Discrete Optimization in Vision and Graphics

Ramin Zabih
Cornell University

EUSIPCO 2008 tutorial

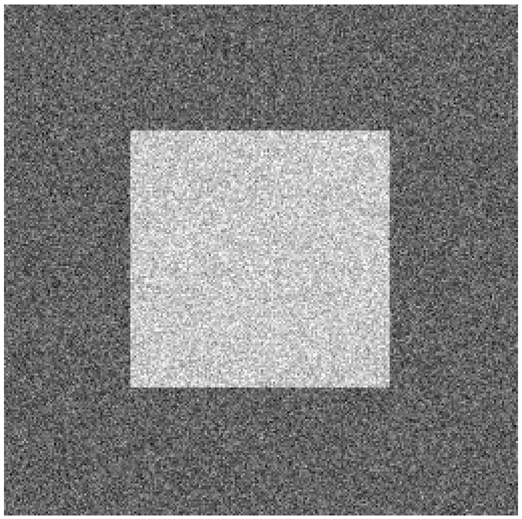
Lausanne, Switzerland, August 2008

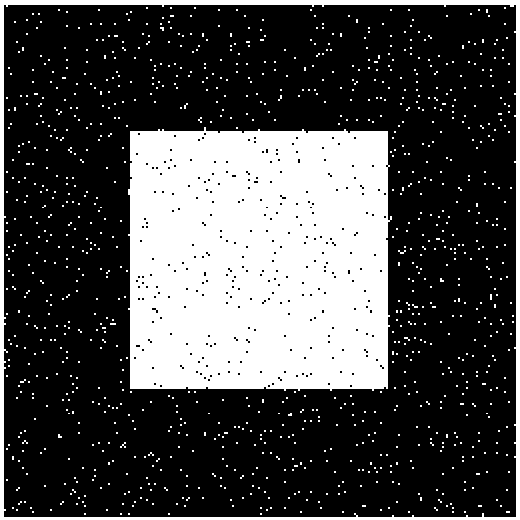
Outline

- MAP estimation of Markov Random Fields
- Flow-based algorithms ("graph cuts") for optimization
- A few sample applications
- Current research problems

Motivating example

- Suppose we want to find a bright object against a dark background
  - But some of the pixel values are slightly wrong





Bayesian view

- Find image x with the highest posterior probability, given observed data y

$$\arg\max \Pr(x|y) = \arg\min[-\ln \Pr(y|x) - \ln \Pr(x)]$$

- Assuming independent noise at each pixel:
  - First term in minimization can be written

$$\sum_p D_p(x_p), \text{ where } D_p(x_p) = -\ln \Pr(y_p|x_p)$$

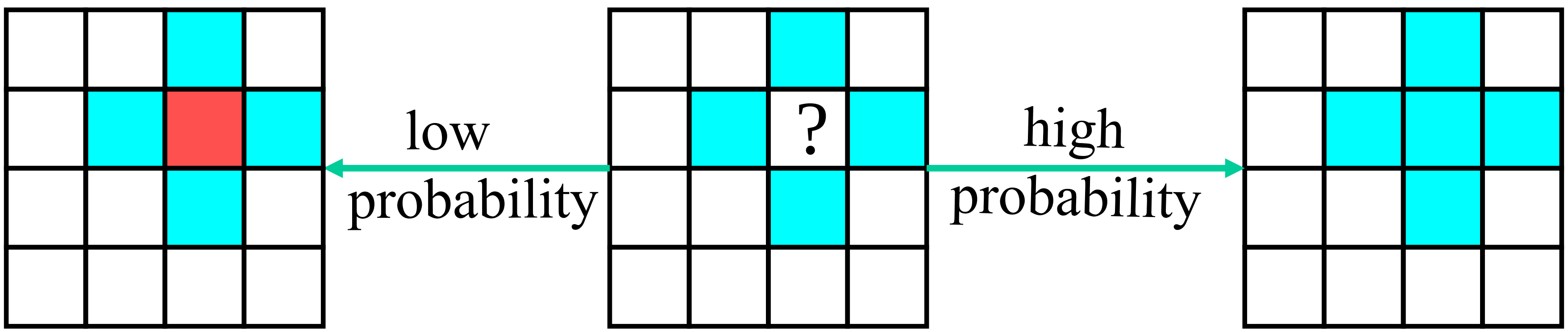
- Write the prior as $\Pr(x) \propto \exp(-G(x))$
- Energy to minimize: $E(x) = \sum_p D_p(x_p) + G(x)$

The key problem

- What is the right prior? What method do we use to minimize the energy?
  - These two issues are **NOT** independent
- Specialized optimization algorithms tend to be better than general-purpose ones
  - Any completely general energy minimization algorithm is equivalent to exhaustive search!
- For a very natural class of priors, there are now powerful specialized optimization methods based on network flow

Markov Random Field priors

- Suppose we want a purely local prior
  - Directly depends only on immediate neighbors



  - See, e.g., Li's book on MRF's for a review

MRF energy function

- In an MRF, $G(x) = \sum_{p,q} V_{p,q}(x_p, x_q)$
  - Think of V as the cost for two adjacent pixels to have these particular labels
  - For binary images, the natural cost is uniform
- MAP-MRF energy function:

$$E(x_1, \ldots, x_n) = \sum_p D_p(x_p) + \sum_{p,q} V_{p,q}(x_p, x_q)$$

Alternate view: optimization

- Find best (least expensive) binary image
  - Costs: C1 (labeling) and C2 (boundary)
- C1: Labeling a dark pixel as foreground
  - Or, a bright pixel as background
- If we only had labeling costs, the cheapest solution is the thresholded output
- C2: The length of the boundary between foreground and background
  - Penalizes isolated pixels or ragged boundaries

Generalizations

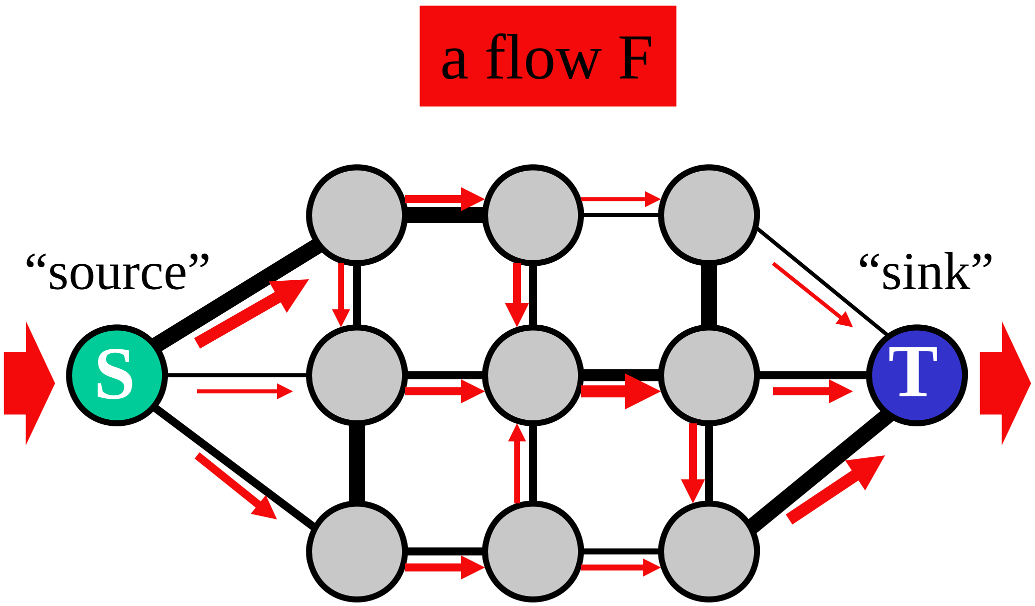
- Many vision problems have this form
  - Assign every pixel a label from a discrete set
  - Each pixel has a cost for every label
  - Information at individual pixels isn't enough!
    - Need a spatial prior
- The optimization techniques are specific to these energy function, but not to images
  - See: [Kleinberg & Tardos JACM 02]
  - Metric labeling problem: the MAP-MRF energy, where V is a metric

2. Flow-based algorithms ("graph cuts")

Network flow can help

- For two labels, natural V is uniform
  - Ising model
- The minimization problem can be solved exactly using network flow
  - Construction due to [Hammer 65]
  - First applied to images by [Greig et al. 86]
- Classical Computer Science problem reduction
  - Turn a new problem into a problem we can solve!

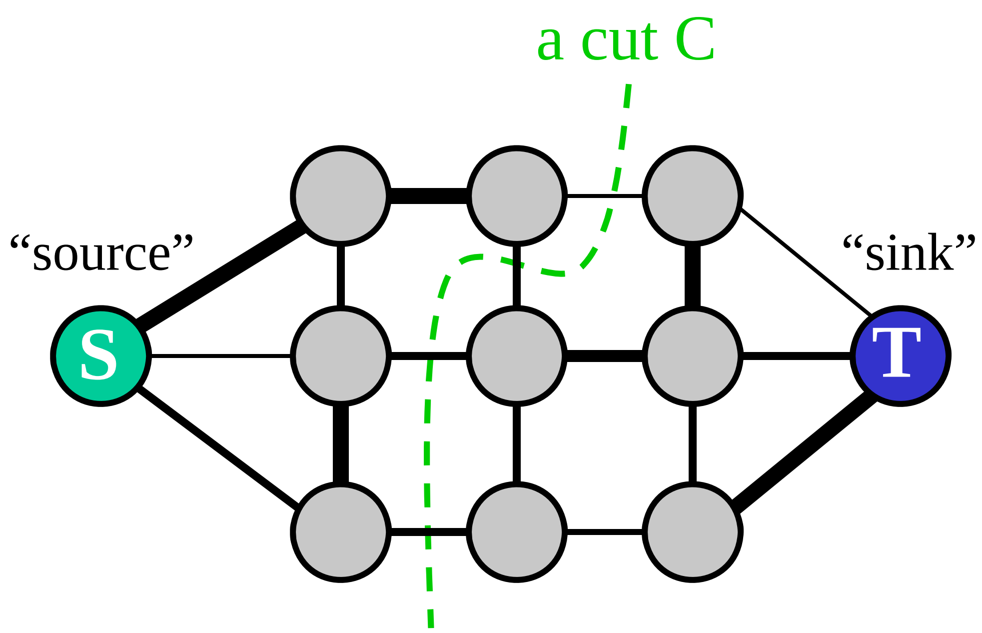
Maximum flow problem



A graph with two terminals

- Max flow problem:
  - Each edge is a "pipe"
  - Find the largest flow F of "water" that can be sent from the "source" to the "sink" along the pipes
  - Source output = sink input = flow value
  - Edge weights give the pipe's capacity

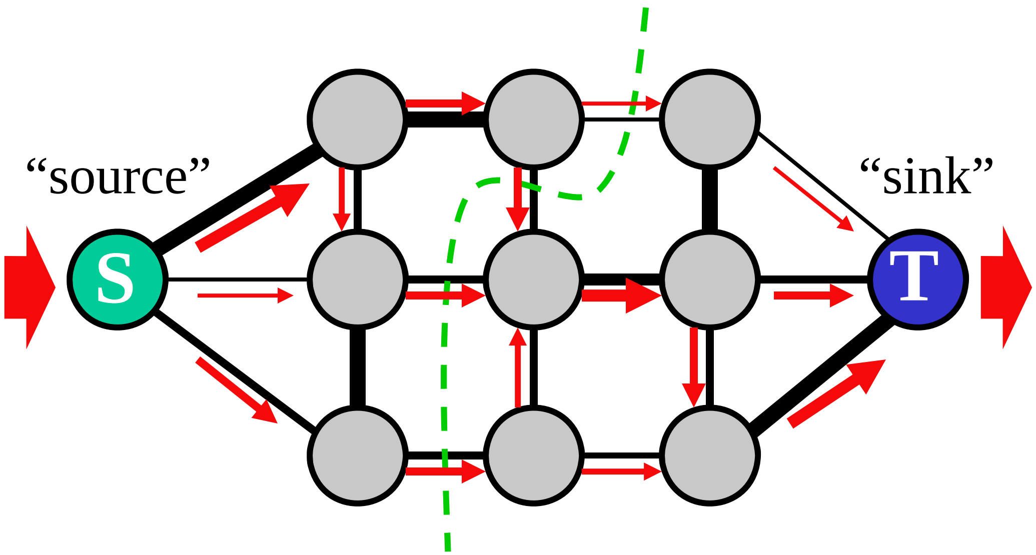
Minimum cut problem



A graph with two terminals

- Min cut problem:
  - Find the cheapest way to cut the edges so that the "source" is separated from the "sink"
  - Cut edges going from source side to sink side
  - Edge weights now represent cutting "costs"

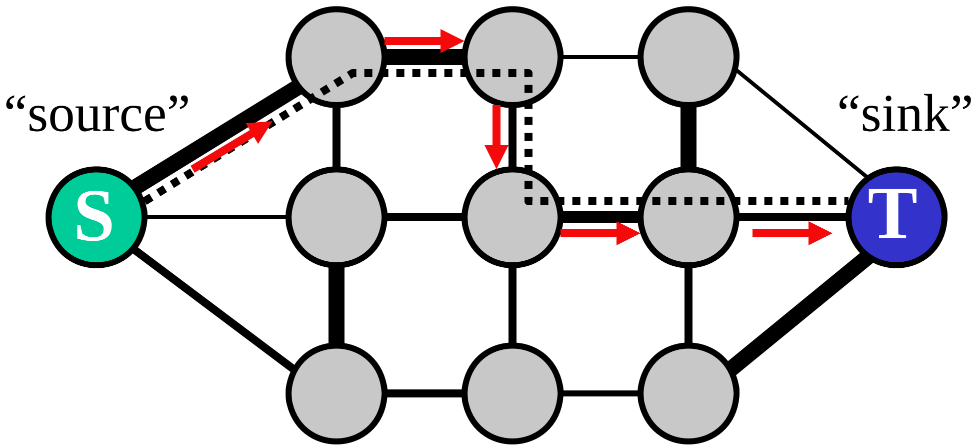
Max flow / Min cut theorem



A graph with two terminals

- Max Flow = Min Cut:
  - Proof sketch: value of a flow is value over any cut
  - Maximum flow saturates the edges along the minimum cut
    - Ford and Fulkerson, 1962
    - Problem reduction!
- Ford and Fulkerson gave first polynomial time algorithm for globally optimal solution

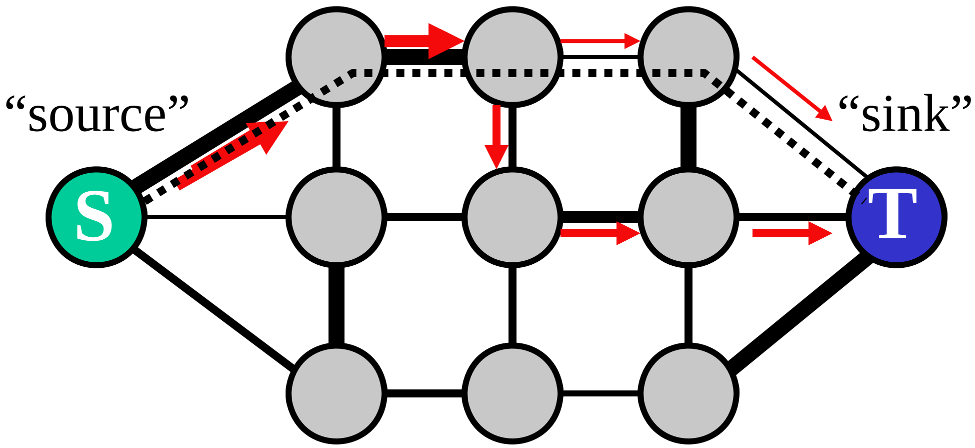
"Augmenting Path" algorithms



A graph with two terminals

- Find a path from S to T along non-saturated edges
- Increase flow along this path until some edge saturates

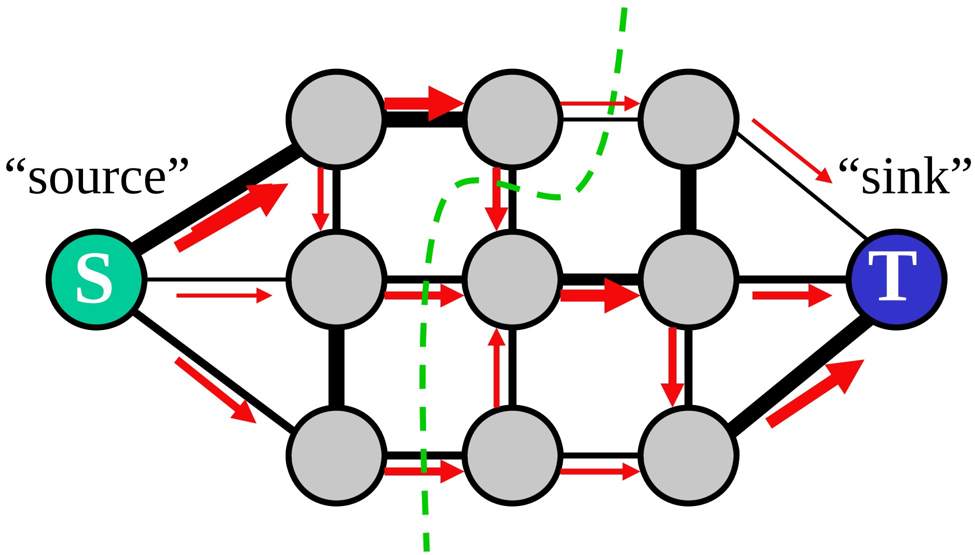
“Augmenting Path” algorithms



A graph with two terminals

- Find a path from S to T along non-saturated edges
- Increase flow along this path until some edge saturates
- Find next path…
- Increase flow…

“Augmenting Path” algorithms



A graph with two terminals

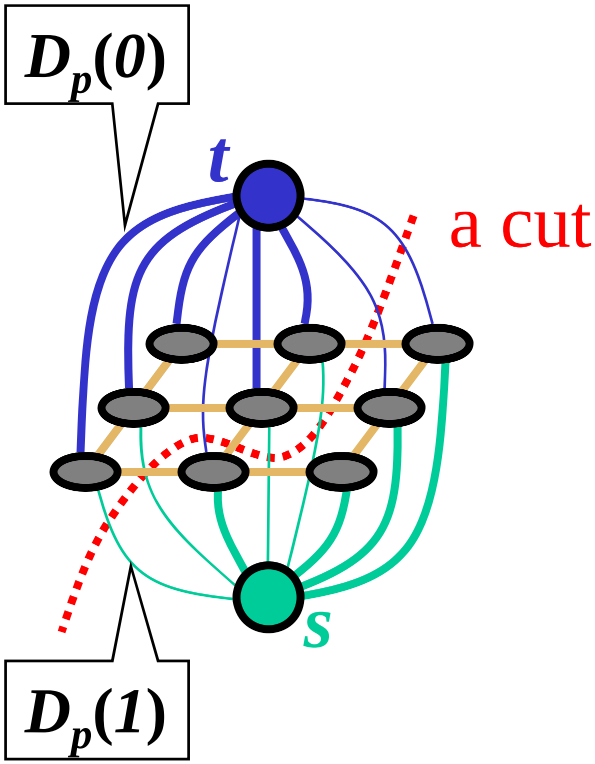
- Find a path from S to T along non-saturated edges
- Increase flow along this path until some edge saturates

Iterate until … all paths from S to T have at least one saturated edge

Cornell University

Basic graph cut construction

- One non-terminal vertex per pixel
- Each pixel connects directly to *s*,*t*
  - Severing these edges corresponds to giving labels 0,1 to the pixel
- Cost of cut is the cost of the entire labeling



Important properties

- Very efficient in practice
  - Lots of short paths, so roughly linear
  - Edmonds-Karp max flow algorithm finds augmenting paths in breadth-first order
- Construction is symmetric (0 vs 1)
- Specific to 2 labels
  - Min cut with >2 labels is NP-hard

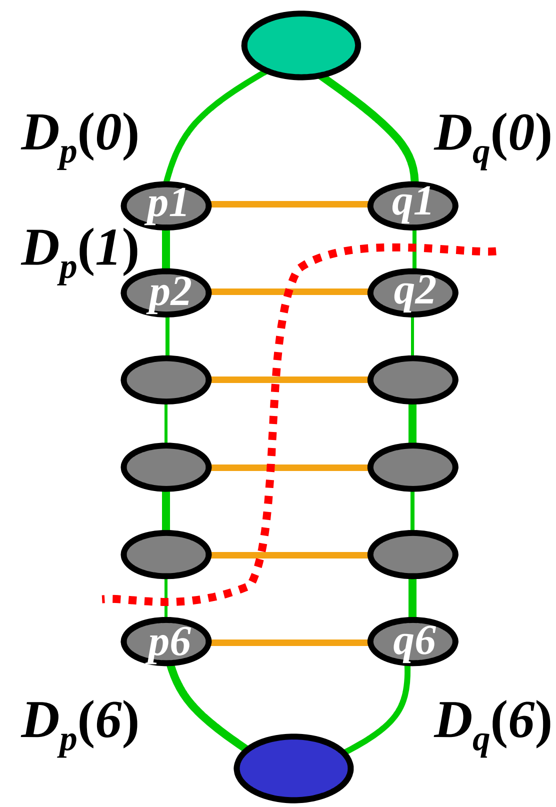
Can this be generalized?

- NP-hard for Potts model [K/BVZ 01]
- Two main approaches

1. Exact solution [Ishikawa 03]
   - Large graph, convex V (arbitrary D)
   - Not the considered the right prior for vision
2. Approximate solutions [BVZ 01]
   - Solve a binary labeling problem, repeatedly
   - Expansion move algorithm

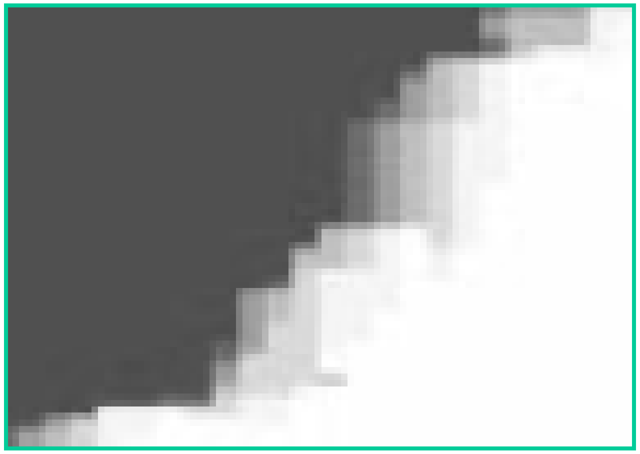
Exact construction for L1 distance

- Graph for 2 pixels, 7 labels:
  - 6 non-terminal vertices per pixel (6 = 7 – 1)
  - Certain edges (vertical green in the figure) correspond to different labels for a pixel
    - If we cut these edges, the right number of horizontal edges will also be cut
- Can be generalized



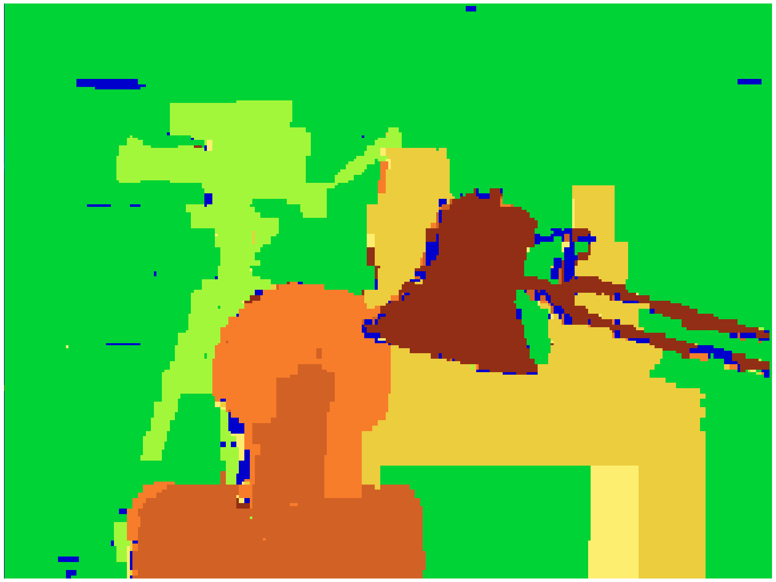
Convex over-smoothing

- Convex priors are widely viewed in vision as inappropriate ("non-robust")
  - These priors prefer globally smooth images
    - Which is almost never suitable
- This is not just a theoretical argument
  - It's observed in practice, even at global min



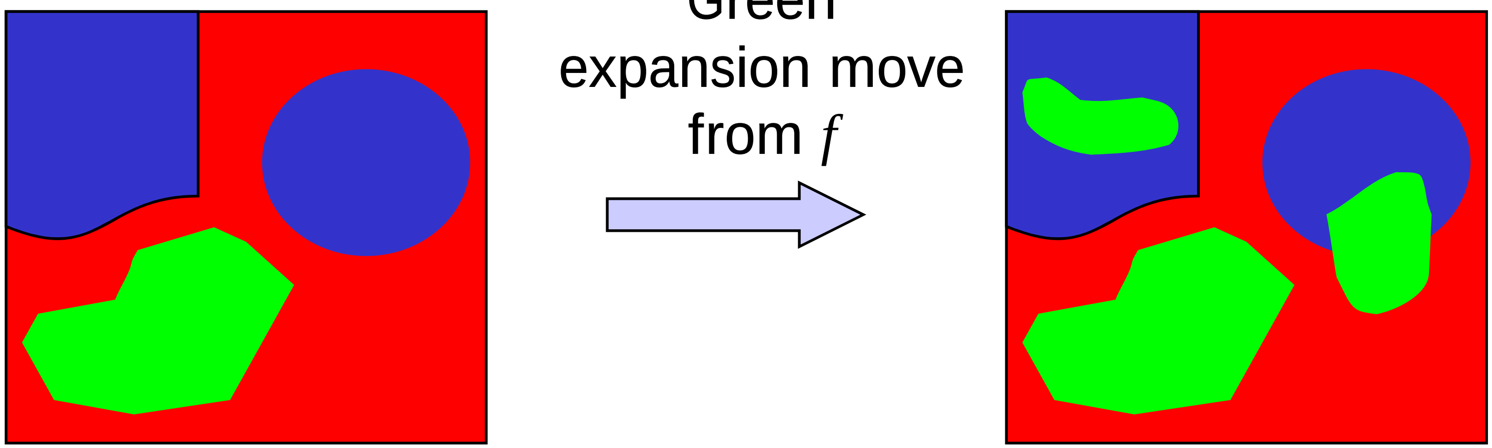
Getting the boundaries right

Right answers



Graph cuts

Expansion move algorithm

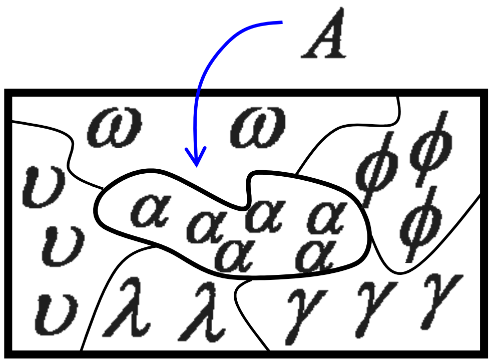


- Make green expansion move that most decreases E
  - Then make the best blue expansion move, etc
  - Done when no α-expansion move decreases the energy, for any label α
  - See [BVZ 01] for details

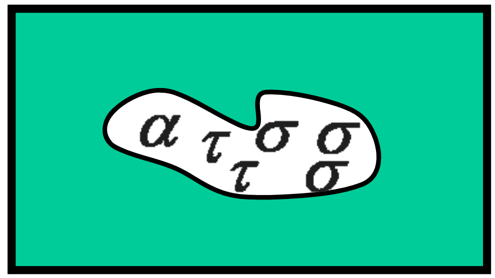
Local improvement vs. Graph cuts

- Continuous vs. discrete
  - No floating point with graph cuts
- Local min in line search vs. global min
- Minimize over a line vs. hypersurface
  - Containing $O(2^n)$ candidates
- Local minimum: weak vs. strong
  - Within 1% of global min on benchmarks!
  - Theoretical guarantees concerning distance from global minimum
    - 2-approximation for a common choice of E

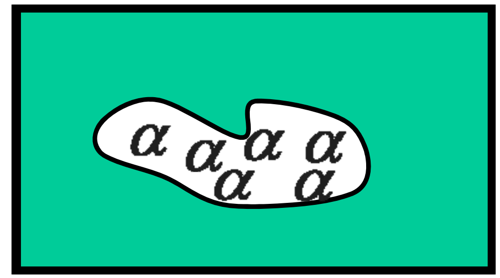
2-approximation for Potts model



optimal solution f^*



local minimum $\hat{f}$



$$f^{\alpha} = \begin{cases} \alpha & p \in A \\ \hat{f}_p & p \notin A \end{cases}$$

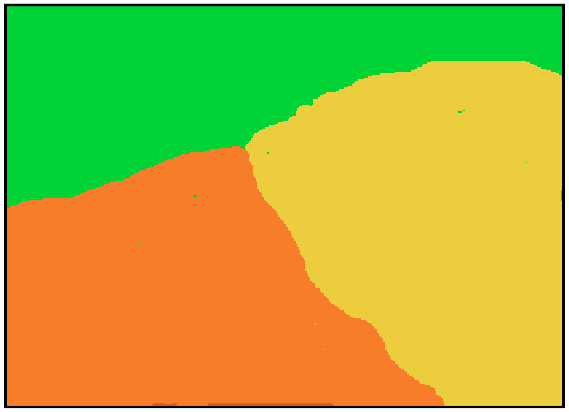
$$\hat{f} \mapsto f^{\alpha} \quad \Longrightarrow \quad E(\hat{f}) \leq E(f^{\alpha})$$

[illegible]

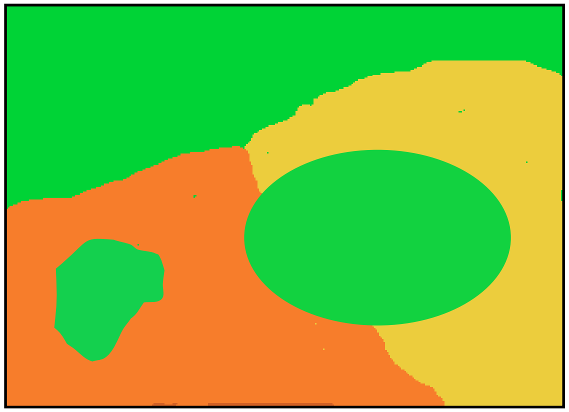
Summing up over all labels:

[illegible]

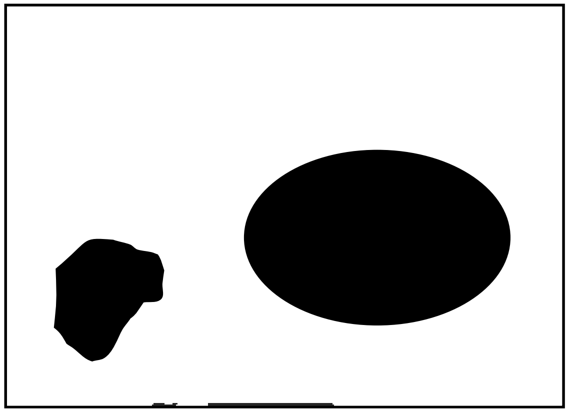
Binary sub-problem



Input labeling

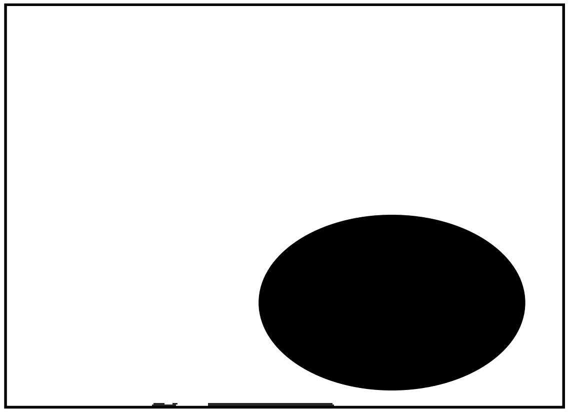


Expansion move



Binary image

Expansion move energy



Goal: find the binary image with lowest energy

Binary image energy **E*(*b*)*** is restricted version of original **E**

Depends on ***f***, α

Regularity

- The binary energy function

$$\sum_p B_p(x_p) + \sum_{p,q} B_{p,q}(x_p, x_q)$$

is *regular* [KZ 04] if

$$B_{p,q}(0,0) + B_{p,q}(1,1) \leq B_{p,q}(0,1) + B_{p,q}(1,0)$$

- Special case of submodularity, which is intimately tied to minimum cuts

When is binary energy regular?

- Can find cheapest α-expansion from f if

$$V(\alpha,\alpha) + V(f(p),f(q)) \leq V(f(p),\alpha) + V(\alpha,f(q))$$

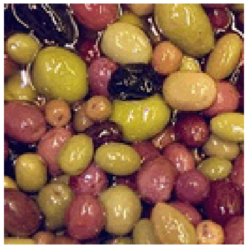
- This is a Monge property
  - It holds if V is a metric
  - A few other cases also
- Until fairly recently, applications of graph cuts required this assumption

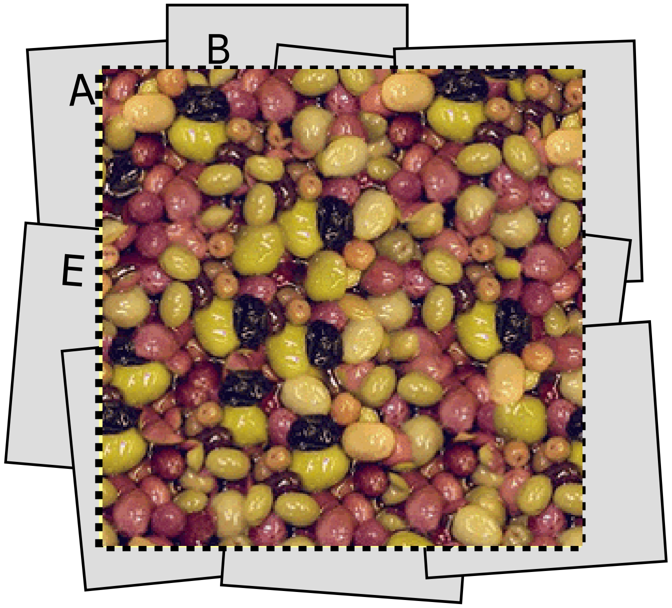
3. Sample applications

Some important applications

- Computer vision
  - Stereo and its variants, segmentation, etc.
- Computer graphics
  - Texture synthesis
  - Creating panoramas
  - Digital photomontage
- Theoretical computer science
  - Metric Labeling Problem
- Industrial applications
  - Microsoft, Google, Siemens

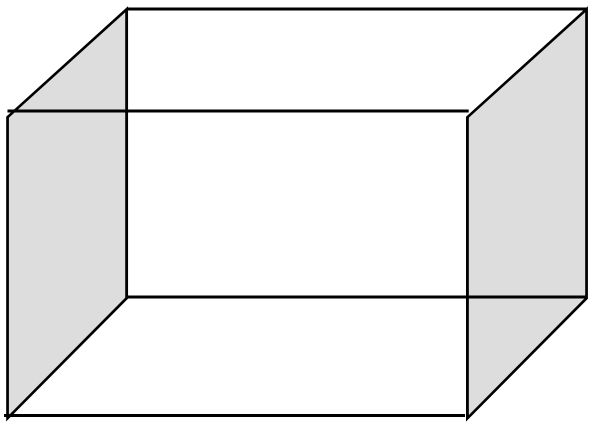
Application: texture synthesis



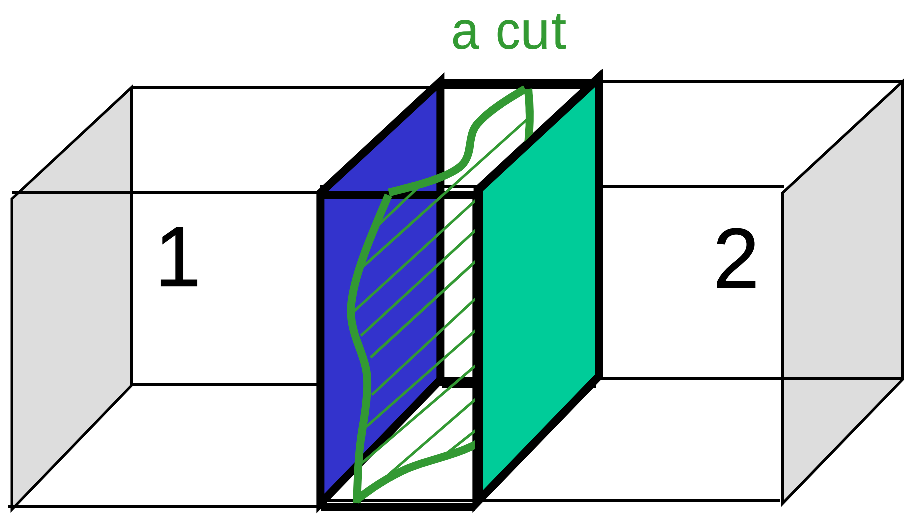


"Graphcut textures" [Kwatra et al 03]

Graphcuts video textures



Short video clip



Long video clip

Another example

original short clip

synthetic infinite texture

Interactive Digital Photomontage

Aseem Agarwala, Mira Dontcheva,

Maneesh Agrawala, Steven Drucker, Alex Colburn,

Brian Curless, David Salesin, Michael Cohen

University of Washington & Microsoft Research

4. Current research problems

Much ongoing work

- Two examples of the kind of work that is currently underway
  - Beyond MAP estimation: measuring uncertainty in graph cut solution [Kohli & Torr ECCV06]
  - Beyond regularity: solving linear inverse systems with graph cuts [Raj et al. MRM07]

Beyond MAP

- [Kohli & Torr] consider this optimization problem: fix the label of a particular pixel, and find the lowest energy labeling
  - "Min-marginal energy"
  - Compute this for all pixel/label pairs, then normalize over each pixel
  - Closely related to the max-marginal probability
    - Maximum probability of all MRF configurations where this pixel has this label

Dynamic graph cuts

- [Kohli & Torr] show how to compute the min-marginal energies fast
  - For all pixel/label pairs!
- They use dynamic graph cuts
  - The graph changes very little when we change the label of a particular pixel
  - We can re-use many of the old flows, which means we can compute the cut efficiently
  - Simple example: if capacities that are not saturated increase, the cut doesn't change

Graph cuts and relaxations

- While graph cuts have been very successful, the regularity constraint has been a major limitation
  – Without it, the binary subproblem (computing the optimal expansion move) is NP-hard
- Much current work uses relaxations to create better methods
  – Some of the nicest work is by Komodakis
  – We will describe how to use a graph cut relaxation for linear inverse systems

Relaxations

- Common sense example: find the cheapest French-made item at a big store
  - Requires exhaustive search
- Consider relaxing the French constraint
  - Easy to solve (look at price list)
  - Suppose it's an item costing €0.10
- If that item is French we are done
- If not, what do we know?
  - Cheapest French item costs €0.10 or more!

Relaxation idea

- Minimize same function over a bigger set
  - If the answer lies in your original set, done!
  - If not, you have a lower bound
  - Standard example: LP-relaxation
- Why care about lower bounds?
  - Provides confidence measure on output
  - Occasionally proves global minimum for NP-hard problems
  - Useful for branch-and-bound algorithms

Linear inverse systems

- Originally, we assumed that each intensity was independently affected by noise
  - This is implicit in the first term of the energy
    - Sum over individual pixels
  - Suppose known linear system $\boldsymbol{H}$ is applied first
- First term costs are $\|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}\|$, which is

$$\sum_{\substack{p,p' \text{ s.t.} \\ H(p,p') \neq 0}} D_p(x(p)) + D_{p'}(x(p')) + d_{pp'} \cdot x(p) \cdot x(p')$$

for appropriate functions $\boldsymbol{D}$, constants $\boldsymbol{d}$

Our energy function

$$E(x) = \|y - Hx\| + \lambda \sum_{(p,q)\in\mathcal{N}} V(x_p - x_q)$$

General problem: edge-preserving solution of linear inverse systems

Regularity is a challenge

- For non-negative $\boldsymbol{H}$, the binary energy function is regular iff

$$f_p \leq \alpha \leq f_q$$

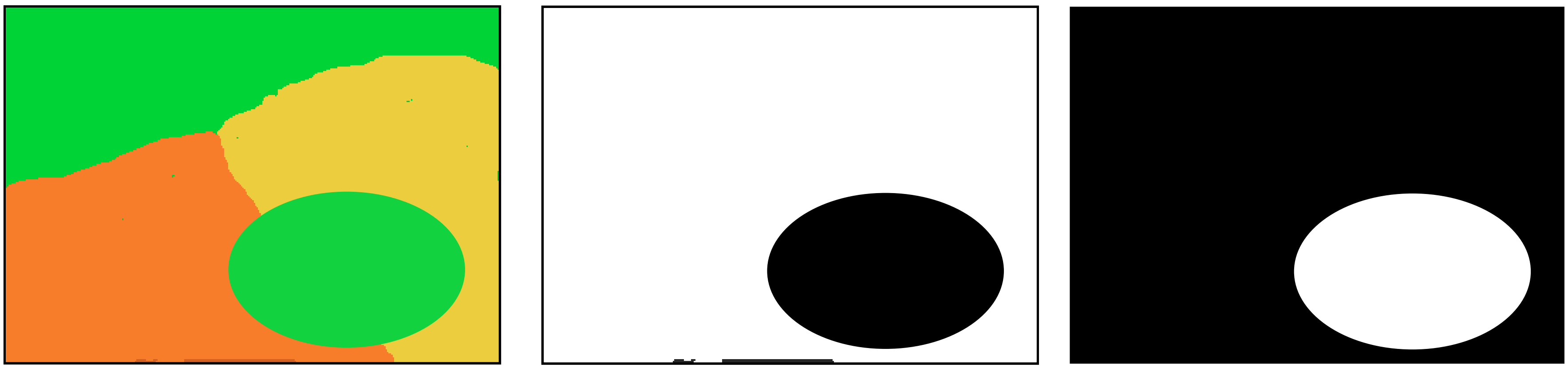
- Can compute the optimal α-expansion move for a pixel below α where all its neighbors are above α (or vice-versa)
  - This is true for very few pixels!

Applying roof duality

- [Hammer et al 84] solves binary subproblem
  - Graph construction called “roof duality”
  - Introduced into computer vision by Vladimir Kolmogorov in early 2005
- Basic idea: relaxation with nice properties
  - Directly find a good expansion move
- Even happens when solution
  - For some linear inverse systems, it’s often optimal!

Roof duality relaxation

- Alternate encoding of expansion moves

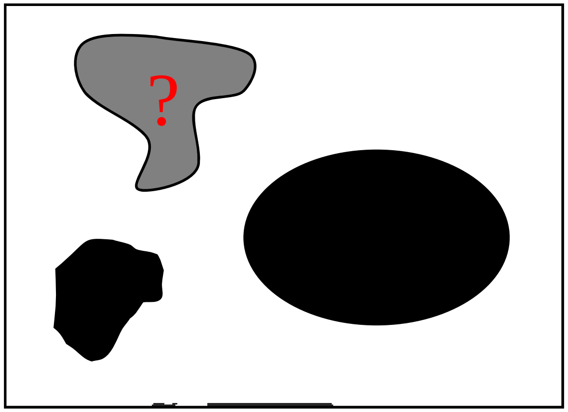


b $\bar{b}$

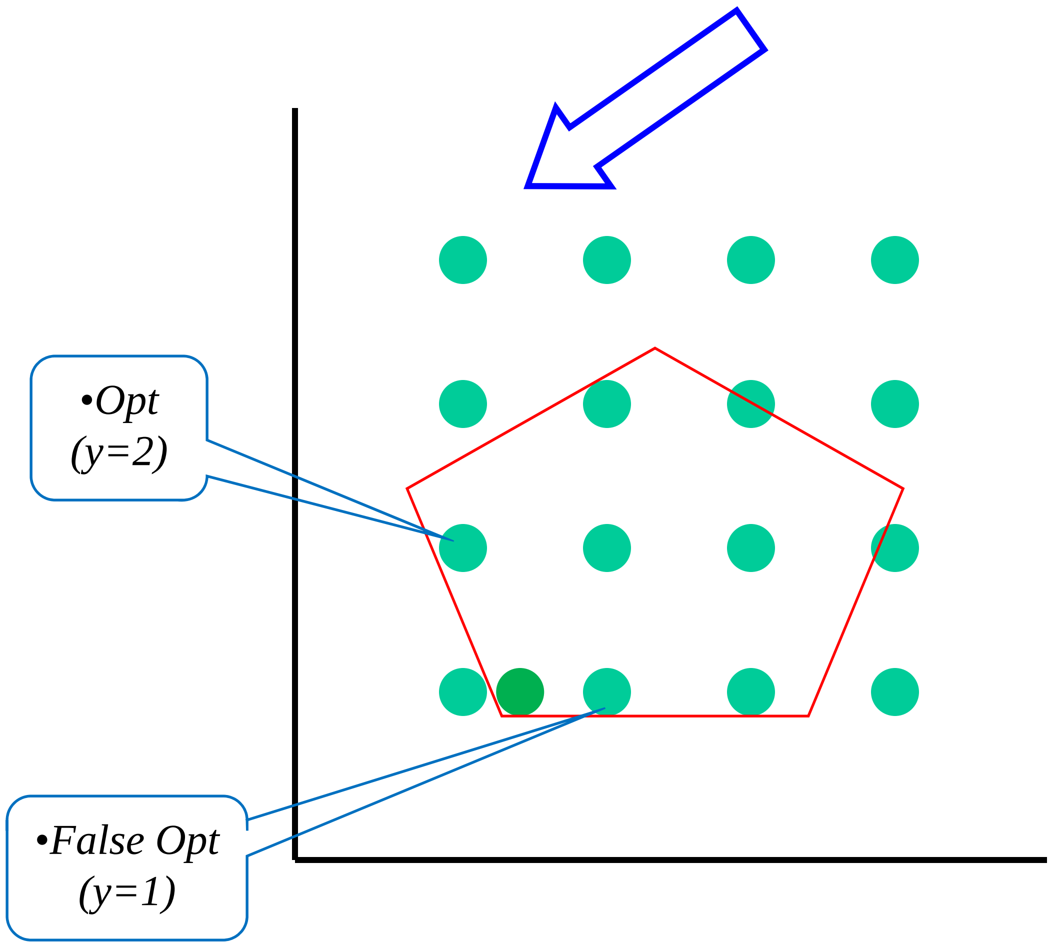
- Can't use graph cuts to minimize $E(b)$
- But we can minimize the relaxation $E'(b, \bar{b})$
  - Note: $E'(b, 1-b) = E(b)$

Theoretical properties

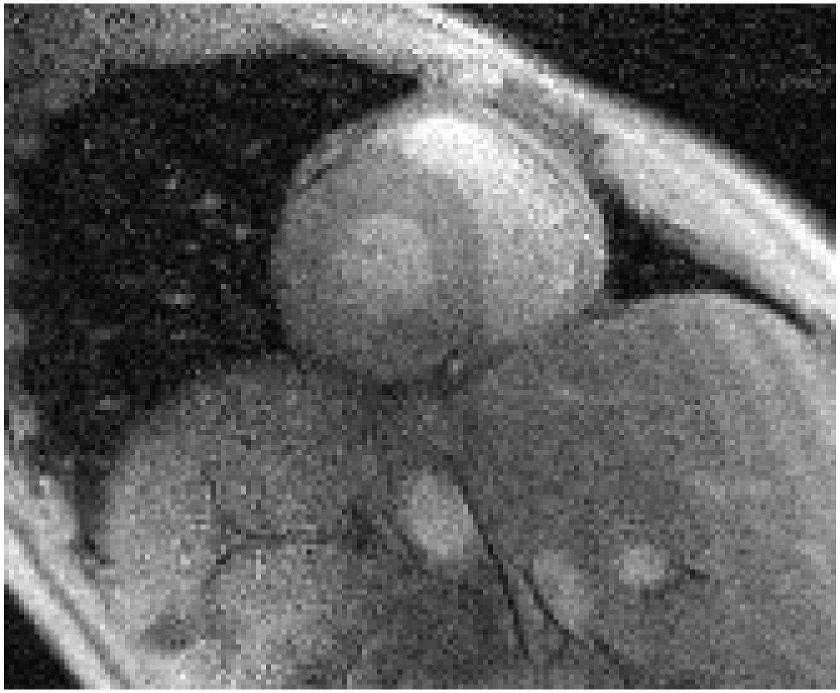
- [Hammer et al 84] show this relaxation has an amazing partial-optimality property
  - Strong persistency: all consistent pixels have the correct label



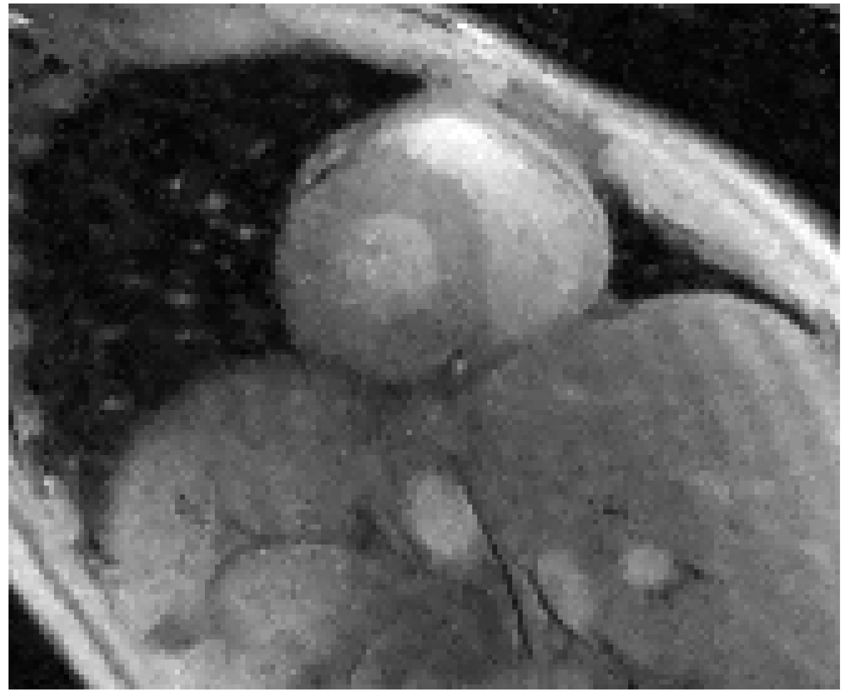
Partial optimality is hard!



MRI results [Raj et al. MRM 07]



SENSE
(= LS)



Graph cuts

Conclusions

- Flow-based algorithms ("graph cuts") are powerful tools for MAP estimation of MRF's
  - Common in computer vision, and elsewhere
- Lots of interest in using these methods for even broader classes of problems
- Graph cuts can give strong results for linear inverse systems with discontinuities
  - There are lots of these (in MR and beyond)