ATTENTIONAL RESOLUTION: THE GRAIN AND LOCUS OF VISUAL AWARENESS

PATRICK CAVANAGH*, SHENG HE° and JAMES INTRILIGATOR§
*Psychology , Harvard University, 33 Kirkland Street, Cambridge, MA 02138, USA
°Psychology , University of Minnesota, Minneapolis, MN, USA
§Neurology, Beth Israel Hospital, Boston, MA, USA

ABSTRACT

Attention is the gateway to visual awareness and it imposes the final limit on what we can consciously experience in our visual world. Our work shows that attention is not limited to a single focus but can have complex spatial distributions with many of the aspects of a spatial image. We have proceeded to evaluate the grain of this image and discovered that the smallest regions which can be isolated by attention are surprisingly coarse. Objects spaced more finely than the limit of attentional resolution can not be individuated for further processing and can only be perceived as a grouped texture. As a result, only part of the spatial and temporal information registered by the early sensory systems is available to conscious perception.. The properties of attentional resolution suggest that the locus of attentional selection is at a stage beyond primary visual cortex.

1. Introduction

We normally pay attention to a visual event, say the appearance of a celebrity within our field of view, by directing our eyes so that the area of interest falls at the center of gaze. The density of receptors is highest at this point so this movement of the eyes brings a great improvement in the resources available for the analysis of the event. However, we can also increase our processing capacity by deploying attention to the event, without moving the eyes — something which can be very important in some social contexts or in situations where there is more than one event of interest. When we "pay attention" to an event we are rewarded with extra detail and richness, but often with the cost of losing awareness of other surrounding events.

One of the most enduring metaphors for attention is that of a "spotlight" (Eriksen & Hoffman, 1972) illuminating a small region of visual space where details are more easily read out — to the detriment of surrounding areas. Some have measured the shape of the local area enjoying the advantages of attention and suggested that it is small but grows with distance from the center of gaze and has a tear-drop shape oriented along the radial lines from fixation (Downing, 1988; LaBerge, 1983)

However, the spatial specificity and the unitary nature of the attentional focus have been strongly challenged recently. A number of experiments have shown that the attention can be directed to noncontiguous locations (Castiello & Umilta, 1992; Kramer & Hahn, 1995) and even used to track as many as four or five randomly moving items simultaneously (Pylyshyn & Storm, 1988). At the same time, other studies have demonstrated that attention is attached to objects, not locations (Driver & Baylis, 1989; Duncan, 1984). When two objects are superimposed but attention is directed to only one of them, attentional facilitation is found to be specific to the object and not its location.

In this paper, we propose that attention is more like an image than a single spotlight and that it has image-like properties of resolution. Figure 1 depicts this description for attention where it is shown as an auxiliary representation which notes the positions of items of interest in the scene. In our second paper in this program, we will go beyond this description to propose that attention constructs the visual world of which we are aware. The resolution properties for attention described in this paper here also hold for the more elaborate model of attention and awareness sketch out in the second paper.

We begin by demonstrating that the grain of attention is coarser than the grain of visual resolution. Moreover, we show that the properties of adaptation to stimuli which cannot be resolved by attention place the substrate of attention beyond the primary visual cortex. Finally, we report some preliminary measures of the resolution of attention as a function of location in the visual field.

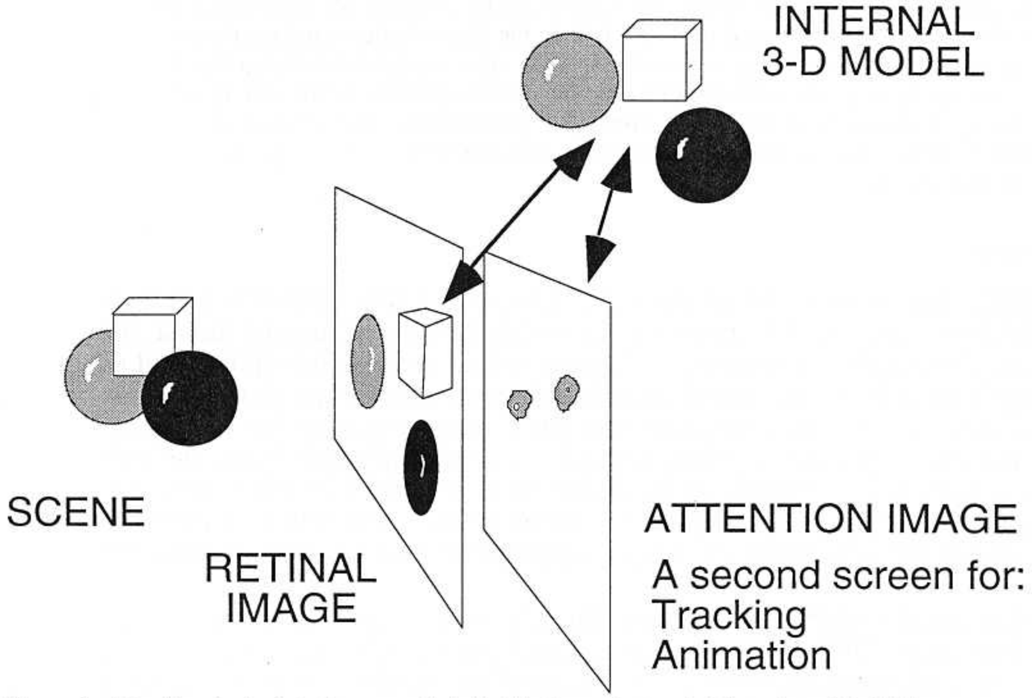

Figure 1. Attention is depicted as a spatial distribution of attended locations. Multiple events can be tracked in this attention image but there is a limiting spatial resolution, the highest density of items for which individual items can continue to be monitored.

2. The grain of attention

Imagine there are two vertical lines presented to the left of a fixation point. An observer can easily attend to one or the other when the two lines are far apart (no eye movements allowed). As the two lines get closer and closer, however, it will become increasingly difficult to focus attention on only a single one — to individuate it. We will discuss how closely spaced items can be and still be isolated by attention. As we will see, the price of being too close is to lose access to the features of individual item — they seem to get mixed up with those of its neighbors.

Surprisingly, this appears to happen well before the two items start to fuse visually and appear as a single line.

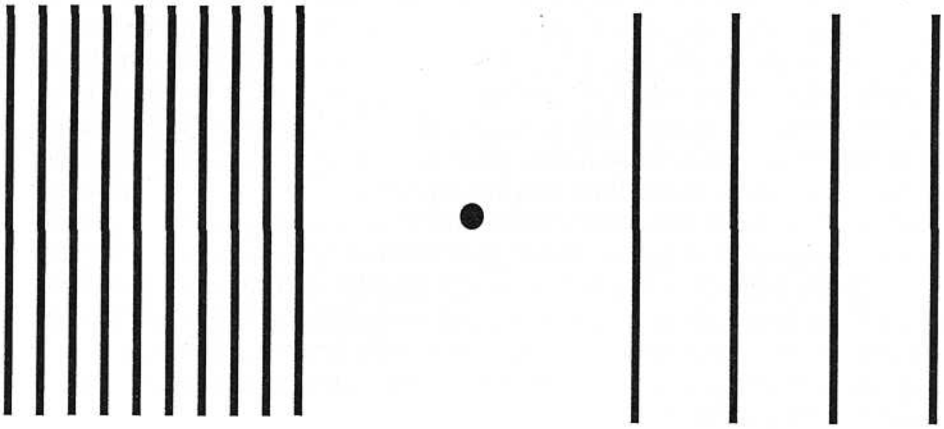

Figure 2. Fixate the central dot and attempt to pay attention to each line in turn. A simple way to do this is to try to count them or step through them one at a time. On the left, it is very difficult to isolate individual bars. We say that the density of this display exceeds the resolution limit of attention. On the right, however, it is easier and so we claim that this display is within the resolution limits of attention.

Early studies using letters with flanking masks addressed exactly this question. Eriksen and his colleagues (Eriksen & Hoffman, 1972) proposed that the attention spotlight has the size of 1 degree, which means that if two items are spaced less than this distance, they can no longer be accessed individually by attention. Usai, Umilta and Nicoletti (1995) found that the benefits of attention spread into gaps intruding into attended objects if the gap was small enough.

Since we will compare attentional resolution to visual resolution, let us begin with a simple description of visual resolution — the finest spacing of visual detail that can be seen. Visual resolution is conventionally measured as the finest sinusoidal grating that can be seen when presented at maximum (100%) contrast (Campbell & Gubisch, 1966). A observer can normally resolve up to 55 cycles per degree (equivalent to about 110 lines alternating between light and dark covering your index fingernail held at arm's length) under bright illumination, at the center of gaze. "Resolve" in this context means either that the observer can tell the grating patch apart from a uniform field (detection task), or can tell its orientation (discrimination task).

We can define attentional resolution in a similar manner. What is the finest spacing of the bars of a grating at maximum contrast that allows the observer to index each bar. A simple test of this ability to index each bar is to attempt to count them. Figure 2 demonstrates the difference between the conventional visual resolution and the attentional resolution. While fixating the central dot, we can clearly see the grating on the left and report that there are several fine bars vertically oriented. However, it is much more difficult to individuate and count the bars on the left (again while fixating the central dot). We resolve the grating visually as a texture but we cannot access the individual elements. In contrast on the right, the bars can be accessed individually, counted, and inspected.

3. Studies on lateral inhibition or crowding

In the grating above, the cost of not being able to individuate the bars is relatively small — we do not know how many there are but we know they are all vertical and thin. If each bar were subtly different, however, we might not be able to report the features of individual bars. For example, in Figure 3 below, we know while looking at the plus sign on the left that there are several disks arrayed out to the right. It is hard, however, to individuate the third and fourth disk and to report the orientation of the bars within these disks. Note that the outermost disk is easier to isolate and the orientation of its bars can be reported. This effect of adjacent elements has been addressed in the extensive literature on crowding tested typically with multiple letters presented in a row, either as a word or a non-word letter string (Bouma, 1970). As in Figure 3, the items in the middle of the array are more difficult to report than those close to fixation, and surprisingly, more difficult to report than the item at the outermost position, even with prolonged viewing time (Townsend, Taylor, & Brown, 1971). Clearly distance from the center of gaze is not sufficient in explaining this result.

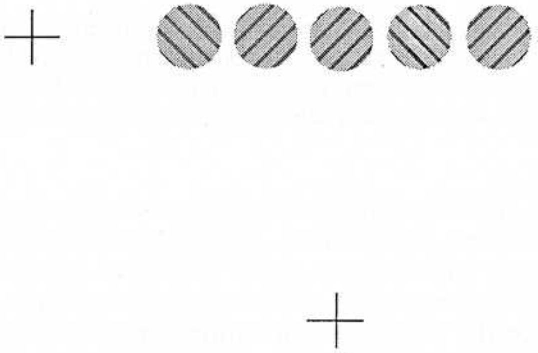

Figure 3. Fixate the plus sign on the left and examine each disk in turn, attempting to identify whether the bars within each disk are titled to the left or right. The first and second disks are relatively easy to scrutinize. The third and fourth are more difficult. And yet, the outermost disk is relatively easier again, arguing against eccentricity as the sole explanation for the difficulty in attending to the third and fourth disks. (This effect is accentuated by rotating the page so that the disks are arrayed in a lign above the cross.) Next try fixating on the plus sign below the disks. Now all of the disks are relatively easy to individuate even though the third, for example, is at the same distance from the center of gaze as before. This observation suggests that the smallest area within which attention can be deployed is not circular; it is narrower (affording more resolution) in the tangent direction than in the radial direction..

The lateral interference effect measured in these experiments is not limited to letter stimuli. In general, when a target is embedded in multiple distractors, it is very difficult to perceive the target. It is our claim, outlined below, that this difficulty reflects the limited spatial resolution of our attentional mechanism.

Interestingly, the crowding effect depends critically on the similarity between the distractors and the target to be detected, and their relative depth (Andriessen & Bouma, 1976; Kooi, Toet, Tripathy, & Levi, 1994). Grouping between the distractors and the target also significantly changes the crowding effect (Banks & Prinzmetal, 1976). These results indicate that effects must emerge at a fairly high level.

4. Crowding, attentional resolution, and the locus of awareness

In severe crowding, a target embedded in a dense array of distractors can not be selected by attention, hence can not be consciously scrutinized, independently of the distractors. The claim that the crowding effect is due to insufficient spatial attentional resolution is further supported by the next experiment and those described in the following sections.

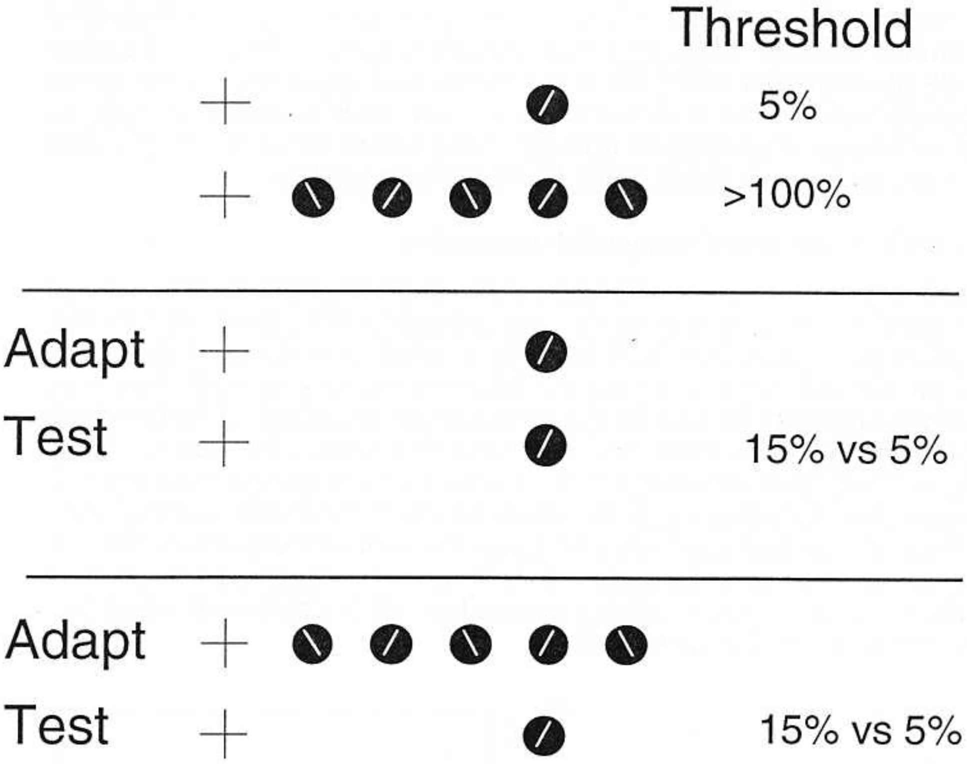

Figure 4. The first panel shows the contrast of the bars required for their orientation to be judged with 75% accuracy. In the crowded case, responses were basically at chance levels. The orientation information was inaccessible to conscious report. Following adaptation, we measured the contrast required to report the orientation of a single grating at the location of the adaptation. Three times more contrast was required if the test had the same orientation as the adapting grating. The same orientation specific effect was found whether the observer was aware of the orientation (single adapt, middle panel) or not (crowded adaptation, bottom panel). Adapted from He, Cavanagh, and Intriligator (1996) but depicted for simplicity as single bars within a circular disk arrayed horizontally. In the experiment the entire disk was filled with uniformly spaced bars and the disks were arrayed vertically.

To measure the consequences of crowding, we (He, Cavanagh, & Intriligator, 1996) used an effect of adaptation which we know to be mediated by structures in the visual system located in primary visual cortex or beyond. Specifically, adaptation to a grating in a particular orientation will raise the threshold for subsequently detecting a grating in that same orientation much more than detecting a grating in the orthogonal orientation. This orientation selective adaptation has been linked to neurons at the primary visual cortex and beyond (Movshon &

Lennie, 1979). In our study , we adapted our observers either to a single grating, or a crowded grating array. In this latter, crowded case, the adapting grating at the target position was unavailable to conscious perception (see Figure 4, top), and identification of the orientation of the grating was at chance levels. However, we found the same magnitude of orientation selective adaptation effect in both the single and crowded adaptation conditions (Figure 4, middle and bottom).

A grating that is unavailable to our conscious perception (due to crowding) is nevertheless as powerful as a single, clearly visible grating in producing orientation specific adaptation effects. This suggests that crowding must happen at a stage beyond the site of adaptation which itself must be at least at primary visual cortex (no orientation analysis occurs at earlier levels). This result is consistent with the proposal that activation of neurons in primary visual cortex alone is not sufficient for visual awareness (Crick & Koch, 1995; Koch & Tootell, 1996).

5. Upper/lower field and radial/tangential asymmetry

Visual resolution is not uniform over the visual field being dramatically better at the center of gaze. Other than this very strong inhomogeneity, there are no other marked effects of position in the visual field (for example, resolution is similar on the left and right, top and bottom, once eccentricity is taken into account). However, the same does not appear to be true for the resolution of attention? A lower visual field advantage was found in many different tasks that require focused attention. For example, a lower field advantage was evident in a attentional tracking task where observers were asked to track two discs among 9 randomly moving discs (Figure 5). Since the tracked discs are physically identical to the nontracked ones, the performance in this task depends critically on the ability of attention to individuate the selected targets, resolving targets from distractors even when they sometimes moved very close to each other.

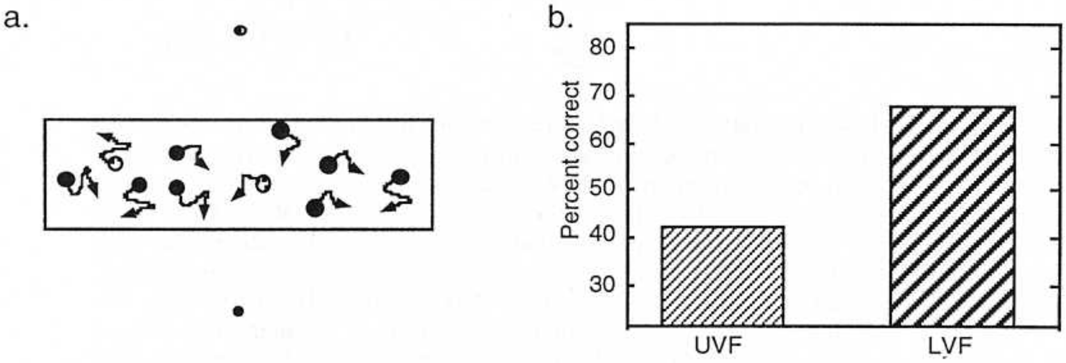

Figure 5. a) Nine disks moved randomly around a rectangular area, bouncing off the perimeter walls and bouncing off invisible cushions around each other. Two of the balls were briefly colored red to indicate that they were the targets to follow. They then returned to the same color as the others and the observers had to keep track of those two disks. After 5 seconds of tracking, the display stopped and observers indicated which were the two tracked disks. For the test of the upper field, the observers performed the task while fixating the lower dot. For the test of the lower field, they fixated the upper dot. b) Performance in the lower field was markedly better than in the upper field.

If the more severe crowding seen in the upper field is really a consequence of attentional processes, we might expect that tasks which require very little attention would show less of a difference between upper and lower fields. When we compared a task requiring significant attention — visual search for conjunctions — to one which required little attention — a pop-out visual search — we found the expected pattern. The conjunction search had a large difference between upper and lower fields, whereas the feature search had little or none.

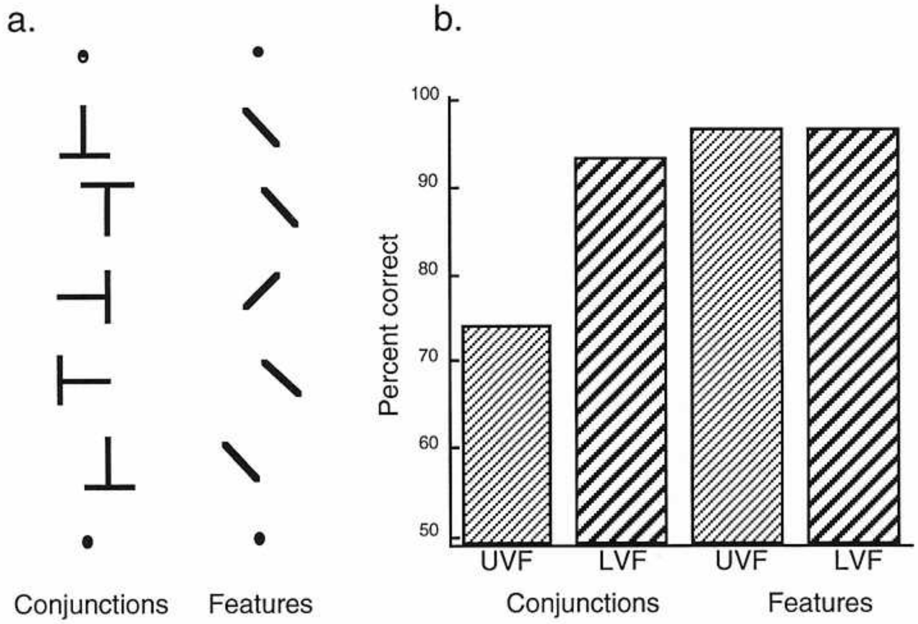

Figure 6. a) Observers reported the orientation of the middle T (a difficult task requiring attention), or the orientation of the middle tilted line (an easy task requiring little attention). The Ts all have both vertical and horizontal bars and it is the conjunction of the two bars which produces the T at a particular orientation. The target single line was always the same as the distractors or it was the one unique orientation — an easy task b) Performance in the lower field was markedly better than in the upper field for the conjunction task, almost reaching the level for the easy feature task which itself was equally good in both fields.

The lower visual field advantage in attentional resolution may be partly due to the fact that lower field is represented in the upper part of the primary visual cortex, which is anatomically adjacent to and projects more heavily (Maunsell & Newsome, 1987) to the occipital-parietal regions that are often linked to spatial attentional control (Gazzaniga & Ladavas, 1987; Posner, Walker, Friedrich, & Rafal, 1987). In other words, these results suggested again that the locus of the attentional bottleneck on resolution lies beyond the primary visual cortex which itself has relatively similar representations of the upper and lower visual fields.

As shown in Figure 3, attentional resolution is anisotropic in the radial/tangential dimension too. There is significantly less interference between items when they are arranged tangentially than radially relative to the eyes' fixation (also shown by Chambers & Wolford, 1983; Toet & Levi, 1992). It seems that the smallest region to which we can attend is ellipsoidal, elongated along the radial line from fixation.

6. What is the grain of attention?

We have demonstrated that the resolution of attention is coarser than visual resolution and that its inhomogeneities go beyond those of primary visual cortex. In this section we will characterize the resolution of attention in more detail.

6.1 Tracking tasks

In our lab, James Intriligator (Intriligator, Nakayama, & Cavanagh, 1991) used an attentional tracking task like that described in Figure 5 to measure the limiting resolution of tracking with attention. Again there were 9 randomly moving disks but now anywhere from 1 to 4 of them could be highlighted as targets for the tracking period and observers were allowed free viewing so they typically tracked the center of the group of target disks. The viewing distance was changed to make the display increasingly smaller and performance was measured at each distance for each of the 1 to 4 targets. At the furthest viewing distance, the display subtended less than a degree of visual angle, about the size of the fingernail on the index finger when held at arms length. At this distance, the disks were clearly visible but it was impossible to track even one of them at better than chance levels. Just as in Figure 1 here, they could be seen as a texture (now teeming with motion rather than static) but they could not be individuated. The data from this experiment suggested that the resolution of attention was about one tenth as good as visual resolution. That is, tracking became difficult when the interdisk spacing was about ten times that required to see that there were two disks not one at any given location.

6.1 Counting tasks

The variation of disk spacings in the tracking task is unpredictable and we cannot know whether a moment's inattention or a close encounter between two disks actually impaired tracking. To get a better estimate of the resolution of attention as a function of location in the visual field, James Intriligator next developed a individuation task where a circumferential ring or a radial row of static disks was presented. A starting position was cued by turning a disk briefly red and then back to the same color as the other disks. The computer then instructed the observer to move his or her attention to the disc on the left or right (without eye movements) six or seven times in sequence. In this attentional walk, the observer's attention should have ended up on a specific disk if he or she was capable of individuating each of the disks. By varying the density of the disks and their eccentricity, Intriligator measured the critical density all over the visual field.

The final result is captured by the display in Figure 7 and shows that at least 40 and as many as 80 items can be packed into the visual field while still permitting each to be accessed individually by attention. This number has two immediate implications. First, it is far larger than the span of apprehension (typically 4 or 5), the number of items which can be apprehended in a glance. Clearly, the attentional stepping task is not accomplished by perceiving the layout in a glance but by stepping through the items one at a time. Second, it is far smaller than the number of hypercolumns in primary visual cortex, often suggested as the elementary module of visual processing (there are about 1100 hypercolumns in macaque monkey and probably a somewhat smaller number in humans).

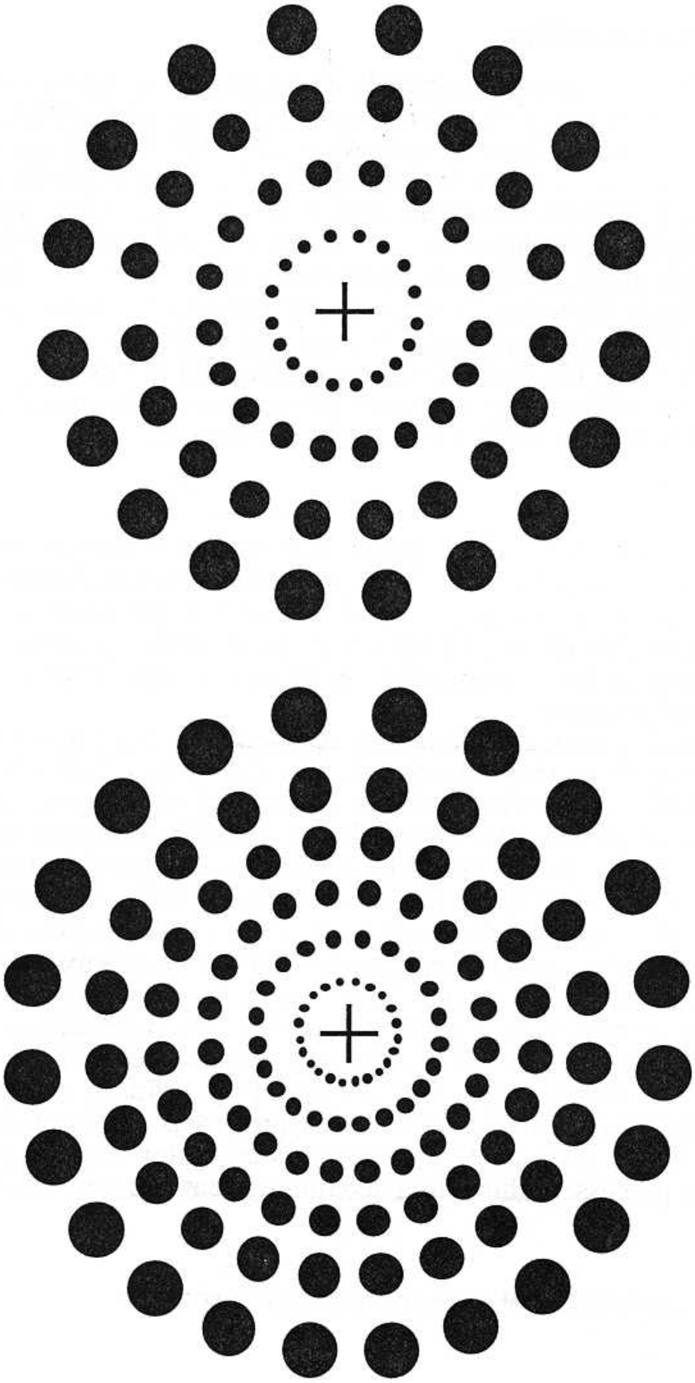

Figure 7. While fixating the cross on the top pattern notice that it is fairly easy to let your attention select any item in the four concentric arrays. Each disk is spaced at just less than the critical local density for access by attention. The pattern on the bottom has a density that exceeds the resolution limit of attention. Consequently, while fixating the bottom cross you may find it difficult to move your attention from item to item especially within the central rings

8. Temporal resolution of attention

The limited resolution of attention in spatially crowded displays has a very close analog in the time domain. We are able to sense very brief changes in visual patterns and we can notice the flicker in lights up to a rate of about 60 flashes per second. However, well below that rate we lose the ability to individuate each flash. This phenomenon was studied long ago as Gestalt flicker. When a light turns on and off at a slow rate, it is seen as on and then gone. At a faster rate, about 5 to 7 times per second, we no longer are aware of the individual events but only that the light has continual existence and it is flickering. In this case, the property of flicker is similar to the spatial texture in Figure 2 where the individual bars could not be isolated but their pattern was clearly seen. The flicker is a temporal texture of this kind. Several studies in our lab and elsewhere have supported this notion of a temporal limit to attentional access and that limit is about 5 to 7 events per second.

9. Conclusions

We have argued that attention has spatial and temporal resolution limits analogous to those of visual resolution but which are an order of magnitude coarser. These limits have direct and profound consequences for the content of our conscious perception specifically because we are aware of only the things we are attending to or things that capture our attention. In this sense, attentional resolution is the bottleneck of visual awareness.

Why would the visual system encode details at resolutions finer than attention can process? One answer is that attention is not the only client for visual information and that issue will be dealt with in our second paper here. Another answer is that small things are not often targets for attention but are more likely the elements in a surface texture which is useful in identifying a larger object to which it belongs. Finally, it is important to point out that this resolution limit applies to dense arrays of targets, blocking attention's access to the individual targets. A single element, remote from any others, can be individuated and scrutinized by attention.

Lastly, a central question in the study of attention is the neuronal substrate of attention and the visual awareness to which it gives rise. We have addressed this question by examining the fate of the information that does not get past the limit of attentional resolution. Our results argue that activity in the primary visual cortex cannot on its own be the site of visual awareness. The properties of the resolution of attention suggest that higher level centers, in the parietal cortex in particular, are more likely sites for the processes which limit determine attentional resolution.

Acknowledgements

This research was supported by NIH grants to PC and SH.

References

Andriessen, J.J., and Bouma, H. (1976) "Eccentric vision: Adverse interactions between line segments", *Vision Research* 16(1):71-78.

Banks, W.P., and Prinzmetal, W. (1976) "Configurational effects in visual information processing", *Perception and Psychophysics* 19(4):361-367.

Bouma, H. (1970) "Interaction effects in parafoveal letter recognition", *Nature,* 226(4):177-178.

Campbell, F.W., and Gubisch, R.W. (1966) "Optical quality of the human eye", *Journal of Physiology* 186:558-578.

Castiello, U., and Umilta, C. (1992) "Splitting focal attention", *Journal of Experimental Psychology: Human Perception and Performance* 18(3):837-848.

Chambers, L., and Wolford, G. (1983) "Lateral masking vertically and horizontally", *Bulletin of the Psychonomic Society* 21(6):459-461.

Crick, F., and Koch, C. (1995) "Are we aware of neural activity in primary visual cortex?", *Nature* 375(6527):121-123.

Driver, J., and Baylis, G.C. (1989) "Movement and visual attention: The spotlight metaphor breaks down", *Journal of Experimental Psychology: Human Perception and Performance* 15(3):448-456.

Duncan, J. (1984) "Selective attention and the organization of visual information", *Journal of Experimental Psychology: General* 113(4): 501-517.

Eriksen, C.W.W., and Hoffman, J.E. (1972) "Temporal and spatial characteristics of selective encoding from visual displays", *Perception and Psychophysics* 12(2-B): 201-204.

Gazzaniga, M.S., and Ladavas, E. (1987) "Disturbances in spatial attention following lesion or disconnection of the right parietal lobe", in: *Neurophysiological and neuropsychological aspects of spatial neglect, Vol. 45,* M. Jeannerod, ed., Amsterdam, Netherlands: Elsevier Science Publishing Co, Inc., pp. 203-213.

He, S., Cavanagh, P., and Intriligator, J. (1996) "Attentional resolution and the locus of visual awareness", *Nature* 383(6598): 334-337.

Intriligator, J., Nakayama. K., and Cavanagh, P. (1991) "Attentive tracking of multiple moving objects at different scales", *Investigative Ophthalmology and Visual Science* 32, 1040.

Koch, C., and Tootell, R.B. (1996) "Visual psychophysics: stimulating brain but not mind [news; comment]", *Nature* 383(6598): 301, 303.

Kooi, F.L., Toet, A., Tripathy, S.P., and Levi, D.M. (1994) "The effect of similarity and duration on spatial interaction in peripheral vision. Special Issue: In honor of Bela Julesz: II. Texture and visual search", *Spatial Vision* 8(2): 255-279.

Kramer, A.F., and Hahn, S. (1995) "Splitting the beam: Distribution of attention over noncontiguous regions of the visual field", *Psychological Science* 6(6): 381-386.

LaBerge, D. (1983) "Spatial extent of attention to letters and words", *Journal of Experimental Psychology: Human Perception and Performance* 9(3): 371-379.

Maunsell, J.H., and Newsome, W.T. (1987) "Visual processing in monkey extrastriate cortex", *Annual Review of Neuroscience* 10, 363-401.

Movshon, J.A., and Lennie, P. (1979) "Pattern-selective adaptation in visual cortical neurones", *Nature* 278(5707): 850-852.

Posner, M.I., Walker, J.A., Friedrich, F.A., and Rafal, R.D. (1987) "How do the parietal lobes direct covert attention? Special Issue: Selective visual attention", *Neuropsychologia* 25(1-A): 135-145.

Pylyshyn, Z.W., and Storm, R.W. (1988) "Tracking multiple independent targets: Evidence for a parallel tracking mechanism", *Spatial Vision* 3(3): 179-197.

Toet, A., and Levi, D.M. (1992) "The two-dimensional shape of spatial interaction zones in the parafovea", *Vision Research* 32(7): 1349-1357.

Townsend, J.T., Taylor, S.G., and Brown, D.R. (1971) "Lateral masking for letters with unlimited viewing time", *Perception and Psychophysics* 10(5): 375-378.

Usai, M.C., Umilta, C., and Nicoletti, R. (1995) "Limits in controlling the focus of attention", *European Journal of Cognitive Psychology* 7(4): 411-439.