TEMPORAL SAMPLING REQUIREMENTS FOR STEREOSCOPIC DISPLAYS

MARC GREEN
COMPUTER STUDIES
TRENT UNIVERSITY
PETERBOROUGH, ONTARIO K9J 7B8

1. ABSTRACT

A series of three experiments examined temporal aliasing in stereoscopic displays. The first experiment compared aliasing in frontal plane motion of different disparities, while the second compared aliasing for motion in different depth directions. The results showed little effect of viewing conditions on perceived aliasing. The third experiment tested whether there was a binocular motion mechanism which integrated temporal sampling in the two eyes. The results were consistent with the first two studies in suggesting that aliasing is generated only by monocular motion signals. The data have both practical and theoretical implications: 1) motion produced by means of LCD glasses will require double the sampling rate needed for motion created by anaglyph methods and 2) the short-range motion system is monocular.

2. TEMPORAL ALIASING

Digital displays present spatially and temporally sampled versions of continuous images. Sampling, however, can create "aliasing" artifacts which result in unwanted image distortions and reduced fidelity. Because aliasing is such an important problem in image quality, there have been many studies examining the effects of spatial (e. g., Nyman and Laurinen, 1982; Nyman and Laurinen, 1985) and temporal (e.g., Watson *et al.*, 1986; Green 1992a) sampling in video displays. This is an important research topic because knowledge visual mechanisms which interpolate sampled signals may suggest new methods of image compression.

The advent of stereoscopic displays presents a new opportunity to improve the realism of video displays but also opens a new set of questions about sampling and aliasing. There has been some research on spatial sampling (e. g., Tzelgov *et al.*, 1990) in stereoscopic displays but no studies on temporal sampling requirements.

There are two possible ways that stereo displays might alter temporal sampling requirements. First, viewers usually verge on the plane of the image, so nonstereo displays present information to corresponding points on the two retinae. In stereo displays, however, the information may fall on noncorresponding points. This activates cortical disparity detectors in which are not involved in nonstereo viewing. There is clear psychophysical evidence that humans possess stereomotion detectors with very different properties from mechanisms for frontal plane motion (Regan, 1991). Moreover, there is other evidence (e. g., Tyler, 1971) that frontal plane and stereo motion mechanisms interact. Second, stereoscopic displays afford different ways to distribute the temporal sampling between the two eyes. In nonstereo displays, each eye receives images sampled at the same

0-8194-0823-9/92/$4.00

time. Stereo displays permit the monocular sampling to be staggered, effectively producing a higher binocular sampling rate.

The study described below compares temporal aliasing in motion through different depth planes and directions. Of primary interest is the relation of stereomotion, motion produced by changing image disparity in the two eyes, and frontal plane motion, motion in the absence of disparity change. The major goal was to determine 1) how motion signals from the two eyes combine and 2) whether aliasing in stereomotion is greater, equal or less than in nonstereo (monocular) motion.

Below, I will first outline the effects of sampling on the spatiotemporal spectrum of a moving image. The subsequent section briefly reviews previous research about the stereo motion system and discusses possible relationships between frontal plane and stereo motion. Next, I describe a series of experiments which test the possible predictions. The final section discusses the practical and theoretical implications of the data.

3. SPATIOTEMPORAL SPECTRA OF MOVING OBJECTS

According to the sampling theorem (Shannon, 1949), a sampled continuous function can be completely reconstructed if 1) the sampling rate is equal or greater than the Nyquist rate, twice the highest frequency and 2) the replicas are filtered away, leaving only the original spectrum. Anti-aliasing is the process correcting the image through removing the spurious spatiotemporal components.

The filtering model is applicable to both spatial and temporal anti-aliasing because digital video images are sampled in both space and time. Spatial sampling is due to the necessity of displaying the image by an array of discrete pixels or lines. Temporal sampling is limited by the 60 Hz update rate of standard display monitors.

The interaction of spatial and temporal aliasing is best seen by representing objects in a spatiotemporal space (Figure 1A). The Y axis shows spatial frequency, and the X axis represents temporal frequency. The quadrants of the space show direction of motion. Even quadrants correspond to rightward motion and odd quadrants to leftward motion.

The figure 1A shows the spectrum of a continuously moving spatial impulse, an infinitely thin line which has equal spectral energy at all spatial frequencies. The velocity of the impulse is -1/slope and direction is rightward. Because temporal frequency is equal to velocity divided by spatial frequency, each spatial frequency of a constant velocity object is moving at different temporal frequency.
The spectrum of a temporally stroboscopically sampled impulse in shown in Figure 1B. Sampling creates "replicas" of the original spectrum at various location along the X axis. The separation of the replicas is proportional to the sampling rate. If the sampling rate is increased, the original and replicas become more widely separated and distinct.

The dashed square shows that human spatiotemporal sensitivity can be described as a filter having a bandwidth of 60 Hz temporally and 60 c/deg. spatially. Since the replicas fall within the visual filter, they can be detected and produce the jitter seen in coarsely sampled motions. If the sampling rate is increased, as shown in Figure 1C, the replicas are pushed further from the origin. When the replicas fall entirely outside the filter and are undetectable, then smooth and sampled motion will be indistinguishable (Watson *et al.*, 1986; Green, 1992a). See Watson *et al.*, (1986) for a more complete discussion.

The "staircase motion" more typical of digital displays can be similarly analyzed into frequency components. Morgan (1980) depicts motion as a function in space-time, rather than spatiotemporal frequency (Figure 2). Smooth motion will produce a ramp while sampled motion is a staircase, which can be decomposed into a ramp plus sawtooth-wave components. The period of each sawtooth-wave component corresponds to the period that the line stands at each "station." Smooth motion is perceived as long as the sawtooth components are not detectable, i. e., lie beyond the sensitivity of the human temporal filter.

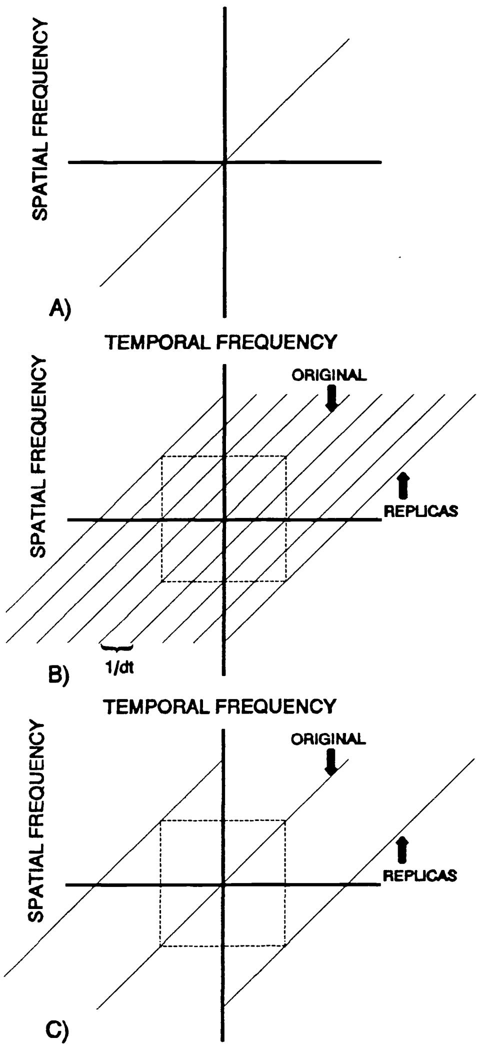


Figure 1

The spatiotemporal spectrum is merely an image representation, not a theory of aliasing. (In fact, the term "aliasing" is highly ambiguous, since it has been used to describe both a mathematical computation and a perceptual experience). In order to derive predictions about aliasing, the spatiotemporal representation must be combined with a psychophysical model of human motion perception. The next section very briefly outlines some relevant psychophysical models and uses them, along with the spatiotemporal spectrum, to produce predictions of sampling requirements in stereoscopic displays.

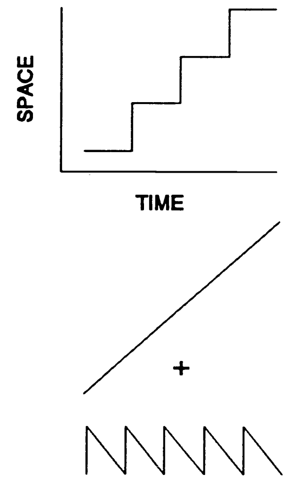
4. COMPARISON OF FRONTAL AND STEREO MOTION

Psychophysicists have postulated several distinct motion detection mechanisms. A popular theory (Braddick, 1974) suggests that there are two distinct mechanisms for detecting frontal plane motion. One is called the "short-range system" because it detects motion which is smooth or has small discontinuities. It presumably reflects the operation of motion sensors located very early in the visual system. The "long-range system" system detects motion with very large displacements and seems to involve higher level, perhaps even cognitive, mechanisms (Exner, 1888).

This distinction is relevant here because it has been frequently suggested that the while the long-range system is binocular, the short-range system does not combine information from the two eyes. For example, studies (Braddick, 1974; Green and Blake, 1981) have shown that a sequence of two image frames which produce motion when presented sequentially to one eye, produce only flicker when presented to left and right eyes. The spatiotemporal filtering approach presumably reflects only the short-range, low-level system.

However, there has been some recent controversy because Shadlen and Carney (1986) found evidence for binocular integration. Georgeson and Shackleton (1989) appeared to rebut this evidence, but the validity of the short/long range distinction is still unsettled.

Another dichotomy divides mechanisms for motion in the frontal plane from stereomotion produced by changing disparities in the two eyes. Regan (see 1991 review) postulates that motion is initially coded in each eye by monocular mechanisms tuned to left and right frontal motion. These combine cortically in various pairs to create 4 distinct stereomotion detectors, each broadly tuned to a different direction in depth.



(AFTER MORGAN, 1980)

Figure 2

The stereo and frontal mechanisms seem to have several major differences. For example, several studies (e.g., White and Odom, 1985; Regan, 1991) have found that the stereomotion system has a much lower temporal frequency response than the frontal motion detectors. This insensitivity to high frequencies may be necessary to prevent changes in apparent depth due to small, rapid shifts in vergence (Ogle and Weil, 1962).

Tyler (1971) further found that sensitivity to stereomotion was lower than for frontal plane motion: observers could perceive frontal motion when the left eye saw a line displaced laterally and the right eye a blank field, but not when the right eye simultaneously viewed a similar line moving in the opposite direction. Tyler hypothesized the result was due to "stereo suppression" of monocular motion sensitivity by stereo mechanisms.

Unfortunately, these psychophysical mechanisms lead to ambiguous predictions about the relationship between frontal plane and stereo motion aliasing. Based on previous research, it is possible to make a reasonable argument that aliasing in stereomotion with be equal to, greater than or less than aliasing in frontal plane motion! Here are the three lines of logic:

1) **Equal aliasing**: The aliasing signal could arise from either the frontal plane motion detectors or from the stereomotion detectors into which they feed. The stereomotion detectors have a much lower temporal frequency response and would not be as sensitive to the high frequency artifacts, the sawtooth components in Figure 2, created by sampling. This suggests that if we have perceptual access to the monocular signal, aliasing should be governed by the more sensitive frontal plane detectors regardless of whether there is stereomotion or not. This would be consistent with the view that the short-range mechanism is monocular.

2) **Lower aliasing**: The converse argument says that aliasing does not arise until after inputs from the two eyes combine. Since stereomotion detectors are less sensitive to high frequencies, they will produce less apparent aliasing. Moreover, Tyler's (1971) work on stereo suppression suggests that the stereo mechanisms may actually inhibit frontal plane motion detectors.

3) **Higher aliasing**: The spatiotemporal model suggests that the prime determinant of aliasing is the frequency of sampling in space. The question is open, however, whether space is defined in two or three dimensions. Objects with very similar 2-D motions may have a variety of 3-D paths, depending on relative disparities. If there is greater disparity, then the apparent 3-D path length increases although the 2-D path (and motion on the retina) remains roughly constant. This produces a simple set of predictions: 1) if the spatial metric is 2-D, sampling requirements would not be much affected by a longer path due to motion in depth or 2) if the metric is 3-D, sampling rates would increase with the longer apparent path created by increased disparity.

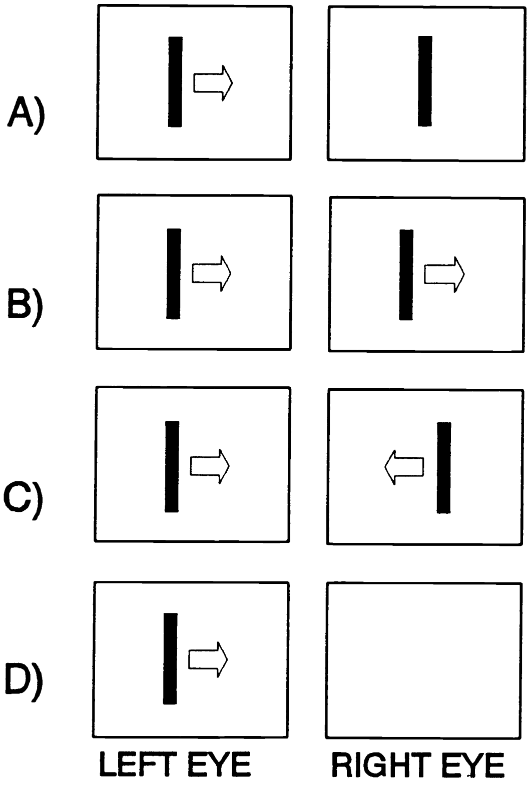


Figure 3

The question of whether the metric is 2- or 3-D is unresolved (see review by Green, 1992b), but it has already been shown (Green and Odom, 1986) that the visual system uses a 3-D metric for performing motion correspondence matching. Similarly, Attneave and Block (1973) found that optimal interstimulus interval between successive stations increases when apparent path length is longer due to monocularly induced motion in depth. This suggests that Korte's Third Law (Korte, 1915) is also based on a 3-D spatial metric. However, both of these studies probably stimulated long-range motion detectors.

The experiments described below compared temporal sampling requirements for objects moving in the a variety of depth planes and directions. Results are generally consistent with the view that aliasing arises from the monocular motion signal rather than from binocular mechanisms. This adds support to the view that the short-range motion system is strictly monocular. It appears that stereoscopic displays offer no means for reducing temporal sampling requirements in digital displays. On the contrary, the use of LCD glasses will likely increase sampling requirements because there is no binocular integration of motion samples.

5. EXPERIMENTS

5.1 Methodology

5.1.2 Images

The observers viewed images generated by an IPS Spectrum Master Graphics controller and displayed on a Mitsubishi 60 Hz, noninterlaced monitor. Disparity was created by dividing the screen into left and right "windows" covered with orthogonally oriented polarizing filters. Observers viewed the windows through a second pair of polarizing lens in front of a phoropter. Eye polarization was arranged so that the left eye saw only the right image, and their right eye saw only the left image.

The windows were 32 cd/m^2 rectangular areas of 2.5 wide X 2.0 degrees high. They appeared on a black background so that the bright edges of the windows provided a strong fusion cue. Each window also contained a red fixation mark in the center to further aid fusion and to minimize motion tracking. In addition, the windows also contained a texture, randomly alternating light and thin lines at 0 disparity, across the top and bottom. As Regan (1991) noted, perception of stereomotion is poor unless there is stationary context in the image.

Two types of stimuli were used as moving objects. In the first two experiments, the stimuli were bright bars with a width of 0.2° and a luminance of 42 cd/m^2. In the third experiment, I used "DOGS", difference of Gaussians, with a sigma of 12 pixels for the positive Gaussian and 18 for the negative Gaussian. The contrast (L_{max}-L_{min}/2xmean) was 18%. The use of DOGs had no special relevance for the studies reported here.

The experiments employed the 4 possible types of motion shown in Figure 3. Left and right panels depict images seen by left and right eyes. The top row shows the left object moving while the right object remains stationary. This produced the appearance of an object moving in a circular path in depth with both frontal and stereo motion. The second row shows the case where both objects moved in the same direction to create purely frontal plane motion. The next pair depicts objects moving in opposite direction, producing only stereomotion: the bar/DOG appeared to move straight toward and away from the viewer. The last condition was a monocular control. Note that in all conditions, the left eye saw exactly the same motion.

In conditions where disparity changed (2A and 2C), the maximum disparities achieved by the motion were +/- 0.24 degrees. Observers were formally tested at velocities of 0.5 and 1.0 °/sec. The lower velocity produced unidirectional (left or away) motion while at the higher velocity, objects to moved back and forth (left-right, away-toward). Informal tests 2 °/sec indicated similar results.

5.1.3 Procedures

The same paradigm was used in all experiments. The observer fused the stereo images and viewed a series of two-alternative, forced-choice trials. Each interval contained a sequence of 60 frames, each presented for 1/60 of a second. In one interval, the motion was updated at the 60 Hz rate. At the velocities used, the 60 Hz sampling rate created motion which was indistinguishable from continuous motion. The other interval contained the same class of motion sampled at a lower rate. After each trial, the observer pressed a switch to indicate which interval contained the lower sampling rate. In practice, the observers performed this task by looking for the interval with "jitter" in the motion.

5.2 Experiment I: Frontal motion with different disparities

The first experiment examined aliasing in frontal plane motion at different disparities, i. e., disparity between eyes, but no *change* in disparity. Before considering the effect of changing disparity (stereomotion), it is worthwhile to perform the control experiment of determining whether disparity *pe se* has any effect on aliasing in frontal plane motion. Virtually previous all studies of frontal plane motion used 0 disparity, and there are no relevant data on whether disparity will affect perceived smoothness of sampled motion.

Observers were tested with frontal plane motions (Figure 3B) with 0, 0.12 or 0.24 degrees of crossed disparity between eyes. In each block of trials disparity conditions were randomly crossed with sampling rate. A Third variable, velocity, was constant for each trial block.

Figure 4 shows the results for these conditions in addition to a monocular control. Not surprisingly aliasing was easier to detect for higher velocities and lower sampling rates. However, there is no systematic difference among the various conditions. Disparity has no affect on frontal motion aliasing.

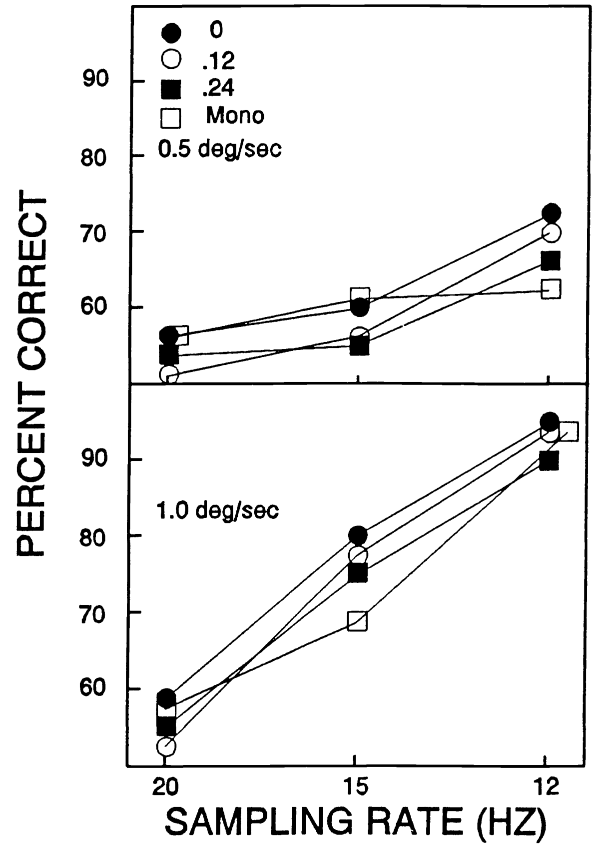


Figure 4

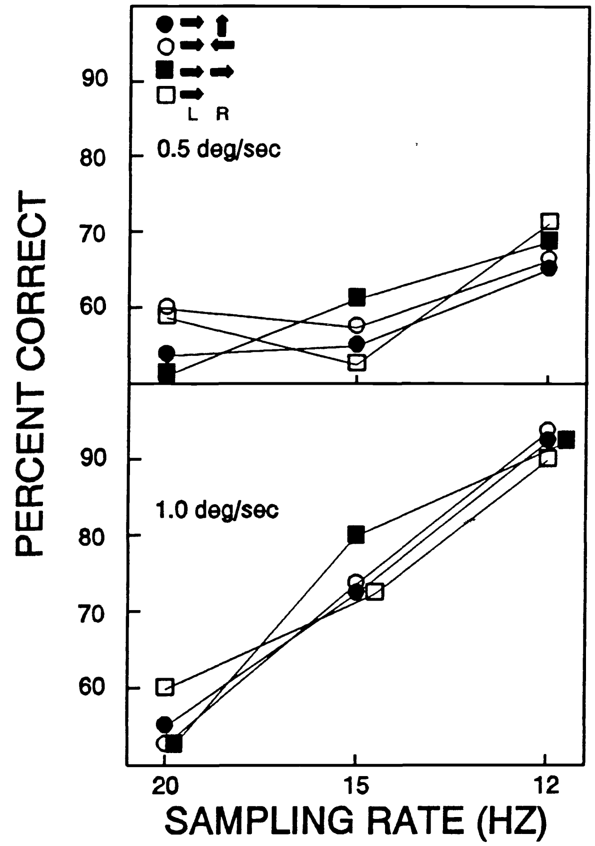


Figure 5

5.3 Experiment II: Motion in different depth planes

The second experiment directly compared the 4 types of motion depicted in Figure 3: pure (3B) frontal plane motion, pure stereomotion (3C), a combination of the two (3A) and a monocular control (3D). The four conditions were randomly crossed with 3 sampling rates within each block of trials.

Figure 5 shows the results. The data again show the unsurprising trend toward more aliasing at higher velocities and lower sampling rates. As in the first experiment, there is no evidence of different degrees of aliasing with different types of motion.

5.4 Experiment III: Monocular vs. binocular interpolation

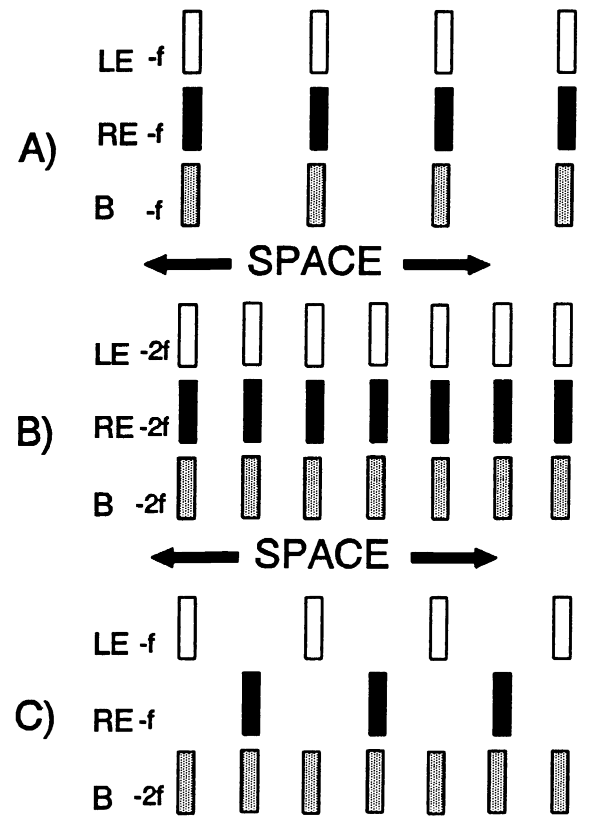


Figure 6

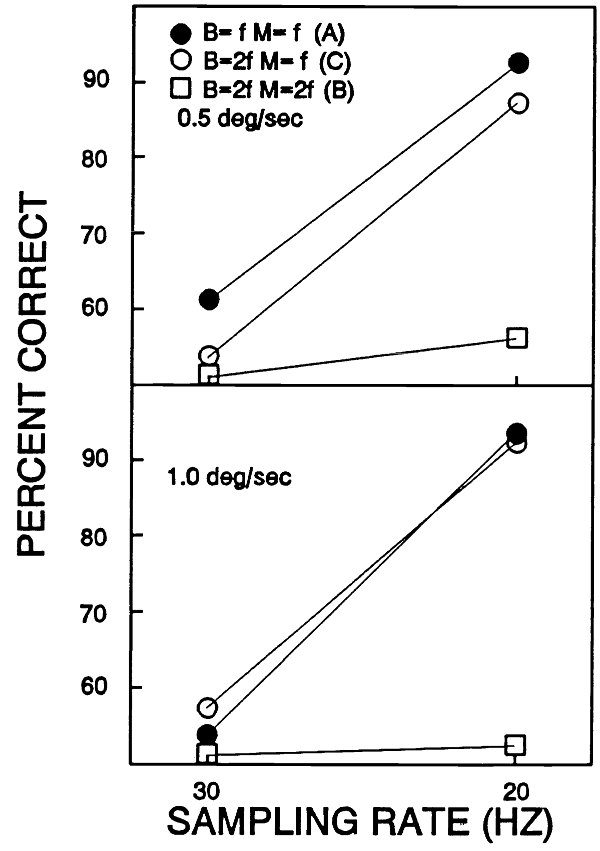


Figure 7

The previous experiments showed no difference between aliasing in frontal plane and stereo motion. Stereo displays are typically used to produce depth, but they also offer new ways to produce frontal plane motion. When sampling motion, for example, there are two ways that the updating can be allocated between the two eyes. This experiment examined whether different allocations of motion sampling might produce different degrees of aliasing. The underlying question is whether binocular mechanisms can integrate monocular samples.

Figure 6 schematically shows 3 different temporal sampling schemes. Each panel shows the left eye, right eye and binocular views of an object's motion as it steps across the display. In the A) panels, the two eyes are updated at the same sampling rate, f, and at the same time. The binocular view sees object sampled at rate f. If rate f produced jitter, then the sampling rate might be doubled to 2f, as shown in the B) panels. Here, again, the monocular and binocular samplings rates are identical. The lower panels shows the case where each eye is sampled at rate f, but the updating is "out of phase." This produces a binocular sampling rate of 2f. The viewing condition in C) is effectively what happens when stereo is created by LCD glasses which only allow alternate eye viewing.

Note that A) and C) have the same monocular sampling rate but different binocular rates. The B) and C) conditions have the same binocular rates but different monocular rates. This allows the following predictions: 1) if motion interpolation is strictly monocular, as suggested by many short-range motion theorists (e. g., Georgeson and Shackleton, 1989), then the A) and C) motions should produce similar results - relatively high jitter - compared to the B) condition. If aliasing is due to binocular sensors which look out of both eyes, then B) and C) conditions should produce similar results - less jitter than in the A) condition. If the second prediction were true, then the update rate for motion in each eye could be halved, permitting a great savings in image compression.

The third experiment tested these predictions by means of DOGs which were viewed in the 3 conditions shown in Figure 6. The results (Figure 7) again show that apparent aliasing is greater with lower sampling rates. More importantly, however, the data show that performance depended almost almost exclusively on the monocular sampling rate. At the lower velocity, binocular sampling rates made only a small difference, which was entirely absent at the higher velocity. There appears to be little or no binocular integration of samples.

6. DISCUSSION

The results of the three experiments paint a very clear picture. The first two studies show that the direction in depth makes no difference in apparent aliasing. This suggests that aliasing is due soley to frontal plane motion. The third experiment showed that there is no binocular interpolation, demonstrating that these signals are monocular in origin.

6.1 Practical Implications

The pragmatic consequences of these results are disappointing. When I started this research, I had hoped to find that stereoscopic displays might offer image compression by reducing the update rate needed to eliminate temporal aliasing. This would have been the case if 1) the aliasing signal were due to less sensitive stereomotion mechanisms or 2) the stereo signal suppressed frontal plane aliasing (Tyler, 1971). However, results of the first two studies failed to support either possibility. Reduced sampling, even in frontal plane motion, could also have been achieved if binocular mechanisms could interpolate samples presented alternately to the two eyes. Unfortunately, the third experiment failed to find significant binocular interpolation.

This result has major implications for display systems. One common system creates stereo by means of LCD glasses which show alternate frames to alternate eyes. Another common method uses polarizing or colored spectacles to create image separation, allowing both eyes to see all frames. Experiment 3 suggests that the temporal sampling rate for images viewed with LCD glasses will be double those using anaglyphs.

However, the viewing conditions used here were not identical to those produced by LCD glasses, which also create a luminance flicker. The third experiment did not assess these flicker effects, but several studies (Braddick, 1974; Green, 1984) have shown that luminance flicker can be highly disruptive to motion perception. This aspect of LCD glasses requires more study.

6.2 Theoretical Implications

The failure to find binocular interpolation of monocular samples supports for the belief that the short-range system is monocular (Braddick, 1974; Green and Blake, 1981; Georgeson and Shackleton, 1989). Apparently, the short-range system cannot interpolate the rapid signals from the two eyes.

However, it is clear that there are binocular motion mechanisms capable of integrating inputs from the two eyes. No one doubts that the long-range system can perform motion interpolation when images are presented to left and right eyes. The results of the third experiment therefore mean that either the long-range system was not fast enough to follow the motion or that apparent aliasing is produced solely by the short-range motion detectors. The data do not distinguish these alternative explanations for the failure to find reduced aliasing through binocular interpolation.

7. ACKNOWLEDGEMENTS

This research was funded by grant #46297 from the National Science and Engineering Research Council of Canada.

8. REFERENCES

Attneave, F. & Block, G. (1973) Apparent movement in tridimensional space. *Perception & Psychophysics*, **13**, 301-307.

Braddick, O. (1974) A short-range process in apparent movement. *Vision research*, **21, 519-527.**

Exner, S. (1888). Uber optische Bewegungsemfindungen. *Biologisches Centralblatt*, **8**, 437-448.

Georgeson, M. & Shackleton, T. (1989). Monocular motion sensing, binocular motion perception. *Vision Research*, **29**, 1511-1523.

Green, M. (1984). Masking by light and the sustained-transient dichotomy. *Perception & Psychophysics*, **35**, 519-535.

Green, M. (1992a). Perceptual basis of temporal aliasing and anti-aliasing. *Proceeding SPIE Conference on Human Vision, Visual Processing and Digital Displays III.* Forthcoming

Green, M. (1992b) Correspondence matching in apparent motion: Computation and psychophysics. Forthcoming.

Green, M. and Blake, R. (1981). Phase effects in monoptic and dichoptic temporal integration. *Vision research*, **21**, 365-372.

Green, M. & Odom, J. V. (1986). Correspondence matching in apparent motion: Evidence for three-dimensional spatial representation. *Science*, **233**, 1427-1429.

Korte, A. (1915) Kinematoskopische Untersuchungen. *Zeitschrift fur Psychologie*, **72**, 193-296.

Morgan , M. (1980). Spatiotemporal filtering and the interpolation effect in apparent motion. *Perception*, 161-174.

Nyman, G. and Laurinen, P. (1982). Reconstruction of spatial information in the human visual system. *Nature*, **297**, 324-325.

Nyman, G. and Laurinen, P. (1985). Visual undersampling in raster sampled images. *IEEE Transactions on Man, Systems and Cybernetics*, **SMC-15**, 655-659.

Ogle, K. and Weil, M. (1962) Spatial localization through binocular vision. In (H. Davson, ed.) *The Eye: Vol. 4, Visual Optics and the Optical space sense, 271-324*, 271-324. New York: Academic Press.

Regan, D. (1991). Depth from motion and motion-in-depth. In (J. R. Cronly-Dillon, series ed., D. Regan, volume ed.) *Vision and Visual Dysfunction: Binocular Vision*, **vol. 9**, 137-169. Toronto: MacMillan.

Shannon, C. (1949). Communication in the presence of noise. *Proc. IEEE IRE*, **37**, 10-21.

Tyler, C. W. (1971). Stereoscopic depth movement: Two eyes less sensitive than one. *Science*, **174**, 958-961.

Tzelgov, J., Henik, A., Dinstein, I and Rabany, J. (1990). Performance consequences of two types of stereo picture compression. *Human Factors*, **32**, 173-182.

White, K. and Odom, J. V. (1985). Temporal integration in global stereopsis. *Perception and Psychophysics*, **37**, 139-14.

Watson, A., Ahumada A. and Farrell, J. (1985). Window of visibility: a psychophysical theory of fidelity in time-sampled visual motion displays. *Journal of the Optical Society of America A*, **3**, 300-307.