

# Das Äquator-Mikrofon-Array für binaurale Wiedergabe

Jens Ahrens<sup>1</sup>, Hannes Helmholz<sup>1</sup>, Sebastià V. Amengual Garí<sup>2</sup>, David L. Alon<sup>2</sup>

<sup>1</sup> Chalmers University of Technology, 412 96 Gothenburg, Sweden, E-Mail: jens.ahrens@chalmers.se

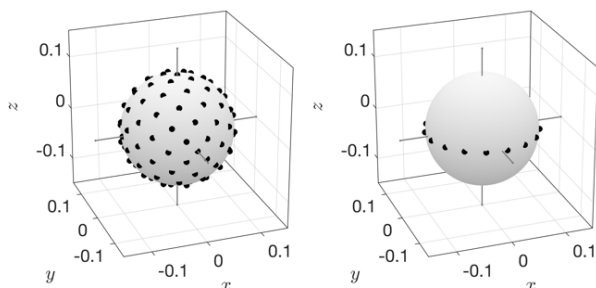
<sup>2</sup> Facebook Reality Labs Research, Facebook, 1 Hacker Way, Menlo Park, California 94025, USA

## Einleitung

Kugelförmige Mikrofon-Arrays mit schallharten Streukörpern stellen ein etabliertes Verfahren für die Aufnahme von Schallfeldern dar. Die binaurale Wiedergabe mit Nachführung der momentanen Kopforientierung der Hörerin/des Hörers ist das typischste Wiedergabeverfahren. Ein zentrales Merkmal der Verarbeitungskette ist eine Repräsentation des aufgenommenen Schallfeldes als Koeffizienten einer Zerlegung in Kugelflächenfunktionen (engl. *spherical harmonics*, SHs).

Kugelförmige Mikrofon-Arrays weisen Eigenschaften auf, die unabhängig von der Einfallrichtung des Schalles sind. Ein signifikanter Nachteil ist, dass sie eine Vielzahl an Mikrofonen erfordern (typischerweise dutzende). Unter der Annahme von rein horizontalem Schalleinfall kann die Mikrofonanordnung derart vereinfacht werden, dass lediglich Mikrofone entlang des Äquators des schallharten Streukörpers benötigt werden (vgl. Abb. 1), wodurch die benötigte Mindestanzahl drastisch sinkt [1]. Ist eine Zerlegung des einfallenden Schallfeldes  $N$ -ter SH-Ordnung beabsichtigt, so erfordern herkömmliche kugelförmige Mikrofon-Arrays (*spherical microphone arrays*, SMAs) mindestens  $(N + 1)^2$  Mikrofone, während das Äquator-Array (*equatorial microphone array*, EMA)  $2N + 1$  erfordert. Vgl. Tab. 1.

Die äquatoriale Mikrofonanordnung wurde bereits z.B. für Beamforming, Quellenlokalisierung o.ä. verwendet, jedoch nicht für die Zerlegungen in Kugelflächenfunktionen. Wir verweisen die Leserin/den Leser auf [1] für eine Zusammenfassung der Literatur. Der vorliegende Beitrag basiert auf [1] und [3] und fasst die wichtigsten Eigenschaften des Äquator-Arrays zusammen.



**Abbildung 1:** Links: Herkömmliches kugelförmiges Mikrofon-Array 8. Ordnung mit 110 Mikrofonen in einer Lebedev-Anordnung. Rechts: Äquator-Array 8. Ordnung mit 17 äquiangular positionierten Mikrofonen.

**Tabelle 1:** Benötigte Mindestanzahl an Mikrofonen für eine Schallfeldzerlegung  $N$ -ter SH-Ordnung für konventionelle Kugel-Arrays (SMA) und das vorgeschlagene Äquator-Array (EMA).

$N$	SMA	EMA
2	9	5
4	25	9
8	81	17
16	289	33
32	1098	65

## Schallfeldzerlegung

Die betrachtete äquatoriale Mikrofonanordnung gestattet es nicht, beliebige Information über das einfallende Schallfeld zu extrahieren. Dies zeigt sich in Mehrdeutigkeiten in der mathematischen Beschreibung und verhindert eine allgemeingültige Zerlegung des einfallenden Schallfeldes in Kugelflächenfunktionen. Es zeigt sich jedoch, dass das Äquator-Array die Koeffizienten der Zerlegung in Kugelflächenfunktionen eines beliebigen höheninvarianten Schallfeldes extrahieren kann. Die in [1] vorgeschlagene Lösung basiert auf dieser Annahme.

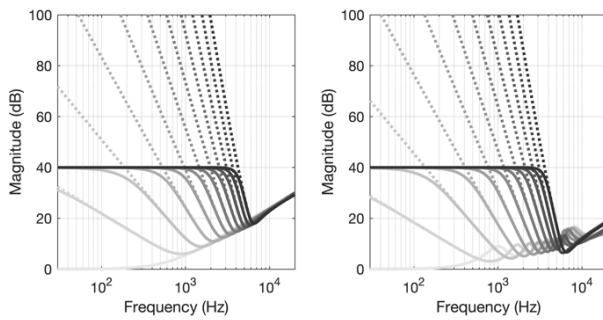
Ist die Annahme der Höheninvarianz erfüllt, dann sind die Schallfeldrepräsentationen, die SMA und EMA gleicher Ordnung extrahieren, gleichwertig. Es ergeben sich kleine Unterschiede in den Eigenschaften des räumlich Aliasing bei hohen Frequenzen. Wird die Annahme der Höheninvarianz des einfallenden Schallfeldes verletzt, so ergeben sich Abweichungen der binauralen Ausgangssignale der Verarbeitungs-Pipeline von den gewünschten binauralen Signalen. Dies wird in den folgenden Abschnitten diskutiert.

## Evaluierung

Die folgende Evaluierung bezieht sich auf die in Abb. 1 dargestellten Arrays, deren Streukörper einen Radius von  $R = 0.0875$  m aufweisen. Die Aliasing-Frequenz beider Arrays liegt bei ca. 5 kHz.

### Radialfilter

Die Verarbeitung der Signale von EMAs erfordert die Anwendung eines Äquivalentes der so genannten Radialfilter, wie sie bei SMAs auftreten. Es zeigt sich, dass die Radialfilter von EMA und SMA die gleichen qualitativen Eigenschaften aufweisen (vgl. Abb. 2). EMA und SMA sind daher ähnlich robust gegenüber Eigenrauschen der Mikrofone sowie Fehlern in deren Positionierung.

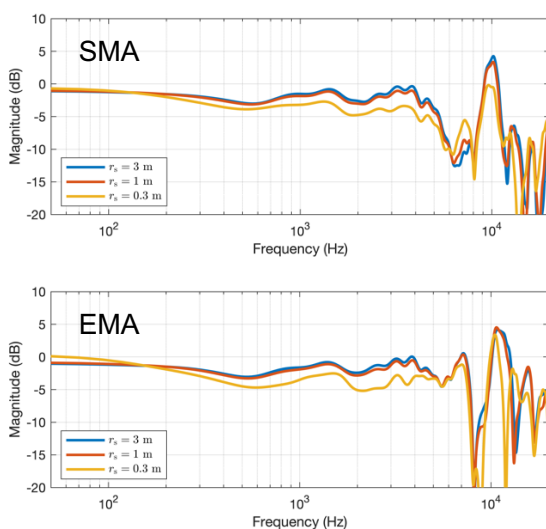


**Abbildung 2:** Beträge der Transferfunktionen der Radialfilter von SMA (links) und EMA (rechts). Gestrichelt: Theoretischer Verlauf. Durchgezogen: Mit Soft-Clipping [2] bei 40 dB über dem Gleichanteil der 0. Ordnung.

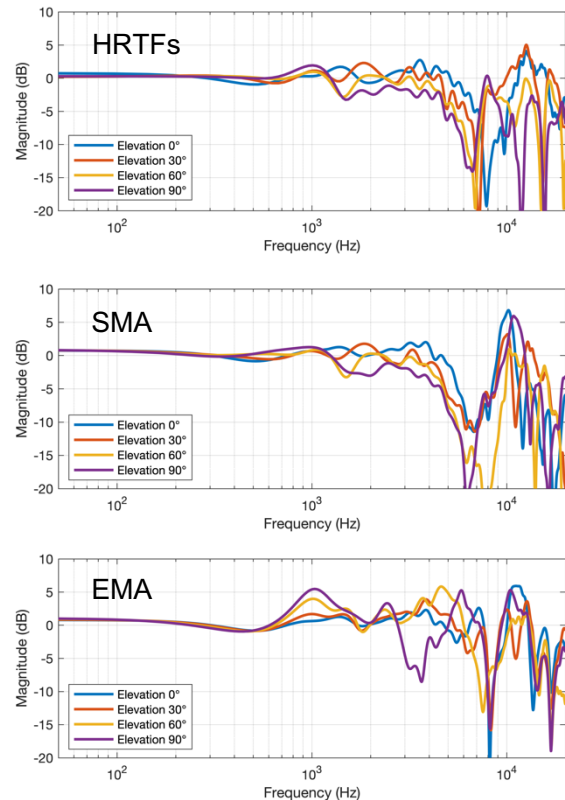
### Binaurale Transferfunktionen

Wie bereits erwähnt, sind die *binauralen Transferfunktionen* (BTFs) von SMA und EMA gleich genau, solange das einfallende Schallfeld höheninvariant ist. Dieser Fall wird hier nicht weiter diskutiert. Wir verweisen den Leser/die Leserin auf [1].

Punktquellen in der Horizontalebene in endlicher Entfernung strahlen Schallfelder ab, die eine schwache Höhenvarianz am Mikrofon-Array aufweisen. Die Höhenvarianz ist stärker ausgeprägt, je näher sich die Quelle am Mikrofon-Array befindet. Abb. 3 vergleicht die BTFs von SMA und EMA für diesen Fall. Ein direkter Vergleich der BTFs mit den entsprechenden Außenohrübertragungsfunktionen (*head-related transfer functions*, HRTFs) der Quellpositionen ist aufgrund nicht zur Verfügung stehender Messdaten nicht möglich. Die SMA-BTFs können unterhalb der Aliasing-Frequenz von 5 kHz als exakt angenommen werden. Abb. 3 zeigt lediglich sehr geringe bis vernachlässigbare Abweichungen der BTFs des EMA von jenen des SMA unterhalb der Aliasing-Frequenz, sogar für die sehr nahe Quellposition im Abstand  $r_s = 0.3$  m.



**Abbildung 3:** Normalisierte Beträge der BTFs von SMA und EMA für Punktquellen in der Horizontalebene in Blickrichtung der Hörerin/des Hörers in unterschiedlichen Abständen  $r_s$ .



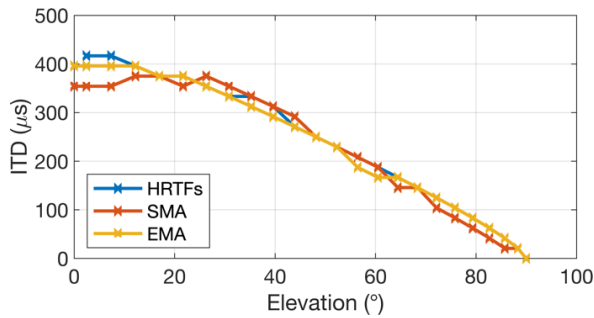
**Abbildung 4:** Beträge der HRTFs wie der BTFs von SMA und EMA für den Einfall ebener Wellen aus nicht-horizontalen Richtungen aus der Blickrichtung des Hörers/der Hörerin.

Abb. 4 zeigt die BTFs für den Einfall ebener Wellen aus nicht-horizontalen Richtungen aus der Blickrichtung des Hörers/der Hörerin. Die entsprechenden HRTFs sind ebenfalls dargestellt. Eine ideale Array-Verarbeitungs-Pipeline würde BTFs aufweisen, die den HRTFs der Situation entsprechen.

Oberhalb der Aliasing-Frequenz von 5 kHz weisen die BTFs sowohl des SMA als auch des EMA, wie zu erwarten, große Abweichungen von den HRTFs auf. Die BTFs des EMA können auch unterhalb der Aliasing-Frequenz um einige dB abweichen. Dies bedeutet, dass monaurale auditorische Lokalisationsmerkmale der Elevation nicht korrekt abgebildet werden. Interessanterweise bleiben jedoch binaurale Lokalisationsmerkmale erhalten, wie in Abb. 5 dargestellt ist. Diese zeigt die interaurale Zeitdifferenz für Schalleinfall aus  $45^\circ$  Azimut und steigender Elevation. Ähnliche Beobachtungen lassen sich auch für die interaurale Pegeldifferenz zeigen [3].

### Zusammenfassung und Diskussion

Das vorgestellte Äquator-Mikrofon-Array erlaubt es ähnliche binaurale Signale auf Basis eines aufgenommenen Schallfeldes zu erzeugen wie mit herkömmlichen Kugel-Arrays. Die dabei benötigte Anzahl von Mikrofonen ist deutlich niedriger (vgl. Tab. 1). Für nicht-horizontalen Schalleinfall ergeben sich Abweichungen der Betragsfrequenzgänge der binauralen Signale des Äquator-



**Abbildung 5:** Interaurale Zeitdifferenz von HRTFs und BTFs des SMA und EMA für Schalleinfall aus 45° Azimut und unterschiedlicher Elevation.

Mikrofon-Arrays im Vergleich zu den idealen Signalen, so dass monaurale auditorische Lokalisationsmerkmale verloren gehen. Binaurale Lokalisationsmerkmale der Elevation bleiben weitgehend erhalten.

Gerade letzterer Umstand ergibt interessante Perspektiven: Zum einen wurde u.a. in [4] gezeigt, dass binaurale Elevationsmerkmale ausreichen können, um eine Wahrnehmung der Elevation zu erzielen. Zum anderen wurde noch nicht gezeigt, dass herkömmliche Kugel-Arrays monaurale Elevationsmerkmale perceptiv korrekt darstellen, da diese hauptsächlich in den Frequenzbereich fallen, der durch räumliches Aliasing beeinträchtigt ist. Mit anderen Worten, es ist unklar, ob herkömmliche Kugel-Arrays trotz der deutlich höheren Anzahl an verwendeten Mikrofonen die Einschränkungen, die Äquator-Mikrofon-Arrays ausweisen, überwinden können. Äquator-Mikrofon-Arrays könnten somit eine kostengünstige Alternative darstellen.

Simulierte Hörbeispiele für Äquator-Mikrofon-Array-Aufnahmen sind verfügbar in [5].

Das Konzept des Äquator-Mikrofon-Arrays wurde in [3,6] auf nicht-kugelförmige Streukörper erweitert und erlaubt somit die Integration eines solchen Arrays in ein Augmented-Reality-Headset, wobei der Kopf der Trägerin/des Trägers des Headsets als Streukörper verwendet wird.

## Danksagung

Wir danken Facebook Reality Labs für die Finanzierung der vorgestellten Arbeit.

## Literatur

- [1] J. Ahrens, H. Helmholtz, D. L. Alon, S. V. Amengual Gari, "Spherical Harmonic Decomposition of a Sound Field Based on Observations Along the Equator of a Rigid Spherical Scatterer" in *J. Acoust. Soc. Am.* 150(2), 2021.
- [2] B. Bernschütz, "Microphone arrays and sound field decomposition for dynamic binaural recording," Ph.D. thesis, Technische Universität Berlin, 2016.
- [3] J. Ahrens, H. Helmholtz, D. L. Alon, S. V. Amengual Gari, "A Head-Mounted Microphone Array for Binaural Rendering," in *Proc. of I3DA*, Bologna, Italy, Sept. 2021.
- [4] D. Ackermann, F. Fiedler, F. Brinkmann, M. Schneider, and S. Weinzierl, "On the acoustic qualities of dynamic pseudo-binaural recordings," *JAES*, vol. 68, no. 6, pp. 418–427, Jul. 2020.
- [5] J. Ahrens, H. Helmholtz, D. Alon, and S. Amengual Gari, "Audio examples accompanying the article 'Spherical harmonic decomposition of a sound field based on observations along the equator of a rigid spherical scatterer', *JASA* (2021)," <https://zenodo.org/record/4805266/export/hx#.YPWDuehKhPY> (Last viewed May 31, 2021).
- [6] J. Ahrens, H. Helmholtz, D. L. Alon, S. V. Amengual Gari, "Spherical Harmonic Decomposition of a Sound Field Based on Microphones Around the Circumference of a Human Head," in *Proc. of IEEE WASPAA*, New Platz, NY, USA, Oct. 2021.