What visual information is used for stereoscopic depth displacement discrimination?

Submitted in revised form to Perception

Harold T Nefs & Julie M Harris

University of St Andrews

School of Psychology

St Andrews, United Kingdom

All correspondence should be addressed to:

Dr Harold T Nefs

Delft University of Technology

Faculty of EEMSC

Man-Machine Interaction Group

Mekelweg 4

2628 CD Delft

The Netherlands

☎ +31 (0)15 2783914

🗐 +31 (0)15 2787141

✉ H.T.Nefs@TUDelft.nl

Abstract

There are two ways to detect a displacement in stereoscopic depth, namely by monitoring the change in disparity over time (CDOT) or by monitoring the inter-ocular velocity difference (IOVD). Though previous studies have attempted to understand which cue is most significant for the visual system, none have designed stimuli that provide a comparison in terms of relative efficiency between them. Here we used two-frame motion and random dot noise to deliver equivalent strengths of CDOT and IOVD information to the visual system. Using three kinds of random dot stimuli, we were able to isolate CDOT or IOVD or deliver both simultaneously. The proportion of dots delivering CDOT or IOVD signals could be varied, and we defined discrimination threshold as the proportion needed to detect the direction of displacement (towards or away)[1]. Thresholds were similar for stimuli containing CDOT only, and containing both CDOT and IOVD, but only one participant was able to consistently perceive the displacement for stimuli containing only IOVD. We also investigated the effect of disparity pedestals on discrimination. Performance was best when the displacement crossed the reference plane, but was not significantly different for stimuli containing CDOT only, or containing both CDOT and IOVD. When stimuli are specifically designed to provide equivalent two-frame motion or disparity-change, few participants can reliably detect displacement when IOVD is the only cue. This challenges the notion that IOVD is involved in the discrimination of direction of displacement in two-frame motion displays.

---

[1] Note that we refer here to displacement in depth, not motion in depth. There is an ongoing debate about the extent to which motion in depth is processed via explicit motion-sensitive mechanisms, or via mechanisms sensitive to position that are monitored for changes over time. We are agnostic on that point here.

1. Introduction

Previous studies have not yet produced a conclusive answer about what binocular mechanisms underlie the discrimination of motion in depth (Harris, Nefs & Grafton, 2008). One specific open question asks about the order of processing of depth displacements across time. There are at least two possibilities: A) the visual system first computes binocular disparities and next evaluates the change in disparity over time (CDOT), or B) the visual system first calculates the lateral displacements in both eyes and next calculates the difference between signals from the two eyes. The latter is often referred to as an inter-ocular velocity difference (IOVD) detector[2]. At the mathematical level, the two alternatives are perfectly equivalent as shown in Equation 1 except for the order of operations (e.g., Regan, 1993; Cumming & Parker, 1994).

Equation 1:

$$\frac{\delta(r-l)}{\delta t} = \frac{\delta r}{\delta t} - \frac{\delta l}{\delta t}$$

In this Equation, r and l are retinal positions of a target in the right and left eye respectively and δt signifies the time lapse. Since subtraction and taking the derivative are both linear operators, they are by definition commutative. In the real world these two cues, changing disparities and inter-ocular velocity differences, are invariably coupled. The question, however, is not about the mathematics, but about the order of implementation in the visual system. Cumming and Parker (1994) were probably the first to investigate this particular problem experimentally, although Rashbass and Wertheimer (1961) did, if only as a side note, describe both alternatives over three decades earlier.

If one takes a brief glance at the literature, it appears there is substantial evidence of the use of IOVD for motion in depth perception (e.g., Brooks, 2001, 2002a, 2002b; Brooks & Mather, 2000; Brooks and Stone, 2004, 2006; Fernandez & Farell, 2005, 2006; Maeda et al., 1999; Shioiri, Saisho & Yaguchi, 2000, 2008). However, when surveyed in detail, as we will show below, the

---

[2] Explicit coding of velocity is not in principle necessary for displacement detection. Displacement in depth equals the linear sum of $(r_1-r_2) - (l_1-l_2)$, where the suffix number indicates the moment in time and r and l are positions in the right and left eye respectively. The order of elements in this linear sum is arbitrary. In this paper we consider only two-frame motion and because speed is ill defined in this stimulus, we use the term depth displacement rather than motion.

few studies that have measured discrimination have done so in very different ways, and found very different results from one another, making it difficult to interpret what the *relative* contribution of CDOT and IOVD sensitive mechanisms are to the perception of motion in depth and displacement in depth.

It has been shown using dynamic random dot stereograms (DRDS) that CDOT in itself is a sufficient cue for the perception of motion in depth (Julesz, 1971; Norcia & Tyler, 1984). In contrast to an ordinary random dot stereogram (RDS), in a DRDS the dots have a single-frame lifetime, after which they are replaced by a new set of dots at different random locations. In a DRDS there is thus no coherent motion signal delivered to each eye separately. Mechanisms that are sensitive to differences in lateral displacements of dots between the two eyes can therefore not contribute to depth displacement perception in a DRDS.

Similarly, a random dot stereogram can be created that is uncorrelated between the eyes, but correlated over time. We will call this a TCRDS (time correlated random dot stereogram). Shioiri, Saisho and Yaguchi (2000) showed that participants were able to correctly identify the direction of motion in depth from two-frame TCRDS stimuli. For a fixed depth displacement, they measured direction discrimination as a function of stimulus contrast. In their experiment 1, performance was statistically significantly above chance level, but they also reported a high level of binocular rivalry in the stimulus. Cumming and Parker (1994) did not find evidence for an IOVD-specific mechanism. They found that depth displacement detection is as good for DRDS as it is for RDS in a multiple frame stimulus that oscillated back and forth in depth. Their study therefore did not address displacement discrimination using supra-threshold depth steps. Several authors have used less direct methods to demonstrate that IOVD cues might contribute to the perception of motion in depth. Fernandez and Farell (2005), for example, showed that adaptation to fronto-parallel motion facilitates the discrimination of direction of motion in depth. These studies suggest the involvement of IOVD sensitive mechanisms in motion in depth discrimination. The studies do not tell us about the relative contributions that CDOT and IOVD sensitive mechanisms make when both cues are present. We have recently reviewed the literature on this topic extensively elsewhere, and concluded that there is tentative evidence in favor of the use of IOVD in motion in depth perception (Harris et al., 2008). What remains unclear is how much CDOT and IOVD cues each contribute to the perception of displacement in depth and motion in depth. Even though some studies are suggestive of IOVD sensitive mechanisms in motion in depth perception, their true importance *relative* to CDOT mechanisms remains elusive.

Previous research has focused on the respective roles of CDOT and IOVD for the detection and discrimination of depth displacements as a function of displacement amplitude (e.g., Cumming and Parker, 1994) or contrast (Shioiri et al, 2000). We take a different approach to the CDOT / IOVD problem in this paper, which allows us to make a direct comparison between the use of CDOT and IOVD information for the perception of motion in depth. We ask the question whether CDOT or IOVD sensitive mechanisms are used to detect average depth displacement of a region containing both signal dots (with CDOT or IOVD information, or both) and noise dots (containing no useful signal). There is a substantial conceptual difference between depth displacement discrimination, as determined by amplitude, and by the proportion of signal amongst noise. The question here is how much coherence there must be in a stimulus, over time and/or over the two eyes, in order to see the stimulus change position in depth. An important advantage of measuring thresholds in this way, over minimal displacement discrimination thresholds, is that the displacements used can be supra-threshold in terms of amplitude. The relative importance of CDOT and IOVD that was found in depth displacement detection and discrimination in terms of minimum displacements does not tell us about their relative importance for discrimination of depth changes when displacements are larger and corrupted by noise. This approach has been used earlier to determine, for example, the statistical efficiency of stereopsis (Harris and Parker, 1992, 1994a, 1994b), where the performance of human participants was compared with an 'ideal observer' model. Wallace and Mamassian (2003, 2004) have constructed ideal participant models for both motion detection and static disparity detection. Similar approaches have also been used to study processing channels underlying stereopsis (Cormack, Stevenson & Schor, 1993) and how disparity information drives vergence (Stevenson, Cormack & Schor, 1994).

A key feature of our design is that we use stimuli that consist of two frames in time for each eye. In this way, we can make a direct comparison between the use of IOVD and CDOT cues. In this comparison the amount of potential visual information in a DRDS and a TCRDS is then equivalent: there are two time frames, and two 'eye frames' (see for example Figure 2). Our design was specifically aimed at delivering (as closely as possible) the equivalent visual information in the *stimulus*, for the two stimulus conditions. A 2-eye by 2-frame design achieves this aim. Thus, an 'ideal observer' would receive the same visual information in each case. We avoid the need to compare human performance with an ideal observer, because we are interested in the comparison between the two conditions (and because the ideal observer would take the same form for the two conditions). We therefore obtain a measure of relative efficiency of the two sources of information by simply comparing coherence thresholds for the two conditions.

We could have taken a different approach because we do already know a lot about real human systems, for example that 2-frames is not optimal for motion mechanisms and that short durations are not optimal for stereo mechanisms. We did not wish to do that here because such an approach would have led to unwanted complexity and assumptions in the experimental design. However, our design could have favoured one or other source of information. The implications of this in the light of our data will be detailed in the Discussion

Many previous studies have been biased in some way towards one or the other of the two cues. For example, by using DRDS stimuli, the number of frames across time can be made infinitely large, whereas the number of eyes is necessarily restricted to two. There is thus more information in such a DRDS stimulus than in a TCRDS. In this paper, we raise the question of how much CDOT and IOVD-based mechanisms contribute each to 2-frame depth displacement discrimination. In Experiment 1, we measure depth displacement discrimination in RDS, DRDS and TCRDS in a straightforward fashion. In Experiment 2, we explore the effects of a disparity pedestal on the direction discrimination threshold for RDS and DRDS stimuli. We also tested the effect of a pedestal on lateral motion for RDS and TCRDS stimuli.

2. Experiment 1

2.1. Method

Participants. Two male and five female participants contributed data to this experiment. Ages ranged between 19 and 38 years. Participant HTN is the first author. Participants VFD and CG are semi-naive coworkers. The other participants were naive with respect to the hypotheses. All participants had normal or corrected-to-normal spatial acuity for near and far distances, accurate color vision as determined with the Ishihara tests for colour blindness (Kanehara Shuppan Co, Ltd), and a stereo-acuity of 40 seconds of arc or better as determined with the stereo butterfly test (Stereo Optical Co., Inc). All experiments reported here were approved by the University Teaching and Research Ethics Committee of the University of St. Andrews and carried out in accordance with the convention of Helsinki. Written informed consent was obtained from all participants.

Materials. All stimuli were presented on an Iiyama 22-inches Vision-Master Pro 514 CRT-screen. The vertical refresh rate was set to 160 Hz with a resolution of 1024 (H) X 768 (V) pixels. Stereoscopic depth was obtained using Crystal-Eyes shutter glasses that were synchronized to the vertical refresh signal of the monitor. There was thus a refresh rate of 80Hz per eye. Luminance of the foreground, as measured through the shutter glasses, was 1.05 $cd.m^{-2}$ for both eyes. Background

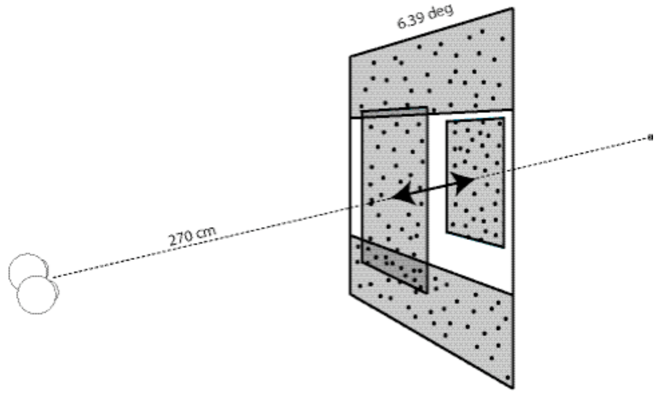

Figure 1. *Experimental set-up. The centre square that is located between the two horizontal bars jumps towards or away from the observer. The observer has to indicate the direction of displacement (away / towards). The centre square does not change its size on screen when it jumps away or towards the observers. Depending on the condition only the location of the dots in the half images changes.*

luminance was smaller than the resolution of our luminance meter (<0.01 $cd.m^{-2}$), as was cross-talk. There was no direct illumination in the room except for the computer monitor. Walls, floor, and ceiling were painted black. There was no visible outline of the computer monitor against the background. Stimuli were presented in red on a black background because red delivers the least crosstalk between the half-images of the stereograms. There was no visible crosstalk. Stimulus presentation was under control of a dual processor Pentium-IV personal computer with an ATI FireGL-3100V graphics card. Stimuli were presented with the psychophysics toolbox for Matlab® (Brainard, 1997; Pelli, 1997).

Procedure. Participants were seated in a dark room at a distance of 2.70 meters from the computer monitor with their heads resting on a chin rest. Responses were recorded from the computer keyboard. At no time during the experiment did we provide feedback to the participants. The entire stimulus pattern subtended a visual angle of 6.36 degrees (H) and 6.36 degrees (V). A schematic representation of the stimulus is shown in Figure 1. The stimulus consisted of two frames separated in time for each eye. Each frame consisted of a centre square measuring (3.18 x 3.18 deg) and two horizontally oriented rectangles (6.36 x 1.59 degrees) as a surround. We presented 100 dots at random locations within the boundary of the inner square in each eye. Dot density was the same in the surround as in the centre square. The random dot pattern in the surround was identical in the two eyes, defining a zero disparity reference frame, and static for the duration of a trial. The centre square was displaced in stereoscopic depth across the two frames. Participants had to indicate the direction of the displacement (towards or away from them).

There were three experimental stimulus conditions for the centre square: a random dot stereogram (RDS), a dynamic random dot stereogram (DRDS) and a time correlated random dot stereogram (TCRDS). The stimulus arrangements for these three experimental conditions are illustrated in Figure 2. For each condition, a proportion of dots were designated as 'signal dots' and the others were designated as 'noise dots'. Signal dots have notional partners in each half-image and across time in a RDS, in the two half-images only but not across time in a DRDS, and across time but not in the two half images in a TCRDS. The noise dots do not have partners, neither across half images nor across time. In the RDS condition, signal dots remained in the same horizontal and vertical position on screen, but changed their disparity with respect to the surround from –9.9 minutes of arc to +9.9 minutes of arc (or the other way around depending on the direction of displacement). Thus the disparity step-size was 19.8 minutes of arc and the corresponding displacement at each eye was 9.9 minutes of arc. In the DRDS condition, signal dots changed their disparity as in the RDS condition, but they were also randomly displaced in horizontal and vertical positions within the centre square. As a result there was no coherent displacement in horizontal position across time in each eye, but there was a change in disparity across time. In the TCRDS condition, signal dots were randomly displaced across the two half images, but they were displaced laterally in opposite directions between the two frames. As a result there was no consistent change in binocular disparity since there was no consistent stereo-correspondence between the two eyes, but there was a change in lateral displacement across time in each eye.

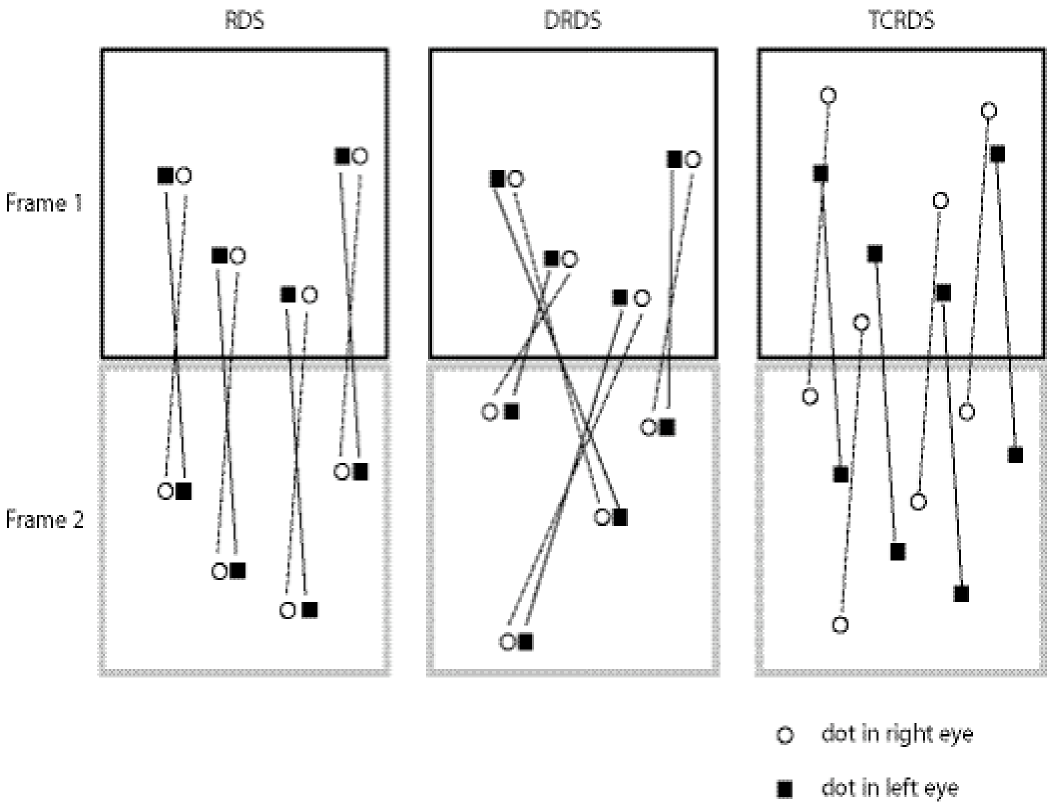

Figure 2. *Schematic illustration of three stimuli: Left column: Random Dot Stereogram (RDS), middle column: Dynamic Random Dot Stereogram (DRDS), and right column: Time Correlation Random Dot Stereogram (TCRDS). Half-dots for the left eye are indicated with squares, and half-dots for the right eye are indicated with open circles. In a RDS squares and circles are always paired next to each other and all pairs have the same disparity. All pairs change their disparities but not their positions from frame 1 to frame 2. In a DRDS all half dots are paired again and change their disparity from frame 1 to frame 2, but also change their position within the frame. The connections between the dots in frame 1 and 2 of the DRDS are drawn at random; in principle every dot in frame 1 can be connected to any dot in frame 2. In a TCRDS, there are no consistent matches between the eyes in either frame 1 or frame 2. All dots shift horizontally from frame 1 to frame 2 and in opposite directions for the two eyes.*

We measured the proportion correct for identification of the displacement direction (towards / away), with the method of constant stimuli, as a function of the proportion of signal dots for the three experimental stimulus conditions (RDS, DRDS & TCRDS) at four different presentation durations (125, 250, 500 and 1000 milliseconds). The first stimulus frame therefore lasted 62.5, 125, 250 and 500ms (5, 10, 20 and 40 CRT frames per eye containing the stimulus, interleaved with the same number of empty monitor frames while the stimulus was presented to the other eye). The second stimulus frame was the same duration as the first stimulus frame. There was no inter-stimulus interval between the first and second stimulus frame. Hereafter we wish to imply 'stimulus frame' when using the word 'frame'. We used 11 different test values ranging from 0 to 100% signal dots in equal step sizes and 40 repetitions per test value. The direction of displacement (towards / away) was counterbalanced across the experiment. All trials were randomly interleaved. The experiment was conducted in 4-5 one-hour sessions. Short breaks of a few minutes were taken every 10 minutes. Sessions were separated by at least a few hours.

2.2. Results

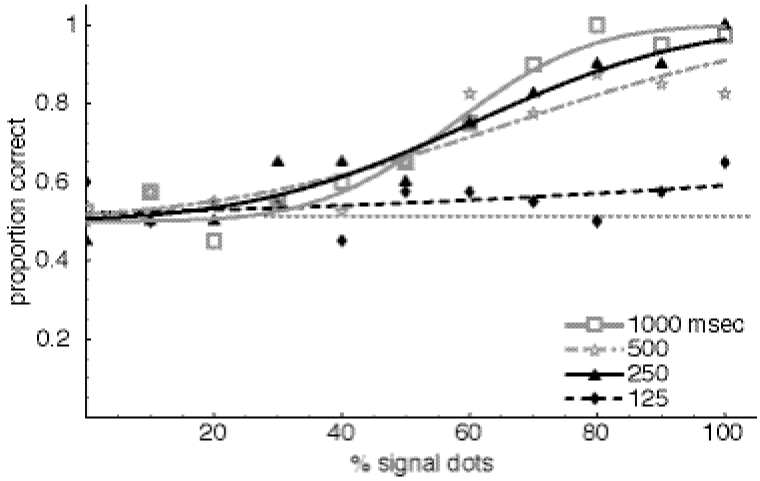

Figure 3. *Examples of fitted psychometric functions taken from observer HN for RDS stimuli. Different curves represent the psychometric functions for different presentation durations. All data points are obtained from 40 trials.*

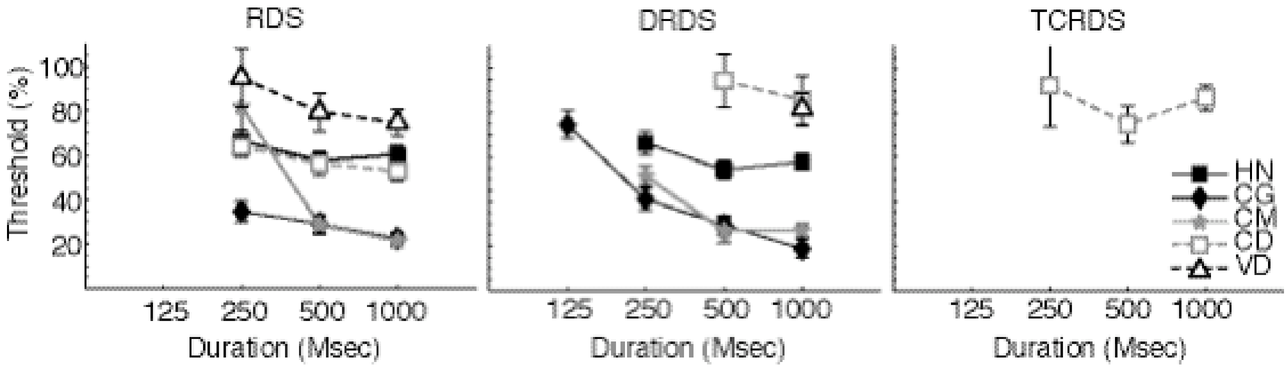

Figure 4. *Direction detection thresholds in percentage signal dots as a function of stimulus duration for all five observers for RDS, DRDS, and TCRDS. The two observers who could not do the task are not included.*

We fitted cumulative Gaussian functions to the data with the Levenberg-Marquardt algorithm (e.g., Marquardt, 1963) with the mean and slope as free parameters. We defined the discrimination threshold as the proportion of signal dots that would give a correct discrimination rate of 0.75. Typical psychometric functions are plotted in Figure 3 for participant HN. The slopes of the psychometric functions are in general quite shallow. Two observers delivered data that did not depart from chance for all conditions. Their data are not shown. Discrimination thresholds for the other 5 observers are plotted in Figure 4. The error bars indicate the standard deviation of the discrimination threshold as calculated with a Monte-Carlo simulation (1000 simulations) on the estimated psychometric function.

The first major point of interest is that only one participant (CD) was marginally able to see the correct direction of displacement in a TCRDS. It was not possible to estimate thresholds for any other participant because full psychometric functions could not be obtained. Data from these 'missing threshold' conditions are plotted in the right column of figure 5. Notice that for most observers, for most conditions, the observer responses were close to chance. Although individual data points can be as high as 75% or as low as 25%, there is no systematic increase in performance as signal strength increases. Occasionally, we found observers (in particular CM, grey circles) whose performance suggested a small decrease as signal strength was increased. Second, we found several observers whose performance was so poor for some conditions that threshold could not be obtained when viewing RDS stimuli (data are shown in the left column of figure 5) and DRDS stimuli (data are sown in the centre column of figure 5).

At the other end of the performance range, thresholds for one participant (CG) could be as low as 25% signal dots. This surprisingly poor performance in general is only a little worse than performance found for a similar task in Shioiri et al (2000), who measured only percent correct and did not attempt to extract threshold estimates from their raw data.

In Figure 6, we plot discrimination thresholds in RDS stimuli against the discrimination thresholds for DRDS for the four participants who had discrimination thresholds below 100%, for both kinds of stimuli. A straight line can approximate the relationship between the discrimination thresholds for the two kinds of stimuli across participants. We fitted a straight line through twelve valid pairs of RDS/DRDS thresholds, using a type II linear regression. R^2 was 0.52 and the slope of the fitted line was 1.30, which was not significantly different from 1. We therefore find no evidence that RDS discrimination thresholds are different from DRDS discrimination thresholds; all data points in Figure 6 are evenly distributed around the "RDS = DRDS line". We tested the mean difference between RDS and DRDS thresholds also with a paired T-Test, but this was not significant. This result is consistent with discrimination data from Cumming & Parker (1994).

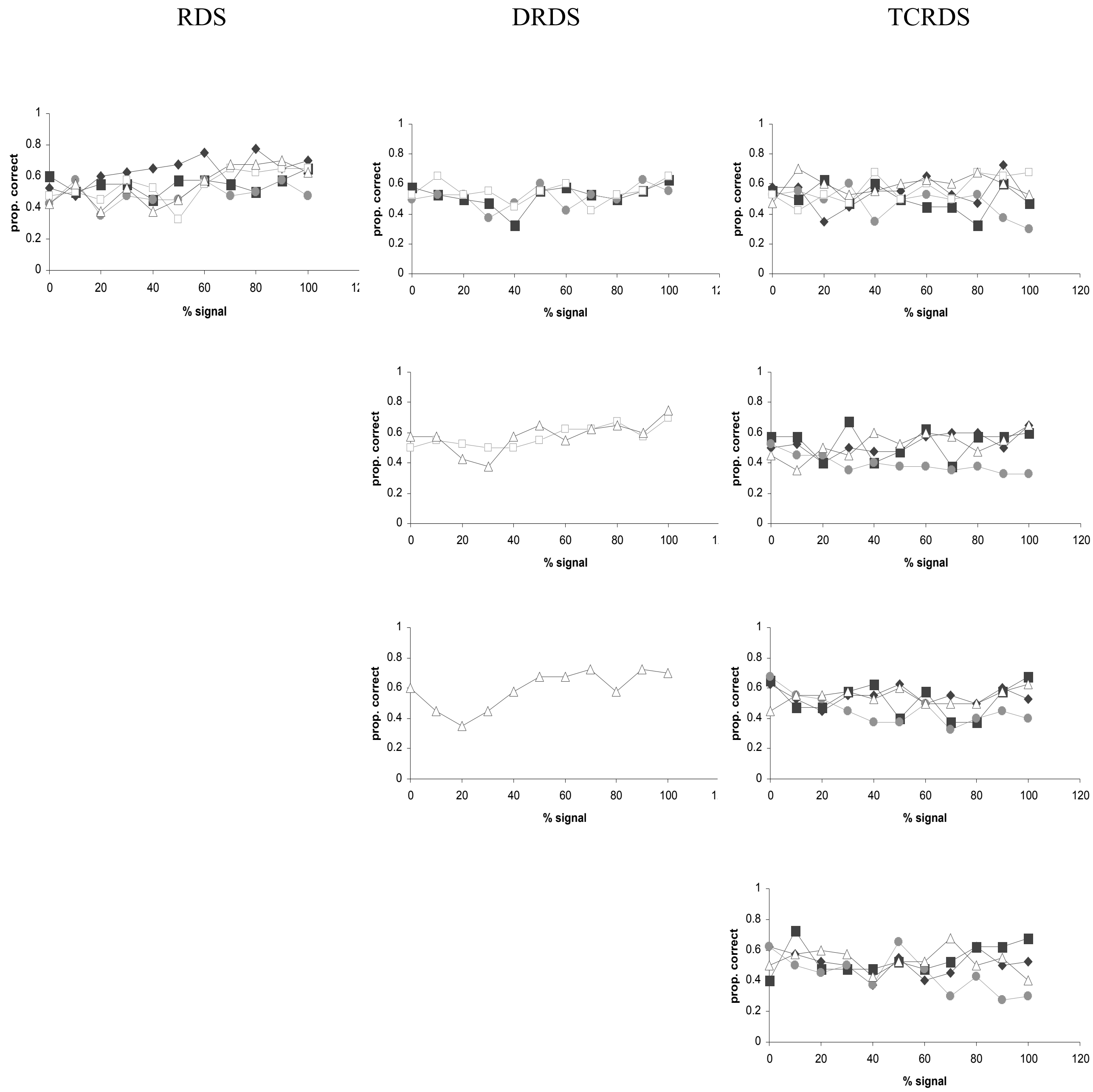

Figure 5. *Proportion correct as a function of % signal strength for RDS (left column), DRDS (centre column) and TCRDS (right column). Each row corresponds to data from a different duration (125ms, 250ms, 500ms, 1000ms from top to bottom). Each symbol represents data from a single observer (symbols as for figure 4). The proportion correct data shown are only for the conditions where we were unable to measure a threshold (threshold data for the other conditions are shown in Figure 4).*

As expected based on earlier research, performance was low in the 125 ms second presentation because it was lower than the minimum time needed for successfully forming depth perceptions from stereo-images (e.g., Cumming and Parker, 1994; Tyler, 1971). Only in one case

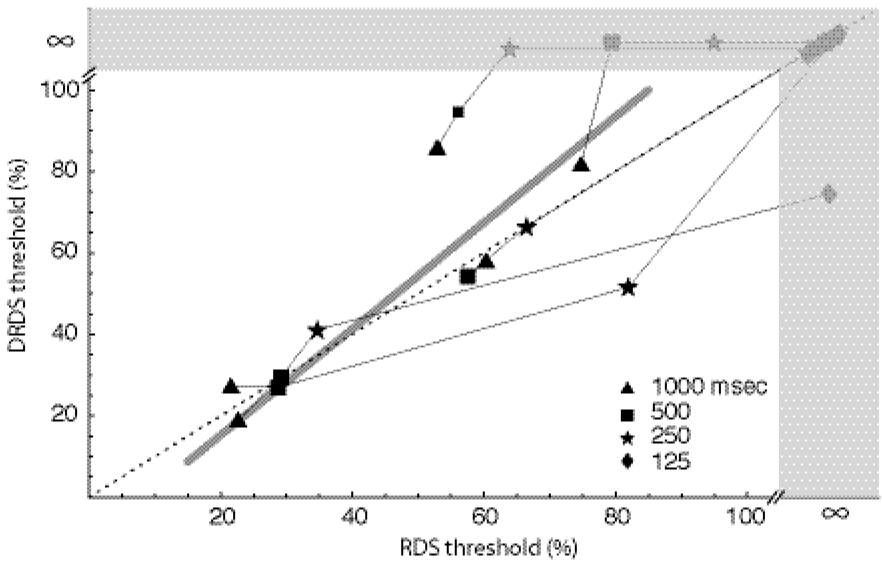

Figure 6. *A scatter plot of DRDS thresholds against RDS thresholds for all observers and all four presentation durations. Different symbols indicate the different presentation durations. Thresholds for the same observer are connected with solid thin lines. Cases where thresholds could not be established below 100% signal intensity are shown in the gray zones with a slight random horizontal or vertical displacement within the gray zone to prevent symbols and lines overlapping each other exactly. The stipple line indicates the "RDS = DRDS" line. The thick solid line is the best-fitting linear regression line (excluding missing values).*

(participant CG in DRDS) was the correct direction of displacement detected above threshold for presentation durations of 125 milliseconds. For the other three presentation durations (250, 500 & 1000 milliseconds), the average discrimination thresholds decrease with increasing duration for both RDS and DRDS stimuli. To quantify this effect, we first normalized the data by dividing the discrimination thresholds by the average discrimination thresholds across conditions for all individual participants to eliminate individual differences obscuring trends within participants. Next we fitted a straight line through the thus obtained normalized thresholds as a function of presentation duration with a type I regression. We found R^2's of 0.42 and 0.56 for RDS and DRDS stimuli respectively. For both RDS and DRDS stimuli we found that the slope was significantly decreasing ($t(13) = 3.09$, $p < .008$ and $t((11) = 3.77$, $p < .003$ respectively).

In summary, we found a surprisingly wide range of abilities between participants when detecting the correct direction of displacement in depth in RDS and DRDS stimuli, with performance generally getting better with longer presentation durations. Although we established that all participants had good stereopsis, some still failed to perform this task to threshold level. There was no significant difference in performance between RDS and DRDS stimuli. Performance for the TCRDS stimuli was so poor that a threshold could not be measured for most participants, and at best marginally for one participant. For our stimulus conditions, when stimuli contain the same amount of potential information as used to compare the discrimination thresholds for CDOT and IOVD, most human visual systems are much worse at using the IOVD information for the discrimination of direction of depth displacement.

3. Experiment 2

Previous research has suggested that, for tasks involving judgments of the speed of motion in depth, there is a performance advantage when stimuli are presented on a depth pedestal, away from the fixation plane (Portfors-Yeomans & Regan, 1996; though see Brooks & Stone, 2004, for contradictory evidence). In contrast, minimum motion in depth discrimination is worse in the presence of a disparity pedestal (Cumming, 1995). In Experiment 2A, we look at the effects of a disparity pedestal on the direction discrimination thresholds for RDS and DRDS stimuli. First we wanted to discover whether RDS thresholds benefit from having opposite lateral displacements in the two eyes when the displacement does not cross the plane of the reference frame. Second, we wanted to explore in more detail how the discrimination of each frame in time contributed to the discrimination of displacement in depth. For example, it is conceivable that participants could see one frame at some test levels but not the other and made their response based on a single frame rather than two, thereby performing better or worse depending on their strategy. If their strategy were to always respond "away" when the first frame was seen in front of the reference plane, their thresholds would be lower than if they were to always respond "towards" in that case. Strategies such as this are possible because we did not provide any feedback during the experiments.

In Experiment 2B, we look at the effects of a lateral displacement pedestal on the thresholds for RDS and TCRDS thresholds. Note that a lateral displacement pedestal in a TCRDS stimulus is symmetrical to the addition of a disparity pedestal in a DRDS stimulus. As participant CD was the only one capable of perceiving displacement using the TCRDS stimuli, we only tested her in this part of the experiment. Similarly to Experiment 2A, we wanted to find out if performance, when viewing a RDS, benefits from having a change in disparity relative to TCRDS stimuli, when the displacement is not directly towards/away from the participant, but at an angle, and likewise if the displacement in a single eye contributes to performance as much as the difference in displacement.

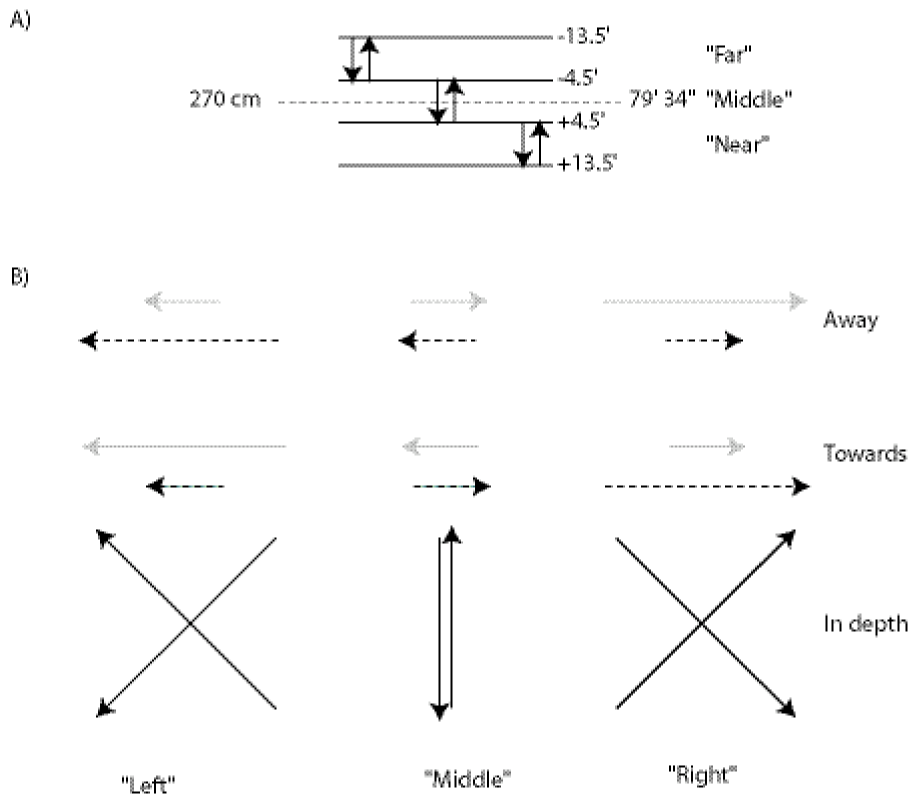

Figure 7. *A) Schematic illustration of the three depth conditions in Experiment 2 for RDS and DRDS stimuli. B) The three motion conditions as tested with observer CD with RDS and TCRDS stimuli.*

A depth displacement towards the participant is always associated with a leftward displacement in the left eye and rightward displacement in the right eye. It is conceivable that the displacement in one eye is detected at lower test levels than in the other eye, and depending on the strategy this might lead to higher or lower thresholds. We knew already that, except perhaps for observer CD, none of the other participants were able to reliably use that cue to perform the task because their responses were below threshold.

3.1. Method

Participants. Seven participants (two male and five female) contributed data to this experiment. Ages ranged between 19 and 33 years. Other particulars were as for Experiment 1. Participants TC and DL only completed about 3/4 of the experiment since it was already clear by then, that they could not do the task as required. Participants HN, CG, CM and CD also participated in Experiment 1, the others did not.

Materials. As described in Experiment 1.

Procedure. Experimental set-up was as in Experiment 1. There were three displacement conditions in Experiment 2A: from –13.5' to –4.5', from –4.5' to +4.5', and from +4.5` to +13.5 minutes of arc, both away and towards the participant. These conditions are illustrated in Figure 7A. We used half of the depth displacement step of Experiment 1 in order to remain within the area that can easily be fused binocularly for all disparity pedestals. We used only one presentation duration, namely 1000ms. We tested both RDS and DRDS stimuli. We did not test with TCRDS stimuli, because the different disparity pedestals do not deliver different stimuli in a TCRDS. That is, the experimental factor has no effect on TCRDS stimuli because the two half images are uncorrelated. All other aspects of the procedure are as described in Experiment 1.

We tested participant CD from Experiment 1, in Experiment 2B with RDS and TCRDS stimuli in which a lateral displacement pedestal was added to the stimulus. This manipulation is illustrated in Figure 7B. We did not test DRDS stimuli for this condition because the lateral pedestal does not have an effect in these stimuli. Pedestal sizes for lateral motion were the same as for disparity in angular units. Other particulars were also as in Experiment 2A.

3.2. Results

Experiment 2A

We fitted cumulative Gaussian functions to the data for all conditions and all participants, and calculated thresholds based on the 75% point, as in Experiment 1. Thresholds are shown in Figure 8A for all participants except participants DL and TC who did not have thresholds below 100% for any of the conditions. Participants HN and CD did not have thresholds below 100% signal dots in DRDS stimuli when the depth displacement did not cross the plane of the reference frame. For all participants for whom thresholds could be measured, thresholds were higher when the displacement did not cross the plane of the reference frame compared to the case when it did cross that plane.

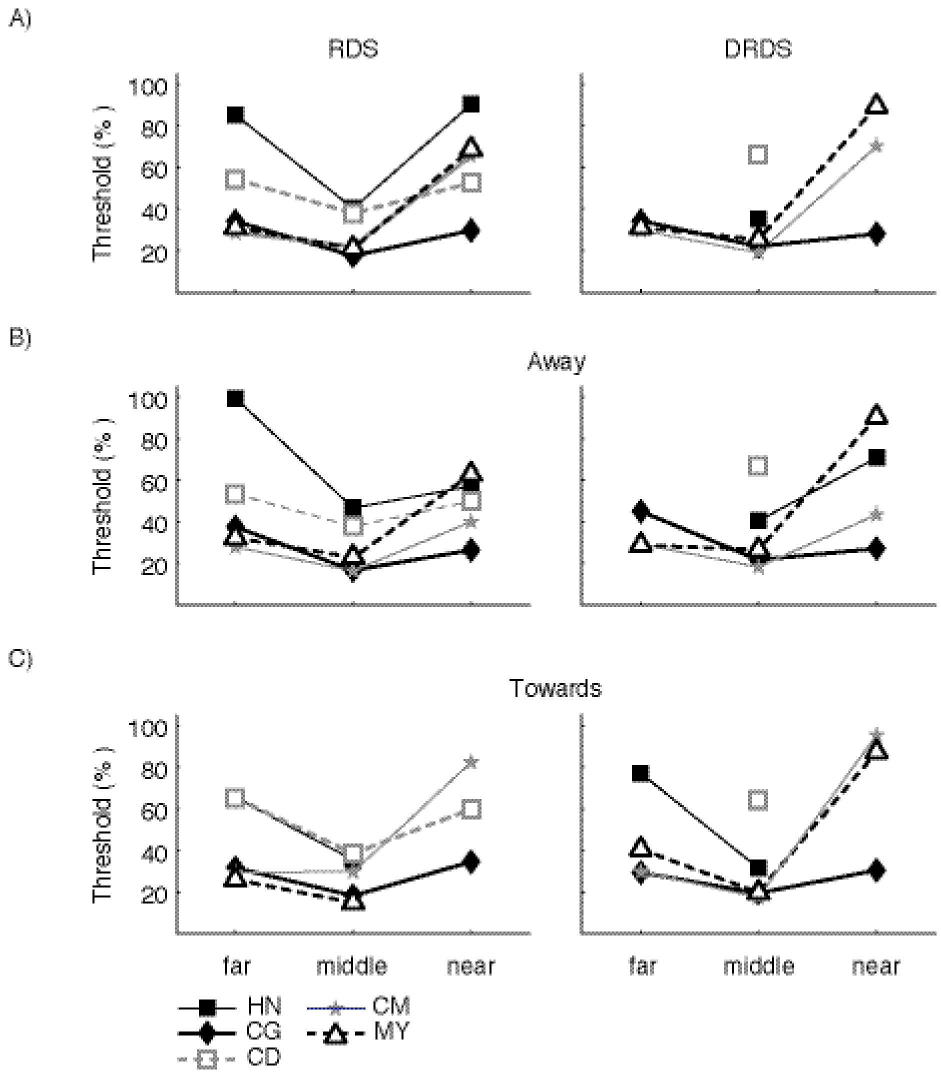

Figure 8 *The thresholds for all observers irrespective of direction are shown in the top row (A). Thresholds for motion away from the observers are shown in the middle row (B) and motion towards the observer are shown in the bottom row (C). The left column is for RDS stimuli and the right column for DRDS stimuli. Different symbols indicate different observers.*

In Figure 9 we plotted DRDS thresholds against RDS thresholds. Thresholds from the same observer in different conditions are connected with thin lines. We calculated the best fitting straight line with a type II regression on all valid pairs of RDS and DRDS thresholds. The slope was not significantly different from 1. The mean thresholds for RDS and DRDS were also not statistically different when tested with a paired T-test. Note, however, that there are several data pairs, for which RDS thresholds were not accompanied by DRDS thresholds, leading to missing values. In these missing values cases, RDS thresholds were thus always lower than DRDS thresholds. If we

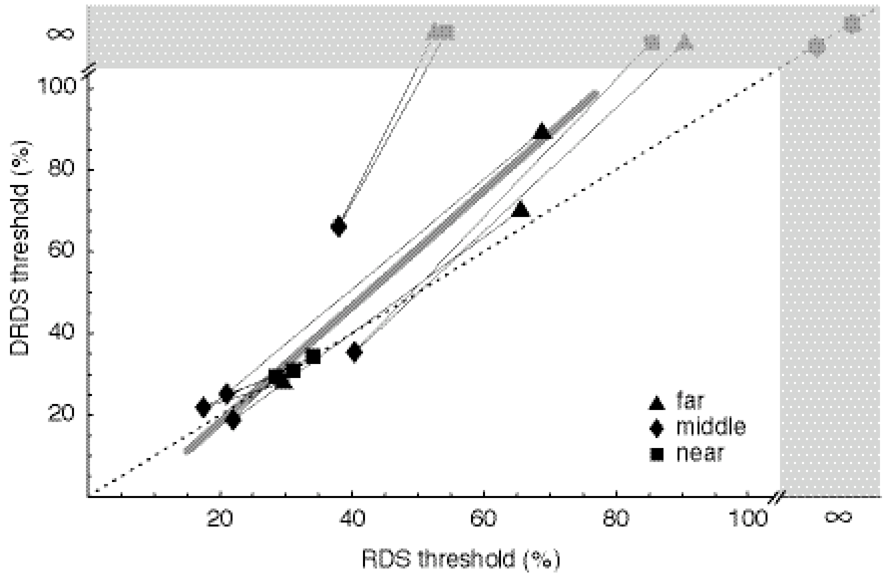

Figure 9. *Scatter plots of DRDS thresholds against RDS thresholds for all observers and all three disparity pedestals. Different symbols indicate the different pedestals. Thresholds for the same observer are connected with solid lines. The stipple line indicates the "RDS = DRDS" line. The gray solid line is the best-fitting linear regression line (excluding missing values).*

were to replace the missing values with the maximum possible performance score (ie 100%) we would obtain a significant difference between RDS and DRDS ($T(20) = 8.16$, $p < .01$). There is therefore a hint here that performance using DRDS stimuli is poorer than when using RDS stimuli.

It would be possible to obtain better than chance performance on this task if participants saw only the first, or only the second frame, and responded according to whether the frame seen was in front of, or behind, the stationary reference. In order to reveal if participants used the depth of the first or second frame as cue to the direction of displacement, we split all data in two sub-conditions, namely the directions "away" and the directions "towards" the participant. For example, when the square is in front of the reference plane in the first time frame, the bias is to say "away" which should lead to a higher threshold for motion towards when the motion is indeed "towards" as in the "near" condition when the displacement is even further towards to participant. In the far condition, the opposite is the case; if the first frame is behind the reference frame and the bias is then to say towards, this will lead to a higher threshold for "away" and a lower threshold for "towards" trials. We fitted cumulative Gaussians to these datasets. We estimated the response bias in these two data sets by looking at the probability of responding "away" across all trials where the stimuli contained

0% signal dots. In an unbiased participant the expected probability should be 0.5. The response bias is then taken as the lower asymptote of the cumulative Gaussian function. The discrimination threshold is now defined as the percentage of signal dot that gives a probability of correct discrimination of the direction of displacement midway between the lower asymptote and 1. Results are shown in Figures 8B (displacement away) and 8C (displacement towards) for all participants.

The data suggests that at least for some participants, discrimination of the depths of the first and second frames are not detected equally well. For example participant HN (black squares in Figure 8) shows a high threshold when the motion is away from the participant and lower when the motion is towards for the far condition, and the reverse for the "near" condition. This suggests that this participant takes the first and/or second frame into account when making perceptual decisions.

Looking at the psychometric functions in detail we can see behavior that is consistent with using only the first or second time frame for direction discrimination in participants HN and CM. When the percentage of signal dots is low and a decision is made based on one frame only, this results in a part of the psychometric function showing performance that is consistently worse than chance. When the percentage of signal dots further increases, these participants pick up the second frame as well, resulting in increasing performance. Although this pattern of behaviour is by no means present or clear in most participants, it is an illustrative feature in a few psychometric functions from observers HN and CM that needs mentioning as a side note. In Figure 10A we have shown a clean example of a psychometric function of participant CM that shows a characteristic dip below chance level in the psychometric function for both RDS and DRDS stimuli, but only for the "far and away condition". Participant HN shows these dips for RDS and DRDS conditions, but only for the "near and towards" condition. Such dips are not observed in any of the other conditions for this participant, nor in any conditions for other participants. In Equation 2 and illustrated in Figure 10B, we have shown how these dips below chance could be understood as the simple sum of the chances of seeing the first and/or second frame.

<u>Equation 2:</u> $$P(\text{“Away”} \mid (\text{far \& away})) = 0.5 + (P1 * P2 + P1*(1-P2) - (1-P1)*P2)/2$$

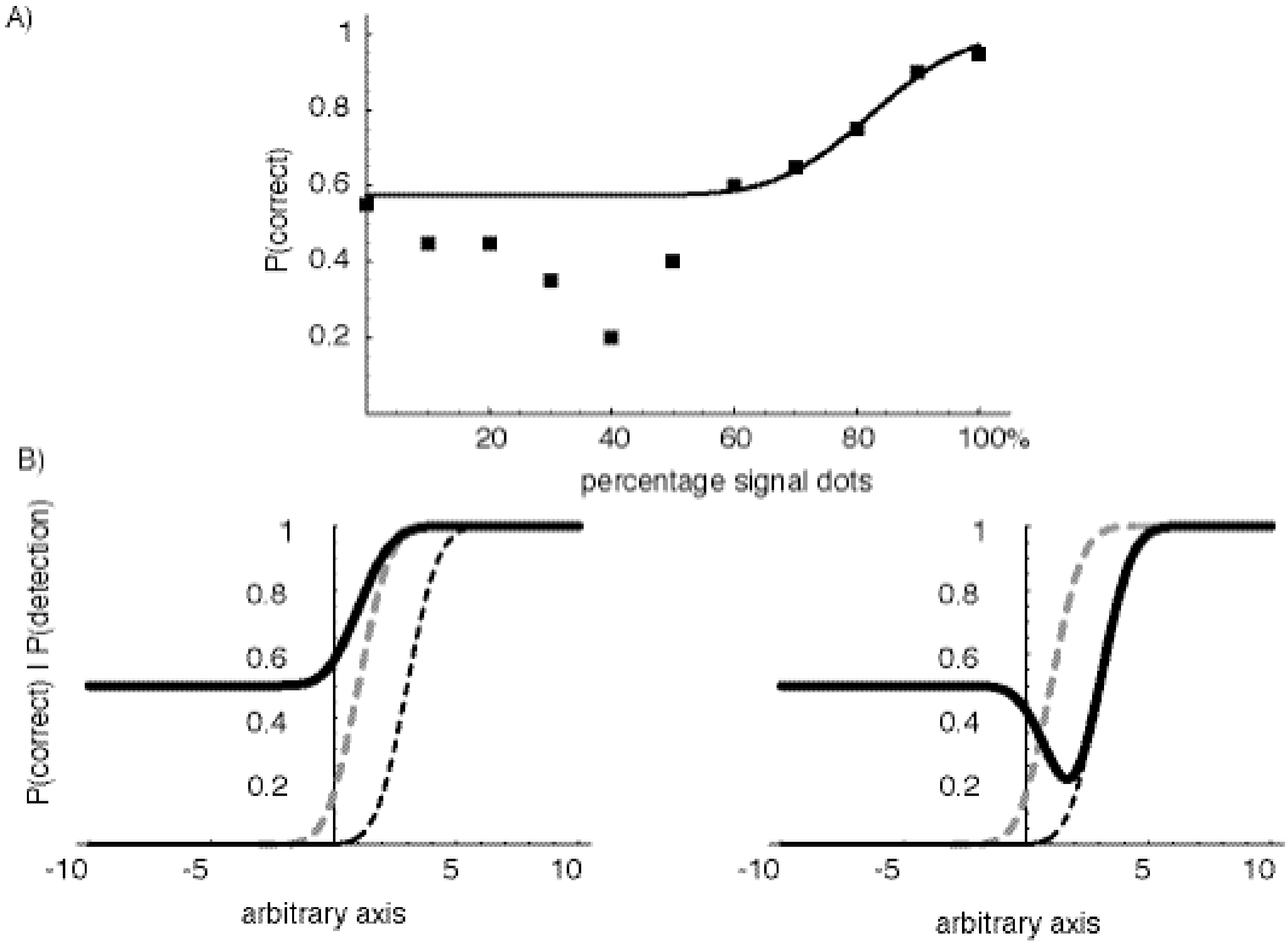

Figure 10. *A) The psychometric function for observer CM in the "far and away" condition. B) Illustration of the psychometric functions stippled black and gray) for detecting the depth in the stimulus in the first and second frame respectively as a function of number of signal dots. The thick black psychometric function is the predicted performance in a 2AFC task as a function of the percentage signal dots when the depth of the first frame is indicative of the direction of displacement in depth (left panel) and when it is counter-indicative (right panel).*

In this equation, P1 is the probability of seeing depth in the first frame and P2 is the probability of seeing depth in the second frame as a function of the number of signal dots. We take cumulative Gaussian distributions here as the basis for the psychometric functions. Both probability functions are restricted to the range 0 to 1. We have divided the sum, of changes by two and added 0.5 to make it into a 2AFC task. The probability of seeing both the first and the second frame of the stimulus is positively correlated with identifying correctly that the displacement is “away”. Likewise, if the depth of only the first frame is detected and the heuristic is to say away if the depth in the first frame is “front”, then this factor will also contribute positively to the probability of saying away in the far condition. The probability of seeing only the second frame will lead to the wrong decision if the heuristic is to say away if the depth in the second frame is “behind”. This factor is therefore given a negative weight in Equation 2. Of crucial importance for the dip in the psychometric function to occur is that the probability of seeing the second frame is lower than seeing the first frame. This condition is satisfied since we have found that the displacements that do not cross the reference frame are more difficult to discriminate than the ones that do (see Figure 8A). That is, for this condition, the second frame was harder to see because it was further away from the reference plane. We did not take different chance levels for the two functions into account in this explanation of the ‘dips’.

Experiment 2B

Results for participant CD, for RDS and TCRDS stimuli with lateral motion pedestals are summarized in Figure 11. Similar to experiment 2A, we would expect a specific switching here between high and low thresholds for left and rightward pedestals if the task was done on the bases of the lateral motion in right or left eye alone. We do not see indications that this participant uses lateral motion in one eye to do the task. There is no specific switching between high and low thresholds for motion towards and away for the lateral conditions similar to the effects explained in Experiment 2A. There does however seem to be a conspicuous absence of right thresholds, we have no idea why. We did however establish again that this participant was able to use the IOVD cue in isolation in some conditions.

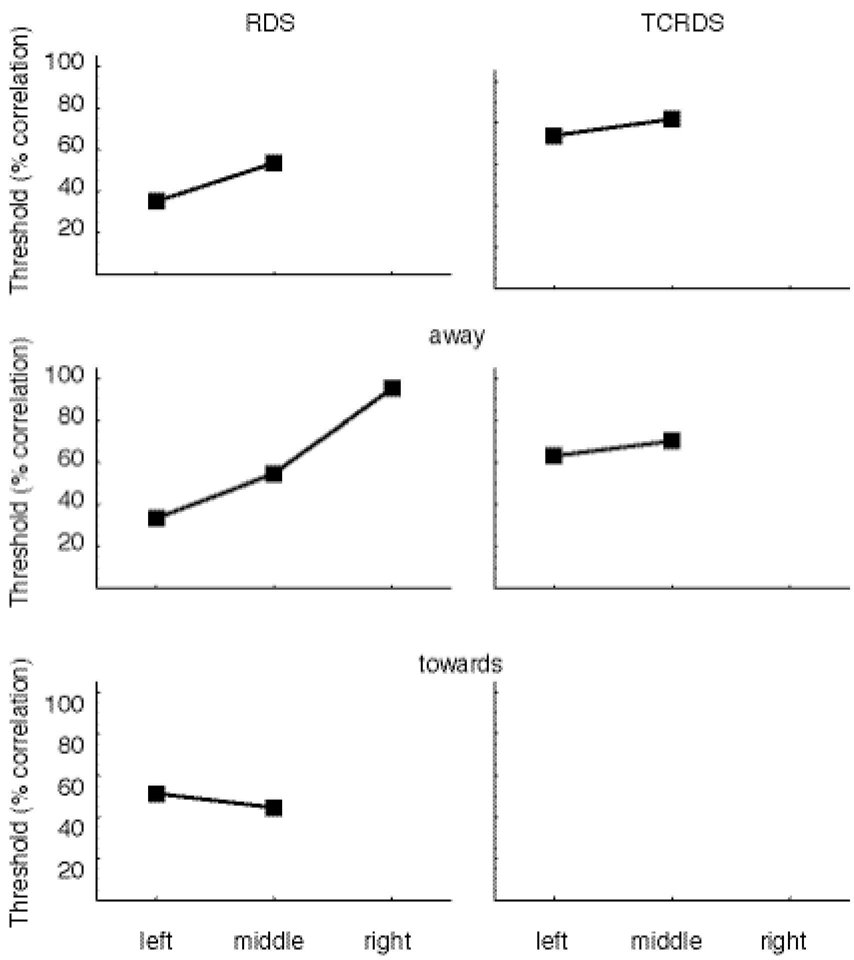

Figure 11. *Thresholds for observer CD irrespective of direction are shown in the top row. Thresholds for motion away from the observer are shown in the middle row and motion towards the observer are shown in the bottom row. The left column is for RDS stimuli and the right column for TCRDS stimuli.*

4. Discussion

4.1 Comparing CDOT and IOVD.

The first experiment was designed to be a direct comparison between CDOT and an IOVD based mechanisms in terms of presenting equivalent information within the stimulus. Only one out of seven participants was able to use the IOVD cue well enough to reach threshold and even for her, close to 100% signal dots were required to detect the direction of displacement at 75% discrimination probability. Considering the group data, we did not find significant differences in performance between RDS and DRDS stimuli, suggesting that for the stimulus conditions used here, the IOVD information in a RDS does not improve discrimination performance over that provided by CDOT information. However, the one observer who showed consistently above threshold performance for the TCRDS stimulus (CD), did show worse performance for the DRDS stimulus, suggesting that occasional observers may benefit from the use of an IOVD mechanism. Overall, we conclude that, for most observers there is thus not much evidence for the involvement of the IOVD mechanism in the discrimination of depth displacements in our 2-frame stimulus. Observer CD may represent a sub-set of observers who do rely more on the IOVD mechanism. An experimental design like this, with only 5 observers, is not able to estimate how common such individuals may be. Further work on a much larger population of observers will be needed to determine how common the pattern observed with observer CD is exhibited across the population (See Nefs, O'Hare & Harris, 2009).

Somewhat surprisingly, discrimination thresholds in RDS and DRDS conditions were high as well, even though all participants had 40 seconds of arc stereoscopic resolution or better. Two out of seven participants were not able to perform the task to our threshold criterion. On the other hand, participant CG required as little as 25% signal dots in the stimulus for reliable discrimination. We thus found very large individual differences in the extent to which participants can make use of the CDOT cue.

The participants' performance seems poor in our experiments at first glance. However, note that in Shioiri et al.'s (2000) study (their Experiment 1) performance in terms of percentage correct identification of the direction of motion in depth in TCRDS stimuli was never larger than 90% for any stimulus condition, and performance in that study was in many conditions so low that it would not have been considered above threshold by our definition (this compares with some observers' performance in figure 5, where data are shown for the conditions where threshold was not reached). Furthermore, Shioiri et al provided the observers with some feedback which might have aided performance. Even though Shioiri et al.'s study shows the use of the IOVD cue, and we confirmed the use of IOVD information here with one of our participants, we have demonstrated that performance on DRDS when an 'equivalent' amount of information is present as in a TCRDS, is far superior.

4.2. A fair test?

One important issue here is the extent to which our stimuli provide a fair comparison of the CDOT and IOVD mechanisms. As a reminder, we would like to emphasise again that our use of a 2 eye x 2 time frame design was intended so that we could compare performance in the two conditions, as if via an ideal observer model. Were an ideal observer model to be developed, the ideal IOVD observer (which would optimally match points across time and then combine binocularly) would be *equivalent* to the ideal CDOT observer (which would optimally match points across eye, and then combine across time). This equivalence arises only because of the 2 eyes x 2 time frames. Only the order of processing of the stimuli would be different. Because of this symmetry in the design, it was not necessary to develop an explict ideal observer model. Instead, we can obtain the relative efficiency between the two conditions by directly comparing human CD with human IOVD performance.

Although the physical stimuli were designed to be exactly equivalent, the extent to which they might suite CDOT and IOVD mechanisms remains an open issue. The comparison made in this paper between IOVD and CDOT used conditions that might have been adverse for the IOVD mechanism, because we know that motion mechanisms require a longer stimulus duration and a larger number of frames (around 5) to operate optimally (McKee, 1981). Our findings here certainly should not be generalized unquestioningly to other stimulus domains such as higher or lower contrasts, displacement step size, luminance, dot density and many more. IOVD might prove to be more useful under different conditions. How processing efficiency for CDOT and IOVD develops over the rest of the possible parameter space remains to be seen.

4.3 Performance at different locations in depth

In the second experiment, we found that discrimination was poorer when the stereoscopic displacement did not cross the plane of the reference frame. This was true for all participants for whom thresholds could be established. This suggests greatest sensitivity around fixation and is what we would expect for a mechanism based on CDOT (Badcock & Schor, 1985; Blakemore, 1970). Performance between RDS and DRDS was not significantly different at zero disparity pedestals, but we found indications that this might not be the case when the stereoscopic displacement did not cross the plane of the reference frame. Detecting the stereoscopic displacement in DRDS stimuli required more signal dots than in RDS stimuli in those cases. Due to missing values (some participants were so poor at the task that thresholds could not be estimated) this comparison did not reach significance. Two participants had thresholds below 100% for RDS stimuli, but not for DRDS stimuli, suggesting a possible advantage when IOVD information accompanies that from CDOT.

In Experiment 2B, we tested one participant (CD) for RDS and TCRDS stimuli with three different lateral motion pedestals. The results did not demonstrate a clear pattern of responding, but they were not consistent with a response pattern that showed a dependency on the lateral motion component in one of the two eyes alone. We did however find, in a new set of data, that this participant was able to pick up the IOVD cue in some conditions. This suggests that the results of Experiment 1 were not an incidental fluke for this participant. It has been argued by Alison, Howard and Howard (1998) that TCRDS stimuli may contain incidental binocular disparity changes over time. It is possible that participant CD used this cue to do the task. She was however not exceptionally good at RDS or DRDS stimuli, which leads us to doubt this explanation. Furthermore, the dot density was relatively low which makes spurious disparity matches less of an issue than in high-density stimuli. It is also worth mentioning that with the issue of spurious disparity matches it is assumed that the same disparity matches are maintained from one time frame to the next. However, a spurious match in one time frame is not necessarily the best match in the next frame. There is no evidence in the current study that other participants were able to pick up that cue since thresholds for TCRDS stimuli could only be established for a single participant.

4.4 Does the task require displacement discrimination?

A critical question that we must ask ourselves is what the discrimination threshold in RDS and DRDS stimuli means. Ability to perform the task above chance level does not necessarily mean that participants can genuinely detect a displacement in a particular direction in depth. We explored this issue in Experiment 2, by manipulating the start and end points of the displacement in depth, to test whether participants could sometimes be responding to depth in just one of the stimulus frames. When the data of Experiment 2A was divided into two subsets for displacement away and towards the participant, we found indications that participants used static depth in the first or second frame as a cue to motion in depth. The finding of "dips" below chance level in the psychometric functions as shown in Figure 10A complemented this finding, although this latter pattern was by no means present or clear for most participants. We showed that this behavior could be expected if participants were relying on both the stereoscopic displacement for some trials, as well as the first or second frame alone for other trials. The general trend in the data of Experiment 2 suggests that the discrimination of the direction of displacement in depth is more difficult when the displacement does not cross the plane of the reference frame than when the displacement does cross the plane of the reference frame. The discrimination of disparity in both the first and the second frame is necessary to detect disparity displacement at threshold level. To further clarify the importance of this finding, the discrimination of depth in the first and second frame precedes (in terms of noise level) the discrimination of the *displacement* in depth. If depth discrimination and displacement in depth were mediated by different mechanisms, this would not necessarily have been the case.

4.5 Motion or displacement?

Another interesting point is whether the mechanisms that make use of the IOVD stimulus are explicitly motion systems whereas those processing CDOT are displacement systems. In other words, because we have used 2-frame motion, our stimulus may favour mechanisms responsive to position, rather than to motion. Some studies suggest that IOVD information is important for speed perception (e.g. Brooks, 2002a) and that IOVD stimuli stimulate the motion brain area MT (Rockers et al, 2009). Others have shown that CD and IOVD mechanisms appear distinct and that they have different temporal frequency tuning (Shioiri et al, 2008). We have recently reviewed this issue in detail elsewhere (Harris et al., 2008). It is quite possible that IOVD performance in the stimulus used in this manuscript is inherently limited because there is no well defined speed in the stimulus. It could be that the better performance for the CDOT stimuli occurred because the CDOT mechanism is not an explicit motion mechanism (ie. one that genuinely responds to motion, rather than being able to monitor that position is different at different times). We cannot address this issue here, because we used only 2-frame motion, rather than a more elaborate multi-frame stimulus.

4.6 Conclusions.

On the whole the evidence in the literature for an IOVD based motion-in-depth mechanism is present, but is limited, and most studies have used indirect indicators to the speed component of motion in depth, such as motion aftereffects and adaptation (e.g. Fernandez & Farell, 2005, 2006). The most direct evidence for an IOVD based mechanism comes from Shioiri et al. (2000). The sizes of the effects found were not very large, as we discussed above. Even with 100% percent correlated stimuli, discrimination rates did not reach 1, as they also did not (for all but one observer) in our experiments. Here we found only one participant who had some significant sensitivity to the IOVD cue as well. In terms of their relative use: the extent to which CDOT and IOVD mechanisms can make of the available cues for detecting displacements in depth, we have found that, for the 2-frame displacement conditions used here, the IOVD mechanism is clearly less useful than the CD mechanism. In summary, we found some direct evidence for an IOVD mechanism in at least some people, but a large IOVD contribution to the discrimination of stereoscopic depth displacement is rare.

Acknowledgments

This work was supported by a grant from the EPSRC.

References

Allison,_R.S,, Howard_I.P. & Howard_A. (1998). Motion in depth can be elicited by dichoptically uncorrelated textures. *Perception,* **27,** ECVP Abstract Supplement

Badcock, D. R. and Schor, C. M. (1985). Depth-increment detection function for individual spatial channels, *J. Opt. Soc. Amer. A* **2**, 1211–1216

Blakemore, C. (1970). The range and scope of binocular depth discrimination in man, *J. Physiol*. **211**,599–622

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10,* 433-436

Brooks, K. (2001). Stereomotion speed perception in contrast dependent. *Perception, 30,* 725-731

Brooks, K.R. (2002a). Interocular velocity difference contributes to stereomotion speed perception. *Journal of Vision, 2,* 218-231

Brooks, K.R. (2002b). Monocular motion adaptation affects the perceived trajectory of stereomotion. *Journal of Experimental Psychology: Human Perception and Performance, 28 (6),* 1470-1482

Brooks, K. & Mather, G. (2000). Perceived speed of motion in depth is reduced in the periphery. *Vision Research, 40,* 3507-3516

Brooks, K.R. & Stone, L.S. (2004). Stereomotion speed perception: contributions from both changing disparity and interocular velocity difference over a range of relative disparities. *Journal of Vision, 4,* 1061-1079

Brooks, K.R. & Stone, L.S. (2006). Stereomotion suppression and the perception of speed: accuracy and precision as a function of 3D trajectory. *Journal of Vision, 6 (11),* 1214-1223

Cormack, L.K., Stevenson, S.B. & Schor, C.M. (1993). Disparity-tuned channels of the human visual system. *Visual Neuroscience, 10 (4),* 585-596

Cumming, B.G. & Parker, A.J. (1994). Binocular mechanisms for detecting motion-in-depth. *Vision Research, 34 (4),* 483-495

Fernandez, J.M. & Farell, B. (2005). Seeing motion in depth using inter-ocular velocity differences. *Vision Research, 45,* 2786-2798

Fernandez, J.M. & Farell, B. (2006). Motion in depth from interocular velocity differences revealed by differential motion aftereffect. *Vision Research, 46,* 1307-1317

Harris, J.M., Nefs, H.T. & Grafton, C.E. (2008). Binocular vision and motion in depth perception. *Spatial Vision, 21 (6),* 531-547

Harris, J.M. & Parker, A.J. (1992). Efficiency of stereopsis in random-dot stereograms. *Journal of the Optical Society of America A, 9 (1),* 14-24

Harris, J.M. & Parker, A.J. (1994a). Constraints on human stereo dot matching. *Vision Research, 34 (20),* 2761-2772

Harris, J.M. & Parker, A.J. (1994b). Objective evaluation of human and computational stereoscopic visual systems. *Vision Research, 34 (20),* 2761-2772

Maeda, M., Sato, M., Ohmura, T., Miyazaki, Y., Wang, A. & Awaya, S. (1999). Binocular depth-from-motion in infantile and late-onset esotropia patients with poor stereopsis. *Investigative Ophthalmology and Visual Science, 40 (12),* 3031-3036

Marquardt, D.W. (1963). An algorithm for least-squares estimation of non-linear parameters. *Journal of the Society for Industrial and Applied Mathematics, 11 (2),* 431-441

McKee, S. P. (1981). A local mechanism for differential velocity detection. *Vision Research, 21,* 491-500

Nefs, H.T, O'Hare, L., & Harris, J.M (2009). Individual differences reveal two independent motion-in-depth mechanisms. *Journal of Vision, 9*(8):627

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10,* 437-442

Rashbass, C. & Westheimer, G. (1961). Disjunctive eye movements. *Journal of Physiology, 159,* 339-360

Shioiri S, , Nakajima T , Kakehi D, Yaguchi H (2008). Differences in temporal frequency tuning between the two binocular mechanisms for seeing motion in depth. *Journal of the Optical Society of America A: Optics, Image Science and Vision, 25 (7),* 1574-1585

Shioiri, S., Saisho, H. & Yaguchi, H. (2000). Motion in depth based on inter-ocular velocity differences. *Vision Research, 40,* 2565-2572

Stevenson, S.B., Cormack, L.K., & Schor, C.M. (1994). The effect of stimulus contrast and interocular correlation on disparity vergence. *Vision Research, 34 (3),* 383-396

Wallace, J.M. & Mamassian, P. (2003). The efficiency of speed discrimination for coherent and transparant motion. *Vision Research, 43,* 2795-2810

Wallace, J.M. & Mamassian, P. (2004). The efficiency of depth discrimination for non-transparant and transparant stereoscopic surfaces. *Vision Research, 44,* 2253-2267

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