Available online at www.sciencedirect.com
SCIENCE DIRECT
Vision Research 46 (2006) 2421–2431
Vision Research
www.elsevier.com/locate/visres

A positive correlation between fixation instability and the strength of illusory motion in a static display

Ikuya Murakami [a,b,*], Akiyoshi Kitaoka [c], Hiroshi Ashida [d]

[a] Human and Information Science Laboratory, NTT Communication Science Laboratories, NTT Corporation, Japan
[b] Department of Life Sciences, University of Tokyo, Japan
[c] Department of Psychology, Ritsumeikan University, Japan
[d] Graduate School of Letters, Kyoto University, Japan

Received 19 July 2005; received in revised form 22 November 2005

Abstract

A stationary pattern with asymmetrical luminance gradients can appear to move. We hypothesized that the source signal of this illusion originates in retinal image motions due to fixational eye movements. We investigated the inter-subject correlation between fixation instability and illusion strength. First, we demonstrated that the strength of the illusion can be quantified by the nulling technique. Second, we concurrently measured cancellation velocity and fixation instability for each subject, and found a positive correlation between them. The same relationship was also found within a single observer when the visual stimulus was artificially moved in the simulation of fixation instability. Third, we confirmed the same correlation with eye movements for a wider variety of illusory displays. These results suggest that fixational eye movements indeed play a relevant role in generating this motion illusion.

Keywords: Visual motion; Fixational eye movements; Illusion; Correlation

1. Introduction

It has been pointed out that illusory motion can be seen in a completely static figure (e.g., a figure printed on a sheet of paper) (Faubert & Herbert, 1999; Fraser & Wilcox, 1979; Naor-Raz & Sekuler, 2000). The underlying mechanism of this illusion is currently under debate. Recently, one of the authors (A.K.) has created variants of the illusory figures (bitmap figures are currently available on the internet; see http://www.ritsumei.ac.jp/~akitaoka/index-e.html), an example of which is shown in Fig. 1. We found that the perceived motion was as vigorous as real motion, and wondered what neural mechanism is responsible. The design rule is simple: construct an array of four regions having different luminances in the systematic order, "black," "dark gray," "white," and "light gray" (Ashida & Kitaoka, 2003; Kitaoka & Ashida, 2003). Slow illusory motion is seen between adjacent regions; the perceived direction is in the order described above, but never the other way around (Fig. 2A). Also, the illusion is clearer when the array pattern is repeated in a row such that the patterns comprise a circle (see Fig. 1), and when it is viewed peripherally.

Several investigators have previously paid attention to the motion illusion seen in such an asymmetric luminance gradient, but they have typically used a saw-tooth luminance profile, not the above-mentioned profile. Fraser and Wilcox (1979) noted a large individual difference (and also a genetic similarity) in perceptual strength of the illusion. Faubert and Herbert (1999) reported that the illusion was more vigorous after an eye blink or a saccade. Naor-Raz and Sekuler (2000) found that the illusion became more vigorous in more peripheral observations, for longer durations, and at higher luminance contrasts,

* Corresponding author. Fax: +81 3 5454 6979.
E-mail address: ikuya@fechner.c.u-tokyo.ac.jp (I. Murakami).

0042-6989/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2006.01.030

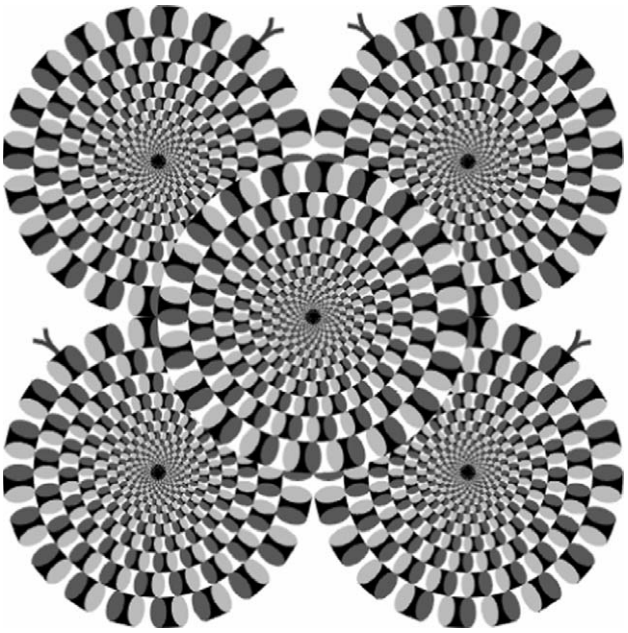

Fig. 1. The "rotating snakes" illusion. When one loosely looks at the center, the outer four disks appear to rotate slowly, with two in the clockwise direction and the other two in the counterclockwise direction.

and also mentioned the possible involvement of involuntary fixational eye movements. In line with these previous investigations, we argue that a physically static figure for this motion illusion is not static at all on the retina, but is instead always moving in random fashion together with fixational eye movements. The amplitude of fixational eye movements is known to be substantially different among subjects (Murakami, 2004), suggesting that inter-subject variability in perceptual strength as reported by Fraser and Wilcox might be a consequence of the variability of eye movements. In addition, fixation is less stable after a saccade than during steady fixation (Rucci & Desbordes, 2003), which implies that Faubert and Herbert's observation is possibly consistent with oculomotor statistics.

However, no previous studies have offered convincing psychophysical evidence for a relationship between the strength of this illusion and eye movements. Our hypothesis is that illusion strength positively correlates with fixation instability: as gaze fluctuates more, the retinal image of a physically static figure fluctuates more, and these jittery motions on the retina consequently lead to more vigorous impressions of illusory motion (we will later discuss how random image fluctuations are converted to smooth motion impression). To test this hypothesis, it would be ideal to vary a subject's fixation instability in a systematic manner and observe how this affects illusion strength. However, it is not easy to manipulate the amplitude of fixational eye movements of an individual. We therefore recruited a relatively large number of subjects and plotted inter-subject scattergrams. As a result, we found a positive correlation between illusion strength and fixation instability.

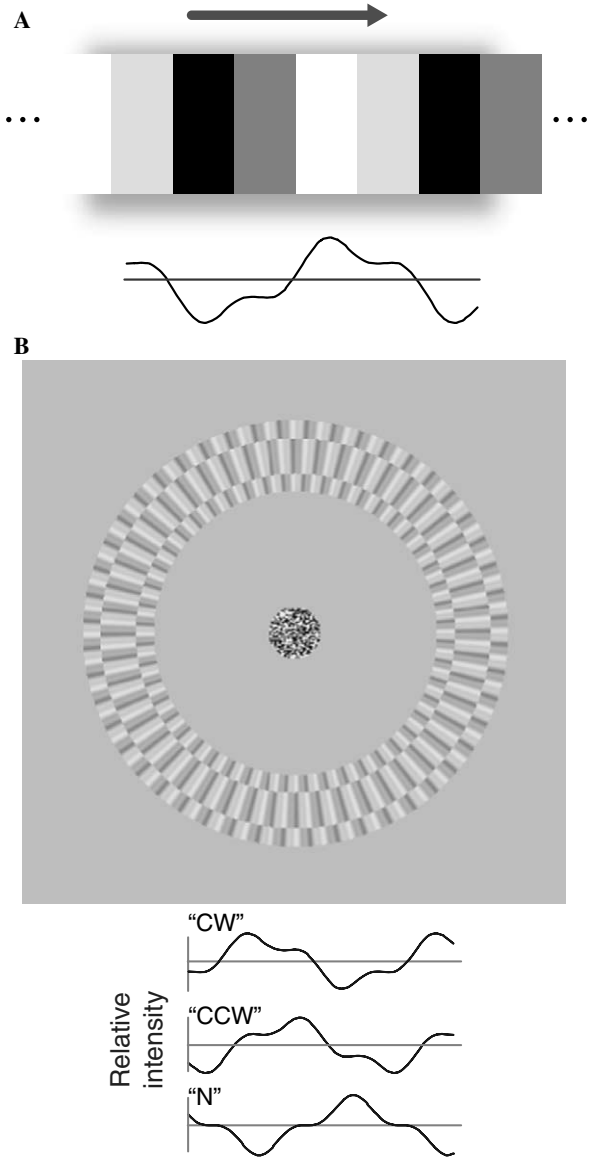

Fig. 2. Schematics of the stimuli. (A) The design rule of the "rotating snakes" illusion. The order of the four luminance levels is critical. This pattern consists of the fundamental and the odd harmonic components. The lower curve shows the luminance profile of the fundamental plus the third harmonic components, with the central straight line indicating the mean luminance. (B) Illustrations of the stimulus configuration and the waveforms used in illusion cancellation. The abscissa indicates polar angle (rightward is clockwise).

The present study is comprised of three experimental sections. In the pilot experiment, we devised a psychophysical method of velocity cancellation to quantify illusion strength in a precisely controlled stimulus. In the main experiment, we concurrently measured cancellation velocity and fixation instability for 22 subjects to see an inter-subject correlation between these quantities. Within a single subject, we also tested the dependence of illusion strength on the magnitude of simulated fixation instability, by actually oscillating the visual stimulus. In the subsidiary experiment, we had 53 subjects rate the perceptual strength of the illusion and also separately measured fixation

instability to duplicate our finding in a wider variety of illusory figures.

2. Pilot experiment: Illusion cancellation

In this experiment, we tested whether the illusion can be quantified in an objective fashion by nulling the illusory motion with opposing real motion. One of the authors (H.A.) and three naïve subjects with normal (or corrected-to-normal) visual and oculomotor functions participated.

2.1. Methods

2.1.1. Apparatus

The stimuli were generated on a Windows PC with the DirectDraw technology, and were displayed on a CRT monitor (NANAO, E57Ts, 1024 × 768 pixels, 0.30 mm/pixel, refresh rate 75 Hz, viewing distance 46 cm constrained by the chinrest). The mean luminance was 36 cd/m^2. The subject binocularly viewed the stimulus in a dim room.

2.1.2. Stimuli

The stimuli were created on the basis of the stepwise luminance rule (Fig. 2A). To allow minute control of stimulus speed using sub-pixel coding, we used a smooth compound waveform that consisted of the first and the third harmonic components of the stepwise stimuli. Luminance (m) as a function of polar angle (θ) is given by

$$m(\theta) = c[\cos(f\theta) + \cos\{3f(\theta + \phi)\}/3], \quad (1)$$

where f is the spatial modulation frequency, ϕ is the relative phase between the two components, and c is the modulation depth. The grating was asymmetric when $\phi = \pm\pi/2$, while it was symmetric when $\phi = 0$. The asymmetric stimuli are hereafter referred to as CW (clockwise) and CCW (counterclockwise) and the symmetric stimuli are referred to as N (neutral). Note that these stimuli were defined by the physical properties but were named to connote expected direction perception (Fig. 2A).

Radial gratings of this waveform were used (Fig. 2B). The width of the ring was 3.6 deg of visual angle, stimulating the retinal locations between 7.4 and 11.0 deg in the periphery. A disk of 1.3 deg radius, which was filled with random black and white dots, was located in the center for gaze control. Since fixational eye movements were supposed essential for this illusion, the subjects were instructed not to fixate hard but to look inside this disk loosely. The other area within the screen was filled with the mean-luminance gray.

The spatial frequency of the gratings (f) was 48 cycles per annulus. It is empirically known that the illusion is enhanced by dividing the figure into thinner stripes, as done by Fraser and Wilcox (1979) and Faubert and Herbert (1999). We therefore divided the ring into a central stripe of 2.4 deg and two flanking stripes of 1.2 deg. The overall spatial phase was randomly set for each trial, and the central phase was always shifted by a half cycle from the flanking ones. The modulation depth (c) was 0.53, so that the Michelson contrast of the symmetric grating was 0.71 and that of the asymmetric grating was 0.66. They drifted either in the clockwise or the counterclockwise direction at the speed of 0.0525–0.2625°/s (0.009–0.044 rpm) by the step of 0.0525°/s, where 1°/s corresponds to one degree of polar angle per second.

2.1.3. Procedure

The stimulus speed that gave the subjective stationary point was measured by the method of constant stimuli. A drifting grating was presented for 0.5 s, and the subject made a forced-choice judgment of the rotation direction and responded using the keyboard. Each trial was initiated by the subject by pressing a key. For each data point, 40 responses were collected through five sessions, except subject Y.S. for whom 32 responses were collected.

The speed and the grating type were randomized within a session. It was impossible during this short period to tell which of the three types was presented.

2.2. Results

The naïve subjects did not find the task difficult. They did not experience competing or transparent percepts of real and illusory motion. The shift of null point of the psychometric function therefore indicates the speed that appeared stationary.

The null velocity was calculated by using the probit analysis. As shown in Fig. 3, the null point was shifted towards the counterclockwise direction for the CW stimulus and in the clockwise direction for the CCW stimulus, which means that the illusion occurred in the expected directions. The N stimulus did not yield significant bias in either direction. The results confirmed that the illusion is quantifiable by nulling it with real motion, which implies that the

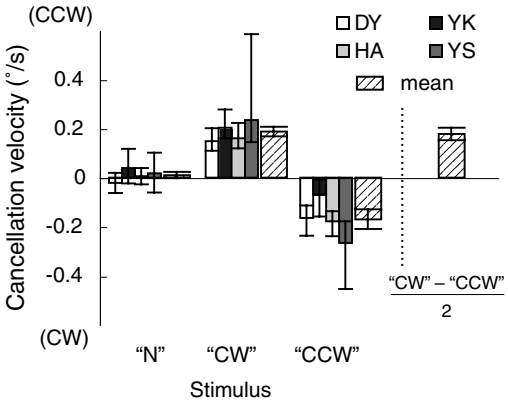

Fig. 3. The results of illusion cancellation for the four subjects. The bars represent cancellation velocity with 95% fiducial limits. Positive and negative values indicate counterclockwise (CCW) and clockwise (CW) directions, respectively. Indicated by the rightmost bar is the differential cancellation velocity between the CW and CCW stimuli.

illusion arises as a consequence of visual motion processing at the sensory level.

Although there were some variations in the cancellation velocity among subjects, this is expected from inter-subject variability of illusion strength that we are addressing in the present study. The fiducial limits for Y.S. are particularly asymmetric (i.e., the true null point could be larger than this estimation, but could not be smaller). Faster test stimuli should have been used for this subject, judging from the psychometric functions. (In the next experiment, where quantitative estimation was more crucial, an adaptive method was used instead of the method of constant stimuli.) Subject Y.K. showed an overall bias towards the counterclockwise direction of the null points, but the difference between responses for the CW and CCW stimuli was nevertheless clear for this subject. To factor out the overall bias, the difference between the cancellation velocities for the CW and CCW stimuli were divided by two and plotted by the rightmost bar. This differential cancellation velocity will be used (and will be simply called the "cancellation velocity") hereafter to represent the illusion strength as quantified by the cancellation method.

3. Main experiment: Cancellation velocity and fixation instability

As the illusion strength was shown to be objectively measurable, its correlation with fixation instability was assessed in the next experiment. While the subject was performing a psychophysical task to obtain the cancellation velocity, his or her eye movements were recorded concurrently. The first author and 21 naïve subjects (19 females and 3 males, aged 20–39) with normal (or corrected-to-normal) visual and oculomotor functions participated.

3.1. Methods

3.1.1. Apparatus and stimuli

The stimulus was presented on a 21-in. color CRT monitor (Sony GDM-F520, 640×480 pixels or 42.7 deg $\times$ 32 deg, refresh rate 75 Hz, viewing distance 54 cm constrained by the chinrest) controlled by a computer (Apple Power Macintosh G3). The stimulus configuration was identical to that used in the Pilot Experiment, except that the luminance contrast was 0.99.

To ask whether fixational eye movements of each subject could vary with experimental conditions, we modified the appearance of the fixation point (but note that these modifications were found to be ineffective). A tiny black disk (10 min in diameter), a larger white disk (55 min in diameter), or no fixation point, was provided during stimulus presentation. In the inter-stimulus interval, the tiny disk was provided constantly to help maintain fixation. The appearance of the fixation point was abruptly changed to the predetermined shape (e.g., to the large white disk) 133 ms before stimulus onset, and was abruptly returned to the tiny disk 13 ms after stimulus offset.

3.1.2. Procedure

One experimental set (90 s) consisted of 20 repeated trials, sandwiched by sequential presentation of 8 calibration dots (for eye-movement recording) at the beginning and end of the set. The subject was given a break for 30–180 s between sets, for ethical and technical reasons. All sets were finished within 60–70 min. In each set, the stimulus was presented for 0.5 s, and was followed by the inter-stimulus interval for 1.5 s, within which the subject had to indicate the perceived direction (clockwise or counterclockwise) of the grating by pressing a key. Each subject had learned the adequate timing of key-pressing over more than 200 practices; our subjects failed to respond in time for only 0.2% of actual trials, in such cases the staircase maintained the same level at the next trial.

The adaptive staircase method with the PEST transition rule (Taylor & Creelman, 1967) was used to find the null point for each of the two stimulus types (CW and CCW) and for each of the three fixation-point types. For each set, the fixation-point type was fixed but the two stimulus types were intermingled randomly from trial to trial. Since each set contained only 20 trials, there were 10 repeated trials for the CW stimulus and 10 for the CCW stimulus. The staircase sequence for each was therefore "inherited" across sets, such that the last trial of the first set and the first trial of the second set should be regarded as though two consecutive trials. As such, a complete staircase (comprised of 60 virtually consecutive trials) was actually a splice of 6 sets. The average of staircase transition points (excluding the first four transitions) was taken as the null point.

3.1.3. Eye-movement recording

Fixational eye movements during a set were recorded concurrently. As the index of fixation instability, we used the variability of horizontal miniature drift eye movements during fixation (Murakami, 2004). The horizontal gaze position was recorded for both eyes by an infrared limbus eye tracker (Iota Orbit 8) with the sampling resolution of 1 kHz. To recalibrate gain and offset parameters for each trial, the recorded eye positions during calibration were fit with the time series of calibration-dot positions; if $r < 0.65$, trials were excluded (8.9% of all), but this criterion did not influence the results.

Data were band-pass-filtered (1–31 Hz) to obtain instantaneous velocity with the effective resolution of 13 ms. A microsaccade was detected by the velocity criterion of 10 deg/s (Bair & O'Keefe, 1998; Snodderly, Kagan, & Gur, 2001) and counted within each stimulus duration. To calculate saccade-free statistics of eye drifts in this duration, data within $\pm$26 ms around each microsaccade were removed from the analysis below. The histogram (with the bin width of 0.1 deg/s) of instantaneous velocities of eye drifts was plotted and fit with a Gaussian, the standard deviation of which was taken as the index of fixation instability. The variance due to machine noise was also measured by using an artificial eye on a dummy head, and

was subtracted from human measurements. Since there was no systematic difference between the left and right eyes, the estimated fixation instability was averaged across eyes for simplicity.

For a control, eye movements during passive observation of a stationary random-dot pattern (a disk-shaped region with the diameter of 7 deg) with a central fixation point were also recorded in intermingled sets (20 s each). Data were analyzed the same way as above.

3.2. Results

We found a positive correlation between illusion strength and fixation instability. We first saw eye-movement data and cancellation velocity for three fixation-point types separately, but found no effect of fixation-point type (repeated-measures ANOVA for fixation instability, $F_{2,21} = 1.48$, *n.s.*; for cancellation velocity, $F_{2,21} = 1.44$, *n.s.*). Therefore, to reduce noise, the cancellation velocity data and fixation instability data were averaged across the three fixation-point types. The inter-subject scattergram is plotted for these averaged data (Fig. 4A). Each point corresponds to each subject. Clearly, a positive correlation between illusion strength and fixation instability was revealed ($r = 0.425$; $t_{20} = 2.10$, $p < 0.05$).

The same analysis was also made between cancellation velocity and the frequency of microsaccades (Fig. 4B); there was no indication of correlation ($r = -0.082$, *n.s.*).

One might notice that, in Fig. 4A, two subjects exhibited negative cancellation velocities. Actually, these data points were not significantly different from zero velocity, hence probably the apparent negative deviation only reflects sampling noise derived from a limited staircase dataset. One might then wonder whether these points artificially led to the above correlation. However, this was not the case. In Fig. 4C, the data were analyzed the same way as Fig. 4A, but we excluded the data in each condition if the cancellation velocity was not statistically significant. After this stringent screening, a positive correlation between illusion strength and fixation instability was not only confirmed but also substantially improved ($r = 0.571$; $t_{15} = 2.69$, $p < 0.05$).

We cannot immediately interpret the correlation we found as the causation that the small eye movements drive the illusion; theoretically the reverse direction is possible, i.e., that the illusion increases fixation instability. This direction is extremely unlikely, however. Among experimental sets, we also intermingled control sets in which each subject's fixation instability was monitored during passive viewing (simple fixation task). The inter-subject variability did not substantially increase during the motion-cancellation task (Fig. 5); there was a high positive correlation in fixation instability between the passive-viewing condition and motion-task condition ($r = 0.868$; $t_{20} = 7.82$, $p \ll 0.0001$), and the data points gathered along the identity function. Therefore, the inter-subject variability would be interpreted as a maintained characteristic inherent in the particular set of participants.

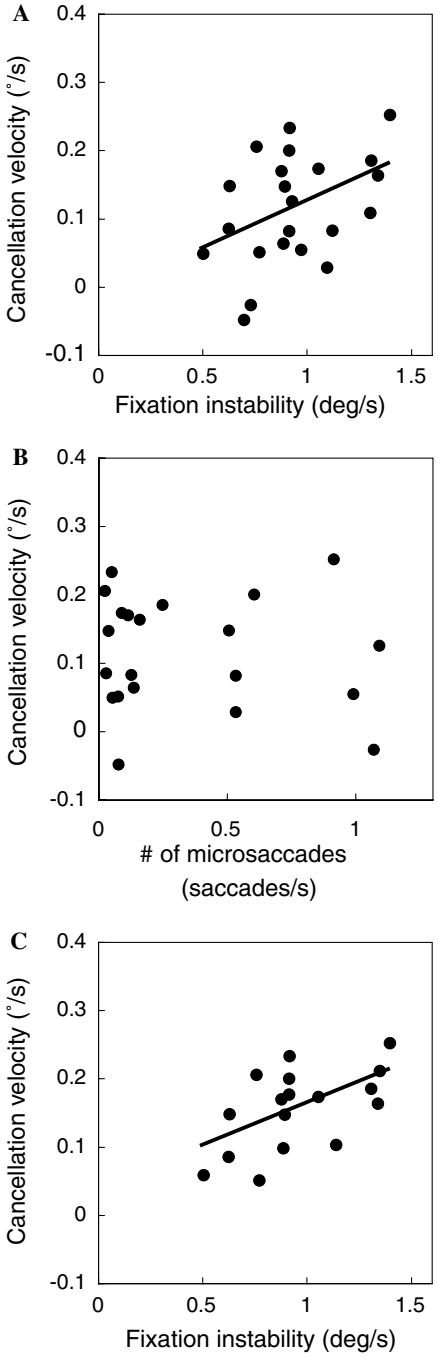

Fig. 4. Scattergram to visualize inter-subject correlation between cancellation velocity and eye-movement statistics. Data from 22 subjects are plotted, one point corresponding to one subject. The data for three fixation-point types were averaged in the ordinate and abscissa. (A) Inter-subject scattergram between cancellation velocity and fixation instability (as defined by the standard deviation of drift velocity). The solid line indicates the linear regression: $y = 0.0185x + 0.002$. (B) Inter-subject scattergram between cancellation velocity and microsaccade frequency (number per second). (C) Inter-subject scattergram between cancellation velocity and fixation instability based on statistically significant cancellation velocities only.

To further elucidate the causal link, we repeated the cancellation experiment on within-subject basis, by applying actual jittery motions at variable speeds to the visual

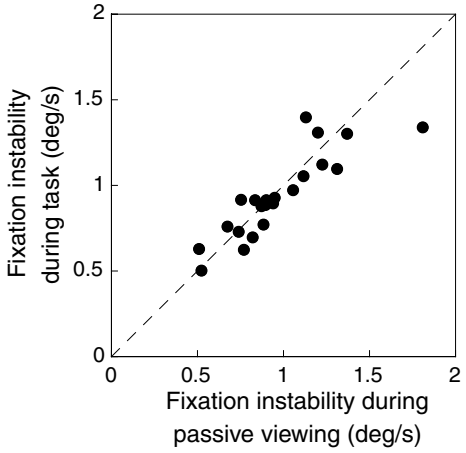

Fig. 5. Scattergram to visualize inter-subject comparison of fixation instability between the motion-task condition and passive-viewing condition. Data from the 22 subjects are plotted, one point corresponding to one subject. The broken line is $y = x$.

stimulus on the monitor. The fixation point remained stationary, whereas the radial grating slightly shifted its overall position, mimicking image motions caused by small eye movements. The instantaneous velocity of the frame-by-frame shift was chosen randomly from a two-dimensional (horizontal × vertical) isotropic Gaussian probability density function. The actual shift was rounded by the display routine at pixel-wise spatial resolution (2.5 min), but was fine enough relative to the visual acuity at the given eccentricity. The standard deviation of the Gaussian, or the "simulated fixation instability," was the independent variable; the differential cancellation velocity was the dependent variable. The simulated fixation instability was fixed within each session, which was comprised of three randomly interleaved PEST sequences (60 consecutive trials for each) for the three stimulus types, CW, N, and CCW (see Fig. 2B). Otherwise, the experimental procedure and derivation of cancellation velocity were identical to the main experiment. Results were based on five repeated staircase sessions for each condition. Our prediction was that the cancellation velocity would become greater as the jitter magnitude increased.

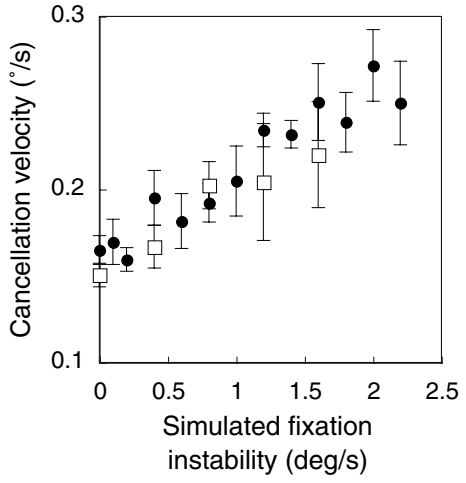

Fig. 6. Differential cancellation velocity plotted as a function of "simulated fixation instability," i.e., standard deviation of velocity distribution. In this experiment, the radial grating on the monitor was actually oscillated according to the specified probability distribution of instantaneous velocity. The error bar indicates the standard error of mean. Solid circles: the first author's data. Open squares: a naïve subject's data.

That was actually the case. The first author's data are shown with solid symbols in Fig. 6. As the simulated fixation instability increases, the cancellation velocity for the same observer increased in a roughly linear fashion (test of linear-regression coefficient, $t_{63} = 8.03$, $p \ll 0.0001$). The same profile was also confirmed for one naïve subject ($t_{23} = 2.66$, $p < 0.02$), whose data are plotted with open symbols.

Therefore, we take all the data described in this section as converging evidence to strengthen the notion that the greater image jitter produces the greater illusion.

4. Subsidiary experiment: Rated illusion strength and fixation instability

The previous experiments used visual stimuli having smooth compound waveforms. We were interested to see whether our finding can be generalized to a wider variety of original illusory figures, such as the illustration in Fig. 1, which lead to a more compelling motion impression than the compound waveform does.

In this subsidiary experiment, 25 such stimulus figures were shown to 53 naïve subjects (39 females and 14 males, aged 20–39), who rated the perceptual strength of motion impression. Fixational eye movements of these subjects were also recorded.

4.1. Methods

4.1.1. Rating of illusory-motion strength

Twenty-five static figures (see the auxiliary file) were selected from a collection of illusory figures designed by one of the authors (A.K.). These static figures all complied with the essential luminance structure that is requisite for the illusion (Fig. 2A), and indeed elicited some impression of slow drifting motion. The stimulus was presented on the color liquid-crystal display of a laptop computer (Apple PowerBook G4). The room was illuminated with office-standard fluorescent lighting. The subject binocularly observed the stimulus from a reading distance (approximately 50 cm) with the head as still as possible. The figures were sequentially presented using graphic presentation software (Microsoft PowerPoint X). Each presentation was 15 s long. The screen remained white in inter-stimulus intervals of 2 s. The subject was requested to give one of the three verbal responses: "moved clearly," "moved faintly," or "did not move at all," which were assigned linear rating scores 1, 0.5, and 0, respectively, in the off-line analysis. There were two conditions, FREE and FIX. In the FREE con-

dition, the subject was allowed to make free eye movements during stimulus presentation. In the FIX condition, the subject was requested to maintain fixation at a greenish stationary spot (with the diameter of 1 cm) at the center of the display. Each subject participated in four experimental sessions, two each for the FREE and FIX conditions, in a counterbalanced order across subjects. We summed actual scores across trials and then divided the sum by the theoretical maximum (50) to obtain the normalized score ranging 0–1.

4.1.2. Eye-movement recording

In a darkroom, eye movements of each subject were recorded in the same day but separately from the rating experiment, under the assumption that fixation instability of each given individual is relatively invariant across separate sessions compared to inter-subject variability (failure of this assumption would only make the data noisier but would not lead to an artifactually significant correlation). The setup was identical to the condition of passive observation in the main experiment.

4.2. Results

We plotted a scattergram (FIX vs. FREE) where one point corresponds to one subject (Fig. 7). Evidently, the figures used in the experiment were more or less effective at eliciting motion impression. The normalized scores for our population of subjects could be reasonably approximated by a Gaussian, making it easy to perform parametric tests.

Illusory motion was more salient in the FREE condition than in the FIX (paired-difference t test, $t_{52} = 3.29$, $p < 0.005$). However, the difference was small; steady fixation did not eliminate the illusion and, for several subjects, resulted in almost the same rating score as in free viewing. Indeed, the averaged ratio of FIX to FREE was 0.934.

Having confirmed fairly compelling illusions under the FIX condition, we plotted an inter-subject scattergram for this condition to see whether illusion strength is correlated with fixation instability (Fig. 8A). A significant positive correlation was found ($r = 0.280$; $t_{50} = 2.07$, $p < 0.05$), provided that the obvious outlier (open symbol) is ignored (if it is included, $r = 0.162$; $t_{51} = 1.17$, *n.s.*).

Of all 53 subjects, all possible combinations of 51 subjects including the outlier always showed much lower r (min 0.084, max 0.227, median 0.159) than all possible combinations of 51 subjects excluding him did (min 0.243, max 0.319, median 0.280), with no overlap between the two distributions. Since his motion-detection performance as assessed by random-dot patterns was normal, his small rating cannot be explained by insensitivity to motion. As each subject has his or her own criterion about illusion strength, the outlier might have been extremely conservative in giving rating scores.

The same analysis was also made between rating score and the frequency of microsaccades (Fig. 8B); there was no indication of correlation if the outlier was included ($r = 0.053$, *n.s.*) or not ($r = 0.127$, *n.s.*).

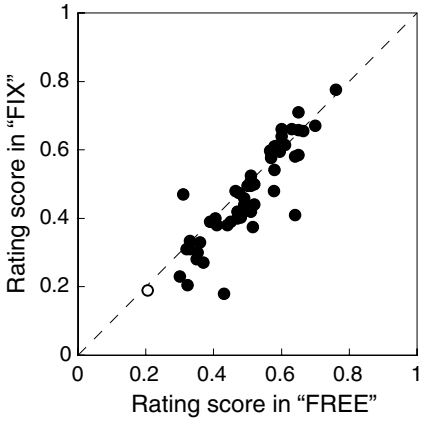

Fig. 7. Results of illusion-strength rating. Data from 53 subjects are plotted in a scattergram, one point corresponding to one subject. The diagonal line indicates $y = x$. The open symbol indicates the point of the outlier plotted in Fig. 8.

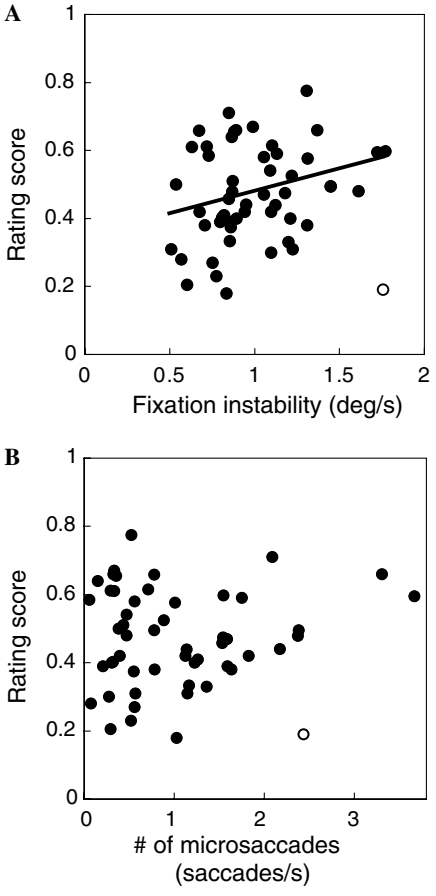

Fig. 8. Scattergram to visualize inter-subject correlation between rating score in the FIX condition and fixation instability. Data from the 53 subjects are plotted, one point corresponding to one subject. The open symbol indicates the outlier. (A) Inter-subject scattergram between rating score and fixation instability. The solid line indicates the linear regression: $y = 0.132x + 0.342$. (B) Inter-subject scattergram between rating score and microsaccade frequency (number per second).

5. Discussion

The present study examined the relationship between the motion illusion perceived in static figures and fixational eye movements. We found a positive correlation such that the poorer fixaters see the stronger illusion.

5.1. Effects of retinal motion due to fixational eye movements

The present findings suggest certain contributions of fixational eye movements to the occurrence of the motion illusion. The fixational eye movements are incessantly making retinal image motions, but we are unaware of such movements of the visual field. Previously we have shown psychophysical evidence for a cortical mechanism by which a stable visual world is maintained despite these eye movements and demonstrated that specially arranged visual stimulation can sometimes confuse this brain mechanism (Ashida, 2002; Kitaoka, 2002; Murakami, 2003; Murakami & Cavanagh, 1998, 2001; Sasaki, Murakami, Cavanagh, & Tootell, 2002). We consider that the illusory slow drift investigated in the present study is also one such rare occasion where our visual system fails to cancel the spurious image motions generated by small eye movements.

If this is true, the illusion should be abolished in the stabilized retinal image. We informally tested this prediction by casual observation of the afterimage of the stimulus. Subjects (the three authors and three naïve subjects in separate sessions) faced on the wall posted with a printed version of a typical illusory figure (Fig. 1) or its control figure (which is based on Fig. 1 except that the luminance relationship is flipped between adjacent local texture elements, so that illusory motion between them cancels each other). In normal illumination, the illusory figure yields a vivid impression of rotary motions for all subjects, whereas the control figure never appears to move. In a total darkness, the experimenter triggered a strobe light (COMET TW-04II) to illuminate the stimulus with an intense flash (decay time <1.5 ms). All subjects subsequently experienced the afterimage of the figure in the dark for around 10 s, but they consistently reported that there was no illusory rotation in the afterimage, or that the illusory and control figures made no difference, even though the afterimage of the figure as a whole could appear to move about synchronously with large-scale eye movements.

Fixational eye movements consist of three major types: drifts, tremors, and microsaccades. In the present study, fixation instability is represented by drifts. As the illusion has been found to correlate with this statistic, eye-drift speed is considered to be one of the major sources of that illusion. What about other two types of small eye movements? The involvement of tremors in visual perception is unclear (Martinez-Conde, Macknik, & Hubel, 2004), but it is highly unlikely that such tiny and rapid oscillations with frequencies over 30 Hz mediate perception of slow illusory motions. The involvement of microsaccades is possible, but our data were negative: there were no significant correlations between the illusion strength and the frequency of microsaccades. We also point out that Zanker et al. recently attempted to find a correlation between the strength of motion illusion in an Op Art figure and the frequency of microsaccades. While contribution of fixational eye movements to illusory motion was demonstrated psychophysically (Zanker, Doyle, & Walker, 2003) and computationally (Zanker, 2004; Zanker & Walker, 2004), their experimental data did not specifically reveal clear dependence of illusory motion strength on the statistics of microsaccades, leaving the question open as to what component of fixational eye movements is really relevant. The present finding suggests that the statistics of fixational eye drifts rather than microsaccades might be a better descriptor of the illusion strength of Op Art figures. However, the Op Art patterns used by Zanker et al. have a fundamentally different layout and therefore do not generate a biased motion illusion as shown in the present study. There might be different mechanisms involved in motion illusions elicited by different types of static patterns.

5.2. Possible accounts for biased motion responses

Our findings implicate retinal image motions as some kind of "power supply" of the illusory motion, but we have not specified what underlying mechanism is responsible for it, or why the illusory motion appears so smooth in spite of random image fluctuation due to small eye movements. We believe that the specific luminance pattern (Fig. 2A) used in the illusory display is the key to these questions. One scenario proposed by Conway et al. is that each small eye movement somehow refreshes retinal stimulation, promoting new onset responses in the stimulated area on the retina (Conway, Kitaoka, Yazdanbakhsh, Pack, & Livingstone, 2005). As the response latency is known to be longer for lower contrasts, the response to the dark (or light) gray region will come later than the response to the black (or white) region. These neighboring pairs of responses will artificially create the "phi" movement (Anstis, 1970). Backus and Oruç also explained the illusion on the basis of differential latency among visual responses to luminance levels, and introduced the effect of adaptation to model the smooth motion perception over many seconds (Backus & Oruç, 2005).

However, this "latency-difference" scenario has a difficulty in explaining why only the fixational drift, but not the fixational saccade, positively correlated with illusion strength in our data, for each microsaccade should be refreshing retinal stimulation and should enhance the illusion. Thus, we must also consider another scenario, in which random retinal velocity due to incessant fixational drifts is converted to smooth perceived motion. In a normal stimulus display, random image motions will be registered as such, in the same instantaneous velocity everywhere. The motion processing system perhaps dismisses such common image motions under the ecological assumption that uniform motions originate in eye

movements in most cases. The illusory display with the specific luminance pattern, however, confuses early motion detectors, such that a particular direction of motion is artificially boosted whereas other directions are attenuated (Murakami, Kitaoka, & Ashida, 2004). One such possibility is argued below.

The illusory figure comprises asymmetric gradients. In Fig. 2A, the luminance steps are always larger on the left of the highest-contrast parts (i.e., white and black) than on their right. This spatial asymmetry results in a temporal asymmetry by eye movements. At around these large steps, the local contrast increases and decreases when the pattern moves leftward and rightward, respectively (see also Fig. 10). We can explain the effects of such asymmetry by the gradient scheme of motion detection, in which velocity is approximated by dividing the temporal derivative by the spatial derivative of intensity at each location (e.g., Benton & Johnston, 2001; Harris, 1986; Johnston, McOwan, & Buxton, 1992; Marr & Ullman, 1981; Mather, 1984; Bruce, Green, & Georgeson, 2003). In this scheme, overestimation of temporal gradient directly results in overestimation of velocity.

A negative bias in the derivative operator would give rise to such overestimation (Fig. 9). The temporal gradient might be biologically estimated by a smooth, biphasic temporal impulse response (Fig. 9A, gray curve), but here let us consider the simplest case, i.e., a two-point discrete impulse response (solid bars). Its operation is to calculate the difference in intensity between $\tau 1$ (e.g., the present) and $\tau 2$ (e.g., sometime in the past). Here, the convolution of the image with the differentiator kernel is simply the scalar product of the relative image intensity and the temporally reversed kernel at two peaks. For instance, when the image intensity decreases and approaches the mean luminance (the upper left of Fig. 9B, gray line), the differentiator without bias would simply calculate the difference between times designated as 1 and 2, i.e., the actual slope. Now let us assume that the negative lobe of the kernel is greater than the positive lobe, so that the relative intensity at $\tau 2$ is slightly "boosted" before subtraction. As a result, the estimated slope becomes artificially steeper (black line). The same overestimation of temporal gradient also happens when the relative intensity increases from darker to zero (lower left). Note that in both cases, the contrast is decreasing with time. On the other hand, the opposite thing happens when the contrast increases with time (upper right and lower right): the estimated slope becomes shallower than actual, resulting in underestimation of temporal gradient.

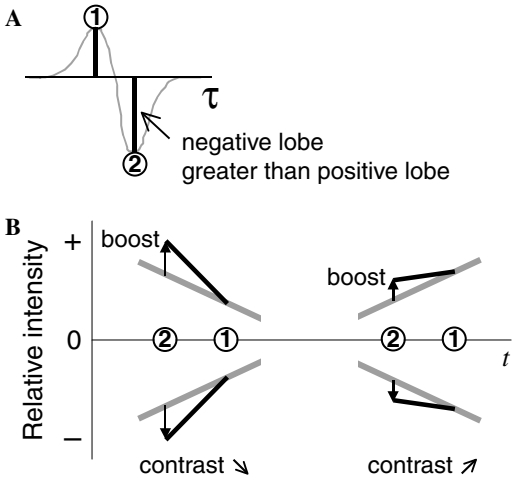

Fig. 9. Schematics of the negatively biased temporal kernel and its behavior. (A) The simplified differentiator. The temporal impulse response is indicated by the solid bars, whereas a biologically more plausible smooth function is drawn in the gray curve. (B) The temporal gradients estimated by the negatively biased kernel. Actual intensity changes are indicated by the gray lines, whereas estimated slopes are indicated by the black lines. The term "relative intensity" refers to the image intensity relative to the mean luminance. The term "contrast" refers to the absolute deviation from the mean luminance. For the sake of simplicity of schematic, the unit is unspecified for these quantities.

Fig. 10 further illustrates how the biased temporal kernel overestimates rightward motion in the pattern of Fig. 2A. Due to small eye movements, the retinal image of the static stimulus is assumed to move sometimes to the left, and sometimes to the right (top row). Velocity estimation is assumed to take place in a spatially low-pass-filtered version of the image (second row) and, in particular, at those black curves where spatial gradients are steep enough (i.e., where the denominator in the velocity calculation is the most reliable; see Marr & Ullman (1981)). The third and fourth rows show estimated spatial and temporal gradients, respectively. The open circles indicate the veridical estimates. As they are symmetric about vertical between the left and right panels, the estimated velocities are of the same size with opposite signs. If the temporal kernel is negatively biased, however, the estimation of temporal gradient is biased, as indicated by the filled circles. For the leftward motion, the local contrast increases with time, so the temporal gradients are underestimated. The opposite thing happens for the rightward motion. Therefore, the velocity estimates are no longer symmetric:

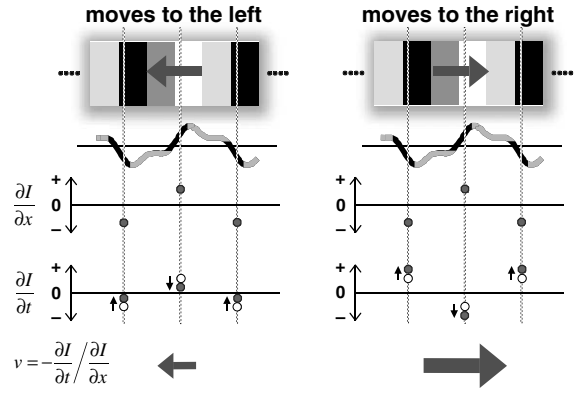

Fig. 10. Explanation of the illusion by the gradient model with the negatively biased temporal kernel. Variables I, x, and t denote intensity, horizontal dimension, and time, respectively.

it becomes smaller for the left and larger for the right. This imbalance accords with the illusion of rightward motion.

The bias in the temporal derivative operator is biologically plausible, because perfect balance would be difficult to assure in biologically implemented causal filters. Such a negative bias in the temporal impulse response has been implicitly suggested by some studies (e.g., Burr & Morrone, 1993; Kelly, 1961).

Though we tentatively propose the gradient model as a possible account, similar accounts based on other motion-detection schemes such as the motion energy model (Adelson & Bergen, 1985) might not be impossible (Mather, 1994). We are also aware of several important points that are yet to be determined. First, the feasibility of our explanation is still wanting in convincing empirical tests. Second, the model does not explain why the illusion is more salient in peripheral viewing. Third, the output might depend on the choice of sampling locations of spatial and temporal derivatives. To address these issues, we are currently engaged in elaboration of the model with parametric computer simulations and phenomenological observations. Finally, it is still an open question as to how these elementary velocity representations in low-level motion processing are spatially and temporally integrated within an image and segregated from its surround to support our perception of global illusory motion. Our ultimate goal is to propose a unified theory incorporating the present study with recent physiological (Ölveczky, Baccus, & Meister, 2003), psychophysical (Murakami & Cavanagh, 1998), and computational (Fermüller, Pless, & Aloimonos, 1997; Zanker, 2004) studies examining the relationship between perception and fixation instability.

Acknowledgments

A.K. and H.A. were supported by a Grant-in-Aid (the Heisei 14th grant entitled ''Internet edutainment using anomalous motion illusions'') awarded from the Hayao Nakayama Foundation for Science, Technology and Culture. A.K. was supported by a Grant-in-Aid for Scientific Research, Ministry of Education, Science and Culture, Nos. 14310045 and 16530480. I.M. was supported by the Center for Evolutionary Cognitive Sciences at the University of Tokyo and by Nissan Science Foundation.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.visres.2006.01.030.

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.

Anstis, S. M. (1970). Phi movement as a subtraction process. *Vision Research, 10*, 1411–1430.

Ashida, H. (2002). Spatial frequency tuning of the Ouchi illusion and its dependence on stimulus size. *Vision Research, 42*, 1413–1420.

Ashida, H., & Kitaoka, A. (2003). A gradient-based model of the peripheral drift illusion. *Perception, 32 Suppl.*, 106.

Backus, B. T., & Oruç, I. (2005). Illusory motion from change over time in the response to contrast and luminance. *Journal of Vision, 5*, 1055–1069.

Bair, W., & O'Keefe, L. P. (1998). The influence of fixational eye movements on the response of neurons in area MT of the macaque. *Visual Neuroscience, 15*, 779–786.

Benton, C. P., & Johnston, A. (2001). A new approach to analysing texture-defined motion. *Proceedings of Royal Society of London. B. Biological Sciences, 268*, 2435–2443.

Bruce, V., Green, P. R., & Georgeson, M. A. (2003). *Visual perception* (4th ed.). Hove: Psychology Press.

Burr, D. C., & Morrone, M. C. (1993). Impulse–response functions for chromatic and achromatic stimuli. *Journal of the Optical Society of America A, 10*, 1706–1713.

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.

Faubert, J., & Herbert, A. M. (1999). The peripheral drift illusion: A motion illusion in the visual periphery. *Perception, 28*, 617–621.

Fermüller, C., Pless, R., & Aloimonos, Y. (1997). Families of stationary patterns producing illusory movement: Insights into the visual system. *Proceedings of Royal Society of London. B. Biological Sciences, 264*, 795–806.

Fraser, A., & Wilcox, K. J. (1979). Perception of illusory movement. *Nature, 281*, 565–566.

Harris, M. G. (1986). The perception of moving stimuli: A model of spatiotemporal coding in human vision. *Vision Research, 26*, 1281–1287.

Johnston, A., McOwan, P. W., & Buxton, H. (1992). A computational model of the analysis of some first-order and second-order motion patterns by simple and complex cells. *Proceedings of Royal Society of London. B. Biological Sciences, 250*, 297–306.

Kelly, D. H. (1961). Visual responses to time-dependent stimuli. II. Single-channel model of the photopic visual system. *Journal of the Optical Society of America, 51*, 747–754.

Kitaoka, A. (2002). A classification of anomalous motion illusions. *Visual Localization in Space-Time 2002* (p. 22). Brighton.

Kitaoka, A., & Ashida, H. (2003). Phenomenal characteristics of the peripheral drift illusion. *VISION, 15*, 261–262.

Marr, D., & Ullman, S. (1981). Directional selectivity and its use in early visual processing. *Proceedings of the Royal Society of London B, 211*, 151–180.

Martinez-Conde, S., Macknik, S. L., & Hubel, D. H. (2004). The role of fixational eye movements in visual perception. *Nature Reviews Neuroscience, 5*, 229–240.

Mather, G. (1984). Luminance change generates apparent movement: Implications for models of directional specificity in the human visual system. *Vision Research, 24*, 1399–1405.

Mather, G. (1994). Motion detector models: Psychophysical evidence. In A. T. Smith & R. J. Snowden (Eds.), *Visual detection of motion* (pp. 117–143). London: Academic Press.

Murakami, I. (2003). Illusory jitter in a static stimulus surrounded by a synchronously flickering pattern. *Vision Research, 43*, 957–969.

Murakami, I. (2004). Correlations between fixation stability and visual motion sensitivity. *Vision Research, 44*, 751–761.

Murakami, I., & Cavanagh, P. (1998). A jitter after-effect reveals motion-based stabilization of vision. *Nature, 395*, 798–801.

Murakami, I., & Cavanagh, P. (2001). Visual jitter: Evidence for visual-motion-based compensation of retinal slip due to small eye movements. *Vision Research, 41*, 173–186.

Murakami, I., Kitaoka, A., & Ashida, H. (2004). The amplitude of small eye movements correlates with the saliency of the peripheral drift illusion. *Society for Neuroscience Annual Meeting Abstracts, 30*, 302.

Naor-Raz, G., & Sekuler, R. (2000). Perceptual dimorphism in visual motion from stationary patterns. *Perception, 29*, 325–335.

Ölveczky, B. P., Baccus, S. A., & Meister, M. (2003). Segregation of object and background motion in the retina. *Nature, 423*, 401–408.

Rucci, M., & Desbordes, G. (2003). Contributions of fixational eye movements to the discrimination of briefly presented stimuli. *Journal of Vision, 3*, 852–864.

Sasaki, Y., Murakami, I., Cavanagh, P., & Tootell, R. B. H. (2002). Human brain activity during illusory visual jitter as revealed by functional magnetic resonance imaging. *Neuron, 35*, 1147–1156.

Snodderly, D. M., Kagan, I., & Gur, M. (2001). Selective activation of visual cortex neurons by fixational eye movements: Implications for neural coding. *Visual Neuroscience, 18*, 259–277.

Taylor, M. M., & Creelman, C. D. (1967). PEST: Efficient estimates on probability functions. *Journal of the Acoustical Society of America, 41*, 782–787.

Zanker, J. M., Doyle, M., & Walker, R. (2003). Gaze stability of observers watching Op Art pictures. *Perception, 32*, 1037–1049.

Zanker, J. M., & Walker, R. (2004). A new look at Op art: Towards a simple explanation of illusory motion. *Naturwissenschaften, 91*, 149–156.

Zanker, Z. (2004). Looking at Op Art from a computational viewpoint. *Spatial Vision, 17*, 75–94.