Luminance spatial scale facilitates stereoscopic depth segmentation

Frederick A. A. Kingdom, Lynn R. Ziegler, and Robert F. Hess

McGill Vision Research Unit, Room H4-14, 687 Pine Avenue West, Montréal, Québec H3A 1A1, Canada

Received September 6, 2000; accepted October 26, 2000

Are differences in luminance spatial frequency between surfaces that overlap in depth useful for surface segmentation? We examined this question, using a novel stimulus termed a dual-surface disparity grating. The dual-surface grating was made from Gabor micropatterns and consisted of two superimposed sinusoidal disparity gratings of identical disparity-modulation spatial frequency and orientation but of opposite spatial phase. Corrugation amplitude thresholds for discrimination of the orientation of the dual-surface grating were obtained as a function of the difference in Gabor (luminance) spatial frequency between the two surfaces. When the Gabor micropatterns on the two surfaces were identical in spatial frequency, thresholds were very high and in some instances impossible to obtain. However, with as little as a 1-octave difference in spatial frequency between the surfaces, thresholds fell sharply to near-asymptotic levels. The fall in thresholds paralleled a change in the appearance of the stimulus from one of irregular depth to stereo transparency. The most parsimonious explanation for this finding is that the introduction of a between-surface luminance spatial-frequency difference reduces the number of spurious cross-surface binocular matches, thus helping to reveal the three-dimensional structure of the stimulus.

OCIS codes: 330.0330, 330.1400, 330.5510, 330.7310.

1. INTRODUCTION

Stereoscopic vision is an important means of determining the three-dimensional structure of the world around us. Primate stereopsis probably originated amid relatively dense vegetation—a visual diet of many overlapping, densely textured surfaces. In such environments stereopsis is particularly useful for perceptually breaking up a scene into its component structures and surfaces. A good example of this is Fig. 1. Fusion of the two stereo halves reveals branches of the fern largely hidden in the monocular view. Densely textured scenes with overlapping depth surfaces present the visual system with a particularly acute correspondence problem. For every feature in one eye's view there are not only a number of possible matches in the other eye's view of the same surface but also possible matches from the other surfaces.[1–5] It is therefore of special interest to know whether we possess mechanisms that help overcome this correspondence problem.

Relevant to this issue is the question as to what stimulus characteristics help to facilitate stereoscopic segmentation in stimuli such as in Fig. 1. This question has been addressed primarily in the context of the study of stereo transparency, in which a typical paradigm has been to measure the degree of perceived stereo transparency between two surfaces separated in depth (see Ref. 5 for a review). It has been shown, for example, that, when the features on the two surfaces differ in color,[2] motion contrast,[6] or contrast polarity,[4] the perception of stereo transparency appears to be enhanced. In this study we consider the effectiveness of a between-surface difference in luminance spatial frequency, not so much in the perception of stereo transparency as in the identification of the three-dimensional structure of the stimulus.

Much evidence points to the importance of luminance spatial frequency for stereopsis. In a seminal study, Julesz and Miller[7] found that random dot stereograms survived the effects of added uncorrelated noise if signal and noise differed in luminance spatial frequency by 2 or more octaves. Other psychophysical studies have since confirmed the importance of spatial frequency selectivity in stereopsis,[8–12] supported by studies of single-unit recordings in primate cortex.[13] However the importance of luminance spatial frequency for stereo segmentation has received comparatively little attention. Rohaly and Wilson[14] studied the effects of spatial frequency differences in the phenomenon known as disparity averaging.[5,14–17] This is the phenomenon whereby two overlapping stimuli at different depths often appear at a single intermediate depth. They found that two overlapping cosine gratings separated by 112 arc sec required a spatial frequency difference of up to 3.5 octaves (a factor of approximately 11) before they appeared as two transparent layers separated in depth. This suggests that disparity averaging is relatively broadly tuned to luminance spatial frequency. However, in the present study we employed random-element stereograms, and it has been argued that with random-element stereograms correspondence noise rather than disparity averaging is the factor most likely to limit the perception of depth transparency.[5] Therefore the effectiveness of between-surface luminance spatial-frequency differences in revealing the structure of multisurface random-element stimuli has yet to be determined.

To ensure that the subjects' task was to detect structure rather than to rate stereotransparency, we employed an objective technique involving a novel stimulus based on the disparity grating introduced by Tyler.[18] A dispar-

Fig. 1. Fusion by crossed disparity reveals branches of the fem at different depths that are largely camouflaged in the monocular view. The more distant branches will have a slightly higher retinal spatial frequency composition. Could this difference facilitate stereoscopic segmentation?

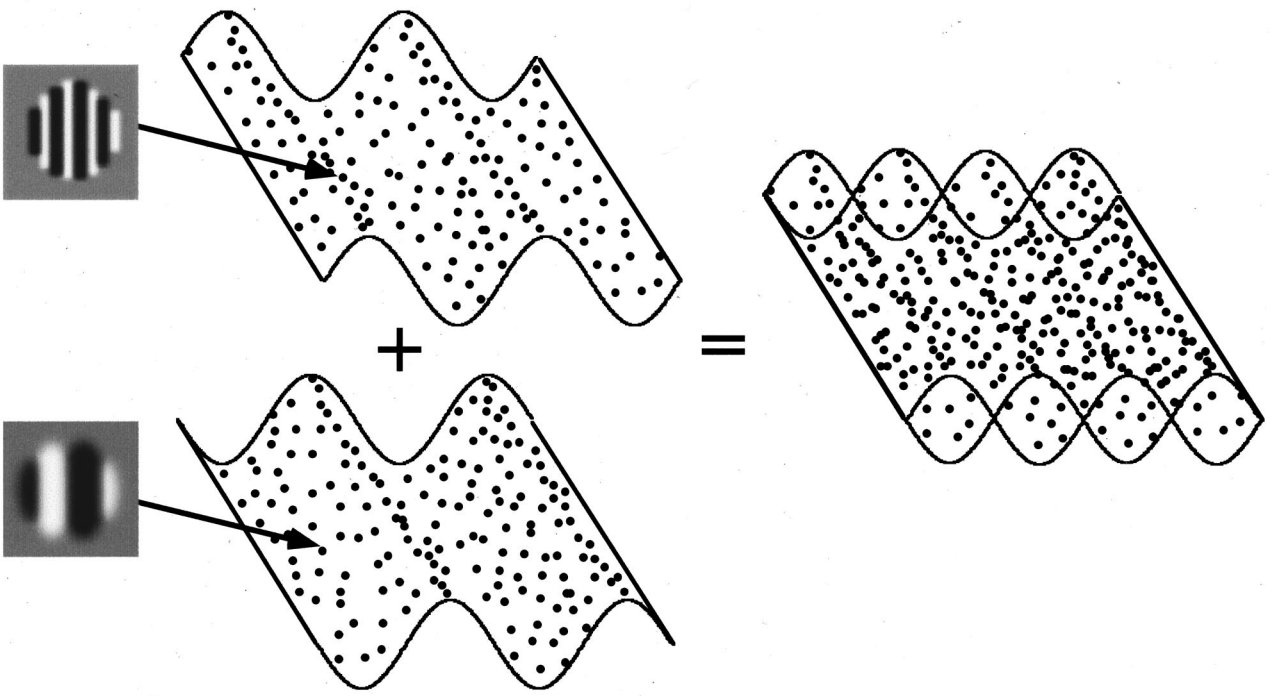

Fig. 2. Method of construction of a dual-surface disparity grating from two types of Gabor micropatterns.

ity grating is a densely textured stimulus modulated in disparity according to a particular waveform, most typically a sinusoid, and appears as a corrugation in depth. Our stimulus consists of two superimposed disparity gratings with the same corrugation orientation and spatial frequency but 180 deg out of phase. We refer to this stimulus as a dual-surface disparity grating. Its method of construction is illustrated in Fig. 2, and two example stereo pairs are shown in Fig. 3. Our dual-surface disparity gratings are composed of dense arrays of randomly positioned Gabor micropatterns whose surfaces can either differ in Gabor spatial frequency, as in Figs. 3(a) (see also Fig. 2), or be the same, as in Fig. 3(b). In all experiments, subjects were required to judge the orientation of the depth corrugations (which could be left or right oblique). Thus we obtained a criterion-free estimate of the ability of the visual system to extract the three-dimensional (3-D) structure of a multisurface stimulus, defined by the same depth distribution of local disparities. As with single-surface disparity gratings, we estimated

the threshold amplitude of disparity modulation, $d_{\min}$, for our dual-surface gratings.

In Fig. 3(a) the two depth corrugations are easily seen, appearing to weave in and out of each other and giving rise to a strong perception of segmentation and transparency. In Fig. 3(b), on the other hand, the stimulus appears lacy and irregular in depth and has no discernible structure or transparency. Our quantitative measurements of $d_{\min}$ parallel these differences in appearance. A brief report of these findings has been given elsewhere.[19]

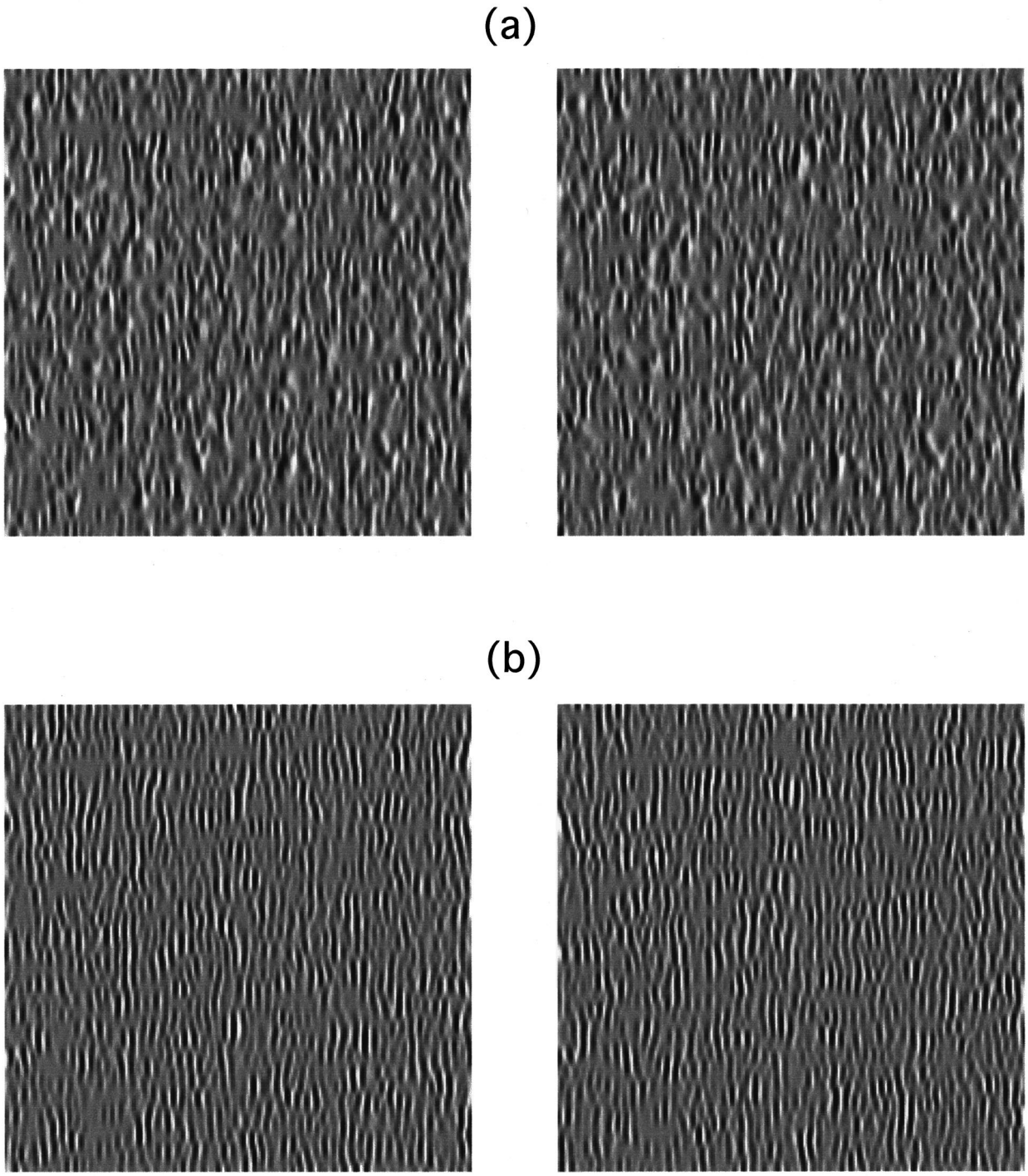

Fig. 3. Example dual-surface disparity gratings. Fusion of the two stereo halves in (a) reveals a 3-D pattern with an obliquely oriented corrugated structure. On closer inspection one can see two interwoven surfaces and an impression of depth transparency. In (a) the micropatterns are of the same size but differ in luminance spatial frequency by 1 octave, or a factor of 2. In (b) the two surfaces are made from the same micropatterns, and the corrugated structure is very difficult to perceive in the fused image. In the main experiment subjects were required to judge the orientation of the corrugations in the stimulus as a function of the difference in Gabor micropattern spatial frequency between the two surfaces (± 45 deg).

2. METHODS

A. Observers

Two of the authors, FK and LZ, acted as subjects. FK was emmetropic, and LZ used his optical correction.

B. Stimuli

1. Generation and Display

The stereograms in all experiments were generated by a Silicon Graphics 02 workstation with screen resolutions of 1024×1280 pixels. In the first experiment they were backprojected onto a large translucent screen by an Electrohome ECP-4100 projector. The screen subtended 50×57 deg at the viewing distance of 114 cm and had a mean luminance of 1.2 cd/m^3. In the remaining experiments the stereo images were displayed on a Sony GDM-20E21 video monitor. The monitor display subtended 28×38 deg at a viewing distance of 57 cm and had a mean luminance of 6.0 cd/m^2 when measured through the stereo glasses. The two stereo half-images were displayed on alternate frames of the display device at 120 Hz and were projected to the two eyes by liquid-crystal shutter glasses (Stereographics, Inc., CrystalEyes) synchronized to the projector frame rate. It is well known that interocular cross talk can occur when liquid-crystal shutters are used to separate stereo half-images. Therefore we used relatively low-contrast stimuli (as measured on the Sony video monitor) to minimize such cross talk ($\leqslant$33.0%). The exception was FK's highest Gabor spatial frequency condition [3.36 cycles per degree (cpd)] in the first experiment, where a contrast of 78% was employed. However, the contrast of this high-spatial-frequency Gabor micropattern (hereafter referred to as Gabor) was substantially attenuated by the translucent screen (thus the need for such a high contrast; see below). There was nevertheless some residual cross talk in our displays in the form of a very-low-contrast "ghost" signal superimposed on the much-higher-contrast "true" signal in each eye. There is good evidence, however, that low-contrast stereoscopic signals have little effect on the perceived depth of superimposed higher-contrast signals[20] and good evidence for a contrast-similarity constraint on stereo matching.[21] We are therefore confident that any cross talk did not significantly affect stereoscopic performance in this study.

2. Gabor Micropatterns

The Gabor micropatterns were generated by use of the function

$$L(x) = M[1 + c \sin(2\pi f x)\exp(-x^2/2\sigma^2)],$$

where M is the mean luminance (see above), c is contrast, f is the spatial frequency, and σ is the space constant of the envelope. The values of the Gabor parameters c, f, and σ will be given with each experiment. The Gabors were all vertically oriented. The carrier modulation was in the sine, or "odd," phase to ensure that the Gabor's space-average luminance was the same as that of the background.

3. Disparity Gratings

Disparity gratings, whether single-surface or dual-surface, contained 1200 Gabors, randomly positioned on the screen. The disparities of the Gabors were defined with subpixel accuracy with a resolution of 128th of a pixel, or 1.25 arc sec or less (see Ref. 22 for details). The disparity was sinusoidally modulated. The orientation of disparity modulation was left or right oblique ($-45°$ or $+45°$). The phase of disparity modulation was randomized on each trial. When Gabors fell on top of each other, their amplitudes but not mean luminances were added. At the density of Gabors employed in all experiments the number of instances in which the chance addition of Gabors resulted in luminance outside the normal range was negligible.

C. Procedure

1. Contrast Matching

Because we employed Gabors with different spatial frequencies, we decided to equate them for apparent contrast. This was particularly important for the first experiment that employed a translucent screen because the screen significantly attenuated high-luminance spatial frequencies. Equating the Gabors for apparent contrast is arguably preferable to equating them for detectability, because the function that relates suprathreshold apparent contrast with spatial frequency does not parallel that with detection.[23,24] For this procedure we positioned two Gabors in the middle of the screen. One was a standard 0.42-cpd Gabor patch at 14% contrast; the other, a test Gabor patch of variable spatial frequency and contrast. Subjects adjusted the contrast of the test Gabor patch until it matched the perceived contrast of the standard. For experiment 1, in which the stimuli were backprojected onto a translucent screen, the resulting matched Gabor contrasts are shown in Table 1. As can be seen, as Gabor spatial frequency increased there was an overall increase in contrast required for the match. To obtain contrasts for Gabor spatial frequencies that lay between those measured, we interpolated the contrast between adjacent spatial frequencies. In the remaining experiments in which the stimuli were displayed on a monitor, we found no significant differences in perceived contrasts between the test and the standard, so all Gabors were presented at 33% contrast.

2. Disparity Thresholds

A standard two-up one-down staircase procedure was used to measure $d_{\min}$, the threshold amplitude of dispar-

Table 1. Contrasts of Various Gabor Spatial Frequencies Obtained by Matching the Perceived Contrast of Each Gabor to a 0.42 cpd Standard Gabor at 14% Contrast[a]

	cpd						
Subject	0.45	0.59	0.84	1.19	1.68	2.38	3.36
FK	14	14	14	14	18	33	78
LZ	14	12	11	16	18	20	–

[a] These contrasts were used in experiment 1.

ity modulation at the 71% correct level. Between trials subjects viewed a blank screen with the same mean luminance as the stimulus. A small fixation cross was present between stimulus presentations. The stimulus was presented for a maximum of 2.7 s, but the subjects were encouraged to respond by key press within 1 s of stimulus onset. The response turned off the stimulus, and a tone indicated an incorrect response. Each session began with a high amplitude of disparity modulation, but within the range for being able to detect the stereoscopic shape of each component surface. The procedure terminated after twelve reversals and the geometric mean amplitude of disparity modulation was calculated over the last eight reversals.

3. EXPERIMENTS AND RESULTS

A. Experiment 1. Effect of Luminance Spatial Frequency on $d_{\min}$ for Single- and Dual-Surface Disparity Gratings

In Fig. 3(a), fusion of the two stereo halves reveals a 3-D corrugated structure that is oriented obliquely. Close inspection of the fused image reveals two corrugated surfaces interwoven in depth, creating an impression of a segmented image in stereo transparency. Each of the two surfaces in Fig. 3(a) is made up of Gabor micropatterns with the same space constant but with an octave difference in spatial frequency. In Fig. 3(b) the micropatterns on the two surfaces are identical, and fusion of the two stereo halves fails to reveal the oriented, corrugated structure. To quantify these observations we measured $d_{\min}$ for dual-surface disparity gratings as a function of Δf, the difference in center spatial frequency between the Gabor micropatterns on the two surfaces. For this experiment the spatial frequency of disparity modulation was 0.039 cpd, producing three cycles diagonally on the screen.

A potential problem with measuring $d_{\min}$ for our dual-surface gratings is that a change in the micropattern composition on one surface could simply make that surface more detectable, producing a fall in thresholds unrelated to the dual-surface nature of the stimulus. To minimize this possibility we first equated the detectability of each surface in our dual-surface stimuli. After setting the contrasts of the Gabor micropatterns to make them equal in apparent contrast (see Table 1), we measured $d_{\min}$ for single-surface disparity gratings made from each Gabor spatial frequency. Then the amplitudes of disparity modulation of each of the surfaces of the dual-surface grating were set to the ratio of their individually measured $d_{\min}$.

For the main part of the experiment we held the spatial frequency of the Gabors on one of the surfaces constant and varied Gabor spatial frequency on the other. The σ of the Gabors on both surfaces was 0.72°. Figure 4(a) shows the results from the two subjects. The two curves on each graph are for two different values of fixed f: 0.42 and 1.68 cpd for LZ and 0.42 and 2.38 cpd for FK. When $\Delta f = 0$ (identical Gabors on each surface), thresholds are very high; for the fixed $f = 1.68$ cpd condition LZ was unable to obtain any threshold, and for the fixed $f = 2.38$ cpd, FK reported an almost complete absence of any sensation of 3-D structure. However, for Δf values of ~1 octave, both subjects' thresholds fell to near-asymptotic levels. For comparison, in Fig. 4(b) we show thresholds for single-surface disparity gratings as a function of Gabor spatial frequency. Figure 4 shows that, although dual-surface thresholds never quite fall to those of the single-surface stimuli, they get close.

B. Experiment 2. Segregated versus Nonsegregated Stimuli

Could the result shown in Fig. 4 simply be a consequence of changing the stimulus from having one Gabor spatial frequency to two? The presence of more than one Gabor spatial frequency might, for example, reduce the "false-target" or stereo-correspondence problem in the stimulus *as a whole*, irrespective of whether the two types of Gabor were segregated between surfaces. Put simply, two types of Gabor might provide a more feature-rich stimulus than one. To test this possibility we conducted a control experiment, using dual-surface disparity gratings made from equal numbers of two Gabor spatial frequencies. In the *segregated* condition the two Gabor spatial

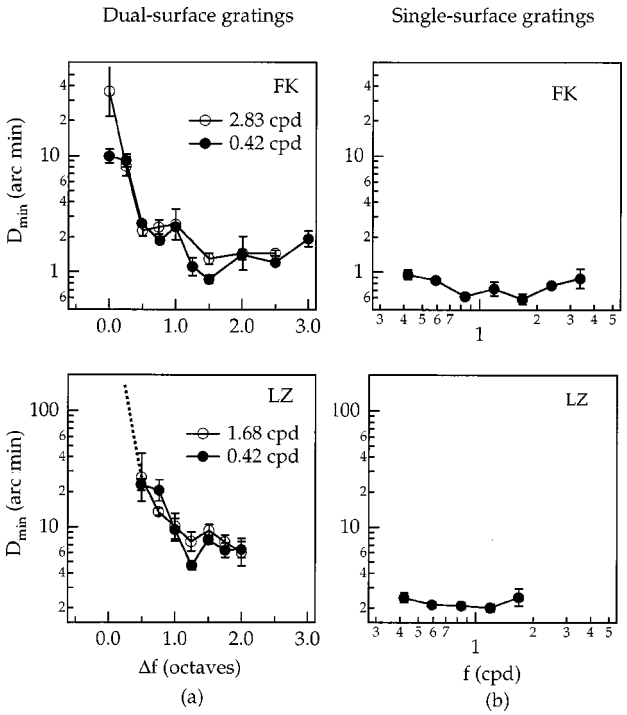

Fig. 4. Results of experiment 1. The graphs show $d_{\min}$, the threshold amplitude of disparity modulation, for identifying the orientation of a dual-surface disparity grating, as a function of the difference in spatial frequency, Δf, between the micropatterns on its two surfaces. The value of zero on the abscissa implies identical micropatterns on the two surfaces. The two curves in each graph are for two different values of fixed micropattern spatial frequency f. For the fixed $f = 0.42$ cpd condition, variable f was always higher, whereas, for the other fixed f conditions variable f was always lower. The dotted curve in LZ's data finishing at $\Delta f = 0.5$ indicates that this subject was unable to obtain a threshold below this point. On the right, thresholds for single-surface disparity gratings are shown as a function of micropattern spatial frequency.

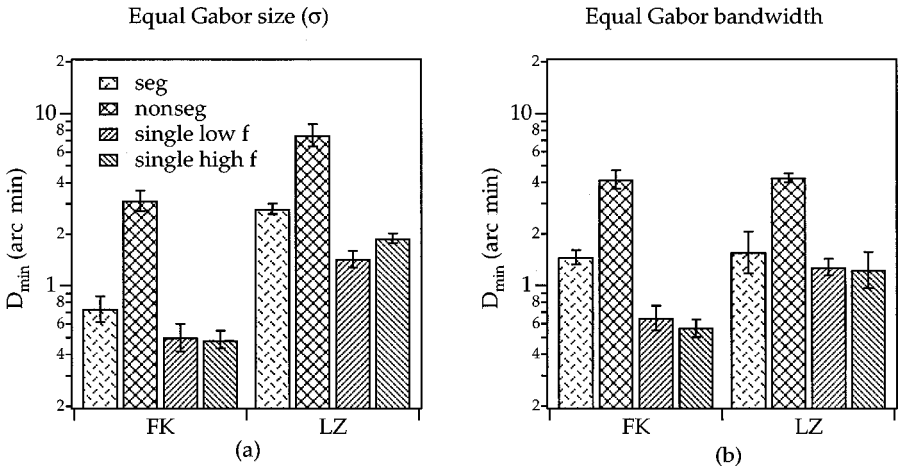

Fig. 5. Results from experiment 2. D_{min} is shown for both dual-surface and single-surface disparity gratings, with all stimuli constructed from just two types of Gabors with spatial frequencies 1.0 and 3.0 cpd. (a) Gabors with equal size, σ; (b) Gabors with equal bandwidth (σ inversely proportional to f). Seg, Gabors segregated between surfaces; nonseg, Gabors nonsegregated between surfaces. Single low f, single-surface grating with 1.0 cpd Gabors; single high f, single-surface grating with 3.0 cpd Gabors.

frequencies were separated onto the different surfaces as before. In the *nonsegregated* condition the two Gabor spatial frequencies were distributed equally onto both surfaces. Thus the segregated and nonsegregated conditions had an identical *overall* micropattern composition. If the reduction in d_{min} in Experiment 1 was due to the presence of two types of Gabor rather than one, we would expect no difference in thresholds between the segregated and the nonsegregated conditions. On the other hand, if the reduction in d_{min} was due to segregation on the basis of Gabor spatial frequency, we would expect thresholds in the segregated condition to be significantly lower. We used Gabor spatial frequencies a factor of 3 apart: 1.0 and 3.0 cpd, with σ of 0.36 deg. The disparity-modulation frequency was set to 0.085 cpd, producing four cycles diagonally on the screen. The results are shown in Fig. 5(a), along with thresholds for single-surface disparity gratings made from each type of Gabor alone. As the figure shows, thresholds for the segregated condition are much lower than for the nonsegregated condition, the former being only slightly higher than the single-surface thresholds. The average ratio of nonsegregated to segregated thresholds is 3.4 across the two subjects.

In Fig. 5(b) we show results for the same experiment using Gabors with equal bandwidth rather than equal size (σ). We set Gabor size to be inversely proportional to Gabor spatial frequency, resulting in σ of 0.36 and 0.12 for the 1.0- and 3.0-cpd Gabors. The results in Fig. 5(b) are similar to those of Fig. 5(a). The ratio of nonsegregated to segregated thresholds is on average 2.8 for the equal-bandwidth condition. Again, the single-surface thresholds are slightly lower than the segregated, dual-surface thresholds. It is interesting to note that the absolute performance levels of the two subjects with the dual-surface disparity gratings in the equal-bandwidth condition are very similar, in marked contrast to the previous experiments in which FK's overall sensitivity was significantly higher than LZ's. One possible reason for this was that LZ was more susceptible to the effects of correspondence noise. This effect was likely to be greatest in stimuli made from Gabors with small bandwidths, i.e., containing multiple carrier cycles; this was the case for the high-spatial-frequency Gabors used in all except the present equal-bandwidth experiment.

These two experiments confirm that the lower d_{min} in segregated-Gabor dual-surface gratings is a result of segregation by Gabor type and not feature enrichment. Moreover, the effects are due to Gabor spatial frequency and not bandwidth.

4. DISCUSSION

Before considering the theoretical significance of the findings of this study, consider whether d_{min} for identifying the orientation of a dual-surface disparity grating is a true reflection of the ability of human observers to segment a multisurface stimulus. What alternative methods could have been used? An obvious technique would have been to measure d_{min} for two superimposed flat surfaces. However, subjects could in principle have simply discriminated the range of disparities in the stimuli. Requiring subjects to identify the orientation of a dual-surface disparity grating forces them to detect the 3-D structure of the stimulus and not just the presence of multiple disparities.

D_{min} for a dual-surface disparity grating is not, however, a direct measure of the degree of perceived stereo transparency. Nevertheless, it is reasonable to suppose that in the segregated-by-Gabor stimuli, such as in Fig. 3(a), a close relationship exists between the perception of depth transparency and the lower measured d_{min}. Why then is it easier to perceive both the corrugated structure and the transparency in Fig. 3(a) than Fig. 3(b)? In Section 1 we suggested that densely textured overlapping surfaces pose a particular correspondence problem for the visual system, and therefore one obvious explanation is that the correspondence problem is significantly reduced

in our Gabor-segregated stimulus. However, before considering this explanation in more detail, we must first consider a number of other possibilities.

First, one can rule out the possibility that dual-surface disparity gratings made from two, as opposed to one, type of Gabor have lower $d_{\min}$ simply because they are more feature rich. In our second experiment we compared $d_{\min}$ for segregated and nonsegregated dual-surface gratings, both of which were made from equal numbers of two types of Gabor, and found lower $d_{\min}$ for the segregated condition. This result demonstrates that it is the *segregation* of the two surfaces by Gabor type, rather than the fact of having two Gabor types, that is the critical attribute of our stimuli.

A second possibility is that shape-from-disparity channels are narrowly tuned for luminance spatial frequency. By a shape-from-disparity channel we mean a mechanism that is sensitive to a particular rate of change of disparity or disparity-modulation spatial frequency. Psychophysical evidence supports the existence of such channels.[25–27] If a disparity-shape channel were stimulated by a dual-surface disparity grating made from a single Gabor spatial frequency, we would expect the two superimposed out-of-phase modulations to at least partially cancel within the channel, raising thresholds significantly (disparity-surface nulling/masking). On the other hand if the two modulations were made from different Gabor spatial frequencies, they might be detected by different disparity-shape channels and hence not cancel. Only a few studies have examined the relationship between disparity-modulation detection and luminance spatial frequency.[22,28,29] The findings of these studies have been contradictory, and none has provided conclusive evidence that disparity-shape channels are luminance spatial frequency specific. The study that employed stimuli most similar to the one here was that of Hess *et al.*[22] They measured disparity-modulation functions ($d_{\min}$ as a function of disparity spatial frequency) for single-surface disparity gratings made from Gabors of various spatial frequencies. Hess *et al.* found a marked dependence of the shape of the disparity-modulation function on luminance spatial frequency at high, but not low, disparity spatial frequencies. Although these results are consistent with the tuning of disparity-shape channels for luminance spatial frequency, as are also the results of the present study, the evidence is only indirect. Nevertheless, the possibility exists that disparity-shape-channel luminance-spatial-frequency specificity could underlie our results.

A third possible explanation is disparity averaging.[5,14–17] Two similar features lying at different depths along the same line of sight often appear to lie at a common intermediate depth. We mentioned in Section 1 the study by Rohaly and Wilson,[14] who showed that with overlapping cosine gratings, a full release from disparity averaging required a between-surface luminance spatial frequency difference of up to 3.5 octaves, suggesting that disparity averaging is a relatively broadband phenomenon. In a comprehensive review of disparity averaging, Howard and Rogers[5] concluded that much of the evidence for disparity averaging in random-element stereograms (as opposed to, for example, the continuous cosine-grating stereograms employed by Rohaly and Wilson), is better understood in terms of correspondence noise. Specifically, Howard and Rogers proposed that, in random-element stereograms with two overlapping surfaces, elements that belong to one depth surface are often paired with nearest-neighbor elements that belong to the other. This generates a mishmash of randomly matched elements and an impression of lacy depth (see below), which could easily be mistaken for disparity averaging. Our finding that a between-surface luminance-spatial-frequency difference of just 1 octave released our stimulus from its elevated $d_{\min}$ and lack of transparency reinforces the idea that it is correspondence noise rather than disparity averaging that is the limiting factor with our random-element stimuli, since the release that we observed showed narrower tuning than that found by Rohaly and Wilson. Thus, whereas we cannot rule out that disparity averaging contributes to our pattern, it seems unlikely to be the principal cause.

The fourth, and most parsimonious, explanation for our results is correspondence noise. Following Howard and Rogers, the nearest-neighbor rule produces spurious cross-surface matches in the single-Gabor and nonsegregated-Gabor dual-surface disparity gratings but does not with segregated-Gabor stimuli because only like Gabors can be matched with like. The like-with-like constraint is known as the similarity rule. Together with the uniqueness constraint, which states that a given element in one eye's view can be matched only with one element in the other, the nearest-neighbor and similarity rules are probably sufficient to account for the general pattern of data in our experiments (for reviews of all the constraints believed to underlie the solution of the correspondence problem, see Refs. 5 and 30). We have not conducted a computer simulation to test whether in combination these constraints are sufficient for predicting the pattern of results with our dual-surface stimuli. However, Fig. 6 illustrates how in principle they are sufficient. Each part of Fig. 6 represents both left eye (LE) and right eye (RE) views of a horizontal slice through the 3-D space of a half-cycle of a dual-surface disparity grating. The region bordered by the LE and RE stimuli represents the space within which the array of disparity-sensitive cells reside. The two types of element (large open circles and small filled circles) correspond to two Gabor spatial frequencies.

Consider first Fig. 6(a), where both depth surfaces are composed of the same elements, i.e., identical Gabors. This is the configuration of the stimuli with identical Gabors in experiment 1 and in Fig. 3(b). In principle, any one of the elements in the left eye's view could be matched with any one in the right eye's view, but we have chosen only those matches that together satisfy the uniqueness, nearest-neighbor, and similarity rules. The nearest-neighbor rule results in a number of spurious cross-surface matches, the effect being to collapse the two surfaces into a single, albeit irregular, depth surface in the cyclopean view. In Figs. 6(b) and 6(c), two types of Gabor are present. Figure 6(b) represents the configuration of the nonsegregated condition in experiment 2. Figure 6(c) represents the configuration of the segregated conditions in both experiments 1 and 2 and Fig. 3(a). We assume that the similarity constraint restricts matches only to

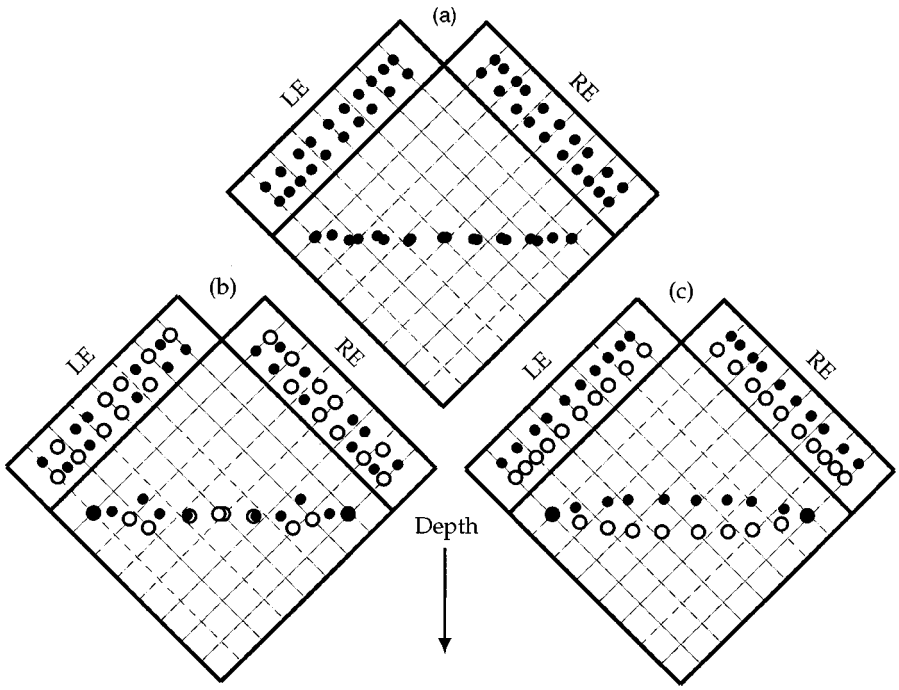

Fig. 6. Correspondence-noise model of the results. Each figure shows left eye (LE) and right eye (RE) representations of a horizontal slice through a half-cycle of a dual-surface disparity grating. For each eye the set of elements making up the two surfaces are shown one above the other for convenience. Small solid circles and large open circles represent Gabors of different spatial frequencies. Binocular matches are shown along the fixation plane running horizontally between the corners of each figure. The nearest-neighbor rule finds those matches that minimize disparity with respect to the plane of fixation, but the similarity rule restricts those matches to like Gabors. The uniqueness rule allows only one match per element. (a) Single-Gabor stimulus [as in Fig. 3(b)], (b) nonsegregated two-Gabor stimulus, (c) segregated two-Gabor stimulus [as in Fig. 3(a)]. Only in (c) is the dual-surface, sinusoidal structure of the stimulus revealed.

like Gabors but that the other constraints operate in the same way as in Fig. 6(a). The result shown in Fig. 6(b) is a series of matches somewhat more jumbled in depth than in Fig. 6(a) but that still do not form two distinct depth surfaces because there are still instances of cross-surface matching. In Fig. 6(c), however, the matched elements form two distinct depth surfaces, as no cross-surface matching is allowed.

Interestingly, this analysis might suggest that we could account for the qualitative trends in our data by merely supposing that subjects responded to the interval in each forced-choice pair with the most perceived *variability* in depth, which is greatest in Fig. 6(c) and least in Fig. 6(a). However, one must remember that the task that we employed required identification of the orientation of the depth corrugations, i.e., the 3-D structure of the dual-surface stimulus. It is only in Fig. 6(c) that the full 3-D structure of the stimulus is revealed.

What exactly is "similarity" as applied to pairs of Gabors of different spatial frequencies? One common measure of similarity is the cross-correlation function. We calculated the peak of the (horizontal) cross-correlation function between the pairs of Gabors employed in our first experiment and found that it declines by a factor of 10 for an approximately 1.4-octave difference in peak Gabor spatial frequency. This decline is similar to the rate at which the thresholds fall in Fig. 4(a). Consistent with the importance of the similarity constraint based on luminance spatial frequency is the recent study by Ziegler and Hess.[31] They showed a complete absence of any perception of stereo shape when they viewed a single-surface disparity grating presented dichoptically with a 2-octave difference in Gabor spatial frequency between the two eyes' views (see also Ref. 32).

Before leaving the discussion of correspondence noise, we must mention the surface-smoothness constraint. This rule posits that only those matches forming surfaces that smoothly vary in depth are chosen. A number of models[33–36] have used the surface-smoothness constraint in conjunction with other constraints to solve the false-target problem in random-dot stereograms. In the best known of these models, that of Marr and Poggio,[35] the surface-smoothness constraint was implemented by a network of facilitatory and inhibitory interactions between local disparity detectors: facilitation in the two-dimensional (frontoparallel) plane and inhibition in the 3-D (i.e., depth) plane. The phenomenon of depth transparency has been used as an argument against this model and the surface-smoothness constraint that it embodies,[30] as the inhibitory interactions should eliminate those matches that give rise to multiple depths along similar lines of sight. However, provided that the surface-smoothness constraint is implemented in conjunction with the similarity constraint (as in the Marr–Poggio model), we see no reason why in principle this might not also reduce the number of false targets within each surface of a Gabor-segregated dual-surface disparity grating, such as in Fig. 3(a).

To summarize: Either luminance-spatial-frequency channels remain segregated up to the level of the depth

corrugation channels or they remain separate up to the level of the mechanisms that implement the similarity constraint for matching. Further experiments will be needed before we can decide which of these mechanisms underlies our results. Further experiments will also establish whether other feature attributes besides luminance spatial frequency, such as orientation, color, and luminance polarity, are salient for stereo-depth segmentation in dual-surface disparity gratings.

Finally, under what circumstances might differences in luminance spatial frequency be useful for stereo-depth segmentation? Consider again Fig. 1. The fern has two physically similar fronds lying in different depths, and these will have slightly different center luminance spatial frequencies in their retinal-image projections. This difference in spatial frequency content could facilitate stereoscopic segmentation. The two fronds may of course also differ in orientation, luminance polarity, or color, but such differences are not inevitable. Unlike these properties, spatial scale differences will tend to covary consistently with disparity. The importance of gradients in spatial scale that are due to perspective in the monocular retinal image has been appreciated by vision scientists since Gibson's publication[37] (e.g., see Refs. 38–40). We have demonstrated that stereopsis can also exploit such spatial scale differences when one is segmenting the depths of complex textured scenes.

ACKNOWLEDGMENTS

This research was supported by a Canadian National Science and Engineering Research Council grant to F. A. A. Kingdom and a Canadian MRC grant to R. F. Hess.

F. A. A. Kingdom's e-mail address is fkingd@po-box.mcgill.ca.

REFERENCES AND NOTES

1. K. Prazdny, "Detection of binocular disparities," Biol. Cybern. **52**, 93–99 (1985).
2. R. A. Akerstrom and J. T. Todd, "The perception of stereoscopic transparency," Percept. Psychophys. **44**, 421–432 (1988).
3. D. Weinshall, "Perception of multiple transparent planes in stereo vision," Nature **341**, 737–739 (1989).
4. S. Gephstein and A. Cooperman, "Stereoscopic transparency: A test for binocular disambiguating power," Vision Res. **38**, 2913–2932 (1998).
5. I. P. Howard and B. J. Rogers, *Binocular Vision and Stereopsis* (Oxford U. Press, Oxford, UK, 1995). For a discussion of disparity averaging and its limitations as an explanatory concept, see pp. 230–234. For a discussion of the constraints on stereo matching, see pp. 216–229.
6. J. M. Lankheet and M. Palmen, "Stereoscopic segregation of transparent surfaces and the effect of motion contrast," Vision Res. **38**, 659–668 (1998).
7. B. Julesz and J. E. Miller, "Independent spatial frequency tuned channels in binocular fusion and rivalry," Perception **4**, 125–143 (1975).
8. J. E. W. Mayhew and J. P. Frisby, "Rivalrous texture stereograms," Nature **264**, 53–56 (1976).
9. C. M. Schor and I. Wood, "Disparity range for local stereopsis as a function of luminance spatial frequency," Vision Res. **23**, 1649–1654 (1983).
10. H. R. Wilson, R. Blake, and D. L. Halpern, "Coarse spatial scales constrain the range of fusion of fine spatial scales," J. Opt. Soc. Am. A **8**, 229–236 (1991).
11. Y. Yang and R. Blake, "Spatial frequency tuning of human stereopsis," Vision Res. **31**, 1177–1189 (1991).
12. H. S. Smallman and D. I. A. MacLeod, "Size-disparity correlation in stereopsis at contrast threshold," J. Opt. Soc. Am. A **11**, 2169–2183 (1994).
13. G. C. DeAngelis, I. Ohzawa, and R. D. Freeman, "Neuronal mechanisms underlying stereopsis: How do simple cells in the visual cortex encode binocular disparity?" Perception **24**, 3–31 (1995).
14. A. M. Rohaly and H. R. Wilson, "Disparity averaging across spatial scales," Vision Res. **34**, 1315–1325 (1994).
15. B. Julesz and S. C. Johnson, "Stereograms portraying ambiguous perceivable surfaces," Proc. Natl. Soc. **61**, 437–441 (1968).
16. A. J. Parker and Y. Yang, "Spatial properties of disparity pooling in human stereo vision," Vision Res. **29**, 1525–1538 (1989).
17. S. B. Stevenson, L. K. Cormack, and C. M. Schor, "Depth attraction and repulsion in random dot stereograms," Vision Res. **31**, 805–813 (1991).
18. C. W. Tyler, "Depth perception in disparity gratings," Nature **251**, 140–142 (1974).
19. F. A. A. Kingdom, L. R. Ziegler, and R. F. Hess, "The role of spatial scale in stereoscopic segmentation," Perception Suppl. **27**, 21 (1998).
20. K. Boothroyd and R. Blake, "Stereopsis from disparity of complex grating patterns," Vision Res. **24**, 1205–1222 (1984).
21. H. S. Smallman and S. P. McKee, "A contrast ratio constraint on stereo matching," Proc. R. Soc. London Ser. B **260**, 265–271 (1995).
22. R. F. Hess, F. A. A. Kingdom, and L. R. Ziegler, "On the relationship between the spatial channels for luminance and disparity processing," Vision Res. **39**, 559–568 (1999).
23. M. A. Georgeson and G. D. Sullivan, "Contrast constancy: Deblurring in human vision by spatial frequency channels," J. Physiol. (London) **252**, 677–656 (1975).
24. N. Brady and D. J. Field, "What's constant in contrast constancy? The effect of scaling on the perceived contrast of bandpass patterns," Vision Res. **35**, 739–756 (1995).
25. R. A. Schumer and L. Ganz, "Independent stereoscopic channels for different extents of spatial pooling," Vision Res. **19**, 1303–1314 (1979).
26. C. W. Tyler, "Sensory processing of binocular disparity," in *Vergence Eye Movements: Basic and Clinical Aspects*, M. C. Schor and K. J. Ciuffreda, eds. (Butterworth, Boston, Mass., 1983), pp. 199–296.
27. A. B. Cobo-Lewis and Y. Y. Yeh, "Selectivity of cyclopean masking for the spatial frequency of disparity modulation," Vision Res. **34**, 607–620 (1994).
28. K. Pulliam, "Spatial frequency analysis of three-dimensional vision," in *Visual Simulation and Image Realism II*, K. S. Setty, ed., Proc. SPIE **303**, 71–77 (1981).
29. B. Lee and B. Rogers, "Disparity modulation sensitivity for narrow-band-filtered stereograms," Vision Res. **37**, 1769–1778 (1997).
30. R. Blake and H. R. Wilson, "Neural models of stereoscopic vision," Trends Neurosci. **14**, 445–452 (1991).
31. L. R. Ziegler and R. F. Hess, "Stereoscopic depth but not shape perception from second-order stimuli," Vision Res. **39**, 1491–1507 (1999).
32. L. R. Ziegler, F. A. A. Kingdom, and R. F. Hess, "Local luminance factors that determine the maximum disparity for seeing cyclopean surface shape," Vision Res. **40**, 1157–1165 (2000).
33. G. Sperling, "Binocular vision: A physical and a neural theory," Am. J. Psychol. **83**, 461–534 (1970).
34. J. L. Nelson, "Globality and stereoscopic fusion in binocular vision," J. Theor. Biol. **49**, 1–88 (1975).
35. D. Marr and T. Poggio, "Cooperative computation of stereo disparity," Science **194**, 283–287 (1976).

36. J. E. W. Mayhew and J. P. Frisby, "The computation of binocular edges," Perception **9**, 69–86 (1980).
37. J. Gibson, *The Perception of the Visual World* (Houghton-Mifflin, Boston, Mass., 1950).
38. J. Cutting and R. Millard, "Three gradients and the perception of flat and curved surfaces," J. Exp. Psychol. Gen. **113**, 198–216 (1984).
39. J. Todd and R. Akerstrom, "Perception of three-dimensional form from patterns of optical texture," J. Exp. Psychol. Hum. Percep. **13**, 242–255 (1987).
40. K. Stevens and A. Brookes, "Integrating stereopsis with monocular interpretations of planar surfaces," Vision Res. **28**, 371–386 (1988).