

Perzeptive Evaluation binauraler Auralisationen

Christoph Moldrzyk¹, Tobias Lentz², Stefan Weinzierl³

¹VISAURAL, D - 10713 Berlin, Email: aural@prz.tu-berlin.de

²Institut für Technische Akustik, RWTH Aachen, D - 52066 Aachen, Email: tle@akustik.rwth-aachen.de

³Fachgebiet Kommunikationswissenschaft, TU Berlin, D - 10623 Berlin, Email: stefan.weinzierl@tu-berlin.de

Einleitung

Ein entscheidendes Kriterium für die Leistungsfähigkeit binauraler Simulation ist die perzeptive Glaubwürdigkeit des simulierten Hörereignisses, im Idealfall die Nichtunterscheidbarkeit vom realen Hörereignis. Die Übereinstimmung von auf der Grundlage von Raumsimulationen berechneten akustischen Kriterien mit im realen Raum gemessenen Werten wurde in mehreren Studien überprüft ([1]), ebenso wurden Anforderungen an einzelne Parameter wie die Detailgenauigkeit von Raummodellen [2] und die Wiedergabelatenz [3] aufgrund von Hörversuchen formuliert. Eine umfassende Bewertung der perzeptiven Plausibilität binauraler Simulationen mit Methoden der Psychoakustik steht jedoch noch aus.

Eine längere Forschungstradition hat die Forschung zu Immersion und Präsenz in visuellen VR-Systemen mit einem breiten Spektrum objektiver und subjektiver Bewertungskriterien.[4] Ein direkter Paarvergleich zwischen Realität und Simulation ist im visuellen Bereich schwierig, da sich das visuelle Interface für die „dahinterliegende“ Realität kaum ausreichend transparent machen lässt. Im akustischen Bereich dagegen ist ein solcher Vergleich möglich, wenn der benutzte Kopfhörer einerseits ausreichend schalldurchlässig ist und andererseits die verbleibende Abschattung des äußeren Schallfelds bei der für die Simulation benutzten Außenohrübertragungsfunktion berücksichtigt ist. Für die vorliegende Untersuchung wurden daher im Rahmen eines Paarvergleichs die wahrnehmbaren Unterschiede zwischen einer realen Schallquelle (1), einer Faltung der im Raum gemessenen, binauralen Impulsantworten (2) sowie einer Faltung der im Raummodell berechneten binauralen Raumimpulsantworten (3) ermittelt. Ein akustisch besonders transparenter Kopfhörer (Stax Lambda Pro) blieb während des ganzen Versuchs aufgesetzt, entsprechend wurden mit aufgesetztem Kopfhörer vermessene Raumimpulsantworten verwendet.

System

Für den Hörversuch im Audimax der TU Berlin (Abb. 1) mit einer Nachhallzeit T_{60} von 2,5 s wurden Messdaten eines im Hals über dem Torso drehbaren Kunstkopfsystems [5] verwendet. Mit diesem Kunstkopf wurden mit einem auf der Bühne des Audimax stehenden Biegestrahler MBL 101 als Quelle am Platz des Hörers binaurale Raumimpulsantworten (BRIR) in einer Winkelauflösung von horizontal $\frac{1}{2}^\circ$ gemessen. Ebenso wurden insgesamt 52000 Übertragungsfunktionen des Kunstkopfes im $\frac{1}{2}^\circ$ -Abstand horizontal und 5° -Abstand vertikal im reflexionsfreien Raum des ITA Aachen gemessen und in das Simulationsprogramm EASE zur Auralisation simulierter BRIR implementiert. Das in EASE verwendete Raummodell des Audimax besteht aus 520 Flächen und 33 verschiedenen Wandmaterialien. Ein an dem ITA Aachen entwickeltes Faltungsprogramm mit geringer Latenz erlaubt eine Echtzeitfaltung in Verbindung mit einem Ultraschall-Headtracker. Horizontale Kopfbewegungen der Versuchspersonen wurden in einem Bereich von $\pm 75^\circ$ nachgeführt.

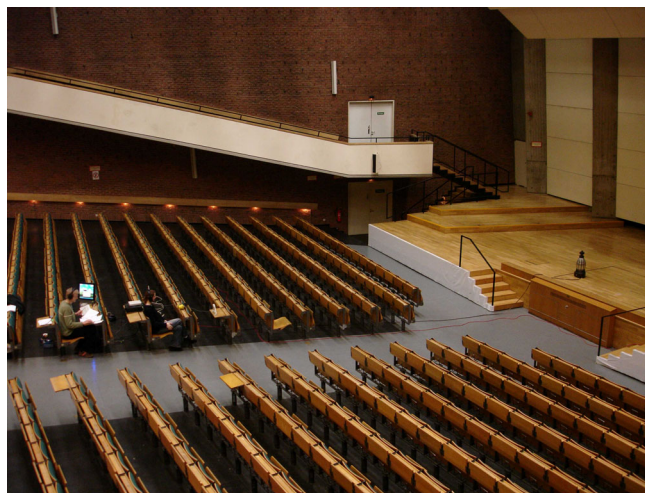


Abbildung 1: Hörversuch im Audimax der TU Berlin

Hörversuch

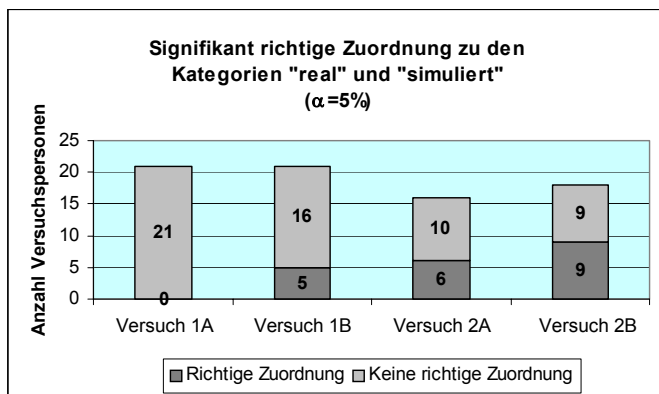
Zur subjektiven Evaluation der perzeptiven Plausibilität wurde ein zweiteiliger Hörtest durchgeführt. Im ersten Teil (Versuch 1) wurden den Probanden 20 Signalpaare in zufälliger Reihenfolge dargeboten, bestehend aus dem Signal der realen Schallquelle (Lautsprecher auf der Bühne) und einer Faltung der gemessenen, binauralen Raumimpulsantworten mit dem nachhallfreien Quellsignal. Der Kopfhörer blieb während des ganzen Versuchs aufgesetzt. Im zweiten Teil (Versuch 2) wurden den Probanden 20 Signalpaare dargeboten, bestehend aus der Faltung der (diesmal ohne aufgesetzten Kopfhörer) gemessenen, binauralen Raumimpulsantworten und der Faltung der im Modell berechneten binauralen Raumimpulsantworten. Für beide Versuche wurde ein Sprachsignal verwendet (männlicher Sprecher, englischsprachig, Archimedes CD track 5, Ausschnitt von 15 s).

Um den Einfluss visueller Informationen sowie einer beim Betreten des Raums stattfindenden Adaption an die Akustik des Raums auf die Unterscheidbarkeit zwischen realer und simulierter Schallquelle zu untersuchen, wurden beide Versuche zunächst mit verbundenen Augen durchgeführt; die Probanden wurden mit Gehörschützern in den Raum geführt (Teil A). Am darauffolgenden Tag wurde der Versuch mit denselben Probanden ohne Gehör- und Sichtschutz wiederholt (Teil B).

Insgesamt standen 39 Studenten der Kommunikationswissenschaft im Hauptstudium als Testpersonen zur Verfügung (13 weiblich, 26 männlich). Protokolliert wurde die durch Handzeichen signalisierte Zuordnung des Signals zu den Kategorien „real“ oder „simuliert“. Bei Versuch 2 wurde die Kategorie „real“ der gemessenen, „simuliert“ der im Modell berechneten Impulsantwort zugeordnet. In einem Fragebogen wurde außerdem auf einer dreistufigen Skala die Vertrautheit mit dem Testraum abgefragt, die Hörerfahrung mit binauralen Simulationen sowie die Bewertung der wahrgenommenen Signalunterschiede in den Kategorien Klangfarbe, Hörereignisrichtung und -entfernung sowie Halligkeit ($u=0$: kein Unterschied, 1: geringer Unterschied, 2: deutlicher Unterschied).

Ergebnisse

Auf einem Signifikanzniveau von 5% wurde die Hypothese getestet, dass der Unterschied zwischen realer und simulierter Schallquelle von den Versuchspersonen erkannt und richtig zugeordnet werden kann. Auf der Grundlage einer Binomialverteilung musste die Nullhypothese (Unterschied nicht erkennbar, $p=0.5$) verworfen werden, wenn mindestens 15 von 20 Signalpaaren richtig zugeordnet wurden.



Versuch	Signal 1 („real“)	Signal 2 („simuliert“)
1 A	blind, ohne Adaption	Lautsprecher gemessene BRIR
1 B	sehend, mit Adaption	
2 A	blind, ohne Adaption	gemessene BRIR im Modell berechnete BRIR
2 B	sehend, mit Adaption	

Bei Versuch 1 war mit Sichtschutz und ohne akustische Adaption an den Testraum keine der 21 VP in der Lage, das simulierte Signal zuverlässig richtig zuzuordnen. Bei der Wiederholung ohne Sichtschutz und nach vorangehender Adaption an den Raum gelang dies 5 von 21 VP. Bei Versuch 2 waren mit Sichtschutz und ohne akustische Adaption an den Testraum 6 von 16 VP in der Lage, das simulierte Signal richtig zuzuordnen. Bei der Wiederholung ohne Sichtschutz und nach vorangehender Adaption gelang dies 9 von 18 VP.

		Geschlecht	Vertrautheit mit Audimax	Hörerfahrung virtuelle Akustik	Richtige Antworten Versuch A (%)	Richtige Antworten Versuch B (%)
Geschlecht	Pearson Correlation	1,000	,058	,309	,408**	,371*
	Sig. (2-tailed)	.	,737	,067	,010	,020
	N	39	36	36	39	39
Vertrautheit mit Audimax	Pearson Correlation	,058	1,000	-,058	,085	,079
	Sig. (2-tailed)	,737	.	,737	,622	,647
	N	36	36	36	36	36
Hörerfahrung virtuelle Akustik	Pearson Correlation	,309	-,058	1,000	,069	,246
	Sig. (2-tailed)	,067	,737	.	,688	,148
	N	36	36	36	36	36
Richtige Antworten Versuch A (%)	Pearson Correlation	,408**	,085	,069	1,000	,391*
	Sig. (2-tailed)	,010	,622	,688	.	,014
	N	39	36	36	39	39
Richtige Antworten Versuch B (%)	Pearson Correlation	,371*	,079	,246	,391*	1,000
	Sig. (2-tailed)	,020	,647	,148	,014	.
	N	39	36	36	39	39

Zusammenhang zwischen Vertrautheit mit dem Testraum, Hörerfahrung mit akustischen Simulationen, Geschlecht der VP mit dem Prozentsatz korrekter Zuordnungen

Bei der Bewertung der verschiedenen Kategorien ordneten die VP die größten wahrgenommenen Unterschiede den Kategorien Klangfarbe ($\bar{u}=1,11$) und Nachhall ($\bar{u}=1,01$) zu, kleinere Unterschiede wurden bei Richtung ($\bar{u}=0,78$) und Entfernung ($\bar{u}=0,40$) der Schallquelle wahrgenommen. Keine dieser Bewertungen korreliert jedoch signifikant mit der korrekten Zuordnung der Signale. Ebenso wenig zeigte sich eine Korrelation zwischen dem Prozentsatz korrekter Zuordnungen mit der im

Fragebogen angegebenen Vertrautheit mit der Akustik des Testraums aufgrund früherer Besuche, mit der im Fragebogen angegebenen Hörerfahrung mit binauralen Wiedergabesystemen (Tab.1) oder mit dem Alter der VP. Ein schwacher aber signifikanter Zusammenhang bestand jedoch mit dem Geschlecht der VP ($r=0,41$ für Versuch A, $r=0,37$ für Versuch B: Der Unterschied zwischen Realität und Simulation wurde von männlichen VP signifikant häufiger erkannt als von weiblichen VP.

Diskussion

Bei der Evaluation eines kopfbezogenen, akustischen Simulationssystems erwies sich die perzeptive Plausibilität bei der rein akustischen Wiedergabe mit gemessenen binauralen Raumimpulsantworten als so hoch, dass keine Versuchsperson in der Lage war, den Unterschied zwischen realer Schallquelle und Simulation zuverlässig zu erkennen. Die Unterschiede zwischen der Simulation mit gemessenen und in einem Raummodell berechneten Impulsantworten waren dagegen so groß, dass etwa die Hälfte der VP in der Lage waren, die im Modell berechnete Impulsantwort zuverlässig als „simuliert“, d.h. als nicht real zu erkennen. Wenn die VP die Möglichkeit hatten, die Simulation mit einem unmittelbar vor dem Hörversuch ausgebildeten akustischen und einem während des Versuchs vorliegenden visuellen Referenzmuster zu vergleichen, wurde die Zuordnung deutlich erleichtert. Beide Faktoren erhöhen offensichtlich die Sensibilität für die Plausibilität der akustischen Wiedergabe, eine isolierte Untersuchung der beiden Faktoren bei zukünftigen Versuchen scheint daher angezeigt. Die aufgrund weiter zurückliegender Besuche gespeicherte Erinnerung an die akustischen Eigenschaften des Raums scheint dagegen weniger bedeutend zu sein, ebenso eine spezifische Hörerfahrung mit akustischen Simulationen. Ob sich die im vorliegenden Versuch festgestellte höhere Sensibilität von Männern für die Plausibilität von Raumsimulationen auch bei Versuchen mit einer höheren Anzahl an Testpersonen bestätigt, bleibt abzuwarten.

Der Versuch zeigt, dass auf dem gegenwärtigen Stand der Technologie zur Raumsimulation, d. h. mit hochaufgelösten binauralen Impulsantworten und Faltung geringer Latenz, selbst für „kritische“ Räume mit langer Nachhallzeit ein Signal generiert werden kann, das eine überwiegende Anzahl von VP nicht mehr zuverlässig als simuliert, d.h. als nicht natürlichen Ursprungs erkennt. Als entscheidender Faktor für die Unterscheidbarkeit von Realität und Simulation hat sich jedoch das Vorhandensein einer akustischen und visuellen Referenz erwiesen. Die Ermittlung der für die unterschiedliche Diskriminierungsfähigkeit maßgeblichen, interpersonellen Unterschiede wird Gegenstand zukünftiger Hörversuche sein. Hierfür hat sich die Methode des direkten Paarvergleichs mit der Realschallquelle als taugliches Instrument erwiesen.

[1] Bork, I.: *Simulation and Measurement of Auditorium Acoustics - The Round Robins on Room Acoustical Computer Simulation*, Proceedings of the Institute of Acoustics, Volume 24, Part 4, 2002

[2] Pompetzki, Wulf: *Psychoakustische Verifikation von Computermodellen zur binauralen Raumsimulation*, Dissertation, RU Bochum 1993

[3] Mackensen, Philip: *Auditive Lokalisation. Kopfbewegungen als zusätzliches Lokalisationsmerkmal*, Dissertation, TU Berlin 2004

[4] IJsselsteijn, W.A., de Ridder, H., Freeman, J. & Avons, S.E.: *Presence: Concept, determinants and measurement*, Proceedings of the SPIE, Human Vision and Electronic Imaging V, 3959-76 (2000)

[5] Moldrzyk, C., Ahnert, W., Feistel, S., Lentz, T., Weinzierl, S.: *Head-Tracked Auralization of Acoustical Simulation*, 117th AES Convention, San Francisco 2004