Available online at www.sciencedirect.com

ScienceDirect

Vision Research 47 (2007) 937–945

Vision Research

www.elsevier.com/locate/visres

Dynamical evolution of motion perception

Ryota Kanai [a,d,*,1], Bhavin R. Sheth [b,c,1], Shinsuke Shimojo [d,e]

[a] *Universiteit Utrecht, Helmholtz Research Institute, Psychonomics Division, Heidelberglaan 2, NL 3584 CS Utrecht, The Netherlands*
[b] *Department of Electrical and Computer Engineering, University of Houston, Houston, TX 77204-4005, USA*
[c] *Center for NeuroEngineering and Cognitive Sciences, University of Houston, Houston, TX 77204-4005, USA*
[d] *Division of Biology and Computation and Neural Systems Program, California Institute of Technology, Pasadena, CA 91125, USA*
[e] *Human and Information Science Laboratory, NTT Communication Science Laboratories, Atsugi, Kanagawa 243-0198, Japan*

Received 2 May 2006; received in revised form 12 December 2006

Abstract

Motion is defined as a sequence of positional changes over time. However, in perception, spatial position and motion dynamically interact with each other. This reciprocal interaction suggests that the perception of a moving object itself may dynamically evolve following the onset of motion. Here, we show evidence that the percept of a moving object systematically changes over time. In experiments, we introduced a transient gap in the motion sequence or a brief change in some feature (e.g., color or shape) of an otherwise smoothly moving target stimulus. Observers were highly sensitive to the gap or transient change if it occurred soon after motion onset ($\leqslant$200 ms), but significantly less so if it occurred later ($\geqslant$300 ms). Our findings suggest that the moving stimulus is initially perceived as a time series of discrete potentially isolatable frames; later failures to perceive change suggests that over time, the stimulus begins to be perceived as a single, indivisible gestalt integrated over space as well as time, which could well be the signature of an emergent stable motion percept.

Keywords: Motion; Integration; Shape; Color; Continuity; Gestalt

1. Introduction

In visual perception, position and motion interact reciprocally. Motion is a directional sequence of change in object position; therefore, it is not unnatural for the percept of stimulus position to affect one's perception of its motion. Apparent motion is a prominent example of a phenomenon in which positional information gives rise to a vivid percept of motion: the successive presentation of two (or more) discrete visual stimuli in distinct locations in space yields a vivid percept of movement (Wertheimer, 1912) from the first to the second stimulus after both stimuli have been presented.

The emerging apparent motion percept itself changes how other objects are perceived. Instantaneous positional information is created, or interpolated, based on the information available from temporally neighboring frames (Fahle and Poggio, 1981; Morgan, 1979, 1980). The interpolation in apparent motion is not merely a subjective impression (Burr, 1979), but influences the visibility of objects presented along the path of apparent motion (Yantis and Nakama, 1998). Recent neurophysiological and imaging studies (Jancke, Chavane, Naaman, and Grinvald, 2004; Muckli, Kohler, Kriegeskorte, and Singer, 2005) have found that the interpolation in apparent motion is associated with the activation of neurons representing the interpolated space in early visual cortex, which contains a neural representation of spatial position.

In light of the reciprocal interactions described above between position and motion in visual perception and the emergence of a motion percept after a delay, it appears likely that the perceptual representation of a moving object gradually emerges via dynamical recurrent processes in the brain. This would imply that the motion percept achieves equilibrium only after some amount of processing time. Indeed, there is some psychophysical

---

[*] Corresponding author. Fax: +1 626 792 8583.
E-mail address: kanair@caltech.edu (R. Kanai).
[1] These authors contributed equally to the work.

0042-6989/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2006.12.009

evidence that the quality of visual motion changes over a short duration following motion onset: The perceived image of a moving object becomes less blurred as the duration of motion increases (Burr, 1980, see also Chen, Bedell, and Ogmen, 1995), and the subjective blur asymptotes to a minimum only after about 120 ms after motion onset. A different set of studies on motion dynamics found that the ability to discern a subtle spatial or temporal offset between two bars moving in parallel improves with increasing stimulus duration (Morgan and Watt, 1982; Morgan, Watt, and McKee, 1983). This improvement continues for at least the first 150 ms following motion onset, and saturates at a constant level a little later. What is common to the two sets of studies is that both show that perception changes, and in fact improves, over the first few hundred milliseconds of a motion stimulus.

The goal of the present study is to further explore how perception changes over the first few hundreds of milliseconds of motion. Specifically, we explored how our perceptions of transient discontinuities within a smoothly moving stimulus might vary by comparing how well observers detected a brief change in an otherwise smooth motion sequence at different times following motion onset. As a stimulus continues to move and our subjective impression of its motion improves with it, our brains might switch from perceiving its motion as a series of stimuli at discrete positions in space to that of a single, indivisible gestalt. This, in turn, would increasingly impair, not improve, our ability to detect a transient change within the motion sequence. Indeed, we found that over a wide range of conditions and stimulus features, our ability to detect a transient change in some feature of a moving stimulus deteriorates during the course of its motion.

2. Experiment 1: Gap detection

2.1. Methods

2.1.1. Observers

Four observers participated in this experiment. One of the observers was one of the authors (R.K.). All observers had normal or corrected-to-normal vision.

2.1.2. Apparatus

The stimuli were controlled by an Apple Macintosh PowerPC running Matlab and presented on a 22 in. CRT monitor with a resolution of 1024 × 768 pixels at a 75 Hz refresh rate. The viewing distance was 57 cm and head movements were restrained by a chinrest. For the eye movement control experiment, we recorded eye position from the left eye at a sampling rate of 500 Hz using a SMI-Eyelink II system (SR Research, Inc.). Eye movements were recorded only in Experiment 1. The same set of apparatus was used in all other experiments unless stated otherwise.

2.1.3. Stimuli

The stimulus is schematically illustrated in Fig. 1a. The stimulus was a black bar (0.64° long and 0.16° wide) on a gray background (11.3 cd/m^2). Its motion (speed = 12 deg/s) consisted of a sequential presentation of 37 frames, each frame lasting for 13.3 ms (total duration = 493 ms). Successive frames physically abutted.

The bar initially appeared at a location 5.3° below, and 2.96° left or right of the fixation marker, and immediately began moving. On a given trial, the stimulus drifted either left or right. Trials of each direction were randomly intermixed. When the bar drifted right, it initially appeared 2.96° to the left of fixation, and when it drifted left, it initially appeared 2.96° to the right of fixation. Thus, in both cases, the bar passed just below the fixation marker on the 19th frame.

On a single trial, a one frame wide (0.16°) gap occurred at the 5th, 11th, 27th, or 33rd frame of the motion sequence. These correspond to stimulus onset asynchronies (SOAs) of 67, 147, 356 and 440 ms, respectively, from the onset of the motion. The gap was present on half of all trials. Note that the 5th and 33rd frames appeared at iso-eccentric positions with respect to fixation, as did the 11th and 27th frames. Because the motion was left or right on an equal proportion of trials, the spatial location relative to fixation of the two members of each iso-eccentric pair was identical across trials.

2.1.4. Task

The observer's task was to report the presence or absence of a gap in the motion. Each subject completed 20 trials for each of four gap positions in one block, with 10 trials each in the leftward and rightward directions. On half the trials, the gap was absent; these trials provide data on the false alarm rate. Each observer ran 160 (= 20 [trials] × 4 [SOAs] × 2 [gap present/absent]) trials. Trials of the different conditions were randomly intermixed. Feedback, in the form of an audible beep, was provided on error trials (false alarm or miss).

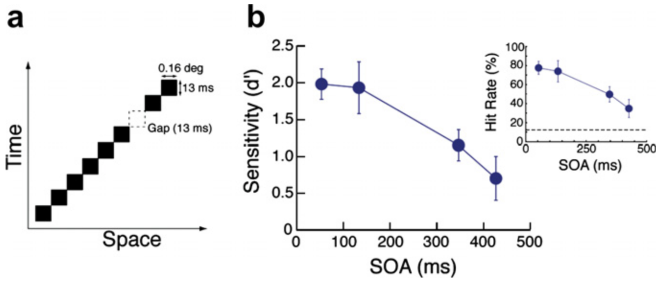

Fig. 1. Gap detection experiment. (a) A schematic space–time diagram of the stimuli. (b) Results of the standard gap detection task. The mean d' data from four observers is plotted (solid blue line, filled circles) against the duration of motion before the occurrence of the break (gap). The inset shows the hit rate (HR; blue solid circles) and false alarm rate (FAR; dotted line) before the conversion to d'. (For interpretation of color mentioned in the figure legend the reader is referred to the web version of the article.)

2.1.5. Eye movement control

Eye movements were recorded from all subjects during the experiment. Trials in which eye position deviated by more than 1° of visual angle from the center of a fixation cross during the stimulus presentation were discarded from the analysis. Under this criterion, 18.1 ± 8.4 (SEM) % of trials were discarded.

2.1.6. A control experiment with stationary stimuli

The stimulus was a stationary black bar (0.64° long and 0.16° wide) on a gray background. The bar appeared at one of the four locations used for the gap in the gap detection task, and stayed on for 493 ms. On half the trials, the bar was absent for one frame (13.3 ms) at a variable SOA (67, 147, 356 and 440 ms). The same four observers as in the gap detection experiment participated and had to judge if there was a temporal gap in the stimulus. For each condition, 40 trials were performed (20 trials with a temporal gap and 20 catch trials). Thus, there were a total of 640 (=40 trials × 4 positions × 4 SOAs) trials. The data from different positions were pooled and analyzed as a function of SOA.

2.2. Results

The results are shown in Fig. 1b. Change detection performance (sensitivity or d') is plotted as a function of SOA, that is, motion duration prior to the gap. Clearly, performance was not uniform across SOA ($F(3,9) = 6.30$, $p < 0.01$); Performance deteriorated with increase in SOA (Fig. 1b, Spearman $R = -0.728$, $p < 0.01$). Since we excluded trials on which fixation was poor (see Methods), the decrease in sensitivity is not attributable to eye movements, but must reflect a dynamic perceptual change associated with the preceding motion stimulus.

The results of the control experiment shows that performance (d') at detecting the temporal gap in the stationary stimulus was largely independent of SOA ($F(3,9) = 1.151$, $p = 0.343$, Spearman's $R = -0.213$, $p = 0.214$). Thus, the deterioration in performance over time following motion onset was unique to spatiotemporal characteristics of motion.

3. Experiment 2: Motion trajectory disruption

In order to examine the robustness of the effect, we had observers perform the same task, but under various conditions. Specifically, we introduced a disruption to the motion stimulus either by changing the luminance polarity of the bar (from black to white, or from white to black) or by placing an occluder in the motion path (on the same plane as, behind, or in front of the moving bar). One might imagine that the introduction of a discontinuity in the motion stimulus would interrupt the motion percept from being firmly established and might even restart the perceptual process all over again; this would leave the ability to detect the transient gap undiminished.

3.1. Methods

3.1.1. Observers and apparatus

Three new observers and one of the authors (R.K.) participated in the luminance-contrast reversal experiment and four new naïve observers participated in the occlusion experiments described below. The apparatus was identical to Experiment 1. In the occlusion experiment, the stimuli were viewed through stereoscope mirrors. The viewing distance through the mirrors was set at 57 cm. In the polarity change experiment, the stereoscope was not used, and the viewing distance was 57 cm. In both cases, the head movements were restrained using a chinrest.

3.1.2. Stimuli

Stimuli were similar to the main condition in Experiment 1 except that the motion stimulus was disrupted as it passed the middle (19th) frame either from an abrupt reversal in its luminance contrast relative to the gray background or by an occluder. On half the trials of the experiment with contrast reversals, the moving bar was initially black and changed its color on the 19th frame to white, and on the other half, it changed color from white to black. As in Experiment 1, a gap was present at one of the four locations on half of the trials. On the other half, there was no gap in the motion sequence. As before, four SOAs (67, 147, 356 and 440 ms) were tested. Thus, there were a total of 160 trials per observer (10 [trials] × 4 [SOAs] × 2 [gap present/absent] × 2 [luminance polarities]). Feedback was provided on error trials.

As for the occlusion experiment, the stimuli were again similar to Experiment 1 and the moving bar was always displayed in black. The occluder was a white stationary rectangle (3.2° long and 0.16° wide) centered at the middle (19th) frame. Three levels of disparities were tested in separate sessions; (1) The occluder was placed in front, (2) in the same depth plane, or (3) behind the moving bar. On conditions 1 and 3, the disparity was 0.16°. On condition 2, which was the condition with zero disparity, the moving bar was not depicted when it passed the position of the occluder.

3.2. Results and discussion

The results are shown in Fig. 2. The performance decline in detecting a gap was observed even when the bar reversed luminance contrast midway through its motion ($F(3,15) = 19.00$, $p < 0.001$, Spearman's $R = -0.68$, $p < 0.01$).

The performance decline was not affected by the addition of an occluder. A similar effect was obtained when the moving bar passed behind the occluder ($F(3,9) = 11.28$, $p < 0.001$, Spearman's $R = -0.584$, $p < 0.01$) as well as when it passed in front of the occluder ($F(3,9) = 10.63$, $p < 0.001$, Spearman's $R = -0.488$, $p < 0.05$). Although the performance was generally worse when the occluder was placed in the same depth plane as

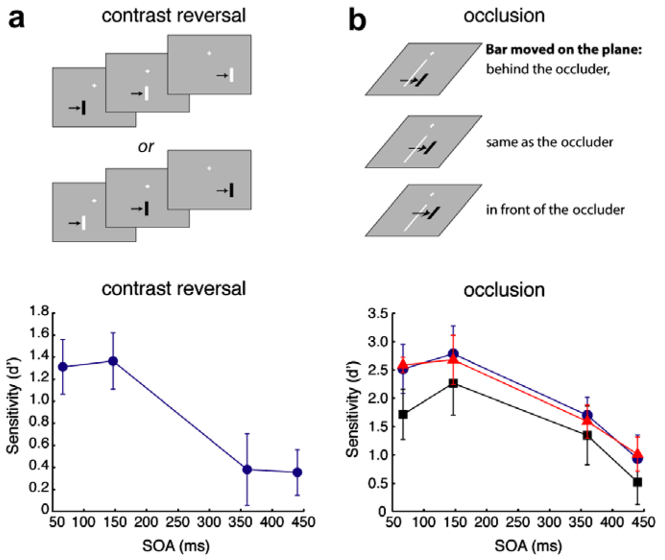

Fig. 2. Disruption experiments. (a) Contrast reversal experiment. Group mean ($n = 6$) performance on the gap detection when the moving bar reversed its luminance polarity during the movement. (b) Occlusion experiment. Group mean ($n = 4$) performance of gap detection when the moving bar traveled behind the occluding bar (red triangles), on the same plane as the occluder (black squares), and in front of the occluder (blue circles). (For interpretation of color mentioned in the figure legend the reader is referred to the web version of the article.)

the moving bar, the same decline pattern was observed in this condition as well ($F(3,9) = 12.75$, $p < 0.001$, Spearman's $R = -0.46$, $p < 0.05$). While it is unclear why the sensitivity differed in this condition, it appears to be due to a higher false alarm rate for the same plane condition (20.6% ± 7.0%, SEM) compared with the other conditions (near, 6.6% ± 2.7%; far, 6.6% ± 2.5%).

Taken together, the introduction of a local discontinuity in the motion trajectory failed to prevent the decline in detection performance with increasing SOA, demonstrating the robustness of the decline.

4. Experiment 3: Variable motion duration before or after the gap

In the experiments above, the total duration of motion was constant across trials of a given task. Therefore, the duration of motion before and after the gap varied together. In other words, as the duration of the motion before the change increased, the duration of the motion after the change decreased. In this experiment, we teased apart the contributions of the motion duration before and after the change to the decline in change detection ability.

4.1. Methods

4.1.1. Observers and apparatus

Nine new observers and one of the authors participated ($n = 10$) in this experiment. The experiment was conducted on a CRT monitor with a refresh rate of 100 Hz.

4.1.2. Stimuli and procedure

The bar was 0.08° by 0.96° in size and moved at a speed of 7.69 deg/s along a horizontal path that was 3.85° below fixation. The direction (leftwards or rightwards) of motion was randomized across trials. On half of the trials, a gap was inserted into the motion sequence. On separate blocks, either the duration of motion before or after the gap was varied (100, 150, 200, 250, 300, 350 or 400 ms), and respectively, the duration after, or before, the gap was fixed at 120 ms. As before, observers ($n = 10$) indicated if there was a gap in the motion. There was a total of 840 trials (=30 (trials per condition) × 7 (motion durations) × 2 (before/after) × 2 (target present/absent)) per observer. Feedback was provided on trials on which the subject erred (false alarm or miss).

4.2. Results and discussion

Fig. 3 illustrates the results. As expected, increase in the duration of motion prior to the gap impaired detection ($F(6,54) = 2.4564$, $p = 0.0651$; Spearman's $R = -0.2990$, $p = 0.0065$), whereas variation in the duration of motion after the gap had little effect on detection ($F(6,54) = 0.6181$, Spearman's $R = -0.0308$, $p = 0.3991$). This result confirms our findings above and further indicates that events that occur after the transient change do not affect the likelihood that the gap is detected. This finding thus argues against accounts of the present experiments based on off-line processing (Eagleman and Sejnowski, 2000; Kanai, Sheth, and Shimojo, 2004; Sheth, Nijhawan, and Shimojo, 2000; Sheth and Shimojo, 2003), which emphasize the influence of events occurring after a target event on perception. Instead, our results clearly argue that it is the motion trajectory prior to the critical transient event (gap) that has an influence on one's perception of it (Chappell and Hine, 2004).

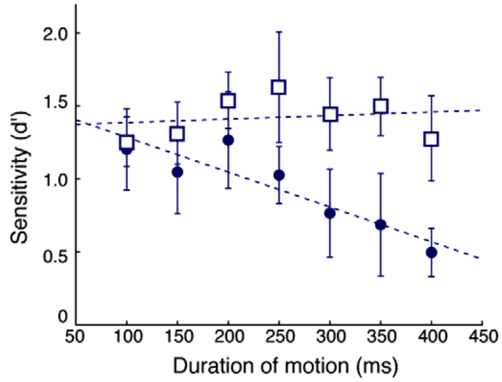

Fig. 3. Variable motion duration before or after the gap. Group mean ($n = 10$) sensitivity (d') is plotted as a function of motion duration before (blue solid circles) or after (blue open squares) the gap. The straight lines represent the optimal (in the least-squares sense) linear fits to each set of data (slopes: -0.0024 [variable duration before], 0.0002 [variable duration after]). Error bars indicate one SEM. (For interpretation of color mentioned in this figure the reader is referred to the web version of the article.)

5. Experiment 4: Detection of a brief change in shape

Thus far, we have shown that the detection of a gap becomes increasingly more difficult with increasing motion duration before the gap. In this experiment, we ask whether the impairment is specific to the gap, which could be interpolated by motion. Alternatively, detection of other types of brief changes, which is not as easily accountable by interpolation of motion, could also be impaired with a similar time course. To test this, we inserted a brief change in shape of the moving bar in the form of an elongation or widening. Observers now had to detect the small shape change instead of a gap.

5.1. Methods

5.1.1. Observers and apparatus

Three new observers and one of the authors (R.K.) participated. The apparatus was identical to Experiment 1, but eye movements were not monitored in this experiment.

5.1.2. Stimuli and procedure

The stimuli were identical to Experiment 1 with the exception that the gap was replaced with a shape change. The detection target was a thicker (parallel to the direction of motion) or longer (orthogonal to the direction of motion) bar embedded in the motion sequence. The thicker bar target was 0.48° wide (200% increase from the original). The longer bar target was 0.48° long (200% increase from the original). For each of the two shape change conditions, four SOAs (67, 147, 356 and 440 ms) were tested. For each condition, the observer ran 20 "target present" trials per SOA, and 80 "target absent" trials, for a total of 160 ($=20 \times 4 + 80$) trials. Task order between the thick bar detection and long bar detection was counterbalanced across observers. Feedback was provided on trials on which the subject erred (false alarm or miss).

5.2. Results

The results are shown in Fig. 4. On both the shape change conditions, change detection performance declined with increase in SOA (width change, $F(3,9) = 11.99$, $p < 0.001$; Spearman's $R = -0.37$, $p < 0.05$; length change, $F(3,9) = 8.05$, $p < 0.01$; Spearman's $R = -0.31$, $p = 0.06$). Observers were less capable at detecting the change when the change was along the direction of stimulus motion (i.e., width change) than in the direction orthogonal to it, $F(1,3) = 39.69$, $p < 0.01$. Of more importance is that these results indicate that the decline in change detection ability with increasing SOA is not confined merely to the detection of a gap, but is observed for the detection of other visual features as well.

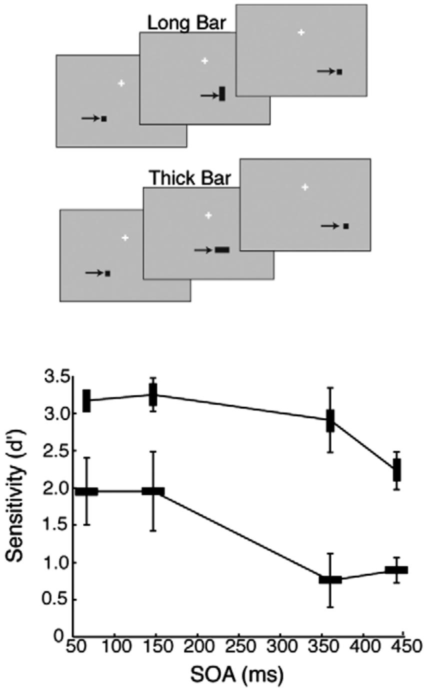

Fig. 4. The shape change experiment. Group mean ($n = 4$) performance on the length change detection (vertically elongated rectangles; the change is along a direction orthogonal to the motion) and width change detection (horizontally elongated rectangles; the change is along the direction of the motion) task is plotted as a function of motion duration prior to the respective changes. Motion (indicated by horizontal black arrow in the figure) is in the horizontal direction (left → right or right → left, depending on the trial). Error bars indicate one SEM.

6. Experiment 5: Color fusion depends on the duration of prior motion

We wondered about the limits of the effect of prior motion on the observer's ability to identify a brief change. Specifically, we wondered if the decline in the ability to detect a transient change described in the experiments above extends to features like color that, unlike shape and position, do not change along physical dimensions of space. To this end, we tapped into the phenomenon of motion-induced color fusion (Nishida, Watanabe, Tachi, and Kuriki, 2004): When the color of a moving stimulus alternates between two colors (e.g., red and green), the color of the moving stimulus is perceived as the mixture of the two (i.e., yellow). In this experiment, we examined the effect of prior motion duration on the perception of color mixture.

6.1. Methods

6.1.1. Observers

Four naïve observers participated.

6.1.2. Stimuli and procedure

The stimulus was a moving bar that alternated in color between equiluminant red ($x = 0.612$, $y = 0.343$) and green

($x = 0.298$, $y = 0.586$). But the luminance of the moving bar was not adjusted to the equiluminant level of the background. The center of the bar was at the same height as the fixation point. Four motion durations (80, 213, 347, and 480 ms, corresponding to 6, 16, 26 or 36 frames, respectively) were tested across trials. The spatial location of the last frame was fixed (8° left of fixation). This eccentricity was chosen on the basis of pilot experiments that were carried out to determine an appropriate level of task difficulty. On separate, randomly intermixed trials, the last frame of the motion was, with equal probability, either red or green R/G; (10 trials each), or yellow (Y; 20 trials), which was a fusion of red and green (Fig. 5a). This resulted in a total of 160 trials (= 40 [trials] × 4 [SOAs]). The observer's task was to report the color of the last frame (R/G or Y). Feedback was provided on trials on which the subject erred.

6.2. Results and discussion

The results are shown in Fig. 5. As the motion duration on the R/G trials increased, observers increasingly reported the color of the last frame as being yellow (Fig. 5b; $F(3,9) = 23.3$, $p < 0.001$, Spearman's $R = -0.41$, $p < 0.05$), implying that it became increasingly difficult for the observers to individuate the color of the final frame. This decline in performance with SOA is in line with our findings from previous experiments. That is, one becomes increasingly insensitive to transient change occurring within some feature of the moving stimulus. On the basis of these findings, it is tempting to speculate that over time, individual frames of the motion trajectory gradually become indivisible perceptually and motion is perceived as a single, coherent gestalt.

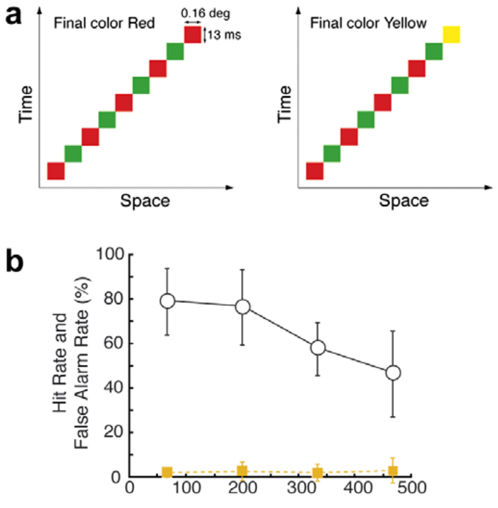

Fig. 5. The color fusion experiment. (a) A schematic space–time diagram of the stimuli. A bar moved from left to right alternating between red and green colors on every frame (13 ms). (b) The mean hit rate (open black circles, $n = 4$) and false alarm rate are plotted as a function of motion duration before the last frame. A hit is defined if the last frame is red or green in color and is perceived as such. A false alarm is defined if the last frame is yellow, but is perceived to be either red or green. Error bars indicate one SEM. (For interpretation of color mentioned in this figure the reader is referred to the web version of the article.)

7. Experiment 6: Motion blur

The present findings suggest that there is a critical difference in how motion is perceived in its early versus later, more mature stage. Early in the motion, the observer is sensitive to subtle changes within the motion sequence, but later in the motion, the observer becomes far less sensitive. As described earlier, the phenomenon of motion deblurring or interpolation demonstrates the dynamical nature of motion perception (Burr, 1980; Burr, Ross, and Morrone, 1986a, Burr, Ross, and Morrone, 1986b; Morgan and Watt, 1982; Morgan et al., 1983). In this experiment, we investigate if the decline in change detection performance with SOA parallels the time course of motion deblurring. Previous studies have shown motion blur decreases over time and reaches a stable minimum level at about 120–150 ms following motion onset (Burr, 1980). While this suggests that blur decreases to a minimum level earlier in time following motion onset than the decline in change detection (~300 ms), we nonetheless investigated the temporal properties of motion deblurring on our set of stimulus parameters and task conditions.

7.1. Methods

7.1.1. Observers and apparatus

Four observers participated.

7.1.2. Stimuli and procedure

Stimulus parameters were the same as in the Experiment 1 except that the motion terminated after a variable duration (67, 147, 356 or 440 ms). The direction of motion was randomized across trials. After the presentation of the motion stimulus, there was a 800 ms blank interval, and then a probe rectangle was presented at the position of the last frame of the motion. Observers had to judge if the width of the moving bar (including the streak) was wider or thinner than that of the probe. Each observer performed six staircases interleaved across trials. Half of the staircases started 0.12° thicker than the moving bar, and the other half started 0.12° thinner. The width of the probe was changed in steps of size 0.04° (one pixel). Each staircase stopped after ten reversals, and the mean of the last six points of reversal was used as a measure of the perceived blur.

7.2. Results and discussion

The perceived blur of a moving stimulus is shown in Fig. 6 as a function of motion duration prior to the point of judgment. Consistent with the literature (Burr, 1980), our results show that perceived blur decreased rapidly over time and fell to a stable minimum already at 147 ms following motion onset (Fig. 6). Thus, the early decline indicates

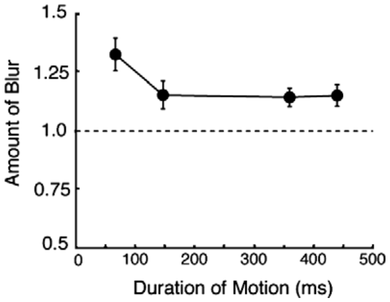

Fig. 6. The group mean ($n = 4$) subjective blur versus motion duration is shown on a normalized scale where 1.0 on the ordinate corresponds to the actual thickness of the moving bar and zero blur, and values increasingly greater than one indicate increasing amounts of blur. Error bars indicate one SEM.

that the temporal dynamics of motion blur and our change detection decline do not coincide. Furthermore, a decline in the ability to detect transient change within the motion sequence with time following motion onset is more likely to correspond to increase, and not decrease, in blur. The present findings of change detection dynamics share with studies of motion blur the theme that the perception of motion evolves over the initial few hundreds of milliseconds. Unlike motion blur however, in the case of change detection, the perception of motion transients drifts further away from veridical with time.

8. General discussion

We have shown that the perception of the features of a moving stimulus significantly deteriorates with increasing motion duration. A transient change in a feature (e.g., presence, size, color) of a moving stimulus was easily noticeable if it occurred early in the motion (<300 ms), but less noticeable if it occurred later in the motion. The effect was not attributable to eye movements or motion blur, but appeared to stem from the unique dynamics of motion perception instead. As motion duration increased, individual frames of the motion trajectory were no longer perceived as distinct, discrete entities, but as components of an indivisible whole (see also Kanai, Paffen, Gerbino, and Verstraten, 2004).

8.1. A computational model of the gap detection task

Previously, many models on spatiotemporal integration were proposed to account for the detection of image motion (Adelson and Bergen, 1985; Pantle and Hicks, 1985; Reichardt, 1961; van Santen and Sperling, 1984, 1985; Watson and Ahumada, 1985). Burr and colleague (1986a, 1986b) applied the idea of spatiotemporal receptive fields to explain the perceptual interpolation in apparent motion. If we assume that a position of a moving stimulus is coded in the population response of neurons with a spatiotemporal receptive field oriented in the motion response, the neurons centered at a gap would be sufficiently activated both by the preceding and succeeding motion trajectory (see the bottom row of Fig. 7), thereby filling-in the gap in the original stimulus.

However, as the present study has shown (see also Burr, 1979), the interpolation starts 200–300 ms following motion onset. In order for the spatiotemporal interpolation to be effective, the filter set for positional estimation needs to be tuned for motion speed. However, when a motion stimulus is first presented, the speed is still uncertain for the system. Therefore, in our model, we assumed that the spatiotemporal receptive fields of the model neurons for positional coding are initially tuned for a stationary stimulus, and its orientation approaches the actual speed exponentially with a time constant of 200 ms following the onset of motion.

We modeled the response of the ith neuron $r_i(t)$ contributing to positional estimation at time t as the following:

$$r_i(t) = \int_{-\infty}^{t} \int_{-\infty}^{\infty} s(x,t') f^{(i,t)}(x, t - t') \mathrm{d}x \mathrm{d}t'$$

where $f^{(i,t)}(x,t)$ is the spatiotemporal receptive field of the ith neuron at time t, and $s(x,t)$ is the stimulus intensity. The stimulus had a value of 1 in our model, and the background had a value of 0. The receptive field was modeled as an oriented Gabor in space and time, and the spatiotemporal unit was scaled such that an orientation of 45° corresponded to the speed of stimulus motion. The orientation of the Gabor was initially flat, assuming that the speed estimate is stationary (see Weiss, Simoncelli, and Adelson, 2002) and then it asymptoted to a value of 45° as a function of time: $\theta(t) = 45(1 - \mathrm{e}^{(-t/\tau)})$, where the time constant τ was 200 ms. Note that in this model, we are not suggesting that the receptive fields of individual neurons dynamically change as a function of time. Instead, we assumed that the selection of neural population for positional estimation be re-adjusted online as the estimate of speed becomes more veridical. The spatiotemporal center frequency of the Gabor was 0.03 ms^{-1} and sigma 12 ms (in the normalized unit space). The Gabor filters were normalized such that the summation over space and time equaled 1. The responses were rectified at a threshold of 0.2.

The results of the simulation are shown in Fig. 7. We have tested the four gap positions used in Experiment 1 (5th, 11th, 27th and 33th with each frame lasting for 13 ms). As can be seen, the gap is clearly visible in the population response of the model neurons when presented early in the motion (Fig. 7a). As the spatiotemporal receptive fields become increasingly tuned to the motion speed over time, the gap becomes more interpolated in the population response. Our model qualitatively reproduced the decline in the gap detection performance. In addition, the model reproduced the results of the wide bar detection task and motion deblurring experiment using the same parameter set (data not shown).

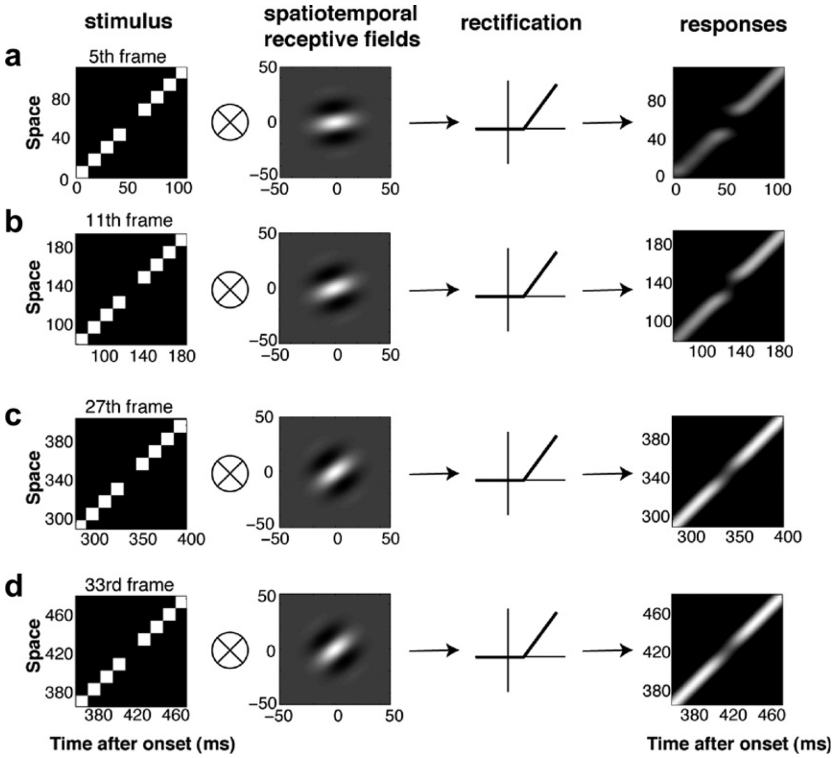

Fig. 7. Simulation results. The leftmost column shows the input stimulus with a gap at different temporal positions (a) 5th frame, (b) 11th frame, (c) 27th frame and (d) 33rd frame). In the second column, the orientation of the spatiotemporal receptive field is shown for the gap position. Results of the convolution were rectified, as depicted in the third column from the left, and the final rectified outputs are shown in the rightmost column.

8.2. Psychological implications

In addition to not being able to detect a gap late in the motion, observers fail to notice other kinds of brief changes occurring to the moving object as well. Insensitivity to transient change can be regarded as a difficulty in individuating a discrete event in a rapidly changing visual environment. Temporal decline in the ability to detect such change suggests that early in the motion when a stable motion percept is not yet fully formed, the events or motion frames occurring in succession are processed as discrete entities making it easy to detect change in each; as the motion percept evolves, identical events are perceptually integrated into a single, continuous object, thereby rendering individual events indistinguishable.

In line with this idea, it is known that the perception of motion between two discrete events binds them into a single perceptual entity making them harder to distinguish from one another. When apparent motion is perceived between two visual events, a single object is perceived (Anstis, 1980; Wertheimer, 1912). Chun and Cavanagh (1997) have gone further by showing a reciprocal relationship between the perception of motion and the perceptual individuation of the physical events that constitute the motion percept. In their experiments, when apparent motion was perceived between two discrete stimuli, observers failed to individuate successive events of the same type—a phenomenon known as repetition blindness. On the other hand, when apparent motion was not perceived in the same stimulus configuration, the two stimuli were perceived as two separate events, and repetition blindness was drastically reduced. These findings support the notion that motion perception is involved in the retention of spatiotemporal continuity of an object.

In the case of gap detection, frames are initially processed as separate events, but when motion percept is established for the object, information from the frames neighboring the critical gap is integrated, thus filling in the missing information (i.e., the gap). The same idea holds for shape-change detection, where the shapes at different frames are integrated, which will result in less sensitivity to a transient change in the shape of an individual frame of the moving object.

In conclusion, the experimental results bolster earlier ideas that the perception of motion itself is dynamic, that our early percept of a moving object is of a stroboscopic nature somewhat resembling that of a patient with akinetopsia (Rizzo, Nawrot, and Zihl, 1995), and that a Gestalt perception of motion arises only later during the motion. The decline in our ability to detect transient change within a smooth motion trajectory 200–300 ms following motion onset could paradoxically herald the emergence of a stable motion percept.

Acknowledgments

We thank Haluk Ögmen, Saumil Patel and Daw-An Wu for insightful discussion and comments on an earlier version of the manuscript. We thank David Burr, Romi Nijhawan and an anonymous reviewer for their candid and insightful criticisms.

References

Adelson, E. H., & Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America A, 2*(2), 284–299.

Anstis, S. M. (1980). The perception of apparent movement. *Philosophical Transactions of the Royal Society of London. Series B. Biological Sciences, 290*(1038), 153–168.

Burr, D. C. (1979). Acuity for apparent vernier offset. *Vision Research, 19*, 835–837.

Burr, D. (1980). Motion smear. *Nature, 284*(5752), 164–165.

Burr, D. C., Ross, J., & Morrone, M. C. (1986a). Seeing objects in motion. *Philosophical Transactions of the Royal Society of London. Series B. Biological Sciences, 227*(1247), 249–265.

Burr, D. C., Ross, J., & Morrone, M. C. (1986b). Smooth and sampled motion. *Vision Research, 26*(4), 643–652.

Chappell, M., & Hine, T. J. (2004). Events before the flash do influence the flash-lag magnitude. *Vision Research, 44*(3), 235–239.

Chen, S., Bedell, H. E., & Ogmen, H. (1995). A target in real motion appears blurred in the absence of other proximal moving targets. *Vision Research, 35*(16), 2315–2328.

Chun, M. M., & Cavanagh, P. (1997). Seeing two as one: linking apparent motion and repetition blindness. *Psychological Science, 8*, 74–79.

Eagleman, D. M., & Sejnowski, T. J. (2000). Motion integration and postdiction in visual awareness. *Science, 287*(5460), 2036–2038.

Fahle, M., & Poggio, T. (1981). Visual hyperacuity:spatiotemporal interpolation in human vision. *Philosophical Transactions of the Royal Society of London. Series B. Biological Sciences, 213*(1193), 451–477.

Jancke, D., Chavane, F., Naaman, S., & Grinvald, A. (2004). Imaging cortical correlates of illusion in early visual cortex. *Nature, 428*(6981), 423–426.

Kanai, R., Paffen, C. L., Gerbino, W., & Verstraten, F. A. (2004). Blindness to inconsistent local signals in motion transparency from oscillating dots. *Vision Research, 44*(19), 2207–2212.

Kanai, R., Sheth, B. R., & Shimojo, S. (2004). Stopping the motion and sleuthing the flash-lag effect: spatial uncertainty is the key to perceptual mislocalization. *Vision Research, 44*(22), 2605–2619.

Morgan, M. J. (1979). Perception of continuity in stroboscopic motion: a temporal frequency analysis. *Vision Research, 19*(5), 491–500.

Morgan, M. J. (1980). Spatiotemporal filtering and the interpolation effect in apparent motion. *Perception, 9*(2), 161–174.

Morgan, M. J., & Watt, R. J. (1982). Effect of motion sweep duration and number of stations upon interpolation in discontinuous motion. *Vision Research, 22*(10), 1277–1284.

Morgan, M. J., Watt, R. J., & McKee, S. P. (1983). Exposure duration affects the sensitivity of vernier acuity to target motion. *Vision Research, 23*(5), 541–546.

Muckli, L., Kohler, A., Kriegeskorte, N., & Singer, W. (2005). Primary visual cortex activity along the apparent-motion trace reflects illusory perception. *PLoS Biology, 3*(8), e265.

Nishida, S., Watanabe, J., Tachi, S., & Kuriki, I. (2004). Motion-induced colour mixture. *Journal of Vision, 4*(8), 162a.

Pantle, A., & Hicks, K. (1985). Using low-level filters to encode spatial displacements of visual stimuli. *Spatial Vision, 1*(1), 69–82.

Reichardt, W. (1961). Autocorrelation, a principle for the evaluation of sensory information by the central nervous system. In W. A. Rosenblith (Ed.), *Sensory communication* (pp. 303–317). New York: Wiley.

Rizzo, M., Nawrot, M., & Zihl, J. (1995). Motion and shape perception in cerebral akinetopsia. *Brain, 118*(Pt 5), 1105–1127.

van Santen, J. P. H., & Sperling, G. (1984). Temporal covariance model of human motion perception. *Journal of the Optical Society of America A, 1*, 451–473.

van Santen, J. P. H., & Sperling, G. (1985). Elaborated Reichardt detectors. *Journal of the Optical Society of America A, 2*, 300–321.

Sheth, B. R., Nijhawan, R., & Shimojo, S. (2000). Changing objects lead briefly flashed ones. *Nature Neuroscience, 3*(5), 489–495.

Sheth, B. R., & Shimojo, S. (2003). Signal strength determines the nature of the relationship between perception and working memory. *Journal of Cognitive Neuroscience, 15*(2), 173–184.

Watson, A. B., & Ahumada, A. J. Jr., (1985). Model of human visual-motion sensing. *Journal of the Optical Society of America A, 2*, 322–341.

Weiss, Y., Simoncelli, E. P., & Adelson, E. H. (2002). Motion illusions as optimal percepts. *Nature Neuroscience, 5*, 598–604.

Wertheimer, M. (1912). Experimentelle Studien über das Sehen von Bewegung. *Zeitschrift für Psychologie, 61*, 161–265.

Yantis, S., & Nakama, T. (1998). Visual interactions in the path of apparent motion. *Nature Neuroscience, 1*(6), 508–512.