Segmentation with Pairwise Attraction and Repulsion

Stella X. Yu Jianbo Shi

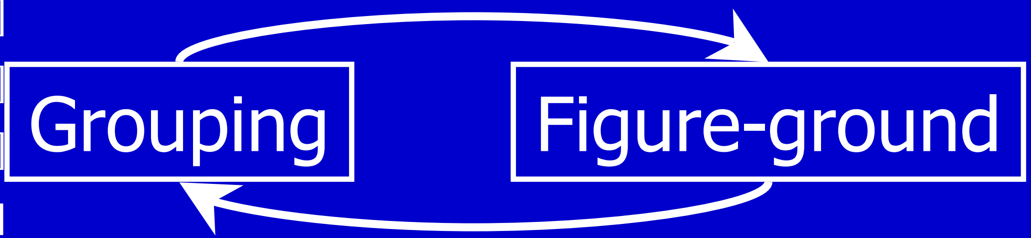
Robotics Institute
Carnegie Mellon University
{stella.yu, jshi}@cs.cmu.edu

Funded by DARPA HumanID ONR N00014-00-1-0915

- Grouping: decomposition into regions of coherent properties
- Figure-ground: decomposition into foreground and background

[Nitzberg et al 93]
[Grossberg & Mingolla 85]
[Belhumeur 96]

[Canny 86]
[Horowitz & Pavlidis 74]
[Geman & Geman 84]
[Blake & Zisserman 87]
[Zhu & Yullie, 95]
[Shi & Malik 97]

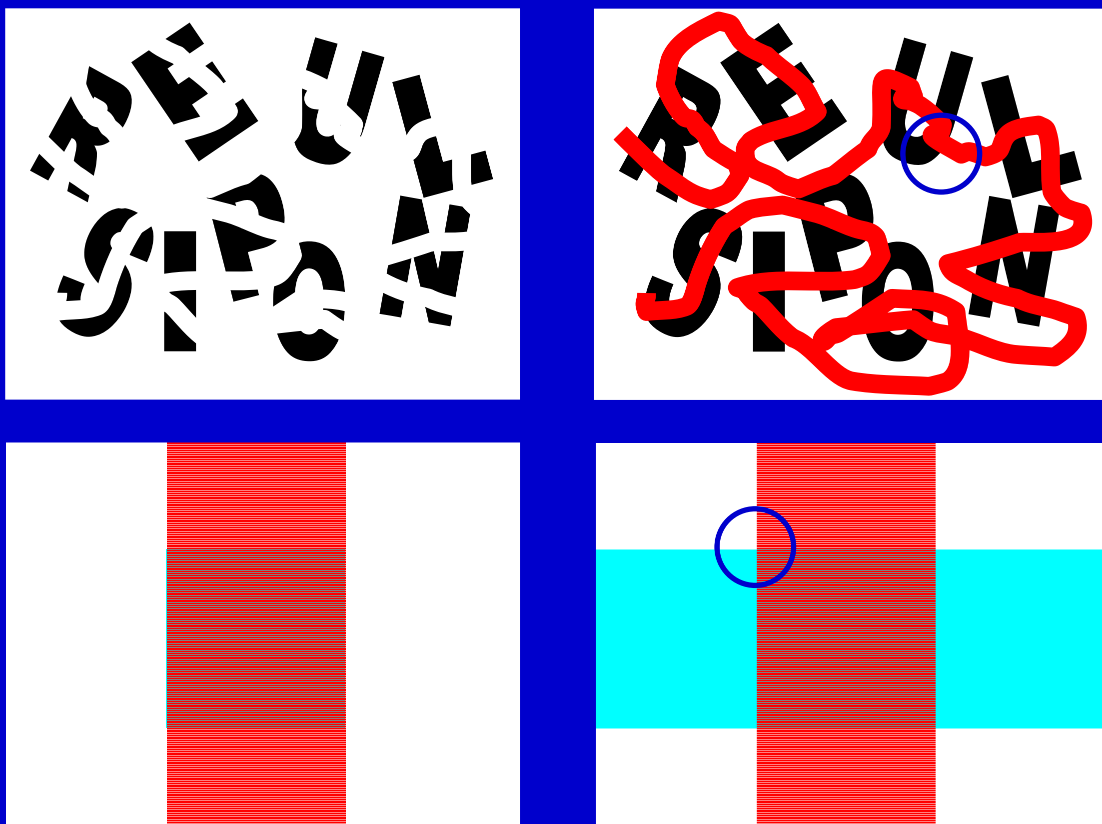


[Herault and Horaud, 93]
[Chou, 95]
[Geiger & Kumaran, 96]
[Williams & Jocobs, 97]
[Amir & Lindenbaum, 98
[Liu & Wang, 99]

- Issues
  - Fundamental: segmentation and figure-ground reasoning is circular
  - Practical: lack of efficient computational techniques

Figure-ground cues for segmentation

- Figure-ground organization □ Segmentation + Figure-ground labelling
- Between-region relations can influence the internal organization of a region



Roles of figure-ground cues:

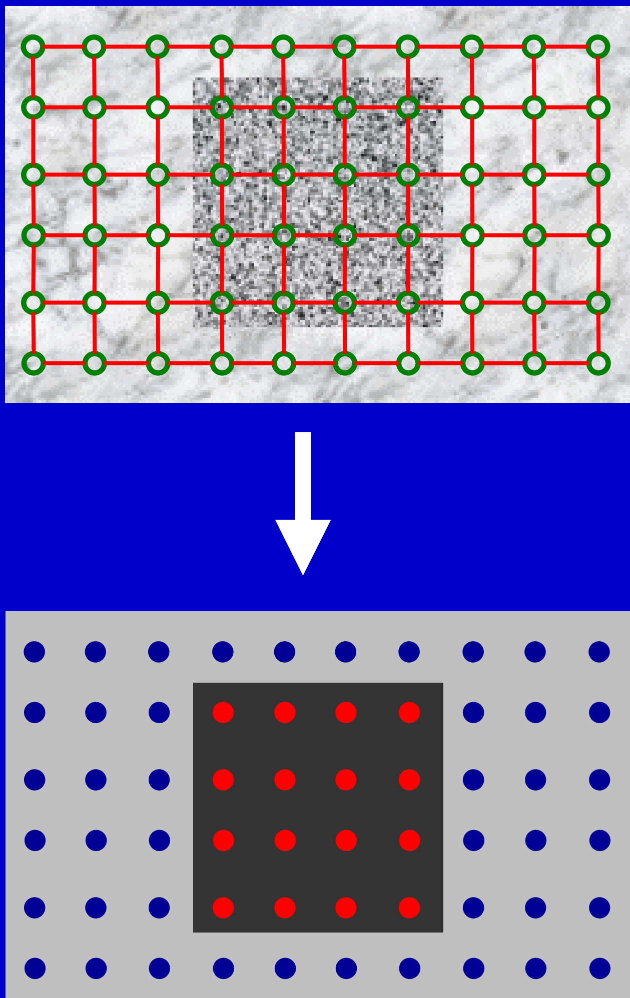
✓ Link incoherent background or foreground together

✓ Global binding through local cues

Talk overview

- Segmentation in a graph framework
  - Ordered graph partitioning
- Representing pairwise relationships
  - Attraction and repulsion
- Criteria
  - Optimize global configuration of figure vs ground
- Energy formulation
  - Rayleigh quotient of Hermitian matrices
- Efficient solution technique
  - Sparse matrix eigendecomposition
- Results

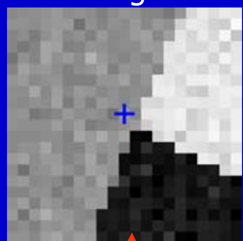
Segmentation in a graph framework



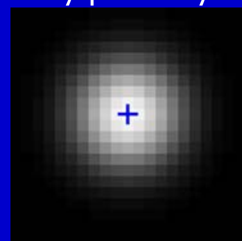
- G=(V, E, W)
  - ➔ V: each node denotes a pixel
  - ➔ E: each edge denotes a pixel-pixel relationship
  - ➔ W: each weight measures pairwise affinity
- Segmentation = vertex partitioning
  - ➔ break V into disjoint sets V_1 , V_2
- Pairwise affinity, W:
  - ➔ similarity of pixel attributes
  - ➔ relative order of figure-ground
- Goal: ordered partitioning
  - ➔ V_1= figure, V_2 = ground
- Criteria:
  - ➔ within-region similarity: high
  - ➔ between-region similarity: low
  - ➔ order from figure to ground: high

Generalized affinity: attraction and repulsion

image

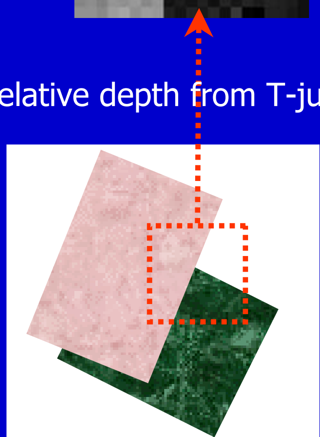


A by proxmity

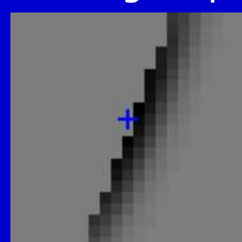


A by brightness similarity

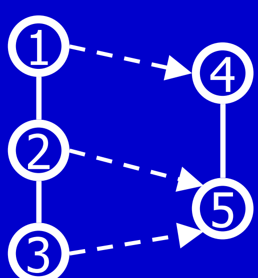
Relative depth from T-junctions



R for a figure pixel



R for a ground pixel



A is symmetric

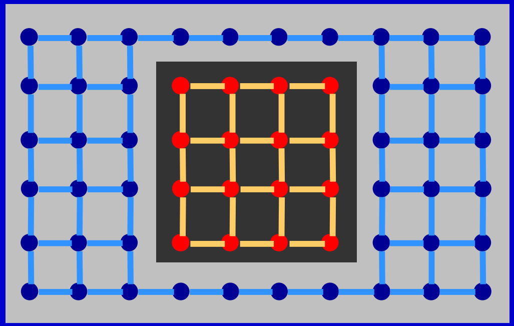
$$A = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}$$

R is skew-symmetric

$$R = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 & 0 \\ 0 & -1 & -1 & 0 & 0 \end{bmatrix}$$

- Attraction A
  - ➔ feature similarity
  - ➔ feature = location, brightness, textureness, motion ,…
  - ➔ the larger the A_{jk}, the more likely j and k in one region
- Directional repulsion R
  - ➔ relative order of figure-ground
  - ➔ the larger the $|R_{jk}|$, the more likely j and k in different regions
- Generalized Affinity W = A + *i* R
  - ➔ A 5-node graph example
  - ➔ W is Hermitian

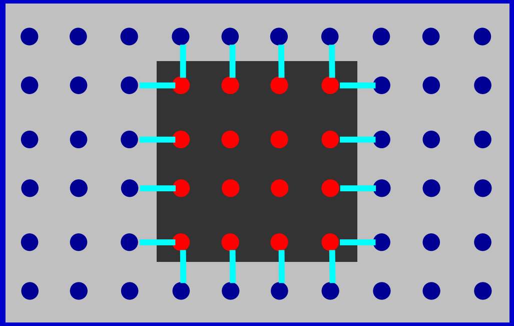
Criteria: cuts, associations and difference flows



- within-region similarity

= within-group attraction
= associations

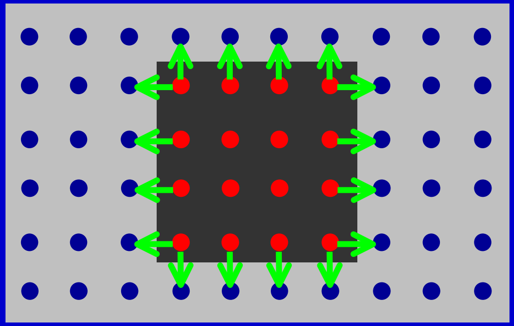
$$assoc(P) = \sum_{u \in P, v \in P} A(u,v)$$



- between-region similarity

= between-group attraction
= cuts

$$cut(P,Q) = \sum_{u \in P, v \in Q} A(u,v), \quad P \cap Q = \Phi$$



- order from figure to ground

= figure-to-ground repulsion
= difference flow

$$difflow(P,Q) = \sum_{u \in P, v \in Q} R(u,v), \quad P \cap Q = \Phi$$

Criteria: generalized normalized association

- Maximize within-region similarity and figure-to-ground order

$$\text{Nassoc} = \frac{\text{within figure similarity}}{\text{total figure affinity}} + \frac{\text{within ground similarity}}{\text{total ground affinity}} + \frac{\text{figure} \sim \text{ground order}}{\text{average total affinity}}$$

$$Nassoc(V_1, V_2) = \frac{assoc(V_1)}{\deg(V_1)} + \frac{assoc(V_2)}{\deg(V_2)} + \frac{2 \cdot difflow\,(V_1, V_2)}{\sqrt{\deg(V_1) \cdot \deg(V_2)}}$$

Criteria: generalized normalized cuts

- Minimize between-region similarity and ground-to-figure order

$$\text{Ncut} = \frac{\text{figure} \sim \text{ground similarity}}{\text{total figure affinity}} + \frac{\text{figure} \sim \text{ground similarity}}{\text{total ground affinity}} + \frac{\text{ground} - \text{figure order}}{\text{average total affinity}}$$

$$Ncut(V_1, V_2) = \frac{cut(V_1, V_2)}{\deg(V_1)} + \frac{cut(V_2, V_1)}{\deg(V_2)} + \frac{2 \cdot difflow(V_2, V_1)}{\sqrt{\deg(V_1) \cdot \deg(V_2)}}$$

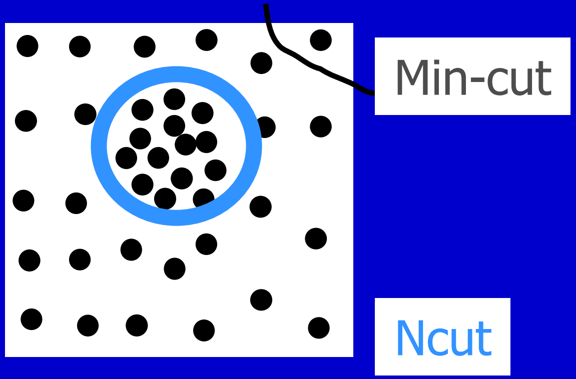
Properties of the criteria

- Duality to achieve both goals at the same time

$$\text{Ncut} + \text{Nassoc} = \text{constant}$$

$$\arg\min Ncut = \arg\max Nassoc$$

- Normalization for global structures



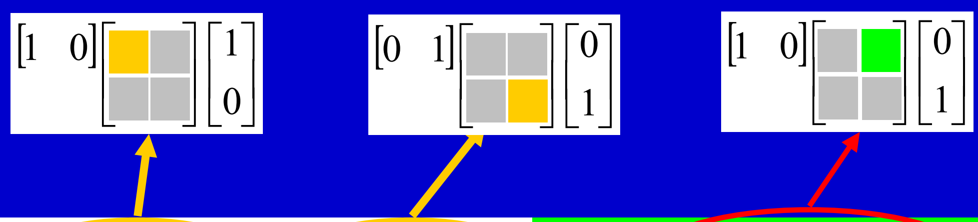
Energy formulation

- Variables: group indicators

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

$$X_1 = \begin{bmatrix} 1_{n_1 \times 1} \\ 0_{n_2 \times 1} \end{bmatrix}, \quad X_2 = \begin{bmatrix} 0_{n_1 \times 1} \\ 1_{n_2 \times 1} \end{bmatrix}$$

- Energy functions of indicator variables



$$Nassoc(X_1, X_2) = \frac{X_1^T A X_1}{X_1^T D X_1} + \frac{X_2^T A X_2}{X_2^T D X_2} + \frac{2 X_1^T R X_2}{\sqrt{X_1^T D X_1 \cdot X_2^T D X_2}}$$

Energy formulation

$$Nassoc(X_1, X_2) = \frac{X_1^T A X_1}{X_1^T D X_1} + \frac{X_2^T A X_2}{X_2^T D X_2} + \frac{2 X_1^T R X_2}{\sqrt{X_1^T D X_1 \cdot X_2^T D X_2}}$$

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

- Change of variables

$$Nassoc(y) = \frac{y^H (A + i \cdot R) y}{y^H D y}$$

$$s.t. \quad y^{2T} D1 = 0$$

- Rayleigh quotient

Three aspects of solutions

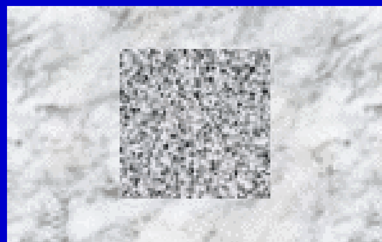
- Efficient solutions in the continuous domain:
  - ➔ Eigenvector corresponding to the largest eigenvalue of (A+*i* R, D)

$$y_{opt} = \arg\max \frac{y^H (A + i \cdot R) y}{y^H D y} \Rightarrow (A + i \cdot R) y_{opt} = \lambda_1 D y_{opt}$$

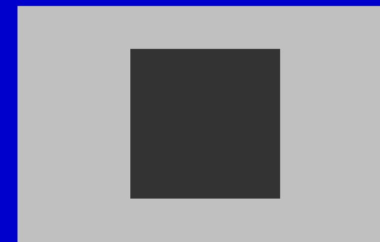
- Little increase in complexity
  - ➔ Hermitian matrices preserve the most important properties of real symmetric matrices in eigendecomposition.
  - ➔ Sparse matrix eigendecomposition $O(n^{3/2})$, n = # of pixels

- Interpretation of complex-valued solutions
  - ➔ Separation in phase plane indicates the separation of groups
  - ➔ Relative phase advance indicates figure-ground ordering

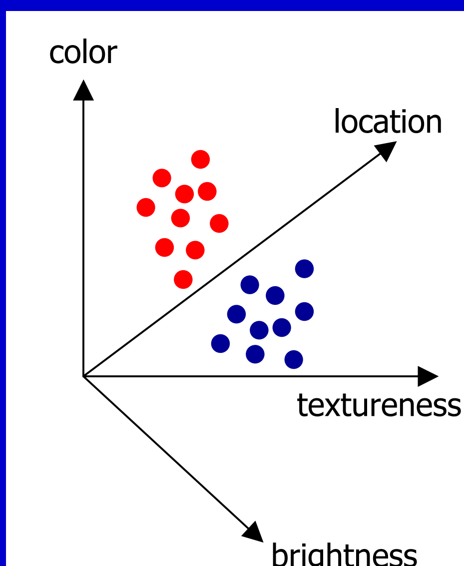
Segmentation as embedding



Assign region identity.



Every pixel corresponds to a point in a high dimensional feature space.



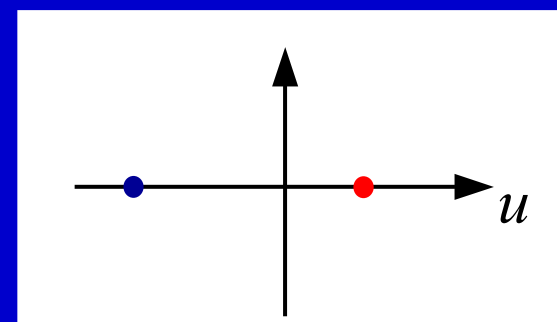
$$y_{opt} = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}$$

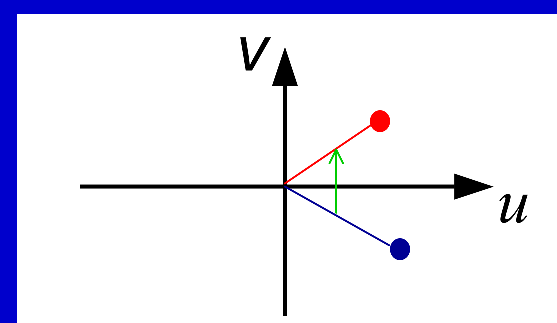
grouping

figure-ground

$$y_{opt} = \begin{bmatrix} u_1 + i \cdot v_1 \\ \vdots \\ u_n + i \cdot v_n \end{bmatrix}$$

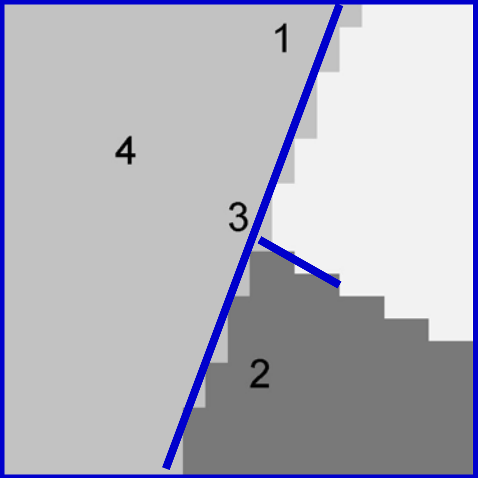
Segmentation is about finding an embedding of manifolds into some 1D or 2D space.



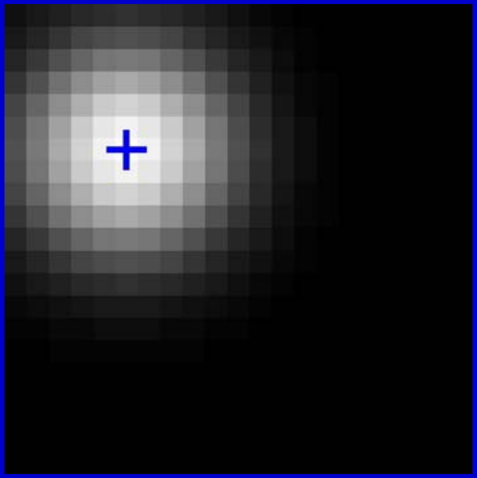


Interaction of attraction and repulsion

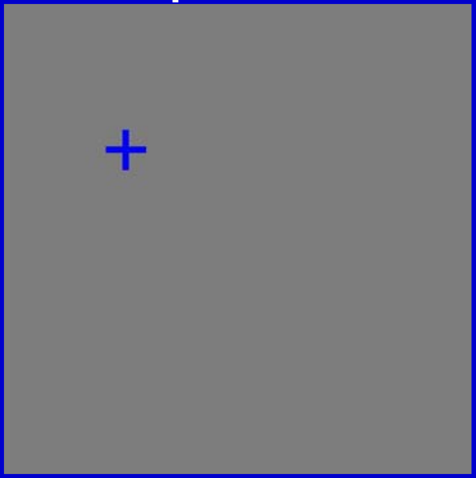
Image



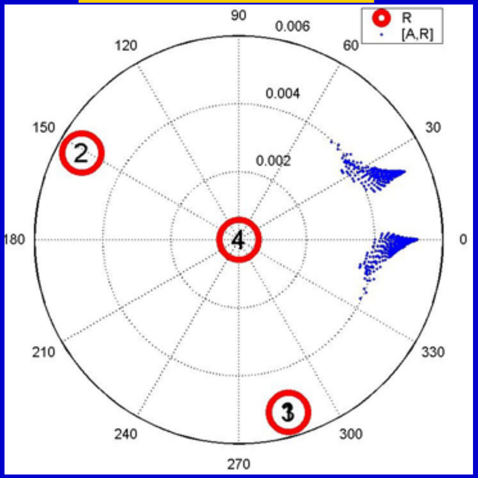
Attraction



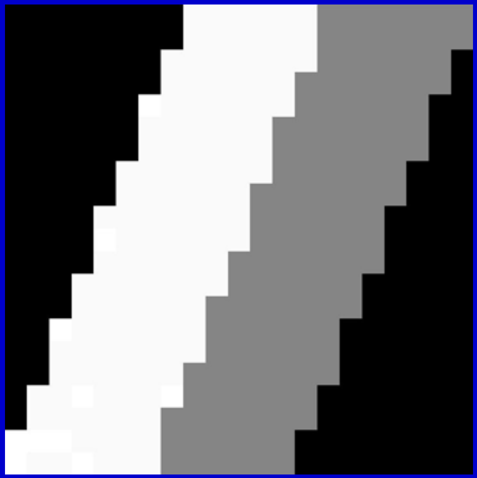
Repulsion



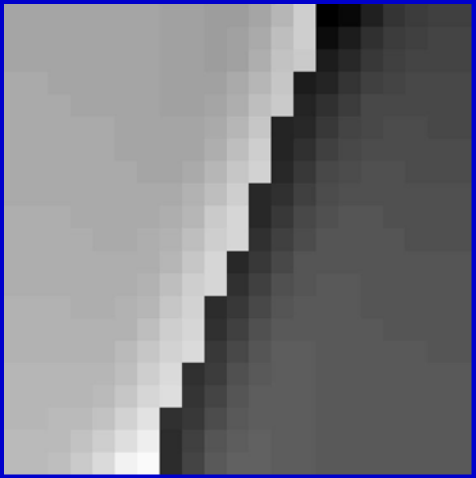
Phase Plot



Result: R



Result: A and R



Objects ordered in depth

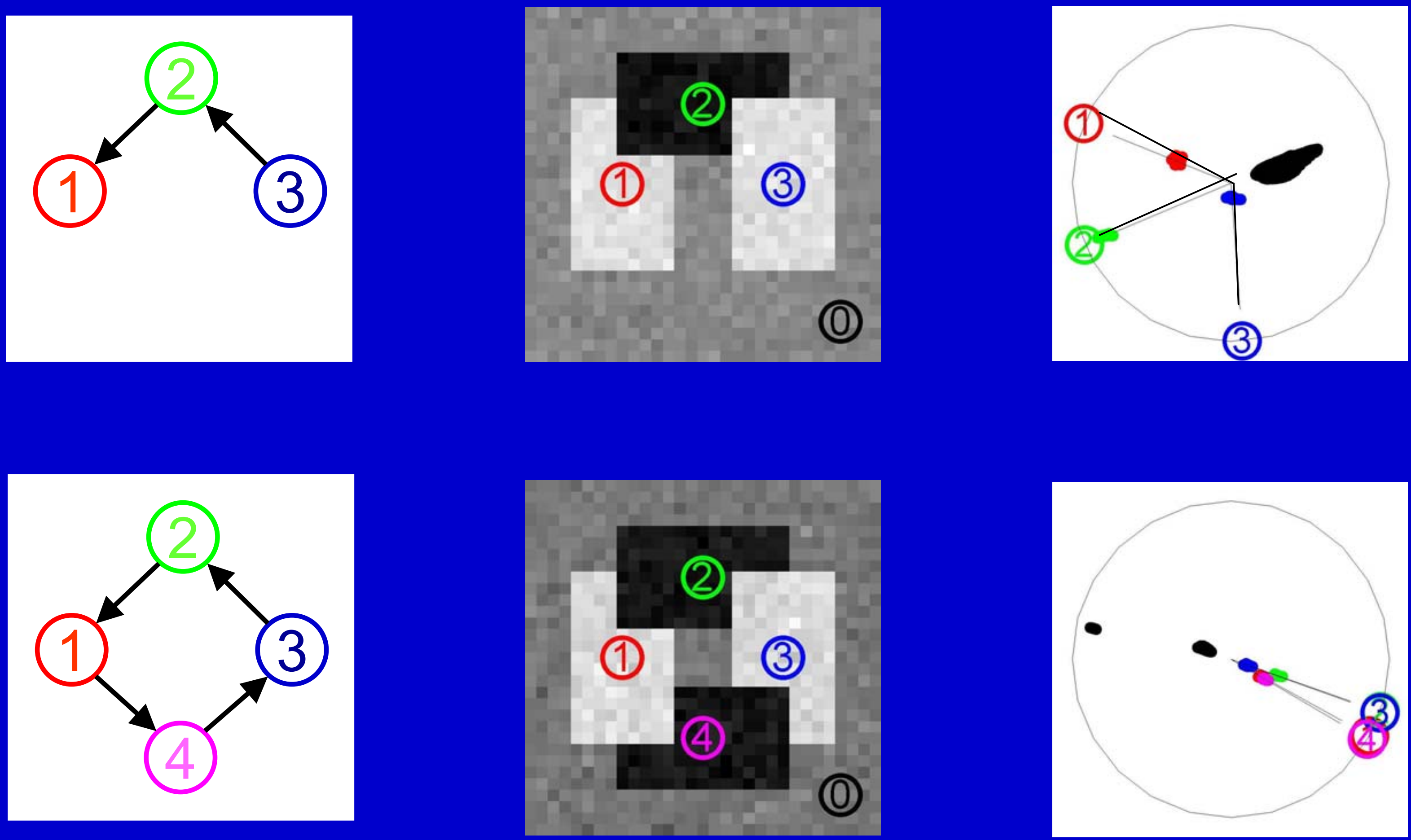
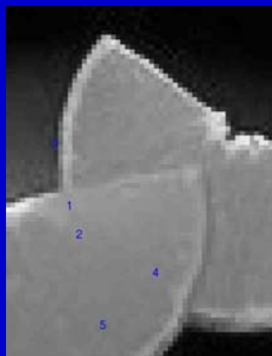
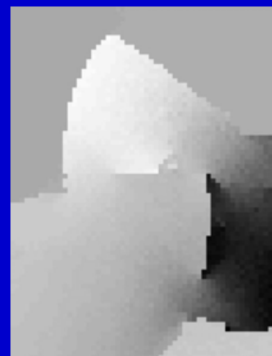


Figure-ground example

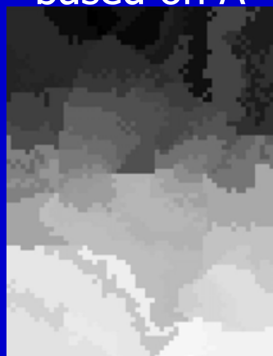
Image



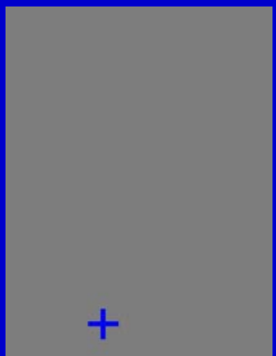
Solutions: A



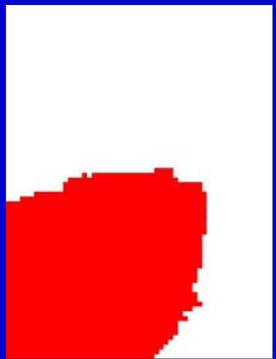
Oversegmentation based on A



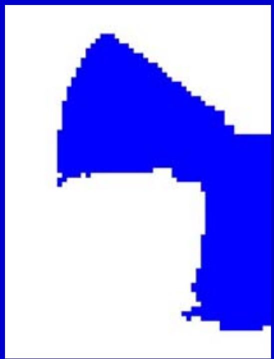
R



Figure

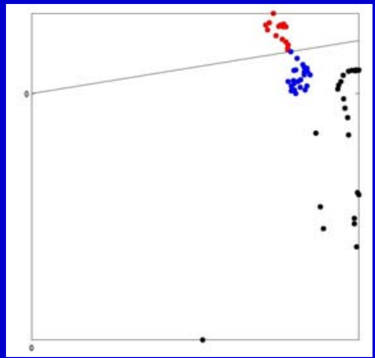


Ground 1



Ground 2

Phase plot



Conclusions

- Representation
  - Attraction + directional repulsion
- Formulation
  - Ordered graph partitioning for figure and ground
  - Rayleigh quotient of Hermitian matrices
- Solution
  - Efficient algorithm by eigendecomposition
  - Embedding in the phase plane of complex numbers
- Benefits
  - Grouping and figure-ground in one step
  - Computational efficiency