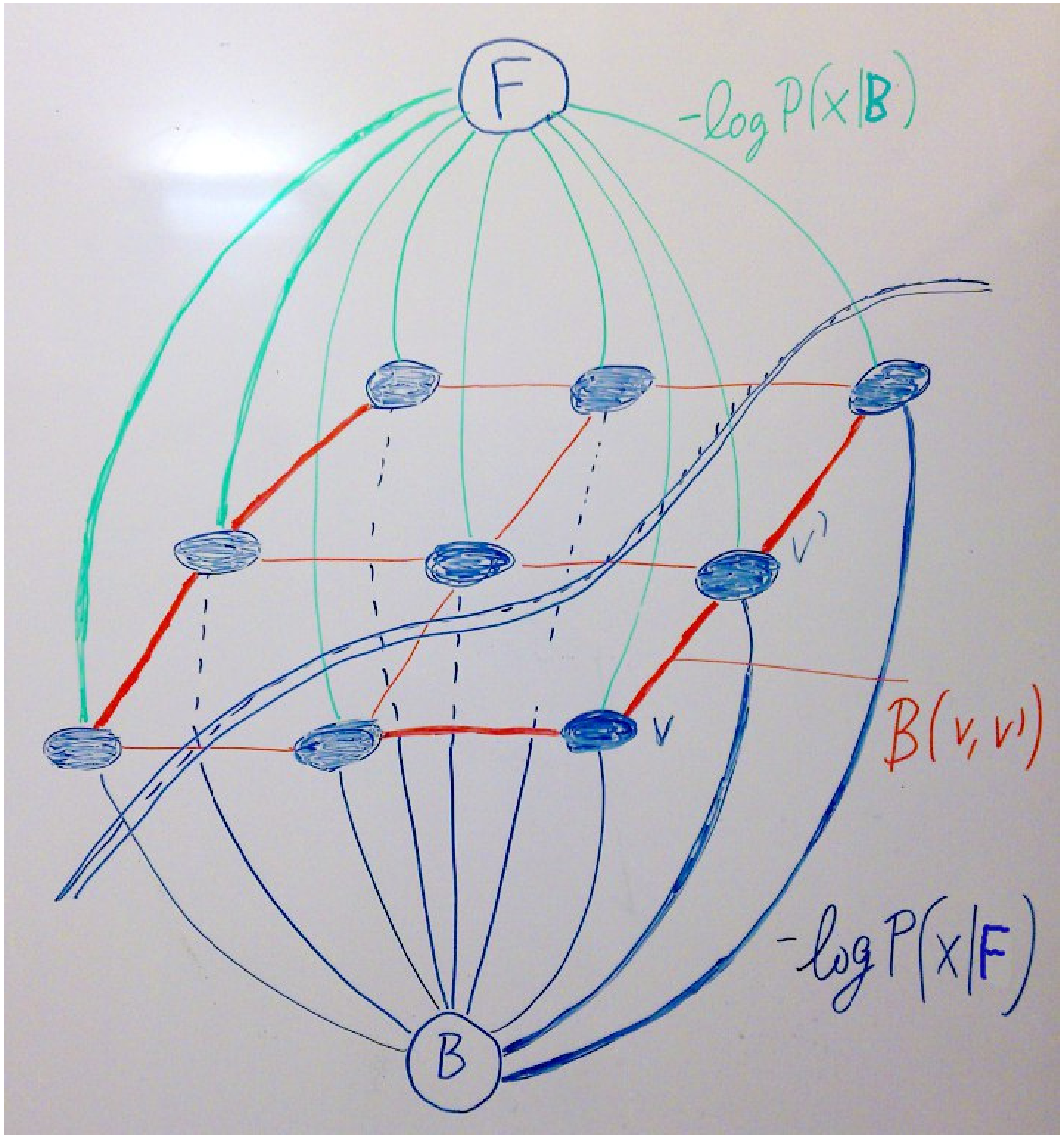
Edges, Data term only



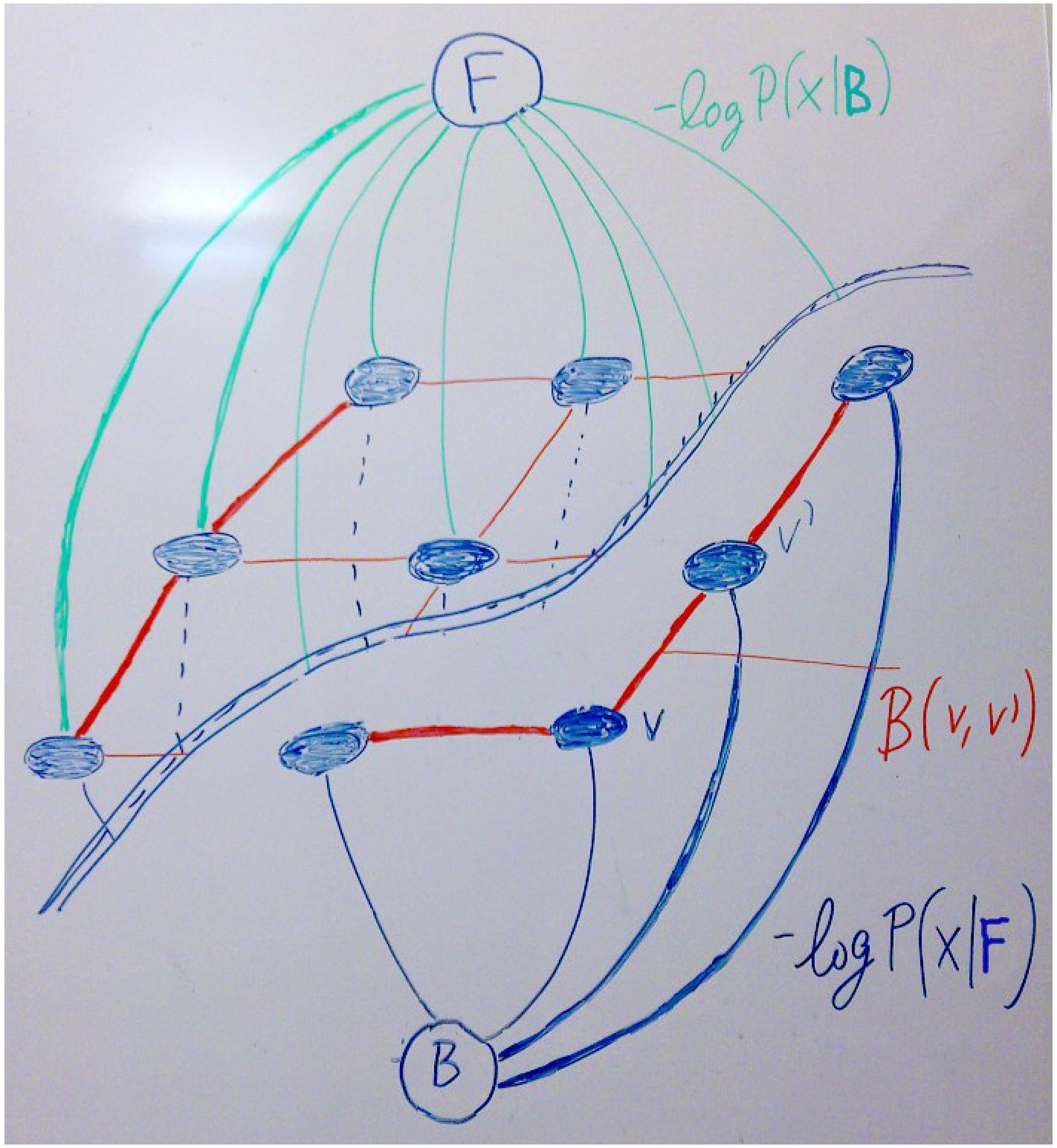
Edges, adding pixel connections



Edges, cutting graph



Edges, cut – Segmentation



Mincut formulation equivalence

$$U(f) = \sum_{i \in \mathcal{S}} V_i(f_i) + \sum_{i \in \mathcal{S}} \sum_{i' \in \mathcal{N}_i} W_{ii'}(f_i, f_{i'})$$

Unary costs

$$V_i(0) = w_{iF}, \quad V_i(1) = w_{iB}$$

Binary costs

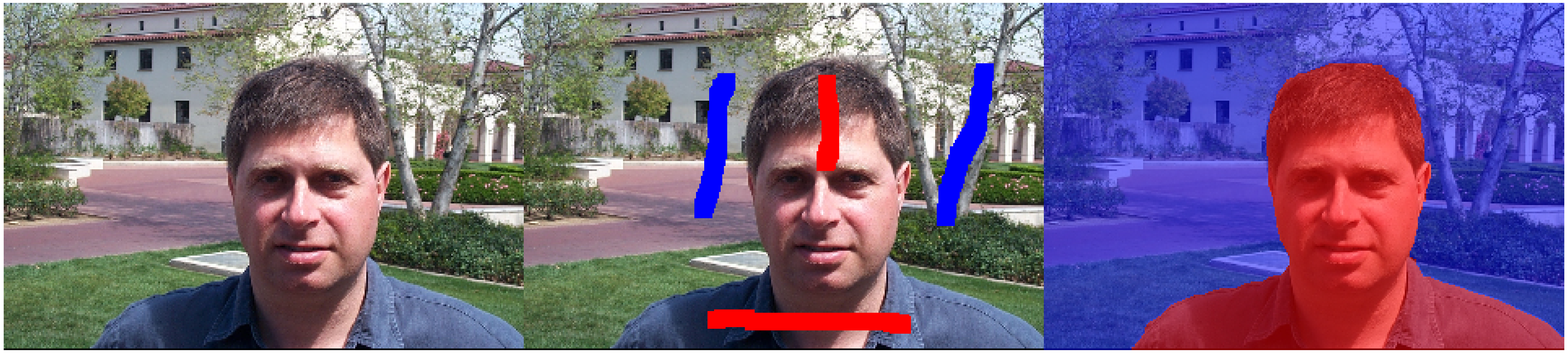
$$W_{ii'} = w_{ii'} \delta(f_i, f_{i'}) = w_{ii'} [\![f_i \neq f_{i'}]\!] \quad \text{(Potts model)}$$

Cost can depend on the image (edge term)

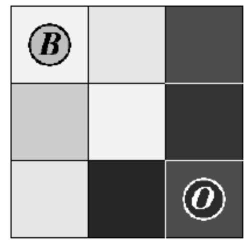
$$w_{ii'} = c\, \mathrm{e}^{-\beta \left| x_i - x_{i'} \right|}$$

Segmentation with seeds – Interactive GraphCuts

Idea: denote few pixels that trully belongs to object or background and than refine (grow) by using soft constraints.



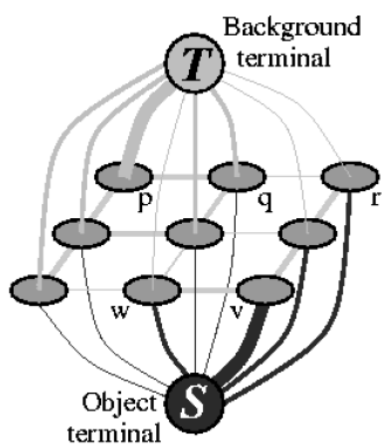
Segmentation with seeds – graph



(a) Image with seeds.

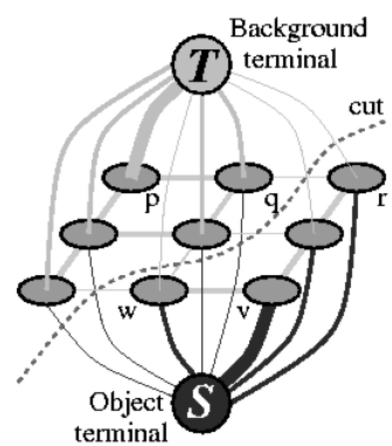
(d) Segmentation results.

⇓ ⇑



(b) Graph.

⇒



(c) Cut.

What data term?

$$\sum_{i \in \mathcal{S}} V_i(f_i)$$

The better matching pixel intensity/color, the lower energy (penalty).

- background, foreground (object) pixels
- intensity or color distributions
- histograms
- parametric models: GMM – Gaussian Mixture Model

Image intensities - 1D GMM

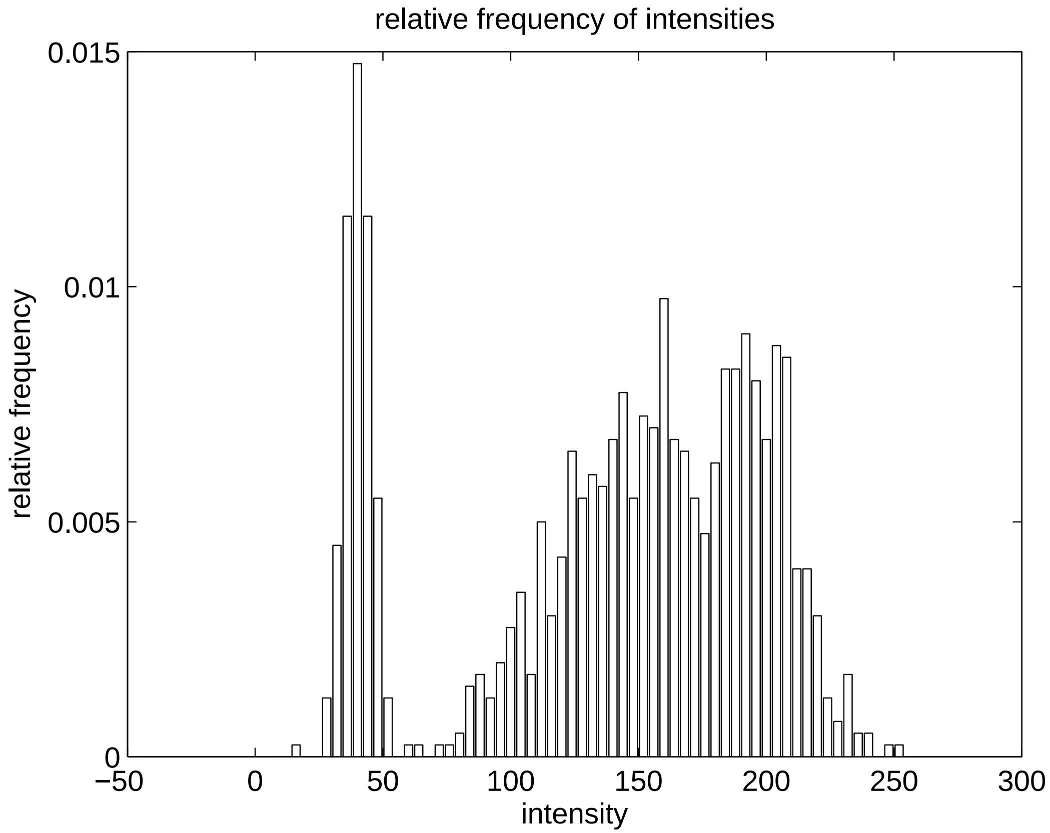


Image intensities - 1D GMM

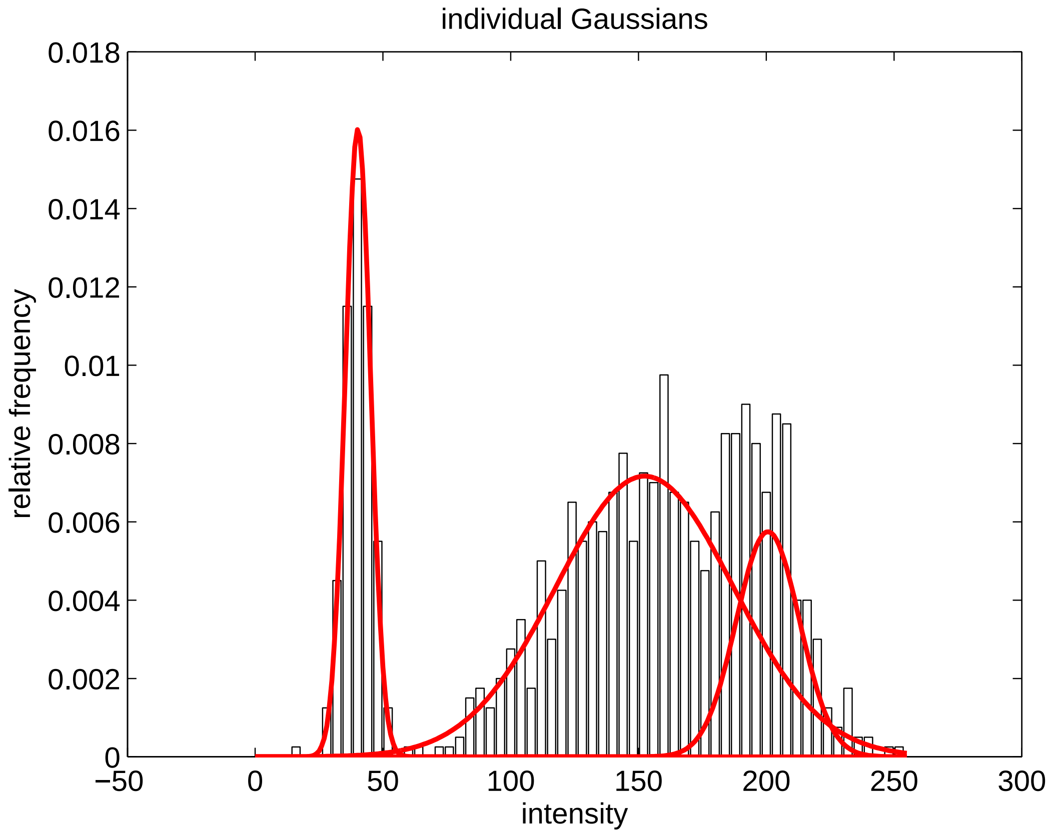
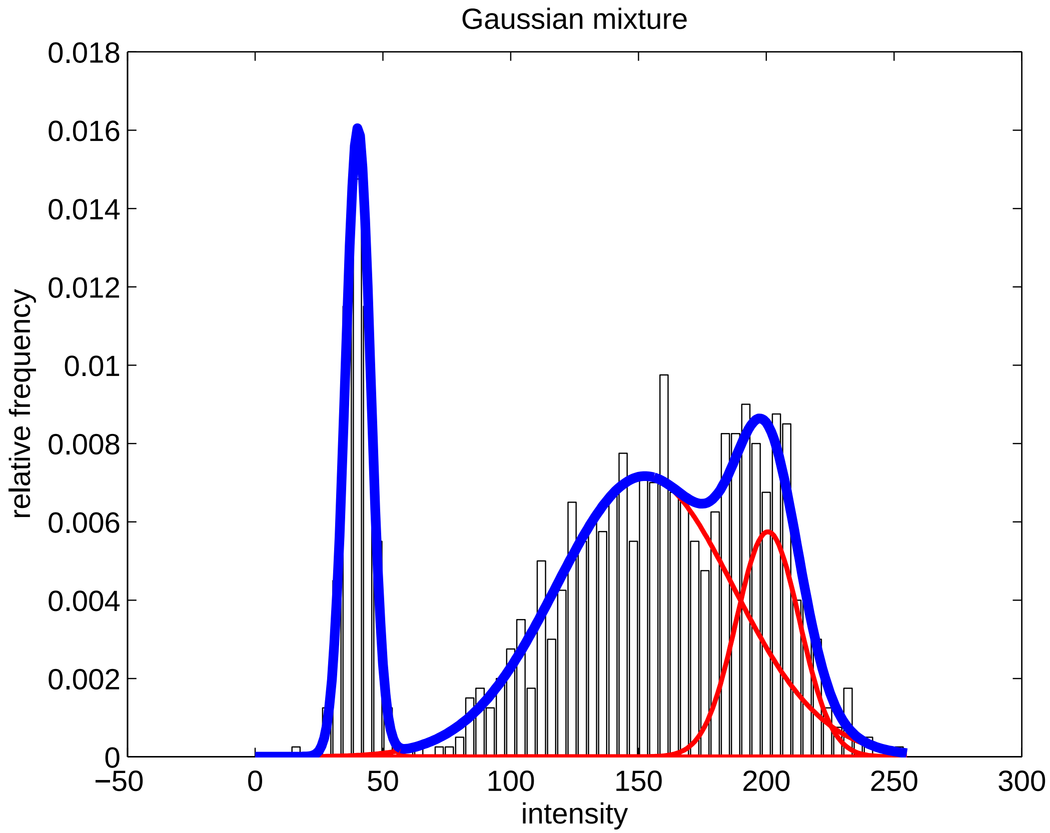
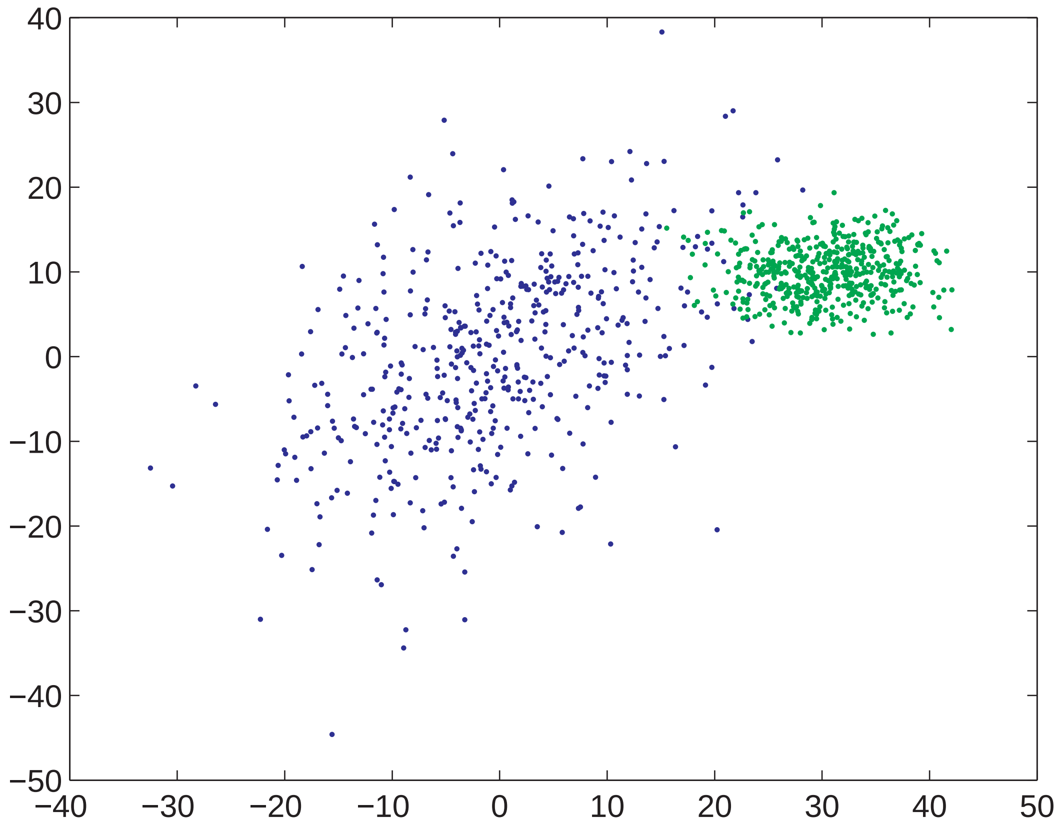


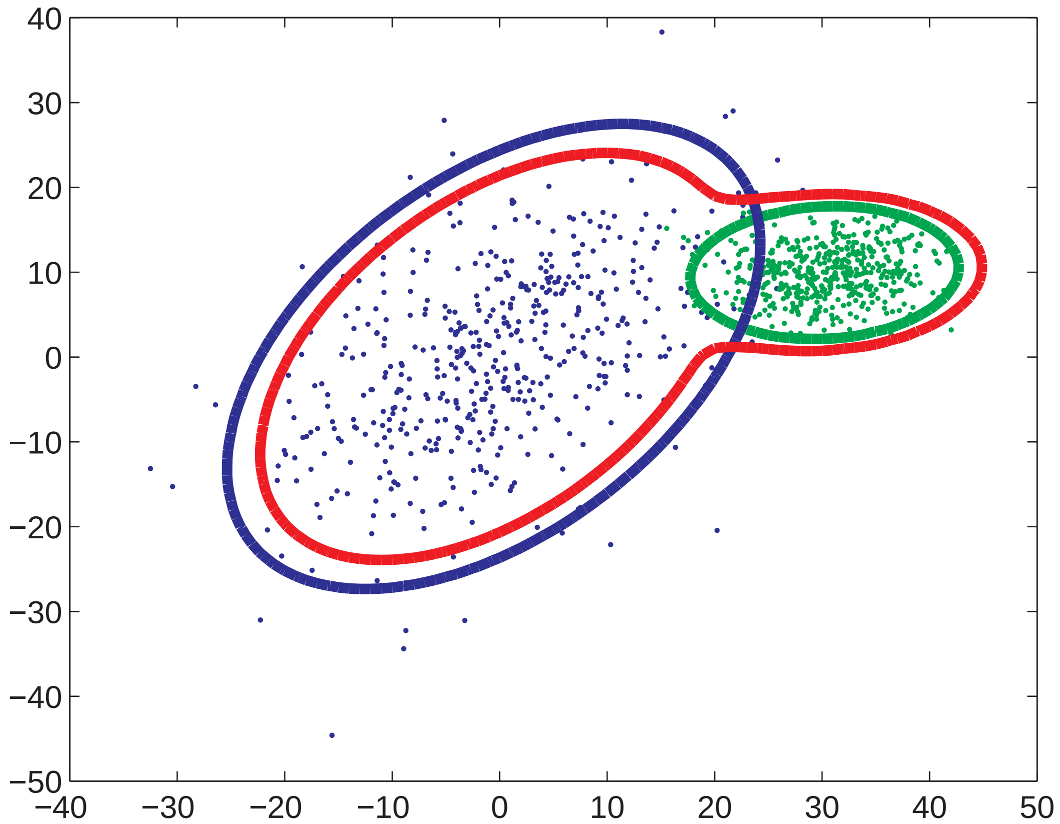
Image intensities - 1D GMM



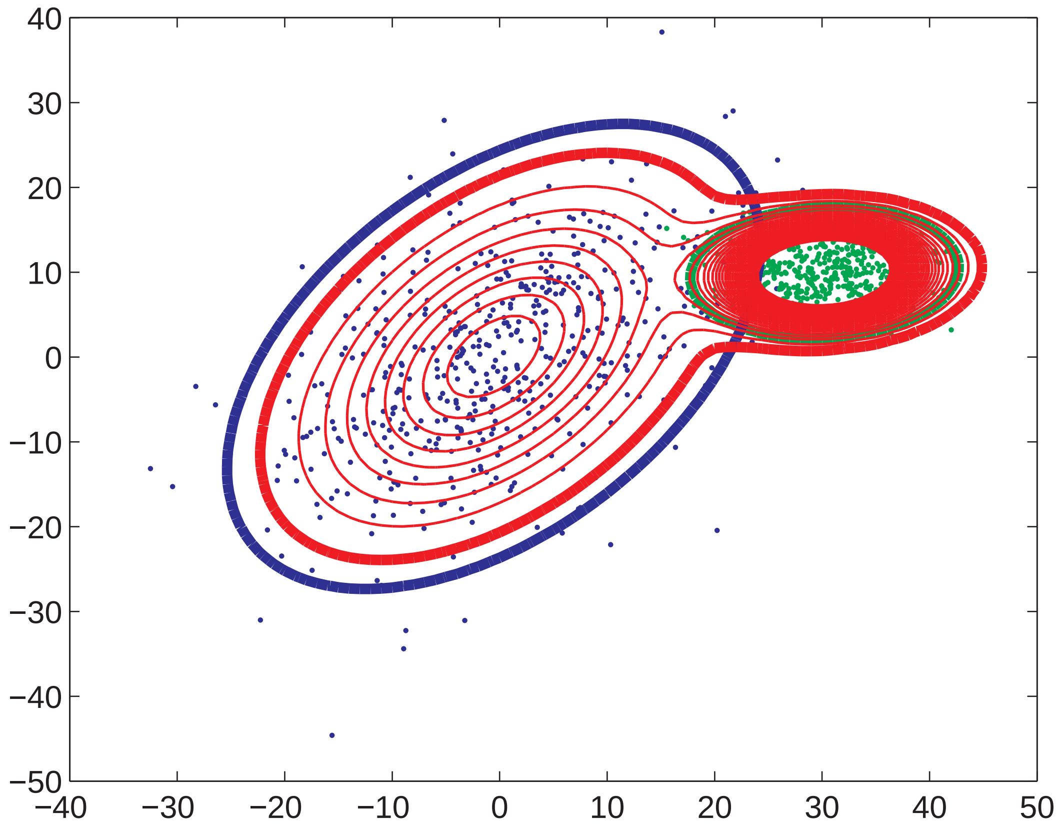
2D GMM



2D GMM



2D GMM



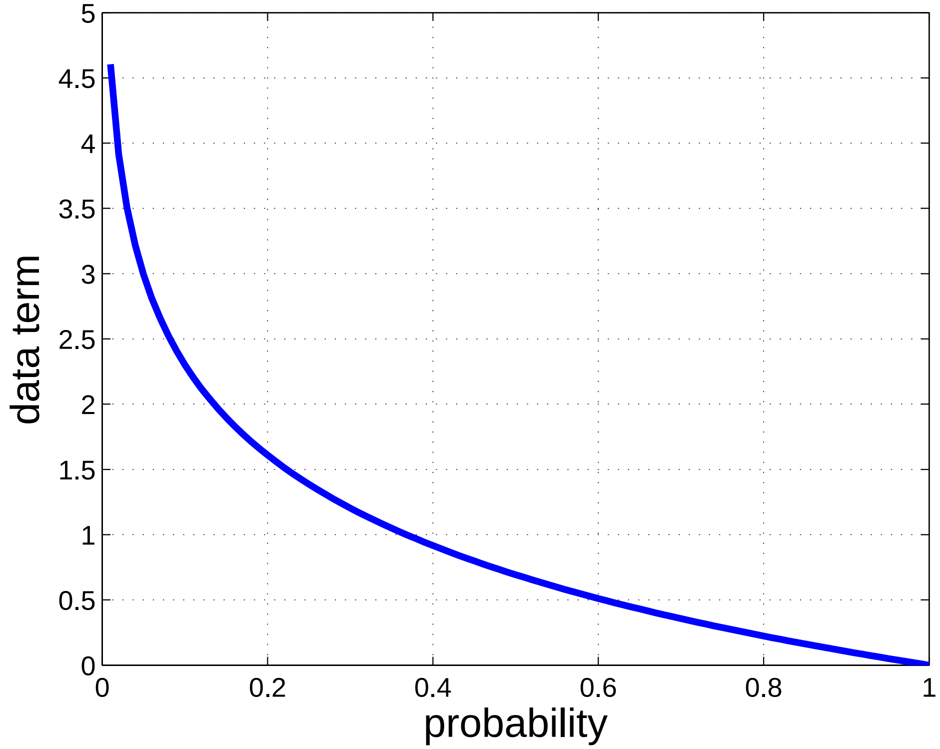
Data term by GMM – math summary

$$p(\mathbf{x}|f_i) = \sum_{k=1}^{K} w_k^{f_i} \frac{1}{(2\pi)^{\frac{d}{2}} \left|\Sigma_k^{f_i}\right|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(\mathbf{x} - \mu_k^{f_i})^T \, {\Sigma_k^{f_i}}^{-1} (\mathbf{x} - \mu_k^{f_i})\right) ,$$

where d is the dimension.

- K number of Gaussians. User defined.
- for each label $\mathcal{L} = \{\mathrm{obj}, \mathrm{bck}\}$ different w_k, μ_k, Σ_k estimated from the data (seeds)
- $\mathbf{x}$ pixel value, can be intensity, color vector, . . .

Data term



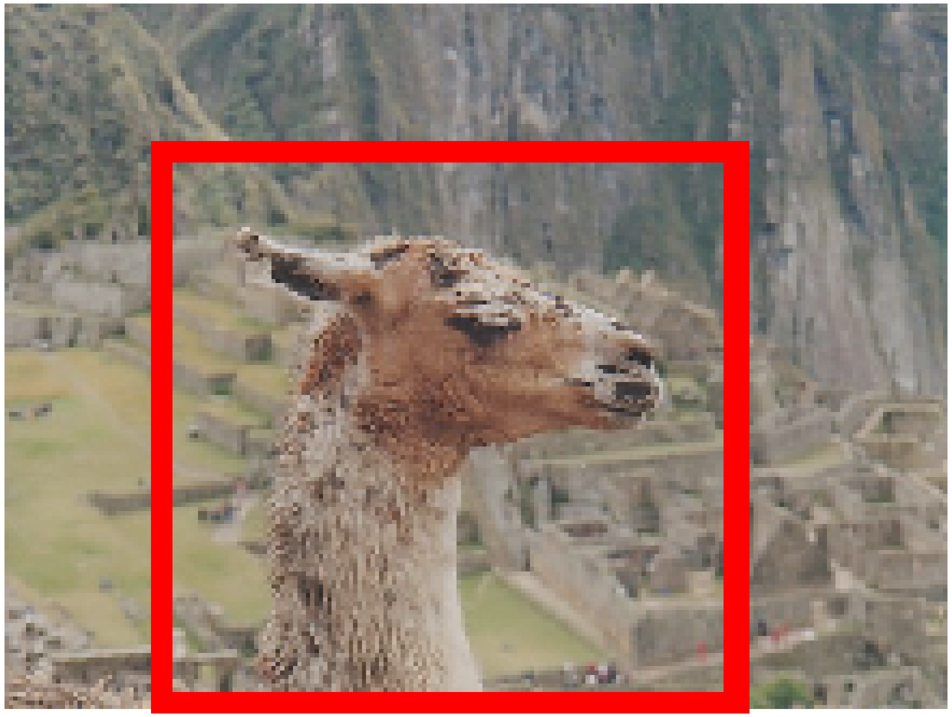
$$V_i(1) = -\ln p(\mathbf{x}|\mathrm{obj})$$
$$V_i(0) = -\ln p(\mathbf{x}|\mathrm{bck})$$

Other alternatives

- single Gaussian
- p estimated directly from a histogram
- train a classifier
- multidimensional features (color, neighborhood, texture. . .)

GrabCut

Iterate the GraphCut and update the data model in each iteration. Stop if the energy (penalty) does not decrease.



GrabCut – algorithm

Init: Specify background/foreground/unknown pixels (bounding box, scribbles)

Iterative minimization:

1. learn GMM parameters from data
2. segment by using Graphcut
3. repeat until convergence

Optional: add constraints and repeat the min-cut.

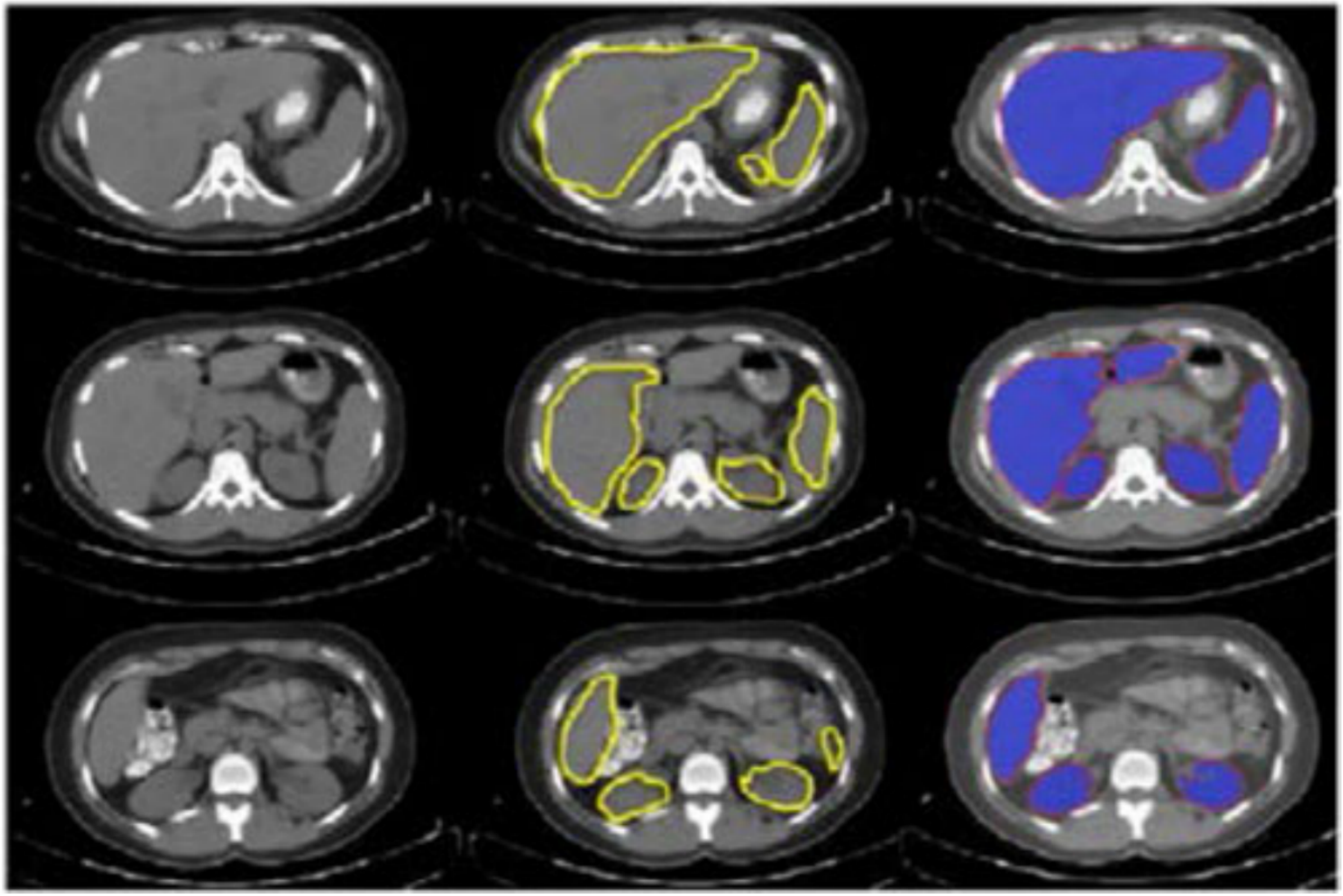
Multiclass segmentation

(a) (b) (c) (d)

- greedy pixel changes $a \rightarrow b$
- $\alpha - \beta$ swap $a \rightarrow c$
- α expansion $a \rightarrow d$
- Convert to binary problems
- Stop after first unsuccessful cycle.

α-β swap and α-expansion.

Multiorgan segmentation

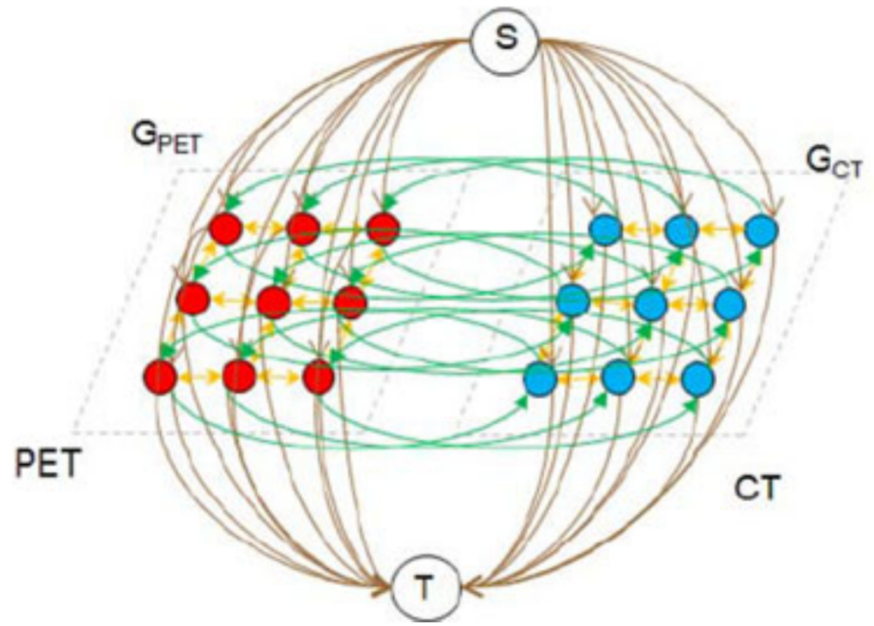


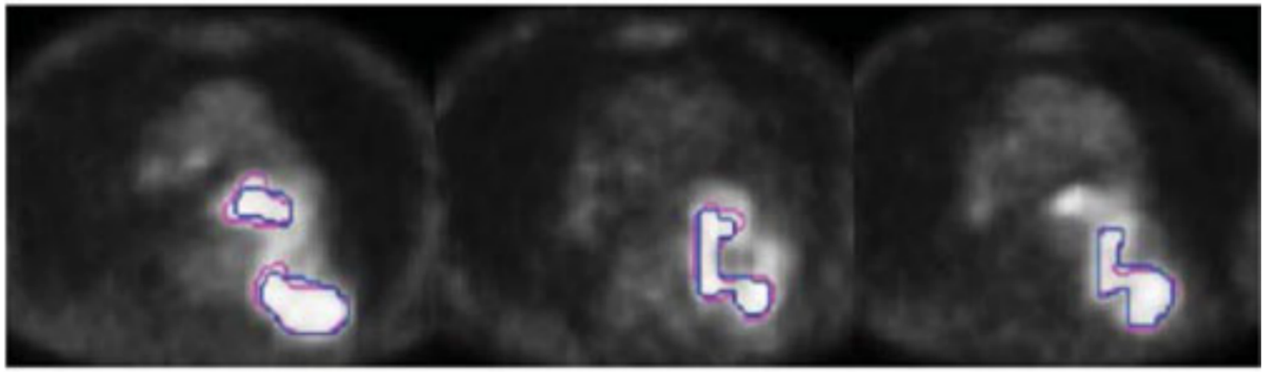
Active shape model for recognition, GraphCut for delineation

X. Chen and U. Bagci, “3D automatic anatomy segmentation based on iterative graph-cut-ASM,” Med. Phys., vol. 38, pp. 4610–4622, 2011.

Multimodality segmentation

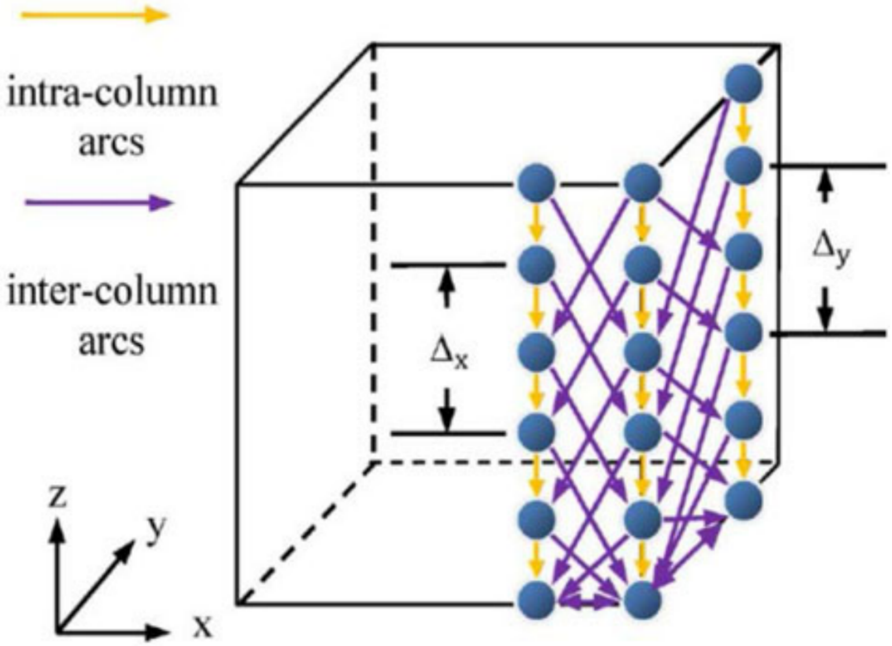
Edges to penalize difference between modalities

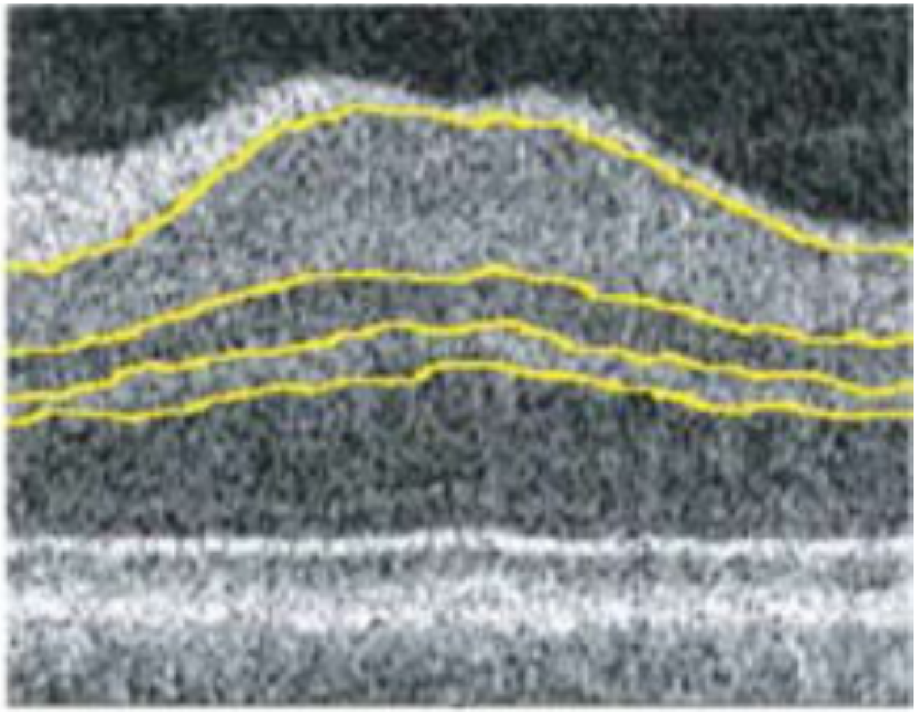




X. Chen and L. Pan, "A Survey of Graph Cuts/Graph Search Based Medical Image Segmentation," in IEEE Reviews in Biomedical Engineering, 2018

Surface detection





Combining CNN and GraphCuts