Citation: Morgan, M. (2023). Stereopsis for rapidly moving targets. Philosophical Transactions of the Royal Society B: Biological Sciences, 378(1869), 20210462. doi: 10.1098/rstb.2021.0462

Stereopsis for rapidly moving targets

Michael Morgan
Division of Optometry
City, University of London

Word Count 2879

Correspondence address:
michaelmorgan9331@gmail.com

Abstract

Stereoscopic depth perception is possible with luminance-defined target velocities at least as high as 600 deg/sec, up to the limit of 30 Hz imposed by the high-temporal frequency cut-off of the eye. . The limitation for perceiving depth from stereo disparity of moving targets is not their velocity but the temporal frequency bandwidth of the eye, which is affected by adaption state. Stereoacuity for a depth shift in a horizontally moving grating depends not on spatial disparity between corresponding luminance points in spatial units of arc min, but on the spatial shift as a fixed proportion of the period of the grating, in other words, on the phase angle difference between the two eyes, as is also the case for obliquely-orientated, stationary gratings. Phase differences explain not only the classic Pulfrich stereophenomenon but its equivalent with dynamic visual noise, and a new effect in which depth results from interocular phase differences in luminance modulation.

Jenny Read's fascinating contribution to this volume makes it clear that prey catching in a variety of animals could involve stereopsis. But stereopsis for a rapidly moving target raises several problems, not least the problem of correspondence, given the relatively sluggish temporal response of photoreceptors. The typical impulse response of cones (their response to a brief flash of light) rises gradually to a peak, and then subsides. Crucially, the time to reach the peak depends on the adaptation states of the receptor and thus on the background light level. The time to peak can be anything from tens of milliseconds at high light levels to hundreds of milliseconds at low levels (Baylor & Hodgkin, 1974; Donner, 2021). This means that a fast-moving target will leave behind it a trail of activity without any precise location. Nevertheless, human stereoacuity thresholds for periodic gratings (vertically oriented, horizontally moving from left to right) is not degraded by velocities up to 640 deg/sec provided that their temporal frequency does not exceed 30 Hz (Morgan & Castet, 1995). The spatial frequency of a grating is the number of cycles of the grating in a degree of visual angle. The temporal frequency is the number of bars of the grating that will pass a fixed point in a second. If we measure the luminance of the grating at a fixed point as it passes, the luminance will cycle sinusoidally. The angle of the sinusoidal modulation is refered to as the phase angle. If there is a disparity between the gratings in the two eyes this will produce a difference in phase angle, which can have a maximal value of 180 deg. It turns out that the best way of characterising the limits of stereopsis for gratings so that it is independent of spatial frequency and velocity, is the phase angle difference between the eyes. This comes out at about 5 deg, which can -depending on spatial frequency and velocity- correspond to an inter-ocular temporal delay of as little as 450 microseconds (Morgan & Castet, 1995).

Although the ~5 deg phase difference limit has been measured for periodic sinusoidal gratings, the same principal can be presumed to apply to a single moving bar, or any other stimulus that can be decomposed by Fourier Analysis into its component sinusoids. With a square wave grating, for example, it is likely that the upper temporal frequency limit would be the same as for its lowest frequency component (the fundamental), and would be in the region of 30 Hz, depending on adaptation level.

It may be helpful to clarify that the upper temporal frequency limit we are discussing here is for luminance defined targets such as moving sine-wave gratings, not the limit for temporal modulations in disparity itself, which are much lower, for example, the 6 Hz limit for oscillation in depth of a grating defined by a random-dot stereogram (Norcia & Tyler, 1984), or the 8 Hz limit for a moving triangle-wave modulation of disparity, also in a random-dot sterogram (Kane, Guan & Banks, 2014). These much lower limits reflect the temporal frequency limits of hte process underlying the computation of the interocular cross-correlation underlying the detection of disparity (Kane, Guan & Banks, 2014).

One dimensional motion is tilted in axes {x,t}. Similarly, if we consider stereoacuity for static gratings that are tilted in space, we have axes {x,y} and once again, in this space the limit for stereoacuity is set by the phase angle difference, not just the horizontal component of the disparity. Over a range of tilts for a Gabor patch ranging from 0 to 80 deg, the stereoacuity threshold was found to be constant over tilt angle when expressed as a phase shift, though not when expressed as a horizontal disparity (Morgan & Castet, 1997).

Stereopsis from inter-ocular temporal delay is demonstrated by the Pulfrich (1922) stereophenomenon, in which a neutral density filter over one eye causes the perceived position of a moving target to be shifted in depth. Pulfrich (who incidentally had only one functioning eye) accepted the explanation by the optical engineer Fertsch at Carl-Zeiss, that the filter caused a transmission delay in the covered eye by delaying the response of the photoreceptors in the covered eye. The delay meant that the covered eye was signalling an earlier position of the target along its trajectory, and thus an instantaneous disparity. Pulfrich quoted from Richard Wagner's "Parsifal": "*Du siehst mein Sohn zum Raum wird hir die Zeit"* (You see my son, that here time is changed into space).

The key word here is "instantaneous". The instantaneous position of a moving object is an abstraction, and cannot be measured by any physical system, let alone the eye with its relatively sluggish response time in the millisecond order of magnitude. The neutral density filter complicates this by decreasing the retinal illumination and delaying the peak response still further, as the eye attempts to maintain sensitivity by integrating photons over time.

The Pulfrich effect could thus be due to greater persistence of the signal in the covered eye, which would push its average position at any time back along the trajectory. The effect of differential persistence can be mimicked by making the target in one eye larger in one eye along the trajectory, but with its leading edge still aligned with the target in the other eye. This produces a depth effect, which can be cancelled by a neutral density in the opposite eye (Morgan, 1977).

However, depth-from-delay can also be demonstrated without differential persistence, using a haploscope with control of the phase angle between the two eyes. Initially, this was done with continuously moving targets, but David Lee pointed out that a depth shift could also be seen with a stroboscopically illuminated target (Lee, 1970), where the target moves in a series of discrete steps. This raises a problem for the simple Fertsch-Pulfrich 'time to space' theory, because delay cannot alter the retinal position of the target, only the time taken for that position to be signalled to the brain. A possible explanation is that the brain would fuse the n'th position of the target in the delayed eye with the (n-1)th position in the other eye. To see if this was the explanation, Morgan & Thompson (1975) used sinusoidally moving , stroboscopically illuminated targets and counteracted the depth effect caused by a ND filter with a phase difference between the eyes. Below stroboscopic sampling rates of about 30 msec, the equivalent temporal delay required to null the depth effect was less than half the sampling interval, contradicting the predictions of the 'shifted pairing' hypothesis. It thus appeared as if the relevant measure of disparity is not the position of the samples but the spatial disparity between the underlying motion trajectories. Using both a vernier and a stereo method, Morgan (1976) showed that targets in apparent (stroboscopic) motion appeared aligned when they had the same underlying motion trajectory, even if that trajectory were sampled at different times and places, provided that the temporal sampling interval did not fall below ~ 40 Hz (depending on adaptation state, Morgan, 1979). Burr & Ross (1975) elegantly confirmed this by measuring a depth effect from temporal delays of as little as 160 microseconds or 2 arcsec, in spatially corresponding positions.

Interpolation of the perceived position of a target in sampled motion would be expected in any system with a limited temporal frequency response. A moving target sampled at 1000 Hz could not in any way be distinguished by the visual system from one sampled at 100 Hz, provided the eye is not tracking the target, which was controlled in at least some experiments (Morgan, 1976, 1980) . In Fourier terms a discretely sampled moving target contains the spatio-temporal spectrum of the underlying trajectory, with the addition of components arising from the sampling, The higher the sampling rate the greater the separation between the trajectory components and the sampling replicas (Fahle & Poggio, 1981). If the sampling replicas are of sufficiently high frequency, the human visual system will not be sensitive to them, and only the underlying trajectory will visible. However, if the replicas fall within the visible range of frequencies they will interfere with the original signal and interpolation will break down. This happens in human vision when the sampling rate falls below a critical value, in the region of 20 Hz at typical experimental background luminance levels . When interpolation fails, it can be restored by external filtering to remove the closest replicas, or by reducing the mean luminance of the signal (for a review of these and other methods see a previous article in this journal; Morgan, 1980).

An exception to the interpolation rule is the little-noticed report of perceived depth in a moving bar target that had the same motion trajectory in the two eyes (Morgan & Watt, 1982). The moving bar was subjected to a sinusoidally modulation of its luminance at ~ 10 Hz, with the modulation being out of phase in the two eyes. Thus, the instantaneous spatial positions of the bar was the same in the two eyes, but one eye was temporally lagging in its luminance modulation relative to the other. This luminance modulation produced a perceive depth shift of the moving bar, as if the temporally lagging bar (in its luminance modulation) were spatially leading its partner in the other eye. Once again, the threshold phase angle for seeing the effect was about 5 deg. To form a picture of this phenomenon, imagine that the targets had been photographed with the camera shutter open for several seconds.. Each target would produce a sinusoidal grating with the phase-lagging grating shifted in the direction of motion. (Fig. 1) The fused image would be seen as shifted in depth in the same direction as the previously moving target. Temporal smearing in the eye and brain means that it is a kind of storage device, so it is not entirely surprising that a phase lag in luminance modulation produces depth. A full model for the effect has yet to be developed (it would be a good student project), but it is easy to see by example that a mechanism sensitive to interocular phase would be affected by temporal luminance modulation. Consider, for example, a quadrature pair of receptive fields, in corresponding retinal positions, which would respond maximally to a 90 deg disparity of a stationary Gabor target (deAngelis, Ohzawa & Freeman, 1991). This would also respond with greater amplitude to 90 deg out-of-phase temporal modulation than to in-phase (0 deg). See Fig. 1.

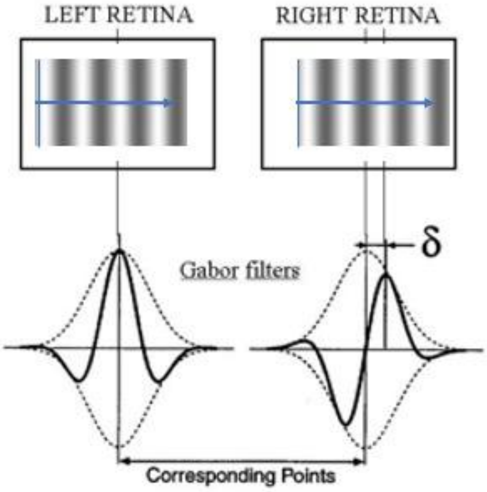

Figure 1: Morgan & Watt (1982) used separate moving bars to the left and right eyes in a Dichoptic display. The movement vector of the bars is shown by arrows. These vectors were identical in the two eyes, and so was the instantaneous position of the bars. However, the luminance of the two bars was modulated sinusoidally so that that over time they would have traced out two sinusoidal gratings, with a phase difference between the eyes. Apparently, the brain was sensitive to the temporal modulation, because the phase difference produced an apparent depth shift in the moving bar. A pair of quadrature receptive fields in the two eyes (deAngelis, Ohzawa & Freeman, 1991) would respond with greater modulation depth to out-of-phase (90 deg) luminance modulation than to in-phase (0 deg).

An interocular phase limit for stereopsis also appears using dynamic visual noise as a stimulus. Tyler (1974) and Ross (1974) reported that dynamic visual noise viewed with an interocular delay appeared to swirl around in depth like a transparent cylinder rotating around a vertical axis. At first sight this is mysterious since the noise does not appear to be horizontally moving until the delay is applied. But since the spatio-temporal spectrum of the noise is broad-band, it must contain horizontally moving components, which could in theory be susceptible to stereo-tuned mechanisms. Horizontal movement of a horizontal Fourier component is invisible, so we can predict that noise containing only horizontal components will not support the stereophenomenon, while noise containing vertical components can. This was verified by Morgan & Tyler (1995) using a cylindrical lens as a filter (a piece of shower stall, as it happens).

To measure the frequency dependence of the noise stereophenomenon, Morgan & Fahle (2000) used band-limited dynamic visual noise composed of a checkerboard array, the elements of which were sinusoidally modulated in luminance over time, with the phase of each element randomly chosen from a uniform distribution (Fig. 1). The threshold phase angle for reporting the direction of rotation (Clockwise or Anticlockwise) was in the region of 5-10 deg, which corresponded to temporal delays in the millisecond region at frequencies of ~10 Hz. As in Morgan & Castet's experiment there was an upper temporal frequency limit for the luminance modulation above which the depth effect could not be seen. In Morgan & Castet's experiment this was ~ 30 Hz; in the dynamic noise stereophenomenon it was lower: ~ 6 Hz in one observer and ~ 19 Hz in the other. The different limits in the two experiments is not remarkable. The dynamic visual noise stimulus contains many different spatio-temporal components and orientations, of which the relevant component for stereopsis is only a weak part. Measurements of contrast sensitivity for flickering gratings show a rapid fall off above 10 Hz. The observers in the dynamic visual noise experiment were presumably too insensitive to higher temporal frequency components to be able to use them for stereopsis.

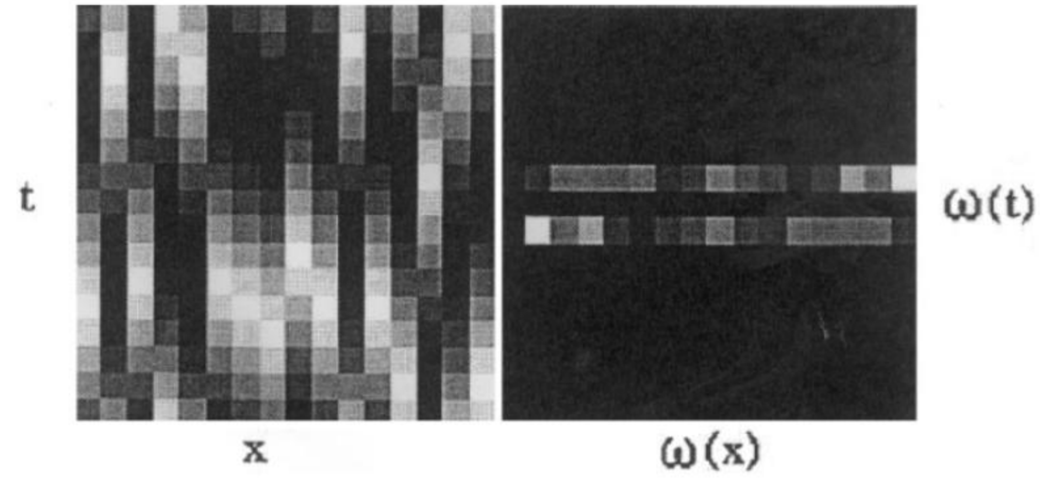

Fig. 1. The left-hand panel is an *x–t* plot of a single row of the random-noise stimulus used in the experiments. The figure shows how a single row of 2-D white noise changes over time (vertical axis). The right-hand panel shows the corresponding power spectrum, which contains temporal frequencies of $\pm\omega$ but a broad-band of spatial frequencies. Thus there are as many velocity components in the stimulus as spatial frequency components.

Sensitivity to high temporal frequencies increases with retinal illuminance. This is assumed to be because the speed of the phototransducation response increases with retinal illuminance (Baylor & Hodgkin, 1974) . It is clear from the temporal frequency limits on dynamic stereoacuity that prey-catching will therefore be best at high levels of illumination. Hovering birds of prey, looking for the tiniest movement of their prey on the ground, prefer to hunt on sunny days. Temporal frequency limits also explain why bad light stops play in Cricket. Bad light does not slow down the action of the bowler, but it does slow down the batsman's retina. The only remaining mystery is why some fielders persist in wearing fashionable sunglasses late into the evening.

References

Baylor, DA & Hodgkin, AL (1974) Changes in time scale and sensitivity in turtle photoreceptors. J. Physiol., 242, 729-758.

Burr, DC & Ross, J (1979) How does binocular delay give information about depth? Vision Research, 19, 523-532

deAngelis, D.G., Ohzawa, I. & Freeman, R. D. (1991) Depth is encoded in the visual cortex by a specialised receptive field structure, Nature, 352-156, 159.

Donner, K. (2021) Temporal vision: Measures, Mechanisms and Meaning. J. Exp. Biol., 224 (15) jeb222679.

Fahle, M. & Poggio, T. (1981) Visual hypearcutiy: spatiotemporal interpolation in human vision. Proc. Roy. Soc. B, 231,451-477

Kane, D., Guan, P., & Banks, M. S. (2014). The limits of human stereopsis in space and time. Journal of Neuroscience, 34(4), 1397-1408.

Lee D N, 1970b "A stroboscopic stereophenomenon" Vision Research 10 587-593

Morgan, M. J. and Thompson, P. (1975). Apparent motion and the Pulfrich effect. Perception 4, 3-18.

Morgan, M. J. (1976). Pulfrich effect and the filling in of apparent motion. Perception 5, 187-195.

Morgan, M. J. (1977). Differential visual persistence between the two eyes: a model for the Fertsch-Pulfrich effect. J.exp.Psychol. 3, 484-495.

Morgan, M. J. (1979). Perception of continuity in stroboscopic motion: a temporal frequency analysis. Vision Research 19, 491-500.

Morgan, M. J. (1980). Analogue models of motion perception. Phil.Trans.Roy.Soc. B 290, 117-135.

Morgan, M. J. and Watt, R. M. (1982). Hyperacuity for luminance phase angle. Vision Research 22, 863-866.

Morgan, M.J. & Castet, E. (1995) Stereoscopic depth perception at high velocities. *Nature, 378*, 380-383.

Morgan,M.J. and Castet, E. (1997) The aperture problem in stereopsis. *Vision Research, 37*, 2737-2744

Morgan,M.J. & Tyler, C.W. (1995) Mechanisms for dynamic stereomotion respond selectively to horizontal velocity components. *Proc. Roy.Soc. Lond.B*, 262, 371-376.

Morgan, M.J. & Fahle, M. (2000). Motion-stereo mechanisms sensitive to interocular phase. *Vision Research 40, 1667-1675.*

Norcia, AM & Tyler, CW (1984) Temporal frequency limits for stereoscopic apparent motion process. Vision Research, 24, 395-401.

Pulfrich, C.(1922) Naturwissenschaft 10, 553-564

Ross, J, (1974) Stereopsis by binocular delay. Nature, 248,363-4

Tyler, CW (1974) Stereopsis with dynamic visual noise Nature, 250, 363-4.