Available online at www.sciencedirect.com
SCIENCE DIRECT®
Vision Research 45 (2005) 759–764
Vision Research
www.elsevier.com/locate/visres

# No effect of spatial phase randomisation on direction discrimination in dense random element patterns ☆

George Mather *, Andrew K. Daniell

*Department of Psychology, School of Life Sciences, University of Sussex, East Sussex, Brighton BN1 9QG, United Kingdom*

Received 28 May 2004; received in revised form 1 October 2004

**Abstract**

Two computational strategies have been proposed for motion analysis in the human visual system. Energy-based schemes involve detection of spatiotemporal Fourier energy in the frequency components comprising a moving pattern. Edge-based schemes track shifts in the position of local edges in the pattern over time. This paper describes a stimulus manipulation, spatial phase randomisation, that acts as a diagnostic test for the involvement of energy-based processes, and describes the results of two experiments which apply the manipulation to random element patterns. Both experiments compared direction discrimination performance in patterns before and after the spatial phase of their components was randomised in the Fourier domain. For dense patterns, there was no effect of phase randomisation on the maximum displacement supporting reliable direction discrimination, indicating that energy-based responses were dominant. For sparse patterns, a significant effect of phase randomisation was obtained, indicating a greater role for edge-based responses.


## 1. Introduction

There has been a long-running debate in the motion research literature concerning the computational strategy used by early visual processes to detect retinal image motion. Two general schemes have been proposed. One scheme, *energy-based*, involves detection of spatiotemporal Fourier energy in the frequency components comprising the pattern (e.g. Adelson & Bergen, 1985). The other scheme, *edge-based*, involves tracking shifts in the position of local edges in the pattern over time (e.g. Morgan, 1992). A number of psychophysical studies have found support for both schemes, but there is still some uncertainty concerning the stimulus conditions under which they operate. Early theories argued that they operated in different stimulus conditions (e.g. Braddick, 1974): energy-based ('short-range') detection is confined to small spatial displacements of dense patterns over short inter-stimulus intervals; edge-based ('long-range') detection occurs using sparse patterns displacing over large spatial distances at long inter-stimulus intervals. More recent papers (e.g. Smith & Ledgeway, 2001) favour the view that the two processes can operate simultaneously.

A perennial problem in studies of early motion processing is how to attribute a given set of psychophysical data to responses in one or the other of these processes. An empirically-based distinction (Braddick, 1974) was called into question by Cavanagh and Mather (1989). The first- versus second-order distinction does not map simply onto qualitatively different visual processes. In this paper we describe a stimulus manipulation based on spatial phase randomisation, that acts as a

☆ An early report of these experiments was presented at the European Conference on Visual Perception, Glasgow, 2002 (Mather & Daniell, 2002).

* Corresponding author. Tel.: +44 1273 678342.
*E-mail address:* g.mather@sussex.ac.uk (G. Mather).

0042-6989/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2004.10.001

diagnostic test for the involvement of energy-based processes in a psychophysical task, and describe the results of two experiments which apply the manipulation to random element patterns.

## 2. Spatial phase randomisation

The fundamental image primitive used in edge-based motion detecting schemes is a steep change in image intensity over space—an intensity edge. Shifts in edge position are used to encode motion. The fundamental primitive used in the energy-based scheme is a spatial frequency component of the pattern. Phase shifts over time in individual frequency components create energy signals for motion detection. This crucial difference between the two schemes offers a simple way to discriminate between them.

The edge structure of a pattern is determined by the spatial phase relationship between the pattern's frequency components. A single edge is created by aligning the phase of many frequency components so that, at a single point in space, the steepest luminance gradient in each component coincides. A complex pattern is defined partly by the particular phase relationships present in its many frequency components. Consequently, if component spatial phase is randomised, while retaining the original amplitude of each component, the coherent spatial structure of the pattern is destroyed. A visual process that depended on this structure, namely an edge-based motion-detecting scheme, would be severely disrupted by this manipulation, which effectively produces the densest pattern possible. (Note that this argument also applies to schemes based on other spatial primitives, such as luminance peaks, since these also require phase alignment between frequency components.)

On the other hand, energy-based motion detecting processes should be immune to spatial phase randomisation, since the motion energy content of the pattern is given by phase shifts within components rather than phase relationships between components.

To illustrate this difference between the schemes, Fig. 1(top) shows two one-dimensional random element patterns. The top-most pattern is binary (black–white), while the second pattern contains multiple grey-levels. Using the scheme depicted on the right, it is possible to create phase-randomised versions of these patterns, shown in Fig. 1(bottom). If a two-frame motion display is created from each of these patterns by introducing a short displacement, it is possible to compute motion energy as follows. First, the spatiotemporal Fourier transform of the two-frame display is computed; then motion energy is calculated by taking the ratio of summed energy in the rightward quadrants versus summed energy in the leftward quadrants. Fig. 2a shows the result of such a computation for a fixed displacement of 4 elements (the same pattern of results was obtained at other displacements). As expected, there is no difference in motion energy between the four patterns. On this basis one would predict no difference between the four patterns in psychophysical performance during a motion direction discrimination task.

Fig. 2b shows the mean separation between nearest-neighbour like-signed edges in the four patterns. This statistic has been used to define the information available for motion detection by edge-based processes (Morgan, 1992). The figure shows results using two edge-finding procedures. The open bars show the mean separation of zero-crossings in the second spatial derivative of each pattern. The filled bars show the mean separation of zero-crossings in the output of a band-pass spatial filter (CF 3.5 cpd, HHBW 2 octaves), corre-

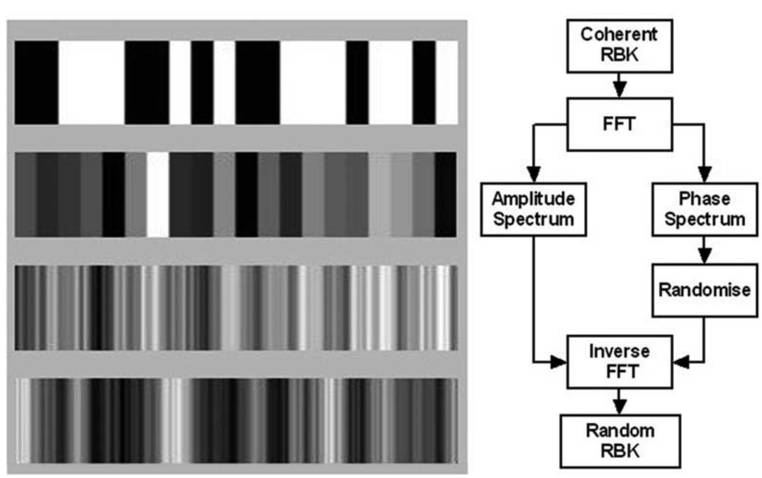


Fig. 1. Left: One-dimensional random element patterns of the kind used in the experiments. From top to bottom: a coherent binary pattern; a coherent multiple grey-level pattern; a phase-randomised binary pattern; a phase-randomised multiple grey-level pattern. Right: Flow chart illustrating the sequence of operations used to transform a coherent pattern into a phase-randomised pattern.

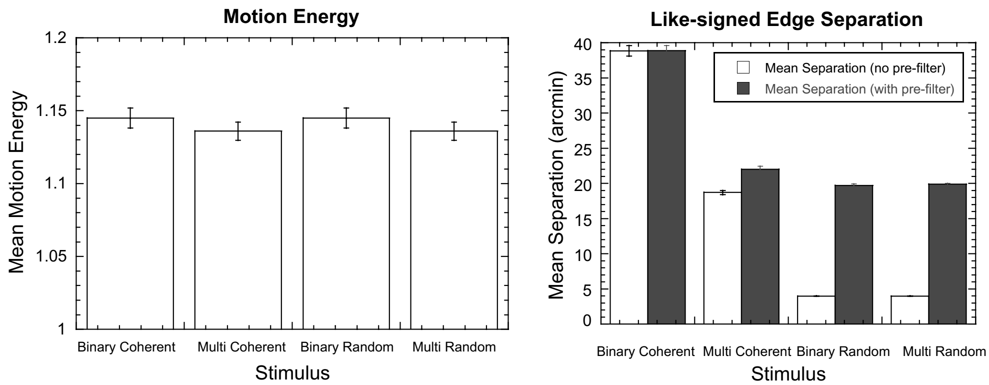


Fig. 2. Left: Motion energy for two-frame displays at a fixed displacement, computed using each of the four patterns illustrated in Fig. 1. Energy is given by the ratio of summed energy in the leftward quadrants and rightward quadrants of the spatiotemporal Fourier transform. Values represent the mean (±1 SD) of 20 repetitions of the computation using different patterns. Values >1 correspond to motion energy in the direction of displacement. Right: Mean separation between like-signed edges in the four patterns illustrated in Fig. 1, computed from zero-crossings in the second spatial derivative of the pattern either before (open bars) or after (filled bars) application of a 'pre-filter' (centre frequency 3.5 cpd, half-height bandwidth 2 octaves).

sponding to the 'pre-filter' adopted in the edge-based literature (e.g. Morgan, 1992). In both cases it is clear that edges are at least twice as far apart in the binary pattern than in the other three patterns. Phase-randomised patterns have the shortest edge separation (equivalent to highest density). On this basis one would predict reliable psychophysical performance in a motion direction discrimination task at larger displacements for the binary pattern than for the other three patterns.

Experiment 1 was designed to test the predictions shown in Fig. 2. Note that predictions are based on the presence or absence of statistically significant differences in psychophysical performance using different patterns, not on precise values of $D_{max}$.

## 3. Experiment 1

### 3.1. Method

#### 3.1.1. Subjects

One author and three naïve observers participated in the experiment.

#### 3.1.2. Apparatus

Stimuli were generated on a Sony Trinitron G400 monitor by a CRS VSG2/5 graphics system. Display linearity was established using a look-up table.

#### 3.1.3. Stimuli

The stimulus comprised four one-dimensional (1-D) strips of random elements (similar to those shown in Fig. 1) arranged to form a square against a uniform background, with a central fixation spot. Each strip subtended $3.21 \times 0.64$ deg arc ($160 \times 32$ screen pixels). This arrangement was adopted for two reasons: 1-D strips avoid the off-axis directional signals present in motion displays containing 2-D random element patterns; and the annular shape reduces the effect of retinal inhomogeneity on performance.

Four types of 1-D spatial pattern were used:

*Binary coherent:* Random binary elements (22.1 and 44.5 cd/sq m) at three element widths in different conditions (9.65, 19.3, 38.6 min arc; 8, 16, 32 screen pixels).
*Multi-coherent:* Each element was randomly assigned a luminance value from a Gaussian distribution (mean 33.9 cd/sq m; standard deviation 6.92 cd/sq m), with three element widths in different conditions (9.65, 19.3, 38.6 min arc).
*Binary random:* Phase-randomised versions of binary coherent patterns (mean luminance 33.9 cd/sq m; standard deviation 6.7 cd/sq m).
*Multi-random:* Phase-randomised versions of Multi-coherent patterns (mean luminance 33.9 cd/sq m; standard deviation 6.92 cd/sq m).

Stimuli were computed off-line as a library of 640-element arrays of 8-bit numbers. Each array defined a single experimental stimulus (four 160-pixel strips). Phase randomisation was performed using *Mathematica* as follows: (i) application of a discrete Fourier transform (DFT) to a given stimulus array; (ii) randomisation of the phase of the real part of the transform, such that any of the 0–2$\pi$ possible phase values were equally likely (the imaginary counterpart of each real component took on the complex complement of that component's phase); (iii) application of an inverse DFT. Amplitude was unaffected by the manipulation.

#### 3.1.4. Procedure

Subjects viewed the display using a chin rest from a viewing distance of 114 cm. Each trial consisted of a two-frame motion stimulus (FD 40 ms, no ISI), in which the random element pattern was displaced a fixed distance from frame 1 to frame 2. Four different frame-to-frame displacements were presented in different trials for each stimulus condition (in the range 19.3–77.2 min-arc). Displacement direction was selected randomly from trial to trial. In half of the trials the elements in each strip shifted to create an overall clockwise direction in the annulus. In the remaining trials the elements shifted anticlockwise. The subject pressed one of two response keys to indicate perceived direction of motion. A 1-sec interval separated successive trials, during which the screen was uniform (33.9 cd/sq m) except for the central fixation spot.

Each of the 48 stimuli was presented 40 times in total over a number of sessions (4 stimulus conditions × 3 element widths × 4 displacements). Stimulus order was pseudo-random: No stimulus condition was presented $n + 1$ times until all conditions had been presented $n$ times. No more than three successive trials were allowed to present the same stimulus direction.

### 3.2. Results

Each of the 12 combinations of stimulus condition and element width produced a psychometric function showing percentage correct direction discrimination as a function of displacement. Maximum displacement ($D_{max}$) supporting direction discrimination was defined as the displacement yielding 80% correct discrimination, and was calculated from the data by linear interpolation. Fig. 3 shows the mean and standard error of $D_{max}$ in each of the 12 conditions. The monotonic increase in $D_{max}$ with element size was highly significant ($F = 69.1$; df 2, 6; $p < 0.0001$). There was no significant difference in performance due to phase randomisation ($F = 3.47$; df 1,3; $p = 0.533$), and no interaction between element size and randomisation ($F = 0.064$; df 1, 3; $p = 0.817$).

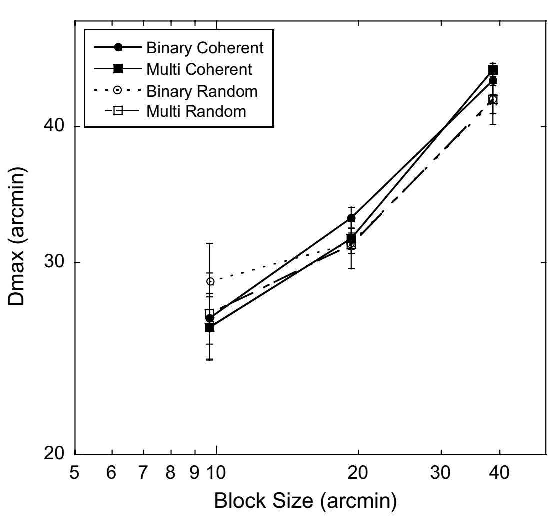


Fig. 3. Results of Experiment 1. Lines show mean $D_{max}$ (±1 SEM) as a function of element size, for each stimulus condition.

### 3.3. Discussion

Since there was no effect of spatial phase randomisation on performance, results are inconsistent with the operation of a motion process based on edge displacement. In defense of edge-based schemes, it could be argued that the predictions shown in Fig. 1 depend on the choice of pre-filter, and phase randomisation would be ineffective if a larger pre-filter was used. The use of different element sizes in the experiment, including relatively large blocks, was an attempt to avoid this problem. As Morgan (1992) showed, element size manipulations can be used to estimate pre-filter width. $D_{max}$ is constant for elements below the resolution of the pre-filter (individual elements cannot be resolved by the pre-filter), but increases in proportion with element size for elements larger than the pre-filter width (each element is resolved by the pre-filter). The fact that $D_{max}$ in our experiment did increase with element width indicates that individual elements could be resolved by the pre-filter. Yet there was no effect of phase randomisation even at larger block sizes.

Although this experiment found support only for energy-based motion processing of random element patterns, previous research indicates that evidence for edge-based processing can be found at low pattern densities (e.g. Baker & Hess, 1998; Boulton & Baker, 1993). Experiment 2 therefore investigated whether pattern density influences the effect of phase randomisation.

## 4. Experiment 2

Fig. 4(left) shows computed motion energy as a function of density. There is a small predicted effect of density on motion energy, but no effect of phase randomisation. Fig. 4(right) shows computed edge separation as a function of density. In coherent patterns, edge separation is reduced by half at 50% density compared to 12.5% density. In phase-randomised patterns there is no effect of density, but edge separation is reduced by a factor of between 4 and 40 compared to coherent patterns, depending on the use of a pre-filter.

To test whether performance agreed with energy-based predictions or with edge-based predictions, binary coherent and binary random patterns were used at one block size (19.3 min arc), and three densities (12.5%, 25%, 50%).

### 4.1. Method

#### 4.1.1. Subjects

Four subjects participated in the experiment, being the same subjects who had participated in Experiment 1.

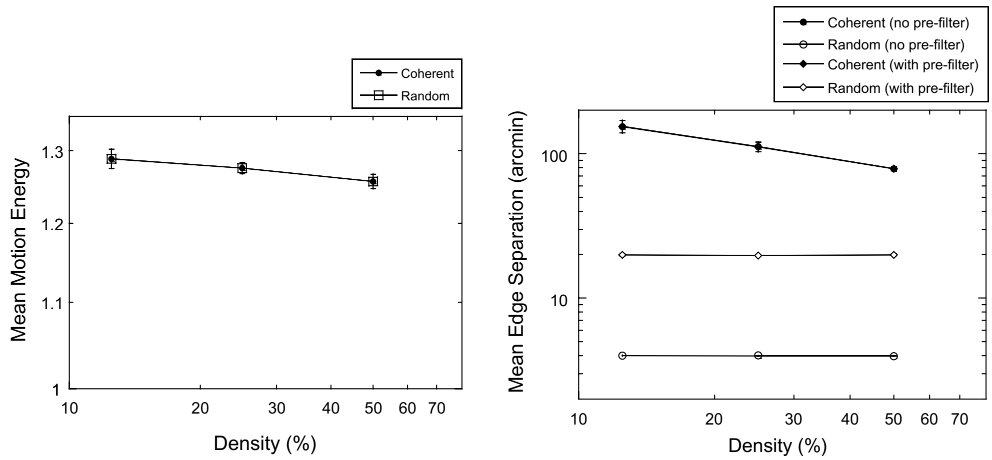


Fig. 4. Predicted motion energy (left) and edge separation as a function of pattern element density, computed in the same way as for Fig. 2.

*4.1.2. Stimuli, apparatus, and procedure*

All details were the same as in Experiment 1, with the exception that only one element size was used (19.3 min-arc), and three pattern densities (12.5%, 25%, 50%).

*4.2. Results and discussion*

As in Experiment 1, $D_{max}$ was computed from the psychometric function for each condition. Fig. 5 shows the mean and standard error of $D_{max}$ in each condition. There were significant effects of density ($F = 15.06$; df 2, 6; $p = 0.005$), and of phase randomisation ($F = 16.32$; df 1, 3; $p = 0.027$). However, Fig. 5 shows that the significant effect of phase randomisation can be attributed to the lowest pattern density, as confirmed by the presence of a significant interaction between density and randomisation ($F = 8.22$; df 2, 6; $p = 0.019$).

The energy-based scheme cannot account for the difference in psychophysical performance between coherent and randomised patterns at low densities. Results therefore support the view that direction discrimination is dominated by energy-based processes at 50% density, but edge-based processes at 12.5% density. Note that this conclusion is based on the assumption that the energy-based process simply sums energy linearly in the Fourier domain, though there is some support for this assumption (Schrater, Knill, & Simoncelli, 2000).

Recently, Bex and Dakin (2003) have used different techniques to reach similar conclusions about the involvement of both energy-and edge-based processes. In one experiment, they band-pass filtered noise patterns at two frequency bands, one low and the other high. They then combined the filtered patterns, using either the same source pattern, or different source patterns. In the former case (same-source), edges are correlated across the two filtered patterns. In the latter case (different-source), edges are uncorrelated across the two patterns. The lack of correlation in different-source images introduces many more edges in the pattern compared to same-source images. However, Bex and Dakin (2003) found no difference in $D_{max}$ between same-source and different-source dense noise patterns. This result is consistent with our finding that phase randomisation is ineffective in relatively dense patterns.

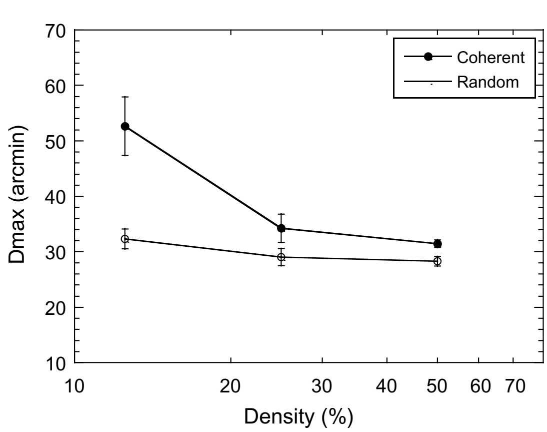


Fig. 5. Results of Experiment 2, showing mean $D_{max}$ (±1 SEM) as a function of pattern element density for two stimulus conditions.

## 5. General discussion

This paper has described a technique—spatial phase randomisation—that can be used to test for the involvement of energy-based processes in motion tasks. Immunity to phase randomisation indicates that performance is mediated by energy-based processes, whereas reduced performance using phase-randomised patterns indicates the involvement of edge-based processes. Two experiments indicated that pure energy-based processes mediate direction discrimination performance in high-density random element patterns, but edge-based processes become important at low densities.

Advocates of edge-based schemes may argue that all current models of edge detection involve an initial stage of filtering by spatial frequency-selective filters (Georgeson, 1992; Morrone & Burr, 1988; Watt & Morgan, 1985). A suitable combination of these filter outputs during edge detection may be able to accommodate immunity to phase randomisation. However, such immunity would only be possible if information in different frequency bands was treated independently. One could, for example, ignore all but the lowest spatial frequencies. However, an argument against low-frequency pre-filters was presented in the discussion of Experiment 1. More generally, it would be difficult to characterise any process behaving in this way as 'edge-based', since the essence of edges is their phase coherence across frequency.

It is remarkable that, although dense coherent and phase-randomised patterns differ so markedly in appearance (see Fig. 1), direction discrimination performance using the two patterns is so similar (see Fig. 3). This observation indicates that at the lowest levels of visual analysis, motion processing relies only on decomposing a pattern into frequency components (detecting phase shifts within frequency bands), and is not affected by the appearance of the pattern (governed largely by phase relationships between frequency components).

## Acknowledgements

This research was supported by a grant from the Wellcome Trust.

## References

Adelson, E. H., & Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America, A2*, 284–299.

Baker, C. L. J., & Hess, R. F. (1998). Two mechanisms underlie processing of stochastic motion stimuli. *Vision Research, 38*, 1211–1222.

Bex, P. J., & Dakin, S. C. (2003). Motion detection and the coincidence of structure at high and low spatial frequencies. *Vision Research, 43*, 371–383.

Boulton, J. C., & Baker, C. L. J. (1993). Different parameters control motion perception above and below a critical density. *Vision Research, 33*, 1803–1811.

Braddick, O. J. (1974). A short-range process in apparent motion. *Vision Research, 14*, 519–527.

Cavanagh, P., & Mather, G. (1989). Motion: The long and short of it. *Spatial Vision, 4*, 103–129.

Georgeson, M. A. (1992). Human vision combines oriented filters to compute edges. *Proceedings of the Royal Society of London, B249*, 235–245.

Mather, G., & Daniell, A. K. (2002). Separating energy-based and feature-based accounts of motion discrimination in random-dot kinematograms. *Perception, 31*(Suppl.), 99.

Morgan, M. J. (1992). Spatial filtering precedes motion detection. *Nature, 355*, 344–346.

Morrone, M. C., & Burr, D. C. (1988). Feature detection in human vision: A phase-dependent energy model. *Proceedings of the Royal Society of London, B235*, 221–245.

Schrater, P. R., Knill, D. C., & Simoncelli, E. P. (2000). Mechanisms of visual motion detection. *Nature Neuroscience, 3*, 9–11.

Smith, A. T., & Ledgeway, T. (2001). Motion detection in human vision: A unifying approach based on energy and features. *Proceedings of the Royal Society, B* 268, 1889–1899.

Watt, R. J., & Morgan, M. J. (1985). A theory of the primitive spatial code in human vision. *Vision Research, 25*, 1661–1674.