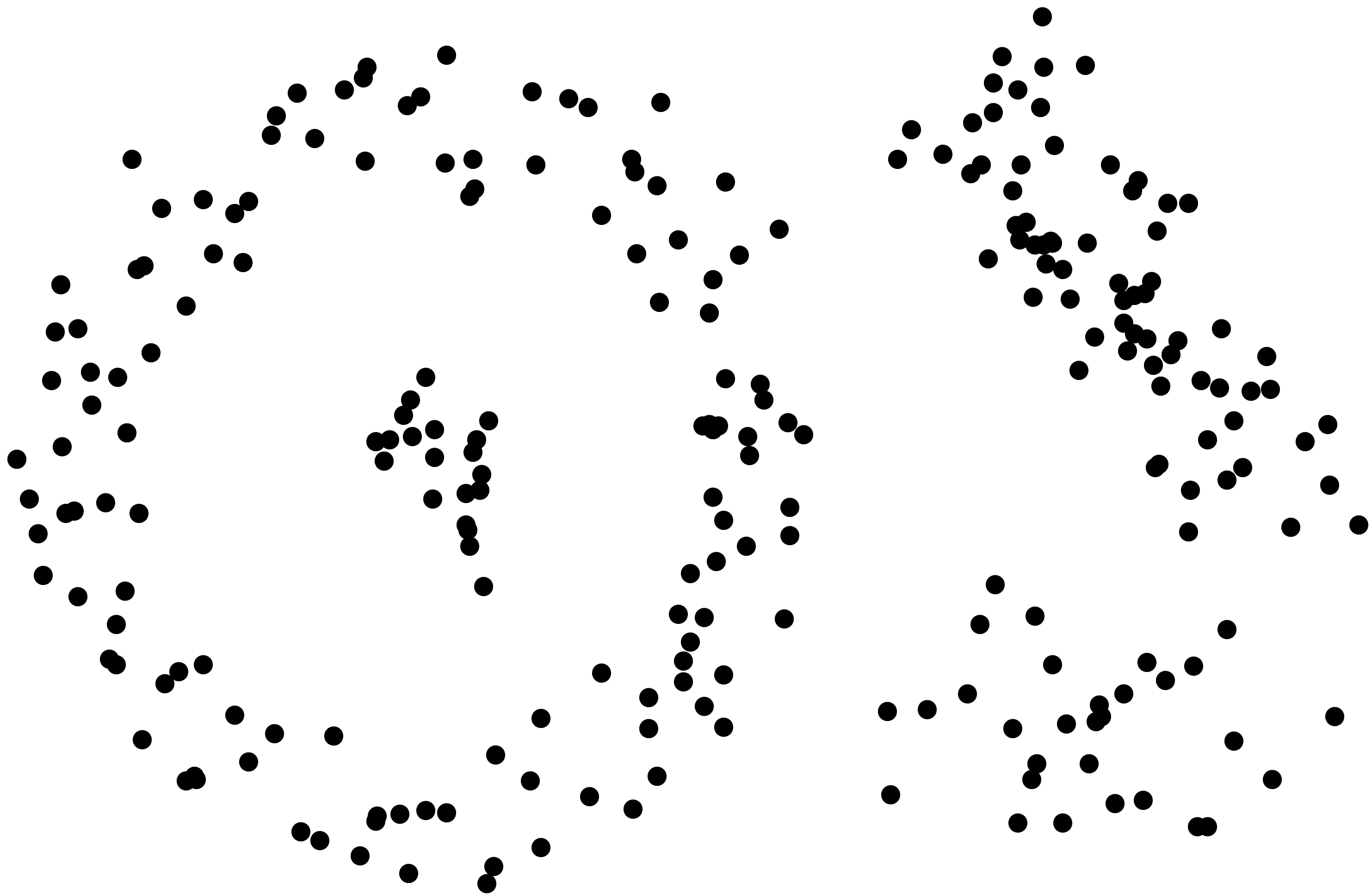
Computational Models of Perceptual Organization

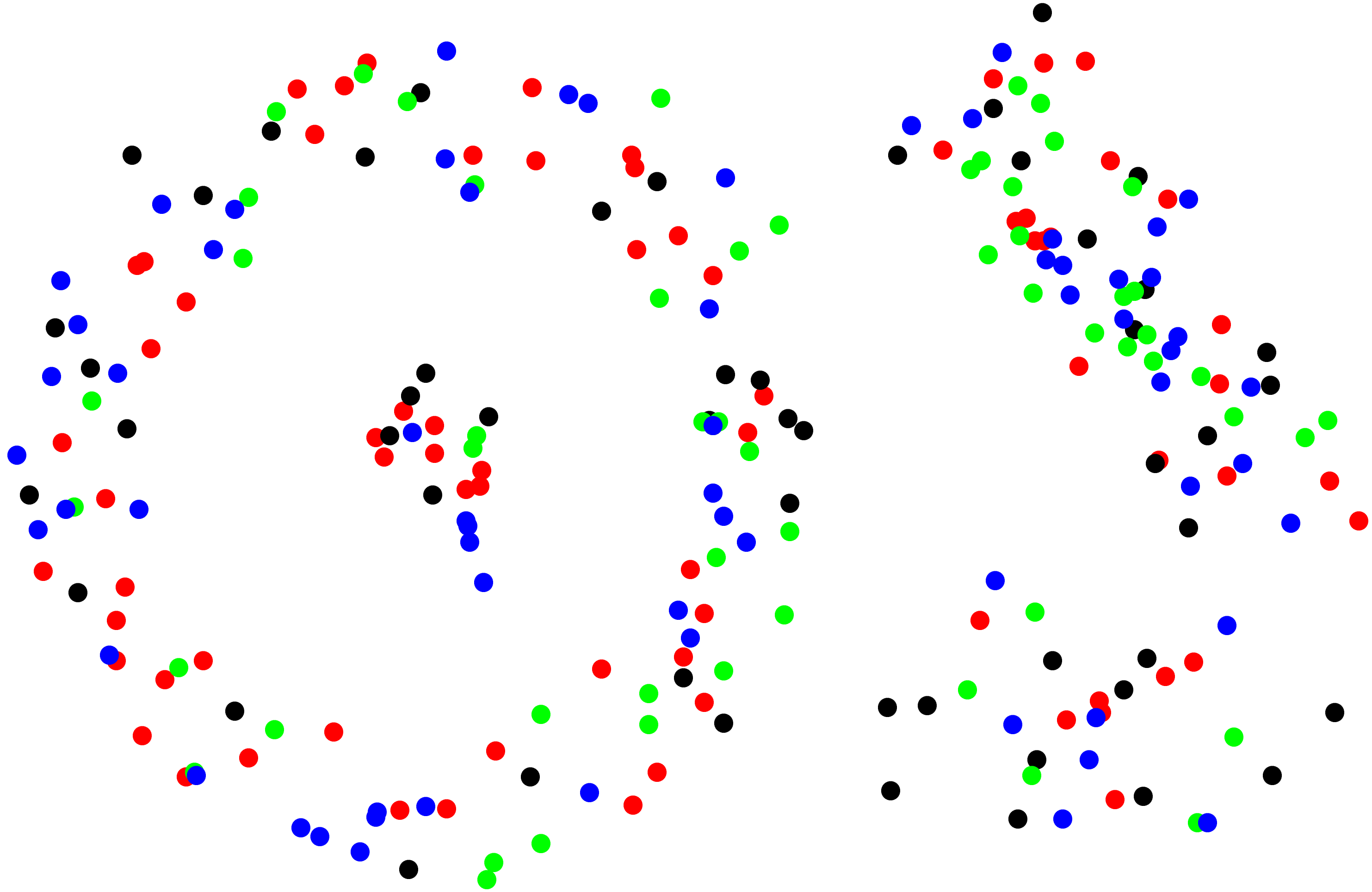
Stella X. Yu

Robotics Institute
Carnegie Mellon University
Center for the Neural Basis of Cognition

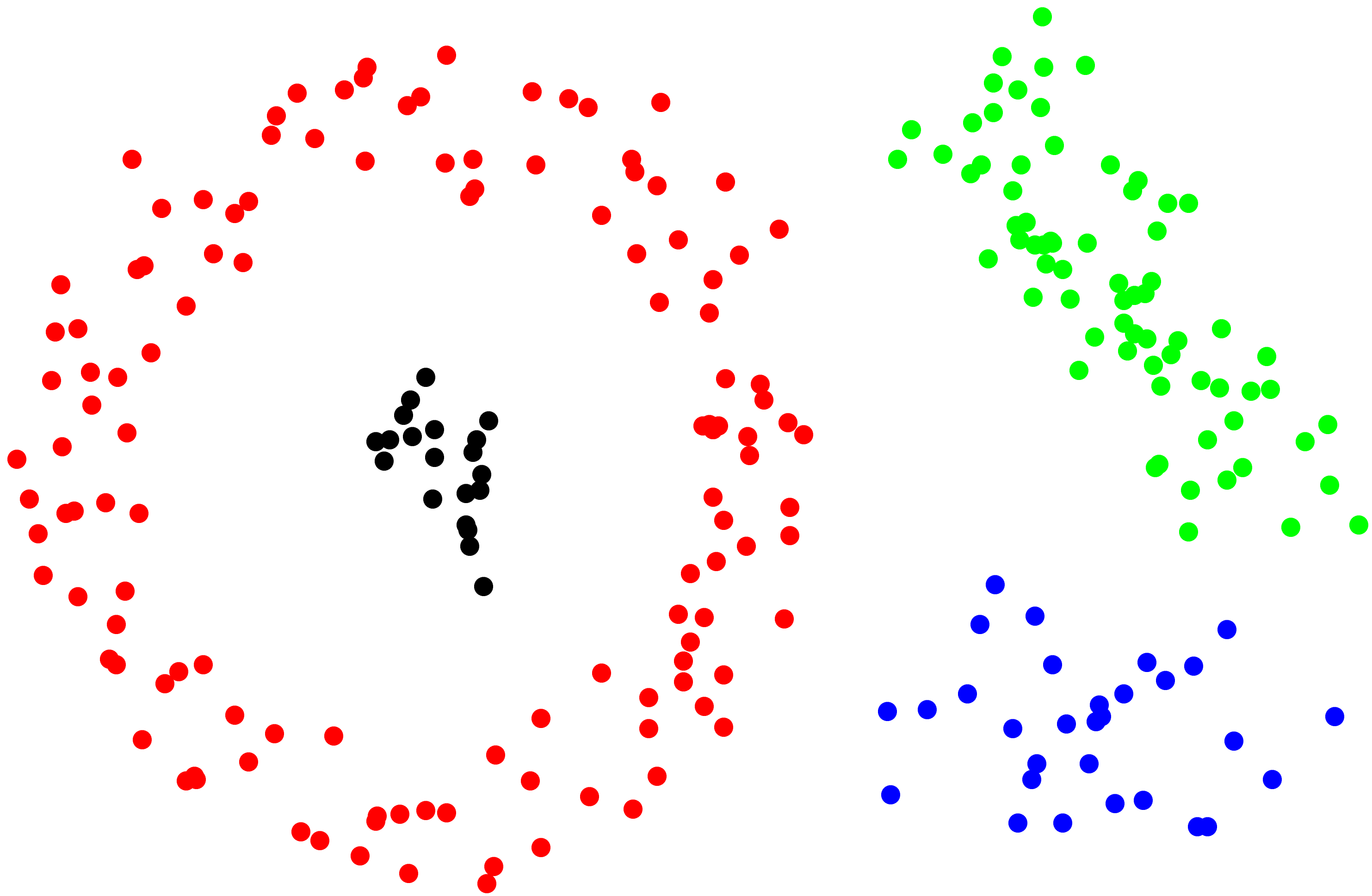
What Is Perceptual Organization



What Is Perceptual Organization

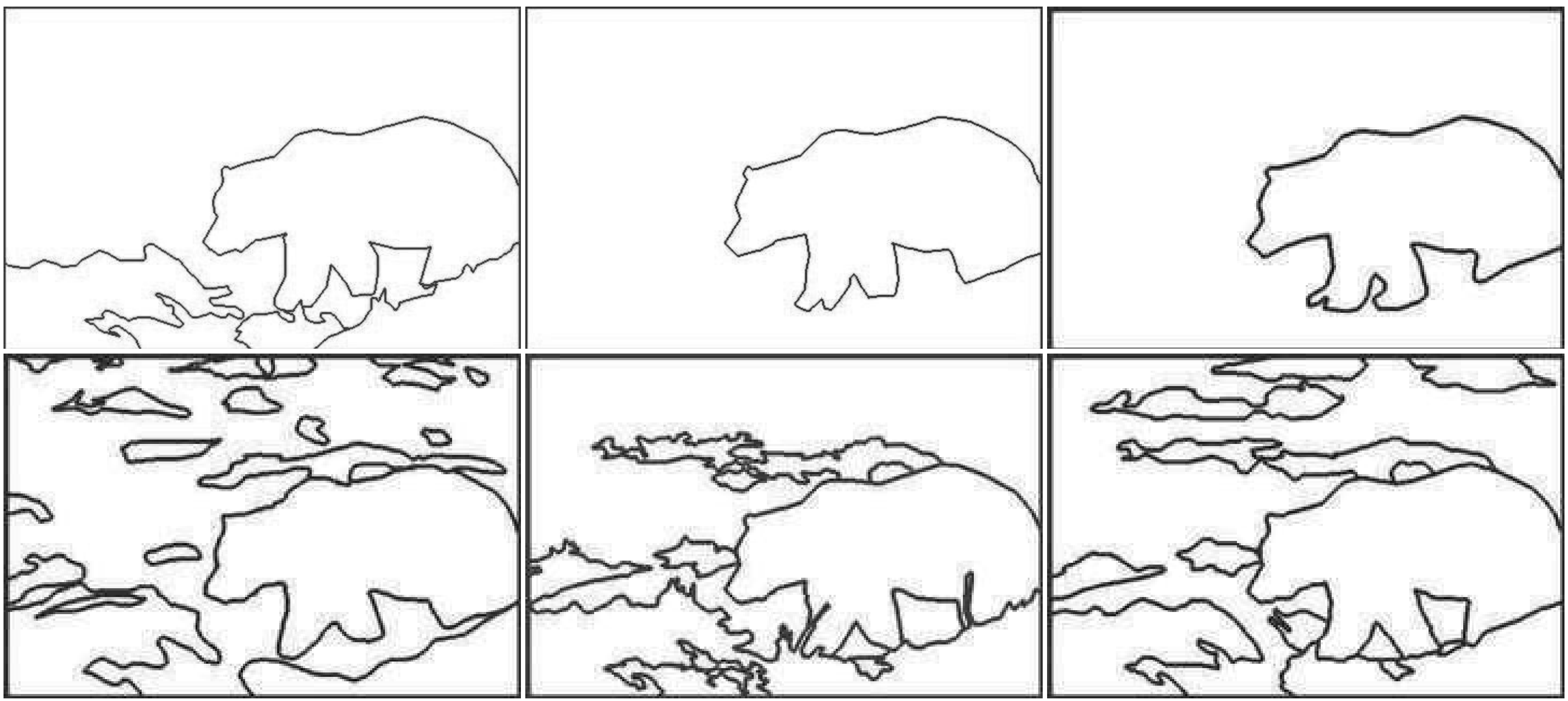


What Is Perceptual Organization



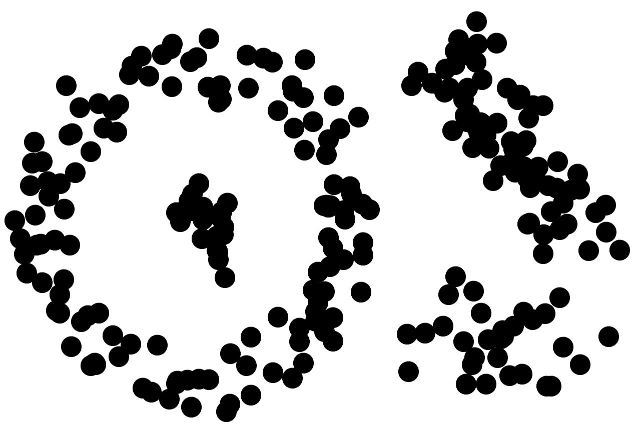
What Is Perceptual Organization

(Martin et al)



What Is Perceptual Organization

- multiple choices
- a variety of features
- content-dependent



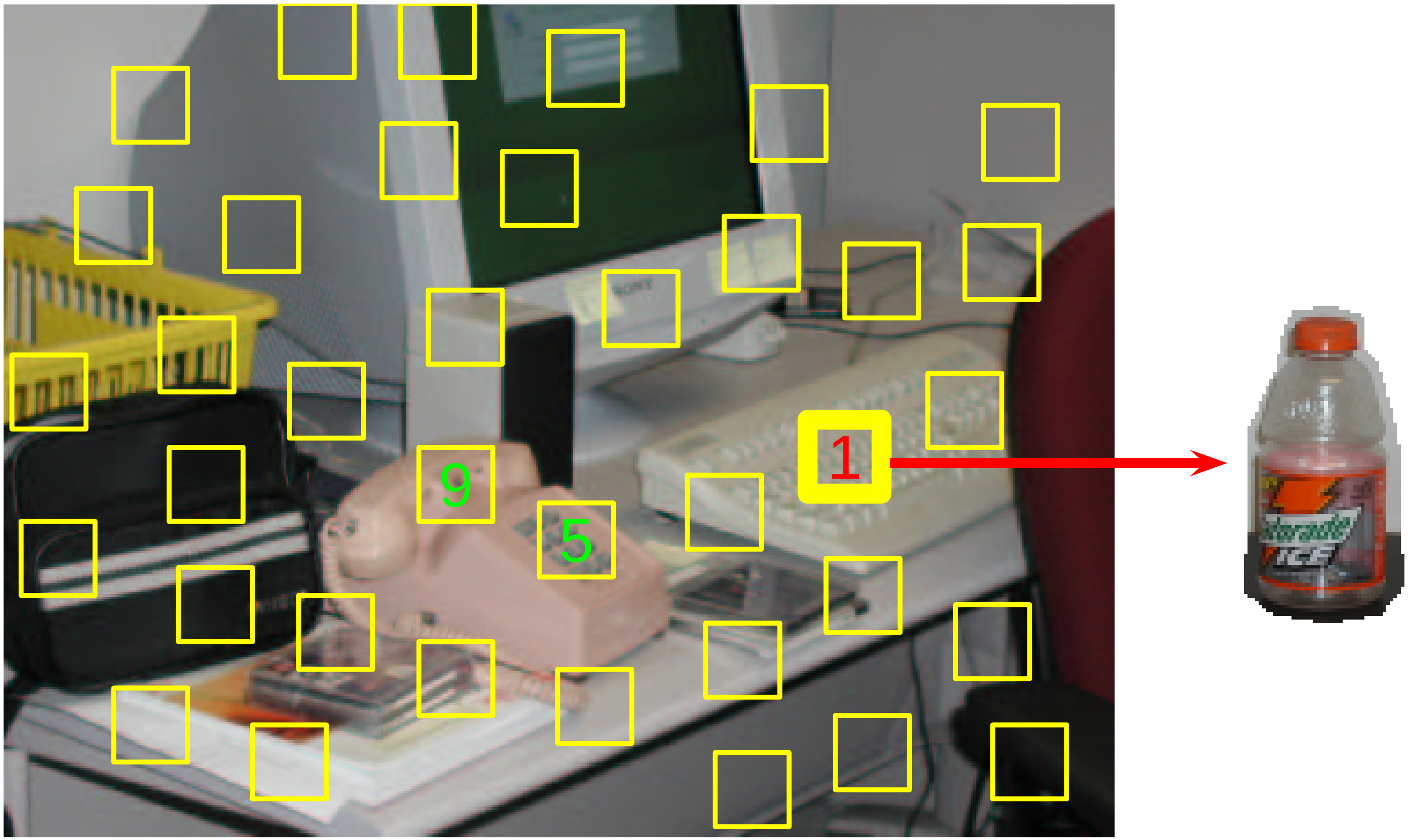
- one choice
- single feature
- content-free

Why Perceptual Organization

recognition

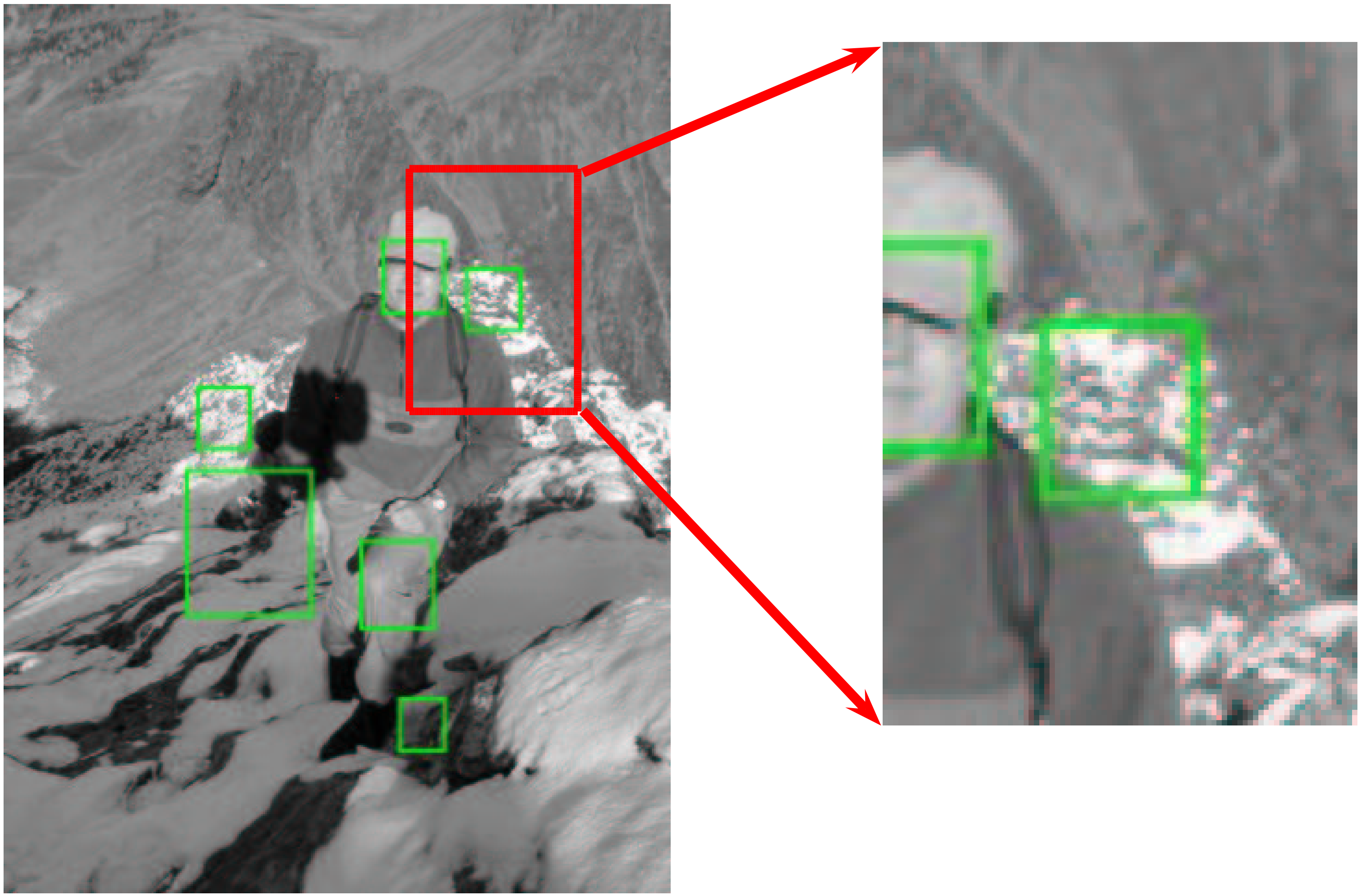
image

Why Perceptual Organization



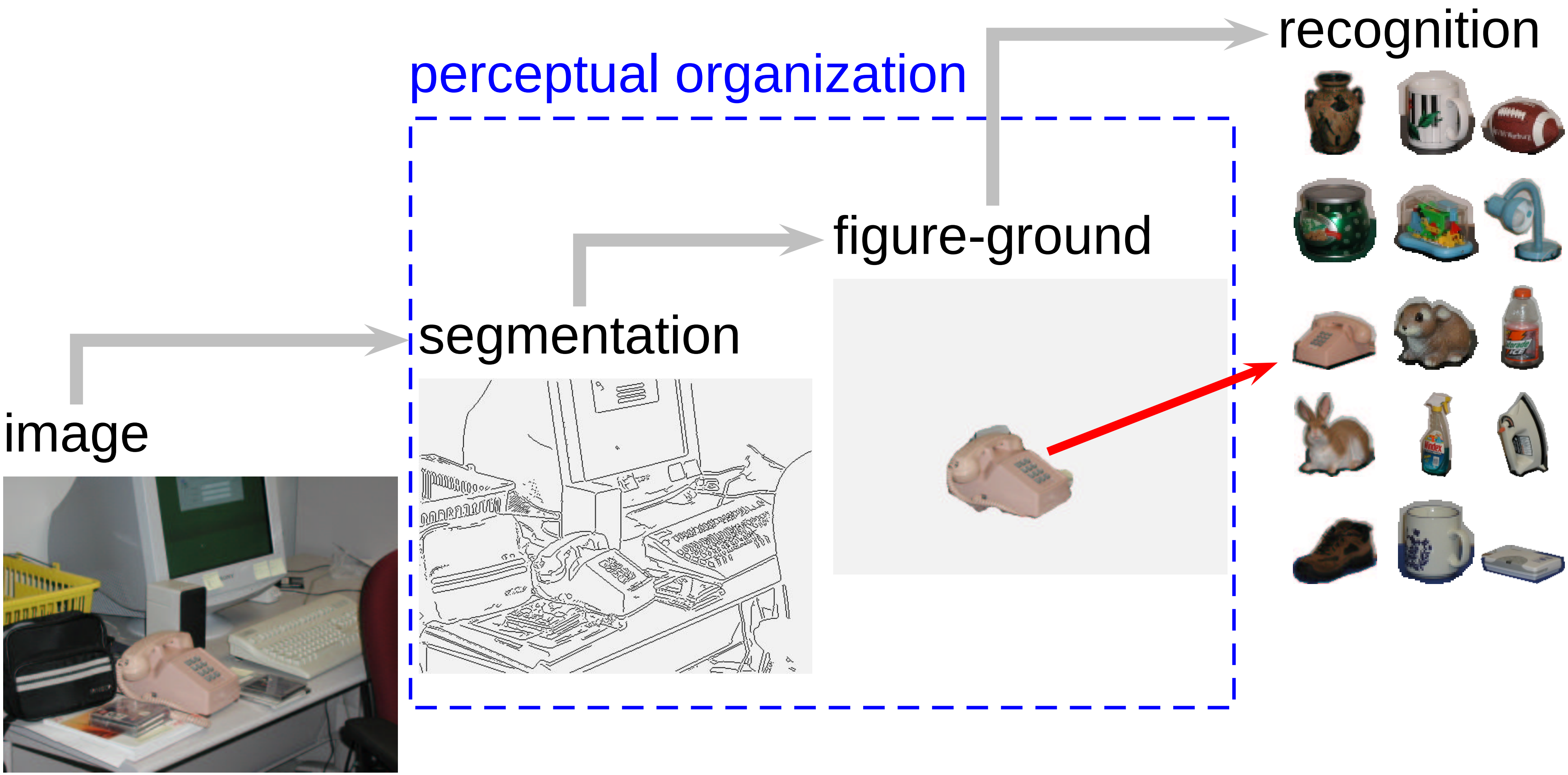
Mahamud multi-object detector

Why Perceptual Organization



Schneiderman face detector

Traditional Use of Perceptual Organization



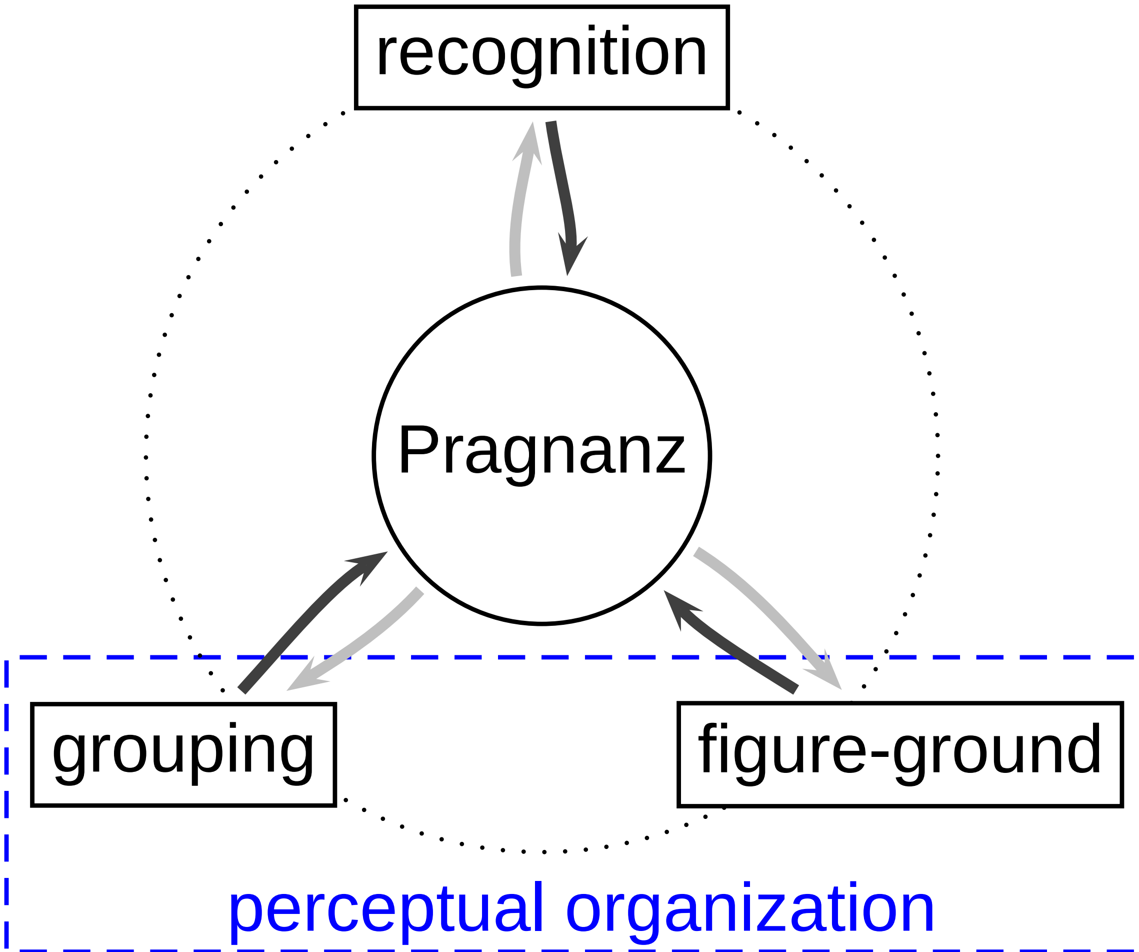
sequential processing (Marr, Lowe, Witkin, Tenenbaum, ...)

Perceptual Organization without Object Knowledge

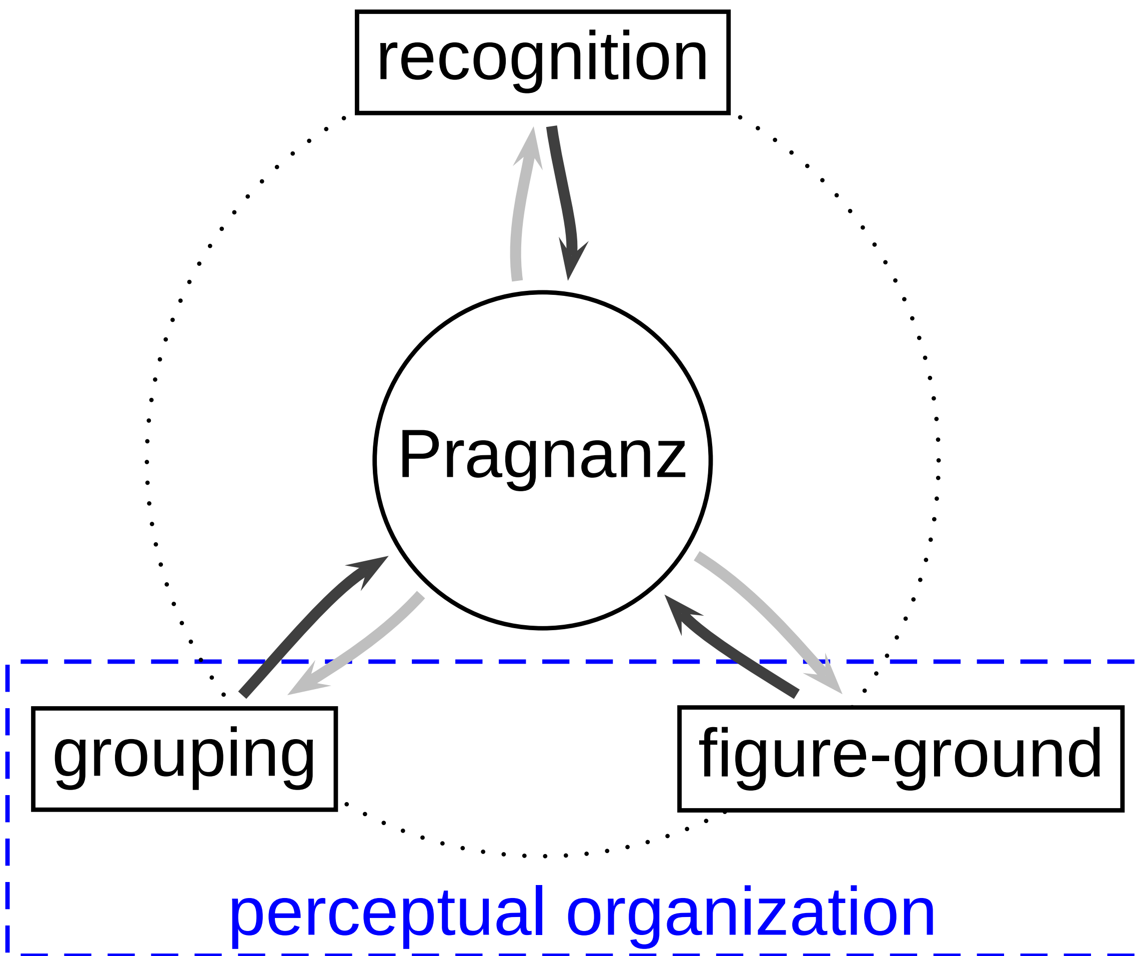
difficult and brittle

(Canny, Geman & Geman, Shah & Mumford, Witkin, Jacobs, ...)

Our Overall Approach

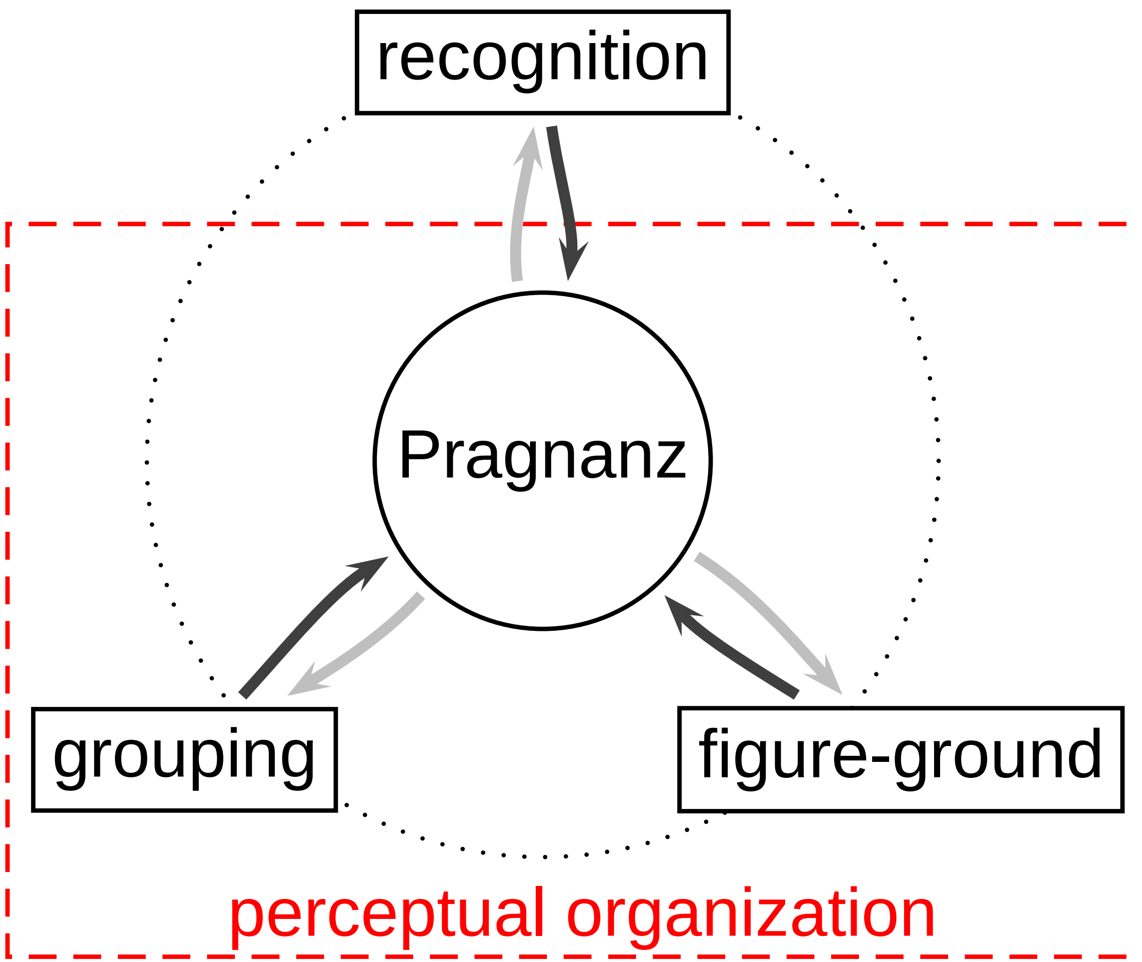


Our Overall Approach



interactive processing (Grossberg, McClelland,Grenandar,Mumford,Lee,...)

Our Overall Approach



A criterion

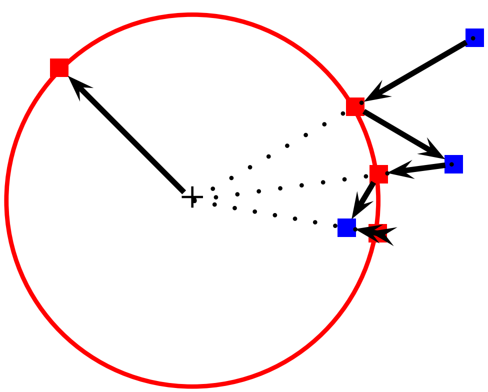
A fast solution

A wide range of images

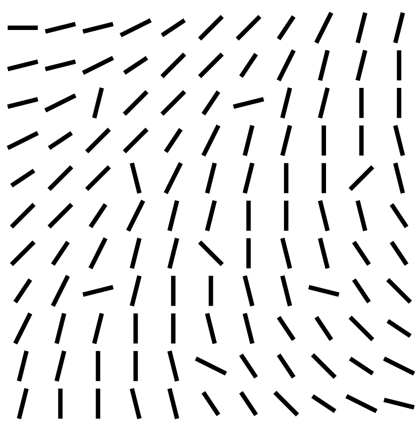
interactive processing (Grossberg, McClelland,Grenandar,Mumford,Lee,...)

Outline

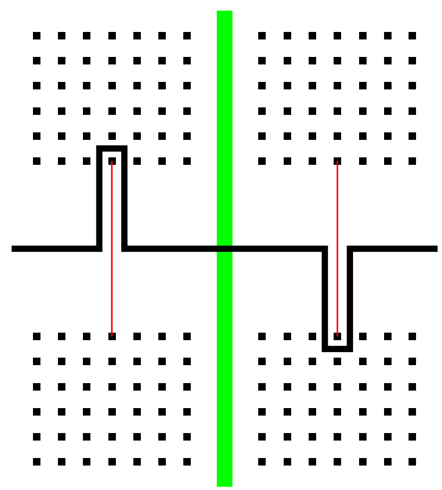
1. Computational framework: spectral clustering
2. Expand the repertoire of grouping cues: dissimilarity
3. Guide grouping with partial cues
4. Guide grouping with object knowledge
5. Summary and future work



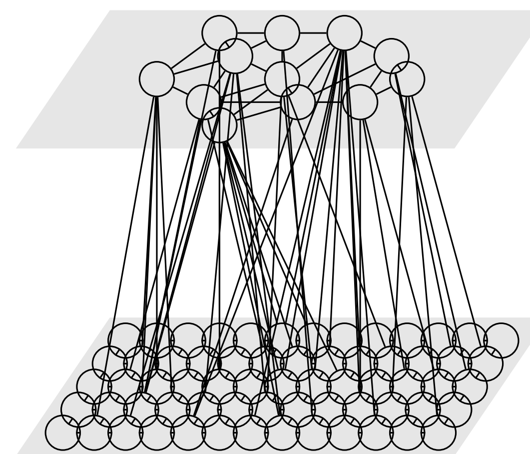
1



2

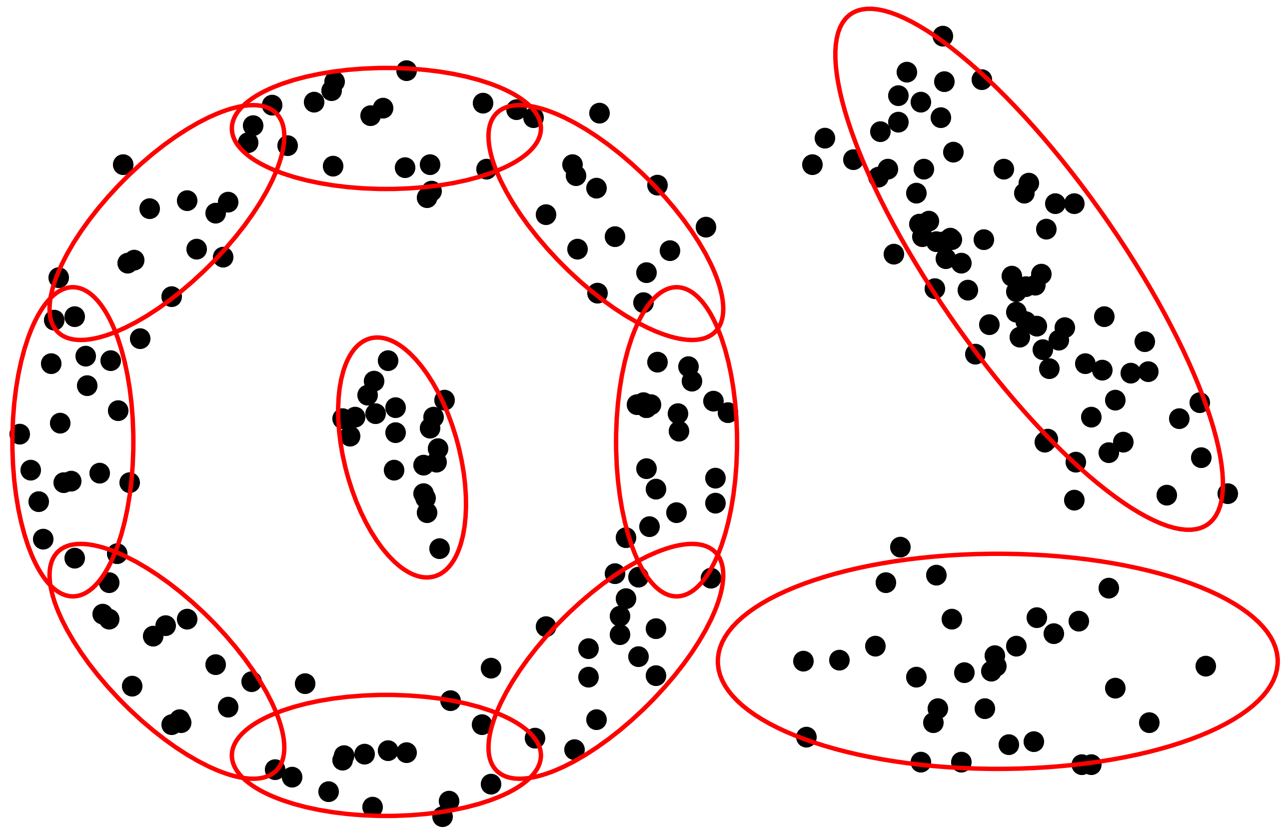


3



4

Generative Approach for Data Clustering

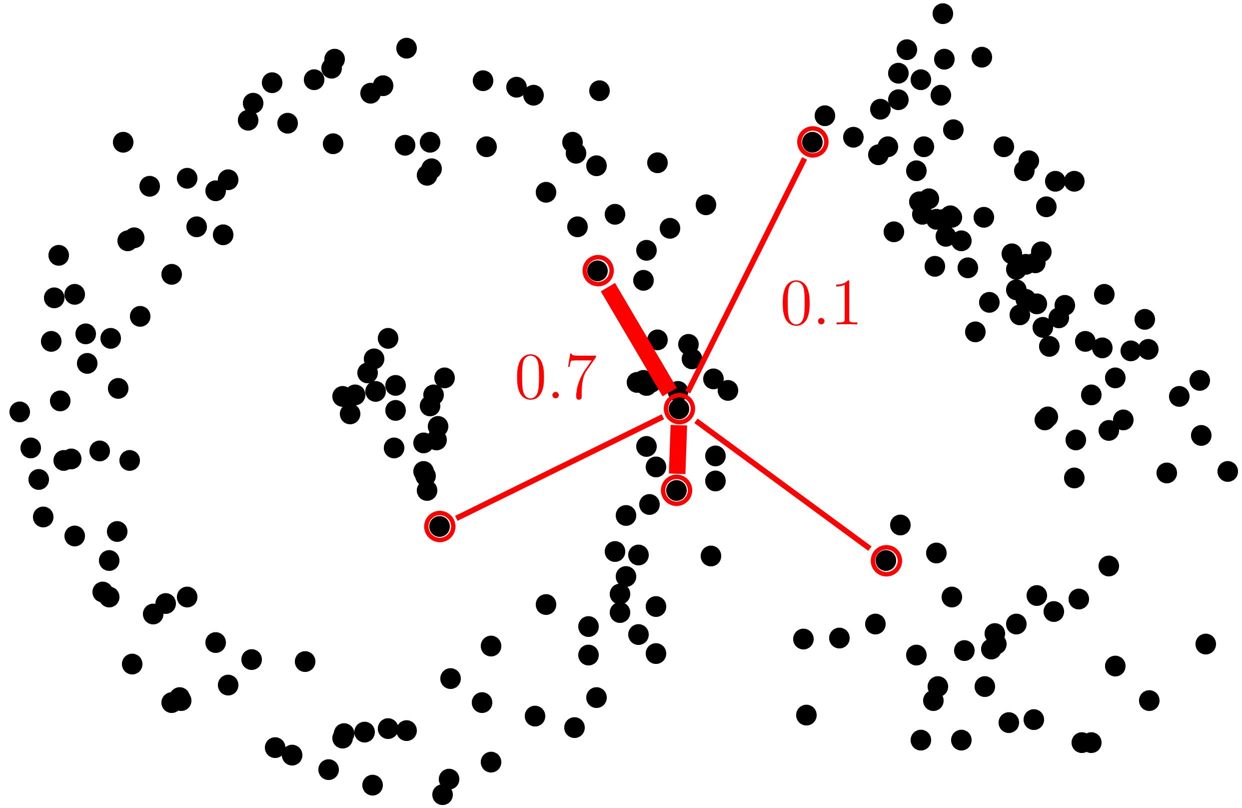


Key: Assumptions on the global structure of the data

Pros: Intuitive interpretation; analysis = synthesis

Cons: Model inadequacy and computational intractability

Discriminative Approach for Clustering

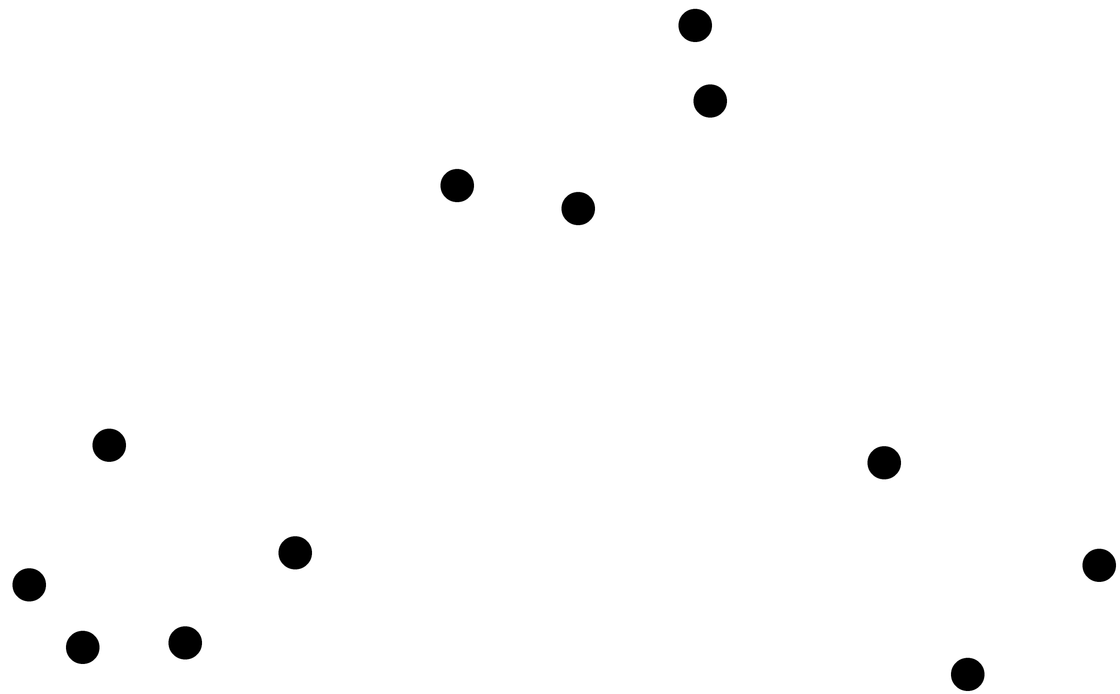


Key: Same group or not

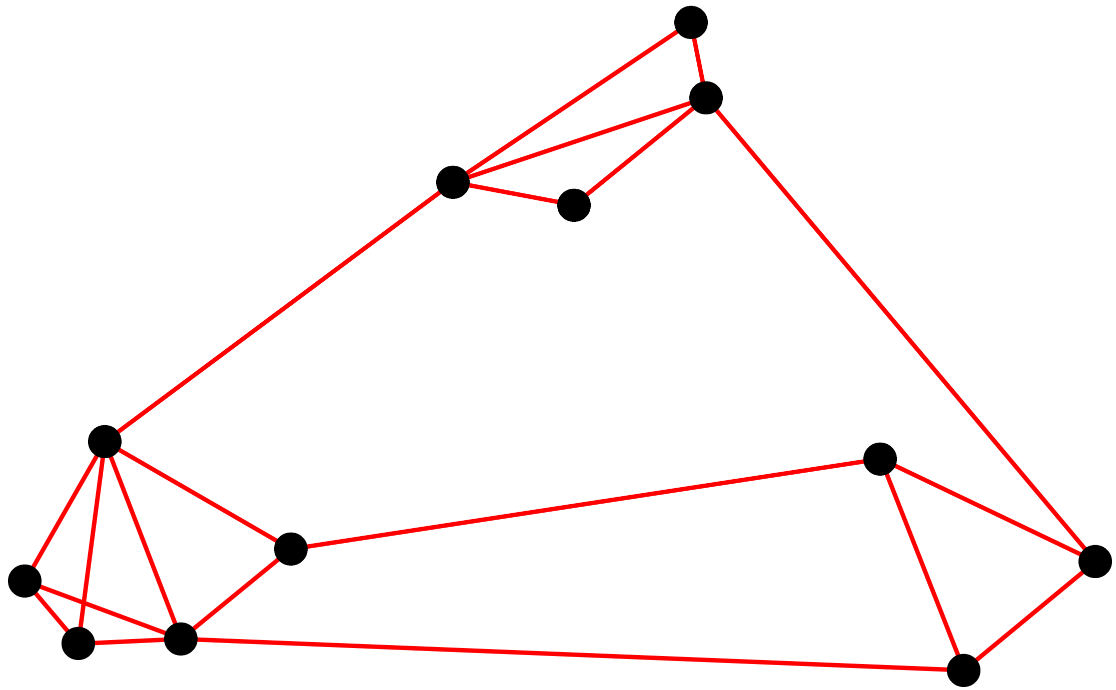
Pros: Adaptable to all data structures; tractable computation

Cons: No interpretation of the groups

Grouping in a Graph-Theoretic Framework

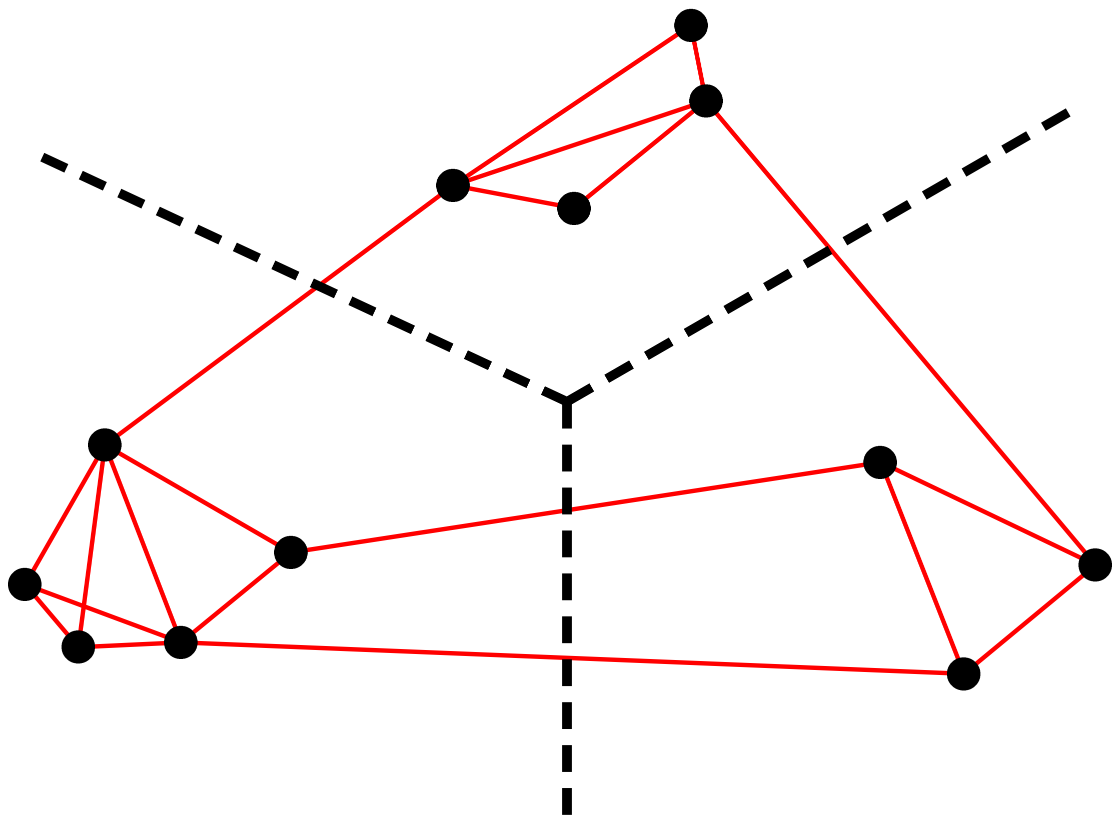


Grouping in a Graph-Theoretic Framework



Representation: $\mathbb{G} = \{\mathbb{V}, \mathbb{E}, W\} = \{$ nodes, edges, weights $\}$

Grouping in a Graph-Theoretic Framework

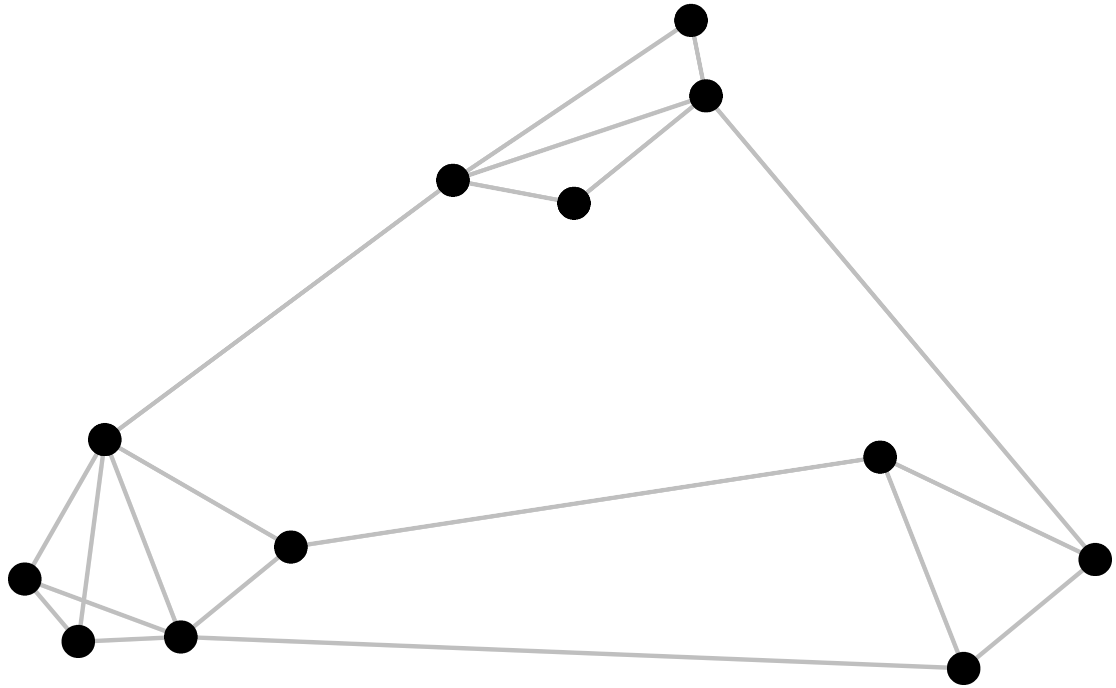


Representation: $\mathbb{G} = \{\mathbb{V}, \mathbb{E}, W\} = \{$ nodes, edges, weights $\}$

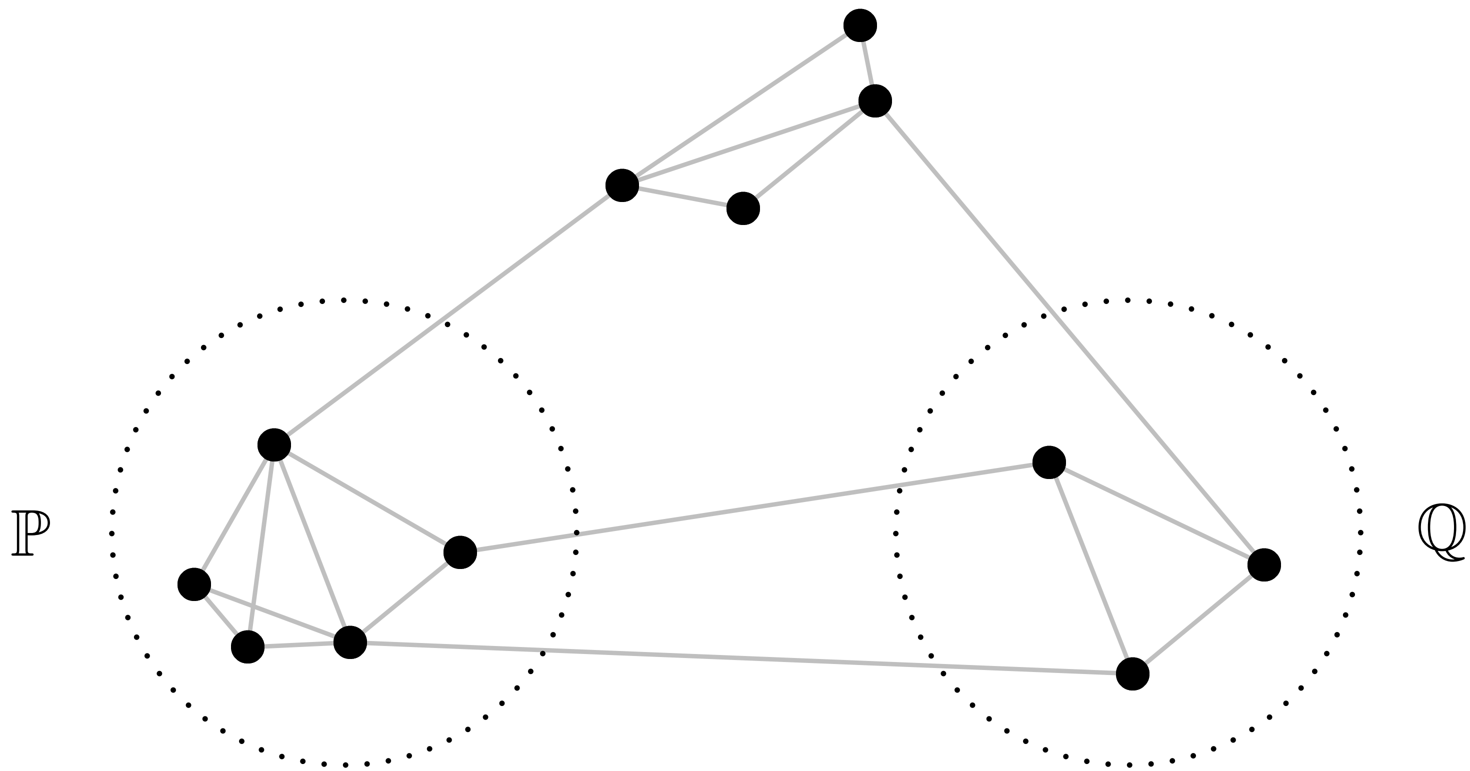
Clustering: $\Gamma_{\mathbb{V}}^{K} = \{\mathbb{V}_1, \ldots, \mathbb{V}_K\} = K$-way node partitioning

(Shi & Malik, Zabih, Boykov, Veksler, Kolmogorov,...)

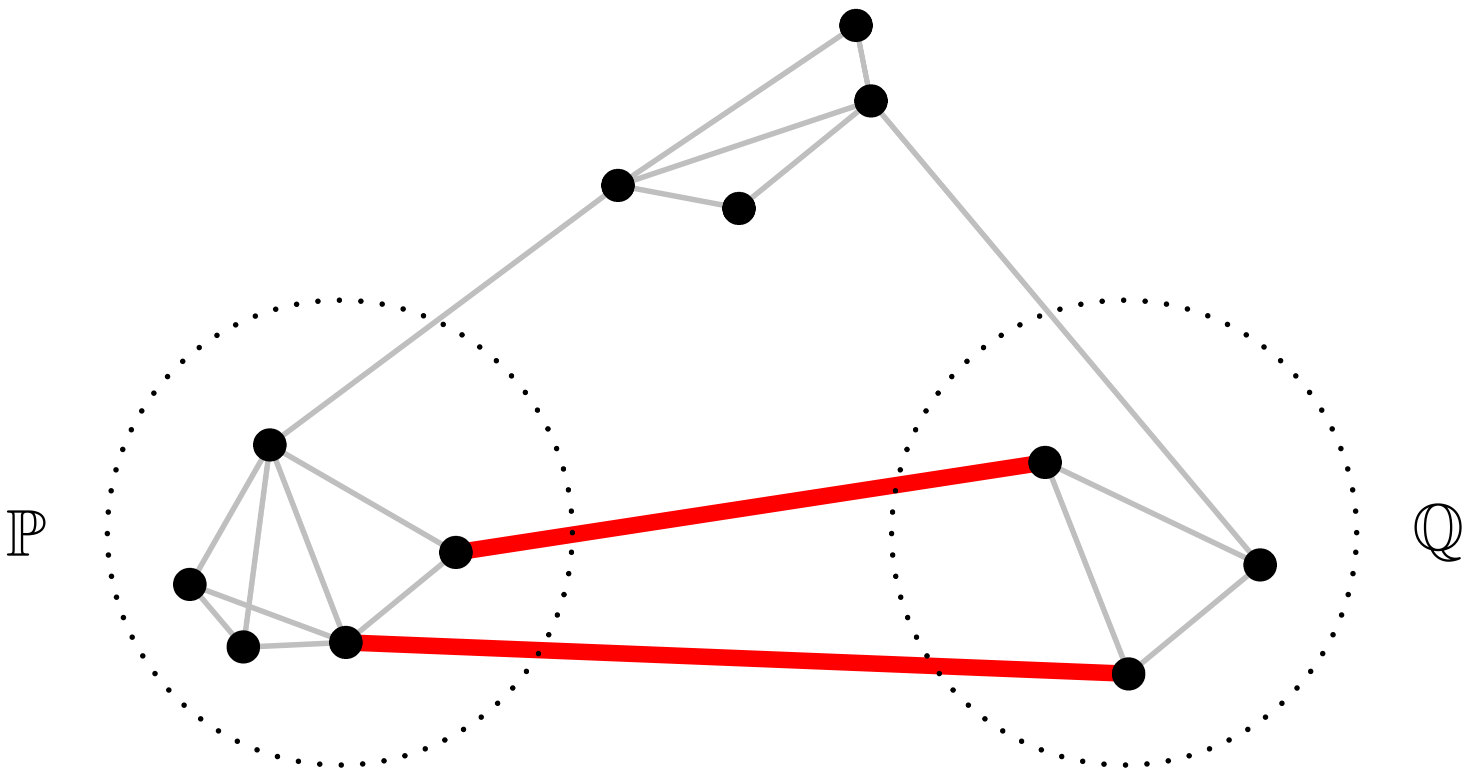
Links in Graph Cuts



Links in Graph Cuts

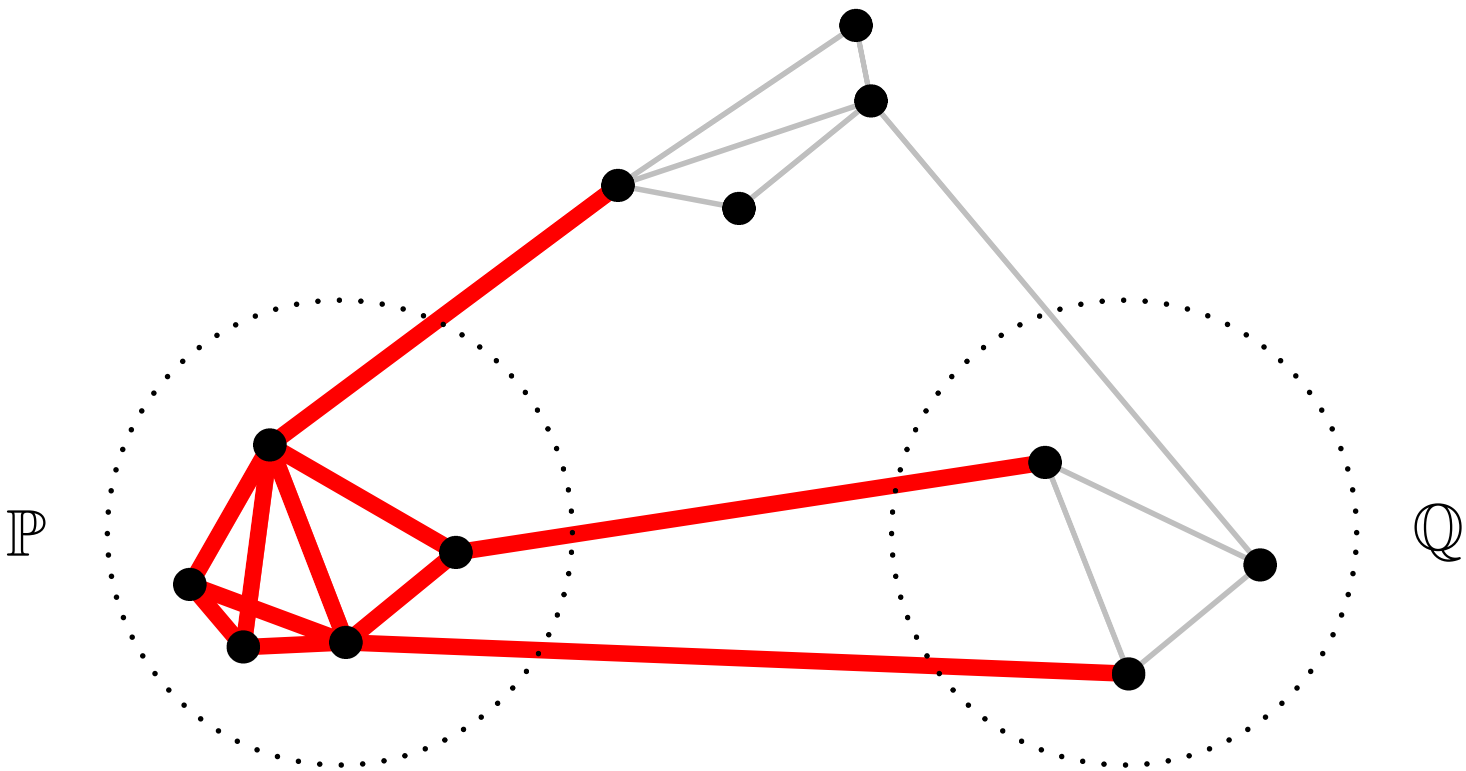


Links in Graph Cuts



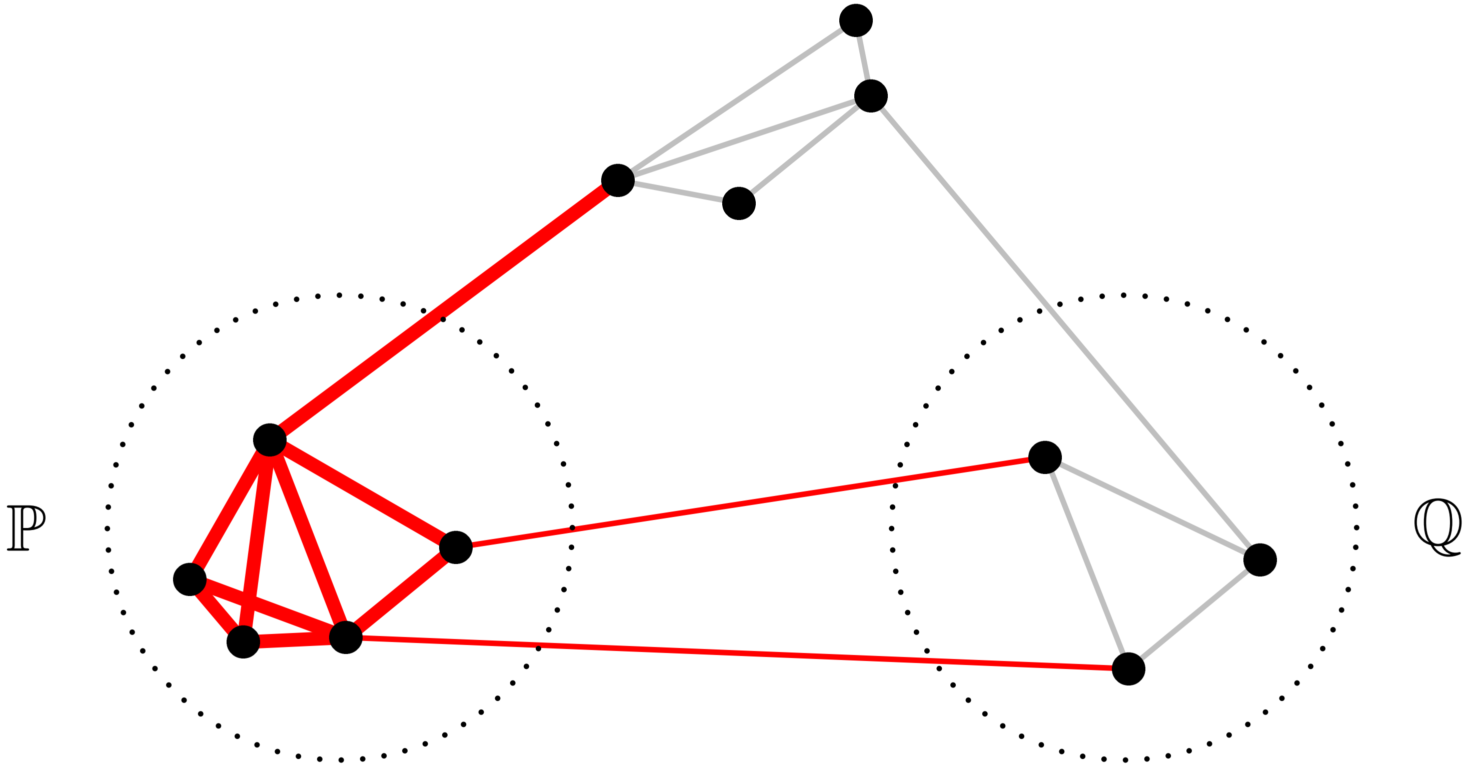
$$\text{links}(\mathbb{P}, \mathbb{Q}) = \sum_{p \in \mathbb{P},\, q \in \mathbb{Q}} W(p, q)$$

Degree in Graph Cuts



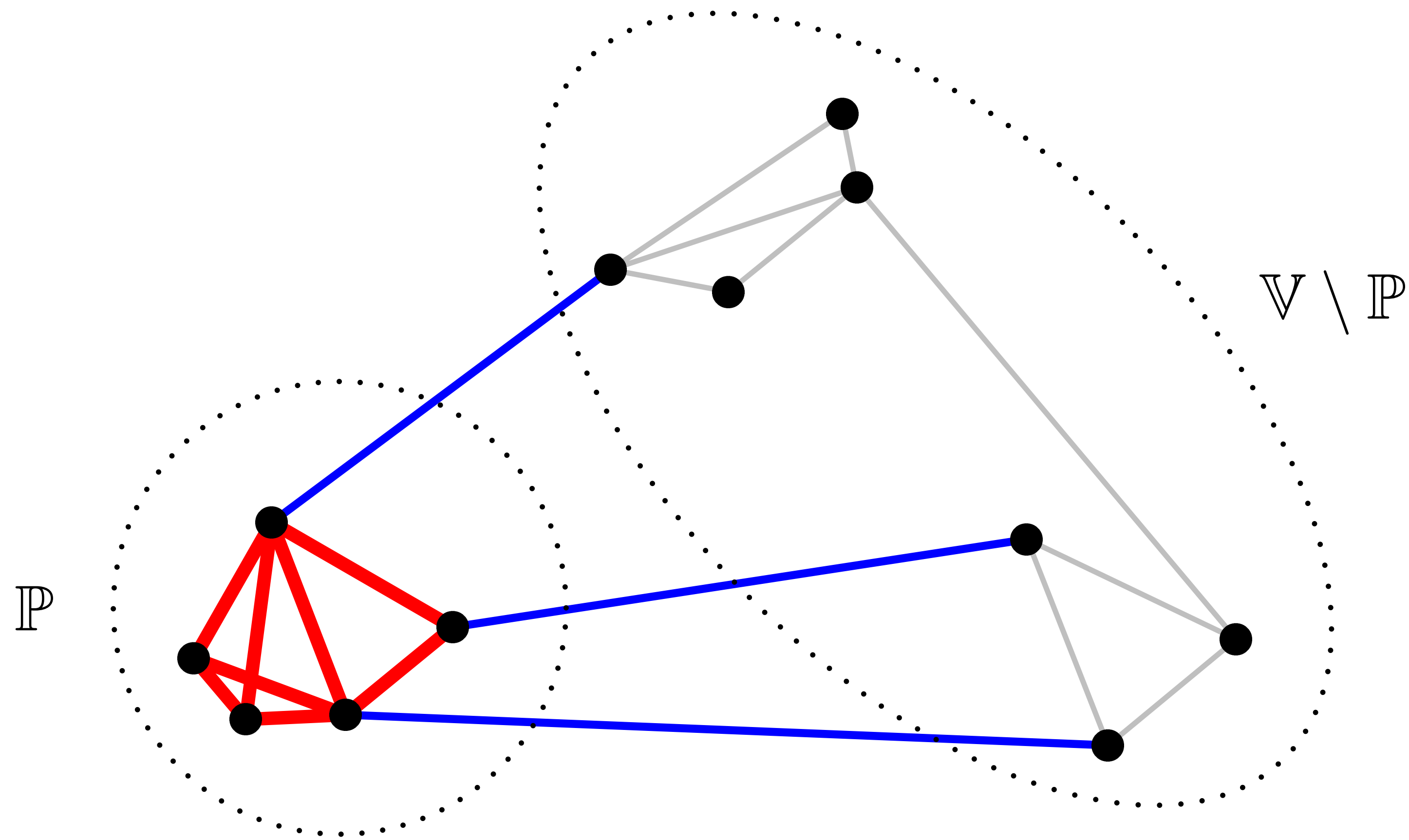
$$\text{degree}(\mathbb{P}) = \text{links}(\mathbb{P}, \mathbb{V})$$

Linkratio in Graph Cuts



$$\text{linkratio}(\mathbb{P}, \mathbb{Q}) = \frac{\text{links}(\mathbb{P}, \mathbb{Q})}{\text{degree}(\mathbb{P})}$$

Goodness of Grouping in Graph Cuts

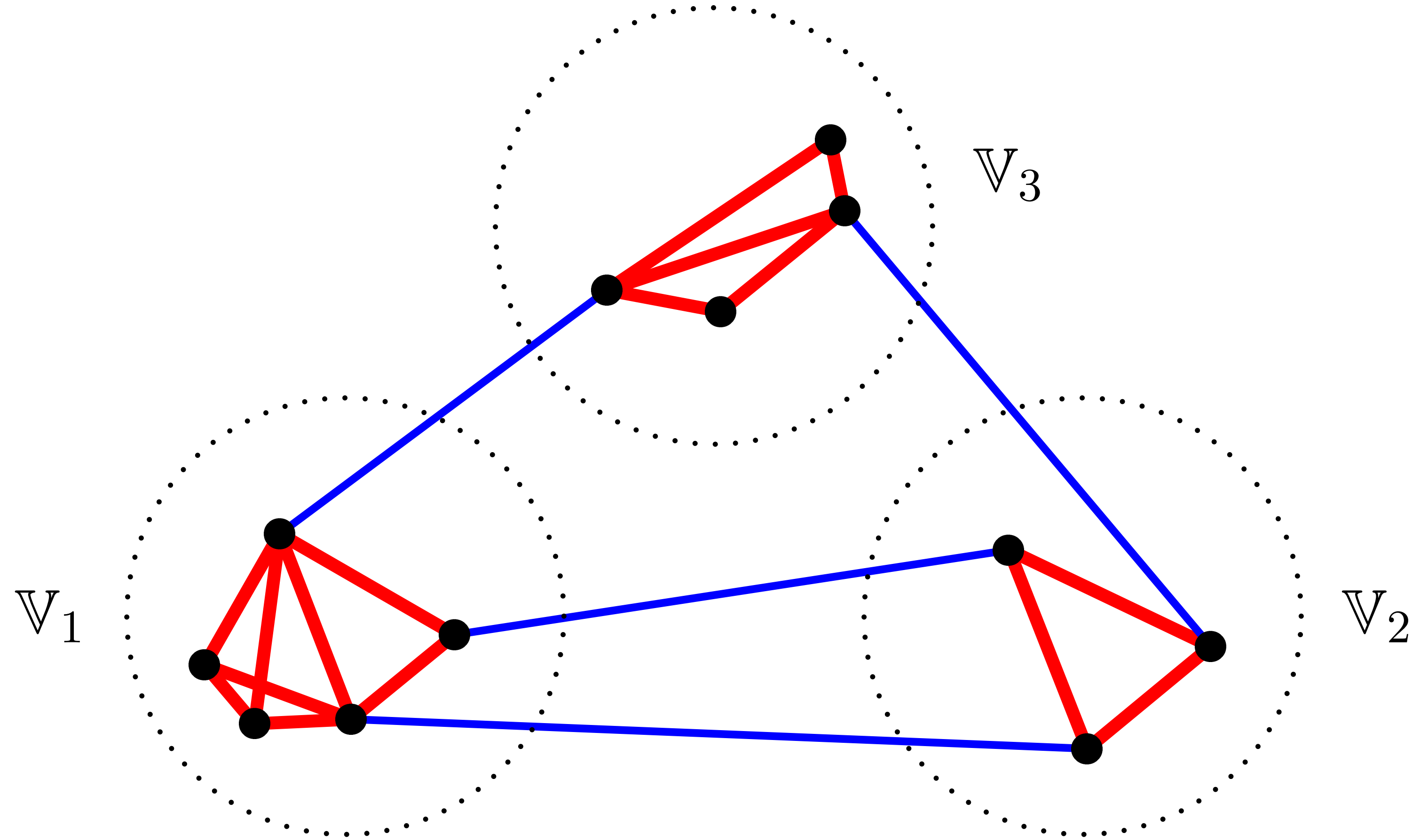


Maximize within-group connections: $\mathrm{linkratio}(\mathbb{P}, \mathbb{P})$

Minimize between-group connections: $\mathrm{linkratio}(\mathbb{P}, \mathbb{V} \setminus \mathbb{P})$

Equivalent: $\mathrm{linkratio}(\mathbb{P}, \mathbb{P}) + \mathrm{linkratio}(\mathbb{P}, \mathbb{V} \setminus \mathbb{P}) = 1$

K-Way Normalized Cuts



$$\max \quad \text{knassoc}(\Gamma_{\mathbb{V}}^{K}) = \frac{1}{K}\sum_{l=1}^{K} \text{linkratio}(\mathbb{V}_l, \mathbb{V}_l)$$

$$\min \quad \text{kncuts}(\Gamma_{\mathbb{V}}^{K}) = \frac{1}{K}\sum_{l=1}^{K} \text{linkratio}(\mathbb{V}_l, \mathbb{V} \setminus \mathbb{V}_l)$$

A Principled Solution to Normalized Cuts

$$\max \quad \mathrm{knassoc}(\Gamma_{\mathbb{V}}^{K}) = \frac{1}{K}\sum_{l=1}^{K} \mathrm{linkratio}(\mathbb{V}_l, \mathbb{V}_l)$$

A Principled Solution to Normalized Cuts

$$\max \quad \mathrm{knassoc}(\Gamma_{\mathbb{V}}^{K}) = \frac{1}{K} \sum_{l=1}^{K} \mathrm{linkratio}(\mathbb{V}_l, \mathbb{V}_l)$$

NP complete even for $K = 2$ and planar graphs

A Principled Solution to Normalized Cuts

$$\max \quad \operatorname{knassoc}(\Gamma_{\mathbb{V}}^{K}) = \frac{1}{K} \sum_{l=1}^{K} \operatorname{linkratio}(\mathbb{V}_l, \mathbb{V}_l)$$

NP complete even for $K = 2$ and planar graphs

Fast solution to find near-global optima:

A Principled Solution to Normalized Cuts

$$\max \quad \mathrm{knassoc}(\Gamma_{\mathbb{V}}^{K}) = \frac{1}{K}\sum_{l=1}^{K} \mathrm{linkratio}(\mathbb{V}_l, \mathbb{V}_l)$$

NP complete even for $K = 2$ and planar graphs

Fast solution to find near-global optima:

1. Find global optima in the relaxed continuous domain

   optima = eigenvectors of $(W, D) \times$ rotations

A Principled Solution to Normalized Cuts

$$\max \quad \mathrm{knassoc}(\Gamma_{\mathbb{V}}^{K}) = \frac{1}{K}\sum_{l=1}^{K} \mathrm{linkratio}(\mathbb{V}_l, \mathbb{V}_l)$$

NP complete even for $K = 2$ and planar graphs

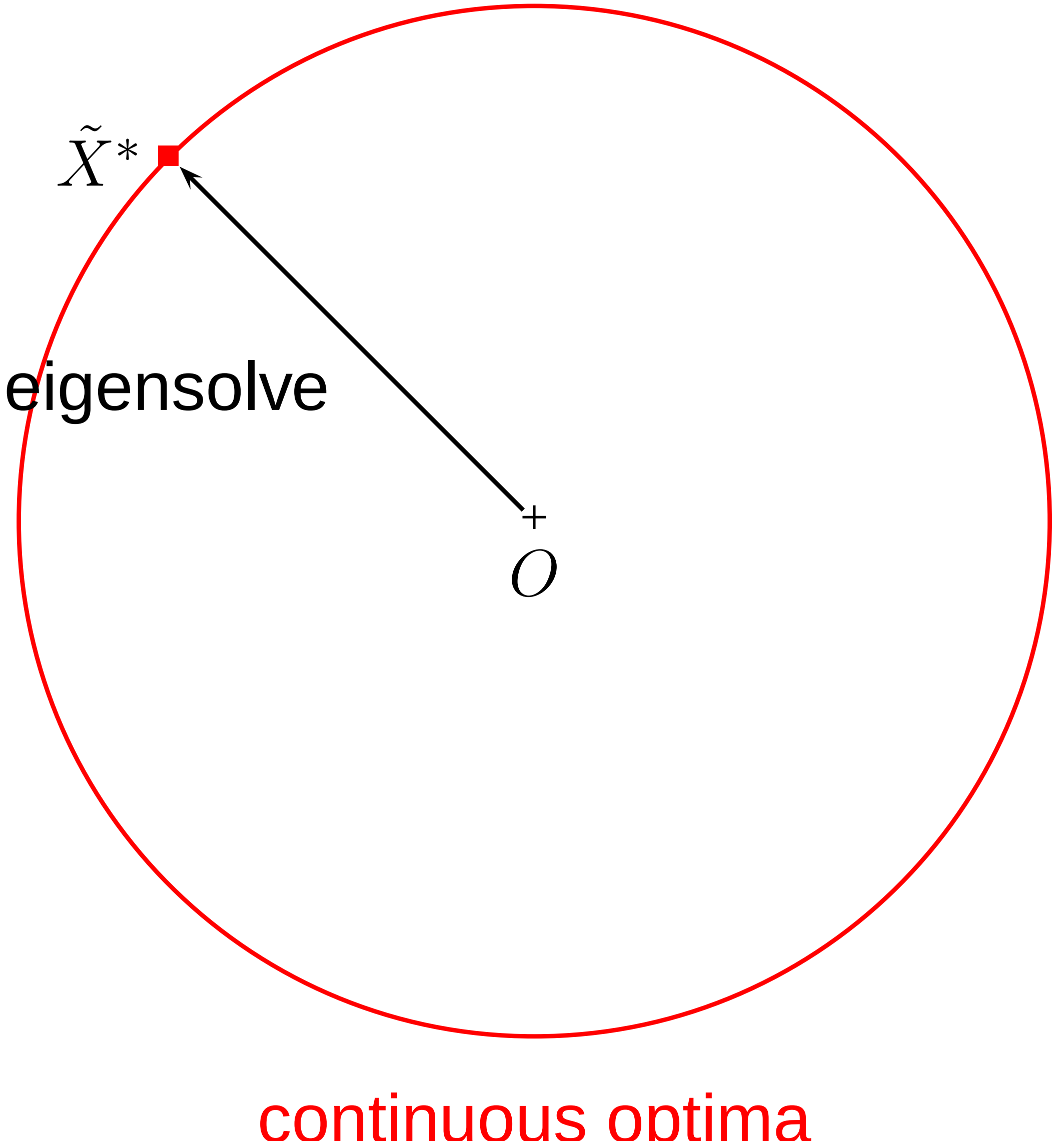
Fast solution to find near-global optima:

1. Find global optima in the relaxed continuous domain
   optima = eigenvectors of $(W, D) \times$ rotations
2. Find a discrete solution closest to continuous optima
   closeness = measured in L_2 norm between solutions

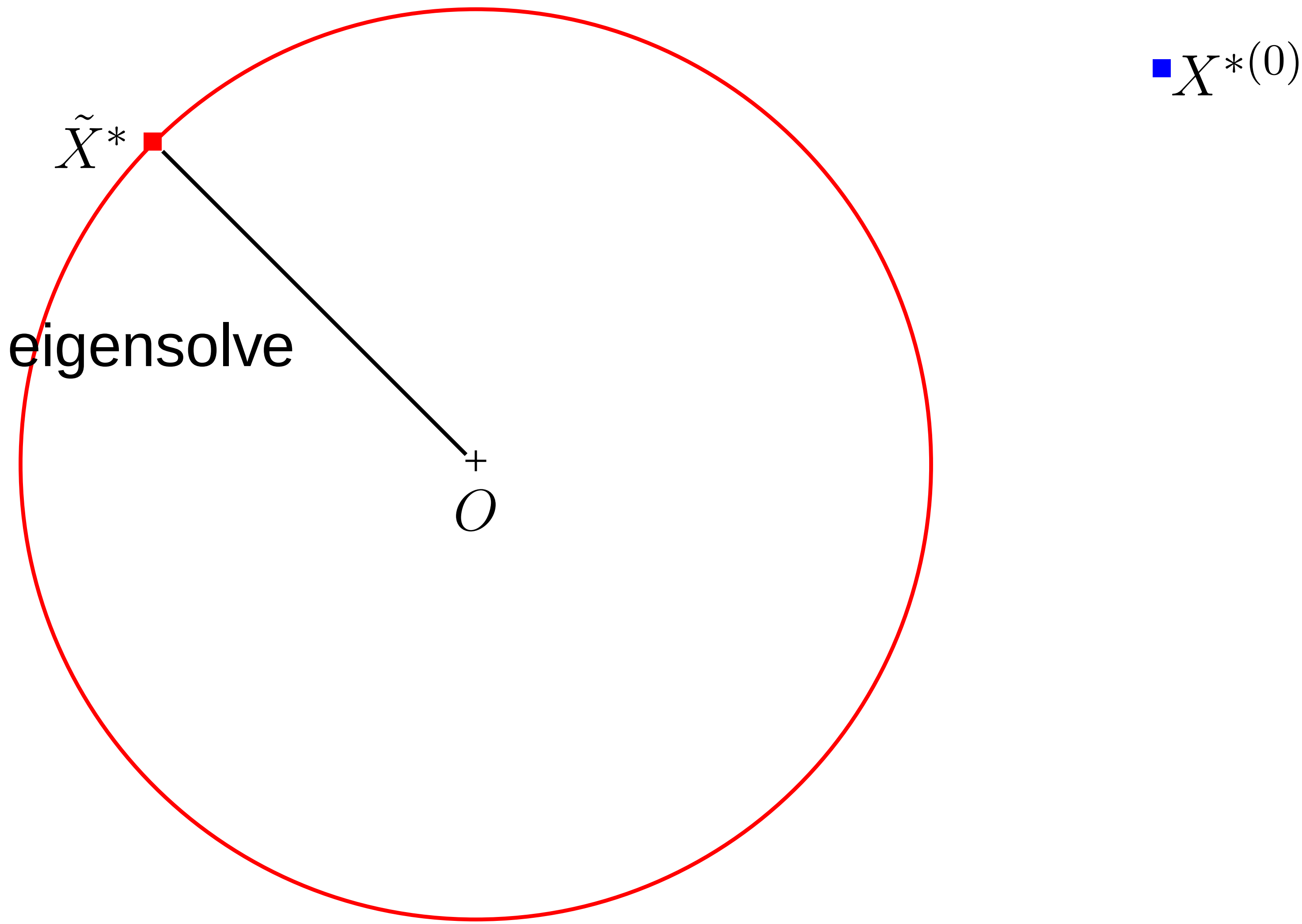
Step 1: Find Continuous Global Optima

$$\overset{+}{O}$$

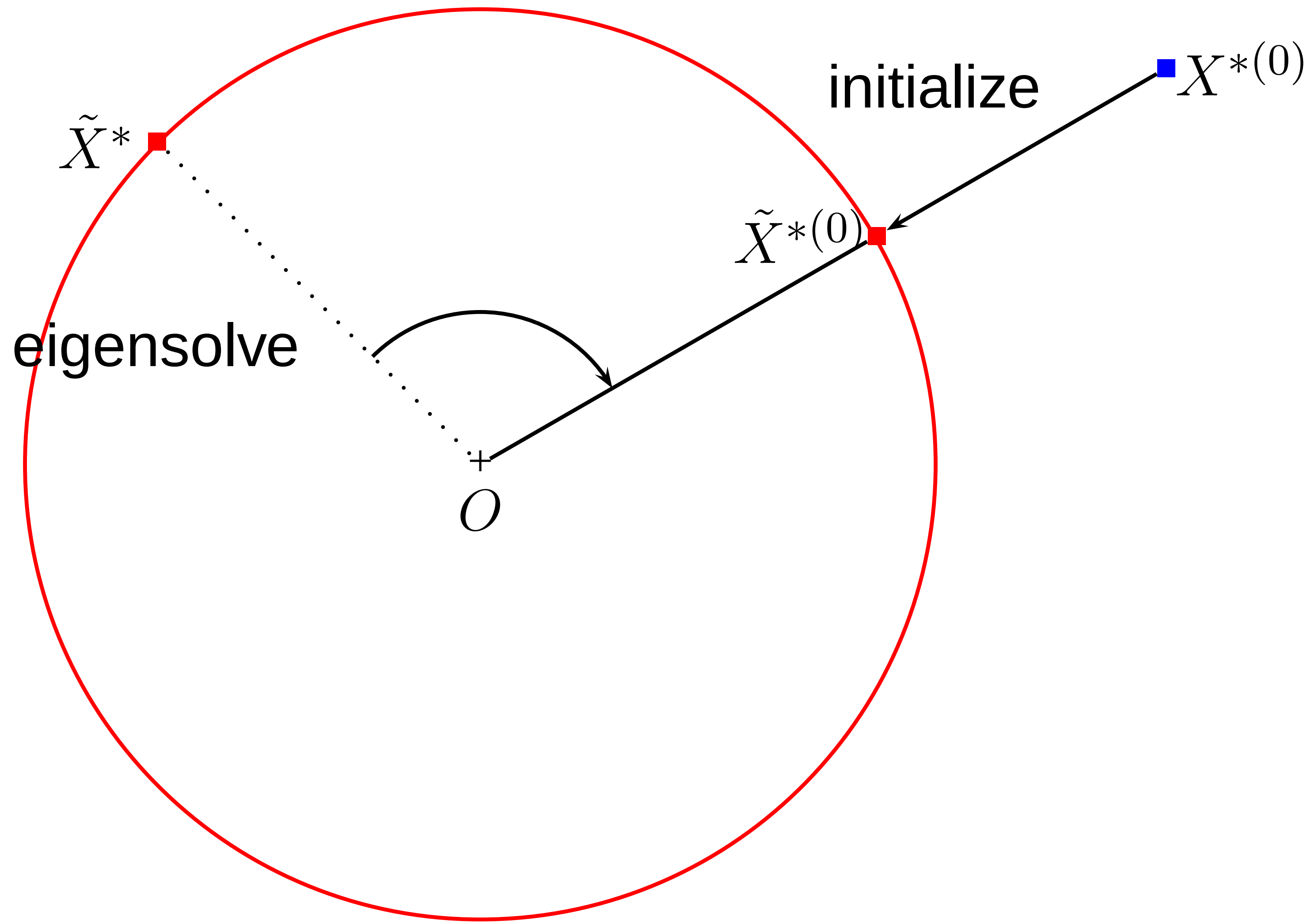
Step 1: Find Continuous Global Optima



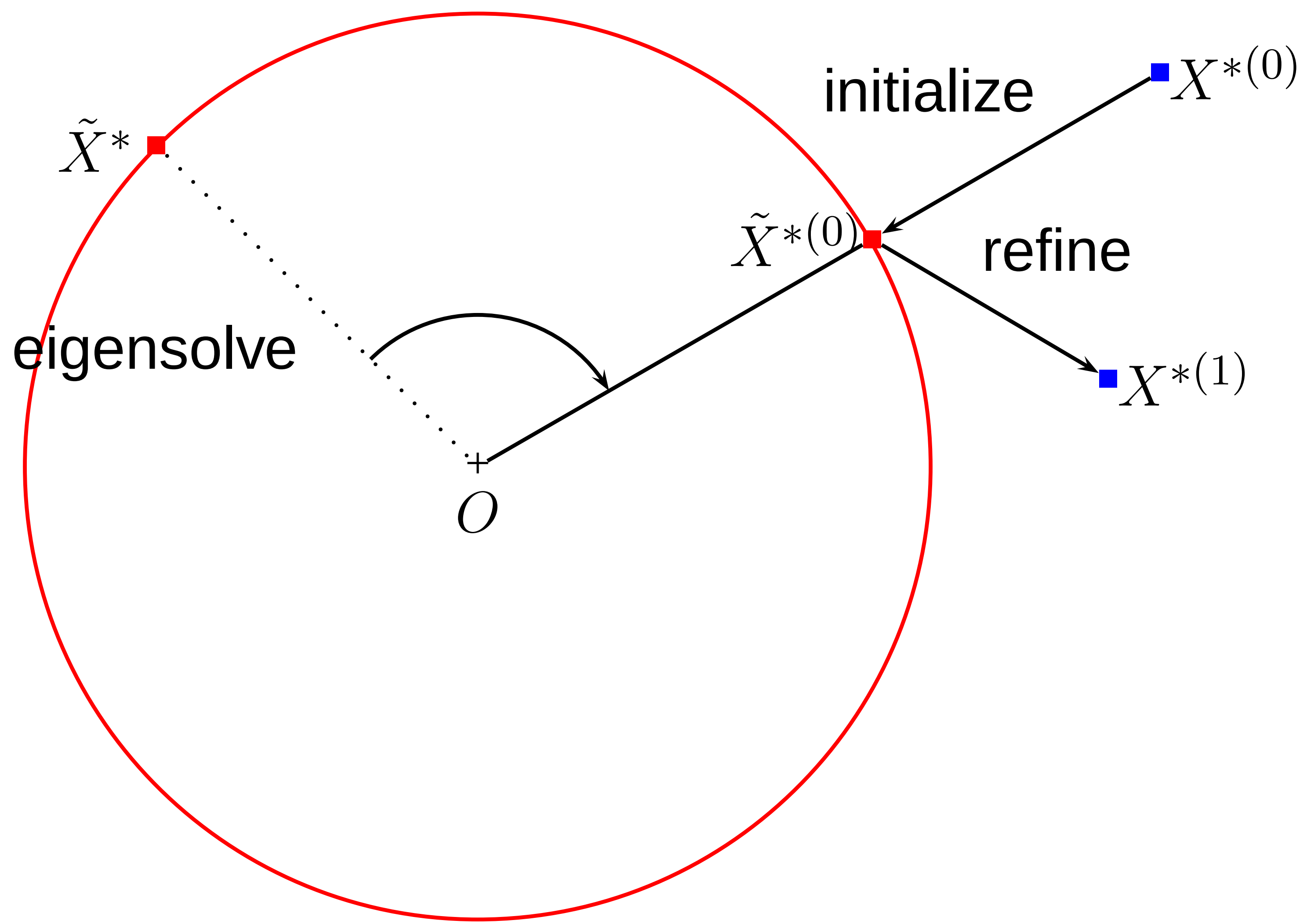
Step 2: Discretize Continuous Optima



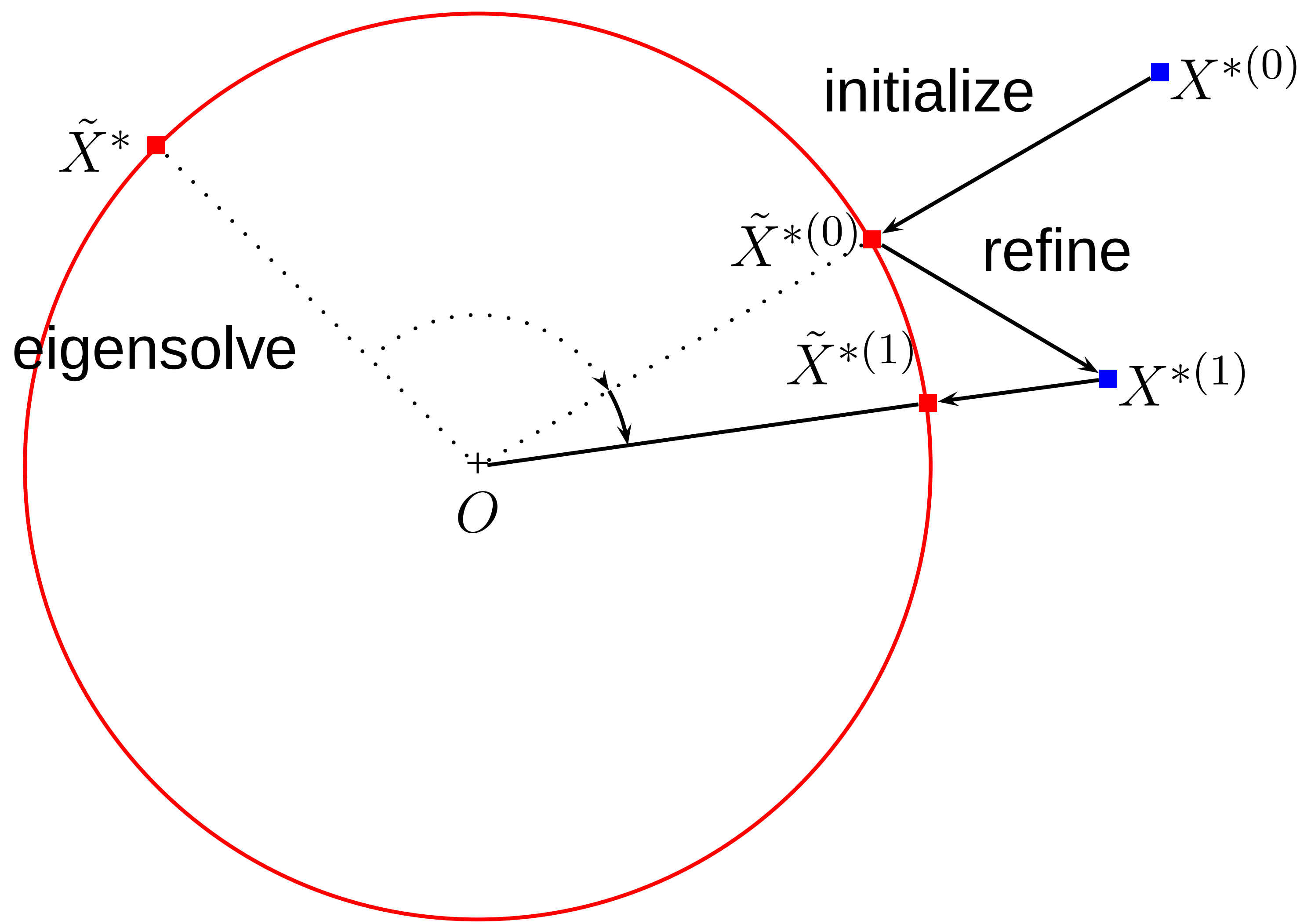
Step 2: Discretize Continuous Optima



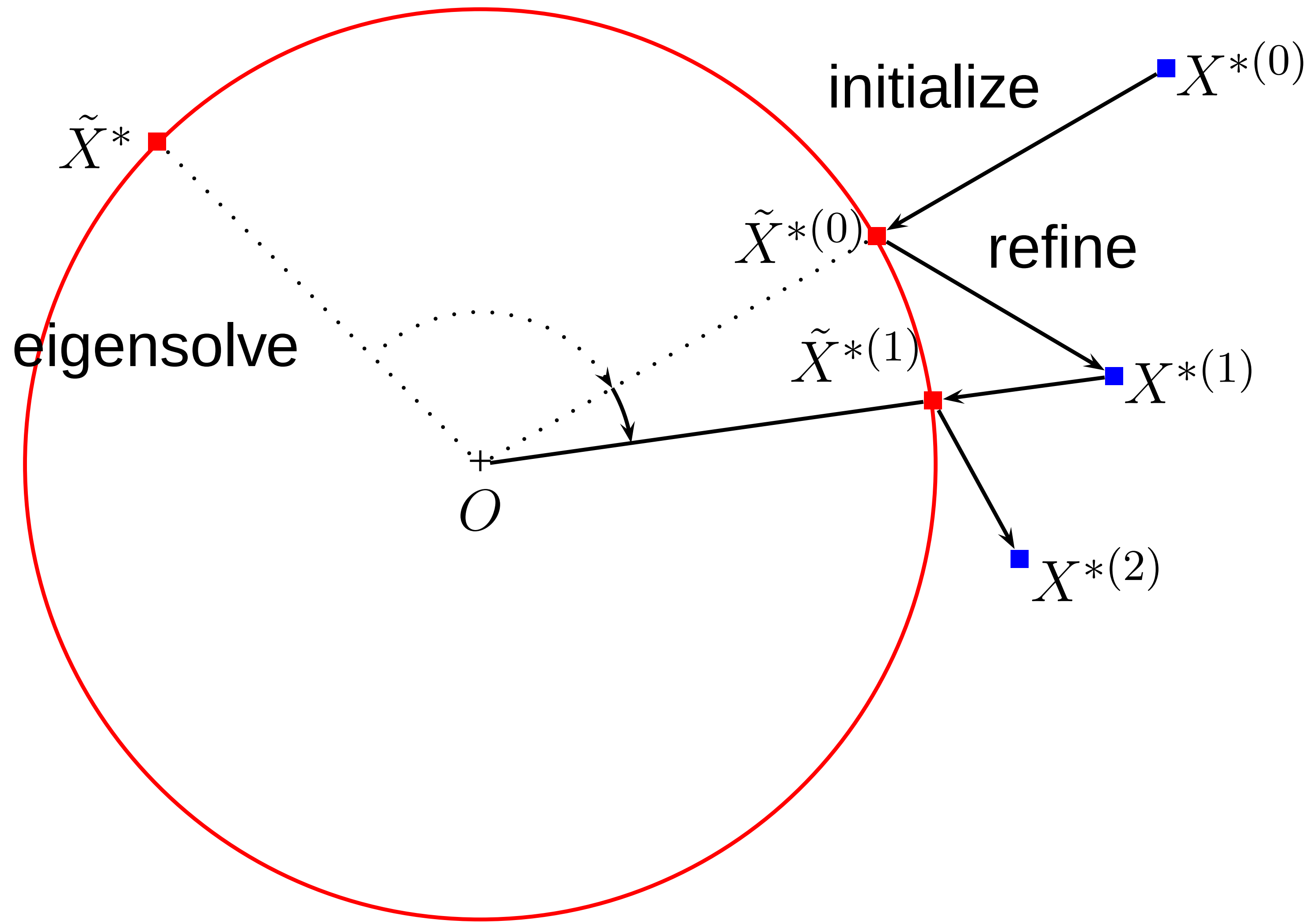
Step 2: Discretize Continuous Optima



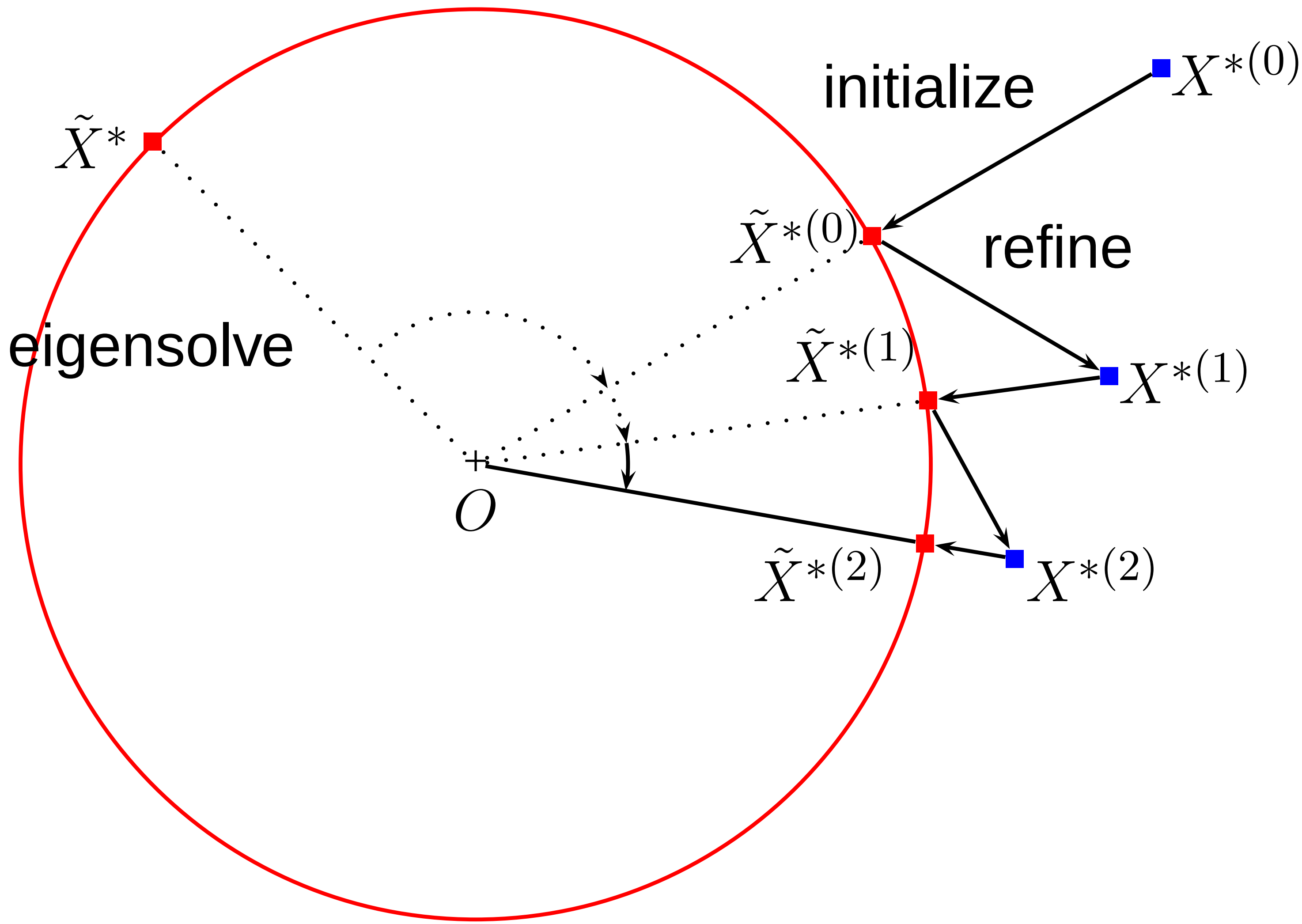
Step 2: Discretize Continuous Optima



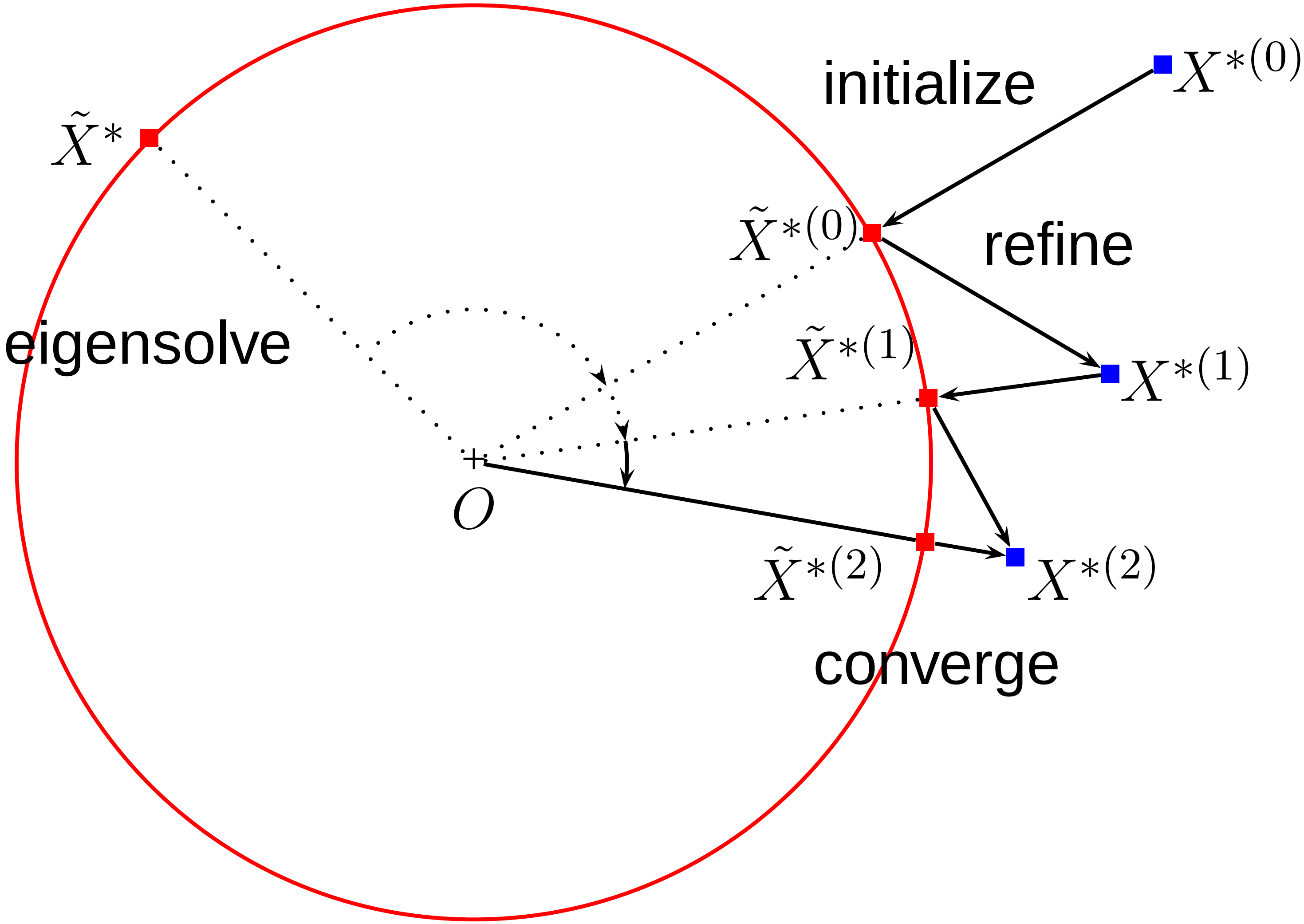
Step 2: Discretize Continuous Optima



Step 2: Discretize Continuous Optima



Step 2: Discretize Continuous Optima



Final solution: $(X^{*(2)}, \tilde{X}^{*(2)})$

Pixel Similarity based on Intensity Edges

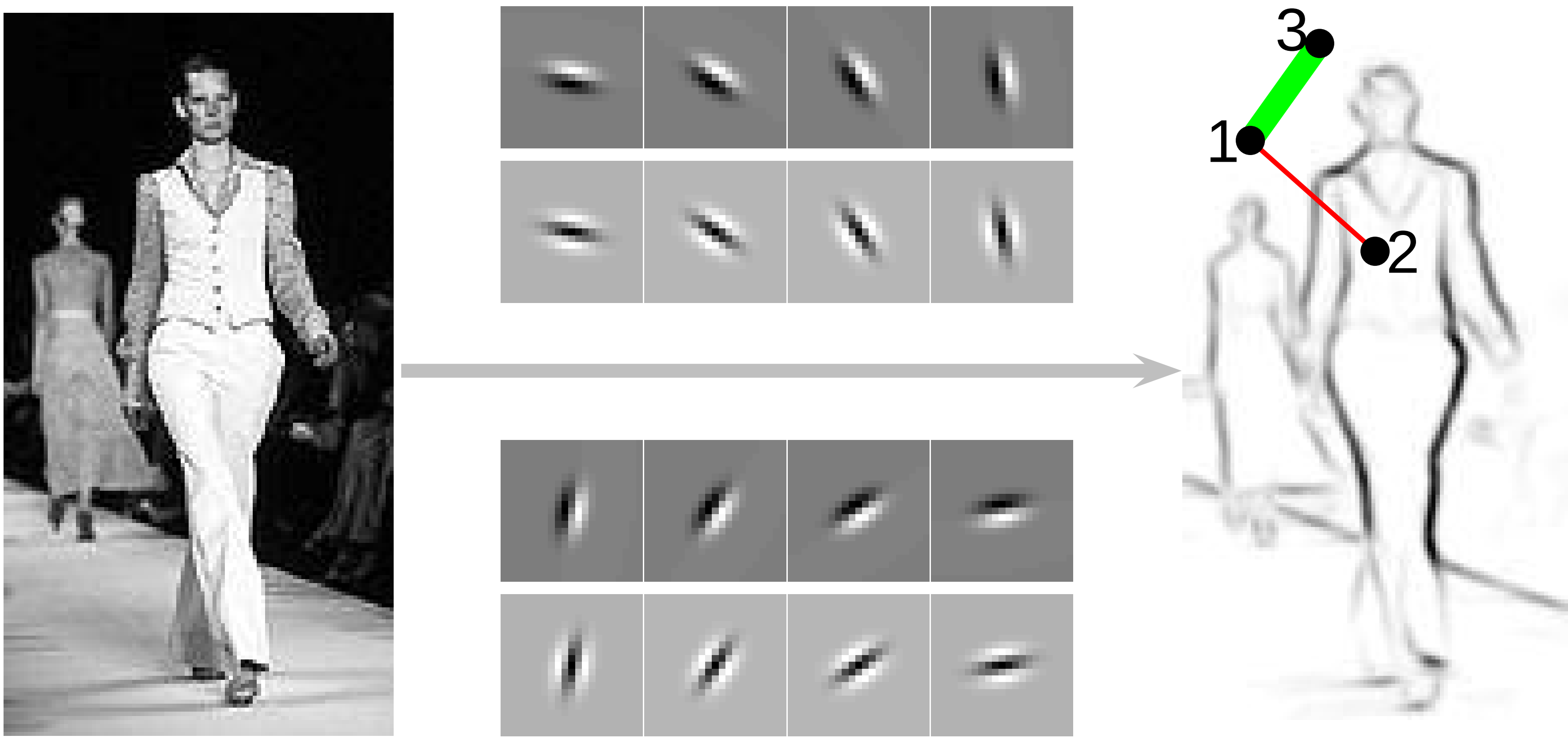


image oriented filter pairs edge magnitudes

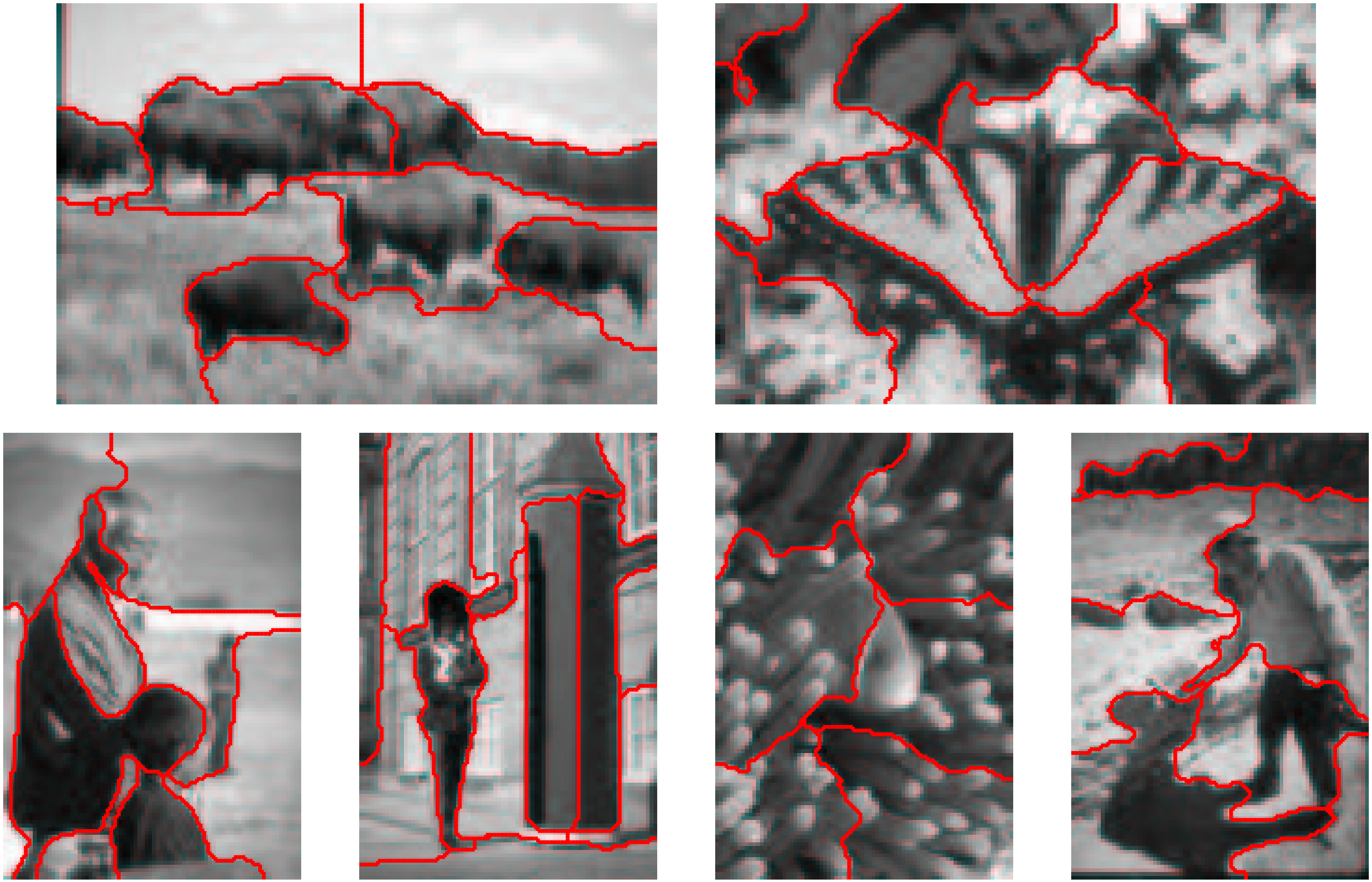
Discrete Optima Generated by Eigenvectors

Discrete Optima Generated by Eigenvectors

$K = 4:\quad 0.9901 \quad 0.9899 \quad 0.9881$

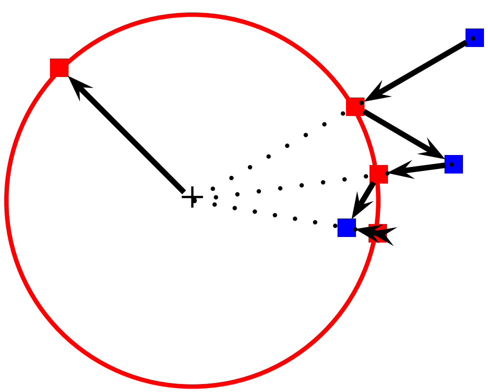
Not many local discrete optima, all good quality

Multiclass Real Image Segmentation

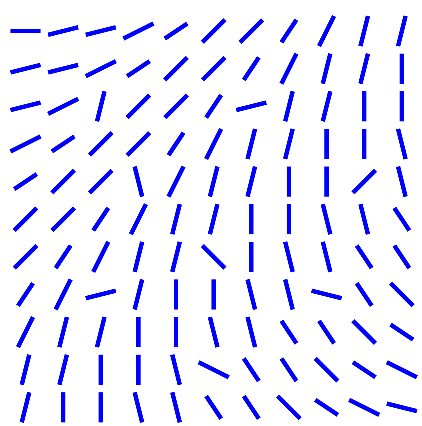


Outline

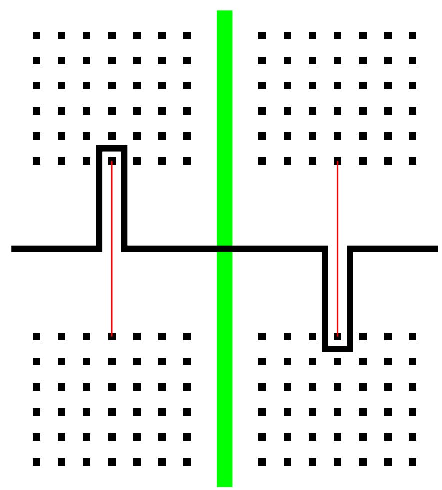
1. Computational framework: spectral clustering
2. Expand the repertoire of grouping cues: dissimilarity
3. Guide grouping with partial cues
4. Guide grouping with object knowledge
5. Summary and future work



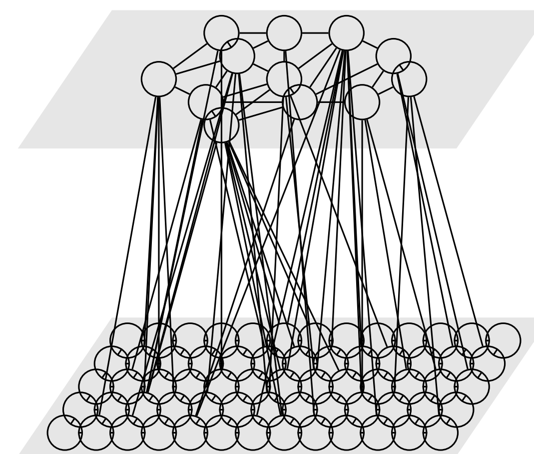
1



2

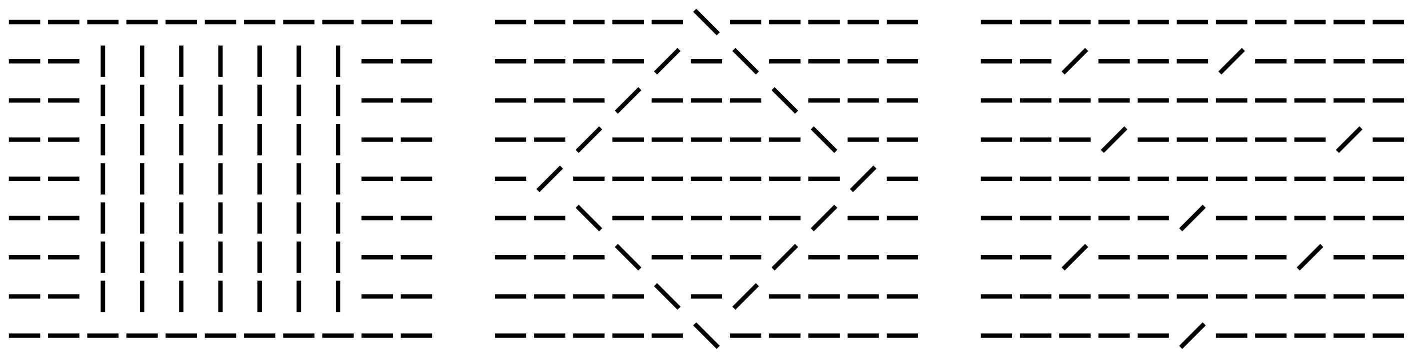


3

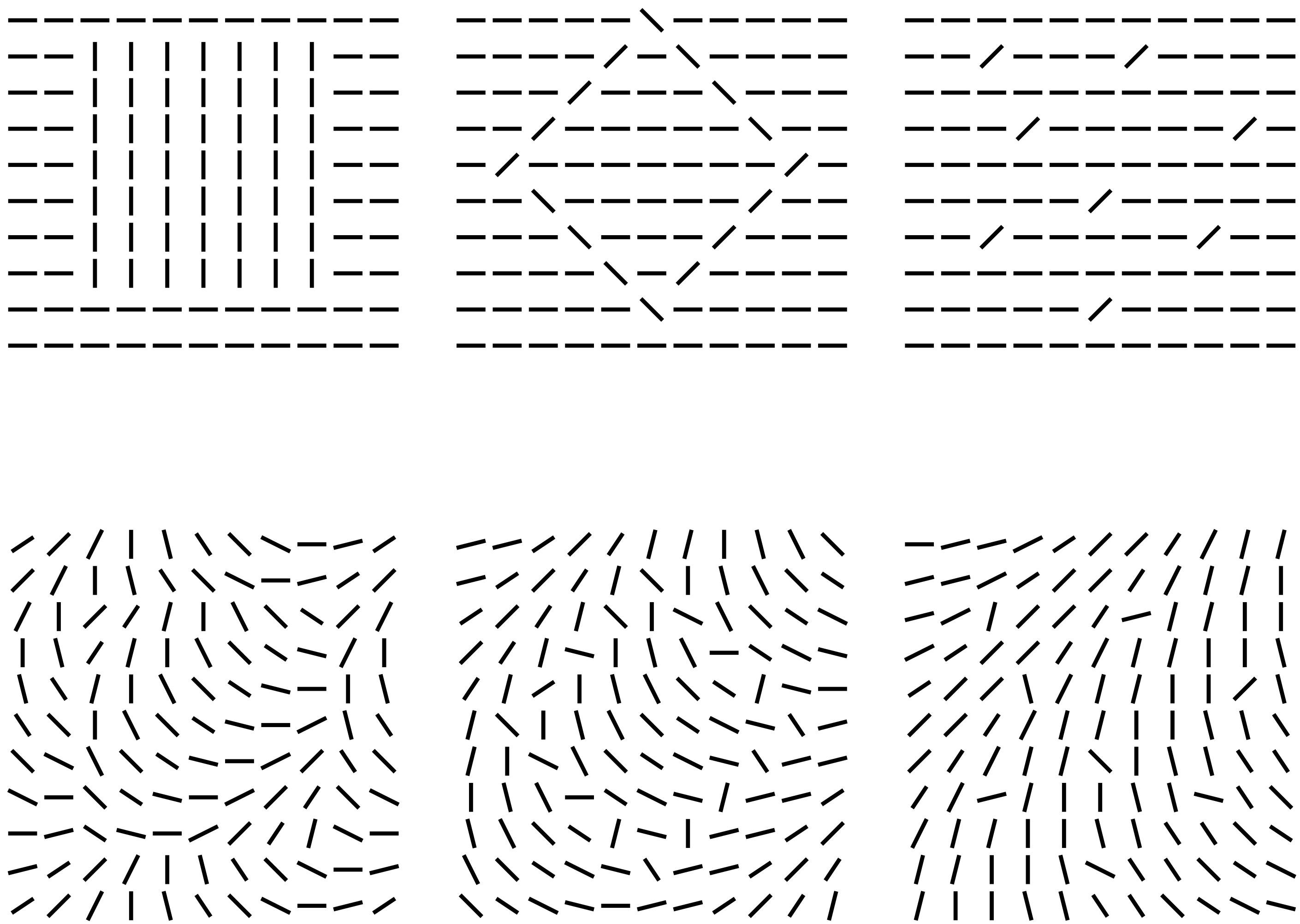


4

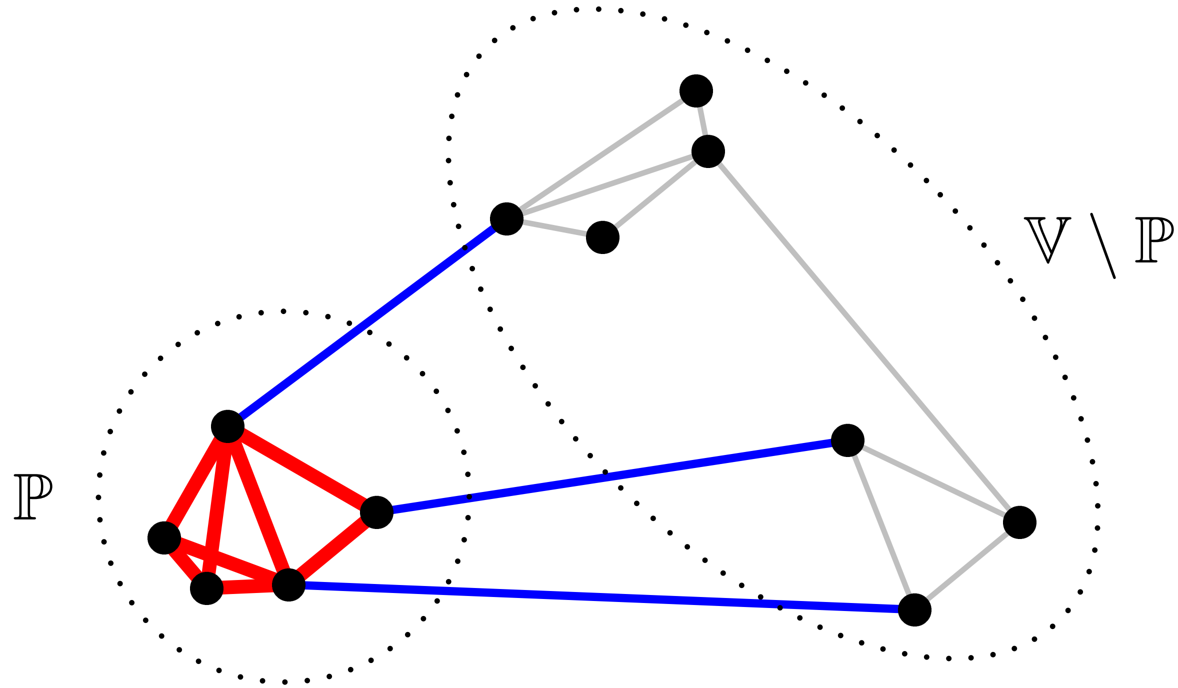
Perceptual Popout



Perceptual Popout



Goodness of Grouping: Attraction and Repulsion

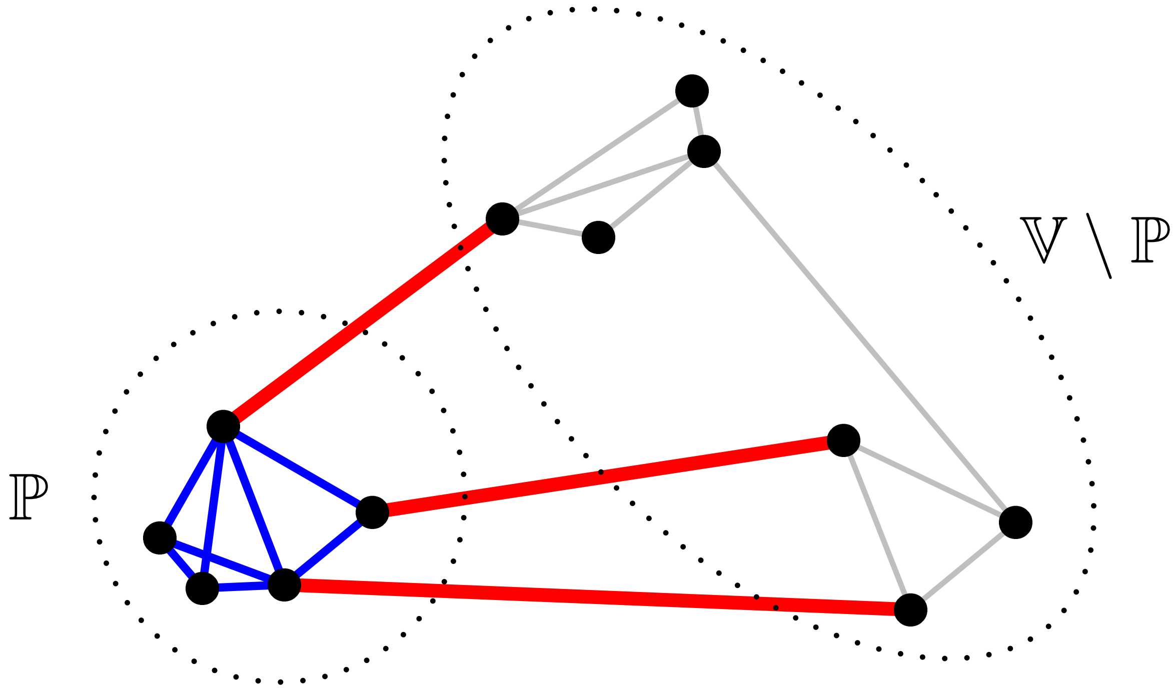


Maximize within-group attraction: $\text{linkratio}(\mathbb{P}, \mathbb{P}; A)$

Minimize between-group attraction: $\text{linkratio}(\mathbb{P}, \mathbb{V} \setminus \mathbb{P}; A)$

Equivalent: $\text{linkratio}(\mathbb{P}, \mathbb{P}; A) + \text{linkratio}(\mathbb{P}, \mathbb{V} \setminus \mathbb{P}; A) = 1$

Goodness of Grouping: Attraction and Repulsion



Maximize between-group repulsion: $\mathrm{linkratio}(\mathbb{P}, \mathbb{V} \setminus \mathbb{P}; R)$

Minimize within-group repulsion: $\mathrm{linkratio}(\mathbb{P}, \mathbb{P}; R)$

Equivalent: $\mathrm{linkratio}(\mathbb{P}, \mathbb{P}; R) + \mathrm{linkratio}(\mathbb{P}, \mathbb{V} \setminus \mathbb{P}; R) = 1$

Normalized Cuts with Attraction and Repulsion

- Criteria

$$\mathrm{knassoc}(\Gamma_{\mathbb{V}}^{K}) = \frac{1}{K}\sum_{l=1}^{K}\frac{\mathrm{links}(\mathbb{V}_l, \mathbb{V}_l; A) + \mathrm{links}(\mathbb{V}_l, \mathbb{V}\setminus\mathbb{V}_l; R)}{\mathrm{degree}(\mathbb{V}_l; A) + \mathrm{degree}(\mathbb{V}_l; R)}$$

$$\mathrm{kncuts}(\Gamma_{\mathbb{V}}^{K}) = \frac{1}{K}\sum_{l=1}^{K}\frac{\mathrm{links}(\mathbb{V}_l, \mathbb{V}\setminus\mathbb{V}_l; A) + \mathrm{links}(\mathbb{V}_l, \mathbb{V}_l; R)}{\mathrm{degree}(\mathbb{V}_l; A) + \mathrm{degree}(\mathbb{V}_l; R)}$$

Normalized Cuts with Attraction and Repulsion

- Criteria

$$\mathrm{knassoc}(\Gamma_{\mathbb{V}}^{K}) = \frac{1}{K}\sum_{l=1}^{K}\frac{\mathrm{links}(\mathbb{V}_l,\mathbb{V}_l;A)+\mathrm{links}(\mathbb{V}_l,\mathbb{V}\setminus\mathbb{V}_l;R)}{\mathrm{degree}(\mathbb{V}_l;A)+\mathrm{degree}(\mathbb{V}_l;R)}$$

$$\mathrm{kncuts}(\Gamma_{\mathbb{V}}^{K}) = \frac{1}{K}\sum_{l=1}^{K}\frac{\mathrm{links}(\mathbb{V}_l,\mathbb{V}\setminus\mathbb{V}_l;A)+\mathrm{links}(\mathbb{V}_l,\mathbb{V}_l;R)}{\mathrm{degree}(\mathbb{V}_l;A)+\mathrm{degree}(\mathbb{V}_l;R)}$$

- Equivalent weight matrix and degree matrix

$$\begin{aligned}\hat{W} &= A - R + D_R\\ \hat{D} &= D_A + D_R\end{aligned}$$

Negative Weights and Regularization

- Negative weights:

$$
\begin{aligned}
W &= A - R \\
&= \text{(positive entries + offset)} - \text{(negative entries + offset)}
\end{aligned}
$$

Negative Weights and Regularization

- Negative weights:

$$W = A - R$$
$$= \text{(positive entries + offset)} - \text{(negative entries + offset)}$$

- Equivalent matrices: $(\hat{W} + D_{\text{offset}},\ \hat{D} + 2D_{\text{offset}})$

Negative Weights and Regularization

- Negative weights:

  $$W = A - R$$

  $$= \text{(positive entries + offset)} - \text{(negative entries + offset)}$$

- Equivalent matrices: $(\hat{W} + D_{\text{offset}},\ \hat{D} + 2D_{\text{offset}})$

- Regularization: increase the degrees of nodes without changing the sizes of weights between two nodes.

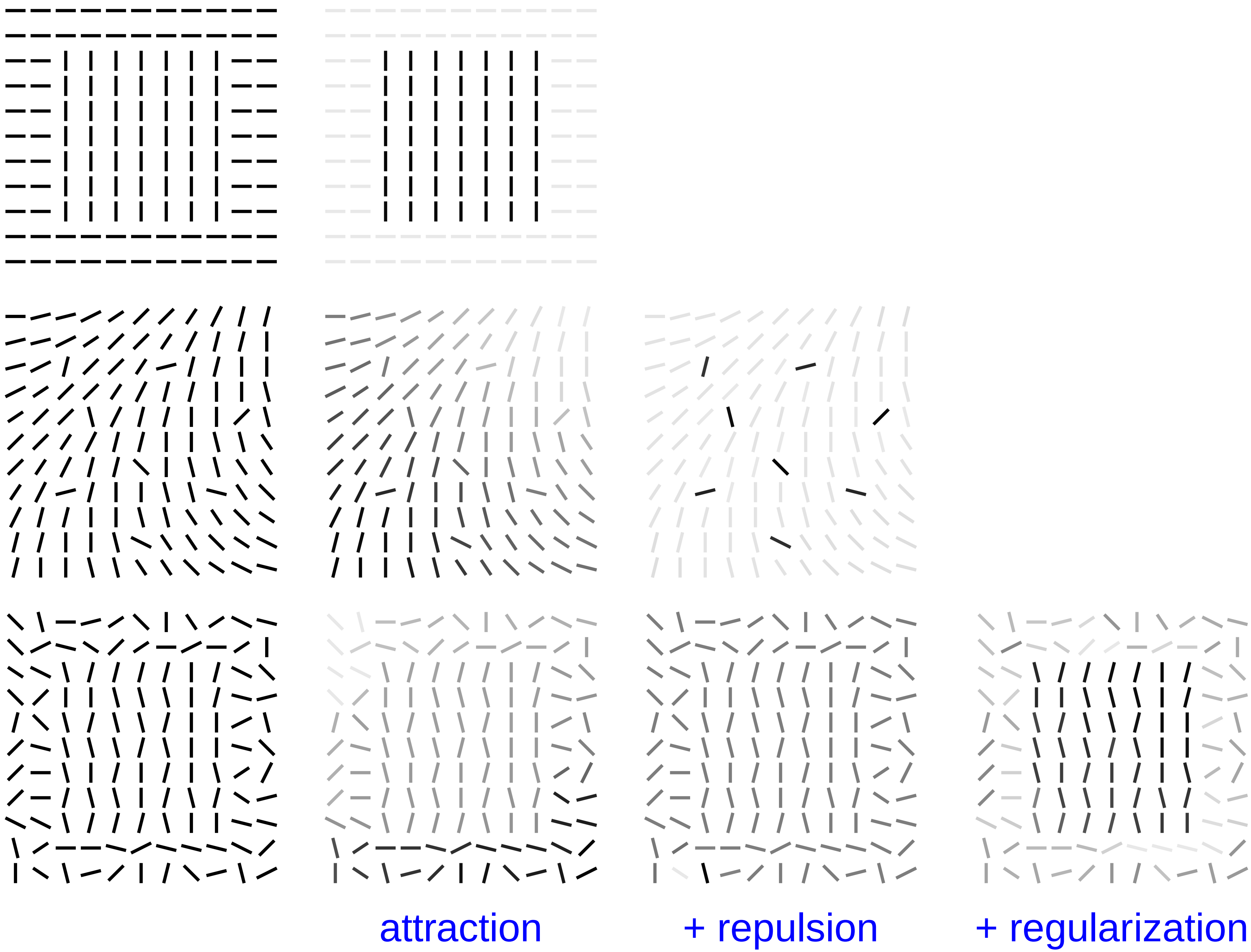
Negative Weights and Regularization

- Negative weights:

  $$W = A - R$$
  $$= \text{(positive entries + offset)} - \text{(negative entries + offset)}$$

- Equivalent matrices: $(\hat{W} + D_{\text{offset}},\ \hat{D} + 2D_{\text{offset}})$
- Regularization: increase the degrees of nodes without changing the sizes of weights between two nodes.
- Decrease the sensitivity of $\mathrm{linkratio}$ for nodes with little connections.

Roles of Attraction, Repulsion, Regularization



Segmentation with Repulsion and Regularization

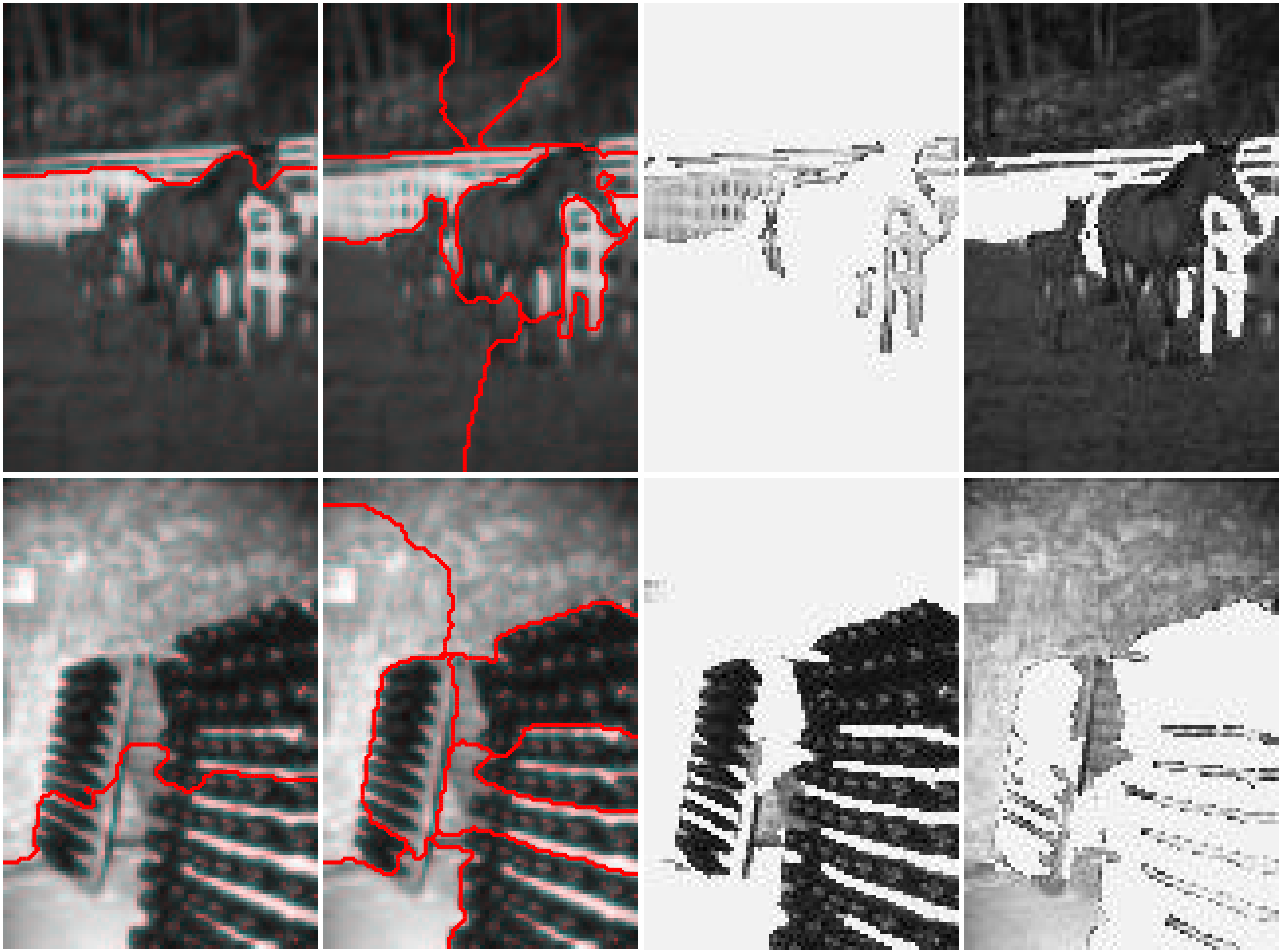
attraction

attraction, repulsion and regularization

Segmentation with Repulsion and Regularization

attraction attraction, repulsion and regularization

Segmentation with Repulsion and Regularization

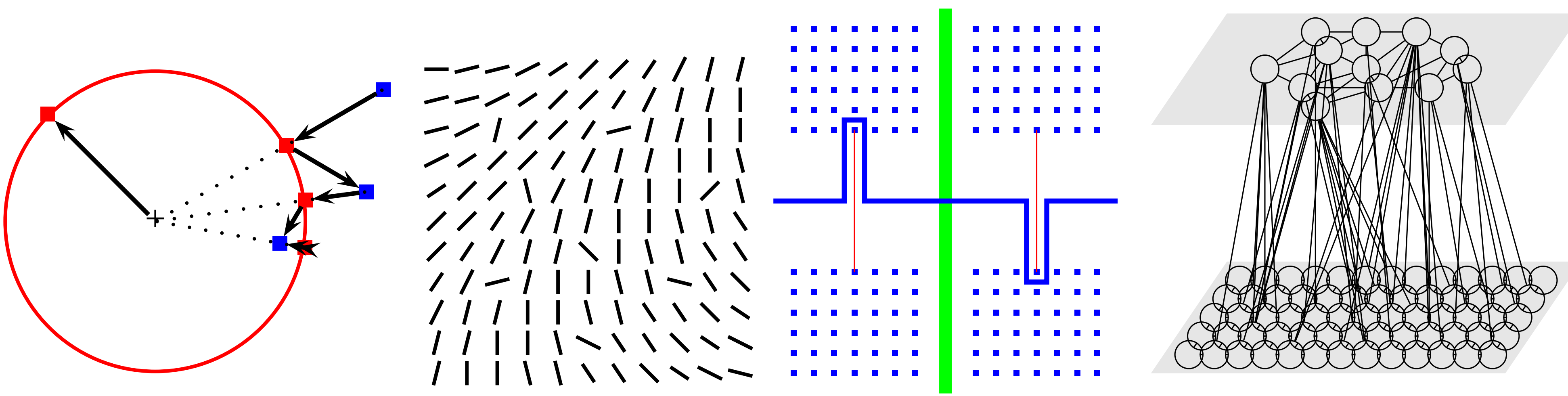


attraction

attraction, repulsion and regularization

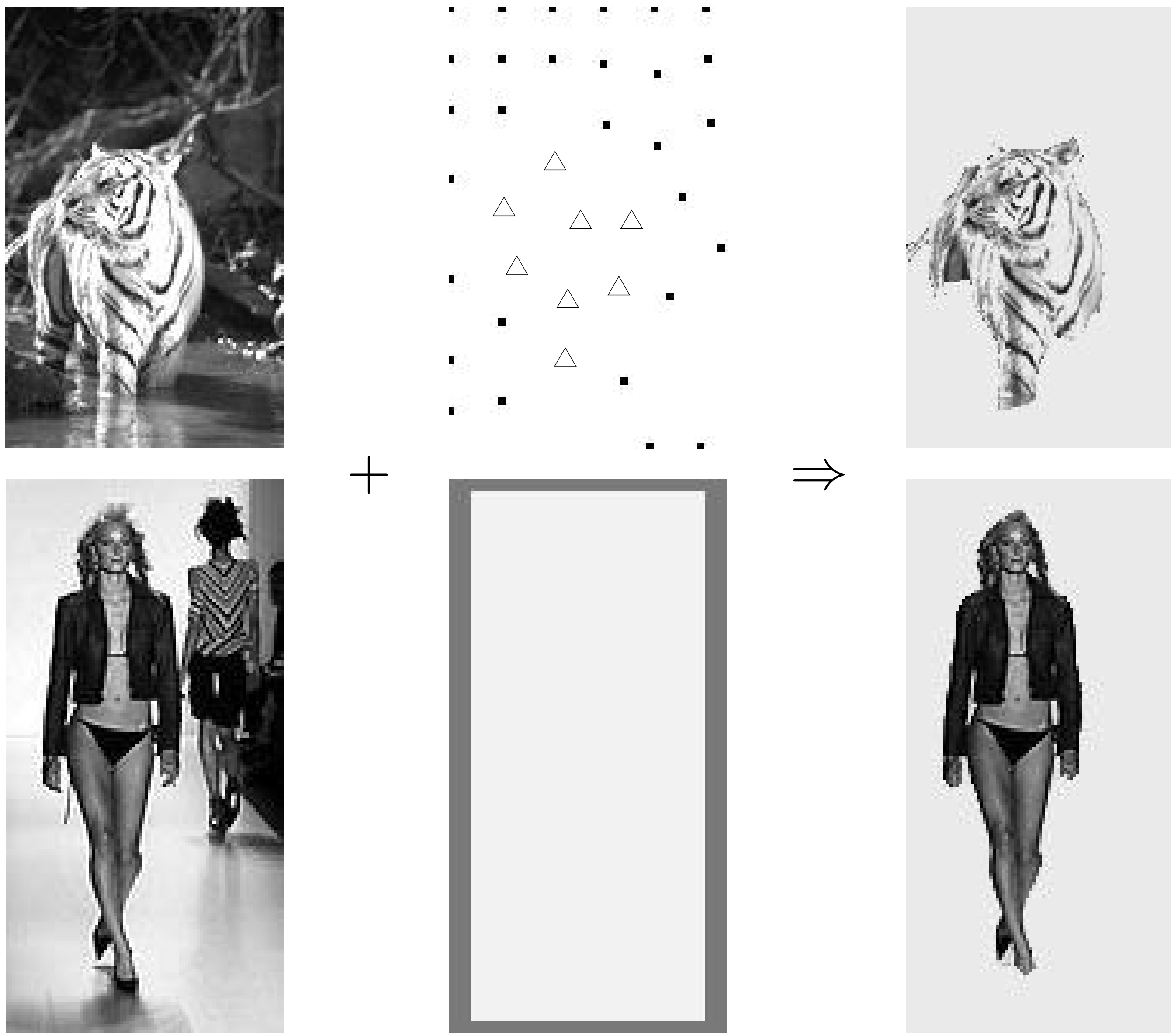
Outline

1. Computational framework: spectral clustering
2. Expand the repertoire of grouping cues: dissimilarity
3. Guide grouping with partial cues
4. Guide grouping with object knowledge
5. Summary and future work

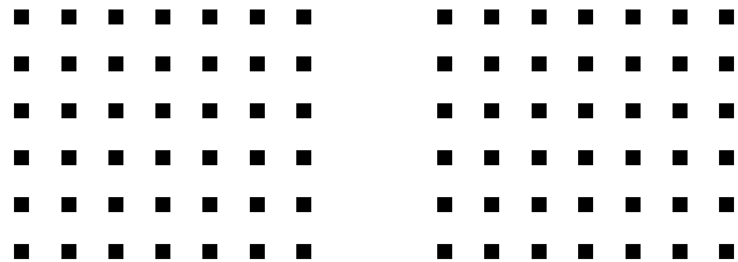


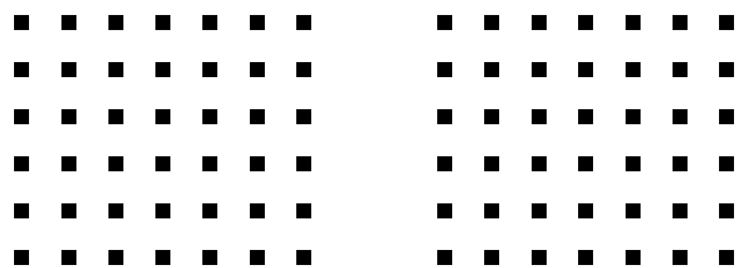
1 2 3 4

Grouping with Partial Cues



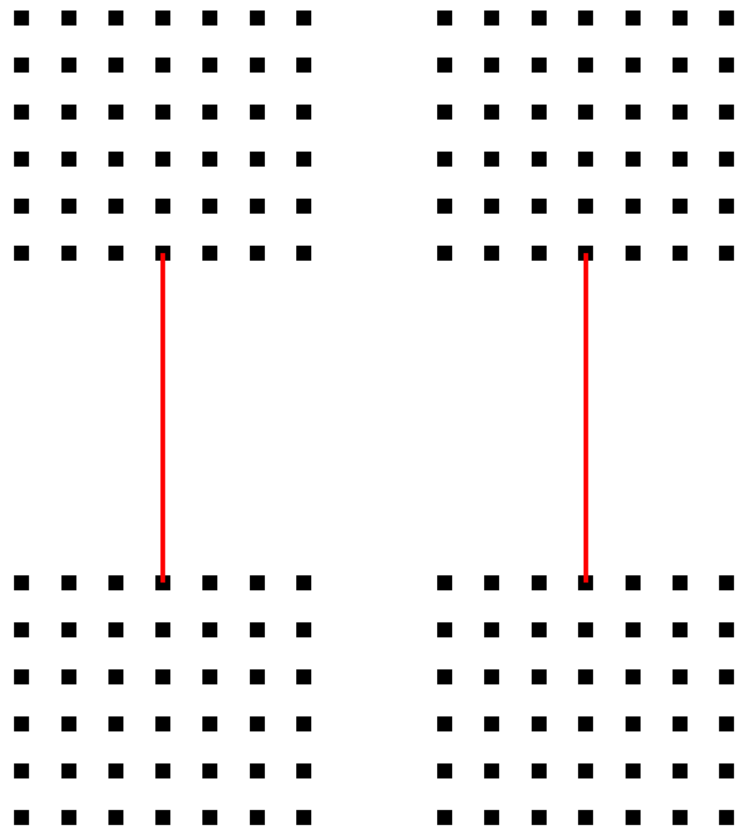
Basic Formulation: Grouping with Constraints





maximize $\varepsilon(\Gamma_{\mathbb{V}}^{K})$

Basic Formulation: Grouping with Constraints



maximize $\varepsilon(\Gamma_{\mathbb{V}}^{K})$

subject to $X(i,l) = X(j,l)$

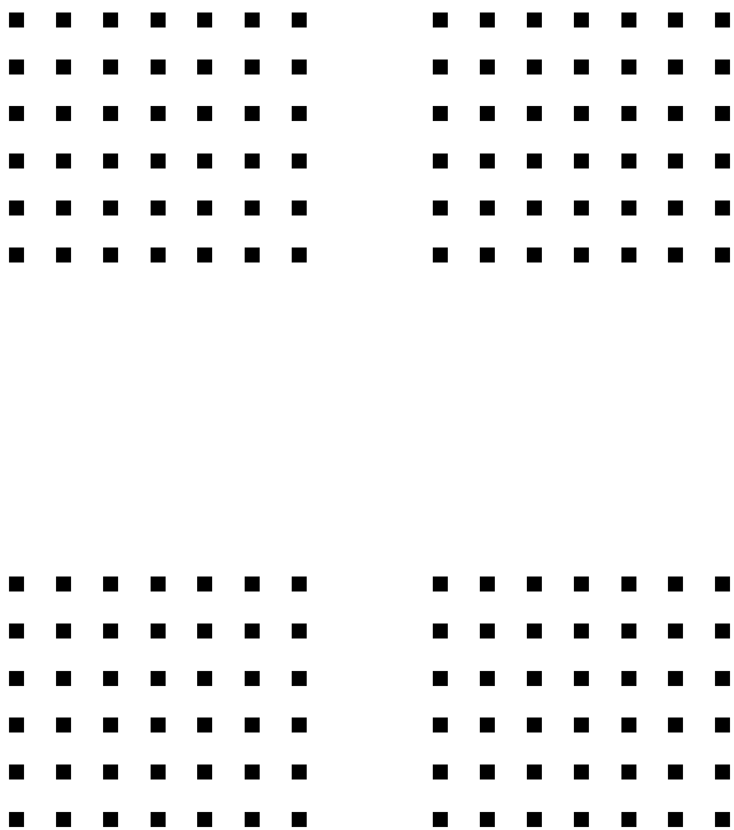
Computing Constrained Normalized Cuts

- Constrained eigenvalue problem

- Efficient solution using a projector onto the feasible space

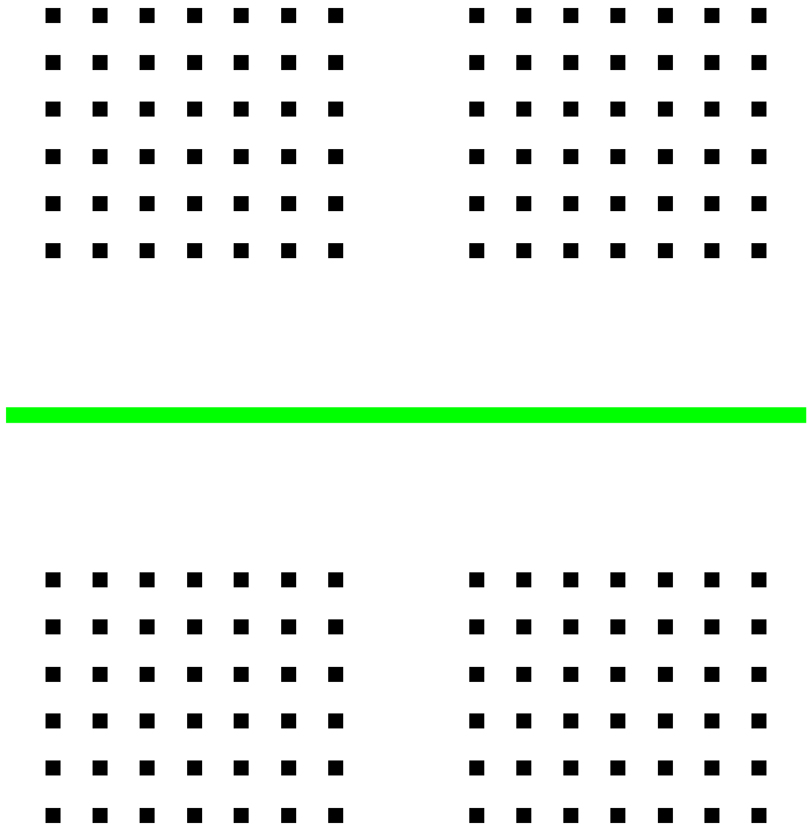
- Generalize Rayleigh-Ritz theorem to projected matrices

Why Simple Constraints Are Insufficient



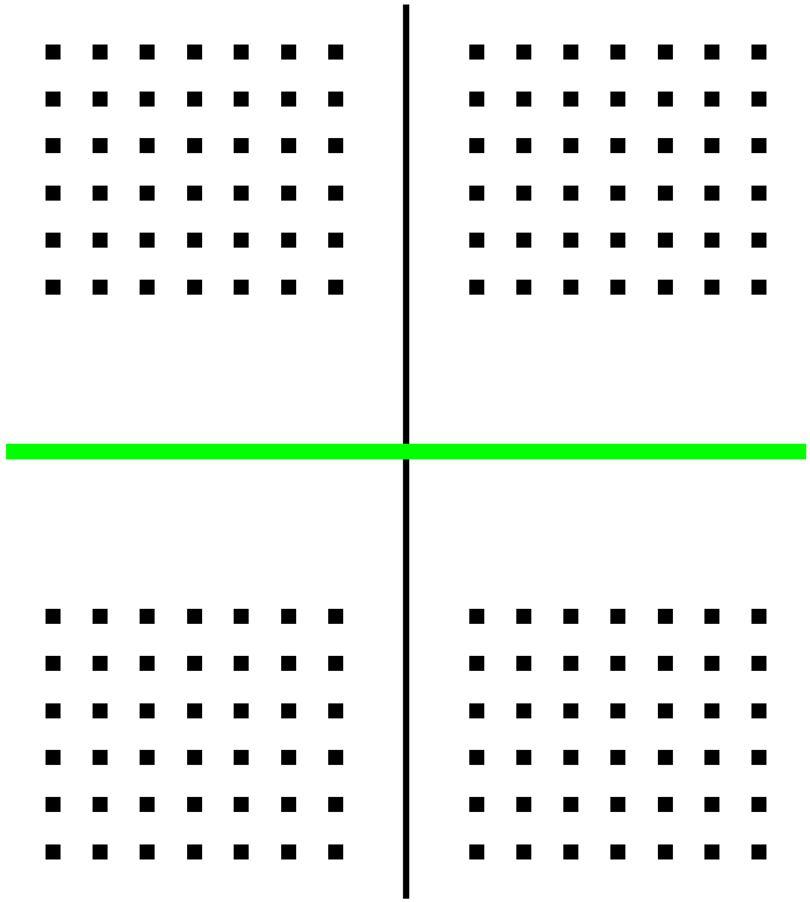
natural grouping

Why Simple Constraints Are Insufficient



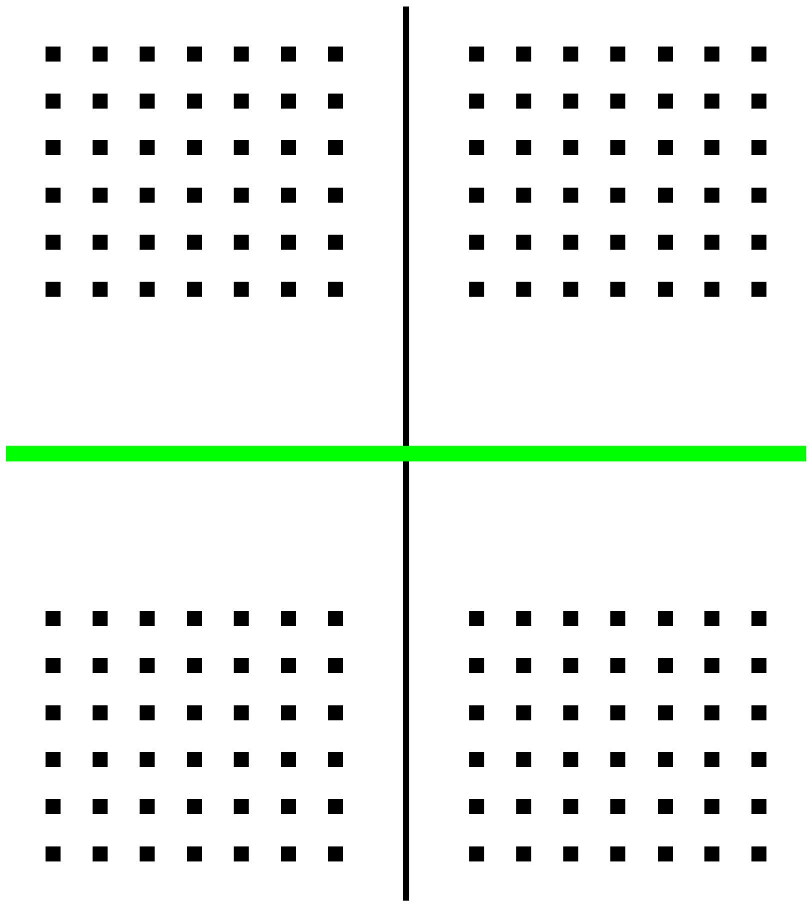
natural grouping

Why Simple Constraints Are Insufficient

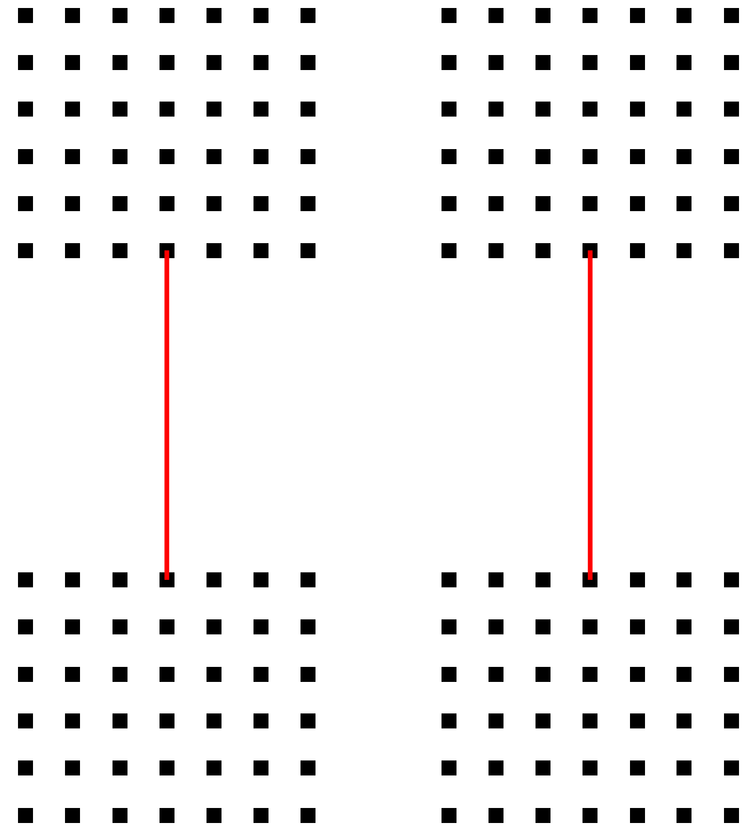


natural grouping

Why Simple Constraints Are Insufficient

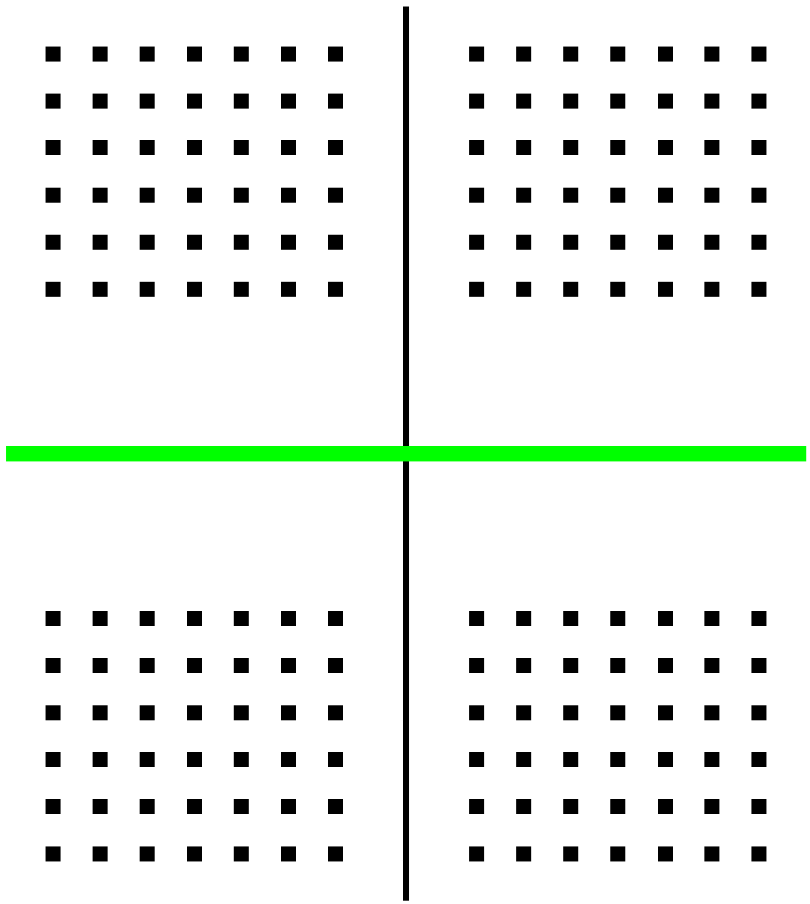


natural grouping

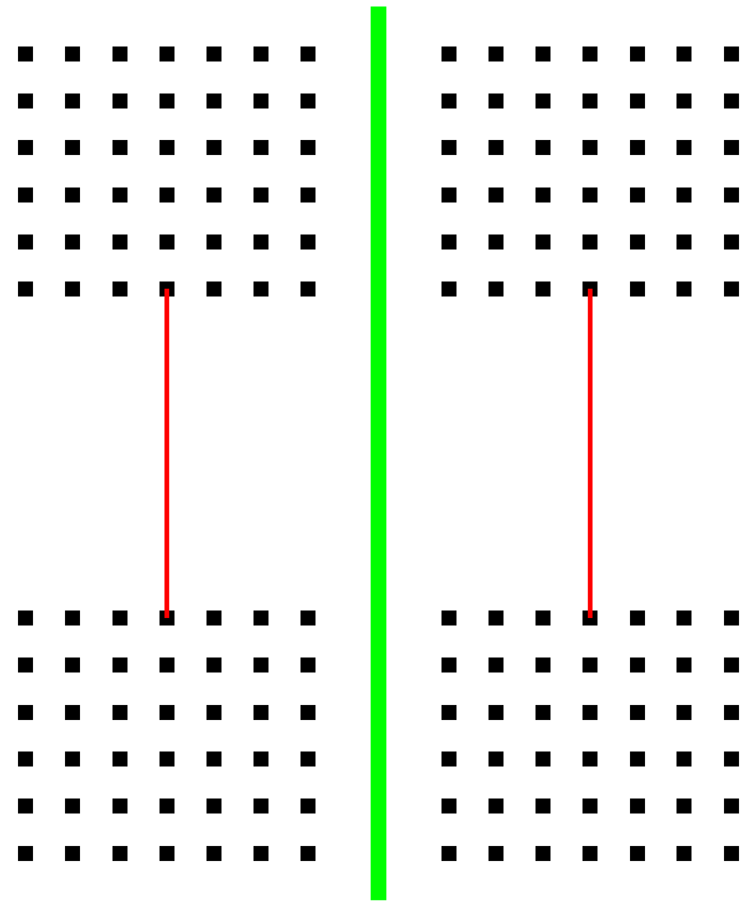


constrained grouping

Why Simple Constraints Are Insufficient

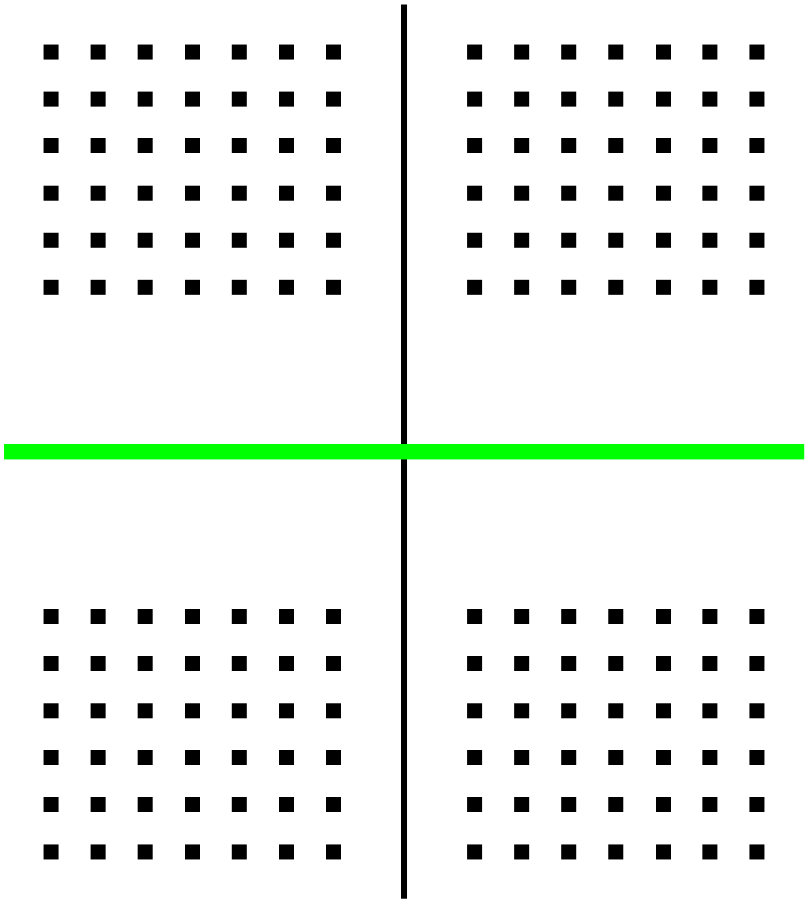


natural grouping

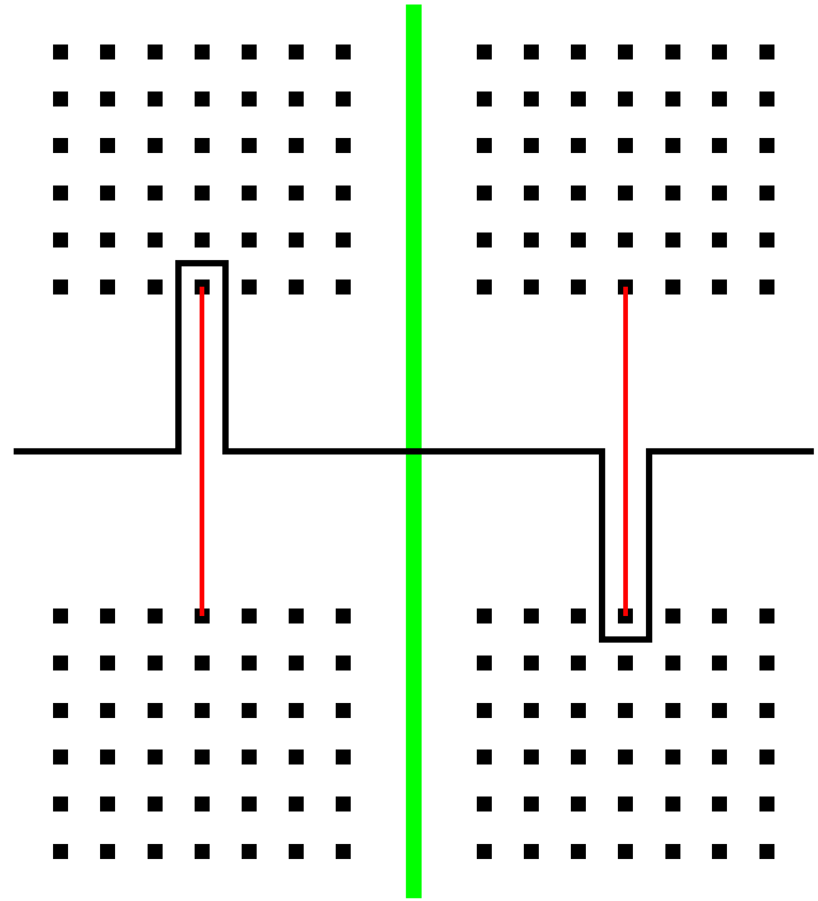


constrained grouping

Why Simple Constraints Are Insufficient



natural grouping



constrained grouping

Remedy: Propagate Constraints

- General formulation:

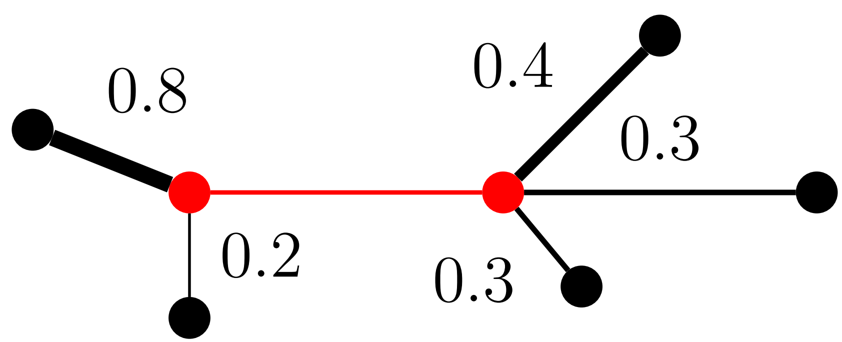
$$\begin{aligned} \text{maximize} \quad & \varepsilon(\Gamma_{\mathbb{V}}^{K}) \\ \text{subject to} \quad & \color{red}{S} \cdot X(i,l) = \color{red}{S} \cdot X(j,l) \end{aligned}$$

Remedy: Propagate Constraints

- General formulation:

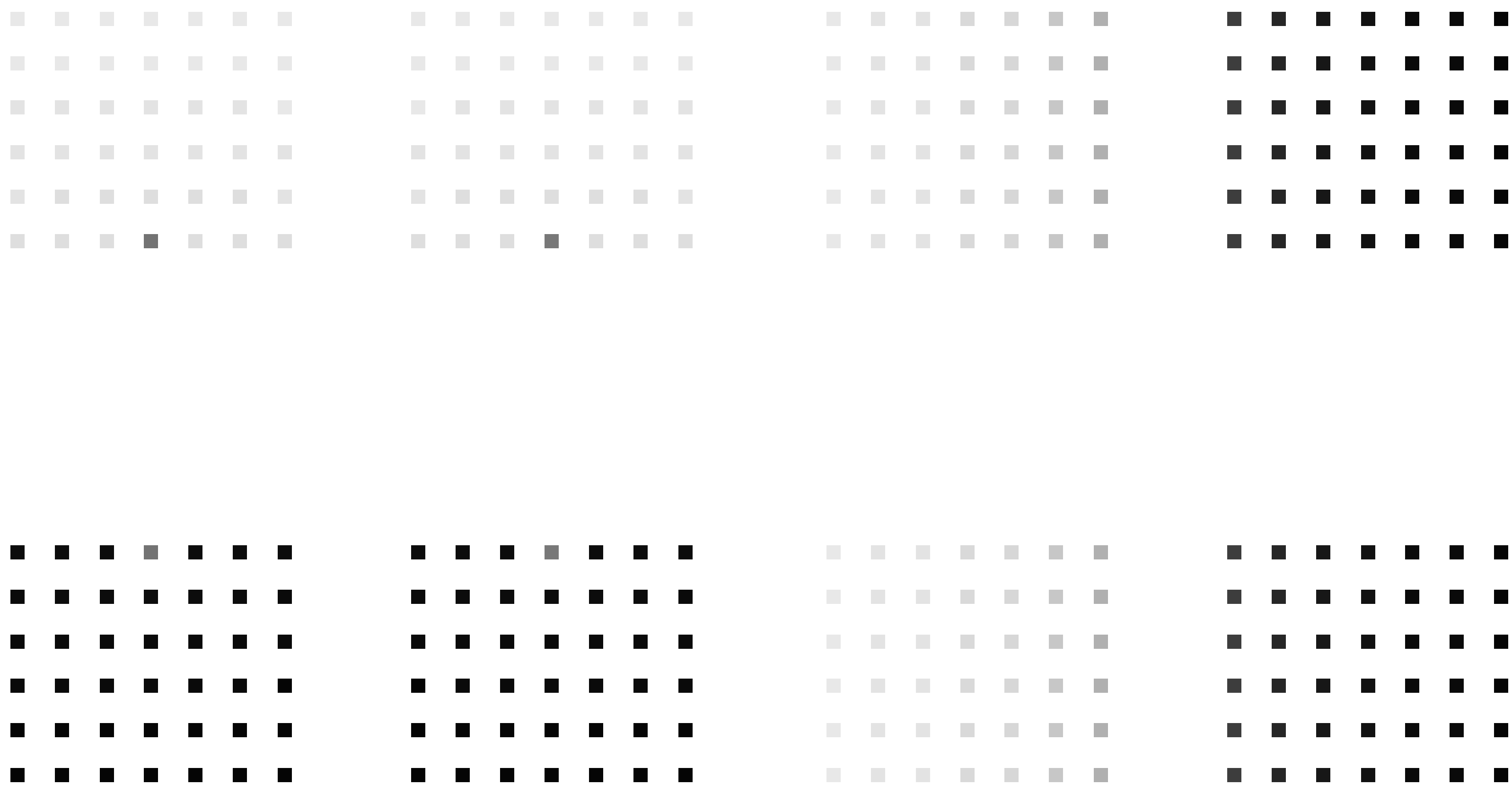
$$\begin{aligned} \text{maximize} \quad & \varepsilon(\Gamma_{\mathbb{V}}^{K}) \\ \text{subject to} \quad & S \cdot X(i,l) = S \cdot X(j,l) \end{aligned}$$

- Normalized cuts:



$$\begin{aligned} S &= P, \quad \text{or} \\ \sum_k P_{ik} X(k,l) &= \sum_k P_{jk} X(k,l), \quad \text{or} (P^T U)^T X = 0 \end{aligned}$$

Clustering Points with Sparse Grouping Cues



simple bias

smoothed bias

Image Segmentation with Biased Grouping

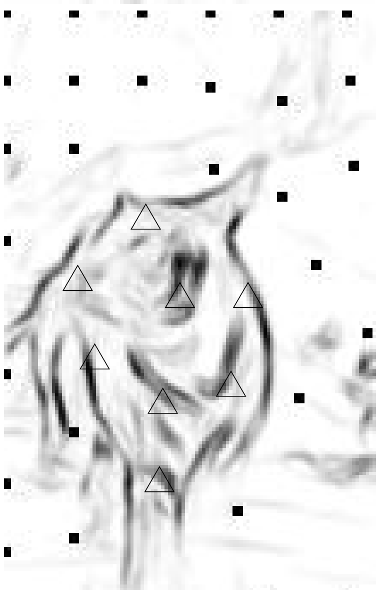
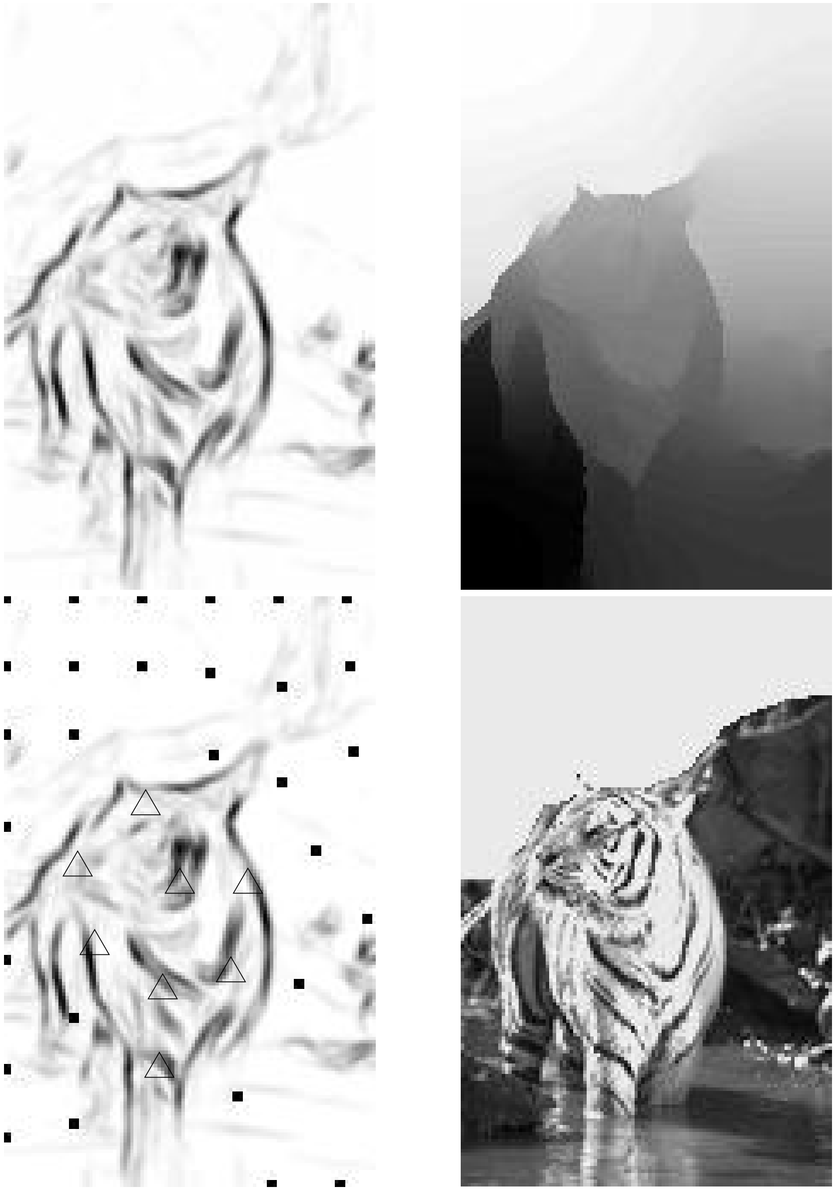
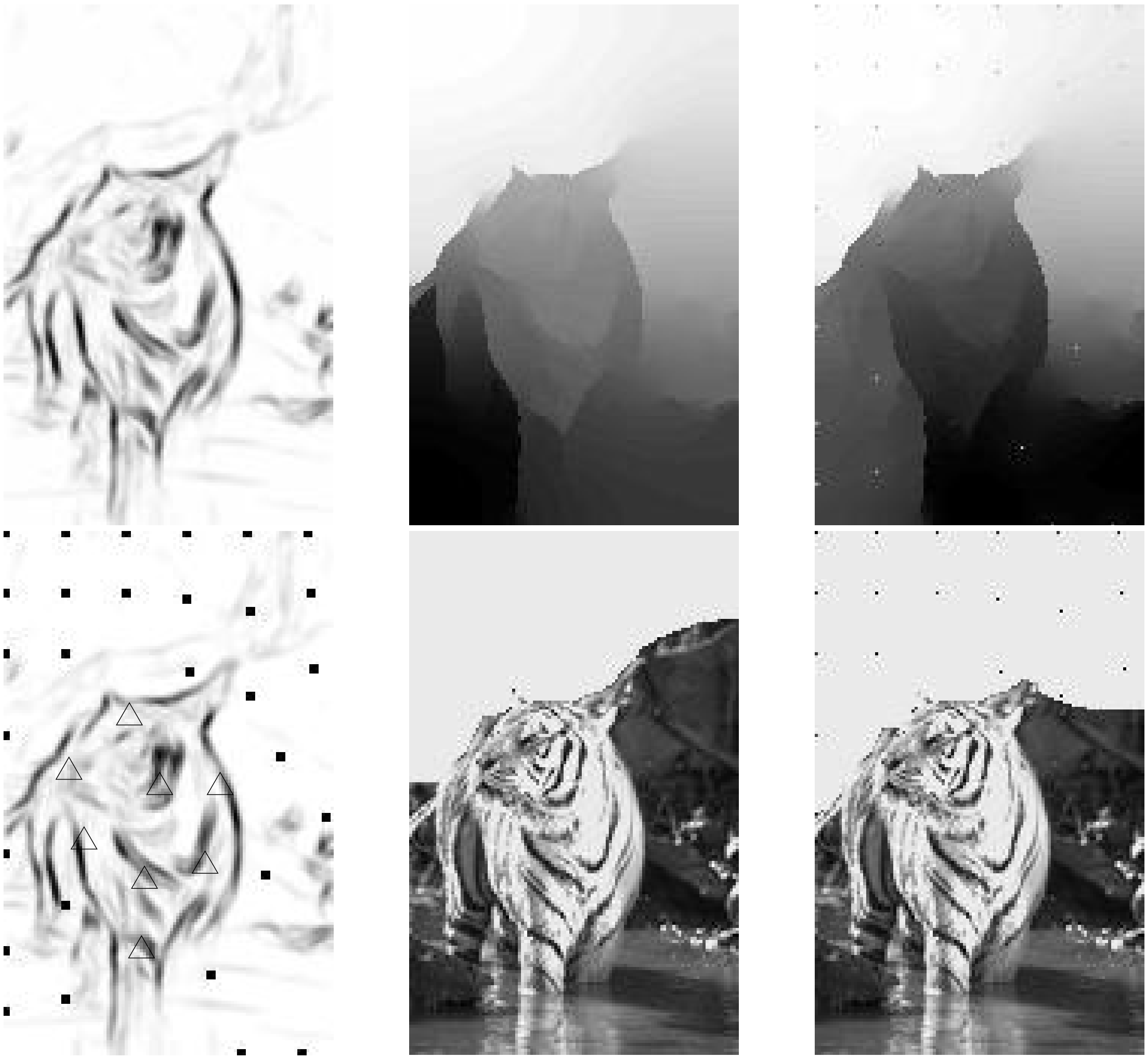


Image Segmentation with Biased Grouping



no bias

Image Segmentation with Biased Grouping



no bias simple bias

Image Segmentation with Biased Grouping

no bias   simple bias   smoothed bias

Image Segmentation with Biased Grouping

Image Segmentation with Biased Grouping

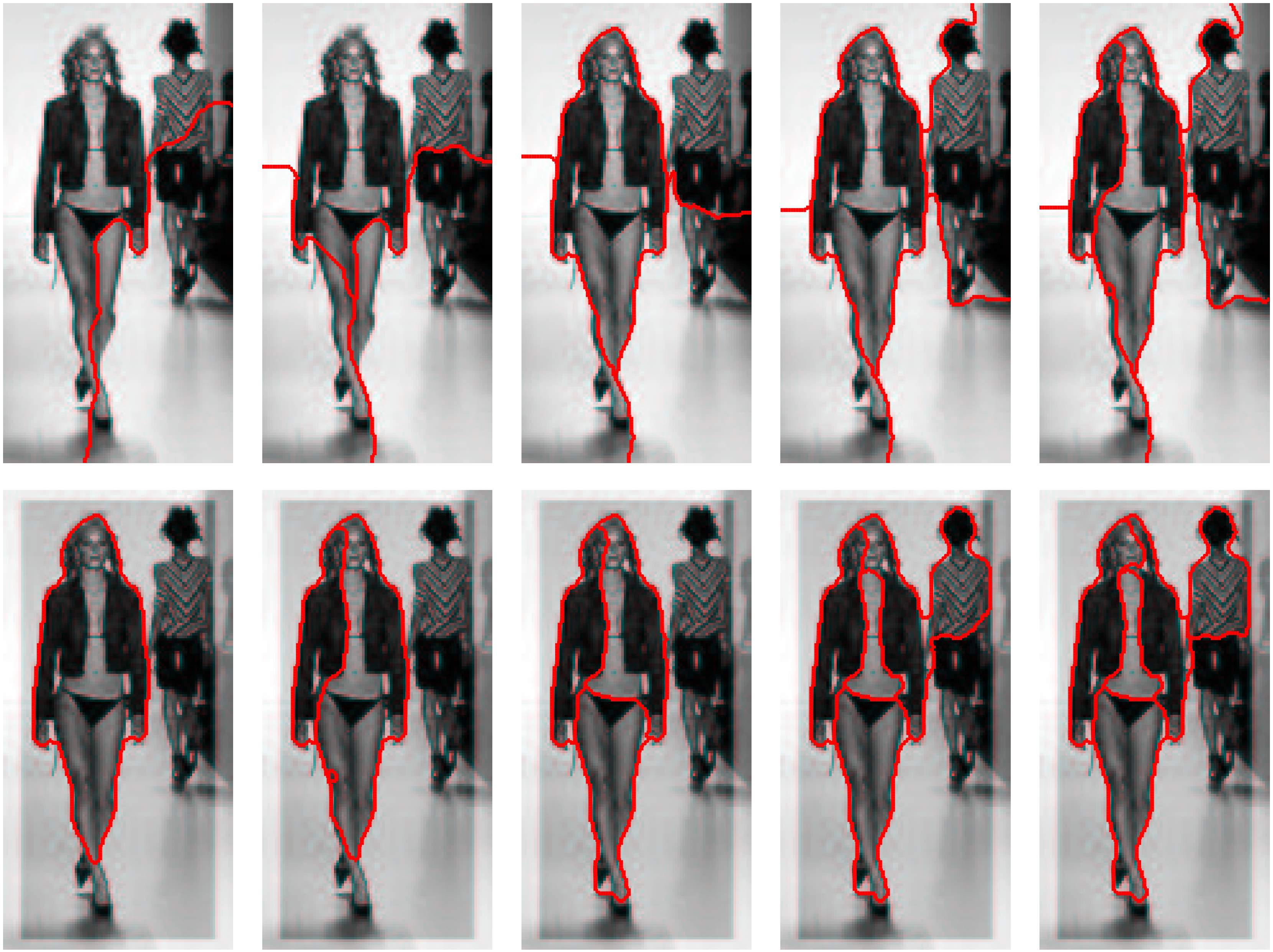
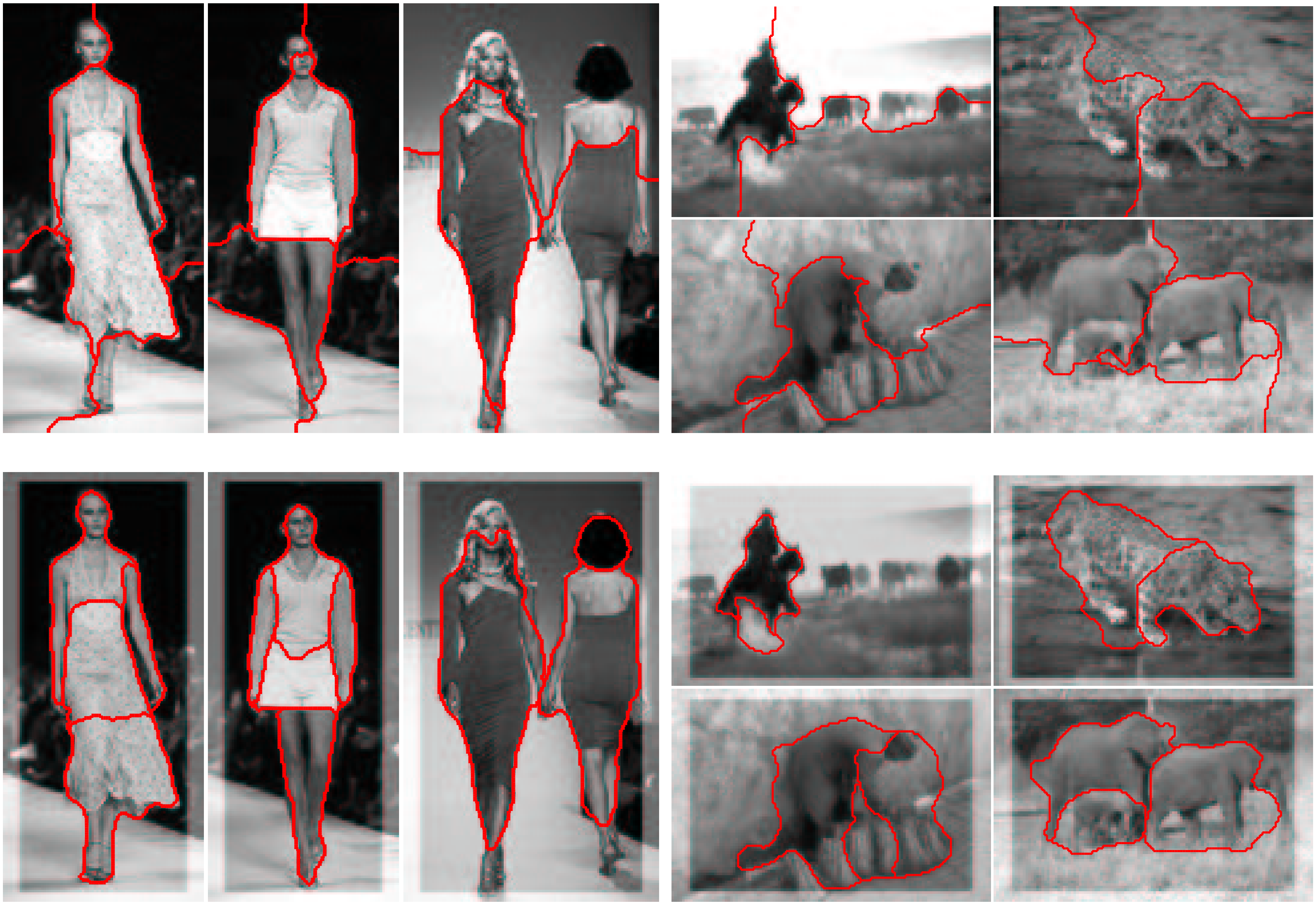


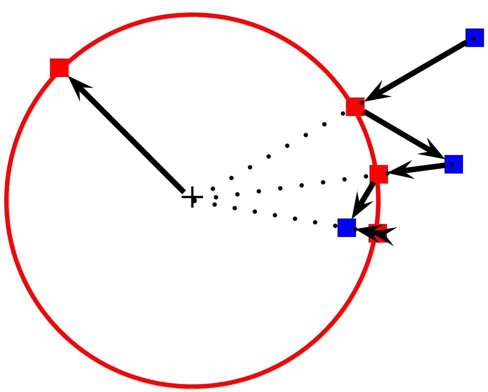
Image Segmentation with Spatial Attention

Image Segmentation with Spatial Attention

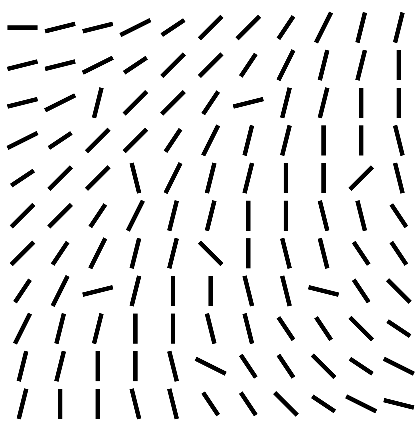


Outline

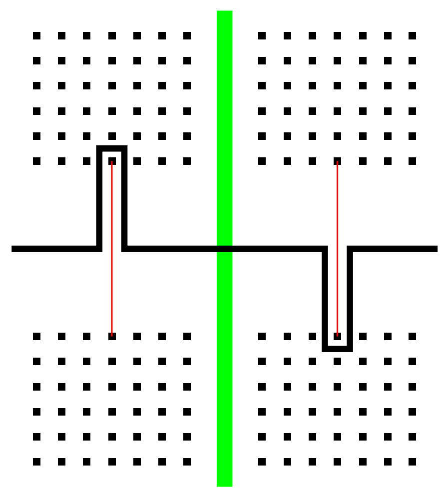
1. Computational framework: spectral clustering
2. Expand the repertoire of grouping cues: dissimilarity
3. Guide grouping with partial cues
4. Guide grouping with object knowledge
5. Summary and future work



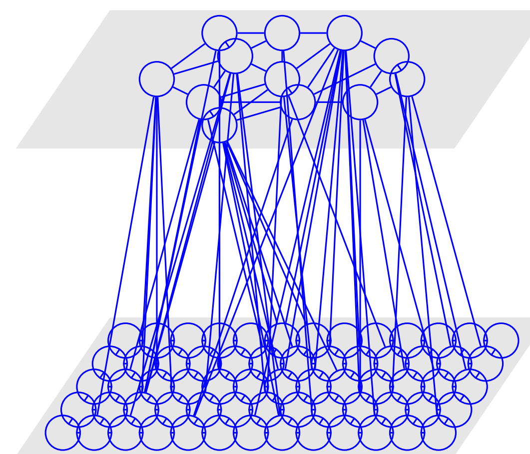
1



2

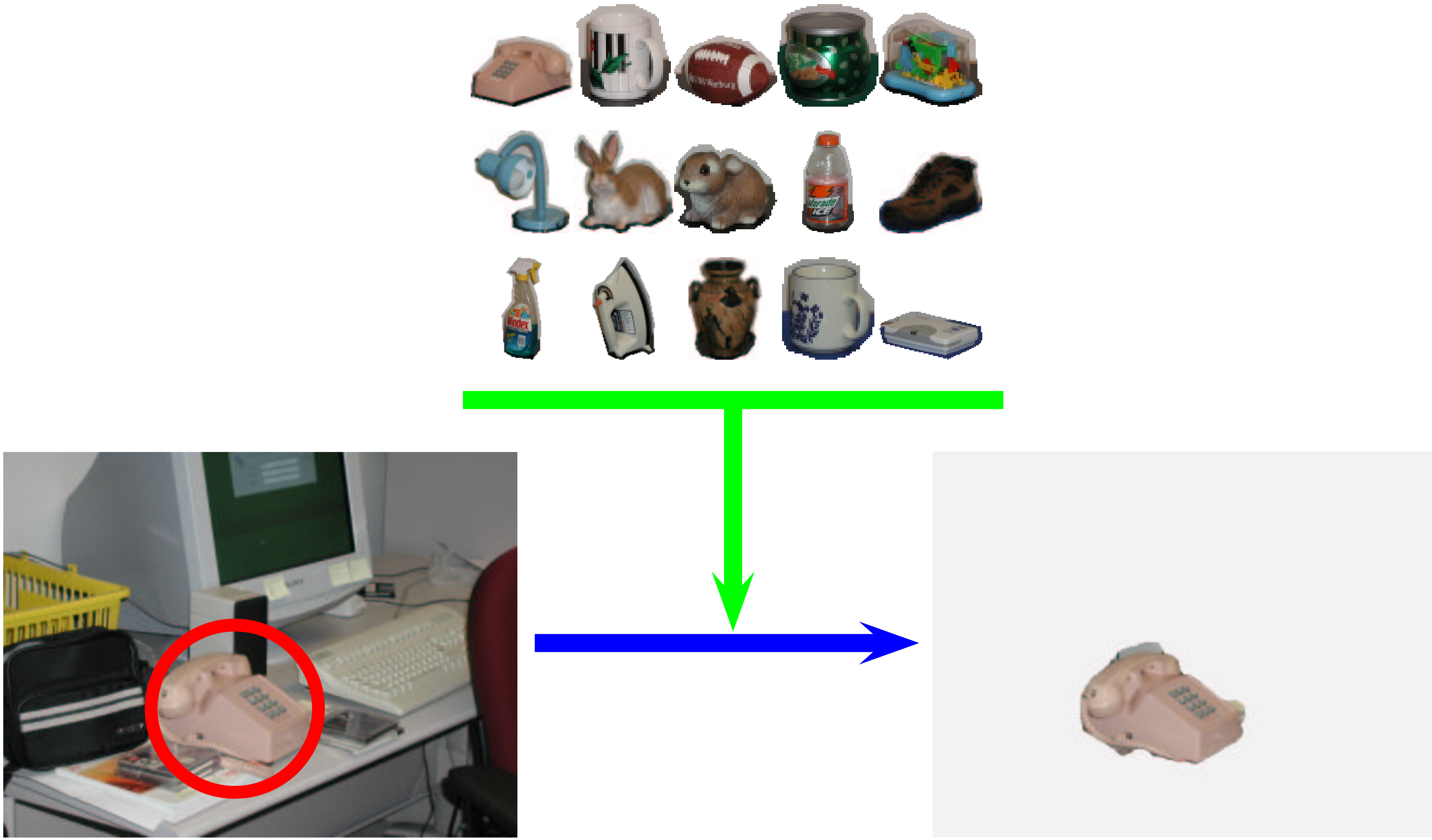


3



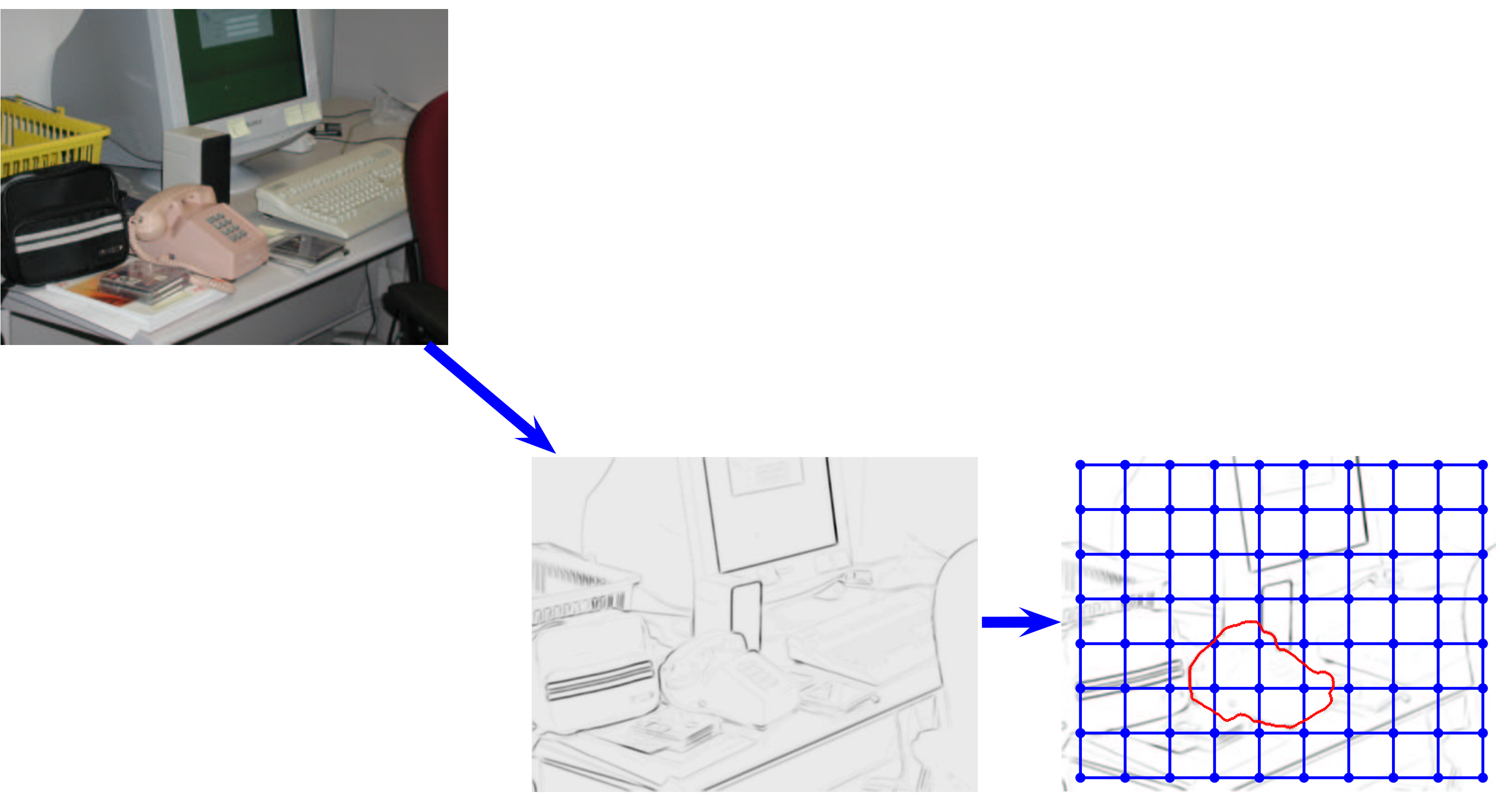
4

Object Segmentation

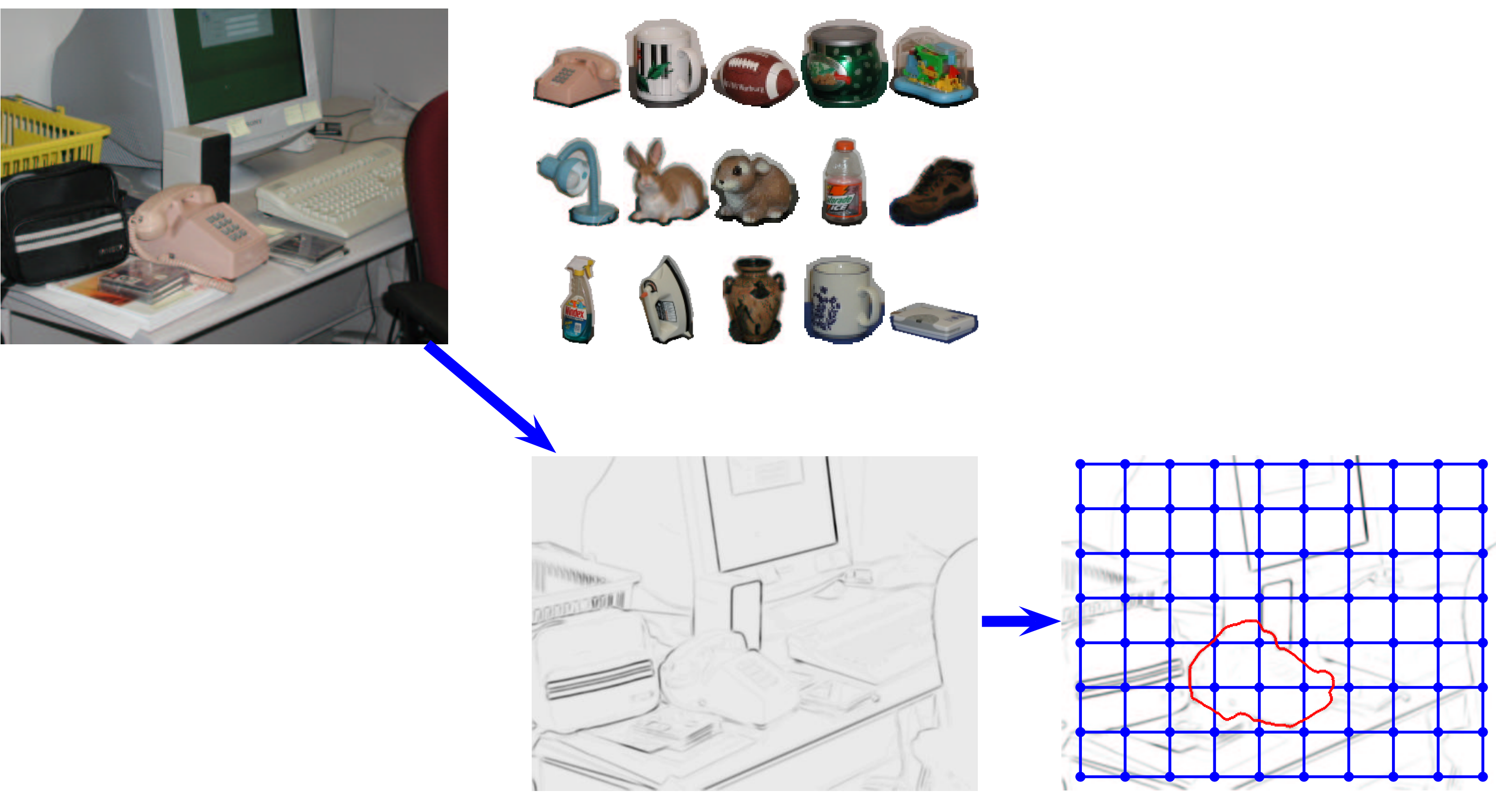


Our Approach to Object Segmentation

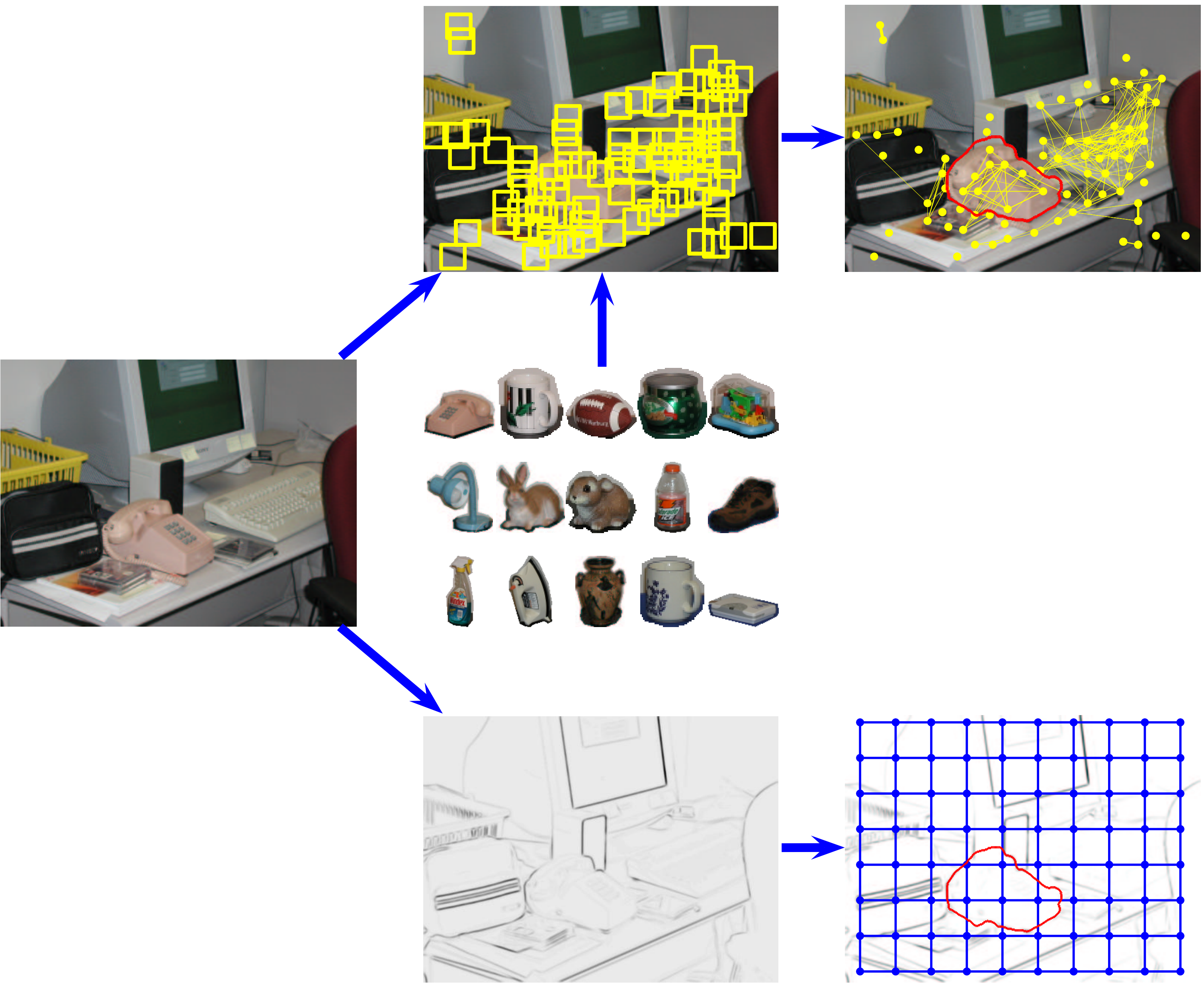
Our Approach to Object Segmentation



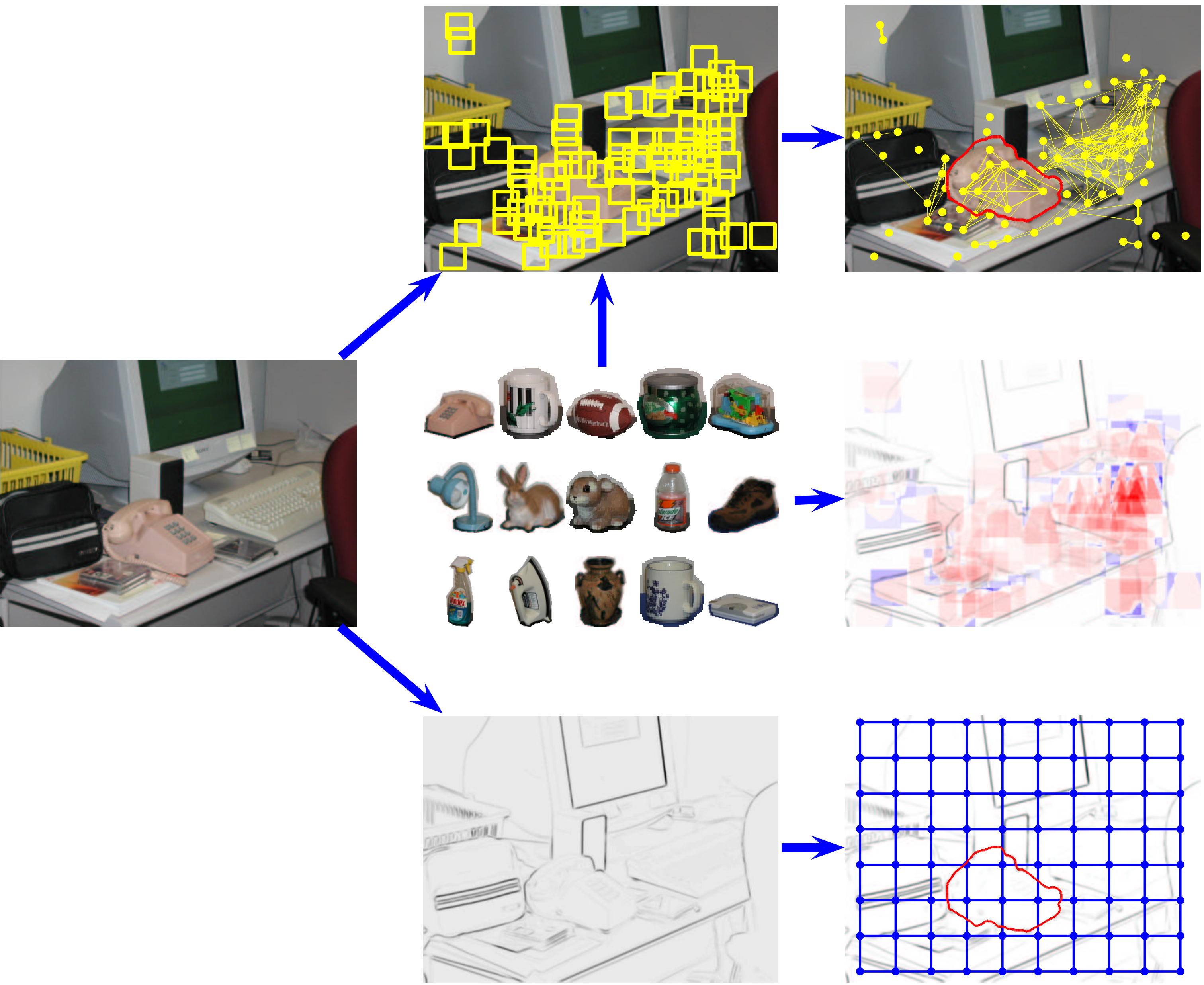
Our Approach to Object Segmentation



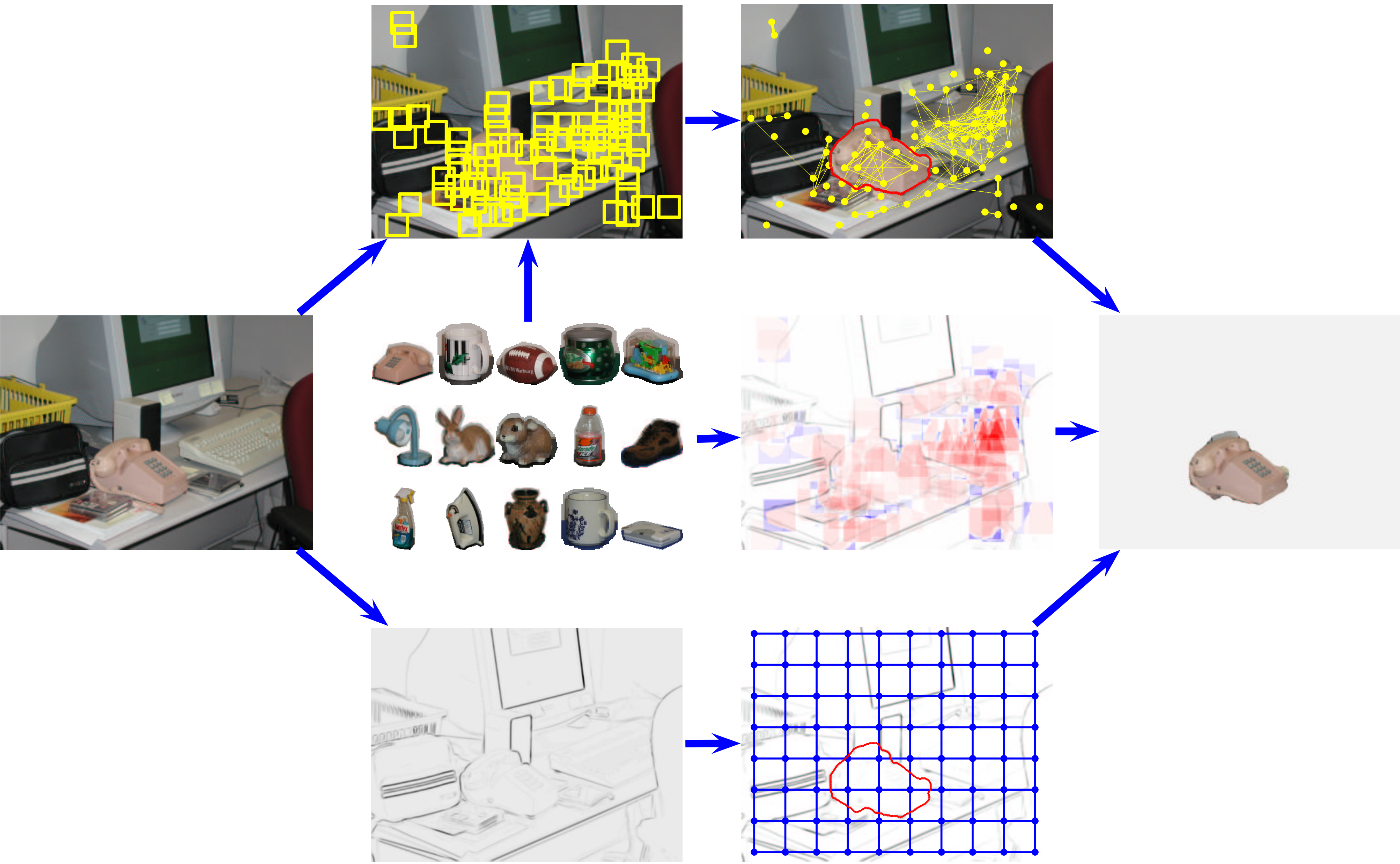
Our Approach to Object Segmentation



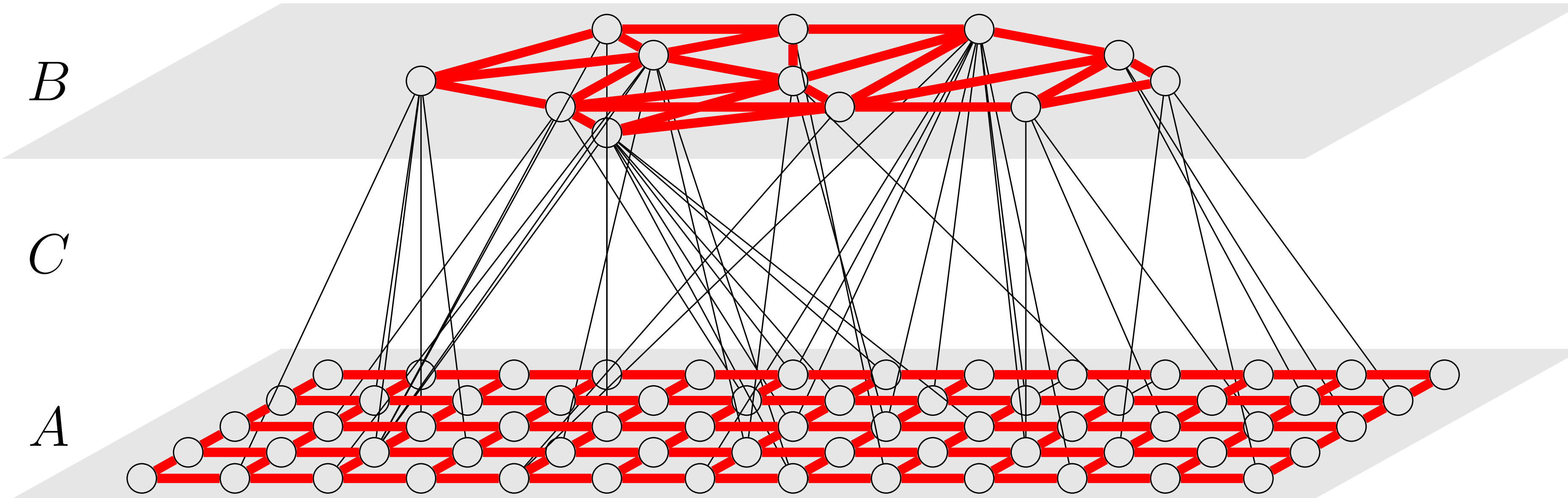
Our Approach to Object Segmentation



Our Approach to Object Segmentation



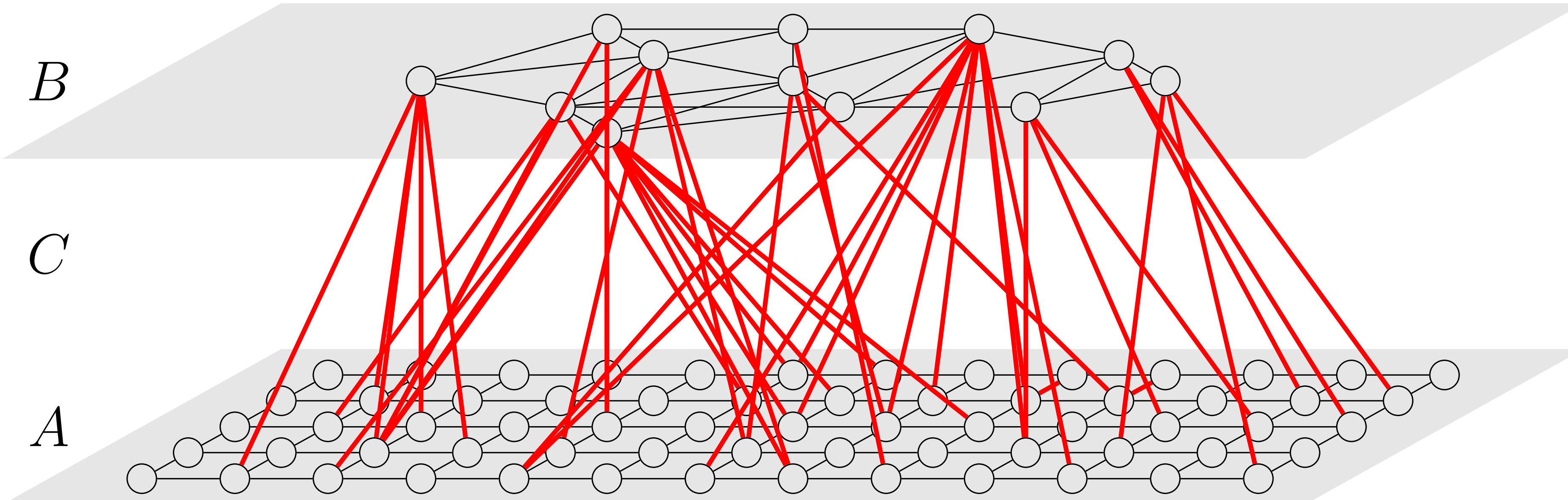
Joint Pixel-Patch Grouping: Criterion



$$\bar{\varepsilon}(\Gamma_{\mathbb{V}}^{K}, \Gamma_{\mathbb{U}}^{K}; A, B) = \frac{1}{K}\sum_{l=1}^{K} \frac{\text{linkratio}(\mathbb{U}_l, \mathbb{U}_l; B) \cdot \text{degree}(\mathbb{U}_l; B)}{\text{degree}(\mathbb{V}_l; A) + \text{degree}(\mathbb{U}_l; B)}$$

$$+ \frac{1}{K}\sum_{l=1}^{K} \frac{\text{linkratio}(\mathbb{V}_l, \mathbb{V}_l; A) \cdot \text{degree}(\mathbb{V}_l; A)}{\text{degree}(\mathbb{V}_l; A) + \text{degree}(\mathbb{U}_l; B)}$$

Joint Pixel-Patch Grouping: Consistency

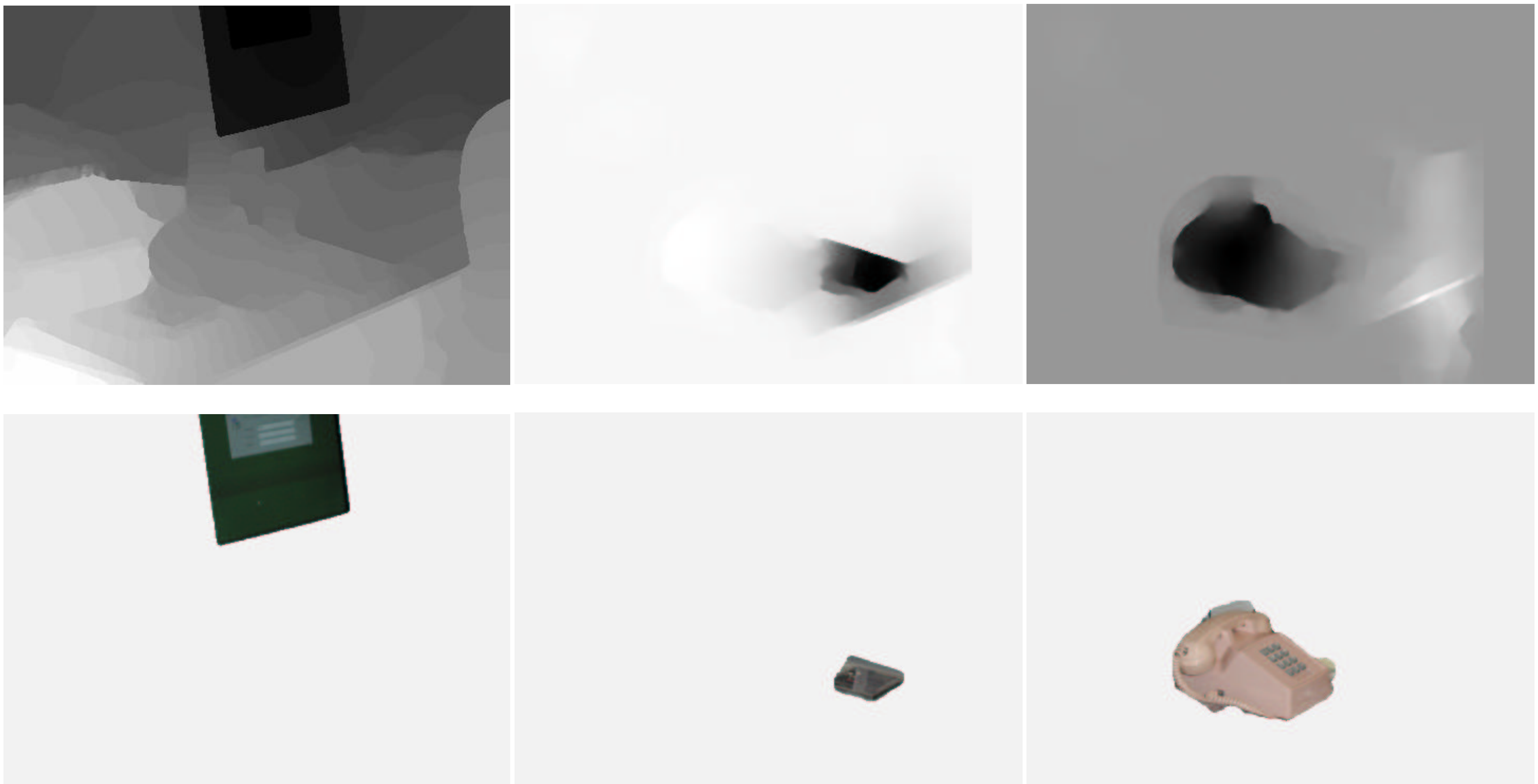


$$\Gamma_{\mathbb{U}}^{K} = \{\mathbb{U}_1, \ldots, \mathbb{U}_K\}, \qquad \Gamma_{\mathbb{V}}^{K} = \{\mathbb{V}_1, \ldots, \mathbb{V}_K\}$$

Bias linking patches with their pixels

How Object Knowledge Helps Segmentation

pixel only	pixel w/ ROI	pixel-patch



541s	150s	110s

How Segmentation Helps Object Detection

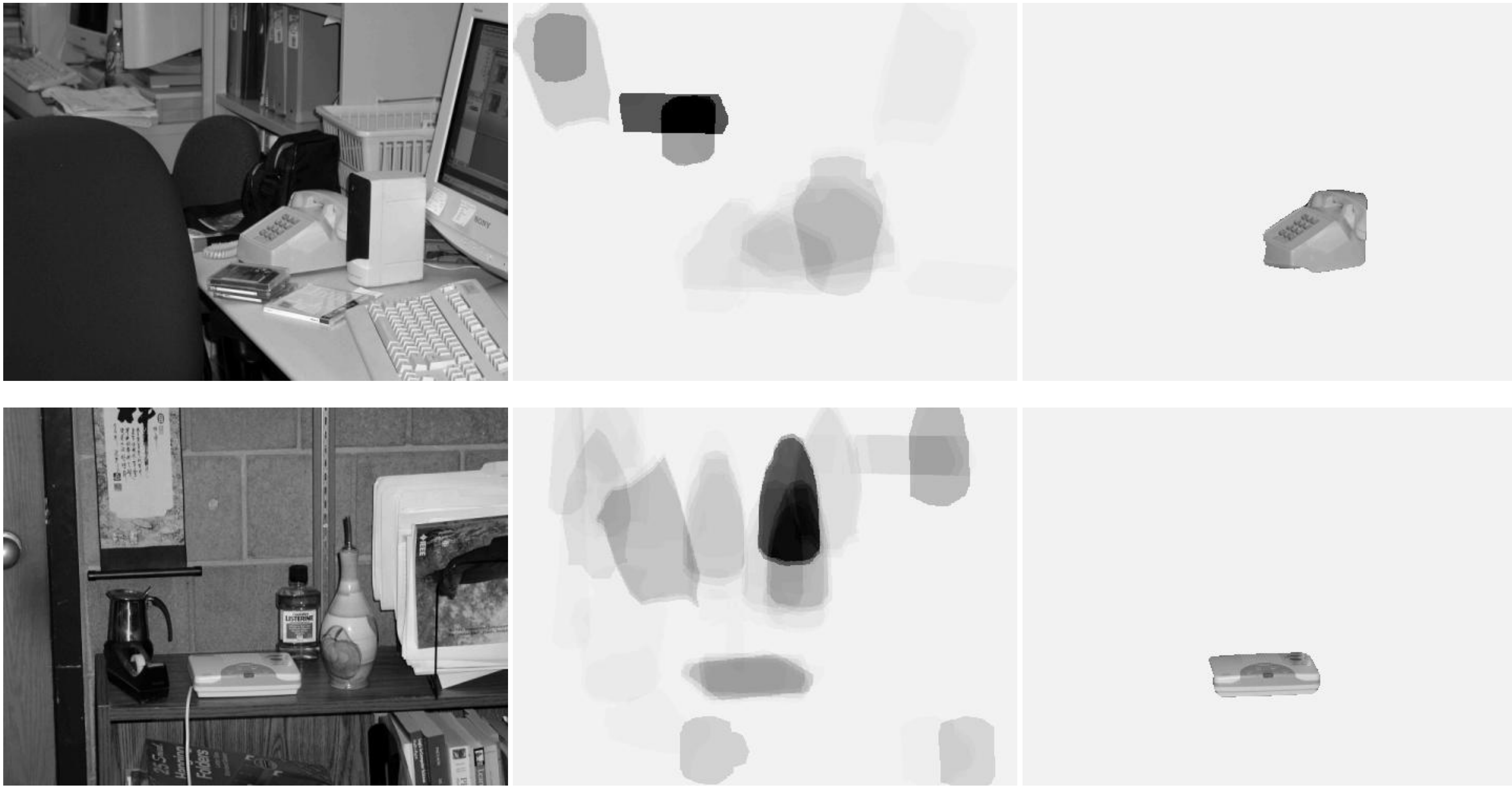


image | patch density | segmentation

When Does Our Method Fail

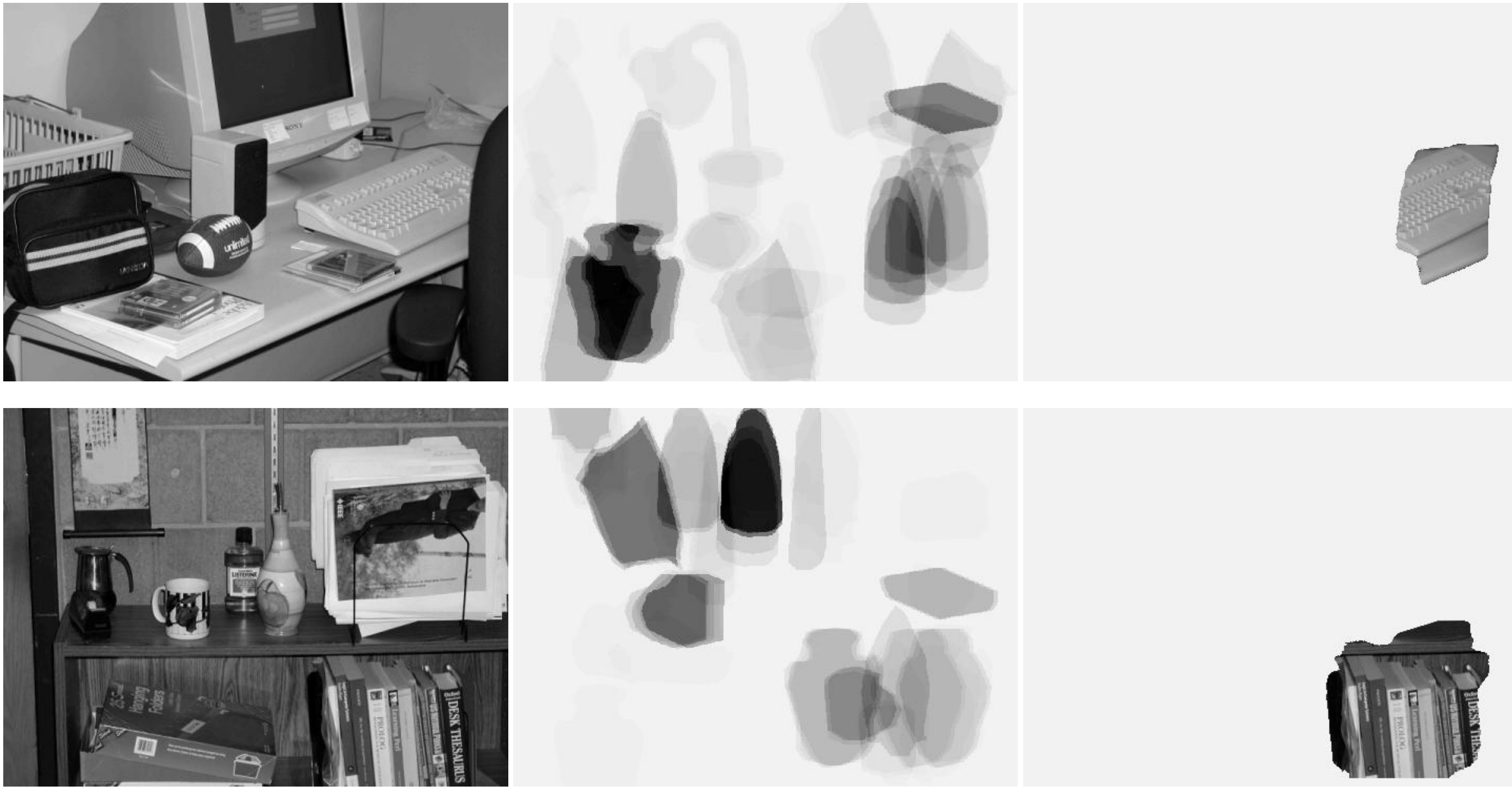
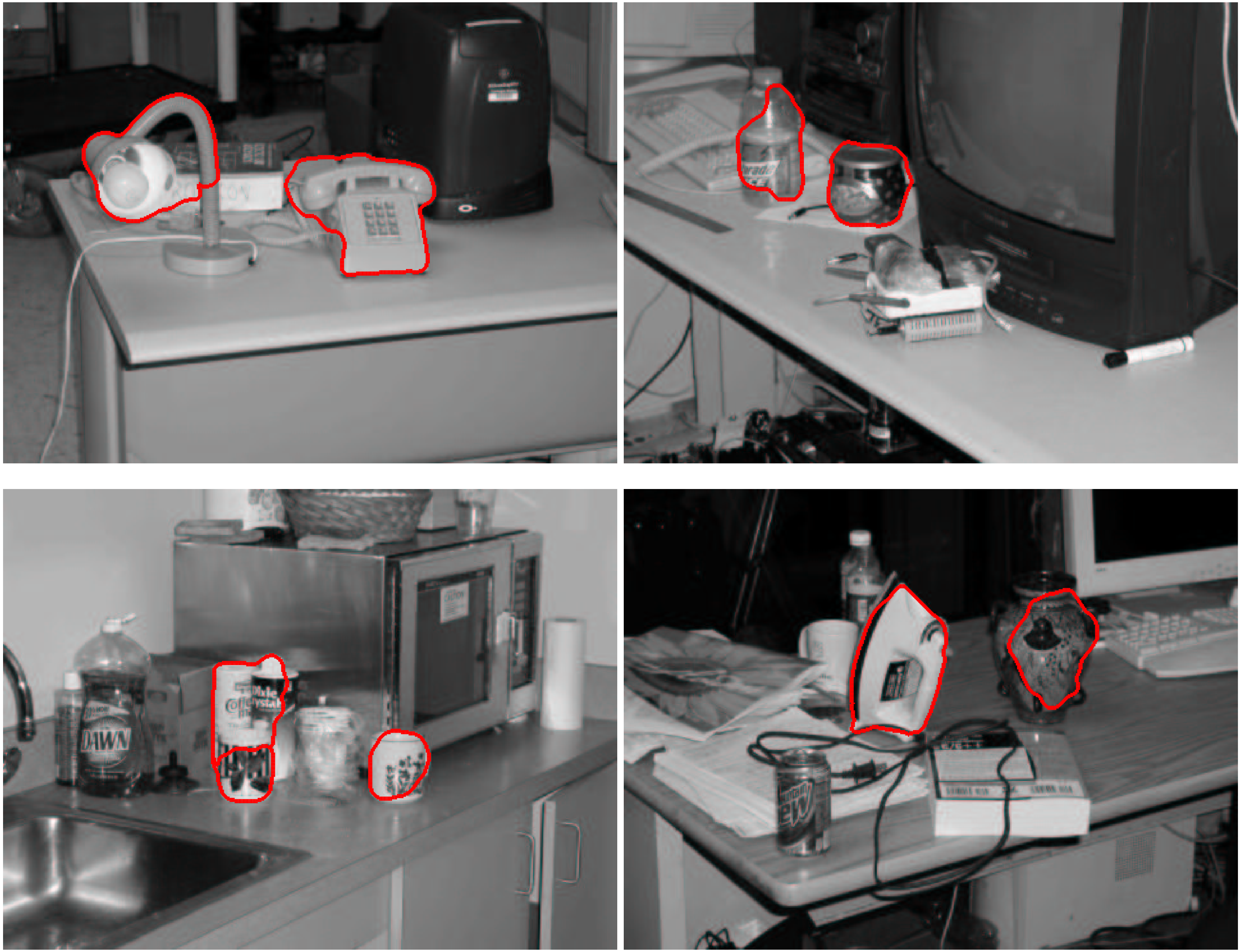


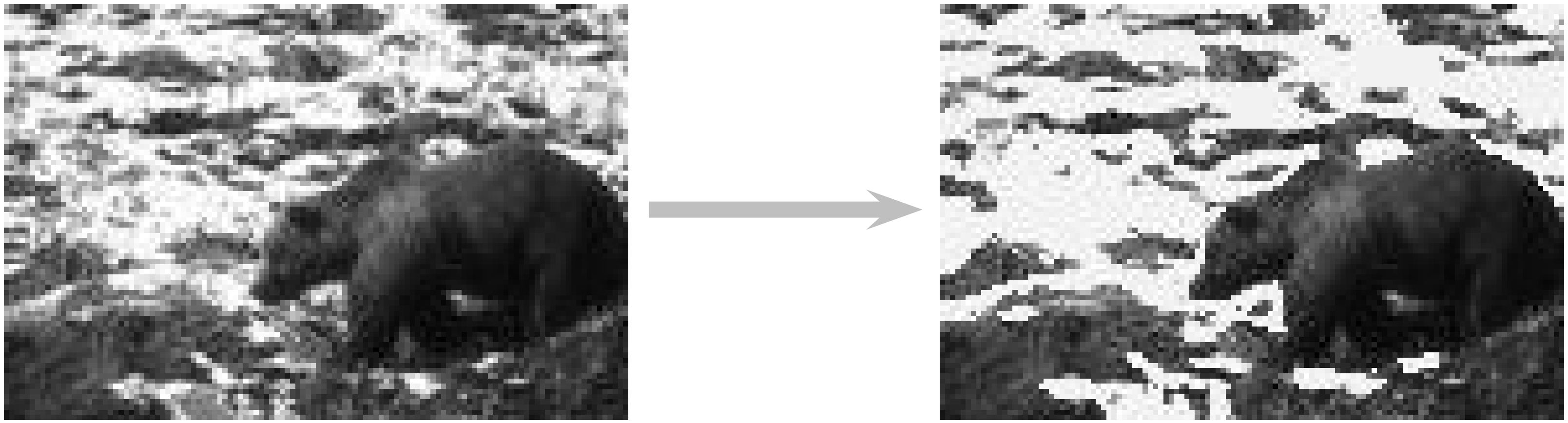
image patch density segmentation

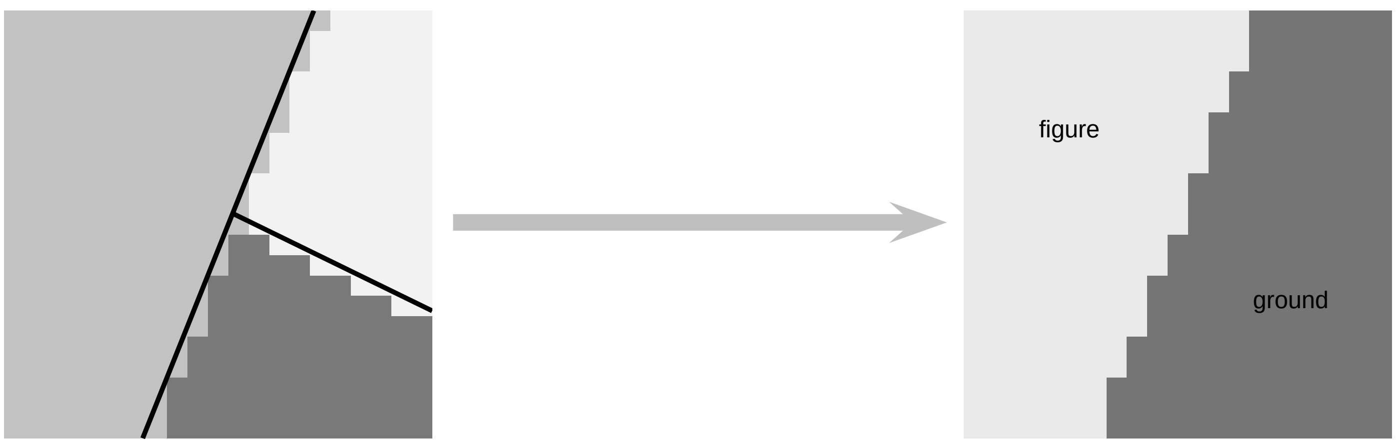
Equally Applicable to Multiple Objects



Contributions to Perceptual Organization

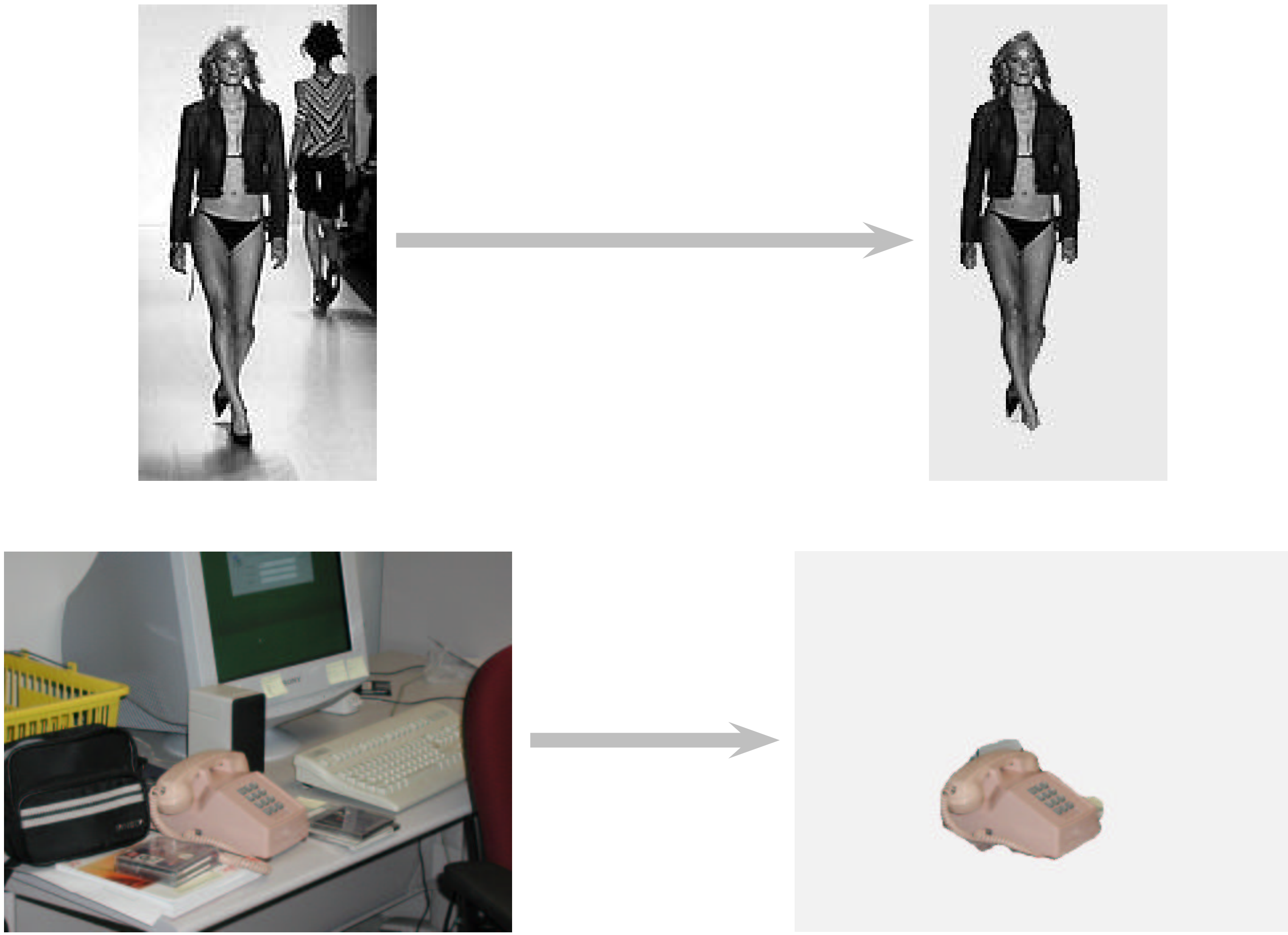
1. grouping and figure-ground in one framework





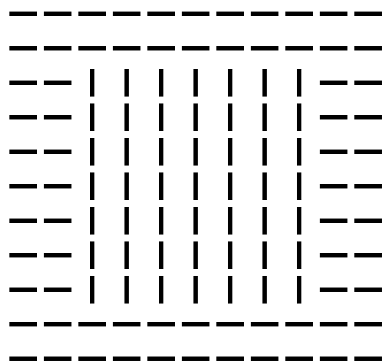
Contributions to Perceptual Organization

2. grouping integrated with spatial and object attention

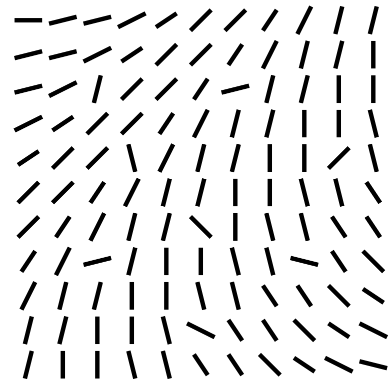


Contributions to Graph Theory

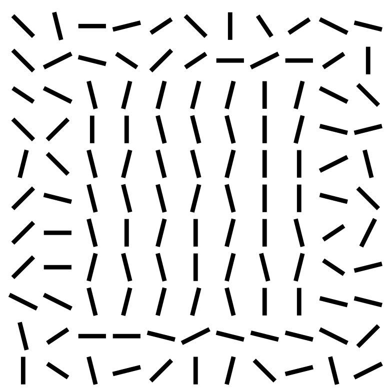
1. new grouping cues



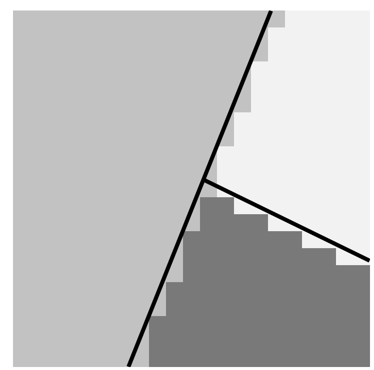
attraction



repulsion



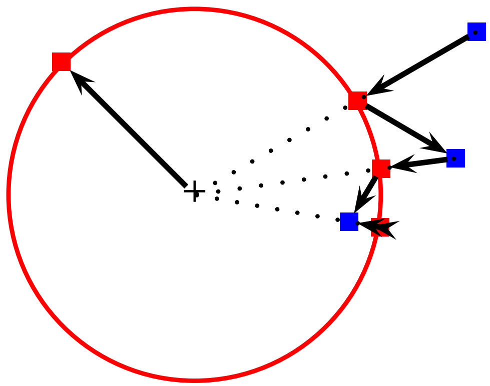
regularization



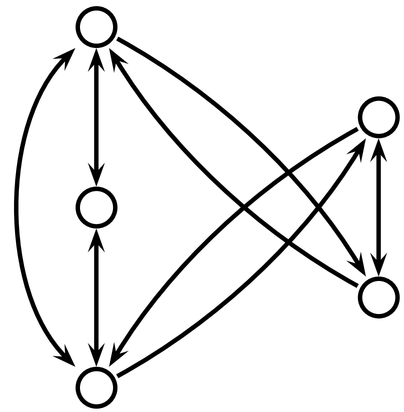
depth

Contributions to Graph Theory

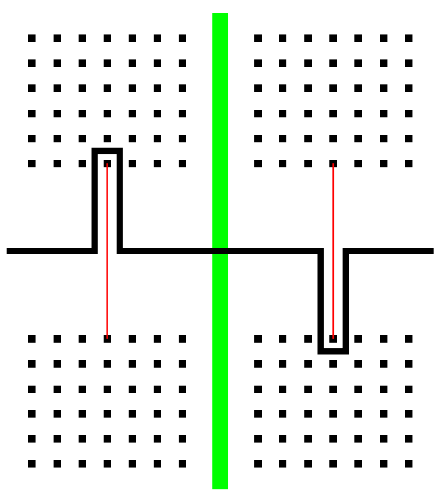
2. new graph partitioning techniques



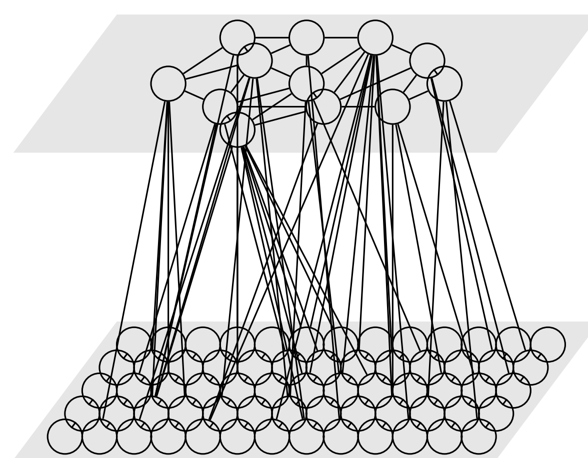
K-way cuts



directed cuts



biased cuts



joint cuts

Future Work

1. Automatic selection of the number of classes.
2. A model-based view on spectral clustering.
3. A criterion for comparing two segmentations.
4. Closing a feedback loop.
5. Object representation.
6. Scaling up.

Acknowledgements

- Jianbo Shi, Tai Sing Lee, Takeo Kanade
  Shyjan Mahamud, David Tolliver
  Jing Xiao, Vandi Verma

- CMU HumanID Lab
  CMU Hebert Lab

- ONR N00014-00-1-0915
  NSF IRI-9817496
  NSF LIS 9720350
  NSF CAREER 9984706
  NIH EY 08098