Imputation of direction of motion in one dimension

Miguel A. García-Pérez[†] and Eli Peli[‡]

[†] Departamento de Metodología, Facultad de Psicología, Universidad Complutense, Campus de Somosaguas, 28223 Madrid, SPAIN

[‡] Schepens Eye Research Institute, Harvard Medical School, 20 Staniford Street, Boston MA 02114-2500, USA

Abstract

Under many conditions below the Nyquist limit, a drifting grating briefly displayed on a CRT is consistently perceived moving in the opposite direction. Taking into account the sample-and-hold operation of CRTs we have derived the temporal-frequency spectra of the displayed gratings and found that they have a broad range of components moving in either direction. We have measured the perceived direction of motion of over 150 different short-duration stimuli, and we have studied the relation that performance bears to narrow-band power imbalance—the normalized difference between power in the positive- and negative-frequency half-lines within a specific band. Perceived direction of motion is highly related to power imbalance in 1-Hz-wide bands centered between 10 and 15 Hz, but none of these bands alone can account for more than 84-91% of the variance of the data, and each band ostensibly fails to explain data from a subset of the stimuli. When broadband power imbalance is determined by weighting the spectrum with an inverted-U-shaped function peaking at ~12 Hz, explained variance increases to 91-97%. Our results suggest that the imputation of direction of motion to stimuli with complex spectra is based on broadband power imbalance determined after weighting the temporal-frequency spectrum with an inverted-U-shaped function.

1. INTRODUCTION

Most scenes in the natural environment consist of a number of opaque objects each having a unique relative-motion vector in three-dimensional space, and our visual system seems to be well equipped to tell apart each of these objects and their trajectories. Yet, in laboratory conditions with artificial stimuli, the visual system is often confronted with the simple task of elucidating a global direction of motion, occasionally embedded in noise.

Extensive research has shown that our visual system is rarely capable of telling apart the individual motion vectors of superimposed gratings that are differently oriented in two-dimensional space. Instead, a rigid object (a plaid) is often perceived moving at a velocity different from that of either component.[1] In other words, the visual system seems to dispose of the individual objects and motion vectors so as to gauge a single object with a unique motion vector. We will refer to the process of collapsing diverse motion information into a single perceived motion vector as the *imputation of direction of motion*, but note that this process is not always operative: under some conditions superimposed gratings are perceived as transparently sliding over one another, each with its veridical motion.[1-3] This paper is concerned with the imputation of direction of motion to stimuli consisting of a diversity of components signaling motion in both directions in one-dimensional space (i.e., components differing as to speed and direction of motion but not orientation in two-dimensional space). We will focus on the decision rules implied in the process, and not on the conditions that trigger it.

Early work in this area has implicitly or explicitly assumed that motion imputation is determined by what Dosher et al.[4] called *net directional power*, the difference of power in the leftward- and rightward-motion areas of the (visible) spatiotemporal-frequency spectrum of the stimulus. Chubb and Sperling[5] discussed a number of situations where this general principle fails, something that has basically led to motion models where so-called non-Fourier channels coexist with their Fourier counterparts. Here we present further evidence that motion imputation is not based on such global characteristic as net directional power, and we also describe a preliminary attempt at characterizing the rules governing motion imputation in Fourier terms.

To further illustrate the problem we are addressing, Fig. 1 schematizes a cross-section of the temporal-frequency spectrum of a hypothetical stimulus consisting of two components moving in opposite directions in one-dimensional space, each at a different speed. In this stimulus, net power is identical in the positive- and the negative-frequency half-lines and, then, a mechanism that imputes direction of motion through net directional power will be unable to see any motion in it. Yet, suppose this stimulus is always perceived moving to the right (corresponding to positive temporal frequencies) with no trace of leftward motion. One could

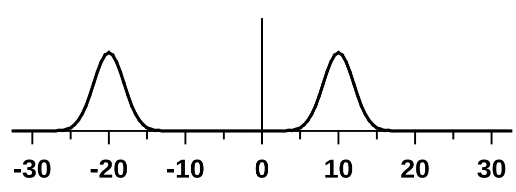

Figure 1. Temporal-frequency cross-section of the spatiotemporal-frequency spectrum of a hypothetical stimulus made of two components, one involving rightward motion with a temporal-frequency band around 10 Hz and the other involving leftward motion with a band around 20 Hz. The cross-section shows the profile at the nominal spatial frequency and orientation of the stimulus in the positive spatial frequency half-space; the corresponding cross-section in the negative half-space (not shown) is identical but reversed, as is the case for all physically realizable signals. By convention, negative and positive temporal frequencies respectively correspond to leftward and rightward motion.

then conclude that the low-frequency band prevails for motion imputation, and the opposite would hold if the stimulus was always perceived moving to the left. On the other hand, if both bands were about equally relevant for the imputation of direction of motion, their competition would result in a 50% chance of seeing the stimulus moving either way (as is the case for all multistable stimuli), or their effects would cancel out and no motion would be perceived (as is the case for flickering stimuli). As it turns out, in many cases stimuli thus defined evoke a motion percept: either both objects (gratings) are seen moving transparently in opposite directions, or a single object (a compound grating) is perceived whose direction of motion alternates over time. However, a rigid object with a unique direction of motion is rarely perceived in these stimuli.

In this research we have used a type of stimulus whose variants do not have these limitations. This stimulus is the usual Gabor patch with a static aperture and a moving carrier. Its characteristics that are relevant for our purposes become obvious at short durations, and they are a result of the way that stimuli are rendered on CRTs. Figure 2 shows temporal-frequency slices of the spatiotemporal-frequency spectra of some of these stimuli, whose space-time plots are also shown in the insets (see the Appendix for derivations). Stimuli consist of either two or three frames (frame duration: 8.16 ms) each of which displays a snapshot with a constant contrast throughout its duration and, when replaced, a discontinuity occurs and the phase of the carrier shifts abruptly by an amount that depends on the nominal speed of motion.[6]

The numeral in each panel of Fig. 2 indicates the percentage of times (out of 50 trials) that an observer (subject MA) perceived that stimulus moving to the right. Although these and other data will be thoroughly described below, it is worth advancing them here to illustrate the potential of our stimuli. Consider first the 2-frame sequences in the left column. At the lowest nominal velocity (top panel), the spectral composition of the stimulus is only slightly biased towards positive frequencies, something that may explain the less-than-perfect performance of the observer. As velocity increases, the spectrum shifts to the right and a notch occurs in the (visible) negative frequencies, resulting in more power in the positive-frequency side and thus perhaps explaining the prevalence of perceived rightward motion as indicated by the 100%-correct performance in the center and bottom panels.

Now consider the 3-frame stimuli in the middle column, each of which differs from its 2-frame counterpart simply by addition of one more frame. One might surmise that this modification can only improve performance, for the additional frame provides the observer with more motion information than is available in two frames. This seems to be

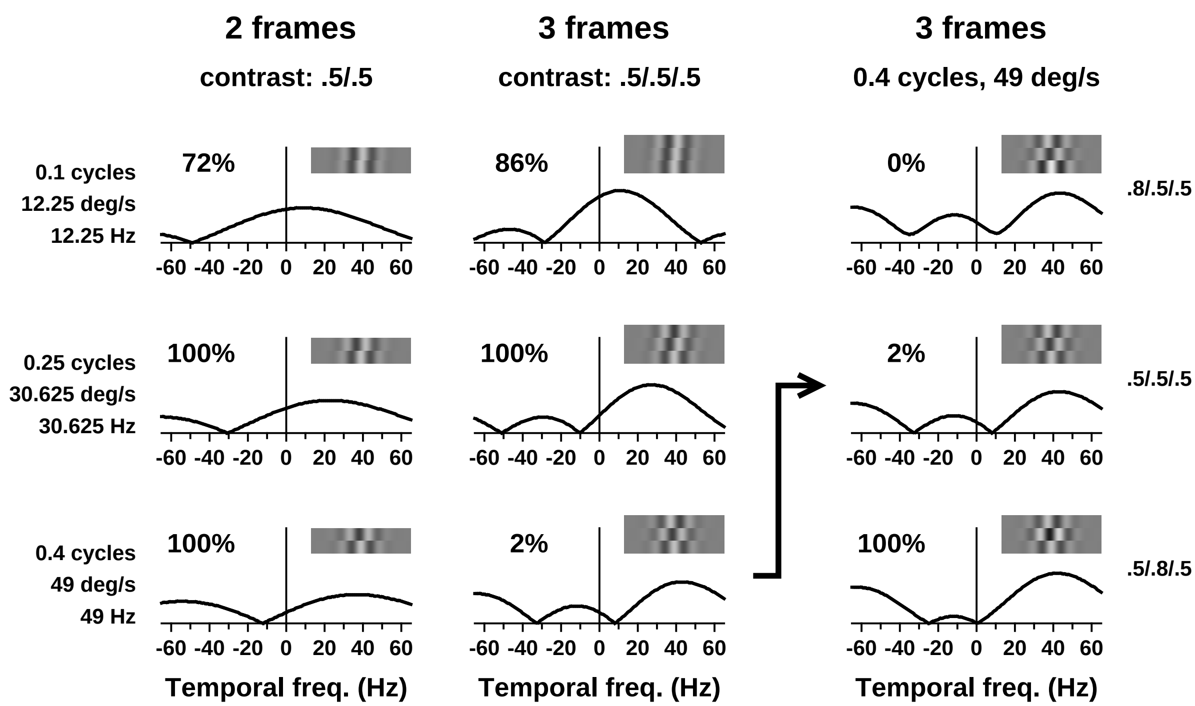

Figure 2. Temporal-frequency profiles of the spatiotemporal-frequency spectrum of sample stimuli used in the experiments. Space-time plots of the corresponding stimuli are shown in the insets; in these, the horizontal dimension is horizontal space (width: 5 deg) at the vertical center of the Gabor patch and the vertical dimension is time, showing two (left column) or three frames. The first frame is at the bottom in each image, and frame duration is 8.16 ms. The left column shows 2-frame presentations in which, top to bottom, the inter-frame phase shift amounts to 0.1, 0.25 and 0.4 cycles of the carrier (a 1 c/deg grating), thus resulting in nominal velocities of 12.25, 30.625 and 49 deg/s or, equivalently, nominal temporal modulations of 12.25, 30.625 and 49 Hz. The image displayed in each frame always had a Michelson contrast of 0.5. The middle column shows thoroughly analogous stimuli except for the addition of a third frame. The right column shows 3-frame presentations and in all three cases the phase shift is 0.4 cycles (as in the bottom panel in the middle column), but the contrast of one of the frames (the first in the top panel and the second in the bottom panel) is 0.8 rather than 0.5. The numeral on the left-hand side of each panel gives the percentage of times (out of 50 trials) that a subject reported perceiving rightward motion.

true at the lowest velocity (top panel), and it cannot show at the intermediate velocity (center panel) because performance was perfect with two frames. Yet, the additional piece of information has a dramatically adverse effect at the higher velocity (bottom panel), transforming perceived rightward motion in the 2-frame condition into a robust percept of leftward motion (as indicated by the 2%-correct performance). Comparing the spectra in the middle column with those in the left column, it is clear that addition of the third frame narrows the bandwidth of the stimulus in a way that a notch occurs closer to the peak. At low and intermediate velocities, this notch sits on the negative-frequency side, thus reinforcing the already prevalent power in the positive side; at higher velocities, the notch enters the positive-frequency half-line creating an imbalance in the low frequencies, for which there is more power now on the "wrong" side of the spectrum. Given that the subject perceived rightward motion in only one out of 50 trials with this stimulus, one has to accept that high temporal frequencies (at which there is still more power in the positive-frequency side) weigh less than low frequencies in the imputation of a direction of motion.

Note that the dramatic effect that addition of a frame has on performance with this stimulus is not a result of aliasing. All velocities implied in Fig. 2 result in nominal temporal frequencies that are below the Nyquist limit at the refresh rate of 122.5 Hz. All remaining conditions being identical, variations in the number of frames have no consequence on aliasing.

The experiments reported here involved variations in the duration (number of frames) of the stimulus as a means to modifying the distribution of temporal-frequency information in order to assess its quantitative effect on perceived direction of motion. We also varied contrast across frames, something that further affects the temporal-frequency spectrum as illustrated in the right column of Fig. 2. In the top panel, the first frame has a Michelson contrast of 0.8, and in the bottom panel it is the second frame that has a contrast of 0.8. Corresponding changes in the spectrum are easy to note by comparison with the constant-contrast case in the center panel: increasing the contrast of the first frame increases the amplitude of all frequencies and slightly shifts the spectrum to the right; increasing the contrast of the second frame (which restores the perception of rightward motion; note the 100%-correct performance) virtually removes all negative-sided low-frequency components.

2. METHODS

A. Apparatus

Stimuli were displayed at a mean luminance of 28 cd/m^2 on a 21-in EIZO FlexScan FX-E7 monitor at a frame rate of 122.5 Hz (frame duration: 8.16 ms). Throughout the course of the experiment, the monitor reset to a uniform field of this same luminance whenever a stimulus was not present. The monitor response was linearized by gamma correction, and all experimental events were under computer control. A Pentium-based PC platform with VisionWorks™ (Vision Research Graphics, Durham, NH) hardware and software was used for this purpose.

B. Stimuli

The nominal stimuli were spatiotemporal Gabor patches consisting of a vertical 1-c/deg carrier moving within a static circular gaussian aperture with a space constant of 0.65 deg [cf. eqn. (A1) in the Appendix]. The actual stimuli rendered on the monitor were obtained by displaying a static Gabor patch of the appropriate phase and contrast in each of the frames along the presentation period. The actual stimuli were thus described by eqn. (A2) in the Appendix. Presentation durations ranged from 16.33 ms (2 frames) to 163.27 ms (20 frames). The inter-frame phase shift was constant within each stimulus, but across stimuli it ranged from 0.05 cycles of the carrier (which results in a nominal velocity of 6.125 deg/s and a nominal temporal frequency of 6.125 Hz) to 0.5 cycles (61.25 deg/s; 61.25 Hz), in 0.05-cycle steps. Note that a 0.5-cycle jump renders a stimulus that flickers in square-wave counterphase and, thus, cannot elicit a direction-of-motion response. Its inclusion served to check for subject bias. The Michelson contrast of all frames was usually 0.5, but some stimuli were used in which one of the frames differed in contrast. All stimuli were clearly visible for all observers at the contrasts used in the experiments.

Stimuli were created as 150×150-pixel arrays which subtended 3.31×4.13 deg. Within the arrays, a cycle of the carrier spans 45.3 pixels, and the gaussian space constant is 29.4 pixels. Arrays were displayed on the center of the 1000×600-pixel image area and blended with the uniform surround.

Sketches of sample stimuli have been shown in Fig. 2, and note that they are not different from what in other contexts are summarily described as moving Gabor patches despite short presentation durations similar to ours.[7]

C. Procedure

Subjects sat 1 m away from the monitor. Their head was not restrained, but they were asked to maintain the same viewing distance within and across sessions. Note, however, that variations in viewing distance (which would have been small) slightly change the spatial frequency and velocity of the carrier, but not the temporal-frequency content of the patch. A cross (line length: 0.18 deg; luminance: 56 cd/m^2) overlaid at the center of the monitor served as a fixation aid.

Perceived direction of motion was determined with a 2AFC procedure and a direction-discrimination task. Each forced-choice trial consisted of the sequential presentation of two stimuli that were identical except that the inter-frame phase shift occurred in opposite directions in each one. Thus, the stimulus moved to the right in one of the intervals (chosen at random) and it moved at the same speed but to the left in the other. The subject had to indicate which interval displayed rightward motion. Intervals had the duration of the stimulus, each was marked by an audible tone, and they were separated by 498 ms (61 frames). Subjects could self-pace the experiment by rushing or delaying their responses, since the next trial did not start until a response had been given.

Each stimulus was presented 50 times, and the percentage of correct responses was computed. Thus, 100% correct implies perceived direction of motion matching the actual direction of motion, 50% correct indicates a failure to perceive any direction of motion, and 0% correct indicates perceived direction of motion opposite the actual direction of motion.

Data were collected in separate sessions each of which lasted 10-30 min. Each session consisted of 500 trials (order randomized) resulting from 50 presentations of each of 10 stimuli that differed only in speed. Each subject went through 19 such sessions to provide data at seven presentation durations (2, 3, 4, 5, 6, 10 and 20 frames) and 12 temporal contrast envelopes defined for 3-frame presentations by setting the Michelson contrast of either the first or the second frame at 0.0, 0.2, 0.4, 0.5, 0.6 or 0.8. Each subject completed these sessions in a newly randomized order.

D. Subjects

Three subjects with normal or corrected-to-normal vision participated in the study, including one of the authors who is an experienced psychophysical observer. The two other subjects were inexperienced observers and they were also naive as to the goals of the experiment.

3. RESULTS

Each panel in Fig. 3 shows results for one of the seven constant-contrast conditions defined by varying the number of frames, as a function of the nominal velocity (or temporal frequency) of the stimulus.[8] At the shortest duration (2 frames), perceived direction of motion is always veridical although it is impaired at low velocities. Veridical perception also occurs at the longest durations (6-20 frames), although in this case performance is impaired at high velocities. Yet, at intermediate durations (3-5 frames) veridical perception breaks down: in 3-frame sequences, reversed motion is perceived at the two highest velocities; in 4-frame sequences, no motion is perceived at intermediate velocities (as given by 50%-correct performance levels); and in 5-frame sequences one of the observers (subject MA) perceived reversed motion at the highest velocity.

Figure 4 shows results for the six 3-frame conditions in which contrast of the second frame varied, as a function of the nominal temporal frequency of the stimulus. As contrast of the second frame increases, the range of velocities at which reversed motion is perceived shifts upwards and narrows, eventually giving way to veridical perception. Note that direction of motion is aliased on the monitor in the .5/.0/.5 condition (left column, top panel), and all subjects' perception is veridical when this is taken into account. Note also that data for the .5/.5/.5 condition (left column, bottom panel) pertain to the same stimuli for

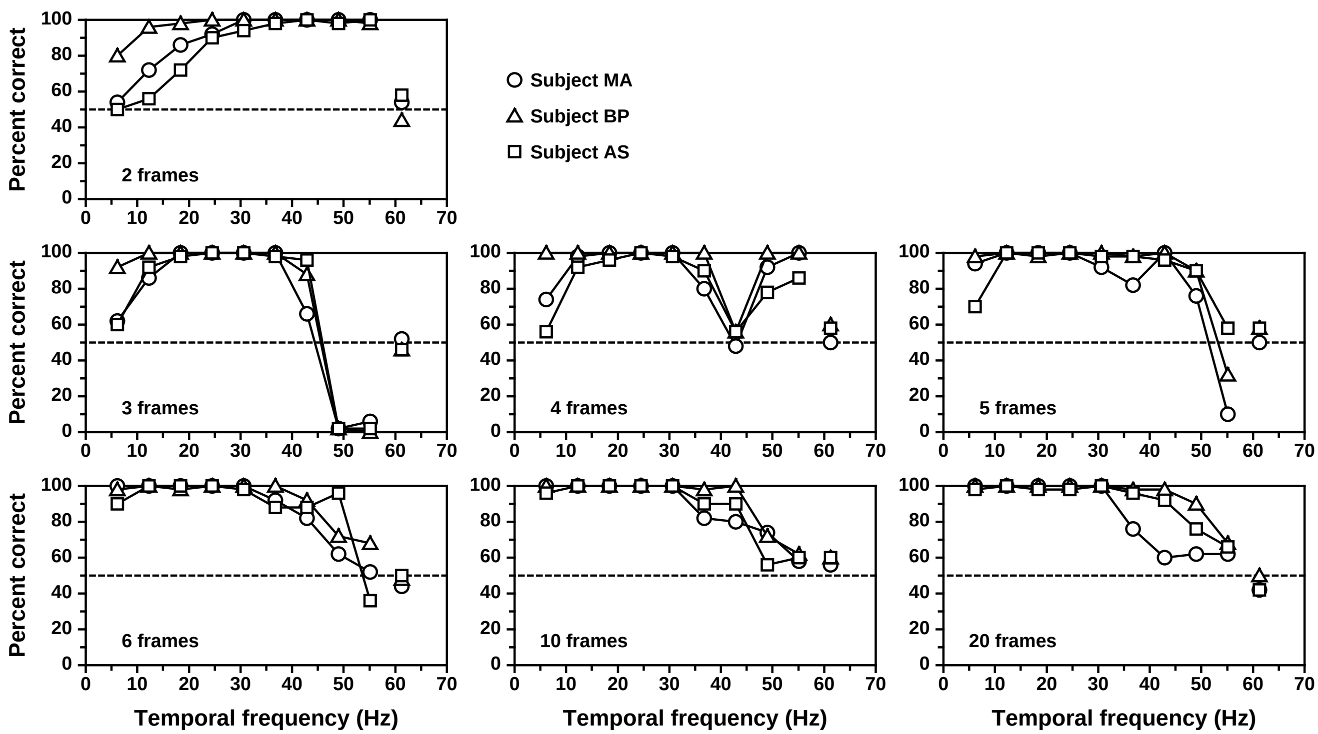

Figure 3. Percentage of correct responses for three subjects in the direction-discrimination task. Data from different stimulus durations are presented in separate panels, as labelled. Each panel presents data for different temporal frequencies (or, equivalently, velocities) as indicated in the abscissa. At the frame rate of 122.5 Hz, the highest temporal frequency used (61.25 Hz; 61.25 deg/s) corresponds to square-wave counterphase flicker rather than motion and, thus, serves as a control condition where performance should be at the 50% level if there is no perceptual bias. Performance levels above the dashed horizontal line at 50% indicate veridical perception of the direction of motion, while performance levels below it indicate that perceived direction of motion is opposite the actual direction of motion.

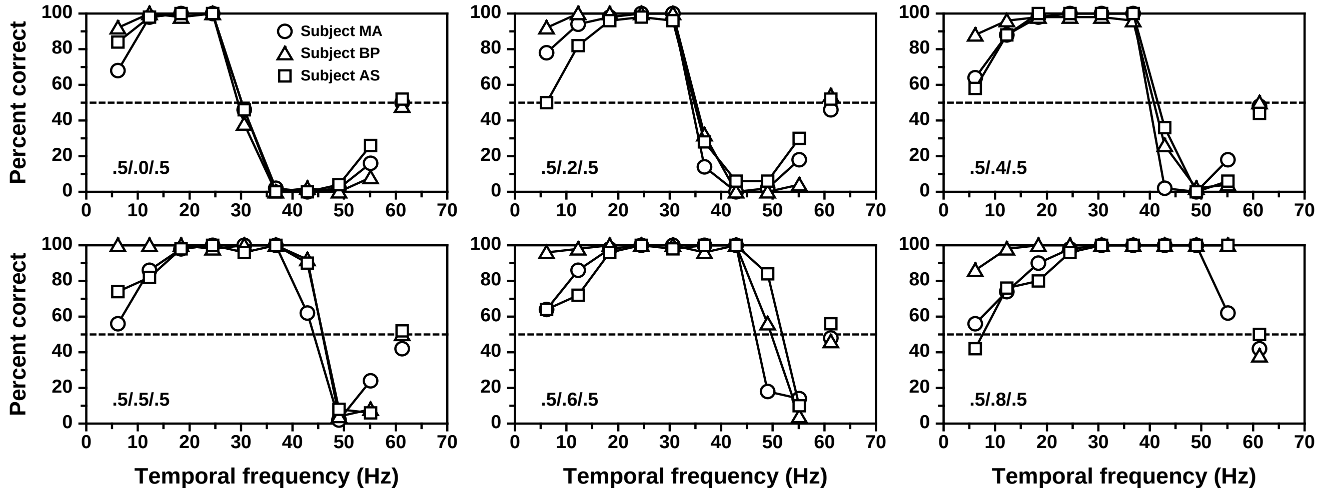

Figure 4. Similar to Fig. 3 except that data pertain to 3-frame presentations in which the first and third frames had a contrast of 0.5 and the second frame had a contrast valued as labelled in each panel. All graphical conventions as in Fig. 3.

which data were presented in the 3-frame panel in Fig. 3. The data are new for Fig. 4, and thus constitute a replication of the corresponding condition in Fig. 3.

Finally, Fig. 5 shows results for 3-frame presentations in which the contrast of the first frame varied, also as a function of the nominal temporal frequency of the stimulus. These results also show instances of veridical and non-veridical perception of the direction of motion. Note that the bottom panel in the left column is another replication of the 3-frame constant-contrast condition, and that the top panel in the left column is a replication of the 2-frame condition of Fig. 3.

We have also carried out a number of controls in which several characteristics of the Gabor patches varied while the power distribution in their temporal-frequency spectrum remained identical. These included (a) variations in contrast for the constant-contrast conditions displayed in Fig. 3, (b) variations in spatial frequency and, consequently, speed of the carrier, and (c) variations in patch size. In all cases the results were indistinguishable from those in Figs. 3-5, suggesting that only the distribution of information in the temporal-frequency spectrum matters.

In Fig. 2 we showed the spectrum of some of these stimuli, and we also mentioned how the data could be used to assess the relative weight of different temporal-frequency bands in the imputation of direction of motion. In the next section we address this issue in detail.

4. IMPUTATION OF DIRECTION OF MOTION

While data presented in Figs. 3-5 clearly indicate that the spectral content of our stimuli is critical for the perceived direction of motion, it is not obvious what mechanism may be responsible for that decision and how the balance of information on either side of the spectrum is used to arrive at it. The analysis we present here is more a first step towards understanding the stimulus information that is used by such mechanism than a final characterization of the rules governing its behavior. It should also be kept in mind that we are only considering stimuli that elicit either the perception of a global direction of motion or no motion perception

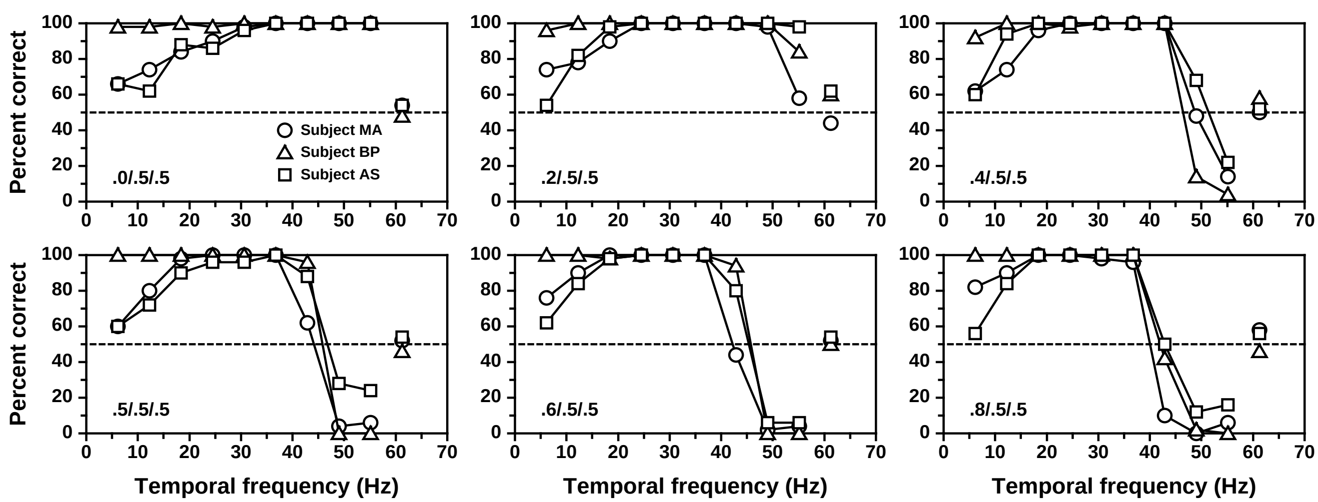

Figure 5. Similar to Fig. 4 except that it is now the first frame that varied in contrast, as labelled in each panel. All graphical conventions as in Fig. 3.

whatsoever.

With our stimuli, direction of motion is not aliased on the monitor except for the .5/.0/.5 condition in Fig. 4, which we have discussed already. Thus, the net power in rightward-moving stimuli is always larger in the positive- than in the negative-frequency side of the spectrum. Since our subjects clearly perceived leftward motion in many of these stimuli, it follows that imputation of motion is not based on net directional power.[9]

To check the validity of the notion that the local imbalance of information on either side of the temporal-frequency spectrum is relevant for the imputation of direction of motion, we have first carried out a simple analysis aimed at determining which temporal-frequency band could grossly account for the direction of motion perceived by our subjects. Suppose direction of motion were determined from information in a single, narrow temporal-frequency band in such a way that if the power in that band were larger on the positive side of the spectrum then rightward motion would be perceived, and leftward motion would be perceived if the opposite were to occur. Then, one could compute an index of power imbalance within that band and determine its relation to observed performance across stimuli. Moreover, variations in the strength of the relationship with band center would indicate their relevance: if power imbalance in some band were highly related to performance then that band must be important for the determination of the global direction of motion; if the relation were small, information in that band must not be used by the mechanism that imputes direction of motion.

Now suppose the power in some band is larger in the positive side. An imbalance index should evaluate to a positive value in this case, and it should attain the opposite value if the powers simply reversed. An index with these characteristics, similar to that used to measure direction selectivity in visual cortical cells, may be computed as

$$PI_i = \frac{P_i^+ - P_i^-}{P_i^+ + P_i^-}, \qquad (1)$$

where PI_i is the power-imbalance index at the i-th band and P_i^+ and P_i^- respectively stand for the power in the positive and negative sides of the spectrum at that band. This index is bounded between −1 and 1, attaining its limiting values when there is no power on one of the sides, and attaining a null value when power is the same on both sides. We expect performance to be monotonic with this index: a value of −1 (indicating absence of power on the positive side) should relate to leftward-motion perception as represented by a performance level of 0% correct; a value of 0 (indicating balance of power on either side of the spectrum) should be associated with performance at the 50% level; and a value of 1 (indicating absence of negative-sided power) should be associated with rightward motion (100%-correct performance). Figure 6 illustrates how this index varies across bands for a given stimulus, with bands defined to have a (linear) width of 1 Hz and centered between 1 and 65 Hz in 1-Hz steps. For this stimulus, power imbalance at frequencies below 18 Hz is consistent with the perception of leftward motion, while at higher frequencies it is consistent with rightward-motion perception. We have computed the distribution of power imbalance for each of the stimuli for which we presented data in Figs. 3-5 and then, separately for each subject, we have analyzed the relation that performance bears to power imbalance at each of these bands.

Performance is plotted against power imbalance at two bands (3 and 11 Hz) in Fig. 7a, revealing strong monotonic relations. We have measured the strength of the relation at each band by fitting a constrained, one-parameter logistic function of the form

$$\text{PC} = \frac{100}{1 + \exp\left(-b\,PI_i\right)}, \qquad (2)$$

where PC is percent correct, PI_i is power imbalance at the corresponding band [from eqn. (1)] and b is a parameter to be estimated. The constraint implicit in eqn. (2) is that we force the logistic function to cross the 50%-correct level at null imbalance. Our choice of a logistic function is based on its symmetry, a characteristic of our imbalance index and our performance measure that the fitted function should accommodate.

The panels in Fig. 7a also show the best-fitting function in each case. Note that there are 190 data points in each panel (performance at 10 nominal velocities within each of 7 duration conditions and 12 contrast conditions,[11] plotted against power imbalance in the corresponding stimulus at

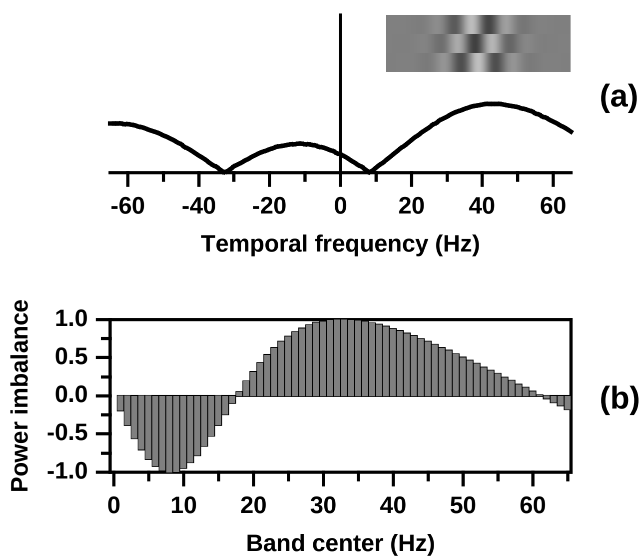

Figure 6. Illustration of the approach to analyzing the relevance of each temporal-frequency band for the imputation of direction of motion. A sample stimulus is shown in (a) along with its temporal-frequency spectrum. This panel is the same as shown at the bottom of the middle column in Fig. 2. The power imbalance in that stimulus is shown in (b) as a function of band center. Imbalance is computed through eqn. (1) for each of 65 bands with a (linear) width of 1 Hz and centered at integral temporal frequencies between 1 and 65 Hz. Note that imbalance is negative at frequencies below 18 Hz and above 61 Hz (implying that there is more power on the negative side of the spectrum within the corresponding bands), while it is positive at temporal frequencies between 18 and 61 Hz inclusive.

the selected band) and we aim at describing them all with a one-parameter function. The percentage of variance that the fitted function explains is given in each panel, and Fig. 7b shows summary results for all bands separately for each subject. Explained variance is largest in the 10-15 Hz range, and it becomes trivially small only for bands above ~30 Hz. Clearly, these percentages can be taken as gross indices of the relative relevance of different bands for the imputation of direction of motion. Yet, as discussed next, this imputation is more likely made on the basis of information collected across the entire spectrum than from whatever is encountered within a single, narrow band.

The panels in Fig. 7a suggest that this must be the case: despite the tight packing of data, all panels contain a few stray data points indicating perception of rightward motion (performance at or near 100% correct) when power imbalance is negative and large. These data points pertain to the same stimuli across subjects, indicating that their location on the plots is not without reason. Note also that stray data points occur mostly for imbalance measured at the 3-Hz band (left panels), thus resulting in less variance explained despite the fact that the remaining data points are more tightly packed around the sigmoid than when imbalance at the 11-Hz band is considered (right panels). Performance on this small subset of stimuli must then be related to power imbalance at some other band. Yet, no single-band relationship displayed better characteristics than those in Fig. 7a, suggesting that narrow-band approaches are insufficient for a satisfactory account of performance on all stimuli.

The obvious next step is to consider that perceived direction of motion depends on broadband weighted power imbalance. Yet, broadband schemes are countless and it is not clear how experiments could be designed that would unequivocally indicate the actual approach taken by our visual system. Thus, we have simply tried out various linear schemes which wound up differing enormously as to their success. We will next describe some of them, including the one that appears to be most successful.

Our approach is as follows. Let $|F|$ be the amplitude spectrum of the stimulus, let w be an even-symmetric weighting function, define broadband weighted power in the positive and negative temporal-frequency half-lines respectively as

$$P^- = \int_{-\infty}^{0} \Big[|F(\rho_0, 0, \omega)|\, w(\omega) \Big]^2 d\omega \tag{3}$$

and

$$P^+ = \int_{0}^{\infty} \Big[|F(\rho_0, 0, \omega)|\, w(\omega) \Big]^2 d\omega, \tag{4}$$

and compute overall power imbalance by inserting P^+ and P^- into a straightforwardly modified eqn. (1). We have tested several functions including unrestricted constant weights [i.e., $w(\omega) = 1$], constant weights applied only within the temporal-frequency bounds of the window of visibility[10] [i.e., $w(\omega) = 1$ if $|\omega| \le 30$ Hz, and 0 otherwise] and weights inversely related to temporal frequency [i.e., $w(\omega) = 1/\omega$], but inverted-U-shaped weighting functions

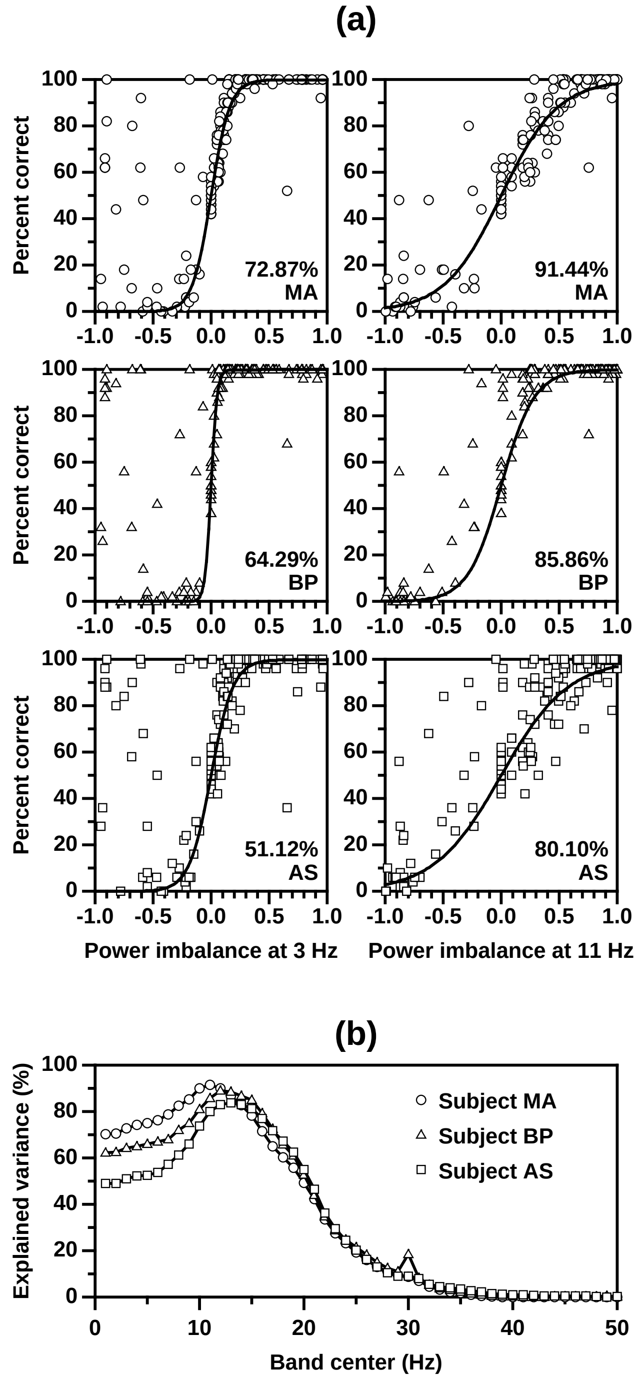

Figure 7. (a) Performance from three subjects (rows) plotted against the power imbalance at each of two frequency bands (columns). Each panel contains 190 data points whose ordinates are the same across all panels pertaining to the same subject (i.e., across rows), representing performance of the corresponding subject across conditions in the panels of Figs. 3-5. The abscissa for each point is the imbalance at 3 Hz (left column) or 11 Hz (right column) computed for the corresponding stimulus as illustrated in Fig. 6. Continuous curves are one-parameter, constrained logistic functions [cf. eqn. (2)] separately fitted to the data in each panel, and the numeral in each panel is the percentage of variance explained by the fitted function. (b) Summary plot of the percentage of variance explained by functions fitted as illustrated in (a) at each of the temporal-frequency bands for which power imbalance was computed. Explained variance is only trivially different from zero for all bands beyond 40 Hz, including the range 50-65 Hz not shown in the figure.

worked best.

Results for these four weighting functions are shown in Fig. 8. With a constant weighting function, power is computed for each stimulus as is. As a result, power imbalance is always non-negative[12] (since all stimuli move to the right) and in no way is it related to performance despite a spurious 35-44% of variance explained by a fitted logistic function (see Fig. 8c). Limiting the constant weighting function to the temporal-frequency bounds of the window of visibility does not do a much better job (see the departure between the scatter of data and the path of the fitted function in Fig. 8d), despite a marked (but also spurious) increase in explained variance. A weighting function of inverse temporal frequency does a better job (see Fig. 8e) in terms both of explained variance (which is larger on a subject by subject basis than it is for constant weights) and match between the scatter of data and the path of the fitted function. Yet, there are still data points at large vertical

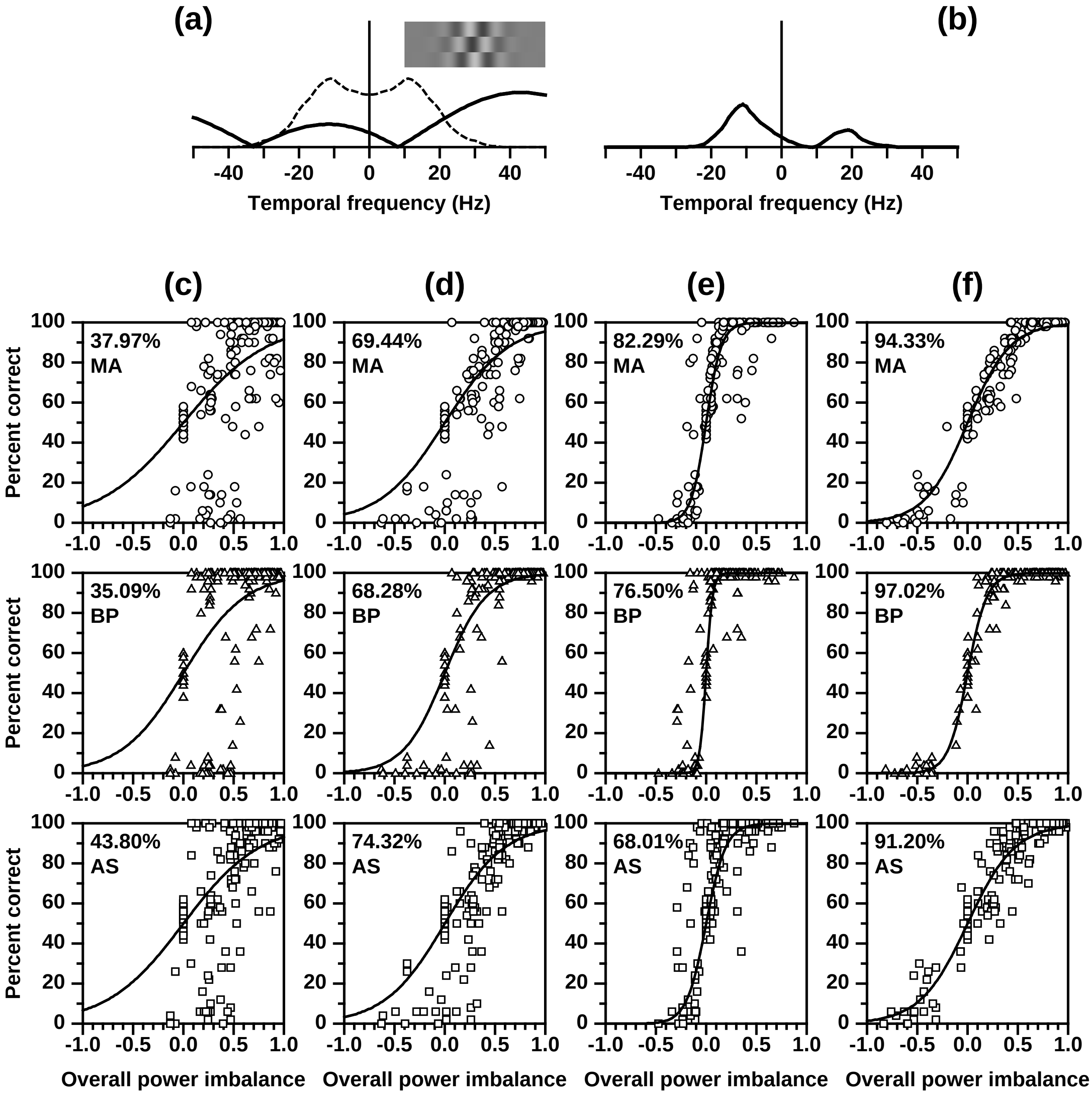

Figure 8. Performance plotted as a function of overall power imbalance defined by first multiplying the amplitude spectrum of each stimulus by a weighting function of temporal frequency, and then computing power imbalance through eqn. (1) but using the weighted power on the positive and negative frequency sides as defined in eqns. (3) and (4). Panel (a) shows a stimulus (inset) along with its spectrum (thick line) and a weighting function (dashed line) identical in shape to the function describing variance explained for subject MA in Fig. 7b. Panel (b) shows the resultant weighted spectrum. The panels underneath show plots of performance (separately for each subject) against overall power imbalance computed using (c) a constant weighting function, (d) a constant weighting function that applies null weight beyond the temporal-frequency bounds of the window of visibility ($|\omega| \leq 30$ Hz), (e) a function of inverse temporal frequency, and (f) an inverted-U-shaped function that has the same exact form as the narrow-band explained-variance function shown for the corresponding subject in Fig. 7b. Each panel also shows the best-fitting one-parameter logistic function and the percentage of variance that the latter explains.

offsets from the fitted curve, just as was the case in most of the narrow-band fits shown in Fig. 7a. Note also that none of the weighting functions considered so far has more explanatory power than the best narrow-band approach that can be identified for each subject in Fig. 7.

Figure 8f finally shows near-optimal results obtained when the spectra of the stimuli are weighted by the function displayed in Fig. 7b for the corresponding subject (see illustration in Figs. 8a and 8b). The fit is very good: there are no points in any subject's data that depart vertically from the path of the fitted function.[13] Note also that the percentage of variance explained by the fitted function varies from 91% to 97%. These two characteristics reveal that a broadband weighted approach is more explanatory than any of the individual narrow-band attempts considered in Fig. 7a.

It is somewhat striking that the functions in Fig. 7b, relating explained variance to power imbalance within narrow bands, would do such a good job in the weighted scheme of Fig. 8. Indeed, there is no statistical reason why explained variance should give the weight of a variable for a subsequent linear or nonlinear combination. It may, then, be a simple coincidence that the functions in Fig. 7b have a shape that we indeed wanted to consider: larger weights at intermediate frequencies that fall off more rapidly towards high than towards low frequencies, describing a shape similar to that of the temporal contrast sensitivity function (CSF).

Beyond this point, one could try to search for other functions that provide a quantitatively better fit, perhaps by estimating the weight of each frequency band instead of using weights from Fig. 7b and simply fitting a one-parameter function to the data, or by using the temporal CSF of each subject as a weighting function. However appealing, this latter approach may not be worth pursuing: first, the CSF describes performance at threshold and it is not clear that its inverted-U shape describes appropriately the filtering characteristics of the visual system in the supra-threshold conditions of our direction-discrimination task;[4] secondly, the shape of the temporal CSF varies with the duration of the stimulus, and since our explanatory approach requires using a single weighting function for all stimuli regardless of their varying durations, it is not clear which duration should the stimuli used to determine the temporal CSF have in order to appropriately predict performance in our direction-discrimination task. Considering the already large amount of variance explained by our simple approach and given the internal variability of the data, this extra effort can only produce a measly improvement in fit at the cost of estimating a substantial number of additional parameters.

5. CONCLUSIONS

We have shown that a potentially harmful artifact of the sample-and-hold operation of CRTs can be exploited to study basic properties of visual motion analysis, namely, how conflicting information about one-dimensional motion is integrated to impute a unique direction of motion. Our results, arising from separate analyses of data from three subjects, indicate that perceived direction of motion is determined by the overall imbalance of power on either side of the weighted temporal-frequency spectrum of the stimulus, where the weighting function has an inverted-U shape. Large imbalance results in a strong percept of motion, while small or null imbalance results in percepts of temporally-varying stimuli that nevertheless lack a direction of motion.

It should be stressed that this picture applies only under the conditions of our study, i.e, when a single object is perceived with a unique direction of motion in one-dimensional space. Obviously, whether a stimulus has this appearance is not a precondition but a consequence of visual processing. Yet, our state of knowledge is insufficient to explain why some stimuli (like ours) are perceived as a single object moving rigidly in one direction while others (like the superimposed gratings idealized in our Fig. 1) fail to elicit the perception of rigid motion.

Our results support the hypothesis that integration of diverse and conflicting temporal-frequency information occurs in a specific way for the imputation of direction of motion in one-dimensional space. Further work is necessary to determine where the accountable mechanism lies in a hierarchy of local spatiotemporal mechanisms for motion analysis and representation, and also to determine the conditions that trigger it. Experimental work is also needed to study whether some form of integration also occurs in the imputation of a unique direction of rigid motion in two-dimensional space.

APPENDIX

This appendix derives the spatiotemporal-frequency spectrum of a moving Gabor patch as rendered by a CRT that operates at a frame rate of ω_s Hz and, thus, has an image update rate of $\Delta = \omega_s^{-1}$ seconds. (A more general description of sample-and-hold motion can be found in Appendix B of Ref. [14].) Since we are mostly concerned with temporal-frequency artifacts, we will regard the two spatial dimensions as continuous. Let the nominal Gabor patch be

$$f_N(x,y,t) = L_0\left[1 + m(t)\exp\left(-\frac{x^2+y^2}{2\sigma^2}\right)\cos\big(2\pi\rho_0(x - v_0 t) - \phi_0\big)\right], \tag{A1}$$

where L_0 is mean luminance, m is a temporal contrast envelope function, σ is the space constant of the circular gaussian aperture, ρ_0 is the spatial frequency of the carrier, v_0 its velocity, and ϕ_0 its initial phase. Note that eqn. (A1) describes a moving Gabor pattern that extends from negative to positive time given some arbitrary origin.

For display on a CRT, the origin of time is set at stimulus onset and the patch is presented for a short time, an operation that can be represented as multiplication of eqn. (A1) by a temporal rectangular window spanning n frames and,

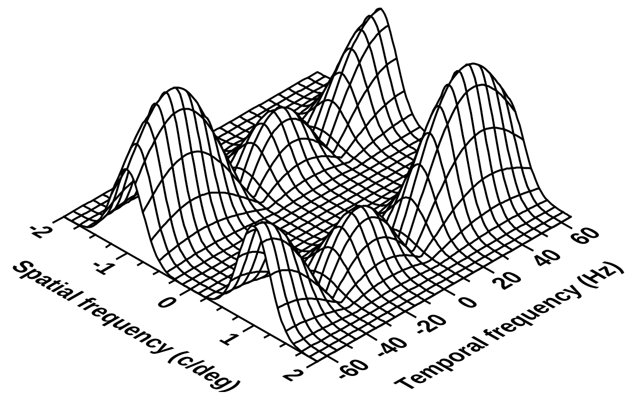

Figure A1. Two-dimensional slice (at $\rho_y = 0$ c/deg) of the amplitude spectrum of a 3-frame stimulus with parameters $\rho_0 = 1$ c/deg, $\sigma = 0.65$ deg, $v_0 = 49$ deg/s, $\phi_0 = 0$ rad, and $m(t) = 0.5$, and presented on a monitor with an image update rate $\Delta = 122.5$ s [cf. eqns. (A1) and (A2)]. A space-time plot of this stimulus is shown as an inset in the bottom panel in the middle column of Fig. 2, which also shows a one-dimensional cross-section of the amplitude spectrum taken at $\rho_y = 0$ c/deg and $\rho_x = 1$ c/deg.

thus, whose duration is $n\Delta = n\omega_s^{-1}$ seconds. Further, the k-th frame ($k = 0, ..., n-1$) presents a static image resulting from sampling the function in eqn. (A1) at $t = k\Delta$ and holding that two-dimensional spatial sample for the entire frame duration. This sample-and-hold operation can formally be represented by the product with a sampling function of period Δ seconds followed by convolution with a rectangular window of duration Δ seconds. This results in the actual image displayed on the CRT being

$$f(x,y,t) = \Pi_{n\Delta}(t)\left[f_N(x,y,t)\,Ш(t) * \Pi_\Delta(t)\right], \tag{A2}$$

where the star denotes convolution, Ш is the sampling function,[15] and

$$\Pi_\tau(t) = \begin{cases} 1 & \text{if} \quad 0 \le t \le \tau \\ 0 & \text{otherwise} \end{cases}. \tag{A3}$$

The Fourier transform, F, of eqn. (A2) is easily shown to be

$$\begin{aligned} F(\rho_x,\rho_y,\omega) = {} & L_0\,{}^3\delta(\rho_x,\rho_y,\omega) + L_0\frac{\sin(\pi\omega\Delta)}{2\pi\omega}\times \\ & \sum_{k=0}^{n-1} m(k\Delta)\left\{\exp\left[-2\pi^2\sigma^2\left((\rho_x-\rho_0)^2+\rho_y^2\right)\right]\exp\left[-i2\pi\Delta\left(\left(k+\tfrac{1}{2}\right)\omega-kv_0\rho_0\right)\right]+\right. \\ & \left.\exp\left[-2\pi^2\sigma^2\left((\rho_x+\rho_0)^2+\rho_y^2\right)\right]\exp\left[-i2\pi\Delta\left(\left(k+\tfrac{1}{2}\right)\omega+kv_0\rho_0\right)\right]\right\}, \end{aligned} \tag{A4}$$

with $i^2 = -1$, from which the amplitude spectrum can readily be obtained. Figure A1 shows a slice (at $\rho_y = 0$ c/deg) of the amplitude spectrum of the 3-frame stimulus in the bottom panel in the middle column of Fig. 2. For simplicity, elsewhere in this paper we are showing one-dimensional profiles taken at $\rho_x = \rho_0 = 1$ c/deg and $\rho_y = 0$ c/deg, and we have omitted the function ${}^3\delta$ at the origin because it does not affect power imbalance.

ACKNOWLEDGEMENTS

This work was carried out at The Schepens Eye Research Institute, where MAGP was a *Research to Prevent Blindness International Research Scholar.* MAGP was supported by DGES grant PB96-0597. EP was supported by NIH grants EY05957 and EY10285, and by NASA contract NCC-2-1039. We thank Vicente Sierra-Vázquez for his help with the derivations in the Appendix.

REFERENCES AND NOTES

1. E. H. Adelson and J. A. Movshon, “Phenomenal coherence of moving visual patterns,” Nature **300**, 523–525 (1982).
2. A. V. van den Berg and W. A. van de Grind, “Conditions for the detection of coherent motion,” Vision Res. **31**, 1039–1051 (1991).
3. J. Kim and H. R. Wilson, “Dependence of plaid motion coherence on component grating directions,” Vision Res. **33**, 2479–2489 (1993).
4. B. A. Dosher, M. S. Landy and G. Sperling, “Kinetic depth effect and optic flow—I. 3D shape from Fourier motion,” Vision Res. **29**, 1789–1813 (1989).
5. C. Chubb and G. Sperling, “Drift-balanced random stimuli: a general basis for studying non-Fourier motion perception,” J. Opt. Soc. Am. A **5**, 1986–2007 (1988).
6. This description is only a convenient simplification. Line raster, phosphor response and display bandwidth all contribute to making the actual image displayed in each frame slightly different from what we just described.
7. A. B. Watson and K. Turano, “The optimal motion stimulus,” Vision Res. **35**, 325–336 (1995).
8. Data at the highest nominal velocity (isolated points at the right in each panel) serve as a control condition of flicker and, not surprisingly, performance is near the 50% (guessing) level in all cases. We will no longer refer to these data.
9. As will become clear in our discussion around Fig. 8, net directional power cannot account for our results even if it is computed within the temporal-frequency bounds of the window of visibility.[10]
10. A. B. Watson, A. J. Ahumada, Jr. and J. Farrell, “Window of visibility: a psychophysical theory of fidelity in time-sampled visual motion displays,” J. Opt. Soc.

Am. A **3**, 300–307 (1986).

11. Recall, however, that only 16 of the 19 conditions were actually different, since one of them (constant-contrast 3-frame presentations) was replicated twice and another (constant-contrast 2-frame presentations) was replicated once.
12. There are a few instances of stimuli that come out with small negative power imbalance values, but this is only an artifact caused by our using temporal frequencies only up to 65 Hz in the computation of P^- and P^+ through eqns. (3) and (4).
13. The inherent variability in the data should be taken into account for this judgement. Data from our conditions with flickering stimuli give an estimate of this variability, and these data can easily be seen in the panels of Fig. 8c as a string of vertically aligned points at an abscissa of 0 and around an ordinate of 50. The vertical spread of these strings gives an indication of the response variability that one has to put up with in these plots.
14. D. J. Fleet and K. Langley, "Computational analysis of non-Fourier motion," Vision Res. **34**, 3057–3079 (1994).
15. R. N. Bracewell, The Fourier transform and its applications. (McGraw-Hill, New York, 1978).