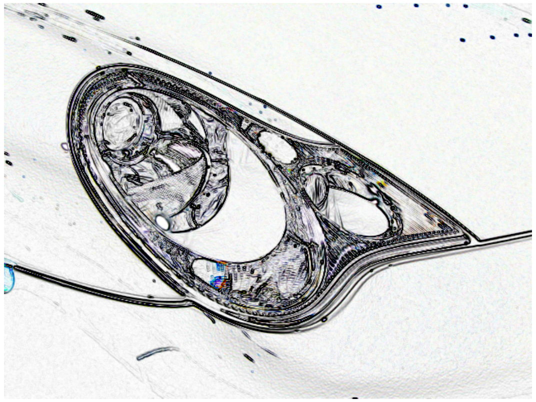
CSE578: Computer Vision

Spring 2016:

Examples of MRF Segmentation

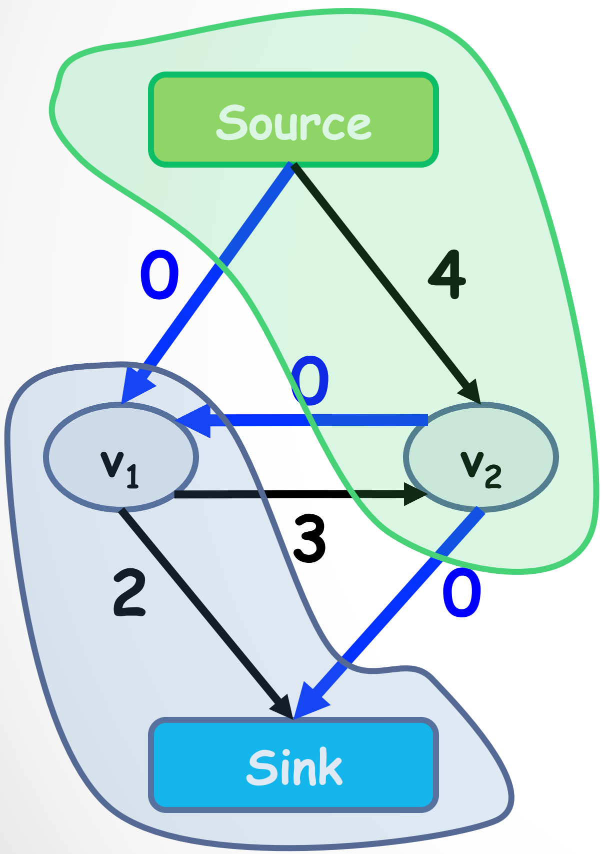
Anoop M. Namboodiri

Center for Visual Information Technology

IIIT Hyderabad, INDIA

Maxflow Algorithms

Final Result



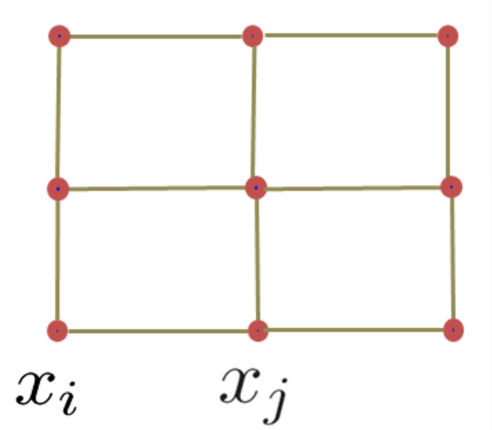
Augmenting Path Based Algorithms

1. Find path from source to sink with positive capacity
2. Push maximum possible flow through this path
3. Repeat until no path can be found

Algorithms assume non-negative capacity

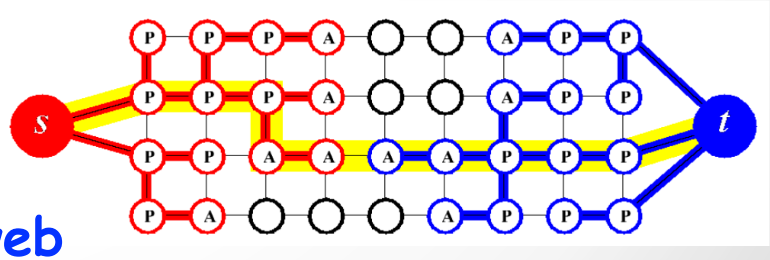
Maxflow in Computer Vision

- Specialized algorithms for vision problems
  - Grid graphs
  - Low connectivity (m ~ O(n))



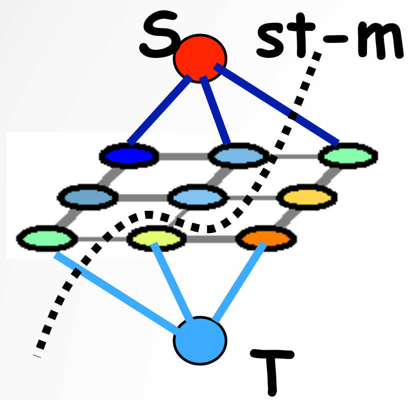
- Dual search tree augmenting path algorithm

  [Boykov and Kolmogorov PAMI 2004]
  - Finds approximate shortest augmenting paths efficiently
  - High worst-case time complexity
  - Empirically outperforms other algorithms on vision problems
  - Efficient code available on the web



- http://pub.ist.ac.at/~vnk/software.html

St-mincut and Energy Minimization



Minimizing a Qudratic Pseudoboolean function E(x)

Functions of boolean variables

$$E: \{0,1\}^n \rightarrow R$$

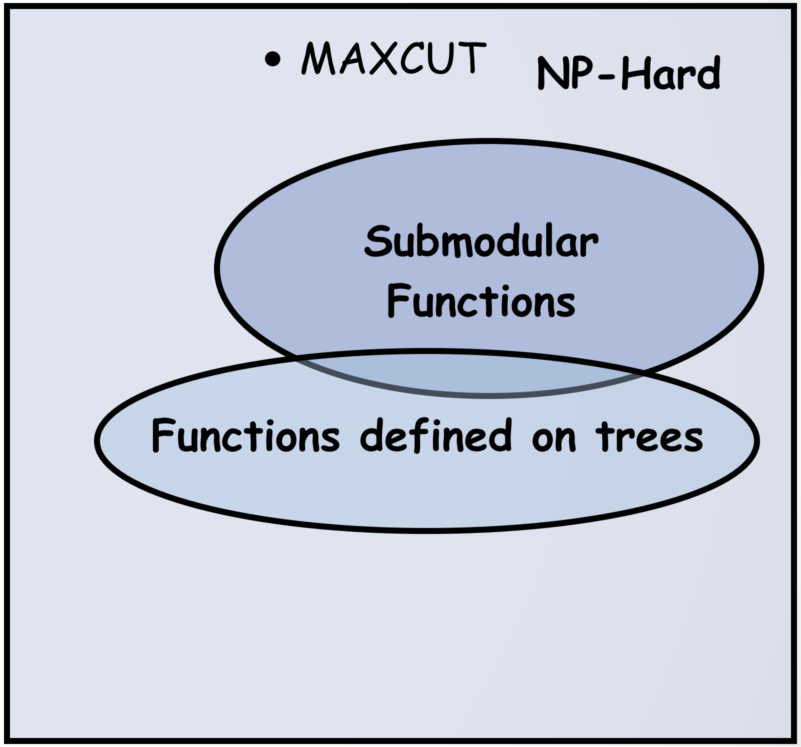
$$E(x) = \sum_i c_i x_i + \sum_{i,j} c_{ij} x_i(1-x_j)$$

$c_{ij} \geq 0$

Polynomial time st-mincut algorithms require non-negative edge weights

Minimizing Energy Functions

- General Energy Functions
  - NP-hard to minimize
  - Only approximate minimization possible
- Easy energy functions
  - Solvable in polynomial time
  - Submodular ~ $O(n^6)$



Space of Function Minimization Problems

GrabCut

Interactive Foreground Extraction using Iterated Graph Cuts

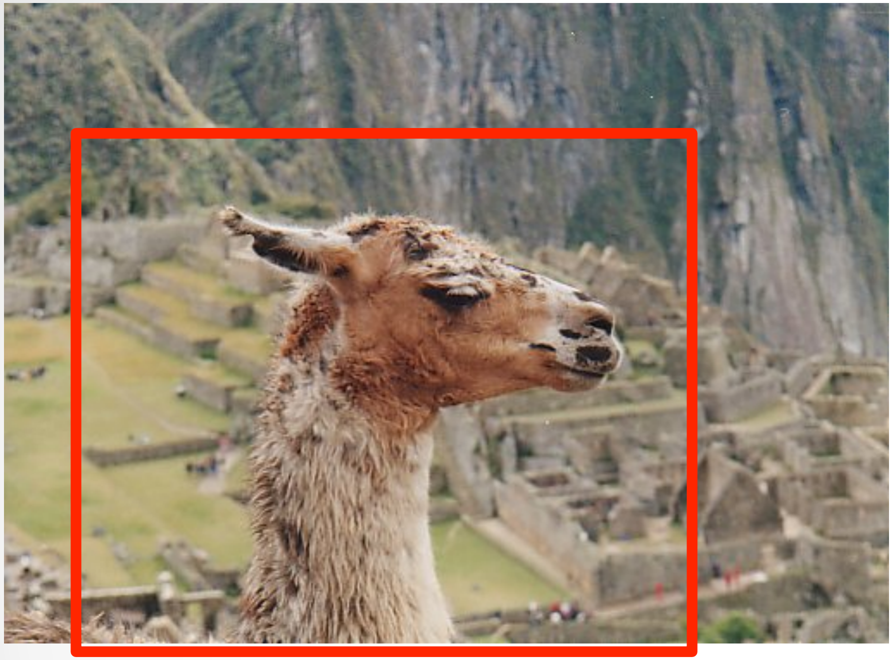
Carsten Rother

Vladimir Kolmogorov

Andrew Blake

Microsoft Research Cambridge-UK

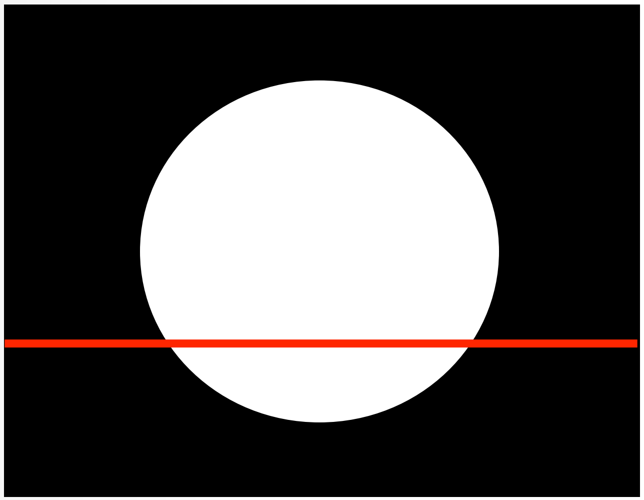
The Problem

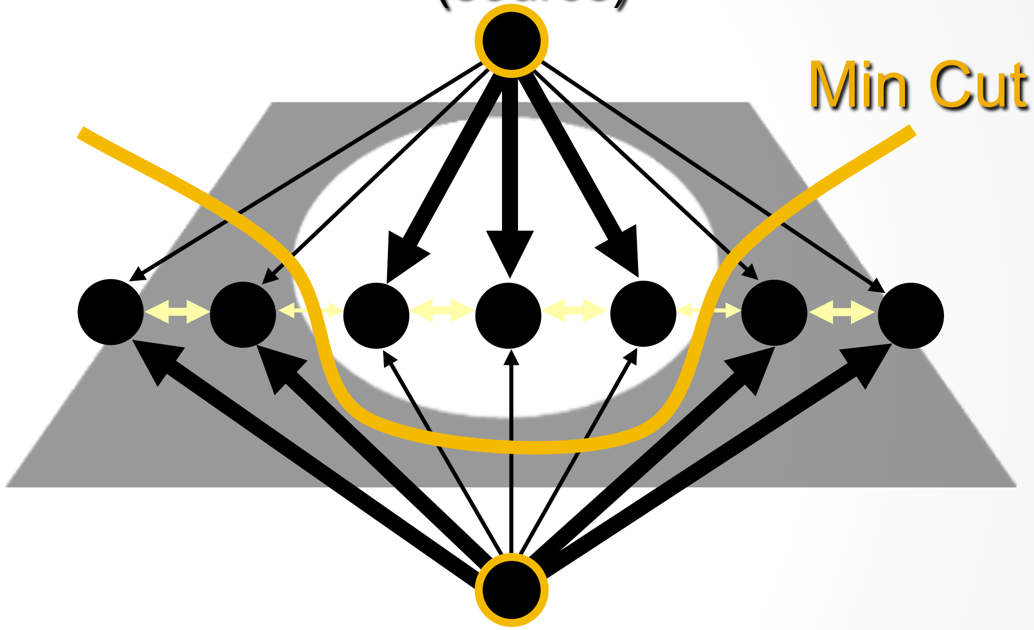


Fast &
Accurate ?

- Less user input: only rectangle
- Handle color
- Extract matte as post-process

Use Basic Graph Cut for Segmentation



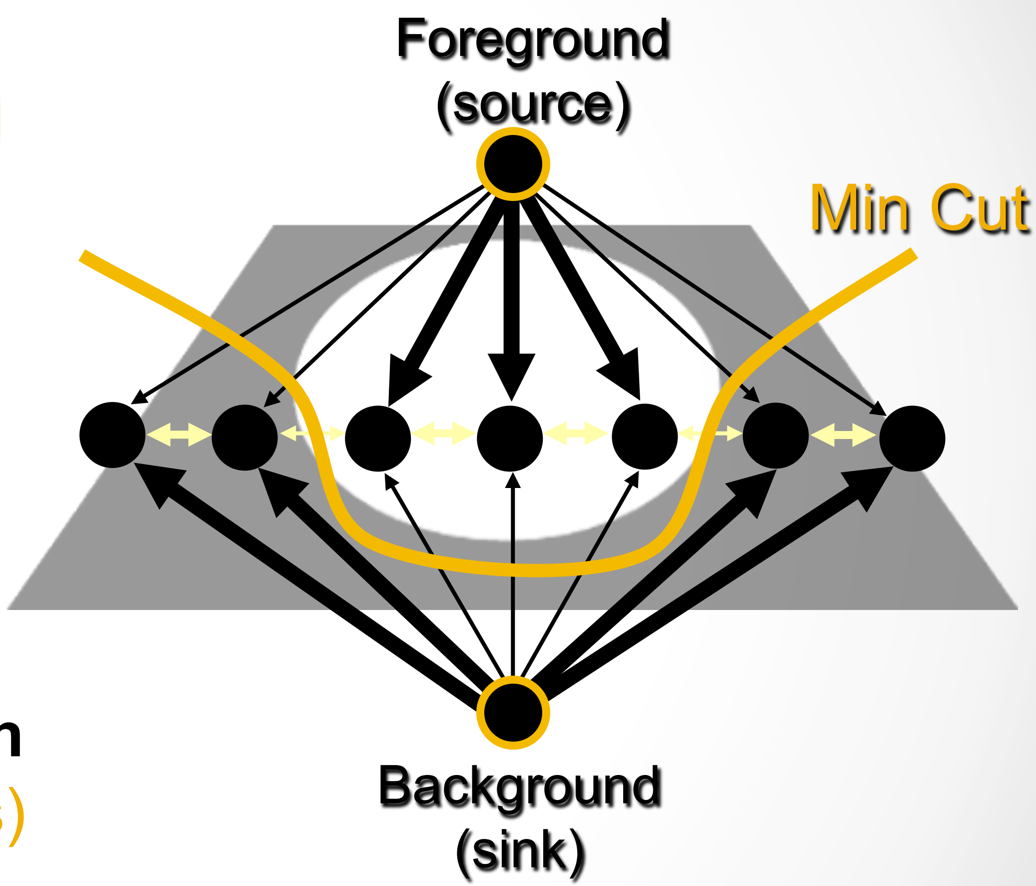


Cut: separating source and sink; Energy: collection of edges

Min Cut: Global minimal enegry in polynomial time

Graph Cuts for Foreground Extraction

Assume we know foreground is **white** and background is **black**

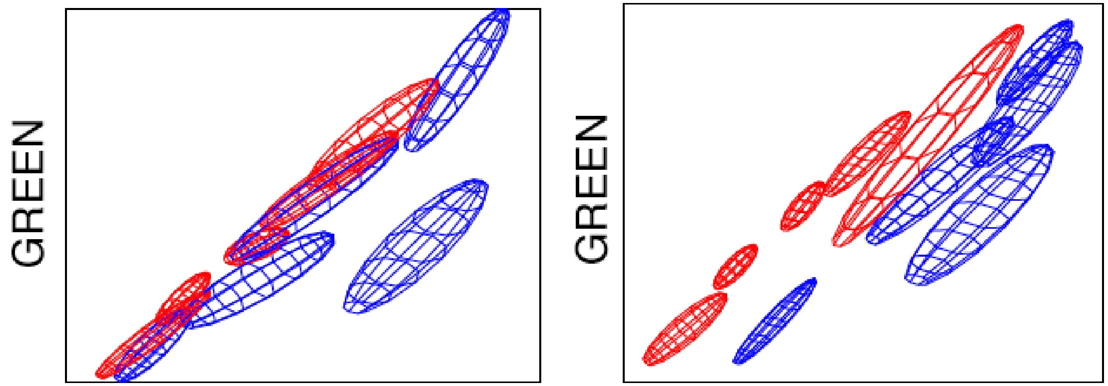


Data term = **whiteness**
(cost of assigning label)

Regularization = **color match**
(cost of separating neighbors)

Color Data Term

- Model 3D color histogram with Gaussians
  - Because brute force histogram would be sparse
  - Gaussian Mixture Model (GMM)
  - Just means histogram = sum of Gaussians
    - They advise 5 Gaussians



(b) (c)

Iterated Graph Cuts

User Initialisation

Learn foreground color model

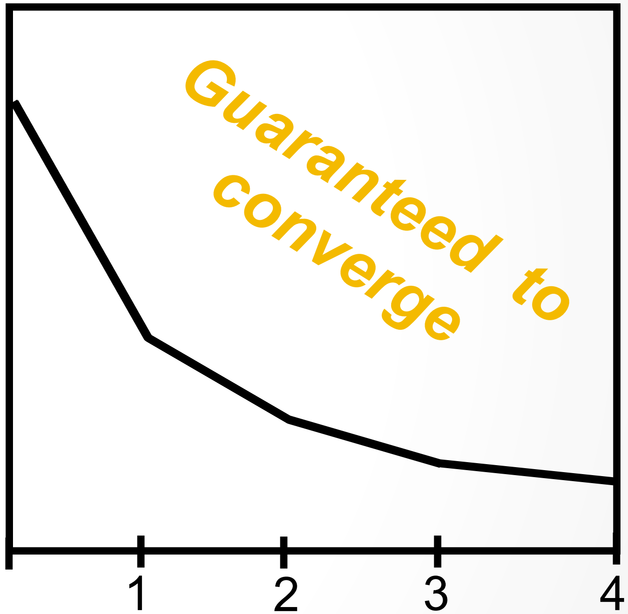
Graph cuts to infer the foreground

Grabcut: Iterative approach

- Initialize
  - Background with rectangle boundary pixels
  - Foreground with the interior of rectangle
- Iterate until convergence
  - Compute color probabilities (GMM) of each region
  - Perform graphcut segmentation
- Apply matting at boundary
- Potentially, user edits to correct mistakes

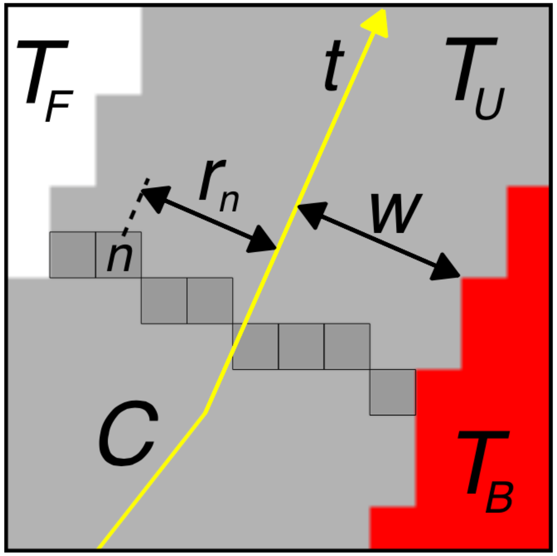
Iterated Graph Cuts

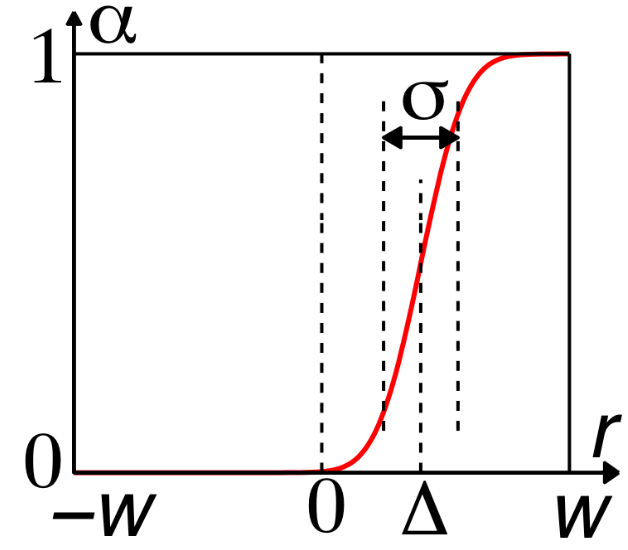
Result



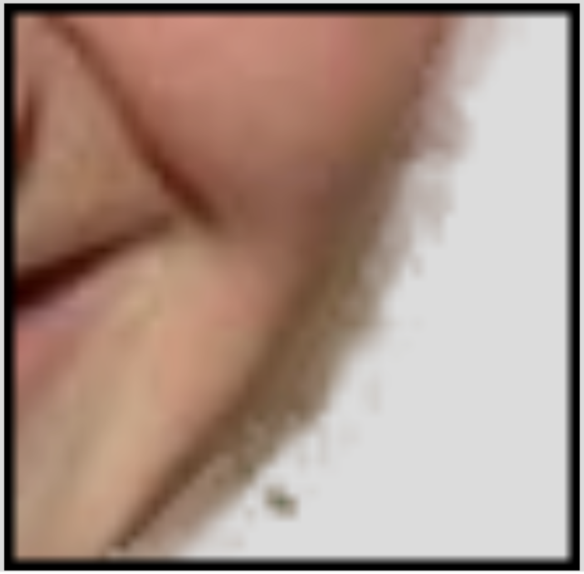
Energy after each Iteration

Border Matting

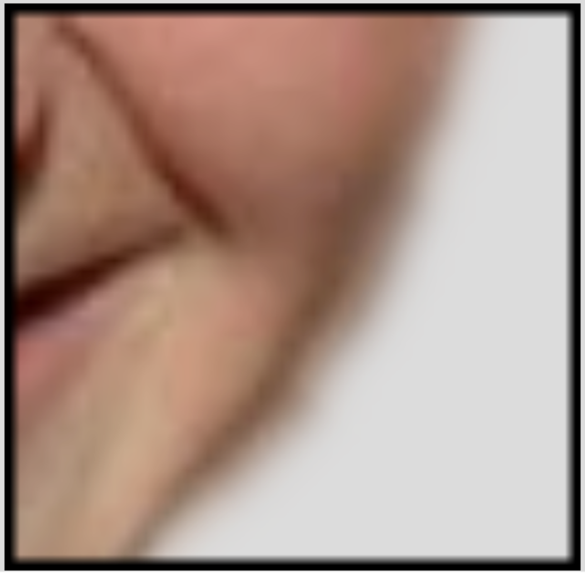




Knockout 2



Bayes Matte



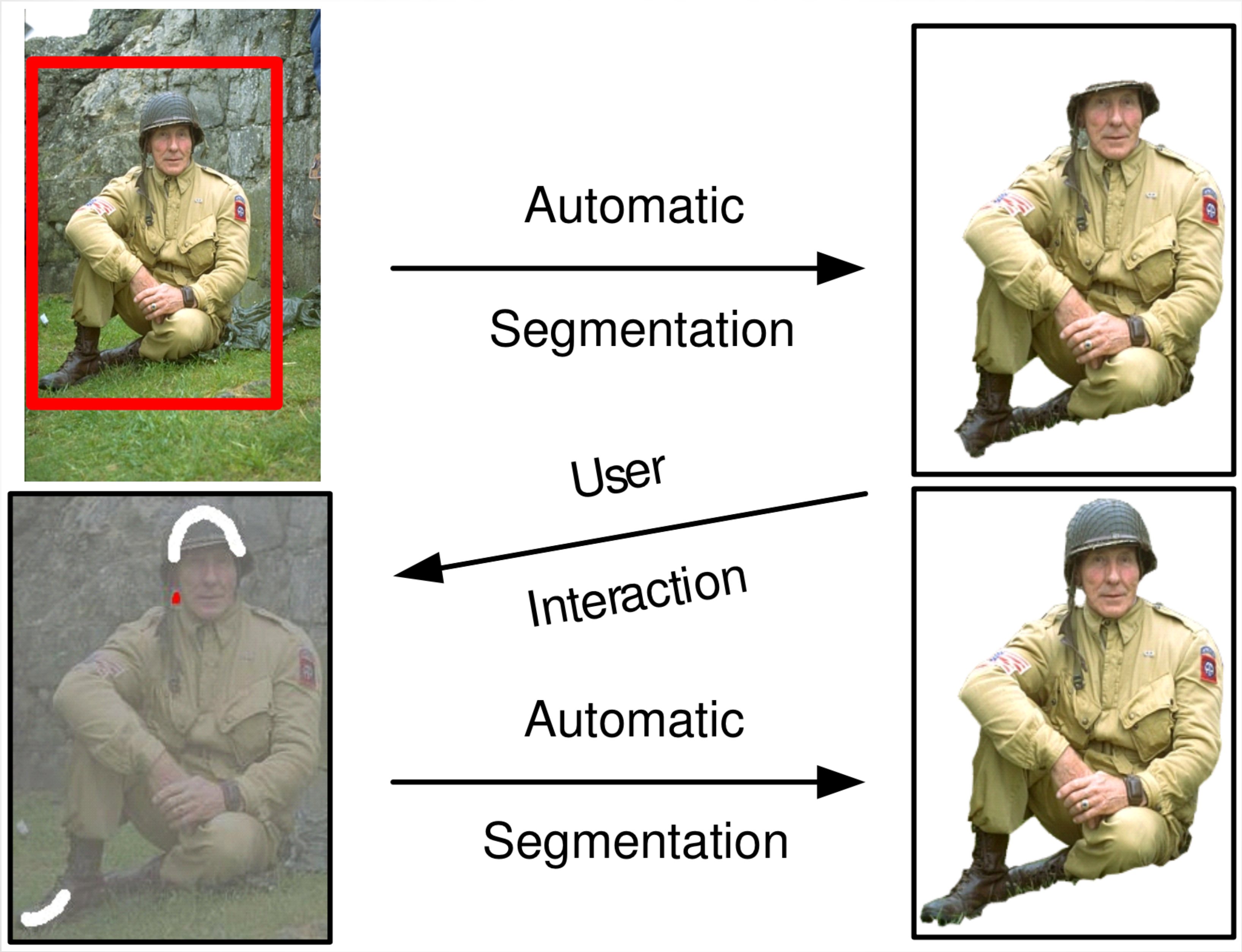
GrabCut

Border Matting Results

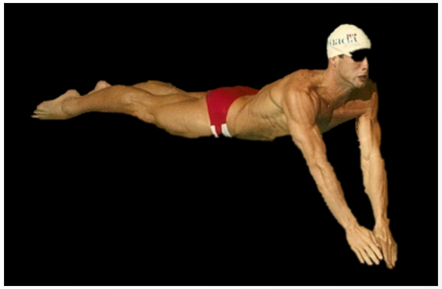
Examples: Moderate Complexity

… GrabCut completes automatically

User Interaction



Examples: Difficult Ones

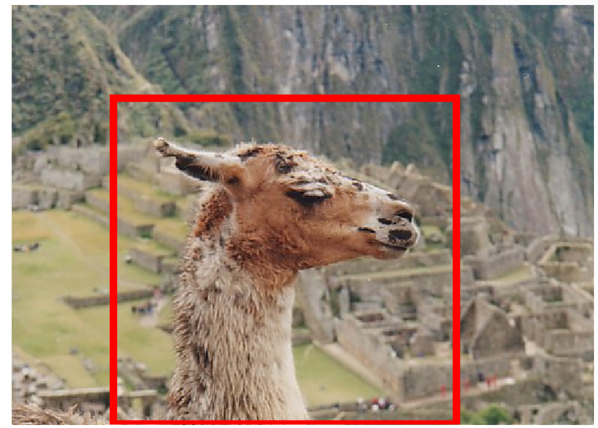
	Camouflage & Low Contrast	Fine structure	No telepathy
Initial Rectangle			
Initial Result			

GraphCut vs. GrabCut

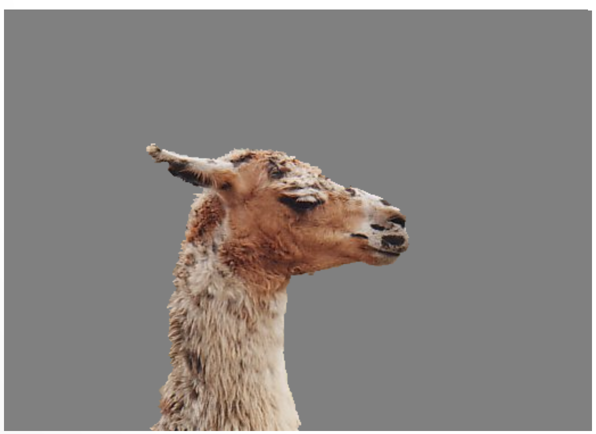
Boykov and Jolly (2001)

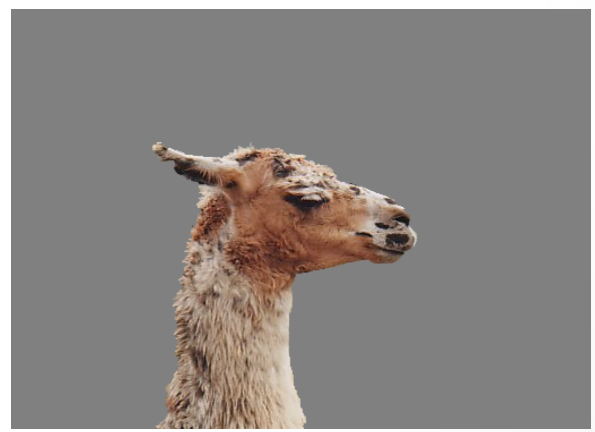
GrabCut

User Input



Result





Error Rate: 0.72%

Error Rate: 0.72%

Usage of MRF: An Example

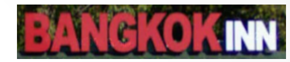
Binarizing Natural Scene Text

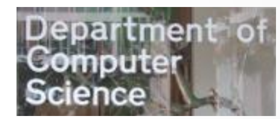
Anand Mishra, Karteek Alahari and C. V. Jawahar

"An MRF Model for Binarization of Natural Scene Text",

Proc. ICDAR, Sept. 2011, Beijing, China

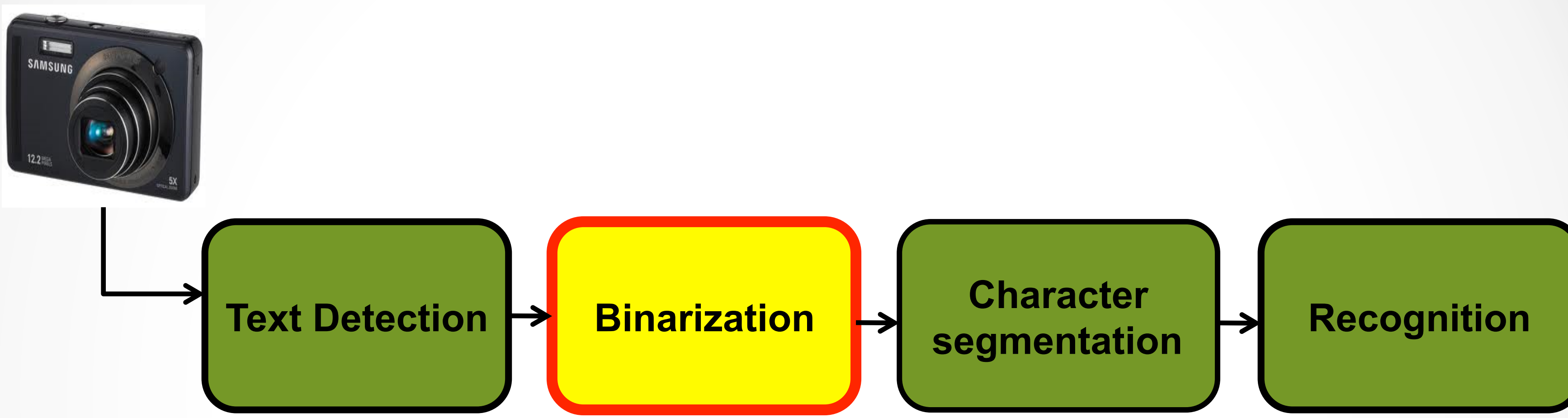
Natural Scene Text

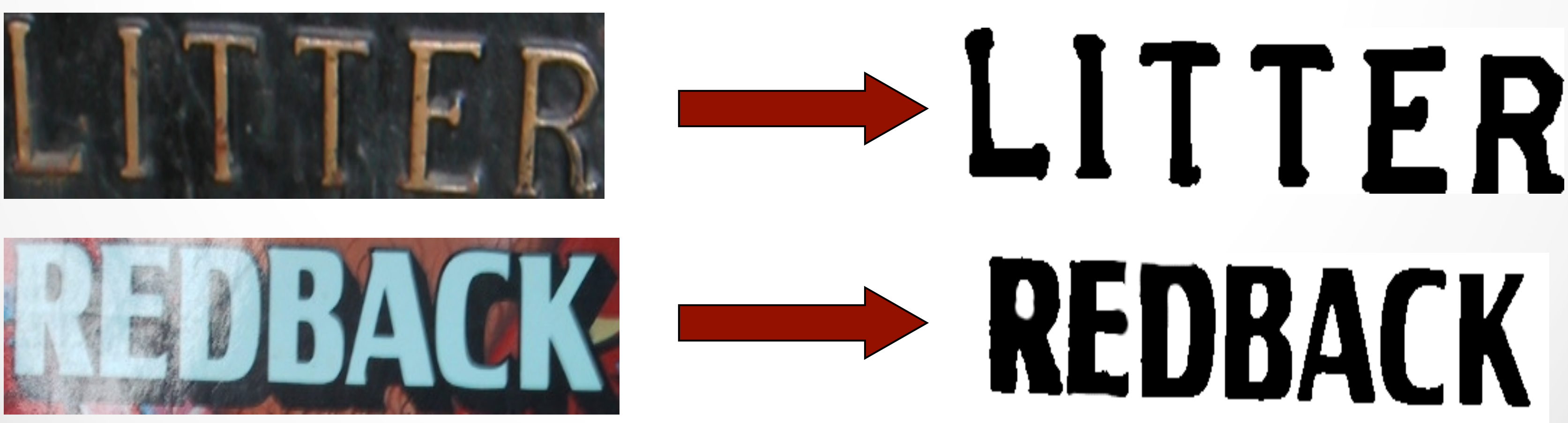




- Challenging ICDAR 2003 and Street View Text (SVT) datasets

Natural Scene Text Recognition





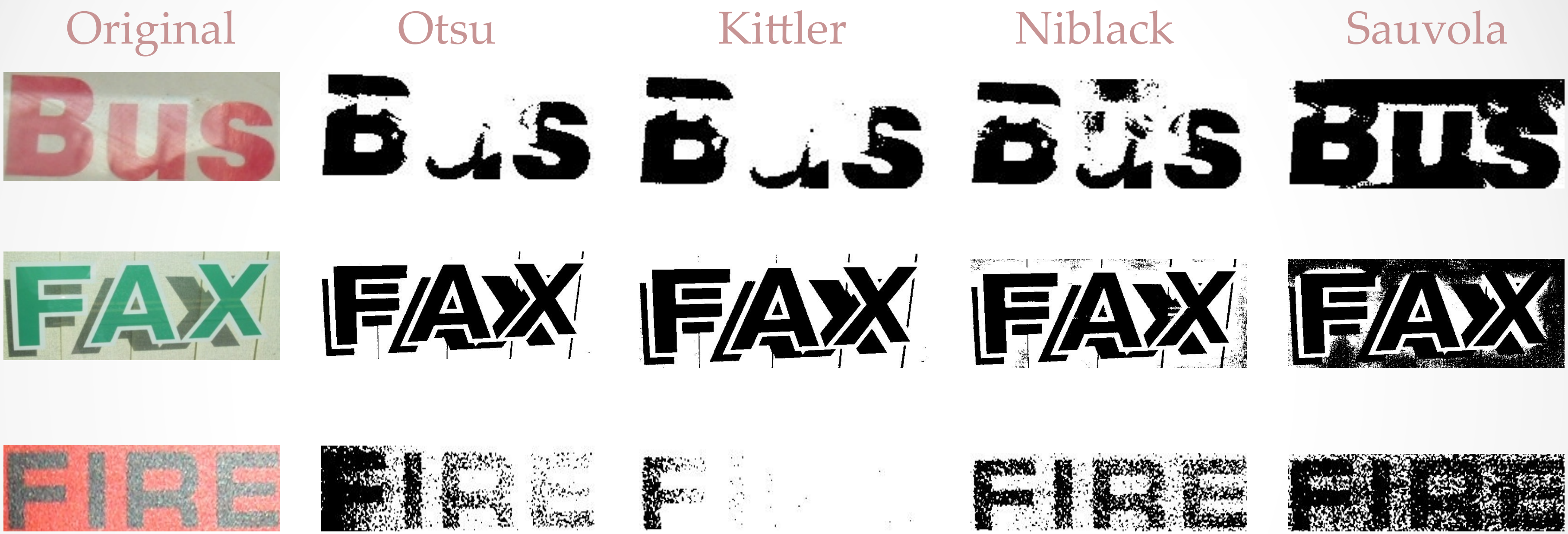
Scene Text Binarization: Challenges

- Reflections/ Transparency
- Illumination/Shadows

- Specularity
- Low Contrast
- Complex BG

- Noise
- Blur

Failure of Existing Methods



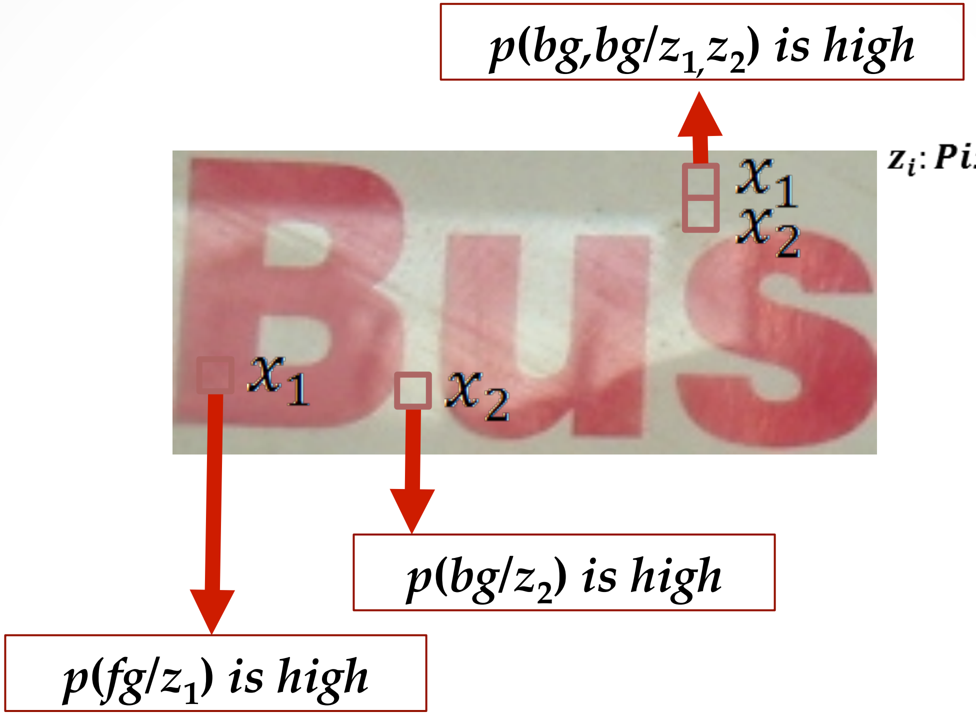
- ICDAR 2003 Robust Word Recognition Competition
  - A.Mishra et al. (ICDAR 2011): State-of-the-art
  - Y. Zhou et al. (ICDAR 2013): +2%, spotting using a small lexicon

An MRF based Binarization

Assign a label to each pixel from L = {Text (0), Background(1)}

Many labelings possible, we are interested in "the optimal" one

MRF Formulation: Potentials



z_i: *Pixel colour at pixel position* i
fg: *foreground (text)*
bg: *background*

Minimize

$$E(x) = -\sum_i \log p(x_i|z_i) + \lambda_1 \sum_{i,j \in N} \exp\left(-\beta \, ||z_i - z_j||^2\right)$$

Unary (data) Term ***Pair wise (smoothness) Term***

MRF Formulation: Gradient Potential

Gradient magnitude at pixel position i

Pair wise term $= \lambda_1 \sum_{i,j \in N} \exp\left(-\beta \, ||z_i - z_j||^2\right) + \lambda_2 \sum_{i,j \in N} \exp\left(-\beta \, ||w_i - w_j||^2\right)$

Edginess Term

An MRF based Binarization

The problem is to minimize following energy (MRF energy):

$$E(x) = Unary\ term + Pairwise\ term$$

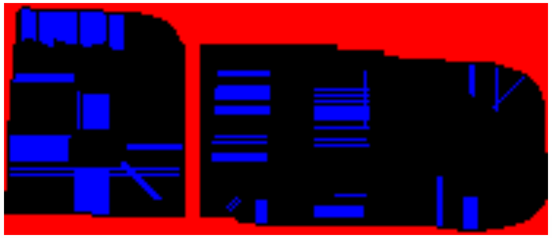
Two questions:

1)How to learn the probabilities $p(x_i|z_i)$ used to compute the unary term?
2)How to find the minima of above energy?

Learning Probabilities

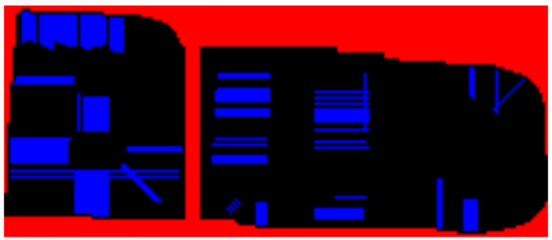
Canny Edge operator

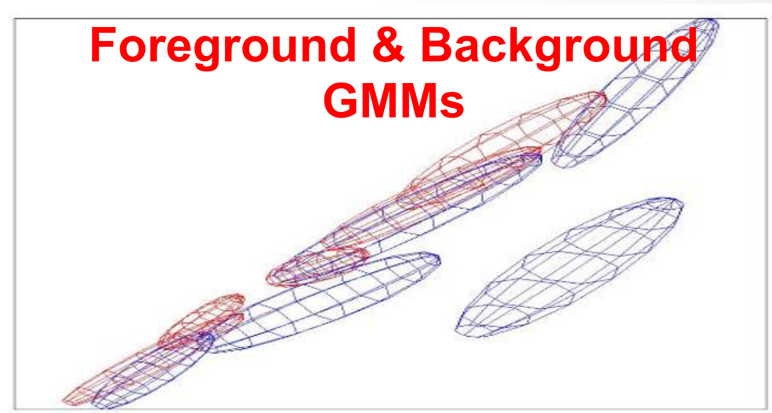
Find foreground - background seeds



Blue colour: Foreground
Red colour: Background

Color Modeling through GMMs





Unary term is calculated based on the probability of a pixel colour belonging to one of the GMM components

GrabCut (Carsten Rother et al., SIGGRAPH 2004) uses similar modelling for object segmentation problems

An MRF based Binarization

The problem is to minimize following energy (MRF energy):

$$E(x) = Unary\,term + Pair\,wise\,term$$

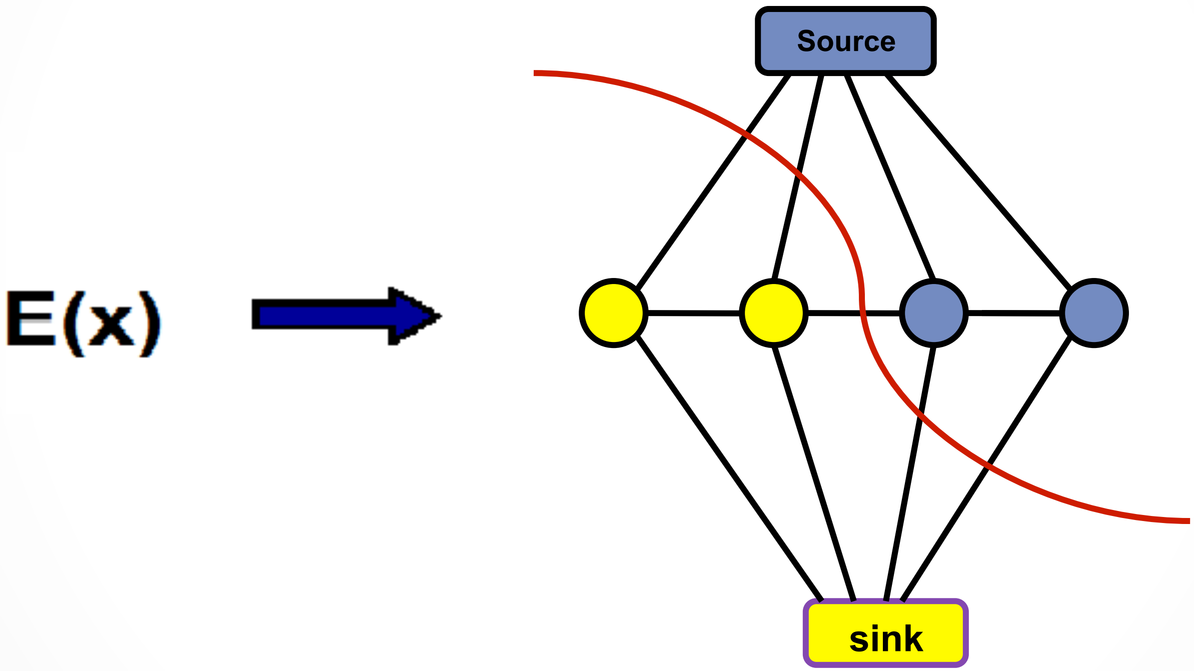
Two questions:

1) **How to learn the probabilities** $p(x_i|z_i)$ **used to compute the unary term?**

2) **How to find the minima of above energy?**

Graph Cut

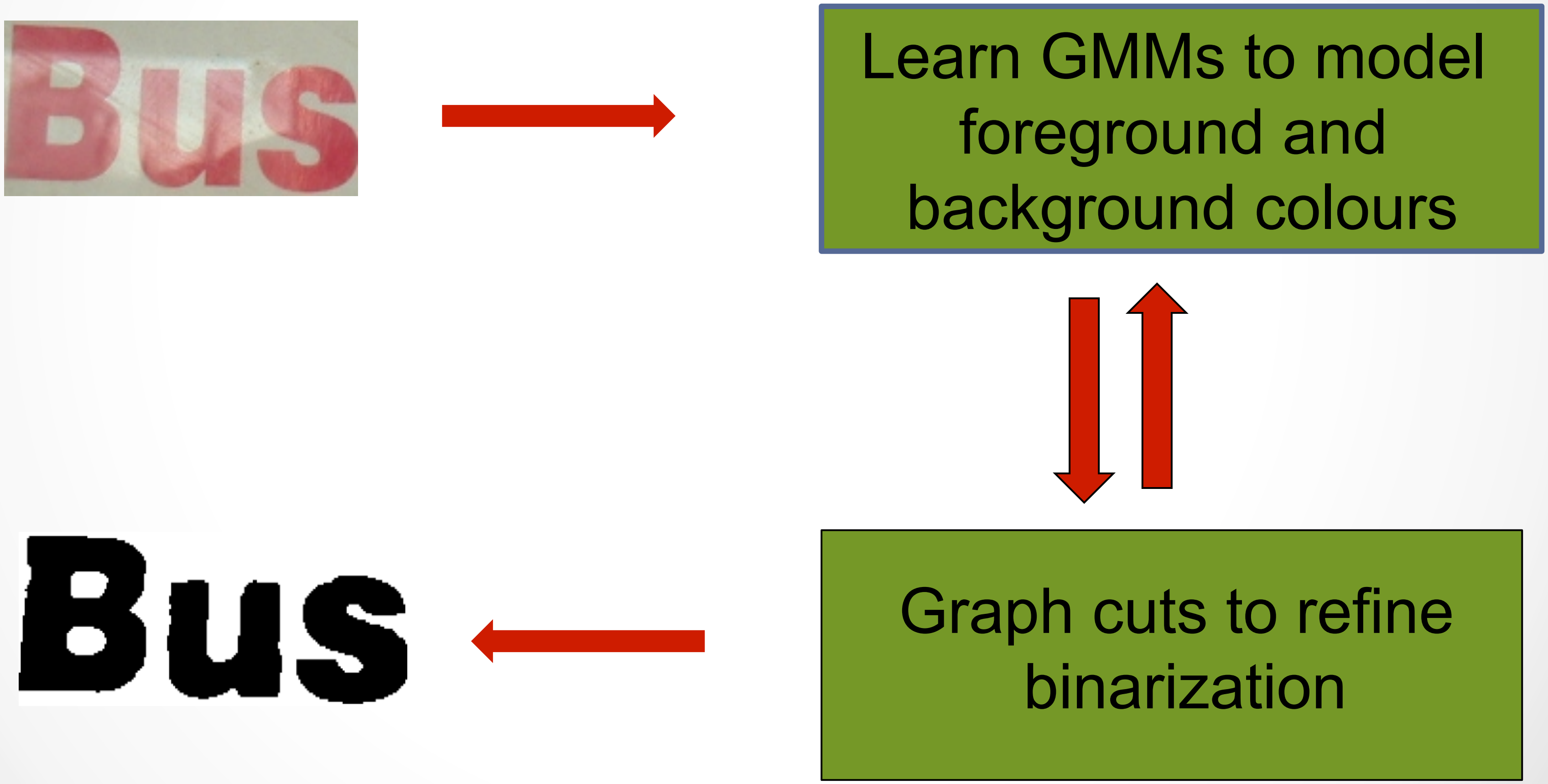
Minimum of MRF energy = min cut of graph



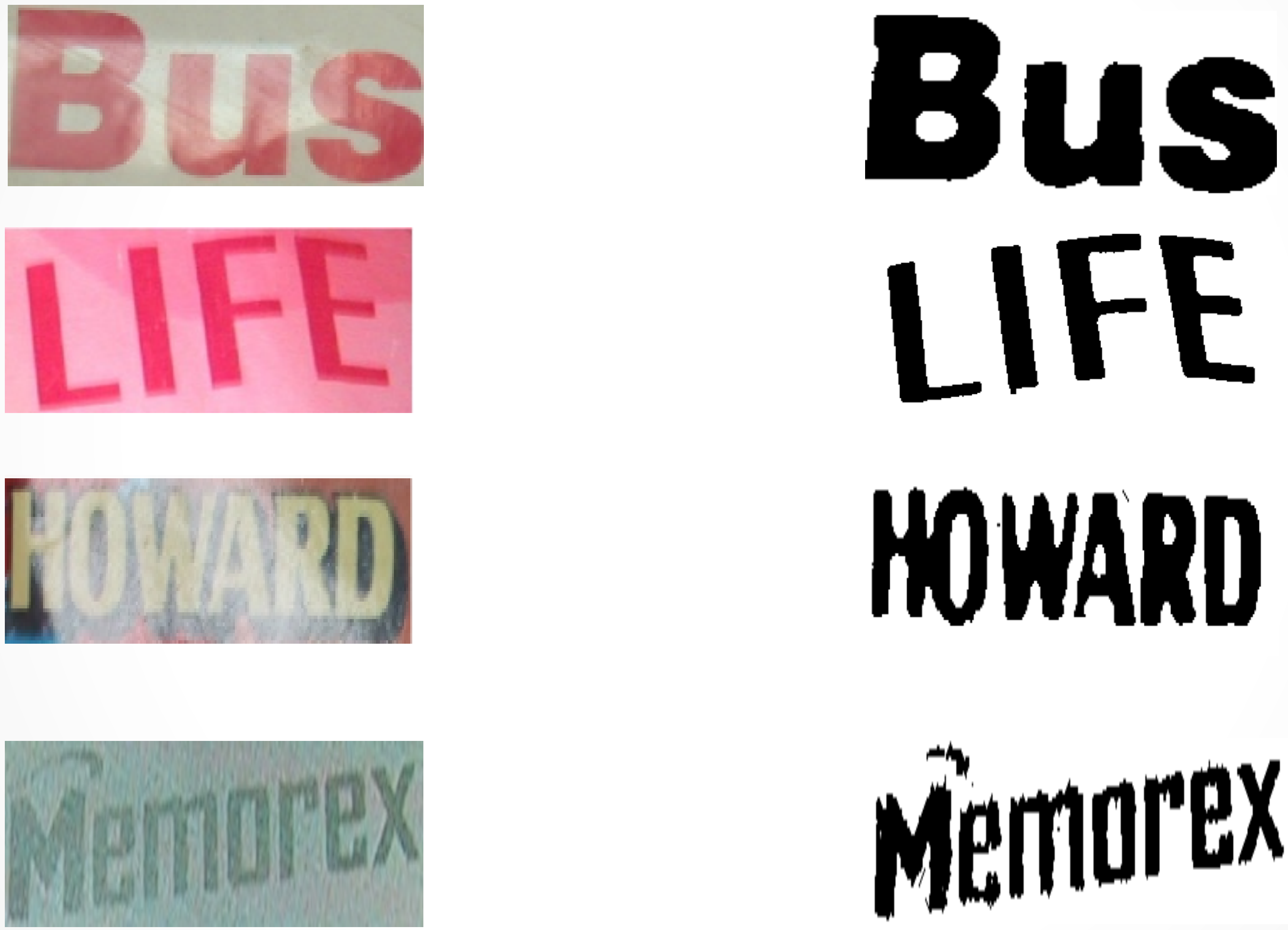
Efficient codes available to compute min cut of such graph

Vladimir Kolmogorov and Ramin Zabih, "What Energy Functions can be minimized via Graph Cut", PAMI 2004

An Iterative Graph Cut based Approach



Qualitative Results

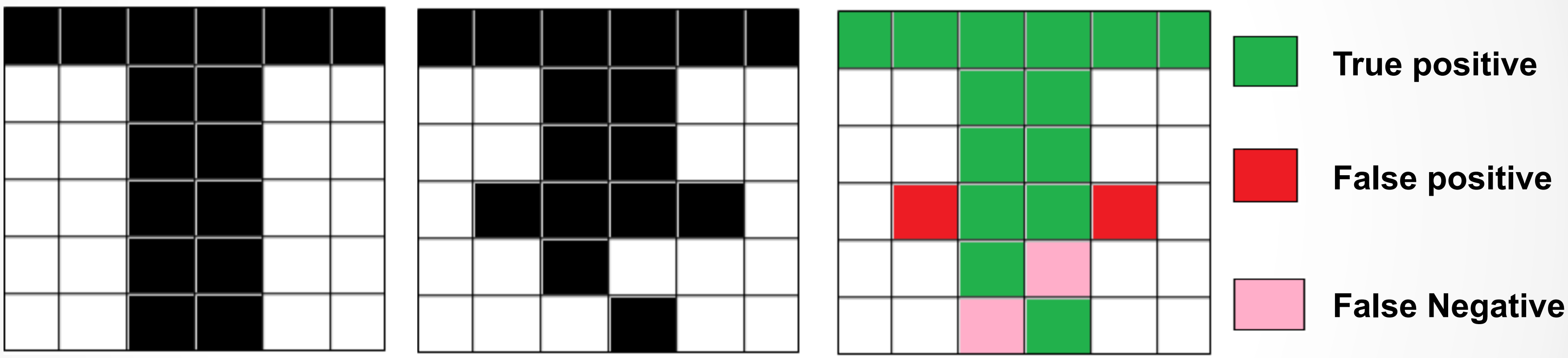


Qualitative Results

Quantitative Results

- OCR accuracy
- Pixel level accuracy

ABBYY



Ground truth　　Binarization result　　Evaluation result

$$precision = \frac{Total\ number\ of\ green\ boxes}{Total\ number\ of\ green\ boxes + Total\ number\ of\ red\ boxes}$$

$$recall = \frac{Total\ number\ of\ green\ boxes}{Total\ number\ of\ green\ boxes + Total\ number\ of\ pink\ boxes}$$

$$f - score = \frac{2\ \times precison\ \times recall}{precision + recall} \times 100$$

Results (ABBYY OCR Accuracy)

Method	Word Accuracy (%)	Character Accuracy (%)
Otsu	41.52	51.74
Sauvola	39.77	51.63
Niblack	39.18	42.31
Kittler	41.12	49.88
Otsu + CT	45.03	51.98
MRF (without edginess diff.)	49.12	55.94
MRF (with edginess diff.)	**52.04**	**60.14**

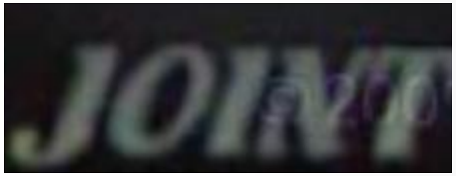
Results (Pixel Level Accuracy)

Method	f-score (%)
Otsu	79.32
Sauvola	73.87
Niblack	76.86
Kittler	72.89
Otsu + CT	78.12
MRF (without edginess diff.)	87.84
MRF (with edginess diff.)	**88.64**

More Results

Results based on Street View Text Dataset

Method	**Word Recognition accuracy (%)**
ABBYY	32.61%
MRF Binarization + ABBYY	**42.81%**



Kai Wang and Serge Belongie (ECCV 2010) have introduced a challenging Street View Text (SVT) dataset

When can it Fail?

- Colour may not characterize the chatacter.
- Failing to learn text-BG probabilities

Reading Material

- Required
  - "Normalized Cuts and Image Segmentation", by Jianbo Shi and Jitendra Malik, TPAMI, Aug 2000
  - "Grabcut: Interactive Foreground Extraction using Iterated Graph Cuts", by Rother, Kolmogorov and Blake, Siggraph 2004.
  - "An MRF Model for Binarization of Natural Scene Text", by Anand Mishra, Karteek Alahari and C.V. Jawahar, ICDAR 2011.
- Additional
  - "OBJ CUT", by M. Pawan Kumar P.H.S. Torr and A. Zisserman, CVPR 2005.