

De betekenis van het begrip HiFi (updated english version [here](#))

Deel 1, Transparantie en Perceptieve Meettechnieken

Met de introductie van audiocompressie (MP3, AAC, WMA) aan de ene kant en Super Audio CD aan de andere kant van het HiFi spectrum kun je jezelf afvragen wat betekent HiFi tegenwoordig nog? Als we dan ook nog eens worden geconfronteerd met tientallen “sound enhancers” op onze MP3 spelers en geluidskaarten raken we helemaal de draad kwijt en gaan sommigen zelfs weer terug naar de LP. Tegenwoordig is alles mogelijk en dus moeten we onszelf afvragen “wat willen we eigenlijk?”. Voordat we die vraag kunnen beantwoorden moeten we beginnen met de betekenis van het begrip “High Fidelity” wat in de zestiger jaren wel als “werkelijkheidsweergave” werd vertaald en nu misschien beter als “natuurgetrouwheid”. Maar willen we dat, natuurgetrouwheid? En als we dat al zouden willen, hoe kun je dan meten of iets natuurgetrouw wordt weergegeven. Moeilijke vragen die ik in deze serie van acht artikelen ga beantwoorden.

We beginnen met het begrip transparantie. In technische zin moet veel HiFi apparatuur transparant zijn, dwz er moet niets aan het geluid worden toegevoegd en er moet ook niets worden weggelaten. Bij veel HiFi apparatuur gaat het signaal er elektrisch in, komt er weer elektrisch uit, en kunnen we simpelweg het signaal wat we erin stoppen aftrekken van wat eruit komt, als de uitkomst nul is, eventueel na amplitude schaling, dan is het apparaat perfect. Is het verschil niet nul dan kan het nog steeds onhoorbaar zijn en dus kan het apparaat, vanuit perceptief standpunt gezien, nog steeds perfect transparant zijn. Voor bijna alle moderne audio apparatuur met een elektrische in/uitgang geldt dat ze nagenoeg perfect (kunnen) zijn. Opslag, transport en versterking zijn met de introductie van de compact disc (CD), met zijn 44.1 kHz/16 bit digitale signaal representatie, geen probleem meer en betere signaal representaties, met hogere sample frequenties en/of meer bits per sample, leveren geen significante verbetering op [1]. Ondanks de “voodoo” audio verhalen is de CD echt een nagenoeg perfect opslag medium, in ieder geval veel beter dan de LP . Voor mensen die dat niet geloven, die zijn er nog steeds, vooral onder muzikanten geef ik twee argumenten.

Als eerste een wiskundig argument gerelateerd aan de 44.1 kHz sampling en 16 bit kwantisatie van tijd en amplitude zoals die wordt gebruikt bij de CD. Welke vervormingen introduceert dit en hoe verhouden deze zich tot vervormingen die we kennen van analoge apparatuur? Voor wat betreft de sampling van de tijd-as zegt het theorema van Nyquist-Shannon dat bij een sample frequentie die twee maal zo hoog is als de hoogst weer te geven frequentie het signaal exact te reconstrueren is. Dat betekent dus dat als we de sample frequentie maar hoog genoeg kiezen de kwantisering van de tijdschaal geen problemen oplevert. Omdat er geen mens is die nog geluiden kan horen boven de 25 kHz moet een sample frequentie rond de 50 kHz genoeg zijn voor perfecte audio kwaliteit. Voor wat betreft de kwantisering van de amplitude schaal ligt de zaak wat moeilijker omdat we hier te maken krijgen met een statistische identiteit tussen het analoge en digitale domein. Het is met een wiskundige techniek, bekend onder de naam dithering, mogelijk om een digitale representatie te maken waarvan het foutsignaal een ruis signaal is waarvan de eigenschappen gelijk zijn aan die van analoge ruis zoals die is te vinden op tape en plaat (ik zou het Roberts- Lipschitz dithering willen noemen). Voldoe je aan het Nyquist-Shannon criterium en gebruik je de juiste Roberts- Lipschitz dithering, dan heeft het digitale signaal exact dezelfde eigenschappen als een band begrensd analogoog signaal waar wat

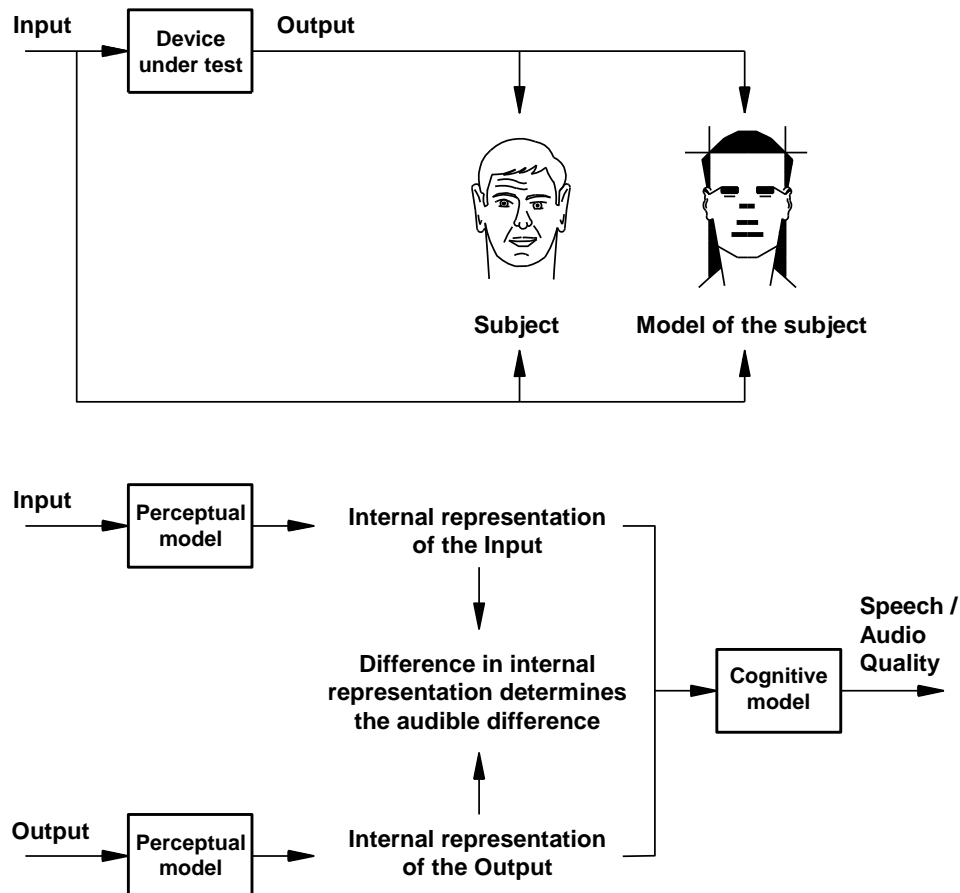
ruis op zit. Deze analyse toont aan dat er een exacte equivalentie bestaat tussen het digitale domein en het analoge domein. Als we een analoge representatie van het signaal maken met dezelfde kwaliteit als die van een CD hebben we een audio bandbreedte van 20 kHz nodig met een signaal ruisafstand van ca. 95 dB. Zo'n prachtige ouderwetse vinyl plaat kan misschien nog wel 20 kHz weergeven, maar de signaal ruisafstand komt niet veel verder dan een dB of 60 (afhankelijk van hoe je het precies meet). En als je nog niet overtuigd bent luister dan naar een LP met pianomuziek en huiver van de zweving (wow en flutter) en zet ter vergelijking een CD op met zijn perfect strakke piano aanslagen met onhoorbare zweving.

Als tweede argument om aan te tonen dat de CD superieur is kun je die ideale plaat opnemen op een CD en vervolgens naar het verschil luisteren tussen de plaat en de CD opname van die plaat. Ik kan garanderen dat het waarneembare verschil zo goed als nul is. Als je echter de CD op een plaat opneemt en luistert naar het verschil zul je onmiddellijk ruis, spetters, krassen., rumble en zweving kunnen horen. Het zal duidelijk zijn dat je een CD wel kunt laten klinken als een ouderwetse vinyl plaat maar dat je het niet lukt om het vinyl te laten klinken als een CD. Het transparantie ideaal is dus met een CD veel beter te benaderen dan met een ouderwetse plaat, sterker nog als we signaal representaties gebruiken die nog beter zijn dan het CD formaat, met meer bits en/of hogere sample frequenties, blijkt dat dit geen significante verbetering oplevert [1].

Als we de beschikking hebben over een perfect opslagmedium kunnen we de ideale transparantie test uitvoeren. Speel een set audio signalen, opgeslagen op het perfecte medium, af over het te testen audio apparaat en luister naar verschillen tussen het signaal wat erin gaat en het signaal wat eruit komt. Hoor je bij geen enkel signaal verschil dan is het apparaat voor jouw perceptief transparant. Je moet zo'n test wel dubbel blind uitvoeren, zowel de proefpersoon als de begeleider van het experiment mogen niet op de hoogte zijn wat er gespeeld wordt. Ook moeten de signaal niveaus binnen 0.2 dB gelijk worden gemaakt om identificatie op grond van een verschil in luidheid te voorkomen. Verder moet de proefpersoon naar keuze kunnen luisteren naar het ingangssignaal om vervolgens met een synchrone loop te kunnen bepalen wat van de twee aangeboden signalen de ingang is en wat de uitgang.

Een goed alternatief voor een duur en moeilijk uit te voeren subjectief experiment is het gebruik van een perceptieve meettechniek. We maken met een perceptief model een afbeelding van het audio signaal zoals dat in onze hersenen beschikbaar is, een interne representatie van het geluid. Dat doen we voor het ingangssignaal en het uitgangssignaal. Als we vervolgens die afbeeldingen van elkaar aftrekken en het verschil is nul, dan is het geteste apparaat perceptief transparant. Als het verschil niet nul is moeten we met een cognitief model dit verschil interpreteren en afbeelden naar een kwaliteitsmaat die overeenkomt met de ervaren kwaliteit door een proefpersoon (zie Figuur 1). Dat lijkt moeilijk, maar een simpele oplossing van dit probleem voor het meten van spraakkwaliteit, die ik begin 90-er jaren bij KPN Research heb ontwikkeld [1], is door de International Telecommunication Union (ITU) geëvalueerd waarbij correlaties tussen subjectieve en objectieve resultaten op onbekende data van boven de 0.97 werden gehaald. Deze methode, [PSQM](#) [2], is in 1996 geaccepteerd als ITU Recommendation P.861, de eerste wereldstandaard voor het meten van spraakkwaliteit. Na een aantal uitbreidingen is dit model in 2001 en 2011 opnieuw geaccepteerd als wereldstandaard P.862 [PESQ](#) [3] [4] [5] (voor telefoonband spraak)

en P.863 [POLQA](#) [6] [7] [8] (voor HiFi spraak). Tevens is een aangepaste versie gebruikt als basis voor het ontwikkelen van een audiokwaliteitsmaat voor muziek compressie [9] [10] [11] (ITU Recommendation BS.1387, [PEAQ](#), 1998). Helaas is er vanuit de commercie weinig belangstelling voor het meten van HiFi audiokwaliteit. Voor het meten van spraakkwaliteit is in 2011 de 3^e generatie meetmethode gestandaardiseerd (P.863 POLQA) terwijl er voor het meten van HiFi muziekkwaliteit geen updates zijn geweest gedurende de laatste 15 jaar, de 1^e generatie meetstandard (BS.1387 PEAQ) is ondanks zware tekortkomingen nooit meer verbeterd.



Figuur 1. Basisopzet voor de perceptieve meettechniek. Boven: equivalentie tussen een subjectieve test en een objectieve perceptieve meting met behulp van een model van de proefpersoon. Onder: de interne representatie van de ingang wordt vergeleken met de interne representatie van de uitgang met behulp van een cognitief model. Als het verschil in interne representatie nul is dan is het geteste apparaat perceptief transparant (perfecte reproductie).

Met het basis begrip transparantie, en de daaruit afgeleide symmetrische transparantietest, kun je al veel audio mythes te lijf maar transparantie is een te simpel ideaal om de vraag mee te beantwoorden of iets goed klinkt. Moet, of liever gezegd kan, een luidspreker transparant zijn?

[1] E.B. Meyer and D. R. Moran, "Audibility of a CD-Standard A/D/A Loop Inserted

into High-Resolution Audio Playback," J.Audio Eng. Soc., vol. 55, pp. 775-779, (2007 Sep.).

[2] J. G. Beerends and J. A. Stemerding, "A Perceptual Speech Quality Measure Based on a Psychoacoustic Sound Representation," J. Audio Eng. Soc., vol. 42, pp. 115-123, (1994 March).

[3] ITU-T Recommendation P.861, Objective Quality Measurement of Telephoneband (300-3400 Hz) Speech Codecs (1996 Aug.).

[4] A. W. Rix, M. P. Hollier, A. P. Hekstra and J. G. Beerends, "PESQ, the New ITU Standard for Objective Measurement of Perceived Speech Quality, Part I - Time Alignment," J. Audio Eng. Soc., vol. 50, pp. 755-764 (2002 Oct.).

[5] J. G. Beerends, A. P. Hekstra, A. W. Rix and M. P. Hollier, "PESQ, the New ITU Standard for Objective Measurement of Perceived Speech Quality, Part II - Perceptual Model," J. Audio Eng. Soc., vol. 50, pp. 765-778 (2002 Oct.).

[6] ITU-T Rec. P.862, "Perceptual Evaluation Of Speech Quality (PESQ): An Objective Method for End-to-end Speech Quality Assessment of Narrow-band Telephone Networks and Speech Codecs," International Telecommunication Union, Geneva, Switzerland (2001 Feb.).

[7] J. G. Beerends, C. Schmidmer, J. Berger, M. Obermann, R. Ullman, J. Pomy and M. Keyhl, "Perceptual Objective Listening Quality Assessment (POLQA), The Third Generation ITU-T Standard for End-to-End Speech Quality Measurement Part I – Temporal Alignment," J. Audio Eng. Soc., submitted 2011 December.

[8] J. G. Beerends, C. Schmidmer, J. Berger, M. Obermann, R. Ullman, J. Pomy and M. Keyhl, "Perceptual Objective Listening Quality Assessment (POLQA), The Third Generation ITU-T Standard for End-to-End Speech Quality Measurement Part II – Perceptual Model," J. Audio Eng. Soc., submitted 2011 December.

[9] ITU-T Rec. P.863, "Perceptual Objective Listening Quality Assessment," Geneva, Switzerland (2011 Jan.).

[10] J. G. Beerends and J. A. Stemerding, "A Perceptual Audio Quality Measure Based on a psychoacoustic sound representation," J. Audio Eng. Soc., vol. 40, pp. 963-978 (1992 Dec.)

[11] T. Thiede, W. C. Treurniet, R. Bitto, C. Schmidmer, T. Sporer, J. G. Beerends, C. Colomes, M. Keyhl, G. Stoll, K. Brandenburg, B. Feiten, "PEAQ - The ITU-Standard for Objective Measurement of Perceived Audio Quality," J. Audio Eng. Soc., vol. 48, pp. 3-29 (2000 Jan./Feb.).

[12] ITU-R Rec. BS.1387, "Method for Objective Measurements of Perceived Audio Quality," International Telecommunication Union, Geneva, Switzerland (2001 Nov.).

John G. Beerends

Gepubliceerd in Hifi Video Test 10/2007, herzien april 2012.