Introduction to Image Segmentation:

Part 2: multi-label segmentation

Yuri Boykov

University of Western Ontario

Multi-label segmentation and high-order constraints

- Basic energies of image labelings
- Move making (and other) algorithms
- Geometric constraints on multi-labelings

Submodular functions

- Edmonds 1970

Lattice $(\mathcal{L},\wedge,\vee)$ - set of elements with *inf* and *sup* operations

$$S,T\in\mathcal{L} \quad\Rightarrow\quad S\wedge T\in\mathcal{L} \qquad S\vee T\in\mathcal{L}$$

Function $\mathrm{E}:\mathcal{L}\rightarrow\Re$ s called **submodular** if for any $S,T\in\mathcal{L}$

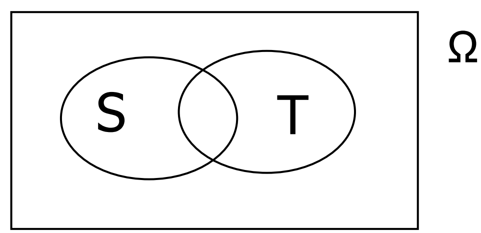
$$E(S\wedge T)+E(S\vee T)\leq E(S)+E(T)$$

Submodular set functions

Assume set Ω, then $(2^{\Omega}, \cap, \cup)$ is a lattice of subsets

Set function $E : 2^{\Omega} \rightarrow \Re$ is **submodular** if for any $S, T \subseteq \Omega$

$$E(S \cap T) + E(S \cup T) \leq E(S) + E(T)$$

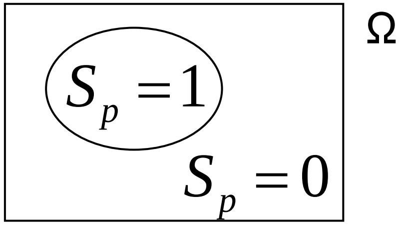


Significance: any submodular set function can be globally optimized in polynomial time [Grotschel et al.1981,88, Schrijver 2000] $O\left(|\Omega|^{9}\right)$

Submodular set functions

Sets are conveniently represented by binary indicator variables

$$S \subset \Omega \quad \leftrightarrow \quad \{ S_p \in \{0,1\} \mid p \in \Omega \}$$



Thus, set functions $E : 2^{\Omega} \rightarrow \Re$ can be represented as

$$E(S) = E(S_1, S_2, \ldots, S_{|\Omega|})$$

Define $S_A = \{ S_p \mid p \in A \}$, a *restriction* of S to any subset $A \subseteq \Omega$ and consider *projections* $E(S_A \mid S_{\Omega \setminus A})$ of energy E onto subsets A

Set function $E(S)$ is **submodular** iff for any pair $p, q \in \Omega$

$$E(\mathbf{0},\mathbf{0} \mid S_{\Omega \setminus pq}) + E(\mathbf{1},\mathbf{1} \mid S_{\Omega \setminus pq}) \leq E(\mathbf{1},\mathbf{0} \mid S_{\Omega \setminus pq}) + E(\mathbf{0},\mathbf{1} \mid S_{\Omega \setminus pq})$$

Graph cuts for minimization of submodular set functions

Assume set Ω and *2nd*-order (quadratic) function

$$E(S) = \sum_{(pq)\in N} E_{pq}(S_p, S_q) \qquad S_p, S_q \in \{0,1\}$$

Indicator variables

Function *E(S)* is **submodular** if for any $(p,q)\in N$

$$E_{pq}(\mathbf{0},\mathbf{0}) + E_{pq}(\mathbf{1},\mathbf{1}) \leq E_{pq}(\mathbf{1},\mathbf{0}) + E_{pq}(\mathbf{0},\mathbf{1})$$

Significance: submodular 2^{nd}-order boolean (set) function can be globally optimized in polynomial time by **graph cuts**

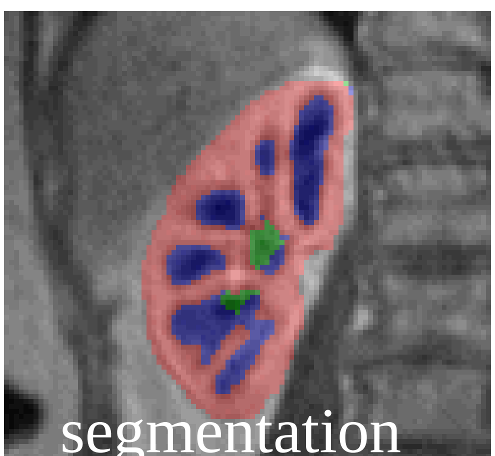
[Hammer 1968, Pickard&Ratliff 1973] $O\left(|N|\cdot|\Omega|^2\right)$

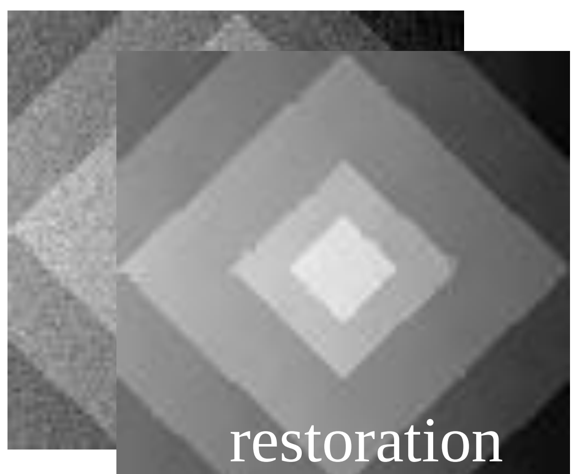
[Boros&Hammer 2000, Kolmogorov&Zabih2003]

Submodular labeling energies

Labelings $\boldsymbol{L}: \Omega \rightarrow \Lambda$

examples of image labelings (non-binary)



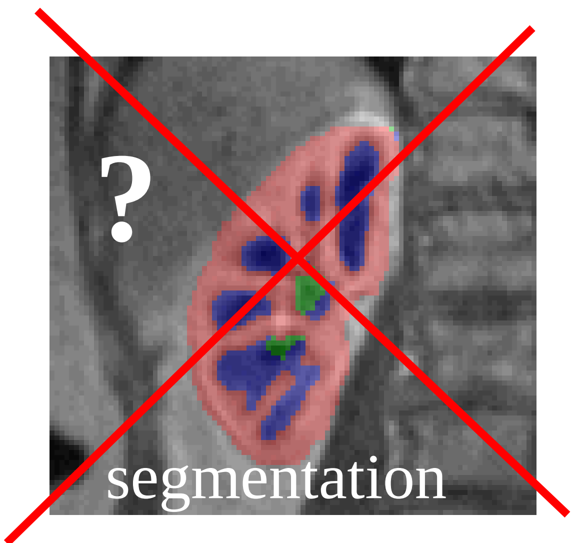


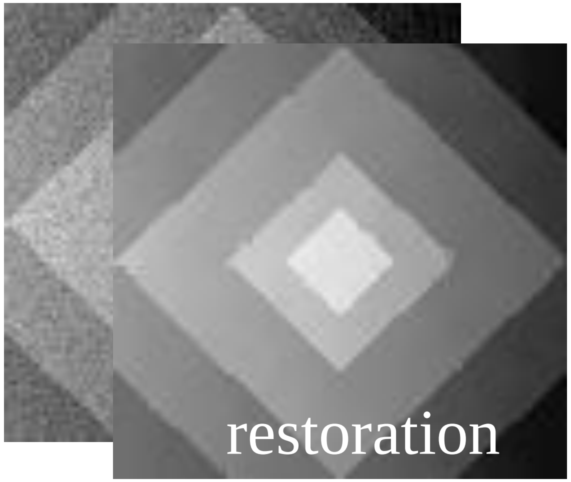
Submodular labeling energies

Labelings $\boldsymbol{L}: \Omega \rightarrow \Lambda$ form a lattice $(\Lambda^{\Omega}, \wedge, \vee)$ for **strictly ordered** labels Λ, e.g. for $\Lambda=\{1,..,n\}$

$$\boldsymbol{L} = (L_p) = \{L_p \mid p \text{ in } \Omega\}$$

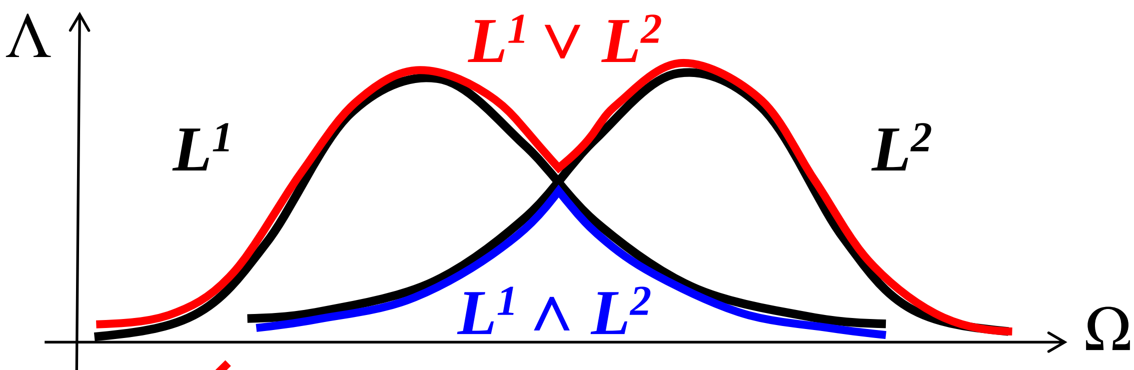
$$(L_p^1) \wedge (L_p^2) = (L_p^1 \wedge L_p^2) \qquad (L_p^1) \vee (L_p^2) = (L_p^1 \vee L_p^2)$$

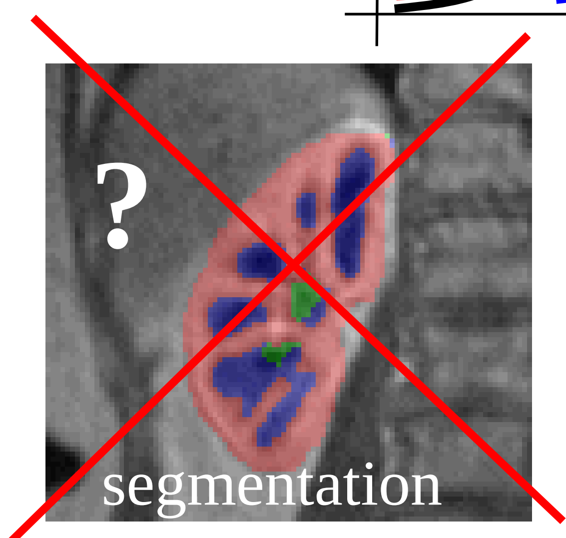


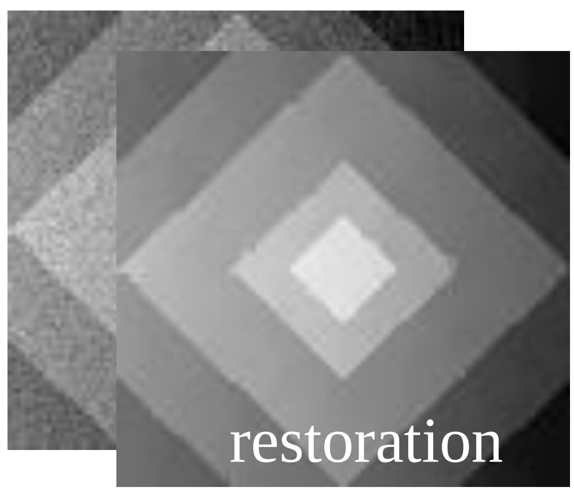


Submodular labeling energies

Labelings $L: \Omega \rightarrow \Lambda$ form a lattice $(\Lambda^{\Omega}, \wedge, \vee)$
for **strictly ordered** labels Λ, e.g. for $\Lambda=\{1,..,n\}$

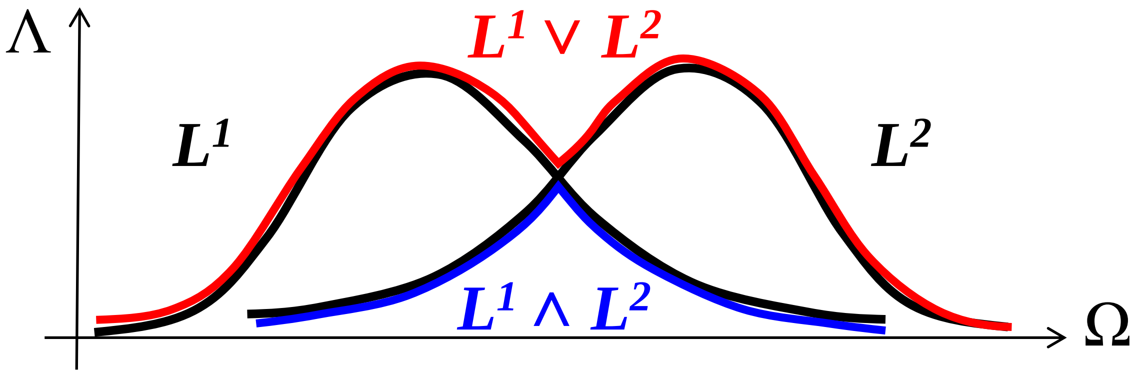






Submodular labeling energies

Labelings $L: \Omega \rightarrow \Lambda$ form a lattice $(\Lambda^{\Omega}, \wedge, \vee)$ for **strictly ordered** labels Λ, e.g. for $\Lambda=\{1,..,n\}$



Energy $E(L)$ is **submodular** if for any two labelings

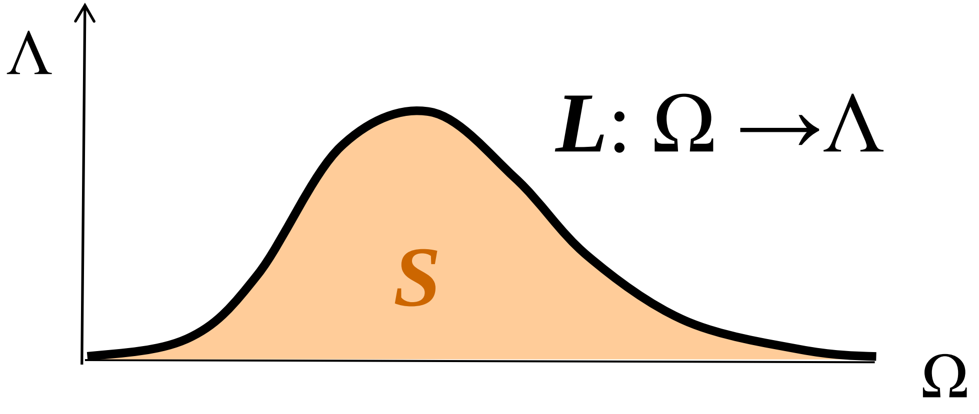
$$E(L^1 \wedge L^2) + E(L^1 \vee L^2) \leq E(L^1) + E(L^2)$$

Reducing to set functions

Theorem [Birkhoff, 1937]: any distrib. lattice $(\mathcal{L},\wedge,\vee)$ is isomorphic to a set lattice $(2^{\Omega},\cap,\cup)$ for some Ω.

Example [e.g. Ishikawa 1999]: labelings in Λ^{Ω} for strictly ordered set of labels Λ can be represented as subsets of ΩxΛ.

$S \subset \Omega \times \Lambda$



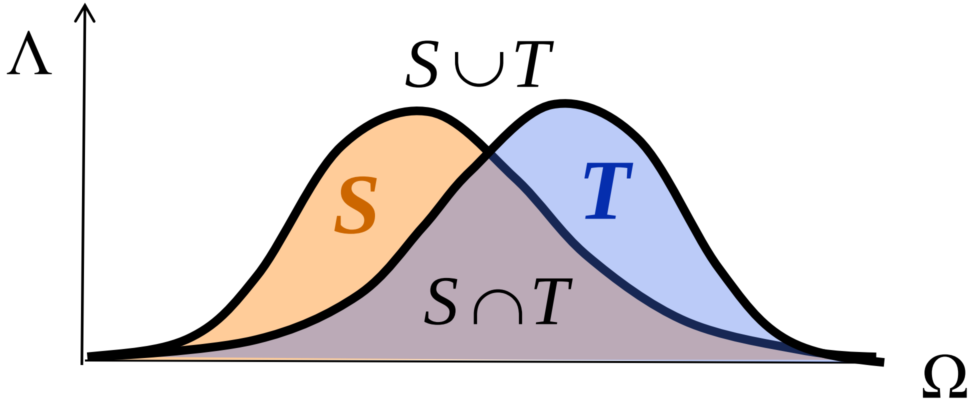
Reducing to set functions

Note: submodular energy $E(L)$ of labelings L in Λ^Ω gives submodular set function $E(S)= E(L)$.

$$E(L^1 \wedge L^2) + E(L^1 \vee L^2) \leq E(L^1) + E(L^2)$$

$$E(S \cap T) + E(S \cup T) \leq E(S) + E(T)$$

$$S \subset \Omega \times \Lambda$$



Graph cuts for minimization of submodular pairwise labeling energies

$$E(L) = \sum_{p\in\Omega} E_p(L_p) + \sum_{(pq)\in N} E_{pq}(L_p, L_q) \qquad L_p \in \Lambda$$

strictly ordered

Function $E(L)$ is **submodular** if for any $(p,q)\in N$

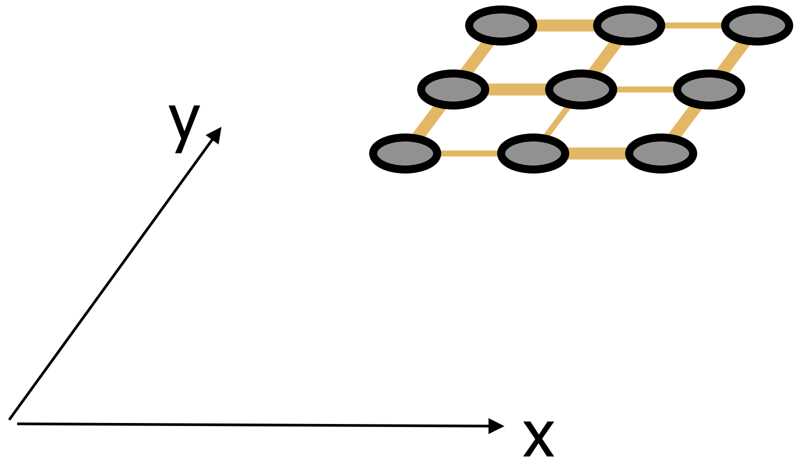
$$E_{pq}(a_1 \wedge a_2, b_1 \wedge b_2) + E_{pq}(a_1 \vee a_2, b_1 \vee b_2) \le E_{pq}(a_1, b_1) + E_{pq}(a_2, b_2)$$

$$E_{pq}(a,b) = g(a-b)$$ for some **convex** function $g(\,)$

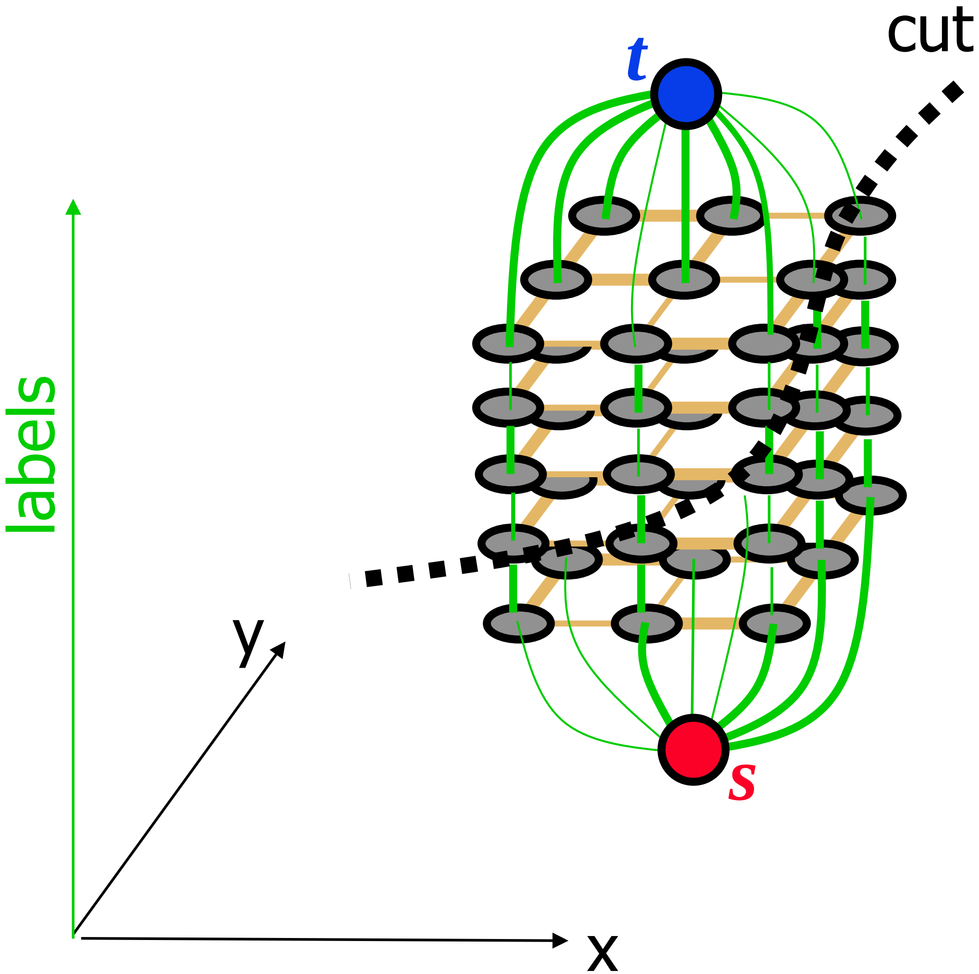
[Ishikawa, PAMI 2003]

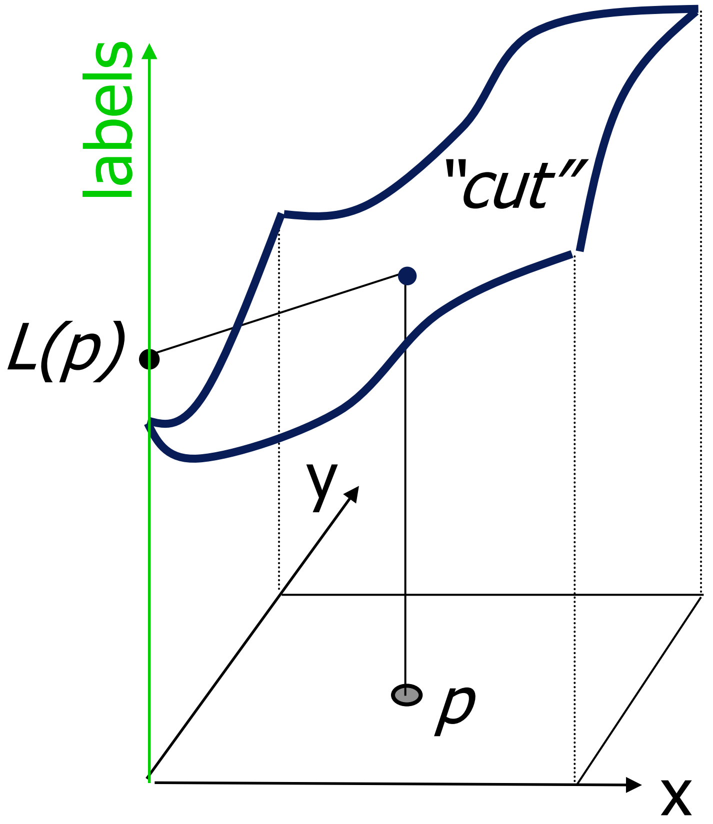
can be globally minimized with graph cuts

Optimizing labelings with *s*-*t* graph cuts [Roy&Cox'98,Ishikawa'98]



Optimizing labelings with *s*-*t* graph cuts [Roy&Cox'98,Ishikawa'98]



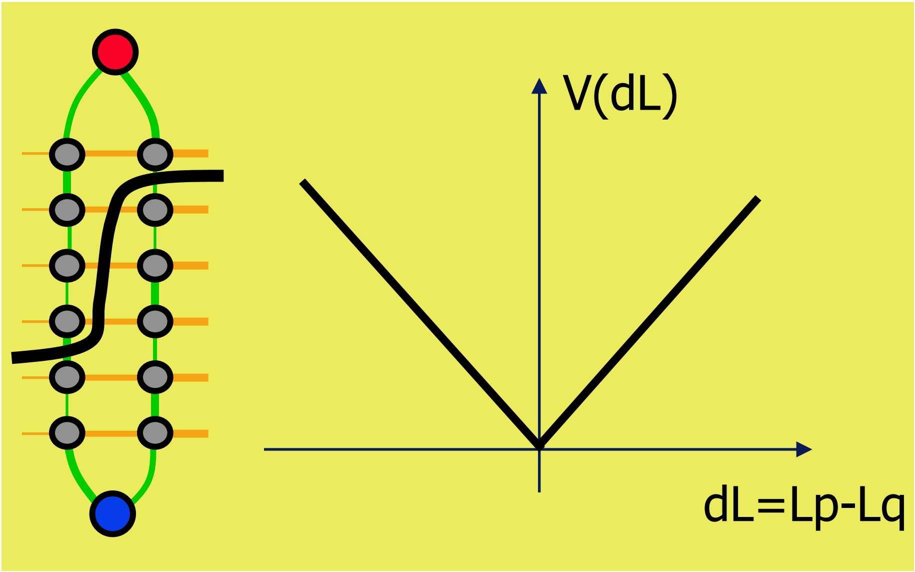


s-*t* graph-cuts for multi-label energy minimization

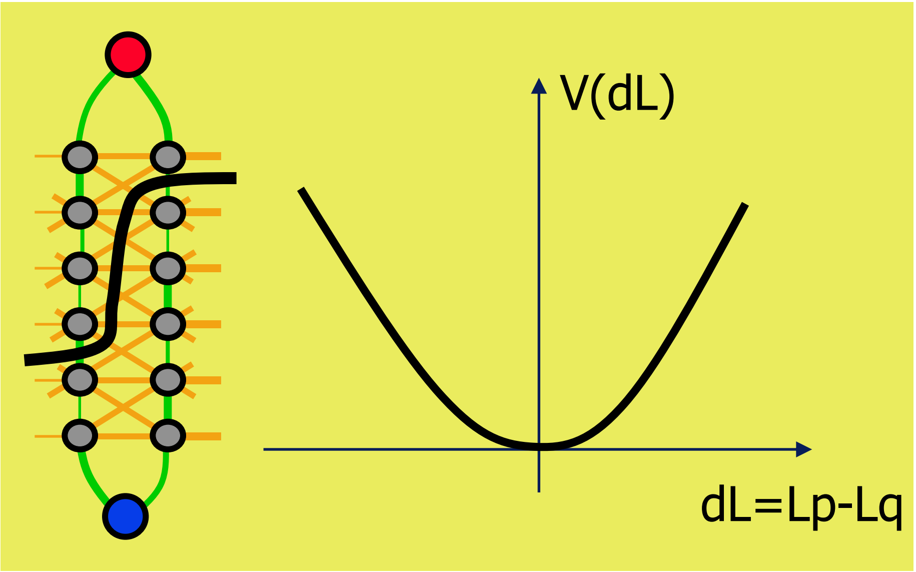
- Ishikawa 1998, 2000, 2003
- Modification of construction by Roy&Cox 1998

$$E(L) = \sum_{p} -D_p(L_p) + \sum_{pq \in N} V(L_p, L_q) \qquad L_p \in R^1$$

Linear interactions



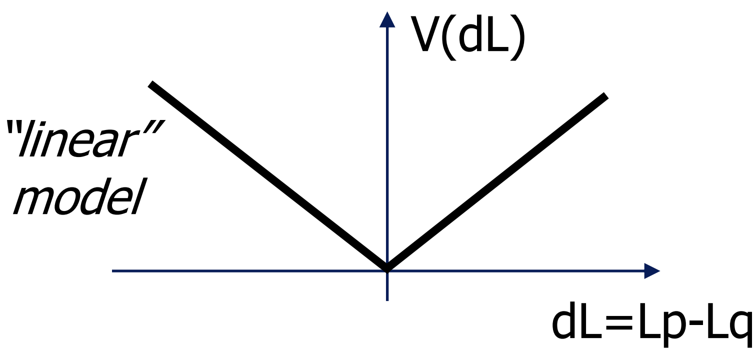
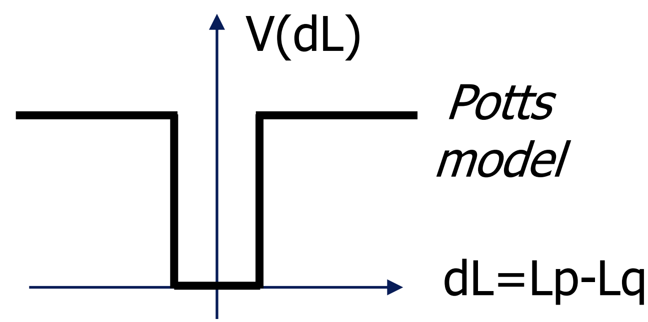
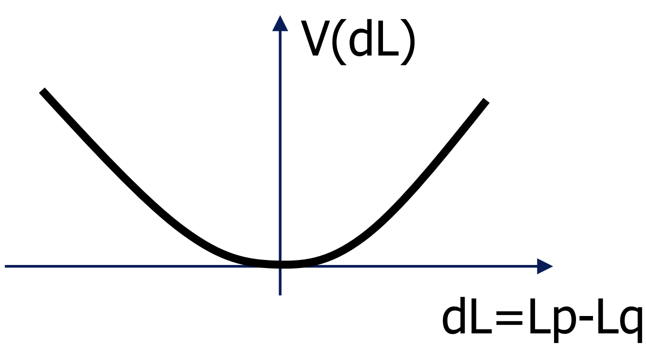
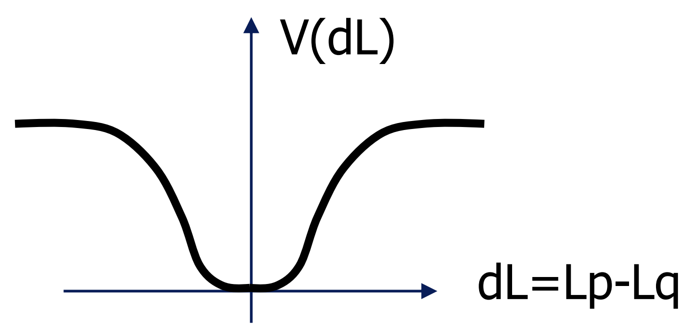
"Convex" interactions



Pixel interactions *V*: “*convex*” vs. “*discontinuity-preserving*”

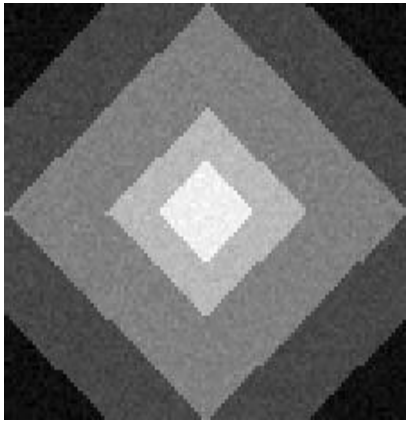
Tractable

NP hard (>2labels)

“Convex” Interactions *V*	Robust “discontinuity preserving” Interactions *V*
 	
 	

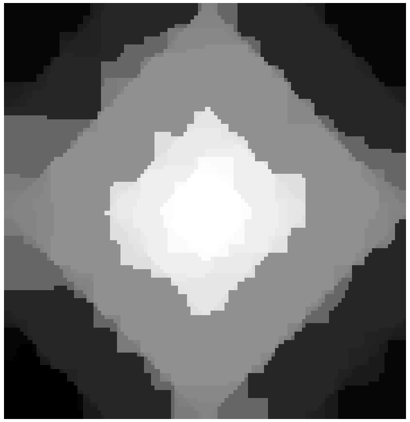
[Blake & Zisserman, 1986]

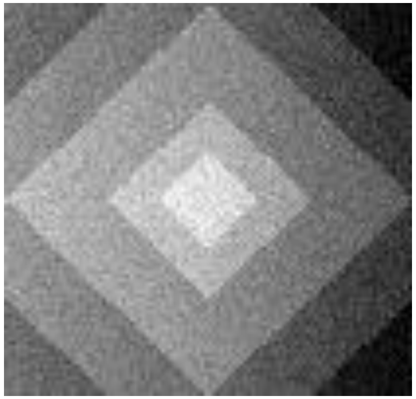
Pixel interactions: “*convex*” vs. “*discontinuity-preserving*”



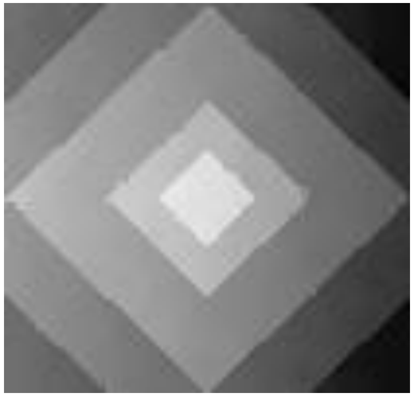
“linear” V

stair-casing





truncated “linear” V



Robust interactions

code

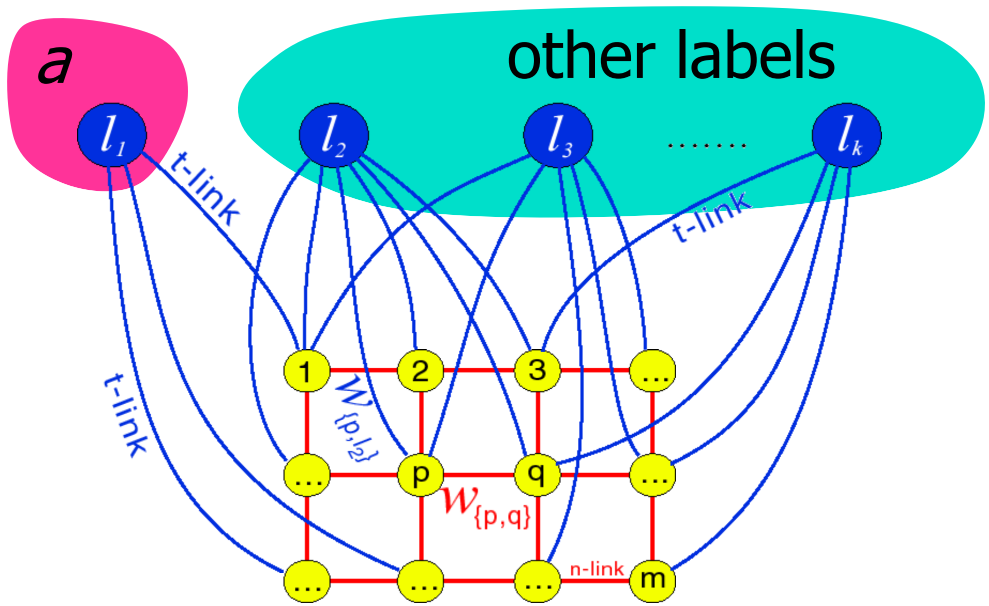
- NP-hard problem (3 or more labels)
  - two labels can be solved via *s-t* cuts
- *a-expansion* approximation algorithm

  (Boykov, Veksler, Zabih 1998, 2001)
  - guaranteed approximation quality (Veksler, thesis 2001)
    - within a factor of 2 from the global minima (Potts model)
- Many other (small or large) move making algorithms

  - a/b swap, jump moves, range moves, fusion moves, etc.
- LP relaxations, message passing, e.g. (TRWS)
- Many other MRF techniques (talk by Ying Niar Wu)
- Variational methods (talk by Mila Nikolova, Daniel Cremers)

a-expansion move

Basic idea is motivated by methods for multi-way cut problem (similar to Potts model)

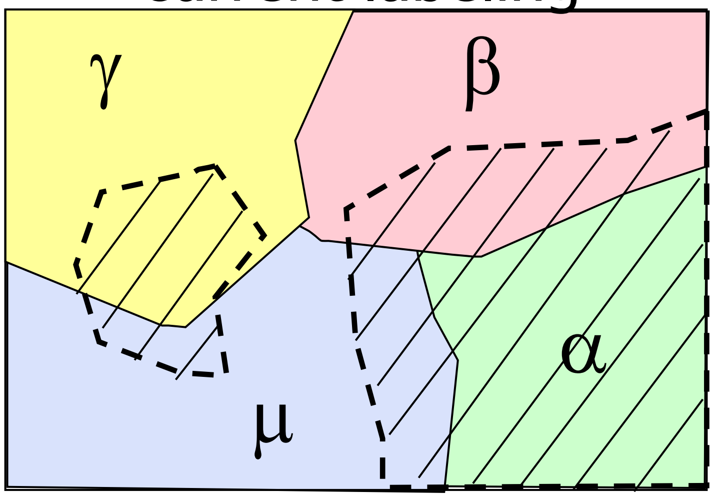


Break computation into a **sequence of binary *s-t* cuts**

a-expansion (binary move) optimizies sumbodular set function

$$\boldsymbol{L} = \{L_p \mid p \in \Omega\}$$

current labeling



expansions correspond to subsets (shaded area)

$$S \subset \Omega$$

$$L'_p(S_p) = \alpha \cdot S_p + L_p \cdot \bar{S}_p$$

$$\bar{S}_p = 1 - S_p$$

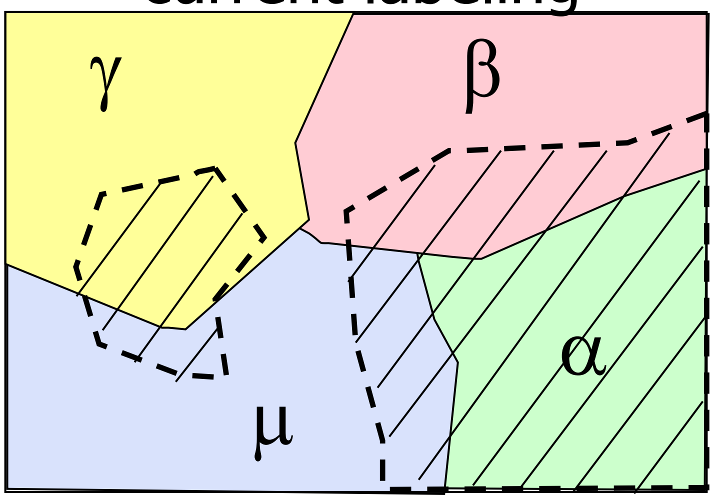
$$\hat{E}(S) = E\big(L'(S)\big) = \sum_p E_p(L'_p) + \sum_{(pq)\in N} E_{pq}(L'_p, L'_q)$$

$$\hat{E}_p(S_p) \qquad \hat{E}_{pq}(S_p, S_q)$$

a-expansion (binary move) optimizies sumbodular set function

$$\boldsymbol{L} = \{L_p \mid p \in \Omega\}$$

current labeling



S_p	0	1
	$E_p(L_p)$	$E_p(\alpha)$

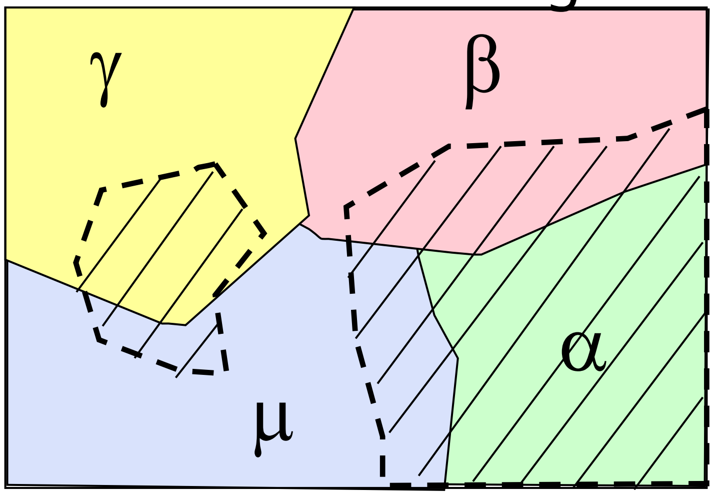
$$\hat{E}(S) = E\big(L'(S)\big) = \sum_p E_p(L'_p) + \sum_{(pq)\in N} E_{pq}(L'_p, L'_q)$$

$\hat{E}_p(S_p)$ $\hat{E}_{pq}(S_p, S_q)$

a-expansion (binary move) optimizies sumbodular set function

$$\boldsymbol{L} = \{L_p \mid p \in \Omega\}$$

current labeling



$S_q \backslash S_p$	0	1
0	$E_{pq}(L_p, L_q)$	$E_{pq}(\alpha, L_q)$
1	$E_{pq}(L_p, \alpha)$	$E_{pq}(\alpha, \alpha)$

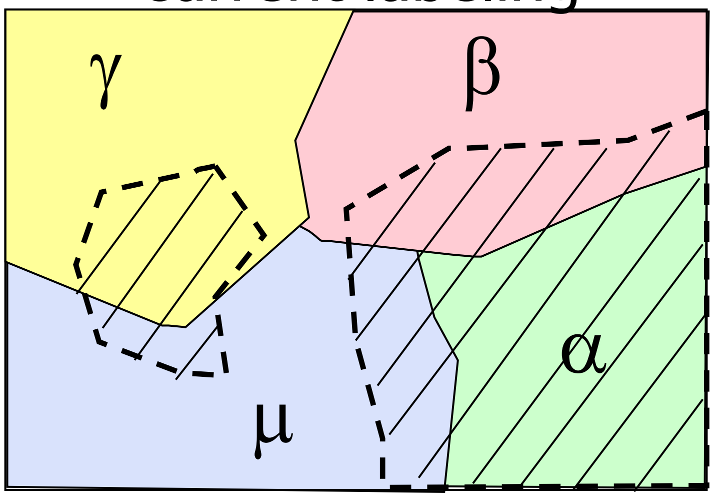
$$\hat{E}(S) = E(L'(S)) = \sum_p E_p(L'_p) + \sum_{(pq)\in N} E_{pq}(L'_p, L'_q)$$

$$\hat{E}_p(S_p) \qquad \hat{E}_{pq}(S_p, S_q)$$

a-expansion (binary move) optimizies sumbodular set function

$$\boldsymbol{L} = \{L_p \mid p \in \Omega\}$$

current labeling



S_q \ S_p	0	1
0	$E_{pq}(L_p, L_q)$	$E_{pq}(\alpha, L_q)$
1	$E_{pq}(L_p, \alpha)$	$E_{pq}(\alpha, \alpha)$

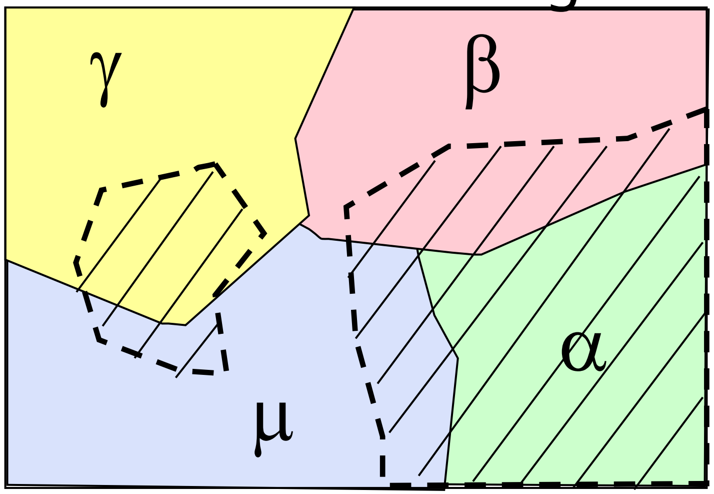
Set function $\hat{E}(S)$ is **submodular** if

$$\hat{E}_{pq}(1,1) + \hat{E}_{pq}(0,0) \le \hat{E}_{pq}(0,1) + \hat{E}_{pq}(1,0)$$

a-expansion (binary move) optimizies sumbodular set function

$\boldsymbol{L} = \{L_p \mid p \in \Omega\}$

current labeling



S_q \ S_p	0	1
0	$E_{pq}(L_p, L_q)$	$E_{pq}(\alpha, L_q)$
1	$E_{pq}(L_p, \alpha)$	$E_{pq}(\alpha, \alpha)$

Set function $\hat{E}(S)$ is **submodular** if

$$\cancel{E_{pq}(\alpha,\alpha)} + E_{pq}(L_p, L_q) \le E_{pq}(L_p, \alpha) + E_{pq}(\alpha, L_q)$$

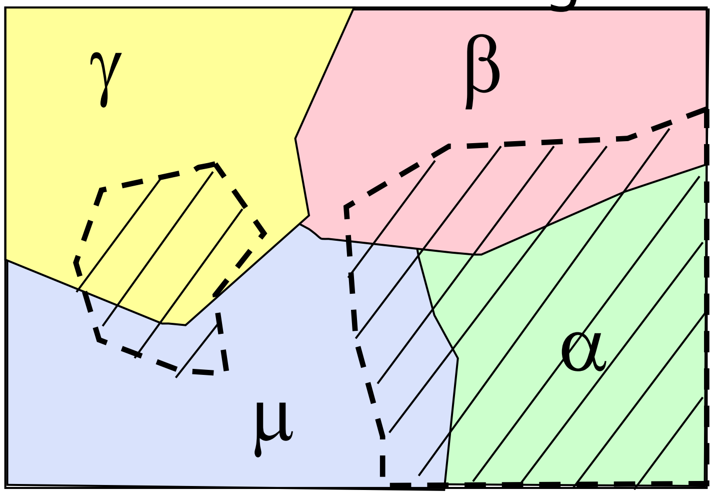
$\cancel{E_{pq}(\alpha,\alpha)}$ = 0

triangular inequality for ||a-b||=E(a,b)

a-expansion (binary move) optimizies sumbodular set function

$$\boldsymbol{L} = \{L_p \mid p \in \Omega\}$$

current labeling



S_q \ S_p	0	1
0	$E_{pq}(L_p, L_q)$	$E_{pq}(\alpha, L_q)$
1	$E_{pq}(L_p, \alpha)$	$E_{pq}(\alpha, \alpha)$

a-expansion moves are **submodular** if $E_{pq}(a,b)$ is a **metric** on the space of labels

[Boykov, Veksler, Zabih, PAMI 2001]

a-expansion moves

In each *a*-expansion a given label "*a*" grabs space from other labels

initial solution

-expansion

-expansion

-expansion

-expansion

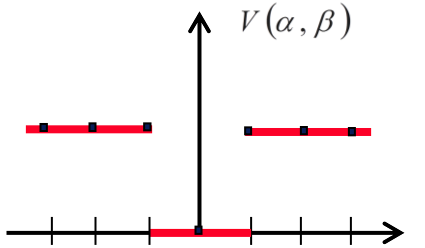
-expansion

-expansion

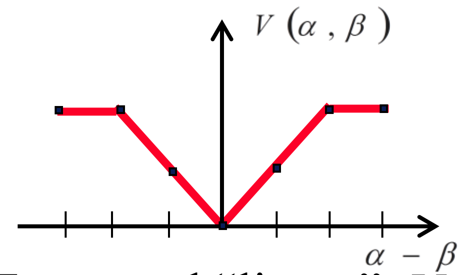
-expansion

For each move we choose expansion that gives the largest decrease in the energy: **binary optimization problem**

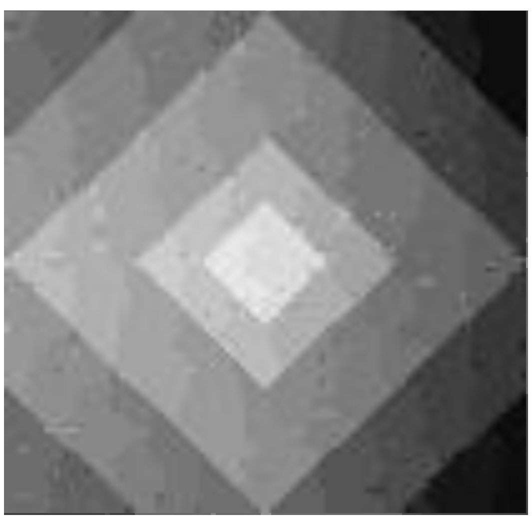
a-expansions: examples of *metric* interactions

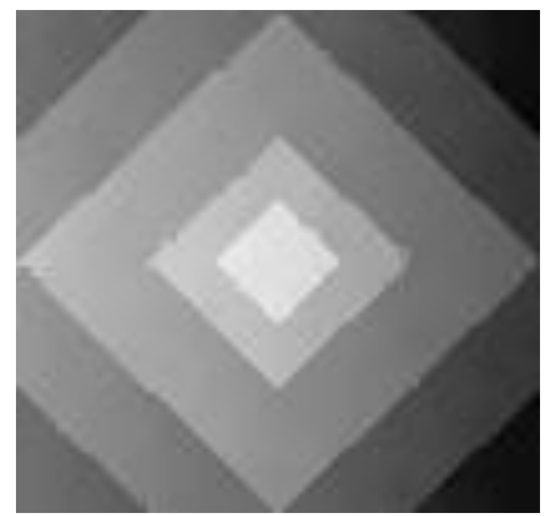


Potts V



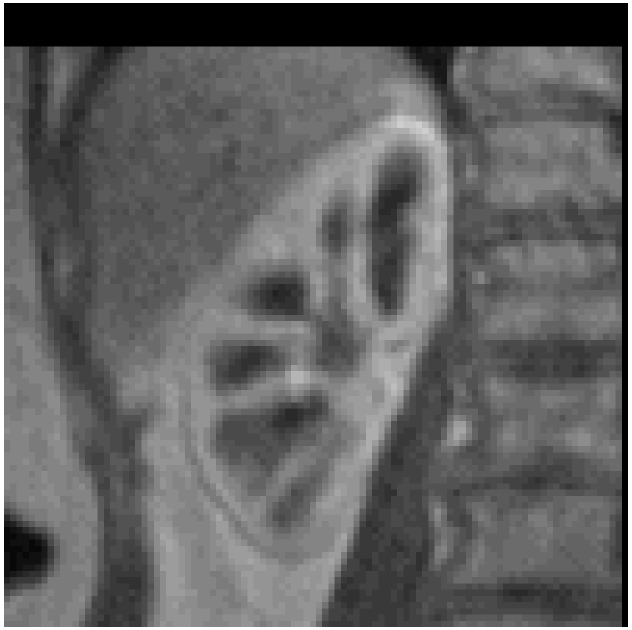
Truncated "linear" V

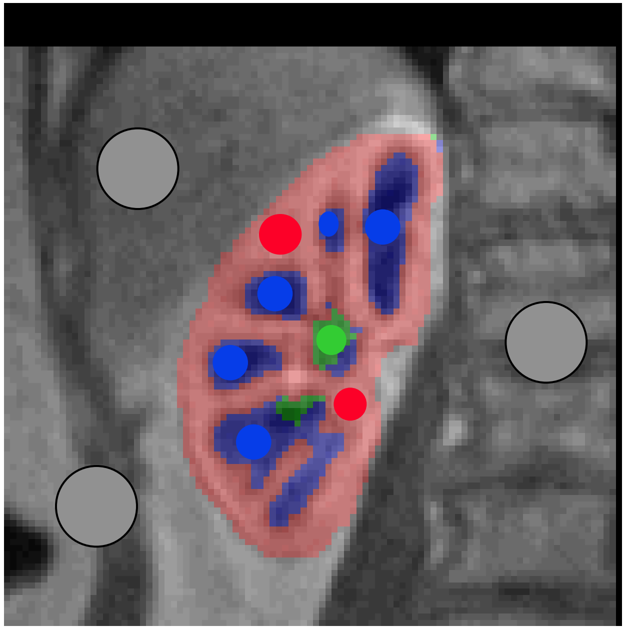




Multi-way graph cuts

Multi-object Extraction





Obvious generalization of binary object extraction technique
(Boykov, Jolly, Funkalea 2004)

Multi-way graph cuts

stereo vision

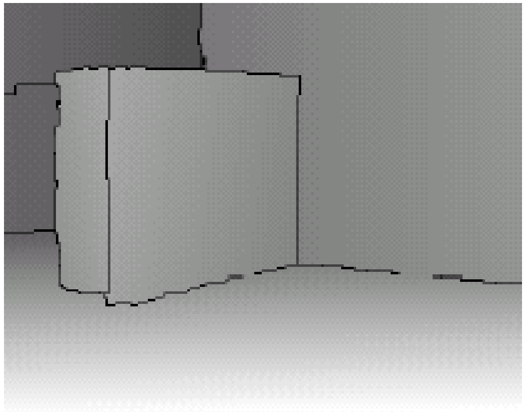
original pair of "stereo" images

depth map

Multi-way graph cuts

Stereo/Motion with slanted surfaces

(Birchfield &Tomasi 1999)



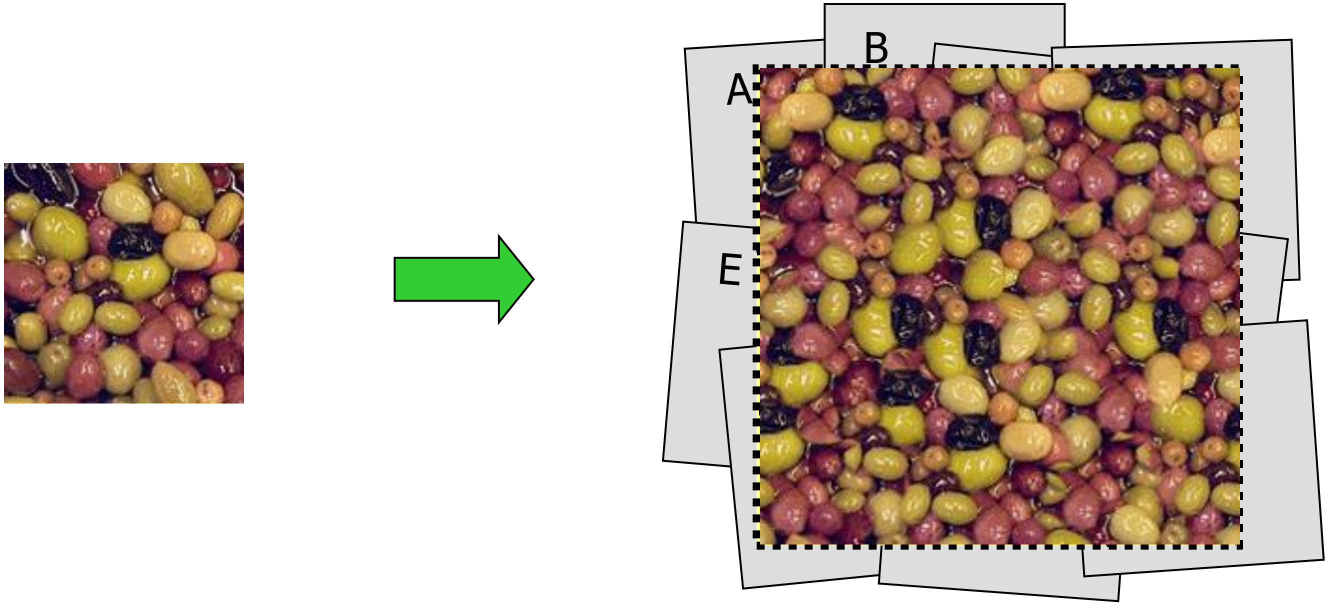
Labels = parameterized surfaces

Block-coordinate descent: models <> segments

Multi-way graph cuts

Graph-cut textures

(Kwatra, Schodl, Essa, Bobick 2003)

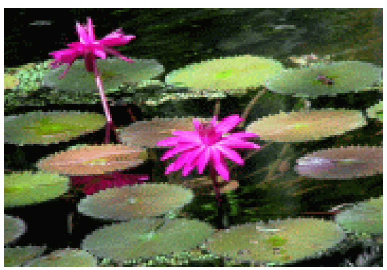


similar to "**image-quilting**" (Efros & Freeman, 2001)

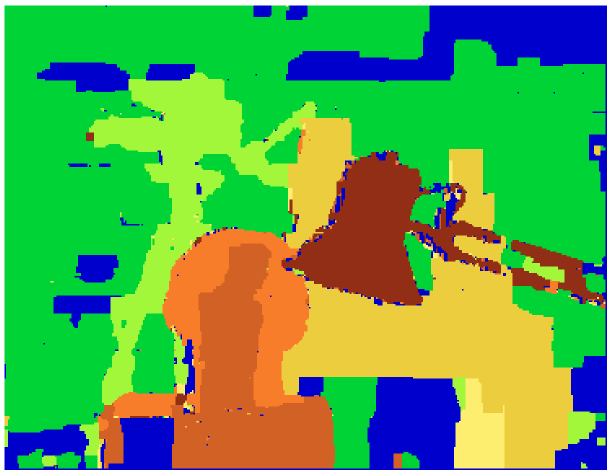
Multi-way graph cuts

Graph-cut textures

(Kwatra, Schodl, Essa, Bobick 2003)



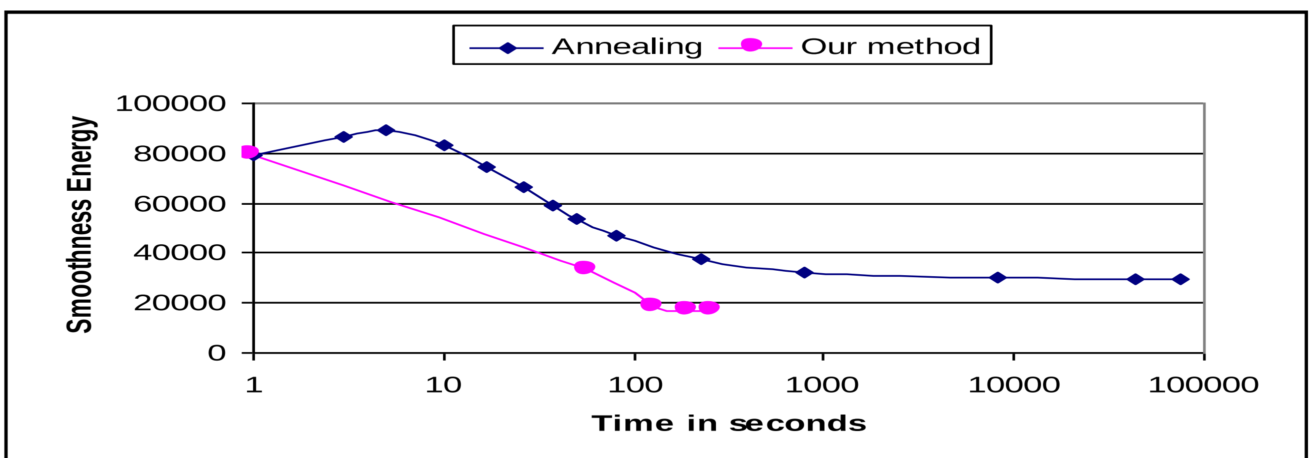
a-expansions vs. simulated annealing



normalized correlation, start for annealing, 24.7% err

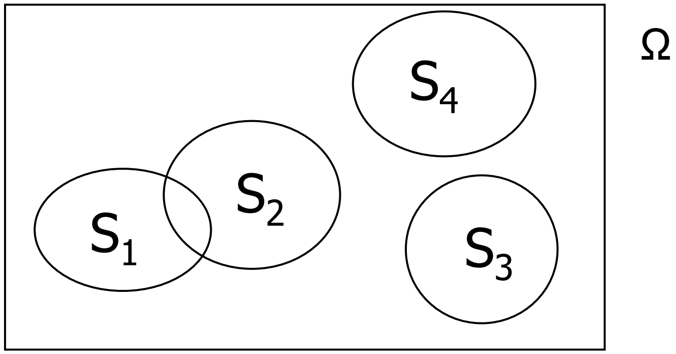
simulated annealing, 19 hours, 20.3% err

a-expansions (BVZ 89,01)
90 seconds, 5.8% err



Multi-set lattices and multi-set functions

Assume set Ω, then $(2^\Omega \times \ldots \times 2^\Omega, \wedge, \vee)$ is a lattice of multi-sets $(S_i) := (S_i)_{i=1}^n$ where each $S_i \subset \Omega$ and



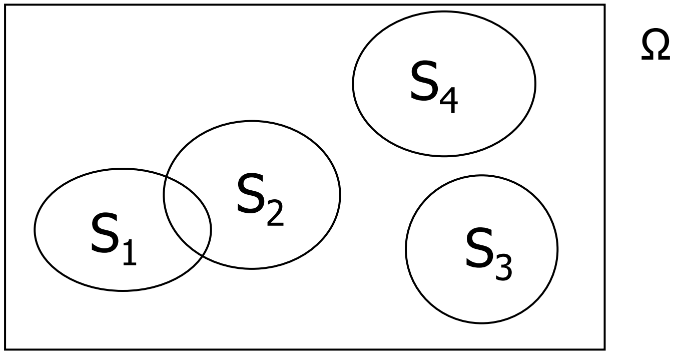
$$\forall (S_i),(T_i) \in 2^\Omega \times \ldots \times 2^\Omega$$

$$(S_i) \wedge (T_i) = (S_i \cap T_i)$$

$$(S_i) \vee (T_i) = (S_i \cup T_i)$$

Multi-set lattices and multi-set functions

Multi-set function $E(S_1, \dots S_n)$ is a mapping $\mathrm{E}: 2^\Omega \times \dots \times 2^\Omega \rightarrow \Re$

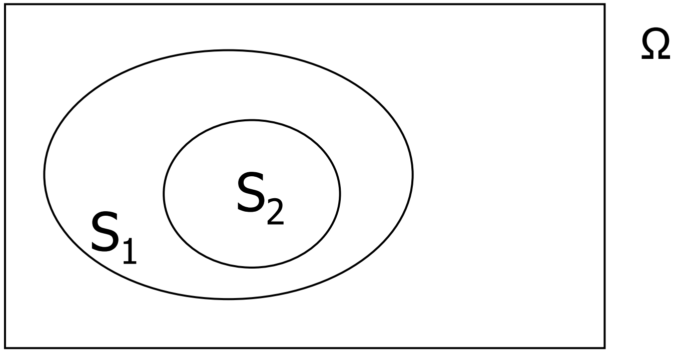


$E(S_1, \dots S_n)$ *is* **submodular** if for any $(S_i),(T_i) \in 2^\Omega \times \dots \times 2^\Omega$

$$E((S_i) \wedge (T_i)) + E((S_i) \vee (T_i)) \leq E((S_i)) + E((T_i))$$

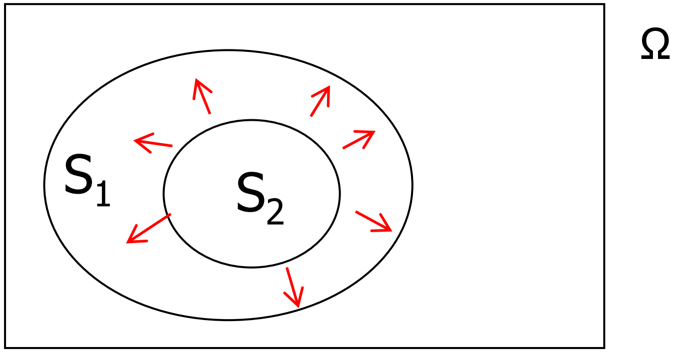
Submodular multi-set functions

Inclusion constraint



Minimum margin constraint

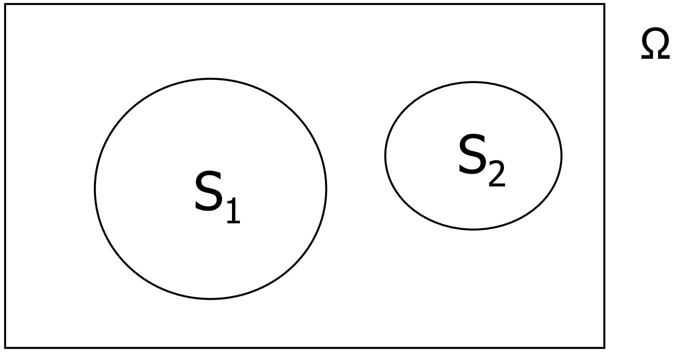
or elastic repulsion



Boundary smoothness (Potts)

Non submodular multi-set functions

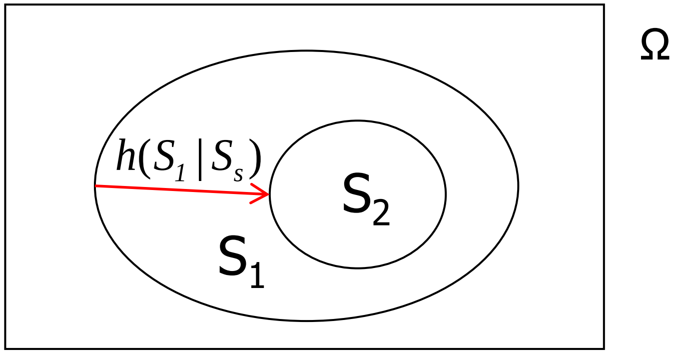
exclusion constraint



Maximum Hausdorf distance constraint

$$h(S_1 \mid S_s) \leq \mathrm{T}$$

or elastic attraction



Reducing to set functions

Theorem [Birkhoff, 1937]: any distrib. lattice $(\mathcal{L}, \wedge, \vee)$ is isomorphic to a set lattice $(2^{\Omega}, \cap, \cup)$ for some Ω.

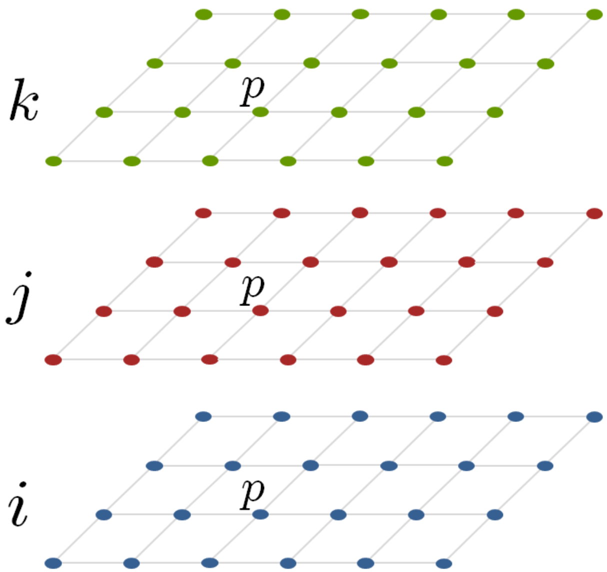
Multi-set functions via graph cuts

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

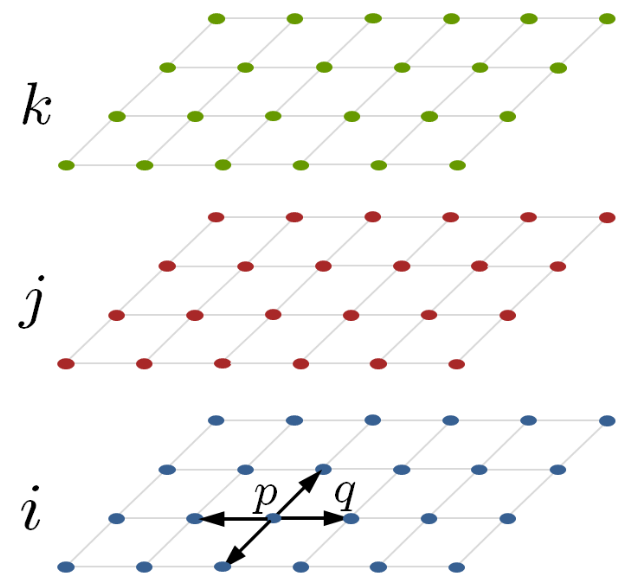
Multi-set functions via graph cuts

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

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



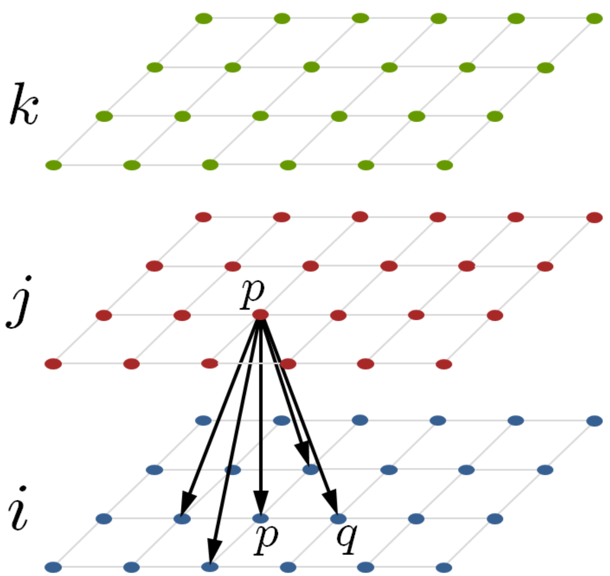
Multi-set functions via graph cuts

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

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



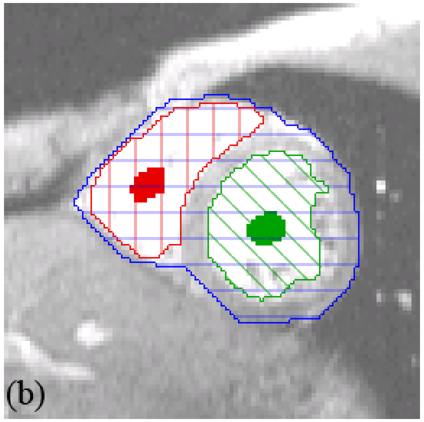
So what *can* we do with graph cuts?

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

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

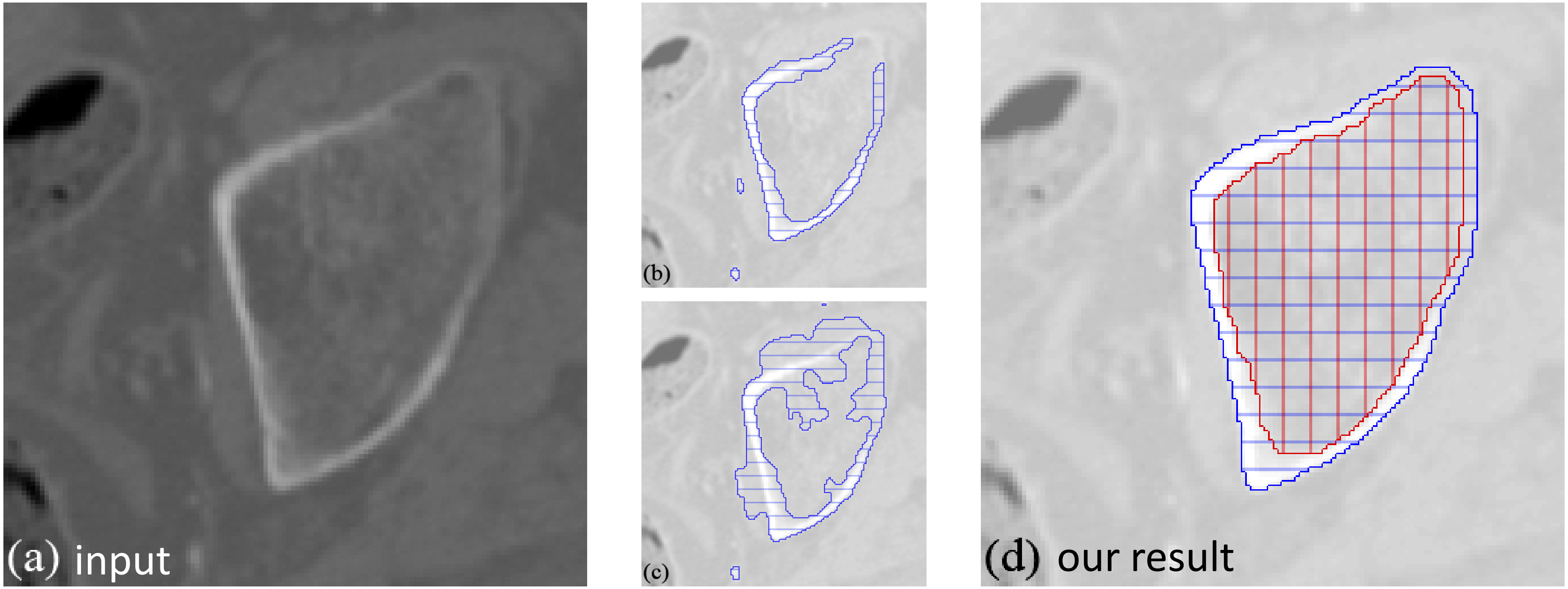
Applications

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

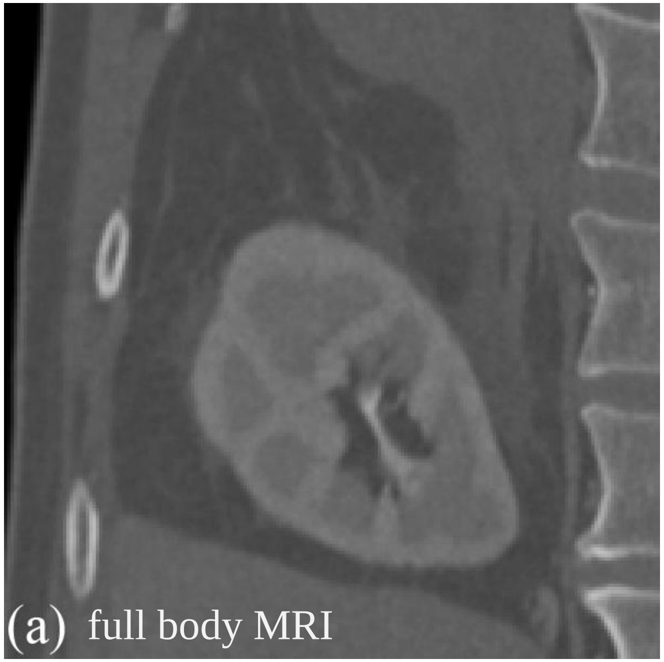


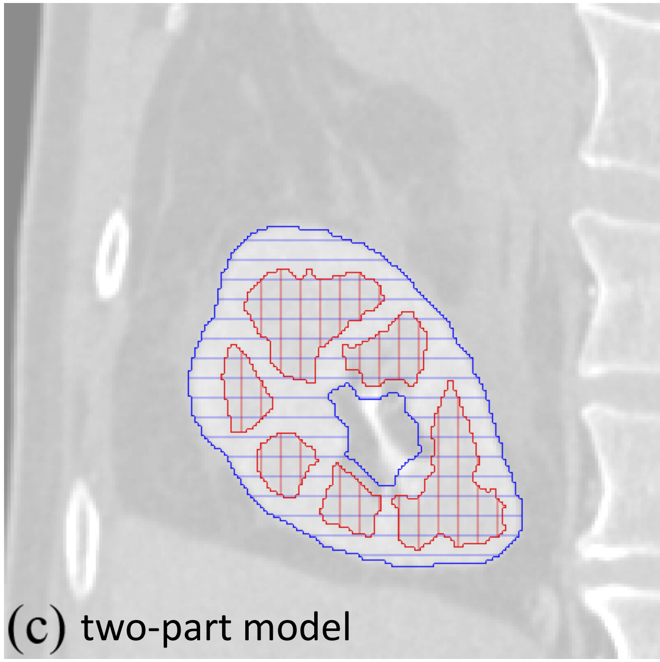
- Scene Layout Estimation
  - Basically just regularize Hoiem-style data terms [4]

Application: Medical

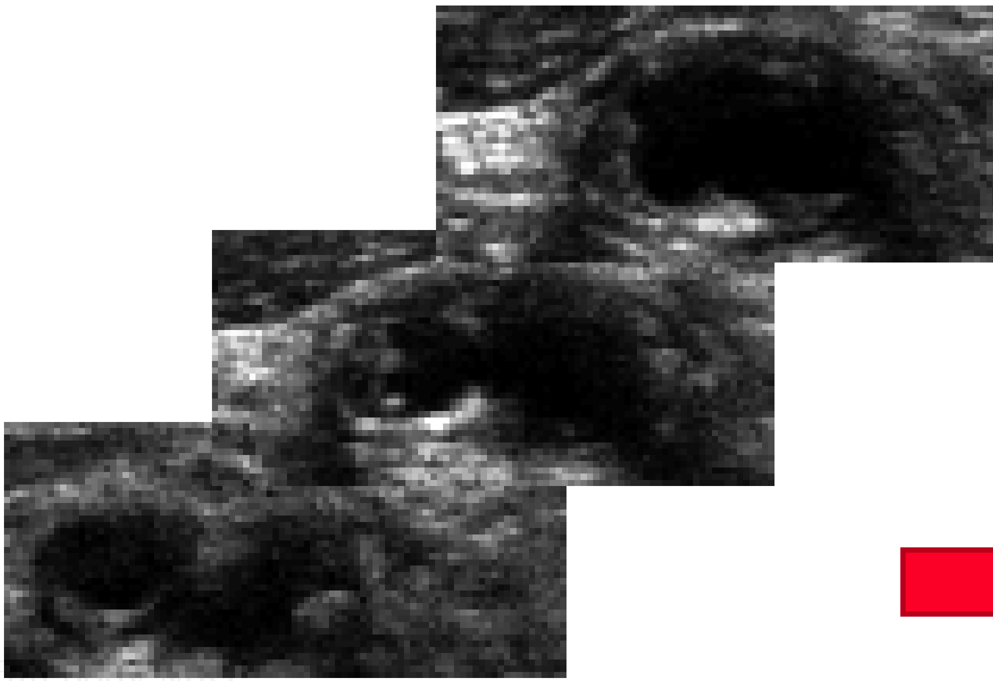


Application: Medical





Application: Medical



full body MRI