ICCV 2007 tutorial on

Discrete Optimization Methods in Computer Vision

part I

Basic overview of graph cuts

Disclaimer

- Can not possibly cover all discrete optimization methods widely used in computer vision in the last 30 years

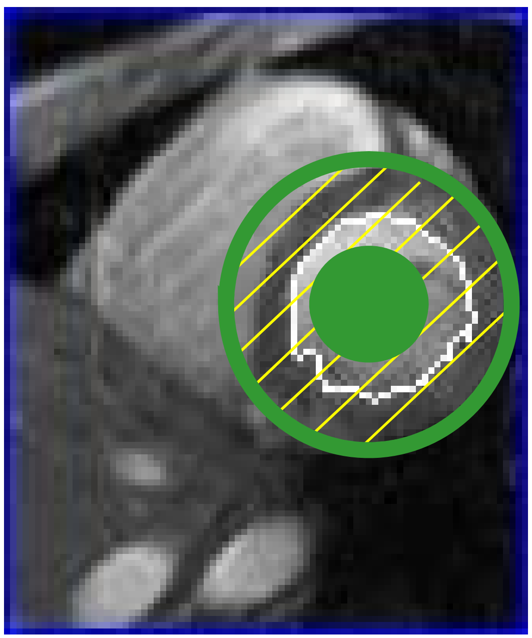
- We mainly concentrate on
  - Discrete energy minimization methods that can be applied to Markov Random Fields with binary or n-labels
    - applicable to a wide spectrum of problems in vision
  - Methods motivated by LP relaxations
    - good bounds on the solutions

Discrete Optimization Methods in Computer Vision

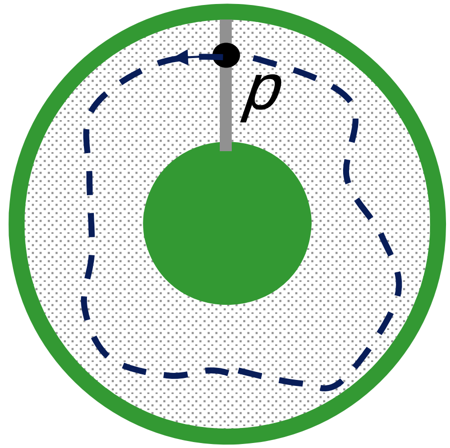
- Part I: **basic overview of graph cuts**
  - binary labeling
    - a few basic examples
    - energy optimization
      - submodularity (discrete view)
      - continuous functionals (geometric view)
      - posterior MRF energy (statistical view)
  - extensions to multi-label problems
    - interactions: convex, robust, metric
    - move-based optimization

2D Graph cut ⇔ *shortest path on a graph*

Example:
find the shortest closed contour in a given domain of a graph

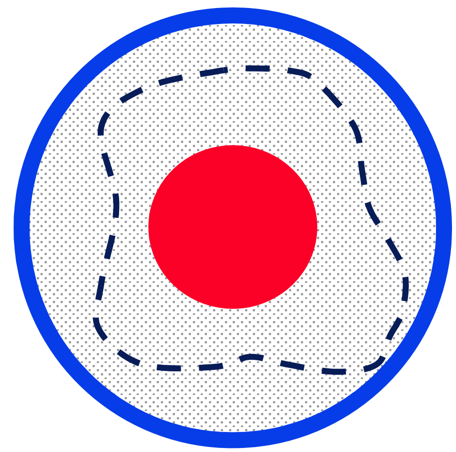


Shortest paths approach
(*live wire*, *intelligent scissors*)



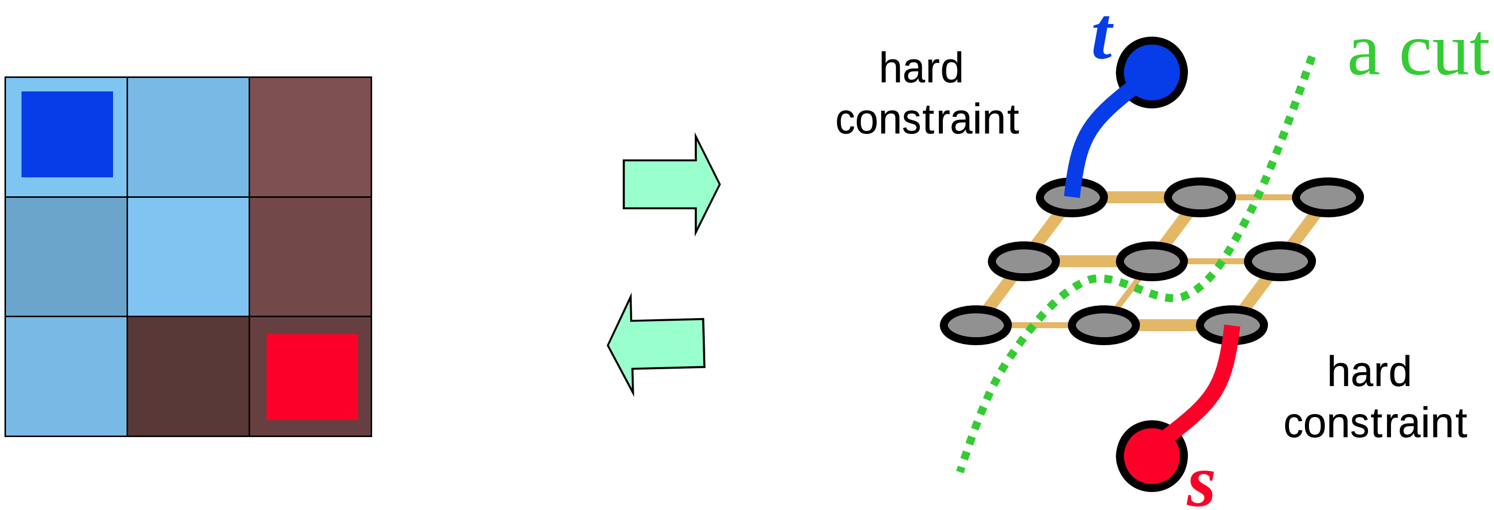
Compute the *shortest path* *p -> p* for a point *p*. Repeat for all points on the gray line. Then choose the optimal contour.

Graph Cuts approach



Compute the *minimum cut* that separates red region from blue region

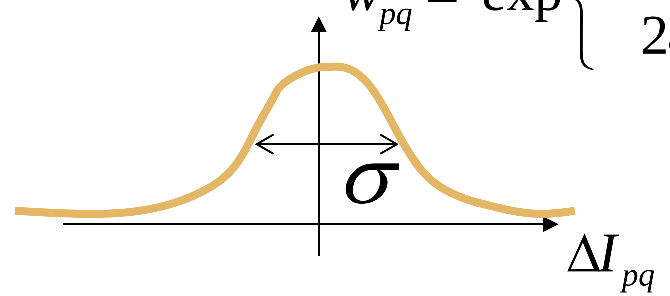
Graph cuts for optimal boundary detection
(B&J, ICCV'01)



Minimum cost cut can be computed in polynomial time

(max-flow/min-cut algorithms)

$$w_{pq} = \exp\left\{-\frac{\Delta I_{pq}}{2\sigma^2}\right\}$$



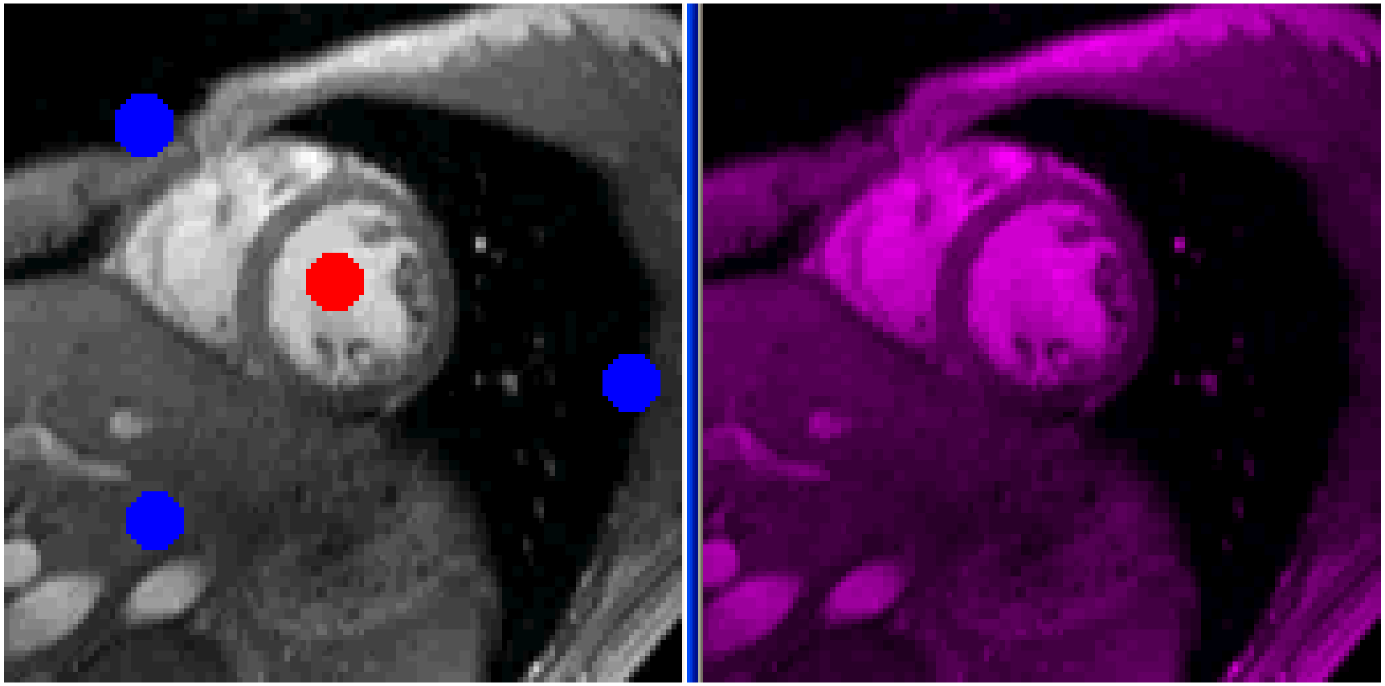
Standard minimum *s-t* cuts algorithms

- Augmenting paths [Ford & Fulkerson, 1962]
- Push-relabel [Goldberg-Tarjan, 1986]

adapted to N-D grids used in computer vision

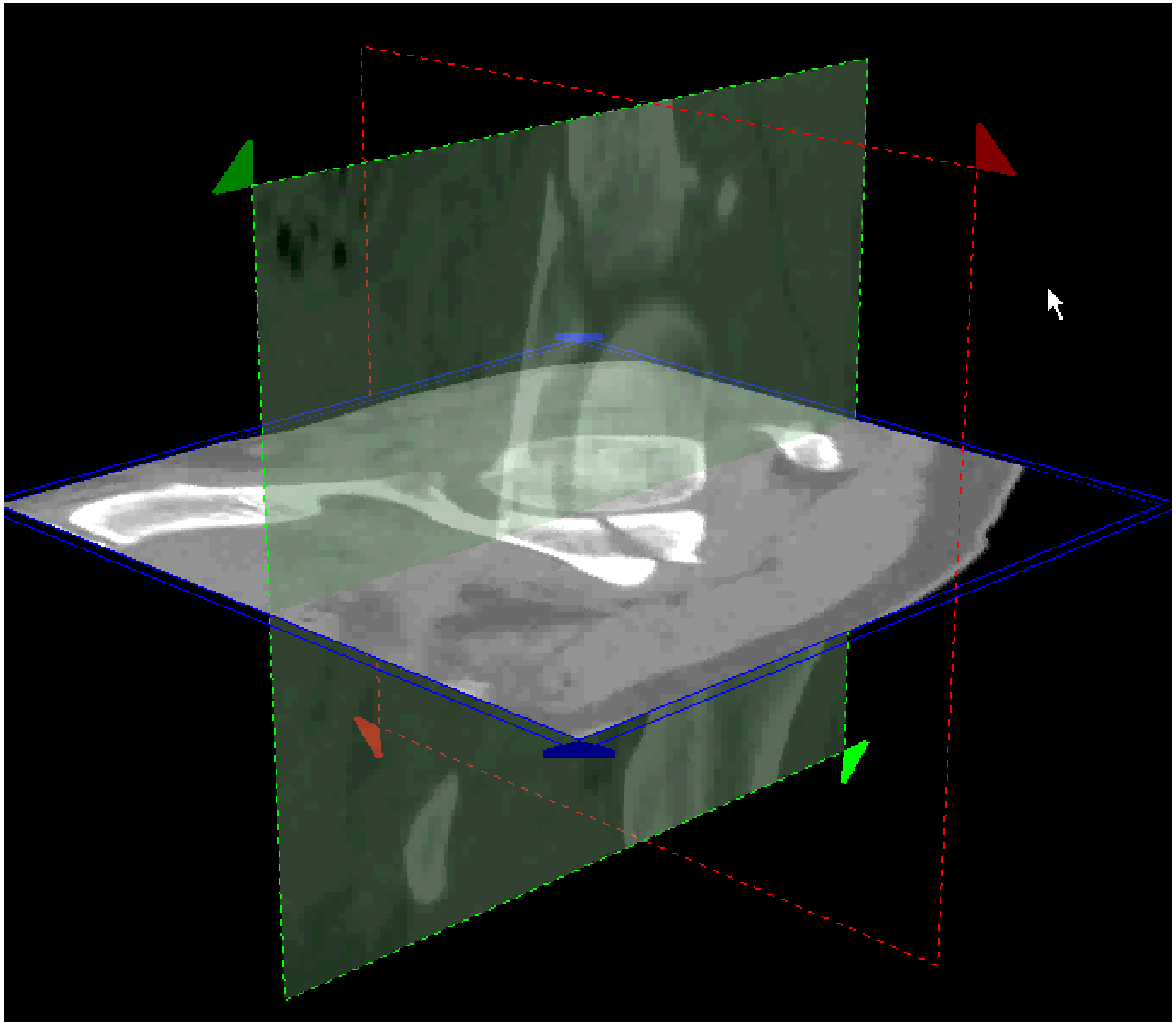
- Tree recycling (dynamic trees) [B&K, 2004]
- Flow recycling (*dynamic cuts*) [Kohli & Torr, 2005]
- Cut recycling (*active cuts*) [Juan & Boykov, 2006]
- Hierarchical methods
  - in search space [Lombaert et al., CVPR 2005]
  - in edge weights (*capacity scaling*) [Juan et al., ICCV07]

Optimal boundary in 2D



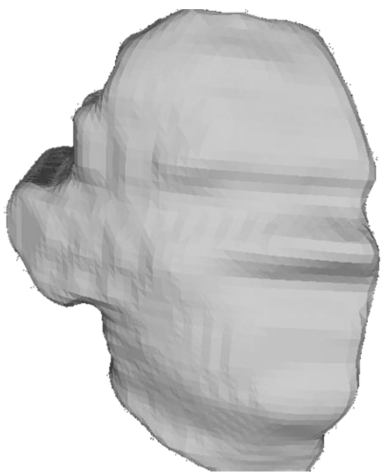
“max-flow = min-cut”

Optimal boundary in 3D



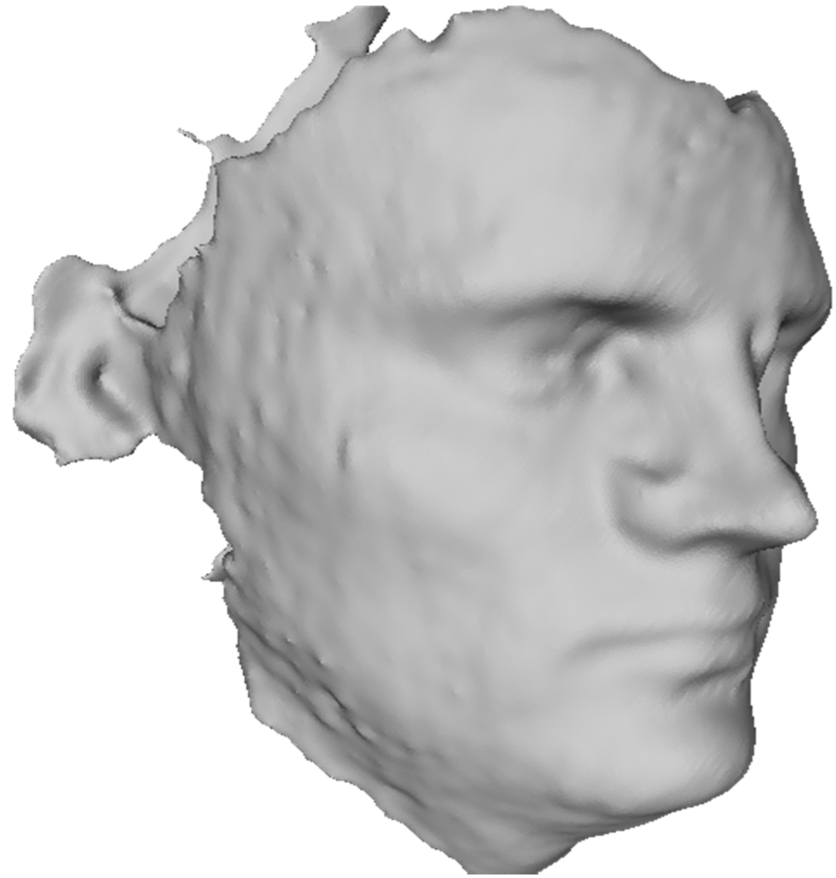
3D bone segmentation (real time screen capture)

Graph cuts applied to multi-view reconstruction



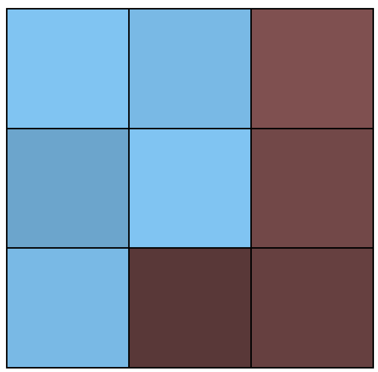
visual hull
(silhouettes)

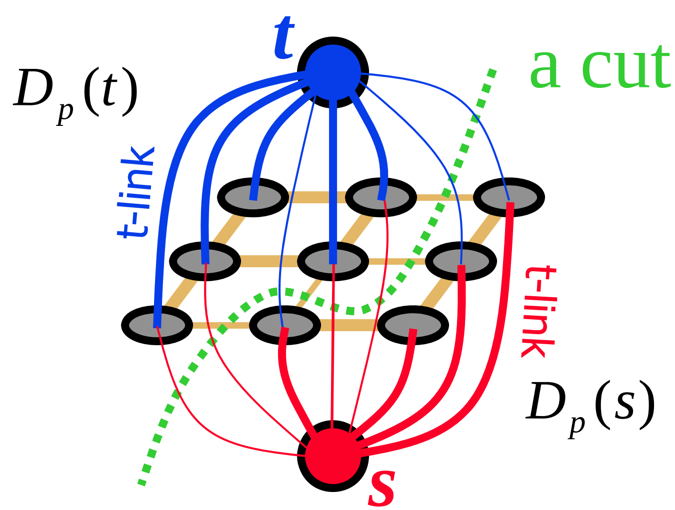
surface of good photoconsistency



CVPR’05 slides from Vogiatzis, Torr, Cippola

Adding regional properties
(B&J, ICCV'01)





regional bias example

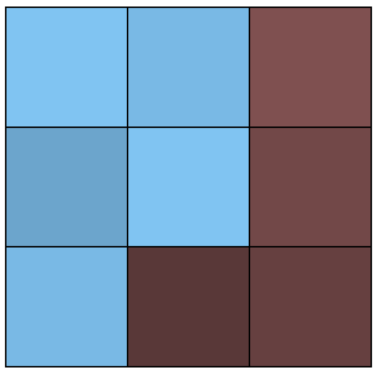
suppose I^s and I^t are given
"expected" intensities
of object and background

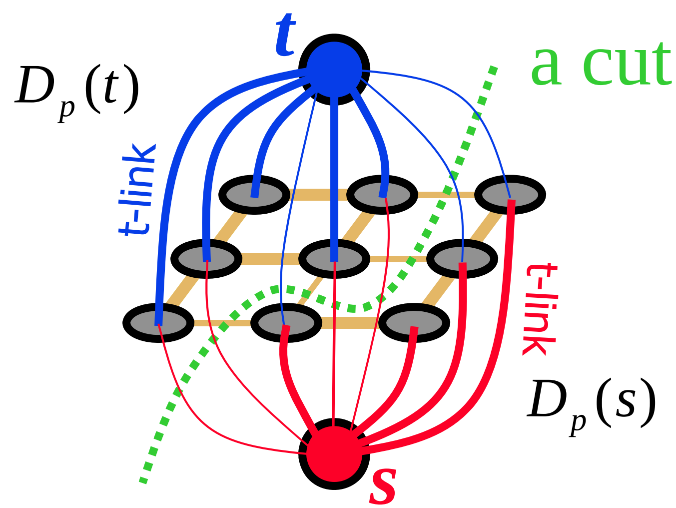
$$D_p(s) \propto \exp\left(-\| I_p - I^s \|^2 / 2\sigma^2\right)$$

$$D_p(t) \propto \exp\left(-\| I_p - I^t \|^2 / 2\sigma^2\right)$$

NOTE: hard constrains are not required, in general.

Adding regional properties
(B&J, ICCV'01)





"expected" intensities of
object and background
I^s and I^t
can be re-estimated

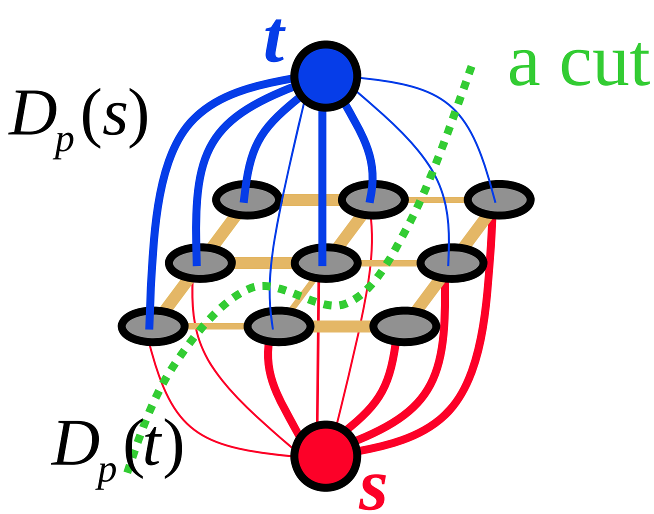
$$D_p(s) \propto \exp\left(-\| I_p - I^s \|^2 / 2\sigma^2\right)$$

$$D_p(t) \propto \exp\left(-\| I_p - I^t \|^2 / 2\sigma^2\right)$$

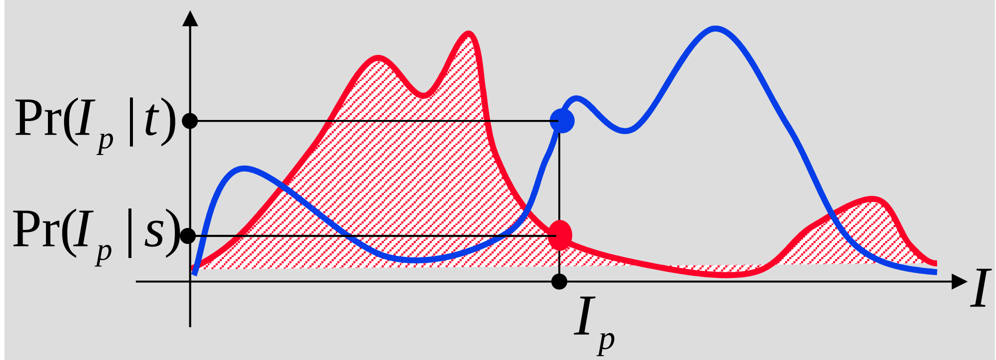
EM-style optimization of piece-vice constant *Mumford-Shah* model

Adding regional properties
(B&J, ICCV'01)

More generally, regional bias can be based on any intensity models of object and background

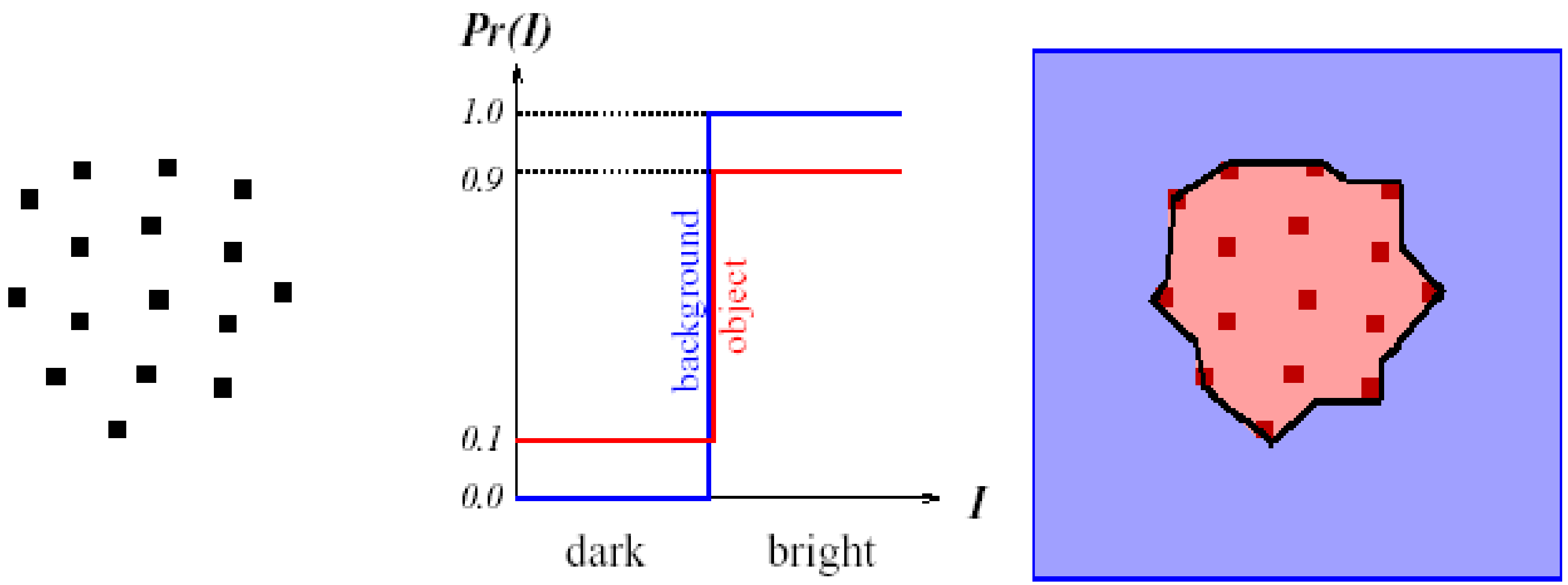


$$D_p(L_p) = -\ln \Pr(I_p \mid L_p)$$



given object and background intensity histograms

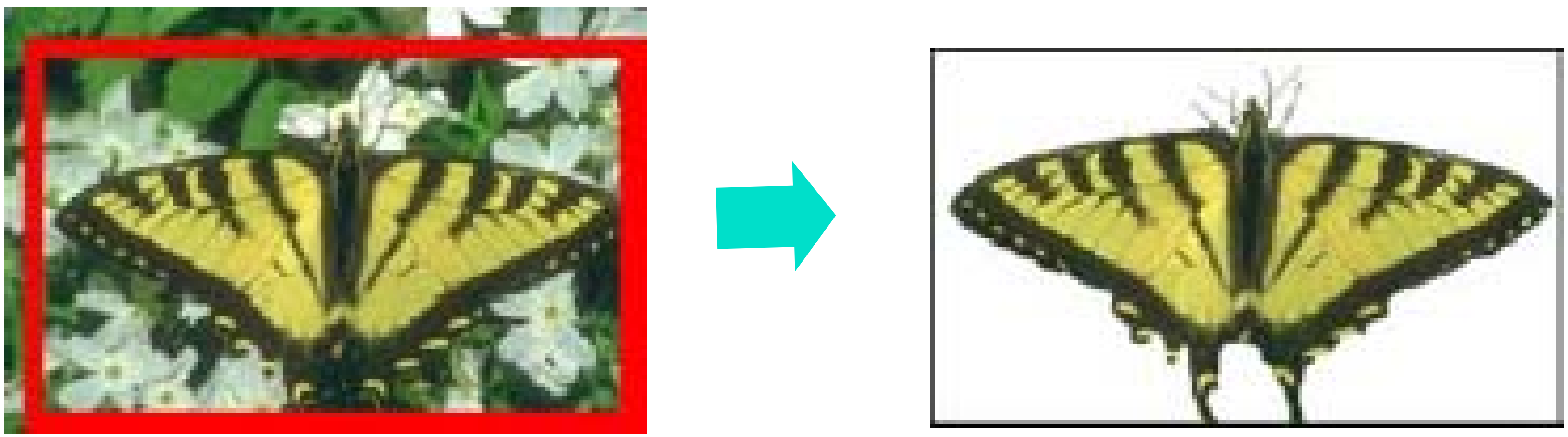
Adding regional properties
(B&J, ICCV'01)



(a) Original image (b) Intensity histograms (c) Optimal segmentation

Iterative learning of regional color-models

- GMMRF cuts (Blake et al., ECCV04)
- Grab-cut (Rother et al., SIGGRAPH 04)



parametric regional model – Gaussian Mixture (GM)
designed to guarantee convergence

At least three ways to look at energy of graph cuts

I: Binary submodular energy

II: Approximating continuous surface functionals

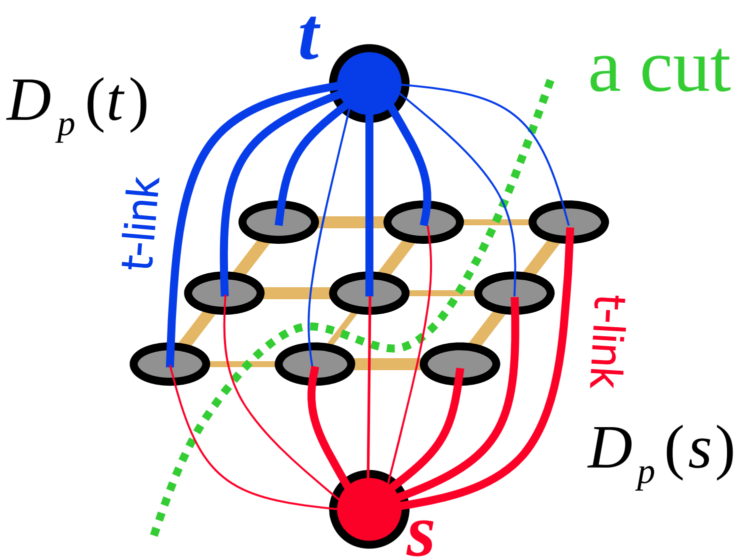
III: Posterior energy (MAP-MRF)

Simple example of energy

Regional term Boundary term

$$E(L) = \sum_p D_p(L_p) + \sum_{pq \in N} w_{pq} \cdot \delta(L_p \neq L_q)$$

t-links **n-links**



$$L_p \in \{s, t\}$$

binary object segmentation

Graph cuts for minimization of submodular <u>binary</u> energies I

Regional term Boundary term

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

t-links **n-links** $L_p \in \{s,t\}$

- Characterization of **binary** energies that can be globally minimized by *s*-*t* graph cuts [Boros&Hummer, 2002, K&Z 2004]

E(L) can be minimized by *s-t* graph cuts	$\Leftrightarrow$	$E(s,s)+E(t,t) \le E(s,t)+E(t,s)$
		Submodularity ("convexity")

- **Non-submodular cases** can be addressed with some optimality guarantees, e.g. ***QPBO*** algorithm
  - (see Boros&Hummer, 2002, Tavares et al. 06, Rother et al. 07)

Graph cuts for minimization of continuous surface functionals II

$$E(C) = \int_C g(\cdot)\, ds + \int_C \langle \vec{\mathbf{N}}, \vec{\mathbf{v}}_x \rangle\, ds + \int_{\Omega(C)} f(x)\, dp$$

Geometric length
any convex,
symmetric metric ***g***
e.g. Riemannian

Flux
any vector field **v**

Regional bias
any scalar function ***f***

- Characterization of energies of <u>**binary**</u> cuts *C* as functionals of continuous surfaces

[B&K, ICCV 2003]

[K&B, ICCV 2005]

One extension using parametric max-flow methods

- optimization of ratio functionals

$$E(C) = \frac{\int_C \left\langle \vec{\mathbf{N}}, \vec{\mathbf{v}}_x \right\rangle ds}{\int_C g(\cdot)\, ds}$$

$$E(C) = \frac{\int_C g(\cdot)\, ds}{\int_{\Omega(C)} f(x)\, dp}$$

- In 2D can use DP [Cox et al'96, Jermyn&Ishikawa'01]
- In 3D, see a poster on Tuesday (Kolmogorov, Boykov, Rother)

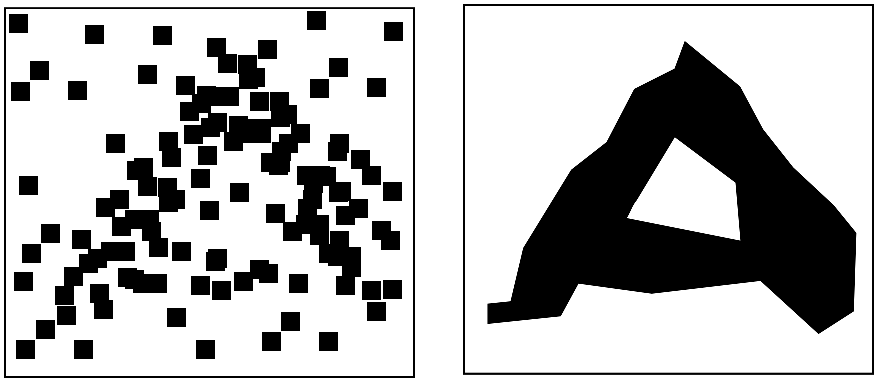
Graph cuts for minimization of posterior energy III

- Greig at al. [IJRSS, 1989]
  - Posterior energy (MRF, Ising model)

$$E(L) \;=\; \sum_{p} -\ln \Pr(D_p \mid L_p) \;+\; \sum_{pq \in N} V_{pq}(L_p, L_q)$$

Likelihood (data term) **Spatial prior (regularization)**

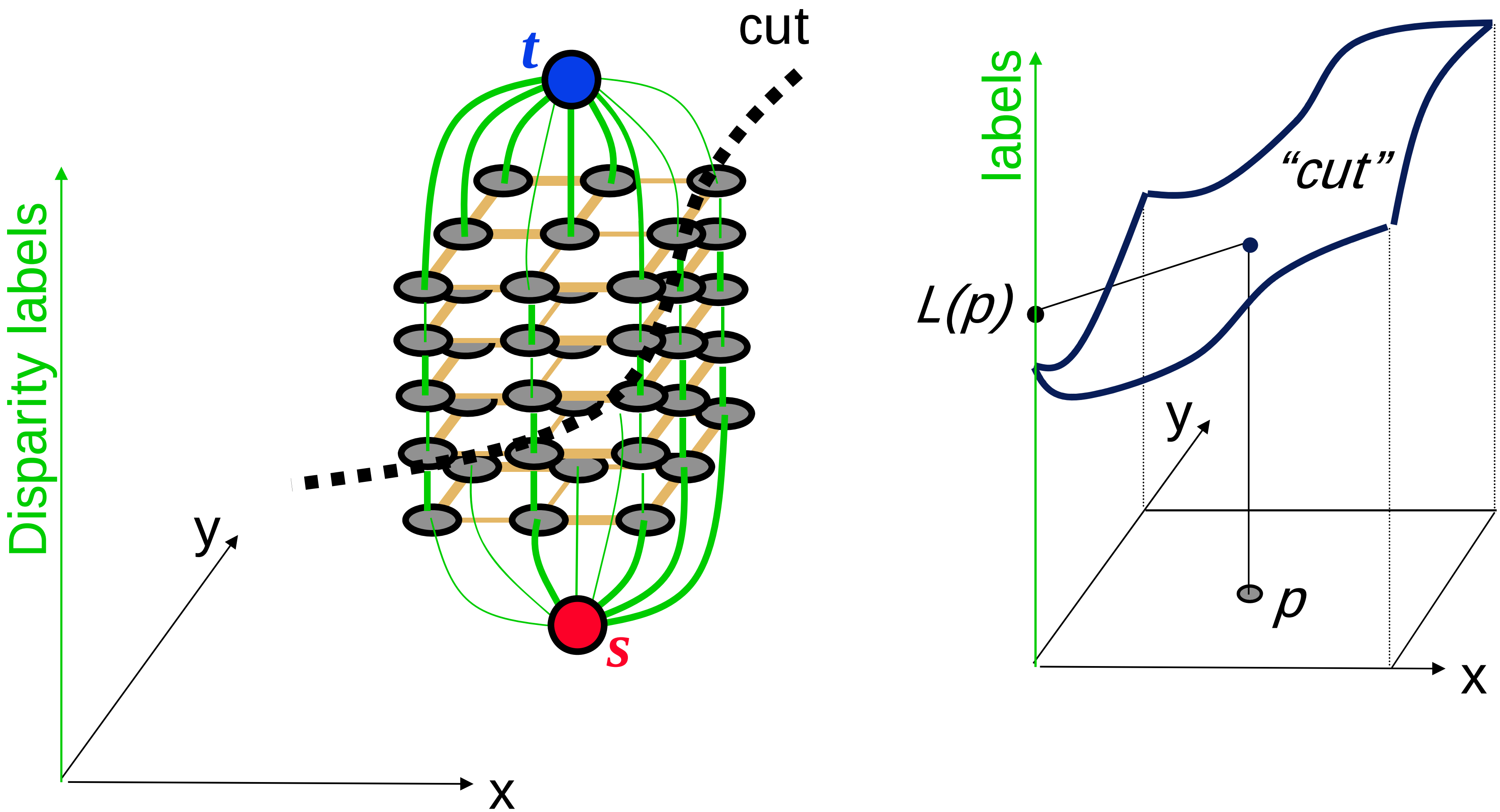
$$L_p \in \{s, t\}$$



Example: binary image restoration

Graph cuts algorithms can minimize multi-label energies as well

Multi-scan-line stereo with *s*-*t* graph cuts (Roy&Cox'98)

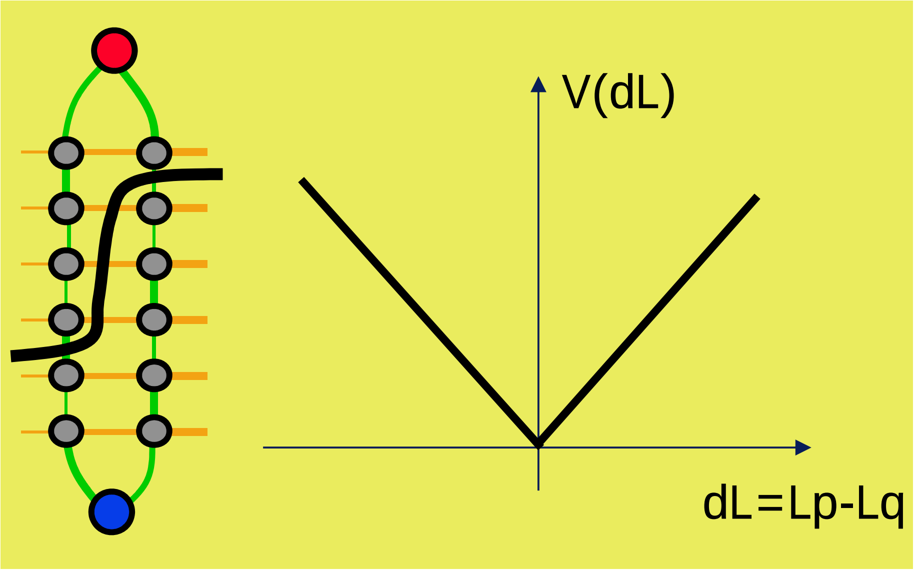


s-*t* graph-cuts for multi-label energy minimization

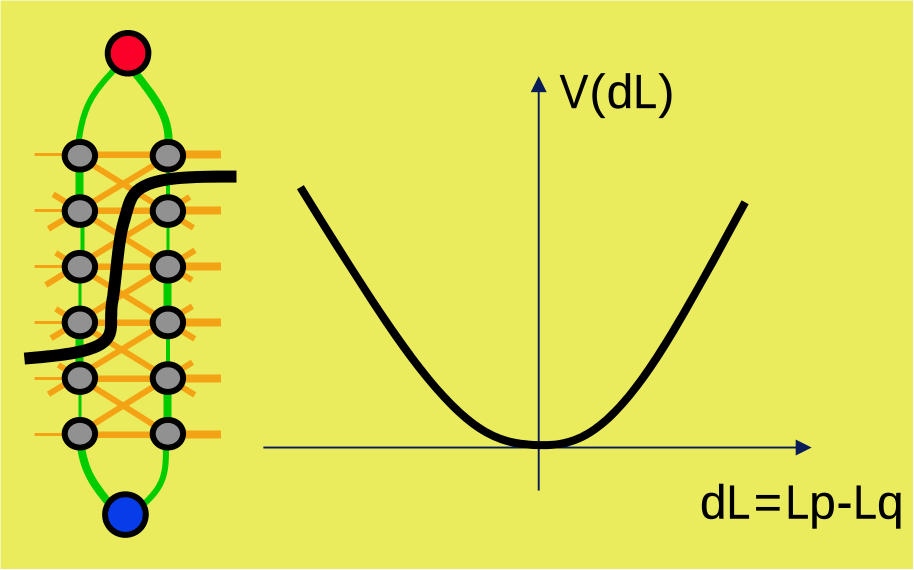
- Ishikawa 1998, 2000, 2003
- Generalization of construction by Roy&Cox 1998

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

Linear interactions

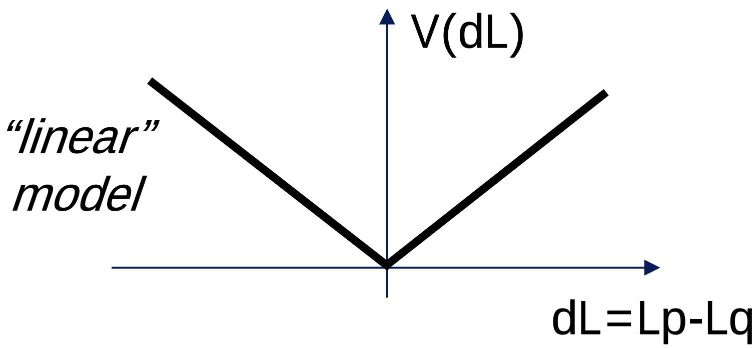


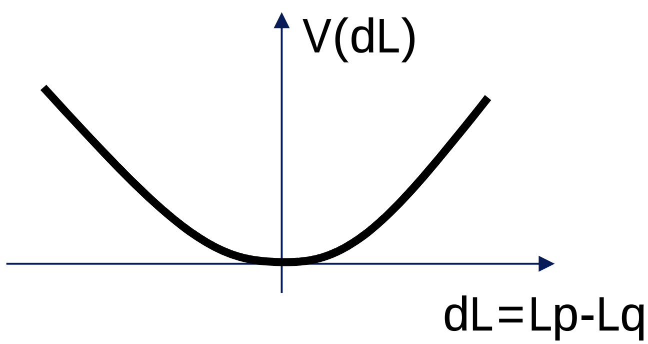
"Convex" interactions

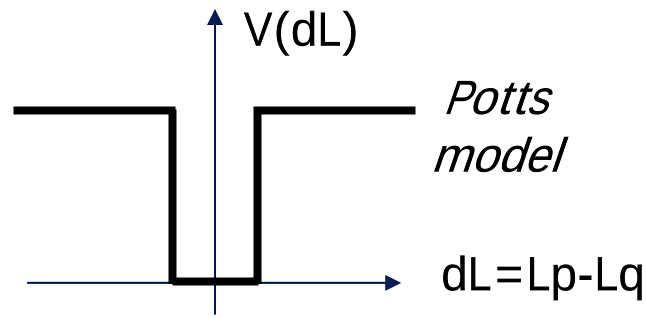


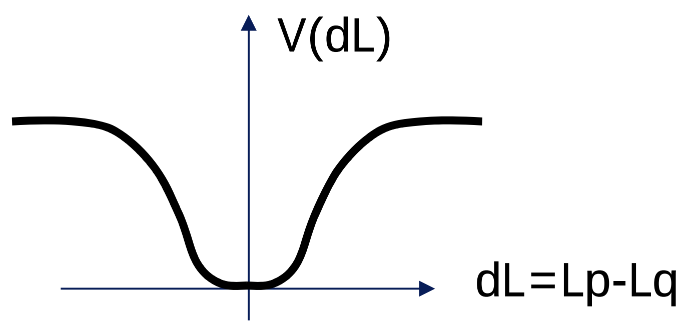
Pixel interactions *V*:
“*convex*” vs. “*discontinuity-preserving*”

“Convex” Interactions *V*	Robust or “discontinuity preserving” Interactions *V* (weak membrane models, see a book by Blake and Zisserman, 87)

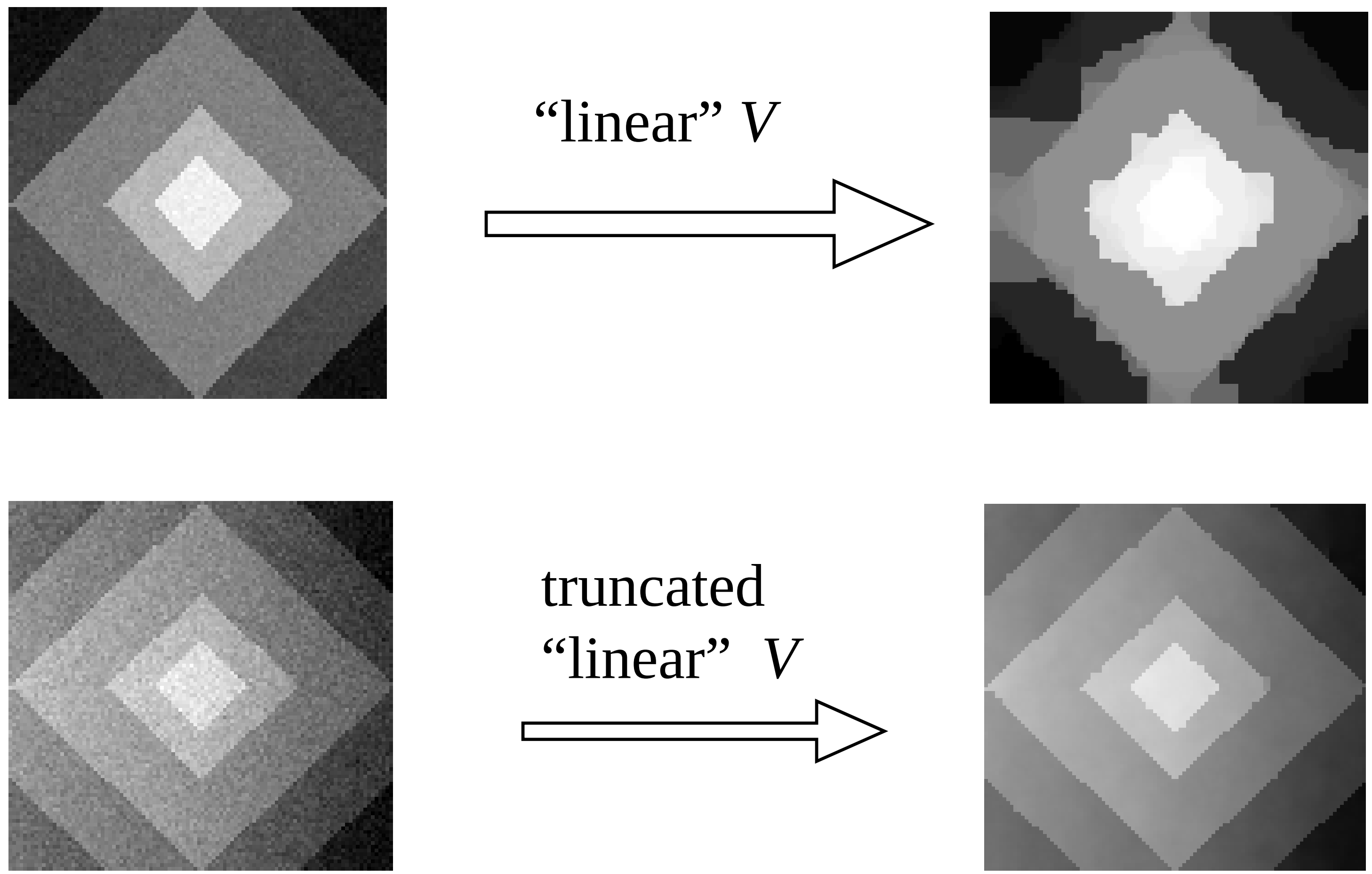








Pixel interactions:
“*convex*” vs. “*discontinuity-preserving*”



Robust interactions

- NP-hard problem (3 or more labels)
  - two labels can be solved via *s-t* cuts (Greig at. al., 1989)
- *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)
  - applies to a wide class of energies with robust interactions
    - Potts model (BVZ 1989)
    - "metric" interactions (BVZ 2001)
    - can be extended to arbitrary interactions with weaker guarantees
      - truncation (Kolmogorov et al. 2005)
      - QPBO (Boros and Hummer, 2002)
- Other "move" algorithms (e.g. *a-b* swap, jump-moves)
- More is coming later in this tutorial

a-expansion algorithm

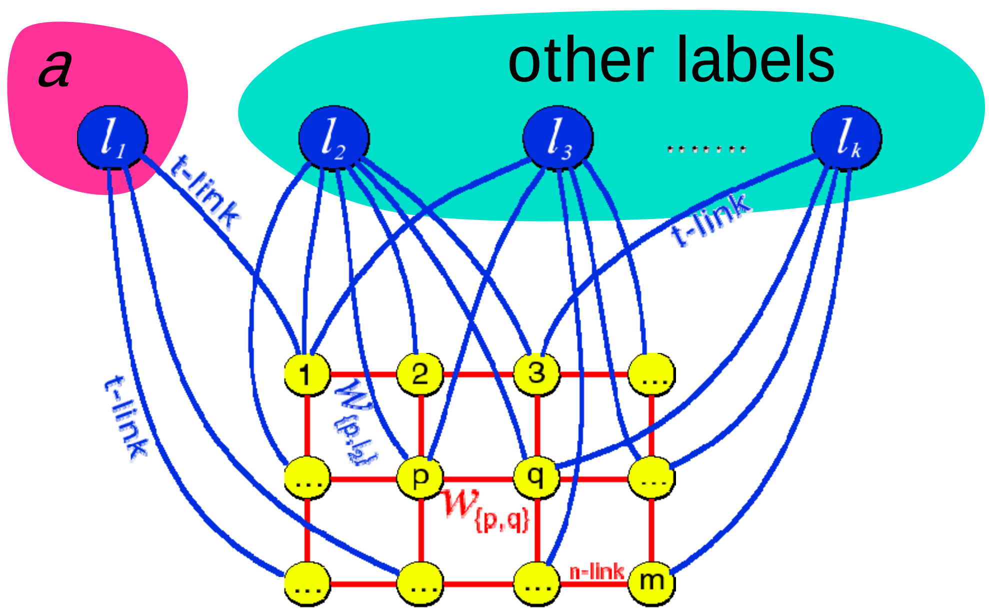
1. Start with any initial solution
2. For each label "*a*" in any (e.g. random) order

   1. *Compute optimal a-expansion move (s-t graph cuts)*
   2. *Decline the move if there is no energy decrease*

3. *Stop when no expansion move would decrease energy*

a-expansion move

Basic idea: break multi-way cut computation into a **sequence of binary *s-t* cuts**



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**

Metric interactions

$V(a,b)=0$ iff $a=b$

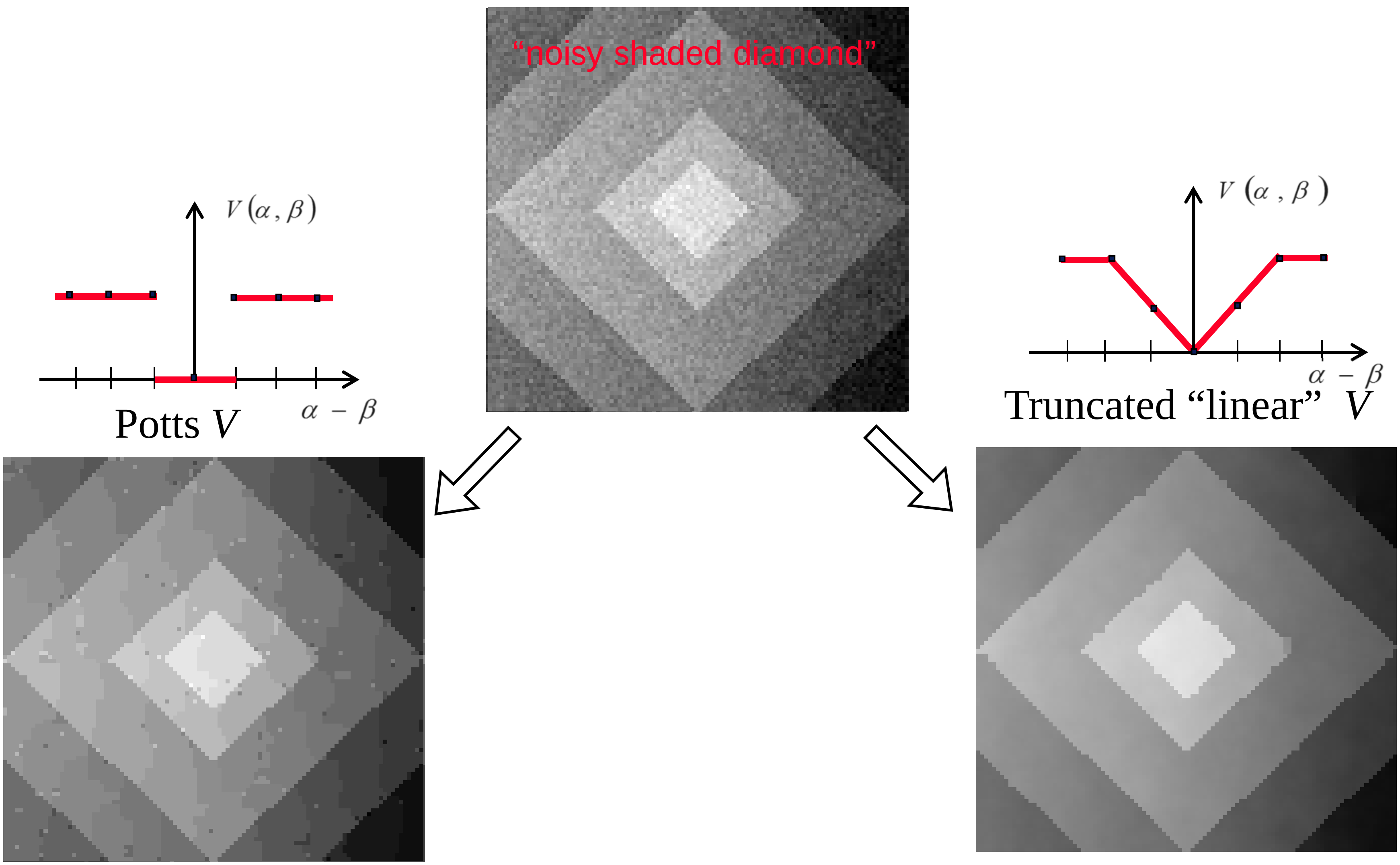
$V(a,b) = V(b,a) >= 0$

$V(a,c) <= V(a,b)+V(b,c)$

Triangular inequality

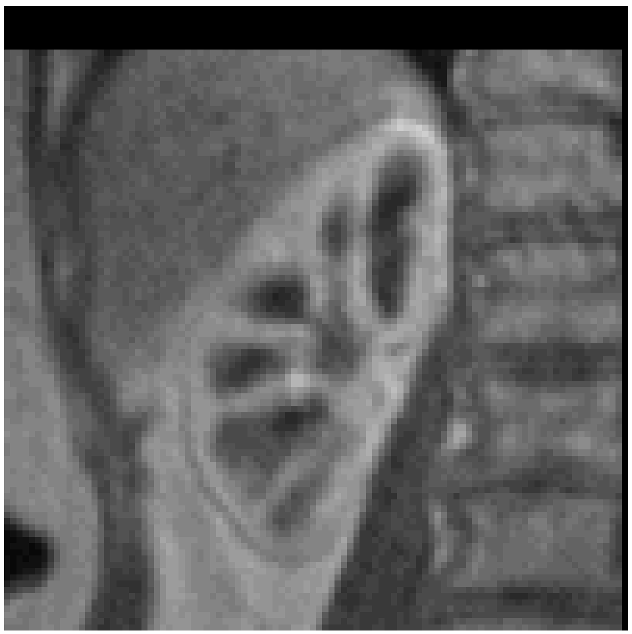
Implies that every expansion move (a binary problem) is submodular

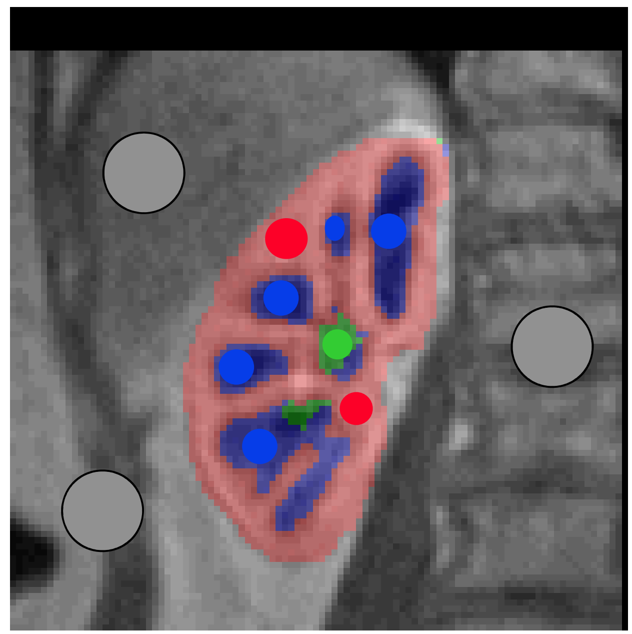
a-expansions: examples of *metric* interactions



Multi-way graph cuts

Multi-object Extraction

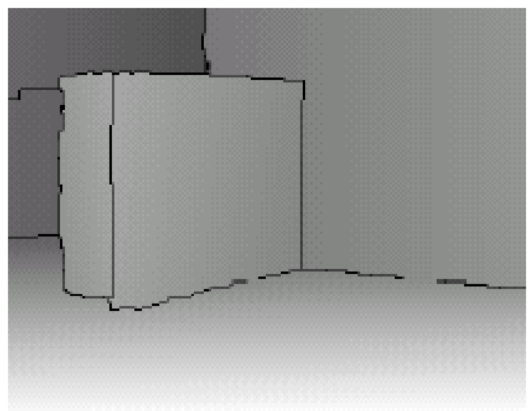




Multi-way graph cuts

Stereo/Motion with slanted surfaces

(Birchfield &Tomasi 1999)



Labels = parameterized surfaces

EM based: E step = compute surface boundaries
M step = re-estimate surface parameters

Multi-way graph cuts

stereo vision

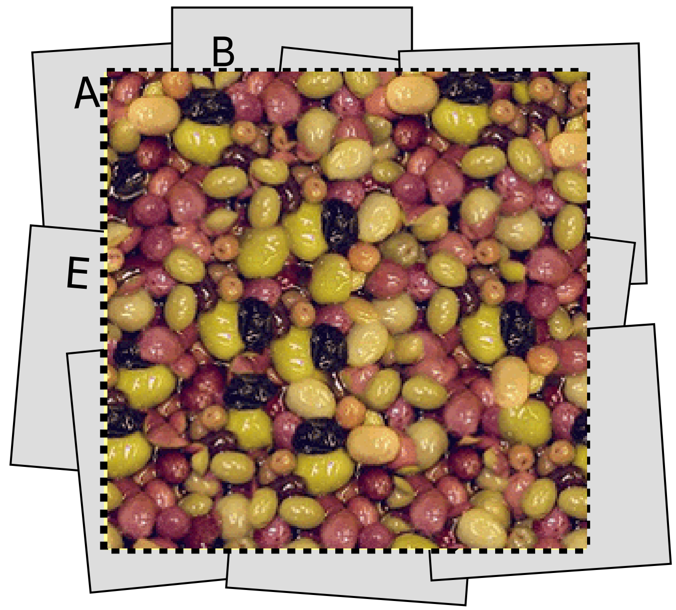
original pair of "stereo" images

depth map

Multi-way graph cuts

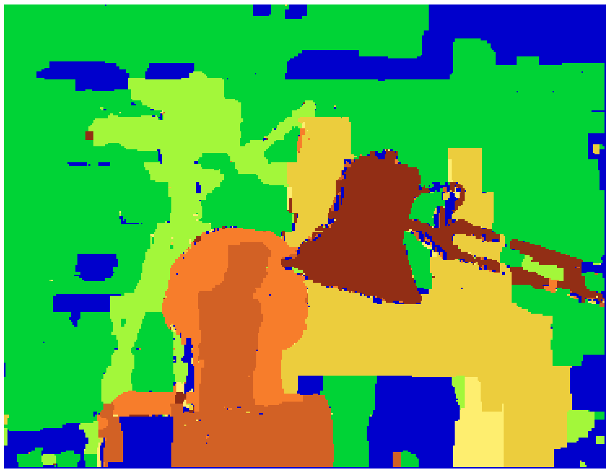
Graph-cut textures

(Kwatra, Schodl, Essa, Bobick 2003)



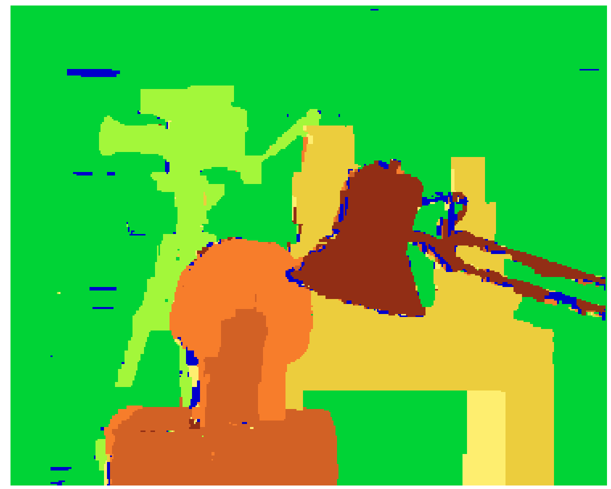
similar to “**image-quilting**” (Efros & Freeman, 2001)

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