Effective signaling of surface boundaries by L-vertices reflect the consistency of their contrast in natural images

Edward A. Vessel
Department of Neuroscience, Max Planck Institute for Empirical Aesthetics, Frankfurt am Main, Germany

Irving Biederman
Department of Psychology & Neuroscience Program, University of Southern California, Los Angeles, CA, USA

Suresh Subramaniam
Whizz Systems, Santa Clara, CA, USA

Michelle R. Greene
Department of Computer Science, Stanford University, Stanford, CA, USA

An L-vertex, the point at which two contours coterminate, provides highly reliable evidence that a surface terminates at that vertex, thus providing the strongest constraint on the extraction of shape from images (Guzman, 1968). Such vertices are pervasive in our visual world but the importance of a statistical regularity about them has been underappreciated: The contours defining the vertex are (almost) always of the same direction of contrast with respect to the background (i.e., both darker or both lighter). Here we show that when the two contours are of different directions of contrast, the capacity of the L-vertex to signal the termination of a surface, as reflected in object recognition, is markedly reduced. Although image statistics have been implicated in determining the connectivity in the earliest cortical visual stage (V1) and in grouping during visual search, this finding provides evidence that such statistics are involved in later stages where object representations are derived from two-dimensional images.

Introduction

Robust statistical properties of natural images are generally believed to determine the tuning and neural connectivity among cells in early cortical visual areas (e.g., V1; Field, 1987; Olshausen, 1996), and there is growing evidence that the tuning in later visual areas to variations in shape is similarly determined (Elder & Goldberg, 2002; Geisler & Perry, 2009; Geisler, Perry, Super, & Gallogly, 2001; Ramachandra & Mel, 2013). However, while the relationships between second order properties of contour segments such as proximity, good continuation (e.g., angle), and similarity (intensity, contrast) have been well characterized, the key role of L-vertices in the construction of object representations has been underappreciated. An L-vertex (Figure 1), produced by the cotermination of two contours at different orientations, is of critical importance for signaling the termination of a surface, and hence, often, a discontinuity in depth. Computer vision algorithms for parsing scenes from contour features suggest that the L-vertex likely imposes the strongest constraint, compared to other vertices and contours, in assigning surfaces to objects (e.g., Guzman, 1968; Waltz, 1975).

The effectiveness for such signaling by L-vertices in human perception was shown by Donnelly, Humphreys, and Riddoch (1991) and modeled by Hummel and Biederman (1992). Hummel and Biederman's model provided a neurocomputational explanation of Blickle's (1989) account of Bregman's (1981) classical demonstration that a display of otherwise uninterpretable random-appearing fragments can be interpreted as a set of four block-letter *Bs* when an occluding mass, resembling spilled ink, is added to the display covering portions of the *Bs*. Blickle showed that the difficulty in interpretation of the fragments was completely accounted for by the addition of the ink's contours to the fragments of the *Bs*, producing spurious L-vertices that inhibited the smooth continuation that would have joined the fragments into readily interpretable *Bs*. Thus removing *all* the contours of the spilled ink rendered the *Bs* even more readily interpretable than when they were shown under the occluding ink (Hummel & Biederman, figure 42, p. 512). Blickle also showed that

Citation: Vessel, E. A., Biederman, I., Subramaniam, S., & Greene, M. R. (2016). Effective signaling of surface boundaries by L-vertices reflect the consistency of their contrast in natural images. *Journal of Vision*, *16*(9):15, 1–10, doi:10.1167/16.9.15.

doi: 10.1167/16.9.15 Received September 5, 2015; published July 29, 2016 ISSN 1534-7362

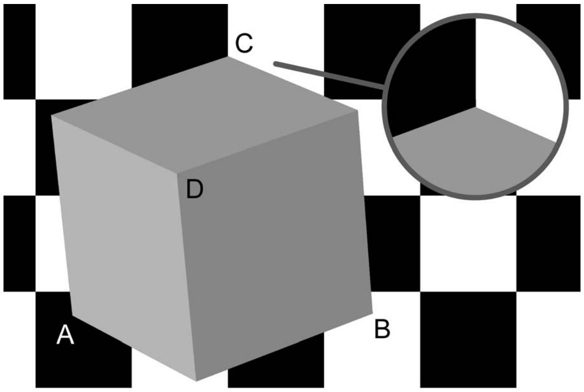

Figure 1. The two legs of an L-vertex are almost always of the same contrast polarity with respect to the background—that is, both lighter (as in A) or both darker (as in B). The uppermost corner of this cube (C) illustrates a highly unlikely, accidental view in which a corner of a midgray surface appears exactly aligned with the change from black to white in the background, producing an accidental Y (or fork) vertex in which three contours coterminate with none of the angles greater than 180°. A slight change in the observer's viewpoint, the position of the object, or the position of the background would eliminate this alignment, leaving only an L-vertex with two legs of consistent contrast polarity at the top of the cube.

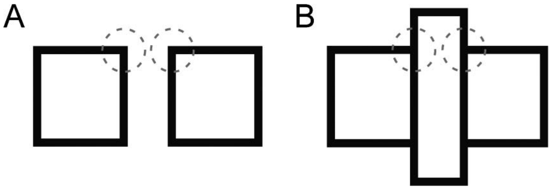

Figure 2. (A) L-vertices are evidence for the termination of a surface and often a depth discontinuity at the boundary of a surface. (B) T-vertices are evidence for occlusion.

the identical explanation could account for Leeper's (1935) demonstration that otherwise uninterpretable blotchy fragments of drawings of common objects (e.g., an elephant) could be readily interpreted if the blotches, which produced spurious L-vertices that inhibited grouping, were eliminated leaving only the object's contours.

Cells tuned to L-vertices are found in cortical area V4 (Pasupathy & Connor, 1999), where they contribute critically to a contour-based population code for shape (Pasupathy & Connor, 2002). As opposed to the relatively high likelihood that two segments of an extended contour are of opposite contrast polarity (e.g., Elder & Goldberg, 2002; Geisler & Perry, 2009), it is extremely rare in natural images for the two legs of an L-vertex to be of different directions of contrast, such that one leg is darker and the other leg lighter than the background. Indeed, there exists experimental evidence that contrast reversals at regions of high curvature, as opposed to along extended contours, adversely affects perceptual closure as a cue in visual search (Elder & Zucker, 1993; Spehar, 2002) as well as the fMRI BOLD response in early visual areas (Schira & Spehar, 2011). Here we show that the effectiveness of an L-vertex to signal the termination of a surface requires that its legs be of the same direction of contrast, thus reflecting the statistics of natural images, including those containing manufactured objects. That this sensitivity to direction of contrast is not just to regions of high curvature but specific to vertex type is shown by our finding of reduced sensitivity to direction of contrast in T-vertices (relative to L-vertices), which also can be regarded as containing regions of high curvature.

Consider Figure 1, which shows a gray cube against a black and white checkerboard background. There are two L-vertices, with A and B representing the expected contrast relations where the two coterminating contours at A are lighter than the black background square, and at B both are darker than the white background square. Both A and B are effective in conveying that the surface terminates in front of (and hence occludes) the background. Vertex C, which would also be an L-vertex if considered solely from the perspective of the cube independent of the background, is "accidentally" aligned with the border between a black and white square, such that one leg is darker and the other leg lighter than the background in the region of the vertex. Locally, vertex C, unlike A and B, does not convey that the gray surface is in front of the black and white surfaces (although that inference might be made from the rest of the figure, aided by vertices A and B). Instead, if we zoom in on vertex C (inset of Figure 1), the interpretation is of a Y- (or fork) vertex defining a junction of *three* surfaces (one black, one white, and the third gray with none of the angles between the three segments greater than 180°). A fork vertex can define a convex three-dimensional corner of a cube similar to vertex D. (The local percept at vertex C is bistable, sometimes appearing convex, as vertex D, and sometimes concave, as when viewing the junction of two walls and the floor inside a room.)

L-vertices signal the termination of a surface and, often, a discontinuity in depth, as illustrated in Figure 2A (and Figure 1). T-vertices, by contrast, in which a contour terminates along the length of another contour, as illustrated in Figure 2B, provide evidence for occlusion (Guzman, 1968). The contour that terminates (the stem of the T) is interpreted as being behind the contour that is continuous at the junction (the top of the T). Matched T-vertices, in which a pair of terminating segments have collinear stems, are interpreted by the visual system as evidence that the stems define an edge of a surface that continues behind the occluder, as in Figure 2B (Guzman, 1968). In natural scenes, in general, the stem and top of a T-

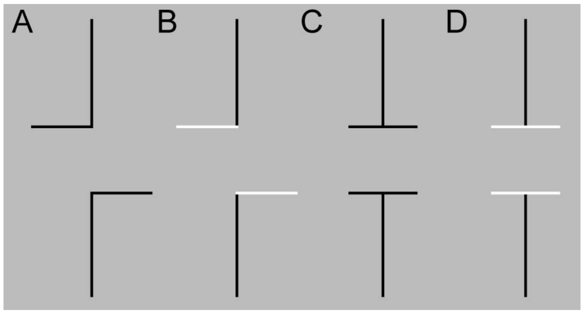

Figure 3. The junction types produced by changing the alignment and contrast polarity of line segments added across a gap in an extended contour. (A) L-junctions with legs of the same contrast polarity, (B) L-junctions with legs of different contrast polarity (C) T-junctions with legs of the same contrast polarity, and (D) T-junctions with legs of different contrast polarity, Gaps terminating in L-junctions as illustrated in (A) would be expected to most effectively suppress smooth continuation and, therefore, result in a greater decrement in object recognition performance than (B). The contrast polarity for T-junctions should have little or no effect.

vertex (i.e., the occluding and occluded surfaces) can have any contrast relation with respect to each other or the background. Therefore, unlike the signaling of a surface termination by L-vertices, the signaling of continuation through an occluder by T-vertices should be less affected by contrast polarity.

We tested whether the effectiveness of an L-vertex to signal the termination of a surface, as reflected in the speed and accuracy of object recognition, is affected by whether its two segments were of the same or different contrast polarity. Given the statistics of natural images in which the segments comprising an L-vertex are almost always of the same direction of contrast, if these statistics are incorporated into the effectiveness of an L-vertex to signal the termination of a surface, we would expect L-vertices with segments of different directions of contrast to be less effective in such signaling. It would be quite common, however, for the head and stem of a T-vertex to be of different directions of contrast, because an occluder could be of any luminance relative to the surface it occludes. We therefore expect that T-vertices would be less affected by variations in the contrast polarity of their segments.

Experimental strategy

Subjects named, as quickly as possible, line drawings of common objects or animals in which multiple gaps were introduced along the longer contours of the drawing. A pair of line segments, perpendicular to the object's contours at the gap, was inserted at each end of the gap to produce either L-vertices or T-vertices (Figure 3). The segments forming the two L-vertices were oriented in opposite directions from each other. The objects in Experiment 1 were drawn as black lines on a gray background or as white lines on a gray background, and the line segments in the gap were either black or white. Figure 3 (Panels A–D) depicts the four combinations of gap vertex and contrast polarity. The vertical contours represent a section of an object's contour with the segment, here horizontal at each gap, producing either an L- or a T-vertex that could be of the same or reversed contrast with respect to the object's contours.

If both segments of an L-vertex must have the same direction of contrast against the background in order to signal the termination of a surface, then gaps configured as in Panel A in Figure 3 should produce greater difficulty in object recognition than Panel B—because the L-vertices signal the (inappropriate) termination of the contour and thus suppress the smooth continuation across the gap that would otherwise have occurred without the L-vertices (Donnelly et al., 1991; Hummel & Biederman, 1992). For T-vertices, Panels C and D, no effect of contrast polarity would be expected, aside from a potential separation of the added segments from the object contour on the basis of luminance grouping alone, which may reduce or eliminate the interference effects of the added segments in the opposite polarity condition (tested in Experiment 2). Figure 4 illustrates these conditions applied to a single object.

Experiment 1: Drawings of constant base polarity

Given that L-vertices in natural images almost always have legs of the same contrast polarity, would their capacity to signal the termination of a surface be dependent on their having the same contrast polarity? Specifically, would L-vertices with legs of the same direction of contrast, when inserted in gaps in the midsections of object contours, be more effective in suppressing smooth continuation through the gaps, thus resulting in greater degradation in object recognition than L-vertices with legs of opposite direction of contrast? Observers viewed briefly presented line drawings of common objects and named them aloud. The contours of the line drawings were of a constant polarity (either all black or all white on a midgray background) and were interrupted with gaps. Black or white line segments were added to both sides of each gap to produce matched L- or T-vertices with legs of the same or different contrast polarity as the object's contours (as illustrated in Figure 4).

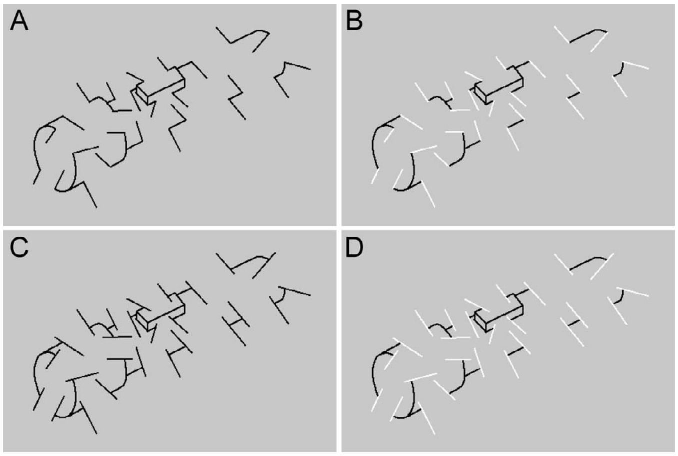

Figure 4. Examples of the stimulus conditions for a black drawing of Experiment 1. Line drawings of common objects were interrupted by gaps, and line segments were added on either side of the gaps to produce (A) L-vertices with legs of the same contrast polarity, (B) L-vertices with legs of different contrast polarity, (C) T-vertices with legs of the same contrast polarity, and (D) T-vertices with legs of different contrast polarity. The variations to produce the conditions for the white line drawings were done in the same manner. Note the greater subjective difficulty in identifying the object in A compared to (B) or (C) but the minimal difference between (B) and (D). Experiment 2 examined the potential role of a global luminance separation mechanism in rendering the identification of (B) easier than (A).

Method

Stimuli

Stimuli were 48 line drawings of common objects and animals rendered as black (0 cd/m^2) or white (61.8 cd/m^2) lines on a midgray (3.46 cd/m^2) background. The objects had high agreement across subjects in their basic-level labeling. All line drawings were modified in Adobe Photoshop and resized such that at a presentation distance of 36 in., the objects subtended between 5° and 10° of visual angle. Approximately one third (35.4% ± 7.0%) of the total contour was then deleted by removing sections of extended contour of variable sizes. The gap size was scaled according to object and part size, and varied between 0.14° and 2.12° of visual angle. Lines of the same contrast were then added to each side of the gap to produce two paired T-vertices. The length of the added segments was generally longer than the gap size, except where such a long segment was unfeasible (such as in crowded sections of the drawing). Importantly, the L-vertices were created by simply removing the shorter leg of the T-vertex; as a consequence, there were more extraneous "noise" pixels in the T-vertex condition as illustrated in Figure 4. Where possible, the two segments associated with one gap were offset in different directions and were of slightly different length. Some of the images (such as an airplane and telephone) had areas composed of close parallel lines (e.g., the edges of the upper and lower surfaces of the airplane wing). In these cases, the added segments were of slightly longer length and bridged both lines—in the L-vertex condition, this segment created an L with one line and a T with the second line.

The remaining conditions were constructed by varying the contrast of either the added segments or the base contour to produce a total of eight conditions: two vertex types (T vs. L) by two base image contrasts (dark vs. light) by two segment contrast polarities (same vs. different; Figure 4).

Subjects

Thirty-two subjects with normal or corrected-to-normal vision between the ages of 17 and 28 (mean age 20 ± 2.1 years, 24 female, two left-handed) participated for credit in psychology courses at the University of Southern California or were paid volunteers. Because participation required the rapid naming of objects, all subjects were native English speakers. None of the subjects had previously seen any of the stimuli, or were aware of the purpose of the manipulations.

Procedure

Subjects viewed the stimuli on a Sony Trinitron Multiscan500PS 19-in. monitor (1280 × 960 pixels, refresh rate 75 Hz; Sony Corp., Tokyo, Japan), controlled by a Macintosh G3 computer (Apple Inc., Cupertino, CA) running MacProbe (Hunt, 1994). Subjects were told that they were participating in an experiment designed to investigate the perception of pictures presented under difficult viewing conditions. All instructions were presented on the computer screen after adaptation to the dark room. The subjects were instructed to name the object as quickly and accurately as possible and to avoid prevocalizations. Twelve practice trials were provided to familiarize the subjects with the procedure, the types of images that they would be seeing, and the sensitivity of the microphone. Images used in the practice trials were not reused in the experiment.

Each subject viewed all 48 objects once, six in each of the eight conditions of vertex type, contrast polarity and base image polarity. A group of eight subjects would be required to have all 48 objects appear in each of the eight stimulus conditions. There were thus four instances of each object in each stimulus condition. The presentation and condition sequences were counterbalanced across subjects such that each object appeared equally often in every condition, and the average serial position of each object and condition was the same (by reversing the stimulus presentation order across subjects). The relatively large number of subjects (32) for a

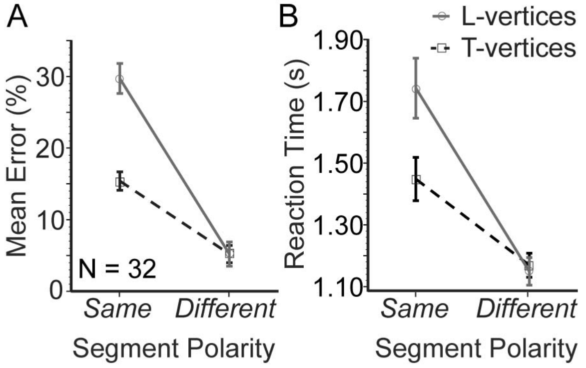

Figure 5. The naming of line drawings of objects had higher error rates (Panel A) and longer RTs (Panel B) when the added contour segments were of the same-contrast polarity as the base drawing and formed L-vertices than when they formed T-vertices. Performance improved and did not differ by vertex type when the added segments were of different-contrast polarity than the base drawing. Error bars are standard errors of the mean (after removal of between-observer mean differences).

reaction time (RT) experiment was necessitated by the constraint that each subject saw each of the 48 objects (in one of its eight conditions) only once. Each of the 48 trials for a given subject began with the subject hitting a key to initiate the presentation sequence. A fixation dot was then displayed in the center of the screen for 667 ms, followed by the image. The image remained on the screen until the subject responded or 10 s had elapsed with no response.

Naming RTs were recorded from the onset of the test image using a voice key and were counted as valid if they fell within a time window between 300 ms and 10 s. Naming errors were recorded by the experimenter, as were any microphone or recording malfunctions (which were excluded from the data analysis). No observations were excluded from error rate calculations. Analysis of RTs was confined to correct trials only. Of these, 11 of 1,323 trials (less than 1%) were excluded for not falling in the valid reaction time window or because of microphone error.

Following each trial, the subject was visually shown his or her reaction time and given verbal feedback from the experimenter as to the name of the object.

Results

The addition of segments of the same contrast polarity to form L-vertices dramatically impaired object naming compared to the addition of different polarity segments forming L-vertices and to T-vertices of either polarity (Figure 5). As no effect of base image color (black vs. white) was found for either error rates or RTs ($F < 1.00$ for both), results were collapsed across this variable. Naming errors for images with same-polarity additions compared to different-polarity additions was much larger when those segments formed L-vertices than when they formed T-vertices (a 24.5% increase for L-vertices but only a 10.2% increase for T-vertices; Figure 5a), leading to a highly significant interaction between vertex type and contrast polarity, $F(1, 31) = 17.59$, $p < 3 \times 10^{-4}$, $\eta^2_p = 0.36$. In fact, there was no difference between L- and T-vertices when the added segments were of different contrast polarity.

The same pattern of results was found for naming RTs (Figure 5b), suggesting that a speed–accuracy tradeoff could not account for the error data. Observers took longer to name images disrupted by same-polarity L-vertices than images disrupted by same-polarity T-vertices although there was no effect of vertex type when the segments were of different contrast polarities, leading to a significant contrast polarity by vertex type interaction, $F(1, 31) = 6.26$, $p < 0.02$, $\eta^2_p = 0.17$. Overall, they took longer to name images with same-polarity segments than images with different-polarity segments, $F(1, 31) = 47.94$, $p < 1 \times 10^{-6}$, $\eta^2_p = 0.61$ and L-vertices than T-vertices, $F(1, 31) = 5.57$, $p < 0.03$, $\eta^2_p = 0.15$. With RTs, as with errors, the greater interference of an L-vertex over a T-vertex, with both segments of the same-contrast polarity, was completely eliminated when the legs of the L-vertex were of different directions of contrast polarity.

An image-based analysis revealed that several of the images used in this experiment generated very long RTs. Given the strong effect on error rates, it is unlikely that the observed RT effects were entirely dependent on these most difficult objects. To confirm this, we performed a control analysis in which we removed images whose average reaction time across all conditions and subjects were identified as outliers (Thompson Tau, critical probability $\alpha = 0.05$; seven of 48 images removed plus two images with no correct trials in one or more conditions) and performed an analysis of variance with images as the random effect. The interaction of vertex type and contrast polarity remained, $F(1, 38) = 7.77$, $p < 0.01$, $\eta^2_p = 0.17$.

Discussion

It is clear that for both error rates and RTs, inserting an L-vertex with legs of the same contrast polarity prevented smooth continuation (i.e., grouping, across the gap in the base contour) resulting in markedly greater difficulty in object identification compared to when the segments of the L-vertices were of different contrast polarity. The implication is that L-vertices with segments of different polarity are ineffective in signaling the termination of a surface. When the added segments were extended past the object contour to form T-vertices, the interference was greatly reduced. This was the case even though the T-vertex segments were composed of *more* pixels than the L-vertex segments. This suggests that the interference was not simply due to the addition of uninformative contour to the image, but was instead a specific consequence of the disruption of contour integration and subsequent surface segmentation. The disruption of contour integration only occurred when the added segments were of the same contrast, and occurred to a much greater degree when the segments formed L-vertices than T-vertices.

Although it would not explain the inferior naming performance for same-polarity line drawings interrupted by L-vertices compared to T-vertices, an alternative explanation for the improved performance in the different polarity condition is that the observers were able to globally ignore all of the opposite-polarity segments. Experiment 2 was designed to test this possibility by reversing the polarity of adjacent contours of the base object (i.e., "candy striping" the contours) thus preventing any potential boost in performance in the different polarity condition through the use of a global luminance separation strategy.

Experiment 2: Drawings of mixed base polarity

In Experiment 1, naming performance for line drawings of objects was impaired when segments interrupting the contours of the drawing were of the same contrast polarity compared to when they were of the opposite polarity, and this effect was greatest when the added segments formed L-vertices with the base image contour. In order to test whether this effect could be attributed to selective attention to high versus low luminance contours as opposed to local inhibition of contour completion by L-vertices of constant contrast, a second experiment was conducted using base images with mixed contrast polarity. These objects, which appeared "candy striped," were then modified in a manner similar to Experiment 1 to include added segments of same or different *local* contrast polarity as the base image contour.

Method

Stimuli

The same 48 line drawings of objects used in Experiment 1 were used in Experiment 2. Unlike Experiment 1, the base line drawings were altered to introduce variations in the polarity of the contrast between the drawing and the gray background, before the addition of gaps and added segments. Gaps were then introduced (in a manner similar to Experiment 1) by removing sections of contour within constant polarity segments (22% ± 5% of the base contour, on average).

In general, the nature of the gaps and added segments in Experiment 2 were the same as those in Experiment 1. As detailed below, there were several slight modifications that were made due to the reversals introduced in the base contour in an effort to prevent local regions of the experimental images from being erroneously grouped on the basis of line luminance alone. The length of the added segments was slightly reduced, the length of the segments used to create T- and L-vertices were more similar, and the added segments of the L-vertices tended to point in the same direction rather than in opposite directions.

Contrast reversals never occurred across a gap, but only on line segments between gaps. The gap size was scaled according to object and part size, and varied between 0.18° and 1.48° of visual angle. Lines of the same *local* contrast were then added to each side of the gap to produce two paired T-vertices. The length of the added segments was approximately the same as the gap size, except where such a long segment was unfeasible (such as in very crowded sections of the drawing). Unlike Experiment 1 where the L-vertices were created by simply removing the shorter end of the T-vertices, in Experiment 2, where possible, the added bit of the L-vertex was approximately equal to the length of the added bit for the matching T-vertex. This would be equivalent to simply sliding the bar of the T-vertex until it became an L-vertex. Some of the images had areas composed of very close parallel lines (as in the airplane wing in Figure 6). In these cases, the added segments were of slightly longer length and bridged across both lines—in the L-vertex condition, this segment created an L with one line and a T with the second line.

As in Experiment 1, the full set of experimental conditions were constructed by varying the contrast of either the added segments or the base contour to produce a total of eight conditions: two vertex types (T vs. L) by two base image versions (complimentary regions of dark and light) by two segment contrast polarities (same vs. different; Figure 6).

Subjects

Fifty-five subjects with normal or corrected-to-normal vision between the ages of 17 and 36 (mean age 20.6 ± 3.5 years, 17 male, one left-handed, one ambidextrous) participated for credit in psychology courses at the University of Southern California or were paid volunteers. A larger number of subjects were required because, again, each subject could see only one

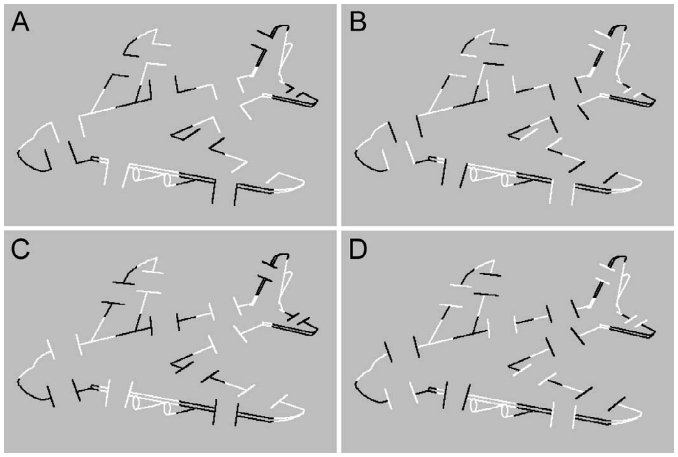

Figure 6. Examples of the mixed-polarity stimulus conditions in Experiment 2. The contrast polarity of contour segments of line drawings was reversed to produce a mixed-polarity (candy striped) base image. Constant polarity sections of contour were then interrupted by gaps, and line segments were added on either side of the gaps to produce (A) L-vertices with legs of the same contrast polarity, (B) L-vertices with legs of different contrast polarity, (C) T-vertices with legs of the same contrast polarity, and (D) T-vertices with legs of different contrast polarity. A set of conditions with the reversed base image polarity was also included.

of the 48 objects once and the candy striping subjectively reduced magnitudes of the perceptual effects.

Procedure

The procedure for Experiment 2 was identical to that of Experiment 1. Ten observations (of 2,460 trials) were excluded from error rate calculations due to microphone errors resulting in premature removal of the image from the screen (less than 0.5%). Analysis of RTs was confined to correct trials only. Of these, 130 of 2,306 trials (5.6%) were excluded for not falling in the valid RT window (between 300 ms and 10 s) or because of microphone error.

Results

Experiment 2 yielded overall lower error rates and shorter RTs than Experiment 1 but the general pattern of results was replicated. As in Experiment 1, observers' naming performance was worse for images whose contours were interrupted with segments of the same local contrast that formed L-vertices (Figure 7). Observers made more errors when the segments were of the same local contrast polarity than when they were of different contrast polarity, $F(1, 54) = 17.7$, $p < 1 \times 10^{-4}$, $\eta^2_p = 0.25$. Post hoc t tests revealed that the error rates for same-contrast L-vertices were reliably greater than different-contrast L-vertices, $t(54) = 3.55$, $p < 0.001$, Cohen's $d = 0.48$, but same-contrast T-vertices were not reliably greater than different-contrast T-vertices, $t(54) = 1.50$, $p = 0.14$, Cohen's $d = 0.20$. Importantly, the reduction in error rates and RTs with L-vertices with different (vs. same) contrast polarity was apparent even with the elimination of a global luminance separation strategy that could have been employed in Experiment 1. Overall, observers made more errors when the segments formed L-vertices than when they formed T-vertices, $F(1, 54) = 13.3$, $p < 6 \times 10^{-4}$, $\eta^2_p = 0.20$, and although it was in the same direction as in Experiment 1, the interaction between contrast polarity and vertex type fell short of significance, $F(1, 54) = 1.91$, $p = 0.17$, $\eta^2_p = 0.034$. The main effects of segment polarity and of vertex type were also significant for RTs, $F(1, 54) = 4.65$, $p < 0.04$, $\eta^2_p = 0.079$, and $F(1, 54) = 4.51$, $p < 0.04$, $\eta^2_p = 0.077$, respectively, suggesting that there was no speed–accuracy tradeoff. The interaction of segment polarity and vertex type for RTs did not reach significance although it was in the same direction as in Experiment 1 which is what would be expected from same polarity L-vertices producing greater difficulty in object naming than different polarity L-vertices.

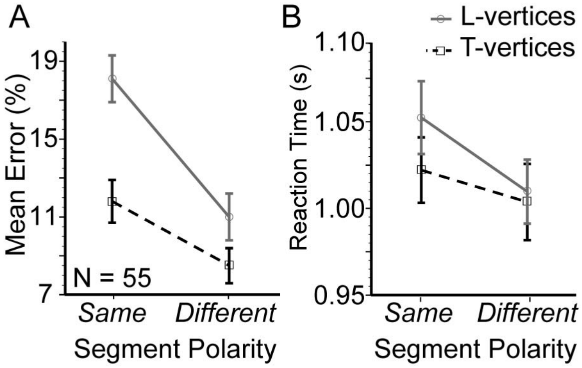

Figure 7. Object naming performance for candy-striped contour deleted images bounded by same- or different-contrast L- or T-vertices. As in Experiment 1, the greatest impairment in performance was with images with L-vertices of the same local contrast. (A) Observers naming line drawings made more errors when the added contour segments formed L-vertices than when they formed T-vertices, and when segments formed vertices with legs of the same local contrast polarity than of different-contrast polarity. (B) RTs showed a similar, though weaker pattern, ruling out any speed–accuracy tradeoff. Error bars are standard errors of the mean (after removal of between-observer mean differences).

An examination of the data for specific images revealed three as having very long RTs (Thompson Tau, critical probability $\alpha = 0.05$). After removing these (plus one image with no correct trials in one or more conditions), the effect of contrast polarity remained $F(1, 43) = 7.38$, $p < 0.01$, $\eta^2_p = 0.15$, though the effect of vertex type was reduced $F(1, 43) = 2.34$, $p = 0.13$, $\eta^2_p = 0.052$ (analysis of variance with images as a random effect).

General discussion

Perceptual grouping of two edge segments into an L-vertex signaling the termination of a surface is contrast dependent. In both experiments, adding line segments to gaps to form L-vertices with legs of the same contrast polarity resulted in greater difficulty in object naming compared to L-vertices with opposite directions of contrast, or T-vertices of either contrast. When the segments of the L-vertices were of different contrast polarities, the L was not effective in suppressing smooth continuation across the gap, resulting in a level of naming performance that was essentially equivalent to the T-vertices of different contrast (although the interaction fell short of significance in Experiment 2).

Given that occluding and occluded contours can be of any contrast relations, why was there a cost of contrast homogeneity with T-vertices? T-vertices, like L-vertices, imposed a cost (though smaller than with the Ls) on recognition performance when the two segments comprising the T were of the same contrast polarity (vs. different contrast polarity). We hypothesize that because the T-vertices always have a short—at times very short—contour that distinguishes a T from an L, additional processing would be required before that short contour is detected which would allow smooth continuation to complete the contour through the T. By this account, some of the T-vertices are initially construed as L-vertices, thus inhibiting smooth continuation and, consequently, object completion/recognition. With additional time (or scrutiny) the short segment is detected—defining now a T—allowing smooth continuation to complete the object's contours and define the objects' parts. This phenomenon can be appreciated by viewing Panel C of Figure 4, which shows a flashlight whose gaps are bridged by Ts. But to determine that they are Ts and not Ls, their short segments have to be detected. Once the T-vertices are perceived, the contours can be completed. Consider the elongated cylinder comprising the main barrel of the flashlight. It takes some time before the Ts can be resolved and smooth continuation can define the entirety of the barrel. The subjective effect is that it takes a bit of time for the parts to be perceived compared to Panel D where color (luminance) separation allows rapid and effortless appreciation of the object's parts. This never occurs in two of the three parts shown in Panel A where the L-vertices, with segments of the same polarity, stop smooth continuation. (The brick constituting the switch did not have sufficient deletion to produce noticeable difficulty in its recognition.)

The interaction between polarity (same vs. different) and vertex type (T vs. L) which was significant in Experiment 1 fell short of significance in Experiment 2 although the effects were qualitatively the same (i.e., relatively greater costs for L- than T-vertices with same polarity segments, versus different polarity segments). A possible explanation for this is suggested from the somewhat surprising superior performance—lower error rates and RTs—in Experiment 1 compared to Experiment 2, despite the candy striping that would have been otherwise expected to exact a cost on performance. The variation in the manner in which the T- and L-vertices were produced in Experiment 2—in which the segments added to the gaps for both kinds of vertices were always parallel, in the same direction, and of identical length for both T- and L-vertices—appeared to define a perceptual group somewhat independent of the object's shape, which could then be simply ignored through selective inattention. This effect can be appreciated by a comparison of the objects in Panel A in Figures 4 (Experiment 1) and 6 (Experiment 2). Both illustrate same direction of contrast L-vertices, but in Figure 6 the added segments at each gap appear to define a perceptual group that is independent of the object's contours. In Figure 4, it is less clear which contours of the L-vertices are part of the object and which are added noise.

In Experiment 1 a global luminance separation function, perhaps implemented by attentional selection to the luminance of the object's contours, could have reduced the interference effects of the added segments in the different-contrast condition, facilitating performance relative to objects with same-contrast T-vertices. However, such a separation function would not have accounted for the higher error rates and RTs of same-contrast polarity L-vertices relative to same-contrast T-vertices. Experiment 2, with candy-striped base objects that alternated between black and white along their extended contours, was designed to eliminate the effectiveness of such a selection function. But Experiment 2 showed essentially the same pattern of results as Experiment 1, with a strong mean effect of contrast, suggesting that the observed effect of same- versus different- polarity segments in Experiment 1 was due to local interactions, not a global process. Any global attentional selection function would appear to have had only a minimal effect—that is, the 2% (nonsignificant) increase in error rates in Experiment 2 of L- over T-vertices when they are of different contrast polarity.

The observed contrast dependence of the detection of L-vertices is particularly important when contrasted with the invariance to contrast polarity that is observed for smooth continuation along extended contours (Gilchrist, Humphreys, Riddoch, & Neumann, 1997), another key feature of midlevel vision. The ability of humans to detect extended contours, even when those contours are extensively interrupted by gaps and are composed of segments of alternating contrast polarity (e.g., Field, Hayes, & Hess, 1993; Sivasubramaniam, 1998) is likely supported by (a) the distribution of long-range horizontal connections observed between columns of V1 (see Gilbert, 1998; Kapadia, Ito, Gilbert, & Westheimer, 1995) and (b) the contrast invariance of complex cells (Ringach, 2002; Sary, Vogels, & Orban, 1993). These neural properties, in turn, are likely a consequence of the statistical likelihood of changes in contrast polarity along extended surface boundaries in natural images (Elder & Goldberg, 2002; Geisler & Perry, 2009).

This interpretation is also consistent with existing evidence that perceptual closure of two-dimensional shapes is impaired when changes in contrast polarity occur right at corners, but not when those changes occur along extended contours (Elder & Zucker, 1993; Spehar, 2002). Furthermore, our findings suggest that local vertex structure is critical, and interacts with contrast polarity.

The results of the present experiments represent strong evidence for the incorporation of natural image statistics into the representation of object shape. Neurons tuned to sharp convexities such as L-vertices have been found in macaque V4 (Pasupathy & Connor, 1999), form a population code in V4 for shape contour (Pasupathy & Connor, 2002), and likely form a basis for a parts-based encoding of shape in macaque IT that explicitly encodes relations (Brincat & Connor, 2004). Our findings suggest that such tuning in macaque V4 would be reduced, if not eliminated, if the two legs of the vertices were of different contrast polarity.

Keywords: L-vertices, natural image statistics, contrast polarity, direction of contrast

Acknowledgments

Supported by Office of Naval Research 202-98-K-1089 and Army Research Office DAAH04-94-G-0065 to IB. The authors would like to thank Robert Shapley for feedback on an earlier version of this manuscript.

Commercial relationships: none.

Corresponding author: Irving Biederman.
Email: bieder@usc.edu.
Address: Neuroscience Program and Department of Psychology, University of Southern California, Los Angeles, CA, USA.

References

Blickle, T. W. (1989). *Recognition of contour deleted images* (Unpublished doctoral dissertation). State University of New York, Buffalo, NY.

Brincat, S. L., & Connor, C. E. (2004). Underlying principles of visual shape selectivity in posterior inferotemporal cortex. *Nature Neuroscience*, *7*(8), 880–886, doi:10.1038/nn1278.

Bregman, A. S. (1981). Asking the "what for" question in auditory perception. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization* (pp. 106–107). Hillsdale, NJ: Erlbaum.

Donnelly, N., Humphreys, G. W., & Riddoch, M. J. (1991). Parallel computation of primitive shape descriptions. *Journal of Experimental Psychology: Human Perception and Performance*, *17*(2), 561–570.

Elder, J. H., & Goldberg, R. M. (2002). Ecological statistics of Gestalt laws for the perceptual organization of contours. *Journal of Vision*, *2*(5):4, 324–353, doi:10.1167/2.5.4. [PubMed] [Article]

Elder, J. H., & Zucker, S. (1993). The effect of contour closure on the rapid discrimination of 2-D shapes. *Vision Research*, *33*, 981–991.

Field, D. J. (1987). Relations between the statistics of natural images and the response properties of cortical cells. *Journal of the Optical Society of America A*, *4*(12), 2379–2394, doi:10.1364/JOSAA.4.002379.

Field, D. J., Hayes, A., & Hess, R. F. (1993). Contour integration by the human visual-system: Evidence for a local "association field." *Vision Research*, *33*(2), 173–193.

Geisler, W. S., & Perry, J. S. (2009). Contour statistics in natural images: Grouping across occlusions. *Visual Neuroscience*, *26*, 109–121.

Geisler, W. S., Perry, J. S., Super, B. J., & Gallogly, D. P. (2001). Edge co-occurrence in natural images predicts contour grouping performance. *Vision Research*, *41*(6), 711–724.

Gilbert, C. D. (1998). Adult cortical dynamics. *Physiological Reviews*, *78*(2), 467–485.

Gilchrist, I. D., Humphreys, G. W., Riddoch, M. J., & Neumann, H. (1997). Luminance and edge information in grouping: a study using visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *23*(2), 464–480.

Guzman, A. (1968). Computer recognition of three-dimensional objects in a visual scene. *Proceedings AFIPS '68 (Fall, part I)*. 291–304. New York: ACM.

Hummel, J. E., & Biederman, I. (1992). Dynamic binding in a neural network for shape recognition. *Psychological Review*, *99*(3), 480–517.

Hunt, S. M. J. (1994). MacProbe: A Macintosh-based experimenter's workstation for the cognitive sciences. *Behavior Research Methods, Instruments, & Computers*, *26*(3), 345–361, doi:10.3758/BF03204643.

Kapadia, M. K., Ito, M., Gilbert, C. D., & Westheimer, G. (1995). Improvement in visual sensitivity by changes in local context: Parallel studies in human observers and in V1 of alert monkeys. *Neuron*, *15*(4), 843–856.

Leeper, R. (1935). A study of a neglected portion of the field of learning: The development of sensory organization. *Journal of Genetic Psychology*, *46*, 41–75.

Olshausen, B. A. (1996). Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*, *381*(6583), 607–609.

Pasupathy, A., & Connor, C. E. (1999). Responses to contour features in macaque area V4. *Journal of Neurophysiology*, *82*(5), 2490–2502.

Pasupathy, A., & Connor, C. E. (2002). Population coding of shape in area V4. *Nature Neuroscience*, *5*(12), 1332–1338, doi:10.1038/nn972.

Ramachandra, C. A., & Mel, B. W. (2013). Computing local edge probability in natural scenes from a population of oriented simple cells. *Journal of Vision*, *13*(14):19, 1–22, doi:10.1167/13.14.19. [PubMed] [Article]

Ringach, D. L. (2002). Spatial structure and symmetry of simple-cell receptive fields in macaque primary visual cortex. *Journal of Neurophysiology*, *88*, 455–463.

Sary, G., Vogels, R., & Orban, G. A. (1993). Cue-invariant shape selectivity of macaque inferior temporal neurons. *Science*, *260*(5110), 995–997.

Schira, M. M., & Spehar, B. (2011). Differential effect of contrast polarity reversals in closed squares and open L-junctions. *Frontiers in Psychology*, *2*(47), 1–11. doi:10.3389/fpsyg.2011.00047

Sivasubramaniam, S. (1998). *Shape representations: Coding of irregularities and inconsistencies* (Dissertation). University of Southern California, Los Angeles, CA.

Spehar, B. (2002). The role of contrast polarity in perceptual closure. *Vision Research*, *42*(3), 343–350.

Waltz, D. (1975). Understanding line drawings of scenes with shadows. In P. H. Winston (Ed.), *The psychology of computer vision* (pp. 19–92). New York: McGraw-Hill.