Grouping with Directed Relationships

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

[1] Robotics Institute
Carnegie Mellon University
[2] Center for the Neural Basis of Cognition
5000 Forbes Ave, Pittsburgh, PA 15213-3890
{stella.yu, jshi}@cs.cmu.edu

Abstract. Grouping is a global partitioning process that integrates local cues distributed over the entire image. We identify four types of pairwise relationships, attraction and repulsion, each of which can be symmetric or asymmetric. We represent these relationships with two directed graphs. We generalize the normalized cuts criteria to partitioning on directed graphs. Our formulation results in Rayleigh quotients on Hermitian matrices, where the real part describes undirected relationships, with positive numbers for attraction, negative numbers for repulsion, and the imaginary part describes directed relationships. Globally optimal solutions can be obtained by eigendecomposition. The eigenvectors characterize the optimal partitioning in the complex phase plane, with phase angle separation determining the partitioning of vertices and the relative phase advance indicating the ordering of partitions. We use directed repulsion relationships to encode relative depth cues and demonstrate that our method leads to simultaneous image segmentation and depth segregation.

1 Introduction

The grouping problem emerges from several practical applications including image segmentation, text analysis and data mining. In its basic form, the problem consists of extracting from a large number of data points, i.e., pixels, words and documents, the overall organization structures that can be used to summarize the data. This allows one to make sense of extremely large sets of data. In human perception, this ability to group objects and detect patterns is called perceptual organization. It has been clearly demonstrated in various perceptual modalities such as vision, audition and somatosensation [6].

To understand the grouping problem, we need to answer two basic questions: 1) what is the right criterion for grouping? 2) how to achieve the criterion computationally? At an abstract level, the criterion for grouping seems to be clear. We would like to partition the data so that elements are well related within groups but decoupled between groups. Furthermore, we prefer grouping mechanisms that provide a clear organization structure of the data. This means to extract big pictures of the data first and then refine them.

To achieve this goal, a number of computational approaches have been proposed, such as clustering analysis through agglomerative and divisive algorithms [5], greedy region growing, relaxation labeling [13], Markov random fields (MRF) [4] and variational formulations [2, 7, 9]. While the greedy algorithms are computationally efficient, they can only achieve locally optimal solutions. Since grouping is about finding the global structures of the data, they fall short of this goal. MRF formulations, on the other hand, provide a global cost function incorporating all local clique potentials evaluated on nearby data points. These clique potentials can encode a variety of configuration constraints and probability distributions [18]. One shortcoming of these approaches is a lack of efficient computational solutions.

Recently we have seen a set of computational grouping methods using local pairwise relationships to compute global grouping structures [1, 3, 12, 11, 14, 15, 16]. These methods share a similar goal of grouping with MRF approaches, but they have efficient computational solutions. It has been demonstrated that they work successfully in the segmentation of complex natural images [8].

However, these grouping approaches are somewhat handicapped by the very representation that makes them computationally tractable. For example, in graph formulations [16, 15, 3, 11], negative correlations are avoided because negative edge weights are problematic for most graph algorithms. In addition, asymmetric relationships such as those that arise from figure-ground cues in image segmentation and web-document connections in data mining cannot be considered because of the difficulty in formulating a global criterion with efficient solutions.

In this paper, we develop a grouping method in the graph framework that incorporates pairwise negative correlation as well as asymmetric relationships. We propose a representation in which all possible pairwise relationships are characterized in two types of directed graphs, each encoding positive and negative correlations between data points. We generalize the dual grouping formulation of normalized cuts and associations to capture directed grouping constraints. We show that globally optimal solutions can be obtained by solving generalized eigenvectors of Hermitian weight matrices in the complex domain. The real and imaginary parts of Hermitian matrices encode undirected and directed relationships respectively. The phase angle separation defined by the eigenvectors in the complex plane determines the partitioning of data points, and the relative phase advance indicates the ordering of partitions.

The rest of the paper is organized as follows. Section 2 gives a brief review of segmentation with undirected graphs in the normalized cuts formulation. Section 3 expands our grouping method in detail. Section 4 illustrates our ideas and methods on synthetic data. Section 5 concludes the paper.

2 Review on Grouping on One Undirected Graph

The key principles of grouping can often be illustrated in the context of image segmentation. In graph methods for image segmentation, an image is described by an undirected weighted graph $\mathsf{G} = (\mathsf{V}, \mathsf{E}, W)$, where each pixel is a vertex in V and the likelihood of two pixels belonging to one group is described by a weight in W associated with the edge in E between two vertices. The weights are computed from a pairwise similarity function of image attributes such as intensity, color and motion profiles. Such similarity relationships are symmetric and can be considered as mutual attraction between vertices.

After an image is transcribed into a graph, image segmentation becomes a vertex partitioning problem. A good segmentation is the optimal partitioning scheme according to some partitioning energy functions, evaluating how heavily each group is internally connected (associations) and/or how weakly those between-group connections (cuts) are. We are particularly interested in the normalized associations and cuts criteria [15], for they form a duality pair such that the maximization of associations automatically leads to the minimization of cuts and vice versa.

A vertex bipartitioning $(\mathsf{V}_1, \mathsf{V}_2)$ on graph $\mathsf{G} = (\mathsf{V}, \mathsf{E})$ has $\mathsf{V} = \mathsf{V}_1 \cup \mathsf{V}_2$ and $\mathsf{V}_1 \cap \mathsf{V}_2 = \varnothing$. Given weight matrix W and two vertex sets P and Q, let $\mathcal{C}_W(P, Q)$ denote the total W connections from P to Q,

$$\mathcal{C}_W(P, Q) = \sum_{j \in P, k \in Q} W(j, k).$$

In particular, $\mathcal{C}_W(\mathsf{V}_1, \mathsf{V}_2)$ is the total weights cut by the bipartitioning, whereas $\mathcal{C}_W(\mathsf{V}_l, \mathsf{V}_l)$ is the total association among vertices in V_l, $l = 1, 2$. Let $\mathcal{D}_W(P)$ denote the total outdegree of P,

$$\mathcal{D}_W(P) = \mathcal{C}_W(P, \mathsf{V}),$$

which is the total weights connected to all vertices in a set P. Let $\mathcal{S}_W(P, Q)$ denote the connection ratio from P to Q,

$$\mathcal{S}_W(P, Q) = \frac{\mathcal{C}_W(P, Q)}{\mathcal{D}_W(P)}.$$

In particular, $\mathcal{S}_W(\mathsf{V}_l, \mathsf{V}_l)$ is called the *normalized association* of vertex set V_l as it is the association normalized by its degree of connections. Likewise, $\mathcal{S}_W(\mathsf{V}_1, \mathsf{V}_2)$ is called the *normalized cuts* between V_1 and V_2. The sum of these ratios respectively over two partitions are denoted by

$$\begin{aligned} \epsilon_a &= \textstyle\sum_{l=1}^{2} \mathcal{S}_W(\mathsf{V}_l, \mathsf{V}_l), \\ \epsilon_c &= \textstyle\sum_{l=1}^{2} \mathcal{S}_W(\mathsf{V}_l, \mathsf{V} \setminus \mathsf{V}_l). \end{aligned}$$

ϵ_a and ϵ_c are called normalized associations and cuts criteria. Since $\forall l$, $\mathcal{S}_W(\mathsf{V}_l, \mathsf{V}_l) + \mathcal{S}_W(\mathsf{V}_l, \mathsf{V} \setminus \mathsf{V}_l) = 1$, $\epsilon_a + \epsilon_c = 2$, thus ϵ_a and ϵ_c are dual criteria: maximizing ϵ_a is equivalent to minimizing ϵ_c. We seek the optimal solution maximizing ϵ_a such that within-group associations are maximized and between-group cuts are minimized.

The above criteria can be written as Rayleigh quotients of partitioning variables. Let X_l be a membership indicator vector for group l, $l = 1, 2$, where $X_l(j)$

assumes 1 if vertex j belongs to group l and 0 otherwise. Let D_W be the diagonal degree matrix of the weight matrix W, $D_W(j,j) = \sum_k W_{jk}, \forall j$. Let $\mathsf{1}$ denote the all-one vector. Let k denote the degree ratio of V_1: $k = \frac{X_1^T D_W X_1}{\mathsf{1}^T D_W \mathsf{1}}$. We define $y = (1-k)\,X_1 - k\,X_2$. Therefore, the optimization problem becomes:

$$\max \epsilon_a = \frac{y^T W y}{y^T D_W y} + 1; \qquad \min \epsilon_c = \frac{y^T (D_W - W) y}{y^T D_W y},$$
$$\text{s. t.} \quad y^T D_W \mathsf{1} = 0; \qquad y_j \in \{1-k, -k\}, \forall j.$$

When the discreteness constraint is relaxed, the second largest generalized eigenvector of (W, D_W) maximizes ϵ_a subject to the zero-sum constraint $y^T D_W \mathsf{1} = 0$. For eigensystem $M_1 y = \lambda M_2 y$ of a matrix pair (M_1, M_2), let $\lambda(M_1, M_2)$ be the set of distinctive generalized eigenvalues λ and $\Upsilon(M_1, M_2, \lambda)$ be the eigenspace of y. It can be shown that $\forall \lambda \in \lambda(W, D_W)$, $|\lambda| \le 1$. Let λ_k denote the k-th largest eigenvalue, then $\lambda_1 = 1$ and $\mathsf{1} \in \Upsilon(M_1, M_2, \lambda_1)$. Thus the optimal solution is:

$$\epsilon_a(y_{opt}) = 1 + \lambda_2, \qquad y_{opt} \in \Upsilon(W, D_W, \lambda_2).$$

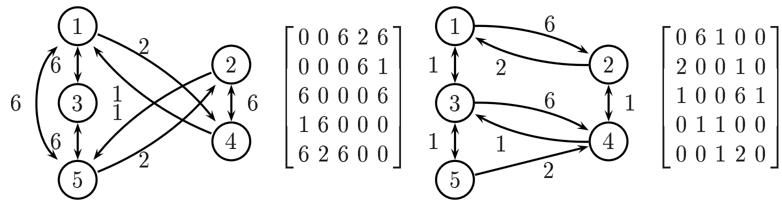
3 Grouping on Two Directed Graphs

The above formulation addresses the grouping problem in a context where we can estimate the *similarity* between a pair of pixels. This set of relationships arises naturally in color, texture and motion segmentation. However, a richer set of pairwise relationships exists in a variety of settings. For example, relative depth cues suggest that two pixels should not belong to the same group; in fact, one of them is more likely to be figure and the other is then the ground. Compared to the similarity measures, this example encapsulates two other distinct attributes in pairwise relationships: repulsion and asymmetry. This leads to a generalization of the above grouping model in two ways. One is to have dual measures of attraction and repulsion, rather than attraction alone; the other is to have directed graph partitioning, rather than symmetric undirected graph partitioning.

3.1 Representation

We generalize the single undirected graph representation for an image to two directed graph representations $\mathsf{G} = \{\mathsf{G_A}, \mathsf{G_R}\}$: $\mathsf{G_A} = (\mathsf{V}, \mathsf{E_A}, A)$, $\mathsf{G_R} = (\mathsf{V}, \mathsf{E_R}, R)$, encoding pairwise attraction and repulsion relationships respectively. Both A and R are *nonnegative* weight matrices. Since $\mathsf{G_A}$ and $\mathsf{G_R}$ are directed, A and R can be asymmetric. An example is given in Fig. 1.

Whereas directed repulsion can capture the asymmetry between figure and ground, directed attraction can capture the general compatibility between two pixels. For example, a reliable structure at one pixel location might have a higher affinity with a structure at another location, meaning the presence of the former is more likely to attract the latter to the same group, but not the other way around.



a. $\mathsf{G_A} = (\mathsf{V}, \mathsf{E_A}, A)$ b. $\mathsf{G_R} = (\mathsf{V}, \mathsf{E_R}, R)$

Fig. 1. Two directed graph representation of an image. a. Didrected graph with nonnegative asymmetric weights for attraction. b. Directed graph with nonnegative asymmetric weights for repulsion.

3.2 Criteria

To generalize the criteria on nondirectional attraction to directed dual measures of attraction and repulsion, we must address three issues.

1. Attraction vs. repulsion: how do we capture the semantic difference between attraction and repulsion? For attraction A, we desire the association within groups to be as large as possible; whereas for repulsion R, we ask the segregation by between-group repulsion to be as large as possible.
2. Undirected vs. directed: how do we characterize a partitioning that favors between-group relationships in one direction but not the other? There are two aspects of this problem. The first is that we need to evaluate within-group connections regardless of the asymmetry of internal connections. This can be done by partitioning based on its undirected version so that within-group associations are maximized. The second is that we need to reflect our directional bias on between-group connections. The bias favoring weights associated with edges pointing from V_1 to V_2 is introduced by an asymmetric term that appreciate connections in $\mathcal{C}_W(\mathsf{V}_1, \mathsf{V}_2)$ but discourage those in $\mathcal{C}_W(\mathsf{V}_2, \mathsf{V}_1)$. For these two purposes, we decompose $2\,W$ into two terms:

$$2\,W = W_u + W_d, \qquad W_u = (W + W^T), \qquad W_d = (W - W^T).$$

   W_u is an undirected version of graph $\mathsf{G_W}$, where each edge is associated with the sum of the W weights in both directions. The total degree of the connections for an asymmetric W is measured exactly by the outdegree of W_u. W_d is a skew-symmetric matrix representation of W, where each edge is associated with the weight difference of W edges pointing in opposite directions. Their links to W are formally stated below:

$$\begin{aligned} \mathcal{C}_{W_u}(P,Q) &= \mathcal{C}_W(P,Q) + \mathcal{C}_W(Q,P) = \quad \mathcal{C}_{W_u}(Q,P), \\ \mathcal{C}_{W_d}(P,Q) &= \mathcal{C}_W(P,Q) - \mathcal{C}_W(Q,P) = -\mathcal{C}_{W_d}(Q,P). \end{aligned}$$

   This decomposition essentially turns our original graph partitioning on two directed graphs of attraction and repulsion into a simultaneous partitioning on four graphs of nondirectional attraction and repulsion, and directional attraction and repulsion.

3. Integration: how do we integrate the partitioning on such four graphs into one criterion? We couple connection ratios on these graphs through linear combinations. The connection ratios of undirected graphs of A and R are first combined by linear weighting with their total degrees of connections. The connection ratios of directed graphs are defined by the cuts normalized by the geometrical average of the degrees of two vertex sets. The total energy function is then the convex combination of two types of connection ratios for undirected and directed partitioning, with a parameter β determining their relative importance.

With directed relationships, we seek an ordered bipartitioning $(\mathsf{V}_1, \mathsf{V}_2)$ such that the net directed edge flow from V_1 to V_2 is maximized. The above considerations lead to the following formulation of our criteria.

$$\begin{aligned}
\epsilon_a(A, R; \beta) &= 2\beta \cdot \sum_{l=1}^{2} \frac{\mathcal{S}_{A_u}(\mathsf{V}_l, \mathsf{V}_l)\mathcal{D}_{A_u}(\mathsf{V}_l) + \mathcal{S}_{R_u}(\mathsf{V}_l, \mathsf{V} \setminus \mathsf{V}_l)\mathcal{D}_{R_u}(\mathsf{V}_l)}{\mathcal{D}_{A_u}(\mathsf{V}_l) + \mathcal{D}_{R_u}(\mathsf{V}_l)} \\
&\quad + 2(1-\beta) \cdot \frac{\mathcal{C}_{A_d+R_d}(\mathsf{V}_1, \mathsf{V}_2) - \mathcal{C}_{A_d+R_d}(\mathsf{V}_2, \mathsf{V}_1)}{\sqrt{[\mathcal{D}_{A_u}(\mathsf{V}_1) + \mathcal{D}_{R_u}(\mathsf{V}_1)] \cdot [\mathcal{D}_{A_u}(\mathsf{V}_2) + \mathcal{D}_{R_u}(\mathsf{V}_2)]}}, \\
\epsilon_c(A, R; \beta) &= 2\beta \cdot \sum_{l=1}^{2} \frac{\mathcal{S}_{A_u}(\mathsf{V}_l, \mathsf{V} \setminus \mathsf{V}_l)\mathcal{D}_{A_u}(\mathsf{V}_l) + \mathcal{S}_{R_u}(\mathsf{V}_l, \mathsf{V}_l)\mathcal{D}_{R_u}(\mathsf{V}_l)}{\mathcal{D}_{A_u}(\mathsf{V}_l) + \mathcal{D}_{R_u}(\mathsf{V}_l)} \\
&\quad + 2(1-\beta) \cdot \frac{\mathcal{C}_{A_d+R_d}(\mathsf{V}_2, \mathsf{V}_1) - \mathcal{C}_{A_d+R_d}(\mathsf{V}_1, \mathsf{V}_2)}{\sqrt{[\mathcal{D}_{A_u}(\mathsf{V}_1) + \mathcal{D}_{R_u}(\mathsf{V}_1)] \cdot [\mathcal{D}_{A_u}(\mathsf{V}_2) + \mathcal{D}_{R_u}(\mathsf{V}_2)]}}.
\end{aligned}$$

Note that the duality between ϵ_a and ϵ_c is maintained as $\epsilon_a + \epsilon_c = 4\beta$.

For undirected graphs, $\mathcal{S}_{A_u}(\mathsf{V}_l, \mathsf{V}_l)$ is the old normalized association by attraction of set V_l; $\mathcal{S}_{R_u}(\mathsf{V}_l, \mathsf{V}\setminus\mathsf{V}_l)$ is the *normalized dissociation by repulsion* of set V_l. They are summed up using weights from their total degrees of connections: $\mathcal{D}_{A_u}(\mathsf{V}_l)$ and $\mathcal{D}_{R_u}(\mathsf{V}_l)$.

For directed graphs, only the asymmetry of 'connections matters. We sum up the cross connections regardless of attraction and repulsion: $\mathcal{C}_{A_d+R_d}(\mathsf{V}_1, \mathsf{V}_2) - \mathcal{C}_{A_d+R_d}(\mathsf{V}_2, \mathsf{V}_1)$, normalized by the geometrical average of the degrees of the two involved sets. Similar to $\mathcal{S}_W(P, Q)$, this again is a unitless connection ratio.

We write the partitioning energy as functions of (A, R) to reflect the fact that for this pair of directed graphs, we favor both attractive and repulsive edge flow from V_1 to V_2. They can also be decoupled. For example, the ordered partitioning based on $\epsilon_a(A^T, R; \beta)$ favors repulsion flow from V_1 to V_2, but attraction flow from V_2 to V_1.

Finally, we sum up the two terms for undirected and directed relationships by their convex combination, with the parameter β determining their relative importance. When $\beta = 1$, the partitioning ignores the asymmetry in connection weights, while when $\beta = 0$, the partitioning only cares about the asymmetry in graph weights. When $\beta = 0.5$, both graphs are considered equally. The factor 2 is to introduced to make sure that the formula are identical to those in Section 2 for $A = A^T$ and $R = 0$, i.e., $\epsilon_a(A, 0; 0.5) + \epsilon_c(A, 0; 0.5) = 2$.

3.3 Computational solutions

It turns out that our criteria lead to Rayleigh quotients of Hermitian matrices. Let $i = \sqrt{-1}$. Let * and ^{H} denote the conjugate and conjugate transpose operators respectively. We define an equivalent degree matrix D_{eq} and equivalent Hermitian weight matrix W_{eq}, which combines symmetric weight matrix U for an equivalent undirected graph and skew-symmetric weight matrix V for an equivalent directed graph into one matrix:

$$\begin{aligned} D_{eq} &= D_{A_u} + D_{R_u}, & U &= 2\beta \cdot (A_u - R_u + D_{R_u}) = \ \ U^T, \\ W_{eq} &= U + i \cdot V = W_{eq}^H, & V &= 2(1-\beta) \cdot (A_d + R_d) \ \ = -V^T. \end{aligned}$$

We then have:

$$\epsilon_a = \sum_{l=1}^{2} \frac{X_l^T U X_l}{X_l^T D_{eq} X_l} + \frac{X_1^T V X_2 - X_2^T V X_1}{\sqrt{X_1^T D_{eq} X_1 \cdot X_2^T D_{eq} X_2}}.$$

We can see clearly what directed relationships provide in the energy terms. The first term is for undirected graph partitioning, which measures the symmetric connections within groups, while the second term is for directed graph partitioning, which measures the skew-symmetric connections between groups. Such complementary and orthogonal pairings allow us to write the criterion in a quadratic form of one matrix by using complex numbers. Let k denote degree ratio of V_1: $k = \frac{X_1^T D_{eq} X_1}{1^T D_{eq} 1}$. We define a complex vector z, the square of which becomes a real vector we used in the single graph partitioning:

$$z = \sqrt{1-k}\, X_1 - i \cdot \sqrt{k}\, X_2, \qquad z^2 = (1-k)\, X_1 - k\, X_2.$$

It can be verified that:

$$\epsilon_a = 2 \frac{z^H W_{eq} z}{z^H D_{eq} z}, \qquad \epsilon_c = 2 \frac{z^H (2\beta D_{eq} - W_{eq}) z}{z^H D_{eq} z},$$

subject to the zero-sum constraint of $(z^2)^T D_{eq} 1 = 0$. Ideally, a good segmentation seeks the solution of the following optimization problem,

$$\begin{aligned} & z_{opt} = \arg\max_z \frac{z^H W_{eq} z}{z^H D_{eq} z} \\ & \text{s.t.} \qquad (z^2)^T D_{eq} 1 = 0, \qquad \forall j, z_j \in \{\sqrt{1-k}, -i\sqrt{k}\}. \end{aligned}$$

The above formulations show that repulsion can be regarded as the extension of attraction measures to negative numbers, whereas directed measures complement undirected measures along an orthogonal dimension. This generalizes graph partitioning on a nonnegative symmetric weight matrix to an arbitrary Hermitian weight matrix.

We find an approximate solution by relaxing the discreteness and zero-sum constraints. W_{eq} being Hermitian guarantees that when z is relaxed to take any complex values, ϵ_a is always a real number. It can be shown that $\forall \lambda \in \lambda(W_{eq}, D_{eq})$, $|\lambda| \leq 3$, and

$$\epsilon_a(z_{opt}) = 2\,\lambda_1, \qquad z_{opt} \in \Upsilon(W_{eq}, D_{eq}, \lambda_1).$$

As all eigenvalues of an Hermitian matrix are real, the eigenvalues can still be ordered in sizes rather than magnitudes. Because $1 \in \Upsilon(W_{eq}, D_{eq}, \lambda_1)$ when and only when $R = 0$ and $A_d = 0$, the zero-sum constraint $z^2 W_{eq} 1 = 0$ is not, in general, automatically satisfied.

3.4 Phase plane embedding of an ordered partitioning

In order to understand how an ordered partitioning is encoded in the above model, we need to study the labeling vector z. We illustrate the ideas in the language of figure-ground segregation. If we consider R encoding relative depths with $R_d(j,k) > 0$ for j in front of k, the ordered partitioning based on $\epsilon_a(A, R; \beta)$ identifies V_1 as a group in front (figure) and V_2 as a group in the back (ground).

There are two properties of z that are relevant to partitioning: magnitudes and phases. For complex number $c = a + i\,b$, where a and b are both real, its magnitude is defined to be $|c| = \sqrt{a^2 + b^2}$ and its phase is defined to be the angle of point (a, b) in a 2D plane: $\angle c = \arctan \frac{b}{a}$. As $z = \sqrt{1-k}\,X_1 - i\,\sqrt{k}\,X_2$, where k is the degree ratio of the figure, the ideal solution assigns real number $\sqrt{1-k}$ to figure and assigns imaginary number $-i\,\sqrt{k}$ to ground. Therefore, the magnitudes of elements in z indicate sizes of partitions: the larger the magnitude of z_j, the smaller the connection ratio of its own group; whereas the relative phases indicate the figure-ground relationships: $\angle z_j - \angle z_k = 0°$ means that j and k are in the same group, $90°$ (phase advance) for j in front of k, $-90°$ (phase lag) for j behind k. This interpretation remains valid when z is scaled by any complex number c. Therefore, the crucial partitioning information is captured in the phase angles of z rather than the magnitudes as they can become not indicative at all when the connection ratios of two partitions are the same.

When the elements of z are squared, we get $z^2 = (1-k)\,X_1 - k\,X_2$. Two groups become antiphase ($180°$) in z^2 labels. Though the same partitioning remains, the figure-ground information could be lost in cz for constant scaling on z. This fact is most obvious when $A_d + R_d = 0$, where both z and z^* correspond to the same partitioning energy ϵ_a. This pair of solutions suggests two possibilities: V_1 is figure or ground. In other words, the ordering of partitions is created by directed graphs. When we do not care about the direction, z^2 contains the necessary information for partitioning. Indeed, we can show that

$$\epsilon_a = \frac{z^2 W_{eq} z^2}{z^2 D_{eq} z^2} + \frac{1^T W_{eq} 1}{1^T D_{eq} 1}, \qquad \text{if } W_{eq} = 2\beta(A_u - R_u + D_{R_u}).$$

Note that W_{eq} now becomes a real symmetric matrix.

The phase-plane partitioning remains valid in the relaxed solution space. Let $W = D_{eq}^{-\frac{1}{2}} W_{eq} D_{eq}^{-\frac{1}{2}}$, the eigenvectors of which are equivalent (related by $D_{eq}^{\frac{1}{2}}$) to those of (W_{eq}, D_{eq}). Let U and V denote the real and imaginary parts of W: $W = U + i\,V$, where U is symmetric and V is skew-symmetric. We consider U_{jk} (the net effect of attraction A and repulsion R) repulsion if it is negative, otherwise as attraction. For any vector z, we have:

$$\begin{aligned} z^H W z &= \sum_{j,k} |z_j| \cdot |z_k| \cdot \Big(U_{jk} \cos(\angle z_j - \angle z_k) + V_{jk} \sin(\angle z_j - \angle z_k) \Big) \\ &= 2 \sum_{j<k} |z_j| \cdot |z_k| \cdot |W_{jk}| \cdot \cos(\angle z_j - \angle z_k - \angle W_{jk}) + \sum_j |z_j|^2 \cdot U_{jj}. \end{aligned}$$

We see that $z^H W z$ is maximized when $\angle z_j - \angle z_k$ matches $\angle W_{jk}$. Therefore, attraction encourages a phase difference of $0°$, whereas repulsion encourages a phase difference of $180°$, and still directed edge flow encourages a phase difference of $90°$. The optimal solution results from a trade-off between these three processes. If $V_{jk} > 0$ means that j is figural, then the optimal solution tends to have $\angle z_j > \angle z_k$ (phase advance less than $90°$) if U_{jk} is attraction, but phase advance more than $90°$ if it is repulsion. Hence, when there is pairwise repulsion, the relaxed solution in the continuous domain has no longer the ideal bimodal vertex valuation and as a result the zero-sum constraint cannot be satisfied. Nevertheless, phase advance still indicates figure-to-ground relationships.

3.5 Algorithm

The complete algorithm is summarized below. Given attraction measure A and repulsion R, we try to find an ordered partitioning $(\mathsf{V}_1, \mathsf{V}_2)$ to maximize $\epsilon_a(A, R; \beta)$.

Step 1: $A_u = A + A^T$, $A_d = A - A^T$; $R_u = R + R^T$, $R_d = R - R^T$.
Step 2: $D_{eq} = D_{A_u} + D_{R_u}$.
Step 3: $W_{eq} = 2\beta \cdot (A_u - R_u + D_{R_u}) + i \cdot 2(1-\beta) \cdot (A_d + R_d)$.
Step 4: Compute the eigenvectors of (W_{eq}, D_{eq}).
Step 5: Find a discrete solution by partitioning eigenvectors in the phase plane.

4 Results

We first illustrate our ideas and methods using the simple example in Fig. 1. The two directed graphs are decomposed into a symmetric part and a skew-symmetric part (Fig. 2).

This example has clear division of figure as $\{1, 3, 5\}$ and ground as $\{2, 4\}$ because: within-group connections are stronger for nondirectional attraction A_u; between-group connections are stronger for nondirectional repulsion R_u; there are only between-group connections pointing from figure to ground for both directional attraction A_d and directional repulsion R_d.

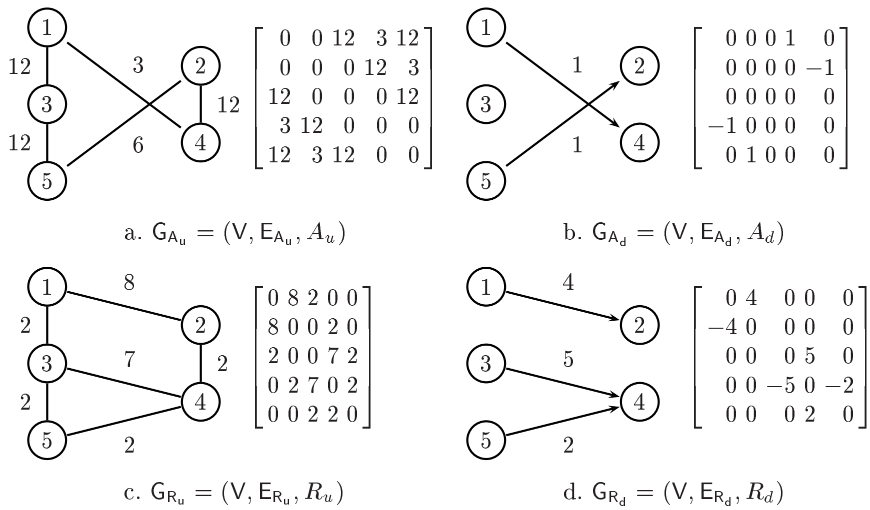


Fig. 2. Decomposition of directed graphs in Fig. 1. a. Nondirectional attraction A_u. b. Directional attraction A_d. c. Nondirectional repulsion R_u. d. Directional repulsion R_d.

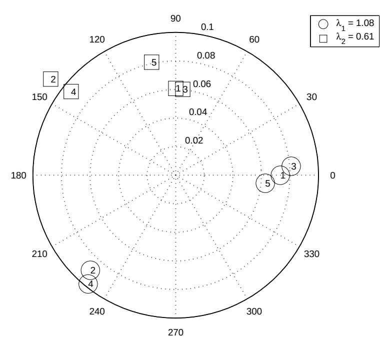
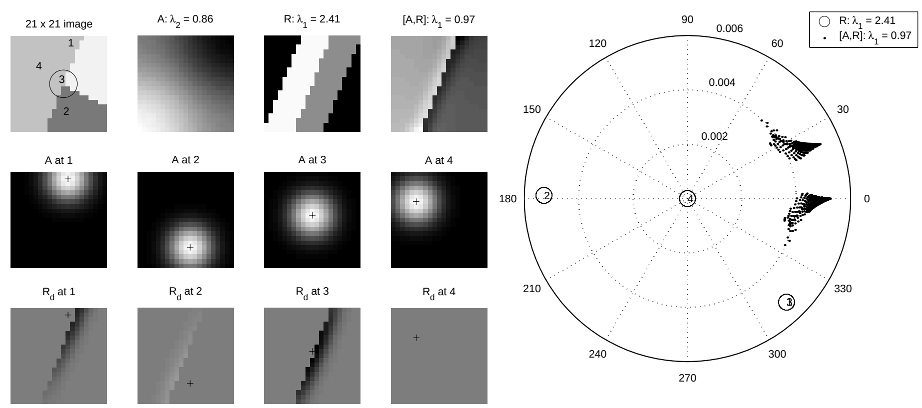


Fig. 3. Partitioning of eigenvectors in the phase-plane. Here we plot the first two eigenvectors of (W_{eq}, D_{eq}) for the example in Fig. 1. The points of the first (second) eigenvector are marked in circles (squares). Both of their phases suggest partitioning the five vertices into $\{1, 3, 5\}$ and $\{2, 4\}$ but with opposite orderings. The first chooses $\{1, 3, 5\}$ as figure for it advances $\{2, 4\}$ by about $120°$, while the second chooses it as ground for it lags $\{2, 4\}$ by about $60°$. The first scheme has a much larger partitioning energy as indicated by the eigenvalues.

The equivalent degree matrix and weight matrix for $\beta = 0.5$ are:

$$D_{eq} = \begin{bmatrix} 37 & 0 & 0 & 0 & 0 \\ 0 & 25 & 0 & 0 & 0 \\ 0 & 0 & 35 & 0 & 0 \\ 0 & 0 & 0 & 26 & 0 \\ 0 & 0 & 0 & 0 & 31 \end{bmatrix}, W_{eq} = \begin{bmatrix} 10 & -8 & 10 & 3 & 12 \\ -8 & 10 & 0 & 10 & 3 \\ 10 & 0 & 11 & -7 & 10 \\ 3 & 10 & -7 & 11 & -2 \\ 12 & 3 & 10 & -2 & 4 \end{bmatrix} + i \cdot \begin{bmatrix} 0 & 4 & 0 & 1 & 0 \\ -4 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 5 & 0 \\ -1 & 0 & -5 & 0 & -2 \\ 0 & 1 & 0 & 2 & 0 \end{bmatrix}.$$

We expect that the first eigenvector of (W_{eq}, D_{eq}) on $\{1, 3, 5\}$ has phase advance with respect to $\{2, 4\}$. This is verified in Fig. 3.



a. b.

Fig. 4. Interaction of attraction and repulsion. a). The first row shows the image and the segmentation results with attraction A (the second eigenvector), repulsion R, both A and R (the first eigenvectors). The 2nd and 3rd rows are the attraction and repulsion fields at the four locations indicated by the markers in the image. The attraction is determined by proximity, so it is the same for all four locations. The repulsion is determined by the T-junction at the center. Most repulsion is zero, while pixels of lighter(darker) values are in front of (behind) the pixel under scrutiny. Attraction result is not indicative at all since no segmentation cues are encoded in attraction. Repulsion only makes boundaries stand out; while working with the non-informative attraction, the segmentation is carried over to the interiors of regions. b). Figure-ground segregation upon directional repulsion. Here are the phase plots of the first eigenvectors for R and A, R. The numbers in the circles correspond to those in the image shown in a). We rotate the eigenvector for A, R so that the right-lower corner of the image gets phase $0°$. Both cases give the correct direction at boundaries. However, only with A and R together, all image regions are segmented appropriately. The attraction also reduces the figure-to-ground phase advance from $135°$ to $30°$.

Fig. 4a shows that how attraction and repulsion complement each other and their interaction gives a better segmentation. We use spatial proximity for attraction. Since the intensity similarity is not considered, we cannot possibly segment this image with attraction alone. Repulsion is determined by relative depths suggested by the T-junction at the center. The repulsion strength falls off exponentially along the direction perpendicular to the T-arms. We can see that repulsion pushes two regions apart at the boundary, while attraction carries this force further to the interior of each region thanks to its transitivity (Fig.4b). Real image segmentation with T-junctions can be found in [17].

Since nondirectional repulsion is a continuation of attraction measures into negative numbers, we calculate the affinity between two $d-$dimensional features using a Mexican hat function of their difference. It is implemented as the difference of two Gaussian functions:

$$\begin{aligned} h(X;\Sigma_1,\Sigma_2) &= g(X;0,\Sigma_1) - g(X;0,\Sigma_2), \\ g(X;\mu,\Sigma) &= \frac{1}{(2\pi)^{\frac{d}{2}}|\Sigma|^{\frac{1}{2}}}\exp^{-\frac{1}{2}(X-\mu)^T\Sigma^{-1}(X-\mu)}, \end{aligned}$$

where Σ's are $d \times d$ covariance matrices. The evaluation signals pairwise attraction if positive, repulsion if negative and neutral if zero. Assuming $\Sigma_2 = \gamma^2\Sigma_1$, we can calculate two critical radii, r_0, where affinity changes from attraction to repulsion and r_-, where affinity is maximum repulsion:

$$r_0(\gamma,d) = \sqrt{\frac{2d\ln(\gamma)}{1-\gamma^{-2}}}, \qquad r_-(\gamma,d) = \sqrt{2+d}\cdot r_0(\gamma,d).$$

The case of $d = 1$ is illustrated in Fig. 5. With this simple change from Gaussian functions [15, 8, 11] measuring attraction to Mexican hat functions measuring both attraction and repulsion, we will show that negative weights play a very effective role in graph partitioning.

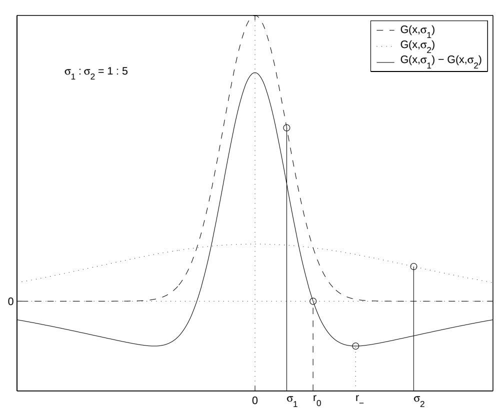
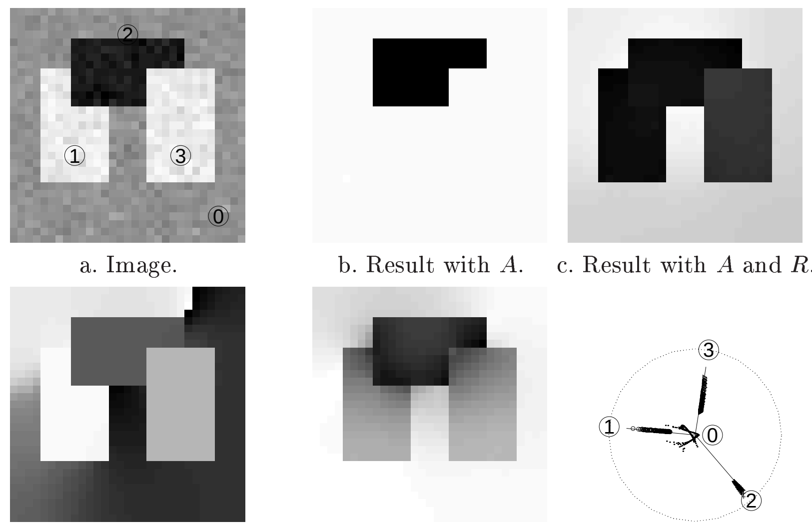


Fig. 5. Calculate pairwise affinity using Mexican hat functions based on difference of Gaussians. When two features are identical, it has maximum attraction; when feature difference is r_0, it is neutral; when feature difference is r_-, it has maximum repulsion.

Fig. 6 shows three objects ordered in depth. We compute pairwise affinity based on proximity and intensity similarity. We see that partitioning with attraction measures finds a dominant group by picking up the object of the highest contrast; with the additional repulsion measures, all objects against a common background are grouped together. If we add in directional repulsion measures based on occlusion cues, the three objects are further segregated in depth.

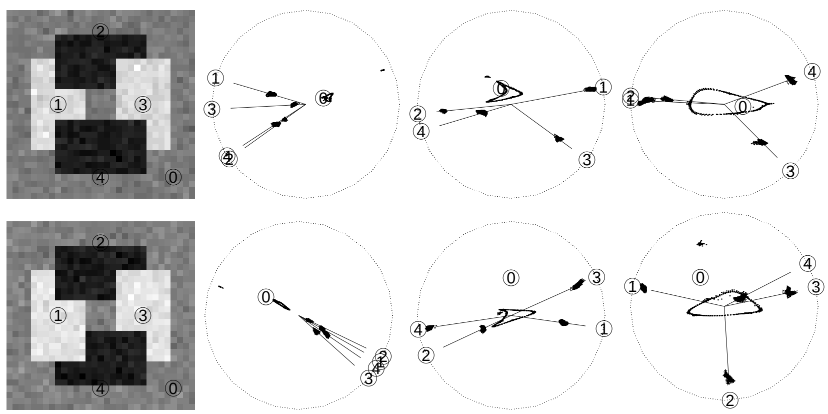


a. Image. b. Result with A. c. Result with A and R.

d. Result with directed R: angle, e. magnitude, f. plot in phase plane.

Fig. 6. The distinct roles of repulsion in grouping. a) 31×31 image. The background and three objects are marked from 0 to 3. They have average intensity values of 0.6, 0.9, 0.2 and 0.9. Gaussian noise with standard deviation of 0.03 is added to the image. Object 2 has slightly higher contrast against background than objects 1 and 3. Attraction and nondirectional repulsion are measured by Mexican hat functions of pixel distance and intensity difference with σ's of 10 and 0.1 respectively. The neighborhood radius is 3 and $\gamma = 3$. b) Segmentation result with attraction alone. c) Segmentation result with both attraction and repulsion. d), e) and f) show the result when directional repulsion based on relative depth cues at T-junctions are incorporated. With nondirectional repulsion, objects that repel a common ground are bound together in one group. With directional repulsion, objects can be further segregated in depth.

Unlike attraction, repulsion is not an equivalence relationship as it is not transitive. If object 3 is in front of object 2, which is in front of object 1, object 3 is not necessarily in front of object 1. In fact, the conclusion we can draw from the phase plot in Fig. 6 is that when relative depth cues between object 3 and 1 are missing, object 1 is in front of object 3 instead. When there are multiple objects in an image, the generalized eigenvectors subsequently give multiple hypotheses about their relative depths, as shown in Fig. 7.



a. b.

Image. $\Upsilon(W_{eq}, D_{eq}, \lambda_1)$ $\Upsilon(W_{eq}, D_{eq}, \lambda_2)$ $\Upsilon(W_{eq}, D_{eq}, \lambda_3)$

Fig. 7. Depth segregation with multiple objects. Each row shows an image and the phase plots of the first three eigenvectors obtained on cues of proximity, intensity similarity and relative depths. All four objects have the same degree of contrast against the background. The average intensity value of background is 0.5, while that of objects is either 0.2 or 0.8. The same parameters for noise and weight matrices as in Fig. 6 are used. The first row shows four objects ordered in depth layers. The second row shows four objects in a looped depth configuration. Repulsion has no transitivity, so object pair 1 and 3, 2 and 4 tend to be grouped together in the phase plane. The magnitudes indicate the reliability of phase angle estimation. The comparison of the two rows also shows the influence of local depth cues on global depth configuration.

These examples illustrate that partitioning with directed relationships can automatically encode border ownerships [10] in the phase plane embedding.

5 Summary

In this paper, we develop a computational method for grouping based on symmetric and asymmetric relationships between pairs of data points. We formulate the problem in a graph partitioning framework using two directed graphs to encode attraction and repulsion measures. In this framework, directed graphs capture the asymmetry of relationships and repulsion complements attraction in measuring dissociation.

We generalize normalized cuts and associations criteria to such a pair of directed graphs. Our formulation leads to Rayleigh quotients of Hermitian matrices, where the imaginary part encodes directed relationships, and the real part encodes undirected relationships with positive numbers for attraction and negative numbers for repulsion. The optimal solutions in the continuous domain can thus be computed by eigendecomposition, with the ordered partitioning embedded in the phases of eigenvectors: the angle separation determines the partitioning, while the relative phase advance indicates the ordering.

We illustrate our method in image segmentation. We show that surface cues and depth cues can be treated equally in one framework and thus segmentation and figure-ground segregation can be obtained in one computational step.

6 Acknowledgements

This research is supported by (DARPA HumanID) ONR N00014-00-1-0915 and NSF IRI-9817496. Yu has also been supported in part by NSF LIS 9720350 and NSF 9984706.

References

[1] A. Amir and M. Lindenbaum. Quantitative analysis of grouping process. In *European Conference on Computer Vision*, pages 371–84, 1996.
[2] A. Blake and A. Zisserman. *Visual Reconstruction*. MIT Press, Cambridge, MA, 1987.
[3] Y. Gdalyahu, D. Weinshall, and M. Werman. A randomized algorithm for pairwise clustering. In *Neural Information Processing Systems*, pages 424–30, 1998.
[4] S. Geman and D. Geman. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6(6):721–41, 1984.
[5] A. K. Jain and R. C. Dubes. *Algorithms for Clustering Data*. Prentice Hall, 1988.
[6] K. Koffka. *Principles of Gestalt Psychology*. A Harbinger Book, Harcourt Brace & World Inc., 1935.
[7] Y. G. Leclerc. Constructing simple stable descriptions for image partitioning. *International Journal of Computer Vision*, 3:73–102, 1989.

[8] J. Malik, S. Belongie, T. Leung, and J. Shi. Contour and texture analysis for image segmentation. *International Journal of Computer Vision*, 2001.

[9] D. Mumford and J. Shah. Boundary detection by minimizing functionals. In *IEEE Conference on Computer Vision and Pattern Recognition*, 1985.

[10] K. Nakayama, Z. J. He, and S. Shimojo. *Vision. In Invitation to Cognitive Science*, chapter Visual surface representation: a critical link between lower-level and higher level vision, pages 1–70. MIT Press, 1995.

[11] P. Perona and W. Freeman. A factorization approach to grouping. In *European Conference on Computer Vision*, pages 655–70, 1998.

[12] J. Puzicha, T. Hofmann, and J. Buhmann. Unsupervised texture segmentation in a deterministic annealing framework. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(8):803–18, 1998.

[13] A. Rosenfeld, R. A. Hummel, and S. W. Zucker. Scene labeling by relaxation operations. *IEEE Transactions on Systems, Man, and Cybernetics*, 6(6):173–84, 1976.

[14] E. Sharon, A. Brandt, and R. Basri. Fast multiscale image segmentation. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 70–7, 2000.

[15] J. Shi and J. Malik. Normalized cuts and image segmentation. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 731–7, June 1997.

[16] Z. Wu and R. Leahy. An optimal graph theoretic approach to data clustering: Theory and its application to image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11:1101–13, 1993.

[17] S. X. Yu and J. Shi. Segmentation with pairwise attraction and repulsion. In *International Conference on Computer Vision*, 2001.

[18] S. C. Zhu, Y. N. Wu, and D. Mumford. Filters, random field and maximum entropy: — towards a unified theory for texture modeling. *International Journal of Computer Vision*, 27(2):1–20, 1998.