CS664 Lecture #9: Binary labeling problems and their applications

Some material taken from:

- **Carsten Rother *et al*, Microsoft Research**
  http://research.microsoft.com/vision/cambridge/i3l/segmentation/GrabCut.htm
- **Olga Veksler, University of Western Ontario**
  http://www.csd.uwo.ca/faculty/olga/

Announcements

- PS1 is on the web, and debugged (?)
  - Due on 10/6
- Quiz 2 on Tuesday 9/27
  - Coverage through last lecture

Recap

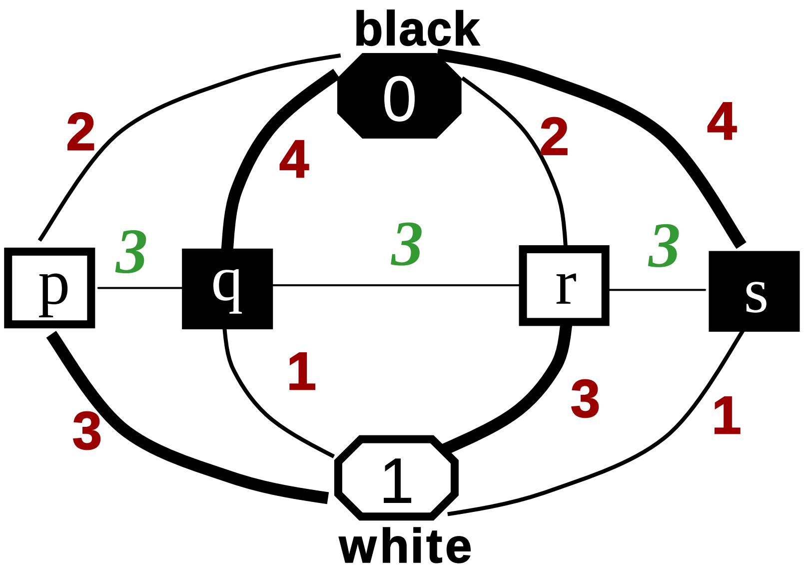
- We want to minimize the energy $E(f)$

$$\arg\min_f \underbrace{\sum_{p \in \mathcal{S}} D_p(f_p)}_{\text{assignment costs}} + \underbrace{\sum_{p,q \in \mathcal{N}} V(f_p, f_q)}_{\text{separation costs}}$$

- Binary labeling problem
  – Also known as Ising model (1924)

Graph construction

- Binary labeling via graph cuts example:
  - $\lambda = 3$
  - $D_p(0) = 1$ and $D_p(1) = 4$ whenever $I(p) = 0$
  - $D_p(0) = 3$ and $D_p(1) = 2$ whenever $I(p) = 1$



s-t Graph Cuts for General Binary Labeling

- Binary Labeling problem:
  - We have 2 labels, let's call them ***s*** and ***t*** (associated with the source and sink terminals, respectively)
  - We have a set of sites (image pixels, volume voxels, so on) to which we wish to assign these labels
  - Each site ***p*** prefers one label more than another
  - We wish to find a labeling ***L*** (assignment of labels ***s*** and ***t*** to sites) such that
    - Each site is assigned a label it prefers, as much as possible
    - Nearby pixels tend to have similar labels

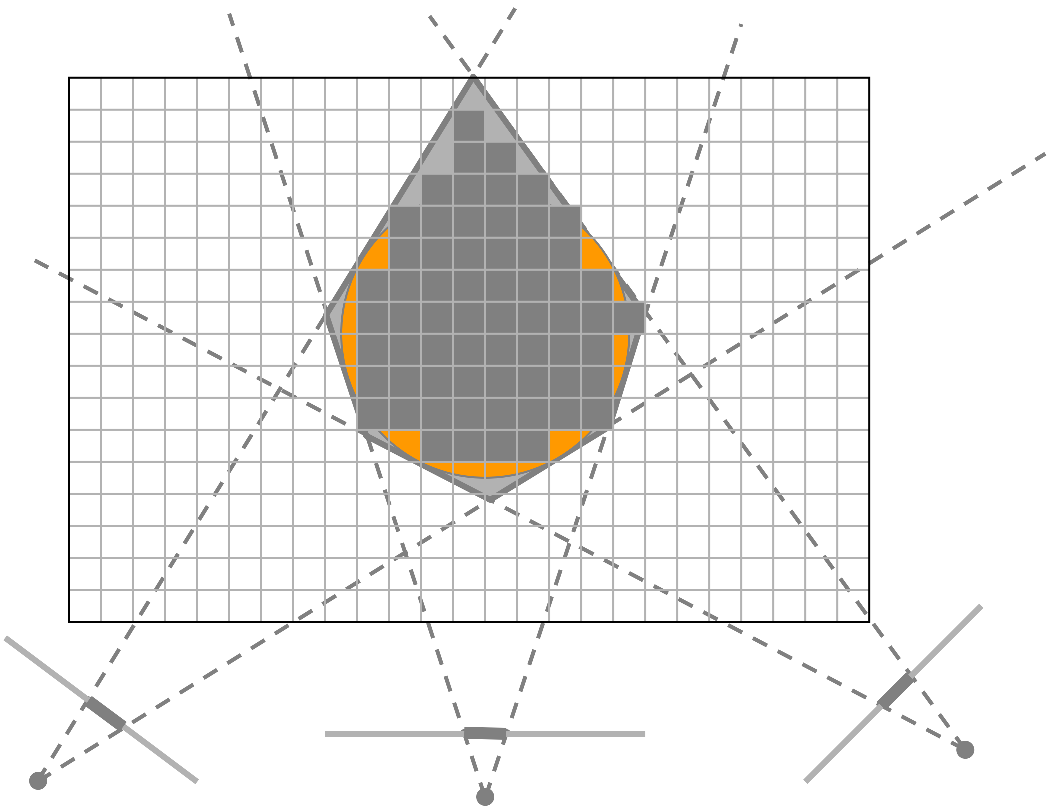
Why should anyone care about binary labeling problems?

- Remarkably powerful and common
  - Now that we can solve them fast!
  - You can use graph cuts to evaluate simulated annealing, and also to determine if the minimum is actually what you want
    - Pretty good, but not quite (why?)
- 3 applications today:
  - 3D shape reconstruction (CVPR 2000)
  - Medical image segmentation (ICCV 2001)
  - Object “cut and paste” (SIGGRAPH 2004)

Application 1: shape reconstruction

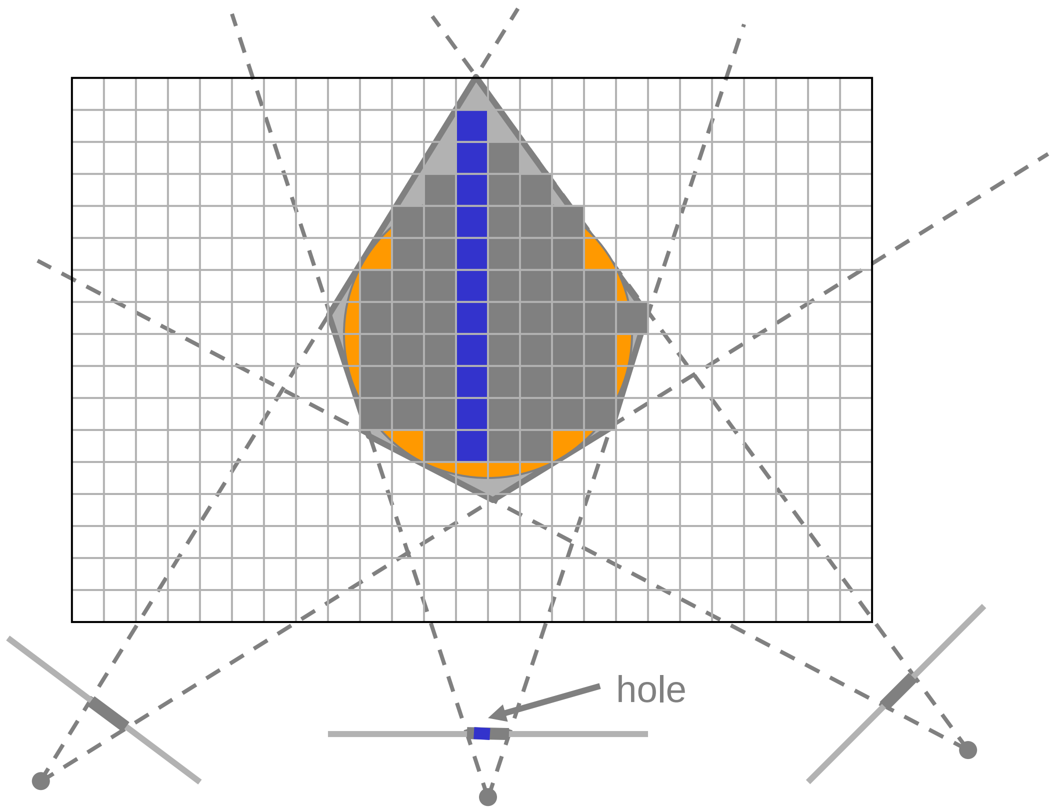
- Snow, Viola & Zabih, CVPR00
  - Create a 3D shape from multiple cameras
- Generalization of “image differencing”
  - Take two consecutive images
  - Subtract and threshold
- Easy to build a simple motion detector
  - Common in “security webcams”
  - Where does this fail?
- Consider an object surrounded by cameras
  - Or, on a turntable

Voxel algorithm for volume intersection



- Color voxel gray if on silhouette in every image

Voxel algorithm for volume intersection



- Color voxel gray if on silhouette in every image

Silhouette intersection

- Early hard decisions are a problem
  - Errors in silhouette identification
  - Can lead to “holes”
- Can we add spatial smoothness with energy minimization?
- More importantly, can we do it fast?

Problem formalization

- We are given a set of voxels $p \in P$ which must be assigned binary labels f_p
  - 1 if filled, 0 if empty
- A pixel in an image corresponds to a set of voxels in a line
  - If there is a large change in intensity (versus background image), these voxels "prefer" to be 1's. Similarly for 0's.
- Note what we don't represent:
  - Surfaces, photoconsistency, etc.

Energy function

- The cost for the voxel p to have the label f_p will be written $D_p(f_p)$
- We seek the labeling f that minimizes

$$E(f) = \sum_p D_p(f_p) + E_{smooth}(f)$$

for some smoothness term $E_{smooth}(f)$

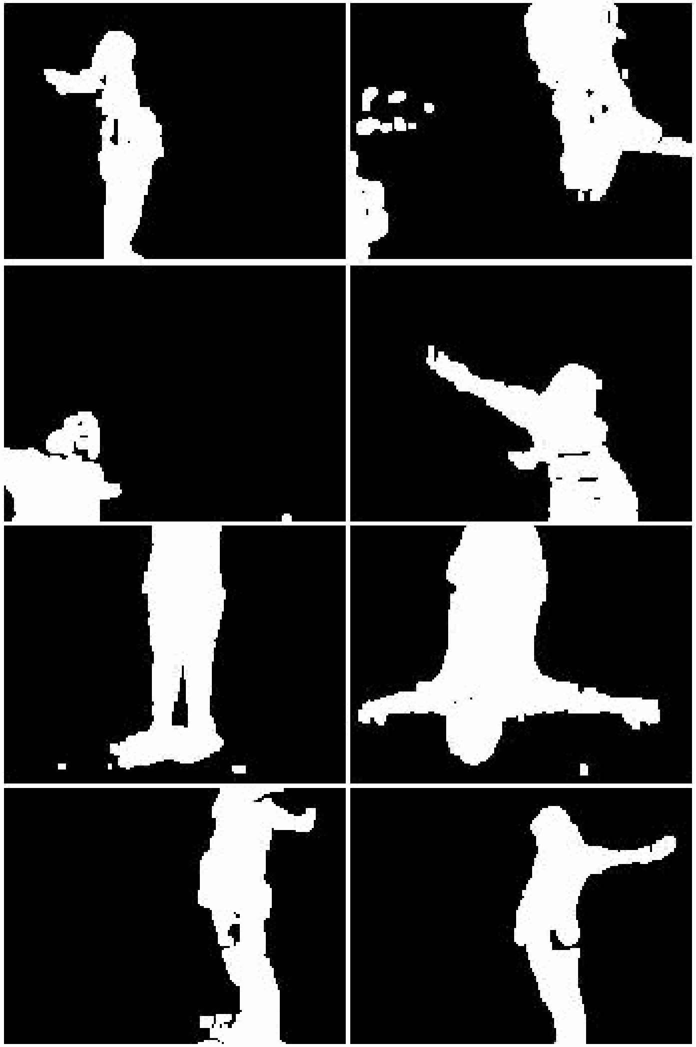
What smoothness term?

- It is easy to imagine lots of clever smoothness terms
  - That we have no idea how to minimize!
- Here is a natural one where we can efficiently find the global minimum!

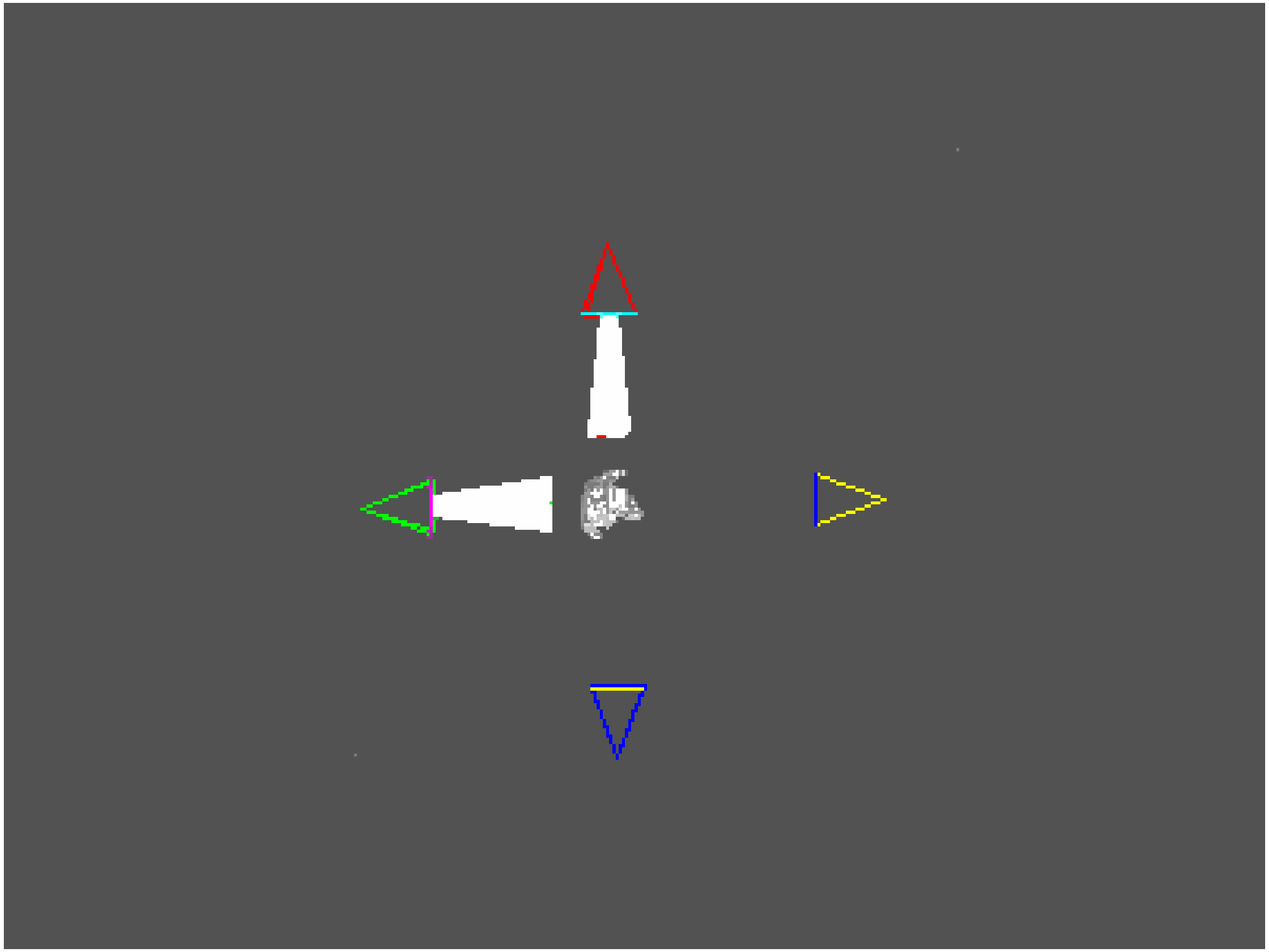
$$E_{smooth}(f) = \sum_{p,q \in N} V(f_p, f_q)$$

$$V(f_p, f_q) = \mathbf{T}[f_p \neq f_q]$$

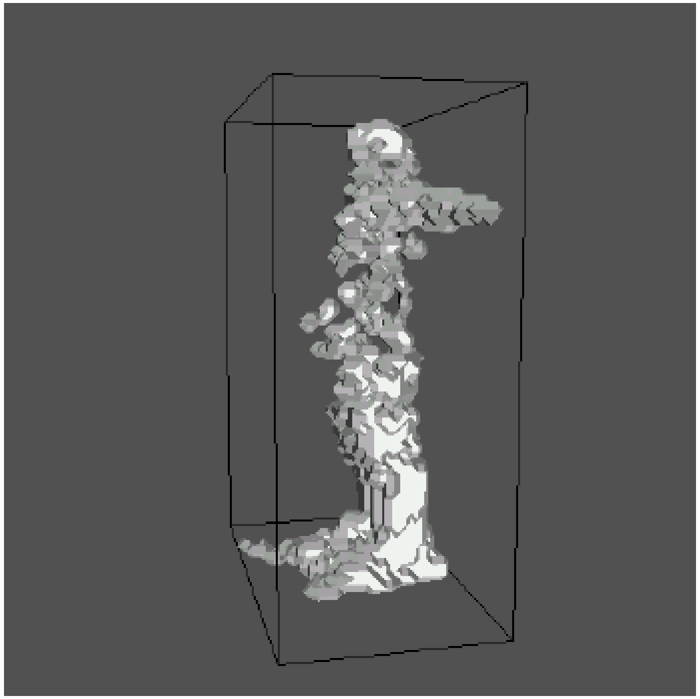
Examples on real data



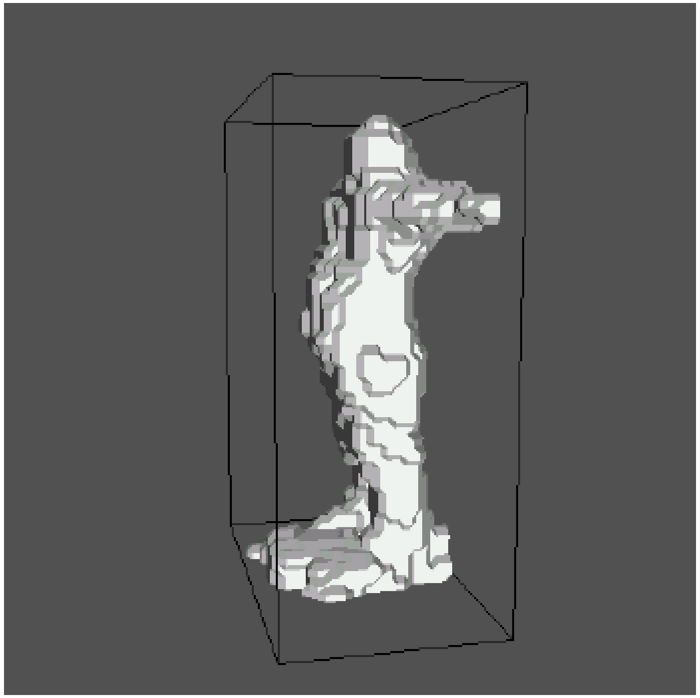
Visualization



Another example



Conventional



Energy minimization

Fly-throughs

- Synthetic example (cylinder)
  – silhouette intersection reconstruction
  – graph cuts reconstruction
- Sara
  – graph cuts reconstruction
- Paul
  – silhouette intersection reconstruction
  – graph cuts reconstruction

Application 2: medical image segmentation

- Boykov & Jolly, ICCV01
  - Interactive tool to segment objects

s-t Graph Cut for General Binary Labeling

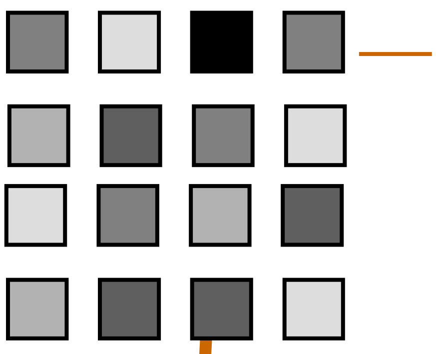
- For the fax image problem, we minimized:

$$E(L) = \sum_p D_P(L_p) \; + \lambda \sum_{pq \in N} \mathbf{T}(L_p \neq L_q)$$

- More general energy function:

$$E(L) = \sum_p D_P(L_p) \; + \sum_{pq \in N} w_{pq} \mathbf{T}(L_p \neq L_q)$$

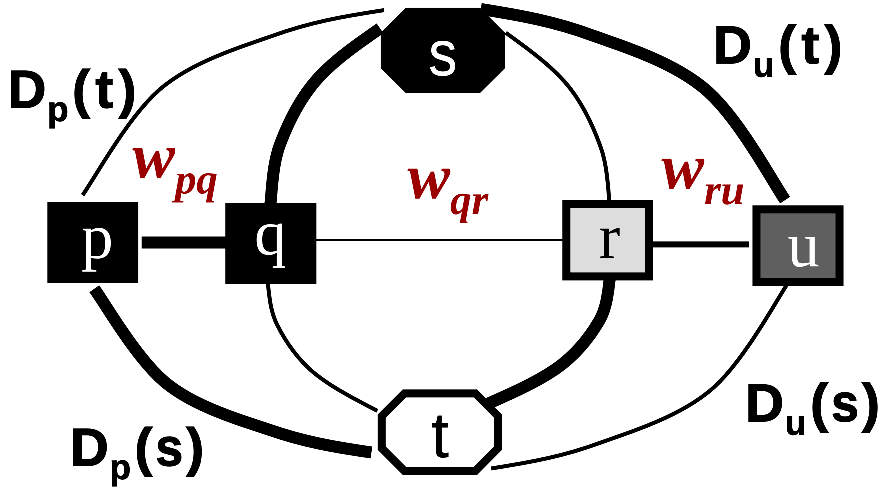
  - The weight of discontinuity is $\boldsymbol{w_{pq}}$ and it depends on a particular pair of pixels $\boldsymbol{p}$ and $\boldsymbol{q}$.



In grayscale (or color) images, discontinuities between similarly colored pixels cost more than discontinuities between differently colored pixels; object boundaries tend to coincide with intensity boundaries

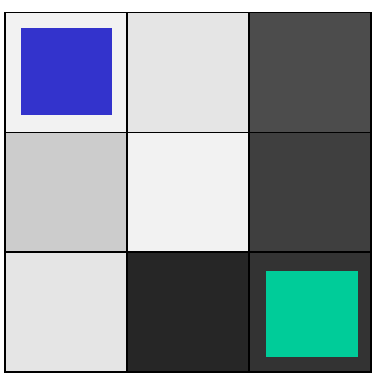
s-t Graph Cut for General Binary Labeling

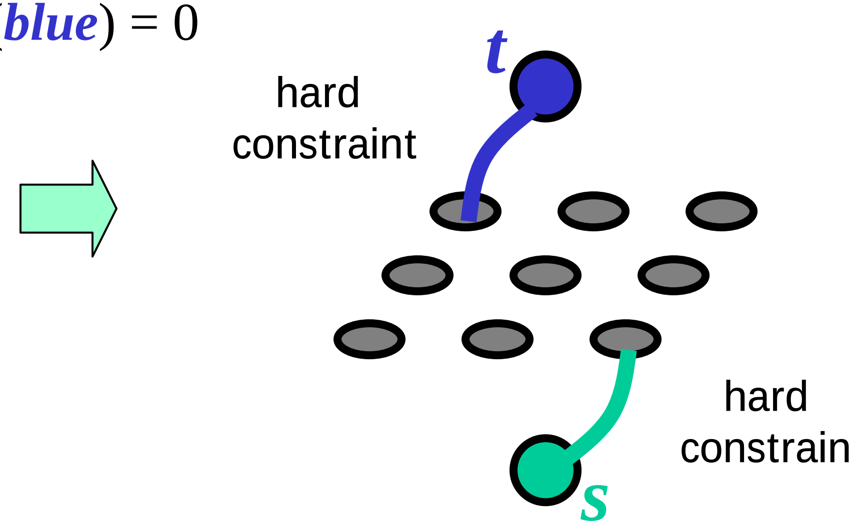
- How does our graph construction change?
  - Cost of t-links does not change
  - Instead of λ, make the weight of n-link between pixels p and q to be w_{pq}



Interactive Object Extraction

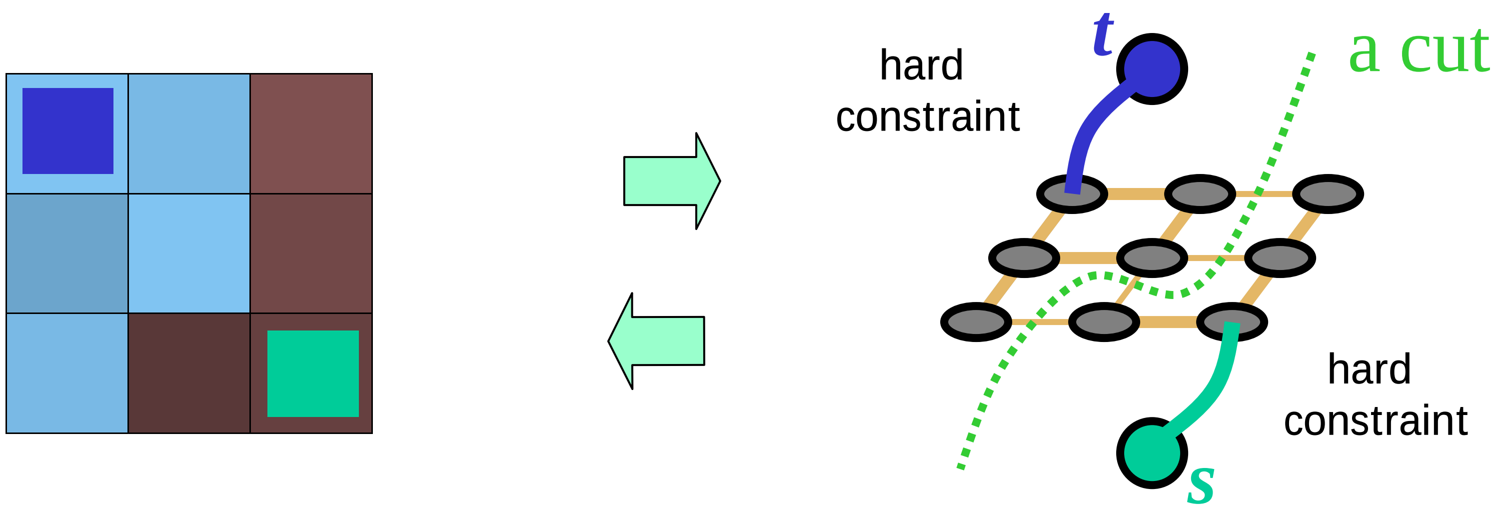
- 2 labels: background (blue), object (red)
- User clicks on a few pixels with blue and red paintbrushes
- Data term
  - $D_p(blue)$ = infinity (that is a very large number) if the user clicked on pixel p with a red paintbrush
    - ✓ Hard constraint: pixel ***p*** must be in the background
  - $D_p(red)$ = infinity (that is a very large number) if the user clicked on pixel p with a blue paintbrush
    - ✓ Hard constraint: pixel ***p*** must be in the object
  - Otherwise $D_p(red) = D_p(blue) = 0$

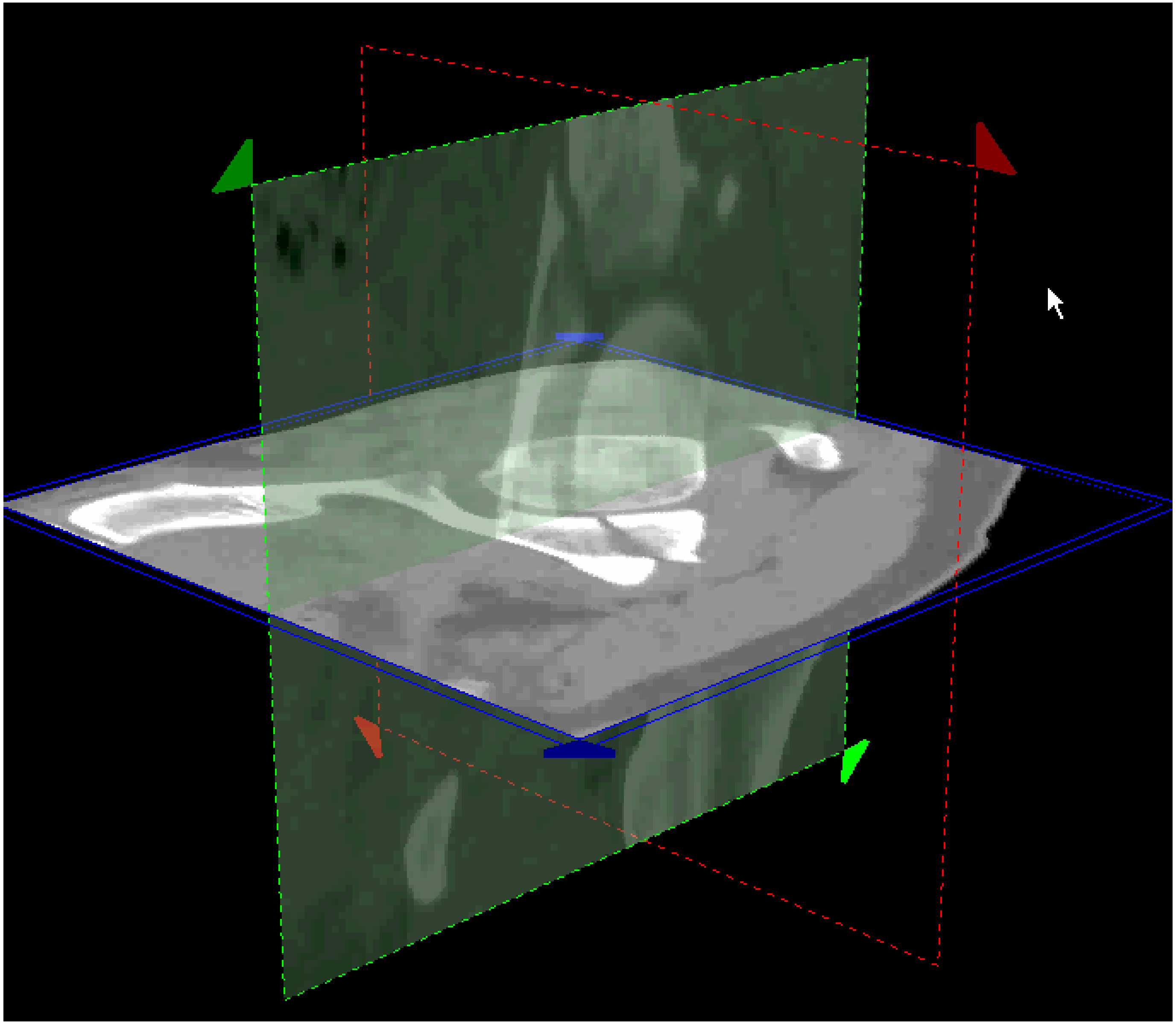




Interactive Object Extraction

- Prior term
  - Object boundaries tend to have sharp intensity edge
  - Encourage cutting around sharp intensity edge
  - Make w_{pq} cheaper for larger intensity difference between p and q

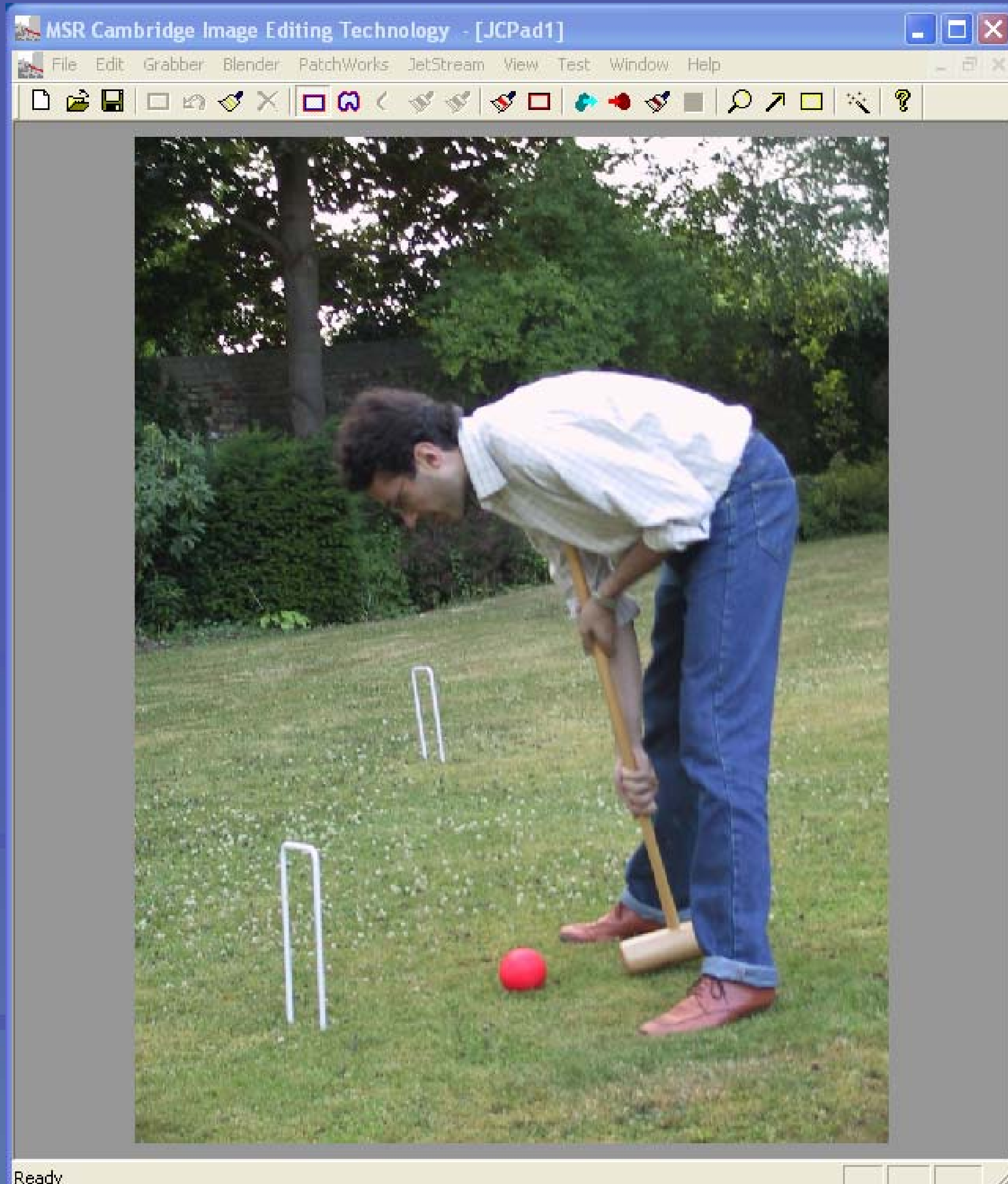




Application 3: object “cut and paste”

- Rother, Kolmogorov, Blake SIGGRAPH04

Photomontage



SIGGRAPH2004

Problem

SIGGRAPH2004

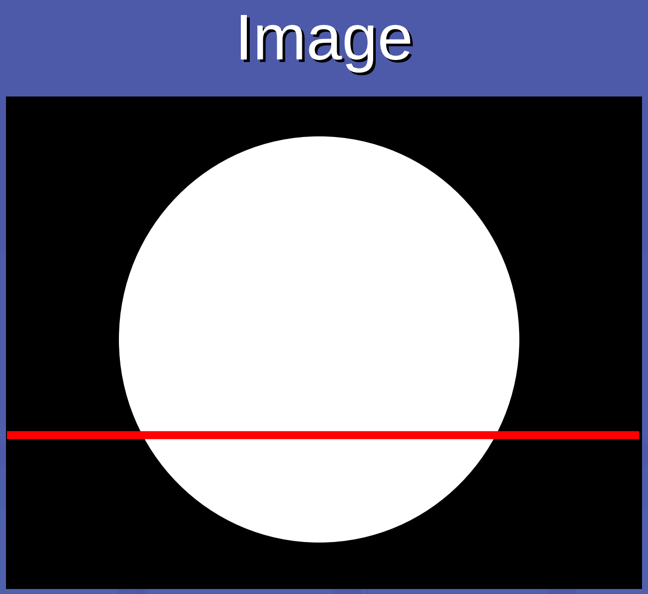
Fast &
Accurate ?

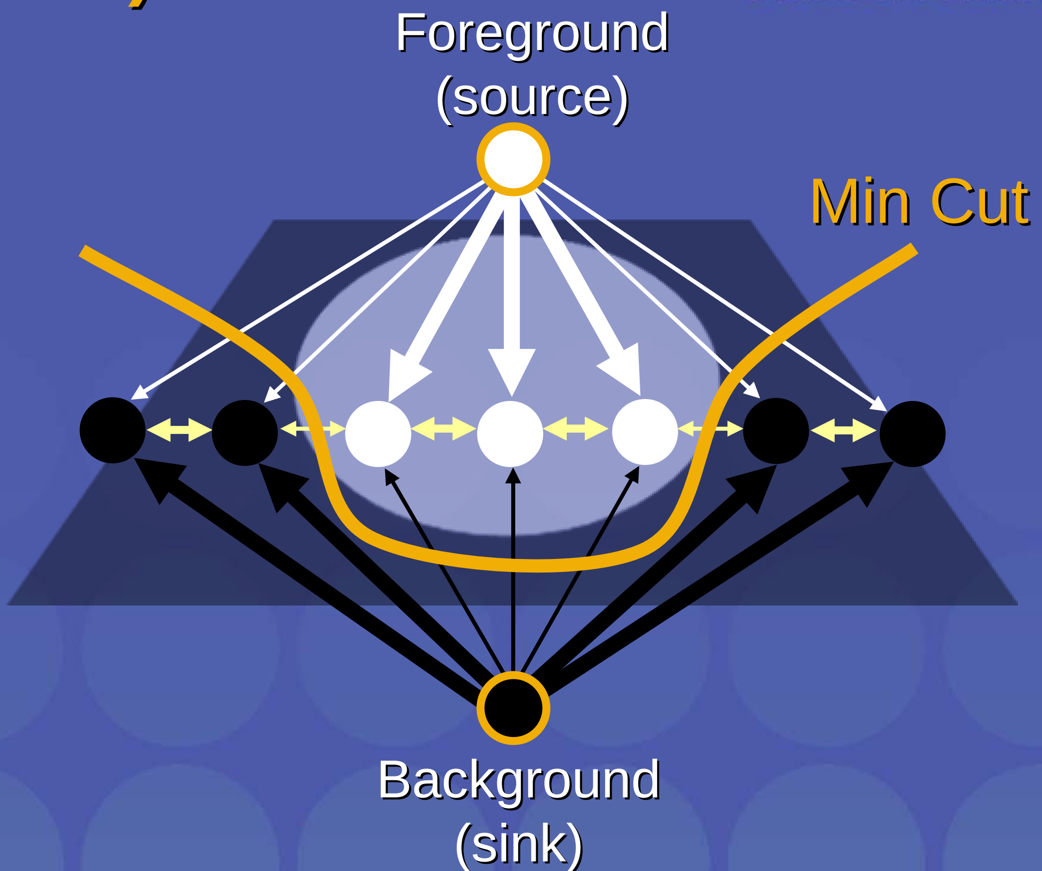
Framework

- **Input:** Image $\mathbf{x} \in \{\mathrm{R}, \mathrm{G}, \mathrm{B}\}^{\mathbf{n}}$
- **Output:** Segmentation $\mathbf{S} \in \{\mathbf{0}, \mathbf{1}\}^{\mathbf{n}}$
- **Parameters:** Colour $\boldsymbol{\Theta}$,Coherence λ
- **Energy:** $E(\boldsymbol{\Theta}, \mathbf{S}, \mathbf{x}, \lambda) = E_{Col} + E_{Coh}$
- **Optimization:** $\arg\min_{\mathbf{S}, \boldsymbol{\Theta}} E(\mathbf{S}, \boldsymbol{\Theta}, \mathbf{x}, \lambda)$

Graph Cuts

Boykov and Jolly (2001)





Cut: separating source and sink; Energy: collection of edges

Min Cut: Global minimal energy in polynomial time

Iterated Graph Cut

User Initialisation

$$\arg\min_{\Theta} E(\mathrm{S}, \Theta, \mathrm{x}, \lambda)$$

$$\arg\min_{\mathrm{S}} E(\mathrm{S}, \Theta, \mathrm{x}, \lambda)$$

K-means for learning colour distributions

Graph cuts to infer the segmentation

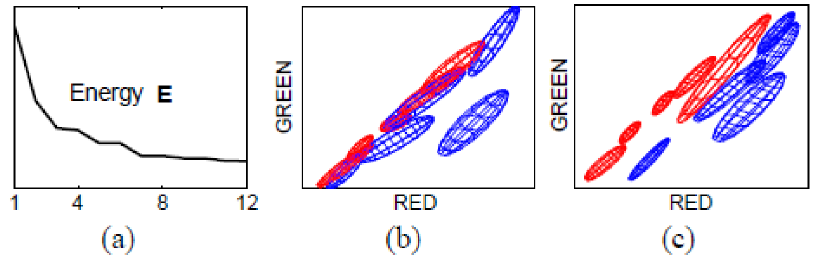
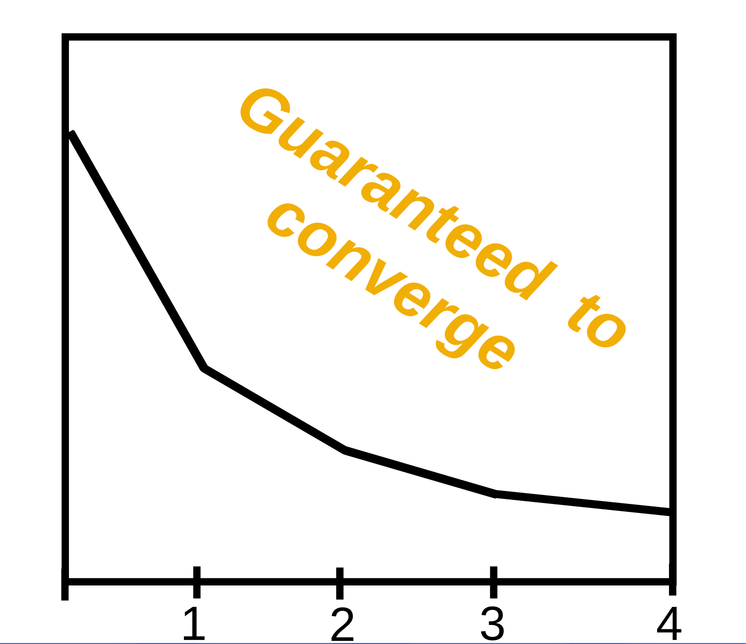


Figure 4: **Convergence of iterative minimization** for the data of fig. 2f. (a) The energy **E** for the llama example converges over 12 iterations. The GMM in RGB colour space (side-view showing R,G) at initialization (b) and after convergence (c). $K = 5$ mixture components were used for both background (red) and foreground (blue). Initially (b) both GMMs overlap considerably, but are better separated after convergence (c), as the foreground/background labelling has become accurate.

Iterated Graph Cuts

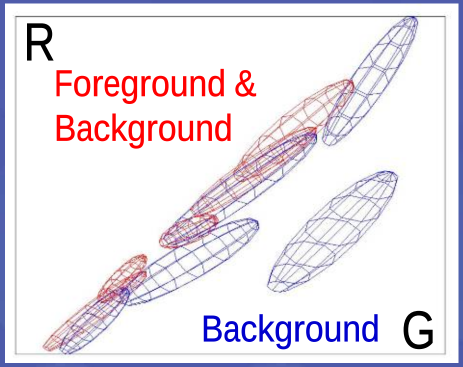
SIGGRAPH2004

Result

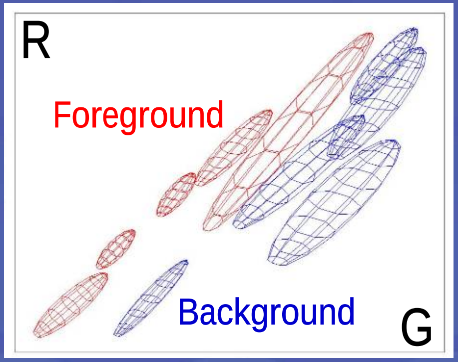


Energy after each Iteration

Colour Model



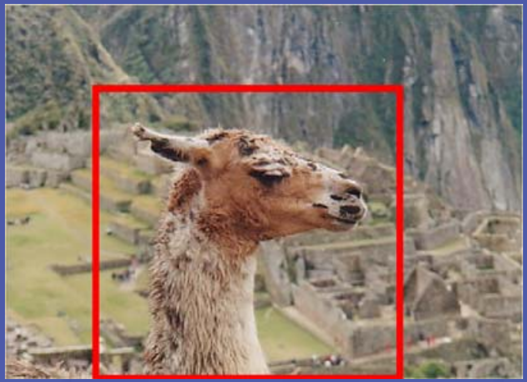
Iterated graph cut



Gaussian Mixture Model (typically 5-8 components)

$$E_{Col}(\Theta, \mathrm{S}, \mathrm{x}) = \sum_{\mathrm{n}} \mathrm{D}(\mathrm{S_n}, \Theta, \mathrm{x_n})$$

Coherence Model



An object is a coherent set of pixels:

$$E_{coh}(\mathbf{S},\mathbf{x},\lambda) = \lambda \sum_{i,j \text{ adj.}} (S_i \neq S_j) \; exp\{-\frac{1}{2\sigma^2}||x_i - x_j||^2\}$$

$\lambda = 0$

$\lambda = 50$

$\lambda = 1000$

Blake et al. (2004): Learn Θ, λ jointly

Moderately straightforward examples

… GrabCut completes automatically

Difficult Examples

SIGGRAPH2004

Camouflage & Low Contrast | **Fine structure** | **No telepathy**

Initial Rectangle

Initial Result

Evaluation – Labelled Database

Available online: http://research.microsoft.com/vision/cambridge/segmentation/