Available online at www.sciencedirect.com
SCIENCE DIRECT®
Vision Research xxx (2006) xxx–xxx
ELSEVIER
Vision Research
www.elsevier.com/locate/visres

The infinite regress illusion reveals faulty integration of local and global motion signals

P.U. Tse *, P.-J. Hsieh

Department of Psychological and Brain Sciences, Moore Hall, Dartmouth College, Hanover, NH 03755, USA

Received 28 April 2006; received in revised form 12 June 2006

Abstract

We report a new visual illusion, where a global shape appears to continually move away from fixation, even though it remains a fixed distance from fixation. The illusion occurs because local motion signals within the object indicate motion away from fixation, and are incorrectly attributed by the visual system to the motion trajectory of the global object. A simple weighted vector summation of global and local motion signals, while a reasonable first approximation, cannot fully account for our data. We show that the faster the local motion signal, the more it biases judgments of global motion direction. We propose that local and global motion signals are summed non-linearly for this stimulus because as local motion speed increases, moving luminance blobs are visible for less time, affording less time to inhibit inappropriate component motion signals. This effect reveals the degree to which the visual system can incorrectly combine local and global motion signals belonging to a single object.

Keywords: Visual motion perception; Motion vector summation; Infinite regress illusion

1. Introduction

The infinite regress illusion (IRI) belongs to a class of illusions where motion signals lead to errors in spatial localization or determination of motion amplitude or direction. There have been numerous examples where position is mislocalized because of motion signals. The initial (Frohlich, 1929) and final (Freyd & Finke, 1987) position of an object can be mislocalized in the direction of either surrounding motion or even non-adjacent motion elsewhere in the visual field (Whitney & Cavanagh, 2000). There have also been examples such as ‘motion capture’ (Ramachandran, 1987) and ‘induced motion’ (Duncker, 1929) where motion at one location can influence the motion perceived elsewhere in the image. The present example is different from either motion capture or motion induction because in these phenomena the motion of one object is influenced by the motion of some other object. In the IRI, in contrast, the motion that is misperceived is that of a single object that has both a local motion component in one direction and a global motion component in another direction. Faulty combination of these local and global motion signals leads to the remarkable illusion that an object can appear to move continually away from the point of fixation without in fact moving away from it at all.

The present motion effect builds upon a positional illusion first noted by Devalois and Devalois (1991; [compare also Ramachandran and Anstis, 1990]), where stationary patches containing drifting Gabor gratings (cosine gratings tapered in *X* and *Y* by Gaussians) appear to be spatially shifted in the direction of drift motion. When two such patches are vertically aligned, each containing, respectively, Gabor gratings drifting in opposite directions, the two stationary patches seem to be vertically misaligned. Here we continuously move a group of such Gabor patches, all drifting at the same speed in the same direction. Words cannot do the illusion justice, and the reader is encouraged to view the effect themselves at

---

* Corresponding author.
E-mail addresses: peter.u.tse@dartmouth.edu (P.U. Tse), Po-Jang.Hsieh@dartmouth.edu (P.-J. Hsieh).

0042-6989/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2006.06.010

either http://ees.elsevier.com/vr/ or http://illusioncontest.neuralcorrelate.com/index.php?module = pagemaster&PAGE_user_op = view_page&PAGE_id = 69. Individual Gabors drifted in the horizontal direction away from fixation, while the group of these Gabors moved together in the vertical direction. We independently varied the speed of the global vertical motion of the group of patches, as well as the speed of the local horizontal motion of the drifting Gabor gratings.

We find, to a rough first approximation, that the motion that is perceived is a weighted sum of these local and global components, where the global motion is weighted more heavily than the local motion signals. However, weighted vector summation cannot fully account for our data because the bias in the global motion percept that is introduced by local motion signals increases as local motion speed increases. In other words, the faster the speed of the local motion signal, the greater the illusion that global motion is biased in the direction of the local motion signal, at least within the range of speeds tested. Whereas De Valois and De Valois (1991) concluded that motion signals can lead to misjudgments of spatial location, we conclude that the problem is more general. The visual system combines local and global motion signals incorrectly, by misattributing some of the local motion signal to the global motion. Misjudgments of position would arise integrally from such a mechanism if perceived position were encoded with a weighting by motion input, as appears to be the case (e.g. Whitney & Cavanagh, 2000).

2. Materials and methods

2.1. Observers

Six subjects (five naïve and one author, age range: 20–28) carried out the experiments. All of them had normal or corrected-to-normal vision. All of our observers were experienced psychophysical observers. All of them were capable of alternating attention between the two regions (physical stimulus and adjustment area). Before each experiment, the subjects practiced several training trials until they were accustomed to the experimental procedure and were capable of fixating while conducting hand movements.

2.2. Stimuli and procedures

The stimulus configuration and experimental procedure used in the first experiment are shown in Fig. 1. The fixation spot was a blue (luminance: 285 lumen/m^2; CIE, $x = 0.402$, $y = 0.517$; measured using a Minolta 100LS colorimeter) square that subtended 0.05 of visual angle and centered 18 visual degrees to the left of the screen center. The target stimulus was composed of 15 'elongated' Gabor gratings (1 cycles/degree), each subtending 1.75° of visual angle in width and 3° of visual angle in height. All Gabor gratings moved up and down coherently as a group as they simultaneously cycled at a constant rate to the right within an 11° × 24° window on a gray (39.5 lumen/m^2) background. The maximum luminance value within the Gabor patch was white (239 lumen/m^2) and the minimum was black (0.5 lumen/m^2) background. A black bar (the to-be-adjusted stimulus'), subtending 0.3° of visual angle in width and 7° of visual angle height, was continuously present on the screen and centered 15 visual degrees to the left and 9 visual degrees below the screen center. All the stimuli were viewed with both eyes. The total size of the visual field was 40 cm × 30 cm, viewed from a distance of 57 cm. The monitor thus subtended 40° vertical visual angle and 30° horizontal visual angle. Subjects had their chin in a chin rest. The visual stimulator was a 2 GHz Dell workstation running Windows 2000. The stimuli were presented on a 23-in. SONY CRT gamma-corrected monitor with 1600 × 1200 pixels resolution and 85 Hz frame rate.

Fig. 1. The stimuli consisted of a group of 15 Gabor patches drifting to the right. The entire group moved up and down continuously at a constant speed. Subjects fixated on the fixation spot to the left, and manipulated the orientation of the black bar to match that of the perceived group motion. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this paper.)

In each trial, the stimulus was identical to the default values described above except that the vertical (up/down) speed of the global motion was randomly assigned to be one of the following values (8, 10, or 12 visual degrees/s) and the horizontal (rightward) speed of the local drifting sine-wave motion was randomly assigned to be one of the following values (3, 4.5, or 6 visual degrees/s). Subjects were required to adjust the orientation of the bar to match the perceived direction of the stimulus as the stimulus was moving downward. Eye movements were monitored using a head-mounted eyetracker (Eyelink2, SR research, Ontario, Canada). Trials during which the subject's monitored left eye was outside a fixation window of 1.5 visual degrees radius were excluded and repeated later in the experiment. Thus all data reported here were carried out under conditions of fixation. All conditions were randomized and counterbalanced across 27 trials.

3. Results

Our results, shown in Fig. 2 and summarized in Table 1, reveal that the perceived direction of motion of the group of Gabors, relative to the true vertical direction of group motion, increases as the horizontal (Gabor drift) speed increases at a given vertical (group) speed ($F = 37.66$, $P < 0.001$). In contrast, the perceived angle decreases as the vertical speed increases at a given horizontal speed ($F = 8.50$, $P < 0.007$). There was no interaction between local (horizontal) and global (vertical) motion directions.

4. Discussion

It is commonly assumed that global motion signals should dominate local motion signals, since otherwise, local motion signals, say from a tiger's legs moving backwards relative to its torso during a leap forward, could lead to the

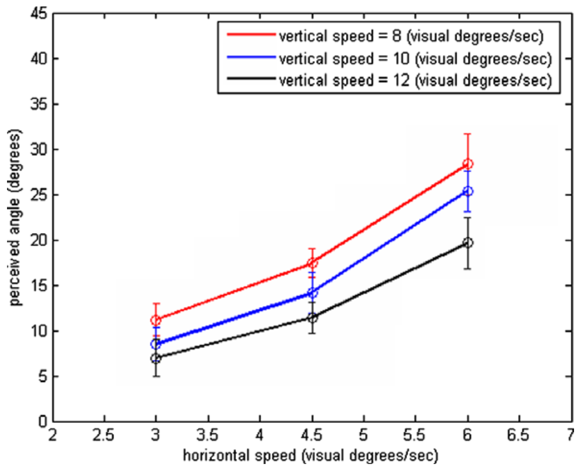

Fig. 2. Perceived angle away from vertical as a function of the speed of the horizontal Gabor drift speed ($n = 6$) and the vertical group motion speed.

Table 1
Repeated measures ANOVA revealing main effects of horizontal (local) and vertical (global) motions, with no interactions between these factors

Source	SS	*df*	Mean square	*F*	Sig.
Horizontal speed	2247.2	2	1123.6	37.66	0.000
Vertical speed	359.7	2	179.877	8.5018	0.007
Vertical speed × Horizontal speed	39.025	4	9.756	1.2993	0.304
Horizontal speed × Subject	298.3516	10	29.8352		
Vertical speed × Subject	211.5745	10	21.1574		
H × *V* × Subject	150.1754	20	7.5088		

Dependent variable: Perceived angle.

misperception of the global direction of motion (the direction that the tiger is in fact moving). In the extreme case, the forward-moving tiger would appear to move backwards because its legs were momentarily moving backwards. Clearly, making such an error could lead to possibly fatal judgments about the direction the tiger was moving. Yet, this is exactly the type of error that we find the visual system to be making. In short, we report here a striking new illusion that makes apparent the degree to which local and global motion signals are miscombined within a single object to create a coherent but incorrect percept of the direction of object motion.

Past work, involving motion capture (Ramachandran, 1987; where the random jumps of small dots are captured by the motion of a low spatial frequency envelope or background object, creating the illusion that the dots are moving in the same direction as the envelope), suggests that motion can be mistakenly attributed from one object to another occupying the same spatial location. Motion induction (Duncker, 1929), where motion at one location can influence the motion perceived elsewhere in the image, implies that motion is not computed only in terms of local motion measurements. Motion induction, like motion capture, is an example where the motion of one object alters the motion perceived over another object. Here, in contrast, we report how local and global motion components of a single object are miscombined.

To a first approximation, our results can be characterized by a simple weighted vector sum model, depicted in Fig. 3, according to which the perceived direction of global motion is given by the vector sum of the true global motion vector ‘*V*’ plus a constant ‘*b*’ times the true local motion vector ‘*H*’. The value of the constant ‘*b*’ that best characterizes the present data set is approximately $b = 0.53$. This means that the global motion vector is weighted roughly twice as much as the local motion vector in this presumed process of motion vector summation. The data from this model are shown in Fig. 4.

This model is, however, inadequate, because the slopes of the predicted data are not high enough. The actual weighting factors for factor ‘*b*’ described above, are shown in Fig. 5. The fact that factor ‘*b*’ increases with increasing local (horizontal) motion speed ($F = 5.2793$, $p = 0.0272$), but not with increasing global (vertical) motion speed (see Table 2), means that the local motion signal distorts the

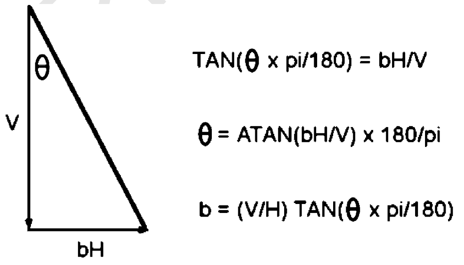

Fig. 3. According to the simplest possible model, the perceived angle of illusory object motion is given by theta, and theta results from some proportion of the true horizontal motion signal, contributing to the global motion percept. The factor ‘*b*’ by which the horizontal (local) motion vector is multiplied before vector summation is given by the formula shown.

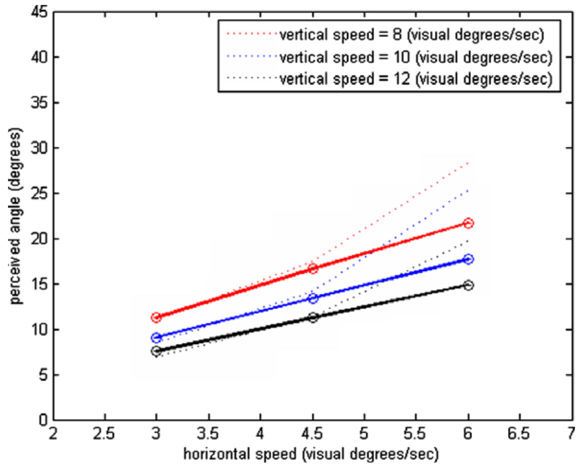

Fig. 4. The data that should arise according to a simple vector summation model where the weighting on the horizontal or local component vector is $b = 0.53$ are shown as solid lines. The actual data from Fig. 2 are shown as dotted lines for comparison. This value of b was chosen to make the leftmost red point have the same perceived angle as the one that was in fact measured, shown in Fig. 2. The model is explained in the text and in Fig. 3. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this paper.)

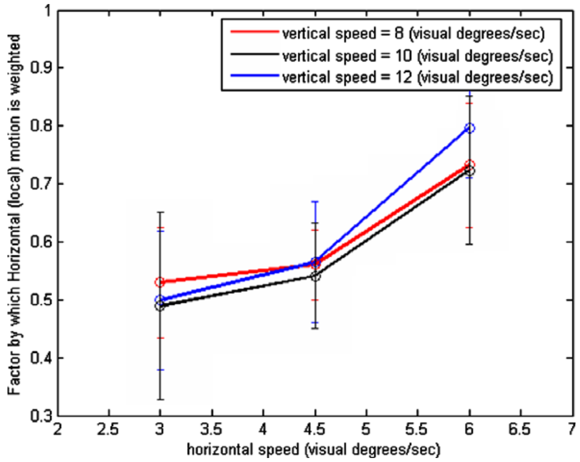

Fig. 5. The factor by which the horizontal motion component must be multiplied to account for the motion percept according to the simple weighted vector summation model shown in Fig. 3 and explained in the text. As the horizontal (local) motion component increases in speed, the degree to which it contributes to the global motion percept direction increases.

Table 2
Repeated measures ANOVA revealing a main effect of speed of horizontal (local) motion on the *b* factor, with no main effect for vertical (i.e., global) motion, and no interactions between these factors

Source	SS	*df*	Mean square	*F*	Sig.
Horizontal speed	0.6057	2	0.3028	5.2793	0.0272
Vertical speed	0.0119	2	0.0059	0.1386	0.8722
Vertical speed × Horizontal speed	0.0147	4	0.0037	0.3525	0.8392
Horizontal speed × Subject	0.5736	10	0.0574		
Vertical speed × Subject	0.4289	10	0.0429		
H × *V* × Subject	0.2091	20	0.0105		

Dependent variable: Magnitude of '*b*' factor.

perception of global motion more as local motion speed increases. The non-linear, increasing contribution of local motion signal to the global motion percept is solely a function of local motion speed, and is independent of global motion speed.

This model is different from other models that have attempted to explain the misperception of position because of the influence of nearby or distant motion signals on the basis of differential latencies (Purushothaman, Patel, Bedell, & Ogmen, 1998; Whitney & Murakami, 1998; Whitney, Murakami, & Cavanagh, 2000), extrapolation (Khurana & Nijhawan, 1995;Nijhawan, 1994), attentional shifts (Baldo & Klein, 1995), anticipatory retinal responses (Berry, Brivanlou, Jordan, & Meister, 1999), or integration of motion signals within a brief temporal window (Eagleman & Sejnowski, 2000). Indeed, it is not a model of positional mislocalization at all. It is, rather, a model of how local and global motion signals are combined (or miscombined) to create a percept of a global motion direction and magnitude. It could account for positional mislocalization due to motion signals if the computation of position takes global and local motion signals as an input, as Whitney and Cavanagh (2000) have argued.

Why might local and global motion signals be combined in this inappropriate way? It is possible that global and local motion signals are integrated because of what is called the 'aperture problem'. Not all cells in motion processing areas will have the whole display available to them in their receptive fields. Rather, they will only respond to a portion of the display that activates them at any moment. Imagine a cell that only responded to a single Gabor. The motion it would code would be that of a luminance blob moving upward or downward and to the right, since this is in fact the trajectory of any given luminance blob viewed from within a small aperture. If the global motion that is perceived arises from the population response of many cells, some of which suffer from this aperture problem because of relatively small receptive field size, then the global motion percept will be biased incorrectly toward a rightward motion. Another possibility is that such luminance blobs are treated as features that are tracked. Such features would also have a locally rightward motion component.

Why might the apparent summation of local and global motion signals be non-linear? As the rightward, local motion of the drifting Gabors increases in speed, luminance blobs not only move faster, they are available for less time before disappearing from the Gaussian envelope that defines the Gabor patch. They would also be available for less time within an aperture, such as a receptive field, of a fixed size. We hypothesize that this affords less time for the inhibition of component motion signals using non-component motion signals. Component motion or motion energy-driven solutions appear to dominate within the first one to two hundred milliseconds following motion onset. Lorenceau and colleagues found that the motion perceived in a field of moving bars, for example, is initially perpendicular to the orientation of the bars, rotating to the actual direction of motion within about 200 ms (Lorenceau, Shiffrar, Wells, & Castet, 1993). After this initial integration period, however, the motions of the intrinsic terminators belonging to the moving bars largely determine the perceived direction of bar motion (Lorenceau & Shiffrar, 1992; Wallach, 1935). It has also been shown that neurons in MT in the macaque will initially respond to the direction of motion that is perpendicular (component direction) to a moving line, independent of the actual direction of motion (Pack & Born, 2001). These same neurons will, over a period of ~60 ms, shift their response properties so that they respond to the true motion of the line independent of its orientation, suggesting that the unambiguously moving endpoints of the line are quickly, but not instantaneously, exploited to generate a veridical motion solution (see also Pack, Gartland, & Born, 2004). The response properties of these neurons match behavioral data that show that initial pursuit eye-movements will be in the direction perpendicular to the moving line, and then rapidly adapt to follow the direction of veridical motion as defined by line terminators

(Pack & Born, 2001). It is likely that motion processing areas generate at least two motion solutions for a moving stimulus; a local one consistent with the component direction of motion that one would obtain by viewing the stimulus through the “aperture” of a receptive field in primary visual cortex, and a global one consistent with the motion of global form cues, such as terminators. Usually the intrinsic terminator solution wins within about 200 ms, as shown by Lorenceau et al. (1993), and the aperture or component solution loses, generating an unambiguous percept of motion in the direction specified by the intrinsic terminators. However, when the luminance blobs are only visible for a very short time, either because they are present in the stimulus briefly, or because they pass quickly through a receptive field, the component contribution to perceived motion will remain high, because inhibition of component motion signals takes on the order of one to two hundred milliseconds. If a luminance blob disappears before the intrinsic terminator and/or global form-based motion signals have had a chance to dominate component signals, the component solution will not be fully suppressed. In the absence of sufficient suppression of component motion signals by non-component motion signals, the visual system may integrate component and non-component motion signals to generate a perception of motion. The contribution of component signals to perceived motion should be greater the shorter the duration of object motion, because of the finite time required to inhibit component motion signals. Because faster local motion signals are present in the stimulus for a shorter duration, the component motion contribution to perceived motion direction will be greater, accounting for the non-linear nature of summation of local and global motion signals.

We conclude that the infinite regress illusion occurs because there is a contribution of local motion signals from within an object to the global motion direction computed for that object as a whole. We have shown here that this local/global motion signal summation is non-linear, and that this non-linearity can be accounted for by a relatively simple model that takes into account the finite duration required for inhibition of local component motion signals. The influence of local motion signal on the global motion that is perceived is a potentially serious error made by the human visual system, in that it can lead perceivers to believe that an object, such as a tiger, is moving in a direction that it is not. If the current model is correct, such errors typically do not arise when observing tigers or other objects, because real-world objects are typically visible long enough for the inhibition of inappropriate local component motion signals.

Acknowledgment

This project was funded by NIH R03 MH0609660-01 grant to P.U.T.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.visres.2006.06.010.

References

Baldo, M. V., & Klein, S. A. (1995). Extrapolation or attention shift? *Nature, 378*, 565–566.

Berry, M. J., II, Brivanlou, I. H., Jordan, T. A., & Meister, M. (1999). Anticipation of moving stimuli by the retina. *Nature, 398*, 334–338.

De Valois, R. L., & De Valois, K. K. (1991). Vernier acuity with stationary moving Gabors. *Vision Research, 31*(9), 1619–1626.

Duncker, K. In *Source book of gestalt psychology* (ed. and trans. Ellis, W.D.) 161–172 (Routledge & Kegan Paul, London, 1938). [Reprinted from Psychologische Forschung, 12, 180–259, 1929].

Eagleman, D. M., & Sejnowski, T. J. (2000). Motion integration and postdiction in visual awareness. *Science, 287*, 2036–2038.

Freyd, J. J., & Finke, R. A. (1987). Representational momentum. *Journal of Experimental Psychology, 19*, 369–401.

Frohlich, F. W. (1929). *Die empfindungszeit*. Verlag von Gustav Fischer, Jena.

Khurana, B., & Nijhawan, R. (1995). Extrapolation or attention shift? Reply. *Nature, 378*, 566.

Lorenceau, J., & Shiffrar, M. (1992). The influence of terminators on motion integration across space. *Vision Research, 32*, 263–273.

Lorenceau, J., Shiffrar, M., Wells, N., & Castet, E. (1993). Different motion sensitive units are involved in recovering the direction of moving lines. *Vision Research, 33*, 1207–1217.

Nijhawan, R. (1994). Motion extrapolation in catching. *Nature, 370*, 256–257.

Pack, C. C. & Born, R. T. (2001). Temporal dynamics of a neural solution to the aperture problem in visual area MT of macaque brain. *Nature*, 22;409(6823):1040–1042.

Pack, C. C., Gartland, A. J., & Born, R. T. (2004). Integration of contour and terminator signals in visual area MT of alert macaque. *The Journal of Neuroscience, 24*(13), 3268–3280.

Purushothaman, G., Patel, S. S., Bedell, H. E., & Ogmen, H. (1998). Moving ahead through differential visual latency. *Nature, 396*, 424.

Ramachandran, V. S. (1987). Interaction between colour and motion in human vision. *Nature, 328*, 645–647.

Ramachandran, V. S., & Anstis, S. M. (1990). Illusory displacement of equiluminous kinetic edges. *Perception, 19*, 611–616.

Wallach, H. (1935). Ueber visuell whargenommene bewegungrichtung. *Psychologische Forschung, 20*, 325–380.

Whitney, D., & Cavanagh, P. (2000). Motion distorts visual space: Shifting the perceived position of remote stationary objects. *Nature Neuroscience, 3*(9), 954–959.

Whitney, D., & Murakami, I. (1998). Difference, not spatial extrapolation. *Nature Neuroscience, 1*, 656–657.

Whitney, D., Murakami, I., & Cavanagh, P. (2000). Illusory spatial offset of a flash relative to a moving stimulus is caused by differential latencies for moving and flashed stimuli. *Vision Research, 40*, 137–149.