

Aantekening Vocoding (Sound Design H1)

Bij vocoding gaat het om twee audio-signalen: bijv. stem en gitaar.

Om met het uiteindelijke, resulterende effect te beginnen: de gitaarklank wordt "in de mond gelegd" alsof het de stembanden waren; zo lijkt het alsof je de gitaar hoort praten.

In Sound Design H1 wordt bij vocoding uitgelegd dat spraak ontstaat doordat een toongenerator (stembanden) klinkt in een akoestische ruimte (spraakkanaal). Omdat elke akoestische ruimte de werking heeft van een filter kun je verstaanbare spraak dus opvatten als afkomstig van een filter waarvan de eigenschappen voortdurend veranderen.

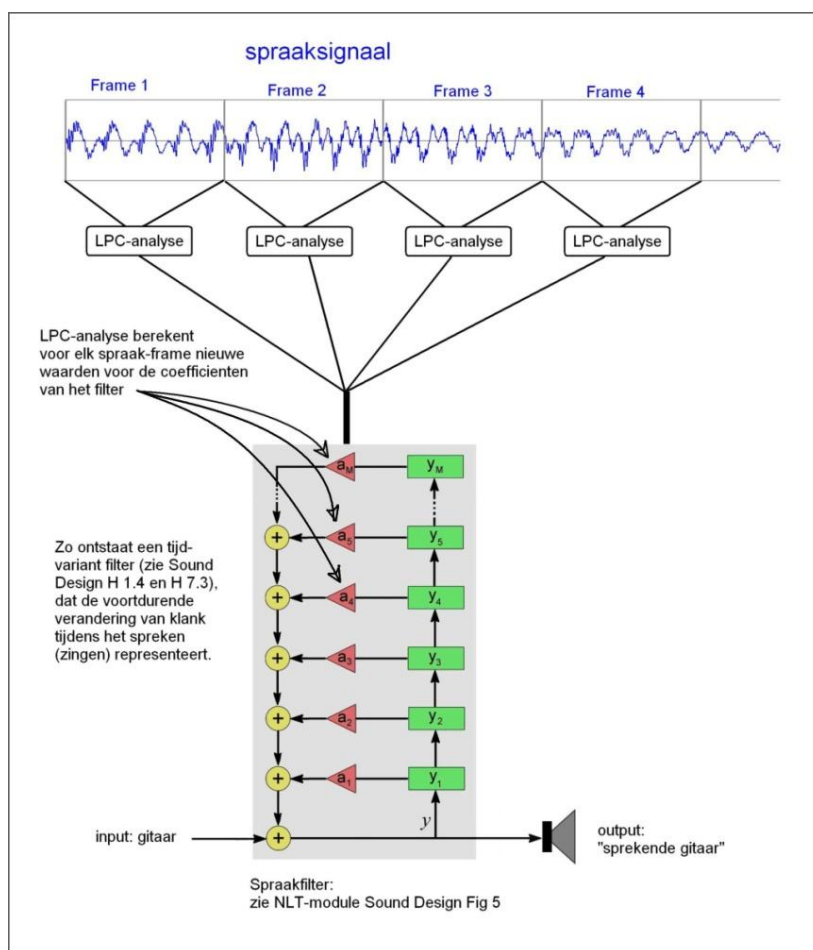
Bij vocoding wordt een (real time) stem-analyse gemaakt. Het spraak/zangsignaal wordt opgedeeld in fragmentjes van ongeveer 15 - 50 msec. Elk fragmentje wordt geanalyseerd met een formant-tracking algoritme. Daarvoor bestaan globaal twee methodes die alletwee in H1 worden genoemd of kort besproken:

(1)

Fourier-gebaseerd, bijvoorbeeld (a) "Vrouwenstem verandert in mannenstem" van Sound Design H1 paragraaf 2. Dat hier sprake is van een Fourier-techniek blijkt uit wat Carla er zelf over zegt: (mix van 128 sinussen). Het geluid van Carla's stem levert een reeks spectra op, precies zoals je die ziet in de Spectrum Analyzer van WaveWizard. Ook van de gitaar wordt een spectrumreeks berekend. Vervolgens worden beide spectra, frame voor frame, met elkaar vermenigvuldigd. Zo ontstaat een filterwerking, want een filterwerking is een vermenigvuldiging "in het frequentiedomein" - zoals besproken in H5.

(2)

Filter-gebaseerd. Zie ook pag. 15-18. Het doel is om de samples van een spraakfragment om te zetten in een set filtercoëf's zodat het "eigengeluid" (pulsresponsie!) van het filter dezelfde klank oplevert als die van het spraakfragment. In voorbeeld "Maan, zaag, Fien" wordt dit toegelicht. Deze techniek heet Linear Predictive Coding (LPC). In de eerste seconden van "Maan, zaag,..." hoor je de pulsresponsies van een aantal afzonderlijke filters. Het effect hiervan kun je zelf gemakkelijk simuleren door tegen je wang te tikken en daarbij je mond geleidelijk in verschillende spraakstanden te zetten. Dan fungeert je vinger als "pulsgenerator" en het geluid dat je hoort zijn de pulsresponsies van de afzonderlijke mondstanden. Omdat de pulsen in "Maan, zaag..." elkaar steeds sneller opvolgen begin je geleidelijk een toon te horen en ontstaat vloeiende spraak. Dus in "Maan, zaag,..." ontstaat spraak door een reeks elkaar opvolgende filters "aan te stoten" (input) met pulsen. In plaats daarvan kun je ook een gitaarsignaal door de filterreeks sturen. Dan krijg je het effect van een sprekende gitaar.



In bijgaande plaatje wordt de LPC-vocoder nader toegelicht.

Er komt een vervolgmodule op Sound Design. Daarin gaan we nader in op vocoders, LPC, spraak- en zangsynthese.