

# Illusory motion from change over time in the response to contrast and luminance

Benjamin T. Backus

Department of Psychology, University of Pennsylvania,  
Philadelphia, PA, USA



İpek Oruç

Department of Psychology, University of Pennsylvania,  
Philadelphia, PA, USA



A striking illusion of motion is generated by static repeated asymmetric patterns (RAPs) such as Kitaoka's (2003) "Rotating Snakes" and Fraser and Wilcox's (1979) peripheral drift illusion. How do RAPs generate spurious motion signals, and what critical difference between RAPs and natural static scenes prevents the latter from appearing to move? Small involuntary eye movements during fixation have been suspected to play a critical role in these illusions, but here we give an account that does not depend on fixation jitter. We propose that these illusions result primarily from fast and slow changes over time in the neuronal representation of contrast ("contrast-driven RAPs") or luminance ("luminance-driven RAPs"). We show that temporal phase advance in the neural response at high contrast can account for the early, fast motion in contrast-driven RAPs (such as "Rotating Snakes") after each fixation change. An essential part of this explanation is that motion detectors fail to compensate for the dynamics of neuronal encoding. We argue that static natural patterns also generate local gain changes, but that these signals do not often trigger illusory motion because they are not usually aligned to drive global motion detectors. Movies in which real luminance changes over time, to mimic the proposed neuronal adaptations to contrast and luminance, evoke qualitatively similar percepts of motion. Experimental data are consistent with the explanation. Color and overall contrast both enhance the illusion.

Keywords: contrast adaptation, global motion, illusory motion, luminance adaptation, peripheral drift illusion, Rotating Snakes

## Introduction

Repeated asymmetric patterns (RAPs) cause many peoples' visual systems to infer the presence of motion where there is none. "Rotating Snakes" (Kitaoka, 2003) (Figure 1) and Judy Chicago's "Through the Flower" (Chicago, 1973) (Auxiliary Figure 1) are examples: Most people see a rotary movement that runs in the black-blue-white-yellow direction for "Rotating Snakes" and in the gradual dark-to-light direction for repeated single gradients (Auxiliary Figures 1 and 2). The cause of these striking illusions has remained mysterious since the peripheral drift illusion was described a quarter century ago (Fraser & Wilcox, 1979).

One starting point to explain an illusion is Helmholtz's assertion that "*such objects are always imagined as being present in the field of vision as would have to be there in order to produce the same impression on the nervous mechanism [italics original]*" (Helmholtz, 1925). If we take "impression on the nervous mechanism" to mean not only neural activity that occurs at transduction but also activity that occurs after some processing, we might restate this principle by asserting that *the visual system constructs percepts that represent that which would most likely have evoked the same pattern of sensory neural*

*activity*, where sensory neural activity could include several postretinal processing stages. Thus, one can explain the illusion by showing how "Rotating Snakes" and real motion are expected to evoke similar activity in the neurons (such as cells in the lateral geniculate nucleus, LGN) that innervate direction-selective neurons (such as V1 cells).

The paper is organized as follows. We start with some informal observations about the illusion and discuss their implications. We then give our "high-contrast phase-advance" explanation of contrast-driven motion for "Rotating Snakes" and discuss how our model-based approach differs from that of Conway et al. (2005), which is based on a similar idea. We point out the need for a separate account of luminance-driven motion to explain the peripheral drift illusion, and identify, qualitatively, a single compound adaptation function that could drive both the contrast- and luminance-based illusions. We then describe experiments that measured the strength of the illusion and consider implications of the data from the experiments. In the Discussion section, we consider why some people see more illusory motion than others, why there might be few costs to building a global motion system that fails to compensate for the dynamics of neural coding, and consider how use of the neural code for contrast must differ between global motion perception, relative motion perception, and pattern perception.

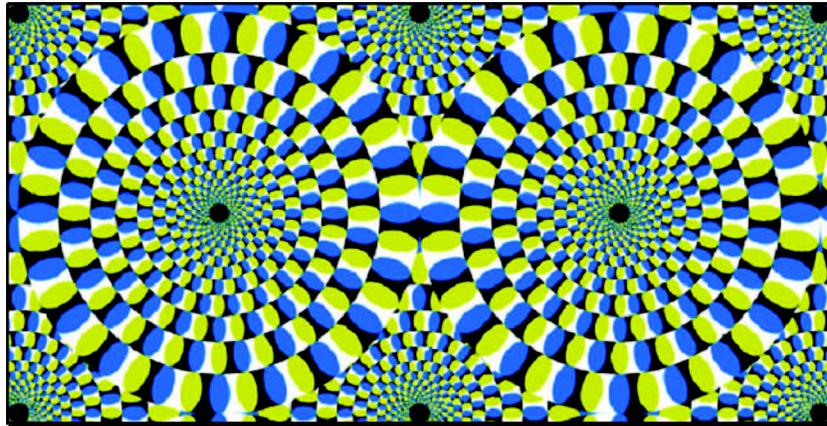


Figure 1. A part of Kitaoka's "Rotating Snakes" illusion. Most people see clockwise rotation in the right disk, especially when they fixate elsewhere.

## Some initial observations about RAP illusory motion

Some properties that may be important to capture in a complete model of RAP illusory motion are as follows: (1) the direction of motion depends on the order of the colors within the RAP; (2) motion stops after about 6–8 s of steady fixation; (3) when re-fixating a previously fixated point, the time required for motion to stop increases with time away from the point; (4) the eyes need not move to refresh the motion (it suffices to move the pattern, see [Auxiliary Videos 1 and 2](#)); (5) motion is restarted by moderate eye movements, but not by very small ones nor by small amounts of image jitter ([Auxiliary Video 3](#)); (6) the motion can be speeded, stopped, or reversed by preadaptation to specific high-contrast patterns ([Auxiliary Video 4](#)); (7) motion stoppage after monocular viewing does not completely transfer to the other eye ([Auxiliary Figure 3](#)); (8) the motion is more pronounced for binocular than monocular viewing; (9) RAPs evoke a negative motion adaptation aftereffect (Ashida & Kitaoka, 2003); (10) different people see different speeds, and for some RAPs, different directions of motion (Fraser & Wilcox, 1979; Naor-Raz & Sekuler, 2000) and the individual differences are to some extent heritable (Fraser & Wilcox, 1979); (11) the motion falls off rapidly with contrast (Naor-Raz & Sekuler, 2000, [Auxiliary Figure 4](#)); (12) the motion can be enhanced by color for some observers ([Auxiliary Figure 5](#)); (13) the motion is most compelling when the repeated elements of the RAP are configured such that individual local motions, as might be generated by each element, contribute to the same motion within a large image region ([Auxiliary Figure 6](#)); (14) motion in crisp images is most compelling in noncentral vision (Fraser & Wilcox, 1979; Faubert & Herbert, 1999, but see [Auxiliary Figure 7](#); Naor-Raz & Sekuler, 2000);

(15) blur reduces the motion of "Rotating Snakes" in noncentral vision while increasing it in central vision ([Auxiliary Figure 8](#)); and (16) motion is visible in printed RAPs in sunlight.

Property 2 suggested to us that adaptation of some sort, reaching asymptote in about 6 s, drives the illusion, and Property 3 is consistent with recovery from adaptation. Properties 4–6 imply that the adaptation occurs largely within a retinotopic representation of the image rather than reflecting adaptation to regional contrast levels (Chubb, Sperling, & Solomon, 1989) *per se*. Property 7 implicates an early locus for the adaptation, prior to the loss of separate representations for each eye in cortex. We have no explanation for Property 8 at the level of neural mechanisms, but computationally additional sense data from any source would provide the system with additional evidence for motion; this would trade against the no-motion prior probability (Weiss, Simoncelli, & Adelson, 2002). Some of the slowing during fixation is explained by Property 9, but motion aftereffects cannot explain why the motion stops completely after 6–8 s because real motion (matched for apparent speed) appeared to move indefinitely (informal observations). Property 10 implicates a biological difference in the adapting mechanism(s) between individuals. Property 11 suggests that nonlinearity in the neural response to contrast, for example, saturation or faster responses at high contrast, plays a role. Property 12 suggests that either color-sensitive motion mechanisms (Hawken, Gegenfurtner, & Tang, 1994) play a role or color affects the neural representation of achromatic contrast used by motion detectors, for example, by affecting contrast gain control. Properties 13 and 14 suggest that RAPs are particularly effective at driving the global motion system (Cavanagh & Favreau, 1980; Nakayama & Tyler, 1981; Williams & Sekuler, 1984). Property 15 suggests that the amount of perceived motion depends on multiple processes operating

at different spatial scales (see also [Discussion](#)). Many people first encountered “Rotating Snakes” on a computer display; Property 16 shows that pulsation at 60–80 Hz is not necessary.

## Motion model

At the heart of the model ([Figure 2](#)) is our assumption that an observer will see global motion, such as real rotation when a wheel turns or the illusory rotation of a disk in “Rotating Snakes”, when a global motion detector is activated by an appropriate set of local velocity detectors that comprise its subunits (Bex, Metha, & Makous, 1998; Williams & Sekuler, 1984). An alternative formulation of the model is possible using model V1 cells (Heeger, 1993) as subunits, but because the supposed contrast and luminance adaptations occur before this stage, and because predictions for the speed of the illusory motion are therefore similar in both cases, we have chosen to use velocity detectors in the model for convenience of exposition. These local velocity mechanisms are tuned to spatial frequency. Thus, they report the speed and direction of sinusoidal grating components (or gabor) derived from the image. We assume that the velocity detector can report rate of change in the phase of the component over time, independent of any changes in overall contrast that may also be occurring.

In the model, velocity is estimated not from the optical pattern of contrast itself, but rather from a “neural image” of contrast, in which high-contrast points in the optical image are registered before low-contrast points (in neurons this is called phase advance; Albrecht, Geisler, Frazor, & Crane, 2002; Georgeson, 1987; Shapley & Victor, 1978; Tolhurst, Walker, Thompson, & Dean, 1980). As a result of this choice, the ratio of neural responses in the high- and low-contrast areas of the neural image changes over time, and this gives rise to a shift in the phase of the sinusoid (or gabor) that best fits the pattern. It is this shift that is monitored by the velocity detectors and reported to global motion detectors.

The model’s neural image of contrast (output of Step 2 in [Figure 2](#)) is created by assuming that the time course of the response at each point in the neural image is proportional to the instantaneous mean firing rate of the average neuron in macaque primary visual cortex as measured by Albrecht, Geisler, Frazor, and Crane (2002). That study measured the responses of cortical neurons to abrupt-onset static gratings of various contrasts, at the neurons’ preferred spatial frequencies and orientations, out to 200 ms after onset. (Measurements out to several seconds would be needed to fully test the model, but these were not available, so we have used the model to work backwards to infer the neural activity

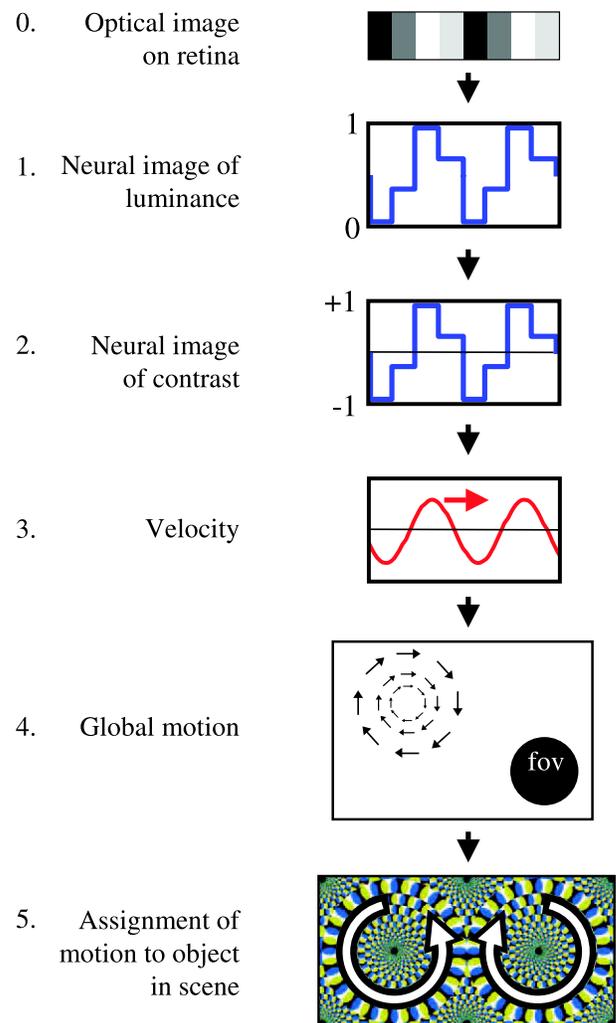


Figure 2. Model of illusory motion from RAPs. The optical image of a small piece (2 cycles) of a RAP are shown as Step 0. This is registered as a neural image of luminance (Step 1) and converted to a neural image of contrast (Step 2). Changes in either of these representations may be detected by local velocity detectors (Step 3), which are selective for spatial frequency. Step 4 depicts local velocities within the visual field for one disk in the “Rotating Snakes” illusion; “fov” depicts the fovea. Steps 0–4 occur in a retinotopic coordinate system. Step 5 depicts the perceived motion, after the global rotation motion has been tied perceptually to the pattern and attributed to a location in the world.

from the perceived motion.) The data of Albrecht et al. (2002) are neuronal firing rates, so they encode local contrast (across space) as a positive number. This can be thought of as activity in a luminance-balanced filter. To use these numbers in the model’s pointwise neural image of contrast, we restore the sign (positive or negative) to the local contrast depending on whether the optical image is above or below mean luminance at that point.

It may seem odd to use data from orientation and spatial-frequency-selective cells to model pointwise

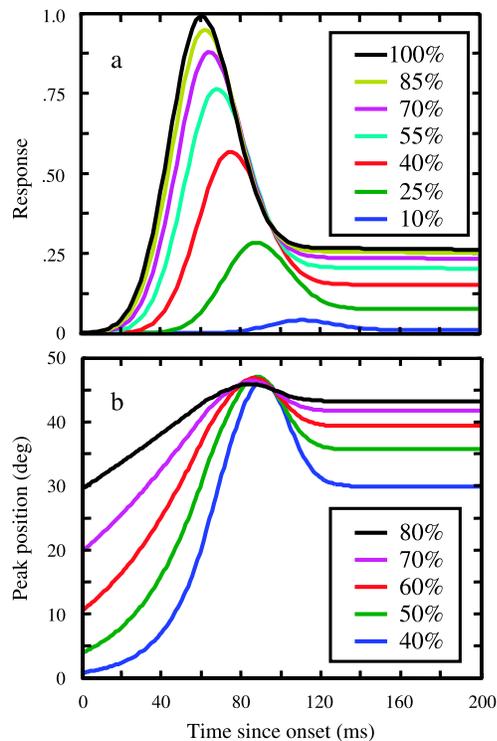


Figure 3. Model response to contrast, and resulting changes over time in the neural image. (a) Pointwise responses to contrast in the neural image are shown as a function of time since pattern onset. The seven curves show different levels of contrast. The curves are simply scaled versions of the response of the average cortical neuron, as measured by Albrecht et al. (2002). (b) Model result for spatial phase (peak position) of the sinusoid of fundamental spatial frequency, in the neural image, as a function of time since pattern onset, for a pattern that repeats black, dark gray, white, light gray (as in Figure 2, Step 0). The black and white regions are assumed to have a contrast of 1.0 (i.e., signed contrasts of  $-1.0$  and  $1.0$ ). The different curves show model responses for different (shared) contrast in the dark and light gray regions, from 0.4 to 0.8. The ordinate shows the spatial phase in degrees within a single 360 deg cycle of the RAP.

responses at the neural image stage of our model, given that the model's neural image more closely resembles the representation of contrast in the retina or LGN. We made this choice because the model parameters estimated by Albrecht et al. (2002) are particularly convenient to use, and because it seems plausible to us that the inputs to cortical motion mechanisms are described by similar response functions. In any case, a substantial part of the low-contrast delay and saturation in the contrast responses of cortical neurons is in fact inherited from neurons earlier in the visual pathway (Carandini, Heeger, & Movshon, 1999; Shapley & Victor, 1978). The model is simple relative to known physiology and neural responses to RAP stimuli have not yet been measured, so we cannot do better than to demonstrate the proof of

principle using a plausible approximation for the model's neural image.

The model predicts illusory motion in RAPs because low-contrast points in the optical image are registered later within the neural image than high-contrast points. This feature of the data of Albrecht et al. (2002) is illustrated in Figure 3. For our model, it is critical that this delay arises initially within units that operate on a smaller spatial scale than the spatial frequency to which the velocity detectors are tuned because after the image is filtered at the spatial frequency of the velocity detector it becomes a repeated symmetric pattern (namely, a cosinusoid) to the detector. If the low-contrast delay occurred after spatial frequency filtering, there would be no basis for a balanced velocity mechanism to infer a change in the phase of the sinusoid over time. In our model, we avoid this by using a pointwise representation of contrast for the neural image, but a bank of units with small circle-symmetric receptive fields could work equally well.

Additional stages of processing at the front and back end are necessary to convert an optical (luminance) image into perceived rotation (Steps 1 and 3–5 in Figure 2). To explain illusory motion in single-gradient RAPs, we will need to describe changes over time in the neural image for luminance, so it is included as a separate stage (Step 1). The neural image of luminance is not needed to explain contrast-driven RAP illusory motion (such as in "Rotating Snakes"), and we do not have neuronal data to separately estimate the changes that occur over time in the luminance image (unlike the contrast image, for which we use the data of Albrecht et al., 2002). In other words, Step 1 is included because it is part of the big picture for understanding RAP illusory motion in general.

Conceptually, we think of the neural image of luminance as being transformed into the neural image of contrast by subtracting mean activity and renormalizing (Step 2). Velocity is then extracted locally throughout the image in Step 3. After this, global motion is detected (Step 4), and then a final process attributes retinotopic global motion to the motion of an object in the world.

To recapitulate, the data of Albrecht et al. (2002) captured only the first 200 ms of the response to static contrast patterns. To explain why the illusion lasts an additional 6 s, we must infer that neurons continue to adapt in a manner that causes the local peaks in activity in the neural image to slowly drift during this time. Slow adaptation by luminance and contrast gain controls (Brown & Masland, 2001; Müller et al., 1999) might effect such a drift. We do not ascribe a functional significance to the mechanism that is responsible for the drift, given that saccades normally occur several times per second, and that some people with good vision do not see illusory motion in RAPs.

Based on the assumption that change in the neural image of contrast is responsible for the illusory motion in RAPs, we can categorize RAPs into two classes: those that appear to move due to change in the neural representation of contrast per se, and those that appear to move due to change in the neural representation of luminance (and hence also contrast).

### Contrast-based RAP illusions

Figure 3a shows the response (firing rate) of the average primate V1 neuron to an abrupt-onset stimulus as a function of time, based on the fitted data from the invariant response descriptive model of Albrecht et al. (2002). The different curves show responses at seven different levels of contrast. When these curves are used to describe the neural image's pointwise response to the static RAP in Step 0 of Figure 2, the sinusoidal grating that best fits the neural image drifts over time. Figure 3b shows the position of the peak in degrees (where 360 deg equals one cycle of the RAP). The different curves show model responses at five different values of contrast for the light and dark gray regions of the RAP; the other regions were black and white (contrast = 100%) for every curve.

The model predicts that the RAP will appear to rotate by 15–45 deg (1/24 to 1/8 RAP cycle) in the first 90 ms. This rapid rotation is followed by a small reversal and subsequent stabilization between 90 and 120 ms. The integration time for luminance changes has been estimated at 100–200 ms (Rashbass, 1970; Watson, 1979) and at threshold global motion integration can exceed 2 s in human (Britten et al., 1992). Because these integrations are not in the model, we would not expect perceived global motion to track the rapid fluctuation of position shown in Figure 3b at 60–120 ms. The model still predicts early fast perceived rotation in the correct direction if integration is taken into account.

What are we to make of the fact that RAPs continue their illusory rotation for 6 s? When viewing a RAP, the firing rates of neurons that code for patches of the RAP should be in ratios that change over time as needed to explain the motion, as for real first-order motion (Thompson, 1982). Albrecht et al. (2002) used a broad contrast-invariant half-gaussian to fit the slow decay in their neurons' tonic firing after the phasic response, but given how slowly the tonic rates decay, and that Albrecht et al. recorded for only 200 ms, we might work in reverse, using the model to infer what form these tails actually take. One possibility is exponential decay with a time constant that depends on contrast. One could add a gain control step to the model to normalize spatial contrast within the neural image over time, which would correspond in cortical neurons to keeping mean neuronal firing rates above some minimum to maintain the representa-

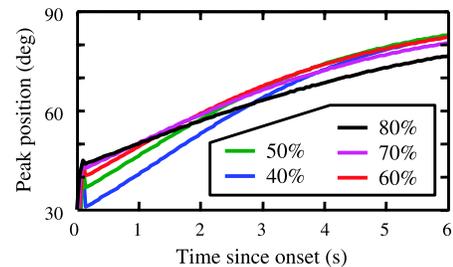


Figure 4. Response of the model out to 6 s, when response to contrast in the neural image is based on a hypothetical contrast-dependent exponential decay in the sustained portion of cortical neurons' response to contrast. The format is the same as Figure 3b. See text.

tion. For simplicity, we omit this step and compute changes in spatial phase over time as if there were no noise.

Figure 4 shows the consequence of making the contrast-dependent decay assumption, using time constants that vary linearly with contrast from 6.0 s at 0% contrast to 1.5 s at 100% contrast. The format is the same as Figure 3b, except that the abscissa now extends to 6 s. During this time, the RAP can rotate nearly 1/4 of a RAP cycle, depending on the contrast of the light/dark gray regions. This shows it is possible for the model to accommodate the gradually slowing motion in RAPs. The model makes a prediction that could be tested experimentally, namely, that in an experiment like that of Albrecht et al. (2002), the ratio between the neural responses to low and high contrasts should increase for several seconds. This might even derive from a response to low contrast that eventually exceeds the response to high contrast, a situation that also occurs for some neurons during parts of the first 200 ms of their response (Figure 1 in Albrecht et al., 2002).

Conway et al. (2005) recently published an account of the illusory motion in “Rotating Snakes” that has, in common with our explanation, an important role for high-contrast phase advance. They measured the response of macaque V1 and MT cells to flashed bars and found that peak responses occurred 10–20 ms earlier for white and black bars than for light gray and dark gray bars, respectively. They also found that direction-selective cells responded to simultaneously presented bars when one bar had higher contrast than the other; furthermore, many of these cells responded when one bar was dark and the other was light, if they differed in contrast magnitude, which would be consistent with a motion contribution from the reverse-phi phenomenon (Anstis, 1970).

Like Conway et al. (2005), we rely on differential time courses in the response to high and low contrasts, prior to motion detection, to explain the illusory motion in “Rotating Snakes” (Backus & Oruç, 2004). In our explanation here, we supposed that the motion is measured by mechanisms tuned to spatial frequency,

which enabled us to model the magnitude of the illusion and to give an account as to why the illusory motion lasts for many seconds during a single fixation.

## Luminance-based RAP illusions

We now have seen that, for a RAP like the grayscale “Rotating Snakes” (Figure 2), a pointwise adaptation whose rate depends on contrast has the effect of expanding the dynamic range devoted to low contrasts relative to high over time, which can account for the illusion of motion. But this account cannot explain the motion in all RAPs. In particular, it cannot explain the motion of single-gradient RAPs such as Auxiliary Figure 2 or the escalator illusion (Fraser & Wilcox, 1979). The theory does not work: A compressive nonlinearity applied to contrast over time does not cause any phase shift in the best-fitting sinusoid for these patterns.

Secondly, single-gradient RAPs usually move in the gradual dark-to-light direction, whether that gradient is from black to white, black to gray, or gray to white. For example, the “dual-gradient” RAP in Figure 7b (left side) appears to rotate counterclockwise. It also rotates counterclockwise if the white-to-gray gradients are all replaced by uniform gray. But it changes direction—and appears to rotate clockwise—if the black-to-gray gradients are all replaced by uniform gray (see also Auxiliary Figures 2–3 and 9–10). The model as we have developed it up to this point does not predict this. Instead, the model would predict that the black-to-gray and white-to-gray gradients appear move in the same direction.

Finally, many people experience a qualitative difference in the motion for “Rotating Snakes” and dual-gradient RAPs, as compared to single-gradient RAPs. The former typically give rise to rapid motion that starts immediately upon each refixation; the latter often start moving more gradually. This suggests that different processes of adaptation may be driving motion in the two illusions.

How would the internal representation of contrast have to change over time to account for these new facts? Within the framework of the model, the single-gradient illusions imply that mid-level grays come to be represented as darker over time. This is the change that would cause the best-fitting sinusoid to shift in the dark-to-light direction. S. Anstis and M. Becker (personal communication, 2005) have found an illusion of motion in large single-gradient (nonrepeating) patterns, consistent with this effect. They report that the direction of perceived motion reverses for very high luminance displays, which further implicates adaptation to luminance (occurring prior to motion measurement) as a factor that contributes to these illusions.

Figure 5a shows how an appropriate compound adaptation could account for both types of illusion.

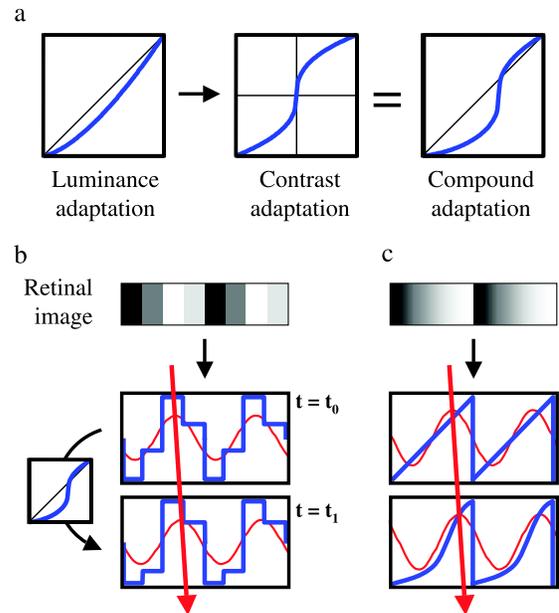


Figure 5. Luminance/contrast adaptation account of illusory motion for two RAPs. (a) The shape of a compound adapting function that accounts qualitatively for slow illusory motion in RAPs is approximated by separate adaptations to luminance and contrast. Input is from 0 (black) to 1 (white) on the x-axis, and output at asymptote (after 6 s) is from 0 to 1 on the y-axis. (b and c) The retinal images of two RAPs are transformed onto a normalized internal representation of contrast (thick blue lines) in the manner of Figure 2, and adaptation over the course of 6 s causes the representation to change (bottom). There is rightward motion at the fundamental spatial frequency (thin red curve) in both cases.

The first adaptation is an expansive nonlinearity (over time) for luminance.<sup>1</sup> This is followed in series by a compressive nonlinearity for contrast. A useful level of abstraction is obtained by supposing an initial (monotonic) mapping from luminance in the retinal image onto a neural image of luminance, represented on a scale from 0 (*black or minimum*) to 1 (*white or maximum*). We conceive of adaptation for luminance as occurring within this representation. The neural image of contrast is then computed from the neural image of luminance. If we suppose that contrast is continually normalized to fill the range 0–1 (Lu & Sperling, 1996; Snippe, Poot, & van Hateren, 2000; Albrecht et al., 2002), we can represent both adaptations with the single transformation shown in Figure 5a.

Accordingly, we can classify RAPs into two groups: those in which the illusion of motion is driven primarily by adaptation to contrast, and those in which it is driven primarily by adaptation to luminance. Figures 5b and c show the effect of applying the compound adaptation to “Rotating Snakes” (contrast driven) and to the single-gradient illusion (luminance driven), respectively. Further psychophysical experiments may be

able to establish whether there really are dissociable effects in the illusion that are due to adaptations to luminance and contrast, for example, by measuring their time courses.

## Movies with real changes in luminance

We use a simple model because it is nontrivial to build a realistic model of perceived speed. The perceived speed of real motion sometimes does depend on contrast (Hurlimann, Kiper, & Carandini, 2002; Stone & Thompson, 1992), has a complicated relationship with perceived position (Gregory & Heard, 1983; Snowden, 1998), depends on various gain control mechanisms (e.g., Lu & Sperling, 1996), and to calculate it the weight of the zero motion prior must be estimated (Weiss, Simoncelli, & Adelson, 2002). A better model than ours would also incorporate some scheme for weighting motion signals at different spatial frequencies rather than looking only at the fundamental frequency of the pattern. However, accepting the basic framework does allow us to make a testable prediction, which is that movies in which real luminance ratios change over time ought to appear to move similarly to RAPs, a prediction that was pointed out to us by Arthur Shapiro (Shapiro et al., 2004).

Auxiliary Videos 5–12 illustrate these phenomena. Videos 5 and 6 show how expansive (as in the model) or compressive adaptation to luminance causes a single-gradient illusion to move. Video 7 shows that expansive luminance adaptation has little effect on the motion of dual-gradient RAP. Videos 8 and 9 show how compressive (realistic) or expansive (backwards) adaptation to contrast causes a dual-gradient illusion to move. Video 10 shows that compressive contrast adaptation has little effect on the motion of a single-gradient RAP. Videos 11 and 12 show contrast compression and expansion in alternation in the grayscale and full color “Rotating Snakes” images, respectively.

These manipulations of luminance cause compelling illusions of motion that are perceptually similar to (and, we suppose, indistinguishable from) the perceived motion in RAPs. They show that if the neural representation of contrast changes in the right way over time, and the motion system fails to compensate, illusory motion of the sort seen in “Rotating Snakes” would follow.

## Experimental measurement of illusion strength

We quantified the strength of the illusory motion in several RAPs by asking observers to compare the apparent

speeds of real and illusory motion stimuli. The resulting data were used to confirm that the magnitude and time course of the illusion are consistent with the model, and to quantify how the illusion is enhanced by color and weakened at low contrast. Figure 6 gives a hypothetical time course of the neural image for the stimulus we used in most of the experiments, which was a dual-gradient RAP. This pattern evokes the same fast initial rotation seen in the four-value “Rotating Snakes” pattern and would be driven primarily by contrast rather than luminance adaptation.

## Methods

To measure the effect of display duration on the magnitude of illusory motion, we used an extended set of RAPs based on the “dual-gradient” illusion (Ashida & Kitaoka, 2003). Figures 7a–b show the stimuli and

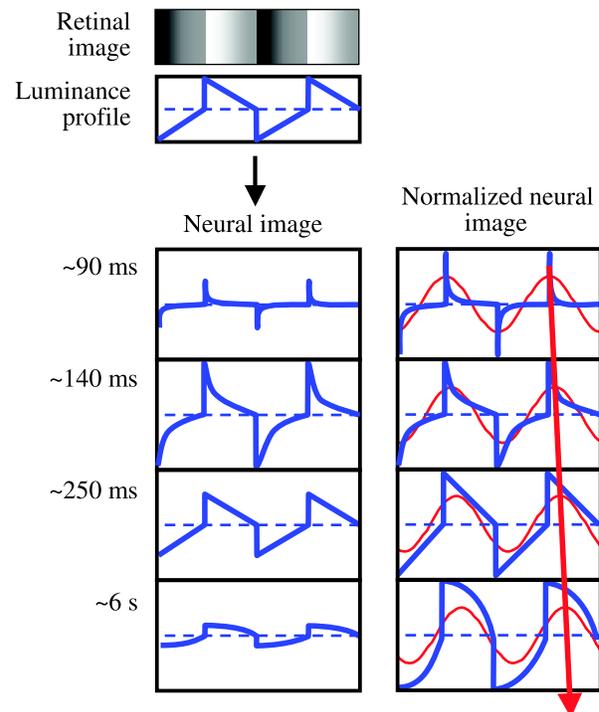


Figure 6. Contrast–response account of illusory motion for the dual-gradient RAP. Two cycles of the RAP are shown at top, with a graph of their luminance profile. Below that are cartoons showing the internal representation of the stimulus, as it might appear to a mechanism that does not take the transient nature of neuronal responses into account. The internal representation is shown for four times after stimulus onset (blue curves). Fast registration at the high-contrast edges is followed by a slower registration of lower contrast regions and then adaptation towards baseline firing rates. At right, the internal representations have been normalized. A sinusoid fit to this pattern at its fundamental frequency (red curve) moves rightward over time (red arrow).

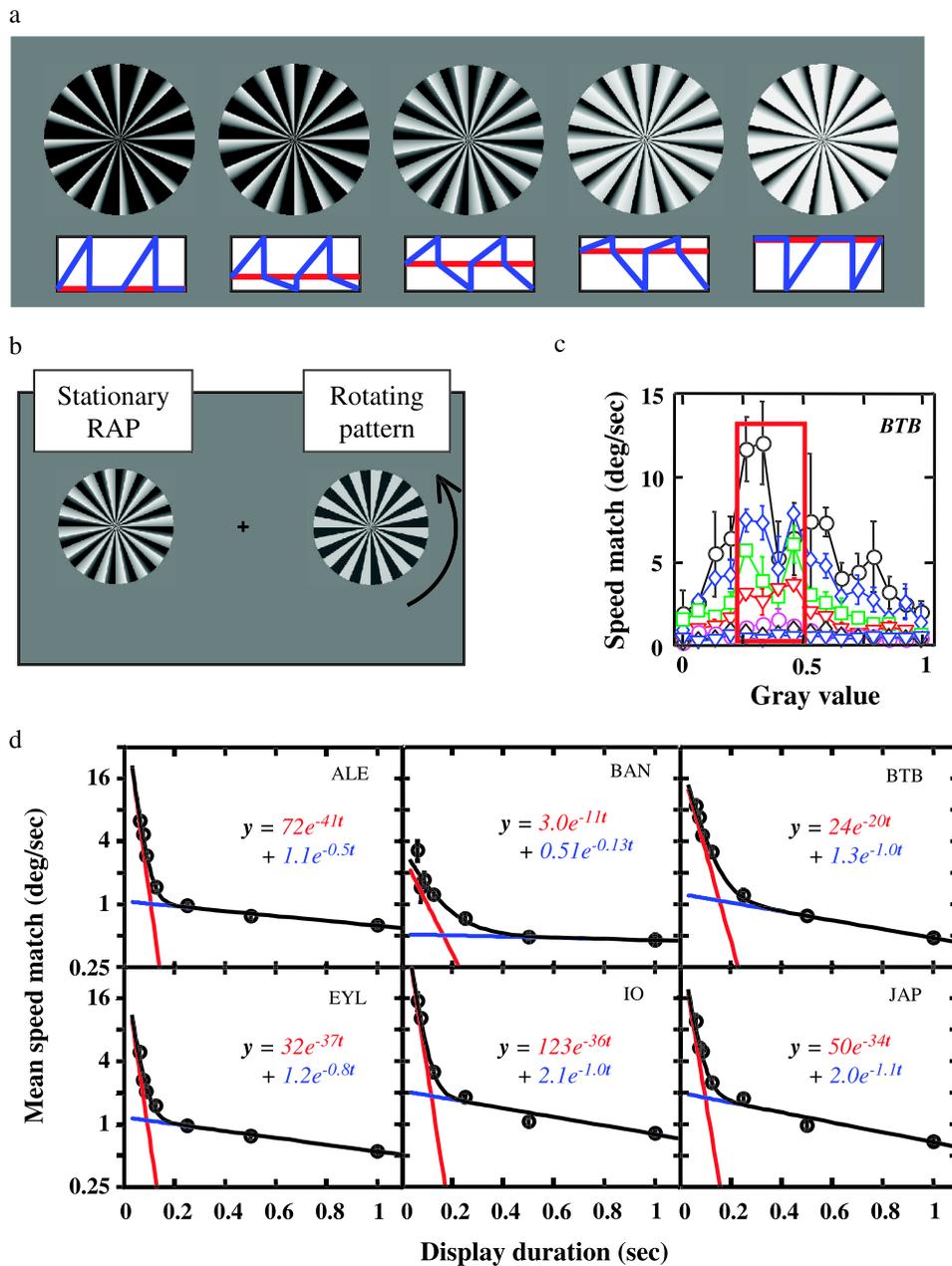


Figure 7. Speed matching experiment. (a) Dual-gradient stimuli. Each RAP repeats gray-to-white and gray-to-black gradients. They differ according to the value of the gradients' common gray endpoint. Luminance profiles are plotted below the RAPs. (b) Depiction of stimuli. A gray screen with fixation mark was followed by the stimulus (fixation mark, RAP, and real rotation), followed by the gray screen. The observer indicated which side of fixation contained faster rotation. (c) Matching speeds for one observer as a function of gray level in the RAP. The series are data for display durations of (top to bottom) 3, 4, 5, 7, 15, 30, and 60 video frames, using a DLP projector running at 60 Hz. The red rectangle shows which data were used to generate the speed versus duration graphs of panel d. (d) Mean speed matches as a function of display duration for six observers. Data are fit by the sum (black curve) of two exponentials (red and blue curves) by minimizing variance-weighted squared error.

experimental paradigm. The observer fixated the plus sign. On each trial, two stimuli appeared, assigned randomly to either side of fixation: one was a stationary dual-gradient RAP and the other was a nonillusory pattern that actually rotated. The observer indicated which side appeared to rotate faster, and a psychophysical staircase

procedure measured how much real rotation was needed to match the apparent speed of the RAP.

Stimuli were constructed using Matlab software and the experiment was controlled using Matlab and the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997) on a PC. Stimuli were shown using a DLP projector

( $1024 \times 768$  pixels) that was mounted behind and above the observer, who sat 2 m from the display screen and used a numeric keypad for responses. The testing room was otherwise dark. The entire image subtended 31 deg wide  $\times$  23 deg tall. The two disks within the display subtended 7.3 deg and were centered 16.5 deg on either side of the central fixation mark. Each disk contained 15 identical repeated wedge patterns, each with its vertex at the center and covering 24 deg of the disk (12 deg each for the black-to-gray and white-to-gray gradients, or, for real motion, 12 deg each for the light and dark wedges). Images were gamma-corrected to linear luminance, and their mean luminance was  $155 \text{ cd/m}^2$ . The illusory-motion and real-motion stimuli had Michelson contrasts of 1.0 and 0.41, respectively.

The two stimuli were always arranged so as to appear to rotate in opposite directions (so the edges closest to fixation both appeared to go up or down) and the direction of rotation was alternated from trial to trial to minimize motion adaptation. The dual-gradient RAP stimulus and real-rotation stimulus were displayed for the same amount of time on any given trial. Stimuli were presented in blocks of constant display duration. Two interleaved QUEST staircase procedures (Watson & Pelli, 1983) were used to estimate the amount of real rotation that perceptually matched illusory rotation in the RAP based on 30 speed judgments per RAP at a given duration.

## Results

Speed matching data for one observer are plotted in [Figure 7c](#): The abscissa is the gray value at which the white-gray and black-gray gradients terminated within the RAP, and the data series are different display durations. To measure the effect of display duration, data from the eight patterns with four darker-than-average gray levels were averaged at each display duration, and these are plotted in [Figure 7d](#) for six observers. A single exponential or power function does not fit these data, but they are well fit by the sum of two exponentials that can be associated with a fast “kick-start” component lasting about 250 ms and a separate slow component that lasts several seconds, respectively (the slow component may be dissociable from the fast component: see [Auxiliary Video 1](#)). Although the magnitude of the illusion differed across observers (as measured by the perceptually matched real motion), all six of our observers saw illusory motion in this experiment.

Two points are worth mentioning in connection with our analysis of these data. First, the fact that two exponentials are needed to fit observers’ data makes it natural to describe the adaptation processes as occurring in two phases. This is not an artifact of simple temporal

integration within the observer: There is no combination of integration window size and a single exponential decay curve that can give a reasonable fit to the data for any of our observers. Second, this experimental design was unsuited to measuring very slow perceived motions because observers simply indicated which side of the display had greater motion, without indicating the direction of the motion. A consequence is that our procedure overestimated the illusory motion when it was very small (i.e., close to threshold). The data graphed in [Figure 7d](#) are not affected by this potential artifact; however, [Figure 7c](#) illustrates that for the extreme RAPs, our procedure produced matched speeds that were small compared to the RAPs used in [Figure 7d](#).

Two separate experiments confirmed that low-contrast RAPs are less effective than high-contrast RAPs at evoking illusory motion, and that color has an enhancing effect on the illusory motion for some observers of the “Rotating Snakes” illusion ([Auxiliary Figures 4 and 5](#), respectively).

## Quantitative check of model plausibility

In most RAPs, high- and low-contrast components are spaced a quarter cycle apart. This puts an upper bound on the total motion the model is capable of generating during a single fixation. Yet observers experience a great deal of motion during a single fixation of “Rotating Snakes”, or when viewing our dual-gradient stimuli. Is it simply that slow rotation of a large object in peripheral vision is particularly salient? Or it could be that the total perceived rotation is so large as to pose a problem for the model? A dual-gradient pattern might conceivably generate more than one quarter cycle of motion, if the best-fitting peaks move more than half way from black and white to gray and gray (see middle pattern in [Figure 7a](#)), but the perceived motion of the RAP still could not exceed one-half period of rotation.

The total real motion required to match the illusory motion can be estimated by integrating matched speed from  $t = 0$  to infinity. Using the fitted double exponentials in [Figure 7d](#) to approximate matched speed, this integration gives total rotary motions of 4.0, 4.2, 2.5, 2.4, 5.5, and 3.3 deg for the six observers in our main experiment, respectively. These values range from 10% to 23% of the 24 deg RAP period of the stimulus. We conclude that the model is in fact capable of explaining total motion in our experiment. The total motion seen during a fixation of the RAP, as measured by matching to real motion, is actually quite small. This highlights one of the lovely features of “Rotating Snakes”: its exploitation of object identity across fixations, as the same disk is seen to rotate again and again every time the disk moves onto a new part of the visual field.

## Discussion

### Motion energy, known physiology, and the model

Which anatomical sites should be identified with parts of the model? One could identify photoreceptors with the model's neural image of luminance; other retinal cells for which the surround only partially balances center with both the luminance and contrast images; LGN cells with the contrast image; direction-selective cells in cortical areas V1, V2, and MT with velocity extraction; and neurons in MT and MST (or V5/MT+ in humans) with global motion detection. We have no suggestion as to where one might find the neurons responsible for tying retinotopic global motion across saccades to fixed locations in the world, but clearly they must exist because one can watch the same disk rotating ad infinitum as one repositions one's eyes over a RAP.

At the heart of the model is the estimation of local velocities from the neural image. We described this as feature tracking. The model could estimate velocity using low-level (or "short-range" or "first-order") motion energy mechanisms instead (Braddick, 1974) (Reichardt, 1957; van Santen & Sperling, 1984; Adelson & Bergen, 1985; Watson & Ahumada, 1985). Velocity detectors built from motion energy units are insensitive to changes in overall contrast because computing velocity is equivalent to finding the orientation of the best-fitting plane through the origin in wavelet-transformed  $x$ - $y$ - $t$  frequency space (Adelson & Bergen, 1985; Heeger, 1987; Grzywacz & Yuille, 1991). This orientation is independent of overall contrast. Thus, a velocity mechanism based on motion energy would in principle extract velocity correctly from a fading neural image, so it would also generate illusory motion from high-contrast phase advance.

Motion mechanisms do not, however, literally extract velocity from a fading neural image of contrast. An obvious deficiency in this account is that the phasic responses of early visual neurons do not serve usefully as temporal filters for motion detection. Instead, it is left to the velocity extraction mechanism to do all of its own temporal filtering. This choice helped make it clear how phase advance can cause illusory motion in contrast-driven RAPs, but it makes the model unrealistic as a general model of motion extraction. The main point here is that a more realistic mechanism would face the same problem from phase advance.

Due to saccades, natural vision is based on a sequence of abrupt-onset, largely static retinal images. The question remains as to how motion energy units in cortex normally deal with high-contrast phase advance during burst of activity following each saccade. We suppose that, in addition to being incapable of distinguishing a change

in spatial frequency from a change in temporal frequency or contrast (Heeger, 1987), *a given motion energy unit is incapable of distinguishing between a delayed onset of activity in one or more of its subunits that is due to real motion, and a delayed onset that is due to a difference in contrast.* For example, using units with gabor-shaped receptive fields (Jones & Palmer, 1987), one can build a motion detector by introducing a delay between two units in quadrature phase (Adelson & Bergen, 1985). This detector, or one equivalent to it, is presumably the basis for the "motion without movement" illusion (Freeman, Adelson, & Heeger, 1991). One cannot help but notice that the contrast pattern in "Rotating Snakes" is very well suited to exploit phase advance to fool such a detector.

Very few studies have measured the responses of visual neurons to abrupt-onset static stimuli. Dynamic stimuli have been used in the majority of physiological studies of contrast adaptation because neuronal firing rates are higher for these stimuli. One study (other than Albrecht et al., 2002) that used static stimuli found that cortical neurons in anesthetized monkeys with immobile eyes adapted within 1 s to static patterns (Müller et al., 1999), but the retinal images in that study were not jittered, as they would be in an awake animal due to ocular microtremor. Adaptations with longer time courses may occur in LGN (Hawken, Shapley, & Grosof, 1996).

### Are global motion estimators peculiarly vulnerable to raps?

By pooling inputs across large regions of space, the global motion system efficiently detects global motion (Morrone, Burr, & Vaina, 1995; Burr, Morrone, & Vaina, 1998). We suspect that there is normally no cost to having global motion detectors that fail to account for the temporal dynamics of contrast coding. Presumably the subunits of global motion estimators report spurious local velocities in natural stimuli a great deal of the time. But for most patterned objects in the world, such signals would not be collectively consistent with a single global (i.e., regionally rigid) motion. If RAPs are statistically unlikely in the natural world, then a Bayesian ideal observer (Geisler et al., 1991) looking at the spatial pattern of spurious local velocity signals evoked by a static natural image would seldom infer the presence of global motion when there is none.<sup>2</sup>

Consistent with this account is the fact that the isolated elements in a RAP do not appear to move very much when every other element is reversed (Auxiliary Figure 6). Contrast can have a dramatic effect on apparent speed: When two parallel gratings moving at the same speed are presented simultaneously, the lower-contrast grating appears slower (Johnston, Benton, & Morgan, 1999; Shioiri et al., 2002; Stone & Thompson, 1992; Thompson,

1982; Thompson, Stone, & Swash, 1996). This may reflect a structural deficiency in the local motion mechanism, but it is also expected as a consequence of optimal motion estimation in any system that knows that local speed estimates are noisy at low contrast, and that slower local motions are more likely to occur than faster ones (Weiss, Simoncelli, & Adelson, 2002). In the alternating-element RAP, each element could still generate a weak local velocity signal, but each of these signals would have to be evaluated on its own merits. No separate motion template exists to detect and measure this flow field; the absence of such a template instantiates the system's belief that the observed flow field is statistically unlikely to result from real motion and that, accordingly, no motion percept should be constructed. Unlike global rotation, the alternating-element RAP does not give rise to a pattern of neural activity that is known, a priori, to be a reliable indicator of motion in the world.

### Illusory motion from sensors for dimming and brightening

In our model, motion is detected from changes over time in the neural image of contrast. Sign-labeled activity in the neural image is treated like an optical image by the velocity detector. However, motion can also be computed from the rate of change in luminance at separate locations in the image, and there exist spatially localized mechanisms early in the human visual pathway that are sensitive to steadily increasing (and decreasing) luminance per se (Anstis, 1967). These mechanisms contribute to perceived motion (Anstis, 1990), which makes them candidate building blocks for explaining illusory motion in RAPs. Motion could be detected within a neural image at which each point represents not contrast, but rate of change in luminance.

We do not know of neurophysiological data that can constrain such a model to make specific predictions, as the model in Figure 2 can from the data of Albrecht et al. (2002). We cannot rule out the possibility that local luminance change mechanisms contribute to RAP illusory motion. A model that predicts an illusion can be constructed by placing local adaptation to luminance (i.e., the process by which all luminances come to appear the same as mean luminance) before the luminance change detectors. If the detectors do not compensate for luminance adaptation, but instead respond to it, and if the rate of luminance adaptation is disproportionately faster at high contrast, then the luminance change detectors would respond to a static image of “Rotating Snakes” the same way they would to a movie of “Rotating Snakes” in which the black and white regions become gray at a faster rate than the dark and light regions. The static image would therefore look like real motion to a temporal gradient-based velocity detector, and to subsequent processing steps such as those shown in Figure 2 after Step 3.

### Does ocular microtremor do more than maintain static contrast at edges?

We have proposed that the illusory motion in a RAP comes to a stop after 6 s because that is when nonlinear local adaptations to luminance and contrast are complete. A different suggestion for RAP motion is that the oscillations in local retinal illumination, caused by small eye movements during fixation (Eizenman, Hallett, & Frecker, 1985), stimulate asymmetric responses in temporal gradient sensors that are interpreted as motion (Ashida & Kitaoka, 2003). In this case, it is not the illusory motion itself but rather the *slowing* of the illusory motion over the course of 6 s that is attributed to adaptation. The slowing would occur as the temporal gradient sensors cease to respond, or their asymmetry is equalized, or both. A physiological basis for this mechanism could be that presumptive P ganglion cells in primates adapt slowly to temporal contrast in vitro (Chander & Chichilnisky, 2001), although this was not confirmed at the level of the LGN in vivo (Solomon et al., 2004).

Whether a phenomenon along these lines contributes to the illusion may be difficult to determine. The obvious experiment would ask whether motion is seen in stabilized retinal images, but stabilized images stimulate the visual system weakly and they fade within 2–3 s because small eye movements during fixation are essential to the static representation of contrast (Ditchburn, 1987). As a result, the reduction or abolition of illusory motion for RAPs when they are stabilized is predicted by both accounts.

One reason to suspect that the illusion is caused directly by adaptation over time, rather than being driven by eye movements, is that change over time in the adapted states of visual neurons is clearly sufficient to evoke a strong percept of motion (Petrov & Popple, 2002). [Auxiliary Flash Demonstration 1](#) illustrates this. In the demonstration, a single-gradient RAP alternates with a plain white (or black) background. The afterimage rotates at definite speed, either clockwise (on the white background) or counterclockwise (on the black background), as it fades over the course of ~1 s (Naor-Raz & Sekuler, 2000). A fading afterimage is, of course, a neural representation that changes over time for strictly internal reasons, independent of eye movements. Thus, an explanation based on small eye movements cannot explain that illusion, and at this point there is no compelling reason to appeal to small eye movements to explain RAP illusory motion either.

### Effect of blur

In the authors' experience, blur causes the “Rotating Snakes” image to move more slowly in peripheral vision, but more quickly in central vision ([Auxiliary Figure 8](#)). Is this consistent with the model? According to the model,

blur is expected to reduce the illusory motion because blur makes the luminance profile more nearly sinusoidal. One way to think of this is that blur defeats the pattern's ability to generate distinct low- and high-contrast components with different phases (but the same spatial frequency) in the neural image. Clearly, we could not filter an image like [Figure 2](#) (Step 0) ahead of time at the RAP's spatial frequency and expect the illusion to work—its luminance profile would have become a single sinusoid.

Thus, in peripheral vision, blur reduces the illusion by removing high spatial frequencies from the neural image. But in foveal vision, it appears that blurring has an unmasking effect. A plausible explanation is that velocity estimators in foveal vision give greater weight to high spatial frequency mechanisms, and because edges do not move within the neural image during evolution of the image, these high spatial frequency mechanisms report the absence of motion, unless the edges are removed by optical blurring. An analogous phenomenon is well known for pattern perception: A pattern that is recognizable from its low spatial frequency content becomes unrecognizable when it also contains high spatial frequencies. An example is the well-known block Lincoln picture of Harmon and Julesz (1973).

## Individual differences

One of Fraser and Wilcox's (1979) most intriguing discoveries was that a genetic component accounted for much of the variance across individuals in their susceptibility to illusory motion. Different people saw rotation in different directions, with relatives tending to report perceived rotation in the same direction. A significant number of people do not see motion in "Rotating Snakes". There is no shortage of loci within the model at which individual differences in visual function might have this effect. For example, adaptation of the luminance image over time might depend on very specific properties of ON and OFF bipolar and retinal ganglion cells. The genetically controlled level of expression of a single protein could easily change the time course of activity in one of these cell types.

The contrast image could be further affected by specific properties of LGN cell responses. In cortex, signal-to-noise considerations might dictate that motion energy detectors have to exceed different activity thresholds before contributing to velocity detectors, or the global motion detectors themselves might be constructed in ways that differ slightly from one person to another.

Kitaoka has created hundreds of variations on his theme of static patterns that appear to move. They are in part the result of an "evolutionary process": Patterns that gave rise to the maximum illusory motion were selected, and new patterns were made by varying them in an iterative cycle. The patterns evolved to have a maximum effect in a very complicated environment, namely, the human visual

system. As a general proposition, entities that evolve in a complicated environment may come to exploit their environment in complicated ways, so it is highly probable that "Rotating Snakes" exploits more than one visual mechanism to achieve its effect. No single theory is likely to explain all of the illusion, and individual variation in several separate mechanisms could cause the perception of motion in RAPs to differ between individuals.

A sensible way to proceed might be to exploit these differences, and to correlate various measures of susceptibility to RAP illusory motion and performance on a variety of psychophysical tasks, ideally using tasks for which performance is limited by known mechanisms.

## Conclusion

Neural adaptations are a ubiquitous feature of the visual system. Myriad controlled adaptations at many levels enhance sensitivity to important stimulus features, reduce bias, and save energy (Brown & Masland, 2001; Laughlin, de Ruyter van Steveninck, & Anderson, 1998; Laughlin & Sejnowski, 2003; Solomon et al., 2004). This raises the question: To what extent do later stages of neural processing take into account adaptation at the earlier stages? A Bayesian ideal observer (Geisler et al., 1991) looking at a dynamic pattern of neural activity could infer that a pattern is static, if it knows how neural activity usually changes over time for static patterns. Accordingly, we suggest that most people's visual systems infer the presence of motion in static RAPs because static RAPs evoke a pattern of neural activity that normally occurs only when an object really is moving. In this formulation, the converse is no paradox: Complicated dynamic patterns of neural firing can and often do give rise to static percepts. Indeed, what seems amazing is how seldom we see illusory motion—especially after saccades—in static patterns. Thus, while neural adaptations have been invoked to explain a variety of illusions, it is not trivial to predict when an adaptation will give rise to an illusion because later stages of processing sometimes do, and sometimes do not, take adaptations into account.

## Acknowledgments

We gratefully acknowledge Akiyoshi Kitaoka for discussion and permission to reproduce his artwork, and Judy Chicago for permission to reproduce her artwork. We thank two anonymous reviewers, David Brainard, Jack Nachmias, Larry Palmer and Peter Sterling for comments on the paper, and Jesse Frumkin and Richard Pater for serving as observers. This work was supported by NIH grants EY013988 and P30 EY001583.

Commercial relationships: none.

Corresponding author: Benjamin T. Backus.

Email: backus@psych.upenn.edu.

Address: Department of Psychology, University of Pennsylvania, 3401 Walnut St. C-Wing Room 302-C, Philadelphia, PA 19104-6228.

## Footnotes

<sup>1</sup>Contrast adaptation should depend on luminance in the image because contrast is coded as deviation from a reference luminance. The reference is probably approximated by mean luminance. This leads to an interesting prediction. For a gray-to-white gradient on a white background, a diagram like that of [Figure 4](#) predicts illusory motion in the gray-to-white direction, as grays become darker (due to both luminance and contrast adaptation). But on a black background, there should be less motion or perhaps even motion in the white-to-gray direction, as contrast adaptation causes gray to become lighter. This is the case ([Auxiliary Figures 10g](#) and [i](#)).

<sup>2</sup>The most compelling natural RAP we have seen is a closed Venetian blind in the window of a shaded brick wall. Shading causes each slat in the blind to give rise to a dark-to-light gradient (in the downward direction). The window thus appears to drift down relative to the wall. This illusion is subtle compared with “Rotating Snakes”, but it is nevertheless quite obvious once it is pointed out. On the other hand, a black-dark-white-light sequence that is *not* repeated in a natural image typically fails to evoke illusory motion; the same sequence evokes motion when repeated ([Auxiliary Figure 11](#)).

## References

- Adelson, E. H., & Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America. A*, *2*, 284–299. [[PubMed](#)] [[Article](#)]
- Albrecht, D. G., Geisler, W. S., Frazor, R. A., & Crane, A. M. (2002). Visual cortex neurons of monkeys and cats: Temporal dynamics of the contrast response function. *Journal of Neurophysiology*, *88*, 888–913. [[PubMed](#)] [[Article](#)]
- Anstis, S. (1990). Motion aftereffects from a motionless stimulus. *Perception*, *19*, 301–306. [[PubMed](#)]
- Anstis, S. M. (1967). Visual adaptation to gradual change of intensity. *Science*, *155*, 710–712. [[PubMed](#)]
- Anstis, S. M. (1970). Phi movement as a subtraction process. *Vision Research*, *10*, 1411–1430. [[PubMed](#)]
- Anstis, S., & Becker, M. (2005). Personal communication.
- Ashida, H., & Kitaoka, A. (2003). A gradient-based model of the peripheral drift illusion [[Abstract](#)]. *Perception*, *32*(ECPV supplement), 106.
- Backus, B. T., & Oruç, I. (2004). Rotating snakes and the failure of motion mechanisms to compensate for early adaptation to luminance [[Abstract](#)]. *Journal of Vision*, *4*(8), 85a, <http://journalofvision.org/4/8/85/>, doi:10.1167/4.8.85.
- Bex, P. J., Metha, A. B., & Makous, W. (1998). Psychophysical evidence for a functional hierarchy of motion processing mechanisms. *Journal of the Optical Society of America. A, Optics, Image Science, and Vision*, *15*, 769–776. [[PubMed](#)]
- Braddick, O. (1974). A short-range process in apparent motion. *Vision Research*, *14*, 519–527. [[PubMed](#)]
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*, 433–436. [[PubMed](#)] [[Article](#)]
- Britten, K. H., Shadlen, M. N., Newsome, W. T., & Movshon, J. A. (1992). The analysis of visual motion: A comparison of neuronal and psychophysical performance. *Journal of Neuroscience*, *12*, 4745–4765. [[PubMed](#)] [[Article](#)]
- Brown, S. P., & Masland, R. H. (2001). Spatial scale and cellular substrate of contrast adaptation by retinal ganglion cells. *Nature Neuroscience*, *4*, 44–51. [[PubMed](#)] [[Article](#)]
- Burr, D. C., Morrone, M. C., & Vaina, L. M. (1998). Large receptive fields for optic flow detection in humans. *Vision Research*, *38*, 1731–1743. [[PubMed](#)] [[Article](#)]
- Carandini, M., Heeger, D. J., & Movshon, J. A. (1999). Linearity and gain control in V1 simple cells. In P. Ulinski, E. Jones, & A. Peters (Eds.), *Cerebral cortex, Vol. 13: Models of cortical function* (pp. 401–443). New York: Plenum.
- Cavanagh, P., & Favreau, O. E. (1980). Motion aftereffect: A global mechanism for the perception of rotation. *Perception*, *9*, 175–182. [[PubMed](#)]
- Chander, D., & Chichilnisky, E. J. (2001). Adaptation to temporal contrast in primate and salamander retina. *Journal of Neuroscience*, *21*, 9904–9916. [[PubMed](#)] [[Article](#)]
- Chicago, J. (Artist). (1973). *Through the Flower* [Acrylic on canvas, 5 ft. × 5 ft.].
- Chubb, C., Sperling, G., & Solomon, J. A. (1989). Texture interactions determine perceived contrast. *Proceedings of the National Academy of Sciences of the United States of America*, *86*, 9631–9635. [[PubMed](#)] [[Article](#)]
- Conway, B. R., Kitaoka, A., Yazdanbakhsh, A., Pack, C. C., & Livingstone, M. S. (2005). Neural basis for a powerful static motion illusion. *Journal of Neuroscience*, *25*, 5651–5656. [[PubMed](#)]

- Ditchburn, R. W. (1987). What is psychophysically perfect image stabilization? Do perfectly stabilized images always disappear?: Comment. *Journal of the Optical Society of America. A*, *4*, 405–406. [PubMed]
- Eizenman, M., Hallett, P. E., & Frecker, R. C. (1985). Power spectra for ocular drift and tremor. *Vision Research*, *25*, 1635–1640. [PubMed]
- Faubert, J., & Herbert, A. M. (1999). The peripheral drift illusion: A motion illusion in the visual periphery. *Perception*, *28*, 617–621. [PubMed]
- Fraser, A., & Wilcox, K. J. (1979). Perception of illusory movement. *Nature*, *281*, 565–566. [PubMed]
- Freeman, W. T., Adelson, E. H., & Heeger, D. J. (1991). Motion without movement. *ACM Computer Graphics (SIGGRAPH Conference)*, *25*, 27–30. [Article]
- Geisler, W. S., Albrecht, D. G., Salvi, R. J., & Saunders, S. S. (1991). Discrimination performance of single neurons: Rate and temporal-pattern information. *Journal of Neurophysiology*, *66*, 334–362. [PubMed]
- Georgeson, M. A. (1987). Temporal properties of spatial contrast vision. *Vision Research*, *27*, 765–780. [PubMed]
- Gregory, R. L., & Heard, P. F. (1983). Visual dissociations of movement, position, and stereo depth: Some phenomenal phenomena. *The Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology*, *35*(Pt. 1), 217–237. [PubMed] [Article]
- Grzywacz, N. M., & Yuille, A. L. (1991). Theories for the visual perception of local velocity and coherent motion. In M. S. Landy & J. A. Movshon (Eds.), *Computational models of visual processing* (pp. 231–252). Cambridge, Massachusetts: MIT Press.
- Harmon, L. D., & Julesz, B. (1973). Masking in visual recognition: Effects of two-dimensional filtered noise. *Science*, *180*, 1194–1197. [PubMed]
- Hawken, M. J., Gegenfurtner, K. R., & Tang, C. (1994). Contrast dependence of colour and luminance motion mechanisms in human vision. *Nature*, *367*, 268–270. [PubMed]
- Hawken, M. J., Shapley, R. M., & Gross, D. H. (1996). Temporal-frequency selectivity in monkey visual cortex. *Visual Neuroscience*, *13*, 477–492. [PubMed]
- Heeger, D. J. (1987). Model for the extraction of image flow. *Journal of the Optical Society of America. A*, *4*, 1455–1471. [PubMed]
- Heeger, D. J. (1993). Modeling simple-cell direction selectivity with normalized, half-squared, linear operators. *Journal of Neurophysiology*, *70*, 1885–1898. [PubMed]
- Helmholtz, H. v. (1925). *Treatise on physiological optics* (Vol. III) (J. P. C. S. o. G. 1910, Trans.). New York: Optical Society of America. [Book]
- Hurlimann, F., Kiper, D. C., & Carandini, M. (2002). Testing the Bayesian model of perceived speed. *Vision Research*, *42*, 2253–2257. [PubMed]
- Johnston, A., Benton, C. P., & Morgan, M. J. (1999). Concurrent measurement of perceived speed and speed discrimination threshold using the method of single stimuli. *Vision Research*, *39*, 3849–3854. [PubMed]
- Jones, J. P., & Palmer, L. A. (1987). The two-dimensional spatial structure of simple receptive fields in cat striate cortex. *Journal of Neurophysiology*, *58*, 1187–1211. [PubMed]
- Kitaoka, A. (Artist). (2003). *Rotating Snakes* [JPEG image].
- Laughlin, S. B., de Ruyter van Steveninck, R. R., & Anderson, J. C. (1998). The metabolic cost of neural information. *Nature Neuroscience*, *1*, 36–41. [PubMed] [Article]
- Laughlin, S. B., & Sejnowski, T. J. (2003). Communication in neuronal networks. *Science*, *301*, 1870–1874. [PubMed]
- Lu, Z. L., & Sperling, G. (1996). Contrast gain control in first- and second-order motion perception. *Journal of the Optical Society of America. A*, *13*, 2305–2318. [PubMed]
- Morrone, M. C., Burr, D. C., & Vaina, L. M. (1995). Two stages of visual processing for radial and circular motion. *Nature*, *376*, 507–509. [PubMed]
- Müller, J. R., Metha, A. B., Krauskopf, J., & Lennie, P. (1999). Rapid adaptation in visual cortex to the structure of images. *Science*, *285*, 1405–1408. [PubMed]
- Nakayama, K., & Tyler, C. W. (1981). Psychophysical isolation of movement sensitivity by removal of familiar position cues. *Vision Research*, *21*, 427–433. [PubMed]
- Naor-Raz, G., & Sekuler, R. (2000). Perceptual dimorphism in visual motion from stationary patterns. *Perception*, *29*, 325–335. [PubMed]
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*, 437–442. [PubMed]
- Petrov, Y. A., & Popple, A. V. (2002). Effects of negative afterimages in visual illusions. *Journal of the Optical Society of America. A*, *19*, 1107–1111. [PubMed]
- Rashbass, C. (1970). The visibility of transient changes of luminance. *Journal of Physiology*, *210*, 165–186. [PubMed]
- Reichardt, W. (1957). Autokorrelationsauswertung als Funktionsprinzip des Zentralnervensystems. *Zeitschrift für Naturforschung. Teil B*, *12*, 447–457.
- Shapiro, A. G., D'Antona, A. D., Charles, J. P., Belano, L. A., Smith, J. B., & Shear-Heyman, M. (2004). Induced contrast asynchronies. *Journal of Vision*, *4*(6), 459–468, <http://journalofvision.org/4/6/5/>, doi:10.1167/4.6.5. [PubMed] [Article]

- Shapley, R. M., & Victor, J. D. (1978). The effect of contrast on the transfer properties of cat retinal ganglion cells. *Journal of Physiology*, 285, 275–298. [[PubMed](#)]
- Shioiri, S., Ito, S., Sakurai, K., & Yaguchi, H. (2002). Detection of relative and uniform motion. *Journal of the Optical Society of America. A, Optics, Image Science, and Vision*, 19, 2169–2179. [[PubMed](#)]
- Snippe, H. P., Poot, L., & van Hateren, J. H. (2000). A temporal model for early vision that explains detection thresholds for light pulses on flickering backgrounds. *Vision Neurosciences*, 17, 449–462. [[PubMed](#)]
- Snowden, R. J. (1998). Shifts in perceived position following adaptation to visual motion. *Current Biology*, 8, 1343–1345. [[PubMed](#)] [[Article](#)]
- Solomon, S. G., Peirce, J. W., Dhruv, N. T., & Lennie, P. (2004). Profound contrast adaptation early in the visual pathway. *Neuron*, 42, 155–162. [[PubMed](#)] [[Article](#)]
- Stone, L. S., & Thompson, P. (1992). Human speed perception is contrast dependent. *Vision Research*, 32, 1535–1549. [[PubMed](#)] [[Article](#)]
- Thompson, P. (1982). Perceived rate of movement depends on contrast. *Vision Research*, 22, 377–380. [[PubMed](#)] [[Article](#)]
- Thompson, P., Stone, L. S., & Swash, S. (1996). Speed estimates from grating patches are not contrast-normalized. *Vision Research*, 36, 667–674. [[PubMed](#)] [[Article](#)]
- Tolhurst, D. J., Walker, N. S., Thompson, I. D., & Dean, A. F. (1980). Non-linearities of temporal summation in neurones in area 17 of the cat. *Experimental Brain Research*, 38, 431–435. [[PubMed](#)]
- van Santen, J. P., & Sperling, G. (1984). Temporal covariance model of human motion perception. *Journal of the Optical Society of America. A*, 1, 451–473. [[PubMed](#)]
- Watson, A. B. (1979). Probability summation over time. *Vision Research*, 19, 515–522. [[PubMed](#)] [[Article](#)]
- Watson, A. B., & Ahumada, A. J., Jr. (1985). Model of human visual-motion sensing. *Journal of the Optical Society of America. A*, 2, 322–341. [[PubMed](#)] [[Article](#)]
- Watson, A. B., & Pelli, D. G. (1983). QUEST: A Bayesian adaptive psychometric method. *Perception & Psychophysics*, 33, 113–120. [[PubMed](#)] [[Article](#)]
- Weiss, Y., Simoncelli, E. P., & Adelson, E. H. (2002). Motion illusions as optimal percepts. *Nature Neuroscience*, 5, 598–604. [[PubMed](#)] [[Article](#)]
- Williams, D. W., & Sekuler, R. (1984). Coherent global motion percepts from stochastic local motions. *Vision Research*, 24, 55–62. [[PubMed](#)]