DISSERTATIONES PSYCHOLOGICAE UNIVERSITATIS TARTUENSIS

3

MEASUREMENT OF ELEMENTARY MOVEMENT VECTORS IN HUMAN VISUAL SYSTEM

ALEKSANDER PULVER

TARTU 1995

DISSERTATIONES PSYCHOLOGICAE UNIVERSITATIS TARTUENSIS

3

DISSERTATIONES PSYCHOLOGICAE UNIVERSITATIS TARTUENSIS
3

MEASUREMENT OF ELEMENTARY MOVEMENT VECTORS IN HUMAN VISUAL SYSTEM

ALEKSANDER PULVER

TARTU UNIVERSITY
PRESS

Department of Psychology, University of Tartu, Tartu, Estonia

Dissertation is accepted for commencement of the degree of Doctor of Philosophy (in Psychology) on April 4, 1995 by the Doctoral Committee of the Department of Psychology, University of Tartu

Opponents:	Talis Bachmann, Ph.D. (Tallinn)
	Piet Vos, Ph.D. (Nijmegen)

Commencement: May 11, 1995

Publication of this dissertation is granted by the Estonian Science Foundation

Tartu Ülikooli Kirjastuse trükikoda
Tiigi 78, EE2400 Tartu
Tellimus nr. 90.

CONTENTS

LIST OF ORIGINAL PUBLICATIONS

This dissertation is based on the following original publications which will be referred to in the text by their respective Roman numerals.

I Allik, J., Tepp, M., Livshits [Pulver], A. Detection of temporal phase by directionally sensitive units in the human visual system. *Acta et Commentationes Universitatis Tartuensis*, 1977, 429, 107–120.

II Pulver, A., Allik, J. Motion direction identification in linear random cinematograms. *Acta et Commentationes Universitatis Tartuensis*, 1986, 753, 18–35 (*in Russian*)

III Allik, J., Pulver, A. Timing of visual events for motion discrimination. *Vision Research*, 1994, 34, 1585–1594.

IV Allik, J., Pulver, A. Magnitude of luminance modulation specifies amplitude of perceived movement. *Perception and Psychophysics*, 1995, 57, 27–34.

V Allik, J., Pulver, A. Contrast response of a movement-encoding system. *Journal of the Optical Society of America*, 1995, 12A (6), (*in press*).

VI Pulver, A., Allik, J. When a sequence of two opposite polarity flashes is perceived to move in the reversed direction. *Vision Research*, 1995 (*submitted*).

HISTORICAL BACKGROUND

The perception was not considered as a problem deserving its own explanation until the second half of the nineteenth century. It was a common belief that all perceptual phenomena concerning perception of motion can be explained on the basis of other more fundamental perceptual processes, particularly the persecution of space and time. As in physics, perceived motion was believed to be a derivative of space and time. Two motion phenomena — motion after-effect and stroboscopic motion — posed some difficulties to this concept but appeared to be explained by some particular mechanism without doubting in the whole concept itself. In particular, the movement after-effect was typically explained by eye movements and 'filling in' blank intervals between two stroboscopic flashes by the prolonged after-image.[1]

A turning point in the study of the motion perception is marked by studies of Sigmund Exner. He claimed, contrary to the belief of the scientific establishment, that motion is neither logically nor even psychologically deduced from more basic and elementary sensations of space and time (Exner, 1875, 1876, 1888). One of his key arguments was founded on his discovery that the human observer is able to discriminate amazingly short time intervals of about 10 msec or even less between two asynchronous electric sparks elicited at two neighboring locations in the visual field. The observer was able to see motion between these two nearby flashes even though they could not be spatially or temporally resolved. For example, these two flashes exposed simultaneously were perceived as a single object. Two asynchronous flashes separated considerably from each other in space were perceived as two separate but simultaneous objects. This means that motion is a sensation (*Bewegungsempfindung*) like sensation of color, contrast and location and not necessarily higher-order perceptual phenomena (*Bewegungswahrnehmung*). This difference is the same as in case of seeing the movement of second hand and minute hand of watch.[2]

Next important event in the scientific explanation of motion perception is connected with Max Wertheimer (1912). It is a well-known episode that in 1910 on his vacation trip, Wertheimer unexpectedly decided to leave the train in Frankfurt where he purchased a simple toy stroboscope in store. He realized that when two objects are presented in a succession at some distance from each other the observer could perceive the first presented object traveling through the space towards the object presented later in time. In some cases 'pure motion' without moving the object itself was perceived (it was named phi-

---

[1] "...Wenn eine Stelle der Netzhaut von periodisch veränderlichem und regelmäßig in derselben Weise wiederkehredem Lichte getroffen wird ... so entsteht ein kontinuierlicher Eindruck ..." (von Helmholtz, 1911/1860, s. 174)

[2] "Im zweiten Falle haben wir die Bewegungen dadurch erkannt, daß wir den Zeiger zu verschiedenen Zeiten an verschiedenen Stellen fanden. Im ersten Falle haben wir einen unmittelbaren Eindruck von der Bewegung selbst" (Exner, 1888, s. 438).

phenomenon or *f-Phänomen*). Opposite to the Exner position, Wertheimer claimed that that type of perceived motion had no reference in physical reality and therefore could not be sensation as such. The perceived motion is an invention of the human observer and exists only in his mind. As a result, the language for the description of motion perception was believed to proceed successfully only in subjective terms such as subjective speed, apparent movement, good movement etc. For this very reason, the motion perception became a paradigm description of the whole school, Gestalt psychology (Koffka, 1935). The shift towards phenomenological language of description ceased attempts to understand and describe basic mechanisms underlying the motion perception.

The sixties mark a remakable change in the understanding of functional architecture of the visual system. Hubel and Wiesel (1959) discovered the so-called 'orientation detectors' in the cat striate cortex. At the same time, Lettvin et al. (1959) discovered motion detecting cells in the frog's eye. These new exciting discoveries seemed to cover the existing gap between neuronal activity and subjectively experienced sensation. Barlow (1972) formulated a neurone doctrine for perceptual psychology, claiming that the perception corresponds to the activity of only a small number of active neurons, each of which are coding a particular aspect of visual stimulation. Psychophysicists did not wait long with responding to these developments in their area. Sekuler and Ganz (1963) exposed to the observer unidirectionally moving high-contrast gratings for a considerably long period of time. After this adaptation contrast thresholds elevated only for gratings moving in the same direction. This selective adaptation indicates indeed that at least movements in opposite directions are coded by two separate populations of neurones. **Study I** demonstrates that the direction specific adaptation can be revealed not only by measuring elevation of contrast thresholds but kinematic ones as well.

Beetles (*Chlorophanus*) play an important role in the explanation of motion perception. Werner Reichardt invented an elegant way how to study motion perception of this simple animal. Many insects react to optical stimulation by moving the eyes, head, or even the whole body. These reactions are called 'optomotor responses' because they are elicited by visual stimulation. Reichardt (1957) puts *Chlorophanus* inside a hollow cilinder which is composed of perpendicular black and white stripes. When the cylinder rotates, the animal tries to follow the movement. For the insect, this response to movement reduces the relative speed of the surrounding "world". In order to measure these responses, the beetle's back is glued to a piece of cardboard which is fixed to a static stand. Then the Y-maze globe was given to the beetle suspending freely in air. The Y-maze globe consisted of six pieces of curved straws that join at four points to form Y-like junctions. When the beetle started to walk it remained fixed, but Y-maze held by the beetle's legs performs the negative of movement the beetle would perform if it were walking freely. The number of turns with and against the direction of cilinder motion characterizes the magnitude of the optomotoric reaction. On the basis of this ingenious experiment, Reichardt (1957) was able to formulate a simple correlational model which quantitatively predicts the beetle's optomotoric reactions to the cilinder motion.

The Reichardt model was soon generalized to other experimental situations and other animals including man (van Santen, Sperling, 1984). Mainly due to the efforts of Reichardt and his followers motion perception appears to be changed into a unique area of perceptual studies where instead of a large variety of equally plausible models only a few basic explanations exist. The remarkable convergence was primarily achieved by theoretical works demonstrating that many different explanations are in fact variants of the same basic scheme of explanation.

The history of scientific research of the motion perception has lasted for nearly 120 years. The starting point was the lecture about motion sensation held by Sigmund Exner in *Kaiserliche Akademie der Wissenschaften* at July 15, 1875. A lot of empirical and theoretical studies accumulated over this period, however, leaves no doubt that motion is a fundamental visual dimension. The current *Zeitgeist* demands the construction of explanation for motion perception in terms of sophisticated formal models. Nevertheless the purpose of scientific research in this field has remained the same — description of basic principles which lie behind the process of motion perception.

EXTRACTION OF ELEMENTARY MOVEMENT VECTORS

The movement of an object is not immediately given to the perceiving organism. All motion parameters, including the direction of motion, must be inferred from a time-varying light flux which is caused by the translation of a luminous pattern across the surface of receptors. Logically, the measurement of movement can be divided into two stages: first, the extraction of local motion signals and labeling these signals to local points in a visual image, indicating at what speed and in what direction this point moves; and second, the use of these local motion signals for various perceptual decisions, for example for decision about the three-dimensional layout of the environment (e.g. Hildreth, 1984; Hildreth, Koch, 1987). Although both of these stages have been studied rather intensively during last few decades, the measurement of elementary movement vectors has been neglected to some extent.

Before motion can be used for recognition of objects or the structure of three- dimensional objects, the visual system must first reliably measure motion signals in time-varying luminance flux. What are the primitives used to detect and measure motion? What is metrics for the motion perception? What type of schemes have been used for the initial measurement of motion? What are the most general properties of these motion measurements?

According to the layman's concept of motion, movement is an intrinsic property of an object. Ordinary people and many scientists among them are convinced that a necessary condition for movement is the successive passing of adjacent loci. This concept was seriously shaken by Wertheimer's (1912) famous paper on f-movement demonstrating that there is no need for continuous displacement to perceive motion. For the physicist, however, motion has always been a quality attributed to an object by an observer: the object can be decided to be in motion only if it is observed at two different instants and it is seen to be in two different positions at those two instants. Therefore, f-motion may simply indicate that the movement experience requires a perceptible change in the position of some stimulus element with respect to another. However, the displacement of some stimulus elements with respect to others cannot be regarded as a **necessary** condition for perception of movement. A distinct impression of movement can be elicited by patterns containing no spatially displaced elements. The perceived movement can be evoked by changes of light flux at different retinal locations. Johansson (1950, 1978) described "wandering motion" seen between two or more spatially adjacent bright objects modulated in brightness (w-motion). What is particular to this and other analogous visual demonstrations (Anstis, 1967; 1986, 1990; Bülthoff and Götz, 1979; Gregory and Heard, 1983; Mather, 1984; Mastebroek and Zaagman, 1988) is that the perceived movement is generated by stimuli in which elements do not change their relative spatial position and usually remain continuously visible. These findings are surprising only if the detection of motion is ultimately regarded as a matching process comparing two spatial luminance patterns at two instants in time. The results of **Study IV** demonstrate that the spatial displacement of individual elements is not a necessary condition for motion perception. A distinct movement impression can be elicited by a relative change in the light flux at two spatial locations. It was demonstrated that alternating modulation of luminance of two adjacent dots produced perceived oscillatory w-motion which could not be perceptually distinguished from that produced by real oscillation. Thus, motion encoding system does not need to establish correspondence between similar individual spatial features in a motion sequence. The correspondence problem can be simply ignored by measuring the asymmetry in the change of the luminance flux at two sampled locations.

According to traditional viewpoint the strongest motion, real or apparent, occurs over short interelement distances (Ullman, 1979; Burt and Sperling, 1981; Shechter, Hochstein and Hillman, 1988; Shechter, Hochstein, 1989; Werkhoven, Snippe and Koenderink, 1990; Werkhoven and Snippe, 1990; Miller and Shepard, 1993). It is more natural, however, to assume that larger displacements convey more information for the presence of an object motion than smaller displacements which are, for example, difficult to separate from displacements caused by involuntary eye movements. The results of **Study IV** show that it is easier to elicit motion between two elements with larger spatial separation than those with smaller spatial separation. Many other psychophysical data, including kinematic thresholds and the detection of motion onset or instantaneous displacement, also require for their proper explanation an assumption that the motion weighting function increases with the displacement magnitude (Dzhafarov, 1992; Dzhafarov, Sekuler and Allik, 1993).

The most general property of any motion measurement system is that underlying operation must be nonlinear. As was convincingly proved by Poggio and Reichardt (1973), linear operation cannot extract the direction of a moving stimulus. Thus, the measurement of motion in human, and any other perceiving organism, must be based on the comparison of information extracted from at least two disparate spatial locations and the comparison process itself must have multiplication-like properties. Indeed, the change in luminance flux at **only one** location cannot reveal the presence of motion in the visual scene. The minimal requirement is the identical or at least analogous change in some other location at some preceding or delayed moment of time. In other words, only a conjoint change in luminance at two locations, A **and** B, can reveal the presence of motion.

All motion detection schemes fall broadly into two big classes: (1) correlation and (2) gradient schemes.

Correlation schemes. A moving object produces typically almost identical but delayed luminance variations at two neighboring sites along the motion path. Thus, the presence of motion in a particular direction can be recovered by measuring similarity between two local luminance fluxes along certain trajectory. Based on this kind of observations, Reichardt (1957) proposed an attractively simple delay-and-multiply scheme for the detection of motion in the visual system of beetle, *Chlorophanus*. According to this scheme, the basic operation for the motion detection is the multiplication of a signal from one spatial location with a delayed signal from another adjacent spatial location. Reichardt assumed that the detection of motion is solved by a big number of bilocal elementary motion encoders each of which is composed of two mirror-symmetrical component subunits tuned to motion in opposite directions. These subunits share two input channels that sample the visual field at two adjacent point-shape areas in space. The delay operation can be implemented by a linear low-pass temporal filter. Each subunit detects motion by delaying the temporal luminance pattern in one input channel and multiplying it with non-delayed pattern in the other input channel. The response of one subunit is algebraically subtracted from that of the complementary unit, the sign of the subtracted signal determining the perceived direction of motion. The subtraction is followed by infinite time averaging. Because of a big number of parallel operating elementary motion encoders, it is necessary to have a rule how outputs of all these encoders are combined into the final decision about motion. Typically, either of the two most simple rules of combination, the sum or maximum of all encoders outputs, are assumed not to alter the basic properties of an elementary bilocal motion encoder.

Gradient schemes. Gradient schemes rely on the relation between the spatial and temporal gradients of image intensity. The method was originally proposed by Limb and Murphy (1975; see also Fennema, Thompson, 1976). In the case of one-dimensional movement of an intensity profile $L(x, t)$ over small displacement dx in time dt, the ratio of temporal and spatial derivatives of image intensity is equivalent to the image velocity. Gradient schemes suffer from several disadvantages, particularly, they require computation of derivatives of intensity values what are contaminated with the sensory noise. Another problem is the inability to discriminate relevant and irrelevant changes in luminance. Most of these problems, however, were removed by the gradient scheme proposed by Marr and Ullman (1981). This model operates on not all locations of the luminance profile but on these where the light intensity changes significantly. Marr and Hildreth (1980) demonstrated that zero-crossings, that is locations where the Laplacian of image is zero, correspond closely to luminance changes that are perceived as edges by the human observer. The basic idea of the proposed gradient scheme is to detect temporal derivatives at the locations of zero-crossings, that is at locations of luminance changes corresponding to edges. Although the gradient scheme looks very different from the correlation scheme, it can be proved that, at least from mathematical point of view, they are equivalent (Hildreth, Koch, 1987).

BASIC PROPERTIES OF ELEMENTARY MOVEMENT MEASUREMENTS

The delay-and-multiply scheme proposed by Reichardt (1957) measures, in fact, the amount of luminous energy concentrated along a certain spatiotemporal motion path and performs, consequently, a local spatiotemporal Fourier analysis. This basic idea is behind various recent explanations of human motion perception that employ motion- sensitive units, with receptive fields that are elongated and oriented in space-time, and thus are tuned to a movement along a definite trajectory. Thus, it is not surprising that various explanations of human motion detection demonstrate a remarkable convergence, becoming essentially identical to only one basic model, the Elaborated Reichardt Detectors (ERD) (van Santen, Sperling, 1984, 1985). In the ERD, point-shaped receptive fields, appropriate for insect facet eyes, were replaced with spatially extended ones which perform linear spatial filtering of the input image. After this amendment ERD became fully equivalent, at least formally, to an elaborated motion detector of Watson and Ahumada (1985) and the spatiotemporal energy model of Adelson and Bergen (1985; see also Burr, Ross and Morrone, 1986).

A sequence of two light stimuli, impinging on two adjacent locations of the retina, is obviously the elementary event that could evoke the perception of motion because it is matched exactly to the structure of an elementary motion encoding operation. If the spatial and temporal interval between two flashes is not too small or too large, it produces a very clear and vivid impression of motion. All correlation-type models make several straightforward predictions for this type of motion stimuli which can be easily subjected to an experimental falsification. Three most important properties of any correlation-type models are the direct consequence of multiplication operation on which these models are supposed to be based:

1. Monotonicity. The strength of the motion response is proportional to the product of amplitudes of the two stimuli and, consequently, the motion direction identification performance must increase monotonically as a function of product of two amplitudes;
2. Commutability. The probability of motion detection does not change when the sequence of two stimuli with unequal amplitudes is reversed. Due to commutability of multiplication, the exchange of spatiotemporal positions of two stimuli cannot be noticed by elementary motion encoders;
3. Sign reversal. If two luminance flashes have opposite polarity, then, following the rule of algebraic sign multiplication, the predicted movement direction is opposite to the actual succession of flashes.

Despite amenability to direct experimental examination only a few attempts to test these predictions more rigorously have been undertaken so far. In particular, van Santen and Sperling (1984) tested directly two of these three predictions, the monotonicity and sign reversal, finding a good agreement between their data and the prediction of the temporal covariance models. They found, in particular, that over a large range of 48:1, the percent of correct motion-direction identification of near-threshold pulses is a monotonically increasing function of the product of the pulses' amplitudes (Experiment 3). At variance from their study, Morgan and Cleary (1992) found, contrary to the monotonicity principle, that direction identification first improves with the increase of the contrast, and then, at higher contrasts, falls again. This near-threshold pedestal effect, well-known from luminance discrimination experiments, obviously violates the prediction of correlation type models. These data suggest a need for further study of this property of elementary movement encoders.

Van Santen and Sperling (1984) found nearly perfect reversal of the perceived motion direction with stimuli of opposite contrast, as it was predicted by both original and elaborated Reichardt models (Experiment 2). Although these confirmations look very encouraging, there are several observations demonstrating their limits. It is usually assumed that the reversed phi motion observed in multi-element cinematograms is an evidence of the sign-sensitive operations in motion measurements. There are, however, two completely different reasons for the reversed motion. If two identical patterns are presented one after another, with a slight spatial shift between them, the motion is perceived in the direction of the later stimulus. But if one of the two patterns is a photographic negative of the other, the motion is perceived in the direction of the earlier stimulus, that is in the opposite direction to the actual displacement (Anstis, 1970; Anstis, Rogers, 1975; Marr, Ullman, 1981; Sato, 1989). This phenomenon was called 'reversed phi motion' and was initially reported for multi-element random-dot patterns. Reversed phi motion is an extreme case of a more general direction reversal phenomenon which can be obtained when each element in a spatially shifted pattern is more likely to reverse its contrast than preserve it (Allik, Dzhafarov, 1984). As a result of contrast reversal the number of element pairs with the same contrast polarity in the direction of actual displacement decreases and becomes less than that in the opposite displacement direction. Thus, the perceived motion direction can be simply determined by the number of potentially displacing elements maintaining their contrast. This explanation does not need an assumption that stimulus elements with the opposite polarity contribute to the perception of motion. The original explanation of the reversed-phi proposed by Anstis (1970) was based on similar statistical considerations. Later, however, he proposed another explanation based on the perceived localization of different luminance distributions (Anstis, Rogers, 1975; Anstis, 1978). Thus, there is no need to explain the motion reversal phenomenon in multi-element patterns in terms of interaction between opposite polarity luminance fluxes. Indeed, one of the most striking properties of the reversed phi motion created by multi-element patterns, both visual and acoustical, is the asymmetry between direct and reversed motion: the identification of the reversed motion is worse than the identification of direct motion (Allik, Dzhafarov, 1984; **Study II**; Sato, 1989). A detailed analysis of this asymmetry leads to the conclusion that only elements with the same polarity contribute to the perception of motion; the contribution of element pairs with the opposite polarity is very small if not negligible. It is also important to notice that this conclusion is generalizable to two-point stimulation, the most preferable type of stimulation of the classical apparent movement paradigm (**Study II**). Recently Edwards and Badcock (1994) also found that dots that change luminance polarity do not effectively drive the global-motion mechanism. Besides these observations some other results suggest that the motion impression can be created by stimulus elements of opposite polarity which direction corresponds not to the reversed but to the actual succession of elements (Livingston, Hubel, 1987; Shechter, Hochstein, 1989; Wehrhahn, Rapf, 1992).

One of the main goals of this series of sudies was to test these three predictions of the Reichardt models.

Monotonicity. **Study V** demonstrates a complete failure of maintaining monotonical relation between the movement direction identification probability and the product of amplitudes. The movement detection performance deviates from the monotonical relation not in particulars or details but in principle. Two separate experiments (experiments 2 and 3) undoubtedly demonstrated that the motion direction discrimination performance becomes almost completely independent of the higher of the two amplitudes provided that it already exceeded a certain critical amplitude level. The success of previous studies (e.g. van Santen and Sperling (1984) and experiment 1 of their study) can be attributed to a fortunate choice of the range of contrast. Indeed, when both signals have relatively low amplitudes the multiplication rule can be used as a satisfactory predictor of the movement discrimination performance. But it fails completely as soon as the amplitude of one of the two signals becomes high enough. Potential difficulties with the monotonicity rule have been realized earlier. In particular, van Santen and Sperling (1985) wrote that the fast saturation of responses to contrast made testing of the constant-product rule extremely problematic. In any case, the saturation of contrast response makes the predictions of both the original and the Elaborated Reichardt Detectors wrong.

Commutability. Contrary to the reported evidences, commutability appears to be valid for the motion encoding, in the first approximation at least (**Study V**). We were not able to find a noticeable asymmetry between low-high and high-low contrast sequences of two unequal luminance amplitudes. As it was already mentioned above, Morgan and Cleary (1992) found that their subjects had a tendency to recognize movement direction in low-high contrast sequence of two-frame random patterns more easily than the reversed high-low sequence. The most simple explanation for this asymmetry is the observer's bias to report more frequently the movement towards more luminous element, irrespective of the temporal order between successive stimuli. Although commutability appears to be valid for two-frame patterns, more complicated stimuli demonstrate clear violation of commutability. In particular, when the observer's task is to identify the temporal order of two adjacent luminance excursions, one of which was a step-function and the other was a linear increase in luminance starting from zero and reaching various final amplitude A after some period of time D, the probability of the movement identification does not remain the same when the sequence of stimuli is reversed (**Study III**). The asymmetry is caused by the fact that perceptual latency which determines timing of visual events used for the determination of motion direction **depends** on the amplitude and waveform of luminance change at a given location.

Sign reversal. Study VI clarifies the controversial issue of the polarity-sensitive operations in the extraction of elementary movement vectors. When a sequence of two opposite polarity flashes is presented, the reversed motion can be perceived only when both spatial separation and temporal asynchrony between two opposite polarity flashes are small. When spatial separation or temporal asynchrony between two flashes is large enough, the movement is perceived in the direction of actual stimulus succession, that is from the stimulus presented earlier in time towards stimulus presented later in time. This implies the existence of the crossover from the reversed to the direct motion direction at some intermediate spatial and temporal separations. The cross-over happens at about 60 msec in the temporal, and about 6–9 min of arc at the spatial domain.

IMPLICATIONS FOR MODELS OF MOTION ANALYSIS

The results of this series of investigations have the following implications for theories of motion analysis.

1. Contrary to the prediction of the Reichardt model, the strength of the motion response is **not** proportional to the product of amplitudes of the two stimuli. The movement detection performance deviates from the monotonical relation not in particulars or details but in principle. The only possibility is to assume, as it is usually done, a fast saturation of responses to contrast which makes testing the constant-product rule extremely problematic.
2. The probability of motion detection **changes** when the sequence of two stimuli is reversed. As it was shown, when two time-varying signals have unequal temporal waveforms, the probability of motion identification changes when these two signals exchange their relative positions.
3. In spite of the rule of algebraic sign multiplication, two luminance flashes of opposite polarity **do not** necessarily appear to move in the direction opposite to the actual succession of flashes. For this very reason the Reichardt model, the original (Reichardt, 1957) or the elaborated one (van Santen, Sperling, 1984), cannot be regarded as a general model of the movement perception: there are many cases when the perceived movement direction is opposite to that predicted by the cross-correlation product (cf. Chubb, Sperling, 1988). For the same reason, all models proposing that the squaring or rectifying operations, that are invariably applied to the input signal, cannot be regarded as sufficiently general (Heeger, 1987, 1991; Werkhoven, Sperling, Chubb, 1993).

Thus, in all three cases the Reichardt model, which is regarded as a universal representation of all known models, appears to be wrong. Egelhaaf and Borst (1989) proposed a new version of the correlation-type movement encoding system in order to account for the transient and steady-state responses of movement-sensitive interneurons in the fly's brain. Their elaboration modified the general scheme of a movement encoding system in the following way: (1) the mean luminance is subtracted from the input signal before it is subjected to a nonlinear compression; and (2) saturation characteristics are inserted into both branches of the two mirror-symmetric motion detection sub-units before the multiplication of the input signals is performed. Our own data based on the psychophysical studies of human observers support these two modifications of the general movement encoding scheme. First, we demonstrated (**Study V**) that the movement detection performance depends on luminance increment (DL) alone, indicating that the background luminance is in some way subtracted out by the nervous system and is affecting perception only by way of retinal adaptation. Second, in order to account for violations of monotonicity it is necessary to assume a rapid compressive nonlinearity before multiplication of the input signals. Only a rapid saturation can explain almost complete independence of the movement direction identification performance from the amplitude of the higher of the two signals when it becomes clearly visible (**Study V**). In addition to modifications proposed by Egelhaaf and Borst (1989), it is also necessary to assume spatio-temporal limits of the sign-sensitive multiplication.

ACKNOWLEDGEMENTS

I would like to take this opportunity to express my sincere thanks to:

Jüri Allik, my teacher, colleague and friend for his long-lasting support and help and as a person who transformed my *apparent movement* towards fundamental problems in science into reality.

Marika, Tiia I, Tiia II, Aavo and Tiit, team of "the vision people" for their cooperation and helpfulness.

All colleagues at the Department of Psychology for their patience and support.

Many people at Universitas Tartuensis who formed my conviction that science is *l'art pour l'art*.

Marvi, Mario and Tanel for their never-ending support, help and love and genuine surprise that something is going to happen, at last ... .

REFERENCES

Adelson, E. H., Bergen, J. R. (1985). Spatio-temporal energy models for the perception of motion. Journal of the Optical Society of America, A2, 284–299.

Allik, J., Dzhafarov, E. N. (1984). Motion direction identification in random cinematograms: A general model. *Journal of Experimental Psychology: Human Perception and Performance*, 10, 378–393.

Anstis, S. M. (1967). Visual adaptation to gradual change of intensity. *Sciene*, 155, 710–712.

Anstis, S. M. (1970). Phi movement as a subtraction process. *Vision Research*, 10, 1411–1430.

Anstis, S. M. (1978). Apparent movement. R. Held, H. W. Leibowitz, H. L. Teuber (Eds.), *Handbook of sensory physiology*. Vol. VIII: Perception, (pp. 656–673). Berlin: Springer-Verlag.

Anstis, S. (1986). Recovering motion information from luminance. *Vision Research,* 26(1), 147–159

Anstis, S. M. (1990). Motion aftereffects from a motionless stimulus. *Perception*, 19, 301–306.

Anstis, S. M., Mather, G. (1985). Effect of luminance and contrast on direction of apparent motion. *Perception*, 14, 167–179.

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

Barlow, H. B. (1972). Single units and sensation: A neuron doctrine for perceptual psychology? *Perception,* 1, 371–394.

Bülthoff, H., Götz, K. G. (1979). Analogous motion illusion in man and fly. *Nature,* 278, 636–637.

Burr, D. C., Ross, J., Morrone, M. C. (1986). Seeing object in motion. *Proceedings of the Royal Society of London*, B227, 249–265.

Burt, P., Sperling, G. (1981). Time, distance, and feature trade-offs in visual apparent motion. *Psychological Review,* 88, 171–195.

Chubb, C., Sperling, G. (1988). Drift-balanced random stimuli: a general basis for studying non-Fourier motion perception. *Journal of the Optical Society of America*, A5, 1986–2007.

Dzhafarov, E. N. (1992). The structure of simple reaction time to step-function signals. *Journal of Mathematical Psychology,* 36, 235–268.

Dzhafarov, E. N., Sekuler, R., Allik, J. (1993). Detection of changes in speed and direction of motion: reaction time analysis. *Perception and Psychophysics*, 54, 733–750.

Edwards, M., Badcock, D. R. (1994). Global motion perception: Interaction of the ON and OFF pathways. *Vision Research*, 34, 2849–2858.

Egelhaaf, M., Borst, A. (1989). Transient and steady-state response properties of movement detectors. *Journal of the Optical Society of America*, A6, 116–127.

Exner, S. (1875). Experimentelle Untersuchung der einfachsten psychischen Processe. III. Abhandlungen. *Archiv für die gesammte physiologie des Menschen und die Thiere*, 11, 403–432.

Exner, S. (1876) Über das Sehen von Bewegungen und die Theorie des zusammengesetzten Auges. *Sitzungsberichte der Kaiserlichen Akademie der Wissenschaften. Mathematisch-Naturwissenschaftlichen Classe.* 72, 156–190.

Exner, S. (1888). Über optische Bewegunsempfindungen. *Biologisches Centralblatt*, 8, 437–448.

Fennema, C. L. Thompson, W.B. (1976). Velocity determination in scenes containing several moving objects. *Computer Graphics and Image Processing*, 9, 301–315.

Gregory, R. L., Heard, P. F. (1983). Visual dissociations of movement, position and stereo depth: some phenomenal phenomena. *Quarterly Journal of Experimental Psychology,* 35A, 217–237.

Heeger, D. J. (1987). Model for the extraction of image flow. *Journal of the Optical Society of America,* A8, 1455–1454.

Heeger, D. J. (1991). Nonlinear model of neural responses in cat visual cortex. M. Landy, A. Movshon (Eds.), *Computational models of visual processing,* (Chap. 9). Cambridge, Mass.: MIT Press.

von Helmoltz, H. (1911). *Handbuch der Physiologischen Optik.* Bd. II, (3. Auflage). Hamburg und Leipzig: Leopold Voss.

Hildreth, E. C. (1984). *The measurement of visual motion.* Cambridge, Ma.: MIT.

Hildreth, E. C., Koch, C. (1987). The analysis of visual motion: from computational theory to neural mechanisms. *Annual Review of Neurosciences,* 10, 477–533.

Hubel, D. H., Wiesel, T. N. (1959). Receptive fields of single neurones in the cat's striate cortex. *Journal of Physiology,* 148, 574–591.

Johansson, G. (1950). *Configurations in event perception.* Uppsala: Almquist & Wiksell.

Johansson, G. (1978). Visual event perception. R. Held, H. W. Leibowitz, H. L. Teuber (Eds.), *Handbook of sensory physiology.* Vol. VIII: Perception (pp. 675–711). Berlin: Springer-Verlag.

Koffka, K. (1935). *Principles of Gestalt psychology.* New York: Harcourt, Brace.

Lettvin, J. Y., Maturana, H. R., McCulloch, W. S., Pitts, W. H. (1959). What the frog's eye tells the frog's brain. *Proceedings of I.R.E.,* 47, 1940–1951.

Limb, J. O., Murphy, J. A. (1975). Estimating the velocity of moving images in television signals. *Computer Graphics and Image Processing,* 4, 311–327.

Livingstone, M. S., Hubel, D. H. (1987). Psychophysical evidence for separate channels for the perception of form, color, movement, and depth. *Journal of Neuroscience,* 7, 3416–3468.

Marr, D., Hildreth, E. C. (1980). Theory of edge detection. *Proceedings of the Royal Society of London,* 207B, 187–217.

Marr, D., Ullman, S. (1981). Directional selectivity and its use in early visual processing. *Proceedings of the Royal Society of London,* 211B, 151–180.

Mastebroek, H. A. K., Zaagman, W. H. (1988). Apparent movement induced by luminance modulation: a model study. *Perception,* 17, 667–679.

Mather, G. (1984) Luminance change generates apparent movement: Implications for models of directional specificity in the human visual system. *Vision Research,* 24, 1399–1405.

Miller, G. F., Shepard, R. N. (1993). An objective criterion for apparent motion based on phase discrimination. *Journal of Experimental Psychology: Human Perception and Performance,* 19, 48–62.

Morgan, M. J., Cleary, R. (1992). Effect of contrast substitutions upon motion detection in spatially random patterns, *Vision Research,* 32, 639–643.

Poggio, T., Reichardt, W. (1973). Considerations on models of movement detection. *Kybernetiks*, 13, 223–227.

Reichardt, W. (1957). Autokorrelationsauswertung als Funktionsprinzip des Zentralnervensystems. *Zeitschrift für Naturforschung*, B12, 448–457.

Sato, T. (1989). Reversed apparent motion with random dot patterns. *Vision Research,* 29, 1749–1758.

Sekuler, R., Ganz, L. (1963). A new after-effect of seen movement with a stabilized retinal image. *Science,* 139, 419–420.

Shechter, S., Hochstein, S., Hillman, P. (1988). Shape similarity and distance disparity as apparent motion correspondence cues. *Vision Research,* 28, 1013–1021.

Shechter, S., Hochstein, S. (1989). Size, flux and luminance effects in the apparent motion correspondence process. *Vision Research,* 29, 579–591.

Ullman, S. (1979). *The interpretation of visual motion.* Cambridge, MA: MIT Press.

van Santen, J. P. H., Sperling, G. (1984). Temporal covariance model of human motion perception. *Journal of the Optical Society of America*, A1, 451–473.

van Santen, J. P. H., Sperling, G. (1985). Elaborated Reichard Detectors. *Journal of the Optical Society of America*, A2, 300–321.

Watson, A. B., Ahumada, A. J. Jr. (1985). Model of human visual-motion sensing. *Journal of the Optical Society of America,* A2, 322–342.

Wehrhahn, C., Rapf, D. (1992). ON- and OFF-pathways form separate neural substrates for motion perception: psychophysic evidence. *Journal of Neuroscience*, 12, 2247–2250.

Werkhoven, P., Snippe, H. P., Koenderink, J. J. (1990). Metrics for the strength of low-level motion perception. *Journal of Visual Communication and Image Representation,* 1, 176–188.

Werkhoven, P., Snippe, H. P. (1990). Effects of element orientation on apparent motion perception. *Perception & Psychophysics,* 47, 509–525.

Werkhoven, P., Sperling, G., Chubb, C. (1993). The dimensionality of texture-defined motion: A single channel theory. *Vision Research*, 33, 463–485.

Wertheimer, M. (1912). Experimentelle Studien über das Sehen von Bewegung. *Zeitschrift für Psychologie,* 61, 161–265.

ELEMENTAARSE LIIKUMISVEKTORI KODEERIMINE INIMESE NÄGEMISSÜSTEEMIS

KOKKUVÕTE

Dissertatsioon on pühendatud laiemalt sõnastades tajupsühholoogia ühele fundamentaalprobleemile: millised on baasoperatsioonid, mida rakendatakse kujutise tajumisel inimese nägemissüsteemis. Vaatluse alla on võetud üks baasoperatsioonidest — liikumise kirjeldamine ehk elementaarse liikumisvektori eraldamine heleduse aeg-ruumilises jaotuvuses. Lähtudes psühhofüüsikalisest ja komputatsioonilisest traditsioonist, uuritakse empiiriliselt erinevate mudelite omadusi ja nende paikapidavust. Dissertatsioonis esitatakse olulisemad ajaloolised etapid liikumistaju teemas, vaadeldakse lähemalt liikumisanalüsaatori kirjeldamiseks pakutud mudelite klasse ja analüüsitakse nende mudelite fundamentaalseid printsiipe. Empiiriliste uurimuste tsükli tulemusena formuleeritakse need põhimõttelised omadused, mida tuleb arvestada liikumisanalüsaatori töö kirjeldamisel.

Dissertatsioonis esitatud järeldused põhinevad järgmistel töödel:

I töö (üldse esimene selletaoline empiiriline uurimus liikumistaju-alases kirjanduses) on pühendatud kinemaatiliste lävede uurimisele selektiivse adaptatsiooni tingimustes. Tulemused lubavad väita, et nägemissüsteemis eksisteerivad bilokaalsed liikumisdetektorid, mis kodeerivad liikumise suunda teineteisest sõltumalt.

II töö on pühendatud liikumise suuna kodeerimise uurimisele stohhastilistes kinematogrammides. Stohhastilised kinematogrammid on üldistatud stiimulite klass, mis hõlmab kõiki kinematogrammi tüüpe. Lähtudes Allik-Dzhafarov'i Dipoolide Kontributsiooni Mudelist (1984), uuritakse liikumisvektori eristamise seaduspärasusi sõltuvalt ajalis-ruumilistest parameetritest. Tööst selgub, et reversiivse liikumise suuna paradoks on seletatav erisuunaliste dipoolide suhtega. Töös näidatakse, et Dipoolide Kontributsiooni Mudel on laiendatav ka lineaarsetele kinematogrammidele ja on võimeline tegema kvantitatiivseid ennustusi liikumise suunavektori kodeerimisprotsessi kohta.

III töös demonstreeritakse, et liikumisanalüsaator ignoreerib kiireid muutusi heleduse aegruumilises jaotuvuses. Tulemuste interpreteerimisel formuleeritakse reeglid, mis seostavad signaali amplituudi liikumissuuna avastamisega.

IV töö näitab, et liikumisvektori arvutamise aluseks on heleduse modulatsioon, mitte aga ruumilise asukoha muutus ajas. Selle töö raames esitatakse andmed, mis näitavad, et liikumisanalüsaatori sisendis on signaal allutatud mittelineaarsele teisendusele.

V töös formuleeritakse korrelatiivsete mudelite 3 baasomadust ja kontrollitakse eksperimentaalselt kahte neist: kommutatiivsust ja monotoonsust. Antud töö tulemused lubavad väita, et nii kommutatiivsuse kui ka monotoonsuse omaduse kehtivusega on suuri probleeme. Näidatakse, et monotoonsus on kehtiv ainult lävelähedaste heleduste diapasoonis. Antud töös esitatakse need põhimõttelised täiendused, mida on vaja arvestada liikumisvektori arvutamise mudeli kirjeldamiseks. Esitatakse järeldus, et liikumise detektsioon heleduse aeg-ruumilises jaotuvuses on kontrasti detektsiooni erijuhtum.

VI töö käsitleb liikumissuuna kodeerimist sõltuvalt stiimulite kontrasti märgist, s.t. siin kontrollitakse kolmandat korrelatiivsete mudelite baasomadust. Selles töös antakse ülevaade reversiivse liikumise kahest tüübist ja analüüsitakse vastavaid empiirilisi tulemusi. Eksperimentide põhjal on leitud need aeg-ruumilised piirid, milles kehtib korrelatiivse mudeli ennustus erimärgilise kontrastiga stiimulitele. Tulemused näitavad, et mudelis peab kajastamist leidma süsteemi tundlikkus kontrasti märgile sõltuvalt aeg-ruumilistest parameetritest. Töös esitatakse hüpotees, et neuronaalsel tasemel eksisteerivad spetsiifilise retseptiivväljaga funktsionaalsed ühikud.

Dissertatsiooni üheks põhitulemuseks on järeldus, et korrelatiivsete mudelite klass vajab edasist täiendamist, sest ükski mudelite kolmest baasomadusest ei kehti. Nii teoreetiline kui ka empiiriline analüüs lubab väita, et liikumist kirjeldav baasoperatsioon ei ole midagi muud, kui kontrasti avastamise eriline juhtum. Saadud tulemused lubavad esitada mõned olulised täiendused liikumisanalüsaatori mudelile. Peamised nendest on: fooni keskmise heleduse lahutamine signaalist ja mittelineaarse teisenduse rakendamine signaalile.

PUBLICATIONS

I

Acta et Commentationes Universitatis Tartuensis, 1977, 429, 107–120.

DETECTION OF TEMPORAL PHASE BY DIRECTIONALLY SENSITIVE UNITS IN THE HUMAN VISUAL SYSTEM

J.Allik, M.Tepp, A. Livshits

1. Introduction

Asynchronous change in the luminance of the two disparate areas in the visual field is the most simple form of stimulation which gives rise to a visually perceived movement /Exner, 1875; Thorson, Lange and Biederman-Thorson, 1969/. The temporal phase between some features of the two temporally varying luminance contours activates the directionally sensitive units in the human visual system. As it is known from the responses of the directionally sensitive units to the threshold contrast, one subset of these units responds solely or predominantly to one direction of movement and does not respond to movement in the opposite direction at all. Consequently the same stimulus input is analysed by the two antisymmetric subsystems each of them tuned to a different sign of the temporal phase.

The directionally sensitive units, which are of primary interest in this study, are revealed by the ability to detect the smallest available temporal phase between two relatively short rectangular flashes. Vertical rectangular bars facing each other are flashed for the short time and the minimal temporal phase needed for a detection of the direction of the displacement is recorded. As the amplitude of the luminance transients greatly exceeds many times the threshold value, then the suprathreshold contrast response of the movement coding mechanism has been measured in this study. The subject is asked to match the perceived direction of the stimulus configuration displacement without notions about the subjective quality or "goodness" of the perceived movement. In this paper we have studied the minimal temporal phase measure in order to answer a question: in what way the directionally sensitive units are composed in the human visual system?

We should like give a reference to a previously published theses on the related experiments reported here /Allik et al., 1976/.

2. Method. Apparatus and stimulation. The stimulus patterns used in the experiments described in this paper are formed by the appropriate aperture on the face of the green solid state light emitters. The light emitters have an emission maximum at about 510 nm. The luminance of the flashes is estimated at about 5 nt. A three-channel mirror tachistoscope is used for the presentating of the stimuli on a black background. The modulating voltage of the stimuli is controlled by a programmed stimulator which gives out any requested duration and phase of stimuli, with a precision of 4% and with summation of the minimal time step. The temporal width of the time step is 2.5 ms. One channel of the stimulator is controlled continuously. The time intervals of this channel are monitored by a frequency-meter 0-5080.

The spatial configuration of the stimuli is shown in Fig.1. From one to four closely spaced rectangular bars are presented for a short time. Each bar was 6.5 mm wide and 8.0 mm high, subtending 0.22° and 0.27° at a viewing distance of 172 cm. The stimuli are presented parafoveally in the upper part of the visual field. A light emitter diode is used as the fixation point at a distance of 50 mm below the principal stimuli, subtending 1.67° at the menteoned viewing distance. Let the bars from the left to right be named A, B, C and D. The duration of all stimuli is constant and is equivalent to 60 ms during the whole session of the experiment.Let t_1, t_2, t_3 and t_4 represent the onset time of the stimuli A,B,C and D respectively.Now one can give a definition for the temporal phases f_1,f_2 and f_3 as $f_1 = t_1-t_2$, $f_2 = t_2-t_3$ and $f_3=t_3-t_4$. If the temporal phases have positive values,then the perceived displacement of the whole pattern to the right.

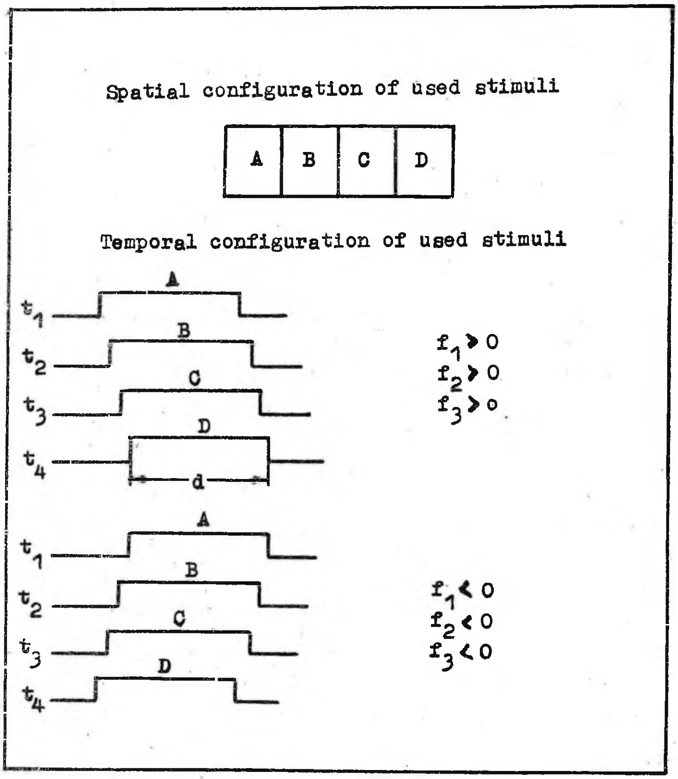


FIGURE 1. The spatial and temporal configuration of used stimuli. A, B, C, D : stimuli used in the experiments; t_1 , t_2, t_3,t_4: the onset time of the stimuli A,B,C,D respectively; f_1, f_2, f_3: temporal phases of stimuli presentation, $f_1 = t_1 - t_2$, $f_2 = t_2 - t_3$ $f_3 = t_3 - t_4$.

One phase between two neighbours is smoothly controlled by the experimenter. Initially the stimuli are locked into the zero phase in order to eliminate any visible displacement. The phase threshold giving visible displacement to the right or to the left was determined by an adjustment procedure. The temporal phase is continously increased or decreased from one test to another up to the time the threshold level is reached. The direction of the phase change is varied from trial to trial in random sequence, so that the subject lacks prior kbowledge about the displacement direction. In some cases the initial position of the physical phases has not to be equal to zero. The subjective point of anequilibrum is determined by means of the subject's match.

Subjects. Four subjects were used. Two female, T.L. (23 years old) and M.P. (24 years old) were naive in respect to the experimental purposes. One female, M.T. (23 years old) and one male A.L. (20 years old) are the authors of this article. T.L., M.P., A.L. were well corrected myopes, M.T. had normal vision. All subjects have had many hours' practice in experiments of this kind.

Adaption to displacement. There are several similarities between the successive effects in the spatial-frequency, orientation and movement domains. Viewing a grating with specific spatial frequency, orientation and direction of movement for some time may cause several aftereffects. First, the detection threshold elevation is specific with respect to spatial frequency, orientation as well direction of movement. Second, after the adaptation subsequently viewed stimuli are distorted in their apparent spatial frequency, orientation and direction of movement. Third, the compound adapting pattern consisting of two or more separate components is much less effective as suprathreshold adapting stimuli than would be predicted from their effects viewed in isolation. This weakening of the adaptation effect is explained as an inhibitory interaction between sensory channels in spatial frequency, orientation or movement domain.

Two types of adaptation patterns were used in these experiments. First, two bars B and C were used as the adaptation pattern. The temporal asynchrony between B and C was $f_2=+30$ ms (perceived displacement to the left) or $f_2 = -30$ ms (perceived displacement to the right). These temporal values were experimentally selected as the most agreeable phase between onset of the two flashes for the detecting of the displacement by the subject. Second, three bars were presented. The temporal phases f_1 and f_2 between the stimuli A,B and C were chosen with the opposite sign so that the perceived displacement of the components was in the opposite direction. If $f_1 = +30$ ms, then $f_2 = -30$ ms, or vice versa: $f_1 = -30$ ms and $f_2 = +30$ ms. In the other words, the second adaptation condition was a sum of the two simple oppositely directed adaptation patterns. The adaptation period, in both cases, lasted for 60 s. Just after the adaptation period the subject matched the phase required for the detecting of the displacement to the right or left. Before and after the series with adaptation , control threshold matching was performed. The main results are presented in Table 1. The threshold values of f_2 in positive and negative directions are shown depedently on the adaptation conditions. The threshold elevation is significant only if the direction of the adapting and testing displacement is the same. Thus, as was to be expected, the adaptation is specific to the direction of movement: rightward movement elevates the thresholds to the right and leftward movement elevates the thresholds to the left leaving the opposite thresholdsunchanged. As is shown (Table 1) by the one-way analyses of variance, the magnitude of the threshold elevation is highly significant. The threshold after the adaptation is twice compared with the control series. The adaptation to the compound pattern, sum of the two simple components, significantly elevates the temporal thresholds in both directions , except in the case of AL when the thresholds of the leftward displacement were not elevated significantly. The threshold elevation is remarkably reduced compared with the adaptation to the simple unidirectional pattern. t-test shows the existence of a significant difference between displacement thresholds of the two adaptation conditions. First, the difference between the mean values of the thresholds is significant on the level of probability p .05 ($t = 4.7$, subject TL; $t = 2.7$, subject AL) for the leftward displacement-testing; and the same for the rightward testing p 0.5 ($t=2.0$, subject TL; $t = 1.1$, subject AL). Consequently, the adapting pattern, as a sum of two unidirectional adaptation patterns , is a much less effective adapting stimulus than its compo -

nents in isolation. One should keep in mind that the simple-pattern adaptation reveals the isolation between mechanisms coding the movement to the left and to the right. The adaptation to the movement in one direction does not affect the thresholds in the opposite direction. Similar results are found in the spatial frequency domain /Tolhurst, 1972/ and in the movement direction domain /Levinson, Sekuler, 1975/ although the near threshold contrast responses of the spatial frequency and movement detecting channels are measured in the papers referred to.

TABLE 1. Magnitude of the threshold phases $+f_2$ (displacement to the left) and $-f_2$ (displacement to the right) in milliseconds for two subjects TL and AL depedently on the various conditions of adaptation. The means, standard deviations and the results of the one-way analyses of variance compared with the control series are presented

Subjects and thresholds		Control	Conditions of the adaptation		
			Simple pattern		Compound pattern
			$+f_2$ (left)	$-f_2$ (right)	$-f_1+f_2$ or $+f_1-f_2$ (left + right)
TL					
$+f_2$ (left)	$\bar{x}$	7.4	15.3	7.7	10.6
	s	1.4	2.2	1.2	1.8
			F = 98.9		F=79.7
			p .001		p .001
$-f_2$ (right)	$\bar{x}$	4.9	6.0	12.8	10.8
	s	1.5	1.8	2.0	1.5
				F =149.9	F= 45.1
				p .001	p .001
AL					
$+f_2$ (left)	$\bar{x}$	4.0	11.0	4.2	5.0
	s	1.4	1.6	0.7	1.0
			F=90.7		F=2,8
			p .001		-
$-f_2$ (right)	$\bar{x}$	4.0	4.2	8.0	6.9
	s	1.7	1.0	1.2	1.0
				F=107.6	F=57.3
				p .001	p .001

Thresholds dependent upon the background-tendency

The spatial lay-out of the stimuli A, B, C and D is areally quite limited. It should be proposed that this configuration covers the spatial sampling area of only one functional unit which extracts the information about movement from the stimulated retinal region. The receptive field subserving the motion detection in the foveal region of man has a diameter of about 3.5° inferred from the measure of the minimum contrast needed to create an aftereffect of movement (Richards,1971). This inferred value is substantially larger than the spatial extent of the maximum stimuli configuration in the present experiments. These facts of the case are the reason that nobody can see the stimuli A,B, C and D moving or shifting in two opposite directions simultaneously. The temporal phases having the opposite signs are subtracted one from another and the difference, not necessarily the ratio, is displayed as the result of the transformation. Let us call C and D the test stimuli and f_3 the test-phase of the test stimuli. The threshold of the test-phase is altered by adding an additional stimulus B or/and A to the principal test stimuli C and D. Let us define f_1 and f_2 as a background-phase. Now the thresholds of the test-phase should be measured as the function of the sign and the degree of the background-phase.

The results are shown in Figures 2 and 3. Figures 2A and 2B show the data using three-stimuli configuration, more exactly B, C and D. Figures 3A and 3B show the results of combination from four stimuli A, B, C and D. In these Figures the temporal threshold of the test-phase f_3 is expressed as a function of the sign, positive value corresponds the leftward movement and negative value corresponds to the rightward movement, and the magnitude of the background temporal phase. The upper row of the experimentally determined points and the straight diagonal line which is a least-squares estimate of the experimental results indicate the thresholds for leftward displacement. The lower row and its linear approximation in all these figures represent the threshold for rightward displacement. Vertical bars indicate ± 1S.E. of the mean (n = 10). The filled triangles settled in the line of $f_2 = 0$ or $f_1+f_2=0$ indicate the thresholds for leftward displacement (upper

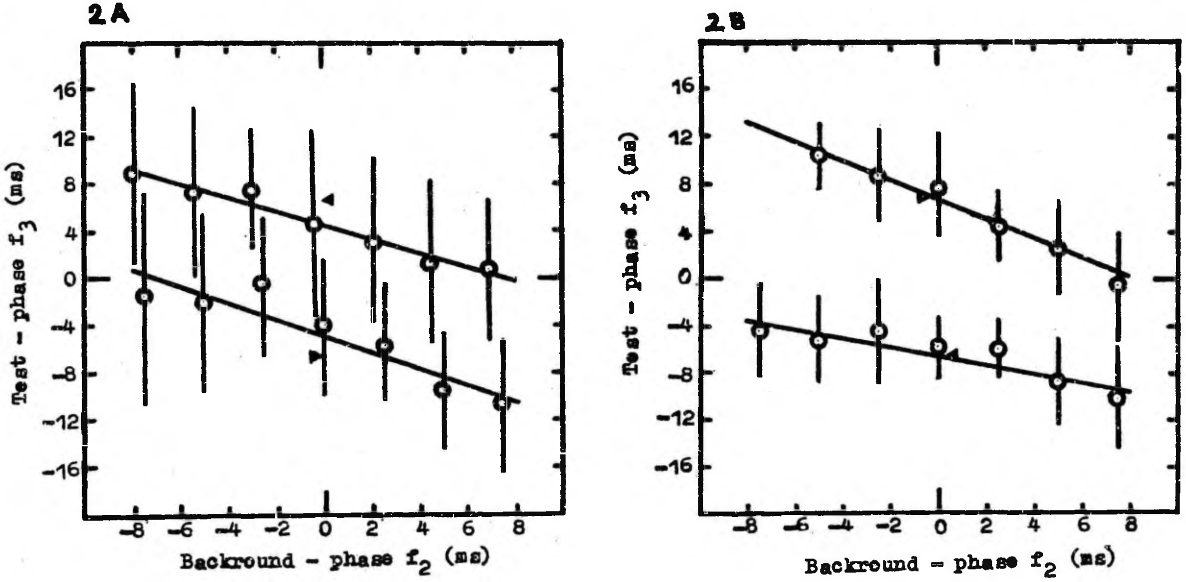


FIGURE 2. The data using three stimuli configuration B, C and D respectively. 2A: results of the subject M.P.; 2B: results of the subject M.T. Upper row of the experimentally determined points and its linear approximation indicate the thresholds for leftward displacement as a function of the sign and the degree of the backround-phase, lower row for rightward displacement respectively. The filled triangles indicate the thresholds for displacements without the additional backward stimulus.

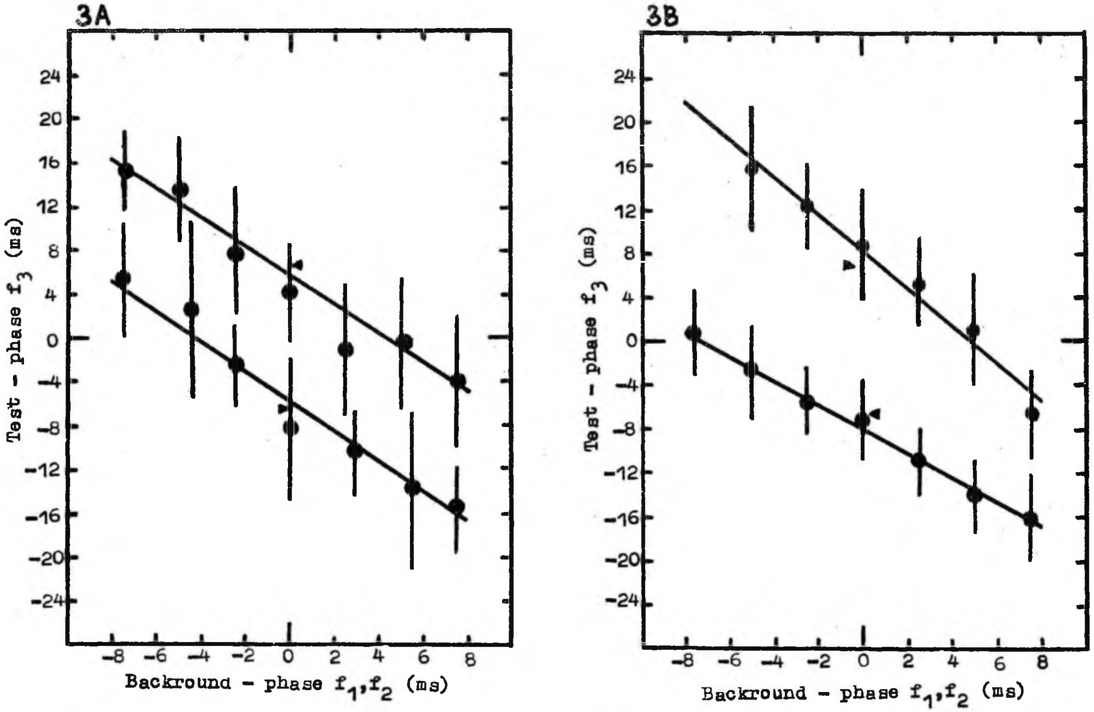


FIGURE 3. The data using four stimuli configurations A, B, C and D. The thresholds of the test phase are expressed as a function of the sign and degree of the background phase. Upper row indicates the thresholds for leftward displacement, lower row for rightward displacement respectively. The filled triangles indicate the thresholds for displacements without the additional backward stimulus. 3A: results of the subject M.P., 3B: results of the subject M.P.

triangles) and for rightward displacement (lower triangles) without the additional background stimulus.The numerical values of the least-squares approximations are shown in Table 2.

<u>TABLE 2</u>. Parameters of the least-squares linear approximation of the experimental results indicated in Figures 2 and 3. The values of the linear regression $y = ax+b$ and the coefficient of correlation r are shown for both subjects M.T. and M.P.

Figure and subject	Direction of the displacement	a	b	r
Fig. 2A	left	-0.61	4.81	.98
M.P.	right	-0.70	-4.79	.95
Fig. 2B	left	-0.86	6.50	.98
M.T.	right	-0.36	-6.49	.89
Fig. 3A	left	-1.35	5.78	.99
MFP.	right	-1.35	-5.76	.96
Fig. 3B	left	-1.71	8.13	.98
M.T.	right	-1.10	-8.10	.99

If the backround has a tendency to be displaced to the left, it causes a proportional displacement of the thresholds to the right.The coefficient "a" shows the exact value of this proportionality.If the background drifts to the right then it is identical to the shifting of the rightward and leftward thresholds in the opposite,i.e. left, direction. Therefore the detection of the displacement direction in the suprathreshold compound counterphase pattern is linearly related to the magnitude of the opposite direction.If the background-phase has the same sign as the phase of the test displacement the threshold level is proportionally reduced up to the test-phase sign inversion (e.g.Figure 3-A;B). The reduction of the thresholds due to the same sign of the test-phase and background-phase has the same value as in the opposite sign case. So there is a partical additivity between background and test-phase as "a"<1 for the Figure 2 and "a"< 2 for the Figure 3. There exist strong linear relationships between test and background components because of the high linear correlation from .89 to high as .99. Finally we should like to draw attention to the deadzone between the two thresholds lines where the movement perception is lacking. The width of this dead-zone is about 8-17 ms and with numbers of bars in the pattern it is remarkably increased.

4. Discussion

The thresholds measured in this study have some pecularities which make the present thresholds determination essentially different from a convenient psychophysical sensitivity determination. As said above, the subject was required to match the direction of the displacement. We supposed that the identification of the displacement direction, i.e. the discrimination of the temporal asymmetry in the stimulation, is sufficient and required provision to specify the directionally sensitive units in the human visual system . The correct identification of the direction is regarded as a measure of the directionally sensitive units and not as a response of the local flicker detection mechanism.The suprathreshold contrast response of the directionally sensitive units is measured in this study. The contrast of the stimuli remained constant during the whole session of the experiments having a 100% value. Under these conditions the temporal resolution limits of the directionally sensitive units are revealed. The finest temporal asynchrony between onset of the two flashes spatially facing each otherм required for the identification of the displacement direction is determined as a limit of the temporal resolution the mentioned system. This limit of the temporal resolution established for the directionally sensitive units one may compare with the classical measure of the critical flicker frequency.

The main results of this paper are the following. First, the adaptation to the displacement in one direction elevates the thresholds only in the direction of the adaptation displacement. In its nature this directionally sensitive adaptation effect is similar to the contrast-threshold elevation after an adaptation to a unidirectionally moving grating /Sekuler and Ganz,1963/. So the selective adaptation shows the strong isolation between mechanisms coding the opposite directions of the displacement. The adaptation affects only one direction, leaving the opposite direction unchanged at all. This suggests that there must be at least two different mechanisms,each of them activated by one temporal sequence and not activated by another temporal sequence. Second, the

ДЕТЕКЦИЯ ВРЕМЕННОГО СДВИГА ЕДИНИЦАМИ, ЧУВСТВИТЕЛЬНЫМИ К НАПРАВЛЕНИЮ В ЗРИТЕЛЬНОЙ СИСТЕМЕ ЧЕЛОВЕКА

Аллик Ю., Тепп М., Лившиц А.

Резюме

В данной работе изучали детекцию временного сдвига стимулов единицами, чувствительными к направлению в зрительной системе человека. Главные результаты экспериментов следующие:

I. Существует селективная адаптация по отношению к направлению сдвига, что дает основание предполагать наличие изолированных механизмов, кодирующих противоположные направления.

2. Адаптация на изображение, сдвигающееся одновременно в двух противоположных направлениях менее эффективна, по сравнению с адаптацией на сдвиги в одном направлении. Такой результат говорит в пользу существования тормозных связей между единицами, кодирующими направление в зрительной системе человека.

3. Фон, сдвигающийся в противоположном направлении, по сравнению с тестом, повышает пороги сдвига теста, в то время как при использовании фона и сдвигающегося в одном направлении теста, обнаруживается понижение порогов сдвига теста.

Очевидно для обнаружения направления сдвига нужен такой энергетический уровень теста, который мог бы превышать инертность всей стимульной конфигурации.

II

Acta et Commentationes Universitatis Tartuensis, 1986, 753, 18–35 (*in Russian*).

ИДЕНТИФИКАЦИЯ НАПРАВЛЕНИЯ ДВИЖЕНИЯ В СЛУЧАЙНЫХ ЛИНЕЙНЫХ КИНЕМАТОГРАММАХ

А.Пульвер, Ю.Аллик

В качестве стимулов использовался линейный ряд элементов, каждый из которых мог находиться в одном из двух состояний. В течение одного кадра (I00 мсек) состояние элементов оставалось неизменным. В то же время, состояние каждого элемента в данном кадре повторяло с вероятностью P (вероятность сохранения состояния) состояние его ближайшего левого и правого элемента в предыдущем кадре. Вероятность идентификаций сдвига зависела от числа элементов в кадре и от числа кадров. Результаты согласуются с данными, полученными для циркулярных кинематограмм, а также подтверждают состоятельность так называемой модели дипольных вкладов (Allik, Dzhafarov ,I984), Согласно этой модели, идентификация сдвига основывается на подсчете кратчайших диполей (пара пространственно неидентичных соседних элементов, принадлежащих разным кадрам). Однако для 3 испытуемых из 5, вероятность подсчета кратчайших диполей при числе элементов 4 и менее – для двух испытуемых, или при числе кадров 4 и менее – для одного испытуемого, отличается от вероятности подсчета кратчайших диполей для остальных комбинаций числа элементов и кадров. Приводятся соображения о причинах такого отклонения.

I. Введение

Стохастически организованные динамические многоэлементные паттерны (кинематограммы) являются одним из основных средств исследования восприятия направления движения (Anstis , I970, I978, I986; Julesz, I97I; Hochberg ,I968; Lee, I972; Braddick ,I973; Lappin, Bell, I976; Bishof, Groner, I985; Baker, Braddick, I985; Nakayama, I985 и др.). Как правило, кинематограммы состоят из набора геометрически идентичных элементов, каждый из которых может находиться в одном из конечного числа возможных состояний. Для предъявления элементы кинематограммы организуются в определенную пространственную конфигурацию. Из числа одномерных конфигураций линейные и циркулярные являются наиболее типичными. Во временном измерении кинематограмма, как правило, разбивается на последовательность равнодлительных интервалов времени (кадров), в течение которых состояние элементов не меняется. Для порождения впечатления движения между состояниями элементов, принадлежащих разным кадрам, устанавливается определенная зависимость. В наиболее простом случае элементы предыдущего кадра сдвигаются в одном определенном направлении на некоторое число элементов, состояние которых реплицируется в соответствующих элементах последующего кадра. В обобщенном случае состояние элемента из предшествующего кадра не должно обязательно детерминистически отображаться в состоянии парируемого элемента некоторого последующего кадра. При стохастической схеме парирование состояния связанных между собой элементов повторяется с некоторой вероятностью P, далее сокращенно названной вероятностью сохранения состояния (ВСС).

Данная работа посвящена исследованию способности наблюдателя идентифицировать вектор движения в линейных стохастических кинематограммах (ЛСК) с варьируемым числом элементов, кадров и ВСС.

2. Метод

Стимулы. ЛСК формировались на горизонтальном ряду 30 светоизлучающих диодов (АЛ102В) красного цвета. На расстоянии 171 см диаметр каждого светодиода равнялся 8 угловым минутам, а весь горизонтальный ряд светящихся элементов занимал 5 град (пробел между двумя соседними светодиодами равнялся 2 угловым минутам). Весь ряд светодиодов находился на прямоугольном фоне (7 x 9 град), который подсвечивался равномерным красным цветом. Голова наблюдателя удерживалась в неподвижном состоянии при помощи подбородника. Эксперимент проходил в затемненном помещении.

Состояние светодиодов управлялось ЭВМ "Электроника ДЗ-28" через специальный интерфейс. Предъявление ЛСК было разделено на последовательность равнодлительных интервалов времени (кадров) с длительностью 100 мс, в течение которой состояние (включенное или выключенное) элементов оставалось неизменным. Межкадровый интервал времени, требуемый для смены кадров, не превышал 0,6 мс. Эксперимент был разделен на 8 сессий, каждая из которых соответствовала одному определенному типу ЛСК (E,F), где E - число задействованных элементов в одном кадре, а F - число последующих кадров. Используемые ЛСК были следующими: ЛСК (15, 2), ЛСК (15,4),

ЛСК(15,8), ЛСК(15,16), ЛСК(2.15), ЛСК(4,15), ЛСК (8,15), ЛСК(30,15). В каждой сессии выбиралось 7-9 значений ВСС (P), последовательность которых была рандомизирована. Перед каждой отдельной пробой в случайном порядке определялось направление сдвига. Вероятность включения элемента в первом кадре ЛСК определялась с безусловной вероятностью 0,5. Каждый последующий кадр генерировался из предыдущего следующим образом: весь ряд сдвигался на один элемент влево или вправо и состояние элемента в новой позиции сохранялось с вероятностью P.

Задача наблюдателя заключалась в каждой отдельной пробе в идентификации направления сдвига. Ответы наблюдателя регистрировались и обрабатывались ЭВМ. Каждая комбинация элементов (E), кадров (F) и ВСС (P) повторялась в эксперименте 100 раз (по 50 в каждом из двух направлений). Последовательность сессий в эксперименте была рандомизированной. В эксперименте участвовало 6 наблюдателей (3 мужчин и 3 женщины) в возрасте от 22 до 34 лет с нормальным или корригированным до нормального зрением. В ходе эксперимента пришлось отказаться от данных одного наблюдателя, которые оказались очень неустойчивыми, хотя по своему характеру не отличались от данных, полученных от 5 остальных наблюдателей.

3. Результаты

Полученные психометрические кривые идентификации направления движения в ЛСК показаны на рис.1-5, где на оси абсцисс отложена ВСС , а на ординате - вероятность идентификации направления истинного сдвига ЛСК. В самом общем плане характер поведения психометрических кривых совпадает с тем, что было описано для идентификации вектора поворота для круговых случайных кинематограмм (см. Allik, Dzhafarov,1984). Наиболее характерные свойства полученных психометрических кривых:

1) С увеличением числа элементов (E) и/или кадров (F) психометрические значения увеличиваются, т.е. с увеличением общего числа элементов ЛСК при одном и том же значении ВСС психометрические значения все более отличаются от уровня случайного угадывания 0,5;

2) При значениях $P<0,5$ (т.е. в тех случаях, когда вероятность несохранения состояния элемента больше ВСС), наблюдатель с большей вероятностью указывает на направле-

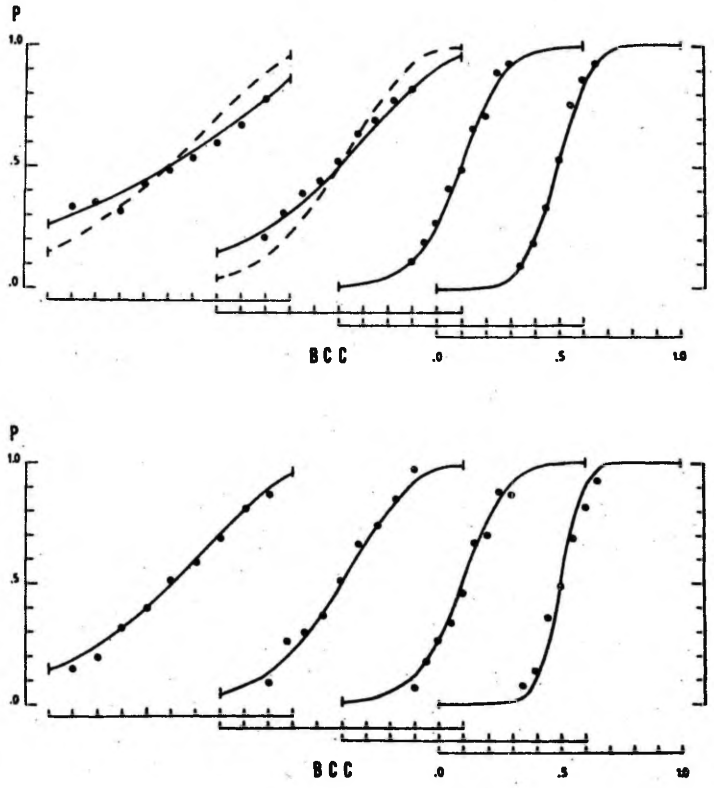


Рис. I. Вероятность идентификации направления сдвига от ВСС, числа кадров (F) и числа элементов (E). Ордината: вероятность идентификации направления сдвига. Абсцисса: вероятность сохранения состояния. ЛСК (E,F). Вверху слева направо: ЛСК (I5,2), ЛСК (I5,4), ЛСК (I5,8), ЛСК (I5,I6). Внизу слева направо: ЛСК (2,I5), ЛСК(4,I5), ЛСК(8,I5), ЛСК(30,I5). Испытуемый Т.Л. Каждое значение вероятности идентификации направления сдвига основывается на IOO измерениях.

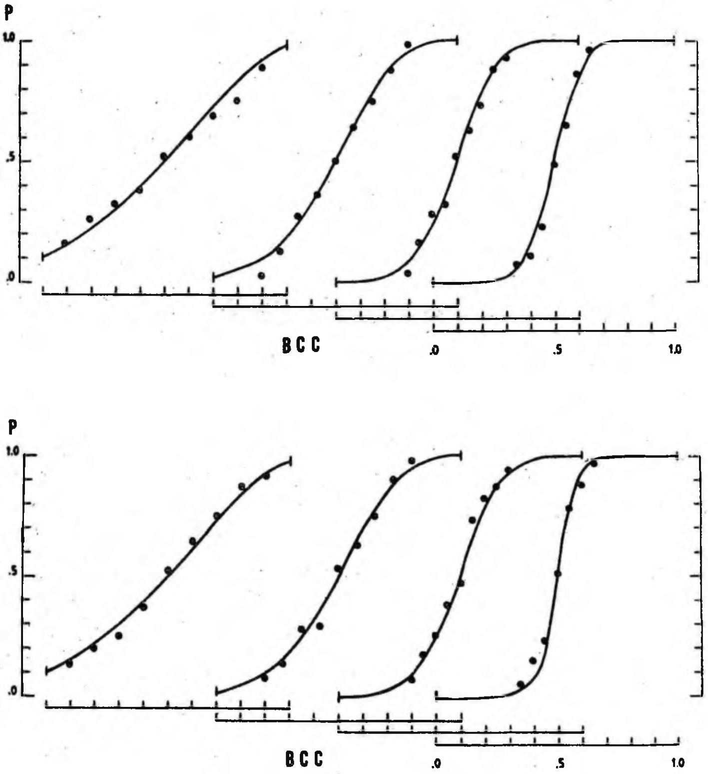


Рис. 2. Вероятность идентификации направления сдвига от ВСС, числа кадров (F) и числа элементов (E). Ордината: вероятность идентификации направления сдвига. Абсцисса: вероятность сохранения состояния.ЛСК(E,F). Вверху слева направо: ЛСК(I5,2), ЛСК(I5,4), ЛСК(I5,8), ЛСК(I5,I6). Внизу слева направо: ЛСК(2,I5), ЛСК(4,I5), ЛСК(8,I5), ЛСК(30,I5). Испытуемый А.П. Каждое значение вероятности идентификации направления сдвига основывается на IOO измерениях.

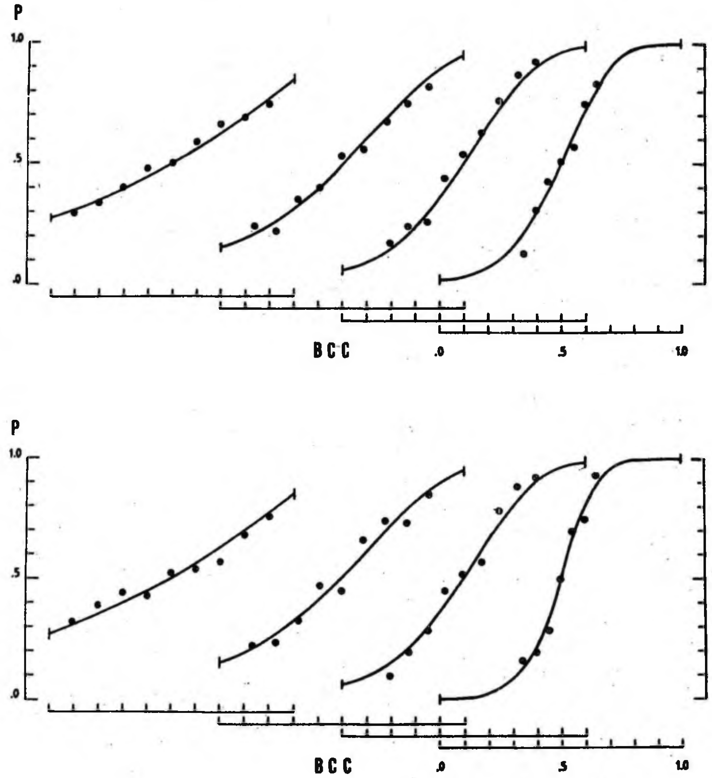


Рис. 3. Вероятность идентификации направления сдвига от ВСС, числа кадров (F) и числа элементов (E). Ордината: вероятность идентификации направления сдвига. Абсцисса: вероятность сохранения состояния. ЛСК(E,F).

Вверху слева направо: ЛСК(I5,2), ЛСК (I5,4), ЛСК(I5,8), ЛСК(I5,I6).

Внизу слева направо: ЛСК(2,I5), ЛСК (4,I5), ЛСК(8, I5), ЛСК(30,I5).

Испытуемый И.М. Каждое значение вероятности идентификации направления сдвига основывается на 100 измерениях.

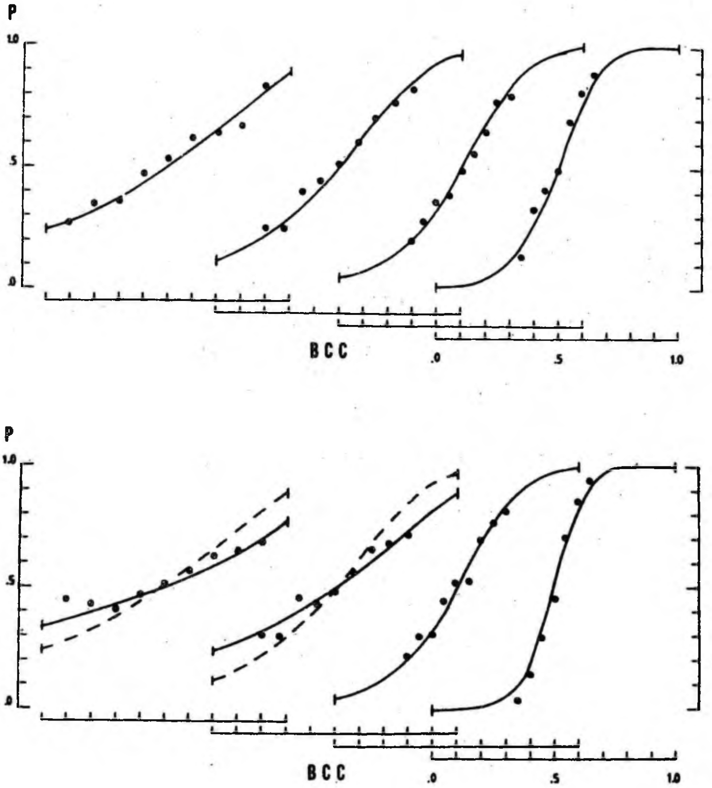


Рис. 4. Вероятность идентификации направления сдвига от ВСС, числа кадров (F) и числа элементов (E). Ордината: вероятность идентификации направления сдвига. Абсцисса: вероятность сохранения состояния. ЛСК(E, F).
Вверху слева направо: ЛСК(I5,2), ЛСК (I5,4), ЛСК(I5,8), ЛСК(I5,16).
Внизу слева направо: ЛСК (2,I5), ЛСК (4,I5), ЛСК(8,I5), ЛСК(30,I5).
Испытуемый Т.Т. Каждое значение вероятности идентификации направления сдвига основывается на 100 измерениях.

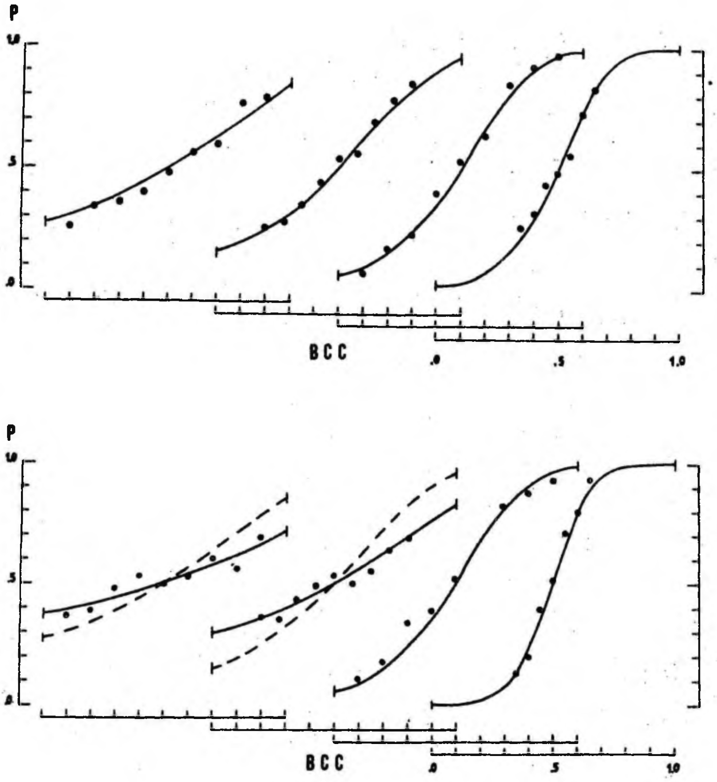


Рис. 5. Вероятность идентификации направления сдвига от ВСС, числа кадров (F) и числа элементов (Е). Ордината: вероятность идентификации направления сдвига. Абсцисса: вероятность сохранения состояния. ЛСК(Е, F).

Вверху слева направо: ЛСК(I5,2), ЛСК (I5,4), ЛСК(I5,8), ЛСК(I5,I6).

Внизу слева направо: ЛСК(2,I5), ЛСК (4,I5), ЛСК(8,I5), ЛСК(30,I5).

Испытуемый Э.К. Каждое значение вероятности идентификации направления сдвига основывается на I00 измерениях.

ние, противоположное истинному вектору сдвига, положенное в основу генерации данной ЛСК. В работе Allik, Dzhafarov (1984) это свойство психометрических кривых было названо феноменом реверсии направления, являющимся обобщенным вариантом реверсивного фи-движения, описанного Anstis(1970, 1986); Anstis, Rogers (1975); Gregory, Heard (1983);

3) Характерным свойством всех психометрических кривых является их асимметричность относительно точки (0,5; 0,5). Смысл этой асимметричности заключается в том, что психометрические значения идентификации реверсии движения находятся ближе к уровню ответов 0,5, чем соответствующие психометрические значения идентификации прямого движения. По всем пяти наблюдателям и типам ЛСК можно организовать 150 пар психофизических значений /Prob(P);Prob(1-P)/ предполагая, что $P > 0,5$. Очевидно, что показателем асимметричности является неравенство

$$Prob(P) + Prob(1 - P) > 1 \qquad (1)$$

которое в настоящем случае выполняется в 113 случаях, из общего числа 150. Вероятность того, что такое число соблюдения неравенства возникло случайно исключительно мала (по крайней мере, $P < 0,001$). Следует напомнить, что аналогичная асимметричность наблюдается и при восприятии круговых стохастических кинематограмм.

Для объяснения выделения наблюдателем направления движения была предложена общая модель, смысл которой вкратце сводится к следующей простой идее. Предполагается, что для объяснения суждений, выносимых наблюдателем, необходимо выявить набор элементарных структур кинематограммы (т.е. определенную конфигурацию элементов, составляющих кинематограмму) и правила их комбинаций при вынесении суждений наблюдателем. Наиболее элементарной из возможных структур является диполь - пара пространственно неидентичных элементов, принадлежащих разным кадрам кинематограммы. Каждый диполь, в зависимости от его вектора (направленного отрезка, соединяющего два элемента кинематограммы) и формы (для бинарных кинематограмм диполь может быть когерентным, если оба элемента находятся в одинаковом состоянии — /0,0/, /1,1/ —, или некогерентным, в случае /0,1/ и /1,0/), имеет свой вес и с некоторой вероятностью P учитывается наблюдателем при вынесении суждения. Поскольку суммарный вклад (TC - Total contribution) складывается из большого числа элементарных вкладов, необходимо решить вопрос о комбинации элементарных вкладов в суммарный результат. Естественно начать исследование с некоторого числа самых простых предположений о том, как элементарные вклады комбинируются. В работе Allik, Dzhafarov (1984) было сделано предположение, что комбинация элементарных контрибуций в суждение, выносимое наблюдателем, подчиняется следующим простым правилам:

I) Гомогенность. Вклад каждого диполя зависит только от его формы – когерентной или некогерентной – и не зависит от положения диполя в пространстве и времени;

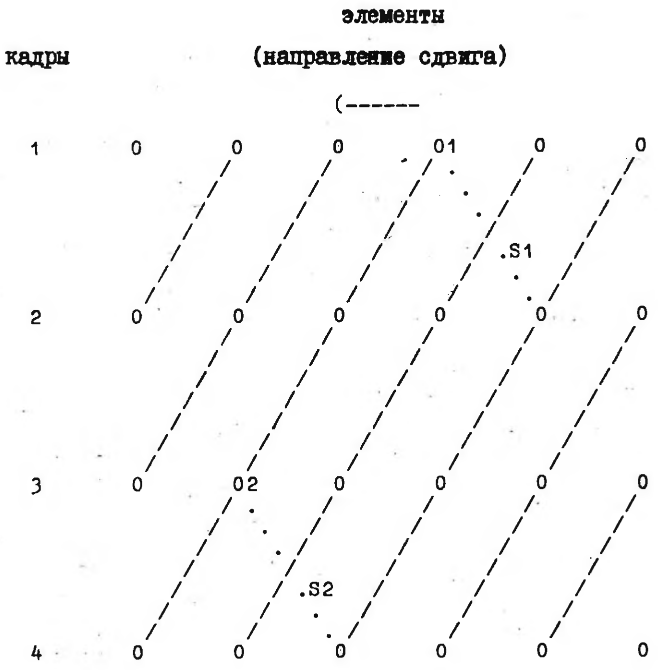
2) Симметричность. Вклад двух симметричных (разнонаправленных) диполей имеет симметричное распределение вокруг нуля (т.е. вклады отличаются только по знаку);

3) Независимость. Вклады любых двух диполей S1 и S2 являются стохастически независимыми.

Оказалось, что этих простых правил комбинации и диполя в качестве элементарной структуры достаточно для объяснения выделения вектора движения в круговых стохастических кинематограммах (Allik, Dzhafarov , 1984). Следует обратить внимание на то обстоятельство, что после выделения элементарной структуры изображения и правил комбинации этих элементарных структур, формальная структура модели и вычислительные формулы выводятся практически автоматически.

Несмотря на элементарность диполя как структуры изображения, число различных структур (диполей разной длины и разного направления) окажется достаточно большим даже при небольшом числе составных элементов и кадров кинематограммы. В своей общей форме так называемая модель дипольных вкладов (или контрибуций) не является моделью в буквальном смысле этого слова. Это скорее определенный язык описания, в котором число свободных параметров слишком велико для того,чтобы быть эмпирически проверяемым. К счастью, как было показано, при вынесении суждения о направлении движения в большинстве случаев наблюдатель принимает в учет лишь наиболее короткие диполи, т.е. диполи, соединяющие соседние элементы из двух последующих кадров. Восприятие ротации циркулярных стохастических кинематограмм оказалось легко интерпретировать в терминах счета: наблюдатель с единичной вероятностью подсчитывает число когерентных диполей, прибавляя к этому числу с некоторой постоянной вероятностью Beta число некогерентных диполей, ошибочно принятых за когерентные; то направление движения, в сторону которого подсчитано наибольшее число когерентных диполей (сокращенно скачков), выбирается наблюдателем в качестве ответа.

Для получения вычислительной формулы модели необходимо знать две математические величины: (I) математическое ожидание числа скачков кратчайшей величины в разных направлениях и (2) их дисперсию. Единственная формальная трудность связана с вычислением дисперсии числа когерентных диполей (скачков), так как диполи не являются статистически независимыми, т.е. в кинематограмме, полученной по вышеописанной схеме, существуют такие пары диполей, при которых вероятность того, что они обе находятся в когерентном состоянии, не равняется произведению их индивидуальных вероятностей нахождения в когерентной форме. Формально говоря, если P(S1) и P(S2) являются вероятностью того, что два диполя S1 и S2 находятся в когерентной форме, то P(S1) (S2) не равняется P(S1)·P(S2). При короткоамплитудном варианте модели имеется только одна форма коварьирующих диполей, которая показана на следующей схеме:



Вероятность того, что диполь 0I–02, с длиной вектора к кадров, является скачком, можно выразить суммой (Sum)

$$P' = \mathrm{Sum}_{I:\ 2I<K} \binom{K}{2I} \cdot P^{K-2I} \cdot Q^{2I}$$

где Q=I–P (суммирование по I с шагом 2 до К). Зная значение P', можно найти вероятность того, что диполи S1 и S2 одновременно являются когерентными (скачками) $P(S1)(S2) = \left[P'^2 + (1-P')^2\right]/2$. Не вдаваясь в остальные математические подробности, которые достаточно подробно описаны в приложениях к статье Allik, Dzhafarov (I984), приведем сразу окончательные вычислительные формулы, которые использовались для аппроксимации экспериментальных данных:

$$E(TC) = N \cdot (P - 0{,}5) \tag{2}$$

$$Var(TC) = N \cdot \left[Par1 \cdot (P+0{,}5) + Par2 \cdot (Q+0{,}5) + (PQ+0{,}25)\right] \tag{3}$$

$$Cov = 2 \cdot \mathrm{Sum}\left[P(S1)(S2) - 0{,}25\right] \tag{4}$$

где N – число наиболее коротких диполей, P– вероятность того, что элемент сохранит состояние парируемого элемента, Sum – символ суммирования (суммирование осуществляется по всем парам S1 и S2). Par1 и Par2 два свободных параметра модели, которые по своему содержанию являются дисперсией вкладов, вносимых скачками и нескачками соответственно, нормированной по математическому ожиданию разницы вкладов скачков и нескачков. Формула (2) является числителем, а квадратный корень суммы формул (3) и (4) является знаменателем верхнего предела стандартного нормального распределения:

$$Prob = \Phi\left\{E(TC) / \mathrm{Sqr}\left[Var(TC) + Cov\right]\right\} \tag{5}$$

где Ф – знак стандартного нормального интеграла.

В этой форме модель с двумя свободными параметрами была приложена к полученным психометрическим кривым. Прежде всего следует отметить, что аппроксимирующая программа для всех наблюдателей установила Par1 близкой к нулю. Действительно, аппроксимация при условии Par1=0 лишь на I–2% ухудшала качество приближения в смысле минимума хи-квадрата, а оставшийся свободный параметр Par2 практически не менялся. Этот результат полностью согласуется с данными, полученными при исследовании круговых стохастических кинематограмм, и показывает, что из множества возможных короткоамплитудных моделей, справедливой оказывается модель "счета нескачков" (Nonjump Counting Model). Смысл Par1=0 заключается в том, что вклад когерентных диполей /I,I/ и /0,0/ равняется некоторому детерминистическому значению W или говоря в терминах счета, когерентные диполи учитываются с вероятностью I и с весом W. Если предположить, что вес W одинаков как для когерентных, так и некогерентных диполей (это может означать, что, нескачки иногда просто принимаются за скачки в соответствующем направлении), то можно проинтерпретировать параметр Par2 в терминах счета. Пусть Вета - вероятность того, что некоторый самый короткий нескачок смешивается со скачком, в этом случае Вета и Par2 связаны следующим образом:

$$\text{Вета} = \text{Par2}/\,(1 + \text{Par2}) \qquad (6)$$

Итак, мы имеем модель с одним-единственным параметром Par2 (Вета), имеющую достаточно ясную содержательную интерпретацию, которая должна предсказать 68 разных психометрических значений, полученных при разных комбинациях E, F и P. Если учесть только сами значения хи-квадрата, то качество аппроксимации можно считать удовлетворительным для всех 5 наблюдателей. Однако анализ теоретических кривых показывает, что по крайней мере для трех наблюдателей невозможно найти единственное значение Par2, которое достаточно хорошо описывает поведение психометрических кривых при разных значениях E и F. При малом числе элементов (E) или кадров (F) значения Par2 начинают существенно отличаться от Par2 полученных при большом числе элементов и кадров. При этом наблюдается весьма любопытное разделение между наблюдателями: у наблюдателя Т.Л. отклонения наблюдаются в том случае, если число кадров является малым (2 или 4), в то время, как наблюдатели Т.Т. и Э.К. испытывали трудности с ЛСК с малым числом элементов (так же 2 или 4). Особо следует подчеркнуть, что такая картина является очень устойчивой для данного конкретного наблюдателя и практически не меняется в ходе продолжительной тренировки, включающей обратную связь. Резюмируя положение, можно сказать, что для аппроксимации данных Т.Т., Т.Л. и Э.К. необходимо использовать два разных значения Par2, так как восприятие ЛСК с малым числом элементов или кадров отличается от многоэлементных ЛСК, по крайней мере, количественно. Результаты анализа предоставлены в следующей таблице:

E	P		Т.Л.	А.П.	И.М.	Т.Т.	Э.К.
все		Par2	I.92	I,29	6.06	4.32	5.86
		Beta	.66	.56	.85	.8I	.85
I5	2	Par2*	5.55	–	–	–	–
I5	4	Beta*	.88	–	–	–	–
2	I5	Par2*	–	–	–	I2.86	2I.80
4	I5	Beta*	–	–	–	.93	.96
хи-квадрат(68,6800)			II.7I	I0.26	II.96	II.3I	II.I2
P > 0.05			+	+	+	+	+

Звездочкой (*) помеченные параметры обрабатывались отдельно от других ЛСК. Соответствующие приведенным параметрам психометрические кривые показаны на рис.I–5 непрерывными кривыми. При отклоняющихся данных (*) прерывистой кривой показана теоретическая кривая, соответствующая аппроксимации всех остальных кривых. По степени расхождения можно судить, насколько восприятие малоэлементных/кадровых ЛСК отличается от восприятия остальных ЛСК

Следует указать, что по параметру хи-квадрат аппроксимация данных оказалась исключительно хорошей: при 68 степенях свободы (N=6800) эмпирические значения лежат значительно ниже критического значения. Удовлетворяет и общий вид аппроксимации: нет систематических трендов, показывающих различие между теоретическим предсказанием и эмпирическими данными.

4. Обсуждение

Общая характеристика модели. Имеет смысл коротко изложить суть предложенной модели. Модель строится на предположении, что в ЛСК имеется некоторая элементарная структура, на основе которой наблюдатель принимает решения о направлении движения кинематограммы. Наиболее простым кандидатом в качестве такой структуры является диполь, т.е. пара элементов, принадлежащих разным кадрам. Анализ данных показывает, что для объяснения полученных психометрических кривых достаточно учитывать лишь наиболее короткие диполи, т.е. такие, которые образуются из соседних элементов двух следующих друг за другом кадров, суждения о направлении сдвига принимаются на основе подсчета числа наиболее коротких скачков (диполей когерентной формы /I,I/ и /0,0/) в том и другом направлении. К этому числу с некоторой постоянной вероятностью Вета прибавляется некоторое число нескачков (некогерентные диполи /I,0/ и /0,I/), которые ошибочно воспринимаются как скачки в том же направлении. В качестве ответа – "движение влево" или "движение вправо" – принимается то направление, в котором подсчитано большее число скачков. Надо отметить, что понятие "скачок" является техническим и не связано с субъективным впечатлением о перемещении некоторого объекта из одного пространственного положения в другое. Предложенная формулировка модели может показаться достаточно правомерной, если не учитывать тот поразительный факт, что оптимальное приближение было достигнуто при условии, что Par1 =0. Асимметричность психометрических кривых сама по себе не удивляет и, очевидно, может достигаться только при условии, что Par1 < Par2 ,однако из этого не следует, что Par1должен равняться нулю, а не некоторому другому положительному значению.

Отклонения при малом числе кадров или элементов. В работе Allik, Dzhafarov (I984) число элементов в одном кадре оставалось всегда постоянным. В настоящей работе E и F систематически варьировались от минимальных до достаточно больших значений. При такой вариации было обнаружено, что часть наблюдателей начитает испытывать трудности в выделении вектора движения, если число кадров или (в смысле исключающего) элементов меньше или равно 4. С чисто формальной точки зрения это означает, что модель,одним основным постулатом которой является пространственно-временная гомогенность вносимых диполями вкладов, неверна,и может претендовать лишь на первое, достаточно грубое описание эмпирических данных. С другой стороны, данная модель, как впрочем и любая другая, делает ряд априорных допущений о способностях наблюдателя, которые в прямой форме не отражаются в формамальной структуре модели. Одним таким допущением,например, является предположение о том, что наблюдатель обладает способностью восстанавливать (точнее интерполировать) положение выключенного элемента. Дело в том, что выключенный элемент ЯСК фактически не отличается от остального фона. В принятой версии модели предполагается, что психологические вклады гомогенных диполей /I,I/ и /0,0/ не отличаются друг от друга. Очевидно, для выполнения этого постулата наблюдатель должен обладать способностью вычислять положение "дырок" в каждом отдельном кадре. Средняя вероятность включения любого элемента ЛСК всегда равняется 0,5, в силу чего вероятность того, что все четыре ближайших соседа по пространству и времени также окажутся выключенными, достаточно мала. А следовательно положение "дырки" можно достаточно хорошо восстановить по включенным соседним элементам. Если, однако, число элементов в кадре мало или число кадров не составляет более 2–4, то у большого числа элементов нет достаточного числа соседей по пространству или по времени. Оставаясь полностью на уровне спекуляций, мы предполагаем, что выпадение крайних точек, ясно наблюдаемое у трех из 5 наблюдателей, может объясняться именно отказом механизма интерполяции. Эти нестрогие рассуждения приводятся тут лишь с целью указать на возможность сохранения логики самой модели при дальнейшей разработке предполагаемого "базиса", который пока оставался в неявно выявленной форме. Мы полностью осознаем, что высказанные соображения окажутся достаточно убедительным лишь после изыскания независимого экспериментального способа их проверки.

<u>Место данной работы в исследованиях движения.</u> Классическая парадигма исследования фи-движения исходила из узкой модели стимульной ситуации. Как правило, в качестве стимулов использовались лишь двухкадровые изображения, отличающиеся друг от друга в основном по пространственному сдвигу геометрически идентичного объекта. Следует обратить внимание на то, что ЛСК с $E=2$ соответствует именно наиболее популярной стимульной конфигурации крассической парадигмы. Изобретение многоэлементных кинематограмм явилось существенной генерализацией стимульных условий, однако, как отмечалось выше, все известные нам случаи применения кинематограммы использовали детерминистическую схему генерирования кинематограммы, соответствующую двум экстремальным случаям ЛСК, а именно $P=1$ и $P=0$. Класс стохастических кинематограмм, круговых или линейных, является расширением класса кинематограмм, так как охватывает все значения ВСС, находящиеся между 0 и 1. В целом данные, полученные на ЛСК, совпадают с данными, полученными при использовании круговых стохастических кинематограмм. Это позволяет заключить, что основные закономерности восприятия стохастических кинематограмм (вариант модели "счета нескачков", наличие феномена

реверсии направления и т.д.) существенно не зависят от конкретной конфигурации кинематограммы. Кроме того, предложенный формальный язык является до сих пор единственным, позволяющим не только предсказывать качественные аспекты кажущегося движения, но и в строгой количественной форме предсказывать поведение психометрических кривых идентификации направления движения в случайных многоэлементных кинематограммах.

Литература

1. Allik J., Dzhafarov E.N. Motion direction identification in random cinematograms: a general model // Journal of Experimental Psychology: Human Perception and Performance-1984.-Vol.10.-P.378-393.
2. Anstis S. Phi movement as subtraction process // Vision Research-1970.-Vol.10.-P.1411-1430.
3. Anstis S. Apparent movement // Handbook of sensory physiology / Eds. R.Held, H.W. Leibowits, H.L. Teuber.- Berlin: Springer, 1978.-Vol.8: Perception. -P.656-673.
4. Anstis S. Recovering motion information from luminance // Vision Research-1986.-Vol.26.-P.147-160.
5. Anstis S.M., Rogers B.J. Illusory reversal of depth and movement during changes of contrast // Vision Research-1975.-Vol.15.-P.957-961.
6. Baker C.L., Braddick O.J. Temporal properties of the short-range process in apparent motion // Perception-1985.-Vol.14.-P.181-192.
7. Bishof W.F., Groner M. Beyond the displacement limit: an analysis of short-range process in apparent motion // Vision Research-1985.-Vol.25.-P.839-848.
8. Braddick O. The masking of apparent motion in random-dot patterns // Vision Research-1973.-Vol.13.-P. 355-369.
9. Gregory R.L., Heard R.F. Visual dissociations of movement,position, and stereo depth: some phenomenal phenomena // Quarterly Journal of Experimental Psychology-1983.-Vol.35A.-P.217-237.
10. Hochberg J. In the mind's eye // Contemporary theory and research in visual perception / Ed. R.N. Haber. -New York: Holt,Rinehart & Winston, 1968.-P.309-331.

11. Julesz B. Foundation of cyclopean perception.-Chicago: Chicago University Press, 1971.-315 p.
12. Lappin J.S., Bell H.H. The detection of coherence in moving random-dot patterns // Vision Research - 1976.-Vol.16.-P.161-168.
13. Lee D.N. Stimulus pairing in sequential phi motion // Perception-1972.-Vol.1.-P.85-91.
14. Nakayama K. Biological image motion processing: a review // Vision Research-1985.-Vol.25.-P.625-660.

MOTION DIRECTION IDENTIFICATION IN LINEAR RANDOM CINEMATOGRAMS

A.Pulver J.Allik

Abstract. The linear array of two-state elements presented in a sequence of 100-ms frames was used as a stimulus. The state of every element depends from the state of its left or right neighbour-element from the preceding frame,repeating the latter's state with the probability P (State Repetition Probability). The shift direction identification functions were obtained as a function of number of elements in the array (E) and frames (F). The results confirm general findings obtained with circular random cinematograms and show the applicability of the Dipole Contribution Model (Allik & Dzhafarov, 1984). According to this model the observer's decisions are based on the counts of the shortest dipoles (pair of neighbour-elements from the successive frames) of the different types. It was found that three from five observers could not maintain the same dipoles counting probability when the number of elements (two observers) or frames (one observer) became less or equal 4. Some speculations about reasons of this oddity are presented.

III

Vision Research, 1994, 34, 1585–1594.

Pergamon

0042-6989(93)E0014-X

Vision Res. Vol. 34, No. 12, pp. 1585–1594, 1994

0042-6989/94 $7.00 + 0.00

Timing of Visual Events for Motion Discrimination

JÜRI ALLIK,* ALEKSANDER PULVER*

Received 12 May 1993; in revised form 21 October 1993

The observer's task was to identify the temporal order of the two adjacent luminance excursions one of which was a step-function and the other a linear increase in luminance starting from zero and reaching various final amplitude A after some period of time D. The interstimulus delay, Δt, between these two transitions was determined at which they appeared isochronous. The point of the subjective equality (PSE) depended on both ramp parameters, the rise-time duration D and its amplitude A. All of the data can be accounted for by supposing that judgements about the temporal order are based on the comparison of a simple attribute of these luminance excursions, the time moments when the luminous energy concentrated on low temporal frequencies exceeds some level. The perceived temporal order, which was experienced as a leftward or rightward displacement of the whole pattern, was determined by the sequence in which low-frequency portion of these two luminance excursions reached the threshold value. The implications of this simple contrast detection explanation for theories of motion analysis are discussed.

Temporal order perception   Motion analysis   Timing of visual events   Perceptual latency   Visual masking

INTRODUCTION

In many cases, the human observer can easily tell which of two events occurred first. The capacity to make deliberate discrimination of temporal order of two perceptual events drops to chance when the difference between their onsets is about 40–50 msec (Exner, 1875; Hirsh & Sherrick, 1961; Allan, 1975; Westheimer & McKee, 1977b). On the other hand, the subject can easily distinguish sounds, flashes, and vibrations that differ only in the order in which two component stimuli occurred at a fraction of the interstimulus time interval at which he or she can explicitly specify their order (Efron, 1973). Analogously, the temporal order of two adjacent flashes can be correctly identified when one is delayed by as little as 3–5 msec (Sweet, 1953; Westheimer & McKee, 1977b). The observer's sensitivity to the difference between two asynchronies is even better than the detection of temporal order *per se*. The smallest difference in the asynchrony of two flashes which can be reliably detected is about 2 msec (Burr, 1979; McKee, 1981). Two spatially proximate asynchronous flashes create a clear impression of motion which direction depends on the order of appearing of two flashes. Motion can be seen even though two stimuli could not be spatially resolved when they are exposed simultaneously (Exner, 1876; Thorson, Lange & Biederman-Thorson, 1969), Evidently, the observer's exquisite timing accuracy of visual events at small spatial separation is a consequence of the requirement to resolve visual motion.

The prediction of the temporal order of two identical visual events appears to be an elementary task logically at least: if the temporal asynchrony between two visual events is long enough their order can be determined on the basis of their succession. The situation becomes more complicated, however, when events are not completely identical. There is no more a single and intuitively transparent rule how to predict the apparent perceptual order of two arbitrary luminance excursions. The moving target leaves behind a translating pattern of changes in the luminance flux across the retinal surface. Such a translation produces typically almost identical but delayed luminance variations at two neighbouring sites along the motion path. Following this line of reasoning, Reichardt (1957) proposed an attractively simple delay-and-multiply scheme for the detection of motion. According to this scheme, the basic operation for the motion detection is the multiplication of a signal received from one spatial location with a delayed signal from another adjacent spatial location. Although the response of the bilocal analyzers is maximized when two time-varying input signals have identical form and their relative delay corresponds to the internal delay, motion can be perceived when two signals have different forms.

In the present study, we investigated the ability to discriminate temporal order of two luminance transitions varied in the duration of their onset time. The onset time defined as the time needed for the amplitude

---

*Department of Psychology, University of Tartu, Tartu, Estonia EE-2400.

17

of the signal to reach its final steady-state level. The delay, Δt, between two transitions was varied in the manner that produced a full psychometric function—i.e. with the change of Δt the choice probability of one of the two opposite motion direction increased smoothly from about zero to about one. The location on the psychometric function at which the opposite movement directions were chosen with equal 0.5 probability is the point of subjective equality (PSE). At that stimulus onset asynchrony (SOA) two luminance transitions seem to begin simultaneously and, as a consequence, the whole stimulus pattern appears stationary on the average. Various luminance transitions were compared with one and the same reference signal and, consequently, all differences in delays needed to synchronize the test stimuli corresponded exclusively to their relative perceptual onset latencies. Our goal, then, was to find a precisely specified *property* of luminance changes from which it will be possible to predict the exact perceptual simultaneity of two luminance onsets.

METHODS

Stimuli

A spatiotemporal representation of stimuli is shown in Fig. 1. A schematic spatial view of the stimulus (lower-left) and its temporal course (upper-right) are shown. Two rectangular elements 0.06 deg in height and 0.03 deg in width were formed by two red light-emitting diodes (LEDs) attached to the back of the milk-film screen forming the background for these elements. The spatial separation between two rectangles was 0.03 deg. The emission maximum of diodes was at 690 nm. The background had the form of a square, the size of which subtended 5.7 deg at the 2 m viewing distance, and was uniformly illuminated through a colour filter with approximately same spectrum. The mean luminance of the background was 0.25 cd/m^2 measured by an electronic photometer. The luminance of LEDs was controlled by a personal computer via two 12-bit digital-to-analogue converter (DAC) outputs at a sample rate of 1 kHz. All stimuli were linear onset ramps starting from zero and reaching a steady-state amplitude A after D msec. They can be described by the equation

$$L(t) = A \cdot k(t) \begin{cases} k(t) = 0 & \text{if } t < 0 \\ k(t) = t/D & \text{if } 0 \leqslant t \leqslant D \\ k(t) = 1 & \text{if } t > D \end{cases} \quad (1)$$

where $k(t)$ is the temporal waveform, D is the duration of the linear rise, and A is the final steady-state amplitude reached by the signal. Five different rise-time durations D of 0, 30, 60, 120, and 240 msec with three different amplitudes A of 6, 12, and 24 cd/m^2 were used. One of the two waveforms, $L_r(t)$, served as a **reference** being the same in all trials. It had the instantaneous rising time ($D = 0$ msec) and the amplitude $A = 6$ cd/m^2. The **test** stimulus $L_t(t)$ was one of 15 possible waveforms (5 rise times by 3 amplitudes) which started with **Δt delay** relative to the reference stimulus onset (positive values of Δt mean that the luminance rise of the test stimulus started *before* the reference luminance jump occurred). For brevity the motion direction towards the test stimulus position will be called **positive**.

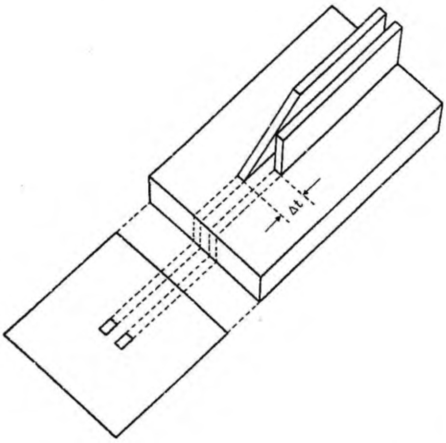


FIGURE 1. Schematic spatial view of the stimulus (lower-left) and its temporal course (upper-right) used in this experiment. The luminance of two small adjacent rectangular areas perturbed from a uniform background. One of these perturbations (right) was a step-function and the other (left) was a gradual increase in luminance. The observer's task was to judge at a given SOA (Δt) value in which direction, to the left or right, the whole stimulus appears to move.

Procedure

The whole experiment was divided into 15 sessions corresponding to a certain test stimulus waveform with a fixed amplitude A and rise time duration D. The particular value of the temporal delay Δt and the spatial position of the test stimulus were chosen before each trial quasi-randomly. The observer's task was in each trial to press one of the two buttons indicating in which direction the whole pattern appeared to move. The responses were coded in terms of the positive direction choice probability, that is choosing the displacement direction from the position of the reference stimulus towards the test stimulus position. Data points in the empirical psychometric function were approximated by a cumulative Gaussian function using the downhill simplex method of minimization.

Two measures characterizing the *sensitivity* and *bias* were derived from these functions. The slope of the function, measured as the distance between 0.5 and 0.75 points, indicates the observer's ability to discriminate temporal order of two luminance onsets and corresponds to the classical concept of the just noticeable difference (JND). The horizontal position of the psychometric function was measured by the location of 0.5 point. At that delay value the PSE is reached since the opposite displacement directions were chosen with equal

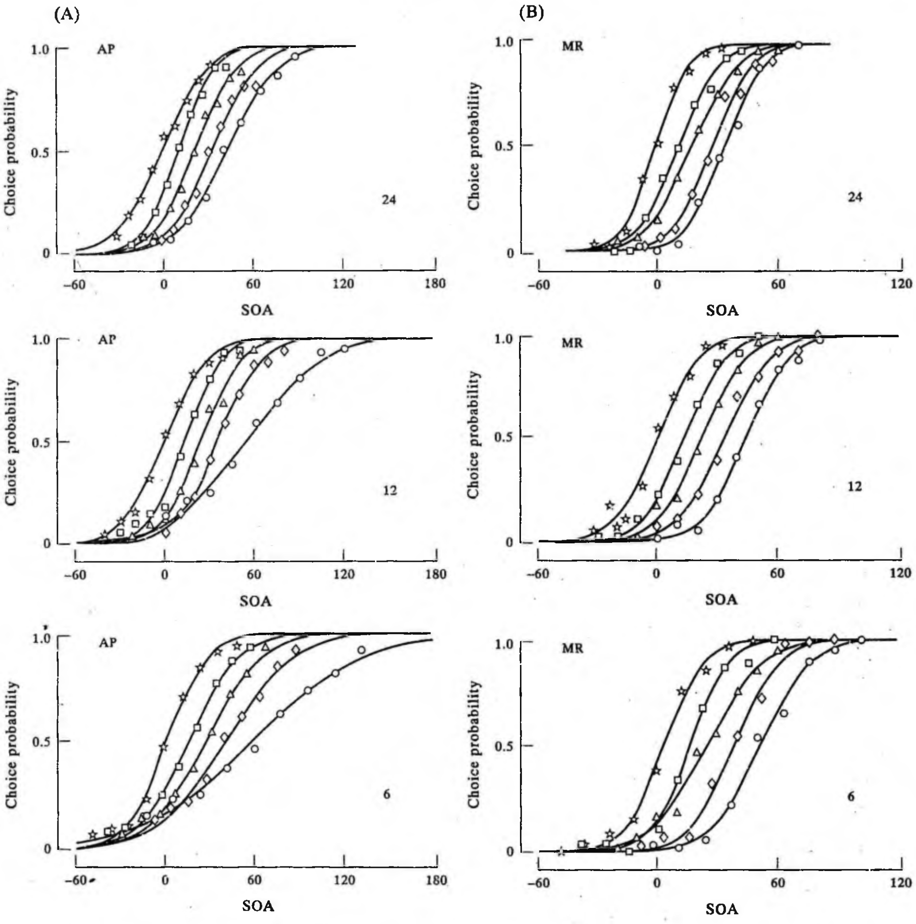


FIGURE 2. Choice probability of the positive motion direction as function of SOA (Δt) in msec, the amplitude, $A = 6$, 12, 24 cd/m^2, and the luminance onset rise-time duration, $D = 0$ (stars), 30 (squares), 60 (triangles), 120 (rombs) and 240 (circles) msec. Two observers: AP (A) and MR (B). Each data point is an average of 60 (MR) or 100 (AP) trials. The best-fitting psychometric functions are shown by continous curves.

probability. The goodness of fit was estimated by comparing the mean approximation error (the mean deviation of a data point from the best fitting psychometric function) with the mean standard error of data points. If the approximation error was smaller than or approximately equal to the mean standard error, the fit was regarded as acceptable.

Subjects

Two experienced psychophysical observers with normal (MR) or corrected myopic (AP) vision participated in the experiments. Viewing was binocular in a semidarkened room.

RESULTS AND ANALYSIS

The choice probabilities of two observers are plotted in Fig. 2(A) (AP) and (B) (MR), as a function of the delay Δt or the SOA in msec. Each panel corresponds to one test amplitude A and each set of data corresponds to one duration of the rise time D. Continuous curves are the best fits to the experimental data. A typical mean approximation error was about 3.3% (MR) and 2.8% (AP) which was smaller than an average standard error of means (4.3 and 3.9% respectively). The fit was satisfactory in all but two cases when the mean approximation error was slightly larger than the standard error of the mean. In these two deviating cases the mean approximation error was 4.81 vs 4.68% (observer MR,

$A = 12\,\text{cd/m}^2$, $D = 0$ msec) and 3.73 vs 3.65% (observer AP, $A = 12\,\text{cd/m}^2$, $D = 60$ msec).

The slope of the psychometric functions remained virtually independent of the rise time D and contrast A. Only on the longest ramp duration, $D = 240$ msec, a slight flattering of the psychometric functions of the observer AP can be noticed. The average JND was 11.8 and 19.3 msec for MR and AP respectively. Thus, both observers were able to discriminate reliably 10–20 msec time difference between two luminance onsets. These values are slightly larger than the highest accuracy that can be achieved when two lines are separated no more than 6 min arc and flashed for few milliseconds (Westheimer & McKee, 1977b; Burr, 1979; McKee & Taylor, 1984) but still in the range of realistic estimates (cf. Exner, 1876). Beside temporal waveform another possible reason of the discrepancy is the lower luminance of targets used in this study (cf. Westheimer & McKee, 1977a). It is suggested that unvarying sensitivity is an indication of the identical processing principles that are utilized on different occasions (Morgan, Hole & Glennerster, 1990). If it is so, the temporal order judgements are relatively invulnerable to the increase of the rise-time of the contrast onset up to 120 msec at least. It is interesting to notice that the precision of the temporal order discrimination of two tones generally follows Weber's law: the differential threshold of the temporal order increases in the proportion of the rise-time duration (Pastore, Harris & Kaplan, 1982; Smurzyński & Houtsma, 1989).

If the duration of the onset time D increased, the whole psychometric function shifted to the right: it was necessary to delay the beginning of the instantaneous reference luminance step to make it apparently simultaneous with the gradual test onset. The horizontal position of the psychometric function also changes with the change of the test stimulus amplitude A. Figure 3 shows more comprehensibly how PSE value, Δt^*, depends on the rise-time duration D and the amplitude A. If the amplitude of the test transition was smaller ($A = 6\,\text{cd/m}^2$) than that of the reference stimulus all PSE values became larger. On the other hand, if the amplitude of the test transition became two times higher ($A = 24\,\text{cd/m}^2$) than the reference transition, it was necessary to reduce the delay Δt^* that makes two luminance transients look simultaneous.

It has been known for more than a hundred years that energetically weaker signals need more time to be noticed than energetically stronger ones (Cattell, 1885; Piéron, 1922; Grice, 1968; Sanford, 1974; Rutchmann, 1966; Roufs, 1963, 1974; Dzhafarov, 1992). Many well-known perceptual phenomena like masking, metacontrast, Pulfrich and Hess effects can be explained by differences in perceptual latency of objects with different intensity (Exner, 1868; Hess, 1904; Pulfrich, 1922; Stigler, 1910). Starting from this well established general rule that the perceptual latency increases with the decrease of contrast it is possible to consider two possibilities: (1) the perceptual latency is *independent* of the temporal waveform $k(t)$ depending only on the total amplitude A; or (2) the perceptual latency *depends* not only on the amplitude A but on the waveform $k(t)$ of the luminance modulation as well. The independence hypothesis predicts that three curves portrayed in Fig. 3 form a parallel shifted family of curves. This means, in particular, that all differences between curves remain constant at all rise-time values. In fact, even a casual inspection reveals that the curves corresponding to different amplitudes fan out with the increase of the rise-time duration. For example, the latency difference was about 2–4 msec when both transitions were instantaneous and about 6–12 msec when the onset lasted 240 msec. Thus, our data provide strong support for the concept that the perceptual latency **depends** indeed on the stimulus waveform $k(t)$. This result can be regarded as more expected than erratic. So far it was well documented that the perceptual latency depends on the spatial frequency (Tartaglione, Goff & Benton, 1975; Lupp, Hauske & Wolf, 1976; Vassilev & Mitov, 1976; Ejima & Ohtani, 1987) increasing by an average 21 msec over a range of 0.5–9 c/deg or approx. 5 msec per octave (Parker & Dutch, 1987). Therefore it is not surprising that the perceptual latency increases with the decrease of the temporal gradient of the contrast change.

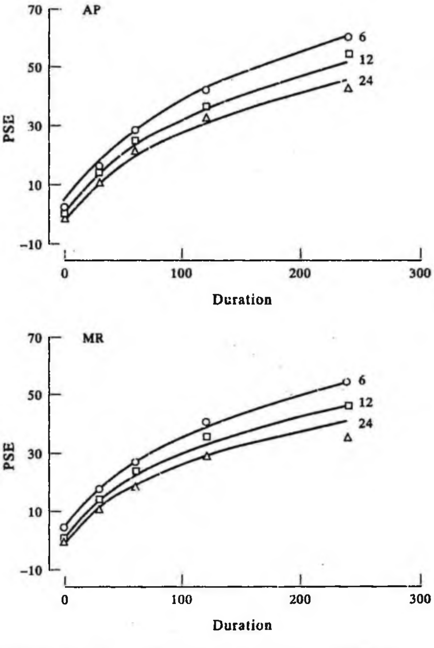


FIGURE 3. PSE as a function of the luminance onset rise-time duration D in msec and the amplitude $A = 6$, 12, and $24\,\text{cd/m}^2$ for the observers MR and AP. PSE is the value of SOA at which the psychometric function is equal to 0.5. The best fitting functions shown by continous curves are explained in the text.

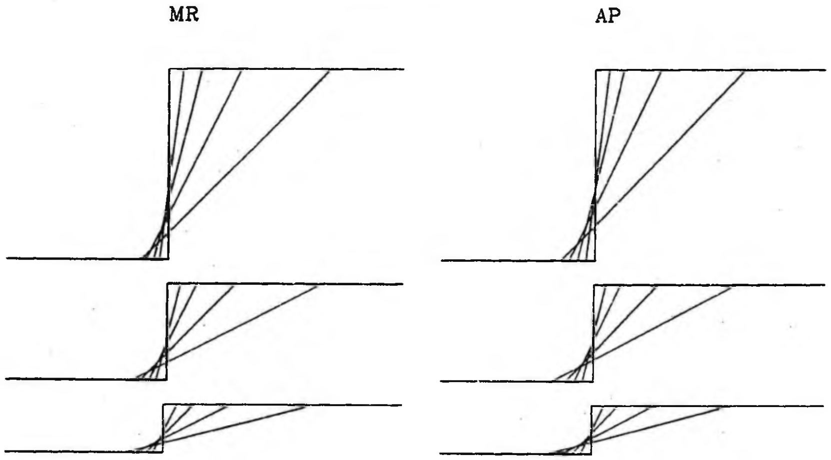


FIGURE 4. Perceptually isochronous waveforms for MR and AP. Each waveform is shifted by the respective PSE value shown in Fig. 3 and then all waveforms with the same amplitude are superimposed on one another.

Another comprehensive way of presenting data is to compute and portray perceptually equivalent or isochronous waveforms (see Fig. 4).The procedure of determining isochronous waveforms was the following. Each waveform was shifted by the respective PSE value, Δt^*, and then superimposed on one another. Since all linear ramps needed for their apparent simultaneity to start before the instantaneous step-function they intersected with it forming an onset triangle which base is Δt^* and the height is equal to $A^* = A \cdot \Delta t^*/D$. The height A^* indicates the amplitude of the luminance increase at which a linear luminance increase becomes subjectively isochronous with the step function. For brevity we call it the threshold contrast. Analysing these onset triangles and Fig. 3 it is possible to formulate three empirical rules.

(1) The temporal shift Δt^* that was required to equalize a gradual onset with the step-function is a compressive function of the rise-time duration D which can be in the first approximation described by a power function with an exponent of about $\frac{1}{2}$ (approximate square root rule).

(2) The threshold amplitude A^* is a decreasing function of the rise-time duration D and the increasing function of the amplitude A. A conspicuous property of all patterns portrayed in Fig. 4 is that as the rise-time duration of the test signal increases earlier it intersects with the step-function. The threshold contrast increases also with the final steady-state amplitude A. For example, 240 msec luminance ramp with 6 cd/m^2 amplitude became noticeable at about 1.3–1.5 cd/m^2 whereas larger transition with the final 24 cd/m^2 steady-state amplitude became visible at about three times higher amplitude (3.3–4.3 cd/m^2).

(3) In the first approximation, the threshold energy (area of the onset triangles) of waveforms with the same amplitude is about the same but increasing about proportionally with the increase of the final steady-state amplitude.

These empirical rules considerably restrain the range of plausible explanations. For example, three most simple explanations shown in Fig. 5 clearly contradict to at least one of these empirically observed regularities.

(1) *Cross-correlation.* According to the Reichardt motion detection scheme the motion direction is determined on the basis of the two time-varying signals X and Y recorded from two adjacent locations. These two signals are delayed on a fixed amount of time and each of the delayed temporal patterns X' and $Y'$are multiplied with the original nondelayed signal recorded from another spatial location. The difference between these two time averaged products, $\bar{X}'\bar{Y}$ and $\bar{X}\bar{Y}'$, determines the direction of recorded motion. Consequently, such a delay Δt must exist for which the products $\bar{X}'\bar{Y}$ and $\bar{X}\bar{Y}'$ are equal yielding to the apparent simultaneity of two luminance transitions. The basic behaviour of the Reichardt detectors does not change when more temporal filters are added: the output response will remain proportional to the cross-correlation' product of the input signals. In Fig. 5(A) the equivalent luminance waveforms predicted on the basis of the cross-correlation products of the gradual luminance increase and the step-function are shown. The predicted pattern has a certain resemblance with empirically determined

patterns shown in Fig. 4: (i) the predicted temporal shift Δt^* is approximately the square-root function of the rise-time duration, and (ii) the predicted threshold amplitude A^* increases proportionally with the increase of the final steady-state amplitude A. There is, however, one crucial discrepancy between the prediction and empirical data which makes the cross-correlation an unlikely candidate for the satisfactory explanation. On the basis of cross-correlation products alone it is impossible to predict the amplitude dependent shortening of perceptual latency which is clearly visible on Fig. 3.

(2) *Threshold amplitude*. Seeing that the cross-products cannot explain amplitude dependent perceptual delay other variants of temporal order determination can be considered. Suppose that the temporal order of two luminance transitions is determined on the basis of time moments when they become individually visible. The luminance change is posited to be noticed as soon as the time-varying luminance function $L(t)$ reaches some critical amplitude A^* and the temporal sequence in which this threshold level is reached defines the perceived motion direction: the displacement is perceived towards the position where the luminance change was detected later [see Fig. 5(B)]. According to this threshold amplitude scheme all equivalent luminance waveforms pass through one single point which clearly contradicts the actual pattern of data.

(3) *Threshold energy*. The threshold model can be also formulated in terms of energy. According to the threshold energy model each signal is noticed when the luminous energy contained in the transition exceeds a certain value W. The predicted equivalent luminance waveforms shown in Fig. 5(C) produce a set of the equal-area triangles (except for very short rise-times not shown on the figure). This model nicely predicts a characteristic hyperbolic arrangement of the onset triangles but fails to describe the threshold energy dependence from the final steady-state amplitude. Two ramps with exactly the same luminance increase rate became visible at two completely different time moments. For example, the same luminance increase rate of 200 cd/m^2 per sec can be obtained when the 6 cd/m^2 final amplitude is reached during 30 msec rise-time and when the 24 cd/m^2 final amplitude is reached during 120 msec rise-time. PSEs for these two cases are very different: 16.6 vs 27.1 msec for MR and 16.2 vs 32.7 for AP respectively. This approximately twice difference in the onset times corresponds to about four times difference in the amount of luminous energy required to notice stimulus appearance. Thus, the amount of minimum energy the observer needed to detect an onset of a visual stimulus depends upon the final amplitude of the luminance increase.

Thus, none of the three simple models is able accurately to predict the empirical data. One probable reason for that failure is the limited use of available information by the observer. From the entire stimulus energy only a certain fraction is effectively employed. For example, it is unlikely that the visual system is able to follow very rapid luminance changes. The approximate square-root dependence of PSEs on the rise-time duration may indicate that the visual system integrates changes in contrast over a considerable time period. Let us suppose, then, that the observer is deciding the appearance of the luminance onset not on the basis of the whole luminous energy but on that part of it which is concentrated on low temporal frequencies. The

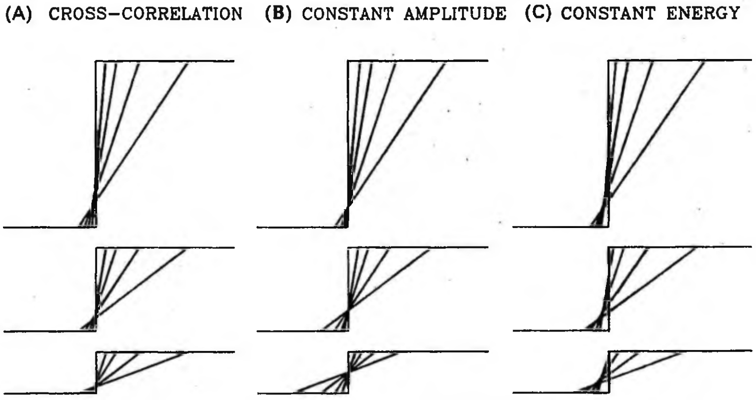


FIGURE 5. Predicted isochronous waveforms of three most simple explanations based on the cross-correlation, constant amplitude and constant energy models.

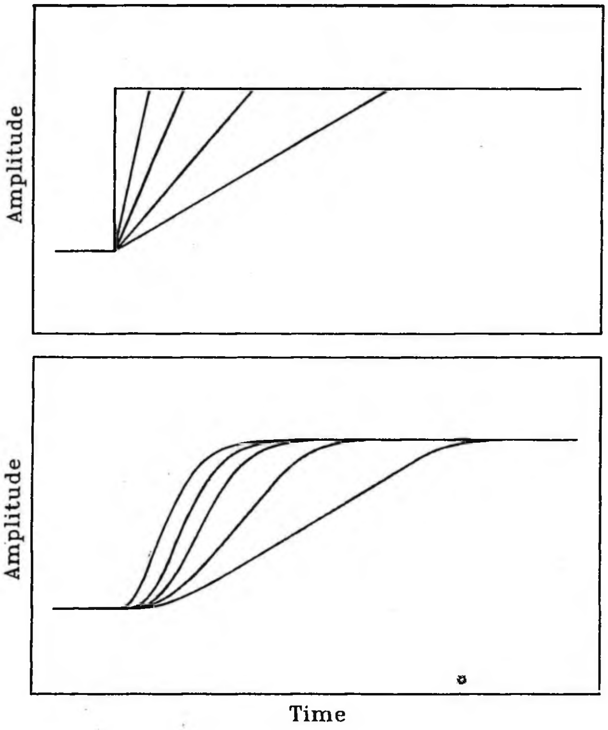


FIGURE 6. The effect of low-pass filtering operation on the luminance transitions used in this study. The fourth-stage ($n = 4$) linear filter with time constant $\tau = 20$ msec was applied to onset ramps. According to the proposed explanation each luminance transition is noticed as soon its transformed function reaches a fixed threshold value, the perceived temporal order corresponds to the sequence in which this amplitude level is reached.

impulse response $h(t)$ of an n-stage low-pass filter with time constant τ is

$$h(t) = L(t)\frac{\tau^{n-1}}{(n-1)!}(t/\tau)^{n-1}e^{-t/\tau}. \qquad (2)$$

For example, if the luminance onsets used in this study passed through such a filter ($n = 4$ and $\tau = 20$ msec) the result would be as shown in Fig. 6. The low-pass filter integrates the input signal but leaks at a rate proportional to the amount of accumulation. After removing high frequency components from the step-function it became more similar to ramps with gradual increase of luminance. The simplest link between these low-pass filtered signals and the observer's judgements about temporal order is of course a threshold mechanism. The only required modification in the above formulated threshold amplitude model is the substitution of the time-varying luminance function $L(t)$ with its low-pass filtered version. In other words, each luminance excursion is noticed as soon as its low-pass filtered transformation exceeds a threshold value, the order of reaching threshold value determines the perceived movement direction.

A satisfactory fit to data was achieved by using the fourth-order low-pass filter ($n = 4$) with time constants $\tau = 32$ msec (AP) and $\tau = 34$ msec (MR). The amplitude of the transformed signal was posited to be proportional to the logarithm of the input signal amplitude. The best approximation of PSEs is shown by continuous curves in Fig. 3. The mean approximation error was 1.6 and 1.8 msec for AP and MR respectively. The correlation between the observed and predicted data was 0.994 and 0.996 for MR and AP respectively which left < 1.2 and 0.8% of total variation unexplained. It is certainly not a coincidence that these low-pass filter parameter values are typical of the human impulse response function obtained with point-like stimuli on a relatively low background (cf. Roufs & Blommaert, 1981).

DISCUSSION

It is very satisfying that such a simple stimulus attribute, the luminous energy concentrated on low temporal frequencies, can be used for the prediction of the perceived order of two luminance onsets. According to this explanation, from all the energy contained in the intensity onset only a relatively small fraction is effectively used. We assumed that the change in luminance at any spatial location cannot be noticed until the amplitude of the low-pass filtered signal reaches a given threshold value. The displacement direction is judged by comparing luminance changes at two spatial locations: the motion is judged in the direction of that location in which the change in luminance was detected later in time. So far we have carefully avoided mentioning any internal processes although it is quite obvious how to reformulate the whole exposition in the terminology of internal representation. Indeed, the insensitivity to high temporal frequencies means that the visual system is lazy when rapid changes occur in illumination. The internal signal resulting from a sudden intensity step rises slowly before it reaches peak value. This means, in particular, that there is always some irreducible time period between the beginning of the stimulus and the moment when the elicited response exceeds a given amplitude level. At a greater intensity step the amplitude of the elicited response is proportionally greater and a preset amplitude value is reached earlier. Thus, if the amplitude threshold of visibility exists a perceptual lag must exist whose duration reduces with the increase of the stimulus amplitude (Roufs, 1974). It also explains why, for example, two uniformly moving lines placed one above the other appear to move in different phases if their intensity differs sufficiently (Hess, 1904). Quantitative parameters of temporal sensitivity and visual latency are usually estimated either from flicker or flash thresholds experiments. As we already noticed, the impulse response to a point-like stimulus on a relatively low background can be described by the fourth-stage linear filter with the time constant of about 20–40 msec (Roufs, 1974; Roufs & Blommaert, 1981). The remarkable agreement between these estimates and the explanation proposed in this study leads to an important theoretical convergence: the judgement of temporal order of two adjacent luminance excursions can be explained by exactly the same fundamental mechanisms that underly the detection of flicker and flashes. The temporal order judgements reported here possess little or nothing which cannot be explained in terms of detection of two isolated contrast changes at two different retinal locations.

This study provides some support for an approach in which the detection of motion is a special case of contrast discrimination. Therefore it is not surprising that our data have a remarkable resemblance with temporal masking experiments. In the masking paradigm the contrast threshold of a probe impulse is measured in the presence of a masking stimulus presented before, during, or after the probe. Sperling established some basic rules characterizing the impulse masking (Sperling, 1965): (i) the probe impulse–contrast threshold is proportional to the energy of the masking stimulus (Weber's law); (ii) masking flash causes a very considerable change in threshold for tests not only presented during the masking flash but also preceding it; (iii) only the first portion of a long masking flash has an impact upon the test contrast thresholds. These rules have implication to our data assuming that the gradual onset can be considered consisting of two parts: the initial part constitutes a target and the subsequent light from the onset may be regarded as a background which appears after the target. The minimum energy that is needed to detect the test stimulus is proportional to the masking energy: four time increase of the masking energy, for example, leads to about four time increase of the required test energy. In other words, the present data can be easily reformulated in the language of temporal masking. This reformulation leads, however, to an apparently paradoxical result: the property of an event (minimum luminance energy required to see the beginning of the onset) occurring earlier in time depends on the property (final steady state amplitude) of another event occurring later in time. However, it is well known that the minimum energy that is needed to detect a visual stimulus depends upon what else has been, is, and will be present in the visual field. More generally speaking, the information that specifies a property localized in time need not be localized in the same way as the property ["Localization Fallacy" of Hilbert (1987, p. 67)]. Indeed, after the integration every momentary value is determined on the basis of information distributed over a considerable time interval. Thus, the apparent backwards referral in time may result from an elementary low-pass filtering operation.

The idea of regarding motion perception as a special case of contrast discrimination is not a new one. For example, Morgan and Cleary (1992) adopted this theoretical position in the study of the contrast effects on two-frame random kinematograms. They found that when the contrast of both frames is too low to permit directional discrimination, increasing the contrast of either the first or the second frame alone makes directional discrimination possible. The direction discrimination first improves with the increase of the relative contrast, and then, at the higher contrasts, collapses again. This near-threshold "Pedestal Effect" is well known from luminance discrimination experiments (Leshowitz, Taub & Raab, 1968; Whittle, 1986) and there seems to be no need to invoke something extra to explain the similar effect in the motion discrimination task. Morgan and Cleary claim, however, that the existence of the pedestal effect in the direction discrimination strongly supports models of the Reichardt type although its non-monotonicity clearly violates the multiplicative law of amplitudes on which the Reichardt model is based. Another indisputable violation of the multiplicative law reported in their study is a tendency that motion is harder to detect in a high–low contrast sequence than in a low–high contrast sequence (Morgan & Cleary, 1992). Thus, two descriptions, in terms of the near-threshold contrast discrimination and the multiplicative model, are not necessarily compatible.

The present findings have some implications for theories of motion analysis. Motion perception appears to be a unique area of perceptual studies where instead of a large variety of equally plausible models only few basic explanations exist. This remarkable convergence was primarily achieved by theoretical works demonstrating that the spatiotemporal energy models (Adelson & Bergen, 1985; Burr, Ross & Morrone, 1986) and linear motion sensor explanation proposed by Watson and Ahumada (1985) are fully or virtually equivalent to the elaborated Reichardt model (van Santen & Sperling, 1984, 1985). Thus, the correlation model is now regarded as almost a universal solution of the motion detection problem (Poggio & Reichardt, 1973; Reichardt, 1987). If the visual task examined in this study implied involvement of elementary motion encoders, our results could seriously challenge the universality of the correlation model. The multiplication operation, the essential part of all correlation schemes, cannot resolve changes exercised separately by each multiplied component. According to commutability, it is indifferent which of two amplitudes is increased or diminished if the product of amplitudes remains the same. The relative perceptual delay, however, changes its sign dependently on which of the two amplitudes is modified. It is possible, as proponents of the Reichardt model usually do, to suppose that *before* subjecting input signals to ordinary correlational analysis, some transformation is applied to the visual input signals. The assumption of a transformation of unknown complexity before the scope of the model makes its falsification very problematic. In fact, a likely candidate for that transformation was proposed in this study. We found, however, that after such a transformation (low-pass filtering) there is no more need for additional computations: the judgement of the displacement direction can be predicted from observing a very simple property of these transformed time-varying luminance functions themselves. The sequence in which these single attributes, reaching some amplitude level, occurred in time predicts the perceived displacement of the composite stimulus. Probably the data of the present study are also compatible with some elaborated variant of the Reichardt model. But even in that case the present explanation, using exclusively terms of the contrast detection, remains more simple and economic.

Another popular motion detection scheme, the gradient model, appears to face some problems when it is applied to the present visual task (Marr & Ullman, 1981; for a review see Hildreth & Koch, 1987). The gradient model is based on determination of the time derivative of the intensity at locations where the Laplacian of the image is zero. The time derivative remains the same during the whole period of the onset rise and it is problematic how the precise timing of two adjacent onsets can be performed. The indefinite does not mean, of course, impossible. We cannot exclude a possibility that after some elaboration the gradient scheme, analogously to the correlation scheme, can account for visual tasks analogous to those studied in this report but we seriously doubt if this description will be more simple and conceptually transparent.

REFERENCES

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

Allan, L. G. (1975). Temporal order psychometric functions based on confidence rating data. *Perception & Psychophysics, 18,* 523–534.

Burr, D. (1979). Acuity for apparent vernier offset. *Vision Research, 19,* 835–837.

Burr, D. C., Ross, J. & Morrone, M. C. (1986). Seeing objects inmotion. *Proceedings of the Royal Society of London B, 227,* 249–265.

Cattell, J. McK. (1885). The influence of the intensity of the stimulus on the length of the reaction time. *Brain, 8,* 511–515.

Dzhafarov, E. N. (1992). The structure of simple reaction time to step-function signals. *Journal of Mathematical Psychology, 36,* 235–268.

Efron, R. (1973). Conservation of temporal information by perceptual system. *Perception & Psychophysics, 14,* 518–530.

Ejima, Y. & Ohtani, Y. (1987). Simple reaction time to sinusoidal grating and perceptual integration time: Contributions of perceptual and response processes. *Vision Research, 27,* 269–276.

Exner, S. (1868). Über die zu einer Gesichtswahrnehmung nöthige Zeit. *Sitzungsberichte der Kaiserlichen Akademie der Wissenschaften, 58,* 601–632.

Exner, S, (1875). Experimentelle Untersuchung der einfachsten psychischen Processe, III. *Archiv für die gesammte Physiologie des Menschen und der Thiere, 11,* 403–432.

Exner, S. (1876). Über das Sehen von Bewegungen und die Theorie des zusammengesetzten Auges. *Sitzungsberichte der Kaiserlichen Akademie der Wissenschaften, 72,* 156–190.

Grice, G. R. (1968). Stimulus intensity and reponse evocation. *Psychological Review, 75,* 359–373.

Hess, C. (1904). Untersuchungen über den Erregungsvorgang im Sehorgan bei kurz- und bei längerdauernder Reizung. *Archiv für die gesammte Physiologie des Menschen und der Thiere, 101,* 226–262.

Hilbert, D. R. (1987). *Colour and colour perception. A study in anthropocentric realism.* Stanford, Calif.: Stanford University Press.

Hildreth, E. C. & Koch, C. (1987). The analysis of motion: From computational theory to neural mechanisms. *Annual Review of Neuroscience, 10,* 477–533.

Hirsh, I. J. & Sherrick, C. E. Jr (1961). Perceived order in different sense modalities. *Journal of Experimental Psychology, 62,* 423–432.

Leshowitz, B., Taub, H. B. & Raab, D. H. (1968). Visual detection of signals in the presence of continous and pulsed backgrounds. *Perception & Psychophysics, 4,* 207–213.

Lupp, U., Hauske, G. & Wolf, W. (1976). Perceptual latencies to sinusoidal gratings. *Vision Research, 16,* 969–972.

Marr, D. & Ullman, S. (1981). Directional selectivity and its use in early visual processing. *Proceedings of the Royal Society of London B, 211,* 151–180.

McKee, S. P, (1981). A local mechanism for differential velocity detection. *Vision Research, 21,* 491–500.

McKee, S. P. & Taylor, D. G. (1984). Discrimination of time: Comparison of foveal and peripheral sensitivity. *Journal of the Optical Society of America, A, 1,* 620–627.

Morgan, M. J. & Cleary, R. (1992). The effect of contrast substitutions upon motion detection in spatially random patterns. *Vision Research, 32,* 639–643.

Morgan, M. J., Hole, G. J. & Glennerster, A. (1990). Biases and sensitivities in geometrical illusions. *Vision Research, 30,* 1793–1810.

Parker, D. M. & Dutch, S. (1987). Perceptual latency and spatial frequency. *Vision Research, 22,* 1279–1283.

Pastore R. E., Harris, L. B. & Kaplan, J. K. (1982). Temporal order identification: Some parameter dependencies. *Journal of the Acoustical Society of America, 71,* 430–436.

Piéron, H. (1922). Nouvelles recherhes sur l'analyse du temps de latence sensorielle et sur la loi qui relie ce temps a l'intensité de l'excitation. *L'Année Psychologique*, *22*, 58–142.

Poggio, T. & Reichardt, W. (1973). Considerations on models of movement detection. *Kybernetik*, *13*, 223–227.

Pulfrich, C (1922). Die Stereoskopie im Dienste der isochromen und heterochromen Photometrie. *Die Naturwissenschaften*, *10*, 553–564.

Reichardt, W. (1957). Autokorrelationsauswertung als Funktionsprinzip des Zentralnervensystems. *Zeitschrift für Naturforschung*, *B12*, 448–457.

Reichardt, W. (1987) Evaluation of optical motion information by movement detectors. *Journal of Comparative Physiology*, *161A*, 533–547.

Roufs, J. A. J. (1963). Perception lag as a function of stimulus luminance. *Vision Research*, *3*, 81–91.

Roufs, J. A. J. (1974). Dynamic properties of vision—V. Perception lag and reaction time in relation to flicker and flash thresholds. *Vision Research*, *14*, 853–869.

Roufs, J. A. J. & Blommaert, F. J. J. (1981). Temporal impulse and step responses of the human eye obtained psychophysically by means of a drift-correcting perturbation technique. *Vision Research*, *21*, 1203–1221.

Rutchmann, R. (1966). Perception of temporal order and relative visual latency. *Science*, *152*, 1099–1101.

Sanford, A. J. (1974). Effects of changes in intensity of white noise on simultaneity judgements and simple reaction time. *Quarterly Journal of Experimental Psychology*, *23*, 296–303.

van Santen, J. P. H. & Sperling, G. (1984). Temporal covariance model of human motion perception. *Journal of the Optical Society of America A*, *1*, 451–473.

van Santen, J. P. H. & Sperling, G. (1985). Elaborated Reichardt detectors. *Journal of the Optical Society of America A*, *2*, 300–321.

Smurzyński, J. & Houtsma, A. J. M. (1989). Auditory discrimination of tone-pulse onsets. *Perception & Psychophysics*, *45*, 2–9.

Sperling, G. (1965). Temporal and spatial visual masking—I. Masking by impulse flashes. *Journal of the Optical Society of America*, *55*, 541–559.

Stigler, R. (1910). Chronophotische Studien über den Umgebungskontrast. *Pflüger's Archiv für die gesammte Physiologie des Menschen und der Tiere*, *134*, 365–435.

Sweet, A. L. (1953). Temporal discrimination by the human eye. *American Journal of Psychology*, *66*, 185–198.

Tartaglione, A., Goff, D. P. & Benton, A. L. (1975). Reaction time to square-wave gratings as a function of spatial frequency. *Brain Research*, *100*, 111–120.

Thorson, J., Lange, G. D. & Biederman-Thorson, M. (1969). Objective measure of the dynamics of a visual movement illusion. *Science*, *164*, 231–248.

Vassilev, A. & Mitov, D. (1976). Perception time and spatial frequency. *Vision Research*, *16*, 89–92.

Watson, A. & Ahumada, A. J. (1985). Model of human visual-motion sensing. *Journal of the Optical Society of America A*, *2*, 322–342.

Westheimer, G. & McKee, S. P. (1977a). Integration region for visual hyperacuity. *Vision Research*, *17*, 89–93.

Westheimer, G. & McKee, S. P. (1977b). Perception of temporal order in adjacent visual stimuli. *Vision Research*, *17*, 887–892.

Whittle, P. (1986). Increments and decrements: Luminance discrimination. *Vision Research*, *26*, 1677–1691.

IV

Perception and Psychophysics, 1995, 57, 27–34.

Perception & Psychophysics
1995, 57 (1), 27-34

Magnitude of luminance modulation specifies amplitude of perceived movement

JÜRI ALLIK and ALEKSANDER PULVER
University of Tartu, Tartu, Estonia

A compelling impression of movement, which is perceptually indistinguishable from a real displacement, can be elicited by patterns containing no spatially displaced elements. An apparent oscillation, w-movement, was generated by a stationary pattern containing a large number of horizontal pairs of spatially adjacent dots modulated in brightness. The observer's task was to adjust the perceived amplitude of the w-motion to match the amplitude of a real oscillation. All of the data can be accounted for by a simple rule: If the relative change in the luminance, $W = \Delta L/L$, between two adjacent stationary dots is kept constant, the distance over which these dots appeared to travel in space comprises a fixed fraction of the total distance by which they are separated. The apparent amplitude of the w-motion increases strictly in proportion with luminance contrast, provided that the contrast is represented in the motion-encoding system by a rapidly saturating compressive Weibull transformation. These findings can be explained in terms of bilocal motion encoders comparing two luminance modulations occurring at two different locations.

It is somewhat astonishing that when Wertheimer's (1912) famous paper on ϕ-movement was published, the fact that a vivid impression of motion can be produced by a sequence of stationary stimuli was widely known. Simple toy stroboscopes were available in stores, and Wertheimer had no difficulty purchasing one after his sudden decision to leave a train in Frankfurt 2 years earlier. But he probably was the first to realize that ϕ-motion violates the layman's concept of motion. According to this concept, movement is an intrinsic property of an object, and encountering a situation in which a clear impression of motion is elicited without that property must come as a big surprise. For the physicist, however, motion appears to be a quality attributed to an object by an observer: The object can be decided to be in motion only if it is observed at two different instants and it is seen to be in two different positions at those two instants. Therefore, ϕ-motion may simply indicate that the movement experience requires a perceptible change in the position of one stimulus element with respect to another.

However, the displacement of some stimulus elements with respect to others cannot be regarded as a *necessary* condition for perception of movement. A distinct impression of movement can be elicited by patterns containing no spatially displaced elements. The perceived movement can be evoked by changes of light flux at different retinal locations. Johansson (1950, 1978) described the "wandering motion" seen between two or more spatially adjacent bright objects modulated in brightness (w-motion). What is particular to this and other analogous visual demonstrations (Anstis, 1967, 1986, 1990; Bülthoff & Götz, 1979; Gregory & Heard, 1983; Mastebroek & Zaagman, 1988; Mather, 1984) is that the perceived movement is generated by stimuli in which the elements do not change their relative spatial position and usually remain continuously visible. These findings are surprising only if the detection of motion is ultimately regarded as a matching process comparing two spatial luminance patterns at two instants in time. Most current theories of movement perception, on the contrary, regard motion as comparing two luminance modulations that occur at two different locations (Reichardt, 1957, 1987; van Santen & Sperling, 1984, 1985). Consequently, w-motion suggests that the appropriate stimulus for motion is a relative change in light flux at two spatial locations—not the spatial displacement tracking of some visual elements after they have been individually recognized.

It is impossible to distinguish an object moving in discrete jumps from a continuously moving object, provided that the time between jumps and their amplitude is not too great (Burr, Ross, & Morrone, 1986b; Morgan, 1979, 1980; Watson, Ahumada, & Farrell, 1986). The sequence of discrete jumps that occurs at rates greater than about 30 Hz is indistinguishable from smooth continuous motion because both provide the same effective stimulus to the visual system. In the present study, we present evidence that w-motion can be perceptually indistinguishable from real displacements. This means that despite their physical difference, w-motion and ϕ-motion are metameric, and they both rely on an identical underlying mechanism in the nervous system. Many

---

This research was supported by a James S. McDonnell Foundation grant and Estonian Science Foundation Grant 450 to J.A. We would like to thank Mary Hayhoe and three anonymous referees for their comments on an earlier draft of this paper. Address correspondence to J. Allik, Department of Psychology, University of Tartu, 78 Tiigi Street, Tartu EE-2400, Estonia (e-mail: allik@psych.ut.ee).

current theories of motion perception assume that the visual system employs motion-encoding units with receptive fields extended over space and time that are tuned to movement along a particular trajectory (Adelson & Bergen, 1985; Burr, Ross, & Morrone, 1986a; van Santen & Sperling, 1984, 1985; Watson & Ahumada, 1985; Wilson, 1985). These units, measuring the amount of luminous energy in some spatiotemporal volume, are indifferent to whether this change in the luminance flux is produced by a moving object traveling from one location to another, or by the luminance modulation of two stationary objects at these two locations.

METHOD

Subjects. Two observers, M.R. (female) and A.P. (male; one of the authors), participated. One of the subjects was naive, having no knowledge of the way visual motion was generated in the experiment.

Procedure. The observer decided whether two display areas, the central part and its surround, showed identical movement. The impression of movement in the central part was generated by luminance modulation of stationary patterns (w-movement). The perceived movement of the surround was produced by spatial displacement of the elements—that is, by their stroboscopic displacement, or ϕ-movement. Thus, the observer's task was to adjust the *perceived amplitude* of periodic oscillation of a stationary

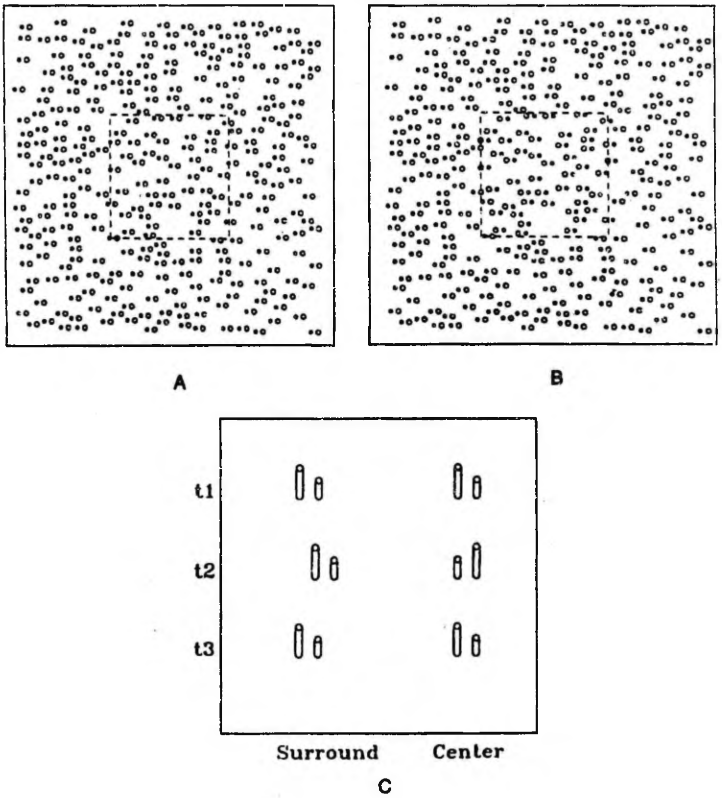


Figure 1. An illustration of one stimulus composed of two patterns—the original (A) and its slightly modified replica (B)—exposed in an alternation rate of 3.3 Hz. Each pattern consisted of a large number of horizontal pairs of dots (dipoles); one was dark with a fixed luminance L (small circles), and the second one was light with adjusted luminance $L + \Delta L$ (large circles). In the central area (dashed rectangles), movement was produced by luminance modulation; all the dark dots became light, and, in turn, all the light dots became dark. Surround movement was elicited by a real displacement of all dipoles without exchange of position between light and dark members within a dipole. (C) A magnified picture of two dipoles in the surround and central area from three subsequent frames—t1, t2, and t3. The height of the cylinders represents luminance.

pattern to the amplitude of *real* oscillation. In this experiment, stimuli were composed of 1,500 micropatterns distributed randomly within a rectangular area that, viewed from 250 cm, had a size of about 5.3° × 3.5° (see Figure 1). The central area (indicated by a dashed rectangle), within which the movement was produced by luminance modulation, was approximately 2.75° wide and 2.07° high. Each micropattern consisted of a horizontal pair of dots (dipoles), separated from each other by a spatial distance, *d*. The dot size was 1 pixel, or about 0.0084° of arc (about half of a minute). Special care was taken to avoid overlap between micropatterns by applying a rule prohibiting any two micropatterns from being closer to each other than 5′. One of the two dots in each dipole had a fixed luminance, L (dark dots), and the second one had a variable contrast, $L + \Delta L$ (light dots), which could be adjusted by the observer. The dark dots served as a standard, and the light ones served as a test. The motion stimulus was generated by endless cycling of a given stimulus pattern and its slightly modified replica. These two patterns, the original one and its slightly modified duplicate, were presented in alternation at the rate of 3.3 Hz on the screen of an Amstrad color monitor. Thus, each pattern remained visible for 300 msec and was thereafter instantaneously replaced between two frames with the second pattern.

In the central part of the display, all the dipoles remained stationary; only dark (with a fixed luminance, L) and light (with a variable luminance, $L + \Delta L$) dots exchanged their spatial positions. In the first and every subsequent odd frame, all the left members of the dipoles were dark, and all the right members were light. In the second and every subsequent even frame, the left element became light and the right element became dark. If the luminance difference ΔL between the two types of dots was small, no motion of the central area could be seen. Above a certain luminance increase, however, the coherent horizontal oscillation of the whole central area began. Shortly, a luminance increment, ΔL, was alternatively added to the left and the right dots, which produced cyclical w-motion of the central portion of the display. With the increase of the luminance modulation, ΔL, the perceived amplitude of oscillation increases. In the surround area, there was no exchange of positions between dark and light elements of dipoles; their relative spatial positions remained the same. Instead, all the dipoles were uniformly displaced by a distance, *s*, to the right in the second and every subsequent even frame, and back to the left on the third and every subsequent odd frame. This displacement produced a coherent to and fro ϕ-motion of the surround area. In most cases, it was phenomenologically difficult, if not impossible, to tell whether the motion was induced by luminance modulation or by real displacement, provided that the perceived amplitudes of both movements were equalized.

The observer was instructed to adjust the luminance increment ΔL until the movement of the central part of the display appeared to be identical to that of the surround area. The adjustment procedure was as follows. The luminance of the two types of dots, dark and light, were tuned to be equal, and the observer started to increase, by revolving a multirevolution knob, luminance increment ΔL, added to all the light dots. After reaching the luminance value that was necessary for equalizing apparent movement in the central and surround areas, the trial was stopped, and the ΔL value

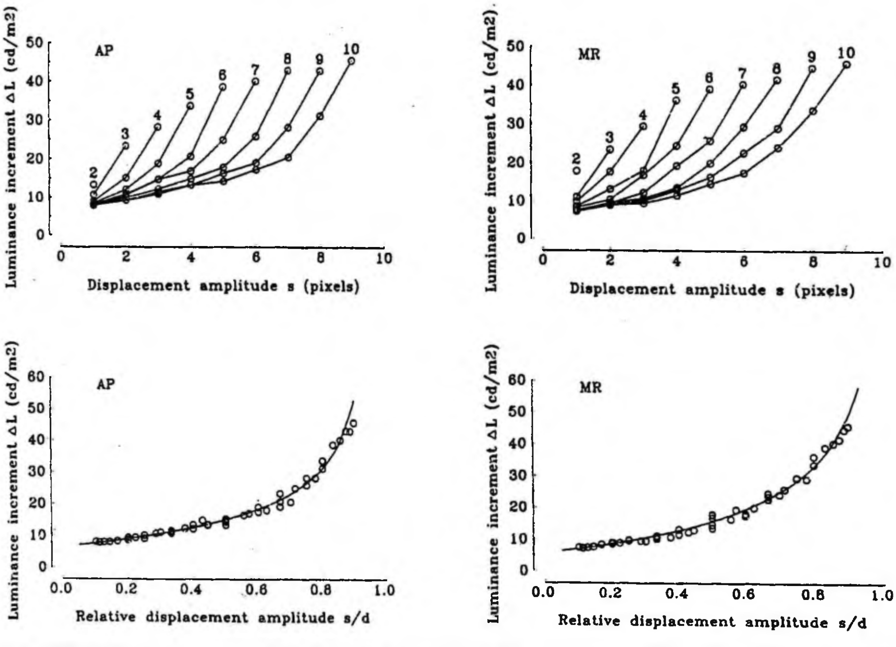


Figure 2. Upper panels: The luminance increment ΔL (cd/m^2) required to make the apparent oscillation of the central part perceptually indistinguishable from the surround oscillation with the displacement amplitude *s* (in pixels), for Subjects A.P. (left panel) and M.R. (right panel), for nine different interdot separations (2, 3, . . . , 10), *d*. Lower panels: The same data replotted as the function of the relative displacement amplitude *s*/*d* (in proportion to interdot separation).

was stored. Although the adjustment time was not limited, usually it took only 5–6 sec to reach a satisfactory ΔL value.

There were two different experiments. In each experimental session, one of the fixed reference luminance values, L, was selected. There were one ($L = 6$ cd/m^2) and three ($L = 3, 6, 12$ cd/m^2) different referent luminance values in the first and second experiments, respectively. Before each trial in the first experiment, one of the interelement separations, d, was selected from nine interelement separations ($d = 2, 3, \ldots, 10$ pixels). In the second experiment, the interdot separation was $d = 6$. Before each trial in both experiments, one of the displacement values (s) was randomly selected. Since the perceived amplitude of w-motion never exceeded interelement separation d, the amplitude of ϕ-motion s was always smaller than d.

In both experiments viewing was binocular, without head fixation, in a semidarkened room. The adjustment was repeated at various combinations of L, d, and s for 5 (M.R.) or 10 (A.P.) times.

RESULTS

Figure 2 (upper panels) shows the luminance increment ΔL required to make the luminance-modulated w-motion perceptually equivalent to the surround movement produced by a given displacement s, for Subjects A.P. (left panel) and M.R. (right panel), for nine different interdot separations, d. The reference luminance was $L = 6$ cd/m^2. Each set of data, corresponding to a given interdot separation d, formed a function clearly distinct from other functions. Two empirical rules can be noticed in these data:

1. The luminance increment ΔL that was required to equalize w-motion in the central area with a real displacement in the surround area increased monotonically with the increase of the stroboscopic displacement amplitude s. This means, in particular, that even when the spatial separation between the luminance-modulated dots remained the same, the perceived amplitude of w-motion increased with the luminance modulation amplitude ΔL.

2. The luminance increment ΔL that was required to match a given stroboscopic jump s of the surround area was smaller for small interdot separations and became progressively larger with the increase of the interdot separations, d. As the separation between the dots increased, less incremental energy flux was needed to produce w-motion that had the same perceived displacement amplitude. This means that the same amount of the luminance modulation ΔL over a larger spatial separation conveys more evidence for the presence of motion than those over a smaller spatial separation.

Figure 2 (lower panels) shows the same data, but normalized with respect to the displacement distance. In the lower panels; the luminance increment ΔL is plotted against the relative rather than the absolute displacement

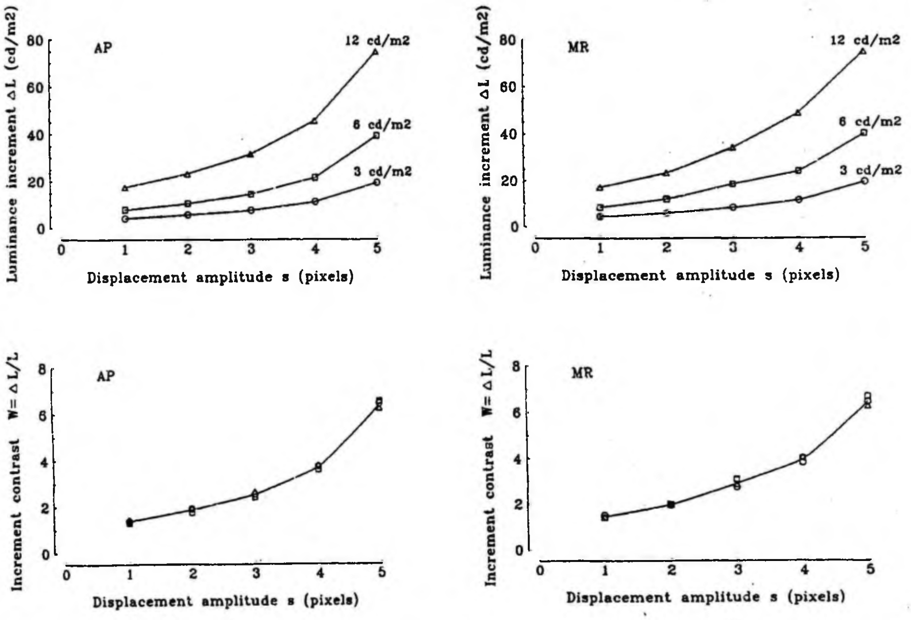


Figure 3. Upper panels: The required luminance increase ΔL (cd/m^2) as a function of the displacement amplitude, s (in pixels), for three different reference luminances, $L = 3$ (circles), 6 (squares), and 12 (triangles) cd/m^2, for 2 observers—A.P. (left panel) and M.R. (right panel). Lower panels: The same data expressed in terms of the relative contrast $W = \Delta L/L$. Symbols are the same as those in the upper panels.

distance. The displacement amplitude is expressed in terms of the proportion to the interdot separation, or simply *s*/*d*. After this transformation, all nine, clearly separate data sets come together to form one single functional relation. This makes it clear that almost the same pattern of results holds for all interdot separations, provided that the amplitude of the adjusted stroboscopic jump is appropriately scaled. Thus, any given luminance difference ΔL between two dots produces an apparent movement whose amplitude is a fixed fraction of spatial separation. This result means, in particular, that the perceived movement is not scaled in terms of velocity; there could be two completely different velocities corresponding to one *s*/*d* value, provided that the transition time remains constant.

Figure 3 shows the results of the second experiment, in which the luminance increment ΔL, required to equalize w- and ϕ-motions, was measured as a function of the reference dot luminance L. Three different reference luminance values (L = 3, 6, and 12 cd/m^2) at one fixed interdot separation (d = 6 pixels; equivalent to 0.05° of arc) were used. As the reference luminance L increased, more luminance modulation ΔL was needed to produce w-motion with the same perceived amplitude. In the lower panels of Figure 3 the same data are replotted, but they are normalized with respect to the luminance modulation. The luminance modulation is expressed in terms of the Weber fraction $W = \Delta L/L$. As a result of this normalization, all the data became almost exactly superimposed. Thus, at a fixed distance between two dots, any given luminance contrast modulation $\Delta L/L$ between these dots produces an apparent displacement of the same amplitude.

The almost perfect constancy of $\Delta L/L$ is a little bit surprising. Usually, photopic luminance discrimination thresholds are measured in conditions in which two spatially separate objects, typically two squares, appear on a large uniform background. The observer's task in the luminance discrimination experiments is to indicate which of these two objects is darker or lighter. Spatially separate stimuli are used to make it more likely that the results will be related to the responses, both subjective and neural, that each stimulus would produce on its own. In these conditions, the luminance difference between two separate objects is noticed as soon as their relative contrast—the ratio between increment or decrement and the standard luminance—reaches a constant threshold value (Whittle, 1986). Unlike in the luminance discrimination task, in the present study, two stimulus dots were always adjacent. They were so close to each other that it was impossible to compare their separate appearances. Instead of telling which of the two dots was darker or lighter, the observer estimated the apparent amplitude of displacement, not of a single micropattern, but of the whole stimulus area. Despite these essential differences between the two psychophysical tasks, all the data obey the same Weber's law: $\Delta L/L$ = constant perceptual outcome.

Many independent psychophysical researchers have indicated that the response of the human motion encoders saturates at low contrast (Campbell & Maffei, 1981; Derrington & Goddard, 1989; Derrington & Henning, 1987; Keck, Palella, & Pantle, 1976; Nakayama & Silverman, 1985; Stone, Watson, & Mulligan, 1990; Thompson, 1982). It has been proposed that the input signals undergo an amplitude-distorting nonlinearity before the motion information is determined. One function that saturates rapidly to a constant value as the signal amplitude increases is the Weibull function:

$$f(w) = 1 - e^{-\left(\frac{w}{k_1}\right)^{k_2}}, \qquad (1)$$

where $W = \Delta L/L$ (Weber's fraction) and k_1 and k_2 are two free parameters of the contrast compression function. We searched for such a function, f, which would allow us to present the adjusted luminance contrast W as a linear function of the relative distance between the two dipole elements. The optimal-fit values were $k_1 = 1.11$ and $k_2 = 0.68$ for A.P., and $k_1 = 1.56$ and $k_2 = 0.68$ for M.R. These estimates are close to $k_1 = 1.99$ and $k_2 = 0.76$ obtained by Stone et al. (1990) in a completely different psychophysical setting. Figure 4 shows the transformed luminance contrast $f(W)$ as a function of the adjusted displacement amplitude, expressed as a fraction of the interdot separation. The correlation coefficients for the

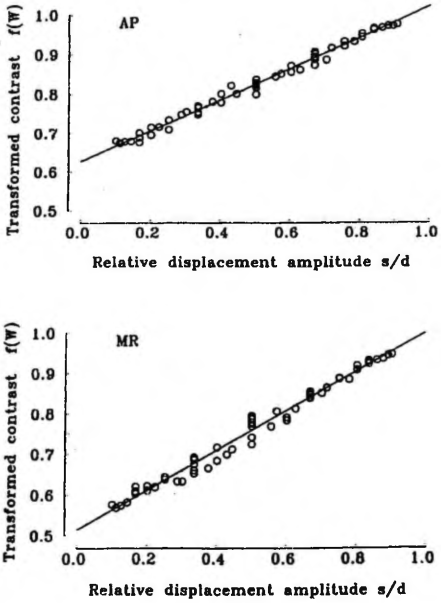


Figure 4. Compressively transformed contrast $f(W)$ as a function of the relative displacement amplitude *s*/*d*, for 2 observers—A.P. (upper panel) and M.R. (lower panel). The data are from the two experiments shown in Figures 2 and 3.

best-fitting functions were highly significant in both cases ($r = .995$ and $r = .991$, respectively). Thus, the proposed linearization function accounts for approximately 98%–99% of the variance in data.

DISCUSSION

The results of the present study demonstrated that spatial displacement of individual elements is not a necessary condition for motion perception. A distinct movement impression can be elicited by a relative change in the light flux at two spatial locations. It was demonstrated that alternating the modulation of luminance of two adjacent dots produced perceived oscillatory w-motion that could not be perceptually distinguished from that produced by a real oscillation. Due to luminance modulation, two stationary dots appeared to oscillate with an amplitude that was easy to match to the amplitude of a real displacement. All of the data can be accounted for by a simple rule: If the relative change in the luminance $W = \Delta L/L$ of two adjacent stationary dots is kept constant, the distance over which these dots appeared to travel in space comprises a fixed fraction of the total distance by which these dots are separated. This result appears to be at variance with the fine-grain movement illusion on human periphery, in which two very closely spaced subsequent flashes produce the impression of movement over a path whose extent considerably exceeds the spatial separation between flashes (Foster, Gravano, & Tomoszek, 1989; Foster, Thorson, McIlwain, & Biederman-Thorson, 1981). Assuming a rapidly saturating luminance contrast compression, it was possible to present the luminance modulation amplitude as a linear function of the relative distance between dots. This may mean, in particular, that exactly the same amount of increase in effective luminance contrast causes exactly the same proportion of the apparent displacement. The established equivalence between the effective relative luminance increment W and the perceived amplitude of displacement suggests that models that posit motion encoding based on the matching of two spatial patterns are not suitable for this particular situation (Dawson, 1991; Lappin & Bell, 1976; Ullman, 1979). The appropriate stimulus for motion is a relative change in light flux at two spatial locations.

Reichardt's (1957) elegant work on the insect movement analyzing system made clear that the simplest operation to detect motion involves the comparison of a signal registered from one spatial location with a delayed signal from another adjacent spatial location. The most general property of any motion-discrimination system is that the comparison process must be nonlinear; multiplication is the minimal operation required to accomplish this comparison (Buchner, 1976; Poggio & Reichardt, 1973; Reichardt, 1987). As a consequence of the multiplication, motion-detection systems based on correlation cannot reliably measure velocity, since their output depends on the contrast and spatial structure of moving patterns. Like insects, the human observer is not able to estimate the perceived velocity of a moving pattern independently of its spatial frequency (Diener, Wist, Dichgans, & Brandt, 1976) and contrast. Thompson (1982) found, for example, that low-contrast gratings appear to move more slowly than a high-contrast reference moving at the same speed. This contrast dependence also implies that the perceived motion direction of a composite pattern can be considerably changed by selectively increasing the luminance of some components of this composite pattern (Allik, 1992; Stone et al., 1990). The results of the present experiment appear to reveal the same property of the underlying motion-encoding operation: The perceived amplitude of w-motion increases monotonically with relative contrast W. Many previous studies have proposed that the correspondence strength between two elements involved in motion increases with luminance flux (Burt & Sperling, 1981; Nishida & Takeuchi, 1990; Shechter & Hochstein, 1989; van Santen & Sperling, 1984; Werkhoven, Snippe, & Koenderink, 1990b). These studies, however, were mainly concerned with the problem of estimating the likelihood that two separate spatial elements form an elementary motion path, rather than with the perceived properties of that path. Correspondence strength, by itself, is ambiguous concerning the output velocity or displacement amplitude. The main advantage of the method equalizing w- and ϕ-motion is that this approach allows the expression of motion strength not only in terms of dimensionless probability of discrimination of direction of motion, but also in metrical units of spatial displacement.

Another consequence of the correlation-type movement-encoding systems concerns the perception of motion without spatial displacement. A motion-encoding system does not need to establish correspondence between similar individual spatial features in a motion sequence. Bilocal motion encoders can ignore the correspondence problem by measuring the asymmetry in the change of the luminance flux at two sampled locations. The bilocal encoding model is indifferent to whether this change in the luminance flux is produced by a moving object traveling from one sample point to another, or by the luminance modulation of two stationary objects at these sample points. Despite obvious physical differences, the motion-encoding system is not able to distinguish these two cases. This explains why w-motion caused by luminance modulation is perceptually indistinguishable from motion evoked by a real displacement. Many current theories of motion perception, which have been shown to be formally equivalent to the elaborated Reichardt model (van Santen & Sperling, 1985), assume that the visual system employs motion-encoding units with receptive fields extended over space and time that are tuned to movement along a particular trajectory (Adelson & Bergen, 1985; Burr et al., 1986a; Watson & Ahumada, 1985; Wilson, 1985). These units measure the amount of luminous energy in some spatiotemporal volume irrespective of the distribution of the luminous energy in that volume. That is why the luminance increment ΔL added alternatively to two stationary objects evokes the perceived motion that is indistinguishable from the impression of motion caused by an object moving from one location to another.

The extended-in-space-time receptive fields means, in particular, that during motion encoding some part of the stimulus information is discarded. For example, when different local motions are spatially superimposed or given within a sufficiently small region, information about individual motion components will be lost and the region is perceived to move in the direction representing a resultant combination of these individual components (Mather & Moulden, 1980; Williams & Sekuler, 1984; Williams, Tweten, & Sekuler, 1991). Similarly, motion encoders seem to ignore the absolute luminance values and respond to the ratio of luminance fluxes, $W = \Delta L/L$, at two sampled locations: Two different pairs of dots with different absolute distance but the same luminance ratio W produce exactly the same magnitude of w-motion.

The results of our experiment suggest that it is easier to elicit motion between two elements with larger spatial separation than between those with smaller spatial separation. As is shown in Figure 2 (upper panels), less modulation in the luminance flux is needed to evoke motion with a required displacement amplitude for a larger interdot separation compared with a smaller one. This finding contradicts the traditional viewpoint that the strongest apparent motion occurs over short inter-element distances (Burt & Sperling, 1981; Miller & Shepard, 1993; Shechter & Hochstein, 1989; Shechter, Hochstein, & Hillman, 1988; Ullman, 1979; Werkhoven, Snippe, & Koenderink, 1990a, 1990b). It is more natural, however, to assume that larger displacements convey more information for the presence of object motion than small displacements, which are, for example, difficult to separate from displacements caused by involuntary eye movements. Many other psychophysical data, including kinematic thresholds and the detection of motion onset or instantaneous displacement, also require for their proper explanation an assumption that the motion-weighting function increases with the displacement magnitude (Allik, 1992; Allik & Dzhafarov, 1984; Dzhafarov, 1992; Dzhafarov & Allik, 1984; Dzhafarov, Sekuler, & Allik, 1993). In order to avoid dependence on a variable motion-weighting function, we analyzed s/d as a fraction of interdot separation. After this normalization, all the curves converged to a single functional relationship, specifying exactly the perceived amplitude of w-motion. For any two values, the luminance modulation increment ΔL and the interdot separation d, there is only one amplitude of the perceived oscillation.

Finally, the idea that the motion-encoding system subjects the input signal to a nonlinear compression is not a new one. The existence of such a compressive operation has been suggested in various contexts (e.g., Bülthoff & Götz, 1979; Chubb & Sperling, 1988, 1991; Egelhaaf & Borst, 1989; Stone et al., 1990; Thorson, 1966). The rapid contrast saturation seems to be an inevitable consequence of a motion-encoding scheme based on the computation of correlation between two input signals. As already noted, this scheme has an intrinsic difficulty with estimating the velocity of a moving object. A simple solution, for a system based on correlation but at the same time not very dependent on stimulus contrast, is to apply the input signal to a rapidly saturating compressive transformation. In that case, only near-threshold low-contrast stimuli are vulnerable to luminance-dependent changes in perceived velocity.

REFERENCES

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

Allik, J. (1992). Competing motion paths in sequence of random dot patterns. *Vision Research*, **32**, 157-165.

Allik, J., & Dzhafarov, E. N. (1984). Reaction time to motion onset: Local dispersion model analysis. *Vision Research*, **24**, 99-101.

Anstis, S. M. (1967). Visual adaptation to gradual change of intensity. *Science*, **155**, 710-712.

Anstis, S. [M.] (1986). Recovering motion information from luminance. *Vision Research*, **26**, 147-159.

Anstis, S. [M.] (1990). Motion aftereffects from a motionless stimulus. *Perception*, **19**, 301-306.

Buchner, E. (1976). Elementary movement detectors in an insect visual system. *Biological Cybernetics*, **24**, 85-101.

Bülthoff, H., & Götz, K. G. (1979). Analogous motion illusion in man and fly. *Nature*, **278**, 636-637.

Burr, D. C., Ross, J., & Morrone, M. C. (1986a). Seeing objects in motion. *Proceedings of the Royal Society of London: Series B*, **227**, 249-265.

Burr, D. C., Ross, J., & Morrone, M. C. (1986b). Smooth and sampled motion. *Vision Research*, **26**, 643-652.

Burt, P., & Sperling, G. (1981). Time, distance, and feature tradeoffs in visual apparent motion. *Psychological Review*, **88**, 171-195.

Campbell, F. W., & Maffei, L. (1981). The influence of spatial frequency and contrast on the perception of moving pattern. *Vision Research*, **21**, 713-721.

Chubb, C., & Sperling, G. (1988). Drift-balanced random stimuli: A general basis for studying non-Fourier motion perception. *Journal of the Optical Society of America A*, **5**, 1986-2007.

Chubb, C., & Sperling, G. (1991). Texture quilts: Basic tools for studying motion-from-texture. *Journal of Mathematical Psychology*, **35**, 411-442.

Dawson, M. R. W. (1991). The how and why of what went where in apparent motion: Modeling solutions to the motion correspondence problem. *Psychological Review*, **98**, 569-603.

Derrington, A. M., & Goddard, P. A. (1989). Failure of motion discrimination at high contrasts: Evidence for saturation. *Vision Research*, **29**, 1767-1776.

Derrington, A. M., & Henning, G. B. (1987). Errors in direction-of-motion discrimination with complex stimuli. *Vision Research*, **27**, 61-75.

Diener, H. C., Wist, E. R., Dichgans, J., & Brandt, T. (1976). The spatial frequency effect on perceived velocity. *Vision Research*, **16**, 169-176.

Dzhafarov, E. N. (1992). The structure of simple reaction time to step-function signals. *Journal of Mathematical Psychology*, **36**, 235-268.

Dzhafarov, E. N., & Allik, J. (1984). A general theory of motion detection. In M. Rauk (Ed.), *Computational models of hearing and vision* (pp. 77-84). Tallinn: Estonian Academy of Sciences.

Dzhafarov, E. N., Sekuler, R., & Allik, J. (1993). Detection of changes in speed and direction of motion: Reaction time analysis. *Perception & Psychophysics*, **54**, 733-750.

Egelhaaf, M., & Borst, A. (1989). Transient and steady-state response properties of movement detectors. *Journal of the Optical Society of America A*, **6**, 116-127.

Foster, D. H., Gravano, S., & Tomoszek, A. (1989). Acuity for fine-grain motion and for two-dot spacing as a function of retinal eccentricity: Differences in specialization of the central and peripheral retina. *Vision Research*, **29**, 1017-1031.

Foster, D. H., Thorson, J., McIlwain, J. T., & Biederman-Thorson, M. (1981). The fine-grain movement illusion: A perceptual probe of neuronal connectivity in the human visual system. *Vision Research*, **21**, 1123-1128.

Gregory, R. L., & Heard, P. F. (1983). Visual dissociations of movement, position and stereo depth: Some phenomenal phenomena. *Quarterly Journal of Experimental Psychology*, **35A**, 217-237.

Johansson, G. (1950). *Configurations in event perception*. Uppsala: Almquist & Wiksell.

Johansson, G. (1978). Visual event perception. In R. Held, H. W. Leibowitz, & H. L. Teuber (Eds.), *Handbook of sensory physiology: Vol. VIII. Perception* (pp. 675-711). Berlin: Springer-Verlag.

Keck, M. J., Palella, T. D., & Pantle, A. (1976). Motion aftereffect as a function of the contrast of sinusoidal gratings. *Vision Research*, **16**, 187-191.

Lappin, J. S., & Bell, H. H. (1976). The detection of coherence in moving random-dot patterns. *Vision Research*, **16**, 161-168.

Mastebroek, H. A. K., & Zaagman, W. H. (1988). Apparent movement induced by luminance modulation: A model study. *Perception*, **17**, 667-679.

Mather, G. (1984). Luminance change generates apparent movement: Implications for models of directional specificity in the human visual system. *Vision Research*, **24**, 1399-1405.

Mather, G., & Moulden, B. (1980). A simultaneous shift in apparent direction: Further evidence for a "distribution shift" model of direction coding. *Quarterly Journal of Experimental Psychology*, **32**, 325-333.

Miller, G. F., & Shepard, R. N. (1993). An objective criterion for apparent motion based on phase discrimination. *Journal of Experimental Psychology: Human Perception & Performance*, **19**, 48-62.

Morgan, M. J. (1979). Perception of continuity in stroboscopic motion: A temporal frequency analysis. *Vision Research*, **19**, 491-500.

Morgan, M. J. (1980). Spatiotemporal filtering and the interpolation effect in apparent motion. *Perception*, **9**, 161-174.

Nakayama, K., & Silverman, G. H. (1985). Detection and discrimination of sinusoidal grating displacements. *Journal of the Optical Society of America A*, **2**, 267-274.

Nishida, S., & Takeuchi, T. (1990). The effect of luminance on affinity of apparent motion. *Vision Research*, **30**, 709-721.

Poggio, T., & Reichardt, W. (1973). Considerations on models of movement detection. *Kybernetik*, **13**, 223-227.

Reichardt, W. (1957). Autokorrelations-Auswertung als Funktionsprinzip des Zentralnervensystem. *Zeitschrift für Naturforschung*, **12B**, 448-457.

Reichardt, W. (1987). Evaluation of optical motion information by movement detectors. *Journal of Comparative Physiology*, **B211**, 533-547.

Shechter, S., & Hochstein, S. (1989). Size, flux and luminance effects in the apparent motion correspondence process. *Vision Research*, **29**, 579-591.

Shechter, S., Hochstein, S., & Hillman, P. (1988). Shape similarity and distance disparity as apparent motion correspondence cues. *Vision Research*, **28**, 1013-1021.

Stone, L. S., Watson, A. B., & Mulligan, J. B. (1990). Effect of contrast on the perceived direction of a moving plaid. *Vision Research*, **30**, 1049-1067.

Thompson, P. (1982). Perceived rate of movement depends on contrast. *Vision Research*, **22**, 377-380.

Thorson, J. (1966). Small-signal analysis of a reflex in the locust: I. Input parameters. *Kybernetik*, **3**, 41-53.

Ullman, S. (1979). *The interpretation of visual motion*. Cambridge, MA: MIT Press.

van Santen, J. P. H., & Sperling, G. (1984). Temporal covariance model of human motion perception. *Journal of the Optical Society of America A*, **1**, 451-473.

van Santen, J. P. H., & Sperling, G. (1985). Elaborated Reichardt detectors. *Journal of the Optical Society of America A*, **2**, 300-321.

Watson, A. B., & Ahumada, A. J., Jr. (1985). Model of human visual-motion sensing. *Journal of the Optical Society of America A*, **2**, 322-342.

Watson, A. B., Ahumada, A. J., Jr., & Farrell, J. E. (1986). Windows of visibility: A psychophysical theory of fidelity in time-sampled visual motion display. *Journal of the Optical Society of America A*, **3**, 300-307.

Werkhoven, P., Snippe, H. P., & Koenderink, J. J. (1990a). Effects of element orientation on apparent motion perception. *Perception & Psychophysics*, **47**, 509-525.

Werkhoven, P., Snippe, H. P., & Koenderink, J. J. (1990b). Metrics for the strength of low-level motion perception. *Journal of Visual Communication & Image Representation*, **1**, 176-188.

Wertheimer, M. (1912). Experimentelle Studien über das Sehen von Bewegung. *Zeitschrift für Psychologie*, **61**, 161-265.

Whittle, P. (1986). Increments and decrements: Luminance discrimination. *Vision Research*, **26**, 1677-1691.

Williams, D., Tweten, S., & Sekuler, R. (1991). Using metamers to explore motion perception. *Vision Research*, **31**, 275-286.

Williams, D. W., & Sekuler, R. (1984). Coherent global motion from stochastic local motions. *Vision Research*, **24**, 275-286.

Wilson, H. R. (1985). A model for direction selectivity in threshold motion perception. *Biological Cybernetics*, **51**, 213-222.

(Manuscript received November 15, 1993;
revision accepted for publication March 15, 1994.)

V

Journal of the Optical Society of America, 1995, 12A (6), (*in press*).

Contrast response of a movement-encoding system

Jüri Allik and Aleksander Pulver

University of Tartu, Department of Psychology, Tiigi 78, Tartu EE-2400, Estonia

Received February 25, 1994; revised manuscript received August 26, 1994; accepted January 20, 1995

The ability to identify the direction of apparent motion in a sequence of two short light pulses of different amplitudes at separate spatial locations was studied. The product of pulse amplitudes is a very poor predictor of such performance when one of the two signals is much higher in amplitude than the other: above a certain amplitude the probability of correct identification becomes virtually independent of the amplitude of the larger pulse. There was no noticeable difference in performance between low–high and high–low contrast sequences. Both the direction identification and the simple contrast-detection probabilities can be represented by the same psychometric function of the luminance increment ΔL, provided that ΔL is normalized by the nth power of the background luminance level, L_b. These results suggest that the general Reichardt-type scheme of movement encoding should be modified in the manner proposed for the fly visual system [J. Opt. Soc. Am. A 6, 116 (1989)]: (1) the mean luminance is subtracted from the input signal before the signal is subjected to a nonlinear compression, and (2) saturation characteristics are inserted into both branches of the two mirror-symmetric motion-detection subunits before multiplication of the input signals. The identical metric of the contrast response suggests that movement discrimination and luminance detection are two different special-purpose computations performed on the output of the same encoding network.

1. INTRODUCTION

The movement of an object is not immediately obvious to the perceiving organism. All motion parameters, including the direction of motion, must be inferred from a time-varying light flux caused by the translation of a luminous pattern across the surface of receptors. Such a translation typically produces almost identical but delayed luminance variations at two neighboring sites along the motion path. Thus the presence of motion in a particular direction can be recovered by measurement of the similarity between two local luminance fluxes along a certain trajectory. On the basis of this kind of observation, Hassenstein and Reichardt[1] proposed an attractively simple delay-and-multiply scheme for the detection of motion in the visual system of the beetle, *Chlorophanus*. According to this scheme the basic operation for motion detection is the multiplication of a signal from one spatial location by a delayed signal from another, adjacent spatial location. It is assumed that motion is detected by a large number of bilocal elementary motion encoders, each of which is composed of two mirror-symmetric component subunits tuned to motion in opposite directions. These subunits share two input channels that sample the visual field at two adjacent point-shaped areas in space. The delay operation can be implemented by a linear low-pass temporal filter. Each subunit detects motion by delaying the temporal luminance pattern in one input channel and multiplying it by the nondelayed pattern in the other input channel. The response of one subunit is algebraically subtracted from that of the complementary unit; the sign of the subtracted signal determines the perceived direction of motion. The subtraction is followed by infinite time averaging. Because of the large number of parallel operating elementary motion encoders, it is necessary to have a rule indicating how outputs of all these encoders are combined into the final decision about motion. Typically, neither of the two simplest rules of combination, the sum and the maximum of all encoder outputs, is assumed to alter the basic properties of an elementary bilocal motion encoder.[2]

In fact, the delay-and-multiply scheme measures the amount of luminous energy concentrated along a certain spatiotemporal motion path and consequently performs a local spatiotemporal Fourier analysis.[3] This basic idea is behind various recent explanations of human motion perception that employ motion-sensitive units, with receptive fields that are elongated and oriented in space–time, and thus are tuned to movement along a definite trajectory.[4] Motion is like orientation in the space–time domain, and motion perception can be regarded as selective recording of orientation in this domain. Thus it is not surprising that various explanations of human motion detection demonstrate a remarkable convergence, becoming essentially identical to only one basic model, the elaborated Reichardt detector.[2,3] In this model, point-shaped receptive fields, appropriate for insect facet eyes, were replaced with spatially extended receptive fields that performed linear spatial filtering of the input image. After this amendment, the elaborated Reichardt detector became fully equivalent, at least formally, to the elaborated motion detector of Watson and Ahumada and the spatiotemporal energy model of Adelson and Bergen.[4]

A sequence of two light stimuli impinging on two adjacent locations of the retina is obviously the elementary event that could evoke the perception of motion, because it is matched exactly to the structure of an elementary motion-encoding operation. If the spatiotemporal interval between two flashes is not too small or too large, the sequence produces a very clear and vivid impression of motion in man[5] and a strong activity of the motion detection neurons in the optic lobe of flies.[6] The most general property of any motion-discrimination system is that the underlying operation comparing the luminance

0740-3232/95/060000-00$06.00

modulation at these two adjacent locations is nonlinear.[7] No linear interaction between these two sample points can reveal the presence of motion. Algebraic multiplication is the simplest operation capable of selective motion evaluation. All correlation-type models make several straightforward predictions for two-flash motion stimuli that can be easily subjected to a experimental falsification. The three most important properties of any correlation-type model are a direct consequence of the multiplication operation on which these models are based:

1. *Monotonicity.* The strength of the motion response is proportional to the product of the amplitudes of the two stimuli, and consequently the accuracy of motion-direction-identification performance must increase monotonically as a function of the product of the two amplitudes;
2. *Commutability.* The probability of accurate motion-detection performance does not change when the order of the two stimuli is reversed. Because of the commutability of multiplication, the exchange of spatiotemporal positions of two stimuli cannot be noticed by elementary motion encoders even if the stimuli have unequal amplitudes;
3. *Sign reversal.* If two flashes have opposite polarity, then, following the rule of algebraic sign multiplication, the predicted movement direction is opposite to the actual order of the flashes.

Only a few attempts to test these predictions more rigorously have been undertaken so far. In particular, van Santen and Sperling[2] found that over a large range of 48:1 the percent correct of motion-direction identification of near-threshold pulses is a monotonically increasing function of the product of the pulses' amplitudes (see experiment 3 below). Although performance was at chance level when pulses were at 2.3% contrast level, performance improved considerably when either of the two amplitudes was increased. In another study Morgan and Cleary[8] also found that when the contrast of both frames in two-frame random cinematograms is too low to permit direction identification, increasing the contrast of either the first or the second frame alone makes confident directional discrimination possible. However, contrary to the monotonicity principle, direction identification first improves with the increase in the contrast and then falls again at higher contrasts. This means that for this low-contrast frame it is more difficult to discriminate movement direction when it is paired with a higher-contrast frame than when it is paired with another lower-contrast one. This near-threshold pedestal effect, well known from luminance-discrimination experiments,[9] obviously violates the monotonicity prediction of the correlation-type models.[10]

Since the product does not depend on the order of multiplicands, it is irrelevant in which sequence, low–high or high–low, two flashes with unequal amplitude are presented. We know of only a few studies of this critical prediction. Morgan and Cleary[8] reported a tendency toward more-difficult motion-direction identification in high–low contrast sequences than in low–high contrast sequences, which evidently violates the law of commutability. Unfortunately, they did not report whether this tendency was significant, and they failed to give a possible reason for this asymmetry. In another study van Santen and Sperling[3] assigned unequal contrast amplitudes to the first and second frames of two-frame sine-wave stimuli. The motion-direction-discrimination thresholds were equally independent of the assignment: exactly the same amount of contrast was required for movement-direction identification in low–high and in high–low stimulus sequences.

Finally, both the original and the elaborated Reichardt models predict the direction reversal of opposite polarity signals. Van Santen and Sperling,[2] who tested this prediction directly, obtained nearly perfect reversal of motion direction when one of the two stimuli had the opposite contrast (see experiment 2 below). Although this result looks solid, there are some observations that appear to limit the generality of their result. In particular, many studies show that opposite polarity signals are perceived as moving in the actual direction, not in the reversed direction as predicted by sign-sensitive multiplication.[11] Besides the lack of reversal itself, the motion signal generated by a pair of opposite-polarity elements is considerably smaller[11] or completely negligible[12] compared with that of two elements with the same polarity. In addition to these observations, several types of drift-balanced stimuli exist that are consistently perceived as moving in a fixed direction and for which no movement can be detected by the standard Reichardt motion-encoding schemes.[13] One typical explanation of these deviant moving phenomena is a nonlinear transformation, such as squaring or full-wave rectification, that is routinely applied to the visual input signal before it is subjected to ordinary temporal correlation analysis. This means, in particular, that after these transformations the information about the polarity of the signals is no longer available for movement encoders, and, as a consequence, opposite-polarity stimuli should appear to move in the actual, not in the reversed, direction, which again violates predictions of the correlation models.

As a consequence of multiplication, movement-encoding elements employing correlation for extraction of movement information cannot reliably measure velocity, since their output depends on contrast. Like insects, the human observer confuses contrast of a moving pattern with its movement attributes. Thompson[14] found, for example, that low-contrast gratings appear to move more slowly than a high-contrast reference grating moving at the same speed. This contrast dependence also implies that the perceived movement direction of a composite pattern can be altered by a selective increase or decrease in the luminance of some components of this composite pattern.[15,16] All these contrast-dependent movement perceptions appear, however, to saturate at a very low contrast level. The observer's ability to estimate movement parameters increases sharply with contrast when the contrast is close to threshold, but further increase in contrast above 2–5% produces no additional increase in the performance.[14,15,17] This compressive nonlinearity is neglected by both the original and the elaborated Reichardt models, which are composed mainly of linear filters if not to mention multiplication itself. Au. queried A few attempts have been made to elaborate the Reichardt model in order to account for empirically observed rapid contrast saturation.[18,19] However, as noticed by van Santen and Sperling,[3] these modifications of the Reichardt model make testing quantitative predictions of its responses to different contrasts highly problematic.

In both insects[19] and mammals,[20] visual neurons responsible for the coding of motion information demonstrate saturation of the same form. Typically, neurons in the middle temporal area of the extrastriate visual cortex, which is specialized for motion processing,[21] reach half of their maximal response at only 7–8% contrast.[20] On average, movement-selective neurons in the middle temporal area saturate at much lower contrast than do striate (V1), lateral-geniculate-nucleus, or retinal neurons.[22] These parallel findings from neurophysiology and psychophysics suggest that the contrast information that is available for the motion-analyzing system is different from that which can be attained by the visual system in general. Compared with the luminance-discrimination system, the movement-encoding system appears to be much more limited in its ability to resolve contrast. In the context of luminance discrimination, saturation is generally regarded as a limitation of the system because of its insufficient dynamic range. In the context of motion, on the contrary, rapid contrast saturation may have a useful function. Rapid contrast saturation seems to be a valuable property of every motion-encoding scheme that is based on the computation of the correlation between two input signals. As already noted above, this scheme has an intrinsic difficulty with estimating the velocity of a moving object. The simplest solution, for a system based on correlation but at the same time not very dependent on stimulus contrast, is to restrict the dynamic range by subjecting the input signal to a rapidly saturating compressive transformation. This saturation could confine contrast-dependent misperceptions to the extreme low end of the contrast scale. It is therefore likely that motion- and pattern-detection mechanisms may rely on two different neural networks, each with its own contrast transfer function. In spite of its plausibility, we found no support for this hypothesis.

In this study we investigated the contrast response of the elementary motion encoders by presenting a sequence of two short light pulses of different contrast above a prevailing luminance level at two adjacent spatial locations. Two basic predictions of all correlation models, monotonicity and commutability, were tested. The ability to discriminate the movement direction of two flashes was found to be independent of their order, but the product of their amplitudes is a very poor predictor of performance. All data can be accounted for by assumption of a very rapid contrast-saturation transformation, which is performed following the stage that eliminates the background light intensity from the input signal. It was demonstrated that two different judgments, the temporal order of two flashes and the presence of a single flash, have a common underlying metric of contrast, which suggests that motion perception and contrast detection can be regarded as two different computations performed on the output of the same encoding network.

2. GENERAL METHODS

A schematic spatiotemporal view of the stimuli is shown in Fig. 1. A sequence of two short luminance increments at two adjacent spatial locations appeared above a prevailing background luminance level. The luminance increments, ΔL, were generated by two red circular LED's 0.06° in diameter, attached to the back of the semitransparent milk film forming the background for these elements. The horizontal gap between the two disks was 0.04°, and consequently the center-to-center separation was ~0.1°. The emission maximum of the diodes was at 690 nm. In order to eliminate possible position cues, we varied the horizontal location of the pair of diodes from trial to trial by equally spacing a horizontal row of six diodes, which permitted presentation of the sequence of the two adjacent luminance increments in five different locations. The horizontal row of diodes, 0.53° in length, was positioned in the center of a uniformly illuminated background field that had the form of a circle, the diameter of which subtended 2.4° at a 3-m viewing distance. The background was illuminated through a red color filter with approximately the same spectrum as the diodes. The mean luminance of the background, L_b, was calibrated with a photomultiplier-based visual spot photometer. In the first two experiments L_b was constantly equal to 0.25 cd/m^2. Only in experiment 3 were three different background luminances used: $L_b = 0.25$, 1.0, and 4.0 cd/m^2. LED luminance was controlled by an IBM AT computer by means of six 12-bit digital-to-analog converter current outputs. A lookup-table method coupled to calibration was used to compensate for small discrepancies in the LED's luminance. Motion stimuli were generated by a sequence of two 15-ms rectangular pulses with amplitudes $\Delta L_{\max}$ and $\Delta L_{\min}$, the larger and the smaller of the two amplitudes, respectively, separated by a ±80-ms interstimulus interval. This time interval was chosen on the basis of our own preliminary experiments and literature findings,[23] which demonstrate that it is close to optimal: at approximately this interstimulus interval value the smallest-amplitude ΔL is required for determining correctly the direction of perceived movement of two subsequent flashes. On each trial the appearance of visual stimuli was preceded by a short tone indicating the beginning of the presentation. The motion direction

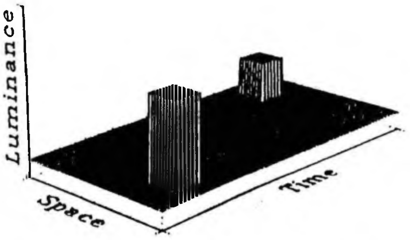


Fig. 1. Schematic space–time view of the stimulus used in this experiment. The luminance cross section of two small adjacent regions perturbing at two succeeding moments in time from a prevailing uniform background luminance is shown. The observer's task was to judge in which direction the whole stimulus pattern appeared to move.

toward the stimulus position presented later in time will be called positive.

The particular values of $\Delta L_{\max}$ and $\Delta L_{\min}$, their assignment to the left and the right position respectively, the horizontal position of the whole stimulus pattern, and the temporal order were chosen before each trial quasi-randomly. Each data point presented in the figures is an average of at least 200 trials. Perception of motion was defined operationally as the ability to discriminate opposite movement directions. The observer's task on each trial was to press one of the two buttons to indicate in which direction, left or right, the whole pattern appeared to move. If the amplitudes of luminance increments were too small to elicit any conspicuous visual impression, the observer was told to guess the movement direction. No feedback about the correctness of choices was provided. All responses were coded in terms of the probability positive-direction choice that is, the probability of choosing the direction corresponding to the actual sequence of the stimuli.

Four psychophysically experienced subjects, the two authors (JA and AP) and two other subjects, naïve regarding the research purpose (TL and TT), participated in this study. Observer TT had emmetropic vision; the three other subjects were corrected myopes. Viewing was binocular in a semidarkened room.

3. RESULTS

A. Experiment 1: Multiplication of Amplitudes

The first experiment was designed to test the multiplicative law: The accuracy of motion-direction identification should increase monotonically as the product of two amplitudes, $m = \Delta L_{\max}\Delta L_{\min}$, increases. That is, except for small statistical fluctuations, whenever the product of the amplitudes of one of the two double-flash stimuli becomes larger than that of the other ($m_i > m_j$), the probability of reporting correct movement direction for the larger of two stimuli, $P(m_i)$, also increases; or, when the difference between m_i and m_j is small, it does not decrease: $P(m_i) \geq P(m_j)$. This means, in particular, that when direction-discrimination performance is plotted as a function of the product of amplitudes m, all the data should lie along a single monotonically increasing curve. Figure 2 shows the probability of the correct (positive) motion-direction identification (in z units) as a function of the product of amplitudes ($\log_{10} m$) for each of the four subjects. In this series of experiments a set of four near-threshold pulse amplitudes were selected from which all possible pairs were used as stimuli. In order to cover the range of movement-detection probabilities from 0.5 to ~0.85, the stimulus values presented to each observer were varied. Different symbols identify a fixed $\Delta L_{\min}$ value. The leftmost point of each symbol set corresponds to the condition $\Delta L_{\max} = \Delta L_{\min}$. The best linear fit of the data points is shown by the straight lines. In general, for this near-threshold stimulus range, the multiplicative law seems to hold: When the product of the amplitudes increased, the accuracy of motion-direction-identification performance also increased. The Pearson product–moment correlations were $r = 0.953$, 0.816, 0.958, and 0.967 for JA, TT, TL, and AP, respectively.

In this experiment, as in the study by van Santen and Sperling,[2] the multiplicative law was tested under the condition that both signals were only barely visible. Next we tested the multiplicative law for the sequence

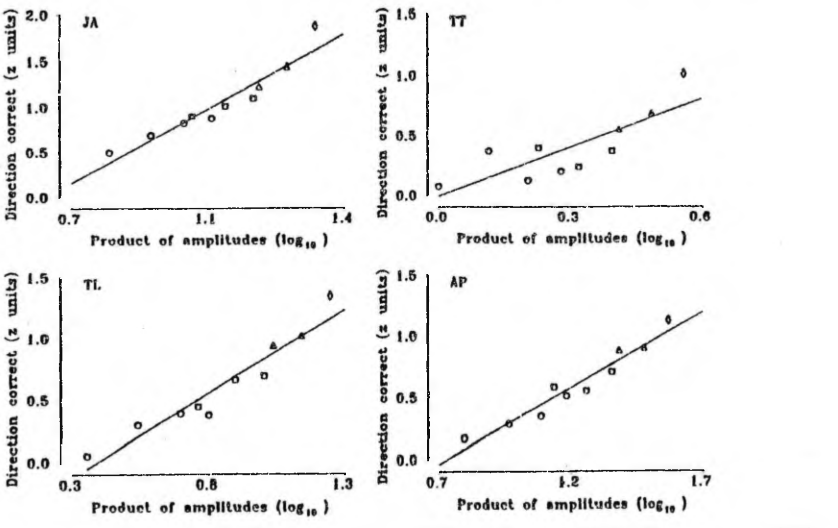


Fig. 2. Probability of correct motion-direction identification (in units of normal distribution) as a function of the product of amplitudes $\log_{10} m$ of two near-threshold flashes. The background luminance $L_b = 0.25$ cd/m^2. Separate panels are shown for the four subjects. Each set of symbols (circles, squares, triangles, and diamonds), corresponds to a particular $\Delta L_{\min}$ value, which increases in the order as listed. Straight lines are the best-fitting linear regressions. Each data point is an average of 200 trials.

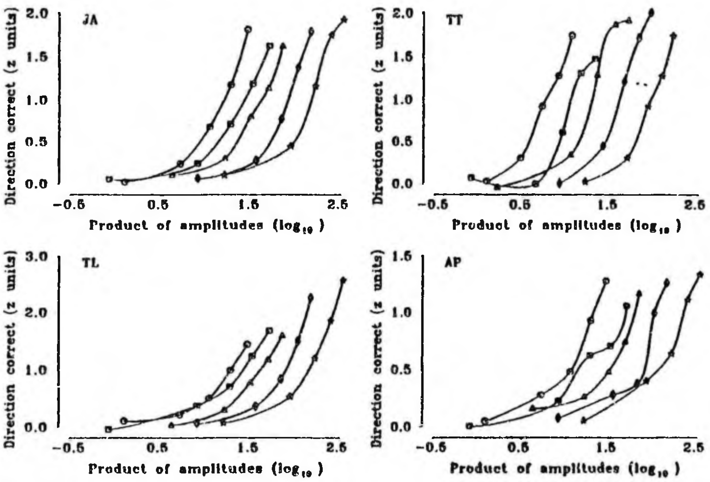


Fig. 3. Probability of correct motion-direction identification (in units of normal distribution) as a function of the product of amplitudes $\log_{10} m$ of two flashes: the smaller had amplitude ΔL and the higher had $\Delta L + L_c$. Each set of data connected by a spline function corresponds to a particular L_c value: 2 (circles), 4 (squares), 8 (triangles), 16 (diamonds), and 32 (stars) cd/m². Each data point is an average of 400 trials.

of two pulses, one of which was clearly above the detection threshold. For that test we defined the smaller of two flashes as $\Delta L_{\min} = \Delta L$ and the larger as $\Delta L_{\max} = \Delta L + \Delta L_c$. Thus a constant amount of energy ΔL_c was added to one of the two stimuli ($\Delta L_{\max}$), with various amounts of incremental energy ΔL added to both flashes to produce a psychometric function covering the full range from ~50% random guessing to ~100% correct movement-direction choices. Figure 3 shows the probability of the correct direction choice (in z units) as a function of the product of the amplitudes m (cd/m²). The curve parameter is ΔL_c, the constant amount of luminous energy added to $\Delta L_{\max}$. For clarity, data points corresponding to one particular ΔL_c value were connected by a spline function: $\Delta L_c = 2$ (circles), 4 (squares), 8 (triangles), 16 (diamonds), and 32 (stars) cd/m². According to the multiplicative rule, all data points should lie on a single monotonically increasing function; this is very different from the obtained result. There are five clearly separate curves, each corresponding to one ΔL_c value. Thus the product of the amplitudes is a very poor predictor of performance in motion-direction identification when one of the two signals has a remarkably higher amplitude than the other. The addition of a constant amount of luminous energy ΔL_c to one of the two increments shifts the empirical psychometric function to the right. On the logarithmic coordinate all shifts appear to be approximately equal; a twofold increase of L_c and thus of the product m shifts the psychometric function ~0.6 logarithmic units to the right. This means that the probability of correct movement identification is virtually independent of the higher of the two amplitudes. The perception of movement direction appears to depend only on the weaker of two signals and can be correctly identified if only the weaker signal is detected. Figure 4 shows the data plotted in Fig. 3, with raw probabilities used as a function of ΔL, the amplitude of the smaller signal. Data points in the empirical psychometric function were approximated by a cumulative Gaussian function with use of the downhill simplex method to minimize the squared deviation between the observed and the predicted movement-direction-identification probabilities. In this format, five different psychometric functions, corresponding to one specific ΔL_c value, can be approximated by a single psychometric function. Although the approximation is far from perfect, the minimum of two amplitudes, $\Delta L_{\min}$, is definitely a better predictor of movement-direction-identification performance than is the product of the amplitudes, m.

Two empirical rules can be formulated on the basis of the data presented in Figs. 2 and 4: (1) for two near-threshold amplitude flashes, the probability of correct movement-direction identification is approximately proportional to the product of their amplitudes, and (2) if one of the two luminance increments becomes considerably higher than the other, movement discrimination depends only on the amplitude of the smaller increment. How can these two seemingly contradictory rules be combined into one overall explanation? One possibility is that a nonlinear transformation is applied to the visual input signal to generate a new signal, which is then subjected to ordinary correlational analysis. As mentioned in Section 1, many data indicate that the response of motion encoders saturates rapidly with an increase in contrast. One suitable function that saturates rapidly to a constant value as the contrast amplitude increases is the Weibull function,

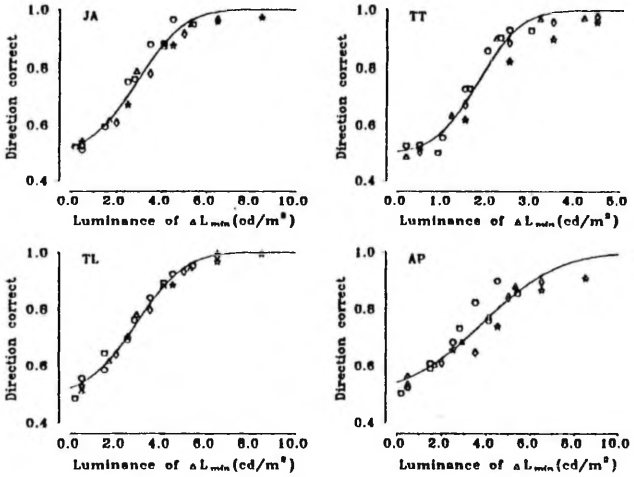


Fig. 4. Data of Fig. 3 replotted in terms of the smaller of two amplitudes $\Delta L_{\min}$ and ΔL (cd/m²). The best-fitting psychometric functions are shown by solid curves.

$$f = 1 - 0.5 \exp[-(\Delta L/k_1)^{k_2}], \qquad (1)$$

where k_1 and k_2 are parameters that determine the shape of the contrast-dependent weighting function.[24] This function provides a satisfactory description of the contrast response of the motion-encoding system, which has been estimated in various psychophysical settings and tasks.[15] How the Weibull function can accommodate, at least qualitatively, the two empirical rules formulated above is illustrated in Fig. 5.

Figure 5 illustrates how signals from two input channels, the left and the right, are combined into the output response of a hypothetical bilocal motion encoder. The two input channels are posited to saturate rapidly as the amplitude of the input signals, ΔL_{left} and ΔL_{right}, increases. Against this particular background level the channels' responses, f_{left} and f_{right}, saturate at approximately 5–6 cd/m² (cf. Fig. 4). The product of transformed responses, $m^* = f_{\text{left}} f_{\text{right}}$, forms a shieldlike surface with a flat roof resting on two smooth slanted walls that meet each other at the point closest to the viewer, where they meet in a rounded corner. If both luminance increments, ΔL_{left} and ΔL_{right}, are below this saturation level, the product m^* increases with the increase in either of the two amplitudes (rule 1). This range corresponds to the corner area between the two slanted walls. When one of the two increments has an amplitude that clearly exceeds the saturation level, the product of responses depends mainly on the amplitude of the smaller signal (rule 2). This minimum-of-two-amplitudes rule corresponds to the two slanted walls that are parallel to the left and the right input axes. Finally, if both amplitudes are approximately 6 cd/m², the output response becomes fully saturated and does not depend on a further increase in either of the two amplitudes. This condition is represented in the figure by the flat roof of the response function. Thus Weibull-type saturation can describe both the multiplicative rule for low-amplitude flashes and the minimum rule for a pair of flashes with considerably different amplitudes.

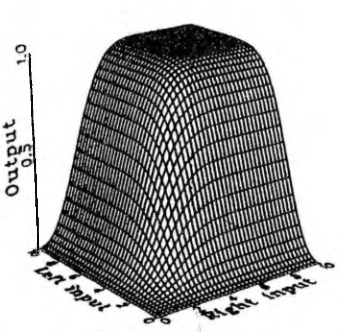


Fig. 5. Perspective view of a two-dimensional response function of a bilocal motion encoder in which two input signals, f_{left} and f_{right}, are posited to saturate rapidly, according to a Weibull function, before they are multiplied together.

Table 1. Best-Fitting Values of Eq. (1)[a]

Subject	k_1	k_2	r
JA	2.80	2.54	0.954
TT	1.60	4.38	0.966
TL	2.89	2.07	0.974
AP	3.53	2.31	0.940

[a] k_1 and k_2, parameters of the Weibull function; r, Pearson product–moment correlation between observed and predicted choice probabilities.

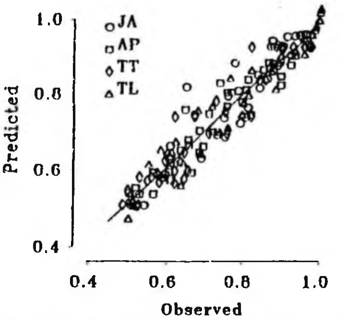


Fig. 6. Predicted (ordinate) and observed (abscissa) probabilities of correct movement-direction identification provided that the input signals are subjected to a compressive Weibull transformation before they are multiplied. This plot contains data for all four subjects. The average correlation between the observed and the predicted data was $r = 0.963$. The data are from Figs. 2 and 3.

Using the downhill simplex method, we obtained the best-fitting values of the parameters of k_1 and k_2 to permit the presentation of motion-direction-identification probability as a function of the product of two Weibull transforms m^*, appropriately corrected for random guessing $(1 + m^*)/2$. The results of this approximation are shown in Table 1. It is interesting to note that Stone *et al.*[16] found the best fit to a diverse set of psychophysical data, including their own, with k_2 very close to 2.

Figure 6 shows the relation between the predicted and the observed motion-direction-identification probabilities for all four observers. The average product–moment correlation between the observed and the predicted data was $r = 0.963$, which accounts for approximately 93% of the total variation in the data; this is remarkable, considering the relatively large random scatter of empirical psychometric functions and the notable individual differences. It is important to note that the values of k_1 and k_2 presented in Table 1 lead to the fewest number of violations of the monotonicity relation as well. Thus our data support the idea that if the input signals from two disparate locations are multiplied by motion encoders they must be subjected to a compressive transformation before the multiplication operation is carried out.

B. Experiment 2: Commutability

In experiment 2 the commutability of motion encoders was tested. Two unequal-amplitude flashes can be presented in either a low–high or a high–low sequence. If the process used by motion encoders to compare two signals has properties that are similar to that of multiplication, the low–high and the high–low sequences should be indistinguishable. In Fig. 7 the probability of correct motion-direction identification is shown as a function of the luminance increment ΔL. As in the second series of experiment 1, the smaller of the two signals was equal to ΔL and the larger was equal to $\Delta L + \Delta L_c$, the sum of the variable-luminance and the constant-luminance increments. Two constant-luminance increment values ΔL_c were used in this experiment: 4 cd/m^2 (circles) and 16 cd/m^2 (triangles). The low–high sequences are shown by filled symbols and the high–low sequences by

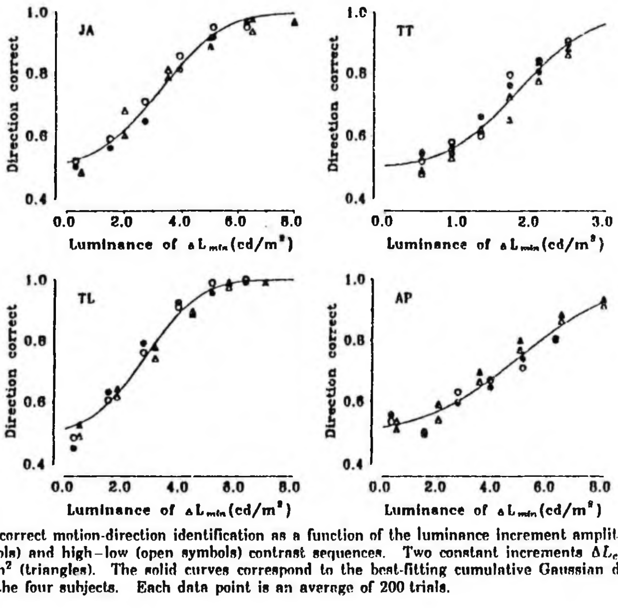


Fig. 7. Probability of correct motion-direction identification as a function of the luminance increment amplitude ΔL_{min} (cd/m^2) for low–high (filled symbols) and high–low (open symbols) contrast sequences. Two constant increments ΔL_c were added to L_{max}: 4 (circles) and 16 cd/m^2 (triangles). The solid curves correspond to the best-fitting cumulative Gaussian distribution. Separate panels are shown for the four subjects. Each data point is an average of 200 trials.

unfilled symbols. Four different data sets (two L_c values by two stimulus sequences) can be represented by a single cumulative Gaussian distribution, characterized by two parameters, the mean and the variance (slope) of the functions. The mean or 75%-correct motion-direction choices was $\Delta L_{0.75} = 3.3$, 1.8, 2.8, and 5.0 cd/m^2 for JA, TT, TL, and AP, respectively, with corresponding slopes (standard deviations) of the psychometric functions $\sigma = 1.76$, 0.80, 1.47, and 2.78, respectively. The absence of systematic trends for open and filled symbols indicates fairly good commutability of the comparison process that underlies motion encoding. Thus the low–high sequence can be approximated by a psychometric function that gives a good fit to the high–low sequence as well. The fit was satisfactory ($p > 0.05$), $\chi^2_{23} = 5.68$ (JA), 6.99 (TT), 6.58 (TL), and 6.86 (AP). Thus this experiment provides additional evidence for the rule that starting at a certain amplitude the addition of luminous energy to the higher of the two stimuli does not change the probability of correct motion-direction identification.

C. Experiment 3: Adaptation to the Background Luminance

Although luminance contrast $\Delta L/L_b$ appears to be the first plausible candidate as the effective stimulus for movement-encoding units, it cannot be taken for granted. The complete independence of contrast thresholds from mean illuminance or, Weber's law, can be observed only at the highest background luminance, typically at ~200 cd/m^2 or higher. Below this, the luminance of the background can be only partly subtracted from the test stimulus increment when amplitude threshold is increasing as a function of L_b.[25] The subtractive adaptation process, also called discounting the background,[26] guarantees that the visual system will discriminate both increments and decrements mainly on the basis of changes from the prevailing luminance level ΔL, regardless of the absolute luminance value $L_b \pm \Delta L$. Specifically, this means that the visual system responds approximately symmetrically to small deviations above and below the prevailing luminance level.[27] Analogously, the motion-analyzing system appears to discard the mean illuminance. Anstis and Mather[28] found that the midgray indifference luminance required for cancellation of apparent movement of two opposite polarity signals when they are pitted against each other lay at the arithmetic mean almost exactly halfway between the luminance of black and white moving stimuli. This symmetry indicates that the background luminance is in some way subtracted from the input signal by the nervous system before the extraction of motion information. In spite of the analogy between detection of luminance increments and cancellation of movement, this symmetry does not prove that exactly the same contrast metric is used by both contrast-detection and motion-discrimination systems. Experiment 3 was designed to examine this possibility more thoroughly.

There were two tasks in this experiment. In the first task, the minimal luminance increment ΔL required for identifying the direction in a sequence of two equal-amplitude flashes dependent on background luminance L_b was studied. The observer was instructed, as in experiments 1 and 2, to indicate in which of two possible directions the presented stimuli appeared to move. This task implies the recognition of the temporal order of the two flash presentations. In the second task, judgments about the presence or absence of a single luminance increment were required. Each trial was divided into two successive observation periods, marked by short sound signals separated by an 800-ms time interval. Only one luminance increment was presented, either in the first or in the second time interval. The observer was instructed to indicate in which interval, the first or the second, the luminance increment appeared. In all other respects—background luminance, duration and increment amplitudes, and spatial uncertainty—conditions for movement-direction identification and contrast detection were identical.

Results of this experiment for two observers are shown in Fig. 8. Three background levels were used: 0.25, 1.0, and 4.0 cd/m^2, shown by circles, triangles, and diamonds, respectively. Data points representing the direction-identification task are connected by solid curves, and data points corresponding to the contrast-detection task are connected by dashed curves. The upper panels of Fig. 8 show probabilities of correct direction and interval identification as a function of $\log_{10}$ luminance ΔL (cd/m^2) for observers AP (left-hand panel) and TT (right-hand panel). The psychometric functions of contrast detection (dashed curves) are systematically shifted to the right with respect to motion discrimination functions. This means that the probability of specifying the temporal order of two flashes was better than the probability of seeing each of theses two flashes when each was presented alone.[29]

With the increase in the background luminance L_b, both motion- and contrast-discrimination functions shift nearly parallel to the right, meaning that on a more luminous background higher-increment amplitudes are required for the same percentage of correct answers. The approximately equal distances among these three sets of psychometric functions suggest that the visibility of the luminance increment ΔL was determined by its ratio to a constant power n of the background luminance L_b, that is, $W = \Delta L/L_b^n$ (for a theoretical explanation of this relation see Kingdom and Moulden[30]). Each data set, one corresponding to motion discrimination and one corresponding to contrast detection, was approximated by a cumulative Gaussian distribution. The best-fitting values for the mean of the psychometric function, $W_{0.75}$, the standard deviation σ, and the power n of the background luminance L_b were searched simultaneously by the downhill simplex method. Table 2 shows the best-fitting values along with the Pearson product–moment correlation between predicted and observed values.

An extremely good fit can be obtained when both motion-discrimination and contrast-detection psychometric functions are expressed in terms of the normalized contrast W, as shown in the lower panels of Fig. 8 for observers AP and TT.[31] After this normalization all three sets of data, each corresponding to one of the three different background levels, obviously bunched together around one single psychometric function. For the sake of presentation clarity, the contrast-detection data (dashed curves) were shifted 0.5 logarithmic unit to the right. The optimal normalization factor, n, was practically identical

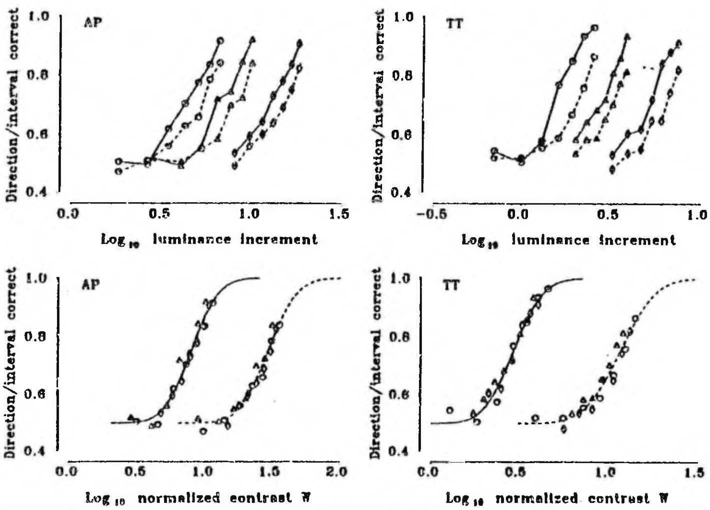


Fig. 8. Probabilities of correct motion-direction identification (solid curves) luminance contrast detection (dashed curves) as a function of luminance increment amplitude ΔL ($\log_{10}$ cd/m^2) (upper panels) and $\log_{10}$ normalized contrast W (lower panels) for subjects AP and TT. Each set of data corresponds to one of three levels of background luminance L_b: 0.25 (circles), 1.0 (triangles), and 4.0 (diamonds) cd/m^2. Each data point is an average of 200 trials.

Table 2. Best-Fitting Values for the Psychometric Functions Shown in Fig. 8 (Lower Panels) and the Normalizing Background Luminance $L_b{}^n$

Subject	Task	$W_{0.75}$	σ	n	r
AP	Motion	8.00	2.76	0.386	0.977
	Contrast	9.36	3.05	0.385	0.978
TT	Motion	2.98	0.89	0.437	0.986
	Contrast	3.68	1.14	0.416	0.967

with the movement-direction-identification and contrast-detection tasks: 0.386 versus 0.384 for AP and 0.437 versus 0.416 for TT. This almost-perfect agreement between normalizing factors $L_b{}^n$ for the motion-direction-identification and the contrast-detection data suggests a common underlying metric for contrast discrimination in these two different psychophysical tasks. In both cases the magnitude of an internal sensation evoked by a luminance increment ΔL added to a steady-state adaptation background is a function of the luminance increment, with the slope of that function being inversely proportional to the nth power of the background luminance level. The luminance increment ΔL required for recognition of the temporal order of two flashes or detection of the presence of a flash increases only slightly less rapidly than a square-root function of L_b, which is typical of the medium (de Vreis–Rose) range of mean illuminance.

4. DISCUSSION

In this study we tested two basic predictions, monotonicity and commutability, of all correlation-type models that are derived from the underlying working principle of the multiplication of two input signals recorded at two nearby locations. Since all other operations of motion encoding are assumed to be accomplished by linear filters, the output response of the Reichardt-type motion encoders can always be factored in such a manner that the response can be presented proportionally to the product of the two input signals. This means that the testing of the correlation model is, in fact, an examination of whether the properties of algebraic multiplication hold in the motion-discrimination tasks. **Au. queried.** First, we were not able to find a noticeable asymmetry between low–high and high–low contrast sequences of two flashes with unequal amplitudes. The lack of asymmetry indicates that commutability holds for the movement-encoding operation. Second, compared with the success in the commutability test, the failure to maintain a monotonic relation between the probability of correct movement-direction identification and the product of amplitudes was obvious. The movement-detection performance deviated from the monotonic relation neither in particulars nor in details but in principle. Two experiments (experiment 1 and 2) demonstrated that motion-direction-discrimination performance becomes almost completely independent of the higher of the two amplitudes, provided that it has already reached a certain amplitude level. It seems that the success of previous studies[2] can be attributed, at least partly, to a fortunate choice of the range of contrast. Indeed, we found in this study that when both signals have relatively low amplitude the multiplication rule can predict movement discrimination performance (Fig. 2); but the product rule fails completely as soon as one of the two amplitudes becomes high enough (Fig. 3). This failure suggests that the neural network underlying the extraction of movement information is somewhat different from that described by both the original[1] and the elaborated[3] Reichardt-type models and that these models need some further revision.

Egelhaaf and Borst[19] proposed a new version of the correlation-type movement-encoding system in order to account for the transient and the steady-state responses of the movement-sensitive interneurons in the fly's visual ganglia. Their elaboration is based on a detailed analysis of the consequences of the saturation characteristics for contrast in the movement-encoding system. It is obvious that the saturation of the response with an increase in contrast is difficult to avoid in any neuronal system. The nonlinear saturation of the input signals, however, does not guarantee that the movement response based on the multiplication of these saturated signals will also be saturated with an increase in contrast amplitude. Quite the opposite: the accuracy of the resulting movement-detection response will increase steeply with increasing modulation amplitude of the input signal. The only way in which saturation of the movement-encoding system can be achieved is through elimination of the mean level of the input signal before the signal is passed through the compressive nonlinearity process. On the basis of these theoretical considerations and the experimentel data, the general scheme of the movement-encoding system was modified in the following ways:

1. The mean luminance is subtracted from the input signal before the signal is subjected to a nonlinear compression.
2. Saturation characteristics are inserted into both branches of the two mirror-symmetric motion-detection subunits before multiplication of the input signals.

Psychophysical data presented in this study support these two modifications of the general scheme of movement encoding. First, our data showed that the background luminance is removed in some way, at least partly, from the effective input signal. This removal can be accomplished by a subtractive process (e.g., high-pass temporal filtering) that decreases the effectiveness of prolonged lights. The movement-encoding system appears to behave as if the background is not present, except that all light increments are scaled by a factor proportional to a power of the mean illuminance (experiment 3). Second, the violation of monotonicity (experiment 1) suggested an essential compressive nonlinearity of the response of movement encoders. We obtained a satisfactory explanation that suggested that the input signals are transformed by a rapidly saturating nonlinear function before they are multiplied. On the basis of our data alone, it was impossible to decide whether the nonlinear compression is performed before or after filtering of the background luminance from the input signal but, as shown by Egelhaaf and Borst,[19] a saturationlike contrast dependence of the movement encoders' output can be achieved only if the mean level of the input signal is eliminated before it is passed through the sigmoidal nonlinearity.

Although these two modifications of the Reichardt model may look rather obvious, they have some important consequences. As a result of saturationlike nonlinearity, the multiplicative scheme loses some of the basic properties characterizing all unmodified Reichardt-type schemes. As was shown above, the monotonicity between the product of the amplitudes and the response strength of the motion encoders does not hold. Thus the distinction between Fourier and non–Fourier motion-perception mechanisms[13] becomes less obvious, because the concept of drift-balanced stimuli is based on the expected response of any Reichardt detector without these two modifications. Another result of these two modifications is that the movement-detection scheme is brought closer to current explanations of contrast perception. For example, subtractive mechanisms have been known for a long time and have been proposed as an explanation for a number of phenomena in brightness and color perception.[26,32] As in the motion-discrimination data, the contrast-detection data also suggest that the subtractive adaptation process must occur before the saturationlike nonlinearity is applied to the input signal. Although there are two different traditions for describing compressive nonlinearity—the contrast-detection tradition uses the hyperbolic tangent (Naka–Rushton) function,[24] and the movement-detection tradition favors the Weibull function[15]—their differences are superficial, at least from a practical point of view. Thus after these two modifications the movement-detection scheme becomes much more similar to a typical luminance-detection model, at least with respect to the qualitative similarity of the subtractive adaptation and the saturationlike nonlinearity modules. This similarity may, of course, be only superficial. Much physiological and psychophysiological evidence demonstrates that different components of visual information processing are segregated into largely independent parallel pathways.[33] This segregation suggests the possibility that, although two components of the visual system can perform the same general function (e.g., nonlinear compression of the input signal), their operating characteristics may be very different.

A direct comparison between the movement-discrimination and the contrast-detection systems demonstrates, however, that the similarity between the two systems goes far beyond qualitative aspects. For example, experiment 3 of this study showed that parameters characterizing a subtractive adaptation process, the purpose of which is to eliminate the average illuminance from the input signal, are virtually identical for both movement-discrimination and contrast-detection tasks. In addition to the identical-contrast metric, several other lines of evidence provide some further support for the view that movement discrimination and contrast detection share a common underlying neural network. We have mentioned above the near-threshold pedestal effect that can be observed not only in the luminance contrast-detection task but also in the movement-direction identification task.[8] In a previous study[34] we demonstrated that the observer's ability to judge the perceived temporal order of the two adjacent luminance excursions, which is experienced as a displacement of the whole pattern, can be explained by exactly the same fundamental process that underlies flicker detection, flash visibility, and temporal masking. In all these perceptual tasks the observer's performance can be inferred from the impulse-response function of the visual system. This means, in particular, that all empirical rules characterizing the contrast thresholds of a probe impulse presented in the presence of a masking stimulus[35] can be easily applied to a situation in which the observer has to decide in which direction two adjacent luminance excursions appear to move.[34] Thus the identical-contrast metric, together with the similar temporal response properties, suggests that questions about the presence of changes in the illuminance and in the direction of motion are answered by some computations that are performed on the outputs of a common neural network.

Sharing a common input does not mean that motion detection is just another instance of contrast detection. In this respect it would be instructive to compare the probability that an observer will be able to identify the temporal order of two flashes with the probability that the observer will see each of these two flashes presented in isolation. For this particular stimulus configuration the temporal order of two flashes was consistently identified more frequently than just the presence of each flash when it was presented alone. This is quite a paradoxical situation: The observer can more confidently identify the relation between two visual events than detect the presence of each of these two events. It is not enough to discover the presence of only one stimulus in a two-flash stimulus to see movement; it is necessary to register both events in their respective spatiotemporal positions. If we assume that the detection of these two events is statistically independent, then the probability of their conjoint detection must be equal to the product of their individual detection probabilities. Consequently, it is impossible that the probability of the conjoint detection can be higher than the probabilities of detecting these individual events. The only solution to this paradox is that the events are not independent but are highly correlated. This correlation process must be very effective, since the probability of identifying the temporal order of two flashes is only slightly worse au: queried than the simple detection of their presence without a need for specification of their exact succession.[29] Cortical neurons with receptive fields that are elongated and oriented in space–time can be regarded as an implementation of this correlation process.[4] Thus if motion detection can be regarded as a special case of contrast detection, then in a special sense it can be regarded as a process establishing a correlation between two visual events at two separate space–time locations.

Perhaps it is not so surprising that motion-direction-identification performance saturates very rapidly and becomes independent of further increase in contrast at a very low level of contrast. The estimation of some other motion attributes such as velocity may be, in principle at least, much less restricted in the contrast-response range, saturating at a considerably higher contrast level. The Reichardt model was initially devised as an explanation of the visual system's ability to determine movement direction under an arbitrary time-varying retinal illumination. It explains how the observer's responses—for example, the direction in which the fly turns its head indicates the perceived movement direction—can be predicted from parameters of a display, but it cannot explain, even in principle, how the velocity can be computed.[3] Although velocity discrimination requires information other than that which is provided by Reichardt-type elementary movement encoders, a number of studies indicate that all other contrast-dependent misperceptions of motion are also confined to a near-threshold contrast range.[14,15,17]

Recently Dzhafarov *et al.*[36] proposed a possible mechanism by which the human visual system can detect changes in the movement of visual targets. According to this explanation the movement-detection task can be regarded as consisting of two parts, encoding and detection. Motion encoding is posited to be a general-purpose computation that is relatively task independent and is performed by an array of elementary bilocal movement encoders.[37] The dynamic array of elementary activations contains sufficiently rich information about motion that specific questions about motion can be answered by means of special-purpose and task-specific computations performed on the outputs of the encoding network. This approach regards motion detection as a computation over the distributed pattern of activations within the network of elementary motion encoders. In the primary encoding network, every elementary movement analyzer samples two different spatial areas at two moments in time. A variety of regularities in psychophysical data, such as the specific form of kinematic thresholds for oscillatory motion[38] or reaction times to onset of motion[39] or to changes in speed,[37] can be accounted for by the assumption that every elementary movement analyzer can be in only one of two possible states: activated or nonactivated. The outputs of elementary movement encoders provide only the most basic information about motion: its presence or absence. Nevertheless, this most primitive information is enough for an array of these encoders to provide sufficient information about all relevant attributes of a moving target. Thus the observed rapid contrast saturation seems to play a constructive role by limiting the output of the elementary movement encoders to two possible states, activated and nonactivated. In other words, what was described so far as a nonlinear saturation can be simply understood as a threshold function. Along with this interpretation the Weibull function obtains another meaning, simply describing the probability with which a certain event can be detected. As soon as an event is detected, the accuracy of the observer's responses reaches its maximum and does not improve further. The elementary movement encoders become, in turn, more like logical devices that activate only if two events are detected at two specified spatial locations at two different moments of time. This interpretation also provides a logical explanation of the striking similarity between movement-discrimination and contrast-detection thresholds. In any case, the dramatic reduction of information at these early stages of visual computation may be inevitable for a system that is based on correlation and therefore is extremely prone to any changes in stimulus contrast.

ACKNOWLEDGMENTS

This research was supported by a James S. McDonnell Foundation grant and Estonian Science Foundation grant 450 to J. Allik. We thank Stanley Klein, Marion Cohn, and two anonymous referees for their valuable comments on an earlier draft of this paper. Part of this work was

presented at the Sixteenth European Conference on Visual Perception, Edinburgh, Scotland, August 1993.

REFERENCES AND NOTES

1. B. Hassenstein and W. Reichardt, "Systemtheoretische Analyse der Zeit-, Reihenfolgen- und Vorzeichenauswertung bei der Bewegungsperzeption des Rüsselkäfers *Chlorophanus*," Z. Naturforsch. Teil B 11, 513–524 (1956); W. Reichardt, "Autokorrelations-Auswertung als Funktionsprinzip des Zentralnervensystems (bei der optischen Wahrnehmung eines Insektes)," Z. Naturforsch. Teil B 12, 488–457 (1957).
2. J. P. H. van Santen and G. Sperling, "Temporal covariance model of human motion perception," J. Opt. Soc. Am. A 1, 451–473 (1984).
3. J. P. H. van Santen and G. Sperling, "Elaborated Reichardt detectors," J. Opt. Soc. Am. A 2, 300–321 (1985).
4. E. H. Adelson and J. R. Bergen, "Spatio-temporal energy models for the perception of motion," J. Opt. Soc. Am. A 2, 284–299 (1985); A. B. Watson and A. J. Ahumada, Jr., "Model of human visual-motion sensing," J. Opt. Soc. Am. A 2, 322–342 (1985); D. C. Burr, J. Ross, and M. C. Morrone, "Seeing objects in motion," Proc. R. Soc. London B 227, 249–265 (1986).
5. S. Exner, "Über das Sehen von Bewegungen und die Theorie des zusammengesetzten Auges," Sitzungsber. Kaiser. Akad. Wiss. 72, III, 156–190 (1876); J. Thorson, G. D. Lange, and M. Biederman-Thorson, "Objective measure of dynamics of a visual movement illusion," Science 164, 1087–1088 (1969); G. Westheimer and S. McKee, "Perception of temporal order in adjacent visual stimuli," Vision Res. 17, 887–892 (1977).
6. G. D. McCann, "The fundamental mechanism of motion detection in the insect visual system," Kybernetik 12, 64–73 (1973).
7. T. Poggio and W. Reichardt, "Considerations on models of movement detection," Kybernetik 13, 223–227 (1973); E. Buchner, "Elementary movement detectors in an insect visual system," Biol. Cybern. 24, 85–101 (1976); W. Reichardt, "Evaluation of optical motion information by movement detectors," J. Comp. Physiol. A 161, 533–547 (1987).
8. M. J. Morgan and R. Cleary, "Effect of contrast substitutions upon motion detection in spatially random patterns," Vision Res. 32, 639–643 (1992).
9. B. Leshowitz, H. B. Taub, and D. H. Raab, "Visual detection of signals in the presence of continuous and pulsed backgrounds," Percept. Psychophys. 4, 207–213 (1968); J. Nachmias and E. C. Kocher, "Visual detection and discrimination of luminance increments," J. Opt. Soc. Am. 60, 382–389 (1970); G. E. Legge and J. M. Foley, "Contrast masking in human vision," J. Opt. Soc. Am. 70, 1458–1471 (1980); P. Whittle, "Increments and decrements: luminance discrimination," Vision Res. 26, 1677–1691 (1986).
10. This effect is contrary to the claim of Morgan and Cleary[8] that van Santen and Sperling[2] found a similar effect and that the pedestal effect provides strong support for models of the Reichardt type.
11. S. Shechter and S. Hochstein, "On and off pathway contributions to apparent motion perception," Vision Res. 30, 1189–1204 (1990); C. Wehrhahn and D. Rapf, "ON- and OFF-pathways form separate neural substrates for motion perception: psychophysics evidence," J. Neurosci. 12, 2247–2260 (1992).
12. J. Allik and E. Dzhafarov, "Motion direction identification in random cinematograms: a general model," J. Exp. Psychol.: Human Percept. Perf. 10, 378–393 (1984). It is important to note that the interaction between elements of opposite polarity is not necessarily the cause of the reversed phi motion phenomenon that can be observed when two identical random-dot patterns are presented one after the other, with a slight spatial shift between them, when one of these two patterns is the photographic negative of the other [S. M. Anstis, "Phi movement as a subtraction process," Vision Res. 10, 1411–1430 (1970); S. M. Anstis and B. J. Rogers, "Illusory reversal of visual depth and movement during changes of contrast," Vision Res. 15, 957–961 (1975); T. Sato, "Reversed apparent motion with random dot patterns," Vision Res. 29, 1749–1758 (1989)]. As a result of the contrast reversal, the potential number of displaced elements that preserve their contrast polarity becomes larger in the opposite than in the actual direction of image displacement. Thus the reversed phi motion in multielement patterns does not require sign-sensitive multiplication for its explanation.
13. 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); C. Chubb and G. Sperling, "Texture quilts: basic tools for studying motion-from-texture," J. Math. Psychol. 35, 411–442 (1991).
14. P. Thompson, "Perceived rate of movement depends on contrast," Vision Res. 22, 377–380 (1982).
15. L. S. Stone, A. B. Watson, and J. B. Mulligan, "Effect of contrast on the perceived direction of a moving plaid," Vision Res. 30, 1049–1067 (1990).
16. J. Allik, "Competing motion paths in sequence of random dot patterns," Vision Res. 32, 157–165 (1992).
17. A. Pantle and R. Sekuler, "Contrast response of human visual mechanisms sensitive to orientation and direction of motion," Vision Res. 9, 397–406 (1969); M. J. Keck, T. D. Palella, and A. Pantle, "Motion aftereffect as a function of the contrast of sinusoidal gratings," Vision Res. 16, 187–191 (1976). K. Nakayama and G. H. Silverman, "Detection and discrimination of sinusoidal grating displacements," J. Opt. Soc. Am. A 2, 267–274 (1985); A. Derrington and B. Henning, "Errors in direction-of-motion discrimination with complex stimuli," Vision Res. 27, 61–75 (1987); A. Derrington and P. A. Goddard, "Failure of motion discrimination at high contrast: evidence for saturation," Vision Res. 29, 1767–1776 (1989).
18. H. Bülthoff and K. G. Götz, "Analogous motion illusion in man and fly," Nature 278, 636–637 (1979).
19. M. Egelhaaf and A. Borst, "Transient and steady-state response properties of movement detectors," J. Opt. Soc. Am. A 6, 116–127 (1989).
20. D. G. Albrecht and D. B. Hamilton, "Striate cortex of monkey and cat: contrast response function," J. Neurophysiol. 48, 217–237 (1982).
21. J. H. R. Maunsell and D. van Essen, "Functional properties of neurons in middle temporal visual area of the macaque monkey. I. Selectivity for stimulus direction, speed and orientation," J. Neurophysiol. 49, 1127–1147 (1983); A. Mikami, W. T. Newsome, and R. H. Wurtz, "Motion selectivity in macaque visual cortex. I. Mechanisms of direction and speed selectivity in extrastriate area MT." J. Neurophysiol. 55, 1308–1327 (1986); H. R. Rodman and T. D. Albright, "Coding of visual stimulus velocity in area MT of the macaque," Vision Res. 27, 2035–2048 (1987).
22. G. Sclar, J. Maunsell, and P. Lennie, "Coding the image contrast in central visual pathways of the macaque monkeys," Vision Res. 30, 1–10 (1990); M. Kaplan and R. M. Shapley, "The primate retina contains two types of ganglion cells, with high and low contrast sensitivity," Proc. Natl. Acad. Sci. (USA) 16, 187–191 (1986).
23. D. Foster, S. Gravano, and A. Tomoszek, "Acuity for fine-grain motion and for two-dot spacing as a function of retinal eccentricity: differences in the specialization of the central and peripheral retina," Vision Res. 29, 1017–1031 (1989).
24. The choice of the Weibull function was motivated mainly by its previous successful application by Stone and his colleagues.[15] More physiologically motivated is a hyperbolic function $R(\Delta I) = \Delta I^n/(\Delta I^n + \sigma^n)$, where R is the response amplitude determined by the flash amplitude ΔI, an exponent n, and the semisaturation constant σ, that is, the intensity at which half of the maximum response is achieved. This function provides a satisfactory fit to the contrast response of vertebrate visual neurones [cf. K. I. Naka and W. A. H. Rushton, "S-potential from colour units in the retina of fish (*Cyprinidae*)," J. Physiol. 185, 536–555 (1966)]. An appropriate selection of n and σ values always guarantees a satisfactory fit to a particular Weibull function.
25. H. R. Blackwell, "Contrast thresholds of the human eye," J. Opt. Soc. Am. 36, 624–643 (1946).

26. J. Walraven, "Discounting the background—the missing link in the explanation of chromatic induction," Vision Res. **16**, 289–295 (1976).
27. D. A. Burkhardt, J. Gottesman, D. Kersten, and G. E. Legge, "Symmetry and constancy in the perception of negative and positive luminance contrast," J. Opt. Soc. Am. A **1**, 309–316 (1984).
28. S. M. Anstis and G. Mather, "Effect of luminance and contrast on direction of ambiguous apparent movement," Perception **14**, 167–179 (1985). See also S. Shechter and S. Hochstein, "Characteristics of the flux dimension cue in apparent movement correspondence," Spatial Vis. **6**, 11–24 (1992).
29. The temporal order of two flashes was not discriminated as well as these two flashes were detected. In an additional experiment, results of which are not reported here, we found that the presence of two flashes was detected slightly better than the temporal order in which they were exposed.
30. F. Kingdom and B. Moulden, "A model for contrast discrimination with incremental and decremental test patches," Vision Res. **31**, 851–858 (1991).
31. A slightly worse fit was obtained with a Stiles threshold-versus-illuminance function $\Delta L/(Lb + L^*)$, which assumes an internal source of excitation (dark light), L^*, which acts like real light and adds to it [H. B. Barlow, "Dark and light adaptation: psychophysics," in *Handbook of Sensory Physiology*, D. Jameson and L. M. Hurvich, eds. (Springer-Verlag, New York, 1972), Vol. VII/4, pp. 1–28. Satisfactory approximation to our data was obtained when L^* was equal to 1.7 and 1.4 cd/m^2 for the movement-direction identification and the contrast-detection tasks, respectively.
32. E. H. Adelson, "Saturation and adaptation in the rod system," Vision Res. **22**, 1299–1312 (1982); W. S. Geisler, "Mechanisms of visual sensitivity: backgrounds and early dark adaptation," Vision Res. **23**, 1423–1432 (1983); M. M. Hayhoe, N. I. Benimoff, and D. C. Hood, "The time-course of multiplicative and subtractive adaptation process," Vision Res. **27**, 1981–1996 (1987); N. Graham and D. C. Hood, "Modeling the dynamic of light adaptation: the merging of two traditions," Vision Res. **32**, 1373–1393 (1992).
33. M. S. Livingstone and D. H. Hubel, "Psychophysical evidence for separate channels for the perception of form, color, movement, and depth," J. Neurosci. **7**, 3416–3468 (1987).
34. J. Allik and A. Pulver, "Timing of visual events for motion discrimination," Vision Res. **34**, 1585–1594 (1994).
35. G. Sperling, "Temporal and spatial visual masking. I. Masking by impulse flashes," J. Opt. Soc. Am. **55**, 541–559 (1965).
36. E. N. Dzhafarov, R. Sekuler and J. Allik, "Detection of changes in speed and direction of motion: reaction time analysis," Percept. Psychophys. **54**, 733–750 (1993).
37. W. A. van de Grind, J. J. Koenderink, and A. J. van Doorn, "Detection of coherent movement in peripherally viewed random-dot patterns," J. Opt. Soc. Am. **73**, 1674–1683 (1983); A. J. van Doorn and J. J. Koenderink, "The structure of the human motion detection system," IEEE Trans. Systems, Man, Cybern. **SMC-13**, 916–922 (1983).
38. E. N. Dzhafarov and J. Allik, "A general theory of motion detection," in *Computational Models in Hearing and Vision*, M. Rauk, ed. (Estonian Academy of Sciences, Tallinn, Estonia, 1994), pp. 77–84.
39. J. Allik and E. Dzhafarov, "Reaction time to motion onset: local dispersion model analysis," Vision Res. **24**, 99–101 (1984).

VI

Vision Research, 1995 (*submitted*).

When A Sequence Of Two Opposite Polarity Flashes Is Perceived To Move In The Reversed Direction

ALEKSANDER PULVER, JÜRI ALLIK

The observer's ability to identify the movement direction of two short luminance increment or decrement pulses at various spatiotemporal separation between them was investigated. At variance of the same polarity pulses that are always perceived moving in the direction of their actual succession, two pulses with the opposite polarity are perceived moving in the direct or reversed directions depending on their spatiotemporal separation. The reversal of the perceived motion direction, that follows the rule of algebraic sign multiplication, is observed only when temporal asynchrony between two succeeding pulses is less than about 60 msec and the spatial separation between them is less than 1.3-1.5 mm of cortical distance. Beyond this spatiotemporal range two pulses with opposite polarity of contrast are always perceived moving in the direction of their actual succession. The coexistence of the reversed and direct movements suggests that neither the original Reichardt model nor any models that assume insensitivity to the contrast polarity can be alone a complete description of the motion encoding process in the human visual system. It was proposed that movement encoding in human visual system is accomplished by encoders with a complex spatiotemporal receptive field the central portion of which exhibits properties of algebraic sign-sensitive multiplication and the periphery is chacacterized by the rectification of the input signal before subjecting it to the standard correlation analysis.

Motion detection; reversal of motion direction; ON and OFF pathways; the Reichardt model of motion detection

INTRODUCTION

Separate ON and OFF pathways in the visual system

In the mammalian eye, ON-center and OFF-center retinal ganglion cells (Kuffler, 1953) form inputs to two major pathways which division is maintained at the striate cortex level (Schiller, 1982, 1992). By using the glutamate neurotransmitter analogue 2-amino-4-phosphonobutyrate (Slaughter & Miller, 1981), it is possible to block selectively ON responses in retinal ganglion cells, as a result of which the detection of light increments but not of light decrements is severely impaired (Schiller, Sandell & Maunsell, 1986). This finding indicates that the ON and OFF pathways remain independent and light increments and decrements are coded separately in the visual system (e.g. Burton, Nagshineh & Ruddock, 1977; Watt & Morgan, 1985). Although many psychophysical evidences also support the division between ON and OFF pathways the extent and maintenance of their separation is debatable. There are several convincing psychophysical evidences suggesting a considerable overlap between mechanisms coding luminance changes in two opposite directions. Tolhurst and Dealy (1975) demonstrated, for example, that changes in contrast can be detected better than the polarity of these changes can be identified. This fact means that not all contrast changes are labelled according to their polarity and it is possible to have some information about any change independent of the polarity of contrast (cf. Klein, 1985). One can also think about this situation in terms of ON and OFF mechanisms sensitive not only to their appropriate polarity but the opposite one as well, that is in terms of their at least partial blindness to the polarity of contrast.

Algebraic sign-sensitive multiplication

A sequence of two luminance changes occurring at two adjacent locations is obviously the most elementary event that could elicit a directional motion response. The most general property of any motion discrimination system is that the underlying operation that compares the luminance modulation at these two sample locations must be nonlinear (Poggio & Reichardt, 1973; Buchner, 1976; Reichardt, 1987). No linear interaction (summation or subtraction) between signals recorded at these two sample points can reveal the presence of directed motion. One of the simplest operations which is able to reveal the presence of directed movement is, of course, algebraic multiplication. This means, however, that if the motion encoding system is using multiplication-like operation, it must obey the rules of the algebraic sign multiplication. In particular, the reversal of the sign of contrast in one of the two luminance pulses must lead to the reversal of the perceived motion direction. A sequence of two flashes which have opposite polarities must be perceived to move in the direction opposite to the actual stimulation succession. Both, the original (Reichardt, 1957) and the elaborated Reichardt models (van Santen & Sperling, 1985), predict that a sequence of two opposite polarity fluxes elicits the perception of motion in the reversed direction. This prediction is physiologically plausible. Beside the optomotoric reaction of insects, complex cells of mammals' visual cortex appear to exhibit the productlike properties required by the Reichardt model. Emerson et al. (1987, Figure 5) obtained the reversed motion when two opposite polarity bars were presented in sequence at two adjacent locations in the receptive fields of the cat complex cells. It is somewhat astonishing that this fundamental prediction of all correlation-like schemes has been tested experimentally in a few psychophysical studies only. To our knowledge, the only direct test of this prediction in human vision was undertaken by van Santen and Sperling (1984, Experiment 2). They applied a periodic travelling train of pulses to a field which was divided into adjacent bars. When the polarity of the pulse function was reversed in each even bar subjects saw near-perfect reversal of the perceived motion direction as it is predicted by the Reichardt models. It is important to notice, that bars containing same-sign pulses were phased in such a way that they did not contain direction information. This is an essential difference from many other studies in which the reversal of luminance contrast of multielement random patterns is used to produce the reversed apparent motion.

Two types of direction reversal

There are two different reasons for the reversed motion, one related to sign-sensitive multiplication and the other to statistical properties of multielement pattern. If two identical patterns are presented one after another, with a slight spatial shift between them, then motion is perceived in the direction of the later stimulus. But if one of the two patterns is the photographic negative of the other, then motion is perceived in the direction of the earlier stimulus, that is in the opposite direction to the actual displacement (Anstis, 1970, 1986; Anstis & Rogers, 1975; Marr & Ullman, 1981; Sato, 1989). This phenomenon was called 'reversed phi motion' and was initially reported for multi-element random-dot patterns. Reversed phi motion is an extreme case of a more general direction reversal phenomenon which can be obtained when each element in a spatially shifted pattern is more likely to reverse its contrast than preserve it (Allik & Dzhafarov, 1984). As a result of contrast reversal, the number of element pairs with the same contrast polarity in the direction of actual displacement decreases and becomes less than that in the opposite displacement direction. Thus, the perceived motion direction can be simply determined by the number of potentially displacing elements maintaining their contrast. This explanation does not need an assumption that stimulus elements with the opposite polarity contribute to the perception of motion. The original explanation of the reversed-phi proposed by Anstis (1970) was based on similar statistical considerations. Later, however, he proposed another explanation based on the perceived localization of different luminance distributions (Anstis & Rogers, 1975; Anstis, 1978). Purely brightness explanation of the reversed phi motion is doubtful at least. The reversed pitch motion, an acoustic analogue of the revered phi motion, is a clear indication of the limitations of a purely brightness explanation (Allik, Dzhafarov, Houtsma, Ross & Versfeld, 1989). Thus, there is no need to explain the motion reversal phenomenon in multi-element patterns in terms of interaction between opposite polarity luminance fluxes. Indeed, one of the most striking properties of the reversed phi motion created by multi-element patterns, both visual and acoustical, is the asymmetry between direct and reversed motion: the identification of the reversed motion is worse than the identification of direct motion (Allik & Dzhafarov, 1984; Allik et al., 1989; Sato, 1989). A detailed analysis of this asymmetry leads to the conclusion that only elements with the same polarity contribute to the perception of motion; the contribution of element pairs with the opposite polarity is very small if not negligible. Green (1989) found that in ambigous situations the observer has a clear preference to see motion in the direction that preserved the sign of element contrast. Recently, Edwards and Badcock (1994) also found that dots that change luminance polarity do not effectively drive the global-motion mechanism. The same general rule, that no or only weak correspondence is established between two elements with the opposite contrast, appears to be valid for the perception of global orientation flow and depth in multielement patterns (cf. Allik, 1992).

Two parallel systems

In many cases, however, the motion impression can be created by the stimulus elements of the opposite polarity which direction corresponds not to the reversed but to the actual succession of elements (Livingston & Hubel, 1987; Shechter & Hochstein, 1990; Wehrhahn & Rapf, 1992). In short, two contradictory evidences exist about the ability of the movement encoding system to resolve the polarity of contrast: one group of evidences demonstrates the visual system's ability to take into account the polarity of the luminance changes and the other group of evidences demonstrates no such ability. These apparently contradictory data support a very popular theme running through recent studies of two parallel movement processing systems, usually called "first-order" and "second-order" motion systems (Cavanagh & Mather, 1989). Chubb & Sperling (1988, 1991) characterize the first-order system as one which behavior can be predicted from the spatiotemporal Fourier transform of the stimulus. There are many models including the Elaborated Reichardt Detectors (van Santen & Sperling, 1984, 1985), motion energy models (Adelson & Bergen, 1985; Burr, Ross & Morrone, 1986), and scalar motion sensors (Watson & Ahumada, 1985) which are, despite of their different implementations, variants of the same abstract correlational model that essentially performs local spatiotemporal Fourier analysis. The second-order system is able to record a motion which has no overall directional component in the stimulus Fourier domain. One standard explanation is that the visual system performs a rectifying transformation on the visual input signal prior to subjecting it to the standard motion analysis. The coexistence of data revealing two diametrically different properties is interpreted as an evidence of two parallel movement processing systems. It is posited that motion computations are carried out in parallel by standard ("first-order" or Fourier) and non-standard ("second-order" or non-Fourier) computations that employ, in addition to the standard spatiotemporal Fourier analysis, rectification of the input signal. It was proposed that the second-order system may be, in turn, divided into two subsystems one of which is employing full-wave and the other half-wave rectification (Solomon & Sperling, 1994). In case of the half-wave rectification no correspondence will be established between elements with the opposite luminance polarity.

Time- and space-dependent direction reversal

It is almost inevitable that if Fourier and non-Fourier systems have even slightly different spatiotemporal sensitivity, the same stimulus configuration can be seen to move in two opposite directions dependently on the temporal and spatial separation between stimulus elements. Indeed, the most dramatic difference in predictions of Fourier and non-Fourier systems is in their predicted responses to the sequence of flashes with the opposite polarity. The standard Fourier computation that employs algebraic sign-sensitive multiplication predicts the direction reversal of the perceived movement and the non-standard computation that employs rectification prior to the extraction of

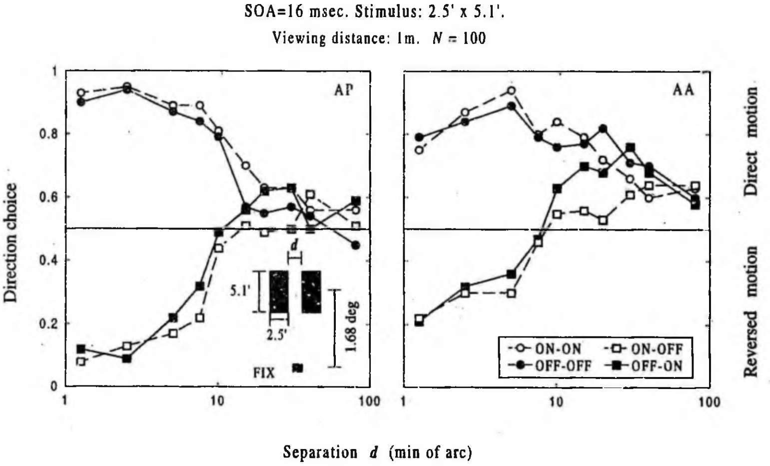


FIGURE 1. The motion direction identification probability as a function of separation d ($\log_{10}$ min of arc) and combination of polarity. There are four different types of flash sequences: ON-ON (two light flashes), OFF-OFF (two dark flashes), ON-OFF (light and dark flashes) and OFF-ON (dark and light flashes). Spatial configuration of stimuli is represented in the left panel. Each data point is an average of 100 trials. Data for two subjects: AP and AA.

movement information predicts no such reversal. If Fourier and non-Fourier systems are stimulated by the same sequence of opposite polarity flashes then they must contradict in their reports about motion directions: the motion is seen in the direction which of the two competing signals is stronger. Several space- and/or time-dependent direction reversal phenomena seem to support this competition hypothesis. There are at least four distinctive experimental situations in which the perceived movement direction reverses or changes with the change of stimulus temporal parameters:

(1) When the missing-fundamental square-wave grating is displaced one quarter cycle of the fundamental frequency to the left or right, the perceived movement direction depends on the time interval between two presentations. The subject saw consistently reversed direction at short ISI's, and motion mainly in the actual displacement direction at longer ISIs, with a cross-over in the region of 40 ms (Georgeson & Harris, 1990);

(2) Motion plaids composed of two superimposed sine-wave gratings at different orientation and moving in different direction appear to move in the vector sum direction for brief presentations and only approach the intersection of contraits direction after 150 msec (Yo & Wilson, 1992; Wilson & Kim, 1994). This change in the perceived movement direction was explained by an assumption that a time delay is necessary for the rectification of the input signal;

(3) The direction of apparent motion in a complex pattern comprising a static 1 cycle/degree and a moving 3 cycles/degree grating changes with stimulus duration. At short durations the perceived motion direction is opposite to actual displacement. At longer than about 150 msec durations, however, motion is seen almost veridically (Derrington & Henning, 1987; Henning & Derrington, 1988);

(4) Two-frame apparent motion stimulus made up of Gabor function micropatterns appears to move in the opposite direction from that of the actual displacement when stimulus onset asynchrony is short, whereas in the 'correct' direction for long asynchronies (about 120 msec) (Boulton & Baker, 1993).

Quite analogously, the direction reversal was reported to depend on distance between stimulus elemens. Anstis and Mather (1985) found, for example, that the same-polarity signals are always preferred over the opposite polarity signals when the displacement was less than 0.25 deg while for larger displacements, the movement between opposite-polarity signals was reported almost as often as same-polarity jumps. They concluded that short-range motion is sensitive to contrast polarity while the long-range motion is insensitive to contrast polarity. Analogously, fovea and periphery differ in their ability to resolve polarity of contrast. For example, a side-ways stepping, contrast-reversing grating appears to move in the reversed direction when viewed in central vision and in the stepping direction when viewed peripherally (Chubb & Sperling, 1989; see also Mather, Cavanagh & Anstis, 1985).

These direction reversal phenomena appear to suggest that the first-order (Fourier) system is operating over short spatiotemporal distances and the second-order (non-Fourier) system is operating over considerably larger spatial separation and longer temporal asynchronies. Typically, the postulated Fourier and non-Fourier motion systems are characterized by the use of different stimuli (Cavanagh& Mather, 1989) which always leaves open a possibility that it is a stimulus dichotomy rather than difference in underlying mechanisms (cf. Boulton & Baker, 1993). In this study, we used a very simple stimulus which is evidently processed by both systems but they differ in the predicted motion direction: Fourier system predicts that a sequence of opposite polarity flashes is perceived to move in the reversed direction, whereas non-Fourier system predicts that it is perceived to move in the direction of factual stimulus succession. We are presenting an evidence that the coexistence of the reversed and direct movement can be explained, in principle at least, by a single movement encoding mechanism with a complex spatiotemporal receptive field different parts of which exhibit properties of either Fourier or non-Fourier system.

METHODS

Stimuli

In the present study, two adjacent rectangular areas of a uniform luminous field increased or decreased their luminance in succession for a short period of time. In each trial the observer's task was in each trial to indicate in which direction the succession of two flashes appeared to move. There are four possible types of flash sequences: two homogeneous, (ON-ON and OFF-OFF) and two inhomogeneous (ON-OFF and OFF-ON) pairs of flashes. All four combinations of luminance changes were presented at a random order. The observer's responses were coded in terms of the factual direction choice probability, that is choosing the displacement direction from the position of the stimulus that appeared earlier towards the stimulus position presented later. Stimuli were generated on the screen of a VGA monochrome monitor. The background had the form of a rectangle 235 mm in width and 176 mm in height. The mean luminance of the background was L_b=24.6 cd/m^2 measured by an electronic photometer. The luminance increments and decrements consisted of two subsequent refreshes separated by a time interval of 16.2 msec. The temporal asynchrony between two flashes was varied by steps that were multiple to 16.2 msec. Before the beginning of the main experiment the detectability of the luminance increments and decrements was established. Typically, luminance decrements are more easily detected than increments. The amplitude of the luminance increment ΔL was 41.5 cd/m² and the amplitude of the luminance decrement was -24.3 cd/m². These amplitudes allow to detect both luminance changes with about 0.9 probability.

In order to obtain better temporal resolution, in one series a horizontal row of six adjacent rectangular segments of red light emitting diodes (*Quality Technology*) were used for stimulus presentation. Each rectangular element was 0.14° in height and 0.06° in width at the 2.0 m viewing distance and separated from their neighbours by a 0.02° gap. They all had the same steady state luminance L_0=30 cd/m². The luminance of each diode was controlled by a personal computer via six 12-bit digital-to-analog converters outputs. The duration of the brief rectangular increments and decrements ΔL from the steady reference luminance was 10 msec. The amplitudes of increment and decrement was determined on the basis of psychometric detection functions for each subject. This was done in order to ensure the equal visibility of luminance increment and decrement (about level of 0.9 probability) . The amplitude of the luminance increment and decrement ΔL was 37.8 cd/m² and -30.0 cd/m² (for subject AP) and 26.4 cd/m² and -22.5 cd/m² (for subject TT).

Procedure

Experiment is divided into three main series: investigation of spatial parameters on movement direction choice, investigation of temporal parameters on movement direction choice and investigation of retinal eccentricity on movement direction choice. All series are divided into sessions with randomized presentation of stimuli. Before the main experiment all subject participated in extensive trainig sessions.

Viewing was binocular with natural pupils. The observer's head was stabilized using a chin rest. The subjects could fixate at a small point which was located at 1.68 deg below stimuli in the monitor or in the display with LED.

The observer pressed one of the two buttons indicating in which direction, to the left or to the right, the sequence of two flashes appeared to move. No feedback was provided.

The experiment was performed in a semidarkened room. Each session was preceded by a 20 min period of dark adaptation.

Subjects

Three observers with normal (AA and TT) and corrected myopic vision (AP) participated in the experiment. Two observers, AA and TT, were naive concerning the purpose of this study.

RESULTS

Spatial separation

For the first series of measurements, which were aimed at studying the role of spatial separation between stimulus elements, two flashes with the temporal asynchrony of 16 msec were presented. Stimuli were presented in the center of the uniform background and the observer was instructed to fixate at a small point below stimuli (1.68 deg). Spatial separation (gap) *d* between two adjacent rectangular

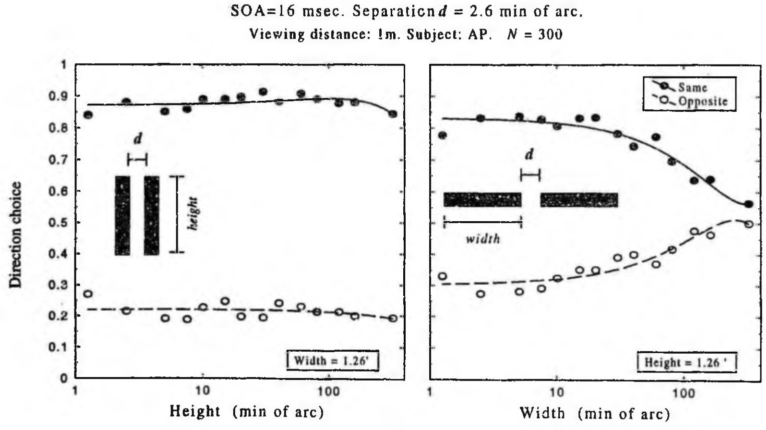


FIGURE 2. The motion direction identification probability as a function of stimuli height (left panel) or width (right panel). Separation ***d*** is fixed (2.6 min of arc). Results for homogeneous flashes (ON-ON or OFF-OFF, filled circles) and for inhomogeneous flashes (ON-OFF or OFF-ON, unfilled cirlces). Spatial configuration of stimuli used in this experiment is also represented . Each data point is an average of 300 trials. Data for subject AP.

areas of the size 0.042 by 0.085 degree was varied on 11 levels from 1 to 90 min of arc (for the stimulus configuration see inset in Figure 1). Viewing distance was 1 meter. Figure 1 shows results for two observers, AP and AA. The probability of the factual direction choice is plotted as a function of spatial separation ***d*** between two succeeding flashes for four types of contrast combination. The sequence of two same polarity flashes, ON-ON and OFF-OFF, is shown by unfilled and filled circles respectively. The sequence of two flashes of the opposite polarity, ON-OFF and OFF-ON, is shown by unfilled and filled squares respectively.

A general finding is that the direction identification performance slightly improves and after reaching an optimal value at about 3-5 min of arc starts to decrease beyond this optimal separation. These data are in good agreement with data reported particularly by Westheimer and McKee (1977) who also found 6′ separation optimal for the determination of temporal order of two adjacent stimuli.

Analysing Figure 1 two regularities can be noticed in these data:

(1) There was no difference between different types of homogeneous and inhomogeneous pairs of flashes, at small spatial separations at least. The movement direction identification in the sequence of two bright flashes (ON-ON) was as good as that in the sequence of two dark flashes (OFF-OFF). Analogously, the order of the inhomogeneous flashes, ON-OFF and OFF-ON, does not seem to affect the motion direction identification probability below 10 min of arc. On larger separations OFF-ON sequence was identified slightly better than ON-OFF sequence. In spite of these minor differences, results of this experiment are generally in agreement with the commutability principle according to which the reversal of order of flashes does not affect the motion direction identification probability (Allik & Pulver, 1995).

(2) There was an essential difference between homogeneous and inhomogeneous sequence of flashes. Two successive flashes with the same polarity were always perceived to move in the direction of the actual displacement, that is from the stimulus position presented earlier in time towards the stimulus position presented later. At variance from the homogeneous pair of flashes, two opposite-polarity flashes appeared to move in both directions dependently on their spatial separation. When the separation between two flashes was smaller than about 10 min of arc, the stimulus pattern appeared to move in the reversed direction. At about 10 min of

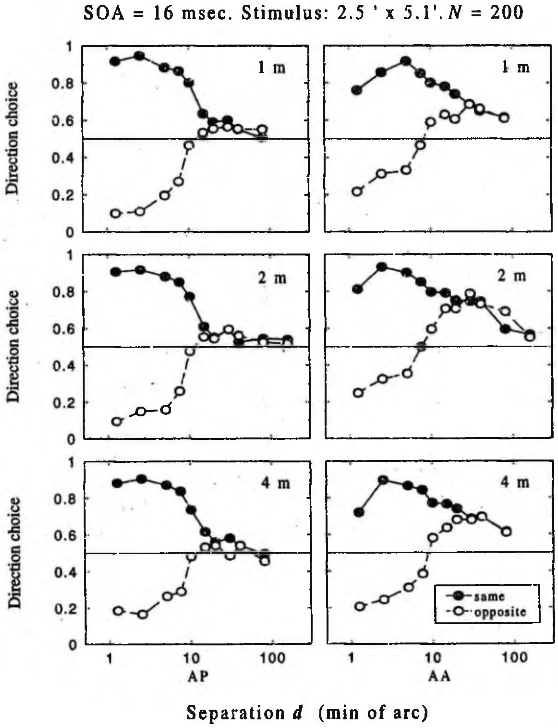


FIGURE 3. The motion direction identification probability as a function of separation ***d*** (log$_{10}$ min of arc) for three different viewing distances (1m , 2m and 4m). Angular size of stimuli is fixed. Results for homogeneous flashes (ON-ON or OFF-OFF, filled circles) and for inhomogeneous flashes (ON-OFF or OFF-ON, unfilled cirlces). Each data point is an average of 200 trials. Data for subjects AP (left panel) and AA (right panel).

arc a cross-over took place and at larger separations the sequence of the opposite polarity signals appeared to move in the actual, not reversed motion direction. At the cross-over region ON-OFF and OFF-ON flash sequences appeared to move with about equal probability in the factual and reversed motion direction so that on average they carry no information about whatever dispacement direction. These findings are in agreement with the reported data. Van Santen and Sperling (1984) obtained a clear direction reversal with the opposite polarity pulses in the conditions where spatial separation between adjacent bars was zero and temporal asynchrony was 30 msec. In all other studies, where direct, not reversed movement is reported either spatial separation or temporal asynchrony was considerably larger (Shechter & Hochstein, 1990).

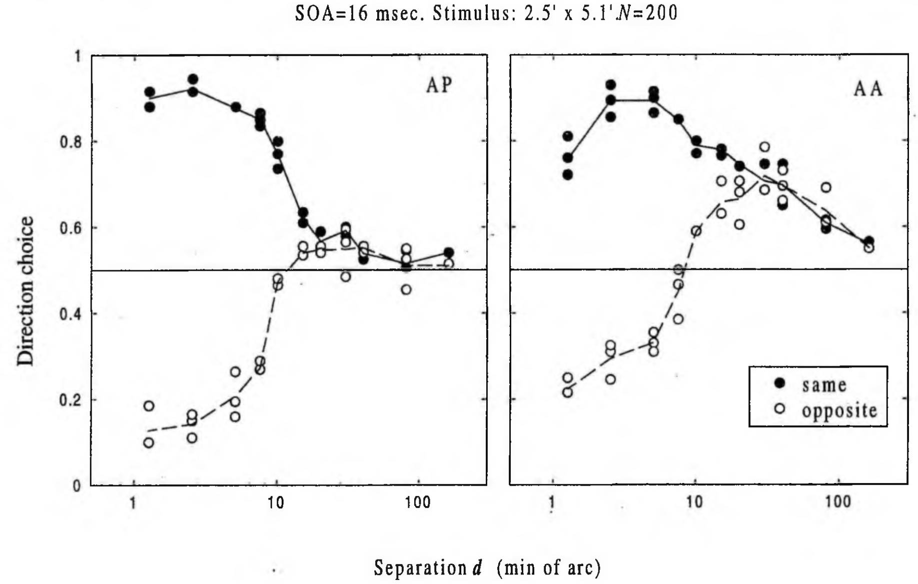


FIGURE 4. The motion direction identification probability as a function of separation d ($\log_{10}$ min of arc). Data from Figure 3. The lines connect points of average probability over three viewing distances. Results for homogeneous flashes (ON-ON or OFF-OFF, filled circles) and for inhomogeneous flashes (ON-OFF or OFF-ON, unfilled cirlces). Left panel (subject AP) and right panel (subject AA).

With a fixed size of stimulus elements it is impossible to tell which stimulus parameter actually predicts curves shown in Fig. 1. It is possible that not the gap between two adjacent elements but their center-to-center distance, for example, is essential for the dependence plotted in Fig. 1. In additional experiment, we varied either width or height of the stimulus element, while the gap between rectangular elemets was fixed at 0.043 deg (about two and half minutes). The subject is instructed to fixate the center of **display**.

Figure 2 (left panel) shows the results for 0.021° width stripe which height was systematically varied. For both types of flash sequences (the same and opposite polarity) the vertical length of the stimulus did not affect the motion direction identification probability. This result is in good accordance with data reported by Westheimer and McKee (1977, Fig. 2) who found no changes in the threshold for the temporal order of two lines separated by 4′ as a function of their length. The width of stimulus elements (Fig. 2, right panel) did not affect remarkably the performance until it was below of about 15-20 min of arc, after that limit the identification probability started to fall approching virtually chance level when two 3° horizonal strips were presented. Because the cross-over from the reversed to direct movement for the opposite polarity flashes happened at about 9-10′ we can conclude that the gap (edge-to-edge separation), not the center-to-center separation predicts the position of function represented in Fig. 1. These findings are in agreement with other data (Shechter, Hochstein & Hillman, 1988; Grossberg & Rudd, 1992) demonstrating that the distance between moving elements is measured in terms of the spatial gap (edge-to-edge), not in terms of their center-to-center distance. On the basis of these findings we will present all spatial separation measures in terms of the spatial gaps.

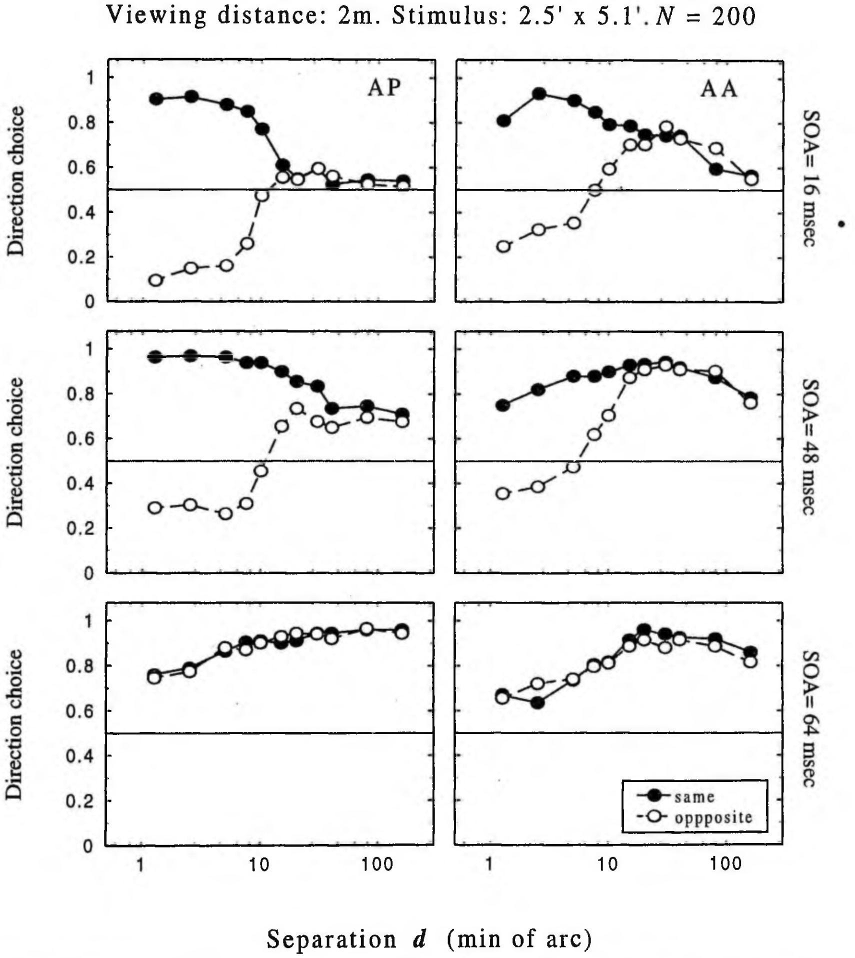


FIGURE 5. The motion direction identification probability as a function of separation ***d*** (log10 min of arc) for three different SOA (16 msec, 48 msec and 64 msec). Results for homogeneous flashes (ON-ON or OFF-OFF, filled circles) and for inhomogeneous flashes (ON-OFF or OFF-ON, unfilled cirlces). Each data point is an average of 200 trials. Data for subjects AP (left panel) and AA (right panel).

Viewing distance

Although results of the previous section demonstrated that the distance between stimulus elements is measured in terms of the edge-to-edge separation, it is unclear whether this separation is defined in absolute metrical or retinal distance units. In this experiment we varied the viewing distance (1, 2, and 4 m) compensating the size of stimulus configuration appropriately for each viewing dictance. This means that in terms of angular size all stimulus configurations were identical (elements were 5.1′ in height and 2.5′ in width). Results of this experiment are shown in Figure 3 for subjects AP (left panel) and AA (right panel). The data for viewing distance of 1m are the same that in the Figure 1. All curves corresponding to different viewing distances look very similar. Figure 4 shows the same data, but three separate sets of data are now superimposed forming virtually one single functional relation for the same (filled circles) and opposite (unfilled circles) polarity pairs of flashes respectively. This means that the viewing distance is irrelevant and the retinal distance between stimulus elements is the main factor determining the motion perception.

Onset asynchrony

With a fixed stimulus element size (2.5′ x 5.1′) and viewing distance (2 m), we next varied stimulus onset asynchrony (SOA) between two successive flashes. Spatial separation (gap) ***d*** between two adjacent rectangular areas was varied on 11 levels from 1 to 161 min of arc and SOA on 3 levels: 16, 48 and 64 msec. Figure 5 shows the results for two observers and three SOA values. Two regularities can be noticed in these data: (1) With the increase of SOA the optimal spatial separation at which the motion direction was identified with the highest probability, shifts towards larger spatial separations. At short asynchronies (16 msec) the optimum is about 5 min of arc and shifts clearly above 10 min of arc at 48 and 64 msec. (2) With the increase of the temporal asynchrony the reversed motion created by the opposite polarity flashes starts to disappear. When SOA was 48 msec it was clearly attenuated compared with the shortest 16 msec asynchrony. At the 64 msec asynchrony the difference between homogeneous (the same polarity) and heterogeneous (the opposite polarity) flash sequences disappeared completely: all pairs of flashes were perceived to move in the direction of the stimulus actual succession irrespective of their polarity combination.

The used display allowed to vary SOA only with 16 msec steps. In order to obtain better temporal resolution we reproduced the basic stimulus configuration with the help of LED display. Figure 6 demonstrates the direction choice probability as a function of SOA for three different spatial separations (***d***=1.4, 6.2, and 11 min of arc). SOA was varied on 11 levels from 15 to 165 msec. The results confirm our previous results: (1) The reversed motion can be perceived only when both spatial separation and temporal asynchrony between two opposite polarity flashes are small; (2) In the temporal domain the crossover from the reversed to the direct motion direction happens at about 40 msec; (3) In the spatial domain the cross-over point is about 6 min of arc which is slightly smaller than was obtained with the raster display. Thus, the direction reversal of the

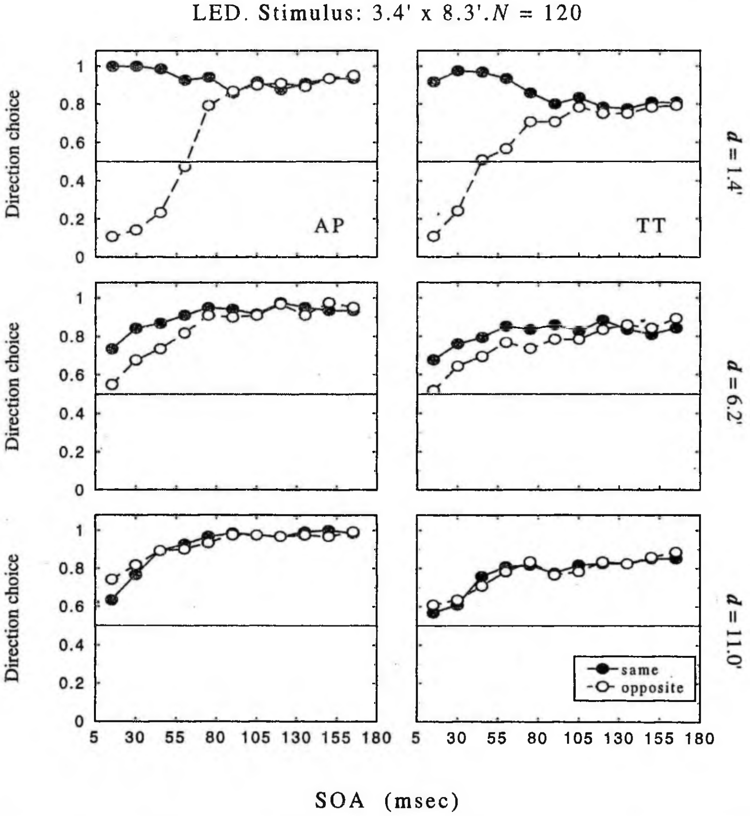


FIGURE 6. The motion direction identification probability as a function of SOA (msec) for three different separations ***d*** (1.4 , 6.2 and 11.0 min of arc). LED display. Viewing distance: 2 m. Results for homogeneous flashes (ON-ON or OFF-OFF, filled circles) and for inhomogeneous flashes (ON-OFF or OFF-ON, unfilled cirlces). Each data point is an average of 120 trials. Data for subjects AP (left panel) and TT (right panel).

opposite polarity flashes was observed only in a very limited spatiotemporal range.

Retinal eccentricity

In this series of measurements we varied retinal eccentricity of the stimulus configuration. The sequential pair of flashes was presented quasirandomly on the left or on the right from fixation point at 2.10, 3.03, 5.32 and 10.21 degrees (for the stimulus configuration see inset in Figure 7A and 7B). Spatial separation ***d*** between two adjacent rectangular areas of the size 0.08 by 0.17 degree was varied on 11 levels from 1 to 60 min of arc. Viewing distance was 0.5 m and SOA=16 msec.

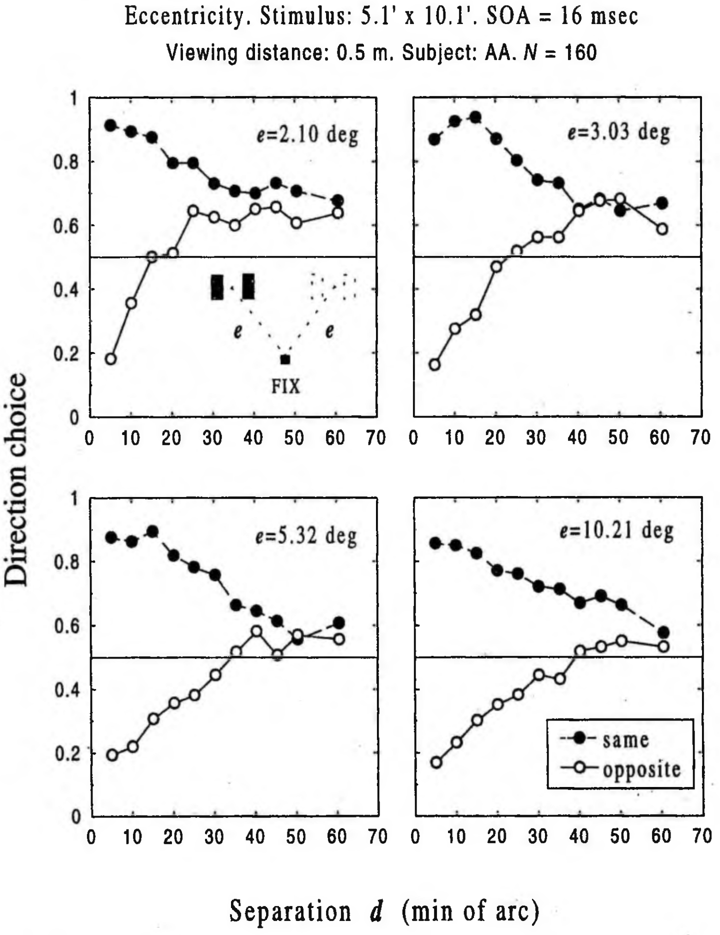


FIGURE 7A. The motion direction identification probability as a function of separation ***d*** (min of arc) for four different retinal eccentricities (2.10, 3.03, 5.32 and 10.21 deg). Eccentricity ***e*** is defined as the angular distance between the fixation point and the center of separation ***d*** . Results for homogeneous flashes (ON-ON or OFF-OFF, filled circles) and for inhomogeneous flashes (ON-OFF or OFF-ON, unfilled cirlces). Each data point is an average of 240 trials. Data for the subject AP.

Figure 7A and 7B demonstrates the motion direction choice probability as a function of the spatial separation ***d*** between two flashes at four different eccentricities. Even a preliminary inspection reveals that the cross-over point from the reversed to the direct motion for the opposite polarity flashes shifts to the right with the increase of retinal eccentricity. This tendency is more conspicuous in Figure 8 (B, the right panel) where critical separation $d_{0.5}$ (a separation at which 0.5 choice probability is reached) is plotted as a function of retinal eccentricity for the observer AP (filled circles) and AA (unfilled circles). In logarithmic coordinates $d_{0.5}$ is approximately a linear function of retinal eccentricity. As we already know from previous measurements the identification of motion direction becomes progressively worse with the spatial separation between flashes. However, contrary to the opposite polarity flashes, the direction identification versus spatial separation curves appear to change only a little with eccentricity. In order to make numerical comparisons we determined the spatial separation $d_{0.75}$ at which the motion direction of two same polarity flashes was identified in 75% of all cases. This dependence is shown in Figure 8 (A, the left panel) for two observers, AP and AA

This difference needs a few words for comments. At spatial separations above about 50′ the difference between the same and opposite polarity flashes virtually disappears. On average, they both appear to move in the direction of actual stimulus succession but the probability with which it can be identified is only slighly above the chance level. This indicates that the same genaral rule is applicable to both types of signals, homogeneous and heterogeneous pairs of flashes: there is an optimal spatial separation between two flashes to see them in motion and beyond this optimal separation the extraction of motion information becomes progressively less efficient.

Within this general tendency the opposite polarity flashes change their perceived motion direction. The spatial separation at which two opposite polarity signals are changing their perceived motion direction is scaled in terms of retinal eccentricity. If we normalize spatial separation at each retinal eccentricity by the respective $d_{0.5}$ value, we can present all the functions in terms of normalized distance. We compute an empirical scaling function for all our data including results of series with eccentricity 1.68 degree and normalize all our data to this eccentricity. Indeed, Figure 9 demonstrates that all curves corresponding to the opposite polarity flashes seen at different retinal eccentricities come together to form one single functional relation.

The increase of the spatial scale accompanied by the retinal eccentricity is usually interpretated in terms of cortical magnification factor (M-factor). Cortical magnification refers to the correspondence between some measure of the amount of cortical substance and some spatial measure of visual field (Daniel & Whitteridge, 1961; Cowey & Rolls, 1974; Drasdo, 1991). Most researhers used linear cortical magnification which is defined as the distance on striate cortex surface in millimeters that corresponds to one degree of visual angle. Variation of the cortical representation factor is caused first of all by a variation of retinal ganglion cells density with eccentricity (Drasdo, 1977). The exact estimation of retinal cells densities is, however, a complicated task, especially in foveal area, which leads to quite different estimations (e.g. Schein & DeMonasterio, 1987; Schein, 1988; Wässle et al., 1990). In spite of the certain degree of controversy, the concept of cortical magnification factor has been a useful concept accountig for changes in performance accompanied with the increase of eccentricity in various psychophysical tasks including acuity (Cowey & Rolls, 1974), contrast sensitivity (Virsu & Rovamo, 1979), lower threshold of motion (Johnston & Wright, 1983;

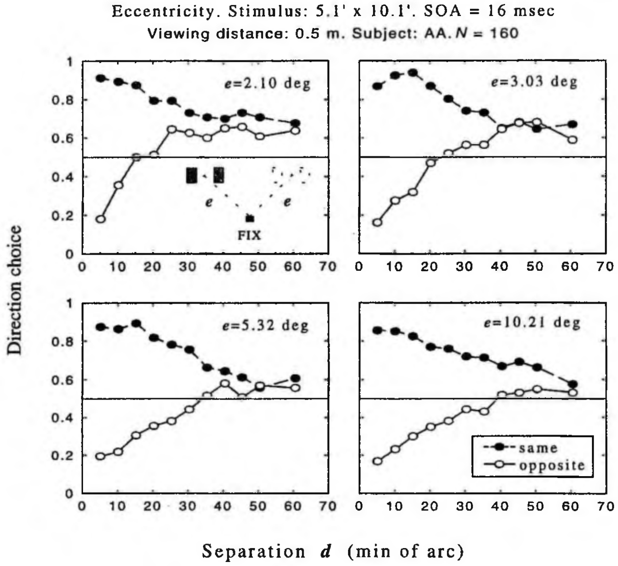


FIGURE 7B. The motion direction identification probability as a function of separation d (min of arc) for four different retinal eccentricities (2.10, 3.03, 5.32 and 10.21 deg). Eccentricity e is defined as the angular distance between the fixation point and the center of separation d . Results for homogeneous flashes (ON-ON or OFF-OFF, filled circles) and for inhomogeneous flashes (ON-OFF or OFF-ON, unfilled cirlces). Each data point is an average of 160 trials. Data for the subject AA.

1985) and detection of coherent motion in dynamic random dot patterns (van der Grind et al., 1983).

Typically, the gradient of some visual performance with retinal eccentricity is compared with the inverse cortical magnification which rises approximately linearly with eccentricity. Such a linear function of the inverse cortical magnification (M^{-1}) would be predicted by the complex logarithmic mapping function (Schwartz, 1977; 1980). The basic idea behind the concept of the cortical magnification is that visual performance is homogenous everywhere across visual field if the stimuli are appropriately scaled. But not all psychophysical measures such as vernier acuity (Westheimer, 1982), relative motion (Levi et al., 1984; McKee & Nakayama, 1984) and the spatial separation threshold for fine-grain movement (Foster et al., 1989) are invariant with M-scaling. One reason for these discrepancies is the division of visual system into two separate anatomical pathways, magno- and parvocellular systems (Livingstone & Hubel, 1987), with their own cortical magnification factors (Drasdo, 1989). Estimating the density of A and B types of ganglion cells (Leventhal et al., 1981), which differ from each other anatomically and physiologically corresponding to magno- and parvocellular systems respectively, Drasdo (1989) computed the gradient of cortical magnification for parvocellular and magnocellular systems. The main result was that magnocellular inverse magnification M^{-1} rises approximately **four times** slower with eccentricity than the parvocellular one. It is important to notice that the most popular cortical magnification estimation used in psychophysical studies (Rovamo & Virsu, 1979; Virsu & Rovamo, 1979) has a gradient which lays approximately midway between parvocellular and magnocellular ones. Magnocellular system exhibits systematic selectivity for movement and projects to the middle temporal area (MT), an extrastriate area specialized for the motion analysis (Maunsell & Newsome, 1987). For example, MT neurons maintain direction selectivity over spatial intervals between successive flashes that are, on the average, three times as large as those for V1 (striate cortex) neurons with a tendency to increase slightly with eccentricity (Mikami, Newsome & Wurtz, 1986).

The separation between these two systems, while pronounced, is not compelete. Therefore different psychophysical tasks probably tap both systems to a different extent and the proportion of activation between these two systems does not neccesarily remain constant at different retinal eccentricities. In Figure 10 the performance in various psychophysical tasks normalized by the perfomance demonstrated at the eccentricity of 2 degrees is shown as a function of retinal eccentricity. The following empirical results are reproduced: values of $d_{\min}$ and $d_{\max}$ averaged over all subjects from Baker and Braddick (1985, Figure 3, p.806) in range of eccentricity 0–10 deg; spatial thresholds of fine-grain movement illusion (FGM) and minimal angle of resolution (MAR) averaged over all subjects from Foster, Gravano & Tomoszek (1989, Figure 3, p. 1024) in range of eccentricity 5-25 deg; upper and lower limits of spatial range for good temporal order detection from Westheimer (1983, Figure 3, p.762) in range of eccentricity 0-20 deg; and minimal angle of resolution (MAR) from Westheimer (1982, Figure 2, p. 160) in range of eccentricity 0-10 deg. These experimental data are placed into the

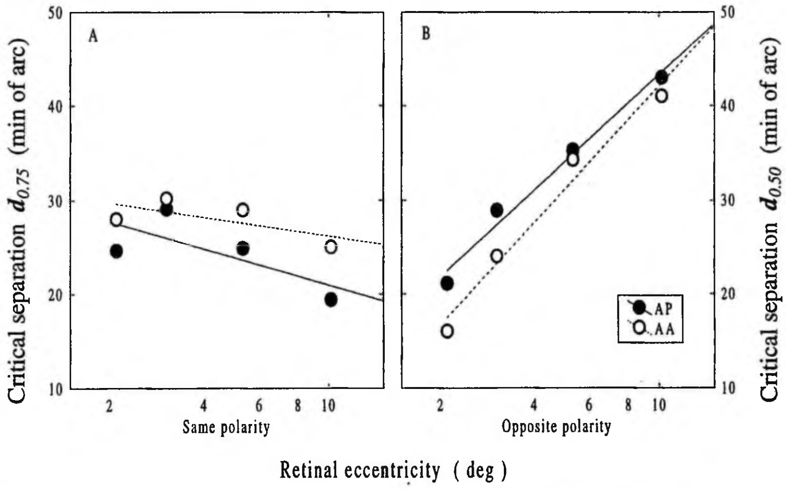


FIGURE 8. Critical spatial separation as a retinal eccentricity. Left panel: critical spatial separation for homogeneous flashes. Right panel: critical spatial separation for inhomogeneous flashes. Data for the subjects AP (filled circles) and AA (unfilled circles). See detailes in the text.

coordinates of the magnification gradient for magnocellular and parvocellular systems (Drasdo, 1989, Figure 2, p. 986) which are shown by continous and dotted curves respectively. The most popular Rovamo-Virsu M-function, used in many psychophysical studies for the scaling of visual stimuli, is also shown by a broken curve. In general, gradient of $d_{0.5}$ value is very close MAR (Westheimer, 1982), the lower and upper spatial limits for temporal resolution (Westheimer, 1983), and $d_{\min}$ for the discrimination of movement direction in random-dot patterns (Baker & Braddick, 1985). It is clear, however, that the gradient of performance for FGM (Foster, Gravano & Tomoszek, 1989) is less steep and the gradient of $d_{\max}$ is, in turn, much more steep than the distance $d_{0.5}$ at which the reversed motion switches over to the direct motion. It is also remarkable that, contrary to usual claims, the performance gradients for visual acuity-type tasks are not very different from those that can be observed in motion discrimination tasks. In general, the relative scale factor for $d_{0.5}$ is approximately in the range constrained by the parvocellular gradient from above and magnocellular gradient from below. This means, in particular, that it can be possible to find statistical estimation for the upper limit of contrast-polarity sensitive computation in cortex.

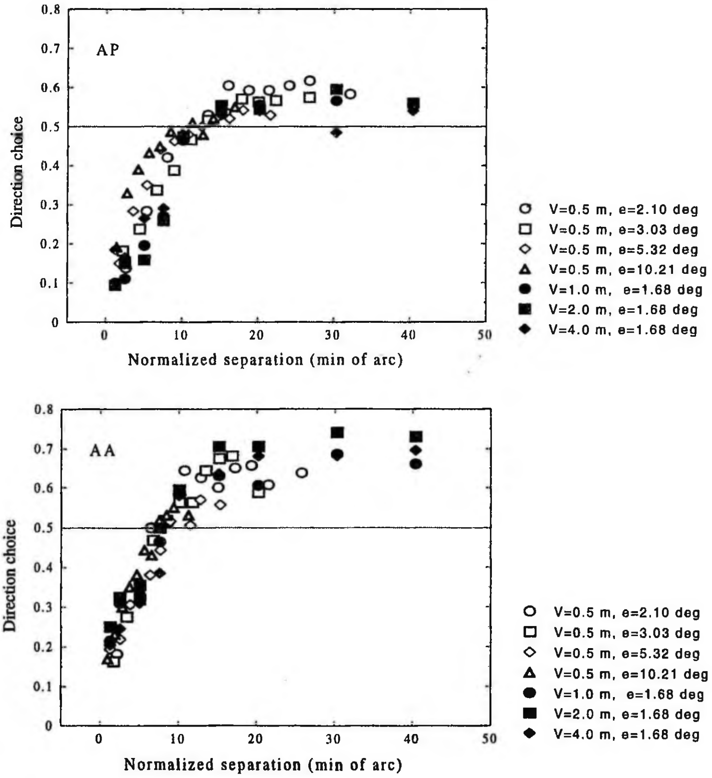


FIGURE 9. Normalized spatial separation for opposite polarity stimuli. All spatial separations are normalized to eccentricity of 1.68 degree by empirical scaling factor. Data from Figure 3, Figure 7A and 7B for the subject AP (upper panel) and for the subject AA (lower panel). See detailes in the text.

We transform our empirical $d_{0.5}$ values to cortical distances using estimates of M from monkeys and humans. M can be predicted from the equation (Dow et al., 1981; Levi et al., 1985, Van Essen et al., 1984; Drasdo, 1991):

$$M = (1/k)*(E+E_2)^{-1},$$

where E is the eccentricity of the stimulus in deg, E_2 is the eccentricity in deg at which cortical magnification has dropped to $0.5*M_f$ (the foveal magnification factor in mm/deg), and $k = (M_f*E_2)^{-1}$. For the rhesus monkey the following parameters have been reported: $k = 0.12$ mm^{-1}, M_f=10.4 mm/deg and E_2=0.80 deg . The human striate cortex can be considered to be a scaled up version of the monkey striate cortex with a linear scaling factor approximately of 1.6 (Drasdo, 1991). On the basis of this consideration we scale up all computed values by factor 1.6. Values of $d_{0.5}$ for data of viewing distance series (Figure 3) and eccentricity series (Figure 7A and 7B) yield the following estimates. The values of $d_{0.5}$ in cortical distance are between .9 and 1.7 mm with an average of $1.2 \pm 0.3_{0.95}$ mm (subject AP) and between .7 and 1.5 mm with an average of $1.0 \pm 0.3_{0.95}$ mm (subject AA). Thus by a very crude estimation the contrast-polarity sensitive computation extends no longer than about 1.3-1.5 mm of the cortical distance. If we take into consideration the approximate size of hypercolumns in the monkey primary visual cortex (Hubel, Wiesel, 1977) this results may be interpreted as meaning that contrast-polarity sensitive computation is carried out only within elementary functional cortical units.

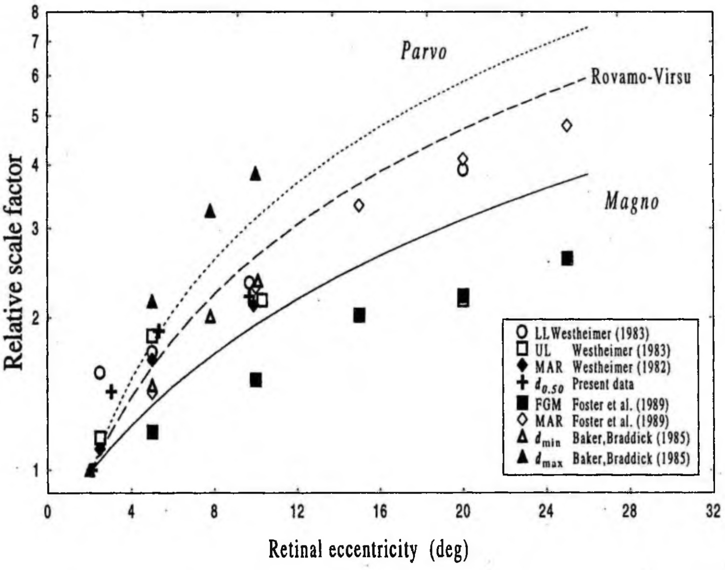


FIGURE 10. Relative scale factor. Ordinate: computed relative scale factor ($\log_{10}$); abscissa: retinal eccentricity (deg). Besides our empirical values of $d_{0.5}$ (averaged over two subject) the following empirical results are reproduced: values of d_{min} and d_{max} averaged over all subjects from Baker and Braddick (1985, Figure 3, p.806) ; spatial thresholds of fine-grain movement illusion (FGM) and minimal angle of resolution (MAR) averaged over all subjects from Foster, Gravano & Tomoszek (1989, Figure 3, p. 1024); upper (UL) and lower (LL) limits of spatial range for good temporal order detection from Westheimer (1983, Figure 3, p.762); and minimal angle of resolution (MAR) from Westheimer (1982, Figure 2, p. 160). The magnification function for magnocellular and parvocellular systems also as Rovamo-Virsu M-function are shown (from Drasdo, 1989, Figure 2, p. 986). All relative scale factors are computed in the mean of performance in 2 degree of retinal eccentricity.

GENERAL DISCUSSION

Basic findings

The principal finding of this study is that the polarity-sensitive operations extend over only limited spatial and temporal separation. This limit is about 40 msec in the temporal domain and about 10 min of arc of foveal distance. Because this spatial limit is linearly scaled with the respect of retinal eccentricity, it is possible to assume that the limit is imposed in terms of a fixed cortical distance: the polarity of two compared signals is taken into account only if their representation is not beyond 1.3-1.5 mm of cortical distance. Beyond this limit the polarity of signals is discarded. All these results can be summarized in a plot which shows the probability of the direction choice as a two-dimensional function of the cortical distance and asyncrony between two opposite polarity flashes (Figure 11). The approximating function was computed on the basis of all data presented in Figures 3, 5, 6, and 7. Light areas represent spatiotemporal separation between two opposite polarity flashes which produces the reveresed motion and dark areas correspond to the direct motion. There is only a small lower-left corner in which a clear impression of the reversed motion can be evoked. It is interesting to notice that contours of isoprobability for the observer AP are obliquely oriented, with a negative slope. This correlation means that his responses to two opposite-polarity flashes as a function of the spatial and temporal separation between the two flashes are not space-time separable. Alternatively, the response function shown for the observer AA is much more circular or elliptical which suggests that her responses are more clearly separable in space-time.

These findings are in an agreement with several previous studies and can provide a simple explanation for several known perceptual phenomena. In particular:

1.It explains why stimulus patterns, composed from elements of both positive and negative contrast, have a tendency to perceive moving in the reversed direction when the temporal interval between their presentation is small and in the direction of actual presentation succession when they are considerably separated in time;

2. It explains why closely spaced opposite polarity flashes produce the impression of reversed motion (van Santen & Sperling, 1984) and the impression of the direct motion if the separation between opposite polarity flashes is beyond a critical limit (Shechter & Hochstein, 1990; Wehrhahn & Rapf, 1992).;

3. It also explains why a side-ways stepping, contrast reversing grating appears to change direction when viewed at different distances or in periphery (Chubb & Sperling, 1989). Because the critical area for the reversed motion is defined in retinal coordinates a stimulus which is out of this range can be brought into the limit by the increase of viewing distance and the corresponding reduction of retinal size. Analogously, a stimulus that appears to move in the direction of the stimulus actual succession when it is viewed foveally reverses its direction and appears to move in the reversed direction when it is viewed peripherally. It is explained by the fact that the range of the reversed motion $d_{0.5}$ is larger in periphery than in the central vision.

The limited spatiotemporal range of operations taking into account the polarity of signals is not limited with movement computations. For example, it was found that vernier acuity with stimuli of opposite-contrast polarity in which one target was brighter than the background, and the other was darker, is about half as good as with the same polarity stimuli. At extremes of spatial separation between targets, acuities of the same and opposite polarity targets became similar; mainly because the same polarity acuities degrade to the level of that of the opposite contrast stimuli. Thus, like motion perception, the finest judgments of spatial position seem to arise within a very limited spatiotemporal range in which light and dark stimuli are treated independently (O'Shea & Mitchell, 1990).

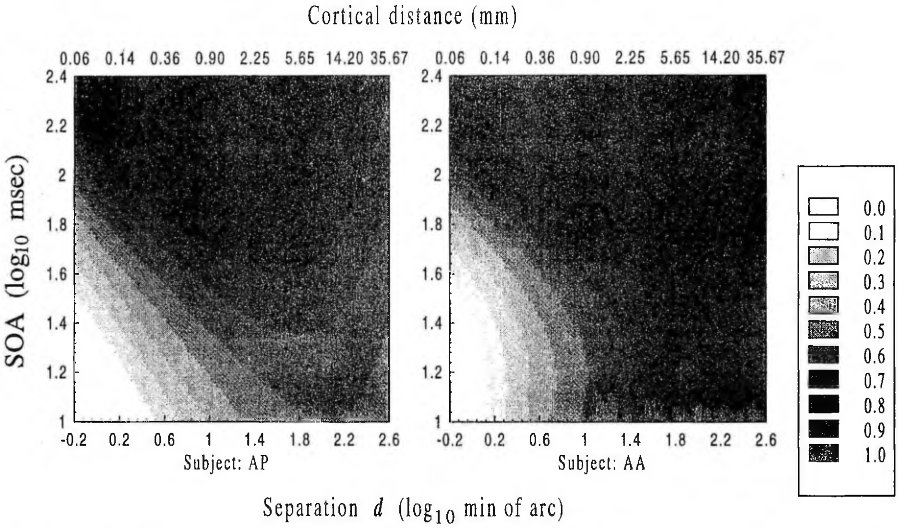


FIGURE 11. Isoprobability contour plot for motion direction identification from spatial separation and temporal asynchrony. The probability values are shown on the legend. The approximating splain function was computed on the basis of all data presented in Figures 3, 5, 6, and 7. Light areas represent spatiotemporal separation between two opposite polarity flashes which produces the reveresed motion and dark areas correspond to the direct motion. Cortical distance (values for eccentricity of 1.68 degree) which corresponds to spatial separation d is also shown. Data for subjects AP (left panel) and AA (right panel).

Implications for the models of motion analysis

Results of this study have some implications for the models of motion analysis. All models proposing that the squaring or rectifying operations as an operation that is invariably applied to the input signal cannot be regarded as sufficiently general (Heeger, 1987, 1991; Werkhoven, Sperling & Chubb, 1993). The direction reversal demonstrated in this and many other studies indicates that at least on some occasions the perceived motion direction is determined by the algebraic sign-sensitive multiplication. For the same reason the Reichardt model, the original (Reichardt, 1957) or the elaborated one (van Santen & Sperling, 1984), cannot be regarded as a general model of the movement perception: there are many cases when the perceived movement direction is opposite to that predicted by the cross-correlation product (cf. Chubb & Sperling, 1988). In search of a compromise, it can be proposed that there are two separate and parallel pathways in the visual system, one for Fourier and the second for non-Fourier motion signals. The non-Fourier pathway, in turn, may be comprised of two separate subchannels that employ full-wave and half-wave rectification respectively (Solomon & Sperling, 1994).

The half-wave rectification assumes that positive and negative contrast signals are kept separate by two tandem pathways without considerable cross-talk across the polarity. Again, even if the half-wave rectification is used in the movement system it is not a universal principle of movement processing. This study demonstrates that two opposite polarity flashes can generate a very clear and consistent impression of motion although its strength may be weaker compared with the motion signal produced by the same polarity signals. Dependently on the spatiotemporal separation between the opposite polarity flashes motion is perceived either in the direction of stimulus element succession or in the direction of stimulation succession as it is predicted by the Reichardt model. In both cases, however, a considerable cross-talk across the polarity of contrast must be assumed. Another possibility is to propose that rectification or any other operation destroying the information about polarity needs time to develop (Wilson & Kim, 1994). Indeed, in at least five different experimental paradigms the direction reversal is obtained when time between two samples is increased. This can be explained by a time delay elapsed before rectification starts to operate. Even if the delay of rectification exists it is not a universal principle in motion computations. This study demonstrates that if such a delay exists it is applicable only for short spatial separations between sample points. When two distant locations are compared the rectification starts without any delay: all signals irrespective of their temporal asynchrony and luminance polarity are perceived to move in the direct, not reversed direction. Thus, none of the above-mentioned operations – sign-sensitive algebraic multiplication, rectification, and delay of rectification – belongs to the category of computational procedures that are invariably employed in all movement computations.

Single or multiple systems

This fact that specific computations are selectively applied to particular spatiotemporal regions provide a temptation to postulate multiple parallel movement analysing systems differing from one another by the underlying computational processes. Although the very popular dichotomy between short-range and long-range motion phenomena demonstrates signs of recession because it primarily reflects difference in stimuli used rather than underlying processes (cf. Cavanagh and Mather, 1989), the same illusive logic is in the basis of separation of Fourier and non-Fourier systems. The basic distinction between these two systems is based on the construction of the special drift-balanced stimuli which are supposed to stimulate only non-Fourier system or Fourier system. Unfortunately, even an abrupt reversal of the perceived movement direction observed in this study can not tell us whether **univariant** stimulus attribute mediates discrimination or judgements are made on the basis of multiple representations (cf. Palmer, 1986). On the basis of our present data alone it is impossible to decide between these two options. According to the univariant representation hypothesis two pairs of flashes that are equivalent in performance remain equivalent after any psychological manipulation, adaptation to movement for example. The existence of the switch-over from the reversed to actual motion direction with the increase of spatiotemporal distance between flashes cannot be regarded as an evidence of two parallel systems. It is possible to regard the response function such as shown in Figure 11 as a representation of a complex spatiotemporal receptive field of a hypothetical cortical neuron which demonstrates Fourier-like properties near the origin of the coordinate system and non-Fourier-like properties on its periphery.

ACKNOWLEDGEMENTS

This research was supported by a James S. McDonnell Foundation grant and Estonian Science Foundation grant #450 to Jüri Allik.

REFERENCES

Adelson, E.H, & Bergen, J.R. (1985). Spatio-temporal energy models for the perception of motion. *Journal of the Optical Society of America*, A2, 284-299.

Allik, J. (1992). Resolving ambiguities in orientation, motion, and depth domains. *Perception*, 21, 731-746.

Allik, J., Dzhafarov, E. N., Houtsma, A. J. M., Ross, J., & Versfeld, N. J. (1989). Pitch motion with random chord sequences. *Perception and Psychophysics*, 46, 513-527.

Allik, J., & Dzhafarov, E. N. (1984). Motion direction identification in random cinematograms: A general model. *Journal of Experimental Psychology Human Perception and Performance*, 10, 378- 393.

Allik, J. & Pulver, A. (1995). Contrast response of movement encoding system. *Journal of the Optical Society of America* (in press).

Anstis, S. M. (1970). Phi movement as a subtraction process. *Vision Research*, 10, 1411-1430.

Anstis, S.M. (1978). Apparent movement. In Held, R., Leibowitz, H.L. & Teuber, H.L. (Eds), *Handbook of sensory physiology.*. Vol. 8: Perception (pp. 656-673). Berlin: Springer.

Anstis, S.M. (1986). Motion perception in the frontal plane. In Boff, K.R., Kaufman, L. & Thomas, J.P. (Eds), *Handbook of perception and human performance*. (Vol. 1). New York: Wiley.

Anstis, S.M. & Mather, G. (1985). Effect of luminance and contrast on direction of apparent motion. *Perception*, 14, 167-179.

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

Baker, Jr., C.L. & Braddick, O.J. (1985) Eccentricity-dependent scaling of the limits for short-range apparent motion perception. *Vision Research*, 25, 803-812

Boulton, J.C. & Baker, C.L. (1993). Dependence on stimulus onset asynchrony in apparent motion: evidence for two mechanisms. *Vision Research*, 33, 2013-2019

Buchner, E. (1976). Elementary movement detectors in an insect visual system. *Biological Cybernetics*, 24, 85-101.

Burr, D.C., Ross, J. & Morrone, M.C. (1986). Seeing object in motion. *Proceedings of the Royal Society of London*, B227, 249-265.

Burton, G. J., Nagshineh, S., & Ruddock, K. H. (1977). Processing by the human visual system of the light and dark contrast components of the retinal image. *Biological Cybernetics*, 27, 189-197.

Cavanagh, P. & Mather, G. (1989). Motion: The long and short of it. *Spatial Vision*, 4, 103-129.

Chubb, C., & Sperling, G. (1988). Drift-balanced random stimuli: a general basis for studying non-Fourier motion perception. *Journal of the Optical Society of America*, A5, 1986-2007.

Chubb, C. & Sperling, G. (1989). Two motion mechanisms revealed through distance-driven reversal apparent motion. *Proceedings of the National Academy of Science, U.S.A.*, 86, 2985-2989.

Chubb, C., & Sperling, G. (1991). Texture quilts: Basic tools for studying motion-from-texture. *Journal of Mathematical Psychology*, 35, 411-442.

Cowey, A. & Rolls, E.T. (1974) Human cortical magnification factor and its relation to visual acuity. *Experimental Brain Research*, 21, 447-454.

Daniel, P.M. & Whitteridge, D. (1961) The representation of the visual field on the cerebral cortex in monkey. *Journal of Physiology*, London, 159, 203-221.

Derrington, A.M. & Henning, G.B. (1987). Errors in direction-of-motion discrimination with complex stimuli. *Vision Research*, 27, 61-75.

Dow, B.M., Snyder, R.G., Vautin, R.G. & Bauer, R. (1981). Magnification factor and receptive field size in foveal striate cortex of the monkey. *Experimental Brain Research*, 44, 213-228.

Drasdo, N. (1977) The neural representation of visual space. *Nature*, 266, 554-556.

Drasdo, N. (1989) Receptive field densities of the ganglion cells of the human retina. *Vision Research*, 29, 985-988.

Drasdo, N. (1991). Neural substrates and threshold gradients of peripheral vision. In Kulikowski, J.J., Murray, I.J. & Walsh, V. (Eds), *Vision and visual dysfunction* (Vol. 5, pp. 250-264). New York: Macmillan.

Edwards, M. & Badcock, D.R. (1994). Global motion perception: Interaction of the ON and OFF pathways. *Vision Research*, 34, 2849-2858.

Emerson, R.C., Citron, M.C., Vaughn, W.J., & Klein, S.A. (1987). Nonlinear directionally selective subunits in complex cells of cat striate cortex. *Journal of Neurophysiology*, 58, 33-65.

Foster, D.H., Gravano, S. & Tomoszek, A. (1989). Acuity for fine-grain motion and for two-dot spacing as a function of retinal eccentricity: differences in specialization of the central and peripheral retina. *Vision Research*, 29, 1017-1031.

Georgeson, M.A. & Harris, M.G. (1990). The temporal range of motion sensing and motion perception. *Vision Research*, 30, 615-619.

Green, M. (1989). Color correspondence in apparent motion. *Perception and Psychophysics*, 45, 15-20.

van der Grind, W.A., van Doorn, A.J. & Koenderink, J.J. (1983) Detection of coherent movement in peripherally viewed random-dot patterns. *Journal of the Optical Society of America*, 73, 1674-1683.

Grossberg, S. & Rudd, M.E. (1992). Cortical dynamics of visual motion perception: Short-range and long-range apparent motion. *Psychological Review*, 99, 78-121.

Heeger, D.J. (1987). Model for the extraction of image flow. *Journal of the Optical Society of America,* A8, 1455-1454.

Heeger, D. J. (1991). Nonlinear model of neural responses in cat visual cortex. M. Landy, & A. Movshon (Eds.), *Computational models of visual processing*, (Chap. 9). Cambridge, Mass.: MIT Press.

Henning, G.B. & Derrington, A.M. (1988). Direction-of-motion discrimination with complex patterns: further observations. *Journal of the Optical Society of America,* A5, 1759-1766.

Hubel, D.H, & Wiesel, T.N. (1977). Ferrier Lecture: Functional architecture of macaque monkey visual cortex. *Proceedings of the Royal Society of London*, B198, 1-59.

Johnston, A. & Wright, M.J. (1983) Visual motion and cortical velocity. *Nature*, 304, 436-438.

Johnston, A. & Wright, M.J. (1985) Lower thresholds of motion for gratings as a function of eccentricity and contrast. *Vision Research*, 25, 179-185.

Klein, S. A. (1985). Double-judgement psychophysics: problems and solutions. *Journal of the Optical Society of America*, A2, 1560-1585.

Kuffler, S. W. (1953). Discharge patterns and functional organization of mammalian retina. *Journal of Neurophysiology*, 16, 37-68.

Leventhal, A.G., Rodieck, R.W. & Dreher, B. (1981). Retinal ganglion cell classes in the Old World monkey: Morphology and central projections. *Science*, 213, 360-362.

Levi , D.M., Klein, S.A. & Aitsebaomo, A.P. (1985). Vernier acuity, crowding and cortical magnification. *Vision Research*, 25, 963-977.

Levi, D.M., Klein, S.A & Aitsebaomo, A.P. (1984). Detection and discrimination of the direction of motion in central and peripheral vision of normal and amblyopic observers. *Vision Research*, 24, 789-800.

Livingstone, M.S. & Hubel, D.H. (1987). Psychophysical evidence for separate channels for the perception of form, color, movement, and depth. *Journal of Neuroscience*, 7, 3416-3468.

Marr, D., & Ullman, S. (1981). Directional selectivity and its use in early visual processing. *Proceedings of the Royal Society of London*, B211, 151-180.

Mather, G., Cavanagh, P. & Anstis, S.M. (1985). A moving display which opposes short-range and long-range signals. *Perception*, 14, 163-166.

Maunsell, J.H.R. & Newsome, W.T. (1987). Visual processing in monkey extrastriate cortex. *Annual Review of Neuroscience*, 10, 363-401.

McKee, S.P. & Nakayama, K. (1984) The detection of motion in the peripheral visual field. *Vision Research*, 24, 25-32.

Mikami, A., Newsome, W.T., & Wurtz, R.H. (1986). Motion selectivity in macaque visual cortex: II. Spatio-temporal range of directional interaction in MT and V1. *Journal of Neurophysiology*, 55, 1328-1339.

O'Shea, R.P. & Mitchell, D.E. (1990). Vernier acuity with opposite-contrast stimuli. *Perception*, 19, 207-221.

Palmer, J. (1986). Mechanism of displacement discrimination with a visual reference. *Vision Research*, 12, 1939-1947.

Poggio, T. & Reichardt, W. (1973). Considerations on models of movement detection. *Kybernetik*, 13, 223-227.

Reichardt, W. (1987). Evaluation of optical motion information by movement detectors. *Journal of Comparative Physiology*, A161, 533-547.

Reichardt, W. (1957). Autokorrelations-Auswertung als Funktionsprinzip des Zentralnervensystem. *Zeitschrift für Naturforschung*, 12B, 448-457.

Rovamo, J. & Virsu, V. (1979) An estimation and application of the human cortical magnification factor. *Experimental Brain Research*, 37, 495-510.

Sato, T. (1989). Reversed apparent motion with random dot patterns. *Vision Research*, 29, 1749-1758.

Shechter, S., & Hochstein, S. (1990). On and off pathway contributions to apparent motion perception. *Vision Research*, 30, 1189-1204.

Shechter, S., Hochstein, S., & Hillman, P. (1988). Shape similarity and distance disparity as apparent motion correspondence cues. *Vision Research*, 28, 1013-1021.

Slaughter, M.M. & Miller, R.F. (1981). 2-Amino-4-phosphonobutyric acid: A new pharmacological tool for retina reserach. Science, 211, 182-184.

Schein, S.J. (1988). Anatomy of macaque fovea and spatial densities of neurons in foveal representation. *Journal of Comparative Neurology*, 269, 479-505.

Schein, S.J. & DeMonasterio, F.M. (1987) Mapping of retinal and geniculate neurons onto striate cortex of macaque. *Journal of Neuroscience*, 7, 996-1009.

Schiller, P. H. (1982). Central connections of the retinal On and Off pathways. *Nature*, 297, 580-583.

Schiller, P.H. (1992). The ON and OFF channels of the visual system. Trends in Neurosciences, 15, 86-92.

Schiller, P. H., Sandell, J. H., & Maunsell, J. H. R. (1986). Function of the ON and OFF channels of the visual system. *Nature*, 332, 824-825.

Schwartz, E.L. (1977) Spatial mapping in the primate sensory projection: analytic structure and relevance to perception. *Biological Cybernetic*, 25, 181-194.

Schwartz, E.L. (1980). Computational anatomy and functional architecture of striate cortex: a spatial mapping approach to perceptual coding. *Vision Research*, 20, 645-669.

Solomon, J.A. & Sperling, G. (1994). Full-wave and half-wave rectification in second-order motion perception. *Vision Research*, 34, 2239-2257.

Tolhurst, D. J., & Dealy, R. S. (1975). The detection and identification of lines and edges. *Vision Research*, 15, 1367-1372.

van Essen, D.C., Newsome, W.T. & Maunsell, J.H.R. (1984). The visual field representation in the striate cortex of the macaque monkey: Asymmetries, anisotropies and individual variability. *Vision Research*, 24, 429-448.

van Santen, J.P.H., & Sperling, G. (1985). Elaborated Reichardt detectors. *Journal of the Optical Society of America,* A2, 300-321.

van Santen, J.P.H., & Sperling, G. (1984). Temporal covariance model of human motion perception. *Journal of the Optical Society of America,* A1, 451-473.

Virsu, V. & Rovamo, J. (1979) Visual resolution, contrast sensitivity and the cortical magnification factor. *Experimental Brain Research*, 37, 475-494.

Wässle, H., Grünert, U., Röhrenbeck, J., Boycott, B.B. (1990) Retinal ganglion cell density and cortical magnification factor in the primate. *Vision Research*, 30, 1897-1911.

Watson, A.B. & Ahumada, A.J. (1985). Model of human visual-motion sensing. *Journal of the Optical Society of America*, A2, 332-342.

Watt, R. J., & Morgan, M. J. (1985). A theory of the primitive spatial code in human vision. *Vision Research*, 25, 1661-1674.

Wehrhahn, C. & Rapf, D. (1992). ON- and OFF-pathways form separate neural substrates for motion perception: psychophysic evidence. *Journal of Neuroscience*, 12, 2247-2250.

Werkhoven, P., Sperling, G., & Chubb, C. (1993). The dimensionality of texture-defined motion: A single channel theory. *Vison Research*, 33, 463-485.

Westheimer, G. (1982) The spatial grain of the perifoveal visual field. *Vision Research*, 22, 157-162.

Westheimer, G. (1983) temporal order detection for foveal and peripheral visual stimuli. *Vision Research*, 23, 759-763.

Westheimer, G. & McKee, S.P. (1977). Perception of temporal order in adjacent visual stimuli. *Vision Research*, 17, 887-892.

Wilson, H.R. & Kim, J. (1994). Perceived motion in the vector sum direction. *Vision Research*, 34, 1835-1842.

Yo, C. & Wilson, H.R. (1992). Perceived direction of moving two-dimensional pattern depends on duration, contrast and eccentricity. *Vision Research,* 32, 135-147.

CURRICULUM VITAE

ALEKSANDER PULVER

Citizenship:	Estonian
Date of birth:	January 5, 1956
Marital status:	Married, two children
Address:	Department of Psychology, University of Tartu, 78 Tiigi Street, EE2400, Tartu, Estonia
Telephone:	(+37 2 7) 43 00 63
Fax:	(+37 2 7) 43 05 41
Internet:	sassa@psych.ut.ee

EDUCATION

University of Tartu, Department of Psychology	1975–1980

PROFESSIONAL EMPLOYMENT

Research Scientist, Department of Psychology, University of Tartu	1983–
Lecturer, Department of Psychology, University of Tartu	1980–1983

Lecture courses:

- social psychology
- general psychology
- interpersonal communication
- psychology of family

SCIENTIFIC ACTIVITY

The main fields of interest: visual perception and psychometrics
The number of scientific publications: 22

CURRICULUM VITAE

ALEKSANDER PULVER

Kodakondsus: Eesti Vabariik
Sünniaeg: 5. jaanuar 1956
Perekonnaseis: Abielus, peres on kaks last
Aadress: Psühholoogia osakond, Tartu Ülikool, Tiigi 78, EE2400, Tartu.
Telefon: (+37 2 7) 43 00 63 (tööl)
Fax: (+37 2 7) 43 05 41
Internet: sassa@psych.ut.ee

HARIDUS

Tartu Ülikool, psühholoogia osakond	1975–1980

ERIALANE TEENISTUSKÄIK

Tartu Ülikooli psühholoogia osakonna teadur	1983–
Tartu Ülikooli psühholoogia osakonna lektor	1980–1983

Loengukursused:

- sotsiaalpsühholoogia
- üldpsühholoogia
- suhtlemispsühholoogia
- peresuhete psühholoogia

TEADUSLIK TEGEVUS

Peamised uurimisvaldkonnad: nägemistaju ja psühhomeetria
Teaduslike publikatsioonide üldarv: 22

DISSERTATIONES PSYCHOLOGICAE UNIVERSITATIS TARTUENSIS

1. Jüri Kruusvall. Environmental and social influence on human acitivity. Tartu 1994, 135 p.
2. Dagmar Kutsar. Transformation in Estonia as reflected in families: Insight into social stress and poverty. Tartu 1995, 171 p.