

Modell der Signallaufzeiten für die binaurale Wiedergabe virtueller Schallquellen

Harald Ziegelwanger^{1,2}, Piotr Majdak¹, Bernhard Laback¹

¹ *Institut für Schallforschung, Österreichische Akademie der Wissenschaften, 1040 Wien, Austria, Email: h.ziegelwanger@me.com*

² *Institut für elektronische Musik und Akustik, Kunstuniversität Graz, 8010 Graz, Austria*

Einleitung

Beidohrige Außenohrübertragungsfunktionen (engl. head-related transfer functions, HRTFs) können als minimalphasige Systeme mit zugehöriger breitbandiger interauraler Zeitdifferenz (engl. interaural time difference, ITD) modelliert werden [1]. Diese Approximation stellt eine gängige Methode zur Reduzierung der Datenmenge [2] oder zur Veränderung der ITD in psychoakustischen Experimenten [3, 4] dar. Dabei wird die ITD für eine räumliche Richtung direkt aus den HRTFs der beiden Ohren geschätzt. Die Schätzung erfolgt auf Basis von korrelativer Techniken im Zeitbereich [3, 5], Regression der abgewinkelten Phase [6] oder direkter Auswertung der Gruppenlaufzeit [7]. Diese Verfahren bearbeiten jeweils nur eine Richtung - eine perzeptiv relevante Fehlschätzung ist häufig und muss manuell korrigiert werden [8]. Durch die direkte Schätzung der interauralen Information bleibt die monaurale Signallaufzeit (engl. time of arrival, TOA) unbekannt, was im Falle multipler Schallquellen zum monauralen Zeitversatz zwischen den Richtungen führen kann. In dieser Arbeit wird ein Verfahren zur direkten räumlichen Modellierung der TOAs vorgestellt. Das Verfahren arbeitet automatisiert und bezieht HRTFs für mehrere Richtungen ein, um perzeptiv korrekte TOAs ohne manuelle Nachbearbeitung modellieren zu können.

TOA für eine Richtung

Die TOA $\hat{\tau}$ für eine Richtung (φ, θ) ist gegeben durch:

$$\hat{\tau}(\varphi, \theta) = \arg \max_k \left\{ \sum_{n=0}^{N-1} h[n] h_{min}[n-k] \right\}$$

mit φ als Azimut- und θ als Elevationwinkel der Schallquelle, wobei $h[n]$ der HRTF im Zeitbereich und $h_{min}[n]$ dem Minimalphasenanteil der $h[n]$ mit der Länge N für die Richtung (φ, θ) entsprechen [5]. Abb. 1 (schwarze Linie) zeigt die TOAs für die Horizontalebenen $H(\varphi, 0^\circ)$ und $H(\varphi, 55^\circ)$, basierend auf HRTFs einer Person. Die sprunghaftigen Abweichungen finden keine physikalische Entsprechung und sind als Fehler einzustufen. Eines der Ziele des vorgestellten Modells ist die automatische Erkennung und Korrektur dieser Fehler.

Räumliches Modell der TOA

Das Modell bildet die Schalllaufzeit zwischen einer Schallquelle und einem auf der Kugel platziertem Ohr nach, wobei die Position der Kugel im Raum und des Ohrs auf der Kugel als freie Parameter berücksichtigt werden (Abb. 2, Kugelmittelpunkt $\mathbf{M} = (x_M, y_M, z_M)$,

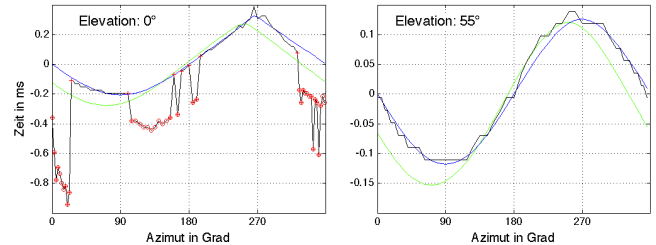


Abbildung 1: TOA für zwei Horizontalebenen. **Schwarz:** $\hat{\tau}$, **rot:** erkannte Fehler, **grün/blau:** $\hat{\tau}$ modelliert ohne/mit Fehlererkennung.

Ohrposition (φ_o, θ_o) , Kugelradius r , Schallgeschwindigkeit c , richtungsunabhängiger Anteil der TOA $\bar{\tau}$). Für eine Schalleinfallrichtung (φ, θ) werden die Schallwege s_1, s_2 und s_3 auf zwei unterschiedlichen Großkreisen über geometrische Beziehungen berechnet:

$$s_1 = \begin{cases} r(1 - \cos \alpha) & \text{für } \alpha \leq \frac{\pi}{2} \\ r(1 + \alpha - \frac{\pi}{2}) & \text{für } \alpha > \frac{\pi}{2} \end{cases}$$

$$s_2 = -r + |\mathbf{M}| \cos \beta + \sqrt{r^2 + |\mathbf{M}|^2 \cos^2 \beta + 2|\mathbf{M}|r}$$

$$s_3 = \frac{|\mathbf{M}| \sin \beta}{2(|\mathbf{M}| + r) \tan(\frac{\gamma}{2})},$$

mit:

$$\alpha = \cos^{-1} \{ \sin \theta_o \sin \theta + \cos \theta_o \cos \theta \cos(\varphi_o - \varphi) \},$$

$$\beta = \cos^{-1} \{ -(x_M \cos \varphi + y_M \sin \varphi) \cos \theta - z_M \sin \theta \},$$

$$\gamma = \pi - \beta - \cos^{-1} \left(1 - \frac{2rs_2 + s_2^2}{2|\mathbf{M}|^2 + 2|\mathbf{M}|r} \right).$$

Die modellierte TOA $\tilde{\tau}(\varphi, \theta)$ ist nun proportional zur Summe der drei Schalllaufwege:

$$\tilde{\tau}(\varphi, \theta) = \frac{1}{c} (s_1 + s_2 + s_3) + \bar{\tau}.$$

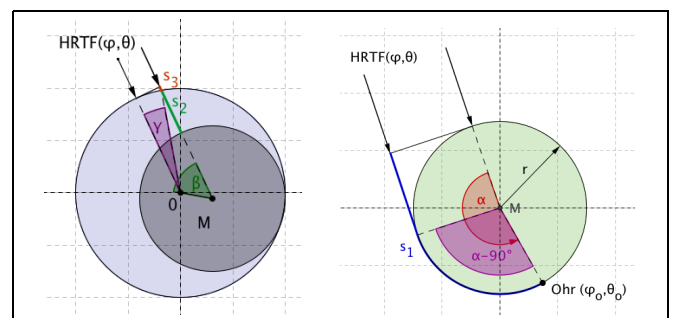


Abbildung 2: Räumliches TOA-Modell. Der grüne Kreis (rechts) ist ein Großkreis der dunkelgrauen Kugel (links).

Parameteranpassung

Die Modellparameter werden aus HRTFs einer Person geschätzt, wobei nur fehlerbereinigte TOAs für die Anpassung verwendet werden. Die Fehlererkennung analysiert die TOAs $\hat{\tau}$ und erstellt die Menge der fehlerbereinigten TOAs $\hat{\tau}_K$ in 3 Schritten: 1) Da TOAs in Horizontalebene glatte Funktionen sein sollten, werden alle Richtungen die zu starken Fluktuationen führen entfernt. 2) Da TOAs in Sagittalebene beinahe konstant sein sollten, werden Richtungen zu Sagittalebene gruppiert, und alle Richtungen die zu großer Varianz innerhalb jeweils einer Sagittalebene führen entfernt. 3) Nach einer vorläufigen Modellanpassung werden alle Richtungen außerhalb eines 2-fachen quadratischen Fehlers entfernt, sowie alle Richtungen innerhalb des halben quadratischen Fehlers, die aber in früheren Schritten als Fehler markiert wurden, wieder in die Modellierung inkludiert.

Basierend auf den fehlerbereinigten TOAs $\hat{\tau}_K$ werden die Parameter des Modells $\tilde{\tau}(\varphi, \theta)$ mit Hilfe einer nicht-linearen Anpassung geschätzt. Die Abb. 1 zeigt beispielhaft die Ergebnisse der Modellierung mit (blau) und ohne (grün) der Fehlererkennung. Der Verlauf der TOA ist nun glatt und Fehlschätzungen einzelner Positionen wurden drastisch reduziert. Der Effekt der Fehlererkennung ist anhand der grünen Linie ersichtlich, die von der blauen Linie deutlich abweicht.

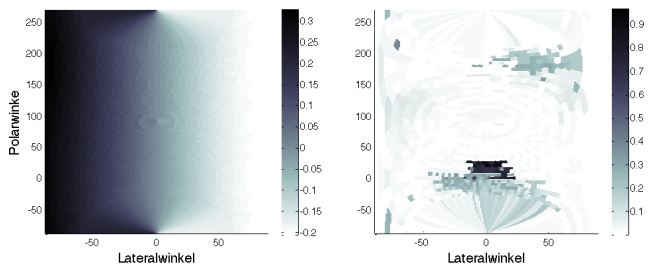


Abbildung 3: Links: Modellierte TOA $\tilde{\tau}$. Rechts: Abweichung $\tilde{\tau} - \hat{\tau}$.

Resultate

Das vorgestellte Modell wurde an HRTFs aus der ARI-Datenbank (57 Personen) und CIPIC-Datenbank (25 Personen) angewandt. Abb. 3 zeigt die modellierten TOAs $\tilde{\tau}$ (links), sowie deren Abweichung zu den geschätzten TOAs $\hat{\tau}$ (rechts) für eine typische Person und alle verfügbaren Richtungen. Die Anwendung des Modells an Daten aller Personen liefert ebenfalls anthropometrisch plausible Resultate. Zum Beispiel zeigen die Kopfradien bei getrennter Modellierung für das linke und rechte Ohr nur eine geringe Abweichung für eine Person (Abb. 4). Besonders erwähnenswert ist - bei der TOA, im Gegensatz zur ITD - die Relevanz der Modellierung der Kopfposition \mathbf{M} bei der HRTF-Messung: wird die seitlich Auslenkung y_M bei der Modellierung nicht berücksichtigt, steigt die Abweichung zwischen den modellierten Kopfradien der beiden Ohr.

Die Anwendungen unseres Modells sind vielseitig. Zum Beispiel kann die Kopfposition bei der HRTF-Messung

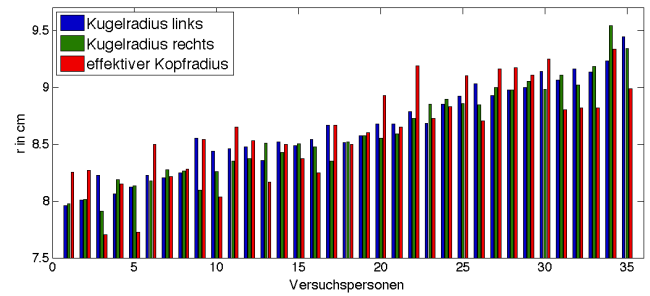


Abbildung 4: Vergleich Kugelradius (links und rechts) und anthropometrische Daten, aufsteigend geordnet für den Mittelwert der Kugelradien.

nachträglich bestimmt und deren Richtungen korrigiert werden. Die genaue Schätzung der anthropometrischen Daten kann aber auch dazu genutzt werden, um bessere Modelle der HRTFs zu erzielen. Somit kann als Beispiel der Zusammenhang zwischen dem resultierenden Kugelradius und den anthropometrischen Daten Kopfbreite b , Kopfhöhe h , sowie Kopftiefe t untersucht werden [9]. Mit Hilfe der HRTFs von 10 Personen der ARI- und 25 Personen der CIPIC-Datenbank konnte über die lineare Regression zwischen den anthropometrischen Daten b , h , t und dem modellierten Kugelradius r der Kopfradius \tilde{r} modelliert werden (Pearson's $R = 0.83, p < 10^{-9}$):

$$\tilde{r} = 0.53 \cdot b + 0.22 \cdot t + 0.24 \cdot h$$

Durch die Verallgemeinerung des ITD-Modells auf TOAs an einem Ohr und die Erweiterung dieses Modells auf mehrere Freiheitsgrade, können TOAs nicht nur mit anthropometrischen Daten in Beziehung gesetzt werden, sondern auch die genaue Kopfposition der Personen während der Messung nachträglich bestimmt werden. Ebenfalls können die TOAs aus den HRTFs entfernt und mittels weniger Parameter exakt rekonstruiert werden, ohne einen störenden Zeitversatz bei binauraler Wiedergabe multipler Schallquellen zu erzeugen.

Literatur

- [1] S. Mehrgardt; V. Mellert (1977): Transformation characteristics of the external human ear. *J. Acoust. Soc. Am.*, 61:1567-76
- [2] L. Wang; F. Yin; Chen Z. (2009): A hybrid compression method for head-related transfer functions. *Applied Acoustics*, 70 (2009):1212-1218
- [3] Kulkarni, A.; Isabelle, S. K.; Colburn, H. S. (1999): Sensitivity of human subjects to head-related transfer-function phase spectra. *J. Acoust. Soc. Am.*, 105(5):2821-2840
- [4] Wightman, F.L. and Kistler, D.J. (1992): A model of head-related transfer functions based on principal components analysis and minimum-phase reconstruction. *J. Acoust. Soc. Am.*, 91(3):1637-1647
- [5] Nam, J.; Abel, Jonathan S.; Smith, Julius O. (2008): A Method for Estimating Interaural Time Difference for Binaural Synthesis, Proc. 125th AES Convention (San Francisco), preprint 7612, 2008
- [6] Jot, J.-M.; Larcher, V.; Warusfel, O. (1995): Digital signal processing issues in the context of binaural and transaural stereophony, in the 98th Audio Engineering Society (AES) Convention, preprint no. 3980
- [7] Minnaar, P.; Plogsties, J.; Olesen, S. K.; Christensen, F.; Møller, H. (2000): The interaural time difference in binaural synthesis, Proceedings of the 108th Audio Engineering Society Conventions. 1-20
- [8] Busson, S.; Nicol, R.; Katz, B.F.G. (2005): Subjective investigations of the interaural time difference in the horizontal plane, Proc. 118th AES Convention, Barcelona, mai-2005
- [9] Algazi, V. Ralph; Avendano, C. D. R. O. (2001): Estimation of a spherical-head model from anthropometry. *J. Audio Eng. Soc.*, 49(6):472-479