Interaction between motion scales: when performance in motion discrimination is worse for compound stimuli than for its integrating components.

Raúl Luna[1] & Ignacio Serrano-Pedraza[1,2],

[1] Faculty of Psychology. Complutense University of Madrid, Madrid, 28223, Spain
[2] Institute of Neuroscience. Newcastle University, Newcastle upon Tyne, NE2 4HH, UK

E-mail: raluna01@ucm.es, iserrano@ucm.es

Abstract

Motion direction discrimination becomes impaired when combinations of drifting high spatial frequency (HSF) and static low spatial frequency (LSF) patterns are merged into a compound stimulus. Such impairment has been suggested to occur due to an interaction between motion sensors tuned to coarse and fine scale spatial patterns. This interaction is modulated by different stimulus parameters like temporal frequency, size, the spectral components mixed, and their relative contrast. The present research precisely aims to explore in a deeper way the interaction's dependency upon the spatial frequency and the relative contrast of the components when both move coherently. Two experiments were therefore performed measuring duration thresholds (Experiment 1) and proportion of correct responses (Experiment 2) in a motion direction discrimination task. Stimuli were vertical Gabor patches of 4 deg diameter horizontally drifting with a speed of 2 deg/sec. Simple LSF and HSF stimuli as well as complex stimuli where both components moved coherently (LSFm+HSFm) were used. These were grouped in the following LSF and HSF pairs: 0.25-0.75, 0.5-1.5, 1-3 and 2-6 c/deg. Each component had a Michelson contrast of 28% or 7%, giving rise to different relative contrast combinations. Most interestingly, the results show a decrease in performance for complex stimuli with respect to each of their simple components when the LSF component has a lower contrast than the HSF one. The decrease depends on the particular spatial frequencies mixed in a stimulus. Further knowledge about the inhibitory mechanism is thus provided, revealing its joint dependency upon contrast and spatial frequency.

Keywords: motion perception, contrast, spatial frequency, interaction between fine and coarse scales, motion sensors, inhibition.**Introduction**

Errors in motion direction discrimination had previously been reported to systematically arise in humans under specific stimulus conditions. These conditions namely refer to drifting stimuli which add together coarse and fine scale patterns (Derrington & Henning, 1987; Henning & Derrington, 1988; Derrington, Fine, & Henning, 1993; Nishida, Yanagi, & Sato, 1995; Serrano-Pedraza, Goddard, & Derrington, 2007; Serrano-Pedraza & Derrington, 2010; Serrano-Pedraza, Gamonoso-Cruz, Sierra-Vazquez, & Derrington, 2013; Gekas, Masson & Mamassian, 2017; Luna & Serrano-Pedraza, 2018; see also the ''Interaction across different spatial scales'' section in Nishida, 2011). The systematic errors are evident when, at short stimulus presentation durations, a low spatial frequency static pattern is added to a moving high spatial frequency one, giving rise to a compound stimulus (Derrington & Henning, 1987). The occurrence of this systematic error phenomenon has been regarded to the existence, in later stages of motion processing, of an inhibitory mechanism producing an interaction between motion sensors tuned to coarse and fine features (Serrano-Pedraza, Goddard, & Derrington, 2007).

Energy models of human visual motion sensing conceive motion as computed by motion sensors operating in parallel that are spatially localized and are spatial frequency, temporal frequency and orientation tuned (Levinson & Sekuler, 1975; Adelson & Movshon, 1982; Adelson & Bergen, 1985; Anderson & Burr, 1985, 1989, 1991; Anderson, Burr, & Morrone, 1991; Cameron, Baker & Boulton, 1992). The aforesaid energy models have been successful in explaining different phenomena which had been reported in the literature. These include reverse Phi, apparent motion, the missing-fundamental illusion, etc., (Adelson & Bergen, 1985; Watson & Ahumada, 1985; van Santen & Sperling, 1985). With that being said, the relevance of the findings reporting systematic errors in motion direction discrimination lies in the fact that they challenge the classical notion about parallel motion processing by the different motion sensors. Conversely, they postulate a later stage in which non-linearities between motion sensors tuned to coarse and fine spatial scales take place due to the action of an inhibitory mechanism.

We know from previous research that this inhibitory mechanism produces the strongest interactions between coarse and fine scales at short stimulus presentation durations (lower than 100 msec). However, a strong effect can still be observed for longer presentations (Serrano-Pedraza, Goddard, & Derrington, 2007). The interaction between motion sensors tuned to coarse and fine spatial frequency components has been mainly reported to occur in compound stimuli integrated by component pairs with spatial frequencies between 1 c/deg

and 3 c/deg, but also a strong effect can be appreciated when the low spatial frequency component is lower than 0.75-1 c/deg and the high spatial frequency component is higher than 1.5-3 c/deg (Derrington & Henning, 1987; Henning & Derrington, 1988; Luna & Serrano-Pedraza, 2018). The interaction is stronger for large stimuli (Serrano-Pedraza & Derrington, 2010) and is dependent upon the relative contrast of the high and low spatial frequency components (Henning & Derrington, 1988; Serrano-Pedraza & Derrington, 2010). Also, the relative phase of the spatial frequency components in the compound stimuli has no effect on the interaction (Henning & Derrington, 1988). Likewise, the effect of the interaction on motion discrimination is smaller than its effect on motion surround suppression (Serrano-Pedraza et al., 2013). Finally, Serrano-Pedraza et al., (2007) and Serrano-Pedraza & Derrington (2010), using anisotropic noise, interestingly found that performance for motion discrimination was impaired for complex stimuli composed of low and high spatial frequency patterns moving coherently (i.e. same direction and same speed). This result was found for a particular pair of spatial frequency components, both having the same contrast energy. In this vein, Burge & Gesiler (2015) reported an impairment in speed discrimination for moving naturalistic images (broadband stimuli with $1/f$ amplitude spectra) when compared to simpler drifting sinewave gratings. Additionally, Jogan & Stocker (2015) found different discrimination thresholds and matching speeds for drifting compound stimuli relative to simple gratings.

In the present study we want to explore in a deeper manner the results from Serrano-Pedraza et al., (2007) and Serrano-Pedraza & Derrington (2010). We want to do so by extending the number of spatial frequency components tested and by changing their relative contrast. Measurements of duration thresholds and proportion of correct responses in a motion direction discrimination task have been performed along two experiments. These measurements are done on different horizontally drifting complex stimuli made out of two components: a low spatial frequency (LSF) and a high spatial frequency (HSF) one. Also, the LSF and HSF components are tested individually for comparison. Furthermore, different spatial frequency combinations for the LSF and HSF components are used as well as different relative contrasts. In this regard, the present study offers more exhaustive data about the inhibitory interaction between motion sensors tuned to low and high spatial frequency scales and how it depends on the spectral content and the contrast of the stimuli.

Method

Subjects

Five human subjects (aged 21-42), 2 males and 3 females, took part in Experiment 1. RLV and ISP were aware of the purpose of the study. Regarding Experiment 2, 6 subjects (aged 20-24), 2 males and 4 females took part in its first part. In its second part, 5 subjects took part (aged 20-28), 2 males and 3 females. For both parts of Experiment 2, the only subject aware of the purpose of the study was RLV. The normal visual acuity of the participants was assessed using the SLOAN ETDRS 2000 letter series (Precision Vision, USA, precision-vision.com) (used at a distance of 40 cm and 3 m). Their stereoscopic visual acuity was also tested, this time using the Frisby stereotest (60 cm distance). Only those subjects with a visual acuity lower than logMAR=0.5 (for both eyes) and a stereoacuity with a value lower than 500 arcsec would participate in the experiments. Every subject needed to sign an informed consent before completing the experiments. These were conducted in a dark room. A chin rest (UHCOTech HeadSpot, Houston, TX) was used in order to control the stimulus observation distance and to stabilize the subjects' head. So that tracking eye movements could be minimized, subjects were asked to fixate, before stimulus presentation, on a small cross (0.258º x 0.258º) located in the center of the screen. Experimental procedures were approved by the Universidad Complutense de Madrid Ethics Committee.

Equipment

For the presentation of stimuli, a gamma-corrected 17-in. Eizo Flex Scan T565 monitor (Eizo Corp., Hakusan, Japan) controlled by a Mac Pro was used. Further, Matlab (MathWorks, Natick, MA) was run making use of the the Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997; Kleiner, Brainard, & Pelli, 2007; www. psychtoolbox.org). The stimuli were displayed using an "Nvidia GForce 8800GT" graphic card that had its output processed by the DataPixx Lite video processor (Vpixx Technologies Inc., Montreal, Canada), letting us have a 16-bit video DACs to produce low contrast stimuli with a high precision. A Gamma correction was made on the monitor through the use of a Minolta LS-100 photometer (Konica Minolta Optics, Inc., Osaka, Japan). In order to automate this calibration process, routines were programed in Matlab (MathWorks, Natick, MA). The resolution of the monitor was of 800 × 600 pixels (horizontal × vertical), it had a vertical frame rate of 148 Hz, and its mean luminance was 49.1 cd/m^2..

Stimuli

The stimuli were generated using Matlab (The MathWorks, Inc, Natick, MA). The images of the stimuli had a resolution of 512 × 512 pixels. They were presented in the centre of the monitor in a square of 19.5 × 19.5 cm and were viewed at a distance of 100 cm, therefore subtending a visual angle of 11.3 × 11.3 deg. The spatial resolution of the display was 45.9 pixels per degree. The stimuli were 2D vertical Gabor patches (i.e. sinusoidal gratings with a 2D Gaussian function as spatial window). The size of the stimuli was 2 × σ_{xy} (spatial standard deviation, $\sigma_{xy} = \sigma_x = \sigma_y = 2$ deg) and they drifted horizontally (either direction, rightward or leftward, was equally likely). The spatial frequency bandwidths in octaves (B_{oct}) and the orientation bandwidth (O_{deg}) in degrees (full-width at half-height) of our stimuli were 0.25 c/deg (B_{oct}=1.13 octaves and O_{deg}=41.09 degrees); 0.5 c/deg (B_{oct}=0.547 octaves and O_{deg}=21.23 degrees); 0.75 c/deg (B_{oct}=0.362 octaves and O_{deg}=14.24 degrees); 1 c/deg (B_{oct}=0.271 octaves and O_{deg}=10.7 degrees); 1.5 c/deg (B_{oct}=0.18 octaves and O_{deg}=7.15 degrees); 2 c/deg (B_{oct}=0.135 octaves and O_{deg}=5.36 degrees); 3 c/deg (B_{oct}=0.09 octaves and O_{deg}=3.58 degrees); and 6 c/deg (B_{oct}=0.045 octaves and O_{deg}=1.79 degrees). These values were obtained with the next equations:

$$B_{oct}(\rho_i) = log_2\left[\frac{\sqrt{2}\pi\sigma_x\rho_i+\sqrt{ln2}}{\sqrt{2}\pi\sigma_x\rho_i-\sqrt{ln2}}\right]; \qquad (1)$$

$$O_{deg}(\rho_i) = 2\tan^{-1}\left[\frac{\sqrt{ln2}}{\sqrt{2}\pi\sigma_y\rho_i}\right], \qquad (2)$$

where ρ_i is the spatial frequency of a 2D Gabor patch.

Two types of stimuli were used. These were simple moving Gabor patches and complex Gabor patches which resulted from adding two simple Gabors with different spatial frequency components (both spatial frequency components moving with the same direction at the same speed). The speed of the stimuli was always 2 deg/sec. The spatial phases of the components set randomly in all experiments. As will be shown, contrast and spatial frequency varied from one experiment to another.

A simple drifting Gabor patch can be formally described as follows:

$$\mathrm{L}(\mathrm{x},\mathrm{y},\mathrm{t}) = \mathrm{L}_0\left[1 + m(\mathrm{t})\exp\left\{-\frac{\hat{\mathrm{x}}^2}{2\sigma_{\hat{\mathrm{x}}}^2} - \frac{\hat{\mathrm{y}}^2}{2\sigma_{\hat{\mathrm{y}}}^2}\right\} \times \cos(2\pi\rho_1(\hat{\mathrm{x}} - \mathrm{vt}) + \phi_1)\right] \qquad (3)$$

The following equation describes a complex moving stimulus where both components are moving at the same speed:

$$L(x,y,t) = L_0 \left[\begin{array}{c} 1 + \exp\left\{ -\frac{\hat{x}^2}{2\sigma_x^2} - \frac{\hat{y}^2}{2\sigma_y^2} \right\} \\ \times \left[m_1(t)\cos(2\pi\rho_1(\hat{x} - vt) + \phi_1) + m_2(t)\cos(2\pi\rho_2(\hat{x} - vt) + \phi_2) \right] \end{array} \right] \quad (4)$$

where $\hat{x} = x\cos(\theta) + y\sin(\theta)$ and $\hat{y} = -x\sin(\theta) + y\cos(\theta)$; x and y are the on-screen positions; θ is the orientation expressed in degrees (all stimuli were vertically oriented), $\theta = 0^o$; L_0 is the mean luminance ($L_0 = 49.1\ cd/m^2$); ρ_1 and ρ_2 refer to the spatial frequencies of the Gabors, in cycles per degree (c/deg); ϕ_1 and ϕ_2 are the their phases, expressed in radians (the phases of the spatial frequency components were set at random in each trial of all experiments); σ_x and σ_y stand for the spatial standard deviations of the 2D spatial Gaussian window, in terms of degrees of visual angle (for all stimuli σ_x = σ_y = 2 deg); v is the speed of the moving Gabor patch, in degrees per second (v = 2 deg/s); and m stands for the Michelson contrast as a function of time, that function being the following:

$$m(t) = M \times \exp\left\{ -t^2 / (2\sigma_t^2) \right\} \quad (5)$$

where t is time, in milliseconds (msec); M is the peak contrast, and σ_t is the temporal standard deviation. Regarding these last two parameters: the contrast and the temporal standard deviations; simple Gabor patches had a peak Michelson contrast either of 0.28 (28%) or 0.28/4 (7%). Moreover, the final contrast of the complex Gabor stimuli was the result of adding two simple Gabors with different combinations of contrasts. To finish with, we will describe in the forthcoming procedure section the Bayesian staircase method that served to control the parameter σ_t from which duration thresholds ($2 \times \sigma_{t0}$) were obtained. See Figure 1 for examples of the complex stimuli used in the experiments.

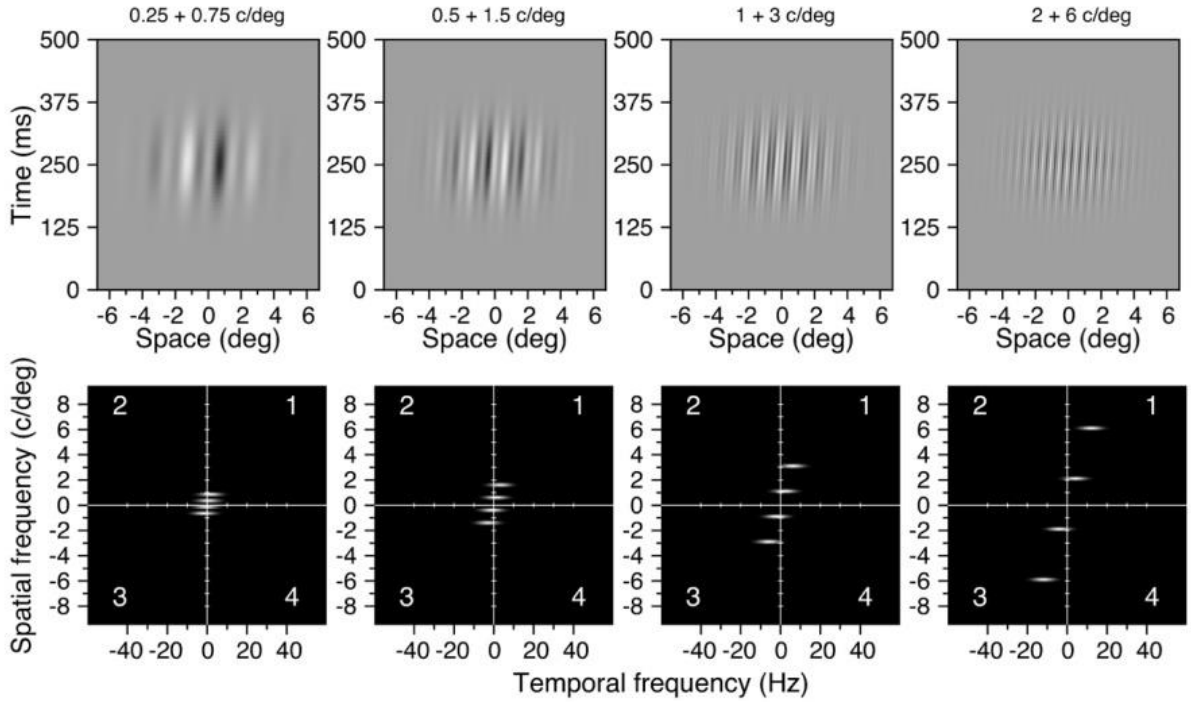

Figure 1. Examples of complex stimuli used in the experiments. The upper panels show the spatiotemporal plots of complex stimuli moving rightwards. In the lower panels the corresponding spatiotemporal amplitude Fourier spectra are displayed. In these examples, the Michelson contrast of each component is 45%, the window diameter is 4 deg ($2 \times \sigma_{xy}$), the speed is 2 deg/sec and the duration of the Gaussian temporal window is 50 msec ($2 \times \sigma_t$).

In Experiment 1 we measured duration thresholds in a forced-choice motion discrimination task. We tested four complex stimuli composed of pairs of spatial frequency patterns that have a 1.58 octave separation: 0.25-0.75 c/deg; 0.5-1.5 c/deg; 1-3 c/deg; and 2-6 c/deg (see examples in Figure 1). Note that the low (LSF) and high (HSF) spatial frequency components in these complex stimuli move at the same time and with the same speed (LSFm+HSFm). We also tested individually the components of the complex stimuli. For these simple stimuli we used two Michelson contrasts (m(t), see equation 5), one of 28% (peak contrast M=0.28, see equation 5) and another of 7% (peak contrast M=0.28/4). For the complex stimuli, we tested three different conditions where different contrast manipulations were made. In the first condition, the Michelson contrast was set to 28% for both components of the complex stimulus (cLSF=cHSF). In the second condition, the lowest spatial frequency components of each pair (0.25, 0.5, 1 and 2 c/deg) had a contrast of 28% while the highest spatial frequency patterns of each pair (0.75, 1.5, 3 and 6 c/deg) had a contrast of 7% (cLSF>cHSF). Finally, in the third condition, the lowest spatial frequency components had a contrast of 7% while the highest ones had a contrast of 28% (cLSF<cHSF).

In Experiment 2 we measured the probability of correct responses for three temporal durations ($2 \times \sigma_t$; 25, 50 and 100 msec) along which the stimuli drifted. We tested the same simple stimuli with the same spatial frequency content used in Experiment 1. Also, we tested the same four complex stimuli used in the previous experiment (composed of the spatial frequency pairs 0.25-0.75 c/deg; 0.5-1.5 c/deg; 1-3 c/deg; and 2-6 c/deg) for the condition where both components had the same contrast (cLSF=cHSF). Contrast manipulations were performed in the same fashion as in Experiment 1 (cLSF=cHSF, cLSF>cHSF, cLSF<cHSF). However, these were only tested for the 1 c/deg and the 3 c/deg spatial frequency components, and the complex stimulus resulting from their addition.

Procedure

In Experiment 1 duration thresholds were estimated. These are the minimum presentation time needed in order for the direction of motion of a drifting stimulus to be discernedwith 82% correct performance. In each trial, a black fixation cross (0.37 × 0.37 deg) was displayed in the center of the screen during a temporal interval of 500 msec. Its luminance was temporally modulated with a Gaussian function that had a standard deviation of 80 msec. After the cross disappeared, the stimuli were presented in an interval of 1000 msec. Their contrast was temporally modulated with a Gaussian function (see equation 5) with a standard deviation (σ_t) controlled by a Bayesian staircase (Treutwein, 1995). Moving stimuli had a fixed speed of v=2 deg/sec. The direction of motion (either leftward or rightward) was randomized in each trial and subjects needed to report it on the ResponsePixx Handheld (VPixx Technologies Inc. Montreal Canada), simply by pressing its right or left button. A new trial started after the response and no feedback was provided about its correctness. The characteristics of the Bayesian staircases were the following: (a) A uniform prior probability-density function for the thresholds was used (Pentland, 1980; Emerson, 1986). This function had a starting duration of 200 msec ($\sigma_t = 100$ msec). (b) The model function used was the logistic one. Its spread value was set to 1 (the delta parameter was 0.01, the lapse rate was 0.01, and the guess rate was 0.5). (c) A Gaussian temporal window determined the the temporal presentation of the stimuli. It was used to control their contrast and their duration (see equation 5). The temporal standard deviation (σ_t) of the Gaussian temporal window in each trial had as its value the mean of the posterior probability distribution (King-Smith, Grigsby, Vingrys, Benes, & Supowit, 1994). (d) The staircase that would provide a measurement of a duration threshold stopped after 40 trials (Anderson, 2003). (e) The mean of the final probability-density function (in log-decimal values) was

taken as an estimation of the final duration threshold (σ_{t0}).. A duration threshold had a temporal duration specified as twice the standard deviation of the temporal Gaussian function ($2 \times \sigma_{t0}$). Finally, from 3 to 6 threshold estimates per condition were obtained for each subject.

In Experiment 2, proportion of correct responses were obtained through the method of constant stimuli using the same forced-choice direction-discrimination task used in Experiment 1. However, in this experiment the stimulus presentation durations were fixed to 25, 50 and 100 msec ($2 \times \sigma_t$, σ_t=12.5, 25, and 50 msec). For each condition and stimulus duration, the proportion of correct responses was calculated across 40 trials. Some stimulus conditions were repeated across the different contrast conditions. In these cases, the proportion of correct responses was averaged across 80 trials.

Results

Experiment 1: Effect of contrast and spatial frequency on motion discrimination measuring duration thresholds

Here we measured duration thresholds for motion discrimination of simple and complex stimuli. In particular, we tested different compound stimuli composed by spatial frequency pairs (i.e. LSF and HSF components) separated by 1.58 octaves. And also, we tested each one of those component spatial frequencies individually. Finally, different manipulations of the contrast for the LSF and HSF components were made. Figure 2 shows the results of the experiment. Each panel displays the duration thresholds as a function of the spatial frequency patterns that compose the complex stimuli. Figure 2A shows the results where both patterns have the same contrast (i.e. cLSF=cHSF); Figure 2B displays the results for the condition cLSF>cHSF; and Figure 2C shows the results for the cLSF<cHSF condition. Generally, in panels 2A and 2B one can see that the duration thresholds for the complex stimuli (red triangles) are higher than the duration thresholds for the HSF (black squares) components. Also, the performance for the LSFm+HSFm stimuli matches the performance for the LSF simple stimuli. Interestingly, although the components of the complex stimuli are presumably processed in parallel by different motion sensors (i.e. the spatial frequencies are set at a 1.58 octaves distance), the visual system does not seem to take advantage of having a motion mechanism that processes the HSF component in a faster manner than the LSF component (i.e. giving rise to a lower duration threshold). Figures 2A

and 2B suggest that the LSF component dominates motion discrimination as performance for the complex stimuli is similar to that for the LSF components. However, a different pattern can be appreciated in Figure 2C. This panel shows that the duration thresholds for the complex stimuli 0.5-1.5, 1-3, and 2-6 c/deg (red triangles) are much higher than the thresholds for the individual components of each compound stimulus (black squares and blue dots). An impairment in motion direction discrimination with respect to each integrating component takes place in agreement with previous results (Serrano-Pedraza et al., 2007). Thus, the hypothesis that there is an inhibitory interaction between motion mechanisms tuned to different spatial scales is supported.

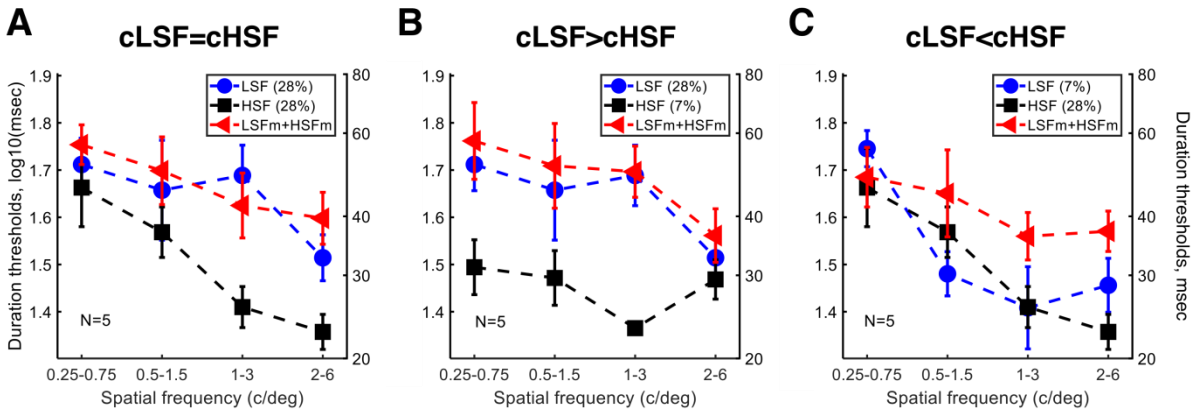

Figure 2. Duration thresholds averaged for all 5 subjects, for simple LSF stimuli (blue), HSF stimuli (black) and compound LSFm+HSFm stimuli (red), for the spatial frequency pairs 0.25-0-75, 0.5-1.5, 1-3, and 2-6 c/deg, represented on the x-axis. Each panel shows different contrast manipulations done on the LSF (0.25, 0.5, 1, and 2 c/deg) and HSF (0.75, 1.5, 3, and 6 c/deg) components. **A:** Both LSF and HSF components have an equal Michelson contrast of 28% (cLSF=cHSF). **B:** The LSF component's contrast is 28% while the contrast of the HSF component has been reduced to 7% (cLSF>cHSF). **C:** The HSF component's contrast is 28% while the contrast of the LSF component has been lowered to 7% (cLSF<cHSF). The mean and the SEM were calculated using the logarithmic values of the duration thresholds. See Appendix A (Figure 1) for the individual data.

Experiment 2: Effect of contrast and spatial frequency on motion discrimination measuring proportion of correct responses

In this experiment we measured the proportion of correct responses in a motion direction discrimination task for simple and complex stimuli presented for durations of 25, 50 and 100 msec ($2 \times \sigma_t$, σ_t=12.5, 25, and 50 msec). In the first part of the experiment, we

tested compound stimuli made out of different pairs of spatial frequency patterns (i.e. LSF and HSF components), and we also tested each of those spatial frequency patterns presented in isolation. The Michelson contrast used was 28% for all stimuli (cLSF=cHSF). Note that this first part of Experiment 2 varies with respect to Experiment 1 in that now we are measuring proportion of correct responses instead of duration thresholds. This way we are able to assess the performance in motion discrimination for different durations rather than only those at the threshold level. To test the effect of different combinations of contrasts, in the second part of this experiment we tested the spatial frequency pair 1-3 c/deg under the same contrast conditions in Experiment 1. These conditions were: one in which both components had a Michelson contrast of 28% (cLSF=cHSF), one in which the contrast of the HSF component was lowered to 7% (cLSF>cHSF) and another one in which it was the contrast of the LSF component that was lowered to 7% (cLSF<cHSF).

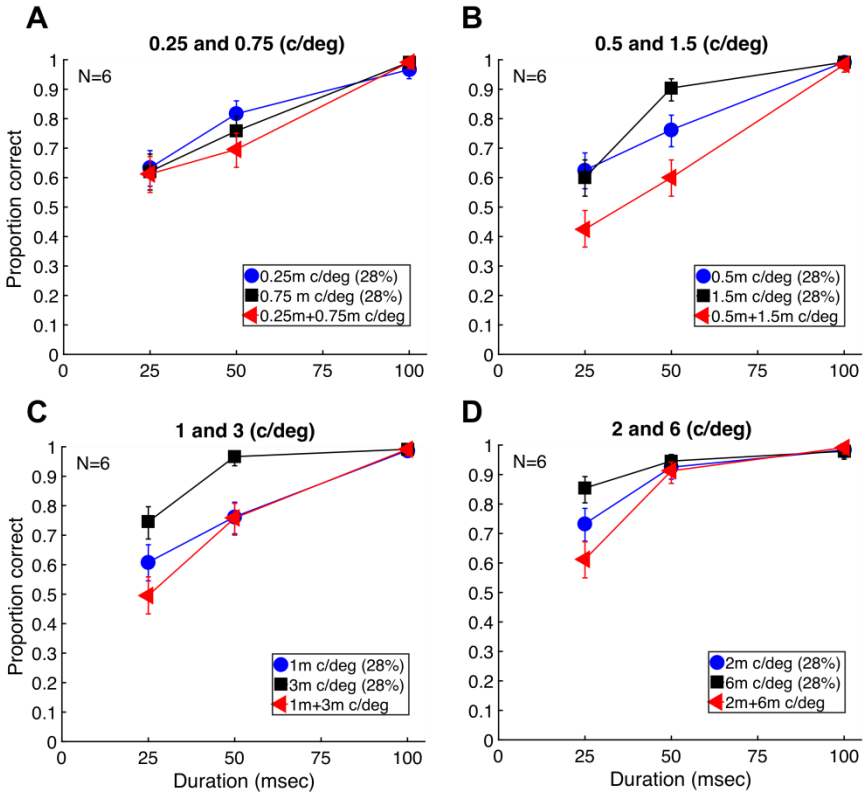

Figure 3. Proportion of correct responses for 6 subjects as a function of the stimulus presentation duration: 25, 50, and 100 msec. Results are shown for the following spatial frequency pairs: **A**: 0.25-0.75 c/deg, **B**: 0.5-1.5 c/deg, **C**: 1-3 c/deg and **D**: 2-6 c/deg. Within each spatial frequency pair, the proportion of correct responses is shown in 3 different scenarios: The case in which only the moving low spatial frequency component is presented (LSF, blue dots), the case in which only the moving high spatial frequency component is presented (HSF, black squares) and the case in which both, low and high spatial frequency patterns, move in the same direction with the same speed, creating a complex stimulus (LSFm+HSFm, red triangles). The Michelson contrast used for all spatial frequency components is 28%. Symbols indicate the mean of the proportion of correct responses while error bars show the 95% score confidence intervals. See Appendix A (Figure 2) for the individual data.

Figure 3 shows the results from the first part of Experiment 2. The panels display the proportion of correct responses as a function of the stimulus presentation duration. Each panel shows the data for different spatial frequency combinations. For the spatial frequency painr, 0.25-0.75 c/deg, the results are somewhat similar for the simple and complex stimuli (see Figure 3A). However, in agreement with our results from Experiment 1, for the pairs 0.5-1.5; 1-3; and 2-6 c/deg, performance becomes impaired for the complex stimuli (red triangles) at short durations (25 msec) with respect to the simple stimuli (see Figure 3B, 3C, and 3D). As expected, in all panels and for every stimulus type, the proportion of correct responses decreases for the shortest durations. The decrease in performance for simple

stimuli at short durations can be easily explained through their spatiotemporal spectra (see Figure 1A). Shorter presentation times not only display motion energy in the actual drifting direction, but also some in the opposite one. This effect vanishes with longer presentation times, for which motion energy becomes restricted to a unique direction of motion; making the task easier. However, the stronger reduction in proportion of correct responses with decreasing duration evidenced for the complex stimuli suggests the existence of an inhibitory interaction between motion mechanisms tuned to different spatial scales.

Figure 4 shows the results coming from the second part of Experiment 2 (apart from subject RLV, who participated in both parts of the experiment, the subjects were different from those participating in the first part). In particular, Figure 4A depicts the contrast condition cLSF=cHSF (same condition as in Figure 3C). One may appreciate how, for the shortest durations (25 and 50 msec), performance for the simple HSF stimulus (3 c/deg) is better than the one for the LSF stimulus (1 c/deg). This is in agreement with the data presented in Figure 2A and Figure 3C. The same as in Figure 3C, performance for the complex stimulus (red triangles) is a bit worse than that for the simple components, at least for the 50 msec presentation duration. Figure 4B shows the results for the contrast condition cLSF>cHSF. In this case, performance for the complex stimulus clearly resembles the results for the LSF component. Provided the existence of two motion mechanisms processing the complex stimulus -one tuned to the LSF component and the other to the HSF component- then, the mechanism processing the LSF component is suggested to be the one mediating motion discrimination.

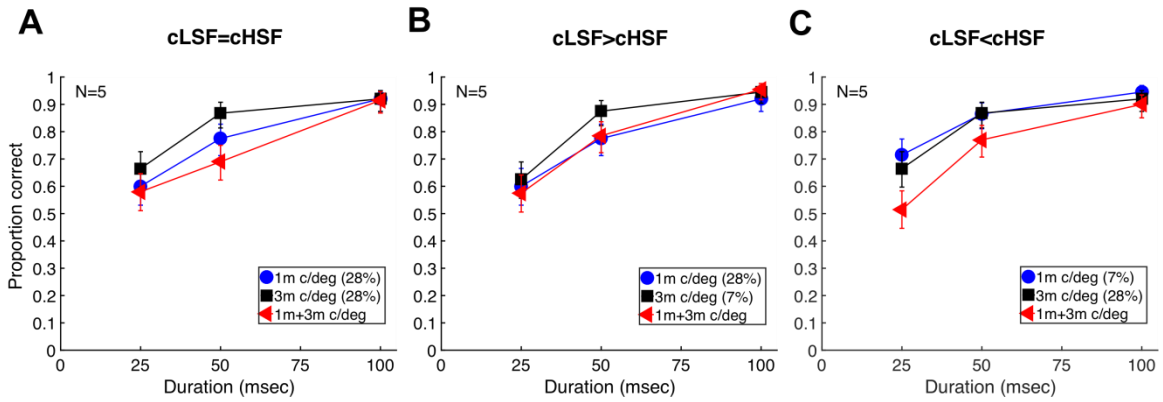

Figure 4. Proportion of correct responses for 5 subjects as a function of the stimulus presentation duration: 25, 50, and 100 msec. Results are shown for the spatial frequency pair 1-3 c/deg. The panels display the results for different stimulus conditions: a moving low spatial frequency stimulus (1 c/deg, blue dots), a moving high spatial frequency one (3 c/deg, black squares), and both, low and high spatial frequency patterns, moving with the same direction and with the same speed (2 deg/sec). This last condition is the one of a moving complex stimulus (1m+3m c/deg, red triangles). Each panel shows the results for different contrast conditions: **A**: All spatial frequency components having a contrast 28% (cLSF=cHSF). **B**: The contrast of the HSF stimulus being lowered to 7% (cLSF>cHSF). **C**: The contrast of the LSF stimulus being lowered to 7% (cLSF<cHSF). Symbols indicate the mean of the proportion of correct responses while error bars show the 95% score confidence intervals. See Appendix A (Figure 3) for the individual data.

On the contrary, for the contrast condition cLSF<cHSF (see Figure 4C), performance is similar for each simple component (i.e. HSF and LSF). Moreover, for the complex stimulus, the proportion of correct responses is lower than the one for each of the simple components at 25 and 50 msec presentation durations. Following our previous reasoning, if there exist two motion mechanisms tuned to LSF and HSF components responding in parallel, then, given that the performance for each component is better than the performance for the compound stimulus, an inhibitory interaction between these two mechanisms must be taking place. This interaction is responsible for the aforesaid impairment seen for the compound stimulus. Additionally, the reported interaction seems to be stronger when both motion mechanisms respond with the same strength. Figures 2C and 4C in fact show that when performance for the LSF and HSF components is similar, then, motion discrimination for the complex stimulus becomes much worse than the performance for each component in isolation.

Model simulations

In this section we test whether the simple model proposed by Serrano-Pedraza et al. (2007) can account for our main results; namely, the impairment in motion direction discrimination of a complex stimulus when the contrast of the low spatial frequency component is reduced. The model is based on the Energy Model described in Adelson & Bergen (1985, their Figure 18b). The motion sensors in our model are tuned to low and high spatial frequencies (i.e. in our case 1 and 3 c/deg). Each motion sensor contains a quadrature pair of spatial weighting functions (e.g. 2D Gabor functions) and includes two quadrature pairs of temporal impulse responses taken from Watson & Ahumada (1985, their equation 12 and 13). The model calculates the oriented energy (left or right) for low (LF) and high (HF) spatial frequency patterns. Therefore, four outputs: L_{LF}, L_{HF}, R_{LF}, and R_{HF}, are generated. The high frequency response (L_{HF} or R_{HF}) is then subtracted from the low frequency response (L_{LF} or R_{LF}) (and vice versa) for the same direction of motion (right or left). After, the resulting outputs become half-wave rectified. The final response is obtained taking the motion sensor with the highest difference between leftward and rightward energy. This highest difference is converted into a direction index and finally, in order to obtain a proportion of correct responses, the direction index is converted into a performance score using a sigmoidal function. A detailed description of the model, including its parameters, can be found in Luna & Serrano-Pedraza (2018, their appendix). For the simulations presented here we have only changed the gain of the spatial sensor with a low spatial frequency tuning. Thereby, the gains of the sensors are 0.9 for the low spatial frequency tuned (c/deg) sensor and 1 for the high spatial frequency tuned (3 c/deg) sensor.

We have simulated the results from Experiment 2 using the same stimuli. In particular, we wanted to compare the results from the simulations with the results presented in Figure 4. The simulation results are shown in Figure 5. Each panel displays the simulation results for each contrast condition and the three stimulus conditions: blue dots (1m c/deg), black squares (3m c/deg), and red triangles (1m+3m c/deg). Our simulations generally show that performance for the single components (blue dots and black squares) and the compound stimuli (red triangles) is impaired at short presentation durations with respect to larger durations. On the other hand, Figures 5A and 5B show that the proportion of correct responses for the complex stimuli (red triangles) are similar to the results obtained for the LSF component (blue dots), although in panel 5A the performance is a bit worse. Figure 5C shows the most interesting result: when the contrast of the LSF component is lower than the

HSF component, then, the performance for the complex stimulus is impaired in agreement with the results presented in Figure 4. Thus, we can say our simple model is able to explain our main findings. According to it, by manipulating the contrast of the stimuli, different imbalances between the motion sensor energies can be obtained. These imbalances further lead to mild or strong inhibitions between sensors.

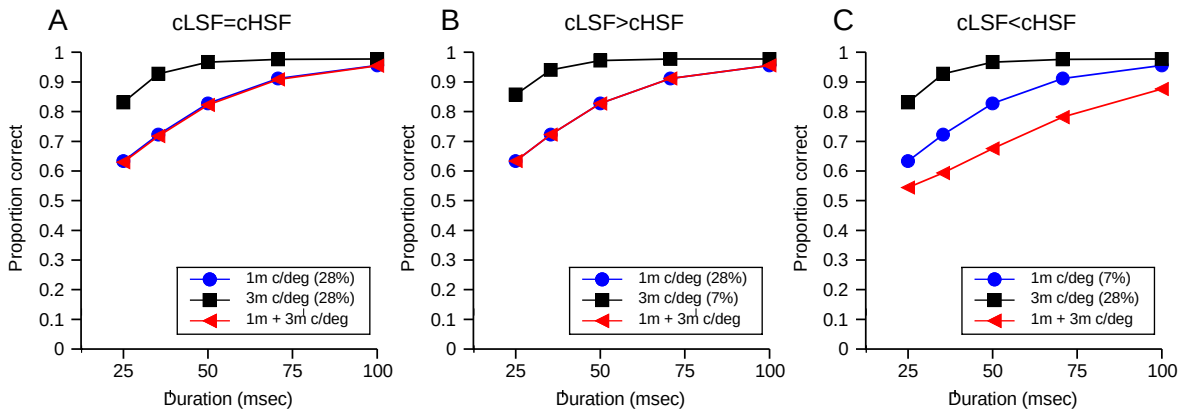

Figure 5. Model simulations for the spatial frequency pair 1-3 c/deg. The figure shows the output of the model for the conditions presented in Figure 4. Each panel shows the results for different contrast conditions: **A**: All spatial frequency components having a contrast 28% (cLSF=cHSF). **B**: The contrast of the HSF stimulus being lowered to 7% (cLSF>cHSF). **C**: The contrast of the LSF stimulus being lowered to 7% (cLSF<cHSF).

Discussion

We know from previous research that systematic errors in motion direction discrimination of a complex stimulus take place when that stimulus is comprised of a moving fine scale component added to a static coarse scale component (Derrington & Henning, 1987). Such effect has been postulated to occur due to an inhibitory interaction between motion sensors tuned to coarse and fine scale patterns (Serrano-Pedraza, et al., 2007). This interaction has been characterized as a function of different stimulus parameters such as presentation duration, size, relative contrast, spatial frequency, and temporal frequency (Derrington & Henning, 1987, Henning & Derrington, 1988; Derrington et al., 1993; Serrano-Pedraza et al., 2007; Luna & Serrano-Pedraza, 2018). Finally, a model which implements the idea of the aforesaid interaction has been successful in replicating some of the psychophysical results obtained in this respect (Serrano-Pedraza et al., 2007; 2010; 2013).

Interestingly, this interaction has been seen to take place even when two components - low and a high spatial frequency ones- move coherently (i.e. same direction and same speed)

(Serrano-Pedraza et al., 2007; Serrano-Pedraza & Derrington, 2010). The experiments cited here were performed using anisotropic noise characterized by Gaussian functions with their maxima at component spatial frequencies of 1 and 3 c/deg. The results obtained showed that performance for the compound or complex stimuli was impaired when compared to the performance of each component in isolation. Therefore, in the present study we have performed two experiments in order to shed light, in a more exhaustive way, on the effect of the spectral content and the relative contrast on motion direction discrimination of complex stimuli. In these, both LSF and HSF components moved with the same direction and with the same speed.

Experiment 1 shows how duration thresholds for the LSF and HSF stimuli, as well as their LSFm+HSFm complex counterparts, vary depending on the spatial frequency content and its contrast. Let us focus on the condition where the contrast of the LSF component was equal the HSF one (cLSF=cHSF) and the condition where the contrast of the HSF component had been lowered (cLSF>cHSF). In these cases, for every spatial frequency pair tested, duration thresholds for the LSF stimulus are close to those of the LSFm+HSFm stimulus. Let us now take the case where the LSF component had its contrast lowered (cLSF<cHSF). On the contrary here, for the spatial frequency pairs 0.5-1.5 c/deg, 1-3 c/deg and 2-6 c/deg, performance for the complex stimulus is impaired with respect to performance for each of its simple components. By manipulating the contrast of the LSF and HSF components we could act upon the energies of the motion sensors respectively tuned to coarse and fine scales (Serrano-Pedraza & Derrington, 2010). This way, we meant to create an imbalance between them. A determined relative contrast combination that caused the maximal interaction between the motion sensors could perhaps be found. In fact, the point of maximal interaction is evidenced when the LSF component has had its contrast lowered with respect to the HSF component. It is in this case where we observe the aforementioned impairment pattern for the LSFm+HSFm stimuli.

As already said, no impairment pattern of the LSFm+HSFm stimulus with respect to each simple component is observed on the cLSF=cHSF and the cLSF>cHSF contrast conditions. However, there is still something interesting to notice about them: Even though there is information in the LSFm+HSFm stimulus about a HSF component (for which performance is good), the visual system seems not to take advantage of it. Instead, performance approximately matches that of the LSF component. It therefore looks like performance for complex stimuli is mediated by the LSF component, as had already been noticed (Hayashi, Sugita, Nishida, & Kawano, 2010; Hayashi, Watanabe, Yokoyama, &

Nishida, 2017). Besides, one can see how at low contrasts, performance for motion discrimination of simple stimuli becomes generally improved. However, this is nothing new provided the already reported action of motion facilitation at low contrasts where duration thresholds decrease with increasing size at low contrasts (Tadin, Lappin, Gilroy & Blake, 2003; Tadin & Lappin, 2005; Glasser & Tadin, 2010; Serrano-Pedraza et al., 2011, Serrano-Pedraza et al., 2013). In our case, given our large fixed size of approximately 4 deg, our results replicate previous ones. That is, performance improves by lowering the contrast.

Experiment 2 has measured the proportion of correct responses in a motion direction discrimination task for different durations. These include presentation durations above the threshold estimated through the Bayesian staircases procedure. The results coming from it converge with the results from Experiment 1 in that Experiment 2 also acknowledges the modulation of the inhibitory interaction by contrast. Like in Experiment 1, the maximal impairment occurs when the contrast of the LSF component is lower than the HSF component (cLSF<cHSF) at brief stimulus presentation durations. We also measured performance for different spatial frequency pairs. However, we only tested for the contrast condition where all components had equal contrast. In this case, one can see how different performances arise depending on the spatial frequency pairs being tested. For example, an obvious impairment for the complex stimuli from the spatial frequency pairs 0.5-1.5, 1-3 and 2-6 c/deg can be observed, but not for the spatial frequency pair 0.25-0.75 c/deg. Bear in mind that these patterns at equal contrast for the LSF and HSF components do not match those obtained in Experiment 1, also at equal contrasts. However, one must consider that by fixing the stimulus presentation time, information different than that gathered using duration thresholds was obtained. The reduced effect for the pair 0.25-0.75 c/deg may be explained by the bandwidth of the underlying channels processing those spatial frequency components. Losada & Mullen (1995) found that the bandwidths (full-width at half-height) for all luminance spatial channels is about 1.3 octaves (suggesting that 0.25 and 0.75 c/deg are processed separately by two channels). However, other evidence shows that the bandwidths of the channels detecting those spatial frequency patterns are wider that 2 octaves (Wilson, et al., 1983; Anderson & Burr, 1985; Solomon, 2000). Therefore, the degree of overlapping of the underlying channels for 0.25 and 0.75 c/deg could reduce the strength of the interaction given that each spatial frequency component could strongly activate both channels. There is also a potential confound caused by the different spatial-frequency bandwidths in octaves of our stimuli (not in linear scale, see the spatiotemporal spectrum in Figure 1). Our stimuli have all the same size, so they present different bandwidths in octaves that could have influenced

our results. However, we think that this is not the case given that previous studies using the same bandwidth in octaves also found a strong interaction (Serrano-Pedraza et al., 2007 and Serrano-Pedraza & Derrington, 2010).

Not strikingly, performance in motion direction discrimination for longer presentation durations is generally better than performance for shorter presentation times (taking a look at the Fourier spectra, shorter presentation times introduce more motion energy in the opposite direction of motion, which makes the task harder). It is worth discussing that in the case of the 100 msec presentation duration, performance for the simple and complex stimuli is fairly similar. More precisely, it generally lies on 100% correct responses. The absence of any measurable trace from the inhibitory interaction for this larger presentation duration can be easily understood through the computational model by Serrano-Pedraza et al. (2007). As already said, short presentation durations display Fourier spectra in which some motion energy also lies in the opposite direction of motion. Conversely, motion energy becomes more restricted to the veridic direction of motion as presentation duration increases. That being said, Serrano-Pedraza et al. (2007) (See the "An explanation of the perceptual reversals using space–time plots and spatiotemporal frequency spectra") suggest that a reciprocal inhibition takes place between those coarsely and finely tuned motion sensors that display the largest responses in presence of a compound stimulus. These sensors are namely the ones sensitive to its veridic direction of motion. Provided short stimulus durations, there is also motion energy in the opposite drifting direction for which uninhibited motion sensors with an opposite direction selectivity are respondent. Because a reciprocal inhibition has taken place between the sensors responding to the veridical motion, this gives rise to misjudgments about the direction of motion of complex stimuli. The fact that no impairment patterns for compound stimuli are present at long presentation durations is a consequence of little or no energy signaling an uninhibited opposite direction of motion. In this sense, the absence of such patterns may not necessarily involve the absence of an inhibitory interaction. Short stimulus presentation durations are just a scenario in which the aforesaid interactions become behaviorally manifested. However, these may take place at very different durations. In fact, they have been behaviorally evidenced at longer presentation times with complex stimuli conformed of a flickering LSF component and a moving HSF one (Serrano-Pedraza et al., 2007).

In short, Experiments 1 and 2 provide complementary information supporting the notion that the inhibitory interaction between motion sensors tuned to coarse and fine scale patterns is modulated by the relative contrast. Also, its action seems to be mediated by the

tuning scale of the motion sensors involved, as evidenced by the differential results depending on the relative spatial frequencies of the components in a compound stimulus. In the Model Simulation section, we have implemented the computational model from Serrano-Pedraza et al. (2007) in order to explain the main findings of our paper. In their paper, this model already predicted that in the compound stimulus condition the proportion of correct responses would be reduced in comparison with the proportion of correct responses obtained for each individual component. Here, we used the same parameters and temporal impulse function of the model described in Luna & Serrano-Pedraza (2018, see their Appendix), but changed the gain of the low spatial frequency tuned sensor (1 c/deg). This model successfully accounts for the main results of Experiment 2, with complex stimuli composed of 1 c/deg and a 3 c/deg spatial frequency patterns which move together with the same speed. Figure 5C shows the stronger impairment (i.e. reduced proportion of correct responses) when the contrast of the LSF component is lower than the HSF component. The same model can also be implemented for other spatial frequency combinations. Different gains of the spatial sensors and probably, other parameters like their bandwidths, or the parameters of the temporal impulse response functions should be changed to reproduce all the results presented here.

Thus, the results described in the present study are in consonance with an inhibitory mechanism for which an interaction between motion sensors tuned to low and high spatial scales is postulated. There is additional evidence supporting the existence of such mechanism. To take a case in point, the motion aftereffect tuning functions with a characteristic asymmetric shape reported by Ledgeway & Hutchinson (2009), or the reverted perceived direction of motion found by Cropper, Kvansakul, & Johnston (2009) in second-order modulations could be explained by the action of the inhibitory mechanism. As already said, the interaction mechanism under study could also shed light on why signals from low spatial frequency components seem to prevail at the time of making judgements on motion direction (Hayashi et al., 2010; Hayashi et al., 2017). On the other hand, our results about an inhibitory interaction between motion sensors could be related to those from Priebe, Cassanello & Lisberger (2003), who found a non-linear behavior in MT cells when exposed to drifting gratings composed of two spatial frequency components. More concisely, they evidenced a shift in the speed tuning of MT neurons towards slower speeds than those predicted by the linear summation of a cell's response to each simple component. Also, performance in motion direction discrimination decreases for slow speeds (Lapin, Tadin, Nyquist & Corn, 2009), with such decrease also being observed for compound stimuli with

coherently drifting components (Luna & Serrano-Pedraza, 2018). That being said, provided a tuning shift towards slower speeds for compound gratings, an impairment in motion discrimination for such stimuli could be a feasible consequence. In addition, our results, and particularly the ones from Priebe et al., (2003) can be further linked to Gekas, Meso, Masson & Mammasian (2017). These researchers found a non-linear behavior between motion channels when these were exposed to a kind of coherently drifting compound stimuli which they termed motion clouds. Namely, the perceived speed for these stimuli was also shifted towards slow ones. In consonance with the tuning shift towards slower speeds in MT cells found by Priebe et al., (2003), Gekas et al., (2017) suggest that the interaction found comprises a neural implementation of a slow speed prior, in agreement with previous research (Jogan & Stocker, 2015; Stocker & Simoncelli, 2006). Nonetheless, caution must be taken when linking the results from the different studies, leaving the possibility open that not all of them could be talking about the same non-linear mechanism. In this sense, Priebe et al., (2003) regard the interaction to be of an excitatory nature, through a feeding from V1 to MT. This makes some MT cells to respond in a speed tuned manner, in consonance with Heeger, Simoncelli & Movshon (1996) and Simoncelli & Heeger (1998). On the contrary, an explanation for our results considers inhibitory interactions (rather than excitatory) between motion sensors tuned to different scales (Serrano-Pedraza et al., 2007; Luna & Serrano-Pedraza., 2018). And also, Gekas et al., (2017) propose lateral inhibition processes as an explanation for their results. Finally, our data show that the inhibitory interaction between motion sensors depends on the spatial frequency of the components. However, the results from Priebe et al (2003) show that the speed-tuning nonlinearity of MT neurons is form-invariant. That is, independent of the spatial frequency of the components.

A parallelism could also be drawn between our interaction and a computation reported by Peronne & Thiele (2002) which achieves a speed tuning in MT neurons. The authors suggest that this speed tuning can be accomplished by taking the absolute value of the subtraction between V1 transient and sustained neurons. More accurately, a particular speed tuning can be best obtained when transient and sustained neurons have a spatial frequency tuning which is slightly shifted respect each other in its frequency peak. This may be reminiscent of our low and high spatial frequency sensor interactions. However, we believe that the spatial frequency tuning displacement in Peronne & Thiele (2002) is too small to match the inhibitory mechanism here under study.

With all this being said, still the ecological function for the interaction mechanism remains fairly unexplored. Possible ecological functions for it arise provided a possible link

between our results and those from Priebe et al., (2003) and Gekas et al., (2017). Namely, the mechanism could be involved in some aspects of speed perception; as is achieving an impression of cohesiveness for complex optic flows integrated by several spatial and temporal frequency components (Gekas et al., 2017). This is in consonance with a sharpening in the speed tuning bandwidth of MT neurons when they are exposed to moving random dot stimuli provided that these group a wide range of coherently moving spatial frequency patterns (Priebe et al., 2003).

Conclusions

Further evidence consistent with a motion sensing inhibitory mechanism through which motion sensors tuned to coarse and fine spatial scales interact is provided. Interestingly, the action of this mechanism is revealed to jointly depend on the spatial frequencies and relative contrast of the spatial components present in a compound stimulus designed to activate both coarsely and finely tuned sensors at the same time. In this regard, contrast is shown to be an important factor modulating the inhibitory interaction, with maximal interactions taking place when the contrast of the LSF component is lowered with respect to the HSF component. This main finding of our study is, indeed, replicated by a simple energy model that includes a reciprocal inhibition stage between motion sensors tuned to low and high spatial frequency components. Finally, the effect of contrast on the mechanism under study varies depending on the spectral content of a given compound stimulus.

Acknowledgments

Supported by grant PGC2018-093406-B-I00 from Ministerio de Ciencia, Innovación y Universidades (Spain) to ISP. RLV is supported by the predoctoral fellowship BES-2015-074077 from Ministerio de Economía y Competitividad (Spain), co-funded by the European Social Fund.

References

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

Adelson, E. H. & Movshon, J.A. (1982). Phenomenal coherence of moving patterns. *Nature*, 300, 5892, 523-525

Anderson, A. J. (2003). Utility of a dynamic termination criterion in the ZEST adaptive threshold method. *Vision Research*, 43, 165–170.

Anderson, S. J. & Burr, D. C. (1985) Spatial and temporal selectivity of the human motion detection system. *Vision Research*. 25(8), 1147-1154.

Anderson, S. J., & Burr, D. C. (1989). Receptive field properties of human motion detector units inferred from spatial frequency masking. *Vision Research*, 29, 1343–1358.

Anderson, S. J., & Burr, D. C. (1991). Spatial summation properties of directionally selective mechanisms in human vision. *Journal of the Optical Society of America A*, 8, 1330–1339.

Anderson, S. J., Burr, D. C., & Morrone, M. C. (1991). Two-dimensional spatial and spatial-frequency selectivity of motion-sensitive mechanisms in human vision. *Journal of the Optical Society of America A*, 8, 1340–1351.

Brainard, D. H. (1997). The Psychophysics Toolbox. Spatial Vision, 10, 433–436.

Burge, J., & Geisler, W. S. (2015). Optimal speed estimation in natural image movies predicts human performance. *Nature communications,* 6, 7900.

Cameron, E. L., Baker Jr, C. L., & Boulton, J. C. (1992). Spatial frequency selective mechanisms underlying the motion aftereffect. *Vision research*, 32(3), 561-568.

Cropper, S. J., Kvansakul, J. G. S., & Johnston, A. (2009). The detection of the motion of contrast modulation: A parametric study. *Attention, Perception, & Psychophysics*, 71, 757–782.

Derrington, A. M., Fine, I., & Henning, G. B. (1993). Errors in direction-of-motion discrimination with dichoptically viewed stimuli. *Vision Research*, 33, 1491–1494.

Derrington, A. M., & Henning, G. B. (1987). Errors in direction-of-motion discrimination with complex stimuli. *Vision Research*, 27, 61–75.

Emerson, P. L. (1986). Observations on maximum likelihood and Bayesian methods of forced-*choice sequential threshold estimation. Perception & Psychophysics, 39, 151–153.*

Gekas, N., Meso, A. I., Masson, G. S., & Mamassian, P. (2017). A normalization mechanism for estimating visual motion across speeds and scales. *Current Biology*, 27(10), 1514-1520.

Glasser, D. M., & Tadin, D. (2010). Low-level mechanisms do not explain paradoxical motion percepts. *Journal of Vision*, 10(4), 20-20.

Hayashi, R., Sugita, Y., Nishida, S. Y., & Kawano, K. (2010). How motion signals are integrated across frequencies: Study on motion perception and ocular following responses using multiple-slit stimuli. *Journal of Neurophysiology*, 103(1), 230-243.

Hayashi, R., Watanabe, O., Yokoyama, H., & Nishida, S. (2017). A new analytical method for characterizing nonlinear visual processes with stimuli of arbitrary distribution: Theory and applications. *Journal of Vision*, 17(6):14, 1–20, doi:10.1167/17.6.14.

Heeger, D. J., Simoncelli, E. P., & Movshon, J. A. (1996). Computational models of cortical visual processing. *Proceedings of the National Academy of Sciences*, 93(2), 623-627.

Henning, G. B., & Derrington, A. M. (1988). Direction-of-motion discrimination with complex patterns: Further observations. *Journal of the Optical Society of America A*, 5(10), 1759-1766.

Jogan, M., & Stocker, A. A. (2015). Signal integration in human visual speed perception. *Journal of Neuroscience*, 35(25), 9381-9390.

King-Smith, P. E., Grigsby, S. S., Vingrys, A. J., Benes, S. C., & Supowit, A. (1994). Efficient and

unbiased modifications of the QUEST threshold method: Theory, simulations, experimental evaluation and practical implementation. *Vision Research*, 34, 885– 912.

Kleiner, M., Brainard, D. H., & Pelli, D. G. (2007). What's new in Psychtoolbox-3? *Perception*, 36 (ECVP Abstract Supplement).

Lappin, J. S., Tadin, D., Nyquist, J. B., & Corn, A. L. (2009). Spatial and temporal limits of motion perception across variations in speed, eccentricity, and low vision. *Journal of Vision*, 9 (1): 30, 1–14,https://doi.org/10.1167/9.1.30.

Ledgeway, T., & Hutchinson, C. V. (2009). Visual adaptation reveals asymmetric spatial frequency tuning for motion. *Journal of Vision*, 9(1):4, 1–9, http:// www.journalofvision.org/content/9/1/4, doi:10.1167/ 9.1.4.

Levinson, E. & Sekuler, R. (1975). The independence of channels in human vision selective for direction of movement. *Journal of Physiology*, 250, 347-366.

Losada, M. A. y Mullen, K. T. (1995). Color and luminance spatial tuning estimated by noise masking in the absence of off-frequency looking. *Journal of the Optical Society of America A, 12,* 250–260.

Luna, R. & Serrano-Pedraza, I. (2018) Temporal frequency modulates the strength of the inhibitory interaction between motion sensors tuned to coarse and fine scales. *Journal of Vision,* 3;18(13):17. doi: 10.1167/18.13.17.

Nishida, S. (2011). Advancement of motion psychophysics: Review 2001–2010. *Journal of Vision*, 11(5):11, 1–53, http://www.journalofvision.org/ content/11/5/11, doi:10.1167/11.5.11.

Nishida, S., Yanagi, J., & Sato, T. (1995). Motion Assimilation and Contrast in Superimposed Gratings: Effects of Spatiotemporal Frequencies. Annual meeting of ARVO, FT Lauderdale, FL

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, 10, 437–442.

Pentland, A. (1980). Maximum likelihood estimation: The best PEST. *Perception & Psychophysics,* 28, 377–379.

Perrone, J. A., & Thiele, A. (2002). A model of speed tuning in MT neurons. *Vision research,* 42(8), 1035-1051.

Priebe, N. J., Cassanello, C. R., & Lisberger, S. G. (2003). The neural representation of speed in macaque area MT/V5. *Journal of Neuroscience,* 23(13), 5650-5661.

Serrano-Pedraza, I., & Derrington, A. M. (2010). Antagonism between fine and coarse motion sensors depends on stimulus size and contrast. *Journal of Vision*, 10(18):18, 1–12, http://www. journalofvision.org/content/10/18/18, doi:10.1167/ 10.18.18.

Serrano-Pedraza, I., Gamonoso-Cruz, M. J., Sierra-Vázquez, V., & Derrington, A. M. (2013). Comparing the effect of the interaction between fine and coarse scales and surround suppression on motion discrimination. *Journal of vision*, 13(11), 5-5.

Serrano-Pedraza, I., Goddard, P., & Derrington, A. M. (2007). Evidence for reciprocal antagonism between motion sensors tuned to coarse and fine features. *Journal of Vision*, 7(12):8, 1–14, http://www. journalofvision.org/content/7/12/8, doi:10.1167/7. 12.8.

Serrano-Pedraza, I., Hogg, L., & Read, J.C.A. (2011). Spatial non-homogeneity of the antagonistic surround in motion perception. *Journal of Vision*, 11(2): 3, 1–9,

http://www.journalofvision.org/content/11/2/3, doi:10.1167/11.2.3.

Simoncelli, E. P., & Heeger, D. J. (1998). A model of neuronal responses in visual area MT. *Vision research*, 38(5), 743-761.

Solomon, J. A. (2000). Channel selection with non-white-noise mask. *Journal of the Optical Society of America A, 17*, 986–993.

Stocker, A. A., & Simoncelli, E. P. (2006). Noise characteristics and prior expectations in human visual speed perception. *Nature neuroscience*, 9(4), 578.

Tadin, D., Lappin, J. S., Gilroy, L. A., & Blake, R. (2003). Perceptual consequences of centre–surround antagonism in visual motion processing. *Nature*, 424(6946), 312.

Tadin, D., & Lappin, J. S. (2005). Optimal size for perceiving motion decreases with contrast. *Vision research*, 45(16), 2059-2064.

Treutwein, B. (1995). Adaptive psychophysical procedures. *Vision Research*, 35(17), 2503–2522.

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

Watson, A. B., & Ahumada, A. J. (1985). Model of human visual-motion sensing. *Journal of the Optical Society of America A*, 2(2), 322-342.

Wilson, H. R., McFarlane, D. K. y Phillips, G. C. (1983). Spatial frequency tuning of orientation selective units estimated by oblique masking. *Vision Research, 23*, 873–882.

Appendix A

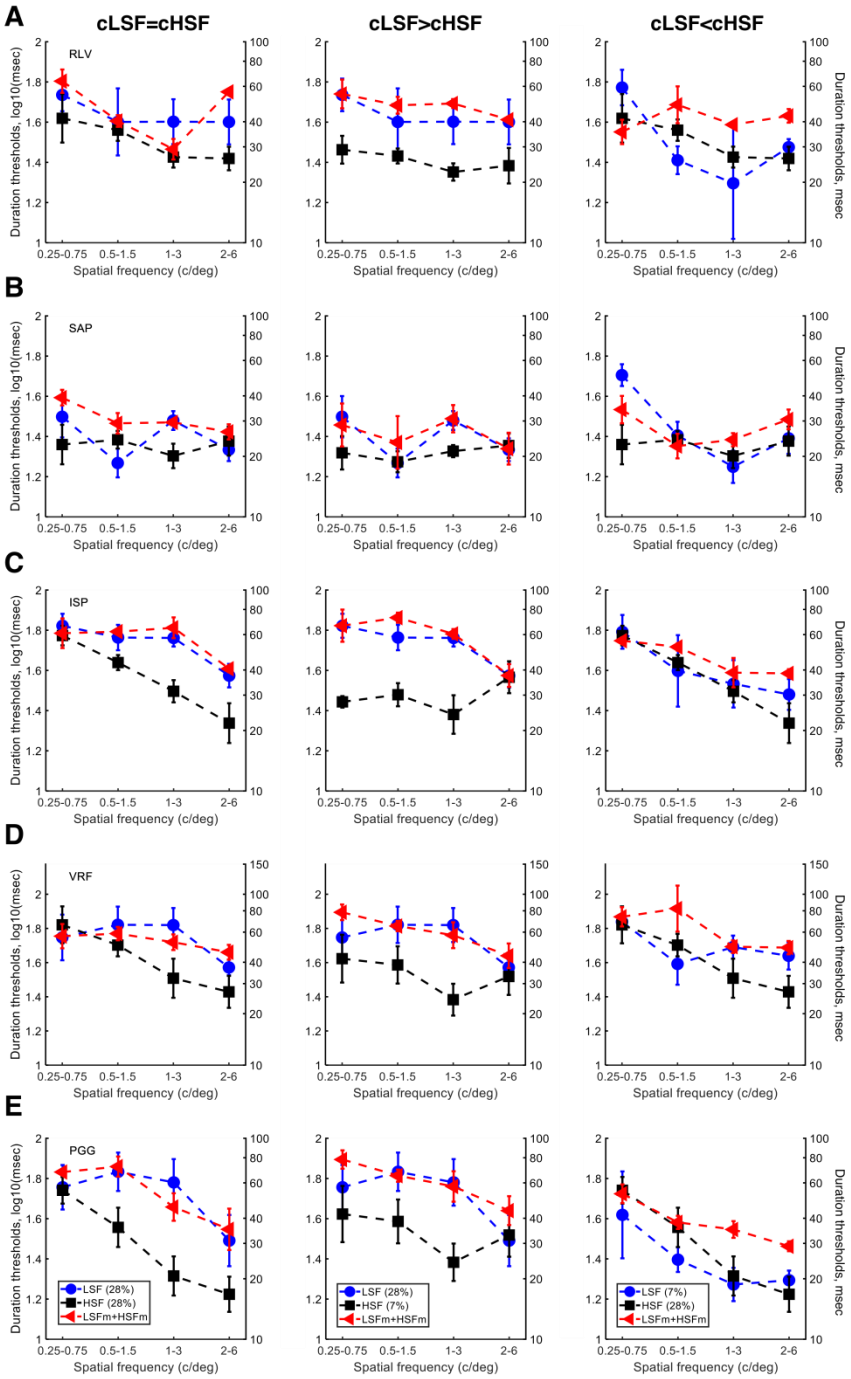

Figure 1. Duration thresholds represented for each of 5 subjects individually (**A**, **B**, **C**, **D** & **E**), for simple LSF stimuli (blue), HSF stimuli (black) and compound stimuli LSFm+HSFm (red). The spatial frequency pairs tested -0.25-0-75, 0.5-1.5, 1-3, and 2-6 c/deg- are represented on the x-axis. Left, middle and right panels show different contrast manipulations done on the LSF (0.25, 0.5, 1, and 2 c/deg) and HSF (0.75, 1.5, 3, and 6 c/deg) components. **Left panels:** Both LSF and HSF components have an equal Michelson contrast of 28% (cLSF=cHSF). **Middle panels:** The LSF component's contrast is 28% while the contrast of the HSF component has been reduced to 7% (cLSF>cHSF). **Right panels:** The HSF component's contrast is 28% while the contrast of the LSF component has been lowered to 7% (cLSF<cHSF). The mean and the SEM were calculated using the logarithmic values of the duration thresholds.

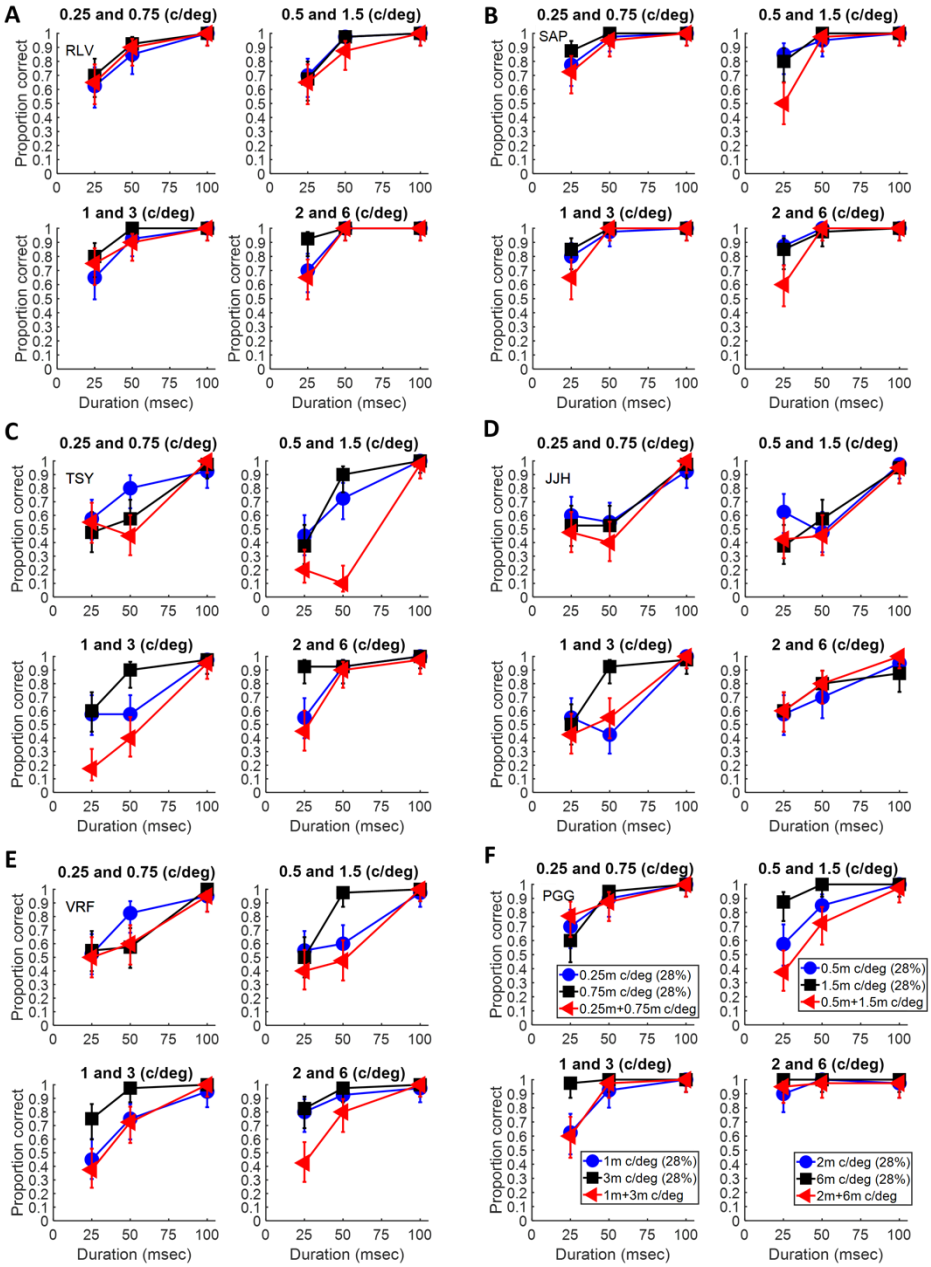

Figure 2. Proportion of correct responses represented for each of 6 subjects individually (**A**, **B**, **C**, **D**, **E** & **F**) as a function of the stimulus presentation duration: 25, 50, and 100 msec. Results are shown for the following spatial frequency pairs: **Top left panels** (within each letter): 0.25-0.75 c/ded, **Top right panels** (within each letter): 0.5-1.5 c/deg, **Bottom left panels** (within each letter): 1-3 c/deg and **Bottom right panels** (within each letter): 2-6 c/deg. Within each spatial frequency pair, the proportion of correct responses is shown in 3 different scenarios: The case in which only the moving low spatial frequency component is presented (LSF, blue dots), the case in which only the moving high spatial frequency component is presented (HSF, black squares) and the case in which both, low and high spatial frequency components, move in the same direction with the same speed, creating a complex stimulus (LSFm+HSFm, red triangles). The Michelson contrast used for all spatial frequency components was 28%. Symbols indicate the mean of the proportion of correct responses while error bars show the 95% score confidence intervals

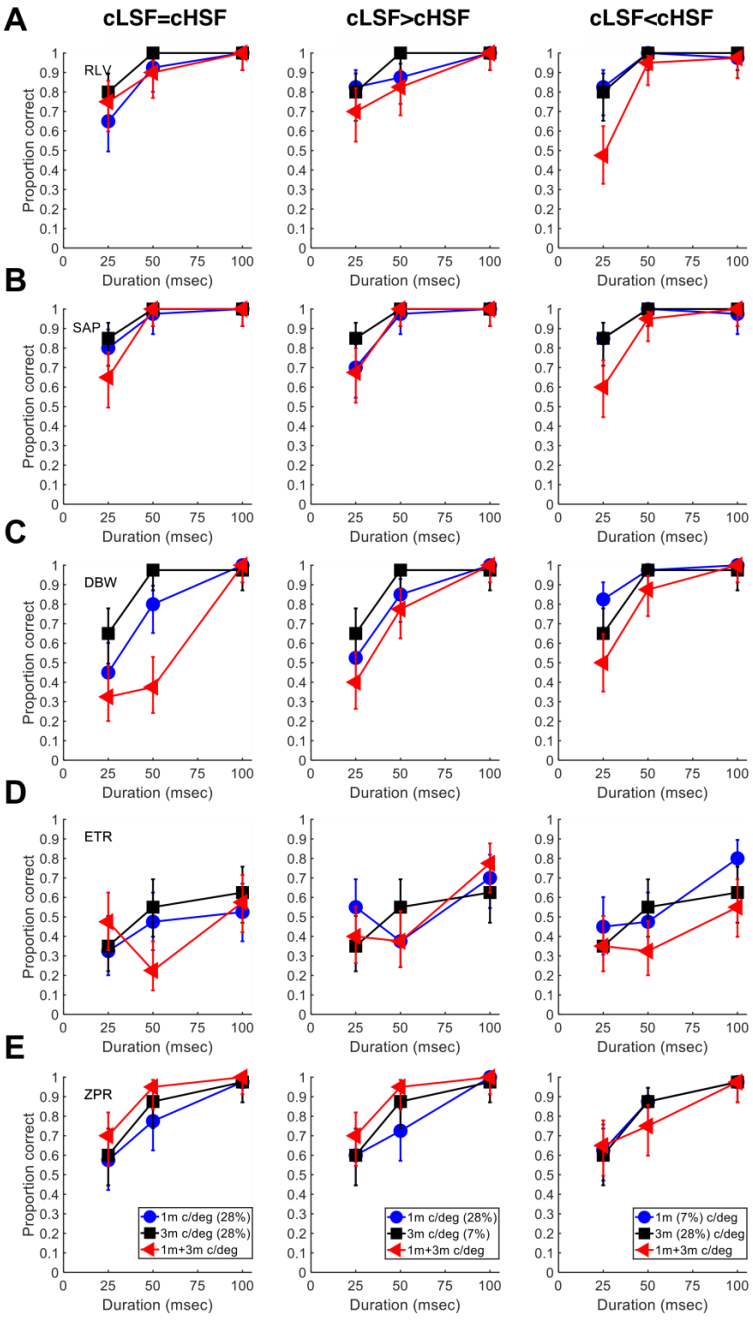

Figure 3. Proportion of correct responses represented for each of 5 subjects individually (**A**, **B**, **C**, **D** & **E**) as a function of the stimulus presentation duration: 25, 50, and 100 msec. Results are shown for the spatial frequency pair 1-3 c/deg. The panels display the results for different stimulus conditions: a moving low spatial frequency component (1 c/deg, blue dots), a moving high spatial frequency component (3 c/deg, black squares), and both, low and high spatial frequency components, moving with the same direction and with the same speed (2 deg/sec). This last condition is the one of a moving complex stimulus (1m+3m c/deg, red triangles). The results for different contrast conditions are shown: **Left panels:** All spatial frequency components having a contrast 28% (cLSF=cHSF). **Middle panels:** The contrast of the HSF component being lowered to 7% (cLSF>cHSF). **Left panels:** The contrast of the LSF component being lowered to 7% (cLSF<cHSF). Symbols indicate the mean of the proportion of correct responses while error bars show the 95% score confidence intervals.