

A theory of a saliency map in primary visual cortex (V1) tested by psychophysics of color-orientation interference in texture segmentation^{*}

Published in *Visual Cognition*, 2006, 14 (4/5/6/7/8), 911-933.

Li Zhaoping^a and Robert J Snowden^b

^a*Dept. of Psychology, University College London, London WC1E 6BT, UK, email: z.li@ucl.ac.uk*

^b*School of Psychology, Cardiff University, Cardiff, CF10 3YG, UKG, UK*

Abstract

It has been proposed that V1 creates a bottom-up saliency map, where saliency of any location increases with the firing rate of the most active V1 output cell responding to it, regardless the feature selectivity of the cell. Thus, a red vertical bar may have its saliency signalled by a cell tuned to red color, or one tuned to vertical orientation, whichever cell is the most active. This theory predicts interference between color and orientation features in texture segmentation tasks where bottom-up processes are significant. The theory not only explains existing data, but also provides a prediction. A subsequent psychophysical test confirmed the prediction by showing that segmentation of textures of oriented bars became more difficult as the colors of the bars were randomly drawn from more color categories.

Key words:

Saliency map, color-orientation interference, texture segmentation, primary visual cortex

1 Introduction

A saliency map aids the selection of visual inputs for further processing given limited computational resources. To better understand the selection, we sepa-

^{*} Work supported in part by the Gatsby Charitable Foundation

rate bottom-up from top-down mechanisms (Wolfe, Cave, and Franzel, 1989, Cave, 1999), and consider a saliency map of the visual field constructed by bottom-up mechanisms only, such that a location with a higher scalar value in this map is more likely to attract attention and be further processed. It has been proposed (Li 1999a, Li 2002) that the primary visual cortex creates a saliency map from direct visual input defined mainly by contrast, and that the saliency of a visual location increases with the firing rate of the most responsive V1 output cell to that location, regardless of the feature tuning of the cell. The primary visual cortex receives many top-down inputs from higher visual areas. Hence, the proposed bottom-up saliency map in V1 is an idealization when the top-down influences are ineffective, such as very shortly after visual presentation onset and without specific top-down knowledge, or when the animal is under anesthesia. The condition of “shortly after stimulus onset” should not be viewed as a severe restriction, since, computationally, bottom-up selections must be fast, and, should be less necessary after the initial selection or long after the corresponding stimulus onset. Furthermore, the saliency value should be such that it is regardless of the visual features like color and orientation (Treisman and Gelade 1980) such that, e.g., the saliency of a red dot can be compared with that of a vertical moving bar (Nothdurft 2000). This desirable property may have led to a common belief, as implicitly or explicitly expressed in previous works (Treisman and Gelade 1980, Wolfe et al 1989, Koch and Ullman 1985, Itti and Koch 2000) on saliency maps, that saliency must be signalled by cells untuned to features, such as cells in parietal cortex (Gottlieb et al 1998, Itti and Koch 2001) and that the saliency map must be outside V1 whose cells are feature tuned. However, just like the purchasing power of an UK sterling is regardless of the holder’s nationality or gender, the firing rate of V1 cells could be an universal currency for saliency with or without simultaneously decoding the input features from them. Thus, in principle, the read out from the bottom-up saliency signal could be feature-blind.

Indeed, V1 sends outputs directly to superior colliculus which is involved in generating saccades (Shipp 2004) and could be viewed as reading out the V1 saliency map. It is recently shown that electrical micro-stimulation of neurons in V1, with currents as low as $2 \mu A$, can evoke saccades by monkeys towards the receptive field locations of the stimulated cells (Tehovnik, Slocum and Carvey 2003). Meanwhile, V1.s role for bottom up saliency does not preclude it from serving other roles such as contributing to object recognition.

Physiologically, the response of a V1 cell to inputs within its classical receptive field (CRF) can be influenced by contextual inputs near but outside the CRF, due to the long but finite range intra-cortical interactions linking nearby cells (Knierim and van Essen 1992, Sillito et al 1995, Kapadia et al 1995). Hence, by our proposal that the V1 response dictates saliency, the saliency of a location is determined by *both* the input strength (or contrast) at that location *and* by its context, as expected (Nothdurft 2000). For instance, a vertical bar will pop out of a background of horizontal bars, but not from a background of vertical bars. This is because the evoked response to the vertical bar (in a cell tuned to vertical orientation) would be much higher than responses to the background horizontal bars (in cells tuned to horizontal orientation), making the vertical bar most salient in the input image. The contextual influence processes most responsible in this case is iso-orientation suppression (Knierim and van Essen 1992), the observation that a cell's response to the optimally oriented bar within its CRF is suppressed by up to 80% when the bar is surrounded by bars of the same orientation outside the CRF. Cells tuned to the horizontal orientation responding to the horizontal bars in the background experience iso-orientation suppression, whereas the cell tuned to vertical bar responding to the single vertical bar does not (Li 1999a,b, Sillito et al 1995). Hence, iso-feature suppression is the neural basis for pop out. Long range connections indeed tend to link cells tuned to similar features (Li and Li 1994). In particular, iso-color suppression has also been physiologically observed (Wachtler, Sejnowski, and Albright 2003), and, from the analysis above, should be responsible for color pop out, say a red among greens.

Physiologically, a visual location containing one or more (overlapping) visual items can evoke responses from many V1 cells whose CRFs overlap. For instance, a small vertical red bar may excite cells tuned to vertical orientation, or cells tuned to red but untuned to orientation, or cells whose optimal orientation is 5 degrees from vertical and whose tuning width is 15° , etc. According to our theory (Li 2002), the saliency of a location is determined by the firing rate of the most responsive V1 output cell to it, regardless of the cell's optimal feature value. Hence, the saliency of the red vertical bar are likely signalled by a cell tuned to vertical, or a cell tuned to red, but less likely by a cell tuned to 10 degrees from vertical. (For simplicity of our argument without loss of generality, we ignore cells simultaneously or conjunctively tuned to both color and orientation in most of this paper. We will show later that including the conjunction cells does not change our main conclusions qualitatively.) Which

cell will be the most active to signal the saliency of this red vertical bar will depend on the contextual stimuli of this bar. For instance, if this red vertical bar is surrounded by a background of green vertical bars, the bar will pop out psychophysically due to its unique color. Physiologically, this bar would excite a cell tuned to red color much more than it does to a cell tuned to vertical orientation at the same CRF location, assuming that the activity levels are comparable in color tuned and orientation tuned cells responding to the background bars (otherwise, when the color signals are too weak, color pop-out ceases anyway). This is because the responding cell tuned to vertical orientation would experience iso-orientation suppression (due to a background of vertical bars) while the cell tuned to red color would not experience iso-color suppression (the background has only green but not red bars). Analogously, a red vertical bar in a background of red horizontal bars will pop out by its unique orientation, and would have its saliency signalled by a cell tuned to vertical orientation.

In this paper, we will apply our theory of the V1 saliency map to understand and predict the phenomena of color-orientation interference in texture segmentation. It is common knowledge that visual search for a target bar of unique orientation, i.e., orientation singleton, oriented at 45° among 135° distractor bars is easy, so is the segmentation between two textures of uniformly oriented bars at 45° and 135° respectively. Snowden (1998) observed that the texture segmentation became difficult when each stimulus bar was randomly assigned a color from two choices (say red and green), whereas the orientation singleton search remained easy under the same color randomization (see Fig. (2)A-D). Note that in both tasks only orientation feature should matter for the task decision. Nevertheless, color feature seems to interfere in the performance. (Nothdurft (1997) also observed interference of luminance variations on orientation based texture segmentation.) Although the tasks, often performed under brief visual presentation and/or under time pressure, require final decisions based on the orientation feature, they also require an initial and presumably dominant task component, namely selection (of the orientation singleton or border) by bottom-up saliency, which is feature blind. The observed color interference is comprehensible in our framework by noting the changes in the saliencies of the orientation singleton or the texture border under the color randomization. It essentially arises because (1) saliency of a colored bar may be signalled by a color tuned cell or an orientation tuned cell depending on the stimulus context, and (2) the saliency of one colored bar

signalled by a color tuned cell can be compared with the saliency of another signalled by an orientation tuned cell to judge which bar is more salient to compete for visual selection. In this paper, a V1 model is used to implement the saliency map theory to account for such color interference. Furthermore, we show a theoretical prediction that color interference in orientation texture segmentation should increase with more color categories in the color randomization (Fig. (2)B,D, F). A psychophysical experiment is then carried out to test and confirm the prediction.

There have been previous models of saliency maps, in particular, those by Koch and Ullman (1985) and by Itti and Koch (2000). These works differ from our theory (and model) in the following ways. First, previous works either do not specify the cortical area for the saliency map, or presume explicitly or implicitly that the saliency map is outside V1. Whereas our theory explicitly states V1 as the locus of the map. Second, the previous works require separate feature maps which extract local spatial discontinuities in the features such as color and orientation. So for instance, a color feature map can help to extract the color pop out features. These separate feature maps are then combined by summation or weighted summation of their outputs to a master saliency map so that the saliency of a visual location is encoded irrespective of the particular feature which is responsible for this location to be salient. In contrast, our theory (to signal saliency by the activity of the most responsive cell to a visual location regardless of its feature tunings) does not require separate feature maps, nor any subsequent combinations of them. Indeed many V1 cells are known to be tuned simultaneously to different features such as orientation and motion direction, they could also signal saliency and it would be impossible to have separate feature maps unless cells tuned to the conjunctive feature form yet another, e.g., orientation-motion conjunction feature map — in which case the arguments for the single feature maps become weaker. Furthermore, not requiring combinations of any maps gets rid of a computational complexity. After all, physiologically, numerous V1 cells respond to a single visual item; many overlapping but non-identical and unequal-sized receptive fields cover a single visual location. It is computationally complex and expensive, perhaps too complex and expensive for the initial bottom up selection purpose before object recognition, to decide which cells should contribute in which way to the saliency of a visual location. To select a location to attend to, it is computationally cheap (perhaps cheapest) to simply attend to the receptive field location of the most responsive output cell in V1, no matter which features the

cell is tuned to and which other cells respond (though less vigorously) to the same visual location. In this sense, saliency signals in V1 are represented by a field of neural activity levels from all V1 output cells, from which the location of the highest single cell activity is selected for further visual processing.

If each V1 cell is tuned to no more than one feature dimension, and one cell's response to one feature value does not affect another's response to another feature, then V1 may be viewed as composed of separate cell groups, each is defined by their common preferred feature value and forms a single feature map. In this case, our proposed selection process may be viewed as a maximum operation over *both* feature maps *and* spatial locations. This is in contrast to the (weighted) *summation* operation over feature maps and maximum operation over spatial locations in traditional saliency models (Treisman and Gelade 1980, Koch and Ullman 1985, Itti and Koch 2000, Wolfe et al 1989, Muller et al 1995, Krummenacher et al 2001). Note that the selection process of maximum operation over both feature maps and spatial locations means a maximum operation over all cells, thereby rendering the feature maps irrelevant for bottom-up saliency computation even if they exist. This also means that, it is not necessary to separately represent or replicate (in V1 or elsewhere) the maximum response over feature maps or cell groups for a given visual location. (Within V1, it is doubtful that separate feature maps exist since there is interaction between cells responding to different features via mechanisms of feature unspecific surround suppression between nearby neurons regardless of their feature preferences (Heeger 1992) and lateral connections linking cells tuned to non-identical feature values (Knierim and van Essen 1992, Kapadia et al 1995). However, for simplicity of presentation (as used sometimes in this paper), non-interaction between V1 responses to different feature values or feature dimensions can be viewed as a crude approximation.) We will discuss in the last section how our experimental findings relate to different models of saliency maps. In particular, we will show that, if the previous models of saliency maps should be adopted, then our psychophysical observations will constrain the combination rules of the separate feature maps into a master saliency map.

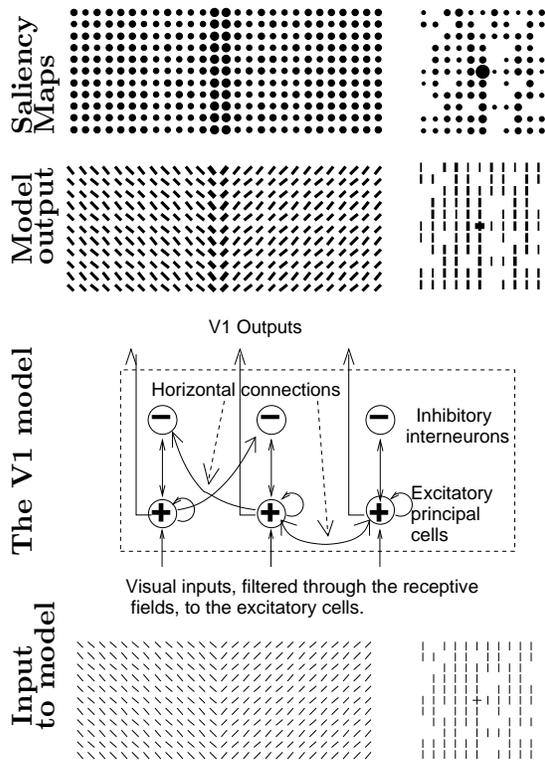
In the next section, we will simulate the saliency map using a physiologically and anatomically based, and computationally designed, V1 model. This simulation is a necessary substitution of physiological observations since current experimental techniques can not yet provide sufficient information on a real V1 saliency map which requires sufficiently dense or high spatial resolution

recordings of V1 responses covering whole scenes of visual search or segmentation stimulus patterns. In section 3, we apply the model saliency map to the stimulus with both color and orientation features, accounting for existing data on color-orientation interference and providing testable predictions. In section 4, we describe the psychophysical experiment testing the prediction and the confirming results. We end with summary and discussions in section 5.

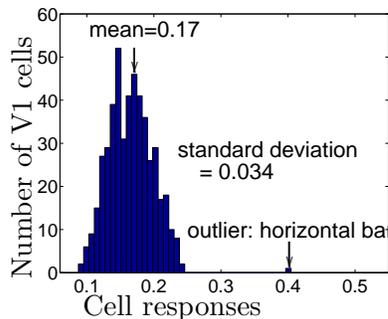
2 Implementing the saliency map by a V1 model

A V1 model (Li 1999b, Li 2000), based on physiological and anatomical data (Rockland and Lund 1983, Gilbert and Wiesel 1983, Hirsch and Gilbert 1992, Fitzpatrick 1996, Knierim and van Essen 1992, Kapadia et al 1995, Sillito et al 1995, Nothdurft, Gallant and van Essen 2000), has been used to validate the saliency map theory. For clarity, we start by describing a model (Li 1999b, Li 2000) which includes only cells tuned to orientation. The model (Fig. 1A) focuses on layer 2-3 of V1 where there is prevalent intra-cortical connections that are responsible for the contextual influences to determine the saliency map. Each model pyramidal cell receives direct visual inputs within its classical receptive field (CRF). CRFs are distributed in space. Each spatial location has twelve model pyramidal cells with overlapping CRFs, roughly modeling a hypercolumn. Each pyramidal cell is tuned to one of the twelve different orientations spanning 180° , with a half tuning width of roughly 15° . Local inhibitory interneurons form reciprocal interactions with the local excitatory pyramidal cells. Longer range interactions between pyramidal cells are mediated by the horizontal collaterals from the pyramidal cells (Rockland and Lund 1983, Gilbert and Wiesel 1983), reaching a distance of about a few CRF sizes in the model (or a few millimeters in the real cortex). These connections enable a pyramidal model cell to excite another mono-synaptically, or inhibit another di-synaptically via inhibitory interneurons. They link cells tuned to similar (not necessarily identical, see Fig. (1C)) orientations (Fitzpatrick 1996), and tend to be mono-synaptic excitatory when the two CRFs are colinear or align to potentially form parts of a smooth curve, and di-synaptic inhibitory otherwise (Fig. (1C)). Orientation unspecific local surround suppression is also implemented phenomenologically as local activity normalization (Nothdurft, Gallant, and van Essen 1999, Heeger 1992). The response to each bar depends on both the input contrast and the contextual stimuli via the lateral connections. The model produces the usual contextual influences observed physio-

A: V1 model, its input, and outputs



B: The histogram of V1 responses to the image of cross in bars in A



C: Lateral connection schematics. Thick solid bar: (oriented) CRF of the pre-synaptic pyramidal. Thin solid (dashed) bars: post-synaptic pyramidal CRFs via mono-synaptic excitation (di-synaptic inhibition)

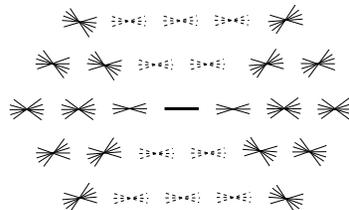


Fig. 1. The V1 model. **A:** A principal pyramidal cell receives direct visual contrast input from a bar within its CRF, comparable in size with the bars in the two input images. Pyramidal cells interact reciprocally with the local interneurons, and interact mono-synaptic-excitatorily and di-synaptic-inhibitorily with other pyramidal cells tuned to similar orientations within a distance of a few CRFs (see C). Each input/output/saliency map image plotted is only a small part of a large extended input/output/saliency map image. The thickness of the bars are plotted as proportional to the input/output strengths of bars for visualization (as in other figures). For example, all input bars in the two input images have the same input contrast, but evoke different V1 response levels (from the pyramidal cells) due to input context via intracortical interactions in the model. In the saliency maps, the radius of a disk is plotted proportionally to the firing rate of the most active cell responding to that visual location. **B:** The histogram of all non-zero model responses (regardless of the preferred orientation of the cells) to the right input image (cross in bars) in A. The model responses are always within range $[0,1]$. **C:** The schematics of the lateral connection pattern between pyramidal cells.

logically. In particular, a cell's response to an optimally oriented bar within its CRF is suppressed if the CRF is surrounded by contextual bars, with the strongest suppression from surrounding bars oriented parallelly to the central bar within the cell's CRF (termed iso-orientation suppression) and weakest suppression from surrounding bars oriented orthogonally to the central bar (Knierim and vanEssen 1992, Sillito et al 1995, Nothdurft et al 1999). The

cell’s response can be enhanced when contextual bars align with the central bar to form a smooth contour — colinear facilitation (Kapadia et al 1995). The model parameters are designed (Li 2001) to be consistent with the known physiology and anatomical data, and to avoid generating spontaneous or hallucinating model outputs not generated by inputs. All the model details and parameters (including activity normalization procedures, lateral connection strengths, etc) of this model are available in (Li 1999b) for interested readers to reproduce the results.

When the model is presented with visual stimuli resembling those in visual search and texture segmentation experiments, the strongest responses are located at or near the pop out targets or texture boundaries, as observed physiologically (Knierim and van Essen 1992, Nothdurft et al 1999, 2000). In the example in Fig (1A) on the right, the cross pops out among the bars since its horizontal bar, the only one that does not experience any iso-orientation suppression from other (vertical) bars in the image, evokes the highest response in the image. Hence, iso-feature suppression is the neural basis for the ease of feature search, when the target has a feature (horizontal orientation) not present in the distractors (Treisman and Gelade 1980). Being the maximum response at its location, the response to the horizontal bar signals the saliency of the location of the cross. If we pick the dominant response to each visual location regardless of the cell’s preferred feature, a histogram of the non-zero dominant responses will show the response outlier to the horizontal bar away from the population responses to the background (Fig. (1B)). The degree to which this response as an outlier could be quantitatively measured by a z score (Li 1999a) and phenomenologically linked to the perceptual saliency. Meanwhile, a sufficient orientation contrast at the border between two textures of uniformly oriented bars (Fig. (1A) left) can pop out because a border bar, having half as many iso-oriented contextual neighbors as those of bars away from the border, experiences reduced iso-orientation suppression and evokes relatively higher responses. This V1 model has been applied to many other visual stimulus patterns similar to those used in visual search and segmentation tasks, giving results consistent with human visual behavior (Li 1999a,b, 2000), such as the quantitative dependence of the strength of orientation pop-out on the quantitative orientation contrast in the stimulus. It elucidates how the just-noticeable-difference for orientation pop-out relates to the structure of the lateral connections. Although all kinds of contextual influences, including iso-feature suppression, colinear facilitation, and general feature unspecific

surround suppression, contribute to the final V1 outputs and the saliency map (Li 1999a,b, 2000), iso-feature suppression is usually the dominant one, and is the most responsible and relevant for the stimulus examples in this paper. Hence, for simplicity of presentation in this paper, we sometimes estimate a cell’s response by the number of iso-feature contextual neighbors for the optimal feature value and dimension of the cell. However, no such approximation is used in the actual model simulation or the figures, where a cell’s response is quantitatively influenced by specific spatial configurations and quantitative feature values of stimulus at and around the CRF and by the level of the local neural activities.

3 Extending the model to the color-orientation feature spaces

Fig. (2, see next page) caption: (A-G): Colour interference in orientation feature based visual search or texture segmentation tasks as demonstrated in the stimuli in the left column, and simulated by the V1 model in the other columns. Each stimulus/response pattern is a small portion of a spatially extended pattern. Second column: the model responses from color tuned cells and from orientation tuned cells. The responses from the color tuned cells are plotted as proportional to the sizes of the colored disks of the corresponding colors. The same ratios, bar thickness:response and disk radius:response, are used in all plots in this column, and they are such that the response levels of the orientation tuned and color tuned cells to the background bars are similar in A. In the saliency maps (third column), the color of a dot is the preferred color of the cell signalling the saliency. If it is black, an orientation tuned cell signals the saliency. The radius:response ratio is the same as that used in the second column. Fourth and fifth columns: the model outputs and saliency maps (using the same size:response ratios as in previous columns) when the model includes cells tuned to both color and orientation, i.e., the conjunction cells. The half-length of the minor axis of the colored ellipses is proportional to the responses from the conjunction cells tuned to the corresponding color and orientation, and the ratio (half) minor axis:response is the same as the radius:response for color tuned cells. Note that color disks, which indicate the responses from color tuned cells, are invisible when the superposing ellipses are larger, indicating stronger responses from the conjunction cells. In the saliency maps (fifth column), when a bar’s saliency is signalled by a conjunction cell, the color of the disk is dark-red, dark-blue, or dark-green to correspond to the color tunings of the conjunction cells and distinguish from the saliency disks (lighter colored) signalled by the color tuned cells. Note that color randomization increases mainly the responses in color but not the orientation tuned cells. Note that the target or border saliency is the highest in A, B, C, and E but not in D, F, and G, indicating significant color interference in D, F, and G. Note that this conclusion holds with or without the conjunction cells in the model.

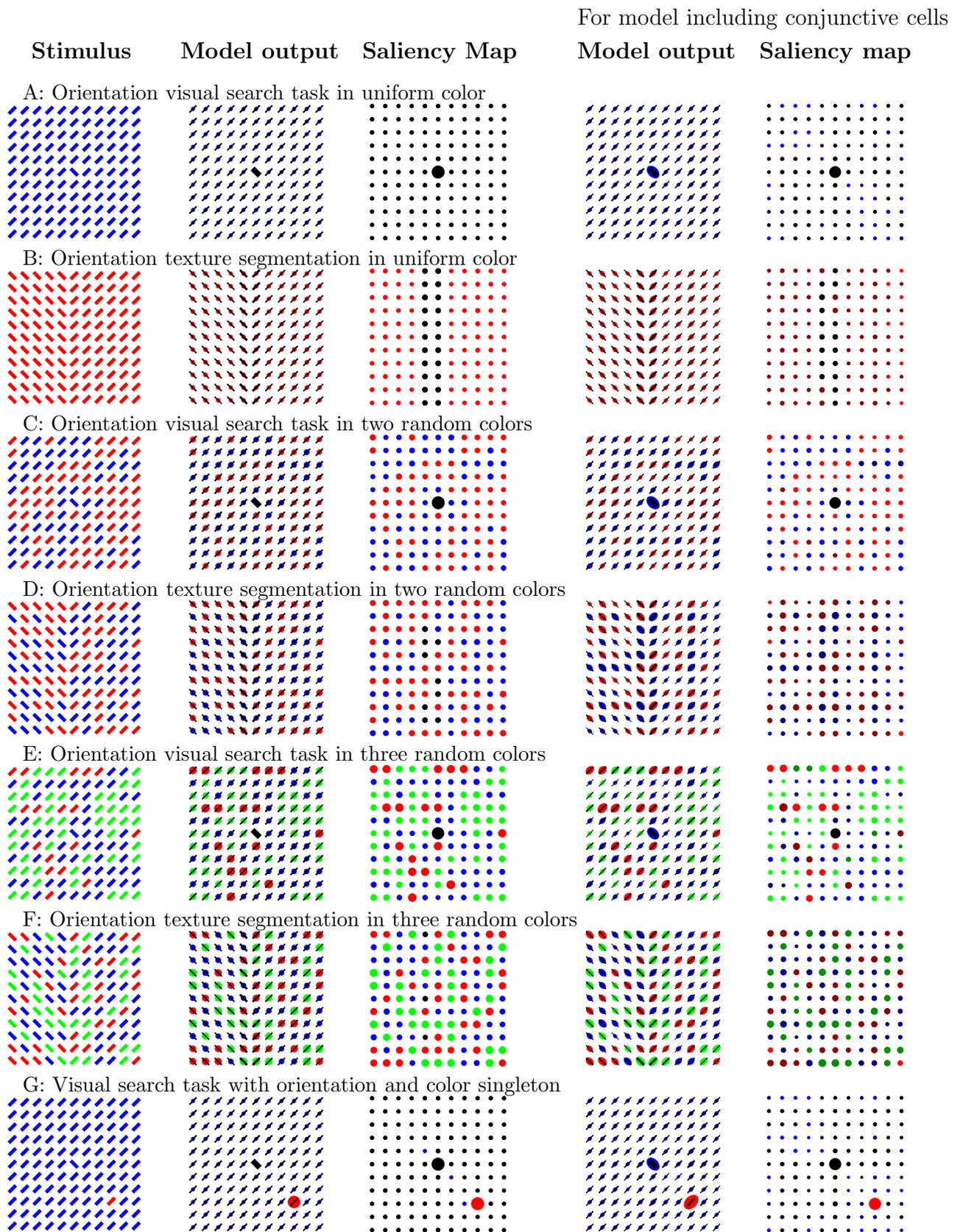


Fig. 2. , see the previous page for caption.

To account for the color-orientation interference, the original V1 model was augmented by adding the color feature dimension without changing any parameters in the original model. At each hypercolumn location, three model cell units, one each tuned to red, green, and blue colors respectively, are added to the set of orientation tuned model units. They have the same pyramidal-interneuron interactions as the orientation tuned units, thus the same input response properties. In this simplified model, a cell is either tuned to color or to orientation but not to both. The mutual suppressive interactions between neighboring units tuned to the same color, i.e., iso-color suppression, as observed physiologically (Wachtler et al 2003), is implemented analogously to the iso-orientation suppression between the orientation tuned cells, except that, since color tuned cells are untuned to orientation, iso-color suppression depends only on the magnitude but not the direction of the displacement between the pre- and post-synaptic cells. Like the orientation tuned cells, each color unit is subject to the same local activity normalization (see Li 1999b for details), thus the feature unspecific local suppression from neighboring (color or orientation tuned) cells. Besides the activity normalization which introduces interactions between cells tuned to different colors or different feature dimensions, we make the simplest assumption that no additional interaction exist between cells tuned to different colors or different dimensions. For the ease of analysis and discussion in this paper, without loss of generality, each color unit is modelled as tuned to a primary color feature of either red, or green, or blue, rather than to the opponent color features (such as red-green opponency) or other hues in real V1 (Gegenfurtner 2003). (Broad tuning to the wavelength of light enables the model cells to respond to non-primary color as well.)

A colored bar evokes response O from an orientation tuned cell and response C from a color tuned cell, in the same hypercolumn, giving a saliency signal $\text{Max}(O, C)$, the maximum of the responses. Let a neuron's response to a bar of preferred orientation be O_{single} when the bar is an orientation singleton, and O_{ground} when the bar is one of the elements in an iso-orientation texture. Let C_{single} and C_{ground} be the analogous responses from color tuned cells to a color singleton and a color element in a uniform color texture. Iso-feature suppression means $O_{single} > O_{ground}$ and $C_{single} > C_{ground}$. In typical stimulus situations with colored bars, color singleton and orientation singleton pop out, implying that both O_{single} and C_{single} are significantly larger than both O_{ground} and C_{ground} . For illustration and in our simula-

tions in Figs. (2) and (3), we use examples in which $O_{single} \approx C_{single}$ and $O_{ground} \approx C_{ground}$, meaning that the color and orientation dimensions have comparable input stimulus strength (for all bars in the figures) and comparable intra-cortical interaction strengths. With uniform color stimuli as in Fig. (2)AB, all bars evoke suppressed responses C_{ground} from color tuned cells, while only the background bars evoke suppressed responses O_{ground} from orientation tuned cells. The orientation singleton and texture border bars, with no or fewer iso-orientation neighbors, evoke higher responses O_{single} and O_{border} respectively, with $O_{single} > O_{border} (> O_{ground})$ since the singleton is the only one with no iso-orientation neighbors. They pop out when their saliencies, $\text{Max}(O_{single}, C_{ground})$ and $\text{Max}(O_{border}, C_{ground})$ respectively, are significantly higher than the background saliency $\text{Max}(O_{ground}, C_{ground})$. This is the case when $C_{ground} \sim O_{ground}$ or when C_{ground} is significantly weaker than O_{single} and O_{border} . Note that the singleton is always more salient than the texture border as long as $C_{ground} < O_{single}$. When each bar randomly takes one of the two colors as in Fig. (2)CD, the number of iso-color neighbors of any bar is halved on average. Thus the color tuned cell gives a less suppressed response $C_{random} > C_{ground}$. If $C_{random} > O_{ground}$ (which happens in stimulus of sufficient color strengths and contrasts), the background saliency signal will increase from $\text{Max}(O_{ground}, C_{ground})$ to $\text{Max}(O_{ground}, C_{random}) = C_{random}$. The texture border, less salient than the singleton, is more likely submerged by the background saliency to weaken its pop-out strength, and explaining the observations by Snowden (1998). In particular, $C_{random} \sim O_{border}$ when $O_{single} \approx C_{single}$ and $O_{ground} \approx C_{ground}$, since on average a border bar has as many iso-orientation neighbors as a colored bar has iso-color neighbors. Thus the responses O_{border} to the border from the orientation tuned cells are submerged by the background responses C_{random} from the color tuned cells. Meanwhile, the orientation singleton pop out is not impaired since its evoked response O_{single} remains the most vigorous against the less-suppressed background responses C_{random} , and the excited cell is the only one tuned to either feature dimension to escape iso-feature suppressions. See Fig. (2C). Therefore, assuming that object saliency rather than subject scrutiny (for feature values) plays a dominant role in such tasks (often performed under brief visual presentations and under time pressure), the essential reason for the color interference is the following: saliency is regardless of the feature dimension(s) of cells signalling it — hence the activity of a color tuned cell signalling saliency of one bar is compared with the activity of an orientation tuned cell signaling saliency of another to see which bar is more salient for visual selection. Increased color

responses from the background make the orientation responses to border no longer response outliers (i.e., the z score is lower). In other words: being regardless of features, saliency allows color feature to interfere in an orientation feature based task.

From the analysis above, one can arrive at another observation. Suppose that only one distractor bar in an orientation singleton search stimulus has a different color from other bars (Fig. (2)G), then, this color (distractor) singleton also pops out when its evoked color response C_{single} is significantly higher than the background responses O_{ground} and C_{ground} . When $C_{single} \sim O_{single}$, the orientation singleton target and the color singleton distractor have comparable saliencies $Max(O_{single}, C_{ground}) = O_{single}$ and $Max(O_{ground}, C_{single}) = C_{single}$ for visual selection. This impairs the visual search task for the orientation singleton, as has indeed been observed (Pashler 1988, Theeuwes and Burger 1998).

We can also predict phenomena that have not been previously observed. Suppose that the number of color categories available to randomly assign to each stimulus bar increases from two to three, the texture segmentation task should become even more difficult. This is because each bar now has on average only $1/3$ of its neighbors of the same color, further reducing the iso-color suppression on the color tuned cells. This gives further elevated response C_{random} and background saliency, see Fig. (2)EF. Next section describes a psychophysics experiment to test this prediction.

Our simple assumption so far of no conjunction cells (i.e., cells tuned conjunctively to color and orientation) agrees more with some physiological data (Livingstone and Hubel 1984, Ts'o and Gilbert 1988, Conway 2001) and less with other data (Leventhal et al 1995, Gegenfurtner et al 1996, Johnson et al 2001, Friedman et al 2003, Hegde and Felleman 2003). While there is a tendency for color tuned cells to be less tuned to orientation and vice versa, as seen from data (see Fig. 5 in the review by Gegenfurtner 2003) and argued from efficient coding considerations (Li and Atick 1994), we show here that including the conjunction cells in the model will not change our conclusions qualitatively. In each model hypercolumn, six conjunction units are added, each tuned to one of the six color-orientation conjunctions (red, green, blue) \times ($45^\circ, -45^\circ$) relevant in our stimulus, and each has the same input response properties as that of other model units. The pattern of the intra-cortical connections between conjunction cells tuned to similar conjunction features is modelled as a

weighted average of that between color tuned cells and that between the orientation tuned cells. Such an interaction pattern and strength are consistent with the limited physiological data (Hedge and Felleman 2003) so far about the contextual influences using both feature dimensions. The conjunction cells are also modeled to suppress and be suppressed by the single feature tuned cells preferring similar orientation or color, but with a 90% reduction in suppression strength compared to the iso-feature suppression within a single cell class. This interaction between the single feature tuned cells and conjunctive feature tuned cells is included so that, as in behavioral data (e.g., Treisman and Gelade 1980), the color-orientation conjunction search tasks may not be as easy as the feature search tasks. The conjunction cells also participate in the feature unspecific surround suppression. The response from a conjunction cell to a colored bar can be estimated approximately by the number of the bar’s neighbors of the same conjunction feature, due to the iso-conjunction-feature suppression. Analogous to notations for other cell types, CO_{single} , CO_{ground} , CO_{random} , and CO_{border} , respectively, are the responses from a conjunction cell to a preferred conjunction bar presented as a conjunction singleton, surrounded by iso-(conjunction) feature neighbors, in a color randomized iso-orientation field, and at an orientation texture border, respectively. Then, to an iso-orientation background texture stimulus with or without color randomization, conjunction cells and the color tuned cells respond with comparable strengths $C_{ground} \sim CO_{ground}$ and $C_{random} \sim CO_{random}$, since the orientation feature is uniform. To an orientation singleton, the conjunction cell and the orientation cell respond comparably $O_{single} \sim CO_{single}$ since both cells experience no iso-feature suppression. Hence, introducing the conjunction cells does not qualitatively change the saliency of the orientation singleton, whose pop-out remains insensitive to color-interference. To an orientation texture border, the responses from the conjunction cells should be comparable to or higher than those from the orientation tuned cells, $CO_{border} \gtrsim O_{border}$, when the stimulus have uniform or random colors, since a border bar has the same or fewer iso-conjunction-feature neighbors as iso-orientation neighbors. Hence, conjunction cells make the border saliency higher or roughly unchanged when the stimulus color is or is not randomized. However, these enhanced border responses from the conjunction cells are not sufficient to offset the increased mean and variances of the responses to the background also due to color randomization. This means, orientation texture segmentation should remain susceptible to color interference when the conjunction cells are included. This is confirmed in model simulation shown in the two right columns in Fig. (2). Increasing

color randomization makes the texture border increasingly submerged by responses to the background, while the orientation singleton continues to evoke the highest response against the background.

4 Testing the model prediction of increased color-orientation interference

4.1 Methods

Stimuli: Texture patterns consisted of a grid of 12 by 12 elements that occupied a square of side 6.6 cm (6.6 deg), thus the average separation between elements was 0.60 deg. The actual position of each individual element within this grid was randomly jittered in both the horizontal and vertical axis by up to 0.135 cm (deg) using a flat probability profile. Most of the elements were assigned to be background whilst an area of 2 by 8 elements was designated to be foreground (or target, like that in Fig. (3)). The position of this target area was random save that it was not allowed to intrude into the outermost elements so that it was always surrounded by background elements. The orientation of the target area (vertical or horizontal) was randomly chosen from trial to trial, and this was the judgement that the observer was required to make. To define this target area the orientation of the target elements was 90 deg different to that of the background elements. On each trial the orientation of all the target elements was randomly chosen to be either +45 deg or -45 deg with respect to the vertical, and thus the background elements were either -45 or +45 deg respectively.

On each trial the colour of any element could be red, green or blue simply defined here as the output of each of the three colour guns in isolation. The salience of the elements was adjusted to be equal in pilot experiments that matches the apparent brightness of the elements in three observers. (We do *not* mean to imply by this that the differently colored elements were equiluminant.) The average brightness matches of the observers were then used throughout all the experiments. Each color was approximately 8 cd/m², the background was a dark grey at 0.1 cd/m². For reasons not expounded further here the elements could have a length of either 3 or 6 mm, and a width of 0.6 or 1.2 mm (on any single trial all elements had the same length and width, chosen randomly for each trial). The data from these manipulations are not analysed

further in this paper and data were simply collapsed across these conditions.

In all 5 conditions were run (Red, Green, Blue, two-colours (Red & Green) and three-colours (Red, Green & Blue)). Each single colour condition was run in order to check that there were no overall differences due to the different colors used. Each observer received 100 trials in each of these conditions. In the next condition each element (both target and background) was randomly assigned to be either red or green we term this the two-colour condition. The two-colour condition does not include other color pairs in order to shorten the total duration of an experimental block. In the final condition each element (both target and background) was randomly assigned to be either red, green or blue we term this the three-colour condition. Observers received 300 of each of these last two conditions. Thus in total each observer saw 900 trials which were presented within a single block of trials in a random order.

Procedure: Each subject was given instructions as to the nature of the task and had a small practice session (approx. 30 trials) before any data were gathered. The observers were told to perform as quickly as possible whilst trying not to commit errors. Each trial consisted of a blank interval (375 ± 125 ms), followed by a fixation cross (375 ± 125 ms) placed at the centre of where the stimulus array would be presented (i.e. the centre of the screen). Following the extinguishing of the fixation cross the stimulus array appeared. The observer then made a speeded two-option forced choice (vertical or horizontal) as to the orientation of the target area. Stimuli were presented until the observer responded and then the screen was blanked. Reaction times and correctness of response were recorded by the computer and feedback was given via an auditory signal for incorrect responses. Observers were given breaks (approximately 5-10 minutes) every three hundred trials or that they could stop at any time if they so wished (but none did).

Observers: Nine observers were used (7 female) with an age range of 19-38. Eight of them were naive to the aims of the experiment whilst one was one of the authors (RS). All were screened for colour deficiencies using the Ishihara test.

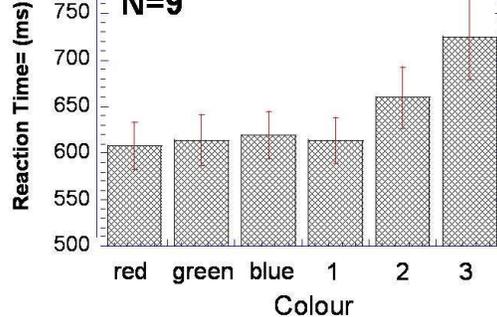
Data analysis: Reaction times were first screened for extreme outliers (> 3000 ms or less than 100 ms) and these were removed ($< 1\%$ for any observer). As RT scores do not form a normal distribution the median RT was then calculated for each observer in each of the 5 colour conditions using only

trials on which the observer was correct. The percentage of trials on which the observer was incorrect was also noted. Differences in RTs and errors were examined via a series of *a priori* planned comparisons using t-tests.

4.2 Results

Figure (3)A displays the means and standard error of measurement for the 5 conditions. Comparison of the data from each of the three single colour conditions revealed that performance was not statistically different for each of the colours (2-way t-tests; RvG $t(8) = -0.08$, ns; RvB $t(8) = -1.22$, ns; GvB $t(8) = -0.59$, ns) and thus data from these conditions were collapsed so as to form a single variable termed 'one colour'. As predicted the two-colour condition gave reaction times that were significantly longer than the one colour condition (one-way t-test: $t(8) = 2.89$, $p = 0.010$), and the three-colour condition was slower than the two-colour condition (one-way t-test: $t(8) = 2.61$, $p = 0.015$). Not surprisingly the three-colour was significantly slower than the one-colour condition (one-way t-test: $t(8) = 2.89$, $p = 0.005$). Error rates were always quite low (0.2 - 8.1%) for all subjects tested. Statistical analysis (ANOVA) did not reveal any significant effects ($p > 0.05$).

Figure (3)B shows the model simulations which agree qualitatively with the experimental outcome. For simplicity, only the simulations in a model without conjunction cells were shown, since those including conjunction cells are qualitatively the same. Note that although the stimuli are mainly of a texture segmentation type, they have an element of orientation search stimuli in it since the foreground region has only 2x8 elements. (When the foreground has only one element the stimulus pattern is the same as that in the singleton search task). This is manifested in the most vigorous responses to the two of the corners of the foreground since they have the fewest iso-orientation neighbors, and, they have a particular spatial configuration of the bars around them to enable favorable contextual influences. Hence, the rate of deterioration of the foreground saliency with increasing color randomization is somewhat between those for the texture segmentation stimulus and the orientation singleton stimulus in Fig. (2). In realistic stimulus, this is common since a texture region has to be spatially bounded with the extreme corners, unlike in the model simulation where the space boundary is wrapped around to simulate an infinitely large spatial texture. Note that although color randomization may



B: Model behavior on the test stimulus patterns

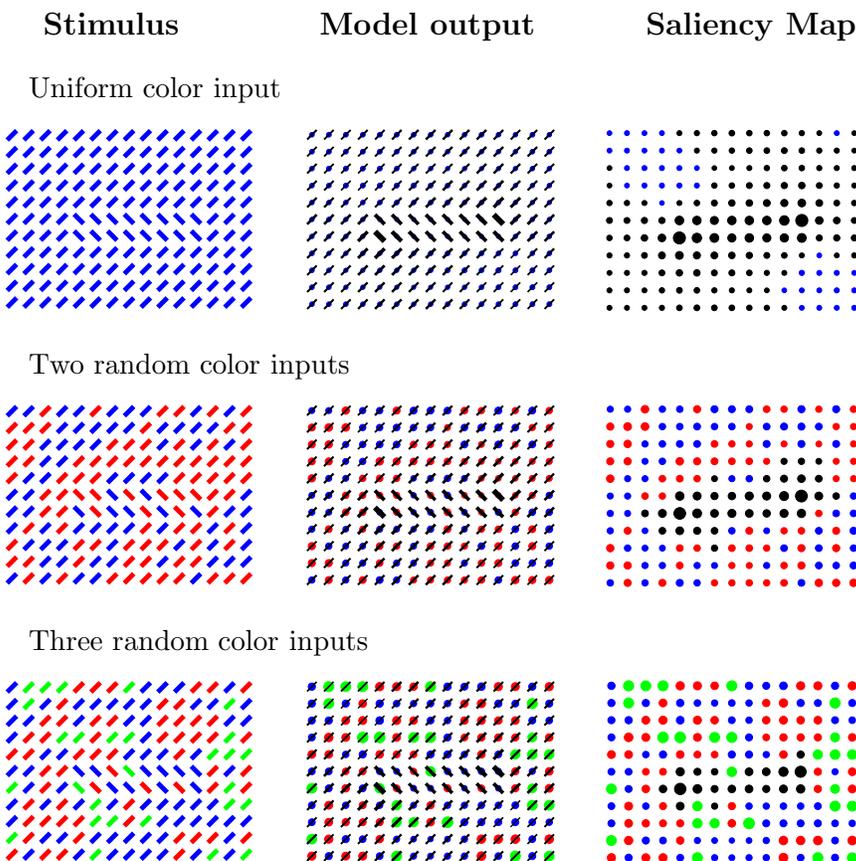


Fig. 3. A: Data from experimental test of the model predictions: more random color features should introduce more color interference on orientation feature based segmentation tasks. Nine subjects performed the tasks of identifying the orientation (vertical or horizontal) of a 2x8 textured foreground area. Shown are averaged reaction times (and their error bars, 900 trials for each subject), as measurements of the difficulties of the tasks. Shortest reactions were for cases when all bars are of uniform color (red, green, or blue, averaged in “1 colour”, the third column from right), longer reactions are for colors of bars randomly drawn from red and green (second column from right), longest reactions were for bars randomly drawn from the three colors (the right most column). B: model simulation of the tasks. Note that the averages and variations of the responses of the color tuned cells to the background increase as more varieties of color features are introduced randomly to the bars, submerging responses to the foreground.

shift the most salient image location from foreground to background, such as in Fig. (2)DF and Figure (3)B, this does not mean that the foreground becomes invisible. Visibility and saliency pop out are two different things. Less salient targets simply require longer reaction times, to allow visual attention to visit more salient locations first before being registered in the awareness (i.e., become visible).

5 Summary and Discussion

We applied the theory of a saliency map in V1 to the color-orientation interference phenomena using a model of V1. This theory is unique in proposing that saliency of a visual location or an object (e.g., a short colored bar) is signalled by the activity level of the most active V1 cell responding to it, and is regardless of the feature tuning properties of the V1 cells signalling it. Hence, the activity of a color tuned cell signalling saliency of a colored bar can be compared with the activity of an orientation tuned cell signalling saliency of another bar to see which bar is more salient. Meanwhile, whether a color tuned or an orientation tuned cell signals the saliency of a colored bar depends on (1) the input strength in color and orientation and (2) the color and orientation contextual input of the bar concerned. This contextual dependence is computed by the intra-cortical interactions in V1 linking neighboring cells, and demonstrated in this paper by a V1 model.

This model explained the existing data (Pashler 1988, Snowden 1998) on the interference of the color features in orientation feature based tasks, in particular, the orientation singleton search and orientation texture segmentation. The ease of these tasks, especially when performed under time pressure, depend largely on the saliencies of visual objects. Specifically, our framework explains why randomizing the colors of the bars impairs performance in some segmentation tasks but not others. In addition, we predict from our model that the color interference should increase with increasing color categories used in the colors of the oriented bars in the stimulus. A psychophysics experiment was carried out to test the prediction, and reaction times in texture segmentation task was shown to increase with increasing color categories, confirming the model prediction.

Nothdurft has observed that non-isoluminant stimuli could contribute to in-

interference in orientation feature based segmentation (Nothdurft 1997). Since our experiments used different luminance values for bars of different colors, it could happen that luminance variations in the background texture contributed to the observed interference. However, we believe that such a luminance factor is minimal or non-significant in our experiment since the reaction times of the orientation segmentation under uniform color are similar (Fig. 3A) for all three colors, implying that the saliencies of differently colored bars are comparable in uniform textures without color randomization. This implies that any changes of saliency values in color randomized stimuli are caused by changes in color feature values rather than the luminance values.

One could ask if the previous models of bottom-up saliency map (Koch and Ullman 1985, Itti and Koch 2000) could also explain our experimental data. These models assume that separate feature maps, such as color feature maps and orientation feature maps, are constructed to obtain the saliency highlights in separate feature dimensions, and these separate feature maps are then subsequently combined into a master, feature unspecific, saliency map. Without diving into details of how separate feature maps are constructed, these separate maps obtain as highlights feature discontinuities such as singletons or borders in each feature dimension (Koch and Ullman 1985, Itti and Koch 2000). These mechanisms correspond to contextual interactions between cells tuned to the same dimension in our V1 model. These models then assume some mechanism that combines the results from different feature maps. The combination rule used has been to sum (or feature weighted sum) the outcomes from the feature maps into a scalar master map (Muller Heller and Ziegler 1995, Itti and Koch 2000). Given that the orientation feature map gives a highlight at the texture border while the color feature map has its high lights spatially distributed in the whole image, the summation of the two maps would produce saliency highlights still along the texture border. This means, the orientation segmentation task should not be sensitive to color interference. Arguably, it is possible that some degree of feature interference could still be predicted by this summation rule, considering that the color feature map gives a very inhomogeneous highlights due to the random assignments of colors to the bars. This creates a noisy background in the master map and this noisy background makes the border highlight not as conspicuous as it is under uniform color (Duncan and Humphrey 1989, Rubenstein and Sagi 1990). Alternatively, if the feature map combination rule into a master map is the maximum rather than summation rule, i.e., the master map takes the maximum among the feature maps at

each spatial location, the outcome would be qualitatively equivalent to our model without the conjunction cells. Hence, the color-orientation interference phenomena argue that if the previous models of the saliency map were to be adopted, the combination rule from separate feature maps to the master map is more likely a maximum or winner-take-all rule, in which case, separation of feature maps is no longer relevant. Meanwhile, it has been observed that double feature search or searching for a dimensionally redundant target, such as searching for a vertical red bar among horizontal green bars, is faster than either of the single feature searches in either dimension alone, e.g., a vertical bar among horizontal bars, or a red among green (Krummenacher, Muller, and Heller 2001, Nothdurft 2000). This observation can not be explained by a maximum rule of the feature maps since it would argue that a double feature search should not be easier than the easiest of the single feature searches. In particular, the maximum rule means that the reaction time to the double feature target should be the shorter one of the two single feature search reaction times. This is called the race model inequality, i.e., taking the reaction time of one of the two single feature dimensions that reaches the task decision first in a race between the two feature dimensions, see Krummenacher et al (2001). However, in our theory, where there is no separate feature maps nor any combinations of them, the faster double feature search can be explained by introducing the conjunction cells (Li 2002). This introduces a third racing element in the reaction time race which is now between three competitors: the color tuned cell, the orientation tuned cell, and the conjunction tuned cell. While each of the former two competitors determine the reaction time for the single feature search, the double feature search can be faster when the conjunction cell's signal becomes the race winner in some of the trials.

We also note that top-down mechanisms (Wolfe et al 1989, Muller et al 1995, Cave 1999), whether or not to construct the saliency map, could not explain the interference phenomena in this paper. This is simply because top-down control would argue instead that, color, being the irrelevant feature for the tasks, should not interfere.

To summarize, we applied the theory and model of the saliency map in V1 to the phenomena of color interference in orientation feature based tasks. Our framework explains the existing data, provided a prediction which is tested and confirmed psychophysically. While our results here favour the V1 saliency map theory over others, they do not conclusively rule out any particular theory. However, our findings and our analysis of the scenarios here can motivate

future investigations to better distinguish the alternative theories. One of these new investigations (Zhaoping and May 2004) has already generated additional findings that strengthen our conclusions here and will be reported in a future paper.

Acknowledgement This work is supported in part (L.Z.) by the Gatsby Charitable Foundation.

References

- [1] Cave K.R. “The FeatureGate model of visual selection” *Psychological Research* 62:182-194 (1999).
- [2] Conway BR (2001). Spatial structure of cone inputs to color cells in alert macaque primary visual cortex (V1), *J. Neurosci.* 21(8):2768-83
- [3] Duncan J. Humphreys G. “Visual search and stimulus similarity”. *Psychological Review* 96: p1-26, (1989).
- [4] Fitzpatrick D. (1996) The functional organization of local circuits in visual cortex: insights from the study of tree shrew striate cortex. *Cerebral Cortex* 6(3):329-41
- [5] Friedman HS, Zhou H, Heydt Rv R. “The coding of uniform color figures in monkey visual cortex” *J. Physiol* 548(Pt 2):593-613. Epub (2003).
- [6] Gegenfurtner KR, Kiper DC, Fenstemaker SB. “Processing of color, form, and motion in macaque area V2” *Vis Neurosci.* 13(1):161-72 (1996)
- [7] Gegenfurtner KR “Cortical mechanisms of colour vision” *Nat. Rev. Neurosci.* 4(7):563-72 (2003)
- [8] Gilbert CD, Wiesel TN. (1983) Clustered intrinsic connections in cat visual cortex. *Journal of Neuroscience* 3(5):1116-33
- [9] Gottlieb, J.P., Kusunoki M. and Goldberg M.E. “The representation of visual salience in monkey parietal cortex ” *Nature* 391(6666), 481-484 (1998)
- [10] D. J. Heeger “Normalization of cell responses in cat striate cortex” *Visual Neurosci.* 9, 181-197 (1992).
- [11] Hegde J. and Felleman DJ “How selective are V1 cells for pop-out stimuli?” *J. Neurosci.* 23(31):9968-80, (2003)
- [12] Hirsch JA, Gilbert CD. (1991) Synaptic physiology of horizontal connections in the cat’s visual cortex. *J. Neurosci.* 11(6):1800-9.
- [13] Itti L. and Koch C. “A saliency-based search mechanism for overt and covert shifts of visual attention” *Vision Res.* 40(10-12):1489-506 (2000)

- [14] Itti L. and Koch C. “Computational modelling of visual attention” *Nature Reviews Neuroscience* 2(3) 194-203, 2001
- [15] Johnson EN, Hawken MJ, Shapley R. “The spatial transformation of color in the primary visual cortex of the macaque monkey”. *Nat Neurosci* 4(4):409-16 (2001).
- [16] Kapadia, M. K., Ito, M. , Gilbert, C. D., and Westheimer G. “Improvement in visual sensitivity by changes in local context: parallel studies in human observers and in V1 of alert monkeys”. *Neuron*. 15(4), 843-56. (1995)
- [17] Knierim J.J. and van Essen D. C. “Neuronal responses to static texture patterns ion area V1 of the alert macaque monkeys.” *J. Neurophysiol.* 67, 961-980 (1992).
- [18] Koch, C. and Ullman, S. “Shifts in selective visual attention: towards the underlying neural circuitry”. *Hum. Neurobiol.* 4, 219-227 (1985).
- [19] Krummenacher J., Muller HJ, and Heller D. “Visual search for dimensionally redundant pop-out target: evidence for parallel-coactive processing of dimensions”. *Percept Psychophys.* 63(5):901-17 (2001).
- [20] Leventhal AG, Thompson KG, Liu D, Zhou Y, Ault SJ. “Concomitant sensitivity to orientation, direction, and color of cells in layers 2, 3, and 4 of monkey striate cortex.” *J. Neurosci.* 15(3 Pt 1):1808-18 (1995).
- [21] Li, Z. “Contextual influences in V1 as a basis for pop out and asymmetry in visual search” *Proc. Natl. Acad. Sci. U.S. A.* 96, 10530-10535 (1999a).
- [22] Li, Z. “Visual segmentation by contextual influences via intracortical interactions in primary visual cortex”, In *Network: computation in neural systems* 10, 187-212. (1999b)
- [23] Li Z. “Pre-attentive segmentation in the primary visual cortex” *Spatial Vision* 13(1): 25-50. (2000)
- [24] Li, Z. “Computational design and nonlinear dynamics of a recurrent network model of the primary visual cortex” *Neural Computation* 13/8, p.1749-1780. (2001).
- [25] Li, Z. “A saliency map in primary visual cortex” *Trends in Cog. Sci.* 6(1)9-16. (2002)
- [26] Li.Z and Atick J. “Towards a theory of striate cortex” *Neural Computation* 6, 127-146 (1994).
- [27] Li, C.Y., and Li, W. “Extensive integration field beyond the classical receptive field of cat’s striate cortical neurons — classification and tuning properties.” *Vision Res.* 34(18), 2337-55 (1994)
- [28] Livingstone M. S. and Hubel, D. H. “Anatomy and physiology of a color system in the primate visual cortex”. *J. Neurosci.* 4(1): 309-356 (1984).

- [29] Muller HJ, Heller D, Ziegler J. "Visual search for singleton feature targets within and across feature dimensions." *Perception and Psychophysics* 57, 1-17 (1995).
- [30] Nothdurft HC, Gallant JL, Van Essen DC. "Response profiles to texture border patterns in area V1" *Vis Neurosci.* 17(3):421-36 (2000).
- [31] Nothdurft HC, Gallant JL, Van Essen DC "Response modulation by texture surround in primate area V1: correlates of 'pop-out' under anesthesia." *Vis Neurosci.* 16(1):15-34 (1999)
- [32] Nothdurft HC "Different approaches to the coding of visual segmentation." In: Harris L. and Jenkins M. *Computational and psychophysical mechanisms of visual coding* p. 20-43. Cambridge University Press, New York (1997).
- [33] Nothdurft H.C. "Saliency from feature contrast: additivity across dimensions." *Vis. Res.* 40, 1183-1202 (2000).
- [34] Pashler H. "Cross-dimensional interaction and texture segregation" *Perception & Psychophysics* 43: 307-318 (1988).
- [35] Rockland KS, Lund JS. (1983). Intrinsic laminar lattice connections in primate visual cortex. *J Comparative Neurology* 216(3):303-18.
- [36] Rubenstein B. and Sagi D. "Spatial variability as a limiting factor in texture discrimination tasks: implications for performance asymmetries" *J. Opt. Soc. Am. A* 9: 1632-1643 (1990).
- [37] Shipp S. "The brain circuitry of attention" *Trends Cogn. Sci.* 8(5):223-230. (2004)
- [38] Sillito, A. M. Grieve, K.L., Jones, H.E. Cudeiro, J. and Davis J. (1995) "Visual cortical mechanisms detecting focal orientation discontinuities" *Nature*, 378(6556), 492-6 (1995)
- [39] Snowden, R.J. "Texture segregation and visual search: a comparison of the effects of random variations along irrelevant dimensions." *Journal of Experimental Psychology: Human Perception and Performance*, 24, 1354-1367 (1998).
- [40] Theeuwes J, Burger R. "Attentional control during visual search: the effect of irrelevant singletons." *J Exp Psychol Hum Percept Perform.* 24(5):1342-53 (1998)
- [41] Tehovnik E.J. Slocum W.M. and Carvey C.E. "Behavioural state affects saccadic eye movements evoked by micro-stimulation of striate cortex." *Eur J. Neurosci.* 18(4): 969-79, (2003).
- [42] Ts'o D. and Gilbert C. "The organization of chromatic and spatial interactions in the primate striate cortex". *J. Neurosci.* 8: 1712-27 (1988).

- [43] Treisman, A. and Gelade, G. “A feature integration theory of attention.” *Cogn. Psychol.* 12, 97-136 (1980).
- [44] Wachtler T, Sejnowski TJ, Albright TD. “Representation of color stimuli in awake macaque primary visual cortex” *Neuron* 37(4):681-91 (2003)
- [45] Wolfe, J. M. Cave K.R. and Franzel S. L. “Guided search: an alternative to the feature integration model for visual search”. *J. Exp. Psychol.* 15, 419-433 (1989).
- [46] Zhaoping L. and May K.A. (2004) Irrelevance of feature maps for bottom up visual saliency in segmentation and search tasks. Program No. 20.1. 2004 Abstract Viewer/Itinerary Planner. Washington, DC: Society for Neuroscience, 2004. Online.