Vision Research 146–147 (2018) 18–31

Contents lists available at ScienceDirect

Vision Research

journal homepage: www.elsevier.com/locate/visres

Binocular contrast-gain control for natural scenes: Image structure and phase alignment

Pi-Chun Huang*, Yu-Ming Dai

Department of Psychology, National Cheng Kung University, Tainan, Taiwan

ARTICLE INFO

Number of Reviews = 2

Keywords:
Natural image
Binocular contrast-gain control
Dichoptic
Phase-scrambled

ABSTRACT

In the context of natural scenes, we applied the pattern-masking paradigm to investigate how image structure and phase alignment affect contrast-gain control in binocular vision. We measured the discrimination thresholds of bandpass-filtered natural-scene images (targets) under various types of pedestals. Our first experiment had four pedestal types: bandpass-filtered pedestals, unfiltered pedestals, notch-filtered pedestals (which enabled removal of the spatial frequency), and misaligned pedestals (which involved rotation of unfiltered pedestals). Our second experiment featured six types of pedestals: bandpass-filtered, unfiltered, and notch-filtered pedestals, and the corresponding phase-scrambled pedestals. The thresholds were compared for monocular, binocular, and dichoptic viewing configurations. The bandpass-filtered pedestal and unfiltered pedestals showed classic dipper shapes; the dipper shapes of the notch-filtered, misaligned, and phase-scrambled pedestals were weak. We adopted a two-stage binocular contrast-gain control model to describe our results. We deduced that the phase-alignment information influenced the contrast-gain control mechanism before the binocular summation stage and that the phase-alignment information and structural misalignment information caused relatively strong divisive inhibition in the monocular and interocular suppression stages. When the pedestals were phase-scrambled, the elimination of the interocular suppression processing was the most convincing explanation of the results. Thus, our results indicated that both phase-alignment information and similar image structures cause strong interocular suppression.

1. Introduction

Both physiological and psychophysical studies have shown that the primary visual cortex decomposes images from the retina into various channels, such as those tuned to spatial-frequency, orientation, and phase information (Braddick, Campbell, & Atkinson, 1978; Campbell & Maffei, 1974; De Valois, De Valois, & Yund, 1979; Hubel & Wiesel, 1959, 1968; Sekuler, 1974). Thus, researchers have commonly used sine-wave grating patterns or Gabor patterns to study visual perception in the laboratory and to predict visual processing for natural-scene images. Natural-scene stimuli are close to what people see in daily life; these stimuli contain broad distributions of spatial frequency, orientations, luminance, and contrast (Field & Chandler, 2012). However, researchers have also shown that, during the processing of these channels, interactions occur within or between locations, at both near-detection and suprathreshold levels (Bex, Mareschal, & Dakin, 2007; Bex, Solomon, & Dakin, 2009; Bonds, 1989; Polat & Sagi, 1993; Cannon & Fullenkamp, 1991; Snowden & Hammett, 1998). No simple relationship exists between gratings and real scenes in terms of either perception or neural response (Bex & Makous, 2002; Bex et al., 2009; David, Vinje, & Gallant, 2004; Gallant, Connor, & Van Essen, 1998). Such findings have led to debates regarding how the visual system processes these natural scenes (Felsen & Dan, 2005; Olshausen & Field, 2005; Rust, Schwartz, Movshon, & Simoncelli, 2005) and regarding whether grating results can be used to predict natural-scene results.

The visual system detects changes in luminance and interprets them as contrast; thus, researchers must understand how the visual system responds to various levels of contrast; a contrast response function (CRF) can describe this relationship. In behavioral studies, researchers have often used the pattern-masking paradigm to derive the CRF because of the limitation that a CRF cannot be measured directly, as in neurophysiological studies. Such a pattern-masking paradigm can be used to obtain a target-threshold versus pedestal-contrast (or TvC) function by measuring a target's detectability in the presence of pedestals with various contrast levels. A TvC function can be used to derive its underlying CRF. That is, a TvC function is approximately proportional to the inverse of the derivative of the CRF. When the target and pedestal are of the same type (e.g., Gabor) aside from the contrast, a

* Corresponding author.
E-mail address: pichun_huang@mail.ncku.edu.tw (P.-C. Huang).

https://doi.org/10.1016/j.visres.2018.02.012
Received 25 August 2017; Received in revised form 12 February 2018; Accepted 12 February 2018
0042-6989/

TvC function typically exhibits a dipper shape in which the pedestal facilitates the target's detectability at low pedestal-contrast levels, but the pedestal suppresses target detection at high pedestal-contrast levels. When the target and pedestal are not of the same type, the facilitation disappears, and the weaker suppression effects are shown at high pedestal-contrast levels. Typical explanations categorize the TvC function as being a consequence of stimulus uncertainty (Pelli, 1985), of contrast-gain control (Foley, 1994; Foley & Chen, 1999), or of a contrast transducer (Foley & Legge, 1981; Legge & Foley, 1980).

Bex et al. (2007) adopted the pattern-masking paradigm and employed natural scenes as stimuli to investigate the visual system's contrast responses. They found that the TvC function for bandpass-filtered natural images (for both target and pedestal) had typical dipper functions consistent with the results for grating stimuli (Chen & Tyler, 2001; Foley, 1994). When the pedestals are unfiltered natural images, the TvC functions show stronger masking effects at high pedestal-contrast levels if the structures of the pedestal and target are aligned. However, when the structure of the unfiltered pedestals is not aligned with the target—whether by position shifting, image mirroring, or image rotating—the suppression effect at a high level of pedestal-contrast is reduced, and the strength of the masking is similar to that of the narrow-band-filtered pedestal. These results suggest that the level of masking depends on the correlation between the edge features of the target and pedestal. Within a limited region of space, the phase structure at remote spatial-frequency scales controls the contrast response to natural scenes.

The visual system receives information from both eyes and integrates that information to form a coherent precept. Scholars of psychophysics have examined binocular contrast interactions extensively using sine-wave gratings within a variety of detection, discrimination, and masking paradigms. However, relatively little is known about binocular contrast interactions for natural-image stimuli. A square-wave grating (with sharp edges) has also been shown to cause stronger dichoptic masking than a phase-scrambled square-wave grating, indicating that the edges formed by phase alignment and the structure of the images influences the strength of the interocular suppression (Huang, Maehara, May, & Hess, 2012; Maehara, Huang, & Hess, 2009). Given that the level of masking is a function of the correlation between the target and pedestal patterns' edge features under binocular viewing conditions (Bex et al., 2007), natural images are the appropriate stimuli for investigating not only the CRF of a broadband stimulus but also the effects that phase alignment and image structure have on interocular suppression and contrast-gain control in visual processing. The phenomenon of dichoptic blur suppression is the suppression of a blurred stimulus in one eye by a well-defined sharp stimulus in the other eye; investigating binocular contrast-gain control using a natural scene could improve the understanding of the underlying dichoptic blur suppression mechanism. One possible explanation for this involves low-level pattern masking in which the sharp image produced by phase-aligned information suppresses the other eye's information through interocular suppression. By using natural images as stimuli, we can further investigate how image-structure and phase-alignment information affect blur suppression.

The pattern-masking paradigm has been widely applied to monocular conditions, dichoptic conditions, and binocular conditions to investigate various properties of monocular pathways and interactions between the two eyes (Legge, 1984a, 1984b; Maehara & Goryo, 2005; Meese, Georgeson, & Baker, 2006). The results for grating stimuli have shown them to have the following characteristics: first, classic dipper shapes can be produced by monocular and by binocular TvC functions (Legge & Foley, 1980; Wilson, 1980). Facilitation has been shown to occur at low pedestal-contrast levels; at high pedestal-contrast levels, masking has been shown to occur (Legge & Foley, 1980; Nachmias & Sansbury, 1974; Stromeyer & Klein, 1974). The monocular and binocular functions have revealed binocular summation at very low pedestal-contrast levels for which the detection thresholds in the binocular viewing condition are √2 times lower than those in the monocular viewing condition (Campbell & Green, 1965; Legge, 1979, 1984a, 1984b; Meese & Hess, 2004, 2005). The binocular advantage disappears at relatively high stimulus contrast levels. Second, in the dichoptic viewing condition, the target and pedestal are presented to different eyes, and the pedestal elevates target thresholds more effectively than in either the monocular or binocular conditions. The dichoptic pedestal increases the discrimination thresholds almost linearly as the pedestal-contrast levels are increased. Recently, scholars have proposed several models of binocular interaction by adapting contrast-gain control (Ding & Sperling, 2006; Maehara & Goryo, 2005; Meese & Hess, 2004; Meese et al., 2006). In general, these models include a nonlinear contrast transducer for each monocular input that also receives suppression from the other eye. Then, the nonlinear contrast transducers from both eyes are combined. Meese et al. published a two-stage binocular contrast-gain control model (Meese et al., 2006); over the course of several publications, that model has successfully explained the results of various tasks: contrast matching, contrast detection, and contrast discrimination (Baker, Meese, & Georgeson, 2007; Baker, Meese, Mansouri, & Hess, 2007; Baker, Meese, & Summers, 2007; Meese et al., 2006). For this study, we modified this robust model; the details are in the Model section.

In this study, we adopted the pattern-masking paradigm using natural scenes as stimuli and compared the TvC functions under monocular, binocular, and dichoptic viewing conditions. We addressed four main topics. First, we explored whether the contrast discrimination ability for the natural scenes is similar to that of gratings. Whereas the aforementioned two-stage model of binocular contrast-gain control had originally been used for grating-discrimination thresholds, we used it to investigate whether the results for natural scenes can be quantitatively described. Second, we investigated how the structure and phase-alignment information from the natural scenes can influence the contrast-gain control by comparing the results for the phase-aligned pedestal to the results for the rotated pedestal (Experiment 1) and to the results for the phase-scrambled pedestals (Experiment 2). If the image's structural information and phase-alignment information are critical to the target-detection mechanism, we would expect that changes in the image structure and phase alignment would exhibit a weak masking effect. Furthermore, we explored the underlying mechanism for blur suppression. Third, we investigated how various types of spatial frequency information influence discrimination performance by considering both phase-aligned and phase-scrambled information. To manage this problem, we adopted three pedestals and three phase-scrambled versions of the pedestals: bandpass-filtered, unfiltered, and notch-filtered. The bandpass-filtered pedestal contained an image structure that was formed using the same frequency as the target; the unfiltered pedestal contained an image structure formed by the broadband spatial-frequency information, and the notch-filtered pedestal contained an image structure with only a small amount of the same spatial-frequency information as that of the target. The phase-scrambled version of the pedestals destroyed the phase-relationship and structural information of the image. Thus, by comparing the results for these six types of pedestals, we also clarified the interactions of the within- and between-channel masking as well as the effects of phase-aligned information. Furthermore, by comparing the masking effects for monocular conditions, binocular conditions, and dichoptic conditions, we investigated whether the between-channel masking occurred before or after the binocular interaction stage. If the cross-channel spatial frequency occurred only after binocular summation, we would expect the monocular and dichoptic viewing conditions to have the same discrimination performance. Finally, regarding the two-stage model of binocular contrast-gain control, we considered whether it could be modified to describe the pattern-masking results when using the aforementioned natural-scene pedestals.