Dynamics of Visual Motion Processing

Uwe J. Ilg · Guillaume S. Masson
Editors

Dynamics of Visual Motion Processing

Neuronal, Behavioral, and Computational Approaches

Springer

Editors
Uwe J. Ilg
Department of Cognitive Neurology
Hertie-Institute of Clinical Brain Research
University of Tuebingen
Otfried-Mueller Str 27
Tuebingen 72076, Germany
Uwe.Ilg@uni-tuebingen.de

Guillaume S. Masson
Team Dynamics of Visual Perception
and Action
Institut de Neurosciences Cognitives
de la Méditerranée
CNRS & Université de la Méditerranée
31 Chemin Joseph Aiguier
13402 Marseille, France
guillaume.masson@incm.cnrs-mrs.fr

ISBN 978-1-4419-0780-6 e-ISBN 978-1-4419-0781-3
DOI 10.1007/978-1-4419-0781-3
Springer New York Dordrecht Heidelberg London

Library of Congress Control Number: 2009930365

Cover illustration: Dynamical extraction and diffusion 2D motion using a particle filter. Blue and red color illustrate the direction of local motion at early and late time steps. See Perrinet & Masson, CoSyne Annual Meeting 2009. Courtesy of Laurent U. Perrinet.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

Biological motion is an essential piece of sensory information for living organism and therefore motion processing units, from simple elementary motion detectors to dedicated motion sensitive cortical areas, have been identified over a broad spectrum of animals. Biological visual motion systems are among the ones having been the most scrutinized at many different levels from microcircuits to perception (see Born and Bradley 2005; Bartels et al. 2008; Callaway 2005; Sincich and Horton 2005; Demb 2007; Britten 2008; Bradley and Goyal 2008; Kourtzi et al. 2008; Orban 2008 for recent reviews). In parallel, since the early work of Reichardt (1961), theoretical approaches of motion detection have been always tightly linked with experimental work so that nowadays, most experiments are conducted within rather well-defined theoretical frameworks (e.g. Carandini et al. 2005). Visual motion has thus become a representative of system neurosciences where similar approaches can be applied across very different levels of brain organization. In particular, neuronal activity at both single-cell and population levels can be accurately linked to simple action system driven by visual motion such as tracking eye movements (Lisberger et al. 1987) as well as motion perception (Parker and Newsome 1998). This integrative approach is rooted on decades of psychophysics that have explored human motion perception (Nakayama 1985; Vaina 1998; Snowden and Freeman 2004; Lu and Sperling 2001). Visual psychophysics provides all of us with a large class of calibrated motion stimuli that can be used to dissect out the different aspects of motion integration and segmentation as needed to accurately measure the velocity of an object that is, the direction and speed of its movement. We decided to open this book with a review paper describing what are the different classes of visual stimuli and what aspects of biological motion processing each of them can unveil. Focusing of low-level motion, Lorenceau presents in great details the different elements of this artillery and how they can be used at both behavioral and neurophysiological levels. By doing so, he set the stage over which most of the work presented inside this book will take place. As for the other chapters, corresponding movies can be found in the DVD joined to the book. However, Lorenceau also stresses out that a motion perception most often involves a tight link between form and motion cues. Such form–motion interactions will be illustrated by other contributions, further demonstrating that biological motion processing escapes the strict modular approach and call for a more integrative view as needed to understand the root of the problem: how to measure the motion of an object, usually represented as a visual surface, singled out from its complex environment. The following chapters will survey how this can be performed at cellular and network levels, with either static or moving eyes.

Dynamics of Neural Mechanisms

Surprisingly, several key aspects of motion perception have not been emphasized over the years. First, although a few psychophysical studies had pointed out that perceived motion undergoes complex time course when human subjects are presented with ambiguous motion signals (e.g. Yo and Wilson 1992; Castet et al. 1993; Lorenceau et al. 1993), it is only very recently that temporal dynamics of motion processing has received attention from physiologists. Before the pioneering work of Pack and colleagues, neurons were classified between those who solve the aperture problem and those who do not. This selection was based on the steady-state properties of their direction selectivity tuning when presented with bars, gratings, or plaid patterns (Movshon et al. 1985). Pack and Born (2001) presented MT neurons with sets of tilted bars, the neuronal counterpart of the elongated moving bars used in psychophysical experiments, and analyzed the time course of direction selectivity of the single-unit responses. They found that such a basic response property of MT neuron is indeed not static. Instead, early part of their tuning showed interactions between direction and orientation while ~100 ms after response onset, optimal direction became independent of lines orientation. Several studies, largely summarized here in the chapters by Pack et al. and Smith et al., have looked at the dynamics of direction selectivity in macaque area MT in response to various 2D motions such as plaid patterns, barber poles, or lines. Although there is a common agreement on the similar timing characteristics across motion stimuli (see chapter by Smith et al.), the origin of such neuronal dynamics is still highly controversial, carrying on a long debate about which, and how, local features are extracted from the image flow. Born and coworkers favor an explanation based on feature-extraction mechanisms such as end-stopped cells found in area V1 (Hubel and Wiesel 1968; Pack et al. 2003). On the other hand, Smith and coworkers argue for a filter-based approach where global motion is computed by merging excitatory and inhibitory inputs from different spatio-temporal channels (see Rust et al. 2006). Within these two frameworks, the dynamics can be seen either as a result of a delayed feature-extraction mechanisms, as the by-product of the different signal strength between channels or by the time course of contextual modulation such as implemented by center-surround interactions or recurrent circuits. The book offers the opportunity for these different views to be presented back to back.

Motion information is extracted locally, but there is many evidence that the brain pools information to solve the aperture problem, to improve signal-to-noise ratio or to normalize inputs across the image, to take a few examples of motion integration. Since all these different aspects involved the diffusion of information between neighboring neurons, there is an urgent need to explore the neural dynamics at population level.

Frégnac and coworkers introduce the concepts and techniques used to investigate the relationships between fundamental properties of individual neurons such as orientation- and direction-selective cells in primary visual cortex and the dynamics of their surrounding network. They point out that descriptive tuning functions in fact reflect the very large diversity of inputs that a single neuron would receive through feed-forward, lateral, and recurrent connectivity. This message is to keep in mind in the design of detailed biophysical models at both cellular and network levels. It remains coherent with the current view that direction selectivity emerges from the convergence of many different feed-forward inputs (both excitatory and inhibitory) covering a very broad range of the spatiotemporal spectrum in Fourier space (see Rust et al. 2006; Lennie and Movshon 2005). However, the evidence gathered by intracellular recordings that responses dynamics of V1 neurons reflect non-iso-oriented inputs (Monier et al. 2003) from distant part of the cortex (Bringuier et al. 1999) urge us to take into account the role of intra- and intercortical connections. The fact that they all have different timing shall help us in constraining dynamical models of motion integration.

Linking population dynamics and integrative properties of individual neurons will be most certainly a future challenge in sensory neuroscience. Visual motion once again shall offer an excellent approach. Jancke, Chavane, and Grinvald provide one very attractive insight into this perspective. Using different and complementary techniques such as voltage-sensitive dye optical imaging and population reconstruction from extracellular recordings, they propose a fresh look at how motion information is represented. In particular, their approach stresses one point often ignored in most electrophysiological, and psychophysical, studies. Motion is primarily a displacement in the visual space and therefore a moving object will elicit a traveling wave along the cortical representation of its trajectory. Moreover, linear and nonlinear interactions along such cortical trajectories can be identified in cat area V1 (Jancke et al. 2004). Most certainly, future work will be able to relate such population dynamics to single-unit activity within direct projection areas such as V2 or MT as well as with perceptual performance in primates (Chen et al. 2006).

Overall, looking at the temporal dynamics of contextual biological motion processing, as well as for other elementary aspects of image features extractions such as orientation, texture, and illusory contours has reinvigorated the investigations on the underpinning neural mechanisms. The results gathered might turned out to be important to decipher which theoretical approach is more closely related to cortical computation. They might also force us to finally take into account the different connectivity rules, operating at different spatial and temporal scales, which are important to compute global object motion.

Visual Motion and Eye Movements

Measuring speed and direction of a moving object is an essential step for many sensorimotor transformations, in particular when controlling eye movements. The impact of low-to-high level motion processing onto the dynamics of oculomotor behavior is reviewed in several chapters. Sheliga and Miles summarize their seminal work in elucidating the basic properties of motion detection in the context of triggering reflexive eye movements at ultrashort latency. Their work illustrates how much can be learned about the spatial and temporal characteristics of the earliest, preattentive stage of local motion extraction when using very accurate behavioral probes. From this, as well as from the work of other groups, it becomes possible to sketch a detailed model of early, and fast, motion processing that incorporates many aspects investigated previously at psychophysical and physiological levels: how is motion information extracted by luminance-based motion detectors, how are their activity normalized across directions and so on and so forth. More than simply confirming what was learned from other approaches, the experiments conducted on ocular following responses unveil functional consequences of such linear and nonlinear processing such as automatic motion segmentation and integration (see Miles et al. 2004; Masson 2004). If tracking eye movements are primarily driven by luminance-based local motion detection, this so-called first-order motion mechanisms might not be the only one contributing to a nearly perfect pursuit performance under a wide range of conditions. Other types of motion information can be extracted under constant luminance conditions, either at preattentive or at attentive stages. System view of the primate motion system postulates the existence of three different motion systems, called first order, second order, and third order (see Lu and Sperling 2001 for a review). The exact contribution of second- and third-order motion information to perceptual performance is still a matter of debate and it is unclear where and how they are computed by the primate brain. Chapter by Ilg and Churan reviews the existing evidence, supporting the idea that second-order motion is indeed extracted within posterior parietal areas. The authors point out, however, that investigating second-order motion, as well as pattern motion, had defeated the simplistic view that global motion is computed once for all in area MT and therefore that area MT must be seen as the key, if not unique area responsible for motion perception in both human and nonhuman primates (see Ilg 2008 for a review).

Once motion is locally extracted, several processing steps are still necessary to reconstruct speed and direction of the object to be pursued. Chapters presenting new results on motion integration, obtained at both psychophysical and physiological levels, have introduced the idea that the integration stage presents complex dynamics. This approach is further extended in the chapter by Masson and colleagues, showing such dynamics of motion integration can have a major impact on how the brain controls action. Taking advantage of the fast visuomotor transformation underlying that pursuit eye movements as well as their smooth acceleration, the oculomotor system can trigger tracking responses based only on the coarse estimate of motion direction that arises from the feed-forward motion pathway but then gradually correct the pursuit direction by taking into account features motion extracted at a finer scale. Thus, time course is closely related to the temporal dynamics of motion integration that we have discussed above. In return, this work stresses the fact that eye movements are an exquisite tool to probe the properties of early motion processing stages, since initial eye acceleration reflects visual velocity signals encoded at the level of macaque areas MT and MST (Krauzlis 2004; Masson 2004).

However, it is well known since the early 1970s that pursuit responses depend on both visual and nonvisual signals, the later being related to eye velocity memory (Yasui and Young 1975; Miles and Fuller 1975). Moreover, the perceived direction of oriented after-images presented during on-going pursuit is always biased toward the axis normal to the orientation of the flashed bars (Goltz et al. 2003). This intriguing result suggests first that the aperture problem contaminates egocentric motion and second that more is yet to learn about motion integration during eye movements (Murakami 2004). Indeed, motion integration tasks such as introduced by Lorenceau offer a great deal to investigate the link between perception and action, as well as the dependency or the lack of dependency of early visual stages upon cognitive factors such as prediction or anticipation. Masson and colleagues report results arguing for a mere independence between low-level vision and higher cognitive processing such as engaged in anticipating future motion events or predicting target trajectory. They suggest that low-level motion integration and spatial reconstruction of target motion acts more or less independently, as illustrated by the difference observed between neuronal responses in either areas MT or MSTl/FEF when using complex line drawing stimuli avoiding the center of the receptive field (Ilg and Thier 2003). These latter experiments suggest that pursuit-related neurons in the lateral part of macaque area MST (also called visual-tracking neurons) integrate both visual and nonvisual information (see Ilg 2008 for a review). Whether these neurons compute the motion-in-space of a pursued target (Newsome et al. 1988) or reflect the existence of a more abstract representation of the inferred motion already emerging at the level of area MT (Assad and Maunsell 1995; Schlack and Albright 2007) is still a matter of debate. Recording activities of MSTl neurons during tracking of different line-drawing objects is one piece of evidence. Furthermore, looking at the dynamics of direction selectivity using tilted bars that are transiently occluded (see Assad and Maunsell 1995 for a similar paradigm although with a simple spot) might also largely contribute to a better understanding about what and how information is represented at various stages along the motion pathway. Clearly, more investigations are needed about the dynamical interactions between posterior parietal and prefrontal cortical areas for motion integration in the context of pursuit eye movements, as well as perception (see Pasternak and Greenlee 2005). However, once again, these studies point out how using simple motion stimuli such as designed for psychophysics can highlight the mechanisms of sensorimotor transformation when the biological motion stage is not collapsed into a simple black box extracting retinal velocity in some unspecified way. Obviously, there is need for models of oculomotor behavior with a more complex front-end dynamics.

In the aforementioned chapters, motion is seen as the source of information for driving perception or simple actions such as tracking responses. Although active vision has been a very productive field of research trying to understand how visual information is actively extracted by means of our eye movements, much more attention has been paid to saccadic eye movements rather than smooth pursuit in this context (Findlay and Gilchrist 2003). Tracking an object aims at stabilizing its image onto the retina, but a mere consequence of the eyeball rotation is a steady continuous retinal sweep of the background image. Dozens of studies have been conducted to understand how such background motion can be either eliminated to perceive a stable world during tracking or on the contrary taken into account to compute object motion in a spatial frame of reference (see Abadi and Kulikowski 2008). Hafed and Krauzlis take a different approach, trying to demonstrate that smooth eye movements can be useful to resolve perceptual ambiguities. This approach is rooted on the motion stimuli, and psychophysical paradigms described by Lorenceau but offer a fresh view of the fascinating problem of perception–action coupling. Their experimental work, summarized in Chap. 9, shows that partially occluded objects can be perceived coherently thanks to the pattern of eye movements produced by human subjects. This seminal study opens the door to a closer examination to the interaction between perception and action using both well-defined behavior and calibrated tasks where retinal flows can be matched between pursuit and fixation condition.

Visual motion processing is not only related to the execution of pursuit eye movements. Both saccadic and pursuit eye movements introduce major changes in the retinal images. However, how motion perception and eye movements are coupled with respect to saccades has been a matter of intense debates over the last decades. One acute example is a phenomenon called "saccadic suppression" (see Ross et al. 2001). That visual perception is largely reduced during saccades is a well-documented phenomenon that everyone can experience everyday. Indeed, psychophysical studies have convincingly demonstrated that intrasaccadic detection thresholds are strongly deteriorated at the time of a saccade (e.g. Volkmann 1986; Burr et al. 1994). Several recent physiological studies have demonstrated that some, but not all direction-selective cells in macaque area MT are consistently inhibited during saccade. On the contrary, some cells also show a strong postsaccadic rebound of activity that could be correlated to the postsaccadic enhancement originally reported by Miles and colleagues when recording ocular following responses (Miles et al. 1986). In Chap. 8, Mike Ibbotson summarizes these studies and relates these saccade-related changes in activity at the level of area MT with the changes in perceptual performance described earlier in human subjects. However, the use of the term "suppression" has led to the stronger, but wrong, assumption that vision is prevented during saccades. Textbooks and nonspecialist review articles have even further cartooned this saying that the entire visual system, not only visual perception, is turned off during saccadic flight. The chapter by Castet offers a very helpful re-examination of the different perceptual changes that occur before, during, and after a saccade. He points out the difficulty in interpreting a wide diversity of perceptual phenomena within the unique, stringent hypothesis of an active (i.e. extraretinal) suppression or modulation of visual inflow at its earliest stages (Ross et al. 2001). One goal of this book was to publish back-to-back articles offering different, sometimes even opposite, standpoints onto a specific aspect of motion processing. The chapters by Ibbotson on one hand and Castet on the other hand give such an opportunity and remind us that solving controversies in neuroscience often needs first to (re)clarify key concepts as often popular ideas drift far away from the conclusions that were drawn from the original experimental results.

Modeling Visual Motion: From Natural Scenes Statistics to Motion Recognition

Listing the existing computational models of visual motion would probably take a couple of pages. Computer as well as biological vision researches have produced a huge number of models, based on many different theoretical approaches such as linear filtering, probabilistic inference, or dynamical systems. Several recent books are available from the library shelves (see Blake 1998; Paragios et al. 2005; Stocker 2004; Watanabe 1998 for a few recent examples) to explore these different aspects. There is, however, clearly the need for a more theoretical approach unifying all these computational efforts. Herein, we have preferred to highlight some key aspects of visual motion information processing. First, Dong summarizes the statistical approach trying to understand what is the critical information in sequences of natural images. Relating the window of visibility, and its neuronal counterpart defined as a set of optimal filters, to the statistics of still natural images has been an intensive area of research over the last two decades. The same approach is now conducted using movies of the image flow experienced by an observer moving in complex, natural environments. From these, Dong demonstrates that spatial–temporal contrast sensitivity of human observers is tuned to extract the most pertinent and reliable motion information that is mainly low temporal frequencies.

A second aspect of motion processing is integration, which involves diffusion of information over neighboring parts of the images to reconstruct the global motion of the object of interest and single it out from its surround. Grossberg summarizes the work conducted by his group in implementing dynamical models of motion segmentation and integration. His solution relies on a strong interplay between modules extracting either form (i.e. features) or motion. Diffusion of information is done by means of recurrent connectivity between areas working at different spatial scales. Once again, this class of model reminds us that motion pathways are highly recurrent and that we absolutely need to better understand how feed-forward and feedback flows of information interplay to solve problems such as motion binding. The model reviewed here sums up a decade of progressive improvement of the class of models developed by himself and his group. Clearly, this approach highlights the interest of computational principles that can be implemented by set of differential equations. The cost is then to overview the detailed connectivity rules corresponding to the actual cortical mechanisms. But we clearly need such a more generic approach, complementary to the more detailed, but also more focused, models proposed by others. Lastly, Grossberg introduces one new aspect of the dynamical approach. The brain takes decision about the incoming stimulus speed or direction. It model succeeds in simulating the time course of such a decision, as seen in parietal cortices (e.g. Britten et al. 1992; Shadlen and Newsome 2001; Hanks et al. 2006; Huk and Shadlen 2005) but also question on what information processing such decision is taken. This links to a rapidly growing field of research about sensory decisions along the motion pathways. Recent reviews about this topic can be found elsewhere (Shadlen 2002; Rorie and Newsome 2005; Gold and Shadlen 2007).

Motion is a useful source of information not only for controlling our basic actions but also to solve highest cognitive tasks such as face recognition (see Roark et al. 2003) or biological action recognition (see Blake and Shiffrar 2007). Understanding how biological motion is analyzed by dedicated brain loci within the superior temporal sulcus (STS) for instance has been the focus of a vast literature. However, biological motion stimuli carry information not only the type of action being executed but also more fine-grained, cognitive cues that are used for social interactions. Giese and coworkers detailed their recent modeling work asking how human emotions can be recognized for sequences of point-light walkers. Here again, a key point is to be able to extract remarkable features such as joint-angle trajectories using sparse feature learning. This approach not only defines a compact visual representation for complex information but depart from more classical models assuming that visual recognition involves the activation of motor representations. Instead, this model demonstrates that human subjects can nearly optimally use the visual information extracted from joint trajectories.

Features, trajectories, dynamic motion integration: these terms have been found in nearly all chapters of this book. By highlighting a few recent approaches, the contributors have shown how much an integrative approach can be useful to understand how the brain computes global motion of an object, being a simple line or a full body. Some of these issues still remain controversial and we want to thank the different contributors to have accepted that chapters with different views are presented back to back. We hope that our colleagues and their students will consider this book for what it was originally proposed: an incentive to bridge approaches across levels and models, using tasks and stimuli as an Ariadne's thread.

Marseille, France Guillaume S. Masson
Tuebingen, Germany Uwe J. Ilg

References

Abadi RV, Kulikowski JJ (2008) Perceptual stability: going with the flow. Perception 37(9):1461–1463

Assad JA, Maunsell JHR (1995) Neuronal correlates of inferred motion in primate posterior parietal cortex. Nature 373:518–521

Bartels A, Logothetis NK, Moutoussis K (2008) fMRI and its interpretations: an illustration on directional selectivity in area V5/MT. Trends Neurosci 31(9):444–453

Blake A (1998) Active contours: the application of techniques from graphics, vision, control theory and statistics to visual tracking of shapes in motion. Springer, Berlin

Blake R, Shiffrar, M (2007) Perception of human motion. Annu Rev Psychol 58:47–73

Bradley DC, Goyal MS (2008) Velocity computation in the primate visual system. Nat Rev Neurosci 9(9):686–695

Britten KH (2008) Mechanisms of self-motion perception. Annu Rev Neurosci 31:389–410

Britten KH, Shadlen MN, Newsome WT, Movshon JA (1992) The analysis of visual motion: a comparison of neuronal and psychophysical performance. J Neurosci 12(12):4745–4765

Born RT, Bradley DC (2005) Structure and function of visual area MT. Annu Rev Neurosci 28:157–189

Bringuier V, Chavane F, Glaeser L, Frégnac Y (1999) Horizontal propagation of visual activity in the synaptic integration field of area 17 neurons. Science 283(5402):695–699

Britten KH (2008) Mechanisms of self-motion perception. Annu Rev Neurosci 31:389–410

Burr DC, Morrone MC, Ross J (1994) Selective suppression of the magnocellular visaul pathway during saccadic eye movements. Nature 371:511–513

Callaway EM (2005) Neural substrates within primary visual cortex for interactions between parallel visual pathways. Prog Brain Res 149:59–64

Carandini M, Demb JB, Mante V, Tolhurst DJ, Dan Y, Olshausen BA, Gallant JL, Rust NC (2005) Do we know what the early visual system does? J Neurosci 25(46):10577–10597

Castet E, Lorenceau J, Shiffrar M, Bonnet C (1993) Perceived speed of moving lines depends on orientation, length, speed and luminance. Vision Res 33(14):1921–1936

Chen Y, Geilser WS, Seidemann E (2006) Optimal decoding of correlated neural population responses in the primate visual cortex. Nat Neurosci 9(11):1412–1420

Demb JB (2007) Cellular mechanisms for direction selectivity in the retina. Neuron 55(2): 179–286

Dobbins A, Zucker SW, Cynader MS (1987) Endstopped neurons in the visual cortex as a substrate for calculating curvature. Nature 329:438–441

Findlay JM, Gilchrist ID (2003) Active vision. The psychology of looking and seeing. Oxford University Press, Oxford

Gold JI, Shadlen MN (2007) The neural basis of decision making. Annu Rev Neurosci 30:535–574

Goltz HC, DeSouza JF, Menon RS, Tweed DB, Vilis T (2003) Interactions of retinal image and eye velocity in motion perception. Neuron 39(3):569–579

Hanks TD, Ditterich J, Shadlen MN (2006) Microstimulation of macaque area LIP affects decision making in a motion discrimination task. Nat Neurosci 9(5):682–689

Hubel DH, Wiesel TN (1968) Receptive fields and functional architecture of monkey striate cortex. J Physiol (Lond) 195(1):215–243

Huk AC, Shadlen MN (2005) Neural activity in macaque parietal cortex reflects temporal integration of visual motion signals during perceptual decision making. J Neurosci 25(45): 10420–10436

Ilg UJ (2008) The role of areas MT and MST in coding of visual motion underlying the execution of smooth pursuit. Vision Res 48(20):2062–2069

Ilg UJ, Thier P (2003) Visual tracking neurons in primate area MST are activated by smooth-pursuit eye movements of an “imaginary” target. J Neurophysiol 90(3):1489–1502

Jancke D, Chavane F, Naaman S, Grinvald A (2004) Imaging cortical correlates of illusion in early visual cortex. Nature 428(6981):423–426

Kourtzi Z, Krekelberg B, Van Wezel RJ (2008) Linking form and motion in the primate brain. Trends Cogn Sci 12(6):230–236

Krauzlis RJ (2004) Recasting the smooth pursuit eye movement system. J Neurophysiol 91(2):591–603

Lennie P, Movshon, JA (2005) Coding of color and form in the geniculostriate visual pathway. J Opt Soc Am A 22(10):2013–2033

Lisberger SG, Morris EJ, Tychsen L (1987) Visual motion processing and sensory-motor integration for smooth pursuit eye movements. Annu Rev Neurosci 10:97–129

Lorenceau J, Shiffrar M, Wells N, Castet E (1993) Different motion sensitive units are involved in recovering the direction of moving lines. Vision Res 33(9):1207–1217

Lu ZL, Sperling G (2001) Three-systems theory of human visual motion perception: review and update. J Opt Soc Am A 18(9):2331–2370

Masson GS (2004) From 1D to 2D via 3D: dynamics of surface motion segmentation for ocular tracking in primates. J Physiol (Paris) 1–3:35–52

Miles FA, Fuller JH (1975) Visual tracking and the primate flocculus. Science 189:1000–1002

Miles FA, Kawano K, Optican LM (1986) Short-latency ocular following responses of monkey. I. Dependence on temporospatial properties of visual input. J Neurophysiol 56(5):1321–1354

Miles FA, Busettini C, Masson GS, Yang D-Y (2004) Short-latency eye movements: evidence for parallel processing of optic flow. In: Vaina L, Beardsley SA, Rushton S (eds) Optic flow and beyond. Kluwer, New York, pp 70–103

Monier C, Chavane F, Baudot P, Graham LJ, Frégnac Y (2003) Orientation and direction selectivity of synaptic inputs in visual cortex neurons: a diversity of combinations produces spike tuning. Neuron 37(4):663–680

Murakami I (2004) The aperture problem in egocentric motion. Trends Neurosci 27(4):174–177

Nakayama K (1985) Biological image motion processing: a review. Vision Res 25(5):625–660

Newsome WT, Wurtz RH, Komatsu H (1988) Relation of cortical areas MT and PST to pursuit eye movements. II. Differentiation of retinal from extraretinal inputs. J Neurophysiol 60:604–620

Orban GS (2008) Higher-order visual processing in macaque extrastriate cortex. Physiol Rev 88(1):59–89

Pack CC, Born RT (2001) Temporal dynamics of a neural solution to the aperture problem in visual area MT of macaque brain. Nature 409:1040–1042

Pack CC, Livingstone MS, Duffy KR, Born RT (2003) End-stopping and the aperture problem: two-dimensional motion signals in macaque V1. Neuron 39(4):671–680

Paragios N, Chen Y, Faugeras O (2005) Handbook of mathematical models in computer vision. Springer, Berlin

Parker AJ, Newsome WT (1998) Sense and the single neuron: probing the physiology of perception. Annu Rev Neurosci 21:227–277

Pasternak T, Greenlee MW (2005) Working memory in primate sensory systems. Nat Rev Neurosci 6(2):97–107

Reichardt W (1961). Autocorrelation, a principle for evaluation of sensory information by the central nervous system. In: Rosenblith WA (ed) Sensory communication (p. 303). Wiley, New York, pp 303–317

Roark DA, Barrett SE, Spence MJ, Abdi T, O'Toole AJ (2003) Psychological and neural perspectives on the role of motion in face recognition. Behav Cogn Neurosci Rev 2(1):15–46

Rorie AE, Newsome WT (2005) A general mechanism for decision-making in the human brain? Trends Neurosci 9(2):363–375

Ross J, Morrone MC, Goldberg ME, Burr DC (2001) Changes in visual perception at the time of saccades. Trends Neurosci 24(2):113–121

Rust NC, Movshon JA (2005) In praise of artifice. Nature Neuroscience 8(12):1647–1650

Rust NC, Mante V, Simoncelli EP, Movshon JA (2006) How MT cells analyze the motion of visual patterns. Nature Neuroscience 9(11):1421–1431

Schlack A, Albright TD (2007) Remembering visual motion: neural correlates of associative plasticity and motion recall in cortical area MT. Neuron 53:881–890

Shadlen MN (2002) Pursuing commitments. Nat Neurosci 5(9):819–821

Shadlen MN, Newsome WT (2001) Neural basis of a perceptual decision in the parietal cortex (areas LIP) of the rhesus monkey. J Neurophysiol 86(4):1916–1936

Sincich LC, Horton JC (2005) The circuitry of V1 and V2: integration of color, form and motion. Annu Rev Neurosci 28:303–326

Snowden RJ, Freeman TC (2004) The visual perception of motion. Curr Biol 14(9):R828–R831

Vaina LM (1998) Complex motion perception and its deficits. Curr Opin Neurobiol 8(4):494–502

Volkmann FC (1986) Human visual suppression. Vision Res 26(9):1401–1416

Watanabe T (1998) High-level motion processing: computational, neurobiological and psychophysical perspective. MIT Press, Cambridge, MA

Yasui S, Young LR (1975) Perceived visual motion as effective visual stimulus for pursuit eye movement system. Science 190:906–908

Yo C, Wilson HR (1992) Perceived direction of moving two-dimensional patterns depends on duration, contrast and eccentricity. Vision Res 32(1):135–147

Contents

Part I Low-level Cortical Dynamic Motion Processing

Part II Active Vision, Pursuit and Motion Perception

Part III Modeling Dynamic Processing

Contributors

Pierre Baudot
Institut des Systèmes Complexes, CNRS UMR7656, 57/59 rue Lhomond, 75005 Paris, France
baudot@math.jussieu.fr

Richard T. Born
Department of Neurobiology, Harvard Medical School, 220 Longwood Avenue, Boston, MA 02115, USA
rborn@hms.harvard.edu

Predo V. Carelli
Unité de Neurosciences Intégratives et Computationnelles (UNIC), CNRS UPR 2191, 1 Avenue de la Terrasse Bat 32/33, 91198 Gif-sur-Yvette, France
carelli@unic.cnrs-gif.fr

Eric Castet
Dynamics of Visual Perception and Action, Institut de Neurosciences Cognitives de la Méditerranée, CNRS & Université de la Méditerranée, 31 Chemin Joseph Aiguier, 13402 Marseille, France
eric.castet@incm.cnrs-mrs.fr

Fréderic Chavane
Dynamics of Visual Perception and Action, Institut de Neurosciences Cognitives de la Méditerranée, CNRS & Université de la Méditerranée, 31 Chemin Joseph Aiguier, Marseille 13402, France
frederic.chavane@incm.cnrs-mrs.fr

Jan Churan
Department of Neurology and Neurosurgery, Montréal Neurological Institute, McGill University, 3801 University Street, Montréal, QC, Canada H3A2B4

Dawei W. Dong
Center for Complex Systems and Brain Sciences, Florida Atlantic University,
777 Glades Road, Boca Raton, FL 33431, USA
dawei@ccs.fau.edu

Yves Fregnac
Unité de Neurosciences Intégratives et Computationnelles (UNIC), CNRS UPR
2191, 1 Avenue de la Terrasse Bat 32/33, 91198 Gif-sur-Yvette, France
fregnac@unic.cnrs-gif.fr

Martin A. Giese
Section for Computational Sensomotorics, Department of Cognitive Neurology,
Hertie Institute for Clinical Brain Research & Center for Integrative
Nueroscience, Frondsbergstr. 23, 72074 Tuebingen, Germany
martin.giese@uni-tuebingen.de

Stephen Grossberg
Cognitive & Neural Systems, Boston University, 677 Beacon Street,
Boston, MA 02215, USA
steve@cns.bu.edu

Amiram Grinvald
Department of Neurobiology, Weizmann Institute of Science,
Rehovot 76100, Israel
Amiram.Grinvald@weizmann.ac.il

Ziad M. Hafed
The Salk Institute for Biological Studies, 10010 North Torrey Pines Road,
La Jolla, CA 92037, USA
zhafed@salk.edu

Michael Ibbotson
Visual Sciences, Research School of Biological Sciences, Australian National
University, Canberra, Australia
Michael.Ibbotson@anu.edu.au

Uwe J. Ilg
Department of Cognitive Neurology, Hertie-Institute of Clinical Brain Research,
University of Tuebingen, Otfried-Mueller Str 27, Tuebingen 72076, Germany
Uwe.Ilg@uni-tuebingen.de

Dirk Jancke
Department of Neurobiology, ND 7/72, Ruhr-University Bochum,
Bochum 44780, Germany
jancke@neurobiologie.rub.de

Richard J. Krauzlis
The Salk Institute for Biological Studies, 10010 North Torrey Pines Road,
La Jolla, CA 92037, USA
rich@salk.edu

Jean Lorenceau
Equipe Cogimage, UPMC Univ Paris 06, CNRS UMR 7225,
Inserm UMR_S 975, CRICM 47 boulevard de l'Hôpital, Paris, F-75013, France
jean.lorenceau@chups.jussieu.fr

Najib Majaj
McGovern Institute for Brain Research, Massachusetts Institute of Technology,
in Cambridge, 46-6161, 77 Massachusetts Avenue, MA 02139, USA
majaj@mit.edu

Olivier Marre
Unité de Neurosciences Intégratives et Computationnelles (UNIC), CNRS UPR
2191, 1 Avenue de la Terrasse Bat 32/33, 91198 Gif-sur-Yvette, France
marre@unic.cnrs-gif.fr

Guillaume S. Masson
Dynamics of Visual Perception and Action, Institut de Neurosciences Cognitives
de la Méditerranée, CNRS & Université de la Méditerranée, 31 Chemin Joseph
Aiguier, 13402 Marseille, France
guillaume.masson@incm.cnrs-mrs.fr

Anna Montagnini
Dynamics of Visual Perception and Action, Institut de Neurosciences Cognitives
de la Méditerranée, CNRS & Université de la Méditerranée, 31 Chemin Joseph
Aiguier, 13402 Marseille, France
Anna.Montagnini@incm.cnrs-mrs.fr

Frederick A. Miles
Laboratory of Sensorimotor Research, National Eye Institute/NIH,
Bldg 49 Rm 2A50, Bethesda, MD 20892, USA
fam@lsr.nei.nih.gov

Cyril Monier
Unité de Neurosciences Intégratives et Computationnelles (UNIC), CNRS UPR
2191, 1 Avenue de la Terrasse Bat 32/33, 91198 Gif-sur-Yvette, France
monier@unic.cnrs-gif.fr

J. Anthony Movshon
Center for Neural Science, New York University, 4 Washington Place, Room 809,
New York, NY 10003, USA
movshon@nyu.edu

Lars Omlor
Section for Computational Sensomotorics, Department of Cognitive Neurology,
Hertie Institute for Clinical Brain Research & Center for Integrative
Nueroscience, Frondsbergstr. 23, 72074 Tuebingen, Germany
lars.omlor@medizin.uni-tuebingen.de

Christopher C. Pack
Montréal Neurological Institute, McGill University, 3801 University Street,
Montréal, QC, Canada H3A2B4
christopher.pack@mcgill.ca

Marc Pananceau
Unité de Neurosciences Intégratives et Computationnelles (UNIC), CNRS UPR
2191, 1 Avenue de la Terrasse Bat 32/33, 91198 Gif-sur-Yvette, France
pananceau@unic.cnrs-gif.fr

Claire L. Roether
Section for Computational Sensomotorics, Department of Cognitive Neurology,
Hertie Institute for Clinical Brain Research & Center for Integrative
Nueroscience, Frondsbergstr. 23, 72074 Tuebingen, Germany
claire.roether@medizin.uni-tuebingen.de

Gérard Sadoc
Unité de Neurosciences Intégratives et Computationnelles (UNIC), CNRS UPR
2191, 1 Avenue de la Terrasse Bat 32/33, 91198 Gif-sur-Yvette, France
sadoc@unic.cnrs-gif.fr

Boris M. Sheliga
Laboratory of Sensorimotor Research, National Eye Institute/NIH,
Bldg 49, Rm 2A50, Bethesda, MD 20892, USA
bms@lsr.nei.nih.gov

Matthew A. Smith
Center for Neural Basis of Cognition, University of Pittsburgh,
4400 Fifth Avenue, Mellon Institute, Room 115, Pittsburgh, PA 15213, USA
masmith@cnbc.cmu.edu

James M.G. Tsui
Montréal Neurological Institute, McGill University,
3801 University Street, Montréal, QC H3A2B4, Canada
james.tsui@mail.mcgill.ca

Part I

Low-Level Cortical Dynamic Motion Processing

Chapter 1
From Moving Contours to Object Motion: Functional Networks for Visual Form/Motion Processing

Jean Lorenceau

Abstract Recovering visual object motion, an essential function for living organisms to survive, remains a matter of experimental work aiming at understanding how the eye–brain system overcomes ambiguities and uncertainties, some intimately related to the sampling of the retinal image by neurons with spatially restricted receptive fields. Over the years, perceptual and electrophysiological recordings during active vision of a variety of motion patterns, together with modeling efforts, have partially uncovered the dynamics of the functional cortical networks underlying motion integration, segmentation and selection. In the following chapter, I shall review a subset of the large amount of available experimental data, and attempt to offer a comprehensive view of the building up of the unitary perception of moving forms.

1.1 Introduction

An oriented slit of moving light, a microelectrode and an amplifier! Such were Hubel and Wiesel's scalpel used during the 1960s (1959–1968) to uncover the properties of the visual brain of cat and monkey. A very simple visual stimulus indeed which, coupled with electrophysiological techniques, nevertheless allowed the analysis of many fundamental aspects of the functional architecture of primary visual cortex in mammals: distribution of orientation and direction selective neurons in layers, columns and hyper columns, discovery of simple, complex and hyper complex cells, distribution of ocular dominance bands, retinotopic organization of striate visual areas, etc.

J. Lorenceau (✉)
Equipe Cogimage, UPMC Univ Paris 06, CNRS UMR 7225, Inserm UMR_S 975,
CRICM 47 boulevard de l'Hôpital, Paris, F-75013, France
e-mail: jean.lorenceau@chups.jussieu.fr

U.J. Ilg and G.S. Masson (eds.), *Dynamics of Visual Motion Processing: Neuronal, Behavioral, and Computational Approaches*,
DOI 10.1007/978-1-4419-0781-3_1, © Springer Science+Business Media, LLC 2010

Equipped with the elementary brick of information processing – the oriented receptive fields – the house of vision was ready to be built up and the Nobel price was in view. However, recording isolated neurons with a microelectrode might, for a while, have been the tree hiding the forest. If an oriented slit of moving light optimally gets a neuron to fire spikes, how many neighboring neurons also fire in response to that stimulus? What is the size and functional role of the neuronal population presumably recruited by this simple stimulus? Is there more than redundancy? An indirect answer to this question is inverse engineering: what are the requirements for recovering the direction of motion of a slit of moving light, e.g. a moving contour?

Fennema and Thompson (1979), Horn and Schunk (1981) and Hildreth (1984) raised the question and uncovered intrinsic difficulties in answering it, as many problems paved the way, like the "correspondence" and "aperture" problems, also identified on experimental grounds by Henry and Bishop (1971)[1]. Imagine two frames of a movie describing the motion of a 1D homogeneous contour (Fig. 1.1a): what part of the contour in the first frame should be associated to its counterpart in the second frame? The shortest path, corresponding to the motion vector orthogonal to the contour orientation seems the obvious answer, but may not correspond to the distal direction of motion. Applying the shortest path rule between two successive frames – a priority towards low speed – might leave parts of the contour unpaired thus facing a "correspondence" problem. Recovering the direction of an infinite 1D contour soon appeared as an ill posed problem, as an infinity of directions of motion are compatible with a single "local" measurement – e.g. through a biological or artificial motion sensor with a spatially restricted field of "view" (Fig. 1.1b).

In order to overcome this "aperture" problem, a solution is to combine at least two measurements from two 1D contours at different orientations. Amongst the large family of possible motion vectors associated to each contour motion, only one is compatible with both and may therefore correspond to the searched solution (Adelson and Movshon 1982). According to this scheme, motion processing would require two processing stages: the first one would extract local – ambiguous – directions and these measurements would be combined at a second stage. Numerous models (Nowlan and Sejnowski 1995; Liden and Pack 1999; Simoncelli and Heeger 1998; Wilson and Kim 1994) rely on this idea: the small receptive fields of V1 cells would first calculate motion energy locally (Adelson and Bergen 1985), followed by the integration of these local responses at a larger spatial scale at a second stage, which has been associated to area MT on experimental grounds

---

[1] "Although aware that units may be direction selective, Hubel and Wiesel have not emphasized this property and it is not considered in any of their tables. In this connection, however, it is interesting to note that, for an extended edge or slit much longer than the dimensions of the receptive field there are only two possible directions of movement, namely the two at right angles to the orientation. This is simply a geometrical necessity. Although orientation necessarily determines the direction of stimulus movement, which of the two possible directions will be effective is independent of the orientation". Bishop et al. (1971). See also Henry and Bishop (1971).

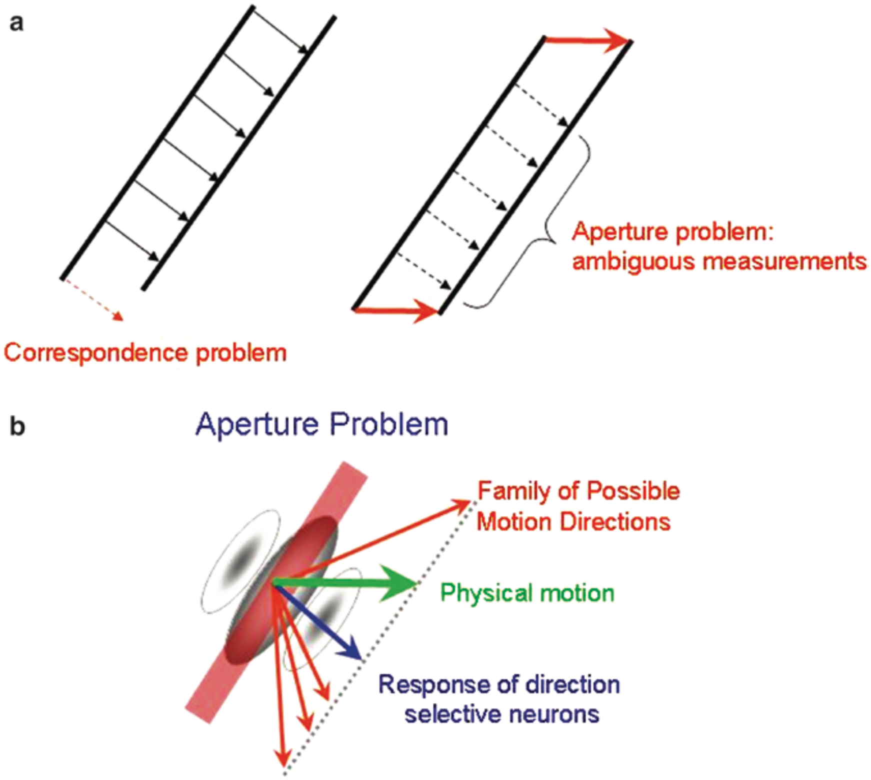

Fig. 1.1 Aperture and correspondence problem. *Top*: Illustration of the correspondence problem. Two frames of a bar moving horizontally are shown. A correspondence established over time using the shortest path – i.e. the lowest speed – leaves parts of the contour unpaired. . *Bottom*: A straight contour crossing the receptive field of a single direction selective neuron elicits the same response for a large family of physical motions. The cell responds only to the motion vector orthogonal to the cell preferred orientation

(Movshon et al. 1986; but see Majaj et al. 2007). Do, however, these two contours belong to a single translating shape or object, conditions required to justify the combination of both measurements, or do they belong to two different shapes or objects, in which case combining these motion measurements would distort the physical stimulus and yield a false solution? Answering the question clearly requires additional constraints for this calculation to be functionally relevant, a point addressed later on.

Another way of solving the "aperture problem" is to use the motion energy available at 2D singularities such as the line-endings of a finite unitary contour. These singularities can be seen as local geometrical features in visual space but are also characterized by their spatial frequency spectrum. As a matter of fact, these singularities of limited spatial extent have a wide energy spectrum in the Fourier plane with a large distribution of orientations and spatial frequencies with different power and phase.

As visual neurons behave as spatial frequency filters (Campbell and Robson 1968; De Valois 1979), singularities provide a rich set of possible motion measurements to spatial frequency and orientation tuned sensors, whose combination can signal the veridical direction of motion, at least for translations in the fronto-parallel plane. In addition or alternately, these local features can be matched or tracked from one position to the next, offering a "simple" solution to the correspondence problem through a feature matching process.

The influence of line-ends or terminators on motion interpretation was first analyzed by Wallach (1935; but also see Silverman and Nakayama 1988; Shimojo et al. 1989) who found that the perceived direction of a moving contour was strongly determined by the apparent direction of line-ends motion, whether these were real line-ends intrinsically belonging to the contour itself or spurious line-ends extrinsically defined by occluders. One question remains: what happens to the measurements of motion performed by direction selective neurons stimulated by the inner part of a moving contour? Consider the following alternatives:

1. Each motion signal from a single neuron is an independent labeled line on which the brain relies to infer the distribution of movements in the outside world. Under this assumption, a single moving contour would appear to break into the different directions of motion signaled by different direction selective neurons. This would not favor the survival of organisms endowed with such an apparatus!
2. Ambiguous motion signals that may not carry reliable information about the physical direction of contour motion are ignored or discarded. Only motion signals corresponding to line-endings are taken into consideration. Under this assumption, what would be the neuronal response that substantiates the contour unity? In addition, discriminating a translation from an expansion would be difficult if each line-end was processed independently.
3. All neurons have the same status regarding the encoding of stimulus direction that is; each response to a moving bar is considered an equal "vote" in favor of a particular direction of motion. Under this assumption, the resulting direction, processed through some kind of averaging of neuronal responses, would not necessarily correspond to the physical motion.

How then is it ever possible to recover the direction of a contour moving in the outside world? One possibility is to weigh these different votes according to some criterion, such as their reliability or salience (Perrone 1990). But again, what homunculus decides that this particular "vote" has less or more weight than the other one, especially if the "voters" are neurons whose receptive fields have similar spatio-temporal structure and function, like the simple and complex direction selective cells discovered by Hubel and Wiesel and thoroughly studied since?

Hildreth (1984) proposed an alternative according to which the ambiguous responses of neurons confronted with the aperture problem would be constrained so as to match the reliable measurements at 2D singularities. She offered a "smoothness constraint" rule – whereby information from measurements at singularities "propagates" along a contour – and elaborated a computational model that recovers the velocity of curved contours. However, the neural implementation of the mechanisms underlying the propagation process along contours still remains an open issue. Others (Nowlan and Sejnowski 1995) developed computational models that implement a selective integration through a weighting process in which the reliability of a measure results from an estimation procedure. However, it remains unclear how this estimation might be implemented in the brain.

Thus, although the seminal work of Hubel and Wiesel helped us to understand what the trees are, we still need to understand what is the forest, which, in modern terms, is captured by the concept of "functional assembly," still to be constrained by experimental data to fully characterize what constitute the "unitary representation" of a moving contour and the activity within a neuronal assembly that provides a "signature" of this unity.

More generally, central questions that should be answered to understand how biological organisms recover the velocity – speed and direction – of objects are the following:

1. When should individual neuronal responses be "linked" into functional assemblies?
2. What is the functional anatomy that underlies this linking, or binding, process?
3. Are the mechanisms identical throughout the visual system, or are there specific solutions at, or between, different processing stages?
4. What are the rules to select, and mechanisms used to select weight and combine the responses of direction selective neurons?

In the following, I briefly review experimental data that suggest a possible neuronal mechanism to smoothing, analyze the dynamics of contour processing and its contrast dependency, address the issue of motion integration, segmentation and selection across moving contours and describe how form constraints are involved in these processes. In the end, I'll attempt to ascribe a functional network to these different processes.

1.2 Propagating Waves Through Contour Neurons: Dynamics Within Association Fields

Neighboring positions in the visual field are analyzed by neighboring neurons in the primary visual cortex, acting as a parallel distributed spatio-temporal processor. However, distant neurons with non-overlapping receptive fields but tuned to similar orientations aligned in the visual field do not process incoming information independently. Instead, these neurons may form a "perceptual association field" linking local orientations into an extended contour. Reminiscent of the Gestalt rule of good continuity and closure, its characteristics were experimentally specified by Field et al. (1993) and Polat and Sagi (1993), although with different paradigms. The particular structure of association fields fits well with the architectony of long-range connections running horizontally in primary visual cortex over long distances (up to 8 mm, Gilbert and Wiesel 1989; Sincich and Blasdel 2001). Moreover electrophysiological responses to contextual stimuli (Kapadia et al. 1995, 2000; Bringuier et al. 1999) suggest that horizontal connectivity is functionally relevant for contour processing (Seriès et al. 2003 for a review).

In addition, optical imaging (Jancke et al. 2004) and intracellular recordings (Bringuier et al. 1999) bring support to the idea that lateral interactions through long-range horizontal connections propagate across the cortex with propagation speeds ranging between 0.1 and 0.5 m/s, which corresponds to speeds around 50–100 °/s in visual space. Recent work in psychophysics, modeling and intracellular recordings further suggest that these slow dynamics can influence the perception of motion (Georges et al. 2002; Seriès et al. 2002; Lorenceau et al. 2002; Alais and Lorenceau 2002; Cass and Alais 2006; Frégnac et al., this volume).

This is for instance the case with the Ternus display (Fig. 1.2) in which, the perception of group or element motion can be seen in a two frames movie, depending upon, amongst many other parameters, the time delay between frames. Alais and Lorenceau (2002) observed that for a given delay, group motion is seen more frequently when the Ternus elements are collinear and aligned as compared to non-oriented or non-aligned elements. This finding indicates that "links" between

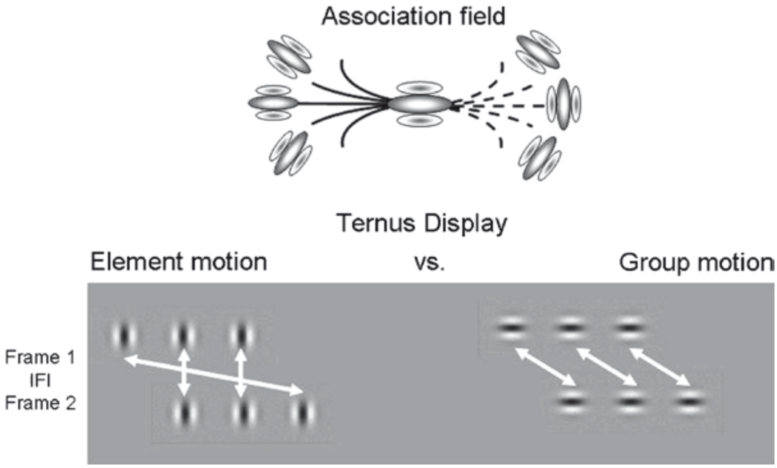

Fig. 1.2 Illustration of the "association field" depicting the spatial configurations that can (*left*) or cannot (*right*) be detected in an array of randomly oriented contour elements. This perceptual "association field" is presumably implemented in the network of long-range horizontal connections running horizontally within V1 (Gilbert and Wiesel 1995; Sincich and Blasdel 2001). In this figure, schematic oriented receptive fields interact through facilitatory long-range horizontal connections when the gestalt criterion of good continuity is met (*black lines*). When it is not (*dashed lines*), these long-range connections may be absent, ineffective or suppressive, a point that is still debated. *Bottom*: Illustration of the Ternus display of Alais and Lorenceau (2002) consisting in three oriented elements presented successively in a two frames movie. When the oriented elements are aligned and collinear (*right*), group motion is seen more often than when they are not (*left*). In this case element motion is seen more often. It is proposed that these different percepts of group and element motion reflect the links established between collinear and aligned element trough long-range associations

elements defining a pseudo continuous contour have been established and that strengthen the association between elements then considered a “whole.”

A possible explanation is that horizontal connections provide a mean to bind individual neuronal responses into a functional assembly signaling a unitary contour moving as an ensemble in a single direction. This mechanism would have the advantage of being highly flexible, such that a functional assembly would easily adapt, within limits, to contours of varying length and curvature. An open issue is whether and how the association field characterized with static stimuli is used in motion processing. In this regard, it should be noted that eye movements of different kinds constantly shift the image on the retina such that different neurons, forming different assemblies, are recruited, even with static images. Thus, a continuous updating of the links to the incoming stimulus is required for “static” images as well as for moving stimuli, raising the possibility that association fields are relevant in motion processing as well.

1.3 Dynamics of Contour Motion Perception

Up to now, the need for combining motion measurements across space and time to recover a contour direction stems from theoretical considerations related to the initial sampling of the retinal image by cells with restricted receptive fields. If true, the computation of a global solution – e.g. Hildreth’s smoothing process – may not be instantaneous and could take time. The finding that indeed the perception of a moving contour smoothly develops and unfolds over a period of time in a measurable way (Yo and Wilson 1992; Lorenceau et al. 1993) brings support to the idea that recovering the direction of moving contours involves an integration process endowed with a slow dynamics. In psychophysical experiments, Lorenceau et al. (1993) found that an oblique contour moving along a horizontal axis first appears to move in a direction orthogonal to contour orientation which smoothly shifts over tens of milliseconds towards the real contour direction (Fig. 1.3a, see Movie 1). This perceptual dynamics was found to depend on contour length and contrast such that a biased direction was seen for a longer time with longer contours and lower contrasts. In the framework described above, the effect of contour length is expected as it can be accounted for by the recruitment of a larger population of cells facing the aperture problem relative to those processing line-ends, thereby contributing to a strong bias toward an orthogonal perceived direction (Fig. 1.3b), that takes time to overcome.

The larger bias observed at low contrasts remains a matter of debate, although there is an agreement to consider that the sensitivity to the direction of 2D singularities the grating’s or contour’s line-ends is in cause. As mentioned above, these singularities are characterized by a broad spatial frequency and orientation spectrum. Decreasing contrast may therefore bring some frequencies close or below detection threshold in which case cells tuned to spatial frequencies and orientations with weak energy would respond poorly and with long latencies, thus degrading the global directional response or slowing down its recovery (Majaj et al. 2002). A model based on a parallel neuronal filtering through V1 receptive fields, followed by response pooling by MT neurons could thus account for the contrast effect.

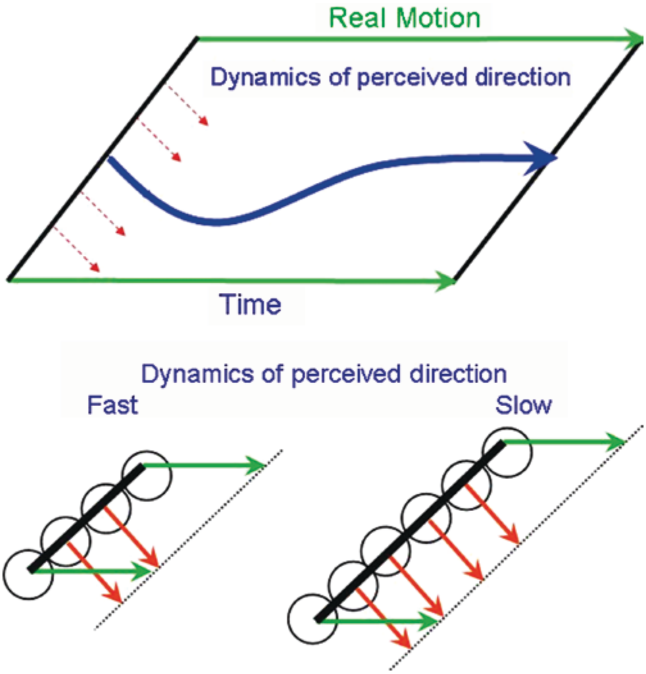

Fig. 1.3 Perceptual dynamics of an oblique bar moving horizontally. The perceived direction at motion onset is orthogonal the segment orientation and then smoothly tunes with the physical motion. The dynamics of the directional shift depends on the contour length (*bottom*), presumably because of the imbalance between the number of cells recruited by the inner part of the contour and its line-ends. The dependence of the dynamics on contrast may reflect a lower sensitivity to line-ends. (Lorenceau et al. 1993; see Movie 1). These perceived shifts are well correlated to the dynamics of MT cell response (Pack and Born 2001) are found in ocular following itself in and pursuit eye movements (Masson 2004)

A second possibility is that these singularities are analyzed by neurons with center-surround organization, often referred to as hyper complex or end-stopped cells (Hubel and Wielsel 1968) whose structure and function make them well suited for the processing of line-endings' motion or local curvature (Dobkins et al. 1987; Pack et al. 2003a, b; Lalanne and Lorenceau 2006). Sceniak et al. (1999) recorded such neurons in macaque V1 and observed that the end-stopping behavior found at a high contrast is decreased at low contrast, such that their capability to process line-ends' motion is degraded. This pattern of response could explain the longer integration time found at low contrast. Interestingly, this type of neurons mostly lies in the superficial layer of V1 where response latencies are longer than in other intermediate layers (Maunsell and Gibson 1992), suggesting that their contribution to motion computation is delayed relative to the simple direction selective neurons of layer 4.

In an attempt to decipher between these explanations (although different mechanisms could be simultaneously at work), Lalanne and Lorenceau (2006) used a Barber pole stimulus – an oblique drifting grating seen as moving in the direction of the line-ends present at the grating's borders. A localized adaptation paradigm was used in order to selectively decrease the sensitivity of the putative neurons underlying line-ends processing. Decreasing the contribution of these neurons to the global motion computation should increase the directional biases toward orthogonal motion thus allowing to isolate the spatial location and structure of the adapting stimulus that entails the largest biases. To get insights into the neuronal substrate at work, high contrast adapters were positioned in different locations at the border or within the grating and their effects on the subsequent grating's perceived direction measured. The results show that the largest directional biases are produced by adapters located within the grating itself and not at the line-endings positions. Although this "remote" effect of adaptation may seem surprising at first sight, it is compatible with a model in which the difference in response of two simple cells gives rise to end-stopping (Dobkins et al. 1987), but at odd with the idea that line-ends' direction is recovered by the parallel filtering of V1 receptive fields at line-ends positions (e.g. Löffler and Orbach 1999).

Neuronal counterparts of the perceptual dynamics underlying the recovery of moving contours described above have been found in macaque MT (Pack and Born 2001; Majaj et al. 2002; Born et al. this issue). In addition, ocular following was also found to manifest similar dynamical directional biases during its early phase with pursuit being deviated towards the normal to contour orientation (Masson et al. 2000; Barthélemy et al. 2008; see Chap. 8).

Altogether these psychophysical, behavioral and electrophysiological results indicate that recovering the motion of the simple moving bar used by Hubel and Wiesel in the sixties is a complex time consuming process that involves a large population of neurons distributed across visual cortex and endowed with different functional characteristics. As complex objects are generally composed of a number of contours at different orientations, understanding how biological systems overcome the aperture problem when processing objects' motion should take these findings into account.

1.4 Integration, Segmentation and Selection of Contour Motions

As stated above, the combination of responses to multiple component motions offers a way to overcome the aperture problem so as to recover object motion (e.g. Fennema and Thompson 1979; Adelson and Movshon 1982). In order to assess the underlying perceptual processes several classes of stimuli have been used to:

1. Measure the global perceived velocity and to determine the computational rules involved in motion integration
2. Evaluate the conditions under which component motions can, or cannot, be bound into a whole
3. Identify the neural substrate and physiological mechanisms that implement these perceptual processes

The numerous kinds of stimuli used to explore these issues can be broadly divided in three classes: Plaids, Random Dot Kinematograms (RDKs) and “aperture” stimuli. Before trying to offer a synthetic view of the results, let us spend some time discussing the appearance and relative significance of these different stimuli (Fig. 1.4).

Made of two extended overlapping gratings at different orientations, drifting plaids can be seen as a single moving surface or as two sliding transparent surfaces, depending on their coherency. As plaids are well defined in the Fourier plane by their

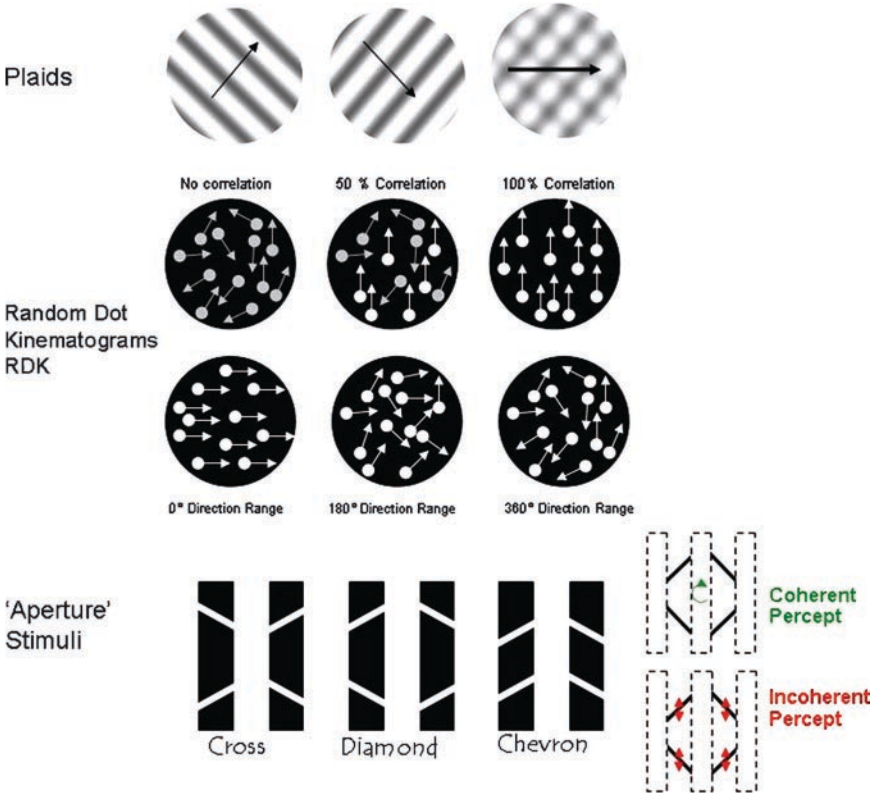

Fig. 1.4 Different stimuli used to probe contour integration. *Top*: Plaid patterns made of two superimposed gratings. Changes of relative orientation, contrast, speed, spatial frequency have been used to determine the conditions of perceived coherence, the perceived direction and speed and the nature of the underlying combination rule. *Middle*: Two types of random dot kinematograms (RDKs). In one, the percentage of coherently moving dot is used to assess motion sensitivity. In the second, dot directions are chosen amongst a distribution of direction varying in width to characterize directional – and or speed – integration. *Bottom*: “Aperture” stimuli where a moving geometrical figure is partially visible behind aperture or masks. Each figure segment appears to move up and down. Recovering figure motion requires the spatio-temporal integration of segment motions. Changing figure contrast or shape, aperture visibility or luminance, duration, eccentricity deeply influences perceived rigidity and coherence and may impair the ability to recover object motion