
The combination of filters in early spatial vision: a retrospective analysis of the MIRAGE model

Michael J Morgan

Institute of Ophthalmology, University College London, Bath Street, EC1V 9EL, UK;
e-mail: m.j.morgan@ucl.ac.uk

Roger J Watt

Department of Psychology, University of Stirling, Stirling FK9 4LA, UK
Received 6 December 1996, in revised form 19 May 1997

Abstract. Since the discovery of spatial-frequency-tuned channels in the visual system, most theories attempting to account for pattern encoding have assumed that the filters can be independently accessed and flexibly combined. We review here an alternative model, 'MIRAGE', in which the filters are inflexibly combined before pattern analysis. In the MIRAGE model the half-wave rectified outputs of all spatial-frequency channels are combined before locating spatial zero-bounded regions in the neural image, which serve as the spatial primitives for pattern analysis. We describe the evidence that led to this model, and review recent evidence on the rules of filter combination.

1 Introduction

Psychophysical and physiological evidence tells us that there are mechanisms in the visual system tuned to specific bands of spatial frequency (Campbell and Robson 1964; De Valois et al 1982). These mechanisms have since formed the central theme of most models of the early stages of human vision. We adopt the normal terminology of describing each spatial-frequency-selective mechanism as a filter. Each of the different filter mechanisms could produce a potential response at every point in visual space, subject to its coverage of the retina. Any difference between two images will result in differences in the response values of the outputs of various filters at various places in visual space. There is thus a massive set of data available to the visual system with which to analyse any given image or to distinguish differences between two images. Subject only to certain boundary conditions of spatial and luminance/contrast resolution, there are no image differences that would not be represented as filter output differences. This massive set of data should be regarded as nothing more than a registration of the stimulating image. There are several general approaches to the issue that follows of how the data can be used for the performance of visual tasks.

This paper revisits a specific proposal that we made concerning this question. The MIRAGE model (Watt and Morgan 1985) is rather different from others that have been proposed since, and we wish to start by explaining the key conceptual differences.

1.1 *Approach 1: Random access in spatial frequency and space*

The simplest approach to conceive is one in which the output of any filter at any point in space can be taken, and if need be combined with any others, to produce a measurement that leads to a response. Thus, for example, the model of Wilson and Gelb (1984) for the discrimination of spatial separation (for small separations) is able to demonstrate that the limits on performance found psychophysically are close to what would be expected from known limits of differential responsiveness of the individual filters. The model shows that just discriminable stimuli cause just detectable variations in filter outputs. An identical approach is used by Wilson and Richards (1989) to account for aspects of line curvature discrimination. The measurement applied to the bank of filter outputs is different, but the limitations on performance are found to be similar: just discriminable stimuli cause just detectable variations in filter outputs.

There are two very significant assumptions of this class of model.

The first is that the bank of filter outputs, sampled in visual space, is the appropriate database for use. In essence what is assumed is that each filter output at each point in space is independently accessible.

The second assumption is that there is a high, perhaps unlimited, degree of flexibility in how the task can become compiled into a measurement on that database.

1.2 *Approach 2: Random access in spatial frequency, constrained access in space*

An alternative had previously been proposed that sought to explain the appearance of stimuli and to relate these early stages to visual computations more generally. The computational approach of Marr and Hildreth (1980) led to the proposition that only certain types of spatial configuration of filter response (zero-crossings) would be represented and therefore used at later stages of the visual system. The configurations were seen as edges and could be used, it was argued, as a sufficiently powerful representation of the grey-level image for most purposes.

Although this aspect was not well-developed in the theory, it seems clear that Marr and Hildreth envisaged unconstrained access to different spatial-frequency outputs. These could be used for mutual confirmation (zero-crossing alignment constraint), but for this to be possible they had to exist independently. Independent access to different spatial scales was also a feature of the second Marr-Poggio theory of stereo-correspondence (Marr and Poggio 1979) in which matches at a coarse scale guided matches at fine scales.

There have been at least two more recent variants on the same lines. The MIDAAS model of Kingdom and Moulden (1992) [recently extended to two dimensions by Moulden and MacArthur (in preparation)] is similar to Marr and Hildreth's model in that it starts by locating features in each spatial-frequency band. The key difference from MIRAGE is that the interpretation rules are applied to each channel before combination, and the results of the symbolic analysis are combined. In other words, if each channel finds an edge, in different places, there will be two edges in the final combination.

The local-energy model (Morrone and Burr 1990; Burr and Morrone 1994) employs a multi-scale analysis, and each scale separately convolves the image with matched pairs of odd- and even-symmetric filters. The output of these pairs is combined by Pythagorean sum, and peaks in the output are used to locate edges and bars, depending on the relative strengths of the odd- and even-symmetric filters.

1.3 *Approach 3: Constrained approach in spatial frequency and space*

A different approach to the issue of how the filter outputs are used by the visual system was proposed by us (Watt and Morgan 1985). It was already known that there were defects in the view that the different spatial frequencies might be detected independently (Henning et al 1975). We accumulated a range of data indicating that, in other respects, pattern at some spatial frequencies could not be processed independently of irrelevant pattern at other spatial frequencies (Morgan and Watt 1982; Watt and Morgan 1983, 1984, 1985).

To account for these data and others we constructed a model in which the outputs of different spatial-frequency filters were combined before they could be accessed for performing tasks. As will be described below, a key feature of the combination was an essential nonlinearity that was needed to preserve high spatial resolution.

A second distinguishing aspect of the MIRAGE model is the proposition we made that only certain spatial configurations were available for later stages of visual processing. In the same spirit as (and inspired by) the Marr-and-Hildreth approach, we proposed that zero-bounded regions of response were to be treated as indivisible spatial primitives. Thus, performance of visual tasks had to be based on certain specified simple descriptors of zero-bounded regions.

