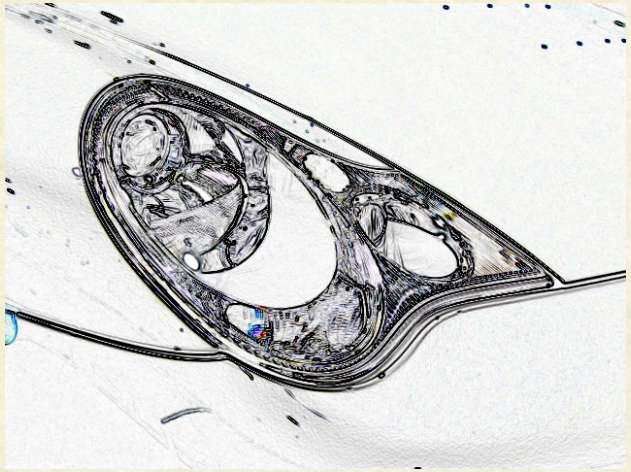
CS7.505: Computer Vision

Spring 2022: Multi-Label MRF

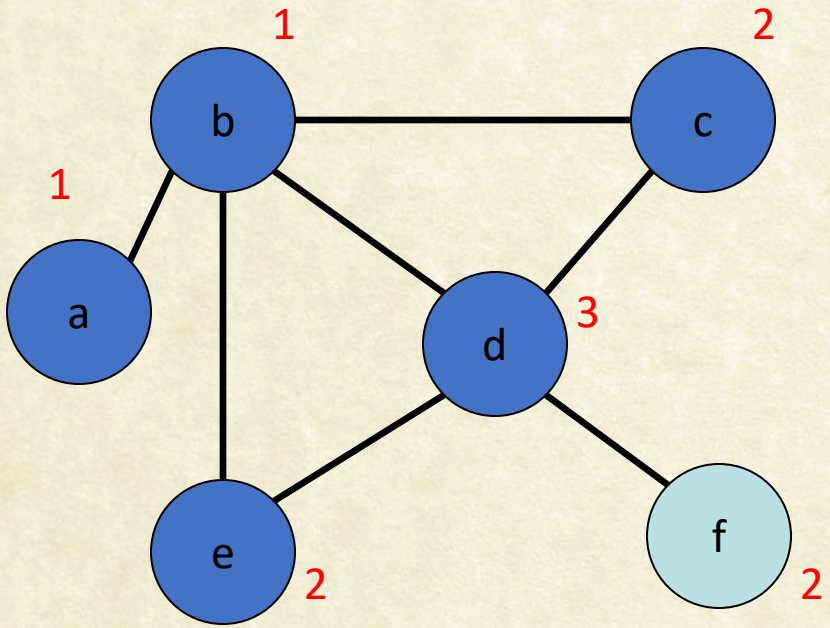




Anoop M. Namboodiri

Biometrics and Secure ID Lab, CVIT,

IIIT Hyderabad

The General Problem



Graph G = (V, E)

Discrete label set L = {1,2,…,h}

Assign a label to each vertex

f: V → L

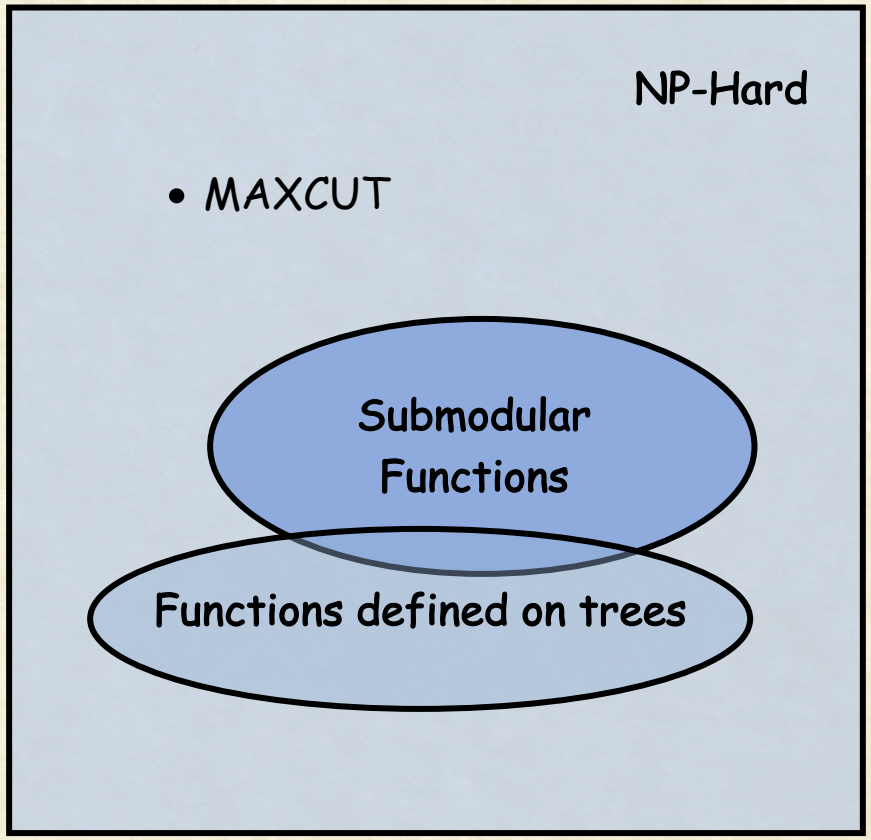
Cost of a labelling Q(f)

Unary Cost Pairwise Cost

Find f* = arg min Q(f)

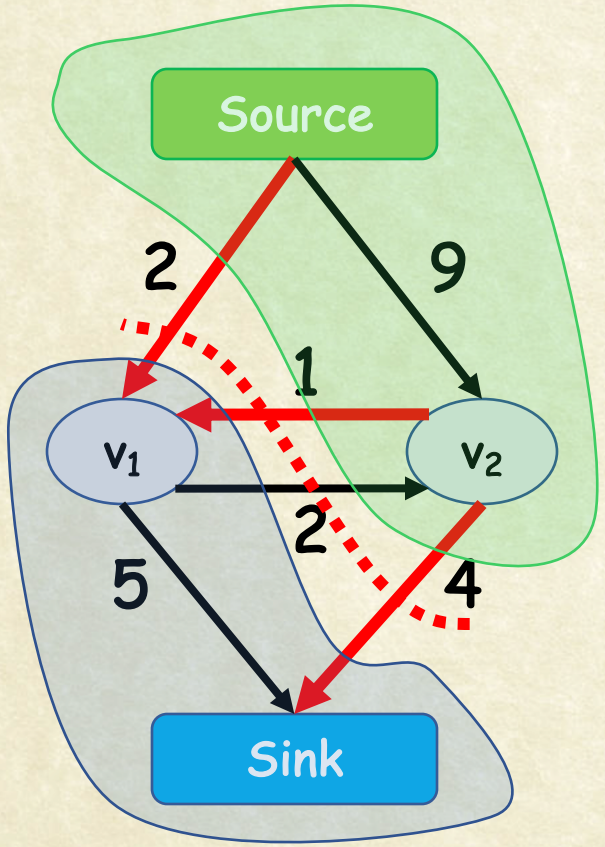
Minimizing Energy Functions

- **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

The st-Mincut Problem



2 + 1 + 4 = 7

What is an st-cut?

An st-cut (S,T) divides the nodes between source and sink.

What is the cost of an st-cut?

Sum of cost of all edges going from S to T

What is the st-mincut?

st-cut with the minimum cost

Robust Interactions

- NP-hard problem (3 or more labels)
  - two labels can be solved via *s-t* cuts (Greig at. al., 1989)
- *α-expansion* approximation algorithm

  (Boykov, Veksler, Zabih 1998, 2001)
  - Guaranteed approximation quality (Veksler, thesis 2001)
    - within a factor of 2 from the global minima (Potts model)
  - Applies to a wide class of energies with robust interactions
    - Potts model (BVZ 1998)
    - "Metric" interactions (BVZ 2001)
    - Can be extended to arbitrary interactions with weaker guarantees
      - truncation (Kolmogorov et al. 2005)
      - QPBO (Boros and Hummer, 2002)
- Other "move" algorithms (e.g. α–β swap, jump-moves)

Metric and Semimetric Interactions

$$E(f) = \sum_{\{p,q\} \in \mathcal{N}} V_{p,q}(f_p, f_q) + \sum_{p \in \mathcal{P}} D_p(f_p)$$

$$\begin{aligned} V(\alpha, \beta) = 0 \quad &\Leftrightarrow \quad \alpha = \beta, \\ V(\alpha, \beta) \quad &= \quad V(\beta, \alpha) \geq 0, \\ V(\alpha, \beta) \quad &\leq \quad V(\alpha, \gamma) + V(\gamma, \beta), \end{aligned}$$

- where α,β, and γ are labels.
- Metric: If all three conditions are satisfied.
- Semi-metric: If only the first two are satisfied.
- Examples?

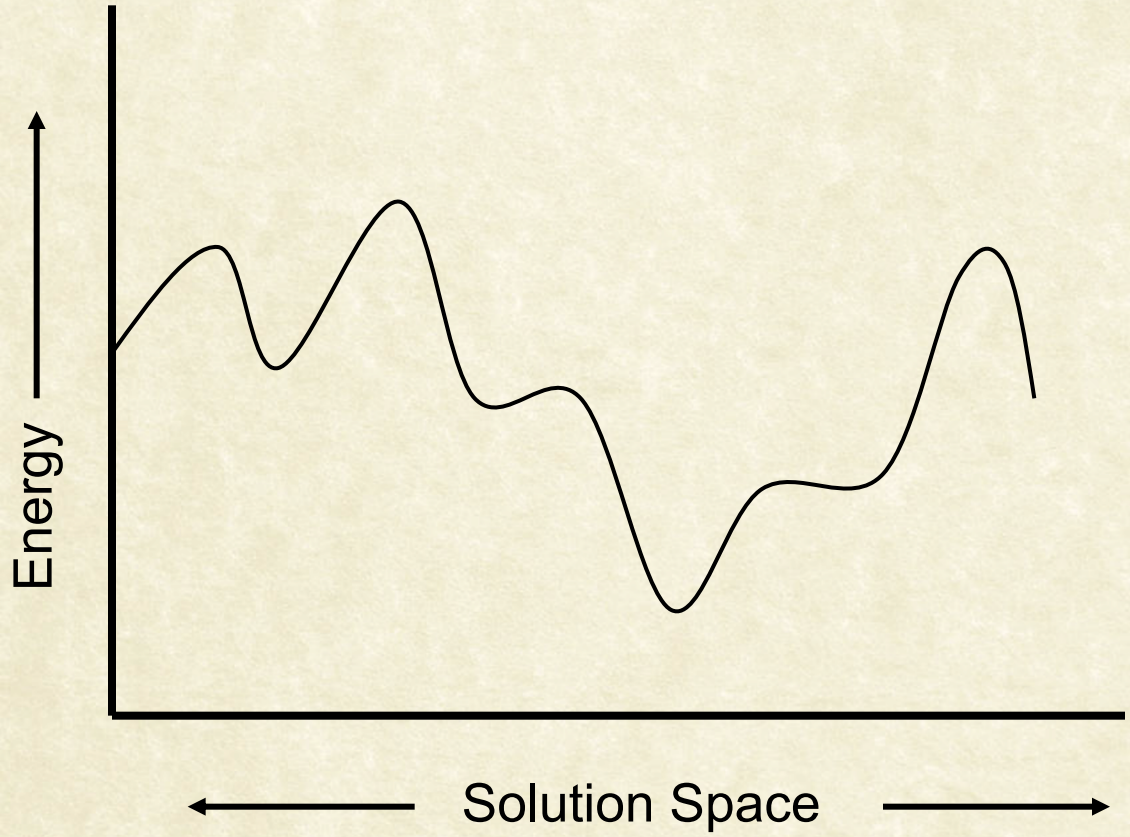
Multi-label Graphcut: Move Making

$$E(x) = \sum_i \theta_i(x_i) + \sum_{i,j} \theta_{ij}(x_i, x_j)$$

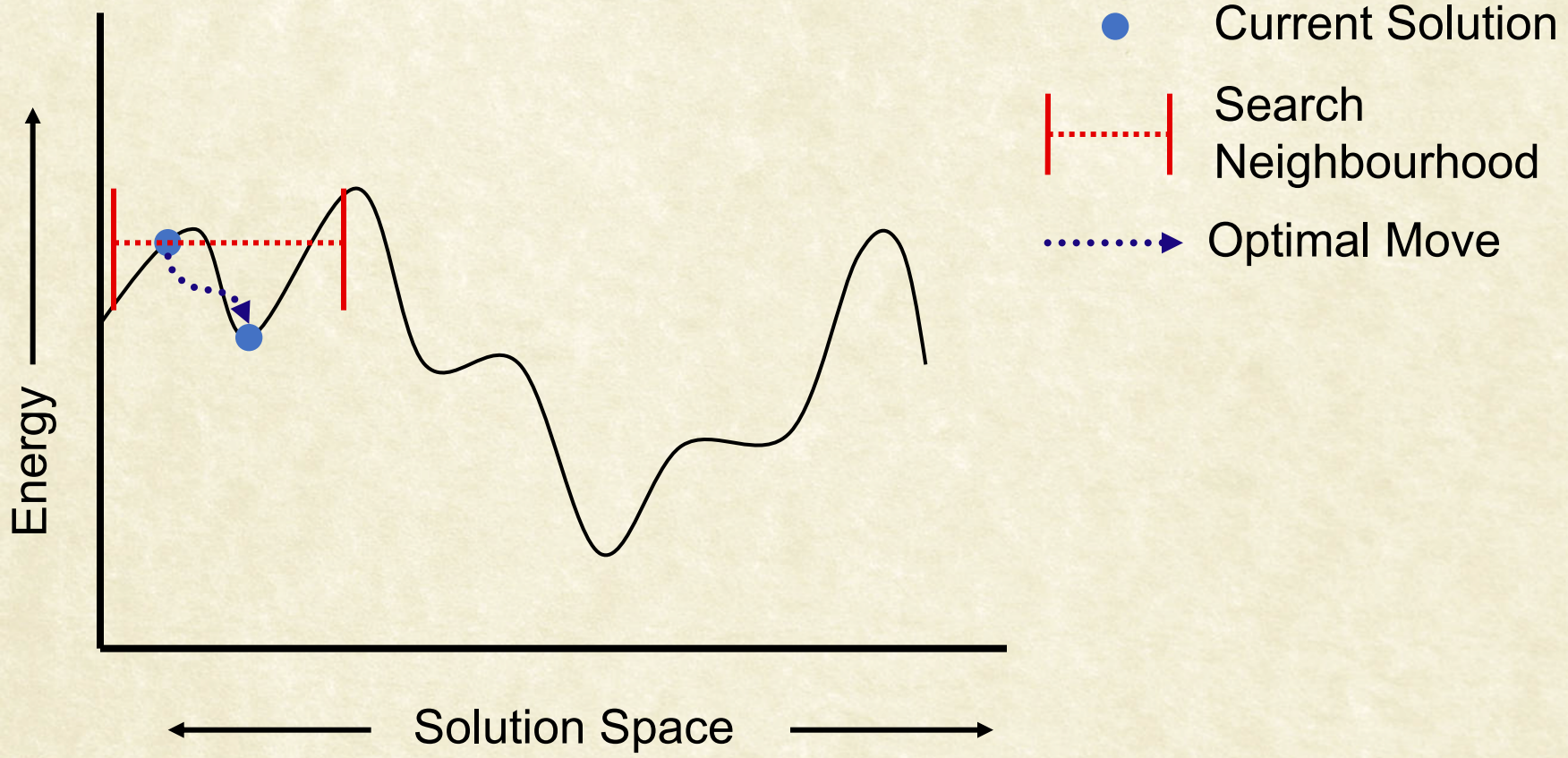
$x \in$ Labels $L = \{l_1, l_2, \ldots, l_k\}$

- Commonly used for solving **non-submodular** multi-label problems
- Extremely efficient and produce good solutions
- Not Exact: Produce local optima

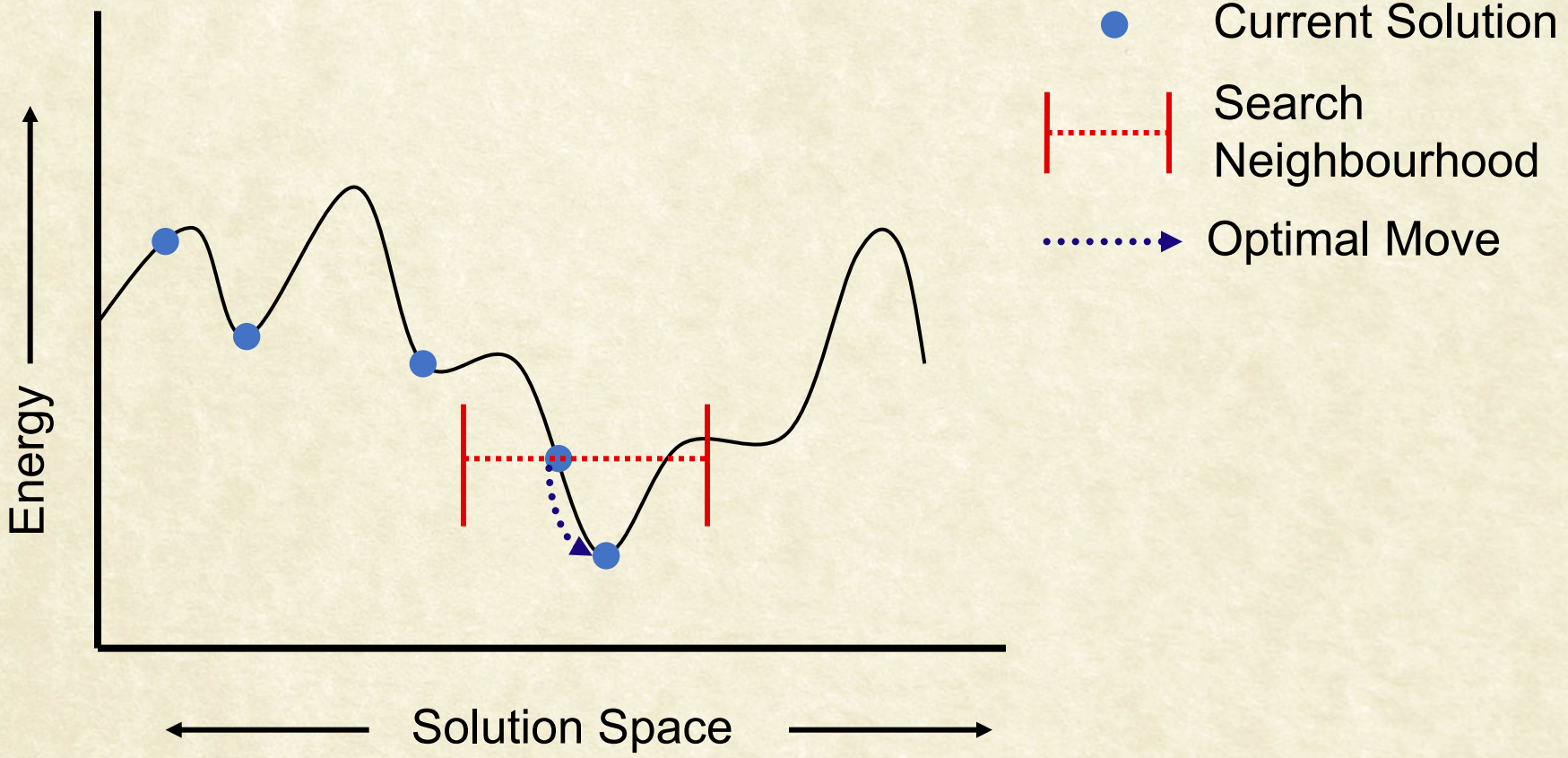
Move Making Algorithms



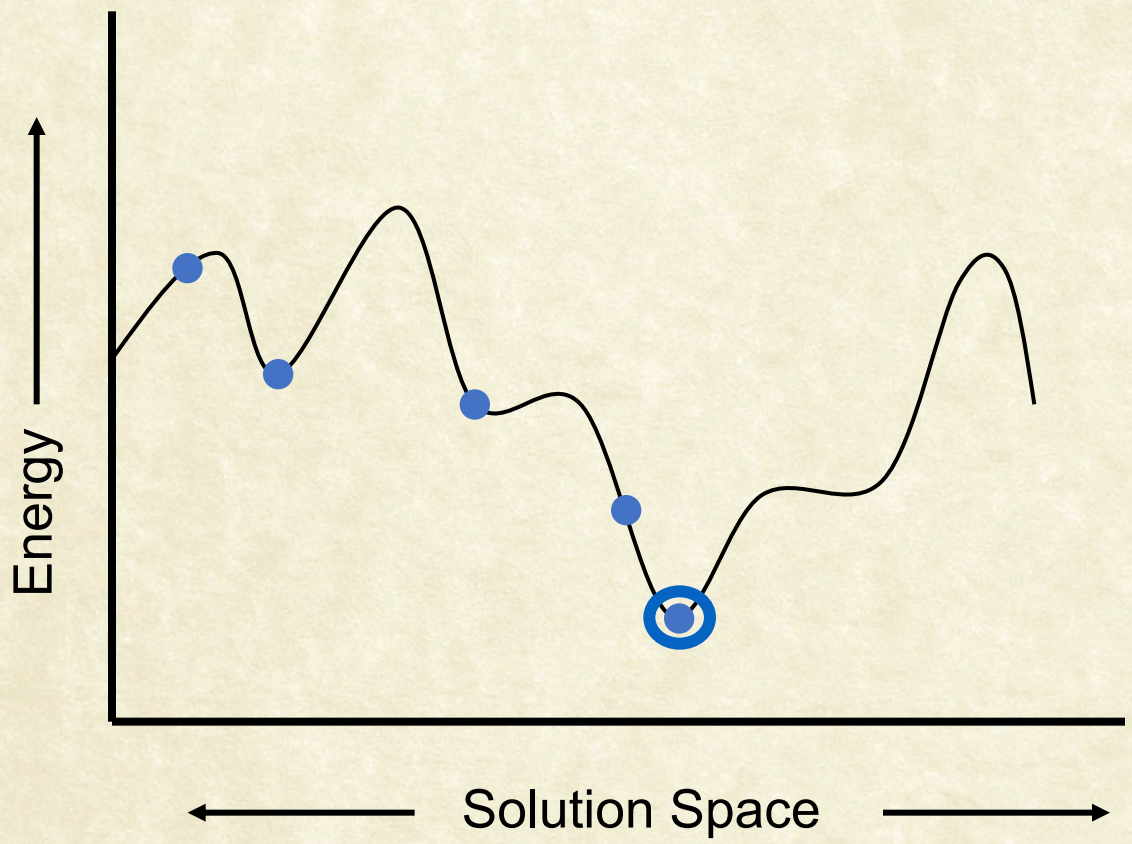
Move Making Algorithms



Move Making Algorithms



Move Making Algorithms

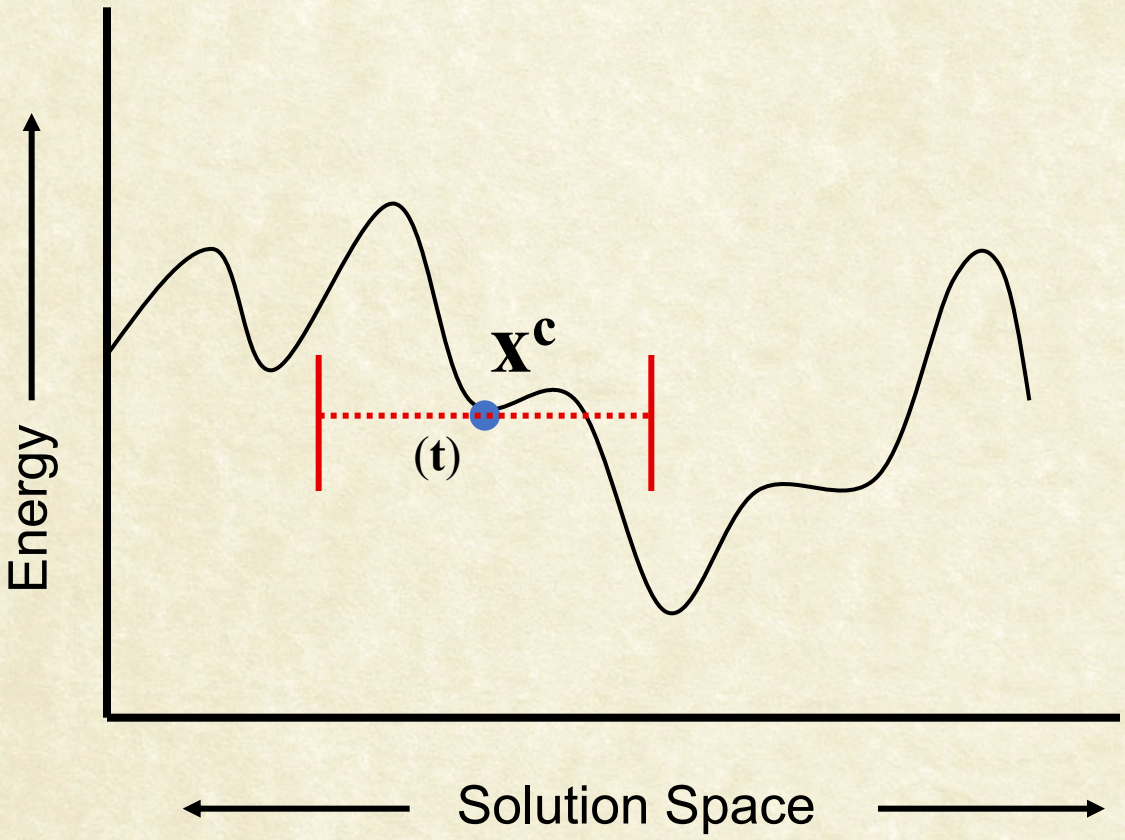


Current Solution

Search Neighbourhood

Optimal Move

Computing the Optimal Move



Current Solution

Search Neighbourhood

Optimal Move

Key Idea

Move Space

Bigger move space → • Better solutions
• Finding the optimal move hard

What are we doing?

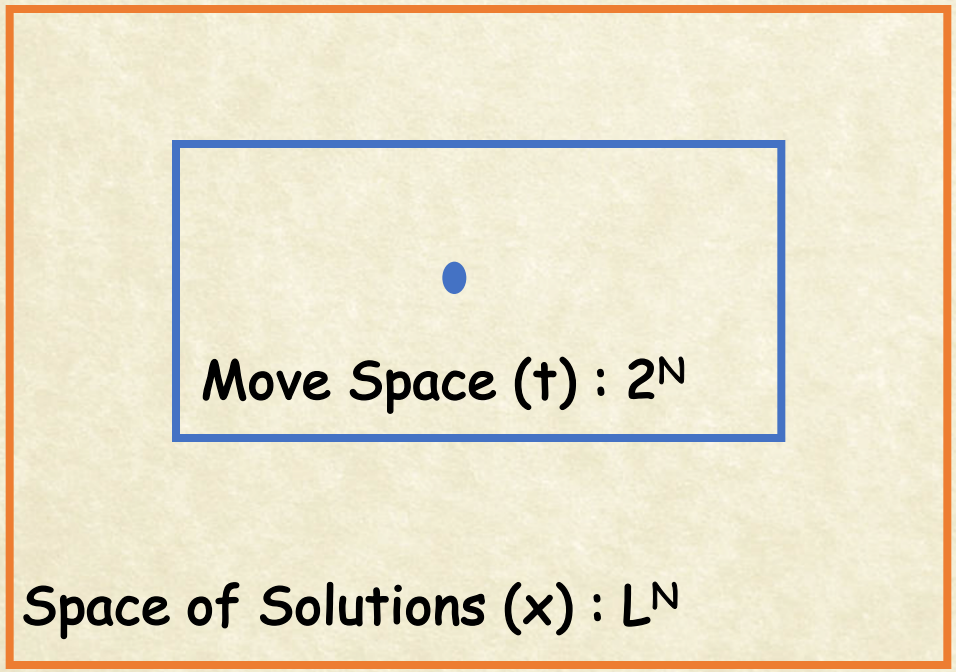
- Cast segmentation as a generic optimization prob. on a graph
- The graph optimization problem is generic:
  - Assign any one of k labels $\mathrm{L} = \{l_1, l_2, \ldots, l_k\}$ to each vertex of a graph
  - Cost of assigning labels in terms of unary and pair-wise costs
  - Derive algorithms that minimize the total cost
- The labels need not be object membership
  - Depth values, Class labels, Instance labels, Motion
- Unary potential need not be based on image pixel values
  - Stereo disparity

Moves using Graph Cuts

Expansion and Swap move algorithms

[Boykov, Veksler and Zabih, PAMI 2001]

- **Makes a series of changes to the solution (moves)**
- **Each move results in a solution with smaller energy**



Current Solution

Search Neighbourhood

N Number of Variables

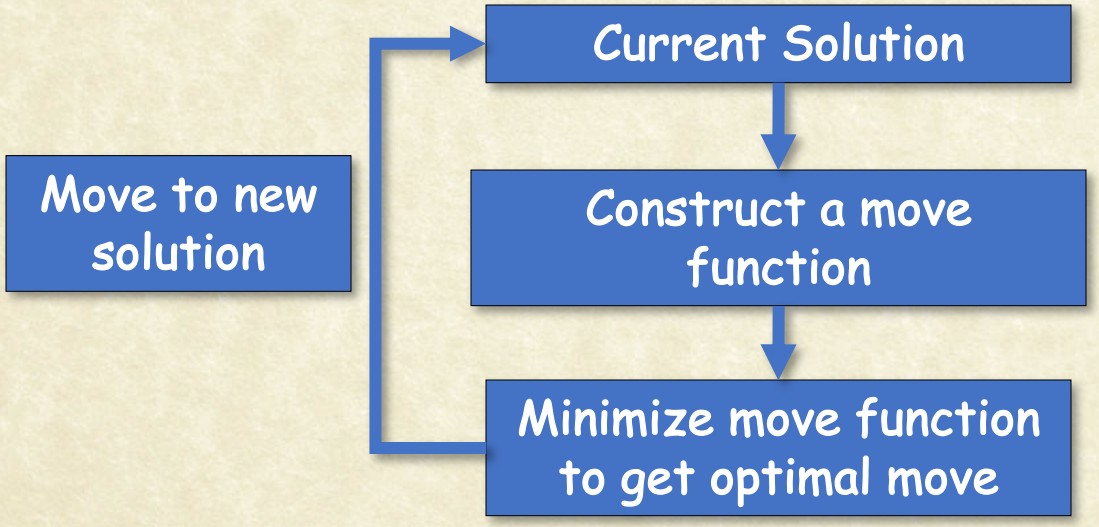
L Number of Labels

Moves using Graph Cuts

Expansion and Swap move algorithms

[Boykov, Veksler and Zabih, PAMI 2001]

- Makes a series of changes to the solution (moves)
- Each move results in a solution with smaller energy



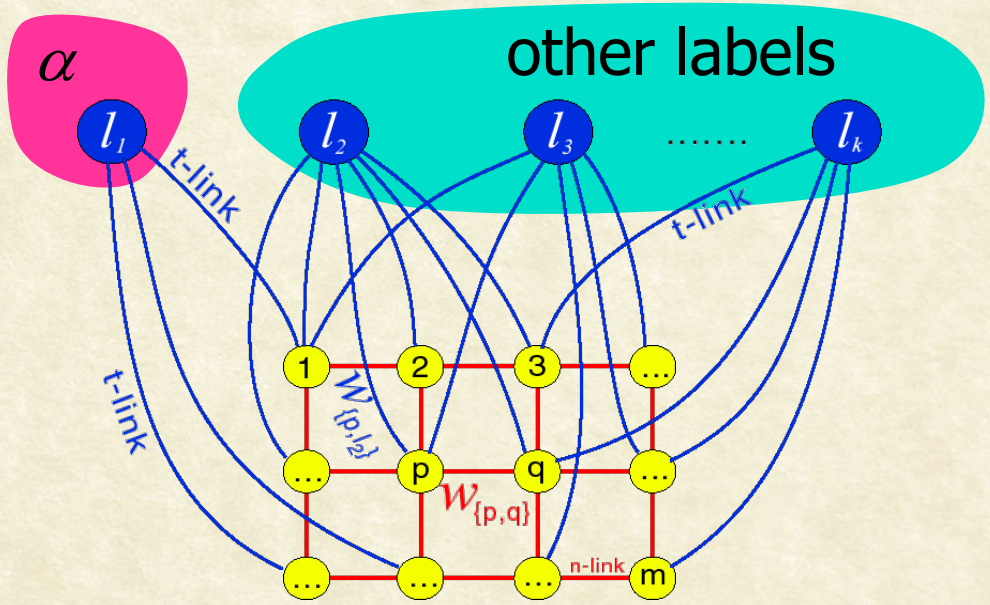
How to minimize move functions?

α-expansion algorithm

1. Start with any initial solution
2. For each label "α" in any (e.g. random) order
    1. *Compute optimal α-expansion move (s-t graph cut)*
    2. *Decline the move if there is no energy decrease*
3. *Stop when no expansion move would decrease energy*

α-expansion move

- Basic Idea: Break multi-way cut computation into a **sequence of binary *s-t* cuts**



α-expansion move

In each α-expansion a given label "α" grabs space from other labels

initial solution

-expansion

-expansion

-expansion

-expansion

-expansion

-expansion

-expansion

For each move we choose expansion that gives the largest decrease in the energy: **binary optimization problem**

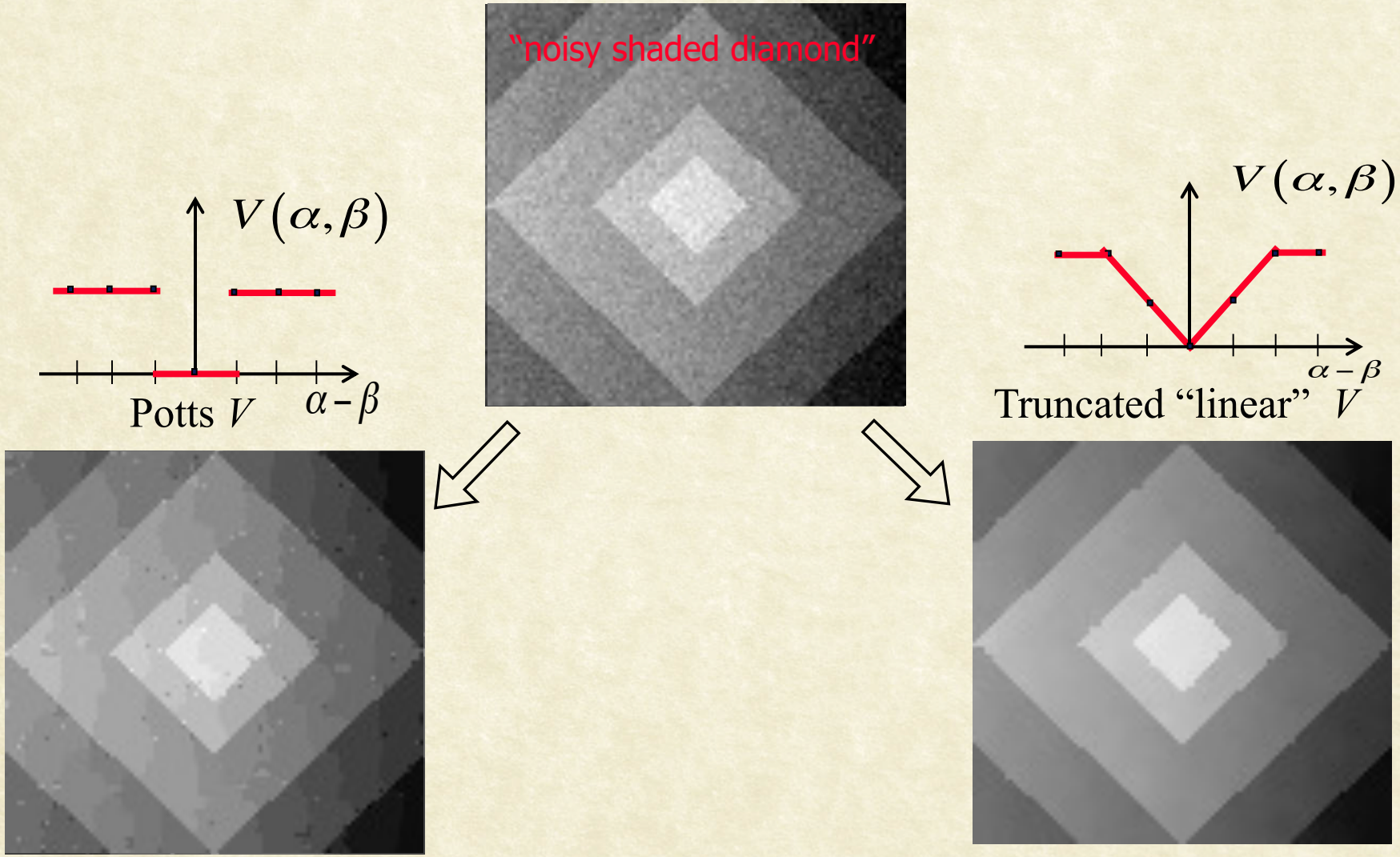
Metric and Semimetric Interactions

$$E(f) = \sum_{\{p,q\} \in \mathcal{N}} V_{p,q}(f_p, f_q) + \sum_{p \in \mathcal{P}} D_p(f_p)$$

$$\begin{aligned} V(\alpha, \beta) = 0 \quad &\Leftrightarrow \quad \alpha = \beta, \\ V(\alpha, \beta) \quad &= \quad V(\beta, \alpha) \geq 0, \\ V(\alpha, \beta) \quad &\leq \quad V(\alpha, \gamma) + V(\gamma, \beta), \end{aligned}$$

- where α,β, and γ are labels.
- Metric: If all three conditions are satisfied.
- Semimetric: If only the first two are satisfied.
- Examples?

α-expansions: Examples of *metric* interactions



Summary, Additional Reading

- Binary Labelling Problems: Exact Solution (s-t cut)
- Multi-label MRF: Efficient Approximate Solutions
  - α-expansion
  - Belief Propagation, Loopy BP, etc.
- Resources:
  - Vladimir Kolmogorov and Ramin Zabih, “What Energy Functions Can Be Minimized via Graph Cuts?”, TPAMI-2004
  - Yuri Boykov, Olga Veksler, Ramin Zabih, “Fast Approximate Energy Minimization via Graph Cuts”, ICCV 1999
  - Course on PGM: Daphne Koller, Stanford

Ink-Bleed Removal

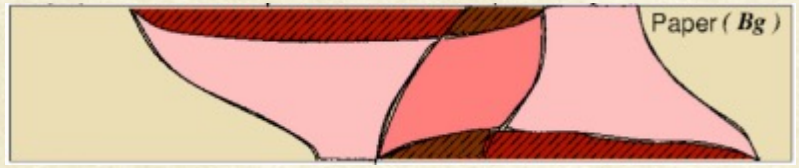
Yi Huang Michael S. Brown Dong Xu

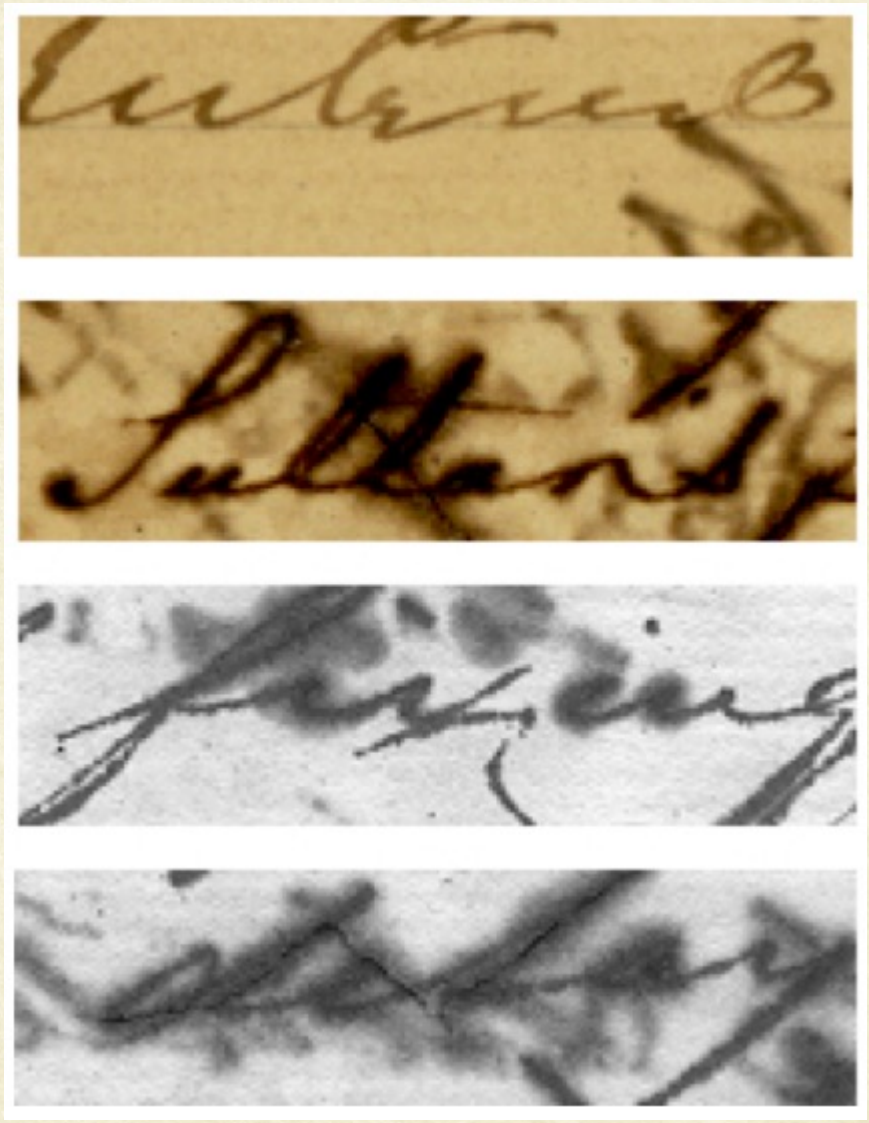
"A Framework for Reducing Ink-Bleed in Old Documents",

Proc. CVPR, June 2008, Anchorage, AK, USA.

What is Ink Bleed?

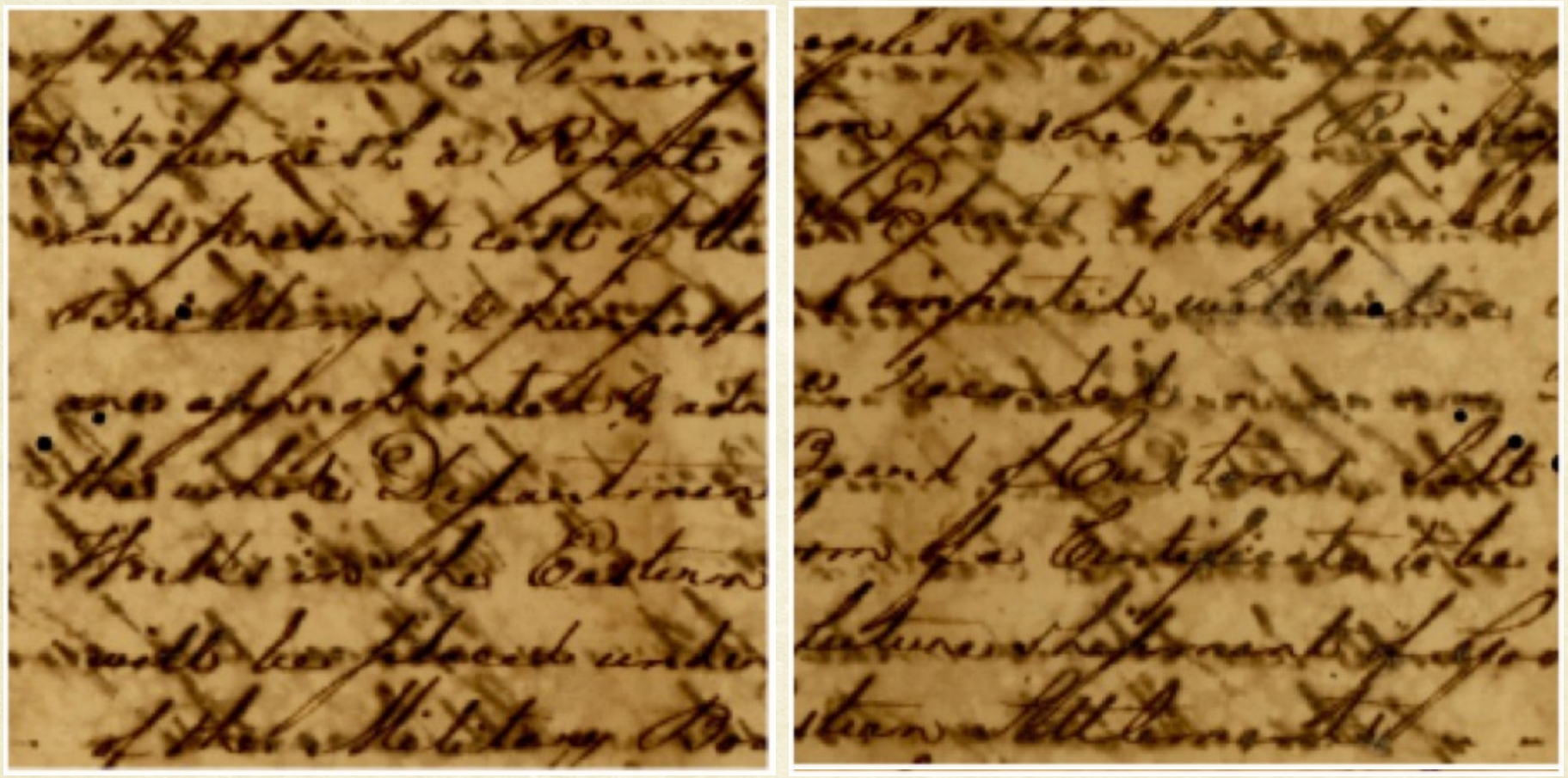
- Ink bloats through paper and appears on the reverse side





Recto and Verso Images

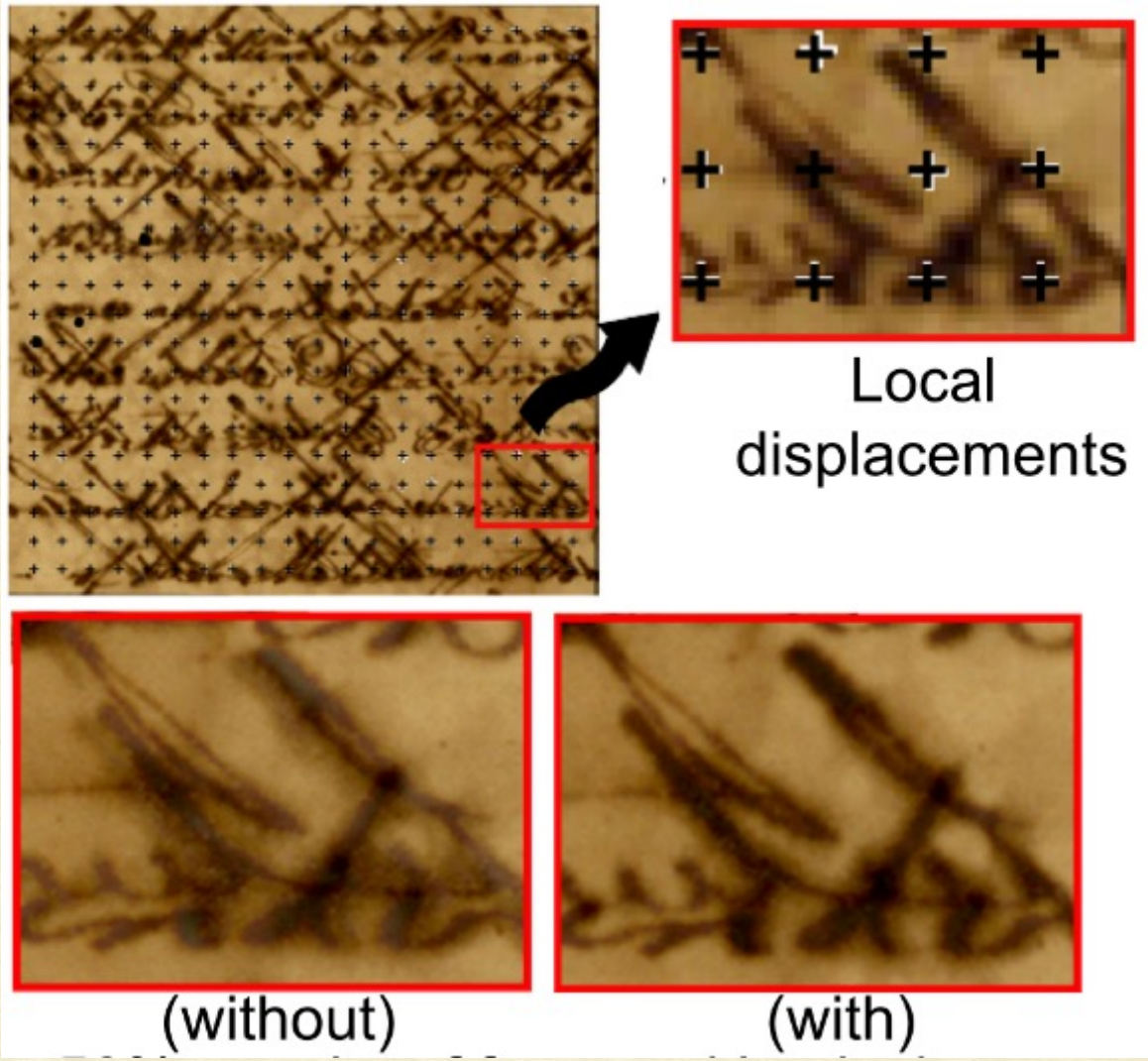
Recto (front) Side ⇔ Verso (back) Side



- Treat the two restoration problems together

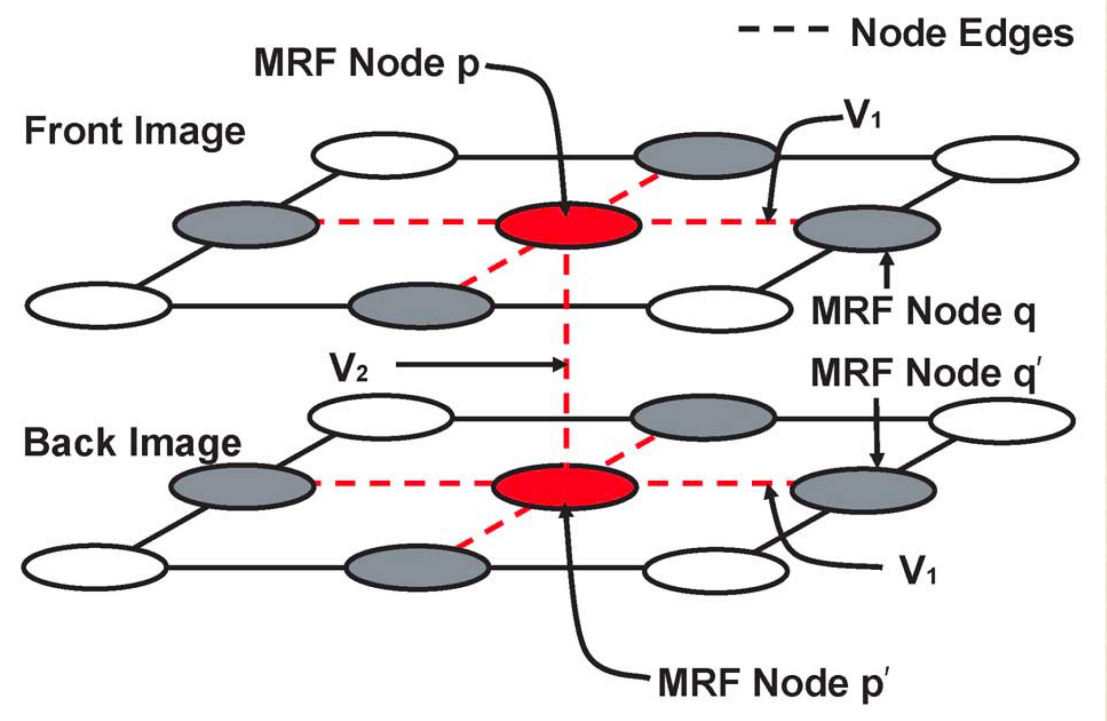
Local Alignment

- Compute correlations at a set of grid points: 60x60 patch; [-10,10] displacement
- Smoothen the local displacements (TPS)
- Warp the verso image



Dual Layer MRF

- Consider two sets of nodes, one for recto and one for verso
- The pixels/nodes of the images are aligned
- Connect corresponding pixels from either side.



Defining Potentials: Data Term

- Class Similarity:

$$S_{\mathcal{F}} = \sum_{m \in \pi^{\mathcal{F}}} \exp(-{}^{d_{pm}^2}\!/_{d_p^2})$$

$$S_{\mathcal{J}} = \sum_{m \in \pi^{\mathcal{J}}} \exp(-{}^{d_{pm}^2}\!/_{d_p^2})$$

$$S_{\mathcal{B}} = \sum_{m \in \pi^{\mathcal{B}}} \exp(-{}^{d_{pm}^2}\!/_{d_p^2}),$$

where d_{pm} is the distance from pixel p to the m^{th} cluster center and d_p is the average distance to K nearest cluster centers in training data.

- Unary:

$$E_d(l_p = \mathcal{F}) = \frac{S_{\mathcal{J}} + S_{\mathcal{B}}}{2 \times (S_{\mathcal{F}} + S_{\mathcal{J}} + S_{\mathcal{B}})}$$

$$E_d(l_p = \mathcal{J}) = \frac{S_{\mathcal{F}} + S_{\mathcal{B}}}{2 \times (S_{\mathcal{F}} + S_{\mathcal{J}} + S_{\mathcal{B}})}$$

$$E_d(l_p = \mathcal{B}) = \frac{S_{\mathcal{F}} + S_{\mathcal{J}}}{2 \times (S_{\mathcal{F}} + S_{\mathcal{J}} + S_{\mathcal{B}})}$$

Defining Potentials: Smoothness Term

- Pairwise:

$$\mathrm{E}_s = \sum_{p,q \in \mathcal{N}} \mathrm{V}_1(l_p, l_q) + \sum_{p,p' \in \mathcal{M}} \mathrm{V}_2(l_p, l_{p'})$$

- Intra-Layer:

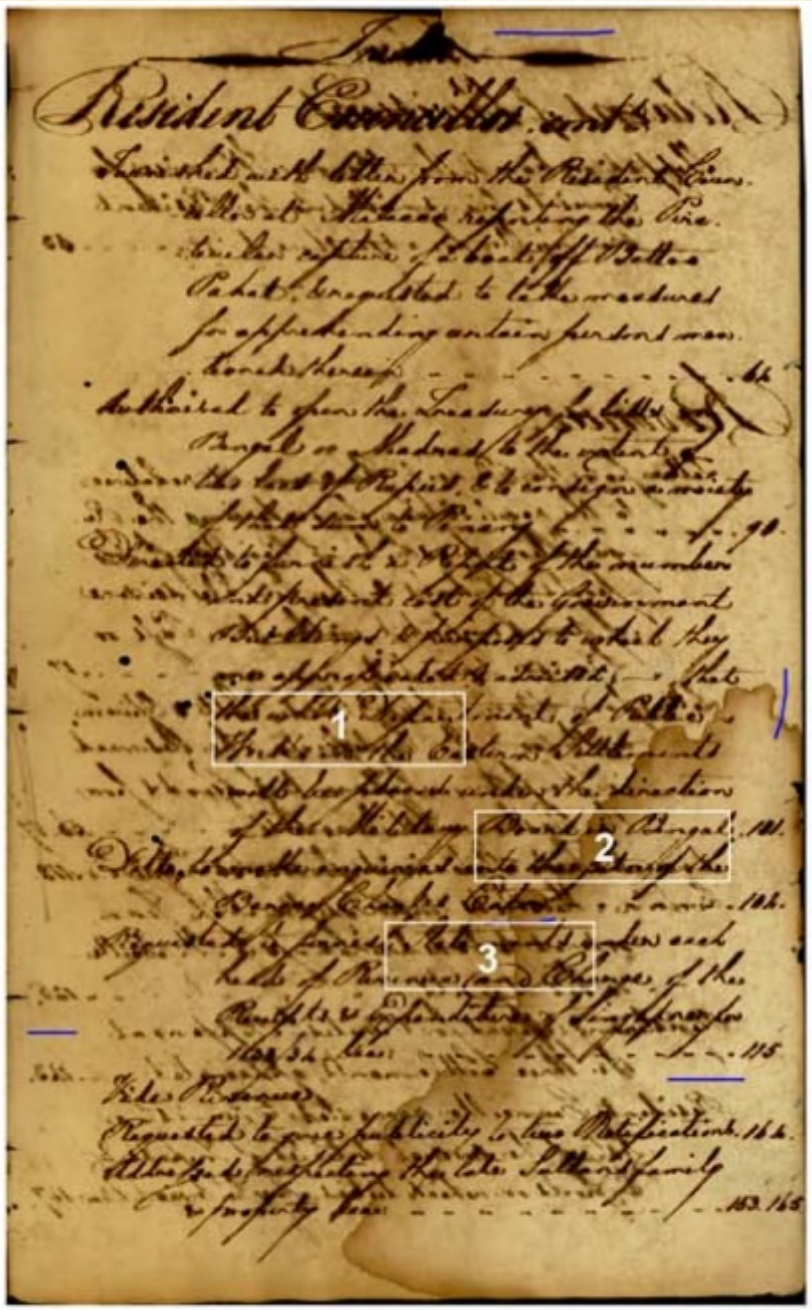
$$\mathrm{V}_1(l_p, l_q) = \frac{1}{1 + (\xi_{pq})^2}$$

ξ_{pq}:

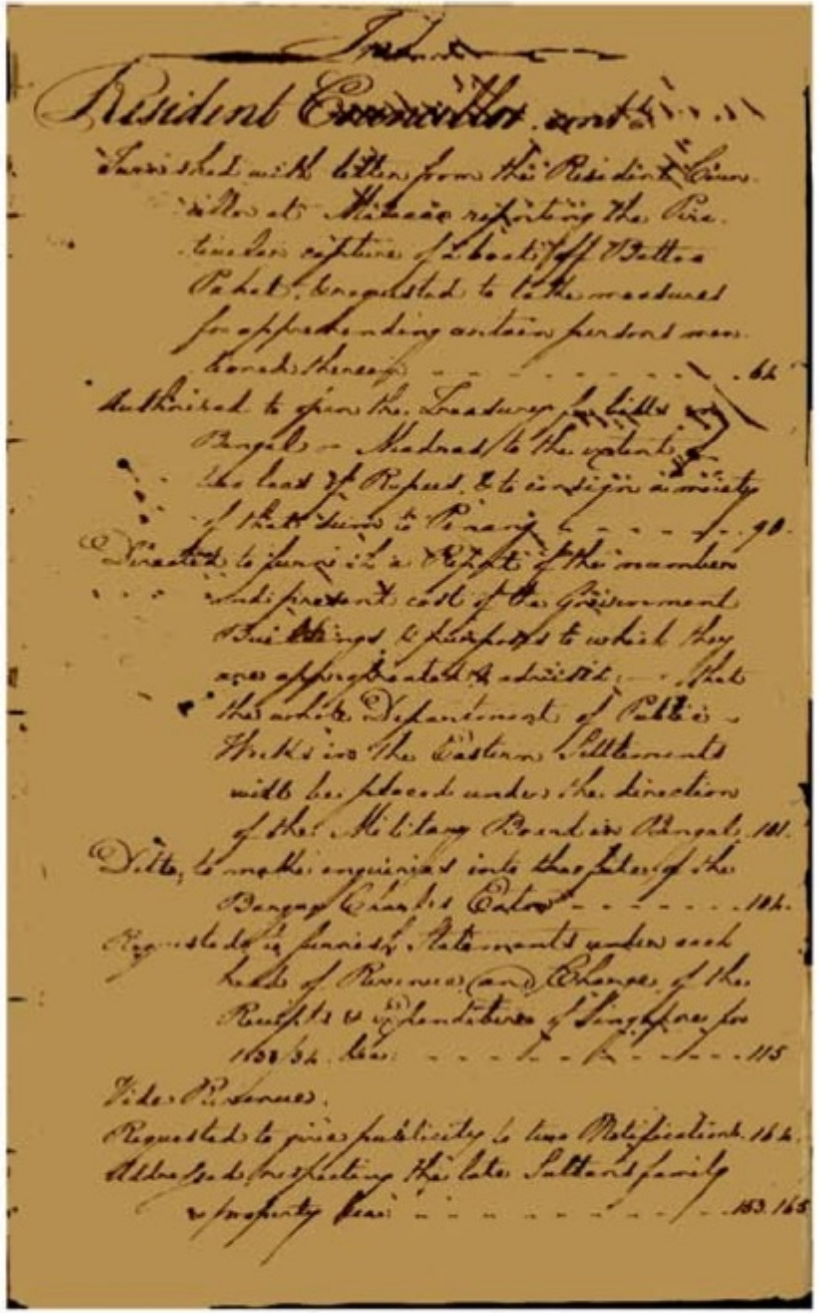
l_p	l_q		
	Foreground	Ink-Bleed	Background
Foreground	∞	d^{ρ}_{pq}	d^{c}_{pq}
Ink-Bleed	d^{ρ}_{pq}	∞	d^{ρ}_{pq}
Background	d^{c}_{pq}	d^{ρ}_{pq}	∞

- Inter-Layer:

l_p	$l_{p'}$		
	Foreground	Ink-Bleed	Background
Foreground	0	0	0
Ink-Bleed	0	∞	∞
Background	0	∞	2ω

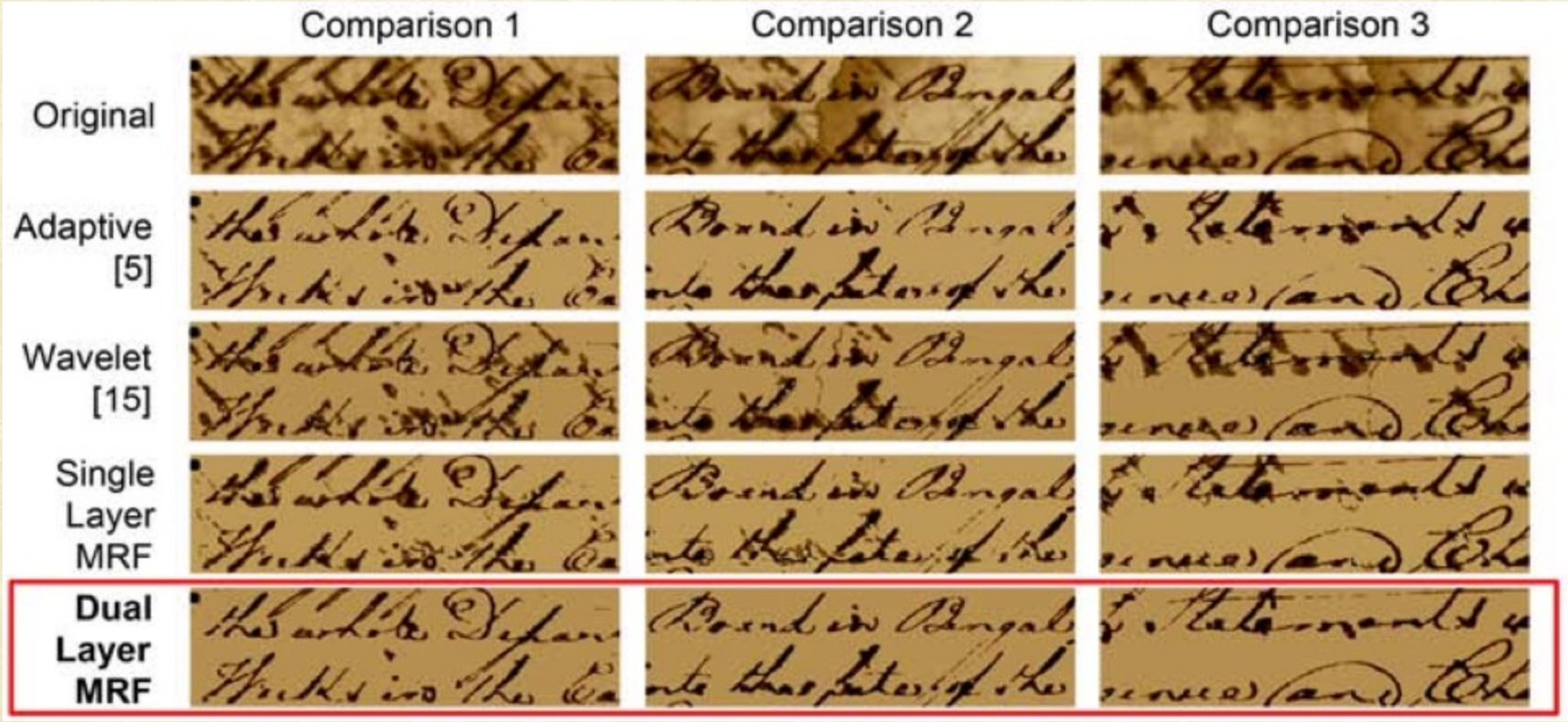


Front+Markup

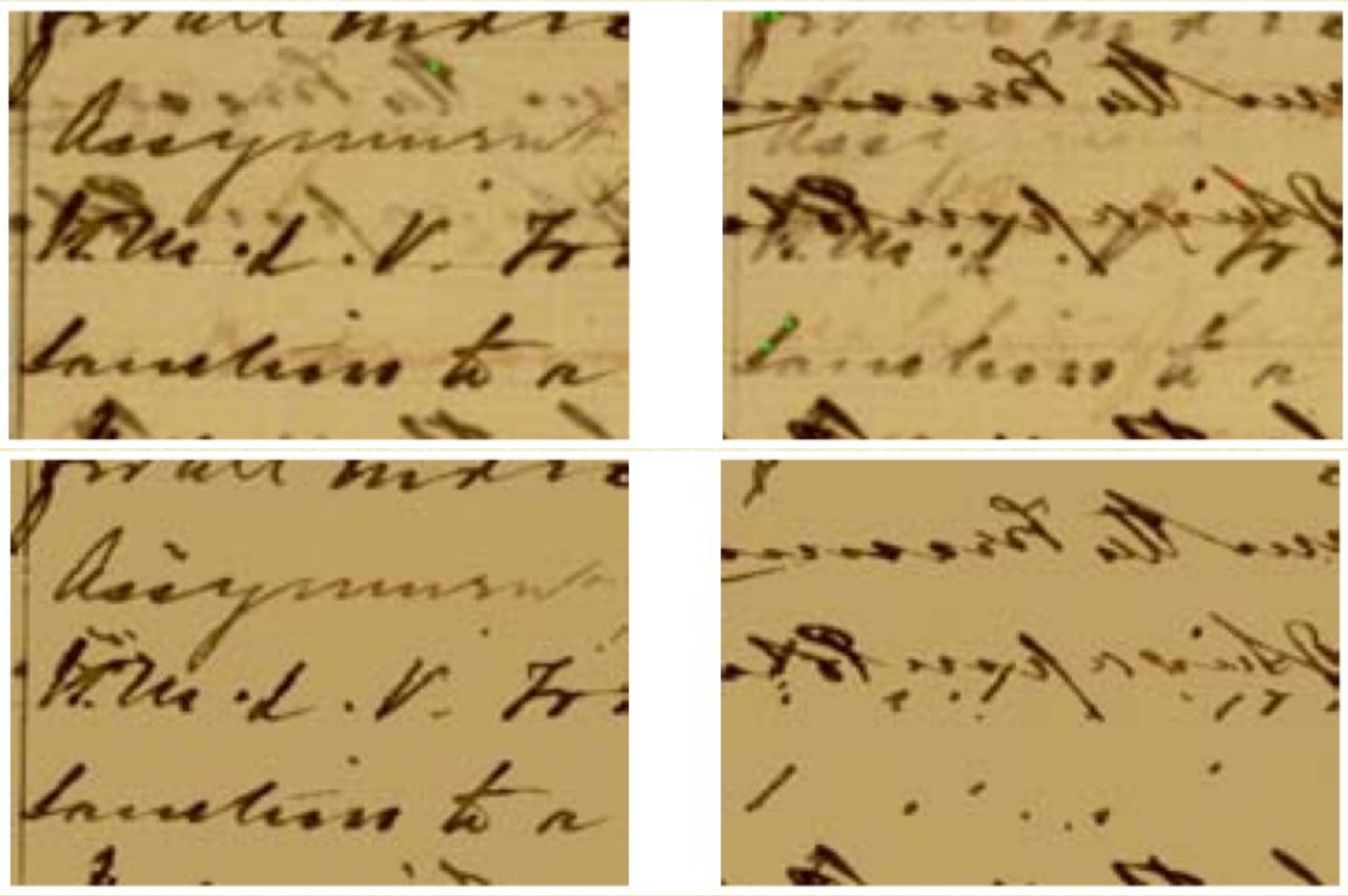


Corrected

Results: Comparison

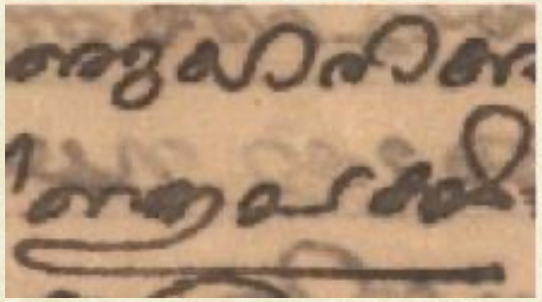


A Closer Look



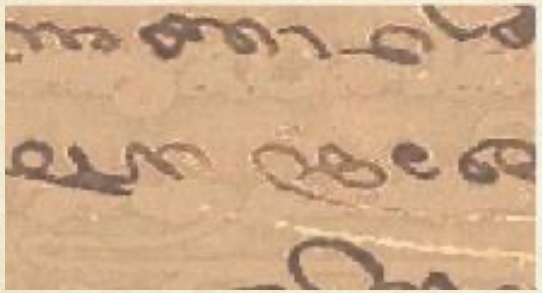
Going Beyond Segmentation

- Can we separate the layers or de-bleed the document?

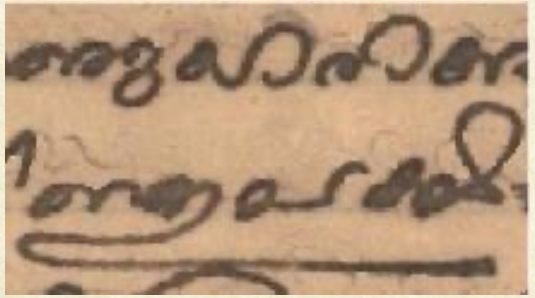


Input Image

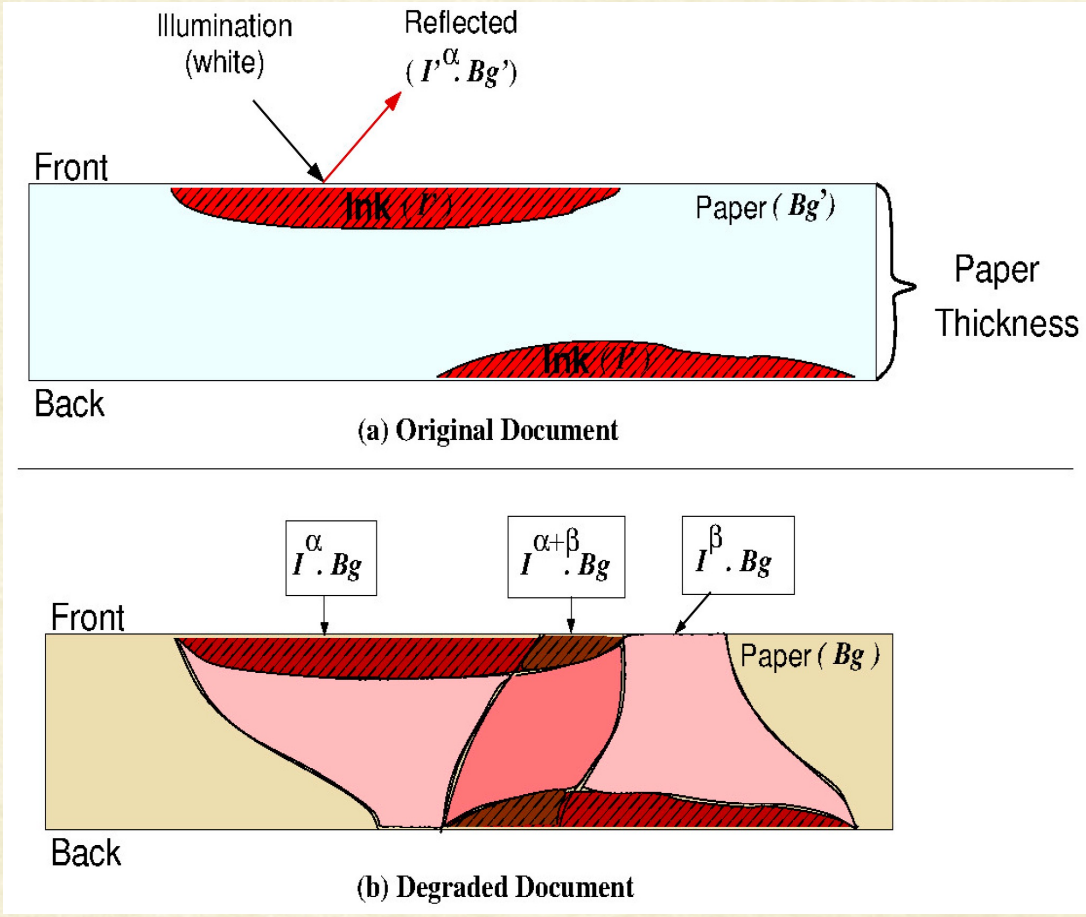
Background Layer



Bleed Layer



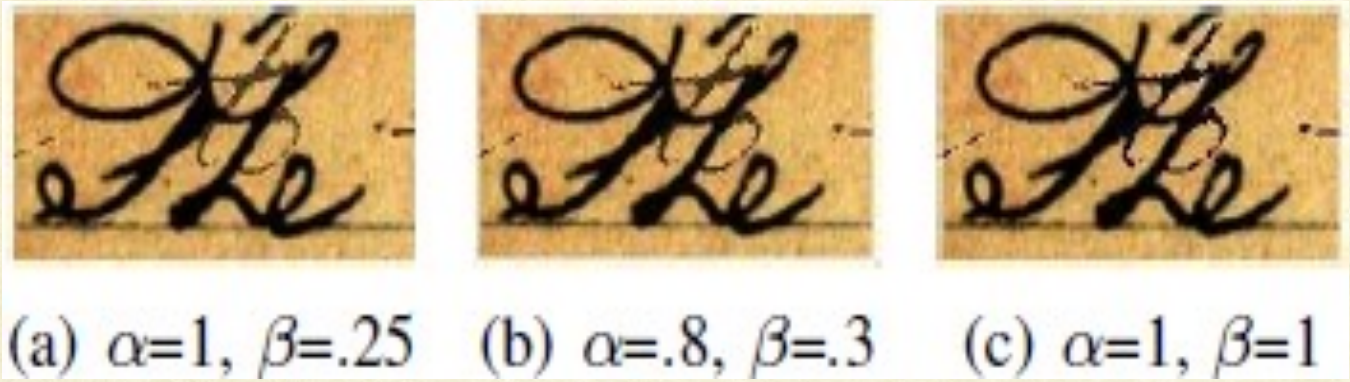
Output(Ink Layer)



Each pixel contain different levels of ink, bleed and background.

Shrikant Baronia and Anoop M. Namboodiri, “Ink-Bleed Reduction using Layer Separation”, Proc. ICDAR 2013.

Multi-Label Classification Model



(a) α=1, β=.25 (b) α=.8, β=.3 (c) α=1, β=1

$$Ink_p = I^{(\alpha_p + \beta_p)} Bg_p$$

- Assume α and β are discrete {0,0.25,0.5,0.75,1}
- Reduces to 20 possible labels for a pixel.
- Define the energy terms and solve for labels.

Defining Potentials

- Total Energy:

$$E = E_d + \phi E_s$$

- Unary:

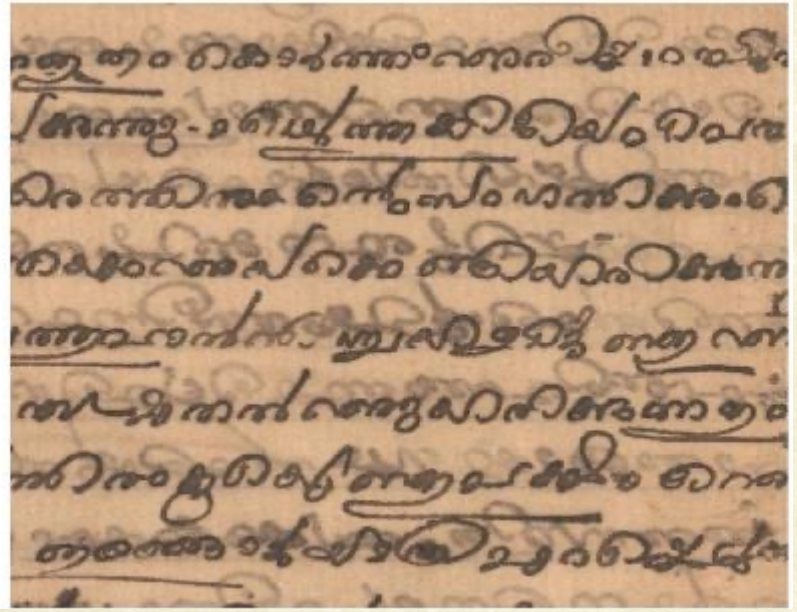
$$E_d(l_p = l^i) = \frac{\sum_j S^j - S^i}{19 \times \sum_j S^j}$$

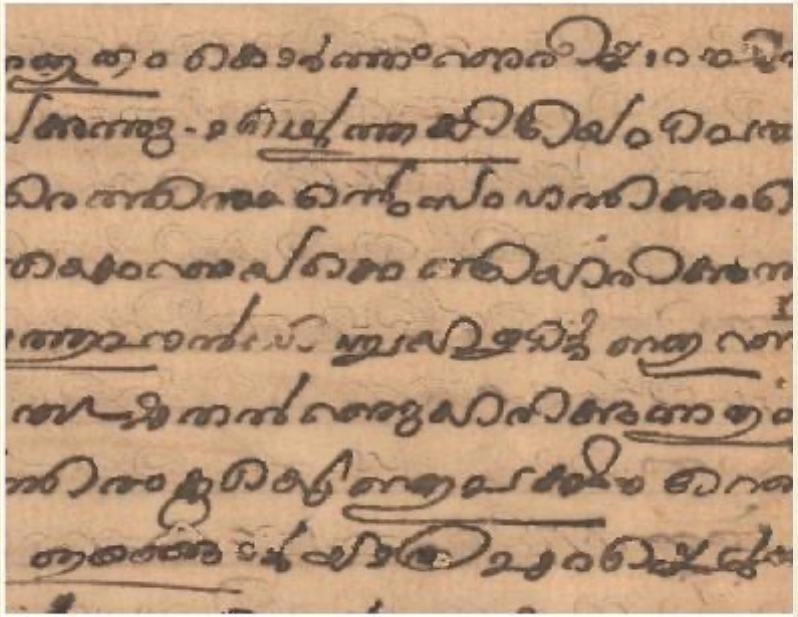
- Pairwise:

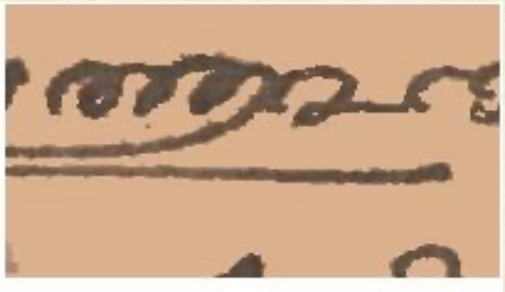
$$\mathrm{E}_s = \sum_{p,q \in \mathcal{N}} \mathrm{V}_1(l_p, l_q) + \sum_{p,p' \in \mathcal{M}} \mathrm{V}_2(l_p, l_{p'})$$

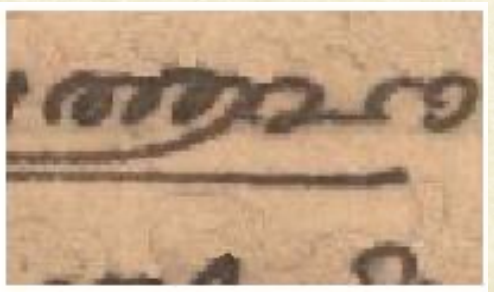
- Solved using α-expansion move of graph cuts.

Results









Input

Binary DMRF

Multilabel DMRF

Questions?