The flash-lag effect as a spatiotemporal correlation structure

Ikuya Murakami

Human and Information Science Laboratory, NTT Communication Science Laboratories, NTT Corporation, Atsugi, Kanagawa, Japan

The flash-lag effect refers to the phenomenon in which a flash adjacent to a continuously moving object is perceived to lag behind it. Phenomenally, the flash appears to be spatially shifted relative to the moving stimulus, and the amount of lag has often been quantified as the flash's nulling position, which is the physical spatial offset needed to establish perceptual alignment. The present study offers a better way to summarize flash-lag data. Instead of plotting data in terms of space, the psychometric function of the observer's relative-position judgment is drawn on spatiotemporal plot. The psychological process underlying illusory lag is formulated as spatiotemporal bias and uncertainty and their estimate as a spatiotemporal convolution kernel that best explains the spatiotemporal psychometric function. Two empirical procedures of kernel estimation are described. One procedure is to fit the free parameters of the kernel to experimental data for continuous motion trajectory. The second is to give an analytical solution to the kernel using experimental data for random motion trajectory. The two procedures yield similar kernels, with negligible spatial bias and uncertainty and substantial temporal bias and uncertainty. In addition, it is demonstrated that an experimental manipulation of temporal predictability of the flash can change the temporal bias in the estimated kernel. The results of this novel analysis reveal that the flash-lag effect is viewed as a spatiotemporal correlation structure, which is largely characterized by the tendency to compare the position of the flash in the past with the position of the moving item in the present.

Keywords: flash-lag effect, relative position, psychometric function, bias, uncertainty, spatiotemporal kernel

Introduction

When a brief flash is presented adjacent to a moving stimulus, the flash appears to lag behind (for review, see Krekelberg & Lappe, 2001). This illusion has been replicated in various stimulus configurations (Baldo & Klein, 1995; Khurana & Nijhawan, 1995; Nijhawan, 1997; Kirschfeld & Kammer, 1999; Brenner & Smeets, 2000; Eagleman & Sejnowski, 2000a; Khurana, Watanabe, & Nijhawan, 2000). As illustrated in Figure 1, one feels a vivid impression that the flash's position relative to the moving stimulus is phenomenally shifted. As such, the amount of flash-lag has been operationally defined as the amount of illusory spatial offset. To measure this amount in psychophysical experiments, the nulling method has usually been used: the observer adjusts the physical position of the flash in the direction opposite the illusory spatial lag, just to cancel it.

Technically speaking, there is nothing problematic in using the nulling method. However, interpretations in terms of purely spatial account can be quite misleading. For example, the theory of “motion extrapolation” argues that the visual system actively shifts the spatial representation of the bar toward the direction of motion (Nijhawan, 1994, 1997; Khurana & Nijhawan, 1995). On the contrary, several follow-up studies have found contradictory results, casting doubt on this strong hypothesis (Baldo & Klein, 1995; Lappe & Krekelberg, 1998; Purushothaman, Patel, Bedell, & Ogmen, 1998; Whitney & Murakami, 1998; Brenner & Smeets, 2000; Eagleman & Sejnowski, 2000a; Whitney, Cavanagh, & Murakami, 2000a; Whitney, Murakami, & Cavanagh, 2000b). Some researchers proposed the alternative model that the flash simply requires a longer latency than the moving stimulus, such that the apparent spatial offset is actually a temporal one (Mateeff & Hohnsbein, 1988; Baldo & Klein, 1995).
depending on the flash’s predictability.

Two empirical approaches to extracting the spatiotemporal relative position; the internal bias and fluctuation of perceptual alignment are described in terms of convolution kernel. Two empirical approaches to extracting the kernel structure out of a spatiotemporal psychometric function are proposed. It is also shown that the kernel changes its shape depending on the flash’s predictability.

**Methods**

This study followed Declaration of Helsinki guidelines and was approved by NTT Communication Science Laboratories Research Ethics Committee. Informed consent was obtained from observers after explanation of the nature and possible consequences of the study.

Stimuli were presented on a 21-inch color CRT monitor (Sony GDM-F500, 1024 pixels × 768 pixels, refresh rate 119.8 Hz) controlled by a computer (Apple Power Macintosh). For the sake of clarity, all the spatial terms are described in pixels (1 pixel = 2.5 arcmin) and time is in video frames (1 frame = 8.35 ms). In a darkened room, the observer’s head was immobilized with a chinrest; the viewing distance was 54 cm. The right eye was used. The fixation point was provided throughout the experiment.

The schematic of the stimulus configuration is shown in Figure 1. The moving bar was actually a pair of upright rectangles (each 4 pixels × 16 pixels; 11.2-cd/m² gray on the 24.4-cd/m² white background) arranged collinearly with the gap of 20 pixels (they always moved synchronously as though a single bar; the singular form “bar” will be used to refer to these two rectangles). In each trial, the moving bar made a horizontal translation at a constant speed along a line 2 degrees below the fixation point (leftward or rightward chosen randomly at each trial). At a random timing, another upright rectangle (4 pixels × 16 pixels; its brightness having been subjectively equated with that of the moving bar) or the flash was briefly presented for one frame, with its vertical position between the two rectangles comprising the moving bar. The flash’s horizontal position was chosen randomly from a range of appropriate position levels for making a psychometric function. The observer judged whether the flash was seen to the left or right of the moving bar. The best-fit cumulative Gaussian was estimated by the maximum likelihood method.

Six different speeds (0.25, 0.5, 0.75, 1, 1.5, 2 pixels/frame) of the moving bar were tested in separate sessions.

**Results**

As the first step of analysis, let us visualize data in a conventional way (Figure 2). Only the author’s data will be shown, but naive observers exhibited quantitatively similar results. For each speed of the moving bar, the percentage of “right” (i.e., as opposed to “left”) responses is plotted as a function of the physical position of the flash (data in the leftward-motion condition were flipped and merged to the rightward-motion data). From this data structure, one can confirm at least three well-known aspects of the flash-lag effect. First, there was only a negligible percentage of “right” responses at the spatial offset of zero, indicating that the flash that was physically aligned with the moving bar had a strong tendency to be seen spatially offset in the opposite direction. As mentioned earlier, this is the definition of the flash-lag effect (Nijhawan, 1994). Second, the point of subjective equality (PSE), or the flash’s nulling position, was found to be a few pixels in the positive direction, indicating that the flash has to be physically shifted in the direction of motion in order to appear aligned with the moving bar. As the psychometric function exhibits quite systematic behavior, the nulling procedure seems useful in quantifying the effect, as many laboratories have already reported (citations in the “Introduction”). Third, the PSE increased with increasing speed (Figure 2H).

Thus far, there is nothing new. However, the data in Figure 2 indicate another important aspect: with increasing speed, the PSE not only becomes larger, but the slope of the psychometric function systematically becomes shallower (Figure 2A), and equivalently, the standard deviation of the function becomes greater (error bars of Figure 2H). Why does this happen? To answer this, a speculative theory could be created. For example, one could propose a spatial-offset mechanism with some sort of gain control, so that when the speed of motion is decreased, illusory spatial offset should be represented in higher and higher resolution, and so the offset should be less and less noisy. Instead, the aim of the present study is to offer a methodological solution to the most parsimonious conclusion from only a few acceptable assumptions about our psychological process.

**Spatial Bias and Uncertainty**

First of all, what is the theoretical background of a sigmoidal psychometric function, Ψ(x) (Figure 3A)? Clearly, it stems from the bias and uncertainty that must occur whenever the visual system encodes signals from the outer world. If the observer is perfectly noise-free such that the flash’s position and the moving bar’s position are represented with perfect accuracy and precision, the psychometric function is simply reduced to a step function, Φ(x) (Figure 3B). In a realistic system, however, some bias and uncertainty are inevitable. Suppose that for some reason a flash spatially shifted in the
Figure 2. Typical psychometric functions in the continuous-motion experiment. The percentage of “right” responses is plotted as a function of the physical position of the flash presented for one frame while the moving bar was in rightward motion. A. All data. B-G. The same data replotted separately for each speed condition. Note that different spatial scales are used across panels. In each panel, the green curve indicates the best-fit cumulative Gaussian found independently for each speed condition. The red curve indicates the result of the novel analysis in the present study: convolution of $Φ(x − t)$ by $K(x, t)$. Note that exactly the same shape of $K$ was used irrespective of speed. The four parameters of $K$ are indicated at the bottom. H. Spatial PSE as a function of speed. The $μ_x$ of the best-fit cumulative Gaussian (green curves of B-G) is plotted as a function of speed, with $σ_x$ as the error bar. The blue line indicates the linear regression.

Figure 3. Theoretical background of a psychometric function. The abscissa of each panel indicates the flash’s position relative to the moving bar. The ordinate indicates probability (%). $Ψ(x)$ denotes an observed psychometric function, which is characterized by a cumulative Gaussian. $Φ(x)$ denotes the psychometric function of a hypothetical perfect observer who knows positions of all stimuli with perfect accuracy and precision. $K(x)$ denotes the probability density function of perceptual alignment. Operator * denotes convolution.
for all differences between the perfect psychometric function \( (\Phi) \) and the one that is observed \( (\Psi) \). In any event, the key concept of this section is that \( \Psi = \Phi \ast K \).

**Spatiotemporal Bias and Uncertainty**

\( \Psi \) was observed. \( \Phi \) is given. \( K \) is the solution to the behavior of the “black box” process. Unfortunately, however, the kernel \( K \) shown in Figure 3C is not the only correct solution, but just one of many. Why? Because bias and uncertainty can also occur along time.

In Figure 4A, the data in Figure 2E are reproduced as a spatiotemporal plot. The diagonal line indicates the trajectory of the moving bar (rightward at 1 pixel/frame). Each colored point is the observer's response to the flash; the origin is set at the moving bar that was presented simultaneously with the flash. At first glance, one might be disappointed with the apparent scarcity of data points. What will be seen if the flash is presented somewhere other than along the abscissa? Why not get more data?

Actually, all data are already available. Recall that the moving bar was in continuous translation and that the flash was presented at a random timing during the translation of the moving bar. To the observer, there was no visible indication (e.g., change in direction) of the moving bar at the frame when the flash was presented. Therefore, the moving bar at the origin of Figure 4A has no special meaning. One can plot the observer's response to the flash that was presented at, for instance, +5 frames with respect to the coordinates in Figure 4A, simply by looking at the data structure of Figure 4A from the viewpoint of the moving bar 5 frames before the flash's presentation. In Figure 4B, these responses are plotted simply by shifting the spatiotemporal coordinates and setting their origin at the moving bar 5 frames before the flash. Relative to this particular moving bar, all the flashes were presented at 5 frames in the future. One can repeat the same procedure by setting the origin at each spatiotemporal instance of the moving bar and by plotting responses according to the new coordinates. The superposition of these plots is shown in Figure 4C. Note that in plotting them there is neither interpolation nor extrapolation of raw data: each point is not an expected hypothetical result that would have been obtained if measured actually, but a result of actual measurement at each spatiotemporal position. By applying the best-fit cumulative Gaussian shown as the green curve in Figure 2E, one gets a smoothly curved surface shown in Figure 5A. Each point on this surface indicates the percentage of “right” responses to the flash presented at each spatiotemporal position. Let us call this surface a spatiotemporal psychometric function.

The profiles shown in Figure 4A and B are the spatiotemporal events that happen in actual space-time. The difference between these figures is that only the origin of space-time is set at a different instance of the moving bar. Specifically, the color profile in Figure 4A can be written as \( p = f(x, t) \), where \( f \) indicates the percentage of “right” responses to the flash at a spatiotemporal point \((x, t)\). Likewise, the profile in Figure 4B is \( p = f(x - 5, t - 5) \), plotted relative to the moving bar presented at...
Figure 5. A spatiotemporal version of the relationship illustrated in Figure 3. The data for the speed of 1 pixel/frame are chosen as an example. A. A colored-surface plot of the observed psychometric function. The horizontal sigmoid shown in Figure 2E (green curve) is applied to the chart shown in Figure 4C. B. The psychometric function of a hypothetical perfect observer who knows spatiotemporal positions of all stimuli with perfect accuracy and precision. C. The probability density function of perceptual spatiotemporal alignment, or the kernel. Probability density is plotted by a color scale with an arbitrary gain. The shape of the kernel is not uniquely determined from the data shown in A. This shape is the result of a fitting procedure: the most likely estimate of the kernel that best explains data for all speed conditions. See text for details.
the flash-lag effect is viewed as a spatiotemporal correlation structure, the particular data set used in Figure 5 is not informative enough to determine the shape of \( K \).

**Finding the Best-Fit Spatiotemporal Kernel**

The difficulty is, however, solved practically by comparing data across other speed conditions. In particular, let us focus on the slopes of spatiotemporal psychometric functions. For example, the spatiotemporal psychometric function for the speed of 2 pixels/frame is shown in Figure 6. Note that in the original chart for this condition (Figure 2G), the psychometric function had a shallower slope. As a result, the spatiotemporal psychometric function also looks elongated spatially. In contrast, the sigmoidal shape along the time axis does not seem much different from that of Figure 5. It follows from this observation that a common kernel with a lot of temporal rather than spatial uncertainty may better explain the data in both speed conditions.

Under the assumption that kernel \( K \) maintains a common shape irrespective of the moving bar’s speed, one can find the best-fit parameters of \( \mu, \sigma, \mu_t, \) and \( \sigma_t \) that minimize the residual between the model \( (\Phi * K) \) and the observed data \( (\Psi) \) for all speed conditions. Using the 66 (11 position levels \( \times \) 6 speeds) points as the data set, Levenberg-Marquardt nonlinear optimization yielded \( (\mu, \sigma, \mu_t, \sigma_t) = (2.10 \text{ pixels}, 1.75 \text{ pixels}, -4.95 \text{ frames}, 3.38 \text{ frames}, \text{respectively}) \). The spatiotemporal plot of the kernel is shown in Figure 5C. Convolution of the perfect psychometric function with this best-fit spatiotemporal kernel resulted in a fairly good approximation to the actual data; the resulting theoretical profiles are drawn as the red curves in Figure 2. Importantly, exactly the same kernel is used throughout the six different speed conditions. These sigmoidal shapes are virtually indistinguishable from the one-dimensional cumulative-Gaussian fit applied separately for each condition (green curves in Figure 2).

This analysis strongly suggests that the flash-lag effect is to a large extent a temporal rather than spatial misjudgment. The best-fit kernel has its peak at roughly 5 frames past \( (\mu_t) \), indicating that the observer somehow compared the relative position between a moving bar at the present and a flash in the past, as though they were stimuli seen simultaneously. As the kernel is elongated temporally \( (\sigma_t) \), the perceptual simultaneity between the moving bar and flash has a large temporal uncertainty. The kernel also shows a little amount of spatial uncertainty \( (\sigma_t) \), but it is in fact in an excellent agreement with the sensitivity of spatial vernier acuity between a flash and a stationary bar at the tested eccentricity range. (A control experiment measured this acuity by performing the same test with the speed of the bar at 0 pixels/frame. The \( \sigma_t \) in this condition was found to be 1.75 pixels.) Finally, the kernel has a slight spatial offset in the direction of motion \( (\mu_t) \). Thus, the observer somehow judged the moving bar and the flash with 2-pixels offset as perceptually aligned.

**The Flash-Lag Effect in Random Motion**

With the understanding that the flash-lag effect is viewed as a spatiotemporal correlation structure, the next step is to seek a better stimulus configuration that is more suitable for estimating the parameters of the internal kernel. Up to the previous section, continuous motions with various speeds have been used for the sake of consistency with previous studies. However, the diagonal autocorrelation structure of the moving bar itself always complicates the shape of the observed spatiotemporal psychometric function. In that situation, estimating the kernel’s spatial component is always confounded by its temporal component, and vice versa. In the previous section, decorrelation of \( \Psi \) to a combination of \( \Phi \) and \( K \) was made possible only by preparing multiple samples from different speed conditions and by finding best-fit parameters.

Things become simpler by making the moving item’s autocorrelation simpler. The best way to accomplish this is to use a
moving stimulus whose trajectory has no spatiotemporal correlation of its own. This section attempts to summarize my own previous psychophysical experiment and analysis, where the flash-lag effect was found to occur even though the moving item is in completely random motion (Murakami, 2001).

In that experiment, the moving bar’s trajectory was such that the bar was horizontally displaced every 20 frames to a randomly chosen position along the horizontal meridian (within ±30 pixels around the fovea), and it stayed there until the next jump. At a random timing, the flash was briefly presented for one frame at a randomly chosen horizontal position (within ±10 pixels around the fovea). Other details were identical to the present experiment.

Because there is no correlation between successive presentations of the jumping bar, its autocorrelation remains quite flat (at the chance level) except for the perfect correlation at its current presentation (the white rectangle spanning the duration of 20 frames in Figure 7B). As a result, the perfect psychometric function Φ is quite rectangular: the percentage of “right” responses should be 0% on the left of the current jumping bar, 100% on the right of it, and should remain at chance otherwise (Figure 7B). The observed percentage of “right” responses for an actual observer is also plotted in a form of a spatiotemporal correlogram. Each response is plotted at each spatiotemporal position of the flash relative to the spatiotemporal position of the current jumping bar, which is always located at the center of the correlogram (Figure 7A).

The question is how to find the internal kernel $K$ that satisfies the relationship $Ψ = Φ ∗ K$. One could solve this by maximum likelihood estimation of the four free parameters of 2D Gaussian, as was done in the previous section. However, the advantage of the randomness of the jumping bar greatly helps break down the question to a few easier ones. Specifically, the randomness makes spatial and temporal correlation structures orthogonal to each other, which means that the spatial and temporal components of $K$ can be estimated separately.

First, let us consider the spatial component of $K$. This is the only source that brings about the horizontal sigmoid near the center of $Ψ$ because however the temporal component of $K$ may change, it would only vertically distort $Ψ$. Therefore, the spatial component of $K$ can be estimated as the one that satisfies the relationship $∫ Ψ(x, t) dt = ∫ Φ(x, t) dt ∗ ∫ K(x, t) dt$. The temporal summation of $Ψ$ is plotted in Figure 8, as $Ψ(x)$, and the best-fit cumulative Gaussian is superimposed; the perfect psychometric function for it is shown as $Φ(x)$. $K(x)$ is the deconvolution of $Ψ(x)$ and $Φ(x)$, but in this case it can be calculated simply as the first-order derivative of $Ψ(x)$. The result was a Gaussian with parameters $(μ, σ) = (0.042$ pixels, $1.08$ pixels). $μ$ was extremely close to zero, which means that there was no response bias in space. $σ$ was as small as 1 pixel, indicating this observer’s good vernier-acuity performance around perceptual alignment.

Next, let us move on to the temporal component of $K$. Convolution of $Φ(x, t)$ with $K(x)$ estimated above would only produce a little bit of spatial blur at the transition between 0%
and 100%, leaving the temporal structure unchanged. Thus, following the same logic as above, all temporal shift and blur observed in $\Psi(x, t)$ remain to be explained by the temporal component of $K$. The spatial summation of $\Psi$ is plotted as $\Psi(t)$ (with the responses in the negative portion of space flipped and merged to the positive portion, and with data close to the ordinate excluded), and its low-pass-filtered curve is superimposed; the perfect psychometric function for it is shown as $\Phi(t)$. As $K(t)$ is the deconvolution of $\Psi(t)$ and $\Phi(t)$, it was calculated as the division of $\Psi(t)$ by $\Phi(t)$ in Fourier domain. The result of deconvolution is shown by the green curve. Another way to find $K$ is to estimate the best-fit pair of $\mu$ and $\sigma$ by the maximum likelihood method, which was also tested. The result of fit, $(\mu, \sigma) \approx (-7.97 \text{ frames}, 6.29 \text{ frames})$, is overlaid (blue). The two estimated profiles of $K(t)$ were in good agreement with each other. $K$ was found to be biased toward the past, with considerable side lobes in time (Murakami, 2001).

Putting them together, the 2D shape of $K$ can be reconstructed as multiplication of the two orthogonal components (Figure 7C). This shape seems to share several aspects with the estimation in the previous section. First, its peak is located 5-8 frames in the past. Second, temporal uncertainty is so large as to span more than 10-20 frames. Third, spatial uncertainty is within the range of vernier-acuity sensitivity measured in stationary stimuli. However, the spatial offset as found in the previous continuous-motion condition is absent (it should be so; see “Discussion”).

**Effects of the Temporal Predictability of the Flash**

The above analysis of the flash-lag effect in random motion used my previous psychophysical data (Murakami, 2001). This section attempts to apply the same analysis to a new set of data from a slightly modified experiment in which the temporal predictability of the flash was manipulated. In the original experiment, the flash’s presentation timing was unpredictable to the observer because the interval between successive flashes was randomly chosen from the range of 360 ± 120 frames (Murakami, 2001). In the new experiment, the inter-flash interval was systematically manipulated to see its effect on the shape of $K$. The stimulus and procedure were otherwise identical to the original experiment described in the previous section.

The experiment started with the presentation schedule of variable interval, i.e., the inter-flash interval was randomly chosen from the range of 359 ± 120 frames (equal to 2997 ± 1002 ms) excluding the central range of 359 ± 20 frames. After 10 ± 2 flashes were successively presented in this fashion, the schedule was changed so that the inter-flash interval for the next 10 ± 2 flashes was strictly fixed at 359 frames. (The number of repeated flashes within each schedule fluctuated randomly.) These two schedules were alternated seamlessly 20 times within a single experimental session so that the instant of schedule change was least noticeable. No instruction about the schedules was given to naive observers. Importantly, the manipulation of the inter-flash interval only changed the temporal predictability of the flash, leaving the spatial predictability untouched: the flash’s position was still chosen at random.

The responses to flashes were sorted according to the temporal order of the flash relative to the change of schedule, and the spatiotemporal kernel $K(x, t)$ was estimated independently at each phase along the time series around the schedule change. The results exhibited no systematic change in the estimated spatial component of $K$, which was still comparable to the range of vernier-acuity sensitivity, indicating that the overall task difficulty was effectively kept constant across schedules. However, a small but significant ($t$ test, $p < 0.05$) change in the estimated temporal bias $\mu$ was observed, as clearly shown in Figure 9. Specifically, the estimated values of $\mu$ for the fixed-interval (predictable) phases were smaller than those for the variable-interval (unpredictable) phases by roughly one frame, although the absolute baseline of $\mu$ varies across observers. (The decrease in the temporal uncertainty $\sigma$ was also significant for the author’s data but not for other observers.)

The results clearly and more convincingly confirm the previous notion that the amount of flash-lag is reduced for a predictable flash (Brenner & Smeets, 2000; Eagleman & Sejnowski, 2000b). In the previous studies, the predictable flash was such that it was always presented at a known time or at a known position. Either information is mathematically sufficient for uniquely determining the time and position of the flash at physical alignment with the moving object, with its

![Figure 8. Estimation of the spatial and temporal components of the kernel in the random-motion experiment. $\Psi(x, t)$ is identical to Figure 7A. $\Psi(x)$ is its temporal summation. $\Psi(t)$ is the spatial summation, with the percentage of “right” responses in the left half of $\Psi(x, t)$ flipped and merged to those in the right half (data around the ordinate are excluded because they are already distorted by spatial uncertainty). See text for details.](https://jov.arvojournals.org/pdfaccess.ashx?url=/data/journals/jov/932816/)
Responses to the portions of the abscissa. For each relative phase to fixed interval, the error bars indicate 0.1 phase relative to the change of schedule from variable interval, whereas the positive indicates temporally predictable flashes presented after a fixed interval. The flat horizontal lines indicate that the observer did not know the location of the next random position of the abscissa. For each relative phase \( p \), the responses to the \( (p - 1)^n \), \( p^n \), and \( (p + 1)^n \) flashes after the change of schedule were gathered to draw the spatiotemporal correlogram for the phase \( p \), from which the spatiotemporal kernel was estimated independently of other phases.

Figure 9. Effects of the temporal predictability of the flash. The author’s (I.M.) and a naive observer’s (R.M.) data are shown. The estimated temporal bias, \( \mu_t \), is plotted against the flash’s phase relative to the change of schedule from variable interval to fixed interval. The error bars indicate 0.1 \( \times \) \( \sigma_t \). Along the abscissa, the negative (zero inclusive) indicates temporally unpredictable flashes presented after a variable inter-flash interval, whereas the positive indicates temporally predictable flashes presented after a fixed interval. The flat horizontal lines indicate the \( \mu_t \) averaged over each of the negative and positive portions of the abscissa. For each relative phase \( p \), the responses to the \( (p - 1)^n \), \( p^n \), and \( (p + 1)^n \) flashes after the change of schedule were gathered to draw the spatiotemporal correlogram for the phase \( p \), from which the spatiotemporal kernel was estimated independently of other phases.

Trajectory also known. Suppose for example, the flash is to come exactly when the moving bar in constant rotation about the fixation point appears exactly horizontal. Then, an insight-ful observer might allege the flash along the horizontal meridian is the PSE, paying no more attention to the moving bar, but this is not what the nulling method is meant to be! The present experiment, however, escapes this potential problem because even though the flash was temporally predictable, the observer did not know the location of the next random position of the flash relative to the randomly jumping bar. That is to say, knowledge did not help accomplish the task. Instead, temporal predictability did help reduce the perceptual time lag between the flash and the jumping bar.

Possibly, however, “predictability” is not even a correct word. For example, is the first flash after the schedule change predictable or unpredictable? The flash could be predictable because it comes on after the fixed inter-flash interval just as 50% of all flashes do, but could also be unpredictable because the schedule change per se is an unpredictable event. The data in Figure 9 already show a fairly good reduction at first flash, suggesting that the observer relies on something other than the trend of only a few recent flashes. It may be the case that, through a number of repeated trials, the observer has learned to focus on the most probable inter-flash interval of 359 frames, paying less attention to other random intervals. If so, “probabilistic likelihood” may better capture the nature of the determinantal factor.

Discussion

In the flash-lag effect, the impression of illusory spatial offset of the flash relative to the moving item is so vivid that it seems intuitive to quantify its magnitude in spatial terms. However, many researchers have noticed that the effect can also be expressed in terms of time — illusory temporal offset of the flash toward the past. This has sometimes been referred to as “differential latency” (Whitney & Murakami, 1998; Whitney et al., 2000a, 2000b), the “time delay” (Baldo & Klein, 1995), the “temporal lead of moving segment” (Furushothaman et al., 1998), the “difference of delay” (Kirschfeld & Kammer, 1999), and the “equivalent delay” (Krekelberg & Lappe, 1999). As implied in the word “equivalent,” however, these lines of terminology per se only indicate that the spatial PSE and temporal PSE are interchangeable by the relationship distance = time \( \times \) speed. So far, there has been no attempt to identify what is really making the (spatial) psychometric function more gently sloping with increasing speed (see Figure 2A). The present study offers a great improvement of methodology in analyzing the flash-lag effect, whereby all behavior of psychophysical data, including speed dependence of the PSE and slope, is explained as the spatiotemporal bias and uncertainty in our psychological process. Everything is concisely described as the interaction between the perfect performance and the observer’s perceptual spatiotemporal alignment that does not particularly depend on stimulus parameters such as speed.

The kernel estimated from the continuous-motion experiment had a spatial bias of approximately 2 pixels in the direction of motion and a temporal bias of approximately 5 frames past (Figure 5). In particular, the direction of spatial bias happens to be consistent with what the “motion extrapolation” theory proposed: in the retinotopic map, the moving bar’s position is spatially shifted toward the direction of motion; therefore, it is judged as perceptually aligned with the flash presented there. However, the estimated spatial bias is too small to explain the amount of lag in various speed conditions. For example, the lag was as great as 10 pixels for the speed of 2 pixels/frame. Moreover, even in the 0.25 pixels/frame condition, the lag (approximately 3 pixels) was still greater than what the spatial bias of 2 pixels can predict. On the other hand, the amount of the estimated spatial bias is consistent with the y-intercept of Figure 2H, where the spatial PSE is plotted as a function of the moving bar’s speed. Whereas a purely temporal account of the flash-lag effect predicts a linear regression that exactly passes the origin, the actual data lie slightly above this prediction. Thus, the result suggests that somehow the observer tended to judge a slightly (i.e., about half of the width of each rectangular stimulus, see Figure 1) overreached flash as perceptually aligned with the moving bar, irrespective of speed. It is questionable, however, whether this is really one of the general characteristics of the flash-lag effect. Previous studies have sometimes shown similar
positive y-intercepts (Krekelberg & Lappe, 1999), but not always (Nijhawan, 1994; Kirschfeld & Kammer, 1999). Repeatability across observers and across experimental situations is, therefore, open to future investigations.

In contrast, the kernel estimated from the random-motion experiment had no spatial bias toward the direction of motion (Figure 7C). On one hand, this is reasonable because the randomly jumping bar itself was devoid of any coherent direction of motion over frames. On the other hand, the lack of spatial bias is also very reasonable because if there is any effect of motion direction, it is canceled out from the spatiotemporal correlogram shown in Figure 7A. The origin of the correlogram is set at the spatiotemporal position of each instance of the jumping bar; it may be on the way of a leftward jump or a rightward jump with a variety of instantaneous velocities. No matter how the jumping bar is moved, responses are dumbly accumulated upon the same chart. In other words, this correlogram only visualizes the first-order correlation between the jumping bar’s spatiotemporal position and the response at the flash’s spatiotemporal position. Similarly, the analysis in the present study simply focuses on estimating the first-order kernel. However, this limitation does not weaken the importance of the proposed methodology of correlation analysis. If necessary, the current approach could be extended to visualization of higher-order correlation structures (Sutter, 2001).

In both experiments, the kernel was found to have a negative temporal bias and a substantial temporal uncertainty. In perceptual terms, this means that the flash presented a few frames past appears simultaneous with the moving item at present, and that the time lag between the flash and moving item to establish perceptual simultaneity fluctuates substantially. In a separate study, I have found that a numerical simulation using a kernel of similar shape successfully mimics all previous psychophysical data on the flash-lag effect in various other situations (Murakami, 2001). Thus, the observed temporal bias and uncertainty are not a queer product of a mathematical game but a plausible model of a great variety of the phenomenon. (As for its plausibility, however, I would like to note that the assumption of a temporal Gaussian might oversimplify the representation of time because uncertainty in the past and future might actually be asymmetric, so that some skewed probability density function would better reflect the reality.) An especially interesting point is that the same explanation framework applies to the flash-lag effect in random motion as well as the effects observed in constant motion. In this respect, it is conceivable that the flash-lag effect should rather be described as the perceptual time lag of an abruptly flashed object relative to a continuously visible object, whether the latter may be moving constantly or randomly.

The estimated shape of the internal kernel quantitatively differed across experiments. Specifically, the kernel in the random-motion case was thinner in space and taller in time. The reason for the decrease of spatial uncertainty probably comes from the difference in eccentricity. The flash in the first experiment was presented somewhere along a line 2 degrees below fixation, whereas the flash in the second was somewhere along the horizontal meridian. Also, the reason for the increases of temporal uncertainty as well as bias is probably due to the presentation schedule. In the first experiment, the observer could have a rough expectation about timing: the flash was eventually to come when the moving bar reached a position approximately under the fixation point. In contrast, the presentation timing of the flash in the second experiment was random and completely independent of the incessant horizontal jumps of the moving bar. It is likely that these experimental manipulations of bias and uncertainty influence the estimated shape of the kernel. Indeed, the last experiment clearly demonstrated that the temporal bias can significantly decrease with increasing temporal predictability of the flash.

With various manipulations in motion trajectory, such as directional reversal and speed change, several psychophysical studies have traced the spatiotemporal PSE of the flash relative to the moving stimulus, essentially with the same conclusion that the observer somehow compares the flash’s position with the moving item in the future (Whitney & Murakami, 1998; Brenner & Smeets, 2000; Eagleman & Sejnowski, 2000a; Krekelberg & Lappe, 2000; Whitney et al., 2000a, 2000b). The present finding is entirely consistent with this general idea. However, the present finding does not necessarily support or reject those researchers’ hypotheses about the underlying mechanisms of such time lag. For example, a strong statement of the “differential latency” model is that the flash requires a longer neural latency than the moving stimulus (Purushothaman et al., 1998; Whitney & Murakami, 1998). Another idea, “postdiction” framework by Eagleman and Sejnowski, states that the flash’s position is compared with the moving item in the future because the perceived position of the moving item at the present is represented as the positional average of the motion trajectory in the future (Eagleman & Sejnowski, 2000a). It is expected that detailed simulations in line with the present approach resolve which idea is most likely (Murakami, 2001).

Conclusions

The analysis in this study is meant to provide a transparent methodology to visualize flash-lag data as a spatiotemporal correlation structure and to extract the spatiotemporal bias and uncertainty in the visual system that give rise to observed data structure. As a result, it was found (1) that spatial bias is negligible compared to the magnitude of flash-lag, (2) that spatial uncertainty is comparable to that of a stationary vernier acuity task, (3) that temporal bias is negative such that the flash in the past is compared to the moving item at present, and (4) that temporal uncertainty is substantial, meaning that the observer is not very sensitive to perceptual simultaneity.

Finally, this methodological proposal should not be viewed as specific to the domain of the flash-lag effect. The same problem and solution may apply to any psychophysical situation in which the perceived spatiotemporal position of a suddenly presented object is concerned. Such situations may include temporal order judgment (Shimojo, Miyachi, &
Hikosaka, 1997; Allik & Kreegipuu, 1998), perisaccadic mislocalization of flash/world (Matin & Pearce, 1965; Cai, Pouget, Schlag-Rey, & Schlag, 1997), the paradigm of rapid serial visual presentation (c.f., Shapiro, Arnell, & Raymond, 1997), etc. In cases where the horizontal as well as vertical position of the briefly flashed object is in question, it would probably be necessary to consider a three-dimensional psychometric function, \( \Psi(x, y, t) \).

Acknowledgments

I thank Shin’ya Nishida for valuable comments on a preliminary version of the manuscript, and Kenichirou Ishii and Tatsuya Hirahara of NTT Communication Science Laboratories for their administrative support. Commercial Relationships: None.

References