OPTICAL REVIEW Vol. 1, No. 2 (1994) 311-313

Spatial Frequency Channels for Stereoscopic Depth Perception

Satoshi SHIOIRI, Tomomichi HATORI, Hirohisa YAGUCHI and Souichi KUBO

Department of Image Science, Chiba University, 1-33, Yayoi, Inage-ku, Chiba, 267 Japan

(Accepted August 26, 1994)

We investigated spatial frequency tuning of the mechanisms for stereoscopic depth perception, using a masking technique, and compared these tunings with those for luminance pattern detection. Observers discriminated depth direction in random-dot stereograms with various contrasts of a masking pattern, which was dichoptic stimulus of uncorrelated random dots. The strength of masking effect as a function of spatial frequency (i.e., masking function) was measured: masking function approximates the spatial frequency tuning of the mechanism that detects the depth in the test pattern. The masking functions for stereoscopic depth were found to be similar to either of the three of the six spatial frequency channels proposed by Wilson and Gelb (1984) for the detection of luminance patterns.

Key words: spatial frequency channel, spatial frequency tuning, masking, stereoscopic depth, binocular disparity

1. Introduction

Spatial frequency channels in the visual system have been widely investigated for the detection of luminance patterns. There are a number of reports that show that the visual system has several channels narrowly tuned in spatial frequency domain.[1,2] These studies suggest that the visual system has a limited number of spatial frequency channels to process luminance distribution on the retinas. Indeed, Wilson and Gelb proposed six spatial frequency channels for the human visual system based on psychophysical experiment.[1]

However, much less is known for the spatial frequency channels for other mechanisms than the luminance mechanism, while it is often assumed that there are several independent processes in the early stage of visual processing that are specialized for different types of image processing such as luminance, color, stereopsis, motion and so on[3] (although they may not be completely independent[4]). There are a few studies of spatial frequency channels for the mechanisms of stereoscopic depth[5] and for motion.[6] However, it is not clear whether or not the spatial frequency channels for stereoscopic depth or for motion are the same as those for luminance. The purpose of the present study is to examine whether the luminance process and the depth process have common spatial frequency channels. If the luminance process and the depth process have common channels, it will be suggested that these channels exist at the stage before the visual information separates for luminance and depth processing. We investigated spatial frequency tuning of the mechanisms for stereoscopic depth detection and compared them with those for luminance pattern detection in the literature.

2. Experiment

2.1 Method

We used a masking technique to measure the spatial frequency tuning. In a masking technique, a mask stimulus with variable contrast was added to a test stimulus with a constant contrast: we used twice of the threshold in the present experiment. By measuring the contrast of the mask stimulus with which the stereoscopic depth of the test stimulus is just detectable for several spatial frequencies of mask, the sensitivity of the mechanism that detect test stimulus can be obtained as a function of spatial frequency. In the present experiment test stimulus was stereogram to measure the channels for stereoscopic depth perception.

Observers discriminated depth direction in random-dot stereograms with various contrasts of masking patterns (Fig. 1) in the method of constant stimuli. Both the test and the mask random-dot patterns were bandpass filtered to limit the spatial frequency content (0.4 octave half-height width). The central spatial frequency of the filtered pattern was chosen between 1.3 and 8.5 cpd for the test pattern (1.3, 1.7, 2.2, 2.9, 3.7, 4.9, 6.5, 8.5 cpd). The test pattern was made up of two regions of opposite depth positioned vertically one above the other in a square random-dot field. One of the halves had crossed disparity (appeared to be closer to the observer) and the other half had uncrossed disparity (appeared to be further). Masking patterns were dichoptic stimulus of uncorrelated random dots to deteriorate the depth perception in the test pattern. Ten pairs of test patterns and 10 pairs of mask patterns were prepared and used in random combinations. While contrast of the patterns was calculated as Michelson's contrast ($(L_{\max}-L_{\min})/(L_{\max}+L_{\min})$, where $L_{\max}$ and $L_{\min}$ are maximum and minimum luminance values of a pattern), we used the maximum and the minimum luminance values averaged over the 10 pairs of random-dot patterns. The average luminance of the stimulus was 45.7 cd/m^2. The binocular disparity between the two regions was $4'$ ($2'$ for uncrossed and $2'$ for crossed) and the size of random-dot fields was $4.3^\circ \times 4.3^\circ$ (128 pixels × 128 pixels). The presentation of the stimulus (i.e., test plus mask) was initialized by the observer's keystroke. A uniform field of 45.7 cd/m^2 with a square frame was displayed before and after the presentation. The square frame was to help the fusion of the left and right images prior to the stimulus presentation. The stimulus presentation duration was 1.5 s and the observer responded the region that appeared to be closer by pressing one of two keys, each of which was assigned

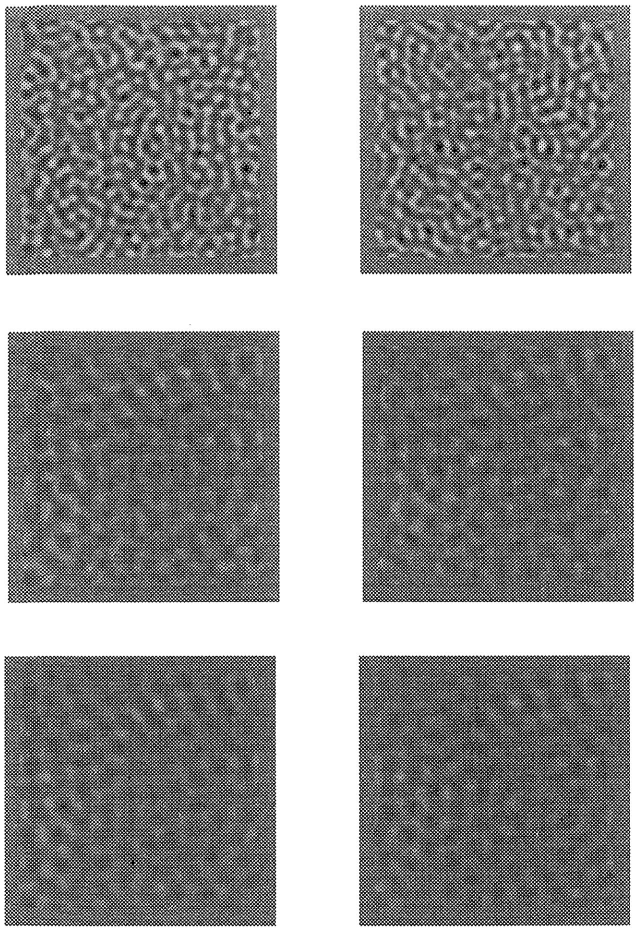

Fig. 1. Examples of stimulus patterns. A test stereogram with three levels of mask contrast; mask contrast is 0 in the top pattern, half of the test contrast in the middle pattern, and the same of the test contrast in the bottom pattern.

to either region. The percentages of correct responses were measured as a function of mask contrast from 40 presentations for each mask contrast. Two observers with normal or corrected-to-normal acuity participated in the experiment.

2.2 *Data Analysis*

We obtained the mask contrast for just detectable depth of the test stimulus as a function of mask spatial frequency to estimate the sensitivity tuning function of the mechanisms that detect the depth of the test stimulus. This function is called as masking function and assumed to be the sensitivity tuning function. The mask contrast for just detectable depth was determined as the contrast with which the observer responded 75% of correct responses from the psychometric function (i.e., percentage of correct responses for depth detection as a function of mask contrast) by Probit analysis. The mask contrast for 75% correct response was plotted as a function of spatial frequency of the mask (masking function) for each test spatial frequency.

3. Results and Discussion

Similar masking functions are obtained for different test spatial frequencies. They can be grouped to two or three tuning functions (half-height bandwidth is about 2 octaves). Masking functions that are similar in shape are shown in a same panel in Fig. 2 for each observer separately. The peak of the function is around 2.0 cpd for test frequencies of 1.7 and 2.9 cpd, around 3.0 cpd for test frequency of 3.7 cpd, and around 4 cpd for test frequencies of 6.5 and 8.5 cpd for SSU. The peak is around 1.5 cpd for test frequencies between 1.3 and 2.9 cpd, and around 3.0 cpd for test frequencies between 3.7 and 8.5 cpd for TH.

As shown in Fig. 2, the similar masking function is obtained for different test spatial frequencies. This indicates that the visual system has a limited number of mechanisms with a narrow spatial frequency tuning (spatial frequency channels) for the stereoscopic depth process. The number of the channels found in the present experiment is three (although two of them seem to be enough to explain the results of observer TH). The tuning curve of the lowest channel peaks at 1.5 or 2.0 cpd, that of the second lowest channel peaks at around 3.0 cpd, and that of the highest channel peaks at around 4.0 cpd. Other channels may or may not exist and might be found if wider range of test frequencies is used. However, the results suggest that these three are the primary channels for depth process, at least for the condition of the present experiment (temporal frequency, disparity, stimulus size and so on may change the results). The results shows that the stereoscopic depth process does not seem to have signal from spatial frequency channels that tuned lower than 1.5 cpd or higher than 4.0 cpd. This is because the results suggest that the lowest test frequency used (1.3 cpd) was detected by the channel with the peak of 1.5 cpd and the highest test spatial frequency used (8.5 cpd) is detected by the channel with the peak of 4.0 cpd. That is, even if the channels with a peak of lower than 1.5 cpd or higher than 8.5 cpd contribute the stereoscopic depth, the amount of the contribution should be smaller than that of the three channels found in our experimental condition. This contrast to the results for luminance pattern detection, which shows that the peaks of the masking function is always around test spatial frequency.[2]

The questions here is whether the channels for stereoscopic depth are the same as for luminance process (i.e., those measured for the detection of luminance pattern). Since luminance and stereoscopic depth are often assumed to be processed independently,[3] spatial frequency channels may also be independent for the two processes. On the other hand, the two processes may have common channels because there are cortical cells that differ in spatial frequency tuning in the first stage of cortical visual processing, after which the division of processing for luminance and stereoscopic depth seems to occur.

The comparison of spatial frequency channels for luminance and those for stereoscopic depth suggests a) that the number of the channels are smaller for the depth process (although the change of disparity temporal condition or stimulus sizes may change the number), and b) that the two processes have common spatial frequency channels. The number of spatial frequency channels of three for the depth process is smaller than that for the luminance process. Wilson and Gelb proposed six spatial

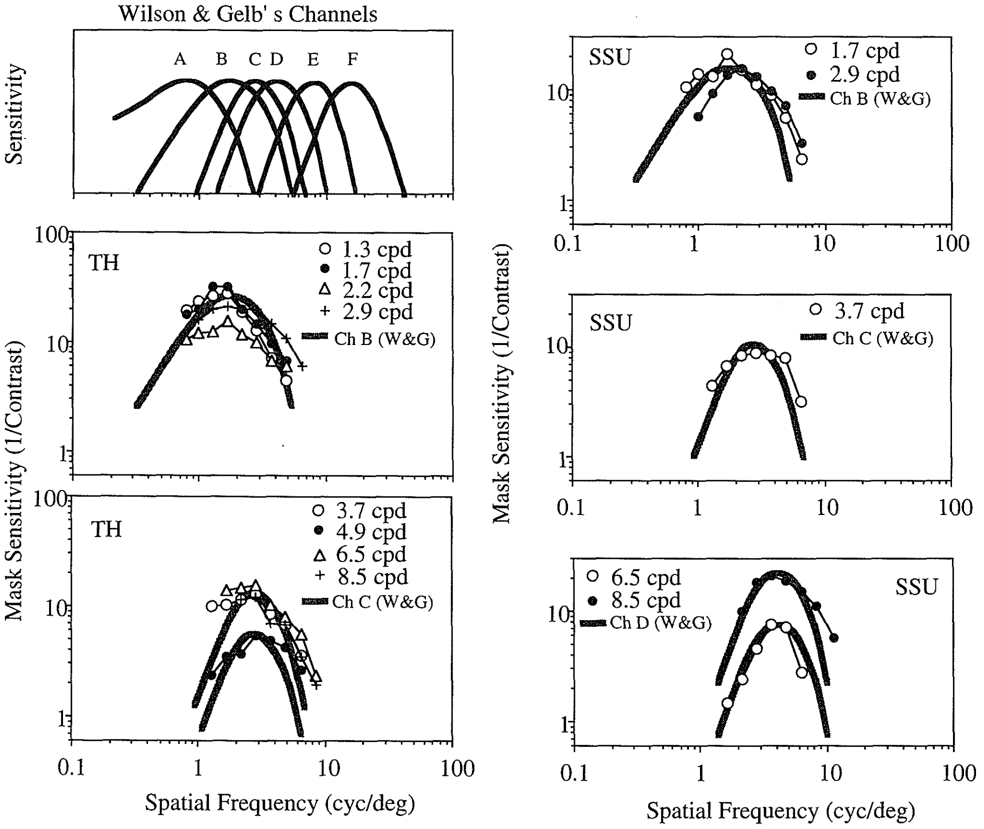

Fig. 2. Masking functions for stereoscopic depth and spatial frequency channels proposed by Wilson and Gelb. The gray line shows the spatial frequency channel that peaks at similar spatial frequency of the peak of the masking functions in each panel.

frequency channels for luminance (Fig. 2)[1] and there is a report that indicates another channels in addition to the six.[7] This suggests that the channels for luminance are not identical to those for stereopsis. However, each group of the masking functions obtained here is similar to one of the channels for luminance. The gray line in each panel in Fig. 2 shows Wilson and Gelb's channel for luminance that peaks at similar spatial frequency of the peak of the masking functions. The masking functions of test spatial frequencies between 1.3 cpd and 2.9 cpd for both observers are similar to that of the second lowest channel (channel B) of Wilson and Gelb, those of test spatial frequencies between 3.7 cpd and 8.5 cpd for TH and that of test spatial frequency of 3.7 cpd for SSU are similar to that of the third lowest channel (channel C) of Wilson and Gelb, and those of test spatial frequencies of 6.5 cpd and 8.5 cpd for SSU are similar to that of the fourth lowest channel (channel D) of Wilson and Gelb. The agreement between the masking functions of the present experiment and Wilson and Gelb's channels is fairly well for all cases. This suggests that the same spatial frequency channels underlie the luminance process and the stereoscopic depth process. This also suggests that three of the spatial frequency channels are mainly used for stereoscopic depth processing at least in the present experimental condition, although all of them are used for luminance pattern processing.

Yang and Blake also concluded based on masking experiments that there is less number of spatial frequency channels for the stereoscopic depth process than for luminance.[5] They found two spatial frequency channels with peaks of 3 cpd and 5 cpd. These peaks are close to 3 cpd and 4 cpd of the masking functions obtained from the present experiment. The reason why Yang and Blake's results did not show spatial frequency channel with the peak of 1.5 cpd or 2.0 cpd may have been because of the stimulus size. The scale of the disparate region in their experiment was 2.2° × 1.1° in the 4.4° × 4.4° background. This is one fourth of the disparate region of our stimulus (each half was 4.3° × 2.2°). Detection of larger feature could require the processing of lower spatial frequency channels.

In summary, it is suggested that the visual system has at least three spatial frequency channels for the stereoscopic depth process and that they are the three of spatial frequency channels for the luminance process. This supports the idea that spatial frequency channels exist prior to the division of processing for luminance and stereoscopic depth.

References

1) e.g. H.R. Wilson and D.J. Gelb: J. Opt. Soc. Am. A 1 (1984) 124.
2) H.R. Wilson, D.K. McFarland and G.C. Phillips: Vision Res. **23** (1983) 873.
3) P. Cavanagh: Comput. Vision, Graphics Image Process. **37** (1987) 171.
4) E.A. DeYoe and D.C. Van Essen: Trends Neurosci. **11** (1988) 219.
5) Y. Yang and R. Blake: Vision Res. **31** (1991) 1177.
6) S.J. Anderson and D.C. Burr: Vision Res. **25** (1985) 1147.
7) S.R. Lehky: J. Opt. Soc. Am. A 2 (1985) 1260.