Perceptual compression of space through position integration

Barrie W. Roulston*, Matt W. Self† and Semir Zeki

Department of Anatomy and Developmental Biology, University College London, London WC1E 6BT, UK

The mechanism of positional localization has recently been debated due to interest in the flash-lag effect, which occurs when a briefly flashed stationary stimulus is perceived to lag behind a spatially aligned moving stimulus. Here we report positional localization observed at motion offsets as well as at onsets. In the ‘flash-lead’ effect, a moving object is perceived to be behind a spatially concurrent stationary flash before the two disappear. With ‘reverse-repmo’, subjects mis-localize the final position of a moving bar in the direction opposite to the trajectory of motion. Finally, we demonstrate that simultaneous onset and offset effects lead to a perceived compression of visual space. By characterizing illusory effects observed at motion offsets as well as at onsets, we provide evidence that the perceived position of a moving object is the result of an averaging process over a short time period, weighted towards the most recent positions. Our account explains a variety of motion illusions, including the compression of moving shapes when viewed through apertures.

Keywords: positional localization; aperture viewing; flash-lag; motion

1. INTRODUCTION

Positional localization of a moving object presents an apparent paradox, namely that when a particular location is frozen at a given moment in time, the attribute of motion cannot exist at all (Salmon 1970; for relevance to psychophysics see Morgan 2003). This issue has been the subject of more recent debate, mainly due to interest in the ‘flash-lag effect’, which occurs when a flashed stationary object is perceived to lag behind a spatially aligned moving object (Mackay 1958; Nijhawan 1994). Despite much research over the last decade, there is still no established consensus on the correct explanation for this phenomenon; motion extrapolation (Nijhawan 1994), attention (Baldo & Klein 1995), differential latency (Whitney & Murakami 1998), positional averaging (Krekelberg & Lappe 2000) and post-diction (Eagleman & Sejnowski 2000a) are all potential candidates.

Previous studies of the flash-lag effect have mainly focused on the relative spatial localization of moving objects at motion onsets or during continuous motion. Similarly, in the Fröhlich effect, subjects mis-localize the initial position of a moving object in the direction of motion (Fröhlich 1929). Here we investigate relative position localization at both motion offsets and onsets. Using the conventional ‘flash-lag’ paradigm and the relative positioning of a continuously moving bar, we found that subjects mis-localize the final position of a moving object prior to its disappearance in the direction opposite to its trajectory. This is consistent with the positional averaging hypothesis (Morgan 1975; Krekelberg & Lappe 2000): the veridical position is averaged over a time period preceding the disappearance, thus causing it to be seen to disappear before its veridical final position. Furthermore, this offset effect is smaller than onset effects observed with either the Fröhlich effect or the flash-lag effect, suggesting that this process is weighted to the most recently sampled positions. Finally, we demonstrate that the influence of both motion onset and offset effects leads to the perceived asymmetrical compression of a moving dot array. We are thus also able to account for the compression of moving shapes when viewed through apertures, a seemingly unrelated illusion that has been unresolved since originally described by Helmholtz (1862/1867) and Zöllner (1862). These results provide strong evidence for the proposal that the perceived position of a moving object is the result of the weighted integration of signals relating to position over a short time window.

2. MATERIAL AND METHODS

All experiments were done in a darkened experimental room using a Sony high-resolution CRT monitor with a refresh rate of 100 Hz and a viewing distance of 30 cm. Stimuli were constructed using a PC running MatLab (Mathworks Inc.) and the COGENT GRAPHICS package (http://www.vislab.ucl.ac.uk/Cogent/index.html). Responses were recorded via key press on a computer keyboard. Points of subjective simultaneity were derived by using the method of constant stimuli in a 2AFC task and fitting a Weibull function (2-parameter model: shape and 50% point) to the data. Each subject had normal or corrected-to-normal vision. Informed written consent was obtained before experiments. Three naive subjects and one of the authors (three male) participated in all experiments. During each trial subjects fixated a central square.
(a) **Experiment 1: flash-lead effect**

We modified the flash-lag paradigm to determine the perceived position of a moving stimulus at motion offset. We refer to this as the ‘flash-lead’ effect. During flash-lead trials, a moving hollow circle (radius: 1.67°) appeared after a random interval of under one second from the beginning of each trial and moved in a circular trajectory at a speed of 200° s⁻¹ around the fixation point. At a random position, the moving circle was briefly filled with a stationary achromatic flash of one of three luminance values (108, 34, 12.5 cd m⁻²). In the following frame, both the moving circle and the flash disappeared (figure 1a). Note that this differs from most previous flash-terminated cycles (e.g. Eagleman & Sejnowski 2000a) as the moving target actually disappears at the end of its trajectory instead of stopping. Subjects had to indicate whether the flash appeared above or below the moving circle. These trials were randomly mixed with trials measuring the conventional flash-lag effect, using the identical set-up described, except that the moving circle continued to move after the flash occurred. By varying the position of the flash relative to the moving circle, the psychometric function of the position of the flash was calculated.

(b) **Experiment 2: reverse-repmo**

We modified the classical Fröhlich experiment, in which a bar moves in a given direction and subjects have to specify its position when it appears, in two ways: instead of having one bar, we introduced a second bar and asked subjects to determine the relative position of the two bars, at both onset and offset. We refer to the latter as the ‘reverse-representation moment’ or reverse-repmo. More specifically, two bars (5.7×0.67°), both presented either to the left or right side of fixation (randomized trial by trial), moved towards each other vertically (speeds ranging from 6 to 49° s⁻¹) from the bottom and the top of the screen, respectively, towards the mid-point where they were horizontally spatially aligned for one frame before disappearing. The bars did not overlap, although there was no gap between them when side-by-side (figure 1b). Subjects then indicated whether the bars passed each other or not prior to disappearance. We calculated the psychometric function of the relative positions by varying the position at which the two bars disappeared relative to each other. In a separate experiment we measured the classical Fröhlich effect by using an identical set-up. In this case the bars were presented aligned at the mid-point and moved vertically in opposite directions. Subjects had to indicate whether the bars appeared above or below relative to each other. The psychometric function was calculated by varying the position at which the two bars appeared relative to each other.

(c) **Experiment 3: compression of space**

The above set of experiments led us to ask whether these mislocalizations may sum and account for compression of moving shapes when viewed through apertures (reviewed by Rock 1981). In the first set of experiments, we measured the amount of compression that subjects perceived a moving diamond shaped array (test array) composed of four dots to undergo. Each dot was 0.95° in radius. The vertically aligned dots were 2° apart, whereas the horizontally aligned dots were 2.9° apart (see figure 1c). In the mid-condition, the test array briefly appeared as if moving rightwards from under an invisible occluder, being completely exposed before moving underneath a second invisible occluder on the right, at a constant speed of 40.6° s⁻¹. Note that the dots therefore just disappeared without being covered by a visible occluder. This set-up was reminiscent of the aperture-viewing experiments originally performed by Helmholtz (1962/1867) and Zöllner (1862), where an object moves through a narrow slit, only exposing a small portion at a time. Observers report seeing the stimulus integrated as whole, although compressed in the direction of its motion. In our experiment the aperture was slightly wider than the moving stimulus (3.4° width), thereby briefly exposing the whole array at the same time. If the theory of positional averaging is correct, we predicted that the array should still appear compressed in size, as reverse-repmo at the leading dot should occur simultaneously with the Fröhlich effect at the trailing dot. Since the latter effect is larger, our second prediction was that the compression effect would be asymmetrical, with the trailing side of the array appearing most compressed.

The test array randomly appeared either above or below fixation on each trial. In addition to this was the reference array, which matched the exact temporal profile of the dots in the moving array, but did not move and appeared on the opposite side of fixation to it. We varied the width of the test array by adjusting the distance of the two horizontally aligned dots relative to the centre, to derive the psychometric function of the test array’s compression. Subjects then indicated in a forced choice if the moving test array was narrower or wider than the reference array. Three additional conditions were randomly interleaved with the main one described above; differential offsets, where the whole array appeared simultaneously but had differential offsets by disappearing underneath the invisible occluder; differential onsets, where the whole array had differential onsets by appearing from underneath the invisible occluder, but disappeared simultaneously; and simultaneous onset/offsets when the whole array appeared and disappeared simultaneously. This last condition acted as a control since motion smear (Burr 1980).
could cause subjects to bias responses, overshadowing the effect of differential on-/offsets.

The second set of experiments measured the asymmetry of the arrays, and took the same structure as those described above. There was no reference array and subjects were asked to indicate what side (left or right) appeared most compressed. To derive the psychometric curve we multiplied the $x$ coordinates using

$$x' = (1 + Kx^p).$$

where $K$ is a constant with 13 values from $-12$ to $12$. This manipulation distorted the layout of the dot array so that it appeared compressed on either the left or right sides. We also quantified the asymmetry of the compression effect using a set-up similar to those traditionally used in aperture viewing experiments (see the review by Rock 1981). Two achromatic rectangles (both 8.5 by 6.7°) were displayed on either side of a central fixation spot leaving a slit of 1.3°. The diamond-shaped dot array moved across this slit from underneath the occluding rectangles either above or below the fixation spot at a speed of 40.6° s$^{-1}$. Subjects had to indicate which side appeared most compressed, as before.

### 3. RESULTS

#### (a) Experiment 1: flash-lead effect

Subjects perceived the moving object to lag behind the flash before both stimuli disappeared (figure 1a). This is consistent with the positional averaging hypothesis: the veridical position is averaged over a time period of 100 ms or so, as previously suggested by integration times for detecting moving objects (Burr 1981) preceding the disappearance, causing it to be seen to disappear before its veridical final position. A repeated measures ANOVA revealed a significant effect of flash luminance ($F_{2,8} = 12.46, p < 0.01$), occurring with the maximum amplitude (mean: $-1.33°$) if the luminance is high, consistent with previous psychophysical evidence suggesting that higher luminance reduces perceptual latencies (Roufs 1963). The earlier in time the flash is perceived relative to the moving cursor because of higher luminance, the earlier in the course of the trajectory of the moving cursor is the flash co-localized with it. But the lead is significant even at low flash luminance. We can infer the existence of such a lead effect from previous data (Müsseler et al. 2002), although it was attributed to response bias and not interpreted to be a genuine perceptual effect. However, the effect of flash luminance makes that interpretation improbable.

#### (b) Experiment 2: reverse-repmo

All subjects mis-localized the final relative position of the two bars in the direction opposite to their motion trajectories prior to disappearance (figures 1b and 2b), with the magnitude increasing with stimulus speed over the range used. The same effect was observed with a moving stimulus when one of the bars was stationary and aligned with the final position of the moving bar (mean: 0.31°, one sample $t$-test across subjects at 95% significance: $p < 0.0001$). We quantified the Fröhlich effect using this same stimulus set-up and found it to be substantially larger than reverse-repmo (figure 2a).

#### (c) Experiment 3: compression of space

The results given above show that the trajectory of a moving stimulus appears to start later and end earlier than it does veridically, causing the perceived length of its trajectory to be compressed. Can both of these effects be observed simultaneously? Our results show significant compression effects over the control condition in both the differential on-/offsets conditions and the simultaneous offsets condition (figure 3a). The simultaneous onsets condition gave a small insignificant compression effect. This is consistent with the larger mis-localizations observed at motion onsets than at motion offsets (figure 2).

However, the compression effects observed in either the simultaneous onset or simultaneous offset conditions were not large enough to be able to account for the compression effect observed in the main condition alone, although the sum of them approximates it in each subject, suggesting that this is due to a contribution from both onset and offset effects. To quantify the asymmetry of compression, subjects performed a forced choice test to report which side of the array appeared to be the most compressed, i.e. on which side the middle dots appeared closest to the outer dot. Consistent with our predictions, the main condition showed a slight asymmetry in the compression.
to the trailing side compared with the control, which we presume is due to the larger influence of onset effects (figure 3b). The identical result (mean: 6.1°, one sample t-test at 95% significance across subjects: $p < 0.0001$) was also observed with a stimulus set-up similar to those traditionally used in aperture viewing experiments and, to our knowledge, has not previously been reported. This was also apparent in the second condition when all the dots disappeared simultaneously and only differential onsets had an influence. In the third condition with only differential offsets, a significant asymmetry in the array was found, although this time the leading side was judged more compressed compared to the control condition, as expected by the small backwards mis-localization observed at motion offsets of the leading dots.

4. DISCUSSION

The results that we report provide strong evidence in favour of the averaging hypothesis for localizing the position of a moving object at a given moment in time. The new experiments that we have introduced have provided results which show that (i) when a moving object disappears, it is perceived to lag behind a spatially concurrent flash; (ii) two moving objects aligned at their final positions with respect to each other are perceived to be mis-aligned in a direction opposite to their direction of motion; (iii) a moving shape with both leading and trailing edges is perceived to be asymmetrically compressed. Together with the variants that we have given in §3, the most plausible interpretation of these results is that the brain averages the positions occupied during the last 100 ms or so. Note, however, that the magnitude of reverse-repmo is much smaller than the Fröhlich effect. We suggest that although positional averaging is taking place, it is weighted towards the most recently sampled positions.

The position averaging theory may also explain the flash-lag effect if we assume that during a period of time after the flash, corresponding to the integration window, the average position of the moving object is calculated, resulting in the moving object being seen after the flash. This process may also take into account positions of the moving object that occurred prior to the flash, in line with data showing that the flash-initiated cycle can produce a larger flash-lag effect than during continuous motion (Ogmen et al. 2004). Similarly, the Fröhlich effect can be explained if, after a certain time period from the appearance of the moving object, its average position is calculated, causing the object to be seen ahead of the veridical initial position. This enables us to account for all four of these effects within a single overall framework.

Our results cannot be accounted for by alternative explanations of the flash-lag effect, because offset effects for moving objects place tight constraints on possible explanations of positional localization. The motion extrapolation (Nijhawan 1994), attention (Baldo & Klein 1995), latency difference (Whitney & Murakami 1998; Purushothaman et al. 1998) and ‘post-diction’ hypotheses (Eagleman & Sejnowski 2000a), at least in their simplest forms, would all predict that the flash should be perceived to lag, or at the very least be concurrent, with the moving cursor. The flash-lag effect described here shows that these predictions do not hold true. Furthermore, the revised version of post-diction (Eagleman & Sejnowski 2000b) would predict that the relationship between the size of the lead effect and the luminance values of the flash should be reversed, as the more salient the flash, the more the positions of the moving cursor before it should be devalued, thereby producing a reduced or annulled flash-lag effect. A recent version of the differential latency hypothesis (Ogmen et al. 2004) is able to accommodate the flash-lag effect with a ramped response function for the position computation of moving stimuli. However, our second experiment shows that this mis-localization is not specific to moving versus stationary objects. Their model would predict that reverse-repmo should not exist: the two bars would eventually be seen to disappear at their veridical positions and be spatially aligned, given the assumption that they both have equal perceptual latencies.
It is important to distinguish the present results from related previous findings. The backwards mis-localization at motion offsets is surprising, given the previous ‘representational momentum’ literature which shows that subjects mis-locate manually the position of a moving object to a point beyond the final position of the moving object (Freyd & Finke 1984). A series of experiments by Kerzel and colleagues (Kerzel 2003; reviewed in Kerzel 2005) has established that under steady fixation representational momentum is due to an anticipatory overshoot in attention in the direction of the moving target. We find that an attentional overshoot, or even undershoot, does not occur with smooth motion foveally (see electronic supplementary material). Therefore such an attentional effect is not responsible for the effects described in our main experiments. However, we did find such an overshoot to occur with smooth moving peripheral targets. This may account for the apparent discrepancy between the present findings and recent results that show a forward mis-localization of the moving object in a flash-terminated condition where both the moving and flashed object disappear in the frame after the flash (Kanai et al. 2004). Their effect was contingent on the moving stimulus presented peripherally, which does cause an attentional overshoot (see electronic supplementary material), and could therefore be due to the same attentional overshoot which underlies representational momentum. The lack of lead effect for foveally presented stimuli by Kanai et al. (2004) may be explained by the lack of sensitivity in using thicker stimuli: narrower stimuli, like the moving cursor used in the present experiment, allow greater vernier sensitivity (cf. Levi et al. 2000) and therefore greater sensitivity in detecting this subtle lead effect. Our second supplementary experiment confirms this hypothesis. Furthermore, we believe the present flash-lead effect to be distinct from the motion extrapolation effect described by Fu et al. (2001). First, as acknowledged by the authors, it is clear that their effect depends on the blurring of targets, unlike the sharp edges used in our experiments, which they account for by the properties of biphasic temporal responses of visual neurones. Second, their effect peaked at low speeds; this is unlike the present effect, which increases with velocity over a comparable range of values. Finally, their sharp-edged targets remained at the final location for 100 ms. According to our explanation, this would result in no mis-localization in either direction, which is what they found; the crucial difference is that the present stimuli disappeared in the frame after reaching their final position.

We further extended these observations with a novel dot array stimulus to investigate the simultaneous effects of motion onsets and offsets. Consistent with our predictions, we found that the array appears to be asymmetrically compressed. Moreover, exclusively differential motion onsets produced larger compression effects than offsets, and did so on the trailing side of the array, whereas exclusively differential motion offsets had a smaller effect and compression occurred on the leading side of the array. The larger effects of motion onsets than offsets may explain the asymmetry in the differential on-offsets condition, and the asymmetry we observed under normal aperture viewing conditions. Therefore the weighted averaging of position is also able to account for how the sequence of parts of an object viewed through an aperture is integrated into a whole and why the image appears compressed (Rock 1981). Each position of the component of the object is averaged over 100 ms or so, thereby allowing the image to be integrated as a whole. This same integration causes the mis-localization at motion onsets and offsets effect to sum, as demonstrated by the present experiments, which leads to the overall compression of the shape. Moreover, we are able to confirm our novel prediction, which directly leads from the weighted averaging hypothesis, that this compression effect is asymmetrical. Our explanation also accounts for the increase in compression previously observed (Morgan et al. 1982) with object speed and the narrowing of aperture size. We suggest that the averaging process corresponds to the hypothetical ‘post-retinal storage mechanism’ proposed to account for the effects of aperture viewing (Parks 1965), in contrast to a ‘retinal painting effect’ caused by eye-movements, since the compression effects observed here occurred under steady fixation. The mechanism of positional averaging is unknown, although one possibility is that it is due to the perceived deblurring of moving objects (Burr 1980; Burr & Morgan 1997) over the relatively long integration time of human vision (Barlow 1958; Burr 1981).

This work is funded by the Wellcome Trust. B. W. R. is supported by a UCL Graduate School Scholarship. We thank M. Morgan for encouragement.

REFERENCES


