Parallel processing in high-level categorization of natural images

Guillaume A. Rousselet, Michèle Fabre-Thorpe and Simon J. Thorpe

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

Correspondence should be addressed to G.A.R.(guillau@cerco.ups-tlse.fr)

Published online: 28 May 2002, doi:10.1038/nn866

Models of visual processing often include an initial parallel stage that is restricted to relatively low-level features, whereas activation of higher-level object descriptions is generally assumed to require attention[1–4]. Here we report that even high-level object representations can be accessed in parallel: in a rapid animal versus non-animal categorization task, both behavioral and electrophysiological data show that human subjects were as fast at responding to two simultaneously presented natural images as they were to a single one. The implication is that even complex natural images can be processed in parallel without the need for sequential focal attention.

High-order representations, up to the semantic level, can be accessed very rapidly from brief picture presentations[5,6]. Event-related potential (ERP) experiments show that complex processing of natural scenes is achieved 150 ms after stimulus onset[7]. Thus, when humans are asked to decide whether a briefly presented photograph contains an animal, the ERPs in response to targets and distractors diverge sharply from 150 ms. There is evidence that these differences reflect a real visual decision rather than physical differences between stimulus categories[8]. The scenes used in such experiments typically contain several objects, suggesting that there is at least some degree of parallelism in the underlying processing. To explore this issue, we analyzed whether processing speed is affected when subjects are asked to process two pictures simultaneously.

Twenty subjects (mean age, 32.5 ± 10.9) performed a modified version of the animal versus non-animal go/no-go task used in previous studies[7,8] (see **Supplementary Fig. 1** and **Supplementary Methods** online). In 20 blocks of 96 trials, single brief presentations (20 ms) of one image appearing 3.6° to the left or right of a central fixation point were randomly mixed with the same number of dual presentations in which two images were flashed simultaneously at the same eccentricities. In both conditions, an animal target was presented on half of the trials. Target location (left versus right hemifield) was equiprobable.

Notably, subjects were able to process dual and single presentations at the same speed (**Fig. 1a**). This is shown by both the median reaction times (RTs, 390 versus 391 ms, respectively) and by the latencies of the earliest responses which were equal or shorter with two images than with one image (means of 255 versus 260 ms, respectively; see **Supplementary Table 1** online).

Subjects tended to be more accurate in the one-image condition (90.4%) than with dual images (86.7%). This accuracy decrease was predicted by a simple parallel model of processing (**Fig. 1b**) in which each of two simultaneously presented images is processed by a separate and independent mechanism, and both mechanisms eventually converge on a single output system (see **Supplementary Methods**). Further support for a parallel processing model comes from the tight fit between the experimental and the predicted cumulative performance accuracy (*d'*) curves (**Fig. 1b**).

The similarity in processing speed between the two conditions was confirmed by electrophysiological data (**Fig. 2**). Associated ERPs were averaged off-line for each condition and difference waves were obtained by subtracting the ERP for correct distractor trials from the ERP for correct target trials. Differential activation, probably generated within high-order extrastriate visual areas[9], was clearly seen at both occipito-temporal and frontal sites (see **Supplementary Fig. 2** online). There was no effect of image condition on the onset of the differential occipital activity. Target and distractor signals diverged sharply around 140–150 ms after stimulus onset with an enhanced occipital negativity on target trials. This differential occipital activity became significant at similar latencies in both

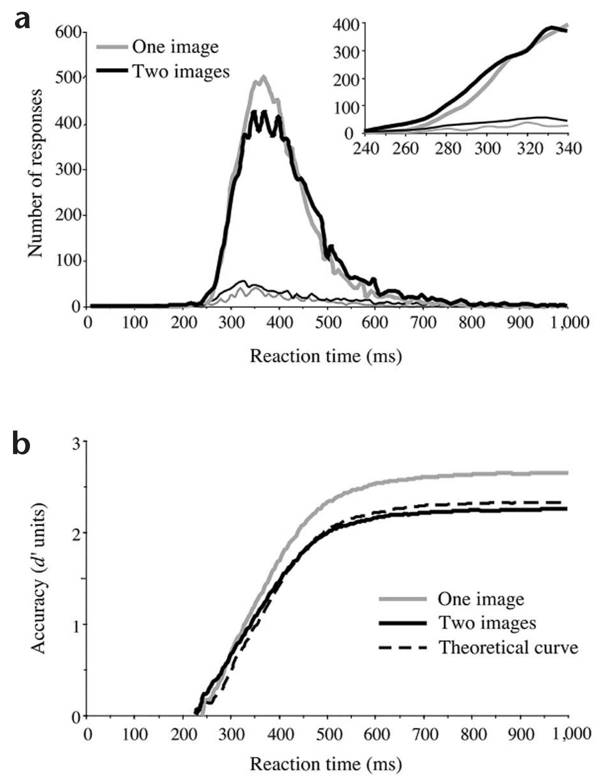

Fig. 1. Behavioral results. (**a**) Reaction time distributions. Number of responses are expressed over time, with time bins of 5 ms. Correct responses or 'hits' (thick top curves) are shown for the one target alone (gray) or for the target flanked by distractor (black). False alarms (thin bottom curves) are shown for the one distractor alone (gray) or for the two distractors (black). (**b**) Performance time course functions and predictions of a parallel model of processing. Average performance accuracy (in *d'* units) is plotted as a function of processing time (in ms) for one image (gray curve) and for two images (black curve). The dynamic *d'* was calculated from the cumulative number of hits and false alarms at each successive 10 ms time step. The predicted curve from the model was calculated using the probabilities of hits and false alarms calculated from the experimental data in the one-image condition. A global fall in accuracy from 90.4% in the one-image condition to 87.7% in the two-image condition was predicted by our model (see **Supplementary Methods**). The experimental procedures were authorized by the local ethical committee (CCPPRB No. 9614003) and all subjects gave informed consent to participate.

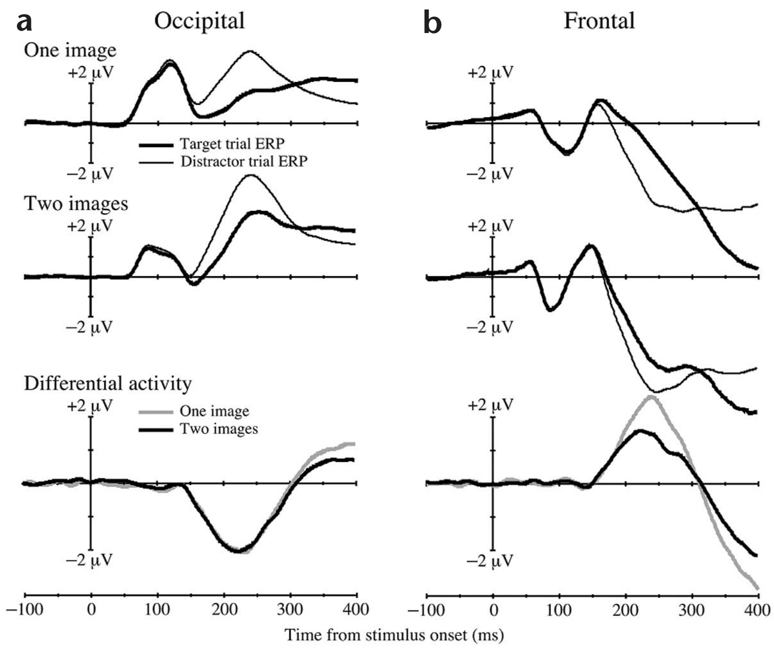

Fig. 2. Grand average ERPs and associated differential activities. Grand average ERPs are plotted for correct target trials (thick line) and for correct distractor trials (thin line). Results are shown for contralateral occipital electrodes (**a**) and all frontal electrodes (**b**), for the one-image (top panel) and the two-image (middle panel) conditions. Bottom, differential activity between the one- (gray) and two-image (black) conditions.

conditions (152 ms with one image versus 150 ms with two images, $P < 0.0005$) and then developed at the same rate, with the same slope and amplitude in both conditions.

Differential activity was also seen at frontal sites (**Fig. 2b**), starting at around 160–170 ms in both conditions and becoming significant ($P < 0.0005$) at about the same latency: 173 ms (one image) and 175 ms (two images). At 190 ms after stimulus onset, the differential activity recorded in the one-image condition began to diverge from that of the two-image condition, developing with a steeper slope and finally reaching a higher amplitude.

These behavioral and electrophysiological results provide strong evidence that processing speed is unchanged between the one- and two-image conditions. Furthermore, the slight accuracy impairment (<4%) with two images can be explained using a very simple model in which the two images are processed by separate mechanisms that pool their outputs. The brief image presentations and initial lateralization of visual inputs to the contralateral striate visual cortex indicate that each hemisphere could work in parallel on a different visual scene. This interpretation is strengthened by the high lateralization of the differential occipital activity.

The RT distributions (**Fig. 1a**) show that the number of 'go' responses in the two-image condition, although initially similar to that seen in the one-image condition, was considerably lower around the mean RTs. This effect might be explained by some form of competitive process occurring in the two-image condition. Given the strong similarity between the occipito-temporal differential activity in the two conditions (**Fig. 2a**), it seems unlikely that this competition affects the initial visual processing. Competition is more likely to occur later on at the point of 'sensorimotor decision'[10]. Evidence for a late competitive process at frontal sites comes from the late divergence seen between the one- and two-image conditions after 190 ms. High-level representations in occipito-temporal visual areas would be activated independently in each hemisphere. At frontal sites, by contrast, when integration of the outputs of the two cerebral hemispheres is needed for decision-making, competition could result from frontal processes related either to category-specific decision-making[11] or to response inhibition on no-go trials[12].

These physiological results directly support psychological models in which the competitive bottleneck is situated at a high level of integration[13,14]. Further support comes from behavioral findings that show that this animal versus non-animal categorization task can be done simultaneously with another attentionally demanding task[15]. We have no evidence, however, to suggest that the ability to process more than one image at the same time extends beyond the specific case of two images presented simultaneously in the two hemifields. Further experiments will be required to explicitly test whether the system can process other stimulus arrangements in parallel, such as two images presented within the same hemifield or four images presented simultaneously.

Taken together, our data show that high-level object categorization of natural scenes can be done in parallel very rapidly and without the need for sequential focal attention. Whereas classic models of allocation of attentional resources consider 'early' vision as being early in complexity and restrict low-level vision to the lower part of the cortical hierarchy (namely V1 and V2), early vision might more appropriately be considered as processing that is early in time.

Note: Supplementary information is available on the Nature Neuroscience website.

Acknowledgments

This work was supported by the Cognitique program (COG35 and 35b). Financial support was provided to G.A.R. by a PhD grant from the French Government.

Competing interests statement

The authors declare that they have no competing financial interests.

RECEIVED 24 JANUARY; ACCEPTED 29 APRIL 2002

1. Treisman, A. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **353**, 1295–1306 (1998).
2. Wolfe, J. M. *Vis. Res.* **34**, 1187–1195 (1994).
3. Kinchla, R. A. *Annu. Rev. Psychol.* **43**, 711–742 (1992).
4. McElree, B. & Carrasco, M. *J. Exp. Psychol. Hum. Percept. Perform.* **25**, 1517–1539 (1999).
5. Potter, M. C. *J. Exp. Psychol. [Hum. Learn.]* **2**, 509–522 (1976).
6. Biederman, I. *Science* **177**, 77–80 (1972).
7. Thorpe, S., Fize, D. & Marlot, C. *Nature* **381**, 520–522 (1996).
8. VanRullen, R. & Thorpe, S. J. *J. Cogn. Neurosci.* **13**, 454–461 (2001).
9. Fize, D. *et al. Neuroimage* **11**, 634–643 (2000).
10. Schall, J. D. *Nat. Rev. Neurosci.* **2**, 33–42 (2001).
11. Freedman, D. J., Riesenhuber, M., Poggio, T. & Miller, E. K. *Science* **291**, 312–316 (2001).
12. Sasaki, K., Gemba, H., Nambu, A. & Matsuzaki, R. *Neurosci. Res.* **18**, 249–252 (1993).
13. Duncan, J. *Psychol. Rev.* **87**, 272–300 (1980).
14. Chun, M. M. & Potter, M. C. *J. Exp. Psychol. Hum. Percept. Perform.* **21**, 109–127 (1995).
15. Li, F. F., VanRullen, R., Koch, C. & Perona, P. *Proc. Natl. Acad. Sci. USA* (in press).