

Lokalisation der Trajektorien-Endpunkte einer bewegten Schallquelle in Rauschen

Norbert F. Bischof & Bernhard U. Seeber

*Audio-Signalverarbeitung, Technische Universität München, 80333 München
{norbert.bischof, seeber}@tum.de*

Einleitung

In vielen Alltagssituationen, wie zum Beispiel im Straßenverkehr oder auch bei einem geselligen Abend in einem gut besuchten Restaurant, müssen Schallquellen in Rauschen lokalisiert und daher auch detektiert werden. Interaurale Zeit- und Pegeldifferenzen (ITDs & ILDs) ermöglichen dem Menschen, akkurat in der Horizontalebene Schallereignisse zu orten [1]. Dabei werden für tiefe Frequenzen bis ca. 1500 Hz vorwiegend ITDs und für höhere Frequenzen zusätzlich ILDs zur Lokalisation einer Schallquelle ausgewertet [1,2,3]. Seeber [4] zeigte zudem, dass für eine natürliche Kombination der interauralen Merkmale, wie sie im Freifeld auftreten, die Auswertung der ITDs auch für Signale mit vorwiegend höheren Frequenzen dominiert. Für tonale Signale lässt sich die ITD dabei direkt in eine Phasendifferenz überführen. Die interaurale Phasendifferenz (IPD) eines Zielschalls kann in binaural korreliertem Rauschen die Detektionsschwellen des Testsignals positiv beeinflussen, was auch als binaurale Entmaskierung bekannt ist [5,6]. Auch wenn die Lokalisation von Schallen und deren binaurale Entmaskierung in Rauschen auf ähnlichen binauralen Merkmalen basieren, können sich die zugrundeliegenden Mechanismen im auditorischen System maßgeblich voneinander unterscheiden. Im Gegensatz zu Untersuchungen mit Kopfhörern ist die Schalldarbietung im Freifeld stets binaural, weshalb die Lokalisation in Rauschen auf zwei grundlegenden Mechanismen zur selben Zeit basiert. Einerseits wird das Testsignal durch eine Änderung in der interauralen Korrelation binaural entmaskiert und dadurch hörbar. Die Lokalisation eines bewegten Testtons unter 1500 Hz basiert überwiegend jedoch nicht ausschließlich auf der Auswertung der ITDs [7]. Da beide Mechanismen auf denselben Merkmalen des Testsignals beruhen, kann es zu einer Interaktion der beiden Mechanismen kommen. Das Ziel dieser Pilotstudie ist es, für sich bewegende Schalle in Rauschen den Zusammenhang der Mechanismen für die Lokalisation und für die Entmaskierung näher zu untersuchen.

In vielen Alltagssituationen sind jedoch die zu lokalisierenden Schallquellen nicht statisch, sondern bewegen sich im Raum. Die sich durch die Bewegung zeitlich ändernden binauralen Merkmale können zu deutlich weniger Entmaskierung führen als statisch dargebotene Signale mit denselben maximalen interauralen Differenzen [8]. Hierbei spricht man von einer Trägheit des auditorischen Systems bei sich zeitlich ändernden binauralen Merkmalen [9]. Aus der Literatur ist zudem bekannt, dass sowohl Lokalisation als auch Entmaskierung bewegter Quellen Trägheit zeigen. Der minimal wahrnehmbare Winkel, den eine bewegte Quelle zurücklegt, nimmt mit zunehmender Geschwindigkeit zu [10,11]. Sich schnell ändernde IPDs führen zu einer Reduktion der binauralen Entmaskierung [9] welche auch bei

sich bewegenden Quellen im Freifeld beobachtet werden kann [8] und somit auf eine Trägheit des zugrundeliegenden Mechanismus deutet.

Das hier vorgestellte Experiment basiert dabei auf beiden Mechanismen, um eine sich im Freifeld bewegende Schallquelle in Rauschen zu lokalisieren. Bisherige Studien fokussierten sich mehr auf die Lokalisation von statischen Quellen nahe der Detektionsschwelle [12-15]. In dieser Studie wurde die Lokalisation für einen sich bewegenden harmonisch komplexen Ton um 500 Hz nahe der Detektionsschwelle in einem Oktavbandrauschen von vorne gemessen. Dabei ist die zentrale Fragestellung der Studie, ob der Ort der Lokalisation dem Ort der Entmaskierung entspricht und eine dynamische Interaktion der beiden zugrundeliegenden Mechanismen vorliegt.

Methoden

Stimuli

In diesem Experiment wurde ein harmonischer komplexer Ton mit einer Mittenfrequenz von 500 Hz und einer Grundfrequenz von 100 Hz als Zielsignal verwendet. Es wurden die vierte bis sechste Harmonische verwendet, wodurch jede harmonische Komponente in einer separaten Frequenzgruppe zu liegen kommt. Als Maskierungsrauschen wurde ein Oktavbandrauschen mit einer Mittenfrequenz von 500 Hz und einem konstanten Schalldruckpegel von 60 dB an der Hörerposition verwendet. Der Testschall wies eine effektive Länge von 300 ms auf, das Rauschen von 600 ms, jeweils mit 10 ms Anstiegs- und Abfallzeit. Das Testsignal wurde zeitlich zentriert im Maskierungsrauschen dargeboten.

Es wurden sechs konstante Winkelgeschwindigkeiten für den bewegten Teststimulus verwendet: 10°/s, 30°/s, 50°/s, 100°/s, 150°/s und 200°/s. Diese Geschwindigkeiten korrespondieren mit einem maximalen lateralen Winkel von 3°, 9°, 15°, 30°, 45° oder 60°, der in der effektiven Dauer von 300 ms und einem Startwinkel von 0° erreicht werden kann. Die Startposition der Trajektorie, die mit dem Beginn der effektiven Dauer des Teststimulus übereinstimmt, war immer auf 0° von vorne, wobei der Stimulus eine laterale Bewegungsrichtung nach links oder rechts aufwies. Der Schalldruckpegel der einzelnen Teststimuli wurde 5 dB über den in Kolotzek & Seeber [8] gemessenen Detektionsschwellen für einen bewegten Stimulus gesetzt. Die hier verwendeten Trajektorien stimmen mit denen in [8] überein. Somit war der Pegel der Testsignale abhängig von der getesteten Winkelgeschwindigkeit. Der Schalldruckpegel wurde jedoch für die gesamte Dauer eines jeden Stimulus konstant gehalten. Dies führt zu einer teilweisen Maskierung des bewegten Stimulus. Der Verlauf des Pegels relativ zum Detektionspegel (Empfindungspegel) ist exemplarisch für eine Trajektorie mit 200°/s in Abbildung 1 dargestellt.

Alle Teststimuli wurden mit Ambisonics 17. Ordnung generiert und über die 36 horizontalen Lautsprecher des Simulated Open Field Environment (SOFE v4) [16] in einem reflexionsarmen Raum wiedergegeben. Das Rauschen wurde von einem einzigen Lautsprecher an 0° von vorne abgespielt, was zu einem annähernd interaural korrelierten Rauschen führt. Die Lautsprecher wurden sowohl im Frequenzgang als auch im Phasengang entzerrt.

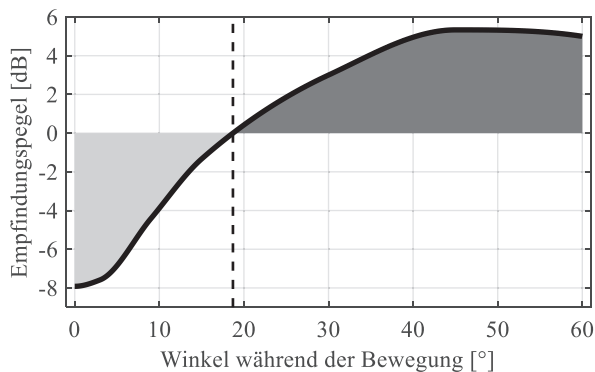


Abbildung 1: Empfindungspegel des Teststimulus für eine konstante Winkelgeschwindigkeit von $200^\circ/\text{s}$. Der Stimulus durchläuft hierbei in 300 ms effektiver Dauer Winkel von 0° bis 60° mit einem konstanten Pegel 5 dB über dem gemessenen Detektionspegel in [5]. Die hellgrau markierten Winkel liegen dabei unterhalb der Detektionsschwelle und der Stimulus sollte dabei vom Rauschen maskiert sein, im dunkelgrau markierten Teil der Trajektorie ist der Stimulus entmaskiert. Die gestrichelte Linie gibt den Winkel an, an dem der Teststimulus detektierbar wird.

Teilnehmer

Es wurden Pilotdaten von vier männliche Probanden im Alter von 23 bis 31 Jahren gesammelt. Die Teilnehmer sind normalhörend mit einem Hörverlust von weniger als 15 dB bis 8 kHz. Alle Teilnehmer gab eine schriftliche Einwilligung zur Teilnahme ab und wurden nicht für die Teilnahme bezahlt. Die Studie wurde von der Ethikkommission 65/18S der TUM genehmigt.

Experimenteller Ablauf

Alle Teilnehmer saßen im komplett abgedunkelten reflexionsarmen Raum in der Mitte des Lautsprecherrings. Sie wurden ausdrücklich angewiesen, den Kopf während der Schallwiedergabe nicht zu bewegen und nach vorne zu schauen. Nach der Wiedergabe sollten die Teilnehmer zunächst den Startpunkt der wahrgenommenen Trajektorie anzeigen. In einem weiteren Durchgang wurde dann der wahrgenommene Endpunkt des bewegten Stimulus abgefragt. Die Teilnehmer zeigten die lokalisierten Positionen mit der *ProDePo* Methode an [17]. Hierfür wurde ein roter Lichtpunkt zufällig im Bereich $\pm 30^\circ$ um den zu lokalisierenden Winkel auf akustisch transparenten Leinwänden vor den Lautsprechern projiziert. Die Teilnehmer bewegten nach der Schallwiedergabe den Lichtpunkt mit einem Trackball an die wahrgenommene Start- bzw. Endposition der Bewegungstrajektorie und bestätigten die Eingabe. Jede Geschwindigkeit wurde insgesamt sechs Mal wiederholt. Da die Lokalisationsdaten für eine rechtslaufende und linkslaufende laterale Bewegungstrajektorie annähernd symmetrisch waren, wurden

die Daten der linken und rechten Hemisphäre für die weitere Analyse zusammengelegt.

Ergebnisse des Experimentes

Lokalisierte Startposition der Trajektorie

Abbildung 2 zeigt die Mediane der lokalisierten Startposition des bewegten Stimulus in Rauschen in Abhängigkeit der dargebotenen Winkelgeschwindigkeit. Wie bereits in den Methoden beschrieben, wurde als Startposition immer 0° gewählt. Da jedoch Teile der Trajektorie abhängig von der maximalen lateralen Endposition und somit auch von der Geschwindigkeit unterhalb der Maskierungsschwelle liegen, wurde erwartet, dass die Teilnehmer für den Startpunkt den Winkel anzeigen, an dem der Stimulus gerade entmaskiert ist (obere gestrichelte Linie in Abbildung 1). Für geringe Geschwindigkeiten bis $50^\circ/\text{s}$ entspricht die wahrgenommene Startposition in etwa der Position, an dem der Stimulus entmaskiert ist. Für $10^\circ/\text{s}$ bzw. $30^\circ/\text{s}$ ist die gesamte Trajektorie entmaskiert, bei $50^\circ/\text{s}$ erst ab 6° . Lediglich bei $30^\circ/\text{s}$ fällt auf, dass die Startposition um 4.7° überschätzt wird. Für höhere Geschwindigkeiten wird die Startposition um ca. 0° lokalisiert. Dies entspricht zwar der realen Startposition, jedoch ist für diese Geschwindigkeiten der vordere Teil der Trajektorie maskiert. Die maskierten Bereiche sind in Abbildung 2 in hellgrau hervorgehoben.

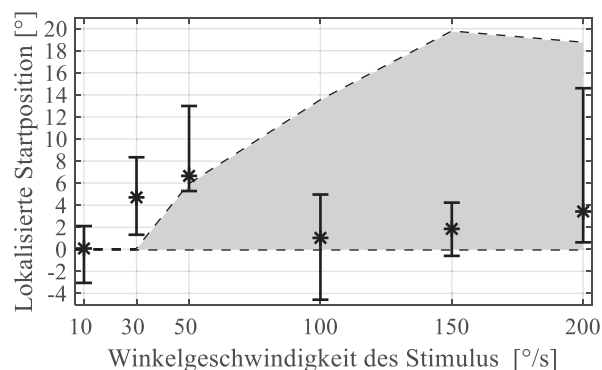


Abbildung 2: Mediane der lokalisierten Startposition einer lateralen Bewegungstrajektorie in Rauschen nahe der Detektionsschwelle. Die laterale Bewegungstrajektorie hatte einen fixen Startpunkt auf 0° . In Abhängigkeit der Geschwindigkeit waren Teile der Trajektorie vom Rauschen maskiert (hellgrau markierte Fläche). Für geringe Geschwindigkeiten entsprechen die lokalisierten Startpositionen in etwa den entmaskierten Startpositionen der Trajektorie. Für höhere Geschwindigkeiten wird der Startpunkt tendenziell nahe 0° lokalisiert.

Lokalisierte Endposition der Trajektorie

In Abbildung 3 sind die Mediane der wahrgenommenen Endposition des bewegten Stimulus in Rauschen für die dargebotenen Winkelgeschwindigkeiten dargestellt. Die maximale laterale Endposition aller Bewegungstrajektorien lag oberhalb der Detektionsschwelle ($+5$ dB Empfindungspegel) und sollte daher für alle dargebotenen Trajektorien hörbar sein. Wie bereits für die Startposition zu beobachten war, stimmen für geringe Winkelgeschwindigkeiten bis $50^\circ/\text{s}$ auch die lokalisierten Endpositionen mit den tatsächlich dargebotenen Trajektorien-Endpositionen weitestgehend überein (diagonal gestrichelte

Linie in Abbildung 3). Jedoch ist hier zu erwähnen, dass die lokalisierte Start- und Endposition eng zusammenliegen und durch die sich deutlich überlappenden Quartile ein signifikanter Unterschied nicht erwartet wird. Für höhere Winkelgeschwindigkeiten wird die Endposition in etwa am Ort der beginnenden Entmaskierung lokalisiert (mittlere gestrichelte Linie in Abbildung 3). Eine Ausnahme bildet hier eine Geschwindigkeit von 200°/s. Jedoch wird auch hier die entmaskierte Endposition an 60° deutlich unterschätzt und tendenziell an der ersten entmaskierten Position lokalisiert.

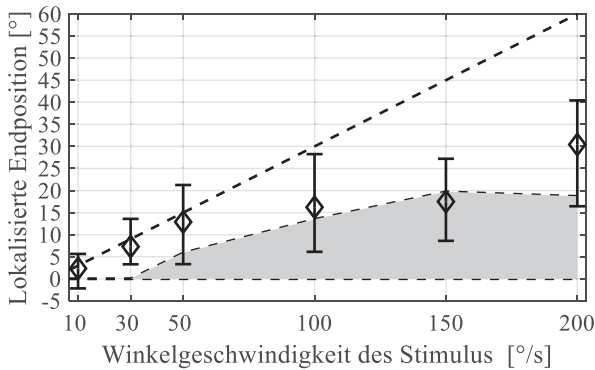


Abbildung 3: Mediane der lokalisierten Endposition einer lateralen Bewegungstrajektorie in Rauschen nahe der Detektionsschwelle. Die laterale Bewegungstrajektorie hatte einen fixen Startpunkt auf 0°. In Anhängigkeit der Geschwindigkeit waren Teile des Trajektorie vom Rauschen maskiert (hellgrau markierte Fläche). Für geringe Geschwindigkeiten entsprechen die lokalisierten Endpositionen in etwa den maximal lateral erreichten Endpositionen der Bewegungstrajektorie. Für höhere Geschwindigkeiten wird der Endpunkt sehr unterschätzt und tendenziell nahe der entmaskierten Startposition lokalisiert.

Schlussfolgerung

Das Ziel dieser Studie war es zu untersuchen, ob der Ort der Lokalisation dem Ort der Entmaskierung für bewegte Schalle nahe der Maskierungsschwelle entspricht und eine dynamische Interaktion der beiden zugrundeliegenden Mechanismen vorliegt. Für einen bewegten harmonisch komplexen Ton mit einer Mittenfrequenz von 500 Hz, der sich mit geringen Winkelgeschwindigkeiten von bis 50°/s in binaural korreliertem Rauschen bewegt, entsprachen die lokalisierten Start- und Endpositionen den entmaskierten Start- und Endpositionen der Bewegungstrajektorie. Für höhere Winkelgeschwindigkeiten wurde die Endposition der Trajektorie deutlich unterschätzt. Die Endposition wurde tendenziell am Ort der Entmaskierung wahrgenommen. Getzmann & Lewald [18] zeigten, dass die Endposition einer Bewegungstrajektorie überschätzt wird. In ihrer Studie verwendeten sie einen bandbegrenzten Stimulus ohne Maskierungsrauschen. In dem vorliegenden Experiment konnte jedoch das Gegenteil beobachtet werden. Die Beobachtung kann auch mit der Unterschätzung der statischen Schallposition im Beisein eines Rauschens und einer Clusterbildung der Lokalisation in Rauschen [13] erklärt werden. Diese Clusterbildung könnte auch die Lokalisation der Startposition bei höheren Winkelgeschwindigkeiten zu 0° erklären, obwohl diese vom Rauschen maskiert ist. Möglich ist hier, dass zwar der Stimulus erst ab dem Ort der

Entmaskierung wahrnehmbar ist, die Clusterbildung der Lokalisation jedoch den Ort der Lokalisation auf 0° rückt.

Danksagung

Diese Forschungsarbeit wurde durch die TUM und durch die Deutsche Forschungsgemeinschaft (DFG) – Projektnummer 352015383 – SFB 1330 C5 finanziert. Die rtSOFE wurde aus dem BMBF 01 GQ 1004B finanziert.

Literatur

- [1] Blauert, J.: Spatial hearing: the psychology of human sound localization, Cambridge, Mass, MIT Press, 1999.
- [2] Wightman, F.L. & Kistler, D.J.: Resolution of front-back ambiguity in spatial hearing by listener and source movement, *J. Acoust. Soc. Am.* 105 (1999), 2841-2853.
- [3] Wiggins, I.M., Seeber, B.U.: Dynamic-range compression affects the lateral position of sounds. *J. Acoust. Soc. Am.* 130 (2011), 3939-3953.
- [4] Seeber, B.U.: The duplex-theory of localization investigated under natural conditions. In: Calvo-Manzano A, Pérez-López A, Santiago S, editors. Proc ICA 2007, 19th Int Congress on Acoustics, Madrid, Spain, 2-7092007: Int. Commission on Acoustics; p. PPA-05-005-IP (2007), 1-6.
- [5] Jeffress, L.A., Blodgett, H.C., Deatherage, B.H.: The Masking of Tone by White Noise as a Function of Interaural Time or Phase Displacement. *J. Acoust. Soc. Am.* 25 (1953), 190.
- [6] Breebart, J., van de Par, S., Kohlrausch, A.: The contribution of static and dynamic varying ITDs and IIDs to binaural detection, *J. Acoust. Soc. Am.* 106 (1999), 979-992.
- [7] Seeber, B.U., Baumann, U., Fastl, H.: Localization ability with bimodal hearing aids and bilateral cochlear implants, *J. Acoust. Soc. Am.* 116 (2004), 1698-1709.
- [8] Kolotzek, N. & Seeber, B.U.: Spatial unmasking of circular moving sound sources in the free field, Proc. of 23rd International Congress on Acoustics, integrating 4th EAA Euroregio 2019, Aachen, Germany, German Acoustical Society e.V. (DEGA) (2019), 1609-1614.
- [9] Grantham, D.W. & Wightman, F.L.: Detectability of a pulsed tone in the presence of a masker with time-varying interaural correlation, *J. Acoust. Soc. Am.*, 65 (1979), 1509-1517.
- [10] Chandler, D.W. & Grantham, D.W.: Minimum audible movement angle in the horizontal plane as a function of stimulus frequency and bandwidth, source azimuth, and velocity, *J. Acoust. Soc. Am.* 91 (1992), 1624-1636.
- [11] Saberi, K. & Perrott, D.R.: Minimum audible movement angle as a function of sound source trajectory, *J. Audiol. Res.* 88 (1990), 2639-2644.
- [12] Altshuler, M.W. & Comalli, P.E.: Effects of stimulus intensity and frequency on median horizontal plane sound localization, *J. Audio. Res.* 15 (1975), 262-265.
- [13] Breebaart, J.: Auditory Localization and Detection in Multiple-Sound-Source Scenarios, Dissertation, Ruhr Universität Bochum, Bochum, 2001.
- [14] Sabin, T.A., Macpherson, E.A., Middlebrooks, J.C.: Homan sound localization at near threshold levels, *Hear. Res.* 199 (2005), 124-134.

- [15] Wiggings, I.M.: Effects of Dynamic-Range Compression in Spatial Hearing, Dissertation, University of Nottingham, Nottingham, UK, 2013.
- [16] Seeber, B.U., Kerber, S., Hafter, E.R.: A system to simulate and reproduce audio-visual environments for spatial hearing research, *Hear. Res.* 260 (2010), 1-10.
- [17] Seeber, B.U.: A New Method for Localization Studies, *Acta Acustica United with Acustica* 88 (2002), 446-450.
- [18] Getzmann, S. & Lewald, J.: Localization of moving sound, *Perception & Psychophysica* 69 (2007), 1022-1034.