Binauralisierung

3D-Audio über Kopfhörer

Autor und Fotos: Peter Kaminski



Immersives Audio wird auf der Consumer-Seite überwiegend über Kopfhörer gehört. Aber auch im Bereich der Produktion von 3D-Audio wird es immer wichtiger, zum Beispiel bei mobilen Produktionen oder kleines Produktions-Suites, wo kein großes 3D-Audio-Lautsprecher-Setup zur Verfügung steht oder bei der Dolby-Atmos-Produktion um Binauralisierungs-Parameter zu überprüfen. Wir möchten daher einmal ein Grundlagenartikel zu diesem Thema anbieten, der auf alle Aspekte der Binauralisierung und das Hören über Kopfhörer eingeht.

Richtungshören

Um Binauralisierung besser zu verstehen, wollen wir uns erst einmal mit der Wahrnehmung und dem menschlichen Gehör auseinandersetzen. Zur Ortung einer Schallquelle werden Pegel- und Laufzeitunterschiede vom linken und rechten Gehör ausgewertet. Damit ließe sich aber nur feststellen, ob eine Schallquelle mehr links oder rechts oder sich eben in der Mitte von unserer Position befindet.

Phoca PDF

Um festzustellen, ob sich eine Schallquelle vor oder hinter uns, bzw. oben oder unten befindet analysiert das Gehirn das Spektrum das wir wahrnehmen. Der eintreffende Schall wird durch die Ohren reflektiert und durch die Geometrie der Ohren wird das Spektrum, je nach dem aus welcher Richtung der Schall kommt, verändert. Aber auch die Kopfform und der Oberkörper haben Einfluss auf den Schall und verändern das Spektrum durch Reflektion und Absorbierung. Da wir unterbewußt gelernt haben, wie sich der Schall, bzw. dessen Spektrum sich bei bestimmten Einfallswinkeln verändert, kann das Gehirn die Richtung der Schallquelle prognostizieren. Man spricht in diesem Zusammenhang auch von richtungsabhängigem Spektrum.

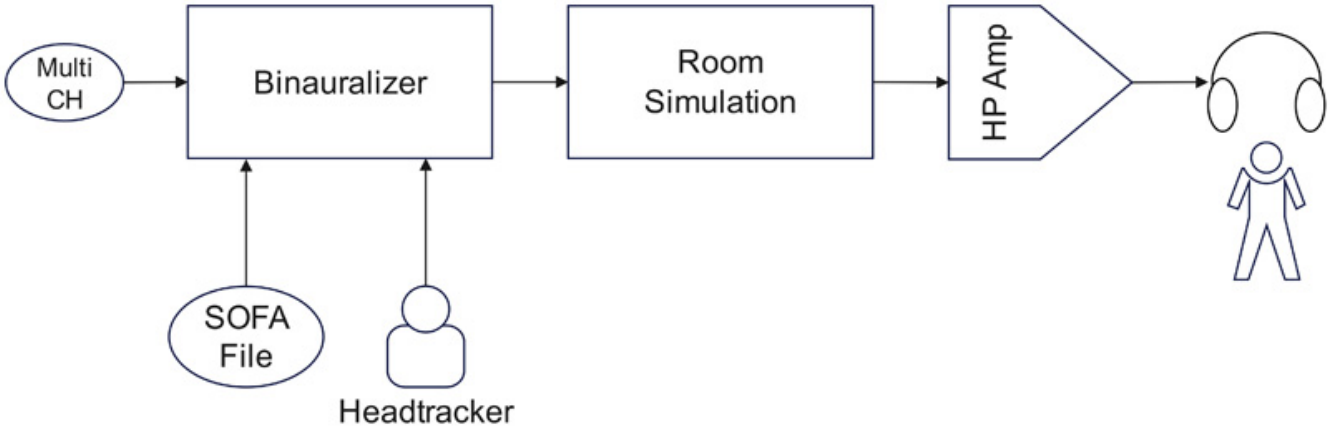
Mathematisch spricht man von der individuellen Außerohr-Übertragungsfunktion (Englisch: Head Related Transfer Function, Abk.: HRTF), die beschreibt, wie Ohrmuschel, Kopf und Oberkörper den Schall reflektieren, beugen und absorbieren und darüber das richtungsabhängig das wahrgenommene Spektrum verändern.

Bei reinen Sinustönen haben wir wegen dem fehlenden Obertonspektrum nur die Pegel- und Laufzeitdifferenz als Bewertungskriterien für die Richtung - und die Schallquelle kann deshalb nur sehr ungenau lokalisiert werden. Bei Schallquellen mit sehr obertonreichem Spektrum ist dagegen eine deutlich präzisere Ortung möglich, da sich hier das Spektrum besser nutzen lässt. Das bedeutet im Umkehrschluss, dass die Ortungspräzision auch von der spektralen Zusammensetzung der Schallquelle abhängt.

Auch die menschliche Fähigkeit das wahrgenommene Spektrum auszuwerten sind durchaus unterschiedlich. Ursachen hierfür liegen in der Verarbeitung im Gehirn, denn eigentlich hören wir ja mit dem Gehirn und das Gehör ist lediglich die Sensorik. Auch die optische Auswertung der Umgebung fließt in den Wahrnehmungsprozess ein und unterstützt ggf. die Schallortung. Fehlt dieser Eindruck bei einer Virtualisierung, dann kann das die Ortungspräzision durchaus negativ beeinflusst werden. Gerade im Bereich der Musikproduktion fehlt ja in der Regel ein visueller Eindruck.

Signalkette

Bei Binauralisierung-Software, bzw. Plug-Ins werden die zuvor beschriebenen Effekte genutzt, um diskrete 3D-Audiosignale, wie ein 7.1.4- oder 9.1.6-Signal oder Dolby-Atmos-Datenströme auf zwei Kanäle so zu kodieren, dass man über Kopfhörer im optimalen Fall das gleiche immersive Erlebnis wahrnimmt, wie bei einer Wiedergabe über Lautsprecher. Je nach Qualität oder Konzept ist man mehr oder weniger weit von diesem optimalen Ergebnis entfernt.



Wenn wir auf die mögliche Signalkette des gesamten Binauralisierungs-Prozesses schauen, dann gibt es da einige Punkte, die wir noch nicht angesprochen haben. Ein Punkt ist, dass auch der Raum in dem das Schallereignis stattgefunden hat Teil der Virtualisierung ist.

Generische vs. Individual HRTF

Es gibt zwei Möglichkeiten eine HRTF in den Binauralisierungs-Prozess zu nutzen: Und zwar einmal die HRTF des Nutzers oder eine generische HRTF. Bei generischen HRTFs mittelt man HRTF-Daten oder man synthetisiert sie über Modelle. Generische HRTF haben den Anspruch einer möglichst großen Anwendergruppe eine realistische Virtualisierung mit einer nicht personalisierten HRTF zu gewährleisten. Dies geht natürlich zu Lasten der Qualität beim Klang und Richtungsempfinden. Anwender kommen somit unterschiedlich gut mit einer generischen HRTF zurecht. Das Optimum ist immer eine HRTF, die auf der eigenen Geometrie der Ohrmuschel und des Kopfes und Oberkörper zu Grunde liegt.

SOFA-Datei

Die Daten, die die individuelle HRTF beschreiben, müssen als Datei oder Profil hinterlegt werden, um sie in der Binauralisierungs-Software, bzw. Plug-In verwenden zu können. Das verwendete Dateiformat kann dabei entweder proprietär, also System-bezogen, oder in einem Standardformat abgelegt werden.

System-, bzw. Produkt-übergreifend wird die individuelle Außerohr-Übertragungsfunktion einer Person in der Regel in Form einer sogenannten SOFA-Datei gespeichert. Dieses Format ist über den Standard AES69 seit 2015 auch definiert. Der Standard wurde im Jahr 2020 nochmal überarbeitet und wird gelegentlich auch als SOFA 2.0 bezeichnet. Für die einzelnen Abtastraten, wie 44,1, 48 und 96 kHz, müssen jeweils einzelne SOFA-Dateien verfügbar sein.

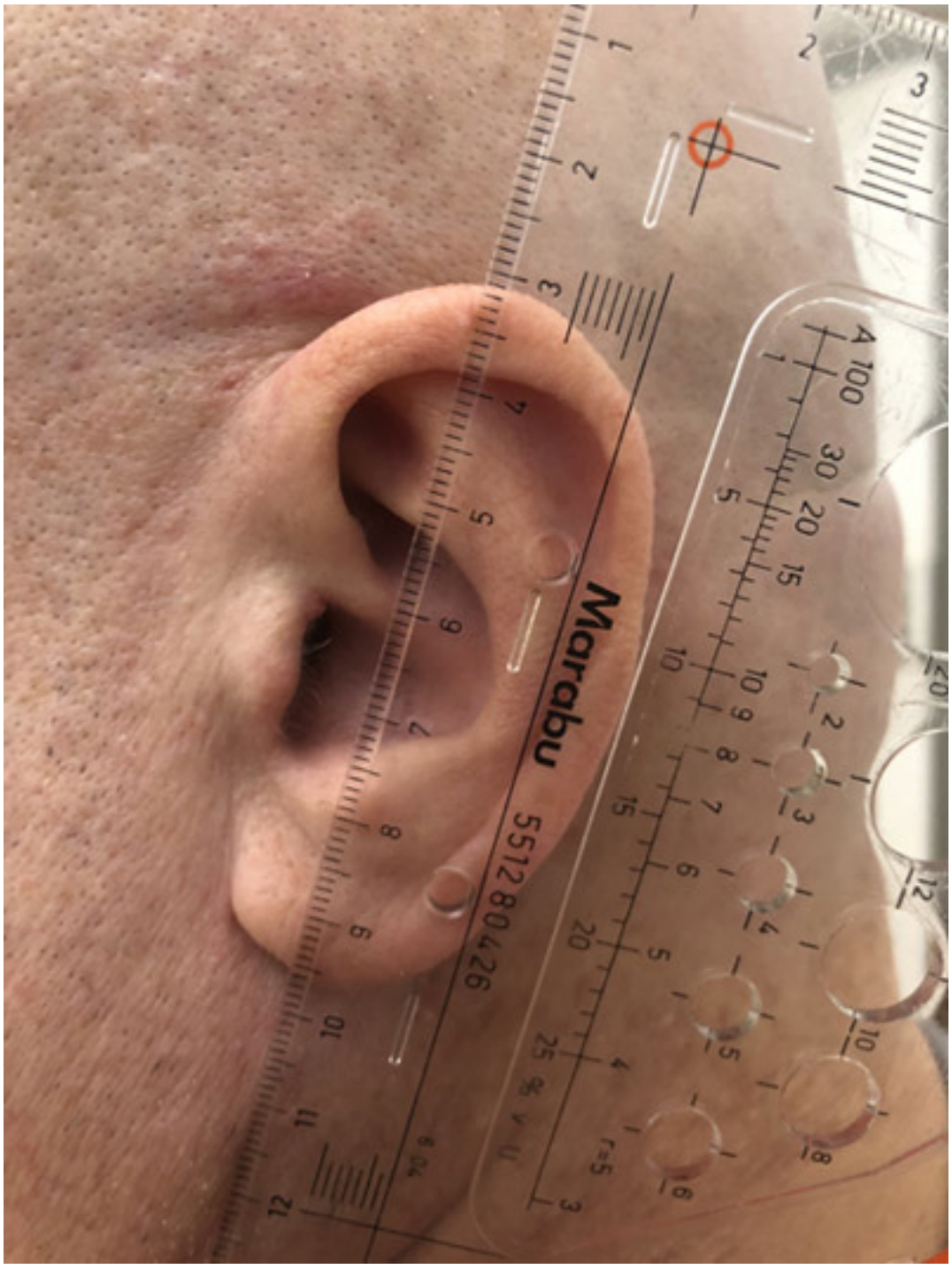
Eine HRTF nach dem Common-Acoustical-Pole and Zero Modell (CAPZ) besteht eigentlich aus zwei Transferfunktionen und zwar aus der Common Transfer Function

(CTF) und der Directional Transfer Function (DTF). Die DTF ist dabei für die Lokalisation verantwortlich. Wer es ganz genau wissen möchte, dem empfehle ich den Beitrag "Common-Acoustical-Pole and Zero Modeling of Head-Related Transfer Functions", der von der IEEE veröffentlicht wurde und im Internet als PDF heruntergeladen werden kann.

Erfassung einer HRTF

Es gibt prinzipiell zwei Methoden eine individuelle HRTF zu erfassen und zwar einmal über eine Messung. Dazu wird im Ohr im Gehörgang möglichst nahe am Trommelfell jeweils ein Miniaturmikrofon platziert. Um die Person herum, wird nun aus einzelnen Lautsprechern mit unterschiedlichen Positionen entweder ein Sweep oder ein Rauschsignal eingespielt und das Signal des Mikrofons im Ohr aufgenommen. Eine große Anzahl an Lautsprechern ist hier sinnvoll. Aus den verschiedenen einzelnen Aufnahmen und dem aufgezeichneten Spektrum wird eine HRTF berechnet und in eine HRTF-Datei - im Idealfall im SOFA-Format - gespeichert.

Wichtig zu wissen ist, dass bei einer HRTF-Erfassung über eine Messung der Raum, in dem gemessen wird, mit in die HRTF einfließt. Je nach Anwendungszweck nutzt man entweder gut klingende Räume oder man versucht den Raumeinfluss durch eine nahe Lautsprecheraufstellung - also im Nahfeld - zu minimieren.



Ein anderer Weg eine HRTF zu generieren ist über Fotos und Videos der Ohrmuscheln und des Kopfes, bzw. Oberkörpers. Dazu werden häufig Referenzen mit aufgenommen, wie ein Lineal mit Millimeterskala oder man muss in der Software den Ohrabstand oder den Kopfumfang mit manuell eingeben. Aus diesem Material wird in der Regel eine virtuelle Körper-, bzw. Ohrmuschelgeometrie generiert und daraus dann eine HRTF errechnet.

Die Prozesse eine HRTF über die Körpergeometrie zu berechnen werden immer besser, aber qualitativ kommen sie zurzeit nicht an die Qualität einer Messung

Phoca PDF

heran, die die Realität besser abbildet. Der Vorteil ist aber, dass man mit relativ einfachen Mittel ohne eine Messung eine HRTF generieren kann.

Headtracking

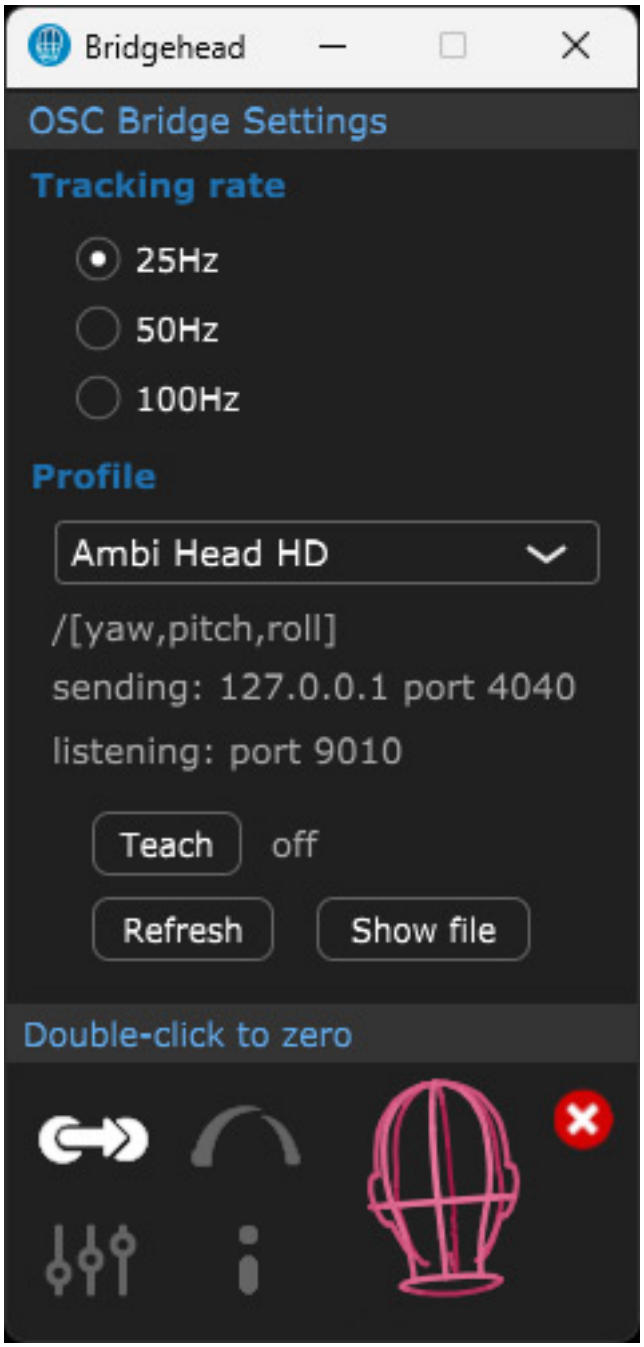


Wir justieren und präzisieren - zum Teil bewusst und zum Teil unterbewusst - unsere Ortungswahrnehmung durch mehr oder weniger große Kopfbewegung. Bewusst zum Beispiel in dem man sich mit dem Kopf in die vermutete Richtung einer Schallquelle dreht um diese eben zu bestätigen oder ggf. zu verwerfen. Aber auch unterbewusst durch sehr kleine Kopfbewegung, quasi wie eine Art der Kalibrierung.

Es ist auch so, dass wenn man sein Kopf bewegt, sich der Eindruck des Schallfeldes bei der Wahrnehmung entsprechend anpasst. Ohne eine Kompensation dreht sich beim Hören mit Kopfhörerhören einer Kopfdrehung das gesamte Schallfeld mit und entspricht somit nicht der Wahrnehmung in der Realität. Dies führt dazu, dass das Gehirn immer wieder daran erinnert wird, dass die Virtualisierung nicht mit dem Realitätsempfinden übereinstimmt.

Um dies im Binauralisierungs-Prozess zu berücksichtigen gibt es sogenannte Headtracker. Über Sensoren wird die Kopfposition, bzw. die Lage des Kopfes im Raum analysiert und über eine Software-Schnittstelle dem Binauralisierungs-Plug-In mitgeteilt, dass dann das Schallfeld entsprechend der Kopfposition anpasst.

Phoca PDF



Wir haben schon einen umfangreichen Test über den Supperware Headtracker 1 geschrieben, der zurzeit der am weitesten verbreitete Headtracker im Produktionsbereich ist. Er bestehend aus der Sensor-Hardware in Form eines Bandes mit verschiedenen Sensoren, was über USB-Kabel an den Rechner angeschlossen wird und die Daten dort über die Software Bridgehead via OSC-Protokoll dem Binauralisierer bereitgestellt wird.

Wahrnehmung und Erwartungshaltung

Ein wichtiger Punkt ist, dass nicht nur technische Parameter das Ergebnis einer Binauralisierung beeinflussen, sondern auch die Wahrnehmung des Menschen einen

Phoca PDF

meistens unterschätzten Einfluss hat. Die Visualisierung einer Aufführung haben wir ja schon kurz angesprochen, aber es geht auch um den Raum in dem man sitzt. Korreliert der Raum und dessen Raumakustik in der Binauralisierung nicht mit dem Raum in dem man das Abhören gewohnt ist, so kann das zu negativem Einfluss in der Wahrnehmung führen, weil eine Erwartungshaltung unter Umständen nicht erfüllt wird. Hier ergibt sich unter Umständen ein Vorteil bei Binauralisierung über HRTFs, die nicht auf akustische Messungen beruhen, da man hier ggf. mit einem synthetischen Raum und Veränderungen von Parametern sich entsprechend anpassen kann. Bei HRTFs die auf Messungen basieren ist der Raum ja schon ein Teil der HRTF.

Kopfhörer

Ein weiterer Punkt in der Signalkette ist der Kopfhörerverstärker - vor allem jedoch der Kopfhörer selbst. Auch Kopfhörerverstärker haben ihren eigenen Klang und beeinflussen das Spektrum. Die Einflüsse sind aber nicht so groß wie die des Kopfhörers, die zum Teil doch sehr unterschiedlich klingen. Daher vernachlässigt man in der Regel den Einfluss des Kopfhörerverstärkers bei der Virtualisierung.

Bei den Kopfhörern hingegen kann man häufig über eine Bibliothek den genutzten Kopfhörer auswählen. Über das gespeicherte Profil wird versucht seine klanglichen Eigenschaften zu kompensieren, damit sie den Binauralisierungs-Prozess nicht negativ beeinflussen. Das hinterlegte Profil des Kopfhörers basiert aber in der Regel auf den Verlauf des Übertragungsverlauf. Komplexes dynamisches Verhalten der Kopfhörer wird nicht berücksichtigt.

Man müsste sich auch die Frage stellen, welcher Kopfhörer für die Wiedergabe von binauralen Signalen geeigneter ist als andere. Die Frage ist leider nicht ganz so einfach zu beantworten. Jedes Kopfhörermodell hat eine akustische Korrekturkurve implementiert, damit er so klingt wie man es möchte. Bekannt ist zum Beispiel die sogenannte Harman-Kurve. Aber es gibt auch Kurven, die stark davon abweichen. Kimio Hamasaki hat zum Beispiel auch empirisch eine Korrekturkurve entwickelt, die speziell der Optimierung der Richtungswahrnehmung zum Ziel hat (s. VDT-Magazin 2024 Ausgabe 1, S. 32 ff.). Diese wird aber kaum eingesetzt. Spezielle Studio-Kopfhörer für die optimale binaurale Wiedergabe gibt es zurzeit nicht. Man kann prinzipiell akustische offene Kopfhörer empfehlen. Es gibt auch Kopfhörer, deren Treiber leicht verdreht eingebaut sind, um so die Außerkopflokalisation zu fördern - aber diese Kopfhörer sind eher ungeeignet für eine binaurale Wiedergabe.

Fazit

Bei den Consumer ist unter anderem dank Apple Music das Hören von 3D-Audio über personalisierte Binauralisierung ein Regelfall. Im Studio ist es eher die Regel, dass entsprechende Lautsprecher-Setup genutzt werden. Mit Hilfe von Binauralisierern ist aber mittlerweile auch durchaus ein Produzieren von immersiven Inhalten über Kopfhörer möglich. Die Qualitäten sind zum Teil sehr unterschiedlich zu bewerten. Dabei spielt auch die individuelle Wahrnehmung eines Menschen eine

Phoca PDF

große Rolle, insbesondere was generische HRTFs angeht. Es gibt Zuhörer die mit solchen generischen HRTFs mehr oder weniger gut zurechtkommen. Was die Realitätsnähe angeht, so sind Verfahren, die auf individuellen, gemessenen HRTFs basieren im Vorteil. Berechnete HRTFs haben wiederum andere Vorteile, wie eben eine höhere Flexibilität über eine Raumsimulation.

Unsere folgenden Tests von Produkten geben einen weiteren Einblick in die Welt der binauralen Kopfhörerwiedergabe:

https://www.proaudio.de/tests/apl-virtuoso-binauralisierer
https://www.proaudio.de/tests/neumann-rime
https://www.proaudio.de/tests/supperware-headtracker
https://www.proaudio.de/tests/noise-makers-binauralizer-v2
https://www.proaudio.de/tests/genelec-aural-id-individuelle-hrtf

Phoca PDF