



Non-retinotopic feature processing in the absence of retinotopic spatial layout and the construction of perceptual space from motion

Mehmet N. Ağaoğlu^a, Michael H. Herzog^b, Haluk Ögmen^{a,c,*}

^a Department of Electrical and Computer Engineering, University of Houston, Houston, TX 77024-4005, USA

^b Laboratory of Psychophysics, Brain Mind Institute, Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland

^c Center for Neuro-Engineering and Cognitive Science, University of Houston, Houston, TX, USA

ARTICLE INFO

Article history:

Received 24 February 2012

Received in revised form 18 July 2012

Available online 21 August 2012

Keywords:

Non-retinotopic perception

Feature attribution

Feature binding

Moving object perception

Anorthoscopic perception

ABSTRACT

The spatial representation of a visual scene in the early visual system is well known. The optics of the eye map the three-dimensional environment onto two-dimensional images on the retina. These retinotopic representations are preserved in the early visual system. Retinotopic representations and processing are among the most prevalent concepts in visual neuroscience. However, it has long been known that a retinotopic representation of the stimulus is neither sufficient nor necessary for perception. Saccadic Stimulus Presentation Paradigm and the Ternus–Pikler displays have been used to investigate non-retinotopic processes with and without eye movements, respectively. However, neither of these paradigms eliminates the retinotopic representation of the spatial layout of the stimulus. Here, we investigated how stimulus features are processed in the absence of a retinotopic layout and in the presence of retinotopic conflict. We used anorthoscopic viewing (slit viewing) and pitted a retinotopic feature-processing hypothesis against a non-retinotopic feature-processing hypothesis. Our results support the predictions of the non-retinotopic feature-processing hypothesis and demonstrate the ability of the visual system to operate non-retinotopically at a fine feature processing level in the absence of a retinotopic spatial layout. Our results suggest that perceptual space is actively constructed from the perceptual dimension of motion. The implications of these findings for normal ecological viewing conditions are discussed.

© 2012 Elsevier Ltd. All rights reserved.

1. Introduction

The spatial representation of a visual scene in the early visual system is well known. First, the optics of the eye map the three-dimensional environment into two-dimensional images on the retina. The projections from retina to early visual areas preserve neighborhood relations to generate retinotopic representations of the stimulus (e.g., Sereno et al., 1995; Tootell et al., 1982). Retinotopic representations and processing are among the most prevalent concepts in current visual neuroscience. Most theories of vision involve computations where retinotopically-based receptive fields extract features, such as oriented boundaries, texture, and color to synthesize the various attributes of the stimulus. On the other hand, it has long been known that a retinotopic representation of the stimulus is neither sufficient nor necessary for perception. In metacontrast masking (rev.: Bachmann, 1994; Breitmeyer et al., 2006), a retinotopically non-overlapping “mask” stimulus can render a target stimulus completely invisible, showing that a retinotopic representation

of the stimulus is not a sufficient condition for its perception. Anorthoscopic perception occurs when a stimulus moves behind a narrow slit, as shown in Fig. 1 (e.g., Morgan, Findlay, & Watt, 1982; Parks, 1965; Rock, 1981; Zöllner, 1862).

Under this viewing condition, all information about the stimulus falls onto the same narrow retinotopic strip over time. As a result, there is no spatially extended retinotopic representation of the stimulus. In other words, at any instant of time, the spatial layout of the stimulus does not have a complete retinotopic representation. However, observers report perceiving a spatially extended coherent form instead of a rapid succession of stimulus fragments confined into the area of the narrow slit (e.g., Fendrich, Rieger, & Heinze, 2005; Morgan, Findlay, & Watt, 1982; Parks, 1965; Rock, 1981; Zöllner, 1862).

On the one hand, one can appreciate the importance of retinotopic representations by considering tasks that are naturally suited for these representations. For example, a retinotopic representation provides a natural reference frame to drive eye movements. The retinotopic distance of a target from the fovea provides a direct position error signal that can be used to position the fovea on a select target (e.g., Orban de Xivry & Lefèvre, 2007). Retinotopic representations also provide information about occlusions among stimuli

* Corresponding author at: Department of Electrical and Computer Engineering, University of Houston, Houston, TX 77024-4005, USA. Fax: +1 713 743 4444.

E-mail address: ogmen@uh.edu (H. Ögmen).

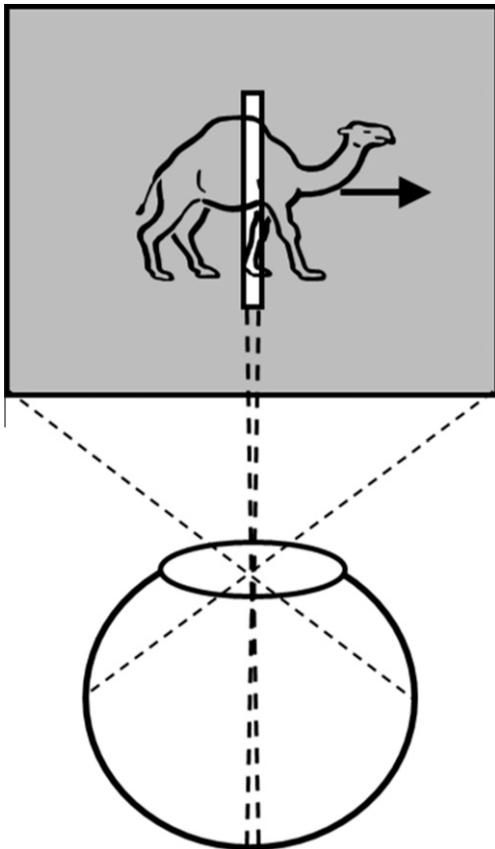


Fig. 1. Anorthoscopic perception. A moving figure (the camel) is viewed behind a narrow slit. In the actual stimulus, the gray area is completely opaque so that only the part of the figure confined to the interior of the slit is visible. All information about the form of the moving stimulus falls on the same narrow retinotopic region. There is no spatially extended retinotopic representation of the moving figure. However, observers report seeing a spatially extended coherent form instead of incoherently moving fragments confined into the narrow area inside the slit.

and as such they can be used as a visibility map that distinguishes between modal and amodal percepts. Finally, basic computations in retinotopic representations, such as motion grouping, can provide the necessary transformational signals to build non-retinotopic representations.

On the other hand, given that retinotopic representations are neither necessary nor sufficient to support our perception, a fundamental question in vision science is to determine non-retinotopic bases of information processing in the visual system. The traditional experimental method to distinguish between retinotopic and non-retinotopic processing is the “Saccadic Stimulus Presentation Paradigm” (SSPP). Here, the observer is asked to make a saccade and two stimuli are presented, one before the saccade and the second after the saccade.

As shown in Fig. 2, retinotopic and spatiotopic representations differ in the way the two stimuli match: According to the retinotopic representation of the stimulus, the ring surrounds the letter C, while according to spatiotopic representation of the stimulus (e.g., a coordinate system based in space at the center of the display monitor), the ring surrounds the letter B. Although SSPP continues to be an informative paradigm in probing non-retinotopic representations and processes, it cannot be directly applied to cases where the eyes are stationary. This is because a variety of processes related to saccadic eye movements, such as saccadic suppression and efferent copy signaling, are active in SSPP. As a result, one cannot distinguish between these eye-movement related processes and other non-retinotopic processes that may be operating

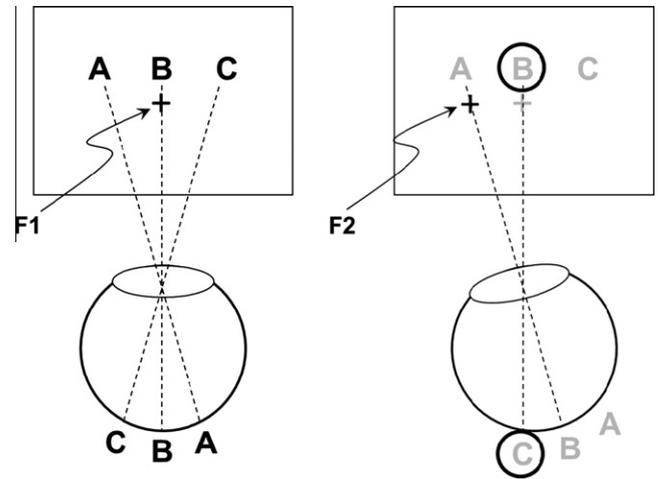


Fig. 2. Saccadic Stimulus Presentation Paradigm (SSPP). The observer first fixates on F1 (left) and is asked to make a saccadic eye movement to F2 (right). Just before the saccade, a stimulus is presented briefly (letters A B C). After the completion of the saccade a second stimulus is presented (ring). The letters A B C in light gray in the right panel are not actually presented but are shown as a reference to compare the relative positions of the ring according to retinotopic and spatiotopic reference frames (from Boi et al. (2009)).

independently from eye movements. In order to investigate non-retinotopic processes in the absence of eye movements, we have introduced an alternative stimulus paradigm (Boi et al., 2009; Ogmen, Otto, & Herzog, 2006) based on the Ternus–Pikler display (Pikler, 1917; Ternus, 1926).

Fig. 3A shows a basic Ternus–Pikler display. The first frame contains three elements. After an inter-stimulus interval (ISI), these three elements are shifted to the right by one inter-element distance so that two of the elements overlap retinotopically across the two frames. For short ISIs, observers report seeing the leftmost element of the first frame move to the rightmost element of the second frame, while the other two elements appear stationary (Fig. 3B). This percept is called “element motion”. For longer ISIs, all three elements move in tandem to the right as a group (Fig. 3C). This percept is called “group motion” (Pantle & Picciano, 1976). In order to study retinotopic versus non-retinotopic processes, we inserted a Vernier offset, called the “probe Vernier” to the central element of the first frame as shown in Fig. 3A. We asked observers to report the perceived offset direction for elements in the second frame, numbered 1, 2, and 3 in Fig. 3A. None of these elements contained a Vernier offset and naïve observers did not know where the probe Vernier was located. According to retino-

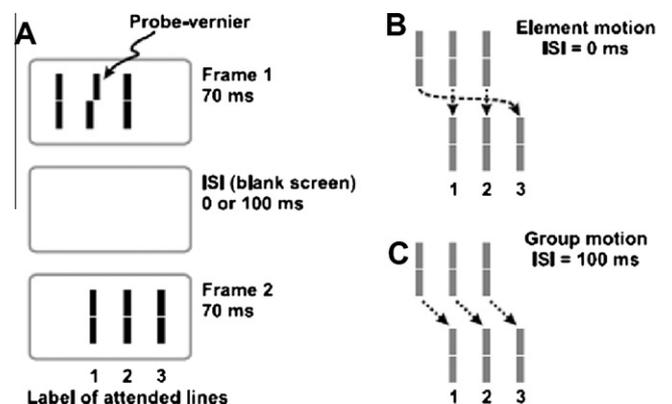


Fig. 3. The Ternus–Pikler display (A) and the associated percepts of “element motion” (B) and “group motion” (C). From Ogmen et al. (2006).

topic relations, we expect the probe Vernier to be integrated with Element 1 in the second frame for values of ISI that are within the window of temporal integration. On the other hand, if feature processing and integration take place according to motion grouping relations (Fig. 3B and C), instead of retinotopic relations, one would expect the probe Vernier to integrate with Element 1 in the case of element motion and with Element 2 in the case of group motion. Our results supported the predictions of the grouping-based non-retinotopic hypothesis (Fig. 4).

This grouping-based non-retinotopic feature processing provides strong support for the role of non-retinotopic mechanisms in processing dynamic stimuli. Based on this and several other studies using the Ternus–Pikler stimulus, we suggested that the visual system attributes features according to motion grouping relations across space and time (Boi et al., 2009; Ogmen et al., 2006; Ogmen & Herzog, 2010).

However, neither the SSPP nor the Ternus–Pikler paradigm completely rules out the contribution of retinotopic processing. In fact, in both SSPP and Ternus–Pikler displays, each stimulus frame contains a complete spatial layout of the stimulus: All elements and their spatial relations are represented in the retinotopic space. Retinotopic conflicts are created either by eye movements (SSPP) or by the retinotopically overlapping motion of the stimulus (Ternus–Pikler). Since retinotopic processes can take place during the presentation of each frame (e.g., retinotopic lateral interactions between elements may establish a spatial grouping of the ele-

ments), these paradigms do not completely eliminate potential contributions of retinotopic processes. In contrast, anorthoscopic viewing eliminates all retinotopic representations with the exception of a very narrow slit region. In addition to minimizing retinotopic representations, anorthoscopic stimuli also create retinotopic conflict in that *different parts* of the stimulus fall on the *same retinotopic area* (the interior of the slit) as the stimulus moves behind the slit. By using anorthoscopic viewing paradigm, the goal of this study was to investigate how stimulus features are processed and attributed in the absence of a retinotopic spatial layout and in the presence of retinotopic conflicts.

2. Methods

2.1. Participants

Six observers (four males) including one of the authors participated in this study. The age of the participants ranged from 24 to 37 years and all participants had normal or corrected-to-normal vision. Experiments followed a protocol approved by the University of Houston Committee for the Protection of Human Subjects. Each observer gave written consent before the experiments and volunteers were paid \$10/h for their participation.

2.2. Apparatus and stimuli

Visual stimuli were created via visual stimulus generator card VSG2/3 (Cambridge Research Systems) and displayed at a resolution of 800×600 with a refresh rate of 100 Hz. A head/chin rest was used. Observers reported their responses by a joystick. The stimulus configurations used are shown in Fig. 5. These stimuli were presented behind a narrow slit as in Fig. 1. The luminance of the opaque region was 20 cd/m^2 while the slit area was 40 cd/m^2 . The stimuli were black (1 cd/m^2). The vertical and horizontal dimensions of the slit were 7.1° and $17'$, respectively. For each condition, the display consisted of three elements, each having two abutting lines with a small ($6.2'$) vertical gap between them. The size of each abutting line was $38.7'$ for the longest element, $29.4'$ for the medium sized element and $20.1'$ for the shortest. The thickness of each line was $7.7'$ and inter-element distance was $51'$. In each trial, a Vernier offset was inserted randomly to one of the elements (probe Vernier). The direction of the offset, i.e. position of the lower segment of a given element with respect to the upper one, was also random in each trial. The experiments consisted of “Connected Lines”, “Basic Display”, and the “Flash” conditions. The rationale of stimulus design for each of these conditions and their

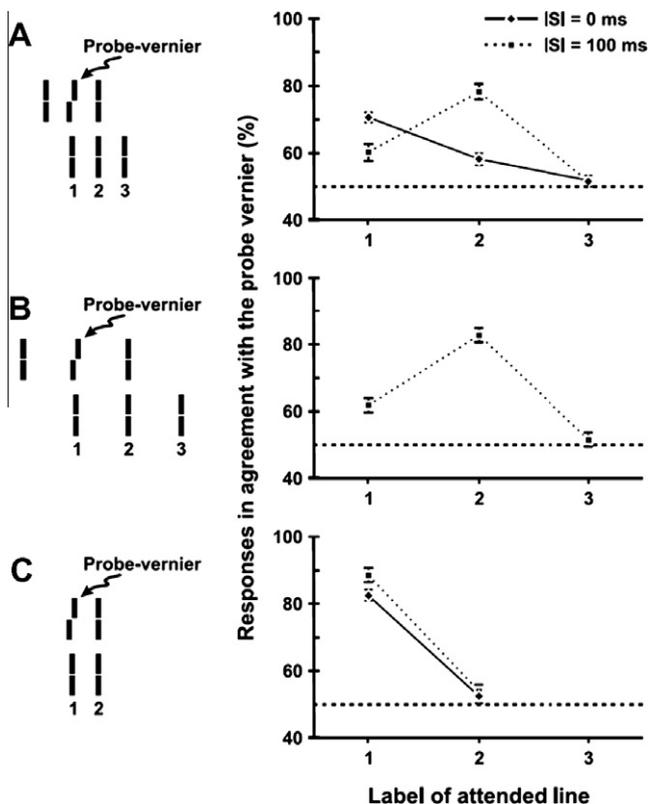


Fig. 4. The percentage of responses in agreement with the probe Vernier as a function of the attended line for the Ternus–Pikler display shown in Fig. 3. In the case of element motion (ISI = 0 ms), the probe Vernier is integrated with Element 1 in the second frame (solid line in A). However, in the case of group motion (ISI = 100 ms), the probe Vernier is integrated with Element 2 in the second frame (dashed line in A). This non-retinotopic integration also occurs when the separation between the elements is increased (B). When flanking elements are removed from the stimulus, no motion is perceived and the probe Vernier is integrated with Element 1 in the second frame for both ISI = 0 ms and ISI = 100 ms (C). From Ogmen et al. (2006).

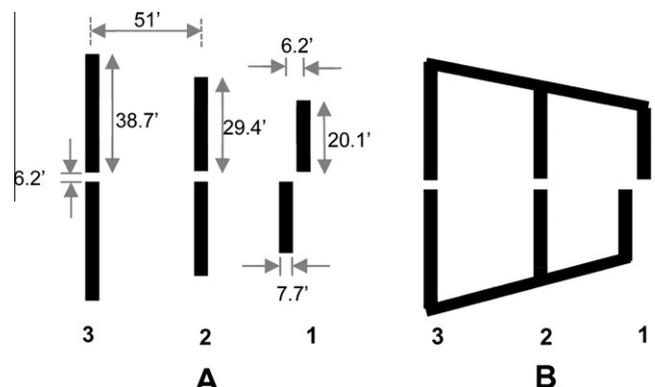


Fig. 5. Stimulus configurations presented behind the slit. (A) “Basic Display” and (B) “Connected Lines” conditions. These two stimuli are identical with the exception of the oblique lines connecting the three vertical line segments. As in the Ternus–Pikler experiments, a probe Vernier was inserted in one of the elements.

predictions are discussed in the next section. Vernier offset sizes were selected to obtain about 85% or higher correct responses when the Vernier offset was inserted to the element attended by the subject. This yielded a Vernier offset of 6.2' for "Connected Lines" and "Basic Display" conditions and 3.1' for the "Flash" condition. The speed of the stimulus behind the slit was 5.1°/s. In the "Flash" condition, the duration of the flash period (70 ms) for each element and the inter-stimulus interval (100 ms) between the flashes were chosen to mimic the timings in the other two conditions. The direction of motion (rightward or leftward) was randomized from trial to trial. In all conditions, the shortest element (Element 1 in Fig. 5) was always presented first inside the slit; in other words, the configurations shown in Fig. 5 were used for rightward motion and their horizontally mirror-symmetric versions were used for leftward motion.

2.3. Procedures

The task of the observers was to attend to a given element and report the perceived direction of the Vernier offset for that element by pressing a button. Observers attended to the same pre-determined element within a given block. Performance was evaluated as the percentage of responses in agreement with the actual direction of the probe Vernier, regardless of which element contained the probe Vernier. The next trial started by pressing another button. There were 120 trials in a block. Within each block, the same stimulus configuration was used and the number of trials in which the probe Vernier is assigned to a particular element was 40. All conditions were presented three times and the order of conditions was randomized. Training sessions were run before the actual experiment in order to familiarize the observers with the equipment and the task. The results of training sessions were not included in the data analysis.

3. Rationale of stimulus design and predictions

We have designed three stimulus conditions based on the degree to which they produced an anorthoscopic percept. Motion information is crucial in generating anorthoscopic percepts (e.g., Aydin, Herzog, & Ogmen, 2008; McCloskey & Watkins, 1978; Morgan, Findlay, & Watt, 1982; Rieger et al., 2007; Shimojo & Richards, 1986). A stationary stimulus behind the slit, or a stimulus which is flashed behind the slit, generate a percept of stimulus fragments confined to the retinotopic area of the slit. On the other hand, when the stimulus moves behind the slit, a process akin to amodal completion (Michotte, Thinès, & Crabbé, 1964) takes place, where different parts of the stimulus are perceived in a spatially integrated form revealing the complete spatial layout of the stimulus.

Our first condition, the "Flash" condition, was a baseline control condition in that no anorthoscopic percept was generated: We flashed the three line segments shown in Fig. 5A, centered in the region of the slit one after the other. The temporal order of presentation was from the shortest to the longest segment.

Our second condition, the "Connecting Lines" condition was designed to generate a strong anorthoscopic percept. Here, the three line segments were connected by oblique lines to generate the spatial layout shown in Fig. 5B. This figure was moved behind the slit. Having line segments of different vertical length allowed us to connect them with oblique rather than horizontal lines. The advantage of an oblique over a horizontal line comes from the constraint imposed by the slit: If a horizontal line is moved behind a vertical slit, motion signals will be generated only when the line enters or leaves the slit. On the other hand, an oblique line generates a motion signal during all time instants when it moves behind the slit. As suggested by previous research (e.g., Sohmiya & Sohmiya, 1994), composite figures, such as our "Connecting Lines" configu-

ration, that generate both vertical and horizontal motion signals are best suited for generating strong anorthoscopic percepts.

Finally, a third condition, the "Basic Display" condition was created as an intermediate between the previous two conditions. As shown in Fig. 5A, the stimulus was identical to that of Fig. 5B with the exception of the removal of oblique connecting lines. When this stimulus was moved behind the slit, the continuous motion signal generated by oblique lines was no longer present. Instead, only a brief horizontal motion signal was generated when each line segment passed through the slit. Thus anorthoscopic percept was either absent or very weak. In this sense, this condition is similar to the baseline "Flash Condition" with the exception that, the line segments moved in a way identical to the "Connecting Lines" conditions instead of being flashed. Our predictions were that, if present, non-retinotopic processes should occur prominently in the "Connecting Lines" condition but not in the "Flash" condition, where no anorthoscopic percept is generated. In the "Basic Display" condition, anorthoscopic percept is either weak or absent, therefore we expected weak, if any, non-retinotopic processing. Demos 1, 2, and 3 illustrate different conditions used in the experiment.

To measure non-retinotopic processes, we assessed the perception of the Vernier offset, as it was done in our previous study using the Ternus–Pikler paradigm (Ogmen et al., 2006) discussed in the preceding section. Predictions of the retinotopic and non-retinotopic feature processing and integrations are depicted in Fig. 6 for the stimuli used in this study. According to the retinotopic hypothesis, Vernier offset of a stimulus should be integrated with those stimuli that are presented at the same retinotopic region within the temporal integration period (ca. 120 ms) (Coltheart, 1980). The integration is graded, strongest for short ISIs and weaker for longer ISIs. Accordingly, the prediction is that the accordance of observers' responses with the Vernier offset direction will be highest when the observers attend to the location of the element containing the probe Vernier and should be lower but significantly better than chance for other elements presented within approximately 120 ms of the probe Vernier (Fig. 6A). The agreement between observers' responses and the probe Vernier should be at chance for elements presented outside the temporal integration window. The retinotopic processing hypothesis predicts that the same pattern of results will be obtained in all three conditions ("Connecting Lines", "Basic Display", and "Flash"). Note that in our depictions of the predictions, for simplicity, we assumed an equal level of performance for the attended line when the physical location of the probe Vernier is 1, 2, or 3. Moreover, when Element 2 is attended, we assumed an equally effective integration for the cases where the probe Vernier is located in Element 1 and Element 3. However, these assumptions may not hold since the Elements 1, 2, and 3 had different lengths and the effectiveness of integration may be influenced by temporal order (Scharnowski, Hermens, & Herzog, 2007). It is difficult to quantify a priori the potential contributions of these factors for our stimuli. Nevertheless, as we will discuss below, the exact shapes of these functions are not critical in comparing the predictions of the retinotopic versus non-retinotopic hypotheses.

According to non-retinotopic processing hypothesis, retinotopic integration should be prevented when the anorthoscopic percept indicates that the elements are spatially distinct parts of a spatially extended figure. In this case, the accordance of the observers' responses with the direction of the probe Vernier should be high when observers attend the element containing the probe Vernier and it should be at chance for all other elements (Fig. 6B). Since a strong anorthoscopic percept is generated in the "Connecting Lines" condition, the non-retinotopic hypothesis predicts the outcome shown in Fig. 6B for this condition. For the "Flash" condition, it predicts the pattern shown in Fig. 6A since no anorthoscopic percept is generated

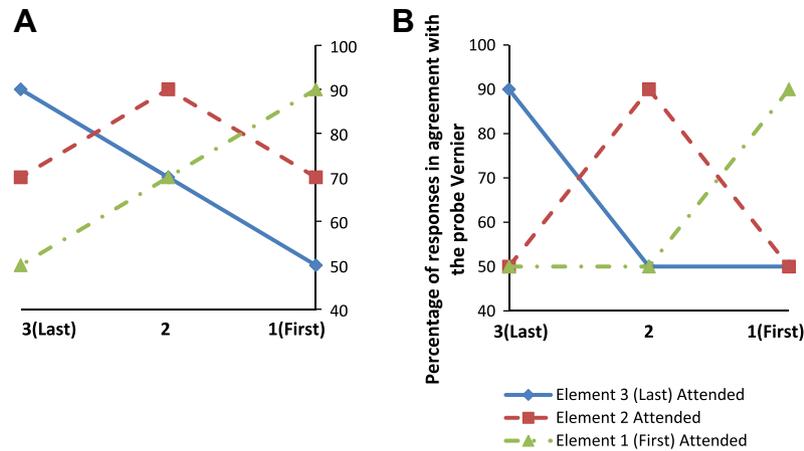


Fig. 6. Predictions of the retinotopic and non-retinotopic feature processing hypotheses. The percentage of responses in agreement with the offset direction of the probe Vernier is plotted as a function of the physical position of the probe Vernier. Each curve corresponds to a case where a specific line is attended by the subject. The retinotopic hypothesis predicts the pattern of results shown in (A) for all three conditions (“Flash”, “Basic Display”, and “Connecting Lines”). The non-retinotopic hypothesis predicts the pattern shown in (A) for the “Flash” condition and the pattern shown in (B) for the “Connecting Lines” condition. It also predicts that the “Basic Display” condition should be more similar to the pattern in (A) than the pattern in (B).

in this case. Moreover, since the anorthoscopic percept is weak or absent in the “Basic Display” condition, a pattern more similar to that in Fig. 6A than that in Fig. 6B is predicted. Therefore, the critical comparison for our study is to determine whether the “Connecting Lines” condition is significantly different from the other two conditions and whether the pattern of results for this condition is more similar to Fig. 6B compared to Fig. 6A.

4. Results

Results for all conditions are given in Fig. 7. Percentages of responses in accordance with the actual direction of the Vernier offsets are plotted against the physical location of the Vernier offsets. Different line styles (colors) represent different attended lines. Both retinotopic and non-retinotopic processing hypotheses pre-

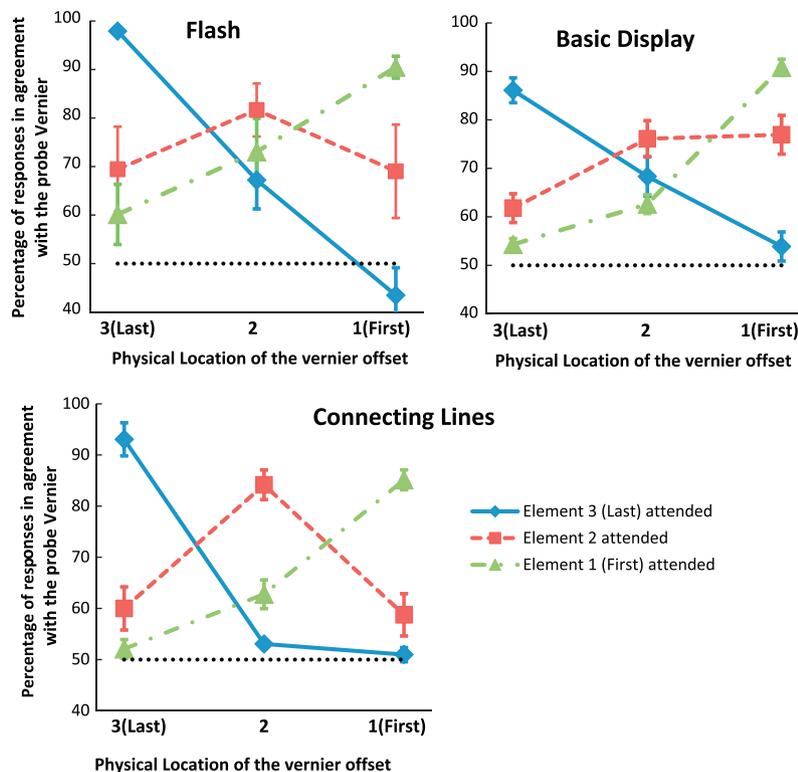


Fig. 7. Experimental results for all stimulus configurations. The horizontal axes represent the actual physical location of the probe Vernier and the vertical axes show the performance evaluated as the percentage of responses in agreement with the actual offset direction of the probe Vernier regardless which element contained the probe Vernier. In terms of temporal order, Element 1 (the shortest element, see Fig. 5) was always presented first. This is indicated on the abscissa. Different colors indicate different attended lines and horizontal dashed lines mark the chance level. Error bars represent \pm SEM across observers ($N = 6$). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

dict that the pattern of results for “Flash” and “Basic Display” conditions should be similar to the one shown in Fig. 6A. For these two conditions, when Element 3 (the longest element which always comes last in the sequence) is attended, we observe a decreasing performance as a function of the physical location of the probe Vernier. Similarly, when Element 1 (the shortest element that always appears first) is attended, we observe an increasing performance as a function of the physical location of the probe Vernier, in agreement with the qualitative predictions depicted in Fig. 6A. When Element 2 is attended, the resulting function is an inverse-V function, as depicted in Fig. 6A, for the Flash condition, but not in the Basic Display condition, possibly due to the involvement of factors that may be influenced by the size and temporal order of the stimuli.

The critical comparison to distinguish between the retinotopic and non-retinotopic feature processing hypotheses is the comparison of the results for “Connected Lines” condition to Fig. 6A (retinotopic hypothesis prediction) versus Fig. 6B (non-retinotopic hypothesis prediction). Visual inspection suggests that the “Connected Lines” results are much more similar to Fig. 6B than to Fig. 6A and therefore supports the non-retinotopic hypothesis. A repeated-measures ANOVA with conditions (Flash, Basic Display and Connecting Lines), attended lines and physical locations of the probe Vernier as main factors showed that the physical location of the probe Vernier and the location of the attended line have no significant main effects ($F(2, 10) = 0.125, p = 0.884$ and $F(2, 10) = 0.937, p = 0.424$, respectively). However, as can be seen from Fig. 7, there is a significant interaction between the physical location of the probe Vernier and the location of the attended line ($F(4, 20) = 111.525, p < 0.0001$). As mentioned in the previous section, the retinotopic hypothesis predicts that the results will be essentially the same for all three conditions while the non-retinotopic hypothesis predicts that the results for the “Flash” and “Basic Display” conditions should be similar to each other but different from the “Connecting Lines” condition. A repeated-measures ANOVA showed that stimulus condition (“Flash”, “Basic Display”, and “Connecting Lines”) does not have a significant main effect ($F(2, 10) = 2.274, p = 0.153$). This is not an unexpected result since ANOVA lumps all points in each plot shown in Fig. 7 and compares these averaged values. According to both retinotopic and non-retinotopic hypotheses, averages should be similar, if not equal.

Whereas we found clear evidence for non-retinotopic processing, this does not completely rule out the involvement of retinotopic processing. To investigate retinotopic contributions, we chose the spatial locations in the displays at which retinotopic and non-retinotopic hypotheses predict different performances. There are four points (combination of factors) in each condition that meet this criterion: (a) When observers attended to either Element 1 or Element 3 while the probe Vernier was at Element 2 and (b) When Element 2 was attended while either Element 1 or Element 3 carried the probe Vernier. Deviation of performance from chance level at these points signals the existence of residual retinotopic integration. Moreover, performance when an element carries the probe Vernier is attended (peak points in Fig. 7) depends on the temporal order of the element and visual masking that this element undergoes resulting in different peak values in Fig. 7. For instance, observers consistently perform slightly worse when they attend Element 2 while it also carried the probe Vernier (red dashed lines in Fig. 7) than when they attend Element 1 or Element 3 and probe Vernier is at the same element. Therefore we calculated the relative retinotopy as the ratio of deviation of performance from chance level and the difference of peak performance from chance level for that specific physical location of probe Vernier. Fig. 8 shows relative retinotopies for all conditions. As predicted by the non-retinotopic hypothesis, retinotopic integration is very effective in the Flash condition and it is reduced by the

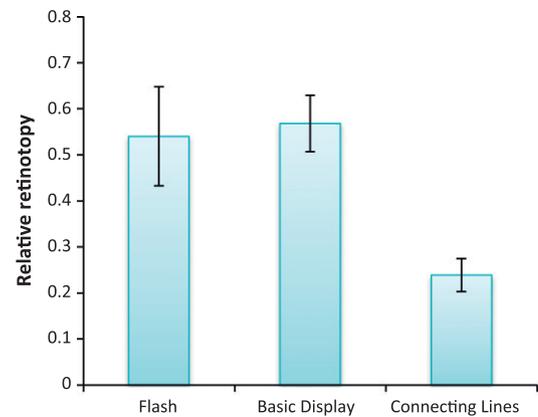


Fig. 8. Relative retinotopies for different experimental conditions. Error bars represent \pm SEM across observers ($N = 6$).

anorthoscopic percept in the Connecting Lines condition. A one-way repeated measures ANOVA shows a significant effect of condition on retinotopic integration ($F(2, 10) = 5.728, p = 0.022$). Post hoc multiple comparisons (without any correction) showed that Basic Display and Connecting Lines conditions are significantly different (mean difference = 0.330, std. error = 0.088, $p = 0.014$) and the difference between Flash and Connecting Lines conditions is only marginally significant (mean difference = 0.302, std. error = 0.119, $p = 0.052$). Considering the spatiotemporal layout of the stimulus display in this experiment, the difference between Basic Display and Connecting Lines conditions stress prominently the effect of non-retinotopic representations. In these conditions, everything else was equal except the two oblique lines that enhance the anorthoscopic percept. On the other hand, as opposed to predictions of the pure non-retinotopic hypothesis, there is some residual retinotopy left in the Connecting Lines condition. This is in agreement with our previous experiments which also showed residual retinotopy. This can be seen, for example, in the data from (Ogmen et al., 2006) study reproduced in Fig. 4. Responses in agreement with the probe Vernier for Element 1 are larger than 50% (top and middle panels in Fig. 4).

5. Discussion

Retinotopic organization of the early visual system is well established. Most theories of vision are based on this retinotopic organization: Features are extracted through retinotopically localized receptive fields of neurons and these features are integrated through connections between neurons positioned at various retinotopic locations. However, as discussed in Section 1, metacontrast masking and anorthoscopic perception demonstrate that retinotopic representations are neither sufficient nor necessary for the perception of spatially extended form. Several recent studies provided evidence for non-retinotopic bases for the computation of stimulus features such as form (Ogmen et al., 2006), color (Nishida et al., 2007), luminance (Shimozaki, Eckstein, & Thomas, 1999), size (Kawabe, 2008), motion (Boi et al., 2009; Melcher & Morrone, 2003), and position (Fischer, Spotswood, & Whitney, 2011). However, in all these studies retinotopic representations of the stimuli were available for visual processing. As a consequence, although these studies showed non-retinotopic processing, they did not completely exclude contributions of retinotopic processes. In this study, by using anorthoscopic perception, we minimized the retinotopic stimulus to a very narrow area corresponding to the inside region of the slit. Under this condition, we tested whether simple feature processing would follow retinotopic or non-retinotopic

rules. Our results support the operation of non-retinotopic processes even when retinotopic representations are minimized.

To probe the broader significance of these findings, one may pose two inter-related questions:

- (1) If retinotopic representations are not necessary for form perception, why is the early visual system organized retinotopically?
- (2) If the early visual system is organized retinotopically, why does the visual system use non-retinotopic processes and representations?

We suggest that retinotopic organization emerges, on the one hand, from the physics of image formation apparatus and, on the other hand, from the fact that spatial contiguity is a common property of objects in our environment. Based on the physics of image formation, the retinotopic representations can be used to control eye movements by providing error signals of targets with respect to fovea. The spatial contiguity of physical objects and the attendant neighborhood relations are preserved by the optical characteristics of the eyes. In retinotopic representations, breakdown of neighborhood relations often correspond to occlusions; as a result retinotopic representations can be used as a visibility map to distinguish between modal and amodal percepts. However, the visual system is forced to transform these retinotopic representations into non-retinotopic ones as a result of ecological constraints. Under normal viewing conditions, our eyes undergo frequent movements making retinotopic representations highly unstable. Moreover, many objects in the environment are in motion and, given the visible persistence of vision, these objects would generate highly blurred percepts if the computations were to take place in retinotopic representations (Ogmen, 2007). Finally, occlusions are ubiquitous in our environment and occlusions often blank out significant parts of retinotopic representations both in a static and dynamic manner. Therefore the visual system needs to create representations where objects can be processed and represented in a way that is invariant with respect to dynamic changes and occlusions. The achievement of object continuity and permanence is a significant milestone in cognitive development (Piaget, 1976) and indicates the central role non-retinotopic processes play both in perception and cognition.

Anorthoscopic perception provides a window through which one can investigate how these non-retinotopic representations are constructed in the visual system. In particular, how is the dimension of space constructed to allow spatially extended percepts from stimulations that fall in a very limited retinotopic region? “Retinal painting” was one of the first hypotheses to explain anorthoscopic percepts (Helmholtz, 1867). The explanation was based on the assumption that eye movements enable the spatial expansion of stimuli from the slit region to a broader retinotopic region: If the eyes move while viewing the stimulus, then successive parts of the stimulus fall on adjacent retinotopic loci thereby “painting” a retinotopic spatially extended picture of the figure. However, several studies showed that, while retinal painting can occur when the eyes move in a coordinated fashion, it cannot explain anorthoscopic perception in general: Measurement of eye movements and studies using retinal stabilization provide evidence that anorthoscopic perception *does also occur* without the contribution of eye movements (Fendrich, Rieger, & Heinze, 2005; Morgan, Findlay, & Watt, 1982). Another theory, proposed by Parks (1965), suggests that a post-retinal mechanism stores in memory the information available through the slit and reconstructs the spatial layout of the figure according to a “time-of-arrival coding”. Fig. 9 shows a stimulus which was used to test retinal painting and the time-of-arrival reconstruction theories (McCloskey & Watkins, 1978; Sohmiya & Sohmiya, 1992, 1994). The stimulus consists of two triangular

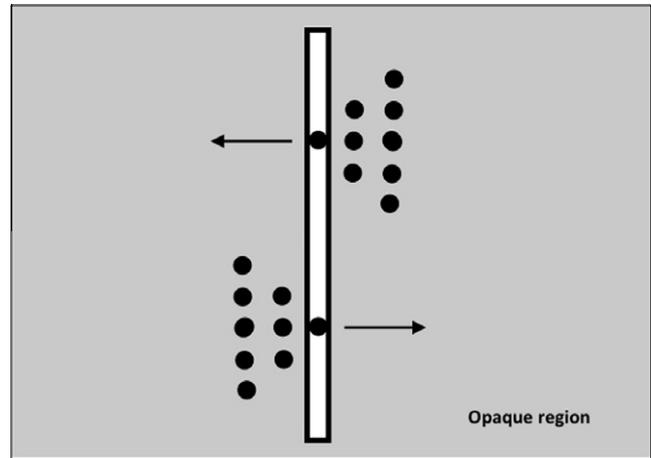


Fig. 9. Experimental stimulus used to test retinal painting and time-of-arrival coding theories of anorthoscopic perception.

shapes made up of dots moving in opposite directions. The tips of the triangles pass through the slit simultaneously, followed by the middle segments and finally the longest segments. Since the tips, the middle segments and the bases of the two triangles arrive to the slit region at the same time instants, the time-of-arrival coding theory predicts that observers will perceive two identical, instead of mirror image, triangles. Similarly, since eye movements affect the two triangles exactly the same way, the retinal painting theory makes the same prediction. In contradiction to this prediction, observers' perception corresponds to the actual stimulus configuration, i.e. two mirror image triangles (McCloskey & Watkins, 1978; Sohmiya & Sohmiya, 1992, 1994). Therefore, the space dimension is created neither through eye movements nor through a transformation of time dimension. Instead, the direction of motion information is critical in constructing spatial representations. If the direction of motion is not known, the stimulus is ambiguous in that a leftward moving image and its mirror-symmetric version moving rightward generate identical patterns in the slit. Therefore, the determination of the direction of motion is critical for anorthoscopic perception. We have proposed that the space dimension is created from motion dimension (Ogmen, 2007). This proposal has been supported by a recent study where spatial distortions in anorthoscopic perception were examined (Aydin, Herzog, & Ogmen, 2008). The figure moving behind the slit typically appears spatially compressed in the direction of its motion (e.g., Anstis & Atkinson, 1967; Aydin et al., 2009; Haber & Nathanson, 1968; Helmholtz, 1867/1962; McCloskey & Watkins, 1978; Morgan, Findlay, & Watt, 1982; Parks, 1965; Rock, 1981; Rock & Sigman, 1973; Zöllner, 1862). Aydin, Herzog, and Ogmen (2008) showed that this spatial compression can be explained by the differences in the perceived speeds of the leading and trailing parts of the anorthoscopic figure. If indeed, space is constructed from motion, one would expect the spatial metric (perceived distances and sizes) to depend on the motion metric (perceived speeds).

Taken together, these studies suggest that to process and represent information about the environment during ecological viewing conditions, the visual system uses non-retinotopic representations where space is constructed from motion. While the single slit configuration used in this study may appear an extreme form of occlusion, that spatial form is constructed by using motion signals has also been demonstrated in the case of multiple slits (Nishida, 2004). As shown in the current study, the visual system is also capable of preserving figural, as opposed to purely retinotopic, relationships as features are processed in these motion-based non-retinotopic representations.

Acknowledgments

This work was supported in part by Award Number R01 EY018165 from the National Institutes of Health (NIH). M.H. was supported by the Swiss National Science Foundation (SNF) Project “Basics of visual processing”. The content is solely the responsibility of the authors and does not necessarily represent the official views of NIH. We thank the anonymous reviewers for helpful comments and suggestions.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.visres.2012.08.009>.

References

- Anstis, S. M., & Atkinson, J. (1967). Distortions in moving figures viewed through a stationary slit. *American Journal of Psychology*, *80*, 572–585.
- Aydin, M., Herzog, M. H., & Ogmen, H. (2009). Shape distortions and Gestalt grouping in anorthoscopic perception. *Journal of Vision*, *9*(3), 1–8. <<http://journalofvision.org/9/3/8/>> <http://dx.doi.org/10.1167/9.3.8> (article no. 8).
- Aydin, M., Herzog, M. H., & Ogmen, H. (2008). Perceived speed differences explain apparent compression in slit viewing. *Vision Research*, *48*, 1603–1612. <http://dx.doi.org/10.1016/j.visres.2008.04.020>.
- Bachmann, T. (1994). *Psychophysiology of visual masking: The fine structure of conscious experience*. Commack, NY: Nova Science.
- Boi, M., Ogmen, H., Krümmenacher, J., Otto, T. U., & Herzog, M. H. (2009). A (fascinating) litmus test for human retino- vs. non-retinotopic processing. *Journal of Vision*, *9*, 1–11 (article no. 5).
- Breitmeyer, B. G., Kafaligonul, H., Ogmen, H., Mardon, L., Todd, S., & Ziegler, R. (2006). Meta- and paracontrast reveal differences between contour- and brightness-processing mechanisms. *Vision Research*, *46*, 2645–2658. <http://dx.doi.org/10.1016/j.visres.2005.10.020>.
- Coltheart, M. (1980). Iconic memory and visible persistence. *Perception and Psychophysics*, *27*, 183–228. <http://dx.doi.org/10.3758/bf03204258>.
- Fendrich, R., Rieger, J. W., & Heinze, H. J. (2005). The effect of retinal stabilization on anorthoscopic percepts under free-viewing conditions. *Vision Research*, *45*, 567–582. <http://dx.doi.org/10.1016/j.visres.2004.09.025>.
- Fischer, J., Spotswood, N., & Whitney, D. (2011). The emergence of perceived position in the visual system. *Journal of Cognitive Neuroscience*, *23*, 119–136. <http://dx.doi.org/10.1162/jocn.2010.21417>.
- Haber, R. N., & Nathanson, L. S. (1968). Post-retinal storage? Some further observations on Parks' camel as seen through the eye of a needle. *Perception and Psychophysics*, *3*, 349–355.
- Helmholtz, H. von (1867/1962). *Handbook of physiological optics: JS Southhall translation* (3rd ed.). New York: Dover.
- Kawabe, T. (2008). Spatiotemporal feature attribution for the perception of visual size. *Journal of Vision*, *8*. <http://dx.doi.org/10.1167/8.8.7>.
- McCloskey, M., & Watkins, M. J. (1978). Seeing-more-than-is-there phenomenon – Implications for locus of iconic storage. *Journal of Experimental Psychology – Human Perception and Performance*, *4*, 553–564. <http://dx.doi.org/10.1037/0096-1523.4.4.553>.
- Melcher, D., & Morrone, M. C. (2003). Spatiotopic temporal integration of visual motion across saccadic eye movements. *Nature Neuroscience*, *6*, 877–881. <http://dx.doi.org/10.1038/nn1098>.
- Michotte, A., Thinès, G., & Crabbé, G. (1964). *Les compléments amodaux des structures perceptives*, *Studia Psychologica*. Louvain: Publications Universitaires de Louvain.
- Morgan, M. J., Findlay, J. M., & Watt, R. J. (1982). Aperture viewing: A review and a synthesis. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*, *34*, 211–233.
- Nishida, S. (2004). Motion-based analysis of spatial patterns by the human visual system. *Current Biology*, *14*, 830–839.
- Nishida, S., Watanabe, J., Kuriki, I., & Tokimoto, T. (2007). Human visual system integrates color signals along a motion trajectory. *Current Biology*, *17*, 366–372. <http://dx.doi.org/10.1016/j.cub.2006.12.041>.
- Ogmen, H. (2007). A theory of moving form perception: Synergy between masking, perceptual grouping, and motion computation in retinotopic and non-retinotopic representations. *Advances in Cognitive Psychology*, *3*, 67–84.
- Ogmen, H., Breitmeyer, B. G., Todd, S., & Mardon, L. (2006). Target recovery in metacontrast: The effect of contrast. *Vision Research*, *46*, 4726–4734. <http://dx.doi.org/10.1016/j.visres.2006.09.006>.
- Ogmen, H., & Herzog, M. H. (2010). The geometry of visual perception: Retinotopic and nonretinotopic representations in the human visual system. *Proceedings of the IEEE*, *98*, 479–492.
- Ogmen, H., Otto, T. U., & Herzog, M. H. (2006). Perceptual grouping induces non-retinotopic feature attribution in human vision. *Vision Research*, *46*, 3234–3242. <http://dx.doi.org/10.1016/j.visres.2006.04.007>.
- Orban de Vivry, J. -J., & Lefèvre, P. (2007). Saccades and pursuit: Two outcomes of a single sensorimotor process. *The Journal of Physiology* *584*, 11–23.
- Pantle, A., & Picciano, L. (1976). A multistable movement display: Evidence for two separate motion systems in human vision. *Science (New York, NY)*, *193*, 500–502. <http://dx.doi.org/10.1126/science.941023>.
- Parks, T. E. (1965). Post-retinal visual storage. *The American Journal of Psychology*, *78*, 145–147. <http://dx.doi.org/10.2307/1421101>.
- Piaget, J. (1976). *The child and the reality: Problems of genetic psychology*. New York: Penguin Books.
- Pikler, J. (1917). *Sinnesphysiologische Untersuchungen*. Leipzig, Germany: Barth.
- Rieger, J. W., Gruschow, M., Heinze, H. J., & Fendrich, R. (2007). The appearance of figures seen through a narrow aperture under free viewing conditions: Effects of spontaneous eye motions. *Journal of Vision*, *7*. <http://dx.doi.org/10.1167/7.6.10>.
- Rock, I. (1981). Anorthoscopic perception. *Scientific American*, *244*, 145–153.
- Rock, I., & Sigman, E. (1973). Intelligence factors in the perception of form through a moving slit. *Perception*, *2*, 357–369.
- Scharnowski, F., Hermens, F., & Herzog, M. H. (2007). Bloch's law and the dynamics of feature fusion. *Vision Research*, *47*, 2444–2452. <http://dx.doi.org/10.1016/j.visres.2007.05.004>.
- Sereno, M. I., Dale, A. M., Reppas, J. B., Kwong, K. K., Belliveau, J. W., Brady, T. J., et al. (1995). Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science*, *268*, 889–893. <http://dx.doi.org/10.1126/science.7754376>.
- Shimojo, S., & Richards, W. (1986). Seeing shapes that are almost totally occluded – A new look at Park's camel. *Perception and Psychophysics*, *39*, 418–426. <http://dx.doi.org/10.3758/bf03207070>.
- Shimozaki, S. S., Eckstein, M., & Thomas, J. P. (1999). The maintenance of apparent luminance of an object. *Journal of Experimental Psychology – Human Perception and Performance*, *25*, 1433–1453. <http://dx.doi.org/10.1037/0096-1523.25.5.1433>.
- Sohmiya, T., & Sohmiya, K. (1992). Where does an anorthoscopic image appear? *Perceptual and Motor Skills*, *75*, 707–714.
- Sohmiya, T., & Sohmiya, K. (1994). What is a crucial determinant in anorthoscopic perception? *Perceptual and Motor Skills*, *78*, 987–998.
- Ternus, J. (1926). Experimentelle Untersuchungen über phänomenale Identität. *Psychologische Research Psychologische Forschung*, *7*, 81–136.
- Tootell, R. B., Silverman, M. S., Switkes, E., & De Valois, R. L. (1982). Deoxyglucose analysis of retinotopic organization in primate striate cortex. *Science (New York, NY)*, *218*, 902–904. <http://dx.doi.org/10.1126/science.7134981>.
- Zöllner, F. (1862). Über eine neue Art anorthoscopischer Zerrbilder. *Annalen der Physik und Chemie: Poggendorffs Annalen*, *117*, 477–484.