In 7th IEEE International Conference on Computer Vision, Sep. 1999, Kerkyra, Greece

Measuring Convexity for Figure/Ground Separation *

Hsing-Kuo Pao
Courant Institute
New York University
New York, NY 10012
hsingkuo@cs.nyu.edu

Davi Geiger
Courant Institute
New York University
New York, NY 10012
geiger@cs.nyu.edu

Nava Rubin
Center for Neural Science
New York University
New York, NY 10012
nava@cns.nyu.edu

Abstract

In human perception, convex surfaces have a strong tendency to be perceived as the "figure". Convexity has a stronger influence on figural organization than other global shape properties, such as symmetry ([9]). And yet, there has been very little work on convexity properties in computer vision.

We present a model for figure/ground segregatation which exhibits a preference for convex regions as the figure (i.e., the foreground). The model also shows a preference for smaller regions to be selected as figures, which is also known to hold for human visual perception (e.g., Koffka [11]). The model is based on the machinery of Markov random fields/random walks/diffusion processes, so that the global shape properties are obtained via local and stochastic computations. Experimental results demonstrate that our model performs well on ambiguous figure/ground displays which were not captured before. In particular, in ambiguous displays where neither region is strictly convex, the model shows preference to the "more convex" region, thus offering a continuous measure of convexity in agreement with human perception.

1. Introduction

The selection of salient surfaces in ambiguous figure/ground displays such as those shown in Figure 1 ([9]) reveals important properties of the human visual system. For these figures, the detection of closed boundaries is a straightforward problem, as exact and complete surface boundaries can be obtained by computing the level sets (the location of change from black to white colors).

And yet, these figures present a challenging image segmentation problem, since a decision needs to be made as to which of the regions is seen as the foreground and which as the background. Thus, we consider figure/ground separation to be an integral part of the process of image segmentation. One can formulate the figure-ground problem as a *border ownership* problem ([17], [18], [14]): given the complete description of the boundary contours, how does a system decide on which side of the boundary is the surface that gives rise to that border? The resolution of the border-ownership ambiguity and the determination of the salient regions in the image are thus two interrelated problems. Consider, for example, Figure 1(a): most observers report that the white regions are seen as figures, with the black regions completing behind them in the background. Equivalently we may say that white/black boundaries belong (perceptually) to the white shapes, and the black regions (perceived as background) terminate along the boundaries merely because of occlusion.

1.1. A Continuous Measure of Convexity

The factors that determine which regions are perceived as figure (or foreground) in Figures 1(a-d) must be related to the *shape* of the regions, and not by their contrast polarity, or any other lightness or texture property [1].

What, then, determines the figural organization in Figures 1(a-d)? These figures were developed by Kanizsa ([9]), to illustrate the role of *convexity* in determining border ownership and figure/ground organization: the regions perceived as foreground are more convex than those assigned as background. More specifically, Figures 1(a-d) set off two different global shape properties against each other: convexity vs. symmetry. Note that the regions perceived here as background are perfectly mirror symmetric – a global shape property which was suggested by several Gestalt researchers to strongly bias a surface to be perceived as foreground (e.g., Koffka [11]). The fact that those regions were nevertheless judged as background in Figure 1

*This work was supported by NSF CAREER award and the Sloan Foundation

[1] Contrast polarity and/or other surface-quality factors can play larger role in figure/ground assignment, but not in the cases discussed here.

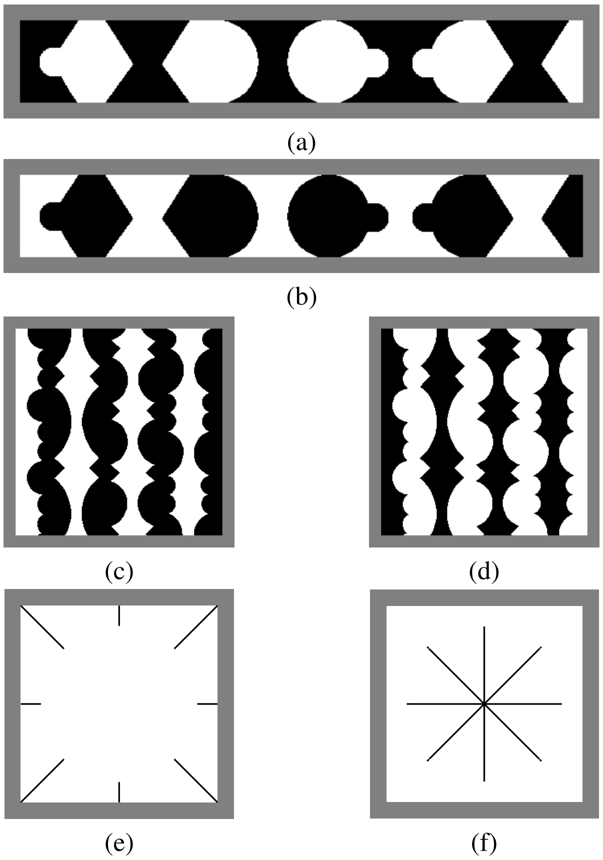


Figure 1. Convexity as a determinant of perceptual saliency (adapted from Kanizsa [9]), observers report perceiving the white regions as the figures in (a) and (d), whereas in (b) and (c) the black regions are the ones perceived as figures. This suggests convexity plays a much larger role in determining figural organization than other factors, such as contrast polarity and symmetry. The effect is also present on illusory figures, where in (e), the convex illusory circle is strongly perceived as foreground, but on (f), the illusory white flat plane with a concave circular hole is hardly perceived as foreground. Here each figure and ground regions are designed to own the same size.

suggests convexity played a larger role than symmetry in determining figural organization. Note, however, that most of the perceptually salient regions in Figure 1 are not convex in the strict mathematical sense, which is an all-or-none definition. Instead, *perceptual* convexity behaves in a continuous manner, where regions can be "more or less convex".

Figure 2. Without convexity, small size objects tend to be perceived as foreground and large size objects as background, regardless of the polarity (adapted from Koffka [11].)

We presents a segmentation model which exhibits a similar preference to select convex regions as foreground. Moreover, the model detects boundaries when they are not readily available (see Figure 1(e)–(f).)

The model is based on local (pixel-to-pixel) computations in a stochastic two-dimensional network. The preference for global shape properties (specifically, convexity) is obtained by propagation of the local interactions, given by the diffusion-like processes. The model, like human perception, selects regions which are intuitively (or perceptually) more convex even in cases when neither region obeys the exact mathematical definition of convexity. The model therefore provides a continuous measure of convexity. Effectively, the diffusion process groups intensity edges to produce closed regions, without any restriction on the number of regions and with topological freedom for admitting any number of holes in each region.

There has been few previous work on using convexity for recognition, so they need to be pointed out ([8], [7] and [20]). Our effort differs from them in that we are extending the notion to a continuous measure of convexity used in studying vision perception. Moreover, our stochastic formulation is unique in this context.

Finally, our model also shows a preference for *smaller* regions to be selected as figures, which is known to hold for human visual perception as well (especially when all other factors are held equal; see Figure 2, adapted from Koffka [11]).

2. Figure-Ground and Entropy Criteria

We start by thinking a shape-figure as a "two dimensional cave" and the shape-figure-boundary as a source of "heat", diffusing inwardly while the background has the source of "heat" diffusing outwardly. In many cases, the partition of the image in regions is not known (Figure 1 (e)-(f)), the source of "heat" are then sparse, at the so called *inducers* such as corners, T-junctions, end lines, and possibly all available intensity edges. After the diffusion process is done, an "entropy" criteria is used to select the best figure-ground configuration. This framework is inspired by Kumaran et al. [12], However, (i) their work concentrated more on the illusory figures. When the figure/ground boundary detection become trivial (Figure 1 (a)-(d)), the closed door diffusion give no difference for each totally blocked region. We, on the other hand, use "leaking energy diffusion process" to overcome this minor (Section 3); (ii) Their sources/inducers must be special intensity change places, such as junctions or corners, while we consider any intensity change as a potential source. Our approach will work even if a curved edge is not detected as a corner. Thus, we put no extra burdens on the feature detection process.

We emphasize here, that we do not address the problem of the combinatorial explosion of multiple hypothesis. Rather, we focus on formulating a criteria to select the best hypothesis.

2.1. Naming the Variables

The input image is defined on a discrete lattice of size N by N,

$$\mathbf{I}^2 \equiv \{k = 1, ..., N^2\}\ . \qquad \langle \text{ set of pixels } \rangle$$

Let us indicate a shape-figure $\mathbf{S}$ of size S pixels in the image either by

$$\mathbf{S} \equiv \{k_s :\ s = 1, ..., S\} \qquad \langle \text{ shape representation } \rangle$$

with area $\mathcal{A}(\mathbf{S}) = \#\mathbf{S} = S$ pixels, or simply by

$$\mathcal{C}(\mathbf{S})\ . \qquad \langle \text{ shape boundary } \rangle$$

Image wise, closed boundary like to "sit" on intensity edges, but boundary points may not have an intensity edge. We then define the edge-perimeter of $\mathbf{S}$ as the size of the boundary set weighted by the normalized intensity edges, i.e.,

$$\mathcal{P}(\mathbf{S}, I) = \sum_{k \in \mathcal{C}(\mathbf{S})} e(k)\ , \qquad \langle \text{ edge-perimeter } \rangle$$

where $e(k) \in [0, 1]$ is the intensity edge (magnitude of intensity gradient) at pixel k, normalized by the largest intensity edge on a given image I. All pixels k of the perimeter belong to $\mathbf{S}$ and since $0 \leq e(k) \leq 1$, we have

$$S = \mathcal{A}(\mathbf{S}) > \#\mathcal{C}(\mathbf{S}) \geq \mathcal{P}(\mathbf{S}, I)\ .$$

Let us assign the ("heat") source value at the *inducer* regions as (see Figure 3b.)

$$\sigma_0(k) = \begin{cases} e(k) & \text{if } k \text{ is figure} \\ -e(k) & \text{if } k \text{ is background}\ . \end{cases}$$

$$\langle \text{ hypothesis at inducers } \rangle$$

At every heat source a local choice of figure/background is thus required. These are the local hypotheses and an entropy measure will select the best hypothesis (best set of local hypotheses). Let us also define

$$P(k)\ :\ \mathbf{I}^2 \to \mathbf{R}\,, \qquad \langle \text{ diffusion field } \rangle$$

a function to be evaluated, at every pixel k, by a diffusion network of the local hypotheses.

2.2. Variational Model

We create a diffusion process model for the field $P(k)$ by formulating it as a variational problem.

Local Hypotheses and Data Fitting: We prefer the field $P(k)$ fitting the local hypotheses, where they are available. In our model they are available at all intensity edge pixels according to its strength ($e(k)$), i.e., we want $P(k)$ to minimize

$$E(P|\sigma_0) = \sum_{k=1}^{N^2} e(k)\,(P(k) - \sigma_0(k))^2\,.$$

Smoothness: In order to obtain a diffusion process, from a minimization standing point, we insert a smoothness constraint on $P(k)$. A simple one minimizes the square of the length of the gradient vector $(\frac{\partial P(x,y)}{\partial x}, \frac{\partial P(x,y)}{\partial y})$, or in the discrete setting we write

$$Smooth(\{P\}) = \mu \sum_{k=1}^{N^2} \sum_{k' \in N_k} (1 - e(k,k'))\,(P(k) - P(k'))^2\,,$$

where μ is the smoothness coefficient, $N_k = \{k+1, k-1, k-N, k+N\}$ are the four neighbors of pixel k, and $e(k,k')$ is the magnitude of the intensity change from pixel k to pixel k', normalized to the largest $e(k,k')$. Note that we can define $e(k) = \max_{k' \in N_k} e(k,k')$.

Energy Model and Level Set: With the smoothing criteria, the total cost function becomes

$$E(P) = \sum_{k=1}^{N^2} \Big[e(k)\,(P(k) - \sigma_0(k))^2 + \sum_{k' \in N_k} \mu(1 - e(k,k'))(P(k) - P(k'))^2 \Big]\,. \quad (1)$$

Thus, the optimal solution $P^*(k)$ balances fitting the local hypothesis and smoothing. It is clear (e.g., [12]) that $P^*(k)$ is bounded by the maximum and minimum values of σ_0, i.e., ± 1. This gives a diffusion property to this process.

2.3. Shape-Figure and Entropy

In order to select the set of hypotheses that produce the "best figure" (best shape), we consider an entropy measure. After all, given a source, we do not know which side takes $\sigma_0(k) = e(k)$ or $-e(k)$. At junctions the multiplicity of hypotheses grows (see [12], [3]). We first, for simplicity, convert $-1 \leq P^*(k) \leq 1$ into a probability distribution at each pixel, via the linear mapping

$$p^*(k) = \frac{1}{2}(1 + P^*(k)) .$$

Thus, the entropy criteria becomes

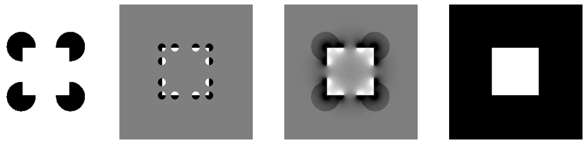
$$\mathcal{S}(p^*) = -\frac{1}{S} \sum_{k \in \mathbf{S}} p^*(k) log p^*(k) + (1 - p^*(k)) log(1 - p^*(k)) ,$$

where $\mathbf{S}$ is defined as the set of pixels k such that $p^*(k) \geq 0.5$ (that means $P^*(k) \geq 0$). For the background entropy (that needs to be compared against), one can also compute

$$\overline{\mathcal{S}}(p^*) = -\frac{1}{\overline{S}} \sum_{k \notin \mathbf{S}} p^*(k) log p^*(k) + (1 - p^*(k)) log(1 - p^*(k)) ,$$

where $\overline{S} = N^2 - S$. The sharper is the diffusion, the closer to 1 is $P^*(k)$ inside $\mathbf{S}$, the better is the figure perception, i.e., the lower the entropy the more salient is the region. Note that the entropy is a per pixel entropy or the total entropy normalized by the number of pixels.

Problem: This criteria favors closed contour that "track" the intensity edges, e.g., the black-and-white four pack men on Figure 3. In this case, with all $e(k, j) = 0, 1$ (edges or non-edges), the four pack man solution gives $P^*(k) = \pm 1$ everywhere, i.e., $\mathcal{S}(p^*) = 0$. This is the best solution one can hope. Moreover, with this criteria we can not distinguish between the black and white regions of Figures 1(a)-(d) as both regions will give zero entropy. Thus, we need to devise a criteria that would favor geometric properties as well as intensity edges, that would favor the Kanizsa square solution and convex regions. We now address the trust of our work: How can we fully characterize a desired shape via an entropy criteria and how can we measure convexity?



(a) (b) (c) (d)

Figure 3. (a) the Kanizsa square. (b) local hypothesis where $\sigma_0 = 1$ is white (figure) and $\sigma_0 = -1$ is black (background). (c) $P^*(k)$-diffusion of all the local hypotheses. (d) threshold (level set) at $P = 0$. The entropy of the diffusion is $\mathcal{S} = 0.75$.

3. Decay Process and Convexity

In order to balance intensity edges (number of inducers) and shape we introduce a "decay" process outside the inducers (sources). We require non-source pixels k to be a non-commitment between figure and ground, i.e., we require $P(k)$ to be zero ("neutral"). This idea can be implemented in the variational approach by adding an energy term $E_{decay} = \sum_{k=1}^{N^2} \nu(1 - e(k)) P^2(k)$. The energy(1) becomes

$$E(P) = \sum_{k=1}^{N^2} \Big[\lambda_k (P(k) - \sigma_0(k))^2 + \sum_{k' \in N_k} \mu_{k,k'} (P(k) - P(k'))^2 \Big] , \qquad (2)$$

where $\mu_{k,k'} = \mu(1 - e(k, k'))$ and now

$$\lambda_k = \max(e(k), \nu(1 - e(k))) \quad \text{and} \quad \sigma_0(k) = \pm\, e(k) ,$$

where $e(k) \in [0, 1]$ and ν is the decay coefficient. In Appendix A, we show that adding this energy term is equivalent to adding a decay to the random process associated to the previous energy.

The solution to equation (2) can be written (e.g., [12]) as

$$\mathbf{P}^* = \mathbf{D}^{-1}(\mu, \nu)\, \lambda \sigma_0 . \qquad (3)$$

where the symmetric and band limited matrix $\mathbf{D}(\mu, \nu)$ have the following structure

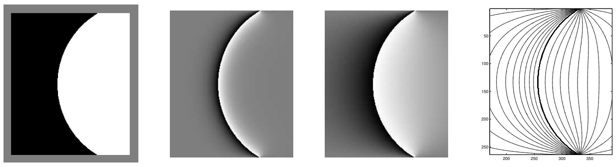
(i) The diagonal element $D_{k,k} = \lambda_k + \mu\, b(k, N_k)$,

(ii) The non-zero off-diagonal elements are $D_{k,k'} = -\mu\,(1 - e(k, k'))$ for $k' \in N_k$,

where $b(k, N_k) = \sum_{k' \in N_k} (1 - e(k, k'))$. Results of this minimization are shown in Figure 3.

In order to obtain the figure $\mathbf{S}$, we consider all pixels $k \in \mathbf{I}$ such that $P^*(k) \geq 0$. The background pixels are obtained as pixels $k \in \mathbf{I}$ such that $P^*(k) < 0$. The level sets $k \in \mathbf{I}$ such that $P^*(k) = 0$, represent the closed contours $\mathcal{C}(\mathbf{S})$. The result of Figure 3 d. (and [12], [3]) suggests that the shapes obtained are "roughly" in agreement to perception. Finally in the Appendix A, we derive the solution $P^*(k)$ and show that it indicates an "average" over the number of random walks starting at the sources and passing through pixel k.

We can now examine and discriminate closed intensity regions, since the associated entropies are no longer zero. We note that "clearly" convex regions have lower entropy than concave ones (see figure 4).



(a) (b) (c) (d)

Figure 4. (a) is an image with two regions, one is convex and concave is another. Both of them own the same area and inducers. (b), (c) are the maps according to the typical hypothesis that assumes the white region is salient. We set $\nu = 5 \times 10^{-4}, 1 \times 10^{-5}$ with (b) & (c) respectively. The maps show decay when it is away from the arc. The entropy (convex / concave or $\mathcal{S}$ / $\overline{\mathcal{S}}$) for them are (b) 0.972 / 0.974 and (c) 0.880 / 0.916. (d) the isocontour (iso-$P^*(k)$) for the map (c). It shows how the diffusion expand larger distance on the convex side.

3.1. Convexity Measurement

Let us show how the entropy measure is capturing convexity (and capturing, simultaneously, size). Let us start with the definition of convexity for a given shape $\mathbf{S}$.

Convexity: *A shape* $\mathbf{S}$ *is convex if and only if for any pair of points inside* $\mathbf{S}$ *the line segment connecting this pair of points is completely inside* $\mathbf{S}$.

Let us represent $\overline{kj}$ to be the Euclidean distance between pixels k and j. Let us denote the shortest legal path between k and j **within** the shape $\mathbf{S}$ by $d(k,j)$. The distinction between convex and concave regions is that for convex regions, $d(k,j) = \overline{kj} \quad \forall k,j \in \mathbf{S}$ (by definition) while for concave ones there are $k,j \in \mathbf{S}$ such that $d(k,j) > \overline{kj}$.

Consider Figure 5 where two regions, one convex and the other concave, have the same area and same inducers. Given a source pixel p (along the boundary) and given a pixel q at the convex side of the figure, there is a pixel q' equally distant to p (Euclidean distance) in the concave side. There are many pixels q' with such property. In the concave region the shortest legal walk between p and q' traverse a distance $d(p,q') > \overline{pq'}$ while in the convex region, the shortest legal walk between p and q traverse a distance $d(p,q) = \overline{pq} = \overline{pq'}$.

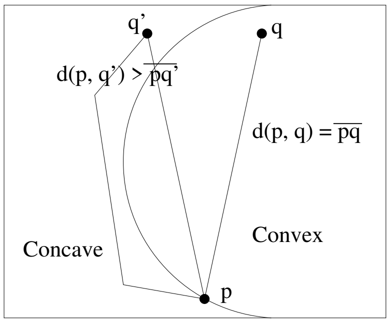


Figure 5. We examine how convexity is captured by the model. Consider all the path starting at the boundary pixel p. In the concave region the shortest random walk from p to q is given by $d(p,q') > \overline{pq'}$ while the shortest random walk to q' is $d(p,q) = \overline{pq} = \overline{pq'}$. The shorter are the random walks the more support is obtained by the source, otherwise the random walk tend to decay. Thus, concave regions get weaker response. The reflectivity makes random walks even longer at concave regions.

We now invoke the theory of random walks (see Appendix A) to suggest that the entropy is lower for convex regions than for concave ones.

The contribution to the the final solution $P^*(q)$ and $P^*(q')$ of the source p to q and q' is directly dependent on $d(p,q)$ and $d(p,q')$ respectively. This is because with the decay term, only the pixels with shorter distance can be reached by the random walk. The longer is the path to reach a pixel the more likely it will die (yielding the decay). More precisely, from the Markov theory the probability of a particle dying when going in the shortest path from p to q is exactly $\sum_{s=1}^{d(p,q)} (Q^s_{pa(s)} R_a)$, where the sum is over all pixels a in the shortest path from p to q, parameterized by $s = 1, \ldots, d(p,q)$. Thus, the longer is the path, the more terms in the sum over a, and the smaller is the probability to reach q. Thus, a point q in a convex side will be reached by more random walks than a point q' in the concave side. This implies that $P^*(q)$ will be more defined to 1 or -1 (to figure or ground) than $P^*(q')$ and the entropy will satisfy $S^*(q) < S^*(q')$. Thus, convexity is encouraged by the entropy criteria.

3.2. Size

The entropy is lower for smaller size objects. Indeed, neglecting the smoothing in (3), and for shapes with inducers on all edges ($e(k) = 1$), we have $P^*(k) = \sigma_0(k)$ and $S(P^*) = log2\frac{(\mathcal{A}(\mathbf{S})-\mathcal{P}(\mathbf{S},\mathbf{I}))}{\mathcal{A}(\mathbf{S})} = log2(1-\frac{Perimeter}{Area})$. Note, that the perimeter is adapted to the number of inducers. The more inducers the smaller is our "size" criteria.

3.3. Reflectivity

In order to enhance the convexity effect, we studied a "reflectivity" term that bounces the random walk perpendicular to the tangent to the boundary shape once it reaches the boundary. Thus, the diffusion tends to go along perpendicular directions to the tangent of the boundary contour.

Let us consider a pixel k at the boundary $\mathcal{C}(\mathbf{S})$. The reflective can be considered by increasing the probability of a jump from k to the neighbor pixel, $f(k)$, in the direction perpendicular to the tangent to the curve. In a corner there may be two such directions, $f1(k)$ and $f2(k)$, so, we consider the set $F_k = \{f1(k), f2(k)\}$ to contain both directions.

Within the theory of random walks, these modifications are simple to implement. More precisely, we modify Q to

$$Q_{k,k'} = \begin{cases} \frac{2\mu}{\lambda_k+\mu(b(k,F_k)+b(k,N_k))} & \forall k' \in F_k \\ \frac{\mu}{\lambda_k+\mu(b(k,F_k)+b(k,N_k))} & \forall k' \in N_k - F_k\,. \\ 0 & \text{otherwise} \end{cases}$$

We are also forced to modify $R_k = \frac{\lambda_k}{\lambda_k+\mu(b(k,F_k)+b(k,N_k))}$ so that M is always stochastic ($\sum_{j=0}^{N^2} M_{k,j} = 1$). This modification of Q only affects pixels near the boundary, where the set $F(k)$ is not empty (where $b(k,F_k) \neq 0$). We note that there is an associated energy to this process

$$E_{reflectivity}(P) = \mu \sum_{k=1}^{N^2} \sum_{k' \in F_k} \mu_{k,k'} (P(k) - P(k'))^2\,.$$

In order to keep the symmetry of the process we request that if $k' \in F_k$ then $k \in F_{k'}$.

In Figure 5 the reflectivity enhances $d(p,q')$ making the difference $d(p,q') - d(p,q)$ larger.

4. Experiments

The first series (see Figure 6) is set for examining the convexity effect. We test our model on a series of images where black and white regions own the same size. The degree of the angle θ in those images are $n\pi/36$, where $n = 6, 7, ..., 36$. The difference in entropy decreases as the angle becomes wider (Figure 6). We then test other images (see Figure 7,8).

We set the smoothness coefficient $\mu = 10^{-1}$ and the decay coefficient $\nu = 5 \times 10^{-4}$ through all our experiments (Except in Figure 4). The numerical method we apply here for the optimization process is conjugate gradient descent method.

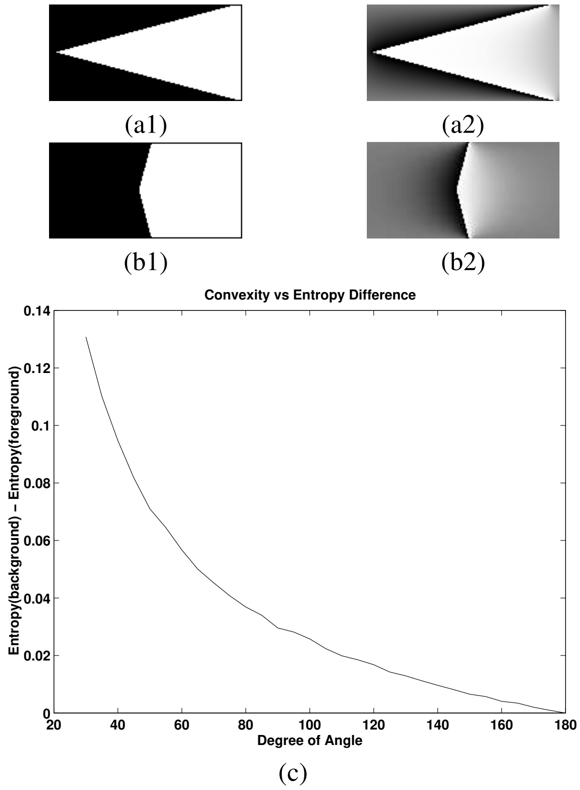


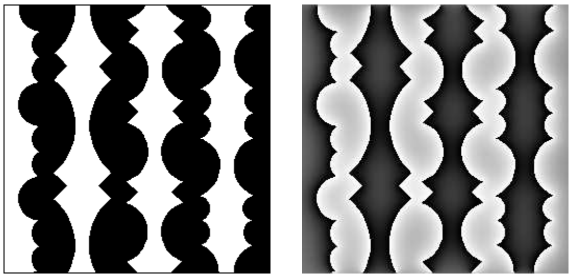
Figure 6. (a1), (b1), The parallel pentagon image with angle of $\pi/6$ and $5\pi/6$ degree, (a2), (b2) the saliency map of them. The entropy values for (a2) and (b2) are (convex/concave or $\mathcal{S}$ / $\overline{\mathcal{S}}$) 0.721 / 0.870 and 0.944 / 0.950 respectively, with $\nu = 1 \times 10^{-5}$. (c) The difference between the entropy for the convex and concave region as a function of the angle ("inverse of convexity").

A. Random Walk Formulation

Let us now bring the theory of Markov chains to understand our approach as "an average over random walks" or as a diffusion. This view will allow us to modify the model and construct a measure of convexity and size.

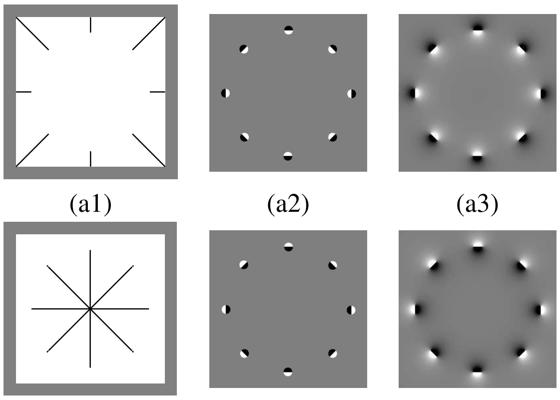
The matrix $\mathbf{D}$ from equation (3) can be written as

$$\mathbf{D} = \mathcal{D}(\mathbf{I} - \mathbf{Q})\,,$$



(a) (b)

Figure 7. Inducers are placed along all the perimeter. The entropy values are $\mathcal{S} = 0.521$ / $\overline{\mathcal{S}} = 0.534$ for the black/white regions.



(a1) (a2) (a3)

(b1) (b2) (b3)

Figure 8. (a1), (b1), initial figures, (a2), (b2), sparse data and (a3), (b3), the diffusion results. The figure entropy values ($P^*(k) > 0$) are (a3)$\mathcal{S} = 0.968$, (b3)$\mathcal{S} = 0.975$.

with $\mathcal{D}$ being the diagonal matrix with elements $\mathcal{D}_{k,k} = \lambda_k + \mu b(k,N_k)$ ($\mathcal{D}_{k,k'} = 0$ for $k \neq k'$.) Therefore $\mathbf{Q}$ must have four off-diagonals non-zero and positive i.e.,

$$Q_{k,k'} = \begin{cases} \frac{\mu}{\lambda_k + \mu b(k,N_k)} & \forall k' \in N_k \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

The inverse of $\mathbf{D}$ is

$$\mathbf{D}^{-1} = (\sum_{n=1}^{\infty} \mathbf{Q}^n)\mathcal{D}^{-1},$$

and so,

$$P^* = (\sum_{n=1}^{\infty} \mathbf{Q}^n)g, \tag{5}$$

with $g = \mathcal{D}^{-1}\lambda\sigma_0$, i.e., $g_k = \frac{\lambda_k}{\mathcal{D}_{kk}}\sigma_{\mathbf{0}}(k)$. We can think g as our "new" input data. The matrix $\mathbf{Q}$ is the "transient" part of an "absorbing Markov chain" $\mathbf{M}$, where $\mathbf{M}$ is

$$\mathbf{M} = \begin{pmatrix} 1 & 000... \\ R & Q \end{pmatrix}, \tag{6}$$

with R being a $N \times 1$ matrix with elements

$$R_k = 1 - \sum_j Q_{kj} = \frac{\lambda_k}{\lambda_k + \mu b(k,N_k)}, \tag{7}$$

so that $\sum_{j=0}^{N} M_{ij} = 1 \quad \forall i$. The stochastic matrix $\mathbf{M}$ is of size $N+1$ by $N+1$. It is clear that

$$\mathbf{M}^n = \begin{pmatrix} 1 & 000... \\ R' & Q^n \end{pmatrix},$$

where $R' = \sum_{\alpha=0}^{n} \mathbf{Q}^{\alpha} R$. $\mathbf{M}$ is an absorbing Markov chain in the sense that if one starts with any state s_0, by successively applying $\mathbf{M}$, one ends up in the state $s_{final} = (1,0,0,...,0)$, i.e., $s_{final} = \lim_{n\to\infty} s_0 \mathbf{M}^n$ (we are applying the matrix to the vector to the "left") This is clear from the fact that $\lim_{n\to\infty} Q^n = 0$, since all the elements of Q are less than 1 and sum less than 1 for each row. This is why Q is the transient part of $\mathbf{M}$. The interpretation of the state $s = (\pm 1,0,...,0)$, is of "death" of the random walk, since the random walk jumped to the zero-th coordinate that is not a pixel in the lattice. Moreover, once it reaches the zero-th coordinate it never leaves, i.e., $s = (\pm 1,0,0,....,0)$ is an eigenstate of $\mathbf{M}$ or $s\mathbf{M} = s$ (or equivalently, to the "right" $s^\dagger$ is an eigenstate of $\mathbf{M}^\dagger$, with $\mathbf{M}^\dagger s^\dagger = s^\dagger$. This is the absorbing state.

Random Walk View: Let us clarify the random walk interpretation of the matrices.

> $(Q^n)_{k,j}$- probability that a random walk that started at pixel k will reach pixel j in n steps.
>
> R_k- probability that a random walk starting at pixel k will die (move to the zero-th state coordinate.)
>
> $R' = \sum_{\alpha=0}^{n} \mathbf{Q}^{\alpha} R$ - probability that the random walk of n+1 steps has died at any stage before reaching pixel k.
>
> $(\sum_{k=1}^{n} \mathbf{Q}^n)_{k,j}$-probability that the random walk started at pixel k and reached pixel j in any number of steps.

In one step, for example, the random walk starting at k has probability non-zero, $Q_{k,j} = \frac{\mu}{\lambda_k + \mu b(k,N_k)}$, to move to one of its $j \in N_k$ neighbors and probability $R_k = \frac{\lambda_k}{\lambda_k + \mu b(k,N_k)}$ to die. Note that inside the shape $\mathbf{S}$ we have $\lambda_k = 0$ and therefore the random walk can not die there. It will only die, possibly, when reaching back the boundary.

The initial state $g = \mathcal{D}^{-1}\lambda\sigma_0$ can be written as

$$g(j) = \sum_{k \in \mathcal{C}(\mathbf{S})} \frac{1}{1+\mu b(k,N_k)} e_k(j)\,,$$

where $e_k(j) = \delta_{jk}$, with $\delta_{jk} = 1,0$ and is 1 only if $j = k$. Thus, we conclude this section with the random walk view interpretation of the solution P^* (equation 3) :

The solution P^* correspond to the average of multiple source averaged over the random walks, where each random walk starts at a boundary pixel $k \in \mathcal{C}(\mathbf{S})$ and carry the weight $g(k) = \frac{1}{1+\mu b(k,N_k)}$ (i.e., corners tend to weight more).

So far, we had $R_k = \frac{\lambda_k}{\lambda_k+\mu b(k,N_k)} = 0$ inside the shape, since $\lambda_k = 0$. We now modify it to $R_k = \frac{\nu}{\nu+\mu b(k,N_k)}$ where $1 > \nu > 0$, so that the probability of jumping to the zero-th coordinate (and dying) is small, depending how small ν is, but not zero.

References

[1] M. Brady and W. E. L. Grimson. The perception of subjective surfaces. A.I. Memo No. 666, AI Lab., MIT, Nov. 1982.

[2] J. Elder and S. W. Zucker. A measure of Closure. *Vision Research*, Vol. 34 (24), pp.3361-3369, 1994.

[3] D. Geiger, H. Pao and N. Rubin. Salient and multiple illusory surfaces. *Computer Vision and Pattern recognition.*, June. 1998.

[4] G.Guy and G. Medioni. Inferring global perceptual contours from local features. In *Proc. IU Workshop DARPA*, Sept. 1992.

[5] S. Grossberg and E. Mingolla. Neural dynamics of perceptual grouping: textures, boundaries and emergent segmentations. *Perception & Psychophysics*, 38(2):141–170, 1985.

[6] F. Heitger and R. von der Heydt. A computational model of neural contour processing: Figure-ground segregation and illusory contours. *Proceedings of the IEEE*, 1993

[7] D. Huttenlocher and P. Wayner, P., "Finding Convex Edge Groupings in an Image," *International Journal of Computer Vision*, **8**(1):7-29, 1992.

[8] D. Jacobs. Robust and efficient detection of convex groups" *IEEE Trans. PAMI* 1995.

[9] G. Kanizsa. *Organization in Vision*. Praeger, New York, 1979.

[10] B. Kimia, A. Tannenbaum, S. Zucker. "Shapes, Shocks, and Deformations I: The components of two-dimensional shape and the reaction-diffusion space", *Int. J. Comp. Vis.***1**: 189-224, 1995.

[11] K. Koffka. *Principles of Gestalt Psychology*. New York: Harcourst. 1935.

[12] K. Kumaran, D. Geiger, and L. Gurvits. Illusory surfaces and visual organization. *Network: Comput. in Neural Syst.*, 7(1), Feb. 1996.

[13] D. Mumford. Elastica and computer vision. In C. L. Bajaj, editor, *Algebraic Geometry and Its Applications*. Springer-Verlag, New York, 1993.

[14] K. Nakayama, Z. J. He, S. Shimojo, "Visual surface representation: a critical link between lower-level and higher-level vision", in *Visual Cognition Eds S M Kosslyn and D N Osherson* (Cambridge, MA: Cambridge, MA), pp. 1-70, 1995.

[15] M. Nitzberg and D. Mumford. The 2.1-d sketch. In *ICCV*, pages 138–144. 1990.

[16] S. Parent and S. W. Zucker, "Trace inference, curvature consistency and curve detection", *IEEE PAMI*, Vol. 11, No. 8, pp. 823-839, 1989.

[17] E. Rubin, *Visuell wahrgenommene Figuren*, (Copenhagen: Gyldendals), 1921.

[18] N. Rubin, K. Nakayama, R. Shapley, "Enhanced perception of illusory contours in the lower versus upper visual hemifields", *Science*, 271, pp.651-653, 1996

[19] S. Ullman. Filling in the gaps: The shape of subjective contours and a model for their generation. *Biological Cybernetics*, 25:1–6, 1976.

[20] Y. Weiss. Interpreting images by propagating Bayesian beliefs. In M. Mozer, M. Jordan and T. Petsche, editors, *Advances in Neural Information Processing System 9*, 1996.