Markov Random Field Models and Optimization in image processing and vision

~~(MLSS, Tuebingen, 2007)~~

short version

TuebingenAB07.pdf

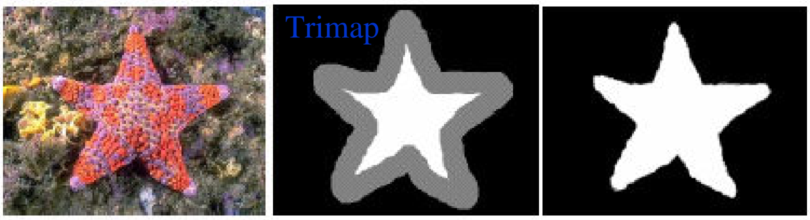
on

http://www.research.microsoft.com/~ablake/talks

Spatial inference problems in vision and image processing

- Image restoration
- Foreground segmentation
- Edge detection
- Matting (interactive graphics)
- Stereo matching and occlusion
- 3D surface reconstruction

Segmentation



- Image $\mathbf{z} = (z_1, \ldots z_N),\ z_n \in \mathbf{R}^3$
- Opacity $\mathbf{x} = (x_1, \ldots x_N),\ x_n \in \mathbf{R}$
- Hard segmentation $x_n \in \{0, 1\}$

coherent foreground

Simple segmentation --- Ising prior

Markov Random Field -- MRF

$$p(\mathbf{x} \mid \mathbf{z}) \propto \exp -E \quad \text{with} \quad E = U + V$$

where

$$U(\mathbf{x}, \mathbf{z}) = \sum_k f_k(z_k \mid x_k)$$

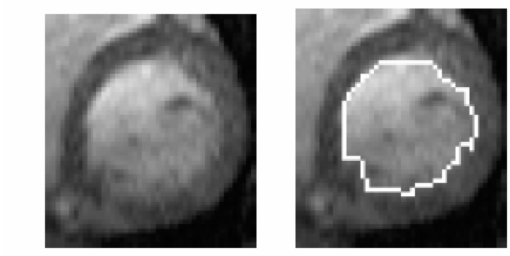
and

$$V(\mathbf{x}) = \gamma \sum_{k,k'} |x_k - x_{k'}|$$

(-ve) log-prior $V(\boldsymbol{x})$

Segmentation artefacts --- Ising prior

(Boykov and Kolmogorov ICCV 2003)



?? How to overcome artefacts

Boykov-Jolly contrast-sensitive segmentation

(Boykov and Jolly 2001; Rother et al. 2004; Li, Shum et al. 2004)

- Conditional Random Field -- CRF

  (Lafferty et al. 2001; Kumar and Hebert 2003)

$$p(\mathbf{x} \mid \mathbf{z}) \propto \exp -E \quad \text{with} \quad E = U + V$$

where now

$$V = \gamma \sum_{k,k'} |x_k - x_{k'}| \left(\exp - \frac{\|z_k - z_{k'}\|^2}{2\sigma^2} \right)$$

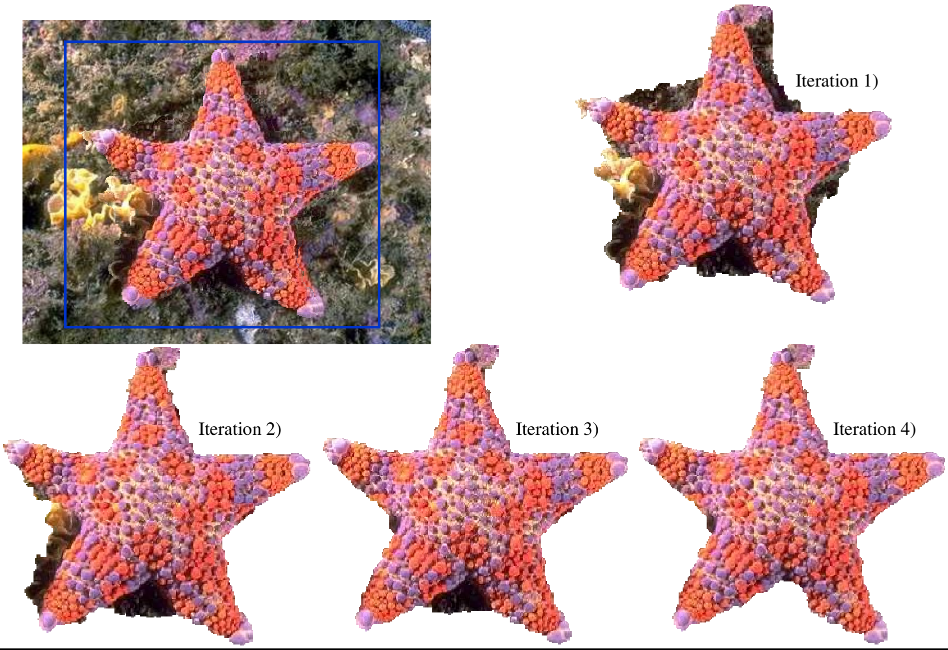
log-"prior" *V*(***x***,***z***)

data-dependence

?? How to compute $\max_{\mathbf{x}} \; p(\mathbf{x} \mid \mathbf{z})$ ie $\min_{\mathbf{x}} \; E(\mathbf{x})$

GrabCut: partially supervised inference

(Rother, Kolmogorov and Blake, Siggraph 2004;)



Optimizing Markov Random Fields

Generally NP-hard, so approximate:

Simulated annealing [Metropolis, Rosenbluth, Rosenbluth, Teller and Teller, 1953]

Gibbs sampling [Geman and Geman 1984]

Iterated conditional Modes ["*On the statistical analysis of dirty pictures",* Besag 1986]

Approximate variational extremum [Mumford and Shah 1985,9]

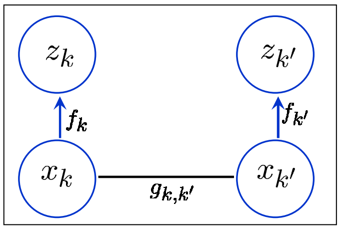
Graduated nonconvexity [Blake and Zisserman 1987]

Graph cut [Greig, Porteous and Seheult, 1989]

Loopy Belief Propagation [Freeman and Pasztor, 1999]

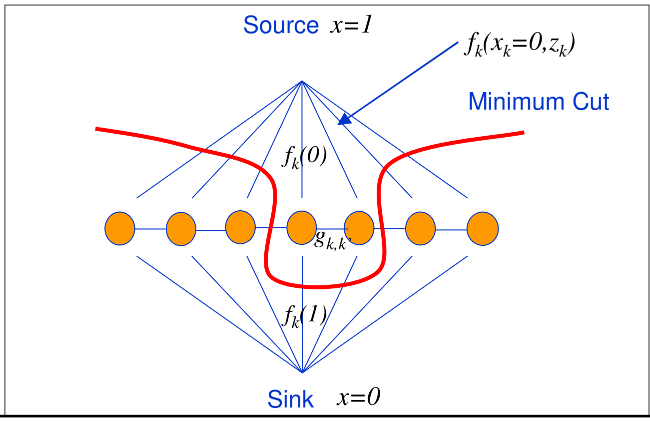
"Modern" graph cut [Boykov, Veksler and Zabih, 2001]

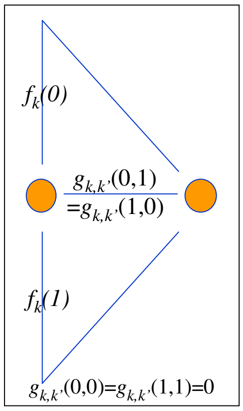
Graph Cut engine for Markov segmentation



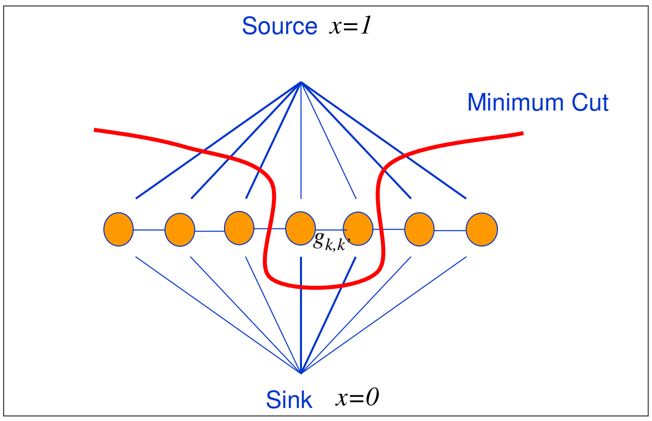
binary optimization: $x_k \in \{0,1\}$

$$E(\mathbf{x},\mathbf{z}) = \sum_k \sum_{k'} g_{k,k'}(x_k, x_{k'}) + \sum_k f_k(x_k, z_k)$$





Ford-Fulkerson Min-cut/Max Flow



* Max flow you can push through the network = min cut
  -- ie capacity of cut with smallest total capacity
* Links saturated by max flow = links separated by min cut

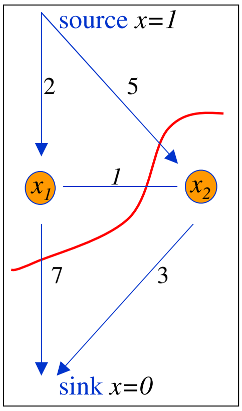
Example: optimization as graph cut problem

Problem: $\min_{\mathbf{x}} \; E = f_1(x_1) + f_2(x_2) + g(x_1,x_2)$

where: $f_1(0)=2; f_1(1)=7; f_2(0)=5; f_2(1)=3;$ [all weights +ve]

$g(0,1)=g(1,0)=1; \; g(0,0)=g(1,1)=0;$ [canonical form]

Graph:

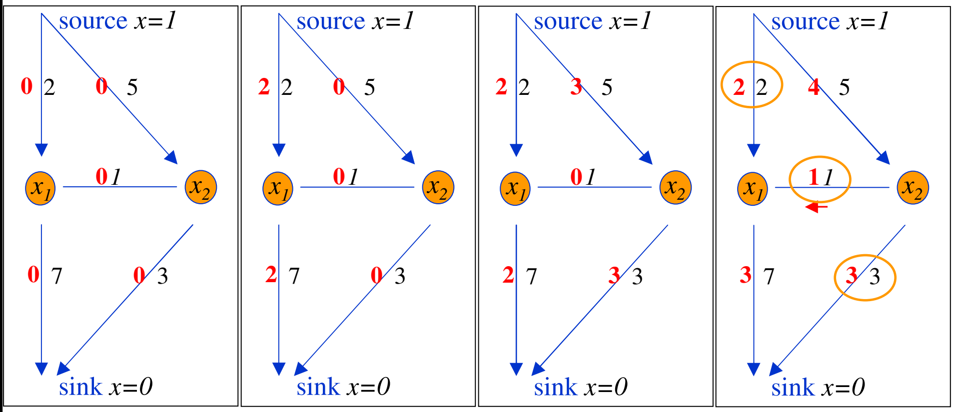


Trial Solution $x_1=1, x_2=0$

Energy: $E(x_1, x_2) = 7+1+5 = 13$

Example: graph cut optimization – min cut/max flow

Solve (exactly) by augmenting flow:



Solution from saturated paths

$x_1=0, x_2=1 \quad E(x_1, x_2) = 2+1+3 = 6$

Augmenting flow

[Boykov and Kolmogorov, PAMI 2004]

-- Breadth-first search for augmenting paths

-- Augment flow along non-saturated paths

-- Termination if weights all positive

-- Typical complexity $O(n^3)$ [Dinic algorithm]

-- Special algorithms for vision problems – wide, shallow graphs

http://www.adastral.ucl.ac.uk/~vladkolm/software.html

Graph cut – regularity (submodularity)

What energy functions E can be minimized by graph cut?

-- Augmenting paths terminates finitely if all costs +ve.

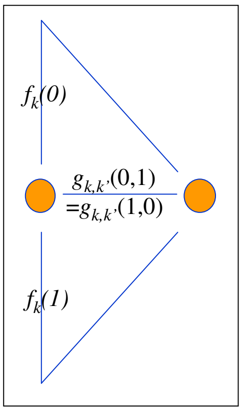
-- WLOG reduce to canonical form

define $D = g_{k,k'}(0,1) + g_{k,k'}(1,0) - g_{k,k'}(0,0) - g_{k,k'}(1,1)$

canonical form: $g_{k,k'}(0,0) = g_{k,k'}(1,1) \rightarrow 0;$

$g_{k,k'}(0,1) = g_{k,k'}(1,0) \rightarrow D/2;$

$f_k(x_k) \geq 0$



-- Solvability if $D \geq 0$ [invariant to transfms of problem]

Regularity [Kolmogorov and Zabih 2004]

Regularity: achieving canonical form

[Kolmogorov and Zabih PAMI 2004]

Simplifying unary $f_k(x_k)$:

$$m = min(f_k(0), f_k(1))$$

if $m<0$

$$f_k(0) := f_k(0) - m$$

$$f_k(1) := f_k(1) - m$$

-- reparameterization

??? Simplifying pairwise $g_{k,k'}(x_k, x_{k'})$:

... Regularity: achieving canonical form

[Kolmogorov and Zabih 2004]

Simplifying pairwise $g_{k,k'}(x_k, x_{k'})$:

$g(0,0)$	$g(0,1)$
$g(1,0)$	$g(1,1)$

→

0	$g(0,1) - g(0,0)$
$g(1,0) - g(0,0)$	$g(1,1) - g(0,0)$

→

0	0
$g(1,0) - g(0,0)$	$g(1,1) - g(0,1)$

→

0	0
$g(1,0) - g(0,0) - g(1,1) + g(0,1)$	0

D

-- reparameterization

Graph cut – regularity (submodularity)

Do they satisfy regularity?

Ising prior $$V(\mathbf{x}) = \gamma \sum_{k,k'} |x_k - x_{k'}|$$

?

$g(x,x')$	$x = 0$	$x = 1$
$x' = 0$	0	γ
$x' = 1$	γ	0

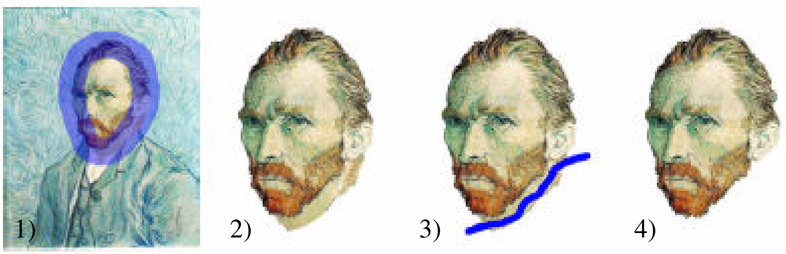
Ising with contrast sensitivity

$$V(\mathbf{x},\mathbf{z}) = \gamma \sum_{k,k'} |x_k - x_{k'}| \left(\exp - \frac{\|z_k - z_{k'}\|^2}{2\sigma^2} \right)$$

?

Dynamic graph cut – Markov editing.

Boykov, Y., Jolly, M.P.: Interactive graph cuts for optimal boundary and region segmentation of objects in N-D images. Proc CVPR 2001.



Dynamic graph cuts

(Boykov and Jolly, 2001; Kohli & Torr ICCV 2005)

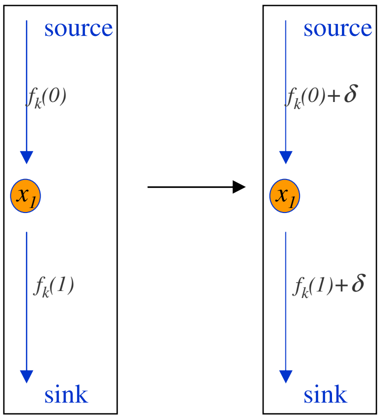
Changing the unaries:

Re-use old flow with new Energy fn ???problem

Where flow exceeds capacity by δ, reparametrise:

$f_k(0) := f_k(0) + \delta$

$f_k(1) := f_k(1) + \delta$



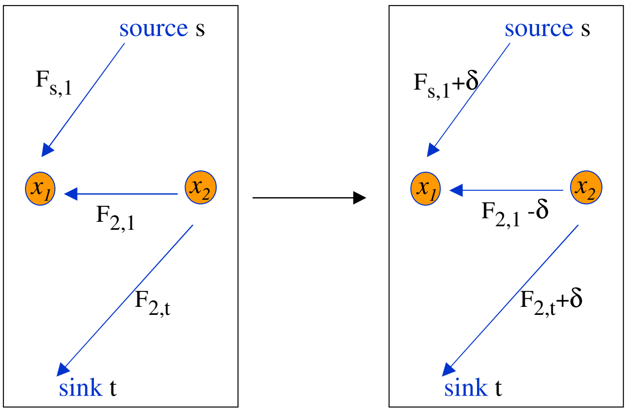
???Changing the pairwise potentials

Dynamic graph cuts

(Boykov and Jolly, 2001; Kohli & Torr ICCV 2005)

Changing the binaries:

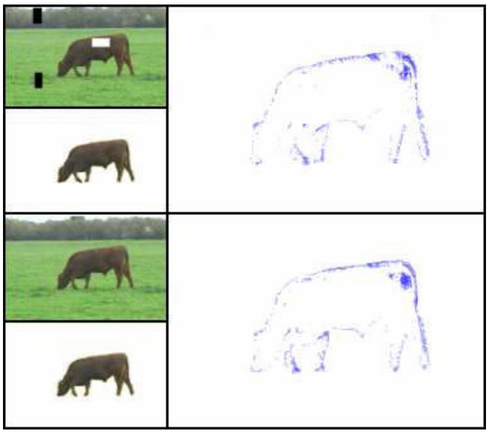
Where flow, eg $F_{2,1}$ exceeds capacity by δ, reparametrise:



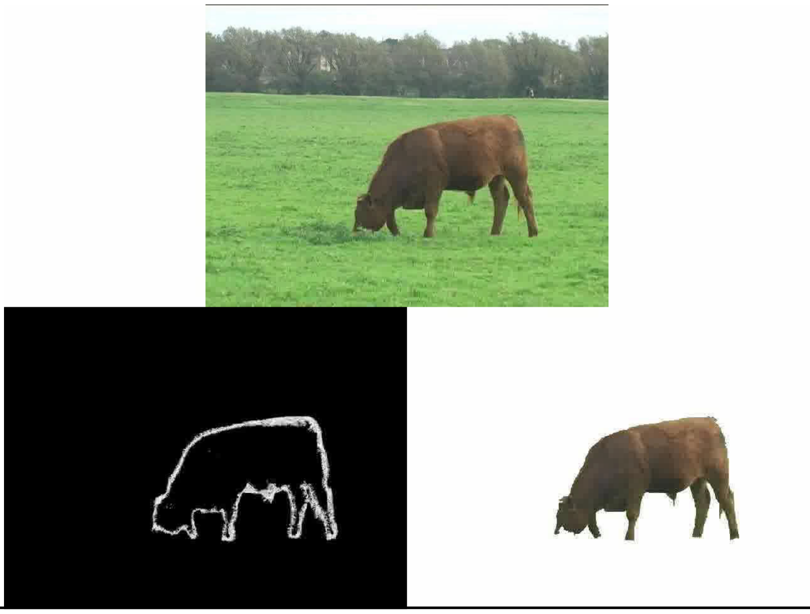
-- then follow procedure for unary flow excess

??? what's this good for

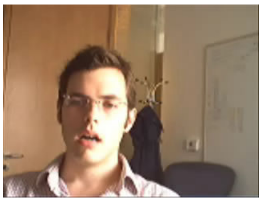
Dynamic graph cuts
(Kohli & Torr ICCV 2005)



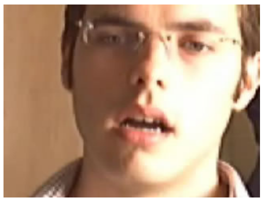
Dynamic graph cuts
(Kohli & Torr ICCV 2005)



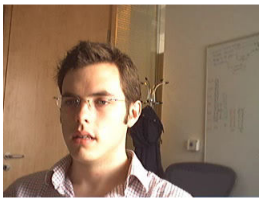
Stereo webcam

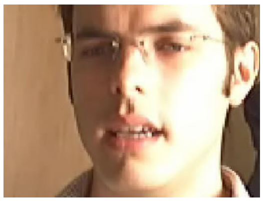


conventional camera
- zoomed out



conventional camera
- zoomed in



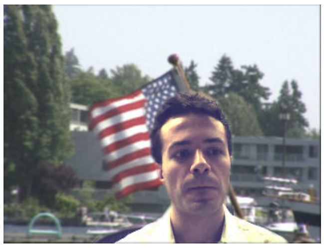


Smart Framing - start

Smart Framing - continue

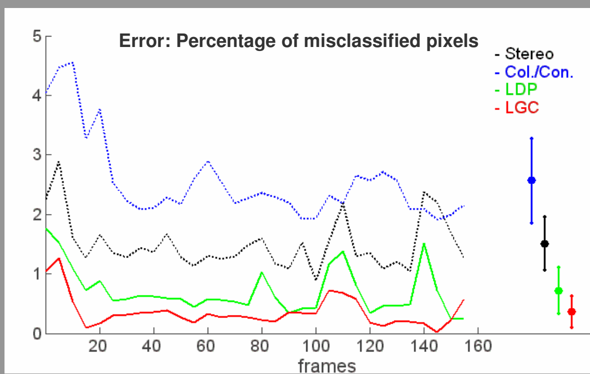
Live Stereo Segmentation

Background substitution

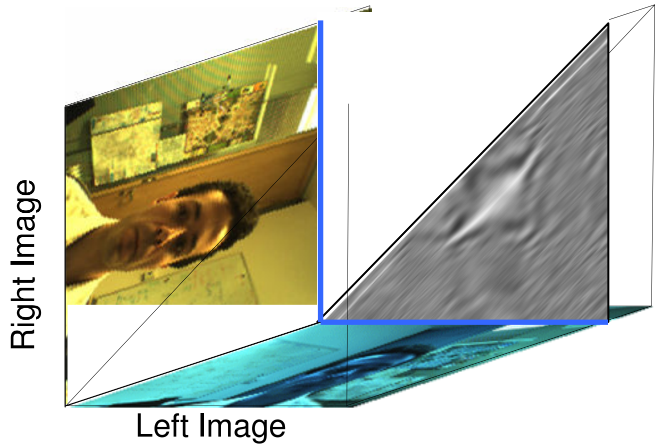


Kolmogorov, Criminisi, Blake, Cross and Rother (CVPR 2005, PAMI 2006)

Segmentation error rates



Stereo matching – full depth



Stereo matching -- cost function

(Ohta & Kanade, 1985; Cox, Hingorani & Rao, 1996; Belhumeur 1996)

Minimize: $\min_{\mathbf{x}} F(\mathbf{x}, \mathbf{z})$

$$F(\mathbf{x}, \mathbf{z}) = \sum_k g_k(x_k, x_{k-1}) + \sum_k f_k(x_k, x_{k-1}, \mathbf{z})$$

log-prior $V(\boldsymbol{x})$ — log-likelihood $U(\boldsymbol{x},z)$

"Potts"

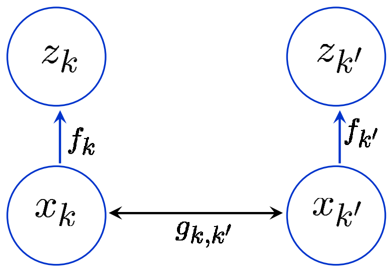
Note: $\boldsymbol{x}$ now integer valued

1D DP: Dijkstra

?? 2D algorithm

Markov Random Field (MRF) for real-valued problems

Integer-valued problems: restoration, stereo, Tapestry, obj-rec with parts.



$x \in \{1,\ldots,L\}$

$$p(\mathbf{x} \mid \mathbf{z}) \propto \exp -E(\mathbf{x},\mathbf{z})$$

$$E(\mathbf{x},\mathbf{z}) = \underbrace{\sum_{(k,k')\in\mathcal{N}} g_{k,k'}(x_k, x_{k'})}_{\text{log-prior } V(\boldsymbol{x})} + \underbrace{\sum_k f_k(x_k, z_k)}_{\text{log-likelihood } U(\boldsymbol{x},z)}$$

log-prior $V(\boldsymbol{x})$ log-likelihood $U(\boldsymbol{x},z)$

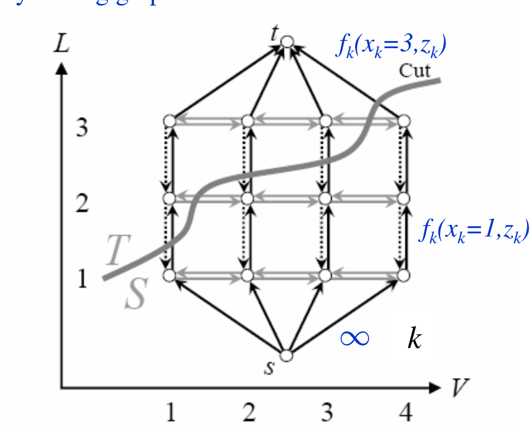
Graph cut – expanded graph

(Roy and Cox, ICCV 1998)

Problem graph of size *n* nodes, with *k* integer levels:

$$E(\mathbf{x},\mathbf{z}) = \sum_{k,k'} g_{k,k'}(x_k, x_{k'}) + \sum_k f_k(x_k, z_k)$$

-- optimize by cutting graph of *nk*+2 nodes



(Ishikawa, 03)

?? g

Graph cut – α-expansion

Optimizer for

$$E(\mathbf{x},\mathbf{z}) = \sum_{k,k'} g_{k,k'}(x_k, x_{k'}) + \sum_k f_k(x_k, z_k)$$

with $x \in \{1,....,L\}$

– repeated binary optimization:

-- cycle over α in the range of x :

$$\min_{\mathbf{y}} E(\mathbf{x},\mathbf{z}) \text{ where } y_k \in \{0,1\} \text{ and } x_k \to y_k\alpha + (1-y_k)x_k$$

?? escapes limitations on g

Regularity – α-expansion

(Kolmogorov and Zabih, PAMI 2004)

Back to integer-valued problems: restoration, stereo, Tapestry, obj-rec with parts.

Given: $y_k \in \{0,1\}$ where $x_k \to y_k\alpha + (1-y_k)x_k$

$$g^y(y,y') = g(y\alpha + (1-y)x,\ y'\alpha + (1-y')x')$$

giving

$g^y(y,y')$		y = 0	y = 1
y'	0	$g(x,x')$	$g(x,\alpha)$
	1	$g(\alpha,x')$	$g(\alpha,\alpha)$

then regularity requires (diagonal subdominance):

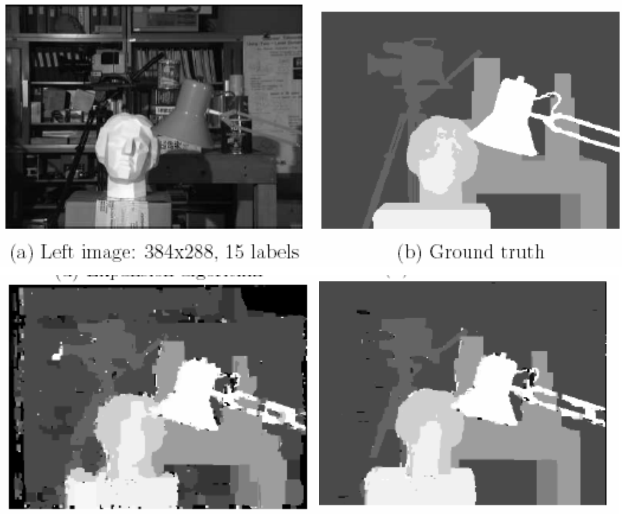
$$g(x,x') + g(\alpha,\alpha) \leq g(x,\alpha) + g(\alpha,x')$$

satisfied by any metric g (since: $g \geq 0$; $g(\alpha,\alpha)=0$; triangle ineq)

-- eg Potts

Graph Cut for stereo (α-expansion)

(Boykov, Veksler and Zabih, IEEE PAMI 2001)

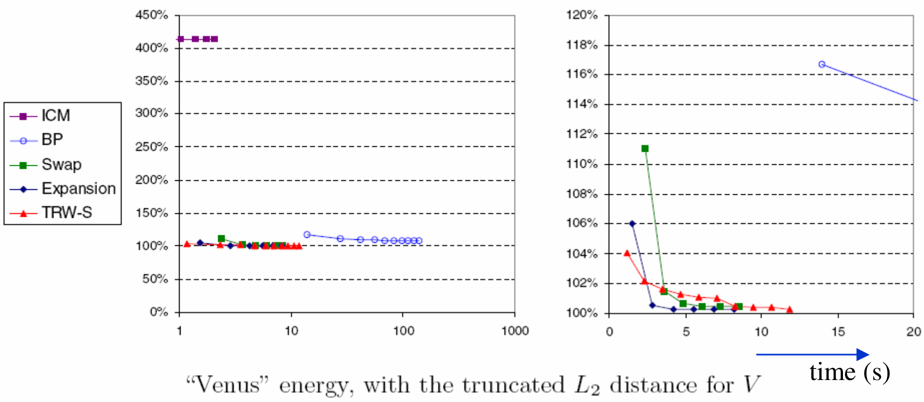


(a) Left image: 384x288, 15 labels (b) Ground truth

(f) Simulated annealing (d) Expansion algorithm

Graph cut integer optimization -- is it an advance?

(Szeliski et al., ECCV 2006, PAMI in press)



“Venus” energy, with the truncated L_2 distance for V

Graph cut variants and issues

Prior modified by data: Boykov, Y., Jolly, M.P.: Interactive graph cuts for optimal boundary and region segmentation of objects in N-D images. Proc CVPR 2001.

Alpha Expansion: Y. Boykov, O. Veksler, and R. Zabih. Fast approximate energy minimization via graph cuts. IEEE Trans. on Pattern Analysis and Machine Intelligence, 23(11), 2001.

Submodularity: V. Kolmogorov and R. Zabih. What energy functions can be minimized via graph cuts? In Proc. European Conf. Computer Vision, pages 65–81, 2002.

Truncation: C. Rother, S. Kumar, V. Kolmogorov, and A. Blake. Digital tapestry. In Proc. Conf. Computer Vision and Pattern Recognition, 2005.

QPBO: P. Hammer, P. Hansen, and B. Simeone. Roof duality, complementation and persistency in quadratic 0-1 optimization. Mathematical Programming, 28:121–155, 1984.

QPBO: V. Kolmogorov and C. Rother. Minimizing non-submodular functions with graph cuts—a review. In IEEE Trans PAMI 29, 7, 2007.

TRW: V. Kolmogorov and M. Wainwright. On the optimality of tree-reweighted max-product message passing. In Conf. Uncertainty in AI, 2005.

Graph-cut vs. BP/TRW: S. Szeliski, R. Zabih, D. Scharstein, O. Veksler, V. Kolmogorov, A. Agarwala, M. Tappen, and C. Rother. A comparative study of energy minimization algorithms for highly connected graphs. In Proc. European Conf. Computer Vision, 2006.