

# Temporal properties of disparity processing revealed by dynamic random-dot stereograms

Elena Gheorghiu, Casper J Erkelens

Helmholtz Institute, Utrecht University, PO Box 80 000, NL 3508 TA Utrecht, The Netherlands;

e-mail: [e.gheorghiu@phys.uu.nl](mailto:e.gheorghiu@phys.uu.nl)

Received 21 November 2003, in revised form 11 January 2005; published online 23 September 2005

**Abstract.** In studies of the temporal flexibility of the stereoscopic system, it has been suggested that two different processes of binocular depth perception could be responsible for the flexibility: tolerance for interocular delays and temporal integration of correlation. None has investigated the relationship between tolerance for delays and temporal integration mechanisms and none has revealed which mechanism is responsible for depth perception in dynamic random-dot stereograms. We address these questions in the present study. Across five experiments, we investigated the temporal properties of stereopsis by varying interocular correlation as a function of time in controlled ways. We presented different types of dynamic random-dot stereograms, each consisting of two pairs of alternating random-dot patterns. Our experimental results demonstrate that (i) disparities from simultaneous monocular inputs dominate those from interocular delayed inputs; (ii) stereopsis is limited by temporal properties of monocular luminance mechanisms; and (iii) depth perception in dynamic random-dot stereograms results from cross-correlation-like operation on two simultaneous monocular inputs that represent the retinal images after having been subjected to a process of monocular temporal integration of luminance.

## 1 Introduction

The human visual system derives depth from small positional differences, or disparities, between the two retinal images. Depth perception derived from disparity is usually called stereopsis (Howard and Rogers 2002). In the literature, two binocular neural processes have been identified that must be active in order to allow stereopsis, which Julesz referred to as 'local' and 'global' stereopsis. The first (local) process is detection of disparities. Detection is through disparity-sensitive cells that are tuned to a specific disparity or range of disparities between corresponding stimuli presented at specific retinal locations (Barlow et al 1967; Poggio and Fisher 1977; DeAngelis et al 1991; Cumming and Parker 1999). The second (global) process is solving the correspondence problem. It is assumed that only the output of a subset of active binocular neurons will reach the level of perception. Binocular neurons, serving the same cyclopean direction but with different disparity tuning, will usually compete for perception. Competition is assumed to be present when the images are composed of random dots, since the dots of one image can be linked to many dots of the other image. A generally accepted view is that global correspondence is achieved by a process that maximises interocular correlation (Howard and Rogers 2002).

Time appears to be a critical factor in both local and global processes, although its role is not yet fully understood. Disparity detection usually occurs between two simultaneously presented monocular stimuli. However, disparity-induced depth can also be perceived when one retinal image is somewhat delayed relative to the other. This phenomenon is called tolerance for interocular delay. Interocular delays can be produced either electronically or with a neutral density filter. Several researchers have investigated the effects of interocular delays (Julesz and White 1969; Mitchell and O'Hagan 1972; Ross and Hogben 1974) and have reported that the stereoscopic system tolerates a time difference between binocular correlated random-dot images of up to 50 ms (Howard and Rogers 2002). Mitchell and O'Hagan (1972) reported that an

interocular delay destroys disparity discrimination for line targets, but the exact value of the interocular delay depends on the magnitude of the target disparity. Using dynamic random-dot stereograms (DRDSs), Julesz and White (1969) investigated the delay hypothesis and reported that a disparity-defined square of 48 min of arc side was not perceived with one frame delay ( $\sim 80$  ms). Cogan et al (1993), using repetitively flashed random-dot stereograms (RDSs) presented with an interocular delay, reported that correlated RDSs lost depth when interocular delays were longer than 45 ms, and anticorrelated RDSs showed robust depth when flashed with interocular delays between some 60 to 90 ms. Cogan et al (1993) explained depth perception from anticorrelated RDSs by the fact that, during the delay time, the image in one eye was paired with the negative afterimage of the contrast-reversed stimulus in the partner eye. More recently, Morgan and Fahle (2000) measured depth from interocular delay using a dynamic random-dot pattern in which the individual dots were modulated sinusoidally in luminance over time. They found that interocular phase lag, not delay, determined depth detection.

The fact that so many studies have reported depth in the presence of interocular delay shows that disparity detection occurs even when one of the monocular stimuli is delayed relative to the other. However, until now the psychophysical results have been explained mainly in terms of local processes, ignoring the fact that the corresponding global solving process may also have specific temporal properties. For instance, Morgan and Fahle (2000) explained their results by assuming that disparity detectors receive their input from pairs of monocular motion-direction-sensitive neurons. In contrast, not much is known about the temporal properties of the corresponding solving process. Neurophysiological data are not yet available. Two psychophysical studies have addressed the problem (Tyler 1991; Cormack et al 1994). These studies reported temporal integration of correlation, implying that a depth figure can be seen in a sequence of partially correlated DRDSs although it cannot be perceived in each individual random-dot pair. These studies were not conclusive as to the maximum integration time of correlation. Tyler (1991) reported a time of about 180 ms, whereas Cormack et al (1994) failed to find a single time of integration.

A problem in drawing conclusions from results of the various experimenters who investigated the above-mentioned temporal mechanisms is that the psychophysical studies were carried out with different types of stimuli (RDSs, DRDSs, line stereograms), and employed different tasks (correlation detection, depth detection, depth discrimination) and limited presentation times of the stimuli. The consequence of using limited presentation times of the stimuli may have been that the values of integration time reported in these studies may not reflect the period over which the visual system integrates the interocular correlation or tolerates delays but instead reflect the minimum time that is needed to detect any perceptual change in the stimulus. Julesz and Tyler (1976) and Tyler and Julesz (1976, 1978) reported that in correlation-detection tasks the subjects judge the appearance of the stimulus (ie the salience of cohesiveness of dots) rather than any impression of the depth plane as such. Correlation-detection tasks are not ideal tasks for the identification of disparity-defined depth. In order to investigate the effects of interocular delays on disparity-induced depth perception, it is desirable and prudent to employ depth-detection or depth-discrimination tasks.

It is clear that, before binocular interaction, visual information is processed monocularly. Thus, interocular delays and temporal integration revealed at a binocular level are probably influenced by temporal properties of monocular luminance mechanisms. Temporal characteristics of luminance processing are described by the impulse-response function in the time domain and by the temporal contrast-sensitivity function in the frequency domain (Burr and Morrone 1993; Shinomori and Werner 2003). By measuring

perceptual effects, numerous psychophysical studies (Ikeda 1965, 1986; Swanson et al 1987; Tyler 1992; Burr and Morrone 1993) have shown that responses to very short luminance impulses are biphasic, with a negative lobe following the initial positive response lasting about 50 ms (Burr and Morrone 1993). The impulse responses reflect the temporal properties of the visual system as a whole. However, electrophysiological studies (Kremers et al 1993; Chichilnisky and Kalmar 2002) have demonstrated that impulse responses of retinal ganglion cells have similar temporal characteristics. Therefore, it is likely that binocular processing is limited by temporal properties of luminance processing.

We have studied the temporal mechanisms of binocular depth perception, considering the properties of both the correspondence-solving process as well as of monocular luminance processing. We used several well-defined types of DRDSs that enabled us to dissociate various hypotheses. The DRDSs used in our experiments contained both spatial as well as (potential) temporal disparities. We employed depth-detection and depth-discrimination tasks. The results demonstrate that disparities from simultaneous monocular inputs dominate those from interocularly delayed inputs. The results also indicate that stereopsis in DRDSs is limited by monocular luminance integration. Thus, stereopsis in DRDSs seems to result from cross-correlation-like operation on two simultaneous, monocular inputs that represent the retinal images after having been subjected to a process of temporal integration of luminance.

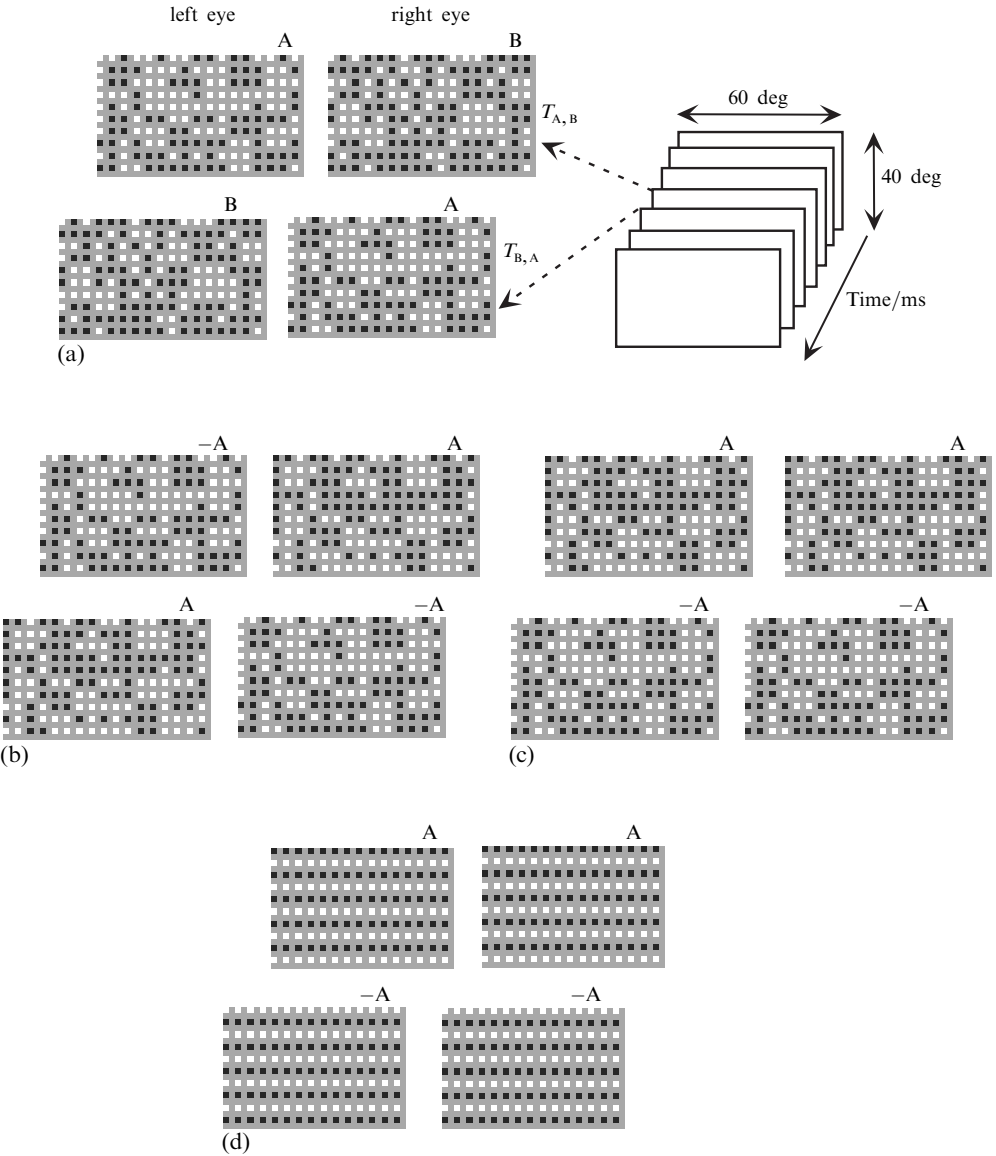
## 2 General methods

### 2.1 Observers

Three subjects (ages between 25 and 50 years) participated in all experiments. None of them showed any visual or oculomotor pathologies other than refraction anomalies. All had normal or corrected-to-normal visual acuity. They were checked for normal stereopsis by means of a recently developed stereoscopic depth-perception test (van Ee and Richards 2002).

### 2.2 Stimuli and procedure

The stimuli were large RDSs of 60 deg (width)  $\times$  40 deg (height). The stereograms consisted of a regular grid of randomly distributed bright and dark square dots on a grey background (see figure 1). The sizes (side lengths) of the dots and the width of the grey lines between the dots were 10 min of arc. The RDSs were generated in real-time at a frequency of 70 Hz by an HP 750 Graphics computer and back-projected onto a frontoparallel translucent screen by a D-ILA projector (JVC DLA-G11E). For this projection, we measured the dynamics of the screen luminance for fast alternations (frame duration 14 ms) between bright and dark dots. The luminance reached its maximum value within 14 ms and decayed in slightly less time than 14 ms (within about 20 ms). We observed that rapidly alternating bright and dark dots were perceived as static dots at an intermediate brightness level, except at edges between neighbouring dots of opposite brightness. Owing to the dynamics of the projection system, these edges were visible as thin lines. To avoid seeing these lines, which could carry disparity information, we isolated the dots from each other over a grey background. Stereoscopic vision was made possible by placing red and green filters in front of the observer's eyes so that each image was visible to one of the eyes only (light separation between red and green filters that were matched to the projector guns was better than 99%). The experiments were performed in an otherwise dark room. Before each trial, the computer generated two stereograms whose image pairs were fully correlated. Both stereograms depicted a square, at the same location, that stood out in depth during binocular viewing. The squares were located on either the left or the right side of the stereograms. The size of the squares was 8 deg  $\times$  8 deg and their disparity 40 min



**Figure 1.** Stimuli used in the four experiments. The stereograms consisted of a grid of grey lines filled with randomly distributed bright and dark dots. The stimuli were DRDSs consisting of the continuous alternation of two image pairs. The images were combined in three different ways in four separate experiments. (a) Experiment 1: stimulus (A, B), (B, A). The patterns A and B were uncorrelated. (b) Experiment 2: stimulus (−A, A), (A, −A). The patterns A and −A were anticorrelated. (c) Experiment 3: stimulus (A, A), (−A, −A). (d) Experiment 4: stimulus (A, A), (−A, −A) in the texture-detection task. The images consisted of alternating rows of bright and dark dots on a grey background. In individual trials, the rows of bright and dark dots were oriented either horizontally or vertically. The stimulus consisted of the repeated presentation of two correlated pairs both containing the vertical or horizontal rows of bright and dark dots.

of arc (crossed disparity). The image seen by each individual eye did not provide any information about the position of the square. In each of the three experiments, the DRDSs consisted of the continuous alternation of two image pairs. We varied the presentation times of the two image pairs. The presentation times were adjustable in steps of 14 ms, between 14 ms and 168 ms (twelve values). Decreasing the step interval (14 ms)

for the presentation times was not possible because the interval was limited by the frame rate of the projector. The computer generated new patterns before each trial. It is important to mention that for each image pair the correlation level was the same everywhere across the image, irrespective of whether an area contained the depth figure or not.

In experiments 1, 2, 3, and 6, we used a depth-detection task and a forced-choice paradigm in which the subjects indicated (by pressing a key) whether the disparity-defined square was presented on the left or the right of the DRDS. The task was easy for all subjects. The subjects were free to make eye movements. Each combination of presentation times was repeated twenty times in ten separate sessions (288 trials in each session). The combinations were presented in a random order within each individual session. We used a level of 75% of correct answers as the criterion for reliable depth detection.

A control experiment (experiment 5) was carried out in order to test whether a depth-discrimination task would produce different results than a depth-detection task for our stimuli. Thus, in experiment 5, we used the same stimuli as those employed in experiments 1, 2, and 3 but in a different task. To demonstrate that the subjects' judgments were based on stereoscopic information, we asked the subjects to judge whether the test square appeared in front (crossed disparity) or behind (uncrossed disparity) the background. The disparity of the square was  $\pm 40$  min of arc.

In all experiments, the DRDSs were presented on the screen until the subject gave his/her answer, by pressing a key.

### 3 Experiment 1

The goal of this experiment was twofold: (i) to reproduce the experiment of Julesz and White (1969) with a slightly different stimulus; (ii) to extend their results to irregular presentation times of the images. In our experiment, we used continuous alternation of only two images instead of a sequence of different images as in the experiment of Julesz and White (1969). We preferred these stimuli because they allowed us to measure luminance effects over time in a controlled way.

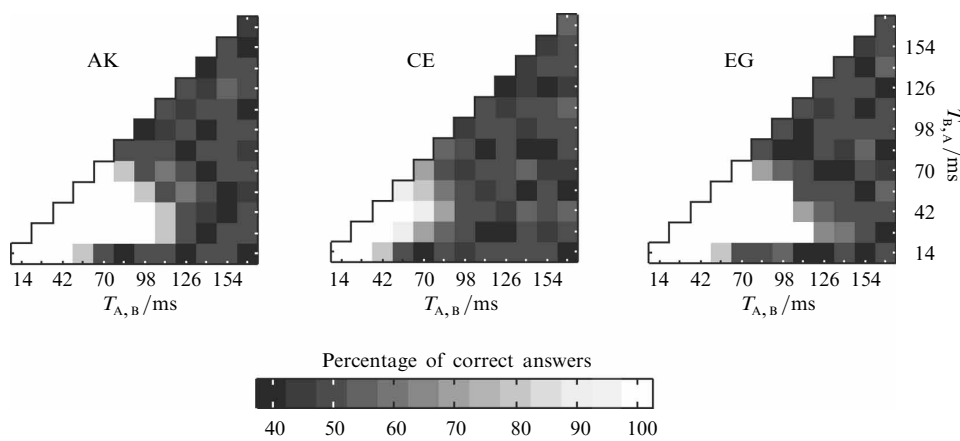
#### 3.1 Method

The computer generated two stereograms whose left and right image pairs (L, R), which we denote by (A, A) and (B, B), were highly correlated. The patterns A and B were uncorrelated. The dynamic stimulus consisted of the continuous alternation of two image pairs: image pair (A, B) and image pair (B, A), respectively (figure 1a). The two alternated images were uncorrelated in two ways: spatially as well as temporally, ie the left and right images were uncorrelated at all times and also the image to each eye consisted of the alternation of two uncorrelated images. The stimulus (A, B), (B, A) constituted a DRDS in which correlation was zero within each of the periods  $T_{A,B}$  and  $T_{B,A}$  and, thus, throughout the presentation time. In this stimulus, the sequence of patterns was identical in the left and the right eye, ie A, B, A, B, ..., but the patterns were out of phase. However, if the visual system would tolerate the delay of the image to one eye by one time period,  $T$ , then this stimulus would become binocularly indistinguishable from the stimulus (A, A), (B, B), resulting in high correlation at all times.

One may think that red-green anaglyphs are not the best choice for studying temporal factors because of differences in processing time for red and green colours. However, we averaged the depth-detection results of combinations ( $T_{A,B}$ ,  $T_{B,A}$ ) with those of ( $T_{B,A}$ ,  $T_{A,B}$ ) and, therefore, the possible differences due to processing time for red and green were cancelled out. This holds for all results in response to different image combinations presented in this paper.

3.2 Results

Depth-detection scores (percentage of correct answers) are presented as a function of the time periods  $T_{A,B}$  and  $T_{B,A}$  for three subjects (figure 2). The graphs have a triangular shape because we averaged the results of combinations  $(T_{A,B}, T_{B,A})$  with those of  $(T_{B,A}, T_{A,B})$ . The only difference between the stimuli associated with these combinations was the image pair, either (A, B) or (B, A), that appeared at the onset of stimulus presentation. Since depth judgments were made after seconds of viewing (typically, 3 to 7 s) we considered both stimuli as equivalent. Statistical analysis did not show any difference between their results. The bright areas in figure 2 indicate combinations of  $T_{A,B}$  and  $T_{B,A}$  for which the subjects detected depth. No depth was detected in the darker areas. The results indicate that the subjects were able to perceive depth for most combinations of  $T_{A,B}$  and  $T_{B,A}$  ranging between 14 ms and about 70 ms. For periods of  $T_{A,B} = T_{B,A}$  longer than about 70 ms, none of the subjects was able to detect depth. Qualitatively, the results were similar in the three subjects.



**Figure 2.** Performances in the depth-detection task for the stimulus (A, B), (B, A). Depth-detection scores (percentage of correct answers) are presented as a function of the presentation time of the two images  $T_{A,B}$  and  $T_{B,A}$ , for presentation times between 14 ms and 168 ms. The white areas indicate combinations of  $T_{A,B}$  and  $T_{B,A}$  that correspond to depth detection. No depth was detected in the darker areas.

The subjects reported that they perceived the DRDSs as static when  $T_{A,B} + T_{B,A} < 42$  ms, flickering when  $42 \text{ ms} < T_{A,B} + T_{B,A} < 168$  ms, and alternating when  $T_{A,B} + T_{B,A} > 168$  ms. When the DRDS was perceived as static, the subjects reported that the texture (dots) of the DRDS contained three brightness levels, whereas each individual pattern (ie image A or B) consisted of only two luminance levels (bright and dark dots). The perceived texture contained bright dots, dark dots, and dots of intermediate brightness (grey).

3.3 Discussion

The present results indicated that delays of about 70 ms between the correlated images are the upper limit for depth detection. Julesz and White (1969) reported that delays of 80 ms were not tolerated. Our results show that by using stimuli that consisted of the continuous alternation of only 2 frames, the tolerated delays have similar values to those reported by authors who used stimuli in which each image pair was different on every frame (Julesz and White 1969; Ross and Hogben 1974).

We can think of three non-exclusive explanations for the fact that depth was perceived in our DRDSs. The first explanation is that the stereoscopic system tolerates time differences between binocularly correlated random-dot images of up to 70 ms.

The second possibility is that the temporal window of the correspondence-solving process is about 140 ms. The third explanation is that seeing depth in fast DRDSs presented with interocular delay is a consequence of the temporal properties of monocular luminance mechanisms. Our stimulus (A, B), (B, A) does not allow us to dissociate between these explanations. The objective of experiment 2 was to use a stimulus that would enable us to reject at least one of the three explanations.

If one uses DRDSs in which the images are different on every frame, then one can study interocular delay but not monocular luminance mechanisms. By using our type of DRDS, we could investigate all three mechanisms. A repeated alternation of only two frames allowed us to measure luminance effects over time in a controlled way.

## 4 Experiment 2

In this experiment, we used a DRDS in which the two alternated images were anticorrelated in two ways: spatially as well as temporally—ie the left and right images were anticorrelated at all times and also the image to each eye consisted of the alternation of two anticorrelated images. In other words, the stimulus consisted of the combination  $(-A, A)$ ,  $(A, -A)$ . If the visual system is able to delay the image to one eye by a certain time period,  $T$ , then this stimulus will be equivalent to the fully correlated stimulus  $(A, A)$ ,  $(-A, -A)$ . As a result, we expected to find similar results as in experiment 1 if temporal tolerance was caused by the interocular delay mechanisms. We also expected similar results if the correspondence-solving process would have a time window of two frames. On the other hand, we anticipated different results if monocular luminance integration was the predominant determinant.

### 4.1 Method

The DRDS consisted of the continuous alternation of two image pairs: image pair  $(-A, A)$  and image pair  $(A, -A)$  respectively (figure 1b). The stimulus  $(-A, A)$ ,  $(A, -A)$  constituted a DRDS in which correlation was  $-1$  at all times. In monocular viewing, successive frames were anticorrelated. The sequence of patterns in the left and right eye was identical over time, ie  $A, -A, A, -A, \dots$ , but the sequences were out of phase. However, if the visual system could delay the image to one eye by one time period  $T_{-A,A}$  or  $T_{A,-A}$ , the stimulus would become indistinguishable from the stimulus  $(A, A)$ ,  $(-A, -A)$ , resulting in a highly correlated stimulus at all times.

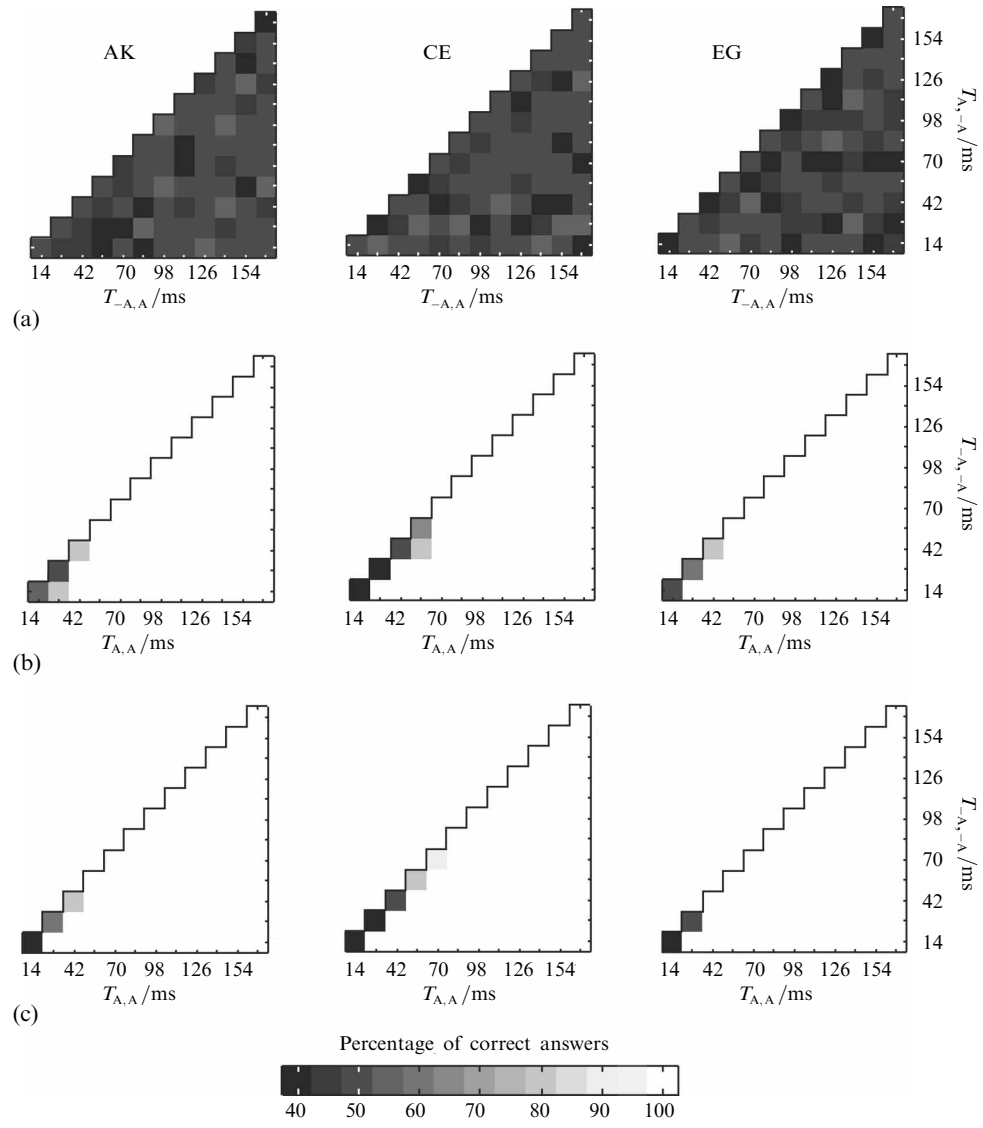
The stimuli  $(-A, A)$ ,  $(A, -A)$ , and  $(A, B)$ ,  $(B, A)$  (experiment 1) were similar to each other in the following respects: (i) the alternation of patterns in the two eyes was out of phase; (ii) in each eye, the two alternated patterns were rivalrous, ie  $-A, A, -A, A, \dots$  and  $A, B, A, B, \dots$ , respectively; (iii) at each instant in time, the images were rivalrous, ie  $(-A, A)$  and  $(A, B)$ , respectively; (iv) if the stereoscopic system could delay the image to one eye by a certain time period, both stimuli would result in highly correlated stimuli.

### 4.2 Results

Depth-detection scores (percentage correct answers) are presented as a function of the time periods  $T_{-A,A}$  and  $T_{A,-A}$  for three subjects (figure 3a). All subjects performed at chance level (50%). None of the subjects was able to perceive depth in this stimulus at any combination of  $T_{-A,A}$  and  $T_{A,-A}$  between 14 ms and 168 ms.

### 4.3 Discussion

The very different results in response to stimuli  $(-A, A)$ ,  $(A, -A)$  and  $(A, B)$ ,  $(B, A)$  (see experiment 1) indicate that the human stereoscopic system cannot detect disparity or cannot solve the correspondence problem between interocular delayed stimuli. In other words, we conclude that, in DRDSs, interocular correlation is computed between simultaneously presented patterns.



**Figure 3.** Performances in the depth-detection task for: (a) stimulus  $(-A, A), (A, -A)$ ; (b) stimulus  $(A, A), (-A, -A)$ . Percentages of correct answers are presented as a function of the presentation time of the two images, for presentation times between 14 ms and 168 ms. The white areas indicate the combinations of  $T_{A,A}$  and  $T_{-A,-A}$  corresponding to depth detection. No depth was detected in the darker areas. (c) Performances in the texture-detection task for stimulus  $(A, A), (-A, -A)$ . Texture-detection scores are presented as a function of the presentation times of the two images  $T_{A,A}$  and  $T_{-A,-A}$ , for presentation times between 14 ms and 168 ms. The white areas indicate the  $T_{A,A}$  and  $T_{-A,-A}$  combinations corresponding to texture detection. Darker areas indicate no texture detection.

The results obtained in response to the dynamic stimulus  $(-A, A), (A, -A)$  are the same as those obtained in response to any static anticorrelated stereogram. It is known that uncorrelated or anticorrelated images produce no sensation of depth (Julesz 1971; Cumming et al 1998). However, many neurons manifest disparity selectivity when stimulated with anticorrelated RDSs (Cumming and Parker 1997). Also, anti-correlated images induce transient vergence eye movements in the opposite direction than correlated images (Masson et al 1997).



One could argue that the impossibility to see any depth in stimulus  $(-A, A)$ ,  $(A, -A)$  for any combination of presentation time  $T_{-A,A}$  and  $T_{A,-A}$ , was not caused by the absence of delay mechanisms but by the fact that, in monocular viewing, the continuous alternations of anticorrelated images might have a detrimental effect on stereopsis. This alternative explanation was tested in experiment 3.

## 5 Experiment 3

The objective of this experiment was to test whether or not stereopsis was negatively influenced by fast monocular alternations of opposite luminance polarity, between the anticorrelated patterns  $A$  and  $-A$ .

### 5.1 Method

We used DRDSs that were similar to those of experiment 2, from a monocular point of view. The difference was that the left and right image pairs were highly correlated at each moment in time instead of being anticorrelated. The stimulus  $(A, A)$ ,  $(-A, -A)$  consisted of the repeated alternation of image pair  $(A, A)$  and image pair  $(-A, -A)$  (figure 1c). The sequence of patterns was identical in the left and right eye and in phase. If monocular repeated presentations of anticorrelated images prevent stereopsis, then we expect similar results to those of experiment 2.

### 5.2 Results

Depth-detection scores (percentage of correct answers) are presented as a function of the time periods  $T_{A,A}$  and  $T_{-A,-A}$  for three subjects (figure 3b). The white areas indicate combinations of  $T_{A,A}$  and  $T_{-A,-A}$  for which the subjects detected depth. No depth was detected in the darker areas. These results indicate that the subjects were able to perceive depth at the level of 100% for all combinations of  $T_{A,A}$  and  $T_{-A,-A}$  between 14 ms and 168 ms, except for a few combinations of short and approximately equal alternation times. Subjects AK and EG were unable to detect depth for the alternation periods of  $T_{A,A} = T_{-A,-A} = 14$  ms, 28 ms, and 42 ms. Subject CE could not detect depth for alternation periods of 56 ms in addition.

It is also important to note that perceived depth was sustained although, in monocular viewing, luminance polarity fluctuated fast as a function of time. For  $T_{A,A} = T_{-A,-A} = 14$  ms and 28 ms, the subjects reported that they perceived a fully grey (intermediate brightness) image without any dots or other textural contours.

### 5.3 Discussion

To our knowledge, our experiment is the first that has studied the effects of opposite luminance polarity on depth detection in DRDSs. The results showed that depth is clearly perceived for presentation times of each image longer than about 40 ms, despite the fact that, in monocular viewing, luminance polarity alternated as a function of time (ie successive monocular images were anticorrelated).

Additional evidence against tolerance to interocular delays in human stereoscopic vision is provided by stimuli  $(-A, A)$ ,  $(A, -A)$  and  $(A, A)$ ,  $(-A, -A)$ . Stimulus  $(-A, A)$ ,  $(A, -A)$  would be indistinguishable from  $(A, A)$ ,  $(-A, -A)$  in binocular viewing if the visual system would tolerate delays between the images to the left and the right eyes. If this were true, the results for stimulus  $(-A, A)$ ,  $(A, -A)$  should resemble those of stimulus  $(A, A)$ ,  $(-A, -A)$ . The fact that in stimulus  $(A, A)$ ,  $(-A, -A)$  the subjects were able to see the depth figure for very short but irregular alternations (for instance,  $T_{A,A} = 14$  ms and  $T_{-A,-A} = 28$  ms or 42 ms, ie white areas on the lateral sides of the darker areas in figure 3b), but none of the subjects perceived depth in stimulus  $(-A, A)$ ,  $(A, -A)$  for the same alternation times (figure 3a), is an indication that even 14 ms of interocular delay between correlated images is not tolerated by the stereoscopic system.

One can conclude that monocular alternations in luminance polarity did not cause the rivalry of stimulus  $(-A, A)$ ,  $(A, -A)$  (experiment 2) for alternation periods longer than about 40 ms. However, the impossibility of seeing disparity-defined shape for very short alternation periods of 14 ms, 28 ms, and 42 ms in both stimuli  $(-A, A)$ ,  $(A, -A)$  and  $(A, A)$ ,  $(-A, -A)$  may be due to the temporal properties of monocular luminance mechanisms. Temporal properties of luminance mechanisms are the likely cause of the depth-detection limits if texture and depth-detection limits coincide in individual subjects.

## 6 Experiment 4

In this experiment, we measured the temporal limits of texture (contrast) formation in our DRDSs.

### 6.1 Method

We used the stimulus configuration  $(A, A)$ ,  $(-A, -A)$  of experiment 2, but this time in a texture-detection task. The texture-detection task required a stimulus in which the textural pattern could easily be recognised by the subjects. Thus, the random-dot patterns used in the previous experiments were not suited to this purpose. We used images that consisted of alternating rows of bright and dark dots on a grey background (figure 1d). Both the dot size (side length) and the width of the grey background in between the rows were 10 min of arc. In individual trials, the rows of bright and dark dots were oriented either horizontally or vertically. Before each trial, the computer generated two image pairs  $(A, A)$  and  $(-A, -A)$ , both containing the vertical or horizontal rows of bright and dark dots. A forced-choice paradigm was used in which the subject's task was to indicate (by pressing a key) whether the rows of bright and dark dots were vertical or horizontal. The task was easy for all subjects. We recorded for which alternations the subjects were able to detect the horizontal or vertical texture of DRDSs. Each combination was presented ten times in ten separate sessions. The combinations were presented in a random order in individual sessions.

### 6.2 Results

Texture-detection scores (percentage of correct answers) are presented as a function of the time periods  $T_{A,A}$  and  $T_{-A,-A}$  for three subjects (figure 3c). The white areas indicate combinations of  $T_{A,A}$  and  $T_{-A,-A}$  for which the subjects detected texture at the level of 100%. No texture was detected in the darker areas.

These results show that the subjects were able to detect horizontal and vertical dotted lines at the level of 100% for all combinations of presentation times of the two images  $T_{A,A}$  and  $T_{-A,-A}$  between 14 ms and 168 ms, except for a few combinations of very short and equal alternation times of 14 and 28 ms. The subjects AK and CE did not detect the lines for alternation periods of 42 ms, in addition. Also, for  $T_{A,A} = T_{-A,-A} = 56$  ms, the texture-detection score of subject CE was not at the 100% level.

### 6.3 Discussion

The significant finding of this experiment is the fact that the areas of no texture detection (darker areas in figure 3c) are very similar to the areas of no depth detection (darker areas in figure 3b) within each subject. This similarity indicates that stereopsis in DRDSs was not possible in the absence of perceived texture or contours, in both the left and right eyes, for short and equal presentation times of the two images  $(A, A)$  and  $(-A, -A)$ . The fact that the depth figure, the square, was visible for very short but unequal alternation times (white areas on the lateral side of the darker areas in figure 3b) is a clear indication that stereopsis in DRDSs requires perceived texture. In short, the present results indicated that the temporal processing of disparity was limited by temporal properties of luminance mechanisms.

## 7 Experiment 5

The goal of this experiment was to test whether our stimuli would produce similar results in a depth-discrimination task.

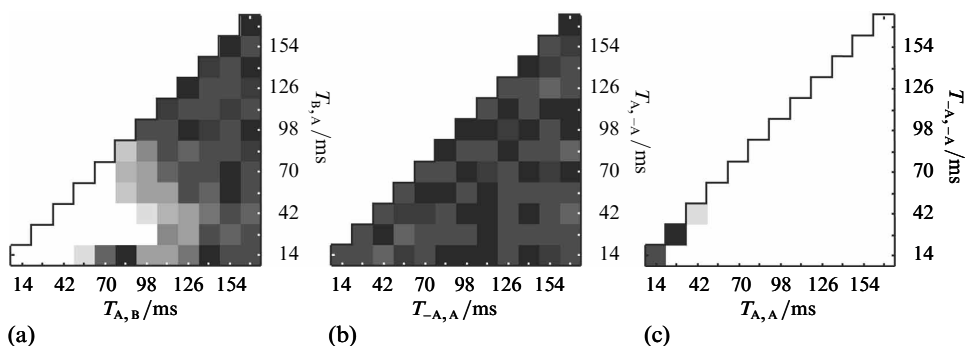
### 7.1 Method

We used the same stimulus configurations as those employed in experiments 1, 2, and 3, namely (i) stimuli (A, B), (B, A); (ii) stimuli (−A, A), (A, −A); and (iii) stimuli (A, A), (−A, −A), but in a depth-discrimination task. Only subject EG participated in this experiment. The task of the subject was to judge whether the test square appeared in front of (crossed disparity) or behind (uncrossed disparity) the background. The disparity of the square was  $\pm 40$  min of arc. Each combination of presentation times of the two images was repeated twenty times in ten separate sessions (288 trials in each session). The combinations were presented in a random order within each individual session.

### 7.2 Results and discussion

Depth-discrimination results (percentage of correct answers) are presented as a function of the time periods,  $T$ , of the two images, for the three stimuli (figure 4). In stimulus (A, B), (B, A), the subject performed at the level of 100% for periods of  $T_{A,B} = T_{B,A}$  shorter than 70 ms (figure 4a). For  $T_{A,B} = T_{B,A}$  longer than about 84 ms, the subject performed at chance level. In stimulus (−A, A), (A, −A), the subject was not able to discriminate between the two disparity-defined shapes and performed at chance level because no disparity-defined shape was perceived regardless of the disparity sign (figure 4b). In stimulus (A, A), (−A, −A), the subject performed at 100% level for all combinations of  $T_{A,A}$  and  $T_{-A,-A}$  between 14 ms and 168 ms, except for  $T_{A,A} = T_{-A,-A} = 14$  and 28 ms, when her performance was at chance. For  $T_{A,A} = T_{-A,-A} = 42$  ms the performance was at the 85% level.

The results show similar trends as those in figures 2, 3a, and 3b, respectively. Both depth-detection and depth-discrimination tasks produced similar results. Therefore, we are confident that the subjects' judgments were based on perceived depth.



**Figure 4.** Performances in the depth-discrimination task for: (a) stimulus (A, B), (B, A); (b) stimulus (−A, A), (A, −A); (c) stimulus (A, A), (−A, −A). Percentages of correct answers (see scale in figure 3) are presented as a function of the presentation time of the two images, for presentation times between 14 ms and 168 ms.

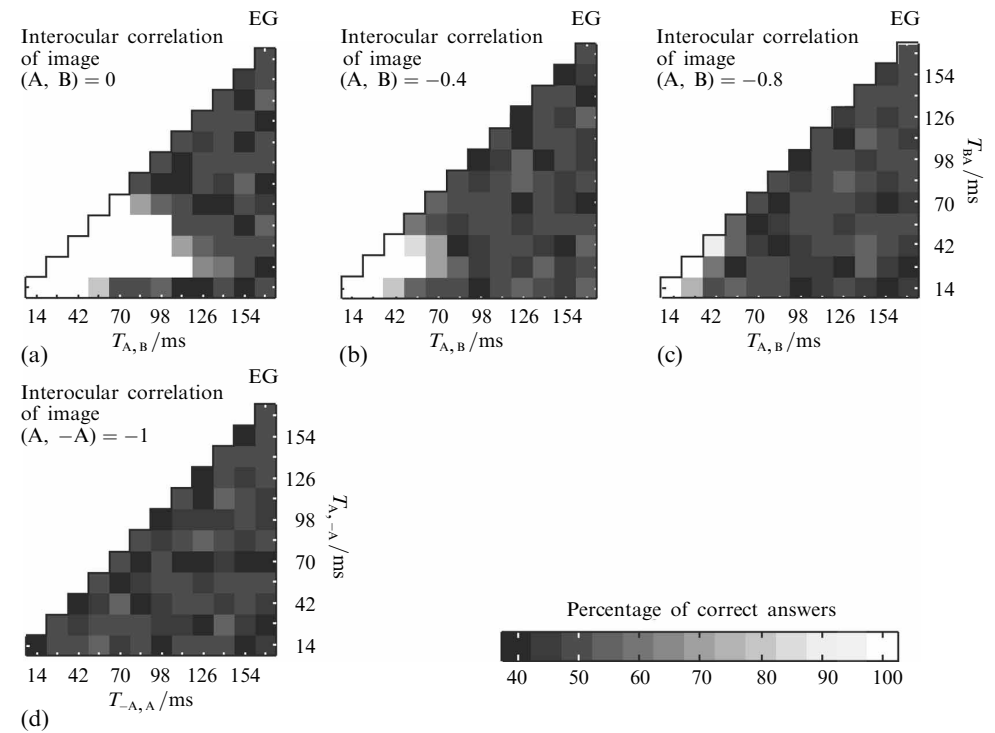
## 8 Experiment 6

The results of experiment 2 and the conclusion that interocular correlation is computed between simultaneously presented images seems to be in conflict with the results of experiment 1 and the tolerance for interocular delay conclusion drawn by Julesz and White (1969). We tested whether the conflict is resolved by considering tolerance for noise in the correspondence-solving process. It is well known that depth is perceived in DRDSs in which interocular correlation is not necessarily unity but is above a certain

threshold. Cormack et al (1994) measured thresholds for detecting depth using images that had a correlation between 0.1 and 0.5. Thus, the fact that depth was perceived in stimulus (A, B), (B, A) for  $T_{A,B} = T_{B,A}$  between 42 ms and 70 ms and also for irregular  $T_{A,B}$  and  $T_{B,A}$  might be explained by the tolerance for noise at the binocular level. To test the tolerance-for-noise hypothesis was the goal of this experiment. To do this, we manipulated the correlation threshold in stimulus (A, B), (B, A). We predicted that decreasing the interocular correlation level between the images A and B would result in the reduction of combinations of  $T_{A,B}$  and  $T_{B,A}$  for which depth was perceived (ie a decrease of the white area in figure 2).

8.1 Method

We used DRDSs similar to those of experiment 1, ie the stimulus (A, B), (B, A). The difference was that the interocular correlation level between the patterns A and B was different. In two separate sessions, we used two different levels of interocular correlation, namely  $-0.4$  (ie 70% of the dots in images A and B have opposite luminance polarity), and  $-0.8$  (ie 90% of the dots in images A and B have opposite luminance polarity). Thus, the stimulus (A, B), (B, A) constituted a DRDS in which correlation was lower than zero ( $-0.4$  and  $-0.8$ , respectively) within each of the periods  $T_{A,B}$  and  $T_{B,A}$  and, thus, throughout the presentation time. In monocular viewing, the continuous alternation of patterns A and B determined the reciprocal masking of 70% and 90% of dots, respectively, in the two images. Only subject EG participated in this experiment. The task and procedure were the same as in experiment 1.



**Figure 5.** Performances in the depth-detection task for the stimulus (A, B), (B, A) when the interocular correlation between the patterns (A) and (B) was (a) zero (ie uncorrelated patterns); (b)  $-0.4$ ; (c)  $-0.8$ ; (d)  $-1$  (ie anticorrelated). Depth-detection scores (percentage of correct answers) are presented as a function of the presentation time of the two images  $T_{A,B}$  and  $T_{B,A}$ , for presentation times between 14 ms and 168 ms. The white areas indicate combinations of  $T_{A,B}$  and  $T_{B,A}$  that correspond to depth detection. No depth was detected in the darker areas.

## 8.2 Results and discussion

Depth-detection scores are presented in figures 5b and 5c as a function of the time periods  $T_{A,B}$  and  $T_{B,A}$  for interocular correlation levels of  $-0.4$  and  $-0.8$ , respectively. Figures 5a and 5d are the same as in experiments 1 and 2, respectively. The bright areas indicate combinations of  $T_{A,B}$  and  $T_{B,A}$  for which the subject detected depth. No depth was detected in the darker areas. The results indicate that combinations of  $T_{A,B}$  and  $T_{B,A}$  for which the subject was able to perceive depth decreased progressively with decreasing interocular correlation level between the images A and B from an upper limit of about 70 ms when interocular correlation was 0 (figure 5a), to about 42 ms when interocular correlation was  $-0.4$  and  $-0.8$  (figures 5b and 5c, respectively), and to no depth perception when the images were anticorrelated (figure 5d). Furthermore, the effect of decreasing correlation level was more prominent for irregular  $T_{A,B}$  and  $T_{B,A}$ . These results are in agreement with predictions of tolerance-for-noise hypotheses.

## 9 General discussion

In this study, experiments were performed to assess the temporal properties and limitations of disparity processing in DRDSs. We measured depth detection in response to three different combinations of DRDSs: (i) (A, B), (B, A); (ii) ( $-A$ , A), (A,  $-A$ ); (iii) (A, A), ( $-A$ ,  $-A$ ). The principal results of these experiments can be summarised as follows. Experiments 1 and 2 indicated that disparities from simultaneous monocular inputs dominate those from interocularly delayed inputs. Experiment 3 [stimulus (A, A), ( $-A$ ,  $-A$ )] showed that alternations in luminance polarity slower than 42 ms per frame could not be the reason for rivalry in stimulus ( $-A$ , A), (A,  $-A$ ). Experiments 3 and 4 showed that for fast and equal alternations of luminance polarity between 14 ms and 42 ms, depth was not perceived in stimulus (A, A), ( $-A$ ,  $-A$ ) owing to temporal properties of monocular luminance mechanisms. These results indicate that stereopsis in DRDSs is limited by monocular luminance integration. Experiment 6 showed that perceived depth in stimulus (A, B), (B, A) for frame durations longer than about 42 ms, and also for unequal presentation times of the two images, can be explained by a tolerance for noise in the stereoscopic system.

It might seem surprising that, in the study of Cogan et al (1993), subjects perceived depth in anticorrelated stereograms presented with interocular delay of 60 to 90 ms. However, in their stimuli, the image in each eye was followed by a blank screen (uniform grey field, G), ie the stimulus configuration was (A, G), (G,  $-A$ ). Cogan et al (1993) explained perceived depth in anticorrelated stereograms presented with interocular delay of some 60 ms by the fact that during the delay time, the image in one eye was paired with the negative afterimage of the contrast-reversed stimulus in the partner eye. The impulse-response function shows that such an afterimage is present between about 50 and 100 ms (Burr and Morrone 1993). Thus, perceptually, Cogan et al's stimulus (A, G), (G,  $-A$ ) became somewhat similar to our stimulus (A, A), ( $-A$ ,  $-A$ ), owing to the monocular luminance mechanisms. For presentation times between 60 to 90 ms, depth was easily perceived in our stimuli (A, A), ( $-A$ ,  $-A$ ). On the basis of these findings, one should take into consideration the effects of luminance mechanisms if one studies the stereoscopic system with stimuli that contain black frames between the images or presentation durations of the stimuli shorter than about 40 ms.

The present experiments allow us to make a few remarks concerning both transient and sustained stereopsis that go beyond the purpose of this paper. To clarify: the terms sustained and transient could refer to the types of stimuli, to the underlying mechanisms, or to the percepts. Recently, Schor and collaborators (Schor et al 1998; Edwards et al 1999; Pope et al 1999) have presented evidence for separate sustained and transient stereo systems that differ in their temporal and spatial sensitivities.

These authors reported that short presentations (200 ms) of dichoptic pairs of opposite-contrast Gaussians induced depth perception, whereas long-lasting presentations (4 s) did not. In their view, the sustained system requires longer stimulus duration. A number of studies have shown that anticorrelated images produce no sensation of depth (Julesz 1971; Cogan et al 1993; Cumming et al 1998). Pope et al (1999) concluded that spatially complex stimuli, like dense anticorrelated RDSs, appear to be processed exclusively by the sustained system, which cannot match opposite-contrast stimuli (Howard and Rogers 2002). Our stimuli were sustained with respect to the presentation times. The resulting depth percepts were also sustained. The tight relationship between these aspects and the findings of Schor and collaborators, allows us to conclude that we investigated the sustained stereo system.

To conclude:

- (i) Disparities from simultaneous monocular inputs dominate those from interocular delayed inputs.
- (ii) Stereopsis is limited by monocular luminance integration mechanisms. In other words, stereopsis in fast DRDSs requires perceived textural pattern.
- (iii) Depth perception in DRDSs results from a sequence of two processes: monocular temporal integration of luminance followed by a cross-correlation-like operation between simultaneous monocular inputs that have each been subjected to the temporal integration of luminance. If the correlation level is above a certain threshold, stereopsis occurs. Otherwise, the images are perceived as rivalrous.

### 9.1 Neurophysiological studies

Physiological studies have shown that disparity-sensitive neurons in V1 and V2 respond to the disparities between the two retinal images. These cells with positionally disparate receptive fields fire when the stimulus disparity matches the receptive-field disparity. Until recently, it was not clear from neurophysiological studies, how disparity-tuned neurons would respond when the stimulus falls on the left and right receptive fields with an interocular delay. In other words, what are the consequences of interocular delays at neuronal level? Very recently, Cumming and Read (2003) recorded responses of disparity-selective V1 neurons using DRDSs in which each image was refreshed for every frame, and showed that an interocular delay of only 1 video frame (14 ms) dramatically reduced the magnitude of the disparity-selective response. Moreover, Cumming and Read suggested that the psychophysical consequences of an interocular delay were even more dramatic. Our results support this suggestion. By using different DRDSs consisting of continuous alternation of two image pairs, we clearly demonstrated that disparities from simultaneously presented monocular inputs dominate those from interocularly delayed inputs. In other words, the human stereoscopic system appears to be very intolerant of interocular time delays between correlated images.

### References

- Barlow H B, Blakemore C, Pettigrew J D, 1967 "The neural mechanism of binocular depth discrimination" *Journal of Physiology* **193** 327–342
- Burr D C, Morrone M C, 1993 "Impulse-response functions for chromatic and achromatic stimuli" *Journal of the Optical Society of America A* **10** 1706–1713
- Chichilnisky E J, Kalmar R S, 2002 "Functional asymmetries in on and off ganglion cells of primate retina" *Journal of Neuroscience* **22** 2737–2747
- Cogan A I, Lomakin A J, Rossi A F, 1993 "Depth in anticorrelated stereograms: effects of spatial density and interocular delay" *Vision Research* **33** 1959–1975
- Cormack L, Stevenson S, Schor C M, 1994 "An upper limit to the binocular combination of stimuli" *Vision Research* **34** 2599–2608
- Cumming B G, Parker A J, 1997 "Responses of primary visual cortical neurons to binocular disparity without depth perception" *Nature* **389** 280–283
- Cumming B G, Parker A J, 1999 "Binocular neurons in V1 of awake monkeys are selective for absolute, not relative disparity" *Journal of Neuroscience* **19** 5602–5618

- 
- Cumming B G, Read J C, 2003 "Sensitivity to interocular delay in binocular V1 neurons" [Abstract] *Journal of Vision* **3**(9) 61a
- Cumming B G, Shapiro S E, Parker A J, 1998 "Disparity detection in anticorrelated stereograms" *Perception* **27** 1367–1377
- DeAngelis G C, Ohzawa I, Freeman R D, 1991 "Depth is encoded in the visual cortex by a specialised receptive field structure" *Nature* **352** 156–159
- Edwards M, Pope D R, Schor C M, 1999 "Orientation tuning of the transient-stereopsis system" *Vision Research* **39** 2717–2727
- Ee R van, Richards W, 2002 "A planar and a volumetric test for stereoanomaly" *Perception* **31** 51–64
- Howard I P, Rogers B J, 2002 *Seeing in Depth* Volume 2 *Depth Perception* (Toronto: I Porteous)
- Ikeda M, 1965 "Temporal summation of positive and negative flashes in the visual system" *Journal of the Optical Society of America* **55** 1527–1534
- Ikeda M, 1986 "Temporal impulse response" *Vision Research* **26** 1431–1440
- Julesz B, 1971 *Foundations of Cyclopean Perception* (Chicago, IL: University of Chicago)
- Julesz B, Tyler C W, 1976 "Neurontropy, an entropy-like measure of neural correlation in binocular fusion and rivalry" *Biological Cybernetics* **23** 25–33
- Julesz B, White B, 1969 "Short term visual memory and the Pulfrich phenomenon" *Nature* **222** 639–641
- Kremers J, Lee B B, Pokorny J, Smith V C, 1993 "Responses of macaque ganglion cells and human observers to compound periodic waveforms" *Vision Research* **33** 1997–2011
- Masson G S, Bussetini C, Milles F A, 1997 "Vergence eye movements in response to binocular disparity without depth perception" *Nature* **389** 283–286
- Mitchell D E, O'Hagan S, 1972 "Accuracy of stereoscopic localization of small line segments that differ in size or orientation for the two eyes" *Vision Research* **12** 437–454
- Morgan M J, Fahle M, 2000 "Motion-stereo mechanisms sensitive to interocular phase" *Vision Research* **40** 1667–1675
- Poggio G F, Fisher B, 1977 "Binocular interaction and depth sensitivity in striate and prestriate cortex of behaving rhesus monkey" *Journal of Neurophysiology* **40** 1392–1405
- Pope D R, Edwards M, Schor C M, 1999 "Extraction of depth from opposite-contrast stimuli: transient system can, sustained system can't" *Vision Research* **39** 4010–4017
- Ross J, Hogben J H, 1974 "Short-term memory in stereopsis" *Vision Research* **14** 1195–1201
- Schor C M, Edwards M, Pope D R, 1998 "Spatial-frequency tuning of the transient-stereo system" *Vision Research* **38** 3057–3068
- Shinomori K, Werner J S, 2003 "Senescence of the temporal impulse response to a luminous pulse" *Vision Research* **43** 617–627
- Swanson W H, Uneno T, Smith V C, Pokorny J, 1987 "Temporal modulation sensitivity and pulse-duration thresholds for chromatic and luminance perturbations" *Journal of the Optical Society of America A* **4** 1992–2005
- Tyler C W, 1991 "Cyclopean vision", in *Binocular Vision* Ed. D Regan, volume 9 of *Vision and Visual Dysfunction* Ed. J R Cronly-Dillon (London: Macmillan) pp 38–74
- Tyler C W, 1992 "Psychophysical derivation of the impulse response through generation of ultra brief responses: complex inverse estimation without minimum-phase assumptions" *Journal of the Optical Society of America A* **9** 1025–1040
- Tyler C W, Julesz B, 1976 "The neural transfer characteristic (neurontropy) for binocular stochastic stimulation" *Biological Cybernetics* **23** 33–37
- Tyler C W, Julesz B, 1978 "Binocular cross-correlation in time and space" *Vision Research* **18** 101–105





ISSN 0301-0066 (print)

ISSN 1468-4233 (electronic)

# PERCEPTION

VOLUME 34 2005

[www.perceptionweb.com](http://www.perceptionweb.com)

**Conditions of use.** This article may be downloaded from the Perception website for personal research by members of subscribing organisations. Authors are entitled to distribute their own article (in printed form or by e-mail) to up to 50 people. This PDF may not be placed on any website (or other online distribution system) without permission of the publisher.