# An occlusion-related mechanism of depth perception based on motion and interocular sequence

Shinsuke Shimojo*†, Gerald H. Silverman* & Ken Nakayama*

* The Smith-Kettlewell Eye Research Institute, 2232 Webster St, San Francisco, California, 94115, USA
† Nagoya University School of Medicine, Tsurumai, Showa-ku, Nagoya, 466 Japan

Objects occlude other objects in natural scenes, and this occlusive relationship increases the spatio-temporal complexity of sensory inputs to the two eyes, especially when objects are moving. We ask whether the visual system can employ clever strategies which make use of real-world constraints on inputs to the eyes to determine the depth of objects. Employing psychophysical methods, we found that occlusion-related geometric rules, which constrain the relationship between the direction of motion and the order and asynchrony of eyes, are implemented at early stages of cortical visual processing.

An example of this phenomenon is shown in Fig. 1a. When a vertical bar is moved horizontally behind a narrow vertical slit, there is a time lag between the stimuli to the two eyes. At one particular depth, the end of the stimulus to one eye coincides with the onset of the stimulus to the other eye. We call this the 'perfect relay case'. The direction of motion of the bar constrains the sequential order of the monocular stimuli to the eyes: rightward motion is visible to the right eye first, then to the left eye. On the other hand, leftward motion is visible to the left eye first, and then to the right eye. These are the only real-world possibilities if the bar is single and moving laterally behind the slit. Thus, it is theoretically possible to recover its depth relative to the occluder from its direction of motion, the interocular order and asynchrony. Simultaneous binocular disparity is unnecessary in theory.

The question is whether the human visual system actually uses these real-world constraints to determine depth. To study this question, we simulated the perfect relay case on a cathode-ray tube stereo display. Note that, apart from the interocular asynchrony (IOA), each eye receives exactly the same motion trajectory on corresponding retinal points. Our preliminary observations revealed that a seamless motion of a single target is in fact perceived as lying at a depth behind the slit (P1 in the

figure). The visual system thus interprets the stimuli which are given sequentially to the two eyes as resulting from one object under smooth motion, rather than from two objects at arbitrary depths.

Reversing the sequence of stimuli to the eyes without changing the direction of motion, however, causes a completely different perception (Fig. 1b). Instead of a single target, the observer sees two bars moving side by side which appear just behind the slit (shown as 'X→X→' at the top of Fig. 1b). Note that if the monocularly viewed target was visually localized on a depth plane in front of the slit, such as P2 in Fig. 1b, it would still violate the geometric constraint posed by occlusion because it should be visible to both the eyes.

To confirm these observations, we tested four subjects (including one naive subject), employing a two-alternative, forced-choice procedure. Two oscillating bars, the perfect relay and the reversed ocular order targets, were presented in the top and the bottom halves of slit in a single test run. Their positions were randomized from trial to trial. The slit width was 20 minutes, and the duration of each monocular stimulus was 50 ms. Thus, the velocity of the targets was $6.7°\,s^{-1}$. The subject's task was to judge which target appeared single, when gazing at a fixation point between them. To avoid possible artifacts related to oculomotor convergence errors, two dots, each presented to one of the two eyes, were added such that when they appeared to align vertically, binocular convergence error would be less than 1.7 min. Subjects were asked to monitor the orientation of these two monocular dots, and to make their judgment only when they were vertically aligned. The percentages of 'single' responses attributed to the perfect relay target were 98% for the naive subject, and 97, 100, and 100% for the other three. These quantitative results agree with our original observations: a target moving far behind the slit is perceived as single only when the direction of motion and the interocular order are in the compatible relationship.

To examine the perceived depth of the slit-motion targets in relation to the conventional stereopsis, a second experiment was carried out. The subject's task here was to adjust the stereoscopic disparity of a comparison target to match the perceived depth of the slit-motion target with various IOAs (Fig. 2). The stereo-matching target was moved sideways in synchrony with the slit-motion target (see inset of Fig. 3) without any occluder so that it was always visible to both the eyes. In the results obtained from two subjects (Fig. 3), two points should be noted. First, the target was not perceived in front of the slit in the reversed case (IOA = −50 ms) or in cases with larger negative IOAs. Second, the perceived depth continued to grow when the IOA
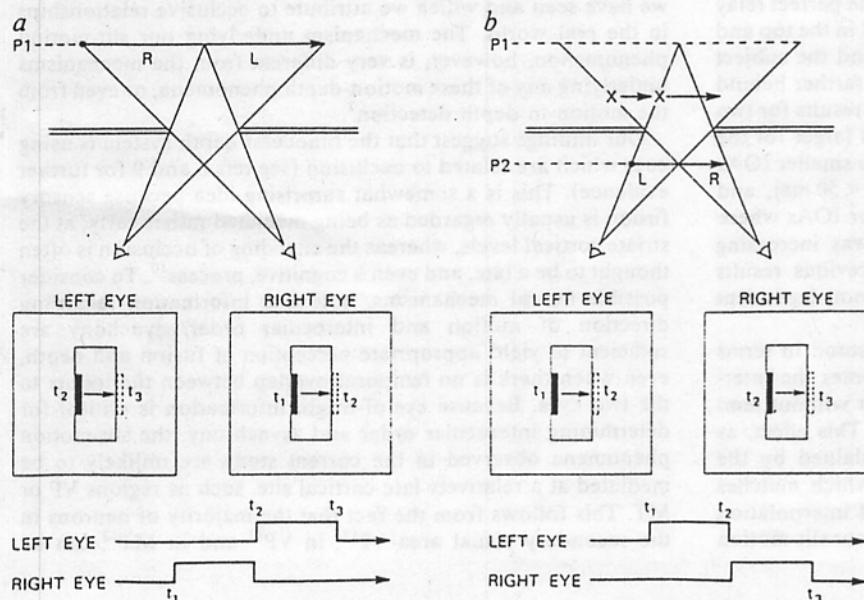


Fig. 1 Schematic diagram of a vertical bar moving behind a vertical slit. The physical structure seen from above (top of figure), the simulated stimulus presented to each eye (middle), and the temporal relationship of monoclular stimuli (bottom) are shown in both Fig. 1a and 1b. a, The perfect relay case where the target appears in the left eye just as the target in the right eye disappears from view. b, The case of reversed ocular order. Smooth motion of a single target at a depth behind the slit was observed in the perfect relay case, whereas two targets moving side by side right behind the slit were observed in the reversed case. In our first experiment, these two stimuli were presented in the top and the bottom halves of the slit. The centre-to-centre separation between the two stimuli was 1.7°, the size of bar target was 32′ × 3.3′ minute, and the slit width was 20′ at the observation distance of 80 cm.
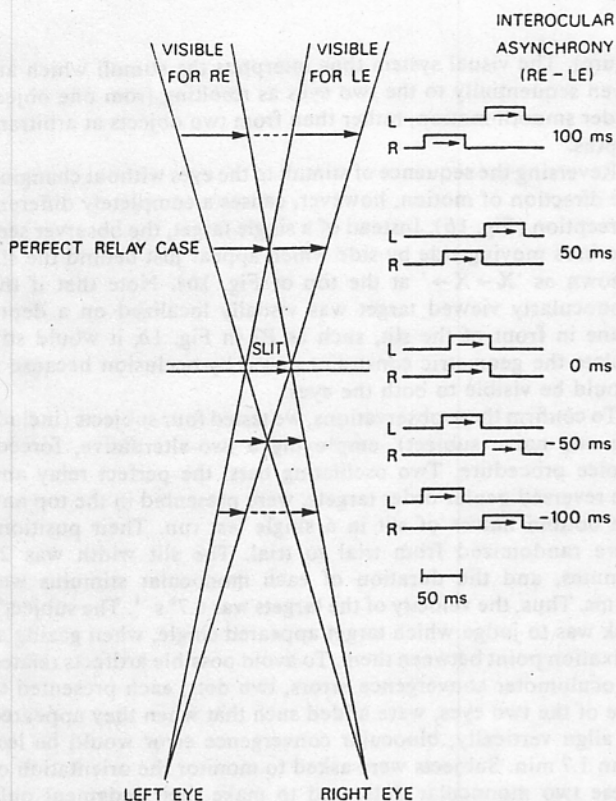
Fig. 2 Relationship between real-world situations of slit view and our simulated motions where an identical target motion is presented within a slit for various degrees of IOA. Note that positive IOAs (right eye followed by left eye when the motion is rightward), including those with no temporal overlap (IOAs greater than the perfect relay case as shown at the top) are compatible with the constraints imposed by occlusion. On the contrary, negative IOAs (left eye followed by right eye) are *not* compatible with real-world situations because targets moving in front of the occluder should always be seen binocularly. RE, right eye; LE, left eye.
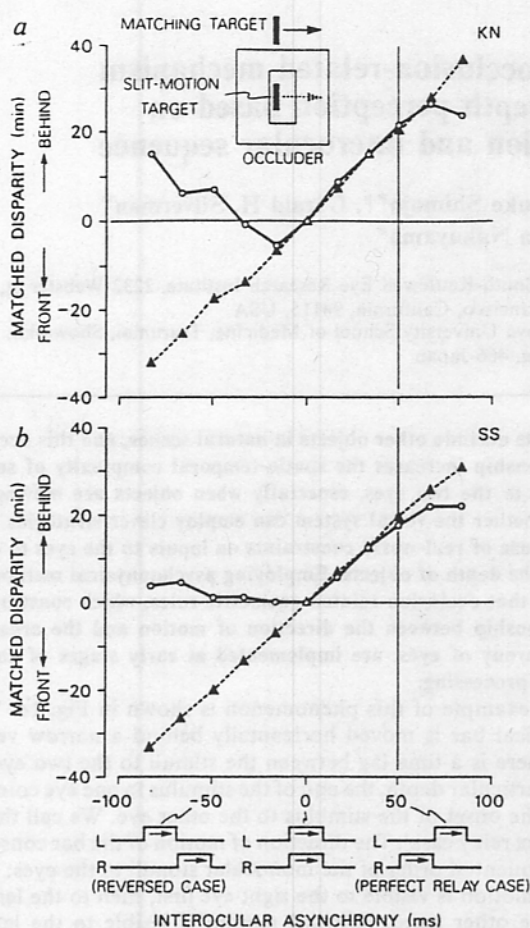
Fig. 3 Perceived depth measured by matching with disparity, as a function of IOA. Two subjects (K.N., top and S.S., bottom) adjusted the disparity of an unoccluded moving target so that it matched the perceived depth of the slit-motion target (open circles). The matching target had a velocity identical to the slit-motion target. Perceived depth of an unoccluded, oscillating stereo target was also measured by the same matching technique (filled triangles). The perfect relay case is marked by a vertical line (IOA = 50 ms). Two findings are noteworthy: first, strong suppressive effects of occlusion cues on negative values of IOA, including the reversed case; second, expanded growth of perceived depth with larger positive values of IOA, even beyond the perfect relay case.

increased beyond that of the perfect relay case so that there was an increasing blank interval between the monocular stimuli (50 ms < IOA < 83 ms). The increase in perceived depth with this increasing blank interval cannot be accounted for by the ordinary mechanism of stereopsis which is based on disparities of simultaneous binocular stimulus.

The depth relationship for increasing IOAs beyond the perfect relay case was examined in more detail by our third experiment, in which slit-motion targets with various positive IOAs (0–117 ms) were paired with the standard target of the perfect relay (IOA = 50 ms). The paired stimuli were presented in the top and the bottom halves of slit (see inset of Fig. 4), and the subject was asked to designate the target that appeared farther behind in a forced-choice paradigm. As shown in Fig. 4, results for two subjects confirmed first that perceived depth was larger for the perfect relay case (IOA = 50 ms) in comparison to smaller IOAs where the monocular stimuli overlapped (IOA < 50 ms), and second that perceived depth was clearly larger for IOAs where there was no overlap and the blank interval was increasing (50 ms < IOA < 100 ms). Thus these and our previous results (Fig. 3) argue for a special mechanism that is not dependent on the processing of fusible disparity.

Our findings taken together cannot be understood in terms of the stereoscopic depth effect which accompanies the interocular delay of a stroboscopically shifting target without temporal overlap of the stimuli to the two eyes[1-3]. This effect, as well as the 'Pulfrich effect'[4], can be fully explained by the conventional, binocular stereo mechanisms[5,6], which matches the neural effects of monocular spatio-temporal interpolation at a central neural site. These phenomena, unlike our slit-motion

phenomenon, are not related to the occlusion constraint, and do not show the asymmetrical effects of disparity reversal which we have seen and which we attribute to occlusive relationships in the real world. The mechanism underlying our slit-motion phenomenon, however, is very different from the mechanisms underlying any of these motion-depth phenomena, or even from the motion in depth detection[7].

Our findings suggest that the binocular depth system is using cues which are related to occlusion (see refs 8 and 9 for further evidence). This is a somewhat surprising idea because sensory fusion is usually regarded as being mediated rather early, at the striate cortical levels, whereas the encoding of occlusion is often thought to be a late, and even a cognitive, process[10]. To consider possible neural mechanisms, note that information regarding direction of motion and interocular order/asynchrony are sufficient to yield appropriate perception of fusion and depth, even when there is no temporal overlap between the inputs to the two eyes. Because eye-of-origin information is critical for determining interocular order and asynchrony, the slit-motion phenomena observed in the current study are unlikely to be mediated at a relatively late cortical site, such as regions VP or MT. This follows from the fact that the majority of neurons in the secondary visual area V2[11], in VP[11] and in MT[12] can be
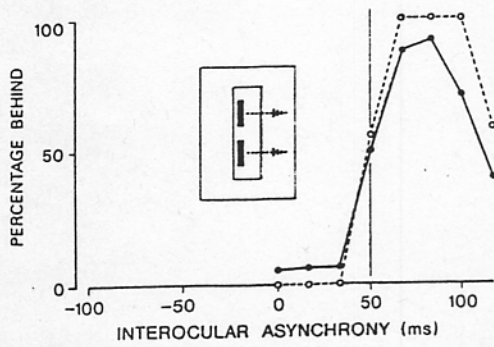
**Fig. 4** Relative depth for different IOAs with the perfect relay case as the standard. For both subjects (o, K.N. and ●, S.S.), increasing percentages of the 'behind' response were found for IOAs greater than the perfect relay case (solid vertical line).

driven by either eye, and do not appear to retain explicit eye-of-origin information. On the other hand, many neurons in primary visual area V1 are selectively responsive to the eye of stimulation and to the direction of motion[13]. Moreover, some neurons in V1 and V2 are apparently selective to interocular order and asynchrony[14], which may play critical roles in the slit-motion phenomenon.

1. Lee, D. *Vision Res.* 10, 65–78 (1970).
2. Morgan, M. J. *Nature* 256, 639–640 (1975).
3. Burr, D. C. & Ross, J. *Vision Res.* 19, 523–532 (1979).
4. Pulfrich, C. *Naturewissenschaften* 10, 533–564, 569–601, 714–722, 735–743, 751–761 (1922).
5. Julesz, B. *Foundations of Cyclopean Perception* (University of Chicago Press, 1971).
6. Marr, D. & Poggio, T. *Proc. R. Soc.* B204, 301–328 (1979).
7. Beverley, K. I. & Regan, D. *J. Physiol., Lond.* 235, 17–29 (1973).
8. Ramachandran, V. S. & Cavanagh, P. *Nature* 317, 527–530 (1985).
9. Gregory, R. L. & Harris, J. P. *Percept. Psychophys.* 15, 411–416 (1974).
10. Rock, I. *The Logic of Perception* (MIT Press, Cambridge, 1983).
11. Burkhalter, A. & Van Essen, D. C. *J. Neurosci.* 6, 2327–2351 (1986).
12. Maunsell, J. H. R. & Van Essen, D. C. *J. Neurophysiol.* 49, 1148–1167 (1983).
13. Hubel, D. H. & Wiesel, T. N. *J. Physiol., Lond.* 195, 215–243 (1968).
14. Gardner, J. C., Douglas, R. M. & Cynader, M. S. *Brain Res.* 328, 154–157 (1985).