Suppression and facilitation of motion perception in humans with single motion stimulus

Tzvetomir Tzvetanov

Anhui Province Key Laboratory of Affective Computing and Advanced Intelligent Machine, School of Computer Science and Information Engineering, Hefei University of Technology, Hefei, P.R. China; current affiliations: NEUROPSYPHY Tzvetomir TZVETANOV EIRL, France & 赐为科学研究（深圳）有限公司, P.R.China.

Emails: t.tzvetanov@neuropsyphy.io, tzvet@qq.com;

Abstract

Research groups use the single motion stimulus design of Dr. D.Tadin (Tadin, Lappin, Gilroy & Blake, 2003) that allows to putatively assess spatio-contrast excitatory and inhibitory effects from duration threshold data of motion perception. The present work presents the correct neurocomputational model for this experimental design and analyses issues related to data and model comparisons, among which: (1) once the full computational model that predicts the psychophysical results is properly defined, it is shown that two low-level models of how inhibition acts on neuronal activity, i.e. "divisive inhibition" and "subtractive inhibition", predict exactly the same thresholds and cannot arguably be claimed that one is better than another one, (2) it is excitatory and inhibitory processes that are the mechanisms shaping threshold changes, i.e. perceptual "facilitation" and "suppression" in the behavioural domain, (3) that this experimental design allows a quantitative comparison and usage of such "contrast–size tuning" data, (4) that such studies must be carefully designed once the model is correctly understood and applied, because of the rather large parameter space (~10-12 variables) necessary to explain the behavioural measures even in such simple experiments.

Introduction

Changes in perception under experimental manipulations are interesting scientific tools in our search to relate visual perception and its putative neural substrate (Spillmann & Dresp, 1995; Spillmann & Ehrenstein, 1996; Spillmann & Werner, 1996; Eagleman, 2001; Albright & Stoner, 2002; Born & Bradley, 2005) (for recent literature, see Kling, Field, Brainard and Chichilnisky (2019); Pasternak and Tadin (2020)). Human motion perception is known to exhibit various astonishing effects, among which very known in the public domain are motion illusions (for some examples among many websites, see https://michaelbach.de/ot/index.html), with even simple motion stimuli providing sometimes counter-intuitive results.

Concerning the percept of a single motion direction, that is, of an object or a collection of objects (as rain drops or dots) moving in a single unidirectional motion in a fixed spatial location, it is considered to be created from the interaction between excitatory, pooling information across space and directions, and inhibitory, suppressing information across space and directions, interactions between motion sensitive neurons. For example, the percept of random-dot kinematograms/patterns (RDK/RDPs) was reported already for more than 50 years to be strongly shaped by the context in which the target motion was presented (Tynan & Sekuler, 1975; Levinson & Sekuler, 1976; Marshak & Sekuler, 1979; Chang & Julesz, 1984; Watamaniuk, Sekuler & Williams, 1989; Nawrot & Sekuler, 1990). Limiting the topic to the spatial spread of the interactions, two interesting studies (Murakami & Shimojo, 1993; Murakami & Shimojo, 1996) showed that size tuning of these centre-surround contextual interactions in motion perception seems to be eccentricity invariant. That is, when rescaling the measures for different eccentricities all results seem to follow a single curve. This recalls the size tuning characteristics of receptive fields and their magnification with visual eccentricity.
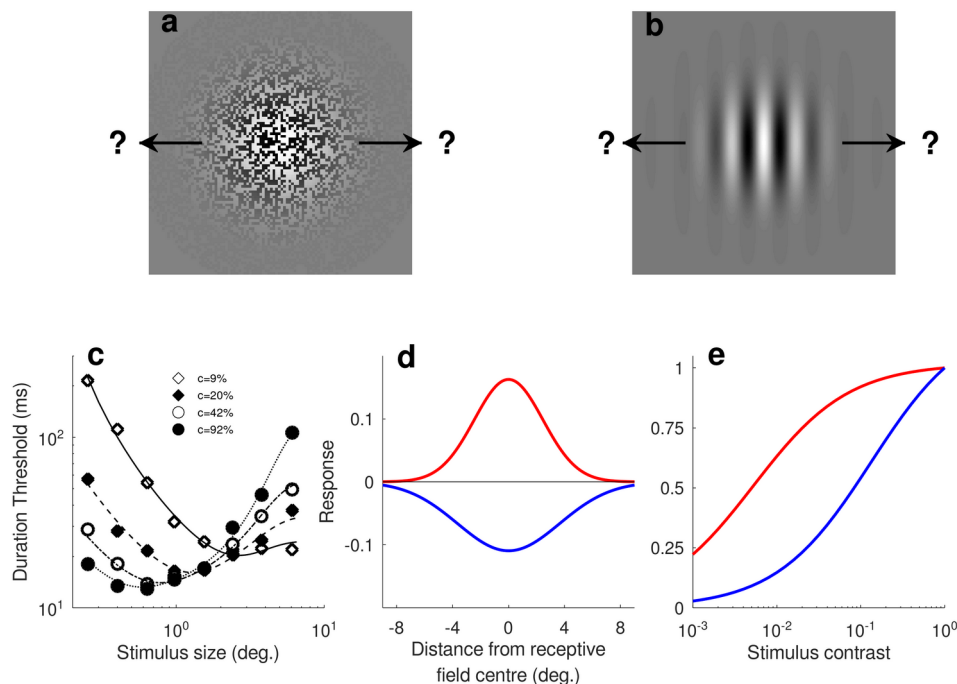
Figure 1: Stimulus examples, example data and model hypothesis. (a,b) Example of stimuli used for measuring duration thresholds of motion perception. (a) A type of random dot kinematogram (RDK) (random light and dark pixels seen through a fixed Gaussian spatial window) whose motion on a given trial could be left- or rightwards (arrows with interrogation mark); remarque: typically only a proportion of the pixels move in one direction from frame to frame, the remaining being random noise (illustration based on the stimulus description in Tadin and Lappin (2005); Tadin et al. (2019)). (b) A Gabor patch stimulus with vertical orientation whose motion could be left- or rightward on a given trial. (c) Report of effects of stimulus contrast and size on duration thresholds (data extracted and replotted from Figure 1A in Tadin and Lappin (2005)). Duration Threshold is the necessary stimulus presentation time in order to correctly perceive its direction of motion. Curves are model fits (see **Modelling** and Equation 21). (d-e) Spatial and contrast tuning characteristics of the receptive field components. (d) illustration of spatial receptive field components, excitatory (red) and inhibitory (blue), with Gaussian profiles. (e) Contrast response functions of excitatory (red) and inhibitory (blue) components of the receptive field.

The authors proposed a simple spatial centre-surround model of receptive fields for explaining their observations.

Yet in another type of psychophysical probes, D.Tadin and colleagues demonstrated that perception of simple moving grating stimuli (Tadin et al., 2003), or a type of RDK (Tadin & Lappin, 2005), have very peculiar perceptual results with respect to size and contrast of the stimuli (see Figure 1a,b for stimulus illustrations). Globally, large and low contrast stimuli needed much less presentation time for being perceived than large and high contrast stimuli, while the opposite was observed for small stimuli. That is, there is an inverted effect on perception as a function of contrast for small and large stimulus sizes (Figure 1c). These observations were attributed to the contrast and size tuning properties of centre-surround motion tuned neuronal populations.

This tuning has a simple form with excitatory and inhibitory components (Figure 1d,e). It has a visuospatial receptive field structure with a relative-to-surround small excitatory centre and spatially larger inhibitory surround (Fig.1d) and contrast responses such that the excitatory component is activated much rapidly at low contrasts (Fig.1e). The final neuronal activity is a combination of the responses of these two drives. It is their precise shape and method to com-

bine them that gives the particular property of the neurons. If one creates stimuli that somewhat match the preferred characteristics of these neurons, then it is considered that these neuronal responses are directly influencing the percept of the stimulus. This gives us a simple way to probe, non-invasively, this particular computational structure of the motion perception system. D. Tadin's insights and results, including later work of collaborators and other researchers, are noteworthy because they showed how we can access the spatial and contrast characteristics of this motion tuned system.

To quantitatively understand these behavioural results, some models were proposed (Tadin & Lappin, 2005; Betts, Sekuler & Bennett, 2012; Schallmo et al., 2018), but they turned out to lack firm neurocomputational grounds (Tzvetanov, 2018) for inferring something about the underlying mechanisms. Furthermore, the present model-based analyses appear necessary because, after a first unsuccessful attempt to peer review and publish this work, my opinion to the editors at that time that more application-oriented studies will use this design seems to be confirmed. A Pubmed search of Tadin et al. (2003) publication gave 21 results for "Cited by" references during the period 2021-2023 (til 31. October), from which about a third to half are comparisons of neurotypical groups (control) versus some groups with differences or correlations with MRS estimates of GABA inhibition (e.g. ageing, depression, adolescence, mental disorders, etc.; including reviews) (e.g. Perani et al. (2021); Song et al. (2021); Ip and Bridge (2022); Liu et al. (2022); Murray and Norton (2023)). Therefore, the foundations for modelling and interpreting the experimental results in this design, and more generally for other features using similar designs, is necessary for researchers interested to say something about the motion system of their groups of subjects, or to explain changes in excitatory/inhibitory effects. After one understands how the model is built from the experimental design and the underlying neurocomputational assumptions, it is easier to interpret, and avoid misinterpretations of, the duration thresholds. This points also to some the difficulties to directly interpret duration thresholds changes between conditions where inhibition or excitation is manipulated (e.g. pharmacologically or between different groups of participants), as now it appears that results are variable across studies (e.g. Ip and Bridge (2022); Murray and Norton (2023)).

In the following, I reanalyse the modelling foundations of this simple experimental design, and demonstrate that it allows to extract information about the putative excitatory and inhibitory processes that shape motion perception in the spatial and contrast domains. First, I recall the psychophysics methods that the modelling must explain. Second, I develop the correct model for predicting motion perception and I make the demonstrations: (1) that such experimental measures and modelling approach cannot dissociate two low-level neuronal models of inhibitory effects, "divisive inhibition" and "subtractive inhibition", because both give exactly the same mathematical prediction for thresholds, (2) that such data can be used for quantitative inference through fitting of excitatory and inhibitory components, and (3) some important consequences concerning what one can claim about low-level models from this design. Furthermore, tests showed impracticability of the current model to quantitatively explain data obtained with grating stimuli.

Results

Psychophysics

Definition of the psychometric function in this design

In this design observers are presented with a single moving stimulus whose contrast, size, presentation duration and direction of motion is varied across trials. The idea is to extract the duration threshold for various stimulus sizes and contrasts. Thus, observers are generally instructed (or learn it during the training before data collection) that (1) on each trial only one of two opposite motion directions can appear (e.g. left- or rightward directions; Fig.1a,b), (2) on some trials the motion may appear so weak, or totally noisy, that it is not clearly perceived and they may need to guess its direction, and (3) they had to report the direction of motion

they thought the presented stimulus had on the trial. These are necessary conditions for measuring the psychometric function in this design, and additionally researchers may, or may not, add more randomness from observer's point of view by measuring thresholds for different contrasts and sizes of the stimulus within the same block. This design is a common "one-stimulus presentation two-alternatives forced choice", let's call it here 1stim2AFC (Morgan, Watamaniuk and McKee (2000) call it Method of Single Stimulus;Kingdom and Prins (2009) call it 1AFC), that can be used for either measuring discrimination thresholds or measuring misperception effects (e.g. motion repulsion,Tzvetanov and Womelsdorf (2008); Tzvetanov (2012); e.g. tilt illusion, Westheimer (1990); Kapadia, Westheimer and Gilbert (2000)).

Here, the experimentalist is interested in the discrimination thresholds of the observers, defined as the necessary duration time of the stimulus for discriminating at some predefined level of correctness (let's say 84%) the direction of motion of the stimulus. It is defined with respect to the midpoint of the psychometric function. The duration of the stimulus is varied between long durations (persons clearly see the motion direction) and short durations (persons have hard time deciding about the direction). The psychometric function spans the two dimensions of stimulus duration (continuous) and direction (binary). Therefore, the resulting data must be analysed as a single psychometric function spanning the full range of left-to-rightward motions (sign of the stimulus) of different durations (intensity of the stimulus). If we assign negative/positive as leftward/rightward motions, absolute intensity as the duration, and percent "rightward" responses as the y-variable one obtains a single psychometric function spanning the full range of 0 to 100% of proportion responses. This definition clearly allows to define also the biases that can appear in this 1stim2AFC design, biases that can be due to response bias in insecure trials (decision to respond with always the same key when not sure; for an explicitly voluntary shift of midpoint see Morgan, Dillenburger, Raphael and Solomon (2012)) or perceptual biases (persons really see something different from "0"). The discrimination threshold is thus $x(p=84\%)-x(p=50\%)$. To extract the thresholds one must take care of the possible biases in at least one of two ways: either through the data analysis procedure described above by fitting a full psychometric function with midpoint and slope (e.g. a logistic of cumulative Gaussian psychometric functions), or either at the experimental methods level by providing feedback to the observers about their response correctness on each trial (e.g. Tadin and Lappin (2005)). This data analysis is important when one has to make model-to-data adjustment, since threshold is related to inverse of the slope of the function.

From these analyses, the threshold data can be used directly to infer something about changes of inhibition and excitation in the visual motion system of the observers. This last point will appear clearly once the model is established and linked to the psychometric function.

About the labels spatial suppression and facilitation

Some literature reports present data obtained from this design by computing a "Size Index" (SI) variable to indicate changes in spatial suppression and facilitation. This variable is defined as the difference of log-thresholds for the smallest and a larger size conditions at a given contrast level (or vice versa). In order for this composite variable to be representative of the concepts of suppression and facilitation it means that changes of thresholds between a small and larger size automatically labels the change as due to only suppression or only summation. This happens in such experiments (Fig.1a-c) when: (1) at very low contrasts inhibition is very weak and thus the results are mainly influenced by spatial summation instantiated by excitatory mechanisms, therefore one should expect a decreasing threshold as a function of stimulus size (see Figure 1c, c=9%); and (2) at very high contrast inhibition is predominant and thus the results are mainly influenced by suppression, therefore one should expect an increasing threshold as a function of size (see Figure 1c, c=92%, but see the two data points for the smallest size). Many psychophysical data do not comply with these rules, which can be seen on a plot of stimulus size versus log-threshold as some "U" shape (Fig.1c, Fig.2), indic-

ating that both excitatory and inhibitory mechanisms act simultaneously in setting observers perceptual thresholds for motion. It is thus re-emphasised that one should directly present and interpret the threshold data. This will appear clearly once the model is presented and how it predicts the data.

Modelling

A note on normalization of neuronal activity

Here, before presenting the exact model, I want to comment on the issue of the "divisive normalization" computational principle/framework that is sometimes advocated (Reynolds & Heeger, 2009; Carandini & Heeger, 2012). It is an important point that concerns how we model perception based on our understanding of neural computations.

How does normalisation appear? Let's look at a general neuronal network with excitatory and inhibitory nodes, respectively noted y_e and y_i. The activity of the excitatory node can be written (similar equation for inhibitory one):

$$\tau_e dy_e/dt \quad = -y_e + F(y_e, y_i, I_e) \tag{1}$$

Here, $F()$ is some function that describes how the nodes are connected and influence each other (through some weights, "firing rate"/"transducer" function, and the inputs to each node I_e and I_i) (Ermentrout, 1998). One general result for neuronal networks is that after a short time of activity the network stabilises itself on a "steady state" (but it could be also a pure oscillatory activity, for a simple introduction to the topic, see Wilson (2005)). What this means is that the activity of a neuron has an initial strong dynamics that typically shows oscillation activity just after input onset (e.g. a burst of firing), and then rapidly stabilises on some steady state activity (for model examples, e.g. Piëch, Li, Reeke and Gilbert (2013), or Wilson (2005)). This steady state activity can (rare) or cannot (in general) be computed in a simple analytical manner, depending on the exact functional form of the above network equations (the function $F()$). This stable point is sometimes considered as a computational principle/framework.

From the above short presentation, it comes natural that neuronal activity stabilisation, aka normalisation, sometimes appearing as "divisive normalization", accounts for the neurophysiological results, but it does so due to the exact functional form of the equations (transducer, excitatory and inhibitory effects...) that the modeller decided to put in it for matching the model to the data. They are carefully considered based on prior knowledge, especially neurophysiological but also behavioural reports. The following two examples of neuronal networks will give different views on the issue.

The first example is a V1-layer 4 cell from the model proposed by Grossberg and Raizada (2000) (their equations 17-18) for explaining contrast and attentional effects observed in neurophysiology of V1. The activity y_{ijk} of these V1-layer 4 cells at equilibrium is:

$$y_{ijk} = \frac{C_{ijk} + \eta^+ x_{ijk} - \sum_{pqr} W^+_{pqrijk} m_{pqr}}{1 + C_{ijk} + \eta^+ x_{ijk} + \sum_{pqr} W^+_{pqrijk} m_{pqr}} \tag{2}$$

where C_{ijk} are processed LGN inputs that are modulated by layer 6 cells activities (x_{ijk}) organised in an excitatory-centre ($\eta^+ x_{ijk}$) and inhibitory-surround ($W^+_{pqrijk} m_{pqr}$) structure, this later being transmitted through a layer 4 inhibitory network activities m_{pqr} (see their equation 19). Leaving aside further considerations about their model, this example shows how nominator and denominator include excitatory and inhibitory effects from various origins. It should be noted that these other feedback activities are themselves function of layer 4's activities and other V1 layers, thus making the exact computation not analytically straightforward and necessitating numerical simulations.

The second example is a model of centre-surround interactions in motion processing taken from Kim and Wilson (1997) (their equation 6). The time dependent activity, C_θ, of one neuron in a centre population sensitive to direction of motion θ is:

$$\tau dC_\theta/dt = -C_\theta + S\Big(E_{C,\theta} - \sum_{i=\theta-120}^{\theta-45} \alpha_i C_i$$
$$- \sum_{i=\theta+45}^{\theta+120} \alpha_i C_i - N f(x) \sum_{k=\theta-60}^{\theta+60} \phi_k A_k \Big) \tag{3}$$

that gives at equilibrium the result:

$$C_\theta = S\Big(E_{C,\theta} - \sum_{i=\theta-120}^{\theta-45} \alpha_i C_i$$
$$- \sum_{i=\theta+45}^{\theta+120} \alpha_i C_i - N f(x) \sum_{k=\theta-60}^{\theta+60} \phi_k A_k \Big) \tag{4}$$

where $S()$ is a simple transducer ($S(x)=x$ if $x>0$, else 0), $E_{C,\theta}$ is the preprocessed pattern motion input to the centre at direction θ, and A_k are the activities of spatially surround neurons that themselves undergo excitation and inhibition in a similar manner (the remaining factors are network connection weights, see their tables 1 and 2). In this example the stabilized activities of the neurons are a simple function of input and subtraction of the inhibitory parts coming from the other cells. The normalisation property here comes from the facts that the input $E_{C,\theta}$ is itself limited between a minimum and a maximum, there are only inhibitory interactions, and the transducer does not allow for negative activities. In this case too, because of the inter-twined interactions through all spatial and motion direction lateral connections, the computation of the stable activities were obtained through numerical simulations.

What did we see from these examples? First, that the exact relation for the neuronal activity at equilibrium can be widely different depending on the exact functional network and connectivities that are assumed by the modeller. The exact terms in the denominator and nominator depend on the network structure the researchers are interested in, based on preliminary knowledge coming from neurophysiology or from behavioural results hinting to a particular structure.

General background for the modelling

Now we can turn to the modelling of the behavioural results in this simple experimental layout. First let's recall the factors of interest in such studies, {*size, contrast, duration*} of the stimulus, and the background hypothesis, that it is the excitatory-centre inhibitory-surround receptive field organisation in area MT/hMT+, which has a typical structure such that similar-to-centre strong motions in the surround reduce the activity due to centre stimulus presentation (Born & Bradley, 2005), that are a direct reason of the perceptual results obtained with motion stimuli.

In predicting psychophysical performance, the models are generally split into two independent stages, the low-level neuronal activities and the high-level decision stage. When combined they must predict the psychometric functions for each combination of size and contrast. The low-level activities are fed to the decision stage that only does one computation – to predict the dependent variable, here percent "rightward" responses.

An important point must be clarified concerning this experimental design and its modelling: the independent variable of the psychometric function is stimulus duration (t_{stim}). Because of this, there is a very strong assumption in the modelling that is made: the neuronal activity, R, of the motion tuned neurons are a direct measure of the duration of the stimulus, analogously to stimulus contrast. That is, there is some monotonic relation between stimulus duration and R, the low-level model of neuronal firing rate. This point is a necessary condition for being able to model the psychometric function that is assumed monotonically increasing with stimulus duration. Let's assume that it has a classic saturating behaviour defined by a hyperbolic ratio equation (Albrecht & Hamilton, 1982) in the time domain:

$$R \propto t_{stim}^n/(t_{stim}^n + t_{50}^n) \tag{5}$$

with $n,t_{50}>0$. This equation states that for an infinitely long presentation of the stimulus the neuronal activity is finite. Furthermore, here it is assumed that this component is multiplicative of the "interaction" components.

Concerning the decision stage model, the two most common, and currently largely advertised and used, models will be considered. The first one is the classic Signal Detection Theory (SDT) (Green & Swets, 1966; Macmillan & Creelman, 2005) applied on neuronal activities (e.g. Britten, Shadlen, Newsome and Movshon (1992)). The second one is the "Drift-Diffusion Model" (DDM) (Link, 1975; Ratcliff, 1978; Luce, 1986; Link, 1992; Smith, 2000; Smith & Ratcliff, 2004; Huk & Shadlen, 2005; Palmer, Huk & Shadlen, 2005; Ratcliff & McKoon, 2008; Forstmann, Ratcliff & Wagenmakers, 2016).

The correct model for D. Tadin's design

We take the low-level processing stage of the model to be represented by a population of neurons coding the two possible motions with opposite directions, whose activities R's depend on all independent variables of stimulus contrast (c), size (s), and duration (t_{stim}). Here, two low-level models, subtractive inhibition (e.g. Tadin and Lappin (2005)) and divisive inhibition (e.g. Betts et al. (2012); Schallmo et al. (2018)), and both decision models (SDT and DDM) are considered, giving as model equation:

$$R_{sub}(c,s,t_{stim}) = R_0 + S\left(E(s,c,t_{stim}) - I(s,c,t_{stim})\right)$$
$$R_{div}(c,s,t_{stim}) = R_0 + \frac{E(s,c,t_{stim})}{\sigma + I(s,c,t_{stim})}$$
$$P_{SDT}(+|c,s,t_{stim}) = \int_0^{+\infty} N\left(\frac{R^+ - R^-}{\sqrt{Var(R^+) + Var(R^-)}}\right) dx \quad (6)$$
$$P_{DDM}(+|c,s,t_{stim}) = \frac{1}{1 + \exp\left(-2A(R^+ - R^-)\right)}$$

with $N(M)$ representing a normal distribution with mean M and unit variance, and R^+ and R^- representing the activity of neurons coding right- and leftward motion directions, respectively; $E()$ and $I()$ are the excitatory and inhibitory drives; R_0 is a "spontaneous firing rate" of the neurons (no input: $c=0$, or $s=0$, or $t_{stim}=0$); $S()$ is the simple transducer $S(x)=x$ if $x>0$, else 0; "+" sign in the psychometric function is defined as the rightward motions; because the experimental design is a 1stim2AFC task with only two possible motion directions, SDT's prediction is based on the difference of activities of the two motion coding populations, and $Var(R)$ is assumed equal to R (Fano factor of one); for the same reason the drift rate μ in the DDM is taken as the difference between the activities of neurons with opposite motion directions.

Important note: in the above Equation 6, activities Rs of the motion tuned neurons may represent the equilibrium state values for the given stimulus input parameters, as discussed in section **A note on normalization of neuronal activity**, or more generally a mean activity across some time window of post-stimulus presentation; starting from equation 6 model presentation can be considered as following the tradition of "pattern analyzers" (Graham, 2011), where it is usual to directly use static mathematical models with predefined pattern sensitive inputs.

What remains to be defined are the excitatory and inhibitory drives $E()$ and $I()$. For the moment we keep a general formulation. Because we assumed that neuronal response to stimulus duration, $R_{dur}(t_{stim})$, is independently pooled from responses to contrast and size, and that it has the same modulation on both components, we can write:

$$E(c,s,t_{stim}) = R_{dur}(t_{stim}) \times R_{exc}(c,s)$$
$$I(c,s,t_{stim}) = R_{dur}(t_{stim}) \times R_{inh}(c,s) \quad (7)$$

To obtain the duration threshold for t_{stim}, from Equation 6 we can write:

$$R(c,s,t_{stim}) = R_0 + F(c,s,t_{stim}) \quad (8)$$

where $F \equiv F(c,s,t_{stim})$ is the general non-specific function of the model. Then, using the SDT model and the definition of threshold as mean differences equal to one standard deviation, we define for a rightward motion input to the model $F^+ \equiv F + R_0$ and $F^- = R_0$, then:

$$F + R_0 - R_0 = \sqrt{F + 2R_0}. \quad (9)$$

After simplification it gives the second order equation and its solution:

$$F^2 - F - 2R_0 = 0$$

$$\Rightarrow \quad F = \frac{1 \pm \sqrt{1+8R_0}}{2} \equiv C_{SDT} \tag{10}$$

In this last equation the negative solution is not biologically meaningful and thus discarded. A similar to the last equation is obtained in the case of the DDM by defining threshold of t_{stim} when $P(t_{stim})$=0.84. Last, with few more mathematical operations one finds for each low-level model considered here that (assuming denominator>0):

Sub: $\quad R_{dur}(t_{stim}^{thr}) = \frac{C_{Decision}}{R_{exc}(c,s) - R_{inh}(c,s)}$

Div: $\quad R_{dur}(t_{stim}^{thr}) = \frac{\sigma C_{Decision}}{R_{exc}(c,s) - C_{Decision} R_{inh}(c,s)}$ (11)

The constants $C_{Decision}$ are dependent on the decision stage models of psychometric functions as follows:

$$C_{Decision} \equiv C_{SDT} = \left(1 + \sqrt{1 + 8R_0}\right)/2$$

$$C_{Decision} \equiv C_{DDM} = (1/2A)\ln(21/4) \tag{12}$$

These mathematical constants are obtained from the full model assumptions and derivations, together with what is defined as threshold for percent responses on the psychometric function (here 84%). Concerning the decision criteria, in SDT case the observer decision criterion is: on a given trial if $R^+-R^->0$ then give response "+", else response "-"; in the case of the DDM the decision criterion is which of the two boundaries is reached (if "+A" then give response "+", if "-A" give response "-"; for further details see the DDM articles and reviews cited earlier).

Making the last step and computing the duration threshold of the theoretical observer, by using equations 5 and 11, one obtains for subtractive inhibition:

$$t_{stim}^{thr} = t_{50}\left(\frac{C_{Decision}}{R_{exc}(c,s) - R_{inh}(c,s) - C_{Decision}}\right)^{1/n} \tag{13}$$

and for divisive inhibition:

$$t_{stim}^{thr} = t_{50}\left(\frac{\sigma C_{Decision}}{R_{exc}(c,s) - C_{Decision} R_{inh}(c,s) - \sigma C_{Decision}}\right)^{1/n} \tag{14}$$

The correct general equations for modelling the psychometric functions of proportion responses as a function of stimulus parameters are given in equations 6, and from considerations about experimental threshold and model parametrisation leads to equations 13-14.

I must re-emphasize that these last equations are the correct model derivations, contrary to some published equations with an unsubstantiated decision stage model, and thus final equation (Tzvetanov, 2018).

Comparing predictions of the two low-level models

Now that the model derivations have been laid down, we can analyse whether these equations make interesting inferences about a particular model, or point to changes in thresholds as a function of excitatory or inhibitory strengths. Because the decision stage of the models make predictions that differ only in the constant $C_{Decision}$, from now on, any application of a full model will be done with SDT. Furthermore, without loss of generality, we fix in the divisive inhibition model σ=1 (amplitude rescaling of $E()$ and $I()$ in Equation 6).

The two models give the predictions of duration threshold with equations 13-14. Interestingly, the condition C_{SDT}=1, which happens when R_0=0, and assuming that R_{inh} for both divisive and subtractive models are the same as well as for R_{exc}, gives exactly the same final prediction from both models. It is independent from the particular functional form of the excitatory and inhibitory drives. That is, despite the very different initial model assumptions (Equations 6), one finds that the final prediction of threshold variation as a function of stimulus size and contrast are equal.

When R_0>0, the exact difference between the two predicted thresholds depends on all three variables of excitatory, inhibitory and spontaneous firing rates. Nevertheless, even when R_0>0, one can see that if the two models differ only in the inhibitory responses R_{inh}, when $R_{inh}(\text{Sub})=C_{SDT}R_{inh}(\text{Div})$ both models still give exactly the same prediction of thresholds

through a simple amplitude scaling of the inhibitory drives. Therefore, if the functions for excitatory and inhibitory effects in the two low-level models are chosen to be of the same mathematical form, then the two models can predict exactly the same thresholds by a proper rescaling of inhibitory amplitude.

To conclude, selecting either of the low-level models in Equation 6 is irrelevant for the final predictions. Consequently, it is the mathematical forms of the excitatory and inhibitory drives that predict the perceptual thresholds and this simple experimental design cannot dissociate between low-level models of divisive inhibition and subtractive inhibition.

On the necessity of centre excitatory tuning width to vary with contrast

Here is analysed the question of what kind of excitatory and inhibitory drives can, or can not, predict some observations in the behavioural results. One noteworthy point from Equations 11-14 is that they allow simple and interesting inferences, inferences that, to the best knowledge of the author, for the particular design of D. Tadin, were missing until now.

(i) Independence of contrast and size effects can not predict behavioural results: simple case

If we make the assumption of independence of stimulus contrast and size effects on each drive, $R_{exc}(c,s)=R_{c,exc}(c)R_{s,exc}(s)$ (similar equation for inhibition), and that inhibitory contrast response function, $R_{c,inh}(c)$, is a scaled version of the excitatory one, that is, $R_{c,inh}(c)=kR_{c,exc}(c)\equiv kR_c(c)$, for all models (from equ.11) one obtains:

$$R_{dur}(t^{thr}_{stim}) \propto 1/R_c(c) \qquad (15)$$

That is, because of the monotonic relation for each function on both sides of the above equation (both $R_{dur}(t_{stim})$ and $R(c)$ are monotonically increasing with their variable), duration thresholds are inversely related to contrast for any size of the stimulus. In fact, the behavioural results for small stimulus sizes follow this relation (see Figure 1), but not for large stimulus sizes. Thus, independent effects of stimulus contrast and size cannot predict the psychophysical observations of inverted relation between duration thresholds and contrast at large stimulus sizes (an effect present for various stimulus types). We can discard this possibility as incompatible with behavioural evidence. This result is a nice counterpart of the neurophysiological findings in low-level visual processing systems, that stimulus contrast and size are somehow strongly intertwined in neuronal responses of areas V1 and MT (Sceniak, Ringach, Hawken & Shapley, 1999; Cavanaugh, Bair & Movshon, 2002; Tsui & Pack, 2011).

(ii) Independent contrast and size effects in the model: different, non-scaled, contrast tuning of excitatory and inhibitory components

Here is further analysed the condition of independent effect of contrast and size. Now it is assumed that $R_{c,inh}(c) \neq kR_{c,exc}(c)$, for example a simple shift of the inhibitory contrast response toward higher contrasts, and we assume that the components describing contrast responses are normalised such that their responses at $c=1$ is equal to one.

Then we can predict the inverted effect of contrast on duration thresholds as a function of stimulus size. The behavioural data show a very strong variation with contrast of the "minimum" (or dip) of the duration threshold as a function of stimulus size. In the data of Tadin and Lappin (2005) this minimum changes from ~2-3 degrees at c=9% down to ~0.5-0.6 degrees at contrasts 42% and 92% (see Figure 1c), a factor change of ~3-6. To demonstrate that this change is hardly predicted with independent contrast and size effects, we take for the size functions the error functions:

$$R(c,s) = R_{c,exc}(c)a_e\text{erf}(\frac{s}{\sigma_e}) - R_{c,inh}(c)a_i\text{erf}(\frac{s}{\sigma_i}) \qquad (16)$$

Because the dip of the effect happens when the response $R(c,s)$ reaches a maximum on the size dimension, we can compute the stimulus size at which the dip appears, s_{dip}, which gives:

$$s_{dip}(c) = \left(\frac{\sigma_e^2\sigma_i^2}{\sigma_i^2 - \sigma_e^2} \ln \left(\frac{R_{c,exc}(c)}{R_{c,inh}(c)} \frac{a_e}{a_i} \frac{\sigma_i}{\sigma_e} \right) \right)^{1/2} \qquad (17)$$

From this result, it is possible to find the necessary ratio of excitatory to inhibitory low-contrast responses for observing a given dip change as a function of the ratio of high-contrast responses. By using the ratio between size dips at low and high contrasts, $r_{dip}=s_{dip}(lc)/s_{dip}(hc)$ (lc – low contrast; hc – high contrast), it is found that:

$$\frac{R_{c,exc}(lc)}{R_{c,inh}(lc)} = \left(\frac{a_e}{a_i}\frac{\sigma_i}{\sigma_e}\right)^{r_{dip}^2-1}\left(\frac{R_{c,exc}(hc)}{R_{c,inh}(hc)}\right)^{r_{dip}^2} \qquad (18)$$

If we take the best case of $R_{c,inh}(hc)=R_{c,exc}(hc)$, e.g. at contrast of one the normalised contrast coding components are equal, and $a_e=a_i$ (maximum surround inhibition), then we have:

$$\frac{R_{c,exc}(lc)}{R_{c,inh}(lc)} = \left(\frac{\sigma_i}{\sigma_e}\right)^{r_{dip}^2-1} \qquad (19)$$

That is, the ratio of excitatory-to-inhibitory low-contrast responses is a function of the ratio inhibitory-to-excitatory size tunings to the power of squared r_{dip} ! This model can predict factor of dip changes of ~3–6 only if the inhibitory responses at low contrasts are essentially zero. This would correspond to a strongly expansive inhibitory contrast response function at intermediate to high contrasts.

To conclude, the model of low-level responses cannot have independent contrast and size effects, as argued by Tadin and Lappin (2005) from their fitting results, and reported neurophysiologically for the V1 visual system (Sceniak et al., 1999; Cavanaugh et al., 2002).

(iii) Necessity of centre excitatory tuning width to vary with contrast

One can further analyse which single component, excitatory or inhibitory, influences the most the position of the dip. We can use Equation 17 (still valid when size tuning is contrast dependent) and rewrite it as:

$$s_{dip}(c) = \sigma_e\left(\frac{k_s^2}{k_s^2-1}\ln\left(\frac{R_{c,exc}(c)}{R_{c,inh}(c)}\frac{a_e}{a_i}k_s\right)\right)^{1/2} \qquad (20)$$

where $k_s=\sigma_i/\sigma_e$ is the ratio of inhibitory-to-excitatory sizes. From this equation it is much easier to see that, at fixed ratio k_s, the dip position is linearly related to σ_e, while at fixed σ_e it is much slowly changing with σ_i. Thus, the behavioural data of duration thresholds showing large changes of the dip position can only be explained if excitatory centre tuning width is contrast dependent.

Application of the model

Until now, it was possible to make model analyses based on general considerations about the functions and observed data. Because the model predicts exact thresholds, and the deduced final equations 13-14 can be fit to the data, now we have to choose specific mathematical forms for the different components. By using Tadin & Lappin's (2005) low-level model, based on prior knowledge, and writing the full equation of the model, we have:

$$t_{stim}^{thr} = t_{50}\left(\frac{C_{Decision}}{R_{exc}(c,s)-R_{inh}(c,s)-C_{Decision}}\right)^{1/n}$$
$$R_{exc}(c,s) = a_e\frac{c^p(1-c_{1/2,e}^p)}{c^p(1-2c_{1/2,e}^p)+c_{1/2,e}^p}\mathrm{erf}\left(\frac{s}{\left(\sigma_e/(1+me^{-k/c})\right)}\right) \qquad (21)$$
$$R_{inh}(c,s) = a_i\frac{c^p(1-c_{1/2,i}^p)}{c^p(1-2c_{1/2,i}^p)+c_{1/2,i}^p}\mathrm{erf}\left(\frac{s}{\sigma_i}\right)$$

where $\{a_e,a_i\}$ are the absolute response amplitudes of the excitatory and inhibitory components, the middle, complex-looking, ratio of contrasts in $R_{exc}()$ and $R_{inh}()$ is the classic hyperbolic ratio for contrast response functions(Albrecht & Hamilton, 1982) but rewritten such that at $c=1$ its response is 1 and at $c=c_{1/2}$ its response is 1/2, and the last part is the error function as a function of size (for the inhibitory part) and size and contrast (for the excitatory part) (Tadin & Lappin, 2005). In the above equations it is assumed that the pure contrast responses for both excitatory and inhibitory drives have the same exponent p but different half-amplitude constants $\{c_{1/2,e}, c_{1/2,i}\}$. The size function, *erf*(), assumes that both receptive field components have a circular Gaussian shape and the space information is simply pooled across the stimulated area (stimulus centred on the receptive field). The model contains a total of 12 paramet-

ers: $\{t_{50}, n, C_{Decision}, a_e, a_i, p, c_{1/2,e}, c_{1/2,i}, \sigma_e, \sigma_i, m, k\}$. Without loss of generality, n and a_e are set to 1, and parameters $\{C_{Decision}, a_i\}$ are interpreted as proportional to the excitatory amplitude a_e. This leaves a total of 10 free parameters.

(i) Example of model fit

A demonstration of fit to the data of Tadin and Lappin (2005) is carried. There are enough data points (n=32) on a sufficiently large range of contrasts and sizes for constraining all 10 parameters. Fit results are: $\{t_{50}=197, C_{Decision}=0.051, a_i=0.876, p=0.75, c_{1/2,e}=0.0044, c_{1/2,i}=0.081, \sigma_e=2.5, \sigma_i=3.6, m=7.78, k=0.305\}$ (see continuous curves in Figure 1c and compare to the data). Although some parameters seem to have stronger influence on a single input dimension (e.g. σ_i on size only), it is the 3D data, *Duration Threshold* vs {*Contrast, Size*}, that constrain all parameters simultaneously through the final non-linear model of Equation 21. Some interesting quantitative points can be derived from the final parameters: (1) inhibitory amplitude is quite strong, nearly 90% of the excitatory amplitude, which is a nice counterpart of neurophysiological results showing that short duration motion stimuli seem to preferentially stimulate neurons with strong surround inhibition (Churan, Khawaja, Tsui & Pack, 2008) ; (2) the constant $C_{Decision}=0.051$ gives, if $a_e=100$, $R_0\sim10$ which is about 10% of a_e, a rather common spontaneous firing rate of neurons; (3) the excitatory half-amplitude constant $c_{1/2,e}$ is low, an interesting counterpart of neurophysiological observations for similar centre-surround designs showing low semi-saturation constants (Tsui & Pack, 2011) ; (4) the power of the contrast response function is very low ($p=0.75$), which was also reported to be the case in neuronal fits of size tuning data of MT (Tsui & Pack, 2011).

In summary, the correct computational model fit to the behavioural data provides quantitative parameters that globally match expectations about parameter values gathered from neurophysiological studies. Thus, one can apply the model on behavioural data to infer putative changes of inhibitory or excitatory mechanisms under experimental manipulations.

(ii) Manipulating individual observer's level of inhibition

What should we expect as an outcome if inhibition within a person is manipulated or is naturally different between two groups of populations? In the model the effects of the inhibitory component are instantiated through three model functions, one of amplitude, one of contrast, and one of size, with their respective parameters $\{a_i, c_{1/2,i}, \sigma_i\}$. The effect of enhanced inhibition can affect all three functions. Thus, there are three simple ways to obtain stronger inhibition: an increase in amplitude a_i, a decrease in contrast half-amplitude constant $c_{1/2,i}$ such that inhibition is activated faster at lower contrasts, and a decrease in tuning width σ_i such that inhibition is activated earlier in spatial domain. Three simple examples are depicted in Figure 2a-c. As one can see, amplitude increase slows very strongly thresholds at large stimulus sizes, inhibitory size decrease affects thresholds for small to medium size stimuli, and half-amplitude constant ($c_{1/2,i}$) decrease affects medium-to-high changes thresholds at medium-to-large size stimuli. One important result is that all three changes increase duration thresholds. This can easily be understood by referring to equations 13-14, where one can see that increasing inhibition ($R_{inh}(c,s)$) decreases the denominator which in turn increases the overall bracket. That is, while the exact threshold variation is dependent on the combined changes of inhibitory parameters, increase in inhibition makes overall thresholds higher.

(iii) Measures of spatial suppression and summation with sine gratings do not conform to the model

The data from the RDKs of Tadin and Lappin (2005) presented in Figure 1c are well fit by the model. In an attempt to explain also other data, it was found that the model has difficulties in fitting them. These data were obtained with moving sine grating stimuli, that were seen through Gaussian envelopes (Gabor patches). The fitting was unsuccessful in matching the model to the data. When the 10 parameters were left free the final best fit gave biologically not plausible values; even with these best parameters, the fit was still difficult to reconcile
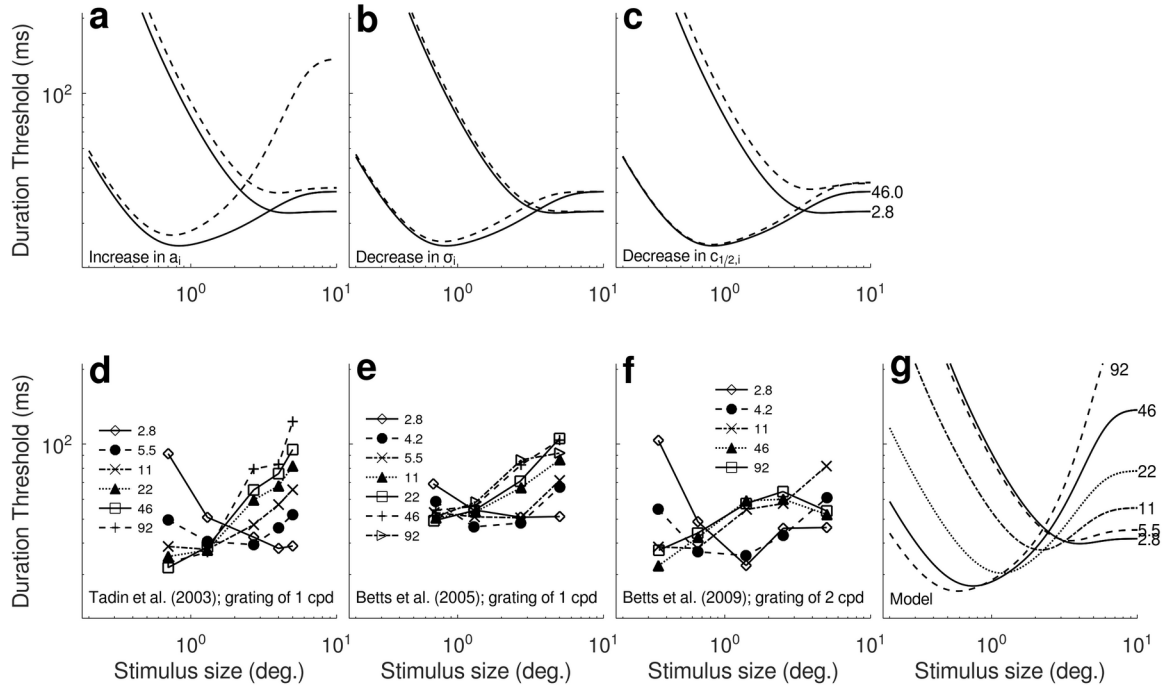
Figure 2 Effects on thresholds when manipulating inhibition (a-c) and duration thresholds data from reports using sine grating type of stimuli (d-f). (a-c) Model prediction for parameters $\{t_{50}=400, C_{Decision}=0.05, a_i=0.5, \sigma_e=2.5, \sigma_i=4, p=0.75, c_{1/2,e}=0.005, c_{1/2,i}=0.08, m=8, k=0.3\}$ (continuous curves) and change of $ai=0.9$ (a), $\sigma_i=3$ (b), and $c_{1/2,i}=0.02$ (c); depicted are changes for stimulus contrasts of 2.8 and 46 for ease of visualisation; (d) Gabor patches results from Figure 1 in Tadin et al. (2003), (e) Gabor patches results for young observers from figure 2 in Betts, Taylor, Sekuler and Bennett (2005), (f) Gabor patches results for young observers from Figure 1 in Betts, Sekuler and Bennett (2009) and (g) model (Equation 21) prediction for parameters $\{t_{50}=400, C_{Decision}=0.05, a_i=0.9, p=0.75, c_{1/2,e}=0.005, c_{1/2,i}=0.08, \sigma_e=2.5, \sigma_i=4, m=8, k=0.3\}$. In (g) contrasts are printed on the rightmost end of each curve.

with the data. The reason of this discrepancy can be seen in Figure 2d-f. Panels (d-f) replot the data from three studies that used moving sine gratings at different contrasts and sizes. As it can be seen, the grating data has a particular behaviour: at low contrasts (~3%) the thresholds have the typical decrease with increasing stimulus size; at slightly higher contrasts (4.2-5.5%) the data show a U-shape with increasing stimulus size; from contrasts above 10% the thresholds exhibit mainly monotonic increasing behaviour with possible plateau effect at very high contrasts (panel (f)). Panel (g) presents the model predictions. It can be seen that the model shows a typical U-shape of thresholds vs. stimulus size for all but the lowest contrasts of 2.8-5.5%, where it is very weak. One can use Equation 20 to see that at sufficiently high contrasts one has approximately:

$$s_{dip}(c) = K\sigma_e(c) \qquad (22)$$

with K approximately constant (k_s increases with c because σ_e decreases with c, and the ratio of contrast functions slowly decreases toward 1, which makes the bracket in Equation 20 to increase very slowly with increasing contrast, in opposite direction of s_{dip} variation). Thus, one has the choices: (i) to change the function relating excitatory centre size and contrast such that at high contrasts it drops much rapidly as in the data, or instead (ii) suppose that the valid fit in Figure 1c for RDK stimuli (Fig.1a) hints to model inadequacy (equ.21) for explaining results with grating type of stimuli. (one can also see the papers of Tadin & Lappin, 2005 and Betts et al., 2012, for different contrast and size low-level models, though the threshold equation is wrong).

Conclusion from modelling

From the above modelling presentation a small summary is made. First, the computational model predicting the duration thresholds for this simple experimental design of motion perception is easy to obtain. From careful considerations about its basic assumptions and functional form of the neuronal activities, it was shown that two competing models of neuronal activities, i.e. divisive inhibition and subtractive inhibition, cannot be dissociated in this design. Instead, it is the exact mathematical form of the excitatory and inhibitory drives that defines the shape of suppressive and facilitative effects on duration thresholds as a function of size and contrast. Second, the computational model was shown to provide quantitative estimates of parameters that globally matched reported neurophysiological counterparts, and issues related to interpreting changes in measures were highlighted.

Discussion

This work analysed the spatial summation and suppression effects that appears in human perception of visual motion, specifically in the very simple stimulus design of D. Tadin (Tadin et al., 2003). It was prompted from a missing correct model derivation and application. Therefore, this study presented the model for predicting the perceptual thresholds in the experimental design and made multiple inferences from it that were presented in the relevant parts. Here are discussed some points about psychophysics and modelling that were not touched upon.

When testing computational models of perception, we match the model prediction to the behavioural measures obtained in a particular experimental design through the use of the psychometric function (PF). Its exact definition depends on the experimental design. In the 1-stimulus-2-Alternatives-Forced-Choice design analysed here, it was argued that the PF should be defined as a function of motion duration and direction, thus representing a typical PF for "discrimination" (discriminating between two possible motion directions) that can be defined in the full range of percent responses (0% to 100%). The current work allowed to model the PF by considering the inputs and outputs that the model should incorporate and predict. It provided the general model equation, where the low-level model responses are presented together with the decision stage levels, that led to predict the perceptual thresholds. The model approach in this work is based on "static models", or "pattern analyzers" (Graham, 2011). It approximates the responses of the neuronal populations to the inputs in specific ways by assuming that all neuronal interactions lead to simple "static" functional forms.

First, the major independent variable stimulus duration was modelled as activating neurons independently from the two other variables of stimulus contrast and size. The results show that it is a sufficient assumption. Furthermore, it was hypothesized that the exact way this variable affects neuronal responses is to modulate independently both excitatory and inhibitory drives. On the contrary, one can hypothesize that the effect is to modulate the combined response of excitatory and inhibitory components. While the "subtractive inhibition" model is not affected by this difference, the "divisive inhibition" model gives different predictions (not shown here).

Second, the model incorporates the "spontaneous firing rate" of the neurons, R_0, as the minimum possible activity. Nevertheless, spontaneous activity of the neurons are known to be lower for surround-suppressed cells in comparison to non-suppressed cells (Churan et al., 2008). This hints to the possibility that, when a moving stimulus is presented to them, the activity of the neurons sensitive to opposite motion directions is in fact lower than R_0, an effect already reported (Snowden, Treue, Erickson & Andersen, 1991; Britten, Shadlen, Newsome & Movshon, 1993) and successfully modelled (Simoncelli & Heeger, 1998), and that it may vary with stimulus size and contrast. The model might, or not, be improved by including such an effect.

Third, it is known that the neuronal responses have contrast and size tuning such that at least one of the drives, excitatory or inhibitory, has intertwined contrast and size response functions. In the current work, assuming a mathematical form for the size tuning drive, it was argued that the spatial spread of the excitatory drive should change with contrast variations. It is possible that instead, with different equations and analyses, one can describe behavioural results by using the contrast tuning components to change with size variations. Such a possibility was used by Tsui and Pack (2011) for modelling their neurophysiological results and one might successfully apply such a model also to the behavioural results, though the current author has not tested it and remains doubtful.

Fourth, if one manipulates neuronal connections with drug intake for example or compares populations with putative differences in connectivities (Schallmo et al., 2018; Perani et al., 2021; Song et al., 2021; Ip & Bridge, 2022) the question naturally arises of how it should be considered in the model. Here, it was argued only through the change of the inhibitory drive model. Since the model approach is based on "static" models, one cannot discard the argument that all the components of the model of "steady state" activity are modified. This is so because in changing connection weights of one type (excitatory or inhibitory) in the system affects the global equilibrium, and this later part is modelled, not the real weights.

Last, because the experimental design analysed here simultaneously measures the stimulus duration with spatial and contrast domain characteristics of motion perception, the model necessarily must incorporate all these factors in its excitatory and inhibitory drives, which naturally leads to high multidimensional parameter space (here 10 parameters) for predicting even such simple data measures. The consequence is that, unless there is prior knowledge, the experimental measures must be carefully designed and carried in order to have data that can correctly constrain all parameters, or at minimum unambiguously give effects associated mainly to inhibition (e.g. very high contrasts) or excitation (e.g. very low contrasts).

While the overall modelling and argumentation was based on the knowledge of motion perception, one should not find it too hard to apply or extend the general part of the model (eqs.6-14 and equ.21) to other basic features, whose neurophysiology was extensively studied during the last 50 years, as for example in V1 for orientation or spatial frequency.

This work showed that we are very successful in providing simple and elegant modelling of behavioural results, results that demonstrate interesting perceptual phenomena that we associate to neurophysiological counterparts. Our modelling must be carefully considered and weighted with respect to prior knowledge. When neurophysiological, computational and behavioural results are combined in appropriate conditions for comparisons, we do gain important insights and knowledge about a given neurophysiological system and how it affects perception. In the case of suppression and facilitation of motion perception in humans, our understanding of how the percept is created and modulated has strong grounds from which we can make interesting inferences and, importantly, allow us to go toward more application oriented studies.

References

Albrecht, D. G. and Hamilton, D. B. (1982). Striate cortex of monkey and cat: contrast response function, Journal of Neurophysiology 48:217-236.

Albright, T. D. and Stoner, G. R. (2002). Contextual influences on visual processing, Annual Review of Neuroscience 25:339-379.

Betts, L. R., Sekuler, A. B. and Bennett, P. J. (2009). Spatial characteristics of center-surround antagonism in younger and older adults., J Vis 9:1-15.

Betts, L. R., Sekuler, A. B. and Bennett, P. J. (2012). Spatial characteristics of motion-sensitive mechanisms change with age and stimulus spatial frequency, Vision Research 53:1-14.

Betts, L. R., Taylor, C. P., Sekuler, A. B. and Bennett, P. J. (2005). Aging reduces center-surround antagonism in visual motion processing., Neuron 45:361-366.

Born, R. and Bradley, D. (2005). Structure and function of visual area MT, Annual Review of Neuroscience 28:157-189.

Britten, K. H., Shadlen, M. N., Newsome, W. T. and Movshon, J. A. (1992). The analysis of visual motion: a comparison of neuronal and psychophysical performance, The Journal of Neuroscience 12:4745-4765.

Britten, K. H., Shadlen, M. N., Newsome, W. T. and Movshon, J. A. (1993). Responses of neurons in macaque MT to stochastic motion signals, Visual Neuroscience 10:1157-1169.

Carandini, M. and Heeger, D. J. (2012). Normalization as a canonical neural computation, Nature Reviews Neuroscience 13:51-62.

Chang, J. J. and Julesz, B. (1984). Cooperative phenomena in apparent movement perception of random-dot cinematograms, Vision Research 24:1781-1788.

Churan, J., Khawaja, F. A., Tsui, J. M. and Pack, C. (2008). Brief motion stimuli preferentially activate surround-suppressed neurons in macaque visual area MT, Current Biology 18:R1051-R1052.

Eagleman, D. M. (2001). Visual illusions and neurobiology, Nature Reviews Neuroscience 2:920-926.

Ermentrout, B. (1998). Neural networks as spatio-temporal pattern-forming systems, Reports on Progress in Physics 61:353-430.

Forstmann, B., Ratcliff, R. and Wagenmakers, E.-J. (2016). Sequential Sampling Models in Cognitive Neuroscience: Advantages, Applications, and Extensions, Annual Review of Psychology 67:641-666.

Graham, N. V. (2011). Beyond multiple pattern analyzers modeled as linear filters (as classical V1 simple cells): Useful additions of the last 25 years, Vision Research 51:1397-1430.

Green, D. M. and Swets, J. A., 1966. Signal detection theory and psychophysics. John Wiley & Sons, Inc..

Grossberg, S. and Raizada, R. (2000). Contrast-sensitive perceptual grouping and object-based attention in the laminar circuits of the primary visual cortex, Vision Research 40:1413-1432.

Huk, A. C. and Shadlen, M. N. (2005). Neural activity in macaque parietal cortex reflects temporal integration of visual motion signals during perceptual decision making., Journal of Neuroscience the Official Journal of the Society for Neuroscience 25:10420.

Ip, I. B. and Bridge, H. (2022). Investigating the neurochemistry of the human visual system using magnetic resonance spectroscopy, Brain Structure and Function 227:1491-1505.

Kapadia, M. K., Westheimer, G. and Gilbert, C. D. (2000). Spatial distribution of contextual interactions in primary visual cortex and in visual perception, Journal of Neurophysiology 84:2048-2062.

Kim, J. and Wilson, H. (1997). Motion integration over space: interaction of the center and surround motion, Vision Research 37:991-1005.

Kingdom, F. A. A. and Prins, N., 2009. Psychophysics: a practical introduction. Academic Press.

Kling, A., Field, G., Brainard, D. and Chichilnisky, E. (2019). Probing Computation in the Primate Visual System atSingle-Cone Resolution, Annual Review of Neuroscience 42:169-186.

Levinson, E. and Sekuler, R. (1976). Adaptation alters perceived direction of motion, Vision Research 16:779-781.

Link, S. W. (1975). The relative judgment theory of two choice response time, Journal of Mathematical Psychology 12:114-135.

Link, S. W., 1992. The wave theory of difference and similarity. Lawrence Erlbaum Associates, Hillsdale, New Jersey.

Liu, D.-Y., Ju, X., Gao, Y., Han, J.-F., Li, Z., Hu, X.-W., Tan, Z.-L., Northoff, G. and Song, X. M. (2022). From Molecular to Behavior: Higher Order Occipital Cortex in Major Depressive Disorder, Cerebral Cortex 32:2129-2139.

Luce, R. D., 1986. Response Times. Oxford University Press.

Macmillan, N. A. and Creelman, C. D., 2005. Detection theory: a user's guide. Lawrence Erlbaum Associates.

Marshak, W. and Sekuler, R. (1979). Mutual repulsion between moving visual targets, Science 205:1399-1401.

Morgan, M., Dillenburger, B., Raphael, S. and Solomon, J. A. (2012). Observers can voluntarily shift their psychometric functions without losing sensitivity, Attention, Perception and Psychophysics 74:185-193.

Morgan, M., Watamaniuk, S. N. and McKee, S. P. (2000). The use of an implicit standard for measuring discrimination thresholds, Vision Research 40:2341-2349.

Murakami, I. and Shimojo, S. (1993). Motion capture changes to induced motion at higher luminance contrasts, smaller eccentricities, and larger inducer sizes, Vision Research 33:2091-2107.

Murakami, I. and Shimojo, S. (1996). Assimilation-type and contrast-type bias of motion induced by the surround in a random-dot display: evidence for center-surround antagonism, Vision Research 36:3629-3639.

Murray, G. E. and Norton, D. J. (2023). Reduced visual context effects in global motion processing in depression., PloS one 18:e0291513.

Nawrot, M. and Sekuler, R. (1990). Assimilation and contrast in motion perception: explorations in cooperativity, Vision Research 30:1439-1451.

Palmer, J., Huk, A. C. and Shadlen, M. N. (2005). The effect of stimulus strength on the speed and accuracy of a perceptual decision, Journal of Vision 5:376-404.

Pasternak, T. and Tadin, D. (2020). Linking Neuronal Direction Selectivity to Perceptual Decisions About Visual Motion, Annual Review of Vision Science 6:335-362.

Perani, D., Scifo, P., Cicchini, G. M., Rosa, P. D., Banfi, C., Mascheretti, S., Falini, A., Marino, C. and Morrone, M. C. (2021). White matter deficits correlate with visual motion perception impairments in dyslexic carriers of the DCDC2 genetic risk variant., Experimental brain research 239:2725-2740.

Piëch, V., Li, W., Reeke, G. N. and Gilbert, C. D. (2013). Network model of top-down influences on local gain and contextual interactions in visual cortex, Proceedings of the National Academy of Science USA 110:E4108-E4117.

Ratcliff, R. (1978). A theory of memory retrieval, Psychological review 85:59-108.

Ratcliff, R. and McKoon, G. (2008). The diffusion decision model: theory and data for two-choice decision tasks, Neural Computation 20:873-922.

Reynolds, J. H. and Heeger, D. J. (2009). The Normalization Model of Attention, Neuron 61:168-185.

Schallmo, M.-P., Kale, A. M., Millin, R., Flevaris, A. V., Brkanac, Z., Edden, R. A., Bernier, R. A. and Murray, S. O. (2018). Suppression and facilitation of human neural responses, eLife 7:e30334.

Simoncelli, EP and Heeger, DJ (1998). A model of neuronal responses in visual area MT., Vision Research 38:743-761.

Smith, P. L. (2000). Stochastic Dynamic Models of Response Time and Accuracy: A Foundational Primer, Journal of Mathematical Psychology 44:408.

Smith, P. L. and Ratcliff, R. (2004). Psychology and neurobiology of simple decisions, Trends in Neurosciences 27:161-168.

Snowden, R., Treue, S., Erickson, R. and Andersen, R. (1991). The response of area MT and V1 neurons to transparent motion, The journal of Neuroscience 11:2768-2785.

Song, X. M., Hu, X.-W., Li, Z., Gao, Y., Ju, X., Liu, D.-Y., Wang, Q.-N., Xue, C., Cai, Y.-C., Bai, R., Tan, Z.-L. and Northoff, G. (2021). Reduction of higher-order occipital GABA and impaired visual perception in acute major depressive disorder., Molecular psychiatry 26:6747-6755.

Spillmann, L. and Dresp, B. (1995). Phenomena of illusory form: can we bridge the gap between levels of explanation?, Perception 24:1333-1364.

Spillmann, L. and Ehrenstein, W. (1996). 43, From neuron to gestalt: Mechanisms of visual perception. In: R. Greger, U. W. (Ed.), Comprehensive Human Physiology, Springer-Verlag Berlin Heidelberg.

Spillmann, L. and Werner, J. (1996). Long-range interactions in visual perception, Trends in Neuroscience 19:428-434.

Tadin, D. and Lappin, J. S. (2005). Optimal size for perceiving motion decreases with contrast, Vision Research 45:2059-2064.

Tadin, D., Lappin, J. S., Gilroy, L. A. and Blake, R. (2003). Perceptual consequences of center-surround antagonism in visual motion processing, Nature 424:312-315.

Tadin, D., Park, W. J., Dieter, K. C., Melnick, M. D., Lappin, J. S. and Blake, R. (2019). Spatial suppression promotes rapid figure-ground segmentation of moving objects, Nature communications 10:2732.

Tsui, J. M. and Pack, C. C. (2011). Contrast sensitivity of MT receptive field centers and surrounds, Journal of Neurophysiology 106:1888.

Tynan, P. and Sekuler, R. (1975). Simultaneous motion contrast: velocity, sensitivity and depth response, Vision Research 15:1231-1238.

Tzvetanov, T. (2012). A single theoretical framework for circular features processing in humans: orientation and direction of motion compared, Frontiers in Computational Neuroscience 6.

Tzvetanov, T. (2018). Suppression and facilitation of motion perception in humans, bioRxiv/465807 DOI: 10.1101/465807.

Tzvetanov, T. and Womelsdorf, T. (2008). Predicting human perceptual decisions by decoding neuronal information profiles, Biological Cybernetics 98:397-411.

Watamaniuk, S., Sekuler, R. and Williams, D. (1989). Direction perception in complex dynamic displays: the integration of direction information, Vision Research 29:47-59.

Westheimer, G. (1990). Simultaneous orientation contrast for lines in the human fovea, Vision Research 30:1913-1921.

Wilson, H. R., 2005. Spikes, decisions, and actions: The dynamical foundations of neuroscience. Oxford University Press.