



# The structure of images: 1984–2021

Jan Koenderink<sup>1</sup>

Received: 26 February 2021 / Accepted: 13 March 2021 / Published online: 28 March 2021  
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2021

## Abstract

I present a personal account of the origin, development and future of a concept that appeared in this journal in 1984. The title was *The Structure of Images*. It became known as “scale space.”

**Keywords** Scale space · Image structure

*Reason why* The editor invited me to comment on a paper from the early 1980s (Koenderink 1984a). I checked Google Scholar <sup>1</sup> today (February 10, 2021): it had 3 618 citations since 1984. Current rate is  $\pm 50$ /annum. The paper was hardly cited before the 1990s. Citations peak about 2010. <sup>2</sup> I'll sketch why I think that is. I freely quote from my own past, because these references sketch the intellectual context.

## 1 Prehistory

*What sparked me off?* It was due to my interests in human awareness, neurophysiology, geometry, philosophy of mind and the visual arts. I've always been struck by the scale invariance of Leibniz Monads Leibniz (1991). As for the sciences, the well-known “Powers of Ten” movie (Eames and Eames 1977) fired my imagination and a remark by Friedrich

Nietzsche alerted me to Boscovich (1762). The problem of the continuum (through Franz Brentano (Brentano 1988; Koenderink et al. 2017a) <sup>3</sup>) kept me awake. I saw similarities between the neurophysiology of Lotze's (1884) *local sign* (Koenderink 1984b,c) (via a genial remark <sup>4</sup> by Helmholtz (1884)) and Čech cohomology (Čech 1932). In the visual arts I was fascinated by John Ruskin's “mystery” of distant details. He drew the first “scale space” I've ever seen (Ruskin 1857).

The academic problems from psychophysics required novel neurophysiological models. Such needs also arose in computer (image) science. At some point this sparked me off.

Note the serendipity. We only see the rivolets but are blind to the stream (science!).

*Mysterious data* During the 1970s and 1980s I worked on extensive perimetric studies of visual abilities such as spatiotemporal contrast luminance and hue detection, movement, etc. (Koenderink et al. 1978; van de Grind et al. 1983; van Esch et al. 1984). In retrospect this huge corpus was mostly ignored. My main satisfaction is that there are many “facts” in recent textbooks that I know to be mistaken. <sup>5</sup> At least I know.

<sup>1</sup> The numbers are volatile. I checked February 10, 2021. If you want to check yourself, try “koenderink scholar” on Google search and make sure you find Jan. *Femius* and *Gijsje* are my children.

<sup>2</sup> I'm well known for contributions that blossom late. A “Koenderink Prize” is awarded annually at the European Conference on Computer Vision for “a paper published ten years ago at that conference which has withstood the test of time.”

Communicated by Benjamin Lindner.

To highlight the scientific impact of our Journal over the last decades, we asked authors of highly influential papers to reflect on the history of their study, the long-term effect it had, and future perspectives of their research. We trust the reader will enjoy these first-person accounts of the history of big ideas in Biological Cybernetics.

✉ Jan Koenderink  
KoenderinkJan@gmail.com

<sup>1</sup> Emeritus Utrecht University, Utrecht, The Netherlands

<sup>3</sup> In comparing dates, please note that novel concepts brew for considerable periods (weeks to decades) before they appear in publications.

<sup>4</sup> “*Ferner ist sehr merkwürdig, dass bei Zahnschmerzen von Beinhautentzündung eines Zahns die Patienten im Anfang gewöhnlich unsicher sind, ob von einem Paar übereinander stehender Zähne der obere oder der untere leidet. ...Soll dies nicht davon herrühren, dass ...immer beide Zähne jedes Paares gleichzeitig starken Druck erleiden?*” This keen observation led to immediate enlightenment! I instantly saw how receptive field overlaps would embody a Čech cohomology. Perhaps unfortunately, I was unable to “sell” that to neurophysiologists.

<sup>5</sup> I sometimes wonder whether science invariably “progresses,” having seen so many (from my perspective) retrograde movements.

These data were *scale independent* (Koenderink and van Doorn 1978; Bijl et al. 1989). It was not predicted by physiological models. This led to self-similar models of the visual system that accounted for the bulk of the data (Koenderink and van Doorn 1982a). Although ignored, they survive in the scale space paradigm.

*Images as geometric data structures* In the early 1980s I was in the physics department of Utrecht University, the Netherlands. Forced to find funds elsewhere, I ended up doing odd-jobs for the American Bureau of Standards and the American Air Force.

The Air Force asked me to report on various laboratories all over the US. At Azriel Rosenfeld's lab (University of Maryland) I got a feeling of how important image science potentially was. That's when I started to think about image structure as an algorithmic problem.

In the years soon after my funds came from European ESPRIT projects. My academic interest rendered me a *Fremdkörper* in the computer science community. They deployed powerful Sun workstations, whereas I ran an Atari 1000 toy. However, I had Marty Veltman's (at my department) SCHOONSCHIP (Veltman and Williams 1993),<sup>6</sup> so I had some formal muscle.

About that time I—having met René Thom (1972) and following tutorials by Michael Berry (1992)—acquired an interest in catastrophe theory. I worked on singularities of optical projections (Koenderink and van Doorn 1979b, 1982b; Koenderink 1990a). This turned out to be crucial in the development of scale space theory.

*The 1984 paper* Sufficient motivation soon yields ideas. I was aware of the “pyramid” data structures (Burt and Adelson 1983; Crowley and Sanderson 1987) through Rosenfeld's lab and I fully understood the problem of “spurious resolution” Strasburger (2018) from my interest in photography and the visual arts.<sup>7</sup>

Hardly surprising that I hit upon the Gaussian kernel as special. Many others did too, like Andy Witkin (1983) whom I visited at Palo Alto in 1982. From my perspective, they had irrelevant reasons.

They failed to grasp the key concept. It is the diffusion equation  $\Delta\Phi = \Phi_t$ , where  $\Phi(x, y, t)$  is image intensity,  $\{x, y\}$  are Cartesian coordinates in the image plane, and  $t$  a scale parameter.<sup>8</sup>

<sup>6</sup> Veltman picked the Dutch word because he was sure no one (except a Dutchman) would be able to pronounce it. SCHOONSCHIP was the forerunner of Mathematica and Maple. It yields algebraic and logical “muscle.”

<sup>7</sup> Nowadays “spurious resolution” is known as “bokeh” (Nasse 2010). The photographic notion is complicated by the fact that one integrates over perspectives, not just images. This is only recorded in the literature around the 1900's.

<sup>8</sup> Much later (end of the 1990s) I was told (Weickert 1999) that I had done nothing novel, since Taizo Iijima apparently did it all in 1959. But

I'm often asked why I “buried” the 1984 paper in this journal. I knew its founder, professor Werner Reichardt (1924–1992), quite well and my interests fitted his journal more naturally than the “expected” computer science journals. But no doubt the bulk of the citations are from the latter field.

*The aftermath* The theory became known as “scale space theory.”<sup>9</sup> It is a standard tool. In medical image processing it is an indispensable part of diagnostic methods. This was important to me, as I was a professor in both physics and the medical faculty during the transition period from silver-based X-ray emulsions to electronic sensors and data storage.

Applications range from the microscopic scale (Midoh et al. 2007) to the cosmic (Schmalzing 1997).

Scale space became one of my “potboilers.” Apart from minor pulp science, there sprouted various diverging threads of academic pursuits (Sect. 3).

## 2 Formalization of the concept

For a classical physicist it is entirely obvious that diffusion cannot generate, but only destroy spatial articulations. It is easy enough to prove that from the diffusion equation (Sect. 1).

I proceeded to capture the essential concepts as a set of simple axioms from which the diffusion equation follows. This is desirable because the axioms can readily be applied in phenomenological models of psychogenesis, as well as models of neural receptive field structures. These topics were of greater interest to me than computer image processing.

The diffusion equation serves to connect scale levels. One may define a vector field whose streamlines capture such connections. It is like the pointers in discrete image pyramids. The streamlines let one track details over finite scale ranges.

The catastrophe notion is crucial. It lets one handle bifurcations of the streamlines. This captures the global causal structure.<sup>10</sup> It enables a discrete (topological) description of “deep” image structure (Koenderink and van Doorn 1986, 1987) and thus symbolical filtering.

Another crucial aspect is that the diffusion equation is a *linear* PDE. So scale space applies to arbitrary partial derivatives (Koenderink and van Doorn 1988). This suggested a principled taxonomy of receptive fields.

then, I did not read Japanese, so I can hardly be accused of plagiarism. In fact, nobody in the Western world was aware. Most people still aren't. Who cares today?

<sup>9</sup> A “scale space” Google search yields 1 370 000 000 hits.

<sup>10</sup> For about a decade there were endless discussions over whether the Gaussian Kernel is unique in establishing a causal structure, or even whether it does at all. From my perspective nothing has changed on this topic since the 1984 paper.

One may follow the evolution of differential invariants over scale (Koenderink and Richards 1988; Koenderink 1993). Images are trivial fiber bundles (Koenderink and van Doorn 2012). The important differential invariants are like geographical objects such as ruts, ridges, peaks, pits and passes (Koenderink and van Doorn 1979a, 1992, 1994). This yields topological, symbolical description. It suggested how cortical circuits might embody differential geometry and calculus.<sup>11</sup>

Most of this is discussed in the 1984 paper. In the remainder of the 1980s and most of the ‘90’s these structures were studied in detail and were generalized in various directions. Some way beyond my horizon.

Certain developments went “too far” from my perspective. If one prunes the development tree accordingly, “scale space theory” proper already reached its mature form in the ‘90s.

Not that developments beyond scale space proper are not interesting. Some (like Perona and Malik 1990) are elegant and useful. But in retrospect one now has a fairly complete overview of the lay of the land.

### 3 Diverging paths

Various people formulated more refined and mathematically elegant accounts than I managed in 1984 (Florack 1997; Griffin 2019). Others wrote text books (Lindeberg 1994; Haar Romeny 2003) that have been instrumental in the acceptance of scale space methods.

This is not a review. I am an outsider today. Nobody should feel offended when I ignore a favorite.

An alternative route to scale space is by way of the Hermite transform. This leads to a valuable extension of the formalism (Martens 2006).

Of course, the theory has been applied to various finite dimensions. Non-trivial are extensions to the temporal domain, because of its causal structure (Koenderink 1988; Lindeberg 2013).

An obvious extension is to consider local histograms instead of mere image intensities. The histograms can be taken over regions with a diameter given by the scale. One has a *local disarray* instead of *blurring* (Koenderink and van Doorn 1999, 2000; Koenderink et al. 2012). This opens up novel perspectives. One application is the perturbation of images for vision research (Koenderink et al. 2017b) and artistic purposes.

There are endless ways to tune or generalize the axioms, to think of special cases that would imply different families of kernels, to consider various ways to complicate the simple

diffusion equation, to apply the formalism to other domains, and so forth. All this—if possible more!—has been done. Perhaps it will be remembered as a cottage industry in theoretical image science for about three decades.

More relevant are implementations of scale space in discrete image processing algorithms. Not a simple matter, but crucial in applications (Lindeberg 1994; Haar Romeny 2003).

The brunt of the task has been completed.

### 4 The future

Scale space<sup>12</sup> is a tool, like the carpenter’s hammer and nails. It is hidden in software packages. Conceptual developments will be breakthroughs, because unexpected. But who knows?

I am interested in psychogenesis, Gestalt creation (Koenderink 2011, 2015; Koenderink et al. 2017c) and models of computational brain structures (Koenderink and van Doorn 1990; Koenderink 1990b; Koenderink et al. 2016, 2017a; Koenderink and van Doorn 2018). That was my main drive in ‘84.

A humbling fact that progress has been less than spectacular.

**Acknowledgements** The 1984 paper was not funded (the reference to ZWO was “political”).

I gratefully thank my spouse and long collaborator Andrea van Doorn, as well as all my students, my colleagues and guests who collaborated on the project at some time in history.

### Declarations

**Conflict of Interest** The author declares that he has no conflict of interest.

### References

- Berry MV (1992) Catastrophe Theory, vol 2. Encyclopaedia Britannica Inc., Chicago, p 948
- Bijl P, Koenderink J, Toet A (1989) Visibility of blobs with A Gaussian Luminance Profile. *Vis Res* 29(4):447–456
- Boscovich PRJ (1762) *Theoria Philosophiæ Naturalis*. Typographia Remondiniana, Venetiis
- Brentano F (1988) Philosophical investigations on space, time and the continuum (Smith, B. transl). Croom Helm, London
- Burt P, Adelson E (1983) The Laplacian Pyramid as a compact image code. *IEEE Trans Commun COM* 31:532–540

<sup>11</sup> Why are all cortices so similar? I speculate that it might be for the same reason that all physics texts look the same to colleagues from the humanities. Neurophysiologists disagree.

<sup>12</sup> Using Google Books NGram Viewer (<https://books.google.com/ngrams>) for “scale space, period 1980–2010, smoothing of 2,” I see a zero count at 1984, then a roughly linear increase to a flat level from 1992 to 2005, after that a steady decrease to one half of the top level. This seems to corroborate my estimate. Scale space’s heydays are over.

- Čech E (1932) Théorie générale de l'homologie dans un espace quelconque. *Fundamenta Math* 19:149–183
- Crowley JL, Sanderson AC (1987) Multiple resolution representation and probabilistic matching of 2-D gray-scale shape. *IEEE Trans Patt Anal Mach Intell* 9(1):113–121
- Eames C, Eames R (directors) (1977 final release date) Powers of Ten. Pyramid Films (final)
- van Esch JA, Koldenhof EE, van Doorn AJ, Koenderink JJ (1984) Spectral sensitivity and wavelength discrimination of the human peripheral visual field. *J Opt Soc Am A* 1(5):443–450
- Florack L (1997) *Image Struct.* Kluwer Academic Publishers, Dordrecht
- Griffin LD (2019) The atlas structure of images. *IEEE Tran Patt Anal Mach Intell* 41(1):234–245
- van de Grind WA, van Doorn AJ, Koenderink JJ (1983) Detection of coherent movement in peripherally viewed random-dot patterns. *J Opt Soc Am* 73(12):1674–1683
- Helmholtz H von (1884) Ueber die Localisation der Empfindungen innerer Organe. In: *Vorträge und Reden (Zweiter Band)*. Verlag von Friedrich Vieweg und Sohn, Braunschweig, p. 252
- Koenderink JJ, Bouman MA, Bueno de Mesquita AE, Slappendel S (1978) Perimetry of contrast detection thresholds of moving spatial sine wave patterns. I–IV. *J Opt Soc Am* 68(6):845–849, 850–854, 854–860, 860–865
- Koenderink J, van Doorn A (1978) Visual detection of spatial contrast; influence of location in the visual field, target extent and illuminance level. *Biol Cybernet* 30:157–167
- Koenderink J, van Doorn A (1979a) The structure of two-dimensional scalar fields with applications to vision. *Biol Cybernet* 33:151–158
- Koenderink J, van Doorn A (1979b) The internal representation of solid shape with respect to vision. *Biol Cybernet* 32:211–216
- Koenderink J, van Doorn A (1982a) Invariant features of contrast detection: an explanation in terms of self-similar detector arrays. *J Opt Soc Am* 72(1):83–87
- Koenderink J, van Doorn A (1982b) The shape of smooth objects and the way contours end. *Perception* 11:129–137
- Koenderink J (1984a) The structure of images. *Biol Cybernet* 50:363–370
- Koenderink J (1984b) Simultaneous order in nervous nets from a functional standpoint. *Biol Cybernet* 50:35–41
- Koenderink J (1984c) Geometrical structures determined by the functional order in nervous nets. *Biol Cybernet* 50:43–50
- Koenderink J, van Doorn A (1986) Dynamic Shape. *Biol Cybernet* 53:383–396
- Koenderink J, van Doorn A (1987) Representation of local geometry in the visual system. *Biol Cybernet* 55:367–375
- Koenderink J, van Doorn A (1988) Operational significance of receptive field assemblies. *Biol Cybernet* 58:163–171
- Koenderink J (1988) Scale-Time. *Biol Cybernet* 58:159–162
- Koenderink J, Richards W (1988) Two-dimensional curvature operators. *J Opt Soc Am A* 5(7):1136–1141
- Koenderink J, van Doorn A (1990) Receptive Field Families. *Biol Cybernet* 63:291–297
- Koenderink J (1990a) *Solid Shape*. The MIT Press, Cambridge
- Koenderink J (1990b) The brain a geometry engine. *Psychol Res* 52:122–127
- Koenderink J, van Doorn A (1992) Receptive field assembly pattern specificity. *J Vis Commun Image Present* 3(1):1–12
- Koenderink J (1993) What is a “Feature”? *J Intell Syst* 3(1):49–82
- Koenderink J, van Doorn A (1994) Two-plus-one-dimensional differential geometry. *Patt Recognit Lett* 15:439–443
- Koenderink J, van Doorn A (1999) The structure of locally orderless images. *Int J Comp Vis* 31(2/3):159–168
- Koenderink J, van Doorn A (2000) Blur and Disorder. *J Vis Commun Image Represent* 11:237–244
- Koenderink J (2011) Gestalts and Pictorial Worlds. *Gestalt Theory* 33(3/4):289–324
- Koenderink J, van Doorn A (2012) Gauge Fields in Pictorial Space. *SIAM J IMAG SCI* 5(4):1213–1233
- Koenderink J, Richards W, van Doorn A (2012) Space-time disarray and visual awareness. *i-Perception* 3:159–165
- Koenderink J (2015) Ontology of the Mirror World. *Gestalt Theory* 37(2):119–140
- Koenderink J (2016) The • of awareness. *Perception* 45(9):969–972
- Koenderink J, van Doorn A, Pinna B, Wagemans J (2016) Boundaries, Transitions and Passages. *Art Percep* 4:185–204
- Koenderink J, van Doorn A, Pinna B (2017a). Plerosis and Atomic Gestalts. *Gestalt Theory*. 39(1):30–53
- Koenderink J, Valsecchi M, van Doorn A, Wagemans J, Gegenfurtner K (2017b) Eidolons: Novel stimuli for vision research. *J Vis* 17(2):7. <https://doi.org/10.1167/17.2.7>
- Koenderink J, Doorn A van, Valsecchi M, Wagemans J, Gegenfurtner K (2017c) Eidolons & Capricious Local Sign. *IS&T International Symposium on Electronic Imaging 2017: Human Vision and Electronic Imaging 2017*, pp. 24–35
- Koenderink J, van Doorn A (2018) Local image structure and procrustes metrics. *SIAM J IMAG SCI* 11(1):293–324
- Leibniz GW (1991) (orig. 1714) *La Monadologie*, Édition établie par E. Boutroux, Paris LGF
- Lindeberg T (1994) *Scale space Theory in Computer Vision*. The Springer International Series in Engineering and Computer Science. Kluwer Academic Publishers, Dordrecht
- Lindeberg T (2013) A computational theory of visual receptive fields. *Biol Cybernet* 107:589–635
- Lotze H (1884) *Mikrokosmos*. Hirzel, Leipzig
- Martens J-B (2006) The Hermite Transform: A Survey. *eurasip J Appl Sig Process*, Vol. 2006. Article ID 26145:1–20
- Midoh Y, Nakamae K, Fujioka H (2007) Object size measurement method from noisy SEM images by utilizing scale space. *Measure Sci Technol* 18:579–591
- Nasse HH (2010) Depth of Field and Bokeh. Carl Zeiss Camera Lens Division, <https://lenspire.zeiss.com/photo/app/uploads/2018/04/Article-Bokeh-2010-EN.pdf> (last checked February 14, 2021), 45 pages
- Perona P, Malik J (1990) Scale space and edge detection using anisotropic diffusion. *IEEE Trans Patt Anal Mach Intell* 12(7):629–639
- Haar Romeny BM ter (2003) *Front-End Vision and Multi-Scale Image Analysis: Multi-scale Computer Vision Theory and Applications*, written in Mathematica. Springer Science & Business Media, ISBN: 978-1-4020-8840-7
- Ruskin J (1857) *The Elements of Drawing*, in *Three Letters to Beginners*. Smith, Elder, & Co., London Figures 21, 22
- Schmalzing J (1997) In: Bender R, Buchert T, Schneider P (eds) *Proc. 2nd SFB workshop on Astro-particle physics*, Ringberg 1996. Report SFB 375/P002 (1997) Koenderink Filters and the Microwave Background
- Strasburger H (2018) Blur Unblurred—A Mini Tutorial. *i-Perception* 9(2), <https://doi.org/10.1177/2041669518765850>
- Thom R (1972) *Stabilité Structurale et Morphogénèse: Essai d'une Théorie Générale des Modèles*. WA Benjamin, Reading
- Veltman MJG, Williams DN (1993, SCHOONSCHIP was developed in 1963) Schoonschip '91. [arXiv:hep-ph/9306228](https://arxiv.org/abs/hep-ph/9306228)
- Weickert J (1999) Linear scale space has first been proposed in Japan. *J Math Imag Vision* 10(3):237–252
- Witkin AP (1983) Scale space filtering. In: *Proc. 8th Int. Joint Conf. Art. Intell.*, Karlsruhe, Germany, pp. 1019–1022