Cambridge University Press
978-1-107-03434-1 - Simplicity in Vision: A Multidisciplinary Account of Perceptual Organization
Peter A. Van Der Helm
Excerpt
More information

Prologue

Levels of vision, description, and evaluation

Scientific research is an endeavor to enable us — via metaphors, theories, and models — to understand and thereby control reality. We may never be able to fully understand reality, however, because the main tool we use to understand reality, namely our brain, is an inextricable part of reality. At best, so it seems, scientific research may arrive at an understanding of reality as we experience it subjectively, that is, acknowledging the workings of the brain. Understanding the latter is the objective of cognitive neuroscience. Even so, cognitive neuroscientists too use their brain as a tool to understand data. Just as in daily life, this is a potential pitfall because, as an abundance of visual illusions are proof of, what you think you see is not always what you look at.

In this context, human vision research is special in that it takes vision not only as mediating instrument but also as the very topic of study (cf. Rock, 1983). It recognizes that vision is a complex yet fast process that organizes meaningless patches of light on the retina into the objects we perceive, that is, objects with potentially meaningful properties such as shape and spatial arrangement of parts (see Fig. P.1). In other words, when we look at a scene, the objects we perceive constitute the output of vision — not its input. This also illustrates the unicity of vision: Whereas every ordered thing in the world eventually degrades into chaos, vision creates order in the chaos of retinal patches of light. How vision solves this "inverse problem" is the central topic in vision research.

Another inversal is the following. Vision is a process that proceeds from input to output, but to understand this process, one has to research one's way from its output back to its input (see Fig. P.2). One has to start from visual experiences and experimental observations, to infer (a) what the nature is of the mental representations of percepts; (b) how

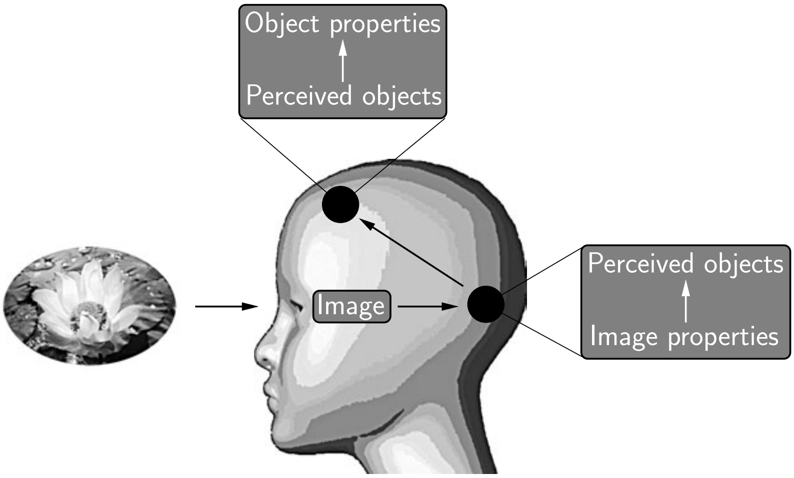

Fig. P.1. Light entering the eyes yields a two-dimensional retinal image of a scene. Properties of this image then are extracted to yield a three-dimensional perceptual organization of the scene in terms of objects arranged in space. Properties of these perceived objects and of their perceived arrangement in space then determine how we might act upon what we think we see.

cognitive processes proceed to yield these representations; and (c) how these processes and representations are neurally realized.

Eventually, vision research may arrive at a "grand unified theory", though I am reluctant to give this the metaphysical or ontological reading of pluralism as "one all-embracing story". That is, even such a theory will probably consist of a set of coherent but still separate subtheories that, each, account for some aspect of vision. This holds more generally in cognitive neuroscience. For instance, neuroscientists may argue that love and near-death experiences result from biochemical interactions between neurons — and they may be right — but this does not do justice to people's conscious experiences, which call for another story. Also for vision, it seems necessary, or at least expedient, to decompose the total scientific question into smaller questions, which may lead eventually to a set of separate but coherent answers to the total question. This explanatory or epistemological method of analysis and synthesis is also promoted in this book. More specifically, I adopt the following decompositions into levels of vision, levels of description, and levels of evaluation.

Cambridge University Press
978-1-107-03434-1 - Simplicity in Vision: A Multidisciplinary Account of Perceptual Organization
Peter A. Van Der Helm
Excerpt
More information

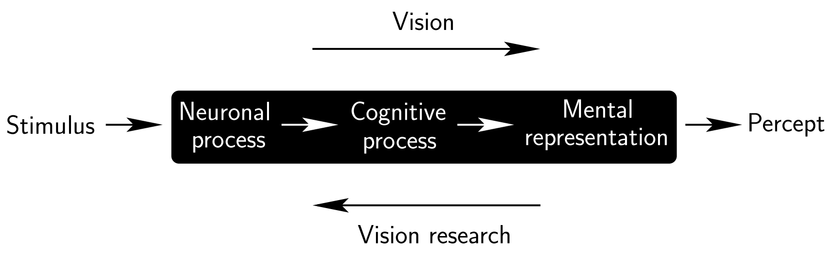

Fig. P.2. The black box of vision. Vision is the process from stimulus to percept. To understand this process, however, vision research has to work back from percepts, via mental representations, to cognitive and neuronal processes. As philosopher Søren Kierkegaard (1813–1855) put it: *"Life can only be understood backward"*.

Levels of vision

In the general field of vision research, topics range from function (what does the visual system do?) to functionality (how do its outcomes influence behavior?). More specifically, one usually distinguishes between three levels of vision (see also Fig. P.1):

1. *Low-level vision*, which concerns feature extraction from retinotopic image information in sensory registers.
2. *Middle-level vision*, which concerns perceptual organization, that is, binding and selection of features into integrated percepts.
3. *High-level vision*, which concerns everyday interactions between perceptual organization and higher cognitive faculties.

This book is dedicated to the pivotal issue of perceptual organization. It focuses therefore on middle-level vision, but it also extends to low-level vision and high-level vision. For instance, Chapter 2 is about the veridicality of vision, which is typically a high-level vision issue. Furthermore, Chapter 4 is about symmetry perception, which is preeminently an issue that calls for inclusion of all three levels of vision.

Levels of description

Fairly independent of the decomposition above is the decomposition into the three complementary levels of description or analysis which Marr (1982/2010) proposed for research on information-processing systems like the visual system. Just as computer programmers have to solve the problem to compute something (the goal) by way of an algorithm (the method) implemented in certain hardware (the means), Marr argued that vision research should distinguish between the goal, the method, and the means of vision. Eventually, compatible descriptions of these three complementary aspects may, together, explain how the goal is reached by a method allowed by the means. More specifically, Marr distinguished:

1. *The computational level*, at which the goal of a system is specified in terms of systematicities in its output as a function of its input; applied to the visual system, this level concerns the question of what logic defines the nature of resulting mental representations of incoming stimuli.
2. *The algorithmic level*, at which the method of a system is specified in terms of the mechanisms that transform its input into its output; applied to the visual system, this level concerns the question of how its input and output are represented and how one is transformed in the other.
3. *The implementational level*, at which the means of a system is specified in terms of the hardware of the system; applied to the visual system, this level concerns the question of how those representations and transformations are neurally realized.

This book starts at the computational level, that is, from ideas about the nature of mental representations. Chapter 4 on symmetry perception, however, reaches out to the other two levels, and Chapter 5 focuses on the algorithmic level by presenting a formal process model of the perceptual organization of symbol strings that may represent visual stimuli. Furthermore, with an eye for metaphors of cognition, Chapter 6 combines all three levels to arrive at a neurally plausible picture of the cognitive architecture of perceptual organization.

Levels of evaluation

Again fairly independent of the decompositions above is a decomposition into three levels of evaluation of theories and models. The meaning of the terms "theory" and "model" varies somewhat with research domains, and the border between theories and models is indeed fuzzy. In general, however, a theory rather is a set of conceptual ideas, whereas a model rather is an applicable elaboration of those ideas. This implies that both theories and models may have predictive power but that falsifiability is an issue which applies to models rather than to theories. Hence, to evaluate theories and models, the conceptual plausibility of ideas formulated in theories can be assessed as such, while to these ideas, models are vehicles that can be tested more directly on predictive power and, not to be forgotten, practical feasibility. This implies three cycles of research to enhance, revise, or reject theories and models (see Fig. P.3):

1. *The theoretical cycle*, which has roots in mathematics; the idea is to formalize ideas and assumptions in theories and models, to see if they can be derived from first principles (i.e., from facts proved earlier).
2. *The empirical cycle*, which has roots in physics; the idea is to conduct controlled experiments to test predictions inferred from theories and models (de Groot, 1961/1969).
3. *The tractability cycle*, which has roots in computer science; the idea is to assess if theories and models allow for feasible implementations in computers or brains (van Rooij, 2008).

The theoretical and tractability cycles are not yet broadly recognized in cognitive neuroscience. That is, though typically suited for a

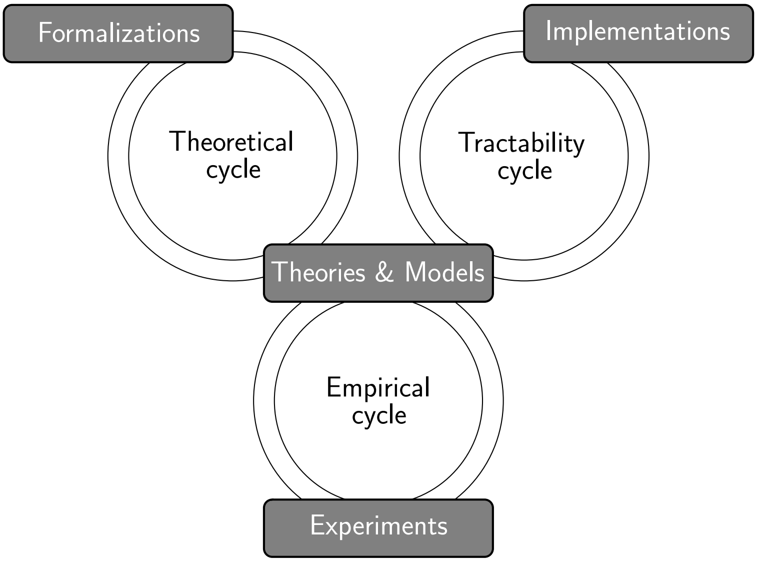

Fig. P.3. The theoretical, empirical, and tractability cycles of research. Theories and models can be enhanced, revised, or rejected on the basis of feedback from attempts to formally underpin conceptual ideas, to test predictions, and to implement tractable mechanisms, respectively.

multidisciplinary approach, cognitive neuroscience is dominated by experimental research and, therefore, by the empirical cycle. It is of course true that experimental data are relevant, but they are often also multi-interpretable. To evaluate a theory or model, evidence from the theoretical and tractability cycles may therefore be equally relevant to arrive at compelling scientific conclusions.

The relevance of all three cycles is acknowledged by the division of this book into three parts:

- In Part I, the focus is on simplicity: The theoretical cycle is passed through to assess the veridicality of the simplest organizations that are proposed to be produced by the visual system. My conclusion will be that — provided one includes both view-dependent and view-independent aspects of vision — simplest organizations are sufficiently veridical to guide us through many different environments.

- In Part II, the focus is on visual regularity: The empirical cycle is passed through to test concrete predictions inferred from the mathematically established hierarchically transparent and holographic nature of the regularities that are proposed to be exploited to arrive at simplest organizations. My conclusion will be that this mathematically unique nature is indeed pertinent in vision and daily life.

- In Part III, the focus is on hierarchy: The tractability cycle is passed through to assess the neural plausibility of the proposed hierarchical organization process. My conclusion will be that this process is mediated by transient, temporarily synchronized, neural assemblies — dubbed gnosons (i.e., fundamental particles of cognition) — whose synchronization is the neural signature of transparallel feature processing. For strings, this form of proccessing — which is feasible on classical computers — is as powerful as quantum computing promises to be, and does justice to the high combinatorial capacity and speed of the human perceptual organization process.

Taken together, these three parts provide, in my view, a firm and coherent underpinning of the proposed role of simplicity in human perceptual organization — not only regarding its empirical adequacy but also regarding its theoretical soundness and its practical feasibility.

A multidisciplinary approach

The field of human vision research is too broad to address all aspects in one book. This book does not deal with color perception, depth perception, and motion perception, for instance. It also does not pretend to provide a "grand unified theory" of vision. Yet, guided by the above distinctions between levels of vision, description, and evaluation, it does address a number of fundamental issues related to the problem of perceptual organization in form and shape perception.

These issues are diverse in character, and indeed, this book promotes a truly multidisciplinary approach to perceptual organization. More specifically, the issues addressed in this book can be situated at the crossroads of experimental psychology, cognitive science, neuroscience, artificial intelligence research, mathematics, computer science, graph theory, evolutionary biology, and philosophy of science. By way of two-way interactions between these research areas, this book aims to contribute to a better understanding of perceptual organization and, thereby, of cognition as a whole. In the Epilogue, I combine my findings to sketch the contours of what might be such a better understanding, and I specify my hope that this advances not only the just-mentioned research areas but, by way of spin-off, also vision-related research and application fields.

Part I
The theoretical cycle

The idea of the theoretical cycle of research is to formalize assumptions, to see if they can be underpinned by derivations from first principles. This method is characteristic of mathematics, in which a theorem usually starts as a conjecture that calls for a proof. The search for a proof may be successful, but may also lead to the conclusion that the conjecture is false or has to be adjusted to be provable. A successful proof means that the correctness of the conjecture can be derived logically from facts proved earlier, and hence, from first principles.

In this first part, Chapter 1 sets the stage by presenting an overview of visual information-processing ideas adhered in structural information theory (SIT). The central idea in SIT is the simplicity principle, which holds that the visual process yields simplest organizations of stimuli. An implicit assumption then is that such organizations have evolutionary survival value in that they are sufficiently veridical to guide us through the world. In Chapter 2, this assumption is addressed in a historical and multidisciplinary setting, using findings from the mathematical domain of algorithmic information theory (AIT). SIT and AIT developed independently of each other, but provide similar modern alternatives for Shannon's (1948) classical selective-information theory.

Notice that Chapter 2 contains mathematical proofs which, however, do not pinpoint the exact degree of veridicality of simplest perceptual organizations (which is probably impossible; see also the Prologue). Yet, as I argue in Chapter 2, these proofs are relevant in that they provide theoretical evidence that simplest perceptual organizations are fairly veridical in many different worlds, possibly including the world we live in.