Grouping with Bias

Stella X. Yu [1,2]   *Jianbo Shi* [1]

Robotics Institute[1]
Carnegie Mellon University
Center for the Neural Basis of Cognition[2]

THE ROBOTICS INSTITUTE

What Is It About?

- Incorporating prior knowledge into grouping
  - Unitary generative model
  - Global configurations: partially labelled data and object models
  - Attention

- Computation
  - Efficient solution in a graph partitioning framework

- Goals
  - Bridge the gap between generative models and discriminative models
  - Bridge the gap between formulation and computation

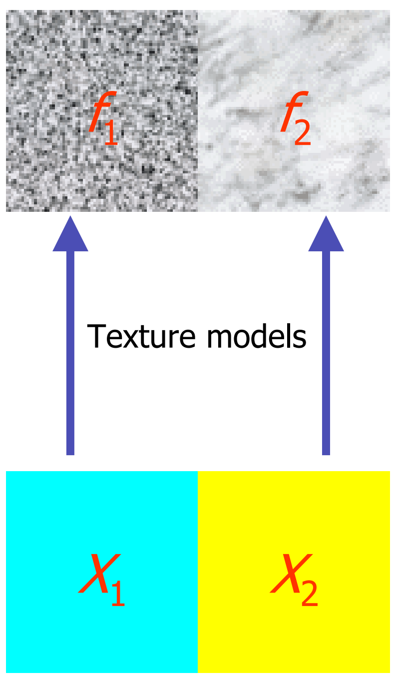
Grouping with Markov Random Fields

MRF: data structure = data generation model + segmentation model

$$\min E(X; f) = -\log p(f \mid X) - \log p(X)$$

Image as observation f

Grouping as state X



Segmentation is to find a partitioning of an image, with generative models explaining each partition.

Generative models constrain the continuous observation data, the segmentation model constrains the discrete states.

The solution sought is the most probable state, or the state of the lowest energy.

[Geman & Geman, 84, …]

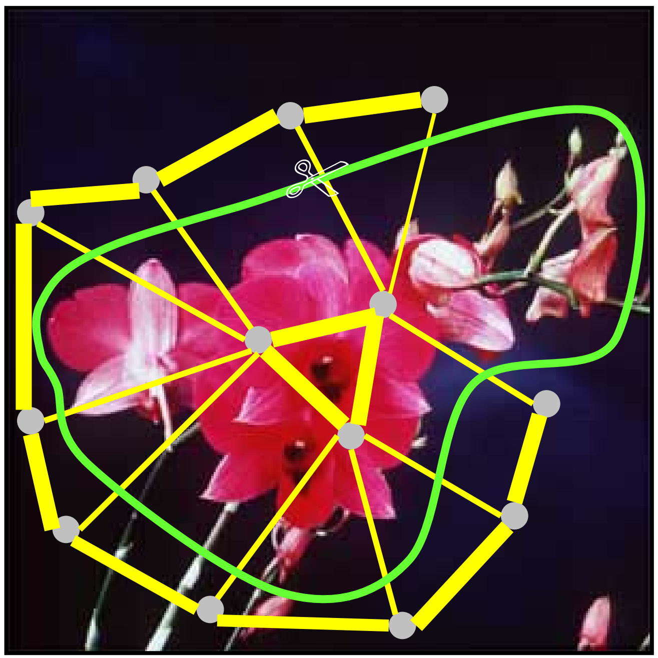
Grouping with Spectral Graph Partitioning

SGP: data structure = a weighted graph, weights describing data affinity

$$\min \; Ncut(V_1, V_2) = \frac{cut(V_1, V_2)}{\deg(V_1) \cdot \deg(V_2)}$$

$$cut(V_1, V_2) = \sum_{p \in V_1} \sum_{q \in V_2} W(p,q)$$

$$\deg(V_1) = \sum_{p \in V_1} \sum_{q \in V} W(p,q)$$



Segmentation is to find a node partitioning of a relational graph, with minimum total cut-off affinity.

Discriminative models are used to evaluate the weights between nodes.

The solution sought is the cuts of the minimum energy.

[Shi & Malik, 97; Perona & Freeman, 98; Malik et al, 01, …]

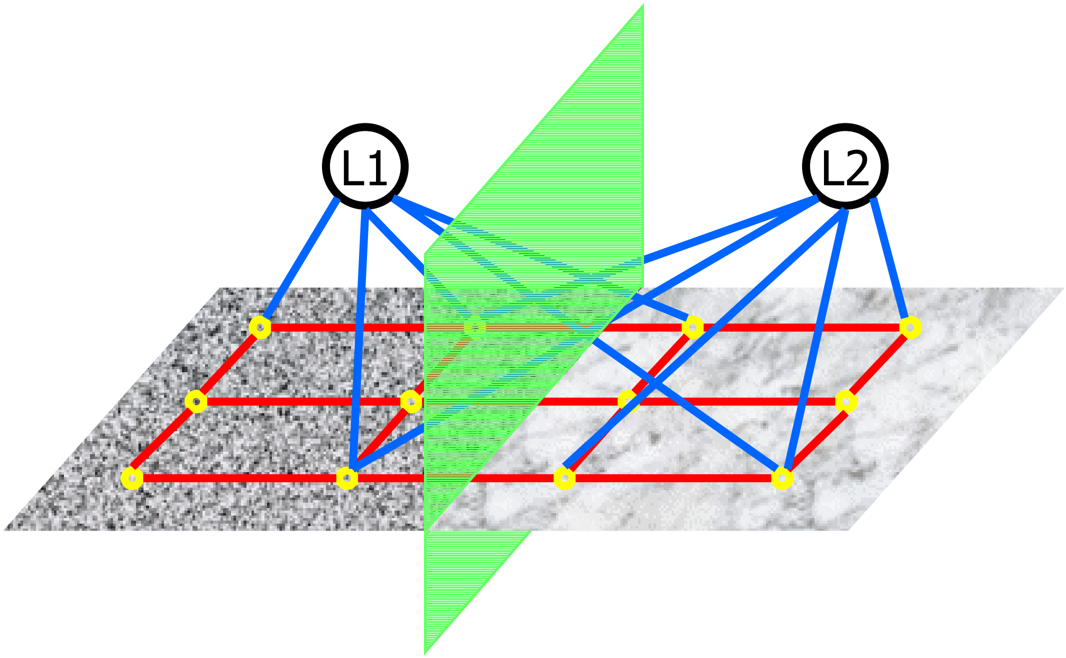
Solving MRF by Graph Partitioning

Some simple MRF models can be translated into graph partitioning

$$\min E(X;f) = \sum_{p} \sum_{q \in N(p)} W_{p,q}(X_p, X_q) + \sum_{p} U_p(X_p, f_p)$$

Binary relationships

Unitary measures



[Greig et al, 89; Ferrari et al, 95; Boykov et al, 98; Roy & Cox, 98; Ishikawa & Geiger, 98, …]

THE ROBOTICS INSTITUTE

Comparison of Two Approaches

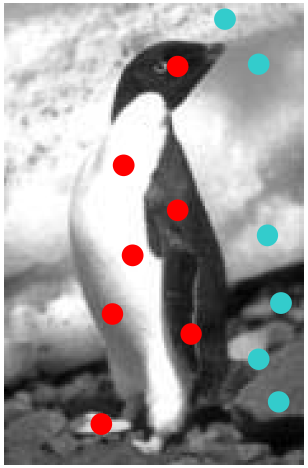
Pros \ Cons	Formulation	Computation
Markov Random Fields	Generative models Bayesian interpretation General local interaction Sensitive to model mismatch	Simulation: e.g. Gibbs sampler Parameter estimation is hard Difficult to compute probability Convergence is very slow Only local optimum
Graph Partitioning	Discriminative models No models required Lack prior to guide grouping	Spectral decomposition Fast and robust Global optimum

THE ROBOTICS INSTITUTE

Prior Knowledge in Grouping

Local Constraints	Global Configuration Constraints	
Unitary generative models	Object models: What to look for	Attention: Where to look for

Red foreground



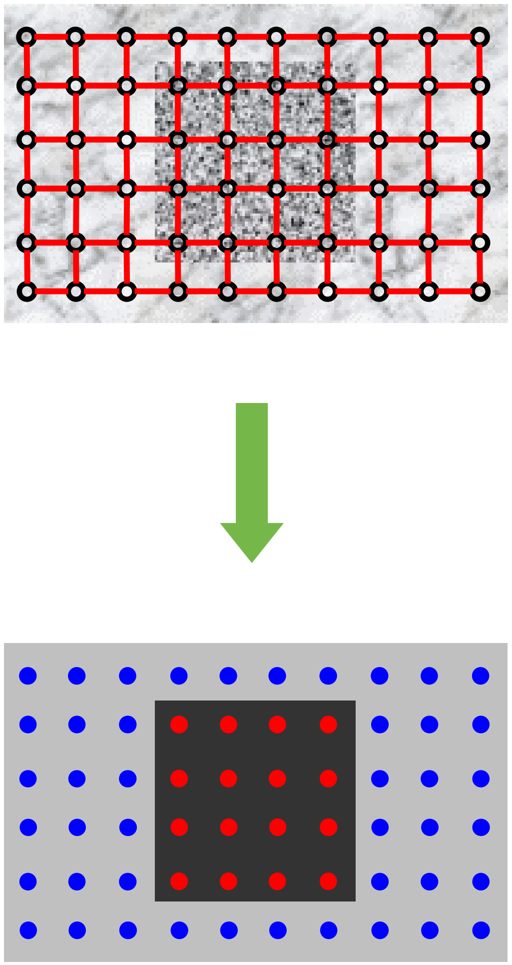
Partial grouping

Spatial attention

➔ How to encode them in discriminative models, e.g. SGP?

THE ROBOTICS INSTITUTE

Review: Segmentation on Relational Graphs



G = (V, E, A, R)

V: each node denotes a pixel

E: each edge denotes a pixel-pixel relationship

A: each weight measures pairwise similarity

R: each weight measures pairwise dissimilarity

Segmentation = node partitioning

break V into disjoint sets V_1 , V_2 ; so that

cut-off attraction is small

cut-off repulsion is large

Dual criteria on dual measures

- ➔ Maximize within-group A and between-group R
- ➔ Minimize between-group A and within-group R

THE ROBOTICS INSTITUTE

Review: Energy Function Formulation

$$X_l(u) = \begin{cases} 1, & u \in V_l \\ 0, & u \notin V_l \end{cases}$$

Group indicators

$$W = A - R + D_R$$

Weight matrix

$$D = D_A + D_R$$

Degree matrix

$$y = (1-\alpha)X_1 - \alpha X_2, \quad \alpha = \frac{\deg(V_1)}{\deg(V)}$$

Change of variables

$$Nassoc(X_1, X_2) = \sum_{t=1}^{2} \frac{X_t^T W X_t}{X_t^T D X_t} = \frac{y^T W y}{y^T D y}$$

Energy function as a Rayleigh quotient

$$\max \frac{y^T W y}{y^T D y} \Rightarrow W y = \lambda_1 D y$$

Eigenvector as solution

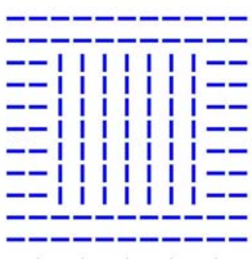
Review: Eigenvector as a Solution

- The derivation holds so long as $X_1 + X_2 = 1$

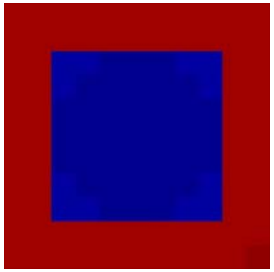
$$y = (1-\alpha)X_1 - \alpha\, X_2 = X_1 - \alpha$$

- The eigenvector solution is a linear transformation, scaled and offset version of the probabilistic membership indicator for one group.

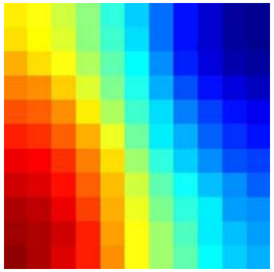
- If y is well separated, then two groups are well defined; otherwise, the separation is ambiguous



stimulus



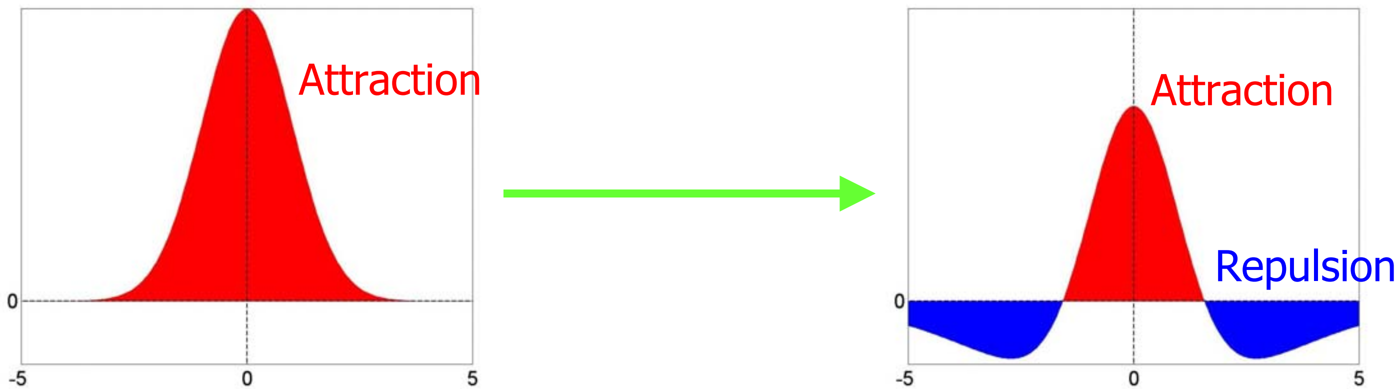
Solution y
well separated



Solution y
ambiguous

Interaction: from Gaussian to Mexican Hat

$$W = A - R + D_R$$

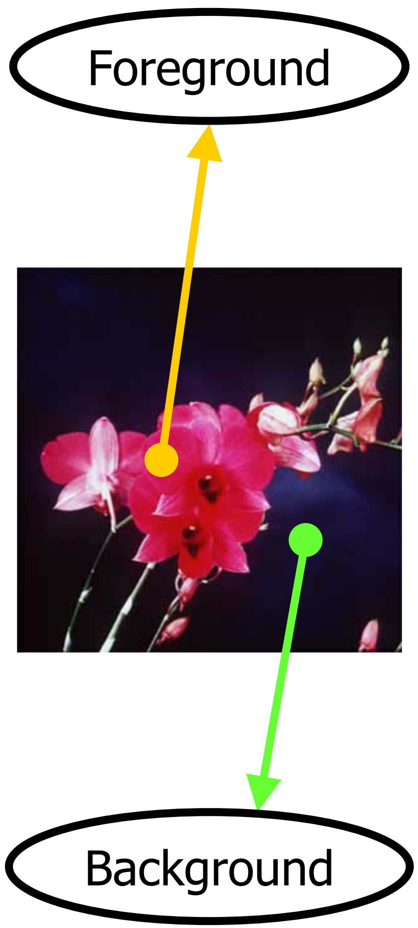


$$W_{ij} = e^{-\frac{(f_i - f_j)^2}{2\sigma_1^2}} - \frac{\sigma_1}{\sigma_2} e^{-\frac{(f_i - f_j)^2}{2\sigma_1^2} \cdot \left(\frac{\sigma_1}{\sigma_2}\right)^2}$$

➔ With repulsion, negative correlations in MRF formulations can be translated into graph partitioning formulations directly.

Encoding Bias: Unitary Preference

Preference of red pixels to be foreground, black pixels to be background



- Introduce dummy nodes
- Expand the node set
- Soft Constraints

$$A := \begin{bmatrix} (1-\gamma)A & \gamma B_a \\ \gamma B_a^T & \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \end{bmatrix}$$

$$R := \begin{bmatrix} (1-\gamma)R & \gamma B_r \\ \gamma B_r^T & \gamma \begin{bmatrix} 0 & r \\ r & 0 \end{bmatrix} \end{bmatrix}$$

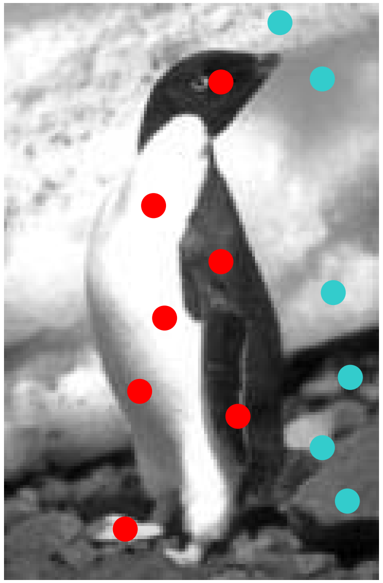
- γ controls the relative weighting between data and preference.

Encoding Bias: Partial Grouping

Manual selection: pixels marked w/ the same color are in one group

Coloring dummy nodes is to assign a particular group label

Foreground



Background

- Introduce partial grouping solution
- Assign a particular group label using dummy nodes
- Hard Constraints

$$X_1(p) = X_1(q), \quad X_2(p) = X_2(q)$$

$$y(p) - y(q) = 0$$

$$m^T y = 0$$

$$m^T = [0, \cdots 0, 1, 0, \cdots, 0, -1, 0, \cdots, 0]$$

Linear: $M^T y = 0 \quad \text{or} \quad y = Q z$

- M is the constraint space
  Q is the reduced solution space.

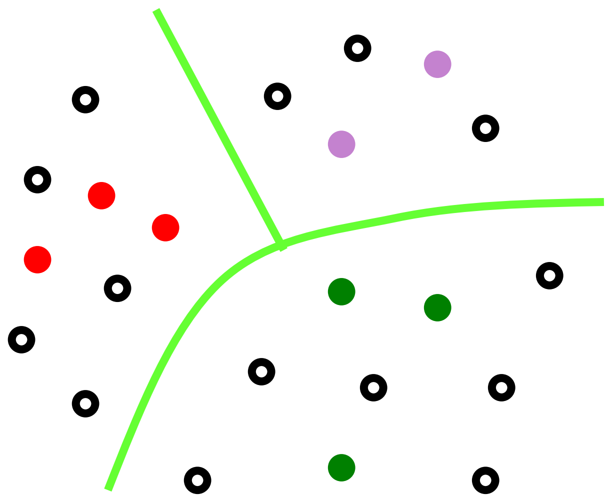
$$M \perp Q, \quad M \oplus Q = All$$

THE ROBOTICS INSTITUTE

Encoding Bias: Constraining Solution Space

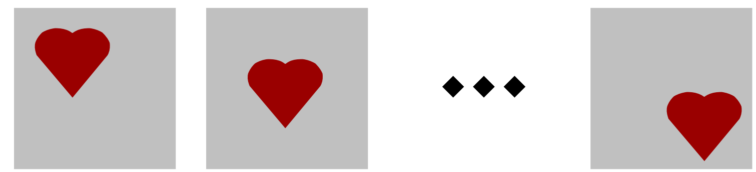
Clustering with
Partially labelled data

$M^T y = 0$



Segmentation with
object models

$y = Q z$



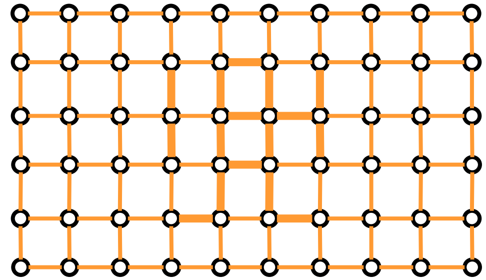
S: translated versions of a shape

- Find the eigenshape Q of S
- Constraining y = Q z allows us to segment out this particular shape in an image.

General form of constraints: $\Psi(y) = 0$

THE ROBOTICS INSTITUTE

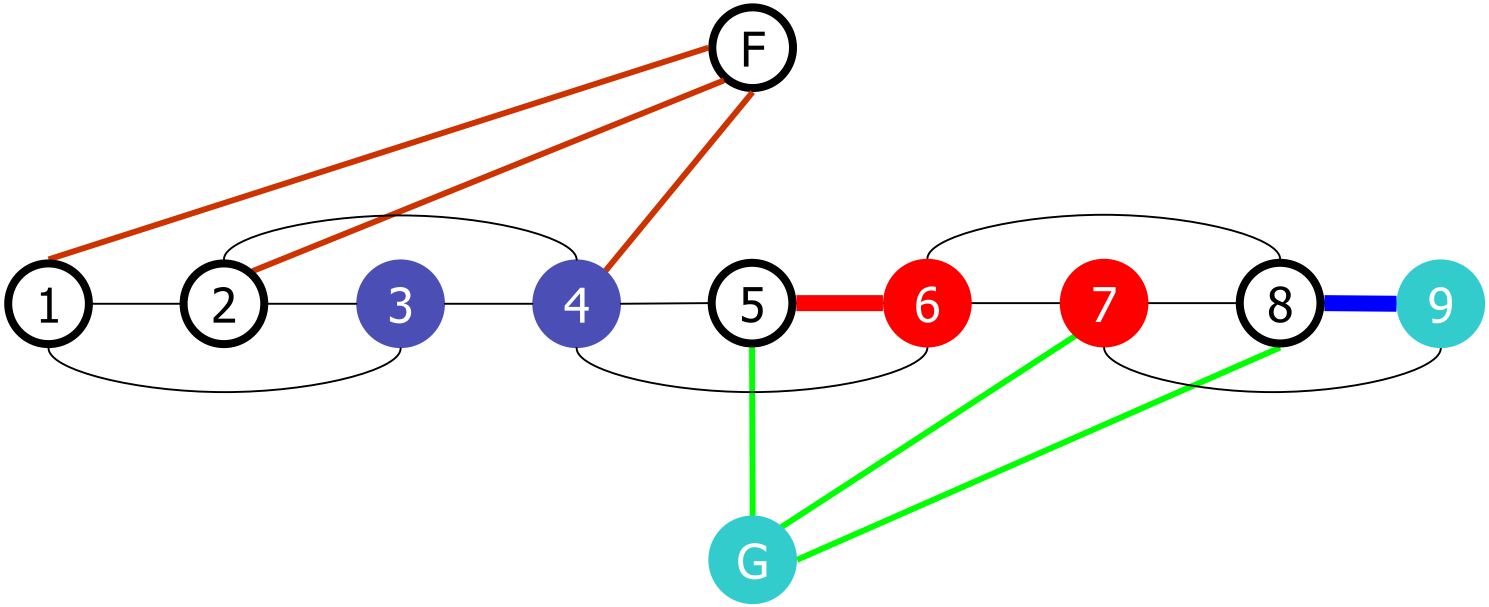
Encoding Bias: Attention



Spatial Attention: center region is analyzed w/ more discrimination

- Modulation
- Connections for some nodes are enhanced / weakened
- Weights at attentional hotspot are less distorted

Representations of Bias



- Partial grouping
  - ➔ {3, 4}, {6, 7}
  - ➔ {9, G}
- Preference
  - ➔ {F, 1,2,4}
  - ➔ {G, 5,7,8}
- Attention
  - ➔ {5, 6}
  - ➔ {8, 9}

Hard constraints

Soft constraints

Modulation

Constrained Optimization

$$Nassoc(y) = \frac{y^T W y}{y^T D y} \qquad s.t. \quad M^T y = 0$$

- Constrained optimization

$$M = U \Sigma V^T$$

$$Q = I - UU^T$$

$$y = Q z$$

- Constraint space
- Feasible solution space

$$Nassoc(z) = \frac{z^T (Q^T W Q) z}{z^T (Q^T D Q) z}$$

- Unconstrained optimization

$$Q^T D^{-1} W \; y = \lambda \; y$$

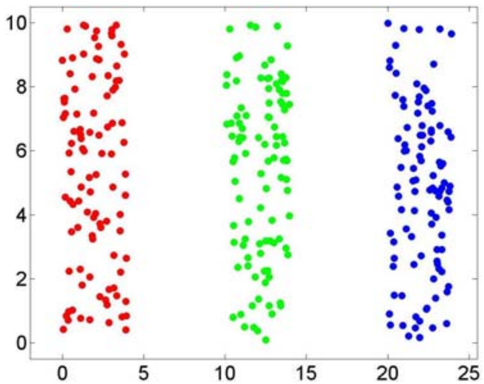
- Eigensolution available

➔ Rank(Q) = # of nodes - # of independent constraints

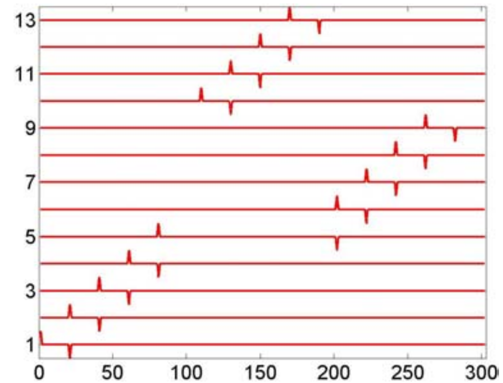
➔ Problem: Unconstrained affinity matrix becomes denser

THE ROBOTICS INSTITUTE

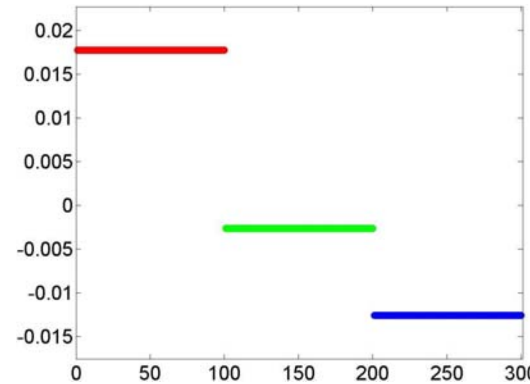
Results: Preference and Partial Grouping



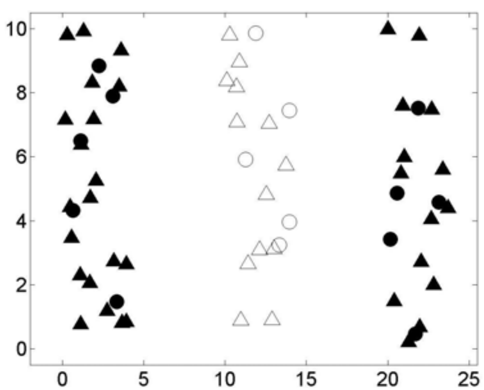
Data: three stripes



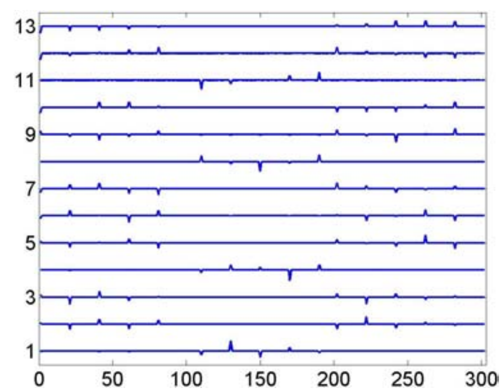
M: Hard constraints



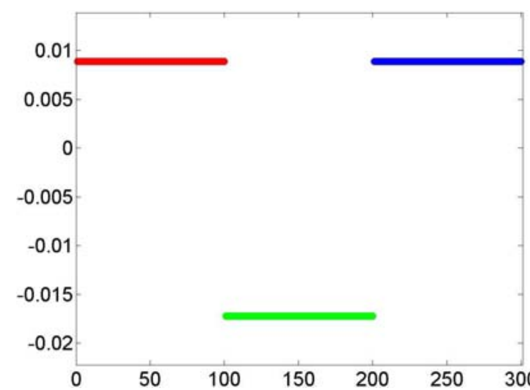
Grouping w/o bias:
Each is one group



O : hard constraints
Δ : soft constraints
F/G : Filled / empty

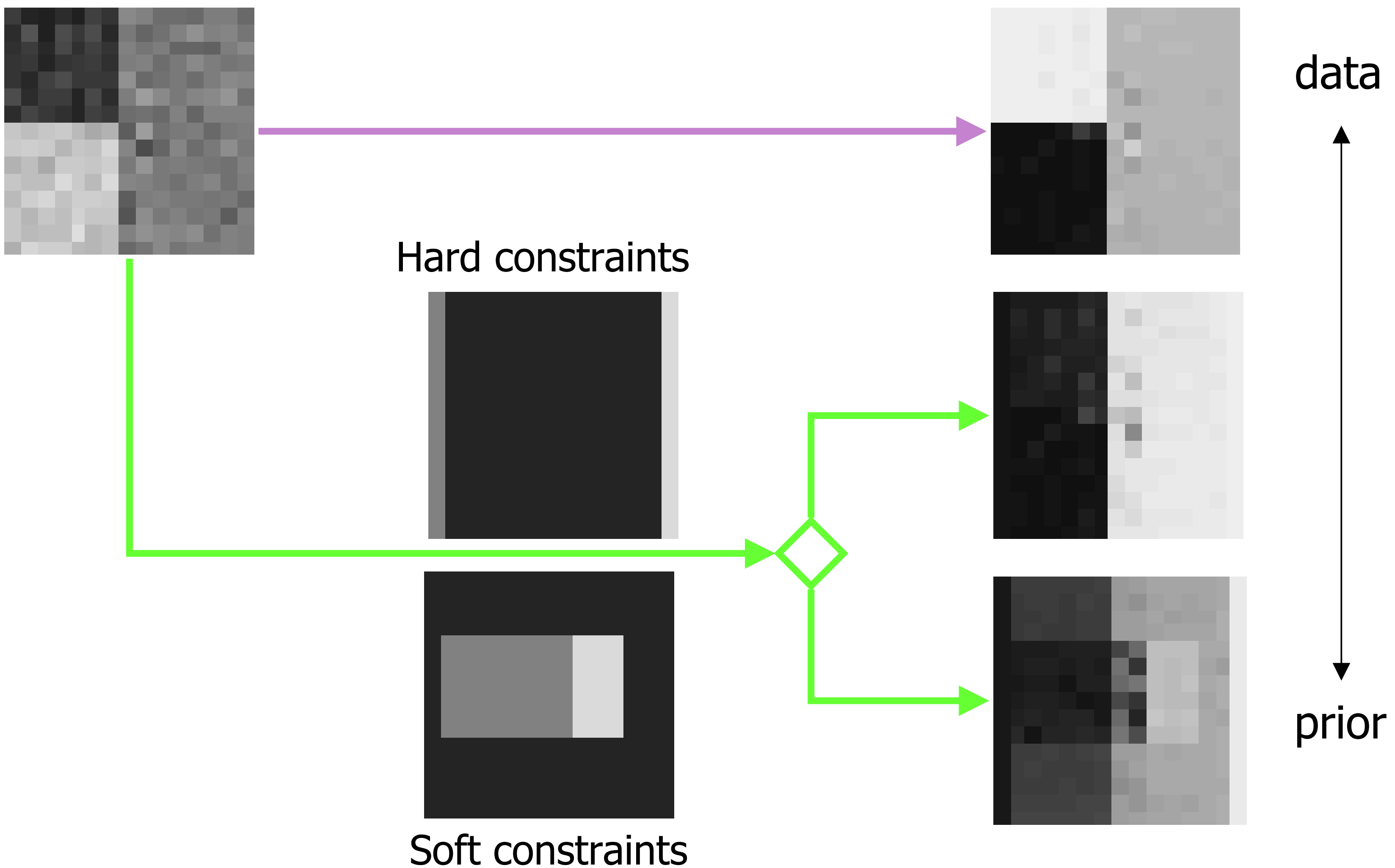


U: Basis of constraint space
$M = U \Sigma V^T$



Grouping w/ bias:
Left and right are one group

Results: Preference and Partial Grouping



Results: Partial Grouping

1st Eigenvector 2nd Eigenvector 3rd Eigenvector

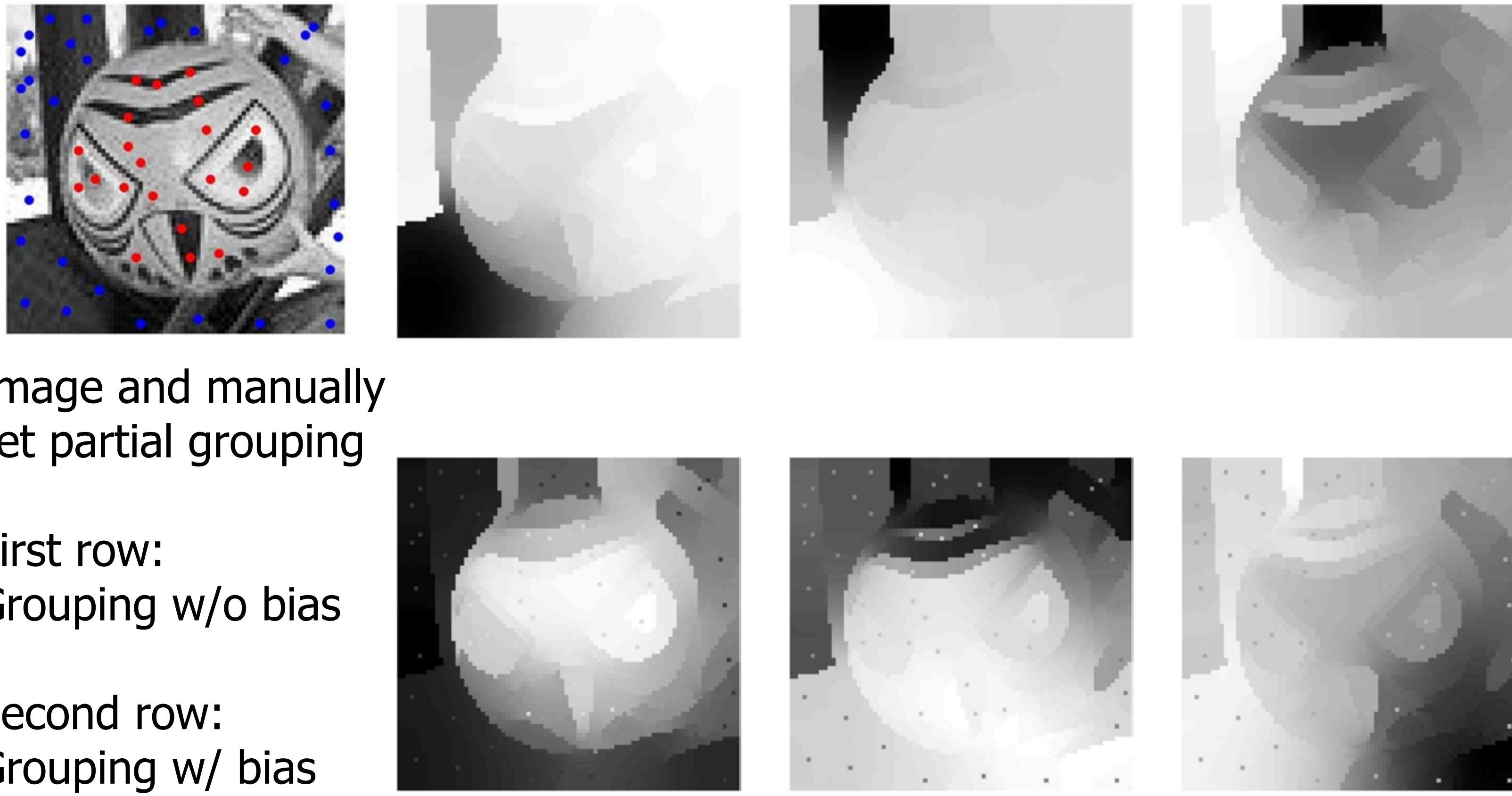


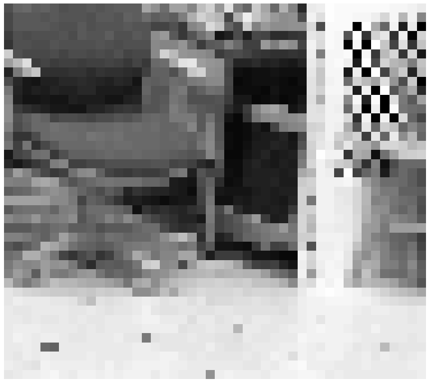
Image and manually set partial grouping

First row: Grouping w/o bias

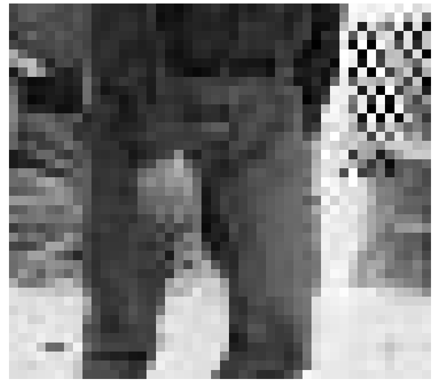
Second row: Grouping w/ bias

The pumpkin starts to emerge as a whole from the background regardless of its surface markings.

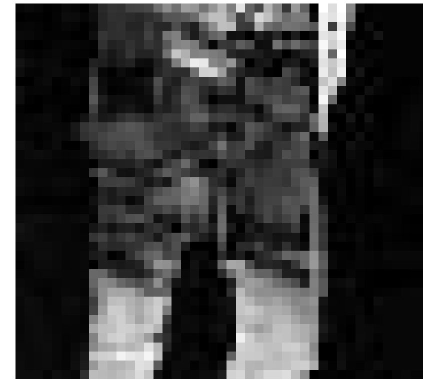
Results: Figure Detection with Soft Constraints



Background



Foreground



Difference

Difference Thresholded used as soft constraints

Grouping of foreground

Grouping of foreground with bias

Results: Spatial Attention