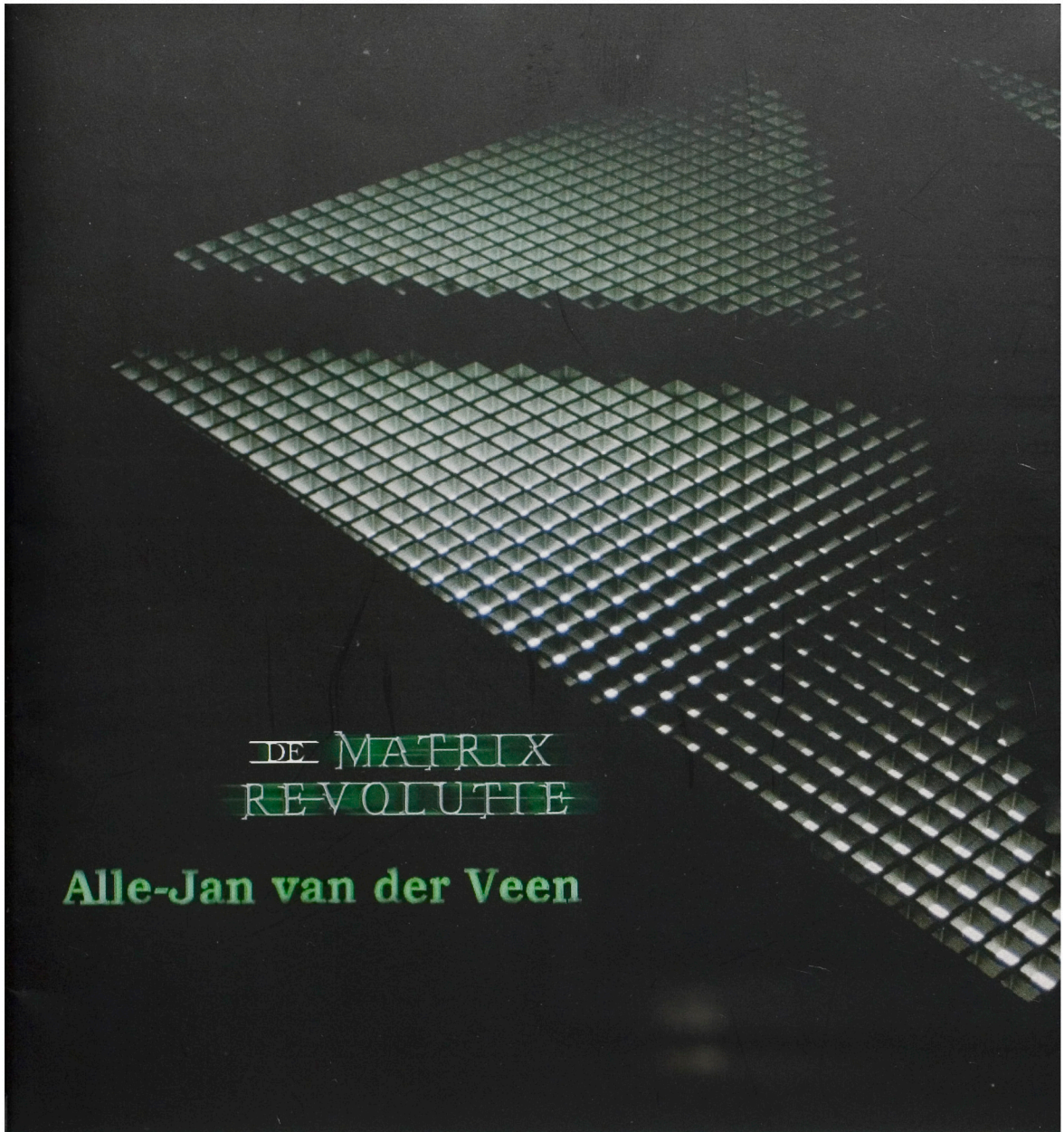


DE MATRIX REVOLUTIE

Alle-Jan van der Veen

Intreerede
18 mei 2005

TU Delft
Technische Universiteit Delft

Faculteit Elektrotechniek, Wiskunde en Informatica

830464
3843548
Red. 2005

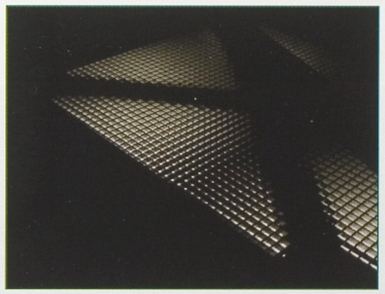
De Matrixrevolutie

Intreerede

Uitgesproken op 18 mei 2005
ter gelegenheid van de aanvaarding van het ambt van
Antoni van Leeuwenhoek hoogleraar
in het vakgebied van Array Signaalbewerking
aan de faculteit Elektrotechniek, Wiskunde en Informatica
van de Technische Universiteit Delft

door

Alle-Jan van der Veen



BIBLIOTHEEK TECHNISCHE UNIVERSITEIT
Prometheusplein 1
2628 ZC
DELFT
Veen-
Red-
2005

© Alle-Jan van der Veen, 2005

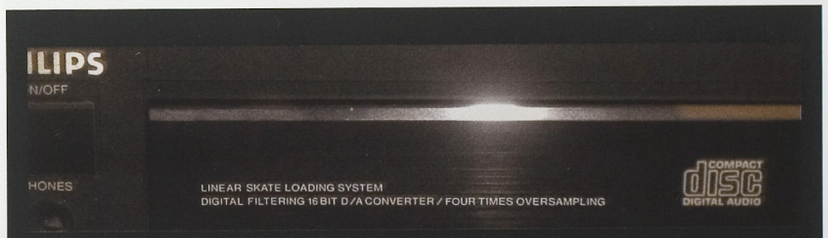
Mijnheer de Rector Magnificus, leden van het College van Bestuur,
Collegae hoogleraren en andere leden van de universitaire gemeenschap,
Zeer gewaardeerde toehoorders,
Dames en heren,

Signaalbewerking is onzichtbaar...

Signaalbewerking is een onzichtbaar vak. Het gemiddelde publiek weet niet wat signaalbewerking is, laat staan dat hij het nodig heeft. Hierin staat signaalbewerking niet alleen, bijvoorbeeld de computerindustrie heeft hetzelfde probleem. Er is een zwarte doos en je weet niet wat er in zit.

Een fabrikant als Intel, producent van microprocessoren, heeft dat opgelost door op iedere computer een sticker te plakken "Intel inside". Bij producten met veel signaalbewerking wordt dit ook wel eens geprobeerd, bijvoorbeeld door op CD spelers te schrijven "4 times oversampling", "16 times oversampling" etcetera. Maar ik denk niet dat er veel mensen zijn die weten wat dat eigenlijk betekent, en waarom een hoog getal nu zoveel beter is (dat hoeft namelijk helemaal niet zo te zijn).

Misschien een betere omschrijving voor signaalbewerking is de slogan van Lucent "We make the things that make things work". In het komende half uur zal ik u laten zien waarom dat zo is.



Figuur 1. Signal processing inside

1. Het marktaandeel van signaalbewerking

Laat ik beginnen met een "executive summary", een marktanalyse. Als signaalbewerking onzichtbaar is, hoe groot (of belangrijk) is signaalbewerking dan eigenlijk? Zo'n vraag is natuurlijk op verschillende manieren te beantwoorden. Om met de tijdgeest mee te gaan zal ik het hier meten aan de hand van publicaties.

De belangrijkste organisatie voor elektrotechnische ingenieurs is de IEEE, het "Institute for Electical and Electronics Engineers". Het aantal leden is meer dan 300.000, wereldwijd. Er zijn 42 ondervereningingen, Signaalbewerking is de drie-na grootste.

De IEEE publiceert een honderdtal tijdschriften, de meeste zijn de belangrijkste op hun vakgebied. Al die tijdschriften zijn tegenwoordig online beschikbaar. Als we kijken welke tijdschriften het meest geraadpleegd worden zien we het volgende (Tabel 1). Het belangrijkste tijdschrift is kennelijk de Transactions on Solid State Circuits, met anderhalf miljoen downloads in 2004; dit tijdschrift gaat grofweg over transistoren. Nummer twee is microgolftechniek, van belang voor mobiele communicatie, op de voet gevolgd door signaalbewerking. De hieropvolgende tijdschriften scoren al beduidend lager. Kennelijk is signaalbewerking dus best wel een belangrijk vakgebied.

Hoe ziet dit vakgebied eruit? Als je kijkt naar het aantal inzendingen naar de belangrijkste conferentie (IEEE ICASSP), of naar het aantal inzendingen naar de IEEE Signal Processing Letters waarvan ik de gegevens heb, dan zie je dat het vakgebied ruwweg in vieren te delen is (zie Tabel 2). Er zijn technieken en toepassingen. De technieken zijn het basis-gereedschap, zoals bijvoorbeeld statistische signaalbewerking, adaptieve filters, en tijd-frequentie-analyse. De toepassingen zijn onder te verdelen in drie ruwweg even grote onderdelen: beeldbewerking en multimedia, spraakbewerking en audio, en "signaalbewerking voor communicatie" tezamen met array signaalbewerking. Dit laatste is het vakgebied dat ik wil beakkeren.

De toepassingen van signaalbewerking zitten onzichtbaar weggewerkt in vrijwel alle high-tech apparaten die we tegenwoordig om ons heen hebben: bijvoorbeeld mobiele telefoons, digitale camera's, MP3 spelers, GPS navigatie, MRI, ultrasound en andere medische scanners, digitale hoortoestellen, etcetera.

Signaalbewerking voor communicatie gaat over het versturen en ontvangen

Tabel 1. Downloads van artikelen uit IEEE tijdschriften, 2004

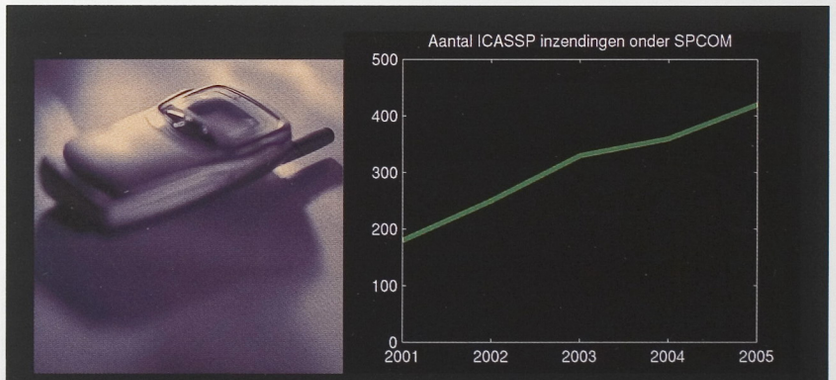
1.	Solid state circuits	1.500.000
2.	Microwave technology	1.030.000
3.	Signal processing	930.000
4.	Magnetics	790.000

Tabel 2. Onderverdeling Signaalbewerking

Technieken	35%
Toepassingen:	
Communicatie en array processing	25%
Beeldbewerking en multimedia	20%
Spraak en audio	20%

van signalen, bijvoorbeeld voor mobiele telefoons, en hoe je dat het beste kan doen. Het is nu al de grootste toepassing binnen signaalbewerking, en als je kijkt naar het aantal inzendingen naar ICASSP groeit het gestaag, ten koste vooral van het aandeel "basis-technieken" (Figuur 2). We zien ongeveer een verdubbeling in de laatste 4 jaar, ik kom hier later op terug.

Hoewel de groei lineair lijkt te zijn en niet exponentieel, roept deze curve toch onmiddelijk suggesties op met de wet van Moore.



Figuur 2. Groei van Signaalbewerking voor Communicatie

2. De wet van Moore

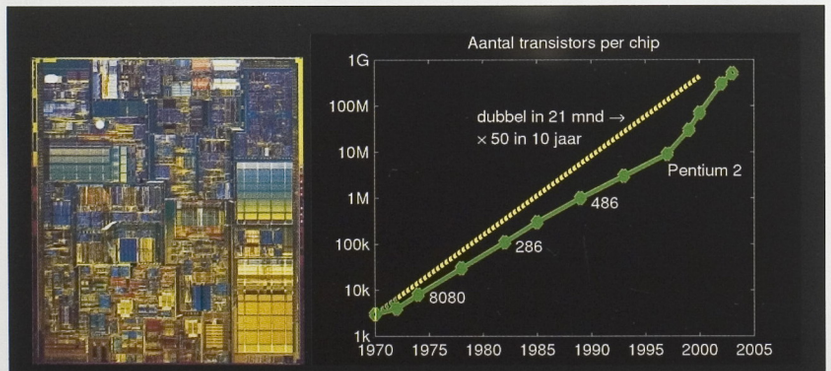
Ik heb zojuist de wet van Moore genoemd, en ik zal dat wat verder toelichten. Gordon Moore is een van de oprichters van Intel. In 1965 stelde hij dat het aantal transistoren per chip ruwweg iedere 18 maanden verdubbelt. Dat komt neer op een factor 100 iedere 10 jaar (Figuur 3) [1].

In de praktijk zien we een iets kleinere groei, een factor 50 iedere 10 jaar, maar dat is nog steeds ongeevenaard. Het is ongelooflijk dat deze groei nu al 35 jaar lang volgehouden is, en naar verwachting nog wel 10 tot 15 jaar doorgaat. Het is deze exponentiele groei die verantwoordelijk is voor de computer-revolutie die we in de laatste 30 jaar meegemaakt hebben, en die minstens zo belangrijk is als de industriele revolutie van de 19e eeuw.

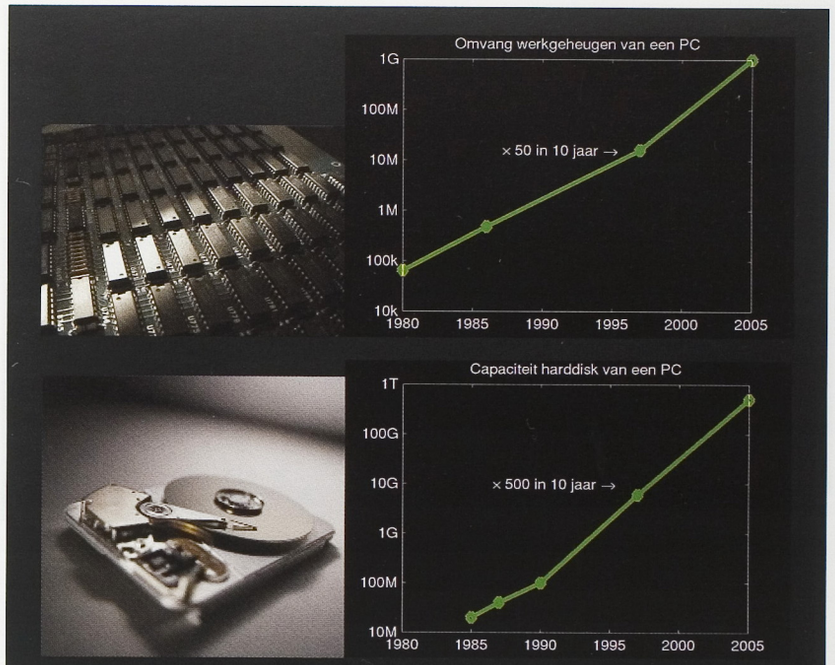
Niet alleen de transistordichtheid verdubbelt iedere anderhalf jaar. Ook bijvoorbeeld werkgeheugens zijn exponentieel gegroeid: van minder dan 100 kB in 1980 tot 1 GB in 2005 (Figuur 4). Computers zijn dus ongelofelijk veel krachtiger dan vroeger. De grote vraag is wat we met deze pure rekenkracht gaan doen.

Wat gaat sneller dan de wet van Moore? Het wordt wel gezegd dat Bill Gates transistoren sneller verbruikt dan Intel ze kan maken [2], zodat bijvoorbeeld Windows tegenwoordig niet sneller opstart dan vroeger. Het is een feit dat magnetische opslag sneller gegroeid is dan de wet van Moore, een verdubbeling van geheugen iedere 9 maanden, of meer dan een factor 500 per 10 jaar. Tegenwoordig zijn harddisks van meer dan 1 TB (ofwel 1000 GB) goed betaalbaar. De nieuwe mogelijkheden worden vooral gebruikt voor de opslag van films, foto's en muziek.

We zien dus dat er een reden is dat de Transactions on Solid State Circuits en de Transactions on Magnetics zo hoog scoren op de download-ranking, respectievelijk nummer 1 en 4. Deze tijdschriften gaan namelijk precies over de technologieen die direct gerelateerd zijn aan de wet van Moore: transistoren en magnetische opslag. De vraag is, waarom staat signaal-



Figuur 3. De wet van Moore



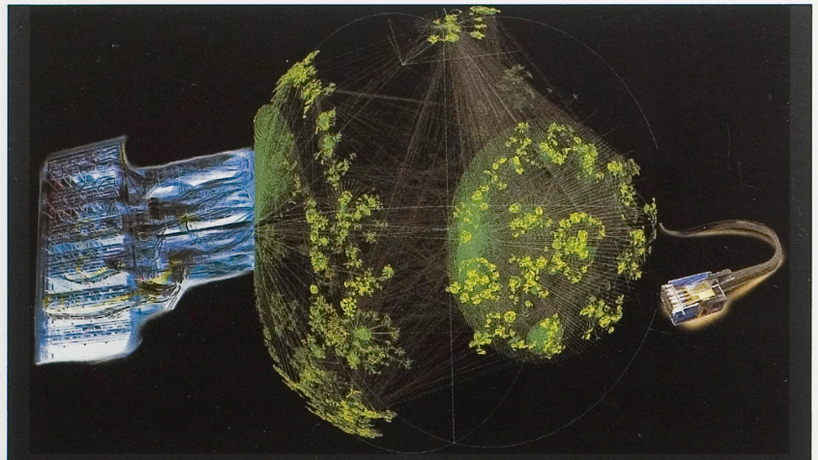
Figuur 4. De wet van Moore voor geheugens en magnetische opslag

bewerking hiertussen, op nummer 3? Volgt signaalbewerking soms ook de wet van Moore? En als dat zo is, waarom dan wel? (Anders gezegd: als je meer transistoren hebt, heb je dan ook meer onderzoek nodig naar signaal-bewerking?)

De netwerkrevolutie

Voordat ik hierop inga, eerst het volgende. De laatste 10-15 jaar heeft nog een andere revolutie laten zien: de netwerkrevolutie. Internet zoals we dat nu kennen bestond 10 jaar geleden nog niet. Tegenwoordig zijn vrijwel alle computers aan het internet gekoppeld. De hoeveelheid informatie die online beschikbaar is groeit sneller dan de wet van Moore, het is namelijk de combinatie van magnetische opslag en het feit dat er steeds meer computers online komen. Het merendeel van die terabytes staat niet bij u thuis, maar ver weg op servers, gedistribueerd over het netwerk, en deze trend zet zich door (Figuur 5) [3].

De overgebleven schakel om deze informatie bij u thuis te krijgen is het communicatienetwerk. Ook hier heeft een revolutie plaatsgevonden, namelijk de ontwikkeling van glasvezeltechnologie, begonnen in 1970. Volgens "de wet van Gilder" [2] groeit de totale capaciteit (bandbreedte) van communicatiesystemen met een factor 3 ieder jaar: dat is sneller dan de wet van Moore, en zelfs sneller dan magnetische geheugens. De huidige stand van de



Figuur 5. Het internet

techniek is dat met behulp van Wave Division Multiplexing (WDM) in het laboratorium meer dan 10 petabit per seconde gehaald is (dat is 10 miljoen Gbps), ofwel de informatie uit ongeveer alle boeken die ooit geschreven zijn per seconde. Dat is voorlopig snel genoeg, en we lijken dus gered.

Helaas. Zoals u zelf vast ook wel ondervonden heeft is de praktijk weerbarstig. U heeft thuis geen 10 Pbps, u heeft zelfs geen Gbps, u heeft waarschijnlijk ADSL met hooguit 1 Mbps, een orde miljard keer lager. Waar zit de bottleneck? Het punt is dat het veel te duur is om uw huis op glasvezel aan te sluiten. Dit heet het last-mile probleem: tot aan het PTT huisje op de hoek van de straat is de bandbreedte onbeperkt, maar de laatste paar honderd meter moet alles door een dun koperdraadje, of misschien zelfs draadloos.

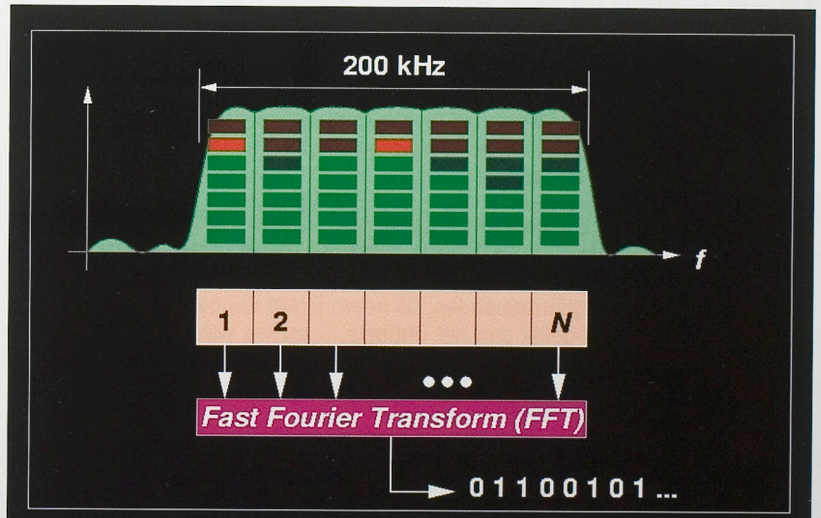
Ook voor draadloos is er een groeiende communicatiebehoefte (Figuur 6) [6]. We kunnen al geen dag meer zonder email, inclusief attachments van een paar megabytes. Zelfs als we op vakantie zijn. We willen draadloos via internet kunnen bellen en chatten, eventueel met beeld. En in de toekomst willen we "telepresence", dat wil zeggen 3-dimensionale video, in grootbeeld [5]. Philips heeft recent de TV technologie hiervoor gepresenteerd. Nu de communicatie nog...

Alles moet dus draadloos, of op z'n hoogst door een heel dun draadje. Wat zijn de ontwikkelingen hier? Dertig jaar geleden waren de eerste modem-verbindingen 300 bps, twintig jaar geleden was dat 2.4 kbps, daarna 56 kbps in de jaren 90, en nu is het 1 Mbps voor ADSL. Een factor 10 tot 20 iedere tien jaar. Vergelijkbaar met de wet van Moore, maar helaas een factor 5 langzamer. Een relatief bescheiden revolutie dus. Wat staat achter deze revolutie, en wat verklaart de verloren factor 5? Dat is nu de 'onzichtbare' signaalbewerking.

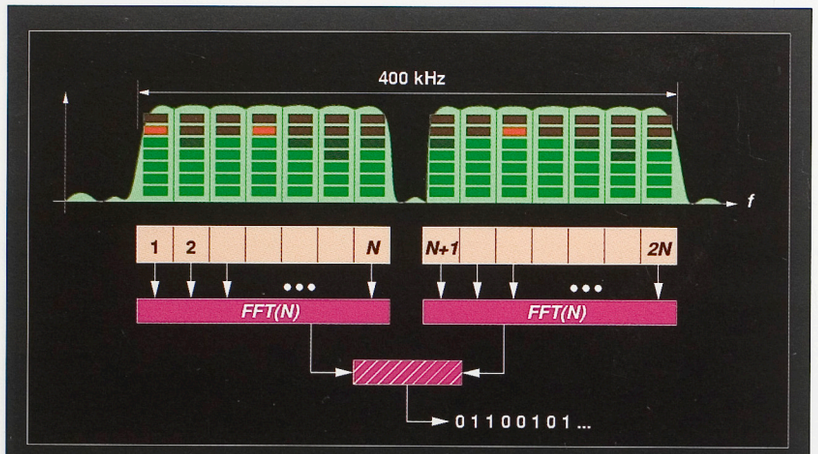
Figuur 6. Nomadic computing

Capaciteit en complexiteit

Signaalbewerking is het rekenen aan signalen. De wet van Moore geeft ons meer rekenkracht, signaalbewerking gaat over het aanwenden van die kracht om nuttige bewerkingen te doen, zoals het verhogen van de communicatie-snelheid. Die factor 5 verlies vergeleken met de wet van Moore betekent dat iedere stap voorwaarts een factor 5 moeilijker is. Ik zal nu uitleggen waarom dat zo is.



Figuur 7. Een capaciteit van orde *N* bits per seconde voor een complexiteit van orde *N* log(*N*) operaties



Figuur 8. Capaciteitsverdubbeling: een capaciteit van 2*N* voor een complexiteit van 2*N* log(2*N*) operaties

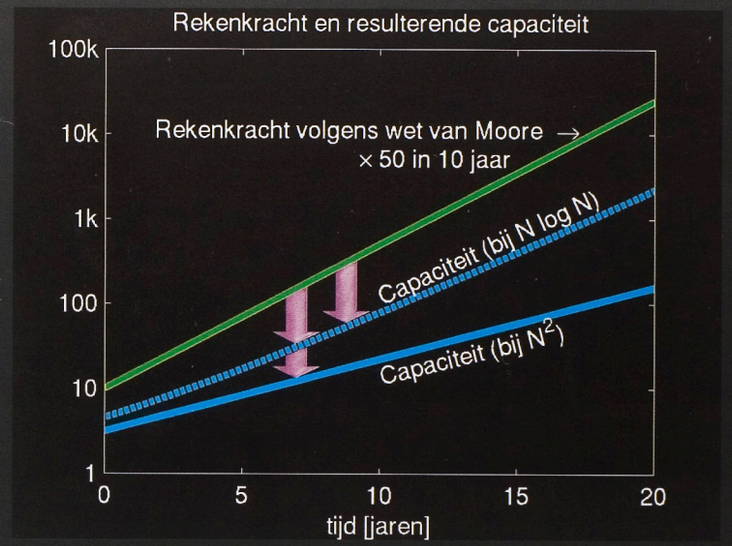
Signalen worden verstuurd in frequentiebanden, en de bandbreedte bepaalt in eerste instantie de hoeveelheid informatie die per seconde kan worden verstuurd (dit is een onderdeel van de wet van Shannon). Om meer informatie over te sturen moet je meerdere frequentiebanden parallel schakelen, en parallel de informatie versturen. Dit is wat we nu op grote schaal zien gebeuren.

De frequentiebanden zijn niet onafhankelijk, ze zijn namelijk onderdeel van hetzelfde fysieke tijdssignaal. Een ontvangen signaal moet eerst opgesplitst worden in frequentiebanden (Figuur 7). U kunt zich dat voorstellen als u thuis een radio heeft met een graphic equalizer, u kunt dan ieder frequentiebandje apart versterken. De digitale techniek voor het splitsen in frequentiebanden heet de Fast Fourier Transform (FFT), (her)uitgevonden in 1965 door Cooley en Tukey [4]. Dit is een van de fundamenten van de signaalbewerking. De complexiteit van deze operatie is van orde $N \log(N)$, voor een capaciteit van N frequentiebanden. Het doel is de capaciteit te verhogen (meer bits per seconde), dat kost dus ook meer operaties.

Hoe kunnen we de capaciteit verdubbelen? Het simpelste is dan natuurlijk om twee sets van N frequentiebanden te nemen (Figuur 8). De complexiteit hiervan is niet precies gelijk aan het dubbele, want we willen terug naar een tijdsignaal en daarom moeten de twee sets gecombineerd worden.

De complexiteit is $2N \log(2N)$, en dat is gelijk aan 2 maal $N \log(N)$, plus nog een rest-term, $2N \log(2)$. De eerste term is het dubbele van wat we voor één set frequenties hadden (dat zijn de twee FFTs), en de rest-term is een verliesterm, de complexiteit van het combineren van de twee.

Je kunt dit veralgemenen. Stel in 10 jaar hebben we processoren die 100 keer krachtiger zijn. Kunnen we dan ook 100 frequentiebanden parallel verwerken?



Figuur 9. Communicatie verliest het van Moore

Nee, want ze zijn onderling afhankelijk. Met 100 keer meer rekenkracht en een complexiteit van $N \log(N)$ kunnen we slechts ongeveer 20 keer de communicatiebandbreedte verhogen ($20 \log(20) \approx 100$), en we verliezen dus een factor 5 aan de meest banale stap in het signaalbewerkingsproces.

Communicatie verliest het van Moore...

Ik heb dit in een grafiek gezet (Figuur 9). De bovenste lijn geeft de rekenkracht volgens de wet van Moore. De lijn eronder is de resulterende capaciteit bij een complexiteit van $N \log(N)$, dat is al een stuk minder. De onderste lijn is de capaciteit bij een complexiteit van N^2, dit is een gebruikelijke complexiteit voor veel van onze algoritmes.

De situatie lijkt dus tamelijk hopeloos: de hoeveelheid informatie die we willen communiceren groeit sneller dan de communicatiesnelheid, mogelijk gemaakt door signaalbewerking. Dat moet een keer fout gaan. Het verlies groeit: in 10 jaar een factor 5, over 20 jaar is dat al een factor 10, etc. We gaan het zo grandioos verliezen.

Want de factor 5 verlies, veroorzaakt door de complexiteit van de FFT, is nog maar een deel van het verhaal. Voor grote bandbreedtes zijn er allerhande complicaties: steeds hogere eisen ten aanzien van mogelijke tijdsverschuivingen, frequentieverschuivingen en tijdstabiliteit. Het analoge deel van een ontvanger kan dit allemaal niet meer corrigeren, en dit moet dan digitaal worden opgelost: het schatten van deze tijdsvarierende parameters, en het corrigeren hiervan. Dit is wat "signaalbewerking voor communicatie" doet, en

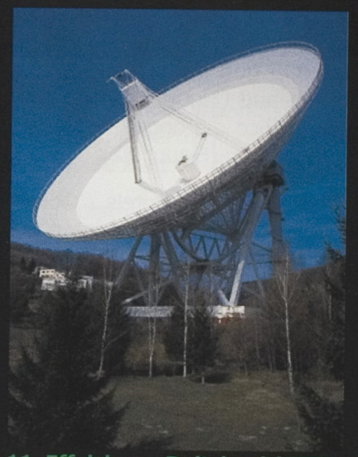
de reden dat de omvang van het onderzoek zo groeit: zoals we eerder zagen een factor 2 in 4 jaar, ofwel als je extrapoleert een factor 5 in 10 jaar!

Het is allemaal geen toeval. Het is nodig om de wet van Moore bij benadering te kunnen bijbenen en het verlies te kunnen opvangen. Over nog eens tien jaar reken ik op een factor 10. Dat is een lineaire groei, maar het gaat hier om mensen. Het betekent dat ik over 10 jaar 5 keer meer mensen in dienst moet hebben, en daarna 10 keer, om aan de verwachtingen te blijven voldoen. Dat stelt dan gelijk de ambitie.

Ik moet u bekennen dat wij in ons laboratorium niet dagelijks met de wet van Moore bezig zijn. Wat we wel doen is communicatiesystemen ontwerpen waar zoveel mogelijk parallel gebeurt. Parallelisme is de sleutel naar het verhogen van de bandbreedte. En we moeten ons hierbij niet beperken tot het parallel gebruik van frequentiebanden: dat is slechts 1 dimensie. We moeten leren twee-dimensionaal te denken, niet alleen de tijd te gebruiken maar ook bijvoorbeeld de ruimte.

Dat brengt mij tot het derde deel van mijn verhaal.

Figuur 10. Idee van Archimedes, 200 v. Chr.



Figuur 11. Effelsberg, Duitsland (1972, 100m)

Figuur 12. Arecibo, Puerto Rico (1960, 305m)

3. Array signaalbewerking

Ik zal u nu een aantal voorbeelden geven van wat wij met ruimte-tijd aan het doen zijn. Het gebruik van ruimte-tijd begint met het gebruik van ruimte, namelijk het gebruik van meer dan 1 antenne. Om dit goed uit te leggen begin ik met een paar eenvoudige voorbeelden.

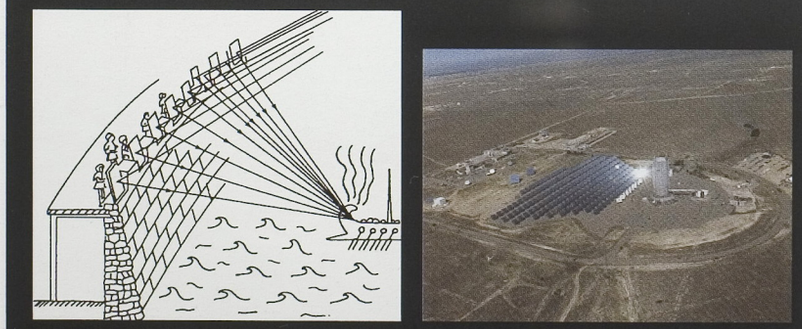
Iedereen kent het gebruik van spiegels. Om licht op een bepaald punt te richten neem je een spiegel. Om de energie verder te concentreren neem je een holle spiegel of een lens. Het gebruik van holle spiegels stamt al uit de tijd van Archimedes, 200 jaar voor Christus, waarin hij bedacht dat dit een manier zou zijn om vijandelijke schepen in brand te steken (Figuur 10).

In de elektrotechniek is het equivalent van een lens of holle spiegel een paraboolschotel. Zo'n schotel wordt gebruikt om zwakke signalen van satellieten op te vangen, of om nog zwakkere signalen van sterren en andere astronomische objecten waar te nemen. In het algemeen geldt hier: groter is beter. Er is echter een grens aan hoe groot je een schotel kunt bouwen. De grootste stuurbare telescoop heeft een diameter van ongeveer 100 m, in Effelsberg bij Bonn in Duitsland (Figuur 11); de grootste "vaste" telescoop staat in een vallei in Puerto Rico die ongeveer de juiste vorm had, en heeft een diameter van ongeveer 300 m (Figuur 12).

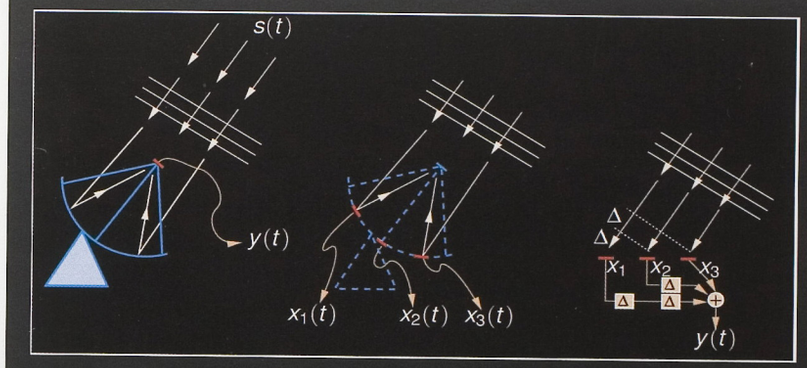
Wat is beter dan één schotel? Twee schotels... Ook dit principe is u bekend. Archimedes stelde al voor om een serie soldaten licht naar een vijandelijk schip te laten reflecteren (Figuur 13(a) [10]). Meer recentelijk hebben we meer civiele toepassingen, bijvoorbeeld het opwekken van energie door het opwarmen van een waterkolom in een toren, met behulp van een matrix aan spiegels (Figuur 13(b)). Dit is dan voor wat betreft licht.

Principe 1: het versterken van signalen

In de elektrotechniek geldt iets soortgelijks (Figuur 14). We kunnen een grote



Figuur 13. (a) Het idee van Archimedes, nu met meer spiegels; (b) Solar Thermal Test Facility, New Mexico. 1978

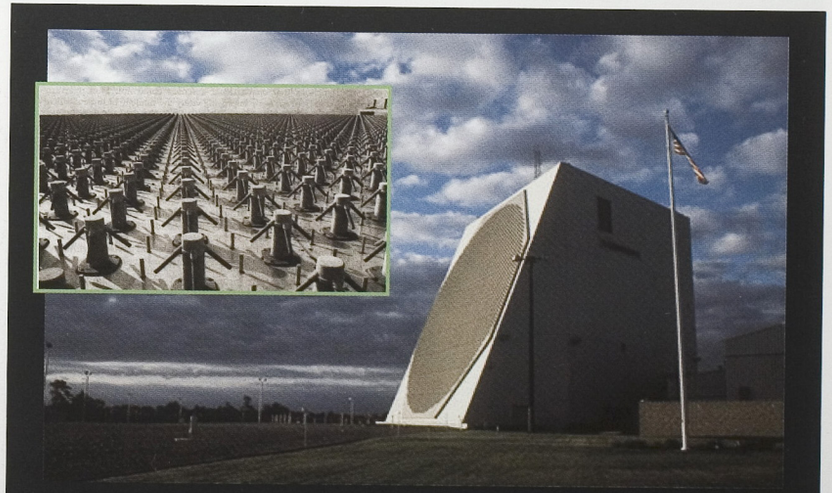


Figuur 14. Vervanging van een paraboolschotel door een serie kleinere antennes (phased array)

schotel opgebouwd denken uit een serie kleintjes, die ieder een deel van het signaal opvangen dat daarna gesommeerd wordt. In tegenstelling tot licht is er een complicatie: de optelling moet wel netjes in fase gebeuren. Het kan dus nodig zijn om de juiste tijdsvertragingen te introduceren, dat zijn de blokjes met een "Delta" erin. Je kunt dit zien als een vorm van focussering.

Je kunt dus heel goed een "platte schotel" bouwen, eentje die niet draaibaar is, maar waarbij alles elektronisch gebeurt. De complexiteit is verplaatst van de mechanica naar de signaalbewerking.

Als voorbeeld hiervan een radarsysteem, misschien wel een van de grootste, bedoeld om binnenkomende raketten op tijd te detecteren (Figuur 15) [10].



Figuur 15. Phased Array Warning System, Cape Cod, 1978

Figuur 16. Very Large Array (VLA), New Mexico, 1978, 27 schotels

Ik schat dat het hier gaat om ongeveer 5000 elementen. Dit lijkt dus op een platte schotel, maar hetzelfde kan er ook anders uitzien. In plaats van simpele antennes kun je ook hele grote nemen, zoals bijvoorbeeld gedaan is met de radiotelescoop in Westerbork, met 14 schotels, of de VLA (Very Large Array) in New Mexico, USA, natuurlijk net iets groter met 27 schotels (Figuur 16). Misschien komt deze u bekend voor, het speelde een rol in de film "Contact", over de ontdekking van buitenaards leven.

Het voordeel van radioastronomie voor ons als toepassing is dat het erg fotogeniek is, mooie foto's van grote schotels. Het is ook een mooie toepassing voor ons omdat alle problemen van signaalbewerking waar wij aan werken er zo mooi in terugkomen. Ik praat hier dan over het vakgebied "array signaalbewerking". Er is hier eigenlijk geen fatsoenlijk Nederlandse vertaling voor, "array" staat voor "een verzameling" antennes. Misschien zou je kunnen zeggen: "matrix signaalbewerking". Dat is ook toepasselijk, omdat alles wat wij doen met matrices te maken heeft: wij doen eigenlijk alleen maar aan lineaire algebra.

Op deze toepassing werken we veel samen met ASTRON, de Stichting Astronomisch Onderzoek in Nederland. Zij beheren ook de radiotelescoop in Westerbork. Zoals bekend speelde die een rol in de verfilming van "De ontdekking van de hemel".

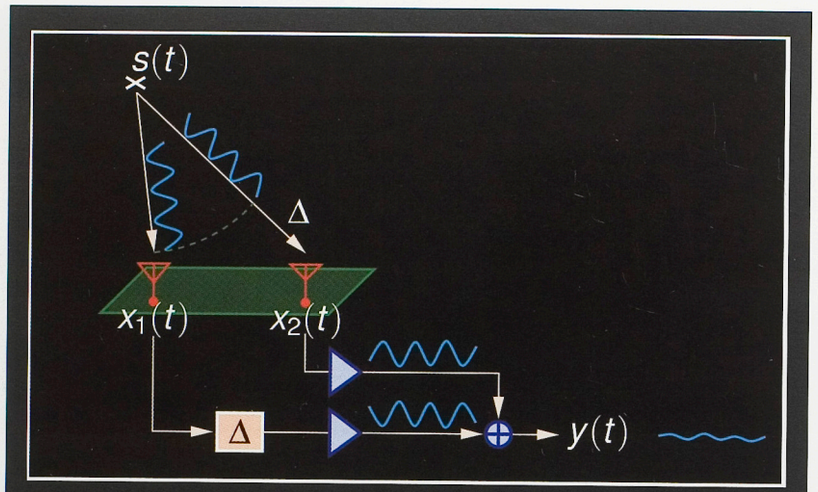
Met behulp van een radiotelescoop kunnen plaatjes gemaakt worden, net zoals voor een optische telescoop maar dan voor veel lagere frequenties. Dat geeft een ander venster op ons heelal, en dus nieuwe inzichten. Cygnus A is een radiobron 600 miljoen lichtjaar van ons vandaan (Figuur 17). Wat we zien

Figuur 17. Cygnus A (een quasar)

is de straling van elektronen die versneld worden door magnetische velden. De elektronen komen uit het centrum, er zit daar een zwart gat. Ik kan er niet veel meer over zeggen, want ik ben geen astronoom. Wat wij doen is helpen met instrumenten bouwen die dit kunnen ontvangen: radio telescopen. Groter is beter: meer resolutie, meer gevoeligheid, meer details.

Het kan groter. Op dit moment wordt op een hoogvlakte in een woestijn in Chili gewerkt aan een radiotelescoop bestaande uit 64 elementen (Figuur 18). Ieder element is een schotel die verrijdbaar is. Het zou klaar moeten zijn in 2011, er staat nu denk ik één schotel, en men is bezig met het aanleggen van een gaspijp vanuit Argentinie om alles van stroom te kunnen voorzien. Het is

Figuur 18. Atacama Large Millimeter Array (ALMA), Chili, gepland voor 2011, 64 verplaatsbare schotels



Figuur 19. Selectief uitdoven

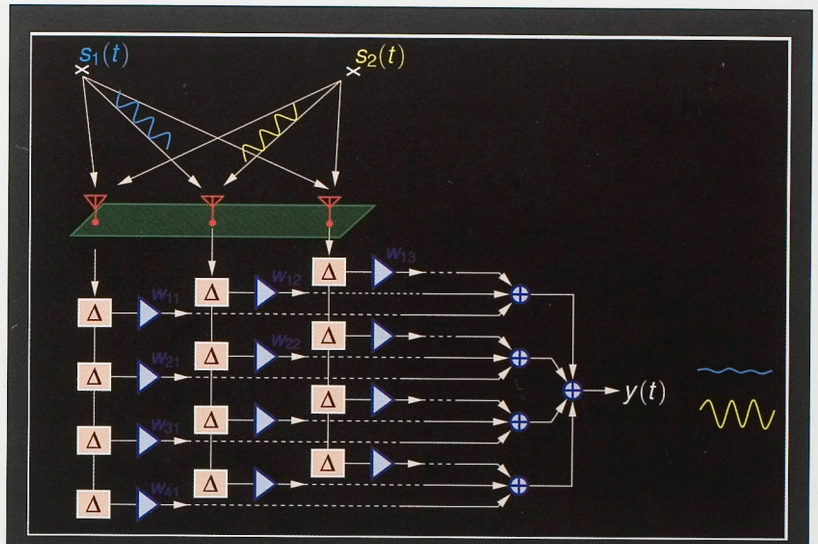
natuurlijk erg onhandig om zover weg te zitten, maar astronomen kozen voor deze locatie omdat de lucht er zo droog en ijl is, dat zorgt ervoor dat de sterren niet zo staan te twinkelen. Bovendien hebben ze in de woestijn minder last van stoorbronnen. Dat brengt mij op het volgende punt.

Principe 2: het nullen van ongewenste signalen

Wat kun je nog meer met twee antennes doen? Je kunt ook proberen een signaal te verzwakken, uit te doven, door het in tegenfase bij zichzelf op te tellen (Figuur 19). Je moet hiervoor precies de juiste tijdvertraging nemen anders lukt het niet. Het resultaat is richtingsafhankelijk: een signaal uit de ene richting wordt verzwakt, maar uit een andere richting juist versterkt. We hebben nu een spatieel-afhankelijke response, een spatieel filter.

De blauwe driehoekjes in Figuur 19 stellen vermenigvuldigingen met filter-coefficienten voor, dat is nodig als je de signalen gewogen bij elkaar wilt optellen. Dat geeft het filter de nodige flexibiliteit. Het is bijna altijd zo dat je deze weegcoefficienten kunt bijregelen, adaptief. Het uitrekenen van deze coefficienten is iets voor signaalbewerking, dat is wat wij doen.

Een spatieel filter is duaal aan een tijd-domein filter. In een simpel digitaal tijd-domein filter wordt een signaal een aantal keren vertraagd, en de vertraagde versies gewogen opgeteld. Signalen met een bepaalde frequentie worden hierdoor versterkt of juist onderdrukt. In een spatieel filter is de rol van frequentie overgenomen door ruimte: signalen uit een bepaalde richting kunnen worden versterkt of onderdrukt. Dat kun je gebruiken om stoor-bronnen weg te filteren.



Figuur 20. Ruimte-tijd filter

Ruimte tijd

Het is ook mogelijk de twee te combineren, we krijgen dan een ruimte-tijd filter (space-time filter), zoals in Figuur 20. De rekenstructuur is nu een twee-dimensionale matrix. Dit is de fundamentele structuur waar wij mee werken: wat wij doen is het uitrekenen van de coefficienten in de matrix zodanig dat we een gewenste response krijgen.

Op zich hoeft deze matrix niet zo groot te zijn, in dit geval 4 bij 3. Het uitrekenen van al deze coefficienten kan wel een hele klus zijn, dat gebeurt ook weer met matrices, maar dan misschien 12 bij 12, of nog groter. De tendens is natuurlijk richting grotere matrices. We hebben tenslotte Moore's law aan onze zijde... Tenminste, zo denken de meeste mensen. Om één bit te ontvangen worden tegenwoordig met gemak 1000 operaties uitgevoerd, vroeger waren dat er maar een paar.

Ik heb al aangegeven dat ik denk dat dit niet goed is, op die manier verliezen we het van de wet van Moore: communicatie blijft dan ver achter bij de toekomstige behoefte. We moeten het slimmer doen. Ik kom daar nu op terug.

Mobiele communicatie

We hebben de twee principes van array signaalbewerking gezien, het kunnen versterken van gewenste signalen, en het uitnullen van ongewenste signalen. Wat kunnen we hiermee? We hebben een nieuwe dimensie gekregen, we kunnen meer dingen in parallel doen. Niet alleen in het frequentiedomein, maar ook ruimtelijk gezien. Als we bijvoorbeeld een zender hebben met twee antennes, kunnen we in twee verschillende richtingen twee verschillende

Figuur 21. Basisstation voor mobiele communicatie

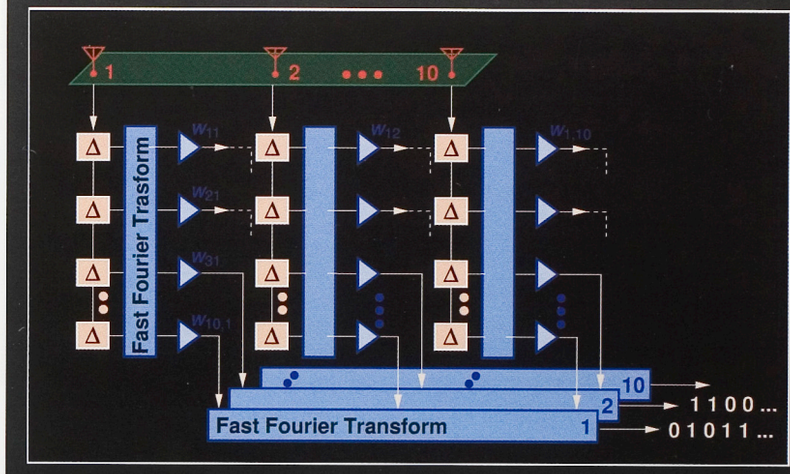
signalen uitzenden, en aan de ontvanger deze twee signalen weer van elkaar scheiden.

Hiermee hebben we de capaciteit verdubbeld met gebruikmaking van dezelfde bandbreedte. Dit is nog toekomst, maar niet zo ver weg als het lijkt, op dit moment wordt er hard gewerkt om tenminste aan de kant van het basisstation meerdere antennes op te hangen (Figuur 21).

Er is de laatste 10 jaar veel onderzoek naar de mogelijkheden hiervoor gedaan, ook door ons hier in Delft. Meestal concentreert men zich op de mogelijke capaciteitswinst: verdubbeld! En met 10 antennes vertienvoudigd! Maar aan complexiteit kost dit een factor 100, dat loopt erg snel uit de hand. Tenzij we slimmere structuren bedenken.

Alternatieve ruimte-tijd filterstructuur

Stel voor het gemak dat we 10 antennes hebben, dan wordt de capaciteit een factor 10 verhoogd. En met 10 frequentiebanden is dat een factor 100. We hebben dan 100 keer meer capaciteit. De uitdaging is om dit voor elkaar te krijgen met orde 100 keer de complexiteit. Lukt ons dat? De gebruikelijke filterstructuur vraagt om een vermenigvuldiging van een vector met een 100 bij 100 matrix: dit geeft een complexiteit van 10.000 operaties. Maar stel dat we de berekening kunnen factoriseren naar tijd en ruimte, en voor iedere factor FFTs kunnen gebruiken, dan krijgen we een complexiteit van ongeveer 650 (Figuur 22). Dat is een heel stuk minder. Het is zelfs minder dan $N \log N$, met N=100 is dat 1300. Het is dus ook minder dan het gebruik van tijd of ruimte apart.



Figuur 22. Alternatieve ruimte-tijd filterstructuur. De complexiteit is 2·10·10 log(10) = 650, niet 10000

De les is dat het gebruik van meerdere dimensies essentieel voordeel oplevert, ten minste als we ook slim gebruik kunnen maken van factorisaties en FFTs. We zullen ook de modulatieschema's hierop moeten aanpassen, daar is nog wel wat onderzoek voor nodig. Dit is het geheim om in de buurt van de wet van Moore te blijven. Het is het geheim van array signaalbewerking. Het is een "geheim" tussen aanhalingstekens, want de principes zijn wel bekend, maar ik denk dat ook binnen array signaalbewerking er weinigen zijn die zich er mee bezig houden. Mijn eigenlijke stelling is dus dat dit in de komende 10 jaar zal moeten veranderen. We kunnen het ons simpelweg niet veroorloven om duizenden bewerkingen per te ontvangen bit te gaan doen. De communicatie-revolutie zal gered moeten worden door de matrix-revolutie!

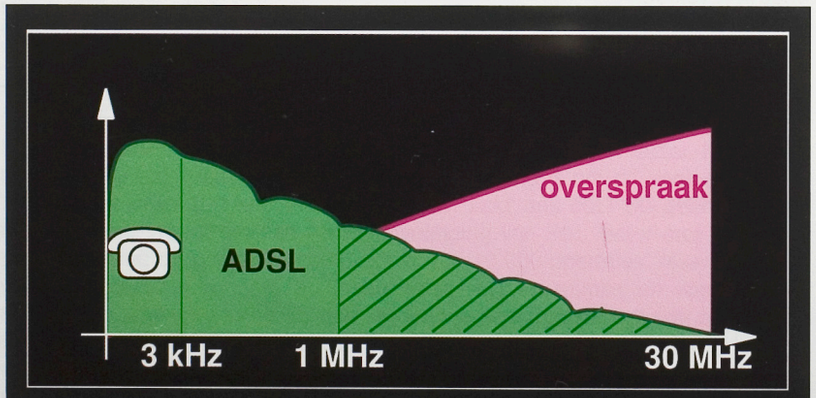
4. De toekomst van array signaalbewerking

Een overvloed aan spectrum

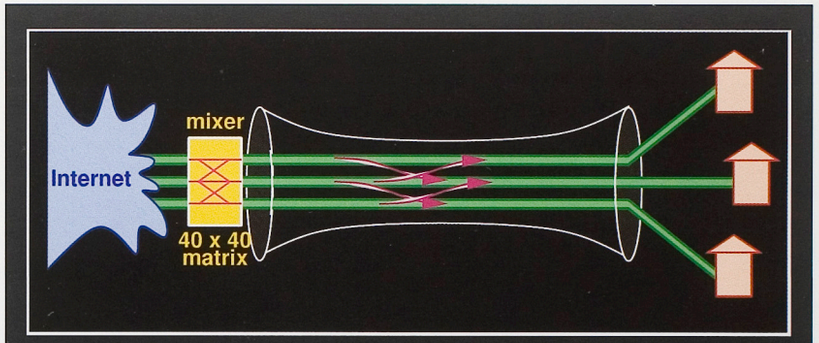
Wat brengt de toekomst? Op het gebied van draadloze communicatie is er een duidelijke verschuiving van paradigma's [7]. Momenteel is het "spectrum" (de frequentiebanden gebruikt voor communicatie) streng gereguleerd: GSM heeft een paar megahertz rond 900 MHz en 1800 MHz, UMTS rond 1900 MHz, etcetera. Wireless LAN (Wi-Fi) gebruikt een licentievrije band rond 2.4 GHz. Spectrum wordt mondjesmaat uitgedeeld en iedere toepassing heeft zijn eigen bandje waar niemand anders van gebruik kan maken. Dit gaat nu veranderen.

Stel je een voetbalstadion voor met 50.000 toeschouwers [8]. Volgens de bestaande spectrumregels mag iedereen maar één voor één praten, want anders is er storing. De praktische ervaring is natuurlijk dat het geen probleem is als iedereen door elkaar heenpraat, er is dan wel veel storing maar het menselijk oor kan signaalscheiding doen en het gewenste signaal uit de storing oppikken. Dit kan voor communicatie over korte afstanden en met de regel dat niemand schreeuwt. Het menselijk oor heeft zelfs geen moeite om de omroepinstallatie te onderdrukken. Voor draadloze communicatie staan we nu op het punt hetzelfde te kunnen.

In het nieuwe concept is spectrum "vrij", en beschikbaar in brede frequentiebanden. Iedereen kan er onder bepaalde voorwaarden gebruik van maken. Momenteel is er 500 MHz beschikbaar in een band rond 5 GHz, en in de USA is nu experimenteel toegestaan om met zeer beperkt vermogen een band van 3 tot 10 GHz te gebruiken: dat is 7 GHz, voldoende voor zeer hoge data transmissiesnelheden. Praktisch gezien is dit vrijwel onbeperkt, in de zin dat niemand nog weet hoe deze enorme bandbreedte technisch gezien ontgonnen kan worden [9]. Wel staat vast dat signaalbewerking hier een belangrijke rol gaat spelen: iedereen "praat" tegelijk, en signaalbewerking tezamen met geschikte modulatieschema's moet zorgen dat de berichten bij



Figuur 23. ADSL "band plan"



Figuur 24. Uitbreiding van ADSL met precompensatie. De capaciteit wordt verhoogd naar mogelijk 100 Mbps

de ontvanger netjes van elkaar gescheiden kunnen worden. Verder moet voorkomen worden dat zenders moeten "schreeuwen": dat kan door berichten niet rechtstreeks naar de beoogde ontvanger te sturen, maar via meerdere kleine stappen ('hops'). In plaats van één centraal basisstation wordt iedereen een knooppunt in het netwerk en geeft berichten van zijn buren door ('mobile ad hoc netwerk', MANET). Het voordeel is dat als er meer mensen zijn, het netwerk beter werkt; dit in tegenstelling tot de oude situatie met een centraal basisstation waar een beperkte vaste capaciteit met iedereen gedeeld moet worden.

Het zal duidelijk zijn dat het ontwerpen van zo'n systeem een enorme uitdaging is, de verandering in paradigma heeft een impact op alle lagen van het communicatiesysteem. Wat de verandering mogelijk maakt is de toegenomen rekenkracht van signaalprocessoren. Er wordt momenteel veel onderzoek naar gedaan. Niet al het onderzoek is realistisch: zoals gezegd denk ik dat we moeten proberen de complexiteit per ontvangen bit in de hand te houden. Met een complexiteit die groeit als N^2 komen we niet ver: we hebben de matrixrevolutie nodig!

Het opschalen van ADSL

Ik zal u nog een voorbeeld geven van wat wij doen. Iedereen van u heeft thuis ADSL. ADSL is ongeveer 1 Mbps. Wij zijn in europees verband bezig met een technologie om dit op te voeren naar 100 Mbps, een factor 100 meer, dat zou genoeg moeten zijn voor de komende 10 jaar. Hoe doen we dit?

ADSL komt bij u binnen over een telefoondraad. De eerste 3 kHz is het telefoonsignaal, dit wordt in een splitter weggefilterd. Wat overblijft is een serie frequentiebanden, tot 1 MHz (Figuur 23). Om een hogere bandbreedte te krijgen kun je twee dingen doen: meer frequentiebanden toevoegen, of meer bits per frequentieband gebruiken (daarvoor is een betere signaal-ruisverhouding nodig). Dat klinkt simpel, maar wat de mogelijkheden beperkt is dat voor hogere frequenties het gewenste signaal zwakker wordt. Tegelijk

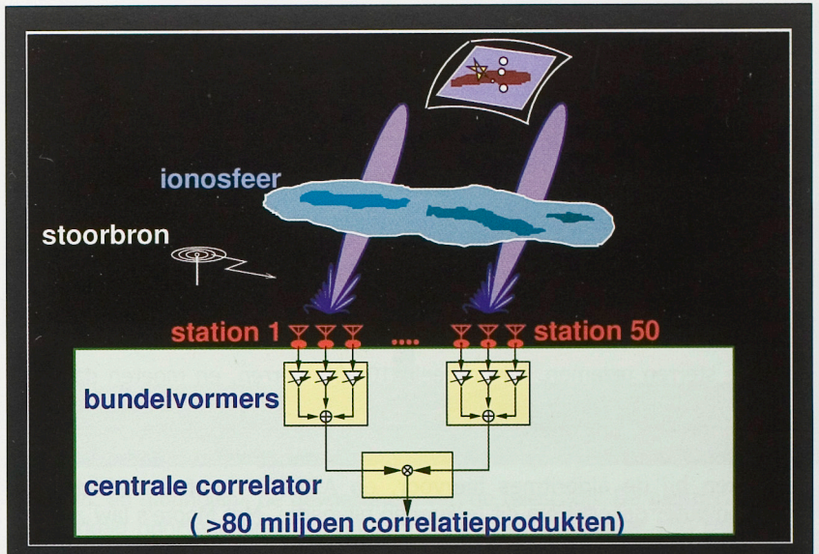
Figuur 25. LOFAR: verspreiding van de stations

krijgt u last van stoorsignalen van uw buurman, die grotendeels over dezelfde dikke telefoonkabel lopen, voor hogere frequenties geven die makkelijker overspraak.

Dit klinkt redelijk hopeloos. U ontvangt een mengsel van storing en gewenst signaal, en heeft maar een dimensie (uw koperdraadje) op het probleem op te lossen. De vraag is: waar zijn de meerdere antennes? Waar is de tweede dimensie?

Wel, de tweede dimensie zit aan de andere kant. Uzelf heeft maar 1 draadje, maar aan de andere kant, het PTT huisje, komt alles samen in een dikke bundel, met draadjes naar ieder huis in de straat. Op dit punt is het net alsof we meerdere antennes hebben, ieder draadje is equivalent aan een antenne. Wat we moeten doen is in het PTT gebouw op ieder draadje een mengsel zetten van gewenst signaal en stoorsignalen (de signalen die naar uw buren gaan), zodanig dat door overspraak bij u thuis precies alle storing uitgedoofd is. Kan dat? Om een lang verhaal kort te maken: ja dat kan (Figuur 24) [11]. We moeten wel precies alle kabels doormeten, inclusief de overspraak van ieder draadje op ieder ander draadje. Met 40 draadjes in een bundel geeft dat een matrix van 40 bij 40 elementen. De zender in het PTT gebouwtje inverteert die matrix en past hem preventief toe. Hierdoor wordt de storing onderdrukt en komt bij de ontvanger alles op z'n pootjes terecht.

Ik zal het verder niet uitwerken maar ook hier speelt complexiteit een grote rol. Een 40 bij 40 matrix lijkt niet zo groot, maar het is per frequentieband, en



Figuur 26. Schematisch overzicht van LOFAR

daarvan zijn er 5000 of meer. Ook hier zijn dus slimme matrixberekeningen nodig.

LOFAR

Verder naar de toekomst. Op het gebied van radio astronomie is Nederland momenteel aan zet, met de ontwikkeling van LOFAR, het Low Frequency Array [13]. Dit is een phased array die bestaat uit zo'n 50 stations (zeg voetbalveldjes) verspreid over Nederland en Duitsland (Figuur 25). De ontwikkeling hiervan is een initiatief van ASTRON. Ieder station is een verzameling van zo'n 200 simpele antennes die tezamen een paraboolschotel simuleren, door middel van de juiste tijdvertragingen. Totaal dus 10.000 antennes, veel meer dan waar we bij communicatie normaal aan denken.

Schematisch ziet het er als volgt uit (Figuur 26). Ieder station simuleert een paraboolschotel, met flexibele bundels die in de gewenste richting kijken, en eventuele storing proberen te onderdrukken. De informatie wordt op een centraal punt verzameld (de universiteit van Groningen) en wordt daar omgezet in plaatjes.

Een probleem is hierbij het feit dat de ionosfeer een onbekende tijdsvertraging introduceert, die ook nog eens plaats- en richtingsafhankelijk is, en verandert in de tijd. Hierdoor gaan de sterren "twinkelen" en haal je niet de scherpte die je zou willen halen.

Om dit op te lossen moet je in principe voor iedere antenne de tijdsvertraging in de richting van iedere ster gaan meten. Dit is een probleem omdat we de sterren niet een voor een aan en uit kunnen zetten: we meten de response

Figuur 27. Proefstation voor LOFAR bij Exloo

van alle sterren tezamen, bijvoorbeeld 10.000 sterren, en moeten dan een matrix van 10.000 bij 10.000 inverteren om het juiste antwoord te vinden. Een enorme klus, ook omdat dit in parallel voor 4000 frequentiebanden moet gebeuren, iedere 10 s opnieuw, we praten dan over teraflops. Wij zijn betrokken bij de algoritmes hiervoor, en ASTRON heeft inmiddels een supercomputer gekocht om dit te kunnen uitvoeren. Met Moore's law zijn die tegenwoordig heel betaalbaar.

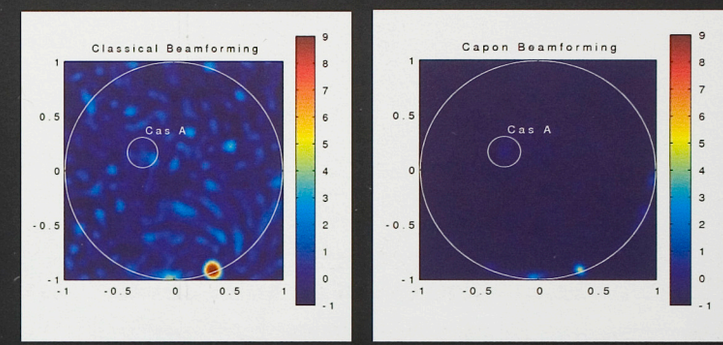
Inmiddels is in Exloo in Drente een begin gemaakt met een proefstation voor LOFAR (Figuur 27). Er staan nu zo'n 100 antennes, hele simpele, gemaakt van rioolbuizen met wat koperdraad erin. De glasvezels liggen er nog niet, iedere dag wordt daar nu nu nog een harddisk met data opgehaald.

Wij hebben intussen de eerste meetsets ontvangen. Omdat er maar 100 antennes staan (dwz 1 station) kun je geen hoge resolutie plaatjes maken, maar je kunt wel de hele hemel in een keer zien. Het linker plaatje in Figuur 28 laat zien wat je dan ziet: niet zoveel, dat komt door een stoorbron aan de horizon. Alle lichtblauwe vlekken zijn rimpelingen (zijlussen) veroorzaakt door die stoorbron, en alle sterren worden hierdoor compleet gemaskeerd. Het rechterplaatje laat zien wat je met eenvoudige array processing al kunt bereiken. De stoorbron wordt netjes geconcentreerd op de horizon, de zijlussen zijn een stuk lager, en dan komt langzaam de eerste ster naar boven: Cassiopeia A, de helderste astronomische radiobron na onze zon [12].

Dit is zo ongeveer wat je kunt halen met 1 station, 1 frequentieband, en 10 s aan data, en geen calibratie. We zijn nu aan het kijken naar calibratie voor het hele array, maar daar heb ik nog geen plaatjes van. Ik ben bang dat we daar dan toch die supercomputer voor nodig gaan hebben.

Square Kilometer Array

Verder in de toekomst: er wordt alweer gedacht aan de volgende generatie telescopen: de Square Kilometer Array, die rond 2015 gereed zou moeten zijn [14]. De ontwikkeling van radiotelescopen in de laatste 50 jaar laat zien dat iedere 10 jaar het ontvangend oppervlak ook vertienvoudigd is, en daarom is gesteld dat voor SKA een ontvangend oppervlak van een miljoen vierkante meter nodig is. Dit oppervlak bepaalt het vermogen dat opgevangen kan



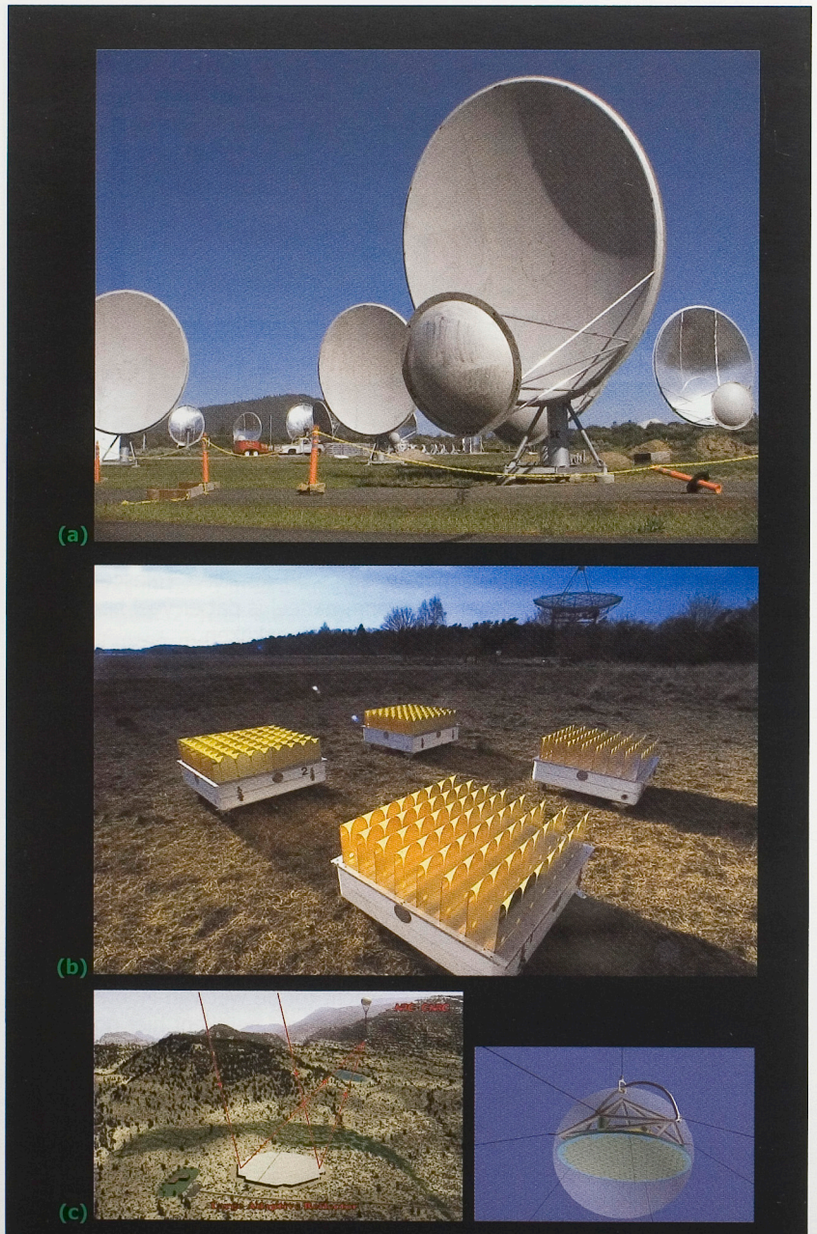
Figuur 28. Snapshot door het LOFAR proefstation. (a) Klassieke beeldvormingstechniek; (b) techniek uit de array signaalbewerking

worden, en uiteindelijk de gevoeligheid van de telescoop. De plannen worden ontwikkeld door een breed internationaal consortium, en er zijn wereldwijd diverse concepten in de maak. Je kunt proberen het te doen met een zeer groot aantal veredelde satellietschotels, zoals bij de Allen Telescope Array (Figuur 29(a)) [15]. Dat is relatief goedkoop, nadeel is dat er veel bewegende onderdelen zijn, en je kunt maar een kant tegelijk opkijken.

De concept-oplossing van ASTRON is een adaptief phased array (Figuur 29(b)) [16]. Hierbij zijn geen bewegende onderdelen, en kun je in parallel meerdere kanten opkijken. Al een aantal jaren geleden is een eerste versie gebouwd van vier keer 1 vierkante meter, kosten 200.000 Euro. Keer 1 miljoen is dit onbetaalbaar.

Er zijn nog andere concepten, zoals een beweegbare reflector op de grond, met de ontvanger in een ballon erboven (Figuur 29(c)). Het is een interessant regeltechniek-probleem om die ontvanger precies op de goede plaats te houden, zelfs als het waait.

Andere landen hebben andere oplossingen, maar allemaal even duur. Voordat er iets gebouwd wordt wordt gewacht tot de wet van Moore dit over 10 jaar wel betaalbaar gemaakt heeft. ("Betaalbaar" betekent hier nog steeds 1 à 2 miljard euro.)



Figuur 29. Testopstellingen en concepten voor SKA: (a) Allen Telescope Array (California); (b) Adaptive Array concept, ASTRON (Nederland); (c) Large Adaptive Reflector (Canada)

5. Conclusie

Ik ga afronden. Mijn vraagstelling was: redt signaalbewerking het? Kunnen we de wet van Moore bijbenen? Mijn conclusies zijn als volgt. (1) Communicatie is de bottleneck. In eerste instantie is het een hopeloos probleem: de hoeveelheid informatie die we willen communiceren groeit sneller dan de communicatiesnelheid, mogelijk gemaakt door signaalbewerking. (2) Het probleem is niet het spectrum, daarvan is er genoeg, maar om het te ontginnen zijn steeds meer bewerkingen nodig per ontvangen bit. Hoewel de communicatiesnelheden toenemen, lijkt het er dus op dat we in de toekomst steeds langer moeten wachten op onze data (films of wat we ook in de toekomst bedenken om de bandbreedte te gebruiken, bijvoorbeeld 3D telepresence). (3) Zoals gezegd ligt de oplossing in het meerdimensionaal denken, slimme factorisaties waardoor de complexiteit bijna lineair groeit met de bandbreedte. Dat is de echte uitdaging. (4) Parallelisme betekent ook veel signalen in parallel versturen, hierdoor is signaalscheiding en storingsonderdrukking nodig. Het interessante is dat we dezelfde technieken kunnen gebruiken voor mobiele communicatie, het opschalen van ADSL, en het kijken naar de sterrenhemel.

Voor het vakgebied is de wet van Moore natuurlijk geen probleem maar een zegen: het belang van signaalbewerking zal alleen maar groeien, het is hier dat de bottleneck zit, en de problemen zullen in de toekomst alleen maar groter worden. Zonder innovativiteit en veel parallelisme lukt het niet. Ik voorspel het vakgebied dan ook een gouden toekomst: de matrixrevolutie is nog maar net begonnen.

Terugkomend op het feit dat signaalbewerking over het algemeen onzichtbaar is: wat kunnen we hieraan doen? Het bestuur van de afdeling Signaalbewerking van de IEEE heeft zich dat een jaar geleden ook afgevraagd en kwam als eerste met het lumineuze idee om het vakgebied dan misschien maar om te dopen in "Information Systems Technology". Dat is een naam die velen niet zo heel veel zegt, en gelukkig is dat idee dan ook weer verlaten, net zoals onze faculteit inmiddels ook weer gewoon Elektrotechniek, Wiskunde en Informatica heet.

Als alternatief zouden we Hollywood kunnen aansporen om wat meer reclame voor ons te maken: de films van tegenwoordig zijn tenslotte zonder signaalbewerking niet meer te maken, niet alleen als decor voor "Contact" of "De ontdekking van de hemel", maar denk bijvoorbeeld ook aan alle computeranimaties, en verder speelt in veel films signaalbewerking en communicatietechnologie ook een grote rol.

In navolging van Intel zouden we dan overal stickers op kunnen plakken: "DSP Inside". We kunnen hiervoor onze promovendi inzetten als "promotieteam", in de andere zin van het woord. De eerste duizend petjes zijn inmiddels uitgedeeld (Figuur 30).

Figuur 30. "Signal Processing Inside" promotieteam

Slotwoord

Ik ben aan het eind gekomen van deze lezing. Ik wil op dit punt nog enige mensen bedanken; ik zal me hier echt beperken. Ten eerste bedank ik mijn ouders, zij stonden aan het begin van mijn opvoeding en het is misschien wel hierdoor dat ik elektrotechniek ben gaan studeren. Ten tweede bedank ik mijn "doctorvater", leermeester en mentor, prof. Patrick Dewilde. Hij heeft op diverse manieren een grote invloed op mijn ontwikkeling gehad, en laat niet na iedere promovendus te herinneren "plus êtes en vous", er zit meer in u. Dat blijkt dan maar weer.

Tot slot wil ik iedereen bedanken die vandaag speciaal voor deze lezing naar de aula getogen is, en in het bijzonder de Roemeense delegatie, mijn schoonvader prof. Vasile Pacurar, en zijn neef Sorin Potra.

Dank u wel.

Ik heb gezegd.

6. Referenties

[1] G. Moore, "No exponential is forever... but we can delay 'forever'", ISSCC, 2 Oct. 2003. http://www.intel.com/research/silicon/mooreslaw

[2] G. Gilder, "The coming software shift", Forbes ASAP, 28 Aug. 1995

[3] L.G. Roberts, "Beyond Moore's law: Internet growth trends", IEEE Computer, **33**, 1, Jan. 2000, p. 117-119

[4] J.W. Cooley, J.W. Tukey, "An algorithm for the machine calculation of complex Fourier series", Math. Comp. **19**, 1965, p. 297-301

[5] J. Gray, "What's next?: A dozen information-technology research goals", Journal of the ACM, **50**, 1, 2003, p. 41-47

[6] J. Rabaey, "Silicon platforms for the next generation wireless systems — what role does reconfigurable hardware play?", Lecture notes in Computer Science, **1896**, 2000, p. 277-285. ISBN 3-540-67899-9

[7] A. Kling, "Moore's bailiff", Tech Central Station, 17 July 2002. http://www.techcentralstation.com/071702C.html

[8] T.J. Shephard, "Spectrum policy task force seeks public comment on issues related to Commission's spectrum policies—comments by Timothy J. Shepard", ET Docket no. 02-135, 6 June 2002

[9] B. Broderson, "A conversation with Teresa Meng", ACM Queue, **2**, 1, Mar. 2004

[10] N. Fourikis, "Phased array-based systems and applications", Wiley interscience, 1997. ISBN 0-471-01212-2

[11] G. Ginis, J.M. Cioffi, "Vectored transmission for digital subscriber line systems", IEEE J. Selected Areas in Comm., **20**, June 2002, p. 1085-1104

[12] S. van der Tol, A.J. van der Veen, "Application of Robust Capon Beamforming to radio astronomical imaging", IEEE ICASSP, Mar. 2005, p. IV 873-876

[13] LOFAR website: http://www.lofar.org, http://www.lofar.nl

[14] SKA website: http://www.skatelescope.org

[15] Allen Telescope Array: http://astron.berkeley.edu/ral/ata/

[16] ASTRON website: http://www.astron.nl

7. Biografie

Prof.dr.ir. Alle-Jan van der Veen is op 18 mei 1966 geboren in Voorschoten. Hij behaalde de ingenieurstitel in de elektrotechniek aan de TU Delft (cum laude) in 1988, en de doctorsgraad (cum laude) in 1993. Hierna verbleef hij een jaar als postdoc aan Stanford University, waar hij nog diverse zomers terugkeerde. Hij verbleef ook enige maanden bij ENST Parijs, en Chalmers University in Goteborg, Zweden. Hij werd benoemd tot universitair hoofddocent aan de TU Delft in 1999, en tot Antoni van Leeuwenhoekhoogleraar in december 2001.

Hij heeft twee IEEE Signal Processing Society (SPS) Young Author Best Paper awards gewonnen (1994, 1997). Hij was voorzitter van de IEEE SPS Signal Processing for Communications Technical Committee, Associate Editor van IEEE Transactions on Signal Processing, en is momenteel Editor-in-Chief van de IEEE Signal Processing Letters. Hij is Fellow van de IEEE en won in 2003 een NWO/STW VICI beurs.

Zijn onderzoek is gericht op signaalbewerking voor communicatie, array signaalbewerking, en meer algemeen systeemtheorie toegepast op signaalbewerking. Hij heeft momenteel projecten op het gebied van UWB, ADSL, Beyond-3G, ad hoc netwerken, en LOFAR.