A neural network theory of binocular combination

Jian Ding and George Sperling

Human Information Processing Laboratory. Department of
Cognitive Sciences and Institute for Mathematical Behavioral
Sciences. University of California, Irvine. CA 92717. U.S.A.

When different stimuli are presented to the left and right eyes, only a single, combined "cyclopean" image is perceived. The numerous theories that have been proposed to describe binocular combination rule typically abstract a parameter from each eye's image (e.g., maximum stimulus contrast or the visual direction of a significant point) and use the two eyes' values to predict the parameter value in the cyclopean image. Here we offer a process theory that, in principle, applies to any image within a spatial frequency band, and predicts, pixel by pixel, the combined cyclopean image.

We propose simply that, in every neighborhood, each eye (1) exerts gain control on the other eye in proportion to the strength of its own input and (2) exerts gain control on the other eye's gain control. We tested the computational predictions theory with superthreshold sinewave gratings, delivered independently to each eye. The gratings were identical except of a difference in phase and contrast. The reason for choosing sine waves is that the sum of two sine waves of the same wavelength is again a sine wave. The relative contribution of each eye to the cyclopean sine wave can be determined from the apparent location of the cyclopean sine between the two ocular inputs. Sine waves were horizontal to make the cyclopean image independent of horizontal vergence angle, and there was a strong, surrounding visual frame to assist the eyes in maintaining vertical vergence. Subjects judged the position of the dark stripe of the cyclopean sine wave relative to adjacent bar markers.

We performed two experiments. In the first experiment, no external noise was added and the sine waves in the left and right eyes were of different contrasts and in different phases. The arithmetic addition of two parallel sine waves produces a sine wave of known phase and amplitude. The data consisted of 192 combinations of contrasts of the left- and right-eye stimuli, and of their relative phases. For stimuli to one eye only, and for equal stimuli to both eyes, the position judgments agreed perfectly with the linear summation model. For all other contrasts in the two eyes, the linear summation model failed. From the judged position, we inferred the strength of the contribution of each eye to the cyclopean image. In general, the stimulus with greater contrast had more weight than predicted by simple linear summation. Without estimating any parameters from the data, the gain-control model accounted for 98 percent of the variance of the data. We conclude that a simple, robust, physiologically plausible model accurately describes binocular combination of parallel sine waves.

In the second experiment, there were sine waves in the left and right eyes; they were always of the same contrast but differed in phase. Bandlimited visual noise was added to the stimulus to one eye only. If there were no added noise, and because the left and right eye stimuli have the same contrast, both linear summation model and gain-control model predict that the cyclopean sine would lie exactly between the sines to each eye. When external noise is added to one (or both) eye, the linear summation model predicts that, except for greater variability, the mean position of the cyclopean image will remain unchanged. If we consider that the added noise reduces the reliability of the input, a Baysian model might predict a lesser weight for the signal from the eye with noise. The gain-control model makes the opposite prediction. The eye with added noise has more total contrast energy; therefore its signals exert more inhibition on the other eye than vice versa; and therefore inputs from the eye with noise should gain greater weight in the cyclopean image.

In the second experiment, a two-dimensional bandpass noise (in one of six frequency bands from 0.34 to 10.9 cpd) was added to the sine grating stimulus. Altogether, 96 combinations of noise contrast amplitude and spatial frequency were tested. The results confirmed the prediction of gain-control model, namely, that the contribution to the cyclopean image would be greater for the eye that had the added noise. Indeed, when the added noise was strong enough, it effectively blinded the other eye. The subject perceived only the image from the noise-containing eye.

In the gain-control model, the total contrast energy for inhibition is summed over the spatial frequency domain. For the sine channel centered at 0.68 cpd (the sine grating), the weight of each spatial frequency channel in the total contrast combination was inferred from the experimental data. With these additional channel-effectiveness parameters, the gain control model again gives an extremely accurate account of the data. We conclude that the total, channel-weighed contrast energy, (not merely the contrast of the single most relevant channel), determines gain-control in binocular combination. So far, only a small subset of relevant stimulus conditions has been tested. Although the results are promising, much remains to be done to further test and develop the model.

Conclusion. A physiologically motivated, neural network model is proposed to account for the contributions to cyclopean perception of the retinal images in each eye. For superthreshold stimuli, when either eye is stimulated alone, the model (correctly) predicts that the cyclopean image would be the same as when both eyes receive the same stimulus. When sine gratings presented to the two eyes differ in contrast or phase, or when noise is added to the stimulus in one eye, the model accurately describes the weights accorded to the left- and right-eyes' contributions to the cyclopean image.

Supported by Air Force Office of Scientific Research, Human Information Processing Program.