PERGAMON

Vision Research 39 (1999) 2335–2347

Vision Research

First and second-order contributions to surface interpolation[1]

Laurie M. Wilcox *

Department of Psychology, York University, 4700 Keele Street, Toronto, Ont. M3J 1P3, Canada

Received 7 January 1998; received in revised form 3 June 1998

Abstract

Comparisons of 1st- and 2nd-order stereopsis have typically employed isolated, or local, narrow-band targets. While these experiments have revealed a great deal about the distinction between these two types of processing, such stimuli are rare in the natural environment. Instead, local disparity signals are more likely to be part of extended surfaces that vary smoothly in depth. The aim of the experiments presented here is to determine the relative contribution of 1st- and 2nd-order stereopsis to the perception of depth-modulated surfaces. Stereothresholds were measured under a range of conditions designed to isolate either 1st- or 2nd-order processing. The results demonstrate that while 2nd-order stereopsis provides local depth estimates for individual texture elements, 1st-order processing is essential to the global interpolation of those estimates across surfaces. © 1999 Elsevier Science Ltd. All rights reserved.

Keywords: Surface interpolation; First-order; Second-order; Local; Global

1. Introduction

Recent experiments have demonstrated that like motion and texture perception, stereoscopic vision involves at least two types of processing: 1st- and 2nd-order[2]. First-order stereopsis is the conventional kind that relies on the spatial frequency content of the stimulus (among others: Julesz, 1971; Julesz & Miller, 1975; Schor & Wood, 1983; Badcock & Schor, 1985; Smallman & MacLeod, 1994; Cormack, Stevenson & Landers, 1997). Second-order stereopsis does not depend on the spatial frequency content, but on the overall scale of the stimulus (Wilcox & Hess, 1995). A number of experiments have demonstrated that the 1st-order system seems to be used by default for fine depth judgements, while the 2nd-order system serves as a back-up system and dominates under two specific circumstances: (1) when the 1st-order disparity signal is ambiguous, or unreliable (Hess & Wilcox, 1994; Wilcox & Hess, 1996; Kovács & Fehér, 1997; Wilcox & Hess, 1997); and (2) when the disparity is large relative to the size of the object[3] (Wilcox & Hess, 1995).

In all but the most recent experiments, studies of 2nd-order stereopsis have used local, isolated narrow-bandwidth stimuli. Although there is good reason to use such stimuli, it is clear that the objects in our natural environment consist of surfaces that usually vary smoothly in depth (Marr, 1982). Thus the aim of the experiments presented here is to determine the relative contributions of 1st- and 2nd-order processing under more global conditions where multiple elements define depth-modulated surfaces.

One of the most consistent findings of previous investigations of 1st- and 2nd-order processing is the dominant role of 1st-order stereopsis under conventional viewing conditions. As noted by Wilcox and Hess

* Fax: +1-416-7365814.

E-mail address: lwilcox@jorku.ca (L.M. Wilcox)

[1] Portions of this work were presented at the Association for Research in Vision and Ophthalmology, 1997.

[2] The terms 1st- and 2nd-order are used here to be consistent with the existing literature concerning 2nd-order motion. Other names that have been applied to these types of processing include Fourier versus non-Fourier (Lin & Wilson, 1995) and linear versus non-linear (Wilcox & Hess, 1995, 1996, 1997).

[3] While it is tempting to argue that the 1st-order/2nd-order distinction described here is simply a restatement of Ogle's (1953) distinction between patent and qualitative stereopsis, this is not the case. The critical difference is that there is no *fixed* disparity above which disparity will be perceived via 2nd-order processing. Instead, the size of the disparity required for 2nd-order processing scales with the size of the stimulus.

PII: S0042-6989(98)00261-2

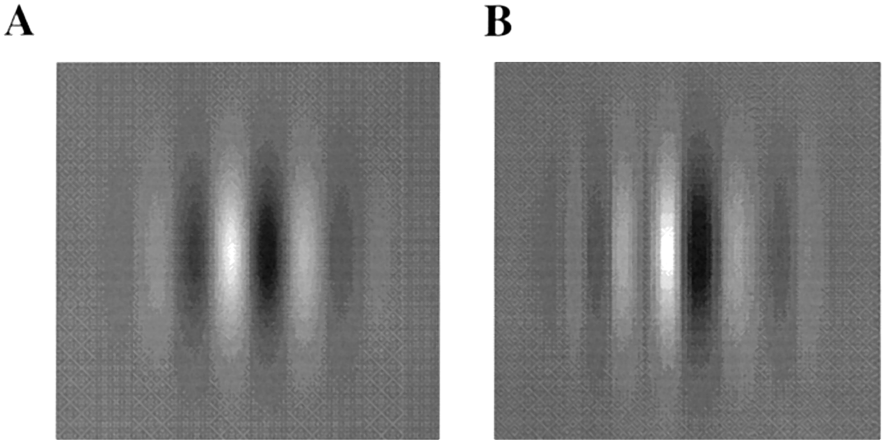

Fig. 1. Two types of stimuli were used to define the modulated surfaces in depth: Gabor patches (A) and Uncorrelated Noise patches (B).

(1996) it appears that whenever a reliable 1st-order disparity signal is present, it is used by default to define perceived depth. In order to force the stereoscopic system to use 2nd-order information the 1st-order information must either be removed or rendered ambiguous. Given the predominance and potential importance of surfaces in the natural environment it would seem reasonable that interpolation of depth estimates from individual elements to form surfaces might be more closely tied to 1st-order processing. However, one could also argue from the existing literature that 1st-order stereopsis specializes in providing high precision depth estimates for fine spatial targets. The relatively coarse nature of the 2nd-order system would make it well-suited to interpolate depth estimates across elements to define a smooth surface.

In the following series of experiments, the 1st- and 2nd-order disparity information were separately manipulated to determine which signal is primarily responsible for the precept of depth in disparity-defined surfaces. The results are consistent in supporting the role of 1st-order processing; while 2nd-order stereopsis can provide depth estimates for individual texture elements, it seems that 1st-order processing is responsible for the smooth interpolation of these estimates across surfaces.

2. General methods

2.1. Subjects and apparatus

For each experiment, extensive measurements were obtained using three experienced subjects. All subjects wore their prescribed optical correction and had excellent stereopsis. All stimuli were presented on a NEC XP17 screen (calibrated using a UDT photometer), with a pixel size of 1.72 min and a frame rate of 160 Hz. A Cambridge Research System, VSG2/3F, graphics card was used to generate and display the stimuli. The mean luminance of the display was approximately 100 cd/m^2, and when viewed through the shutter glasses was reduced to near 49 cd/m^2.

Stereoscopic depth was achieved using Display Tech liquid crystal shutters mounted in trial frames[4]. A ± 10 v signal, supplied via a digital to analogue port, controlled the state of the shutters and was synchronised with the onset of each frame of the display. The stimuli for each eye were presented on alternate frames at a rate of 160 Hz (80 Hz per eye). The reference stimuli were presented with zero disparity on all trials, while the target elements viewed by the two eyes were offset in equal and opposite directions by the required amount.

2.2. Stimulus

Two types of stimulus elements were used to measure stereoacuity: Gabor and 1 D noise patches (see Fig. 1A, B). The Gabor stimuli were generated conventionally by multiplying a vertical sinusoid in sine-phase by a 2 D Gaussian envelope. The equation used to represent this stimulus is:

$$L(x, y) = A^{*}\exp\left(\frac{-((x-x_0)^2+y^2)}{2\sigma^2}\right)(\sin 2\pi f x) + L_0 \quad (1)$$

where f represents spatial frequency of the carrier, L_0 is the mean luminance, x_0 represents the disparity offset

[4] The liquid crystal shutters allow a very fast alternation rate, which can be faster than the decay time of a monitor's phosphor(s). In some situations (e.g. high contrasts) this results in cross-talk, or leakage, between the two eyes views. This problem was avoided ensuring that the stimulus contrasts used in our experiments were well below the threshold for detection of the cross-talk (see also Simmons & Kingdom, 1994).

which was in equal and opposite directions in each eye, and the amplitude term A was chosen so that the Michelson contrast ($2A/L_0$) was 15 dB above the subjects' detection threshold for each condition. The size and spatial frequency of the patch were manipulated by changing the viewing distance to the screen, or by changing the appropriate parameters in the equation used to generate the patches.

The noise stimuli were vertically oriented patches of one dimensional spatial noise multiplied by a 2 D Gaussian envelope. To create the noise pattern a random number was used to select one of 256 grey-levels for each line of the patch. The noise was then multiplied spatially by a two-dimensional Gaussian window. In the experiments described here the noise patches were presented as uncorrelated stereo-pairs, i.e. on each trial the noise patches to be presented to each eye were selected at random from a large set of independently generated patches. This selection procedure combined with the randomization of test disparities required by the method of constants, ensured that there was no consistent disparity information provided by the noise carrier that could signal the direction of the offset of the Gaussian envelope. Similar noise patches were used in a previous study (Wilcox & Hess, 1996) and have been found to effectively isolate 2nd-order processing for local patches.

The stimulus elements described above were used in both local and global test conditions. In the local condition two vertically aligned patches, separated by approximately the diameter of the patches, were presented simultaneously. The upper patch provided a zero disparity reference plane, while the lower patch was shifted in depth relative to that plane. On each trial the horizontal position of the reference patch was displaced quasi-randomly by ± 2–8 min to ensure that the subjects could not use any monocular alignment cues provided by the carrier gratings. In the global condition, many elements were presented simultaneously and used to define a surface in depth. Except where indicated otherwise the density of the elements was fixed at 50% so that on each trial half of the available positions were filled. The disparity modulation was held constant at 1.0 cycle per screen so that one full cycle of the modulated surface was visible within the screen boundaries. Prior to every trial, individual texture elements were positioned quasi-randomly and constrained so that there was no overlap. This was achieved simply by storing the occupied locations in an array, and checking that array when placing each subsequent patch. The surfaces subtended approximately $17 \times 14°$. Details of the procedures used to assess stereothresholds are provided below.

2.3. Procedure

In all experiments reported here stereothresholds were measured using the method of constant stimuli, and a set of 11 test values. The range of test values was chosen individually for each stimulus condition to bracket the point at which the perceived location of the target stimulus changed from being in front of to behind the reference patch. The stimuli were visible for approximately 200 ms. Within a single run each disparity was tested 20 times in random order and the stereothreshold was derived from the resulting psychometric function by fitting the error function (cumulative normal), ERF(x), of the form:

$$P(x) = A(0.5 + 0.5\mathrm{ERF}((x - B)/(\sqrt{2.0}C))) \quad (2)$$

where A is the number of presentations per stimulus condition, B is the offset of the function relative to zero, and C is the standard deviation of the assumed underlying, normally distributed error function. This standard deviation parameter is the measure of the stereothreshold; as it increases stereoacuity deteriorates. Each datum represents the average of at least three such estimates from which the standard error of the mean was derived. Two experimental protocols were used, one- and two-interval forced choice.

2.4. 1-IFC

2.4.1. Single reference

Two stimuli were aligned vertically and separated by the diameter of the patch. Subjects indicated whether the bottom patch was in front of, or behind the zero disparity reference patch.

2.4.2. Modulated surface (Experiment 1 only)

A horizontally-oriented, sinusoidally-modulated surface covered the whole display area and a fixation spot (10 min diameter) was located in the centre of the display at the maximum of a peak or a trough. The subjects were asked to indicate if the fixation spot was on a peak or a trough.

2.5. 2-IFC

In the remaining test conditions a horizontally-oriented, sinusoidally-modulated surface and a 2-IFC paradigm were used to measure stereothresholds. The subjects were presented with two 200 ms intervals both of which contained a field of elements whose positions varied in depth. In one of the intervals the depth variation was systematic and defined a smoothly modulating surface, in the other interval the elements were randomly distributed in depth, with the minimum and maximum disparity values matched to the current test disparity. This constraint ensured that the random depth interval could not be identified by the presence of elements with extreme depth values. In addition, because both crossed and uncrossed disparities were tested within a run the phase of the modulated surface randomly alternated by ±180°. Note that the insets depicting the surface in subsequent figures are for discrimination purposes only and do not accurately represent the characteristics of the modulated surfaces. The observer's task was to chose the interval that contained the surface.

Except where noted otherwise, contrast thresholds were measured prior to testing for all conditions. Subsequently, the test contrast was set at 15 dB above threshold. The method of adjustment was used with a randomized starting point to obtain seven binocular threshold estimates, which were then averaged. When assessing contrast thresholds, the contrast of the composite stimulus was varied, and subjects indicated the point at which it was just detectable.

3. Results and discussion

3.1. Experiment I: size versus spatial frequency dependence

Numerous experiments have demonstrated that stereoacuity for 1st-order stimuli depends on spatial frequency; performance improves with increasing frequency over a broad range of spatial frequencies[5]. In contrast, it has been demonstrated that 2nd-order stereopsis does not depend on the spatial frequency content but on the overall scale of the stimulus (Wilcox & Hess, 1995, 1996). Gabor patches have been used to examine 1st- and 2nd-order contributions to stereopsis because they contain both 1st- and 2nd-order disparity information, which can be separately manipulated, i.e. 1st-order processing will be based on the centre frequency of the patch, and so by manipulating the centre frequency while holding the Gaussian envelope size constant it is possible to determine if there is a 1st-order contribution to a given task (Toet & Koenderink, 1988). However, there is no energy at the frequency of the Gaussian envelope in the Fourier transform of the Gabor patch, it only exists as a distribution of frequencies about the carrier spatial frequency[6]. Thus, by fixing the centre frequency of the patch and varying the size of the Gaussian envelope it is possible to determine if there is a 2nd-order contribution to the same task. Note that the ability to separately vary the size and centre frequency of the Gabor makes it quite different from stimuli which appear quite similar such as difference of Gaussians (doGs) where the size and spatial frequency must covary.

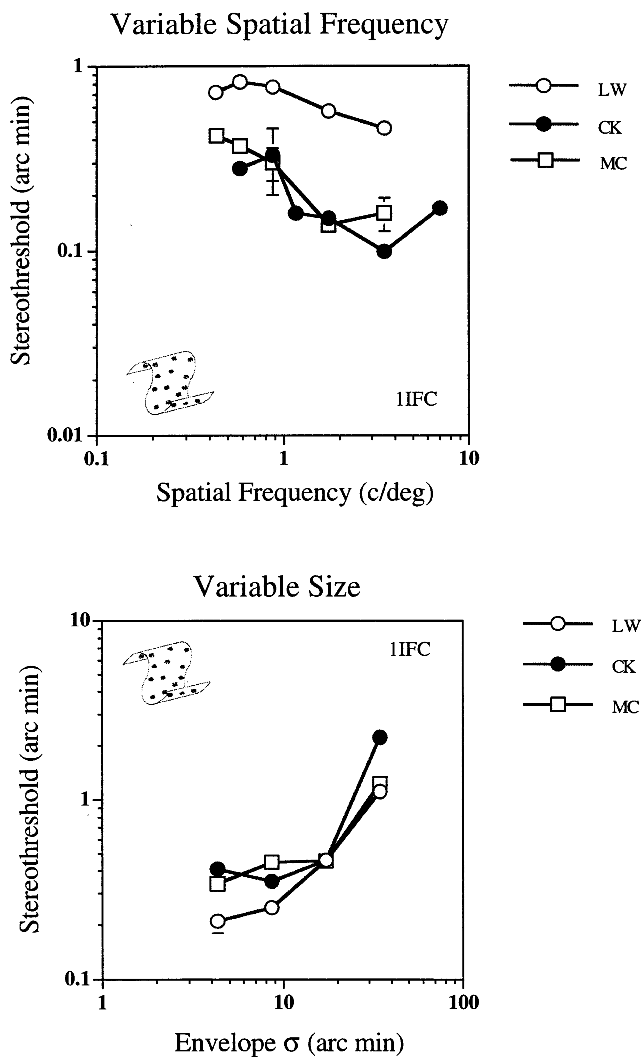

Fig. 2. Stereothresholds are plotted as a function of Gabor carrier spatial frequency (top) and envelope size (bottom) for modulated surfaces, for three subjects. When the centre frequency was varied, σ was fixed at 8.5 (CK) or 17 (LW, MC) min. When the size was varied the centre frequency was held constant at 7 cd. Error bars represent ±1 S.E.M.

[5] While Schor and Wood (1983) have argued that stereoacuity only improves with increasing spatial frequency up to approximately 2 c/deg, there is evidence in the literature that the spatial frequency dependence is more variable. For example, Hess and Wilcox (1994) used Gabor patches to measure stereothresholds and found that performance improved up to 10 c/deg. Similarly, Heckmann and Schor (1989) show continued improvement in stereoacuity well beyond 3 c/deg. The results of Experiment I are also variable; there is evidence of a plateau at 2.5 c/deg for MC which is not present for LW or CK. An alternative interpretation of Schor and Wood's results is that by using stimulus contrasts of 100% they were stimulating a wide range of spatial frequency-selective neurons not only those tuned to the centre frequency of their doGs.

[6] It has been postulated that the contrast envelope of the stimulus is extracted via a non-linear operation such as rectification. This non-linear operation appears to follow a linear filtering stage and occurs prior to binocular combination (Lin & Wilson, 1995; Wilcox & Hess, 1996).

As a first step towards understanding the relative contributions of 1st- and 2nd-order processing to stereoscopically-defined surfaces, stereothresholds were measured for modulated surfaces as a function of size and centre frequency of the Gabor elements that defined the surface, using the 1-IFC procedure described above. If only one of the two types of processing contribute to the perception of these depth defined surfaces then the dependence of stereothresholds on the other should be relatively flat. When the envelope σ was fixed at 8.5 (CK) or 17 (LW, MC) min the centre frequency was varied from 0.6 to 3.5 c/deg (CK was tested at an additional frequency of 7 c/deg). When the centre frequency was fixed at 3.5 c/deg, σ ranged from 4 to 34 min. The effect of both spatial frequency and envelope size on stereothresholds can been seen in Fig. 2.

All three subjects exhibited an effect of both spatial frequency and envelope scale. Performance improved with increasing spatial frequency and degraded with increasing scale. The average slopes were -0.37 and 0.71 for the spatial frequency and envelope size conditions, respectively. While it would be useful to evaluate the difference in slopes using regression analyses it is not valid to test this particular comparison statistically because the x-axis units are different (c/deg vs. width in arc min). Therefore, while the results of this experiment suggest that both 1st- and 2nd-order processing may play a role in the perception of depth-modulated surfaces, they cannot be used to decide upon the relative contribution of each.

To ensure that this pattern of results was not due to the gradual change in disparity in these modulated surfaces, stereothresholds were also measured for each subject as a function of spatial frequency and envelope scale for planar surfaces. In this control experiment the viewing area was divided in half horizontally and on each trial the elements on a side were displaced by the same amount creating a step edge in depth. The subjects' task was to indicated whether the surface on the right was in front or behind the one on the left. The Gabor elements used to define the surfaces were identical to those described above and thresholds were measured using the 1-IFC paradigm and the method of constant stimuli. In the variable spatial frequency condition the envelope σ was held constant at either 8.5 (CK) or 17 (LW, MC) min while in the variable size condition the centre frequency was fixed at 7 c/deg. Comparison of Fig. 2 with Fig. 3 confirms that the subjects showed a similar pattern of results for the two surfaces types with average slopes of -0.28 and 0.90 for the spatial frequency and envelope size conditions, respectively. Thus it seems unlikely that the pattern of results obtained in Experiment I can be attributed to the type of modulated surface using to assess stereothresholds.

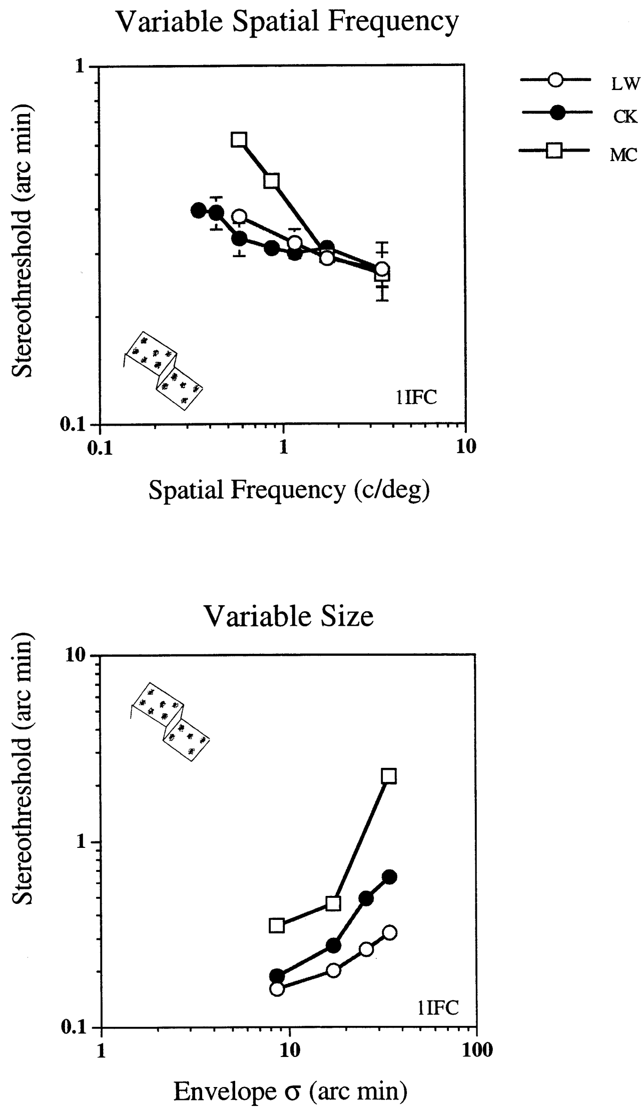

Fig. 3. Stereothresholds are plotted as a function of Gabor carrier spatial frequency (top) and envelope size (bottom) for planar surfaces, for all three subjects. The stimulus parameters are as described for Fig. 2. Error bars represent ± 1 S.E.M.

3.2. Experiment II: stereothresholds for 2nd-order surfaces defined by uncorrelated noise patches

The ideal means of determining the relative contribution of 1st- and 2nd-order processing to surface perception would be to examine their contribution in isolation. While it is not feasible to generate a stimulus with no 2nd-order disparity information, it is possible to isolate 2nd-order processing using Gaussian-windowed noise patches. When different (randomly generated) noise patches are presented stereoscopically the only reliable disparity information available is that provided by the contrast envelope. Previous experiments have shown that the stereoscopic system is able to extract a 2nd-order disparity signal from these stimuli (Wilcox & Hess, 1996). Although performance is approximately a factor of 10 worse than that obtained for correlated noise patches, subjects are consistently able to make reliable depth judgements. In the following experiment stereothresholds were measured using uncorrelated, 1 D noise patches (see Fig. 1B) which defined a modulated surface in depth. For individual elements $\sigma = 8.5$ min and Michelson contrast was fixed at 50%. For comparison, stereothresholds were also measured for identical isolated noise patches. In the latter case, two vertically aligned patches were presented; the upper patch provided a zero-disparity reference and the lower patch was displaced in depth relative to this reference plane. The stimuli were identical to those used in the surface condition, and were separated laterally by an amount equal to the distance between a peak and a trough of the modulated surface. The psychometric functions obtained using the modulated surfaces (2-IFC) and isolated patches (1-IFC) are shown in Fig. 4 for each of the three subjects. The r^2 values obtained when the error function was fit to the local two-element psychometric functions shown in Fig. 4 was 0.99 for all three subjects.

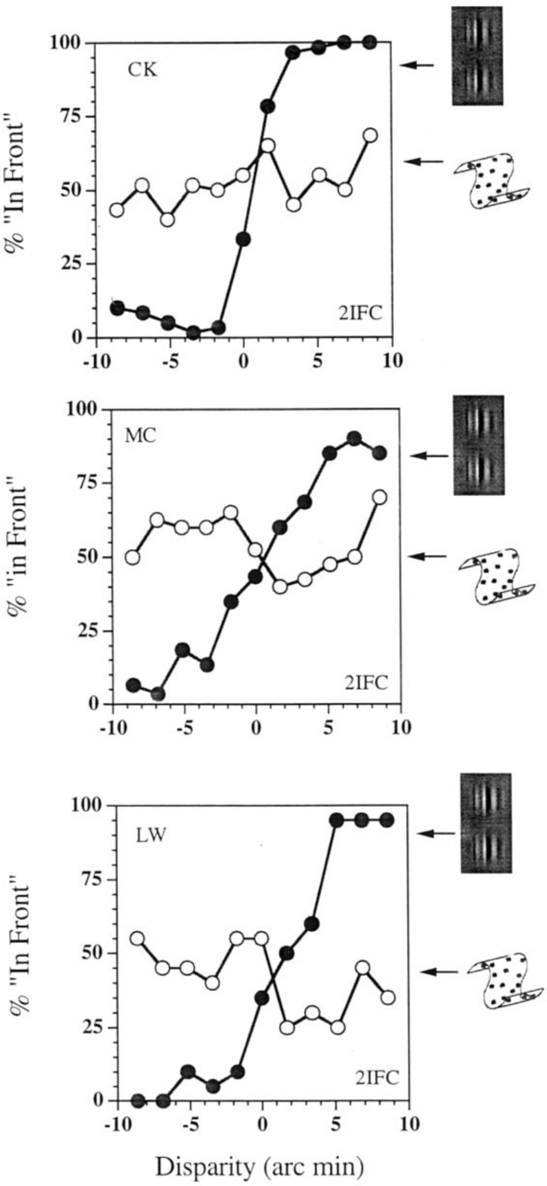

Fig. 4. Psychometric functions are shown here for three subjects using uncorrelated noise patches ($\sigma = 8.5$ min, Michelson contrast = 50%). Two conditions are compared here: (1) two isolated noise patches were presented (●); or (2) multiple elements defined a modulated surface in depth (○). Each data point represents the average of 60 observations.

Stereothresholds for isolated noise patches was identical to those obtained in previous experiments. However, when the uncorrelated noise patches were used to define a surface no depth was perceived. All three subjects reported that in the surface condition there was no sense of any consistent variation in depth; the surface was indistinguishable from the random depth field. Note that this pattern of results was immune to variations in the disparity range tested (additional ranges with upper limits of 0.5–15 arc min were also tested), and to reduction of the number of cycles per screen by half. These results are consistent with data presented by Ziegler and Hess (1997, 1998) who also used uncorrelated noise stimuli to define modulated surfaces, and who reported no evidence of a consistent depth precept. In their experiments Ziegler and Hess tested an impressive range of stimulus parameters to demonstrate convincingly that the task (stereoacuity) was impossible.

An alternative explanation for our inability to see the surfaces defined by uncorrelated noise elements is that this is an impoverished stimulus which provides a relatively weak disparity signal to the 2nd-order system which is unable to support the percept of a smooth surface. To test this possibility, two 1st-order stimuli which produced performance at a level similar to that obtained using the isolated uncorrelated noise patches were identified. One of these was a 1st-order, low-contrast Gaussian ($\sigma = 8.5$ min), and the other a Gabor patch ($\sigma = 8.5$ min, carrier = 7 cd) which contains both sources of information. If the lack of surface perception for the noise patches was due to a poor signal to noise ratio, then one would expect that it would also be impossible to see the modulated surfaces defined by the two stimuli that were perceptually equally degraded. The isolated patch results are shown in Fig. 5, and clearly are quite similar to those shown in Fig. 4. However, the modulated surface results displayed in Fig. 5 are distinctly different from those in Fig. 4. In this case all subjects described the stimulus as a smooth surface, and had no difficulty completing the task. The associated r^2 values for the error function fit to the data shown in Fig. 5 (except for uncorrelated noise, modulated surface condition, where subjects could not do the task) are displayed in Table 1.

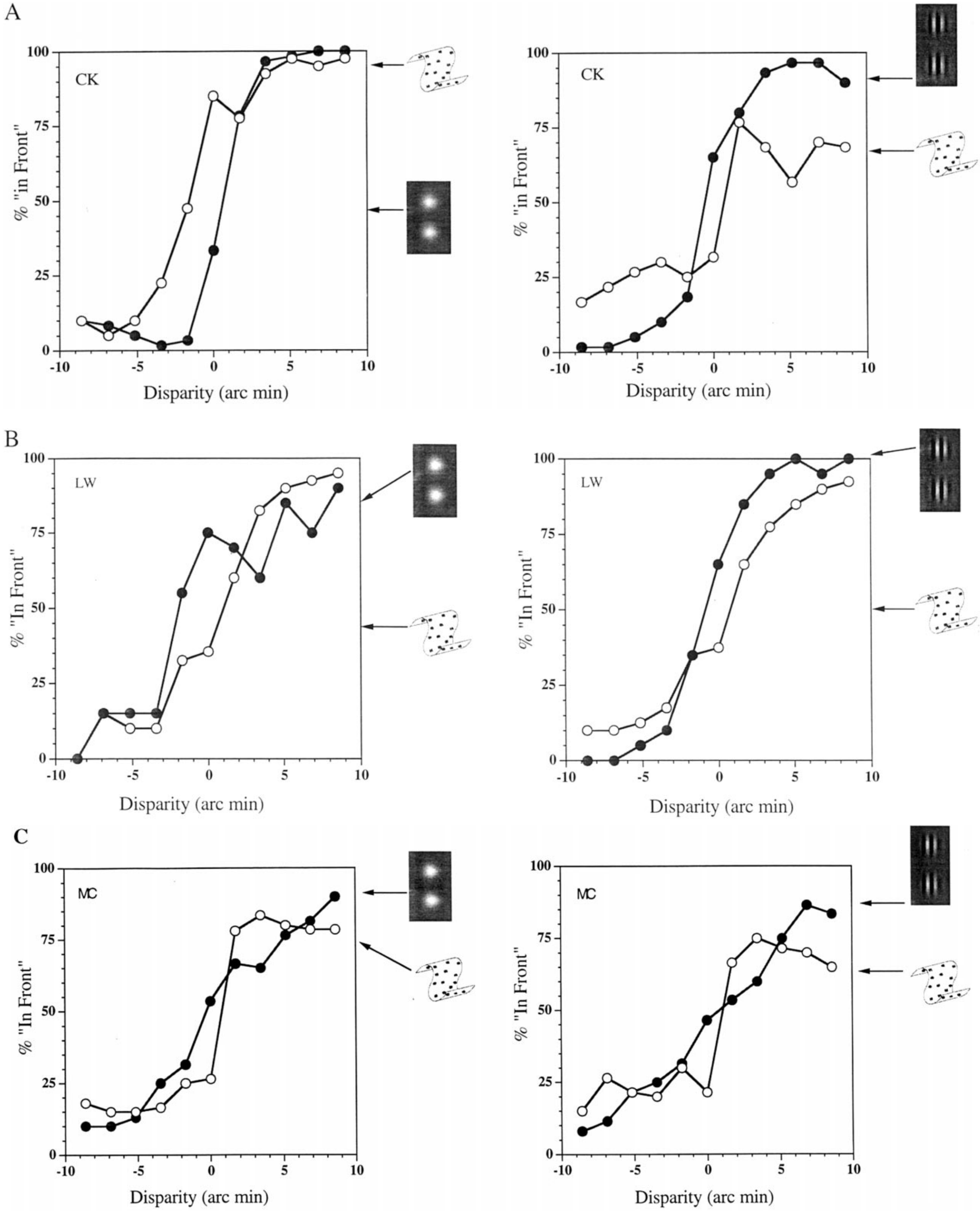

Fig. 5. Psychometric functions are shown here for three subjects (A, B, C) for Gaussian patches (top) and Gabors (bottom). Each graph displays results obtained using local (closed) and global (open) configurations as indicated by the insets. The stimulus size was fixed at $\sigma = 8.5$ min and the centre frequency of the Gabor was 7 cd. The contrasts of the two stimuli were chosen to make performance comparable to performance using local uncorrelated noise patches in the previous study (Fig. 4). Gaussian patches were presented at a Michelson contrast of 5% while the Gabors had 50% contrast. Each data point represents the average of 60 observations.

It is evident from the preceding experiment, and the control studies performed by Ziegler and Hess (1997, 1998) that the failure to see depth when using uncorrelated noise patches to define modulated surfaces is not due to the specific choice of stimulus parameters. It is clear from the two-element condition that the 2nd-order disparity signal does provide a local depth estimate for these uncorrelated patches, but these local estimates are not interpolated when multiple elements define a surface. It could be argued that these results point to the primary role of 1st-order processing in surface interpolation, however the case could also be made that surface interpolation can only occur when *both* sources of disparity information are present. This was not true in the surfaces used here because the 1st-order information was uncorrelated in the two eyes. The following study addresses this issue.

3.3. Experiment III: stereothresholds for 2nd-order surfaces defined by Gabor patches

In the preceding experiment the modulated surfaces were defined by uncorrelated noise patches which isolated 2nd-order processing but added random 1st-order disparity noise. The stimulus used here was designed to provide both 1st- and 2nd-order disparity signals, but allow them to be separately manipulated. Stereothresholds were measured using surfaces defined by Gabor patches ($\sigma = 17$ min; sf $= 3.5$ cd) and the 2-IFC procedure. Three conditions were tested, in the 2nd-order fixed carrier (FC) condition the carrier grating was fixed at zero disparity while the envelope disparity was varied. In the 1st-order fixed-envelope (FE) condition the envelope was fixed at zero disparity while the disparity of the carrier grating was varied. In the normal condition, both components of the stimulus varied simultaneously. The upper graphs in Fig. 6A–C display stereothresholds for three subjects for the three test conditions (FC, FE, normal). The associated r^2 values for the error function fit to the data shown in Fig. 6 (except for the FC narrow range condition where the responses hovered near 50%) are displayed in Table 2.

Table 1
The r^2 values calculated for each of the psychometric functions displayed in Fig. 5[a]

	CK	MC	LW
Gaussian patches			
Two-elements	0.99	0.99	0.96
Surface	0.99	0.95	0.99
Gabor patches			
Two-elements	0.99	0.997	0.99
Surface	0.95	0.94	0.99

[a] These provide an index of the goodness-of-fit of the error function to each data set with the exception of the conditions where the functions were essentially flat.

If 2nd-order processing is an important component of surface interpolation then one would expect similar FC and Normal psychometric functions. If this role is played by 1st-order processing then one would expect very similar functions in the FE and normal test conditions. The data clearly support the latter, and also shows that when forced to used the 2nd-order disparity signal alone, subjects cannot do the task so the psychometric functions remain near 50%. In this respect, the data echo those obtained with uncorrelated noise patches in supporting the failure of 2nd-order disparity signals to support surface perception. However, these data were collected at a fine range of disparity offsets (0.17–0.83). The lower graphs illustrate 2nd-order performance (FC) when the range of test disparities was increased to (1.7–8.6 min). It is evident that 2nd-order disparity information can signal depth-modulation, but the disparities used to define the surface must be large enough to be detectable via 2nd-order processing or in other words, must be at the scale of the contrast envelope.

At the large range of test disparities in the FC condition the difference between the random noise field and the 2nd-order modulated surface is obvious (as evidenced by the smooth psychometric functions in Fig. 6). However, subjects noted that the surface appeared to be different than that seen in the 1st-order conditions. Specifically the 2nd-order surface was not smooth, but varied in a step-like manner, consistent with the abrupt changes in depth that occurred at each new row of elements. In contrast, in the 1st-order condition the surface appeared to vary smoothly. To quantify these subjective observations, stereothresholds were measured using two subjects, one experienced and one naive (who did not participate in any of the other studies presented here), and the 2-IFC paradigm described above using normal and 2nd-order (FC) texture elements. The stimuli and procedures were identical to those described for the preceding study, except that subjects were told to identify the interval that contained the step-like pattern. In the normal condition one interval contained a random depth field while the other contained a smooth surface. The correct response in this case would be the random depth field. If the subjective reports were accurate, then in the 2nd-order condition the stepped surface should provide a more salient step-like pattern than the random field. Therefore in this instance subjects should choose the interval containing the surface, not the random field, resulting in a mirror-reversed psychometric function. If the 2nd-order surface is perceived as smooth, then the psychometric functions obtained in the two conditions should be identical. The psychometric functions are depicted in Fig. 7 for both subjects and both test conditions.

As predicted, the psychometric functions are reversed and as such provide empirical support for the qualitative impression that the 2nd-order surfaces appeared to be stepped rather than smoothly interpolated. The re-

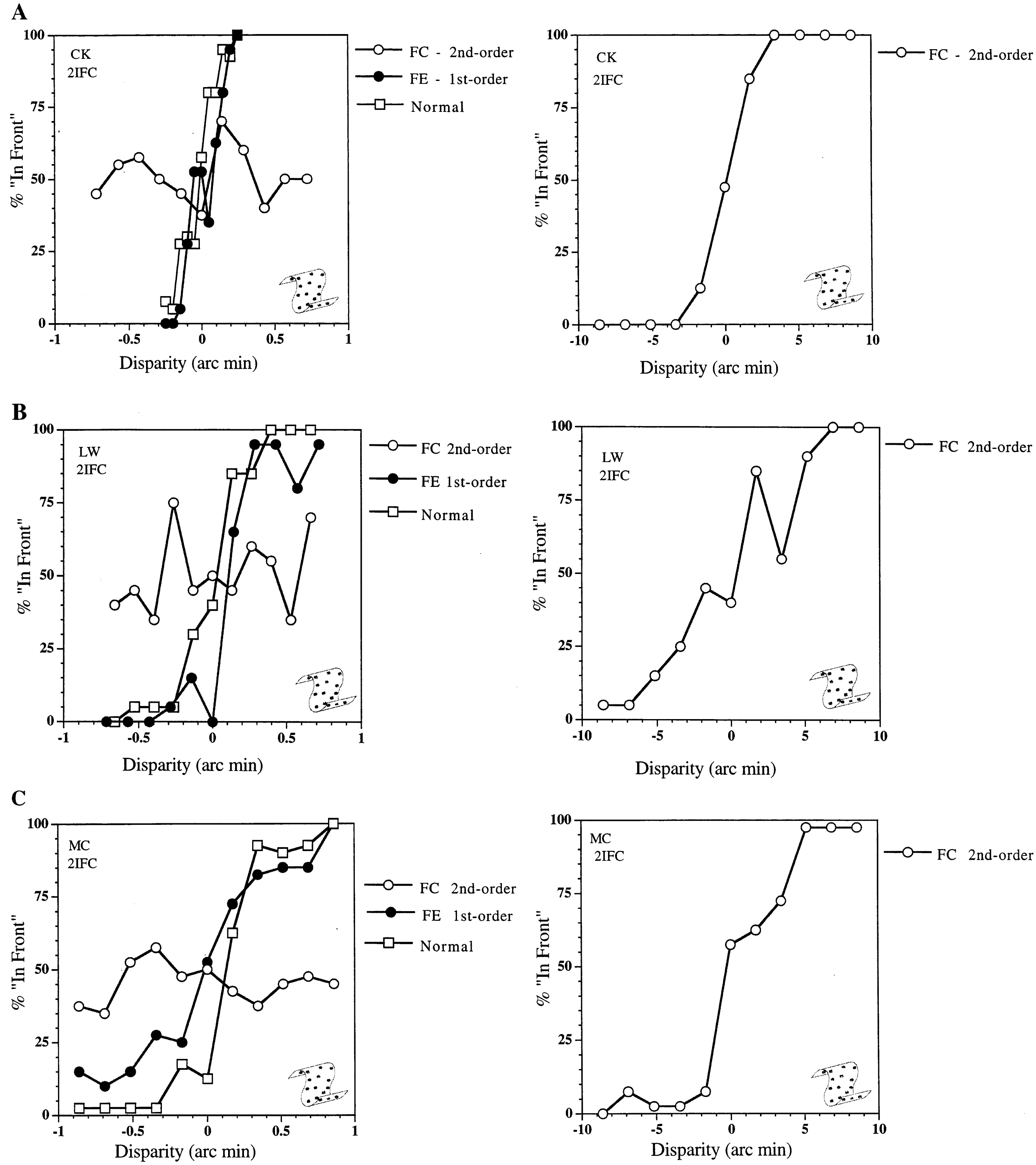

Fig. 6. Psychometric functions are shown here for three subjects (A, B, C) for Gabor patches ($\sigma = 17$ min; sf = 3.5 cd). Three conditions were tested: (1) normal (□); (2) the carrier was fixed at zero disparity (○); and (3) the envelope was fixed at zero disparity (●). In the upper panel both conditions were tested using a narrow range of disparities (0.17–0.86 min), in the lower panel the fixed carrier condition was retested using a larger range of disparities (1.7–8.6 min). Each data point represents the average of 60 observations.

sults of this set of experiments suggest that although depth discontinuities are perceived in the 2nd-order surface, interpolation does not occur, and so the surface does not appear smooth. Apparently in the 1st-order condition the interpolation process does occur, for a strong sense of an extended surface is obtained. These data, considered with the results of Experiment II, provide strong support for the key role played by

Table 2
The r^2 values calculated for each of the psychometric functions displayed in Fig. 6[a]

	CK	MC	LW
FC	0.99	0.98	0.97
FE	0.99	0.95	0.99
Normal	0.99	0.997	0.99

[a] These provide an index of the goodness-of-fit of the error function to each data set with the exception of the conditions where the functions were essentially flat.

1st-order processing in the accurate perception of smoothly modulated surfaces in depth.

There are two aspects of the results displayed in Fig. 6 that support the position that the 2nd-order test conditions described here effectively isolate 2nd-order processing. First, the small range of offsets tested when the carrier was fixed at zero disparity were selected to be at the scale of the carrier. All subjects were able to see reliable depth variation using a normal Gabor patch of the same centre frequency presented at that range of offsets. If there were 1st-order disparity cues available in the 2nd-order condition, then performance should have been good in the small-offset condition and random in the large-offset condition: exactly the opposite took place. Second, the large-offset condition included test disparities beyond the half-cycle limit of the carrier grating. Therefore, if subjects were matching the bars of the grating they should have experienced depth reversals, resulting in periodic psychometric functions. The relatively smooth psychometric functions shown in Fig. 6 argue convincingly against this proposal.

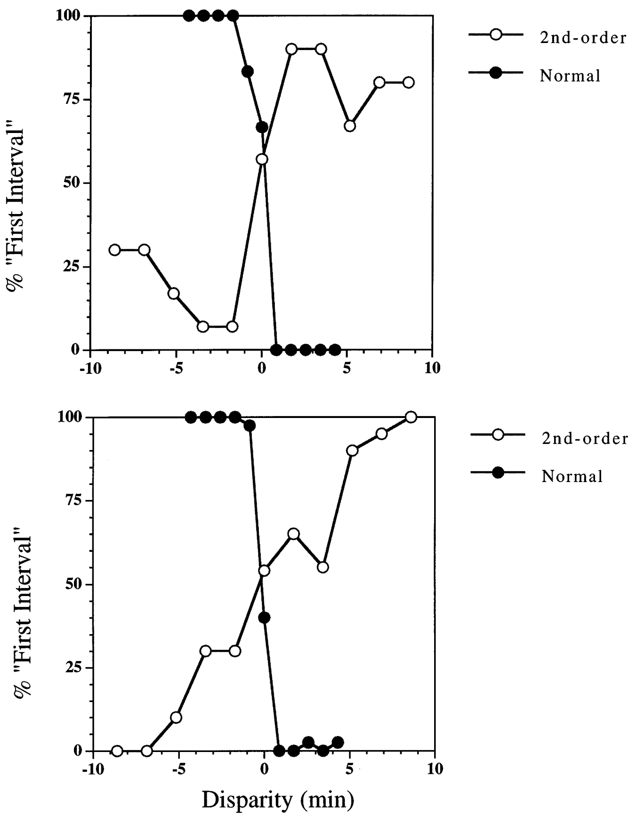

Fig. 7. Psychometric functions are shown here for two subjects (top and bottom) for two conditions, normal (●) and 2nd-order (○). Gabor elements that defined the surfaces had $\sigma = 17$ min and a spatial frequency of 3.5 cd. Each data point represents the average of 40 observations.

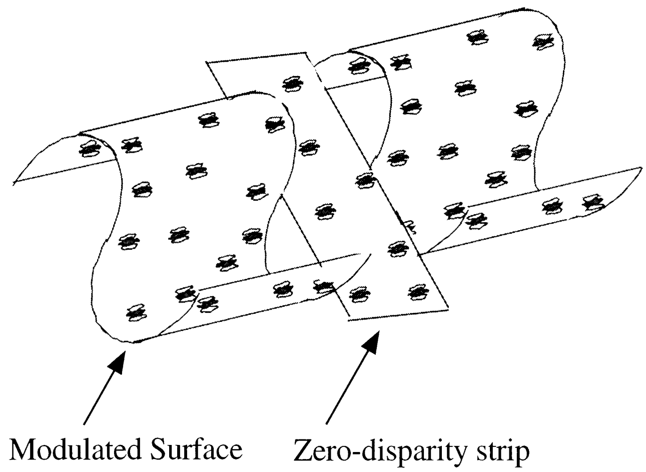

Fig. 8. A schematic illustration of the stimulus used to assess the salience of illusory boundaries created by interrupting a modulated surface with a strip of zero disparity elements. Subjects were asked to indicate (1) if a dark ruler at the bottom of the display was aligned with a boundary or not and (2) whether the perceived boundary was strong or weak. The surfaces were defined using Gabor patches ($\sigma = 8.5$ min, sf = 3.5 cd) and the modulation frequency was fixed at 1 cycle per screen.

3.4. Experiment IV: is surface interpolation performed via 1st-order processing?

The experiments reported thus far point to 1st-order processing as the most likely candidate for surface interpolation. Each of these studies assessed the subjects' ability to accurately perceive the depth modulation of surfaces, in each case the percept of a smoothly modulating surface required the successful interpolation of depth estimates. In this final experiment, the interpolation process, and the role played by 1st-order processing is examined more closely.

When the depth modulation that describes a surface, is interrupted by a strip of zero disparity elements one perceives a clear percept of an illusory boundary between the flat and modulated regions (for a schematic illustration see Fig. 8). This boundary is defined solely by depth discontinuities, and appears continuous in spite of the fact that there are relatively large areas of mean luminance in both the flat strip and the modulating surface.

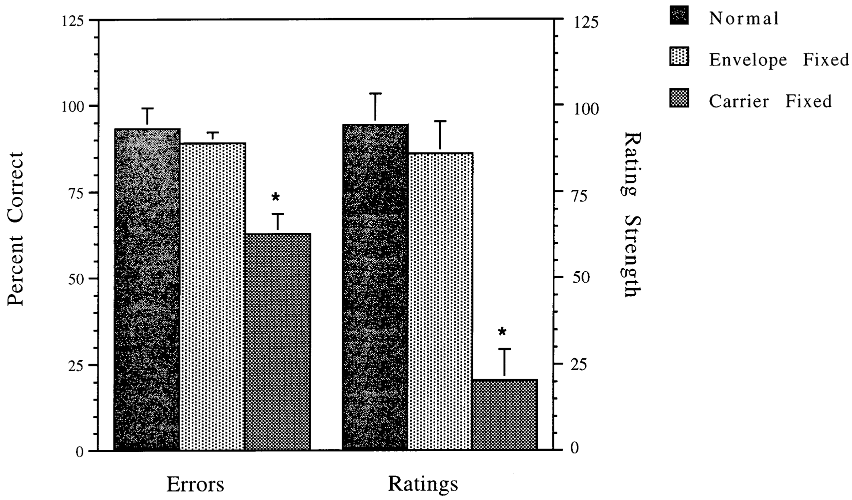

Fig. 9. Averaged results from all three subjects are shown here for the illusory boundary detection and rating task for normal (dark grey), fixed envelope (dotted) and fixed carrier (light grey) conditions. The vertical axis on the left corresponds to the error results shown on the left, while the right-hand axis corresponds to the rating strength results on the right. Error bars represent ± 1 S.E.M. and asterisks designate data that is significantly different from the normal condition.

The percept of the illusory boundaries at the edge of zero-disparity strip can be attributed to the interruption of the interpolation process and so provides a means of evaluating the source of surface interpolation. If 1st-order processing is responsible for interpolation of depth estimates across elements, then the percept of the illusory boundary should be equally strong for normal surfaces with both 1st- and 2nd-order disparity information and surfaces defined only by 1st-order disparity information. This prediction was tested by measuring the detectability and perceived strength of an illusory boundary under three conditions: normal, carrier fixed at zero disparity (PC), and envelope fixed at zero disparity (PE). The modulated surface was identical to that described in Experiment III, and was defined by Gabor patches with a centre frequency of 3.5 cd and $\sigma = 8.5$ min. The modulation frequency was held constant at 1 cycle per screen, the peak disparity of the surface was approximately 4 min, and the width of the zero disparity strip was slightly less than 1°. The location of the strip was varied randomly on each trial but it was constrained not to fall on the edge of the display. On each trial a dark bar appeared along the bottom edge of the screen with a height and width of 3.0 and 0.5°, respectively. On half the trials the bar was aligned with an edge of the zero disparity strip. Following a 250 ms exposure to the stimulus a subject made two responses: the first was to indicate if the ruler was aligned with an edge of the strip, and the second whether the percept of the illusory boundaries was strong or weak. The three conditions were randomly interleaved in the same test block, with at least 100 presentations of each. The average results for all three subjects are shown in Fig. 9.

The illusory boundary was reliably detected in all three conditions, but more errors were made in the 2nd-order only condition (FE). Subjects reported that they were able to identify the depth discontinuity in the 2nd-order condition but that it did not appear as a boundary. This observation is supported by the rating scores where all subjects indicated that the illusory boundary was extremely weak in the 2nd-order condition. In contrast, in the 1st-order (FE) and normal conditions the percept of the boundary was consistently strong making the task trivially easy. Student's directional t-tests support these observations and show that the difference between the normal and FC conditions was statistically significant in both the rating and percent correct data where in both cases $P < 0.01$ ($P =$ 0.003 and 0.004, respectively). The normal and FE conditions, however, were not significantly different, with $P > 0.01$ for both the rating and percent correct tasks ($P = 0.23$ and 0.43, respectively).

When considered in light of the results of the preceding experiments, the similarity of the normal and 1st-order test results, and the poor rating scores for the 2nd-order condition, confirm that 1st-order processing is responsible for surface interpolation. In addition, the fact that subjects could reliably perceive the location of the depth discontinuity in the 2nd-order condition echoes observations made in Experiment III where the depth variation of discrete elements was evident in the absence of interpolation.

4. Discussion

Collectively the results of Experiments II, III and IV show that a reliable 1st-order disparity signal is both necessary and sufficient for perceiving smooth depth-modulated surfaces. However the relatively strong dependence of performance on the stimulus envelope reported in the first experiment suggested that 2nd-order operations do make a significant contribution to surface perception. Another interpretation of these data is that the dependence on envelope size was not due to the change in envelope size per se, but to the concurrent change in the stimulus bandwidth. The centre frequency of the Gabor elements was fixed, therefore, when the size was increased the number of cycles of the carrier in the patch also increased. One consequence of this may have been to introduce 1st-order matching ambiguity which in turn degraded performance. This explanation echoes that provided by Hess and Wilcox (1994) who proposed that increasing the number of cycles of the carrier grating in Gabor patches effectively introduced matching ambiguity which made the 1st-order disparity signal less reliable.

Why was a strong bandwidth dependence only observed when size was varied, and not so evident when spatial frequency was manipulated? Comparison of the octave bandwidths of stimuli in each of the test conditions reveals that much narrower bandwidths were tested in the variable size condition. That is, the bandwidths in the variable size condition ranged from 0.14 to 0.79 octaves, while the narrowest bandwidths in the variable spatial frequency were 0.27 (LW) and 0.55 (CK) octaves. It is possible that the apparently stronger dependence on envelope size reported in Experiment I may be due, at least in part, to the range of bandwidths tested.

The results of these experiments clearly support the primary role of 1st-order stereopsis in interpolating depth estimates across surfaces. The data also suggest that 2nd-order processing is restricted to providing local depth estimates for individual texture elements. This interpretation resolves the initially puzzling result reported in Experiment II, that uncorrelated noise patches provide reliable depth estimates via 2nd-order processing for isolated patches, but not for multi-element surfaces. The 2nd-order depth signal provides the depth estimate for the individual patch, but because the luminance-based disparity information is uncorrelated in the two eyes, the 1st-order process is unable to accurately interpolate the depth estimates. This account also explains the appearance of the modulated surfaces in Experiment III. In the 1st-order condition a smooth surface was clearly visible, while in the 2nd-order condition the surface appeared disconnected and rough. Although it was possible to discriminate the rough surface from the random field because of the regular variation of elements at different depth planes, they did not seem to be connected (or interpolated) to form a surface.

4.1. Relation to previous research

A number of studies have been performed to investigate stereoscopic interpolation, most of which have used broadband stimulus such as dots. Recently, Yang and Blake (1995) modulated strips of binary random-dot noise in depth, using a Gabor profile. Subjects were asked to perform a Vernier task by responding to the relative offset of the peaks of two such strips. Performance was degraded as the number of random dot elements defining the depth was reduced. It is difficult to directly compare their experiments with those reported here since the stimuli, methods and psychophysical task were different, but it is worth noting some interesting differences in the two sets of results.

Yang and Blake (1995) pointed out that the stereoscopic system has a remarkable ability to reconstruct surfaces over disparity discontinuities and the appearance of the modulated surfaces displayed here are entirely consistent with this observation. However, while Yang and Blake (1995) reported an upper interpolation limit of 0.3°, in Experiment I subjects consistently perceived smooth surfaces in the 1st-order test conditions in spite of discontinuous regions as large as 2°. The difference in the size of this range may be attributed to at least two causes. The first of these is the size of the texture elements. Yang and Blake's RDS were defined by dots subtending 2×2 arc min whereas the diameter of the Gabor patches used here was as large as approximately 1°. Further, the RDS patterns generated by Yang & Blake were dense while the texture elements defining the surfaces here were separated by blank regions of variable distance. Thus the difference in the size of the discontinuous region that could be accommodated in the two cases could be attributed to a scaling of the upper limit with stimulus size and spacing.

Another contributor to the difference in the range of interpolation is nature of the discontinuous regions. The discontinuous elements in Yang and Blake's stimuli were set at zero disparity. These zero disparity elements would likely form a second, planar surface, and so create a competing disparity signal. In the experiments reported here the discontinuous region was the mean luminance of the display and contained no texture elements. Presumably the reduce interference in this case would make it easier for subjects to perceive the modulated surfaces.

Mitchison and McKee (1985) (see also Mitchison and McKee (1987a,b)) studied depth interpolation using rows of small isolated dots. They varied the depth of the endmost pairs of dots and assessed the resulting effect on the intervening zero-disparity elements. Their results demonstrate convincingly that depth interpolation occurs and show a compelling effect of inter-dot spacing, i.e. different percepts were obtained depending on whether the spacing was large (> 6–8 arc min) or small (< 6–8 arc min). Large spaces resulted in a percept consistent with discrete dot-by-dot matching, while the percept obtained with small spaces favoured a more global interpolation operation. These results are consistent with the argument posed above, i.e. the large spacing condition would be more likely to activate 2nd-order processing and so produced discrete element matches, as was seen in Experiment III, where 2nd-order surfaces appeared to be formed of discrete elements. The small spacing condition would have been more likely to involve 1st-order processing, and in this case there was evidence of smooth disparity interpolation just as surfaces defined only via 1st-order processing in Experiment III appeared smooth and continuous.

The results reported here further clarify the role of 2nd-order stereopsis in vision in the natural environment. Earlier results suggest that for local stereopsis 1st-order processing is relied upon heavily, and 2nd-order depth signals are used only when the 1st-order processing is unreliable or unavailable. The dominant role of 1st-order processing is also true for the global perception of depth modulated surfaces, and again the coarse depth information provided via 2nd-order processing is available when large disparities are presented and the 1st-order information is degraded.

Acknowledgements

Many thanks to colleagues who commented on earlier versions of this manuscript: David Simmons, Harvey Smallman and Steven Dakin. I would also like to thank two anonymous reviewers for their valuable feedback. This work was supported by the Natural Sciences and Engineering Council of Canada.

References

Badcock, D. R., & Schor, C. M. (1985). Depth-increment detection function for individual spatial channels. *Journal of the Optical Society of America A*, *2*, 1211–1215.

Cormack, L. K., Stevenson, S. B., & Landers, D. D. (1997). Interactions of spatial frequency and unequal monocular contrasts in stereopsis. *Perception*, *26*(9), 1121–1136.

Heckmann, T., & Schor, C. M. (1989). Is edge information for stereoacuity spatially channeled? *Vision Research*, *29*, 593–607.

Hess, R. F., & Wilcox, L. M. (1994). Linear and non-linear contributions to stereopsis. *Vision Research*, *34*, 2431–2438.

Julesz, B. (1971). *Foundations of cyclopean perception*. Chicago: University of Chicago.

Julesz, B., & Miller, J. E. (1975). Independent spatial frequency-tuned channels in binocular fusion and rivalry. *Perception*, *4*, 125–143.

Kovács, I., & Fehér, A. (1997). Non-Fourier information in bandpass noise patterns. *Vision Research*, *37*, 1167–1177.

Lin, L., & Wilson, H. (1995). *Investigative Ophthalmology and Visual Science Supplement*, *36*, 364.

Marr, D. (1982). *Vision*. New York: Freeman.

Mitchison, G. J., & McKee, S. P. (1985). Interpolation in stereoscopic matching. *Nature*, *315*, 402–404t.

Mitchison, G. J., & McKee, S. P. (1987). The resolution of ambiguous stereoscopic matches by interpolation. *Vision Research*, *27*, 285–294.

Mitchison, G. J., & McKee, S. P. (1987). Interpolation and the detection of fine structure in stereoscopic matching. *Vision Research*, *27*, 295–302.

Ogle, K. (1953). Precision and validity of stereoscopic depth perception from double images. *Journal of the Optical Society of America*, *43*, 906–913.

Schor, C. M., & Wood, I. (1983). Disparity range for local stereopsis as a function of luminance spatial frequency. *Vision Research*, *23*, 1649–1654.

Simmons, D., & Kingdom, F. A. A. (1994). Contrast thresholds for stereoscopic depth identification with isoluminant and isochromatic stimuli. *Vision Research*, *34*, 2971–2982.

Smallman, H. S., & MacLeod, D. I. A. (1994). Size-disparity correlation in stereopsis at contrast threshold. *Journal of the Optical Society of America A*, *11*, 2169–2183.

Toet, A., & Koenderink, J. J. (1988). Differential spatial displacement discrimination of Gabor patches. *Vision Research*, *28*, 133–143.

Wilcox, L. M., & Hess, R. F. (1995). Dmax for stereopsis depends on size, not spatial frequency content. *Vision Research*, *36*, 391–399.

Wilcox, L. M., & Hess, R. F. (1996). Is the site of non-linear filtering in stereopsis before or after binocular combination? *Vision Research*, *36*, 391–399.

Wilcox, L. M., & Hess, R. F. (1997). Scale selection for 2nd-order (non-linear) stereopsis. *Vision Research*, *37*, 2981–2992.

Yang, Y., & Blake, R. (1995). On the accuracy of surface reconstruction from disparity interpolation. *Vision Research*, *35*, 949–960.

Ziegler, L., & Hess, R. F. (1997). Linear and non-linear stereoscopic contributions distinguished by task. *Investigative Ophthalmology and Vision Science Supplement*, *38*, 906.

Ziegler, L., & Hess, R. F. (1998). Why is there depth but no shape from uncorrelated micropatterns? *Investigative Ophthalmology and Vision Science* (*Suppl.*), *39*, 615.