Binaurale Raumsynthese mittels Wellenfeldsynthese - Realisierung und Evaluierung

D. Menzel[1], H. Wittek[3,2], H. Fastl[1] und G. Theile[2]

[1] *AG Technische Akustik, MMK, TU München, Arcisstr. 21, D-80333 München, men@mmk.ei.tum.de*

[2] *Audio Systems and Engineering, IRT München*

[3] *Schalltechnik Dr.-Ing. Schoeps GmbH*

Einleitung

Mit Hilfe der Binauralen Raumsynthese (BRS) ist es möglich, stabile virtuelle Schallquellen zu erzeugen. Dazu wird das Nutzsignal mit binauralen Raumimpulsantworten (BRIR) in Echtzeit gefaltet und über Kopfhörer wiedergegeben. Durch den Einsatz eines Headtrackers wird der aktuelle horizontale Kopfwinkel des Zuhörers ermittelt und bei der Faltung berücksichtigt, sodass Probleme wie vorne-hinten Vertauschung oder Im-Kopf-Lokalisation vermieden werden [1]. Es ergibt sich für den Zuhörer der Eindruck einer außerhalb des Kopfes gelegenen virtuellen Schallquelle, die auch bei Kopfdrehungen ihre Position beibehält.

Für Anwendungen, bei denen das Tragen eines Kopfhörers hinderlich oder nicht möglich ist, z. B. im Auto oder in Augmented Reality Installationen, ist ein BRS-System ohne Kopfhörer wünschenswert. Diese Arbeit erweitert daher das bestehende BRS-Konzept um einen „virtuellen Kopfhörer“.

Konzept

Sollen binaurale Signale über normale Lautsprecher abgespielt werden, muss das Übersprechen zwischen linkem Kanal und rechtem Ohr und umgekehrt mittels Crosstalk-Cancellation (XTC) vermieden werden. Dazu müssen die kopfbezogenen Übertragungsfunktionen (HRTF) der beteiligten Lautsprecher bekannt sein [2]. Ändern sich jedoch die HRTF, weil etwa der Kopf gedreht wird, müssen die XTC-Filter angepasst werden, beispielsweise mittels mathematischer Modelle (siehe z. B. [3]).

Eine dynamische Anpassung der XTC-Filter kann entfallen, wenn die HRTF konstant bleiben, sich also die relative Position der Lautsprecher in Bezug zu den Ohren nicht ändert. Die zur Crosstalk-Cancellation verwendeten Schallquellen müssen sich also synchron mit dem Kopf mitdrehen, was mit normalen Lautsprechern natürlich nicht realisierbar ist.

Es wurden daher kopfnahe fokussierte Quellen verwendet, die mittels Wellenfeldsynthese (WFS) erzeugt werden. Diese fokussierten Schallquellen dienen als „Lautsprecher“ zur Übertragung der XTC-gefilterten binauralen Signale. Da die fokussierten Quellen durch ein Lautsprecher-Array generiert werden, können sie durch Anpassung der WFS-Parameter (Verzögerung und Pegel der einzelnen Array-Lautsprecher) einfach im Raum verschoben werden und über einen Headtracker mit der Kopfdrehung synchronisiert werden.

Realisierung

Zur Synthetisierung der fokussierten Quellen wird ein kreisförmiges Lautsprecher-Array verwendet, das 40 cm über dem Kopf des Zuhörers angebracht ist (siehe Abbildung 1). Durch die Kreisform wird gewährleistet, dass auch bei Kopfdrehungen der Abstand zwischen Ohren und Array-Lautsprechern konstant bleibt, was zu einer gleichbleibenden Aliasfrequenz führt und Klangverfärbungen minimiert. Die Montage des Arrays über dem Kopf hat praktische Gründe und ermöglicht eine freie Sicht nach vorne, z. B. auf Computer- oder TV-Bildschirme.

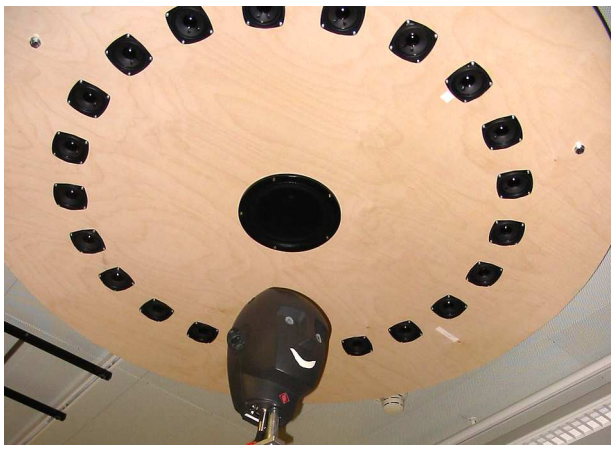

Abbildung 1: Das zur Erzeugung der fokussierten Quellen verwendete kreisförmige Lautsprecher-Array.

Das Array hat einen Durchmesser von 1 m und besteht aus 22 Breitbandlautsprechern und einem Tieftöner, die in eine Schallwand eingebaut sind. Die Signalverarbeitung erfolgt PC-basiert in Echtzeit, das Nutzsignal wird dabei mit vorberechneten WFS- und XTC-Filtern gefaltet. Hierbei sind nur die WFS-Filter vom aktuellen horizontalen Kopfwinkel abhängig, die XTC-Filter sind für alle Kopfrichtungen identisch.

Das Array arbeitet somit als „virtueller Kopfhörer“: Ein zweikanaliges Eingangssignal (z. B. das binaurale Signal eines BRS-Systems [1]) kann ohne nennenswertes Übersprechen dem jeweils korrekten Ohr zugeführt werden, auch wenn der Zuhörer den Kopf dreht.

Objektive Evaluierung

Es zeigte sich in Simulationen und Messungen, dass die unerwünschten Übersprech-Anteile unterhalb von etwa 7 kHz um 10 bis 20 dB gedämpft werden. Oberhalb der Grenzfrequenz von 7 kHz lässt die Dämpfung aufgrund von Aliaseffekten deutlich nach. Da im Moment nur Kopfdrehungen berücksichtigt werden, nimmt die Effizienz der Crosstalk-Cancellation bei Verlassen der zentralen Hörposition unter dem Array schnell ab. Hörversuche und Messungen mit dem Kunstkopf ergaben, dass sich ab einer seitlichen Kopfverschiebung von etwa 8 cm kein sinnvoller Klangeindruck mehr ergibt.

Hörversuche

Wie gut die virtuellen Schallquellen lokalisierbar sind, wurde in Hörversuchen evaluiert. Dabei wurden sechs reale Schallquellen, die in einem Abstand von 1 m rund um die Versuchsperson platziert waren, mit vom Lautsprecher-Array erzeugten virtuellen Quellen verglichen. Es zeigte sich, dass der Azimut der virtuellen Quellen sehr gut wahrgenommen werden kann (Abbildung 2), vergleichbar mit realen Quellen oder kopfhörerbasierter BRS-Wiedergabe [4].

Abweichungen entstehen hauptsächlich bei der Elevationswahrnehmung: Die Quellen werden um etwa 10° zu hoch gehört, außerdem scheint eine Reproduktion von unterhalb der Ohrenhöhe gelegenen Quellen problematisch (siehe Quellen Nr. 3 und 6 in Abbildung 3). Diese Phänomene sind jedoch schon seit längerem bekannt und treten auch bei Kopfhörerwiedergabe binauraler Signale auf [4].

Eine weitere Untersuchung ergab, dass sich die virtuellen Quellen in ihrer Klangfarbe nicht nennenswert unterscheiden. Vier vom Array erzeugte virtuelle Quellen wurden dazu untereinander hinsichtlich ihrer Klangfarbe verglichen. Es ergaben sich minimale Abweichungen, die von den Versuchspersonen als „wahrnehmbar, aber nicht störend“ eingeschätzt wurden.

Zusammenfassung

Das vorgestellte Lautsprecher-Array ist in der Lage, durch Kombination von Wellenfeldsynthese und Crosstalk-Cancellation binaurale Signale wiederzugeben und stabile virtuelle Schallquellen zu erzeugen. Durch die Verwendung eines Headtrackers können sehr gute Ergebnisse bei der horizontalen Lokalisation erzielt werden, sie entsprechen den Ergebnissen, die mit dem BRS-Verfahren bei üblicher Kopfhörerwiedergabe erzielt werden [4]. Probleme ergeben sich vor allem bei der Elevationswahrnehmung, die Reproduktion von Schallquellen unterhalb der Ohrenhöhe ist entsprechend der bekannten Erfahrung mit der Kunstkopftechnik nicht möglich. Klangfarbenunterschiede zwischen den virtuellen Quellen erwiesen sich als unkritisch. Das System ist damit in der Lage, als „virtueller Kopfhörer“ für die Binaurale Raumsynthese eingesetzt zu werden. Es wurde unter dem Namen „Binaural Sky“ erfolgreich zahlreichen Tagungsteilnehmern demonstriert [5].

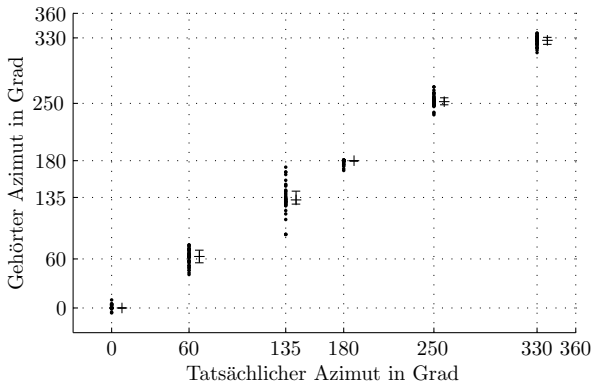

Abbildung 2: Ergebnisse des Hörversuchs zur horizontalen Lokalisation virtueller Schallquellen. Gezeigt sind die Datenpunkte aller Versuchspersonen sowie seitlich versetzt Median und Interquartilbereich.

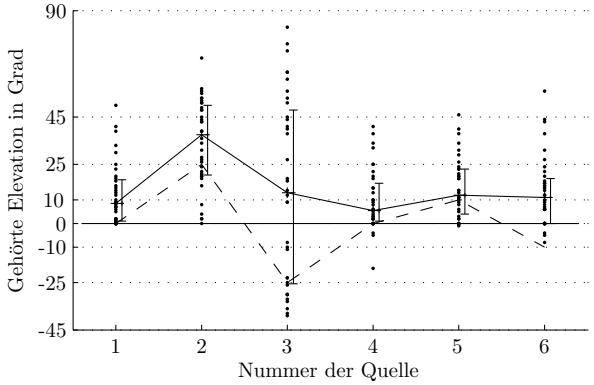

Abbildung 3: Ergebnisse des Hörversuchs zur Elevation virtueller Schallquellen. Gezeigt sind die Datenpunkte aller Versuchspersonen mit Interquartilbereich und Median (durchgezogen) sowie die tatsächlichen Elevationen (gestrichelt).

Literatur

[1] Horbach, U. ; Pellegrini, R. ; Felderhof, U. ; Theile, G. : Ein virtueller Surround Sound Abhörraum im Ü-Wagen. In: 20. Tonmeistertagung, Karlsruhe, Verlag KGSaur, 238-245 (1998)

[2] Bauck, J.; Cooper, D. H.: Generalized Transaural Stereo and Applications. In: J. Audio Eng. Soc. **44**, 683–705 (1996)

[3] Gardner, W. G.: 3-D Audio Using Loudspeakers, Massachusetts Institute of Technology, PhD thesis, September 1997

[4] Spikofski, G. ; Fruhmann, M. : Optimisation of Binaural Room Scanning (BRS): Considering inter-individual HRTF-characteristics. In: AES 19th Int. Conference on Surround Sound, Elmau, 124-135 (2001)

[5] Menzel, D.; Wittek, H.; Theile, G.; Fastl, H.: The Binaural Sky: A Virtual Headphone for Binaural Room Synthesis. In: 1st Int. Tonmeister Symposium, Hohenkammer, 2005. `http://www.tonmeister.de/symposium/2005/np_pdf/R04.pdf`