Seeing grating-textured surface begins at the border

Yong R. Su
Zijiang J. He
Teng Leng Ooi

Two experiments were conducted to reveal that the human visual system represents grating texture surface using a border-to-interior strategy. This strategy dictates that the visual system first registers the surface boundary contour and then sequentially spreads texture from the border to the interior of the image. Our experiments measured the perceived grating texture surface at various stimulus durations after the onset of a grating texture image. We found that the grating texture is initially seen near the boundary contours, with eventual spreading inward to the center of the image. To quantify the observation, the extent of the texture spreading from the boundary contour is measured as a function of the stimulus duration (30–500 ms). This allows us to analyze the texture spreading in retinal and cortical distances, based on human fMRI studies of the cortical magnification factor in cortical areas V1–V4, and to derive the spreading speed. We found that the spreading speed is constant when scaled according to the cortical distance. Similar findings are obtained no matter whether the grating texture image is presented monocularly or dichoptically, suggesting the generality of the border-to-interior strategy for representing surfaces.

Keywords: boundary contour, cortical travel speed, filling-in, surface representation, surface spreading-in, texture

Introduction

We live and interact in an environment where most objects in the natural scenes are defined by surfaces and are themselves often supported by larger background surfaces. This suggests that a significant amount of visual processing must be devoted to representing surfaces (Gibson, 1950; Nakayama, He, & Shimojo, 1995; von der Heydt, 2003). Although we have learned a great deal about how retinal images are coded locally by the early visual processors to obtain simple feature information, we know less about how the local simple feature information is integrated into a global surface representation. For example, for texture regions comprising of lines with different orientations, V1 neurons whose preferred orientation corresponds to the line orientation of either texture region will respond strongly to those lines. There are also some V1 neurons whose activities increase at the border between the texture regions, which reveals an early form of texture boundary contour representation (e.g., Lamme, 1995; Nothdurft, 1992; Nothdurft, Gallant, & van Essen, 2000). It is, however, not entirely clear how V1 and the extrastriate cortices subsequently process the texture information to ultimately construct the global surface representation that supports our perception of texture surfaces. There are reasons to believe that the surface integration process is not an exclusively data-driven operation but one that relies on a number of intrinsic assumptions. These assumptions, or perceptual rules, are largely derived from the regularities embedded in our ecology (niche; Albert & Hoffman, 2000; Geisler, 2008; Gibson, 1950; Nakayama & Shimojo, 1992; Ooi, Wu, & He, 2001; Purves, Lotto, Williams, Nundy, & Yang, 2001; Schwarzkopf & Kourtzi, 2008). Thus, an important approach in research of visual surface representation is discovering the rules that are shaped by the statistical properties of natural images, projection geometry, and visual-motor interactions.

The study reported in this paper tested and extended the proposal that the visual system relies on a border-to-interior strategy to represent texture surfaces (Grossberg & Mingolla, 1985; Paradiso & Nakayama, 1991). With this strategy, the surface boundary contour (BC) of the texture image is first coded, followed by the texture region adjacent to the BC. Presumably, the surface texture adjacent to the BC is then sequentially integrated with the local texture signals of the interior image. The time required to integrate the local texture signals for global surface representation...
depends on the complexity of the texture. In this way, the particular properties of the texture determine the overall speed of surface representation.

It should be noted that the process of representing a texture surface is probably different from the filling-in process for representing a uniform color/achromatic surface. Unlike the representation of a texture surface, the filling-in operation for a uniform surface is mainly based on the color/brightness contrast information at the BC and does not include the local integration of the uniform region enclosed by the BC (von der Heydt, Friedman, & Zhou, 2003). This is because the primary visual cortex receives only either very weak or no signals from the LGN neurons regarding the color/light intensity of the interior uniform region.

An argument for adopting the border-to-interior strategy is related to the fact that surface BCs are crucial cues for surface separation, and they are represented in the early visual cortices (e.g., Bakin, Nakayama, & Gilbert, 2000; Grossberg & Mingolla, 1985; Lamme, 1995; Nakayama et al., 1995; Nothdurft et al., 2000; von der Heydt, 2003; Zhou, Friedman, & von der Heydt, 2000). Furthermore, since rich depth information (T-junctions, etc.) are found in the vicinity of surface BCs, representing 3-D surfaces from 2-D images beginning at the BCs can help overcome some depth ambiguity inherent in the interior surface regions with similar texture information. Thus, we hypothesize that for the display in Figure 1, the BC of the disk (Figure 1a) will be represented first, followed by the spreading-in of the vertical grating texture from the BC (Figure 1b). This contrasts with a pixel-by-pixel strategy that locally integrates individual image patches with the same surface feature (vertical grating) to form a global surface representation (vertical grating disk).

The current study is also motivated by the early theoretical and empirical studies of perceptual fading and filling-in of uniformly colored surfaces and scotoma (e.g., Friedman, Zhou, & von der Heyst, 2003; Gerrits & Timmerman, 1969; Gerrits & Vendrik, 1970; Huang & Paradiso, 2008; Komatsu, 2006; Komatsu, Kinoshita, & Murakami, 2000; Krauskopf, 1963; Meng, Feneyhough, & Tong, 2007; Riggs, Ratliff, Cornsweet, & Cornsweet, 1953; Sasaki & Watanabe, 2004; Troxler, 1804; von der Heydt et al., 2003; Walls, 1954; Yarbus, 1967). For example, when the BC of a uniformly colored figure is stabilized on the retina, the contrast signals at its border/edge become weak. This causes the color of the interior of the figure to fade away, to be replaced (“filling-in”) by the color of the larger surface external to the figure’s border/edge. It is thus hypothesized that the visual system relies on the color contrast signals at the border/edge to fill in the color of the interior region (Cohen & Grossberg, 1984; Gerrits & Vendrik, 1970). Separately, studies have also demonstrated perceptual fading and filling-in of texture patterns for stabilized images and scotoma (De Weerd, Desimone, & Ungerleider, 1998; De Weerd, Gattass, Desimone, & Ungerleider, 1995; Ramachandran & Gregory, 1991; Spillmann & Kurtenbach, 1992; Watanabe & Cavanagh, 1991).

There are also empirical studies other than with perceptual fading and filling-in performances that support the border-to-interior strategy (e.g., Caputo, 1998; Davey, Maddess, & Srinivasan, 1998; De Valois, Webster, De Valois, & Lingelbach, 1986; Motoyoshi, 1999; Paradiso & Hahn, 1996; Paradiso & Nakayama, 1991; Rossi & Paradiso, 1996). More direct and straightforward support for the border-to-interior strategy comes from studies that measured how the perceived image changes over time after the stimulus onset using a masking paradigm (Caputo, 1998; Motoyoshi, 1999; Paradiso & Nakayama, 1991; Stoper & Mansfield, 1978). For example, Paradiso and Nakayama (1991), in one of their experiments, presented the observers with a homogeneous white disk (target) on a black background for 16 ms, followed after a variable ISI, with a small white ring (annular mask) on a black background for 16 ms. Their observers perceived that the brightness of the central area concentric with the annulus changed from black to white as the ISI increased from 0 to 126 ms. According to the border-to-interior strategy, the visual system first codes the BC of the disk after the onset of the white disk, and then spreads the edge contrast signal (brightness) inward until it meets another edge. Thus, in Paradiso and Nakayama’s experiment, when the inter-stimulus interval (ISI) is sufficiently short, the white ring stops the brightness-spreading wave resulting in the central region being seen as black. This explanation predicts that when the spatial separation between the white disk and masking ring is increased (i.e., with larger diameter white disk), it will take a longer time for the spreading wave to reach the location of the white masking ring. This is because the spreading wave, which originates from the BC of the white disk, now has to travel a longer distance before reaching the masking ring. Therefore, a longer ISI is expected to produce the same quality of brightness perception. This was exactly what Paradiso and Nakayama (1991) found. Using a similar masking paradigm, Caputo (1998) and Motoyoshi (1999) presented a texture surface in one eye and an annular or square frame mask in the other eye to reveal the characteristics of texture spreading.

The current study provides further evidence that the visual system follows the border-to-interior strategy for representing textured surfaces using a new psychophysical method that measures the texture spreading during surface representation. This method allows us to directly measure the perceived position of the leading front of the spreading texture. Thus, when measured as a function of time after the stimulus onset we can derive the speed of the spreading wave. Notably, our method of measuring texture spreading, where the judgment depends on the leading front of the spreading wave over time, differs from the masking method (Caputo, 1998; Motoyoshi, 1999; Paradiso & Nakayama, 1991). The masking method is more indirect because it derives the spatiotemporal characteristics of a hypothetical
Figure 1. Stimuli and percepts. (a) A grating disk surrounded by orthogonal grating. (b) The border-to-interior strategy predicts the representation of the grating disk spreads inward from the BC. (c) Dichoptic stimulus with MBC. When the half-images are fused, one perceives a stable vertical grating disk in front of the horizontal grating background. (d, e) Dichoptic stimuli with BBC. For each stimulus, free fusion leads to frequent perceptual alternation between the vertical and horizontal gratings. Time-sequence illustrations of the representation of the MBC grating disk (f) for the MBC stimulus and (g) for the BBC stimuli.
spreading wave based on the perceived brightness within a masking ring. Specifically, its spatiotemporal characteristics are derived from manipulations of the ISI between the target stimulus and the masking ring and the BC distance between the masking ring and the target stimulus. While the outcomes of these studies using the masking method are consistent with the hypothesis that the brightness of a target stimulus spreads inward from the BC, they can also be explained by an alternative hypothesis that the brightness of the entire surface of the target stimulus is represented simultaneously. According to the latter hypothesis, the spatiotemporal characteristics of the perceived brightness inside the masking ring reflect the contour interaction between the masking ring and the target stimulus, which also depends on the BC distance and the ISI (Cornelissen, Wade, Vladusich, Dougherty, & Wandell, 2006; von der Heydt et al., 2003).

Our investigation begins with the phenomenological observations of the textured surfaces in Figures 1c–1e, which suggest that the visual system employs the border-to-interior strategy. One can fuse the dichoptic stimuli in Figures 1c–1e to qualitatively verify the important role of the BC in surface representation. With Figure 1c, one perceives a vertical grating disk floating above the surrounding horizontal grating (Ooi & He, 2006; Su, He, & Ooi, 2009). This demonstrates that the visual system preferentially selects an image (vertical grating) with a BC over a competing image (horizontal grating) without the BC for perception. Such a significant role of the BC cannot be explained by a strategy based on a pixel-by-pixel representation, which predicts that the vertical and horizontal gratings in the two eyes have equal chance to be represented. Therefore, it is not surprising that when the half-images of Figures 1d and 1e carry BCS of similar strengths both the orthogonal gratings associated with the BCS are represented, resulting in perceptual competition between the two gratings.

In a related study (Su, He, & Ooi, 2006), we showed the advantage of the BC by comparing the stimulus duration required to achieve global dominance with the MBC (Figure 1c) and BBC (Figure 1d or 1e) stimuli. We presented each stimulus at various durations (30–150 ms) and asked the observers to report their percepts in three categories: global dominance of the grating disk, piecemeal,
or plaid. We found that global dominance is perceived as early as 30 ms for the MBC stimulus, whereas it is hardly perceived even at 150 ms for the BBC stimuli. The latter finding is consistent with the report by Wolfe (1983) who tested a BBC stimulus with similar procedure. On the other hand, our finding with the MBC stimulus clearly demonstrates that the visual system preferentially selects an image (vertical grating) with a BC over a competing image (horizontal grating) without the BC for perception. This paper takes advantage of the uniqueness of the MBC stimulus to study the early dynamics of texture spreading.

**Experiment 1: Measuring the speed of texture spreading in a dichoptic stimulus**

We sought evidence that the representation of the interior surface spreads inward from the MBC, as shown schematically in Figure 1f. (This differs from the dichoptic stimulus with binocular BCs in Figure 1g.) We created a dichoptic stimulus with a rectangular monocular boundary contour (MBC) containing horizontal grating (Figure 2a). To restrict spreading along the lateral direction, the upper and lower edges of the rectangle were blurred. The stimulus was presented at various durations (30–500 ms). Refer to Figure 2b and Supplementary Movie 1 for the predicted (simulated) percepts as a function of time. It is predicted that at the shortest stimulus duration, a plaid, i.e., mixture of horizontal and vertical gratings, is seen (Wolfe, 1983). Next, the left and right edges of the rectangle are filled with horizontal grating while much of the central region is filled with plaid. We refer to the length of the plaid region as the “gap size.” With longer stimulus duration, as more horizontal grating texture spreads inward from the left and right BCs, the region filled with plaid (gap size) becomes smaller.

As mentioned above, we presented the dichoptic stimulus in Figure 2a at various durations (30–500 ms) during the experiment. For each trial, observers reported the perceived plaid region (gap size, G) using a proportional scale from 0 to 6. From G, we obtained the average length of the perceived spread of the horizontal grating texture from one side of the rectangle, \( S_{\text{ret}} = (L - G)/2 \), where \( L \) is the total MBC rectangle length and \( S_{\text{ret}} \) denotes the distance spread in retinal domain. We tested MBC rectangles of three different lengths (\( L = 1.50^\circ, 2.00^\circ, 2.67^\circ \)).

**Methods**

**Apparatus**

Stimuli generated with MATLAB and Psychophysics Toolbox (Brainard, 1997; Pelli, 1997) on a Macintosh were presented on a flat-screen CRT (1280 × 1024 pixels at 100 Hz). A mirror haploscope system attached to a chin-and-head rest aided fusion (viewing distance = 75 cm).

**Observers**

One author participated in all experiments. Six naive observers with informed consent participated (three in Experiment 1 and three in Experiment 2). All observers had normal or corrected-to-normal visual acuity, stereopsis (<20 sec arc), and no significant sensory eye dominance (Ooi & He, 2001).

**Stimuli and procedures**

The dichoptic MBC stimulus (4.14° × 4.14°) had 4-cpd sinusoidal grating (40.3 cd/m²; 29.1% contrast) surrounded by black (2.77 cd/m²) and white (153 cd/m²) square-framed fusion lock (5.23° × 5.23°, width = 0.23°; Figure 3a). One half-image had vertical grating and the other the same vertical grating and an additional rectangular region of horizontal grating (MBC target). The upper and lower horizontal boundaries of the MBC target were blurred with a Gaussian kernel whose full width at half-maximum (FWHM) was 0.6°. The MBC target length was 1.50°, 2.00°, or 2.67°.

The experiment comprised 18 blocks (2 eyes × 3 MBC lengths/conditions × 3 repeats). Each block had 8 warm-up and 50 test trials (5 durations × 10 repeats). A trial began with the observer aligning his/her eyes on a 0.91° × 0.91° nonius fixation target (153 cd/m²). Then, he/she pressed the spacebar to remove the nonius fixation target. Fifty milliseconds later, the MBC stimulus was presented for 30, 60, 120, 250, or 500 ms, followed by a blank field. The observer indicated the perceived gap size within the MBC rectangle using the numbered keys (0–6). This psychological paradigm allows us to directly measure the actual surface spread.

**Prior training for Experiments 1 and 2**

Observers practiced rating simulated gap sizes (Figure 2b) before the experiments to stabilize their gap size rating criterion. They were shown seven simulated gaps paired with 0–6 digits and scaled according to the proportion of the gap width relative to the overall horizontal length of the MBC rectangle. At the extremes, “0” indicates seeing horizontal grating and “6” seeing a mixture of vertical and horizontal gratings (plaid) over the entire length of the MBC stimulus, while numbers 1–5 represent various gap sizes.

Once the observers were familiar with the rating scale, they rated the simulated gap sizes with the accompanying numeric values removed. To reinforce learning, feedback was given after each trial. A block of trials would end after the observer scored eight consecutive correct responses.
For Experiment 1, the training trials were run for each of the three MBC rectangle lengths used. In this and Experiment 2’s training, the viewing duration of the simulated gap stimuli was unrestricted. We implemented two other training paradigms on the naive observers before Experiment 2. First, we randomly timed the stimulus duration of the simulated gap between 30 and 500 ms. In this way, a gap size of “0” could be presented for 30 ms and a gap size of “6” could be presented for 500 ms, etc. This ensured that the observers did not learn to associate smaller gap sizes with longer stimulus durations. The clarity of the simulated gap stimuli was also varied to alter the difficulty of the task (achieved by manipulating the contrast and horizontal–vertical mix ratio of the texture pattern of the gap). Within every single training block, each combination of gap size and stimulus duration was tested ten times. The trials within a block were fully randomized so that no more than three consecutive trials were the same. The binocular and monocular viewing conditions were conducted in separate blocks. The second training paradigm had “catch trials,” in which the test stimuli were intermingled with the simulated gap stimuli. As in the above, the stimulus duration of the simulated gap stimulus was randomized (30–500 ms). The ratio between the number of catch trials and test trials was fixed at 1:3. Therefore, in each block, 4 catch trials were employed for each level of stimulus duration, while 12 test trials with the same stimulus duration were tested. The performance of the observers was evaluated by both their average simulated gap size rating errors and percentages of correct answers in the catch trials. Formal testing (Experiment 2) only began when the observer’s performance stabilized at a reasonably good level (e.g., average errors across all levels of presentation duration should be less than 0.1° after conversion to the perceived spread in visual angle).

**Results and discussion**

Figure 2c shows that, for each L, \( S_{\text{ret}} \) increases with the stimulus duration [\( F(1.104, 3.311) = 13.54, p = 0.029; \) Greenhouse–Geisser correction] and asymptotes around 250 ms. This result provides direct evidence that surface representation spreads inward from the BCs. Assuming the texture spreading activities involve cortical area V1 and/or beyond, we scaled the average lengths in Figure 2c according to the cortical distances in V1 based...
on the cortical magnification factor obtained from a human fMRI study (Figure 2d) by Engel, Glover, and Wandell (1997). Refer to Appendix A for details of the data analysis and fitting. Remarkably, the data from all three MBC conditions with different horizontal rectangular lengths follow the same linear trajectory ($R^2 = 0.937$), suggesting a constant spreading speed (28.7 cm/s) of surface representation within V1. We also applied cortical magnification factors of V1, V2, V3v, and V4v from a different fMRI study by Sereno et al. (1995) and found that the data could be fitted with linear functions, suggesting constant spreading speeds in the visual cortex (V1: 29.5 cm/s, $R^2 = 0.914$; V2: 48.0 cm/s, $R^2 = 0.927$; V3v: 21.2 cm/s, $R^2 = 0.903$; V4v: 18.6 cm/s, $R^2 = 0.891$). Figure 3 depicts the transformed data according to the V1 and V2 cortical magnification factors.

We further compared our derived speed with those from two other psychophysical studies, albeit with different paradigms and parameters. We first examined the finding by Paradiso and Nakayama (1991) who presented the target (uniform disk) and mask (ring) dichoptically. They estimated the speed of brightness filling-in activities in V1 in the range of 15–40 cm/s. This estimation is close to what we observed above. Second, we applied a model of 3-D depth spreading (Nishina, Yazdanbakhsh, Watanabe, & Kawato, 2007) to fit the first 4 data points of our results (before the data asymptote). The estimated speed based on their model is ~30 deg/s, which is slower than the depth propagation of the illusory (neon color) surface (~95–117 deg/s) studied by Nishina et al. The slower speed we found with our data is likely due to the task and stimulation difference between the two studies. Whereas we measured texture surface representation, Nishina et al. measured spreading of 3-D homogeneous surface (neon color).

There are no directly comparable physiological studies for us to verify our estimate of the spreading speeds from V1 to V4. However, we can make an indirect comparison based on an optical imaging recording study by Grinvald, Lieke, Frostig, and Hildesheim (1994) in the monkey. They observed the speeds of spreading activities that originated from the stimulus location to be 0.10–0.25 m/s for area V1. These speeds are just slightly slower than our estimates for V1 (from our human observers). Likely, with optical imaging or single unit recording, future research can verify our speculation that surface representation activities travel in the visual cortex (V1 and extrastriate cortices) at a constant speed. Further, having a constant speed suggests the homogeneity of the cortical neural network in constructing the representation of surfaces.

The average results in Figure 2c show that more than 90% of the MBC grating surface area is represented by 150 ms (stimulus duration). This is comparable with the study we mentioned in the Introduction section (Su et al., 2006) where we measured the time to achieve global dominance in MBC and BBC stimuli. In that study, we found that with a stimulus size of 1.9 deg and at a stimulus duration of 150 ms, the percentage of seeing global dominance was about 80% with the MBC stimulus but less than 10% with the BBC rivalry stimulus.

Finally, we consider another important factor that could influence our observation. It has been shown that the speed of visual information processing along the early visual pathway increases with retinal eccentricity (e.g., Carrasco, McElree, Denisova, & Giordano, 2003; De Valois & De Valois, 1988). This “eccentricity factor” could arguably have contributed to the observed spreading effect in Figure 2c, because the middle section of the horizontal MBC rectangle was imaged on the fovea while its left and right sides were imaged on the peripheral retina. However, such an eccentricity factor could not be the sole contributor. If it were, we would have found that the (measured) perceived gap sizes were the same regardless of the different lengths of the MBC rectangle. Refuting this, our data show a significant effect of the length of the MBC rectangle on the perceived gap size [main effect of length: $F(1.061, 19.103) = 32.302, p < 0.001$; interaction effect between the presentation duration and length: $F(1.061, 19.103) = 12.961, p = 0.002$; 2-way ANOVA with repeated measures; Greenhouse–Geisser correction]. The two gray square symbols in Figure 2c show sample predictions of the eccentricity factor. (Note that the graph plots the perceived length of the horizontal grating from one side, which is derived from the perceived gap size.) To make the prediction, we used the measured perceived gap size for the MBC rectangle of 1.5-deg length and calculated the hypothetical perceived horizontal grating length per side [(MBC length − gap size)/2] for the MBC rectangle of 2.67-deg length. The two gray squares would overlap with the corresponding actual data (blue) if they had the same perceived gap sizes. However, they do not overlap, thus ruling out the eccentricity factor as the sole factor causing the spreading-in results in Figure 2c.

**Experiment 2: Measuring the speed of texture spreading in a monocular stimulus**

Up until now, our findings support the border-to-interior strategy for texture surface representation, which is consistent with the conclusion by Caputo (1998) and Motoyoshi (1999) based on the dichoptic stimuli (Caputo used a monocular target and binocular mask while Motoyoshi presented the target and mask in different eyes). However, since dichoptic stimuli trigger binocular competition (Ooi & He, 2006; Su et al., 2009), it is possible that our conclusion and those of Caputo (1998) and Motoyoshi (1999) are limited to binocular surface representation that involves binocular inhibitory interaction. To generalize our conclusion, we next tested with a monocular stimulus that
only carries the right MBC half-image of Figure 2a while the other half-image is a homogeneous gray field. We first, following the same test protocol of Experiment 1, only observed a moderate trend of surface spreading for representing the rectangular region with horizontal grating (triangles, mono condition, Figure 4a). However, this seemingly rapid surface representation cannot be easily revealed with our current method (we address this issue further in the General discussion section). We thus modified the test protocol by presenting a homogeneous vertical grating pedestal for 200 ms before the MBC test stimulus (horizontal grating rectangle) is presented. Presumably, at the initial stage of processing, the iconic memory of the prior vertical grating pedestal fuses with the signals of the horizontal grating to generate a plaid pattern. This then retards the spreading wave for representing the MBC grating rectangle.

Methods

Stimuli and procedures

There were four conditions: Mono, Dichoptic, Mono-pre, and Dichoptic-pre. The suffix “pre” indicates trials where a homogeneous vertical pedestal grating was presented for 200 ms before the MBC stimulus was presented. The basic stimulus parameters were the same as in Experiment 1, except the contrast was 40% and the MBC rectangle target length was 2.67°. The stimulus was displayed either dichoptically (similar to Figure 2a) or monocularly. With the monocular presentation, only the half-image containing the MBC target was presented while the other eye viewed a homogenous gray field. A 0.4° × 0.4° green fixation cross (6.6 cd/m², CIE: 0.281, 0.484) was presented either to both eyes (dichoptic conditions) or to the test eye (monocular conditions). The MBC stimulus duration was 30, 50, 100, 150, 200, 250, or 500 ms. The dichoptic conditions were retested here for the purpose of verification and comparison.

The four conditions were tested in different blocks. Each test block consisted of 4 warm-up trials and 70 randomized test trials (7 durations × 10 repeats). The test procedure was the same as in Experiment 1 except for the prior presentation of the vertical pedestal grating in the “pre” conditions.

Results and discussion

As predicted for the mono-pre condition, we found that the representation of the horizontal grating (rectangle) takes a longer time and exhibit the characteristic surface spreading from the BCs (Figure 4a, pink circles). This indicates that the border-to-interior strategy is implemented for monocular surfaces. For comparison, we also tested the same observers using the dichoptic MBC stimulus (Figure 2a) with a preceding binocular vertical grating pedestal, and without a preceding vertical grating pedestal (the latter is essentially the same as in Experiment 1). The average results (dichoptic and dichoptic-pre in Figure 4a) show a similar trend as Experiment 1 (Figure 2c). We then converted the data in Figure 4a to V1 cortical distance (Figure 4b) to estimate the spreading speeds within V1 based on the human fMRI study by Engel et al. (1997). We found that, with the preceding vertical grating pedestal, the spreading speed for the monocular stimulus is 53.9 cm/s ($R^2 = 0.850$), which is faster than the speed for the dichoptic-pre stimulus (35.2 cm/s, $R^2 = 0.891$). The speed of the dichoptic stimulus, 32.1 cm/s ($R^2 = 0.810$), is quite similar to that of the dichoptic-pre stimulus. We attribute the slower speeds with the dichoptic and dichoptic-pre stimuli to the local interocular competition between the dichoptic orthogonal gratings. This is because the orthogonal gratings (conflicting local features) at corresponding retinal locations will initiate
interocular inhibition. At each local area, the dichoptic gratings have equal chance to compete for representation. As such, along its inward path, the horizontal grating spreading wave will encounter a locally dominant image representation, which can be either a horizontal or vertical grating. If it is a horizontal grating, the texture integration will be quick, whereas if it is vertical, it will be slower as time is needed for that local patch of retinal area to revert to horizontal grating dominance. The latter operation requires additional processing and thus extra time, which does not occur in the mono and mono-pre stimulus conditions.

Figure 5 provides the same analysis based on the fMRI study by Sereno et al. (1995). The colored linear regression lines in (b) and (d) that pass through the origin are derived from the data of each of the three stimulus conditions (V1: Mono-pre, $y = 0.495x$, $R^2 = 0.886$; Dichoptic, $y = 0.362x$, $R^2 = 0.828$; Dichoptic-pre: $y = 0.404x$, $R^2 = 0.894$; V2: Mono-pre, $y = 0.783x$, $R^2 = 0.865$; Dichoptic, $y = 0.542x$, $R^2 = 0.798$; Dichoptic-pre: $y = 0.611x$, $R^2 = 0.886$).

We now address the possible factors causing the faster surface representation in the mono condition than mono-pre condition (Figure 4a). We recognize that the mono condition resembles the viewing condition where one first views a blank field and then shifts gaze to a texture surface, whereas the mono-pre condition is similar to the viewing condition of gaze shift from one texture surface to another texture surface. In the mono-pre condition, the test grating signals may initially integrate with the iconic memory signals from the preceding grating to form a plaid pattern neural image representation. With time, the spreading wave of the horizontal test grating originating from the BC overcomes the plaid signals to form the global (test) grating representation. Thus, the integration operation in the mono-pre condition requires extra time compared to the mono condition, which does not generate an initial plaid representation. Similarly, as mentioned in the previous paragraph, the spreading wave from the BC in the dichoptic
conditions needs to integrate with the local texture signals in the interior. However, one key difference between the mono-pre and dichoptic conditions is that the plaid signals in the interior region are generated as a result of temporal interaction (forward masking) in the former but interocular interaction (binocular rivalry) in the latter.

Finally, since the average data from the mono condition are quite close to 100% (spreading) completion at the 30-ms stimulus duration (Figure 4a), one might question whether the monocural surface is always represented according to the border-to-interior strategy. We believe our behavioral method may not be quick enough to reveal the rapid surface spreading in the mono condition for all observers. This is because we noticed that at the shortest stimulus duration tested (30 ms), two of our four observers in the current study (Experiment 2) reliably reported seeing gaps in the MBC rectangle, i.e., the surface representation of the MBC rectangle for these observers was not complete at 30 ms. Our calculations revealed that the percentages of incomplete surface representation for these two observers were, respectively, 21.27 ± 2.32% and 5.02 ± 1.50%. Separately, we have found several other observers showing an incomplete surface representation at 30 ms in an unreported experiment. This thus reflects individual difference in the speed of surface representation. A systematic study with more sensitive psychophysical methods, beyond the scope of the current paper, is needed to reveal the fast speed of surface spreading in the mono condition. Such a future study could provide important insights into the related question of why we do not normally perceive the phenomenon of surface spreading in the everyday visual environment.

**General discussion**

In sum, we provide strong evidence that the visual system employs a border-to-interior strategy to represent a (grating) textured surface. This finding dovetails with the results of earlier studies using different experimental paradigms (e.g., Caputo, 1998; Davey et al., 1998; De Valois et al., 1986; Motoyoshi, 1999; Paradiso & Hahn, 1996; Paradiso & Nakayama, 1991; Rossi & Paradiso, 1996). The border-to-interior strategy may be driven by the fact that most natural surfaces are smooth, with abrupt surface discontinuation or curvature changes mainly occurring at the surface border. This being the case, most crucial information regarding image segmentation is found at the boundary contour (Grossberg & Mingolla, 1985; Nakayama et al., 1995). Moreover, the color and texture of the surface region adjacent to the BC usually provides a good estimate of the entire surface, as the interior of most natural surfaces have common optical properties (color and texture; Elder & Goldberg, 2002; Fine, MacLeod, & Boynton, 2003; Yang & Purves, 2003). There is evidence that our visual system capitalizes on such regularities to represent uniform colored surfaces (Attneave, 1954; Barlow, 1961). With the lateral inhibition mechanism implemented early in the retina, retinal ganglion cells produce vigorous responses to luminance changes (edge contrasts) of the stimulus but little responses to the interior, uniform luminance area of the stimulus (Kuffler, 1953). For representing a textured surface, a similar lateral inhibition mechanism is also implemented at the visual cortex that causes the neurons to be more sensitive to the BC of the textured surface than to the interior textured region of the surface (Knierr & van Essen, 1999; Kourtzi, Tolias, Altmann, Augath, & Logothetis, 2003; Lamme, 1995; Nothdurft et al., 2000; Sillito, Grieve, Jones, Cuderio, & Davis, 1995). Thus, by relatively strengthening the surface BC signals early on, the visual system can facilitate the later surface representation process to select the BC. For example, in Figure 1c, the lateral inhibition mechanism could enhance the monocular BC signal in the left half-image. This then facilitates the MBC disk to be selected for surface representation over the homogeneous grating in the other half-image.

The notion of strong BC signals complements our findings that the visual system uses the border-to-interior strategy to represent a grating texture surface. Since we used a psychophysical method that measures the position of the leading front of the surface-spreading wave at various durations, our findings provide the first direct and explicit behavioral evidence that a spreading-in operation (the isomorphic theory) is involved in textured surface representation.

The observed spreading wave in our study reveals a fundamental operation of surface representation. We speculate two possible underlying neural mechanisms supporting this operation. The first possible mechanism, which was briefly described in the Introduction section, establishes the BC as a reference (Bakin et al., 2000; von der Heydt et al., 2003; Zhou et al., 2000), and then uses the local texture region adjacent to the BC to sequentially integrate with the remaining local texture signals (in the interior image) from the early cortical neurons (with similar selectivity) to form a global surface representation. An advantage of using the BC, which carries rich 3-D surface layout information as a reference, is that the visual system can reduce some 3-D ambiguity when integrating the local texture information (He, Wu, Ooi, Yarbrough, & Wu, 2004; Mitchison & McKee, 1987; Wu, Ooi, & He, 2004). The second possible mechanism is related to a high-level interpolation process that modally creates a texture surface “image” based on the texture information at the BC. The image needs to be created because the interior homogeneous texture signals at the early level are suppressed by lateral inhibition and, thus, do not contribute to the later surface representation stage. An implication of this process is that the textured surface we perceive in the real world is actually an estimate of the global surface whose information is largely contributed by the signals near the BC. At first blush, this conjecture appears to be at odds with our normal experience of seeing a textured surface where we
are able to scrutinize the rich, local structures of the surface in detail. However, it does not need to be, as we propose that to see the local surface structures requires focal attention to be directed to the early local features. With attention, the lateral inhibition on the local texture signals is reduced, thus allowing its contribution to texture surface representation to be realized.

**Appendix A**

**Analysis of the perceived gap size data: Transforming between retinal and cortical domains**

From the observers’ perceived gap size (G, in degree of visual angle) data, we calculated the average spreading from one side of the horizontal grating rectangle (length = L, in degree) in the retinal domain (S\text{ret}, in degree):

\[ S_{\text{ret}} = 0.5(L - G). \tag{A1} \]

The derived average S\text{ret} is then plotted as a function of stimulus duration (D) in Figures 2c and 4a, respectively, for Experiments 1 and 2. To estimate the average spreading in the cortical domain (S\text{cor}), we used the V1 cortical magnification factor, M = 15.87E^{-1}, where E is eccentricity in degree. This factor is taken from Engel et al.’s (1997) fMRI study of the retinotopic organization of the human visual cortex between 2° and 12° of eccentricity:

\[ S_{\text{cor}} = \int_{0.5L-S_{\text{ret}}}^{0.5L} Mde \]
\[ = 15.87[\ln(0.5L) - \ln(0.5L - S_{\text{ret}})]. \tag{A2} \]

The predicted retinal spreading is plotted as solid curves in Figures 2c and 4a, for Experiments 1 and 2, respectively.

For the graphs shown in Figures 3 and 5 (for V1 and V2), we also performed similar analyses based on the cortical magnification factors derived from another fMRI study of human cortical magnification for areas V1, V2, V3v, and V4v between 0.5° and 12° of eccentricity (Seren et al., 1995); M\text{V1} = 20.05(E + 0.08)^{-1.26}, M\text{V2} = 25.19(E + 0.09)^{-1.53}, M\text{V3v} = 18.28(E + 0.24)^{-1.75}, M\text{V4v} = 18.17(E + 0.24)^{-1.55}. We thus obtained the cortical spreading functions for areas V1–V4 as follows:

\[ S_{\text{V1cor}} = 77.12 [(0.5L - S_{\text{V1ret}} + 0.08)^{-0.26} - (0.5L + 0.08)^{-0.26}], \tag{A4} \]
\[ S_{\text{V2cor}} = 47.53 [(0.5L - S_{\text{V2ret}} + 0.09)^{-0.53} - (0.5L + 0.09)^{-0.53}], \tag{A5} \]
\[ S_{\text{V3vcor}} = 24.37 [(0.5L - S_{\text{V3vret}} + 0.24)^{-0.75} - (0.5L + 0.24)^{-0.75}], \tag{A6} \]
\[ S_{\text{V4vcor}} = 33.04 [(0.5L - S_{\text{V4vret}} + 0.24)^{-0.55} - (0.5L + 0.24)^{-0.55}]. \tag{A7} \]

The average results are shown in Figures 3 (Experiment 1) and 5 (Experiment 2). Clearly, this analysis also indicates a constant cortical spreading speed (linear regression lines). Accordingly, using the estimated cortical velocities (slopes of the regression lines), we can model the retinotopic data with the following functions:

\[ S_{\text{est.V1ret}} = 0.5L + 0.08 - [0.013VD + (0.5L + 0.08)^{-0.26}]^{-3.85}, \tag{A8} \]
\[ S_{\text{est.V2ret}} = 0.5L + 0.09 - [0.021VD + (0.5L + 0.09)^{-0.53}]^{-1.89}, \tag{A9} \]
\[ S_{\text{est.V3vret}} = 0.5L + 0.24 - [0.041VD + (0.5L + 0.24)^{-0.75}]^{-1.33}, \tag{A10} \]
\[ S_{\text{est.V4vret}} = 0.5L + 0.24 - [0.030VD + (0.5L + 0.24)^{-0.55}]^{-1.82}. \tag{A11} \]
Note that even though Figures 3 and 5 do not include the fittings for V3v and V4v, an examination of the data suggests that the velocity would also be constant.

Acknowledgments

This study was supported by a grant from the National Institutes of Health (EY015804) to TLO and ZJH.

Commercial relationships: none.
Corresponding authors: Teng Leng Ooi and Zijiang He.
Emails: tlooi@salus.edu; zjhe@louisville.edu.
Address: Department of Basic Sciences, Pennsylvania College of Optometry at Salus University, 8360 Old York Road, Elkins Park, PA 19027, USA; Department of Psychological and Brain Sciences, University of Louisville, Louisville, KY 40292, USA.

References


