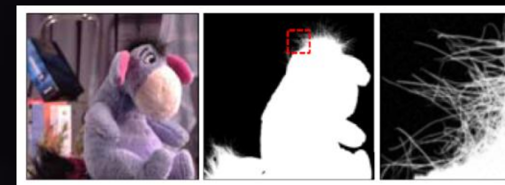
TRACTABLE HIGHER ORDER MODELS IN COMPUTER VISION

Carsten Rother

MSR-INRIA Workshop

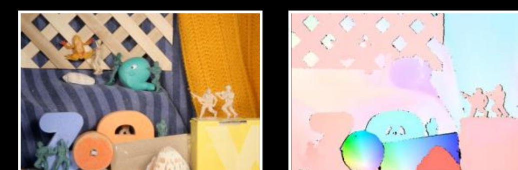
What I am interested in …

Applications:

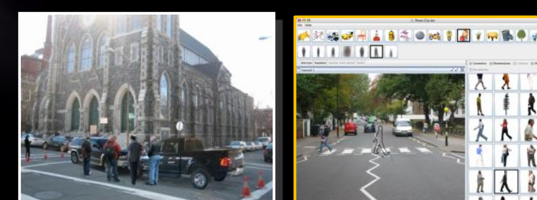


- Segmentation and Matting

- Low Level Vision:
  3D reconstruction, Stereo, Optical Flow, Image de-noising, de-convolution



- Vision for Graphics
- Object recognition



Optimisation:

- Efficient methods for solving MRFs
- Recently: global (higher-order) Models

This talk …

- GrabCut: Interactive Foreground Extraction using Iterative Graph Cut [Rother et al. Siggraph ’04]
- Branch-and-MinCut [Lempitsky et al. ECCV ’08]
- Joint optimization of Segmentation and Appearance Models [Vicente et al. ICCV ‘09]
- A global Perspective on MAP Inference in Low-Level Vision [Woodford et al. ICCV 09]
- Minimizing Sparse Higher Order Energy Functions of Discrete Variables[Rother et al. CVPR 09]

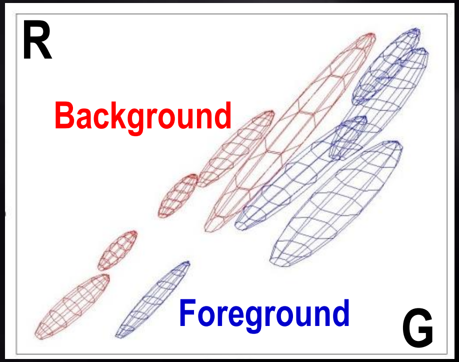
Markov Random Field - Segmentation

[Boykov, Jolly ICCV '01]

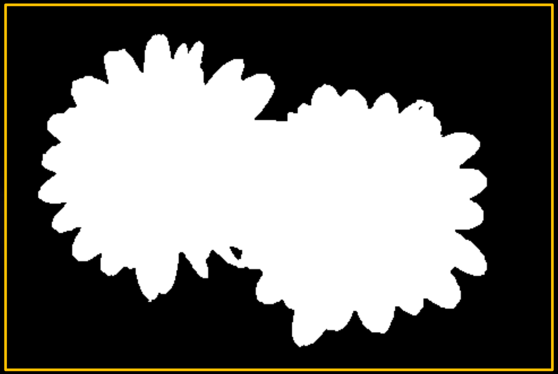
$$E(x) = \sum_{p \in V} F_p(\theta^F) x_p + B_p(\theta^B)(1 - x_p) + \sum_{pq \in E} w_{pq} |x_p - x_q|$$

$$F_p = -\log Pr(z_p | \theta_F) \qquad B_p = -\log Pr(z_p | \theta_B)$$

Image **z**



$\theta^{F/B}$ Gaussian Mixture models

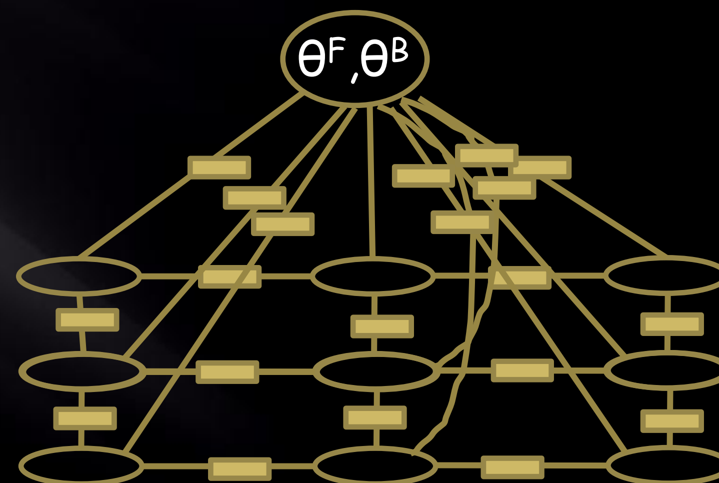


Output **x**

Graph Cut (Min Cut): global optimum in polynomial time

MRF with global Paramaters
[GrabCut; Siggarph '04]

$$E(x,\theta^F,\theta^B) = \sum_{p\in V} F_p(\theta^F)x_p + B_p(\theta^B)(1-x_p) + \sum_{pq\in E} w_{pq}|x_p - x_q|$$



Problem: for unknown θ^F,θ^B the optimization is NP-hard!

GrabCut: Iterated Graph Cuts

[Rother et al. Siggraph '04]

Input

Output

Global Optimum
[Vicente et al. '09]

$$\min_{\theta^F, \theta^B} E(x, \theta^F, \theta^B)$$

$$\min_{x} E(x, \theta^F, \theta^B)$$

Learning of the colour distributions

Graph cut to infer segmentation

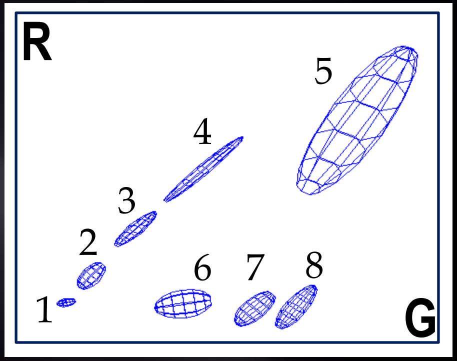
Question: How good is this procedure?

In Office 2010 (PowerPoint, word, …)

… a first attempt to solve it

[Lempisky et al. ECCV '08]

$$E(x,\theta^F,\theta^B) = \sum_{p\in V} F_p(\theta^F)x_p + B_p(\theta^B)(1-x_p) + \sum_{pq\in E} w_{pq}|x_p - x_q|$$



8 Gaussians
whole image

Model a discrete subset:

$w^F = (1,1,0,1,0,0,0,0)$; $w^B = (1,0,0,0,0,0,0,1)$

#solutions: $w^F * w^B = 2^{16}$

Global Optimum:

Exhaustive Search: 65.536 Graph Cuts

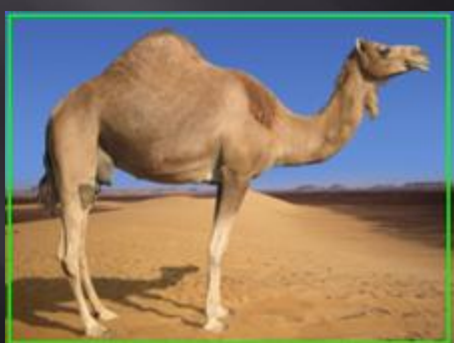
Branch-and-MinCut: ~ 130-500 Graph Cuts (depends on image)

Results …

E=-618
GrabCut

E=-624 *(speed-up 481)*
Branch-and-MinCut

E=-628
Combined



E=-593
GrabCut

E=-584 *(speed-up 141)*
Branch-and-MinCut

E=-607
Combined

Object Recognition & Segmentation

min E(x,w) with: w = Templates x Position

$|w| \sim 2.000.000$

Given exemplar shapes:

Test: *Speed-up ~900; accuracy 98.8%*

… second attempt to solve it

[Vicente et al. ICCV '09]

Eliminate global color model θ^F, θ^B :

$$E'(x) = \min_{\theta^F, \theta^B} E(x, \theta^F, \theta^B)$$

Image

discretized in bins

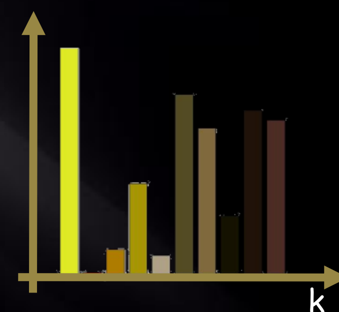
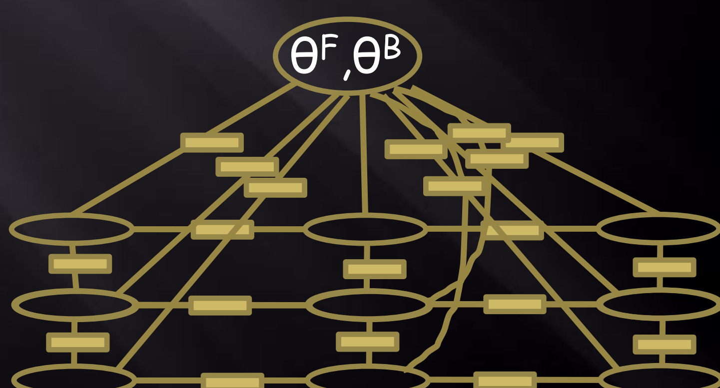
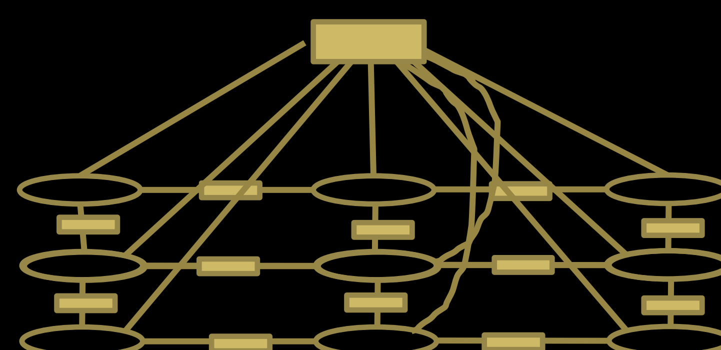


Image histogram

$K = 16^3$





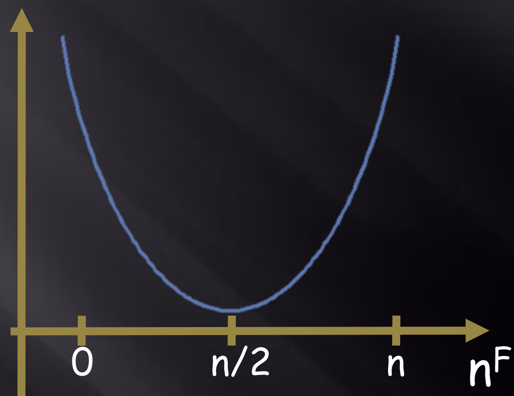
Eliminate color model

$$E'(x) = g(n^F) + \sum_k h_k(n_k^F) + \sum_{pq \in E} w_{pq} |x_p - x_q| \quad \text{with } n_k^F = \sum_{p \in V_k} x_p, \quad n^F = \sum_{p \in V} x_p$$

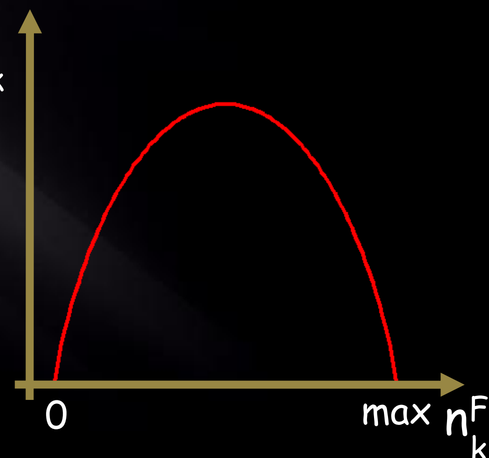
#fgd pixel in bin k

#fgd pixel

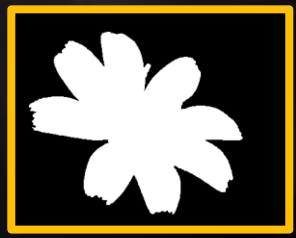
convex g



concave h_k



Prefers “equal area” segmentation



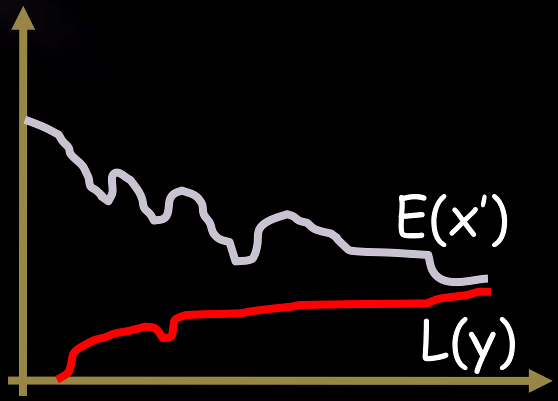
Each color either fore- or background

How to optimize … Dual Decomposition

$$E(x) = \underbrace{g(n^F)}_{E_1(x)} + \underbrace{\sum_k h_k(n^F) + \sum w_{pq}|x_p - x_q|}_{E_2(x)}$$

$$\min_x E(x) = \min_x [\, E_1(x) + y^T x + E_2(x) - y^T x \,]$$
$$\geq \min_{x'} [\, E_1(x') + y^T x' \,] + \min_x [E_2(x) - y^T x] =: L(y)$$

Goal: maximize concave function $L(y)$ using sub-gradient [Shor '70], no guarantees on E (NP-hard)



Some results…

Input | GrabCut | Global Optimum

Global optimum in 61% of cases (GrabCut database)
… but GrabCut works really well

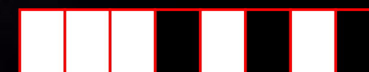
A global perspective on low level vision

[Woodford, Rother, Kolmogorov, ICCV '09]

Training data:  . . . (60% white; 40% black)

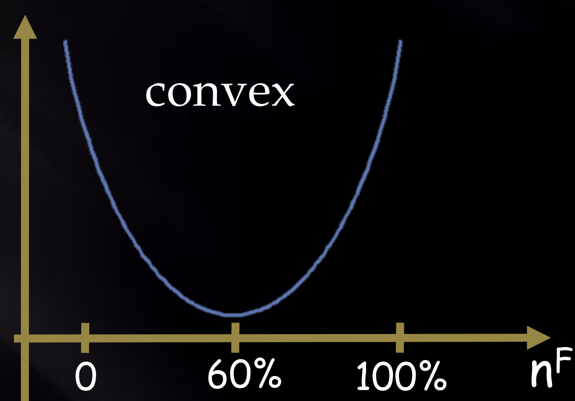
MAP: 

$\text{prior}(x) = 0.6^8 = 0.016$

Others less likely: 

$\text{prior}(x) = 0.6^5 * 0.4^3 = 0.005$

MRF is a bad prior since input marginal statistic is ignored !



Prefers "equal area" segmentation

Marginal Probability Field (MPF)

[Woodford, Rother, Kolmogorov, ICCV '09]

Training: any statistic over terms: unary, pairwise, tripple; also conditional terms.

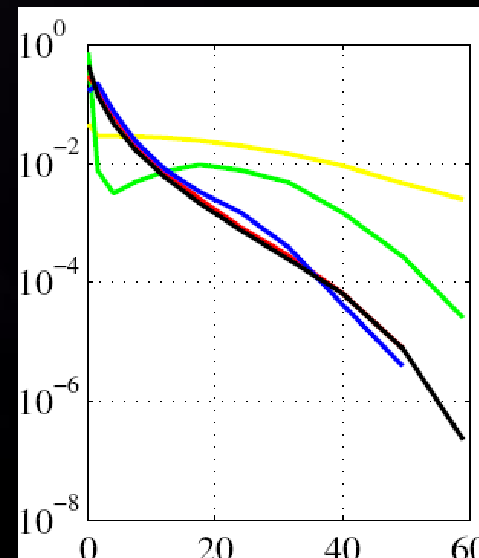
Test: preserve the training statistic

Optimisation: Dual-Decomposition

Image de-noising

original

noisy input



MRF, stronger prior

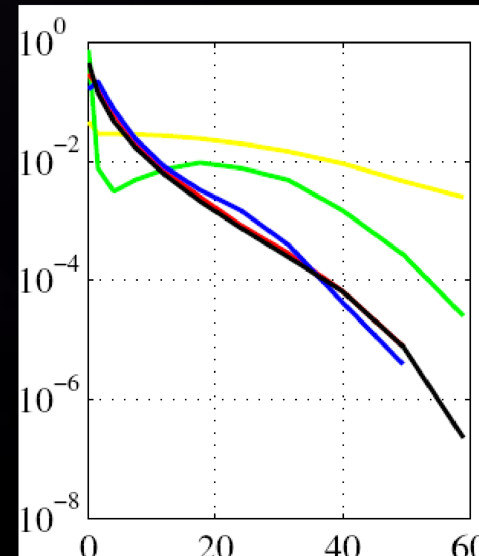
MPF

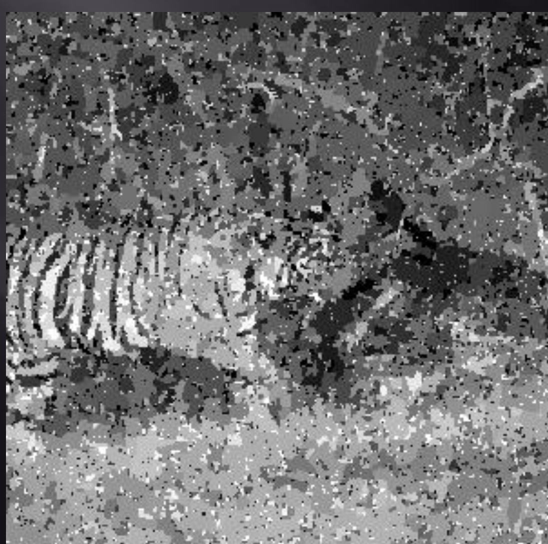
Original – blue
Noisy – yellow
MRF – green
MPF – red
Learned – black

Image de-noising

original

noisy input





MRF

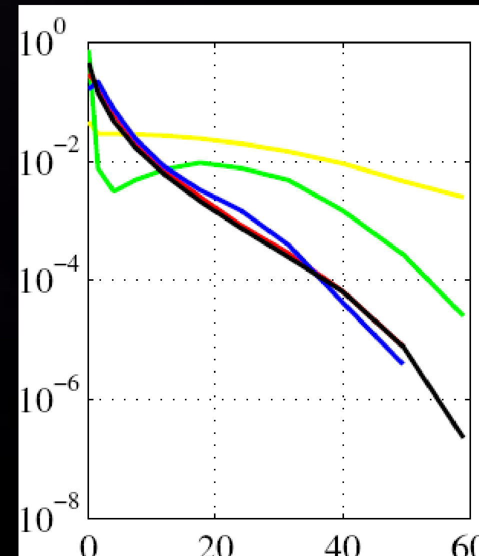
MPF

Original – blue
Noisy – yellow
MRF – green
MPF – red
Learned – black

Image de-noising

original

noisy input



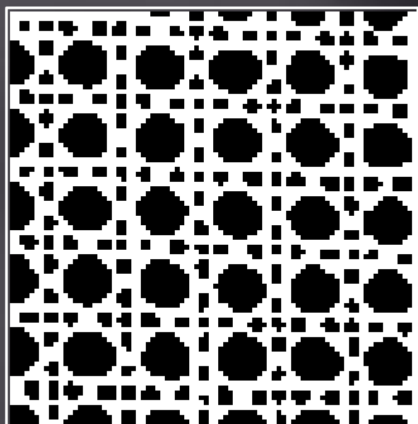
MRF, weaker prior

MPF

Original – blue
Noisy – yellow
MRF – green
MPF – red
Learned – black

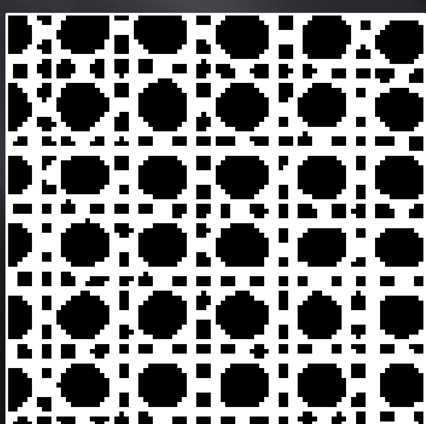
Sparse Higher-Order Function

[Rother, Kohli, Feng, Jiya, CVPR '09]



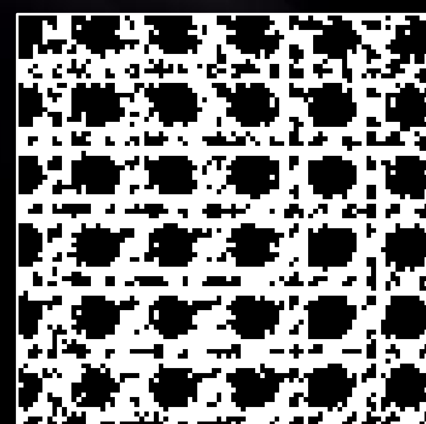
Training Image

Higher Order Structure not Preserved

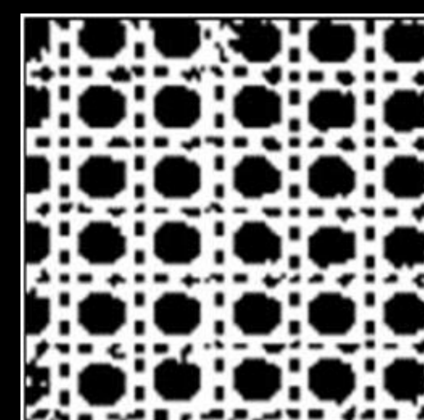


Test Image

Test Image (60% Noise)



Result 15-connected



Higher-order result

Sparse Higher-Order Function

[Rother, Kohli, Feng, Jiya, CVPR ‘09]

Minimize: $E(X) = P(X) + \sum_c h_c (X_c)$

Where: $h_c: \{0,1\}^{|c|} \rightarrow R$

Higher Order Function ($|c| = 10 \times 10 = 100$)

Assigns cost to 2^{100} possible labellings!