Pergamon

PII: S0042-6989(96)00147-2

Vision Res., Vol. 37, No. 1, pp. 45–62, 1997

Printed in Great Britain
0042-6989/97 $17.00 + 0.00

Separate Detection of Moving Luminance and Contrast Modulations: Fact or Artifact?

ANDREW T. SMITH,*‡ TIMOTHY LEDGEWAY*†

Received 17 January 1996; in revised form 25 March 1996

We have investigated first-order artifacts in second-order motion perception. Subjects were required to identify the orientation and direction of a drifting sinusoidal contrast modulation. When the carrier consisted of static two-dimensional noise, performance often reflected the use of first-order artifacts that arise from stochastic local biases in the noise, rather than the detection of the contrast modulation *per se*. This stimulus, which has been used widely for studying second-order motion, therefore appears to be inappropriate for that purpose. In contrast, global distortion products arising from luminance non-linearities do not appear to provide usable artifacts. Two manipulations were employed to eliminate local first-order artifacts: the use of dynamic noise and the use of high-pass filtered static noise. These two manipulations gave similar results, which were quite different from those obtained with broadband static noise. We argue that performance with both of these image types reflects the activity of a true second-order motion mechanism. A characteristic property of this mechanism is that it cannot specify direction at the threshold for detecting orientation. Direction thresholds are around 50% higher than orientation thresholds when first-order artifacts are eliminated. Copyright © 1996 Elsevier Science Ltd

Second-order motion   Contrast modulation   Direction perception

INTRODUCTION

When the contrast of a static pattern is modulated by a drifting sinusoid, motion is easily visible despite the fact that the motion is defined by contrast (a second-order image characteristic) rather than by luminance (a first-order characteristic). Second-order motion stimuli are of interest because they are not expected to activate conventional motion detectors, such as the motion energy detector of Adelson and Bergen (1985). This is because such detectors rely on the presence of spatial Fourier components in the luminance domain which move consistently in one direction. Second-order stimuli contain moving luminance components but these are normally expected to be equal in opposite directions along any given axis, so that any motion-detection mechanism which operates by detecting an imbalance of motion energy in opposite directions will fail to respond. This state of affairs is made clear in the theoretical treatment provided by Chubb and Sperling (1988), who refer to stimuli in which luminance motion energy is equal in both directions along the axis of motion as "drift-balanced".

*Department of Psychology, Royal Holloway College, University of London, Egham, Surrey TW20 0EX, U.K.
†Present address: Department of Experimental Psychology, University of Oxford, South Parks Road, Oxford, U.K.
‡To whom all correspondence should be addressed [*Email* a.t.smith@rhbnc.ac.uk].

A simple type of contrast modulation is an amplitude-modulated (AM) grating. If a sine grating, referred to as the carrier, is multiplied by another sine grating of the same orientation but a much lower spatial frequency, the result is a grating of the same spatial frequency as the carrier whose contrast is modulated sinusoidally at the frequency of the modulating grating. The Fourier spectrum contains just three spatial frequency components, all at the same orientation: one at the carrier frequency and the other two equidistant from the carrier, one at a higher and one at a lower frequency. If the modulating waveform is now made to drift while the carrier remains stationary, the result in Fourier terms is that the carrier temporal frequency remains at zero while the two sidebands take on non-zero temporal frequencies in opposite directions of motion. Since the two sidebands have the same amplitudes, the motion energy in the image is equal in the two directions. The same is true of beat stimuli, which are the same as AM gratings except that they lack the stationary centre frequency, being simply the sum of two oppositely drifting gratings. A number of empirical studies have been conducted using AM gratings and beats [e.g. Badcock & Derrington (1985); Derrington & Badcock (1985); Derrington *et al.* (1993)]. The results of these experiments suggest that this type of second-order motion is detected by a mechanism which is distinct from that which detects first-order motion, raising the possibility of two, perhaps parallel, motion detection systems. Chubb and Sperling (1988) built on this work by presenting a model of second-order motion detection in which the image is rectified so as to introduce luminance components into the neural representation of the image that are correlated with the contrast components in the original image and can be detected by conventional motion energy analysis. Wilson *et al.* (1992) further developed this notion by developing a model of motion perception comprising two parallel pathways, one based on conventional motion energy detection and the other on a non-linear transformation of the image followed by energy detection, the outputs of which are then combined to yield a single motion signal. Direct evidence that second-order motion is normally detected using the motion energy principle comes from experiments using stimuli in which the direction perceived is as predicted by this principle but not by alternative principles (Smith, 1994).

It is believed by most researchers that motion is initially detected by neurons which are tuned for spatial frequency, and this belief is encompassed in conventional models of motion detection. Consequently it is not sufficient, in order to avoid detection by standard motion analysis, for the total motion energy in the image to be equal in opposite directions. It is also necessary for motion energy to be equal in opposite directions at each spatial scale considered alone. In the case of an AM grating, the spatial frequencies of the two sidebands are slightly different. It is in most cases reasonable to assume that they are sufficiently close in frequency that they will both fall within the bandwidth of a single spatial frequency channel and so motion will not be detected. Nonetheless, it may be prudent, for empirical studies, to use an image in which the balance is perfect. Such an image is contrast modulated noise. If the carrier is composed of static noise with a flat spatial frequency spectrum, then contrast modulation results in drifting luminance components that are equal in opposite directions at all spatial frequencies. The noise could in principle be one-dimensional, although the use of one-dimensional noise would necessitate great care to ensure that the noise does indeed have a flat spectrum. In practice, two-dimensional noise is preferable because the presence of many different noise samples in a given spatial dimension greatly reduces the impact of any samples that do not have a flat spectrum, and two-dimensional noise has been used almost exclusively in empirical studies using contrast-modulated noise.

In the past few years, a substantial number of empirical studies of second-order motion have been conducted using contrast-modulated two-dimensional noise images [e.g. Victor & Conte (1992); McCarthy (1993); Smith *et al.* (1994); Smith (1994); Ledgeway & Smith (1994, 1995); Ledgeway (1994); Solomon & Sperling (1995); Johnston & Clifford (1995)]. In all these studies it has been assumed that such stimuli are pure second-order motion stimuli. However, it is possible for first-order motion artifacts to arise in second-order motion images. Normally, these artifacts correlate in direction and speed with the second-order motion characteristics, making it difficult to distinguish them perceptually. There is therefore a continual risk of inadvertently studying first-order motion sensitivity when attempting to study second-order motion sensitivity. In this paper two potential first-order artifacts that may arise in supposedly pure second-order images involving contrast-modulated noise are described. Empirical studies are then presented in which the extent to which these artifacts actually occur in practice and give rise to misleading conclusions is examined.

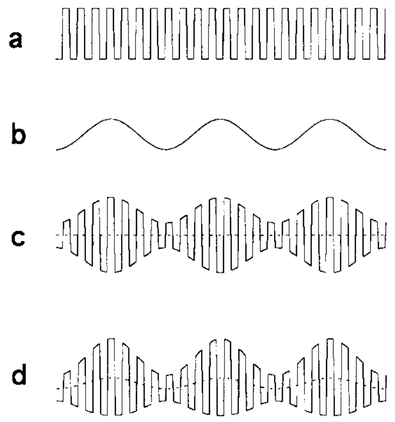

FIGURE 1. Luminance waveforms illustrating global distortion arising from non-linear processing. All plots show luminance as a function of spatial location and represent a slice through a two-dimensional image. (a) A sample of binary noise composed of light and dark pixels (to facilitate comparison with Fig. 2 the noise is represented as a regular array of pixels, not as a true noise waveform). (b) A sinusoid. (c) Contrast-modulated noise formed by multiplying the binary noise shown in (a) by the (raised and scaled) sinusoid in (b). The mean (space-average) contrast is the same as the contrast of the original noise and the modulation depth (determined by the amplitude of the sinusoid) is 50%. The dotted line shows the local mean luminance, averaged over a patch of adjacent light and dark pixels. Because this locally averaged luminance value is the same at all points in the image, there will be no change in its spatial profile when the modulating sinusoid moves. (d) The effect of passing the waveform in (c) through an expansive brightness non-linearity. The dotted line shows locally averaged luminance, as in (c). The locally averaged luminance now has a sinusoidal waveform, which will be present in the Fourier spectrum of the image as a component at the modulation frequency. This waveform will move when the modulating sinusoid moves, giving first-order motion. The amplitude of the waveform (and hence its detectability) is proportional both to the amplitude of the sinusoid and to that of the noise carrier.

Global distortion products

The first potential artifact is already well known. If a contrast-modulated image passes through a processing stage involving a non-linear luminance transformation, then a first-order modulation will be introduced which will have the same waveform as the contrast modulation and will move when it moves (see Fig. 1). This luminance modulation, referred to as a distortion product, will then be visible to conventional motion energy detectors. If psychophysical judgements of the image are based on distortion products then clearly such judgements cannot be taken as reflecting the properties of a second-order motion detection system. Assuming that the non-linearity

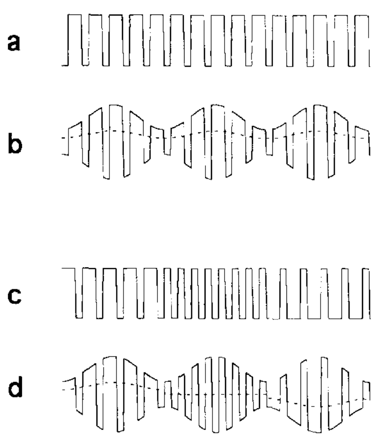

FIGURE 2. Luminance waveforms illustrating distortions arising from the stochastic properties of the carrier. (a) A section through two-dimensional noise in which two-thirds of all pixels are light and one-third dark. For clarity, this is shown as a regular array with two adjacent light pixels followed by one dark, rather than as a noise waveform. (b) The product of the waveform in (a) and a (raised, scaled) sinusoid, together with the locally averaged luminance of the resulting waveform (dotted line). Because there are two light pixels for each dark pixel in any local region of the image, the locally averaged luminance is closer to the local maximum than to the local minimum and varies in proportion to the local contrast (i.e. sinusoidally). When the modulating envelope moves, the locally averaged luminance waveform moves with it, giving first-order motion. In this case (where the noise has an overall bias) the distortion of the locally averaged luminance is global. (c) A noise waveform (again shown as a regular waveform for simplicity) in which there are equal numbers of light and dark pixels overall but there are local biases: in the first segment there are twice as many light pixels as dark, in the middle segment there is no bias and in the last segment there is the opposite bias. (d) The product of the waveform in (c) and a sinusoid together with the locally averaged luminance (dotted line). The latter has a form which varies according to the bias in the original noise. If the waveform in (c) were true noise of varying bias, rather than a regular waveform, then the dotted line in (d) would be a complex waveform that could be considered as a set of local segments of sinusoid of varying amplitude, phase and extent. The first-order motion that results when this waveform moves would not be revealed by global Fourier analysis because, globally, segments of opposite phase are equal and will cancel out. But the segments will be detectable by local patchwise analysis such as that employed by motion energy detectors. All such detectable patches will signal the same direction (that of the envelope) irrespective of their phase.

is applied equally to all points in the image, the distortion product will occur globally and will be reflected in the Fourier spectrum of the distorted image.

Distortion products may occur at any of several levels. Firstly, they may arise in the image itself. Images used in vision research are normally generated by computers and presented on CRT displays. CRT displays are typically characterized by a very substantial expansive luminance non-linearity. In principle the distorting effects of such non-linearities can be eliminated by applying an equal and opposite non-linearity to the input signal (a process referred to as gamma correction). In practice, it is easy to gamma correct an image approximately, normally by using a look-up table (LUT), but impossible to correct it perfectly. This is because the correction factor required depends on the properties of the image and varies to some extent both over space (across the screen) and over time. Thus, there are always distortion products in all CRT images and the question becomes: are they large enough to matter?

The next potential source of distortion products is in the process of luminance transduction in the receptors. Several investigators [e.g. MacLeod *et al.* (1992)] have argued that this process is not perfectly linear. We must therefore assume that distortion products arise in the receptors. These will be very similar in nature to those arising in the display. Finally, distortion products may arise at any stage of visual processing subsequent to luminance transduction in the retina (e.g. in ganglion cells, thalamus or cortex). For example Derrington (1987) has discussed distortion products arising in feline thalamus. Distortion products of this type differ from the other two categories in that they are not applied to a simple luminance signal but to some transformed, filtered or processed representation of the image. This may be an important consideration (see Discussion). Distortion products arising either in the receptors or subsequently, like those arising in the image, will only cause problems in second-order motion experiments if they are large enough to be detectable. This is an empirical question which has been addressed previously in other contexts and will be addressed further in this paper.

Local first-order motion patches

The second type of potential artifact is quite different. It arises locally, not globally, and exists in the undistorted image even with perfectly linear display and processing. It arises whenever the luminance of the carrier varies across space in a way that is not symmetrical about the mean luminance. Figure 2 illustrates the effect of such an asymmetry. Distortions will occur if the carrier is binary noise (i.e., all pixels have one of two values, light and dark) but the proportion of light pixels is not 50% [Fig. 2(b)]. In the case of greyscale noise, it will occur if values above the midpoint of the luminance range are more prevalent than those below, or vice versa. Thus, it is essential to ensure that the noise generator used to produce contrast-modulated noise stimuli is unbiased. However, even when the noise has 50% light and 50% dark pixels in the image as a whole, there will be local regions in which the balance is not perfect, because perfect balance will only occur for very large areas (just as 1000 tosses of an unbiased coin will yield close to 500 heads but sub-samples of 10 tosses will frequently yield 4 or 6). When a local imbalance occurs, there will be a local distortion in the mean luminance of the modulated waveform [Fig. 2(d)], leading to a local first-order motion signal. In fact, the image will be full of variously sized patches of first-order motion of various amplitudes, according to the sign and magnitude of the local bias in the carrier. The sum of such local distortions over the whole image will theoretically be zero, and so the image as a whole may still be drift-balanced. But since motion is detected locally in the visual system, local distortions, if large enough, will be detected individually and may be

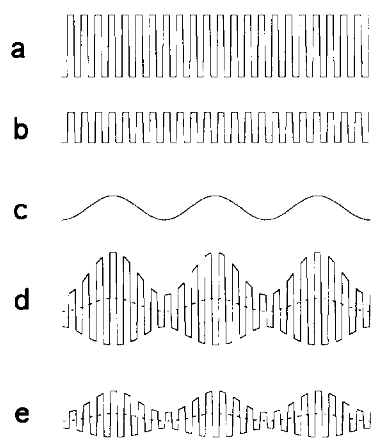

FIGURE 3. Luminance waveforms illustrating the effect of carrier contrast on the amplitude of first-order artifacts. The figure shows the case of global distortion products, but the effect is similar for local artifacts. (a) and (b) Noise samples of different amplitudes (i.e. different carrier contrasts), (a) having twice the amplitude of (b). (c) A sinusoid. (d) The product of (a) and (c) after applying an expansive non-linearity, as shown in Fig. 1. The locally averaged luminance waveform is also shown (dotted line). (e) The product of waveforms (b) and (c) after applying the same non-linearity. The resulting first-order component (dotted line) is similar to its counterpart in (d) but its amplitude is smaller.

used by the subject in psychophysical judgements. All such distortions will signal the same direction and speed as the contrast envelope and so their detection may be mistaken for detection of the envelope.

The experiments in this paper were designed to establish whether and when the global and local first-order artifacts described above occur in contrast-modulated noise images to an extent that contaminates measurements of the properties of putative second-order motion detection mechanisms. This question was examined in the context of detection thresholds for second-order motion. The underlying assumption is that where multiple detection mechanisms exist, then at threshold any pattern is detected by the mechanism that is most sensitive to it. Thus, assuming the existence of a true second-order motion detection mechanism, second-order motion images will be detected by the first-order system if they contain first-order artifacts large enough for that system to be more sensitive than the second-order motion system, and by the second-order system if they do not.

The underlying rationale for all the experiments presented here is as follows. If detection of second-order motion is, in practice, accomplished by detecting first-order artifacts using standard first-order motion detection mechanisms, then:

(i) Threshold sensitivity will show the same properties and dependence on stimulus parameters as threshold sensitivity for pure first-order motion. Specifically, where both the orientation and drift direction of a one-dimensional contrast modulation are to be identified, direction thresholds will be very similar to orientation thresholds except at very low drift speeds, as is the case for first order motion (Watson *et al.*, 1980).

(ii) Thresholds will decline in a predictable fashion as carrier contrast is increased, because the contrast of first-order artifacts (of both types described above) scale with the amplitude of the carrier giving rise to them (see Fig. 3).

If, on the other hand, second-order motion detection is accomplished by a specialized second-order motion detection system, then the relationship between orientation and direction thresholds may or may not be the same as for first-order motion and thresholds may or may not vary in inverse relation to carrier contrast.

To anticipate, we find that when the carrier is static broadband noise, thresholds for orientation and direction are in most cases very similar, as previously reported (Smith *et al.*, 1994). When the carrier used is either dynamic broadband noise or static high-pass-filtered noise, eliminating local first-order artifacts but not global distortion products, thresholds for direction are much higher than those for orientation, suggesting the involvement of a different (second-order) mechanism. Under these circumstances thresholds do not fall as carrier contrast increases. It is concluded that: (a) local first-order motion artifacts are often present and are utilized by subjects when the carrier is static broadband noise, leading to misleading results; (b) global distortion products do not appear to be large enough to provide a basis for perception of motion in these stimuli; (c) when the correct stimulus is used, clear evidence of a mechanism that is truly sensitive to second-order motion is found.

GENERAL METHODS

Subjects

Two subjects were used. AS is one of the authors. TF is an experienced observer who was naïve to the purpose of the experiments.

Stimuli

All stimuli were contrast-modulated two-dimensional noise patterns. They were generated using a Matrox IM-640 image processing system and were displayed on a monochrome monitor with P4 (white) phosphor at a refresh rate of 67 Hz. The mean luminance was 38 cd/m^2. In all cases the modulating waveform was a vertically or horizontally oriented sinusoid of spatial frequency 1 c/deg. Three types of carrier were employed in different experiments: (i) static, spatially broadband, two-dimensional noise; (ii) dynamic, broadband, two-dimensional noise; and (iii) spatially filtered, static, two-dimensional noise. Broadband noise was generated simply by assigning one of two states, light or dark (at random, with equal probability) to each pixel. The noise pixel size could be varied by constraining the noise generation process such that a square region of adjacent screen pixels (e.g. 2×2 or 4×4) was assigned a common luminance state to produce a single noise pixel. In the case of static noise, the same noise sample was used throughout the image sequence; in the case of dynamic noise a new sample was used on each frame. High-pass filtered images were produced by filtering such a binary noise image using conventional Fourier filtering techniques (yielding an 8-bit greyscale image). Filtered carriers were generated off-line and stored on disk; binary noise carriers were generated afresh at the beginning of each run.

The contrast-modulation of the carrier was achieved by drawing or loading a carrier into one frame buffer and drawing a modulation waveform in another. The two were then multiplied together in real time using a statistical LUT. The contrast modulation was always a grating shaped in space by a two-dimensional Gaussian to produce a two-dimensional Gabor patch. The position of the Gaussian envelope remained unchanged over time while the phase of the grating within it was updated to produce motion. Contrast modulations were also shaped in time, by varying the amplitude of the modulating waveform prior to multiplication by the carrier. Except where otherwise stated, each animation sequence lasted for 750 msec. The grating (envelope) amplitude started at zero, was ramped up with an integrated Gaussian waveform over the first 255 msec, remained constant at that level for 240 msec and was then ramped down again over 255 msec. Thus, the carrier appeared abruptly and had a constant mean (space-average) contrast over time but the amplitude of the contrast envelope was shaped in time. A double-buffering technique was used to allow one frame to be displayed while the next was being computed.

All images subtended 4 deg $\times$ 4 deg at the viewing distance of 95 cm and were presented on a uniform background of the same mean luminance. They were very carefully gamma-corrected (using a LUT) to minimize distortion products arising in the screen. Gamma correction was performed individually for each carrier used. A high-modulation-depth version of the contrast modulation to be used in the experiment was drifted slowly across the screen. A spot photometer was used which had a spot size (0.5 deg) large enough to integrate a number of light and dark noise pixels but smaller than the modulation period (1 deg). The gamma correction factor was adjusted so as to minimize the change in the reading that occurred as the contrast of the noise in the spot was modulated by the drifting contrast envelope. Readings were averaged over a number of noise samples. Precautions were taken to minimize changes in the gamma characteristics of the monitor over time (long warm-up period prior to commencement of experimental sessions, cushioned mounting of the monitor to reduce vibration) and the correction factor was checked regularly.

Procedure

Two detection thresholds were measured simultaneously using the method of constant stimuli. These were the modulation depth required to detect the orientation of the modulation and that required to detect its drift direction. The subject was asked to fixate the centre of the screen. Each experimental run commenced with a 30 sec period during which the screen was uniform. In each trial the subject was presented with a single animation sequence of the same mean luminance as the uniform field. The orientation of the drifting contrast modulation could be either horizontal or vertical and it could be drifting in either direction along the axis orthogonal to its orientation. These parameters were determined at random with equal probability. After each trial the subject was required to make two forced-choice responses by pressing buttons on a hand-held response box. The first response was the orientation of the modulation and the second was its direction. The computer recorded the responses in terms of whether they were correct or incorrect. At the end of the trial the image was replaced by a uniform screen of the same mean luminance. After an interval of at least 3 sec the next trial was presented. There were 50 trials in each run, ten at each of five modulation depths chosen on the basis of pilot trials so as to span the orientation and direction detection thresholds. Five identical runs of 100 trials were completed for each experimental condition so as to give a total of 50 responses at each modulation depth (250 responses in total). A sigmoid curve, constrained to asymptote at 50% and 100%, was fitted to the function relating percent correct responses to modulation depth using a least-squares method. The midpoint on this curve (75% correct performance) was taken as the threshold for that condition. Separate functions were fitted to the orientation and direction data to yield two separate threshold values, each based on 250 responses. Sample curve fits may be seen in Fig. 7.

EXPERIMENT 1: EFFECT OF NOISE PIXEL SIZE (STATIC BROADBAND CARRIER)

Varying noise pixel size has no effect on the magnitude of the global distortion product that will result from a luminance non-linearity of a given magnitude. The purpose of Experiment 1 was to address the issue of local biases in carrier luminance (i.e. deviations from the mean luminance of the pattern), as described in the Introduction. Consider a local patch in which a luminance bias exists. The size of the patch that is needed in order to cause an artifact has a minimum corresponding to the minimum fraction of the envelope period that is required for detection of motion. For example, if direction of motion is visible in a drifting first-order grating whose spatial extent is 0.1 spatial cycles or more (corresponding to 0.1 deg for the 1 c/deg envelope used here), then the patch of contrast-modulated noise in which there is an overall bias must be at least that large for detectable first-order artifacts to arise. For small pixels, this means that there must be a bias in a patch which contains many pixels. But for larger pixels sizes, this is not the case. When each pixel is (in this example) 0.1 deg in size or more then it is appropriate to regard a single pixel as a region in which luminance deviates from the global mean and a detectable first-order artifact will occur within each pixel. Thus, very large noise pixels will undoubtedly give

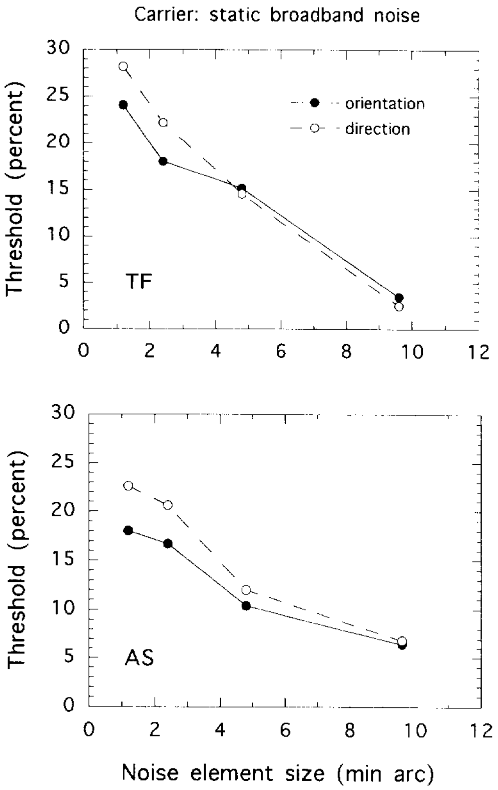

FIGURE 4. Thresholds for detecting the orientation (filled symbols) and direction (open symbols) of a drifting, oriented contrast modulation, as a function of the pixel size of a two-dimensional static noise carrier (mean contrast 21%). Thresholds are expressed in terms of contrast modulation depth (%). The envelope spatial frequency was 1 c/deg and the drift speed was 4 deg/sec. Results are shown separately for two observers, TF and AS.

rise to artifacts, irrespective of the stochastic properties of the noise. Very small pixels can only do so in regions where there is a bias affecting a large cluster of pixels. This depends on the stochastic properties of the noise and may happen only infrequently. The purpose of Experiment 1 was to establish the effect of noise pixel size on direction and orientation thresholds so as to reveal the range of conditions over which artifacts may mediate detection.

A static binary two-dimensional noise carrier was used. Its contrast with zero modulation was 21% (Michelson). With modulation, its contrast varied sinusoidally, with some modulation depth M, about 21% so its mean (space-average) contrast remained 21%. Modulation depth is defined as:

$$M = (C_{max} - C_{min})/(C_{max} + C_{min})$$

where C_{max} and C_{min} are the maximum and minimum contrasts between adjacent pixels of opposite polarity occurring at any point in the image. The independent variable in Experiment 1 was noise pixel size, which varied from 1.2 min arc (one screen pixel) to 9.6 min arc (8 screen pixels). Thresholds were measured separately for each noise pixel size.

Figure 4 shows direction and orientation thresholds for a 1 c/deg contrast envelope drifting at 4 deg/sec as a function of noise pixel size. The 4 deg/sec drift was created by updating the position (phase) of the modulating waveform by 3 screen pixels (3.6 min) every 15 msec. For both subjects, thresholds fall markedly as noise size increases. As stated above, the larger the noise pixel size the greater the probability of obtaining local luminance biases that are large enough to give rise to detectable first-order motion and so the lower the threshold. If this is the explanation of the observed fall in thresholds with increasing noise pixel size, then detection must be based on artifacts over much or all of the range of pixel sizes used because pixel size affects threshold over the whole range of sizes. This applies both to the detection of motion and to the detection of orientation, since both thresholds fall as pixel size increases.

However, it is impossible to know, in the case of the results in Fig. 4, whether the first-order artifacts that are apparent resulted from luminance biases in local patches of noise pixels or, at least for large noise pixels, arose within single noise pixels. Since the modulating waveform had a spatial resolution of one screen pixel (1.2 min) and was updated in increments of three screen pixels, local artifacts arising within noise pixels are expected where the noise pixel size >3 and not an integer multiple of three screen pixels (3.6 min). To eliminate within-pixel artifacts, further experiments were conducted in which the contrast envelope was incremented on each image update by an integer multiple of the noise pixel size. To maintain the drift speed at a constant value (4 deg/sec) the update rate was manipulated. For example, for a carrier whose noise pixel size was 3 screen pixels (3.6 min), the envelope was updated by 3 screen pixels every 15 msec; for a carrier whose noise pixel size was 6 screen pixels (7.2 min) the envelope was updated by 6 pixels every 30 msec, and so on. This eliminated the possibility of artifacts occurring within individual noise pixels. Figure 5 shows the results, which are similar to those in Fig. 4. Thus, it appears that any within-pixel local artifacts that occur are small in comparison with local artifacts arising on a slightly larger scale from stochastic biases in local patches of noise pixels.

A notable feature of the data in both Figs 4 and 5 is that for large pixel sizes, direction and orientation thresholds are similar. Thus, if the spatial structure of the contrast envelope is visible then so is its drift direction. This is the result obtained by others for first-order sine gratings, where direction is always visible at detection threshold except for very low drift rates (Watson *et al.*, 1980; Green, 1983). However, for small noise pixels, thresholds for direction appear to be rather higher than those for orientation. The results of Experiment 1 therefore suggest that for large pixel sizes performance was based on first-order artifacts but they also suggest that for small noise pixels performance may have been based on a different mechanism, perhaps a true second-order motion mechanism, which does not support direction perception at threshold.

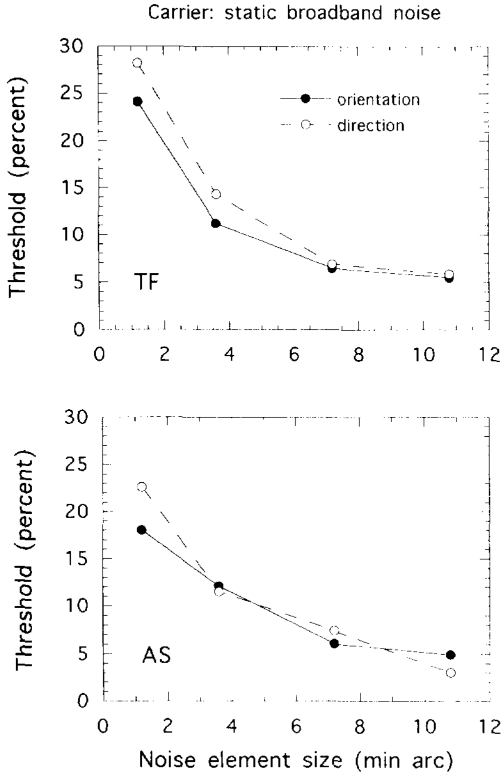

FIGURE 5. Thresholds for detecting the orientation (filled symbols) and direction (open symbols) of a drifting, oriented contrast modulation, as a function of the pixel size of a two-dimensional static noise carrier (mean contrast 21%). The envelope spatial frequency was 1 c/deg and the drift speed was 4 deg/sec. The envelope phase was always updated by an integer multiple of the noise pixel size.

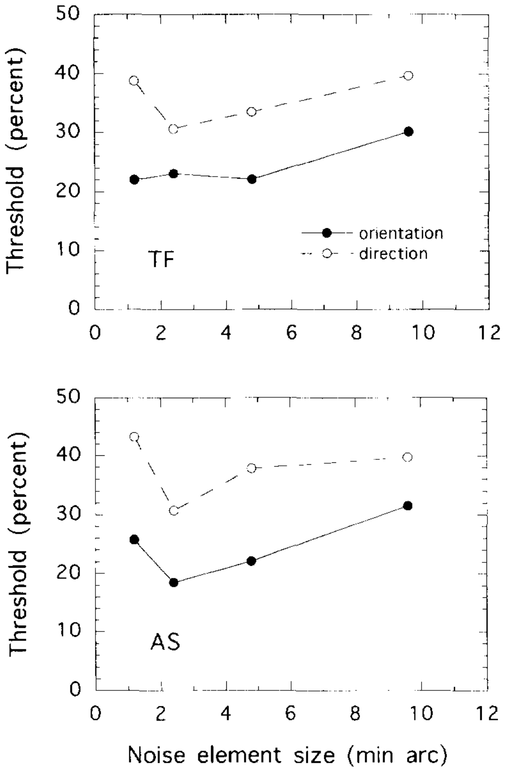

FIGURE 6. Thresholds for detecting the orientation and direction of a drifting contrast modulation, as a function of the pixel size of a two-dimensional dynamic noise carrier. All details are the same as for Fig. 4 except that a dynamic carrier was used.

EXPERIMENT 2: USE OF DYNAMIC, SPATIALLY BROADBAND CARRIERS

The results of Experiment 1 suggest that local first-order artifacts of the kind described in the Introduction are used by subjects in the detection of moving contrast modulations when the carrier is static noise, at least when the noise pixel size is large. An obvious implication is that it is best to use a small pixel size in experiments using stimuli of this type, so as to avoid artifacts. However, it is difficult to know whether this precaution is sufficient. Although local artifacts are reduced, they may still be large enough to be used by subjects. The objective of Experiment 2 was to find and test a way to eliminate local first-order artifacts entirely and then to re-examine orientation and direction thresholds for second-order motion.

A simple way to eliminate the problem is to use a dynamic noise carrier instead of a static noise carrier. Local first-order motion artifacts can only arise when a contrast envelope moves across a region of the carrier which has a luminance bias. At least two consecutive frames with the same bias in the carrier are therefore required. If the noise sample is replaced every time the envelope position is updated (in this case at 67 Hz) then motion of the envelope across a single, biased carrier sample does not occur. There is, of course, a finite probability of displaying two consecutive noise samples which, although different, have the same luminance bias at a given location. However, there is an equal probability of displaying two consecutive samples with opposite biases. In the latter case the resulting first-order motion artifact is analogous to a sine grating which moves a short distance and reverses its phase, resulting in "reversed phi" motion in the direction opposite that of the displacement [e.g. Anstis & Rogers (1975)]. Thus, local first-order motion artifacts will still occur with a dynamic noise carrier, but they will be equally divided between the two directions, whereas with a static carrier they are all in the same direction (that of the envelope motion). It is therefore highly unlikely, in the dynamic noise case, that these artifacts provide the basis for detecting the direction of motion of the envelope. It is conceivable that they might nonetheless provide the basis for detecting its orientation, since they signal the axis of motion of the envelope, from which orientation (always orthogonal to the axis of motion) could be inferred. However, later results (Experiment 5) showing similar performance using other methods of eliminating local artifacts suggest that this is not the case.

Effect of noise pixel size

In an initial experiment, the effect of noise element size was investigated using a method identical to that used in Experiment 1 except that the noise was dynamic i.e. the

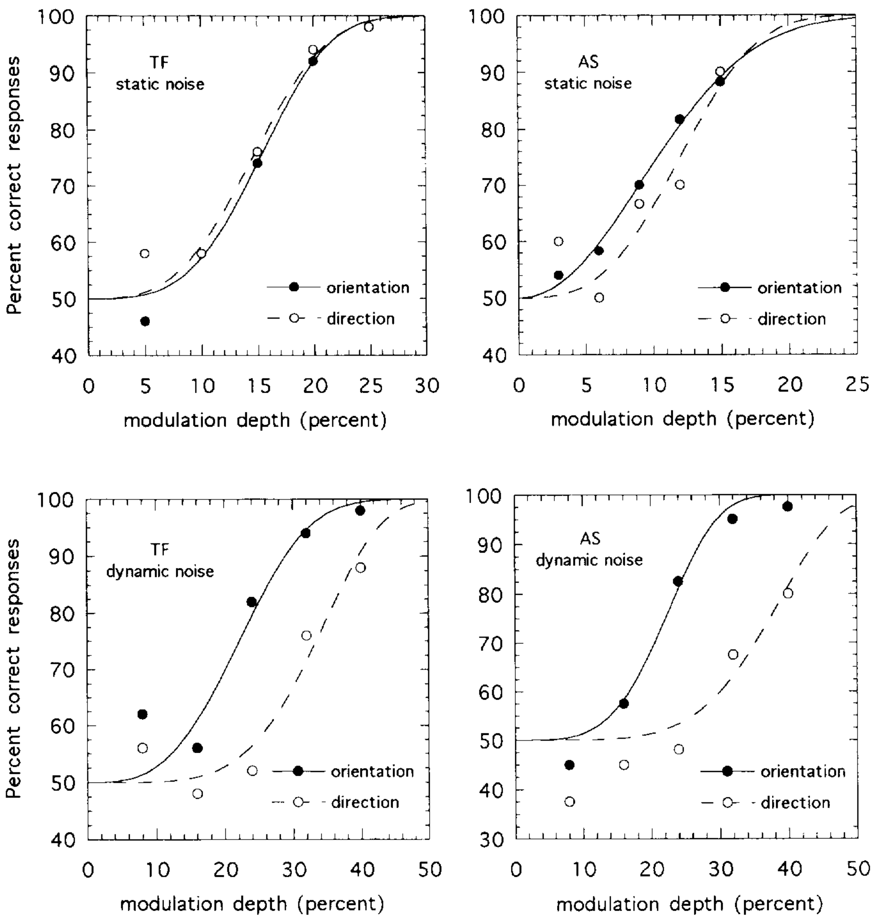

FIGURE 7. Sample psychometric functions illustrating the effect of the use of dynamic noise. The functions show performance as a function of contrast modulation depth and were fitted to the data using a least-squares method. The drift speed was 4 deg/sec in all cases and the carrier was either static noise (top) or dynamic noise (bottom). The noise element size was 5.2 min and the mean carrier contrast was 21%. Results are shown separately for the two observers, TF (left) and AS (right).

noise sample was changed on each image update. Each noise sample was binary as in Experiment 1 and its contrast was again 21%. The envelope spatial frequency was again 1 c/deg and its drift speed was 4 deg/sec. Figure 6 shows, for two subjects, the effect of noise element size on orientation and direction thresholds using a dynamic noise carrier. The pattern of results is quite different from that obtained with static noise (Figs 4 and 5). Firstly, thresholds are generally rather higher than for static noise. Secondly and more importantly, the sharp fall in thresholds as check size increases that was seen for static noise is completely absent for dynamic noise. Thirdly, thresholds for direction are substantially greater than those for orientation for all pixel sizes. The ratio of the two thresholds, averaged over both subjects and all pixel sizes, was 1.53. Thus thresholds for direction are around 50% higher than those for orientation, a much higher figure than is evident using a static noise carrier (Figs 4 and 5), even for the smallest pixel size.

Figure 7 shows the psychometric functions from which the thresholds in Fig. 6 were obtained, for one noise pixel size only (5.2 min). Also shown, for comparison, are the equivalent functions for static noise (Fig. 4). The difference between the functions for orientation and direction, in the case of dynamic but not static noise, is obvious.

Effect of drift speed

In order to see whether the finding of substantially higher thresholds for direction than for orientation holds over a range of speeds, we conducted further experiments using a dynamic noise carrier. This experiment had the additional objective of examining temporal acuity. It has been suggested that the temporal acuity of the second-order motion system may be significantly worse than that of the first-order system since direction detection for second-order motion stimuli is impossible at stimulus durations below about 200 msec (Derrington *et al.*, 1993).

The method was the same as that used for examining the effect of pixel size except that instead of using a fixed drift speed and a range of pixel sizes, pixel size was held constant at 1.2 min and a range of envelope drift speeds was used. Mean carrier contrast was again 21% and envelope spatial frequency was again 1 c/deg.

The results are shown in Fig. 8. Thresholds for direction are again substantially higher than those for orientation and this is the case for all drift speeds. The

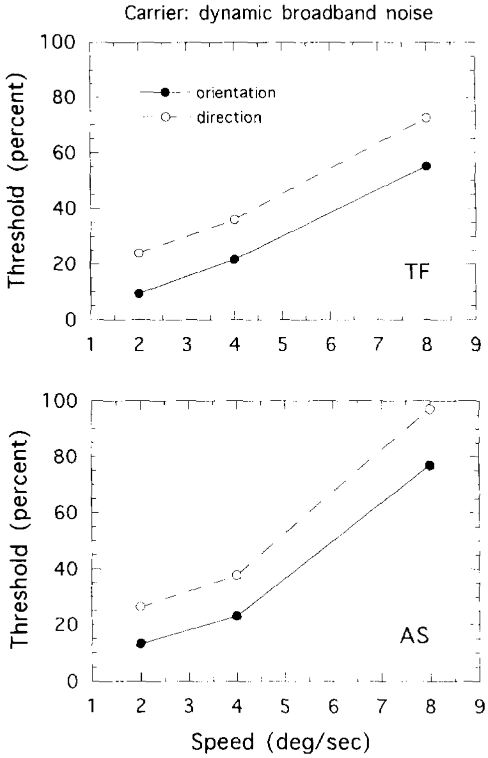

FIGURE 8. Thresholds for detecting the orientation and direction of a drifting, oriented contrast modulation, as a function of its drift speed. The carrier was dynamic noise (pixel size 1.2 min, mean contrast 21%) and the envelope spatial frequency was 1 c/deg.

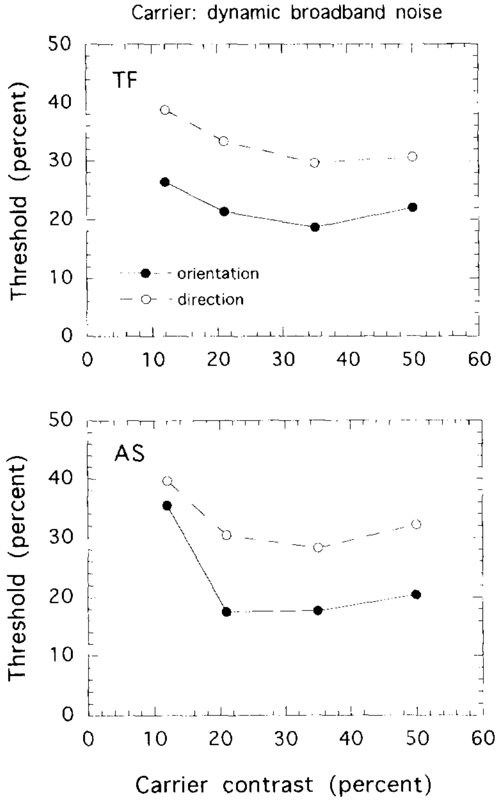

FIGURE 9. Thresholds for detecting the orientation and direction of a drifting, oriented contrast modulation, as a function of the mean contrast a two-dimensional dynamic noise carrier. The envelope spatial frequency was 1 c/deg and the drift speed was 4 deg/sec.

ratio of the two thresholds, averaged across the two subjects and across all drift speeds, was 1.7. The ratio varied considerably with speed, from 2.3 at the lowest speed tested to 1.3 at the highest; however, the difference between the two thresholds was approximately constant (15.8% contrast, averaged across the two subjects and the three drift speeds). Both direction and orientation thresholds increase with speed. The highest drift speed at which thresholds could be measured was 8 deg/sec (8 Hz). Thus temporal acuity for a 1 c/deg contrast modulation appears to be around 8 Hz. However, this figure may not represent temporal acuity for second-order motion in general. Threshold sensitivity is highly dependent on the nature of the carrier. For a 67 Hz dynamic noise carrier the lowest direction thresholds we obtained were around 25% (at 2 deg/sec, see Fig. 8) and it is possible that some other carrier might give greater sensitivity and a correspondingly greater temporal acuity.

Effect of carrier contrast

So far, only local first-order artifacts have been considered. We also wished to examine the role of global artifacts (distortion products). These (in common with local artifacts) are expected to increase in proportion to the carrier contrast (see Fig. 3). In order to examine the effect of the contrast of the dynamic noise carrier, a single noise pixel size (1.2 min) and envelope drift speed (4 deg/sec) were used and thresholds were measured, using the same method as before, for a range of carrier contrasts. The results are shown in Fig. 9. Once again, direction thresholds are much higher than orientation thresholds. The ratio of the two thresholds, averaged across the two subjects and across all carrier contrasts, was 1.51. Rather little effect of carrier contrast is evident. Thresholds are somewhat higher at the lowest carrier contrast than at the others, particularly for subject AS. It is, of course, inevitable that modulation thresholds increase as the carrier approaches its own contrast detection threshold. However, thresholds at 21% carrier contrast are similar to those at 50% (the highest contrast tested).

Discussion

The results obtained with dynamic noise carriers (Experiment 2) bear directly on the question of the involvement of both of the potential first-order artifacts that were described in the Introduction.

They suggest strongly that local first-order artifacts are indeed responsible for the detection of contrast modulations at threshold when the carrier is static noise, at least when large noise pixels are used (Experiment 1). When a dynamic carrier is used, eliminating such local artifacts, thresholds behave quite differently. Firstly, they are generally higher, suggesting the involvement of a less sensitive mechanism. Secondly, and more tellingly, thresholds are higher for direction than for orientation, by a factor of up to 2.3 (average 1.6). This is in contrast to results obtained with a static carrier [Experiment 1; Smith *et al.* (1994)] and to results obtained by others for first-order grating stimuli [e.g. Watson *et al.* (1980)]. The clearcut difference between the two thresholds that was found with dynamic carriers suggests the action of a mechanism with different properties from the first-order motion system. We suggest that this mechanism is a true second-order motion system that responds to moving contrast modulations (and probably other types of second-order motion). Unlike its first-order counterpart, this mechanism is not capable of specifying direction at the threshold for detecting orientation.

The finding of different thresholds for direction and orientation also argues against the involvement of global first-order artifacts (distortion products). Distortion products arising from a luminance non-linearity are expected to be qualitatively the same whether the carrier is static or dynamic (they may be of lower amplitude in the dynamic case because of reduced visual sensitivity to high temporal frequencies). Since they are first-order artifacts, thresholds should behave like those for first-order stimuli if they are used for detection i.e. direction and orientation thresholds should be the same. This is not the case. Further evidence against the involvement of distortion products comes from the results of varying carrier contrast. The amplitude of a distortion product will increase in proportion to carrier contrast (see Fig. 3). If detection of motion in Experiment 2 were based on detection of such distortion products then thresholds should be inversely related to carrier contrast, halving when carrier contrast is doubled. In fact, at least above 20%, a doubling of carrier contrast has very little effect on motion thresholds. This is not to say that distortion products do not exist in the neural representation of the image, only that they are not large enough to be useful for detection of motion in images such as ours.

EXPERIMENT 3: EFFECT OF DURATION (DYNAMIC BROADBAND CARRIER)

In Experiment 2 it was concluded that when an appropriate stimulus is used for studying second-order motion, a true second-order motion detection mechanism is revealed which differs from its first-order counterpart in that thresholds for identifying direction are much higher than those for identifying orientation. In Experiment 3 an alternative interpretation of this difference in thresholds is explored and rejected. Derrington *et al.* (1993) have shown that whereas identification of the direction of a moving luminance modulation (grating) is possible at extremely short stimulus durations, identification of the direction of motion of a contrast modulation (beat) is possible only if the duration of the stimulus exceeds about 200 msec. It seems likely that this result for beats would also apply to contrast-modulated noise, although we have not established this directly. In our experiments, contrast was temporally shaped. This means that, at threshold, it was invisible for part of its 750 msec duration. If it were the case that the pattern was visible for <200 msec then orientation might be detected but direction might not. Only by increasing contrast, so that a larger part of the temporal profile was above threshold, would direction be detectable. This would lead to detection thresholds being higher than orientation thresholds, as observed. If this were the explanation of our results, then it might still be appropriate to infer the existence of a second-order motion mechanism with a poor temporal response, but it would be inappropriate to conclude that this mechanism cannot identify direction at the threshold for orientation. With this consideration in mind, we used in Experiments 1 and 2 a temporal shaping waveform with a plateau at its peak, rather than the more conventional temporal Gaussian. The duration of the plateau was 240 msec (see General Methods section), this value being chosen to exceed the value of 200 msec quoted by Derrington *et al.* (1993). However, in view of the results obtained in Experiment 2, it seemed worthwhile to establish directly that the observed difference between the two thresholds does not arise from the use of a limited duration.

Direction and orientation threshold measurements were obtained using a dynamic noise carrier as in Experiment 2, a carrier contrast of 21% and a 1 c/deg envelope drifting at 4 deg/sec. A range of temporal shaping waveforms was used. The rising and falling

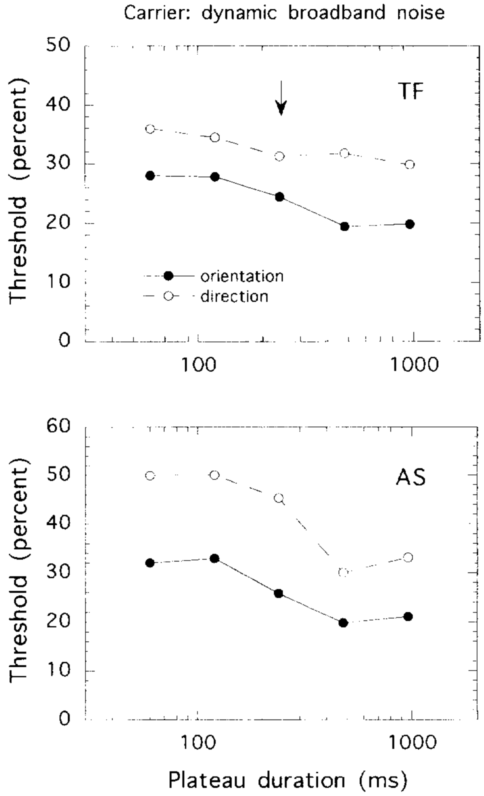

FIGURE 10. Thresholds for detecting the orientation and direction of a drifting, oriented contrast modulation of a two-dimensional dynamic noise carrier, as a function of duration of the stimulus. The contrast modulation was ramped on, held constant for the duration shown on the abscissa and then ramped off. The duration used in all other experiments in this paper is indicated by an arrow. The envelope spatial frequency was 1 c/deg and the drift speed was 4 deg/sec.

portions were identical to those used in the previous experiments (a cumulative Gaussian and its inverse, respectively, each lasting 255 msec). But the duration of the plateau (the time at which contrast modulation depth was at its peak value) was varied from 60 to 960 msec (spanning the 240 msec used in Experiments 1 and 2).

The results are shown in Fig. 10. A modest increase in performance (fall in threshold) occurs with increasing duration. However, there is little change in the ratio of the two thresholds with duration. Even for the longest duration (960 msec) there is a clear difference between the two thresholds (ratio 1.54, averaged across the two subjects). Thus it is not the case that the difference between the two thresholds arises from restricted duration.

The fact that direction is detectable at the shortest plateau durations does not necessarily conflict with the results of Derrington *et al.* (1993), who used abrupt onset and offset and whose figures relate to total duration. Our shortest total duration (i.e. when the plateau was 60 msec) was 570 msec, and it is not impossible that the stimulus was visible for 200 msec or more at threshold.

EXPERIMENT 4: EFFECT OF GAMMA CORRECTION FACTOR (DYNAMIC BROADBAND CARRIER)

The conclusion of Experiment 2 was that local first-order artifacts can contaminate measurements of second-order motion detection but that global first-order distortion products do not, at least for the images we have described. In Experiment 4 a different strategy is used to address the role of distortion products. In this experiment global luminance non-linearities of various magnitudes were deliberately introduced into the image and their effects on direction and orientation thresholds were measured. We have asserted that, for dynamic but not static carriers, motion detection involves a second-order motion mechanism with characteristic properties. If a distortion product is introduced which is correlated with the contrast modulation, then it follows that if the distortion product is made progressively larger it will at some stage be used for detection. When this happens, direction and orientation thresholds should be similar because the first-order motion mechanism can detect direction at the threshold for orientation.

The magnitude of the distortion product arising in the display, which was as close as possible to zero in the preceding experiments, was systematically varied by varying the gamma correction factor that was applied to the image. The screen was first calibrated as described under General Methods. A LUT was used to apply an opposite (compressive) non-linearity to compensate for the expansive non-linearity inherent in the CRT display. Gamma (γ) was defined as follows:

$$L = kV^{\gamma}$$

where V is the voltage applied to the video input of the CRT display (linearly related to the value of intensity in the original, digital representation of the image), L is the luminance generated by the display in response to that voltage, k is a scaling constant and γ is a constant which characterizes the degree of non-linearity of the display. Thus $\gamma = 1$ represents a linear response and any value >1 represents an expansive non-linearity. The gamma value of the display used was found to be close to 1.9 when calibrated using a binary noise carrier of pixel size 1.2 min (other carriers gave somewhat different values). The gamma value varied slightly from time to time (within ± 0.1 over the period of several months during which the present study was conducted).

Thresholds were measured repeatedly for a single image but using various gamma correction factors. Apart from this, the image was identical to the standard image used in Experiment 2. The carrier was dynamic noise (pixel size 1.2 min, mean contrast 21%). The spatial frequency of the envelope was 1 c/deg and its drift speed was 4 deg/sec. The temporal profile was as in Experiment 2. The correction factors used are specified in terms of the gamma factor that they would correct. Thus, a value of 1.9 means that the digital values in the image were raised to the power of the inverse of 1.9 then rescaled before display, which would exactly linearize a display with a gamma factor of 1.9. The range spanning the value of 1.9 was sampled finely and an additional, higher range of more coarsely spaced values was also used. It should be noted that when the gamma correction factor is changed, the mean luminance of the display and the carrier contrast both change. It was therefore necessary to adjust the amplitude of the carrier in the original image in order to achieve a constant carrier contrast of 21% despite changes in gamma.

The results are shown in Fig. 11. The gamma correction factor could be set anywhere in the range 1.7–2.2 without any effect on thresholds. Over this range, direction thresholds are again much higher than orientation thresholds (mean ratio 1.47). Clearly, compensating for a gamma factor of 2.2 must introduce a substantial distortion product into a contrast-modulated image if the

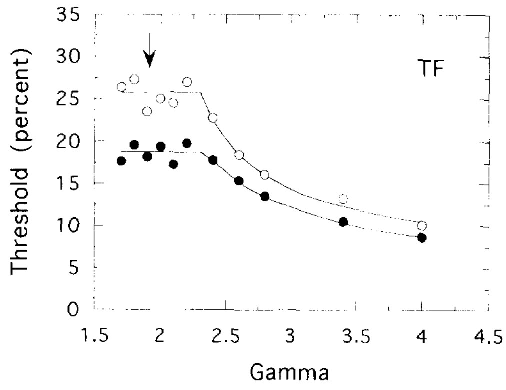

FIGURE 11. Thresholds for detecting the orientation and direction of a drifting, oriented contrast modulation of a two-dimensional dynamic noise carrier. The gamma value for which correction was applied is shown on the abscissa. The actual gamma factor for the display used is indicated by an arrow. The envelope spatial frequency was 1 c/deg and the drift speed was 4 deg/sec. Curve fitting was done by eye.

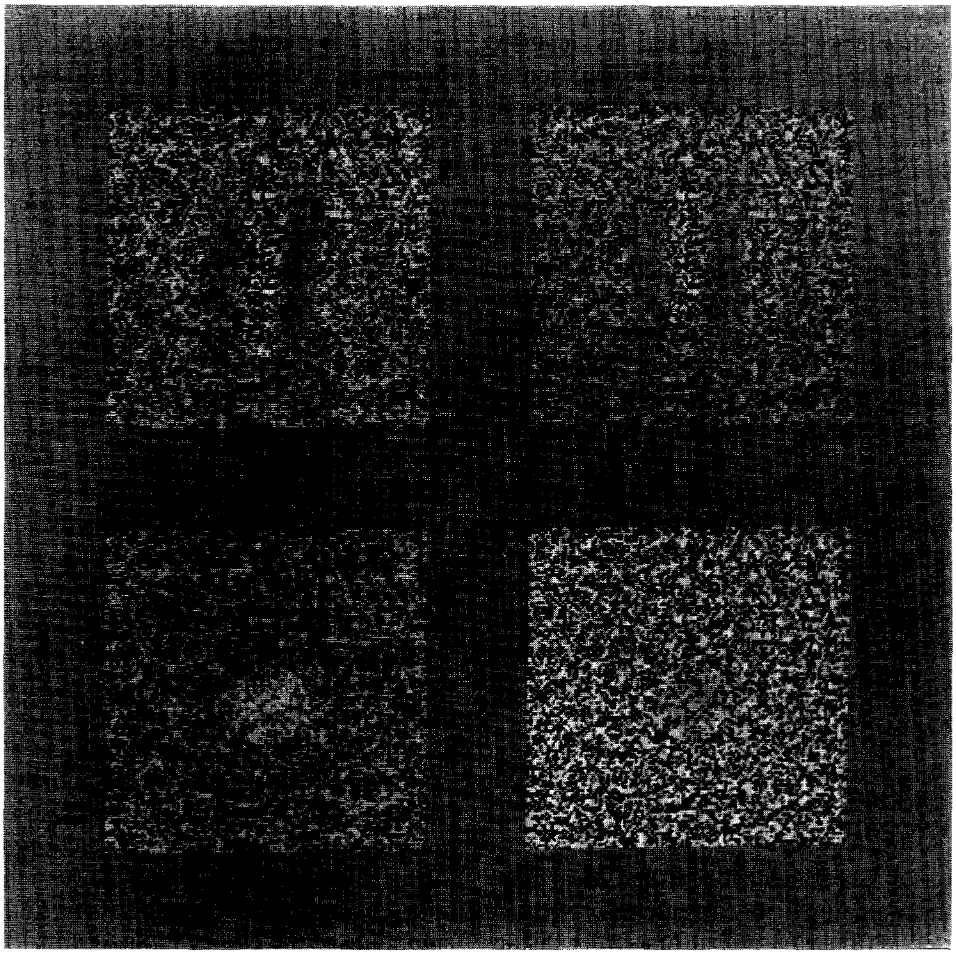

FIGURE 12. Sample images showing the effects of high-pass filtering. Top, an unfiltered two-dimensional noise image (top left) and the same image high-pass filtered using a filter with a cut-off at 2 c/deg (top right). Bottom, an unfiltered noise image (lower left) similar to that shown top left but with a deliberately introduced local patch containing a luminance bias (75% light and 25% dark pixels as opposed to 50% each in the rest of the image) and (lower right) the effect of high-pass filtering the biased image using a filter cut-off of 2 c/deg. The appearance of the noise elements is not greatly affected by the filtering operation, but the luminance bias in the central patch is much reduced. The spatial frequencies quoted assume an image size of 4 deg.

gamma factor of the display is in fact 1.9. Yet thresholds are no lower than for a value of 1.9, suggesting that detection is not based on the distortion product introduced by inappropriate correction. Assuming that detection is always performed by the most sensitive mechanism, this means that, over this range of gamma values, the second-order mechanism is more sensitive to the contrast modulation itself than the first-order mechanism is to the associated distortion product.

As gamma is increased beyond 2.2 thresholds begin to fall. In the falling portion of the curve, presumably the sensitivity of the first-order motion system to the distortion product is greater than that of the second-order system to the contrast modulation and so thresholds fall in inverse proportion to the size of the distortion product. Importantly, the difference between the direction and orientation thresholds declines as detection is based increasingly on first-order motion detection.

Aside from the above theoretical implications, a practical implication of this result is that accurate gamma correction is less critical for experiments of this kind than might be supposed. The drift over time and variations over space that inevitably occur in any display is not sufficient, at least for the display we used (a standard Apple 12" monochrome monitor), to cause usable first-order artifacts.

EXPERIMENT 5: USE OF HIGH-PASS-FILTERED STATIC NOISE CARRIERS

Experiment 2 explored the use of dynamic noise to eliminate local first-order artifacts in contrast-modulated noise. In Experiment 5 a completely different strategy was used for the same purpose. The use of dynamic noise does not remove local artifacts; it simply ensures that they are equal in opposite directions along the axis of motion. An alternative strategy is to retain the use of a static carrier and remove local artifacts by spatially filtering the carrier, prior to modulating its contrast. Removing low spatial frequencies from two-dimensional noise (i.e. high-pass filtering it) tends to make the mean luminance of any patch of the image regress towards the mean luminance of the display as a whole. This is illustrated in Fig. 12. Thus filtering tends to reduce local luminance biases of the type described in the Introduction. If the filter cut-off is sufficiently high it will remove them completely, although there will be a cost in terms of contrast sensitivity because the visual system is relatively insensitive to the high frequencies that will remain. High-pass filtering does not remove global distortion products, although their amplitude may be reduced due to reduced sensitivity to high spatial frequencies.

It should be noted that when the carrier used is high-pass-filtered two-dimensional noise, the image that results when a moving contrast envelope is introduced is not drift-balanced (see the Appendix). However, the departure from drift balance is modest and in fact there is more first-order motion energy in the direction opposite that in which the contrast envelope moves than there is in the same direction. This means that if the imbalance were detected by a motion energy detection system such as that proposed by Adelson and Bergen (1985), which compares energy in opposite directions, then motion would be perceived in the direction opposite that of the moving contrast envelope. This does occur in some circumstances (see Experiment 6) but normally it does not. Whenever motion is seen in the correct direction, detection cannot be based on the directional imbalance in the Fourier spectrum of the image and so the fact that the image is not quite drift-balanced does not provide a means of detecting the direction of a moving contrast modulation.

Several experiments similar in design to those in Experiments 1 and 2 were carried out using static, high-pass filtered carriers.

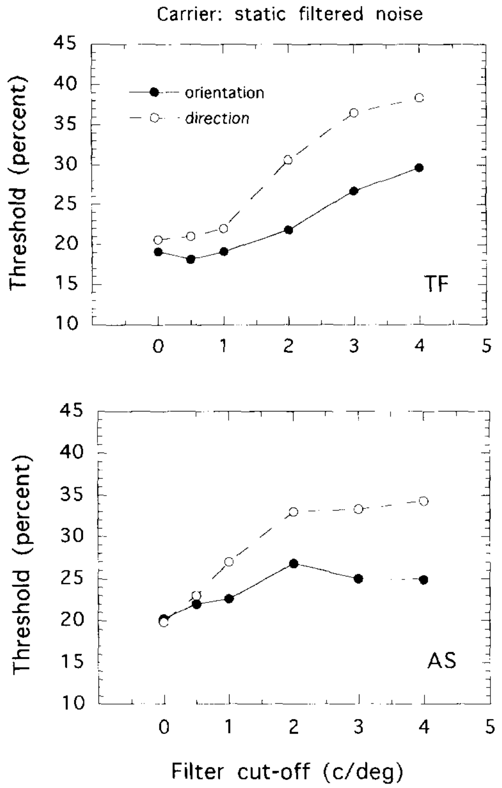

FIGURE 13. Thresholds for detecting the orientation and direction of a drifting, oriented contrast modulation. The carrier was static, high-pass filtered noise and thresholds are plotted as a function of the filter cut-off. The envelope spatial frequency was 1 c/deg and the drift speed was 4 deg/sec.

Effect of filter cut-off frequency

Initially, thresholds for identifying the orientation and direction of a drifting contrast modulation were measured using a single envelope spatial frequency (1 c/deg) and drift speed (4 deg/sec) for various high-pass filtered carriers with different filter cut-off frequencies. First, a carrier consisting of binary, two-dimensional noise was generated. This was then high-pass filtered using conventional Fourier techniques. The same image was filtered using each of several cut-off frequencies ranging from 0.5 to 4 c/deg (see Fig. 12). The filtered images were then scaled such that they all had an r.m.s. contrast of 20%. (High-pass filtering reduces r.m.s. contrast for a given Michelson contrast, so this scaling procedure means that Michelson contrast increased somewhat as the filter cut-off was increased, from 20% for the unfiltered image to 40% in the case of the image filtered with a 4 c/deg cut-off.)

Figure 13 shows the thresholds obtained. Thresholds increase as the cut-off spatial frequency increases. For an unfiltered image (shown as 0 c/deg cut-off) the orientation and direction thresholds are similar, as previously observed (Experiment 1). This is also true for the 0.5 c/deg cut-off condition. But at 1 c/deg (when the filter cut-off coincides with the envelope frequency) the two thresholds begin to diverge, direction thresholds being higher than orientation thresholds. At 2 c/deg and above, there is a substantial difference between the two thresholds. The average ratio of the two thresholds over the range 2–4 c/deg, averaged across the two subjects, is 1.33.

Thus the results obtained with high-pass filtered carriers mirror those obtained with dynamic broadband carriers. In both cases, direction thresholds for a moving contrast modulation are significantly higher than orientation thresholds for the same stimuli. Our interpretation is the same in both cases. Both manipulations remove local luminance artifacts that contaminate threshold measurements obtained with static, broadband carriers, revealing the action of a second-order motion mechanism which cannot identify direction at the threshold for detecting second-order spatial structure.

Effect of carrier contrast

The same predictions for the effects of carrier contrast hold for high-pass filtered carriers as for unfiltered and dynamic carriers. If detection is based on first-order artifacts, whether local or global, thresholds should fall with increasing carrier contrast. The results shown in Fig. 13 suggest that high-pass filtering is effective in removing local first-order motion artifacts. Since thresholds for direction are higher than for orientation, they also suggest that no other first-order artifact is used for

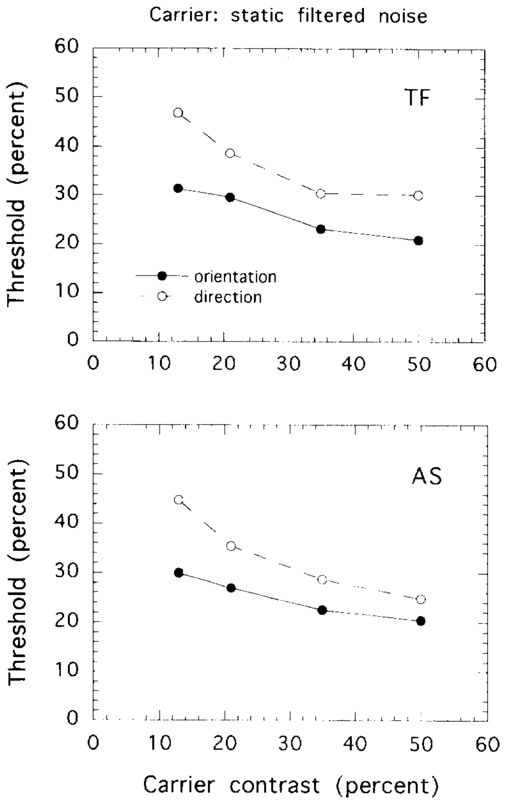

FIGURE 14. Thresholds for detecting the orientation and direction of a drifting, oriented contrast modulation, as a function of the mean contrast of a high-pass filtered, two-dimensional static noise carrier. The filter cut-off used was 2 c/deg. The envelope spatial frequency was 1 c/deg and the drift speed was 4 deg/sec.

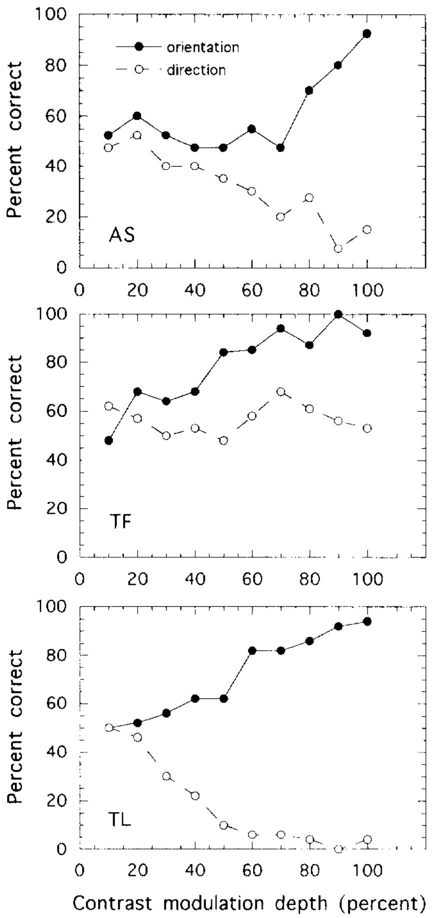

FIGURE 15. Performance (per cent correct responses) for three subjects on orientation identification (filled symbols) and direction identification (open symbols) as a function of contrast modulation depth. The carrier was static two-dimensional noise high-pass filtered with a cut-off of 2 c/deg. The envelope spatial frequency was 1 c/deg and the drift speed was 12 deg/sec. Performance of 50% correct represents chance responding, higher values indicate correct identification and lower values indicate incorrect identification.

detection and hence that global distortion products are not involved. However, to establish with greater certainty that high-pass filtered noise does not result in usable distortion products, thresholds were measured for a variety of carrier contrasts using a high-pass filtered carrier. A single filter cut-off of 2 c/deg (1 octave above the envelope frequency) was used. The speed was again 4 deg/sec, the envelope spatial frequency was 1 c/deg and the temporal profile was the same as in previous experiments.

The results are shown in Fig. 14 and are quite similar to those obtained with a dynamic broadband carrier (Fig. 9). Thresholds for direction are consistently higher than those for orientation (mean ratio 1.36 averaged across the two observers). There is a modest effect of carrier contrast, but much less than predicted by the distortion product hypothesis. This suggests that first-order artifacts are not involved and that the thresholds recorded reflect the activity of a true second-order mechanism.

EXPERIMENT 6: REVERSED MOTION AT 12 DEG/SEC (HIGH-PASS FILTERED CARRIER)

The results of Experiment 5 suggest that the use of high-pass-filtered carriers is an effective means of eliminating local bias and is therefore preferable to the use of unfiltered static carriers in experiments on second-order motion. However, we have so far tested it only using a single drift speed. It is not clear that it provides a universally valid stimulus, particularly in view of the fact that it is not-drift balanced and contains the potential for perception of motion in the direction opposite the direction of the moving contrast modulation (see the Appendix). In all measurements in Experiment 5, direction, when perceived, was perceived correctly. However, we have used a variety of drift speeds and find that at high speeds, direction is perceived in the opposite of the envelope drift direction. The purpose of Experiment 6 was simply to provide a formal demonstration of this phenomenon.

In informal experiments in which drift speed was varied, we identified the highest drift rate that is detectable using a 1 c/deg modulation of a high-pass filtered carrier. We find that motion is visible up to about 12 deg/sec (12 Hz), somewhat higher than was the case for dynamic unfiltered noise (Fig. 8).

Using a carrier contrast of 20% r.m.s. (30% Michelson), a filter cut-off of 2 c/deg and a drift speed of

12 deg/sec, we obtained psychometric functions relating orientation and direction performance to envelope contrast. The two subjects gave different results and so a third subject (TL, the other author) was tested. The results for all three subjects are shown in Fig. 15. For low envelope contrasts, performance is at chance levels (50% correct). In the case of orientation judgements, as the envelope contrast is increased, performance increases, reaching 100% correct at very high modulation depths, for all three subjects. However, in the case of direction judgements, performance for AS and TL falls from chance levels to a level close to zero. That is, at low modulation depths no consistent direction is seen and at high modulation depths motion is consistently seen in the wrong direction. The third subject (TF) performed at chance levels for direction at all modulation depths. Reversed motion does not occur for lower drift speeds. Certainly it never occurred during extensive measurements at 4 deg/sec (Experiment 5) and informal investigation over a range of speeds failed to reveal it below about 8–10 deg/sec.

We interpret the reversed perceived motion demonstrated (for two of three subjects) in Fig. 15 as follows. When the contrast of a high-pass-filtered noise sample is modulated by a drifting sinusoid, the image contains a narrow range of spatial frequencies close to the filter cut-off frequency at which all the motion energy present is in the direction opposite the envelope drift direction (see the Appendix). At all other spatial frequencies, motion energy is equal in both directions (as it is at *all* frequencies in a broadband carrier, whether static or dynamic). We believe that the reversal of perceived direction shown in Experiment 6 reflects detection of this marginal dominance of the reverse direction in the luminance-domain Fourier spectrum. It is well known that a consistent direction is perceived when only a small directional bias exists in a pattern containing motion energy in opposite directions and so it is perhaps not surprising that the marginal direction bias in our image can be detected. However, a second-order motion system sensitive to moving contrast modulations will always signal motion in the opposite direction to that signaled by the first-order system in such circumstances. What direction is perceived then becomes a matter of which of the two systems dominates. Normally it is the second-order system, not because it is inherently dominant but, presumably, because the first-order system is only very weakly stimulated by such a small directional imbalance. At high envelope drift speeds, however, the second-order system is at the limit of its temporal resolution and its sensitivity is much reduced. Here there will be some range of speeds where the first-order system is more sensitive because of its superior temporal acuity, and the direction opposite to the modulation direction will be perceived. We find that it is sometimes possible to see transparent motion at the transition point, although it is never very clear, perhaps because both systems are operating near threshold. Inspection of Fig. 15 shows that the three subjects show variation in the ability to detect the small directional imbalance. Subject TL was most sensitive, seeing it reliably at about 50% modulation depth. AS was less sensitive, requiring 90% modulation depth, and TF failed to detect it at all. If the threshold for perception of reversed motion for AS and TL is taken as the 25% correct level, it can be seen from Fig. 15 that this threshold is no higher than the orientation threshold (in fact it appears to be rather lower). This is consistent with the notion that the first-order system is mediating detection in these circumstances.

Since direction perception is based on first-order motion energy at 12 Hz, it follows that the temporal acuity limit for second-order motion is lower than that. The highest drift rate at which direction is perceived correctly was found in informal observations to be about 8–10 Hz, which is consistent with the earlier experiments using dynamic noise (Fig. 8).

DISCUSSION

There are two main conclusions to be drawn from the work reported in this paper. One is theoretical, the other practical.

Theoretical implications

The first main conclusion is that the human visual system includes a mechanism that is specialized for the detection of second-order motion and that this mechanism cannot specify direction of motion at its absolute detection threshold. A number of empirically supported claims have previously been made for the existence of a second-order mechanism [e.g. Harris & Smith (1992); Mather & West (1993); Werkhoven *et al.* (1993); Ledgeway & Smith (1994)], yet some scepticism remains in some circles. One reason is that the properties of the putative second-order system have much in common with those of the first-order system, leading to the suspicion that both types of motion are detected by the same mechanism. In fact, similarity between the two is predicted by models (Chubb & Sperling, 1988; Wilson *et al.*, 1992) in which the two types of motion are detected using the same principle (motion energy detection), in one case in a linear representation of the image and in the other in a non-linearly transformed version of the image. Nonetheless, for the distinction to be persuasive it is necessary to identify some properties, at least, of the second-order motion mechanism that differ from first-order motion. One such property is inferior temporal acuity. The second-order motion system requires long stimulus durations (Derrington *et al.*, 1993) and cannot detect high drift temporal frequencies [Derrington (1994); our Fig. 8]. To this we add another property. Whereas the first-order motion system is able to specify direction of motion at the threshold for identifying orientation, the second-order motion system is not. In order to identify direction of second-order motion, modulation depth must be some 50% higher than that required to identify orientation.

It is unclear why direction cannot be identified at threshold for second-order motion. It is interesting to note in this context that it has been reported that direction thresholds are also higher than orientation thresholds in the case of motion defined purely by colour modulations (Lindsey & Teller, 1990), although Mullen and Boulton (1992) and Derrington and Henning (1993) have both suggested that the difference is not very marked. It is possible that second-order form (including orientation) perception is based on a separate mechanism from the second-order motion mechanism and that the latter is less sensitive than the former. However, this interpretation seems unlikely in view of the fact that the two thresholds co-vary quite closely, suggesting a common basis for perception of form and motion. In the case of first-order motion, detection of form and motion appear to be based on a common initial filtering stage that feeds both processes. If both processes are highly efficient then the two thresholds are expected to be very close. Perhaps the same applies to second-order stimuli except that motion is extracted less efficiently than form, so that motion and orientation thresholds differ, but by a fairly constant factor.

Practical implications

The second main conclusion is practical. It is that contrast modulation of static two-dimensional noise may not be a suitable stimulus for use in experiments on second-order motion. This is because although the stimulus contains second-order motion which will be detectable by the second-order motion system at sufficiently high modulation depths, it also contains first-order motion arising from local luminance biases. This first-order motion varies in strength across space and may be below threshold at some locations, but our results suggest strongly that, at least at some locations, sensitivity to first-order artifacts is greater than sensitivity to contrast modulation and that such artifacts are used consistently by subjects in detection tasks involving this type of image. It is arguable that static noise composed of small pixels (1–2 min for an envelope of 1 c/deg) provides an artifact-free carrier, but it is much safer to avoid the use of static broadband carriers altogether. We have only considered threshold sensitivity and so we cannot assert that local luminance artifacts have the same contaminating effects in studies of suprathreshold second-order motion perception. However, it seems very possible that they do and so broadband static noise carriers are perhaps best avoided in all experiments on second-order motion. As well as having implications for future work, this conclusion casts some doubt on a number of findings in the existing literature.

Although we find strong evidence that local first-order artifacts occur, we find no evidence of global first-order artifacts (distortion products). The possible contaminating effects of this class of first-order artifacts has been considered before in other contexts [e.g. Henning *et al.* (1975); Derrington & Badcock (1986); Derrington (1987)]. Recently, Derrington (1994) has reached a somewhat different conclusion from ours in the context of AM gratings. He found, using a 1 c/deg modulation of a 5 c/deg sine grating, that global luminance artifacts do contaminate results for high, though not low, carrier contrasts. He measured temporal acuity for AM gratings and found acuity resembling first-order motion at high carrier contrasts (suggesting the involvement of distortion products) but a quite different (inferior) temporal response for low contrast carriers (suggesting the involvement of a second-order motion mechanism). We find no sign of this type of artifact using two-dimensional noise carriers and so it is worthwhile considering why not. Unfortunately it is not possible for us to apply our paradigm to AM gratings because the orientation judgement is only appropriate in the case of a carrier whose luminance varies in two dimensions. We can therefore only speculate. One possible reason for a difference between one-dimensional grating carriers and two-dimensional noise carriers concerns the spatial frequency spectrum of the carrier. It is known that there is a receptoral non-linearity, but suppose that the greatest non-linearity is "late" (post-receptoral). If the non-linearity is preceded by spatial filtering, then what matters is not the contrast of the carrier but the energy present in the carrier at any given spatial scale. In the case of a noise carrier, there is rather little energy at any one spatial frequency, even when the contrast of the image as a whole is high. In the case of a grating carrier, all the energy is at a single spatial frequency and so a relatively large distortion product will arise at that spatial scale. If the non-linearity is "early" (receptoral) then, of course, all that matters is the contrast and all carriers would be affected equally. Our results could therefore be construed as evidence that the most substantial non-linearity is post-receptoral.

So what is the most appropriate carrier for use in the study of second-order motion? Sine grating carriers avoid the problem of local luminance artifacts, provided a carrier of sufficiently high spatial frequency is used. They also have merit in situations where a simple (in Fourier terms) carrier is required i.e. in investigations of the spatial and temporal frequency sensitivity of the second-order motion mechanism. In other circumstances, however, noise carriers may be preferable because of the reduced risk of global distortion products. Since two-dimensional noise is as easy to generate as one-dimensional noise and gives greater scope for experimentation (allowing directional judgements along more than one axis, for example), two-dimensional noise would seem preferable. But if noise is used, it is essential to take precautions against local luminance artifacts. There is rather little to choose between the two methods we have explored in this paper. Both deal effectively with the local artifact problem. Sensitivity to contrast modulations is quite similar in the two cases, at least in the circumstances we have employed (compare Figs 9 and 14). But dynamic noise would seem to have the edge over high-pass filtered noise because the latter is not perfectly drift balanced and this adds an additional source of potential first-order artifacts. These are easily detected (by reversed perception of direction) at high speeds (Fig.

15), but are present at all drift speeds and may have more subtle contaminating effects.

In the foregoing discussion, only contrast modulation has been considered. Other types of second-order motion have also been described (Chubb & Sperling, 1988). In view of the potential artifacts that can arise and the degree of care needed to eliminate them, does it make sense to abandon contrast modulation in favour of other types? In our view, several other classes of second-order motion can be reduced to contrast modulation (Ledgeway & Smith, 1994). One such is texture modulation (e.g. varying the size of texture elements across space and then moving the size-modulation envelope). Because of the shape of the contrast sensitivity function, most changes in texture involve concomitant changes in sensitivity and so, in effect, the encoded contrast is modulated even though the physical contrast is not. Similarly when flicker rate is modulated, although the physical contrast is the same at all points in the image (assuming a display with an adequate temporal modulation transfer function), encoded contrast is not, because of the roll-off in sensitivity at high temporal frequencies. The contrast modulation that is created in this way may form the basis of detection of motion. Once a contrast modulation exists in the representation of the image, it is vulnerable to first-order artifacts. Thus, although there are classes of second-order motion which might avoid the artifacts described in this paper, in practice it would be imprudent to switch to them without first conducting a detailed investigation of whether they truly solve the problem.

REFERENCES

Adelson, E. H. & Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America A, 2*, 284–299.

Anstis, S. M. & Rogers, B. J. (1975). Illusory reversal of visual depth and movement during changes of contrast. *Vision Research, 15*, 957–961.

Badcock, D. R. & Derrington, A. M. (1985). Detecting the displacement of periodic patterns. *Vision Research, 25*, 1253–1258.

Chubb, C. & Sperling, G. (1988). Drift-balanced random stimuli: A general basis for studying non-Fourier motion perception. *Journal of the Optical Society of America A, 5*, 1986–2006.

Derrington, A. M. (1987). Distortion products in geniculate X-cells: A physiological basis for masking by spatially modulated gratings? *Vision Research, 27*, 1377–1386.

Derrington, A. M. (1994). Analysis of the motion of contrast-modulated patterns. *Investigative Ophthalmology and Visual Science, 35*, 1406.

Derrington, A. M. & Badcock, D. R. (1985). Separate detectors for simple and complex patterns? *Vision Research, 25*, 1869–1878.

Derrington, A. M. & Badcock, D. R. (1986). Detection of spatial beats: Non-linearity or contrast increment detection? *Vision Research, 26*, 343–348.

Derrington, A. M., Badcock, D. R. & Henning, G. B. (1993). Discriminating the direction of second-order motion at short stimulus durations. *Vision Research, 33*, 1785–1794.

Derrington, A. M. & Henning, G. B. (1993). Detecting and discriminating the direction of motion of luminance and colour gratings. *Vision Research, 33*, 799–812.

Green, M. (1983). Contrast detection and direction discrimination of drifting gratings. *Vision Research, 23*, 281–289.

Harris, L. R. & Smith, A. T. (1992). Motion defined exclusively by second-order characteristics does not evoke optokinetic nystagmus. *Visual Neuroscience, 9*, 565–570.

Henning, G. B., Hertz, B. G. & Broadbent, D. E. (1975). Some experiments bearing on the hypothesis that the visual system analyses spatial patterns in independent bands of spatial frequency. *Vision Research, 15*, 887–897.

Johnston, A. & Clifford, C. W. G. (1995). Perceived motion of contrast-modulated gratings: Predictions of the Multi-channel Gradient Model and the role of full-wave rectification. *Vision Research, 35*, 1771–1784.

Ledgeway, T. (1994). Adaptation to second-order motion results in a motion aftereffect for directionally-ambiguous test stimuli. *Vision Research, 34*, 2879–2889.

Ledgeway, T. & Smith, A. T. (1994). Evidence for separate motion-detecting mechanisms for first- and second-order motion in human vision. *Vision Research, 34*, 2727–2740.

Ledgeway, T. & Smith, A. T. (1995). The perceived speed of second-order motion and its dependence on stimulus contrast. *Vision Research, 35*, 1421–1434.

Lindsey, D. T. & Teller, D. Y. (1990). Motion at isoluminance: Discrimination/detection ratios for moving isoluminant gratings. *Vision Research, 30*, 1751–1761.

MacLeod, D. I. A., Williams, D. R. & Makous, W. (1992). A visual non-linearity fed by single cones. *Vision Research, 32*, 347–363.

McCarthy, J. E. (1993). Directional adaptation effects with contrast modulated stimuli. *Vision Research, 33*, 2653–2662.

Mather, G. & West, S. (1993). Evidence for second-order motion detectors. *Vision Research, 33*, 1109–1112.

Mullen, K. T. & Boulton, J. C. (1992). Absence of smooth motion perception in color vision. *Vision Research, 32*, 483–488.

Smith, A. T. (1994). Correspondence-based and energy-based detection of second-order motion in human vision. *Journal of the Optical Society of America A, 11*, 1940–1948.

Smith, A. T., Hess, R. F. & Baker, C. L. Jr. (1994). Direction identification thresholds for second-order motion in central and peripheral vision. *Journal of the Optical Society of America A, 11*, 506–514.

Smith, A. T. & Ledgeway, T. (1995). Second-order motion: The carrier is crucial. *Perception, 24*, 28.

Solomon, J. A. & Sperling, G. (1995). 1st- and 2nd-order motion and texture resolution in central and peripheral vision. *Vision Research, 35*, 59–64.

Victor, J. D. & Conte, M. M. (1992). Coherence and transparency of moving plaids composed of Fourier and non-Fourier gratings. *Perception and Psychophysics, 52*, 403–414.

Watson, A. B., Thompson, P. G., Murphy, B. J. & Nachmias, J. (1980). Summation and discrimination of gratings moving in opposite directions. *Vision Research, 20*, 341–347.

Werkhoven, P., Sperling, G. & Chubb, C. (1993). The dimensionality of texture-defined motion: A single channel theory. *Vision Research, 33*, 463–486.

Wilson, H. R., Ferrera, V. P. & Yo, C. (1992). A psychophysically motivated model for two-dimensional motion perception. *Visual Neuroscience, 9*, 79–97.

Acknowledgements—This work has previously been presented in abstract form (Smith & Ledgeway, 1995). It was supported by a project grant from the Science and Engineering Research Council (Image Interpretation Initiative). The authors are grateful to Ms Tracey Francis for her diligence in acting as a subject over many weeks of experimentation and to Professor A. M. Derrington for suggesting the use of high-pass filtered carriers.

APPENDIX

Description of the Fourier Characteristics of Contrast-modulated Filtered Noise

In order to appreciate the effects of high-pass filtering the carrier (Experiments 5 and 6) on the first-order motion signals present in our second-order motion stimuli it is necessary to give a brief description of the Fourier characteristics (in the luminance domain) of contrast-modulated motion stimuli. The Fourier spectrum can be understood intuitively by analogy with an AM grating. If a stationary sine grating carrier with spatial frequency (f_c) is multiplied by a drifting sinusoidal envelope of lower frequency (f_s) then the resulting contrast-modulated grating has three components, one with frequency (f_c) and two sidebands with frequencies ($f_c - f_s$) and ($f_c + f_s$) referred to as the lower and upper sidebands, respectively. The component at (f_c) is stationary. Each sideband has a drift temporal frequency equal to that of the drifting contrast modulation. However, the lower spatial frequency sideband drifts in the opposite direction to the contrast modulation, while the higher frequency sideband drifts in the same direction as the contrast modulation.

Multiplication of a carrier consisting of static broadband noise by a drifting sinusoid of a given spatial frequency (f_s) produces a contrast-modulated image in which for every Fourier component originally in the carrier (f_c) two additional Fourier components (sidebands) are present in the modulated image. The spatial frequencies of the sidebands are equal to ($f_c - f_s$) and ($f_c + f_s$), as for an AM grating. Each has a drift temporal frequency equal to that of the drifting modulation in contrast and again the lower spatial frequency sideband drifts in the opposite direction to the contrast modulation and the higher frequency sideband drifts in the same direction.

Since sidebands are produced for all Fourier components present in the carrier, the net result, in the case of a broadband noise carrier, is that sidebands drifting in both directions are present at all spatial frequencies. In other words, the second-order motion stimulus is drift-balanced (Chubb & Sperling, 1988). In the case of the high-pass filtered second-order motion stimuli employed in Experiments 5 and 6, the same principles apply but the result is that the motion stimuli are no longer strictly drift balanced. They contain more first-order motion components drifting in the opposite direction to the contrast-modulation than in the same direction. Consider a broadband carrier which has been high-pass filtered with a sharp cut-off at 2 c/deg and is then contrast-modulated using a 1 c/deg sinusoidal envelope. The lowest spatial frequency in the filtered carrier, prior to multiplication with the 1 c/deg sinusoid, is 2 c/deg. Following multiplication, sidebands are introduced at various frequencies below 2 c/deg, the lowest being 1 c/deg (the lowest carrier component minus $f_s = 2 - 1 = 1$ c/deg). These sidebands all drift in the opposite direction to the envelope because they are all "lower sidebands" as defined above. The lowest spatial frequency sideband which drifts in the same direction as the modulation is at 3 c/deg (the lowest carrier component plus $f_s = 2 + 1 = 3$ c/deg). Thus, all moving components below 3 c/deg (in this example) drift in the opposite direction to the modulation because they are all lower frequency sidebands. For all components at or above 3 c/deg, each lower sideband will be balanced by an oppositely drifting upper sideband arising from a carrier component at some other frequency, as in the case of an unfiltered carrier. Hence, there is a range of spatial frequencies, centred on the filter cut-off frequency (centre frequency $\pm f_s$) in which all first-order motion components drift in the opposite direction to the envelope with the same temporal frequency as the envelope. There is no motion energy at spatial frequencies below this range and all spatial frequencies above this range are drift-balanced.