Critique: 5/5
Future work: 5/5

Neuro 230 Visual Object Recognition HW1

Comments on *Speed of processing in the human visual system. Nature (1996)*

Chia-Jung Chang

[PAPER SUMMARY]

This paper gave an estimate of processing time for fast visual categorization of animals. While subjects were asked to decide whether a 20-msec flashed natural photograph contains an animal, differential event-related potential (ERP) activity between target trials and distractor trials was observed in frontal recording sites at around 150 msec after stimulus onset. The authors thus claimed that this differential response should rise once necessary visual processing has been completed, and that such rapid visual categorization could mainly rely on feedforward mechanisms. These claims, however, are arguable due to the following issues in their hypothesis and experimental design.

[POSSIBLE ISSUES]

One comment from the beginning: way too long! Many good ideas, but for the future please try to focus on maybe 1-3 for each section that you think are most compelling.

Stimuli selection

First of all, animals can be considered as a special object category for their biological relevance. It is very likely that our brain has developed a specific network to deal with this kind of stimuli, which do not apply to other object categories. Therefore, it will be crucial to demonstrate whether this differential response is observed for other object categories as well.

While this point may limit the generality of the results, is there another reason you would want to do this?

Secondly, although each stimulus was shown only once to subjects, it is unclear how familiar the subjects were with the targets and scenes. If the subject is so familiar with animals, he could have developed a neural network selective to corresponding object category. Therefore, it will be necessary to show whether such processing is biased by familiarity and expertise to the object categories.

Good point - how would you do this?

More importantly, although the authors used a large set of natural color photographs, the systematic difference between animal and non-animal images was not ruled out as the authors assumed. Subjects could implicitly learn the correlation between scenes and target animals. For example, the possibility of animals presented within buildings is extremely low. In this case, even with briefly flashed stimuli, subjects could benefit from the contexts in which animals were presented, which led to a faster process.

Furthermore, the image statistics were not controlled in this study. It is possible that the contrast in photographs with targets was higher than ones with distractors, and thus biased the visual processing. The difference between complexity of photographs with targets and distractors could also lead to the response difference. Therefore, it is necessary to control the feature statistics (regularity, contrast, etc.) for all the stimuli.

Last but not least, the results could have been biased by the animal size, orientation and location, which were not described in this paper. While an animal is always shown with similar size in the middle of a photograph, it is possible to train a classifier using just the outputs of cortical neurons at the initial stage of visual processing, and still get reasonable discrimination ability. Therefore, it is necessary to diversify the animal size, orientation and location in the stimulus set to study visual categorization.

While they don't mention it in the paper, the images used actually did have quite a diversity of sizes, viewpoints, etc.

Trial selection

Although the proportion of correct responses was high across all the subjects in this experiment, the fact that the authors only examined trials with correct reports makes it difficult to separate decision-making process from visual object processing. It is very likely that the visual processing is done in an even faster way and differential response occurs much later afterwards. Therefore, the authors should disassociate between subject reports and visual stimuli to see whether a differential response is purely depending on the image difference, or biased by the decisions.

Besides, authors only looked at the average response across all the trials instead of providing single-trial activity. If the differential response fully reflects when necessary visual processing is completed, there should not be much variance between trials in the initial stage and later stage. We should also be able to observe differential response on single trials because information should be extracted on each trial.

Single trials of EEG are incredibly noisy. It would have been tough, especially with the technology at the time.

Electrode selection

The authors selected 7 electrodes out of 20 electrodes to show the differential ERP between targets and distractors. The average differential ERP curves of 15 subjects had more temporal variability than what the authors claimed. It is thus unclear whether these 7 electrodes were responsible for general visual categorization. If we randomly assign photographs into two different categories, it is possible that the differential response can still be observed by some electrodes. Therefore, it is necessary to show ERP data for all electrodes with a stimulus permutation test.

Good point. However, the 7 electrodes were the same across subjects, and were spatially contiguous, so concern about a "fishing expedition" isn't nearly as big as it might have been otherwise.

Feedback involvement

The differential activity observed in frontal recording sites might not represent the first responses in feedforward streams, but could reflect changes in activity produced by feedback from other cortical areas. It is possible that the differential response in the lower-level of visual system was not detected with ERP for the low spatial resolution. A recent study (Kirchner et al. 2009) demonstrated that visual inputs can reach the frontal eye fields at around 50 msec with intracerebral recordings. As a result, the differential ERP observed in this study could already involve feedback modulation.

Absolutely, it could reflect information reaching frontal regions directly from early visual or subcortical regions (e.g. Moshe Bar's work)

[FUTURE DIRECTIONS]

According to the interesting results observed in this study and the possible issues in their experimental set up, the follow-up experiments to further test the characteristics and limits of rapid visual categorization are described as below.

Object context

The same experiment can be tested with different object categories such as vehicle to rule out the concerns on biological relevance. Objects at basic levels (a car or a bike) or subordinate levels (a race car or a jeep) can also be examined to see whether visual information processing speed changes with detailed features.

Objects can also be categorized with familiar and unfamiliar items to study whether there is a particular cortical network in which experience could shape the strength of mutual interconnections, which make the processing even faster.

The next follow-up experiment should have a set of visual stimuli with similar image statistics. Instead of natural photographs, we can use synthetic stimuli to reduce the bias from the context. For example, we can synthesize images of an animal in a kitchen, or a laptop in a forest. We can even make all the photographs have similar background scenes, and all the foreground objects have similar features. Then, a performance bias depending on whether the context is congruent or not with target would be eliminated.

Besides, the variability of object from a same category should also be included to avoid classifier learning problem, and see whether the differential response corresponding to objects with variations in orientations, sizes, contrasts, and locations.

Attention effect

A question rises with rapid visual categorization is that whether such behavior requires attention engagement or it could rely on object representations without attention. This could be tested whether object categorization could be performed with images presented in visual periphery, or two images presented simultaneously. If similar performance and neural response can still be observed, then we know this rapid visual categorization can be done without focused attention.

Nice idea! The periphery confounds attention with acuity, but of course there are other ways to redirect the deployment of attention.

Processing stage

It is possible that neural activity well before the 150-msec frame can be useful for encoding the category of a stimulus. This could be tested with two images flashed simultaneously in left and right hemi-fields, and saccade to the one with the target. If visual stimulus processing is achieved much earlier than differential ERP is observed, then the saccade would occur earlier than 150 msec.

This has been done!

It would also be helpful to see how much information is extracted in single trials, rather than averaging data across multiple trials. This could be done with bootstrapping the trials, randomizing the trials within a block, or analyzing the variability across trials.

Cortical pathway

While this study suggested a role of frontal areas in inhibiting inappropriate behavioral responses during object categorization, the electrode selection issue and the coarse spatial resolution of recordings limit further interpretations. More experiments with better spatial resolution such as fMRI and MEG would benefit finding the responsible cortical organizations for rapid object categorization. Single-unit recordings could also provide more detailed information on when and where the necessary visual input is processed completely before a cognitive decision is involved.

Another related question would be which pathway our visual system utilizes for rapid object categorization. By varying the image color and intensity contrasts, we can study whether it is magnocellular (M) or parvocellular (P) pathway responsible for this rapid visual object categorization, with assistance of different recording techniques.

This has also been done!