



Motion detection and directional tuning

William A. Simpson *, Aaron Newman

Department of Psychology, University of Winnipeg, 515 Portage Avenue, Winnipeg, Manitoba R3B 2E9, Canada

Received 25 April 1995; received in revised form 24 March 1997; accepted 14 July 1997

Abstract

A random dot pattern was presented which made two jumps in various directions with a variable delay between them. The jumps occurred at the frame transitions of a 3-frame apparent motion sequence. The variation in detectability with the directional difference and the temporal separation of the jumps allows us to make inferences about directional tuning and the temporal response of the motion detection mechanism. The detectability of a pair of jumps was highest when the delay between the jumps was short and the difference in the jump directions was small. In all cases the data were well fitted with a vector version of the speed energy model earlier proposed by Simpson. The model supposes that the two input vectors are temporally filtered, squared and integrated. Using the model, the autocorrelation function of the motion system's temporal impulse response can be found. This function shows the filter to be lowpass. According to the model, the shape of the threshold or d' locus as a function of the difference in the directions of the two jumps does not show the tuning of a motion mechanism. A tuned mechanism will respond well to a jump in its preferred direction, but less well to any other jump. Instead we show that the apparent tuning evident in the threshold and d' loci is due to the way in which the two jump vectors, each fully recovered, are combined in a vector sum. © 1998 Elsevier Science Ltd. All rights reserved.

Keywords: Motion detection; Motion temporal response; Motion direction sensitivity

1. Motion detection directional tuning

Velocity is a vector, having both direction and speed. In this study we examine the directional tuning of the mechanism which underlies speed detection, the discrimination of a zero from a nonzero speed. We present a random dot pattern which performs two sudden jumps (apparent motion). The difference in the directions of the first and second jumps is varied, as is the delay between them. Intuitively, one would expect that the closer together the two jumps are in direction and in time, the more detectable the speed. By presenting a number of delays and directions, the directional and temporal tuning of speed detection can be measured.

Our new technique is a modification of the two-jump method used by Simpson [1], which was in turn based on Rashbass's [2] classic two-flash experiment. Formally, a jump is a speed impulse or Dirac delta function [1]. A pair of jumps is created by a three-frame appar-

ent motion display, with each jump sitting at a frame transition (frame 1/frame 2; frame 2/frame 3). The main modification of the method as used here is that whereas before the jumps were always horizontal, now we systematically vary their direction. The delay between the jumps is also varied. Motion detectability is measured by the minimum jump amplitude (displacement) for discriminating moving from stationary displays, or by d' [3] (pp. 60–66) for fixed size jumps.

In order to get some predictions about the outcome of the two-jump experiments, we now sketch three simple models whose starting point is the speed energy model of Simpson [1]. Simpson originally called this a velocity energy model, but it is properly called a speed energy model since it uses the scalar speed rather than the vector velocity. The model asserts that the speed waveform delivered to the observer is passed through a motion filter (in the motion system), squared and integrated. The motion system consists of those parts of the brain that extract a motion signal from the spatiotemporal luminance distribution that falls on the retina. Symbolically, the model is:

* Corresponding author. Tel.: +1 204 7869406; fax: +1 204 7744134; e-mail: wsimpson@uwinnipeg.ca.

$$E = \int_0^\tau [v(t)*h(t)]^2 dt \tag{1}$$

where E is the velocity energy and $v(t)*h(t)$ is the convolution of the speed signal in the world, $v(t)$, with the impulse response of the motion filter, $h(t)$. The true speed waveform is $v(t)$, but the brain has no direct access to that. The retina receives the spatiotemporal luminance distribution $L(x,y,t)$. The motion filter extracts some version of $v(t)$ from $L(x,y,t)$. The internal representation of $v(t)$ is the convolution of $v(t)$ and $h(t)$. This filtered speed is then squared and integrated over a duration τ that is longer than the stimulus duration. In Simpson’s original experiments the motion was always horizontal and so the stimulus motion was the scalar $v(t)$. Now however the jumps can be in different directions and so the stimulus is the vector $\mathbf{v}(t)$.

The models under consideration differ in the information extracted from $\mathbf{v}(t)$ prior to squaring and integration. The amplitude model simply uses the amplitude (displacement or speed) of each jump, ignoring its direction. The directional tuning model uses the full speed of the first jump, but only the component of the second jump in the same direction as the first. In the vector model, the velocity energy of the full jump vectors (speed and direction) is computed.

1.1. Amplitude model

In this model only the amplitude or speed of each jump is used and no information about the directions of the jumps is used at all. The observer discriminates a pair of jumps from a stationary display. d' is measured as a function of the difference in the directions of the first and second jump and as a function of the delay T between them. It is shown in the Appendix A that the prediction is

$$d' = k[a^2 + b^2 + 2ab R_h(T)] \tag{2}$$

where a and b are the amplitudes of the first and second jumps, k is the reciprocal of the standard deviation of the internal motion noise and $R_h(T)$ is the autocorrelation function of $h(t)$, the motion filter. Since the amplitudes of both jumps are the same, call both a , giving

$$d' = 2a^2k[1 + R_h(T)]. \tag{3}$$

There is no term in Eq. (3) for the direction of the jumps and so d' is a fixed value regardless of the jump directions. In polar coordinates (r,θ) , where r is d' and θ is the difference in directions, the predicted threshold locus is a circle centred on the origin (Fig. 1(a)). (We use the term locus in the same sense as in the expression ‘spectrum locus’: it is a set of points or curve).

At threshold $d' = 1$, so the threshold size of a is

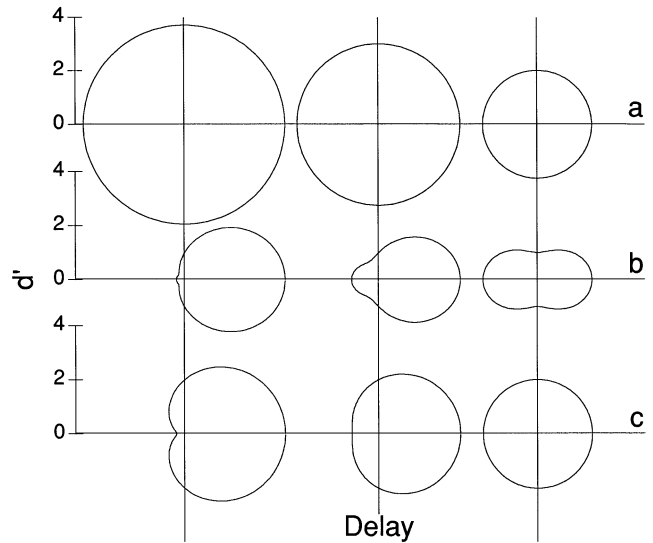


Fig. 1. Polar plots showing the predicted d' (r) as a function of the difference in the direction of two jumps (θ). In the amplitude model (a), the energy of the speeds of the jumps is computed. In the directional tuning model (b), the energy of the speed of the first jump and of the component of the second jump in the same direction as the first is computed. In the vector model (c), the velocity energy of the vectors for jump₁ and jump₂ is computed. For each row the delay between the two jumps increases from left to right.

$$a_{\text{thresh}} = \frac{\sqrt{2}}{\sqrt{k[1 + R_h(T)]}} \tag{4}$$

Direction is not represented in Eq. (4) either. In polar coordinates where r is the threshold and θ is the difference in directions, the prediction is again a circle centred on the origin (Fig. 2(a)).

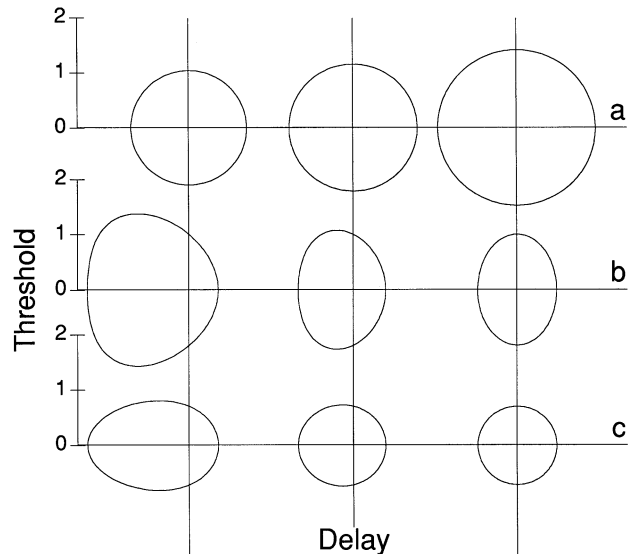


Fig. 2. Polar plots showing the predicted threshold jump amplitude (r) as a function of the difference in the direction of two jumps (θ). Both jumps have the same amplitude; the threshold is for each jump. The predictions of the amplitude model (a), the directional tuning model (b) and the vector model (c) are shown. For each row the delay between the two jumps increases from left to right.

The predictions for d' and threshold loci show absolutely no directional tuning. This is really not surprising, since we assumed that only the jump size matters, not the direction, and this comes out in the model prediction. Since the observer is not required to make a judgement about motion direction, this model and its prediction seem plausible.

1.2. Directional tuning model

Suppose that both jumps stimulate the same motion sensor and that this sensor is tuned to direction. If the jumps are in very different directions the detectability of the pair will be low, since one of the jumps will be in a direction that the sensor does not respond well to. Conversely, if the two jumps are in similar directions the detectability of the pair will be high, since both jumps are near the peak of the sensor's tuning curve. This simple and intuitive way of thinking about the experiment we call the directional tuning model. The motion sensor responds to the first jump completely (or very well), but not so well to the second, since it is in a different direction. To elaborate the model further, we can suppose that only the component of the second jump along the same direction as the first is extracted. If the directions of jump₁ and jump₂ differ by the angle θ , the component of jump₂ in the same direction as jump₁ is just $b \cos \theta$. Substituting into Eq. (2),

$$d' = k[a^2 + (b \cos \theta)^2 + 2ab \cos \theta R_h(T)]. \quad (5)$$

Since a and b are equal, represent both by a

$$d' = a^2 k[1 + \cos^2 \theta + 2 \cos \theta R_h(T)]. \quad (6)$$

As can be seen in Fig. 1(b), d' depends on θ , the difference in directions of jump₁ and jump₂. At short delays between jumps, if the jump directions are nearly the same, the detectability is high. If the directions differ greatly, detectability is low.

At threshold, $d' = 1$, giving

$$a_{\text{thresh}} = \frac{1}{\sqrt{k[1 + \cos^2 \theta + 2 \cos \theta R_h(T)]}} \quad (7)$$

If the difference in jump directions θ is small, the threshold is low; if θ is large, the threshold is high (Fig. 2(b)).

According to this model, the threshold and d' loci show the directional tuning of a motion sensor that responds well to one of the two jumps but less well to the second since it is further away from the direction to which the sensor is tuned.

1.3. Vector model

One last possibility to consider is that the full vectors, not just the speeds or components of the speeds in the same direction, are used in the velocity energy

computation. This model is a vector generalization of Simpson's [1] original speed energy model. We show in the Appendix A that this model gives

$$d' = 2a^2 k[1 + \cos \theta R_h(T)]. \quad (8)$$

The d' locus is shown in Fig. 1(c). Detectability is best when the jumps are in similar directions.

At threshold $d' = 1$, so the threshold jump size is

$$a_{\text{thresh}} = \frac{1}{\sqrt{2k[1 + \cos \theta R_h(T)]}} \quad (9)$$

The threshold locus is shown in Fig. 2c. The threshold is low when the difference in jump directions θ is small and high when the difference in jump directions is large.

The vector model will give threshold and d' loci that give the appearance of directional tuning: pairs of jumps in similar directions will be more detectable than pairs in dissimilar directions. However this behaviour results from essentially perfect recovery of each vector and these vectors are combined essentially perfectly (at short delays).

2. Methods

Here is an overview of the experiments. We measured the temporal response and the directional tuning of the motion detection mechanism with the two-jump technique. A three-frame random dot motion sequence was used. Changing the frame duration varied the delay between the two jumps. The direction of the first jump was fixed, but the direction of the second jump was variable. The thresholds for pairs of jumps for several direction differences at a given delay form the threshold locus. The shape of the locus and how that shape changes with the delay between jumps tells us about the direction tuning and temporal response of motion detection.

We repeated this basic experiment with several variations. We either measured the threshold amplitude of a pair of jumps, or the detectability of a pair of fixed-size jumps. We either used a traditional random dot motion stimulus where all dots moved rigidly together, or we used a statistical motion stimulus. The statistical stimulus forced the observer to use only motion to discriminate between stationary and moving displays, since the flicker and lengths of the time-averaged dot trails for the displays were the same. Finally, we measured the detectability of the motion in displays where the first jump direction was 0° (rightwards), 45° (upwards to the right), or 90° (upwards).

2.1. Subjects

The authors (WS and AN) served as subjects. Both had corrected-to-normal vision. WS was experienced in motion psychophysics; AN was not.

2.2. Apparatus

The displays were presented on a Tektronix 608 oscilloscope with P15 phosphor. P15 has essentially no persistence—if we had used P31 phosphor a veiling luminance would have been necessary to mask the lingering phosphor trails [4]. The oscilloscope was controlled by a point-plotting memory buffer developed at the University of Alberta [5]. This buffer was programmed with an MSDOS computer. Viewing was monocular from a chin-and-forehead rest placed 100 cm from the face of the oscilloscope. A button box collected responses.

2.3. Stimuli

The display consisted of 990 dots which were randomly (uniform) distributed inside a $2.7^\circ \times 2.7^\circ$ square region. Each dot was 2.6 min in diameter. A fixation dot was centred over the top of the square, 59.4 min above the top border. The fixation mark was created by plotting a set of five dots, arranged in a pattern as on a die, twice (the second set superimposed on the first set). There were 1000 dots in total, including the fixation mark, per frame and the buffer displayed the dots at a rate of 1 dot/ μ s, therefore the refresh rate was 1000 Hz (minimum frame duration of 1 ms). The background, as measured in a dark room by a Minolta LS-100 photometer, had a luminance of 0.001 cd/m^{-2} , while the dots had a luminance of 75 cd/m^{-2} . The latter value was measured as follows. A matrix of 44×44 dots was plotted, with the centres of the dots separated by 0.9 min arc. The interdot separation was smaller than the nominal dot diameter, since the dots were Gaussian blobs and we wanted a uniform patch. Extra dots were plotted elsewhere to bring the total number of dots to 2000, yielding a refresh rate of 500 Hz. The luminance reading from the photometer was doubled, since the experiment refresh rate was 1000 Hz.

The motion display consisted of a pair of jumps, generated by a three frame sequence. The first jump was at the transition between frames 1 and 2 and the second jump was at the transition between frames 2 and 3. Frames 1 and 3 were each displayed for 17 ms and the middle frame had a duration of either 20 (17 for WS), 40, 80, 120, or 160 ms. Since a motion impulse (jump) occurs at frame transitions, this duration was also the delay between motion impulses.

There were two basic types of motion stimulus: statistical and nonstatistical. The statistical stimulus has been described by Simpson [1]. Each jump of the moving display contained 90% of the dots moving in one direction and 10% in the opposite direction, giving a nonzero space-averaged speed. In the stationary display, half the dots moved in one direction and half moved in the opposite direction, giving a space-aver-

aged speed of zero. The only difference between stationary and moving statistical displays was in the proportions of dots: 90/10 or 50/50. Of course the individual dots of the statistical stationary display are not stationary. But the space-average is zero or stationary. The subjective appearance of the stationary statistical display is of flicker or chaotic motion with no specifiable direction. Transparency was not seen (perhaps because the whole sequence was quite short).

Let us clarify the composition of the statistical stimuli with an example. Suppose the first jump was rightward (0°) and the second jump was upwards (90°). For the moving statistical display: in the first jump 891 dots moved right and 99 dots moved left; in the second jump the same 891 dots that earlier moved right, move up and the same 99 dots that earlier moved left, move down. So it was 891 right and 99 left followed by 891 up and 99 down. The stationary statistical stimulus in this example: 495 right and 495 left followed by 495 up and 495 down. Again, the dots' lifetimes persisted throughout.

The statistical stimulus forced the observer to use only motion to discriminate between stationary and moving displays, since the flicker and time-averaged dot trails for the displays were the same. The difference between the moving and stationary displays was in the space-averaged speed. Watamaniuk and Duchon [6] have found that humans do such space-averaging of velocity.

The nonstatistical motion stimulus was simple rigid motion of all the dots in the display. In the static comparison stimulus all the dots were motionless.

In most of the experiments, the first jump was rightward (0°); the second jump was any of several directions between 0 and 315° in 45° increments. The size of the second jump was exactly equal to the first if the second jump was at 0, 90, 180, or 270° relative to the first, but slightly different if the second jump was 45, 135, 225, or 315° relative to the first jump. Diagonal movement is a simultaneous jump in both the vertical and horizontal directions. The horizontal and vertical components of this jump, however, can only be integers due to the nature of the apparatus. The jump sizes used in the experiment were chosen so as to keep the quantization error small—less than 0.01 min.

Two control experiments were done to check that the direction of the first jump was not critical. In most of the experiments the direction of the first jump was rightward (0°). In one control experiment the direction of the first jump was at 45° (diagonally upwards to the right) and in another the first jump was at 90° (upwards).

In some experiments the jump amplitude was fixed. For nonstatistical stimuli, this fixed jump amplitude was 0.4 min for WS and 1.4 min for AN. The displacements were individually chosen for the observers so

that the obtained d' 's were not too small or too large. For the statistical stimuli, the jump size was 1.0 min for both observers.

The following experiments were done:

1. Statistical stimulus; first jump at 0° ; threshold jump size measured.
2. Statistical stimulus; first jump direction 0° , fixed jump size; d' measured.
3. Nonstatistical stimulus; first jump direction 0° ; fixed jump size; d' measured.
4. Statistical stimulus; first jump direction 45° ; fixed jump size; d' measured.
5. Statistical stimulus; first jump direction 90° ; fixed jump size; d' measured.

2.4. Procedure

On each trial either a moving or stationary display was presented. If a statistical stimulus was used, the space-averaged speed was either zero (stationary) or nonzero (moving). That is, for statistical stimuli the task is to detect the global motion. The subjects' task was to decide whether a moving or stationary stimulus had been presented (yes–no task) and this response was indicated by a button press. The computer gave a beep if the response was incorrect. The two sorts of stimuli were presented in random order with equal probability. After a trial was complete the display was blank except for the fixation mark. When the subject was ready for the next trial, he pressed the button and the trial started about half a second later.

Trials were run in blocks of 50. Half of the trials contained motion and the other half did not. The direction of jump₂ was constant throughout a block. Each angle for jump₂ was run once at a given delay (in random order) before going on to the next delay. The delays were ordered randomly. Once all delays and jump directions were run, this process was repeated. It is worth stressing that within a block of trials there were only two stimuli to be discriminated: moving or nonmoving. The delay, direction of first jump and direction of second jump were all fixed. The threshold was measured by obtaining d' for each of a number of displacements (jump amplitudes). Both jumps were the same amplitude; the threshold was the amplitude of each jump. A minimum of four blocks of trials were run, each at a displacement around the threshold region. A curve was fit to the (displacement, d') pairs. At threshold d' is 1, so setting d' to 1 and rearranging the best-fitting curve gave the threshold estimate and a 95% confidence interval. This is the classical approach to what is known as a calibration problem [7] (pp. 404–409). If the confidence interval was excessively large, more trials were run to improve the fit, up to a maximum in one case of 12 blocks.

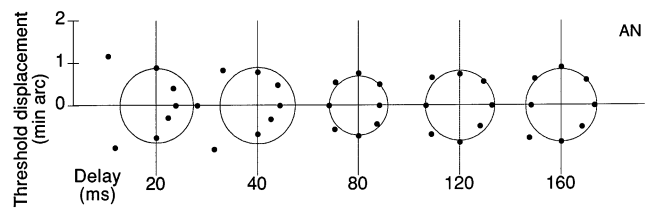


Fig. 3. The threshold displacement for pairs of jumps at each of several delays. Each plot is polar: r is the threshold displacement and θ is the difference in the directions of jump₁ and jump₂. The plotted points show the threshold size of each jump. For example, a threshold of 1 min arc means that jump₁ was 1 min arc and jump₂ was also 1 min arc. Each threshold is based on a minimum of 200 trials. The curves are the least squares fits of the amplitude model—Eq. (4). Subject AN.

For experiments where d' was measured for a fixed jump size, five blocks of 50 trials were run. So each point in the d' locus plots is based on 250 responses.

3. Results and discussion

Figs. 3–8 show the displacement thresholds for the detection of motion in statistical displays. The displacement thresholds are plotted as a function of the angular difference in the directions of the first and second jump. Each plot is polar, with the threshold represented as r and the difference in directions as θ . Plots for several delays between jumps are shown in each figure. The figures show the least squares fits of the amplitude model (Eq. (4)), the directional tuning model (Eq. (7)) and the vector model (Eq. (9)) to the data of each observer.

The amplitude model yields a circular threshold locus and the fit is plainly poor (Figs. 3 and 6). Note that all models predict that the locus becomes nearly circular at long delays between jumps, so the fit at short delays is most important for discriminating the models. The index we use to compare model fits is the sum of the squared errors (SSE). For WS the SSE is 11.745. The threshold for AN at the shortest delay and $\theta = 180^\circ$ was immeasurably high. The amplitude model gives a fitted value that is finite and so computing the error of this point is problematic (the error is infinite since error = infinite observed value – finite fitted value). Ig-

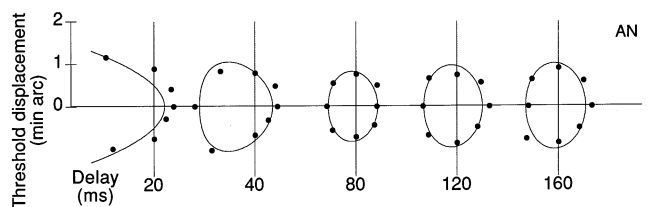


Fig. 4. The threshold displacement for pairs of jumps at each of several delays, as in Fig. 2. The curves are least squares fits of the directional tuning model—Eq. (7). Subject AN.

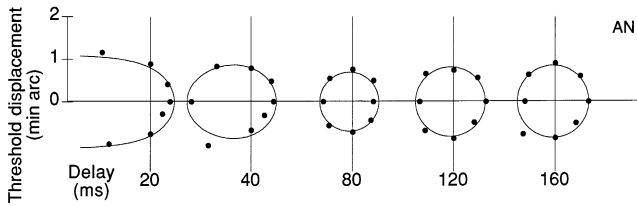


Fig. 5. The threshold displacement for pairs of jumps at each of several delays, as in Fig. 2. The curves are least squares fits of the vector model—Eq. (9). Subject AN.

noring the problematic point we get an SSE of 2.688, though in reality the SSE must be much higher.

The directional tuning model gives somewhat better fits (Figs. 4 and 7). Since the threshold for AN, 20 ms, $\theta = 180^\circ$ was immeasurable, the value of $R_h(T)$ was set to 1.0. At this value the fitted curve correctly does not give a finite value for the point. The SSE of the fit for AN is 0.800. The fit is better than for the amplitude model, though still fairly poor, especially at short delays. This is even more apparent in WS's data, which show a very poor fit at short delays. The SSE is 2.143.

The vector model gives very good fits (Figs. 5 and 8). The SSE for AN is 0.424 and for WS is 0.540. The vector model clearly fits the displacement threshold data best. The SSEs from worst to best are: amplitude model 2.688 (AN), 11.745 (WS); directional tuning model 0.800 (AN), 2.143 (WS); vector model 0.424 (AN), 0.540 (WS). Since the vector model was a clear winner here, in subsequent figures only the fit of the vector model will be given.

The vector model allows us to recover the autocorrelation function of the motion impulse response from these data. For each delay between jumps, $R_h(T)$ is the parameter which controls how egglike the curves are and this can be plotted as a function of the delay T . The resulting autocorrelation function of the impulse response for AN and WS are plotted in Fig. 9. These functions are fitted with a version of the log-logistic cumulative distribution function:

$$R_h(T) = \frac{1}{1 + (T/p)^s} \quad (10)$$

(This function is commonly used in survival analysis—cf Cox and Oakes, [8] (p. 21.) Parameter p is the delay

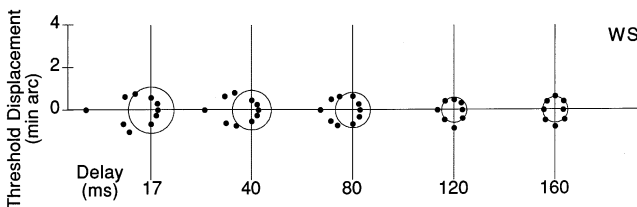


Fig. 6. The threshold displacement for pairs of jumps at each of several delays, as in Fig. 2. The curves are least squares fits of the amplitude model—Eq. (4). Subject WS.

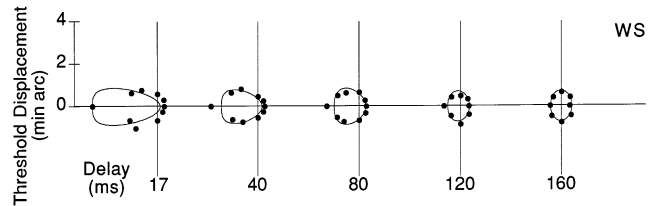


Fig. 7. The threshold displacement for pairs of jumps at each of several delays, as in Fig. 2. The curves are least squares fits of the directional tuning model—Eq. (7). Subject WS.

at which $R_h(T)$ falls to a value of 0.5 (the persistence) and s controls the slope. For AN, the 95% confidence interval for the persistence is 58 ± 17 ms; for WS it is 102 ± 12 . For AN, the 95% confidence interval for the slope is 3 ± 2 ; for WS it is 5 ± 2 . AN's persistence is about half the value obtained for WS. This means that he is able to see higher frequencies of oscillation. Where does this difference in motion function arise in the visual system and how? One possibility is that WS's greater persistence is due to his greater age (38 versus 21 years). The upper cut-off temporal frequency for detection of luminance modulation declines with age [9–11]. If the processing of luminance becomes more sluggish with age, this could have the effect of prolonging the persistence of visual motion.

The plot for WS in Fig. 9 compares the $R_h(T)$ function obtained here with that obtained by Simpson [1] (Fig. 8) with a quite different procedure. In the earlier experiment all the jumps were horizontal and a scalar speed energy model was used. Fitting the earlier data we find a persistence of 117 ± 22 ms and a slope of 1.7 ± 0.9 . So although the slopes of the estimated $R_h(T)$ functions from the two experiments are rather different, the persistence estimates are very similar.

A crude theory of the observer's behaviour (Figs. 3–8) would hold that at long delays (where the threshold locus is a circle) he uses only one of the two jumps and at short delays (where the threshold locus is an egg) he uses both jumps. The crude theory of course does not fit a curve to the data. The theory proposes that the observer uses one of two strategies. This theory can easily be tested against the data in Fig. 9. The theory predicts that $R_h(T)$ will be a step, with a high

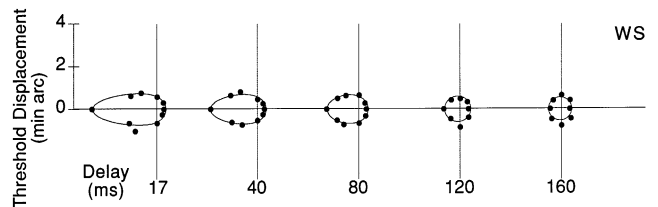


Fig. 8. The threshold displacement for pairs of jumps at each of several delays, as in Fig. 2. The curves are least squares fits of the vector model—Eq. (9). Subject WS.

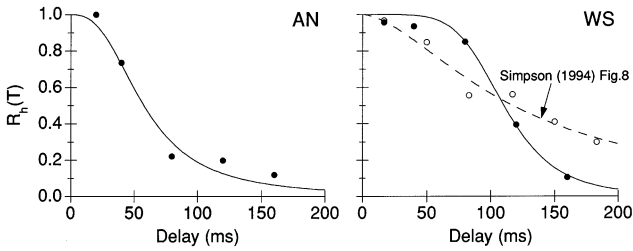


Fig. 9. The autocorrelation of the motion system's temporal impulse response, estimated by the fits of Eq. (9) to the data in Figs. 5 and 8. The curve is a least squares fit of a log-logistic cumulative distribution function (Eq. (10)). For subject WS, the $R_h(T)$ from Simpson [1], Fig. 8, is included for comparison.

value at short delays where both jumps are used and a low value at long delays where only one jump is used. The data decisively refute the crude theory, since the $R_h(T)$ functions smoothly decline with the delay (according to a log-logistic).

Now consider the sensitivity for detecting a pair of jumps of fixed displacement. The data from both observers are shown for the statistical and nonstatistical conditions in Figs. 10 and 11. The data are most similar to the family of curves in Fig. 2 corresponding to the vector model. As well, we found that the threshold loci measured in experiment 1 were best fit by the vector model. Therefore the curves fitted to the data in Figs. 10 and 11 are least squares fits of the vector model, Eq. (8).

The quality of fit is good for both the statistical and nonstatistical conditions. Hence we have generalized the original conclusion about a full vector velocity energy model to a traditional nonstatistical display [12–15]), to d' loci instead of threshold loci and to above threshold stimulus levels.

The purpose of the nonstatistical stimulus is to reduce the chance that performance on the motion detection task is based on cues other than motion. It is theoretically possible that the moving and stationary sequences in a nonstatistical display can be distin-

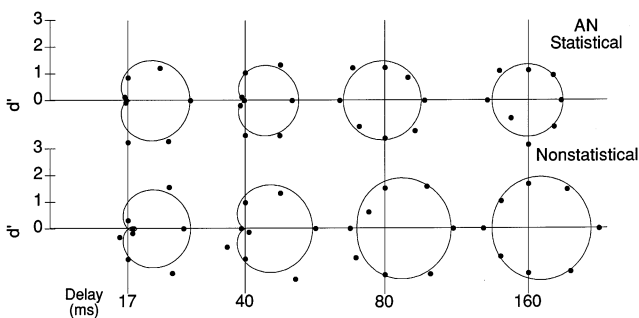


Fig. 10. Polar plot of d' (r) for pairs of jumps in different directions (θ), separated by a variable delay. The curves are least squares fits of the velocity energy model that extracts the jump vectors—Eq. (8). Data from statistical and nonstatistical displays are shown. Each point is based on 250 trials. Observer AN.

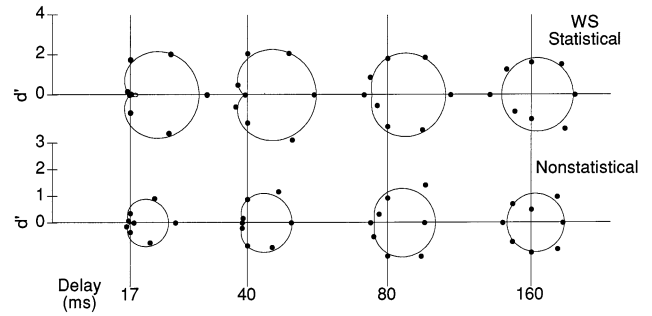


Fig. 11. d' for pairs of jumps in different directions, separated by a variable delay, as in Fig. 10. Note the different scales for statistical and nonstatistical displays. Observer WS.

guished on the basis of spatial or temporal cues alone. For example, a purely spatial discrimination does a time-average of the display. If a dot is moving, the result is a streak; if it is stationary, the dot stays a dot. A purely temporal discrimination uses the different patterns of flicker from moving and stationary displays. We have shown that the results with the statistical and nonstatistical displays are quite similar and so the nonmotion cues potentially available in nonstatistical motion displays do not seem to be used.

The fit of Eq. (8) gives an estimate of the autocorrelation function of the impulse response at each delay, $R_h(T)$. Fig. 12 shows plots of $R_h(T)$ for each observer and each type of display. The reader will note that for short delays the fitted value of $R_h(T)$ is greater than 1. This is problematic, because in the derivation of Eq. (8) we assumed that the energy of the impulse response and of the time-shifted impulse response were equal to 1:

$$\int_0^{\tau} [h(t)]^2 dt = \int_0^{\tau} [h(t - T)]^2 dt = 1$$

It follows immediately that the maximal value of $R_h(T)$, which occurs for a zero delay, will also be 1:

$$R_h(0) = \int_0^{\tau} h(t)h(t - 0) dt = \int_0^{\tau} [h(t)]^2 dt = 1$$

What then shall we make of the fitted values of $R_h(T)$ greater than 1? We earlier assumed that the energy of the impulse response and of the time-shifted impulse response were both equal to 1. Clearly if the

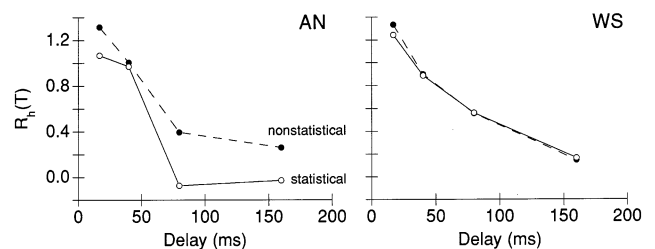


Fig. 12. Estimates of $R_h(T)$ as derived from fits of Eq. (8) to the data in Figs. 10 and 11.

integration of the motion filter's squared response extends to infinity the two integrals will be equal. However, if the integration duration is short, the integrals will not be exactly equal. Part of the time-shifted impulse response will fall outside the integration window. Therefore the computed energy of the second impulse response will be smaller than that of the first impulse response. In Appendix A we repeat the derivation, this time not assuming that the two integrals are equal and we find:

$$d' = a^2 k [1 + c + 2c \cos \theta R_h(T)]. \quad (11)$$

If the energy of the first impulse response really is equal to that of the second, $c = 1$ and Eq. (11) reduces to Eq. (8). But if the computed energy of the second impulse response is smaller due to its persistence past the integration duration, then unreasonably large values for $R_h(T)$ can result using Eq. (8). We attribute the apparently impossible values of $R_h(T)$ greater than 1 plotted in Fig. 12 to values of c in Eq. (11) greater than 1. Unfortunately we cannot simply fit Eq. (11) and find c directly since multiple solutions for k , c and $R_h(T)$ exist for any set of data.

Since the $R_h(T)$ function takes on values greater than 1 in Fig. 12, we do not fit a log-logistic to the data points—the log-logistic has an upper limit of 1. Instead the points are simply joined by line segments. Although we have to take the plots in Fig. 12 with a grain of salt, we can still make a crude assessment of the persistence. For AN the persistence (delay at which $R_h(T)$ falls to 0.5) is around 60–75 ms for the nonstatistical and statistical conditions, respectively. For WS the two functions are essentially identical, with a persistence of about 90 ms. If we compare these persistence values to those obtained in the first experiment (AN: 58 ± 17 ms; WS: 102 ± 12), we see that the agreement is good.

In the experiments described thus far, the first jump was rightward. Some researchers have reported inhomogeneities and anisometries in motion detection e.g. [15,16], so we ran the basic experiment twice more with different directions for the first jump: at 45° (upwards to right) and at 90° (upwards).

The d' loci are shown with least squares fits of Eq. (8) in Fig. 13. The loci look about the same as those obtained when the first jump is at 0° (to the right)—see Fig. 10. The $R_h(T)$ functions for the first jump oriented at 0, 45 and 90° are compared in Fig. 14. The functions for 45 and 90° first jumps are fitted with the log-logistic but the 0° function points are simply joined by line segments.

Inspection of Fig. 14 reveals what appears to be a systematic difference in the obtained $R_h(T)$ functions as the direction of the first jump is varied from 0 to 90° . As the jump becomes more vertical, the persistence becomes shorter and the slope becomes shallower. These differences are not large enough to reach statisti-

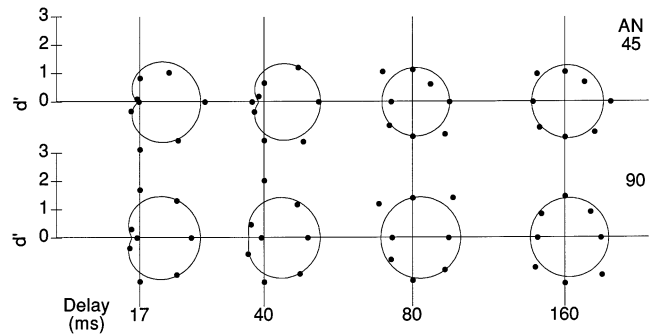


Fig. 13. d' for pairs of jumps in different directions, separated by a variable delay, as in Fig. 10. The direction of the first jump was 45° (upwards to right) in one experiment and 90° (vertical) in the other. Observer AN.

cal significance. The point estimates and 95% confidence intervals for the persistence are about 60 ms for the 0° condition, 52 ± 11 for the 45° condition and 40 ± 15 for the 90° condition. The slope estimates and confidence intervals are 5 ± 4 (45°) and 1.6 ± 0.9 (90°). In summary, the shape of the d' locus and how it varies with the delay do not depend on the direction of the first jump.

4. General discussion

These experiments were designed to show the directional and temporal tuning of the mechanisms underlying motion detection. Let us discuss these aspects in turn.

4.1. Directional tuning

We examined three models that differed in what information they extracted from the stimulus velocity waveform $v(t)$. The amplitude model extracts only the speed, the directional tuning model extracts the speed of the first jump and the speed of the component of the second jump in the same direction and the vector model uses the full vectors of both jumps.

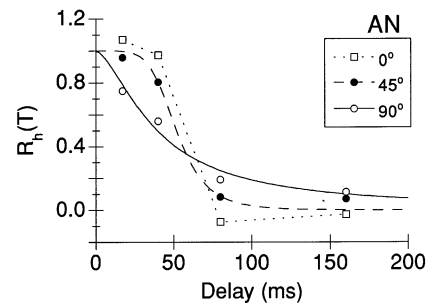


Fig. 14. Estimates of $R_h(T)$ as derived from fits of Eq. (8) to the data in Fig. 13 (45 and 90° first jump) and Fig. 10 (0° , statistical stimulus). Observer AN.

The amplitude model uses only the speed or displacement of the jumps. This corresponds to the idea that displacement detection does not involve motion sensing [17] Section 2.3 [14]. The amplitude model predicts that the direction of the jumps is irrelevant to detectability: the threshold or d' locus is a circle in polar coordinates, where θ is the difference in jump directions. The data speak decisively against this model. Especially at short delays between jumps, detectability is much higher when the jumps are in similar directions.

The directional tuning model uses the first jump's speed and the component of the second jump's speed in the same direction. We can relate this model to direction-tuned motion sensors whose output is proportional to the speed in the preferred direction. Single cells in MT [18–22] are tuned to motion direction. If the speed in the preferred direction is increased, in some MT cells the spike rate declines, in others it rises and in yet others the spike rate rises and falls again [19]. One can think of the directional tuning model as using cells whose spike rate increases with the speed in the preferred direction. A very simple-minded implementation of the model works as follows. The most active cell is the one tuned to the jump direction and the spike rate is proportional to the speed (displacement). The most active cell (or class of cells) is found for the first jump and this cell's output is integrated over the display duration (including the second jump). This cell, since it is tuned to direction, will give a smaller output to the second jump than to the first. (On the idea that a single cell or small group of cells might underlie the psychophysical performance of an observer see Salzman et al. [23]). The greater the separation in jump directions, the smaller the response to the second jump and so the less detectable the pair. Although the results are qualitatively in agreement with this directional tuning model, detailed examination shows that the fit is poor.

Finally, the vector model uses the vectors of both jumps. We can again try to relate the model to direction-tuned motion cells. Suppose that there are n classes of direction-tuned cells. Each jump is coded as $(r_1, r_2, r_3, \dots, r_n)$, the spike rate in each class of direction-tuned cell, just as colour is initially coded as (r, g, b) , the level of activity in the three classes of cone. This is a nonorthogonal way to code the speed and direction (or x and y components) of the jump vector. The two recovered $(r_1, r_2, r_3, \dots, r_n)$ vectors are added. This model fit both the threshold and d' loci better than did the other models by a large margin.

The best expressions of the models are given by Eq. (3), Eq. (4), Eqs. (6)–(9). The interpretations just given in terms of direction-tuned motion cells are not meant to be taken too seriously. We still know very little about brain functioning and exactly how the brain implements the model equations is unknown.

Some readers may be surprised that the vector model fits the data better than does the directional tuning model. The threshold and d' loci may seem to reveal directional tuning, but we have argued that they do not show such tuning at all. We do not mean that the directions of the vectors are completely irrelevant (that would be the amplitude model, which we have decisively discredited). We mean that the full speed and direction of each vector is used in detecting the pair. The directional tuning model, in contrast, asserts that some of the second vector's speed is lost simply because it is not in the sensor's preferred direction. This idea is wrong. The apparent tuning of the detection data is really due to near perfect vector summation of the two vectors (at short delays between jumps), each of which is recovered almost perfectly.

The vector model is a generalization of the speed energy model of Simpson [1]; since it uses the full velocity vectors instead of speed we can call it a velocity energy model. The model asserts that the length of the vector sum of the two temporally filtered jumps is found, squared and integrated (Eq. (A9)). The full information from each vector is used (ignoring the temporal filtering) and the information about the size of the jumps is perfectly combined in a vector sum. The speed energy model for two jumps with speeds a and b is

$$E = a^2 + b^2 + 2ab R_h(T). \quad (12)$$

The velocity energy model for jump vectors \mathbf{a} and \mathbf{b} is

$$E = a^2 + b^2 + 2ab \cos \theta R_h(T). \quad (13)$$

The velocity energy model, besides fitting the present data admirably, also fits the original data of Simpson [1]. In the speed energy model used for the original data the size of each jump, a and b , could be positive or negative, indicating leftward or rightward movement. In the velocity energy model a and b are always positive, the sign being carried by the $\cos \theta$ term. For horizontal motion, then, Eq. (12) and Eq. (13) are equivalent (bearing in mind the remarks on motion sign). The vector generalization of the original speed energy model can successfully deal with pairs of jumps in arbitrary directions and at the same time fits the original data.

The other recent psychophysical studies we know about that have measured directional tuning are Raymond [24] and Zanker and Hüggen [25]. The motion stimulus used by Zanker and Hüggen was quite different ('non-Fourier' motion; [26]) from that used here. Raymond used random dots. Both studies derived direction tuning curves. Raymond's had a full width at half maximum of around 80° and Zanker and Hüggen's curves had a FWHM of around 90° . We have argued that our own results show that motion detection is not directionally tuned at all. By motion detection we

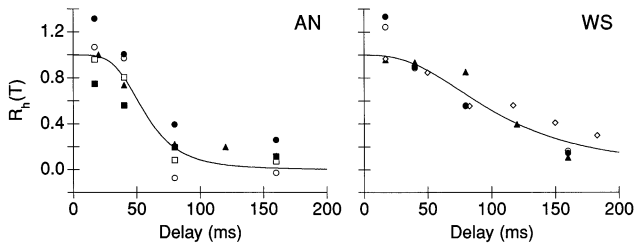


Fig. 15. Summary of $R_h(T)$ functions obtained from several experiments. Filled triangle, experiment 1; empty circle, experiment 2, statistical; filled circle, experiment 2, nonstatistical; empty square, experiment 4, 45° jump; filled square, experiment 5, 90° jump; empty diamond, data from Simpson [1], Fig. 8.

mean discrimination of moving from stationary displays. By saying that motion detection is not directionally tuned, we mean that no matter what the directions of the two jumps, they are summed perfectly. There are directionally tuned mechanisms that extract the direction of each jump and these mechanisms feed the motion detection stage, which itself is not directionally tuned. We will not argue the point here, but we believe that a similar vector summation idea can be applied to Raymond's and Zanker and Hüggen's results.

4.2. Temporal response

Besides generalizing the speed energy model so it can handle vectors instead of scalar speeds, we also generalized the method of measuring the motion system's temporal response. We used pairs of jumps in different directions at various delays. Using a velocity energy model we were able to estimate the autocorrelation function of the temporal impulse response $R_h(T)$. The temporal response so obtained agrees with that originally measured by Simpson [1], who used horizontal jumps. So despite the large differences in the stimuli—any direction versus horizontal only; both jumps same size versus jumps of varying sizes—the measured $R_h(T)$ functions are quite similar, especially in their persistence.

Fig. 15 summarizes the $R_h(T)$ data from all the experiments. The plot for WS also includes the data obtained by [1] for this same subject. Although the different conditions give somewhat different $R_h(T)$ functions, there is consistency in the shape and in the persistence. A log-logistic curve was fitted to all the points. The best fitting persistence with 95% confidence interval is 56 ± 10 ms for AN and 103 ± 16 ms for WS. The points are consistently well described by the log-logistic. The corresponding temporal impulse response is lowpass. Therefore we confirm the common conclusion that the motion system is not very sensitive to visual acceleration [27–30,1,31,32].

Acknowledgements

This work was supported by a research grant from the Natural Science and Engineering Research Council to WS. We thank Walter Bischof and Andrew Derrington for their helpful comments on the manuscript and Uffe Thygesen for his knowledgeable advice on linear systems.

Appendix A. Derivation of model predictions

Amplitude model

In this model the amplitude or speed of each jump is filtered, squared and integrated:

$$E = \int_0^{\tau} [v(t) * h(t)]^2 dt \quad (\text{A1})$$

where E is the speed energy, τ is the duration over which the energy is integrated, $v(t)$ is the speed waveform, $h(t)$ is the motion filter and $*$ denotes convolution. Since a jump is a Dirac delta function, the two-jump stimulus is

$$v(t) = a\delta(t) + b\delta(t - T) \quad (\text{A2})$$

where a and b represent the size (amplitude, speed, or displacement) of the first and second jumps and T is the delay between them. Substituting into Eq. (A1),

$$\begin{aligned} E &= \int_0^{\tau} \{[a\delta(t) + b\delta(t - T)] * h(t)\}^2 dt \\ &= \int_0^{\tau} [ah(t) + bh(t - T)]^2 dt \\ &= a^2 \int_0^{\tau} [h(t)]^2 dt + b^2 \int_0^{\tau} [h(t - T)]^2 dt \\ &\quad + 2ab \int_0^{\tau} h(t)h(t - T) dt. \end{aligned} \quad (\text{A3})$$

Let

$$\int_0^{\tau} [h(t)]^2 dt = \int_0^{\tau} [h(t - T)]^2 dt = 1$$

The first integral is equal to the second since both are areas under time-shifted versions of the same function. Each term represents the energy of the impulse response, a constant which we set equal to 1. By setting the energy of the impulse response equal to 1 we ensure that filtering preserves the energy of the original signal, which is quite a natural property for a filter. This gives us

$$E = a^2 = b^2 + 2ab \int_0^{\tau} h(t)h(t - T) dt. \quad (\text{A4})$$

Since $\int_0^\tau h(t)h(t-T) dt$ is the autocorrelation function of the impulse response, let us represent it by $R_h(T)$, yielding the velocity energy

$$E = a^2 + b^2 + 2ab R_h(T). \tag{A5}$$

We measure the detectability d' . If we suppose that performance is limited by internal normal noise with mean n and standard deviation σ which is added to the stimulus velocity energy, the pair of random variables to be discriminated in a detection task is $N(n, \sigma^2)$ and $N(n + E, \sigma^2)$. d' is the difference of the means of the two random variables divided by their standard deviation:

$$d' = \frac{(n + E) - n}{\sigma} = \frac{E}{\sigma} \tag{A6}$$

Substituting the right-hand side of Eq. (A5) for E we obtain the prediction

$$d' = k[a^2 + b^2 + 2ab R_h(T)], \tag{A7}$$

where k is $1/\sigma$. Eq. (A7) is the predicted detectability for a pair of jumps for the amplitude model. Since the two jumps are the same size we represent both by a

$$d' = k[2a^2 + 2a^2 R_h(T)] = 2a^2k[1 + R_h(T)]. \tag{A8}$$

At threshold $d' = 1$, so the threshold size of a is

$$a_{\text{thresh}} = \frac{\sqrt{2}}{\sqrt{k + [1 + R_h(T)]}}. \tag{A9}$$

Vector model

This model is a vector generalization of Eq. (A1):

$$E = \int_0^\tau [\mathbf{v}(t) * h(t)]^2 dt \tag{A10}$$

where $\mathbf{v}(t)$ is the pair of jumps

$$\mathbf{v}(t) = \mathbf{a}\delta(t) + \mathbf{b}\delta(t - T). \tag{A11}$$

Substituting,

$$\begin{aligned} E &= \int_0^\tau \{[\mathbf{a}\delta(t) + \mathbf{b}\delta(t - T)] * h(t)\}^2 dt \\ &= \int_0^\tau [\mathbf{a}h(t) + \mathbf{b}h(t - T)]^2 dt \\ &= \mathbf{a}^2 \int_0^\tau [h(t)]^2 dt + \mathbf{b}^2 \int_0^\tau [h(t - T)]^2 dt \\ &\quad + 2\mathbf{a}\mathbf{b} \int_0^\tau h(t)h(t - T) dt. \end{aligned} \tag{A12}$$

Setting the energy of the impulse response equal to 1,

$$E = \mathbf{a}^2 + \mathbf{b}^2 + 2\mathbf{a}\mathbf{b} R_h(T). \tag{A13}$$

Taking \mathbf{a}^2 , \mathbf{b}^2 and $\mathbf{a}\mathbf{b}$ as dot products (cf Jeffrey, [33] p. 128) and letting \mathbf{a} and \mathbf{b} denote the moduli (speed, displacement) of \mathbf{a} and \mathbf{b} ,

$$E = a^2 + b^2 + 2ab\cos\theta R_h(T). \tag{A14}$$

d' is proportional to the velocity energy

$$d' = k[a^2 + b^2 + 2ab\cos\theta R_h(T)]. \tag{A15}$$

Since the vector lengths a and b are equal in the experiment, represent both by a :

$$d' = k[2a^2 + 2a^2\cos\theta R_h(T)] = 2a^2k[1 + \cos\theta R_h(T)]. \tag{A16}$$

At threshold $d' = 1$, so the threshold jump size is

$$a_{\text{thresh}} = \frac{1}{\sqrt{2k[1 + \cos\theta R_h(T)]}} \tag{A17}$$

Vector model with unequal energy for first and second impulse response

As we have seen, the vector model gives the energy

$$\begin{aligned} E &= \mathbf{a}^2 \int_0^\tau [h(t)]^2 dt + \mathbf{b}^2 \int_0^\tau [h(t - T)]^2 dt \\ &\quad + 2\mathbf{a}\mathbf{b} \int_0^\tau h(t)h(t - T) dt. \end{aligned} \tag{A18}$$

Taking \mathbf{a}^2 , \mathbf{b}^2 and $\mathbf{a}\mathbf{b}$ as dot products,

$$\begin{aligned} E &= a^2 \int_0^\tau [h(t)]^2 dt + b^2 \int_0^\tau [h(t - T)]^2 dt \\ &\quad + 2ab\cos\theta \int_0^\tau h(t)h(t - T) dt. \end{aligned} \tag{A19}$$

Now we set $\int_0^\tau [h(t)]^2 dt$ to 1, as before, but set $\int_0^\tau [h(t - T)]^2 dt$ to c . As before, $\int_0^\tau [h(t)h(t - T)]^2 dt$ is $R_h(T)$, giving

$$E = a^2 + b^2c + 2ab\cos\theta R_h(T). \tag{A20}$$

The two jumps are the same size, so represent them by the same symbol a

$$E = a^2 + a^2c + 2a^2\cos\theta R_h(T). \tag{A21}$$

d' is proportional to E ,

$$d' = a^2k[1 + c + 2\cos\theta R_h(T)]. \tag{A22}$$

References

- [1] Simpson WA. Temporal summation of visual motion. *Vis Res* 1994;34:2547–59.
- [2] Rashbass C. The visibility of transient changes of luminance. *J Physiol* 1970;210:165–86.
- [3] Green DM, Swets JA. Signal detection theory and psychophysics. CA: Peninsula, Los Altos, 1988.
- [4] Groner R, Groner MT, Müller P, Bischof WF, Di Lollo V. On the confounding effects of phosphor persistence in oscilloscopic displays. *Vis Res* 1993;33:913–7.

- [5] Finley G. A high-speed point plotter for vision research. *Vis Res* 1985;25:1993–7.
- [6] Watamaniuk SNJ, Duchon A. The human visual system averages speed information. *Vis Res* 1992;32:931–41.
- [7] Montgomery DC, Peck EA. Introduction to linear regression analysis. Toronto: Wiley, 1992.
- [8] Cox DR, Oakes D. Analysis of survival data. NY: Chapman and Hall, 1984.
- [9] Kuyk TK, Wesson MD. Aging-related foveal flicker sensitivity losses in normal observers. *Optom Vis Sci* 1991;68:786–9.
- [10] Lachenmayr BJ, Kojetinsky S, Ostermeier N, Angstwurm K, Vivell PMO, Schaumberger M. The different effects of aging on normal sensitivity in flicker and light-sense perimetry. *Invest Ophthalmol Vis Sci* 1994;35:2741–8.
- [11] Tyler CW. Two processes control variations in flicker sensitivity over the lifespan. *J Opt Soc Am A* 1989;6:481–90.
- [12] Freeman T C A, Harris M G. Human sensitivity to expanding and rotating motion: effects of complementary masking and directional structure. *Vis Res* 1992;32:81–7.
- [13] Legge GE, Campbell FW. Displacement detection in human vision. *Vis Res* 1981;21:205–13.
- [14] Scobey RP, Johnson CA. Displacement thresholds for unidirectional and oscillatory movement. *Vis Res* 1981;21:1297–302.
- [15] Scobey RP, van Kan PLE. A horizontal stripe of displacement sensitivity in the human visual field. *Vis Res* 1991;31:99–109.
- [16] van de Grind WA, Koenderink JJ, van Doorn AJ, Milders MV, Voerman H. Inhomogeneities and anisotropies in the monocular visual field of human observers. *Vis Res* 1993;33:1089–107.
- [17] McKee SP, Watamaniuk SNJ. The psychophysics of motion perception. In: Smith AT, Snowden RJ, editors. *Visual Detection of Motion*. New York: Academic Press, 1994:85–114.
- [18] Albright TD. Direction and orientation selectivity of neurons in visual area MT of the macaque. *J Neurophysiol* 1984;52:1107–30.
- [19] Felleman DJ, Kaas JH. Receptive-field properties of neurons in middle temporal visual area (MT) of owl monkeys. *J Neurophysiol* 1984;52:487–513.
- [20] Lagae L, Raiguel S, Orban GA. Speed and direction selectivity of macaque middle temporal neurons. *J Neurophysiol* 1993;69:19–39.
- [21] Maunsell JHR, van Essen DC. Functional properties of neurons in middle temporal area of the macaque monkey. I. Selectivity for stimulus direction, speed and orientation. *J Neurophysiol* 1983;49:1127–47.
- [22] Mikami A, Newsome WT, Wurtz RH. Motion selectivity in macaque visual cortex. I. Mechanisms of direction and speed selectivity in extrastriate area MT. *J Neurophysiol* 1986;55:1308–27.
- [23] Salzman CD, Britten KH, Newsome WT. Cortical microstimulation influences perceptual judgements of motion direction. *Nature* 1990;346:174–7.
- [24] Raymond JE. Movement direction analysers: independence and bandwidth. *Vis Res* 1993;33:767–75.
- [25] Zanker JM, Hüggen IS. Interaction between primary and secondary mechanisms in human motion perception. *Vis Res* 1994;34:1255–66.
- [26] Chubb C, Sperling G. Drift-balanced random stimuli: a general basis for studying non-Fourier motion perception. *J Opt Soc Am A* 1988;5:1986–2006.
- [27] Gottsdanker RM, Frick JW, Lockard RB. Identifying the acceleration of visual targets. *Br J Psychol* 1961;52:31–42.
- [28] Glünder H. Correlative velocity estimation: visual motion analysis, independent of object form, in arrays of velocity-tuned bilocal detectors. *J Opt Soc Am A* 1990;7:255–63.
- [29] Nakayama K. Biological image motion processing: a review. *Vis Res* 1985;25:625–60.
- [30] Regan DM, Kaufman L, Lincoln J. Motion in depth and visual acceleration. In: Boff KR, Kaufman L, Thomas JP, editors. *Handbook of Perception and Human Performance*, Vol. 1. Toronto: Wiley, 1986:19-1–19-46.
- [31] Snippe HP, Werkoven P. Pulse modulation detection in human motion vision. *Vis Res* 1993;33:647–56.
- [32] Werkoven P, Snippe HP, Toet A. Visual processing of optic acceleration. *Vis Res* 1992;32:2313–29.
- [33] Jeffrey A. Mathematics for engineers and scientists. UK: Van Nostrand Reinhold, Wokingham, 1979.