

Studieblad

nr. 2 • 48e jaargang • februari 1993

Themanummer **Audio-codering**



ptt telecom

Studieblad

Inhoud

PTT Telecom Studieblad is een uitgave van PTT Telecom Opleidingen (OT)

Hoofredacteur

drs. Y.M. van der Veen

Redactie

E.J. Boessenkool,

ing. N. Herwig,

A. Welling

Tekstredactie

drs. A. Kok (Info Transfer)

Secretariaat

mw. F. Stulp-Huttema

tel. 050-853732

Correspondentie-adres

PTT Telecom Opleidings-

centrum, Postbus 13000,

9700 EA Groningen

Telefax 050-266355; telex

77053; Memocom NPS 1452

Abonnement

f 18,— per jaar. Voor niet-

PTT-ers f 90,— per jaar.

Verschijnt 11 x per jaar (dubbelnummers voorbehouden)

Vormgeving

Studio Dorèl, Groningen

Druk

Ten Brink, Meppel

Fotografie

Perry Hokke

PTT Research, Thom Segers/
Fred de Jager

Philips

PTT Telecom (KRV)

Tekeningen

Sieger Zuidersma

Pagina 69 **Audio-codering: daar zit muziek in**
*drs. Y.M. van der Veen, ir. J.A. Stemerding,
dr. ing. J.G. Beerends, E.J.E. Koenderink,
P. Morley B.A., B.E., M.Sc., ir. P.T. Pont*

Pagina 109 **Studieblad Kort**

© PTT Telecom

Overname van (gedeelten van)

artikelen alleen na vooraf

verkregen toestemming van de

redactie en met uitdrukkelijke

bronvermelding: auteur, titel,

Studieblad PTT Telecom en

aflevering

ISSN 0165 8913

Bij de omslagfoto

Foto: PTT Research, Thom Segers/Fred de Jager.

In de loop van dit jaar gaat PTT Telecom in dertig grote steden (steden met een kort netnummer) met het Integrated Services Digital Network (ISDN) van start. Medio 1995 zal een aansluiting op dit veelzijdige, digitale basisnetwerk overal in Nederland beschikbaar zijn. Naast diensten als digitale telefonie, hoge-snelheid fax (groep-4), datacommunicatie en beeldtelefonie, zal dit Europese telecommunicatienetwerk ook een kwalitatief hoogwaardig audio-transport kunnen verzorgen. Anders dan bij de huidige telecommunicatienetten (datanet, telexnet en telefoonnet) wordt de specialisatie daarbij niet meer in het netwerk gezocht, maar aan de gebruikerskant. Voor het hoge-kwaliteit audio-transport ten behoeve van met name omroepen betekent dit dat er overal in het land op de live-uitzendlocaties speciale audio-apparatuur nodig zal zijn. Audio-apparatuur die de omvang van de bitstroom met behoud van de veel geroemde CD-kwaliteit zodanig kan terugbrengen, dat transport over 2×64 kbit/s ISDN-aansluitingen mogelijk wordt.

Tot voor kort werd een dergelijk sterk terugdringen van de bitstroom met behoud van de geluidskwaliteit niet mogelijk geacht. Onderzoekers bij PTT Research twijfelden er echter niet aan of dit moest haalbaar zijn. In de tachtiger jaren startten zij daarom het onderzoek naar een audio-codeersysteem dat hoge-kwaliteit audio-transport over 64 kbit/s kanalen kan bewerkstelligen.

In het voor u liggende themanummer vindt u een beschrijving van dit door PTT Research ontwikkelde audio-codeersysteem. Daarnaast zal worden stilgestaan bij de manier waarop een van oorsprong analoog geluidssignaal wordt gedigitaliseerd. Veel aandacht is er vanzelfsprekend voor de problemen die overwonnen moesten worden om de bitstroom tot 64 kbit/s te kunnen reduceren (N.B. De omvang van de bitstroom op een CD is ruim tien keer zo groot, namelijk 705,6 kbit/s). Als uitgangspunt is daarbij niet gekozen voor de in de audio-wereld vertrouwde technische kwaliteitsmaten, maar voor datgene wat wij mensen feitelijk horen. Bepaalde fundamentele eigenschappen van het menselijk gehoor zullen in dat kader worden toegelicht. Ook komt in dit nummer van het Studieblad het huidige en toekomstige gebruik van hoge-kwaliteit audio-diensten aan

de orde. De manier waarop PTT Telecom omroepen nu en in de toekomst bij het maken van live-radiouitzendingen kan ondersteunen, neemt hierbij een centrale plaats in.

Maar niet alleen laten we u schriftelijk kennismaken met de nieuwe audio-technologie. Bij dit nummer van het Studieblad heeft u ongetwijfeld ook de CD gevonden, waarop veel van de hiervoor genoemde zaken worden gedemonstreerd. Hopelijk bent u net als de redactie enthousiast over deze unieke manier van kennismaken met wat straks via het ISDN mogelijk wordt. De voor de technische demonstraties gebruikte muziekfragmenten kunt u op de tracks [35] tot en met [39] van de Studieblad-CD bovendien in hun geheel beluisteren. Naast een educatief deel bevat de CD dus ook aangename ontspanning in de vorm van een half uurtje afwisselende muziek, uiteenlopend van een live opgenomen bluesnummer tot en met een klassiek strijkkwartet. Hoe de CD precies is ingedeeld kunt u lezen in het CD-boekje dat in het hart van het Studieblad is opgenomen.

Om het technische deel van de CD tot zijn recht te laten komen, verdient het aanbeveling niet direct naar uw CD-speler te lopen, maar eerst rustig de tijd te nemen om het artikel *Audio-codering: daar zit muziek in* te lezen. Voor een goed begrip van de demonstraties is de kennis uit het artikel onmisbaar.

Ysbrand van der Veen,
Jan Stemerding,
John Beerends,
Ernest Koenderink,
Patrick Morley,
Peter Pont

Bij live-uitzendingen van radioprogramma's heeft PTT Telecom een belangrijke ondersteunende taak. Naast telefonie en datacommunicatie is ook het transport van hoge-kwaliteit audio een activiteit van het bedrijf. Om de diverse uitzendlocaties overal in het land met de Hilversumse radiostudio's te verbinden, maakt PTT Telecom op dit moment nog van het gewone telefoonnet gebruik. Dankzij speciale analoge transmissietechnieken is de radioluisteraar daarbij altijd van kwalitatief hoogwaardige live-uitzendingen verzekerd. De gehoormatige kwaliteit van de muzieklijnen, de aanvoerroutes naar de radiostudio's, kan binnenkort verder worden verhoogd door nieuwe digitale technieken toe te passen, en dat bovendien tegen lage kosten. Kort gezegd komt het erop neer dat het van oorsprong analoge geluidssignaal al op de uitzendlocatie digitaal wordt bewerkt en gecodeerd. Het signaal is daarna betrekkelijk eenvoudig met CD-kwaliteit richting Hilversum te transporteren via standaard 2 Mbit/s huurlijnen. In een wat verdere toekomst zal dit hoge-kwaliteit audio-transport bovendien mogelijk worden over lijnen met een heel wat beperktere transmissiecapaciteit. Te denken valt in dit verband vooral aan audio-verbindingen via het openbaar geschakelde ISDN-netwerk (2×64 kbit/s). Behalve direct beschikbare muzieklijnen – ieder ISDN-telefoonstopcontact is tevens een 'muziekstopcontact' – kan dit hoge-kwaliteit audio-transport via het Integrated Services Digital Network ook allerlei andere voordelen bieden. Zo liggen er aantrekkelijke diensten in het verschiet, uiteenlopend van 'gesloten muzieknetten' voor supermarkketens tot ISDN-concertlijnen voor bijvoorbeeld thuisabonnementen op de 'Concertserie A' van het Concertgebouworkest of een reeks concerten in het kader van Madonna's Europese tournee.

Als gevolg van het grote succes van de Compact Disc (CD) is het digitaal opnemen en opslaan van muziek in een stroomversnelling geraakt. Dankzij de CD beschikt een breed publiek nu over geluidsopnames van een bijzonder hoge kwaliteit. Het kwaliteitsbewustzijn van het publiek is hierdoor toegenomen, waarmee digitale audio ook voor PTT sterk aan belang wint. Zo zullen digitale technieken bij het hoge-

kwaliteit audio-transport ten behoeve van live-radiouitzendingen steeds belangrijker gaan worden.

Internationaal zijn inmiddels verschillende codeertechnieken ontwikkeld waarmee digitale transmissie van hoge-kwaliteit audio via 2 Mbit/s verbindingen mogelijk is. Zonder allerlei bijzondere voorzieningen in het netwerk, die momenteel voor de analoge muzieklijnen in het telefoonnet nodig zijn, zal met deze digitale technologie op een snelle en eenvoudige manier in de groeiende vraag van omroepen naar audio-verbindingen kunnen worden voorzien. De noodzakelijke codeerapparatuur komt daarbij in vaste opstelling aan de gebruikerskant te staan, zodat de verbindingswegen met Hilversum standaard 2 Mbit/s huurlijnen kunnen zijn.

Deze nieuwe technologie over huurlijnen, vaste verbindingen tussen een punt *A* en een punt *B*, is om economische redenen echter alleen aantrekkelijk voor locaties waar vandaan met grote regelmaat radioprogramma's worden verzorgd. Om ook minder frequent gebruikte locaties zoals een sporthal of vergadercentrum snel en eenvoudig aan de radiostudio's te kunnen koppelen, wordt momenteel hard gewerkt aan de ontwikkeling van een codeertechnologie die de omvang van het oorspronkelijke signaal sterk kan reduceren. Bij voorkeur zelfs zodanig dat hoge-kwaliteit audio via het internationale, geschakelde ISDN-netwerk mogelijk zal zijn. In dit netwerk staan voor het geluidstransport 2 kanalen van 64 kbit/s ter beschikking. Door per stereokanaal binnen deze relatief beperkte transportcapaciteit te blijven zijn bijzonder flexibele audio-verbindingen te realiseren. Uiteraard maakt het daarbij voor het ISDN, het Integrated Services Digital Network, niets uit of de 'eentjes' en 'nulletjes' die worden verzonden nu data, spraak, tekst, beeld of hoge-kwaliteit audio representeren.

Realiseren we ons dat in veel Europese landen momenteel hard wordt gewerkt aan de invoering van ISDN en dat in Nederland halverwege dit jaar ISDN-aansluitingen tegen een aantrekkelijk tarief in de dertig grootste steden beschikbaar zullen zijn (medio 1995 ook landelijk), dan zal het waarschijnlijk niemand verbazen dat geluid van CD-kwaliteit over 2×64 kbit/s verbindingen op dit moment sterk in de belangstelling staat¹.

In dit artikel zal op al deze ontwikkelingen op het gebied van de audio-codering uitgebreid worden ingegaan. De huidige wijze van hoge-kwaliteit audio-transport via het telefoonnet

¹ Overigens is het onderzoek naar dergelijke ingrijpende vormen van datareductie ook om tal van andere redenen interessant. Eén van die redenen is bijvoorbeeld dat hiermee zeer verregaande vormen van miniaturisatie mogelijk zullen zijn. Op een Compact Disc met de grootte van een schijfje citroen zal bijvoorbeeld gemakkelijk een uur muziek opgeslagen kunnen worden, af te spelen op een CD-Walkman ter grootte van een pakje sigaretten. Dat dit geen verre toekomstmuziek meer is, blijkt wel uit de kortgeleden geïntroduceerde Sony Minidisc.

komt eveneens aan de orde. Echt technisch georiënteerde Studiebladlezers zullen in de verdiepingsstof aan het slot van dit artikel bovendien een diepgaande beschrijving vinden van de door PTT Research ontwikkelde codeermethode ATC, Adaptieve Transformatie Codering. Met deze manier van coderen is PTT Research erin geslaagd om binnen de smalle marges van 2×64 kbit/s toch hoge-kwaliteit audio te realiseren. De bij dit nummer van het Studieblad gevoegde Compact Disc laat u dat op de tracks 31–34 overtuigend horen.

Overigens verdient het voor een goed begrip van de demo's 1–30 aanbeveling om eerst rustig door de hele tekst heen te lopen. Het luisterplezier zal er straks aanmerkelijk door worden vergroot.

Geluid van CD-kwaliteit over standaard $n \times 64$ kbit/s verbindingen

Om uit de verschillende nu en in de toekomst op de markt verkrijgbare systemen voor audio-codering de hoorbaar beste te kunnen kiezen, heeft PTT Research in opdracht van PTT Telecom een techniek ontwikkeld waarmee ook zonder peperdure luisterpanels dit soort codeersystemen op hun gehoor-matige kwaliteit te beoordelen is². In eerste instantie ging het daarbij om onderzoek naar audio-codeersystemen ten behoeve van geluidstransport over 2 Mbit/s verbindingen. Uit het onderzoek in de tachtiger jaren bij PTT Research naar deze codeersystemen kwam echter tamelijk snel naar voren dat de omvang van de bitstream met behoud van de gehoor-matige kwaliteit sterk terug te dringen moest zijn, zelfs aanzienlijk verder dan tot voor kort werd verondersteld. Reden voor de onderzoekers op het Neher Laboratorium om een codeertechniek te ontwikkelen die met 2×64 kbit/s verbindingen kan volstaan, uiteraard zonder dat daarbij hoorbare kwaliteitsverliezen optreden. In de verdiepingsstof komt deze techniek, de zogenaamde Adaptieve Transformatie Codering (ATC), meer uitgebreid aan de orde. De klinkende resultaten kunnen op de tracks 31–34 van de demonstratie-CD worden beluisterd, waarbij ook de veelbelovende vooruitzichten op hoge-kwaliteit audio over 2×32 kbit/s voor het voetlicht worden gebracht.

Internationaal werkt men inmiddels in diverse laboratoria aan de ontwikkeling van een met ATC vergelijkbaar type audio-

² J.G. Beerends & J.A. Stemerdink, A perceptual audio quality measure based on a psychoacoustic sound representation, in: *Journal of the Audio Engineering Society*, 40: 963-978, december 1992 (equivalent met PTT Research publ. NT-IN-92-756).

³ MPEG = Motion Pictures Expert Group. Op de door het ISO (International Standards Organisation) erkende MPEG-standaard voor audio-codecs zal in het vervolg van dit artikel nog nader worden ingegaan.

⁴ Belangrijke toepassingsmogelijkheden van ISDN zijn op dit moment met name de snelle, volledig digitale Groep-4 fax, 64 kbit/s beeldtelefonie en tal van zaken op het gebied van datacommunicatie. Zie hiervoor onder andere:

N. Korving en Y.M. van der Veen, *De successtory van de fax*, PTT Telecom Studieblad, oktober 1992, pp. 509-542, R. Wijbrands, *ISDN: digitale toegangsweg tot informatie en informatiebestanden*, PTT Telecom Studieblad, thema-nummer ISDN, juni 1992, pp. 313-335.

⁵ Kort gezegd heeft het AVVC dus twee belangrijke taken: contributie en distributie. De netten voor radio en TV, bijv. Nederland 2 en Radio 3, worden ook wel distributienetten genoemd. Het muzieknet – de (deels tijdelijke) verbindingen voor radiouitzendingen op locatie – is een contributienet.

codec. Een audio-codec die ook via $n \times 64$ kbit/s verbindingen hoge-kwaliteit audio-transport kan realiseren, maar die tevens aansluiting dient te vinden bij de onder meer in de consumentenelektronica veel gebruikte MPEG-standaard³.

Aan deze standaardmethode voor $n \times 64$ kbit/s audio wordt in onderzoekscentra nog altijd intensief gewerkt. De tot dusver behaalde onderzoeksresultaten zijn echter dermate veelbelovend dat de met ATC vergelijkbare, hoge-kwaliteits MPEG-codec niet lang meer op zich zal laten wachten. Stereogeluid van CD-kwaliteit via het internationale, geschakelde ISDN-netwerk zal in de nabije toekomst dan ook breed beschikbaar kunnen komen. Vanzelfsprekend is deze vernieuwing op audio-gebied behalve voor bijvoorbeeld omroepen ook voor de ontwikkeling van het ISDN als geheel van belang. Nieuwe aantrekkelijke diensten komen immers in zicht, waardoor het aantal in een ISDN-aansluiting geïnteresseerde telecommunicatiegebruikers naar verwachting aanzienlijk zal groeien⁴.

Muzieklijnen nu

Iedere dag weer kan kijkend en luisterend Nederland uit een gigantisch aanbod van radio- en televisieprogramma's kiezen. Veel van die programma's worden op locatie opgenomen en bovendien vaak rechtstreeks uitgezonden. Voor de televisieuitzendingen wordt daarbij veelal van tijdelijke straalverbindingen gebruik gemaakt. Gaat het om radio-uitzendingen dan zorgen tijdelijke verbindingen in het telefoonnet (de zogenaamde muzieklijnen) ervoor dat de signalen waar ook vandaan – een dorpscafé in Zeeland, de Tweede Kamer in Den Haag of muziekcentrum Vredenburg te Utrecht – steeds op één centraal punt terecht komen: het PTT Audio en Video Verbindingen Centrum (AVVC) te Hilversum.

Vanuit het AVVC worden deze signalen vervolgens naar de verschillende radiostudio's gezonden, waar ze eventueel kunnen worden bewerkt. Ten slotte worden de diverse programma's wederom via het AVVC naar alle Nederlandse radio- en televisiezenders getransporteerd. Klaar voor uitzending via de ether.

Het AVVC is in het licht van dit alles met recht te beschouwen als het schakelende hart van de omroepdiensten van PTT Telecom⁵.



Bij de audio-verbindingen via het telefoonnet (muzieklijnen) behoren ook altijd één of meerdere 'spreeklijnen'. De programmamaker op locatie onderhoudt via deze spreeklijnen voortdurend contact met de Hilversumse studio. Uiteraard zonder dat de radioluisteraar daar iets van merkt. Het signaal dat voor uitzending bestemd is, wordt namelijk – al dan niet in stereo – via bovengenoemde muzieklijnen apart naar de radiostudio getransporteerd.

Het hoge-kwaliteit audio-signaal wordt daartoe via het lokale telefoonnet direct vanaf de opname-locatie naar de dichtstbijzijnde nummercentrale geleid, om vervolgens over twee aderporen in de knooppuntcentrale terecht te komen. Via draaggolfkabels belandt het signaal vervolgens in een districtsversterkerstation om uiteindelijk naar het AVVC te worden doorgestuurd⁶. Dankzij een speciale truc, de zogenaamde fantoomschakeling, ondervindt het gewone telefoonverkeer van dit muzieknets weinig hinder⁷.

Mochten zich tijdens de uitzending onverhoopt problemen voordoen met de speciale hoge-kwaliteit audio-verbindingen, dan kan het in de praktijk overigens nog wel eens gebeuren dat er tijdelijk op de spreeklijnen wordt teruggevallen. Lijnen met een gewone telefoonkwaliteit (3 kHz spraakband), zoals u tijdens live-verslagen van grote internationale sportevenementen ongetwijfeld wel eens heeft gemerkt.

▲ Foto 1

Het Audio en Video Verbindingen Centrum (AVVC) van PTT Telecom te Hilversum is het schakelende hart van de omroepdiensten van PTT.

⁶ Het eerste deel van het muzieknets (t/m de districtsversterkerstations) wordt voornamelijk op ad-hoc basis gerealiseerd. Voor de verbindingen tussen de versterkerstations en (uiteindelijk) het AVVC kan veelal van permanente lijnen gebruik worden gemaakt.

⁷ Zie voor deze schakeling: E. Boessenkool, H. Koene, Y.M. van der Veen, *Elementaire kennis – Telecommunicatie* (dl. 9: soorten verbindingen), PTT Telecom Studieblad, oktober 1991, pp. 572-589.

Analoog of digitaal?

Aan de analoge technologie die PTT Telecom momenteel voor de muzieklijnen gebruikt, kleeft een aantal bezwaren.

- Het analoge contributienet c.q. het telefoonnet is gevoelig voor overspraak, dat wil zeggen dat de signalen in nabijgelegen koperaders elkaar kunnen beïnvloeden. (Dit betreft vooral storing door asymmetrische signaleringen.)
- Er zijn aan het tot stand brengen van hoge-kwaliteit audio-verbindingen relatief veel inmeetwerkzaamheden verbonden (arbeidsintensief).
- Onderhoud en beschikbaarheid van de analoge apparatuur zullen in de nabije toekomst problemen kunnen geven.
- De kwaliteit van de verbindingen kan, met name wanneer het gaat om live-muziekuitzendingen met een korte voorbereidingstijd, niet wedijveren met de veelgeroemde CD-kwaliteit.

⁸ Op een aantal kabelnetten wordt momenteel al digitale radio doorgegeven. Het gaat hierbij om Digitale Satelliet Radio (DSR), waarvoor vanwege de grote bitstream een brede frequentieband noodzakelijk is. Vanwege het schaars beschikbare spectrum is uitzending van dit breedbandige signaal via de normale FM-kanalen niet mogelijk, vandaar dat voor de distributie van deze uitzendingen naar de satelliet is uitgeweken. Codeersystemen als ATC en de lage bitrate MPEG audio-codec (layer III) zullen het in de toekomst wellicht mogelijk maken dat digitale radio ook via normale aardgebonden zender-netten aangeboden gaat worden.

Op dit moment is de kwaliteit van de via de FM-band analoog uitgezonden radioprogramma's én die van de muzieklijnen vergelijkbaar. In de keten zijn dus alle schakels met elkaar in overeenstemming. Het vierde bezwaar dat het huidige analoge muziektransport zich kwalitatief niet kan meten met de veelgeroemde CD-kwaliteit, zal daarom pas gaan spelen wanneer in Nederland op grote schaal op digitale radiouitzendingen wordt overgegaan⁸.

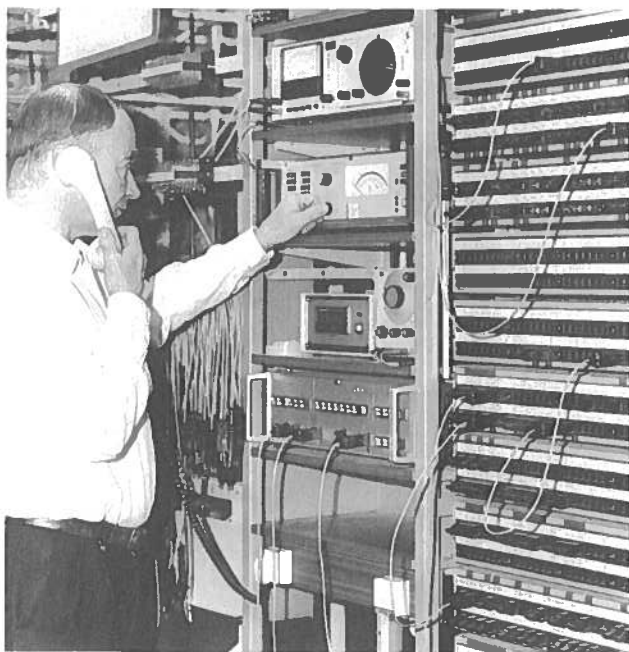
Inmiddels wel hoogst urgent zijn de drie eerstgenoemde bezwaren. Zo is er als gevolg van het incidentele gebruik van de muzieklijnen op dit moment sprake van hoge inmeetkosten en dus van (te) hoge exploitatiekosten. In het directe verlengde hiervan speelt uiteraard ook het noodgedwongen moeten hanteren van levertijden. Klanten dienen een hoge kwaliteit audio-verbinding (muzieklijn) dus van tevoren aan te vragen. Daarnaast staat de overspraakgevoeligheid van het huidige audio-contributienet de internationaal gewenste invoering van een hoger zendniveau⁹ voor op het telefoonnet aangesloten (data)modems in de weg. Dit hoge zendniveau wordt wenselijk geacht om in de toekomst via het telefoonnet ook over langere afstanden onversterkt datacommunicatie te kunnen plegen.

⁹ Van -6 dBm naar 0dBm.

Alle genoemde factoren te zamen maken dat het Netwerkb-

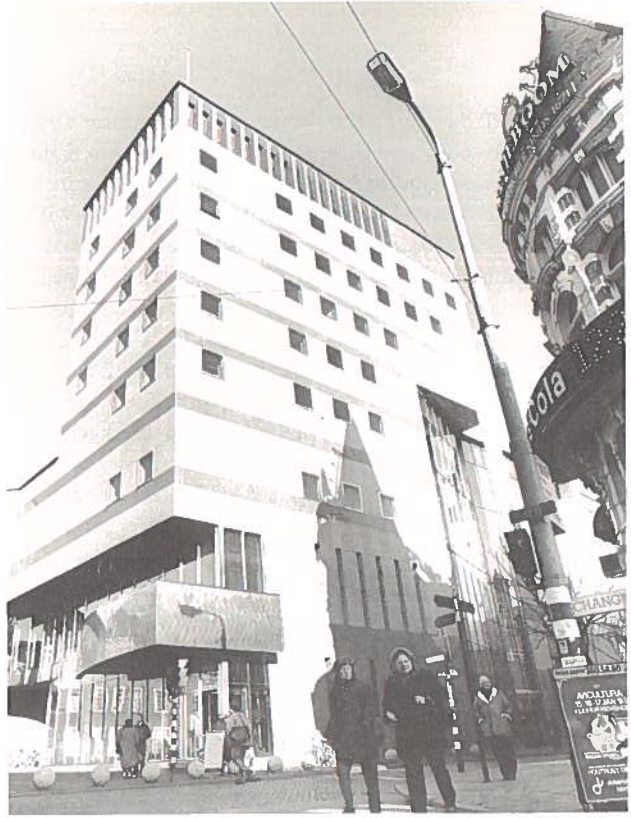
drijf van PTT Telecom op dit moment naarstig naar andere methoden van audio-contributie uitziet. Het streven is daarbij om de analoge audio-contributie zo snel mogelijk uit het telefoonnet te halen. Hangende de internationale ontwikkelingen rond ISDN en het hoge-kwaliteit audio-transport, bestaan er daarom plannen de echt incidentele uitzendlocaties voorlopig via mobiele satellietverbindingen met de radiostudio's te gaan verbinden. Locaties met een meer dan incidenteel gebruik zullen min of meer continu met het AVVC verbonden worden. Speciale apparatuur op deze (naar schatting 50) locaties dient er daarbij voor te zorgen dat het audiosignaal digitaal wordt gecodeerd, waarna het via vaste verbindingen (2 Mbit/s huurlijnen) naar het AVVC in Hilversum getransporteerd kan worden.

PTT Telecom kan voor deze, op de locatie te plaatsen, audioapparatuur binnenkort uit meerdere op de markt verkrijgbare systemen kiezen. Audio-codeersystemen, maar dan van een vorige generatie, worden momenteel al toegepast voor het digitaal coderen van het televisiegeluid in onder andere Engeland, België en de Scandinavische landen.



◀ Foto 2

Om via het telefoonnet een audioverbinding tot stand te kunnen brengen, moeten er in het telecomdistrict heel wat werkzaamheden worden verricht. Een belangrijke rol is daarbij weggelegd voor de versterkerstations die voor de versterking van het signaal en de nodige schakelingen zorgdragen. Via de verschillende versterkerstations arriveert het audio-signaal over de muzieklijnen uiteindelijk in het AVVC te Hilversum, vanwaar het naar de radiostudio's wordt getransporteerd. Op de foto een muziekschakelrek in het Telecomdistrict Den Haag.



► Foto 3

Perscentrum Nieuwspoor te Den Haag is één van de locaties die min of meer continu met het AVVC verbonden zal worden via een 2 Mbit/s verbinding.

Audio

Onder audio verstaan we in de regel geluidssignalen die het totale voor mensen hoorbare bereik bestrijken (ruim genomen van 15 Hz tot 20 kHz), ongeacht de geluidsbron die het signaal heeft voortgebracht. In de meeste gevallen zal de geluidsbron evenwel uit één of meerdere menselijke stemmen en/of muziekinstrumenten bestaan.

Voor het geluid van de menselijke stem geldt dat vrijwel alle informatie zich in de band van 50 tot 7000 Hz zal bevinden. Voor een goed verstaanbare weergave van spraak geldt dat doorgifte van het gedeelte tussen 300 en 3400 Hz al voldoende is. Vandaar dat in de telefonie gebruik wordt gemaakt van een spraakband met een bandbreedte van circa 3000 Hz. Voor de weergave van muziek is een dergelijke bandbreedte echter bij lange na niet toereikend (vergelijk track 1, 2 en 5 van de demonstratie-CD). Bovendien valt een criterium

als verstaanbaarheid voor het luisteren naar muziek niet te hanteren. Om toch van een bepaalde maatstaf uit te kunnen gaan, wordt tegenwoordig daarom veelal de eis gesteld dat alle hoorbare signaaldelen ook daadwerkelijk worden weergegeven.

Uit proeven naar het menselijke gehoor blijkt evenwel dat er voor de meeste mensen geen verschil bestaat tussen een muzieksignaal met een bandbreedte van 15 kHz en datzelfde muzieksignaal met een bandbreedte van 20 kHz. Voor de transmissie van digitale audio-signalen is wereldwijd dan ook gekozen voor een audio-bandbreedte van 15 kHz¹⁰.

Overigens zal deze gehoorgrens bij jonge mensen in de regel een stuk hoger liggen, sommigen kunnen zelfs frequenties boven 20 kHz onderscheiden. Zij zullen dus in ieder geval verschil kunnen waarnemen tussen het oorspronkelijke signaal (demo [1] op de CD) en het signaal dat voldoet aan de internationale 15 kHz standaard voor digitaal audio-transport (demo [2] op de CD).

¹⁰ CCIR rec. 606, vol. XII. De bijbehorende bemonsteringsfrequentie ligt op 32 kHz.

CD: demo bandbreedte-begrenzing [1] – [9]

Op deze tracks van de demonstratie-CD is een gedeelte van een opname van het *Budapest String Quartet* op verschillende bandbreedtes te beluisteren.

Het frequentiebereik van het menselijk gehoor bestrijkt ruim genomen een gebied van 15 Hz tot 20.000 Hz. De meeste mensen zullen het effect van een bandbreedtebegrenzing tot 15.000 Hz echter niet kunnen waarnemen. Hoe dat met uw gehoor zit, kunt u zelf beoordelen aan de hand van de demo's [1] – [6].

Het frequentiebereik van de originele opname [1] zal in principe tot 22 kHz doorlopen, alhoewel er in de regel in het bovenste deel van het spectrum niet veel signaal meer aanwezig is. In de fragmenten [2] – [6] is de bandbreedte vervolgens met behulp van een low-pass filter beperkt tot respectievelijk 15, 10, 8, 4 en 2 kHz.

Om u een indruk te geven wat er bij de bandbreedtebegrenzing precies uit het signaal verdwijnt, zijn de demonstraties [7] – [9] op de Studieblad-CD opgenomen.

[7] is het oorspronkelijke, ongefilterde muziekfragment (0-22 kHz).

8 laat horen wat er uit het signaal is verdwenen bij de bandbreedte-begrenzing tot 4 kHz.

9 maakt duidelijk wat het resultaat is van een bandbreedte-begrenzing tot 8 kHz.

De demo's 8 en 9 zijn vervaardigd met behulp van een zogenaamd high-pass filter.

Digitale audio

Onze wereld van alledag proberen te vertalen in een abstracte wereld van 'enen' en 'nullen' is een ingewikkeld proces. Immers wat wij zien, horen en voelen is niet eenvoudig in zwart/wit-waardes uit te drukken. Zo zijn de bewegingen, kleuren en grijstinten die wij zien oneindig geschakeerd, zoals wij ook in stemmen en muziekinstrumenten een oneindig aantal nuances kunnen onderscheiden.

Blijft de analoge-digitaal omzetting uitsluitend beperkt tot het digitaal registreren/opslaan van een bepaalde analoge werkelijkheid, dan slagen we er tegenwoordig steeds beter in resultaat te boeken. De Compact Disc, de CD-Interactive en de ISDN-beeldtelefoon¹¹ zijn hiervan goede voorbeelden. Alleen kost het vaak erg veel nulletjes en eentjes (en daarmee geld) om dat op een voor mensen aanvaardbaar kwaliteitsniveau te doen¹².

Veel moeilijker wordt het vervolgens wanneer we de computer aan deze digitaal vastgelegde werkelijkheid een bepaalde betekenis willen laten toekennen. Het probleem van een betrouwbare spraakherkenning zoals kortgeleden in het Studieblad geschetst, is hiervan een sprekend voorbeeld¹³. Een ander bekend probleem is dat van de handschriftherkenning, waar PTT Research ten behoeve van PTT Post eveneens intensief onderzoek naar doet.

Laten we echter terugkeren naar de audio-codering, het eigenlijke onderwerp van dit artikel. Zoals we hiervoor al zeiden zijn er enorm veel nulletjes en eentjes voor nodig om een analoge geluidssignaal met een voldoende hoge kwaliteit digitaal vast te leggen. Dat is trouwens niet zo verwonderlijk, wanneer we ons realiseren hoe complex een analoge geluidssignaal in feite is. Het signaal is immers noch in de tijd noch in de waarde die het kan aannemen beperkt. In feite is het analo-

¹¹ Aan de ISDN-beeldtelefoon heeft het Studieblad in juni 1990 een speciaal thema-nummer 'Audiovisuele communicatie' gewijd.

¹² Vergelijk hiervoor: Y.M. van der Veen, *Image processing en multimedia: optische technologie maakt van computer steeds veelzijdiger communicatiemiddel*, PTT Telecom Studieblad, november 1992, pp. 584-616.

¹³ Zie: J. Hendriks, *Praten met de computer: spraaksynthese en spraakherkenning*, PTT Telecom Studieblad, september 1992, pp. 449-466.

ge geluidssignaal oneindig nauwkeurig, iets wat te vergelijken is met de oneindige precisie waarmee de secondewijzer van een ouderwets polshorloge continu de tijd aangeeft.

Een dergelijke nauwkeurigheid is bij de vertaling in 'nullen' en 'enen' ten enen male onmogelijk. Er zullen bij de digitale registratie van een analoog geluidssignaal dus bepaalde knopen moeten worden doorgehakt. Meer specifiek betekent dit voor een digitaal audio-systeem dat er bij de analoog-digitaal omzetting (de zogenaamde A/D-conversie) in eerste instantie twee bewerkingen dienen te worden verricht: bemonsteren en kwantificeren. Bemonstering (sampling) is het omzetten van het analoge signaal (tijd-continu) in een bepaalde tijdreeks van signaalwaarden. Op regelmatige tijdstippen, bijvoorbeeld enkele duizenden malen per seconde, worden van het

▼ Afb. 1



oorspronkelijke signaal geluidsmonsters genomen. De waarde van deze monsters is op dat moment uiteraard nog oneindig nauwkeurig (waarde-continu), wel is door het audiosysteem inmiddels het mes in de tijd gezet.

Door de geluidsmonsters te kwantificeren lost het audiosysteem vervolgens ook het waardeprobleem op. Kern daarbij is dat het systeem de oorspronkelijke continue waarden dient af te ronden. Anders gezegd: bij de waardebepaling kan het aantal cijfers achter de komma niet oneindig groot zijn.

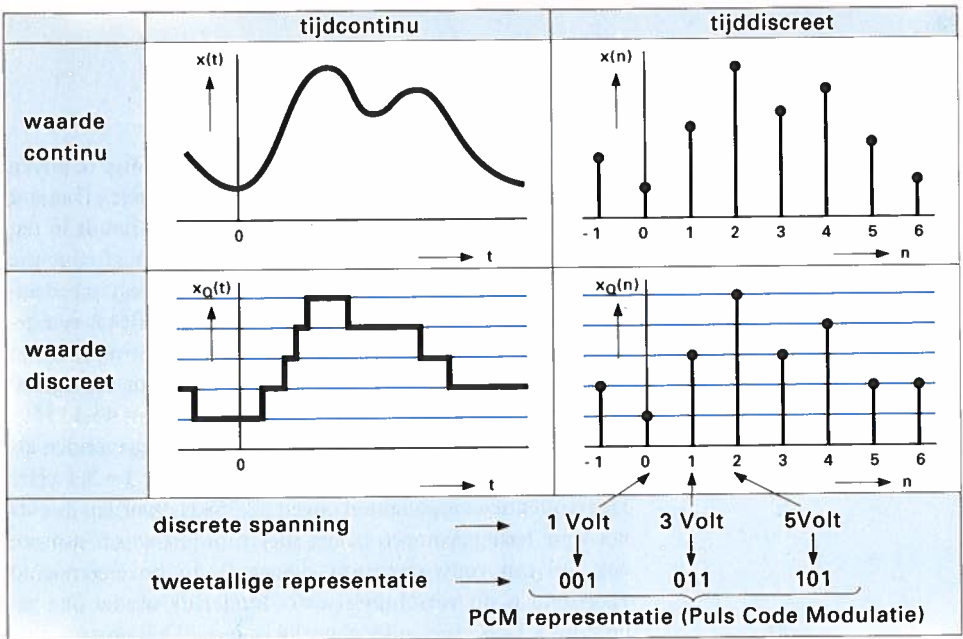
Concreet betekent dit dat de A/D-converter van het audiosysteem bij het kwantificeren bepaalde keuzes moet maken. Daarvoor staat een getallenreeks ter beschikking, die uit voortdurend wisselende combinaties van 'enen' en 'nullen' is opgebouwd (bijv. 0001, 0010, 0011, 0100, 0101 etc.). In de wereld van de digitale techniek worden dit soort getallen discrete getallen genoemd.

Het bij de A/D-converter binnenkomende geluidssignaal (tijd-continu en waarde-continu) wordt op de hierboven beschreven manier omgezet in een bitstream: een reeks discrete signaalwaarden van op bepaalde regelmatige tijdstippen genomen geluidsmonsters.

Bij de weergave zal een digitaal-analoog (D/A) omzetter deze reeks getallen vervolgens weer naar een bepaalde serie signaalwaarden (spanningen) moeten terugvertalen. Een speciaal filter (een zogenaamd reconstructie-filter) zorgt er ten slotte voor dat de signaalwaarden weer in een glad verlopend, tijd-continu signaal worden omgezet. Is de bemonsteringsfrequentie bij de analoog-digitaal omzetting voldoende hoog gekozen (bijvoorbeeld op 44,1 kHz) en is het aantal bits én daarmee de nauwkeurigheid waar de signaalmonsters mee zijn vastgelegd groot genoeg (bijvoorbeeld 16 bits), dan is uit deze bitstream met voldoende geluidskwaliteit weer een analoog signaal te reconstrueren.

Aliasing

De snelheid waarmee de A/D-converter de monsters neemt, noemen we de bemonsteringsfrequentie of sample rate. Over het algemeen wordt deze sample rate uitgedrukt in kHz of het aantal samples per seconde. Het signaal dat op een CD staat is bijvoorbeeld met een frequentie van 44,1 kHz bemonsterd, wat wil zeggen dat het signaal per (stereo)kanaal met 44.100



signaalwaarden per seconde is vastgelegd. De tijdsduur tussen twee opeenvolgende samples is dan $1/44.100 = 22,6 \mu\text{sec}$. In de digitale telefonie is een bemonsteringsfrequentie van 8 kHz gangbaar (3 kHz spraakband), terwijl voor het hoge-kwaliteit audio-transport (bandbreedte 15 kHz) een sample rate van 32 kHz gestandaardiseerd is.

▲ Afb. 2

Bij de weergave wordt met behulp van een reconstructie-filter uit de discrete signaalwaarden weer een tijd-continu signaal gereconstrueerd. Tussen de opeenvolgende samples zal dit speciale filter een soort interpolatie uitvoeren, waardoor het uitgangssignaal van het audio-systeem gehoormatig weer met het oorspronkelijke (analoge) signaal overeen zal stemmen.

Volgens een bepaalde wiskundige stelling, het bemonsteringstheorema van Nyquist, zal het signaal exact uit de samples te reconstrueren zijn, wanneer de sample-frequentie tenminste twee maal zo hoog is als de hoogste in het signaal voorkomende frequentie. Bevatte het bij de A/D-converter binnenkomende signaal dus geen componenten boven 22,05 kHz, dan moet het oorspronkelijke signaal in het geval van de CD weer exact uit de 44.100 samples per seconde te reconstrueren zijn. Voorafgaande aan de bemonstering zal het audio-systeem daartoe wel eerst zorgvuldig alle eventuele hogere frequenties weg dienen te filteren. Pas daarna wordt tot het bemonsteren en kwantificeren overgegaan.

Mocht het signaal ondanks deze filtering toch nog componen-

ten bevatten die de halve bemonsteringsfrequentie te boven gaan, dan treedt een hinderlijk verschijnsel op dat aliasing wordt genoemd. Aliasing of 'alias'-vervorming houdt in dat alle signaaldelen hoger dan de halve bemonsteringsfrequentie bij de reconstructie van het signaal (uit de samples) in een ander deel van het geluidsspectrum terugkomen. Bevat een geluidssignaal bijvoorbeeld een hoogfrequente storing met een frequentie van 41.000 Hz ($f = 41$ kHz) en wordt dit signaal bemonsterd met een frequentie van 44,1 kHz ($f_s = 44,1$ kHz), dan wordt het stoorgeluid na reconstructie teruggevonden als een aliasing produkt met een frequentie van $f_s - f = 3,1$ kHz. De frequentie-componenten boven 22,05 kHz worden dus als het ware teruggevouwen in het spectrum (dit wordt daarom ook wel een vouw-spectrum genoemd). In bovengenoemd voorbeeld is dit verschijnsel extra hinderlijk omdat ons gehoor bij 3,1 kHz bijzonder gevoelig is, terwijl het oorspronkelijke stoorgeluid van 41 kHz voor ons totaal onhoorbaar is. Het is dus zaak ervoor te waken dat alle signaalcomponenten boven 22,05 kHz zorgvuldig zijn weggefilterd, alvorens het signaal wordt bemonsterd. Het filter waarmee dit kan worden gedaan, heet een anti-aliasing filter.

CD: demo bemonstering 10 – 16

Bij het omzetten van een tijd-continu signaal in een tijd-discreet signaal kan hinderlijke vervorming optreden, de zogenaamde 'alias'-vervorming of aliasing. Dit hinderlijke effect doet zich voor wanneer er in het te bemonsteren geluidssignaal frequenties voorkomen die hoger zijn dan de halve bemonsteringsfrequentie. Aan de hand van een live-opname van *Harry Muskee* is op de CD een demonstratie te vinden van dit effect.

10 is het oorspronkelijke, breedbandige muziekfragment (0-22 kHz).

11 is hetzelfde fragment maar nu bemonsterd met 12 kHz en zonder tevoren een anti-aliasing filter toe te passen. Aangezien het signaal vóór de bemonstering aanzienlijke componenten boven de halve sample-frequentie (dus boven 6 kHz) bevatte, zullen deze spectrum-delen teruggevouwen worden naar de band onder 6 kHz. Dit levert een duidelijk hoorbaar aliasing-effect op.

[12] is een gefilterde versie van het oorspronkelijke signaal, waaruit voorafgaand aan de bemonstering met een anti-aliasing filter alle componenten boven 6 kHz weggefilterd zijn.

[13] Het gefilterde signaal van de voorgaande demo is nu bemonsterd met 12 kHz. Door het verdwijnen van het aliasing-effect klinkt [13] beduidend beter dan [11].

[14] Wat er bij de bandbreedte-begrenzing tot 6 kHz allemaal precies is weggefilterd is op deze demo te beluisteren.

[15] Hoe het pure aliasing-product (foutsignaal) klinkt, is hoorbaar gemaakt op track [15] van de CD. Vergelijking van demo [15] en demo [14] maakt duidelijk hoe de tonen hoger dan 6 kHz precies in het spectrum worden teruggevouwen.

[16] In dit fragment is ten slotte te horen wat er gebeurt wanneer het signaal bij weergave niet op de juiste manier door een reconstructie-filter wordt 'glad gestreken'. Normaal zal het reconstructie-filter in uw CD-speler ervoor zorgen dat het blokvormige signaal dat door de bemonstering is ontstaan (verg. afb. 2), weer in een glad verlopend (tijd-continu) signaal wordt omgezet. Door een speciale technische truc (up-sampling van 12 naar 44,1 kHz zonder reconstructie-filter) wordt er op de CD voor gezorgd dat uw CD-speler daar ditmaal niet in slaagt. De stapjes waartussen het reconstructie-filter interpolaties moet uitvoeren zijn hiervoor te groot.

Wat op deze manier hoorbaar wordt gemaakt, is hoe muziek zou klinken wanneer er in uw CD-speler geen reconstructie-filter zou zitten. Alle opnames zouden dan 'schel' klinken zoals hier wordt gedemonstreerd.

Precisie

Onder kwantificeren verstaan we, zoals hiervoor reeds is gezegd, het omzetten van een signaalmonster met een continu waardenbereik in een getal met een discreet waardenbereik. Door uit een beperkte set signaalwaarden (de zogenaamde kwantificerniveaus) steeds de dichtstbijzijnde waarde te kiezen, wordt het oorspronkelijke signaal benaderd. Onvermij-

delijk ontstaat daarbij steeds een afrondingsfout, die ook wel de kwantificeerfout wordt genoemd. Uiteraard zal de gemiddelde kwantificeerfout kleiner zijn naarmate er meer kwantificerniveaus beschikbaar zijn. Echter, het aantal bits dat nodig is voor het coderen van de geluidsmonsters neemt hierdoor toe en daarmee ook de omvang van de bitstream.

Bij de opname van een CD bijvoorbeeld, worden de samples omgezet in getallen van 16 bits waarmee een aantal kwantificerniveaus bereikt wordt van

$$- 2^{15} \dots 2^{15} - 1 \text{ ofwel } - 32768 \dots 32767.$$

Een bereik dus van $2^{16} = 65536$ getallen, of 2^{16} mogelijke combinaties van 16 bits waarmee de samples gecodeerd kunnen worden.

Is er voor het kwantificeren (coderen) een kleiner aantal bits dan 16 beschikbaar, dan moet er vanzelfsprekend sterker worden afgerond. De kwantificeerfout neemt dan toe en de precisie af. De gehoormatige effecten hiervan zijn op de demonstratie-CD te beluisteren op de tracks 17–25.

CD: demo kwantificeer-precisie 17–25

Aan de hand van een opname van de Friese *Brassband de Wâldsang* (wereldkampioen in dit genre) is op de CD het effect te horen van een afnemende kwantificeer-precisie op de weergave. De grondregel is daarbij: hoe minder de kwantificeer-precisie, des te hinderlijker is de vervorming die als gevolg van de afrondingsfouten optreedt. Overigens mag hieruit niet worden afgeleid dat het zinvol zou zijn om met een zeer, zeer grote precisie te kwantificeren. Op zeker moment (boven 16 bits) zal de precisie namelijk zo groot zijn, dat deze door ons gehoor niet meer kan worden waargenomen.

Bij een bemonsteringsfrequentie van 44,1 kHz is het signaal op de demo's 17–25 met een afnemende precisie gekwantificeerd. 17 is het originele 16-bits signaal zoals dat normaal op de CD is opgeslagen. Bij een bemonsteringsfrequentie van 44.1 kHz komt dit overeen met een bitstream van 705,6 kbit/sec per kanaal.

| fragment | aantal bits/sample | interval- grootte | kbits/sec per kanaal |
|----------|-----------------------|----------------------|-------------------------|
| 17 | 16 | 1 | 705,6 |
| 18 | 14 | 4 | 617,4 |
| 19 | 12 | 16 | 519,2 |
| 20 | 10 | 64 | 441,0 |
| 21 | 8 | 256 | 352,8 |
| 22 | 6 | 362 | 264,6 |
| 23 | 4 | 500 | 176,4 |
| 24 | 2 | 725 | 88,2 |
| 25 | 1 | 861 | 44,1 |

Huidige standaarden voor codering van hoge-kwaliteit audio

Tegenwoordig zijn er al verschillende audio-codeertechnieken in gebruik. Belangrijk onderscheid tussen deze codeersystemen is de omvang van de bitstream. Deze wordt bepaald met de eenvoudige formule $f_s \times R$, waarbij f_s de bemonsteringsfrequentie in kHz is en R het aantal bits per monster. Voor bijvoorbeeld de codering van een CD komt dat neer op een netto bitsnelheid (dus voordat er foutprotectie-bits worden toegevoegd) van $44,1 \times 16 = 705,6$ kbit/s per kanaal.

De belangrijkste audio-codeersystemen van het moment zijn:

- Het Compact-Disc systeem dat uitgaat van een bandbreedte van 22 kHz en dat zoals hierboven al is vermeld een bitsnelheid heeft van $44,1 \times 16 = 705,6$ kbit/s per kanaal.
- De Digital Audio Tape (DAT)-recorder met een bitsnelheid van $48 \times 16 = 768$ kbit/s per kanaal.
- Instantaneous A-Law companding met een netto bitsnelheid van 352 kbit/s per kanaal. Het systeem wordt gebruikt voor de transmissie van hoogwaardige digitale audio-signalen over het telefoonnet via 2048 kbit/s (2 Mbit/s) transmissiekanalen. Het systeem wordt onder andere toegepast bij de distributie van radioprogramma's naar FM-zenders en voor het coderen van TV-geluid.

- Het NICAM (Near-Instantaneously Companded Audio Multiplex) systeem van de BBC bestaat in verschillende versies. De bitsnelheden zijn respectievelijk 322, 354 en 323 kbit/s per kanaal. Via 2 Mbit/s verbindingen wordt NICAM in Engeland gebruikt voor de distributie van radioprogramma's naar steunzenders. Recentelijk is er in België, Engeland en de Scandinavische landen een versie van NICAM met 2 kanalen in gebruik genomen voor het transport van het stereo TV-geluid.

Wat bij al deze codeertechnieken opvalt is de relatief hoge bitsnelheid. Recent onderzoek heeft echter uitgewezen dat een aanzienlijke reductie van de bitsnelheid mogelijk is zonder dat dit een duidelijk hoorbare invloed hoeft te hebben op de geluidskwaliteit. Volgens schattingen kan er voor het coderen van audio-signalen worden volstaan met een bemonsteringsfrequentie van 32-48 kHz en een aantal van 3 tot 4 bits per signaalmonster om het geluid zonder hoorbare vervorming te coderen. Dit komt overeen met een bitsnelheid in de buurt van 128 kbit/s per kanaal.

¹⁴ ISO/IEC/JTC1/SC2/WG11 MPEG; Coding of moving pictures and associated audio for digital storage media at up to about 1.5 MBit/s. ISO/MPEG document 3-11172 DIS, november 1992.

- De MPEG audio-codec, die door het ISO is gestandaardiseerd, is een voorbeeld van zo'n efficiënt codeersysteem¹⁴. Deze audio-codec is ontwikkeld in een samenwerkingsverband van onder andere AT&T, Philips en Thomson Consumer Electronics. De codec bestaat uit drie layers, elk met een bit-rate die instelbaar is van 192 tot 32 kbit/s per kanaal. Layer I staat daarbij voor de meest eenvoudige manier van coderen, layer III is de meest complexe en kwalitatief meest hoogwaardige codeermethode. Het onderzoek naar layer III is voor de lagere bitsnelheden nog niet afgerond. De door PTT Research ontwikkelde en op de CD gedemonstreerde Adaptieve Transformatie Codering (ATC) bevindt zich op een niveau dat vergelijkbaar is met MPEG layer III.

- Het codeersysteem dat in de onlangs op de markt verschenen Digitale Compact Cassette (DCC)-recorder wordt toegepast, is afgeleid van ISO/MPEG layer I en werkt op een bitsnelheid van 192 kbit/sec per kanaal.



◀ Foto 4

DCC 900 recorder van Philips.

Redundantie en irrelevantie

Er zijn in het verleden nogal wat experimenten uitgevoerd waarbij men heeft getracht hoge-kwaliteit audio te realiseren met behulp van technieken die in de wereld van de spraakcodering gangbaar zijn. De betreffende spraakcodeertechnieken zijn gebaseerd op bepaalde modellen van de bron (de menselijke stem) en op de daarmee samenhangende karakteristieke eigenschappen van het spraaksignaal. Met behulp van deze technieken is het mogelijk om bijvoorbeeld ten behoeve van digitale autotelefonie (GSM) de bitstroom terug te brengen van 64 kbit/s (ISDN) naar 13 kbit/s¹⁵.

Bij het coderen van muziek blijkt een dergelijke aanpak echter niet te werken. Dit komt omdat van tevoren nauwelijks iets vaststaat over de bron (variërend van een groot symfonieorkest, tot een popgroep, koor of strijkkwartet) noch over de eigenschappen van het signaal. Een codeertechniek als ATC van PTT Research¹⁶ neemt daarom de golfvorm van het oorspronkelijke signaal als uitgangspunt en zal vervolgens proberen om deze met behulp van zo min mogelijk bits vast te leggen c.q. te reproduceren. Op de CD is aan de hand van de

¹⁵ In het Studieblad is aan deze efficiënte manier van spraakcodering (LPC, Linear Predictive Coding) aandacht besteed in het kader van de artikelenreeks over het pan-Europese autotelefoonnet GSM (ATF-4). Zie: W. van Blitterswijk en A.A.M. van der Krogt, *De toekomst van de autotelefoondienst: GSM het vierde generatie autotelefoonnet* (dl.2), PTT Telecom Studieblad, juli/augustus 1990, pp. 367-384.

¹⁶ Maar bijv. ook Sub-Band Coding (SBC) en Adaptive Differential Pulse-Code Modulation (ADPCM).

demo's [31]–[34] te horen dat een dergelijke aanpak in de praktijk goed werkt, zonder dat daarvoor een omvangrijke bitstroom nodig is.

Om tot deze beperkte bitstroom te kunnen komen, moet er tijdens het coderen gebruik gemaakt worden van redundantie en irrelevantie in het audio-signaal. In de volgende twee paragrafen zal kort worden uitgelegd wat er onder deze begrippen wordt verstaan.

Redundantie in het audio-signaal. Audio-signalen bevatten over het algemeen veel zich herhalende componenten, ook wel periodieke componenten genoemd. Componenten die dus regelmatig terugkeren en waarvan het niet nodig is ze tijdens het coderen steeds opnieuw te beschrijven (redundantie). Een sprekend voorbeeld hiervan is een signaal dat uit maar één enkele sinus bestaat. Een dergelijk signaal is met slechts enkele tientallen bits te karakteriseren door de amplitude, de frequentie en de fase te beschrijven. Bij de weergave zal zo'n periodieke component op basis daarvan steeds opnieuw gereconstrueerd kunnen worden. Een belangrijk verschijnsel is in dit verband dat hoe gepiekerd een bepaald spectrum is, hoe meer (lange termijn) redundantie er in het signaal zal voorkomen. Uit afbeelding 3 valt dit duidelijk af lezen. Daarnaast geldt nog dat wanneer we een audio-signaal met een voldoende hoge frequentie bemonsteren, de openvolgende signaalmonsters sterk op elkaar zullen lijken. Dit komt omdat de bemonsteringsfrequentie dan zo hoog is, dat in de buurt van de halve bemonsteringsfrequentie (de snelst mogelijke variaties) in de praktijk slechts zelden signaal voorkomt of alleen met een zeer kleine amplitude. Tussen twee elkaar opvolgende signaalmonsters bestaat als gevolg hiervan vaak een sterke onderlinge overeenkomst. Er is zoals technici dat wel noemen sprake van een hoge auto-correlatie. Dit resulteert in een veelvuldig herhalen van steeds dezelfde informatie, wat er kort gezegd op neerkomt dat de bitstroom een grote (korte termijn) redundantie bevat.

Irrelevantie in het audio-signaal. Irrelevantie is nauw verbonden met het proces van waarneming of anders gezegd met de mogelijkheden en onmogelijkheden van het menselijk gehoor. Onder irrelevantie in audio-signalen verstaat men dan het gedeelte van de informatie dat niet door ons gehoor kan

DEMO 3 KWANTIFICEREN 17-25 (lineaire PCM, 44,1 kHz sampling)

Brassband de Wäldsang – Steal Away

| | bits/sample | max. waarde | stapgrootte | kbit/s kanaal | |
|----|-------------|-------------|-------------|---------------|------|
| 17 | 16 | 32768 | 1 | 705,6 | 0:34 |
| 18 | 14 | 32768 | 4 | 617,4 | 0:34 |
| 19 | 12 | 32768 | 16 | 519,2 | 0:35 |
| 20 | 10 | 32768 | 64 | 441,0 | 0:35 |
| 21 | 8 | 32768 | 256 | 352,8 | 0:35 |
| 22 | 6 | 11600 | 362 | 264,6 | 0:34 |
| 23 | 4 | 4000 | 500 | 176,4 | 0:35 |
| 24 | 2 | 1450 | 725 | 88,2 | 0:35 |
| 25 | 1 | 861 | 861 | 44,1 | 0:36 |

DEMO 4 MASKERING 26-30 (foutsignaal bij ATC, 64 kbit/s per kanaal)

AghArtA – Frozen Echoes

| | | | |
|----|------------------------------------|-----------------------------|------|
| 26 | oorspronkelijke opname | (32 kHz sampling, 0-15 kHz) | 1:04 |
| 27 | foutsignaal zonder noise-shaping | (32 kHz sampling, 0-15 kHz) | 1:04 |
| 28 | ATC gecodeerd zonder noise-shaping | (32 kHz sampling, 0-15 kHz) | 1:04 |
| 29 | foutsignaal met noise-shaping | (32 kHz sampling, 0-15 kHz) | 1:04 |
| 30 | ATC gecodeerd met noise shaping | (32 kHz sampling, 0-15 kHz) | 1:04 |

DEMO 5 ADAPTIEVE TRANSFORMATIE CODERING (ATC) 31-34

Richard Nanes – Sonnet # 7 in D Minor

| | kbit/s kanaal | sample freq. | bandbreedte | | |
|----|----------------|--------------|-------------|----------|------|
| 31 | oorspr. opname | 512 | 32 kHz | 0-15 kHz | 0:48 |
| 32 | ATC gecodeerd | 128 | 32 kHz | 0-15 kHz | 0:48 |
| 33 | ATC gecodeerd | 64 | 32 kHz | 0-15 kHz | 0:48 |
| 34 | ATC gecodeerd | 32 | 32 kHz | 0-15 kHz | 0:48 |

- 35 **AghArtA** – Frozen Echoes (Martin de Boer) 5 : 00
Recorded at Lowland Studio, Roderwolde, Holland
Produced by Martin de Boer and Rene Houwen
Engineer: Rene Houwen
Black 'n White Music BW 92002
- 36 **Brassband de Wäldsang** – Steal away (trad. arr. H. Snell) 4 : 02
Recorded at Pelikaankerk, Leeuwarden
Recorded and mixed at Lowland Studio, Holland
Sound Director: G. Huizinga
Engineer: Rene Houwen
Pink Records PR 26012
- 37 **Richard Nanes** – Sonnet # 7 in D Minor (Richard Nanes) 4 : 05
Piano: C. Bechstein # 176290
Producer, Engineer & Editor: Gregory K. Squires
Delfon CDR 6070
- 38 **The Budapest String Quartet** – String Quartet # 4: Andante (Richard Nanes) 8 : 53
Recorded at MA Film and Sound Studios
Producer, Engineer & Editor: Gregory K. Squires
Delfon CD 7080
- 39 **Harry Muskee** – Brother Booze (Muskee/Kuschel) 5 : 07
Recorded live at Ahoy, Rotterdam 7-2-91
Recorded and mixed at Lowland Studio, Holland
Engineer: Rene Houwen
Black 'n White Music BW 91001

DEMO 1 BANDBREEDTE 1-9

The Budapest String Quartet – Nanes String Quartet # 4: Andante

| | | | | |
|---|------------------------|----------|-------------------------|------|
| 1 | oorspronkelijke opname | 0-22 kHz | (sample freq. 44,1 kHz) | 0:32 |
| 2 | low-pass bandbegrensd | 0-15 kHz | (sample freq. 44,1 kHz) | 0:32 |
| 3 | low-pass bandbegrensd | 0-10 kHz | (sample freq. 44,1 kHz) | 0:32 |
| 4 | low-pass bandbegrensd | 0- 8 kHz | (sample freq. 44,1 kHz) | 0:36 |
| 5 | low-pass bandbegrensd | 0- 4 kHz | (sample freq. 44,1 kHz) | 0:35 |
| 6 | low-pass bandbegrensd | 0- 2 kHz | (sample freq. 44,1 kHz) | 0:31 |
| 7 | oorspronkelijke opname | 0-22 kHz | (sample freq. 44,1 kHz) | 0:33 |
| 8 | high-pass bandbegrensd | 4-22 kHz | (sample freq. 44,1 kHz) | 0:33 |
| 9 | high-pass bandbegrensd | 8-22 kHz | (sample freq. 44,1 kHz) | 0:32 |

DEMO 2 ALIASING 10-16

Harry Muskee – Brother Booze

| | | | | |
|----|---|----------|---------------------|------|
| 10 | oorspronkelijke opname | 0-22 kHz | (44,1 kHz sampling) | 0:44 |
| 11 | aliasing | 0-22 kHz | (12 kHz sampling) | 0:44 |
| 12 | low-pass bandbegrensd | 0- 6 kHz | (44,1 kHz sampling) | 0:44 |
| 13 | low-pass bandbegrensd | 0- 6 kHz | (12 kHz sampling) | 0:44 |
| 14 | high-pass bandbegrensd | 6-22 kHz | (44,1 kHz sampling) | 0:44 |
| 15 | aliasing product | | (12 kHz sampling) | 0:44 |
| 16 | up-sampling 12 kHz naar 44,1 kHz, geen reconstructie-filter | | | 0:44 |

Audio

C O D E R I N G

demonstratie op 128, 64 & 32 kbit/s
per kanaal van ATC, een door PTT Research
ontwikkelde audio-codeertechniek

Projectleiding

Ysbrand van der Veen (PTT Telecom)

Vormgeving

Eddy Cusiel (Studio Dorèl)

Fotografie

Thom Segers, Fred de Jager (PTT Research)

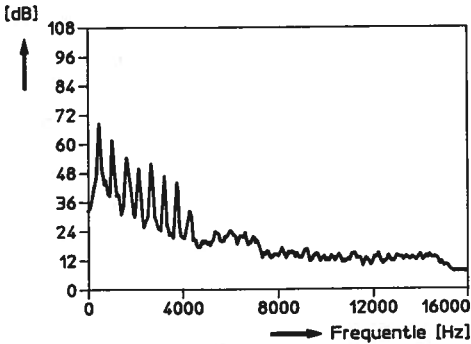
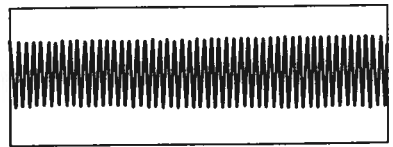
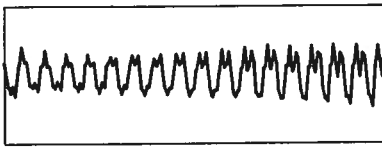
Techniek demo's

Jan Stererdink (PTT Research)

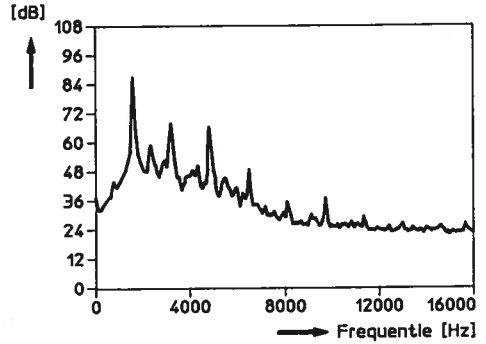
Bijlage bij PTT Telecom Studieblad, februari 1993

Deze CD is tot stand gekomen met medewerking
van Delfon Recording Society, Lowland Studio,
Black 'n White Music, Pink Records.

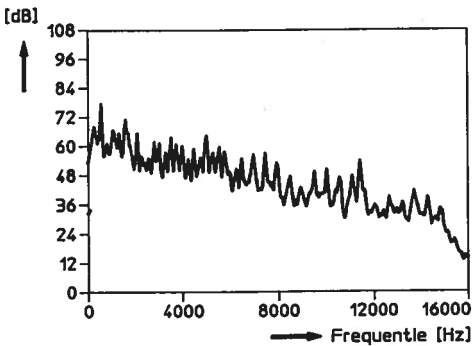
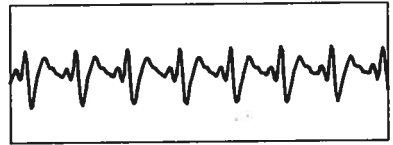
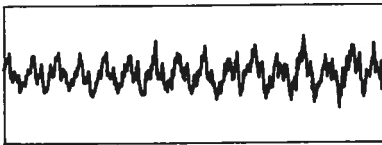
*Special thanks to Eddy Harms (Pink Records),
Rene Houwen (Lowland Studio).*



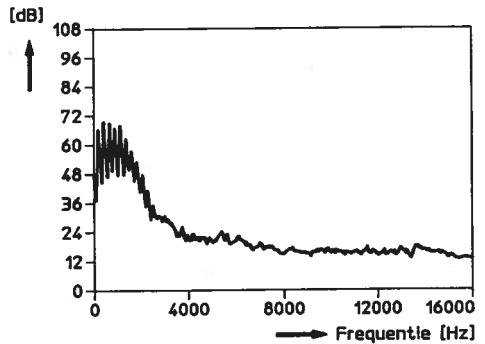
a) banjo



b) fluit



c) spaanse gitaer



d) trombone

worden waargenomen. Dit deel van de informatie kan dus weggelaten worden zonder dat de luisteraar het verschil zal merken.

▲ Afb. 3

Enkele voorbeelden van muziek-signalen in tijd en frequentie.

Een voorbeeld van irrelevantie is een signaalcomponent van 23 kHz. Deze component is voor vrijwel iedereen onhoorbaar en kan zonder dat iemand het zal opmerken worden weggelaten.

Hetzelfde geldt voor een signaal dat met een bijzonder grote precisie (bijvoorbeeld 20 bits) is gekwantificeerd. De laatste (minstens 4) bits geven aanleiding tot een niet-waarneembare nauwkeurigheid van de weergave, waarmee deze bits dus overbodig zijn. Aan de hand van de demo's [17]–[25] van de Studieblad-CD kunt u het effect van deze nauwkeurigheid zelf beoordelen.

Klassieke kwaliteitsmaten

Hoe goed de moderne audio-apparatuur ook is, toch zal er in vergelijking met het oorspronkelijke signaal bij het coderen, transporteren of opslaan van het geluid altijd een zekere kwaliteitsvermindering optreden. Wanneer we bijvoorbeeld stukjes muziek van de Compact Disc op een analoge geluidscassette opnemen, is er bij het afspelen van de cassette extra ruis hoorbaar. De geluidskwaliteit is afgenomen.

Brengen we de omvang van de bitstroom op een Compact Disc met behulp van een bepaalde codeermethode terug, dan zal er meestal ook van een degradatie van de audio-kwaliteit sprake zijn.

Om deze degradatie te kunnen meten wordt er in het algemeen gekeken naar het verschil tussen het originele en het gedegradeerde signaal; het signaal zoals dat aan het audiosysteem is aangeboden en het signaal zoals dit het audiosysteem weer verlaat. Is het verschil tussen beide klein dan is de kwaliteit van het audio-apparaat hoog. Is het verschil groot dan is de kwaliteit laag¹⁷.

Bij de klassieke technische metingen aan audio-apparatuur worden altijd speciale testsignalen gebruikt, er wordt dus bijvoorbeeld niet met stukjes muziek gewerkt. Behalve de meting van de frequentie-overdracht worden meestal ook ruis- en vervormingsmetingen uitgevoerd. Uit het oogpunt van de luisteraar dient daarbij met twee zaken rekening gehouden te worden.

- Het meten van de frequentie-overdracht is alleen belangrijk voor frequenties die ons oor kan waarnemen (20-20.000 Hz).

¹⁷ Ten behoeve van de eenvoud laten we hier 'signal-enhancement' technieken, waarmee door slimme signaalprocessing de kwaliteit van het ingangssignaal verbeterd kan worden, buiten beschouwing.

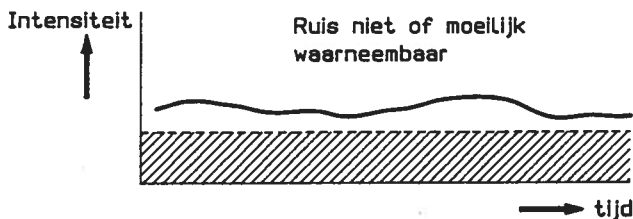
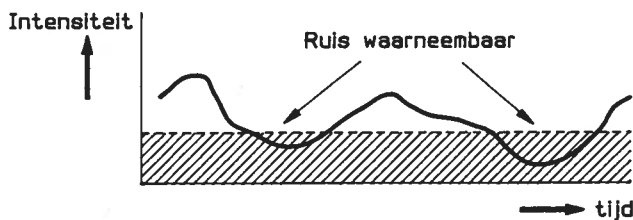
• Bij ruis- en vervormingsmetingen moet rekening worden gehouden met de frequentie-afhankelijke gevoeligheid van ons gehoor¹⁸. Voor een (stoor)signaal met een frequentie van 1000 Hz is het menselijk oor bijvoorbeeld heel wat gevoeliger dan voor een (stoor)signaal van 15.000 Hz.

Is de kwaliteit eenmaal op de klassieke manier bepaald, dan kunnen we voorspellen wat de waargenomen audio-kwaliteit van een apparaat voor een bepaalde toepassing zal zijn. Nemen we bijvoorbeeld een goedkope cassetterecorder zonder ruisreductie-systeem met een slechte signaal/ruis-afstand van pakweg 40 dB, dan weten we dat er in muziekstukken met zachte passages altijd ruis is waar te nemen (afb. 4a). Draaien we op dezelfde cassetterecorder echter een muziekstuk met weinig dynamiek af (afb. 4b) dan zal de ruis veel moeilijker waarneembaar zijn. De hoorbare kwaliteit is dus niet alleen afhankelijk van de kwaliteit van het audio-apparaat zelf, maar is tevens afhankelijk van het type signaal dat wordt aangeboden.

De betekenis van de klassieke kwaliteitsmaten is daarmee dat we de gehoormatige prestaties van audio-apparatuur verantwoord met elkaar kunnen vergelijken. Betrouwbaar voorspellen wanneer een cassetterecorder x of y goed of slecht zal klinken, is mogelijk.

¹⁸ De stoorspanningen moeten bij de bepaling van de signaal/ruis-verhouding perceptief worden gewogen met bijvoorbeeld de A-curve of de CCIR/CCITT-curve. Deze weging houdt rekening met de frequentie-afhankelijke gevoeligheid van ons oor. (CCITT. Measurement of weighted noise in sound-programme circuits, 1989. Recommendation J.16).

▼ Afb. 4
Signaal met veel dynamiek (a) en weinig dynamiek (b).



¹⁹ De signaal/ruis-verhouding in aanwezigheid van muziek (SNR) blijft praktisch gelijk.

Er ontstaan echter problemen wanneer deze klassieke technische kwaliteitsmaten worden toegepast op audio-systemen met bijvoorbeeld ruisonderdrukking. Nemen we dezelfde casetrecorder maar nu voorzien van een ruisreductie-systeem zoals Dolby B, dan zal de gehoormatige of perceptieve kwaliteit bij eenzelfde muziekstuk toenemen. In de klassieke kwaliteitsmaat komt deze hoorbare kwaliteitsverbetering evenwel onvoldoende tot uitdrukking¹⁹. Dit komt omdat de traditionele kwaliteitsmaten nu eenmaal niet ontworpen zijn voor het beoordelen van systemen die bepaalde fundamentele eigenschappen van het menselijk gehoor als uitgangspunt hebben.

Het menselijk gehoor

Uit het bovenstaande blijkt dat voor het beoordelen van 'klassieke' audio-apparaten zoals tape-recorders, tuners, CD-speakers, DAT-recorders en telefoons de vertrouwde technische kwaliteitsmaten zinvol zijn. Voor de beoordeling van moderne codeertechnieken (analoog of digitaal) zijn ze echter niet toereikend. Het verschil tussen technische specificatie en waargenomen geluidskwaliteit is daarvoor te hoog, om niet te zeggen onaanvaardbaar hoog. Met inachtneming van de eigenschappen van het menselijk gehoor is er voor dit probleem dan ook maar één oplossing denkbaar, namelijk de metingen die gebaseerd zijn op sinussen aan te vullen met metingen die gebaseerd zijn op relevante audio-signalen: spraak en muziek.

Eén van de belangrijkste eigenschappen van het menselijke gehoor waar verreweg de meeste technische metingen geen rekening mee houden is maskering. Een sterk signaal kan als gevolg van deze gehoooreigenschap een zwak signaal verdoezelen, zoals we in afbeelding 4 al zagen. Hebben we thuis bijvoorbeeld de radio heel zachtjes aan staan, dan is de kans groot dat we tijdens het stofzuigen de radio niet meer zullen horen. Door het maskeereffect is het zwakkere signaal niet meer waar te nemen.

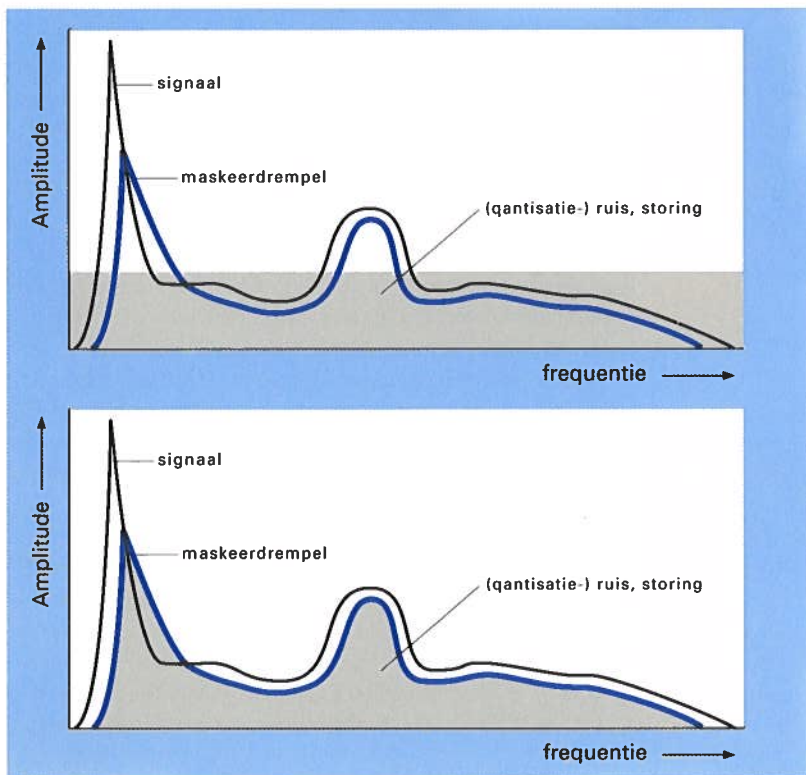
Maskering treedt zowel in het frequentie- als in het tijd-domein op. Zo zijn stoorsignalen die zich in hetzelfde frequentiegebied bevinden als het audio-signaal veel moeilijker waarneembaar dan stoorsignalen die verder weg zitten. In de tijd gezien kunnen stoorsignalen andere signalen maskeren die later (of zelfs eerder) komen. In werkelijkheid zullen deze

effecten echter nooit afzonderlijk optreden. Er is altijd sprake van een tijd/frequentie-domein maskering. In de verdiepingstof wordt hier nader op ingegaan.

Doordat de klassieke technische kwaliteitsmaten geen rekening houden met het verschijnsel maskering, stuiten we op problemen bij het beoordelen van audio-apparatuur die daar wel rekening mee houdt. Stel bijvoorbeeld dat we twee audio-codeersystemen met elkaar willen vergelijken die beide gemiddeld evenveel ruis produceren. Bij één systeem geven we echter door slim te coderen de ruis in het tijd/frequentie-domein een zodanige vorm dat de ruis onder de maskeerdrempel valt. Een normale technische specificatie zou voorspellen dat beide systemen even slecht klinken. De ruis is bij het slimme codeersysteem echter niet hoorbaar, zoals in afbeelding 5 wordt geïllustreerd.

▼ Afb. 5

Wanneer een signaal door een audio-apparaat verwerkt wordt, ontstaat ruis. Dit is principieel onvermijdelijk. Voor digitale audio-apparatuur valt daarbij te denken aan de zgn. kwantificeringsruis. Zonder bijzondere maatregelen heeft deze ruis een vrij vlak verloop en zal op hinderlijke wijze hoorbaar zijn. Door de ruis een zodanige vorm te geven dat deze bij alle frequenties onder de maskeerdrempel valt (noise-shaping), kan ze onhoorbaar worden gemaakt.



Deze codeertechniek, ook wel noise-shaping genoemd, kan zowel in het analoge als in het digitale domein worden toegepast. Ze maakt eens te meer duidelijk dat de klassieke technische specificaties soms bitter weinig zeggen over de waargenomen kwaliteit van moderne audio-apparatuur. Wil PTT in de toekomst verantwoord bepaalde audio-codeersystemen kunnen aanschaffen, dan zijn er dus nieuwe testmethodes nodig waarmee wel perceptief relevante metingen kunnen worden gedaan.

CD: demo maskering [26]–[30]

Alle codeersystemen introduceren vervorming. Deze vervorming kan echter onhoorbaar gemaakt worden door gebruik te maken van het verschijnsel maskering. Op de CD wordt het verschijnsel maskering aan de hand van het door PTT Research ontwikkelde codeersysteem Adaptieve Transformatie Codering (ATC) geïllustreerd in de demonstraties [26]–[30]. De codering vindt plaats met een bemonsteringsfrequentie van 32 kHz (bandbreedte 15 kHz) en een bitsnelheid van 64 kbit/sec per kanaal. Het muziekfragment is afkomstig van een in 1992 verschenen CD van de Nederlandse groep *AghArtA*.

[26] is het originele signaal dat alleen is aangepast aan een bandbreedte van 15 kHz.

[27] coderen we dit signaal met ATC in 64 kbit/sec, dan ontstaat er een foutsignaal dat op de demo geïsoleerd is weergegeven. Dit foutsignaal is duidelijk waarneembaar, de signaal-ruisverhouding bedraagt ca. 45 dB.

[28] In het gecodeerde audio-signaal van deze demo is het foutsignaal uit de vorige track vanzelfsprekend aanwezig. Het foutsignaal is nu echter veel moeilijker te horen omdat het gedeeltelijk (en voor veel mensen zelfs geheel) wordt gemaskeerd door het audio-signaal.

We kunnen het verschijnsel maskering nog veel beter uitbuiten door van noise-shaping gebruik te maken. Het foutsignaal wordt hierbij aangepast aan het eigenlijke audio-signaal.

[29] is weer een geïsoleerd foutsignaal zoals dit optreedt bij codering met ATC bij 64 kbit/sec, echter nu is bij het

coderen noise-shaping toegepast. Dat de kwantificatiefout het signaalspectrum in zekere mate volgt, is duidelijk te horen aan het feit dat het audio-sigitaal ditmaal veel beter te herkennen is dan in demo [27]. Bovendien is het foutsigitaal op demo [29] sterker dan op demo [27], de signaal/ruis-verhouding bedraagt namelijk 40 dB, met andere woorden de ruis is 5 dB sterker.

[30] Doordat het ruisspectrum op deze demo dankzij noise-shaping het signaalspectrum volgt, is het samen met het gecodeerde audio-sigitaal minder hoorbaar. Door een betere maskering is de sterkere ruis op demo [30] minder goed hoorbaar dan de zwakkere ruis op demo [28]. Wie het verschil niet direct hoort, moet maar eens speciaal letten op de hoge frequenties (tamboerijn en bekens).

Perceptief relevant meten

Moderne systemen voor audio-codering zijn signaalafhankelijk werkende systemen. Hun gedrag staat als gevolg hiervan niet bij voorbaat vast, maar is afhankelijk van het signaal dat wordt aangeboden. Dit brengt met zich mee dat deze systemen ons bij de kwaliteitsbepaling voor bijzondere problemen stellen. Zo kunnen we met sinussen (speciale testsignalen) perceptief relevante metingen doen die met alle eigenschappen van het menselijk gehoor rekening houden én toch tegen het probleem aanlopen dat de overeenkomst met de in de praktijk waargenomen audio-kwaliteit bij spraak en muziek laag is. Het kan dus gebeuren dat een audio-apparaat met een goede perceptieve kwaliteit, gemeten met sinussen, bij bepaalde stukken muziek erg slecht klinkt.

De vraag is dan natuurlijk hoe we wel tot verantwoord opgezette, perceptief relevante metingen kunnen komen. Tenminste twee problemen dienen hiervoor allereerst te worden opgelost:

- er moet een relevante set audio-testsignalen worden samengesteld,
- er moet met alle relevante eigenschappen van ons gehoor rekening worden gehouden.

Is eenmaal aan deze voorwaarden voldaan dan kan er worden gemeten²⁰. Een ding weten we op dat moment echter niet,

²⁰ In de verdiegingsstof wordt uiteengezet hoe zo'n perceptief relevante meting er in grote lijnen uitziet.

namelijk of de nieuwe audio-kwaliteitsmaat ook daadwerkelijk overeenstemt met wat mensen als goede kwaliteit ervaren. De nieuw ontwikkelde kwaliteitsmaat zal dus getest moeten worden aan de hand van subjectieve experimenten waarbij proefpersonen uitspraken doen over de waargenomen audio-kwaliteit.

Het wetenschappelijk verantwoord uitvoeren van zo'n test is niet alleen bijzonder moeilijk maar ook heel erg tijdrovend. Dit blijkt onder andere uit de twee grote databases waar PTT Research op het ogenblik mee werkt. De eerste is aangemaakt tijdens de (Europese) standaardisatie van een spraakcodec die gebruikt wordt in het nieuwe pan-Europese autotelefoonnet (GSM)²¹. De tweede database is aangemaakt binnen de ISO/MPEG-groep die zich bezig houdt met het standaardiseren van een audio-codec met CD-kwaliteit en een lage bit-snelheid²². Het totaal aantal kwaliteitsoordelen van de proefpersonen is bij beide databases in de orde van vijftienduizend. Dit grote aantal is noodzakelijk om tot statistisch relevante uitspraken te kunnen komen. De testen in HiFi-tijdschriften waarbij – onder niet wetenschappelijke condities – slechts één of enkele proefpersonen uitspraken doen op grond van een zeer beperkte luister-test, kunnen bij het ontwikkelen van een audio-kwaliteitsmaat niet worden gebruikt. Al met al zal het duidelijk zijn dat het meten van de kwaliteit van audio-apparatuur die naast een signaalafhankelijke werking ook bepaalde eigenschappen van het menselijk gehoor als uitgangspunt heeft, aanzienlijk ingewikkelder is dan het uitvoeren van een frequentie-meting of het bepalen van de signaal/ruis-afstand.

Een eerste probleem bij het meten van de kwaliteit is dat er nauwelijks testsignalen te bedenken zijn om deze apparatuur mee door te meten. Voor een ideale meting zou gemeten moeten worden met alle mogelijke signalen die ooit aan het audio-apparaat kunnen worden aangeboden. In de praktijk wordt getest met een zinvolle selectie van signalen²³, die eveneens wordt gebruikt bij de hiervoor genoemde subjectieve evaluatie van apparatuur door proefpersonen.

Een tweede probleem is dat de goede modellen van ons gehoorsysteem allemaal erg veel rekentijd nodig hebben. Het testen aan de hand van een groot aantal signalen zoals dat in de subjectieve testen wordt gedaan, zou hierdoor bijzonder

²¹ J.E. Natvig, Evaluation of six medium bit-rate coders for the pan-european digital mobile radio system, in: *IEEE Journal on Selec. Areas in Commun.*, 6(2):324-331, februari 1988.

²² MPEG/Audio test report, 1990, ISO/IEC JTC1/SC2/WG11.

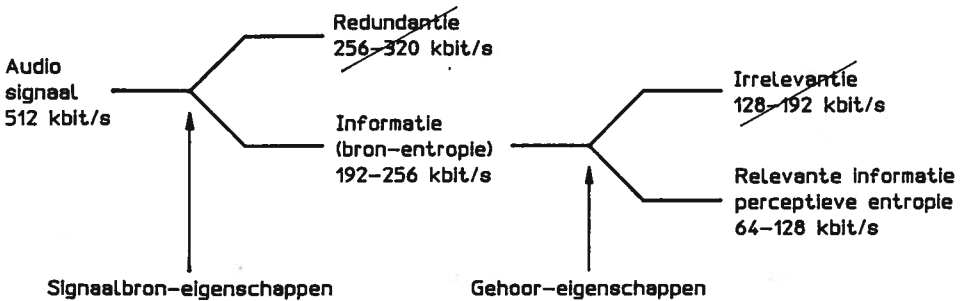
²³ Zie bijvoorbeeld: *Sound quality assessment material, recordings for subjective tests*, 1988, EBU document Techn. 3253.

veel tijd in beslag nemen. Bij PTT Research wordt er op het ogenblik daarom gewerkt aan een model dat met realistische rekentijden grote bestanden van audio-signalen kan doorrekenen. Voor het toetsen van dit model wordt gebruik gemaakt van de eerder genoemde databases. De eerste resultaten laten zien dat met redelijk eenvoudige modellen goede voorspellingen te maken zijn²⁴.

Zuinig coderen

Zoals hiervoor reeds is besproken, bevat een audio-signaal in het algemeen veel overbodige informatie. Overbodige informatie in de vorm van redundantie en irrelevantie, die respectievelijk samenhangt met de eigenschappen van de bron en de eigenschappen van het menselijk gehoor.

Hoeveel informatie er precies overbodig is, wordt in afbeelding 6 aangegeven. Daarin wordt gebruik gemaakt van twee belangrijke termen uit de informatie-theorie. In de informatie-theorie noemt men de minimale hoeveelheid bits die nodig is om een signaal foutloos weer te geven de bron-entropie. Analoog daaraan noemt men de minimale hoeveelheid bits die nodig is om een signaal zonder hoorbare fouten weer te geven de perceptieve entropie.



²⁴ J.G. Beerends & J.A. Stemerdink, A perceptual audio quality measure based on a psycho-acoustic sound representation, in: *Journal of the Audio Engineering Society*, 40:963-978, december 1992 (equivalent met PTT Research publ. NT-IN-92-756).

▼ Afb. 6
Bitreductie door de verwijdering van redundantie en irrelevantie.

We zien in de afbeelding dat een bitsnelheid van 64-128 kbit/s per audio-kanaal voldoende is voor het weergeven van de perceptieve entropie. Door nu apparatuur met een zekere mate van intelligentie aan het audio-systeem toe te voegen – aan de

²⁵ Bij Adaptieve Transformatie Codering wordt het audio-signaal blok voor blok met behulp van een Discrete Fourier Transformatie (DFT) of een Discrete Cosinus Transformatie (DCT) getransformeerd naar het frequentie-domein en opgesplitst in frequentie-componenten. Bij Sub-band Codering gebeurt de opsplitsing door een bank van filters. In het frequentie-domein worden de verschillende componenten vervolgens gecodeerd.

zendzijde ná de A/D-omzetting (coder) en aan de ontvangzijde vóór de D/A-omzetting (decoder) – kan redundantie en irrelevantie worden verwijderd c.q. de bitsnelheid worden gereduceerd.

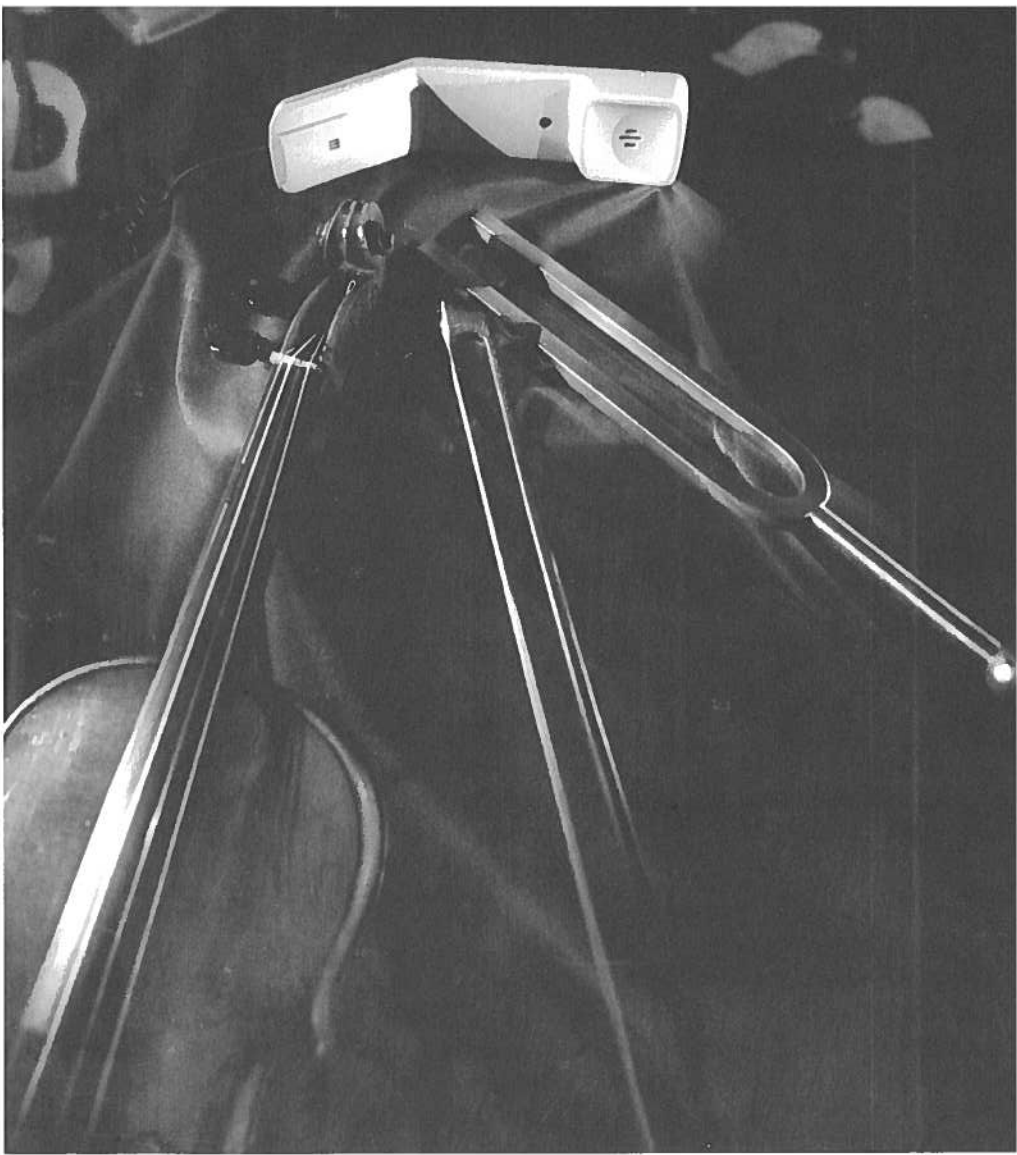
Zowel redundantie als irrelevantie kunnen het eenvoudigst worden verwijderd in het frequentie-domein. Het is dan ook niet verwonderlijk dat de betere audio-codeersystemen, zoals Adaptieve Transformatie Codering en Sub-band Codering, van het frequentie-domein uitgaan²⁵.

Beide technieken zijn in de spraakcodering reeds veelvuldig toegepast. Daarbij is meestal alleen de redundantie uit het signaal verwijderd. Bij codering van audio is echter ook verwijdering van de irrelevantie noodzakelijk, waarbij men zich meestal als doel stelt een signaal met CD-kwaliteit te coderen bij een lage bitsnelheid.

Adaptieve Transformatie Codering (ATC)

Een transmissiecapaciteit van 64 kbit/s per kanaal of zelfs het twee- of drievoud daarvan is bitter weinig wanneer het gaat om het transporteren van een complex signaal als hoge-kwaliteit audio. Nogmaals ter vergelijking: de netto bitsnelheid (dus exclusief de foutcorrectie-bits) waarmee een muziekopname voor Compact Disc wordt gecodeerd bedraagt 705,6 kbit/s per kanaal. Codeerapparatuur die deze informatiestroom zonder hoorbaar kwaliteitsverlies naar 64 kbit/s per kanaal om kan zetten, is internationaal nog niet verkrijgbaar. Vanwege de interessante toepassingsmogelijkheden van dergelijke apparatuur heeft PTT Research daarom besloten onderzoek te doen naar een hoge-kwaliteit audio-codeersysteem dat met een bitsnelheid van $n \times 64$ kbit/s kan volstaan. Dit systeem, de zogenaamde Adaptieve Transformatie Codering (ATC), wordt op de tracks 31–34 van de CD gedemonstreerd.

Het basisprincipe van ATC is eenvoudig. Het oorspronkelijke audio-signaal wordt na het passeren van de analoog/digitaal (A/D-)omzetter opgedeeld in blokken (frames) van bijvoorbeeld 1024 signaalmonsters ($N = 1024$). Met behulp van een wiskundige operatie wordt van ieder frame het frequentiespectrum berekend. Dit spectrum wordt vervolgens zo zuinig mogelijk gecodeerd (redundantie en irrelevantie worden verwijderd) en daarna overgezonden. Aan de kant van de ontvan-



ger wordt het signaal vervolgens gedecodeerd en via een omgekeerde operatie (een zogenaamde inverse transformatie) weer teruggebracht tot een analog signaal. Is daarbij alles goed gegaan dan zal een gemiddelde luisteraar geen verschil kunnen horen tussen het oorspronkelijke signaal en het signaal dat hij via de ATC-codec te horen krijgt. Gehoormatig of perceptief dienen beide signalen identiek te zijn. Op de demonstratie-CD kunt u beoordelen of dat ook daadwerkelijk het geval is.

▲ Foto 5

CD: demo Adaptieve Transformatie Codering [31] – [34]

Op deze tracks van de Studieblad-CD wordt aan de hand van een opname van de Amerikaanse pianist/componist *Richard Nanes* de audio-codec gedemonstreerd die gebaseerd is op de door PTT Research ontwikkelde Adaptieve Transformatie Codering (ATC).

[31] is het oorspronkelijke signaal dat conform de internationale standaard voor audiotransport bemonsterd is met 32 kHz (512 kbit/s). De bandbreedte is 15 kHz. Wat betreft het effect van de bandbreedte-begrenzing kunt u deze demo vergelijken met de originele CD-opname met een bandbreedte van 22 kHz op track [37].

[32] is het met ATC op 128 kbit/sec per kanaal gecodeerde en gedecodeerde signaal. De bandbreedte is 15 kHz.

[33] Eveneens met ATC gecodeerd, maar nu met 64 kbit/sec per kanaal (bandbreedte 15 kHz).

[34] Met ATC gecodeerd met 32 kbit/sec kanaal (bandbreedte 15 kHz).

De meeste mensen zullen in [34] een min of meer hinderlijke storing waarnemen, bij de hogere bitsnelheden van demo [32] en [33] is de storing veel moeilijker waarneembaar.

Conclusies en mogelijke toepassingen

In dit artikel is aandacht besteed aan de werkzaamheden van PTT Research op het gebied van audio-codering en perceptief relevant meten. Met behulp van de ontwikkelde perceptieve kwaliteitsmaat kan van de meest uiteenlopende apparatuur op het gebied van audio-transport (codecs, versterkers, telefoons etc.) de kwaliteit op een subjectief relevante manier worden bepaald. Dat transmissie van hoge-kwaliteit audio via $n \times 64$ kbit/s verbindingen mogelijk is, is aangetoond. De bij dit nummer van het Studieblad gevoegde CD laat u dat zelfs op een unieke manier horen.

Met het hoge-kwaliteit audio-transport over $n \times 64$ kbit/s verbindingen komen vele nieuwe mogelijkheden beschikbaar voor het gebruik van een digitaal basisnetwerk zoals ISDN. Enkele voorbeelden zijn:

- efficiënter gebruik van bandbreedte,

- efficiënter gebruik van geheugenruimte in opslagmedia (bijv. DCC, zie foto 4),
 - de mogelijkheid om muziek met hoge kwaliteit over het ISDN-netwerk te verzenden,
 - de mogelijkheid om via het (ISDN-)netwerk van PTT Telecom nieuwe diensten aan te bieden, zoals een hitlijn-dienst (Dial-a-Disk) met hoge audio-kwaliteit, live-concerten etc.,
 - audio-verbindingen voor radiouitzendingen op locatie (muzieklijnen) kunnen op ieder willekeurig moment en vanaf elke willekeurige plaats gerealiseerd worden,
 - vanwege het Europese karakter van ISDN zijn dergelijke flexibele muzieklijnen ook internationaal op eenvoudige wijze te realiseren, waarbij uiteraard ook het door omroepbedrijven onderling uitwisselen van programma's mogelijk is,
 - de marketing en verkoop van muziek met CD-kwaliteit volgens het teleshopping-concept gaat tot de mogelijkheden behoren,
 - bedrijven, bijvoorbeeld winkelketens, kunnen gesloten muziekdistributiesystemen naar hun filialen opzetten; centraal geproduceerde auditieve reclame-boodschappen kunnen hierdoor razendsnel worden geproduceerd en gedistribueerd.
- Bij al deze voorbeelden is het economisch interessant dat gebruik kan worden gemaakt van een bestaand digitaal basisnetwerk zoals ISDN. Zonder geavanceerde audiocodeertechnieken zouden voor al deze toepassingen aanzienlijk duurdere verbindingen noodzakelijk zijn. Verbindingen die bovendien van tevoren aangevraagd moeten worden en dus door de klanten minder flexibel in te zetten zijn.

Trefwoorden: Audio, ISDN, Audio-codering, Teleshopping-diensten, Telemarketingdiensten, Contributie, Distributie

Drs. Y.M. van der Veen
studeerde Nederlands aan de Universiteit van Amsterdam en specialiseerde zich tijdens en na zijn studie in het naar een breed publiek vertalen van wetenschappelijk verkregen inzichten, resulterend in o.a. onderwijs-

projecten, tentoonstellingen, tal van publicaties en gedramatiseerde documentaires (KRO-radio). Sinds 1987 is de heer Van der Veen in dienst van PTT Telecom, vanaf 1989 als hoofdredacteur van PTT Telecom Studieblad.

Ir. J.A. Stemerink studeerde elektrotechniek aan de TU Twente. Sinds eind 1985 is hij werkzaam bij PTT Research, de afdeling Transmissie en Codering. Zijn werkzaamheden liggen op het gebied van codering van spraak en muziek. Verder werkt hij in samenwerking met dr. ing. J.G. Beerends aan de ontwikkeling van een perceptieve audio-kwaliteitsmaat.

Dr. ing. J.G. Beerends studeerde elektrotechniek aan de gemeentelijke HTS te Den Haag en natuur- en wiskunde aan de Leidse universiteit (RUL). In de periode 1984-1989 verrichtte de heer Beerends een promotie-onderzoek op het gebied van akoestische perceptie aan het Instituut voor Perceptie Onderzoek te Eindhoven, waar hij in 1989 promoveerde op het onderwerp 'toonhoogtewaarneming in simultane complexe tonen'. Vanaf 1989 is dr. Beerends werkzaam bij PTT Research op het gebied van audio-kwaliteit. Samen met ir. J.A. Stemerink ontwikkelde hij een audio-kwaliteitsmaat die binnen de diverse standaardisatie-organen zoals CCITT en CCIR wordt ingebracht als een nieuwe meet-techniek om de subjectieve audio-kwaliteit te kunnen voorspellen aan de hand van objectieve metingen.

E.J.E. Koenderink trad in 1962 in dienst bij PTT, volgde de HPTD opleiding van het bedrijf en is sinds 1969 werkzaam bij PTT Telecom Netwerkbedrijf in het werkveld transmissie, voor een belangrijk deel op het gebied van AV-verbindingen.

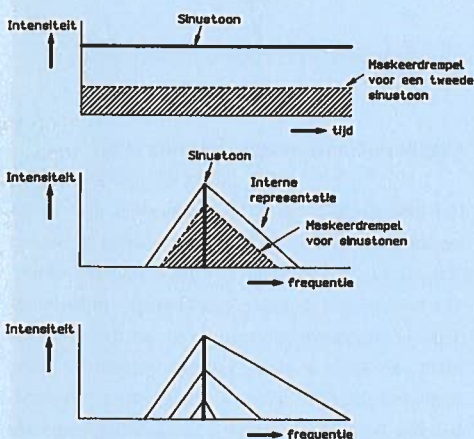
P. Morley B.A., B.E., M.Sc. studeerde wiskunde en elektrotechniek aan Trinity College Dublin. Van 1979 tot 1981 deed hij onderzoek op het gebied van digitale signaalbewerking en doceerde wiskunde bij het Dublin Institute of Technology. In 1981 trad hij in dienst bij Schlumberger International en werkte op boorinstallaties aan seismische, elektromagnetische en radio-actieve metingen t.b.v. gas- en olie-exploratie. In 1986 trad de heer Morley in dienst van PTT en werkte o.a. als recruitment officer bij KPN. Vanaf 1991 was hij als coördinerend projectleider binnen PTT Research betrokken bij het onderzoek naar spraakcodering en transmissiekwaliteit. Momenteel is de heer Morley binnen PTT Research adjunct-hoofd bij de hoofdafdeling Natuurwetenschappen.

Ir. P.T. Pont is afgestudeerd aan de TU Delft, faculteit elektrotechniek, vakgroep informatietheorie. Sinds 1990 is de heer Pont in dienst bij PTT Research als technisch wetenschappelijk medewerker binnen het werkveld audio.

Verdiepingsstof

Maskering

Om inzicht te krijgen in het maskeergedrag in het tijd- en frequentie-domein zijn veel experimenten uitgevoerd. Afbeelding 5 geeft op schematische wijze de resultaten weer voor wat betreft de frequentie-domein maskering van signalen²⁶. In de afbeeldingen is het maskerend vermogen weergegeven van een sinustoon (lokalisatie in het frequentie-domein). Een gelijktijdig klinkende sinus die binnen het gearceerd aangegeven gedeelte zit wordt niet waargenomen.



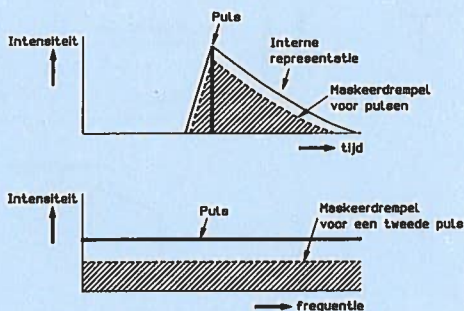
Afb. 7 a) maskeerdrempel van sinustoon in tijd-domein; b) idem in frequentie-domein; c) invloed intensiteit op interne representatie.

Deze maskeereigenschap kan worden geïnterpreteerd als het gevolg van een versmering van het geluidssignaal in het frequentie-domein. De sinustoon wordt in onze hersenen niet waargenomen als een scherpe piek. Het veroorzaakt een interne activiteit in een breed versmeerd frequentie-gebied. Samen met de eindige nauwkeurig-

heid waarmee veranderingen in de interne representatie worden waargenomen, geeft dit het in afbeelding 7b getekende driehoekige maskeerpatroon.

In afbeelding 7c is de interne representatie getekend van een sinustoon met drie verschillende geluidsniveaus. We zien dat de versmering van de interne representatie sterk afhankelijk is van het geluidsniveau²⁷. Vooral de rechterhelling van de interne representatie wordt er sterk door beïnvloed.

Er bestaat ook nog tijddomein-maskering, waarbij een maskeerder een signaal dat later (of zelfs eerder) kan maskeren²⁸. Afbeelding 8 geeft een geïdealiseerde weergave van het maskerend vermogen van een korte geluidspuls (lokalisatie in het tijddomein).



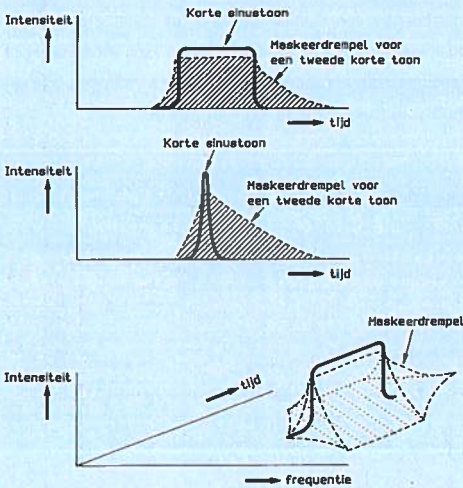
Afb. 8 a) maskerend vermogen van een korte geluidspuls in het tijd-domein; b) idem in het frequentie-domein.

Een correcte beschrijving van het maskeergedrag voor dynamische signalen is alleen te geven in een gecombineerd tijd/frequentie plaatje. Afbeelding 9 geeft een schematische weergave van beide effecten, tijd- en frequentie-domein maskering, voor een korte toonpuls²⁹ (gelijktijdige lokalisatie in tijd- en frequentie-domein).

Perceptief relevante meten

Hoe een perceptief relevante meting er in grote lijnen uitziet, is schematisch weergegeven in afbeelding 10.

Het muzieksignaal dat aan de ingang van het audio-apparaat wordt aangeboden, wordt met behulp van een model van het gehoorsysteem afgebeeld op een interne representatie, een representatie van het geluid die ook in onze hersenen aanwezig is. Het uitgangssignaal van het audio-apparaat wordt eveneens afgebeeld op deze interne representatie en vervolgens wordt het verschil in interne representatie gebruikt om de waarneembare degradatie in kwaliteit te voorspellen.

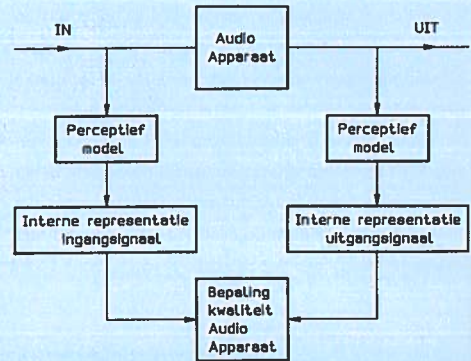


Afb. 9 a) tijd-domein maskering van korte sinustoon; b) idem in frequentie-domein; c) gecombineerde tijd-/frequentie-domein representatie.

Essentieel in deze meetmethode is dat zowel de in- als de uitgang van het audio-apparaat op de interne representatie worden afgebeeld, om vervolgens het verschil in interne representatie te gebruiken voor het definiëren van de kwaliteitsmaat.

Bij de klassieke meetmethoden wordt bijna altijd het verschilsignaal tussen de in- en uitgang met een model op een kwaliteitsmaat afgebeeld (bijvoorbeeld signaal/ruisverhouding). Vanwege het sterk niet-lineaire karakter van

de afbeelding van het signaal op de interne representatie, is het verschil in interne representatie echter niet gelijk aan de interne representatie van het verschil.

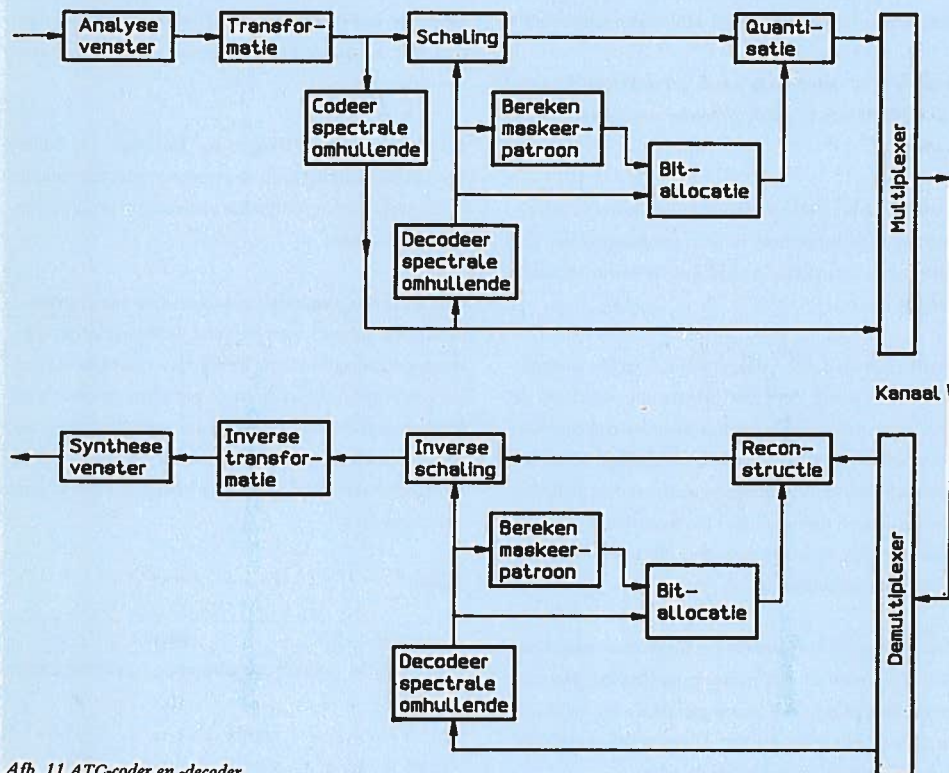


Afb. 10 Opstelling voor perceptief relevante meting.

Adaptieve Transformatie Codering (ATC)

Het basisprincipe van ATC is eenvoudig. Het audiosignaal wordt na de A/D-omzetter opgedeeld in blokken (frames) van bijvoorbeeld $N = 1024$ signaalmonsters. Met behulp van een wiskundige operatie wordt van elk frame het frequentie-spectrum berekend. Dit spectrum wordt vervolgens zo zuinig mogelijk gecodeerd (redundantie en irrelevantie worden verwijderd) en overgezonden. Aan de ontvangkant wordt het signaal gedecodeerd en via de inverse transformatie weer omgezet naar het tijd-domein. Als alles goed is gegaan hoort een luisteraar geen verschil tussen het originele ingangssignaal en het signaal aan de uitgang van de decoder. Perceptief gezien zijn beide signalen identiek.

ATC. Afbeelding 11 geeft een gedetailleerd blokschema van de ATC-coder en -decoder. In het navolgende zullen de verschillende blokken één voor één worden behandeld.



Afb. 11 ATC-coder en -decoder

Analyse-venster De coder bewerkt het signaal frame-voor-frame; het te coderen signaal wordt opgedeeld in tijdframes van bijvoorbeeld 30 milliseec. die afzonderlijk gecodeerd en overgezonden worden. Om een goede resolutie in het frequentie-domein te verkrijgen, dat wil zeggen een goed onderscheid te kunnen maken tussen signaal-componenten van verschillende frequenties, is het noodzakelijk elk signaalframe te vensteren (windo-venen) voorafgaand aan de transformatie. Het ingangssig-naal wordt hierbij vermenigvuldigd met een venster-functie (analyse-window) met geleidelijk aflopende flan-ken.

Transformatie De transformatie verzorgt de omzetting van het signaal uit het tijd-domein naar het frequentie-domein. Het frequentie-spectrum bestaat uit N reële ge-tallen (coëfficiënten), waarbij N gelijk is aan het aantal signaalmonsters in het frame.

Spectrale omhullende Van het frequentie-spectrum wordt een zogenaamde omhullende (enveloppe) bere-kend, een ruwe benadering van de vorm van het spec-trum. Deze omhullende wordt gebruikt als besturings-informatie bij het verdere codeer- en decodeer-proces en wordt als 'zij'-informatie meegezonden naar de decoder.

In de encoder wordt de omhullende gecodeerd en vervolgens weer gedecodeerd om zo over dezelfde omhullende informatie te beschikken als bij de decoder aanwezig is.

Schaling De omhullende wordt gebruikt om de spectraal-coëfficiënten te schalen voordat ze gekwantificeerd worden.

Bereken maskeerpatroon De omhullende wordt ook gebruikt voor de berekening van een maskeerpatroon. Dit geeft voor elke spectraal-coëfficiënt de ruismaskeringsdrempel aan.

Bit-allocatie Uit het maskeerpatroon en de spectrale omhullende wordt door het bit-allocatie algoritme de verdeling van de voor het huidige frame beschikbare hoeveelheid bits bepaald. Voor elke coëfficiënt wordt het aantal bits (de precisie) berekend waarmee die coëfficiënt gekwantificeerd moet worden. De bitverdeling moet zodanig zijn, dat de kwantificatiefout bij elke frequentiecomponent onhoorbaar is.

Kwantificeren Op grond van het toegewezen aantal bits worden de spectraal-coëfficiënten, na schaling, één voor één gekwantificeerd. Te zamen met de zij-informatie (de spectrale omhullende) worden de gekwantificeerde coëfficiënten overgezonden naar de decoder.

In de decoder wordt uit de binnengekomen zij-informatie de spectrale omhullende berekend. Vervolgens wordt, net zoals bij de coder, het maskeerpatroon berekend en de bitverdeling bepaald. De in de coder gekwantificeerde spectraal-coëfficiënten worden dan weer gereconstrueerd en invers geschaald. De inverse transformatie berekent uit het frequentie-spectrum de tijddomein golfvorm van het signaalframe. Na vermenigvuldiging met een venster-functie, het synthese-venster, worden de frames weer aan elkaar gepast.

Het synthese-venster heeft net als het analyse-venster geleidelijk aflopende flanken, waardoor scherpe overgangen tussen de frames worden vermeden. Door de frames

met een zekere overlapping bij elkaar op te tellen, en door de vorm van de venster-functies is het totale systeem van analyse-venster en synthese-venster transparant, met andere woorden het signaal wordt onvervormd doorgelaten.

Bit-reductie in de ATC-codec: een illustratie. De onderstaande beschrijving geeft de essentie van de bit-reductie bij de adaptieve transformatie codering op een inzichtelijke manier weer.

Om het proces van de bit-reductie in een ATC-codec te illustreren, introduceren we eerst de binaire getalrepresentatie signmagnitude. In de regel worden binaire getallen gerepresenteerd in de two's complement notatie (in fixed-point DSP's bijvoorbeeld), maar voor de uitleg is de signmagnitude notatie beter geschikt. Hieronder is in een tabel, voor 4 bits, de relatie tussen de twee stelsels weergegeven.

Tabel 1

| decimale signaalwaarde | binair | |
|---------------------------|------------------|----------------|
| | two's complement | sign/magnitude |
| 7 | 0111 | 0111 |
| 6 | 0110 | 0110 |
| 5 | 0101 | 0101 |
| 4 | 0100 | 0100 |
| 3 | 0011 | 0011 |
| 2 | 0010 | 0010 |
| 1 | 0001 | 0001 |
| 0 | 0000 | 0000 |
| -0 | - | 1000 |
| -1 | 1111 | 1001 |
| -2 | 1110 | 1010 |
| -3 | 1101 | 1011 |
| -4 | 1100 | 1100 |
| -5 | 1011 | 1101 |
| -6 | 1010 | 1110 |
| -7 | 1001 | 1111 |
| -8 | 1000 | - |

Bij sign-and-magnitude geeft het eerste bit het teken (sign) aan. Een 0 voor een positieve waarde en een 1 voor een negatieve waarde. De overige bits geven de grootte (magnitude) aan.

We nemen nu de 16-bits signmagnitude notatie

$$x = [s \ b_{14} \ b_{13} \ b_{12} \ b_{11} \ b_{10} \ b_9 \ b_8 \ b_7 \ b_6 \ b_5 \ b_4 \ b_3 \ b_2 \ b_1 \ b_0]$$

Hierin is de waarde van x gegeven door

$$x = (1 - 2s) \times b_i \cdot 2^i$$

$$= (1 - 2s) \times (b_0 + 2b_1 + 4b_2 + 8b_3 \dots + 16384b_{14})$$

s is hierin het tekenbit. De bits b_i representeren elk een signaalcomponent ter grootte 2^i .

Stel dat l de positie is van het eerste bit, afgezien van het tekenbit s , dat gelijk is aan 1. Dus

$$b_l = 1, \text{ en}$$

$$b_i = 0, \text{ voor } i = 14, 13, \dots, l+1$$

De bits $b_{14} \dots b_{l+1}$ zijn niet-significante nullen. De $l+1$ bits b_l, b_{l-1}, \dots, b_0 zijn de significante magnitude-bits. De niet-significante magnitude-bits geven geen informatie over het getal. Als bij ontvanger van tevoren bekend is hoeveel bits er niet-significant zijn, hoeven deze bits niet overgezonden te worden. Ze zijn redundant.

Stel verder dat we besluiten slechts $R-1$ significante magnitude-bits over te zenden. Dus bij gebrek aan meer bits worden alleen de R bits

$$s, b_l, b_{l-1} \dots b_m \text{ met } m = l - R + 2$$

gecodeerd. De overige m bits $b_{m-1} \dots b_0$ worden weggegooid (in feite is dit kwantificeren met stapgrootte 2^{m-1}).

De gekwantificeerde waarde noemen we \hat{x} en q de kwantificeerfout $q = x - \hat{x}$

De maximale fout die bij het afbreken gemaakt wordt volgt uit

$$|x - \hat{x}|_{\max} = q_{\max} = \frac{1}{2}2^m = 2^{m-1} = 2^{l-R-1}$$

Uit de berekende maskeerdrempel kan worden bepaald, hoe groot de kwantificeerfout (de kwantificeerruis) bij een bepaalde frequentie-component mag zijn. Als het getal x een te coderen frequentie-component voorstelt, is de kwantificeerfout q de kwantificeerruis bij die frequentie en deze ruis moet onder de maskeerdrempel liggen. De maskeerdrempel geeft dus aan hoeveel bits er mogen worden weggegooid. Het geeft de index-waarde m van het laatste mee te nemen bit. De overige m bits $b_{m-1} \dots b_0$ zijn irrelevant en mogen worden weggegooid. Een voorbeeld:

Tabel 2

| x decimaal | teken | redundant | gecodeerd | irrelevant |
|--------------|-------|-----------|-----------|------------|
| 2460 | 0 | 000 | 1001 | 10011100 |

Het aantal significante bits is op een eenvoudige manier gerelateerd aan de logaritme van de magnitude, $\log |x|$. Nemen we de logaritme met het grondtal 2 (hier aangegeven door \log_2), dan vinden we:

$$\lfloor \log_2 |x| \rfloor = l$$

waarin $\lfloor a \rfloor$ de naar beneden afgeronde waarde van a voorstelt. We kunnen heel eenvoudig de logaritmische amplitude van een signaal aflezen door de positie van de eerste 1 te bepalen.

Ook voor de kwantificeerfout q kunnen we een eenvoudige logaritmische benadering bepalen.

Omdat $q_{\max} = 2^{m-1}$, volgt $\log_2 q_{\max} = m - 1$

De positie van het eerste niet-redundante bit en van het laatste gecodeerde bit geven dus een logaritmische benadering van resp. de amplitude van het signaal en de amplitude van de kwantificeerruis.

In afbeelding 12 zijn de bits van de spectraal-coëfficiënten onder elkaar weergegeven in signmagnitude no-

tatie. De teken-bits zijn voor het gemak weergegeven als + of -. Als we nu de posities van de meest significante 1-en bekijken, krijgen we een ruwe indruk van het (logaritmische) amplitude spectrum. We kunnen nu een curve tekenen langs deze posities en krijgen zo een logaritmisch spectrum (na de afbeelding een kwartslag te draaien). Alle bits die boven deze curve liggen, behalve de teken-bits, zijn niet-significante nullen (leading zeros). Zij bevatten geen informatie over het signaal en zijn redundant. Bij het coderen kunnen we deze bits weglaten zonder dat dit invloed heeft op de signaalweergave. Wel moeten we dan op een of andere manier aan de decoder kenbaar maken welke bits zijn weggelaten, zodat deze de weggelaten 0-en weer kan invullen. Dit wordt gedaan door de curve mee te zenden als zij-informatie. Omdat er niet voldoende bits beschikbaar zijn om deze curve gedetailleerd mee te zenden wordt deze benaderd door een ruwe omhullende.

De irrelevantie wordt als volgt verwijderd. Uit de ruwe omhullende van het frequentie-spectrum wordt de maskeerdrempel berekend. Deze geeft aan hoe groot de codeer-fout in het spectrum bij elke frequentie mag zijn, zonder dat deze fout hoorbaar is. Hieruit is af te leiden

met hoeveel bits de spectraal-coëfficiënten gecodeerd moeten worden. De maskeerdrempel is ook in afbeelding 9 weergegeven. De bits die onder de drempel liggen zijn irrelevant. Zij geven aanleiding tot een onnodige precisie in de weergave van de amplitude bij die frequentie. Ze kunnen weggelaten worden, weer onder de voorwaarde dat de decoder uit de zij-informatie kan afleiden welke bits er weggelaten zijn.

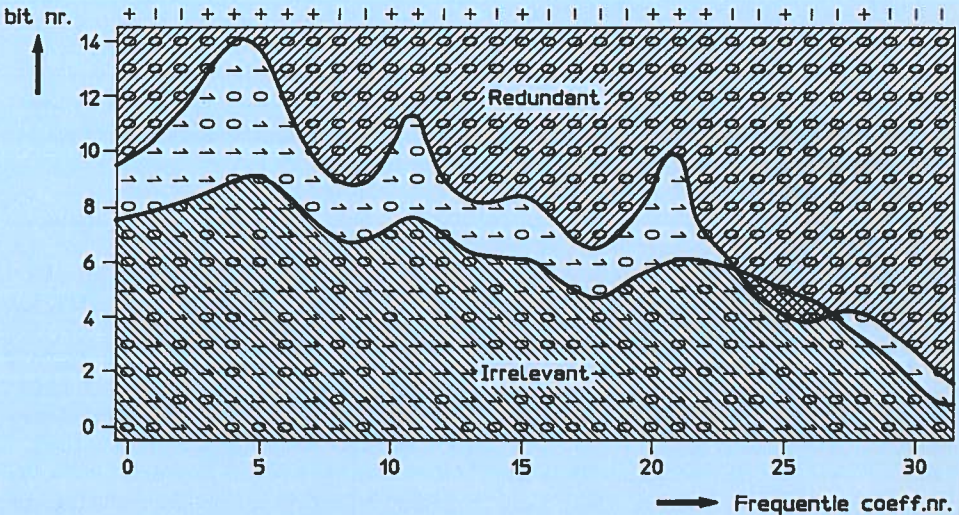
²⁶ B.C.J. Moore, *An introduction to the Psychology of Hearing*, Academic Press, 1982.

B. Scharf & S. Buus, Stimulus, physiology, thresholds, in: L. Kaufman, K.R. Boff and J.P. Thomas (ed.), *Handbook of Perception and Human Performance*, Wiley, New York, 1986.

²⁷ E. Zwicker & R. Feldtkeller, *Das Ohr als Nachrichtenempfänger*, S. Hirzel Verlag, Stuttgart, 1967.

²⁸ B. Scharf & S. Buus, Stimulus, physiology, thresholds, in: L. Kaufman, K.R. Boff and J.P. Thomas (ed.), *Handbook of Perception and Human Performance*, Wiley, New York, 1986.

²⁹ H. Fastl, Temporal masking effects: II. critical band noise masker, in: *Acustica*, 36 (1976), pp. 317-331.



Afb. 12 Bit-reductie in een frame

Studieblad kort

PTT Telecom bouwt landelijk net voor aanbod tv- en radiosignalen

PTT Telecom bouwt een hoogwaardig, digitaal netwerk voor het aanbod van televisie- en radiosignalen aan kabelexploitanten. Dit Breedband VideoNet (BVN) is gebaseerd op bestaande en nieuwe glasvezelverbindingen.

Het netwerk, dat in fasen wordt gebouwd, zal eind 1994 geheel operationeel zijn.

Op dat moment kan PTT Telecom door heel Nederland signalen leveren van een groot aantal tv-stations (waarvan een aantal op studio-kwaliteit, zoals Nederland 1, 2 en 3) en radiostations (waaronder digitale satellietradio, DSR). In eerste instantie gaat het om een assortiment van 24 televisie- en 30 radiokanalen. Ook nieuwe diensten van programma-aanbieders behoren tot de mogelijkheden. De kabelexploitanten bepalen zelf of zij al of niet deze signalen afnemen en in een pakket dat naar eigen keuze kan worden samengesteld.

Deze plannen heeft PTT Telecom voorgelegd aan het bestuur van de Vecai, de Vereniging van kabelexploitanten en machtiginghouders in Nederland. Binnenkort bespreekt PTT Telecom met de grote (regionale) exploitanten op het gebied van kabeltelevisie haar voornemens om bij hen de belangstelling te peilen voor de dienstverlening die zij over het BVN aan hen wil aanbieden en de wijze van samenwerking met hen.

- **Keuzevrijheid:** alternatief voor investeren PTT Telecom haakt met het BVN in op de ontwikkeling van de technische mogelijkheden en sluit tevens aan bij de wens van de exploitanten van kabeltelevisie om zich regionaal te organiseren. Die exploitanten van grote regionale en stadsnetten staan voor grote investeringen op het gebied van hun infrastructuur voor het ontvangen en leveren van kabelsignalen. Het gaat daarbij zowel om vervanging als verbetering van hun netten, maar ook om vergroting van de

capaciteit. PTT Telecom biedt met BVN de kabelexploitanten hiervoor een alternatief. Een alternatief, dat een aanzienlijke verbetering omvat van de kwaliteit van signalen die worden geleverd, zowel van de aardse zendstations als van de satellietzenders.

PTT Telecom brengt met BVN standaardisatie aan in de dienstverlening op het gebied van signaalleverantie, onder voor ieder gelijke condities en met een heldere tariefstructuur. Daarbij staat de keuzevrijheid van de exploitant voorop; die kan zoveel signaal afnemen als hij zelf wil. PTT Telecom biedt de signalen aan in een open concurrentie met andere aanbieders.

PTT Telecom zal de signalen leveren aan wat zij ziet als kernexploitanten op het gebied van kabeltelevisie. Zij kunnen op hun beurt de signalen doorleveren aan lokale exploitanten.

- **Aanbod van programma's:** groot assortiment Naast de levering van een groot assortiment bestaande kanalen, tegen een hogere kwaliteit dan nu, ziet PTT Telecom ook mogelijkheden voor de doorgifte van andere programma's. Daarbij moet bijvoorbeeld worden gedacht aan regionale televisie en nationale kabelprogramma's, zoals bijvoorbeeld een kabelkrant. BVN biedt bovendien de mogelijkheid tot het programmeren van regionale reclameblokken in landelijke programma's. Daarnaast kan PTT Telecom via het BVN bepaalde signalen van regionale aanbieders inkopen om die elders door te leveren.

- **Hoogwaardig net** Het Breedband VideoNet is gebaseerd op de bestaande telecommunicatie-infrastructuur van PTT Telecom in Nederland. Hiernaast zal PTT Telecom, mede ten behoeve van het BVN, een deel van haar zogenaamde aansluitnet van glasvezelkabel voorzien. Tussen 1993 en 2002 wordt in 128 gemeenten met meer dan 10.000 inwoners dit deel van het telecommunicatienetwerk 'verglasd'.

Geavanceerde transmissietechniek, in combinatie met specifiek netwerkmanagement, geven BVN een hoge betrouwbaarheid en een goede

toekomstvastheid. Dubbele uitvoering van het-net en een dubbel aantal kopstations zorgen verder voor een zeer hoge beschikbaarheid.

PTT Telecom Studieblad heeft aan het BVN reeds eerder aandacht besteed in het themanummer Van draadomroep tot breedbandnet (december 1992, pp. 699-732)

(Bron: Persbericht PTT Telecom, 011/1993)

Boekbespreking

Titel: *ISDN en privacy*

Auteur: J.M.A. Berkvens e.a.; onder red. van P. van Hoogstraten

Amsterdam: Cramwinckel, 1992

78 p.

ISBN 90-7189-436-3

Op 28 november 1991 organiseerde PTT Telecom een conferentie over de privacy-aspecten van ISDN. De conferentie werd gehouden in het kader van een intensiever wordende interactie tussen PTT Telecom en de samenleving. Dit boek bevat een bewerking van de voordrachten die er tijdens de conferentie te beluisteren vielen.

- De eerste bijdrage is van P. van Hoogstraten en is een algemene inleiding betreffende ISDN en privacy. Aan de orde komen o.a. gidsinformatie, call records, automatische nummerherkenning en rekeninginformatie.

- In de tweede bijdrage bespreekt J.E. Katz de problemen rond nummeridentificatie in de Verenigde Staten. In de VS worstelt men met de vraag of en hoe het weergeven van het telefoonnummer van de opbeller op het toestel van de gene die gebeld wordt (calling line identificatie, CLI) in telecommunicatiediensten vorm moet krijgen. In de vijftig staten variëren de oplossingen van een rigide afwijzing om redenen van privacy tot een onvoorwaardelijke invoering.

- D. Garbe gaat in op gegevensbescherming en ISDN in enkele Europese landen. Frankrijk, Engeland en Duitsland hebben een privacywetgeving die ook op telecommunicatie van toepassing is. Een van de belangrijkste problemen m.b.t. ISDN en privacy heeft betrekking op de opslag van verkeersgegevens of verbindingsdata en de verrekening van kosten die op deze gegevens gebaseerd zijn. De oplossingen voor dit probleem, voor de gespecificeerde rekening en voor automatische nummerherkenning in Frankrijk, Engeland en Duitsland worden besproken.

- De heer Berkvens probeert antwoord te geven op de vraag wanneer het recht op privacy zwaarder weegt dan het controle- en inzage-recht. Hij baseert zich hierbij op de Nederlandse wetgeving.

- J. Holvast gaat in op het begrip privacy, de beleving van privacy en de knelpunten bij privacy. Voorts worden de gevaren voor de aantasting van de privacy besproken voor consumenten telediensten (bijv. telebankieren, electronic mail, de Telegids) en voor diensten van PTT Telecom.

- J.J.C. Kabel bespreekt ten slotte artikel 15 van de Wet Persoonsregistraties (WPR). Dit artikel roept bedrijfstakingen op tot zelfregulering bij de behandeling van persoonsgebonden gegevens. De wet zelf fungeert als stok achter de deur: het nalaten van zelfregulering kan strafbaar worden gesteld.

Dit boek is geschikt voor iedereen die geïnteresseerd is in de privacy-aspecten van nieuwe telecommunicatiediensten.

(Deze boekbespreking is samengesteld door Genevieve Geppart, BIDATA techniek, in opdracht van de redactie van PTT Telecom Studieblad. PTT-medewerkers kunnen het boek onder vermelding van BIDATA-kenmerk 828184 lenen bij: Koninklijke PTT Nederland, BIDATA, Kamer D 275, Postbus 30.000, 2500 GA Den Haag, tel. 070-3323172)