

Audiotechnik des aixCAVE Virtual Reality-Systems

Frank Wefers, Sönke Pelzer, Ramona Bomhardt,
Markus-Müller Trapet, Michael Vorländer
Institut für Technische Akustik, RWTH Aachen
Kopernikusstraße 5, 52074 Aachen
fwe@akustik.rwth-aachen.de

Einleitung

Im Jahr 2012 wurde an der RWTH Aachen das aixCAVE Virtual Reality System (Abbildung 1) installiert. Dieses System ist gegenwärtig eines der weltweit größten CAVE-Systeme zur Darstellung immersiver, virtueller Umgebungen. Im Zeitraum 2013-2014 wurde die aixCAVE durch das Institut für Technische Akustik um ein Audiosystem für die Echtzeit-Auralisierung in VR-Umgebungen erweitert. Die räumliche Klangwiedergabe geschieht mittels Binauraltechnik. Funkkopfhörer erlauben eine simultane Wiedergabe für mehrere Benutzer gleichzeitig. Die binaurale Wiedergabe kann für einen einzelnen Benutzer auch mittels zwölf Lautsprechern und dynamischer Übersprechkompensation geschehen. Dies hat den Vorteil, dass der Benutzer keine zusätzliche Hardware tragen muss. Für eine ausreichende Schallwiedergabe niedriger Frequenzen wurde ein symmetrisches Array von Subwoofern konzipiert, welches die Anregung mittels ebener Wellen approximiert und so modale Effekte minimiert. Zusätzlich können virtuelle Szenen durch Spracheingabe des Benutzers gesteuert werden, wozu ein Verbund von 22 Mikrofonen in der Decke der Konstruktion verwendet werden kann. Dieser Beitrag erläutert die Entwicklung und Konzeption des Audiosystems und dessen Eigenschaften, sowie die besonderen Herausforderungen aus Sicht des Akustikers.

Technik des Grundsystems

Die aixCAVE ist ein hochauflösendes Projektionssystem, welches es erlaubt, Benutzer in eine sie umgebende stereoskopische visuelle Darstellung einzuhüllen. Das aixCAVE VR-System kann zur dreidimensionalen Visualisierung komplexer Daten (z.B. Strömungssimulationen) und auch zur Erzeugung interaktiver virtueller Szenen genutzt werden (*virtual walk-throughs*).

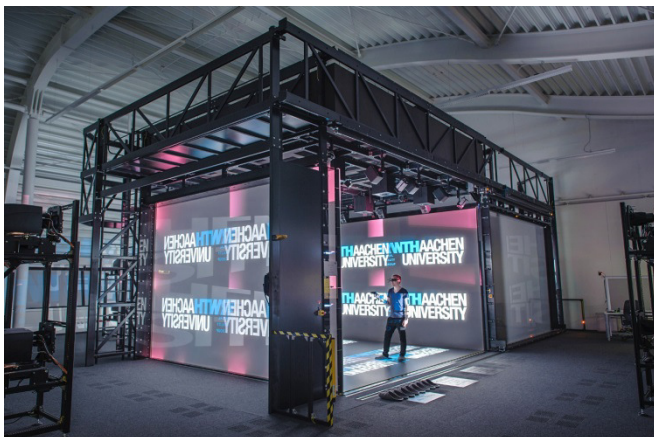


Abbildung 1: aixCAVE Virtual Reality System

Die aixCAVE hat eine Grundfläche von 5×5m und eine Höhe von 3,4m. 24 Projektoren (Barco Galaxy NW12) werfen hochauflösende Bilder auf die nahtlos ineinander übergehenden Projektionsflächen (Quaderraum mit Boden und vier Wände). Diese Projektionsflächen werden rückseitig illuminiert und bestehen aus ca. 15mm dickem, praktisch schallhartem Acrylglas. Stereoskopisches Sehen wird durch Shutterbrillen ermöglicht. Die Eingangstür der CAVE kann für eine umschließende Projektion geschlossen werden. Die Decke ist keine Projektionsfläche. Ein optisches Motion-Tracking-System bestimmt in Echtzeit Position und Orientierung des Benutzers (Kopf) und die Lage weiterer Eingabegeräte.

Ziele und Herausforderungen

Damit die Simulationen solcher Welten als plausibel wahrgenommen werden, sollte die Simulation nicht nur auf einen hochqualitativen visuellen Reiz beschränkt sein, sondern auch die weiteren menschlichen Sinne in Betracht ziehen (Multimodalität). Das Ziel der hier beschriebenen Konstruktionsmaßnahmen war es, die aixCAVE für eine audio-visuelle Wiedergabe zu erweitern. Dadurch können Benutzer die Szenen nicht mehr nur rein visuell erleben, sondern auch auditive Reize wahrnehmen. Die Anforderungen sind hierbei nicht bloß akustische Effekte (z.B. Schallergebnisse aus bestimmten Richtungen zu hören), sondern aufwendige Akustiksimulationen und Schallfeldreproduktionen, welche der physikalische Wirklichkeit nahe kommen und diese adäquat nachbilden. Erst dadurch erschließen sich vielfältige neue Anwendungsmöglichkeiten, wie z.B. interaktive Planung der Raumakustik in der Architektur [1] oder Lärmbewertungen in der Verkehrs- und Städteplanung (*soundscape*s) [2].

Wie auch bei der Computergrafik (*visual rendering*), wird auch die Akustik (*audio rendering*) benutzerzentriert berechnet. Die Auralisierung in Echtzeit wird für die Position und Orientierung der Ohren des Hörers durchgeführt und bei Veränderungen mit geringstmöglicher Latenz aktualisiert. Ein umfassendes Softwaresystem simuliert hierzu die Schallausbreitung in der virtuellen Welt mittels der Methoden der Geometrischen Akustik (GA) [3-6]. In einem hybriden Simulationsansatz werden Spiegelschallquellen (*image sources*) für die Berechnung früher Reflexionen genutzt und akustisches Ray-Tracing für die Bestimmung des Nachhalls. Durch parallele Schallfeldsimulationen und umfangreiche Parallelisierung können die Reaktionszeiten gering gehalten werden. Trotzdem sind die Berechnungen sehr aufwendig und können nach je Aufwand der Szene auf ein PC-Cluster ausgelagert werden [5]. Dieser Artikel befasst sich mit den technischen Aspekten des Akustiksystems. Details zu den Simulationsalgorithmen und Signalverarbeitungstechniken finden sich in [3-6].

Absorber

Eine der ersten Maßnahmen bei der Planung des Akustiksystems bestand in der Verbesserung der Raumakustik, sowohl in der aixCAVE selbst (Reflektionen an den Projektionsflächen), als auch außerhalb (z.B. Störgeräusche der Projektoren). Hierzu wurden, im Rahmen der vorgegebenen Statik und Tragfähigkeit der Konstruktion, ein Maximum an Absorptionsfläche im Deckenbereich geschaffen. Aufgrund der massiven, schallundurchlässigen Projektionsflächen, stellte dies die einzige Möglichkeit dar. Da Schall aller Frequenzbereiche absorbiert werden soll, wurden poröse Absorber aus Polyurethan-Schaumstoff (PU-Schaum) gewählt. Eine spezielle Anordnung vertikal aufgehängter Blöcke (Abbildung 2) stellte sich, im Rahmen der zulässigen Gesamtmasse und des Gesamtvolumens, als effektivste Lösung heraus. Gleichzeitig ermöglicht sie weiterhin die nötige thermische Konvektion und Belüftung. . Insgesamt wurden 120 Blöcke der Abmessungen 1×1×0,15m verbaut. Die alternierende Anordnung der Blöcke wurde aus Gründen der Isotropie gewählt. Alle Blöcke sind oberseitig mit MDF-Trägern verklebt, welche an Stahlseilen aufgehängt und in gleicher Höhe nivelliert sind. Das Gesamtgewicht der Absorberkonstruktion liegt bei ≈600 kg.

Räumliche Klangwiedergabe

Konzeptionell werden in der aixCAVE binaurale Verfahren für eine räumliche Klangwiedergabe eingesetzt. Im einfachsten Fall geschieht dies durch Funkkopfhörer (Sennheiser RS170). Die Position und Orientierung des Hörers ist durch das Motion-Tracking-System bekannt und kann direkt zur interaktiven Auralisierung verwendet werden. Für eine bessere Immersion (Eintauchen des Nutzers in die dargebotene virtuelle Welt) ist es vorteilhaft die vom Benutzer getragene Hardware auf ein Minimum zu reduzieren (keine unnötigen Kabel, Geräte und Gewichtsbelastungen). Kopfhörer und InEar-Hörer isolieren den Benutzer *in* der virtuellen Realität und erschweren die Kommunikation mit den Begleitpersonen (z.B. Operator). Daher werden gezielt Verfahren zur binauralen Wiedergabe mittels Lautsprechern erforscht und konzipiert, welche Kopfhörer verzichtbar machen. Konstruktionsbedingt

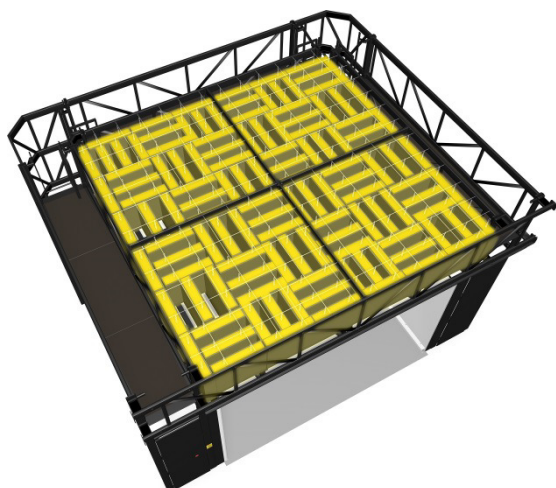
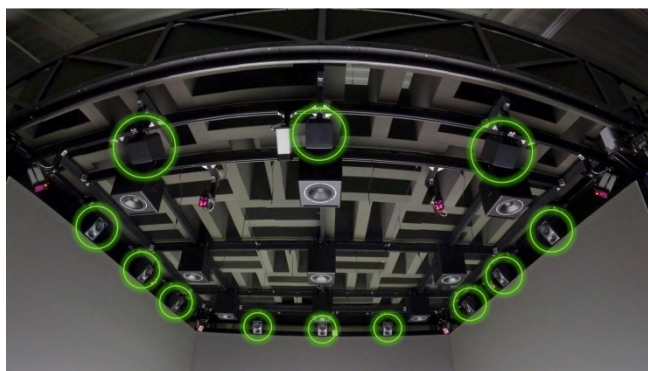
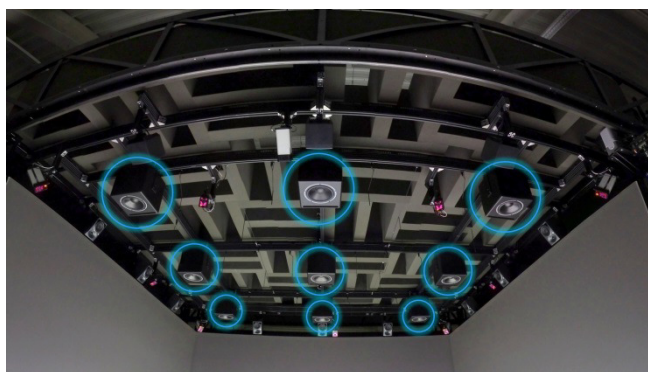


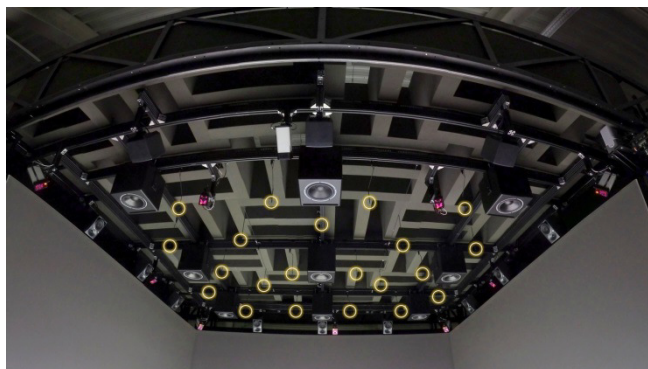
Abbildung 2: Anordnung der PU-Absorberblöcke (gelb markiert) im Dach der aixCAVE.



(a) Lautsprecher (12× Neumann KH120)



(b) Subwoofer (9× ProDipe Pro 10S)



(c) Mikrofone (22× Panasonic WM-61A)

Abbildung 3: Decke der aixCAVE mit installierten Audiokomponenten (Blick von unten nach oben).

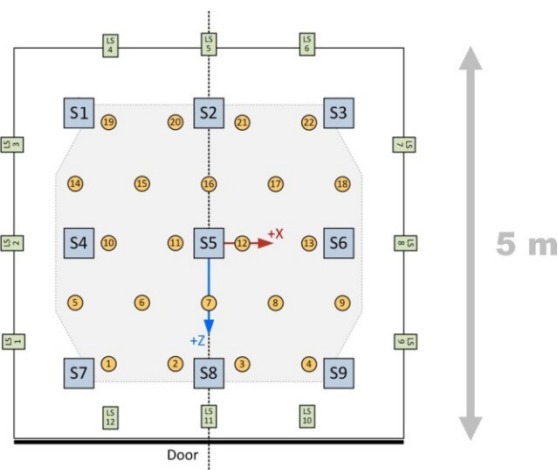


Abbildung 4: Anordnung der Audiokomponenten Draufsicht (grün markiert: Lautsprecher, blau markiert: Subwoofer, gelb markiert: Mikrofone).

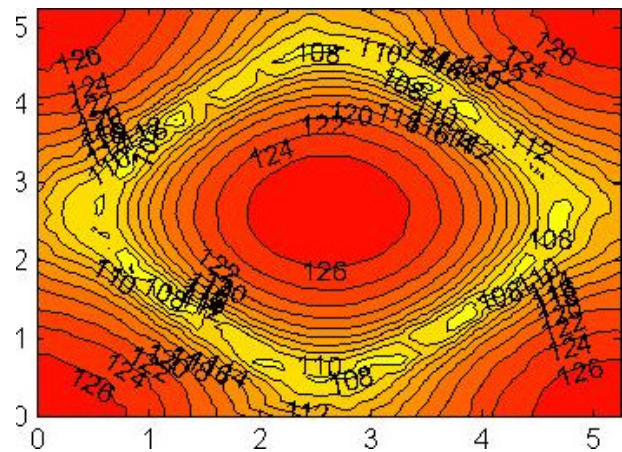
können die Lautsprecher nicht an beliebigen Positionen angebracht werden (z.B. akustische undurchlässige Wände und Verdeckung der Projektion), sondern nur oberhalb der Projektionsflächen. Eine Schallfeldsynthese mittels Ambisonics oder Wellenfeldsynthese (WFS) kommt daher in der aixCAVE nicht in Frage. Stattdessen wurde und wird in den Virtual Reality Systemen der RWTH Aachen eine lautsprecherbasierte binaurale Wiedergabe mittels dynamischer Übersprechkompensation (*dynamic crosstalk cancellation*) eingesetzt [6]. Es wurden zwölf Nahfeld-Monitor-Lautsprecher vom Typ Neumann KH 120 oberhalb der Projektionsflächen angebracht (Abbildung 3a). Die Lautsprecher weisen eine vergleichsweise geringe Richtwirkung auf, was für die Übersprechkompensation vorteilhaft ist. Die zwölf Lautsprecher wurden azimuthal in gleichmäßigen Winkelschritten von 30° angeordnet. Strategien zur optimalen Lautsprecherauswahl für die Übersprech-Kompensation werden gegenwärtig in Aachen erforscht. Weiterhin wird an Methoden geforscht um den negativen Einfluss der Akustik in der aixCAVE auf die dargebotene auralisierte Szene zu minimieren [7].

Tieffrequente Klangwiedergabe

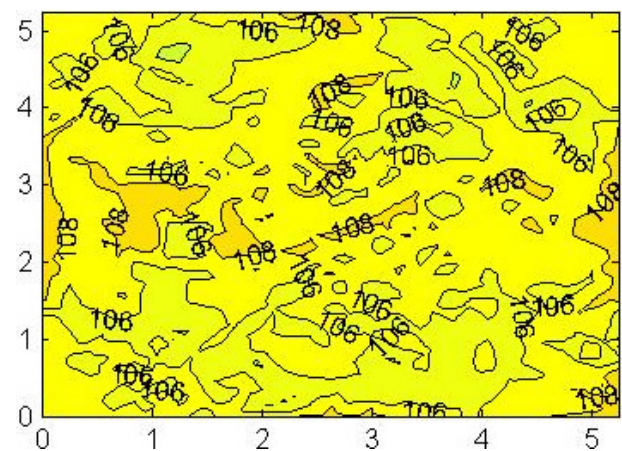
Die zwölf Lautsprecher sind alleine nicht in der Lage bei tiefen Frequenzen (<80 Hz) die gewünschten Schalldruck zu erzeugen, wie sie gerade bei Verkehrslärmsimulationen auftreten. Daher wurde die aixCAVE mit zusätzlichen neun Subwoofern (ProDipe Pro 10S) ausgerüstet. Die nahezu schallharten Wände der aixCAVE führen zu sehr ausgeprägten Modaleffekten im Schallfeld und in Folge dessen zu einem stark ortsabhängigen Schalldruck bei tiefen Frequenzen. Dies ist nicht wünschenswert. Abbildung 5a zeigt den resultierenden Schalldruckpegel (FEM-Simulation, $f=64$ Hz) in horizontaler Ebene auf Ohrhöhe (1,70m) bei Verwendung eines einzelnen, zentral angebrachten Subwoofers (Position SW5 in Abbildung 4). Zu erkennen sind ringförmige Einbrüche des Pegels von mehr als 10 dB. Bei höheren Frequenzen treten diese Fluktuationen örtlich noch dichter auf. Um diese unerwünschten Effekte zu vermindern, werden 3×3 Subwoofer in spiegelsymmetrischer Anordnung kohärent betrieben. Die zugrundeliegende Idee ist die Approximation ebener Wellen [8], welche sich für diese Konstellation unterhalb von 102 Hz ausbilden. Aufgrund der geringen Distanz (ca. 2-2,5m) zwischen den Subwoofern und dem Kopf des Hörers kann eine ebene Wellenfront zwar nur angenähert werden. Trotzdem werden Quermoden in horizontaler Richtung weniger stark angeregt und die Pegel im tieffrequenten Schallfeld sind deutlich ortsunabhängiger (siehe Abbildung 5b). In FEM-Simulationen wurden hierbei Abweichungen im Bereich von ± 2 dB ermittelt. Auch bei weiteren Frequenzen im Bereich von $f < 120$ Hz wurden ähnliche Verbesserungen gefunden.

Mikrofonarray

Die Möglichkeit virtuelle Szenen durch akustische Eingaben zu beeinflussen und zu steuern fördert die Immersion des Benutzers. Eine Spracherkennungssoftware kann verwendet werden um beispielsweise Steuerkommandos aufzuführen (z.B. „lade Szene“, „öffne Tür“). Ferner ermöglicht die entwickelte Akustik-Software die Stimme des Benutzers in Echtzeit im virtuellen Raum zu auralisieren [9]. Um die



(a) Einzelner zentraler Subwoofer (SW5)



(b) Neun symmetrisch angeordnete Subwoofer (3×3)

Abbildung 5: FEM-Simulationen der horizontalen Schalldruckverteilung in Höhe der Ohren (1,7 m), bei monofrequenter Anregung ($f=64$ Hz). Die Isobaren markieren Pegelschritte von 2 dB.

auditiven Eingaben des Benutzers zu erfassen kann dieser ein Nackenbügel-Funkmikrofon (Sennheiser EW300) tragen. Bei der Konzeption des Akustiksystems wurde aber auch die Möglichkeit erörtert hierfür ein Mikrofonarray einzusetzen, um auf das Tragen eines Mikrofons verzichten zu können. Hierzu wurden 22 omnidirektionale Elektretmikrofonkapseln (Panasonic WM-61A) in hexagonaler Anordnung in die Decke der aixCAVE integriert. Der Abstand der Mikrofone beträgt ca. 90cm. Aufgrund dessen ist die Richtwirkung des Arrays physikalisch stark begrenzt. Zur selektiven Trennung einzelner Sprecher und zur Auralisierung auditiver Eingaben (Sprechen, Klatschen, usw.) ist es daher vermutlich nicht hinreichend. Gegenwärtig werden Ansätze untersucht mittels tracker-gestützter Gewichtung der Mikrofonsignale und Delay-and-Sum Beamforming adäquate Nutzsignale zur Spracherkennung zu gewinnen, welche geringere Störanteile des Außenlärms (Projektoren, mögliche Audiowiedergabe) enthalten.

Signalwandlung

Das Audiosystem wurde auf insgesamt 32 Eingabe- und 32 Ausgabekanäle ausgelegt. Vier AD/DA-Wandler des Typs Behringer ADA8000 (jeweils 8 Ausgabekanäle und 8 vorverstärkte Eingangskanäle pro Gerät) werden mittels Lichtwellenleitern im ADAT-Protokoll angesteuert. Ein zentraler Akustikrechner ist mit einer RME RayDat Audiokarte ausgerüstet und ermöglicht eine Audioverarbeitung in Echtzeit, mit geringer Latenz. Drei der AD/DA-Wandler befinden sich im Dach der aixCAVE und steuern die zwölf Lautsprecher und neun Subwoofer an. Zusätzliche Endstufen mussten nicht verbaut werden, da alle Lautsprecher aktiv sind. Alle 22 Mikrofone verfügen über eigene Vorverstärkerschaltungen (integriert im Steckverbinder) und sind direkt mit den Vorverstärkern der Wandler verbunden (Phantomspannung). Der vierte AD/DA-Wandler befindet sich am Arbeitsplatz des Akustikrechners und dient zur Ansteuerung der Kopfhörersysteme, kabellosen Mikrofone (*Nackebügel*), lokaler Abhörmonitore, sowie Kommunikationskanälen (*talkback*). Die Verwendung von Lichtwellenleitern stellte sich angesichts der Kabellängen von über 15m als besonders vorteilhaft heraus.

Zusammenfassung

Mit den durchgeführten Arbeiten wurden die akustischen Grundlagen geschaffen, mit dem aixCAVE VR-System hochqualitative audio-visuelle virtuelle Umgebungen zu generieren. Zum gegenwärtigen Stand befinden sich Aufnahme- und Wiedergabetechniken noch im Entwicklungsstadium. Erste Versuche verliefen vielversprechend und deuten darauf hin, dass die angestrebten Ziele der kopfhörerlosen Binauralwiedergabe trotz der nicht-optimalen akustischen Randbedingungen erreicht werden können. Ein zentrales Ziel ist die audio-visuelle Darbietung, welche den Benutzer vom tragen zusätzlicher Geräte befreit (Kopfhörer, Mikrofone, Send- und Empfangsgeräte). Zukünftige Untersuchungen werden zeigen müssen, inwiefern diese Technik eine Alternative zur binauralen Klangwiedergabe mittels Kopfhörern darstellt und mit welchen Einschränkungen dies einhergeht. Gleiches gilt für Aufnahme akustischer Eingaben (Klänge, Sprache, usw.). Hier sind geeignete Techniken, auch in Kombination mit einer gleichzeitigen lautsprecherbasierten Klangwiedergabe, noch näher zu untersuchen. Erste Anwendungen des neuen Systems finden sich bereits in der Raumakustiksimulation und in der Bewertung von Verkehrslärm. Beispielsweise wurden vollständig echtzeitfähige Raumakustiksimulation von Konzertsälen und einer historisch rekonstruierten spanischen Kirche erstellt [10] und ein System zur interaktiven Auralisierung von Flugverkehrslärm [2].

Referenzen

- [1] M. Vorländer, D. Schröder, S. Pelzer, F. Wefers, "Virtual reality for architectural acoustics". *Journal of Building Performance Simulation*, Vol. 8(1), 15-25, 2015.
- [2] A. K. Sahai, E. Anton, E. Stumpf, F. Wefers, M. Vorländer, „*Interdisciplinary Auralization of Take-off and Landing Procedures for Subjective Assessment in Virtual Reality Environments*“, 18th AIAA/CEAS Aeroacoustics Conference, Colorado Springs, USA, 2012.
- [3] T. Lentz, D. Schröder, M. Vorländer, I. Assenmacher. "Virtual reality system with integrated sound field simulation and reproduction", *EURASIP journal on applied signal processing*, Vol. 2007, 2007.
- [4] D. Schröder, F. Wefers, S. Pelzer, D. Rausch, M. Vorländer, T. Kuhlen, "Virtual Reality System at RWTH Aachen University". In 20th International Congress on Acoustics (ICA2010), Sydney, Australia, 2010.
- [5] F. Wefers, J. Stienen, M. Vorländer, „*Interactive acoustic virtual environments using distributed room acoustics simulations*“, EAA Joint Symposium on Auralization and Ambisonics, Berlin, Germany, 2014.
- [6] D. Schröder, "Physically based real-time auralization of interactive virtual environments". Dissertation, Logos Verlag Berlin GmbH, 2011.
- [7] M. Kohnen, S. Pelzer, M. Vorländer, „*3-D audio reproduction in difficult environments of virtual reality systems*“, Fortschritte der Akustik (DAGA), Nürnberg, 2015.
- [8] A. Goertz, M. Wolff, L. Naumann, "Optimierung der Tieftonwiedergabe in Tonstudios und Abhörräumen", Fortschritte der Akustik (DAGA), Aachen, 2003.
- [9] Frank Wefers, Michael Vorländer. "Interactive acoustic feedback into virtual acoustic scenes", EAA Euroregio 2010, Ljubljana, Slovenia, 2010.
- [10] A. Pedrero, A. Díaz-Chyla, S. Pelzer, M. Pollow, C. Díaz, M. Vorländer, "Auralization of Mozarabic chant in a pre-romanesque church". *Tecnicastica* 2013, Valladolid, Spain, 2013.