

Hans-Joachim Maempel

---

## Die Trennung der Sinne

### Der Virtuelle Konzertsaal als Forschungswerkzeug und Vermittlungsinstrument

Sowohl zur Alltagswelt als auch zu vielen Künsten, etwa der Musik, dem Theater, dem Film oder der multimedialen Installation, ist der wahrnehmende Zugang im wesentlichen ein audiovisueller. So trivial diese Feststellung ist, so grundsätzlich sind die sich aus ihr ergebenden Fragen und so schwierig ist deren Beantwortung mit den Mitteln der empirischen Wissenschaften. Wie beeinflussen uns Stimme und Mimik einer Person? Wie tragen Bildgeschehen und Filmmusik zur Spannung im Film bei? Berührt uns eine musikalische Darbietung auch aufgrund des Gesehenen? Und welchen Anteil haben daran die Akustik und die optische Gestaltung des Aufführungsraums? Um derartige Fragen empirisch beantworten zu können, reicht es nicht aus, Wahrnehmungsexperimente mithilfe von realen Menschen, Filmen oder Konzerten durchzuführen. Vielmehr müssen Testreize zum Einsatz kommen, die wir in unserer natürlichen Umgebung nicht vorfinden. Mit dem Virtuellen Konzertsaal wurde ein Simulationssystem geschaffen, das sowohl die Trennung der Einflussfaktoren Darbietung und Raum als auch die Trennung ihrer optischen und akustischen Eigenschaften erlaubt, und zwar unter weitgehender Aufrechterhaltung derjenigen Reizeigenschaften, die für das Hören und Sehen von Bedeutung sind. Obwohl seine Entwicklung ausschließlich an methodologischen Kriterien der empirischen Forschung orientiert war, verspricht auch seine öffentliche Ausstellung im institutseigenen Musikinstrumenten-Museum gleich in mehrfacher Hinsicht einen Gewinn.

## Problemstellung

Die akustische Forschung hat dem Zusammenspiel der auditiven und visuellen Modalität lange Zeit wenig Beachtung geschenkt. Speziell die Raumakustik ist zudem traditionell eher physikalisch als psychologisch ausgerichtet. Daher wissen wir wenig sowohl über die auditiven und kognitiven Wirkungen raumakustischer Eigenschaften als auch über das Verhältnis von Hören und Sehen bei der Wahrnehmung von Räumen, die ja nicht nur für künstlerische Darbietungen von großer Bedeutung sind. Im Rahmen der DFG-Forschergruppe »Simulation and Evaluation of Acoustical Environments« (SEACEN) und in Kooperation mit dem Fachgebiet Audiokommunikation der Technischen Universität Berlin soll das Projekt »Audio-visual perception of acoustical environments« diese Forschungslücke mithilfe der virtuellen Akustik und unter Anwendung experimenteller Methoden schließen.

Dass sich die geistige Repräsentation der physikalischen Welt wenigstens sowohl auf die auditive als auch auf die visuelle Modalität stützt, wurde für viele Empfindungs-, Wahrnehmungs- und Verarbeitungsvorgänge gezeigt,<sup>1</sup> zum Beispiel die Reizintensitätsbestimmung<sup>2</sup>, die Reizlokalisation<sup>3</sup>, das Zeit- und Synchronitätsempfinden<sup>4</sup>, die Spracherkennung<sup>5</sup>, die Qualitätsbewertung<sup>6</sup> und die Bedeutungsgebung<sup>7</sup>. Dabei wurden für bestimmte Reizkonfigurationen und Wahrnehmungsmerkmale auch crossmodale Effekte nachgewiesen. So klingt etwa ein roter ICE lauter als ein grüner.<sup>8</sup> Ferner wird häufig von einer audiovisuellen Interaktion ausgegangen, wobei die Bedeutung des Begriffs in der Literatur nicht immer klar ist. Crossmodale Effekte und Interaktionseffekte müssen insofern auch für die Wahrnehmung von und in Räumen angenommen werden.

Eine Zusammenschau bestehender Untersuchungen ergab, dass sich bis heute weder eine Theorie noch eine Forschungsstrategie zur audiovisuellen Raumwahrnehmung formiert hat.<sup>9</sup> Vielmehr wurden zum Teil sehr spezifische Merkmale untersucht und verschiedene Paradigmen und Methoden angewandt. Mit Blick auf das geringe Vorwissen in dem Forschungsgebiet sind grundsätzliche Fragen sinnvollerweise zuerst und spezifischere nachfolgend zu untersuchen. Die wichtigsten grundsätzlichen Fragen sind, in welchem Maße Hören und Sehen eigentlich zur Raumwahrnehmung beitragen, etwa zu geometrischen, materialen und ästhetischen Eindrücken, und ob beide Modalitäten interagieren.

Methodische Vorüberlegungen ergaben die Notwendigkeit einer faktischen und terminologischen Unterscheidung zwischen physischen (›akustischen‹ und ›optischen‹) und geistigen (›auditiven‹ und ›visuellen‹) Merkmalen wie auch



1 Für die Simulation ausgewählte Aufführungsräume (v.l.o.n.r.u.): Renaissance-Theater (Berlin), Komische Oper (Berlin), Gewandhaus (Leipzig), Kleiner Saal des Konzerthauses (Berlin), Jesus-Christus-Kirche (Berlin), Basilika des Klosters Eberbach (Eltville am Rhein)

einer Unterscheidung zwischen modalitätsspezifischen und modalitätsunspezifischen Merkmalen. Die wahrgenommene Lautstärke ist zum Beispiel nicht geeignet, visuelle Eindrücke direkt und nicht-metaphorisch zu beschreiben, während der Eindruck von Raumgröße oder das Gefallensurteil sowohl auf auditive als auch auf visuelle Eindrücke angewandt werden können und daher modalitätsunspezifische Merkmale darstellen, die auf beide Sinnessysteme zugreifen.

Weiterhin musste es im Experiment möglich sein, optische und akustische Raumeigenschaften unabhängig voneinander zu variieren, so dass Konfliktreize entstehen. Ein Konfliktreiz, der für die Untersuchung der audiovisuell basierten Lokalisation geeignet wäre, ist zum Beispiel die gleichzeitige Präsentation der optischen Komponente eines Sprechers an einer Position und der akustischen Komponente desselben Sprechers an einer anderen Position. Hier stellt sich das Problem, dass die optoakustischen Konfliktreize nicht mit realen Sprechern herstellbar sind. Vielmehr müssen diese simuliert werden, und das gilt auch im vorliegenden Falle für Räume als Testreize. Natürlich ist experimentell zu überprüfen, inwieweit die entsprechenden empirischen Befunde von der Simulation auf die Realität übertragen werden können.

Damit die Variationsbreite der optischen und der akustischen Raumeigenschaften gleich groß ist, müssen sie denselben Räumen entstammen und auf diese Weise naturgemäß zueinander passen. Die akustischen und optischen Raumeigenschaften werden also jeweils in ihrer Gesamtheit variiert, indem einfach mehrere, in sich stimmige Räume simuliert werden. Für die Simulation



2 Lautsprecher sowie *head and torso simulator* in der Basilika des Klosters Eberbach

bedeutet dies, dass sie reale Räume anstatt im Computer konstruierte Räume abbilden muss, und dies in höchstmöglicher technischer Qualität und in 3D, damit die optischen und akustischen Informationen möglichst vollständig und unverfälscht übertragen werden.

### Entwicklung des Virtuellen Konzertsaals

Die Entwicklung des Virtuellen Konzertsaals war ausschließlich an den dargestellten Forschungsfragen und methodologischen Kriterien orientiert. Er erlaubt die wechselseitig unabhängige Variation nicht nur von optischer und akustischer Komponente, sondern auch von Inhalt (Darbietung) und Übertragungssystem (Raum). In der ersten Ausbaustufe können sechs Aufführungsräume verschiedener Volumina (Spalten) und mittlerer akustischer Absorptionsgrade (Zeilen) dargestellt werden. (*Abb. 1*)

Technische Schritte der Realisation des Virtuellen Konzertsaals waren die Akquisition der Raumeigenschaften, die Aufnahme von künstlerischen Darbie-



3 Schauspielerin im reflexionsarmen Raum

tungen, deren Einbettung in die Räume, der Aufbau eines Wiedergabesystems (virtuelle Umgebung) und die Programmierung der Versuchsablaufsteuerung und -datenerfassung.<sup>10</sup>

Die Akquisition der akustischen Raumeigenschaften erfolgte durch das Abspielen von Messtönen über Lautsprecher, die anstelle der Musiker eines Streichquartetts beziehungsweise anstelle einer Rezitatorin auf der jeweiligen Bühne aufgestellt waren. An einem prominenten Sitzplatz wurden die binauralen Raumimpulsantworten für verschiedene Kopforientierungen über die Im-Ohr-Mikrofone eines Kopf- und -Torso-Simulators aufgenommen.<sup>11</sup> (Abb. 2) In der optischen Domäne wurden aus dem Blickwinkel jedes Auges stereoskopische Fotografien für 360 Kopforientierungen erstellt und von der zirkularen in die equirektangulare Projektionsgeometrie transformiert. Durch das Zusammensetzen der jeweils ein Grad breiten vertikalen Bildstreifen je Auge wurden dann stereoskopische Vollpanoramen erzeugt.

Zur Herstellung der Darbietungen spielten Berufsmusiker in einem reflexionsarmen Raum den zweiten Satz des Streichquartetts g-Moll op. 10 von Claude Debussy. Dabei wurden die Instrumente akustisch separiert und nach-

hallfrei als Mehrspuraufzeichnung aufgenommen. Um auch einen Wortbeitrag für die Experimente zur Verfügung zu haben, rezitierte außerdem eine professionelle Schauspielerin den ersten Absatz der Duineser Elegien von Rainer Maria Rilke. (Abb. 3) Die Darbietungen wurden dann in einem Green-Box-Studio im Vollplayback-Verfahren wiederholt und mit einer stereoskopischen Videokamera im Full-HD-Format aufgezeichnet. Dabei wurde besonderes Augenmerk auf optoakustische Synchronität und Konsistenz der Interpretation gelegt.

Eine besondere technische und zeitliche Herausforderung stellte die Einbettung der Darbietungen in die Räume dar. Die akustischen Darbietungen und Raumakustiken werden erst im Moment der Wiedergabe in Echtzeit durch das Verfahren der dynamischen Binauralsynthese (ursprünglich als *Binaural Room Scanning* bezeichnet)<sup>12</sup> zusammengeführt. Dabei wird die nachhallfreie Audioaufnahme kontinuierlich mit binauralen Raumimpulsantworten gefaltet. Für die Faltung wird in jedem Moment diejenige binaurale Raumimpulsantwort verwendet, die während einer Kopforientierung des *head and torso simulators* im realen Raum aufgenommen wurde, welche der mittels *head tracking* in Echtzeit erfassten aktuellen Kopforientierung des Hörers entspricht. Wiedergegeben wird das resultierende Audiosignal über eine eigens für die Projekte der Forschergruppe angefertigte hochqualitative Kopfhörer-Verstärker-Kombination mit digital linearisierter Übertragungsfunktion.<sup>13</sup> Das angewandte Binauralsynthese-System erlaubt aufgrund weiterer technischer Verbesserungen eine hochplausible dreidimensionale Wiedergabe.<sup>14</sup> Schallquellen und Raumreflexionen werden aus allen Richtungen wahrgenommen und bewegen sich bei Kopfdrehung nicht mit. Die optischen Darbietungen wurden mittels *chroma key compositing* unter Hinzufügung von Schatten und Anwendung von Farbkorrekturen Frame für Frame in die Raumpanoramen montiert. (Abb. 4)

Die aus den Einzelbildern gerenderten stereoskopischen Halbpanorama-Videos werden dem Besucher des Virtuellen Konzertsals durch fünf senkrecht ausgerichtete Projektoren und unter Verwendung einer Warping- und Edge-Blending-Software auf eine halbzyklindrische Leinwand mit einer Höhe von knapp drei Metern und einem Durchmesser von fünf Metern projiziert (Abb. 5), so dass ein Sichtwinkel von mehr als 160 Grad und eine Winkelauflösung, die ungefähr der des menschlichen Auges entspricht, realisiert werden konnten. In einem zweiten Labor wird das Bild auf einem 85-Zoll-Ultra-HD-Flachbildmonitor zwar mit geringerem Sichtwinkel, aber höherer Winkelauflösung gezeigt. Für die verschiedenen Wahrnehmungsexperimente wird das jeweils geeignetere Wiedergabesystem gewählt, und beide Labore können parallel betrieben werden, um die geplante Dichte an Experimenten realisieren zu können.



4 Halbpanorama des Gewandhauses mit einmontiertem Streichquartett

Mit dieser Verbindung aus akquirierten Raumeigenschaften, produzierten Inhalten und Wiedergabesystem ist es erstmals möglich, identische künstlerische Darbietungen in verschiedenen optischen und akustischen Konzerträumen zu präsentieren, und zwar unter Bewahrung fast aller physikalischen Schlüsselmerkmale (*cues*), die für beide Modalitäten perceptiv relevant sind. Da sich bei Kopfdrehung die virtuelle akustische Umgebung aufgrund des interaktiven, dynamischen Auralisationsverfahrens nicht mitdreht, kann der Rezipient die virtuellen Räume auditiv und visuell aktiv erkunden. Etwa 50 Versuchspersonen je Inhaltskategorie (Musik, Sprache) wurden gebeten, jede der 36 möglichen Kombinationen der optischen und akustischen Raumkomponenten sowie alle unimodalen Reize (nur Ton oder Bild) hinsichtlich auditiver, visueller und audiovisueller Merkmale mithilfe eines elektronischen Fragebogens auf einem Tablet-Computer zu beurteilen. Abgefragt wurden unter anderem die Ausprägungen von Lautstärke, Halligkeit, klanglicher Umhüllung, Helligkeit, dominierendem Farbton, Raumgröße, Entfernung der Darbietung, eigener Präsenz im Aufführungsraum, Ton-Bild-Passung oder ästhetischen Merkmalen wie Gefallen. Diese Daten wurden statistisch ausgewertet.

### Erste Ergebnisse

Im Gegensatz zum verbreiteten Postulat einer audiovisuellen Interaktion zeigen erste Ergebnisse generell keine signifikanten oder praktisch bedeutsamen Interaktionseffekte zwischen der optischen und der akustischen Einflussgröße, viel-

mehr wurden reine Haupteffekte beobachtet. Das heißt, dass keine Kombinationen bestimmter Stufen des optischen und des akustischen Faktors gefunden wurden, die sich in anderer Weise auf die erhobenen Merkmale auswirken als die beiden Faktoren selbst. Darüber hinaus wurden keine signifikanten oder praktisch bedeutsamen crossmodalen Effekte beobachtet, also weder eine optische Beeinflussung auditiver Merkmale noch eine akustische Beeinflussung visueller Merkmale. Im Experiment basierte die auditive Wahrnehmung also schlicht auf der akustischen Information und die visuelle Wahrnehmung auf der optischen Information. Die Berücksichtigung einer Voruntersuchung<sup>15</sup> zur Distanz- und Raumgrößenwahrnehmung auf der Grundlage einer schlechteren optischen Raumsimulation mit geringerem Informationsangebot zeigt allerdings, dass die verfügbaren optischen und akustischen Informationen für die Raumwahrnehmung in sehr flexibler Weise genutzt werden. Unter diesen Bedingungen stützte sich beispielsweise die Distanzwahrnehmung überwiegend auf akustische Information. Sind fast alle optischen und akustischen Schlüsselmerkmale gleichermaßen verfügbar, wie im hier beschriebenen Virtuellen Konzertsaal, basieren Distanz- und Raumgrößenwahrnehmung jedenfalls erwartungsgemäß ganz überwiegend auf den optischen Informationen. Im Gegensatz dazu scheinen die ästhetischen Urteile jedoch vorwiegend auf der Auswertung der akustischen Informationen zu beruhen.

### Ausblick

Die zweite Projektphase ist der Entflechtung der derzeit noch komplexen Faktoren und der genaueren Betrachtung bestimmter Wahrnehmungsmerkmale gewidmet.<sup>16</sup> So wird die Wahrnehmung von Distanz, Raumgröße und Raumform im Detail untersucht. Weiterhin soll geklärt werden, ob Menschen in der Lage sind, optische und akustische Räume einander richtig zuzuordnen, und welche Rolle das Gedächtnis dabei spielt. Die Ermittlung des Einflusses einer Vereinfachung der optischen Reize wird von Nutzen für die Interpretation vorliegender experimenteller Untersuchungen sein, in denen die Testreize weniger oder unschärfere Schlüsselmerkmale lieferten oder reale Szenen abstrakter dargestellt wurden. Schließlich wird die Rolle der musikalischen und raumakustischen Expertise der wahrnehmenden Personen untersucht werden, um die Beschränkungen und Geltungsbedingungen einer Theorie zur audiovisuellen Raumwahrnehmung abschätzen zu können.

Mit Blick auf die populärwissenschaftliche Wissensvermittlung ist der Aufbau eines hinsichtlich der technischen Daten reduzierten und insofern kosten-



5 Versuchsperson vor der Projektion eines Streichquartetts im Virtuellen Konzertsaal

günstigen Demonstrationssystems geplant, das den Besucherinnen und Besuchern des Musikinstrumenten-Museums sowohl für begleitete Vorführungen in einem gesondertem Raum als auch als festes Selbstbedienungs-Terminal auf der Ausstellungsfläche zur Verfügung stehen kann. Ein solches Demonstrationssystem würde es erstens erlauben, immobile und daher bislang nicht ausstellbare Objekte, nämlich Räume, wenigstens als technische Replikation erlebbar zu machen. Unter Aufgabe des Kriteriums der Datenbasierung ließen sich künftig auch solche Räume einbeziehen, die nicht mehr oder noch nicht existieren. Der Ansatz einer historischen Rekonstruktion und Auralisation wurde, wengleich ohne optische Simulation, bereits für die verschiedenen baulichen Varianten des Leipziger Gewandhauses erfolgreich umgesetzt.<sup>17</sup> Zweitens trennt speziell der hier dargestellte Virtuelle Konzertsaal Einflussgrößen auf Reizebene, die im Alltag stets miteinander verbunden sind. Indem beispielsweise bei laufender und unveränderter künstlerischer Darbietung die Aufführungsräume umgeschaltet werden können, wird der empirische Befund der Bedeutsamkeit

des Aufführungsraums für bestimmte Wahrnehmungsmerkmale individuell erlebbar. In Verbindung mit einem entsprechenden Begleittext wird drittens die empirische Methodik selbst anschaulich, und zwar sowohl was die Simulationstechnologie (3D-Audio und -Video) als auch was die Bedingungsvariation (Konfliktreizparadigma) und die erhobenen Merkmale (uni- und multimodale) betrifft.

Am Staatlichen Institut für Musikforschung, das den Forschungsauftrag bereits im Namen trägt und dessen Musikinstrumenten-Museum in der Dauer Ausstellung ohnehin auch instrumentenbauliche Verfahren vermittelt, begegnet man der Frage, ob es zulässig ist, den Begriff des Ausstellungsobjekts um wissenschaftliche Befunde und Methoden zu erweitern, jedenfalls offen.

#### Anmerkungen

- 1 Hans-Joachim Maempel: Audiovisuelle Wahrnehmung, in: Helga de la Motte-Haber et al. (Hg.): Lexikon der Systematischen Musikwissenschaft, Laaber 2010, S. 49–53; Armin Kohlrausch/Steven van de Par: Audio-Visual Interaction in the Context of Multi-Media Applications, in: Communication Acoustics, Berlin et al., 2005, S. 109–138.
- 2 Joseph C. Stevens/Lawrence E. Marks: Cross-Modality Matching of Brightness and Loudness, in: *Proceedings of the National Academy of Sciences of the United States of America*, Bd. 54/1965, H. 2, S. 407–411.
- 3 Garth J. Thomas: Experimental Study of the Influence of Vision of Sound Localization, in: *Journal of Experimental Psychology*, Bd. 28/1941, H. 2, S. 163–177; Ian P. Howard/William B. Templeton: Human spatial orientation, London 1966, S. 353–359; Mark B. Gardner: Proximity Image effect in sound localization, in: *Journal of the Acoustical Society of America*, Bd. 43/1968, S. 163.
- 4 Ladan Shams/Yukiyasu Kamitani/Shinsuke Shimojo: Visual illusion induced by sound, in: *Cognitive Brain Research*, Bd. 14/2002, S. 147–152; Tobias S. Andersen/Kaisa Tippiana/Mikko Sams: Factors influencing audiovisual fission and fusion, in: *Cognitive Brain Research*, Bd. 21/2004, H. 3, S. 301–308; Berta Luise Heide/Hans-Joachim Maempel: Die Wahrnehmung audiovisueller Synchronität in elektronischen Medien, in: 26. Tonmeistertagung, Leipzig, November 2010, S. 525–537.
- 5 John MacDonald, Harry McGurk: Visual Influences on Speech Perception Process, in: *Perception and Psychophysics*, Bd. 24/1978, H. 3, S. 253–257.
- 6 John G. Beerends/Frank E. de Caluwe: The Influence of Video Quality on Perceived Audio Quality and Vice Versa, in: *Journal of the Audio Engineering Society*, Bd. 47/1999, H. 5, S. 355–362.
- 7 Claudia Bullerjahn/Markus Guldénring: An empirical investigation of effects of film music using qualitative content analysis, in: *Psychomusicology*, Bd. 13/1994, H. 1/2, S. 99–118; Annabel J. Cohen: How music influences the interpretation of film and video: approaches from experimental psychology, in: R. A. Kendall & R. W. H. Savage (Hg.): *Perspectives in systematic musicology*, Los Angeles 2005, S. 15–36.
- 8 Hugo Fastl: Audio-visual interactions in loudness evaluation, in: *International Conference on Acoustics*, Kyoto 2004, S. II-1161–II-1166.

- 9 Hans-Joachim Maempel: P9: Audio-visual perception of acoustical environments, in: Simulation and Evaluation of Acoustical Environments (SEACEN). Proposal for a DFG Research Unit, Berlin et al., S. 189–221 (DFG-Antrag MA 4343/1-1).
- 10 Hans-Joachim Maempel/Alexander Lindau: Opto-acoustic simulation of concert halls – a data-based approach, in: 27. Tonmeistertagung, Köln 2012, S. 293–309.
- 11 Alexander Lindau / Stefan Weinzierl: FABIAN – An instrument for software-based measurement of binaural room impulse responses in multiple degrees of freedom, in: 24. Tonmeistertagung, Leipzig 2006, 5 S.
- 12 Ulrich Horbach et al.: Design and applications of a data-based auralization system for surround sound, in: 106<sup>th</sup> AES Convention, München 1999, Preprint 4976.
- 13 Vera Erbes/Frank Schultz/Alexander Lindau/Stefan Weinzierl: An extraaural headphone system for optimized binaural reproduction, in: DAGA, 38. Jahrestagung für Akustik, Darmstadt 2012, S. 313–314.
- 14 Alexander Lindau/Stefan Weinzierl: Assessing the Plausibility of Virtual Acoustic Environments, in: Acta Acustica united with Acustica, Bd. 98/2012, H. 5, S. 804–810.
- 15 Hans-Joachim Maempel/Matthias Jentsch: Auditory and visual contribution to egocentric distance and room size perception, in: Building Acoustics, Bd. 20/2013, H. 4, S. 383–402.
- 16 Hans-Joachim Maempel: P9: Audio-visual perception of acoustical environments, in: SEACEN. Renewal Proposal for a DFG Research Unit, Berlin et al., S. 157–175 (DFG-Antrag MA 4343/1-2).
- 17 Stefan Weinzierl/Hagen Rosenheinrich/Johannes Blickensdorff/Michael Horn/Alexander Lindau: Die Akustik der Konzertsäle im Leipziger Gewandhaus. Geschichte, Rekonstruktion und Auralisation, in: DAGA, 36. Jahrestagung für Akustik, Berlin 2010, S. 1045–1046.