THE ROYAL SOCIETY

Published online 19 November 2002

Pre-attentive segmentation and correspondence in stereo

Li Zhaoping

Department of Psychology, University College London, Gower Street, London WC1E 6BT, UK (z.li@ucl.ac.uk)

Traditional stereo grouping models have focused on the problem of stereo correspondence between monocular inputs. Recent physiological data revealed that the disparity selective V2 cells increase their responses when (random-dot stereograms) stimuli within their receptive fields are at or near the boundary of a depth surface. Such highlights to depth (non-luminance) edges are seemingly not computationally required for the correspondence problem. Computationally, these highlights make the boundaries of a depth surface more salient, serving pre-attentive segmentation (between depth planes) and attracting visual attention. In special cases, they enable the psychophysically observed perceptual pop-out of a target from a background of visually identical distractors at a different depth. To achieve the highlights, mutual inhibition between disparity selective cells that are tuned to the same or similar depths is required. However, such mutual inhibition would impede the computation for the correspondence problem, which requires mutual excitation between the same cells. In this work, I introduce a computational model that, I believe, is the first to address both stereo correspondence and pre-attentive stereo segmentation. The computational mechanisms in the model are based on intracortical interactions in V2. I will demonstrate that the model captures the following physiological and psychophysical phenomena: (i) depth-edge highlighting; (ii) disparity capture; (iii) pop-out; and (iv) transparency.

Keywords: stereo pop-out; stereo correspondence; modelling; contextual influences; V2

1. INTRODUCTION

The random dot stereograms of Julesz (1971) convincingly demonstrate that stereo information alone (without other image cues such as colour, shape and shading) suffices to recover 3D depth from two monocular 2D images. Depth can be recovered by correctly matching the corresponding image features (e.g. dots) in the monocular images. This correspondence problem is ill-posed as there are many possible false matches between individual dots, as illustrated in figure 1 for the wallpaper illusion. These observations led to a substantial research focus on solving the correspondence problem in early stereo vision, as exemplified by the well-known model by Marr & Poggio (1976, 1979) in the 1970s and many other models since then (Prazdny 1985; Pollard *et al.* 1985; Qian & Sejnowski 1989; Nasrabadi *et al.* 1989; Geiger *et al.* 1995; Marshall *et al.* 1996; McLoughlin & Grossberg 1998; Watanabe & Fukushima 1999; Read 2002). These models employ cooperative algorithms (or indirectly, using optimizations or Bayesian approaches) that use contextual information to find *true* binocular matches among all possible matches. Each match is signalled by a unit modelling a disparity selective cell with a finite-sized RF in the cortex such as V1 and V2 (Barlow *et al.* 1967; Hubel & Wiesel 1970; Poggio 1992). The experimentally observed recurrent intracortical connections (Rockland & Lund 1983; Gilbert & Wiesel 1983) could mediate the contextual influences in the cooperative algorithms. The two essential ingredients in the model algorithms are

(i) a smoothness constraint encouraging nearby binocular matches that report the same or similar depths to support each other; and
(ii) a uniqueness constraint enabling true matches to suppress the competing false matches that share the same monocular inputs.

Some of the previous models (Nasrabadi *et al.* 1989; Geiger *et al.* 1995; McLoughlin & Grossberg 1998; Watanabe & Fukushima 1999; Read 2002) also include a third or alternative ingredient that takes into account the relatively fewer unmatched monocular images for the perception of occlusions.

By comparison with the focus on the correspondence problem, the segmentation problem in early stereo vision has been relatively neglected. Psychophysically, a single target can spontaneously pop out from a background of visually identical distractors at a different depth (Nakayama & Silverman 1986). Merely solving the correspondence problem does not explain why a lone target of a unique depth should be more salient than any other identical item in the background depth plane. More generally, a boundary between two depth surfaces is perceived sharply and spontaneously in spite of the random-dot nature of the underlying stereogram (Julesz 1971). The spontaneous nature of these phenomena indicates that such segmentation is pre-attentive, with neural correlates probably situated in the early stages of the visual pathway. Indeed, many disparity selective cells in V2 give higher

One contribution of 22 to a Discussion Meeting Issue 'The essential role of the thalamus in cortical functioning'.

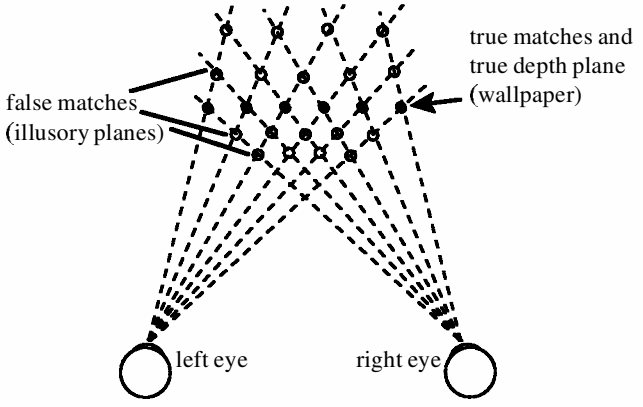
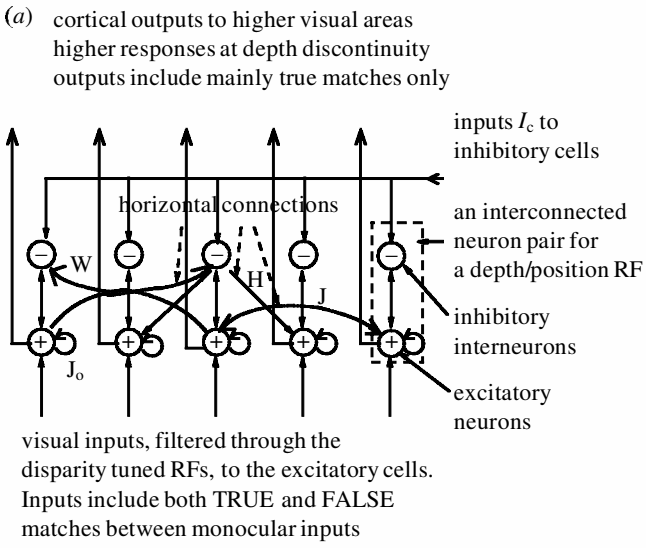


Figure 1. The stereo correspondence problem with the example of the wallpaper effect. The black circles are the visual objects giving rise to the monocular images (of six circles each). The white circles arise from possible false matches between monocular circles. The wallpaper effect is the illusion when the filled circles are identical and equally spaced in an extended depth plane, leading to illusory depth planes nearer or further away.

responses when their RFs are near or at the boundary, rather than the centre, of a random-dot surface of constant depth (Von der Heydt *et al.* 2000). Such higher neural responses to depth discontinuities are not required computationally merely to solve the correspondence problem. However, they certainly help to segment a depth plane from its neighbours. A lone target against a background depth plane is simply a special case of such depth discontinuity.

Computationally, it is not clear whether and how the neural substrates assumed for the correspondence problem are really meant for, or can also be used for, pre-attentive segmentation. Intuitively, there would appear to be a conflict. The smoothness constraint requires that nearby cortical cells tuned to similar depths facilitate each other's activities, as implemented in traditional models (Marr & Poggio 1976; Prazdny 1985; Pollard *et al.* 1985; Qian & Sejnowski 1989; Marshall *et al.* 1996). However, such mutual excitation implies that, for a stimulus composed of a single target (e.g. a dot) from a background of visually identical distractors at a different depth, the lone target dot should evoke a weaker response than any background dot as the cell responding to the lone dot does not receive any mutual excitation affecting cells responding to the background. Such a response pattern would make the pop-out of the unique depth target difficult to achieve. Experimental observations of the depth-edge highlighting (Von der Heydt *et al.* 2000) indicate that nearby cells tuned to similar depths *inhibit* each other, so that cells whose RFs are near the boundary of the depth plane respond more vigorously in comparison to others because they have fewer iso-depth neighbours that actively inhibit them.

In this paper, I introduce a model of V2 that uses the same machinery (based on physiological and anatomical data) to solve the correspondence problem *and* address pre-attentive stereo segmentation, in particular pop-out. It is proposed here that the higher neural responses to depth discontinuities serve pre-attentive segmentation by making depth edges more salient to attract visual attention, and



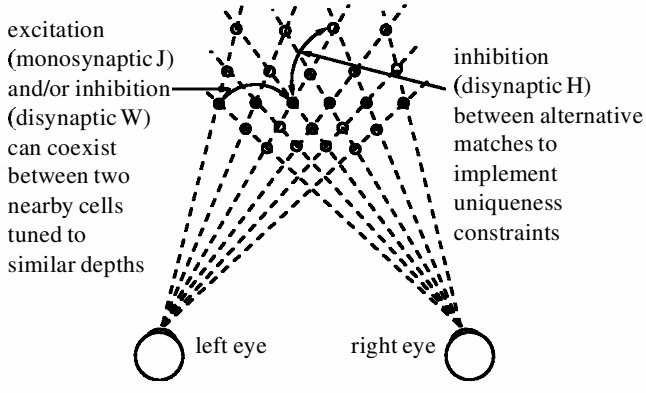


Figure 2. (*a*) The model structure and elements. To avoid clutter, the background inputs I_o to the excitatory cells are omitted in this diagram. W denotes arrows from '+' circles to '−' circles, H denotes arrows from '−' circles to '+' circles and J denotes arrows from '+' circles to other '+' circles. (*b*) The patterns and functions of the intracortical connections, $\mathcal{J}$, W and H, in the model are illustrated using the wallpaper stimuli, each node represents a pyramidal cell responding to a particular (true or false) match from the stimuli. Computationally, the $\mathcal{J}$ connections serve the smoothness constraint, the W connections serve to highlight the depth edges, and the H connections serve the uniqueness constraint.

that they are responsible for the corresponding behavioural pop-out (Li 1999*a*, 2002). The model links physiology with perception and to demonstrate the feasibility of a single neural circuit implementing a cooperative algorithm for an extended computational goal.

2. THE MODEL

(a) *Structure and elements*

The structure and elements of the model are based on biological data (Barlow *et al.* 1967; Hubel & Wiesel 1970; Rockland & Lund 1983, Gilbert & Wiesel 1983; White 1989; Poggio 1992; Von der Heydt *et al.* 2000; Bakin *et al.* 2000) (see figure 2). The model consists of pairs of

principal excitatory (pyramidal) cells and inhibitory interneurons. Each cell is binocular and tuned to a specific depth and a local RF, and models a local cell population in V2. To focus on the stereo domain, we adopt the same simplification as Julesz by omitting other visual features such as orientation, colour and motion of the stimulus and using random-dot stereograms as inputs. The principal cell and the interneuron within a pair are reciprocally connected (White 1989). The principal neurons receive external visual inputs that signal binocular matches at the corresponding locations and depths of the RFs, regardless of whether the matches are true or false. This is in accordance with the data showing that V1 cells, which are disparity selective and provide inputs to V2, respond to both the true and false matches (Cumming & Parker 2000). The responses of the principal cells are the model's outputs. Finite range horizontal connections, represented by $\mathcal{J}$, W and H in figure 2 and the following equations, enable nearby excitatory cells to influence each other, both through monosynaptic excitation and disynaptic inhibition via interneurons, as observed physiologically (Hirsch & Gilbert 1991). The $\mathcal{J}$ and W connections link cells that are tuned to the same or similar depth, as indicated by experimental data (Ts'o & Gilbert 1988). The $\mathcal{J}$ connections mediate monosynaptic excitation between nearby principal cells, whereas the W connections mediate disynaptic inhibition between nearby principal cells via interneurons. Both the monosynaptic excitatory $\mathcal{J}$ connections and the disynaptic inhibitory W connections can exist between a single pair of principal cells tuned to similar depth. We will see later that $\mathcal{J}$ connections enforce the smoothness constraint while W connections enhance depth edges. The pre- and postsynaptic RFs, linked by $\mathcal{J}$ and W, may displace from each other by up to multiple times the RF size. In addition, H connections enable principal cells tuned to different depth values but with the same monocular RF in left or right eyes to inhibit each other via interneurons to implement the uniqueness constraint. The model's response is initialized by the external inputs from V1 (which are sustained after onset) containing both the true and false matches, and is then modified by recurrent interactions. After a transient, the model response is such that mainly the cells corresponding to the true matches give significant responses. In addition, model cells for the depth edges or lone targets give significantly higher responses, even though their external inputs are no stronger than those of the other true matches.

The model's state follows the equations of motion

$$\dot{x}_{id} = -\alpha_x x_{id} - g_y(y_{id}) - \sum_{j,d',d' \neq d} H_{id,jd'} g_y(y_{jd'}) + \mathcal{J}_o g_x(x_{id}) + \sum_{j,d',d' \sim d} \mathcal{J}_{id,jd'} g_x(x_{jd'}) + I_{id} + I_o,$$

$$\dot{y}_{id} = -\alpha_y y_{id} + g_x(x_{id}) + \sum_{j,d',d' \sim d} W_{id,jd'} g_x(x_{jd'}) + I_c.$$

Here, x and y model the states (e.g. membrane potentials) of the principal cells and interneurons, respectively, and $g_x(x)$ and $g_x(y)$ the cells' output activities or firing rates, which are sigmoid-like functions of x and y. The index i,d denotes a cell's RF as at location i in the left monocular image and at depth d. The image location i is distributed in a 2D plane, and, in the examples shown in this paper, the depth takes five values $d \in \{-2, -1, 0, 1, 2\}$. The decay of model states to resting potentials is modelled by the terms $-\alpha_x x_{id}$ and $-\alpha_y y_{id}$. I_o and I_c are background inputs including noise and suppressive normalization of the local neural activities, and I_{id} are the inputs from V1 to x_{id}. The suppressive normalization component $I_{\text{normalization}}(id)$ in, $I_o \equiv I_{\text{normalization}}(id) + \text{constant} + \text{noise}$, to cell id scales with the squared weighted sum of the activities of the neighbouring excitatory cells, i.e. $I_{\text{normalization}}(id) \propto [\Sigma_{jd' \in \text{neighbourhood of } id} f_{jd',id} g_x(x_{jd'})^2]$, where the weights $f_{id',id}$ of the neighbouring activities decays with the distance between the target cell id and its neighbours jd'. Some components and structures of the model are similar to, and others are significantly different from, those of a recent model of intracortical interactions in V1 for pre-attentive segmentation in the texture domain (Li 1999*b*). The quantitative values of the horizontal connections are designed in this model by the constraints of our computational goal as well as the dynamic stability of the model. It is not clear whether the corresponding physiological weights are hardwired or learned. There are, as yet, no known mathematical learning rules for networks of such a dynamic nature and such a computational goal, although there is physiological evidence for experience-driven synaptic plasticities in sensory cortices (Zhang & Poo 2001).

3. MODEL BEHAVIOUR

Figure 3 illustrates the model's behaviour using the example of pop-out. The depth values of the stimuli, or the preferred depths of the cells, are colour coded for visualization purposes. The visual scene (figure 3*a*) contains a random-dot depth plane at depth $d = -2$ and a lone target dot at depth $d = 1$. All stimulus dots have the same input strengths. As the dots are identical, the actual input to the model (figure 3*b*) contains both the true and false matches, distributed in all depth planes $d \in \{-2, -1, 0, 1, 2\}$. Through intracortical interaction, the model evolves from its early states (figure 3*c*) that resemble its inputs, to final states (figure 3*f*) that resemble the perception of the actual scene—the cells that respond are mainly those corresponding to the true matches, and the most responsive cells correspond to the lone pop-out target at $d = 1$ and the dots near the boundary of the depth plane (depth-edge highlighting). In all figures in this paper, the sizes of the dots are plotted to increase with the input or output strengths for visualization purposes.

Closer examination of the evolution of activity in figure 3 reveals that the false matches are removed soon after the onset of the input, whereas the highlighting of the lone target and the depth edges are achieved later, via lateral inhibition W between cells tuned to similar depths. Very soon after the onset of the stimuli at zero time, the model exhibits (figure 3*c*) a strong transient response to all input matches, whether they are true or false. About one membrane time-constant later, mutual excitation between nearby matches of similar depths sustain the responses to true matches in the background depth plane, while inhibition implementing the uniqueness constraint and the general suppression for activity normalization have sup-

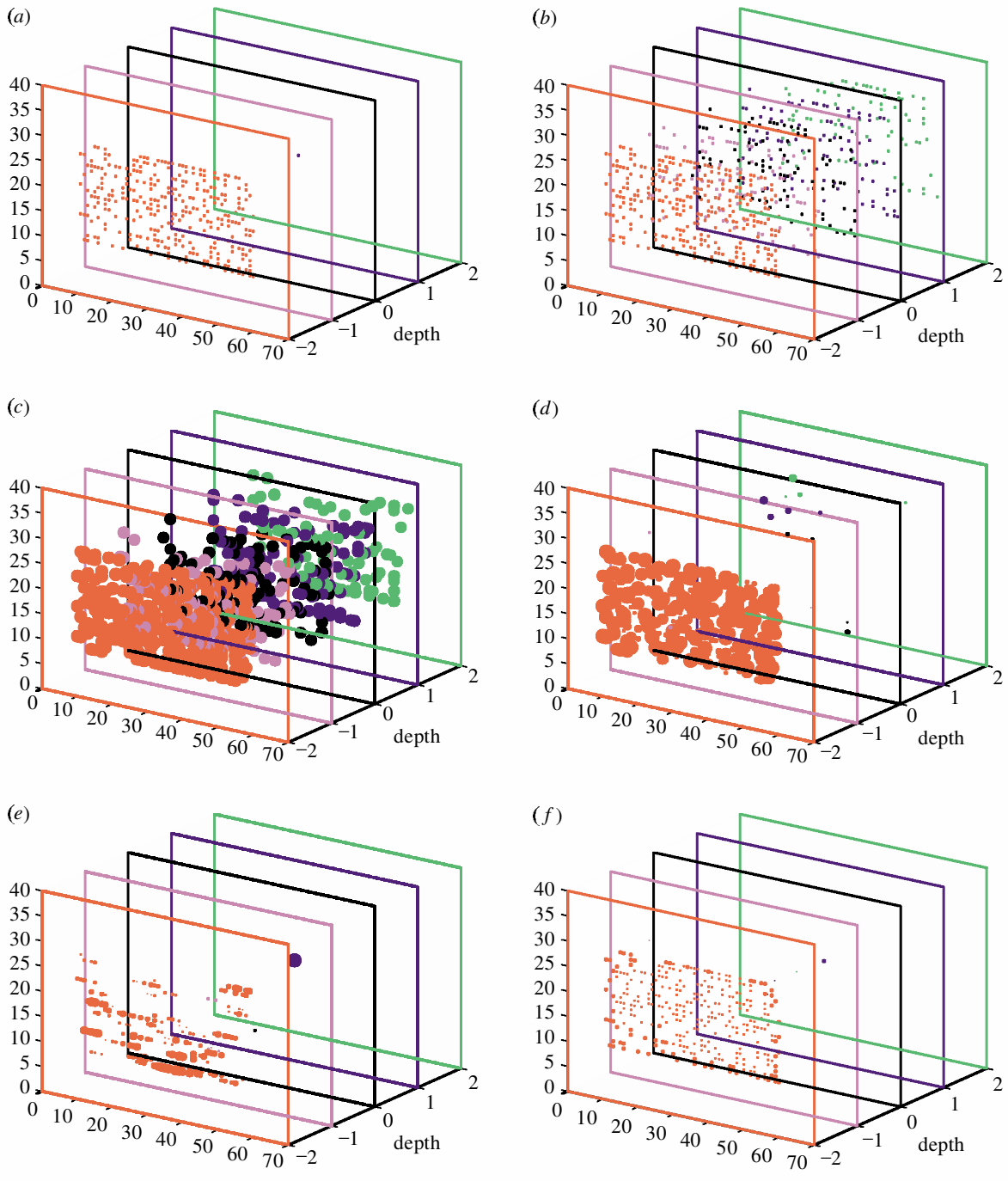


Figure 3. (*a*) Scene, (*b*) model input, (*c*–*e*) model activities at (*c*) $t = 0.6$, (*d*) 1.6, and (*e*) 2.8 membrane time-constant after stimulus onset, and finally, (*f*) the time-averaged model activities after the initial transients. The stimulus onset at $t = 0$. In each plot, the dot coordinates indicate the 3D locations in space, and the sizes of the dots plotted increase monotonously with the output (or input) strengths in a nonlinear fashion for better visualization. The scale of the dot size used in (*f*) is different from that in (*c*)–(*e*) for optimal visualization purposes. The depth values are colour coded for clarity.

pressed the response to almost all false matches as well as the response to the lone target (figure 3*d*). Due to their disynaptic nature, mutual inhibition between cells that are tuned to similar depths become effective only after another membrane time-constant (figure 3*e*), suppressing the response to the background depth plane. This weakens the general suppression on the cell responding to the lone target, and hence its response recovers. Meanwhile, the inhibition H implementing the uniqueness constraint remains effective in suppressing the responses to the false matches. Although the inhibitory H connections for the uniqueness constraint are also disynaptic in nature, they are designed to be much stronger than the disynaptic W connections. Consequently, they act faster and become effective almost immediately after the mutual excitation within the background depth plane takes effect (figure 3*d*). Within a few membrane constants, the network settles into a dynamic cycle in which the network state oscillates between activity patterns resembling figure 3*d*,*e*. This cycle leads to the temporal mean activities shown in

figure 3*f*. In other words, this model can address the stereo correspondence problem and the segmentation problem using the same neural circuit mainly because of the temporal dynamics that are intrinsic in such biological circuits. The monosynaptic $\mathcal{J}$ facilitation between cells tuned to the similar depth acts on a faster time-scale to implement the smoothness needed for stereo matching. The strong activities of the true depth plane help suppress the false matches, as well as the lone true target. This is due to the inhibition that implements the uniqueness constraint and local activity normalization. The disynaptic W inhibition between principal cells at similar depths acts on a slower time-scale (due to the relay via interneurons) to suppress the cells away from the surface boundary and disinhibit the lone target at a different depth. This leads to the saliency highlights for segmentation. The different time-scales of different interactions allow the two computational problems—correspondence and segmentation—to be addressed during two different time-windows within a single dynamic cycle of the neural assembly. Incidentally, these temporal dynamics also help de-synchronize the neural activities that are associated with different depth planes or targets, adding another useful segmentation signature.

The dynamics of the model exhibit similar characteristics, for example, oscillatory activities, synchrony within a depth plane and asynchrony between different depth planes, and highlighting the depth discontinuities, when processing other input stimuli. Figure 4 shows how the model accounts for the physiological observation (Von der Heydt *et al.* 2000) that there are relatively higher responses to the boundaries of depth surfaces. Highlighting the depth discontinuity, or the lone target dot, certainly serves to segment the two depth planes by attracting visual attention to the boundary or the target.

Figure 5 shows how the model accounts for disparity capture, which, like depth-edge highlighting, is another physiologically observed manifestation of contextual influences in V2. The centre of a wallpaper-like grating stimulus (or stimulus of a regular structure such as in figure 5*a*), which is far from the boundary, presents ambiguous depth signals, since many binocular matches are possible (see figures 1 and 5*b*). However, a V2 cell whose RF is exposed to the centre of the grating often responds as if it receives an unambiguous global depth signal from the grating boundary beyond its RF (Bakin *et al.* 2000). The model shows the same behaviour (figure 5*c*).

Figure 6 demonstrates how the model accounts for transparency, i.e. the perceptual segregation of two overlapping depth planes. Again, the responses to the surface boundaries are stronger. In addition, responses to a few false matches did survive; a phenomenon (ghost dots) also observed in human vision for stereo transparency (Weinshall 1989). Note that transparency has been traditionally difficult to model, since it requires the model to accommodate two discrete depth values at a given visual angle. Although the original model by Marr & Poggio (1976) could not account for it because of its particular implementation of the smoothness constraint, more recent stereo models that solve the correspondence problem (Qian & Sejnowski 1989; Marshall *et al.* 1996) do successfully model it.

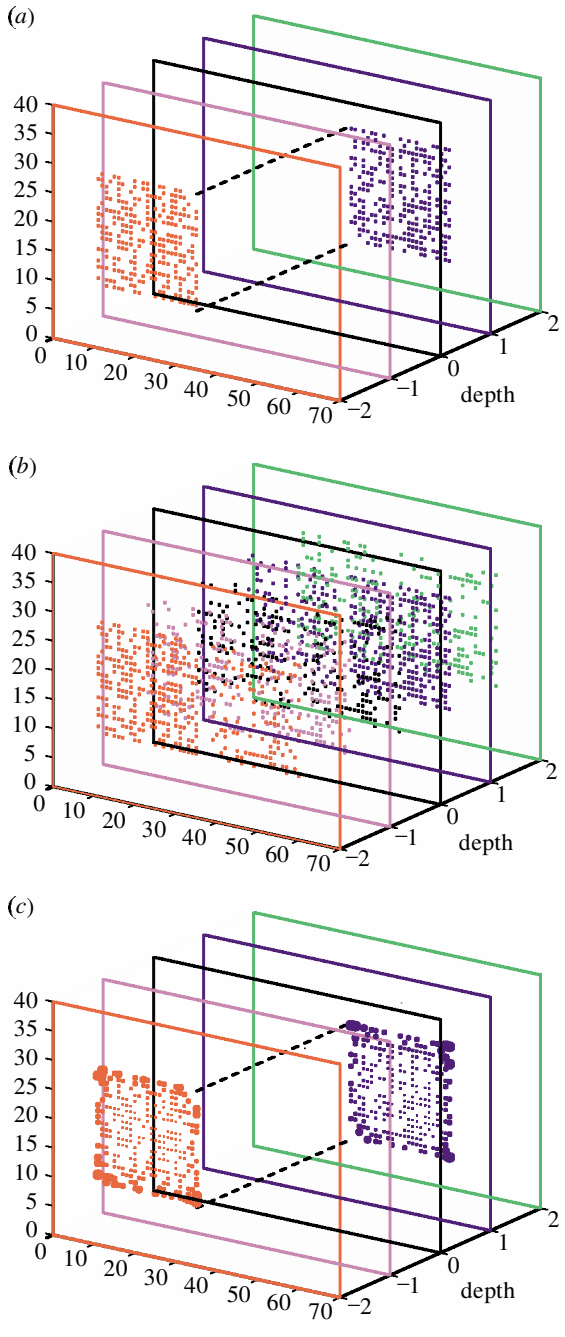


Figure 4. Modelling depth discontinuity in a scene of two depth planes (with depths $d = -2$ and $d = 1$) with a depth edge (or depth step, marked by dashed lines) between them. Note the enhanced response near the depth edge and the boundary of the planes. (*a*) 3D scene, (*b*) actual input to the model and (*c*) time-average model activity.

4. SUMMARY AND DISCUSSION

I propose in this paper that pre-attentive segmentation in the stereo domain is addressed in the brain by awarding relatively stronger neural responses to more salient depth features. Salient features include a target of a unique depth in the scene or a depth discontinuity, and are important as an initial step towards segmentation and for attracting visual attention for further processing. I believe that this model is the first to address computational mechanisms for pre-attentive stereo segmentation in the sense of pop-out or saliency highlights by depth, in addition to

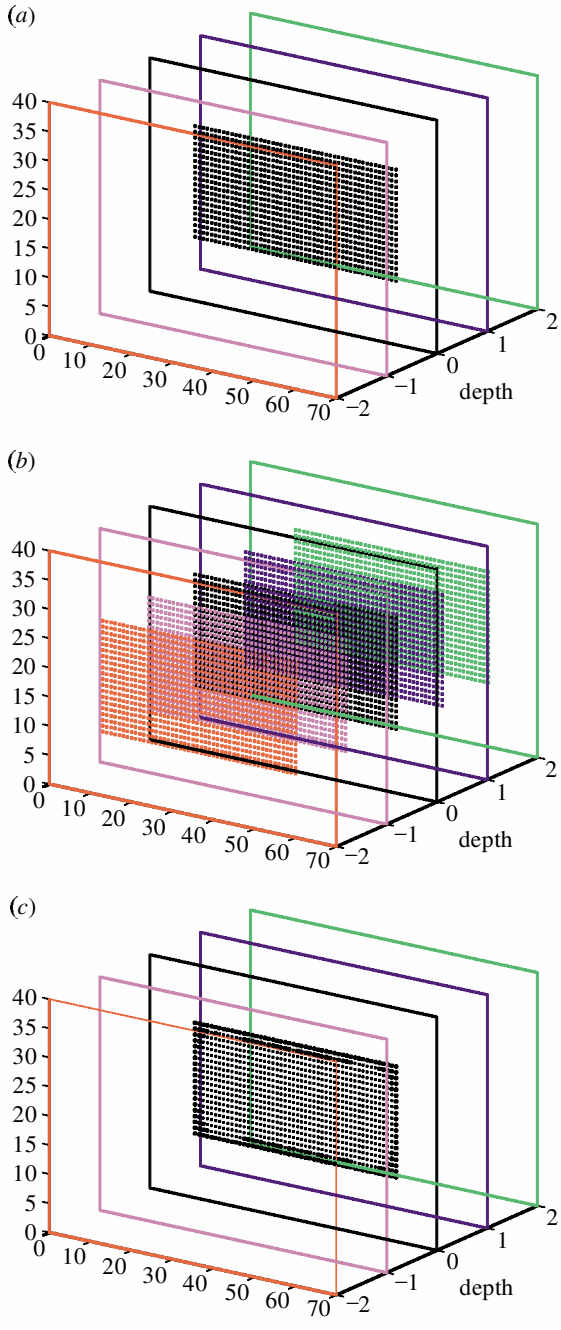


Figure 5. Modelling disparity capture. Note that the actual input to the model contains four illusory depth planes as do occasionally occur in human vision. The outputs are strongest at the boundary of the true depth plane at a depth of $d = 0$. (*a*) 3D scene, (*b*) actual input to the model and (*c*) time-average model activity.

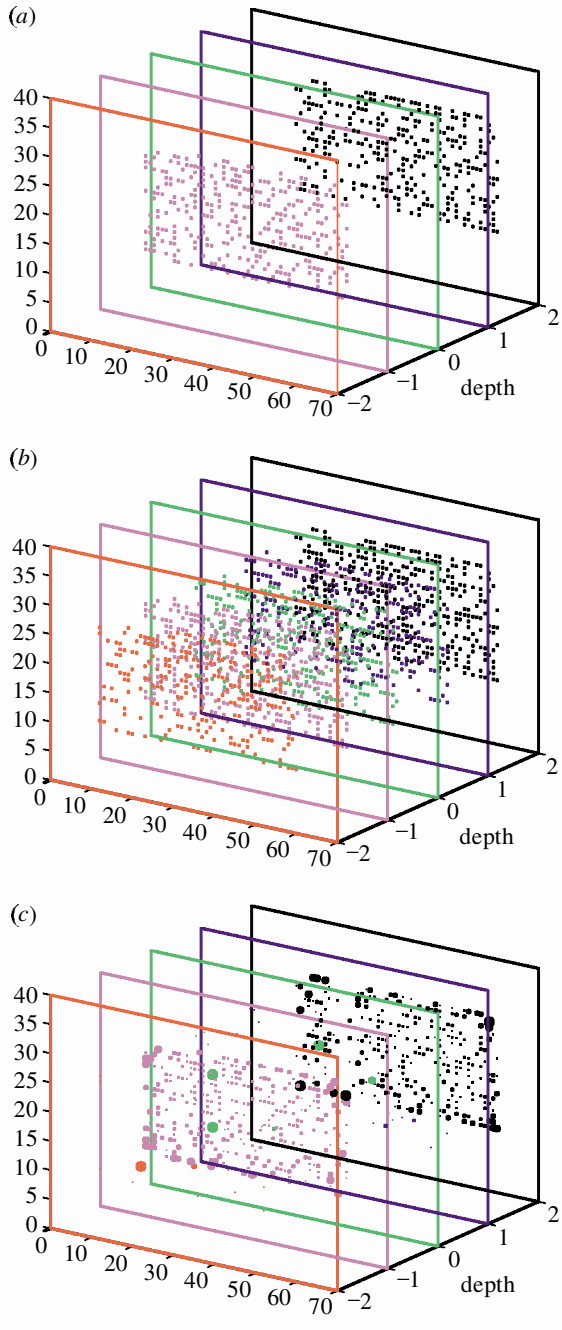


Figure 6. Modelling transparency of two depth planes at depths of $d = -1$ and $d = 2$. Note that there are a few ghost dots outside the two depth planes in the model outputs, as occurs in human vision (Weinshall 1989). (*a*) 3D scene, (*b*) actual input to the model and (*c*) time-average model activity.

solving the stereo correspondence problem that has been extensively studied previously (Marr & Poggio 1976; Prazdny 1985; Pollard *et al.* 1985; Qian & Sejnowski 1989; Marshall *et al.* 1996). Addressing both segmentation and correspondence at the same time is important since there was an apparent conflict at the level of neural mechanism between these two computational goals—segmentation requires mutual inhibition and the correspondence requires mutual excitation between nearby cells that are tuned to similar disparities. We demonstrate that it is possible that the same, biologically based, neural circuit of intracortical interactions can achieve both goals in such a model. In particular, not only does the model find the correct binocular matches given ambiguous inputs, but it also gives response highlights to more salient input features, even though the input strengths for these features are no stronger than those for the less salient ones. The dynamics of the model involve recurrent interactions to address different computational goals, correspondence and segmentation at different timescales. Such dynamics are essential for the model's successful performance. Our model thus accounts for recent physiological data (Von der Heydt *et al.* 2000; Bakin *et al.* 2000; Cumming & Parker 2000) and links them with

psychological behaviour (Julesz 1971; Prazdny 1985; Nakayama & Silverman 1986).

The model predicts that there should be positive temporal correlation at zero time-delay between activities of V2 cells responding to the same depth object (surface), and this correlation should be negligible or negative between cells that respond to different depth objects. The model neural connection structure also predicts that, in addition to intracortical connections linking cells tuned to similar depth (Ts'o & Gilbert 1988), there should be disynaptic inhibitory connections between pyramidal cells that are tuned to different depths but have the same or overlapping monocular RFs.

To focus on the problem of stereo correspondence and pre-attentive depth-edge highlights, our model is highly simplified, and lacks the elements and mechanisms needed to process scale, orientation, colour and motion in visual inputs. However, the conceptual lessons obtained from this focused study are expected be applicable to extended studies in the future to address more complex phenomena and process more realistic inputs such as slanted or curved surfaces. Better computational performances are expected in an extended model that includes other input dimensions through cooperation between different feature dimensions.

I thank Y. Petrov, A. Popple, N. Qian, S. Ullman and especially P. Dayan for discussions, P. Dayan and P. Adams for very helpful comments on an earlier draft, and the Gatsby Charitable Foundation for support.

REFERENCES

Bakin, J. S., Nakayama, K. & Gilbert, C. D. 2000 Visual responses in monkey areas V1 and V2 to three-dimensional surface configurations. *J. Neurosci.* **20**, 8188–8198.

Barlow, H. B., Blakemore, C. & Pettigrew, J. D. 1967 The neural mechanism of binocular depth discrimination. *J. Physiol. (Lond.)* **193**, 327–342.

Cumming, B. G. & Parker, A. J. 2000 Local disparity not perceived depth is signaled by binocular neurons in cortical area V1 of the macaque. *J. Neurosci.* **20**, 4758–4767.

Geiger, D., Ladendorf, B. & Yuille, A. 1995 Occlusions and binocular stereo. *Int. J. Comput. Vis.* **14**, 211–226.

Gilbert, C. D. & Wiesel, T. N. 1983 Clustered intrinsic connections in cat visual cortex. *J. Neurosci.* **3**, 1116–1133.

Hirsch, J. A. & Gilbert, C. D. 1991 Synaptic physiology of horizontal connections in the cat's visual cortex. *J. Neurosci.* **6**, 1800–1809.

Hubel, D. H. & Wiesel, T. N. 1970 Cells sensitive to binocular depth in area 18 of the macaque monkey cortex. *Nature* **225**, 41–42.

Julesz, B. 1971 *Foundations of cyclopean perception*. University of Chicago Press.

Li, Z. 1999*a* Contextual influences in V1 as a basis for pop out and asymmetry in visual search. *Proc. Natl Acad. Sci. USA* **96**, 10 530–10 535.

Li, Z. 1999*b* Visual segmentation by contextual influences via intracortical interactions in primary visual cortex. In *Network, computation in neural systems*, vol. 10, pp. 187–212. See http://www.iop.org/EJ/S/3/37/journal/0954-898X.

Li, Z. 2002 A saliency map in primary visual cortex. *Trends Cogn. Sci.* **6**, 9–16.

McLoughlin, N. P. & Grossberg, S. 1998 Cortical computation of stereo disparity. *Vis. Res.* **38**, 91–99.

Marr, D. & Poggio, T. 1976 Cooperative computation of stereo disparity. *Science* **15**, 283–287.

Marr, D. & Poggio, T. 1979 A theory of human stereopsis. *Proc. R. Soc. Lond.* B **204**, 301–328.

Marshall, J. A., Kalarickal, G. J. & Graves, E. B. 1996 Neural model of visual stereomatching: slant, transparency, and clouds. *Network: Comput. Neural Syst.* 7, 635–669.

Nakayama, K. & Silverman, G. H. 1986 Serial and parallel processing of visual feature conjunctions. *Nature* **320**, 264–265.

Nasrabadi, N. N., Clifford, S. P. & Liu, Y. 1989 Integration of stereo vision and optical flow by using an energy-minimization approach. *J. Opt. Soc. Am.* A **6**, 900–907.

Poggio, G. F. 1992 Physiological basis of stereoscopic vision. In *Vision and visual dysfunction.* 9. *Binocular vision* (ed. J. R. Cronly-Dillon & E. Regan). Boca Raton, FL: CRC Press.

Pollard, S. B., Mayhew, J. E. & Frisby, J. P. 1985 A stereo correspondence algorithm using a disparity gradient limit. *Perception* **14**, 449–470.

Prazdny, K. 1985 Detection of binocular disparities. *Biol. Cybern.* **52**, 93–99.

Qian, N. & Sejnowski, T. J. 1989 Learning to solve random-dot stereograms of dense and transparent surfaces with recurrent backpropagation. In *Proc. 1988 Connectionist Models Summer School* (ed. D. Touretzky, G. E. Hinton & T. J. Sejnowski), pp. 435–443. San Mateo, CA: Morgan Kaufman.

Read, J. C. A. 2002 A bayesian approach to the stereo correspondence problem. *Neural Comput.* **14**, 1371–1392.

Rockland, K. S. & Lund, J. S. 1983 Intrinsic laminar lattice connections in primate visual cortex. *J. Comp. Neurol.* **216**, 303–318.

Ts'o, D. & Gilbert, C. 1988 The organization of chromatic and spatial interactions in the primate striate cortex. *J. Neurosci.* **8**, 1712–1727.

Von der Heydt, R., Zhou, H. & Friedman, H. S. 2000 Representation of stereoscopic edges in monkey visual cortex. *Vis. Res.* **40**, 1955–1967.

Watanabe, O. & Fukushima, K. 1999 Stereo algorithm that extracts a depth cue from interocularly unpaired points. *Neural Networks* **12**, 569–578.

Weinshall, D. 1989 Perception of multiple transparent planes in stereo vision. *Nature* **341**, 737–739.

White, E. L. 1989 *Cortical circuits*. Basel: Birkhäuser.

Zhang, L. I. & Poo, M.-M. 2001 Electrical activity and development of neural circuits. *Nature Neurosci.* **4**, 1207–1214.

GLOSSARY

RF: receptive field