

Op het eerste gezicht lijkt een computerrepresentatie voor beelden die op pixels gebaseerd is de meest natuurlijke en eenvoudigste. Toch kleven er veel nadelen aan pixels. **Kees van Overveld** bekijkt op welke andere betekenisniveaus beelden gerepresenteerd zouden kunnen worden. Hij presenteert op NWD 10 een beschrijving van beelden met zogenaamde dynamische segmenten.

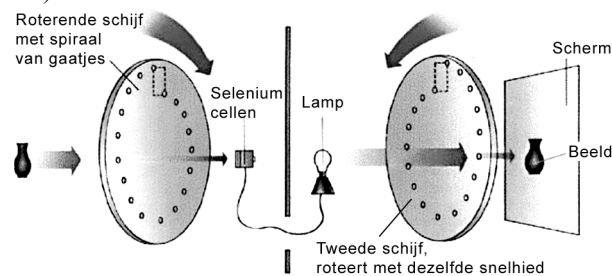
Ontsnappen aan de Nipkow-doctrine

Ooit is door mensenhand de allereerste afbeelding gemaakt. Dat was, met recht, een historisch moment. Het was het moment waarop de geschiedenis van de verbeelding een nieuwe wending nam. Vanaf dat ogenblik waren er niet alleen maar meer beken, bomen, bergen, buffels en Batavieren te zien, maar ook beelden! Uiteraard is de naam van die allereerste beeldende kunstenaar ons reeds lang ontgaan, en ook kunnen we nog slechts speculeren over de ontroering die die allereerste KUNSTmatige afbeelding teweeg moet hebben gebracht bij het allereerste kunstminnend publiek. Maar we hebben wel een redelijk idee wat die eerste afbeelding geweest zou kunnen zijn. Het was waarschijnlijk de afbeelding van een hand. Een modderige hand. Nauwkeuriger gezegd: de afbeelding bestond waarschijnlijk uit de modder op die hand, afgeveegd aan de binnenwand van een grot. Temidden van prehistorische grotschilderingen zijn dergelijke oeroude handafdrukken teruggevonden, en hun grote aantal is misschien een aanwijzing voor het enthousiasme waarmee deze revolutionaire afbeeldingstechniek ontvangen werd door de toenmalige beeldende kunstenaars.

Veel technieken volgden. De handafdrukken werden afbeeldingen van bison, oerossen en mammoeten, opgebouwd uit luttel strepen en vlekken van as, plantensap, of vochtige aarde. In duizenden jaren werden zowel de pigmenten meer geavanceerd, alsook de manieren om die pigmenten op een ondergrond aan te brengen.

Met de vervolmaking van de olieverftechniek in de zeventiende eeuw van onze jaartelling werd een voorlopig hoogtepunt bereikt, maar kunstenaars zochten verder naar nieuwe soorten van beeldelementen. Het laat negentiende-eeuwse pointillisme (letterlijk: 'stippelkunst') bijvoorbeeld, bracht een richting waarin voor het eerst met beeldelementen gewerkt werd die niet meer overeenkwamen met betekenisinhouden in het schilderij: de afgebeelde personen, voorwerpen of landschappen waren het resultaat van een optische zinsbegoocheling die patronen van gekleurde stippels ineen laat smelten tot betekenisvolle eenheden, waarbij het vrijwel onmogelijk wordt om de afzonderlijke stippels nog als zelfstandige, visuele entiteiten te zien.

Het is deze zelfde optische illusie die ten grondslag ligt aan het briljante idee van Paul Nipkow. Een ronddraaiende schijf met gaatjes gerangschikt langs een spiraal, maakt het mogelijk om de tweedimensionale helderheidsverdeling in een beeld om te zetten in een eendimensionale tijdsreeks van helderheidswaarden die, bij benadering¹, dezelfde informatie bevat. Zo'n tijdsreeks kan vervolgens middels een seleniumcel omgezet worden in een reeks elektrische spanningen die door een draad of zelfs draadloos overgebracht kunnen worden. Met behulp van een tweede schijf met gaatjes die aan de ontvangstkant synchronoos met de eerste schijf ronddraait voor een lamp, kan het oorspronkelijke beeld min of meer weer tevoorschijn geroepen worden. Als de lamp namelijk feller of zwakker gaat branden in overeenstemming met de door de seleniumcel ontvangen lichtsterkte, laat de eerdergenoemde optische illusie de afzonderlijk waargenomen lichtflitsjes samensmelten tot een (wazige versie van het) beeld aan de zenderkant.



De uitvinding van deze zogenaamde Nipkowschijf was niet alleen de wegbereider van de televisie (een pretentieuze prijs voor televisieprogramma's heet tot op vandaag nog steeds de 'Nipkowschijf'), zij was ook bepalend voor het denken over beeldrepresentaties in latere technologische contexten.

Toen, begin jaren zeventig van de vorige eeuw, computers voldoende krachtig begonnen te worden om met grote hoeveelheden gegevens te werken, was het dan ook niet onlogisch om een representatie voor beelden voor te stellen waarin deze gerepresenteerd zouden worden als een reeks kleurwaarden of grijswaarden, waarbij elke waarde wordt toegewezen aan één, van te voren bepaalde locatie

in het beeld, analoog aan de tevoren gekozen plaatsen van de gaatjes in de Nipkowschijf. Waar zouden deze locaties beter gekozen kunnen worden dan gerangschikt op een rechthoekig, regelmatig rooster, en ziedaar, de pixel was geboren. Natuurlijk lijkt een rechthoekige, discrete verdeling van grijswaarden in niets op een portret ...



... tenzij het aantal pixels voldoende groot is. Dan zorgt dezelfde optische illusie, die de vinding van Nipkow mogelijk maakte, ervoor dat we geen afzonderlijke pixels meer zien, en in de veronderstelling verkeren naar een ongeschonden beeld te kijken.

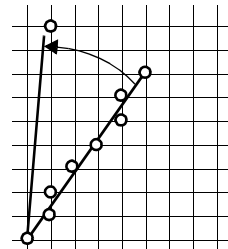
Voor veel generaties beeldtechnologen, computerdeskundigen en televisiemakers was hiermee, tenminste gedurende de jongste drie of vier decennia, de kous af. Dat wil zeggen: in industriële laboratoria en op universiteiten moest een grote hoeveelheid inspanning getroost worden om een aantal opmerkelijke onhebbelijkheden van deze aanpak te onderdrukken – we komen hier zo dadelijk nog op terug – maar al deze ingrepen werden opgevat als ‘noodzakelijk kwaad’, zoals een airbag nu eenmaal een noodzakelijke voorziening is als je met een auto hard wil kunnen rijden en toch de kans wil hebben om botsingen te overleven.

Maar voor wiskundigen bleef er toch iets knagen. Immers, er is een groot aantal betekenisvolle bewerkingen op beelden dat niet mogelijk is op een rooster van pixels. Transleren, anders dan over een geheel veelvoud van pixels, roteren over een andere hoek dan een veelvoud van 90 graden, en schalen met een niet-gehele schaalfactor hebben geen eenduidige betekenis. Toch zijn dergelijke bewerkingen aan de orde van de dag in talloze toepassingen van beeldbewerking.

Laten we als voorbeeld kijken naar het roteren van een lijnstuk, bijvoorbeeld tussen de punten (0,0) en (5,7), over een hoek van 30 graden rondom de oorsprong. Elementaire berekening laat zien dat het geroteerde lijnstuk loopt tussen (0,0) en (0.830127..., 8.562177...). Maar dat betekent dat dit eindpunt niet op een pixel ligt, en dat we moeten afronden naar (1,9). Het lijnstuk is daarmee maar liefst ruim 5% langer geworden – hetgeen een buitengewoon merkwaardig resultaat is voor een rotatie.

We kunnen hier moeiteloos talloze voorbeelden aan toevoegen van elementaire meetkundige doodzonden die veroorzaakt worden doordat we ons moeten beperken tot een rechthoekig rooster van pixels. Bijvoorbeeld: het

middelpunt van een lijnstuk wordt bij rotatie niet afgebeeld op het geroteerde middelpunt, twee evenwijdige lijnstukken kunnen naar believen 0, 1 of willekeurig veel pixels gemeen hebben, en dit geldt ook voor twee snijdende lijnstukken, enzovoort. De gehele Euclidische meetkunde, en daarmee de theoretische grondslag van ons begrip van beeldinhouden, houdt op betekenis te hebben op pixelroosters.



Beeldtechnologen en computerontwerpers, geconfronteerd met deze rariteiten, hadden twee keuzes. Men kon afzien van het idee beelden te representeren met pixels, of men kon proberen de schade te beperken. Misschien onder druk van de markt en het niet onmiddellijk voorhanden zijn van een alternatief voor pixels, koos men voor het laatste. Men ging *filteren*.

In de figuur op de volgende pagina (kolom 1) is hetzelfde lijnstuk te zien van de oorsprong naar (5,7), ditmaal opgevat als een pixelplaatje waarbij elk pixel door een zwart of grijs vierkantje is weergegeven. Als eerste wordt het beeld een beetje gefilterd of uitgesmeerd. Wiskundig is zo'n filtering een *convolutie*: als $F(x,y)$ een functie is die aan een pixel (x,y) een grijswaarde toevoegt, is de convolutie $F_s(x,y)$ van F met een gegeven uitsmeerfunctie G_s gegeven door:

$$F_s(x,y) = \sum_{x'=x-h \dots x+h, y'=y-h \dots y+h} F(x',y') G_s(|x'-x|, |y'-y|). \quad (1)$$

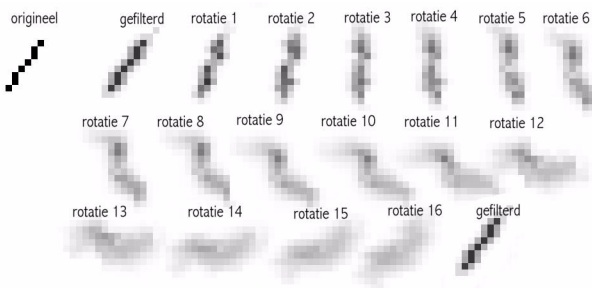
Hierin is $G_s(|x'-x|, |y'-y|)$ een functie die aangeeft hoe groot de invloed is van een pixel in het origineel op locatie (x',y') op het resulterende plaatje in locatie (x,y) . Zo'n functie heet een *filterkern*. Die filterkern $G_s(|x'-x|, |y'-y|)$ bereikt zijn maximum in het punt (0,0) en loopt snel af voor toenemende $|x'-x|$ en $|y'-y|$. G_s kan voorgesteld worden als een soort symmetrische bult². De index s duidt op de breedte van de bult. Daarom zorgt G_s ervoor dat het signaal F uitgesmeerd wordt, en wel des te breder naarmate s groter is. Er gaat dan meer detail in F verloren. Er zijn allerlei vormen van filterkernen denkbaar die allerlei andere bewerkingen op F laten ontstaan; kenmerkend is echter dat in het algemeen meer detail in F verloren gaat naarmate de filterkern breder is. We zullen daar straks meer over horen.

Het effect van filteren door middel van zo'n overall positieve functie G is alsof het beeld wat vager gemaakt wordt, alsof we door onze ooghaartjes naar het beeld kijken.

In alle plaatjes hieronder, genaamd ‘rotatie 1’ ... ‘rotatie 16’ is dit gefilterde beeld vervolgens steeds over (ongeveer) 22,5 graden gedraaid. Omdat het gedraaide beeld echter weer in pixels gerepresenteerd moest worden, moest er steeds een kunstgreep uitgehaald worden. Om namelijk te bepalen wat de nieuwe grijswaarde in een pixel, zeg op plaats (x,y) , moet worden, nemen we een gemiddelde van alle pixels uit het oorspronkelijke beeld. Om dat gemiddelde te berekenen geven we de grijswaarde in enig pixel zeg op (x',y') een weegfactor waarbij we die weegfactor sterk laten afnemen met de afstand tussen $R(x',y')$ en (x,y) . Hierbij stelt R de rotatieafbeelding voor. Ook voor deze afname gebruiken we weer onze functie G_s . Dus een pixel (x',y') dat door R toevallig op of vlakbij (x,y) terecht komt krijgt een grote weegfactor, en telt dus belangrijk mee bij het bepalen van de grijswaarde in (x,y) , terwijl een pixel in (x',y') dat door R ver van (x,y) wordt afgebeeld niet of nauwelijks bijdraagt. Kort gezegd: ook hier passen we weer een convolutie toe die sterk lijkt op (1):

$$F_{s\sim}(x,y) = \sum_{x'=x-h\dots x+h, y'=y-h\dots y+h} F(x',y') G_s(|R_x(x',y') - x|, |R_y(x',y') - y|) \quad (2)$$

We schrijven hier $R_x(p,q)$ voor de x -component van het over R geroteerde punt (p,q) ; idem voor R_y . Ook hebben we dus weer een keuze moeten maken voor de breedte s van de bult G . Als we deze s te klein nemen ontstaat een erg rafelig beeld; we noemen dat aliasing. Het welbekende Moiré-effect is een voorbeeld van aliasing. Als we s voldoende groot nemen opdat geen al te nadrukkelijke rafels ontstaan, wordt het beeld echter bij elke bewerking een stukje vager. We zien dat goed geïllustreerd als we 16 keer achterelkaar diezelfde rotatie uitvoeren. We weten uiteraard dat 16 rotaties achterelkaar over 22,5 graden de identieke afbeelding oplevert. Maar door het steeds weer opnieuw convolueren met (2) blijft er weinig over van het oorspronkelijke beeld (rechtsonder in onderstaande figuur zetten we voor de duidelijkheid het gefilterde lijnstuk in zijn oorspronkelijke oriëntatie nog eens naast rotatie 16).



In de meeste toepassingen van beeldtechnologie is een reeks van zestien bewerkingen achterelkaar tamelijk veel. Bovendien is de resolutie van moderne beeldschermen zo hoog, dat een klein beetje onscherpte niet altijd onmiddellijk opvalt.

Naarmate beelden meer en meer digitaal worden, en naarmate er meer soorten uiteenlopende beeldschermen zullen komen (uiteenlopend van enkele vierkante centimeters tot meerdere vierkante meters), met sterk uiteenlopende resoluties, kunnen we het dilemma tussen aliasing en onscherpte echter niet blijven ontlopen. En zoals we hierboven gezien hebben, is het dilemma niet simpelweg het resultaat van onzorgvuldig geïmplementeerde beeldbewerkingen. In plaats daarvan is het een onvermijdelijk, rechtstreeks resultaat van het feit dat beelden worden gerepresenteerd als reeks (discrete) grijswaarden of kleurwaarden op een rooster dat a priori gekozen is, dat wil zeggen dat geen enkele rekening houdt met de inhoud van het beeld. We zullen dit probleem, dat ontstaan is met de uitvinding van de Nipkowschijf, aanduiden als de Nipkowdoctrine.

De centrale vraag is uiteraard of aan de Nipkowdoctrine te ontsnappen is. Met andere woorden: bestaat er een representatie van beelden die het toestaat om aan beelden te rekenen, om ze in computergeheugen op te slaan, en om er verstandige bewerkingen op te kunnen uitvoeren, zonder dat bij elke bewerking de kwaliteit slechter wordt en zonder dat we last hebben van aliasing.

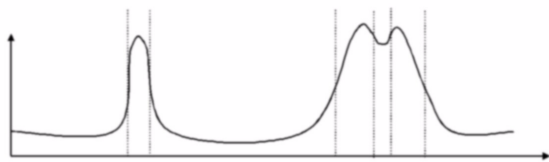
Hoe diepzinnig deze vraag ook schijnen moge, we kunnen inspiratie voor een mogelijk antwoord halen uit een kinderlijk eenvoudige hoek, namelijk uit een tekening. En meer in het bijzonder, een kindertekening.



Links zien we een grijswaardenafbeelding van een tafel, en rechts een kindertekening hiervan. Als we gemakshalve aannemen dat die tekening de meest essentiële elementen van het beeld representeert, zien we dat een heleboel detail verloren is gegaan (of lijnen recht zijn of krom is bijvoorbeeld niet belangrijk). De informatie die is overgebleven, betreft de zogenaamde singulariteiten in het beeld, met name randen en hoeken. Een singulariteit is, in de meest letterlijke zin, een bijzonder punt. Het is een punt waarvoor geldt dat in de omgeving ervan het beeld ‘saai’ is: alleen in de singulariteiten verandert iets wezenlijks. Kennelijk moeten we, als we de essentie van een beeldinhoud willen representeren, op zoek gaan naar de punten waar iets drastisch verandert.

Wiskundig is dit te vertalen naar tekenwisselingen in afgeleiden, met name in tweede afgeleiden. Een tweede afgeleide kan positief zijn of negatief, en de omslag van teken zullen we aanduiden als singulariteit. Beschouw het onderstaande eendimensionale signaal (de stippellijnen

geven de locaties van tekenwisselingen van de tweede afgeleide aan):



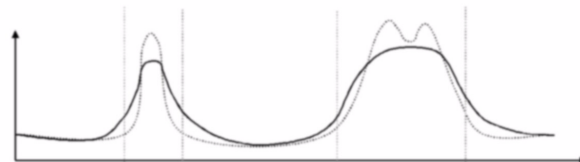
Kwalitatief zien we hier twee ‘objecten’, laten we zeggen ‘bulten’, waarbij de breedste van de twee nog een klein ‘objectje’ omvat, namelijk een ‘kuiltje’. Deze niet-wiskundige noties kunnen we makkelijk formaliseren: *elk paar opvolgende tekenwisselingen van de tweede afgeleide begrenst een object*. Uiteraard hebben we nog wat meer informatie nodig dan alleen de plaats van de tekenwisselingen om de ‘hele’ inhoud van het beeld te beschrijven, maar het is belangrijk om op te merken dat we in ieder geval geen beeldelementen met a priori gegeven plaatsen (zoals pixels) hebben geïntroduceerd. Daardoor vermijden we het duivelse dilemma tussen aliasing en onscherpte. We kunnen nu bijvoorbeeld bewerkingen op het signaal uitvoeren door voor te schrijven op welke manier de singulariteiten bewerkt moeten worden (opgeschoven, of, in een tweedimensionaal signaal, geroteerd of nog anderszins getransformeerd).

We merken nog iets anders op. In het bovenstaande signaal is weinig discussie over de linkerbult als zelfstandig object. Maar hoe zit het nu met de rechterbult? Is dat één brede bult, of zijn het twee smalle bulten op korte afstand, of is dat een bult met een kuiltje erin? Al deze interpretaties zijn verdedigbaar; ze veronderstellen alleen allemaal een verschillende schaal waarop we naar het signaal kijken. Het aardige is dat we ook deze interpretatiekwes- tie elegant in wiskundige termen kunnen uitdrukken. Daar- toe moeten we eerst vertellen hoe die tweede afgeleide berekend kan worden. Zoals bekend is differentiëren niet numeriek stabiel (ruis op het signaal wordt versterkt). Maar we kunnen handig gebruikmaken van een elemen- taire stelling uit de integraalrekening. Voor twee functies, $F(x)$ en $G_s(x)$, waarbij $G(x)$ voldoende snel naar 0 gaat als $|x|$ groot wordt, geldt:

$$\int_{-\infty}^{\infty} F(x) \frac{d}{dx} G_s(x) dx = - \int_{-\infty}^{\infty} G_s(x) \frac{d}{dx} F(x) dx \quad (3)$$

Analoge relaties gelden voor tweede afgeleiden. Als we weer voor G_s onze filterkern denken kunnen we (3) als volgt interpreteren: ‘als we een signaal F convolueren (filteren) met de afgeleide van een filterkern G_s is dat *precies hetzelfde* alsof we de afgeleide van het signaal con- volueren met de oorspronkelijke filterkern G_s ’. Niet al- leen hebben we hier een stabiele manier gekregen om de

tweede afgeleide van een signaal te berekenen, we zien ook dat het berekenen van een tweede afgeleide altijd im- pliciet vergt dat we een keuze moeten doen voor de waar- de van s . Met andere woorden: we moeten zeggen in wel- ke mate van details we geïnteresseerd zijn. Als we de bo- venstaande figuur nog een keer tekenen, maar nu met een andere schaal, zien we duidelijk dat er automatisch een andere toekenning van objecten tevoorschijn komt. Met name zien we dat er op deze grotere schaal nog maar vier van de zes tekenwisselingen van de tweede afgeleide overblijven, hetgeen mooi overeenkomt met onze intuïtie dat er nog maar twee ‘objecten’ op deze schaal te onder- scheiden zijn.



Wij zijn ons ervan bewust dat we veel (lastige en subtiele) details links hebben laten liggen. Het feit dat er in (3) integralen staan terwijl we eerder in (1) en (2) over som- men spraken, is er slechts één van; de generalisatie van dit verhaal naar tweedimensionale signalen is een andere, technisch veel lastigere kwestie, die ons ver buiten de be- doeling van dit artikel zou voeren. Niettemin hopen we een paar wezenlijke aspecten van beeldrepresentatie aan de orde te hebben gesteld: het fundamentele dilemma tus- sen aliasing en onscherpte, veroorzaakt door het gebruik van pixels om beelden te representeren, kan alleen opge- lost worden door een drastisch andere keuze van de repre- sentatiewijze. Discrete singulariteiten lijken recht te doen aan wat wij als mensen ervaren als ‘de essentie’ van een beeld, en om deze singulariteiten te kunnen vinden moe- ten we ons uitspreken over de schaal waarop we naar het beeld kijken. Ofschoon dit in het vakgebied van de beeld- bewerking een nog verre van uitgemaakte zaak is, lijkt het erop alsof een nieuwe technologie, gebaseerd op zo- genaamde *scale spaces*, een interessante kanshebber is. We eindigen daarom ook met de geïnteresseerde lezer door te verwijzen naar het boek *Image Structure* van Luc Florack (Kluwer Academic Publishers, 1997).

Kees van Overveld, Technische Universiteit, Eindhoven

Noten

- [1] Het is precies deze benadering die, bij nader inzien, aanleiding geeft tot de fundamentele problemen bij het reconstrueren van het beeld waar we ons in dit artikel druk over maken.
- [2] Vaak kiezen we hier een Gaussiaan voor: $G_s(p,q) = C \exp(-(p^2+q^2)/s^2)$ waarbij C een goed gekozen normeringsfactor is om te zorgen dat de som van alle G_s waarden één oplevert.