Soft-Limiting der modalen Amplitudenverstärkung bei sphärischen Mikrofonarrays im Plane Wave Decomposition Verfahren

Benjamin Bernschütz[1,3], Christoph Pörschmann[1], Sascha Spors[2], Stefan Weinzierl[3]
[1] FH Köln, Institut für Nachrichtentechnik, Email: benjamin.bernschuetz@fh-koeln.de, christoph.poerschmann@fh-koeln.de
[2] Deutsche Telekom Laboratories / Technische Universität Berlin, Email: sascha.spors@telekom.de
[3] Technische Universität Berlin, Fachgebiet Audiokommunikation, Email: stefan.weinzierl@tu-berlin.de

Einleitung

Beim Plane Wave Decomposition Verfahren wird zunächst das Schallfeld mit einem Mikrofonarray abgetastet und durch räumliche Fouriertransformation in sphärisch harmonische Basisfunktionen verschiedener Moden mit Ordnung n und Modus m zerlegt [1][2]. Wir approximieren bei diskreten räumlichen Abtastpunkten die benötigten Fourierkoeffizienten $P_{nm}(k,r)$ nach:

$$P_{nm}(k,r) \approx \sum_{s=1}^{S} \beta(\Omega_s)\; p(k,r,\Omega_s)\; Y_n^{m^*}(\Omega_s). \tag{1}$$

Dabei ist k die Wellenzahl, r der Arraymessradius im Schallfeld, S die Anzahl der raumdiskreten Abtastpunkte, $\Omega_s \equiv (\vartheta_s, \varphi_s)$ der Raumwinkel bestehend aus Elevation ϑ_s und Azimuth φ_s, und $Y_n^{m^*}$ sind die konjugiert komplexen sphärisch Harmonischen. Der Ausdruck $\beta(\Omega_s)$ beschreibt einen quadraturspezifischen Gewichtungsfaktor. Das Array liefert das Ausgangssignal $y(\Omega_L, k)$:

$$y(\Omega_L,k) = \sum_{n=0}^{N} \sum_{m=-n}^{n} P_{nm}(k,r)\; Y_n^m(\Omega_L)\; \frac{c_n(k)}{b_n(kr)}. \tag{2}$$

Wobei $\Omega_L \equiv (\vartheta_L, \varphi_L)$ die Blickrichtung (Achse) vorgibt und $c_n(k) = 1$ als Aperturgewichtung für eine Plane Wave Decomposition dient. Die Antwort weist eine konstante Richtcharakteristik über der Frequenz auf. Der Ausdruck $b_n(kr)$ beschreibt die modalen Amplituden und hängt von der Beschaffenheit des Mikrofonarrays ab. So ergibt sich beispielsweise für ein *Open Sphere* Array (ohne schallharten Reflektor) und Druckempfänger [3]:

$$b_n(kr) = 4\pi\; i^n\; j_n(kr), \tag{3}$$

mit $i = \sqrt{-1}$ und $j_n(kr)$ als sphärische Besselfunktion der ersten Art. Um die verschiedenen Moden im Ausgangssignal nutzen zu können, müssen ihre Beiträge bei kleinem kr teils erheblich verstärkt werden. Die benötigte Verstärkung ergibt sich aus dem Inversen der modalen Amplituden und wird demnach in Gl. (2) über den Ausdruck $1/b_n(kr)$ erzielt. In realen Anwendungen ist der Signal-Rauschabstand einer Messung begrenzt, sodass eine Verstärkung der Modenbeiträge hoher Ordnung bei kleinem kr überwiegend Rauschen einbringt. Daher muss die Verstärkung der Modenbeiträge begrenzt und an den verfügbaren Signal-Rauschabstand angepasst werden.

Begrenzung der Verstärkung

Um eine Begrenzung der Verstärkung der Modenbeiträge zu erzielen, wird nun der Ausdruck $1/b_n(kr)$ aus Gl. (2) durch eine Filterfunktion $d_n(kr)$ ersetzt:

$$y = \sum_{n=0}^{N} \sum_{m=-n}^{n} P_{nm}(k,r)\; Y_n^m(\Omega_L)\; c_n(k)\; d_n(kr). \tag{4}$$

Sei nun a [dB] die vorgegebene maximal zulässige Verstärkung, die sich aus einem begrenzten Signal-Rauschabstand ergibt. Es soll eine Begrenzung der Verstärkung auf a stattfinden. Es gelte $\alpha = 10^{(a/20)}$. Die Begrenzung sollte jedoch nicht hart erfolgen, vgl. Abb. 1(a), da es sonst zu einem sprungartigen Verhalten in der Arrayantwort kommen würde, vgl. Abb. 2(a). Dies entspricht einer räumlichen Sprungfunktion im Frequenzbereich und führt zu Überschwingern im Ortsbereich. Um dieses zu verhindern, kann die Charakteristik der Begrenzung z.B. unter Anwendung eines Arkustanges verbessert werden:

$$d_n(kr) = \frac{2\;\alpha}{\pi}\; \frac{|b_n(kr)|}{b_n(kr)}\; \arctan\left(\frac{\pi}{2\;\alpha\;|b_n(kr)|}\right). \tag{5}$$

Der Ausdruck führt zu einer Soft-Knee-Charakteristik der Begrenzung, vgl. Abb. 1(b). Infolgedessen entfallen die Kanten in der Arrayantwort, vgl. Abb. 2(b).

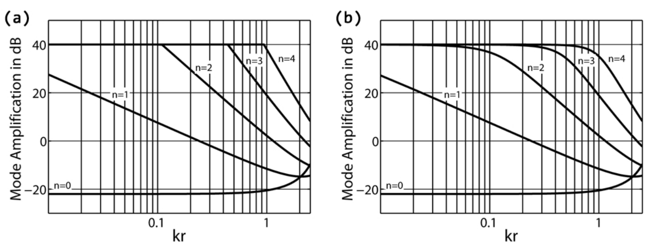

Abbildung 1: Betrag der Modenverstärkung $|d_n(kr)|$ für $n \in \mathbb{N}_0 : n \leq 4$ und $kr \in \mathbb{R} : 0.1 \leq kr \leq 2.5$ und a=+40 dB, (a) mit harter Begrenzung und (b) mit Soft-Limiting.

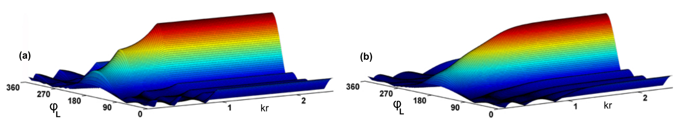

Abbildung 2: Simulierte Arrayantworten $y(\Omega_L, k)$ auf eine ebene Welle aus $\Omega = (90°, 180°)$ mit $N = 4$ für $\Omega_L = (90°, \varphi_L)$, $\varphi_L = [0°, 360°[$ für (a) harte Begrenzung der modalen Verstärkung und (b) Soft-Limiting.

Kompensation des Frequenzgangs

Die Arrayantwort weist durch die stufenweise modale Begrenzung neben der Aufweitung des Beams einen Hochpasscharakter auf. Dieses Verhalten kann je nach Anwendung wünschenswert sein (z.B. bei Sprachsignalen). Es besteht aber auch die Möglichkeit, den Frequenzgang auf Achse durch eine geeignete Filterung der nullten Mode (omnidirektional) ganz oder teilweise zu linearisieren. Bei der nullten Ordnung ist der Signal-Rauschabstand bei kleinem kr nicht nur unproblematisch, sondern sogar besonders günstig (ca. +20 dB). Die modalen Filterfunktionen $d_n^{HC}(kr)$ (HC: Highpass Compensation) ergeben sich zu:

$$d_n^{HC}(kr) = \begin{cases} d_n(kr), & \text{für } n > 0 \\ \xi(kr), & \text{für } n = 0 \end{cases} \tag{6}$$

mit $d_n(kr)$ aus Gl. (5) und

$$\xi(kr) = d_0(kr) + \frac{1}{b_0(kr)} \sum_{n'=0}^{N} (2\,n' + 1)\,(1 - d_{n'}(kr)\; b_{n'}(kr)). \tag{7}$$

Durch Anwendung des Filters aus Gl. (6) wird der Frequenzgang auf Achse linearisiert. Dieses Verhalten ist dann gut geeignet, wenn ein glatter Frequenzgang und eine große Frequenzbandbreite wichtiger als eine hohe Richtwirkung sind (z.B. bei Musikaufnahmen).

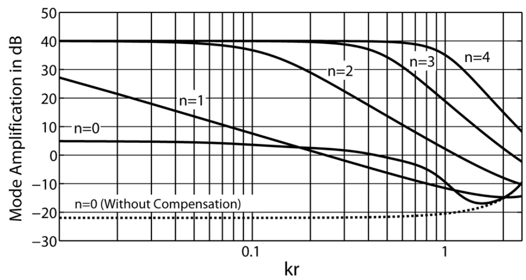

Abbildung 3: Betrag der modalen Verstärkung $|d_n^{HC}(kr)|$ für $n \in \mathbb{N}_0 : n \leq 4$ und $kr \in \mathbb{R} : 0.1 \leq kr \leq 2.5$ und a=+40 dB mit Frequenzgangkompensation auf Achse durch die nullte Mode $n = 0$.

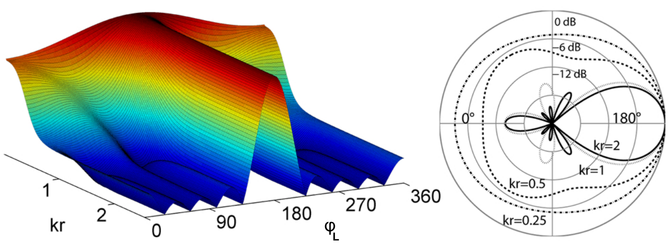

Abbildung 4: Simulierte Arrayantwort $y(\Omega_L, k)$ auf eine ebene Welle aus $\Omega = (90°, 180°)$ für $\Omega_L = (90°, \varphi_L)$, $\varphi_L = [0°, 360°[$, $N = 4$ mit Soft-Limiting und Frequenzgangkompensation durch $d_n^{HC}(kr)$.

Bei *Open Sphere* Arrays mit Druckempfängern darf die Kompensation durch $d_n^{HC}(kr)$ nur im untersten kr-Band für ca. $kr < 2.5$, also unterhalb der Polstellen der nullten Mode eingesetzt werden. Die Koeffizienten müssen für eine bestimmte Arraykonfiguration bei festgelegter maximaler Verstärkung jeweils nur ein Mal berechnet werden und können dann gespeichert werden.

Verifikation

Die Filter wurden in einem definierten realen Szenario getestet. Dazu wurde ein Array (VariSphear) als offene Sphäre mit Niere Großmembran (Microtech Gefell M900) auf 86 Abtastpunkten (Lebedev) und r = 0.25 m im reflexionsarmen Raum einer Wellenfront aus 180° AZ/90° EL ausgesetzt. Die Wellenfront wurde mit einem Lautsprecher erzeugt (AD-Systems Flex15). Die Zerlegung erfolgte auf der Maximalordnung $N = 7$ und als modales Verstärkungslimit wurde a = +5 dB angesetzt.

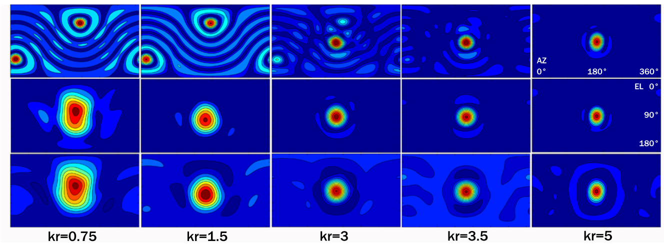

Abbildung 5: Real gemessene Arrayantworten auf eine Wellenfront aus 180° AZ/90° EL. Die obere Reihe zeigt Antworten ohne das vorgestellte Verfahren mit konventioneller Filterung durch $1/b_n(kr)$. Bei $N = 7$ bricht die Arrayantwort für $kr < 3$ auf. Die mittlere Reihe zeigt die Filterung mit $d_n(kr)$ (normalisiert). Die Arrayantwort weitet sich auf, bleibt aber immer stabil. Die untere Reihe zeigt die frequenzganglinearisierte Variante mit Filterung durch $d_n^{HC}(kr)$. Hier erscheinen mehr omnidirektionale Anteile.

Zusammenfassung

Es wurde ein Verfahren für die Behandlung des unteren kr-Bandes sphärischer Arrays vorgestellt und verifiziert. Bei dem Verfahren wird die Verstärkung der Modenbeiträge mit einem Soft-Knee-Verhalten begrenzt und so verhindert, dass mitverstärktes Rauschen zu einer undefinierten Arrayantwort führt. Das Array arbeitet dabei immer automatisch auf der höchstmöglichen Ordnung, die durch eine vorgegebene maximale Modenverstärkung erreichbar ist. Durch die Begrenzung der modalen Verstärkung stellt sich auf Achse eine Hochpasscharakteristik ein, die bei Bedarf durch geeignete dynamische Filterung der nullten Mode zur Frequenzganglinearisierung aufgeholt werden kann. Das Verfahren liefert zudem einen Denkansatz für andere mögliche Filterfunktionen und kann auch bei der Kombination verschiedener Sphärenradien nützlich sein.

Literatur

[1] William, E.G.: Fourier Acoustics. Academic Press, San Diego, 1999

[2] Meyer, J. & Elko, G.: A Highly Scalable Spherical Microphone Array Based on an Orthonormal Decomposition of the Soundfield. IEEE ICASSP 2002, vol. 2, pp. 1781-1784

[3] Rafaely, B.: Analysis and design of spherical microphone arrays. IEEE Transactions on Speech and Audio Processing 2005, vol. 13, no. 1, pp. 135–143