Available online at www.sciencedirect.com
SCIENCE DIRECT®
Vision Research 45 (2005) 3027–3049
Vision Research
www.elsevier.com/locate/visres

Local and global limitations on direction integration assessed using equivalent noise analysis

Steven C. Dakin [a,*], Isabelle Mareschal [a], Peter J. Bex [b]

[a] Department of Visual Science, Institute of Ophthalmology, University College London, 11-43 Bath Street, London EC1V 9EL, UK
[b] Department of Visual Rehabilitation, Institute of Ophthalmology, University College London, 11-43 Bath Street, London EC1V 9EL, UK

Received 8 September 2004; received in revised form 28 June 2005

Abstract

We used an equivalent noise (EN) paradigm to examine how the human visual system pools local estimates of direction across space in order to encode global direction. Observers estimated the mean direction (clockwise or counter-clockwise of vertical) of a field of moving band-pass elements whose directions were drawn from a wrapped normal distribution. By measuring discrimination thresholds for mean direction as a function of directional variance, we were able to infer both the precision of observers' representation of each element's direction (i.e., local noise) as well as how many of these estimates they were averaging (i.e., global pooling). We estimated EN for various numbers of moving elements occupying regions of various sizes. We report that both local and global limits on direction integration are determined by the number of elements present in the display (irrespective of their density or the size of region they occupy), and we go on to show how this dependence can be understood in terms of neural noise. Specifically, we use Monte Carlo simulations to show that a maximum-likelihood operator, operating on pooled directional signals from visual cortex corrupted by Poisson noise, accounts for psychophysical data across all conditions tested, as well as motion coherence thresholds (collected under similar experimental conditions). A population vector-averaging scheme (essentially a special case of ML estimation) produces similar predictions but out-performs subjects at high levels of directional variability and fails to predict motion coherence thresholds.

Keywords: Motion; Motion integration; Direction; Averaging; MT; Local/Global

1. Introduction

Neurons in primary visual cortex (V1) respond selectively to motion within a limited region of visual space (Hubel & Wiesel, 1962). Consequently, the visual system must pool the outputs of V1 cells across the visual field in order to estimate the direction of large moving objects. Computationally, there are sound reasons for averaging motion signals across space. First, ego-motion generates characteristic patterns of retinal motion known as *optic flow* (Gibson, 1979) and pooling local motion is necessary to allow subsequent decomposition of the flow field into its components (Koenderink, 1986). Second, pooling can assist in overcoming poor reliability of sparsely distributed local motion signals (Braddick, 1993) that may be noisy due to, e.g., intrinsic variability in cell spiking rate. Third, pooling direction estimates along a contour can assist in overcoming the *aperture problem* (i.e., that V1 neurons can only signal strength of one direction—perpendicular to their preferred orientation—rendering local motion signals ambiguous). However, spatial averaging of motion can also be disadvantageous, e.g., at the location of a motion boundary or between surfaces undergoing transparent motion. Under these conditions the visual system must both encode the presence of multiple directional sources yet still discount directional variability due to noise.

* Corresponding author. Tel.: +44 207 608 6988.
E-mail address: s.dakin@ucl.ac.uk (S.C. Dakin).

0042-6989/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2005.07.037

Convergent behavioural and electrophysiological data support the hypothesis that increasingly complex motion is processed within a hierarchy from V1 to the middle temporal (MT) and medial superior temporal (MST) areas. MT is demonstrably central to the processing of visual motion: lesions in this area impair monkeys' ability to discriminate motion direction (Newsome & Pare, 1988), and patients suffering damage to the human MT-homologue are rendered "motion blind" (Baker, Hess, & Zihl, 1991). Both anatomical (Lund, 1988) and physiological (Movshon & Newsome, 1996) data indicate that MT integrates directionally-tuned input from V1, as does the finding that responsiveness and directional tuning of MT neurons is greatly reduced after inactivation of V1 (Girard, Salin, & Bullier, 1992; Rodman, Gross, & Albright, 1989). In addition, a number of the requirements for processing complex stimuli are satisfied by MT. First, the directional tuning of MT neurons is broader than the tuning of the V1 neurons from which they receive input; thus they must be receiving input across a range of directions (Movshon & Newsome, 1996). Second, receptive fields of MT neurons are larger than their V1 counterparts and have either antagonistic or facilitative surrounds producing sensitivity to local motion boundaries, or to global motion direction over a large area, respectively (Born & Tootell, 1992). Third, the ambiguity intrinsic to the aperture problem appears to be resolved by area MT; Movshon, Adelson, Gizzi, and Newsome (1985) report that the responses of about 25% of MT neurons to plaids matched the perceived- or pattern-direction and not the directions of the components. Furthermore. Treue, Hol, and Rauber (2000) report that MT neurons do not necessarily produce a bimodal response under conditions leading to perceptual transparency implying that the components of multidirectional displays are encoded not by separate populations of neurons in MT but by the population response of the area as a whole.

In this paper, we concentrate on an attribute of complex visual motion thought to be encoded in area MT/MST: the *global direction of motion*. In the natural environment, we frequently have to judge the global direction of a group of objects each moving in a different direction (e.g., a flock of birds, turbulent water-flow, etc.). Some electrophysiological studies have investigated how the visual system does this. Newsome and co-workers used cortical micro-stimulation in awake primates to perturb MT activity in order to examine if the global-direction code in MT is based on a winner-take-all strategy (WTA; the identity of the most active directionally-tuned channel), or population-vector averaging (PVA; the average of all directions weighted by the magnitude of the neural response to each). The evidence is equivocal: data collected using a forced-choice discrimination task (Salzman, Britten, & Newsome, 1990) favoured WTA, but eye-movement data favoured PVA (Groh, Born, & Newsome, 1997). Nichols and Newsome (2002) propose that these differences arise because discrimination tasks impose perceptual categories on direction estimates, and argue that a task based on "free report" supports PVA over a wide directional range (although behaviour is more consistent with WTA for near-opposite directions).

In terms of computational models of MT function (e.g., Nowlan & Sejnowski, 1995; Simoncelli & Heeger, 1998; Wilson, Ferrera, & Yo, 1992), various strategies for direction encoding have been proposed sharing a number of common features. In particular, many extract local motion using a local energy mechanism (Adelson & Bergen, 1985) and then construct a histogram of the directional energy present using a bank of channels tuned to different directions. How might directional information be "read-off" from such a representation? One approach is the aforementioned PVA. However, PVA is known to be non-ideal under conditions of additive Gaussian noise and with narrowly tuned direction channels (Snippe, 1999). This is because vector averaging a population of noisy directional responses is formally identical to least-squares fitting a *cosine function* to the directional histogram (Seung & Sompolinsky, 1993). Maximum-likelihood estimation (MLE) is superior to PVA (for non-cosinusoidal directional variation) since the variance of its estimate of the global direction approaches the theoretical (Cramer-Rao) lower bound. MLE works by fitting appropriately shaped templates to the neural responses; its superiority results from these templates embodying extra "knowledge" about the noise. Furthermore, although PVA is vulnerable to system non-linearities (Snippe, 1999), Deneve, Latham, and Pouget (1999) propose that these very non-linearities could be a key component of biological instantiations of MLE which, they show, can be implemented using a recurrent network with activation functions undergoing divisive normalisation (a feature of MT neurons; Simoncelli & Heeger, 1998; Snowden, Treue, Erickson, & Andersen, 1991).

A limitation of both PVA and MLE is that they can signal only a single direction. Alternative schemes (Zemel, Dayan, & Pouget, 1998) decompose observed directional structure using a Gaussian mixture model and encode features such as multi-modality (e.g., arising from transparency) and directional variance. Such schemes are supported by the finding that perceived transparency seems to rely on a population code in MT (Treue et al., 2000). However, in terms of modelling psychophysics, such schemes have not addressed how one determines "overall direction" if a stimulus generates multiple perceived directions. This is also a problem for psychophysical investigations of perceived global direction. In the case of plaid stimuli, Wilson and Kim (1994) report that the PVA of the two components determines perceived direction. However, Zohary, Scase, and Braddick (1996) showed that subjects

adjusting the overall direction of random-dot stimuli generated with skewed directional distributions rely on PVA when directional signals were of similar strength, but switched to a more ''winner-take-all'' strategy when one directional signal dominated. The latter conditions may favour representation of multi-modal structure (Zemel et al., 1998).

Many psychophysical studies of global motion perception have relied on *motion coherence* tasks where one measures how many coherently moving dots can be replaced by randomly moving elements, while still supporting reliable discrimination of the overall direction (e.g., ''left'' versus ''right''). This paradigm was introduced by Newsome and colleagues (e.g., Britten, Shadlen, Newsome, & Movshon, 1992; Newsome, Britten, & Movshon, 1989) and has proved influential for both basic and clinical psychophysical studies of motion integration. For example, motion coherence has been used to estimate the directional-bandwidth (Meese & Harris, 2001) and spatial-frequency bandwidth (Bex & Dakin, 2002) of direction discrimination. In a clinical context, elevated motion coherence thresholds have been reported for subjects with visual deficits such as amblyopia (Simmers, Ledgeway, & Hess, 2005; Simmers, Ledgeway, Hess, & McGraw, 2003), retinitis pigmentosa (Turano & Wang, 1992) and glaucoma (Baez et al., 1995; Joffe, Raymond, & Chrichton, 1997) as well as for syndromes as diverse as hemiplegia (Gunn et al., 2002), autism (Milne et al., 2002), Williams syndrome (Atkinson et al., 1997), schizophrenia (Li, 2002), and dyslexia (Talcott, Hansen, Assoku, & Stein, 2000).

It has become commonplace to assume that high motion coherence thresholds indicate poor global pooling of motion signals across space, but this is not necessarily the case. Fig. 1A schematically depicts a typical motion coherence stimulus along with a hypothetical integration region (shaded area). Figs. 1B and C illustrate that an increase in coherence threshold can be caused either by poor estimation of local direction (i.e., of individual dots) or by an inability to pool these direction estimates. A number of authors have suggested that local noise—specifically the false correspondences arising between dots in successive frames of the stimulus—limits motion perception when the jump-size between frames becomes large (i.e., D_{max} paradigms; Braddick, 1974; Eagle & Rogers, 1996; Morgan & Ward, 1980). Barlow and Tripathy (1997) asked if such false correspondences might also limit motion coherence thresholds under varying stimulus conditions (e.g., density, size of region, etc.). Comparing their data to an ideal observer (which computed all possible element-matches between successive frames) led them to conclude that their results were consistent with observers being limited by local ''correspondence noise''. Nevertheless motion-coherence data are still frequently (particularly in a clinical context) used as an unambiguous index of global pooling.

Here we attempt to separate out the influences of local and global limits on motion integration using a novel application of the *equivalent noise* paradigm. Given that observers pool multiple motion signals across space, with each signal limited by local noise, the EN paradigm works by measuring the influence that additional *external noise* (in this case, directional variability) has on ones' ability to judge global/mean direction. The essential difference between this and motion coherence paradigms is that here all elements are ''signal'' elements in that they convey information about the overall direction. At low levels of external noise performance is limited both by local noise and by the extent of pooling, but at higher levels external noise swamps the influence of local noise so that performance comes to depend almost wholly on the extent of global pooling. We use this paradigm to examine the influence of various spatial parameters on the estimation of global direction.

2. General methods

2.1. Rationale

Our paradigm is a variant on one used to study orientation integration (Dakin, 2001). For a patch containing a number of moving elements we consider the effect on motion integration of the following parameters: patch-size, number of elements, and element-density within the patch. To determine which has the greatest influence on integration we cannot vary these parameters individually because varying one necessarily affects another, e.g., changing density within a fixed-size region necessarily changes the number of elements. Fig. 2 illustrates our paradigm, which overcomes this problem by co-varying all possible parameter pairs (e.g., size and number of elements) while keeping the third parameter constant (e.g., density), and then determining which parameter, when fixed, produces the *least* influence on integration across conditions. Our method for assessing motion integration is described next.

2.2. Assessing motion integration: Equivalent noise analysis

As discussed above, motion coherence cannot disentangle the effects of local and global limits on motion perception. To deal with this problem, we adapted the equivalent noise paradigm (Barlow, 1956) to direction integration. EN embodies the idea that visual integration is limited by two factors: *internal noise* and *sampling*. For directional integration of dot patterns these factors translate to: precision at estimating each dot's direction and the number of such estimates one can average over. An EN experiment for direction discrimination involves presenting moving elements whose

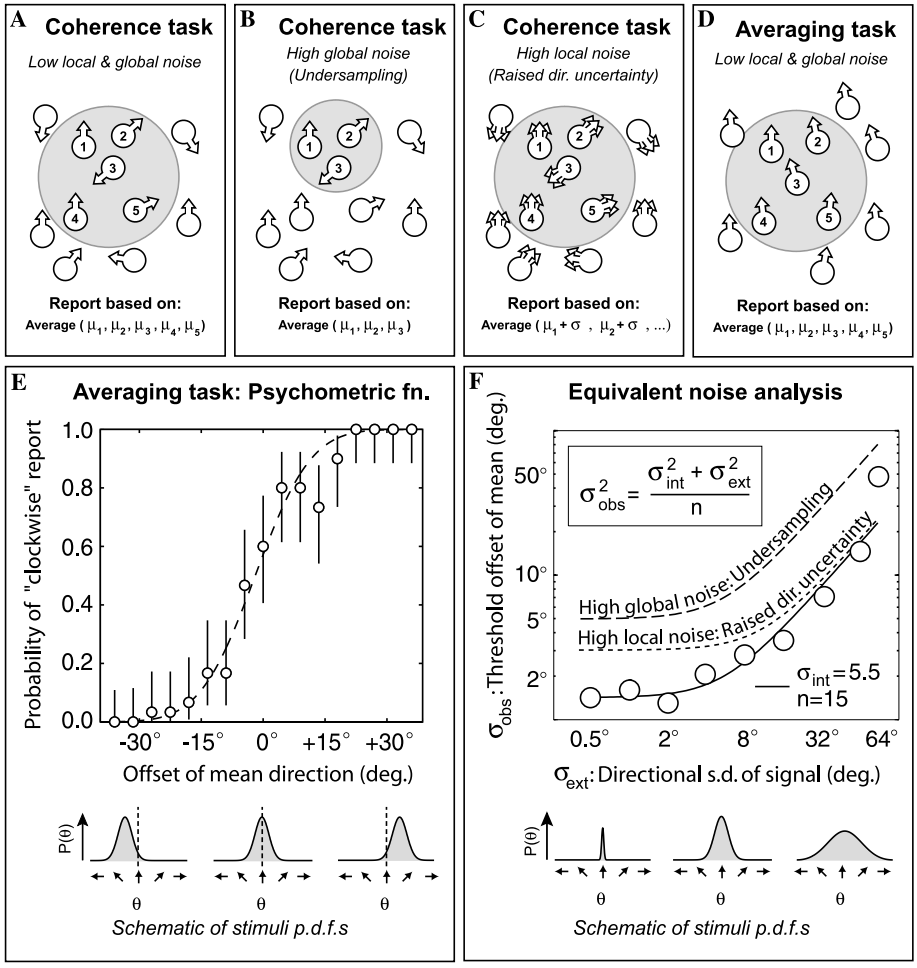

Fig. 1. (A–C) Schematic motion coherence stimuli. Subjects report the direction of a set of target-dots moving coherently (in this case vertically upwards) intermixed with noise dots moving in random directions. The performance measure is the maximum tolerable proportion of noise dots supporting direction discrimination. If observers are combining local motion estimates over some region of space (shaded region) threshold will be affected both by (B) global factors (i.e., the number of motion samples) and (C) local factors (i.e., the precision of each sample). The paradigm illustrated in (D–F) can separate these influences. (D) Now all dots are signal dots and have directions drawn from a wrapped Gaussian distribution. Observers report if the average direction is clockwise or anti-clockwise of vertical-upwards, and we measure a psychometric function (shown in E) for the probability of reporting "clockwise" as a function of the offset in mean direction. The performance measure is then the slope of the best-fitting wrapped cumulative Gaussian to these data—i.e., the threshold offset of the mean direction. We measure threshold as a function of the standard deviation of the directions present in the stimulus. (F) Because we estimate *response variability* (threshold) as a function of *stimulus variability*, equivalent noise exploits additivity of variance to fit the data (boxed equation); the fit-parameters quantify local (σ_{int}) and global (n_{samp}) limits on direction integration.

directions are drawn from a random distribution, and then estimating the smallest discriminable change in mean direction as a function of the amount of directional variability. Because the slope of a psychometric function is effectively a measure of the *variance* (of some internal representation of mean direction) one can exploit additivity of variance to predict the effect of noise on performance. Specifically, it has been shown that one can relate threshold offset in the mean (σ_{obs}) to the external noise (the variability imposed on the stimulus: σ_{ext}), the internal noise (σ_{int}), and the sample size (n_{samp}) by

$$\sigma_{\text{obs}} = \sqrt{\frac{\sigma_{\text{int}}^2 + \sigma_{\text{ext}}^2}{n_{\text{samp}}}}. \quad (1)$$

This formulation is identical to that developed by Watt and Morgan (1984) for estimating intrinsic blur and has been applied by Watamaniuk and Heinen (1999) within the motion domain (specifically, to compare noise for a visual pursuit task compared to a direction discrimination). Here, we identify two potential problems with applying this model to direction discrimination data, both related to the wrapping of directional data at 2π. The first is that one must allow for wrapping when fitting psychometric functions; Appendix A details how the assumption of a standard cumulative Gaussian model (either through fitting or by its assumption within a staircase procedure) can lead to drastic overestimation of threshold σ_{obs}.

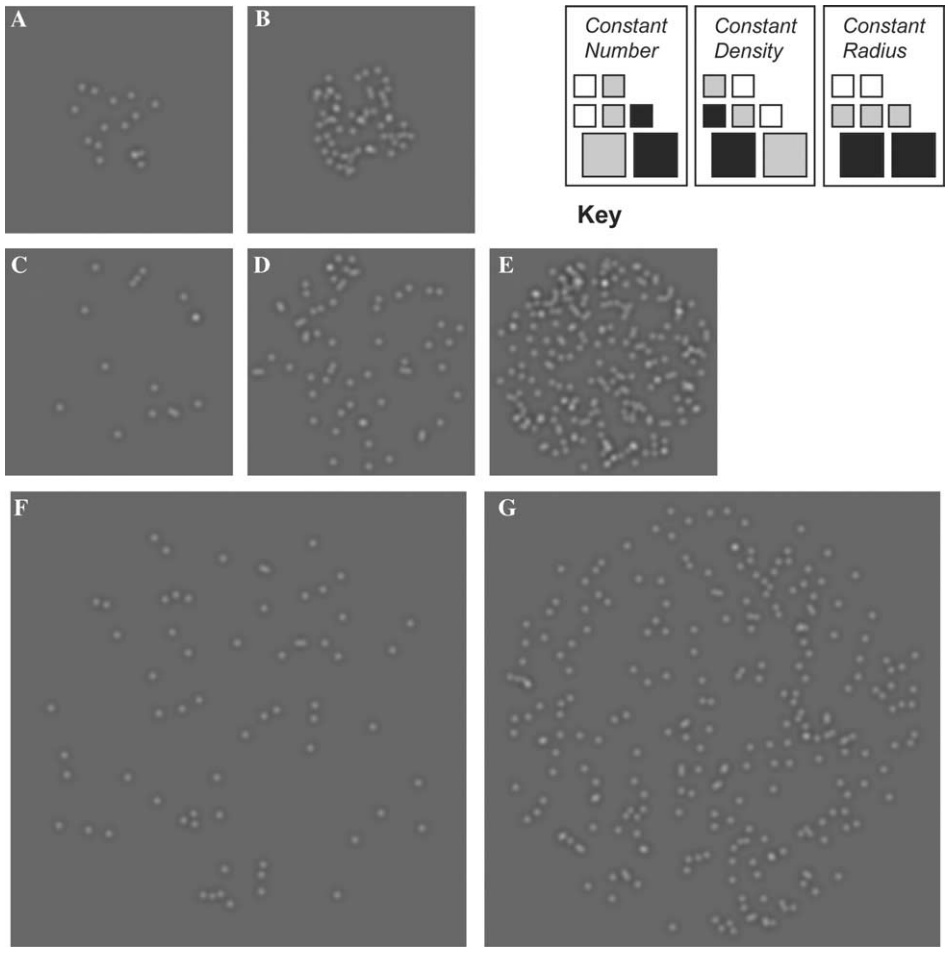

Fig. 2. Single frames from typical stimuli employed in the study. Stimuli contained (A,C) 16, (B,D,F) 64 or (E,G) 256 elements within circular regions of radii (A,B) 2, (C–E) 4 or (F,G) 8 deg.

Appendix B describes a practical procedure for assessing wrapped cumulative normal functions which may be used for fitting psychometric functions. The second problem is that Eq. (1) itself does not take into account wrapping; there is no analytic form of (1) employing circular standard deviation (the standard measure of dispersion for cyclic variables). In practice, this can be overcome with Monte Carlo simulation, as described in Section 2.5

Note that EN uses an un-weighted vector-average to compute global direction which is sub-optimal if the underlying directional p.d.f. is non-cosinusoidal (Salinas & Abbott, 1994) as it was in our experiments. We use this form simply because PVA is the standard comparison observer for this type of task and because we make an explicit comparison of PVA and MLE later in the paper.

2.3. Stimuli

Stimuli were fields of moving, spatially band-pass (Laplacian-of-Gaussian; $\nabla^2 G$) micro-patterns:

$$\nabla^2 G(x,y) = \frac{x^2+y^2-2\sigma^2}{2\pi\sigma^6}\exp\left(-\frac{x^2+y^2}{2\sigma^2}\right), \quad (2)$$

which are spatial-frequency (SF) band-pass with a peak SF (in cycles per image) at

$$f_{\text{peak}} = \frac{1}{\pi\sigma\sqrt{2}}. \quad (3)$$

Our micro-patterns had a peak SF of 1.78 c/deg. Perception of motion in such patterns cannot rely on the low SFs present in random-dot stimuli but requires that local motion signals be integrated over space (Smith, Snowden, & Milne, 1994). Although such patterns are two-dimensional – with complex temporal frequency spectra—their band-pass SF characteristics and fixed jump-size means that there is a unique, optimal motion-energy mechanism sensitive to their motion. $\nabla^2 G$ stimuli are therefore a useful compromise between gratings (SF narrow-band but 1D) and dot patterns (2D but SF broad-band) and furthermore can (like gratings but not dots) be positioned with high (sub-pixel) accuracy (Dakin & Mareschal, 2000).

Movies were 480 ms sequences played at the monitor refresh rate (75 Hz). Each movie frame was repeated three times before being replaced (i.e., each movie frame lasted 40 ms and the sequence contained 12 unique frames). Micro-patterns moved in discrete jumps of 12 arc minutes (i.e., ~35% of one cycle per frame, based on the peak SF of the patch). This corresponds to a continuous velocity of ~5 deg/s. Stimuli were composed of either 16, 64 or 256 elements presented within a circular patch of radius 2.0, 4.0 or 8.0 deg. Sub-pixel accuracy for element positioning was achieved by generating a 10×10 grid of elements with combinations of 0.0–0.9 pixel offsets (in 0.1 pixel steps) in the x- and y-directions. All elements were of "infinite-lifetime" and overlapping elements were added. If an element passed outside of the viewable aperture its position was wrapped to the opposite side of the aperture (i.e., its position was determined by selecting the intersection of a diagonal passing through the point of egress and the centre of the aperture with the opposite side of the aperture). A small amount of positional jitter was added when the dot was reintroduced to minimise the possibility of dots becoming "stuck" when they exited the aperture with a trajectory near-parallel to the aperture edge. This method minimised edge artefacts and maintains an impression of even density over the pattern.

Seven combinations of element-number and patch-radius were tested: 16 elements at 2.0 or 4.0 deg, 64 elements at 2.0, 4.0 or 8.0 deg, and 256 elements at 4.0 or 8.0 deg. This yields a range of conditions with fixed-number (16, 64 or 256 elements), density (0.32, 1.27 or 5.1 elements/deg^2) or radius (2.0, 4.0 or 8.0 deg). Fig. 2 shows single frames from a range of typical stimuli.

Elements moved with directions drawn from a wrapped normal (WN) distribution (a Gaussian wrapped on a circle) defined on the range $\theta \in [0, 2\pi]$ by the p.d.f.

$$f(\theta) = \frac{1}{\sigma\sqrt{2\pi}} \sum_{k=-\infty}^{k=\infty} \exp\left[\frac{-(\theta - \mu - 2\pi k)^2}{2\sigma^2}\right]. \quad (4)$$

We chose this function over alternatives because the convolution of two WN distributions produces another WN distribution (this is not the case for, e.g., a Von Mises distribution). At modest levels of directional variability the variance of the result is equal to the sum of component variances, a property required for EN (but as mentioned above this property breaks down at higher levels of directional variance, a problem we deal with in the Section 2.5). Inasmuch as we are using the WN to represent uncertainty in the cortical representation of direction there is little to chose between models based on various circular Gaussian-like p.d.f.s (Swindale, 1998). Although the WN function does not have a closed form, a simple algorithm for approximating it is given in Appendix B. To compute a WN distributed random variable one simply wraps a standard Gaussian random variable (e.g., within MATLAB: `mod(randn,2*pi)`).

2.4. Procedure

Subjects were presented with a field of moving, band-pass elements and were required to make a judgment of their "overall" direction: either clockwise or anti-clockwise of vertical-upwards motion. They signalled their response using two keys on a computer keypad. Subjects could also press a third key to re-view the stimulus in the case of a lapse in attention but were instructed not to exercise this option when simply uncertain as to the direction (during the experiment, observers exercised this option on less than 1% of trials). Feedback, in the form of an audible beep, was given for incorrect responses.

Within each condition (i.e., for each element-number/patch-radius combination) directions were drawn from WN distributions with 11 different circular standard deviations of either 0.5, 1.0, 2.0, 4.0, 8.0, 16.0, 23.0, 32.0, 45.0, 64.0 or 90 deg. A method of constant stimuli was used to estimate the psychometric function at each SD level. Blocks of 272 trials probed 17 stimulus levels for each condition (i.e., 16 trials per stimulus level). At least two complete runs were undertaken for each subject in all conditions. The 17 baseline stimulus directions tested were -8.0 to $+8.0$ deg in steps of 1.0 deg, where sign indicates whether the stimulus was clockwise or anti-clockwise of vertical-upwards motion. These values were multiplied by a "gain" parameter to provide a stimulus range that adequately bracketed the psychometric function. These gain parameters were determined from short pilot runs conducted with an adaptive method of constant stimuli (APE; Watt & Andrews, 1981) and were chosen to adequately sample the psychometric function for a particular level of directional variance, across the range of all other stimulus parameters to be varied. We then fixed this gain parameter for a particular directional SD across all conditions and subjects, the logic being that this should minimise any discrepancies *between conditions* that could have arisen from differences in the stimulus range tested (e.g., due to under-sampling of psychometric functions). The gains associated with the SDs tested were: 0.25 ($\sigma = 0.5$–2.0 deg), 0.5 ($\sigma = 4.0$ deg), 1.0 ($\sigma = 8.0$ deg), 1.5 ($\sigma = 16.0$–23.0 deg), 2.0 ($\sigma = 32.0$ deg), 3.0 ($\sigma = 45.0$ deg), 4.0 ($\sigma = 64.0$ deg), and 12.0 ($\sigma = 90$ deg). These values provided a good bracketing of the psychometric functions across conditions. We used a method of constant stimuli because we observed in pilot runs that, at higher levels of directional variance, psychometric functions often did not reach 0% and 100%, which renders many adaptive methods unstable.

Raw data were fit using a bootstrap incorporating a cumulative WG model (allowing for *periodicity* in

psychometric functions; details in Appendices A and B). Data plotted are the bootstrapped fit of data pooled across runs for that condition/subject; error bars are the 95% confidence intervals on the fits.

2.5. Model fitting procedure

Fits of the EN model to our psychophysical data were made using the bootstrapped estimates of threshold derived using the procedure just described. Specifically, a dataset was constructed by selecting randomly from the bootstrapped data at each SD level, this set was then fit using a Nelder-Mead/simplex minimisation method (the `fminsearch` function in MATLAB), and the best fitting parameters (σ_{int} and n_{samp}) were recorded. This yielded distributions of the two parameters of the fit for a given data set. This fitting procedure is repeated 1000 times on different datasets derived from the bootstrap. Finally, the peaks of the two parameter-histograms are used to give the best parameters and the spread is used to derive the 95% confidence intervals associated with each parameter.

The model used a least-squares distance metric computed between log-transformed data and log-transformed model, weighted by the logarithm of the confidence intervals on the threshold estimate. The justification for computing goodness-of-fit on log-transformed data is that our estimated confidence intervals were broadly similar across directional SDs on log-axes so that this procedure approximately equally weighted the contribution of each data point to the fit. Using this procedure, subjects' performance will systematically deviate from EN at high levels of directional variability when additivity of variance breaks down due to wrapping (for details see Appendix C). In practice, this is apparent only for the highest directional-variability conditions tested ($\sigma = 64$, 90 deg). We dealt with this problem using a Monte Carlo simulation; model parameters were estimated as described and were then used to determine the global-sampling and local-precision of a simulated observer. Specifically, a simulated threshold was estimated using longer runs (typically at least 15,000 trials) of the same psychophysical procedure (method of constant stimuli), during which directional distributions with various directional offsets relative to vertical were generated. From trial-to-trial the ideal observer estimated mean over all directions ($\bar{\theta}$) using the population vector-average

$$\bar{\theta} = \begin{cases} \tan^{-1}(\bar{S}/\bar{C}), & \bar{S} > 0, \bar{C} > 0, \\ \tan^{-1}(\bar{S}/\bar{C}) + \pi, & \bar{C} < 0, \\ \tan^{-1}(\bar{S}/\bar{C}) + 2\pi, & \bar{S} < 0, \bar{C} > 0, \end{cases} \tag{5}$$

where

$$\bar{C} = \sum_{i=1}^{n} \cos\theta_i, \quad \bar{S} = \sum_{i=1}^{n} \sin\theta_i.$$

(This computation is achieved in one step in MATLAB using a $\texttt{atan2}(\bar{S}, \bar{C})$).

The resulting direction was used to generate a binary response (clockwise or anti-clockwise of vertical), and results were pooled and fit with a psychometric function (using the procedure described in the preceding section). This procedure was repeated at the directional SDs we tested. Because Eq. (5) takes into account wrapping so does the simulated observer. The predictions from this version of EN provide a consistently better fit to the subjects' data than the analytic predictions of EN (on average a 3% reduction in R^2) and are used in all subsequent figures showing EN predictions.

2.6. Apparatus

Stimuli were generated with an Apple Macintosh G4 computer running MATLAB (MathWorks). The programs controlling the experiment incorporated elements of the PsychToolbox (Brainard, 1997). Stimuli were presented on a La Cie 22″ CRT monitor fitted with a video attenuator (Pelli & Zhang, 1991). The attenuated signal was amplified and copied (using a line-splitter) to the three guns of the monitor to generate a monochrome image. The display was calibrated using code from the VideoToolbox (Pelli, 1997) and a Minolta LS 110 photometer, and then linearised using look-up tables in software (to give pseudo 12-bit contrast accuracy). The monitor operated at a resolution of 1024×768 pixels (24 pixels/cm) with a vertical blanking rate of 75 Hz. It was viewed binocularly at a distance of 57 cm and had a mean background luminance of 50 cd/m^2.

2.7. Observers

The three authors (wearing optical correction as necessary) served as observers. All are experienced at psychophysical tasks involving motion perception and conducted practice runs on the task until their performance reached asymptotic levels.

3. Experiment: Effect of density, number and region-size on integration

The goal of this experiment was to determine which spatial attributes of a moving stimulus determine local and global limits on direction integration. A number of papers have reported that the density of a stimulus has only a modest effect on direction discrimination (Barlow & Tripathy, 1997; Scase, Braddick, & Raymond, 1996; Watamaniuk, 1993; Williams & Sekuler, 1984). Watamaniuk (1993), in particular reported that fine direction discrimination improved with increasing exposure duration but did not improve greatly with increasing size of stimulus nor with increasing number

of elements. Watamaniuk (1993) used a d' based methodology for assessing integration. That is, he measured the percent correct discrimination performance for a fixed offset (2 deg) between the average direction of his target patterns and the reference direction. This then allows one to use the ideal observer

$$d' = \frac{\sqrt{n}\Delta\theta}{\sigma},$$

where $\Delta\theta$ is the difference in direction between reference and target, σ is the directional standard deviation, and n is the number of directional vectors being used for the judgment. This model is free from internal noise and as such cannot, like the standard motion coherence paradigm, evaluate how much d', is influenced by local or global noise. Watamaniuk (1993) asserts that a local noise parameter incorporated into the modelling procedure was close to unity for all conditions, i.e., that "it did not capture any significant loss of direction information beyond that captured by S_1" (a global noise parameter). We found this surprising especially given that other authors have found that local noise appears to have a central role in determining performance on motion coherence tasks. Barlow and Tripathy (1997), for example, performed an extensive study of how density, size, and area affect motion integration using motion coherence thresholds. They report that changes in dot density produce only a modest reduction in motion coherence thresholds but that these changes were entirely attributable to "correspondence noise" (mismatches of elements between movie frames). We would have expected such correspondence noise to be present in the stimuli employed by Watamaniuk (1993) and consequently for there to be a role for local noise in accounting for his findings. One possible reason for this discrepancy is that Watamaniuk (1993) used stimuli containing a large number of elements (>300) and it may be that local noise effects saturate at a relatively small number of elements. Our paradigm allowed us to address this issue. We attempted to minimise the role of local noise by employing infinite-lifetime elements with relatively small jump-sizes. A common criticism of such stimuli is that they allow direction judgements to be based on the motion of a single element. If observers had relied on this strategy it would have been reflected in very low estimates of their global sampling (which it was not).

When stimuli become large, eccentricity may affect performance. Unlike others (e.g., Burr, Morrone, & Vaina, 1998) we chose not to scale our stimuli for cortical magnification. We did this for two reasons. First, it is unclear if size-scaled stimuli should also be velocity-scaled. If so this assumes broad velocity-tuning for global direction mechanism. Second, our efforts to produce low local noise produced highly visible stimuli, so that at the maximum eccentricity tested (i.e., 8 deg) discrimination thresholds are barely affected (which we have confirmed using eccentric fixation; Dakin, Mareschal, & Bex, 2004).

3.1. Results

Fig. 3 shows three subjects' direction discrimination performance, plotted as a function of directional variability, for various patch-sizes and numbers of elements. Lines show the fit of the EN model to each condition; the two numbers in parentheses within the legends are the (local) additive noise and (global) sampling parameters, respectively, for the corresponding fit. Note from Fig. 3 that estimated sample size always exceeds one, confirming that averaging over multiple elements is occurring, i.e., that observers were not tracking single elements. Results also support Watamaniuk's (1993) report of high efficiencies (average 35%) for a similar direction discrimination task using broad-band random-dot kinematograms. Our findings also confirm Smith et al. (1994), who showed that global motion perception in band-pass filtered patterns cannot rely wholly on low SFs.

Fig. 3 illustrates the main finding of this experiment; data separate under conditions that maintained either a constant patch radius or a constant density, but superimpose when density and patch radius were co-varied so as to maintain a constant number of elements. Thus it is wholly the number of elements in the display that determines global direction discrimination. Note from the parameters in Fig. 3 (top section) that although the absolute number of global samples increases with increasing number of elements, *local noise also increases*. This illustrates the power of EN to separate out the opposing influence of these two factors. That subjects' thresholds tend to fall as the number of elements increases, is entirely attributable to global pooling, but the pattern of results indicate that the improvement is offset by the *increased* local noise. That is not to say that spatial arrangement of elements is never important: elements arranged in trajectories can be easier to detect (Verghese, Watamaniuk, McKee, & Grzywacz (1999) but see Bex, Simmers, & Dakin (2003)) and elements can be positioned so that their local motions cancel (Qian, Andersen, & Adelson, 1994).

This influence of density, number, and region-size on direction discrimination is summarised in Fig. 4 which plots local-noise and global-pooling parameters for all conditions tested. Notice the increase in local noise (Figs. 4A–C) as the number of elements increases from 16 to 64 but that this increase plateaus (particularly in Fig. 4B) for increases from 64 to 256 elements. This suggests that the influence of local noise may saturate at a relatively small number of elements and may explain why Watamaniuk (1993) did not require local noise to model his dot-density data (since his stimuli contained

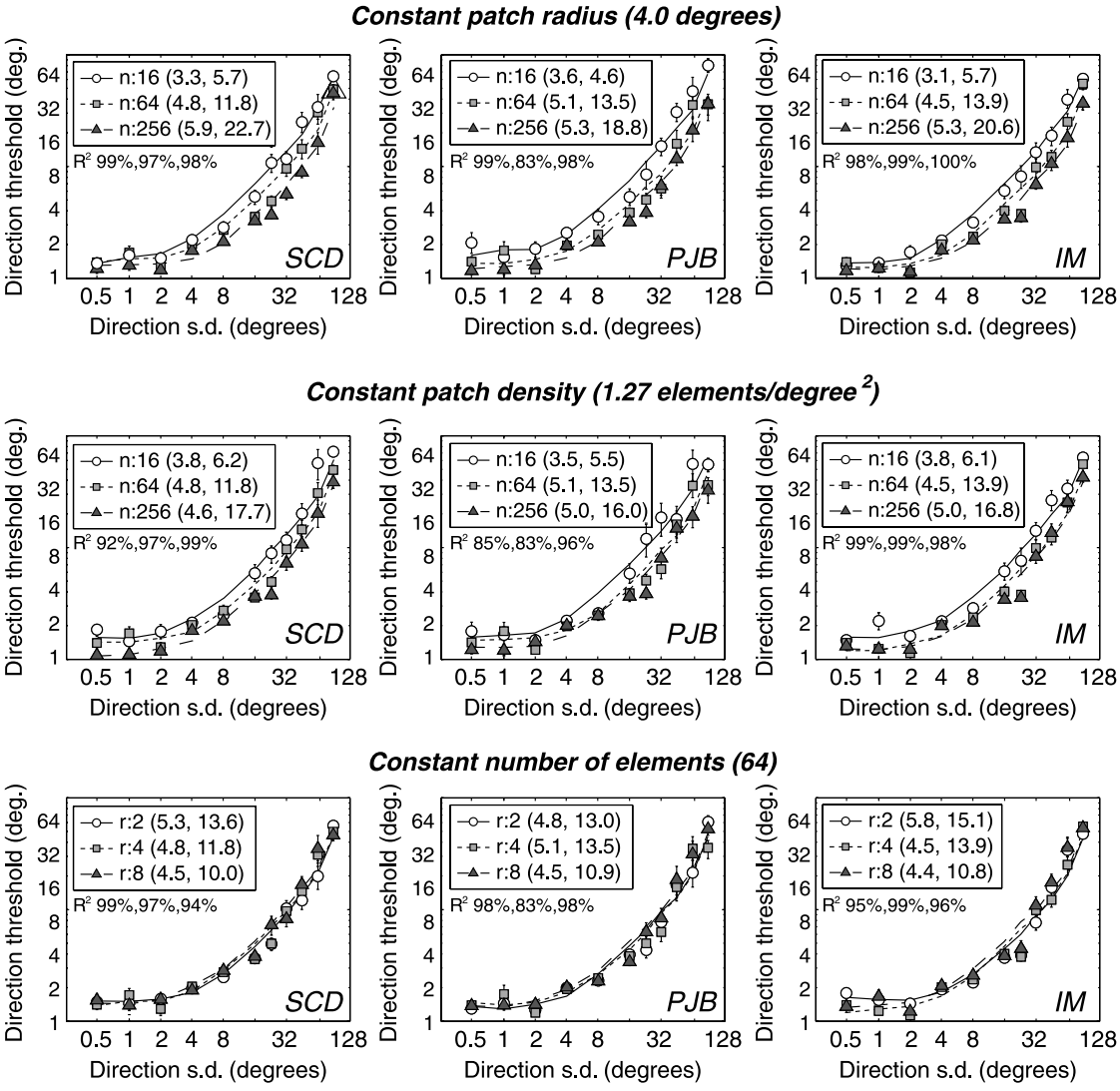

Fig. 3. The directional offset necessary for observers to reliably discriminate the mean direction of a field of moving band-pass elements from vertical, as a function of the range of directions present (the directional standard deviation). Results are shown from three subjects in three conditions (fixed patch-size, patch-density or number of elements). Data were fit with the EN model. Within the boxed legends, "n" refers to the number of elements in the display, "r" to the patch radius (in degrees), and the two number in parentheses give the values of σ_{int} and n_{samp}, respectively. R^2 goodness-of-fit parameters—to the nearest %—are listed for each of the three fits below each legend. Notice that fits are similar when the number of elements is fixed.

>300 elements). We suppose, like Barlow and Tripathy (1997), that this pattern of results reflects intrinsic properties of local motion detectors (i.e., their bandwidth, arrangement across the visual field, etc.).

Global sampling (Figs. 4D–F) shows an even clearer dependence on the number of elements present. Lack of overlap between the error bars indicates a statistically significant difference in a parameter value. Because our results are straight lines on log-linear axes we performed a weighted least-squares fit of these data with a model of the form $y = m\log(x) + c$. The parameters derived for the global sampling estimates (from SCD, PJB, and IM, respectively) were $m = 4.9$, 4.6, 4.6, and $c = 7.9$, 7.7, and 6.9 with corresponding R^2 of 90.4%, 93.4%, and 87.4%. These results are consistent with an earlier study on orientation integration (Dakin, 2001).

It is also informative to compare these findings with an existing statistical model of motion integration. Fig. 4G plots direction discrimination thresholds averaged across subject and across conditions with similar numbers of elements. The fits to these data were generated using a three-parameter version of a pooling model due to Watamaniuk (1993), shown in the boxed caption within Fig. 4G. We generated predicted thresholds from this model by estimating sensitivity d' for a range of directional offsets and then fitting a psychometric function to give the directional offset yielding $d' = 1.35$ (equivalent to an 83% correct threshold). The parameter n in the boxed equation in Fig. 6G refers to the number of moving elements (set according to the condition) and the remaining parameters were determined from a three-parameter least-squares fit to the log-transformed data

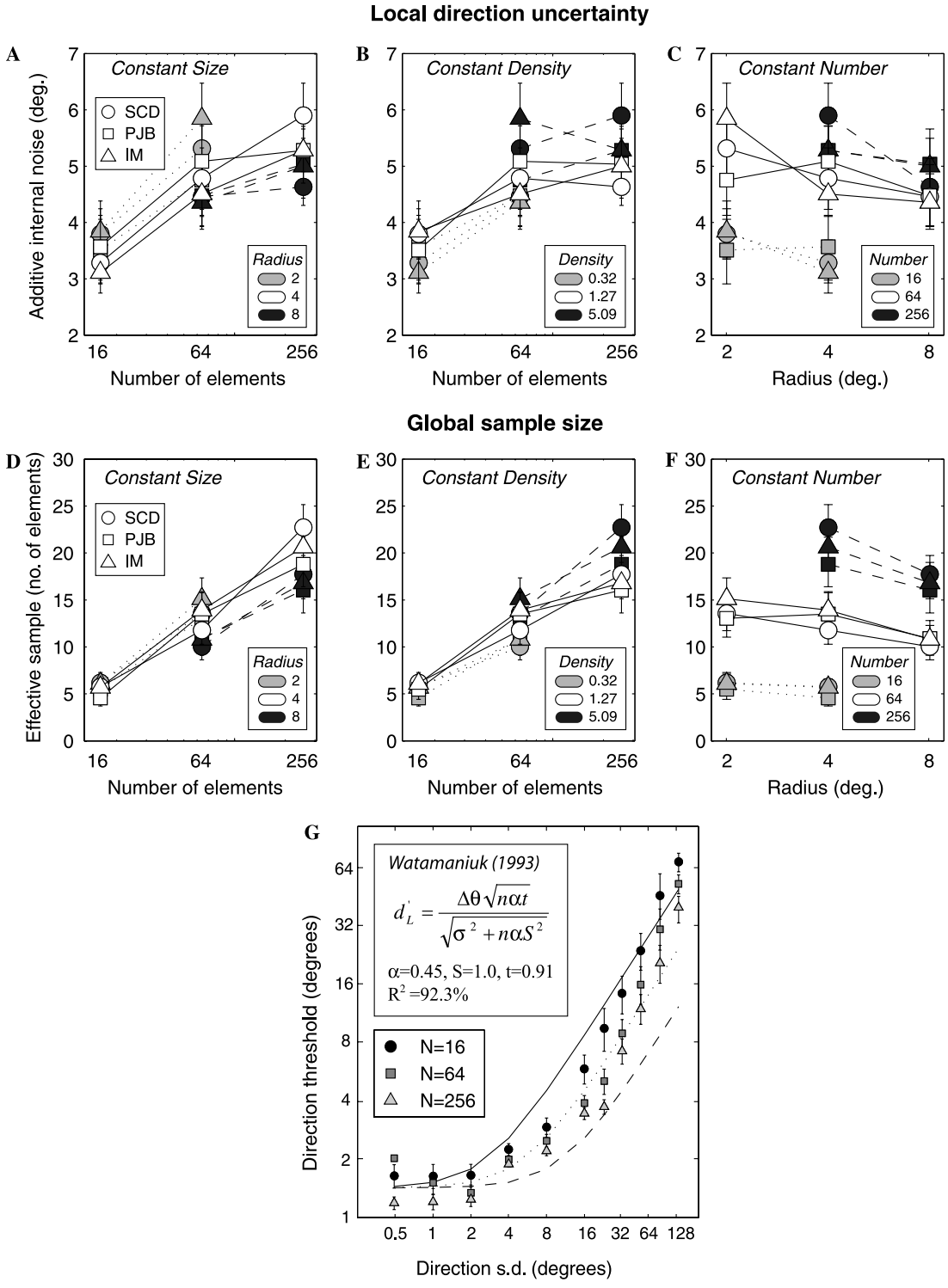

Fig. 4. (A–F) Summary plot of parameters derived from the EN fits. (A–C) Plots of σ_{int} and (D–F) n_{samp} derived from the seven conditions tested. Within each plot symbol shape codes the subject and symbol shade codes condition. Error bars show the 95% confidence intervals on the parameter estimate. The primary determinant of both global sampling and local noise is clearly the number of elements present in the display, irrespective of their arrangement. (G) Symbols show direction discrimination thresholds pooled across conditions with equal numbers of elements and across subjects, along with predictions from the motion integration model of Watamaniuk (1993).

(inversely weighted by the associated confidence intervals). Although capturing the general trend of convergence of thresholds at low directional variability, and divergence at high variability, this model fails to capture the exact dependence of threshold on directional variability.

4. Noise limited model of local-to-global direction encoding

4.1. Overview

In this section, we consider how neural activity constrains the computation of global direction. Specifically we compare our psychophysical results to the performance of two models of local-to-global direction encoding—a population vector average, and a maximum likelihood estimator—operating under conditions of Poisson noise.

The logic of the approach is illustrated in Fig. 5A. We hypothesise that there are local and global stages of motion pooling corresponding to the operation of neurons in areas V1 and MT, respectively. Local motion encoding consists of a series of local motion *sensors* (each organised around a single V1 hyper-column) comprised of a series of *sub-unit*s (V1 neurons, located at or near the same retinotopic location, direction sensitivity spanning 360 deg). Each local sensor computes a single estimate of local direction by integrating the responses from its sub-units using either PVA or MLE. We assume this process can be achieved via divisive inhibitory interactions between neurons (Deneve et al., 1999). These local direction estimates drive a second stage of motion processing where a single global sensor receives input from a series of sub-units (MT neurons, direction sensitivity spanning 360 deg). Global sub-units have large receptive fields so that each responds to information (within its

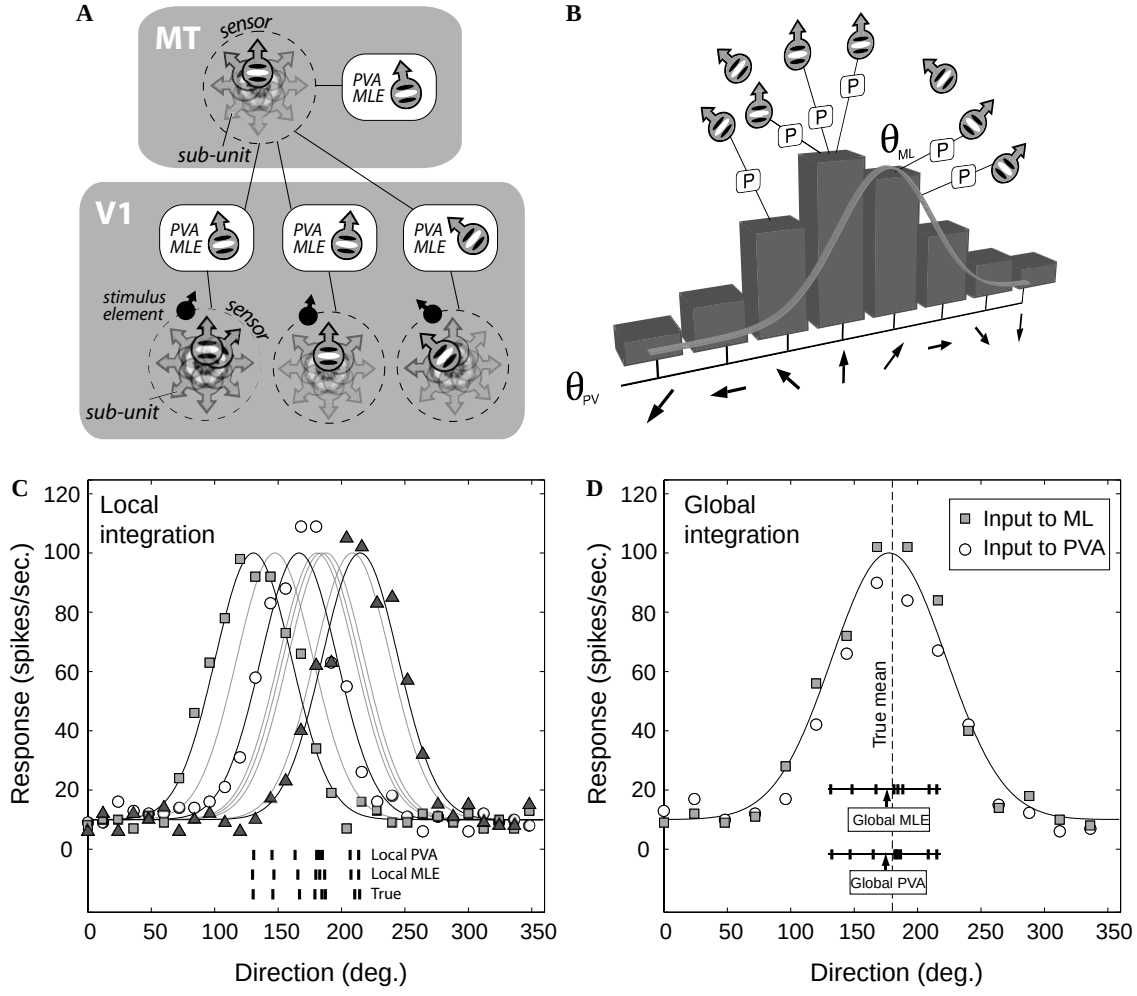

Fig. 5. Local-to-global motion integration model. (A) We suppose that V1 neurons interact to derive a series of unitary *local direction estimates* that are pooled by MT neurons over a wider spatial extent in order to estimate *global direction*. (B) A direction histogram (shaded blocks) is constructed from a set of motion sub-units whose responses are corrupted by noise (*P*). The PVA direction (θ_{PVA}) is the result of averaging each detector's preferred direction in proportion to the neural response associated with it. MLE fits a wrapped Gaussian template to the histogram; the location of the best-fitting template (θ_{MLE}) codes direction. (C) Illustration of the local response to eight elements. Each of the three datasets shows the response of a local motion sensor (comprising of 30 sub-units) to a single moving element. Curves show MLE of the best-matching templates to all eight sets of responses. Local PVA estimates are the average of the sub-unit directions weighted by each ones response. Inset in the lower part of the figure are the true directions along with the estimates derived using PVA and MLE. (D) Response of the global sensor (with 15 sub-units) to the eight local direction estimates from either local PVA (circles) or MLE (squares). The final estimates of global PVA and MLE are shown on the inset.

directional pass-band) generated by many local direction estimates. The global sensor finally computes a single estimate of global direction by integrating responses from global sub-units using either PVA or MLE.

The difference between PVA or MLE estimates of direction is illustrated in Fig. 5B. PVA averages the preferred direction of each sensor weighted by the sensor response. As such the system knows nothing about the nature of the random direction variability (Gaussian in the figure) or of intrinsic variability in response magnitude (Poisson in the figure). The MLE by contrast performs a least-squares fit of the sensor responses to a template whose shape matches the underlying p.d.f. for the direction variability, and is scaled/offset to allow for the known response range/spontaneous activity. The location of the best-fitting template is the MLE estimate of direction.

Note that all models operate on an abstract representation of the stimulus motion vectors. By doing this, and by not constraining the size of MT receptive fields, we are effectively making global direction estimation independent of spatial sampling (in line with our psychophysical findings). We consider the effects of spatial arrangement on local motion estimation (using motion energy) in Section 5.

4.2. Model details

4.2.1. Tuning and response properties of sub-units

We assume that the directional tuning functions of local and global sub-units (notionally, V1 and MT neurons, respectively)—that is the mean number of spikes produced by a neuron as a function of the stimulus direction (θ)—are captured by a wrapped-Gaussian function of the form

$$Q(\theta,\mu,\beta)=\sum_{k=-\infty}^{k=\infty}\exp\left[\frac{-(\theta-\mu-2\pi k)^2}{2\beta^2}\right], \tag{6}$$

where Q refers to the responsivity of a sub-unit (in spikes/s), μ to its preferred direction, and β to its bandwidth (circular SD). Note that we are assuming that one local sensor sees only one moving element, a point we return to Section 5. Additionally, because β was fixed for a simulation of the full dataset (i.e., across directional SD), we are implicitly assuming that directional tuning (local and global) remains invariant with the signal bandwidth. We justify this assumption based on the observation that MT neurons have broad (average width ~90 deg)' directional tuning (Albright, 1984; Britten & Newsome, 1998; Snowden, Treue, & Andersen, 1992) which varies little with stimulus coherence (Britten & Newsome, 1998), a result mirrored by recent psychophysical findings (Fine, Anderson, Boynton, & Dobkins, 2004).

The number of sub-units comprising each local or global sensor is referred to as N_{local} and N_{global}, respectively. We assume that sub-units at both integration stages are sampled evenly around the clock, based on the smooth progression of orientation/direction preference observed in primary visual cortex (e.g., Hubel & Wiesel, 1962). Thus, we define a set of preferred directions for the local sub-units as

$$\mu_{\text{local}}=\frac{2\pi}{N_{\text{local}}}\{1,2,3,\ldots,N_{\text{local}}\}. \tag{7}$$

The response of the nth local sub-unit to direction, θ, is then

$$r_n(\theta)=\kappa+(F-\kappa)Q(\theta,\mu_{\text{local}}(n),\beta_{\text{local}}), \tag{8}$$

where F is the maximum firing rate and κ the spontaneous activity of the sub-unit. We assume that sub-unit noise at all stages is independent and Poisson distributed (i.e., response variance is equal to response) so that the probability of local sub-unit number n producing k spikes in response to a certain direction (θ) is

$$P(k\mid\theta,n)=\frac{|r_n(\theta)|^k\exp(-r_n(\theta))}{k!}. \tag{9}$$

Electrophysiological evidence supports the idea that the mean activity of cortical neurons (showing low spontaneous activity) in V1 (Dean, 1981; Softky & Koch, 1992; Tolhurst, Movshon, & Dean, 1983; Vogels, Spileers, & Orban, 1989) and MT (Britten, Shadlen, Newsome, & Movshon, 1993) approximately equals their response variance.

4.2.2. Computing local direction

The ith element from the stimulus, moving at direction θ_i, generates N_{local} noisy responses from the local sub-units (call this $s_n(\theta_i)$). We compare two ways of estimating local direction from a series of such responses. The estimated PVA direction of the ith element—call it p_i—is defined by Eq. (5) taking these values as input:

$$\bar{C}_i=\sum_{n=1}^{N_{\text{local}}}s_n(\theta_i)\cos\mu_{\text{local}}(n)\quad\text{and}$$

$$\bar{S}_i=\sum_{n=1}^{N_{\text{local}}}s_n(\theta_i)\sin\mu_{\text{local}}(n). \tag{10}$$

Another way of estimating the local direction of the ith element is to use MLE:

$$m_i=\arg\max_{\theta}P(s|\theta).$$

In practice evaluating this expression amounts to template matching; one locates the minimum distance from the observed responses to a template (in this case the underlying channel response function $Q(\theta,\mu,\beta)$). The distance metric employed depends on the distribution of the noise. The MLE of direction for a response set corrupted by Poisson noise is

$$m_i=\arg\min_{\theta}\sum_{n=1}^{N_{\text{local}}}s_n(\theta_i)\log(r_n(\theta_i)). \tag{11}$$

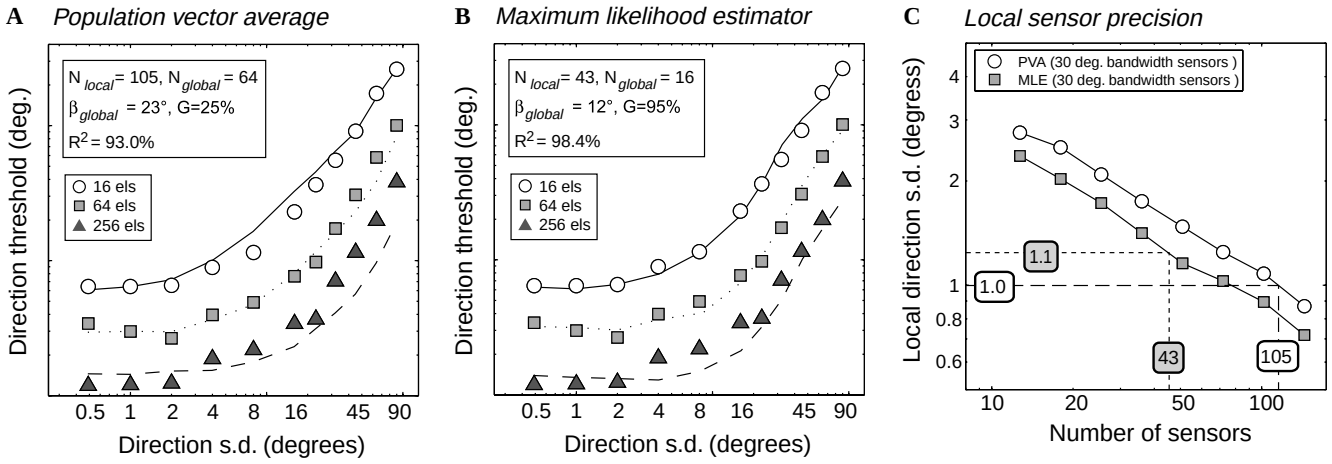

Fig. 6. Symbols show discrimination thresholds (averaged across three subjects) for judging the direction of 16, 64 or 256 elements moving within a 4 deg radius region as a function of directional standard deviation. For the sake of clarity, datasets have been offset from one another and error bars omitted. We assume that this task is mediated by a series of MT neurons each averaging a series of noisy inputs from V1/local motion detectors. (A) Lines show predictions of a model that derives overall direction by vector averaging all inputs. The parameters (inset top left) indicate the number of local and global sub-units (with directions evenly distributed around the clock), the bandwidth of the global sub-units and the proportion of the local motion signals seen by the global integrator. (B) A similar model that computes global direction using maximum likelihood estimation; i.e., it fits a wrapped Gaussian template to the histogram of observed V1 responses. This MLE model produces a generally better fit. (C) Given the parameters derived from the MLE and PVA models one can compute the effective precision of each estimate of local direction given the local sampling density (N_{local}); notice how many more local samples are required to give near-equivalent precision from local PVA compared to MLE.

The MLE estimate for the ith stimulus element was evaluated to the nearest 0.1 deg using a minimisation procedure. Note that the bandwidth of template used for local MLE matches the bandwidth of the directional tuning of the local sub-units; this is the optimal strategy.

We now have local direction estimates for each of the elements of the stimulus. The precision of this set of estimates is determined entirely by two parameters: N_{local} and β_{local} but for fixed β_{local} it depends entirely on N_{local}. This is illustrated in Fig. 6C, which shows the precision of each local sensor (i.e., the circular SD of its direction estimate) as a function of N_{local}, for the PVA and MLE schemes. For a 30 deg bandwidth sub-unit, MLE is approximately 20% more precise than PVA, for the sampling rates tested.

4.2.3. Computing global direction

For a set of local directions estimated using MLE or PVA, we now consider the computation of overall direction. The first issue is how many independent samples of local direction should be employed? We assume that only relevant directional signals feed up to the next level: i.e., the responses of local motion sensors that are not being stimulated are discounted. We then use a fixed proportion of the number of elements present (P)—to give global sample size of G—and set P as a free parameter. In fitting the model we constrain P so that the resulting G is less than or equal to the number of elements present in the pattern. Thus the MLE and PVA schemes generate sets of local direction estimates comprising G samples each; we refer to this set of local estimates as θ_{local}.

Similar to the local motion stage (Eq. (7)) we define μ_{global}, a set of N_{global} preferred directions for the global sub-units. We can then define $R_n(\theta)$, a measure of global sub-unit response to the local direction estimates:

$$R_n = \kappa + \frac{(F-\kappa)}{G}\sum_{i=1}^{G} Q(\theta_{\text{local}}(i), \mu_{\text{global}}(n), \beta_{\text{global}}). \quad (12)$$

Note that the response of the sub-unit is re-scaled by the size of the global sample (G); we support this assumption by the finding that the response of MT neurons saturates quickly as the number of elements moving in its preferred direction increases (Snowden et al., 1991). Using Eq. (12), the response is corrupted by Poisson noise (Eq. (9)) to arrive at a final measure of global sub-unit response $S_n(\theta)$. We can then compute an estimate of global PVA direction by plugging the following values into Eq. (5):

$$\bar{C} = \sum_{n=1}^{N_{\text{global}}} S_n \cos\mu_{\text{global}}(n) \quad \text{and}$$
$$\bar{S} = \sum_{n=1}^{N_{\text{global}}} S_n \sin\mu_{\text{global}}(n). \quad (13)$$

Similarly, the MLE of global direction can be derived by adapting Eq. (11):

$$\theta_{\text{MLE}} = \arg\min_{\theta} \sum_{n=1}^{N_{\text{global}}} S_n \log(R_n(\theta)). \quad (14)$$

We also evaluated the performance of a model that uses the peak of the directional histogram to code direction but thresholds derived using this procedure were

uniformly much worse than human thresholds (typically several orders of magnitude) and for this reason are not reported.

Figs. 5C and D illustrate the operation of the local and global stages of the model given eight moving elements (with a mean direction of 180 deg and a circular SD of 32 deg). Each set of symbols in Fig. 5C shows the response (in spikes/s) to a single moving element of one set of local sub-units (three sensors are shown). The MLE for direction is the location of the best fitting template to the response (shown by solid lines), the PVA is the vector average of sub-unit preferred directions (i.e., the x-location of each symbol) weighted by the sub-unit response (i.e., the y-location of each symbol). The inset at the bottom shows the true directions of the eight moving elements along with PVA and MLE estimates of local direction. These estimates feed into a second stage of pooling, (Fig. 5D). Global sub-units at this stage (15 are shown in the example) have large receptive fields and so base their response on many (in this case, all) of the local direction estimates. This generates the pattern of response shown according to whether the inputs came from the PVA or MLE local estimates. Now, a single estimate of global direction (indicated by the arrows in the inset) can be derived from the global sub-unit responses.

4.3. Summary of model parameters

- F: Peak neural response: 100 spikes/s (corrupted by Poisson response noise).
- κ: Spontaneous activity: 10% peak neural response
- β_{local}: Directional bandwidth of local sub-units (SD of WN response function). Fixed at 30 deg (Based on typical V1 tuning; Albright, 1984),
- β_{global}: Directional bandwidth of global sub-units. **Free parameter**. For MLE, β_{global} is the template SD.
- N_{local}: Number of sub-units comprising local sensor. **Free parameter**.
- N_{global}: Number of sub-units comprising global sensor. **Free parameter**.
- μ_{local} and μ_{global}: The preferred directions of local and global sub-units (determined by N_{local} and N_{global}).
- P: Proportion of local elements contributing to the global direction estimate. **Free parameter**.
- G: Global sample size employed (set by P and the number of elements in the stimulus).

4.4. Simulation methods

The predictions of all models under comparison were generated using Monte Carlo simulations of the experiments. Simulations were conducted in the same way as described above in Section 2.5. The Monte Carlo procedure was repeated at various directional SDs and for the three numbers of elements tested (16, 64 or 256). Note that all models operate without inclusion of a spatial component (reflecting the lack of effect of arrangement in our results). The range of offsets tested for a given directional SDs was identical to the comparable condition with the human observers; i.e., the stimulus-gains associated with the directional SDs tested were similar.

4.5. Simulation results

Figs. 6A and B show the fits of the PVA and MLE models, respectively, with datasets vertically displaced from one another for clarity. Symbols show psychophysical data, pooled across three subjects, for conditions involving the presentation of 16, 64 or 256 elements within a 4 deg region. The PVA model produces a good account of the results ($R^2 = 93.0\%$) with four free parameters. We note, however, that this model does not capture subjects' deteriorating performance at high levels of directional noise for high density patterns. The MLE model on the other hand can capture all of our data ($R^2 = 98.4\%$) with the same number of parameters. The predictions shown were generated using a bandwidth (β_{global}) of 12 deg. For broader bandwidths the MLE model produces performance identical to the PVA model and indeed it has been shown that PVA is optimal and equivalent to MLE under conditions of Poisson noise and broad template bandwidths (Snippe, 1999).

Fig. 6C converts these estimates of local sub-element spacing into estimates of local directional precision. Notice that although more than twice as many PVA elements are required for the fits shown, there is only a modest (10%) increase in equivalent local precision over the MLE model.

5. Discussion

5.1. Summary

We have described how an equivalent noise paradigm can separate the influence of local and global constraints on observers' ability to estimate the overall direction of a moving image. We have also shown that it is the number of elements present, regardless of their density or the size of the region they occupy, that limits both local noise and global sampling. As the number of elements increases, thresholds drop in manner that is consistent with an increase in the global pool of directional samples but also an increase in local noise (that is presumably due to raised correspondence noise). These findings can be modelled by a labelled-line system that integrates the responses of a bank of directionally-tuned mechanisms each of whose response is corrupted by Poisson noise. In this section, we consider the implications of

these findings for how we estimate motion integration psychophysically and how we generate computational models of both local and global motion extraction.

5.2. Relationship to motion coherence

The sampling efficiency parameter of EN allows us to predict motion coherence thresholds, assuming that direction discrimination in these stimuli is based on a coarse +/−180 deg direction discrimination. In the course of conducting these experiments we noted that the sampling efficiencies derived from EN apparently predicted better performance on motion coherence than values reported in the literature. To effect a direct comparison, we collected motion coherence thresholds (up-versus-down) using standard techniques and similar stimuli to those employed in Experiment 1—a 4 deg radius patch, with same dot-speed, dot-density, stimulus lifetime, etc. We used an adaptive staircase procedure (QUEST; Watson & Pelli, 1983) to estimate motion coherence thresholds (defined as the maximum proportion of signal elements that could be replaced with randomly moving elements while supporting direction discrimination at the 83% level). Each threshold estimate was based on the mean of four runs of 45 trials each.

Results from two subjects are plotted in Fig. 7, as a percentage of the total elements present. First, note that motion coherence thresholds are relatively invariant of the element density or stimulus area as reported for fine direction discrimination in the EN experiments; it is primarily the number of elements that determines performance. There is a modest reduction in coherence thresholds as the number of elements increases within a fixed patch radius (in accord with Barlow & Tripathy, 1997).

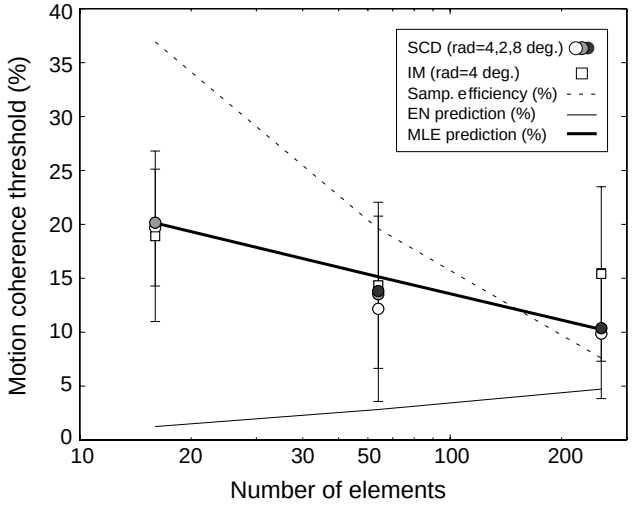

Fig. 7. Motion coherence thresholds plotted as a function of the number of elements in the display. The dashed line shows the number of pooled direction signals (derived from EN) expressed as a percentage of the dots present. The thin solid line shows the predicted performance on the motion coherence task if observers based their judgments on the average direction (with local and global noise constraints set by the equivalent noise parameters derived from Experiment 1). Observers are consistently worse at motion coherence tasks than would be predicted by EN. The thick solid line shows estimates of motion coherence thresholds by the MLE model with no change in the parameters derived from the EN task.

We first ask: how well can the global sampling efficiencies we measured in the main EN experiment predict such thresholds? Taking the average number of samples employed by subjects in comparable conditions (in terms of density and number of elements displayed) in the equivalent noise experiment (i.e., the values given in Fig. 3), and dividing by the total number of elements, we get the predicted motion coherence thresholds shown by the dashed line. Plotted this way, sampling shows a marked decline as a function of the number of elements, and this is consistent with other estimates of sampling as the number of elements increases (Watamaniuk, 1993). We also used a Monte Carlo simulation, employing both local *and* global EN parameters, to predict motion coherence thresholds. For a single trial, we took the directions contained in a typical motion coherence stimulus, corrupted them by additive noise (equal to the average local noise estimate from the comparable EN condition), took a sample of those directions (equal to the average sample size estimate from the comparable EN condition), and used the PVA of these values to make an "up-versus-down" decision. By running many trials over a range of coherence levels we generated the performance shown by the thin solid line in Fig. 7, which predicts lower motion coherence thresholds than are observed experimentally. In short, observers are not as good at motion coherence tasks as they should be based on estimate of local noise and global pooling derived from EN.

Finally, we generated simulated motion coherence thresholds using the PVA and MLE models derived in the last section (using identical parameters to those derived from the EN study). To do this, we generated a series of stimuli over a range of motion coherences, and measured percent correct performance on the basis that "up-versus-down" decisions were based on the direction predicted from the PVA or MLE of the directions present in the display. The PVA model failed to achieve human levels of performance; it predicted thresholds on the order of 80%. The reason that PVA performed so badly is largely attributable to the extremely low values of parameter P (the proportion of local estimates seen by the global stage) that were required for the EN experiment. At large values of P, PVA is highly resistant to local noise and grossly overestimates human performance. By contrast, and with no modification, the MLE model produced a close fit to the psychophysical data (thick line, Fig. 7). The success of the MLE model for fitting both motion coherence and equivalent noise data provides strong evidence that

the computation of global motion uses a large proportion of the number of elements present, but only a limited range of directions can be integrated (consistent with MLE, but not PVA, parameters).

5.3. Extraction of local direction

Fig. 4C indicates that there is a modest increase in internal noise as the stimulus area decreases. We speculate that the observation that noise increases as the inter-element spacing declines represents an increase in the probability of false matches between elements on successive animation frames, known as "correspondence noise". Our modelling so far has not included a parameter for correspondence noise; in this section we examine how the responses of local motion energy mechanisms are affected by such noise. Our use of spatially band-pass elements with a fixed jump-size constrains the type of mechanism that could be used to extract local direction information from the stimulus. Although the temporal spectrum of our stimuli is relatively broad there exists an optimal motion energy mechanism tuned to the spatial frequency and jump-size used.

The top left section of Fig. 8 shows the sum of all frames from a typical stimulus from our experiment; inset and framed-red in the top right of this panel is a single frame from the optimal motion energy mechanism for this stimulus. Because motion energy involves the comparison of filters in quadrature phase, the optimal stimulus for a motion energy mechanism has a jump-size equal to one-quarter of the peak spatial frequency of the stimulus (λ). Our stimuli had a jump size closer to 0.4λ and, as a consequence, there is substantial energy in the opposite direction. This can be seen in the remaining

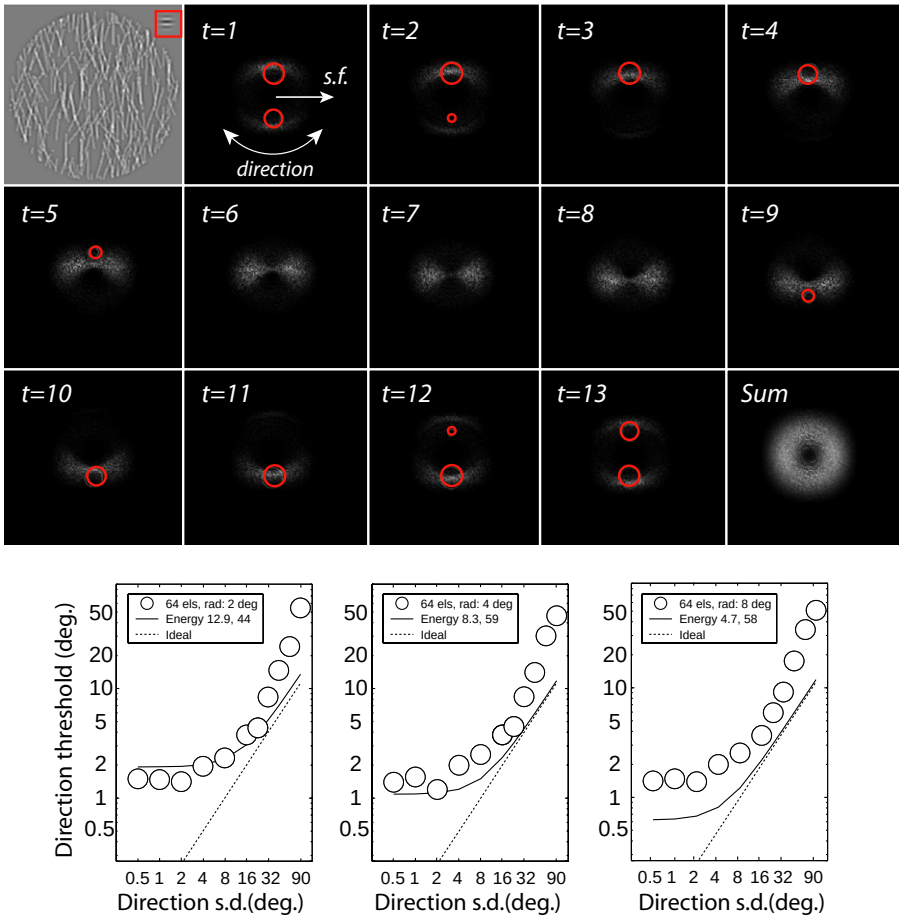

Fig. 8. (Top left) The frames from a typical psychophysical stimulus (circ. SD = 16 deg) have been superimposed; in the top right corner is a single frame from the optimal motion-energy filter for this sequence. Subsequent images show frames from the 3D space–space–time power spectrum of the stimulus (each frame shows energy for single temporal frequency, ω, indicated in cycles per movie). The pass-band of the optimal motion-energy filter is indicated by the ellipses (which mark filter-power >1 SD above mean). Note the presence of energy in the reverse direction due to the non-optimal (i.e., >1/4-cycle) jump-size employed and the stochastic nature of the stimuli. The image in the lower right-hand corner shows that the power summed across temporal frequency for this stimulus is spatially band-pass. (Bottom section) Psychophysical performance (averaged across three subjects) for 64 moving elements occupying regions with radii of 2, 4, and 8 deg. Dotted line is absolute ideal performance (i.e., free of local noise). Solid lines show predictions of a global motion direction estimator using the output of a bank of optimally-tuned motion energy mechanisms. Although no noise has been added to the filter predictions, human performance slightly exceeds motion energy predictions at the densest conditions.

panels in the upper section of Fig. 8 (labelled $t = 1, t = 2$), which show individual frames of the Fourier power spectrum of the stimulus (grey pixels) and the filter (red rings indicate high-energy potions of the filter). Note the presence of two regions of high amplitude in $t = 1, t = 2$ and $t = 12, t = 13$, indicating substantial sensitivity to backwards and forwards directions at high temporal frequencies.

We wondered how an ideal observer relying on the response of the motion energy mechanism would fare on the direction averaging task, compared to human performance. The ideal observer computed global direction based on the response magnitude of a bank of motion energy detectors. In this way, performance is only limited by the combination of intrinsic stimulus noise (directional variability and false correspondences) with the operation of the filter mechanism. To this end, we performed a Monte Carlo simulation with a total of 9856 stimulus movies from our experiments (7 conditions*11 directional SDs*128 unique stimuli). To test the motion energy model we convolved each of the stimulus movies with a bank of 24 directionally-tuned complex Gabor filters (separated by 15 deg; envelope $\sigma_{\text{envelope}} = 0.7\lambda$). For each of the resulting filtered movies we computed energy and then pooled across space and time to generate a single estimate of motion energy at a given direction. We computed the final estimate of direction by computing the population-vector direction from the resulting histogram;[1] the standard deviation of these values (corrected for wrapping) was used as an estimate of orientation threshold.

Representative results, for the three density conditions tested with 64 elements, are depicted in the lower part of Fig. 8. The dotted line shows ideal performance (i.e., of an integrator combining across all elements with no local noise). The solid lines are the predicted direction discrimination performance of the motion energy mechanism. Notice that across densities the model achieves high levels of efficiency at high levels of directional variability, as indicated both by its convergence with the noiseless ideal observer and with the estimated efficiency (the second figure in each entry of the caption gives the estimated sample size, the first, estimated local noise). Notice also that the overall shape of the curves does not match the psychophysical data: therefore, our results cannot be explained by a simple energy filter being pooled across the whole stimulus. However, we do note a reduction in internal noise with increasing area which would seem to indicate that increased correspondence noise, as it affects the operation of a local motion filter, could contribute to the similar effect we observed in the psychophysical results (Fig. 4C). Interestingly, in the highest density condition (leftmost graph) the motion energy cannot quite achieve subjects' levels of performance at low levels of stimulus variability. It therefore appears likely that for fast, dense patterns some other mechanism is required that can signal local second-order image structure. It remains open if existing proposals for such systems based on either "filter-rectify-filter" (Chubb & Sperling, 1988; Wilson et al., 1992) or image gradient approaches (e.g., Benton & Johnston, 2001; Johnston, McOwan, & Buxton, 1992) can achieve this levels of precision.

[1] Note that we also tried a maximum likelihood estimator which produced identical results.

5.4. Extraction of global direction

The estimates of local noise and efficiency we have reported constrain the development of models of global motion processing. The system as described is able not only to discriminate direction crudely in the presence of enormous directional variability ($\sigma = 90$ deg) but also to make extremely fine directional judgments—with thresholds approaching 1 deg—under conditions of low directional variability. We found that the pattern of breakdown in threshold observed for increasing directional variability was consistent with standard equivalent noise (i.e., additive local noise, limited efficiency) and that both local and global noise was largely determined by the number of elements present irrespective of density/region-size.

It is implicit in the EN model that averaging use a population-vector code and this proved adequate to account for our data using the statistical model. However, when we introduced Poisson noise, a local-to-global direction-encoding model using a population vector underestimated subjects thresholds at high levels of directional variability. A maximum likelihood operator could more closely match observers' break down in performance at high levels of directional variability for various numbers of elements. The concordance between model and data were close given the small number of free parameters employed and it is remarkable that such simple models can overcome the effect of Poisson noise and wide channel-bandwidths to produce a reasonable approximation to human performance over the range of stimuli employed. However, the observation that PVA frequently outperformed subjects cannot be taken as strong evidence against the mechanism per se. It is quite possible that an additional source of noise further constrains averaging at high levels of directional noise. In particular, PVA schemes outperform subjects at high directional variability because they are blind to the range of directions present and treat all direction samples equally. In contrast, MLE operators underweight the contribution of directions falling outside of the range of the template being matched. A simple reason that PVA schemes might similarly underweight directions falling far from the mean is *motion opponency* (as discussed above). If PVA were computed in area MT,

then near-opposite directions would be suppressed even when those directions are informative, as is always the case for an averaging task. Thus, MLE-like behaviour could be achieved by PV signal averaging once those signals have undergone directional opponency. The definitive test of this would be to measure MT response to patterns of various directional bandwidths.

Electrophysiological estimates of the bandwidths of primate MT neurons are close to 90 deg (Albright, 1984; Britten & Newsome, 1998; Snowden et al., 1992). If we consider PVA encoding as formally equivalent to least-squares fitting of a cosinusoid, then it is informative to note that the best fitting wrapped Gaussian to a cosinusoid is 93 deg. This would appear to represent strong evidence for PVA encoding in its own right. However, note that although our modelling suggests quite narrow bandwidths for the templates with MLE, one should resist the temptation to relate these figures directly to electrophysiological estimates of bandwidth. The values we report are the effective template bandwidths and are a consequence of the operation of the system as a whole.

Distributed channel codes, such as the one for direction we describe, rely on response-redundancy to compensate for intrinsic noise. However, correlations within that noise are present in the response of neurons in area MT (Bair, Zohary, & Newsome, 2001; Zohary, Shadlen, & Newsome, 1994). The effect of such correlations can affect the information capacity of the code (Abbott & Dayan, 1999) and it remains to be seen how this will affect the predictions of candidate schemes for global direction encoding.

The computational models described here compute a unitary direction upon which they base their 2AFC decision. An additional factor we have not explicitly considered is the role of perceived transparency in motion coherence tasks and in our own experiments. For a fixed sample size, increasing the directional range, or decreasing the strength of a coherent signal increases the probability of local peaks occurring in the direction histogram at directions other than the underlying mean direction. These spurious local peaks could bias the output of the visual system away from the mean. The introduction of such structure could explain why Zohary et al. (1996) found that observers apparently switch strategy depending on the directional structure of the stimulus. When all directions are near equi-probable observers rely on the mean direction but as asymmetries are introduced, they appear to switch to a winner-take-all scheme. A complete account of global direction integration therefore needs to take into account the influence of multi-modal directional structure in the stimulus (e.g., transparency). A reasonable candidate for this would be an efficient population encoding scheme using basis functions of the sort suggested by Zemel et al. (1998). It remains to be seen how the efficiency of such models compares to human performance with complex motion.

6. Conclusions

We conclude that the equivalent noise paradigm is useful for separating the influence of local and global factors in direction discrimination. We report that both local and global limits are largely determined by the number of elements present irrespective of density or arrangement. The changes in local noise that we observe with changes in stimulus density are likely attributable to correspondence noise of the sort reported by Barlow and Tripathy (1997). The dependence of local and global noise on element number can be explained by a simple channel-based model of direction integration that incorporates Poisson noise. When two different strategies for computing direction integration were compared we found that a maximum-likelihood estimator produced a better account of our data than a simple population vector averaging scheme, and had the additional advantage that it could naturally account for motion coherence thresholds (collected with similar stimulus to the EN experiment) without further modification. The most desirable extension of the model as it stands is to incorporate some form of filtering to estimate local image motion. Our initial efforts with motion energy mechanisms were only partially successful and suggested that the estimation of the local direction of two-dimensional patterns with the impressive precision achieved by human observers, is non-trivial in its own right.

Acknowledgments

This project was funded by BBSRC Grant 31/S17766 to SCD and PJB. We thank Alan Johnston, Wyeth Bair and Peter Dayan for useful discussion of this work.

Appendix A. Directional psychometric functions

Although it is usual to fit psychometric functions (PFs) for orientation or direction-discrimination tasks with aperiodic functions (e.g., the cumulative Gaussian; CG), this fails to take into account *stimulus wrapping*. Consider the underlying uncertainty function associated with a directional PF. For a task involving judgment of clockwise versus anti-clockwise relative to vertical-upwards, the observer is uncertain both when the stimulus is near to vertical-upwards motion *but also when it approaches vertical downwards motion*. The shaded functions inset in Figs. 9A–C show the probability of making a clockwise classification (relative to 0 deg), over a

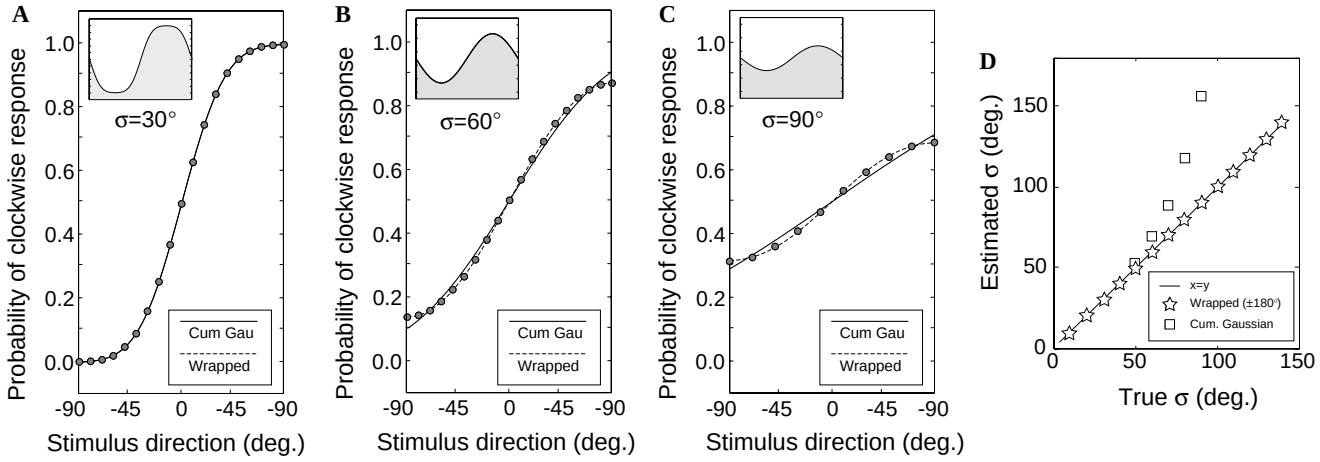

Fig. 9. (A–C) Ideal directional psychometric functions. Filled symbols plot the probability (evaluated using Monte Carlo simulation) of categorising a stimulus as clockwise of horizontal (0 deg) assuming representation of the stimulus with uncertainty determined by a wrapped Gaussian p.d.f. Insets show the full psychometric function, from −180 to +180 deg, for increasing values of σ. Fits from a cumulative Gaussian (solid line) or a wrapped cumulative Gaussian (dashed line) do not appear to differ greatly, but (D) shows that the former greatly overestimates directional uncertainty for σ's above ~50 deg.

stimulus-range of ±180 deg, when the stimulus is represented with increasing uncertainty. We express this uncertainty as the circular standard deviation of a wrapped Gaussian function, and the shaded regions shown are the cumulative probability densities for these functions. The symbols in Figs. 9A–C are drawn from the centre section of the periodic psychometric functions (i.e., for stimuli in the range ±90 deg). Good fits over this range can be made with either standard CG or a wrapped CG functions. However, Fig. 9D shows that fitting an unwrapped CG will lead to overestimation of the true internal circular standard deviation for circular $\sigma > 50$ deg. As a reasonable rule of thumb, if the tails of a PF do not reach 90% (discounting finger errors) over the range ±90 deg then the standard cumulative Gaussian model is inadequate. Note that because one half of the complete underlying uncertainty function is the mirror symmetric opposite of the other there is nothing to be gained from collecting data over a range greater than ±90 deg.

This technique for fitting periodic psychometric functions may have application in other domains involving noisy performance measurements for oriented stimuli (e.g., neurometric modelling). For example, we have refit some of the data a physiological study of MT directional tuning (Britten & Newsome, 1998) and estimate that the standard deviations quoted (derived using unwrapped Gaussians) overestimate bandwidth by approximately 10%.

Appendix B. Evaluating wrapped normal functions

Mardia and Jupp (2000) give a more useful form of the wrapped normal function than Eq. 4 derived using theta functions (Whittaker & Watson, 1944):

$$f(\theta) = \frac{1}{2\pi}\left(1 + 2\sum_{p=1}^{\infty} \rho^{p^2} \cos p(\theta - \mu)\right), \tag{15}$$

where $\rho = \exp(-\sigma^2/2)$ and σ is the circular SD. In practice, the WN may be computed to a desired levels of accuracy using the following MATLAB (MathWorks) function derived from an algorithm given in (Mardia & Jupp, 2000):

```
function p = WrapNormP(t, mu, rho, tol)
var = -2*log (rho);
indl = 1/sqrt (var*2*pi);
k = 0; error = Inf;
p = indl*exp (-((t-mu)^2)/(2*var));
while (error > tol)
  k = k+1;
  LastP = p;
  p = LastP...
  +indl*(exp(-((t-mu+2*pi*k)^2)/
  (2*var)...
  +exp(-((t-mu+2*pi*-k)^2)/(2*var)));
  error = abs(p-LastP);
end
```

For reasons outlined in the last section, it would be desirable to have a cumulative version of this function; i.e., for an angle θ, to evaluate the sum of the probabilities over the range $\theta - \pi$ to θ. A simple way of modifying the cumulative Gaussian (CG) to take into account direction is to transform the stimuli using the `sin()` function. One can then see that clockwise stimuli will now produce positive values, and anticlockwise will produce negative. This procedure works well for very high ($\sigma > 80$ deg) and very low ($\sigma < 50$)

levels of variance but fails at intermediate values, when it cannot predict maximum levels of performance. There is a second problem with this approach and that is how to convert the standard deviation derived from the CG fit to the `sin()` transformed data back to an equivalent σ of a wrapped CG. Standard approaches to approximating the standard deviation of a transformed data set (e.g., for a transform, $z = f(x)$, $\sigma(z) = f'(X)\sigma(x)$; e.g., Barford (1967)) are inapplicable because of the derivative involved. Our solution is to divide the range into two at 80 deg and fit the `sin()` transformed dataset at high variances only. We then work back to the correct σ using a polynomial approximation (parameters are given in the function below). For low variances, one can adequately approximate the wrapping by using three standard CGs: one centred at the mean, one centred at the mean +180 deg, and one centred at the mean −180deg. These two functions give the tails of the parallel uncertainty function falling 180 deg from the centre of stimulus range and may be combined to allow or the effects of wrapping. This approach does not take into account self-wrapping (i.e., the influence of a distribution wrapping around onto itself) but for low levels of variance the effect of this are negligible. The MATLAB function for this is:

```
function p = WrapCumNormP(theta, mu, sd)
theta = theta - mu;
if (sd < 1.4) %~80 degrees
  p = NormCum (theta, 0, sd^2)...
  -NormCum (theta-pi, 0, sd^2)+ ...
  (1-NormCum (theta+pi, 0, sd^2));
else
  sd2 = polyval ([-3.546 39.131 -158.724...
  309.646 -292.966 109.1288], sd);
  p = NormCum (sin(theta), 0, sd2^2);
end
```

where `NormCum` evaluates the cumulative Gaussian function using either MATLAB's built in approximation to the error function (`erf`)—as in the PsychToolbox version; (Brainard, 1997)—or using a polynomial approximation based on Probit analysis (Finney, 1971). Speed considerations make the latter desirable when many fits are required (e.g., for bootstrapping).

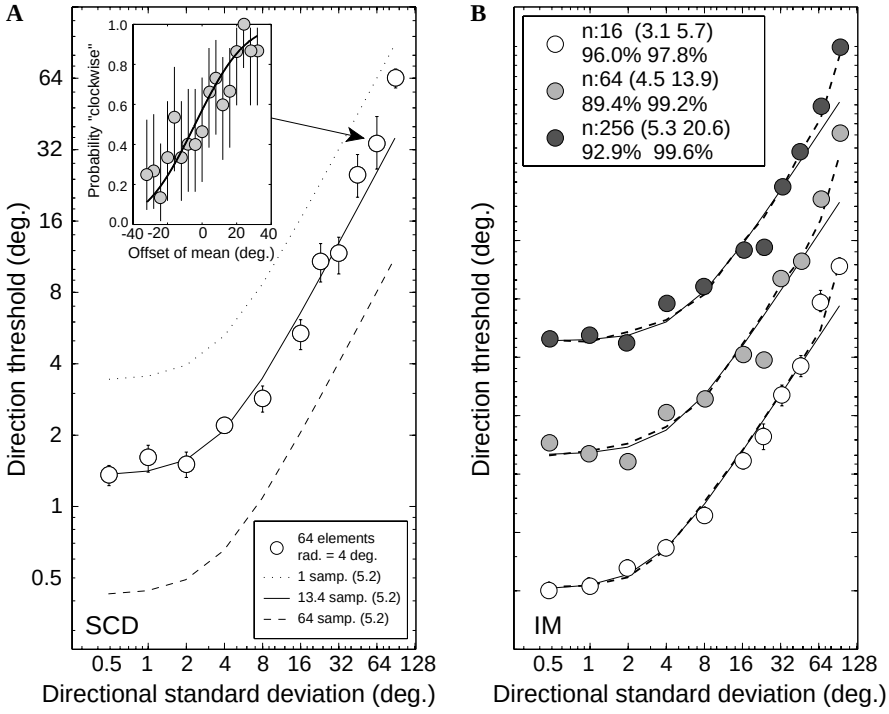

Fig. 10. (A) SCD's thresholds for estimating the mean direction of 64 elements within a patch of radius 4.0 deg. The solid line plots predictions from the unmodified EN model, which indicates that the observer is using 13.4 samples with each having an equivalent directional SD of 5.2 deg. Other lines show predictions of models with similar additive noise, but averaging either all (dashed line) or one (dotted line) of the elements. Note that performance deteriorates faster than EN predicts at high directional SDs. (B) Data from IM (datasets are vertically offset from one another for clarity). The EN model (solid lines) again underestimates threshold at high levels of directional SD Dashed lines show the better fit provided by an EN model employing a simulated observer which bases its decision on the true average direction, (allowing for wrapping). The legend shows number of elements followed by the local noise and global sample from the best fitting EN model, along with the R^2 of the fits to the unmodified and Monte Carlo versions of the model, respectively.

Appendix C. Deviation From EN with high directional variability

Fig. 10A shows typical performance on the directional averaging task and plots the offset of the mean direction of a field of moving elements at threshold as a function of directional variability. Parameters from the unmodified EN model indicate that the ability of SCD to average the direction of 64 elements moving within a 4.0 deg radius window is consistent with his using 13.4 elements, where each element has a local directional SD of 5.2 deg. Note the deviation from the prediction at high directional variance, where the observer's performance is consistent with a much lower level of efficiency (dotted line). The psychometric function inset in Fig. 10 corresponds to the data point at SD = 64 deg and illustrates that this apparent deterioration in performance is not due to a poor fit of psychometric functions. Fig. 10B reiterates this point using data from observer IM (datasets are offset from on another for clarity). Solid lines are the predictions of the unmodified EN model fit using the bootstrapping technique described in combination with a standard least-squares goodness-of-fit measure. The predictions shown by the dashed lines were generated using the same parameters.

Applying the correction for wrapping improved R^2 by an average of 3% compared to unmodified predictions. Given this modification it is remarkable that the predictions of EN hold over such an enormous range of directional variability.

References

Abbott, L. F., & Dayan, P. (1999). The effect of correlated variability on the accuracy of a population code. *Neural Computation, 11*(1), 91–101.

Adelson, E. H., & Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America. A: Optics and Image Science, 2*(2), 284–299.

Albright, T. D. (1984). Direction and orientation selectivity of neurons in visual area MT of the macaque. *Journal of Neurophysiology, 52*(6), 1106–1130.

Atkinson, J., King, J., Braddick, O., Nokes, L., Anker, S., & Braddick, F. (1997). A specific deficit of dorsal stream function in Williams' syndrome. *Neuroreport, 8*(8), 1919–1922.

Baez, K. A., McNaught, A. I., Dowler, J. G., Poinoosawmy, D., Fitzke, F. W., & Hitchings, R. A. (1995). Motion detection threshold and field progression in normal tension glaucoma. *The British Journal of Ophthalmology, 79*(2), 125–128.

Bair, W., Zohary, E., & Newsome, W. T. (2001). Correlated firing in macaque visual area MT: Time scales and relationship to behavior. *Journal of Neuroscience, 21*(5), 1676–1697.

Baker, C. L., Jr., Hess, R. F., & Zihl, J. (1991). Residual motion perception in a ''motion-blind'' patient, assessed with limited-lifetime random dot stimuli. *Journal of Neuroscience, 11*(2), 454–461.

Barford, N. C. (1967). *Experimental measurements: Precision, error and truth*. London: Addison-Wesley.

Barlow, H., & Tripathy, S. P. (1997). Correspondence noise and signal pooling in the detection of coherent visual motion. *Journal of Neuroscience, 17*(20), 7954–7966.

Barlow, H. B. (1956). Retinal noise and absolute threshold. *Journal of the Optical Society of America, 46*, 634–639.

Benton, C. P., & Johnston, A. (2001). A new approach to analysing texture-defined motion. *Proceedings of the Royal Society of London. Series B. Biological sciences, 268*(1484), 2435–2443.

Bex, P. J., & Dakin, S. C. (2002). Comparison of the spatial-frequency selectivity of local and global motion detectors. *Journal of the Optical Society of America A: Optics, image science, and vision, 19*(4), 670–677.

Bex, P. J., Simmers, A. J., & Dakin, S. C. (2003). Grouping local directional signals into moving contours. *Vision Research, 43*(20), 2141–2153.

Born, R. T., & Tootell, R. B. (1992). Segregation of global and local motion processing in primate middle temporal visual area [published erratum appears in Nature 1993 Sep 16;365(6443):279]. *Nature, 357*(6378), 497–499.

Braddick, O. (1974). A short-range process in apparent motion. *Vision Research, 14*(7), 519–527.

Braddick, O. (1993). Segmentation versus integration in visual motion processing. *Trends in Neurosciences, 16*, 263–268.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10*, 433–436.

Britten, K. H., & Newsome, W. T. (1998). Tuning bandwidths for near-threshold stimuli in area MT. *Journal of Neurophysiology, 80*(2), 762–770.

Britten, K. H., Shadlen, M. N., Newsome, W. T., & Movshon, J. A. (1992). The analysis of visual motion: A comparison of neuronal and psychophysical performance. *Journal of Neuroscience, 12*(12), 4745–4765.

Britten, K. H., Shadlen, M. N., Newsome, W. T., & Movshon, J. A. (1993). Responses of neurons in macaque MT to stochastic motion signals. *Visual Neuroscience, 10*(6), 1157–1169.

Burr, D. C., Morrone, M. C., & Vaina, L. M. (1998). Large receptive fields for optic flow detection in humans. *Vision Research, 38*(12), 1731–1743.

Chubb, C., & Sperling, G. (1988). Drift-balanced random stimuli: A general basis for studying non-Fourier motion perception. *Journal of the Optical Society of America A: Optics and image science, 5*(11), 1986–2007.

Dakin, S. C. (2001). Information limit on the spatial integration of local orientation signals. *Journal of the Optical Society of America A: Optics, Image science and vision, 18*(5), 1016–1026.

Dakin, S. C., & Mareschal, I. (2000). The role of relative motion computation in 'direction repulsion'. *Vision Research, 40*(7), 833–841.

Dakin, S. C., Mareschal, I., & Bex, P. J. (2004). Local noise (not efficiency) limits direction integration in the periphery. *ECVP*.

Dean, A. F. (1981). The variability of discharge of simple cells in the car striate cortex. *Experimental Brain Research, 44*, 437–440.

Deneve, S., Latham, P. E., & Pouget, A. (1999). Reading population codes: A neural implementation of ideal observers. *Nature Neuroscience, 2*(8), 740–745.

Eagle, R. A., & Rogers, B. J. (1996). Motion detection is limited by element density not spatial frequency. *Vision Research, 36*(4), 545–558.

Fine, I., Anderson, C. M., Boynton, G. M., & Dobkins, K. R. (2004). The invariance of directional tuning with contrast and coherence. *Vision Research, 44*(9), 903–913.

Finney, D. J. (1971). *Probit analysis*. Cambridge, MA: Cambridge University Press.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.

Girard, P., Salin, P. A., & Bullier, J. (1992). Response selectivity of neurons in area MT of the macaque monkey during reversible inactivation of area V1. *Journal of Neurophysiology, 67*(6), 1437–1446.

Groh, J., Born, R., & Newsome, W. T. (1997). How is a sensory map read out? Effect of microstimulation in visual area MT on saccades

and smooth pursuit eye movements. *Journal of Neuoscience, 17*, 4312–4330.

Gunn, A., Cory, E., Atkinson, J., Braddick, O., Wattam-Bell, J., Guzzetta, A., & Cioni, G. (2002). Dorsal and ventral stream sensitivity in normal development and hemiplegia. *Neuroreport, 13*(6), 843–847.

Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and function architecture in the cat's visual cortex. *Journal of Physiology, 160*, 106–154.

Joffe, K. M., Raymond, J. E., & Chrichton, A. (1997). Motion coherence perimetry in glaucoma and suspected glaucoma. *Vision Research, 37*(7), 955–964.

Johnston, A., McOwan, P. W., & Buxton, H. (1992). A computational model of the analysis of some first-order and second-order motion patterns by simple and complex cells. *Proceedings of the Royal Society of London. Series B. Biological sciences, 250*(1329), 297–306.

Koenderink, J. J. (1986). Optic flow. *Vision Research, 26*(1), 161–179.

Li, C. S. (2002). Impaired detection of visual motion in schizophrenia patients. *Progress in neuro-psychopharmacology and biological psychiatry, 26*(5), 929–934.

Lund, J. S. (1988). Anatomical organization of macaque monkey striate visual cortex. *Annual Review of Neuroscience, 11*, 253–288.

Mardia, K. V., & Jupp, P. E. (2000). *Directional statistics*. Chichester, UK: Wiley.

Meese, T. S., & Harris, M. G. (2001). Broad direction bandwidths for complex motion mechanisms. *Vision Research, 41*(15), 1901–1914.

Milne, E., Swettenham, J., Hansen, P., Campbell, R., Jeffries, H., & Plaisted, K. (2002). High motion coherence thresholds in children with autism. *Journal of Child Psychology and Psychiatry, 43*(2), 255–263.

Morgan, M. J., & Ward, R. (1980). Conditions for motion flow in dynamic visual noise. *Vision Research, 20*(5), 431–435.

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

Movshon, J. A., & Newsome, W. T. (1996). Visual response properties of striate cortical neurons projecting to area MT in macaque monkeys. *Journal of Neuroscience, 16*(23), 7733–7741.

Newsome, W. T., Britten, K. H., & Movshon, J. A. (1989). Neuronal correlates of a perceptual decision. *Nature, 341*(6237), 52–54.

Newsome, W. T., & Pare, E. B. (1988). A selective impairment of motion perception following lesions of the middle temporal visual area MT. *Journal of Neuroscience, 8*(6), 2201–2211.

Nichols, M. J., & Newsome, W. T. (2002). Middle temporal visual area microstimulation influences veridical judgments of motion direction. *Journal of Neuoscience, 22*, 9530–9540.

Nowlan, S. J., & Sejnowski, T. J. (1995). A selection model for motion processing in area MT of primates. *Journal of Neuroscience, 15*(2), 1195–1214.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming number into movies. *Spatial Vision, 10*, 437–442.

Pelli, D. G., & Zhang, L. (1991). Accurate control of contrast on microcomputer displays. *Vision Research, 31*, 1337–1350.

Qian, N., Andersen, R. A., & Adelson, E. H. (1994). Transparent motion perception as detection of unbalanced motion signals. I. Psychophysics. *Journal of Neuroscience, 14*(12), 7357–7366.

Rodman, H. R., Gross, C. G., & Albright, T. D. (1989). Afferent basis of visual response properties in area MT of the macaque. I. Effects of striate cortex removal. *Journal of Neuroscience, 9*(6), 2033–2050.

Salinas, E., & Abbott, L. F. (1994). Vector reconstruction from firing rates. *Journal of Computational Neuroscience, 1*(1–2), 89–107.

Salzman, C. D., Britten, K. H., & Newsome, W. T. (1990). Cortical microstimulation influences perceptual judgements of motion direction [published erratum appears in Nature 1990 Aug 9;346(6284):589] [see comments]. *Nature, 346*(6280), 174–177.

Scase, M. O., Braddick, O. J., & Raymond, J. E. (1996). What is noise for the motion system? *Vision Research, 36*(16), 2579–2586.

Seung, H. S., & Sompolinsky, H. (1993). Simple models for reading neuronal population codes. *Proceedings of the National Academy of Sciences of the United States of America, 90*(22), 10749–10753.

Simmers, A. J., Ledgeway, T., & Hess, R. F. (2005). The influences of visibility and anomalous integration processes on the perception of global spatial form versus motion in human amblyopia. *Vision Research, 45*(4), 449–460.

Simmers, A. J., Ledgeway, T., Hess, R. F., & McGraw, P. V. (2003). Deficits to global motion processing in human amblyopia. *Vision Research, 43*(6), 729–738.

Simoncelli, E. P., & Heeger, D. J. (1998). A model of neuronal responses in visual area MT. *Vision Research, 38*(5), 743–761.

Smith, A. T., Snowden, R. J., & Milne, A. B. (1994). Is global motion really based on spatial integration of local motion signals. *Vision Research, 34*(18), 2425–2430.

Snippe, H. P. (1999). Parameter estimation from population codes: A critical assessment. In L. Abbot & S. J. Sejnowski (Eds.), *Neural codes and distributed representaton* (pp. 85–103). Cambridge, MA: MIT Press.

Snowden, R. J., Treue, S., & Andersen, R. A. (1992). The response of neurons in areas V1 and MT of the alert rhesus monkey to moving random dot patterns. *Experimental Brain Research, 88*(2), 389–400.

Snowden, R. J., Treue, S., Erickson, R. G., & Andersen, R. A. (1991). The response of area MT and V1 neurons to transparent motion. *Journal of Neuroscience, 11*(9), 2768–2785.

Softky, W. L., & Koch, C. (1992). Cortical cells should fire regularly but do not. *Neural Computation, 4*, 643–646.

Swindale, N. V. (1998). Orientation tuning curves: Empirical description and estimation of parameters. *Biological Cybernetics, 78*(1), 45–56.

Talcott, J. B., Hansen, P. C., Assoku, E. L., & Stein, J. F. (2000). Visual motion sensitivity in dyslexia: evidence for temporal and energy integration deficits. *Neuropsychologia, 38*(7), 935–943.

Tolhurst, D. J., Movshon, J. A., & Dean, A. F. (1983). The statistical reliability of signals in single neurons in cat and monkey visual cortex. *Vision Research, 23*(8), 775–785.

Treue, S., Hol, K., & Rauber, H. J. (2000). Seeing multiple directions of motion-physiology and psychophysics. *Nature Neuroscience, 3*(3), 270–276.

Turano, K., & Wang, X. (1992). Motion thresholds in retinitis pigmentosa. *Investigative Ophthalmology and Visual Science, 33*(8), 2411–2422.

Verghese, P., Watamaniuk, S. N., McKee, S. P., & Grzywacz, N. M. (1999). Local motion detectors cannot account for the detectability of an extended trajectory in noise. *Vision Research, 39*(1), 19–30.

Vogels, R., Spileers, W., & Orban, G. A. (1989). The response variability of striate cortical neurons in the behaving monkey. *Experimental Brain Research, 77*(2), 432–436.

Watamaniuk, S. N. (1993). Ideal observer for discrimination of the global direction of dynamic random-dot stimuli. *Journal of the Optical Society of America A: Optics and image science, 10*(1), 16–28.

Watamaniuk, S. N., & Heinen, S. J. (1999). Human smooth pursuit direction discrimination. *Vision Research, 39*(1), 59–70.

Watson, A. B., & Pelli, D. G. (1983). QUEST: A Bayesian adaptive psychometric method. *Perception and Psychophysics, 33*(2), 113–120.

Watt, R. J., & Andrews, D. (1981). APE: Adaptive probit estimation of psychometric functions. *Current Psychological Review, 1*, 205–214.

Watt, R. J., & Morgan, M. J. (1984). Spatial filters and the localization of luminance changes in human vision. *Vision Research, 24*(10), 1387–1397.

Whittaker, E. T., & Watson, G. S. (1944). *A course in modern analysis*. Cambridge, UK: Cambridge Univeristy Press.

Williams, D. W., & Sekuler, R. (1984). Coherent global motion percepts from stochastic local motions. *Vision Research, 24*(1), 55–62.

Wilson, H. R., Ferrera, V. P., & Yo, C. (1992). A psychophysically motivated model for two-dimensional motion perception. *Visual neuroscience, 9*(1), 79–97.

Wilson, H. R., & Kim, J. (1994). Perceived motion in the vector sum direction. *Vision Research, 34*(14), 1835–1842.

Zemel, R. S., Dayan, P., & Pouget, A. (1998). Probabilistic interpretation of population codes. *Neural Computation, 10*(2), 403–430.

Zohary, E., Scase, M. O., & Braddick, O. J. (1996). Integration across directions in dynamic random dot displays: Vector summation or winner take all. *Vision Research, 36*(15), 2321–2331.

Zohary, E., Shadlen, M. N., & Newsome, W. T. (1994). Correlated neuronal discharge rate and its implications for psychophysical performance. *Nature, 370*(6485), 140–143.