Vision Res. Vol. 28, No. 4, pp. 503–519, 1988
Printed in Great Britain

0042-6989/88 \$3.00 + 0.00
Pergamon Press plc

RECIPROCITY BETWEEN LUMINANCE AND DOT DENSITY IN THE PERCEPTION OF BRIGHTNESS

JEFFREY B. MULLIGAN* and DONALD I. A. MACLEOD
Department of Psychology, University of California at San Diego, La Jolla, CA 92093, U.S.A.

(*Received* 13 *November* 1986; *in revised form* 26 *August* 1987)

Abstract—Thresholds were measured for detecting perturbations in a regular lattice of dots by modulating local dot density, local dot luminance, or some combination of the two. For high mean densities (dot spacing $\leqslant$ 15 min of arc), perturbations in local density increase the perceived brightnesses of the individually resolved elements in the more densely filled regions, and appear (at near threshold levels) as modulations of brightness rather than density. This illusory brightness modulation may be nulled by applying a real luminance modulation to make the lattice elements appear equally bright. Once this is done, thresholds for detecting any nonuniformity in the array are elevated compared to thresholds for detecting uncompensated density modulation. This result suggests that uncompensated density modulation is detected via the illusory brightness variations. This interpretation suggests that dot brightness is determined on the basis of the space average luminance of an area a substantial fraction of 1 deg in diameter. To test this hypothesis, thresholds were measured for detecting luminance modulation in a regular array of dots viewed against a comparatively dim background, where the modulation was applied to the dots themselves, to the background alone, or to both the dots and the background in either reinforcing or cancelling relative phase. For small, closely spaced dots, the threshold for modulation of luminance can be predicted on the basis of the amplitude of the Fourier component at the modulation frequency, regardless of whether it is carried by dots, the background, or both. The threshold is greatly elevated when modulation in the dots cancels the background modulation, so that there is contrast modulation of the dots, but no net energy at the fundamental frequency (zero amplitude of the Fourier component). For large, coarsely spaced dots, on the other hand, thresholds for conditions which contain energy at the fundamental modulation frequency are higher. The threshold increase is much greater when the modulation is applied to the dots than when it is applied to the background. This result suggests that the coarsely spaced dots are saturating the response of spatially opponent units. This hypothesis was confirmed by tests using backgrounds with the same luminance as the dots; threshold elevations selective for dots or background were abolished.

Spatial modulation sensitivity   Brightness discrimination

INTRODUCTION

It is well known that the brightness of an object is not determined solely by its luminance. Simultaneous contrast and assimilation are two complementary effects whereby the brightness of an object or region is influenced by the presence of a nearby object or region. Simultaneous contrast refers to the tendency for an object to appear darker when viewed against a light background, and lighter when viewed against a dark background. It is usually interpreted as an expression of the antagonistic influence of surround stimulation in cells with spatially-opponent receptive fields (Ratliff, 1965). Assimilation, on the other hand, describes a situation where the brightness of a region is shifted *towards* the brightness of the perturbing region. Whether an object will have its brightness shifted by contrast or assimilation generally depends on the spatial configuration of the perturbing region, with thin lines producing assimilation and large regions and surrounds producing contrast. The spatial factors governing the two effects have been summarized by Helson (1963).

This paper describes a brightness illusion which, like assimilation, seems to be the result of a synergistic, rather than antagonistic, effect of nearby stimulation. The effect was first noted during an attempt to measure motion thresholds by applying a velocity field to an array of randomly placed dots (Nakayama *et al.*, 1985). When the velocity field was such that the motion of the dots produced changes in the local dot density, Nakayama *et al.* were able to elevate thresholds for detecting the motion by modu-

*Present address: NASA Ames Research Center, Mail Stop 239-3, Moffett Field, CA 94035, U.S.A.

lating the luminances of individual dots in such a way that the space-average luminance of the display remained constant in spite of the fluctuations in local dot density. This showed that the fluctuations in space-average luminance (over a region containing several dots) could serve as a cue for motion independent of motion *per se*. The subjective appearance of such displays can vary widely: small compressive motions without reciprocal luminance modulation can appear as brightness modulation without the perception of motion (Mulligan, unpublished observation); conversely, under other conditions, luminance modulation of a static dot pattern may give rise to the perception of motion (MacLeod, Nakayama and Silverman, unpublished observations).

One striking aspect of the effect is that it seems to hold at virtually all levels of modulation. Dots whose luminance is being modulated by 50% may appear to have a constant brightness when the local density is modulated to maintain a constant space-average luminance. The only way to convince skeptical observers that the dots' luminances are indeed being modulated is to direct them to note the artifactual changes in dot size which most CRT's produce with changing luminance ("blooming").

Studies of preattentive vision (Julesz, 1980; Treisman and Gelade, 1980; Treisman and Schmidt, 1982; Bergen and Julesz, 1983) have led theorists to postulate the existence of "feature maps" which provide a spatial representation of particular stimulus dimensions (Crick, 1984; Nakayama and Silverman, 1986). Such maps may represent features at a lower spatial resolution than the objects that carry them, even though the objects themselves may be perfectly resolvable. There is evidence that many attributes are represented at spatial resolutions far below visual acuity: Nakayama and Tyler (1981) and Nakayama *et al.* (1985) demonstrated that spatial integration is performed on the outputs of low-level motion detectors; Rodieck (1983) has pointed out that flicker is seen with a resolution less than visual acuity, a fact that is exploited in the reduction of subjective television flicker by interlacing. The even scan lines are not seen to flicker when the odd scan lines (which do not substantially overlap the even lines) are flickering with the opposite phase, even though the lines and the flicker can be seen clearly when one set of lines is dark. It is conceivable that the reciprocity observed between dot luminance and density may be due to poor spatial resolution in a "brightness map", and the following experiments are directed towards determining the spatial characteristics of this hypothetical brightness system.

EXPERIMENT I

Procedure

With the idea in mind that the brightness of a single dot is computed on the basis of the light falling on an area much larger than the dot itself, we attempted to measure the size of the hypothetical integration area by determining the maximum dot spacing for which perturbations of density were seen as perturbations of brightness. The stimulus was a regular lattice of dots in which static dot density and/or brightness varied sinusoidally across the pattern. On each trial, subjects were presented with a lattice to which some density modulation was applied; their task was to adjust the amplitude of the luminance modulation to make the dots appear equally bright.

Stimuli were presented on Hewlett Packard display scope (model No. 1300A) with P4 phosphor. Special hardware [designed by Walter Kropfl at Bell Laboratories, and described in detail in Breitmeyer *et al.* (1975)] generated the deflection signals under program control to produce a perfectly regular lattice. The display scope had been modified to accept the additional analog signals which were summed with the x, y and z inputs to produce the perturbations. Digital-to-analog converters (DACs) attached to the computer produced the signals needed to produce the modulations of position and luminance.

If all the dots in a lattice have their position shifted by a given amount, there is no net change in dot density. To produce *variations* in dot density it is necessary to apply different positional shifts to different dots; changes in dot density are proportional to the *spatial derivative* of a perturbation applied to dot position. Therefore, if one wishes to produce a sine phase modulation of dot density, a *cosine* phase signal must be used to perturb the dot positions. Thus, in order to produce modulations of density and luminance that would produce uniform space-average luminance, the two DAC's had to produce signals in quadrature. The absolute phase was chosen so that positional displacements were never applied to the dots at the edges of the pattern, thus ensuring that the overall size of the pattern was constant and independent of modulation depth.

The relative luminance produced for various DAC settings (gamma) was measured with a photodiode; the calibration data were incorporated into the computer software to enable truly sinusoidal variations of luminance to be produced. Regrettably, no absolute calibration of luminance was made.

Three lattice spacings, each with eight amplitudes of density modulation were tested. The mean dot luminance was fixed, so the sparser lattices had correspondingly lower space-average luminances. For each lattice, the amplitude of luminance modulation producing the appearance of uniform brightness was determined using the method of adjustment. The 24 lattices were each presented three times in a random order.

The display was continuously visible, with the amplitude of the luminance modulation being dynamically controlled by a knob under the subject's control. The knob produced a voltage that was read by an analog-to-digital converter attached to the computer. A computer program used this voltage to set the amplitude of the luminance modulation in the lattice. On each trial, the computer added a random offset to the values read from the knob, so that subjects could not use knowledge about the position of the knob in making their settings. Both positive and negative modulations could be produced about zero; the knob position corresponding to zero luminance modulation was always near the middle of the mechanical range of the knob, so that the luminance modulation could either be spatially in-phase with the density modulation or in opposite phase. Subjects were instructed to adjust the level of luminance modulation to a point where the dots in the lattice appeared equally bright. Three experienced psychophysical observers were used as subjects, two of whom (J.R.B. and D.R.W.) were ignorant of the purpose of the experiment.

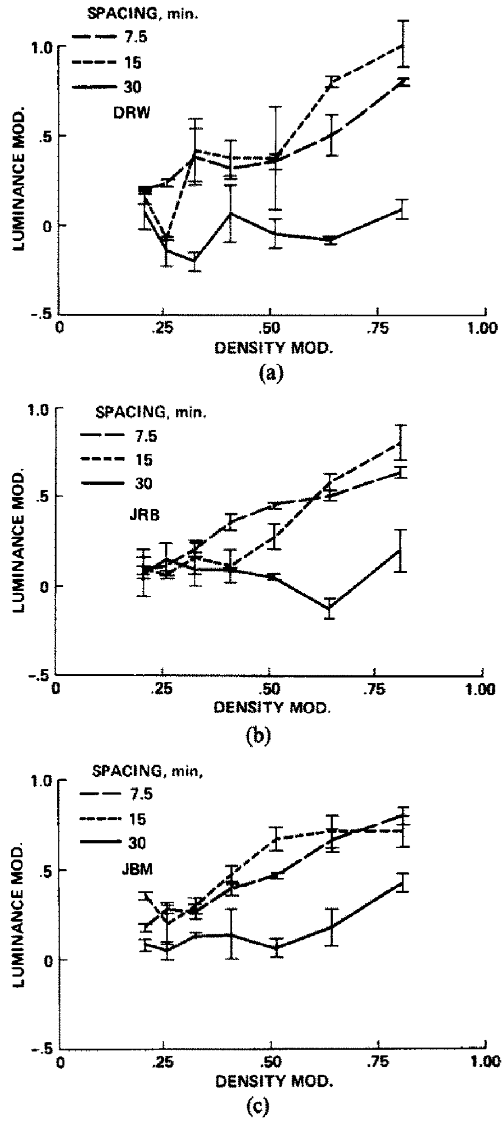

Fig. 1. Amplitude of the luminance modulation required for uniform subjective brightness of dots is plotted as a function of the amplitude of applied density modulation. Positive luminance modulation signifies that dots are made ***brighter*** in regions of reduced density, while negative luminance modulation signifies that the dots are made ***dimmer*** in regions of reduced density. The solid curve is for a mean dot spacing of 30 min arc, dotted curve for 15 min and dashed curve for 7.5 min. For spacings of 7.5 and 15 min of arc (dashed and dotted curves, respectively), luminance and density modulations at the equi-brightness settings are roughly proportional. At a spacing of half a degree (solid curve), dot brightness is only slightly affected by density modulation. The error bars represent ± 1 SEM. (a) Subject D.R.W. (b) Subject J.R.B. (c) Subject J.B.M.

Results

Figure 1 shows the luminance contrast selected for uniformity of subjective brightness as a function of the amplitude of the applied density modulation, for subjects D.R.W. [Fig. 1(a)], J.R.B. [Fig. 1(b)], and J.B.M. [Fig. 1(c)]. Each curve represents a different mean lattice spacing. In these figures, luminance amplitude is defined as positive if the dots are set brighter in regions of reduced density, and negative if these dots are set dimmer. Each data point is the mean of three settings; the error bars represent plus and minus one standard error of the mean. The data show that for mean lattice spacings as large as 15 min of arc the luminance modulation producing a sensation of equal brightness is proportional to the amplitude of the applied density modulation. For a dot spacing of 30 min, however, density modulations have no effect on the subjective brightness of the dots; the corresponding curves in Fig 1(a) and 1(b) have nearly a flat slope. In this case the null settings correspond to the point where the dots all have the same physical luminance; that is, there is no interaction between dots.

Discussion

The results demonstrate that the mean dot spacing can have strong effects on the amount of coupling between dot density and apparent brightness. In order to test the idea that the results are due to brightness signals being integrated over some moderately large area, the experiment was simulated for a simple model for the receptive field for brightness. Brightness signals were assumed to be the outputs of units having spatially bivariate Gaussian receptive fields with standard deviation σ. Such a unit will respond linearly to changes in intensity of the dots falling on its receptive field; it will also respond linearly to changes in density when the density is high, but *not* when it is low.

Figure 2 shows the slope predicted for a curve of the type shown in Fig. 1, as a function of log dot spacing (expressed in units of σ). The curve in Fig. 2 falls from a value near one to a value near zero over a change of about 0.3 log units in dot spacing; this is similar to the results of the experiment, where both observers showed a transition from complete reciprocity at a spacing of 15 min of arc, to no interaction

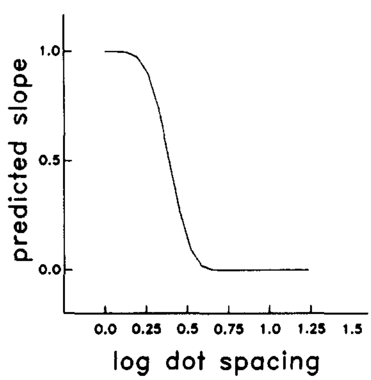

Fig. 2. Results of a simulation in which brightness was modeled as the output of units having spatially bivariate Gaussian receptive fields. The abscissa represents the log of the dot spacing in a regular lattice, expressed in units of the standard deviation of the Gaussian receptive field. The ordinate represents the predicted slope of the curve relating the amplitude of a nulling luminance modulation to a given applied density modulation, shown in Fig. 1.

with a spacing of 30 min. We can estimate σ, the standard deviation of the receptive field for brightness, by assuming that the 30 min spacing corresponds to the point on the abcissa labeled 0.5 in Fig. 2. This results in a value of approximately 10 min for σ.

EXPERIMENT II

Introduction

Experiment II was designed to address many of the same issues as in Experiment I, using a different task. Instead of adjusting luminance to produce uniform brightness, subjects set thresholds for detecting *any* type of nonuniformity in a dot lattice that was modulated either in luminance, in density alone, or in both luminance and density. Thresholds for combinations were measured both in *cancellation phase*, where variations of dot density and brightness exactly cancel to produce a constant space average luminance, and *summation phase*, where the combined effects produced a net modulation of space-average luminance exactly twice that produced by either stimulus dimension alone. Keeping in mind the model of spatial integration of brightness signals suggested by the results of Experiment I, we expected to see two regimes in the data, corresponding to mean dot spacings above and below the size of the integration region. For dot spacings small enough so that dot density interacts with luminance in the generation of the brightness signal, we would expect to find higher thresholds for joint modulations in cancellation phase than in summation phase. For dot spacings larger than the size of the integration area, density modulations should not interact with luminance, and the thresholds for the two phases should be equal.

Procedure

The apparatus for Experiment II is shown schematically in Fig. 3(a), with the waveforms from selected test points shown in Fig. 3(b). A 50 Hz sawtooth was provided by oscilloscope O1 (Tektronix 538A) for the vertical deflection signal. The display was viewed on oscilloscope O2 (Tektronix 502A); a temporal sawtooth horizontal deflection signal was provided by the time base unit of O2, which was set to 20 μsec/cm. Both the vertical and horizontal sawtooth signals were triggered by function generator G1, whose frequency was set to 1550 Hz, producing a raster with 31 lines.

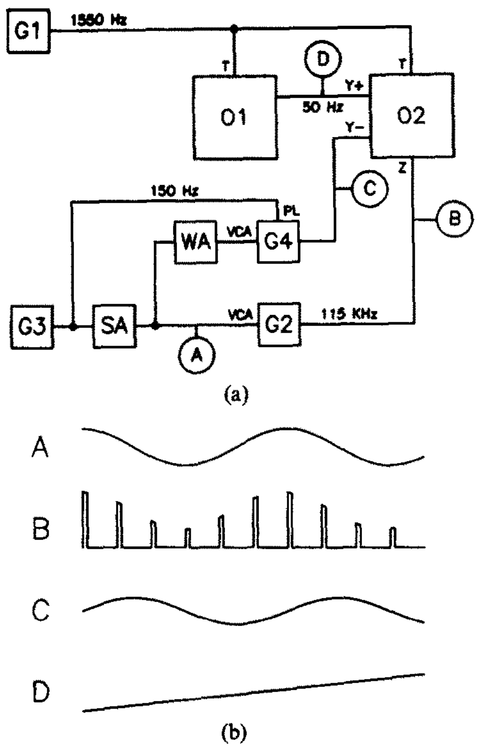

Fig. 3. Block diagram of apparatus for Experiment II. Stimuli are viewed on oscilloscope O_2; horizontal deflection of the raster is provided by O_2's internal circuitry, while the vertical signal comes from oscilloscope O1. Both scopes are triggered by function generator G1. Bright dots are produced on the screen by supplying pulses from function generator G2 to the z-axis input of oscilloscope O2. Modulation of the brightness is accomplished by applying a modulating signal from G3 (attenuated by the subject using attenuator SA) to the voltage-controlled-amplifier (VCA) gain input of G2. Vertical perturbations to the raster are provided by function generator G4, which is phase locked to G3, and whose output amplitude is proportional to the output of the subject's attenuator (SA) through use of the wave analyser (WA), which measures the signal from SA, in conjunction with the VCA input of G4. (b) Typical waveforms (not to scale) are shown for various test points shown in Fig. 2(a). Curve labeled "A" is the brightness modulating signal, a 150 Hz sine wave. Curve labeled "B" is the pulse train whose height is modulated by signal "A". Curve labeled "C" is the position perturbing signal, which is in quadrature with signal "A". Signal "D" is a 50 Hz ramp which provides vertical deflection for the raster; the absolute amplitude of signal "D" is generally much greater than that of signal "C".

Dots were produced by sending pulses from function generator G2 (Wavetek model 186) to the z-axis (brightness) input of scope O2. The pulses were generated at a frequency of 115 kHz, and had a duty cycle of 0.08. The fast (x) sawtooth was used to gate the pulses to achieve blanking. The net result was a lattice consisting of 31 rows, each with 29 dots, which filled the scope face.

Brightness modulation was produced by applying the signal from another function generator (G3) to the voltage-controlled amplifier (VCA) input of generator G2. The frequency of G3 was adjusted close to 150 Hz (three times the vertical rate, so that there were three grating cycles on the screen), but was not phase-locked to any of the other signals. By having G3 free running, the pattern was not stationary on the screen over long intervals of time, but was effectively constant during the time required to make a single setting. The slow drift was considered desirable, as it eliminated the possibility of local changes in the steady-state level of light adaptation. The grating drift frequency was less than 0.05 Hz.

Luminance calibrations were made using a spot meter (Pritchard UBD 1/4°). Space-average luminance was measured by defocusing the spot meter until the (unmodulated) lattice appeared uniform through the viewfinder. The luminance modulation was then calibrated by applying a large modulating signal, and measuring the luminances at the peak and trough of the resulting grating. The variations in space-average luminance resulting from density modulation were also checked this way, although an independent calibration of density modulation was done simply in terms of the deflection voltages. The space average luminance was $0.22\ cd/m^2$ corresponding to about $2\ \mu cd$ per dot.

The luminance modulation signal from G3 was passed through a variable attenuator en route to the VCA input of G2; This attenuator was under the control of the subject, who used it to set thresholds using the method of adjustment. The attenuator consisted of a 10-turn linear potentiometer, equipped with a counting dial; the experimenter read settings directly from the dial.

Modulation of vertical dot position was accomplished by adding a sinewave signal to the vertical deflection signal. This was easily done since oscilloscope O2 provided differential inputs on the vertical amplifier. The modulating signal itself was produced by function generator G4 (Wavetek model 186) which was phase-locked to the luminance signal from G3. Since density modulation is proportional to the first spatial derivative of position modulation, the signal from G4 had to be in quadrature with the signal from G3 to obtain the proper phase relation between modulations of density and luminance.

It was desired that subjects to be able to vary the two modulations with a single control while keeping them in a fixed ratio; to that end, the amplitude of the signal from G4 was made proportional to the luminance modulation depth (the amplitude of the signal from G2) by driving the VCA input of G4 (as well as G2) from the output of G3 as attenuated by the subject. In this case, however, what was needed was not to control the gain with the instantaneous *level* of the signal from G3, but rather with the *amplitude* of that signal. The amplitude of the attenuated signal from G3 was therefore measured by wave analyser WA (General Radio model 1900), set for a 50 Hz acceptance bandwidth. The recorder output of wave analyser WA provided a d.c. level proportional to the amplitude of the signal producing luminance variation and was connected to the VCA input of G4. The VCA gain of G4 was set to its maximum value, and the ratio of the two types of modulation was adjusted using the VCA gain control on generator G2.

Only one subject (DIAM) was run in Experiment II: unfortunately, parts of the apparatus were appropriated by other members of the laboratory for use in other projects before additional subjects could be tested. DIAM is an experienced psychophysical observer with good (uncorrected) acuity.

The display was viewed at a distance of 5 feet, resulting in a mean dot spacing of 6 min of arc of visual angle, a dot diameter of roughly 1 min (effectively approximating a point source), and a modulation frequency of 1 c/deg. The entire lattice subtended a region three degrees square, with three cycles of the modulation visible.

Closer viewing distances were used to obtain sparser dot densities. This had the disadvantage of covarying modulation frequency with dot spacing, but this was deemed unimportant, since the intent of the experiment was not to measure *absolute* sensitivities, but rather relative sensitivities to the different types of modulation. For the coarsest lattice (48 min of arc of visual angle spacing) in addition to using the closest viewing distance (14 in), the horizontal and vertical gains of scope O2 were doubled.

Results

Typical data are shown in Fig. 4. A two-dimensional space is used to represent joint modulation of luminance and density. In the figure, the abscissa represents the contrast of the applied luminance modulation, while the ordinate represents modulation of spacing, or density. The units of density modulation have been chosen in the same way as for luminance, so that a modulation of 1 corresponded to an excursion from zero density to twice the average. For a chosen ratio of density-to-luminance modulations, adjustments of the subject's attenuator produced modulations which lay on straight lines through the origin in this space.

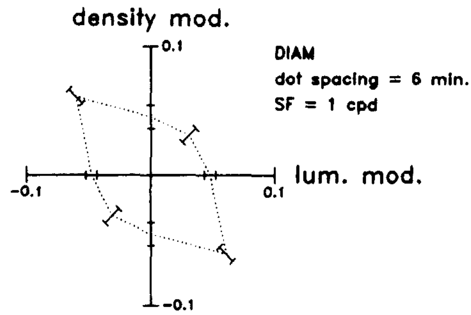

Fig. 4. Thresholds for seeing nonuniformity in a regular lattice are plotted in a two dimensional space where the x-axis represents contrast of luminance modulation and the y-axis represents amplitude of density modulation (normalize so an amplitude of 1 corresponds to variations from zero density to twice the mean). The thresholds for each type of modulation alone are approximately equal, suggesting that a single mechanism sensitive to the amplitude of the Fourier fundamental component of the luminance image is responsible for the detection. In quadrants 1 and 3 (upper right and lower left, respectively), where the luminances of dots are increased in regions of increased dot density (summation phase), thresholds are lower than for either type of modulation alone. In quadrants 2 and 4 (upper left and lower right, respectively), where luminance modulation compensates for density modulation to make all dots appear equally bright (cancellation phase), thresholds are significantly elevated; that is, visible density modulations are rendered invisible by the simultaneous application of luminance modulation in the opposite phase. Four points were measured and then reflected to show the complete threshold contour. Dot spacing was 6 min of arc of visual angle, and the modulation frequency was 1 c/deg. Error bars represent ± 2 SEM, based on variability between sessions.

The results shown in Fig. 4 are for a relatively dense lattice with a dot spacing of 6 min. Four points were measured; each has been reflected in the figure. Several aspects are noteworthy: first, the thresholds for luminance modulation only and density modulation only are approximately equal; second, when luminance and density modulation are added in-phase (so that dots are brighter in regions of increased density) the combination can be seen when each component is at little more than half of the threshold modulation for that component alone. Surprisingly, for each of these thresholds the nonuniformity of the lattice appeared at threshold as a modulation of dot brightness: the modulation of density was not seen as such at threshold.

The thresholds for cancellation phase (quadrants 2 and 4, upper left and lower right, respectively) show that our term is an apt one: the two types of modulation indeed appear to cancel each other, producing a joint threshold that is larger than that for either type alone. At this threshold, the subject could not have been detecting a Fourier component in the intensity profile of the whole display at the fundamental frequency of the modulation, since there was zero amplitude regardless of the modulation depth. This is because the two components of the modulation kept the *local* space-average luminance constant, regardless of the modulation depth. In this case the subject was forced to detect the change in dot spacing (assuming he could not detect the luminance modulation when the dots *appeared* equally bright). The fact that this threshold is higher than for the other cases shows that the spacing cue is unimportant in determining threshold in the other conditions where there was an actual component at the Fourier fundamental frequency. The ratio of the cancellation phase threshold to the summation phase threshold gives us an indication of the extent of reciprocity between dot density and luminance: a threshold ratio greater than 1 indicates that there is some reciprocity of density and luminance in cancellation phase.

In the absence of any reciprocity, that is, for a system that detects modulations of density and luminance completely independently, we would expect the diagonal thresholds to be all the same. The combined thresholds should exhibit luminance and density components slightly smaller than for either type of modulation by itself, either as a result of probability summation, or on the basis of signal detection theory (Green and Swets, 1974). If one assumes that the two types of stimuli stimulate different mechanisms, that a given stimulus evokes an internal response which is normally distributed about a mean value linearly related to the stimulus level, and that the distributions for the the two stimulus dimensions are independent of each other and of both stimulus levels, then it can be shown that threshold should decrease by a factor of the square root of two when the two types of modulation are combined in amounts proportional to the individual thresholds (see Appendix I).

Figure 5 shows a plot of the threshold ratio

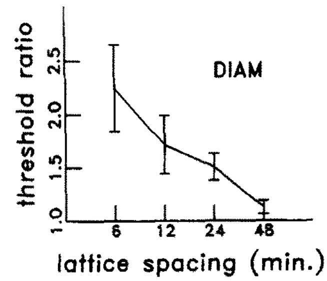

Fig. 5. The interaction between density and luminance is assessed by evaluating the ratio of the cancellation phase threshold (quadrants 2 and 4 of Fig. 3) to the summation phase threshold (quadrants 1 and 3 of Fig. 3); this ratio is plotted as a function of lattice spacing. The interaction persists for dot spacing up to 24 min of arc. Error bars represent ± 2 SEM, based on variability between sessions. Subject D.I.A.M.

versus lattice spacing. Here we see a fairly gradual drop from a high ratio at small lattice spacings to a value of 1, consistent with independent detection of brightness and density modulation at the largest spacing. The value of 1 is not approached until the spacing reaches 48 min of arc. The fact that a significant interaction between density and luminance is still occurring at a mean lattice spacing of 24 min is roughly consistent with the results of Experiment I, where the interaction seemed to vanish at a spacing of 30 min. Exact agreement is perhaps not to be expected given the differing natures of the tasks in the two experiments: it is possible that threshold for seeing density modulation (as measured in Experiment II) might be affected by a physical luminance modulation even in a lattice where density modulation did not produce a change in the subjective brightness of the dots.

EXPERIMENT III

Introduction

One somewhat surprising result of Experiment II was the equality of the thresholds for luminance and density modulations (expressed in terms of the amplitude of the fundamental modulation component). This is surprising because there is evidence for an early transduction non-linearity, possibly in the photoreceptors themselves (Boynton and Whitten, 1970; MacLeod *et al.*, 1985). If the intensities of individual dots were first passed through a compressive nonlinearity, the result should be sensitivity loss for luminance modulation. Although one might think that the small modulations at threshold hold would be little affected by such a nonlinearity, it should be remembered that the individual dots in Experiment II were at 100% contrast with their dark background; the local dot luminance was 9 times higher than the overall mean luminance. Since each dot had a luminance much higher that the space-average mean luminance, any one would be expected to produce a response near the end of the curve. The differential response to small perturbations of a dot's luminance would be proportional to the slope of the response function. Density modulation, on the other hand, affects the space-average luminance through area summation, which is not affected by transduction nonlinearities. Therefore we might expect to find *lower* thresholds for density modulation than for luminance modulation in densely sampled arrays. To test this hypothesis, Experiment II was repeated using a more flexible apparatus that allowed the different stimulus parameters to be manipulated with greater independence.

Procedure

The stimulus for Experiment III was similar to that used in Experiment II, but was generated by computer to simplify modification of parameters and allow a forced-choice method to be used to estimate thresholds. Stimuli were displayed on a cathode ray tube with P4 phosphor and hardware gamma correction (Hewlett-Packard Model No. 1332A, options 604 and 215). The hardware gamma correction provided good local linearity, but overall linearity correction was accomplished in software. A twelve bit digital-to-analog converter (DAC) generated voltages that were applied to the scope inputs. Three sample-and-hold amplifiers combined with a simple counter circuit allowed the single DAC to control all three scope inputs (two deflection or position inputs and z-axis or brightness). The DAC was attached to a digital computer (Digital Equipment Corp. PDP-11/23) which ran the experimental programs.

The relation between the DAC setting and the scope brightness was measured with a photodiode (United Detector Technologies PIN-10) in conjunction with a current-to-voltage amplifier. Quoted modulation values take account of the measured nonlinearity of the scope, which was not pronounced in the range of interest. Absolute measurements of the mean luminance were made with another photometer (EG&G model 450-1, equipped with multiprobe 550-2 and pulse integrator 550-3). The background was left "dark", so the contrast of individual dots was very close to 1.0, being diminished only by small amounts of electron scatter within the CRT, and intraocular light scatter. All experiments were done with the room lights off, to avoid reflections from the scope face. The display had a mean luminance of 16.6 candelas per meter squared, corresponding to 68.6 microcandelas per dot.

Stimuli consisted of regular lattices of 1024 dots (a 32×32 array). From the viewing distance of two meters, the array subtended a visual angle of 1.77 deg, with a corresponding dot spacing of 3.4 min of arc. A change of one least-significant-bit on the DAC generated a positional displacement of 2 sec of arc. Fifty frames of each stimulus were presented at a refresh rate of 57 Hz, for a total stimulus duration of 875 msec. Each stimulus was preceded by a fixation cross which appeared in the center of the screen for 500 msec immediately preceding the stimulus. This served both to alert the subject and to guide fixation and accommodation.

Nonuniformity was introduced into the arrays in one of four ways. The luminances of individual dots could be modulated, always as a one-dimensional sinusoidal function of position. Alternatively, the positions could be modulated, resulting in a modulation of dot density equal to the spatial derivative of the position modulation. Modulations of position and density could also be applied in conjunction, either in similar phase (the denser dots being brighter) or opposite phase (the denser dots being dimmer).

On any given trial, the subjects' task was to report the orientation of the modulation, which was chosen at random to be either horizontal or vertical. Modulation amplitudes on successive trials were determined in accordance with a staircase procedure designed to concentrate the trials about the amplitude corresponding to 71% correct. Each staircase was governed by the "2 to 1" rule; that is, the stimulus would be made more detectable after each incorrect response, whereas *two* consecutive correct responses were required to decrease the detectability. Each run began with a preliminary section during which the staircases used large increments to rapidly home in on the threshold level. The increment for each staircase was initially half of the range, and was decreased by half after each staircase reversal. Data were not recorded during this preliminary phase, which terminated when all staircase increments had reached the minimum level (0.1 log units). Four independent staircases were randomly interleaved in each block.

Two spatial frequencies of modulation were tested: 1.2 c/deg, for which two complete cycles were visible, and 0.6 c/deg, for which only a single cycle was visible. In order that the arrays maintain a constant size for all amplitudes of density modulation, the signal which modulated the dot's position was always applied in sine phase. This ensured that no displacements were ever applied to the dots at the edges of the pattern. This also meant that the luminance modulation was always applied in cosine phase, with the dimmest dots at the edges for positive modulation amplitudes, and the brightest dots at the edges or negative amplitudes. Since only a single grating bar was visible in the 0.6 c/deg condition, it was considered desirable to test both positive and negative modulations for this spatial frequency.

Subjects were undergraduate students with some prior experience at making psychophysical judgments. All were naive with respect to the purpose of the experiment. Thresholds were first measured separately for the two types of modulation (luminance and density). Once these thresholds were known, combined conditions were constructed, with the ratio of luminance modulation to density modulation in approximate proportion to their individual thresholds. Subjects ran at at least three sessions per condition; the threshold estimate for each session was based on either 200 trials (1.2 c/deg) or 100 trials (0.6 c/deg). A normal ogive anchored to 50% at $x = 0$ was fit to the observed probabilities using a weighted least-squares regression. Thresholds were estimated as the modulation amplitude for which the fitted curve assumed a value of 0.75.

To fit the data, the observed probabilities of correct response were first transformed to the corresponding normal deviates (z-scores). (In this representation, the normal ogive has the form of a straight line.) To constrain the regression line to pass through the origin, the reflection through the origin of each point was added to the data set. Ideally, each point should have been assigned the following weight

$$w_i = \frac{n_i \exp(-z_i^2)}{p_i(1-p_i)},$$

where n_i represents the number of observations at the ith stimulus level, p_i the true underlying probability of correct response, and z_i the normal deviate corresponding to p_i. These weights incorporate the fact that the observed probabilities will be distributed about the true underlying probability according to a binomial distribution with variance proportional to $[p_i(1-p)]/n_i$. The exponential term corrects for the fact that we are fitting in z-score space, although we want to minimize the squared errors between the observed and estimated probabilities. A derivation of these weights can be found in McKee *et al.* (1985), who attribute them to Finney (1971).

In practice, these weights cannot be calculated, since the true underlying probabilities are not known. An iterative procedure was adopted to avoid this difficulty. On the first iteration, the observed probabilities were used in place of the underlying probabilities. The estimates of the underlying probabilities formed on the basis of the first regression were used to compute the weights for the second regression. This procedure was repeated until successive estimates converged.

Results

Figures 6 and 7 show thresholds plotted in a space where the horizontal axis represents amplitude of luminance modulation and the vertical axis represents amplitude of density modulation. Negative amplitudes indicate a phase inversion of the dot modulation. The error bars represent plus and minus two standard errors of the mean based on variability between sessions. Figures 6(a) to (c) show the results of three subjects for a modulation frequency of 0.6 c/deg. At this spatial frequency, the screen was filled by only a single modulation cycle, and the luminance modulation thresholds show an unexpected phase dependence, a luminance decrement at the center of the screen (with an increment at the edges) being much less detectable than the reverse. This asymmetry may be explained in terms of the hypothesized compression of the response to the dots, if we assume that the amount of compression is related to adaptational state, which might vary across the stimulus field. Because the area surrounding the dot lattices was dark, the part of the retina which viewed the stimuli was relatively light adapted. Random variations in fixation would blur the edges of the sensitivity profile; since a bit of retina near the edge of the stimulus area would sometimes be in light, sometimes in dark, it would be less light-adapted than if it were always covered by the lattice, and would consequently be more apt to

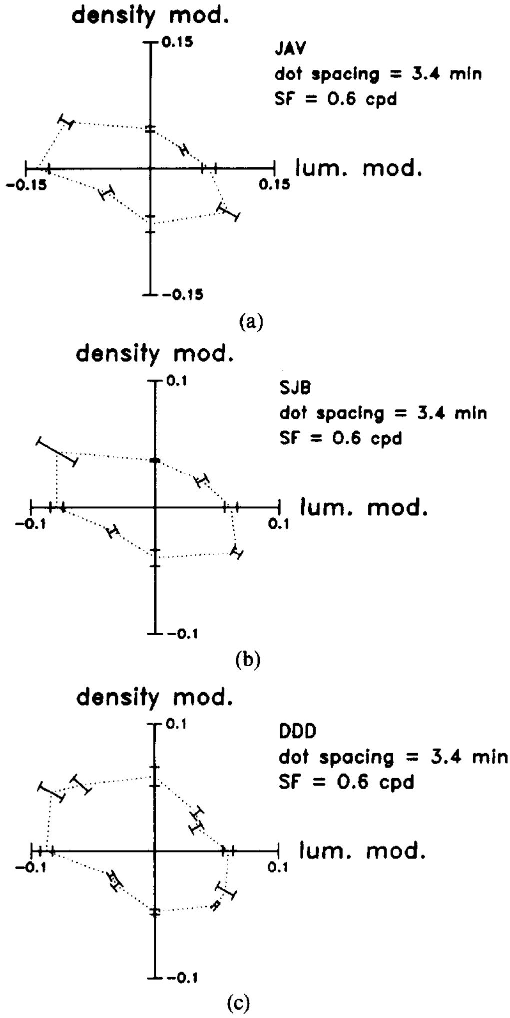

Fig. 6. As in Fig. 3, thresholds for detecting modulation are plotted in a two-dimensional space where the horizontal axis represents luminance modulation (applied to the lattice elements) and the vertical axis represents density modulation of the lattice elements. The modulation frequency was 0.6 c/deg, for which only a single (cosine phase) cycle was visible. Positive modulations represent the case where the center bar had increased brightness [density]. Dot spacing was 3.4 min of arc; error bars represent ± 2 SEM based on variability between sessions. (a) Subject J.A.V. (b) Subject S.J.B. (c) Subject D.D.D.

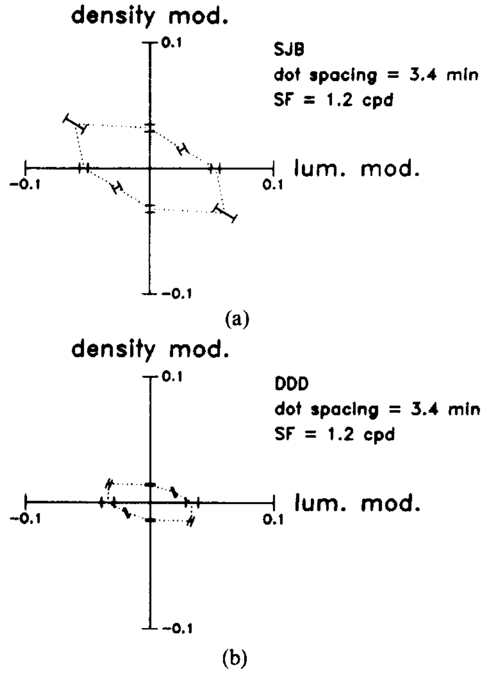

Fig. 7. Same as Fig. 6, but for a modulation frequency of 1.2 c/deg. Since more than one grating cycle was visible, only four points were measured, which were reflected to show the threshold contour. (a) Subject S.J.B. (b) Subject D.D.D.

saturate in response to a small bright dot. Regardless of the source, the asymmetry was never enough to make the luminance modulation threshold as low as the density threshold. Although this asymmetry is most pronounced for the case of actual luminance modulation, there is a hint of a smaller asymmetry in the same direction for the density modulation. The lack of a comparable effect for density modulation is also consistent with our explanation, provided that the nonlinearity precedes the spatial summation which we hypothesize to be responsible for the illusory brightening due to increased density. We suspect that the asymmetry would be reduced if a larger display were used which was capable of displaying more than one grating cycle.

Several other points concerning the data in Fig. 6 are noteworthy: first, the average threshold for seeing luminance modulation is higher (by almost a factor of 2) than that for seeing density modulation; unlike the results of Experiment II, these results are consistent with a compressive nonlinearity in the transduction of luminance. Secondly, modulations of luminance and density combine additively in quadrants 1 and 3, just as we saw in Experiment II. In quadrants 2 and 4, although we do not see quite as much actual cancellation as in Fig. 5, there is still clear sub-additivity: the thresholds are higher than what would be expected from independent encoding of dot luminance and density. Since the degree of cancellation was less than was seen in Fig. 5, (i.e. the components of the cancellation phase threshold stimulus were not significantly above their individual thresholds), it was thought that the optimal cancellation phase stimulus would be one in which the components were combined in proportion to their thresholds, rather than to maintain a constant space-average luminance. If we assume that the individual component thresholds correspond to a single threshold brightness difference, then making a cancellation phase stimulus where the component amplitudes are proportional to their thresholds should produce a stimulus having non-uniform luminance, but constant space-average *brightness*. Stimuli like this augmented the basic stimulus set for subject DDD [Fig. 6(c)]; no significant increase in the degree of cancellation was observed.

Figures 7(a) and (7b) show similar results for a modulation frequency of 1.2 c/deg. Because more than a single grating cycle was present, both phases (i.e. positive and negative modulation amplitudes) were not actually tested; four of the points are merely reflections of the other four. For determination of the on-axis thresholds, negative modulation amplitudes (bright dots at edges) were used to measure the brightness-only thresholds, while positive modulation (sparse dots at edges) was used to measure the density-only thresholds. For the off-axis measurements, negative amplitude brightness modulation was combined with both positive and negative amplitude density modulations, with amplitudes proportional to the individual thresholds. The results are quite similar to those seen for the lower modulation frequency, showing higher thresholds for luminance modulation relative to density modulation, additivity in quadrants 1 and 3, and even stronger cancellation in quadrants 2 and 4.

Discussion

The results of Experiment III have confirmed the interaction of density and luminance at high dot densities; in addition, the relation between thresholds for luminance and density modulation separately has been interpreted as evidence for a compressive nonlinearity in the transduction of luminance. A likely reason that the latter feature was not observed in the results of Experiment II is that the dots in Experiment II were dimmer by a factor of about 1.5 log units.

EXPERIMENT IV

Introduction

The results of Experiments I and II have been seen to be consistent with a model of brightness perception in which signals are gathered from a region having a radius slightly less than half a degree (the dot spacing at which interactions of density and brightness vanish). Experiment IV was performed to determine whether this spatial integration is simply summation over all space, or whether it is a more complicated integration of strictly figural properties, i.e. a *weighted* sum which includes the dot luminances, but not light from the surrounding area. This idea was suggested by an incidental observation in the experiments already described: increasing the local dot density can increase the subjective brightness of dots, but does not cause a dark background to appear subjectively lighter, implying that a separate brightness signal is needed to represent the level of the background. If the brightness at any point (in the background) were influenced similarly to that of the dots by the local space-average luminance at that point, then the dark background should appear brighter in a region populated with more dots. The fact that the (black) background does not appear lighter in regions of increased dot density suggests that separate signals are used to encode the brightnesses of the dots and their surround. Since the brightness of the surround seems unaffected by the luminances of the dots, we wondered whether the brightnesses of the dots would be similarly unaffected by modulations of the surround luminance.

We sought to answer this question by asking observers to detect luminance modulation in a regular lattice of dots where the modulation was applied either to the dots alone, to the dot surround alone, or to both the dots and the surround, in the same or opposite relative spatial phase. If observers could discount the local level of the background and base their judgements solely on the luminance of individual dots, or perhaps local dot contrast, then we would expect that having the background modulated in the opposite phase from the dots would maintain or improve sensitivity. Background modulations in the same phase should have no effect if dot luminance alone is relevant, or a small inhibitory effect if local dot contrast is important. If, on the other hand, observers do this task simply by detecting the Fourier component at the modulation frequency, then we expect in-phase dot and background modulations to combine additively to determine threshold; dot and background modulations that are out-of-phase, with amplitudes in the ratio producing *no* net energy at the fundamental, should produce large threshold elevations.

Stimuli were produced on a color monitor (Tektronix model 690SR), which received video signals from a graphics terminal (Advanced Electronic Devices model 767), which in turn was controlled by computer (Digital Equipment Corp. PDP 11/23). The display was viewed at a distance of 3 m, from which distance it subtended 4 deg of visual angle. The graphics unit provided 8 bit digital-to-analog converters (DAC's) for each phosphor. The 24 bit DAC programming values were stored in a hardware lookup table having 256 entries; this table was addressed by the 8 bit pixel values. Thus, although the displayed colors could be chosen from a set of some 16 million distinct shades (distinct to the program at least), only 256 of these could be displayed simultaneously, due to the size of the pixel word and lookup table. The hardware limitation to a palette of 256 or fewer "colors" did not affect the display of the stimuli in these experiments, since the patterns were modulated in only one dimension and were periodic.

In order to decrease the digital quantization errors in the rendering of the luminance profile, which were limited by the video digital-to-analog converter resolution (8 bits per phosphor), the display was viewed through a red filter (a double-density of Kodak Wratten No. 26). This filter had the effect of selectively attenuating the light from the green phosphor, and making it approximately the same reddish color as the light from the red phosphor; the smallest test modulations could therefore be produced by varying the output of the green phosphor, with a high contrast background modulation provided by light from the red phosphor. In general, the size of the chromatic artifact introduced by this technique depends on the spectral transmission of the viewing filter. The particular case of these phosphors and filter has been quantitatively analysed by Mulligan (1986). Briefly, since the No. 26 filter has a rather sharp cutoff in the red, the chromaticity of the light from the green phosphor is not much different from that of the red (CIE coordinates $x = 0.693$, $y = 0.306$ for the red, $x = 0.669$, $y = 0.330$ for the green, for a single density of Wratten No. 26). In this situation, chromatic differences are visible only between mixtures containing very little light from the red phosphor (since the red light will dominate the filtered mixture). This was never the case for the low-contrast threshold stimuli. Furthermore, the green phosphor was not used to create all of the modulation, but merely a fine correction to the luminance produced by the red phosphor.

The dots had a mean luminance of 20 cd/m^2. The mean background luminance was one tenth of this. Since the background occupied 8 times as much area as the dots, the modulation amplitude of the background was reduced by a factor of 8 relative to the amplitude of the dot modulation in the combined cases. However, since the background mean luminance was lower by a factor of ten, the *contrast* of the background grating was actually 10/8 or 1.25 times higher than the dot grating having the same power at the modulation frequency.

Individual dots were larger than those used in the preceding experiments, and were not in general spatially uniform. The area of each dot was effectively a window through which a continuous grating was seen. An independent grating was windowed by the area surrounding the dots. Each stimulus presentation was preceded by the appearance of a fixation cross in the center of the screen and an auditory signal. The fixation target consisted of a small bright cross on a dark field slightly larger than the cross, embedded in a surround having the same mean luminance as the stimulus background. The fixation stimulus was visible for 500 msec, and was immediately followed by the stimulus which was displayed for 125 msec. Following the stimulus the screen displayed a uniform field with a luminance equal to the space-average luminance of background.

Thresholds were determined by having subjects discriminate between vertical and horizontal modulations. Subjects were shown a single stimulus and had to report the orientation, which was chosen at random. Modulation levels for successive trials were determined in accordance with a staircase procedure. Two staircases were randomly interleaved for each condition to minimize the amount of *a priori* information available to the subjects about the presentation for any given trial. The orientation of the modulation was chosen at random for each trial. Thresholds were estimated from the raw data using the same curve fitting procedure as in Experiment III.

Different modulation frequencies and dot spacings were run in different blocks of trials. The blocks for the different conditions were randomly interleaved, and each of the two subjects ran three blocks for each condition. Thresholds were measured for three modulation frequencies (0.625, 1.25, and 2.5 c/deg), each at three dot spacings (3, 6 and 12 min of arc). The two subjects were both experienced psychophysical observers; only subject KFP was naive concerning the purposes of the experiments.

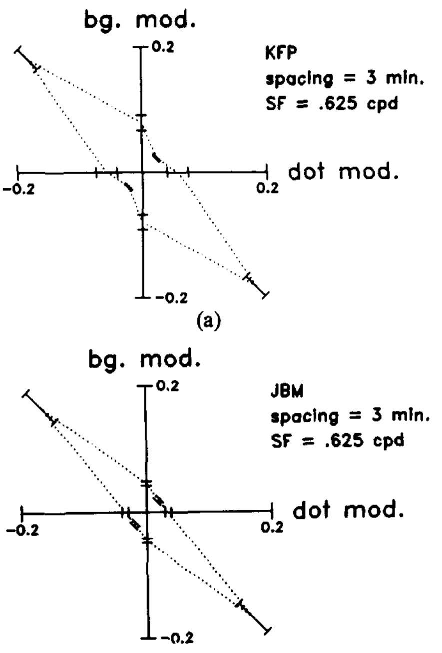

Fig. 8. Modulation thresholds plotted in a space where the horizontal axis represents amplitude of modulation applied to dots, and the vertical axis represents amplitude of modulation applied to the background. Mean background luminance was one tenth that of the dots, but the background occupied eight times the area of the dots. Axes are normalized to reflect area-weighted amplitude of the modulation. Dot spacing was 3 min of arc, modulation frequency 0.625 c/deg. Error bars represent ± 2 SEM based on variability between sessions. (a) Subject K.F.P. (b) Subject J.B.M.

Results

Data are shown in Figs 8–11. Thresholds are plotted in a space where the horizontal axis represents the amplitude of the dot modulation, and the vertical axis the surround modulation. The axes have been normalized to represent equal area-weighted amplitudes of modulation; a dot modulation of 1 means that the dot luminance varied from zero to twice the mean dot luminance. The background modulations have been equated with the dot modulations on the basis of Fourier fundamental amplitude. In all cases except the Nyquist case, this is approximately equivalent to defining background modulation x as having a mean-to-peak amplitude 0.125 times the amplitude of a dot modulation of x, since the background occupied 8 times as much area as the dots. When sampling at the Nyquist rate the power in the sampled waveform depends on the placement of the samples. In our display the samples fell at the peak and trough of the sampled waveform, with the result that the amplitude of the Fourier fundamental

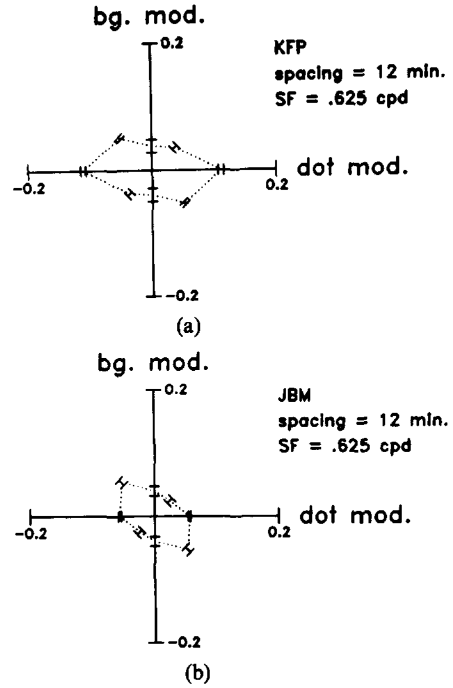

Fig. 9. Same as Fig. 8, but for a dot spacing of 12 min. (a) Subject K.F.P. (b) Subject J.B.M.

is increased by a factor close to 2 (see Appendix II). Because the background was a factor of 10 dimmer than the dots, the largest background modulation that was possible was 0.8.

Figure 8(a) and (b) show two subjects' thresholds for a lattice of high density (3 min element spacing) and a modulation frequency of

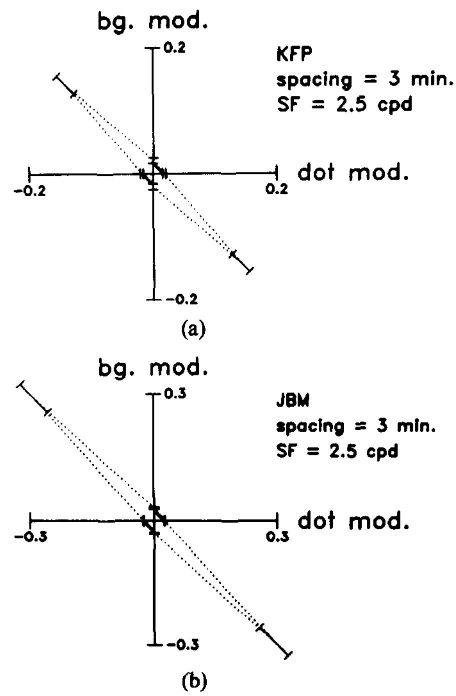

Fig. 10. Same as Fig. 8, but for a modulation frequency of 2.5 c/deg. (a) Subject K.F.P. (b) Subject J.B.M.

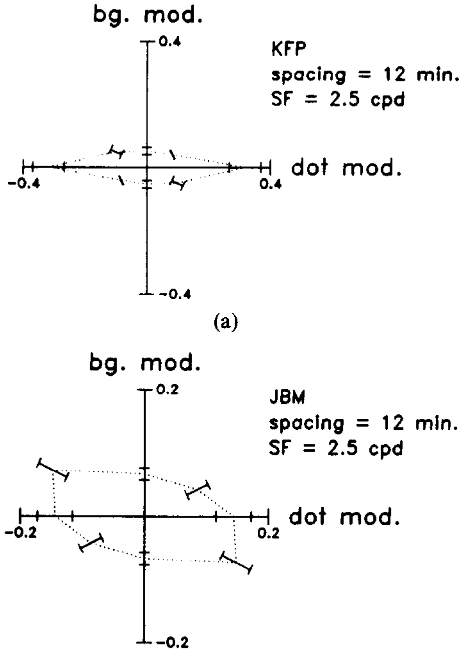

Fig. 11. Same as Fig. 10, but for a dot spacing of 12 min. (a) Subject K.F.P. (b) Subject J.B.M.

0.625 c/deg. We observe that the threshold, with the appropriate scaling, is roughly independent of whether the modulation is applied to the dots or to the background. When the modulation is applied in the same phase to both the dots and the background (quadrants 1 and 3), threshold is well-predicted on the basis of the amplitude of the Fourier component at the moduation frequency in the image; that is, the data points fall on lines of negative unity slope. When the modulations are applied in cancelling phase (quadrants 2 and 4), threshold is greatly elevated; since there is no energy at the fundamental in this case, subjects are detecting a variation in the local dot contrast, but the fact that their threshold is several times higher than the others suggests that in the other cases local dot contrast is irrelevant, and that it is the Fourier component at the modulation frequency that mediates detection.

Figures 9(a) and (b) show analogous results for a dot spacing of 12 min. Subject KFP [Fig. 9(a) now shows an elevation of thresholds for dot modulation compared to background modulation, similar to the results of Mulligan and MacLeod (1984). The thresholds for the combined conditions now show much less of a difference. Although the thresholds in quadrants 2 and 4 are still slightly higher than those in quadrants 1 and 3, all may be predicted fairly well on the basis of the amplitude of the background component. Subject J.B.M. [Fig. 9(b)] shows less of an elevation for the dot modulation only condition, and more residual asymmetry in the oblique thresholds.

Results for a modulation frequency of 2.5 c/deg are shown in Figs 10 and 11. For a dot spacing of 3 min [Fig. 10(a) and (b)] the elliptical threshold contours are even more eccentric than those for the lower modulation frequency. When the dot spacing is increased to 12 min of arc (so that sampling is at the Nyquist rate of two samples per cycle) the effect of this is as much as at 0.625 c/deg: subject K.F.P. [Fig. 11(a)] shows a large increase in the dot modulation thresholds, and the thresholds for the oblique conditions can be predicted on the basis of the amplitude of the background component. The results for subject J.B.M. [Fig. 11(b)] are qualitatively similar, but less pronounced.

Discussion

These results show that under these conditions observers are most sensitive to the amplitude of the fundamental component of luminance, and relatively insensitive to changes in the local dot contrast. They also refute our initial conjecture that spatial integration of brightness signals occurred independently for the dots and the surround, for if that were true the threshold for the cancellation phase condition would not show a large elevation, since the brightnesses of individual dots would be unaffected by modulation of the surround.

As has been noted elsewhere (Mulligan and MacLeod, 1984; Nothdurft, 1985), thresholds for coarsely sampled modulation (dot modulation with large dot spacings) are greatly elevated compared to those obtained with fine sampling or continuous field stimulation; thresholds for surround modulation, on the other hand, are relatively unaffected by the spacing of the superimposed dots. In the present case, the frequency components added by the sampling operation are approximately equal regardless of whether a given small modulation is applied to the dots or the surround (see Appendix). In light of this, the result that the two types of stimuli produce different threshold elevations supports the claim of Mulligan and MacLeod that the elevation for dot modulation under coarse sampling is not just a case of masking by the frequency components in the sampling lattice. Any model in which detection or discrimination depends on the power spectra

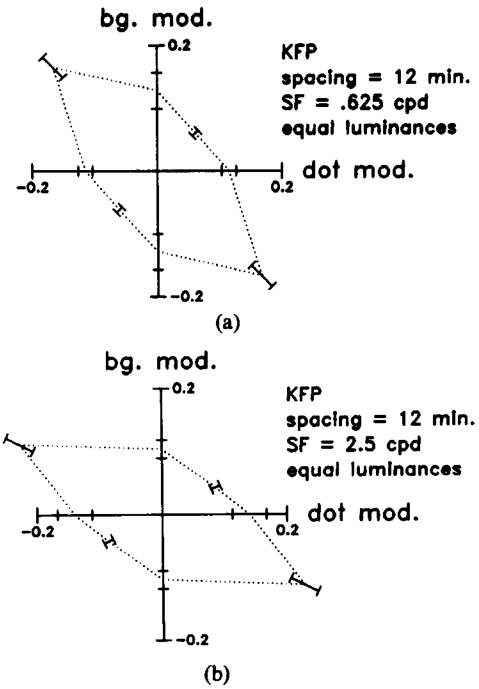

Fig. 12. Similar to Figs 9 and 11, but for a background having the same mean luminance as the dots. Dot spacing 12 min of arc, subject K.F.P. (a) Modulation frequency = 0.625 c/deg. (b) Modulation frequency = 2.5 c/deg.

of the images predicts similar thresholds in the two cases, and is therefore contradicted by the results.

Why is the threshold higher for modulation of the dots only than for modulation of the background only? One interpretation is that the modulation is attenuated at an early stage by saturation of spatially-opponent units. To test this idea, Experiment IV was repeated, but with a background having a luminance matched to the space-average luminance of the dots. If the coarsely spaced dots are saturating spatially-opponent units because of a lack of surround stimulation from the dark background, then using an equiluminous surround should restore the units' sensitivity by restoring the surround input.

The results of this modified experiment are shown in Fig. 12(a) (modulation frequency = 0.625 c/deg) and 12(b) (modulation frequency = 2.5 c/deg), both for a dot spacing of 12 min (the spacing for which the thresholds for dot and background modulation were the most different). As in the case of equal luminance, we see summation in quadrants 1 and 3, and even greater cancellation in quadrants 2 and 4 [compare with Figs 9(a) and 11(a)]. The approximate equality of the thresholds for the background and dot modulations under these conditions is consistent with the hypothesis of saturation of spatially-opponent nerurons when the background is dark or comparatively dim, since making the surround equiluminant would sensitize units saturated by the spots on the dark surround.

GENERAL DISCUSSION

A number of experiments have been described in this paper which suggest that there is a summation area for brightness having a diameter slightly less than 1 deg. The observation made in Experiment I, that the brightness increase caused by increasing dot density seems to be confined to the dots, caused us to believe that this integration might occur independently for figure and ground; the results of Experiment IV, however, showing that dot and surround modulations cancel each other for small dot spacings refuted this conjecture.

In Experiment IV a saturating nonlinearity in a spatially-opponent mechanism has been suggested as the cause of insensitivity to luminance modulation of discrete dots. With that idea in mind, it seems paradoxical that in dense dot arrays, increasing the dot density makes dots appear *brighter;* one might suppose that the primary effect of increasing the dot density would be to increase surround stimulation, thereby lowering net excitation. Sensitization experiments (Westheimer, 1967) suggest that in the central fovea receptive field centers have diameters of only a few minutes of arc, and inhibitory surrounds have diameters of about 10–15 min of arc; the results of Experiments I and II, however, demonstrate summation over a much larger area. It is obvious that the two sets of results cannot be due to the same neural mechanisms. An interesting question to which an answer is not yet forthcoming is whether the relevant mechanisms are located at different stages in a single pathway, or reflect the existence of completely separate (parallel) pathways or channels. However, spatial integration of brightness signals cannot occur before the separation from signals used for the extraction of individual dot contours, since the dots can be sharply resolved even when their brightnesses interact.

In conclusion, we have observed effects where the brightness of distinctly resolved elements are affected by the positions of nearby elements. The results are consistent with a model where brightness signals are integrated over an area slightly under a degree in diameter, a size which is different from that implied by the outwardly similar phenomenon of assimilation. Dramatic decreases in sensitivity in sparse dot arrays are consistent with an early compressive non-linearity, possibly in the photoreceptors themselves. As would be expected from such a model, sensitivity to luminance differences between distinct elements depends on their spacing, as well as on the spatial scale of the luminance modulation.

Acknowledgements—Experiment I was performed in the laboratory of Bela Julesz during a short visit to Bell Laboratories, Murray Hill, NJ; the help of Walter Kropfl was invaluable in taming the equipment in the short time available. Thanks to Bruce Henning and Andrew Stockman for commenting on the manuscript. This work was supported by NIH grant EY-01711 award to DIAM and an IBM graduate fellowship awarded to J.B.M.

REFERENCES

Boynton R. M. and Whitten D. N. (1970) Visual adaptation in monkey cones: recordings of late receptor potentials. *Science, N.Y.* **170,** 1423–1426.

Breitmeyer B. Julesz B. and Kropfl W. (1975) Dynamic random-dot stereograms reveal up-down anisotropy and left-right isotropy between cortical hemifields, *Science, N.Y.* **187,** 269–270.

Crick F. H. C. (1984) Function of the thalamic reticular complex: The searchlight hypothesis. *Proc. natn. Acad. Sci. U.S.A.* **81,** 4586–4590.

Finney D. J. (1971) *Probit Analysis.* Cambridge Univ. Press.

Green D. M. and Swets J. A. (1974) *Signal Detection Theory and Psychophysics.* Wiley, New York.

Helson H. (1963) Studies of anomalous contrast and assimilation. *J. opt. Soc. Am.* **53,** 179–184.

Julesz B. (1980) Spatial nonlinearities in the instantaneous perception of textures with identical power spectra. *Phil. Trans. R. Soc. Lond. Ser. B* **290,** 83–94.

Bergen J. R. and Julesz B. (1983) Parallel versus serial processing in rapid pattern discrimination. *Nature, Lond.* **303,** 696–698.

MacLeod D. I. A., Williams D. R. and Makous W. (1985) Difference frequency gratings above the resolution limit. *Invest. Ophthal. visual Sci.* **26,** 11.

McKee S. P., Klein S. A. and Teller D. Y. (1985) Statistical properties of forced-choice psychometric functions: Implications of probit analysis. *Percept. Psychophys.* **37,** 286–298.

Mulligan J. B. (1986) Optimizing stereo separation in color television anaglyphs. *Perception* **15,** 27–36.

Mulligan J. B. and MacLeod D. I. A. (1984) Pooling of perceived brightness among distinct texture elements. *Invest. Ophthal. visual Sci., Suppl.* **25,** 199.

Nakayama K. and Silverman G. H. (1986), Serial and parallel encoding of feature conjunctions. *Invest. Ophthal. visual Sci., Suppl.* **27,** 182.

Nakayama K., Silverman G. H., MacLeod D. I. A. and Mulligan J. B. (1985) Sensitivity to shearing and compressive motion in random dots. *Perception* **14,** 225–238.

Nakayama K. and Tyler C. W. (1981) Psychophysical isolation of movement sensitivity by removal of familiar position cues. *Vision Res.* **21,** 427–433.

Nothdurft H. C. (1985) Sensitivity for structure gradient in texture discrimination tasks, *Vision Res.* **25,** 1957–1968.

Ratliff F. (1965) *Mach Bands: Quantitative Studies on Neural Networks in the Retina.* Holden-Day, N.Y.

Rentschler I., Hubner M., Caelli T. and Stoerig P. (1986) Structural masking revealed by extrafoveal image recognition. *Invest. Ophthal. visual Sci., Suppl.* **27,** 95.

Rodieck R. W. (1983) Raster-based colour stimulators. In *Colour Vision* (Edited by Mollon J. D. and Sharpe L. T.) pp. 131–144. Academic Press, New York.

Treisman A. M. and Gelade G. (1980) A feature-integration theory of attention *Cogn. Psychol.* **12,** 97–136.

Treisman A. and Schmidt H. (1982) Illusory conjunctions in the perception of objects. *Cogn. Psychol.* **14,** 107–141.

Westheimer G. (1967) Spatial interaction in human cone vision. *J. Physiol., Lond.* **190,** 139–154.

APPENDIX I

Derivation of Predicted Threshold Decrease for Independent Encoding of Luminance and Density

We assume that the stimuli of Experiment II are represented in a perceptual space whose dimensions are local dot luminance and local dot density. We also assume that these dimensions are independent, and that the response to a particular component of a given stimulus is normally distributed about a mean which is proportional to the stimulus level, and with constant variance. We analyse the detection of such stimuli using the framework of signal-detection theory (Green and Swets, 1974); that is, the detection thresholds will be assumed to correspond to some criterion separation of the response distributions for the signal-present and signal-absent cases.

Let us begin by normalizing the units of luminance and density modulation so that each has a value of 1 at threshold. This has the effect of equalizing the variances of the response distributions in the two dimensions. We can represent the response space by a coordinate system where the x-axis represents luminance modulation and the y-axis represents density modulation. The distribution of responses corresponding to no modulation is centered at the origin. Since the units have been scaled to equate the variances in the two dimensions, this distribution is simply a circularly symmetric unit variance Gaussian. Applying modulation has the effect of shifting this distribution away from the origin. The detectability, d', is proportional to the shift and is constant (independent of direction) at threshold.

Now, for combination stimuli where the two components of the modulation are combined in amounts proportional to the individual thresholds, the response distributions will fall on the line through the origin with slope 1. How far out must we go to reach threshold? Since we have normalized the units so that the Gaussian distributions are circularly symmetric, the value of d' will depend only on the distance between the distribution centers. Since thresholds for the individual thresholds have center separations of 1 (by our choice of units), we seek a point on the line of slope 1 that is exactly one unit from the origin: this is the point whose coordinates are $(\sqrt{2}/2 \cdot \sqrt{2}/2)$. Thus, for the combined threshold, the individual components have an amplitude a factor of $\sqrt{2}$ below their individual thresholds. *Q.E.D.*

APPENDIX II

Frequency Analysis for Sample and Background Modulations

One-dimensional case. In this appendix we will analyse the spectral content of sample arrays where modulation is applied to either the samples themselves, or to the background. For simplicity, we will first analyse the one-dimensional case. Let the sampling function in one dimension be represented by $s_1(x)$. Let $S_1(\omega)$ represent the Fourier transform of $s_1(x)$. The luminance profile of an array of uniform samples having luminance I_s on a surround or background having luminance I_b can be written

$$l_0(x) = I_s s_1(x) + I_b[1 - s_1(x)].$$

Now, let us add some modulation at spatial frequency ω_0 to the samples:

$$l_s(x) = I_s s_1(x)[1 + C_s\cos(\omega_0 x)] + I_b[1 - s_1(x)]. \quad (1)$$

The Fourier transform of $l_s(x)$, which we shall refer to as $L_s(\omega)$, can be expressed

$$\begin{aligned} L_s(\omega) &= I_s S_1(\omega)*[\delta(\omega) + \tfrac{1}{2}C_s\delta(\omega - \omega_0) \\ &\quad + \tfrac{1}{2}C_s\delta(\omega + \omega_0)] + I_b\delta(\omega) - I_b S_1(\omega). \\ &= I_b\delta(\omega) + (I_s - I_b)S_1(\omega) + \tfrac{1}{2}I_s C_s S_1(\omega - \omega_0) \\ &\quad + \tfrac{1}{2}I_s C_s S_1(\omega + \omega_0). \quad (2) \end{aligned}$$

We may similarly consider modulating the background:

$$l_b(x) = I_s s_1(x) + I_b[1 - s_1(x)][1 + C_b\cos(\omega_0 x)], \quad (3)$$

$$\begin{aligned} L_b(\omega) &= I_s S_1(\omega) + I_b[\delta(\omega) - S_1(\omega)] \\ &\quad *[\delta(\omega) + \tfrac{1}{2}C_b\delta(\omega - \omega_0) + \tfrac{1}{2}C_b\delta(\omega + \omega_0)]. \\ &= I_b\delta(\omega) + (I_s - I_b)S_1(\omega) \quad (4) \\ &\quad + \tfrac{1}{2}I_b C_b[\delta(\omega - \omega_0) - S_1(\omega - \omega_0)] \\ &\quad + \tfrac{1}{2}I_b C_b[\delta(\omega + \omega_0) - S_1(\omega + \omega_0)]. \end{aligned}$$

Let us consider the values of L_s and L_b at the modulation frequency, $\omega = \omega_0$

$$L_s(\omega_0) = \tfrac{1}{2}I_s C_s S_1(0) + \tfrac{1}{2}I_s C_s S_1(2\omega_0).$$

$$L_b(\omega_0) = \tfrac{1}{2}I_b C_b[1 - S_1(0)] - \tfrac{1}{2}I_b C_b S_1(2\omega_0).$$

We want to normalize C_s relative to C_b so that the two stimuli have the same amplitude of the modulation fundamental frequency, ω_0. By equating the expressions for $L_s(\omega_0)$ and $L_b(\omega_0)$, we obtain the following expression for the ratio of the contrasts

$$\frac{C_s}{C_b} = \frac{I_b}{I_s}\frac{1 - S_1(0) - S_1(2\omega_0)}{S_1(0) + S_1(2\omega_0)} \quad (5)$$

Note that for sampling frequencies above the Nyquist limit, $S_1(2\omega_0) = 0$. For the sampling signals consisting of unity height pulses, $S_1(0)$ is simply the duty cycle of the pulses; thus the quotient on the far right represents the ratios of the areas occupied by the samples and the background (with a correction when sampling at the Nyquist frequency); to simplify notation we will represent this quantity by the symbol k_{bs},

$$k_{bs} = \frac{1 - S_1(0) - S_1(2\omega_0)}{S_1(0) + S_1(2\omega_0)}.$$

We now would like to know what are the relative amplitudes of frequency components other than the modulation fundamental after this normalization. We substitute the value of C_s from equation (5) into equation (2)

$$\begin{aligned} L_s(\omega) &= I_b\delta(\omega) + (I_s - I_b)S_1(\omega) \\ &\quad + \tfrac{1}{2}k_{bs}I_b C_b S_1(\omega - \omega_0) + \tfrac{1}{2}k_{bs}I_b C_b S_1(\omega + \omega_0). \end{aligned}$$

Comparison of equations (4) and (5) shows that the two spectra receive the same direct contributions from the sampling lattice (which do not depend on the modulation contrast); the replicas of the sampling spectrum which are frequency-shifted by the modulation frequency have opposite relative phases in the two stimuli, however. The amplitudes of these sidebands are proportional to the modulation contrast, and are equal only when $k_{bs} = 1$.

We may subtract equation (4) to find the difference energy

$$\begin{aligned} L_s(\omega) - L_b(\omega) &= \tfrac{1}{2}I_b C_b(k_{bs} + 1)S_1(\omega - \omega_0) \\ &\quad + \tfrac{1}{2}I_b C_b(k_{bs} + 1)S_1(\omega + \omega_0). \end{aligned}$$

Two-dimensional case. The above analysis can be generalized to two dimensions as follows: for a regular square lattice of samples, the two-dimensional sampling function, $s_2(x, y)$, can be simply expressed as the product of two one dimensional functions

$$s_2(x, y) = s_1(x)s_1(y),$$

with $S_2(\omega_x, \omega_y)$ representing the Fourier transform of $s_2(x, y)$

$$S_2(\omega_x, \omega_y) = S_1(\omega_x)*S_1(\omega_y).$$

The definition of the two-dimensional sampled grating is similar to equation (1)

$$l_s(x, y) = I_s s_2(x, y)[1 + C_s\cos(\omega_0 x)] + I_b[1 - s_2(x, y)].$$

Similarly

$$l_b(x, y) = I_s s_2(x, y) + I_b[1 - s_2(x, y)][1 + C_b\cos(\omega_0 x)].$$

By following the same development as in the one-dimensional case, it is easy to show that

$$\begin{aligned} L_s(\omega_x, \omega_y) - L_b(\omega_x, \omega_y) &= \tfrac{1}{2}I_b C_b(k_{bs} + 1)S_2(\omega_x - \omega_0, \omega_y) \\ &\quad + \tfrac{1}{2}I_b C_b(k_{bs} + 1])S_2(\omega_x + \omega_0, \omega_y), \end{aligned}$$

where

$$k_{bs} = \frac{1 - S_2(0, 0) - S_2(2\omega_0, 0)}{S_2(0, 0) + S_2(2\omega_0, 0)}.$$

Since $S_2(0, 0) = [S_1(0)]^2$, the numerical value of k_{bs} is larger in the two-dimensional case; in the present experiments, where the duty cycle in one dimension was 1/3, $k_{bs} = 8$ when sampling above the Nyquist frequency. Although this suggests a large difference between the two cases, it should be remembered that the aliased replicas of the sampling spectrum are attenuated by the modulation contrast; additionally, they are weighted by I_b, the background luminance. The magnitude of the unaliased spectrum, on the other hand, does not depend on the modulation contrast, and is weighted by $(I_s - I_b)$. When the background luminance is small compared to the sample luminance, the contrasts of the aliased components will be negligible compared to that of the unaliased components, even when threshold is maximally elevated. The asymmetry does become important when $I_b = I_s$, as in the final phase of Experiment IV (Fig. 12).