

Detection of animals in natural images using far peripheral vision

Simon J. Thorpe,¹ Karl R. Gegenfurtner,² Michèle Fabre-Thorpe¹ and Heinrich H. Bülthoff³

¹Centre de Recherche Cerveau et Cognition (UMR 5549, CNRS-UPS), Faculté de Médecine de Rangueil, 133, route de Narbonne, 31062, Toulouse, France

²Institut für Psychologie, Otto-von-Guericke Universität Magdeburg, PF 4120 39016 Magdeburg, Germany

³Max-Planck Institut für Biologische Kybernetik, Spemannstr. 38, 72076 Tübingen, Germany

Keywords: human performance, large eccentricities, natural images, object vision, peripheral vision

Abstract

It is generally believed that the acuity of the peripheral visual field is too poor to allow accurate object recognition and, that to be identified, most objects need to be brought into foveal vision by using saccadic eye movements. However, most measures of form vision in the periphery have been done at eccentricities below 10° and have used relatively artificial stimuli such as letters, digits and compound Gabor patterns. Little is known about how such data would apply in the case of more naturalistic stimuli. Here humans were required to categorize briefly flashed (28 ms) unmasked photographs of natural scenes (39° high, and 26° across) on the basis of whether or not they contained an animal. The photographs appeared randomly in nine locations across virtually the entire extent of the horizontal visual field. Accuracy was 93.3% for central vision and decreased almost linearly with increasing eccentricity (89.8% at 13°, 76.1% at 44.5° and 71.2% at 57.5°). Even at the most extreme eccentricity, where the images were centred at 70.5°, subjects scored 60.5% correct. No evidence was found for hemispheric specialization. This level of performance was achieved despite the fact that the position of the image was unpredictable, ruling out the use of precued attention to target locations. The results demonstrate that even high-level visual tasks involving object vision can be performed using the relatively coarse information provided by the peripheral retina.

Introduction

It is frequently assumed that, because of the poor quality of extrafoveal vision, object recognition requires a combination of selective attention and successive eye movements that bring items of interest into foveal vision (Liversedge & Findlay, 2000). The poor quality of form vision in the periphery has been documented extensively using artificial patterns such as digits, letters or compound Gabor patterns (Strasburger *et al.*, 1991; Nasanen & O'Leary, 1998; Jüttner & Rentschler, 2000). A very limited number of studies have used drawings of real-world objects (Biederman *et al.*, 1981; Levy *et al.*, 2001) and very little is known about the perception of natural scenes in peripheral vision (though see, for example, Loftus & Mackworth, 1978; Nelson & Loftus, 1980). Furthermore our understanding of peripheral visual functions is limited by the fact that most studies have only used a limited range of eccentricities, often below 10°. Although it seems likely that an ability to detect important categories of natural objects across the entire visual field would be a distinct advantage for animals, almost nothing is known about the performance of the peripheral field in such situations.

In the present experiment, we used a superordinate visual categorization task in which human subjects had to decide whether a natural photograph contained an animal or not. Our aim was to study how naturalistic stimuli can be processed using the peripheral retina up to the far limits of the visual field and how performance

varies with increasing eccentricity when subjects are required to monitor their entire horizontal visual field.

Methods

To address these issues, we used a Go/No-go experimental paradigm introduced by Thorpe and colleagues (1996) in which subjects have to respond by releasing a button when they detect an animal in a briefly flashed (28 ms) colour photograph of a natural scene. The photographs could appear at random locations across practically the entire extent of the horizontal visual field; subjects were seated in the centre of a semicircular panoramic theatre 3.5 m from the screen that allows images to be presented anywhere within a region 180° wide by 50° high, under realistic viewing conditions (see Fig. 1a). The images were very large (39° high, and 26° across) and were presented at nine different positions, spanning much of the horizontal extent of the visual field (see Fig. 1b). Target and distractor images were equally probable, and no attempt was made at rescaling the photographs to compensate for the loss of acuity associated with eccentric presentations. The task presents a major challenge to the visual system because of the high variability of the target and distractor images, the unpredictability of the location for the next stimulus and the use of very short stimulus durations. It is particularly important to note that because the position of the flashed photograph was random, subjects could not direct their attention in advance to a particular location, and were obliged to spread attention across the entire visual field. In addition, the very short presentation time (28 ms) implies that there

Correspondence: Dr Simon Thorpe, as above.
E-mail: thorpe@cercu.ups-tlse.fr

Received 24 January 2001, revised 19 June 2001, accepted 29 June 2001

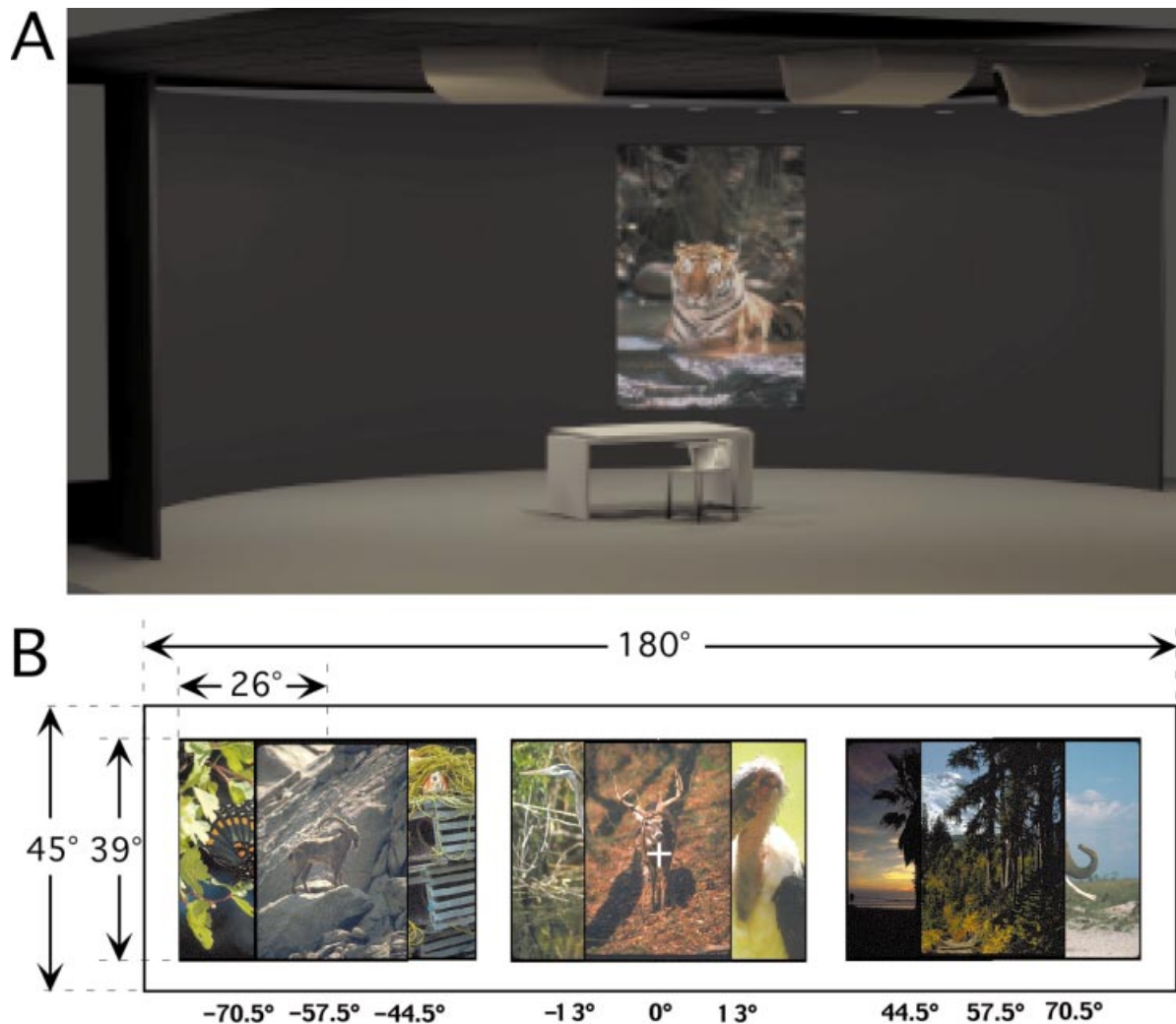


FIG. 1. (A) Experimental set-up. Images were projected on the screen using three Electrohome Marquee 8000 CRT projectors controlled by a Silicon Graphics Infinite Reality II system, each with a resolution of 1280×1024 pixels and each covering 65° horizontally. The overall lighting in the room was dim, the background luminance of the screen was 0.02 cd/m^2 and the average luminance of the images was about 3 cd/m^2 . (B) Arrangement of images. The nine possible positions covered almost the entire horizontal extent of the visual field. Because there was a 6° overlap between each projector the total horizontal range was 3584 pixels. The images were 384 pixels high by 256 wide, but were doubled in size for projection resulting in images that were 39° high by 26° wide. They could appear in nine different places in the visual field, on the left, right or centre of each projector. On the screen, images were centred at -13° , 0° and 13° in the central and near peripheral field and at -70.5° , -57.5° and -44.5° on the left and 44.5° , 57.5° and 70.5° on the right of the far peripheral field. As images were 26° wide there was an overlap of 13° between different positions. Note that, at the most extreme positions, the edge of the image had an eccentricity of 83.5° .

was no possibility of making eye movements and that visual inputs were initially lateralized to one cerebral hemisphere.

Ten subjects participated in the study (seven males and three females, age range 25–38 years). One thousand, four hundred images were taken from the Corel Photo Library. Of these, 700 were targets that contained one or more animals in their natural environments. Targets included mammals, birds, fish and insects (for examples see Figs 1 and 3). The other 700 were a very varied set of distractors that included landscapes, buildings, flowers, and so on (see Fig. 3C). They were organized in 14 blocks of 100 images each (50 targets and 50 distractors). Each subject was shown all 1400 images only once at a given position in the peripheral field, and over the group of 10 subjects, each image was presented once at each of the eight eccentric positions and twice at the 0° position. During the experiment, the subject was requested to fixate a central cross (size $0.5^\circ \times 0.5^\circ$). Then images were presented one at a time for two frames at 72 Hz

(28 ms) and the subject was asked to respond by releasing a mouse button if the image contained an animal. Subjects were allowed 1 s to respond, and the intertrial interval was ≈ 2 s. This go/no-go protocol was used in order to permit direct comparisons with preceding studies and to allow reaction times (RTs) to be as short as possible. The use of a two-button task would almost certainly have resulted in an increase in RT because of the additional delay required to decide between two alternative motor responses.

Results

The results for the 10 observers are summarized in Fig. 2 and Table 1. With centrally presented images, accuracy was 93.3%, a value very close to those reported in previous studies using the same sort of paradigm (Thorpe *et al.*, 1996; Fabre-Thorpe *et al.*, 1998;

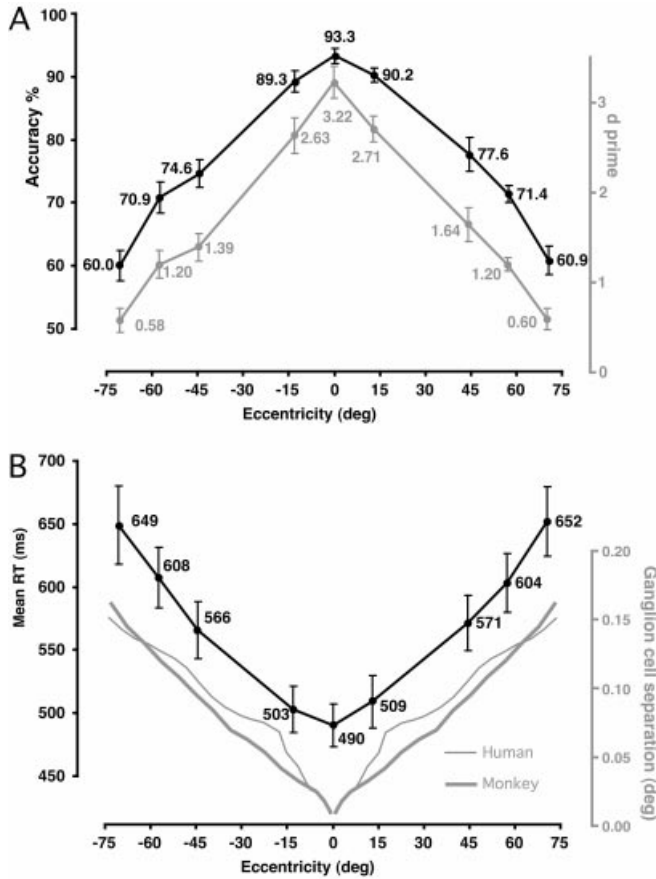


FIG. 2. (A) Performance as a function of eccentricity. The data plotted in black refer to overall accuracy (left hand scale), whereas data in grey refers to d' scores (right hand scale). Values are the means for 10 subjects and the error bars are SEM. (B) The black lines plot mean reaction times (RTs) as a function of eccentricity (left hand scale). Values are means and standard errors for 10 subjects. For comparison, the grey lines plot values for ganglion cell spacing for the nasal retina (the part used for processing of stimuli with high eccentricities). The data have been recalculated from the studies by Wässle *et al.* (1990) in monkeys, and by Curcio & Allen (1990) in humans. Note that in both species, the angular spacing increases roughly linearly with eccentricity.

Antal *et al.*, 2000). As expected, performance drops with increasing eccentricity from an accuracy of 89.8% for images centred at 13° decreasing to 76.1% at 44.5°, 71.1% at 57.5° and 60.4% at 70.5°. However, performance was very highly above chance for all subjects at all eccentricities up to and including 57.5° ($P < 0.0001$, two-tailed t -test). When the presentations were made centred at 70.5° eccentricity, the edge of the image had an eccentricity of 83.5°. But even here, the overall results for the group were also well above chance ($P = 0.0007$, two-tailed t -test), and only three of the 10 subjects failed to reach a significance of $P < 0.02$ (two-tailed t -test). These results indicate that visual information originating in the far peripheral retina can be processed to make superordinate categorizations, such as deciding whether or not a natural photograph contains an animal. The fall in accuracy with increasing eccentricity is remarkably linear; this is illustrated in Fig. 2A, which also shows the same linear decrease in performance using the sensitivity measure d' ($\rho = 0.998$), from a value of 3.22 for centrally presented stimuli to 0.58 at 70.5° in the left visual field.

The accuracy of categorization in peripheral vision varied substantially between subjects, but in the best subjects, performance

was very high. One of them, subject 8, correctly categorized 189 of the 280 images presented at 70.5° eccentricity (72.1%). At 57.5° eccentricity, the best result was observed with subject 7, who correctly categorized 227 of the 280 images, an accuracy level of 81.1%.

Despite the short presentation time that resulted in the lateralization of the visual inputs to one cerebral hemisphere, there was no evidence for hemispheric specialization as neither accuracy nor mean RT differed between the left and right visual fields, confirming previous studies that have failed to find evidence for lateralization (Biederman & Cooper, 1991; Fabre-Thorpe *et al.*, 1998). Overall accuracy on the left was 73.7% compared with 75.0% on the right ($P = 0.24$, paired t -test), while mean RTs were 577 ms on the left and 580 ms on the right ($P = 0.61$, paired t -test).

There was a nearly linear increase in mean RT with increasing eccentricity from 490 ms for central presentations to 650 ms for the most peripheral positions (see Fig. 2B and Table 1). There are various potential explanations for the systematic increase in RT with eccentricity. One factor that is likely to be important is related to the fact that in the present experiments, no attempt was made to adjust the size of the stimuli to compensate for changes in cortical magnification with eccentricity. As a result, stimuli presented close to the fovea will activate a far-larger number of neurons at every level of the visual system than would peripherally presented stimuli.

Indeed, if it were the case that object categorization in periphery required more prolonged processing, one should predict that when subjects responded particularly quickly with peripheral stimuli, processing may not be complete, resulting in lower accuracy. To look for that possibility, we selected, at each eccentricity and for each of the 10 subjects, the 40 trials where they responded with the shortest RTs. Table 1 shows that these fast responses were no less accurate than the overall level of performance. For example, at 70.5° eccentricity, mean accuracy for the fastest responses was 62.5%, which was actually even higher than the overall accuracy level of 60.4%. Accuracy is as high for fast responses than for longer latency responses, which suggests that the longer mean RTs seen with peripheral stimuli do not necessarily reflect a fundamentally different decision strategy from the one used with central stimuli.

Analysing those images that resulted in the shortest RTs also provides information about the sorts of stimuli that can be categorized successfully in peripheral vision. One might expect that this ability would be restricted to a very limited subset of images, ones where the target is particularly large and visible, but we recently reported evidence that this was not the case, at least in central vision (Fabre-Thorpe *et al.*, 2001). To illustrate the variety of images to which subjects could respond with fast RTs, Fig. 3A shows, for subject 7, the images associated with the 30 shortest RTs at 57.5° eccentricity. The set of images is very varied, including mammals, birds and fish seen from a wide range of viewing angles. Subject 7 was the most accurate subject at that eccentricity and the 35 fastest responses were all to targets — a result that would be extremely unlikely by chance.

To get a more complete view of the sorts of targets that were successfully detected in peripheral viewing, we looked at the images that were associated with particularly short RTs across all subjects. One hundred and nine of the 700 targets were identified correctly by all 10 subjects, i.e. at all eccentricities. Figure 3B shows the 20 images associated with the shortest mean RTs when presented at eccentricities of 57.5° and 70.5°, and ranked in order of mean RT. At these extreme eccentricities, the range of forms appears more stereotyped (including many close-ups of one centrally presented animal) than when a single subject is taken into account (as in Fig. 3A) but the range of targets correctly identified is still very wide,



including not just mammals, birds and fish, but also a beetle and a butterfly.

It is clear that if target stimuli differed in some obvious way, then it might be possible to use a relatively low-level decision strategy. Thus, it is equally important to know what sorts of distractor stimuli can be ignored successfully. The set of distractors was chosen to provide a severe challenge to the visual system. Of the 700 distractor images used in the present study, 133 were successfully ignored by all 10 subjects, i.e. at all positions in the visual field. Figure 3C illustrates 20 such stimuli and demonstrates the wide range of distractors that were successfully categorized. Note that such images contain various centrally presented forms, globally similar to some of the target stimuli, and that might well be expected to confuse a visual categorization system that used simple strategies for detecting animals.

On the other hand, the failures of the system can be equally informative. Figure 3D and E illustrates 10 targets and distractors that were identified perfectly in central vision and near periphery (0° and 13° eccentricity), but always induced errors in the far periphery (70° and 57.5° eccentricity). This peripheral fall-off in accuracy could be attributed to target size, as already reported by Biederman and colleagues (1981) and suggested by some of the pictures in Fig. 3D. However this former study used drawings in central vision and near periphery. In our task various factors have been shown to play a role in the central detection of targets in their natural environment (Delorme, 2000). Size is certainly one of them, but an even more critical factor might be the contrast between the target and its background. Further experiments are clearly needed to analyse the characteristics of the natural images classified correctly in far periphery.

Interestingly, there was little evidence that subjects needed to learn the task. Overall accuracy on the first of the 14 blocks was 78.7%, a value that was almost identical for each of the 14 blocks (mean value \pm SEM, $78.15 \pm 0.46\%$) and a paired *t*-test over the 10 subjects showed no significant difference in accuracy performance between the first and the last testing blocks ($t_9 = 0.79$, $P > 0.45$). This

makes it unlikely that the subjects were learning to use a particular *ad hoc* strategy for distinguishing targets from distractors. Rather, it would appear that they used processing strategies that had been learned previously. On the other hand, RTs did show a significant decrease during the first few blocks, but were stable from about the fourth block onwards (mean RT for block 1, 640 ms; for blocks 4–14, 548 ms).

One of the striking features of the current results is that, in extreme peripheral vision, subjects often reported that they did not really see the target in any detail, so that they would not have been able to identify the animal. They had the impression that they were guessing, although performance was well above chance at far eccentricities. Indeed in some pilot experiments, three subjects hardly responded at all when the images were presented at 70.5° – response rates were below 10%. For this reason, the subjects reported here were instructed specifically to rely on their first impressions and respond whether they were sure of their response or not; they were also told that because the probability of a target was 0.5, irrespective of eccentricity, the optimal strategy was to try and respond on roughly 50% of trials. Overall, subjects managed this quite well, as the response rates were close to 50% for all eccentricities up to 57.5° . Only at 70.5° did the response rate drop significantly, with an average of 42%, but individual subjects had response rates as low as 22.1% (see Table 1).

Discussion

The first point that can be made from these results is that superordinate categorization of objects is clearly possible in peripheral vision, even in the absence of foveating eye movements. With images centred at 13° , accuracy drops by about 3% and mean RT increases by 15–20 ms. It is clear that, as we made no attempt at scaling the images, this level of performance is excellent and shows that, had we restricted the study to the eccentricities normally studied, the impairment would have been very mild. However, the results reveal that the ability to perform superordinate object categorization extends well beyond the range of eccentricities generally used in

TABLE 1. Values for accuracy, percentage of responses and reaction time as a function of eccentricity

Eccentricity	Accuracy (%)		Percentage of responses (mean \pm SEM)	Reaction time (ms)	
	Overall (mean \pm SEM)	The fastest 40 responses (mean \pm SEM)		Median	Mean \pm SEM
0°	93.3 ± 1.2	91.8 ± 3.3	50.9 ± 1.1	465	490.2 ± 16.8
15°	89.8 ± 1.2	86.8 ± 2.9	50.9 ± 1.2	483	506.1 ± 19.4
44.5°	76.1 ± 2.2	79.0 ± 3.7	52.2 ± 2.6	547	568.9 ± 22.0
57.5°	71.1 ± 1.8	75.5 ± 3.3	50.5 ± 3.1	585	605.8 ± 23.2
70.5°	60.4 ± 2.1	62.5 ± 3.3	42.2 ± 3.7	615	649.6 ± 28.8

Results for presentation in the left and right visual fields have been merged. Accuracy is given both for the overall level of performance and for the 40 fastest responses for each subject at each eccentricity.

FIG. 3. (A) The 30 images that induced the fastest responses at 57.5° eccentricity for subject 7. A total of 280 images was presented (140 targets and 140 distractors), but all 35 of the fastest reaction times (RTs) were to targets. The numbers underneath each image provide: on the left, the RT for subject 7 (in ms) and on the right, the number of correct detections for the 10 subjects across all eccentricities. (B) The 20 targets that were successfully detected by all 10 subjects and which had the shortest mean RT at 57.5° or 70.5° eccentricity. The images are ranked in order of increasing mean peripheral RT (indicated in ms under each image). (C) Twenty of the 133 distractors (from a pool of 700) that were successfully ignored by all 10 subjects. (D) Ten of the 18 targets that were unusually difficult to detect in peripheral vision. All were detected perfectly at eccentricities of 0° and 13° , but were missed when presented at eccentricities of 57.5° or 70.5° . (E) Ten distractors that systematically resulted in false positives at extreme eccentricities (57.5° or 70.5°). They include the seven such distractors that were always correctly ignored at small eccentricities (0° or 13°) together with three of the five distractors that only produced one false positive out of four presentations at eccentricities of 0° or 13° .

studies of visual perception. Moreover, the nearly perfect linear drop in performance with eccentricity, as measured by d' scores (see Fig. 2A), implies that performance is probably limited by some fairly fundamental characteristic of the visual system. The obvious candidate is the spacing between retinal ganglion cells, which has been shown to increase in a roughly linear way with eccentricity. The angular separation between ganglion cells is about 0.01° at 1° from the fovea, but this value increases to about 0.2° at 90° eccentricity in both monkeys (Wässle *et al.*, 1990) and humans (Curcio & Allen, 1990) (see Fig. 2B). Cortical magnification for primary visual cortex also scales with eccentricity in such a way that about the same amount of cortical surface is devoted to processing a given number of ganglion cells, irrespective of eccentricity. There is a wide range of psychophysical tasks for which performance is known to scale with eccentricity in a way that is consistent with cortical magnification. These include measures of spatial acuity, and contrast sensitivity as function of both spatial frequency and temporal frequency (Koenderink *et al.*, 1978; Drasdo, 1991). However, in tasks involving form vision, peripheral vision appears to be systematically worse than would be predicted on the basis of changes in magnification alone (Strasburger *et al.*, 1994; Chung *et al.*, 1998; Levi *et al.*, 1999). One reason may be that the coding of the spatial phase appears particularly poor in peripheral vision (Rentschler & Treutwein, 1985; Bennett & Banks, 1987; Stephenson *et al.*, 1991). Another may be that contour linking may be totally absent for peripheral processing (Hess & Dakin, 1997). Whatever the precise reasons, it is generally believed that peripheral vision is much more appropriate for signalling potentially interesting transients than for form vision. For example, studies that have examined which aspects of peripherally presented stimuli are important for deciding where the eye will move next during scene exploration have almost invariably failed to demonstrate an impact of the 'meaning' of the stimulus. Instead, eye movement programming seems to depend essentially on relatively low-level visual attributes, such as size, contrast and motion (Henderson & Hollingworth, 1999).

However, in the present experiment subjects scored above 70% when images were centred at 57.5° of eccentricity and were still above 60% at an eccentricity of 70.5° . One reason may lie in the choice of stimuli and task. Previous studies of peripheral form vision have used artificial stimuli, such as letters that require not only high acuity, but also accurate phase information. The large size of the images used here (39° by 26°), together with their wide spatial frequency content may mean that, even at low resolution, there is enough information available for categorization to occur. The use of large stimuli also meant that we were able to test performance at much higher eccentricities than have been used in the past. Moreover, the natural stimuli used here are clearly of vital importance for survival; even a limited ability to detect the presence of an animal in the peripheral visual field would have had a very high survival advantage. We have reported recently that this form of rapid visual categorization is not restricted to biological forms, as processing is just as fast and accurate using 'means of transport' as the target category as with our more standard 'animal' task (VanRullen & Thorpe, 2001a,b). However, it should be noted that those experiments were performed using presentations limited to central vision, and so it remains to be seen whether the ability to categorize objects in peripheral vision extends to nonbiological categories. Until such experiments have been performed we cannot rule out the possibility that the ability of subjects to detect animals in the peripheral visual field demonstrated here reflects the activity of a biologically primitive object processing system that cannot be used for more arbitrary stimulus forms.

Another reason for the good performance reported here might be the use of a task in which subjects were strongly encouraged to rely on their first impression. As noted earlier, subjects were typically unable to report precisely *what* they had seen, even when they were performing the task reliably. Some of our recent results suggest that perceptual and semantic priming can occur at 30° of eccentricity, but that only perceptual priming is observed at 60° eccentricity (Boucart *et al.*, 2001). Such findings are reminiscent of reports of blindsight in brain-damaged subjects who, when forced to guess, can often score well above chance despite having little or no phenomenal perception of the stimulus (Weiskrantz, 1996; Sahraie *et al.*, 1997; Stoerig & Cowey, 1997). However, blindsight is generally considered to be restricted to only a very limited range of visual attributes, including presence or absence of a visual stimulus, and/or discrimination of movement, simple forms and colour. Our task presumably requires more complex processing. The present results can also be related to other studies that have revealed extensive visual processing in the absence of conscious awareness (Mack & Rock, 1998). These have included evidence for object identification in subjects with both parietal neglect (Driver & Mattingley, 1998) and blindsight (Marcel, 1998), as well as unconscious semantic priming in normal subjects (Dehaene *et al.*, 1998) or unconscious priming by faces in prosopagnosia (Young *et al.*, 1988).

The fact that subjects reported not being able to identify the animal they had detected in the image raises another important question. A number of studies have shown that, in central vision, objects are reported initially as instances of a 'basic level' (Rosch *et al.*, 1976) or an 'entry level' (Jolicoeur *et al.*, 1984) category. But to identify faster a penguin as a penguin (rather than as a bird or an animal), perceptual categorization at the 'entry level' probably requires more detailed (foveal?) analysis, whereas coarse cursory processing of visual information might be enough to allow superordinate categorization. Thus, our results can be interpreted in two ways. Either the concept of 'entry level' does not apply equally to central and peripheral vision or our subjects had more information about the animal that they had detected than they reported. This could be tested by further experiments that specifically analyse the ability to perform basic level categorizations in the far periphery.

The view that animal categorization may involve processing that is largely automatic is reinforced by a further aspect of the present results. The use of very brief presentations (28 ms) ruled out any possibility of making controlled eye movements, and because stimulus position was randomized, subjects needed to monitor nearly the entire horizontal extent of the visual field. Nevertheless, there was essentially no performance loss associated with spreading attention so widely in the visual field. This can be inferred from the fact that accuracy for centrally presented images was virtually identical to that reported in earlier studies where stimulus position was totally predictable (Thorpe *et al.*, 1996). In fact, this result confirms an earlier study that compared the standard one-position task with one in which the test image was presented randomly at one of three positions: centrally, or 3.6° to the left or right of the fixation point (Fabre-Thorpe *et al.*, 1998). Here again there was no effect of stimulus position uncertainty. But in the present task, uncertainty about stimulus position was effectively infinite, because the target could be positioned anywhere within the photograph, and the photograph could be presented almost anywhere within the horizontal extent of the visual field.

Such results therefore seem at odds with the conventional view that directed attention is required for high-level form vision. For many visual tasks, directed visual attention appears not only important, but also often essential. For example, while preattentive processing can

allow many features of the visual scene to be processed in parallel, binding those features together to form an object appears to require attention (Wolfe & Bennett, 1997; Treisman, 1998). Furthermore, recent data have argued strongly that the ability of subjects to notice significant changes in a scene depends critically on attention, in that subjects fail to notice even major changes unless they are actively attending to the particular place (Rensink *et al.*, 1997; Mack & Rock, 1998; O'Regan *et al.*, 1999). This might be taken to imply that high-level visual processing *requires* attention to be focused in the right place. Indeed, the idea that attention is used in our task would provide an alternative explanation of the increase in RTs seen with peripheral presentations. Because the images in the present experiments were not masked, they might leave an iconic representation even after they have disappeared. Thus, increasing the eccentricity of the stimulus might result in an increase of the time required to redirect attention.

However, as suggested earlier, the increase in RT with eccentricity can also be explained simply by the smaller number of retinal ganglion cells activated by images presented in the periphery. If this is the case, it might be that directed attention is not required specifically for the sort of high-level categorization tasks used here, and that subjects can still categorize objects when attention is spread diffusely across the entire visual field. According to this view, directed attention would only be required in cases where the visual scene contains many different and competing objects. Attention may specifically be required to *inhibit* the processing of irrelevant information (Kastner *et al.*, 1998; Milliken & Tipper, 1998). In the present series of experiments, although images could be presented at any position in the visual field, only one image was presented at a time. As a consequence, there may be no requirement to selectively filter out information from other positions in the visual field, as might well be the case if the entire field was stimulated simultaneously or if competing images flanking the targets had been flashed simultaneously. However, many of the images contained more than one object of interest so that objects were in competition within a given stimulus. Performance would have probably been even better if we had used isolated animals presented on a uniform background. While the present results suggest that high-level descriptions of images could be reached without specifically directing attention to the location of the stimulus, this does not contradict reports that directed attention is required to consciously register the information.

Together, the present results imply that certain forms of high-level scene analysis may be possible on the basis of primitive processing strategies that can operate on very limited visual information, in the absence of clear conscious perception and quite possibly without a need for spatially directed attention. In this respect, it is interesting to note that the same task can be performed accurately with monochromatic images, ruling out the possibility of a simple colour-based strategy (Delorme *et al.*, 2000). Because the task involves a superordinate categorization of static images in which the shape, the size and the number of targets can vary within a large range, no single feature would be sufficient to explain the results. It may be that this sort of task could be performed using a large number of mechanisms operating in parallel, each looking for a relatively simple combination of features. Such fast and relatively coarse processing could be used as an 'intelligent header' to orient attention and improve further processing of the fine details needed for precise identification. The fact that subjects are able to perform considerably better than they expect has obvious implications for pathologies, such as macular degeneration in which patients have no option other than relying on their residual peripheral visual information. The results of the present experiments suggest that there is considerable potential for high-level visual processing in the peripheral visual field and that

such patients might well be encouraged to make better use of such capacities.

Acknowledgements

This work was supported by the CNRS and by the GIS 'Science de la Cognition'. Karl R. Gegenfurtner was supported by a Heisenberg-Fellowship from the Deutsche Forschungsgemeinschaft (Ge 879/4-1).

Abbreviation

RT, reaction time.

References

- Antal, A., Keri, S., Kovacs, G., Janka, Z. & Gyorgy, B. (2000) Early and late components of visual categorization: an event-related potential study. *Cogn. Brain Res.*, **9**, 117–119.
- Bennett, P.J. & Banks, M.S. (1987) Sensitivity loss in odd-symmetric mechanisms and phase anomalies in peripheral vision. *Nature*, **326**, 873–876.
- Biederman, I. & Cooper, E.E. (1991) Object recognition and laterality: null effects. *Neuropsychologia*, **29**, 685–694.
- Biederman, I., Mezzanotte, R.J., Rabinowitz, J.C., Francolini, C.M. & Plude, D. (1981) Detecting the unexpected in photointerpretation. *Human Factors*, **23**, 153–164.
- Boucart, M., Fabre-Thorpe, M., Thorpe, S., Arndt, C. & Hache, J.-C. (2001) Covert object recognition at large visual eccentricity. Abstract from the conference of the Vision Science Society (2001). *J. Vision*, in press.
- Chung, S.T.L., Mansfield, J.S. & Legge, G.E. (1998) Psychophysics of reading. XVIII. The effect of print size on reading speed in normal peripheral vision. *Vision Res.*, **38**, 2949–2962.
- Curcio, C.A. & Allen, K.A. (1990) Topography of ganglion cells in human retina. *J. Comp. Neurol.*, **300**, 5–25.
- Dehaene, S., Naccache, L., LeClecq, H.G., Koechlin, E., Mueller, M., Dehaene-Lambertz, G., vandeMoortele, P.F. & LeBihan, D. (1998) Imaging unconscious semantic priming. *Nature*, **395**, 597–600.
- Delorme, A. (2000) Traitement visuel de scènes naturelles chez le singe, l'homme et la machine: une vision qui va de l'avant. Unpublished Thesis, Université Paul Sabatier, Toulouse.
- Delorme, A., Richard, G. & Fabre-Thorpe, M. (2000) Ultra-rapid categorisation of natural images does not rely on colour: a study in monkeys and humans. *Vision Res.*, **40**, 2187–2200.
- Drasdo, N. (1991) Neural substrates and threshold gradients of peripheral vision. In Kulikowski, J.J., Walsh, V., Murray, I.J. (eds), *Limits of Vision*, pp. 251–265. Macmillan, London.
- Driver, J. & Mattingley, J.B. (1998) Parietal neglect and visual awareness. *Nature Neurosci.*, **1**, 17–22.
- Fabre-Thorpe, M., Delorme, A., Marlot, C. & Thorpe, S. (2001) A limit to the speed of processing in Ultra-Rapid Visual Categorization of novel natural scenes. *J. Cogn. Neurosci.*, **13**, 171–180.
- Fabre-Thorpe, M., Fize, D., Richard, G. & Thorpe, S.J. (1998) Rapid categorization of extrafoveal natural images: Implications for biological models. In Bower, J. (ed.), *Computational Neuroscience: Trends in Research 1998*. New York, NY, Plenum Press, pp. 7–12.
- Henderson, J.M. & Hollingworth, A. (1999) High-level scene perception. *Annu. Rev. Psychol.*, **50**, 243–271.
- Hess, R.F. & Dakin, S.C. (1997) Absence of contour linking in peripheral vision. *Nature*, **390**, 602–604.
- Jolicoeur, P., Gluck, M.A. & Kosslyn, S.M. (1984) Pictures and names: making the connection. *Cogn. Psychol.*, **16**, 243–275.
- Jüttner, M. & Rentschler, I. (2000) Scale-invariant superiority of foveal vision in perceptual categorization. *Eur. J. Neurosci.*, **12**, 353–359.
- Kastner, S., De Weerd, P., Desimone, R. & Ungerleider, L.G. (1998) Mechanisms of directed attention in the human extrastriate cortex as revealed by functional MRI. *Science*, **282**, 108–111.
- Koenderink, J.J., Bouman, M.A., Bueno de Mesquita, A.E. & Slappendel, S. (1978) Perimetry of contrast detection thresholds of moving spatial sine patterns. II. The far peripheral visual field (eccentricity 0 degrees–50 degrees). *J. Opt. Soc. Am.*, **68**, 850–854.
- Levi, D.M., Klein, S.A. & Sharma, V. (1999) Position jitter and undersampling in pattern perception. *Vision Res.*, **39**, 445–465.
- Levy, I., Hasson, U., Avidan, G., Hendler, T. & Malach, R. (2001) Center-

- periphery organization of human object areas. *Nature Neurosci.*, **4**, 533–539.
- Liversedge, S.P. & Findlay, J.M. (2000) Saccadic eye movements and cognition. *Trends Cogn. Sci.*, **4**, 6–14.
- Loftus, G.R. & Mackworth, N.H. (1978) Cognitive determinants of fixation location during picture viewing. *J. Exp. Psychol. [Hum Percept]*, **4**, 565–572.
- Mack, A. & Rock, I. (1998) *Inattention blindness*. MIT Press, Cambridge, MA.
- Marcel, A.J. (1998) Blindsight and shape perception: deficit of visual consciousness or of visual function? *Brain*, **121**, 1565–1588.
- Milliken, B. & Tipper, S.P. (1998) Attention and inhibition. In Pashler, H. (ed.), *Attention*. Psychology Press, Hove, pp. 191–222.
- Nasanen, R. & O'Leary, C. (1998) Recognition of band-pass filtered handwritten numerals in foveal and peripheral vision. *Vision Res.*, **38**, 3691–3701.
- Nelson, W.W. & Loftus, G.R. (1980) The functional visual field during picture viewing. *J. Exp. Psychol.: Human Percept. Perform.*, **6**, 391–399.
- O'Regan, J.K., Rensink, R.A. & Clark, J.J. (1999) Change-blindness as a result of 'mudsplashes'. *Nature*, **398**, 34.
- Rensink, R.A., O'Regan, J.K. & Clark, J.J. (1997) To see or not to see: The need for attention to perceive changes in scenes. *Psychol. Sci.*, **8**, 368–373.
- Rentschler, I. & Treutwein, B. (1985) Loss of spatial phase relationships in extrafoveal vision. *Nature*, **313**, 308–310.
- Rosch, E., Mervis, C.B., Gray, W.D., Johnson, D.M. & Boyes-Braem, P. (1976) Basic objects in natural categories. *Cogn. Psychol.*, **8**, 382–439.
- Sahraie, A., Weiskrantz, L., Barbur, J.L., Simmons, A., Williams, S.C. & Brammer, M.J. (1997) Pattern of neuronal activity associated with conscious and unconscious processing of visual signals. *Proc. Natl Acad. Sci. USA*, **94**, 9406–9411.
- Stephenson, C.M., Knapp, A.J. & Braddick, O.J. (1991) Discrimination of spatial phase shows a qualitative difference between foveal and peripheral processing. *Vision Res.*, **31**, 1315–1326.
- Stoerig, P. & Cowey, A. (1997) Blindsight in man and monkey. *Brain*, **120**, 535–559.
- Strasburger, H., Harvey, L.O. Jr & Rentschler, I. (1991) Contrast thresholds for identification of numeric characters in direct and eccentric view. *Percept. Psychophys.*, **49**, 495–508.
- Strasburger, H., Rentschler, I. & Harvey, L.O. Jr (1994) Cortical magnification theory fails to predict visual recognition. *Eur. J. Neurosci.*, **6**, 1583–1587.
- Thorpe, S., Fize, D. & Marlot, C. (1996) Speed of processing in the human visual system. *Nature*, **381**, 520–522.
- Treisman, A. (1998) The perception of features and objects. In Wright, R.D. (ed.), *Visual Attention*. Oxford University Press, Oxford, pp. 26–54.
- VanRullen, R. & Thorpe, S.J. (2001a) The time course of visual processing: from early perception to decision making. *J. Cogn. Neurosci.*, **13**, 454–461.
- VanRullen, R. & Thorpe, S.J. (2001b) Is it a bird? Is it a plane? Ultra-rapid categorisation of natural and artificial objects. *Perception*, **30**, 655–668.
- Wässle, H., Grunert, U., Rohrenbeck, J. & Boycott, B.B. (1990) Retinal ganglion cell density and cortical magnification factor in the primate. *Vision Res.*, **30**, 1897–1911.
- Weiskrantz, L. (1996) Blindsight revisited. *Curr. Opin. Neurobiol.*, **6**, 215–220.
- Wolfe, J.M. & Bennett, S.C. (1997) Preattentive object files: shapeless bundles of basic features. *Vision Res.*, **37**, 25–43.
- Young, A.W., Hellawell, D. & De Haan, E.H. (1988) Cross-domain semantic priming in normal subjects and a prosopagnosic patient. *Q. J. Exp. Psychol. [a]*, **40**, 561–580.