



Pooling and segmenting motion signals

David C. Burr^{a,b}, Stefano Baldassi^{a,*}, M. Concetta Morrone^c, Preeti Verghese^d

^a Dipartimento di Psicologia, Università Degli Studi di Firenze, Via S. Nicolò 89, 50125 Florence, Italy

^b Department of Psychology, University of Western Australia, Perth WA 6009, Australia

^c University of Pisa and Scientific Institute Stella Maris, Calambrone, Pisa, Italy

^d Smith-Kettlewell Eye Research Institute, San Francisco, CA 94115, USA

ARTICLE INFO

Article history:

Received 26 June 2008

Received in revised form 14 October 2008

Keywords:

Motion

Attention

Summation

Visual search

ABSTRACT

Humans are extremely sensitive to visual motion, largely because local motion signals can be integrated over a large spatial region. On the other hand, summation is often not advantageous, for example when segmenting a moving stimulus against a stationary or oppositely moving background. In this study we show that the spatial extent of motion integration is not compulsory, but is subject to voluntary attentional control. Measurements of motion coherence sensitivity with summation and search paradigms showed that human observers can combine motion signals from cued regions or patches in an optimal manner, even when the regions are quite distinct and remote from each other. Further measurements of contrast sensitivity reinforce previous studies showing that motion integration is preceded by a local analysis akin to contrast thresholding (or intrinsic uncertainty). The results were well modelled by two standard signal-detection-theory models.

© 2008 Elsevier Ltd. All rights reserved.

1. Introduction

Reliable motion perception requires spatial integration of visual motion signals, largely because averaging over space improves signal-to-noise ratios. Integration of local motion signals along extensive complex flow trajectories can provide the basis for detection of optic flow vectors such as expansion and rotation. However, there are other circumstances when summation is not useful, such as for segmenting moving objects from their background. It is still not clear how the visual system reconciles the conflicting demands of integration and segmentation: it could involve multiple representations of motion in different regions, or adaptive strategies for a particular task (Braddick, 1993).

There is considerable evidence that motion signals are integrated over quite large regions (Yang & Blake, 1994). Perhaps one of the more explicit examples is Morrone, Burr and Vaina's (1995) study showing that coherence sensitivity for motion signals embedded in random noise improved with increasing stimulus area in a fashion that implied linear integration. The result held both for simple translational trajectories and along more complex flow trajectories, circular and radial motion. Integration occurred over very large areas, extending up to 70° of space (Burr, Morrone, & Vaina, 1998), and also over long durations, up to 3 s (Burr & Santoro, 2001). However, when motion performance was measured by

contrast thresholds (minimum contrast to support reliable motion discrimination), sensitivity was almost independent of stimulus area, increasing only by the amount predicted by “probability summation” between independent detectors (Graham, 1977; Pelli, 1985). The results were well modelled by two stages of motion analysis: an early, local, contrast-thresholded stage, followed by one that integrates over larger regions. The early stage could correspond to V1/V2, or possibly MT (Sclar, Maunsell, & Lennie, 1990), while the later stage is more consistent with the properties of cells reported in area MSTd (Duffy & Wurtz, 1991; Orban et al., 1992; Tanaka & Saito, 1989).

The experiments of Morrone et al. (1995) and of Burr et al. (1998) were designed to demonstrate summation of motion signals, but did not test whether this summation was “compulsory”. That is, whether subjects could choose not to integrate, or to choose the region over which to integrate. In this study we investigate the ability of observers to integrate selectively over pre-cued regions. In some conditions there were multiple signal stimuli, in others a single signal-stimulus embedded within noise distractors, akin to a search paradigm. In visual search paradigms, some tasks show set-size dependencies, others not. For example, contrast, size, orientation and speed discrimination show strong dependencies on set size, all well accounted for by signal detection theory (Eckstein, 1998; Morgan, Ward, & Castet, 1998; Palmer, 1994; Palmer, Ames, & Lindsey, 1993; Verghese & Stone, 1995). However, luminance- and colour-contrast detection show no set-size dependency, although contrast increments do (Baldassi & Burr, 2004). Within the framework of signal detection theory, the lack of dependence

* Corresponding author. Address: Department of Psychology, University of Florence, Via della Cittadella, 7, 50125 Florence, Italy.

E-mail address: stefano.baldassi@unifi.it (S. Baldassi).

on set-size can be explained by assuming high intrinsic uncertainty for contrast detection tasks (Pelli, 1985).

The present research had two broad goals: to study whether integration of motion signals can be regulated by cueing attention; and to study set-size effects for motion discrimination, that can provide useful information about underlying mechanisms.

2. Methods

2.1. Stimuli

All stimuli were generated on a Sony Trinitron monitor under the control of a Cambridge Research Systems VSG2/5 framestore at 110 Hz. The stimuli comprised 600 dots randomly positioned within a 20° square on the monitor (24×24 cm, viewed from 68 cm). Each dot subtended $20'$ diameter, randomly black or white against a grey background (30 cd/m^2). Within given regions ($3 \times 3^\circ$ squares), a proportion of dots moved coherently leftwards or rightwards at 13° per/s. The motion was “limited lifetime”, where dots moved for either two or four displacements, after which the dots were reborn in a new random location. One half or one quarter of the dots (or more generally, $1/\text{lifetime}$) died and were reborn on each motion frame. Different dot lifetimes could be used to avoid floor and ceiling effects. Each motion frame corresponded to two video frames (18.2 ms), and the motion sequence lasted for eight motion frames (145 ms). If the dot fell outside the target square, it was wrapped round to the other side. Examples of stimuli are shown in Fig. 1 and in on-line movies.

Two measurements of motion performance were used: coherence sensitivity and contrast sensitivity. For the coherence task a proportion of the dots moved in the coherent direction, while the remainder were plotted in random positions within the target square (dot contrast was 95% for this task). Sensitivity was taken as the inverse of the proportion of coherently moving dots. Contrast sensitivity was measured by varying the contrast of all dots, with sensitivity given by the inverse of Michelson contrast at threshold. In all cases, subjects were required to report whether the direction of motion was left or right, and each trial was scored correct or incorrect. Both the proportion of coherent dots and the contrast were varied dynamically by the QUEST routine (Watson & Pelli, 1983), which homed in near threshold. Thresholds were then calculated offline, by fitting probability of correct response curves with a cumulative Gaussian function that asymptoted at 50% and 100%. In separate analyses, we have fitted the psychometric functions with cumulative Weibull functions (also asymptoting at 50% and 100%), that captured the trend of the data seemingly well; however we report the Gaussian analysis in all the figures in order to be able to compare the present data with previous studies. In all conditions at least four separate QUEST sessions were run, each with 30 trials. Before data collection, at least one session was run in each condition to habituate the subjects to the task. Feedback was given after each trial both during training and data collection, to eliminate the possibility of a direction bias.

2.2. Summation

The summation paradigm was similar to that used by Morrone et al. (1995) except that the target regions could be cued by a partial cueing technique. Coherent dot motion (limited lifetime 2) was confined to one or more of eight possible target squares, each $3 \times 3^\circ$ arranged around a virtual circle abutting the visible square (see Fig. 1). Either 1, 2, 4 or 8 patches contained motion information, the others dynamic noise of matched density. Two summation conditions were run: in one the location of the patches containing motion information was cued with spokes that ap-

peared on motion onset and lasted for the duration of the motion; in the other all eight possible patches were cued (thereby providing no useful information). For both cued and uncued conditions, the target patches were placed symmetrically opposite each other, in any arbitrarily rotated position.

2.3. Search

In the search paradigm coherent dot motion (limited lifetime 4) was confined to only one of eight possible target squares ($3 \times 3^\circ$) arranged around a virtual circle abutting the visible square (see Fig. 1). The location of the target was cued by a partial cueing technique (Baldassi & Burr, 2004; Grindley & Townsend, 1968; Palmer, 1994), where within a particular session 1, 2, 4 or all 8 locations were cued (always including the actual target location). The cues were always arranged in a symmetric spoked pattern, rotated to an arbitrary position. They appeared at the same time as the motion stimulus, and lasted for the lifetime of the motion. The major variation between conditions was whether the target regions were clearly identified by an opaque grey mask covering the background (Fig. 1A), or whether they blended into the noisy background (Fig. 1B). As will be seen, this had a major effect on the results in some conditions.

Three authors, all with corrected to normal vision served as observers for these experiments.

3. Results

3.1. Summation

The summation results are shown in Fig. 2. The blue symbols show the condition where there was no selective cueing (all locations were cued), effectively a replication of Morrone et al.'s. (1995) condition where the non-signal sectors were filled with noise. This condition shows very strong set-size dependency, with sensitivity increasing in direct proportion to the stimulus area, indicated by log-log slopes near unity. This is predicted from an ideal linear integrator that sums linearly the signals, and adds the variance of the stimulus noise. As the noise is effectively constant in all conditions—generated by both the signal and non-signal sectors—sensitivity depends on the signal, and hence should increase linearly with sector number (Morrone et al., 1995).

When the patches containing the motion signals were cued the summation was far less steep (red symbols of Fig. 2). Here the log-log slopes were close to 0.5 (squareroot relationship), suggesting summation of local independent signals. This result is very similar to that obtained by Morrone et al. when the non-signal sectors were physically removed, and shows that cueing can be as effective as physically removing distractor (noise) sectors. Again this result is consistent with linear summation of stimuli. Because there is signal and noise at every location, the signal should increase linearly with set size as, should the variance of the cued (attended) patches. So the signal-to-noise ratio (ratio of signal to root-variance) increases with the squareroot of set-size. This is what was observed in both subjects, suggesting that observers can exclude the non-cued sectors from the analysis, while summing the cued sectors. This result indicates that summation is not obligatory, but clearly under attentional control. But for both variants of the experiment, our data are consistent with observers using the ideal decision rule, which is summation.

3.2. Search

The previous data show that subjects can control the region over which they summate: the only difference between the two

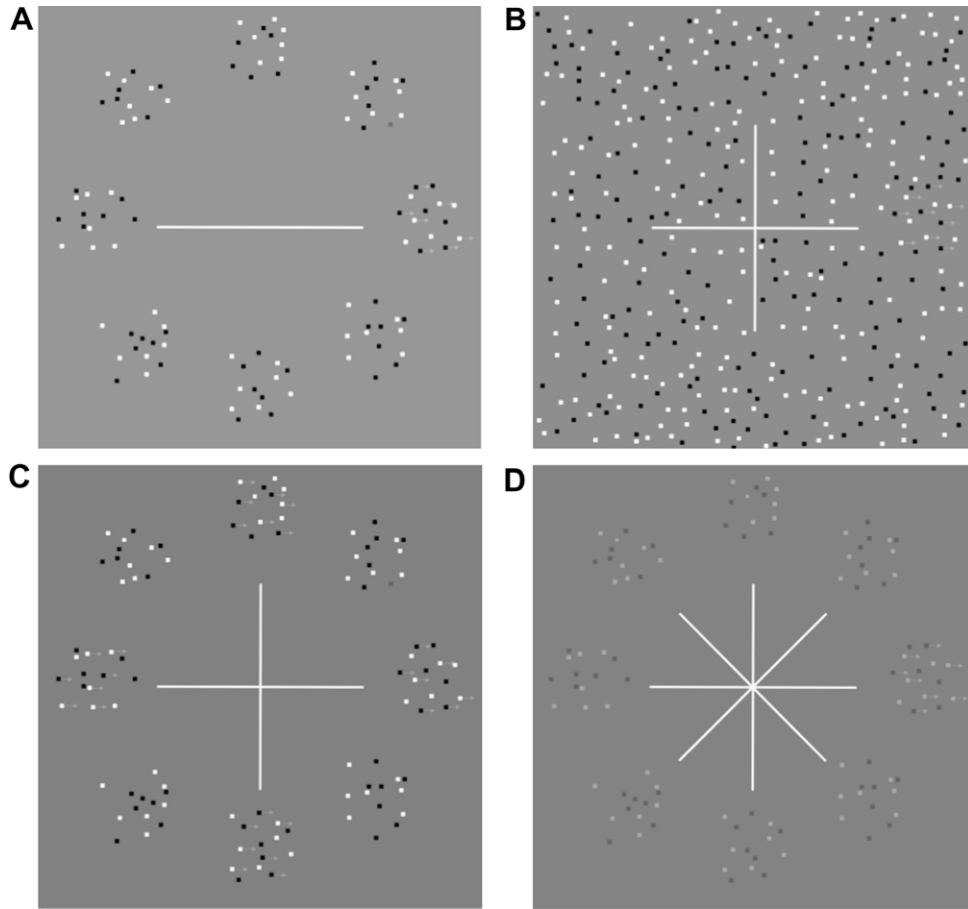


Fig. 1. Examples of stimuli and conditions. In all the trials, a proportion of dots in one or several of the regions pointed translated coherently to the left or to the right. The subject was required to report the direction of translation (but not which patch or patches seemed to contain the motion). (A) Example of the patch condition at set-size 2, in which one of the two clouds of dots cued by the spot contained a variable proportion of coherently translating dots, while all other patches contain dots independently moving in random directions. (B) Example of the region condition at set size 4, where the moving dots were not confined to a clearly visible patch. (C) Example of the cued summation condition, where the 4 patches containing the moving dots were correctly cued. (D) Example of the contrast threshold experiment at set size 8, where the patches were clearly delineated by noise dots of 30% contrast. One of the patches also contained low-contrast dots moving coherently.

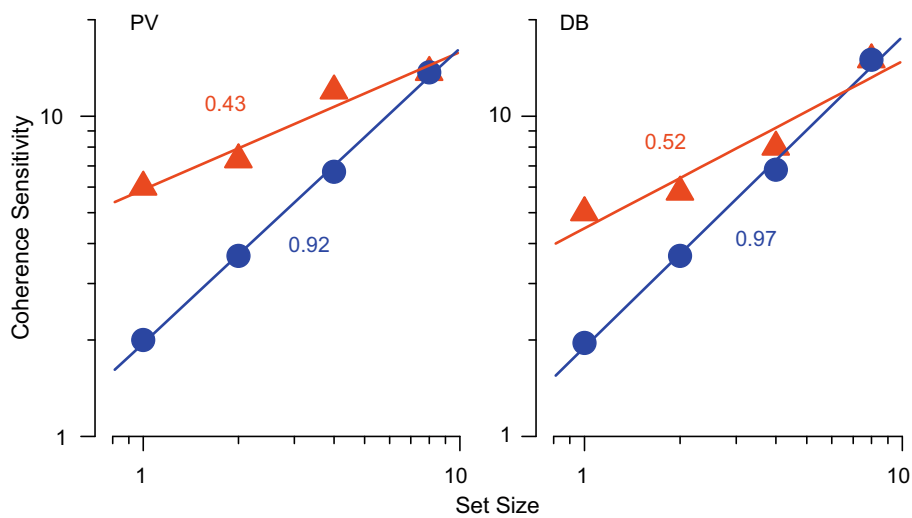


Fig. 2. Coherence sensitivity as a function of set size of two observers for the summation conditions. Circles plot the condition without informative cueing (all 8 patches were cued), triangles the condition with cueing. The lines are linear best fits of the data (on log–log plots) with the slope of the fits indicated by adjacent number. The log–log slope for the cued condition is close to 0.5 for both observers, while that of the uncued condition is near unity. Sensitivity is much better for the cued condition up to set size 8, where the conditions are identical.

sets of data at set-size 1 was whether the target patch was cued, and this caused a factor-of-three difference in sensitivity. To ex-

plore this effect further, we measured coherence thresholds for the condition when only one patch contained the signal, and all

others were noise “distractors” (Fig. 1A: a typical search paradigm). The signal patch was cued by a partial cueing technique, where 1, 2, 4 or 8 patches (always including the signal patch) were cued (Baldassi & Burr, 2004; Grindley & Townsend, 1968; Palmer, 1994). The red symbols of Fig. 3 show the results for similar conditions to those of the summation experiment (both signal and noise confined to distinct patches). Coherency sensitivity showed a strong dependency on cue size, with an average log–log slope near -0.50 . The extremes of these curves are comparable to the results of Fig. 2 at set-size 1 (except for a difference in absolute sensitivity due to different limited lifetimes). In both cases only one signal patch was displayed, which was either selectively cued or not, producing in both cases a three-fold difference in sensitivity.

In the experiments described so far, the motion signals were displayed within clearly delineated patches on a background devoid of dots. The blue symbols of Fig. 3 show results when the background was filled with noise dots at the same density and contrast (Fig. 1B). Sensitivity in this condition was about half that for the “patch” condition, but the dependency on set-size was very similar for the two conditions. Even when the positions of the potential stimuli and distractors were not well segmented from the background, subjects could still attend selectively to the set of potential stimuli.

We next measured contrast sensitivity for motion discrimination, as was done previously for luminance and color (Baldassi & Burr, 2004). The procedure was similar to the previous experiment except that the coherence of the motion was always 100%, and the contrast of the dots varied. In one condition (that we call “region”), the motion dots were displayed on their own, but always confined to the patches. This may seem different from the previous “region” condition, where dots filled the background; however, it was “region-like” in the sense that when the dots fell below threshold, they were invisible and blended into the background (following the procedure of Baldassi & Burr, 2004). In this sense it is analogous to when high contrast dots fell below coherence threshold, blending into the noise field. In the other condition (that we call “patches”), the stimuli were filled with an equal number of random dots of fixed (30%) contrast (Fig. 1D). These extra noise dots acted as a “pedestal” and define the motion regions as visually distinct “patches”, even when the stimuli were at contrast threshold. Otherwise, the “regions” tended to merge with the background at contrast threshold. As before, either 1, 2, 4 or 8 stimulus areas were cued, with the positions of the cues selected at random from trial to trial.

The results, shown in Fig. 4, were quite different from the coherence thresholds. In the presence of high contrast noise dots (that demarcated the signal patches), contrast sensitivity depended on set size, with a log–log dependency near 0.5 as before. Without the noise dots (triangle symbols), however, the dependency on set size was much reduced, nearly zero. This lack of dependency on set size is similar to that observed for contrast sensitivity for detecting Gaussian patches (Baldassi & Burr, 2004).

4. Discussion

One goal of this study was to test whether motion summation could be regulated by attentional control. The summation data show that informative cuing of the patches containing the motion signals increased sensitivity (relative to not cuing) by a factor of about three. Observers could exclude the non-signal patches from the integration: cuing out the non-signal sectors had the same effect as physically removing them, as Morrone et al. (1995) did in their study. This shows that integration of motion signals does not occur simply by summing all signals within a large receptive field, but by more adaptive means. Furthermore, observers were able to choose the area to be integrated within a very brief time-frame: the cues came on with the motion stimulus, and lasted for less than 150 ms. The search results reinforce this finding, showing that when a single signal patch is cued, the uncued noise distractors can be ignored, not considered in the analysis. In the case of set-size four, for example (either summation or search), signals from four diametrically opposite patches are summed, skipping over the non-informative patches between them. This means that the visual system can pool motion signals over large discontinuous areas, supporting previous work by Verghese, Watamaniuk, McKee, and Grzywacz (1999), and agreeing with many other studies of visual search in other visual dimensions (Baldassi & Burr, 2004; Eckstein, 1998; Palmer, 1994; Palmer et al., 1993; Verghese & Stone, 1995, 1996).

4.1. Search at contrast threshold

Unlike the motion coherence measurements, contrast sensitivity for motion discrimination did not show set-size effects, unless the stimuli were clearly segmented from background. Coherence sensitivity showed strong set-size effects for both patches and regions, regardless of whether the stimuli were segregated from the background. The results for contrast sensitivity are reminiscent of

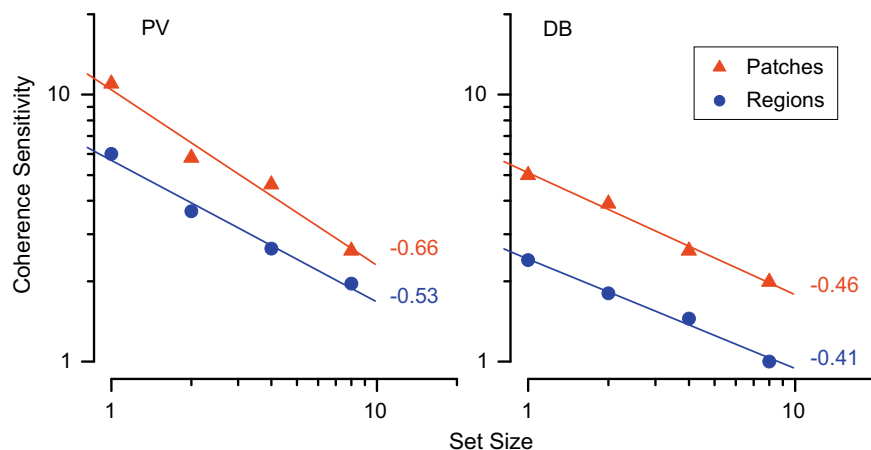


Fig. 3. Coherence sensitivity as a function of set size for two observers for the patches (triangles) and the region (circles) condition in the search experiment. Coherence sensitivity for the patches condition is about twice that of the regions condition at all set sizes measured (presumably because of the adjacent noise being integrated in the region condition). The lines are linear fits and the number report their slope for each condition, close to 0.5 in all cases.

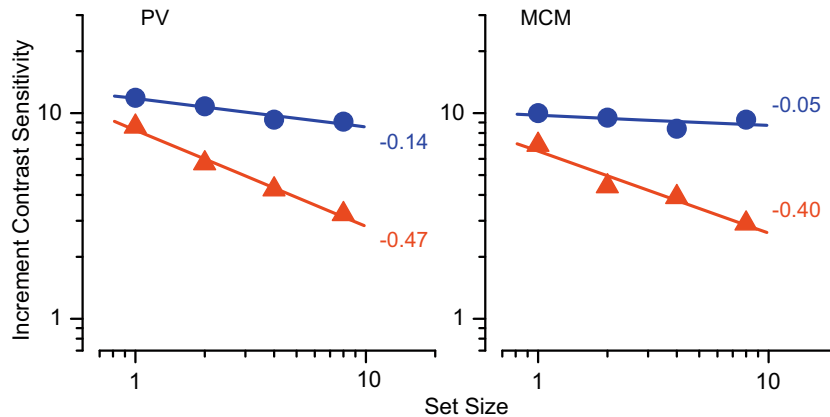


Fig. 4. Contrast sensitivity as a function of set size for two observers in the search experiment. In the regions condition (circles) the patches were not clearly defined. In the patch condition (triangles), each patch comprised 50 random dots of 30% contrast. Only the patch condition showed a set-size dependency. The lines are linear fits and the number report the slope of the set size function for each condition: the log–log slope for the patches condition was not much different from 0.5, but that for the region (no-pedestal) condition was much closer to zero.

Baldassi and Burr’s (2004) results showing no set-size dependency for luminance and colour discriminations when target and distractors merged into the background (“region” condition), but strong set-size effects when the stimuli were superimposed on clearly visible pedestals (analogous to “patches”). Comparisons between the previous and the present results are summarized in Fig. 5. For all “patch” conditions (red bars), where the stimuli and distractors were perceptually distinct from the background, there was a strong set-size dependence, in the order of 0.5 log–log slope, a roughly square-root relationship between set-size and sensitivity. This was true for contrast increment thresholds (both luminance and colour) and for contrast thresholds for motion discrimination, when the stimuli were superimposed on clearly visible noise dots. On the other hand, in the “regions” condition, when the stimuli blended into the background, the log–log slopes were virtually flat for the luminance- and colour-contrast discrimination task (Baldassi & Burr, 2004), and also for contrast thresholds for motion direction. The one exception is the motion coherence task with “re-

gion” stimuli that also blended into the background (Fig. 1B): for these stimuli the set-size effect was large, as large as for the patch condition.

The lack of set-size dependence for search contrast thresholds is reminiscent of Morrone et al.’s summation study showing little or no summation for contrast thresholds. To explain the lack of summation, the authors suggested a two-stage model where stimuli were first subject to contrast thresholding at an early stage of analysis, followed by a pooling stage, that either integrates quasi-linearly, or makes some form of decision. An alternative (but probably not incompatible) description of the lack of set-size dependency is based on the concept of high “intrinsic uncertainty” for contrast detection (Pelli, 1985). The idea is that when the position of the stimulus is not well defined (for example by a “pedestal”), the visual system must monitor many “channels” for each potential stimulus position. This effectively increases the size of the set that needs to be monitored, by a multiplicative constant (equal to the number of channels monitored at each position). The predictions of both models are shown in Fig. 7 and discussed below.

Another characteristic of the data at contrast threshold is that psychometric functions are very steep. Fig. 6 shows the average

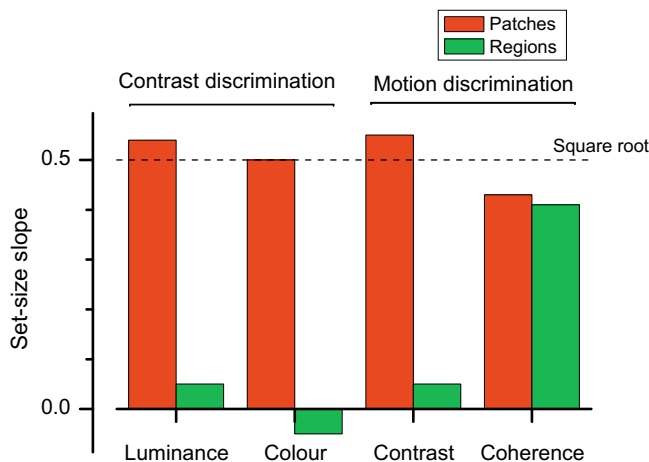


Fig. 5. Comparison of the average log–log slopes of the set size function in this and a previous search study on luminance and color discrimination in which conditions similar to the present patches and regions were used (Baldassi and Burr, 2004). In all but the motion coherence task, the functions measured on a uniform background have near-zero slope, while salient and well segregated distractors cause a set-size dependency with slope near 0.5 (square-root dependency). For motion coherence, however, there was a set-size dependency even in the “region” condition when the stimuli were not salient objects but a uniform field of random dots. This is discussed in the text.

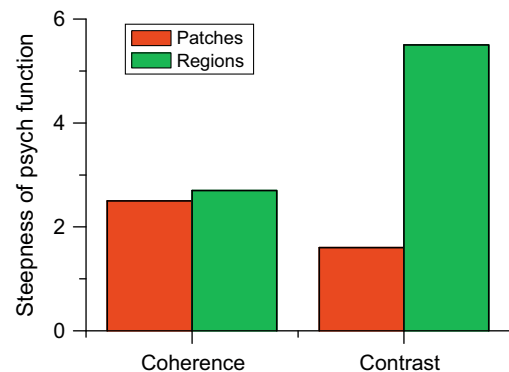


Fig. 6. Average steepness of psychometric functions for set-size 1, for the four conditions of “patches” and “regions”, motion coherence and contrast sensitivity. The data were fitted with Weibull functions.

$$P(c) = 1 - (0.5) \exp \left[\left(-\frac{c}{\alpha} \right)^\beta \right] \tag{1}$$

where P is proportion correct, c is contrast or coherence α contrast or coherence threshold (to support 81.6% correct response) and β the parameter determining steepness.

steepness of psychometric functions for the no-distractor (set-size 1) conditions (obtained by best fits of “Weibull functions”). As expected, the slopes are quite shallow (around 2) for all conditions except the contrast “regions” condition, where it was over 5. Both contrast thresholding and intrinsic uncertainty can account for this if we assume a contrast exponent of 3.5 and an intrinsic uncertainty of 50, respectively.

In practice, it is very difficult to distinguish between thresholding and uncertainty models. They both tend to explain well a great deal of psychophysical data, from the steepness of the psychomet-

ric function to the dipper function for many increment discrimination tasks (Boynton, Demb, Glover, & Heeger, 1999; Henning & Wichmann, 2007; Legge & Foley, 1980; Pelli, 1985). It may be that the two models are just different ways of looking at the same physiological phenomenon. Morgan, Chubb, and Solomon (2008) have recently suggested that this thresholding-like phenomenon may be a general mechanism designed to ensure that that neural noise does not impinge on consciousness (Ross & Burr, 2008). Uncertainty may be the mechanism by which this is achieved. However, this discussion goes well beyond the scope of this study.

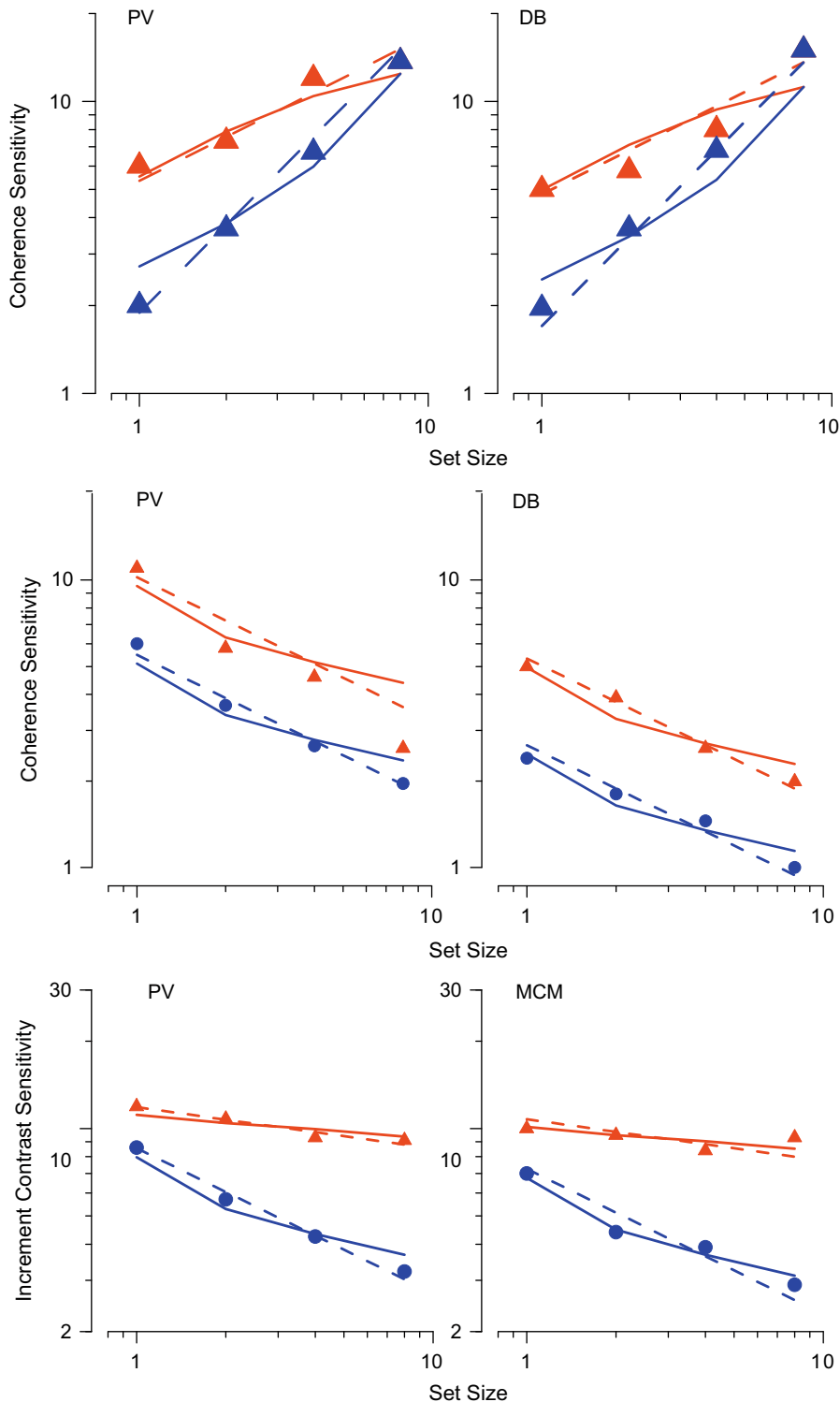


Table 1

Coefficient of determination (R^2) showing the proportion of the variability of the data explained by the two models.

$$R^2 = 1 - \frac{\sum_{i=1}^n (S_i - \hat{S}_i)^2}{\sum_{i=1}^n (S_i - \bar{S})^2} \quad (5)$$

where S_i is log sensitivity of each data point, \hat{S}_i the log sensitivity of the model at that set-size, \bar{S} the average log-sensitivity and n the number of data points (7 for the summation condition, 4 for all the others). In other words the coefficient of variation is $1 - R^2$ – the ratio of the squared residuals to the variance. Note that the low or negative values of R^2 for the no-pedestal contrast condition are because the variance of the data in that condition is small.

Condition	Observer	Signed max	Linear summation
Summation	PV	0.91	0.98
	DB	0.83	0.97
Coherence patches	PV	0.96	0.99
	DB	0.93	0.99
Coherence regions	PV	0.93	0.99
	DB	0.94	0.96
Contrast pedestal	PV	0.96	0.99
	MCM	0.98	0.92
Contrast no-pedestal	PV	0.83	0.95
	MCM	0.38	−0.46

4.2. Signal pooling: summation or decisions?

The data show clear evidence that pooling of signals is under attentional control. At a glance, the slopes of the best log–log fits agree well with ideal linear integration over the selected regions. While summing the local signals may be optimal when each local patch contains signal, it is sub-optimal when only one of the patches contains signal, and the others noise, as is the case for the search condition. An alternate possibility is to compare activity between individual regions and to base a decision only on the region with the largest response. For example, the “signed-max” model (Baldassi & Vergheze, 2002) chooses the largest (absolute) value over the regions monitored (see Baldassi & Vergheze, 2002) for details of model), and has been applied with success to many search paradigms.

The continuous curves of Fig. 7 show the predictions of the signed-max model to our data both for the summation and search paradigms, and the dashed lines show the predictions for the linear integrator. Both models have one degree of freedom: the absolute sensitivity that best fits all the data. The sensitivity for all other set sizes was set to this value. Table 1 shows the “coefficient of determination”, the proportion of variance explained by the two models in all conditions. Both do well, above 80%. For the summation data,

the fit of the linear integrator was generally better than the signed max model, consistent with the fact that that is the optimal model for the task. The search data, however, are equally well-fit by the non-optimal linear integrator and the more ideal signed-max model. Unfortunately, the predictions of these models are very close over the range of set-sizes that we examine here.

To explain the lack of set-size dependency for contrast thresholds we introduced thresholding and intrinsic uncertainty for the integration and signed-max models, respectively. For the summation model, we approximated the effect of thresholding by non-linear pooling after raising to an exponent of 3.5 (see Morrone et al., 1995 for more details). For the signed max model, we assumed that each region comprised 50 independent detectors that need be monitored. Thus the number of detectors monitored is 50 for set-size 1, 100 for set-size 2 etc. The effect of this intrinsic uncertainty is to decrease considerably the effect of set-size on visibility, and the predictions do follow the data well. Both thresholding and intrinsic uncertainty were able to simulate well the trends of the data. Note that the R^2 measures for these fits are not high (Table 1), because the variance in sensitivity measures was low under these conditions.

In a sense, the two models are extremes: one linear—combining all signals with equal weight—the other highly non-linear, choosing only the largest signal. In practice the “max” decision can be approximated by raising each local signal to a large exponent before summing. Between these two extremes lie a wide range of less severe non-linearities, such as response squaring or normalization before integration. Interestingly, a model along these lines captures many of the properties of MT neurons (Rust, Mante, Simoncelli, & Movshon, 2006).

What seems certain from this and previous studies (Burr & Santoro, 2001; Burr et al., 1998; Morrone et al., 1995) is that motion mechanisms can be considered to comprise at least two stages, one limited by a contrast threshold-like mechanism (that can also be considered in terms of intrinsic uncertainty), the other a mechanism that combines input from this first-stage. The pooling of motion signals seems to be optimal, improving sensitivity close to the amount predicted by ideal observer models. But most importantly, the pooling is flexible, under attentive control, so observers can choose which regions to include and which to exclude, and this presumably helps the visual system to reconcile the conflicting demands of integration and segmentation.

Acknowledgments

Supported by the European Commission Framework 6 “NEST” programme (MEMORY) and Framework 7 European Research

Fig. 7. Model fits to the data. The continuous curves show the predictions of the signed-max model (Baldassi and Vergheze, 2002), and the dashed lines the prediction for the linear integrator. In both cases the limits to performance were early visual noise, whose root variance is equal to the threshold at set-size = 1 (although in practice all data points are used in the model fit). The signed max predictions were given by Monte Carlo simulations (5000 repetitions). Basically for each simulation the model looked for the stimulus containing the max absolute amount of motion energy, and scored its direction. The coherence yielding 75% correct was taken as threshold. The linear integrator simply assumed addition of signals and noise variance, producing theoretical log–log slopes of +1 for the uncued summation (Eq. (2)), +0.5 for cued summation (Eq. (3)) and −0.5 for the search conditions (Eq. (4)).

$$\ln(s) = \alpha_1 + \ln(N) \quad (2)$$

$$\ln(s) = \alpha_1 + \ln(8)/2 + 0.5 \ln(N) \quad (3)$$

$$\ln(s) = \alpha_2 - 0.5 \ln(N) \quad (4)$$

where S is sensitivity, N set-size, and α_i a constant that minimizes the residual errors between model and data, corresponding to log sensitivity at set-size 1 (for Eqs. (2) and (4)). For both models all the summation data were fitted together, with the curves anchored at set-size 8, identical for both conditions. Thus the summation data were fit to 7 points, the search to 4 data points. For both models the absolute sensitivity of the curves was adjusted to minimize the squared residuals of the logarithms of sensitivities. Table 1 shows the amount of variance explained by the two models in all conditions. For contrast threshold data (with no pedestal) we introduced thresholding and intrinsic uncertainty for the summation and signed-max models respectively. For the summation model, we approximated the effect of thresholding by non-linear pooling after raising to an exponent of 3.5 (see Morrone et al. 1995 for more details). For the signed max model, we assumed that each region comprised 50 independent detectors that need be monitored. Thus the number of detectors monitored is 50 for set-size 1, 100 for set-size 2 etc. The effect of this intrinsic uncertainty is to decrease considerably the effect of set-size on visibility, so the curve follows well the data.

Council Advanced Grant 229445 “STANIB” and by NASA grant NAG-2-1202 and NSF grant 0642728 to PV.

References

- Baldassi, S., & Burr, D. C. (2004). “Pop-out” of targets modulated in luminance or colour: The effect of intrinsic and extrinsic uncertainty. *Vision Research*, *44*(12), 1227–1233.
- Baldassi, S., & Verghese, P. (2002). Comparing integration rules in visual search. *Journal of vision*, *2*(8), 559–570.
- Boynton, G. M., Demb, J. B., Glover, G. H., & Heeger, D. J. (1999). Neuronal basis of contrast discrimination. *Vision Research*, *39*(2), 257–269.
- Braddick, O. (1993). Segmentation versus integration in visual motion processing. *Trends in Neurosciences*, *16*(7), 263–268.
- 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.
- Burr, D. C., & Santoro, L. (2001). Temporal integration of optic flow, measured by contrast and coherence thresholds. *Vision Research*, *41*(15), 1891–1899.
- Duffy, C. J., & Wurtz, R. H. (1991). Sensitivity of MST neurons to optic flow stimuli. I. A continuum of response selectivity to large-field stimuli. *Journal of Neurophysiology*, *65*(6), 1329–1345.
- Eckstein, M. (1998). The lower visual search efficiency for conjunction is due to noise and not serial attentional processing. *Psychological Science*, *9*, 111–118.
- Graham, N. (1977). Visual detection of aperiodic spatial stimuli by probability summation among narrowband channels. *Vision Research*, *17*(5), 637–652.
- Grindley, G. C., & Townsend, V. (1968). Voluntary attention in peripheral vision and its effects on acuity and differential thresholds. *Quarterly Journal of Experimental Psychology*, *20*, 11–19.
- Henning, G. B., & Wichmann, F. A. (2007). Some observations on the pedestal effect. *Journal of Vision*, *7*(1), 3.
- Legge, G. E., & Foley, J. M. (1980). Contrast masking in human vision. *Journal of the Optical Society of America*, *70*(12), 1458–1471.
- Morgan, M. J., Chubb, C., & Solomon, J. (2008). A ‘Dipper’ function for texture discrimination based on orientation variance. *Journal of Vision*, *8*(11), 1–8. 9.
- Morgan, M. J., Ward, R. M., & Castet, E. (1998). Visual search for a tilted target: Tests of spatial uncertainty models. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*, *51*(2), 347–370.
- Morrone, M. C., Burr, D. C., & Vaina, L. M. (1995). Two stages of visual processing for radial and circular motion. *Nature*, *376*(6540), 507–509.
- Orban, G. A., Lagae, L., Verri, A., Raiguel, S., Xiao, D., Maes, H., et al. (1992). First-order analysis of optical flow in monkey brain. *Proceedings of the National Academy of Sciences of the United States of America*, *89*(7), 2595–2599.
- Palmer, J. (1994). Set-size effects in visual search: The effect of attention is independent of the stimulus for simple tasks. *Vision Research*, *34*(13), 1703–1721.
- Palmer, J., Ames, C. T., & Lindsey, D. T. (1993). Measuring the effect of attention on simple visual search. *Journal of experimental psychology. Human perception and performance*, *19*(1), 108–130.
- Pelli, D. G. (1985). Uncertainty explains many aspects of visual contrast detection and discrimination. *Journal of the Optical Society of America A*, *2*(9), 1508–1532.
- Ross, J., & Burr, D. (2008). The knowing visual self. *Trends in Cognitive Sciences*, *12*(10), 363–364.
- Rust, N. C., Mante, V., Simoncelli, E. P., & Movshon, J. A. (2006). How MT cells analyze the motion of visual patterns. *Nature Neuroscience*, *9*(11), 1421–1431.
- Sclar, G., Maunsell, J. H., & Lennie, P. (1990). Coding of image contrast in central visual pathways of the macaque monkey. *Vision Research*, *30*(1), 1–10.
- Tanaka, K., & Saito, H. (1989). Analysis of motion of the visual field by direction, expansion/contraction, and rotation cells clustered in the dorsal part of the medial superior temporal area of the macaque monkey. *Journal of Neurophysiology*, *62*(3), 626–641.
- Verghese, P., & Stone, L. S. (1995). Combining speed information across space. *Vision Research*, *35*(20), 2811–2823.
- Verghese, P., & Stone, L. S. (1996). Perceived visual speed constrained by image segmentation. *Nature*, *381*(6578), 161–163.
- 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.
- Watson, A. B., & Pelli, D. G. (1983). QUEST: A Bayesian adaptive psychometric method. *Perception & Psychophysics*, *33*(2), 113–120.
- Yang, Y., & Blake, R. (1994). Broad tuning for spatial frequency of neural mechanisms underlying visual perception of coherent motion. *Nature*, *371*(6500), 793–796.