Angular Embedding: from Jarring Intensity Differences to Perceived Luminance

Stella X. Yu

Computer Science Department
Boston College, Chestnut Hill, MA 02467, USA

syu@cs.bc.edu

Abstract

Our goal is to turn an intensity image into its perceived luminance without parsing it into depths, surfaces, or scene illuminations. We start with jarring intensity differences at two scales mixed according to edges, identified by a pixel-centric edge detector. We propose angular embedding as a more robust, efficient, and versatile alternative to LS, LLE, and NCUTS for obtaining a global brightness ordering from local differences. Our model explains a variety of brightness illusions with a single algorithm. Brightness of a pixel can be understood locally as its intensity deviating in the gradient direction and globally as finding its rank relative to others, particularly the lightest and darkest ones.

1. Introduction

The objective intensity of light recorded in an image does not always accord with our subjective experience of it. Measured at equal intensity, a shaded white patch appears brighter than an illuminated gray patch (2 & 3 in Fig. 1).

intensity: $I_1 > I_2 = I_3 > I_4 = I_5 > I_6$
brightness: $B_1 > B_2 > B_3 > B_4 > B_6 > B_5$
lightness: $L_1 = L_2 > L_3 = L_4 = L_6 > L_5$

Figure 1. Distinction of intensity, brightness, and lightness. This image features a 3D box whose top and front surfaces have identical text and background colors. 6 pixels are circled, labeled, and under marked with their intensity values between 0 and 1. Their perceived luminance and reflectance, i.e., brightness and lightness, could be ordered very differently from their measured intensity.

This phenomenon reflects of human vision not a failure, but a remarkable ability to sense stable physical properties under ever changing illumination. Understanding how this subjective experience can be achieved computationally is thus of fundamental importance to effective object recognition, spatial reasoning, and material rendering.

Our goal is to turn the objective intensity of light into its subjective brightness, without parsing the intensity image into depths, surfaces, or scene illuminations.

We first define 3 objective and 2 subjective quantities [1]:

illuminance = amount of light incident on a surface
 luminance = amount of light hitting the sensor from a surface
 reflectance = proportion of incident light reflected by a surface
  lightness = perceived reflectance of a surface
 brightness = perceived luminance from the image itself

That is, image = luminance = illuminance × reflectance. For lightness, the observer is asked to judge the brilliance of physical paint on the surface depicted by the image, and for brightness, the grayscale level in the image itself.

For example, the center of letter *S* on the top and front surfaces in Fig. 1, i.e., 1 & 2, have the same reflectance, and appear of the same lightness. 1 has higher intensity and looks brighter than 2, due to higher illuminance on the top.

When brightness ranks differently from intensity, it tends to agree with lightness. In Fig. 1, the shaded gray patch 6 is darker than the illuminated black patch 5, but it appears brighter as well as on a lighter surface. Such an intimate connection between brightness and lightness makes brightness modeling inseparable from lightness modeling.

Existing models can be roughly divided into two camps with respect to the Helmholtz-Hering debate [11]. While the Hering camp seeks a low-level physiological cause (e.g., lateral inhibition [8], multiscale filtering [4]), the Helmholtz camp seeks a high-level cognitive cause, in an attempt to recover reflectance from luminance with unknown illumination (e.g., Retinex [13, 9], intrinsic images [3, 2, 19, 17]).

Gestalt theory, emphasizing mid-level perceptual organization, challenged Hering's sensory theory and Helmholtz's cognitive theory, in *a series of devastating crucial experiments* (Fig. 15) that *turned the field upside down* [7].

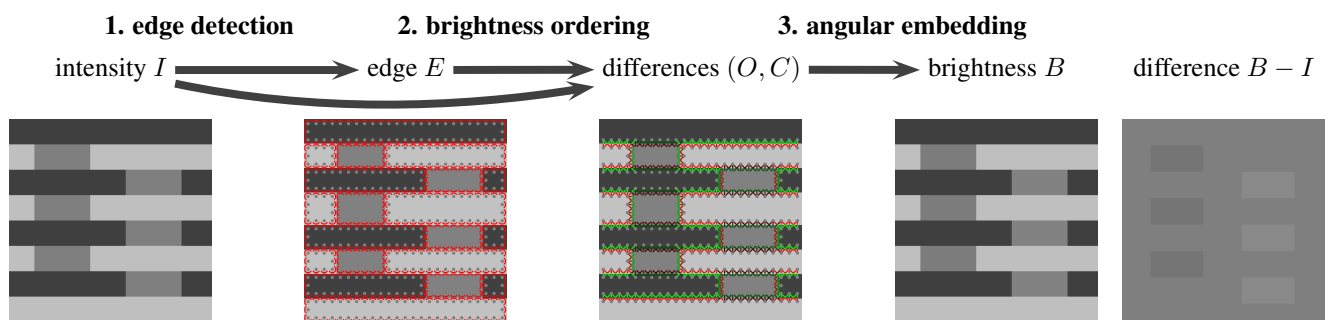There are also models focusing more on replicating the

1. edge detection   **2. brightness ordering**   **3. angular embedding**

intensity I ⟶ edge E ⟶ differences (O, C) ⟶ brightness B difference $B - I$

Figure 3. Model overview. **1)** Pairwise edges E in the intensity image I are detected by a pixel-centric edge detector. **2)** Pairwise brightness ordering O is established with confidence C based on E, I and its smoothed version I_σ. **3)** Global brightness ordering B is computed as the angular embedding of pairwise measurements (O, C). Grays of equal intensity exhibit different brightness desired by the illusion.

neurodynamics of the human visual system without clear computational objectives [8, 12], or establishing mathematical frameworks without specific features [6].

We approach brightness, not as a by-product of lightness computation, but as a pure modeling problem which must account for brightness illusions. Given an intensity image I, we output a brightness image B such that $B(a) > B(b)$ when pixel a appears brighter than b, even if $I(a) \le I(b)$.

Our challenge is that brightness illusions such as Fig. 2 are long-standing puzzles in human vision research. There is no coherent verbal explanation for their seemingly contradictory effects. Simultaneous contrast (SC) says a lighter surround darkens the center. White argues it could also assimilate and brighten the center. Anti-snake shows that SC can be annihilated by simply zigzagging the boundaries of the surround. Snake argues that adding a few remote flanks can overcome the annihilation and make SC even stronger.

a: simultaneous contrast (SC)

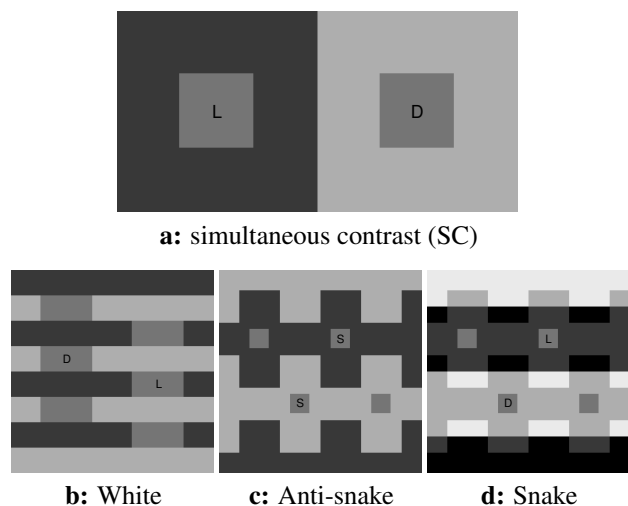**b:** White   **c:** Anti-snake   **d:** Snake

Figure 2. Basic illusions. Each test patch is a neutral gray (0.5) surrounded by darkgray 0.25, lightgray 0.75, black 0, or white 1. Identical test patches in the same image could appear lighter (L), or darker (D), or the same (S). These illusions have not been convincingly explained individually or reconciled collectively.

Brightness perception is global brightness ordering, where brightness judgement between two pixels needs only *their* ordering numbers, not any others'. Local brightness ordering is often easy to establish: For adjacent pixels a and b, if intensity has $I(a) > I(b)$, brightness must have $B(a) > B(b)$. The puzzle then becomes: **If local brightness ordering is consistent with (global) intensity ordering almost everywhere in the image, how come global brightness ordering is not the intensity itself?** Any brightness model thus must address 3 key issues.

1. **Feature**. Where and how local brightness ordering becomes inconsistent with intensity? We discover that (Section 2) this happens at edges, with corners or high curvature places introducing more inconsistency than lines. Local brightness ordering for adjacent pixels is measured by jarring intensity differences at two scales: fine scale for interiors and coarse scale across edges. These short-range cues enable brightness illusions.

2. **Aperture.** How can such an inconsistency at very few local areas extend to the entire image? Our idea is that local measurements do not have the same level of confidence. Local brightness ordering for distant pixels adopt intensity differences of higher confidence, which tend to originate from corners. These long-range cues reinforce brightness illusions.

3. **Integration.** How can multiple local brightness orderings yield a global brightness ordering? We propose *angular embedding*, a new and better alternative to LS [5], NCUTS [16, 20] and LLE [15], for finding a global ordering that maximally fulfills these local orderings in accordance with their confidence levels. The global brightness ordering realizes brightness illusions.

Brightness perception is thus analogous to motion perception, where local motion cues, least ambiguous at corners, need to be integrated to yield a global motion percept.

Our model has 3 steps (Fig. 3). It first requires the detection of edges in the presence of junctions, for which we develop a pixel-centric edge detector (Section 5). We then

derive local brightness ordering and its confidence between pairs of (adjacent and distant) pixels (Section 3). Integrating these pairwise orderings with angular embedding yields a global brightness ordering (Section 4). We present our results on illusions in Section 6, and conclude in Section 7.

2. Key Insight from Simultaneous Contrast

Brightness is so puzzling that there has not even been a satisfactory account for the SC illusion alone (Fig. 2**a**).
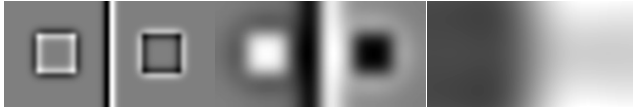

a: scale too small **b:** scale just right **c:** scale too large

Figure 4. Center-surround filtering of the intensity image produces a result that correlates with brightness only at the right scale. The filter is simply the difference of Gaussians: $G(x, \sigma) - G(x, 2\sigma)$.

The textbook explanation of center-surround filtering requires the right scale to produce a result that *roughly* correlates with brightness (Fig. 4**b**). Too small a scale creates halos along boundaries and no difference between interiors (Fig. 4**a**); too large a scale washes out the patches and opposite brightness conclusions can be drawn (Fig. 4**c**).

For the gray in SC to appear brighter on black than on white, the gray-black and gray-white contrast must be enhanced with respect to the black-white contrast (Fig. 5**a**). One way to achieve this, since gray-black or gray-white has a smaller contrast than black-white, is to enhance small differences. This selective enhancement by size explains SC, but not double-decrement SC [7] (Fig. 5**b**).
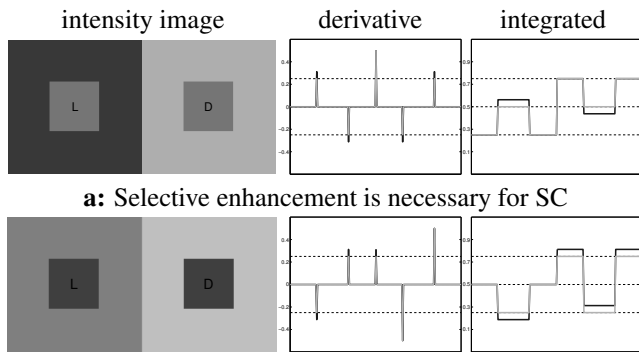
a: Selective enhancement is necessary for SC

b: Selection by size is wrong for double-decrement SC

Figure 5. SC must come from enhancing testpatch-surround differences with respect to surround-surround differences (**Row 1**), and this cannot be achieved by enhancing small differences, which explains SC but not double-decrement SC (**Row 2**). Columns 1-3 are images, derivative profiles along the middle horizontal lines, and profiles upon integrating the derivatives. Light gray for intensity, dark gray for brightness. In Row 2, both patches are decrements to their surrounds, and the left still appears lighter. Enhancing small differences wrongly predicts that the left patch appears darker.
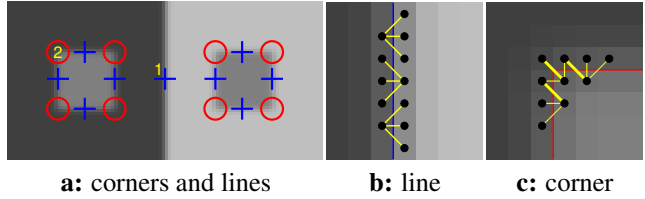

a: corners and lines **b:** line **c:** corner

Figure 6. Intensity differences are enhanced around corners at a coarser scale. **a)** Circles and crosses mark respectively corners and lines in the slightly blurred SC image. **b-c)** Local patches at line 1 and corner 2 marked in **a** respectively. Adjacent pixels (black dots) across (red) corner edges have larger intensity differences (thicker yellow lines) than those across (blue) straight line edges.

Our insight is that local intensity differences change at different rates with the scale, and the right type of enhancement emerges at a coarser scale across edges.

At a slightly coarser scale, while the intensity differences across a line remain uniform and proportional to their fine-scale counterparts, they increase around a corner. This relative enhancement is exactly what SC needs (Fig. 6).

Our way to achieve the selective enhancement of Fig. 5**a** is therefore not by the *size* of edges, but by the *geometry* of edges. The geometry is nevertheless not explicitly analyzed, but implicitly encoded in coarser-scale intensity differences.

3. Pairwise Brightness Ordering

We establish local brightness ordering O and its confidence C, based on detected edges. For image I of n pixels, $E_{n \times n}$ is a binary edge indicator: $E(a, b) = 1$ if an edge exists between adjacent pixels a and b.

3.1. Rationale and Short-Range Cues

Local brightness ordering O between adjacent pixels is a mixture of intensity differences at two scales: coarse scale ΔI_σ in the presence of an edge, and fine scale ΔI in the absence of an edge. Its confidence C increases with the edge magnitude, and 0 in the absence of an edge.

ordering:	$O = \Delta I_\sigma \cdot E + \Delta I \cdot (1 - E)$	(1)
confidence:	$C = |O| \cdot E$	(2)

Δf denotes local differences of f: $\Delta f(a, b) = f(a) - f(b)$, $\forall a, b$. I_σ is I smoothed with a Gaussian with a (fixed) small standard deviation σ, and scaled such that $\Delta I_\sigma(a, b) = \Delta I(a, b)$ for a, b separated by a step edge (Fig. 6b).

The presence of edges alters local measurements. While this is motivated from the computational necessity of transforming intensity into brightness (Fig. 5**a**), it can be understood as discounting ambient illumination across adjacent surfaces, which is not so perceivable on the same surface.

A light surface abutting a dark one in a real scene is often seen with a luminance transition, as ambient lighting cannot

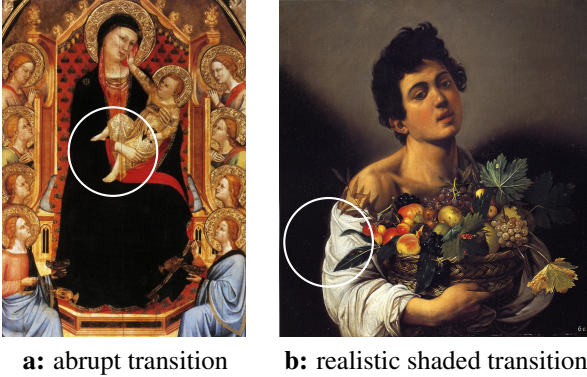**a:** abrupt transition   **b:** realistic shaded transition

Figure 7. Realistic depiction requires value transition even across sharp boundaries. **a)** *Daddi: Madonna and Child*, Tempera on Canvas, 1347. **b)** *Caravaggio: Boy with a Basket of Fruit*, Oil on Canvas, 1593. The light-dark transition in the circled area is more abrupt in **a** than in **b**, which shows convincing depth and volume.

a: junction   **b:** general neighbourhood

c: initialized   **d:** corrected   **e:** propagated

Figure 8. Resolving inconsistency from multiple paths. **Row 1)** We have yellow center path Q, green side path P, red edges and blue non-edges. **a)** Winning path P^* at a T-junction for pixels a, b has the largest single-edge difference. It is used to scale $O(a, m)$ and $O(m, b)$ so that path a, m, b and P^* result in the same $O(a, b)$. **b)** Multiple side paths are used to update $O(a, b)$ for pixel a and its neighbour b separated by an edge, shown here at $90°$ and $-45°$. **Row 2)** $O(a, b)$ is visualized from a towards b in red (positive) or green (negative) lines, darker lines for smaller differences. The elevated (**c**) diagonal contrast of the gray pixel at the T-junction is scaled down (**d**) by the black-white difference of path P^* in **a**, and the correction is propagated (**e**) to other pixels.

drop abruptly. Fig. 7 is a case in point: depictions with abrupt transitions across depths look disturbingly fake.

We formalize the rationale of our local brightness ordering as follows. Lighting L varies relatively slowly and its variation is perceived only at edges E and consequently discounted from brightness B.

$$\text{From } B + L = I \tag{3}$$

$$\text{For } E = 0, \text{ assuming } \Delta L = 0, \text{ we have } \Delta B = \Delta I; \tag{4}$$

$$\text{For } E = 1, \Delta B_\sigma + \Delta L_\sigma = \Delta I_\sigma, \tag{5}$$

$$\text{assuming: } \Delta B = \Delta B_\sigma, \quad \Delta L_\sigma = 0, \tag{6}$$

$$\text{we have: } \Delta B = \Delta B_\sigma = \Delta I_\sigma - \Delta L_\sigma = \Delta I_\sigma. \tag{7}$$

$$\text{That is, } O = \Delta B = \Delta I_\sigma \cdot E + \Delta I \cdot (1 - E). \tag{8}$$

3.2. Aperture and Long-Range Cues

Given O and C for adjacent pixels, we extend them to pixels a, b at distance $d(a, b) > 1$, connected through path $P = (P_1, P_2, \ldots, P_{n+1})$, with $a = P_1, \ldots, P_{n+1} = b$:

$$O(P) = \sum_t O(P_t, P_{t+1}) \tag{9}$$

$$C(P) = \max_t C(P_t, P_{t+1}) \tag{10}$$

However, different P's for the same (a, b) can result in different $O(P)$'s, since O comes from intensity at two scales (Eqn. 1). This is not simply a 2D integrability problem for noisy signals [5, 10], but a modeling problem where signals are overwhelmed by structured "noise". It is the aperture problem frequently encountered in computer vision.

The challenge is to resolve inconsistency over a longer range so that the very few brightness cues enabled at corners and junctions become stronger than the majority of cues which agree with the non-illusion outcome: $B = I$.

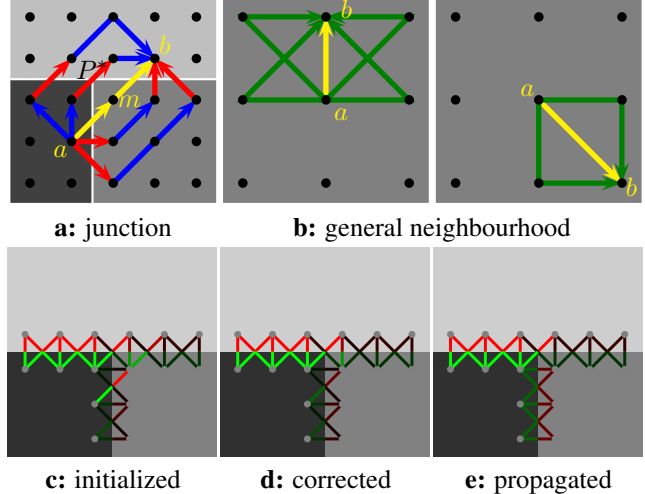Junction study reveals two essential resolution rules:

1. Paths made with edges of the same sign dominate;
2. Paths made with edges of higher confidence dominate.

We select winning path P^* based on these two principles and use $O(P^*)$ and $C(P^*)$ to define $O(a, b)$ and $C(a, b)$.

Formally, let S be the sign of differences at edges, i.e. $S = \text{sign}(O) \cdot E$. Let Q be the shortest path that connects a and b. Initializing $O(a, b) = O(Q)$ and $C(a, b) = C(Q)$, we examine neighbouring path P's to update O, C:

$$O(a, b) = O(P^*) \tag{11}$$

$$C(a, b) = C(P^*) \tag{12}$$

$$P^* = \arg \max_{\substack{E(P) = |S(P)| \\ C_{min}(P) \geq C(Q)}} C_{sum}(P) \tag{13}$$

$$\text{where: } E(P) = \sum_t E(P_t, P_{t+1}) \tag{14}$$

$$S(P) = \sum_t S(P_t, P_{t+1}) \tag{15}$$

$$C_{sum}(P) = \sum_t C(P_t, P_{t+1}) + 1 - E(P_t, P_{t+1}) \tag{16}$$

$$C_{min}(P) = \min_t C(P_t, P_{t+1}) \tag{17}$$

We use the same update rule in two scenarios (Fig. 8). One is at junctions, where $O(P)$ from two edges pointing in the gradient direction, i.e., path $P = (a, m, b)$, is larger than $O(P^*)$ from a single edge (Fig. 8**c**). The confidence $C(P^*)$ is larger, thus $O(a, m)$ and $O(m, b)$ are scaled down

(Fig. 8**d**) to yield $O(P) = O(P^*)$. The other scenario is to update each pixel's local ordering with its edge-separated neighbours. This process helps propagate the correction initiated at junctions (Fig. 8**e**), and it is recursively applied with an increasing neighbourhood radius to obtain consistent long-range cues that reinforce brightness illusions.

4. Angular Embedding

Now we have pairwise orderings in (arbitrary) size O and (nonnegative) confidence C for n elements. We assume $C(a,b) \leq 1$, $C(a,a) = 1$, $O(a,a) = 0$, $\forall a, b$.

The goal of an embedding is to find a global ordering of the n elements where pairwise differences in the global order match those pairwise local ordering measurements.

While conventional embedding realizes the order in the positions of points on a line, our angular embedding (AE) realizes it in the angles of points on a unit circle.

AE is a direct extension of LLE [15] from linear value interpolation to arbitrary neighbour relationships. AE is more robust and efficient than Poisson diffusion or LS [5], as it is a nonlinear criterion with eigensolutions. AE is more versatile than NCUTS with attraction and repulsion [20], as it handles measurements with arbitrary size and confidence.

4.1. Representation and Optimality Criterion

The goal of AE is to find angles, θ, that fulfill pairwise orderings O and C (Fig. 9). Point a on the unit circle can be written as a complex number with its angle $\theta(a)$:

$$z(a) = e^{j\theta(a)}, \quad j = \sqrt{-1} \tag{18}$$

Point b projects a's position to be $e^{j(O(a,b)+\theta(b))}$ with confidence $C(a,b)$. The consensus of a's position from b's can be estimated as a confidence-weighted vector average:

$$\tilde{z}(a) = \sum_b \tilde{C}(a,b) \cdot e^{jO(a,b)} \cdot z(b) \tag{19}$$

$$\tilde{C}(a,b) = \frac{C(a,b)}{\sum_b C(a,b)} \tag{20}$$

Since $\sum_b \tilde{C}(a,b) = 1$ and $\tilde{C}(a,a) > 0$, we have $|\tilde{z}(a)| \leq |z(a)| = 1$. It is straightforward to show that:

$$|z(a) - \tilde{z}(a)| = 0 \tag{21}$$

$$\Leftrightarrow \quad |\tilde{z}(a)| = |z(a)| = 1 \tag{22}$$

$$\Leftrightarrow \quad O(a,b) + \theta(b) = \theta(a), \forall b \in \{b : C(a,b) > 0\}. \tag{23}$$

That is, $z(a)$ and $\tilde{z}(a)$ coincide *if and only if* every local ordering with positive confidence is perfectly fulfilled. $|z - \tilde{z}|$ is thus a good indicator of the quality of an embedding.

We formulate the AE criterion in terms of minimizing the disagreement or maximizing the agreement between z and $\tilde{z}$ for all a's, weighted by their total confidence. The two objectives are dual and thus equivalent:
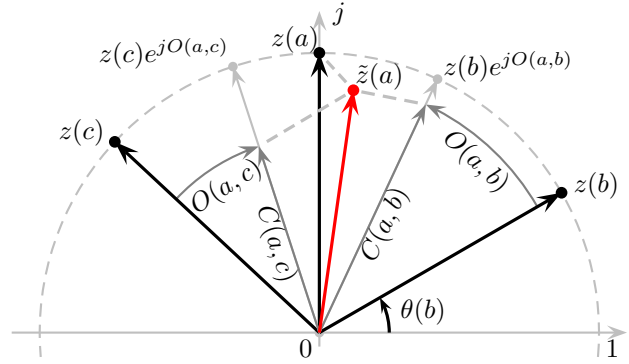

Figure 9. AE places elements on the unit circle such that their angular displacements best match local orderings. Consider element a placed at $z(a)$. a's neighbour b, based on its position $z(b)$ and difference $O(a,b)$ with a, projects a to be $z(b)e^{jO(a,b)}$ with confidence $C(a,b)$, so does neighbour c. The neighbourhood expectation of a's position $\tilde{z}(a)$ is in disagreement with $z(a)$. A desirable embedding minimizes the total disagreement for all a's.

disagreement: $\quad \varepsilon_d = \sum_a D(a) \cdot |z(a) - \tilde{z}(a)|^2 \tag{24}$

agreement: $\varepsilon_a = \sum_a D(a) \left(|z(a)|^2 - |z(a) - \tilde{z}(a)|^2 \right) \tag{25}$

where: $\quad D(a) = \dfrac{\sum_b C(a,b)}{\sum_{a,b} C(a,b)} \tag{26}$

hence: $\quad \varepsilon_d + \varepsilon_a = \sum_a D(a) = 1 \tag{27}$

4.2. Matrix Formulation

We rewrite AE in a matrix form. Let $'$ denote conjugate transposition, $\bullet$ the matrix Hadamard (entrywise) product, e^X the entrywise exponential function of matrix X, $\mathrm{Diag}(\cdot)$ the diagonal matrix formed from its vector argument, I_n the $n \times n$ identity matrix, and 1_n the $n \times 1$ vector of 1's.

	minimize	$\varepsilon_d(\theta; O, C) = z'Wz$	(28)
$\Leftrightarrow$	maximize	$\varepsilon_a(\theta; O, C) = z'(D - W)z$	(29)
	subject to	$z = e^{j\theta}$	(30)

$$\begin{aligned} \text{where} \quad & W = (I_n - M)'D(I_n - M) & (31) \\ & M = \mathrm{Diag}(C1_n)^{-1}C \bullet e^{jO} & (32) \\ & D = \mathrm{Diag}(C1_n \cdot (1_n'C1_n)^{-1}). & (33) \end{aligned}$$

W is a Hermitian error matrix; M is a measurement matrix that combines two aspects of a local measurement, size and (normalized) confidence, in one complex number, and D is a normalized total-confidence degree matrix.

The criterion ε_d in Eqn. 28 and the error matrix W in Eqn. 31 resemble LLE's [15]. It is not surprising, since Eqn. 19 formally extends LLE's scalar average to AE's vector average. Decoupling confidence C from size O allows a principled normalization for AE, which results in a dual objective, ε_a in Eqn. 29, with a sound and robust solution.

4.3. Eigenvector Solution

The optimal embedding for Eqns. (29-30) satisfies:

$$\theta^* = \measuredangle z^*, \quad z^* = \arg \max_{|z|=1_n} z'(D-W)z \qquad (34)$$

$|\cdot|$ denotes the entrywise magnitude of its argument. Relaxing $|z| = 1_n$ to its necessary condition $z'Dz = 1$, we have a Rayleigh quotient maximization problem. The global optimum is the leading eigenvector of $(D-W, D)$:

$$z^* \approx \arg \max_{z'Dz=1} \varepsilon_a(z) = V_1, \qquad \varepsilon_a(z^*) = \lambda_1 \qquad (35)$$

$$\tilde{P} \cdot V_k = \lambda_k \cdot V_k, \quad \forall k, \qquad \lambda_1 \geq \ldots \geq \lambda_n \qquad (36)$$

$$\tilde{P} = D^{-1}(D-W), \qquad (37)$$

where $\tilde{P}$ is akin to the generalized affinity matrix in [20], but admits cues other than attraction and repulsion.

Brightness B as the AE optimum for O and C is thus:

$$\boxed{B = \measuredangle \text{ leading-eigenvector-of } (D-W, D) \qquad (38)}$$

5. Pixel-Centric Edge Detection

Now we come to edge detection, which is crucial for our model. Edge detectors often fail in scenarios that violate the *single step edge* assumption (Fig. 10).

Existing methods assume a pixel is an isotropic Gaussian and an edge is where the intensity difference dominates in both forward and backward directions. Fig. 10 shows that, e.g. in a highlight, pixels on the left have little to do with the edge on the right; in an X junction, pixels in one quadrant have little to do with edges in the other 3 quadrants.

Therefore, some sense of pixel grouping must be built into the earliest stage of edge detection. Our idea is that:

1. Each pixel is a small Gaussian modified by local similarity;
2. Edge sensitivity varies according to local differences;
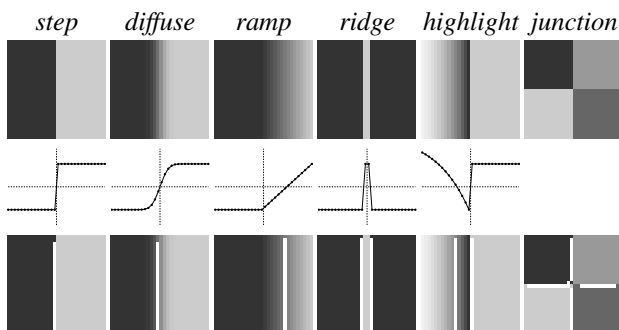3. An edge is a gradient discontinuity along *either* direction.

Figure 10. Intensity and spatial variations can cause displaced or missed edge detections. **Rows 1-3:** Images, intensity profiles along the middle horizontal lines, and edges by a Canny detector. Cases 1-5 are essentially 1D problems. Case 6 is a 2D problem: As multiple edges meet at a junction, weaker edges are easily missed.
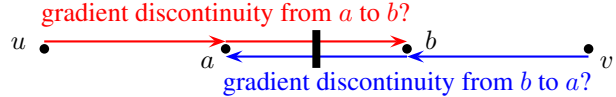

Figure 11. Our edge detector treats gradient discontinuity in two directions separately. $E(a, b) = 1$ if the average intensity in a small neighbourhood changes sharply over $u \to a \to b$ or $v \to b \to a$.

Consider Fig. 11. Pixel a tries to decide if an edge exists between itself and its neighbour b. It evaluates whether there is a discontinuity in the average intensity $\bar{I}$ over $u \to a \to b$ or $v \to b \to a$, the average computed with weights A dependent on distance d and intensity difference $|\Delta I|$. If it changes monotonically, and the difference $\Delta \bar{I}(a, b)$ is much larger than $\Delta \bar{I}(u, a)$ or $\Delta \bar{I}(b, v)$, a concludes $E(a, b) = 1$.

Let $G(\cdot, \sigma)$ be a Gaussian with mean 0 and standard deviation σ. There is a discontinuity over $u \to a \to b$ if:

$$\text{monotonicity: } \Delta\bar{I}(u, a) \cdot \Delta\bar{I}(a, b) > 0 \qquad (39)$$

$$\text{dominance: } |\Delta\bar{I}(a, b)| > \alpha \cdot |\Delta\bar{I}(u, a)| + \beta \qquad (40)$$

$$\Delta\bar{I}(a, b) = \frac{\sum_p (A(a, p) - A(b, p)) \cdot I(p)}{\sum_p |A(a, p) - A(b, p)|} \qquad (41)$$

$$A(a, p) = \frac{G(\Delta I(a, p), \sigma_e) \cdot G(d(a, p), \sigma_d)}{\sum_p G(\Delta I(a, p), \sigma_e) \cdot G(d(a, p), \sigma_d)} \qquad (42)$$

$$\text{adaptive intensity sensitivity: } \sigma_e = \Delta I(a, b)/3 \qquad (43)$$

where σ_d dictates the neighbourhood size, and α and β are magnitude thresholds for determining discontinuity.

This pixel-centric gradient discontinuity detection can be further improved with contextual reasoning such as bilateral gap completion. Note that all we need is edges, which are simply difficult to detect at junctions (Fig. 12). Junctions are neither detected separately nor identified explicitly.
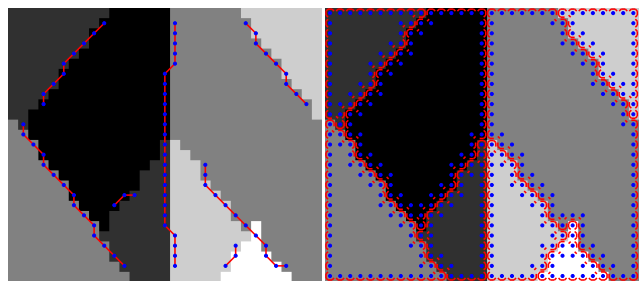
a: Canny with threshold 0   **b:** pixel-centric edge detector

Figure 12. Pixel-centric edge detection does not miss edges around junctions. **a)** Canny edge detector labels a pixel as an edge or not; nearby edge pixels are linked. It fails to detect many weaker edges even with threshold 0. **b)** Our detector finds edges from each pixel's point of view in 8 directions. A small line is drawn for every edge between a pixel and its neighbour.

6. Results and Discussions

Given an intensity image I, we first use our pixel-centric edge detector to find pairwise edges E. We then derive pairwise brightness ordering O and confidence C, from which we compose Hermitian error matrix W and degree matrix D. We compute the leading eigenvector of $(D - W, D)$ and obtain brightness B in the angles of its components.

We scale the range of B to match that of I, i.e., B and I have the same extreme values of e.g. 0 and 1.

We show results on various junctions (Fig. 13), basic (Fig. 14) and additional brightness illusions (Figs. 15-16). The distinction between brightness and intensity is clearer in their difference images or superimposed profile plots.
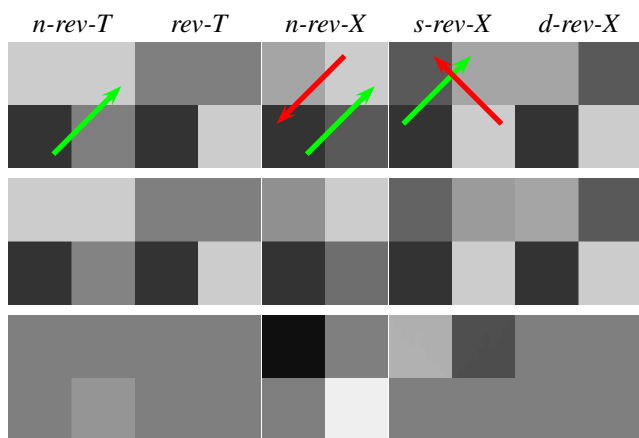

Figure 13. Junction study. **Row 1:** Intensity I. **Row 2:** Brightness B. **Row 3:** Difference $B - I$. There is no illusion induced by reversing T or double-reversing X, but a little by non-reversing T, more by single-reversing X, and most by non-reversing X. All these effects can be singly understood as deviation along the intensity gradient direction (marked by arrows in Row 1).

In Fig. 13, grays need to establish themselves with respect to black and white, which have the lowest and highest brightness. In the non-reversing T, the gray is angled between black and white. Since there is more white than black, the average intensity difference is smaller for gray-white than for gray-black, causing the gray to be closer to white and thus appear lighter. In the reversing T, the gray is not in the way of, but side by side with an equal amount of black and white. Its brightness stays in the middle → no illusion. Similar reasons hold for 3 types of X-junctions.

Our junction study provides a computational account for brightness effects from junctions [18]. In particular, transparency, haze, or clear atmosphere can be predicted without factoring luminance into illuminance and reflectance [1].

Transparency maximally alters brightness from intensity. It is mediated by non-reserving X-junctions, where opposing black and white reduce the difference between the two sandwiched grays, making the light-gray next to white
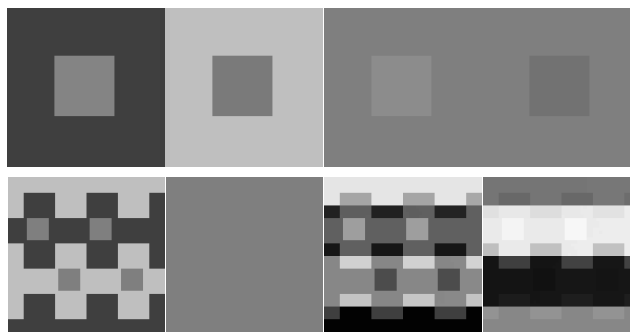

Figure 14. Brightness results for Fig. 2, shown in brightness B and difference $B - I$. See White in Fig. 3. Snake has the largest illusion, with double effects from the corners of test patches and non-reversing X junctions, while Anti-snake has no illusion as balanced local enhancements from SC do not alter the global ranking.
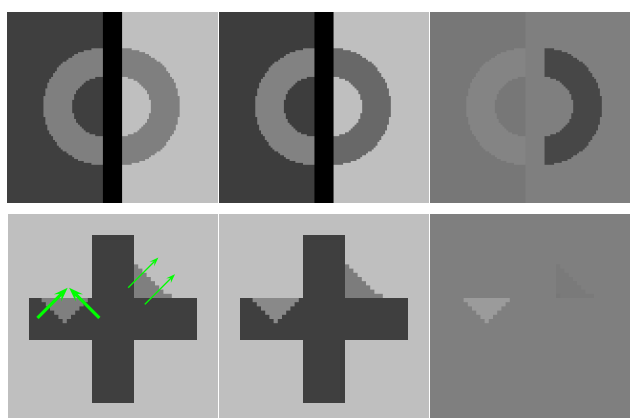

Figure 15. Results for Kanizsa ring (**Row 1**) and Benary illusions (**Row 2**), shown in intensity I, brightness B, and difference $B - I$. In Benary, arrows mark stronger enhancement for the left triangle.

darker and the dark-gray next to black lighter, their distinction increased. **Haze** also makes brightness different from intensity, but to a lesser extent. It is mediated by single-reversing X-junctions, where side-by-side black and white also reduce the difference between abutting grays, giving a more pronounced local contrast between hazy grays and clear black-and-white. **Clear** atmosphere does not make brightness any different from intensity. It is mediated by double-reversing X-junctions of no change effect at all.

The characteristic local brightness effects of corners and junctions can explain many illusions such as SC, White, Snake (Fig. 14), Koffka ring (Fig. 15), double-decrement SC, Articulation, Crisscross, and COCE (Fig. 16).

While brightness can be primarily understood as the consequence of local brightness ordering deviating from global intensity ordering in its gradient direction, it can only be fully appreciated as a Gestalt.

As a global ranking, it is not the absolute sizes but the relative sizes of local contrast that matter. **1)** Anti-snake has the SC illusion completely turned off (Fig. 14), not because
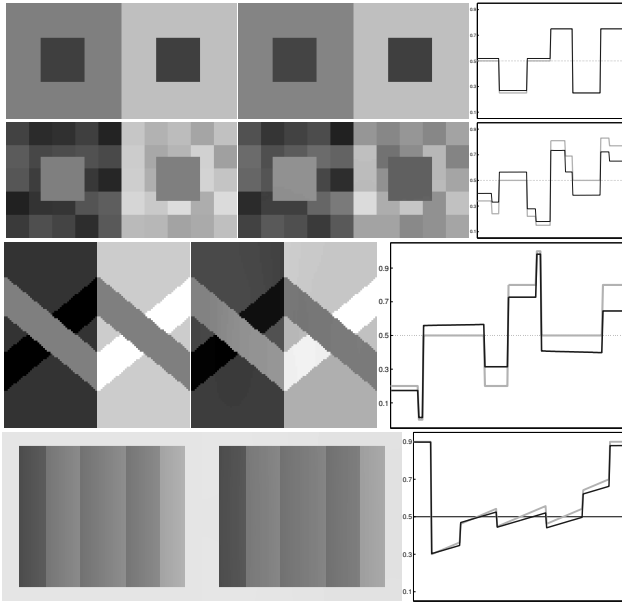
Figure 16. Brightness results for double decrement SC (**Row 1**), Articulation (**Row 2**), Crisscross (**Row 3**), and COCE (**Row 4**), shown in intensity I, brightness B, and their profiles (gray for intensity, black for brightness) along the middle horizontal line

there is no local enhancement from zigzagging boundaries, but because the gray's differences to the two extremes, black and white, are equally large so that they stay remain in the middle rank. **2)** For Benary (Fig. 15), the triangle inside black is lighter than the one next to black, because of the larger enhancement for the former, not because of a lack of enhancement for the latter. **3)** Articulated surrounds not only introduce more junctions for local enhancement, but also force all the grays to spread out on the ranking scale, making each more distinctive (Fig. 16).

7. Summary

We model brightness as an emergent global order from intensity differences at two (fixed) scales mixed according to edges, identified by our pixel-centric edge detector.

Brightness is modeled not by modifying the size of local comparisons as Retinex [13, 9], or by requiring edge classification or segmentation as the selective integration model [14], or by knowing the right scale as the multiscale filtering model [4], but by having local comparisons dependent on the geometry of edges.

We explain a variety of brightness illusions with a single algorithm and a clear intuition. Brightness of a pixel can be understood locally as its intensity deviating in the gradient direction and globally as finding its rank relative to others, particularly the lightest and darkest ones.

We propose AE as a more robust, efficient, and versatile method to integrate pairwise comparisons than LS, LLE, and NCUTS. AE will thus find its use in many areas.

References

[1] E. H. Adelson. Lightness perception and lightness illusions. In M. Gazzaniga, editor, *The cognitive neurosciences*. MIT Press, 1999.

[2] E. H. Adelson and A. Pentland. The perception of shading and reflectance. In Knill and Richards, editors, *Peception as Bayesian Inference*. Cambridge, 1996.

[3] H. G. Barrow and J. M. Tenenbaum. Recovering intrinsic scene characteristics from images. *Computer Vision Systems*, pages 3–26, 1978.

[4] B. Blakeslee, W. Pasieka, and M. McCourt. Oriented multiscale spatial filtering and contrast normalization: a parsimonious model of brightness induction in a continuum of stimuli including White, Howe and simultaneous brightness contrast. *Vision Research*, 2005.

[5] R. T. Frankot and R. Chellappa. A method for enforcing integrability in shape from shading algorithms. *PAMI*, 10(4):439–51, 1988.

[6] T. Georgiev. Covariant derivatives and vision. In *ECCV*, 2006.

[7] A. Gilchrist. *Seeing Black and White*. Oxford University Press, 2006.

[8] S. Grossberg. The quantized geometry of visual space: The coherent computation of depth, form, and lightness. *Behavioral and Brain Sciences*, 6:625–57, 1983.

[9] B. K. P. Horn. Determining lightness from an image. *Computer Graphics and Image Processing*, 1974.

[10] B. K. P. Horn and M. J. Brooks. *Shape from Shading*. MIT Press, 1989.

[11] I. P. Howard. The Helmholtz-Hering debate in retrospect. *Perception*, 1999.

[12] F. Kelly and S. Grossberg. Neural dynamics of 3D surface perception: figure-ground separation and lightness perception. *Perception & Psychophysics*, 2000.

[13] E. H. Land and J. J. McCann. Lightness and retinex theory. *JOSA*, 1971.

[14] W. D. Ross and L. Pessoa. Lightness from contrast: a selective integration model. *Percept Psychophys*, 2000.

[15] S. Roweis and L. Saul. Nonlinear dimensionality reduction by locally linear embedding. *Science*, 2000.

[16] J. Shi and J. Malik. Normalized cuts and image segmentation. In *CVPR*, 1997.

[17] M. Tappen, W. T. Freeman, and E. H. Adelson. Recovering intrinsic images from a single image. In *NIPS*, 2002.

[18] D. Todorovic. Lightness and junctions. *Perception*, 1997.

[19] Y. Weiss. Deriving intrinsic images from image sequences. In *ICCV*, 2001.

[20] S. X. Yu and J. Shi. Segmentation with pairwise attraction and repulsion. In *ICCV*, 2001.