The extrapolation of visual motion trajectories

N. Jeremy Hill
Department of Experiment Psychology,
University of Oxford, Oxford OX1 3UD, UK.
jeremy.hill@psy.ox.ac.uk

Abstract

Owing to the delays incurred by neural transmission and processing, there is an inevitable delay between the occurrence of an event and our ability to perceive or respond to it. Thus, if we attempt to catch (or avoid) a moving object, we must extrapolate its trajectory in order to produce an appropriate response: if we were simply to attempt to respond according to our representation of its instantaneous position, the object would have moved by the time that representation could be computed.

This presentation will concern mechanisms by which the brain might extrapolate motion trajectories. The project examines the spatial relations between the perceived locations of moving, flashed and static stimuli. From developments of the paradigm described by Nijhawan (1994, Nature 370, 256-7), data from three observers suggest a low-level "motion capture" mechanism which may be useful for motion extrapolation. In addition., there appears to be a temporal pooling mechanism which affects the subjective appearance of the stimulus array: although, when measured separately, the apparent locations of the flashed and moving stimuli are affected roughly equally by motion capture, there is still a striking subjective disparity between the two.

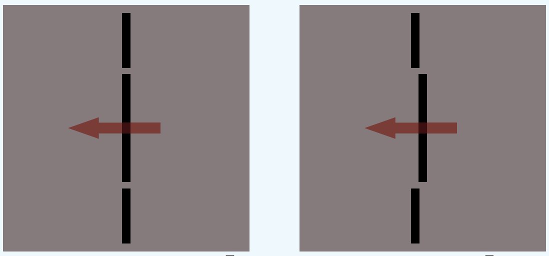
Comparison of uni-directional and bi-directional stimulus arrays suggests that neither mechanism depends crucially upon eye movements.

Introduction

Our perception of events and objects in the world is inevitably constrained by the fact that neural processing takes time. From the onset of visual stimulation at time t, for example, it takes between 30 and 100 msec for activity to reach primary visual cortex (at least in the macaque - Vogels & Orban 1994). This delay alone is enough to allow an object moving at 10 mph to have shifted between 15 and 45 cm from the position it occupied at time t. These figures, of course, do not even allow for the subsequent processing involved in planning and coordinating, let alone executing, motor actions which are coordinated with the motion of objects. In order to achieve such an easy task as catching a ball moving at 10 mph, the human system cannot rely on representations of that object's instantaneous position at any given time t, but that it must rather extrapolate the motion trajectory to the required time after t, based on the motion information available so far. Cavanagh (1997) aptly described this process as "predicting the present."

Processing delays have been with us throughout evolutionary history, so it would not be surprising to find that the brain had developed a range of different solutions to the problem. Such solutions might operate at a number of different "levels," across a spectrum from low-level implementations (for example, motion detectors might locally bias the place-coding of location in retinotopically organized early visual areas) to high-level (such as cognitive mechanisms which operate on object-level representations, and which might apply explicit knowledge about the physics of motion). The current research concentrates on possible mechanisms at the lower end of the spectrum, with specific reference to a phenomenon presented by Nijhawan (1994) as evidence of the brain's motion extrapolation capabilities.

Experiment I - After Nijhawan (1994)



Perceived　　　Actual

- central bar moves repeatedly in the same direction;
- outer markers flash once per motion cycle;
- observers adjust marker position to align the markers with the bar.

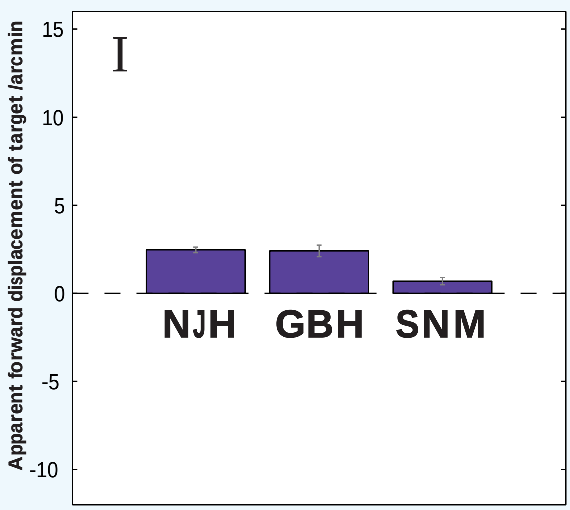
The first experiment aimed to replicate the findings of Nijhawan (1994, 1997) that a briefly flashed stimulus is seen to occur behind a smoothly moving stimulus, when their location in the relevant dimension is in fact the same. The phenomenon can be measured by cancellation. Nijhawan originally demonstrated this using a bright rotating slit, but the aim here is to replicate the effect using a stimulus setup more suitable for the intended manipulations of experiments II & III: horizontal translation of a black bar.

The bar (2.7 x 0.2 deg) moved over a 7.2 deg path at 10.7 deg/sec. During one trial it moved repeatedly in one direction only, either left to right or right to left. Two markers (each 1.4 x 0.2 deg) flashed on for one frame (13.3 msec) at some point during each cycle of the bar's motion. The observers' task on each trial was to adjust the position at which the markers flashed, from their randomly chosen start position, so that they appeared collinear with the bar.

The results for three observers are shown in figure I. Each blue bar shows, for one of the observers, the mean offset between marker position and bar position at the time of the flash. A positive offset means the markers were ahead of the bar, in the direction of motion. Error bars show ± one standard error of the mean, based on 25 trials. ▶

The findings replicate those of Nijhawan, in that the error for all observers was consistently positive. It should be noted that the magnitude of the effect was very small: the offsets reported here correspond to, at most, the distance moved by the bar in 6 msec, as compared with 82 msec in Nijhawan's rotation study. This may be a function of the type of motion: additional experiments with these stimuli have demonstrated that the figure for rotational motion is much closer to the order of magnitude of Nijhawan's effect (around 40 msec, in fact - results not shown).

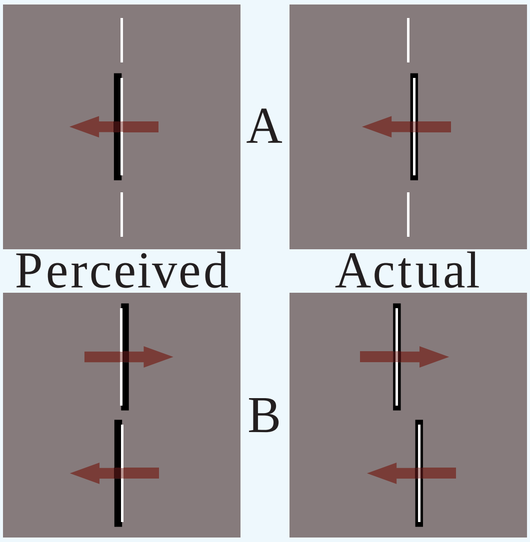
Nijhawan's interpretation of his result was that the motion of the bar, being predictable, allowed the observer to extrapolate in order to calculate the bar's position, whereas the flash, being transient, suffers the full effect of neural processing delay.



jeremy.hill@psy.ox.ac.uk

N. Jeremy Hill
Department of Experiment Psychology,
University of Oxford, Oxford OX1 3UD, UK.

Experiment II - Motion Capture



- central black bar moves;
- outer white markers are static;
- central white stripe flashes (always in centre of the black bar);
- observers can shift bar & flash forwards & back, in order to line up the white stripe with the markers.

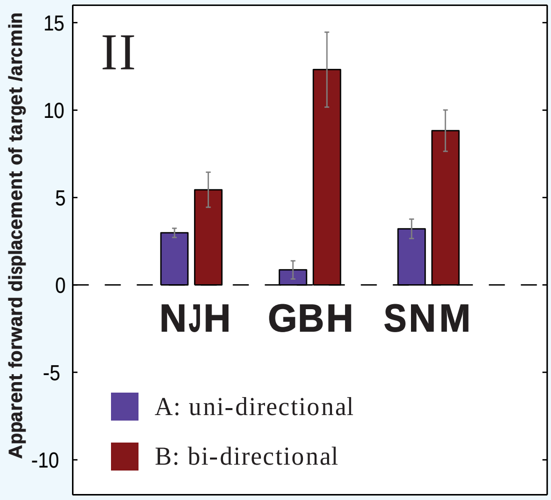
- black bars move in opposite directions;
- white stripes flash simultaneously (always in centre of the black bars);
- observers shift upper bar & flash forwards & back.

Condition A: uni-directional motion

This experiment was designed to test a "low-level" hypothesis about the effect observed in experiment I. The hypothesis is that motion signals cause a bias in the place-coding of stimulus location, regardless of object identity.

The moving stimulus was the same as the black bar in experiment I. This time, however, the transient stimulus, flashed on for 13.3 msec, was a white stripe (2.4 x 0.05 deg) down the middle of the moving bar. The task was once again a Vernier adjustment: observers were required to align the white stripe with two static white markers (1.1 x 0.05 deg). In effect, the moving bar was a distractor, its presence irrelevant to the performance of the task.

Results are shown in the blue bars of figure II: again, the mean and standard errors from 25 trials are shown for each of three observers, and the positive sign of all the results indicates that, when the observers judged the target stripe to be aligned with the markers, the markers were actually ahead of the stripe, in the direction of motion. ▶



The results indicate that the white stripe was seen ahead of its actual location, in the direction of the bar's motion, despite the fact that the stripe itself was not moving: it had effectively been "captured" by the motion of the bar. The effect is reminiscent of the results of DeValois and DeValois (1991) who found that motion of the carrier frequency of a Gabor path shifted the apparent location of the envelope, in a Vernier acuity task.

It seems likely that one of the mechanisms which might contribute to the effect observed in experiment I, and by Nijhawan (1994), is a form of "motion capture": a distortion of apparent location in the direction of a nearby motion signal, regardless of object identity. The motion capture effect does not tell the whole story, however. All observers reported that the white stripe appeared to lag behind the black bar, despite the fact that its physical location was always down the centre of the bar. This echoes the effect in experiment I, in that a transient stimulus is seen to lag behind a smoothly moving one, and the stimulus configuration is very similar to that of Nijhawan's (1997) colour study, in which a lag is also reported.

Note that for observer NJH the shift in the stripe's apparent location is of very similar magnitude to the shift in the apparent location of the bar, as measured in experiment I. For SNM, the stripe shift is even larger than the bar shift. This presents an intriguing paradox: when measured separately, the bar and stripe are found to be perceived at the same location (NJH), or the white stripe is shifted further forward than the bar (SNM), even though both observers' perception of the spatial relationship between the two is that the stripe lags behind. It could be that the perceived *relative* locations of the two stimuli are computed by different mechanisms than those involved in the Vernier task: it is possible that we are seeing the effects of a mechanism which can segregate the two stimuli according to their temporal characteristics, as well as one which cannot. Alternatively, the perceived position of the stimuli relative to the outer markers may depend on which stimulus is attended, which depends on the task, although some recent studies would suggest that this is unlikely (Khurana & Nijhawan 1995, Khurana, Cavanagh & Nijhawan 1996).

Condition B: bi-directional motion

The second part of this experiment aimed to test whether the "motion capture" effect was a result of local stimulus processing, or whether it was a result of a shift in a global coordinate system - for example, an eye movement. Shimojo & Nijhawan (1998) have found that a voluntary tracking eye movement can cause a shift in the apparent location of a flashed stimulus, and while observers in the current study did not deliberately track the bar, it is possible that involuntary eye movements or "oculomotor readiness" (Klein & Pontefract 1994) might cause a shift in a global coordinate frame that might account for the results.

Whether the result of an eye movement, preparation for eye movement, or other shift of spatial attention, a change in global coordinates should affect all stimuli in the same direction. Therefore the hypothesis was tested by replacing the static markers with a second bar, moving in the opposite direction to the first. White stripes flashed simultaneously down the middle of both bars, and the task was to align the two white stripes. The distance between the stripes was measured for each trial, keeping the convention of the previous experiments in that positive numbers indicate that the bars had not yet reached each other, and negative numbers indicate that they had passed. If the effect observed in condition A, above, is the result of a global coordinate shift, we should see a reduction in the magnitude of the effect in this bidirectional condition, because one bar will be shifted forward, and the other will be shifted backwards. If, on the other hand, the effect was the result of purely local processing mechanisms, we should obtain double the magnitude, because both stripes will be shifted forwards.

The results (shown by the red bars) clearly show an increase in the magnitude of the effect for all observers. Therefore it seems that eye movements and spatial attention shifts do not play a crucial role in the motion capture effect.

Acknowledgments

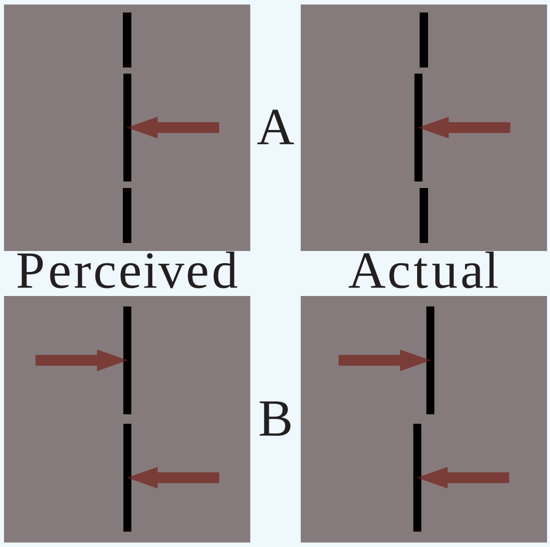
This research was supported by the Christopher Welch Trust and St. Hugh's College, University of Oxford. Thanks to Bruce Henning and David Gilden for their advice.

The extrapolation of visual motion trajectories

jeremy.hill@psy.ox.ac.uk

N. Jeremy Hill
Department of Experiment Psychology,
University of Oxford, Oxford OX1 3UD, UK.

Experiment III - Vanishing Point



- central bar moves, then vanishes;
- outer markers are static, then vanish at the same time as the bar;
- observers adjust the vanishing position of the bar.

- bars move, then vanish simultaneously;
- observers adjust the vanishing position of the upper bar.

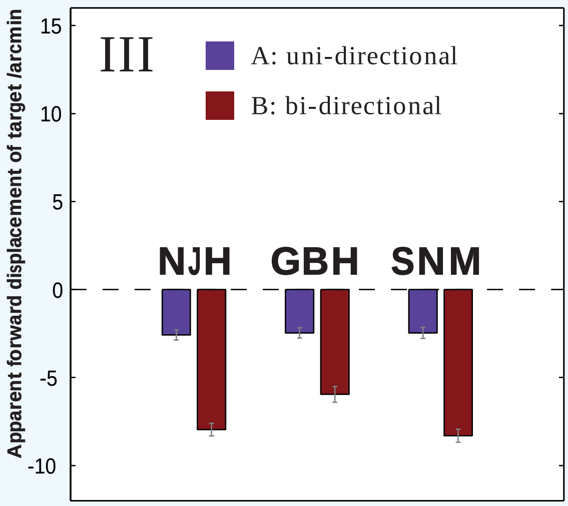
The last two experimental conditions were designed to test a prediction from Nijhawan's original hypothesis: if the position of a moving stimulus is extrapolated forwards, then when that stimulus suddenly disappears, the point at which it subjectively vanished should be shifted forward from the true vanishing point.
Such an effect has been shown when observers' *memory* for the vanishing point is probed, and has been interpreted as evidence for motion extrapolation (Hubbard 1995). Kolers & Touchstone (1965) have also investigated overshoots in phi motion, using a similar paradigm to the following, and with mixed results.

Condition A: uni-directional motion

This time the bar moved repeatedly over the same trajectory, passing between two constant static black markers. At a randomly chosen point near the middle of the cycle, bar and markers vanished simultaneously. Observers were instructed to line up the "last seen" position of the bar with the markers. Specifically, they were told to "adjust the vanishing point so that you can see the bar aligned with the markers, but you never see it go past."

The results, denoted by the blue bars, show exactly the reverse of the effect predicted by a theory of motion extrapolation. As before, the mean and standard error from 25 trials is plotted for each observer, and the convention is that positive values indicate that the markers were lined up ahead of the bar. This time the negative displacements indicate that all three observers allowed the bar to go past the markers. ▶

One possible explanation for this is that the response to any one real frame of the motion sequence is too weak (perhaps due to "motion de-blurring" - Burr 1980) to provide a sensation of where the bar was "last seen," and that this subjective point is in fact determined by a weighted average across the last few frames. Although Nijhawan (1997) argues against temporal pooling as an explanation for the original effect (see experiment I) it is possible that, if the effect is pronounced enough to show up in experiment III, it also contributes to the original effect.



Condition B: bi-directional motion

Again, a control experiment was performed to investigate the effects of eye movements or other global coordinate shifts in the "vanishing point" paradigm. Two bars moved towards each other, and disappeared roughly half-way through each cycle. The observer adjusted the bars so that they were collinear when "last seen."

As before, the results of condition B are in the same direction as, and of greater magnitude than, the results of condition A, suggesting that eye movements do not play a major role in the effect.

Conclusion

There are a number of mechanisms at work in determining the apparent location of a moving object. One of these appears to be a "motion capture" mechanism, by which the perceived location of stimuli is distorted in the direction of a local motion signal, regardless of whether the relevant stimulus is the thing that is moving. Such a mechanism may have a use in coordinating our behaviour with that of moving objects. Another mechanism may be temporal pooling, which, although useless for motion extrapolation, could account for the subjective disparity between a moving stimulus and a flashed stimulus at the same location. The subjective disparity occurs despite the fact that the apparent location of each stimulus, measured separately, is the same. Neither mechanism appears to be dependent on eye movements or global shifts of spatial attention, although there is undoubtedly a role for such mechanisms in motion extrapolation.

References

Burr, D. (1980): Motion smear. Nature 284, 164-165.

Cavanagh, P. (1997) Predicting the present. Nature 386, 19-21.

De Valois, R. L. & De Valois, K. K. (1980): Vernier acuity with stationary moving Gabors.Vision Research 31, 1619-1626.

Hubbard, T. L. (1995): Cognitive representation of motion: evidence for friction and gravity analogues. Journal of Experimental Psychology: Learning, Memory and Cognition 21, 541-254.

Khurana, B. & Nijhawan, R. (1995): Reply to Baldo & Klein. Nature 378, 566.

Khurana, B., Cavanagh, P. & Nijhawan, R. (1996): Are moving objects "corrected" or flashed objects attentionally delayed? Investigative Ophthalmology and Visual Science 37(3) #2421.

Klein, R. M. & Pontefract, A. (1994): Does oculomotor readiness mediate cognitive control of visual attention? Revisited! In Umilta, C. & Moscovitch, M. (Eds): Attention and performance 15, 333-350. Cambridge, MA: MIT Press.

Kolers, P. A. & Touchstone, G. E. (1965): Variations of perceived distance with apparent motion. Quarterly Progress Report No. 79, Research Laboratory of Electronics, Massachusetts Institute of Technology. Reported in P. A. Kolers (1972): Aspects of Motion Perception, pp 142-143. Oxford: Pergamon.

Nijhawan, R. (1997) Visual decomposition of colour through motion extrapolation. Nature 386, 66-69.

Nijhawan, R. (1994): Motion extrapolation in catching. Nature 370, 256-257.

Shimojo, S. & Nijhawan, R. (1998): Spatial localization of double flashes during smooth pursuit eye movement I: perception oriented response. Perception 27(supp), 146.

Vogels, R. & Orban, G. A. (1994): Activity of inferior temporal neurons during orientation discrimination with successively presented gratings. Journal of Neurophysiology 71(4), 1428-1451.