



Published in final edited form as:

Ann N Y Acad Sci. 2005 April ; 1039: 252–259.

Short-Latency Disparity Vergence in Humans: Evidence for Early Spatial Filtering

B.M. SHELIGA, K.J. CHEN, E.J. FITZGIBBON, and F.A. MILES

Laboratory of Sensorimotor Research, National Eye Institute, National Institutes of Health, Bethesda, Maryland 20892, USA

Abstract

Our study was concerned with the disparity detectors underlying the initial disparity vergence responses (DVRs) that are elicited at ultrashort latencies by binocular disparities applied to large images. DVRs were elicited in humans by applying horizontal disparity to vertical square-wave gratings lacking the fundamental (termed here, the “missing fundamental”). In the frequency domain, a pure square wave is composed of odd harmonics—first, third, fifth, seventh, etc.—such that the third, fifth, seventh, etc., have amplitudes that are one-third, one-fifth, one-seventh, etc., that of the first, and the missing fundamental lacks the first harmonic. The patterns seen by the two eyes have a phase difference of one-quarter wavelength, so the disparity of the features and $4n + 1$ harmonics (where $n = \text{integer}$) has one sign (crossed or uncrossed), whereas the $4n - 1$ harmonics—including the strongest Fourier component (the third harmonic)—has the opposite sign (uncrossed or crossed): spatial aliasing. The earliest DVRs, recorded with the search-coil technique, had minimum latencies of 70 to 80 ms and were generally in the direction of the third harmonic, that is, uncrossed disparities resulted in convergent eye movements. In other experiments on the DVRs, one eye saw a missing fundamental and the other saw a pure sine wave with the contrast and wavelength of the third harmonic but differing in phase by one-quarter wavelength. This resulted in short-latency vergence in accordance with matching of the third harmonic. These data all indicate the importance of the Fourier components, consistent with early spatial filtering prior to binocular matching.

Keywords

missing fundamental; binocular disparity; vergence eye movements

DISPARITY DETECTION AND VERGENCE EYE MOVEMENTS

When large random-dot patterns are viewed dichoptically and are suddenly subjected to small binocular misalignments (disparities), corrective vergence eye movements are elicited at ultrashort latencies.¹⁻⁸ Crossed disparities elicit increased convergence and uncrossed disparities elicit decreased convergence, exactly as expected of a depth-tracking servomechanism driven by disparity. However, this mechanism has a very limited range of disparities over which it behaves like a servo-mechanism, that is, the range over which increases in the disparity vergence error result in roughly linear increases in the vergence response is <2 deg. Indeed, disparities >4 deg are without effect at short latency, so only small misalignments of the two eyes can be corrected by this ultrarapid vergence mechanism, commensurate with mediation by disparity detectors that perform only local stereo matches. The clear implication is that this disparity vergence mechanism functions as a low-level automatic (i.e., involuntary) servo and is not involved in high-level operations such as the voluntary transfer of fixation to

new images in new depth planes, a process that requires time-consuming target selections, and may involve the decoding of large disparity errors (>10 deg), with all the attendant correspondence problems.

Vergence responses can also be elicited at ultrashort latencies by disparity stimuli when the dichoptic patterns are anticorrelated, a situation in which the two images have opposite contrast so that regions that are seen as black by one eye are seen as white by the other eye and vice versa.^{3,5} An important feature of these vergence responses to anticorrelated stimuli is that they are in the reverse direction of those to normal correlated stimuli. In two-alternative forced-choice tests, subjects could readily discriminate between crossed and uncrossed disparities when applied to normal correlated patterns, but not when applied to anticorrelated patterns.^{3,9-11} Thus, anticorrelated patterns are seen as rivalrous and do not support depth perception. This indicates that the earliest (i.e., short-latency) vergence responses must derive their visual input from an early stage of cortical processing prior to the level at which depth percepts are elaborated.³

Bilateral lesions of the medial superior temporal area of the cortex (MST) result in major impairments of the earliest DVRs,⁸ and single-unit recordings indicate that neurons in this region discharge in relation to the disparity stimuli used to elicit DVRs.^{5,7} Very few of these individual neurons show a dependence on disparity resembling that for the earliest DVRs, but when the neuronal responses are simply averaged together the resulting dependence on disparity is almost identical to that of the vergence eye movements: population coding.^{5,7} This coding of DVRs by the average activity in MST was equally good for the data obtained with correlated and anticorrelated patterns, and even reproduced the idiosyncrasies of the individual animals. In sum, the initial DVRs are encoded in the population activity in MST and this region of cortex appears to play a critical role in the generation of these eye movements. Recordings from disparity-selective neurons earlier in the cortical visual pathway suggest that their monocular inputs undergo spatial filtering prior to binocular combination,¹² and imply that low-level stereo matching does not depend on the point-to-point correspondence in the raw images, but rather on the common Fourier components in the (filtered versions of the) binocular images. We sought to examine this issue using complex grating patterns in which the disparity of the principal Fourier component was in conflict with the disparity of the features in the raw image.

THE MISSING FUNDAMENTAL STIMULUS: EVIDENCE FOR EARLY SPATIAL FILTERING

We recorded the initial vergence eye movements elicited by disparities applied to the so-called missing fundamental stimulus, constructed by subtracting the fundamental sine-wave component from a square wave. In the frequency domain, a pure square wave is composed entirely of the odd harmonics—the first, third, fifth, seventh, etc.—with progressively decreasing amplitudes such that the third, fifth, seventh, etc., have amplitudes that are one-third, one-fifth, one-seventh, etc., that of the first harmonic. Accordingly, the missing fundamental stimulus lacks the first harmonic and so is composed entirely of the higher odd harmonics, with the third having the lowest spatial frequency and the largest amplitude (see FIG. 1). The missing fundamental stimulus has been used extensively in visual psychophysics and is described only in brief here (see a companion paper, Chen *et al.*,¹³ in this same volume for a more complete description). It has the important property that when displaced one-quarter wavelength, all of its harmonics are displaced by one-quarter of their wavelengths, the $4n + 1$ harmonics (where $n = \text{integer}$) being displaced in the forward direction and the $4n - 1$ harmonics in the backward direction. In the present experiments, a dichoptic viewing arrangement was used to present missing fundamental vertical gratings to the two eyes. These two gratings were identical except for a (horizontal) phase difference of one-quarter wavelength, so that the

disparity of the features and $4n + 1$ harmonics had one sign (crossed or uncrossed), whereas the $4n - 1$ harmonics, including the strongest Fourier component (the third harmonic), had the opposite sign (uncrossed or crossed): spatial aliasing. This is illustrated in FIGURE 2A, which shows a missing fundamental stimulus with an uncrossed disparity, so that the pattern seen by the right eye is a quarter-wavelength to the right of the pattern seen by the left eye. We will see that the principal Fourier component—the third harmonic—is an important determinant of the responses, and its disparity has the opposite sign (crossed): FIG. 2B. We have indicated this matching in FIGURE 2B by encircling one of the peaks in the third harmonic seen by the left eye and linking it by an arrow to its nearest neighbor in the image seen by the right eye—the peak to the immediate left—even though it could have been matched to any of the other peaks seen by that right eye. The reason for this is that the brain seems to give greatest weight to the nearest match (the “nearest neighbor” principle, a form of spatial aliasing).

Visual images were generated on two identical video monitor screens and presented dichoptically in a Wheatstone mirror stereoscope. Trials began with a blank screen except for a centered binocular fixation target. The fixation target(s) then disappeared and were immediately replaced by the missing fundamental patterns, which were identical for the two eyes except for a horizontal phase difference of onequarter wavelength, and these filled the screens for 200 ms. At this point the screens were blanked, marking the end of the trial. Initial DVRs, recorded from three human subjects with the magnetic search-coil technique,¹⁴ had minimum latencies of 70 to 80 ms and were generally in the direction of the third harmonic over a wide range of wavelengths (1 to 24 deg/cycle) and contrasts (2% to 79%). Sample mean vergence velocity profiles from one subject in response to 2-deg uncrossed disparities applied to a missing fundamental grating of wavelength 8 deg/cycle are shown in FIGURE 3 (solid trace labeled “MF,” indicating “Missing Fundamental”) and indicate that the earliest DVRs were convergent. Also shown in FIGURE 3 are the responses to identical 2-deg uncrossed disparities applied to pure sine-wave gratings whose wavelengths were either (1) the same as that of the missing fundamental (8 deg/cycle), resulting in divergent DVRs (solid trace, labeled “F” in FIG. 3), or (2) the same as that of the third harmonic of the missing fundamental (2.67 deg), resulting in convergent DVRs (dashed trace, labeled “3F”). These responses to pure sinusoidal gratings are not surprising—they are exactly as expected from the nearest neighbor principle—and the interesting result is that the vergence responses when quarter-wavelength uncrossed disparities were applied to missing fundamental stimuli were in the opposite direction to the applied disparity, that is, in the direction of the third harmonic. Thus, the DVRs were in the direction expected if the underlying disparity detectors were responding to the principal Fourier component of the disparity, consistent with early spatial filtering prior to binocular matching.

BINOCULAR STIMULI THAT SHARE FOURIER COMPONENTS BUT NOT FEATURES: FURTHER EVIDENCE FOR EARLY SPATIAL FILTERING

We obtained further support for the early spatial filtering hypothesis by recording the DVRs elicited when the two eyes saw patterns that differed in their spectral composition except for some particular Fourier component that differed in phase by onequarter wavelength. For example, one eye saw a missing fundamental grating and the other saw a pure sinusoidal grating that had the same wavelength as the third harmonic of the missing fundamental, but differed from it in phase by one-quarter wavelength (see cartoons in FIG. 4). The two images lack obvious matching binocular features but share a common Fourier component in the spatial frequency domain. When we presented this stimulus combination to three human subjects (MF stimulus: wavelength, 8 deg/cycle and contrast, 74%; 3F stimulus: wavelength, 2.67 deg and contrast, 32%) DVRs were consistently generated at short latency (70 to 80 ms) and were always in the direction expected if the pure sinusoid was being matched to the third harmonic component of the missing fundamental. Sample vergence velocity profiles for one subject are

shown in FIGURE 4: When the matching Fourier components had crossed disparity, the short-latency DVRs were convergent and vice versa. Thus, once more the earliest DVRs were largely determined by the Fourier components of the binocular stimuli, consistent with the idea that monocular signals undergo spatial filtering prior to their binocular combination. Interestingly, it has been shown that when humans see a missing fundamental grating with one eye and a pure sine-wave grating with the other, they can achieve binocular fusion when the sine wave has the spatial frequency of the third harmonic but not when it has the spatial frequency of the fundamental.¹⁵

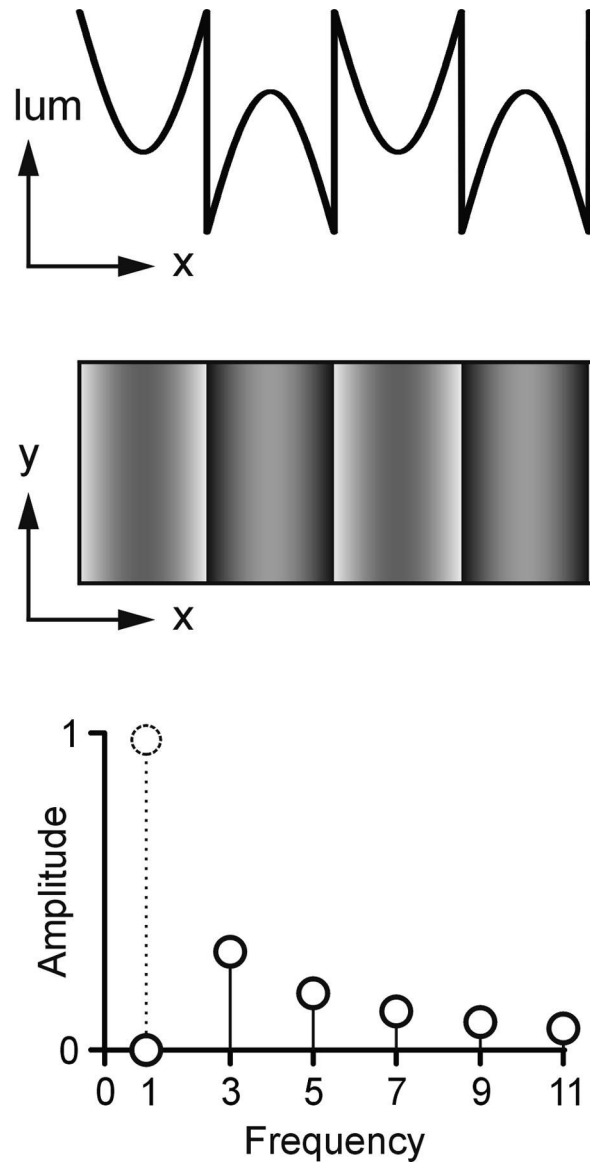
INITIAL DVR: A MODEL SYSTEM FOR STUDYING LOW-LEVEL DISPARITY MECHANISMS?

We have employed complex visual gratings that have long been used by visual psychophysicists to study the mechanisms by which we sense visual motion. These stimuli are relatively new to research on disparity and eye movements, and have revealed that the monocular visual inputs to the disparity detectors mediating the earliest vergence responses undergo spatial filtering prior to their binocular combination. We suggest that these vergence responses provide a promising model system for studying low-level disparity detection mechanisms.

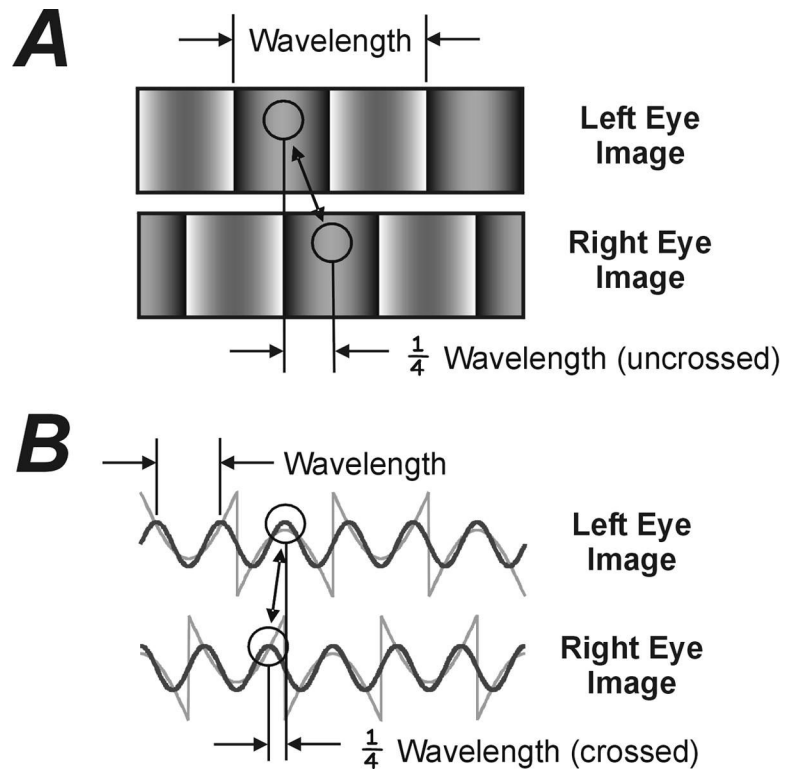
REFERENCES

1. Busetini C, Miles FA, Krauzlis RJ. Short-latency disparity vergence responses and their dependence on a prior saccadic eye movement. *J. Neurophysiol* 1996;75:1392–1410. [PubMed: 8727386]
2. Busetini C, Fitzgibbon EJ, Miles FA. Short-latency disparity vergence in humans. *J. Neurophysiol* 2001;85:1129–1152. [PubMed: 11247983]
3. Masson GS, Busetini C, Miles FA. Vergence eye movements in response to binocular disparity without depth perception. *Nature* 1997;389:283–286. [PubMed: 9305842]
4. Masson GS, Yang DS, Miles FA. Version and vergence eye movements in humans: open-loop dynamics determined by monocular rather than binocular image speed. *Vision Res* 2002;42:2853–2867. [PubMed: 12450510]
5. Takemura A, et al. Single-unit activity in cortical area MST associated with disparity-vergence eye movements: evidence for population coding. *J. Neurophysiol* 2001;85:2245–2266. [PubMed: 11353039]
6. Yang DS, FitzGibbon EJ, Miles FA. Short-latency disparity-vergence eye movements in humans: sensitivity to simulated orthogonal tropias. *Vision Res* 2003;43:431–443. [PubMed: 12536000]
7. Takemura A, Kawano K K, Quaia C, Miles FA. Population coding in cortical area MST. *Ann. N.Y. Acad. Sci* 2002;956:284–296. [PubMed: 11960812]
8. Takemura A, Inoue Y, Kawano K. Visually driven eye movements elicited at ultra-short latency are severely impaired by MST lesions. *Ann. N.Y. Acad. Sci* 2002;956:456–459. [PubMed: 11960839]
9. Cumming BG, Shapiro SE, Parker AJ. Disparity detection in anticorrelated stereograms. *Perception* 1998;27:1367–1377. [PubMed: 10505181]
10. Cogan AI, Lomakin AJ, Rossi AF. Depth in anticorrelated stereograms: effects of spatial density and interocular delay. *Vision Res* 1993;33:1959–1975. [PubMed: 8249313]
11. Cogan AI, et al. Binocular disparity processing with opposite-contrast stimuli. *Perception* 1995;24:33–47. [PubMed: 7617417]
12. Cumming BG, DeAngelis GC. The physiology of stereopsis. *Annu. Rev. Neurosci* 2001;24:203–238. [PubMed: 11283310]
13. Chen KJ, Sheliga BM, FitzGibbon EJ, Miles FA. Initial ocular following in humans depends critically on the Fourier components of the motion stimulus. *Ann. N.Y. Acad. Sci* 2005;1039:260–271. [PubMed: 15826980]
14. Collewyn H, Van Der Mark F, Jansen TC. Precise recording of human eye movements. *Vision Res* 1975;15:447–450. [PubMed: 1136166]

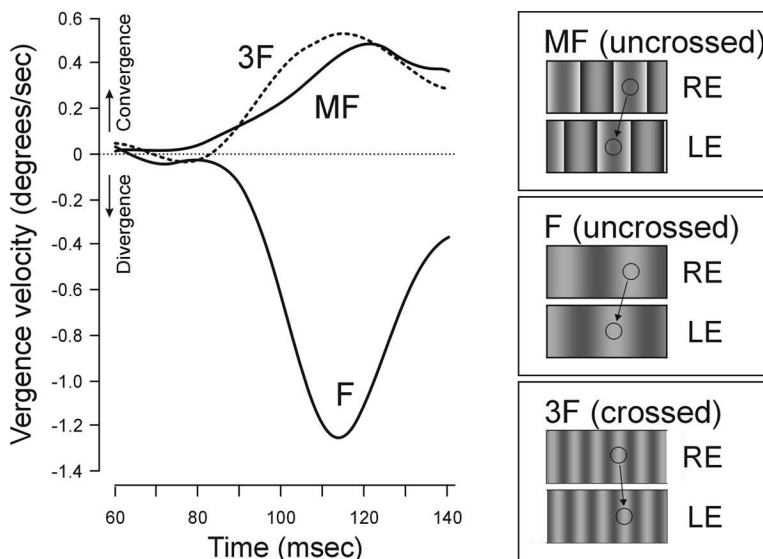
15. Levinson E, Blake R. Stereopsis by harmonic analysis. *Vision Res* 1979;19:73–78. [PubMed: 419703]

**FIGURE 1.**

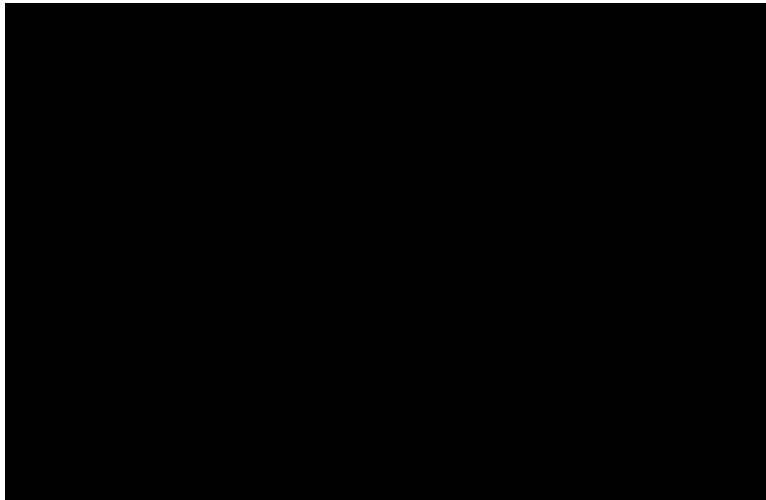
The missing fundamental stimulus. **Top** (*x-lum* plot): luminance plotted as a function of horizontal spatial position (luminance profile). **Middle** (*x-y* plot): luminance as a function of horizontal and vertical position, indicating the visual appearance of the grating pattern on the screen (two-dimensional image). **Bottom**: harmonic content of the spatial stimuli (Fourier spectra). A square wave is made up of the odd harmonics, and the missing fundamental lacks the first harmonic, which is shown in dotted outline.

**FIGURE 2.**

The disparity of the missing fundamental stimulus when defined **(A)** by the matching local features, and **(B)** by the matching principal Fourier components. **(A)** The missing fundamental image seen by the right eye is one-quarter wavelength to the right of that seen by the left eye: an *uncrossed* disparity stimulus. **(B)** The same missing fundamental images as in **(A)**, but now shown as luminance profiles (*gray traces*) along with the principal (third) harmonic (*black traces*). The third harmonic seen by the right eye can be described as *three-quarters wavelength to the right* or *one-quarter wavelength to the left* of that seen by the left eye. The latter relationship is generally the one that prevails because the brain gives precedence to the nearest matches (spatial aliasing).

**FIGURE 3.**

The initial horizontal vergence responses elicited by identical uncrossed disparities applied to various types of grating pattern (sample mean vergence velocity traces from one subject). MF (*solid line*): a 2-deg uncrossed disparity (one-quarter wavelength) applied to a missing fundamental stimulus (wavelength, 8 deg; contrast: 18.5%) elicited a convergent response. F (*solid line*): the same 2-deg uncrossed disparity, when applied to a pure sinusoidal grating (contrast: 8%) that had the same spatial frequency as the fundamental (8 deg), elicited a divergent response. 3F (*dashed line*): the same 2-deg uncrossed disparity, when applied to a pure sinusoidal grating (contrast: 8%) that had the same spatial frequency as the third harmonic (2.67 deg), elicited a convergent response. The cartoons at the right show x - y plots of the stimuli. Upward deflections of vergence velocity correspond to convergent responses. RE, right eye; LE, left eye. Note that the contrast of the MF was arranged to be 18.5%, so that its third harmonic had a contrast that was the same as that of the pure 3F sine wave (8%).

**FIGURE 4.**

The DVRs elicited when one eye saw a missing fundamental grating and the other saw a pure sinusoidal grating that had the same wavelength as the third harmonic of the missing fundamental, but differed from it in phase by one-quarter wavelength (sample vergence velocity profiles for one subject). When the matching Fourier components had crossed disparity the short-latency DVR were convergent (*solid trace*) and when the matching Fourier components had uncrossed disparity the short-latency DVR were divergent (*dashed trace*). The cartoons at the right show *x-luminance* plots of the stimuli. Upward deflections of vergence velocity correspond to convergent responses. RE, right eye; LE, left eye.