

Segmentation with Pairwise Attraction and Repulsion

Stella X. Yu^{†‡} and Jianbo Shi[†]
Robotics Institute[†]
Carnegie Mellon University
Center for the Neural Basis of Cognition[‡]
5000 Forbes Ave, Pittsburgh, PA 15213-3890
{stella.yu, jshi}@cs.cmu.edu

Abstract

We propose a method of image segmentation by integrating pairwise attraction and directional repulsion derived from local grouping and figure-ground cues. These two kinds of pairwise relationships are encoded in the real and imaginary parts of an Hermitian graph weight matrix, through which we can directly generalize the normalized cuts criterion. With bi-graph constructions, this method can be readily extended to handle nondirectional repulsion that captures dissimilarity. We demonstrate the use of repulsion in image segmentation with relative depth cues, which allows segmentation and figure-ground segregation to be computed simultaneously. As a general mechanism to represent the dual measures of attraction and repulsion, this method can also be employed to solve other constraint satisfaction and optimization problems.

1. Introduction

Perceptual organization [11], the structuring of perceptual information into groups, objects and concepts, is an important problem in vision and cognition. It was first studied by Gestaltists, who proposed three parts of grouping: grouping factors [19], figure-ground organization [15], and *Pragnanz*, the last of which, in sharp contrast with atomistic view of visual perception by structuralists, refers to the goodness of overall structure arisen from global interactions from visual stimuli and within visual systems [11].

This view of grouping has led to a computational mechanism of image segmentation as the process of extracting global information from local comparisons between image features (pixels). Gestalt grouping factors, such as proximity, similarity, continuity and symmetry, are encoded and combined in pairwise feature similarity measures [21, 17, 13, 4, 18, 16]. It has been demonstrated on real im-

ages that complex grouping phenomena can emerge from simple computation on these local cues [5, 7].

However, the purpose of grouping is not isolated from figure-ground discrimination. The goodness of figure is evaluated based on both goodness of groups and segregation cues for figure-ground. Grouping and figure-ground are two aspects of one process. They evaluate on the same set of feature dimensions, such as luminance, motion, continuation and symmetry. Closure in grouping is closely related to convexity, occlusion, and surroundness in figure-ground. When a pair of symmetrical lines are grouped together, for example, it essentially implies that the region between contours is the figure and the surrounding area is the ground.

This strong connection between grouping and figure-ground discrimination is not well studied in computer vision. In general, segmentation and depth segregation are often dealt with at separate processing stages [2, 20]. From a computational point of view, this two-step approach is not capable of fully integrating these two types of cues and is prone to errors made in each step. [1] provides a Bayesian approach to binocular stereopsis where local quantities in the scene geometry, which include depth, surface orientation, object boundaries and surface creases, are recovered simultaneously. However, like most formulations in Markov random fields, it suffers from poor computation techniques.

The difficulty of integrating figure-ground cues in the general grouping framework lies in different natures of factors. While grouping factors look at the association by feature similarity, figure-ground emphasizes the segregation by feature dissimilarity, and this dissimilarity can be directional. In fact, figure-ground is closely related to depth segregation [6, 9], since regions in front tend to be smaller, surrounded, occluding and complete, which in turn makes them more likely to exhibit symmetry, parallelism, convexity etc., as our visual world is made of such objects. That a significant number of V1, V2 and V4 cells were found sensitive to distance even in monocular viewing conditions [3]

suggests that depth cues might be intertwined with many early visual processes [10, 14]. The existence of depth-polarity sensitive cells has also been found recently in V1, V2 and V4 [22]. The representation of direction is crucial in discriminating figure from ground.

In this paper, we present a computational grouping method which naturally incorporates both grouping and figure-ground discrimination. We formulate the problem in a directed graph partitioning framework. We represent attraction and directional repulsion in the real and imaginary parts of an Hermitian weight matrix, which we call *generalized affinity*. Segmentation and figure-ground segregation can be encoded together by a complex labeling vector in its phase plane. We generalize the normalized cut criterion to this problem and an analogous eigensystem is used to achieve segmentation and figure-ground in one step.

These results can be extended to nondirectional repulsion with the help of bigraphs. A bigraph is constructed by making two copies of the graph with attraction edges, and representing nondirectional repulsion as directional repulsion in both ways between the two copies of the graph.

The rest of the paper is organized as follows. Section 2 expands our grouping method in detail. Section 3 illustrates our ideas and methods on synthetic data as well as real images. In particular, we will see how repulsion and attraction work together leading to a better segmentation and how we can discover figure-ground information from labeling vectors. Section 4 concludes the paper.

2. Computational algorithms

Image segmentation based on pairwise relationships can be formulated in a graph theoretic framework. In this approach, an image is described by a weighted graph $G = (V, E)$, where a vertex $v \in V$ corresponds to a pixel in image, and an edge $e \in E$ between vertex j and k is associated with a weight which measures both similarity of grouping cues and directional dissimilarity of figure-ground cues. A vertex partitioning (V_+, V_-) on graph $G = (V, E)$, which has $V = V_+ \cup V_-$ and $V_+ \cap V_- = \emptyset$, leads to a figure-ground partitioning of the image.

2.1. Representing directional repulsion

We first extend the definition of conventional graph weight matrices to include directional repulsion. Consider the case illustrated in Fig.1a. We have pairwise attraction between (1, 2), (2, 3) and (4, 5), and repulsion between (1, 4), (2, 5) and (3, 5). The attraction can be represented by a symmetrical weight matrix A_{Ga} (Fig.1c), while the repulsion can be represented by a skew symmetrical weight matrix A_{Gr} (Fig.1d). In the image domain, directional repulsion can arise from relative depth cues such as T-junctions.

We can unify these two types of graph weights nicely in the domain of complex numbers as

$$A_G = A_{Ga} + i \cdot A_{Gr}.$$

We call A_G the *generalized affinity*. We define the degree matrix D_G to be the sum of L_p norm of A_G entries,

$$D_G(j, j) = \sum_k \|A_G(j, k)\|_p, \forall j.$$

Note that the weight matrix for directed edges is different from the conventional nonnegative matrix representation (Fig.1b). With our choice, the generalized affinity A_G becomes Hermitian, which allows us to generalize graph partitioning criteria on real symmetrical matrices.

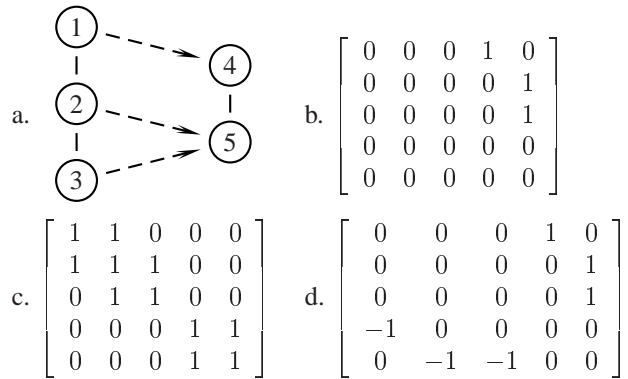


Figure 1: Represent attraction and repulsion in graph weight matrices. a. A graph with mixed mutual attraction (undirected edges) and directional repulsion (directed edges). b. Conventional nonnegative weight matrix for directed edges. c. Represent attraction in a nonnegative and symmetrical matrix A_{Ga} . d. Represent repulsion in a skew symmetrical weight matrix A_{Gr} . This representation is natural for relative depth cues that we can say j is in front of k if $A_{Gr}(j, k) > 0$, j is behind k if $A_{Gr}(j, k) < 0$, j and k are on the same depth layer (or as in affinity, information not available) if $A_{Gr}(i, j) = 0$.

2.2. Normalized generalized association criterion

While figure-ground discrimination has a natural asymmetry favoring the figure in the grouping process [12], the goodness of figure is evaluated based on both goodness of groups and segregation cues for figure-ground. Due to the direction nature of figure-ground repulsion, we need polarized subdivisions to balance figure and ground. Fig. 2 shows an example where subdivisions in the ground layer help to establish the perception of figure.

In general, we have further partitions in both figure $V_+ = V_{+1} \cup V_{+2}$ and ground $V_- = V_{-1} \cup V_{-2}$. Let

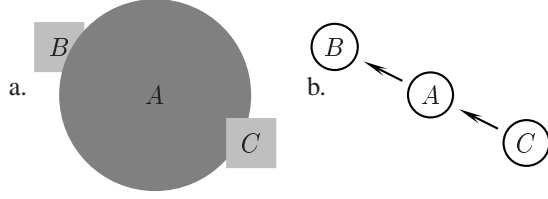


Figure 2: Figure-ground segregation by polarized subdivisions. a. Figure A occludes B but is occluded by C . b. A directed graph representation of example a. In this case, B and C are subdivisions in the ground, they both contribute but with opposite polarities to the perception of figure A .

$\mathcal{A}(P, Q)$, $\mathcal{R}(P, Q)$, $\mathcal{D}(P, Q)$ denote the total attraction, repulsion and degree of connections from P to Q , i.e.

$$\begin{aligned}\mathcal{A}(P, Q) &= \sum_{j \in P, k \in Q} A_{G_a}(j, k), \\ \mathcal{R}(P, Q) &= \sum_{j \in P, k \in Q} A_{G_r}(j, k), \\ \mathcal{D}(P, Q) &= \sum_{j \in P, k \in Q} \|A_G(j, k)\|_p.\end{aligned}$$

Let X, Y denote either figure or ground. We define goodness of grouping as κ_a , goodness of directional segregation as κ_r and degree of connections as κ_d :

$$\begin{aligned}\kappa_a(X, X) &= \sum_{t=1}^2 [\mathcal{A}(X_t, X_t) - \mathcal{A}(X_t, X_{3-t})], \\ \kappa_r(X, Y) &= \sum_{t=1}^2 [\mathcal{R}(X_t, Y_{3-t}) - \mathcal{R}(X_t, Y_t)], \\ \kappa_d(X, X \cup Y) &= \mathcal{D}(X, X \cup Y).\end{aligned}$$

We use the following normalized generalized association measure ϵ_n as our criterion:

$$\begin{aligned}\epsilon_n(V_+, V_-) &= \frac{1}{2} \left[\frac{\kappa_a(V_+, V_+)}{\kappa_d(V_+, V)} + \frac{\kappa_a(V_-, V_-)}{\kappa_d(V_-, V)} \right] \\ &+ \frac{1}{2} \left[\frac{\kappa_r(V_+, V_-) - \kappa_r(V_-, V_+)}{\sqrt{\kappa_d(V_+, V) \cdot \kappa_d(V_-, V)}} \right].\end{aligned}$$

Maximizing ϵ_n is to maximize simultaneously the attraction within both figure and ground, the repulsion from figure to ground, each normalized by involved total connections.

Let x be a vertex membership vector, which assumes values from $\{-1, 1, -i, i\}$, where $\{1, -1\}$ label V_+ and $\{i, -i\}$ label V_- . Let $*$, T and H denote conjugate, transpose and conjugate transpose respectively. With $y = \sqrt{1-k}(x+x^*) - \sqrt{k}(x-x^*)$, where $k = \frac{\mathcal{D}(V_+, V)}{\mathcal{D}(V, V)}$, ϵ_n can be written as a Rayleigh quotient:

$$\epsilon_n(y) = \frac{y^H A_G y}{y^H D_G y}, \quad \forall j, y_j \in \{\pm 2\sqrt{1-k}, \pm 2i\sqrt{k}\}.$$

A good segmentation seeks: $y_{opt} = \arg \max_y \epsilon_n(y)$.

2.3. Computational solution

As y is relaxed to take any complex values, according to the Rayleigh's principle, the above combinatorial optimization problem has an approximate solution by the generalized eigenvectors of (A_G, D_G) . 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 distinct eigenvalues and $\Upsilon(M_1, M_2, \lambda)$ be the eigenspace for a particular eigenvalue λ . We use subscript k to denote the k -th largest eigenvalue, i.e., $\lambda_k(M_1, M_2)$ or λ_k when the context is clear, refers to the k -th largest eigenvalue. Let $V_k(M_1, M_2) \in \Upsilon(M_1, M_2, \lambda_k)$. The optimal solution $y_{opt} = V_1(A_G, D_G)$. In general, $\epsilon_n(y) = \lambda$ for $\forall \lambda \in \lambda(A_G, D_G)$, $y \in \Upsilon(A_G, D_G, \lambda)$. It is well known that all eigenvalues of an Hermitian matrix are real, which ensures us that any vertex valuation by eigenvectors has a real partitioning energy ϵ_n . We discard those y 's with $\lambda < 0$, as an equivalent partitioning but with reversed figure-ground can be obtained by y^* and $\epsilon_n(y^*) > 0$.

2.4. Phase encoding of a segmentation

With slight abuse of notation, let $A_G = D_G^{-\frac{1}{2}} A_G D_G^{-\frac{1}{2}}$. The eigenvectors of this new A_G are equivalent to those of (A_G, D_G) . We show here that segmentation and figure-ground segregation can be encoded simultaneously in the phase plane of labeling vectors. Let $\angle c$ be the angle of complex number c , and we take $\angle c = \theta$ to mean that they are congruent: $\angle c = \theta \pmod{360^\circ}$. It can be shown that $\epsilon(x) = \frac{1}{2} x^H A_G x = \frac{1}{2} \sum_j A_{G_a}(j, j) + \epsilon_{a+}(x) - \epsilon_{a-}(x) + \epsilon_{r+}(x) - \epsilon_{r-}(x)$, where

$$\begin{aligned}\epsilon_{a+}(x) &= \sum_{j < k, \angle x_j - \angle x_k = 0^\circ} A_{G_a}(j, k), \\ \epsilon_{a-}(x) &= \sum_{j < k, \angle x_j - \angle x_k = 180^\circ} A_{G_a}(j, k), \\ \epsilon_{r+}(x) &= \sum_{j < k, \angle x_j - \angle x_k = 90^\circ} A_{G_r}(j, k), \\ \epsilon_{r-}(x) &= \sum_{j < k, \angle x_j - \angle x_k = -90^\circ} A_{G_r}(j, k).\end{aligned}$$

This decomposition is illustrated in Fig.3. There are a few points of interest in this derivation. First, the complementary pairings in attraction and repulsion terms confirm that x being real and imaginary is the criterion to partition vertices into two classes, i.e., $\{1, -1\}$ for one class and $\{i, -i\}$ for the other class. $\epsilon_{a+} - \epsilon_{a-}$ measures within class attraction (κ_a) and $\epsilon_{r+} - \epsilon_{r-}$ measures between class repulsion (κ_r), both of which should be maximized for a good segmentation. Second, as $\epsilon(x) \neq \epsilon(x^*)$ and they only differ in the repulsion terms, the relative phases of x components encode the direction of between class repulsion. For our choice in Fig.1, j is in front of k if x_j has an advance ($\leq 90^\circ$) in phase than x_k . The phase-plane embedding remains valid in the relaxed solution space since the relative phases of y components are invariant to any constant scaling on y .

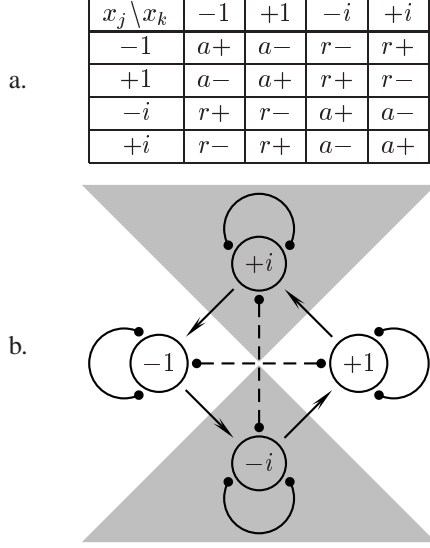


Figure 3: Phase plane embedding of vertex valuation function x . The partitioning energy function $\epsilon(x)$ can be decomposed into four x -dependent terms, ϵ_{a+} , ϵ_{a-} , ϵ_{r+} and ϵ_{r-} . a. $\epsilon(x)$ contribution chart. Attraction and repulsion play complementary roles in summing up pairwise contributions from components of x . b. $\epsilon(x)$ contribution diagram. The undirected and directed edges show attraction and repulsion terms respectively. The shading shows how the phase plane should be divided in order to represent figure and ground clusters. Solid(dotted) undirected lines represent positive(negative) contribution $\epsilon_{a+}(\epsilon_{a-})$. Directed lines represent pairwise repulsion ϵ_{r+} and ϵ_{r-} , and the directions indicate that their contribution to $\epsilon(x)$ is modulated by a particular sign. They form a directed cycle, which suggests relative phases can encode the direction of between class repulsion. For our definition of A_{G_r} in Fig.1d, phase advance means figural. For continuous values of x components, the larger the magnitudes in the phase plane, the more certainty we have about this direction information.

2.5. Representing nondirectional repulsion

For nondirectional repulsion, we transform its undirected graph into a directed graph. One way to achieve this is to duplicate graph G with its nondirectional attraction only and set directional repulsion in both ways between two identical copies of vertices(Fig 4). We call such graphs *bigraphs* and their corresponding vertices *twin nodes*.

Algebraically, for *nonnegative* and *symmetric* nondirectional repulsion A_{G_r} , we construct bigraph \overline{G} , with generalized affinity $A_{\overline{G}}$ and normalization matrix $D_{\overline{G}}$,

$$A_{\overline{G}} = \begin{bmatrix} A_{G_a} & iA_{G_r} \\ -iA_{G_r}^T & A_{G_a} \end{bmatrix}, D_{\overline{G}} = \begin{bmatrix} D_G & \\ & D_G \end{bmatrix}.$$

The eigen-structure of a bigraph is revealed in Theorem 1. Let $m(M_1, M_2)$ be the multiplicities of $\lambda(M_1, M_2)$, and the spectrum be $\Lambda(M_1, M_2) = \begin{bmatrix} \lambda(M_1, M_2) \\ m(M_1, M_2) \end{bmatrix}$.

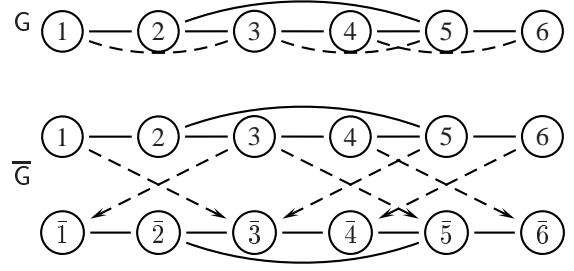


Figure 4: Bigraph to encode nondirectional repulsion. The solid and dashed lines are attraction and repulsion edges respectively. G is a single graph with attraction and nondirectional repulsion. \overline{G} is a bigraph where vertices are cloned together with attraction edges and nondirectional repulsion is represented by directed edges in both ways between two copies of the vertices. The repulsion directions consistently point from one graph to its cloned counterpart.

Theorem 1 Let $A_{G+} = A_{G_a} + A_{G_r}$, $A_{G-} = A_{G_a} - A_{G_r}$.

- $\Lambda(A_{\overline{G}}, D_{\overline{G}}) = [\Lambda(A_{G+}, D_G), \Lambda(A_{G-}, D_G)]$.
- $\forall \lambda \in \lambda(A_{G+}, D_G)$, $y \in \Upsilon(A_{G+}, D_G, \lambda)$ iff $\begin{bmatrix} y \\ -iy \end{bmatrix} \in \Upsilon(A_{\overline{G}}, D_{\overline{G}}, \lambda)$. $\forall \lambda \in \lambda(A_{G-}, D_G)$, $y \in \Upsilon(A_{G-}, D_G, \lambda)$ iff $\begin{bmatrix} y \\ iy \end{bmatrix} \in \Upsilon(A_{\overline{G}}, D_{\overline{G}}, \lambda)$.
- $1 \in \Upsilon(A_{G+}, D_G, 1)$, $\begin{bmatrix} 1 \\ -i \end{bmatrix} \in \Upsilon(A_{\overline{G}}, D_{\overline{G}}, 1)$ iff $\text{trace}(A_{G_a} A_{G_r}^T) = 0$.

Theorem 1 shows that: 1) the spectrum of a bigraph is the combined spectrums of two derived graphs $G+$ and $G-$; 2) the eigenvectors of a bigraph can be obtained from those of $G+$ and $G-$, such that the two copies of G make up figure and ground layers respectively and twin nodes are guaranteed to have either the same or opposite valuation; here we see an example where antiphase components of labeling vectors indicate further segmentation within figure and ground layers; 3) the trivial solution of one graph being figural and the other being ground is a counterpart of $V_1(A_{G_a}, D_{G_a})$ when and only when attraction and repulsion work at different pairings of vertices so that they are completely orthogonal.

To deal with nondirectional repulsion in the framework of directional repulsion, we need to enforce that twin nodes in \overline{G} , which is of the same identity in G , be polarized differently so that within-group repulsion can be counted negatively and between-group repulsion counted positively in ϵ_n . This consideration rules out solutions from A_{G+} . We have proven that the solution for the original nondirectional repulsion problem is equivalent to that for A_{G-} in a rigorous sense.

3. Results

The example in Fig. 5 demonstrates that our representation of *nondirectional* repulsion works in the way that we expect it to be. Indeed, the partitioning gets shifted with where we add repulsion. For example, if we put repulsion between points located at 45° and -135° of the circle, we have a partition boundary along 135° . On the other hand, dots forming a perfect circle have no preference to a cut at a particular angle if only proximity is considered, thus when the points are generated with a different random seed, the partitioning given by attraction alone varies subject to small random changes in the point distribution.

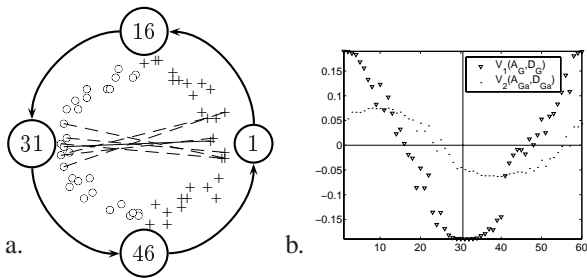


Figure 5: Repulsion can break the transitivity of attraction. a. Random point set. Attraction measures proximity. Now hypothetical *nondirectional* repulsion is added between two solid-line connected points and their 6 neighbours (with linear fall-off in the repulsion strength). These points are numbered counterclockwise as marked in the figure. The points are classified as circles and crosses by thresholding $V_1(A_G, D_G)$ with zero. b. The triangles and dots are $V_1(A_G, D_G)$ and $V_2(A_{Ga}, D_{Ga})$ respectively. When the point set is generated with a different random seed, the former remains the same cut at 90° while the latter changes.

Fig.6 shows that how attraction and repulsion complement each other and their interaction through normalization gives a better segmentation. We use spatial proximity for attraction. Since the intensity similarity is not considered, we cannot possibly segment this image with this 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. Compare the first eigenvectors of iA_{Gr} , (iA_{Gr}, D_{Gr}) , A_G and (A_G, D_G) . 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, so that the repelled boundary does not stand out as a separate group as with repulsion alone case. For this directional repulsion case, we can tell figure vs. ground by examining the relative phases of the labeling vector components (Fig.7).

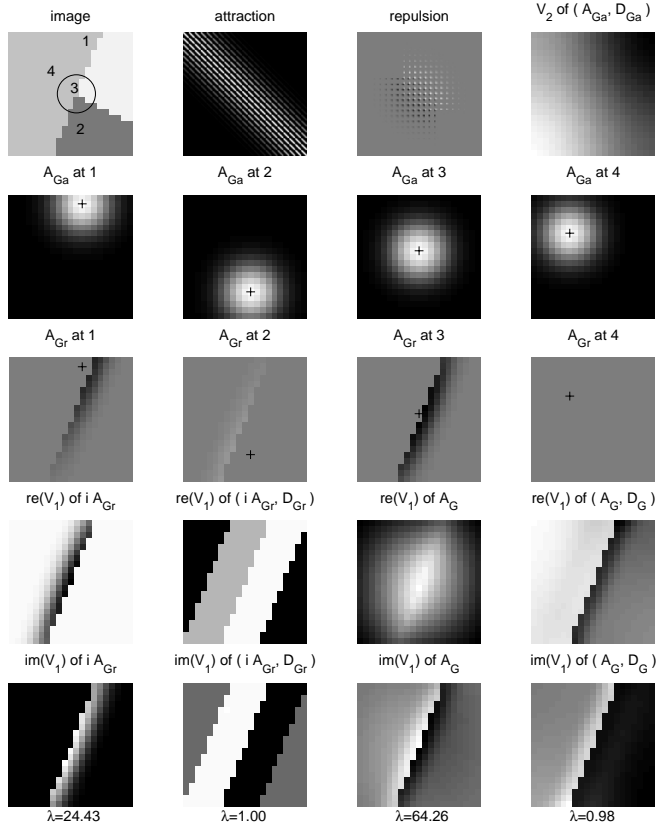


Figure 6: Interaction of attraction and repulsion. The first row shows the image, A_{Ga} , A_{Gr} and the solution $V_2(A_{Ga}, D_{Ga})$ on attraction alone. The 2nd and 3rd rows are the attraction and repulsion fields at the four locations indicated by the markers in the image. The attraction is determined 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. The fourth and fifth rows are the real and imaginary parts of the first eigenvectors of repulsion, normalized repulsion, generalized affinity, normalized generalized affinity respectively. Their eigenvalues are given at the bottom. The normalization equalizes both attraction and repulsion, while the interaction of the two forces leads to a harmonic segmentation at both the boundary and interiors.

For single real images, we consider depth cues arisen from occlusion, e.g., T-junctions. We fit each T-junction by three straight lines and determine the area bound by the smoothest contour (maximum angle between two lines) to be figural. The imprecision of the T-junction localization is accommodated to certain extent by setting up repulsion in a zone out of certain radius of T-junctions(Fig.8). This way of setting up repulsion not only is more robust, but also reflects its spatially complementary role to attraction in grouping.

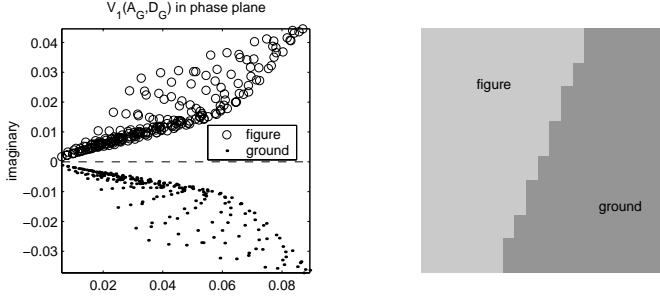


Figure 7: Figure-ground segregation upon directional repulsion. Left figure is the phase plot of $V_1(A_G, D_G)$ shown in Fig.6. It is readily separable into two groups by thresholding its angle with 0° (the dashed line). Since phase advance means figural (Fig.3), the upper group is the figure. On the right we show the results by mapping this classification in the phase plane back to image pixels.

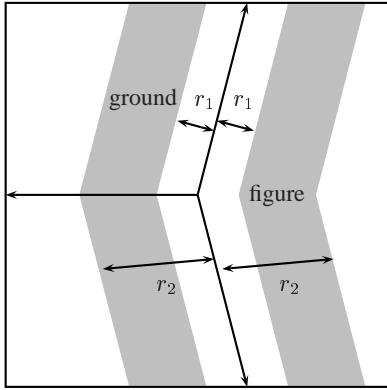


Figure 8: Active repulsion zone. Repulsion is set up within a banded zone aligned with the two arms of the T-junction. The repulsion strength between one pixel in figure and the other in ground falls off exponentially from the maximum of 1 at those pairs adjacent to inner lines to the minimum around 0 at those pairs adjacent to outer lines.

One problem with such depth cues is that they are very sparse. It has been shown that with a few cues at stereoscopic depth boundary, we are able to perceive a surface from random dot displays [8]. This cue sparsity suggests that there is probably some kind of grouping happened before depth cues are actively incorporated. We follow the approach in [7] and find an oversegmentation by using region growing on the first few eigenvectors of (A_{Ga}, D_{Ga}) . We derive A_G on these supernodes through the standard procedure of summation. This lumping procedure not only saves tremendous amount of computation, but also plays a constructive role in propagating sparse depth cues (Fig. 9).

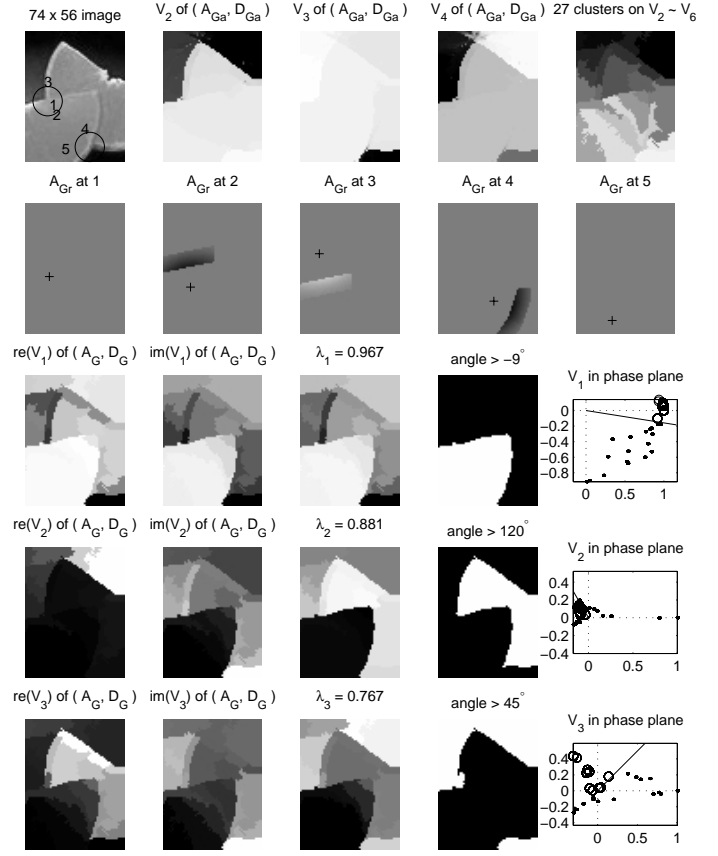


Figure 9: Figure-ground segregation. This image has two clear T-junctions at the places where one orange occludes the other. We add repulsion based on these two cues and its distribution at the five locations marked on the image is given in the second row. The first row also shows the eigenvectors and clustering from normalized cuts on attraction alone. The last three rows shows the first three eigenvectors of (A_G, D_G) . Their third columns are the angles of eigenvectors. By thresholding these angles, in the phase plane (row 3, 4, 5, column 5), we can clearly isolate three parts as figural portions successively (row 3, 4, 5, column 4). The circles in column 5 corresponds to the white region in column 4.

4. Conclusions

In this paper, we propose the necessity of repulsion in characterizing pairwise relationships and generalize the normalized cuts criterion to work on the dual mechanism of attraction and repulsion. Through examples on point sets, synthetic data and real images, we demonstrate that: 1) repulsion complements attraction at different spatial ranges and feature dimensions; 2) repulsion can encode dislikeness to stop the transitivity of similarity cues; 3) sparse repulsion cues can take place of massive and longer attraction; 4) with directional repulsion, we get figure-ground segregation and segmentation at the same time by partitioning generalized

eigenvectors in the phase plane.

As a representation of repulsion, the theoretical benefits are far more than what enables us to segment images with relative depth cues. By likening attraction and repulsion to excitation and inhibition in biological vision systems, we can draw much inspiration on how to use attraction and repulsion for contrast enhancement and competition in grouping. By expanding the graph weight space from nonnegative symmetrical matrices to Hermitian matrices, we can represent a larger class of constraint satisfaction problems, thus many vision problems formulated as such might be solved in this new framework.

5. Acknowledgements

This research is supported by (DARPA HumanID) ONR N00014-00-1-0915 and NSF IRI-9817496. We thank Tai Sing Lee for valuable comments.

References

- [1] P. Belhumeur. A bayesian approach to binocular stereopsis. *International Journal of Computer Vision*, 19(3):237–260, 1996.
- [2] A. Blake and A. Zisserman. *Visual Reconstruction*. MIT Press, Cambridge, MA, 1987.
- [3] A. C. Dobbins, R. M. Jeo, J. Fiser, and J. M. Allman. Distance modulation of neural activity in the visual cortex. *Science*, 281:552–5, 1998.
- [4] Y. Gdalyahu, D. Weinshall, and M. Werman. A randomized algorithm for pairwise clustering. pages 424–30, 1998.
- [5] G. Guy and G. Medioni. Inferring global perceptual contours from local features. *International Journal of Computer Vision*, pages 113–133, 1996.
- [6] G. Kanizsa. *Organization in vision*. Praeger Publishers, 1979.
- [7] J. Malik, S. Belongie, T. Leung, and J. Shi. Contour and texture analysis for image segmentation. *International Journal of Computer Vision*, 2001.
- [8] 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.
- [9] K. Nakayama, S. Shimojo, and G. H. Silverman. Stereoscopic depth: its relation to image segmentation, grouping, and the recognition of occluded objects. *Perception*, 18:55–68, 1989.
- [10] K. Nakayama and G. H. Silverman. Serial and parallel processing of visual feature conjunctions. *Nature*, 320:264–5, 1986.
- [11] S. E. Palmer. *Vision science: from photons to phenomenology*. MIT Press, 1999.
- [12] P. Perona and W. Freeman. A factorization approach to grouping. In *Proceedings of the European Conference on Computer Vision*, pages 655–70, 1998.
- [13] 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.
- [14] I. Rock. *Indirect perception*. MIT Press, 1997.
- [15] E. Rubin. Figure and ground. In D. C. Beardslee and M. Wertheimer, editors, *Readings in perception*, pages 194–203. Van Nostrand, New York, 1958.
- [16] E. Sharon, A. Brandt, and R. Basri. Fast multiscale image segmentation. In *Proceedings of the IEEE Conf. Computer Vision and Pattern Recognition*, pages 70–7, 2000.
- [17] J. Shi and J. Malik. Normalized cuts and image segmentation. In *Proceedings of the IEEE Conf. Computer Vision and Pattern Recognition*, pages 731–7, June 1997.
- [18] Y. Weiss. Segmentation using eigenvectors: a unifying view. In *Proceedings of the International Conference on Computer Vision*, pages 975–82, 1999.
- [19] M. Wertheimer. Laws of organization in perceptual forms (partial translation). In W. B. Ellis, editor, *A sourcebook of Gestalt Psychology*, pages 71–88. Harcourt Brace and company, 1938.
- [20] R. Wildes. Direct recovery of three-dimensional scene geometry from binocular stereo disparity. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 761–74, 1991.
- [21] 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.
- [22] H. Zhou, H. Friedman, and R. von der Heydt. Coding of border ownership in monkey visual cortex. *Journal of Neuroscience*, 20(17):6594–611, 2000.