Visual interactions in the path of apparent motion

Steven Yantis and Takèhiko Nakama

Department of Psychology, The Johns Hopkins University, Baltimore, Maryland 21218–2686, USA

Correspondence should be addressed to S.Y. (yantis@jhu.edu)

When two stationary visual objects appear in alternating sequence, they evoke the perception of a single object moving back and forth between them. This is known as stroboscopic or apparent motion and forms the basis of perceived continuity in, for example, motion pictures. When the spatiotemporal separation between the inducing objects is optimal, the subjective appearance of apparent motion is nearly indistinguishable from that of real motion. Here we report that the detection and identification of a simple visual form in the path of apparent motion is impaired by the illusory perception of an object moving through the empty space between the locations at which the inducing objects are presented. This observation may be a manifestation of perceptual completion or 'filling in' during apparent motion perception. We propose that feedback from higher to lower visual cortical areas activates an explicit neural representation of a moving object, which can then disrupt the representation of visual stimuli in the path of the movement.

It is known that visual perception goes beyond the retinal image to recover properties of the scene that are only implicit[1–6]. For example, the correct shape of a passing dog is effortlessly perceived through the slats of a picket fence despite the fragmentary retinal image. The perceptual completion that compensates for partial occlusion is due in part to visual 'filling in.' Similarly, motion is perceived even when the retinal image is formed by the alternating appearance of two spatially distinct stationary objects, a phenomenon known as stroboscopic or apparent motion[7–11]. The subjective impression of an object moving through the empty space between the motion inducers can be nearly indistinguishable from real motion, and there is indirect evidence for the presence of a sequence of intermediate representations during apparent motion[12,13]. However there is little direct evidence that the visual system explicitly represents an object in the motion path.

Here we report four psychophysical experiments in which the perception of an object is impaired when it appears on the path of apparent motion compared to when it is off the path. This finding suggests that the phenomenal experience of apparent motion is accompanied by an explicit filling-in of the apparent motion path, which can interfere with the perception of objects in that path. We suggest that areas of the brain that are specialized for representing real and apparent visual motion (for instance, the higher visual cortical areas MT or MST) may send feedback to earlier visual areas to perceptually fill in the apparent motion path[12–16] thereby interacting with visual representations of objects in the path.

Results

In the first experiment, observers viewed an arrangement of four disks flashing on and off in alternation as shown in Fig. 1. The direction of apparent motion is ambiguous and can be perceived as either vertical or horizontal [10,17]; the perceptions can be controlled experimentally using an occluding surface that gradually recedes horizontally or vertically to reveal the display (Fig. 2). Observers had to report which of two target letters appeared by pressing an appropriate button; the target appeared on the path of apparent motion on a random half of the trials and off the path on the remaining trials. This stimulus is useful in that either vertical or horizontal motion can be perceived within a single display. Because the on-path and off-path stimulus arrays were physically identical for more than two seconds between the disappearance of the occluder and the appearance of the target letter, any effect of the perceived direction of motion must be attributed to the subjective motion representation and not to any physical property of the display.

Table 1 shows the mean response time (RT) for each combination of perceived motion and target position. The time required to correctly identify the target letter was longer when the target appeared on the motion path (M = 772 ms) than when it was off the path (M = 723 ms, difference = 49 ms, $t(9) = 4.35$, $p < .001$). For example, the time required to identify a target in the left location was 698 ms when horizontal motion was perceived (so the target was not in the motion path), but increased to 728 ms when vertical motion was perceived (placing the target on the motion path, as illustrated in Fig. 2). There was no effect of the placement of the occluder (for example, left versus right side for vertical motion and left target), which rules out the possibility that a visual afterimage of the occluder or a bias to follow the receding edge of the occluder with the eyes contributed to the difference seen in Table 1. Error rates were less than 4% in all conditions.

Table 1 Mean response times for Experiment 1

Motion Direction	Target Position			
	Top	Right	Bottom	Left
Vertical	706	726	779	728
Horizontal	760	711	875	698

mean off-path: 723 mean on-path: 772

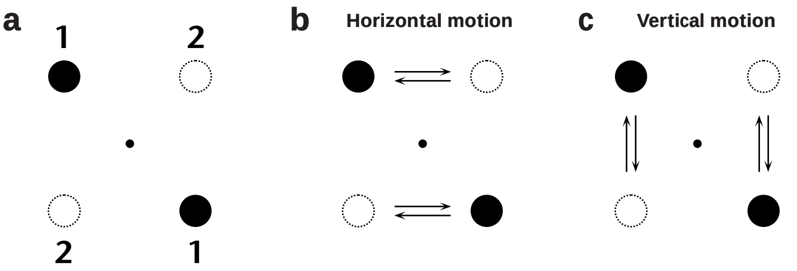

Fig. 1. Bistable apparent motion quartet. **(a)** Motion tokens appear at locations marked '1' for the stipulated frame duration, then disappear during the inter-stimulus interval (ISI), reappear at locations marked '2' for the frame duration and disappear during the ISI. This constitutes a single cycle of motion. Typical displays consist of multiple cycles. **(b, c)** The display is perceptually ambiguous and can yield either horizontal or vertical motion percepts[10].

We suggest that the perception of apparent motion gave rise to an explicit visual representation of a moving element appearing dynamically in the perceived path of motion, and this interacted with the target representation, slowing target identification. We refer to this interaction as 'motion masking.'

It is known that the perceived quality of apparent motion varies with the spatiotemporal properties of the display, such as the spatial separation and the temporal interval (stimulus-onset asynchrony or SOA) between the motion tokens[8,9,18,19]. For example, very short and very long SOAs typically yield reports of 'flicker' and 'succession,' respectively; reports of motion are weak or absent. If the observed interaction reflects an explicit representation of the motion token moving through space, then variations in perceived motion quality should yield corresponding variations in the magnitude of motion masking.

To test this prediction, we used the same bistable apparent motion display as in our first experiment, but we manipulated the quality of apparent motion by varying the SOA. We asked one group of observers to directly rate (on a five-point scale) the quality of apparent motion for each SOA; the mean ratings are shown in Fig. 3a. The perceived quality of motion followed the expected inverted-U pattern: the 'best' motion occurred with an SOA of 350 ms; shorter and longer SOAs yielded motion of poorer perceived quality.

Next, with a new group of observers, we measured the magnitude of visual interaction at each of these SOAs by recording the difference in RT for targets on and off the motion path respectively, using the receding-occluder method of Experiment 1. For SOAs up to 350 ms, RT was slower for targets appearing on the path of apparent motion than off the path, and the magnitude of this slowing closely mirrors the perceived quality of apparent motion (Fig. 3b).

A distinct pattern was observed for the two longest SOAs (as verified by a significant interaction between SOA and on-path versus off-path conditions, $F(4, 88) = 8.70$, $p < .01$). Here identification was somewhat faster for on-path than for off-path targets. We speculate that when motion tokens are integrated over long spatiotemporal intervals, observers may use deliberate attentive tracking[20–22], which could speed target identification in the tracking path. This admittedly speculative account will require further investigation.

Next we sought to manipulate the magnitude of motion masking by exploiting a phenomenon[23] termed 'path-guided motion.' When a dim gray path that links the positions of the two motion tokens is briefly flashed during the blank interstimulus interval (ISI) of an apparent motion display, a strong sensation of motion along the path is reported (Fig. 4). We reasoned that if the motion token follows a curved path around the target, then any interaction produced by the motion should be reduced or eliminated.

We tested this with a bistable apparent motion quartet using an optimal 350 ms SOA with dim grey paths flashed for 17 ms midway through the ISI (Fig. 4); to increase the strength of path-guided motion, we included a dim green disk on the path. All ten observers reported that these displays strongly induced path-guided motion. (No occluder was required because the motion path was unambiguous.) This manipulation virtually eliminated the difference in identification time for 'near-path' ($M = 772$) and off- path ($M = 763$) targets (a difference of 9

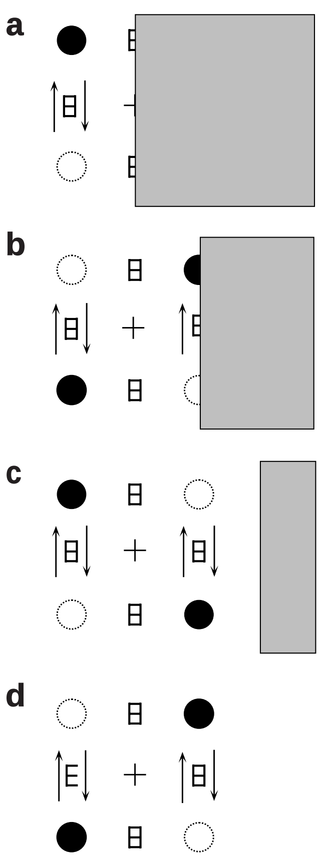

Fig. 2. Stimulus display, Experiments 1 and 2. **(a)** One of the two possible motion percepts (horizontal or vertical) was evoked by occluding two of the four motion token locations at the start of each trial for five cycles, yielding unambiguous apparent motion between the two visible tokens. (In this example, vertical motion is seen between the motion tokens on the left of the display.) **(b, c)** The occluder then gradually receded and finally disappeared. For the relatively rapid motion frequencies used here, the perceived direction of motion rarely changes from what is initially perceived[10,17]. We verified in a separate control experiment that the perception of motion persisted in the same direction as that induced by the occluder for at least 3.5 s after the occluder disappeared. **(d)** After the occluder disappeared, the motion continued for 3–4 cycles, and then a target letter (*E* or *S*) was revealed by removing two line segments from one of the four figure-eight placeholders in the display. In this example, the target letter *E* is in the motion path. All sixteen possible combinations of four target positions and four occluder positions were used equally often; thus, horizontal and vertical motion paths were induced equally often. The perceived motion path either intersected with the target's location (as in this example: left target and vertical motion), or it fell elsewhere (for example, left target and horizontal motion).

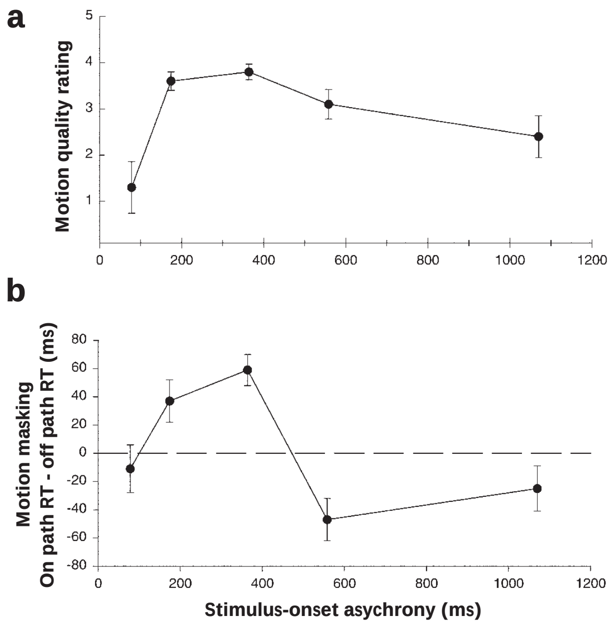

Fig. 3. Data from Experiment 2. **(a)** Mean ratings of motion quality on a five-point scale for each SOA. **(b)** Mean difference in response time for target identification on and off the path of apparent motion. Error bars signify one standard error of the mean.

ms, $t(9) = 0.71$, *ns*). The 9-ms difference was significantly smaller than the 49-ms difference observed in our first experiment ($t(18) = 2.27$, $p < .05$), providing evidence for the spatial specificity of motion masking.

Response time differences can reflect delays in perception, but they can also arise as a result of later decision-related processing. To determine whether motion masking occurs at an early perceptual level, we employed an accuracy measure in our final experiment. We measured the contrast threshold required to discriminate the position of a small gap in a briefly flashed annulus (a 'Landolt C') appearing on or off the path of apparent motion, which was again manipulated using a receding occluder. The luminance required to yield 70.7% correct discriminations was 201 cd/m^2 for on-path targets and 186 cd/m^2 for off-path targets, a significant difference ($t(8) = 2.6$, $p < .02$). Although the target annulus was always visible, the location of the gap was not. This result corroborates our RT findings and suggests that the observed interaction between apparent motion and real stimuli occurs at a relatively early point in the visual pathway.

Discussion

We have described four experiments in which the identification of a target form was impaired when the target appeared in the path of apparent motion compared to when the motion path appeared elsewhere in the display. The physical properties of the display were the same in both on-path and off-path conditions, which rules out an account that appeals to lateral interactions in the retina or early cortex. Two previous studies reported no masking in the path of apparent motion[24,25]. One significant difference between these studies and our own was the quality of the apparent motion produced by the displays used in these earlier studies. In the first study[24], only a single cycle of motion was used on each trial and the target to be detected appeared during the ISI, before apparent motion could be evoked by the appearance of the second motion token[14]. If apparent motion was evoked at all by this stimulus, it is likely to have occurred after the target was presented. In the second study[25], which used a test that is similar in some ways to the one we used, the SOA was 600 ms, a value for which we observed no motion masking (Fig. 3b).

Because the target was equally likely to appear in any of the four possible locations, it is unlikely that spatially selective attention contributed to this finding. If anything, an account asserting that motion draws attention would predict an advantage for targets in the path of motion, which is opposite of what we observed for the strongest motion stimulus (350 ms SOA, Fig. 3). It is more likely that observers distributed attention evenly over the entire display to prepare for the target when it appeared.

Our findings suggest that there is an explicit representation of a motion token in the path of apparent motion, which in turn interacts in a spatially specific manner with representations of visual form, impairing form discrimination (for related evidence, see McNamara, T. P. *et al. Annu. Meeting Psychonomic Soc.*, Nov. 1998). Two possible mechanisms may be implicated. The first is cortical feedback from higher visual areas to lower ones to effect perceptual completion. For example, neural activity in areas such as MT or MST, with relatively large receptive fields, represents apparent motion over the relatively large distances used here[26–28]. These motion areas are known to send significant feedback projections to earlier visual areas[29,30]. These feedback pathways could send reciprocal signals from the motion areas to earlier visual areas to create a neural representation of a moving token where there is no explicit sensory evidence for one. This representation could then interact with cortical responses to the target letter (Exper-

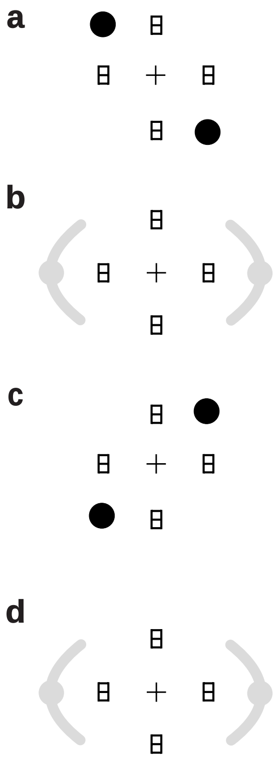

Fig. 4. Motion quartets with paths, Experiment 3. Motion tokens appeared at diagonally opposite corners for 200 ms (**a**), followed by a 150-ms blank interstimulus interval (ISI), in the center of which dim gray paths were flashed for 17 ms (**b**). This was followed by the appearance of the motion tokens in the remaining corners for 200 ms (**c**) and finally by the blank interval and briefly-flashed motion paths (**d**). Several cycles of motion were presented before one of the figure-eight placeholders changed to a letter to be discriminated. The motion tokens are perceived as following the paths[23].

iments 1–3) or target annulus (Experiment 4) in the motion path. Other examples of perceptual feedback are consistent with such a mechanism[31–33]. Alternatively, feedback to suppress motion smear[34,35] could, as a side effect, suppress the neural response to the target letter when it was in the path of apparent motion.

A second possible mechanism involves lateral interactions between the dorsal visual pathway, which is generally thought to represent motion information, and the ventral visual pathway, which is thought to represent form information[36]. Although these are distinct parallel pathways, they are anatomically connected, and interactions between form and motion representations occupying the same location could well occur[29,37–42]. Further psychophysical and neurophysiological investigation will be required to discriminate between these possibilities.

Methods

All experimental protocols were approved by the Institutional Human Subjects Review Board at The Johns Hopkins University. Stimuli were presented on a 21” Nanao F750i monitor refreshed at 60 Hz. The motion tokens were bright green (397 cd/m^2), the placeholders, target and fixation cross were dim blue (44 cd/m^2), and the occluder was dim red (13 cd/m^2) on a black background (2 cd/m^2). The center-to-center horizontal and vertical distance between motion tokens was 5° visual angle from a viewing distance of 66 cm, controlled by a chin rest. Motion tokens subtended 52 min visual angle in diameter. Seven-segment placeholders (13 × 26 min) were positioned at the midpoint of each edge of the motion quartet. The occluder subtended 4.9° × 7.6° at the beginning of each trial in Experiments 1, 2, and 4.

Ten observers participated in Experiment 1. The frame duration was 200 ms and the ISI was 150 ms. An occluder appeared at the beginning of each trial and occluded half the display (two of the four motion tokens); the occluder receded smoothly over an interval of five 700-ms motion cycles (3.5 s). After the occluder disappeared, the unoccluded quartet continued in apparent motion for 3–4.5 additional cycles (randomly selected on each trial), and then the target (*E* or *S*) was revealed by removing two line segments from one of the figure-eight placeholders. The display remained visible, and the motion continued, until the observer responded. Each observer completed 128 trials in a single session. Observers were instructed to report (as quickly as possible while minimizing errors) the identity of the target letter by pressing one of two labeled buttons on a custom response box.

In Experiment 2, one group of 12 observers performed the motion-quality rating task, and a separate group of 23 observers performed the target-identification task. The stimulus conditions were identical to Experiment 1 except the SOA was varied (67, 183, 350, 550, or 1033 ms, selected randomly for each trial). Each observer completed 160 trials.

In Experiment 3, ten observers performed the target-identification task. The stimulus conditions were the same as in Experiment 1, except there was no occluder, and a pair of curved paths (either horizontally or vertically arrayed) appeared for 16.7 ms midway through the ISI (Fig. 4). Each path was constructed by displaying a 120% gray arc (2.4 cd/m^2) with radius 2.5% visual angle centered at one of the four placeholder locations. A dim green disk (13.7 cd/m^2) was placed at the center of each path. Each observer completed 128 trials.

In Experiment 4, the receding occluder procedure of Expt 1 was used to induce vertical or horizontal apparent motion. The target was a low-contrast green annulus (diameter 34 min visual angle) that was flashed for 17 ms at the temporal midpoint of the 150-ms interstimulus interval and at the spatial midpoint of one of the four edges of the motion quartet. The background was dim gray (20 cd/m^2). The target appeared in each of the four possible spatial locations equally often, selected randomly on each trial, and it coincided with the perceived motion path on half the trials. The target annulus contained gap subtending a visual angle of 6 min on its right or left side, and the observer was to report the location of the gap; the response was not speeded.

Eight simultaneous, interleaved adaptive staircases were used to adjust the luminance of the target for each combination of location and on- versus off-path conditions using a two-up, one-down rule; this procedure converges to the luminance required to correctly discriminate the target on 70.7% of the trials.[43] Observers’ adaptive state was equated across the eight conditions by ensuring that an approximately equal number of trials was completed in each condition as the experiment proceeded and by keeping the room moderately lit to prevent dark adaptation. Each staircase was terminated after five reversals.

Acknowledgements

We thank C. Bundesen, C. Connor, V. Di Lollo, H. Egeth, J. Enns and M. Rudd for valuable discussions. This work was supported by grant R01-MH43924 from the U.S. National Institute of Mental Health.

RECEIVED 13 JULY: ACCEPTED 26 AUGUST 1998

1. Kanizsa, G. *Organization in Vision* (Prager, New York, 1979).
2. Kellman, P. J. & Shipley, T. F. A theory of visual interpolation in object perception. *Cogn. Psychol.* 23, 141–221 (1991).
3. Michotte, A., Thinès, G. & Crabbé, G. in *Michotte’s Experimental Phenomenology of Perception* (eds Thinès, G., Costall, A. & Butterworth, G.; trans Miles, E. & Miles, T. R.) 140–167 (Erlbaum, Hillsdale, New Jersey, 1991; originally published 1964)
4. Rock, I. *The Logic of Perception* (MIT Press, Cambridge, Massachusetts, 1983).
5. Shimojo, S. & Nakayama, K. Amodal representation of occluded surfaces: Role of invisible stimuli in apparent motion correspondence. *Perception* 19, 285–299 (1990).
6. von der Heydt, R., Peterhans, E. & Baumgartner, G. Illusory contours and cortical neuron responses. *Science* 224, 1260–1261 (1984).
7. Wertheimer, M. Experimentelle studien über das sehen von bewegung. *Z. Psychol.* 61, 161–265 (1912).
8. Kolers, P. A. *Aspects of Motion Perception* (Pergamon, New York, 1972).
9. Burt, P. & Sperling, G. Time, distance, and feature trade-offs in visual apparent motion. *Psychol. Rev.* 88, 171–195 (1981).
10. Ramachandran, V. S. & Anstis, S. M. Perceptual organization in moving patterns. *Nature* 304, 529– 531 (1983).
11. Albright, T. D. & Stoner, G. R. Visual motion perception. *Proc. Natl. Acad. Sci. USA* 92, 2433–2440 (1995).
12. Shepard, R. N. & Judd, S. A. Perceptual illusion of rotation of three-dimensional objects. *Science* 191, 952–954 (1976).
13. Robins, C. & Shepard, R. N. Spatio-temporal probing of apparent rotational movement. *Percept. Psychophys.* 22, 12–18 (1977).
14. Beck, J., Elsner, A. & Silverstein, C. Position uncertainty and the perception of apparent motion. *Percept. Psychophys.* 21, 33–38 (1977).
15. Farrell, J. E. & Shepard, R. N. Shape, orientation, and apparent rotational motion. *J. Exp. Psychol. Hum. Percept. Perform.* 7, 477–486 (1981).
16. Bundesen, C., Larsen, A. & Farrell, J. E. Visual apparent movement: Transformations of size and orientation. *Perception* 12, 549–558 (1983).
17. Hock, H. S., Kelso, J. A. S. & Schöner, G. Bistability and hysteresis in the organization of apparent motion patterns. *J. Exp. Psychol. Hum. Percept. Perform.* 19, 63–80 (1993).
18. Korte, A. Kinematoskopische untersuchungen. *Z. Psychol.* 72, 194–296 (1915).
19. Neuhaus, W. Experimentelle untersuchung der scheinbewegung. *Arch. Ges. Psychol.* 75, 315–458 (1930).
20. Cavanagh, P. Attention-based motion perception. *Science* 257, 1563–1565 (1992).
21. Yantis, S. Multielement visual tracking: Attention and perceptual organization. *Cogn. Psychol. 24*, 295–340 (1992).
22. Horowitz T. & Treisman, A. Attention and apparent motion. *Spatial Vis.* 8, 193–219 (1994).
23. Shepard, R. N. & Zare, S. L. Path-guided apparent motion. *Science* 220, 632–634 (1983).
24. Kolers, P. A. Some differences between real and apparent visual movement. *Vision Res.* 3, 191–206 (1963).
25. Attneave, F. & Block, G. Absence of masking in the path of apparent movement. *Percept. Psychophys.* 16, 205–207 (1974).
26. Mikami, A., Newsome, W. T. & Wurtz, R. H. Motion selectivity in macaque visual cortex. II. Spatiotemporal range of directional interactions in MT and V1. *J. Neurophysiol.* 55, 1328–1339 (1986).
27. Newsome, W. T., Mikami, A. & Wurtz, R. H. Motion selectivity in macaque visual cortex. III. Psychophysics and physiology of apparent motion. *J. Neurophysiol.* 55, 1340–1351 (1986).
28. Mikami, A. Direction selective neurons respond to short-range and long-range apparent motion stimuli in Macaque visual area MT. *Int. J. Neurosci.* 61, 101–112 (1991).
29. Zeki, S. & Shipp, S. The functional logic of cortical connections. *Nature*

335, 311–317 (1988).

30. Felleman, D. J. & Van Essen, D. C. Distributed hierarchical processing in the primate cerebral cortex. *Cereb. Cortex* 1, 1–47 (1991).
31. Sillito, A. M, Jones, H. E., Gerstein, G. L. & West, D. C. Feature-linked synchronization of thalamic relay cell firing induced by feedback from the visual cortex. *Nature* 369, 479–482 (1994).
32. Ramachandran, V. S. & Gregory, R. L. Perceptual filling in of artificially induced scotomas in human vision. *Nature* 350, 699–702 (1991).
33. De Weerd, P., Gattass, R., Desimone, R. & Ungerleider, L. G. Responses of cells in monkey visual cortex during perceptual filling-in of an artificial scotoma. *Nature* 377, 731–734 (1995).
34. Hogben, J. H. & Di Lollo, V. Suppression of visible persistence in apparent motion. *Percept. Psychophys.* 38, 450–460 (1985).
35. Watamaniuk, S. N. Visible persistence is reduced by fixed-trajectory motion but not by random motion. *Perception* 21, 791–802 (1992).
36. Van Essen, D. C. & Gallant, J. L. Neural mechanisms of form and motion processing in the primate visual system. *Neuron* 13, 1–10 (1994).
37. Breitmeyer, B., Love, R. & Wepman, B. Contour suppression during stroboscopic motion and metacontrast. *Vision Res.* 14, 1451–1456 (1974).
38. Breitmeyer, B. G. *Visual Masking* (Oxford Univ. Press, New York, 1984).
39. Grossberg, S. & Rudd, M. E. Cortical dynamics of visual motion perception: Short-range and long-range apparent motion. *Psychol. Rev.* 99, 78–121 (1992).
40. Baloch, A. A. & Grossberg, S. A neural model of high-level motion processing: line motion and formotion dynamics. *Vision Res.* 37, 3037–3059 (1997).
41. Finkel, L. H. & Edelman, G. M. Integration of distributed cortical systems by reentry: a computer simulation of interactive functionally segregated visual areas. *J. Neurosci.* 9, 3188–3208 (1989).
42. Tononi, G., Sporns, O. & Edelman, G. M. Reentry and the problem of integrating multiple cortical areas: simulation of dynamic integration in the visual system. *Cereb. Cortex* 2, 310–335 (1992).
43. Levitt, H. Transformed up-down methods in psychoacoustics. *J. Acoust. Soc. Am.* 49, 467–477 (1971).