The discrimination of correlated and anti-correlated motion in the human visual system

Emily R. Lindemer*[1], Ali Syed[2]

[1]Department of Neuroscience, McGill University, McIntyre Medical Building, 3655 Promenade Sir William Osler, Montreal, Quebec, Canada, H3G 1Y6
[2]Department of Biochemistry, McGill University, McIntyre Medical Building, 3655 Promenade Sir William Osler, Montreal, Quebec, Canada, H3G 1Y6

ABSTRACT

Introduction: How the brain integrates spatial and temporal information is not known. This issue is referred to as the "binding problem" of visual perception. It has been proposed that groups of neurons which correspond to the same elements of an image become synchronous in order to form a coherent neural representation; however, direct experimental evidence supporting this role for neural synchrony is highly controversial. As our perceptual capabilities are limited by the neural mechanism that supports them, an alternative approach to understanding neural synchrony is to instead characterize our ability to perceive synchrony. Thus, our aim was to demonstrate how correlated (i.e., synchronous) motion is perceived by the brain and how its discrimination can be enhanced or impaired. **Methods**: In this study, we used human psychophysics experiments to characterize the ability of subjects to discriminate synchrony in a moving visual stimulus. **Results**: By varying stimulus length, motion speed, and direction we found that humans were less than optimal in their ability to discriminate correlated motion when compared to an ideal mathematical model. In addition, we found that the length of the entire stimulus was not an important factor, but the length of individual motion pulses which made up the stimulus was crucial to performance. **Discussion**: Overall these results suggest that neural synchrony is likely used by the brain, but its resolution is highly limited compared to an ideal model.

KEYWORDS

Temporal structure, Binding problem, Correlation

*Corresponding author:
emily.lindemer@mail.mcgill.ca
Received: 3 January 2010
Revised: 7 March 2010

INTRODUCTION

To understand images as a whole, the brain must break down incoming visual information into basic categories. Many neurons work together in a complex network to integrate local elements into one coherent image; however, the method that the brain uses to break down visual information and piece it back together is still unclear. While several theories have attempted to reconcile this *binding problem* issue (*1*), one commonly accepted theory is known as *binding by synchrony* (BBS). *Binding by synchrony* suggests that detectors in the brain correspond to different low-level characteristics of an object, such as color, contrast, shape, and spatial and temporal orientation (*2*). Neurons which respond to each of these categories are able to fire in synchrony, integrating to create a meaningful interpretation of the object as a whole.

Synchrony of motion falls into the category of 'temporal orientation' and plays a very important role in visual perception. There are two ways in which this can be interpreted. Motion can either be correlated in the sense of *temporal structure*, or by means of *temporal synchrony*. When motion is correlated via *temporal structure*, the individual elements of a scene follow the same relative motion patterns over time. If, on the other hand, motion is correlated by means of *temporal synchrony*, each element of a scene changes its motion pattern at the exact same time, in the exact same direction. Research has found that the human visual system prefers to use *temporal structure* as a means of integrating synchronous motion when dealing with the binding problem (*3*). Directional motion is also one of the lower-level binding categories within temporal orientation, and is therefore more quickly and easily processed than higher-level properties such as colour (*4*). This study supports the preference of using temporal structure to detect synchronous motion. Stimuli were manipulated to test different aspects of synchronous motion, such as speed and length of presentation in human psychophysics experiments. An optimal model was then developed, which was compared against human performance to determine enhancements or impairments in the discrimination of correlated motion. Our stimuli consisted of dot patches moving together horizontally in pulses of different time lengths. Based on previous studies, it was expected that human performance would be enhanced with either the length or number of motion pulses in the stimuli (*5*). The question also remains whether there is a point at which human performance can no longer increase, and whether this plateau is at the same level as the optimal model.

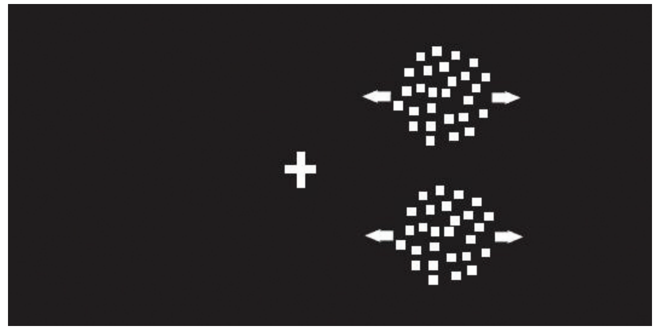

Fig. 1. Dot patches of equal density remained stationary at a fixed, equal distance from the central focal point. The dots exhibited horizontal motion according to the level of correlation specified by the experiment conductor.

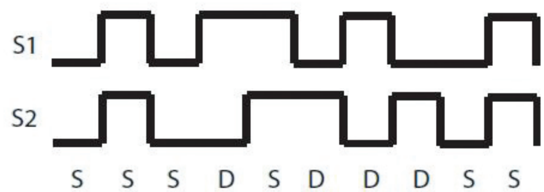

Fig. 2. Two binary sequences were generated for each trial, one corresponding to each dot patch. If the motion of the two patches was the same (S) then the motion patterns were considered "correlated." If the motion of the two patches was different (D) the motion patterns were considered "anti-correlated." Ten of these random pulses were generated between the two patches for each trial, with more S pulses in a trial with a higher correlation level and more D pulses in a trial with a lower correlation level.

METHODS AND MATERIALS

The stimulus consisted of two target patches that were 3 cm in diameter, each made up of random dot patterns of equal densities. The two patches were placed equidistantly with their centers 8 cm to the right of and one 3 cm above the centre focal point of the screen, and one 3 cm below, as shown in Figure 1. Motion sequences were generated for each of the two dot patches using two binary sequences. Each binary value corresponded to the direction of one motion pulse with a value of 1 corresponding to rightward motion and a value of 0 corresponding to leftward motion. Each of the two sequences was assigned to one of the dot patches, whose dots then moved accordingly. Random motion sequences were generated for the dot patches, depending on a level of correlation between -1 and 1 that was indicated by the experiment conductor (*6*). In the motion patterns, all dots in one patch moved together while the patch as a whole remained stationary. The correlation level between the two patches was then observed. A correlation of 1 corresponded to perfect synchrony of motion pattern between the two patches and a correlation of -1 corresponded to motion between the two patches that was perfectly opposite (Fig. 2). A 17-inch CRT monitor was used with 1280x1024 resolution and subjects were set at a distance of 60 cm from the screen. Random pulse sequences between the dot patches were generated and displayed using the Psychophysics Toolbox application for Matlab™.

Data was collected from human subjects using a two-alternative forced choice (2AFC) motion discrimination task in which the subjects were asked to respond to the motion pattern of the two dot patches with an answer of either "correlated" or "anti-correlated." Two sets of preliminary trials were run with all subjects. In the first trial, the lengths of the motion pulses were varied while the number of pulses was held constant. In the second trial, the number of pulses presented was varied while pulse length was held constant at 100 milliseconds. This was done to determine which paradigm produced a greater variation in responses. It was found that responses varied more as the length of the pulses varied, rather than as the number of pulses presented in total varied. From this, a formal experiment was developed.

Four subjects were recruited to participate in this experiment (EL, AG, SA, and LM), three of whom were naïve to the purposes of the study, and one was the author. All subjects had normal or corrected-to-normal vision. In the experiment, subjects were asked to view a block of 200 trials per sitting; each trial lasted for the duration of ten pulses. Pulse lengths were varyingly set at five, ten, or twenty milliseconds for the entire trial block of 200. Throughout the trial block, the level of correlation of the motion between the two dot patches varied randomly from -1 (anti-correlated) to 1 (perfectly correlated) on a 0.2 scale. Each correlation level was presented an equal amount of times within the block. Each subject performed 5 blocks per pulse length, totalling 1000 trials per pulse length and 3000 trials in all. Their responses to each trial of "correlated" and "anti-correlated" were collected at the end of each trial block. Data was stored in structure variables with multiple fields by the Matlab™ program. The results from each subject were separated into three groups according to the trial pulse lengths, creating three sets of 1000 trials for each subject. An average proportion of "correlated" responses to all the levels from -1 to 1 were obtained by counting the number of "correlated" responses for each level and dividing by the total. A psychometric curve, which is a plot that is generated to fit the data of correct subject responses as a function of the changing properties of the stimulus, was then generated for each of the three groups. The Nelder-Mead nonlinear optimization method was used to minimize the maximum likelihood function for a logistic fit for the data points. This method is included in Matlab™ software packages where it is implicated in the "fminsearch" function. Three psychometric curves were generated using the average proportion values for each subject.

A model was developed to analyze what the optimal human response should be to trials of different levels of correlation for a stimulus of ten pulses. In this situation, "optimal" refers to a choice made based solely on the majority of pulses being the same or different. This model was developed by generating long motion pattern sequences at all levels of correlation. Then, the probability of two pulses in these sequences being the same (correlated) or opposite (anti-correlated) was determined. Using the generated probabilities given for each level of correlation, the Binomial Cumulative Distribution Function (Eq. 1) was used to determine the likelihood of an observer responding that a stimulus was "correlated" for each level of correlation from -1 to 1 on a 0.2 scale. A value of ten was inserted into the equation for *n*, corresponding to ten pulses, and a value of five was inserted for x, corresponding to the number of correlated pulses necessary for an observer to consider the majority of the pulses as being "correlated." These likelihoods were also generated for stimuli consisting of other numbers of pulses (Fig. 3).

$$F(x,p,n) = \sum_{i=0}^{x} \binom{n}{i} (p)^i (1-p)^{(n-i)}$$

Eq. 1

RESULTS

No significant amount of variation was found in the preliminary experiments in which subjects viewed stimuli that varied by the numbers of pulses. Therefore, data was only considered from the experiment in which pulse lengths alone varied and the number of pulses was kept constant at ten pulses per trial.

For each of the three psychometric curves generated per subject, the point of most interest on each curve was at 0 correlation, because when a stimulus has 0 correlation, the subject is essentially guessing as to whether it is correlated or anti-correlated. At this point, each psychometric curve had a distinct slope, which could be used to compare the subject's performance to the optimal model's performance. Since pulse length did not factor into the optimal model's performance, and the number of pulses was the only significant characteristic, subjects' performances for different pulse lengths were all compared against the optimal model's performance at ten pulses.

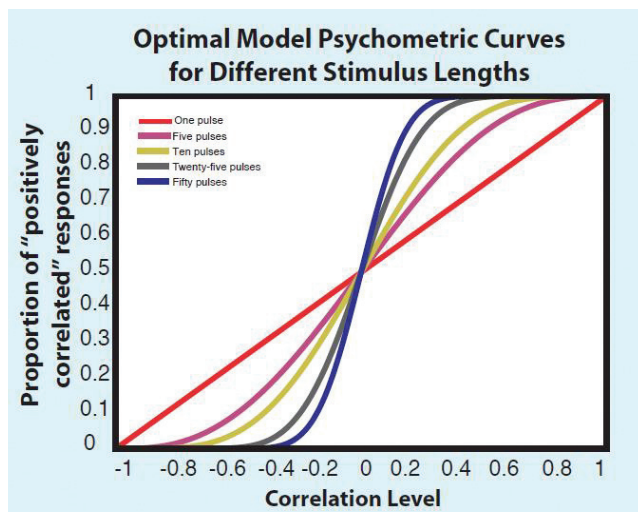

Fig. 3. Optimal model psychometric curves generated for different numbers of pulses. As the number of pulses increased, so did the slope of the curve, as well as the model's accuracy for detecting whether a stimulus was "correlated" or "anti-correlated."

As pulse length increased, subjects' performances became more similar to the optimal model's (Fig. 4). At pulse lengths of 50 ms, the average slope for all subjects at zero correlation was a low

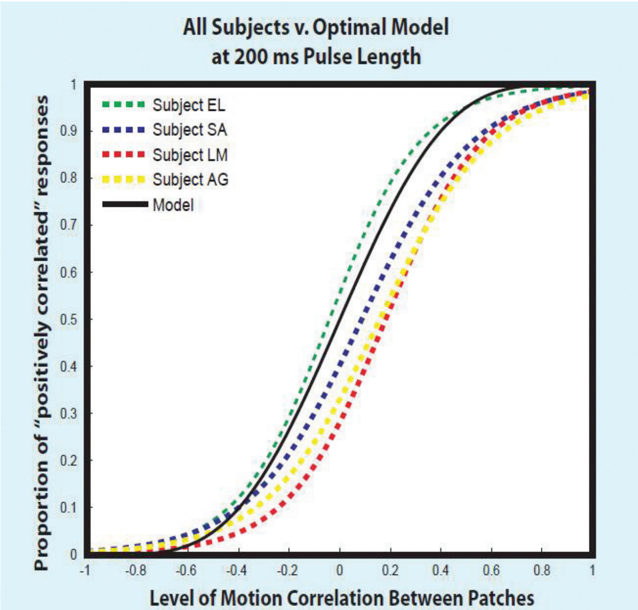

Fig. 4. At the longest experimental pulse length, 200 ms, all subjects' performances became more similar to the optimal model's.

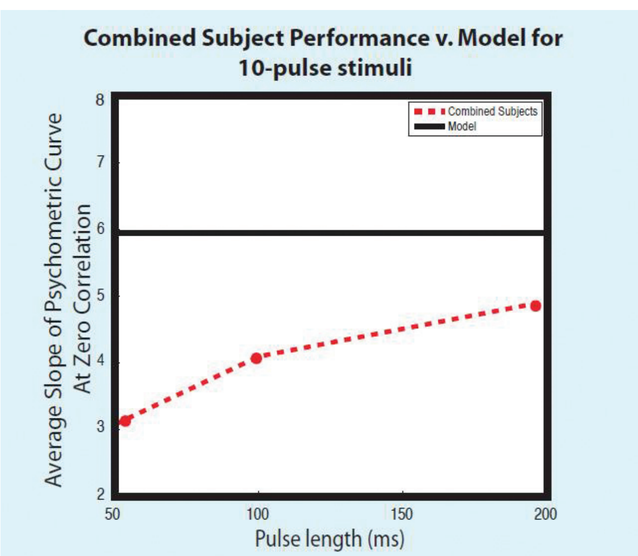

Fig. 5. The slopes of each subjects' psychometric curves were averaged for each experimental pulse length. At 50 ms, there is a great difference between subject slope and the model's slope. As the pulse length increased, the difference between subject slope and model slope became much smaller. Note that the model's slope never changes, as its decision is based solely on majority of correlated pulses, and pulse length is not a deciding factor.

3.09. As the pulse lengths increased, however, so did the slopes of the subjects' psychometric curves. At a length of 200 ms, subjects' average slope was 4.94, in comparison to the optimal model's slope, which at zero correlation for ten pulses was 5.99. The average slopes of all subjects at zero were plotted against the optimal model, with one plot for each of the three pulse lengths (Fig. 5).

DISCUSSION

In this experiment, human performance was determined according to the slope of the psychometric curve during each experimental condition. The slopes were viewed at the correlation level of zero, so as to precisely analyze the differences in response between correlated and anti-correlated stimuli. A higher slope would correspond to a more accurate subject performance, since dependent points on the psychometric curve correspond to proportions of "correlated" responses. At all negative correlation levels, an ideal observer would *never* respond with a "correlated" answer and at all positive correlation levels the response would *always* be "correlated." Human subjects are not ideal observers, however, and statistically the closest that we can come to this ideal observer is shown by the optimal model. With a greater number of pulses, the slope of the optimal model's curve becomes steeper (Fig. 3). When calculating the response probabilities with the Binomial Cumulative Distribution Function, the model's probabilistic response became more accurate with more pulses, leading to steeper curves.

Human performance was expected to approach the level of the optimal model as the number or length of the pulses increased. Preliminary data ruled out the possibility of human performance enhancement through increasing the number of pulses. This result could be due to the brain's limited integration window in perceiving visual stimuli (*5*). The brain is only capable of integrating with a certain level of accuracy at a given speed and pulse length. Therefore, if the motion patterns in the stimulus move at a set speed and a set pulse length, the number of pulses presented may not affect the subject's performance. Based on this result, a stimulus of ten pulses or one hundred pulses could then be presented at that same pulse length and pulse speed without any significant performance change. Although the pilot trials with increased number of pulses increased the total length of the trials, we can rule out the possibility of increased trial length affecting subject response because of the limiting factor of the brain's integration window.

In contrast, experimental data showed significant performance enhancement with increased pulse lengths and a constant pulse number, which does correspond to the second possible method of enhancing human performance stated in the original hypothesis. The comparisons between the subjects' psychometric curves and the optimal model's ten-pulse psychometric curve can be seen in Fig. 4. The slopes of all four subjects' curves ranged from 4.5 to 5.5 at 0 correlation with a 200 ms stimulus. According to the model, optimal performance would exhibit a slope of 5.99 with a ten-pulse stimulus. The increasing slope trend is shown in Fig. 5, with the average across-subject slope approaching that of the optimal model, with all stimuli consisting of ten pulses.

The hypothesis also asked whether a threshold exists at which point human performance can no longer improve and if that threshold would equal the model's performance. Trials in this study tested pulse lengths of 50, 100, and 200 ms, and an increasing performance trend was seen throughout. Human performance approached that of the model's (Fig. 5), although the difference between the average slope and the model at the longest pulse length was still greater than 1, which was found to be a significant difference. Although the data presented here satisfies the hypothesis, further experimental data should be collected to determine if even longer pulse lengths can create greater improvement, and if a threshold pulse length exists where the difference between human performance and model performance is not significant. It is clear from this data, however, that human performance is improved with longer pulse length rather than with a greater number of pulses.

CONCLUSION

The results of this experiment appear to support and expand the *binding by synchrony* theory of the Binding Problem. *Binding by synchrony* explains how neurons that correspond to a certain property of an image fire in synchrony, but it does not investigate *why* the groups of neurons fire together or whether this synchronous firing can be made stronger or weaker.

Our data concerns the property of *temporal structure* in BBS. By manipulating temporal structure by varying stimulus pulse lengths, we have elucidated a specific mechanism within the Binding Problem. Here, we were able to study exactly how a certain property of an image can be altered to further enhance or decrease visual integration performance. The influence of temporal structure is more qualitative than quantitative when used as a binding element in visual processing. This was shown by demonstrating that the quality (length of pulses) was more important than quantity (the number of pulses actually presented) when determining whether motion pulses were correlated or anti-correlated.

This knowledge of how visual perception is enhanced qualitatively will help add to our understanding of how the human brain breaks down and integrates information from a dynamic environment to create complete, coherent images. Now that specific parameters have been established under which human performances become more similar to the optimal model, future research will be required to explore how human performance can change in comparison to an optimal model through manipulation of different stimuli, such as background distracters, contrast changes, shifts in which the two patches are offset by one or two pulses (jitter experiments), and spatial changes of the target patches.

REFERENCES

1. A. K. Engel, P. Konig, A. K. Kreiter, T. B. Schillen, W. Singer, *Trends Neurosci* **15**, 218 (Jun, 1992).
2. S. Cheadle *et al., J Vis* **8**, 12 1 (2008).
3. S. E. Guttman, L. A. Gilroy, R. Blake, *Vision Research* **47**, 219 (2007).
4. A. Bartels, S. Zeki, *Vision Res* **46**, 2280 (Jul, 2006).
5. W. R. Uttal, L. Spillmann, F. Stürzel, A. B. Sekuler, *Vision Research* **40**, 301 (2000).
6. tS. H. Lee, R. Blake, *Science* **284**, 1165 (May 14, 1999).