CellPress
REVIEWS

Series: Seminal Neuroscience Papers 1978–2017

Science & Society

An Enduring Dialogue between Computational and Empirical Vision

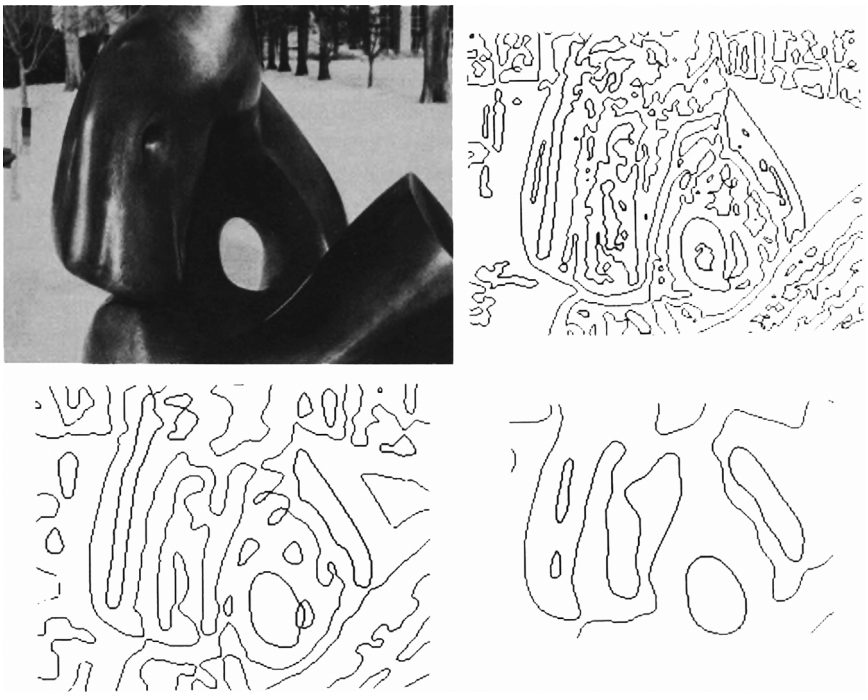
Susana Martinez-Conde,[1,*] Stephen L. Macknik,[1] and David J. Heeger[2]

In the late 1970s, key discoveries in neurophysiology, psychophysics, computer vision, and image processing had reached a tipping point that would shape visual science for decades to come. David Marr and Ellen Hildreth's 'Theory of edge detection', published in 1980, set out to integrate the newly available wealth of data from behavioral, physiological, and computational approaches in a unifying theory. Although their work had wide and enduring ramifications, their most important contribution may have been to consolidate the foundations of the ongoing dialogue between theoretical and empirical vision science.

The late 1970s were an exciting time for vision science. Key discoveries in neurophysiology, psychophysics, computer vision, and image processing had reached a tipping point that would shape the field for decades to come. What was lacking was a unifying theory to integrate the newly available wealth of data from behavioral, physiological, and computational approaches. David Marr and Ellen Hildreth's 'Theory of edge detection', published in 1980 [1], attempted to achieve just that.

Hildreth and Marr started working together in the fall of 1977. At that point Marr had already published his ideas on the primal sketch [2], which he envisioned as a rich, symbolic description of intensity changes in an image, computed in a bottom-up manner (E. Hildreth, personal communication). Around the same time, physiological and psychophysical studies had converged on vision's predisposition for edges and contours, as opposed to diffuse illumination [3,4]. A prevalent proposition was that the visual system was organized as a hierarchy in which neurons in sequential levels extracted increasingly complex image features in an iterative process. Marr had grown fascinated with Hubel and Wiesel's neurophysiological work in the primary visual cortex (L. Vaina, personal communication) as well as disappointed in the performance of ongoing edge detection methods on natural images, which he felt did not capture the known biology closely enough (E. Hildreth, personal communication).

Marr and Hildreth's paper contained two main insights. The first relates to intensity changes in an image across different spatial scales. Specifically, Marr and Hildreth noted that intensity changes in a natural scene occur over a wide range of scales and that such changes can be separately detected at different scales. In mathematical terms a useful operation to identify intensity changes at different scales in an image involves applying 'Gaussian filters' with different bandwidths to it (this is technically called 'convolving' the image with the filters). Marr and Hildreth proposed that wherever an intensity change occurs at a given scale, one will find a zero-crossing in the second derivate of the Gaussian filter (Figure 1). A second



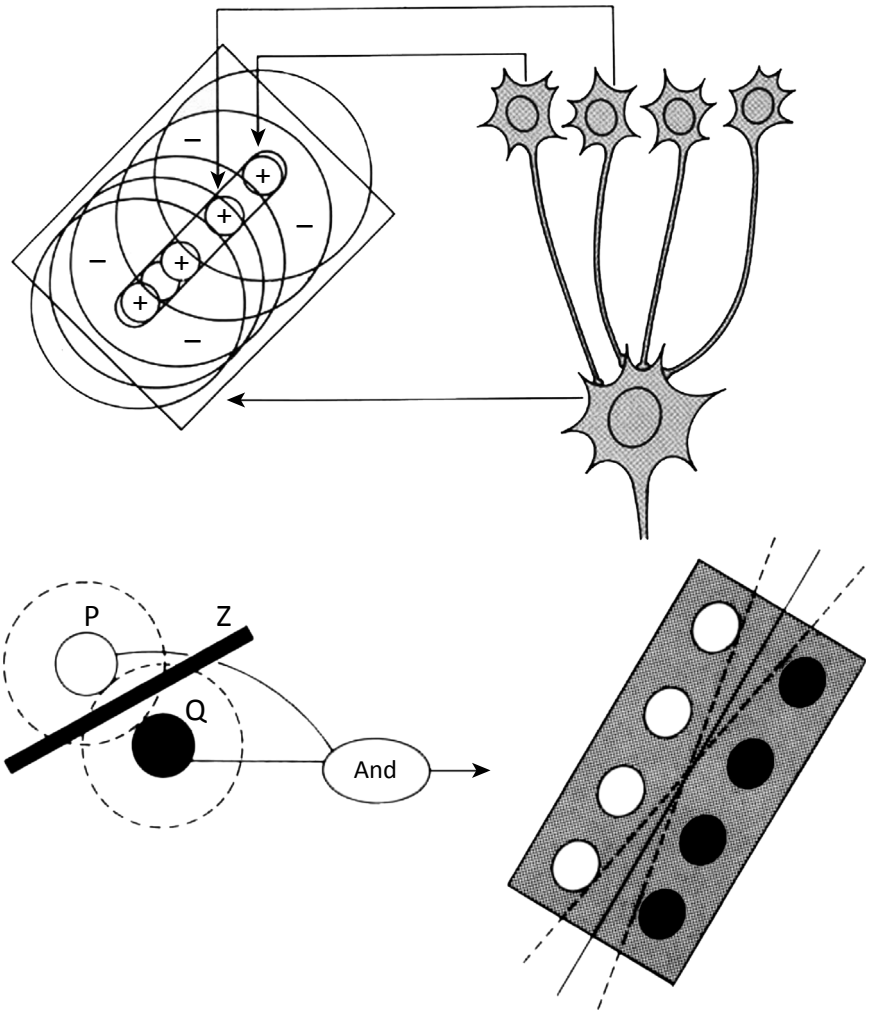
Trends in Neurosciences

Figure 1. Zero-Crossings Detected at Different Spatial Scales of the Same Image. The top-left panel depicts a 'natural' scene used as an example for image processing. The top-right panel and two bottom panels depict the zero-crossings that result from convolving the image with various Gaussian filters, approximating the range of filters that operate in the human fovea. Marr and Hildreth endeavored to combine all of the resulting sets of information into a single description. Reproduced from [1].

key insight in the paper is that zero-crossings at different scales coincide spatially and thus can be combined into primitive 'edge' elements. In other words, agreement across the zero-crossings of two or more independent channels (i.e., different spatial scales that are reasonably separated in the frequency domain) indicates that an edge was present in the image. Conversely, if zero-crossings do not agree across channels, this indicates that they are likely to have resulted from different physical phenomena in the image.

The publication attracted wide interest not just from within the computational field but also from the psychophysics and neurophysiology communities. One important reason for this broad impact was that Marr and Hildreth dedicated the extensive last section of their paper to discussing its 'implications for biology'. This section grounded their proposal in the state of the art of empirical knowledge, made specific predictions for psychophysics, and spelled out consequences for neurophysiology. Marr and Hildreth linked their computational theory to the biology by positing a mechanism by which (at least some) 'simple cells' could detect zero-crossing segments based on the output of earlier geniculate on-center and off-center receptive fields. In doing so Marr and Hildreth lent theoretical support to contemporary physiological models of vision that assigned the function of 'edge detector' to early visual neurons (Figure 2).

Marr and Hildreth's decision to rest their theoretical work on a three-legged stool of computational, physiological, and psychophysical evidence was immensely influential for subsequent vision research. More specifically, their use of multiscale bandpass analyses became a cornerstone of visual neuroscience and stimulated ensuing work on wavelets and multiresolution analysis [6–8]. Exploration of zero-crossings, and the search for their physiological counterparts, also spurred a widespread burst of research for many years following Marr and Hildreth's paper (further boosted by the posthumous publication of Marr's book *Vision* in 1982 [9]). Although zero-crossings have since fallen out of favor – along with the notion that the visual system reconstitutes an image from a strict hierarchy of component parts – Marr and Hildreth's insistence that theory be grounded in biology was principled, prescient, and powerful.



Trends in Neurosciences

Figure 2. Grounding Theory in Biology: Marr and Hildreth's Proposal for How Simple Cells May Function as Edge Detectors. Top: Hubel and Wiesel's simple-cell model. On the right, four lower-order neurons feed into a higher-order neuron. On the left, four center-surround antagonistic receptive fields (corresponding to the four lower-order neurons on the right) line up so that their centers lie along a line. The resulting simple-cell receptive field comprises an excitatory central region with two inhibitory flanks. Bottom: Marr and Hildreth's proposed mechanism by which simple cells detect zero-crossing segments. On the left, P and Q represent two center-surround receptive fields of opposite signs: on- center and off- center. On the right, several pairs of on- and off-center receptive fields are arranged in tandem and connected by 'logical AND gates'. The resulting simple-cell receptive field detects a zero-crossing segment within the orientation bounds approximately defined by the dotted lines. Top panels reproduced from [5], with kind permission from Paul Hubel. Bottom panels reproduced from [1].

From a personal perspective, Marr and Hildreth's work influenced our own thinking about edges, corners, and terminators [10–12] as well as our studies on multiscale image representations [10] and texture perception and texture image synthesis [13]. Although it may not be possible to capture the full breadth of the impact and wider ramifications of Marr and Hildreth's work on the field at large, perhaps one of their most important contributions has been to help establish the firm foundations of the ongoing dialogue between theoretical and empirical vision science.

Acknowledgements

The authors thank Edward Adelson, Ruzena Bajcsy, Geoffrey Hinton, Yann LeCun, and Stephane Mallat for their valuable insight on the state of the art in vision science before Marr and Hildreth's paper and the impact of its publication on the field. We are much indebted to Ellen Hildreth and Lucia Vaina for their inside perspective and candid observations. Rosario Malpica and Daniel Cortes-Rastrollo provided administrative and technical support. This work was supported by an NSF Award (1734887) to S.M-C. and S.L.M.

[1]State University of New York, Downstate Medical Center, Brooklyn, NY 11203, USA
[2]New York University, Center for Neural Science, New York, NY 10003, USA

*Correspondence:
smart@neuralcorrelate.com (S. Martinez-Conde).

https://doi.org/10.1016/j.tins.2018.02.005

References

1. Marr, D. and Hildreth, E. (1980) Theory of edge detection. *Proc. Biol. Sci.* 207, 187–217
2. Marr, D. (1976) Early processing of visual information. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 275, 483–519
3. Hubel, D.H. and Wiesel, T.N. (1959) Receptive fields of single neurons in the cat's striate cortex. *J. Physiol.* 148, 574–591
4. Ratliff, F. and Hartline, H.K. (1959) The responses of *Limulus* optic nerve fibers to patterns of illumination on the receptor mosaic. *J. Gen. Physiol.* 42, 1241–1255
5. Hubel, D. (1988) *Eye, Brain, and Vision,* W.H. Freeman
6. Burt, P. and Adelson, E.H. (1983) The Laplacian pyramid as a compact image code. *IEEE Trans. Commun.* 31, 532–540
7. Mallat, S.G. (1989) A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Trans. Pattern Anal. Mach. Intell.* 11, 674–693
8. Simoncelli, E.P. *et al.* (1992) Shiftable multi-scale transforms. *IEEE Trans. Inf. Theory* 38, 587–607
9. Marr, D. (1982) *Vision,* W.H. Freeman
10. Troncoso, X.G. *et al.* (2005) Novel visual illusions related to Vasarely's 'nested squares' show that corner salience varies with corner angle. *Perception* 34, 409–420
11. Troncoso, X.G. *et al.* (2007) BOLD activation varies parametrically with corner angle throughout human retinotopic cortex. *Perception* 36, 808–820
12. Troncoso, X.G. *et al.* (2009) Corner salience varies linearly with corner angle during flicker-augmented contrast: a general principle of corner perception based on Vasarely's artworks. *Spat. Vis.* 22, 211–224
13. Heeger, D.J. and Bergen, J.R. (1995) Pyramid-based texture analysis/synthesis. In *Proceedings of the 22nd Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH 1995*, pp. 229–238, ACM

Series: Seminal Neuroscience Papers 1978–2017

Science & Society

Retinal Cell Fate Specification

Suqian Wu,[1,2,3] Kun-Che Chang,[1,3,*] and Jeffrey L. Goldberg[1]

How are different neural cell types generated from progenitor cells? In 1990, Turner *et al*. used new lineage tracing techniques to show that different cells in the mammalian retina share their progenitor origin. The findings established a key step toward our understanding of how multipotent progenitor cells give rise to complex circuitry in the retina.

Vision is one of our richest senses and the dedication of our central nervous system (CNS) to visual processing is astounding. In mammals, vision originates with sensory processing in the retina, which demonstrates remarkable evolutionary conservation. Shared molecular and cellular features in development and adulthood from rodents to humans have allowed many basic questions to be modeled across species. One of the key questions relating to the formation of the visual system is how are different retinal cell types generated during development? Three decades ago, there were generally two hypotheses regarding the origin of retinal cells. According to one hypothesis, different cell-type classes – for instance, retinal ganglion cells (RGCs), rods, and cones – originate each from separate precursor cells, with each precursor restricted to produce only one or a few cell types. According to the second hypothesis, different retinal cell types could be differentiated from the same precursor cell. The question at the core of these competing hypotheses is relevant not only for retinal development, but also for other parts of the nervous system, and even other body organs. To track cell lineages, scientists had come up with methods for labeling precursor cells, and then determining what cell types differentiated from labeled precursor or stem cells. In the 1980s, a retrovirus-mediated gene transfer technique was developed for labeling individual precursor cell in the vertebrate CNS by expression of β-galactosidase and tissue staining [1,2]. In 1987, Turner and Cepko [1] reported a study in a postnatal rat retina showing that retrovirus-marked progeny clones differentiated into four different cell types with various overlapping combinations. The study represented an important milestone in clarifying retinal cell lineages. However, since only four of the seven major retinal cell-type classes were labeled, the study did not fully differentiate between the two hypotheses outlined above, and particularly when it comes to the earlier origin of retinal cell types, both theories remained plausible. Differentiating between the two accounts required taking the challenging step of going into earlier stages of development (i.e., the embryonic retina).

To label retinal precursor cells in the embryonic retina, Turner, Snyder, and Cepko [3] conducted *ex utero* surgery, which had not been applied much in