25

Bridging the gap: a model of common neural mechanisms underlying the Fröhlich effect, the flash-lag effect, and the representational momentum effect

DIRK JANCKE AND WOLFRAM ERLHAGEN

Summary

In recent years, the study and interpretation of mislocalization phenomena observed with moving objects have caused an intense debate about the processing mechanisms underlying the encoding of position. We use a neurophysiologically plausible recurrent network model to explain visual illusions that occur at the start, midposition, and end of motion trajectories known as the Fröhlich, the flash-lag, and the representational momentum effect, respectively. The model implements the idea that trajectories are internally represented by a traveling activity wave in position space, which is essentially shaped by local feedback loops within pools of neurons. We first use experimentally observed trajectory representations in the primary visual cortex of cat to adjust the spatial ranges of lateral interactions in the model. We then show that the readout of the activity profile at adequate points in time during the build-up, midphase, and decay of the wave qualitatively and quantitatively explain the known dependence of the mislocalization errors on stimulus attributes such as contrast and speed. We conclude that cooperative mechanisms within the network may be responsible for the three illusions, with a possible intervention of top-down influences that modulate the efficacy of the lateral interactions.

25.1 Introduction

Localizing an object in the presence of motion is a fundamental ability for many species as a moving object often represents danger or food. In recent years, advances in neurophysiology and psychophysics have substantially increased our understanding of how the visual system calculates the present and future positions of moving objects. New insights have been gained in the last couple of years by analyzing systematic mislocalization errors occurring at different points along the motion trajectory. When observers are asked to localize the initial or final position of a moving stimulus they typically judge the position as shifted forward in the direction of motion. These errors are known as the Fröhlich effect (Fröhlich 1923) and the representational momentum effect (Freyd & Finke 1984; Hubbard & Bharucha 1988), respectively. Another well-established form of mislocalization is the flash-lag effect (Metzger 1932; Nijhawan 1994). When a stimulus is briefly flashed in physical alignment with a continuously visible moving object, observers nonetheless perceive the moving object ahead of the flash. In all three illusions, the position percept does not agree with the physical measurements, indicating that the mechanisms underlying the assignment of an object's location go beyond a mere spatiotemporal filtering of retinal information. Multiple hypotheses pertaining to these illusions have been offered in the past several years. Possible explanations include among others attentional mechanisms (Baldo & Klein 1995; Müsseler & Aschersleben 1998; Kirschfeld & Kammer 1999), active trajectory extrapolation (Nijhawan 1994), differential latencies (Whitney, Murakami, et al. 2000; Maiche et al. 2007), some form of temporal averaging (Eagleman & Sejnowski 2000; Krekelberg & Lappe 2000), or a memory shift of a high-level representation (Hubbard & Bharucha 1988). The lack of agreement on the underlying processing principles may in part be explained by the fact that most of the experimental studies focus exclusively on one illusion. Yet new insights may be gained by elucidating the relationship between the different illusory displacements (Müsseler et al. 2002).

The main purpose of this chapter is to present a network model based on well-known neuronal mechanisms. Within the framework of the model we summarize some existing experimental data and discuss the potential commonalities among the three mislocalization errors. The proposed model consists of a network of excitatory and inhibitory neural populations that encode stimulus position. It implements the fundamental idea that local cortical feedback plays a dominant role in shaping the population representation of a motion trajectory. Internal cooperative mechanisms are in general beneficial for the visual system because they allow, for instance, one to cope with noisy or missing afferent information (Douglas et al. 1995). However, as we shall argue here, they generate in some instances a possible substrate for illusory percepts.

In response to an apparent motion display, the network exhibits a stimulus-locked traveling wave of activity. Lateral interactions mediated by excitatory connections result in a preactivation of neurons encoding future positions. The moving object is thus processed more efficiently compared to a flashed object as in a flash-lag display, resulting in a differential processing delay. We have recently reported neural trajectory representation in the primary visual cortex of cat (Jancke et al. 2004b) that showed such a path-dependent facilitation. If the cooperative mechanisms within the network are sufficiently strong, the dynamic transformations sustain for some time upon stimulus offset (Erlhagen & Jancke 2004). As a result, neurons encoding positions displaced forward in the direction of motion become active. The population response thus represents a possible neural substrate for a stimulus position that is perceived but not sensed directly.

To test the idea that recurrent interactions constitute a common low-level mechanism for explaining the illusions, we proceed as follows. We first use the experimentally observed traveling waves in cat primary visual cortex to calibrate the model parameters that represent the spatial ranges of the lateral interactions. Our working hypothesis for discussing the three mislocalization phenomena is that spatiotemporal characteristics of the build-up, middle, and decay phase of the wave are related to the Fröhlich effect, the flash-lag effect, and the representational momentum effect, respectively. We test this hypothesis by directly comparing model predictions and experimental findings when stimulus attributes such as contrast and speed are systematically varied. There is experimental evidence that changes in

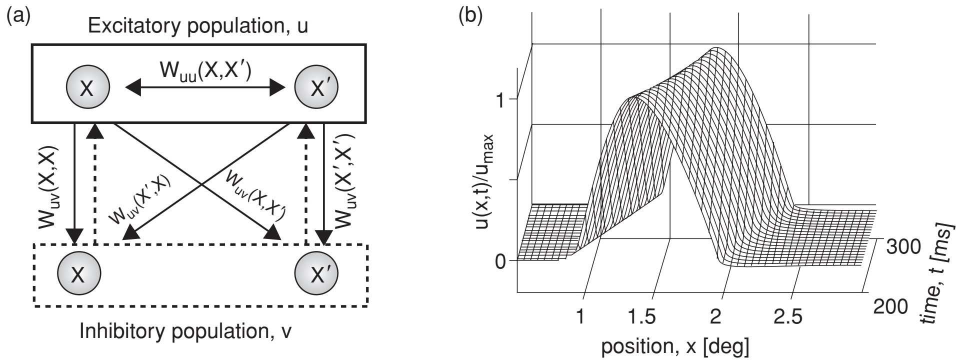

Fig. 25.1 (a) Schematic representation of the connectivity between any two nodes x and x' of the model network consisting of an excitatory (u) and an inhibitory (v) subpopulation of neurons coding for position. For details see the text. (b) Traveling wave in response to a stimulus of width $2\sigma_S = 0.4\,\text{deg}$ and intensity $A_S = 13.2$ moving with the apparent speed $v = \Delta x/\Delta t = 40\,\text{deg}/s$ (spatial displacement $\Delta x = 0.4\,\text{deg}$, frame duration $\Delta t = 10$ ms). Model parameters were: $\sigma_u = 0.3\,\text{deg}$, $\sigma_v = 0.4\,\text{deg}$, $A_u = 4.65$, $A_v = 3.99$, $\beta = 1$, $u_f = u_g = 0$, $h = -3$. To adjust the spatial scale in the model to the experimental units we have chosen 10 pixel = 0.2 deg.

the task demands may alter the magnitude of mislocalization errors (Müsseler et al. 2002; see also the discussion in Kreegipuu & Allik 2003), suggesting that the effects cannot be explained by a feedforward architecture alone. We propose and test a neuroplausible mechanism that allows altering the efficacy of the recurrent interactions based on top-down influences.

25.2 The dynamic model

25.2.1 Model architecture

There are two main hypotheses about the neural mechanisms underlying the processing of stimulus position that have guided our modeling work (for a detailed discussion see Erlhagen et al. 1999; Jancke et al. 1999). First, information about stimulus position is encoded in visual brain areas by the distributed activity pattern of large neural populations, rather than by single neurons. Each cell of a population is tuned to a specific position in visual space, and its level of activity defines the extent to which the information is present. A second hypothesis concerns the role of cortical interactions in shaping the response properties of neural populations. The fact that the largest input to cortical cells comes from neighboring cells rather than from feedforward afferents suggests that massive excitatory feedback counterbalanced by cortical inhibition plays a central role for the processing of stimulus attributes such as position (Douglas et al. 1995).

In the model, neurons split into an excitatory and an inhibitory subpopulation. They are organized as layers of neurons densely covering the stimulus dimension. Because we focus on experiments with horizontal motion displays, the model is one-dimensional.

The connection schema for any two neurons of the network is sketched in Fig. 25.1(a). An excitatory neuron tuned to position x integrates activity from neighboring excitatory neurons via lateral connections and projects to neurons of the inhibitory pool with similar and dissimilar receptive field centers *x* and *x'*, respectively. This model architecture is in line with anatomical and physiological work demonstrating that direct reciprocal connections exist primarily between pairs of excitatory pyramidal cells. Axon collaterals of these cells, on the other hand, may also target inhibitory interneurons (Gilbert 1995). Each inhibitory neuron *x* spatially summates incoming activation from excitatory neurons. For simplicity, we assume that it feeds its activation back to the excitatory pool only locally, that is, to a neuron with the same receptive field center x. Note that the implementation of spatially extended inhibitory feedback would not change the qualitative conclusions of our work.

Like many other models of cortical function (for an overview see Dayan & Abbott 2001), cortical interactions is assumed to depend on the functional distance of the cells and defined by the feature coded by them. The model parameters describing the recurrent interactions are adjusted such that neurons with similar receptive field centers excite each other, whereas inhibition dominates for larger distances. This interaction pattern, known as "Mexican-hat" organization, guarantees a sharply tuned excitation profile as a network response to an afferent input that carries information about the visual location of a stimulus.

Assuming that the number of excitatory and inhibitory neurons is large and that their receptive field centers densely cover the visual field, the mean activity at time *t* of an excitatory neuron and an inhibitory neuron tuned to horizontal position *x* can be described by two continuous functions $u(x, t)$ and $v(x, t)$, respectively. To model the dynamics of the neural population, we use the model class of neural fields first introduced and analyzed by Wilson and Cowan (1973). Neural field models are system-level models adequate to describe the mean activity of large populations of neurons without referring to a detailed level of physiological realism. The following differential equations (Jancke et al. 1999; Erlhagen & Jancke 2004) govern the evolution of the activation variables u and v:

$$\begin{aligned}
\tau \frac{d}{dt} u(x,t) &= -u(x,t) + h + S(x,t) + g(u(x,t)) \\
&\qquad \times \left[\int w_{uu}(x-x') f(u(x',t)) dx' - v(x,t) \right] \\
\tau \frac{d}{dt} v(x,t) &= -v(x,t) + \int w_{uv}(x-x') f(u(x',t)) dx'
\end{aligned}$$

where du/dt and dv/dt represent the changes in mean activity over time. The parameter $\tau > 0$ is used to adjust the time scale of the field dynamics to the experimentally observed time scale. The afferent input $S(x, t)$ to the excitatory population is modeled as a Gaussian profile. Its space constant, σ_s, reflects the half-width of the stimulation, whereas the amplitude, A_S, is assumed to change as a function of the luminance contrast of the external stimulus. The constant h defines the resting level to which the population activity relaxes without external input. The integral terms describe the spatial summation of excitation in the two layers. The spatial interactions fall off with increasing distance between field sites:

$$w_{ui}(x-x') = A_i \exp\left(-\frac{(x-x')^2}{2\sigma_i^2}\right), (i = u, v),$$

P1: JZP Trim: 174mm × 247mm Top: 0.581in Gutter: 0.747in
CUUK929-25 cuuk929/Nijhawan ISBN: 978 0 521 86318 6 October 23, 2009 23:43

where the choice of the relative amplitudes and spatial ranges, $A_u > A_v$ and $\sigma_u < \sigma_v$, implements the Mexican-hat pattern. Only sufficiently activated neurons contribute to the interaction. The nonlinear activation function $f(u)$, which gives the mean firing rate for a given level of activation, is taken as a monotonic function of sigmoid shape going from 0 to 1:

$$f(u) = \frac{1}{1 + \exp(-\beta(u - u_f))}.$$

The parameter u_f determines the position of the maximum slope of the function f, and β controls the value of the maximum slope.

The field dynamics exhibit a threshold behavior. Starting from the stable resting state h, only a sufficiently strong afferent input is able to drive excitation to a level that triggers the self-stabilizing forces within the network. The activation in the excitatory layer is normalized relative to the threshold excitation level $u_{TH} = 0$. The resting state is thus chosen to be negative ($h < 0$), and consequently negative u-values describe subthreshold activity.

An increasing body of experimental evidence suggests that cognitive factors like attention or task demands can alter the efficacy of the lateral connectivity in primary visual brain areas (Li et al. 2004). There is certainly a need for neural mechanisms that allow for flexible visual processing without referring to learning-based synaptic reorganization that is believed to take place on a longer time scale. The neural sources for such top-down influences are not known. A number of architectures have been proposed that, in principle, enable changing the functional properties of recurrent networks in an efficient way (Hahnloser et al. 1999). We have implemented a simple shunting mechanism (for an overview see Grossberg 1988) that can alter the gain of the network response to an afferent stimulus. The recurrent interaction of the u-layer is gated by a nonlinear function $g(u)$ that is also of sigmoid type with parameters u_g and β. We simulate a top-down influence by gradually changing the position of the maximum slope u_g. This allows us to control in a flexible manner the spatial spread of the suprathreshold population response and thus the extrapolation properties of the network (Erlhagen 2003).

25.2.2 *Choice of model parameters*

We study how the neuronal population response, which is shaped by the cooperative mechanisms within the network, interacts with an external input representing a stimulus in motion. When a brief, localized input of adequate intensity is applied, the network develops a localized activity pattern in position space known as an “active transient” (Wilson & Cowan 1973). Due to the recurrent excitation, the activity continues to grow in amplitude and width upon stimulus offset. It reaches a peak value and then decays back to resting level driven by the increasing feedback inhibition. If a flashed stimulus is part of a motion paradigm, the population representation of that stimulus will interact through the lateral connections with representations of preceding and succeeding stimulus frames. As shown in Fig. 25.1(b), the delicate interplay between the excitatory and inhibitory feedback loops may result in an activity wave that propagates with the velocity of the inducing stimulus in the direction of motion. This locking to the stimulus occurs for a whole range of velocities (Ben-Yishai et al. 1997; Giese & Xie 2002), thus permitting the study of the effect of velocity changes on the spatiotemporal characteristics of the wave within a single population. The velocity range is determined by the interaction structure of the network. It covers the speed of a spontaneous wave that evolves in the absence of a time-dependent input for a sufficiently low threshold u_g of the gating mechanism (Erlhagen 2003). Beyond the characteristic range, the wave loses stability and a population response with a quasi-periodic amplitude modulation emerges, suggesting that multiple neural networks with different speed selectivity might exist.

We use the neural trajectory representations, which we have recently described in the primary visual cortex of cat, to adjust the model parameters that define the interaction processes. The velocity range tested in the experiments (4–40 deg/sec) constrains the spatial ranges for excitation and inhibition, $\sigma_u < \sigma_v$. The amplitude ratio $A_u/A_v < 1$ is adjusted to reproduce the mean amplitude and width of the neural waves. It is worth mentioning that the whole set of model parameters used for the present study is in agreement with our previous modeling of neural interaction effects probed with stationary displays (Jancke et al. 1999).

The link to the perceptual mislocalization errors is made by taking the peak position of the traveling wave as an estimate of stimulus position. Our working hypothesis is that the wave model fitted to reproduce the neural data recorded in the visual cortex will allow us to test whether such a self-stabilized representation may qualitatively explain the three visual illusions. To also reproduce the magnitude of the mislocalization errors observed in the psychophysical experiments we make two adjustments to the model parameters. A slightly larger amplitude ratio $A_u/A_v < 1$ is used. This favors the recurrent excitation over the inhibition and results in a stronger facilitation effect along the motion trajectory. A second adjustment concerns the time scale of the field dynamics. It is fixed to $\tau = 15$ ms for the population data, which is in the range of the membrane time constant of a neuron (Abeles 1991). To cover also a perceptually relevant time scale we adapt the value to $\tau = 35$ ms for the modeling of the localization errors. This ensures that the duration of the active transient response to a brief afferent input reflects the persistence of a flashed stimulus in the visual system (100–150 ms, Coltheart 1980). The identical set of field parameters is used for the simulation of the three visual illusions. Only the gating threshold u_g is modified in some simulations to model assumed top-down influences on the efficacy of the recurrent interactions. The stimulus dimension and the spatiotemporal properties of the linear motion displays (frame duration Δt, spatial displacement Δx) are adapted to meet in each case, as close as possible, the conditions of the psychophysical experiments.

25.3 Modeling results

25.3.1 Motion trajectories in primary visual cortex

To address how motion trajectories are represented at the level of primary visual cortex we recently used a population approach that pools spiking activity of many neurons in cat area 17 (Jancke et al. 2004b). Following our earlier study with stationary displays, our working hypothesis was that the observed nonlinear interaction effects between neuronal representations of adjacent stimuli should manifest also when the afferent input is time-dependent. In those experiments we used small squares of light (0.4 deg) that moved with different velocities ranging from 4 to 40 deg/s along a horizontal line. We employed an optimal linear estimator technique that allows reconstructing with high spatial and temporal resolution the stimulus position from a pool of broadly tuned neurons (Salinas & Abbott 1994; Erlhagen et al. 1999). The general idea behind the construction of population distributions is that the activity of each cell in each trial is treated as a vote for its tuning curve (or in case of an optimal estimator a so-called basis function obtained by a template-matching procedure). The summation of all votes weighted by the firing rate gives the population response in parametric space. Despite the fact that adjacent stimuli activate highly overlapping cell populations, the reconstructed motion trajectories reveal that the peak of the population response represents well even small changes in stimulus position.

Two examples of neuronal trajectory representations that differ in velocity and direction are shown in a space–time diagram in Fig. 25.2(a). One striking characteristic of these traveling activity waves is that the mean activation level increases with stimulus velocity. As shown in Fig. 25.2(b), the network model with adequately adjusted spatial ranges of lateral excitation and inhibition can explain this finding. In the model, the spread of excitation is followed by a wave of inhibition that reduces the amplitude as well as the duration of the population response. This suppressive effect increases with lower speed. Moreover, the model predicts, in line with the experimental findings, that the response to the stimulus train with the highest tested velocity reaches approximately the mean activation level of the response to the stimulus flashed in isolation. A second characteristic of the neuronal trajectory representations captured by the dynamic model is that the localized activity profile locks to the stimulus with a speed-dependent spatial lag. To allow for a direct comparison of modeling and experimental results (Fig. 25.2(c)), a constant time interval of 25 ms was added before stimulus onset in all simulations. This time window represents the average temporal delay between the stimulus presentation and the onset of the population response in primary visual cortex (Jancke et al. 1999). There are three findings of particular interest for the discussion of the localization errors. First, a near compensation of processing delays for the slowest speed can be observed. Second, for the range tested the speed dependence of the spatial lag is roughly linear. However, the most important finding is that the peak latency of the population response to the stimulus in motion was shorter by about 16 ms compared to the response when the stimulus was flashed in isolation (Jancke et al. 2004b). The observed differential latency in primary visual cortex gives direct physiological support for explanations of visual illusions that stress the importance of the time at which an object is perceived (for a review see Whitney 2002). It is important to note that this latency reduction is observed even for the stimulus train with highest velocity. Here, the individual stimulus frames do not significantly overlap. This excludes a simple explanation of the differential latency based on the fact that the spatially extended moving stimulus may start to trigger retinal cells earlier than the flash.

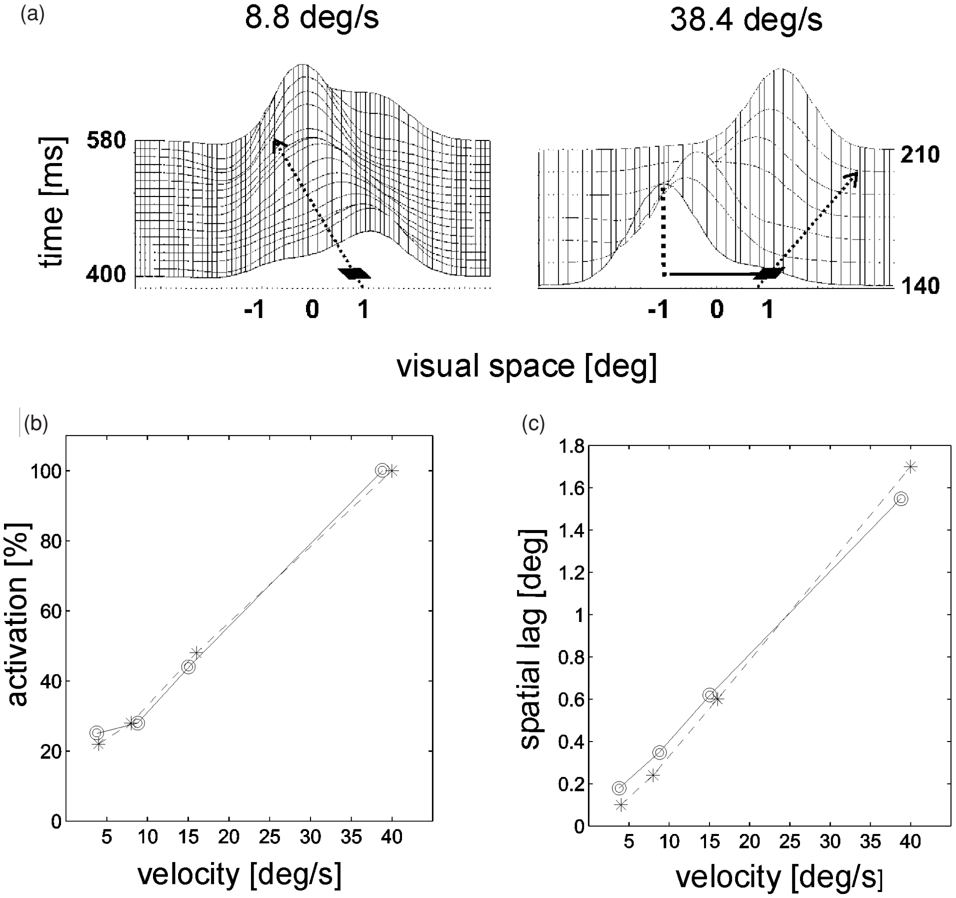

Fig. 25.2 (a) Two examples of cortical trajectory representations: a square stimulus of width 0.4 deg moved with speed 8.8 deg/sec (left) and 38.4 deg/sec (right). The reconstructions of the traveling waves are based on the activity of 178 neurons recorded in the central visual field representation of cat area 17. (b) Comparison of model predictions and experimental findings: dependence of the mean activity level of the wave on stimulus speed. The value averaged over both directions tested in the experiments is shown for the model simulations (asterisk) and the neuronal reconstructions (circle). Different movement velocities were induced by using identical temporal presentation rates $\Delta t = 5$ ms and adequately adapted spatial displacements. Model parameters were like in Fig. 25.1 except $A_u = 4.52$, $A_v = 4.39$, $u_f < u_g = 3$. (c) Dependence of the spatial lag on stimulus speed. The mean spatial displacement between the current stimulus position and the peak position of the traveling wave is plotted as a function of speed for the model simulations (asterisks) and the neuronal representations (circle).

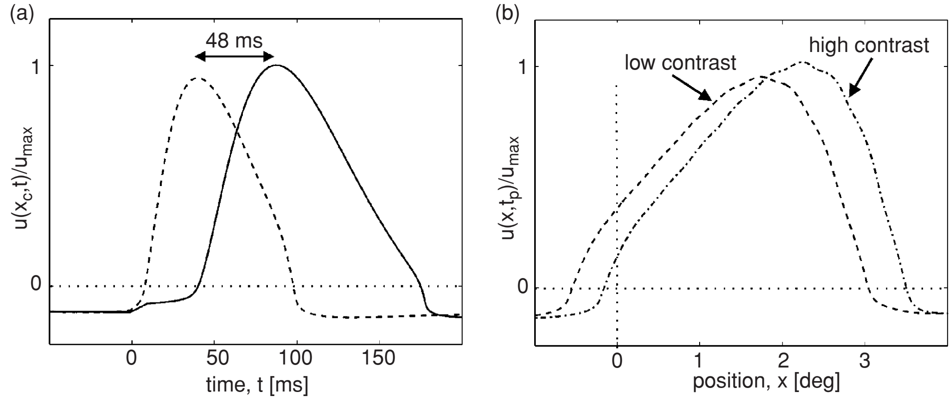

Fig. 25.3 (a) Comparison of the network response to a flashed stimulus ($\Delta t = 10$ ms, $\sigma_S = 0.2$ deg, $A_S = 6.6$) when it is presented at time $t = 0$ in isolation (solid line) or as part of a motion display ($\Delta x = 0.4$ deg, $\Delta t = 10$ ms, dashed line). The time course of the maximal exited neuron at the center position x_c is shown. The activity level $u_{TH} = 0$ indicates the threshold for triggering the active transient response. The same set of model parameters as in Fig. 25.1 was used. (b) Dependence of the flash-lag illusion on the contrast of the moving stimulus. The position of the traveling wave at the time t_p of the peak response to the stationary flash is compared with the wave position when the stimulus intensity A_s of the moving stimulus was increased by a factor of 2, $A_S = 13.2$. The wave appears further ahead of the flash position $x_c = 0$ for the "high contrast" (HC) compared to the "low contrast" (LC) stimulus.

25.3.2 *The flash-lag effect*

The flash-lag effect (FLE) describes a visual illusion wherein a moving object is perceived as being ahead of a stationary flashed object when the two retinal images are physically aligned. It was first discovered some 80 years ago and has often been explained as being due to differential perceptual latencies for the flashed and the moving stimulus (Metzger 1932; Purushothaman et al. 1998; Whitney, Murakami, et al. 2000; Maiche et al. 2007). If the time to perception for the flash were longer, the moving object should appear ahead of the flash position. Taking the peak latency of the flash response as a time maker, the findings of our population study suggest a neural correlate for this explanation because the peak of the traveling wave has already passed the flash position. However, the observed latency difference of 16 ms is not sufficient to explain the full range of flash-lag effects for the majority of experimental studies, which find differences in the range of 45–80 ms (Krekelberg & Lappe 2001). We therefore explored in the model simulations of the impact of a larger amplitude ratio $A_u/A_v < 1$ of the excitation relative to the inhibition on the path-dependent facilitation. An important constraint for the modeling comes from the finding that no significant mislocalization further ahead in motion direction occurs when with the disappearance of the flash the moving object also disappears (Whitney, Murakami, et al. 2000; but see Fu et al. 2001). This means that the cooperative mechanisms within the network should not be strong enough to sustain the wave beyond the vanishing position. In Fig. 25.3(a) we compare the time course of the maximum excited neuron in response to the stimulus flashed in isolation (solid line) and when it is part of a linear motion paradigm (dashed line). Stimulus onset is for both cases t = 0 ms. Due to the spread of subthreshold activity from preceding frames, the stimulus in motion triggers immediately the suprathreshold response, whereas it takes much longer when the stimulus is flashed without spatiotemporal context. Because the network inhibition following the excitation also starts earlier, the peak latency appears to be reduced by 48 ms in this example. This value is in agreement with the 45 ms inferred by Whitney and colleagues from the findings of their linear motion paradigm. Consistent with the differential latency hypothesis, the flash-lag effect depends systematically on the luminance of the object (Purushothaman et al. 1998). The visual latency for a given stimulus is believed to vary inversely with luminance (Lennie 1981). Consequently, the flash-lag effect is predicted to increase if the luminance of the moving object but not that of the flash is increased. Reducing instead the flash latency by applying a higher flash contrast decreases this lead. The network model captures the latency dependence on contrast because stronger afferent inputs reach the threshold u_{TH} for the self-stabilized population response earlier. As depicted in Fig. 25.3(b), the traveling wave model is qualitatively in line with these experimental observations concerning the flash-lag effect. The snapshot of the wave at the peak time of the flash response is further ahead of the flash position for the stimulus with higher intensity A_s.

A latency advantage of about 50 ms represents the maximal value that can be achieved with the present parameter settings. It is important to stress that this advantage is the result of cooperative mechanisms within a single stage of cortical visual processing. Taking the afferent pathway from retina to cortex into account may allow explaining even larger flash-lag effects. Mechanisms like contrast gain control in populations of retinal cells (Berry et al. 1999) or a structured summation over excitatory feedforward input from hierarchically lower processing stages (Baldo & Caticha 2005) have been proposed as explanations for the advance of the moving stimulus over the flash. Although the largest input to cortical cells comes from neighboring cells at the same stage of processing, afferent input may contribute to the preactivation of neurons encoding future stimulus positions, and thus to the path-dependent facilitation. The observation that the flash-lag may turn into a flash-lead for high-contrast flashes (Purushothaman et al. 1998) requires the assumption of differential delays from retina to cortex (Baldo & Caticha 2005). In the present simulations the intensity parameter, A_S, was varied but not the onset of the afferent input, $S(x, t)$, to the excitatory population.

A second stimulus parameter that has been systematically analyzed in experimental studies is stimulus speed. A linear increase of the flash-lag illusion with speed has been reported (Nijhawan 1994; Whitney, Murakami, et al. 2000; for a review see Krekelberg & Lappe 2001). The wave model provides a natural explanation for this finding because the time interval that elapses before the population peak reaches stimulus position is approximately constant for all velocities (compare Fig. 25.2(b)).

25.3.3 The Fröhlich effect

Fröhlich (1923) was the first who studied systematically the phenomenon that a slit of light moving on a track is not seen immediately after it emerges from behind a screen

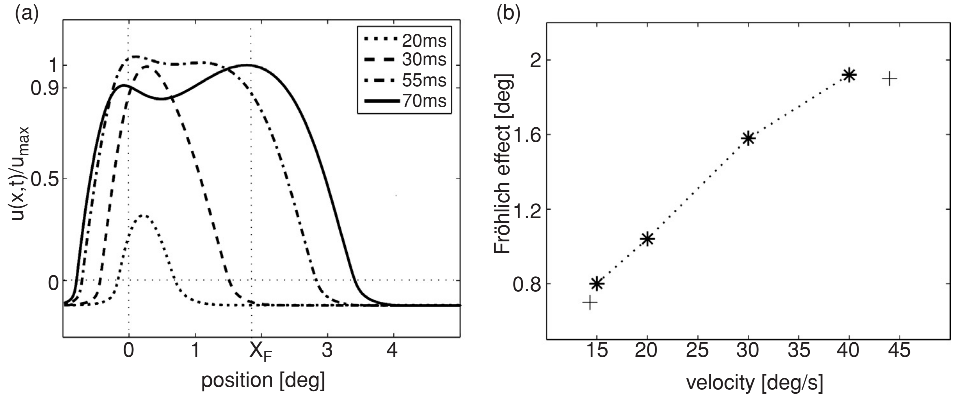

Fig. 25.4 (a) Four snapshots of the buildup phase of the wave used in Fig. 25.3 are plotted. The position x_F indicating the start of a coherently propagating activity pattern is taken as a correlate for the Fröhlich effect. (b) Comparison of model predictions (asterisk) and experimental findings (plus) of the Fröhlich effect as a function of speed. The experimental data were estimated from Fig. 2 in Müsseler & Aschersleben (1998). Model parameters were like in Figs. 1 and 3, stimulus parameters were $\sigma_S = 0.25$ deg, $A_S = 13.2$. Different movement velocities were induced by using a fixed temporal presentations rate $\Delta t = 3$ ms and adequately adapted spatial displacements.

but only after passing through a certain distance. Recently, Müsseler and Aschersleben (1998) established the existence of this illusory displacement of the onset position with a computer-generated motion paradigm. Early temporal accounts to explain this effect were based on the idea that during the time it takes to perceive the object ("sensation time") it has already moved a certain distance (for a detailed discussion see Müsseler & Aschersleben 1998). In the network model, the build-up of the traveling wave is not immediate and occurs at positions subsequent to the object's initial position. The straightforward idea behind the modeling work is thus to qualitatively and quantitatively discuss the when and where of the wave evolution in relation to the Fröhlich effect.

Figure 25.4(a) shows four snapshots describing the build-up of the population response to the motion display used to discuss the flash-lag illusion (see Fig. 25.3). The activity pattern first crosses the threshold for the self-sustaining mechanisms at a position that is shifted forward in motion direction relative to the starting position $x = 0$. Due to the local excitatory loops within the network the activity continues to grow both in amplitude and width. A maximum activation level is reached when the local inhibition starts to dominate the processing. Subsequently, an activity peak evolves that starts to follow the stimulus with the respective speed. Note that it still takes about 30 ms until the balance of excitation and inhibition stabilizes the final shape of the traveling wave. It is important to stress that only at the point in time of the peak formation the spatiotemporal activity pattern carries in a reliable manner the information about a stimulus in motion. As our modeling and neurophysiological data show (Jancke & Erlhagen, unpublished), the build-up within the first 70 ms time window resembles the population response to a briefly flashed, spatially extended bar.

P1: JZP Trim: 174mm × 247mm Top: 0.581in Gutter: 0.747in
CUUK929-25 cuuk929/Nijhawan ISBN: 978 0 521 86318 6 October 23, 2009 23:43

To directly compare model predictions with experimental findings, we read out the peak position at the time when the "tail" of the activity profile has decayed to about 90% of its maximal activation. The dependence of this reference value x_F on stimulus parameters such as contrast and speed qualitatively reflects the experimental observations. Fröhlich (1923, p. 73) pointed out that increasing stimulus luminance reduces the illusory displacement at movement onset. Consistently, position x_F comes closer to the starting position with increasing intensity A_s, indicating an earlier onset of the trajectory representation. We have adjusted this parameter to also quantitatively reproduce the dependence of the Fröhlich effect on speed. In the horizontal motion paradigm of Müsseler and Aschersleben (1998), the stimulus was moved at two different velocities, 14.3 deg/sec and 44 deg/sec, in the two directions. In Fig. 25.4(b) we compare the magnitude of the experimentally observed displacement averaged over both directions (plus) with model predictions for four velocities in the range between 14.3 deg/sec and 44 deg/sec (asterisk). The modeling results reproduce well the observed increase of the effect with speed.

A robust mislocalization in motion direction has been also reported when the first frame of the motion display appears simultaneously with a flash. Because in some studies of this so-called flash-initiated cycle (Khurana & Nijhawan 1995) the shift in onset position appeared to be comparable in magnitude to that of the standard flash-lag illusion, it has been argued that the same processing mechanisms might cause the two localization errors (Khurana & Nijhawan 1995; Eagleman & Sejnowski 2000; see also Nijhawan et al. 2004). The differential latency account would again suggest the flash as a time marker for reading out the population activity in response to the moving stimulus. However, Eagleman and Sejnowski (2000) convincingly showed that presenting the flash 50 ms before the onset of the motion does not affect the error. The authors propose instead that the flash "resets" motion integration. Consequently, like at motion onset an internal position representation (e.g., a traveling wave) has first to be built up. On this view, the flash-lag illusion is just a variant of the Fröhlich effect. Other studies, however, that also investigated directly the relation of the two localization errors (albeit using higher speeds) reported a significant difference in magnitudes (Müsseler et al. 2002; Kreegipuu & Allik 2003), suggesting different underlying mechanisms. The network model makes a clear prediction to decide this open question. With increasing stimulus contrast, the traveling wave starts earlier in space and time. Simultaneously, the wave appears further ahead of the population representation of a physically aligned flash (compare Fig. 25.3(b)). The model thus predicts the opposite effects of changes in stimulus contrast on the Fröhlich and flash-lag illusions.

25.3.4 The representational momentum effect

When observers are asked to remember the final position of an object in motion they typically misremember it as further along the implied trajectory. Analogous to Newton's first law of motion this form of motion extrapolation has been termed "representational momentum" (Freyd & Finke 1984). The momentum metaphor refers to the notion that like a physical object the dynamic internal representation of position cannot be halted immediately

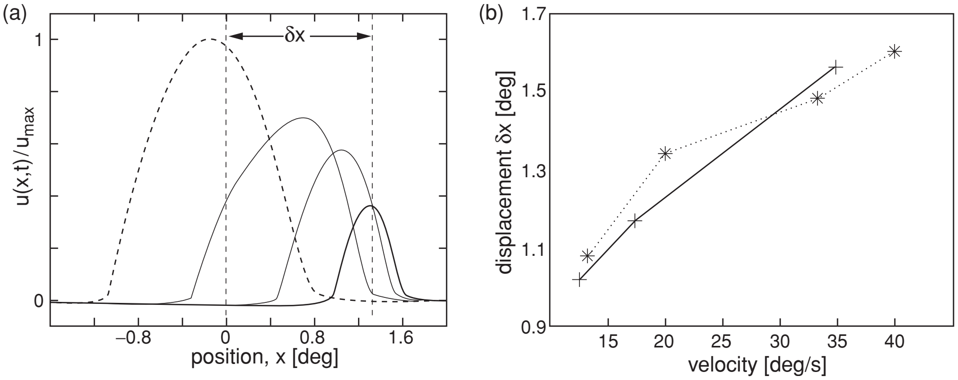

Fig. 25.5 (a) Extrapolation to positions that were not physically stimulated. Four snapshots of the wave at the end of the motion trajectory are shown. The leftmost profile (dashed line) represents the wave at the time of stimulus offset. The forward displacement ∂x is defined as the spatial distance between the actual vanishing position $x = 0$ and the peak position of the wave when it stops to travel (rightmost profile). Model parameters were like in Figs. 25.1 and 25.3 except a lower gating threshold $u_f = 0 > u_g = -0.25$. (b) Comparison of model predictions (asterisk) and experimental findings (circle) of the representational momentum as a function of speed. The experimental data was estimated from Fig. 1 in Hubbard and Bharucha (1988) and Fig. 1 in Hubbard (1990). Model parameters were like in Figs. 25.1 and 25.3, stimulus parameters were $\sigma_S = 0.45$ deg, $A_S = 10$. Different movement velocities were induced by using a fixed frame duration $\Delta t = 3$ ms and adequately adapted spatial displacements.

upon stimulus offset. The effect has been originally found in implied motion paradigms but has been later replicated with continuous motion displays (Hubbard & Bharucha 1988; Hubbard 1990).

Depending on the efficacy of the lateral interactions, the traveling wave may overshoot the vanishing position. The population response thus suggests the presence of a moving object at positions that were never physically occupied. For the simulation example shown in Fig. 25.5 we used the identical set of model parameters as for the simulation of the flash-lag illusion (Fig. 25.3) but reduced the gating threshold u_g for the lateral interactions. As depicted by the snapshots in Fig. 25.5(a), the wave still lags behind stimulus position at the time of stimulus cessation (dashed line). The spatial lag is, however, smaller compared to the simulation shown in Fig. 25.3, giving further support for the notion that lateral interactions are an efficient means to compensate for processing delays. Upon stimulus offset, the wave loses speed and amplitude. Finally, the activity profile stops to travel and decays back to resting level. We use the peak position representing the offset of the continuous propagation as a correlate for the representational momentum. The forward displacement $\delta x > 0$ is defined as the difference between that peak position and the actual vanishing position.

In Fig. 25.5(b) we compare model predictions with the experimental findings reported by Hubbard and Bharucha (1988) and Hubbard (1990) using a horizontal motion display. The experimental data (plus) for the three speeds, 12.5, 17.4 and 34.8 deg/sec, represent

P1: JZP Trim: 174mm × 247mm Top: 0.581in Gutter: 0.747in
CUUK929-25 cuuk929/Nijhawan ISBN: 978 0 521 86318 6 October 23, 2009 23:43

the average value of the observed forward displacement for the four tested subjects. The modeling results (asterisk) for speeds within this range reproduce well the order of magnitude of the effect and the increase of the forward displacement with speed. Note that for the set of parameters used, the range of stimulus velocities that leads to a smoothly traveling wave is between 10 and 40 deg/sec. Beyond this range, the amplitude oscillations of the population response make the overshoot dependent on the exact vanishing position.

There is evidence that stimulus attributes such as contrast or shape (e.g., blurred edges) may affect the overshoot of the vanishing position of a moving target (Fu et al. 2001). Also the population response of the model network shows this dependence on stimulus parameters. A stronger afferent stimulation, for instance, leads to a larger forward displacement. However, the increase is rather small ($< \sigma_s$) and cannot explain alone the order of magnitude of the momentum effect.

In the model simulations, a change in efficacy of the lateral interactions by adapting the gating threshold allows us to explain the seemingly conflicting psychophysical data about position judgment at the vanishing position. Under the condition of the flash-lag paradigm no overshooting is observed. The only, but important, difference to the representational momentum paradigm is the judgment relative to an accompanying flash. In a study designed to directly compare the flash-lag illusion and the representational momentum effect, Müsseler and colleagues (Müsseler et al. 2002) found evidence for some kind of cognitive control over the position judgment. They showed that a verbal manipulation of the task relevance of the accompanying flash ("ignore the flash") may gradually change the pattern of mislocalization at the end of the motion trajectory. When trying to compare the position of the moving stimulus with an unpredictable position in space, the temporal facilitation of subsequent stimuli becomes less important because no latency compensation is needed for this task. One may hypothesize that the instruction cue about the type of position judgment (relative or absolute) generates feedback to lower visual areas, thereby altering the efficacy of the lateral interaction loops before the stimulus is processed (Lamme & Roelfsema 2000).

In this context it is worth mentioning that in Hubbard and Bharucha (1988) all subjects showed the increase of the forward displacement with speed. However, there is a large difference in magnitude across subjects. Assuming that many individual top-down influences (e.g., attention, experience, or context) can alter the efficacy of the lateral interactions (Li et al. 2004) may explain this finding.

25.4 General discussion

We have suggested that the spatiotemporal dynamics of a network model that incorporates plausible assumptions about the local cortical connectivity accounts for perceptual correlates of localization errors observed in various motion paradigms. In its functional architecture the network model reflects converging lines of physiological and anatomical evidence that single neurons are not passive filters but dynamic entities with response properties depending on the collective behavior of large populations of cortical cells (Fitzpatrick

2000). The interplay between afferent inputs and local feedback loops results in a wave of activation that locks to the moving stimulus. The cooperative mechanisms within the network explain the critical dependence of the visual illusions on the stimulus attributes of contrast and velocity.

25.4.1 Facilitation through preactivation

The fundamental finding of our population study in the primary cortex was that the neural trajectory representations reveal a reduced latency when compared to a flashed stimulus. The threshold mechanism that causes such a temporal facilitation in the model network relies on the assumption that spreading subthreshold activation leads to a preactivation of surrounding neuronal populations (Kirschfeld & Kammer 1999). Consequently, the threshold for spike generation is reached faster compared to a population starting the processing of the afferent input from resting level. There is a growing body of experimental evidence from optical imaging and intracellular recording studies that supports the existence of a subthreshold depolarization wave temporally ahead of the afferent input (Grinvald et al. 1994; Bringuier et al. 1999; Jancke et al. 2004a).

One important question discussed in the context of the differential latency account to the flash-lag illusion is whether the temporal facilitation is omnidirectional or restricted to the specific motion trajectory (Nijhawan et al. 2004). Whitney and colleagues (Whitney, Murakami, et al. 2000; Whitney, Cavanagh, et al. 2000) found no significant difference in the effect even during unpredictable changes in direction including motion reversal. Because the interaction structure of the model network is completely symmetric there is no a priori preference for a certain direction. However, an asymmetry is introduced by the fact that the propagating excitation is followed in time and space by an inhibition wave. For the particular case of motion reversal this means that the afferent stimulus interacts with local inhibition caused by the previous stimulation of the same position in visual space. As a result, a coherently traveling wave has first to build up, suggesting that the observed flash-lag may be caused by the same mechanisms as for motion initiation and not by differential latencies. However, for a two-dimensional extension of the model network we expect that the circular spread of excitation mediated by an omnidirectional interaction structure may be sufficient to guarantee a temporal facilitation effect for a whole range of direction changes (excluding the particular case of motion reversal).

25.4.2 Readout of positional information

We have suggested that the traveling wave constitutes a common neural mechanism that relates the three illusory displacements observed with objects in motion (see also Müsseler et al. 2002). The localization at motion initiation and motion offset is defined by the starting and stopping position of the wave, respectively. The Fröhlich effect is explained by the fact that the starting position appears to be shifted forward with respect to the first physical position of the stimulus. The representational momentum effect has its correlate in the

P1: JZP Trim: 174mm × 247mm Top: 0.581in Gutter: 0.747in
CUUK929-25 cuuk929/Nijhawan ISBN: 978 0 521 86318 6 October 23, 2009 23:43

overshooting of the last stimulus position. Finally, the differential latency of the flash and the moving stimulus determines the localization error at a midposition of the trajectory. The "reading out" of the neuronal activity pattern is done from a snapshot for simplicity. It is important to note that an averaging of positions over a whole time interval (Eagleman & Sejnowski 2000) could have been chosen as well without qualitatively changing the overall pattern of results. For instance, using the time window defined by the suprathreshold flash response would explain a larger flash-lag compared to a readout at the time of peak latency. This is due to the systematic asymmetry of the temporal evolution with respect to the peak.

There is no reason to believe that the three position estimates will define visual illusions of the same order of magnitude. In the only study thus far that directly compared the three illusory displacements in a single experimental setting, Müsseler and colleagues (Müsseler et al. 2002) found different sizes for the Fröhlich, the flash-lag, and the representational momentum effect. Inspired by our work the authors also introduced the notion of a wave as a common mechanism that might explain their data.

25.4.3 Motion extrapolation

In the last 10 years, the flash-lag effect has been the object of an intense debate. Nijhawan (1994) renewed interest in this effect by explaining it as an attempt by the visual system to compensate for processing delays. These delays cause the neuronal representation to spatially lag behind the actual position of the moving object. The extrapolation hypothesis states that the visual system compensates for this spatial lag by actively shifting the coded position forward along the trajectory of motion. Consistent with this hypothesis, Berry and colleagues (1999) demonstrated convincingly that the peak of the population response of retinal ganglion cells may lead the stimulus position in visual space. However, they also reported a decrease of this spatial lead with increasing speed of the moving object. The findings of our population study in primary visual cortex suggest that the spatial lead may change into a spatial lag for sufficiently high speeds. Although this dependence on speed does not support the mechanism originally proposed by Nijhawan, the model simulations show that a mechanism based on lateral interactions may explain a substantial compensation of processing delays.

A task-related alteration of the efficacy of the lateral interactions by means of the proposed shunting mechanism allows controlling the extrapolation properties of the network. We propose such a top-down contribution as an explanation for the lack of effect in the flash-terminated cycle. In the flash-lag paradigm the cooperative forces within the network are tuned to prevent the wave from overshooting the vanishing position. For tasks that require the accurate localization of a moving object, the lateral interactions can be tuned to bring the peak of the neural trajectory representation close to the actual stimulus position. However, the system pays the price that this predictive representation overshoots the final position of the object when it abruptly vanishes (compare Fig. 25.5(a)).

Changes in the efficacy of the lateral interactions affect not only the decay but also the buildup of the wave. Consequently, the network model predicts a correlation between the error at the start and the end of the motion trajectory. This correlation has indeed been found in studies in which the subjects were asked to indicate both the perceived onset and offset position of the moving stimulus (Hubbard & Motes 2002; Müsseler et al. 2002; Thornton 2002). The fundamental finding was that the perceptual overshooting at the end of the trajectory is accompanied by the disappearance of the Fröhlich effect. Subjects reported instead a displacement along the path of observed motion. The occurrence of this so-called onset repulsion effect (Thornton 2002) is in agreement with the dynamic properties of the network model. Note that the spatial interactions within the network are omnidirectional. If the cooperative mechanisms are sufficiently strong there is a tendency for "extrapolation" also in the direction opposite to motion. The activity pattern that propagates beyond the point of onset does, however, not resonate with afferent inputs and thus quickly decays to resting level. Nonetheless, a readout of this pattern might explain the onset repulsion effect (see discussion in Thornton 2002).

In conclusion, we have shown that the proposed network model can place seemingly unrelated or even controversial findings about visual illusions observed with objects in motion in one coherent, unifying framework. With the recent advances in neuroscience we are beginning to understand the neural correlates of visual perception. Many systematic misperceptions remain, however, unsolved. They provide a fertile ground for combined modeling and physiological efforts that ultimately will lead to new insights into the complex structure of the visual system.

Acknowledgment

The work was supported by the BMBF (D.J.) and the "Acções Integradas Luso-Alemas/DAAD." We would like to thank two anonymous reviewers and Beena Khurana for their helpful comments on a previous draft.

References

Abeles, M. (1991). *Corticonics*. Cambridge: Cambridge University Press.
Baldo, M. V. C., & Caticha, N. (2005). Computational neurobiology of the flash-lag effect. *Vision Res* **45**: 2620–2630.
Baldo, M. V. C., & Klein, S. A. (1995). Extrapolation or attention shift? *Nature* **378**: 565–566.
Ben-Yishai, R., Hansel, D., & Sompolinsky, H. (1997). Traveling waves and the processing of weakly tuned inputs in a cortical network module. *J Comp Neurosci* **4**: 57–77.
Berry II, M. J., Brivanlou, I. H., Jordan, T. A., & Meister, M. (1999). Anticipation of moving stimuli by the retina. *Nature* **398**: 334–338.
Bringuier, V., Chavane, F., Glaeser, L., & Frégnac, Y. (1999). Horizontal propagation of visual activity in the synaptic integration field of area 17 neurons. *Science* **283**: 695–699.
Coltheart, M. (1980). Iconic memory and visible persistence. *Perception & Psychophysics* **27**: 183–228.

Dayan, P., & Abbott, L. F. (2001). *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems.* Cambridge, MA: MIT Press.

Douglas, R. J., Koch, C., Mahowald, M., Martin, K. A. C., & Suarez, H. (1995). Recurrent excitation in neocortical circuits. *Science* **269**: 981–985.

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

Erlhagen, W. (2003). Internal models for visual perception. *Biol Cybern* **88**: 409–417.

Erlhagen, W., Bastian, A., Jancke, D., Riehle, A., & Schöner, G. (1999). The distribution of neuronal population activation (DPA) as a tool to study interaction and integration in cortical representations. *J Neurosci Methods* **94**: 53–66.

Erlhagen, W., & Jancke, D. (2004). The role of action plans and other cognitive factors in motion extrapolation: A modelling study. *Vis Cogn* **11**(2 /3): 315–340.

Fitzpatrick, D. (2000). Seeing beyond the receptive field in primary visual cortex. *Curr Opin Neurobiol* **10**: 438–443.

Freyd, J. J., & Finke, R. A. (1984). Representational momentum. *J Exp Psych: LMC* **10**: 126–132.

Fröhlich, F. W. (1923). Über die Messung der Empfindungszeit. *Zeitschrift für Sinnesphysiologie* **54**: 58–78.

Fu, Y., Shen, Y., & Yang, D. (2001). Motion-induced perceptual extrapolation of blurred visual targets. *J Neurosci* **21** (RC172): 1–5.

Giese, M., & Xie, X. (2002). Exact solution of the nonlinear dynamics of recurrent neural mechanisms for direction selectivity. *Neurocomp* 44–46: 417–422

Gilbert, S. D. (1995). Dynamic properties of adult visual cortex. In M. Gazzaniga (ed.), *The Cognitive Neurosciences* (73–90). Cambridge: MIT Press.

Grinvald, A., Lieke, E., Frostig, R., & Hildesheim, R. (1994). Real-time optical imaging of naturally evoked electrical activity in intact frog brain. *J Neurosci* **14**: 2545–2568.

Grossberg, S. (1988). Nonlinear neural networks: Principles, mechanisms, and architectures. *Neural Networks* **1**: 17–61.

Hahnloser, R., Douglas, R. J., Mahowald, M., & Hepp, K. (1999). Feedback interactions between neuronal pointers and maps for attentional processing. *Nat Neurosci* **2**(8): 746–752.

Hubbard, T. L. (1990). Cognitive representation of linear motion: Possible direction and gravity effects in judged displacement. *Memory & Cognition* **18**(3): 299–309.

Hubbard, T. L., & Bharucha, J. J. (1988). Judged displacement in apparent vertical and horizontal motion. *Perception & Psychophysics* **44**: 211–221.

Hubbard, T. L., & Motes, M. A. (2002). Does representational momentum reflect a distortion of the length or the endpoint of a trajectory? *Cognition* **82**: B89–B99.

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

Jancke, D., Erlhagen, W., Dinse, H. R., Akhavan, A. C., Giese, M., Steinhage, A., et al. (1999). Parametric population representation of retinal location: Neuronal interaction dynamics in cat primary visual cortex. *J Neurosci* **19**: 9016–9028.

Jancke, D., Erlhagen, W., Schöner, G., & Dinse, H. R. (2004b). Shorter latencies for motion trajectories for flashes in population responses of cat visual cortex. *J Physiol* **556.3**: 971–982.

Khurana, B., & Nijhawan, R. (1995). Extrapolation or attentional shift? *Nature* **378**: 565–566.

Kirschfeld, K., & Kammer, T. (1999). The Fröhlich effect: a consequence of the interaction of visual focal attention and metacontrast. *Vis Res* **39**: 3702–3709.

P1: JZP Trim: 174mm × 247mm Top: 0.581in Gutter: 0.747in
CUUK929-25 cuuk929/Nijhawan ISBN: 978 0 521 86318 6 October 23, 2009 23:43

Kreegipuu, K., & Allik, J. (2003). Perceived onset time and position of a moving stimulus. *Vision Res* **43**: 1625–1635.

Krekelberg, B., & Lappe, M. (2000). A model of the perceived relative positions of moving objects based upon a slow averaging process. *Vision Res* **40**: 201–215.

Krekelberg, B., & Lappe, M. (2001). Neuronal latencies and the position of moving objects. *TINS* **24**: 335–339.

Lamme, V. A. F., & Roelfsema, R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *TINS* **23**: 571–579.

Lennie, P. (1981). The physiological basis of variations of visual latency. *Vision Res.* **21**: 815–824.

Li, W., Piëch, V., & Gilbert, C. D. (2004). Perceptual learning and top-down influences in primary visual cortex. *Nat Neurosci* **7**(6): 651–657.

Maiche, A., Budelli, R., & Gómez-Sena, L. (2007). Spatial facilitation is involved in the flash-lag effect. *Vision Res* **47**: 1655–1661.

Metzger, W. (1932). Versuch einer gemeinsamen Theorie der Phänomene Fröhlichs und Hazelhoffs und Kritik ihrer Verfahren zur Messung der Empfindungszeit. *Psychol Forsch* **16**: 176–200.

Müsseler, J., & Aschersleben, G. (1998). Localizing the first position of a moving stimulus: The Fröhlich effect and an attention-shifting explanation. *Perception & Psychophysics* **60**: 683–695.

Müsseler, J., Stork, S., & Kerzel, D. (2002). Comparing mislocalizations with moving stimuli: The Fröhlich effect, the flash-lag, and representational momentum. *Vis Cog* **9**(1/2): 120–138.

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

Nijhawan, R., Watanabe, K., Khurana, B., & Shimojo, S. (2004). Compensation of neural delays in visual-motor behaviour: No evidence for shorter afferent delays for visual motion. *Vis Cog* **11**(2/3): 275–298.

Purushothaman, G., Patel, S. S., Bedell, H. E., & Ogmen, H. (1998). Moving ahead through differential visual latency. *Nature* **396**: 424.

Salinas, E., & Abbott, L. F. (1994). Vector reconstruction from firing rates. *J Comput Neurosci* **1**: 89–107.

Thornton, I. M. (2002). The onset repulsion effect. *Spatial Vision* **15**: 219–243.

Whitney, D. (2002). The influence of visual motion on perceived position. *TICS* **6**: 211–216.

Whitney, D., Cavanagh, P., & Murakami, I. (2000). Temporal facilitation for moving stimuli is independent of changes in direction. *Vision Res* **40**: 3829–3839.

Whitney, D., Murakami, I., & Cavanagh, P. (2000). Illusory spatial offset of a flash relative to a moving stimulus is caused by differential latencies for moving and flashed stimuli. *Vision Res* **40**: 137–149.

Wilson, H. R., & Cowan, J. D. (1973). A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Kybernetik (Biol Cybern)* **13**: 55–80.