On raising our sights

PETER C. DODWELL
Department of Psychology, Queen's University, Kingston, Ontario, K7L 3N6 Canada

Received for publication 11 October 1993

Abstract—There are some very strong traditions in contemporary research on spatial vision that have arisen, in large measure, through advances in technology. This is especially true of research on the neurophysiology of vision, e.g. in recording from single cells, and in developments in computing that have promoted the study of machine vision. No one would dispute that these are major advances in the field, but one may ask whether they have unduly constrained the sorts of question that we ask about the visual processes. It is argued that they have, and that it is timely to stand back to see what other problems and questions have suffered neglect. Two items seem to be quite ripe for reconsideration: the question of how to understand the enormous adaptability of visual systems to sensory rearrangement (by means of distorting lenses, for example); and the very broad matter of keeping our perceptual processing in close and coherent touch both with stimulation, and with the properties of the physical/biological environment—what Gibson called the problem of veridical perception. It is argued that the two items are closely allied, and can be approached from a single theoretical stance. The one I favour is outlined, and I claim that it is one that has little to do, in the first instance, with either neurophysiology or computation. In that sense tackling the two questions involves raising our sights above the current traditions of vision research, but that of course does not in any way mean attempting to belittle the immense advances that those traditions have brought about. Theoretical and experimental tools are available for tackling both these questions, so one hopes the field may open up a bit to encompass new research on them.

At the end of a reasonably long career as a researcher, perhaps I may be permitted a few remarks about the state of the art in vision research, what I think its strengths and weaknesses are, and where it might be headed. Or, should I rather say, where I think it ought to be headed.

STIMULUS CODING

When I started out, in the 1950s, Gestalt Psychology was in the process of receiving the coup de gràce from Lashley, Sperry, and others, to the extent that their physiological studies showed no evidence of the electrostatic cortical fields purported to be the basis for visual pattern recognition. At the same time Lashley, Pitts and McCulloch, Klüver, and others were searching for a coherent explanation for what used to be called Stimulus Equivalence. The logical bases for this search were not really clearly articulated, but it was accepted without question that an explanation would involve efficient procedures for mapping a variable stimulus input into a relatively small number of categories called patterns, and that this should be done with neuropsychologically plausible resources.

None of those early attempts were particularly successful, but they did lay the groundwork for impressive gains at a later date. A model for stimulus coding devised by Deutsch (1955) was a milestone on the road to understanding visual
pattern recognition. He proposed effective procedures for the input–output mapping, and did so in the context of an exact specification of which aspects of stimulus equivalence were to be accounted for, and also in reasonable agreement with what was then known about brain function in the mammalian visual cortex. The effective procedures were early examples of what are now known as algorithms, and the structural models, as we then called them, were perhaps the very first instances of computational models in vision (Dodwell, 1970, 1992a). Thus the ideas that have dominated much of vision research in recent decades have a different ancestry, and an older tradition, than is commonly recognized.

INFLUENCE OF THEORY AND TECHNOLOGY

Most of the developments in vision research during my career have been driven by advances in technology. I refer especially to the breakthroughs in neuroscience highlighted by the ability to record from single neurons in the CNS, first exploited brilliantly by Hubel and Wiesel (1962), and to the rise of computers in all their glory (e.g., Newell 1991; Dodwell, 1992b). Theoretical advances have been, on the other hand, and relative to those developments, not so widespread or influential. I can say this because, as it happens, the two theoretical giants of the mid-century published their major ideas right at the halfway mark (Hebb, 1949; Gibson, 1950) and before my career started. The work of each one has been enormously influential, and interestingly enough (at the expense of some oversimplification), has provided respectively the engines of growth for the two fields identified above. One might identify them also with the old distinction between Nature and Nurture that dominated theoretical discussion for 100 years or more. That distinction, and the related theoretical disputes, were largely resolved, or at least rendered theoretically less interesting, by the development of technology that allowed for the detailed investigation of how the built-in circuitry of the visual system can be shaped by environmental influences (Blakemore, 1978; Goodman and Jessel, 1992).

The old dichotomy (Nature vs Nurture) has been replaced by another, equally influential and pervasive, and evidently equally difficult to put to rest. That is the distinction between ‘top-down’ and ‘bottom-up’ processes in perception. Again, one can view the latter as the stance of Hebb, insofar as he was a perceptual theorist, and the former as Gibson’s ‘station-point’, as he termed the viewer-centred place of departure for perceptual activity. It is not straining the facts too far to further identify the bottom-up view with what neuroscience can contribute to understanding perception, the top-down with computational vision’s offering. That perhaps requires a word of explanation: although the bottom-up approach has been strongly espoused by some computational researchers—notably Marr—in general there is now a strong move to consider ‘knowledge-based’ systems as the ones most likely to yield fruitful results in pattern and object recognition. There is another sense in which computer vision is indebted to Gibson’s denial of bottom-up strategies; his greatest contribution, and certainly his strongest influence in computational circles, was his insistence on the careful specification of stimulus properties of the (extended) optical array, something that requires much more than point-by-point analysis of physical descriptions of the array. Researchers in this field go beyond Gibson—indeed are quite anti-Gibsonian—in