Genelec Aural ID

Individuelle Erstellung einer HRTF auf Basis von Fotos und Video

Autor: Peter Kaminski



Wir haben ja in verschiedentlich über immersives Audio, bzw. 3D-Audio berichtet und auch schon mehrfach über das Themengebiet der Binauralisierung, bzw. des Hörens von immersivem Audio über Kopfhörer. Hierzu ist eine sogenannte HRTF (Head Releated Transfer Function) erforderlich. Es gibt hier verschiedene Anbieter die persönliche HRTFs als Dienstleistung erzeugen. Auch Genelec bietet seit einiger Zeit einen solchen Service an undnennt es "Genelec Aural ID". Das Endprodukt sind SOFA-Dateien, die man in Software mit SOFA-Import-Möglichkeit nutzen kann. Dazu später im Verlauf den Beitrags mehr.

Head Related Transfer Function

Zunächst noch mal ein paar technische Grundlagen. Die individuelle Außerohr-Übertragungsfunktion einer Person wird ja in der Regel in Form einer sogenannten SOFA-Datei gespeichert wird. Dieses Format ist über den Standard AES69 seit 2015 auch definiert, der im Jahr 2020 nochmal überarbeitet und auch als SOFA 2.0 bezeichnet wird. Wer mehr zum Thema erfahren möchte, der sollte sich ggf. auch mal den Beitrag VST AmbiDecoder anschauen, in dem wir mehr auf die Grundlagen der HRTF eingegangen sind.

Auf einige spezifische Dinge bezüglich der HRTF möchte ich aber auch in diesem Artilel eingehen. Es gibt verschiedenste Verfahren, wie man zu einer HRTF kommt. Von der Umsetzung her muss man zwischen zwei prinzipielle Methoden unterscheiden und zwar:

- Messungen über Miniatur-Mikrofonen im Gehörgang und Testsignale aus einer 3D-Anordnung von Lautsprechern,
- visuelle Erfassung der Kopf/Körper- und Ohr-Geometrie über Fotos und

Video.

Bei einer Messung werden Testsignale aus den einzelnen Lautsprechern ausgegeben und über die Mikrofone aufgenommen. Aus diesern einzelnen Signalen kann man dann direkt die erforderlichen Daten für die HRTF, bzw. das SOFA-File gewinnen. Beim zweiten genannten Verfahren muss die HRTF aus der ermittelten Geometrie aber erst berechnen oder über eine Datenbank eine HRTF ausgewählt werden, die der Körper/Kopf/Ohr-Geometrie möglichst sehr ähnelt.

Es ist so, dass beide Verfahrensweisen möglichst viele und präzise Daten benötigen, um auch eine HRTF-Beschreibung für die SOFA-Datei zu generieren, die der realen HRTF möglichst nahe kommt. Bei den Messungen ist es zum Beispiel hilfreich, möglichst viele Lautsprecher einzusetzen und diese möglichst präzise zu plazieren. Auch die Raumakustik kann Einflüsse auf die Messung und die HRTF haben.

Auch bei den Verfahren die über einen Visualisierung mittels Foto und/oder Video die Geometrie erfassen, ist es hilfreich möglichst viele Daten zu haben, um ein möglichst genaues Abbild der Geometrie des Körpers, bzw. Kopfes und Ohr zu bekommen. Viele Dinge haben Einfluss auf die HRTF und auch nicht nur das Außenohr selbst, sondern auch der Kopf, bzw. die Kopfform bis hin zum Körper selber. Der Vorteil bei der visuellen Methode ist, dass akustische Beeinflussungen nicht auftreten.

Viele Anbieter vereinfachen den Erfassungprozess, um es dem Anwender möglichst einfach zu machen. Das resultiert dann aber auch in weniger Datenmaterial für die Berechnung der Geometrie. Das wiederum führt zu eine generierten HRTF, die von der realen HRTF auch mehr abweicht, also ungenauer ist.

Konzept



Kommen wir nun zum Konzept der Genelec Aural ID. Die Grundlage hierfür sind Fotos und ein Video. Aus den Daten wird ein dreidimensionales Modell des Oberkörpers, des Kopfes und im Detail der Ohren generiert. Anhand dieses Modells lassen sich die einzelnen HRTF-Filter berechnen. Es werden dabei Daten für beide Ohren aus weit über 800 verschiedene Richtungen erzeugt, was eine Menge ist. Das System basiert also nicht auf HRTF-Datenbankvergleiche.

Der Service kann ausschließlich Online direkt bei Genelec bestellt werden. Man sendet die erstellten Fotos und das Video, welches man mit einer hochauflösenden Kamera eines Mobiltelefons aufnehmen kann, zu Genelec und erhält mehrere Dateien zurück und zwar jeweils eine SOFA-Datei für 44,1, 48 und 96 kHz Abtastrate.

Eine HRTF nach dem Common-Acoustical-Pole and Zero Modell (CAPZ) besteht eigentlich aus zwei Transferfunktionen und zwar aus der Common Transfer Function (CTF) und der Directional Transfer Function (DTF). Die DTF ist dabei für die Lokalisation verantwortlich. Genelec liefert neben den SOFA-Dateien in der beide Transferfunktionen enthalten sind auch noch jeweils eine, in der ausschließliche die DTF beschrieben ist. Je nach der Applikation kann es klanglich nützlich sein, die SOFA-Datei ausschließlich mit der DTF zu nutzen. Wer es ganz genau wissen möchte, dem empfehle ich zum Beispiel den Beitrag "Common-Acoustical-Pole and Zero Modeling of Head-Related Transfer Functions", der von der IEEE veröffentlicht wurde.

Phoca PDF

Fotos und Video Erstellung

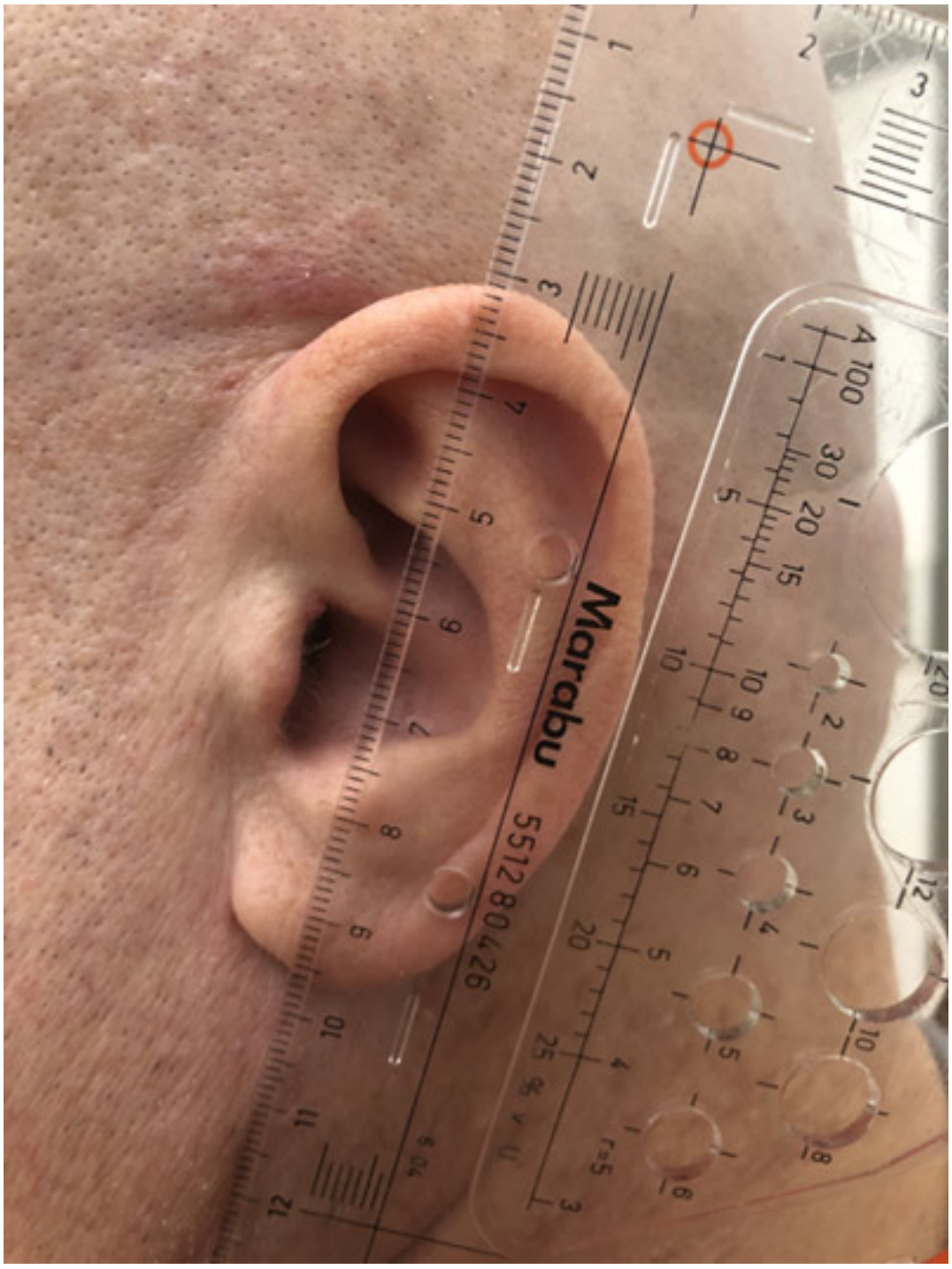
Kommen wir nun zu dem Teil des Prozesses wo der Anwender gefragt ist, nämlich der Erstellung der benötigten Fotos und des Videos. Genelec liefert mit dem Aural ID Operating Manual ein Dokument in Form eines PDFs, welches sehr genaue beschreibt, was genau zu tun ist. Darüber hinaus gibt es noch ein Erklärungsvideo.

Für das Video benötigt man die Hilfe einer weiteren Person. Die Länge des Videos soll drei bis vier Minuten dauern. Man sitzt bei der Erstellung aufrecht auf einem Stuhl und es wird zunächst der gesamten Körper von vorne aufgenommen und der Assistent, der das Video aufnimmt, wandert langsam in Richtung linke Kopfseite, nähert sich dem Ohr und nimmt es von verschiedenen Richtungen auf, um die exakte dreidimensionale Geometrie einzufangen. In einem Anleitungsvideo zur Aural ID wird hierbei auch auf Stellen aufmerksam gemacht, die dort genauer zu erfassen sind.



Danach geht man wieder vom Ohr weiter weg und setzt seinen Rundgang fort, bis zum rechten Ohr, geht dort wieder näher an das Ohr und nimmt auch dieses im Detail auf. Anschließend wird der Abstand wieder vergrößert und man führt den Gang um die Person weiter bis zum Ausgangspunkt fort. Soweit das Ganze einmal grob beschrieben.

Phoca PDF



Es ist jetzt noch jeweils ein Foto vom rechten und linken Ohr zu erstellen und zwar mit einem (s. Abb. oben) Lineal am Ohr, um die Größenverhältnisse zu protokolieren. Das Video und die beiden Fotos werden über den Genelec Community Account, den man ggf. noch anlegen muss, übertragen. Hierzu muss man nach Log-in "Create New Profile" anwählen und die Daten dann übertragen. Die SOFA-Dateien stehen hier auch nach der Anfertigung zum Download bereit. Nach der Bezahlung dauert es keine Woche bis die SOFA-Dateien abrufbar sind.

Praxis

Phoca PDF

Zunächst zum Erstellen des Videos, denn das ist für den Kunden der Aural ID der komplizierteste Prozess. Man sollte hier eine Person zur Unterstützung heranziehen, die etwas Übung im Aufnahmen von Videos hat und auch weiss was man da so eigentlich vor hat und wozu es dient. Sehr hilfreich ist hier das Beispielvideo, was man sich unbedingt vorab anschauen sollte. Hier wird sehr anschaulich auf wichtige Punkte hingewiesen.

Ganz so einfach wie es aussieht ist es dann nämlich nicht. Wir haben schon drei Anläufe gebraucht, um ein für Genelec verwertbares Video zu produzieren. Zwei bis drei Stunden Zeit sollte man sich für die Vorbereitung, Aufbau, Durchführung und Kontrolle des Videos und der beiden Fotos nehmen. Das ist deutlich mehr Zeit, als man zum Beispiel bei anderen Verfahren benötigt.

Durch die verschiedenen Video-Perspektiven an den beiden Ohren bekommt Genelec aber auch genügend Datenmaterial für ein sehr genaues 3D-Model des Ohr und durch den Rundgang um die Person auch über Kopf und Oberkörper. Letztere Informationen werden bei anderen Verfahren häufig gar nicht erfasst und man beschränkt sich auf den Außerohrbereich.

Wir haben mit den Tests zur Aural ID schon Ende 2019 begonnen. Mittlerweile wurde die Generierung der SOFA-Dateien auch optimiert und zwar unter anderem in dem Punkt, dass in der Metadaten-Beschreibung ein paar Abweichungen vom Standard waren, die zu Konflikten mit einigen Applikationen führen konnten. Bei den aktuellen Versionen der Aural ID konnten wir dies nicht mehr beobachten.

Ich hatte schon vor über 20 Jahren mich erstmalig mit HRTFs beschäftigt und an einem Versuch teilgenommen, in dem eine ganze Menge von HRTFs auf ein Musikprogrammsignal angewendet wurde, mit dem Ergebnis, dass ich nur eine HRTF überhaupt brauchbar fand und das war - welcher Zufall - meine eigene, die damals auch gemessen wurde. Seid dem weiss ich, dass ich sehr schnell Abweichungen gegenüber meiner realen HRTF wahrnehme.

Mein Fazit in der Benutzung der Genelec Aural ID ist, dass in den verwendeten binauralen Plug-Ins sowohl die Ortung gut war als auch der gesamte Klangeindruck. Das ist bei HRTFs häufig ein Problem die richtige Balance zu finden. Es gibt HRTFs, da funktioniert die Ortung aber es hört sich gruselig an oder HRTFs wo der Gesamtklangeindruck zwar sehr gut aber keine präzise Ortung möglich ist. Viele HRTFs haben besonders das Problem eine Ortung vorne zu realisieren und dabei auch einen guten Klang zu bieten. Bei der Genelec Aural ID ist das wirklich gut gelungen und ich bin mit der Ortung und dem Gesamtklangeindruck sehr zufrieden. Man muss natürlich sagen, dass vieles bei Thema HRTF subjektiv ist und auch die binauralisierenden Plug-Ins selbst auch noch eine klangliche Rolle spielen.

Fazit

Die Kosten für eine Genelec Aural ID liegen bei knapp über 600 Euro. Das ist nicht

gerade wenig für eine personliche HRTF, bzw. SOFA-Files, aber das Verfahren ist ein sehr guter Kompromiss zwischen Aufwand und Ergebniss. Der Aufwand ist höher als bei vielen anderen Verfahren, aber das Ergebniss ist auch wirklich nachvollziehbar besser, als bei Verfahren wo nur das Ohr fotografiert wird.

Wer also häufig binaural produziert oder binaurales Kopfhörermonitoring für immersive Produktionen nutzt und keine Möglichkeit hat eine HRTF messtechnisch generieren zu lassen (was nicht automatisch heisst, dass das Resultat dann besser ist, denn das hängt vom Messaufbau ab), dem kann man die Genelec Aural ID wirklich sehr empfehlen.

https://www.genelec.com/aural-id

Phoca PDF