



**HAL**  
open science

# The false aperture problem: Global motion perception without integration of local motion signals

Rémy Allard, Angelo Arleo

► **To cite this version:**

Rémy Allard, Angelo Arleo. The false aperture problem: Global motion perception without integration of local motion signals. *Psychological Review*, In press, 10.1037/rev0000347 . hal-03426654

**HAL Id: hal-03426654**

**<https://hal.science/hal-03426654>**

Submitted on 17 Nov 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1                                   **The false aperture problem:**  
2                                   **Global motion perception without**  
3                                   **integration of local motion signals**

4                                   Rémy Allard<sup>1,2\*</sup> & Angelo Arleo<sup>1</sup>

5                                   <sup>1</sup>Sorbonne Universités, UPMC Univ Paris 06, INSERM, CNRS, Institut de la Vision, Paris, 75012 France

6                                   <sup>2</sup>School of optometry, Université de Montréal, Montréal, Canada

7                                   \*Corresponding Author: [remy.allard@umontreal.ca](mailto:remy.allard@umontreal.ca)

8                                   Tel: 514-343-6111 #8807

9                                   <http://www.aging-vision-action.fr>

10 **Author Note**

11 This research was supported by ANR – Essilor SilverSight Chair ANR-14-CHIN-0001.

12 R.A. designed and conducted the experiments, analyzed and interpreted data, and wrote the  
13 article. A.A. provided overall project coordination and proofread the article.

14 The results of the first experiment were presented at the Vision Science Society conference  
15 (Allard & Arleo, 2016b), and the results of the second experiment were presented at the European  
16 Conference on Visual Perception (Allard & Arleo, 2016a). This study was not preregistered and  
17 the data are available upon request to the corresponding author.

18 The authors declare no competing financial interests.

19 © 2021, American Psychological Association. This paper is not the copy of record and may not exactly  
20 replicate the final, authoritative version of the article. Please do not copy or cite without authors' permission.  
21 The final article will be available, upon publication, via its DOI: 10.1037/rev0000347

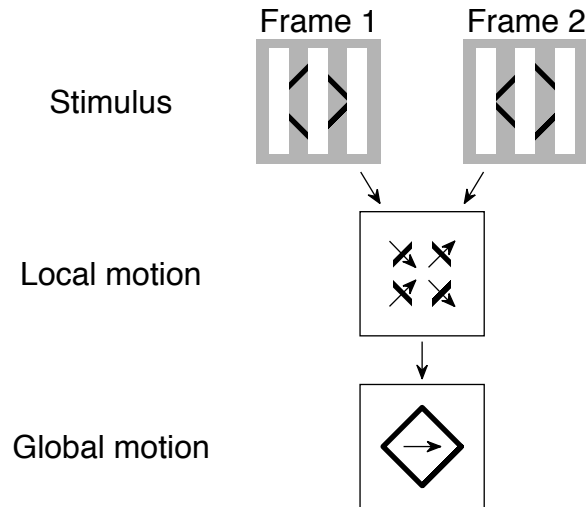
## 22 **Abstract**

23 Early direction-selective neurons in the primary visual cortex are widely considered to be the  
24 main neural basis underlying motion perception even though motion perception can also rely on  
25 attentively tracking the position of objects. Because of their small receptive fields, early  
26 direction-selective neurons suffer from the aperture problem, which is assumed to be overcome  
27 by integrating inputs from many early direction-selective neurons. Because the perceived motion  
28 of objects sometimes depends on static form information and does not always match the mean  
29 direction of local motion signals, the general consensus is that motion integration is form-  
30 dependent and complex. Based on the fact that early direction-selective neurons respond to  
31 motion only within a short temporal window, the current study used stroboscopic motion to test  
32 their contribution to motion perception of objects. For conditions under which the perceived  
33 motion was impaired by stroboscopic motion, the perceived motion matched the mean direction  
34 of local motion signals and was form-independent. For classic conditions under which the  
35 perceived motion could not be explained by a simple form-independent averaging of local motion  
36 signals, neutralizing the contribution of early direction-selective neurons using stroboscopic  
37 motion had little impact on the perceived motion, which demonstrates that the perceived motion  
38 relied on position tracking, not on early direction-selective neurons. When the perceived motion  
39 relies on position tracking, assuming that motion perception relies on early direction-selective  
40 neurons can lead to erroneously postulate the existence of complex or form-dependent integration  
41 of inputs from early direction-selective neurons.

42 **Keywords:** Motion perception; Aperture problem; Direction-selective neurons; Position tracking;  
43 Stroboscopic motion; Global motion

## 44 **Introduction**

45 Since Hubel and Wiesel (Hubel & Wiesel, 1959, 1962) recorded early direction-selective neurons  
46 in the primary visual cortex, such neurons are considered to be the main neural basis underlying  
47 motion perception. Early direction-selective neurons are sensitive to motion energy within their  
48 receptive fields (Adelson & Bergen, 1985; Emerson, Bergen, & Adelson, 1992) and within a short  
49 temporal window (<100 msec; Bair & Movshon, 2004; Baker & Braddick, 1985; Braddick, 1973;  
50 Georgeson & Harris, 1990; Morgan & Ward, 1980). However, because their receptive fields are  
51 small, they suffer from the famous aperture problem (Adelson & Movshon, 1982; Fennema &  
52 Thompson, 1979; Marr & Ullman, 1981; Wallach, 1935; Wuerger, Shapley, & Rubin, 1996): local  
53 motion signals are often ambiguous and do not necessarily reveal the global motion direction of  
54 the object (e.g., Video 1). The general consensus is that this aperture problem is overcome by  
55 integrating local motion signals into a global moving object (e.g., global motion stage in Figure 1),  
56 but the precise nature of this integration remains highly debated (see various models cited below).  
57 However, here we show that the perceived global direction of simple objects that would generally  
58 be assumed to be processed by early direction-selective neurons did not rely on early direction-  
59 selective neurons.



60

61 **Figure 1. Motion processing based on local motion signals. An object moving in a direction (here a partially**  
 62 **occluded black diamond moving from left to right) often induces local motion signals in different directions**  
 63 **(aperture problem, Video 1). Local motion signals can be detected by early direction-selective neurons and**  
 64 **can be integrated at a global motion stage to solve the aperture problem. Various integration strategies have**  
 65 **been suggested (see text).**

66 The simplest motion integration strategy consists in a vector averaging (or summation) of local  
 67 motion signals detected by early direction-selective neurons (Adelson & Movshon, 1982; Amano,  
 68 Edwards, Badcock, & Nishida, 2009; Weiss, Simoncelli, & Adelson, 2002). However, the  
 69 perceived motion direction of an object often does not match the mean direction of local motion  
 70 signals. A classic example is a drifting tilted line (Wallach, 1935; Wuerger et al., 1996), which can  
 71 be perceived as moving in its veridical direction when viewed entirely and perceived as moving in  
 72 a direction orthogonal to the line orientation when viewed through an aperture (Video 1). The  
 73 difference between the perceived motion direction and the mean direction of local motion signals  
 74 has led to the general consensus that, at least under some conditions, the aperture problem is  
 75 overcome by an integration process (i.e., global motion stage in Figure 1) more complex than vector  
 76 averaging. It has been suggested that the visual system could rely on non-ambiguous local motion

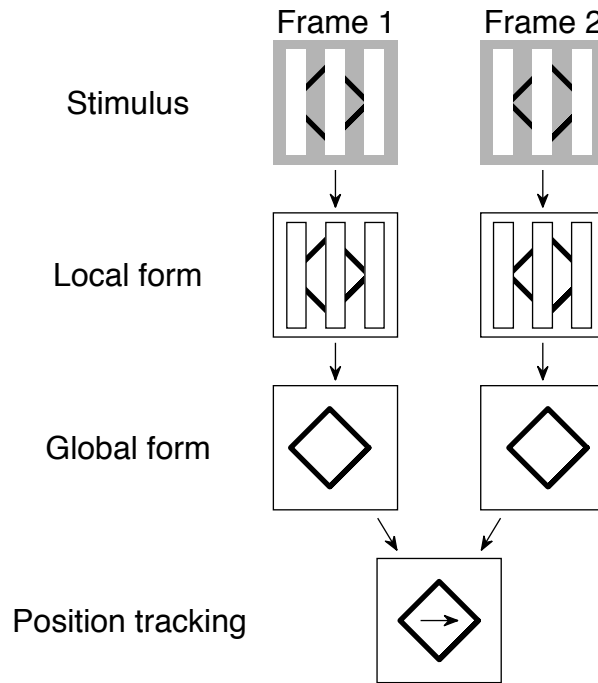
77 signals such as corners or line terminators (extracted by end-stopped cells (Caplovitz & Tse, 2007;  
78 Lorenceau, Shiffrar, Wells, & Castet, 1993; Pack, Gartland, & Born, 2004; Pack, Livingstone,  
79 Duffy, & Born, 2003; Shimojo, Silverman, & Nakayama, 1989; Ullman, 1979)), select the most  
80 reliable local motion signals (Nowlan & Sejnowski, 1995), compute the intersection of constraints  
81 (Adelson & Movshon, 1982; Nakayama & Silverman, 1988), compute differentiation between  
82 local motion signals (Anstis, 2012; Braddick, 1993) or use predictive coding based on some prior  
83 assumptions for the binding of segments (e.g., a regular shape is more probable than a random one)  
84 and local motion signals (Perrinet & Masson, 2012).

85 Furthermore, the perceived motion does not always depend solely on motion information as  
86 different stimuli containing the same motion information and differing only on spatial information  
87 can result into radically different motion percepts. For instance, the perceived motion direction of  
88 a partially occluded object can depend on the disparity (Shimojo et al., 1989), visibility (Lorenceau  
89 & Alais, 2001; Lorenceau & Shiffrar, 1992) and shape (McDermott, Weiss, & Adelson, 2001) of  
90 stationary occluders. Famously, the global motion percept of a rotating diamond having its corners  
91 occluded can be easily perceived when the occluders are visible (e.g., first sequence in Video 4),  
92 but not when they blend with the background (Lorenceau & Alais, 2001; Lorenceau & Shiffrar,  
93 1992) (e.g., first sequence in Video 5). Varying the visibility of the stationary occluders by only  
94 changing their color or luminance can drastically affect the perceived motion. Furthermore, altering  
95 the spatial configuration of moving parts without altering the vector mean of local motion signals  
96 can also drastically impair the global motion percept (Lorenceau & Alais, 2001) (e.g., first  
97 sequence in movie Video 6). These form-motion interactions were taken as evidence that motion  
98 integration (i.e., global motion stage in Figure 1) may not only be complex, it can also be form-  
99 dependent as whether the local motion signals are integrated into a global moving object or not

100 would depend on spatial information (e.g., depth, shape, luminance, color and spatial  
101 configuration).

102 In sum, the common assumption that the perceived global motion relies on inputs from early  
103 direction-selective neurons led to the general consensus that the aperture problem is sometimes  
104 solved by complex or form-dependent integration of local motion signals (Figure 1). Although  
105 early direction-selective neurons are assumed to be the main underlying neural basis of motion  
106 perception, motion perception does not necessarily rely on early direction-selective neurons as the  
107 position of a feature can be attentively tracked (Cavanagh, 1992) and the existence of a position-  
108 based motion system that does not rely on early direction-selective neurons is well established  
109 (Allard & Faubert, 2013a, 2016; Anstis & Mackay, 1980; Cavanagh, 1992; Lu & Sperling, 2001;  
110 Seiffert & Cavanagh, 1998; Smith, 1994; Ullman, 1979). However, tracking a feature faces the  
111 same aperture problem as early direction-selective neurons when the motions of features are  
112 ambiguous (e.g., global direction of the diamond cannot be inferred by tracking an edge in Figure  
113 1). Thus, motion integration would be required to solve the aperture problem for some stimuli (e.g.,  
114 Figure 1) whether ambiguous local motion of features is extracted by early direction-selective  
115 neurons or tracked. However, there is no reason why the position of a simple feature (e.g., edge)  
116 could be tracked, but not the position of a global shape (e.g., diamond, Figure 2). Given that a shape  
117 can be localized in space based only on spatial information (e.g., a static shape can be localized in  
118 space), the position of the global shape could be tracked in which case the perceived global motion  
119 would *not* rely on the integration of local motion signals so there would be no aperture problem to  
120 solve. The first experiment of the current study investigated the contribution of early direction-  
121 selective neurons and position tracking to the perceived motion of a simple object: a spot. The  
122 second experiment investigated which of these motion systems is responsible for the famous form-

123 motion interaction described above and typically taken as evidence of a form-dependent integration  
124 of inputs from early direction-selective neurons.



125

126 **Figure 2. Motion processing based on tracking the position of a global object. The perceived trajectory of an**  
127 **object can be computed without first extracting local motion signals (as in Figure 1). An object can be**  
128 **localized by integrating local form information (e.g., edges) and its global shape can then be tracked.**

## 129 **Experiment 1: Reverse-phi and attentional resolution**

130 The first experiment investigated how we perceive the motion direction of spots rotating around a  
131 fixation point (left panel of Figure 3).



## 132 **Methods**

### 133 **Observers**

134 Seven naïve observers with normal or corrected-to-normal vision participated in each experiment.  
135 Only a few participants were necessary because the goal was to objectively confirm strong  
136 subjective perceptual effects. Note that the reader can directly experience the strong perceptual  
137 effects reported in the results section by viewing the videos provided in supplementary material.  
138 The experimental procedures were approved by the Comité de Protection des Personnes Ile de  
139 France V in accordance with the Code of Ethics of the World Medical Association (Declaration  
140 of Helsinki). Informed consent was obtained from participants prior to the experiment.

### 141 **Apparatus**

142 Stimuli were presented on a 22.5-inch LCD monitor designed for psychophysics (VIEWPixx)  
143 with a refresh rate of 120 Hz. The viewing distance was 50 cm resulting in spatial resolution of  
144 32 pixels per degrees of visual angle (dva). The monitor was the only source of light in the room  
145 ( $50 \text{ cd/m}^2$ ). The output intensity of each color gun was carefully linearized (gamma correction).

### 146 **Stimuli and procedure**

147 Observers were asked to report the rotating direction (clockwise or counterclockwise) of spots  
148 along an annulus of 8 dva (Figure 3) from a fixation point. The spots were blobs following a  
149 windowed cosine function of 1 dva. The ring of spots was rotated by 2.8 degrees (i.e., 0.39 dva)  
150 every 150 msec for a duration of 1200 msec. This saccadic motion at 6.7 Hz (1/150 msec) is  
151 slightly different from continuous motion, but is nevertheless expected to strongly activate early  
152 direction-selective neurons. The lapses between position updates (150 msec) could not be varied

153 over a large range as it needed to be more than 100 msec to enable stroboscopic motion with gaps  
154 of at least 100 msec (see below). The rotating speed was maximized by displacing the spots by  
155 2.8 degrees (0.39 dva) at every step, which is a quarter of the smaller distance between spots at  
156 the highest density (32 spots). A step size of half the distance between spots would have caused  
157 ambiguous motion.

158 To investigate the contribution of the tracking system, the ability to perceive motion was  
159 investigated for different spot densities: 1, 2, 4, 8, 16 and 32 spots uniformly distributed on an  
160 annulus. Position tracking is inefficient when the number of feature to attend is above the spatial  
161 resolution of attention (Allard & Faubert, 2013a, 2016), which corresponds to about 14 spots  
162 distributed on an annulus (Intriligator & Cavanagh, 2001).

163 To investigate the contribution of the energy-based motion system, two manipulations were  
164 performed. The contrast polarity of the spots (black or white) was either constant during a trial  
165 (i.e., phi motion), or alternated at each displacement causing motion energy to be reversed (i.e.,  
166 reversed-phi motion; Anstis, 1970; Anstis & Rogers, 1975), which reverses the preferred motion  
167 direction of early direction-selective neurons (Clark, Bursztyn, Horowitz, Schnitzer, & Clandinin,  
168 2011; Emerson, Citron, Vaughn, & Klein, 1987; Livingstone & Conway, 2003). The second  
169 manipulation consisted in presenting the spots continuously during the trial or stroboscopically  
170 (25 msec followed by a 125 msec temporal gap) as illustrated in Figure 3. In both cases, the spots  
171 were displaced every 150 msec. Early direction-selective neurons are known to be sensitive to  
172 energy motion occurring within their receptive fields and within a temporal window shorter than  
173 about 100 msec (Bair & Movshon, 2004; Baker & Braddick, 1985; Braddick, 1973; Georgeson &

174 Harris, 1990), so a temporal gap of 125 msec neutralizes (or at least severely compromise) their  
175 contribution to motion perception.

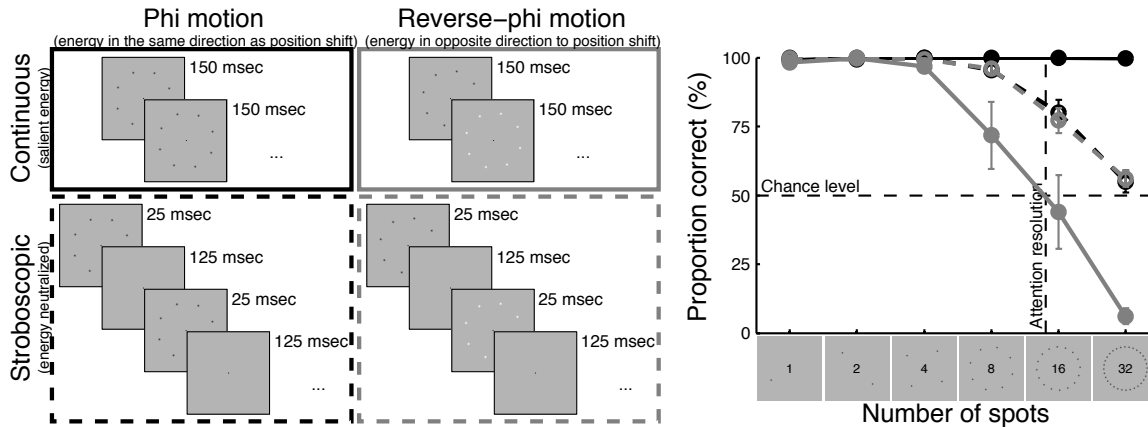
176 The resulting 24 conditions (6 spot densities times 2 types of motion (phi or reversed-phi) times 2  
177 display conditions (continuous or stroboscopic)) were each presented 5 times in a pseudo-random  
178 order within a block. Each observer performed 8 blocks of trials for a total of 40 trials per  
179 condition.

180 This study was not preregistered and the data are available upon request to the corresponding  
181 author.

## 182 ***Results & discussion***

183 When the spot density was considerably below attentional resolution (up to 4 spots, see Figure 3  
184 and Video 2), motion was perceived in their veridical direction: reversing energy direction (i.e.,  
185 reverse-phi motion) and/or compromising the contribution of motion detectors sensitive to  
186 motion only within a short temporal window (i.e., stroboscopic motion) had little impact on the  
187 perceived motion direction (performance near 100% in all cases). Results were drastically  
188 different when the spot density was considerably above attentional resolution (i.e., 32 spots, see  
189 Video 3) in which case reversing energy direction reversed the perceived motion direction  
190 (performance near 0%, filled grey symbols in Figure 3) and stroboscopic motion compromised  
191 motion perception (performance near chance, open symbols). Although the distance between  
192 spots was the lowest in the 32-spot configuration, lateral interaction (e.g., Chemla et al., 2019)  
193 between spots is unlikely the cause because the distance was nevertheless relatively large (1.57  
194 dva) and the results were expected based on the properties of early direction-selective neurons

195 (reverse-phi reverses their preferred response direction and their temporal window is shorter than  
 196 100 msec).



197  
 198 **Figure 3. Perceived motion direction of apparent motion. Left panel represents the four motion conditions,**  
 199 **each tested with 1, 2, 4, 8, 16 and 32 spots uniformly distributed around an annulus of 8 degrees of visual**  
 200 **angle of eccentricity. Right panel represents the results averaged across observers. Filled and open symbols**  
 201 **represent continuous and stroboscopic motion, respectively. Black and grey symbols represent phi motion**  
 202 **(constant contrast polarity) and reverse-phi motion (alternating contrast polarity at each motion**  
 203 **displacement), respectively. Error bars represent standard error of the mean (lower than symbols where not**  
 204 **visible). Perceiving motion in the opposite direction as the position displacement would have resulted in a**  
 205 **performance at 0%. Perceiving no net motion would have resulted in performance near chance (dashed thin**  
 206 **horizontal line). Dashed thin vertical line represents approximated attentional resolution (Intriligator &**  
 207 **Cavanagh, 2001).**

208 The fact that no motion was perceived with 32 spots under stroboscopic stimulation (performance  
 209 near chance and observers reported no motion sensation as can be appreciated in the second and  
 210 fourth sequences of Video 3) implies that the contributions of direction-selective neurons and  
 211 position tracking were both negligible in these conditions. Presenting the same 32-spot stimuli  
 212 continuously led to perceived motion in the energy direction: phi motion was perceived in the

213 veridical direction (first sequence in Video 3, black solid line in Figure 3) and reverse-phi motion  
214 (Anstis, 1970; Anstis & Rogers, 1975) was perceived in the opposite direction (third sequence in  
215 Video 3, grey solid line in Figure 3). The fact that no motion was perceived with 32 spots under  
216 stroboscopic stimulation and motion was perceived in the energy direction under continuous  
217 stimulation, confirms that stroboscopic motion successfully impaired the contribution of early  
218 direction-selective neurons. Indeed, these results are consistent with the facts that early direction-  
219 selective neurons operate within a short temporal window (Bair & Movshon, 2004; Baker &  
220 Braddick, 1985; Braddick, 1973; Georgeson & Harris, 1990; Morgan & Ward, 1980) and that  
221 reverse-phi indeed reverses their direction selectivity (Clark et al., 2011; Emerson et al., 1987;  
222 Livingstone & Conway, 2003). Thus, there is no doubt that early direction-selective neurons were  
223 the neural basis underlying motion perception in the continuous 32-spot conditions (as generally  
224 assumed) and that stroboscopic motion compromised the contribution of early direction-selective  
225 neurons (as expected).

226 Given that early direction-selective neurons did not considerably contribute to motion perception  
227 under stroboscopic stimulation, stroboscopic motion can be used to estimate the contribution of a  
228 position tracking strategy. Position tracking considerably contributed to motion only when the  
229 spot density was considerably below attentional resolution of about 14 elements (Intriligator &  
230 Cavanagh, 2001). Indeed, when the contribution of early direction-selective neurons was  
231 minimized (stroboscopic motion, open symbols in Figure 3), motion was correctly perceived  
232 when the spot density was lower than attentional resolution (i.e., up to 4 or 8 spots depending on  
233 participants; Video 2) and not when density was considerably above attentional resolution (i.e.,  
234 32 spots; Video 3). These results show that position tracking considerably contributed to motion

235 perception only when spot density was markedly below attentional resolution, which is consistent  
236 with previous findings (Allard & Faubert, 2013a, 2016).

237 Consequently, continuous reverse-phi motion was perceived as moving in its veridical direction  
238 up to 4 or 8 spots, and perceived in the opposite direction with 32 spots. With 32 spots, motion  
239 was perceived in the opposite direction undeniably because reverse-phi reverses the direction  
240 selectivity of early direction-selective neurons (Clark et al., 2011; Emerson et al., 1987;  
241 Livingstone & Conway, 2003). Up to 4 or 8 spots, however, the perceived motion relied on  
242 position tracking as motion was perceived in the veridical direction, not the preferred direction of  
243 early direction-selective neurons. These results therefore suggest that an object can be attentively  
244 tracked despite its polarity flickering over time, which is consistent with the fact that tracking can  
245 operate despite the object changing its color and shape (Cavanagh, Arguin, & von Grünau, 1989;  
246 Kolars & Pomerantz, 1971; Kolars & von Grünau, 1976). Around the attentional resolution (i.e.,  
247 16 spots), there was a high inter-subject variability (large error bars in Figure 3) compared with  
248 the stroboscopic conditions. Interestingly, in this condition, the two motion systems would  
249 similarly contribute to motion, but in opposite directions. The high inter-subject variability  
250 suggests different observers had considerably different thresholds at which the contribution of the  
251 two motion systems would be counterbalanced. Note that although performance is near chance  
252 level, it does not imply that no motion was perceived (as for stroboscopic motion at 32 spots) as  
253 motion perception could be perceived in both directions (i.e., transparent motion) or a bistable  
254 motion percept could be perceived (alternating directions).

255 Note that the current experiment was conducted under some specific parameters. The eccentricity  
256 was set to 8 dva, but the findings likely generalize to a wide range of eccentricity for a given

257 number of spots as spatial resolution scales with eccentricity (Intriligator & Cavanagh, 2001; e.g.,  
258 viewing Videos 2 and 3 at different distances has little impact on the perceived motion). Also,  
259 saccadic motion was used instead of continuous motion, but the results confirm that this stimulus  
260 causes a strong motion response in early direction-selective neurons as the observers could easily  
261 discriminate the direction of motion when the contribution of position tracking was neutralized  
262 with 32 spots (performance at 100%). The use of this saccadic motion, which was necessary for  
263 stroboscopic and reverse-phi motion, forced the use of a relatively slow speed (2.6 dva/s) so it is  
264 not possible to know if the perceptual effects would generalize to higher speeds. Nonetheless, we  
265 should expect the contribution of position tracking to decrease with increasing speed.

## 266 **Experiment 2: Motion integration**

267 In the second experiment, we investigated how local information is integrated into a global  
268 moving object using a stimulus classically used to probe motion integration: a rotating diamond  
269 with static occluders hiding its corners (Lorenceanu & Alais, 2001; Lorenceanu & Shiffrar, 1992;  
270 McDermott et al., 2001). As in the first experiment, stroboscopic motion was used to neutralize  
271 the contribution of early direction-selective neurons. To minimize the contribution of a position  
272 tracking strategy, the same stimuli were viewed peripherally where tracking relatively small  
273 position displacements is difficult (Allard & Faubert, 2016).

### 274 ***Method***

#### 275 **Observers**

276 The same observers participated to the second experiment.

277 **Apparatus**

278 The same apparatus was used with the exception that the viewing distance was 100 cm resulting  
279 in spatial resolution of 64 pixels/dva.

280 **Stimuli and procedure**

281 Observers were asked to discriminate the rotating direction (clockwise vs counterclockwise) of  
282 the position of a shape by pressing one of two keyboard keys. The stimuli for continuous motion  
283 were similar to the ones used in previous studies (Lorenceanu & Alais, 2001; Lorenceanu et al.,  
284 1993), but we used smaller stimuli and smaller amplitude of motion to facilitate the neutralization  
285 of position tracking in peripheral vision. The stimulus was visible for 1100 msec during which  
286 the shape rotated by a complete cycle around a radius of 0.0625 dva (i.e.,  $\sim 3$  minutes of arc = 4  
287 pixels), with a random initial global position. The corners of the shape were occluded by three  
288 vertical occluders. Given that each line segment was perceived through a narrow opening  
289 between two occluders, the motion of a single line was ambiguous (i.e., aperture problem, Video  
290 1) and perceiving the global motion therefore required some integration. It has been shown  
291 (Lorenceanu & Alais, 2001) that the perception of the global motion depends on the visibility of  
292 the occluders and the shape of the moving object. The visibility of the occluders was altered by  
293 changing their brightness, which was either grey (50 cd/m<sup>2</sup>, in which case they fused with the  
294 background) or light grey (55 cd/m<sup>2</sup>). The baseline shape was a diamond (more precisely, a  
295 square rotated by 45 degrees) to which the distance between two parallel segments was 2 dva.  
296 Two shapes were used: a closed shape (i.e., diamond) and an open shape (i.e., chevron  
297 (Lorenceanu & Alais, 2001)), which was obtained by swapping the positions of the two bottom  
298 lines. The line segments of the rotating shape were black with a width profile following a  
299 windowed cosine function of 0.3 dva. The size of the occluders was 0.75 by 3.6 dva separated by



300 gaps of 0.67 dva so that the center of the occluders were in line with the mean left, center and  
301 right corner positions.

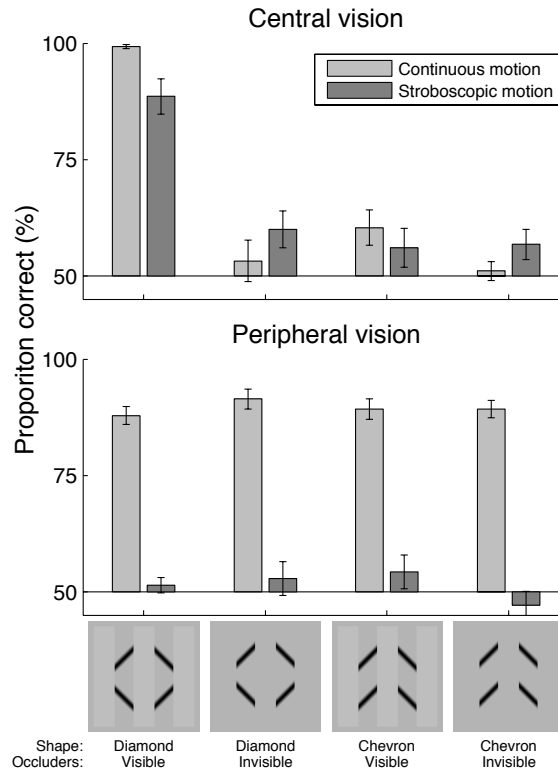
302 To investigate the contribution of energy-based processing, motion was either continuous  
303 (displaced every 33 msec) or stroboscopic (appeared for 33 msec with 100 msec gaps). The  
304 global motion speed was kept constant to 0.9375 Hz in both cases so that the object was moved  
305 by a complete cycle during the 1100 msec presentation.

306 To investigate the contribution of the position tracking motion system, the stimulus was viewed  
307 either at fixation or at 20 degrees of eccentricity to the right of a fixation point. Position tracking  
308 of oscillating motion over a short spatial range is undeniably more difficult in the periphery  
309 (Allard & Faubert, 2016), which was confirmed by the results presented below.

310 In sum, there were 4 binary independent variables resulting into 16 conditions: shape of the  
311 object (diamond/chevron), brightness of the occluders (visible/invisible), motion type  
312 (continuous/stroboscopic) and eccentricity (central/peripheral). To avoid position uncertainty,  
313 eccentricity was constant within a block. To avoid triggering strong transients not related to the  
314 moving stimulus, the occluders were constantly displayed throughout a block (even between  
315 trials) and its brightness did not vary within a block. The other two variables (shape and motion  
316 type) were interlaced within blocks resulting into 4 different blocks (visible/invisible occluders X  
317 central/peripheral) each interlacing 4 different conditions (diamond/chevron X  
318 continuous/stroboscopic). A block was composed of 80 trials in which the 4 conditions were  
319 presented 20 times in a random order. Each of the four blocks was performed twice in a pseudo-  
320 random order, resulting in a total of 40 trials for each of the 16 conditions.

321 ***Results & discussion***

322 The results were consistent with the known form-motion interaction taken as evidence that spatial  
323 information affects motion integration (Lorencean & Alais, 2001; Lorencean & Shiffrar, 1992):  
324 the global motion of the continuously rotating diamond viewed centrally was easily perceived  
325 when the occluders were visible (first light-grey column in Figure 4, first sequence in Video 4),  
326 but not when they were invisible (same brightness as the background) or when the spatial  
327 configuration was altered to form an open shape (second to fourth light-grey columns, first  
328 sequences in Videos 5 to 7; note that for open shapes, discriminating the global motion direction  
329 seems easier in the videos than in the testing conditions with a shorter presentation time and  
330 interlaced with other conditions). Critically, however, a similar form-motion interaction was also  
331 observed under stroboscopic stimulation compromising the contribution of early direction-  
332 selective neurons (dark-grey columns, second sequences of the same movies). Thus, the global  
333 motion percept depended critically on form information (i.e., brightness of occluders and spatial  
334 configuration) and not on the type of motion (i.e., continuous vs stroboscopic).



335

336 **Figure 4. Perceived motion direction requiring integration of local information. Proportion of correct answers**  
 337 **(mean across observers) for judging the rotating direction of various stimuli (diamond and chevron shapes**  
 338 **with visible and invisible occluders; see Videos 4 to 7). Error bars represent standard error of the mean.**  
 339 **Perceiving no motion would have resulted in performance near chance (50%, horizontal line).**

340 The opposite pattern of results was observed when the same stimuli were viewed peripherally  
 341 where tracking relatively small position displacements is difficult (Allard & Faubert, 2016)  
 342 (bottom graph of Figure 4). Global motion was efficiently perceived for continuous motion  
 343 (light-grey columns), but not for stroboscopic motion (dark-grey columns), whether the occluders  
 344 were visible or not and for the two shape configurations. Thus, in central vision, the perception of  
 345 global motion strongly depended on form information for both types of motion, whereas in the  
 346 peripheral vision, the perception of global motion critically depended on the type of motion and  
 347 was independent of form information.

348 This double dissociation confirms that stroboscopic motion impaired the contribution of early  
349 direction-selective neurons and presenting the stimulus in the periphery impaired position  
350 tracking. Indeed, when stroboscopic motion was viewed peripherally, no motion was perceived  
351 (dark-grey columns in bottom graph of Figure 4), which can be appreciated by peripherally  
352 viewing the second sequences of Videos 4 to 7. This absence of motion perception implies that  
353 the contribution of early direction-selective neurons and position tracking to the global motion  
354 percept were both negligible. On the other hand, motion was efficiently perceived when  
355 presenting the same stimuli continuously (light-grey columns, first sequences of same movies),  
356 which indisputably results from the contribution of early direction-selective neurons as generally  
357 assumed. When early direction-selective neurons were not contributing to motion (stroboscopic  
358 motion), rotation of the diamond with visible occluders was easily perceived at fixation (first dark  
359 column, top graph), but not in the periphery (first dark column, bottom graph), implying that the  
360 tracking efficiently operated at fixation, but not in the periphery.

361 Given that some properties of early direction-selective neurons vary with eccentricity, it is  
362 tempting to wonder if these differences could explain the different patterns of results observed  
363 between central and peripheral vision. However, different properties in early direction-selective  
364 neurons cannot explain the differences observed between central and peripheral vision because  
365 the perceived motion in central vision did not rely on early direction-selective neurons. Indeed,  
366 one property of early direction-selective neurons that does not vary with eccentricity is their  
367 temporal integration time (Baker & Braddick, 1985), so that stroboscopic vision with temporal  
368 gaps of at least 100 msec neutralizes the contribution of early direction-selective neurons in  
369 central and peripheral vision (as confirmed in peripheral vision, see above). Indeed, the results  
370 strongly suggest that the perceived motion in central vision relied on position tracking, not early

371 direction-selective neurons. Nevertheless, it could be useful to neutralize position tracking in  
372 central vision where the perceived motion is form-dependent. This could be done using high  
373 temporal frequencies above the temporal resolution of attention (Allard & Faubert, 2013b) or  
374 using a spatially dense stimulus such as concentric diamonds.

375 Note that Lorenceau and Alais (2001) obtained a similar pattern of results with continuous  
376 motion (perceived motion was form-dependent in central vision and not in peripheral vision)  
377 using different stimulus parameters: their stimuli were larger, the rotation amplitude was greater,  
378 different speeds were evaluated and the peripheral stimulus was presented at a lower eccentricity  
379 (12 dva compared with 20 dva in the current study). Thus, although the current study did not  
380 investigate over a large set of parameters, the effects likely generalize over a wide range of  
381 parameters.

## 382 **General discussion**

383 With continuous motion under conditions in which position tracking was expected to efficiently  
384 operate (i.e., up to 4 spots in experiment 1 and in central viewing in experiment 2), the perceived  
385 motion critically depended on form information (brightness of occluders and spatial configuration  
386 in experiment 2) and not on the direction of motion energy (phi vs reverse-phi in experiment 1).  
387 Neutralizing the contribution of early direction-selective neurons using stroboscopic motion  
388 under these conditions had little impact on the perceived motion as the perceived motion also  
389 critically depended on form information and not on the direction of motion energy. These results  
390 reveal the existence of a motion system that does not rely on early direction-selective neurons  
391 (position tracking) and is form-dependent. Conversely, under conditions in which the  
392 contribution of position tracking was expected to be neutralized (32 spots in experiment 1 and

393 peripheral viewing in experiment 2), the perceived motion was independent of form information  
394 (experiment 2) and critically depended on the direction of motion energy (experiment 1). The fact  
395 that no motion was perceived when neutralizing the contribution of early direction-selective  
396 neurons using stroboscopic motion under these conditions (both experiments) confirms that the  
397 visual system relied on early direction-selective neurons as generally assumed.

398 The widely held assumption that the visual system relies on early direction-selective neurons to  
399 judge the motion direction of objects can lead to a false aperture problem when the visual system  
400 rather mainly relies on position tracking: how are inputs from early direction-selective neurons  
401 integrated to explain the perceived motion direction of objects? Forcing a solution to this false  
402 problem could lead to erroneously postulate the existence of some kind of integration of inputs  
403 from early direction-selective neurons. For instance, consider the continuous motion conditions  
404 for which the perceived motion cannot be explained by a simple vector averaging of inputs from  
405 early direction-selective neurons: continuous motion up to 4 spots in experiment 1 and  
406 continuous motion viewed centrally in experiment 2. Assuming that the perceived motion  
407 direction relied on early direction-selective neurons would have led to the conclusion that the  
408 integration of inputs from early direction-selective neurons must be complex (i.e., not a simple  
409 averaging of inputs from early direction-selective neurons) to explain the perceived motion in the  
410 veridical direction even when the motion energy is reversed (experiment 1) and must be form-  
411 dependent to explain that the perceived motion depended on spatial information (experiment 2) as  
412 famously inferred (e.g., a form-based veto on motion integration; Lorenceau & Alais, 2001).  
413 However, the same patterns of results were observed when the contribution of early direction-  
414 selective neurons was neutralized with stroboscopic motion, which reveals the existence of  
415 another motion system (position tracking) that is form-dependent and independent of the

416 direction of motion energy. Before investigating how inputs from early direction-selective  
417 neurons are integrated to explain the perceived motion, it is first necessary to determine if the  
418 perceived motion relies on early direction-selective neurons. The results in the current study  
419 show that, for many simple moving stimuli, the widely held assumption that the perceived motion  
420 relies on early direction-selective neurons is questionable and cannot be taken for granted a  
421 priori. If the perceived motion does not rely on early direction-selective neurons, then the  
422 question of how inputs from early direction-selective neurons are integrated to explain the  
423 perceived motion is a false problem.

424 As mentioned in the introduction, a classical example of the aperture problem is a drifting tilted  
425 line (Video 1). When the line is perceived as moving in its veridical direction (e.g., horizontally  
426 when viewing the entire line in Video 1), the perceived motion cannot be explained by a simple  
427 vector averaging of inputs from early direction-selective neurons, as the mean of local motion  
428 signals is oblique. Indeed, the motion direction of a drifting line viewed through an aperture is  
429 ambiguous as it is consistent with many motion directions ranging over 180 degrees, but the  
430 perceived motion direction through an aperture is usually perpendicular to the line (e.g., oblique  
431 motion is perceived when viewing the line through an aperture in Video 1). Based on the  
432 assumption that the perceived motion relies on inputs from early direction-selective neurons, this  
433 example is typically taken as compelling evidence of an integration process more complex than  
434 vector averaging. However, neutralizing the contribution of early direction-selective neurons with  
435 stroboscopic stimulation has little impact on the perceived motion (see Video 8), which implies  
436 that the veridical motion direction of the line can be perceived by tracking its position. As a  
437 result, there is a motion system that does not rely on early direction-selective neurons and that  
438 can extract the veridical motion direction of the tilted line, so this classical example does not

439 necessarily imply the existence of complex integration of inputs from early direction-selective  
440 neurons.

441 The furrow illusion (Anstis, 2012) has also been taken as evidence of complex integration of  
442 local motion signals. In the furrow illusion, a spot moves in a direction that differs from the mean  
443 local motion signals. In central vision, motion is perceived in the veridical motion direction,  
444 whereas in the periphery, motion is perceived in the direction of the mean local motion signals.  
445 The illusion in the periphery is trivial to explain and is not disputed: the motion percept is the  
446 result of a simple vector averaging of inputs from early direction-selective neurons. For central  
447 vision, Anstis (2012) concluded that the perceived veridical motion must be the result of a  
448 complex integration of inputs from early direction-selective neurons. This conclusion is based on  
449 the implicit assumption that the perceived motion relies on the integration of inputs from early  
450 direction-selective neurons. However, motion can be perceived in the veridical direction even  
451 when neutralizing the contribution of early direction-selective neurons with stroboscopic motion  
452 (Allard & Faubert, 2016), which implies the existence of a motion system that does not rely on  
453 early direction-selective neurons and can extract the veridical motion direction of the spot.  
454 Furthermore, Allard and Faubert (2016) found that using a stimulus configuration that impairs  
455 position tracking (crowding) enhanced the furrow illusion, which suggests that position tracking  
456 plays an important role in perceiving the veridical motion direction. Consequently, in the furrow  
457 illusion, there is no need to postulate the existence of complex integration of inputs from early  
458 direction-selective neurons to explain the perceived veridical motion in central vision because  
459 there is another motion system (position tracking) that can explain it.



460 In sum, in these examples (reverse-phi stimulus up to 4 spots, form-dependent rotating square in  
461 central vision, drifting tilted line, furrow illusion in central vision), similar motion phenomena  
462 were observed when neutralizing the contribution of early direction-selective neurons with  
463 stroboscopic motion. These results demonstrate that there is another motion system that does not  
464 rely on early direction-selective neurons (position tracking) responsible for these motion  
465 phenomena so they do not imply the existence of a particular kind of integration of inputs from  
466 early direction-selective neurons. To prove the existence of some kind of integration of inputs  
467 from early direction-selective neurons, it is necessary to demonstrate that the motion  
468 phenomenon cannot be explained by another motion system (e.g., position tracking). The  
469 perceived motion cannot simply be assumed a priori to rely on early direction-selective neurons.

470 The ideal stimulus to probe the integration of inputs from early direction-selective neurons is  
471 therefore one that cannot be processed by position tracking. Interestingly, the current study did  
472 not find evidence that the integration of inputs from early direction-selective neurons was  
473 complex or form-dependent. Consider the conditions under which only position tracking was  
474 neutralized: continuous motion with 32 spots in experiment 1 and continuous motion with  
475 peripheral viewing in experiment 2. Observers perceived motion in the mean direction of motion  
476 energy (experiment 1) and the perceived global motion was form-independent (experiment 2).  
477 Consequently, when the perceived motion was not based on position tracking, the perceived  
478 motion relied on the integration of inputs from early direction-selective neurons (as generally  
479 assumed), but motion integration was simple (e.g., vector average) and form-independent. In  
480 other words, when the visual system necessarily relied on early direction-selective neurons, the  
481 solution to the aperture problem simply consisted in averaging the inputs from early direction-  
482 selective neurons; we found no evidence of complex or form-dependent integration of inputs

483 from early direction-selective neurons. However, the current study investigated a limited number  
484 of conditions, which is not sufficient to conclude that there is no complex or form-dependent  
485 integration of inputs from early direction-selective neurons. There may be other conditions that  
486 could reveal the existence of complex or form-dependent integration.

487 Although the perceived motion sometimes clearly relies solely on the contribution of early  
488 direction-selective neurons or position tracking, both could contribute to the perceived motion in  
489 other conditions. The combination of both contributions could take different forms. The  
490 contribution of one motion system could dominate (e.g., continuous reverse-phi motion up to 4  
491 spots in experiment 1 is dominated by position tracking). Conflicting motion interpretations could  
492 also result in transparent motion in which both motion percepts are superimposed (e.g., Solomon  
493 & Sperling, 1994) or the motion the perceived motion could be bistable, alternating between the  
494 two possible interpretations (which is the subjective impression we observed when viewing the  
495 reverse-phi stimulus with 16 spots). When the two motion interpretations are not incompatible,  
496 the visual system could combine them (e.g., furrow illusion that gradually shifts from position  
497 tracking in central vision to early direction-selective neurons in peripheral vision (Allard &  
498 Faubert, 2016)). The visual system may also try to combine the two interpretations and apply a  
499 drastic correction when the resulting percept diverges too much from reality (e.g., peripheral  
500 reverse-phi motion, Lorenceau & Cavanagh, 2020).

501 In a recent review, Shiffrar (2017) concludes that the solution to the aperture problem often  
502 involves a form-dependent integration of local motion signals. The strongest evidence for this  
503 claim was the rotating diamond showing that the perceived global motion percept is form-  
504 dependent (i.e., second experiment of the current study with continuous motion in central vision).

505 Based on the implicit assumption that the perceived global motion relied on the integration of  
506 inputs from early direction-selective neurons, this form-dependent motion perception was  
507 therefore taken as evidence of a form-dependent integration of local motion signals: the global  
508 form interpretation of the stimulus (global form stage in Figure 2) would affect the integration of  
509 local motion signals (global motion stage in Figure 1). As described above, however, the current  
510 study found that the perceived form-dependent global motion for a similar stimulus was based on  
511 position tracking, which does not require motion integration, so these results cannot be taken as  
512 evidence of a form-dependent integration of local motion signals.

513 The fact that the global motion percept based on position tracking was form-dependent is not  
514 surprising as the ability to track the position of a global shape (Figure 2) obviously depends on  
515 the ability to perceive the global shape. The diamond was perceived as a global moving shape  
516 when the occluders were visible and as four independent segments when the occluders were  
517 invisible (Videos 4 and 5, respectively), which shows that the visibility of occluders affected the  
518 global *form* interpretation of the stimulus. When the global form integration stage interpreted the  
519 stimulus as a single global shape (diamond; Video 4), then tracking the position of the global  
520 shape was easy, but when the global form integration stage rather interpreted the stimulus as four  
521 independent segments or two pairs of segments (Video 5), then there was no unified global  
522 percept and tracking the position of an unperceived global shape was meaningless (and hard).

523 Consequently, the fact that the perceived motion of a global shape depends on the form  
524 interpretation of the global shape would have been surprising if the visual system relied on early  
525 direction-selective neurons (because it implies that form processing affects motion integration),  
526 but is trivial to explain given that the visual system relied on position tracking (because we  
527 cannot track a global shape that is not perceived).

528 Since tracking the position of a global object does not necessarily require the integration of local  
529 motion signals (e.g., Figure 2), sometimes there may be no aperture problem to solve. Indeed, the  
530 famous aperture problem is known to be a consequence of the properties of early direction-  
531 selective neurons having small receptive fields (i.e., apertures). If the perceived motion of an  
532 object is based on tracking its global shape, then it does not rely on the integration of local  
533 motion signals so there is no aperture problem to solve. In such a case, the perceived motion  
534 would not rely on motion integration, but on form integration enabling to localize an object, and  
535 the critical processing stage regarding the motion perception of objects would not be motion  
536 integration, but form integration. Since we can localize static objects, the existence of a global  
537 form integration process is undeniable, but how objects are localized in space is beyond the scope  
538 of the present study.

539 In conclusion, early direction-selective neurons suffer from the famous aperture problem. Based  
540 on the assumption that early direction-selective neurons are the underlying neural basis for  
541 judging the motion direction of objects, the consensus is that the integration of inputs from early  
542 direction-selective neurons is complex and form-dependent. However, the current study found  
543 that for stimuli classically used to provide evidence of complex and form-dependent integration  
544 of inputs from early direction-selective neurons, neutralizing the contribution of early direction-  
545 selective neurons using stroboscopic motion had little impact on the perceived motion direction,  
546 which demonstrates the existence of a motion system that does not rely on early direction-  
547 selective neurons (e.g., position tracking) and explains the perceived motion. Note that the  
548 current study suggests that the motion system that does not rely on early direction-selective  
549 neurons could simply consists in tracking the position of an object localized by the spatial  
550 system, which we refer to as position tracking (Figure 2). Indeed, given that a shape can be

551 localized in space based only on spatial information (e.g., a static shape can be localized in  
552 space), the position of an object should be easy to track. Nonetheless, other motion models that  
553 are not based on early direction-selective neurons are also possible such as feature tracking (e.g.,  
554 Seiffert & Cavanagh, 1998), attention-based tracking (Cavanagh, 1992), saliency-based (third-  
555 order) motion system (Lu & Sperling, 2001), or long-range suppressive waves (Chemla et al.,  
556 2019). On the other hand, for particular conditions under which motion perception undeniably  
557 relied on early direction-selective neurons, no complex or form-dependent motion integration  
558 was observed, which suggests that, for the stimuli used in the current study, the integration of  
559 inputs from early direction-selective neurons was simple and form-independent. The aperture  
560 problem (how are inputs from early direction-selective neurons integrated to explain the global  
561 motion percept?) is a false problem if the perceived global motion does not rely on the integration  
562 of local motion signals (e.g., tracking the position of an object).

563 **References**

- 564 Adelson, E. H., & Bergen, J. R. (1985). Spatiotemporal energy models for the perception of  
565 motion. *J Opt Soc Am A*, 2(2), 284–299. Retrieved from  
566 <http://www.ncbi.nlm.nih.gov/pubmed/3973762>
- 567 Adelson, E. H., & Movshon, J. a. (1982). Phenomenal coherence of moving visual patterns.  
568 *Nature*, 300(5892), 523–525. <https://doi.org/10.1038/300523a0>
- 569 Allard, R., & Arleo, A. (2016a). Perception of global object motion without integration of local  
570 motion signals. In *European Conference on Visual Perception*.
- 571 Allard, R., & Arleo, A. (2016b). Position-based vs energy-based motion processing. *Journal of*  
572 *Vision*, 16(12), 670. <https://doi.org/10.1167/16.12.670>
- 573 Allard, R., & Faubert, J. (2013a). No dedicated second-order motion system. *Journal of Vision*,  
574 13(11). <https://doi.org/10.1167/13.11.2>
- 575 Allard, R., & Faubert, J. (2013b). No dedicated second-order motion system in the periphery. In  
576 *Vision Sciences Society*.
- 577 Allard, R., & Faubert, J. (2016). The Role of Feature Tracking in the Furrow Illusion. *Frontiers*  
578 *in Human Neuroscience*, 10, 81. <https://doi.org/10.3389/fnhum.2016.00081>
- 579 Amano, K., Edwards, M., Badcock, D. R., & Nishida, S. (2009). Adaptive pooling of visual  
580 motion signals by the human visual system revealed with a novel multi-element stimulus.  
581 *Journal of Vision*, 9(3), 4.1-25. <https://doi.org/10.1167/9.3.4>
- 582 Anstis, S. M. (1970). Phi movement as a subtraction process. *Vision Research*, 10(12), 1411–  
583 1430. [https://doi.org/10.1016/0042-6989\(70\)90092-1](https://doi.org/10.1016/0042-6989(70)90092-1)
- 584 Anstis, S. M. (2012). The furrow illusion: peripheral motion becomes aligned with stationary  
585 contours. *J Vis*, 12(12), 12. <https://doi.org/10.1167/12.12.12>
- 586 Anstis, S. M., & Mackay, D. M. (1980). The Perception of Apparent Movement [and  
587 Discussion]. *Philosophical Transactions of the Royal Society of London. Series B,*  
588 *Biological Sciences*, 290(1038), 153–168. <https://doi.org/10.1098/rstb.1980.0088>
- 589 Anstis, S. M., & Rogers, B. J. (1975). Illusory reversal of visual depth and movement during  
590 changes of contrast. *Vision Research*, 15(8–9). [https://doi.org/10.1016/0042-6989\(75\)90236-](https://doi.org/10.1016/0042-6989(75)90236-9)  
591 9
- 592 Bair, W., & Movshon, J. A. (2004). Adaptive Temporal Integration of Motion in Direction-  
593 Selective Neurons in Macaque Visual Cortex. *Journal of Neuroscience*, 24(33), 9305–9323.  
594 <https://doi.org/10.1523/JNEUROSCI.0554-04.2004>
- 595 Baker, C. L., & Braddick, O. J. (1985). Temporal properties of the short-range process in  
596 apparent motion. *Perception*, 14(2), 181–192. <https://doi.org/10.1068/p140181>

- 597 Braddick, O. J. (1973). The masking of apparent motion in random-dot patterns. *Vision Research*,  
598 13(2), 355–369. [https://doi.org/10.1016/0042-6989\(73\)90113-2](https://doi.org/10.1016/0042-6989(73)90113-2)
- 599 Braddick, O. J. (1993). Segmentation versus integration in visual motion processing. *Trends in*  
600 *Neurosciences*, 16, 263–268. [https://doi.org/10.1016/0166-2236\(93\)90179-P](https://doi.org/10.1016/0166-2236(93)90179-P)
- 601 Caplovitz, G. P., & Tse, P. U. (2007). V3A processes contour curvature as a trackable feature for  
602 the perception of rotational motion. *Cerebral Cortex*, 17(5), 1179–1189.  
603 <https://doi.org/10.1093/cercor/bhl029>
- 604 Cavanagh, P. (1992). Attention-based motion perception. *Science*, 257(5076), 1563–1565.  
605 Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/1523411>
- 606 Cavanagh, P., Arguin, M., & von Grünau, M. (1989). Interattribute apparent motion. *Vision*  
607 *Research*. [https://doi.org/10.1016/0042-6989\(89\)90065-5](https://doi.org/10.1016/0042-6989(89)90065-5)
- 608 Chemla, S., Reynaud, A., Di Volo, M., Zerlaut, Y., Perrinet, L., Destexhe, A., & Chavane, F.  
609 (2019). Suppressive traveling waves shape representations of illusory motion in primary  
610 visual cortex of awake primate. *Journal of Neuroscience*.  
611 <https://doi.org/10.1523/JNEUROSCI.2792-18.2019>
- 612 Clark, D. A., Bursztyn, L., Horowitz, M. A., Schnitzer, M. J., & Clandinin, T. R. (2011).  
613 Defining the Computational Structure of the Motion Detector in *Drosophila*. *Neuron*, 70(6),  
614 1165–1177. <https://doi.org/10.1016/j.neuron.2011.05.023>
- 615 Emerson, R. C., Bergen, J. R., & Adelson, E. H. (1992). Directionally selective complex cells and  
616 the computation of motion energy in cat visual cortex. *Vision Research*, 32(2), 203–218.  
617 [https://doi.org/10.1016/0042-6989\(92\)90130-B](https://doi.org/10.1016/0042-6989(92)90130-B)
- 618 Emerson, R. C., Citron, M. C., Vaughn, W. J., & Klein, S. A. (1987). Nonlinear directionally  
619 selective subunits in complex cells of cat striate cortex. *Journal of Neurophysiology*, 58(1),  
620 33–65.
- 621 Fennema, C. L., & Thompson, W. B. (1979). Velocity determination in scenes containing several  
622 moving objects. *Computer Graphics and Image Processing*, 9(4), 301–315.  
623 [https://doi.org/10.1016/0146-664X\(79\)90097-2](https://doi.org/10.1016/0146-664X(79)90097-2)
- 624 Georgeson, M. A., & Harris, M. G. (1990). The temporal range of motion sensing and motion  
625 perception. *Vision Research*, 30(4), 615–619. [https://doi.org/10.1016/0042-6989\(90\)90072-](https://doi.org/10.1016/0042-6989(90)90072-S)  
626 S
- 627 Hubel, D. H., & Wiesel, T. N. (1959). Receptive fields of single neurones in the cat's striate  
628 cortex. *Journal of Physiology*, 148, 574–591.
- 629 Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional  
630 architecture in the cat's visual cortex. *The Journal of Physiology*, 160(1), 106–154.  
631 <https://doi.org/10.1113/jphysiol.1962.sp006837>
- 632 Intriligator, J., & Cavanagh, P. (2001). The Spatial Resolution of Visual Attention. *Cognitive*

- 633 *Psychology*, 43(3), 171–216. Retrieved from  
634 <http://www.sciencedirect.com/science/article/B6WCR-458W209->  
635 [D/2/26c48622b7646e8f1977fafb85927147](https://doi.org/10.1016/0042-6989(76)90192-9)
- 636 Kolers, P. A., & Pomerantz, J. R. (1971). Figural change in apparent motion. *Journal of*  
637 *Experimental Psychology*. <https://doi.org/10.1037/h0030156>
- 638 Kolers, P. A., & von Grünau, M. (1976). Shape and color in apparent motion. *Vision Research*.  
639 [https://doi.org/10.1016/0042-6989\(76\)90192-9](https://doi.org/10.1016/0042-6989(76)90192-9)
- 640 Livingstone, M. S., & Conway, B. R. (2003). Substructure of direction-selective receptive fields  
641 in macaque V1. *Journal of Neurophysiology*, 89(5), 2743–2759.  
642 <https://doi.org/10.1152/jn.00822.2002>
- 643 Lorenceau, J., & Alais, D. (2001). Form constraints in motion binding. *Nature Neuroscience*,  
644 4(7), 745–751. <https://doi.org/10.1038/89543>
- 645 Lorenceau, J., & Cavanagh, P. (2020). Jumpy and Jerky: When Peripheral Vision Faces Reverse-  
646 Phi. *I-Perception*. <https://doi.org/10.1177/2041669520939107>
- 647 Lorenceau, J., & Shiffrar, M. (1992). The influence of terminators on motion integration across  
648 space. *Vision Research*, 32(2), 263–273. [https://doi.org/10.1016/0042-6989\(92\)90137-8](https://doi.org/10.1016/0042-6989(92)90137-8)
- 649 Lorenceau, J., Shiffrar, M., Wells, N., & Castet, E. (1993). Different motion sensitive units are  
650 involved in recovering the direction of moving lines. *Vision Research*, 33(9), 1207–1217.  
651 [https://doi.org/10.1016/0042-6989\(93\)90209-F](https://doi.org/10.1016/0042-6989(93)90209-F)
- 652 Lu, Z.-L., & Sperling, G. (2001). Three-systems theory of human visual motion perception:  
653 review and update.[see comment][erratum appears in J Opt Soc Am A Opt Image Sci Vis  
654 2002 Feb;19(2):413]. *Journal of the Optical Society of America, A, Optics, Image Science,*  
655 *& Vision*, 18(9), 2331–2370.
- 656 Marr, D., & Ullman, S. (1981). Directional selectivity and its use in early visual processing. *Proc.*  
657 *R. Soc. Lond. B Biol. Sci.*, 211(1183), 151–180.
- 658 McDermott, J., Weiss, Y., & Adelson, E. H. (2001). Beyond junctions: Nonlocal form constraints  
659 on motion interpretation. *Perception*, 30(8), 905–921. <https://doi.org/10.1068/p3219>
- 660 Morgan, M. J., & Ward, R. (1980). Conditions for motion flow in dynamic visual noise. *Vision*  
661 *Research*, 20(5), 431–435. [https://doi.org/10.1016/0042-6989\(80\)90033-4](https://doi.org/10.1016/0042-6989(80)90033-4)
- 662 Nakayama, K., & Silverman, G. H. (1988). The aperture problem-I. Perception of nonrigidity and  
663 motion direction in translating sinusoidal lines. *Vision Research*, 28(6), 739–746.  
664 [https://doi.org/10.1016/0042-6989\(88\)90052-1](https://doi.org/10.1016/0042-6989(88)90052-1)
- 665 Nowlan, S. J., & Sejnowski, T. J. (1995). A selection model for motion processing in area MT of  
666 primates. *The Journal of Neuroscience*, 15(2), 1195 LP – 1214. Retrieved from  
667 <http://www.jneurosci.org/content/15/2/1195.abstract>



- 668 Pack, C. C., Gartland, A. J., & Born, R. T. (2004). Integration of Contour and Terminator Signals  
669 in Visual Area MT of Alert Macaque. *The Journal of Neuroscience: The Official Journal of*  
670 *the Society for Neuroscience*, 24(13), 3268–3280.  
671 <https://doi.org/10.1523/JNEUROSCI.4387-03.2004>
- 672 Pack, C. C., Livingstone, M. S., Duffy, K. R., & Born, R. T. (2003). End-stopping and the  
673 aperture problem: Two-dimensional motion signals in macaque V1. *Neuron*, 39(4), 671–  
674 680. [https://doi.org/10.1016/S0896-6273\(03\)00439-2](https://doi.org/10.1016/S0896-6273(03)00439-2)
- 675 Perrinet, L. U., & Masson, G. S. (2012). Motion-based prediction is sufficient to solve the  
676 aperture problem. *Neural Computation*, 24(10), 2726–2750.  
677 [https://doi.org/10.1162/NECO\\_a\\_00332](https://doi.org/10.1162/NECO_a_00332)
- 678 Seiffert, A. E., & Cavanagh, P. (1998). Position displacement, not velocity, is the cue to motion  
679 detection of second-order stimuli. *Vision Research*, 38(22), 3569–3582.
- 680 Shiffrar, M. (2017). The Aperture Problem: Illusions Arising During the Integration and  
681 Segmentation of Motion Within and Across Apertures. In *The Oxford Compendium of*  
682 *Visual Illusions*. New York: Oxford University Press.  
683 <https://doi.org/10.1093/acprof:oso/9780199794607.003.0076>
- 684 Shimojo, S., Silverman, G. H., & Nakayama, K. (1989). Occlusion and the solution to the  
685 aperture problem for motion. *Vision Research*, 29(5), 619–626.  
686 [https://doi.org/10.1016/0042-6989\(89\)90047-3](https://doi.org/10.1016/0042-6989(89)90047-3)
- 687 Smith, A. T. (1994). Correspondence-based and energy-based detection of second-order motion  
688 in human vision. *Journal of the Optical Society of America, A, Optics, Image Science, &*  
689 *Vision*, 11(7), 1940–1948.
- 690 Solomon, J. A., & Sperling, G. (1994). Full-wave and half-wave rectification in second-order  
691 motion perception. *Vision Research*, 34(17), 2239–2257. Retrieved from  
692 <http://www.ncbi.nlm.nih.gov/pubmed/7941419>
- 693 Ullman, S. (1979). *The Interpretation of Visual Motion. The Interpretation of Visual Motion.*  
694 [https://doi.org/10.1016/0013-4694\(80\)90055-3](https://doi.org/10.1016/0013-4694(80)90055-3)
- 695 Wallach, H. (1935). Über visuell wahrgenommene Bewegungsrichtung. [On the visually  
696 perceived direction of movement.]. *Psychologische Forschung*, 20, 325–380.  
697 <https://doi.org/10.1007/BF02409790>
- 698 Weiss, Y., Simoncelli, E. P., & Adelson, E. H. (2002). Motion illusions as optimal percepts.  
699 *Nature Neuroscience*, 5(6), 598–604. <https://doi.org/10.1038/nn858>
- 700 Wuerger, S., Shapley, R., & Rubin, N. (1996). “On the visually perceived direction of motion” by  
701 Hans Wallach: 60 years later. *Perception*, 25(11), 1317–1368.  
702 <https://doi.org/10.1068/p251317>

703

## **Supplementary Information**

### **Video 1. Aperture problem**

Through an aperture, the motion direction of a line is ambiguous and is usually perceived in a direction orthogonal to its orientation (here diagonally, first sequence). Without the aperture, the line is usually perceived as drifting in its veridical direction (here horizontally, second sequence). Early direction-selective neurons are sensitive to motion within a small receptive field equivalent to a small aperture. Assuming that the perceived motion of the line is based on early direction-selective neurons sensitive to motion orthogonal to the line, the aperture problem consist in determining how the visual system recover the veridical motion direction.

### **Video 2. Four rotating spots**

The motion was either forward phi-motion (first two sequences) or reverse-phi motion (last two sequences), and either continuous (first and third sequences) or stroboscopic (second and fourth sequences). Observers had to report the rotating direction (here, always clockwise).

### **Video 3. Thirty-two rotating spots**

Legend as in Video 2, but for 32 spots.

### **Video 4. Rotating diamond with visible occluders**

The motion was either continuous (first sequence) or stroboscopic (second sequence). Observers had to report the rotating direction of the global shape position (here, always clockwise).

### **Video 5. Rotating diamond with invisible occluders**

Legend as in Video 4, but with occluders the same brightness as the background.

### **Video 6. Rotating chevron with visible occluders**

Legend as in Video 4, but for a chevron shape.

### **Video 7. Rotating chevron with invisible occluders**

Legend as in Video 4, but for a chevron shape with occluders the same brightness as the background.

### **Video 8. Drifting tilted line**

Oblique line drifting to the right. The motion is continuous and stroboscopic (first and second sequence, respectively).