A psychophysical and computational analysis of the spatio-temporal mechanisms underlying the flash-lag effect†

André M Cravo, Marcus V C Baldo
Department of Physiology and Biophysics, Institute of Biomedical Sciences, University of São Paulo,
Av. Prof. Lineu Prestes, 1524, sala 245, 05508-900 São Paulo (SP), Brazil; e-mail: cravo@icb.usp.br
Received 8 January 2008, in revised form 4 May 2008; published online 10 December 2008

Abstract. Several accounts put forth to explain the flash-lag effect (FLE) rely mainly on either spatial or temporal mechanisms. Here we investigated the relationship between these mechanisms by psychophysical and theoretical approaches. In a first experiment we assessed the magnitudes of the FLE and temporal-order judgments performed under identical visual stimulation. The results were interpreted by means of simulations of an artificial neural network, that was also employed to make predictions concerning the FLE. The model predicted that a spatio-temporal mislocalisation would emerge from two, continuous and abrupt-onset, moving stimuli. Additionally, a straightforward prediction of the model revealed that the magnitude of this mislocalisation should be task-dependent, increasing when the use of the abrupt-onset moving stimulus switches from a temporal marker only to both temporal and spatial markers. Our findings confirmed the model's predictions and point to an indissoluble interplay between spatial facilitation and processing delays in the FLE.

1 Introduction

When observers compare the positions of a moving object and a stationary flash presented in the visual field, they usually perceive the moving stimulus as being spatially advanced in relation to the positions of the flash when, in fact, both stimuli happen to be physically aligned with each other in space-time. This phenomenon is called the flash-lag effect (FLE) and, although it was known of for over 70 years, it was rediscovered and has been vigorously discussed over the past 15 years (Metzger 1932; MacKay 1958; Nijhawan 1992, 1994, 2002; Baldo and Klein 1995; Purushothaman et al 1998; Whitney and Murakami 1998; Eagleman and Sejnowski 2000a, 2000b, 2000c, 2007; Krekelberg and Lappe 2000a, 2000b, 2001; Whitney 2002; Namba and Baldo 2004; Öğmen et al 2004; Baldo and Caticha 2005; Chappell et al 2006; Maus and Nijhawan 2006).

According to Eagleman and Sejnowski (2007), there are two main groups of models that attempt to explain the FLE. The difference between them lies in the mechanism responsible for the FLE: whereas one group postulates a misalignment in time, the other postulates errors in spatial localisation of the moving object as the underlying mechanism. The accounts of the temporal models are based mainly on the idea that different kinds of neural signals are processed at different speeds. In its simplest form, the FLE would result from a shorter perceptual latency of the moving stimulus compared to a stationary one (Baldo and Klein 1995; Purushothaman et al 1998; Whitney and Murakami 1998; Patel et al 2000; Murakami 2001). This account has been questioned by several authors largely on the basis of a condition known as 'flash initiated cycle'—FIC (Khurana and Nijhawan 1995; Eagleman and Sejnowski 2000a, 2000b, 2000c; Nijhawan 2002; Nijhawan et al 2004). In this condition, the FLE is observed even when the moving stimulus initiates its motion simultaneously with the presentation of the stationary stimulus (Khurana and Nijhawan 1995). According to Nijhawan (2002), it is difficult to conceive that a moving object, when it first comes suddenly into view, could still be processed faster than a flash. Yet, the FIC produces an undiminished FLE.

† Preliminary reports of this work have been presented at the Vision Science Society (VSS) Meeting in 2004, and the European Conference on Visual Perception (ECVP) Annual Meeting in 2007.

Eagleman and Sejnowski (2000a) argued that the differential-latencies model predicts a 'race' between the processing of a moving stimulus and of a stationary one in our visual system. Therefore, by giving the stationary stimulus a temporal advantage, the flash-lag illusion should disappear under the FIC condition. Their results showed the opposite: even by presenting the stationary stimulus with a 53 ms advantage, the FLE remained. Evidence against the differential-latencies model includes also comparisons between the latencies of moving and stationary stimuli by temporal-order judgments, showing no significant difference between their temporal processing (Baldo and Cravo 2004; Nijhawan et al 2004; Chappell et al 2006; Cravo and Baldo 2007).

The second group of theories proposes that the FLE is due mainly to spatial mechanisms (Nijhawan 1994; Brenner and Smeets 2000; Eagleman and Sejnowski 2000a, 2000b, 2000c, 2007; Krekelberg and Lappe 2000a, 2000b, 2001; Nijhawan et al 2004). According to the extrapolation model, our visual system uses information from the earlier part of the trajectory of the moving object to predict accurately its current position (Nijhawan 1994; Nijhawan et al 2004; Maus and Nijhawan 2006). The temporal-averaging model (Krekelberg and Lappe 2000a, 2000b, 2001) proposes that the visual system collects position signals over time and estimates the position on the basis of the integrated signals. The flash-lag would occur because position information of the flash persists for some time and therefore biases the position estimate towards the last seen position of the flash. In a recent reformulation of the 'postdiction' account, Eagleman and Sejnowski (2007) presented the motion-bias model, in which they posit that, when an observer must answer where a moving object was at a particular moment, a smear of spatial positions must be evaluated into a single, unambiguous answer. The outcome of the localisation is biased in the direction of motion signals that are collected over a small window of time after the flash.

Nevertheless, the class of spatial theories for the FLE has difficulties in explaining findings that involve direct manipulations of temporal aspects of the stationary stimulus. For instance, it has been shown that the magnitude of the FLE depends on the relative luminance between stationary and moving stimuli. Although increase (decrease) of luminance of the moving stimulus could allow a stronger (weaker) motion biasing, changes of luminance of the stationary stimulus should in principle not affect the motion biasing of the moving stimulus. Also, for high enough ratios, the FLE may turn into a flash-lead effect (Purushothaman et al 1998; Öğmen et al 2004). Here again, the spatial models have difficulties in explaining how the increase of luminance of the stationary stimulus could turn the FLE to a flash-lead effect, unless one takes into consideration the latency (temporal aspect) of the stationary stimuli.

Additionally, it has been found that the magnitude of the FLE is also modulated by the spatial predictability of the stationary stimulus (Baldo et al 2002; Baldo and Namba 2002; Namba and Baldo 2004). If temporal aspects, such as the latency of the stationary stimulus, are ignored, spatial models will not fully account for the transformation of a flash-lag into a flash-lead effect by changing the luminance of the stationary stimulus. Space-based models for the FLE could hardly explain the modulation of its magnitude when the focus of visual attention has to be directed to a stationary flash whose presentation site is made unpredictable (Eagleman and Sejnowski 2000b; Baldo et al 2002). Thus far, neither a temporal nor a spatial model alone has been able to accommodate the wide set of findings related to the FLE. In addition, since these models are mainly temporal or spatial in their essence, the potential links between them have been rarely explored.

The objective of the present work was to explore the relationship between tasks involving either visual localisation in space-time or temporal-order judgments by both empirical and computational approaches. In the following section, we report our attempt to reproduce the two main findings usually raised against the differential latency model

(the lack of difference between latencies of moving and stationary stimuli and the FIC). Next, we discuss these results in light of simulations performed by a feedforward neural network model (Baldo and Caticha 2005). Also, we use the computational model to make direct predictions concerning verifiable perceptual outcomes. Then we proceed to report findings from psychophysical experiments that empirically tested the predictions of the model. Finally, we contrast and discuss the present results in order to scrutinise the entanglement between the spatial and temporal components of the FLE.

2 Experiment 1: Comparing spatial localisation and temporal-order judgments

All participants were either undergraduate or graduate students at the University of São Paulo, aged between 18 and 30 years. They all had normal or corrected-to-normal vision. The experiment took place in a dimly lit room with relative sound isolation. In each session the participant sat in front of a 19 inch monitor connected to a PC 486 DX4 (60 MHz). The head of the volunteer was placed in a chin-rest so that the distance between the eyes and the monitor was kept constant at 57 cm. The computational routines were worked out in a specific program for psychophysical experimentation (MEL Professional v2.01—Psychology Software Tools, Inc.).

In experiments 1A and 1B, the stimuli consisted of two squares (0.1 deg $\times$ 0.1 deg), diametrically opposed to each other, presented at 5.5° of visual eccentricity. The moving stimulus rotated around the fixation point with a constant speed of 0.33 rev s^{-1} (figure 1). The stationary stimulus used in all psychophysical experiments and simulations was a luminance step-function and not a pulse stimulus. Baldo and colleagues (2002) found that the magnitude of the FLE is about the same when the temporal marker (flash) is either a pulse or a step-function. Nevertheless, a step-function stimulus emits more energy than a pulse and consequently is more like a moving stimulus. Since the aim of the present paper is to compare differences between the spatial and temporal computation of a moving and stationary stimuli, a step-function seems to be a sensible choice.
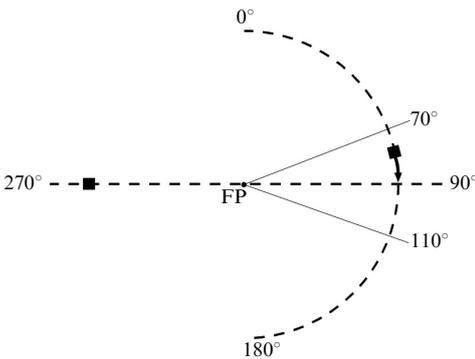
Figure 1. A schematic view of the stimuli configuration. It consisted of a fixation point (FP) (in the centre of the screen) and moving and stationary stimuli. Both stimuli were squares subtending 0.1 deg of visual angle and 5.5° of visual eccentricity. The moving stimulus had a circular trajectory with a constant speed (0.33 rev s^{-1}). The locations (left or right) and order of presentation of the stimuli varied according to the objective of each experiment.

2.1 Experiment 1A

In experiment 1A we assessed the magnitude of the FLE under the present stimulation conditions. Seven naive volunteers participated in this experiment. Each trial started with the presentation of a fixation point and, after an interval randomly chosen from 800 to 1800 ms, both moving and stationary stimuli were presented simultaneously (flash initiated cycle—FIC).

In blocked conditions, the stationary stimulus was always presented either in the right (90°) or in the left hemifield (270°), whereas the moving stimulus was presented in the opposite hemifield, in one out of eleven possible positions between 20° below and 20° above the imaginary line connecting the stationary stimulus and the fixation

point (corresponding to the interval between 70° and 110° when presented in the right hemifield and to the interval between 250° and 290° when presented in the left hemifield). The movement of the moving stimulus could follow either a clockwise or a counterclockwise trajectory (figure 1). The task was to report, at the instant the stationary stimulus was presented, the location of the moving stimulus in relation to the horizontal imaginary line passing through the stationary stimulus and the fixation point. Each volunteer participated in three sessions, the first for training and the other two for data collecting.

Psychometric curves were constructed from the collected data, and the points of subjective equality (PSE) were calculated and submitted to a two-tailed Student's t-test. Although we measured spatial mislocalisation, the perceived misalignments are here expressed in temporal units (temporal misalignment = spatial misalignment/speed), so the results from this experiment can be better compared with the results of experiment 1B, where a true temporal mislocalisation was measured, and with other experiments reported in the literature. The mean PSE was significantly different from zero ($p < 0.02$). This result shows that the present design was able to evoke an FLE under the FIC condition (9.6 ms $\pm$ 2.8 ms—figure 2).
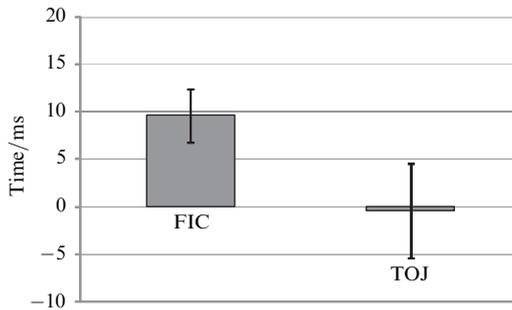
Figure 2. Perceived mislocalisation obtained in a relative localisation task under the flash-initiated cycle (FIC) condition and perceived asynchrony obtained in a temporal-order judgment (TOJ) task comparing the perceived onset of moving and stationary stimuli (mean ± 1 SEM). The perceived misalignments are expressed in temporal units (temporal misalignment = spatial misalignment/speed); positive values indicate that the moving stimulus was perceived as starting its trajectory spatially ahead of the stationary stimulus (FLE effect). The measured FLE effect was significantly different from zero ($p < 0.02$). The perceived asynchrony between moving and stationary stimuli were not significantly different from zero ($p > 0.8$).

2.2 Experiment 1B

Experiment 1B was designed to measure the relative perceptual latencies of both stimuli employed in experiment 1A, evaluated by means of temporal-order judgment (TOJ) tasks. Each trial started with the presentation of the fixation point. After a random interval between 800 and 1800 ms, moving and stationary stimuli were presented in opposite visual hemifields, at 90° and 270° positions (figure 1). The presentation hemifields (right or left) of moving and stationary stimuli were counterbalanced across trials. The temporal onset asynchrony between stationary and moving stimuli was randomly chosen from −167 ms to +167 ms. The task was to report which stimulus was perceived first.

Ten naive volunteers participated in this experiment. Each volunteer participated in three sessions, the first for training, and the other two for data collecting. The PSE extracted from the best-fitted psychometric curves corresponds to the temporal asynchrony between the two stimuli necessary to evoke the perception of simultaneity. The experimental data were submitted to a two-tailed Student's t-test, revealing that the mean PSE (0.4 ms $\pm$ 4.9 ms) was not significantly different from zero ($p > 0.9$).

These results point to similar perceptual latencies for both abrupt-onset, moving and stationary, stimuli. Although our findings agree with the results previously reported by Nijhawan and colleagues (2004) and also Chappell and colleagues (2006), an alternative interpretation will emerge from a joint analysis of psychophysical data and computational simulations.

3 Computational simulations

3.1 Neural network

The feedforward neural network we used to perform the present simulations was originally proposed by Baldo and Caticha (2005). The model was able to reproduce not only the FLE, but also the Fröhlich (1923) effect and the modulation of the FLE by stimulus luminance, trajectory, priming, and spatial predictability (Baldo and Caticha 2005). The essence of the model can be thus summarised: (i) each neuron in the network is modeled by a leaky integrate-and-fire (LIF) unit; (ii) the network consists of a feedforward layered architecture (input, hidden, and output layers); (iii) these layers are linked to each other by convergent and divergent interlayer connections; (iv) every neuron in a hidden layer receives input stemming from a one-dimensional receptive region of neurons in the input layer and projects onto a corresponding region in the output layer; (v) the pattern of connections reproduces centre–surround antagonist receptive fields (figure 3).
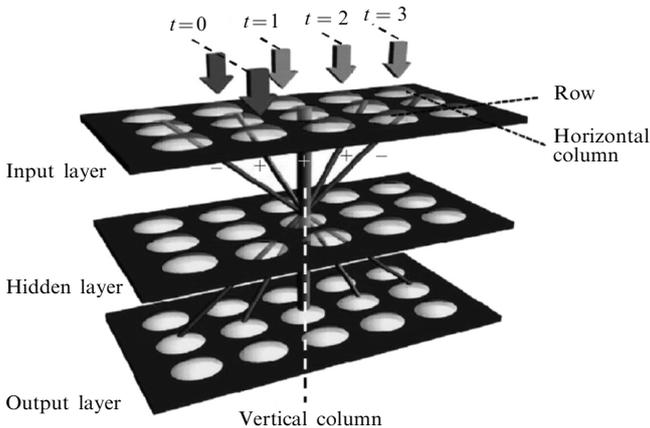
Figure 3. The basic architecture of the neural-network model. An input layer, one hidden layer, and an output layer are connected by convergent and divergent projections (lateral interactions). Each layer is composed of neurons arranged in rows and horizontal columns, linked to neurons located in other layers by means of excitatory (+) and inhibitory (−) connections (the figure shows the connections of only one selected neuron belonging to the hidden layer). A vertical column is defined by the connection between neurons located in different layers but belonging to the same row and the same horizontal column. A moving stimulus consists of a spatio-temporal sequence of inputs along either a row or a horizontal column of neurons.

In figure 3 each unit represents an LIF neuron linked to other neurons in the previous and/or next layer by means of either excitatory of inhibitory unidirectional 'synaptic' connections of strength J_{ij} from neuron j in layer $K - 1$ to neuron i in layer K (only the connections of a single neuron belonging to the hidden layer are fully shown in figure 3).

In neuron i in layer K the graded 'membrane potential' $v_i^K(t)$ at time t evolves according to:

$$v_i^K(t) = (1 - \Omega)v_i^K(t - 1) + I_i(t). \tag{1}$$

The constant Ω (kept between 0 and 1) defines the decaying rate of the leaking voltage. The input current $I_i(t)$ for a given neuron i is obtained by summing over the

neurons belonging to its receptive field; the overall input $I_i(t)$ depends on both the state $v_j^{K-1}(t-1)$ of each neuron j belonging to the layer $K-1$ at the time $t-1$, and the weight of the synaptic connection J_{ij} from neuron j to neuron i:

$$I_i(t) = \sum_{j \in \text{layer}(K-1)} J_{ij} \Theta[v_j^{K-1}(t-1) - \lambda]. \tag{2}$$

The step function Θ yields zero (one) for negative (positive) arguments and implements firing above threshold λ (neurons in the input layer display only graded membrane potentials, lacking a firing threshold, similar to generator potentials in sensory receptors). The connections J_{ij} between neurons j and i were chosen as follows: $J_{ij} = J_{vc} > 0$ (stronger excitatory connection) if j and i belong to the same vertical column (figure 3); $J_{ij} = J_{ilc} > 0$ ($J_{ilc} < J_{vc}$, weaker excitatory connection) if J_{ij} is the innermost lateral connection; and $J_{ij} = J_{olc} < 0$ (inhibitory connection) if J_{ij} is the outermost lateral connection. The parameters adopted in the present simulations were $J_{vc} = 0.4$, $J_{ilc} = 0.2$, and $J_{olc} = -0.2$ for the connections, $\lambda = 0.65$ for the threshold, and $\Omega = 0.6$ for the leaking constant. The intensity of both stationary (I_{stat}) and moving (I_{mov}) stimuli was the same in all simulations. These are precisely the same parameters as those employed by Baldo and Caticha (2005) and there was no attempt to find a particular set of parameters specifically suitable for the present simulations.

The comparisons between 'physical' and 'neural' locations of a given stimulus (either moving or stationary) were made by relating its position in the input layer to the corresponding neural activity generated in the output layer (a single vertical column defines the same coordinates for all layers). The output layer should not be seen, however, as a decisional stage or the percept itself, but rather as a still intermediate stage where the initial stimulation pattern has been transformed according to the dynamics of neural processing. Indeed, if a mismatch between 'physical' and 'neural' locations is observed already in this intermediate level (output layer), a related mismatch would be likely to manifest itself later on, in a further stage of sensory processing, ahead of that represented by the output layer.

3.2 Simulation of TOJ and FIC conditions

The perceptual latency of an abrupt-onset stimulus in our simulations is defined as the time interval between the presentation of a stimulus in the input layer and the first above-threshold activity evoked by the same stimulus in the output layer. For example, the perceptual latency of an abrupt-onset stimulus (either stationary or moving) is the number of time steps between the presentation of the stimulus in the input layer and the moment in which the respective evoked activity in the output layer reaches the threshold. In a similar manner, the perceptual latency of a continuously moving stimulus is defined as the number of time steps between the instant a given location is stimulated in the input layer and the first above-threshold activity brought about at the corresponding location in the output layer. Since the feedforward neural network devised here is a deterministic model, the output of a single computational run represents the average result one would obtain when performing a whole sequence of equivalent empirical trials.

In the present simulation, two abrupt-onset stimuli (one moving and another stationary) were presented in the input layer. Both were presented at the same time and aligned with each other (figure 4). As can be seen, the stationary stimulus presented in the input layer in location $s = 6$ ($t = 0$) reaches the threshold in the output layer four time steps later ($t = 4$) in the same corresponding location ($s = 6$). The moving stimulus presented in location $s = 6$ ($t = 0$) reaches the threshold in the output layer also four time steps later ($t = 4$) but one spatial step ahead, in location $s = 7$.
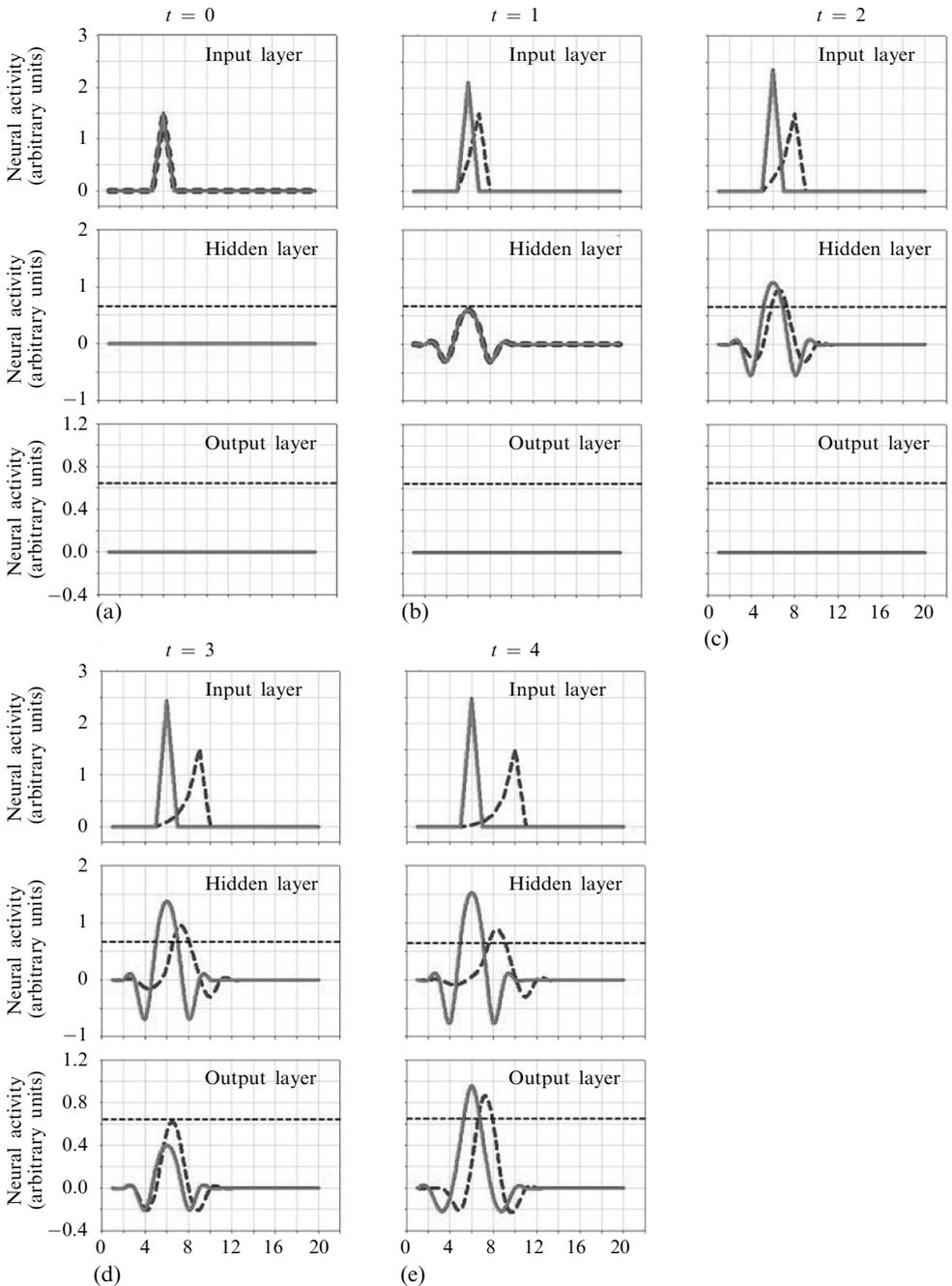
Figure 4. Simulated behaviour of the feedforward neural network. From (a) to (e) the panels show five snapshots ($t = 0$, 1, 2, 3, and 4 arbitrary units of time) of the neural activity taking place at the input (upper), hidden (middle), and output (lower) layers for both moving (dashed line) and stationary (solid line) stimuli. Both stimuli have the same intensity. (a) At $t = 0$ the moving stimulus initiates its movement in the input layer from horizontal column $s = 6$; at the same time a stationary stimulus is presented at the same horizontal column of the input layer in a different row, in spatial alignment with the moving stimulus (see figure 3 for the definitions of horizontal column and row). (b) At $t = 1$, we can notice activity induced by the stationary stimuli in the hidden layer. (c) At $t = 2$, we can notice above-threshold activity in the hidden layer induced by both moving and stationary stimuli (the threshold is indicated by the horizontal

After the transient activation of the network by the moving stimulus, we can observe that the ongoing temporal delay between the activation of a given location in the input layer and the corresponding above-threshold activity in the output layer decreases from four to three time steps. For example, in figure 4b, when the moving stimulus passes over position $s = 7$ of the input layer ($t = 1$), the above-threshold activity in the corresponding position ($s = 7$) of the output layer (figure 4e) occurs after three time steps only ($t = 4$), one time step less than the latency of an abrupt-onset stimulus.

This simulation can be interpreted as either the outcome of a spatial localisation task under the FIC condition (experiment 1A), or the outcome of a TOJ task (experiment 1B). As can be noticed in figure 4e, both stimuli first reach the threshold in the output layer at the same time (equivalent to the perception of simultaneity in a TOJ task), whereas a clear spatial lag arises between stationary and moving stimuli (equivalent to the FLE in a visual localisation task).

The outcome of the model clearly shows that the abrupt-onset moving stimulus has a perceptual latency equal to the latency of an abrupt-onset stationary stimulus (four time steps for both stimuli in the present simulation). Yet, this similarity would occur only at the first instants of the motion onset of the moving stimulus. According to the model, as soon as the second position of the input layer is stimulated, the spatio-temporal facilitation brought about by divergent connections promotes a reduction in processing latency of the moving stimulus, which decreased to three time steps in the present simulation. This view is compatible with that of Bachmann and colleagues (2003), according to which the initial establishment of the visible representation of an object takes longer than it takes to continuously update it with the newly arriving sensory data.

In the FIC condition, the transient dynamics involved in the establishment of a spatio-temporal facilitation, that is responsible for the decreased latency of moving stimuli, leads to the loss of the initial trajectory of an abrupt-onset moving stimulus (being possibly the root of the Fröhlich effect). Because the input of the moving stimulus is at a different position in each time step, there is insufficient activation at any one location in the movement trajectory before a certain threshold is crossed. The threshold can only be crossed after the spatial and temporal connections have provided enough facilitation to the next position of the moving stimulus. Therefore, our simulations suggest that the FLE (under the FIC condition) and the Fröhlich effect might emerge, at least partially, from similar mechanisms.

On the basis of these statements, one might argue that the model should predict noticeably differing outcomes resulting from the FIC condition and from the continuous-motion condition, in which the moving object has been in full view before the onset of the flash. As can be observed in further simulations (figure 5), the model actually predicts that the FLE in the continuous-motion condition has a magnitude of one spatial step, similar to the FLE in the FIC condition. This prediction is in agreement with experimental findings of Nijhawan (2002) and Öğmen et al (2004) who found a magnitude similar in both displays. Therefore, the FLEs simulated in both conditions are similar to each other and due to the same mechanism: spatial and temporal facilitations.

Figure 4 (continued)
dashed line in both hidden and output layers). (d) At $t = 3$ we can notice a still below-threshold activity in the output layer induced by both moving and stationary stimuli. (e) At $t = 4$ the activities generated by both moving and stationary stimuli first cross the threshold in the output layer simultaneously. This simulation is in agreement with the results we obtained in the present TOJ task. The activity generated by the stationary stimulus first crosses the threshold in the output layer at the horizontal column $s = 6$ (the same horizontal column where it has been presented in the input layer); the moving stimulus, however, crosses the threshold at the horizontal column $s = 7$, a clear manifestation of the FLE effect (under the FIC condition).
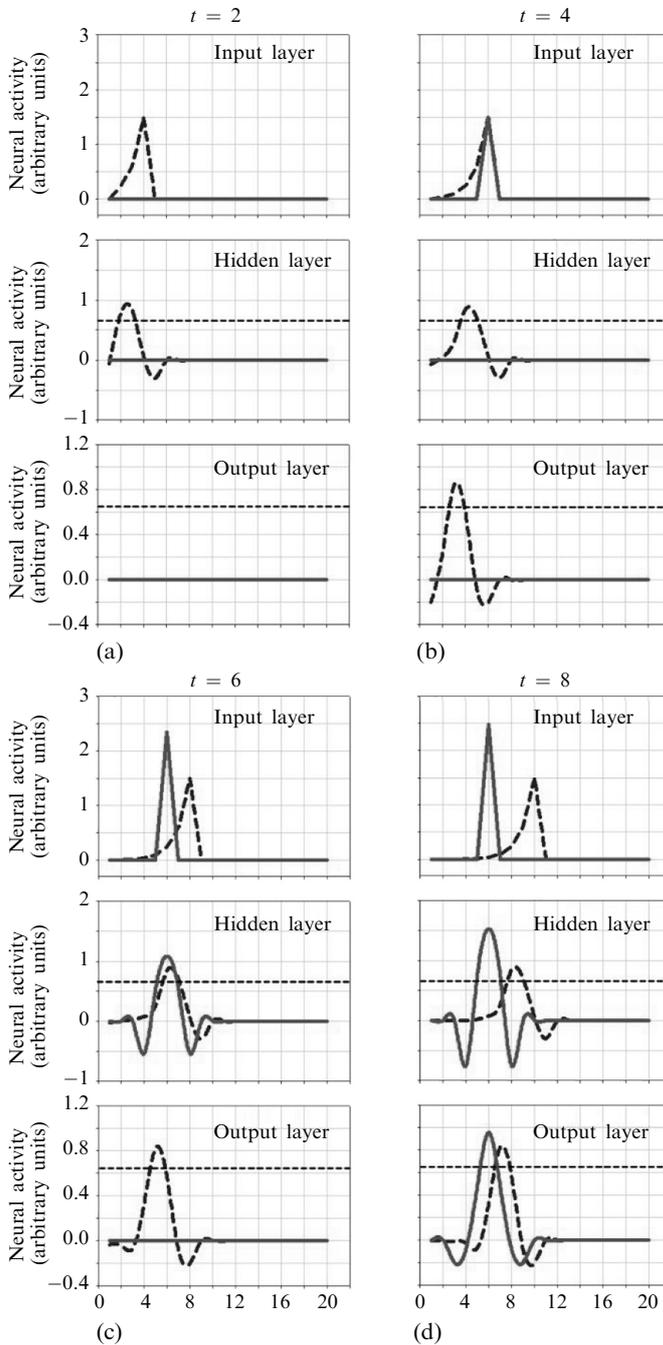
Figure 5. Simulated behaviour of the leaky integrate-and-fire (LIF) network. From (a) to (d) the panels show four snapshots ($t = 2, 4, 6$, and 8 arbitrary units of time) of the neural activity taking place at the input (upper), hidden (middle), and output (lower) layers for both moving (dashed line) and stationary/orthogonal moving (solid line) stimuli. (a) At $t = 2$ the moving stimulus is located, in the input layer, in horizontal column $s = 4$ (the motion was initiated at $t = 0$ in position $s = 2$, and therefore we can observe already an above-threshold activity in the hidden layer; the threshold is indicated by the horizontal dashed line). (b) At $t = 4$, when the moving stimulus is being presented in the horizontal column $s = 6$, the stationary/orthogonal moving stimulus is presented in the same horizontal column of the input layer but in a different row, in spatial

In the FIC condition, the spatial mislocalisation is due mainly to the loss of the initial trajectory of the moving stimulus, resulting from an insufficient, subthreshold facilitation. The same mechanism (spatial facilitation) will result in a decrease of the perceptual latency of the moving stimulus, leading to the standard FLE in the continuous-motion condition. Here again is an example of the indissoluble interplay between spatial and temporal mechanisms in the model proposed, since a spatial facilitation is responsible for the latency decrease (temporal aspect) of the moving stimulus.

Consistent with the empirical findings reported by Eagleman and Sejnowski (2000a), our network model shows that, even when the stationary stimulus is given a temporal advantage, the magnitude of the FLE (under the FIC condition) should not be altered, since this temporal advantage would not interfere with the processing of the moving stimulus. Paradoxically, the FLE should manifest itself even under a condition in which the onset of the stationary stimulus is perceived as occurring before the onset of the moving one. However, as the present simulations show, this situation is only apparently paradoxical, being entirely compatible with the dynamics of the computational network reported here, which helps to clarify the relationship between the temporal and spatial features underlying this perceptual phenomenon.

3.3 Computational predictions

Previous simulations allow us to conclude that the dynamics of temporal and spatial facilitations taking place throughout successive processing layers are responsible for (i) equal primary latencies for both moving and stationary, abrupt-onset, stimuli to reach an above-threshold activity; (ii) a shorter, secondary, perceptual latency of the moving stimulus when the spatio-temporal facilitation is already fully established (responsible for the FLE under the continuous-motion condition); (iii) the loss of the initial trajectory of the moving stimulus due to a transient activation latency during which the spatio-temporal facilitation is not yet fully established (responsible for the FLE under the FIC condition). From these observations it is possible to predict that a FLE would emerge even if the stationary 'flash' is replaced with an abrupt-onset *moving* stimulus. According to experiment 1B and our previous simulations, abrupt-onset stationary and moving stimuli should have similar latencies. As a result, not only could a FLE exist with an abrupt-onset moving stimulus replacing the 'flash', but it should also have a magnitude similar to the standard FLE.

To test these predictions we simulated two conditions wherein an abrupt-onset moving stimulus was perfectly aligned with an ongoing moving stimulus. In the first condition (figure 5) the abrupt-onset moving stimulus had a trajectory orthogonal to the path of the ongoing moving stimulus. In the second one, the trajectories were parallel (figure 6).

As seen in figure 5, this simulation is equivalent to the simulation of a standard continuous-motion FLE. In the continuous-motion FLE, an abrupt-onset stationary stimulus is presented, at a given instant, in the same horizontal column as, but in a different row than, the moving stimulus. Similarly, in this simulation an abrupt-onset stimulus is also presented in the same horizontal column as, but in a different row than, the continuous moving stimulus. In this case, however, the abrupt-onset stimulus is a spatial sequence of inputs along a horizontal column, travelling orthogonally towards, or away

Figure 5 (continued)
alignment with the continuously moving stimulus; by this time we can notice in the output layer an above-threshold activity induced by the moving stimulus. (c) At $t = 6$ the activity generated by the continuously moving stimulus first crosses the threshold in the output layer. (d) At $t = 8$ the activity generated by the stationary/orthogonal moving stimulus first crosses the threshold in the output layer at the horizontal column $s = 6$ (the same horizontal column where it is being presented in the input layer); at this time, however, the corresponding activity of the ongoing moving stimulus is located in the horizontal column $s = 7$, a clear manifestation of the flash-lag effect.
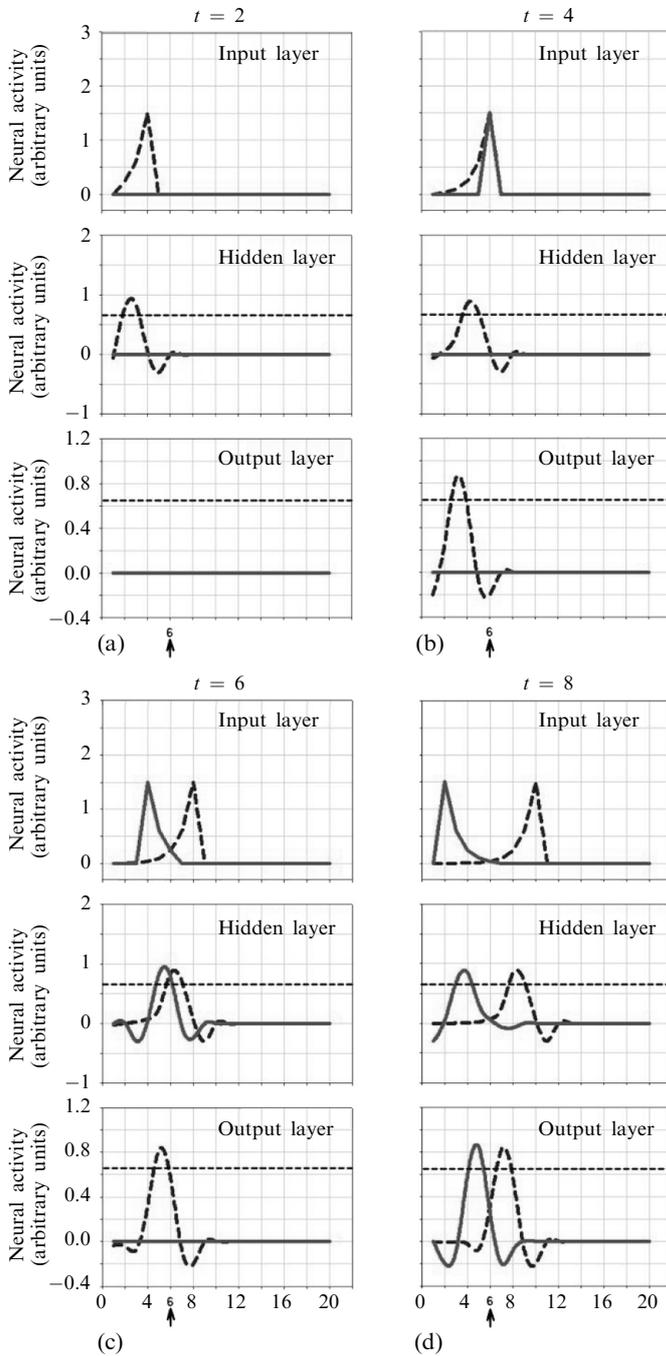
Figure 6. Simulated behaviour of the leaky integrate-and-first (LIF) network. From (a) to (d) the panels show four snapshots ($t = 2$, 4, 6, and 8 arbitrary units of time) of the neural activity taking place at the input (upper), hidden (middle), and output (lower) layers for both rightward moving stimulus (dashed line) and leftward moving stimulus (solid line). (a) At $t = 2$ the rightward moving stimulus is located, in the input layer, in horizontal column $s = 4$ (the motion was initiated at $t = 0$ in position $s = 2$, and therefore we can observe already an above-threshold activity in the hidden layer; the threshold is indicated by the horizontal dashed line). (b) At $t = 4$, when the rightward moving stimulus is being presented in the horizontal column $s = 6$, the leftward moving stimulus initiates its motion in the same horizontal column of the input layer but in a different

from, the trajectory of an ongoing moving stimulus. As seen in figure 5, the model predicts that the magnitude of the FLE generated under both conditions should not differ from that observed in the standard continuous-motion condition.

If the trajectory of the abrupt-onset moving stimulus, rather than being orthogonal, is parallel and opposite (say, leftward) to the path of the ongoing moving stimulus (say, rightward), the model predicts a different outcome. Figure 6 shows that, in this case, the FLE should increase and become noticeably larger than the standard FLE. However, a more careful analysis reveals that the magnitude of the FLE depends on the exact nature of the task. If one is asked to compare the position of the rightward-moving stimulus with the position of the leftward-moving stimulus at the instant the abrupt-onset (leftward) moving stimulus first reaches above-threshold activity in the output layer, then, according to our simulation, this gap would amount to two spatial steps (doubling the magnitude of the standard FLE). However, if the task is to compare the position of the rightward-moving stimulus to a stationary reference stimulus, such as the fixation point ($s = 6$), when the leftward-moving stimulus reaches above-threshold activity in the output layer, then the magnitude of the effect would still remain around one spatial step. The location of the fixation point in position $s = 6$ is not arbitrary, but coincides with the starting position of the leftward-moving, abrupt-onset stimulus.

The critical distinction between these two conditions concerns the use of the abrupt-onset moving stimulus as a temporal marker only, or as both temporal and spatial markers. The condition in which the abrupt-onset moving stimulus displays a trajectory orthogonal to the continuously moving stimuli exemplifies the use of the abrupt-onset moving stimulus as, essentially, a temporal marker. The same applies to the situation in which the spatial reference for the localisation task is the fixation point: the abrupt-onset moving stimulus provides a temporal reference only, irrespective of its (changing) location. Yet, when the localisation task involves the comparison between the locations of two stimuli moving along opposing trajectories, the abrupt-onset moving stimulus acts as both a temporal and a spatial reference marker.

Thus, the main predictions of the model can be summarised as follows: (i) a spatial mislocalisation can occur even if the flash is substituted by an abrupt-onset moving stimulus; (ii) the magnitude of this mislocalisation is similar to the standard FLE as long as the abrupt-onset moving stimulus is used exclusively (or essentially) as a temporal marker; (iii) if the abrupt-onset moving stimulus is used as a temporal as well as a spatial marker, then the observed mislocalisation should be significantly larger for opposing, but not for orthogonal, trajectories. These predictions were tested psychophysically.

4 Empirical verification of the computational predictions
4.1 Participants and methods
Ten volunteers (one author and nine naive subjects) participated in this experiment. Each trial started with the presentation of either a fixation point or a fixation line (see below). After a random interval between 800 and 1800 ms a moving stimulus was presented 5° below the fixation point/fixation line with its trajectory starting at a

Figure 6 (continued)
row, in spatial alignment with the rightward moving stimulus; by this time we can notice in the output layer an above-threshold activity induced by the rightward moving stimulus. (c) At $t = 6$ the activity generated by the leftward moving stimulus first crosses the threshold in the hidden layer. (d) At $t = 8$ the activity generated by the leftward moving stimulus first crosses the threshold in the output layer in the horizontal column $s = 5$; at this time, however, the corresponding activity of the rightward moving stimulus is located in the horizontal column $s = 7$. The arrow indicates the possible location of the fixation point (see experiment 2).

random position on the left of the visual field (figure 7). At a certain point in its trajectory, an abrupt-onset stimulus was presented 5° above the fixation point/fixation line, at the central meridian of the screen. The volunteer's task was to report if the continuously moving stimulus was located before or after a spatial reference mark (which differed from one session to another, as explained below) when the abrupt-onset stimulus was perceived. On each trial, the position of the continuously moving stimulus at the instant the abrupt-onset stimulus appeared was defined by the parameter estimation of sequential testing (PEST) procedure (Taylor and Creelman 1967).
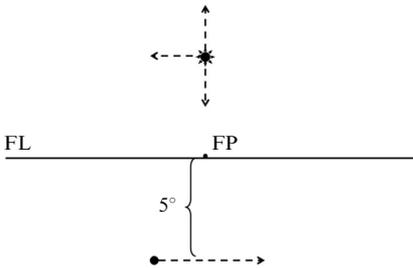
Figure 7. A schematic view of the stimuli configuration. It consisted of a fixation point (FP) or a fixation line (FL) in the centre of the screen and two moving stimuli. Both stimuli were circles (0.1 deg of diameter and 5° of visual eccentricity each). The moving stimulus had a constant speed of approximately 12 deg s^{-1}.

The abrupt-onset stimulus could be stationary or moving, depending on the session. In moving-stimulus sessions, the trajectory of the stimulus could be orthogonal (moving either towards or away from the trajectory of the moving stimulus) or parallel (opposite direction) to the trajectory of the continuously moving stimulus. In this case of parallel trajectories, subjects performed two different tasks, in separate sessions, differing in the judgment in hand (figure 7): (i) the localisation task required the comparison of the continuously moving stimulus with the fixation point at the instant the abrupt-onset moving stimulus (a temporal marker only) was first perceived; (ii) the task involved a comparison between the locations of both the continuously moving and the abrupt-onset moving stimuli at the instant the latter one was perceived, therefore acting as both temporal and spatial markers (a fixation line was employed instead of a fixation point in order to avoid the presence of a stationary reference point in the visual field).

Each subject performed five experimental sessions, which were named after the nature of the abrupt-onset stimuli: stationary (S), orthogonal trajectory towards the continuously moving stimulus (OT); orthogonal trajectory away from the continuously moving stimulus (OA); parallel opposing trajectories with fixation point as spatial reference (PP) (abrupt-onset moving stimulus provides a temporal marker only); parallel opposing trajectories with fixation line (PL) (abrupt-onset moving stimulus provides both, temporal and spatial, markers). Each session consisted of four runs, and the result of the localisation task was estimated for each run by averaging the last five reversals. The run ended when it reached a minimum of 15 reversals and the step size had decreased below 33.3 ms.

4.2 Results

An estimate for each session was calculated by averaging the outputs measured in the four runs. Although we measured spatial mislocalisation, the perceived misalignments are expressed in temporal units (temporal misalignment = spatial misalignment/speed) so that the results from this experiment can be better compared with other results of the literature. The experimental data were submitted to a one–way ANOVA, which showed a significant main effect ($F_{4,36} = 12.94$, $p < 0.001$—figure 8). An a posteriori comparison (Tukey's HSD test) revealed a significant difference between PL session and each one of the other sessions ($p < 0.001$). No other statistically significant difference was found among the remaining conditions ($p > 0.88$).
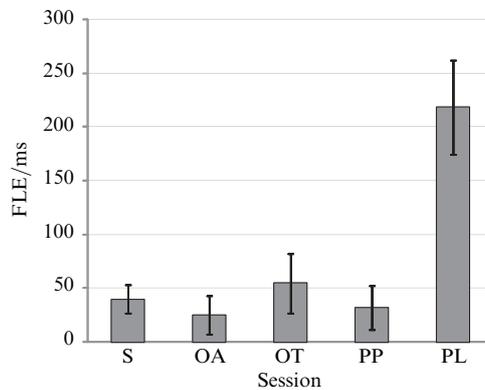
Figure 8. Mean ±1 SEM of the FLE obtained in experiment 2. The perceived misalignments are expressed in temporal units (temporal misalignment = spatial misalignment/speed). S = stationary (40.17 ms ± 12.92 ms); OA = *orthogonal* trajectory *away* from the continuously moving stimulus (25.40 ms ± 16.84 ms); OT = *orthogonal* trajectory *towards* the continuously moving stimulus (54.83 ms ± 28.00 ms); PP = *parallel* opposing trajectories with fixation *point* (abrupt-onset moving stimulus provides temporal marker only; 31.88 ms ± 20.35 ms); PL = *parallel* opposing trajectories with fixation *line* (abrupt-onset moving stimulus provides both, temporal and spatial, markers; 218.44 ms ± 43.74 ms).

The S, OT, OA, and PP sessions all have in common the use of the abrupt-onset stimulus essentially as a temporal marker: in session S, the abrupt-onset stimulus was stationary; in session PP, the spatial reference was given by the fixation point; and in both OT and OA sessions, the orthogonal nature of the trajectory of the abrupt-onset stimulus made its spatial displacement irrelevant to the spatial judgment in hand. Our results showed no significant difference ($p > 0.88$) between these conditions, suggesting not only that an FLE-like mislocalisation can be evoked by two moving stimuli (as long as one of them displays an abrupt onset) but also that the magnitude of the psychophysical effect is comparable to that observed in the standard FLE (session S). Moreover, in agreement with the predictions of our model, the magnitude of the perceptual mislocalisation obtained when the abrupt-onset moving stimulus was used simultaneously as a spatial and a temporal marker (session PL) was markedly larger ($p < 0.001$) in comparison to the condition where it served uniquely as a temporal reference (session PP). According to the prediction of the network model, the FLE should double its magnitude in session PL, whereas in our results the FLE seems larger than that. As can be observed in figure 8, the PL session produced a large variability, which might result from the very demanding task present in this condition. In fact, subjects reported finding it quite hard to compare the position of two moving stimuli. Owing to this variability, it is difficult to estimate the exact magnitude of the FLE; within the range of two standard errors, it can be as small as 130 ms. So far, the present results cannot either confirm or refute the quantitative prediction of the network model, but can validate the qualitative prediction that the magnitude should increase in the PL condition.

5 General discussion

In psychophysical experiments (Baldo and Cravo 2004; Nijhawan 2004; Chappell et al 2006; Cravo and Baldo 2007) it has been shown that abrupt-onset stationary and moving stimuli can be perceived as simultaneous but still leading to the FLE in the FIC condition [a similar result was found by Eagleman and Sejnowski (2000c)]. Here we replicated these empirical findings under close theoretical scrutiny, showing that perceptual latencies of abrupt-onset moving or stationary stimuli are indeed similar to

each other, although the computation of the position of a moving stimulus is additionally shaped by spatial facilitations taking place in the course of sensory processing. In fact, we were able to elicit an FLE-like perceptual mislocalisation by employing two moving stimuli: an ongoing stimulus in continuous motion and another moving stimulus abruptly presented in the visual field [similar results were recently reported by Gauch and Kerzel (2007)]. The magnitude of this perceptual effect depended on the precise nature of the localisation task. When the abrupt-onset moving stimulus was used solely as a temporal marker, the magnitude of the observed effect was comparable to that obtained in a control experiment replicating the standard FLE. However, when the abrupt-onset stimulus was used simultaneously as a temporal and a spatial marker, we found a larger perceptual effect. Remarkably, these empirical results were consistently predicted from computational simulations based on a very simple feed-forward neural network model (Baldo and Caticha 2005).

Although the FLE is fundamentally a spatio-temporal illusion, earlier models were biased towards either temporal (Baldo and Klein 1995; Purushothaman et al 1998; Whitney and Murakami 1998; Krekelberg and Lappe 2001) or spatial (Nijhawan 1994; Maus and Nijhawan 2006; Eagleman and Sejnowski 2007) mechanisms. Growing evidence has suggested robust interactions between these two aspects (Kreegipuu and Allik 2004; Maiche et al 2007). Kreegipuu and Allik (2004) attempted to disentangle confounding stimulus attributes by observing the apparent spatial or temporal lag of the FLE without using a reference flash. Instead, they used as temporal marker the colour change of a stationary reference stimulus. Their results show that subjects were quite accurate when asked to compare the instants moving and stationary stimuli change their colours. Nevertheless, when asked where the moving stimulus was when the reference stationary stimulus changed colours, observers tended to perceive the position of the moving stimulus as being ahead, in a flash-lag-like effect. These findings are in full agreement with the present results and simulations. Our results also revealed an accurate perceived simultaneity in temporal-order judgments together with a perceptual mislocalisation in spatial tasks. Thus, an FLE-like visual mislocalisation might emerge from the comparison between the location of a moving object that is currently visible and any abrupt event, either moving or stationary, such as the presentation of a flash, the sudden onset of motion of an object, or the colour change of a reference stimulus.

Here, we suggest a spatio-temporal framework as an attempt to accommodate results that could not be explained by either model alone. Our model includes in its architecture basic neural properties, such as divergent/convergent connections, which lead to spatial facilitation throughout sensory pathways. This facilitation can be seen as the basis of spatial models, such as motion extrapolation (Nijhawan 1994, 2002) and motion-bias accounts (Eagleman and Sejnowski 2007). Additionally, the existence of a sequence of neural processing stages naturally gives rise to intrinsic temporal delays. These delays lie at the very core of the differential-latencies account (Baldo and Klein 1995; Purushothaman et al 1998; Whitney and Murakami 1998; Patel et al 2000; Murakami 2001; Öğmen et al 2004). In the present account, the relationship between temporal and spatial mechanisms is not arbitrary, nor an ad hoc attempt to combine the explanatory power of both models. On the contrary, this relationship emerges naturally from the structural and functional attributes that characterise the network model offered here. On one hand, a spatial component is intrinsically related to the network architecture, when convergent and divergent connections comprise the building blocks of spatial facilitation mechanisms. On the other hand, the neural processing throughout successive sensory circuits adds temporal delays in either transmitting a signal along synaptic pathways or raising the activity of a pool of neurons towards above-threshold levels.

The neural-network model proposed here had already shown its explanatory capacity by being able to reproduce not only the FLE itself, but also its modulation by stimulus luminance, trajectory, priming, and spatial predictability (Baldo and Caticha 2005). However, in the present work we tested its predictive power by checking empirical results anticipated by computational simulations based on the model. If we take into account the present experimental results and computational simulations, it is possible to conceive spatio-temporal integration as a unifying framework to account for FLE and its relationship with other visual phenomena, such as the Fröhlich effect and temporal-order perception. This spatio-temporal integration would result from both the spatial pattern of convergent/divergent connections and the temporal dynamics of neuronal activation along successive layers of neural processing.

Even though the essential operations of spatio-temporal facilitation and processing delays are taken into account in the present network model, it unquestionably lacks several other physiological mechanisms involved in genuine experimental procedures. Therefore, we should not expect a faultless quantitative agreement between the reported empirical findings and the theoretical predictions offered by a still crude and unfinished computational model. Nonetheless, despite its simplicity, the present model is already able to predict, with a remarkable degree of qualitative accuracy, the outcome of psychophysical procedures involving both temporal-order and spatial-localisation tasks. In addition, it discloses what may turn out to be the very basic mechanisms underlying the FLE and possibly other perceptual phenomena, making it possible to look inside and scrutinise the operations of a running neural network.

References
Bachmann T, Luiga I, Poder E, Kalev K, 2003 "Perceptual acceleration of objects in stream: evidence from flash-lag displays" *Consciousness and Cognition* **122** 279 – 297
Baldo M V C, Caticha N, 2005 "Computational neurobiology of the flash-lag effect" *Vision Research* **4520** 2620 – 2630
Baldo M V C, Cravo A M, 2004 "A misconception about the relationship between the flash-lag effect and temporal order judgments" *Journal of Vision* **4**(8) 579 (abstract)
Baldo M V C, Kihara A H, Namba J, Klein S A, 2002 "Evidence for an attentional component of perceptual misalignment between moving and flashing stimuli" *Perception* **31** 17 – 30
Baldo M V C, Klein S A, 1995 "Extrapolation or attention shift?" *Nature* **378** 565 – 566
Baldo M V C, Namba J, 2002 "The attentional modulation of the flash-lag effect" *Brazilian Journal of Medical and Biological Research* **35** 969 – 972
Brenner E, Smeets J B J, 2000 "Motion extrapolation is not responsible for the flash-lag effect" *Vision Research* **40** 1645 – 1648
Chappell M, Hine T J, Acworth C, Hardwick D W, 2006 "Attention 'capture' by the flash-lag flash" *Vision Research* **46** 3205 – 3213
Cravo A M, Baldo M V C, 2007 "The spatio-temporal basis of the flash-lag effect: Psychophysical and computational approaches" *Perception* **36** Supplement, 87
Eagleman D M, Sejnowski T, 2000a "Motion integration and postdiction in visual awareness" *Science* **287** 2036 – 2038
Eagleman D M, Sejnowski T, 2000b "The position of moving objects (reply to Krekelberg, Lappe, Whitney, Cavanagh)" *Science* **289** 1107a
Eagleman D M, Sejnowski T, 2000c "Flash-lag effect: Differential latency not postdiction (reply to Patel, Öğmen, Bedell, Sampath)" *Science* **290** 1051a
Eagleman D M, Sejnowski T J, 2007 "Motion signals bias localization judgments: a unified explanation for the flash-lag, flash-drag, flash-jump and Fröhlich illusions" *Journal of Vision* **7**(3) 1 – 12, http://www.journalofvision.org/7/4/3
Fröhlich F W, 1923 "Über die Messung der Empfindungszeit" *Zeitschrift für Sinnesphysiologie* **54** 58 – 78
Gauch A, Kerzel D, 2007 "Flashes and abrupt onsets of moving stimuli as probe stimuli in the flash-lag effect" *Perception* **36** Supplement, 90

Khurana B, Nijhawan R, 1995 "Extrapolation or attention shift? (reply to Baldo and Klein)" *Nature* **378** 566

Kreegipuu K, Allik J, 2004 "Confusion of space and time in the flash-lag effect" *Perception* **33** 293 – 306

Krekelberg B, Lappe M, 2000a "The position of moving objects" *Science* **289** 1107a

Krekelberg B, Lappe M, 2000b "A model of the perceived relative positions of moving objects based upon a slow averaging process" *Vision Research* **40** 201 – 215

Krekelberg B, Lappe M, 2001 "Neuronal latencies and the position of moving objects" *Trends in Neurosciences* **23** 335 – 339

MacKay D, 1958 "Perceptual stability of a stroboscopically lit visual field containing self-luminous objects" *Nature* **181** 507 – 508

Maiche A, Budelli R, Gomez-Sena L, 2007 "Spatial facilitation is involved in flash-lag effect" *Vision Research* **47** 1655 – 1661

Maus G W, Nijhawan R, 2006 "Forward displacements of fading objects in motion: The role of transient signals in perceiving position" *Vision Research* **46** 4375 – 4381

Metzger W, 1932 "Versuch einer gemeinsamen Theorie der Phänomene Fröhlichs und Hazelhoffs und Kritik ihrer Verfahren zur Messung der Empfindungszeit" *Psychologische Forschung* **16** 176 – 200

Murakami I, 2001 "A flash-lag effect in random motion" *Vision Research* **41** 3101 – 3119

Namba J, Baldo M V C, 2004 "The modulation of the flash-lag effect by voluntary attention" *Perception* **34** 621 – 631

Nijhawan R, 1992 "Misalignment of contours through the interaction of apparent and real motion systems" *Investigative Ophthalmology and Visual Sciences* **33** 1415

Nijhawan R, 1994 "Motion extrapolation in catching" *Nature* **370** 256 – 257

Nijhawan R, 2002 "Neural delays in visual motion and the flash-lag effect" *Trends in Cognitive Sciences* **6** 387 – 393

Nijhawan R, Watanabe K, Khurana B, Shimojo S, 2004 "Compensation of neural delays in visual-motor behavior: evidence for shorter afferent delays for visual motion" *Visual Cognition* **11** 275 – 298

Öğmen H, Patel S S, Bedell H E, Camuz K, 2004 "Differential latencies and the dynamics of the position computation process for moving targets assessed with the flash-lag effect" *Vision Research* **44** 2109 – 2128

Patel S S, Öğmen H, Bedell H E, Sampath V, 2000 "Flash-lag effect: Differential latency not postdiction" *Science* **290** 1051a

Purushothaman G, Patel S S, Bedell H E, Öğmen H, 1998 "Moving ahead through differential visual latency" *Nature* **396** 424

Taylor M M, Creelman C D, 1967 "PEST: Efficient estimates on probability functions" *Journal of the Acoustical Society of America* **41** 782 – 787

Whitney D, 2002 "The influence of visual motion on perceived position" *Trends in Cognitive Sciences* **6** 211 – 216

Whitney D, Murakami I, 1998 "Latency difference not spatial extrapolation" *Nature Neuroscience* **1** 656 – 657

PERCEPTION