Sensitivity for global shape detection

Rebecca L. Achtman
McGill Vision Research, Department of Ophthalmology, Montreal, Quebec, Canada

Robert F. Hess
McGill Vision Research, Department of Ophthalmology, Montreal, Quebec, Canada

Yi-Zhong Wang
Retina Foundation of the Southwest, Dallas, Texas, USA

In order to understand the nature of the mechanisms responsible for global shape detection, we measured coherence thresholds in a 2IFC task where subjects judged which of two arrays of Gabors contained global circular structure. The stimulus was composed of an array of oriented Gabor patches positioned on a polar grid. Thresholds were obtained for different array parameters (e.g. different area, density, number and positions of elements) as well as for different element parameters (e.g. different carrier spatial frequencies, contrasts, polarities and orientations). Global structure was detected when ~10% of the elements were coherently oriented. Neither the properties of the array (density, area, number or position of elements), nor those of the individual elements (carrier spatial frequency, contrast, polarity) altered coherence thresholds. Varying contrast or carrier spatial frequency within individual arrays also did not alter performance. Sensitivity was invariant to positional perturbations of the array grid. Only jittering the local orientation of elements decreased sensitivity. The underlying mechanisms are broadly tuned for contrast, spatial frequency and the spatial positioning of image samples. Detecting circular structure is a robust process and, in this case, a purely global one. Sensitivity was highest for circular as opposed to radial or spiral shapes.

Keywords: circularity, local, global, orientation, shape

Introduction

The visual system is seamlessly able to detect, discriminate and localize objects, with one of the key aspects of this shape processing being the binding of distributed local features into a whole. The visual system can be thought of as being composed of detectors or filters specialized for particular spatial frequencies, orientations, contrasts, polarities and positions. Shape detection requires the selective integration of the outputs of several such filters across one or more of these processing dimensions.

An important question in this area concerns how the visual system extracts shape information. It is now well established that the primary visual cortex (V1) extracts information about the local orientation, spatial frequency and polarity of contours and edges (Hubel & Wiesel, 1962). Receptive fields of V1 neurons can be thought of as discrete and localized linear filters – each one seeing only a small part of the visual scene. However, there is evidence supporting the idea that lower cortical areas, through facilitative and suppressive interactions, are also involved in global processing (see Allman, Miezin & McGuinness, 1985; Gilbert, 1992; Fitzpatrick, 2000). More complex global processing of visual stimuli is also known to occur further along the visual cortical pathway. Single-unit electrophysiology experiments indicate that neurons at the highest level of the ventral processing stream in the inferotemporal cortex (IT) respond specifically to very complex stimuli and are tuned for highly complex patterns such as faces (Gross, 1992). In these experiments we ask: What local information is pooled, and how is it combined, to enable us to detect global shapes?

Shape detection in general, and global shape detection in particular, comes in many forms. Even limited to the detection of circular structure, there is still a range of different tasks that the visual system must solve which might necessitate different processing solutions. For example, radial frequency (RF) patterns have been used to examine how we perceive deviations from circularity (Wilkinson, Wilson & Habak, 1998; Hess, Wang & Dakin 1999; Achtman, Hess & Wang, 2000). The unmodulated, or perfectly circular stimulus contains energy at all orientations and the slightly deformed stimulus contains altered energy at some orientations. Observers’ excellent sensitivity to these deformations cannot be explained solely by local orientation or curvature analysis (Wilkinson et al., 1998; Hess et al., 1999), but point instead to the global pooling of orientation information at specific locations in the field. Figure 1A illustrates how the first-stage inputs of a putative radial frequency detector might be arranged.

Another example of global curvature detection involves contour integration, where contours are defined by the alignment of Gabor patches embedded in a noise background field composed of similar, but randomly positioned and oriented elements. Field, Hayes and Hess...
Achtman, Hess, & Wang (1993) demonstrated that the detection of contour fragments involves facilitative interactions between neighbouring cells whose orientation preferences conform to simple first-order curves. Figure 1B illustrates the form of the local interactions that have been proposed to underlie performance for contour detection. This task is global in the sense that global parameters of the contour, such as whether it is smooth or jagged, affect its ability to be detected (Pettet, McKee & Grzywacz, 1998). In these two global curvature tasks, RF pattern detection and contour integration, there is a conjoint dependence on orientation and (relative) position.

Glass pattern detection is an altogether different task involving the global detection of circular structure (Glass, 1969). The detection of Glass patterns provides an example of circularity detection in which orientations are pooled without regard for their relative positions. These stimuli are generated by taking a random set of dots and superimposing on it a geometrically transformed copy of itself. These textures are composed of dot pairs locally aligned in the direction of the geometric transformation used to generate them and are perceived as having compelling global spatial structure. The detection of structure in these patterns requires the observer to first locally group the dot pairs and then combine the local orientation information into a global form (Dakin, 1997). While these stimuli require observers to integrate orientation information across space, this can only be done after the initial grouping stage. Figure 1C illustrates the possible layout of the subunits of a putative Glass pattern detector.

We are primarily interested in whether the mechanisms underlyng the detection of globally-distributed circular structure are fundamentally different from those that have been proposed to detect circular structure in the radial frequency pattern or contour fragment cases outlined above (depicted in Figures 1A & B). We have devised a new global form stimulus that involves the detection of globally distributed circular structure like that found in Glass patterns. However, as our stimulus is composed of Gabor elements and not dot pairs, we effectively eliminate the need for the initial local grouping stage. The stimulus is composed of an array of spatial frequency narrowband, one-dimensional, luminance-defined Gabors, a variable proportion of which (corresponding to the coherence threshold) have their local orientations aligned along tangents of notional circles of different radii centred on the screen. We investigate the role of a number of key parameters for the elements comprising this global task. These include carrier spatial frequency, position, orientation, polarity and contrast. The second issue we address is whether some globally distributed shapes (e.g. transformations involving rotation or expansion) are more detectable than others. This bears upon the relative sensitivity of detectors tuned to different global shapes, which is an issue of recent debate in the Glass pattern literature (Wilson, 1998; Wilson, Wilkinson & Asaad, 1997; Dakin & Bex, 2002).

Figure 1. Schematics of putative shape detectors. (A) Radial frequency detector with orientation filtering followed by a rectification stage, along the lines of that proposed by Wilkinson et al. (1998). The spatial positions of the pooled or summated inputs are specified. (B) The association field model originally proposed by Field et al. (1993) represents the specific rules of alignment by which elements in a contour are associated. Grouping of neighbouring elements occurs when the elements conform to simple first-order curves. Elements with alignment like those shown on the left will be ‘associated’, unlike those on the right whose orientations do not conform to the rays extending from the centre element. The models depicted in A & B show joint constraints on position and orientation. (C) Possible layout of the subunits of a Glass pattern detector where the relative position of the initial filters are unimportant as long as the orientation is tangential to a circle centred on the pattern.
Methods

Stimuli

Stimuli were composed of a polar array of oriented Gabors. In the target, some Gabors were oriented tangential to the circumference of circles centred on the pattern (see Figure 2). The global structure in the stimulus is defined by both the position and orientation of the individual Gabor patches. The advantage of these stimuli is that they are spatially narrow-band and well suited to test the integration of the 1st stage filters.

Stimuli were generated digitally in MATLAB (MathWorks, Inc.) and displayed on a gamma-corrected, ViewSonic gray-scale monitor using the Psychophysics Toolbox (Brainard, 1997) that provides high-level access to the C-language VideoToolbox (Pelli, 1997). The mean luminance of the monitor was 92 cd/m². The stimulus screen subtended 22° x 17° at the viewing distance of 1 m.

Psychophysical Procedures

We used a 2-interval forced-choice, method of constant stimuli paradigm where observers were asked to judge which of 2 arrays of Gabors contained circular structure. All Gabor elements were presented on a polar grid. One interval in a trial contained an array of Gabors with random orientations, and the other was an array containing circular structure as defined by the orientation of a subset (the coherence parameter) of the Gabor elements. The duration of each stimulus presentation was 0.5 s. Each session consisted of ten trials for each of five coherence levels with a step size of 1 dB. Audio signals were used to prompt the subject just before and after each trial, but no feedback about the correctness of responses was provided. The resulting data were fit by a Weibull function (Weibull, 1951) using a maximum-likelihood procedure. Coherence thresholds corresponding to 75% correct were interpolated from the Weibull fits. Means and standard deviations were obtained from multiple estimates (typically 5).

We varied the signal-to-noise ratio to determine coherence thresholds for these patterns by measuring the number, or proportion of elements that need to be coherently aligned for correct detection of the interval containing the global form (circular structure). Figure 2J is an example of 100% coherence stimulus, Figure 2C, a 50% coherence stimulus, Figure 2B, a 10% coherence stimulus and Figure 2A a 0% coherence stimulus. Unless otherwise noted, all stimuli were presented on a polar grid for the middle sized area and density tested (array area = 7.5° x 7.5°; element separation = 0.75°; number of elements = -107) with Gabors at 80% contrast (carrier spatial frequency = 6 cpd, circular envelope SD = 0.1 deg, Gabors in sine phase).

Array Parameters

Figures 2D – F are examples of stimuli used to investigate the effect of array parameters.

Area/Density/Number Experiment: We used arrays of different areas while allowing the density and number of elements to covary to determine if one of these variables was essential to the task. The overall area of the array subtended a visual angle of either 5° x 5°, 7.5° x 7.5°, or 10° x 10°. We tested three different element densities for each area with the number of elements in an array ranging from 24 (smallest area, lowest density) to 411 (largest area, highest density). Figure 2D is an example of a low-density stimulus. Spatial Scaling Experiment: Observers performed the task at different viewing distances ranging from 0.25 – 3 m to determine the effect of scaling while keeping the geometry of the stimulus the same. Position Jitter Experiment: Position jitter was introduced into the regular grid on which the elements were positioned. In both intervals, the amount of jitter (Gabor separation) followed a Gaussian distribution where the standard deviation was specified, ranging from 0 (no position jitter), to a position jitter of 1 (where elements could overlap). The orientations of the individual elements were assigned after their positions were jittered so that the orientation of a signal Gabor was tangent to a circle centred on the pattern. Examples of 0% and 50% coherent, position jittered stimuli (SD = 0.7 patch separation) are shown in Figures 2E & F, respectively.

Element Parameters

Contrast Experiment: We used different contrasts – having all Gabors within an array at the same contrast (low = 10% or high = 80%), or randomizing the contrasts of elements within an array (element contrasts were randomly drawn from a rectangular distribution ranging from 10 – 80%). An example of this latter condition is shown in Figure 2G. Spatial Frequency Experiment: We investigated the effect of different carrier spatial frequencies. Again, we did this by either varying the absolute spatial frequency of all the elements of the array (low s.f. = 3 cpd; high s.f. = 12 cpd), or by randomizing the spatial frequency across elements within the array (mean s.f. of 6 cpd with a SD of 2 cpd for Gaussian jittering). An example of this stimulus is shown in Figure 2H. Polarity Experiment: We looked at the effect of reversing the polarity of neighboring elements. Orientation Jitter Experiment: We also jittered the orientations of the individual Gabors defining the circular structure. An example of a 50% coherence stimulus with orientation jitter (Gaussian distribution, SD = 15°) is shown in Figure 2I.
Figure 2. Examples of the Gabor arrays used in the experiment: (A) 0% coherent stimulus – a purely random pattern; (B) 10% coherent stimulus – at our observers’ discrimination threshold; (C) 50% coherent stimulus – more circular structure can be seen in the image; (D) 50% coherent low density stimulus; (E) 0% coherent stimulus where the positions of the individual elements have been jittered with a SD = 0.7 patch separation; (F) 50% coherent stimulus with the same amount of position jitter as in E; (G) 50% coherent stimulus where the individual elements are of random contrasts; (H) 50% coherent stimulus where the individual elements are of varying spatial frequencies; (I) 50% coherent stimulus where the orientation of the individual elements has been jittered with a SD of 15°; (J) 100% coherent stimulus with a circular shape; (K) 100% coherent stimulus where the individual elements align to form a radial pattern; (L) 100% coherent stimulus where the individual elements align to form a spiral pattern.
Shape Parameters

Target Shape Experiment: Finally, we compared sensitivity for different target shapes, namely circular, radial and spiral. Examples of these stimuli at 100% coherence are shown in Figures 2J, K and L, respectively.

Subjects

The subjects (one of the authors and two naive observers) had corrected-to-normal vision and were experienced at the task.

Results

Dependence on Array Parameters

Coherence thresholds for all observers in the midrange of the parameter space that we tested (array area = 7.5° × 7.5°; element separation = 0.75°; number of elements = ~107; element contrast = 80%) were ~7-10%. In general, we found that threshold sensitivity was remarkably invariant to most of our stimulus manipulations. For example, in our experimental manipulation of area/density/number, we found no differential effect of using arrays of different sizes, numbers or densities (data not displayed). Observers needed roughly the same proportion of elements to be aligned in the array whether the array contained 24 or 411 elements independent of its density. We found that the task exhibited scale invariance since thresholds were the same at a range of different viewing distances. This result is shown in Figure 3A where coherence thresholds, in proportion of elements aligned, are plotted against the viewing distance in metres. Initially, measurements were made with the elements positioned on a polar grid, however, later experiments revealed similar thresholds for a rectangular grid arrangement. Indeed, when we specifically jittered the individual element positions within the grid (as in the position jitter experiment), coherence threshold was unaffected. These results are displayed in Figure 3B where coherence threshold is plotted against the standard deviation of the Gaussian positional jitter in units of Gabor patch separation. The results for 3 observers show threshold invariance with positions of the Gabors within the array.

Dependence on Element Parameters

We varied the contrast, spatial frequency, polarity and orientation of array elements. In some cases these were varied across the array as a whole, whereas in other cases we introduced the variation within the array. In the experiment where contrast was manipulated, coherence thresholds remained constant at both low and high contrasts tested, as well as when we used random contrasts within an array. The results are shown in Figure 3C, where coherence thresholds, plotted for two observers, demonstrate that mechanisms used to detect these patterns are broadly tuned for contrast, both within and across arrays. In the experiment where element spatial frequency was manipulated, we similarly found no influence on coherence threshold of either absolute or relative element spatial frequency. Figure 3D displays these results for three observers for the condition where spatial frequency changed across and within the array. Observers were able to integrate information across space from Gabors with a range of different spatial frequencies. In our polarity experiment, we found that alternating the polarity of neighboring elements had no effect on detection thresholds (data not shown). Finally, the results of the orientation jitter experiment, which illustrate the importance of individual element orientation for this task, were to some extent expected. We measured coherence thresholds as a function of the amount of orientation jitter added to the individual elements. Results displayed in Figure 3E, where coherence threshold is plotted against the standard deviation of the orientation jitter in degrees, show that performance remains unaltered over a certain range of orientation jitter. However, coherence thresholds do increase dramatically when the SD of the jitter is greater than ~15°. A performance ceiling is reached at around 35 – 40° of orientation jitter. While the detrimental influence of large jitters was not unexpected (Levi & Klein, 2000; Saarinen & Levi, 2001) we were surprised to find that threshold remained invariant for small to medium jitters.

Dependence on Global Shape

The second issue that we investigated concerned global shape and in particular whether sensitivity is higher for some shapes. The stimuli were constructed the same way as before, however the target structure was circular, radial or spiral. Trials for these patterns were run separately with observers being aware of the target pattern. For 3 observers the coherence thresholds (i.e. the proportion of elements coherently aligned to the target pattern) are plotted for each of the three target shapes (circles, radials and spirals) (Figure 3F). These results demonstrate that observers are most sensitive to circles, then radial patterns, and finally spirals.

To control for the possibility that the presentation of elements on a polar grid biases these results, we performed the position jitter experiment on two subjects for the three shapes and the same position jitter conditions we used earlier (proportion of jitter range: 0-1, in 0.1 step increments). The results (not shown) indicate that there is no differential effect of jitter for detecting the different shapes. That is, the positioning of the elements on a polar grid is not what contributed to lower detection thresholds for circular structure. To do this task, observers pool the oriented carrier information from positions that are unrelated to the overall structure to be detected.
Figure 3. Results of the various experimental conditions: (A) Viewing distance experiment: Coherence thresholds for two observers at viewing distances ranging from 0.25 to 3 m. There is no effect of viewing distance on coherence thresholds; (B) Position jitter experiment: Coherence threshold in proportion of elements aligned to the target pattern plotted as a function of position jitter for three observers. We specified the standard deviation of the position jitter in units of patch separation, ranging from 0 (no jitter) to 1 (possibility of elements overlapping completely). There is no effect of position jitter on coherence thresholds; (C) Contrast experiment: Coherence thresholds for two observers are plotted for three contrast conditions (low contrast = 10%; high contrast = 80%; mixed = random contrasts between 10 - 80% within one array – e.g. Figure 2G); (D) Spatial frequency experiment: Coherence thresholds for three observers are plotted for three carrier spatial frequency (s.f.) conditions (low s.f. = 3 cpd; high s.f. = 12 cpd; mixed = mean s.f. of 6 cpd with SD = 2 cpd of Gaussian jittering – e.g. Figure 2H); (E) Orientation jitter experiment: Coherence thresholds are plotted as a function of orientation jitter for three observers. Orientation jitter is measured as the standard deviation of the carrier orientation in degrees and ranges from 0 (no jitter) to an orientation jitter of 50°. At about 35 – 40° observers can no longer do the task of detecting circular structure regardless of how many Gabor elements are aligned; (F) Target shape experiment: Coherence thresholds for three observers for 3 different target shapes (circular, radial and spiral). The pattern of results with observers being most sensitive to detecting circular structure, then radial and finally spiral structure holds for all observers. All error bars represent one standard deviation.

Discussion

We use a novel global spatial task in which orientation signals are clearly defined and the only constraints are at the level of the global integration process rather than at the level of the initial matching process as in Glass patterns. This spatial task is, in general, similar to a well-accepted global motion task (Morgan & Ward, 1980; Newsome & Pare, 1988) and in particular, equivalent to the stochastic Gabor display of Baker and Hess (1998). While absolute thresholds for this global spatial stimulus are similar to that for its motion counterpart at about 7 - 10% coherence, the spatial task differs dramatically from its motion counterpart in that it exhibits little dependence on various array and element parameters. In the spatial case, coherence threshold was invariant with the area/density/number of elements of
Contrast variation does not affect global threshold frequency patterns (Hess, 1999). The shape to be detected. This is not the case for the insensitive to the position of the elements representing present task, within the range of parameters used, is contour fragments (depicted in Figures 1A & B). The present task, within the range of parameters used, is insensitive to the position of the elements representing the shape to be detected. This is not the case for the discrimination of subtle changes in circularity (Keeble & Hess, 1999) or for the detection of Gabor-sampled radial frequency patterns (Wang & Hess, 2003). In each of these tasks, spatial position plays a strong role and the type of integration is quite different because no noise elements are involved. Our present task is also unlike tasks involving contour detection of strings of aligned Gabors (Hess, McIlhagga & Field, 1997) where the signal and noise elements are spatially segregated. This is suggested by the insensitivity of the present task to element polarity, carrier spatial frequency and element configuration. Nevertheless, we did wonder whether contour integration mechanisms, which are thought to detect small, co-aligned strings of elements that may occur by chance, could play a role in determining global sensitivity for this task. Although this seemed unlikely to be the case with coherence thresholds at 7 - 10%, we directly assessed this by ensuring, in a separate control experiment, that signal elements were evenly distributed within and across the notional circles that they represented. This ensured that there was no clumping of signal elements. We found identical threshold sensitivity for this modified display and concluded that the mechanism underlying this global task and that underlying contour integration are different.

At this point we are unsure of whether the mechanism underlying the detection of Glass patterns is similar to that underlying detection in our global task. The two tasks share a common broad tuning for spatial frequency at the level of the global matching stage, although in the Glass pattern case it is low-pass, not spatial frequency invariant (Dakin & Bex, 2001). Contrast variation does not affect global threshold sensitivity in our task, but does affect the global matching stage of Glass patterns (Dakin & Bex, 2001, Wilson, Switkes & De Valois, 2001); however both tasks share the important distinguishing feature of being insensitive to the spatial layout of the individual local signals. Glass pattern stimuli are usually very dense, containing a large number of element pairs. It would be interesting to know whether coherence thresholds in Glass patterns vary systematically with the number of dot pairs/density/size of the array as we have shown for our task.

The second question that we addressed is whether some purely global shapes are preferred to others. This too has its counterpart in the global motion literature where there is evidence that some forms of optic flow are more detectable than others and that the mechanisms underlying their detection may be different (Morrone, Burr, DiPietro & Stefanelli, 1999). Here we compared thresholds for circular, radial and spiral global shapes. Thresholds were lowest for circular shapes and highest for spiral shapes. It is hard to compare this finding to similar previous comparisons for Glass patterns because there is still some debate concerning whether an observer’s sensitivity depends on the particular transformation involved (Wilson & Wilkinson, 1998; Dakin & Bex, 2002). Since our stimuli were displayed within a square window, the lower thresholds for circular shapes that we find are at odds with those reported for Glass patterns (Dakin & Bex, 2002) and may represent another difference between the present task and its nearby relative, the Glass pattern. However, this difference may be related to the initial matching stage in Glass patterns, as opposed to the later global integration stage common to both tasks.

Although this seemed unlikely to be the case with coherence thresholds at 7 - 10%, we directly assessed this by ensuring, in a separate control experiment, that signal elements were evenly distributed within and across the notional circles that they represented. This ensured that there was no clumping of signal elements. We found identical threshold sensitivity for this modified display and concluded that the mechanism underlying this global task and that underlying contour integration are different.

Site of Mechanism

The purely global nature of this task suggests a detector with subunits that are randomly, but widely spatially distributed. Receptive fields in higher regions of the extra-striate pathway are known to have very large receptive fields that integrate information from multiple, widely distributed subunits, as exemplified in the case of translational motion in area MT (Britten, Shadlen, Newsome & Movshon, 1993) and optic flow (Burr, Morrone & Vaina, 1998) in MSTd (Tanaka & Saito, 1989; Orban, Lagae, Raiguel, Xiao & Maes, 1995). We expect that the mechanism underlying performance for
this task would be located in extra-striate cortex, as shown in a recent fMRI study (Braddick, O'Brien, Wattam-Bell, Atkinson & Turner, 2000) using a similar, though not identical, stimulus. Although we cannot rule out the possibility that multiple areas are involved to provide the generic form of integration that typified our task (Kourtzi, Tolas, Altmann, Augath & Logothetis, 2003).

Acknowledgment

This work was supported by a grant from CIHR (MOP-10818) to RFH. Commercial relationships: none.

References


