Moving from spatially segregated to transparent motion: a modelling approach

Szonya Durant, Alejandra Donoso-Barrera, Sovira Tan and Alan Johnston

Biol. Lett. 2006 **2**, 101-105
doi: 10.1098/rsbl.2005.0379

biology letters

Biol. Lett. (2006) **2**, 101–105
doi:10.1098/rsbl.2005.0379

Moving from spatially segregated to transparent motion: a modelling approach

Szonya Durant[1,*], **Alejandra Donoso-Barrera**[2], **Sovira Tan**[3] and **Alan Johnston**[2,3]

[1]*Colour, Form and Motion Laboratory, Visual Perception Unit, School of Psychology, The University of Sydney, Sydney, NSW 2006, Australia*
[2]*Centre for Mathematics and Physics in the Life Sciences and Experimental Biology (CoMPLEX), University College London, Gower Street, London WC1E 6BT, UK*
[3]*Department of Psychology, University College London, Gower Street, London WC1E 6BT, UK*
**Author for correspondence* (*szonyad@psych.usyd.edu.au*).

Motion transparency, in which patterns of moving elements group together to give the impression of lacy overlapping surfaces, provides an important challenge to models of motion perception. It has been suggested that we perceive transparent motion when the shape of the velocity histogram of the stimulus is bimodal. To investigate this further, random-dot kinematogram motion sequences were created to simulate segregated (perceptually spatially separated) and transparent (perceptually overlapping) motion. The motion sequences were analysed using the *multi-channel gradient model* (McGM) to obtain the speed and direction at every pixel of each frame of the motion sequences. The velocity histograms obtained were found to be quantitatively similar and all were bimodal. However, the spatial and temporal properties of the velocity field differed between segregated and transparent stimuli. Transparent stimuli produced patches of rightward and leftward motion that varied in location over time. This demonstrates that we can successfully differentiate between these two types of motion on the basis of the time varying local velocity field. However, the percept of motion transparency cannot be based simply on the presence of a bimodal velocity histogram.

Keywords: motion transparency; population coding; gradient model

1. INTRODUCTION

One of the fundamental problems in the neural computation of motion is the grouping of velocity signals into surfaces (Braddick 1993). In the case of motion transparency local moving elements appear to be grouped into two or more spatially overlapping surfaces. Thus the challenge for models of motion transparency (Snowden & Verstraten 1999) is to demonstrate how two different motion signals can appear perceptually co-localized in space. This might require a multi-valued representation for each point in the image or the co-localization of more global surface descriptors.

Motion transparency can be generated using random-dot kinematogram (RDK) displays. These are composed of many dots moving with different directions and/or speeds, whose statistical properties can be controlled by the experimenter (Braddick 1997). The aggregate motion of RDKs is difficult to compute because global motion at the scale of the whole random-dot display coexists with the local motion at smaller scales, i.e. that of individual dots or smaller clusters of dots.

Computational approaches to optic flow estimation often incorporate smoothing or averaging of the vector field (Horn & Schunk 1981), which can also result in the cancelling out of opposing motions. This is not what happens in the transparent case, as we in fact see two or more surfaces moving over one another. Other models determine velocity from the mode of the distribution of the velocity vectors (Wilson *et al.* 1992; Yuille & Grzywacz 1988). However, this 'winner-take-all' approach does not give a satisfactory result when we have two or more winners. Alternatively, the perception of transparent motion could be represented by different populations of cells that signal motion independently of each other. Again, this is problematic, since some psychophysical studies (Mather & Moulden 1983; Snowden 1989) have concluded that superimposed patterns suppress each other's detectability, implying the presence of inhibitory connections in the visual system, and Qian *et al.* (1994*a*) have shown that arrays of paired dots moving directly across each other do not give rise to transparency.

A plaid formed by two superimposed sinusoidal gratings of different orientation and direction of motion can appear to move as a single structure in a single direction (Adelson & Movshon 1982). Typically, a population of V1 cells tested for a range of directions produce two peaks in activity to moving plaids, with each peak corresponding to cells with preferred directions of the components of the plaid (Movshon *et al.* 1985). Unlike these component cells, pattern cells (about one-quarter of the cells tested in the middle temporal area, MT), will respond best when the plaid, not the constituent components, is moving in their preferred direction (Movshon *et al.* 1985). Pattern cell responses, therefore, reflect the combination of separate motion signals, whereas the population of component cell responses might provide the neural substrate for transparency. Further work by Qian *et al.* (1994*b*) and Snowden *et al.* (1991) has shown that MT cells give a reduced response when motion in their preferred direction is combined with motion in the non-preferred direction, and that this suppression has a broad direction tuning.

It has been suggested that whether transparency is seen or not depends upon whether the recovered velocity distribution of the stimulus is unimodal or bimodal (Jasinschi *et al.* 1992). Jasinschi *et al.* reported that for two patterns (consisting either of curves, features, straight lines or opaque ellipses), motion coherency, transparency or a mixture of both was perceived depending on whether the shape of the calculated velocity histogram was unimodal, bimodal

The electronic supplementary material is available at http://dx.doi.org/10.1098/rsbl.2005.0379 or via http://www.journals.royalsoc.ac.uk.

Received 25 July 2005
Accepted 9 August 2005

or trimodal. However, the structure of the velocity histogram drawn from the pattern may not be sufficient to determine the percept. Treue *et al.* (2000) have shown that two combined directions of motion give rise to a single peaked population response in area MT if the difference in direction is less than 90° but subjects still report transparency for direction differences of as little as 10° (Mather & Moulden 1980). This indicates that a bimodal population response is not *necessary* for the perception of transparency; but is the presence of a bimodal velocity histogram *sufficient* for generation of a transparent percept?

van Doorn & Koenderink (1982) studied the spatial and temporal parameters affecting the detectability of coherent and transparent motions by presenting RDKs that were divided into horizontal regions of different widths. When the strips contained vertically moving dots the percept changed from segregation through incoherence to transparency as the strip width was reduced. In another experiment adjacent strips contained random-dots moving horizontally in opposite directions. In this case, there was a simple transition from segregation to transparency. They found that if the strips were wide, coherently moving dots with alternating directions for each region were perceived. When the strips were narrow, observers saw two patterns moving across each other transparently. In these two cases the velocities of the dots are the same, but the velocity field is arranged differently.

In the present study, these segregated and transparent patterns were analysed by applying the multichannel gradient model (McGM) optic flow algorithm (Johnston *et al.* 1999*a*; Dale 2002; Johnston *et al.* 2003), which computes the angle and magnitude of the velocity at every pixel in every frame of the stimuli. The McGM is a biologically plausible model that has been successful in predicting perceived motion in a range of studies (Johnston & Clifford 1995*a*; Johnston & Clifford 1995*b*; Johnston *et al.* 1999*b*) and which produces a readily analysable velocity histogram. We assume that the output from the model represents a spatial map of the neural response in area MT of primates. However, we are not making any special claims for the model here—it simply provides a filter-based method of computing local velocity. We expect that other methods for computing local velocity fields would deliver similar results. The velocity fields produced by the McGM were compared and used to construct velocity histograms corresponding to the various strip widths that give rise to the segregated and transparent percepts. It delivers a bimodal population response like that discussed by Treue *et al.* (2000) while also providing us with a spatial and temporal map of the motion signals. Note that this approach differs from biological models that deliver an activation distribution over a population of velocity-tuned elements, each of which overlaps the input space (Simoncelli & Heeger 1998; Watanabe & Kikuchi 2005). Our approach delivers a velocity field map, with a velocity vector for each pixel at each time frame as well as a function that varies over velocity, in our case a velocity histogram. Nevertheless we would expect similar results and conclusions from population code models since in both cases transparency leads to bimodal distributions and there is significant aggregation over space.

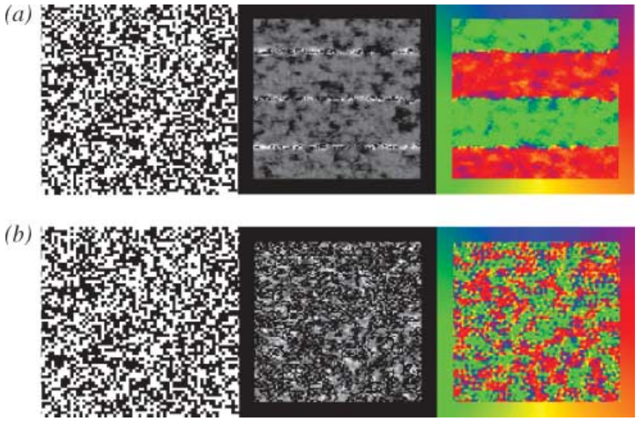

Figure 1. McGM output examples: (*a*) the stimulus is coherent motion consisting of four rows of random-dots moving in opposite directions (first frame). The middle frame shows the speed map, which represents the different speeds at every pixel of stimulus by different shades of grey. The black border is set to a speed of zero; white represents 2 pixels per frame. The third frame shows the plotted direction map. It gives the direction of motion for every pixel of the original stimulus. The direction is coded by the colour. The border around the third frame indicates the correspondence between direction and colour, i.e. red indicates movement to the right, while green, movement to the left. (*b*) Same as (*a*) but for a stimulus perceived as transparent motion consisting of two overlapping random-dot surfaces moving in opposite directions and at the same speed (64 rows of alternating motion). Notice that in the third frame left (red) and right (green) motion predominates.

2. APPARATUS AND STIMULI

Stimuli were sequences of 100 bitmap frames of size 128×128 pixels generated in MATLAB v.6. Each motion sequence generated consisted of RDK displays with horizontal strips of motion of different thicknesses consisting of equal amounts of 2×2 pixel black and white random-dots. The black and white dots were in equal proportion, which along with the dot size (1.1 min arc) was kept consistent with the psychophysical studies of van Doorn & Koenderink (1982). In the model 1 pixel represents 32.5 s arc (Johnston & Clifford 1995*a*). Alternating strips were programmed to move in opposite directions at 1 pixel per frame, which represents $9\,^{\circ}\,\mathrm{s}^{-1}$, chosen from the mid-range of velocities used by van Doorn & Koenderink (1982). In the model one frame represents 7.8 ms (Johnston & Clifford 1995*a*). The McGM parameters of the zero order kernel were $\sigma=1.5$, $\alpha=10$, $\tau=0.25$, with an integration window of 11×11 pixels. See Johnston & Clifford (1995*a*) for an explanation of the parameters.

Four different motion strip widths were chosen to match the percepts described by van Doorn & Koenderink (1982):

(i) 32 pixels. Four rows of random-dots moving in alternating opposite directions at the same speed; this was matched to the case when subjects saw segregated motion.
(ii) 4 pixels. In this case, the percept was intermediate between segregation and transparency; this represented the intermediate case.
(iii) 3 pixels. This was intended to match the perceived transparent motion case, perceived as two superimposed random-dot patterns moving in opposite directions.
(iv) 2 pixels. This was also representing perceived transparent motion.

The frame sequences were analysed using the McGM. For every frame of the stimulus, the model outputs a speed map and a direction map (see figure 1). One movie sequence was produced for each exemplar (see electronic supplementary material for examples of the 32 pixel and 2 pixel cases).

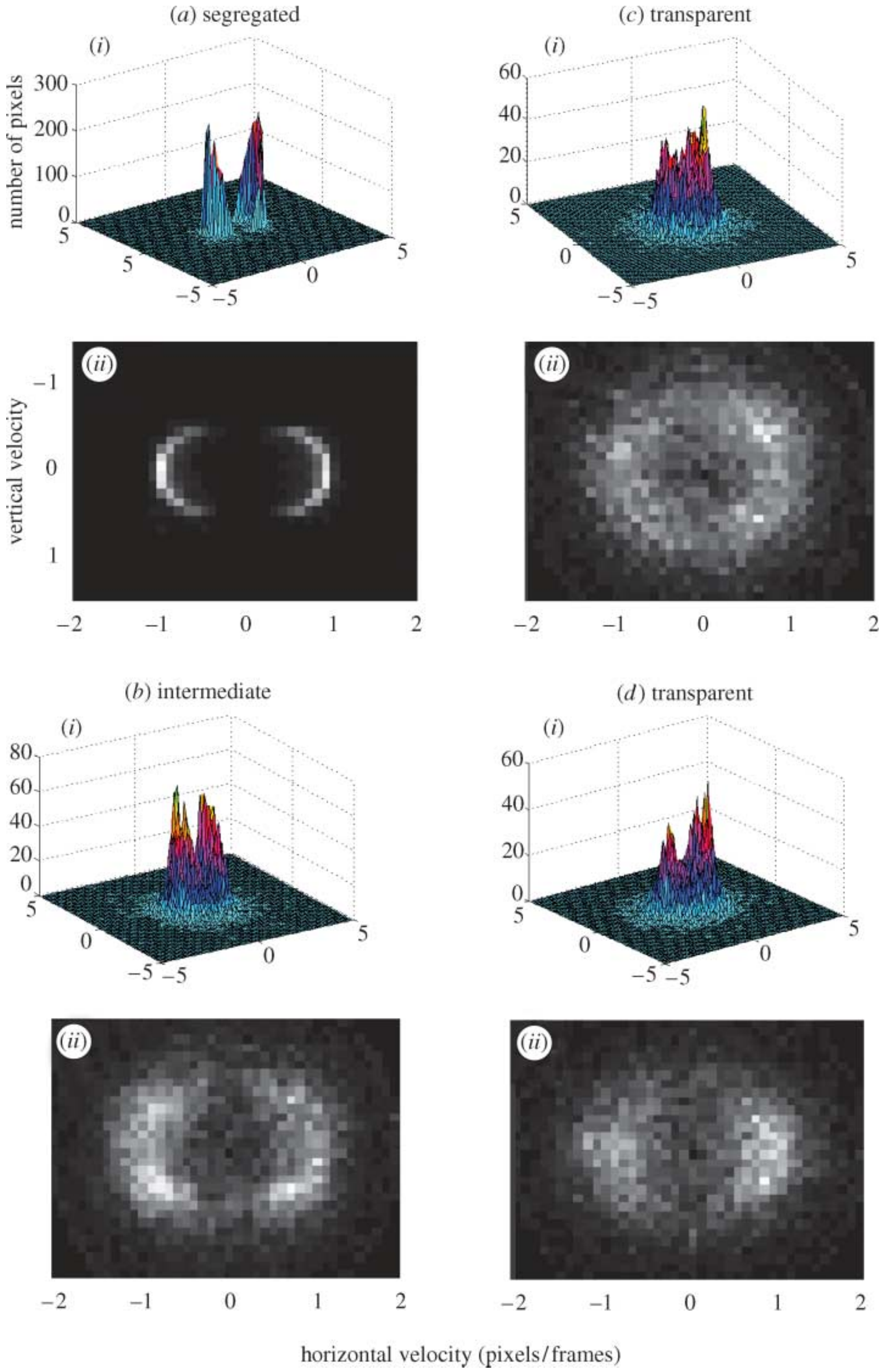

Figure 2. *a*–*d*(*i*) Velocity histograms from one frame of output from segregated (4 rows), intermediate (30 rows) and transparent (40 and 64 rows) motion RDK stimuli, respectively; (*ii*) velocity distribution plots as a grey intensity map for the four different stimuli. Each area corresponds to a velocity; the brighter the area, the more pixels from the output contain that velocity.

3. RESULTS AND DISCUSSION

The speed and angle of direction computed for every pixel of each frame of the stimuli was obtained and the output for the first frame was plotted for the four and 64 row cases (see figure 1). Although both velocity fields contain predominantly leftward and rightward motion, they clearly differ in spatial structure from the coherent to the transparent case. Moreover, the pattern of the velocity field for the coherent motion remains the same over frames, whereas the spatial distribution of the transparent case varies. The McGM output clearly is able to differentiate between the two cases.

Figure 2*a*(*i*)–*d*(*i*) show the velocity histograms from one frame for each stimulus. A velocity histogram is a three-dimensional plot of the number of pixels that correspond to each velocity vector. Figures 2*a*(*ii*)–*d*(*ii*) show the corresponding velocity distribution plots, with brightness indicating the amount of pixels at each velocity.

The velocity histogram and distribution plot for the segregated motion stimulus (*a*)(*i*)–(*ii*) show that the distribution is clearly bimodal and highly symmetric, as expected for two opposing directions of motion. For both stimuli consisting of transparent motion, (*c*) (*i*)–(*ii*) and (*d*)(*i*)–(*ii*), the distribution is bimodal, and also shows some symmetry. For the ambiguous case *b*(*i*)–(*ii*), the distribution is again bimodal. However, there is some noise present in these histograms, revealed by the number of pixels that contain low velocities (towards the centre of the distribution plot) and the tails of the distributions. Comparing the velocity histograms over ten frames of output, we find that for the four row case the mean peak of the velocity distribution is 0.96 (s.e. 0.004) pixels per frame, and with 64 rows of motion the mean peak of the velocity distribution is 0.90 (s.e. 0.017) pixels per frame, a significant drop of 6.25%. The mean fitted Gaussian width of the velocity distribution was 0.23 (s.e. 0.01) pixels per frame for the 4 row and 0.76 (s.e. 0.01) pixels per frame for the 64 row stimulus. Hence, it appears that transparent motion produces lower velocities and a broader distribution of velocities.

4. CONCLUSION

The velocity histograms calculated using the McGM for both transparent motion sequences and segregated motion sequences containing two opposing directions of motion are bimodal. This demonstrates that the presence of a bimodal velocity distribution is not a sufficient condition for the perception of transparency. Treue *et al.* (2000) have previously shown that a bimodal population response is not a necessary condition for transparency. Velocity distributions from the McGM outputs resemble each other; however, the local velocity fields for the segregated and the transparent cases are qualitatively markedly different, demonstrating the ability of the local velocity field to differentiate between the individual cases. It is also worth noting that the pattern of the transparent velocity field reflects the subjective percept of small patches of coherent motion translating in the two opposite directions as part of two layers of motion—this percept is often referred to as lacy transparency, indicating a time-dependent space-variant attribution of foreground and background.

Remarkably, although the blur kernel for the spatio-temporal filters in the McGM extends over 23 pixels spatially and 23 frames (Johnston & Clifford 1995*a*), the system as a whole can resolve 1 pixel wide rows of spatially interleaved motion. The calculation of speed is compromised in this situation. The peak velocities for the ambiguous and transparent stimuli histograms shift to lower velocities and are broader when compared to the coherent motion case. However, MT cells also give a reduced response to their preferred direction, when this signal is part of two overlapping motion signals. This inhibition in individual cells would result in an over all reduced response in MT. Thus the McGM, which is designed as a neuromorphic model of motion perception, is successful in distinguishing between segregated motion and motion transparency and is also successful at reproducing the reduction in the response of MT cells to transparent motion. We have only considered the responses of the model to random-dot motion. Other motion stimuli such as plaids can also sometimes appear transparent, but the degree of transparency is affected by factors other than movement such as whether the brightness of intersections conforms to the physics of luminance transparency (Stoner *et al.* 1990; Noest & Vandenberg 1993). Motion transparency is also influenced by the introduction of a disparity cue to surface segmentation alongside the motion direction cue (Bradley *et al.* 1995; Verstraten *et al.* 1994). However both these issues are beyond the scope of the current model, although eventually one would like to draw parallels between all forms of transparency (Gerbino *et al.* 1990). Our main aim here is to emphasize, through simulation, that the key to motion transparency is not the structure of the velocity histogram but the *variability over time* of the spatial location of areas of consistent motion signals, which leads to the grouping of two overlapping planes of motion, in contrast to the spatial segregation of regions of coherent motion.

This research was supported by a Royal Society International Fellowship to S.D., a Cancer Research UK studentship to A.D.B. and a Human Frontiers Science Program grant to A.J.

Adelson, E. H. & Movshon, J. T. 1982 Phenomenal coherence of moving visual patterns. *Nature* **300**, 523–525. (doi:10.1038/300523a0)

Braddick, O. J. 1993 Segmentation versus integration in visual motion processing. *TINS* **16**, 263–268.

Braddick, O. 1997 Local and global representation of velocity: transparency, opponency, and global direction perception. *Perception* **26**, 995–1010.

Bradley, D. C., Qian, N. & Andersen, R. A. 1995 Integration of motion and stereopsis in middle temporal cortical area of macaques. *Nature* **373**, 609–611. (doi:10.1038/373609a0)

Dale, J. L. 2002 A real-time implementation of a neuromorphic optic flow algorithm. Ph.D. thesis, University College, London.

Gerbino, W., Stultiens, C. I. F. H. J., Troost, J. M. & Deweert, C. M. M. 1990 Transparent layer constancy. *J. Exp. Psychol. Hum. Percept. Perform.* **16**, 3–20.

Horn, B. K. P. & Schunk, B. G. 1981 Determining optic flow. *Artif. Intell.* **17**, 185–204. (doi:10.1016/0004-3702(81)90024-2)

Jasinschi, R., Rosenfeld, A. & Sumi, K. 1992 Perceptual motion transparency: the role of geometrical information. *J. Opt. Soc. Am. A* **9**, 1865–1878.

Johnston, A. & Clifford, C. W. G. 1995*a* A unified account of three apparent motion illusions. *Vision Res.* **35**, 1109–1123. (doi:10.1016/0042-6989(94)00175-L)

Johnston, A. & Clifford, C. W. G. 1995*b* Perceived motion of contrast modulated gratings: predictions of the multichannel gradient model and the role of full-wave rectification. *Vision Res.* **35**, 1771–1783. (doi:10.1016/0042-6989(94)00258-N)

Johnston, A., McOwan, P. W. & Benton, C. P. 1999*a* Robust velocity computation from a biologically motivated model of motion perception. *Proc. R. Soc. B* **266**, 509–518. (doi:10.1098/rspb.1999.0620)

Johnston, A., Benton, C. P. & McOwan, P. W. 1999*b* Induced motion at texture-defined motion boundaries. *Proc. R. Soc. B* **266**, 2441–2450. (doi:10.1098/rspb.1999.0620)

Johnston, A., McOwan, P. W. & Benton, C. P. 2003 Biological computation of image motion from flows over boundaries. *J. Physiol. Paris* **97**, 325–334. (doi:10.1016/j.jphysparis.2003.09.016)

Mather, G. & Moulden, B. 1980 A simultaneous shift in apparent directions: Further evidence for a 'distribution shift' model of direction coding. *Q. J. Exp. Psychol.* **32**, 325–333.

Mather, G. & Moulden, B. 1983 Threshold for movement direction: two directions are less detectable than one. *Q. J. Exp. Psychol.* **35**, 513–518.

Movshon, J. A., Adelson, E. H., Gizzi, M. S. & Newsome, W. T. 1985 The analysis of moving visual patterns. In *Pattern recognition mechanisms* (ed. C. Chagas, R. Gattas & C. G. Gross), pp. 117–151. Berlin: Springer.

Noest, A. J. & Vandenberg, A. V. 1993 The role of early mechanisms in motion transparency and coherence. *Spatial Vision* **7**, 125–147.

Qian, N., Andersen, R. A. & Adelson, E. H. 1994*a* Transparent motion perception as detection of unbalanced motion signals: I. Psychophysics. *J. Neurosci.* **14**, 7357–7366.

Qian, N., Andersen, R. A. & Adelson, E. H. 1994*b* Transparent motion perception as detection of unbalanced motion signal. II. Physiology. *J. Neurosci.* **14**, 7376–7380.

Simoncelli, E. P. & Heeger, D. J. 1998 A model of neuronal responses in visual area MT. *Vision Res.* **38**, 743–761.

Snowden, R. J. 1989 Motions in orthogonal directions are mutually suppressive. *J. Opt. Soc. Am. A* **6**, 1096–1101.

Snowden, R. J. & Verstraten, F. A. 1999 Motion transparency: making models of motion perception transparent. *Trends Cogn. Sci.* **3**, 369–377. (doi:10.1016/S1364-6613(99)01381-9)

Snowden, R. J., Treue, S., Erickson, R. G. & Anderson, R. A. 1991 The response of area MT and V1 neurons to transparent motion. *J. Neurosci.* **11**, 2768–2785.

Stoner, G. R., Albright, T. D. & Ramachandran, V. S. 1990 Transparency and coherence in human motion perception. *Nature* **344**, 153–155. (doi:10.1038/344153a0)

Treue, S., Hol, K. & Rauber, H. J. 2000 Seeing multiple directions of motion—physiology and psychophysics. *Nat. Neurosci.* **3**, 270–276. (doi:10.1038/72985)

van Doorn, A. J. & Koenderink, J. J. 1982 Spatial properties of the visual detectability of moving spatial white noise. *Exp. Brain. Res.* **45**, 189–195.

Verstraten, F. A. J., Verlinde, R., Fredericksen, R. E. & van de Grind, W. A. 1994 A transparent motion aftereffect contingent on binocular disparity. *Perception* **23**, 1181–1188.

Watanabe, O. & Kikuchi, M. 2005 Nonlinearity of the population activity to transparent motion. *Neural Netw.* **18**, 15–22.

Wilson, H. R., Ferrera, V. P. & Yo, C. 1992 A psychophysically motivated model for two-dimensional motion perception. *Visual Neurosci.* **9**, 79–97.

Yuille, A. L. & Grzywacz, N. M. 1988 A computational theory for the perception of coherent visual motion. *Nature* **333**, 71–74. (doi:10.1038/333071a0)