Vision Research 51 (2011) 1397–1430

Contents lists available at ScienceDirect

Vision Research

journal homepage: www.elsevier.com/locate/visres

Review

Beyond multiple pattern analyzers modeled as linear filters (as classical V1 simple cells): Useful additions of the last 25 years

Norma V. Graham *

Dept. of Psychology, Mail Code 5501, Columbia University, NY, NY 10027, USA

ARTICLE INFO

Article history:
Received 3 May 2010
Received in revised form 7 February 2011
Available online 15 February 2011

Keywords:
FRF
Second order
Contrast gain
Contrast normalization
Contrast comparison
Surround suppression
Lateral connections
Contour integration

ABSTRACT

This review briefly discusses processes that have been suggested in the last 25 years as important to the intermediate stages of visual processing of patterns. Five categories of processes are presented: (1) Higher-order processes including *FRF* structures; (2) Divisive contrast nonlinearities including contrast normalization; (3) Subtractive contrast nonlinearities including contrast comparison; (4) Non-classical receptive fields (surround suppression, cross-orientation inhibition); (5) Contour integration.

0. Introduction

We were asked for the 50th anniversary issue of *Vision Research* to highlight new knowledge on important questions open 25 years ago and on which progress had (or had not) been made. In a happy coincidence for me, 25 years ago I had just completed the draft of a book (published as Graham, 1989, summarized in a short paper Graham, 1992). I am reasonably certain, therefore, of what was known 25 years ago about a set of questions in pattern vision, or at least of what I thought was known.

The *simple multiple-analyzers model* shown in Fig. 1 top panel seemed at that time to be a very good model of pattern vision, particularly when you limited your attention to experiments using visual patterns of *near-threshold contrast*. In this model there were multiple analyzers, each of which was selectively sensitive on at least one of the multiple dimensions of pattern vision. These dimensions included spatial frequency, spatial position, orientation, direction of motion, and a number of others. To get from these multiple analyzers to the observer's response the model used a *decision rule* that was just a very simple combination of the multiple analyzers' outputs, e.g.: the observer says the pattern is vertical if and only if the analyzer producing the biggest output is the analyzer having peak sensitivity at the vertical orientation.

An aside about terms and the glossary: Many terms used in the main text without much definition are described more fully in the glossary. These terms appear in italics at least when they are first introduced. (Some italicized terms are not in the glossary but are italicized for momentary emphasis, or because they are titles of other sections in this review, or for other conventional reasons.)

The physiological substrate for an analyzer might be considered to be either a single neuron, or a set of neurons that are homogeneous in some sense (e.g. all sensitive to vertical orientation but in different spatial positions). To minimize blatantly neurophysiological terms when talking about concepts used to explain behavior, the word *unit* will be used here to mean a more abstract entity analogous to a single neuron, and the word *channel* will be used here to mean a more abstract entity analogous to a set of neurons that are homogeneous in some sense. The word *receptive field*, although it has its origin in the neurophysiological literature, is less blatantly neural, and both units/channels and neuron/neurons will be said to have *receptive fields*.

Twenty-five years ago the analyzers were generally based on the classical model of one of the types of neurons Hubel and Wiesel had discovered in cortical area V1 (striate cortex), the type called *simple cells*. According to the classical model, a simple cell adds and subtracts the weighted amount of stimulation of the excitatory and inhibitory areas in its *receptive field*. Since a neuron's output is spikes, and since spike rates lower than zero do not exist, a *half-wave rectification* or similar nonlinearity was assumed to change any below-zero result of the addition and subtraction into zero.

* Fax: +1 212 854 3609.
E-mail address: nvg1@columbia.edu

0042-6989
doi:10.1016/j.visres.2011.02.007

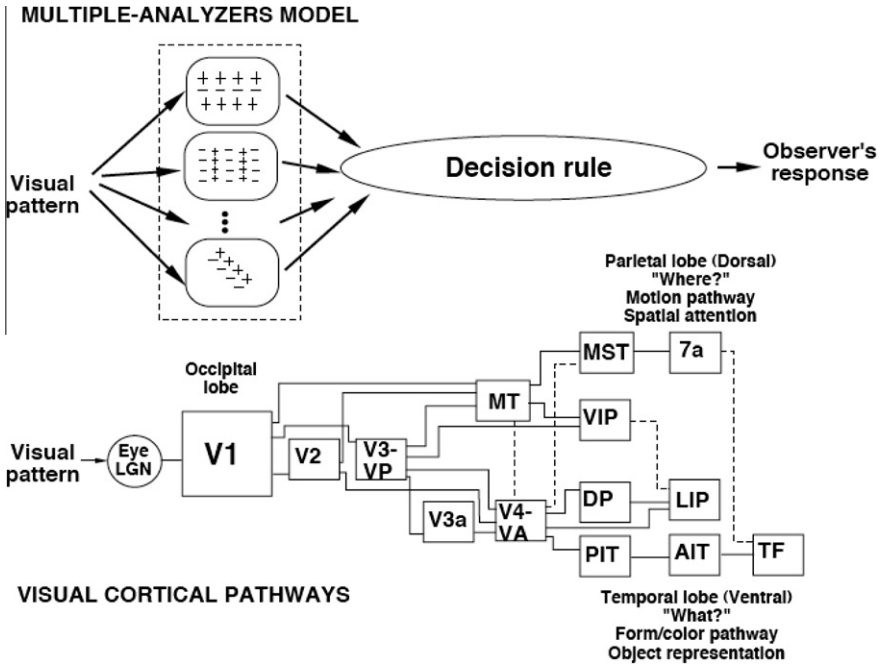

Fig. 1. Simple multiple-analyzers model (top). Simplified sketch of visual pathways (bottom, based on Movshon (1990)).

(This rectification was left implicit frequently, while referring to the model as a linear system. This common practice has led to some confusion.) Thus we will define a *classical V1 simple cell* as a *linear system* (an adding and subtracting device) followed by a *half-wave rectification*.

This classical V1 simple cell model was in decent accord with known physiological results of the time. It turned out NOT to be in complete accord with the physiology, however, as is discussed below. Hence a distinction is made here between a *classical V1 simple cell* (one that is perfectly described by the classical model) and a *simple cell* (any V1 cell that would be classified as a simple cell by the criteria ordinarily used by physiologists of Hubel and Wiesel's time or today).

The simple multiple-analyzers model shown in the top panel of Fig. 1 was and is a very good account, qualitatively and quantitatively, of the results of psychophysical experiments using near-threshold contrasts. And by 1985 there were hundreds of published papers each typically with many such experiments. It was quite clear by that time, however, that area V1 was only one of 10 or more different areas in the cortex devoted to vision. See sketch in Fig. 1 bottom panel. (Lennie (1998) and Hochberg (1998) give an interesting perspective on the complexity and functionality of a subset of these cortical areas, V1 through V4 and MT.) The success of this simple multiple-analyzers model seemed almost magical therefore. How could a model account for so many experimental results when it represented most areas of visual cortex and the whole rest of the brain by a simple decision rule? One possible explanation of the magic is this: In response to near-threshold patterns, only a small proportion of the analyzers are being stimulated above their baseline. Perhaps this sparseness of information going upstream limits the kinds of processing that the higher levels can do, and limits them to being described by simple decision rules because such rules may be close to optimal given the sparseness. It is as if the near-threshold experiments made all higher levels of visual processing transparent, therefore allowing the properties of the low-level analyzers to be seen.

Even for near-threshold experiments, there were hints of extra non-linear inhibition among analyzers (Graham, 1989). And for supra-threshold psychophysical results (although not in the 1989 book, there were many tens of published papers I knew very well) a satisfactory decision rule would either be very complicated or very vague. Given that V1 is one area of many known visual areas in the brain, and that even V1's physiology was known to be more complicated than the classical model of V1 simple cells, this was not very surprising. But it was unclear how to improve the model and yet keep it tractable and useful.

In the last 25 years a number of processes have been suggested as possible additions to the simple multiple-analyzer model of Fig. 1, additions which have the flavor of intermediate stages of visual processing, of stages for which the physiological substrate might be V1 (or perhaps V2 or V3). These stages might be called the "hidden stages" as they are far from the light image that stimulates the eye and far also from both conscious perception and the control of action. Several of these suggested additions to the simple multiple-analyzers model are the substance of this review. They have been suggested as explanations of pattern vision in general, both for psychophysical and neurophysiological results. Here the discussion is focused on the psychophysical side, but the neurophysiological is too intertwined in the history to be ignored entirely. (A multi-author paper from a mini-symposium in the early 2000s – Carandini et al., 2005 – is one convenient source for more about the physiological side as are other of the articles in this volume.)

I will discuss these additional processes as falling into the five categories listed below, and the rest of the article will be organized by these five categories. The general categories are neither mutually exclusive nor exhaustive. There are specific examples in each category, however, which are distinct from examples in other categories and which seem to present distinct computational advantages and to give different perspectives on desirable functionality. The list below is ordered for ease of exposition as I could find no more systematic order (e.g. chronological) that turned out to be satisfactory or useful.

0.1. List of five categories of additional processes

Addition 1. Higher-order processes (including FRF structures).
Addition 2. Divisive contrast nonlinearities (including contrast normalization).