

Understanding the Rotating Snakes illusion

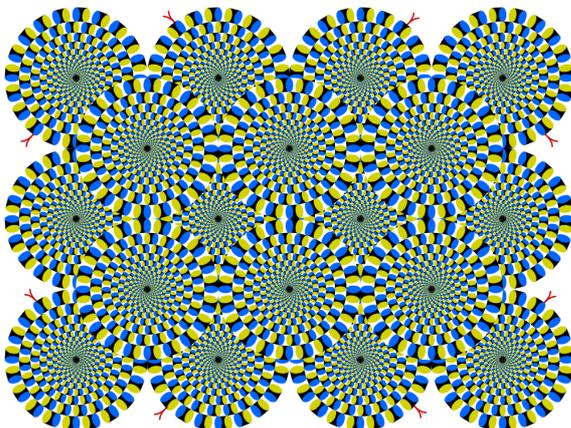
Student: Martin O'Reilly[†]; Supervisor: Alan Johnston

This paper explores potential mechanisms underlying Rotating Snakes - a piece of Op-Art by Kitaoka evoking a striking apparent motion illusion. The Op-Art is based on an illusion developed by Kitaoka & Ashida (2003) and is similar to previous illusions developed by Fraser & Wilcox (1979), Faubert & Herbert (1998) and Naor-Raz & Sekuler (2000). The subsequent illusions, including Rotating Snakes are considered by their authors to be related to the original Fraser & Wilcox "Escalator" illusion and are collectively termed Peripheral Drift illusions. Despite their similarities, none of the mechanisms proposed to date can explain both the "single gradient" members of this class of illusion (exemplified by the Faubert & Herbert illusion) and the "double gradient" members of this class (namely the Rotating Snakes illusion). A mechanism based on luminance-dependent latency proposed by Faubert & Herbert explains the "single gradient" illusions well but does not explain the "double gradient" illusions. It is also not clear how biologically plausible such a mechanism is given the centre-surround structure of retinal ganglion cell receptive fields. Two similar mechanisms based on contrast-dependent latency explain the "double gradient" illusions but fail to explain the single gradient illusions. Conway et al (2005) describe the effect of this latency in terms of forward and reverse phi motion, while Backus and Oruc (2005) explicitly model the temporal response of V1 neurons as a function of contrast. It is likely that these two approaches are describing the same underlying mechanism. Johnston et al (1998; 2003) have developed a high-level Gaussian derivative model of motion, which has successfully detected "illusory" contrast-driven second order motion in moving patterns. The contrast-dependent latency described above could be considered a gradual switch from an input dominated by high-contrast elements to one dominated by low-contrast elements. Therefore a simplified Rotating Snakes pattern was contrast-modulated and passed through the Gaussian derivative model to see if any motion was detected. Despite preliminary theoretical analysis suggesting that the model would detect motion in the direction of the illusion for a stimulus decreasing in contrast over time, the motion reported by the model was excessively large and inconsistent in direction. It seems likely that any perceived motion is an artefact of the model. Further work will be necessary to identify and eliminate any causes of any anomalous motion before conclusions can be drawn regarding whether a Gaussian derivative based motion model will see the Rotating Snakes Illusion.

Introduction

The Rotating Snakes illusion, produced as a piece of Op-Art by Akiyoshi Kitaoka in 2003 and shown in figure 1, belongs to the class of "Peripheral Drift" illusions. These illusions evoke the perception of apparent motion from stationary images viewed in the periphery of the visual field, and the Rotating Snakes illusion is a particularly strong example of its type. The illusion should be viewed in the peripheral vision and can be sustained by eye movement, blinking or flashing the image. The use of colour in the illusion has been reported by Kitaoka to enhance the effect of the illusion, but a greyscale version of the illusion still evokes a strong perception of rotation (see figure 1b). In fact, the only study to formally evaluate the effect of colour on the strength of the illusion reported no significant effect (Conway et al, 2005).

A



B

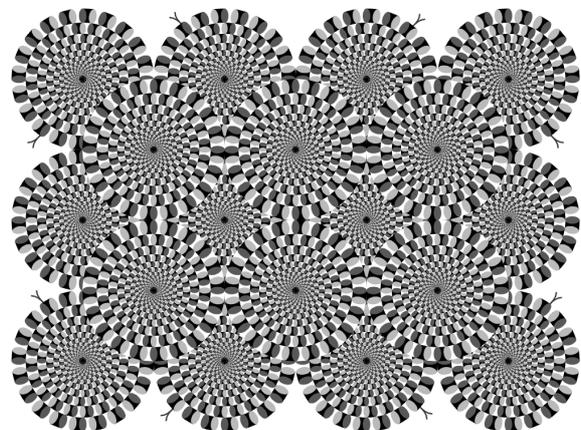


Figure 1: Kitaoka's Rotating Snakes illusion. **A:** Original colour Op-Art. The inclusion of colour appears to strengthen the effect of the illusion. **B:** A greyscale version of the illusion, which still evokes a strong perception of apparent motion. [Source: <http://www.ritsumeai.ac.jp/~akitaoka/rotsnake.gif>]

The fundamental building block of this illusion is a sequence of 4 elements of varying luminance as shown in figure 2. Motion is almost universally perceived in the direction of Black→Dark Grey→White→Light Grey (Black→Blue→White→Green for the colour version). It is presumed that this sequence generates local motion signals in the visual system. However, the large scale organisation of this fundamental building block is critical to this illusion. Informal testing suggests that an isolated block or even a line of repeated blocks do not evoke a strong motion illusion. However, when the element is

[†] MRes student at UCL CoMPLEX (Centre for Mathematics and Physics in the Life Sciences and Experimental Biology).
Email: martin.o'reilly@ucl.ac.uk

repeated in the concentric ring pattern of the full illusion, strong illusory motion is observed. The illusion appears to rely on the aggregation of a collection of local motion cues into coherent global motion. In the case of the concentric ring pattern, the presumed local motion generated by each ring is consistent with the global rotation of a fixed disc.

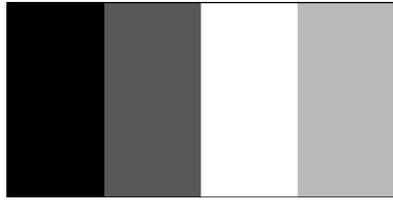
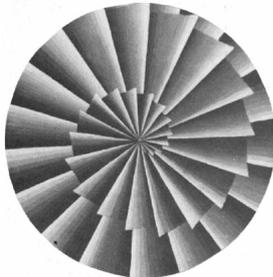


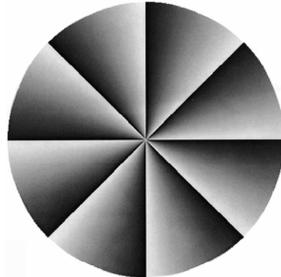
Figure 2: The fundamental building block of the Rotating Snakes illusion. Motion is observed in the direction of Black→Dark Grey→White→Light Grey (left to right in the above example).

Related illusions

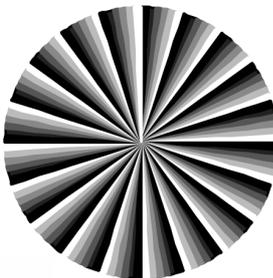
A: Fraser & Wilcox



B: Faubert & Herbert



C: Naor-Raz & Sekuler



D: Kitoaka & Ashida

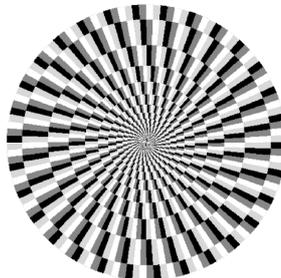


Figure 3: Successive enhancements of the Peripheral Drift illusion. **A:** The original Fraser-Wilcox (1979) Escalator illusion; **B:** The Faubert & Herbert (1998) smooth gradient Peripheral Drift illusion; **C:** The Naor-Raz & Sekuler (2000) stepped gradient Peripheral Drift illusion; **D:** The Kitoaka & Ashida (2003) enhanced Peripheral Drift illusion. In all 4 illusions, clockwise motion is usually perceived. [Source: Images taken from their respective papers as referenced above]

Fraser & Wilcox (1979)

One of the earliest examples of a Peripheral Drift illusion is the "Escalator" illusion described by Fraser & Wilcox (1979), although a similar effect was reported by Bülthoff & Gotz the same year. The illusion consists of a spiralling "staircase" of triangular wedges with a dark to light gradient fill (see figure 3a). Fraser & Wilcox performed a large psychophysics study, encompassing a total of 678 individuals. When presented with the illusion and asked to comment on it, with no priming to expect motion, most subjects (59.0%) perceived rotation in the direction of the dark-to-light gradient. However a significant minority (24.9%) perceived no motion. Yet others perceived motion either in the dark-to-light direction (6.5%) or

sometimes in one direction and sometimes in the other (9.6%).

The results of further tests with rotating patterns indicate that the illusory motion is integrated with real motion to produce a composite perception of apparent motion. The rotating pattern study also suggested that subjects who perceived no motion in the static illusion might simply have a much lower sensitivity to the illusion. This suggests that the strength of perception of the illusion in the populace may span a continuous range. Finally, a parent-offspring survey and twin study provide intriguing evidence for a genetic component to how the illusion is perceived.

Fraser & Wilcox do not suggest a mechanism for the illusion, although they do focus on the direction of the fill gradient as the key feature of the pattern that varies with the direction of perceived motion. However, the direction of the spiral also differs between the two pattern types and could conceivably provide a separate motion cue that might either reinforce or conflict with any cue provided by the direction of the fill gradient.

Faubert & Herbert (1999)

Faubert & Herbert were the first to coin the term "Peripheral Drift" to describe this class of illusion. They simplified the original Fraser-Wilcox "Escalator" by removing the staircase effect and reducing the number of segments (see figure 3b). The removal of the staircase spiral also removed any direction cue that it may have provided and it is interesting to note that all but one of the reported subjects observed motion of this simplified illusion in the direction of the dark-to-light gradient.

The formal psychophysics study performed by Faubert & Hubert is very small scale, comprising only the 2 authors and 3 other subjects. The very small number of independent subjects in this study does not provide a solid evidential basis for drawing generalised conclusions about the strength and consistency of the illusion. However, presentation of the illusion to over 200 conference goers via a poster display provided strong but informal corroboration of the findings of the study. All 5 formal subjects and all but 2 informal subjects perceived motion in the dark to light direction. However, approximately 50% of subjects (including 3 naive formal subjects) did not report motion on their initial viewing. This is almost twice the proportion of subjects who reported seeing no motion in Fraser & Wilcox's study.

The decreased incidence of unprompted perception could be interpreted as evidence that the Faubert & Herbert illusion is actually weaker than the original "Escalator" pattern of Fraser & Wilcox. However the increased consistency in direction of apparent motion when perceived provides evidence in the opposite direction, indicating that the simplified Faubert & Herbert pattern elicits a stronger motion effect. Given the obvious differences in the presentation environment between Fraser & Wilcox's formal psychophysics study and Faubert & Herbert's informal poster testing, it is difficult to draw firm conclusions from the above comparisons.

Naor-Raz & Sekuler (2000)

Naor-Raz & Sekuler's simplification of Fraser & Wilcox's "Escalator" illusion was very similar to that of Faubert & Herbert. The key differences are that their pattern utilises 6 discrete luminance steps rather than

Faubert & Herbert's continuous luminance gradient and retains the same sized wedges as Fraser & Wilcox, resulting in 20 wedges compared to Faubert & Herbert's 8 (see figure 3c).

Like Faubert & Herbert, Naor-Raz & Sekuler report that approximately 50% of subjects failed to perceive motion in their illusion unprompted. Their psychophysics studies were limited in size, with only 22 naïve subjects exposed to the illusion. Of these 68% perceived motion in the dark-to-light direction, with the remaining 32% either seeing no motion or motion in the reverse direction. These proportions are in line with those originally observed by Fraser & Wilcox. In contrast, the consistency of the direction of perceived motion reported by Faubert & Herbert is essentially 100% in the dark-to-light direction.

Given that the majority of the subjects reported in Faubert & Herbert's paper were informally exposed to the illusion via a poster display, it is possible that there was some indication of the expected or "normal" direction of perceived motion within the poster contents. However, given that Faubert & Herbert reported the same proportion of unprompted observers not seeing any motion on their initial viewing, any such indication of expected direction of motion would need to be subtle enough not to prime a naïve observer to expect motion on their first viewing yet strong enough to strongly bias the direction of perceived rotation once primed to expect motion.

Naor-Raz & Sekuler investigated the effect of several properties of stimulus presentation on the perceived strength of the illusion. They found that the consistency of the perceived motion (percentage of time motion was seen) increased with both increasing duration of stimulus display and with increasing peripheral stimulus location. For both properties, the increase in consistency appears to plateau. They also found that subjective ratings of the strength of the illusion increased as the contrast of the pattern was increased. Finally, they ruled out small movements of the image on the retina during refocusing as a cause of the perceived motion by confirming that the illusion was perceived by an aphakic¹ observer.

Kitaoka & Ashida (2003)

Kitaoka & Ashida enhanced the illusion further. They initially simplified Naor-Raz and Sekuler's stepped illusion by reducing the number of luminance levels to 4. They then reversed the order of half the elements and fragmented the long edges to produce a stronger illusion (figure 3d). Later in 2003 Kitaoka further developed this illusion into the Rotating Snakes Op-Art. It should be noted that Kitaoka & Ashida report no detail of any formal or informal testing, instead simply asserting that their successive changes enhance the strength of the illusion. In the absence of any claim to the contrary, it must be assumed that these assertions are supported only by informal, subjective evaluations of illusion strength by a limited number of subjects. However, the final Rotating Snakes illusion was subsequently more formally evaluated by Conway et al (2005); Hisakata & Murakami (2008) and Kuriki et al (2008). These studies are discussed further in the next section.

It should be noted that it has not been conclusively demonstrated that this family of illusions

are all caused by the same neurological mechanism. No proposed mechanism to date can explain all the above illusions. However, the authors reporting each successive illusion have all claimed them to be related to the original Fraser-Wilcox "Escalator" illusion.

Properties of the Rotating Snakes illusion

Conway et al (2005)

The earliest studies to formally examine the Rotating Snakes illusion were reported in a 2005 paper by Conway et al. The first was a psychophysics study which examined the effect of flashing individual combinations of the luminance elements making up the illusion (black; dark-grey; white; light-grey - see figure 4a). Initial trials in which the stimuli were simply flashed on or off did not bias the reported perception of motion in either direction. The main trial used slightly adapted "oscillating" stimuli where alternate frames were offset such that the grey gaps in one frame were replaced by the element pairs in the next. This sequence provides a strong motion cue from frame to frame, but one that is ambiguous in direction. In these trials there was a consistent bias in the direction of the illusion. While there were some element pairs for which the bias was not significant for all subjects, averaged across all subjects and all element pairs the bias was significant overall. This latter trial was repeated using the element pairs from the colour version, and this was found to have no significant effect on the size of the bias (see figure 4b). The size of these studies was relatively low, with 10 naïve observers for both greyscale and colour trials. It should be noted that 6 observers participated in both trials but, with the right controls, could likely still be considered naïve for the second trial.

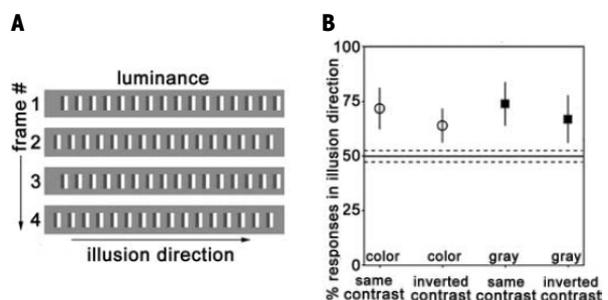


Figure 4: A: Greyscale "oscillating" stimuli used in Conway et al (2005) psychophysics study. B: Summary of the results for this study, showing a consistent perceived motion bias in the direction of the illusion averaged across all element pairs and all subjects. [Source: Conway et al, 2003]

The second study reported in the Conway et al (2005) paper was an electrophysiological study designed to test their hypothesised mechanism for the illusion, which is evaluated in the next section. Recordings were taken from directionally selective cells in V1 and MT in 3 macaque monkeys. When each of the static greyscale luminance elements were presented to the monkeys in isolation, the time from stimulus onset to peak response was shorter for the "high contrast" white and black stimuli than for the "low contrast" light and dark grey stimuli (see figure 5a). The monkeys were then presented with isolated static element pairs. The results indicate that the responses of strongly directionally selective cells were significantly higher for element pairs presented such that the direction of motion in the illusion was in the cell's preferred direction (see figure 5b).

¹ Aphakia is the absence of the eye's lens.

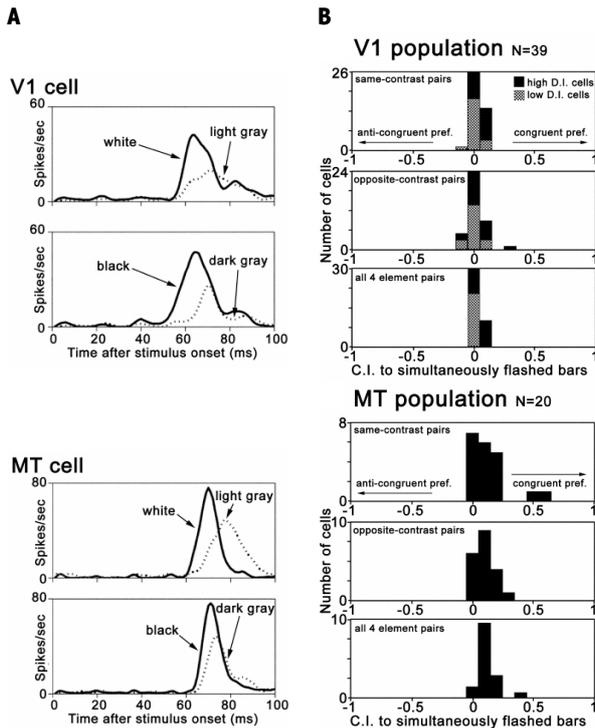


Figure 5: A: Response of sample macaque V1 and MT neurons to static presentations of single luminance elements from the Rotating Snakes illusion **B:** "Congruence Index" (C.I.) of populations of macaque V1 and MT cells to static presentations of pairs of luminance elements. For strongly directionally selective cells (high D.I. cells in plot) there is a significant increase in the response of the cells to element pairs presented such that the direction of illusory motion is in the cell's preferred direction (congruent) compared to the response when the direction of illusory motion is opposite to the cell's preferred direction (anti-congruent). The greater the difference in response, the greater the congruency index, which can range from 0 (no difference) to 1 (response to congruent element pair dwarfs response to anti-congruent element pair). Precise definitions of Congruency Index (C.I.) and Direction Index (D.I.) are in the original paper. [Source: Conway et al, 2005]

Hisakata & Murakami (2008)

A more recent psychophysical study by Hisakata & Murakami explored the effect of eccentricity and stimulus brightness. This study used simplified stimuli rotating at a range of velocities to measure "response functions" for perceived motion in response to a simplified Rotating Snakes pattern. The simplified stimulus and sample response functions can be seen in figure 6. These functions plot the probability of a judgement of counter-clockwise motion against true clockwise and counter-clockwise rotational velocity. The point at which the response curve passes a probability of 0.5 can be considered to be when the pattern was perceived as stationary. Therefore the velocity at that point represents the velocity of real "cancellation" rotation required to counter the illusory motion evoked by the pattern. Therefore, this "cancellation" velocity can be considered a quantitative measure of the strength of the illusion. The stronger the perceived illusory motion, the larger the shift of the response curves left (for counter-clockwise stimulus) or right (for clockwise stimulus). The study found that the illusion was stronger when viewed in the periphery of the visual field and for brighter stimuli.

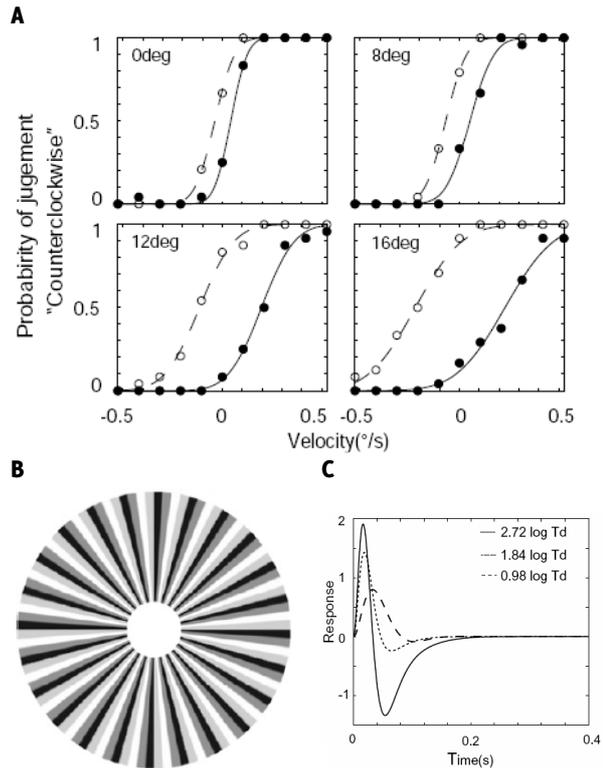


Figure 6: A: Sample "response functions" for perceived motion used to qualitatively measure the strength of illusory motion in Hisakata & Murakami (2008). The stronger the perceived illusory motion, the larger the shift of the response curves left (for counter-clockwise stimulus - open circles) or right (for clockwise stimulus - closed circles). **B:** Sample simplified stimulus used in the study. Illusory motion is in clockwise direction. **C:** Temporal impulse response function for one subject at 3 different levels of stimulus brightness. [Source: Hisakata & Murakami, 2008]

A supplementary study measured the variation in the shape of the temporal impulse response (TIR) function with stimulus brightness. It found that the TIR becomes progressively less biphasic with lower stimulus brightness (see figure 6C). After fitting these measured TIRs to a sum of transient and sustained components, they plotted the power spectra for these components at each brightness level. They found that the relative contribution of the transient component fell with decreasing stimulus brightness, leading them to postulate that the transient component of the TIR contributed to the strength of the Rotating Snakes illusion.

The Hisakata & Murakami studies were well designed, managing to produce a quantitative measure of the strength of the illusion and confirm anecdotal reports that the illusion was stronger in the visual periphery and with brighter stimuli. However, the number of subjects was very limited, comprising only 3 naive observers and one of the authors.

Kuriki et al (2008)

A recent fMRI study by Kuriki et al provides partial support for the earlier electrophysiological results reported by Conway et al (2005). BOLD² signals from 11 subjects, including the 4 authors, were recorded

² The intensity of a pixel in an fMRI image is dependent on the difference between the measured levels of blood oxygenation between two conditions. This measured difference is known as the Blood Oxygen Level Dependent (BOLD) signal.

when the subjects were looking at either the Rotating Snakes illusion or a control stimulus. The control stimulus was constructed by reversing the direction of every other 4 element "building block" in the Rotating Snakes illusion, thus ensuring that local motion signals averaged to zero over the pattern as a whole. Three conditions were explored. Recordings were made from a sub-region of MT which responded significantly to motion (hMT+) and a sub-region of V1 corresponding to the peripheral visual field. In the first, subjects fixated on a central cross and made no eye movements. In this condition, the average BOLD response to Rotating Snakes was not significantly higher than the response to the control in either hMT+ or V1. In the second and third conditions, subjects were asked to make a pre-set sequence of saccades between three fixation points, with the timing prompted by an attentional task (guided eye motion) or under their free control (natural eye motion). In both these conditions, the average BOLD response in hMT+ were significantly stronger for Rotating Snakes than for the control, but BOLD responses in V1 did not significantly differ between the two stimuli. These results are summarised in figure 7, alongside a sample BOLD signal trace for the guided eye motion condition.

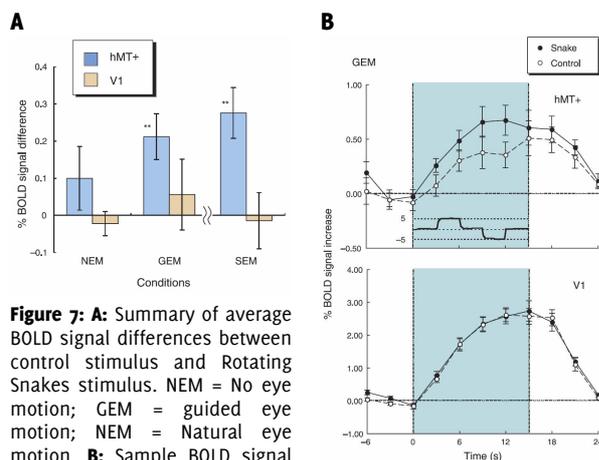


Figure 7: A: Summary of average BOLD signal differences between control stimulus and Rotating Snakes stimulus. NEM = No eye motion; GEM = guided eye motion; SEM = Natural eye motion. **B:** Sample BOLD signal trace for GEM condition [Source: Kuriki et al, 2008]

While these results support Conway et al's electrophysiological finding that some neurons in MT appear to respond to the Rotating Snakes illusion, they do not support their corresponding finding for V1. This lack of a measurable fMRI response in V1 might be due to differences in the organisation of directionally selective cells in V1 compared to hMT+. A clear sub-region of MT (hMT+) was identified that responded significantly to motion. However, it is not clear if Kuriki et al isolated a directionally-selective sub-region of V1. It would seem that the flickering checkerboard stimulus used to determine the peripheral sub-region of V1 would excite orientation detectors in V1 at least as much as motion detectors, if not more so. However, it is difficult to think of a stimulus that would excite low-level motion detectors without also exciting orientation detectors. It is also possible that directionally sensitive cells are not sufficiently localised within V1 to generate a significant fMRI motion response, with directionally selective cells intermingled with non-directionally selective cells. Finally, given the differences observed by Conway et al in the strength of directional activity in V1 cells compared with MT cells, it is perhaps not

surprising that no significant difference in fMRI response was observed. In the Conway study all MT neurons exhibited extremely high directional selectivity. In contrast, approximately 50% of V1 cells were classified as weakly selective and showed no significant difference in their response to congruent and anti-congruent element pairs. Even "strongly selective" V1 cells were much less selective than the MT cells.

Potential mechanisms for the illusion

When discussing potential mechanisms for the Rotating Snakes and other Peripheral Drift illusions, it will be useful to define three related but distinct concepts. The first of these is *luminance*. This is simply the intensity of light falling on the retina from a point in the stimulus. White represents maximum luminance and black represents zero luminance. The second is *"global" contrast*. This is the difference between the luminance at a particular point in the stimulus and the average luminance across the entire stimulus. For a stimulus with an average luminance of mid-grey, both white and black represent maximum global contrast and mid-grey represents zero global contrast. The final concept is *"local" contrast*. This is the difference in luminance between a particular point in the stimulus and its near neighbours.

Retinal ganglion cells carry the output of the retina to various areas of the brain. It is widely accepted that the majority of the retinal ganglion cells feeding into the primary visual pathway have centre-surround antagonistic receptive field structures and fall into two main categories, X and Y (Linsenmeier, 1982; Wässle, 2004). X cells have smaller centres than Y cells and essentially sum their centre and surround linearly, the resulting receptive field being well modelled by a Difference of Gaussians distribution (Rodieck, 1965). X cell responses have an initial transient peak but also a lower sustained response to stimuli that excite them. Y cells have larger centres and, in addition to having conventional linear Difference of Gaussians receptive fields, sum the rectified outputs of nonlinear subunits within this conventional receptive field receptive field. Y cell responses are also faster and more transient than X cell responses. Both X and Y cells come in on-centre and off-centre varieties, where the centre is stimulated or suppressed by light respectively. The spatial arrangement of typical on-centre X and Y cells is shown in figure 8. Off-centre cells simply have the signs of their centre and surround inverted. X cells are thought to correspond to the mammalian parvocellular pathway and Y cells to the magno-cellular pathway.

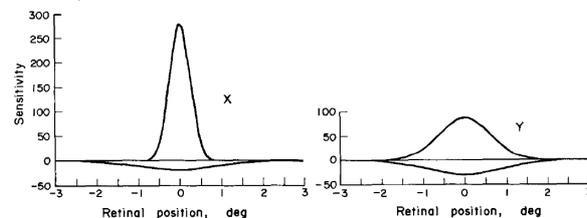


Figure 8: Average sensitivity of receptive field centre and surround in cat X and Y type retinal ganglion cells. [Source: Linsenmeier, 1982]

Given the antagonistic nature of the centre-surround receptive field structure of X and Y ganglion cells, it is clear that the signal that leaves the retina this route encodes some measure of contrast. A local

luminance is being compared to a wider regional luminance. Typically reported ratios of centre to surround size (1:4 for X cells, 1:1.5 for Y cells) suggest that ganglion cells compute a relatively local measure of contrast. If this measure is sufficiently local, then these cells will essentially act as un-oriented edge detectors. However, the size of the receptive fields vary considerably across the retina, being very small in the fovea and becoming increasing larger towards the periphery. Additionally, there appears to be substantial variation in the sizes of receptive fields at the same eccentricity. At some point between the fovea and the extreme periphery there could well be a group of X cells with centres that are perfectly sized to sample a single luminance element of the Rotating Snakes pattern. The centre/surround ratio of 1:4 would then mean that the surround was sampling at least one entire block and these cells could therefore provide an effective measure of "global" contrast for the pattern. The likelihood of such a match is increased by the radial nature of these illusions. This results in the fundamental component of each illusion being present at a range of spatial scales across a range of eccentricities.

It is difficult to find data in the literature that provides a detailed picture of how the receptive fields of ganglion cells vary across the retina. If one does not already exist, and if it is experimentally feasible, it would be a worthwhile study to produce a detailed map of ganglion cell receptive fields across the entire retina.

Luminance-dependent processing latency

The first potential mechanism for a Peripheral Drift illusion was proposed by Faubert & Herbert (1998). They suggested that higher luminance points in the stimulus were processed before lower luminance points, resulting in local motion signals across the pattern. This theory is illustrated in figure 9. Along the smoothly varying luminance gradient, small but consistent local motion signals will be generated in the light-to-dark direction. At the step edge between maximum and minimum luminance, a much stronger local motion signal is generated. While this signal is also in the local dark-to-light direction, it is in the opposite direction to the motion signals evoked by the gradient. When these local motion signals are summed over multiple wedges of the illusion, the overall global motion signal is dominated by the strong local signals at the step edges.

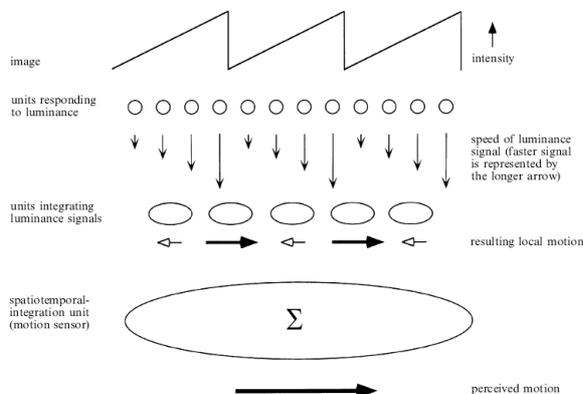


Figure 9: Mechanism proposed by Faubert and Herbert to explain their continuous gradient peripheral drift illusion [Source: Faubert & Herbert, 1998]

Faubert & Herbert cite 4 studies demonstrating that processing of lower luminance stimuli is delayed compared with processing of higher luminance stimuli (Roufs, 1963; Wilson & Anstis, 1969; Parker & Salzen, 1977; von Grunau et al 1995). However, for all but the von Grunau et al study, the lower luminance stimuli also had lower "global" and "local" contrast. These first 3 studies therefore do not provide clear evidence that processing latency increases as luminance decreases. However, the von Grunau et al study directly demonstrated that motion is perceived towards the higher luminance edge of a bar filled with a light-to-dark gradient. Motion was perceived from dark-to-light regardless of whether the stimulus was presented on a dark or light background (see figure 10). On the light background, the lighter end of the bar has lower global and local contrasts and therefore the effect cannot be due to contrast-dependent processing latency. This supports the proposal that there is a luminance-dependent latency in the visual system.



Figure 10: Stimuli used by von Grunau et al to demonstrate the perception of motion towards the higher luminance part of the stimulus. [Source: Adapted from von Grunau et al, 1995]

However, as discussed above, the conventional Difference of Gaussian receptive fields of X and Y ganglion cells cannot encode luminance directly. There is therefore the question as to how luminance information is encoded in the retina so that it is available to motion detectors in the cortex. One possibility is that the non-linear subunits of the Y receptive field preserve luminance information in some way. However, research to date has focused on the role of these subunits in a "contrast gain control" mechanism that rapidly modulates the response of the conventional Y cell receptive field to efficiently encode a wide range of contrasts with a limited output range.

An alternative possibility is that the luminance signal is carried by one of the other 10 types of ganglion cell identified in the retina. Although X cells make up the majority of ganglion cells in mammals (~50% in cat; ~75% in primates), Y cells make up only 2-3%. Yet both types of cell provide complete coverage of the retina. There remain sufficient numbers of non-X, non-Y ganglion cells to permit each of the remaining 10 types to make up over 4% of the total population and therefore it is likely that they too cover the entire retina. The function of some of these additional ganglion cell types has been identified. One cell type appears to directly sense light as well as receiving additional luminance inputs from cone photo-receptors. It is known that these cells are involved in the maintenance of the circadian rhythm, but it is not known if they also project to the primary visual pathway (Wassle, 2004). If they do, then they would be a good candidate to provide the absolute luminance information required for Faubert & Herbert's theory.

However, such a luminance-based mechanism fails to account for the perceived motion in the Rotating Snakes illusion, as can be seen in figure 11 (blue arrows).

Contrast-dependent processing latency

Conway et al (2003) propose a mechanism for the Rotating Snakes illusion in which areas of high global contrast are processed faster than areas of lower global contrast. They cite 3 papers demonstrating that high contrast stimuli are processed faster than low contrast stimuli (Shapley & Victor, 1978; Sestokas & Lehmkuhle, 1986; Mansuelli & Gibson, 1992). However, all three studies utilised sinusoidal gratings as their stimuli and thus, gratings with low global contrast also had low local contrast. Therefore the results of these studies do not provide unambiguous evidence in support of Conway et al's theory.

If it occurs, a lag between processing areas of high and low global contrast would result in local motion signals from the high contrast black and white areas towards the neighbouring lower contrast areas on either side, as shown in figure 11 (green arrows). Assuming that the latency is only sensitive to the magnitude of an area's global contrast, and not its sign, the strength of these opposing motion signals would be identical and no net motion would be observed. Conway et al propose that the "reverse phi" effect (Anstis, 1970) would reverse the perceived direction of apparent motion between areas with contrast of opposite sign. This is best understood by considering the example in figure 2 of Anstis & Rogers (1974). In this example, gradual fading out of a white disc, accompanied by gradual fading in of a black disc to its right, evokes a sensation of motion to the left. Given the proposed delay in processing for lower contrast areas, the borders between areas with oppositely signed contrast could be considered analogous to the above example, with the higher contrast area corresponding to the white disc in the example. This would result in the perceived direction of motion being towards the high contrast area.

However, the explanation of both the forward and reverse phi effect in Anstis & Rogers (1974) relies on the stimuli being processed by a retinal neuron with a Difference of Gaussians receptive field. As discussed earlier, whether such receptive fields detect local or wider regional contrast is crucially dependent on the dimensions of the receptive field centre and surround. However there may well be a point between the fovea and the extreme periphery of the visual field where there are neurons with receptive fields are "just right" to compare the luminance of a single element of the illusion's "building block" against the average luminance across one or more entire blocks. This would permit the computation of the "global" contrast required by Conway et al.

Although Conway et al's "contrast-driven latency + reverse phi" model appears to explain the Rotating Snakes illusion well, it does not predict any perceived motion for Faubert & Herbert's gradient version of the Peripheral Drift illusion. At all points in the pattern where the global contrast changes sign, the magnitude of the contrast of the areas to either side are equal. As a result there is no latency difference and therefore no perceived motion.

Phase advance model

Backus & Oruc explicitly modelled the time course of the response of V1 neurons to the high and low contrast areas of the Rotating Snakes pattern, using data from Albrecht et al (2002). Albrecht et al demonstrated that the peak response of these neurons was delayed for lower contrast stimuli. Again

the stimuli were sinusoidal, so global and local contrast co-varied. Therefore this is not unambiguous evidence for latency varying with global contrast. Backus and Oruc calculated the neural representation of the image at each point in time as a function of the global contrast at each point in the image. They found that the phase of the sinusoid representing the fundamental frequency of the variation in the neural image drifted over time in a direction consistent with the illusion. Assuming that the fundamental sinusoid has the greatest influence on the determination of position by V1 cells, this "phase advance" would result in the perception of motion in a direction consistent with the illusion.

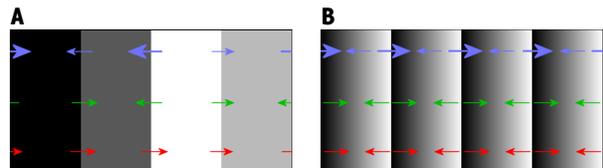


Figure 11: Motion predicted by various models for the Rotating Snakes illusion (A) and Faubert & Herbert's gradient version of the Peripheral Drift illusion (B). Thicker arrows indicate a stronger motion effect. **Blue:** luminance model. **Green:** "global" contrast model with no reverse phi. **Red:** "global" contrast model with reverse phi. As can be seen, none of these models successfully predict the perceived motion for both illusions (which is rightwards in both figures). Predicted motion arrows for the Backus & Oruc "phase-advance" model are not shown in this figure, but the model successfully predicts the perceived motion for the Rotating Snakes illusion. However, it fails to account for the perceived motion in Faubert & Herbert's illusion.

In common with the reverse phi mechanism proposed by Conway et al, this phase advance model fails to explain Faubert & Herbert's gradient version of the Peripheral Drift illusion. Backus & Oruc suggest that a separate luminance based model is required to explain such "single gradient" illusions. Although couched in different terms and claimed to be not reliant on reverse phi motion, Backus & Oruc's phase advance model seems very similar to the model of forward and reverse phi motion put described by Anstis & Rogers (1974), and it is likely that the underlying mechanism is the same.

Gaussian Derivative model

Johnston et al (1998; 2003) proposed a model of visual processing in V1 based on the use of Gaussian derivatives to blur the luminance image and compute spatial and temporal derivatives of various orders. The model has proved effective at accurately detecting real motion. It has also detected some second order motion effects caused by, for example, contrast modulation of moving images. So far the model has not been used to attempt to detect motion

in static images. However, the contrast-dependent latency described above could be considered a gradual switch from an input dominated by high-contrast elements to one dominated by low-contrast elements. A theoretical qualitative analysis of the expected output of the model suggested that it would detect appropriate illusory motion if presented with the Rotating Snakes pattern at gradually increasing contrast (see figure 12). Therefore a simplified Rotating Snakes pattern was contrast-modulated and passed through the Gaussian derivative model to see if any motion was detected.

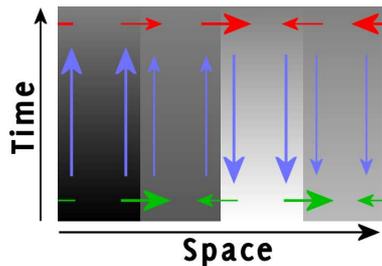


Figure 12: Theoretical qualitative prediction of motion that should be perceived by the Johnston et al Gaussian derivatives model for a Rotating Snakes pattern that decreases in contrast over time. Large arrows of the same colour indicate qualitatively larger derivatives. **Red:** Spatial derivatives along axis of variation. **Blue:** Temporal derivatives. **Green:** Crossed second order space and time derivatives. First order derivatives are positive if luminance is increasing for positive movement along the relevant axis. Purely spatial or temporal derivatives are not considered as they will not give rise to motion cues.

Two initial set of frames were produced and were processed using a version of the model optimised for 2D motion. The first set of frames comprised the stationary "building block" of the Rotating Snakes illusion (see figure 2), increasing and decreasing in contrast on a mid-grey background. The second set of frames comprised the same "building block" drifting from left to right and back again across a mid-grey background at 1 pixel per frame. This second set of frames served as a control.

The results from these initial framesets were very confusing. Intriguingly, motion was detected in the contrast frameset. The direction of any motion perceived at each point of the image was always in the same direction, regardless of whether contrast was increasing or decreasing. Additionally, the magnitude of the detected motion varied significantly across the frameset. More worryingly, the motion detected in the control subset was equally bizarre, with downward or rightward motion reported, regardless of the actual direction of motion of the stimulus. A number of potential causes for these bizarre results were identified, mostly related to stimulus preparation and the reversal of direction cues over a short number of frames.

A new set of contrast and control frames were produced that addressed most of the issues identified. The new contrast sequence contained more frames and only varied the contrast in one direction. Similarly, the new control sequence only moved in one direction. As both the real and illusory motion expected from the stimulus were essentially 1D, it was also decided to run the new framesets through a version of the model optimised for 1D motion.

The output for the control frameset was largely as expected. All detected motion was in the correct direction, and most points were reported as moving at 1-2 pixels/frame, close to the 1px/frame true motion. However, there were some anomalous parts of the frame where speeds of up to 80 pixels/frame were reported (see figure 13c). For the contrast frameset, the reported motion was very strange. Although some frames reported motion in up to 3 elements, most frames either reported no motion or a block of motion in the centre of a single element, as shown in figure 13b. The direction of this motion appeared random, and the reported magnitudes were generally extreme, averaging 776,000 pixels/frame

and reaching 3.6 million pixels/frame. Such staggering velocities are very likely to be artefacts of the model, especially given their random directions. However, it was not possible to evaluate the algorithm underlying the model implementation provided, and there was not sufficient time to replicate the model to investigate the cause of these excessively high reported velocities.

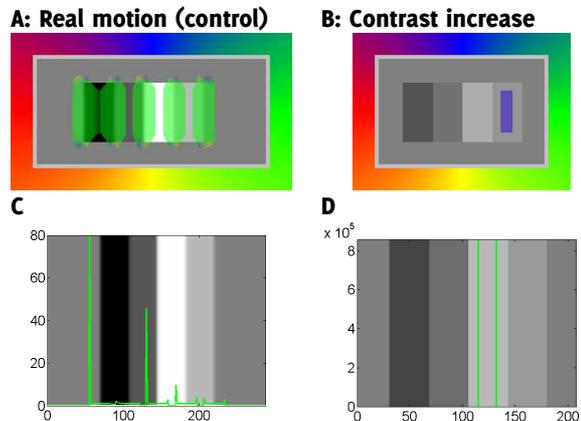


Figure 13: Motion detected by the 1D-optimised Gaussian derivative model for **A:** 1 pixel-per-frame true motion to the left. The model accurately detects motion in the correct direction centred on the element borders. The magnitude of the motion is reported accurately for most image pixels, but there are a few anomalies reporting motion of up to 8 pixels-per-frame. **B:** An increase in contrast of 0.01 per frame. Motion is sporadically detected in 0-3 elements per frame. The direction appears to be random and the reported magnitude is extreme (850,000 pixels-per-frame in this example). **C** and **D:** plots showing the magnitude of motion along a central cross-section of the frames shown in A and B. The background to each plot is the time average of the 23 frames around the frame of interest.

Conclusion

To date, no proposed mechanism can explain both "single gradient" Peripheral Drift illusions such as that reported by Faubert & Herbert and the "dual gradient" Rotating Snakes illusion. However, reasonable mechanisms have been proposed that separately explain each type of illusion. One conclusion is that these illusions do not share an underlying mechanism, but are caused by separate luminance-dependent and contrast-dependent processing latencies respectively. However, all the cells known to transfer retinal output to the primary visual pathway are believed to have antagonistic centre-surround receptive fields. With such fields it is not possible to encode raw luminance. However, suitably sized fields could potentially encode either global or local contrast. Any luminance-dependent latency would have to be encoded by an as yet unknown channel, or perhaps occur prior to the centre-surround layer of the retina. It is not immediately clear how such early luminance-dependent latency would affect the output of the centre-surround ganglion cells. For example, it is possible that the centre-surround processing may mask any information regarding luminance-dependent latency from higher level motion detectors.

It seems clear that different researchers consider the idealised retina to be encoding different properties of the image falling upon it. In order to fully evaluate the proposed mechanisms discussed above, more information is required about the true nature of the "neural image" encoded by the retina.

For example, the retina may encode a very local measure of contrast, essentially performing edge detection. None of the mechanisms proposed to date would appear to work with such a neural image. Should this be the case, one interesting avenue of exploration would be to explore any differences between the on and off channels in the retina. As mentioned previously, retinal ganglion cells are either on-centre or off-centre. These two types of cell are the final stages of two largely separate pathways in the retina, and this separation of the on and off pathways extends to the primary visual cortex (Westheimer, 2007). Are there any timing differences between these pathways that might contribute to these illusions? One intriguing recent finding is that many off-centre ganglion cells in the Salamander briefly switch to act as on-centre cells for 100ms or so following a shift of the peripheral image (Geffen et al, 2007). What effect might this switch back and forth have on the neural representation of these illusions?

To gain greater clarity on the nature of the neural image encoded by the retina, it would be useful to model a range of different centre-surround cells of different types and at different eccentricities, using the receptive field data available in the literature. Combined with measures of the size of illusion elements for various viewing geometries, this should permit reasonable estimates of the retinal image to be produced. These estimates of possible retinal images should help assess the relative likelihood of information on the luminance, global contrast or local contrast of the illusion being available to the post-retinal motion centres of the brain.

At a more abstract level, the preliminary theoretical analysis of the Gaussian derivatives model suggested that it would detect illusory motion in the direction of the illusion for a Rotating Snakes pattern of decreasing contrast. However, the model reported velocities with extreme magnitudes and random directions when presented with this stimulus. It is likely that this output reflects model artefacts. However, it also makes it difficult to reject the theoretical expectation on the basis of these results. After revisiting the theoretical analysis to confirm if it is valid, and establishing the cause of the extreme velocities reported, it would be useful to re-run the stimulus with an artefact-free model. Of course, it is possible that the reported velocities may not be artefacts of the model implementation, but real properties of the Gaussian derivatives model. In this case it would be intriguing to understand how they arise in the model and if similar effects could be expected in biological motion processing systems.

One thing all the proposed mechanisms have in common is that they rely on transient properties of the neural response. This fits with the need to refresh the retinal image to sustain the illusion. However, it would be useful to know to what extent retinal refreshment is necessary to sustain the illusion in the short term. Backus and Oruc note that, when fixated, the strength of the illusion gradually fades over 6-8 seconds of steady fixation. However, this is not supported by a formal study. Informal tests by this author have identified at least one person who claims the illusion ceases immediately upon fixation. If true, this suggests that micro-saccades might play a crucial role in the perception of the illusion. It might be possible to completely remove the effect of eye movement on the retinal image by tracking the observer's eye and redrawing the displayed image so

that the retinal image remains perfectly stationary. Eye tracking is fast and precise enough to move a small sharp section of otherwise fuzzy image to follow normal saccades, giving the observer the illusion that the entire displayed image is perfectly sharp. However, it is not known how much more difficult the timely tracking of micro-saccades would be.

Finally, most of the studies formally exploring these illusions had a limited number of subjects. The conclusions regarding the effects of stimulus duration, luminance, contrast, size, eccentricity etc would be greatly strengthened if further studies were carried out with more significant numbers of participants. Ideally these studies would use the forced choice speed matching paradigm employed by Kuriki et al to produce quantitative measures of how the strength of the illusion varies with each parameter.

Acknowledgements

Many thanks to Alan Johnston and Xuefeng Liang for very useful discussions and the provision of a Matlab implementation of the Gaussian derivatives motion model.

All modelling and analysis was performed using Matlab (The Mathworks, 2007). UCL CoMPLEX is an EPSRC-funded Life Sciences Interface Doctoral Training Centre.

References

- 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. *J Neurophysiol*, *88*, 888-913.
- Anstis, S. (1970). Phi movement as a subtraction process. *Vision Research*, *10*, 1411-1430.
- Anstis, S. M. & Rogers, B. J. (1975). Illusory reversal of visual depth and movement during changes of contrast. *Vision Research*, *15*, 957-961.
- Backus, B. T. & Oruc (2005). Illusory motion from change over time in the response to contrast and luminance. *Journal of Vision*, *5*, 1055-1069.
- Bulthoff, H. & Gotz, K. G. (1979). Analogous motion illusion in man and fly. *Nature*, *278*, 636-638.
- Conway, B. R.; Kitaoka, A.; Yazdanbakhsh, A.; Pack, C. C. & Livingstone, M. S. (2005). Neural Basis for a Powerful Static Motion Illusion. *J. Neurosci.*, *25*, 5651-5656.
- Faubert, J. & Herbert, A. M. (1999). The peripheral drift illusion: A motion illusion in the visual periphery. *Perception*, *28*, 617-622.
- Fraser, A. & Wilcox, K. J. (1979). Perception of illusory movement. *Nature*, *281*, 565-566.
- Geffen, M. N.; de Vries, S. E. J. & Meister, M. (2007). Retinal Ganglion Cells Can Rapidly Change Polarity from Off to On. *PLoS Biology*, *5*, e65.
- von Grunau, M.; Saikali, Z. & Faubert, J. (1995). Processing speed in the motion-induction effect. *Perception*, *24*, 477-490.
- Hisakata, R. & Murakami, I. (2008). The effects of eccentricity and retinal illuminance on the illusory motion seen in a stationary luminance gradient. *Vision Research*, *48*, 1940 - 1948.
- Johnston, A., McOwan, P.W., Benton, C.P. (1999). Robust velocity computation from a biologically motivated model of motion perception. *Proc of Royal Society London B*, *266*, 509-518.
- Johnston, A., McOwan, P.W., Benton, C.P. (2003). Biological computation of image motion from flows over boundaries. *Journal of Physiology*, *97*, 325-334.

CoMPLEX Case essay

Kitaoka, A. & Ashida, H. (2003). Phenomenal Characteristics of the Peripheral Drift Illusion. *Vision (the journal of the Vision Society of Japan)*, 15, 261-262.

Kuriki, I.; Ashida, H.; Murakami, I. & Kitaoka, A. (2008). Functional brain imaging of the Rotating Snakes illusion by fMRI. *J. Vis.*, 8, 1-10.

Linsenmeier, R.; Frishman, L.; Jakiela, H. & Enroth-Cugell, C. (1982). Receptive field properties of X and Y cells in the cat retina derived from contrast sensitivity measurements. *Vision Research*, 22, 1173-1183.

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

Parker, D. & Salzen, E. (1977). Latency changes in the human visual evoked response to sinusoidal gratings. *Vision Research*, 17, 1201-1204.

Rodieck, R. (1965). Quantitative analysis of cat retinal ganglion cell response to visual stimuli. *Vision Research*, 5, 583-601.

Roufs, J. (1963). Perception lag as a function of stimulus luminance. *Vision Research*, 3, 81-91.

Shapley, R. M. & Victor, J. D. (1978). Effect Of Contrast On The Transfer Properties Of Cat Retinal Ganglion-Cells. *Journal Of Physiology-London*, 285, 275-298.

Wassle, H. (2004). Parallel processing in the mammalian retina. *Nat Rev Neurosci*, 5, 747-757.

Westheimer, G. (2007). The ON-OFF dichotomy in visual processing: From receptors to perception. *Progress in Retinal and Eye Research*, 26, 636-648.

Wilson, J. A. & Anstis, S. M. (Sep., 1969). Visual Delay as a Function of Luminance. *The American Journal of Psychology*, 82, 350-358.