

Natural scenes upset the visual applectart

Jochen Braun, Plymouth Institute of Neuroscience, Devon, Great Britain
(achim@pion.ac.uk)

The effortless ease of everyday vision seems to conflict with laboratory findings on the limited capacity of visual attention. It now appears that naturalistic stimuli escape the stringent limitations of attention that apply to seemingly far simpler stimuli. This astonishing result will oblige us to rethink the nature of visual attention and its limited capacity.

Everyday vision differs in many ways from what is commonly studied in vision laboratories. For one, the complex natural scenes that inundate everyday vision are far cry from the meagre fare of geometric primitives (letters, gratings, etc.) served in most laboratories. For another, the stimuli of everyday vision are rarely as repetitious as those used in the laboratory, raising the spectre of narrow skills more highly practised than their everyday counterparts. The list of differences lengthens further when we consider dynamic aspects of vision, integration with eye movements, and so forth. These are sobering thoughts for anyone hoping to understand vision in the real world, for it means that laboratory findings will not necessarily generalise to the domain of ultimate interest.

Li, Van Rullen, and colleagues [1] now report a particularly dramatic failure to generalise from simple to naturalistic stimuli. Their key finding is that observers categorise the gist of natural scenes (e.g., does it contain an animal/vehicle?) even though they cannot, under otherwise identical conditions, distinguish simple geometric shapes. Thus it would appear that natural scenes are processed more efficiently, and with less call on attentional resources, than simple geometric shapes. Before accepting this startling conclusion, one will wish to take a closer look at the details. Li and colleagues flash natural scenes briefly in the visual periphery, followed by a custom-made stochastic mask to limit visual persistence. The effective presentation time (= 80ms) falls in the ‘ultra-rapid categorisation’ range described by Thorpe and others [2-4]. With full attention, observers categorise approximately 3 out of 4 of scenes correctly, that is to say, they correctly distinguish scenes with and without animals/vehicles in about 75% of all trials (Fig. 1A). The same level of performance is obtained for discriminating the simple geometric shapes that serve as a comparison (rotated Ts or Ls, two-coloured disks, Fig. 1BC), although presentation times need to be longer, hinting that this may not be so ‘simple’ after all.

Superiority of natural scenes

A yawning gulf separates natural scenes and geometric shapes, however, when the observer’s attention is coaxed away from either. For natural images continue to be categorised with the same success rate even in the ‘near absence’ of attention, whereas

the discrimination of geometric shapes collapses to chance level. The startling robustness of natural image processing is especially impressive in view of the demonstrable effectiveness of the attentional coaxing [5]. Specifically, observers focussed attention as much as possible on a visual search task near the centre of gaze, while reporting independently on both the central search and the peripheral categorisation/discrimination. To verify this lopsided distribution of attention, central performance was compared against a baseline of performing the central search alone (i.e., ignoring the periphery) and no significant difference was found. Any straying of attention would, of course, have been reflected in a loss of central performance. In addition, attention was engaged for an extended interval, starting well *before* and ending well *after* the peripheral stimulation, so that attention could not have 'swerved' to the periphery either before or after the central search.

Why should natural scenes be more readily distinguishable than geometric shapes? In fact, there are previous indications that natural scenes may trigger a more profound perceptual and neural response than plain geometric shapes. Mack and Rock [6] investigated whether observers are aware of highly visible but entirely unexpected stimuli when their attention focuses elsewhere, a closely related situation the one under discussion. They observed a dramatic failure of awareness with unexpected geometric shape (25% to 80% 'blindness' for shapes of various sorts), but virtually complete awareness with natural scenes (including faces). In fact, no observer failed to notice the unexpected appearance of a natural scene, and most gave fairly accurate descriptions of its contents, in stark contrast to geometric shapes. The underlying reason for this heightened sensitivity to natural scenes may be the nature the neural code for visual information. In area V1 of macaque, natural stimulation increases sparseness and intensity of neuronal responses, coupling intense bursts with lower spiking activity overall and thus increasing information per spike and efficiency per neurone [7,8]. Information transmission by V1 neurones (median values of bits/sec and bits/spike) varies with stimulus type and may be substantially higher for natural scenes than for geometric stimuli such as gratings [8,9]. The informative components of natural scenes are sparsely distributed in space and time [10] and the neural code is thought to take advantage of this fact. If so, the neural representation of natural scenes will be more intense and more intermittent than that of other stimuli that appear 'simple' to us. In short, natural scenes may be intrinsically superior stimuli to the 'simple' geometric shapes commonly used in vision research.

Vision without attention

In addition to the possible special status of natural scenes, the present results also bear on visual capabilities in the 'near absence' of attention. Trained psychophysical observers are aware of salient stimuli 'outside the focus' of attention and discriminate some of their attributes (e.g., contrast, colour, orientation, ...), although more complex attributes remain indistinguishable (e.g., rotated Ts and Ls, colouring of two-coloured disks)[11,12]. Prior familiarity with flashed, masked display turns out to be important to these findings [13,14]. In particular, vision 'outside the focus' of attention is observed with 'expert' but not with 'novice' observers (to both groups, the specific displays in

question are novel, but ‘experts’ benefit from prior experience with other displays). Importantly, Li and colleagues [1] show that their results are not due to prior familiarity or training. Firstly, categorisation in the ‘near absence’ of attention occurs also for unfamiliar scenes, which observers have not seen during training. Secondly, observers can alternate between categorising animals and categorising vehicles without adverse effects on performance, suggesting that categorisation extends even to unfamiliar categories (Fig. 1A). Finally, Li and colleagues show that the discrepancy between natural scenes and geometric shapes persists even after equal training on each stimulus type.

In perceptual research, ‘absence’ of attention remains somewhat controversial with no operational definition universally agreed. One school of thought relies on the performance effects of dividing attention (not necessarily equally) between two conflicting tasks [15]. As one task receives more attention, the other task necessarily receives less, causing the respective performance levels to move in opposite directions. When one task is performed at chance, one may infer ‘absence’ of attention (or ‘near absence’, given the uncertainty of performance measures). Another school of thought holds that some attention is allocated involuntarily [16,17]. In this view, it may not be possible to ignore (i.e., leave entirely unattended) a stimulus as its mere presence may ‘capture’ attention. To account for the results of Li and colleagues, adherents of this view might suppose that natural scenes ‘capture’ more attention than geometric shapes. However, the central search was performed equally with both types of peripheral stimuli, providing no evidence that differential ‘capture’ detracted differentially from the central search. Accordingly, a ‘captured’ attention account would not be consistent with a limited capacity of visual attention.

Neural theory of attention

From a neural point of view, the notion of ‘unattended’ stimuli is less problematic. The dominant neural theory views attention as a dynamic state of activity in a hierarchy of cortical and possibly sub-cortical visual areas [18,19]. In this view, the activity pattern is shaped by visual input signals, by competitive interactions within the hierarchy, and by ‘bias’ signals from higher areas. The pattern that emerges may comprise multiple winning (‘salient’) objects, all of which become available to awareness, short-term memory, voluntary reporting, and so on. Objects that gained saliency from ‘bias’ signals would be considered ‘attended’ in the perceptual terminology, whereas objects benefiting exclusively from bottom-up interactions would be considered ‘unattended’. It follows that conscious vision ‘outside the focus’ of attention should embrace only the most competitive stimuli in the field of view. In the paradigm of Li and colleagues [1], observers would have been aware of the single peripheral stimulus due to its competitive nature. It may be worth pointing out that this awareness should not extend to multiple peripheral stimuli, where the competitive advantage should be lost. In other words, it should not be possible to search an entire array of natural scenes without focussed attention. Rather, visual search with natural scenes should still be ‘serial’ and search times should increase with array size.

Conclusion

For more than twenty years, the effortless ease of everyday vision has perched uneasily next to some cherished laboratory findings on visual capacity limitations (e.g., the ponderous pace of serial visual search). To resolve this paradox, it has even been suggested that our experience of a rich visual world in front of our eyes is a ‘grand illusion’ sustained by memory [20]. On a less grand but still helpful scale, visual capacity was found to be more commodious than at first apparent, thanks to bottom-up interactions mediating conscious access ‘outside the focus’ of attention [11-14]. For about the same time, we have known that observers process natural scenes extremely rapidly, at least to the point of recognising ‘gist’ [2-4]. Li, Van Rullen, and colleagues have now gone even further by demonstrating that natural scenes escape some of the stringent capacity limitations that pertain to geometric shapes [1]. If visual capacity is even more commodious for natural scenes, this may finally reconcile visual psychophysics with our experience of everyday vision. Of course, this reconciliation would oblige us to revisit many results on capacity limitations previously obtained with geometric shapes. In short, everyone working on visual attention and capacity limitations ought to be grateful to Li, Van Rullen, and colleagues for having well and truly upset the apple cart.

REFERENCES

- 1 Li, F.F., Van Rullen, R., Koch, C. and Perona P. (2002) Rapid natural scene categorization in the near absence of attention. *Proc Natl Acad Sci USA* 99, 9596-601.
- 2 Fabre-Thorpe, M., Delorme, A., Marlot, C. and Thorpe, S. (2001) A limit to the speed of processing in ultra-rapid categorization of novel natural scenes. *J Cogn Neurosci* 13, 171-80.
- 3 Potter, M.C. and Levi E.I. (1969) Recognition memory for a rapid sequence of pictures. *J Exp Psychol* 81, 10-5.
- 4 Biederman, I., Mezzanotte, R.J. and Rabinowitz, J.C. (1982) Scene perception: detecting and judging objects undergoing relational violations. *Cognit Psychol* 14, 143-77.
- 5 Lee, D.K., Koch, C. and Braun, J. (1999) Attentional capacity is undifferentiated: concurrent discrimination of form, color, and motion. *Percep Psychophys* 61, 1241-1255.
- 6 Mack, A. and Rock, I. (1998) *Inattentional Blindness*, MIT Press, Cambridge, MA.
- 7 Vinje, W.E. and Gallant, J.L. (2000) Sparse coding and decorrelation in primary visual cortex during natural vision. *Science* 287, 1273-6.

- 8 Vinje, W.E. and Gallant, J.L. (2002) Natural stimulation of the nonclassical receptive field increases information efficiency in V1. *J Neurosci* 22, 2904-15.
- 9 Reich, D.S., Mechler, F. and Victor, J.D. (2000) Formal and attribute-specific information in primary visual cortex. *J Neurophysiol* 57, 132-146.
- 10 Simoncelli, E.P. and Olshausen, B.A. (2001) Natural image statistics and neural representation. *Annu Rev Neurosci* 24, 1193-216.
- 11 Braun, J. and Sagi, D. (1990) Vision outside the focus of attention. *Percept Psychophys* 48, 45-48.
- 12 Braun, J. and Julesz, B. (1998) Withdrawing attention at little or no cost: detection and discrimination tasks. *Percept Psychophys* 60, 1-23.
- 13 Joseph, J.S., Chun, M.M. and Nakayama, K. (1997) Attentional requirements in a 'preattentive' feature search task. *Nature* 387, 805-7.
- 14 Braun, J. (1998) Vision and attention: the role of training. *Nature* 393, 424-5.
- 15 Sperling, G. and Melchner, M.J. (1978) The attention operating characteristic: examples from visual search. *Science* 202, 315-8.
- 16 Simons, D.J. (2000) Attentional capture and inattention blindness. *Trends Cog Sci* 4, 147-155.
- 17 Rensink, R.A. (2002) Change detection. *Annu Rev Psychol* 53: 245-77.
- 18 Reynolds, J., Chelazzi, L. and Desimone, R. (1999) Competitive mechanisms subserve attention in macaque area V2 and V4. *J Neurosci* 19, 1736-53.
- 19 Corchs, S. and Deco, G. (2002) Large-scale neural model for visual attention: integration of experimental single-cell and fMRI data. *Cereb Cortex* 12: 339-48.
- 20 Noe, A. and O'Regan, J.K. (2000) Perception, attention, and the grand illusion. *Psyche* 6, October 2000.

FIGURE LEGEND

Natural scenes and geometric shapes, the two types of stimuli used by Li, Van Rullen, and colleagues [1]. Both types of stimuli were presented briefly in the visual periphery and were followed by a mask to limit visual persistence. Observers focussed attention on a demanding task near the centre of gaze, so that any awareness of the peripheral stimuli arose in the ‘near absence’ of attention. Surprisingly, observers categorised natural scenes successfully, but fared dismally on discriminating geometric shapes. A. Natural scenes with animals, vehicles, and neither. Observers categorised scenes as with or without animals, or as with or without vehicles. B. Rotated Ts and Ls. Observers tried to report letter identity (T or L). C. Two-coloured disks. Observers attempted to report colour order (red-green or green-red).

