

Einfluss von visueller Stimulation auf Distanzwahrnehmung und Externalisierung

Felix Dollack¹, Christina Imbery¹, Steven van de Par², Jörg Bitzer¹

¹ Jade Hochschule, Institut für Hörtechnik und Audiologie, 26121 Oldenburg

² Carl-von-Ossietzky Universität, Akustik, 26129 Oldenburg

Einleitung

Bei der Darbietung von Kopfhörersignalen kann es wünschenswert sein, eine möglichst realistische Hörsituation eines Schallereignisses zu simulieren. Neben einer bestimmten Einfallsrichtung und Distanz sollte das Schallereignis dabei auch außerhalb des Kopfes wahrgenommen werden. Es ist jedoch nicht immer möglich bei der Generierung der Kopfhörersignale die dafür notwendigen individuellen Außenohrübertragungsfunktionen (HRTF = head related transfer function) zu berücksichtigen und das Kopfhörersignal individuell zu entzerren. Hörereignisse außerhalb des Kopfes wahrzunehmen ist auch mit nicht-individuellen HRTFs möglich, allerdings fällt der Grad der Externalisierung bei nicht-individuellen HRTFs häufig geringer aus [4]. In verschiedenen Experimenten wurde eine Steigerung des Grads der Externalisierung durch zusätzliche visuelle Stimulation nachgewiesen [2, 3]. Im Kontext künstlicher visueller Umgebungen ist dieser Effekt bislang noch nicht untersucht worden. In einem psychoakustischen Experiment soll daher der Einfluss visueller Stimuli in Form virtueller Umgebungen auf den Grad der Externalisierung untersucht werden. Da die wahrgenommene Entfernung einer Schallquelle ebenfalls eine Rolle bei der Externalisierung spielt, wird in einem weiteren Experiment der Einfluss visueller Stimuli auf die Distanzwahrnehmung untersucht.

Methode

Um den Einfluss von visueller Stimulation auf die subjektiv wahrgenommene Distanz und den Grad der Externalisierung zu untersuchen, werden den Probanden neben der Schalldarbietung visuelle Stimuli dargeboten.

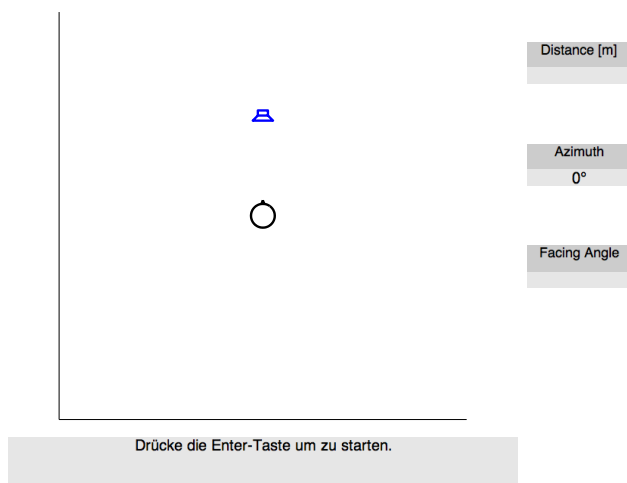


Abbildung 1: Grafische Oberfläche der Referenzkondition GUI.

Die visuelle Stimulation erfolgt auf einem 3D-Monitor mit deaktivierter 3D-Funktion in 2D und mit aktivierter 3D-Funktion in 3D, sowie mit einer Virtual Reality Brille VR. Zusätzlich wird eine Referenzkondition mit der grafischen Benutzeroberfläche GUI aus Abbildung 1 anstelle der visuellen Umgebung durchgeführt. Die Darbietungsreihenfolge der Referenzkondition und der visuellen Konditionen erfolgte randomisiert. Die Probanden wurden angehalten den Kopf während des Experiments so ruhig wie möglich zu halten. Die Darbietung der akustischen Stimuli erfolgen über einen Kopfhörer (HD-800 von Sennheiser).

Versuchsaufbau

Die Experimente wurden in einem kleinen Büro durchgeführt. Der Messaufbau besteht aus zwei Rechnern. Einer der Rechner ist für die Darstellung der visuellen Umgebungen auf der VR-Brille Oculus Rift und einem 3D-Monitor mit einer NVIDIA Shutter-Brille zuständig. Der andere Rechner steuert das Experiment mit Hilfe von Matlab. Das beinhaltet die Kommunikation mit der virtuellen Umgebung und die Darbietung der akustischen Signale. Ihre Antworten können die Probanden mit einer Tastatur und einem Drehregler geben. Abbildung 2 zeigt eine Skizze der wichtigsten Elemente des Aufbaus.

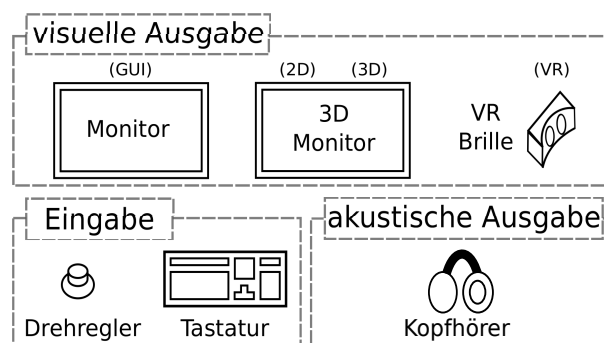


Abbildung 2: Skizze des Versuchsaufbaus. Für die Darbietung der visuellen Konditionen GUI, 2D, 3D und VR wurden verschiedene Ausgabegeräte verwendet. Die akustischen Stimuli wurden mit einem Kopfhörer wiedergegeben. Mit einem Drehregler und einer Tastatur konnten die Antworten eingegeben werden.

Probanden

An den Experimenten nahmen zwei Frauen und sechs Männer mit einem Durchschnittsalter von ca. 29 Jahren teil. Alle Probanden waren Erfahren mit psychoakustischen Tests und hatten nach eigenen Angaben keine Hörbeeinträchtigung.

Einfluss visueller Stimulation auf die subjektive Distanzwahrnehmung

Im ersten Experiment wurde die Distanzwahrnehmung der Probanden untersucht. Dafür wurde die subjektiv wahrgenommene Distanz für sechs Entfernungen zwischen 0.9 m bis 4.9 m erfragt.

Akustischer Stimulus

Die akustischen Stimuli wurden durch Binauralsynthese mit binauralen Raumimpulsantworten (BRIR) eines Kunstkopfes (KEMAR 45BB-3) für 6 Entfernungen generiert. Die BRIRs wurden mit der Audio Talkbox von NTi in einem Vorlesungsraum (Abmessung: 10x7.75x3m, $T_{60}=0.4s$) der Jade Hochschule Oldenburg aufgenommen. Der Aufbau der Messung entspricht der visuellen Umgebung aus Abbildung 3. Der Lautsprecher auf dem Podest und die Ohren des Kunstkopfes befanden sich in einer Höhe von 1.5 m. Als Eingangssignal der Synthese wurden trockene Sprachaufnahmen von 27 Sprechern aus einem Tonstudio verwendet, welche ca. 10 Sekunden der griechischen Fabel Nordwind und Sonne sprachen.

Visueller Stimulus

Als visueller Stimulus wurde der Raum in dem die Aufnahmen der BRIRs stattfanden als maßstabsgetreue virtuelle Umgebung modelliert. Der Kamerawinkel im Experiment beträgt 60 Grad, was dem Standard im Bereich virtueller Realität entspricht. Das entscheidende Element in dieser Umgebung ist ein Lautsprecher auf einem Podest. Abbildung 3 zeigt die visuelle Umgebung. Während dem Experiment wurde diese visuelle Umgebung auf den in Abbildung 2 aufgeführten visuellen Ausgabegeräten gezeigt. Der visuelle Stimulus der Referenzkondition ist die in Abbildung 1 dargestellte GUI. Zusätzlich wird die Entfernung zum Lautsprecher rechts in einem Zahlwert angegeben.



Abbildung 3: Modell des Raums zur Untersuchung des Einflusses eines visuellen Stimulus auf die subjektiv wahrgenommene Distanz.

Durchführung

Zu Beginn jeder visuellen Kondition werden dem Probanden Instruktionen für Steuerung und Ablauf des Experiments gezeigt. Danach wird ein kurzer Trainingslauf absolviert. Dabei wird den Probanden eine Auswahl der späteren Stimuli präsentiert, um sie mit den akustischen

Stimuli und der visuellen Umgebung vertraut zu machen. Im Anschluss an das Training und vor dem Start des eigentlichen Experiments wird den Probanden Ihre Aufgabe erläutert, welche darin besteht, nach der Präsentation des akustischen Stimulus, die Distanz zum Lautsprecher im visuellen Stimulus in Schritten von 10 cm so anzupassen, dass sie der subjektiv akustisch wahrgenommenen Distanz entspricht. Die Probanden hatten die Möglichkeit den akustischen Stimulus beliebig oft zu wiederholen.

Ergebnisse

Durch den maßstabsgetreuen Aufbau des visuellen Stimulus konnten die von den Probanden subjektiv wahrgenommenen Distanzen direkt in Meterangaben umgewandelt werden. Die akustischen Stimuli wurden von den Probanden in allen visuellen Konditionen zwischen 0.5 und 1.5 mal wiederholt. Hierfür zeigen durch eine einfaktorielle Varianzanalyse mit Messwiederholungen bei einem Signifikanzniveau von 5% keine signifikanten Unterschiede ($F(3,188)=0.038$ mit $p=0.99$). Die von den Probanden wahrgenommenen Distanzen sind für alle Konditionen in Abbildung 4 als Boxplots dargestellt. An die Mittelwerte der sechs Distanzen in der Referenzkondition, ganz links in Abbildung 4, wird das Modell aus [1] zur Vorhersage der subjektiv wahrgenommenen Distanz \hat{d} angepasst. Die Verschiebung ϵ_d der Distanzwahrnehmung wird durch die Minimierung der Abweichungen zwischen den Ergebnissen der visuellen Konditionen und dem an die Referenzkondition angepassten Modell berechnet. Die in verschiedenen Experimenten [1, 2, 5] festgestellte Tendenz die Distanz nahe gelegener Schallquellen zu überschätzen und die von entfernten Schallquellen zu unterschätzen, kann in den Ergebnissen aller Konditionen erkannt werden. Dabei zeigt sich eine systematische Überschätzung der Distanz von im Mittel 0.5 m durch den visuellen Stimulus. Zudem weisen die Ergebnisse der visuellen Konditionen 2D ($\epsilon_d = 0.41m$) und 3D ($\epsilon_d = 0.76m$) trotz identischer Kameraeinstellungen verschiedene Abweichungen von der Referenzkondition auf.

Diskussion

Die Ergebnisse der Referenzkondition entsprechen der Literatur, somit scheinen die akustischen Stimuli der erwarteten Distanzwahrnehmung zu entsprechen. In den visuellen Konditionen wird die Wahrnehmung der Distanz von den Probanden hingegen unterschiedlich bewertet. Die akustischen Stimuli können hierbei als Ursache für die veränderte Wahrnehmung vernachlässigt werden, da sie für alle visuellen Konditionen identisch waren. Es scheint, dass die hier dargebotenen visuellen Stimuli die Wahrnehmung der Distanz beeinflussen haben.

Das Modell aus [1], welches zur Auswertung herangezogen wurde, stellt nicht die tatsächlich wahrgenommene Distanz dar. Aus diesem Grund kann auch keine qualitative Aussage über den Wahrnehmungsversatz zwischen den visuellen Konditionen und der Referenzkondition getroffen werden. Der fast doppelt so große Versatz der Wahrnehmung in der Kondition 3D gegenüber der Kondition 2D ohne Shutter-Brille erscheint trotz gleicher vi-

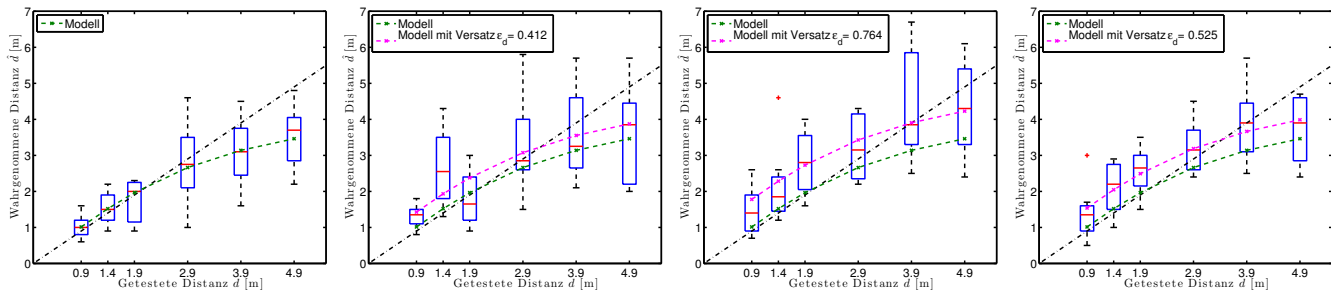


Abbildung 4: Ergebnisse der visuellen Konditionen (von links) GUI, 2D, 3D und VR.

sueller Einstellungen auf den ersten Blick nicht nachvollziehbar. Allerdings wird hier der visuelle Stimulus durch einen proprietären Algorithmus des Grafikkartenherstellers NVIDIA für die stereoskopische Darstellung mit der Shutter-Brille aufbereitet. Der Algorithmus erzeugt einen horizontalen Versatz zwischen den visuellen Stimuli um den Abstand zwischen den Augen zu simulieren. In der Kondition VR geschieht dies durch 2 Kameraobjekte die, wie unsere Augen, einen Abstand zueinander haben. Dieser Abstand bestimmt wie stark sich die visuellen Darstellungen der beiden Augen überlappen und auch in welcher Entfernung der Fokus des Betrachters liegt. Visuelle Objekte die in größerer Entfernung zum Betrachter liegen, befinden sich etwa in der Mitte des Sichtfelds beider Augen. Um hingegen nahe gelegenen Objekten zu fokussieren werden die Augen zur Nase hin gedreht. Dadurch befinden sich die Objekte zwar in der Mitte des wahrgenommenen Sichtfelds, allerdings an den inneren Rändern des Sichtfelds der einzelnen Augen. Bei der Darstellung mit dem 3D-Monitor und der VR-Brille bleibt das Objekt hingegen in der Mitte des Sichtfelds. Es ist daher denkbar, dass die Abweichungen der Konditionen 3D und VR gegenüber der Kondition 2D durch die falsche Augenstellung bzw. Darstellung des visuellen Stimulus erzeugt wird.

Die Unterschiede zwischen der Kondition 3D und VR könnten durch unterschiedliche Abstände zwischen den Augen zu erklären sein. Der Abstand zwischen den Kameras in der Kondition VR betrug etwa 6 cm wohingegen der Abstand in der Kondition 3D über einen Regler am Infrarot-Empfänger der Shutter-Brille auf den Standardwert der Software von 15% justiert wurde.

Einfluss der visuellen Stimulation auf die subjektiv wahrgenommene Externalisierung

Im zweiten Experiment wird untersucht ob audio-visuelle Stimulation den Grad der empfundenen Externalisierung erhöhen kann. Dafür wird der Grad der Externalisierung für 24 verschiedene Schalleinfallswinkel in einer Entfernung von 1.4 m von den Probanden erfragt und auf einer ganzzahligen Skala von 0 bis 5 angegeben. Dabei entspricht ein Wert von 0 keiner Externalisierung und ein Wert von 5 Externalisierung an der Position des Lautsprechers. Die Vorgehensweise ist angelehnt an ein vergleichbares Experiment von [4], jedoch ohne visuelle Stimulation.

Akustischer Stimulus

Die akustischen Stimuli wurden durch Binauralsynthese mit den selben BRIRs wie im Experiment zur Distanzwahrnehmung generiert. Anstelle der Sprachaufnahmen wurde in diesem Experiment eine Rauschsequenz mit gepulstem weißen Rauschen verwendet. Die Pulsdauer betrug 250 ms mit einer Pausendauer von 100 ms zwischen den Pulsen und bei einer Abtastfrequenz von 44.1 kHz.

Visueller Stimulus

Die visuelle Umgebung für das Experiment zur Bestimmung des Grads der Externalisierung besteht aus einem Ring mit 24 Lautsprechern, welche auf Podesten stehen. Neben einigen farbigen Markierungen zur Orientierung befinden sich keine weiteren Gegenstände in der virtuellen Umgebung. Die verwendete visuelle Umgebung ist in Abbildung 5 dargestellt. Der Proband befindet sich während des Experiments im Zentrum des Aufbaus und hat den Blick in Richtung des rot angestrahnten Lautsprechers gerichtet. In der Referenzkondition wird die GUI aus Abbildung 1 verwendet und der Lautsprecher um den Kopf im Zentrum rotiert. Zusätzlich wird die Einfallsrichtung rechts in einem Zahlwert angegeben.

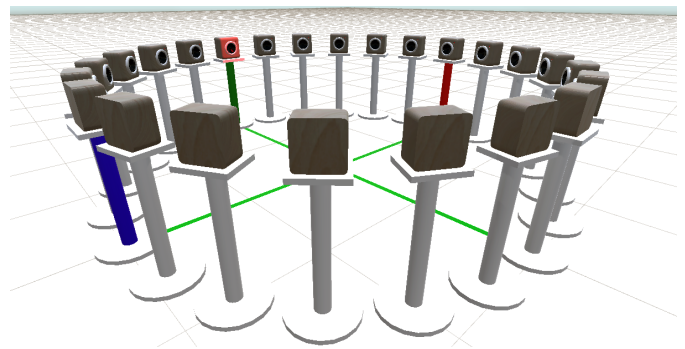


Abbildung 5: Visuelle Umgebungen mit dem Ring aus 24 Lautsprechern. Der Proband befindet sich im Zentrum des Aufbaus mit Blick in Richtung des rot angestrahnten Lautsprechers auf dem grünen Podest.

Durchführung

Zu Beginn jeder visuellen Kondition werden dem Probanden die Instruktionen für Steuerung und Ablauf des Experiments gezeigt. Danach wird ein kurzer Trainingslauf absolviert. Dabei werden den Probanden im Wechsel ein nicht Externalisiertes Monosignal ohne HRTF und ein Externalisiertes binauralsynthetisiertes Signal von einer Auswahl der Einfallsrichtungen präsentiert, um sie mit

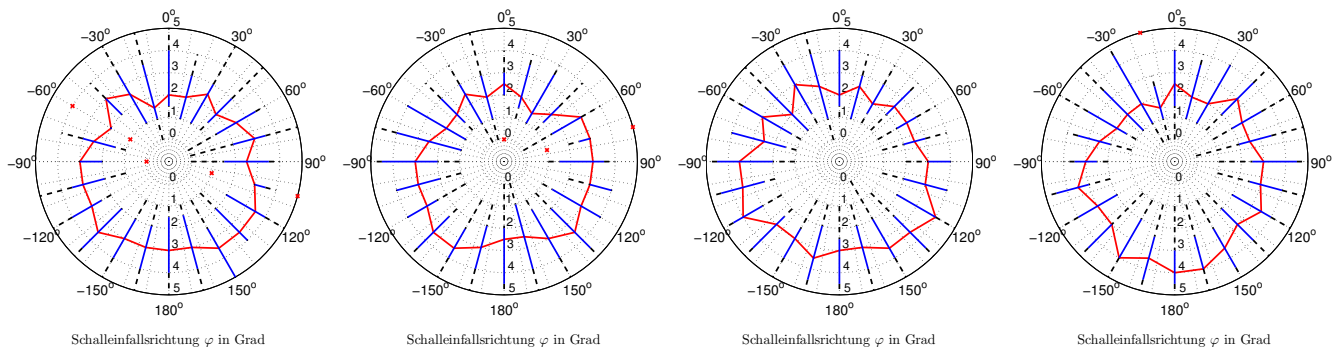


Abbildung 6: Ergebnisse der visuellen Konditionen (von links) GUI, 2D, 3D und VR als Funktion des Schalleinfallswinkels φ .

den akustischen Stimuli, der visuellen Umgebung sowie dem Konzept der Externalisierung vertraut zu machen. Im Anschluss an das Training und vor dem Start des eigentlichen Experiments wird den Probanden Ihre Aufgabe erläutert. Vor jeder Wiedergabe des akustischen Stimulus war es den Probanden erlaubt sich in der virtuellen Umgebung umzusehen, um die Einfallrichtung, welche durch ein rotes Licht angezeigt wurde, zu bestimmen.

Ergebnisse

Mit den Antworten der Probanden in Abhängigkeit der Schalleinfallrichtung wurden die Parameter für Boxplots bestimmt. Diese sind in Abbildung 6 als Funktion des Einfallswinkels φ in einem Polardiagramm dargestellt. Der Grad der Externalisierung liegt für alle Konditionen in der hinteren Hemisphäre im Median bei 3. In der vorderen Hemisphäre, besonders um die Einfallrichtung von 0 Grad, fällt der Grad der Externalisierung auf einen schlechteren Wert von ca. 2. Für diese Einfallrichtung werden die akustischen Stimuli in den visuellen Konditionen GUI und 2D mehr als doppelt so häufig wie andere Einfallrichtungen wiederholt. Gegenüber den visuellen Konditionen 3D und VR werden die akustischen Stimuli der Konditionen GUI und 2D für die 0 Grad Richtung bis zu dreimal so häufig wiederholt. Eine einfaktorielle Varianzanalyse mit Messwiederholungen bei einem Signifikanzniveau von 5% der Stimuli-Wiederholungen zeigt, dass dieser Unterschied signifikant ($F(23,168)=1.96$ mit $p<0.01$) ist.

Diskussion

In den Ergebnissen lässt sich durch die zusätzliche visuelle Darbietung kaum ein Einfluss auf den Grad der Externalisierung erkennen. Die häufige Wiederholung der Stimuli aus der 0 Grad Richtung lässt darauf schließen, dass es den Probanden dort schwerer gefallen ist ein Urteil zu fällen. In den Konditionen 3D und VR musste die 0 Grad Richtung weniger häufig wiederholt werden. Eine mögliche Erklärung hierfür wäre, dass diese beiden Konditionen zusätzliche Information zur räumlichen Tiefe durch Veränderung des visuellen Stimulus bei Kopfbewegungen liefern. Diese waren den Probanden zwar untersagt, können aber ohne Fixation nicht ganz verhindert werden. Dies scheint den Probanden so viel zusätzliche Informationen zu liefern, dass sie den Grad der Externalisierung ohne vermehrte Wiederholung des akustischen

Stimulus angeben können. Eine andere mögliche Ursache ist das Fehlen einer räumlichen Begrenzung. Die verwendeten HRTFs waren die selben wie im Experiment zur Distanzwahrnehmung. Sie beinhalteten Reflektionen der Raumwände, welche in diesem Experiment nicht zu sehen waren. Der Grund hierfür liegt in der Aufnahme der HRTFs bei denen der Kunstkopf rotiert wurde. Mit Darstellung des Raums hätte so die Ausrichtung des Probanden während des Experiments verändert werden müssen, um einen anderen Schalleinfallswinkel zu simulieren.

Zusammenfassung

Die visuelle Stimulation beeinflusst sowohl die Distanzwahrnehmung, was sich durch die systematische Verschiebung der Ergebnisse zeigt, als auch in geringem Maße die Externalisierung von akustischen Signalen. In Hinblick auf 3D und VR könnte die zur Zeit verwendete 3D Darbietung die visuelle Wahrnehmung stärker beeinflussen als vermutet. In weiteren Untersuchungen könnte dieser Einfluss der konventionellen 3D-Darbietung im Vergleich zu einer physiologisch motivierten Ausrichtung der Augen untersucht werden. Der Grad der Externalisierung scheint durch die visuelle Stimuli nur dann beeinflusst zu werden, wenn die Darstellung stereoskopisch ist. Dies könnte durch individuelle Binauralsynthese rein akustisch zu besseren Ergebnissen führen und so den tatsächlichen visuellen Einfluss auf Externalisierung genauer einzugrenzen.

Literatur

- [1] A. W. Bronkhorst und T. Houtgast. „Auditory distance perception in rooms“. In: *NATURE* 397 (Feb. 1999).
- [2] Esteban R Calcagno u. a. „The role of vision in auditory distance perception“. In: *Perception* 41 (2012).
- [3] J. Cubick u. a. „The influence of visual cues on auditory distance perception“. In: *Fortschritte der Akustik DAGA 2015*.
- [4] S.-M. Kim und W. Choi. „On the externalization of virtual sound images in headphone reproduction: A Wiener filter approach“. In: *JASA* 117 (Juni 2005).
- [5] Pavel Zahorik, Douglas S. Brungart und Adelbert W. Bronkhorst. „Auditory Distance Perception in Humans: A Summary of Past and Present Research“. In: *ACTA Acoustica United with America* (2005).