

Luc Florack

*Technische Universiteit Eindhoven
Faculteit Biomedische Technologie
Postbus 513, 5600 MB Eindhoven
L.M.J.Florack@tue.nl*

Overzichtsartikel

Visuele perceptie en digi

Hoe kijkt de mens? Kan een computer interpreteren wat door een camera wordt waargenomen? Kan perceptie op een wiskundige manier worden gemodelleerd? Luc Florack, sinds 2000 universitair docent aan de Faculteit der Biomedische Technologie van de TUE, brengt in dit overzichtsartikel deze problemen met elkaar in verband.

Beeldverwerking — het bewerken of analyseren van beelden met behulp van de computer — is ontstaan uit de behoefte om de interpretatie van digitale beelden te vereenvoudigen met alle mogelijkheden die hedendaagse computers ons bieden. Haar ultieme raison d'être is het feit dat de mens een van nature visueel ingesteld dier is.

Een omvangrijk toepassingsgebied is medische beeldverwerking ter ondersteuning van diagnostiek en therapie. Daarnaast bestaat er een groot aantal terreinen waarin beeldverwerking een toenemende rol van betekenis krijgt, zoals seismiek, satellietwaarneming en robotiek. In al deze gevallen tracht men, gegeven een specifieke taakstelling, tot een interpretatie van een beeld te komen. Om sturing te kunnen geven aan ons gedrag heeft ons visuele systeem eveneens een interpretatie van de optische omgeving nodig zoals die geprojecteerd wordt op ons netvlies.

Voor een doeltreffende interpretatie van een beeld is een ope-

rationele representatie van structuur — los van de betekenis die daar à posteriori aan wordt toegekend — een noodzakelijke voorwaarde. Aan een ongestructureerde datastroom kan immers nooit zinnige betekenis worden ontleend. Een goede structurrepresentatie is een afspiegeling van relevante, objectieve feiten op grond waarvan specifieke (subjectieve) interpretaties tot stand kunnen komen. Anders gezegd, interpretatie moet niet op hallucinaties stoelen en met het beschikbare bewijsmateriaal moet je niet knoeien. Maar hoe kan men dit alles wiskundig formaliseren? Daartoe is inzicht in de fysica van optische waarneming onontbeerlijk. Perceptieonderzoek biedt eveneens aanknopingspunten.

Het visuele systeem

In het visuele systeem blijkt een zekere scheiding te bestaan tussen structurrepresentatie (opslag) enerzijds en interpretatie (perceptie) anderzijds, ofschoon elke poging tot exacte duiding hiervan enigszins arbitrair is. Je zou kunnen stellen dat de retina (oftewel het netvlies), de LGN (laterale geniculate nucleus, een onderdeel van de thalamus) en een gebiedje in de achterste hersenschors aangeduid als V1, een afspiegeling vormen van de optische omgeving in een universeel formaat geschikt voor uitlezing door tal van perceptiemodulen. Als zodanig is dit zogenaamde 'front-

tale beeldverwerking

end' feitelijk blind; van perceptie is hier nog geen sprake.

Het is instructief om naar een aantal electrofysiologische bevindingen bij zoogdieren te kijken. Met name bij apen en katten is er veel gemeten aan de respons van visuele cellen. De meetkarakteristieken hiervan worden 'receptieve velden' genoemd.

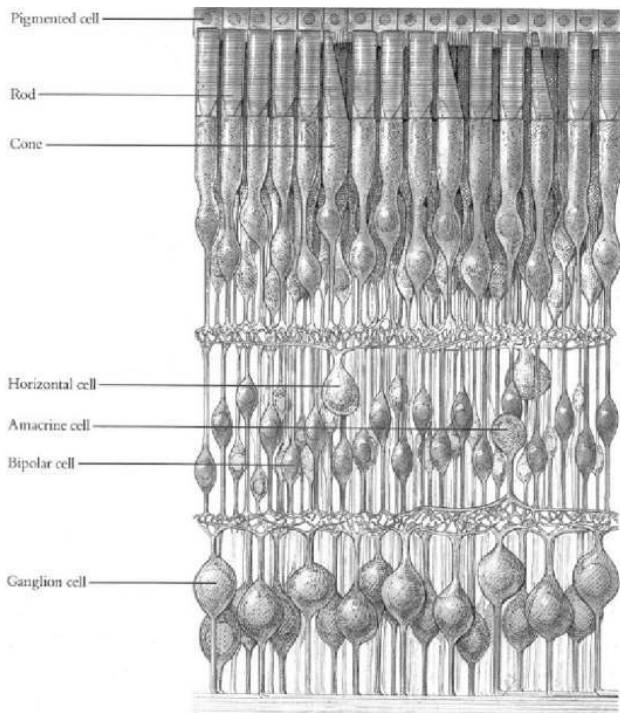
Receptieve velden ontstaan doordat het visuele systeem een aantal naburige fotoreceptoren (lichtgevoelige cellen in het netvlies) mobiliseert tot een legertje dat zijn collectieve bijdrage doorgeeft aan één enkele ganglioncel. Elke receptor legt daarbij zijn eigen gewicht in de schaal. De gewichtsfactor hangt af van de relatieve positie van de receptor en is bovendien gemoduleerd in de tijd. Zodoende ontstaat er een zeker spatiotemporeel weegprofiel. De signalen van de afzonderlijke fotoreceptoren, zeg maar de 'pixels' van het netvliesbeeld, worden dus *als zodanig* genegeerd. Eén enkele receptor kan bijdragen aan een veelheid van over lappende receptieve velden. Omgekeerd ontvangt een receptief veld van één enkele ganglioncel ingangssignalen van (meestal) een groot aantal receptoren. Een ruwe schatting van de gemiddelde convergentiefactor in het menselijke visuele systeem kan afgeleid worden uit het empirische feit dat het netvlies zo'n honderd miljoen receptoren (kegeltjes en staafjes) en ongeveer één

miljoen ganglioncellen telt. Zie figuur 1.

De uitgangssignalen worden via de ganglionaxonen, samengebundeld in de optische zenuw, doorgesluisd naar visuele cellen elders in de hersenschors. Door soortgelijke spatiotemporele weging als bij de retinale ganglionen vindt hier wederom formatie van karakteristieke (meer complexe) receptieve velden plaats. Een fraai geïllustreerd boek over de architectuur van de eerste stadia van het visuele systeem is Rodieck [1].

Men kan de weegprofielen van receptieve velden in principe in kaart brengen door in een geschikte visuele zenuwcel te prikken en haar axonsignaal te bestuderen als functie van de positie van een klein lichtstipje op het netvlies, zie figuur 3. Dit is de welbekende methode van 'reverse engineering' met behulp van een glijdende ' δ -functie', gedefinieerd 'onder de integraal' door $\delta(z) = 0$ voor alle $z \neq 0$ en $\int_{\Omega} \delta(z) dz = 1$ voor elk open volumegebied Ω rond de oorsprong. Alleen als het lichtstipje, bij benadering voor te stellen als een δ -functie, zich in het receptieve veld van de

Der Verstand vermag nichts anzuschauen, und die Sinne nichts zu denken.
Immanuel Kant



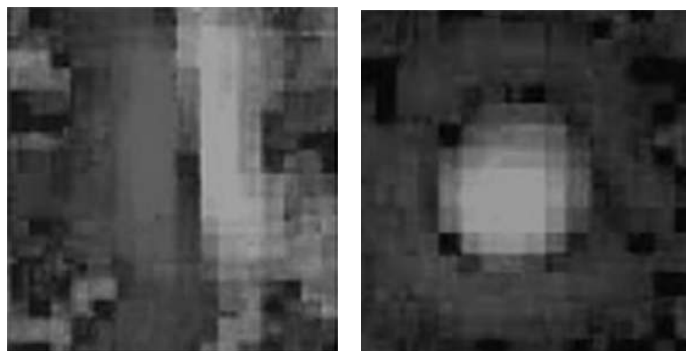
Figuur 1 Lichtgevoelige fotoreceptoren projecteren via diverse typen tussencellen op ganglioncellen. Licht valt, wellicht tegenintuïtief, binnen van beneden. Bron: [3]. From EYE, BRAIN, AND VISION by D.H. Hubel, copyright 1988, 1995 by Scientific American Library. Reprinted by permission of Henry Holt & Co., LLC.

aangepriekte cel bevindt vangt men een significant signaal op, waarvan de sterkte evenredig is met die van het plaatselijk weegprofiel:

$$\delta \star \phi(x) = \int_{\mathbb{R}^n} \delta(z - x) \phi(z) dz = \phi(x).$$

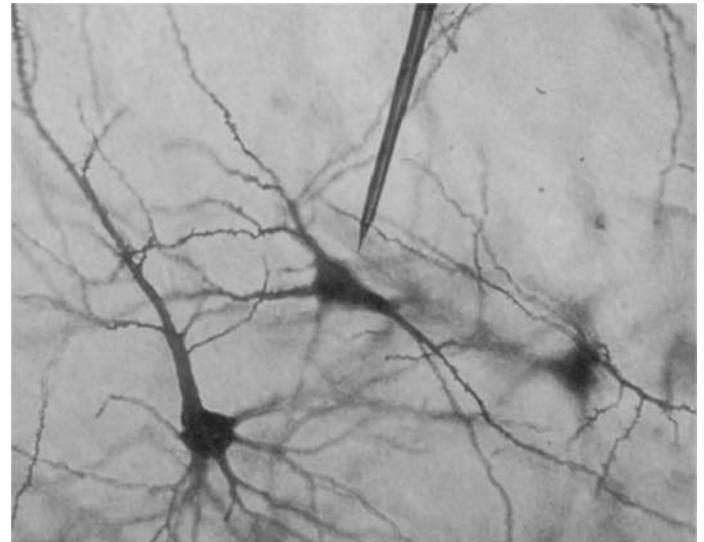
De infix operator \star staat voor correlatie. Neurofysiologisch gezien bepaalt de vuurfrequentie van een cel de sterkte van het signaal. Aangezien een cel in rust een zekere spontane activiteit heeft kan dit zowel negatief als positief zijn. Men spreekt dan van inhibitie of excitatie, als gevolg waarvan de vuurfrequentie lager dan wel hoger wordt dan de ‘nulactiviteit’. Door op deze manier te werk te gaan vindt men receptieve veldprofielen $\phi(x)$ zoals weergegeven in figuur 2. Veel receptieve velden in V1 blijken een zekere oriëntatievoorkeur te hebben, zie figuur 4.

Aanvankelijk werden receptieve velden opgevat als elementaire ‘vormdetectoren’ met namen als ‘edge’, ‘bar’, ‘centre-surround’, ‘grating detector’, enzovoort, al naar gelang de vorm



Figuur 2 Voorbeelden van gemeten receptieve velden. Links. ‘Edge detector’ cel. Georiënteerde profielen zoals deze komt men tegen in V1. Rechts. ‘Centre-surround’ cel. Dit is de typische structuur van een receptief veld van een ganglioncel. Bron: Ralph Freeman Visual Neuroscience Lab, University of California, Berkeley.

van hun profiel. Zo heeft een typische ganglioncel een ‘centre-surround’ structuur (zoiets als een Mexicaanse hoed) bestaande uit een exciterend centrum omgeven door een inhiherende ring, of andersom; de twee varianten worden ook wel aangeduid met ‘on-centre’ en ‘off-centre’ en blijken paarsgewijs gegroepeerd voor te komen. Een ‘edge detector’ is een receptief veld met een exciterend en inhiherend deel aan weerszijden van een lijn, de vermeende ‘edge’, enzovoort. De voorstelling van receptieve velden als elementaire vormdetectoren in deze naïeve gedaante is echter niet erg zinvol en zelfs misleidend [2].



Figuur 3 Met behulp van ‘single cell recording’ kunnen receptieve velden in kaart gebracht worden door te registreren hoe een cel reageert op een gecontroleerde netvliesstimulus. Bron: [3]. From EYE, BRAIN, AND VISION by D.H. Hubel, copyright 1988, 1995 by Scientific American Library. Reprinted by permission of Henry Holt & Co., LLC.

Gezien de overweldigende hoeveelheid ‘wetware’ die in het brein voor handen is zouden we à priori een enorme diversiteit aan receptieve velden verwachten. Afgezien van de grootte van hun spatiële drager — welke inderdaad grote variatie vertoont — is dit echter niet het geval voor wat betreft de vorm van de aange-troffen profielen. Deze blijkt in de praktijk allesbehalve willekeurig. De vraag doet zich voor welk dwingend evolutionair principe ten grondslag zou kunnen liggen aan de kennelijk door de natuur uitverkoren deelverzameling van mogelijke profielen. Om dat te begrijpen volstaat het niet om te meten, daartoe moeten we theoretiseren. Ik zal hieronder trachten om vanuit een aantal ogenschijnlijk verschillende invalshoeken te komen tot een verklarend beginsel.

Algebraïsch principe

Overall in de primaire visuele cortex treft men receptieve velden aan en toch blijkt het aantal wezenlijk verschillende profielen in de praktijk zeer beperkt, om de gedachten te bepalen in de orde van tientallen. In de context van een lineair c.q. gelineariseerd systeem suggereert dit een gesloten algebraïsche structuur, aangezien de neuronale overdracht van een signaal in zo’n geval beschreven kan worden als een keten van convoluties van corresponderende overdrachtsfuncties. Voor de overdracht tussen opeenvolgende receptieve velden ϕ en ψ hebben we, bij gegeven netvliesbeeld f :

$$(f \star \phi) \star \psi = f \star (\phi \star \psi).$$

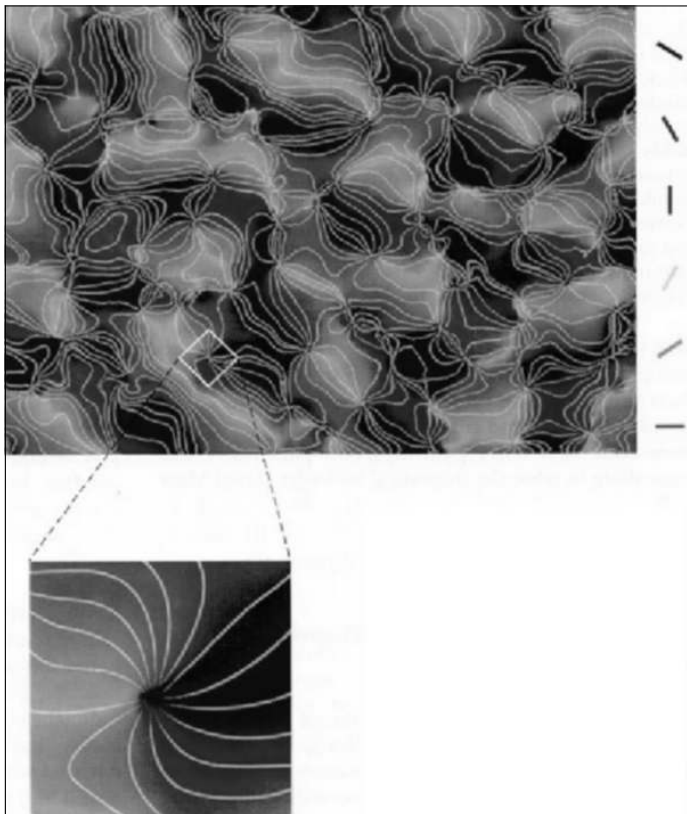
De infix operatoren \star en \ast symboliseren correlatie (lineaire be-monstering), respectievelijk convolutie:

$$f \star \phi(x) = \int_{\mathbf{R}^n} f(z - x) \phi(z) dz,$$

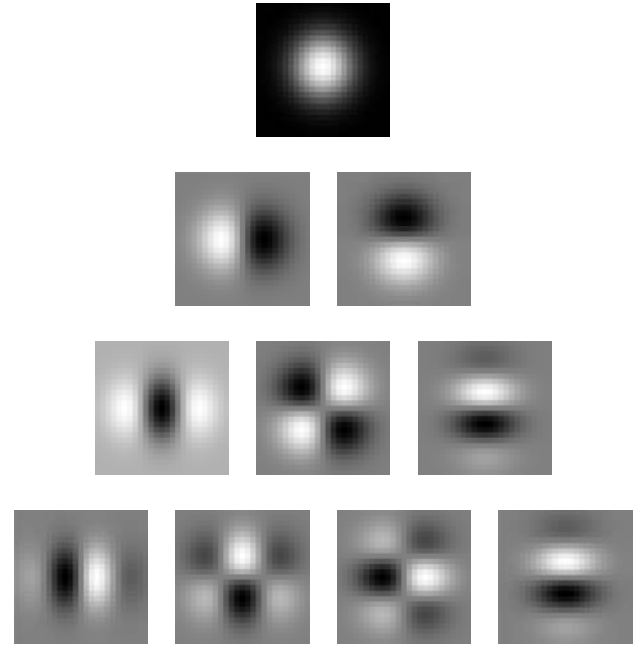
$$\phi \ast \psi(x) = \int_{\mathbf{R}^n} \phi(x - z) \psi(z) dz.$$

Aangezien het effect van twee opeenvolgende correlaties met receptieve velden ϕ en ψ kennelijk equivalent is met dat geïnduceerd door een enkel, effectief receptief veld $\phi \ast \psi$, moet dit laatste, om redenen van consistentie met de electrofysiologische bevindingen, wederom een toelaatbaar profiel hebben. Dat wil zeggen, het moet passen binnen de beperkte klasse van profielen zoals we die in het visuele systeem daadwerkelijk aantreffen. Dit leidt tot een drietal voorwaarden: Ten eerste, als ϕ, ψ receptieve velden zijn, dan moet ook $\phi \ast \psi$ een receptief veld zijn. Ten tweede, de aldus ontstane convolutiealgebra moet ‘voldoende klein’ en, ten derde, fysiologisch plausibel zijn.

Een voor de empirische praktijk voldoende rijke klasse van functies die aan de eerste eis voldoet is de zogenaamde Schwartzruimte van gladde functies, welke ‘voldoende hard naar nul gaan op oneindig’ [5]. Deze functies zijn ‘vrijwel compact’ en er is er altijd wel een geschikt om bij goede benadering, dat wil zeggen binnen de meettolerantie, model te kunnen staan voor een gemeten receptief veld. Echter, juist omdat deze klasse zo buitengewoon



Figuur 4 Oriëntatiegevoelige cellen rond een vast gekozen punt in het visuele veld blijken keurig georganiseerd te zijn in zogenaamde corticale kolommen, waarvan dwarsdoorsneden een opvallende spaakwielstructuur vertonen (uitsnede). Cellen langs de spaken van één zo'n wiel hebben alle dezelfde oriëntatievoorkeur. Deze verloopt continu in transversale richting. Een corticale kolom is het neurofysiologisch analogon van een pixel, dat wil zeggen zij representeert één (kennelijk meerwaardig) beeldpunt van het visuele veld. Corticale kolommen voor naburige punten liggen op hun beurt netjes geordend langs het oppervlak van de visuele cortex (bovenste plaatje), een ontdekking waarvoor Hubel en Wiesel in 1981 de Nobelprijs ontvingen. Bron: [4].



Figuur 5 Schaalruimterepresentatie van een tweedimensionale magnetische resonantie opname van een brein. Schaal (resolutie) neemt van links naar rechts toe (af).

groot is voldoet zij niet aan de overige voorwaarden. Het ligt dus voor de hand dat we ons beperken tot subalgebra's. De kleinste denkbare wordt gegenereerd door een genormeerde Gaussische functie $\phi_{\Sigma}(x), x \in \mathbf{R}^n$, welke volledig bepaald is in termen van een positief definitieve symmetrische parametermatrix Σ :

$$\phi_{\Sigma}(x) = \frac{1}{\sqrt{2\pi}^n \sqrt{\det \Sigma}} \exp \left[-\frac{1}{2} x \cdot \Sigma^{-1} x \right].$$

Deze (positieve) functies genereren weliswaar een subalgebra die aan de tweede eis voldoet — immers $\phi_{\Sigma_1} \ast \phi_{\Sigma_2} = \phi_{\Sigma_1 + \Sigma_2}$ — ze stroken echter niet met niet-positieve cellen zoals die in figuur 2. We kunnen dit echter eenvoudig ondervangen door ook afgeleiden mee te nemen. Voor niet-negatieve geheeltallige multi-indices [6] α, β geldt dan:

$$\nabla_{\alpha} \phi_{\Sigma_1} \ast \nabla_{\beta} \phi_{\Sigma_2} = \nabla_{\alpha + \beta} \phi_{\Sigma_1 + \Sigma_2}.$$

Deze zogenaamde Gaussische familie [7] heeft alle gewenste eigenschappen en blijkt inderdaad goed te correleren met tal van gemeten receptieve velden in netvlies, LGN en V1 van apen en katten en wellicht ook de mens [8–9]. Figuur 5 toont een aantal Gaussische afgeleiden in $n = 2$ dimensies. Orde van differentiatie neemt toe van boven ($|\alpha| = 0$, dus $\alpha = (0, 0)$) naar beneden ($|\alpha| = 3$, dus $\alpha = (3, 0), (2, 1), (1, 2)$ of $(0, 3)$). Merk op dat de V1-cel van figuur 2 aardig overeenkomt met een eerste orde afgeleide, terwijl de ganglioncel eenvoudig verkregen kan worden door superpositie van de niet-gemengde tweede orde afgeleiden (de Laplaciaan).

Regularisatieprincipe

Een geheel andere invalshoek ontstaat door het visuele systeem te beschouwen in het licht van een regularisatieprincipe [10–11]. Aangezien de retinale stimulus f inherent ruisig en buitengewoon complex van structuur is, zou men kunnen opperen dat het visuele systeem hiervan een robuuste en vereenvoudigde afspiegeling tracht te construeren. Bruikbare informatie zit immers per



Figuur 6 Potentiële isofootvlakken door een Morse extremum. Dit laatste genereert volgens de impliciete functiestelling een niet-horizontale curve in de (x, t) -ruimte (t neemt toe naar boven). **Links.** Toelaatbaar. **Rechts.** Ontoelaatbaar.

definitie in die structuren die ondanks ruis toegankelijk zijn en bovendien heeft het visuele systeem nou eenmaal een beperkte verwerkingscapaciteit.

Tikhonov regularisatie gaat uit van het variationeel principe waarbij men door minimalisatie van een geschikt gekozen functionaal ('energie') een geregulariseerde functie g zoekt 'in de buurt van' de oorspronkelijke functie f . Laten we gemakshalve aannemen dat $f \in L^2(\mathbf{R}^n)$ en dat toelaatbare functies g tenminste glad zijn. Aangezien er geen reden is om ons te beperken tot eindige differentieerbaarheid (want welke orde zou om welke reden de voorkeur moeten genieten?) kiezen we een functionaal op $C^\infty(\mathbf{R}^n) \cap \mathbf{H}^\infty(\mathbf{R}^n)$ waarbij $\mathbf{H}^\infty(\mathbf{R}^n) = \bigcap_{k \in \mathbf{N}} \mathbf{H}^k(\mathbf{R}^n)$, de doorsnede van alle Sobolevruimtes van eindige orde. Sterker nog, als $\|g\|$ de standaard $L^2(\mathbf{R}^n)$ -norm aanduidt, beperken we ons tot die functies g waarvoor de volgende norm gedefinieerd is voor alle $t > 0$:

$$\|g\|_t^2 = \sum_{|\alpha| > 0} \frac{t^{|\alpha|}}{\alpha!} \|\nabla_\alpha g\|^2.$$

De parameter t is geïntroduceerd uit fysische overweging, opdat alle termen dimensioneel compatibel zijn [12]: \sqrt{t} is hier een natuurlijke lengte-eenheid.

Vervolgens postuleren we voor de kandidaatfuncties g een geschikte energiefunctie als een superpositie van 'potentiële' (f -afhankelijke) en 'kinetische' regularisatietermen:

$$E_t[g] = \|g - f\|^2 + \|g\|_t^2.$$

De rol van t blijkt onder andere uit inspectie van de limietgevallen $t \downarrow 0$ en $t \rightarrow \infty$. In het eerste geval verdwijnen de regularisatietermen en komt minimalisatie kennelijk overeen met L^2 -projectie op f : $g \sim f$. In het tweede geval zien we dat alleen $g \sim 0$ de energie binnen de perken houdt. Kennelijk is t , behalve een maat voor schaal, tevens een maat voor de regularisatiesterkte.

De algemene oplossing $u(t) = \arg \min_g E_t[g]$ volgt uit de Euler-Lagrange vergelijking en luidt: $u(t) = \exp(t\Delta) f$. Een equivalente formulering wordt verkregen door te differentiëren naar t :

$$\begin{cases} \frac{\partial u}{\partial t} = \Delta u, \\ \lim_{t \downarrow 0} u = f, \end{cases}$$

waarmee het duidelijk wordt dat de expliciete oplossing wederom niets anders is dan correlatie, $u = f \star \phi_t$, met een genormeerde Gaussische (Greense) functie:

$$\phi_t(x) = \frac{1}{\sqrt{4\pi t}^n} \exp\left[-\frac{\|x\|^2}{4t}\right].$$

Deze komt overeen met de eerder besproken variant ϕ_Σ voor het isotrope geval $\Sigma = 2tI$.

Generalisatieprincipe

Als laatste hypothese zouden we kunnen veronderstellen dat ons visuele systeem een soort atlas aanlegt van de optische wereld, waarin het naar believen kan bladeren tussen fjnschalige en grof-schalige 'kaarten' van onze omgeving, al naar gelang onze interesse uitgaat naar de details van een kleinschalig voorwerp dan wel naar de globale structuur van het gehele visuele veld. Omdat onze voorkeursschaal continu kan variëren moeten we hierbij eerder denken aan een continuümstructuur dan aan een atlas met discrete pagina's.

Het principe van een atlas is om voor elke gewenste schaal een overzichtelijke hoeveelheid informatie te presenteren binnen het kader van de bladspiegel. Dit houdt bijvoorbeeld in dat wanneer we willen uitzoomen, zeg van een overzichtskaart van Nederland naar een van heel Europa, we kleinschalige details zullen moeten weglaten, of liever gezegd, generaliseren (samenvatten tot globalere structuren). Zouden we met dezelfde resolutie blijven werken, dan zouden we een onwerkbaar hoeveelheid gegevens op één pagina krijgen. De intrinsieke schaal (inverse resolutie of 'korreligheid') moet evenredig zijn met de gewenste zoomfactor.

De vraag doet zich voor hoe we moeten generaliseren in geval van een (netvlies)beeld f . Een intuïtief noodzakelijke maar enigszins vage voorwaarde is dat er geen structuur mag ontstaan wanneer we resolutie verlagen (schaal vergroten). In één van zijn eerste artikelen over dit onderwerp stelt Koenderink de volgende operationele definitie voor [13]. Wederom gaan we uit van isotropie, hetgeen achteraf eenvoudig te generaliseren valt. Zij $t > 0$ een maat voor de intrinsieke schaal van onze optische 'kaart'. Bij afnemende resolutie, dus toenemende t , mogen isofoten — contouren van constante luminantie — wel verdwijnen, maar niet uit het niets ontstaan. Dit geschiedt uiteraard in extrema (maxima of minima), welke, als functie van t , paden in het (x, t) -domein definiëren. Lopend langs zo'n pad in positieve t -richting manifesteert een extremum zich hetzij als bron (creatiepunt) hetzij als put (annihilatiepunt) voor isofootcontouren — dit alles gaat goed zolang het extremum leeft en van het Morse-type is. Het laatste geval moet, aldus Koenderink, worden uitgesloten op grond van het atlasprincipe. Bijgevolg kunnen isofootvlakken in de (x, t) -ruimte zich alleen naar boven toe sluiten, zie figuur 6.

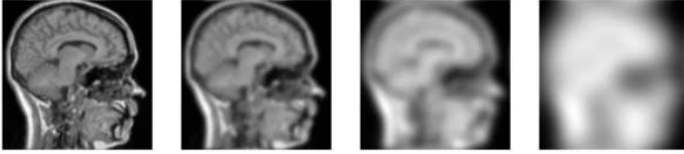
Zij $u(x, t)$ de representatie van het hoge resolutiebeeld $f(x)$ op schaal t . Op grond van bovengenoemde geometrische restrictie is eenvoudig af te leiden dat voor een willekeurig ruimtelijk extremum $x_e \in \mathbf{R}^n$ moet gelden dat

$$\frac{\partial u}{\partial t} \Delta u > 0,$$

aangenomen dat $u \in C_1^2(\mathbf{R}^n \times \mathbf{R}^+)$. De extra voorwaarde dat u lineair afhangt van het netvliesbeeld f levert, onder andere, de door eerder genoemde lineaire diffusievergelijking impliciet gedefinieerde oplossing [16]. Hiermee hebben we weer aansluiting gevonden bij voorgaande invalshoeken.

Biologische overwegingen

Er zijn enkele subtiliteiten die men niet over het hoofd mag zien wanneer men tracht een wiskundig model te relateren aan de werkelijkheid. In het bijzonder is het maar zeer de vraag of de para-



Figuur 7 Schaalruimterepresentatie van een tweedimensionale magnetische resonantie opname van een brein. Schaal (resolutie) neemt van links naar rechts toe (af).

metrisaties van de spatiële variabele x , de schaalparameter t en de intensiteitswaarde u , zoals die in bovenstaand model gehanteerd zijn, de biologisch relevante zijn en het is zelfs niet uitgesloten dat we ook de impliciete veronderstelling van een Euclidische ruimte moeten laten varen. In het algemeen zullen we dus een herparametrisatie moeten uitvoeren en eventueel onze metriek moeten aanpassen. In de diffusievergelijking is dit eenvoudig te verdisconteren door de isotrope Laplace operator te vervangen door een algemenere, zogenaamde Laplace-Beltrami operator op een Riemannse ruimte met metriek $\mathbf{G} = \sum_{\alpha,\beta} \mathbf{g}_{\alpha\beta}(\mathbf{x}) d\mathbf{x}^\alpha \wedge d\mathbf{x}^\beta$, waarvan de keuze der coëfficiënten aan de werkelijkheid moet worden getoetst. Voor de Euclidische ruimte vinden we dan de standaard diffusievergelijking terug. Intensiteit kunnen we herparametriseren middels een inverteerbare C^2 -functie γ , zeg $v = \gamma(u)$.

Het visuele systeem van de mens en andere zoogdieren wordt gekenmerkt door een fovea, een klein centraal gebiedje in het netvlies ten opzichte waarvan resolutie naar de periferie toe afneemt. Het menselijk netvlies is min of meer rotatiesymmetrisch en resolutie blijkt hier ongeveer omgekeerd evenredig te zijn met eccentriciteit. Dit kan redelijk goed gemodelleerd worden middels een conforme, rotatiesymmetrische metriek, die overal vlak is met uitzondering van het (singuliere) middelpunt [17]. Herparametrisatie van intensiteit leidt tot een speciale klasse van niet-lineaire diffusievergelijkingen voor *perceptuele intensiteit*, die zijn terug te voeren tot lineaire:

$$\begin{cases} \frac{\partial u}{\partial t} = \Delta u, \\ \lim_{t \downarrow 0} u = f, \end{cases}$$

$$\begin{cases} \frac{\partial u}{\partial t} = \sum_{\alpha,\beta} \left[\frac{1}{\sqrt{g}} \nabla_\alpha (\sqrt{g} g^{\alpha\beta} \nabla_\beta u) + \mu(u) g^{\alpha\beta} \nabla_\alpha u \nabla_\beta u \right], \\ \lim_{t \downarrow 0} u = f. \end{cases}$$

Hierin is $g = \det \mathbf{G}$. De niet-lineariteit wordt bepaald door de functie $\mu(u)$ die van de fotonoverdrachtsfunctie afhangt: Zij f de fysische fotonflux en $\gamma' > 0$, dan is $\mu = (\ln \gamma)'$. Door geschikte keuze van γ , dus μ , kunnen we ervoor zorgen dat aan fenomenologische wetmatigheden, zoals de Weber-Fechner wet [18], voldaan is [19], terwijl $v = \gamma(u)$ aan de lineaire diffusievergelijking met beginvoorwaarde $g = \gamma(f)$ voldoet. In de beeldverwerking ligt de keuze voor lineariteit en Euclidische metriek echter het meest voor de hand.

Schaalruimte

Naast genoemde zijn er nog andere axiomatische uitgangspunten die leiden tot eenzelfde (x, t) - of *schaalruimterepresentatie* van een ingangsbeeld als die welke natuurlijke evolutie van biologische systemen lijkt te hebben voortgebracht. Figuur 7 toont een viertal schaalniveaus van een magnetische resonantie opname, verkregen door correlaties van het oorspronkelijke hoge-resolutie beeld met een viertal isotrope Gaussische testfuncties.

Het schaalruimtemodel impliceert dat men receptieve velden

kan opvatten als differentiaaloperatoren van eindige schaal en orde, inclusief de niet-triviale orde nul, zoals in figuur 7. Immers, het netvliesbeeld f wordt, zoals bekend, niet als zodanig door het visuele systeem benaderd. In plaats daarvan ontvangt de visuele hersenschors (afgeleiden van) correlaties van dit ingangsbeeld met receptieve velden die kunnen worden gemodelleerd door afgeleiden van Gaussische functies van diverse schalen. Merk nu op dat

$$\nabla_\alpha (f \star \phi_\Sigma) = (-1)^{|\alpha|} f \star \nabla_\alpha \phi_\Sigma.$$

Dit suggereert dat het voorstadium van het visuele systeem als een differentiaalgeometrisch substraat fungeert [20]: Receptieve velden in de vorm van geschaalde differentiaaloperatoren $(-1)^{|\alpha|} \nabla_\alpha \phi_\Sigma$ (rechterlid) genereren een multiresolutie lokale jetbundel van de optische wereld (linkerlid, voor een veelheid van schalen Σ en met $|\alpha| \leq N$ voor zekere [21] N). Dit is het 'universele formaat' waar eerder over gerept werd en waaruit perceptie-modulen elders in de visuele hersenschors vrijelijk kunnen putten om van betekenisloze data een betekenisvol percept te maken. Hoe dit precies in zijn werk gaat blijft vooralsnog een mysterie. Niettemin stelt de theoretische ansatz in dit artikel de wiskundige als geen ander in staat tot gerichte speculatie. Bijvoorbeeld, het ligt voor de hand dat aan het 'uitlezen' van afzonderlijke receptieve velden van vaste schaal een invariantiemechanisme ten grondslag ligt dat signalen (in ons model immers partiële afgeleiden) combineert tot coördinaatonafhankelijke *invarianten*. De voornamelijk laat-negentiende eeuwse literatuur over dit onderwerp is in de beeldverwerking dan ook weer uiterst actueel. Een recent boek over klassieke invariantentheorie is dat van Olver [22].

Wat het schaalruimtemodel echter bijzonder intrigerend maakt is de niet-triviale rol van de schaalparameter(s), waarvoor in de klassieke wiskunde geen pasklare antwoorden zijn te vinden. Het visuele systeem lijkt zijn onovertroffen prestaties te ontleen aan de *simultane* verwerking van de optische wereld *over een geheel interval van schalen* [23].

Relevantie voor beeldverwerking

Heuristiek heeft sedert de beginjaren van het digitale tijdperk ook in de beeldwetenschap tot het besef geleid dat beelden hun informatie slechts prijsgeven bij de gratie van een correcte inschatting van het relevante schaaldomein. Er zijn in de loop der jaren tal van veelbelovende multiresolutiemethoden ontwikkeld voor grof-naar-fijn analyse van beelden. Dit lijkt inderdaad de enige zinvolle aanpak om uiterst complexe data te benaderen. Vrijwel alle voorgestelde algoritmen zijn echter in meer of mindere mate ad hoc en heuristisch van aard. Sinds kort begint hier echter verandering in te komen.

Wanneer je resolutie verlaagt verdwijnen er, per constructie, isofoten uit een beeld. Veranderingen in functiewaarden treden geleidelijk op, veranderingen in functietopologie (kritieke punten en isofoten) echter schoksgewijs. Zulke plotselinge topologische overgangen worden geassocieerd met *singulariteiten* of *catastrofen*. De bijbehorende schalen vormen het singulariteitspectrum. Het bijzondere van de situatie ligt besloten in het feit dat dit alles geschiedt binnen onze schaalruimterepresentatie, dus binnen de oplossingsruimte van de diffusievergelijking bij gegeven beginvoorwaarde.

Damon heeft aangetoond dat er in het isotrope geval slechts twee typen generieke singulariteiten kunnen optreden [24], ervan

uitgaand dat aan het ingansbeeld geen bijzondere eigenschappen toegedicht worden (ruis en meetonzekerheid zorgen ervoor dat dit laatste in de praktijk altijd het geval is; Damon beschouwt overigens ook ontappingsen die wiskundig gezien spannender zijn, maar met kans nul voorkomen). Een singulariteit behelst altijd een morsificatie van Morse kritieke punten in/vanuit een niet-Morse kritiek punt en verloopt altijd volgens het principe van 'ladingsbehoud'. De 'lading' van een Morse kritiek punt is hierbij gedefinieerd als het teken (± 1) van de determinant van de matrix van zijn locale tweede orde afgeleiden. De generieke singulariteiten van Damon vallen uiteen in annihilaties en (als $n > 1$) creaties van paren kritieke punten (in twee dimensies een extremum en een zadel) bij toenemende schaal. De laatste categorie lijkt op het eerste oog tegenintuïtief — er ontstaat in zekere zin structuur ondanks resolutieverlaging, maar is niet in strijd met het beginsel van Koenderink. Creaties doen zich in de praktijk echter wel minder vaak voor dan annihilaties en dat is ook begrijpelijk; als je resolutie maar voldoende verlaagt hou je uiteindelijk immers vrijwel geen structuur meer over. Statistische beschouwingen kunnen uitsluitel geven over de binaire kansverdeling op basis van de 'statistiek van natuurlijke beelden'.

Er is een stelling die zegt dat er bij compact gedragen, positieve, integreerbare functies f een schaal $T > 0$ bestaat zodanig dat er voor $t > T$ precies één kritiek pad overblijft, en wel een maximumpad [25]. De stelling biedt zelfs een scherpe afschatting voor T als functie van de straal R van het convexe omhulsel van de compacte drager van f , namelijk $2T = R^2$ en van het asymptotische gedrag van het onsterfelijke maximum als $t \rightarrow \infty$. Een andere stelling zegt dat kritieke punten niet kunnen ontsnappen uit het convexe omhulsel.

Ook de klassieke differentiaalmeetkunde biedt bruikbare invalshoeken voor de bestudering van schaalruimtereprésentaties. De doorgaans kwalitatieve uitspraken van singulariteitentheorie kunnen we complementeren met geometrische detaillering. Wederom uitgaande van een isotrope schaalruimte representatie $u(x, t)$ is het bijvoorbeeld niet zo moeilijk aan te tonen dat in $n + 1$ schaalruimtedimensies de (generieke) kritieke paden, inclusief de singuliere punten, beschreven kunnen worden door de volgende evolutievergelijking [26]:

$$\frac{d}{ds} (x(s), t(s)) = \left(-\tilde{\mathbf{H}} \nabla \text{tr} \mathbf{H}, \det \mathbf{H} \right),$$

waarin \mathbf{H} de matrix van tweede orde afgeleiden van $u(x, t)$ is, $\tilde{\mathbf{H}}$ zijn cofactor matrix (in het niet-ontaaarde geval hebben we $\tilde{\mathbf{H}} = \det \mathbf{H} \mathbf{H}^{\text{inv}}$) en $\text{tr} \mathbf{H}$ en $\det \mathbf{H}$ spoor respectievelijk determinant. Uit bovenstaande blijkt dat we als beginvoorwaarden de collectie van alle op $t = 0$ bestaande extrema mogen nemen opdat we via integratie alle kritieke paden zullen vinden die het beginvlak bereiken. Let op: er bestaan ook 'vacuümfluctuaties' in de vorm van gesloten lussen welke we op deze manier missen!

Combinatie van statistisch onderzoek — met vragen als: Hoe zijn kritieke punten en hun singulariteiten verdeeld in 'natuurlijke beelden'? —, singulariteitentheorie in de context van de diffusievergelijking, analytische afschattingen voor haar oplossing en differentiaalmeetkundige eigenschappen van kritieke paden en duale isofoten kan leiden tot nieuw inzicht in de 'diepe structuur' van een beeld, tot verifieerbare hypothesen voor de onderliggende mechanismen van ons visuele systeem en mogelijk zelfs tot gedurfde voorspellingen ten aanzien van de grotendeels nog

onbegrepen functionaliteit van tal van visuele stadia. Zo blijkt er bijvoorbeeld een intensieve, grotendeels onbegrepen feedback te zijn van V1 naar LGN. Er gaat méér dataverkeer van de visuele hersenschors richting netvlies dan andersom! De vraag doet zich voor wat hier wiskundig achter zou kunnen steken. Dit soort inzicht zal op zijn beurt nieuwe inspiratie verschaffen voor de ontwikkeling van grof-naar-fijn algoritmen voor beeldverwerking ter vervanging van bestaande heuristiek.

Literatuuroverzicht

Anders dan het geval is met toepassingen is er nog niet zoveel theorie over het schaalruimtebegrip. Een onmisbare bron is de oorspronkelijke publicatiereeks van de hand van Koenderink uit de jaren tachtig van de vorige eeuw [2, 7, 13, 27–33]. Deze lijst is verre van compleet. Een voor wiskundig georiënteerde lezers geschikte inleiding is te vinden in een recent boek [34]. Hierin wordt uitgebreider ingegaan op het dualiteitsbeginsel, op de constructie van differentiaalvarianten en op fundamentele ruimte-tijd aspecten, zoals temporele filters op de causale half-as en op bewegingsdetectie in tijdreeksen. In het boek van Lindeberg [35] ligt de nadruk op axiomatische onderbouwing in de context van discrete voorstellingen (schaal wordt evenwel continu verondersteld). Hierin wordt onder andere aangetoond dat het discrete analogon van de Gaussische autoconvolutiealgebra gegeven wordt door filters van de vorm $T(n, t) = \exp(-t) I_n(t)$, waarin I_n gemodificeerde Besselfuncties van geheeltallige orde zijn (dus géén gediscrètiseerde Gaussische functies). Voor schalen groter dan de roosterconstante verschillen deze overigens nauwelijks van hun continue tegenhangers. Lindeberg bespreekt verder een heuristisch grof-naar-fijn algoritme voor de detectie van 'blobs'. Daarnaast bestaat er een aantal summerschool- en conferentieproceedings met fundamentele en heuristische bijdragen die inspiratie kunnen verschaffen voor verdere wiskundige modellering [36, 38–40]. Tot slot is er een leerzaam boek van Ter Haar Romeny, welk kort na het schrijven van dit artikel ter perse zal gaan [37]. Dit is geheel geschreven in Mathematica [41] en voorzien van CDROM met notebook versie, zodat men naar believen kan experimenteren. Uiteraard is dit beknopte literatuuroverzicht incompleet en niet zonder persoonlijke bias.

Conclusie

Er bestaat een bijzondere verhouding tussen digitale beeldverwerking en perceptieonderzoek. De eerste poogt algoritmen op te stellen voor de extractie van informatie uit beelden, welke — als bijkomstigheid, want dit is slechts zelden het doel — een verklaring zouden kunnen bieden voor de manier waarop bepaalde visuele competenties in het brein gerealiseerd zijn. Omgekeerd kan men in de beeldverwerking lering trekken uit de bevindingen van perceptieonderzoek, welke de architectuur en functionaliteit van het door miljoenen jaren van evolutie geoptimaliseerde visuele systeem in kaart tracht te brengen.

Helaas vindt er nog relatief weinig kruisbestuiving plaats. Verschillen in doelstelling, wetenschappelijke cultuur en vakjargon liggen hier aan ten grondslag. Wiskunde zou hierin, als neutraal en onduubbelzinnig communicatiemiddel, een grotere rol moeten spelen. ◀

Noten en referenties

- 1 R.W. Rodieck, *The First Steps in Seeing*. Sunderland, Massachusetts: Sinauer Associates, Inc., 1998.
- 2 J.J. Koenderink, 'What is a "feature"?', *Journal of Intelligent Systems*, vol. 3, no. 1, pp. 49–82, 1993.
- 3 D.H. Hubel, *Eye, Brain and Vision*, vol. 22 of *Scientific American Library*. New York: Scientific American Press, 1988.
- 4 E.R. Kandel, J.H. Schwartz, and T.M. Jessell, *Principles of Neural Science*. McGraw-Hill, fourth ed., 2000.
- 5 L. Schwartz, *Théorie des Distributions*. Publications de l'Institut Mathématique de l'Université de Strasbourg, Paris: Hermann, second ed., 1966.
- 6 Een multi-index α is een n -tupel $(\alpha_1, \dots, \alpha_n)$ waarvoor een aantal conventies geldt, zoals $|\alpha| = \alpha_1 + \dots + \alpha_n$ (orde), $\alpha! = \alpha_1! \dots \alpha_n!$, $\nabla_\alpha = \frac{\partial^{\alpha_1 + \dots + \alpha_n}}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}}$, enzovoort. Meestal spreekt de notatie voor zich.
- 7 J.J. Koenderink and A. J. van Doorn, 'Receptive field families,' *Biological Cybernetics*, vol. 63, pp. 291–298, 1990.
- 8 G.C. DeAngelis, I. Ohzawa, and R. D. Freeman, 'Depth is encoded in the visual cortex by a specialised receptive field structure,' *Nature*, vol. 352, pp. 156–159, July 1991.
- 9 R.A. Young, 'The Gaussian derivative model for machine vision: Visual cortex simulation,' *Journal of the Optical Society of America*, July 1986.
- 10 A. Kirsch, *An Introduction to the Theory of Mathematical Inverse Problems*, vol. 120 of *Applied Mathematical Sciences*. New York: Springer-Verlag, 1996.
- 11 A. Tikhonov and V.Y. Arseninn, *Solution of Ill-Posed Problems*. New York: John Wiley & Sons, 1977.
- 12 In een spatiotemporeel model kan men ruimte en tijd formeel op gelijke voet behandelen door gebruik te maken van de conventie $x_0 = ct$ analoog aan die in de relativiteitstheorie; $c > 0$ is hier echter geen constante, maar relateert onafhankelijke lengte- en tijdschalen. Gemakshalve ga ik uit van isotropie oftewel rotatie-invariantie.
- 13 J.J. Koenderink, 'The structure of images,' *Biological Cybernetics*, vol. 50, pp. 363–370, 1984.
- 14 R. Duits, L. Florack, J. de Graaf, and B. ter Haar Romeny, 'On the axioms of scale-space theory.' In preparation.
- 15 K. Yosida, *Functional Analysis*. Berlin: Springer-Verlag, sixth ed., 1980.
- 16 Er zijn ook minder voor de hand liggende positiviteit behoudende lineaire p.d.v.'s die aan de eis voldoen, zoals $u_t = -(-\Delta)^\alpha u$ met $\frac{1}{2} < \alpha \leq 1$ [14–15]. Deze laat ik hier verder onbesproken.
- 17 L.M.J. Florack, 'A geometric model for cortical magnification,' in *Biologically Motivated Computer Vision: Proceedings of the First IEEE International Workshop, BMCV 2000, Seoul, Korea, May 2000*, S.-W. Lee, H.H. Bülthoff, and T. Poggio, eds., vol. 1811 of *Lecture Notes in Computer Science*, (Berlin), pp. 574–583, Springer-Verlag, May 2000.
- 18 De Weber-Fechner wet is een fenomenologische wet die stelt dat de kleinste waarneembare verandering in psychofysische respons δu evenredig is met de waarde van de fysische stimulus u , met andere woorden dat log u de psychofysisch relevante grootheid is. Uiteraard geldt dit slechts bij benadering binnen zekere drempel- en verzadigingswaarden.
- 19 L.M.J. Florack, R. Maas, and W. J. Niessen, 'Pseudo-linear scale-space theory,' *International Journal of Computer Vision*, vol. 31, pp. 247–259, April 1999.
- 20 J.J. Koenderink, 'The brain a geometry engine,' *Psychological Research*, vol. 52, pp. 122–127, 1990.
- 21 Er zijn aanwijzingen voor een goede overdekking van het netvlies voor orde $N = 4$ en incidentele rapportages van ordes tot en met $N \approx 12$, al is dit laatste nogal speculatief gezien de uiterst beperkte signaalruis verhouding van celmetingen.
- 22 P.J. Olver, *Classical Invariant Theory*, vol. 44 of *London Mathematical Society Student Texts*. Cambridge: Cambridge University Press, 1999.
- 23 Bij mijn weten geldt dit alleen voor spatiele schaal. Psychofysische experimenten lijken erop te duiden dat het menselijk visuele systeem een vaste tijdschaal van circa 50 milliseconden hanteert.
- 24 J. Damon, 'Local Morse theory for solutions to the heat equation and Gaussian blurring,' *Journal of Differential Equations*, vol. 115, pp. 368–401, January 1995.
- 25 M. Loog, J.J. Duistermaat, and L.M.J. Florack, 'On the behavior of spatial critical points under Gaussian blurring. a folklore theorem and scale-space constraints,' in Kerckhove [40], pp. 183–192.
- 26 L. Florack and A. Kuijper, 'The topological structure of scale-space images,' *Journal of Mathematical Imaging and Vision*, vol. 12, pp. 65–79, February 2000.
- 27 J.J. Koenderink and A.J. v. Doorn, 'Dynamic shape,' *Biological Cybernetics*, vol. 53, pp. 383–396, 1986.
- 28 J.J. Koenderink and A.J. v. Doorn, 'Representation of local geometry in the visual system,' *Biological Cybernetics*, vol. 55, pp. 367–375, 1987.
- 29 J.J. Koenderink and A.J. van Doorn, 'Operational significance of receptive field assemblies,' *Biological Cybernetics*, vol. 58, pp. 163–171, 1988.
- 30 J.J. Koenderink, 'Scale-time,' *Biological Cybernetics*, vol. 58, pp. 159–162, 1988.
- 31 J.J. Koenderink, 'A hitherto unnoticed singularity of scale-space,' *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, pp. 1222–1224, November 1989.
- 32 J.J. Koenderink and A. J. van Doorn, 'Receptive field assembly pattern specificity,' *Journal of Visual Communication and Image Representation*, vol. 3, no. 1, pp. 1–12, 1992.
- 33 J.J. Koenderink and A.J. v. Doorn, 'The structure of locally orderless images,' *International Journal of Computer Vision*, vol. 31, pp. 159–168, April 1999.
- 34 L.M.J. Florack, *Image Structure*, vol. 10 of *Computational Imaging and Vision Series*. Dordrecht, The Netherlands: Kluwer Academic Publishers, 1997.
- 35 T. Lindeberg, *Scale-Space Theory in Computer Vision*. The Kluwer International Series in Engineering and Computer Science, Dordrecht, The Netherlands: Kluwer Academic Publishers, 1994.
- 36 B.M. t. Haar Romeny, L.M.J. Florack, J.J. Koenderink, and M.A. Viergever, eds., *Scale-Space Theory in Computer Vision: Proceedings of the First International Conference, Scale-Space'97, Utrecht, The Netherlands*, vol. 1252 of *Lecture Notes in Computer Science*. Berlin: Springer-Verlag, July 1997.
- 37 B.M. t. Haar Romeny, 'Front-end vision.' Leerboek over schaalruimtemodellen voor digitale beeldverwerking en de eerste stadia van het visuele systeem, geschreven in Mathematica. In voorbereiding.
- 38 M. Nielsen, P. Johansen, O.F. Olsen, and J. Weickert, eds., *Scale-Space Theories in Computer Vision: Proceedings of the Second International Conference, Scale-Space'99, Corfu, Greece*, vol. 1682 of *Lecture Notes in Computer Science*. Berlin: Springer-Verlag, September 1999.
- 39 J. Sporring, M. Nielsen, L.M.J. Florack, and P. Johansen, eds., *Gaussian Scale-Space Theory*, vol. 8 of *Computational Imaging and Vision Series*. Dordrecht, The Netherlands: Kluwer Academic Publishers, 1997.
- 40 M. Kerckhove, ed., *Scale-Space and Morphology in Computer Vision: Proceedings of the Third International Conference, Scale-Space 2001, Vancouver, Canada*, vol. 2106 of *Lecture Notes in Computer Science*. Berlin: Springer-Verlag, July 2001.
- 41 S. Wolfram, *The Mathematica Book*. Wolfram Media/Cambridge University Press, fourth ed., 1999.