

## **A computational analysis of separating motion signals in transparent random dot kinematograms**

JOHANNES M. ZANKER \*

*Department of Psychology, Royal Holloway, University of London, Egham, Surrey TW20 0EX, UK*

Received 25 September 2004; accepted 15 January 2005

**Abstract**—When multiple motion directions are presented simultaneously within the same region of the visual field human observers see motion transparency. This perceptual phenomenon requires from the visual system to separate different motion signal distributions, which are characterised by distinct means that correspond to the different dot directions and variances that are determined by the signal and processing noise. Averaging of local motion signals can be employed to reduce noise components, but such pooling could at the same time lead to the averaging of different directional signal components, arising from spatially adjacent dots moving in different directions, which would reduce the visibility of transparent directions. To study the theoretical limitations of encoding transparent motion by a biologically plausible motion detector network, the distributions of motion directions signalled by a motion detector model (2DMD) were analysed here for Random Dot Kinematograms (RDKs). In sparse dot RDKs with two randomly interleaved motion directions, the angular separation that still allows us to separate two directions is limited by the internal noise in the system. Under the present conditions direction differences down to 30 deg could be separated. Correspondingly, in a transparent motion stimulus containing multiple motion directions, more than eight directions could be separated. When this computational analysis is compared to some published psychophysical data, it appears that the experimental results do not reach the predicted limits. Whereas the computer simulations demonstrate that even an unsophisticated motion detector network would be appropriate to represent a considerable number of motion directions simultaneously within the same region, human observers usually are restricted to seeing not more than two or three directions under comparable conditions. This raises the question why human observers do not make full use of information that could be easily extracted from the representation of motion signals at the early stages of the visual system.

**Keywords:** Vision; motion perception; transparency; direction discrimination; computational modelling; direction distribution; integration.

---

\*E-mail: [j.zanker@rhul.ac.uk](mailto:j.zanker@rhul.ac.uk)

## 1. INTRODUCTION

The basic properties of human motion perception have been analysed in many psychophysical experiments using isolated moving objects or extended homogenous motion patterns (for review, see Sekuler *et al.*, 1990). However, the optic flow generated by moving observers in a three-dimensional world (Gibson, 1979; Nakayama, 1985) is characterised by complex spatial and temporal distributions of motion signals (Zanker and Zeil, 2002). In practice the visual system furthermore has to face dynamic noise superimposed on the pattern of motion signals, which can be due to as many causes as perturbations of the light travelling through optical media, foliage swirling in the wind occluding the view, or the limited transmission fidelity of the neural system itself. A critical test to assess strategies of extracting the behaviourally significant motion information could be the perception of transparent and tessellated motion surfaces, when multiple motion signals can appear in the same or neighbouring regions of the visual field (Smith *et al.*, 1999). The need for well defined psychophysical tasks and stimuli beyond the detection or recognition of motion-defined shapes (Braddick, 1980; Regan, 1991) was partially met by motion-defined stripes, which have been increasingly used in psychophysics during the last 20 years (Van Doorn and Koenderink, 1982; Watson and Eckert, 1994; Zanker, 1996). In a typical configuration these stimuli consist of stripe regions in which randomly distributed dots are moving in alternating directions parallel to the stripe boundaries. Such periodic motion stimuli can be easily segmented by human observers into stripes of coherent motion, as long as the patches are large enough. When the stripes are more narrow, however, they give rise to the sensation of transparent motion, i.e. both motion directions are perceived simultaneously in the complete stimulus field. Finally, with very narrow stripes the stimulus may no longer be discriminated from pure dynamic noise, i.e. dot patterns changing randomly in space and time. The processing of motion-defined gratings in the human visual system was modelled in a previous study (Zanker, 2001) with a biologically plausible motion detector network satisfactorily, predicting the transitions from segmentation to transparency, and from transparency to noise. The same two-dimensional motion detector model (2DMD) has been also used to account for a diverse range of motion processing phenomena (Zeil and Zanker, 1997; Zanker, 2004). Using alternating stripes defined by dots moving in opposite directions, the previous study did not address the issue of how motion many motion directions could be represented simultaneously by the 2DMD model as occupying the same region. This question will be investigated here together with the related question of the minimum difference between two transparent motion directions that still can be separated in the 2DMD output.

When a group of randomly distributed sparse dots is moving coherently in a single direction, one would expect the output of an ideal motion detector network processing this stimulus to be dominated by a sharp peak in the direction distribution at exactly this direction of dot motion. The properties of realistic motion detectors, as well as stimulus ambiguities such as local contour orientation or the interaction

between neighbouring dots, on the other hand, lead to the superposition of a certain amount of external and internal noise to the signal, and to a corresponding broadening of the direction distribution (Zanker, 2001). The variance of the local velocity measurements, determining the width of the direction distribution, will be the limiting factor for the ability to separate two adjacent directions of dot motion, because any mechanism involved in this task would need to discriminate two separate peaks in the joint signal distribution from a situation in which there is only a single peak (Green and Swets, 1966), and to identify the particular directions contributing to the stimulus. In the absence of discrete peaks in the direction distribution, higher-level analysers would be restricted to discrimination of differences in the overall shape of direction distributions, which would be sufficient to solve some transparency tasks (Treue *et al.*, 2000) but would not provide an explicit representation of transparent directions. Therefore the minimum angle between two adjacent directions in a transparent motion stimulus that a biological visual system can separate, as well as the maximum number of directions separable in such a stimulus, is expected to critically depend on the width of the direction distribution for each direction in its neural representation. The stimuli considered here are restricted to Random Dot Kinematograms in which individual dots are not providing orientation cues. This stimulus design restricts the potential of narrow tuning of spatio-temporal filters which has been discussed as possible computational solution to motion transparency ambiguities (Langley, 1999) for the case of sinewave gratings superimposed in plaids that may lead to coherent or transparent motion percepts. Therefore in the present context the separation performance is to be expected to be limited by the width of directional tuning of local motion detectors for moving dots. The minimum tuning width achieved in the human or primate visual system still is an open empirical question, but some electrophysiological evidence suggests that it might be considerable (Treue *et al.*, 2000).

The common strategy to reduce noise by spatial or temporal averaging, by which the coherent motion signals are amplified as compared to uncorrelated noise components, is not possible for transparent motion stimuli. Because groups of dots are moving in different directions in the same region and at the same time, any pooling over the full region would average all signals across the different directions and thus rule out separation. Indeed, the perception of an average direction rather than two transparent motion directions for dots moving in two or more directions in close neighbourhood has been interpreted as result of local integration mechanisms (Zohary *et al.*, 1996). This view is further supported by the finding that in inseparable bi-vectorial RDKs the perceived direction and speed quantitatively reflects the vector average of the components (Curran and Braddick, 2000). Therefore averaging strategies to reduce directional noise, in order to improve the separation of transparent motion, are limited by the constraint that the pooling of overlapping motion signals from different directions has to be kept at a minimum. One way of dealing with this dilemma would be to adopt more

specific pooling strategies, such as averaging motion signals preferentially along the trajectories of moving dots (Verghese *et al.*, 1999). Close proximity of oppositely moving dots, however, destroys the transparency percept entirely (Qian *et al.*, 1994), suggesting that at the very local level motion information — not surprisingly — is averaged indiscriminately of their membership to a particular direction group.

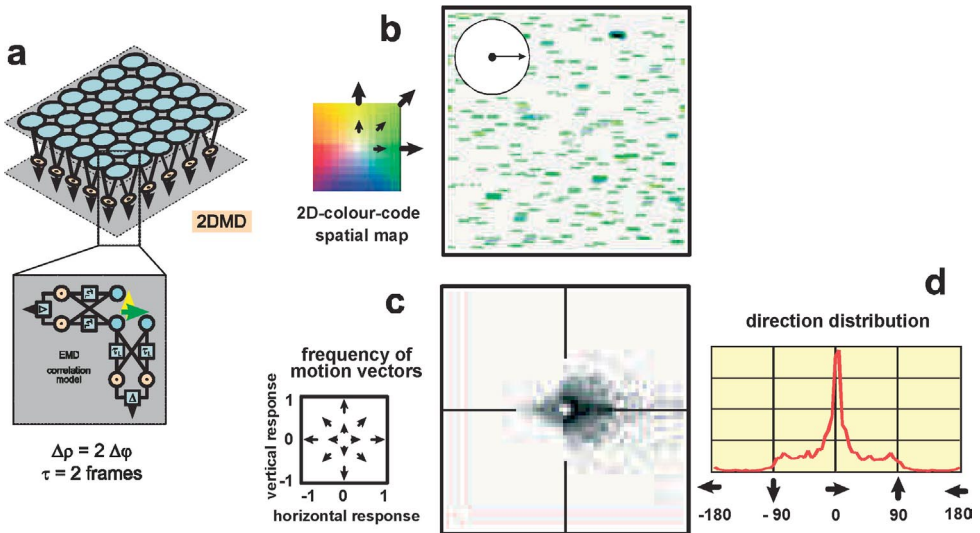
The present study is not presenting data of the actual performance of human observers, but is focussed instead on the question of what motion signals will be extracted in the early visual system by a simple motion detector network (Zanker, 2001), and what kind of information is provided for higher-level mechanisms (see, for instance, Snowden and Verstraten, 1999) to drive behavioural/perceptual decisions based on the motion signal distribution generated by this network. An elementary motion detector (EMD) model of the correlation type is used as basic building block (for review, see Reichardt, 1987), because it can account for a wide range of experimental data ranging from insect vision to human perception (Van Santen and Sperling, 1985; Borst and Egelhaaf, 1989). This model is a representative of luminance based motion detectors and could be replaced by other models without changing the major results (Adelson and Bergen, 1985; Watson and Ahumada, 1985). The crucial aspect of the local mechanism of motion detection is that it responds to spatio-temporally coupled displacements by combining signals from at least two points in space in an asymmetrical non-linear operation after sending one of the two signals through a temporal filter. With regard to these fundamental operations, which under certain conditions the models can be formally equivalent to those of other models (see Hildreth and Koch, 1987), the EMD model provides the minimum local motion information that the visual system needs to acquire when faced with the problem of motion transparency.

## 2. METHODS

The digital simulations (programs were written in Borland C++ 5.0 for a Pentium II PC under Windows 2000) used sequences of 8 image frames as motion stimuli. Each frame consisted of  $256 \times 256$  pixels with 8-bit grey level. A group of 256 bright dots, each covering  $2 \times 2$  pixels, were distributed randomly on top of a dark background; the contrast was 100%. Each dot was displaced by 2 pixels between two consecutive frames along a linear path for each frame of the complete sequence ('infinite lifetime') and was warped around the stimulus field when moving outside the frame boundaries such as to keep dot density constant. In the first set of simulations the dots were randomly allocated to two groups with two different motion directions ( $N_D = 2$ ). The separation between the two motion directions,  $\Delta_D$ , was treated as stimulus variable. In the second set of simulations each dot was randomly allocated to one of multiple motion directions, and the overall number of directions present in the transparent motion display ( $N_D$ ) being treated as stimulus variable. In this case all directions were equally spaced, in other words the angular

separation between neighbouring directions was the same for all groups of dots (i.e.  $\Delta_D = 360/N_D$ ).

The basic building blocks of the 2DMD model were elementary motion detectors of the correlation type (EMDs, see Reichardt, 1987). In a simple implementation (default circuit diagram is sketched at the bottom of Fig. 1a), each EMD receives input from two neighbouring locations of the stimulus patterns, which interact in a non-linear way (here by multiplication) after the temporal filtering of one input. Difference of Gaussians (DOGs) were used as spatial filters in each input line, with their excitatory centre and inhibitory surround balanced so as to remove DC components from the input (Marr and Hildreth, 1980; Srinivasan and Dvorak, 1980), thus effectively operating as bandpass filter on the incoming image. To prevent aliasing, the diameter of the receptive field (as measured between zero-crossings from excitatory to inhibitory regions) was set to about twice the sampling distance between the two inputs (Götz, 1965). This sampling distance was used as a fundamental spatial model parameter, and was kept constant for the simulations presented here at 4 image pixels. A first-order lowpass filter was used in the line leading from one spatially filtered input to the non-linear interaction, while the input to the non-linear operator was not temporally filtered. The time constant of the lowpass filter was treated as the fundamental temporal model parameter, usually



**Figure 1.** (a) Sketch of the 2DMD model consisting of an extended array (top) of orthogonal pairs of local motion detectors (bottom); (b) *motion signal map*, showing the spatial distribution of local motion signal responses for a sparse dot RDK moving from left to right; the hue and intensity of the 2D-colour-code (see inset) indicate the direction and strength of the motion response at a given location in the signal map; (c) *2D motion vector histogram*, representing the frequency of motion signals of a given horizontal and vertical strength by the greylevel of the corresponding bin; (d) the *direction distribution* shows the cumulative strength of all motion signals in the map as function of motion direction.

kept constant at the duration of two stimulus frames. The signal from one input line was multiplied with the temporally filtered signal from the other line, and two antisymmetric units of this kind were subtracted from each other, with equal weights leading to a fully opponent EMD with high directional selectivity (Borst and Egelhaaf, 1989).

Sequences of stimulus frames were processed by two 2D-arrays of orthogonal local motion detectors (two sets of  $256 \times 256$  EMDs, see Fig. 1a) which were either oriented along the horizontal or along the vertical axis of the digital images. This led to a two-dimensional spatial distribution of the 2DMD output, the *motion signal map*, after converting pairs of horizontal and vertical components for each image point into a local two-dimensional direction. In the examples shown here, this raw 2DMD output was subjected to temporal averaging over the eight frames of the stimulus sequence, to enhance the visibility of the motion signal patterns and to minimise dynamic artefacts of the motion detector. In the results section, some examples of these signal maps are plotted in a two-dimensional colour code (see inset of Fig. 1b) in which direction of motion is indicated by hue and strength of the response by saturation of the colour. The 2D-maps of motion responses were converted into *two-dimensional motion vector histograms* by digitising the local EMD responses at 8-bit resolution and then plotting the frequency for a given pair of horizontal and vertical responses at the corresponding  $x$ - $y$  histogram location. The greylevel of a given  $x$ - $y$  histogram bin indicates the number of occurrences for the corresponding horizontal and vertical motion component ( $x$ - and  $y$ -components in the histogram, see Fig. 1c). These histograms are scaled relative to the maximum signal strength so that the majority of response magnitudes are accommodated. The zero histogram bin was excluded from the analysis because it usually contains a huge number of counts from the static image regions which would bias the greylevel scale unnecessarily. These data plots are further simplified into *direction distribution* plots (see Fig. 1d), showing the weighted frequency of all signals observed in the motion signal map for each motion signal direction.

### 3. RESULTS

#### 3.1. Motion direction distributions in coherent RDKs

To demonstrate the function of the 2DMD model and to investigate the baseline condition, in Fig. 1 the simulation result is shown for the translation of 256 dots moving coherently from left to right. The motion signal map in Fig. 1a is characterised by a set of horizontal streaks that are generated by pooling the 2DMD output across all eight stimulus frames, during which the stimulus dots were moving horizontally. Closer inspection of the colour of these streaks reveals that they are not homogeneously green, as they should be if motion signals would indicate pure motion to the right: (i) some of the streaks are tinted slightly into the blue or yellow, typically in the regions where neighbouring dots get close to each other and are

like to interfere in the local motion detectors; (ii) there are fine fringes around each streak with a range of colours, reflecting local distortions of the perceived direction. This variation of local motion direction is reflected by the 2D motion vector histogram in Fig. 1c that shows a widespread distribution of responses around zero with a bias to the right quadrants. In particular, a large number of bins are filled in the neighbourhood of the right horizontal axis, demonstrating that on top of the broad distribution of local motion directions there is a focus on directions pointing rightwards. A simplified and clearer image is given by the direction distribution plotted in Fig. 1d, which shows a clear peak of motion signals for the horizontal right direction (0 deg), with a broad range of directions being observed for all directions with a right ward component (between  $-90$  and  $+90$  deg). This result confirms the expectation that even a coherent motion in one direction leads to a comparatively broad distribution of directions being represented in the motion detector output.

### 3.2. Separating two directions of motion

In the first set of simulations, the limits of separating two motion directions ( $N_D = 2$ , half of the 256 dots were displaced in each of the two directions) in transparent motion stimuli was tested by generating 2DMD outputs for a variation of angular separations ( $\Delta_D$ ) between the two directions. The results for four values of  $\Delta_D$  are shown in Figs 2a to 2d, to be compared with the control condition shown in Fig. 1. The first case, with orthogonal motion directions ( $\Delta_D = 90$  deg), shows two groups of motion streaks in the signal maps (left panel in Fig. 2a); one is oriented at  $+45$  angle in space and is dominated by yellow-green colours; the other one is oriented at  $-45$  deg and dominated by blue-green colours. Whereas these two groups of streaks in their colour and orientation clearly reflect the two groups of moving dots, once again is obvious that there is considerable noise on the signal. This fact is demonstrated by colours being present from virtually the whole palette, corresponding to the full range of motion directions. This pattern is confirmed by the broad distribution of motion signals around zero in the 2D motion vector histogram (right panel), with two regions of higher 2D-vector frequencies around the diagonals representing directions of  $+45$  and  $-45$  deg. Correspondingly, the direction distribution (small inset in right panel) shows two prominent peaks on top of a broad baseline covering the full range of motion directions — the two motion directions present in the stimulus are clearly represented and separable in this response, but they are accompanied by a considerable amount of stimulus-unspecific noise.

What happens when the angular separation  $\Delta_D$  is reduced? Figure 2b shows the signal map and histograms for  $\Delta_D = 60$  deg. For this stimulus, the oriented motion streaks in the signal maps are closer to each other than in Fig. 2a, the colours have less prominent yellow and blue tint, and the oriented bands in the 2D motion vector histograms (while still being very prominent) are getting closer to the horizontal axis. The direction distribution now shows two clearly separate peaks at about  $+30$  and  $-30$  deg, sitting on top of a broad distribution of noise directions. These

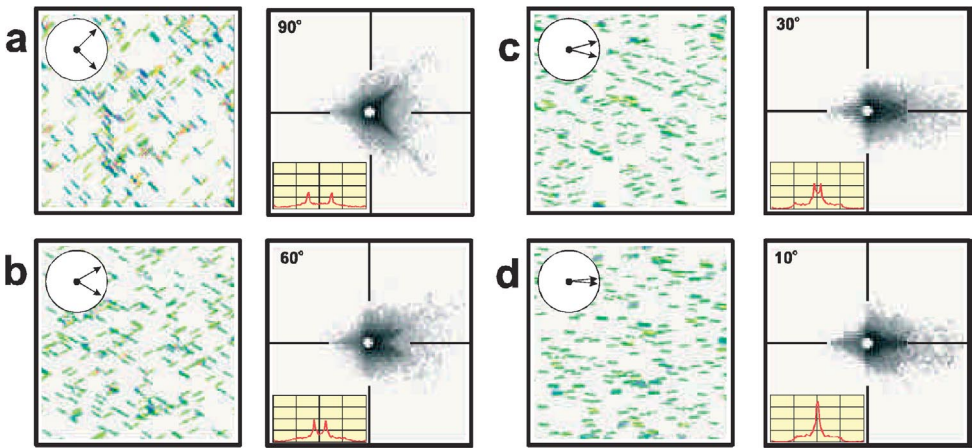
two direction peaks move closer together, but are still separable for  $\Delta_D = 30$  deg (Fig. 2c). However, for this case, the 2D motion vector histogram is already difficult to interpret, and the motion signal map shows a rather homogenous range of greenish motion streaks that still appear to represent to different groups as far as their orientation in space is concerned, but no longer lend themselves to be easily discriminated on the basis of their colour. Finally, for  $\Delta_D = 10$  deg (Fig. 2d), it seems no longer possible to attribute one or the other transparent motion direction to any of the motion streaks on basis of their spatial orientation or colour. In this case, the 2D motion vector histogram has merged into a single-lobed cloud with a centre of gravity on the right horizontal axis, and the direction distribution only shows a single peak at 0 deg, corresponding to rightwards motion. Comparing each of the panels representing the 2DMD motion response for  $\Delta_D = 10$  deg in Fig. 2d with their counterparts in Fig. 1 for a coherent motion stimulus with a single direction, it is obvious that these stimuli would be very hard to discriminate on the basis of the motion detector output. In conclusion, the simulations suggest that the minimum direction difference between two transparent motions that still can be separated under the present conditions would be expected for  $\Delta_D$  somewhere between 30 and 10 deg.

### 3.3. Separating multiple directions of motion

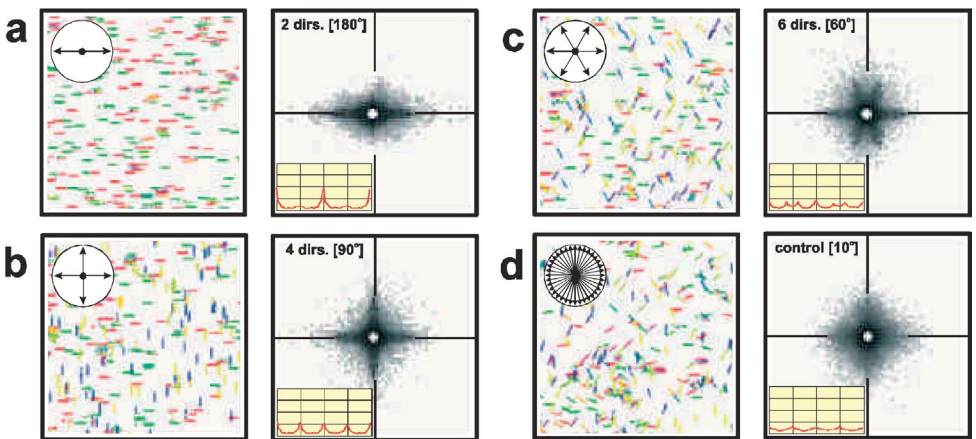
In the second set of simulations, the limits of separating two motion directions was compared to the maximum number of motion directions that can be separated from each other in a transparent motion stimulus. This question was studied by generating 2DMD outputs for transparent RDKs in which the number of motion directions,  $N_D$ , was varied and the angular separations ( $\Delta_D$ ) between directions was adjusted such that all directions were equally spaced. The results for four example settings of  $N_D$  are shown in Figs 3a to 3d.

The first instance, with two opposite motion directions ( $N_D = 2$ ,  $\Delta_D = 180$  deg), shows two groups of motion streaks in the signal maps (left panel in Fig. 3a). Both types of motion streaks are oriented horizontally, but one of them is dominated by green colours, and the other one is dominated by red colours. These two groups of streaks in their colour and orientation clearly reflect the two groups of moving dots, but once again there is considerable noise on the signal, as indicated by the range of colours being present apart from pure red and green. The 2D motion vector histogram (right panel) shows the corresponding broad distribution of motion signals around zero with two lobes of higher 2D-vector frequencies around the horizontal axis in both directions, reflecting the dots moving leftward and rightward. In the direction distribution (small inset in right panel), this has two prominent peaks (at 0 deg and 180 deg, respectively) on top of a broad baseline covering the full range of motion directions — just as in the case of a pair of perpendicular motion directions (Fig. 2a) the two stimulus directions are clearly separable from the 2DMD response, although they are accompanied by substantial noise.





**Figure 2.** Motion signal maps (left panels), 2D motion vector histograms (right panels) and directional distributions (insets) for transparent motion stimuli with dots moving in  $N_D = 2$  directions at variable angular separation  $\Delta_D$  between the two directions. (a)  $\Delta_D = 90$  deg; (b)  $\Delta_D = 60$  deg; (c)  $\Delta_D = 30$  deg; (d)  $\Delta_D = 10$  deg.



**Figure 3.** Motion signal maps (left panels), 2D motion vector histograms (right panels) and directional distributions (insets) for transparent motion stimuli with dots moving in a variable number directions  $N_D$  at equal spacing of angular separation  $\Delta_D$  between the superimposed directions. (a)  $N_D = 2$ ; (b)  $N_D = 4$ ; (c)  $N_D = 6$ ; (d)  $N_D = 32$ .

When the number of directions  $N_D$  is increased and the angular separation  $\Delta_D$  is reduced, a systematic change in the response pattern can be observed. Figure 3b shows the signal map and histograms for  $N_D = 4$  and  $\Delta_D = 90$  deg. For this stimulus, four groups of motion streaks appear in the signal maps with horizontal and vertical orientation, and the four cardinal direction colours, green, yellow, red and blue. The 2D motion vector histograms now have four prominent lobes, along all four cardinal directions (right, up, left, and down), correctly representing the four transparent directions in the stimulus. The direction distributions show the corresponding four peaks at about  $-90$ ,  $0$ ,  $90$  and  $180$  deg, clearly separate but sitting on top of a broad distribution of noise directions. Six peaks at  $60$  deg separation are visible in the direction distribution for  $N_D = 6$  and  $\Delta_D = 30$  deg (Fig. 3c), and the 2D motion vector histogram has lobes in horizontal and oblique directions, and the motion signal map shows a hexagonal set of motion streaks that clearly represent the different groups of moving dots in their orientation and colour. This type of behaviour continues up to a number of at least 8 directions but the separate peaks are difficult to identify for 12 directions or more (data not shown). The pattern for a large number of direction is shown in (Fig. 2d) for  $N_D = 32$  and  $\Delta_D = 10$  deg, where it seems no longer possible to attribute any one of the transparent motion directions contained in the stimulus to any of the motion streaks on basis of their spatial orientation or colour. In this case, the 2D motion vector histogram has merged into a amorphous cloud with a centre of gravity in the origin of the coordinate system, and the direction distribution is almost flat, without any prominent peaks in any relation to the stimulus directions. There is a small tendency to have larger signal frequencies around the cardinal axis which is due to the digitising of the 2DMD responses when they are converted to 2D vector histograms and direction distributions. This situation, however, is not distinguishable from a stimulus in which all dots move in random directions. In conclusion, the simulations suggest that the maximum number of transparent motion directions that could be resolved by the motion detector network under the present conditions would be somewhere between 8 and 12.

#### 4. DISCUSSION

The computational analysis of sparse dot random dot kinematograms, using a biologically plausible motion detector network (2DMD model) revealed three key points which are relevant with respect to the processing of transparent motion signals. (i) The direction distribution arising from the 2DMD response to translation of a coherently moving group of dots is by no means narrowly tuned to the motion direction of the stimulus dots; instead, the dominating response peak at the expected direction is embedded in a rather broad range of response directions. (ii) Two superimposed motion directions can be separated in the 2DMD output, by discriminating peaks in the signal direction distribution, down to a smallest angular separation of approximately  $30$  deg; stimuli with direction differences below

this angle cannot be distinguished from a stimulus containing a single coherent direction. (iii) The number of motion directions that can be simultaneously encoded in the 2DMD output, by discriminable peaks in the signal direction distribution, is clearly going beyond six, and may reach as many as twelve. If more motion directions are superimposed in a transparent motion stimulus, the response cannot be distinguished from random directions (direction noise). This is a consistent pattern of results, because the presence of substantial noise in the network response would predict a severe limitation of the minimum angular distance, and more specifically, the minimum separation of 30 deg would predict a separability of up to  $360/30 = 12$  directions. A previous computational study of the 2DMD model suggested that such an early vision model without extended spatial averaging quite nicely reflects the properties of human perception related to the segmentation of motion-defined gratings and the transition to the percept of transparency and to the percept of dynamic noise (Zanker, 2001). Is the predictive power of this model, despite its simple architecture, similar for the minimum angular separation and maximum number of directions that can be perceived in transparent motion stimuli?

There is some evidence that the minimum distance between the directions of two separable transparent motions goes down to a few tens of degrees, somewhat depending on stimulus conditions and tasks (Mather and Moulden, 1980; Williams and Sekuler, 1984). It is also evident that the directions contained in such a stimulus are not perceived veridically, with a systematic misjudgement of direction being known as 'motion repulsion' (Marshak and Sekuler, 1979; Hiris and Blake, 1996; Rauber and Treue, 1999; Benton and Curran, 2003). On the other hand, the maximum number of separable motion directions for human observers according to the currently available evidence appears to be restricted to a very small number (2–3) of directions when other cues to discriminate the groups of dots (such as orientation depth, colour, brightness) are excluded (Mulligan, 1993; Zanker, 2000). Because in the latter experiments observers were asked to discriminate different numbers of transparent directions from each other, or from noise, at present it is not fully clear whether they would be able to reliably detect the presence of a specific direction in a large number of transparent directions similar to the detection of coherent motion in noise (see Braddick and Qian, 2001). There are indications that separations between transparent motion directions need to be considerably larger than the minimum separation found previously, if direction information is required for the psychophysical task (Smith *et al.*, 1999). Preliminary results (see Felisberti and Zanker, 2004) suggest that, under experimental conditions that require observers to identify a particular direction in a transparent motion stimulus, the number of simultaneously visible motion directions is indeed restricted to 2–3, but furthermore that observers seem also to be less accurate for the minimum angular separation of two transparent directions than previously found with less specific psychophysical tasks (Mather and Moulden, 1980). Indeed, it appears that behavioural responses might be based on broad direction tuning as it is known for MT neurons, which would not allow for high accuracy recovery of direction-specific information for

transparent motion at small angular separations, but does allow to discriminate such transparent stimuli from single direction RDKs by the shape of the direction distribution encoded in the population of these neurons (Treue *et al.*, 2000). If higher level processing mechanism were to convolve the early direction distributions derived in the present study with a broad direction sensitivity function, all the narrow peaks would obviously disappear, leading to differences in shape of the filtered direction distributions as the only means to distinguish different stimuli, indeed.

In a more abstract sense, this pattern of similarities and discrepancies between simulations and psychophysical results suggests that the human capacity of information processing is substantially limited at some stage of the visual system beyond the initial encoding in the early stages of the motion stream. At face value, this suggestion is rarely surprising, but it may also invite an interpretation in the context of visual attention. In motion vision, attention is known not only to influence detection and discrimination thresholds (Ball and Sekuler, 1981; Raymond *et al.*, 1998) and after-effects (Chaudhuri, 1990), to limit the tracking of crowded targets (Pylyshyn and Storm, 1988), but it is also discussed as a mechanism that can facilitate or even generate peculiar types of motion sensations (Cavanagh, 1992; Lu and Sperling, 1995). In the RDK stimulus, neither a distinct object nor a specific location in the visual field is available to attract the focus of attention. Instead, attention is specifically attached to the particular stimulus attribute of motion direction (Treue and Trujillo, 1999), because other stimulus properties to guide attention like colour or spatial structure (Chawla *et al.*, 1999; Braun, 2000) are excluded by the stimulus design. It might be possible that human observers could split attention to multiple directions in transparent motion stimuli, similar to the assignment of attention to non-contiguous regions of the visual field (e.g. Castiello and Umiltà, 1992). The idea that by attending to a particular motion direction it could be possible to detect this direction in a transparent stimulus is currently under investigation. Indeed, some trained observers seem to be able to extract a given direction from many transparent directions if a pre-cue is provided prior to the stimulus (Felisberti and Zanker, 2004). So it might be that the discrepancies between the current simulation results and previous experimental evidence can be resolved by (i) more adequate experimental designs that address specifically the encoding and separation of particular motion directions (which may be related to different levels of processing in the visual stream), instead of testing the discrimination for overall direction distributions, and (ii) the consideration of attentional mechanisms that could affect performance on various levels of motion information processing.

Of course, the simple processing structure put forward here at best can be regarded as one possible component of the actual cortical processing. In particular, for reasons of parsimony, it disregards other kinds of information that are important for the segmentation process such as stereoscopically defined depth or colour composition (Qian and Andersen, 1997; Treue *et al.*, 2000). Indeed, there are various threads of experimental evidence that such information influences the response properties of motion-sensitive MT neurones (see Braddick and Qian,

2001). The emphasis here is on possible strategies to process motion signal distributions in the absence of any other source of information, which surely could be improved using certain additional criteria to select particular signal components, such as stereo or colour information labelling particular groups of dots. Of great interest will be the question whether the separation of a particular motion surface in a transparent motion stimulus can benefit from being exposed to an attentional focus (Raymond, 2000) as discussed above. One of the challenging questions for future research is to identify the neurones involved in such higher-level processing.

In summary, it was shown here by means of model simulations how a simple motion detector network would be able to deal with the complex spatial and spatio-temporal aspects of motion signal distributions arising from transparent motion stimuli, suggesting simple computational strategies to extract multiple directions in transparent motion stimuli. Comparing these computational considerations with results from psychophysical and electrophysiological experiments will give us clues to understand the functional architecture of the visual motion stream, and the specific processing roles of V1, MT and MST in the primate cortex, and may even provide us with new insights into potential mechanisms of attentional modulation of these regions.

### Acknowledgements

I am very grateful to the Vision Research Group at Royal Holloway for critical discussion of various aspects of this work, Fatima Felisberti for helpful comments on an earlier version of this manuscript, and to the BBSRC for financial support (# 111/S17110).

### REFERENCES

- Adelson, E. H. and Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion, *J. Opt. Soc. Amer. A* **2**, 284–299.
- Ball, K. and Sekuler, R. (1981). Cues reduce direction uncertainty and enhance motion detection, *Perception and Psychophysics* **30**, 119–128.
- Benton, C. P. and Curran, W. (2003). Direction repulsion goes global, *Current Biology* **13**, 767–771.
- Borst, A. and Egelhaaf, M. (1989). Principles of visual motion detection, *Trends in Neuroscience* **12**, 297–306.
- Braddick, O. J. (1980). Low-level and high-level processes in apparent motion, *Philosoph. Trans. Roy. Soc. (Biology)* **290**, 137–151.
- Braddick, O. J. and Qian, N. (2001). The organization of global motion and transparency, in: *Motion Vision: Computational, Neural and Ecological Constraints*, Zanker, J. M. and Zeil, J. (Eds), pp. 85–112. Springer, Berlin.
- Braun, J. (2000). Intimate attention, *Nature* **408**, 154–155.
- Castiello, U. and Umiltà, C. (1992). Splitting focal attention, *J. Exp. Psychol.: Human Perception and Performance* **18**, 837–848.
- Cavanagh, P. (1992). Attention-based motion perception, *Science* **257**, 1563–1565.
- Chaudhuri, A. (1990). Modulation of the motion after-effect by selective attention, *Nature* **344**, 60–62.

- Chawla, D., Rees, G. and Friston, K. J. (1999). The physiological basis of attentional modulation in extrastriate visual areas, *Nature Neuroscience* **2**, 671–676.
- Curran, W. and Braddick, O. J. (2000). Speed and direction of locally-paired dot patterns, *Vision Research* **40**, 2115–2124.
- Felisberti, F. M. and Zanker, J. M. (2004). Does attention affect transparent motion perception? *Perception* **33**, S 124.
- Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. Lawrence Erlbaum Associates, Hillsdale, NJ, USA.
- Götz, K. G. (1965). Die optischen Übertragungseigenschaften der Komplexaugen von Drosophila, *Kybernetik* **2**, 215–221.
- Green, D. M. and Swets, J. A. (1966). *Signal Detection Theory and Psychophysics*. Wiley, New York.
- Hildreth, E.-C. and Koch, C. (1987). The analysis of visual motion: From computational theory to neuronal mechanisms, *Ann. Rev. Neurosci.* **10**, 477–533.
- Hiris, E. and Blake, R. (1996). Direction repulsion in motion transparency, *Visual Neuroscience* **13**, 187–197.
- Langley, K. (1999). Computational models of coherent and transparent plaid motion, *Vision Research* **39**, 87–108.
- Lu, Z.-L. and Sperling, G. (1995). Attention-generation apparent motion, *Nature* **377**, 237–239.
- Marr, D. and Hildreth, E.-C. (1980). Theory of edge detection, *Proc. Roy. Soc. London B* **207**, 187–217.
- Marshak, D. W. and Sekuler, R. (1979). Mutual repulsion between moving visual targets, *Science* **205**, 1399–1401.
- Mather, G. and Moulden, B. (1980). A simultaneous shift in apparent directions: further evidence for a distribution-shift model of direction encoding, *Quart. J. Exper. Psychol.* **32**, 325–333.
- Mulligan, J. B. (1993). Motion transparency is restricted to two planes, *Investigat. Ophthalmol. Visual Sci.* **33**, 1049.
- Nakayama, K. (1985). Biological image motion processing: a review, *Vision Research* **25**, 625–660.
- Pylyshyn, Z. W. and Storm, R. W. (1988). Tracking multiple independent targets: evidence for a parallel tracking mechanism, *Spatial Vision* **3**, 179–197.
- Qian, N. and Andersen, R. A. (1997). A physiological for motion-stereo integration and a unified explanation of Pulfrich-like phenomena, *Vision Research* **37**, 1683–1698.
- Qian, N., Andersen, R. A. and Adelson, E. H. (1994). Transparent motion perception as detection of unbalanced motion signals. I. Psychophysics, *J. Neurosci.* **14**, 7357–7366.
- Rauber, H.-J. and Treue, S. (1999). Revisiting motion repulsion: evidence for a general phenomenon? *Vision Research* **39**, 3187–3196.
- Raymond, J. E. (2000). Attentional modulation of visual motion perception, *Trends in Cognitive Sciences* **4**, 42–50.
- Raymond, J. E., O'Donnell, H. L. and Tipper, S. P. (1998). Priming reveals attentional modulation of human motion sensitivity, *Vision Research* **38**, 2863–2867.
- Regan, D. (1991). Spatial vision for objects defined by colour, contrast, binocular disparity and motion parallax, in: *Vision and Visual Dysfunction 10 Spatial Vision*, Regan, D. (Ed.), pp. 135–178. Macmillan Press, Houndsmill.
- Reichardt, W. (1987). Evaluation of optical motion information by movement detectors, *J. Compar. Physiol.* **A161**, 533–547.
- Sekuler, R., Anstis, S. M., Braddick, O. J., Brandt, T., Movshon, J. A. and Orban, G. A. (1990). The perception of motion, in: *Visual Perception. The Neurophysiological Foundations*, Spillmann, L. and Werner, J. S. (Eds), pp. 205–230. Academic Press, San Diego, CA, USA.
- Smith, A. T., Curran, W. and Braddick, O. J. (1999). What motion distributions yield global transparency and spatial segmentation? *Vision Research* **39**, 1121–1132.
- Snowden, R. J. and Verstraten, F. A. J. (1999). Motion transparency: making models of motion transparency transparent, *Trends in Cognitive Sciences* **3**, 369–377.

- Srinivasan, M. V. and Dvorak, D. R. (1980). Spatial processing of visual information in the movement-detecting pathway of the fly, *J. Compar. Physiol.* **140**, 1–23.
- Treue, S. and Trujillo, J. C. M. (1999). Feature-based attention influences motion processing gain in macaque visual cortex, *Nature* **399**, 575–579.
- Treue, S., Hol, K. and Rauber, H.-J. (2000). Seeing multiple directions of motion — physiology and psychophysics, *Nature Neuroscience* **3**, 270–276.
- Van Doorn, A. J. and Koenderink, J. J. (1982). Spatial properties of the visual detectability of moving spatial white noise, *Exp. Brain Res.* **45**, 189–195.
- Van Santen, J. P. H. and Sperling, G. (1985). Elaborated Reichardt detectors, *J. Opt. Soc. Amer. A* **2**, 300–321.
- Verghese, P., Watamaniuk, S. N. J., McKee, S. P. and Grzywacz, N. M. (1999). Local motion detectors cannot account for the detectability of an extended trajectory in noise, *Vision Research* **39**, 19–30.
- Watson, A. B. and Ahumada, A. J. (1985). Model of human visual-motion sensing, *J. Opt. Soc. Amer. A* **2**, 322–342.
- Watson, A. B. and Eckert, M. P. (1994). Motion-contrast sensitivity: visibility of motion gradients of various spatial frequencies, *J. Opt. Soc. Amer. A* **11**, 496–505.
- Williams, D. W. and Sekuler, R. (1984). Coherent global motion percepts from stochastic local motions, *Vision Research* **24**, 55–62.
- Zanker, J. M. (1996). On the elementary mechanism underlying secondary motion processing, *Proc. Roy. Soc. London B* **351**, 1725–1736.
- Zanker, J. M. (2000). Motion transparency and multiple motion directions, *Investigat. Ophthalmol. Visual Sci.* **41**, S 720.
- Zanker, J. M. (2001). Combining local motion signals: a computational study of segmentation and transparency, in: *Motion Vision: Computational, Neural and Ecological Constraints*, Zanker, J. M. and Zeil, J. (Eds), pp. 113–124. Springer, Berlin.
- Zanker, J. M. (2004). Looking at Op Art from a computational viewpoint, *Spatial Vision* **17**, 75–94.
- Zanker, J. M. and Zeil, J. (2002). An analysis of the motion signal distributions emerging from locomotion through a natural environment, in: *Biologically Motivated Computer Vision 2002, Lecture Notes on Computer Vision 2525*, Bülthoff, H. H., Lee, S.-W., Poggio, T. and Wallraven, C. (Eds), pp. 146–156. Springer, Berlin.
- Zeil, J. and Zanker, J. M. (1997). A glimpse into crabworld, *Vision Research* **37**, 3417–3426.
- Zohary, E., Scase, M. O. and Braddick, O. J. (1996). Integration across directions in dynamic random dot displays: vector summation or winner take all? *Vision Research* **36**, 2321–2331.