

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

Deel 8, Telefonie

Ik ben bang dat de titel van dit laatste deel over HiFi een beetje afschrikt. Telefonie, is dat niet de bodem in de wereld van HiFi? Iedereen kent het afgeknepen, bandbegrensd (300-3400 Hz), geluid van een stem over de klassieke telefoonverbinding. Feit is wel dat de “International Telecommunication Union” (ITU) een heel duidelijk ideaal heeft gedefinieerd. En we weten inmiddels dat je kwaliteit alleen kunt meten als je het ideaal kent. Voor onze HiFi wereld hebben we daartoe twee verschillende, onverenigbare, idealen gedefinieerd, het “hier en nu” versus het “daar en toen”. Een opname-weergave keten ontworpen voor “hier en nu” kan nooit een natuurgetrouwe “daar en toen” illusie creëren. Binnen de ITU is de ideale telefoonlink gedefinieerd als een link die overeenkomt met de situatie dat twee mensen op ca. 1 meter van elkaar in een stille, nagalm-arme, kamer met elkaar praten. Deze definitie ligt ergens tussen onze HiFi idealen in en is tot op grote hoogte goed realiseerbaar, je neemt aan beide kanten van de link een microfoon vlak bij je mond en twee luidsprekers vlak bij je oor en je komt een heel eind. Dat moet tegenwoordig toch technisch geen probleem meer zijn.

Maar er zitten twee adders onder het gras. De eerste is de opname/weergave kant waar een slechte akoestiek (galmende kamers) en “natuurlijke” achtergrondruis (vooral bij mobiel bellen) vaak storend werken. Niemand haalt het in zijn hoofd om een HiFi opname van een zanger te maken in een auto, naast een autosnelweg of in een drukke winkel. Maar van een telefoonlink eisen we dat hij onder alle omstandigheden goed werkt. Nu is de contradictie dat onze oude, grote maar handzame, telefoons perfect in en uitkoppelen, vlak bij de mond met een goede microfoon en vlak bij het oor met een goede luidspreker. Bij moderne mobieltjes is die koppeling veel slechter en is ook de luidspreker vaak te klein, terwijl we bij mobiel bellen meer achtergrond lawaai hebben en juist dan veel baat hebben bij een goede akoestische koppeling. Met geavanceerde signaalprocessing (ruisonderdrukking) wordt dan een poging ondernomen om de tekortkomingen van de te kleine handset weer goed te krijgen.

De tweede adder onder het gras zit hem in het feit dat telefonie niet alleen passief luisteren is zoals normale HiFi. Bij telefonie wordt ook nog actief gesproken, en als je praat hoor je jezelf, en het liefst zo goed mogelijk. Dat is in de ITU referentie conditie geen probleem, je hoort jezelf op natuurlijke wijze.

Bij een telefoon krijg je te maken met twee problemen bij het spreken:

- 1) mijn stem wordt aan de andere kant van de link opgepikt door de microfoon en ik kan mezelf terughoren over de link,
- 2) de luidspreker aan mijn oor schermt het natuurlijke, akoestische, pad af en als ik in een lawaaierige omgeving probeer de ander te verstaan door strak te koppelen aan mijn oor hoor ik mijn eigen stem niet meer goed.

In de klassieke telefonie zijn beide spreek problemen op een natuurlijke, bijna triviale manier opgelost. Ik koppel een deel van mijn microfoon signaal terug naar mijn eigen oor en ik houd de vertraging op de link zo laag dat het signaal wat terugkomt alleen bijdraagt aan het goed kunnen horen van mijn eigen stem. In combinatie met onze ouderwetse, goede, akoestische koppeling kun je een werkelijk fantastische, natuurlijk klinkende, telefoonlink bouwen die zelfs bij achtergrondlawaai uitstekend klinkt.

Waarom klinken onze moderne telefoons zo slecht? Gedeeltelijk door de toename van het mobiel gebruik met kleine handsets in lawaaierige omgevingen. Maar er is meer aan de

hand want ook ons huidige vaste net met zijn klassieke telefoondienst levert in toenemende mate slechte kwaliteit. Het is inmiddels zover dat het ouderwetse analoge telefoonnet (dat al bijna niet meer bestaat) als gouden standaard wordt gezien waar iedere service provider naar streeft, maar die niemand meer waar kan maken. Dat komt door onze nieuwe digitale pakkettechnieken die voor extra vertraging zorgen op de lijn. Nu is vertraging op zich niet zo'n groot probleem, waardes tot ca 100 ms (1-weg van A naar B) zijn bijna niet waarneembaar. Maar als je je eigen stem vertraagd terug hoort is dat een ramp, zo erg zelfs dat je bij bepaalde vertragingen en luidheden niet meer normaal kunt spreken. Nu is echo met moderne signaalprocessing goed te onderdrukken maar vooral bij de aanwezigheid van variërende achtergrondruis zijn er vaak storende bijgeluiden te horen. Bekend probleem is het aan- en uitschakelen bij achtergrondruis, iedere keer als ik spreek hoor ik de achtergrondruis van de andere kant schakelen.

Om te komen tot een HiFiPhone moeten we zorgen dat alle drie aspecten van de gesprekskwaliteit perfect zijn:

- 1) de luisterkwaliteit (passief), hoe goed hoor ik de ander, voornamelijk bepaald door ruis en vervorming
- 2) de spreekkwaliteit (actief), hoe goed hoor ik mezelf, voornamelijk bepaald door echo maar ook door achtergrondruis variaties
- 3) en als laatste de interactiekwaliteit (actief), hoe goed kunnen de gesprekspartners elkaar onderbreken en wat gebeurt er als we tegelijkertijd praten.

Dat laatste aspect van de gesprekskwaliteit, de interactiekwaliteit wordt een steeds groter probleem door de toenemende vertraging in het netwerk. Tegenwoordig is niet de snelheid van een digitale verbinding de bottleneck maar de vertraging. Omdat een digitaal netwerk alleen efficiënt is als we de spraak in pakketten transporteren, krijgen we in alle links meer vertraging en, vooral bij koppeling met het oude netwerk, echo. Ook bij heel kleine vertragingen kunnen er al problemen ontstaan (kamfiltering) en naarmate de vertraging toeneemt, wordt het steeds moeilijker om een hoge gesprekskwaliteit te garanderen.

Het mooie van de nieuwe ontwikkelingen is wel dat we op een heel simpele manier de luisterkwaliteit kunnen opschroeven van "benepen telefoon geluid (300-3400 Hz)" naar "CD kwaliteit (20-20000 Hz)". Wel is het zo dat naarmate we de audiobandbreedte opvoeren de echo en achtergrondruis problemen steeds moeilijker in de hand zijn te houden. Onder de 100 Hz en boven de 7 kHz zit vrijwel geen energie in het spraaksignaal terwijl stoorsignalen daar nog ruim vertegenwoordigd zijn. Het onderdrukken van achtergrondlawaai en echo over de volle audiobandbreedte is dan ook vrijwel onmogelijk. Een HiFiPhone?, ondanks de dappere pogingen met HD Voice zal het er wel nooit van komen.

Dr. John Beerends

Gepubliceerd in Hifi Video Test 8/2008, herzien 1/2012