



Hochschule für Angewandte Wissenschaften Hamburg
Hamburg University of Applied Sciences

AW2 Ausarbeitung

David Hemmer

Mobile Augmented Reality Audio Systeme

Inhaltsverzeichnis

| | |
|---|-----------|
| 1 Einführung | 2 |
| 1.1 Einleitung | 2 |
| 2 Vergleichbare Arbeiten | 3 |
| 2.1 Virtual acoustics and 3-D Sound in Multimedia signal processing | 3 |
| 2.2 Augmented Reality Audio for Mobile and Wearable Appliances | 4 |
| 2.3 Augmented Reality Audio Application in Outdoor Use | 7 |
| 2.4 Urban Sound Gardens | 8 |
| 2.5 3D Audio Augmented Reality: Implementation and Experiments | 9 |
| 3 Zusammenfassung und Ausblick | 10 |
| 4 Anhang | 11 |
| 4.1 Texas Instruments - Spatial Audio Technology | 11 |
| Abbildungsverzeichnis | 12 |
| Abkürzungsverzeichnis | 12 |
| Literaturverzeichnis | 13 |

1 Einführung

1.1 Einleitung

Die virtuelle und erweiterte Realität spielt schon heute in der Industrie und im Heimbereich eine große Rolle in unserem Alltag. Viele der Anwendungsbereiche erfordern die Mobilität des Benutzers. Dazu werden immer neue Interaktionsmöglichkeiten mit dem System entwickelt und somit verschimmt die reale Welt noch schneller mit der virtuellen Welt. Heutzutage beschränkt sich die virtuelle und erweiterte Realität noch sehr auf das visuelle Sinnesorgan.

Ein zweites wichtiges Sinnesorgan des Menschen ist das Gehör. Das menschliche Gehör ist bei jeder Wahrnehmung ganz automatisch beteiligt. Es erlaubt uns viele Informationen über unsere Umgebung unterbewusst wahrzunehmen. Dazu zählen Informationen des Raumes, in dem wir uns befinden anhand der Raumakustik. Ebenso können wir die Position und Entfernung von Schallquellen bestimmen. Ein entsprechendes Augmented Reality Audio (ARA¹) System muss die Realität so genau wie möglich nachbilden. Das bedeutet, dass ein solches System die virtuellen Schallquellen in Echtzeit berechnen und in die Realität einbinden muss. Im Bezug auf virtuelle und erweiterte Realitätssysteme darf ein solches System die Mobilität des Benutzers nicht beeinträchtigen.

Durch ein ARA-System wird die reale oder virtuelle Umwelt mit Informationen angereichert. Daraus folgt eine Vielzahl von neuen Einsatzmöglichkeiten für uns Menschen. Zum einen können die bestehenden Systeme erweitert werden, zum anderen ist ein solches System auch ein neuartiges Auditory Display, welches in einer Vielzahl von Anwendungsbereichen eingesetzt werden kann. Mit der Entwicklung eines mobilen Augmented Reality Audio System möchte ich mich im Rahmen meiner Masterarbeit an der HAW Hamburg beschäftigen.

In dieser Ausarbeitung werden vergleichbare Arbeiten vorgestellt und die Erkenntnisse aus Ihnen für mein Projekt bewertet.

¹Abkürzungen sind im Abkürzungsverzeichnis im Anhang erläutert

2 Vergleichbare Arbeiten

In diesem Kapitel werden mehrere vergleichbare Arbeiten aus dem Bereich der mobilen Augmented Reality Audio vorgestellt. Dabei unterscheiden sich die vorgestellten Arbeiten in zwei Bereiche. Zu einem in Arbeiten mit theoretischen Ansätzen und zum anderen in Arbeiten in denen vergleichbare Mobile Augmented Reality Audio Systeme vorgestellt werden.

2.1 Virtual acoustics and 3-D Sound in Multimedia signal processing

Diese Ausarbeitung ist die Dissertation von Jyri Huopaniemi von der Helsinki University of Technology vom 5. November 1999 [Huo99]. Schwerpunkt seiner Dissertation sind die Aspekte von der Echtzeit Modellierung und Synthese für 3D-Sound im Bezug auf digitale Audio, Multimedia und virtuelle Umgebungen.

In der Dissertation werden mehrere Möglichkeiten für Augmented Reality Audio Systeme vorgestellt. Zwei dieser Vorschläge können auch für Mobile ARA Systeme eingesetzt werden. Diese beiden Möglichkeiten sind in Abbildung 1 dargestellt.

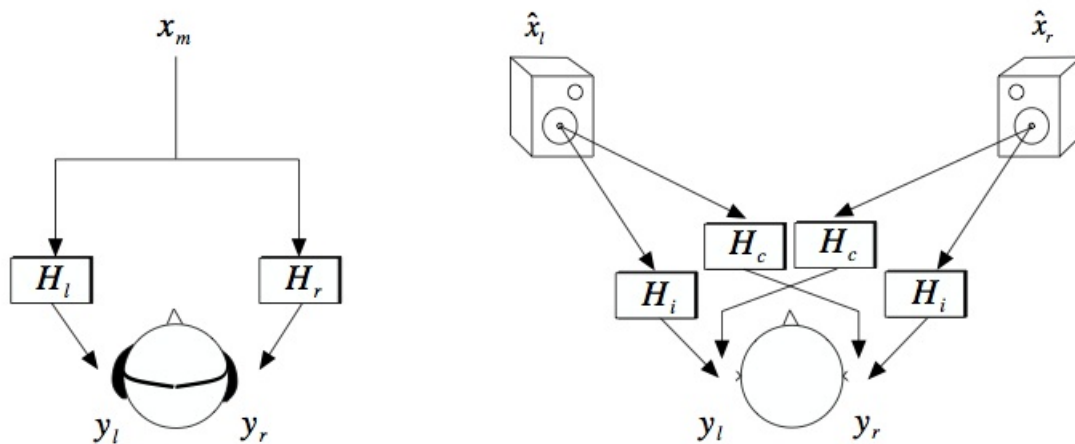


Abbildung 1: Signalflussgraph: links: binaural und rechts: crosstalk canceled binaural audio processing

Der binaurale Ansatz, auch Headphone Reproduction (links) genannt, arbeitet über Kopfhörer. Das Mono-Audiosignal x_m wird über die HRTF-Filter H_l und H_r zur virtuellen Audioquelle gefiltert und kann über die entsprechende Kopfhörerseite abgespielt werden. Dieses Verfahren berücksichtigt nicht die Akustik des Raumes des Zuhörers ebenso wird die Kopfposition des Zuhörers nicht mit einbezogen. Durch Head-Tracking kann die Kopfposition mit in die Audiofilterung einbezogen werden.

Der zweite Ansatz ist die Loudspeaker Reproduction (rechts). Die crosstalk canceled binauralen Signale x_l und x_r werden über die Lautsprecher ausgestrahlt, um die Wunsch-Signale y_l und y_r zu erhalten. Die richtungsabhängigen Lautsprecher-zu-Ohr-Übertragungsfunktionen H_i und H_c müssen beachtet werden, um den Effekt von virtuellen Schallquellen zu erfüllen. Die Filterung kann als ein kaskadierter Prozess verstanden werden. Zum einen die HRTF-Filterung und die separate crosstalk canceling Filterung. Bei diesen Verfahren gibt es zwei wesentliche Grenzen: Die Zuhörerposition und die Raumposition des Zuhörers.

Im Verlauf meiner Ausarbeitung wird nur noch der Ansatz mittels Kopfhörer verfolgt. Der zweite Ansatz ist trotzdem nicht zu vernachlässigen. Für diesen Ansatz gibt es schon ein erstes marktfähiges Produkt der Firma Texas Instruments. Näheres dazu im Anhang 4.1.

Die Dissertation von Jyri Huopaniemi gibt eine gute Übersicht vom Aufgabenfeld der Augmented Reality Audio. Des Weiteren zeigt er in seiner Arbeit, dass eine Vielzahl der benötigten Audio-Filterungen auf Digital Signal Processors (DSPs) bearbeitet werden können. Das ist für die von uns gewählte Entwicklungsplattform eines System on Chip (SoC) auf einem FPGA sehr gut geeignet.

2.2 Augmented Reality Audio for Mobile and Wearable Appliances

Dieses Paper wurde in Zusammenarbeit von der Helsinki University of Technology und dem Nokia Research Center (Media Technologies Lab) im Journal der Audio Engineering Society in der Ausgabe 52 (JAES-52) am 6. August 2003 vorgestellt [HJT⁺04]. Dabei unterteilt es sich in zwei verschiedene Bereiche. Im ersten Bereich werden die Anforderungen an eine mobiles Augmented Reality Audio (MARA) System festgelegt. Im anschließenden Bereich wird ein erstes Augmented Reality Audio System vorgestellt.

Zuerst wird das Konzept der Augmented Reality Audio Technik charakterisiert. Es handelt sich dabei um die Erweiterung der realen Umwelt mit virtuellen auditiven Umgebungen und oder Kommunikations-Szenarien. Dabei werden verschiedene Anforderungen an ein Augmented Reality Audio System festgestellt. Diese gelten ebenso für eine mobiles Augmented Reality System. Es werden folgende Punkte aufgeführt:

- Die virtuelle Audioumgebung verbindet sich mit der Realität
- Mobilität und Benutzerfreundlichkeit des Systems
- Echtzeitfähigkeit
- Raumklang

Mit dem ersten Punkt: "Die virtuelle Audioumgebung verbindet sich mit der Realität" ist der Augmented Reality Ansatz gemeint. Ein ARA-System muss sicherstellen, dass der Zuhörer gleichzeitig die reale Umwelt und die virtuelle erzeugte Umwelt wahrnehmen kann. In diesem Teilbereich von ARA gibt es viele verschiedene Ansätze. Der einfachste und zurzeit noch am meisten genutzte Ansatz, sind die sogenannten Augmented Reality Audio Kopfhörer mit integrierten Mikrofonen und entsprechendem Audio-Mixer. Dabei wird die reale Audioumgebung mittels der Mikrofone aufgenommen und mit den virtuell erzeugten Audiosignalen im AR-Mixer zusammengeführt und über die Kopfhörer abgespielt.

Ein weiterer wichtiger Punkt ist die Mobilität und Benutzerfreundlichkeit eines solchen ARA-Systems. Es muss sichergestellt sein, dass ein solches System auf den Benutzer keinerlei Einschränkungen im Bezug auf die Bewegungsfreiheit und Mobilität hat. Ebenso muss es sich intuitive bedienen lassen. Im besten Fall soll dem Benutzer des Systems gar nicht bewusst sein, dass er ein ARA-System nutzt.

Die Echtzeitfähigkeit des Systems ist ebenso ein wichtiger Punkt. Es muss auf jeden Fall sichergestellt werden, dass das System auf jede Aktion sofort reagieren kann. Wenn der Benutzer sich zum Beispiel bewegt, ändert sich auch gleichzeitig die relative Position zu den virtuellen Audioquellen. Wenn die Positionsänderung nicht sofort wahrgenommen wird und die Audiofilterung nicht angepasst wird, erzeugt dieses Effekte, sodass die Illusion der virtuellen Schallquellen nicht mehr funktioniert. Bei diesem Problem gibt es zwei wichtige Faktoren, die berücksichtigt werden müssen. Zum einen müssen die Bewegungen des Benutzers schnellstmöglich und mit einer hohen Genauigkeit aufgenommen werden, zum anderen muss sichergestellt werden, dass das System sofort auf diese Änderung reagiert.

Der letzte Punkt, der aufgeführt wird, ist der Raumklang. Dabei gibt es zwei unterschiedliche Punkte, die beachtet werden müssen. Zum einen müssen die virtuellen Audioquellen der Umgebung angepasst werden. Befindet sich zum Beispiel der Benutzer in einer großen Halle und die virtuelle Audioquelle weist nur die Eigenschaften eines kleinen Raumes auf, führt diese beim Benutzer zu einer nicht realen Einschätzung und die Audio-Quelle wird nicht als real wahrgenommen. Zum anderen muss auch die Position des Benutzers im Raum eingebunden werden, um eine real wirkende virtuelle Audioumgebung zu erzeugen.

Im zweiten Teil des Papers wurde ein Vergleich von realen und virtuellen Schallquellen gemacht. Der Aufbau des Vergleiches wird in Abbildung 2 dargestellt.

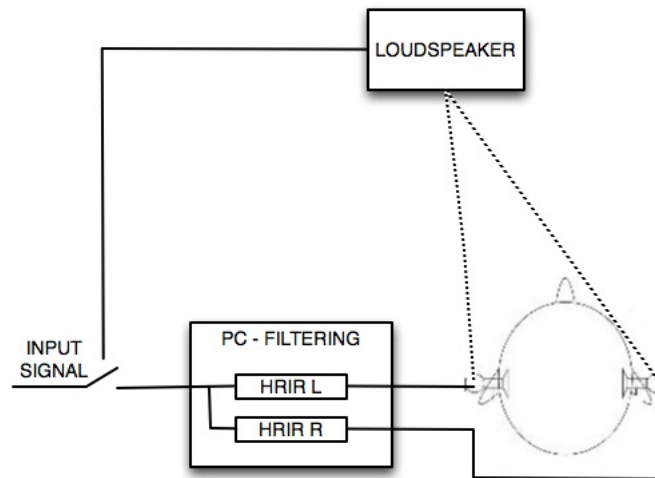


Abbildung 2: Zuhörer Test Aufbau. Vergleich virtueller Audioquellen und realer Audioquellen über AR-Audio Kopfhörer

Die Testpersonen nutzten Kopfhörer mit integrierten Mikrofonen und einem festen ARA-Mixer. Die Head Related Impulse Response (kopfbezogene Impulsantwort) wurde zwischen der Lautsprecherposition und der Kopfposition gemessen. Für den Vergleich wurden die Signale entweder über die Lautsprecher oder über die gemessenen HIRI's in einem PC gefiltert. Zu erwähnen ist, dass die Kopfposition nicht mit berücksichtigt wurde. Somit sind Änderungen der Kopfposition nur für die Lautsprecheridentifikation richtig, was zu einer Verfälschung der Testergebnisse führen kann. Der Vergleichstest zwischen realer Schallquelle und virtueller Schallquelle zeigt, dass es für die meisten Zuhörer nicht möglich ist, zwischen beiden zu unterscheiden. Somit konnte gezeigt werden, dass es theoretisch möglich ist, virtuelle Audioquellen zu simulieren, ohne dass es einen hörbaren Unterschied zwischen realen Soundquellen gibt.

Erkenntnisse für meine Masterarbeit aus diesem Paper sind die Anforderungen an ein MARA-System, sowie dass der Unterschied zwischen realen und virtuellen Audioquellen sehr klein ist. Nachteilige Erkenntnisse für meine Arbeit ist der feste AR-Mixer, sowie die nicht vorhandene Mobilität des Systems.

2.3 Augmented Reality Audio Application in Outdoor Use

Das von Mikko Peltola in seiner Masterarbeit [Pel09] (23.02.2009) an der Helsinki University of Technology entwickelte mobile Augmented Reality Audio System, ist für den Einsatz außerhalb von Gebäuden gedacht. Das System ist in Abbildung 3 dargestellt. Die Audioverarbeitung sowie die nötigen Koordinationsaufgaben werden auf einem handelsüblichen Laptop im Rucksack berechnet. Die Personenbestimmung arbeitet über GPS und mittels eines Head-Trackers, welcher auf einer Cappy befestigt ist. Die Benutzerinteraktion mit dem System erfolgt mittels einer Wii-Remote Fernbedienung.

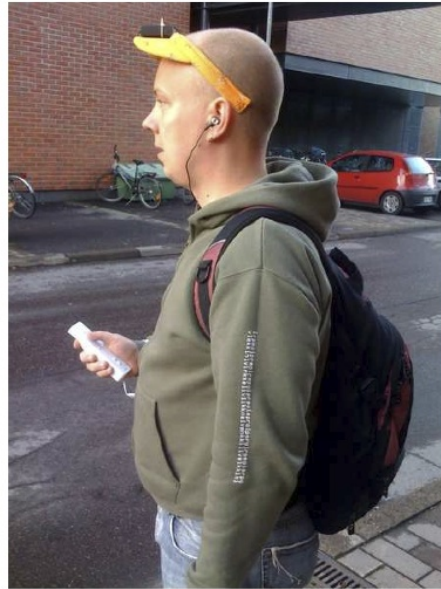


Abbildung 3: Entwickeltes MARA-System

Der AR-Ansatz wurde in dieser Arbeit vernachlässigt. Er verweist für dieses Themenfeld auf die zur gleichen Zeit geschriebene Dissertation “Development and evaluation of Augmented Reality Audio Systems“ von Miikka Tikander. [Tik09].

In seiner Arbeit stellt er verschiedene Möglichkeiten vor, um Zusatzinformationen mit Audiosignalen zu verbinden. Solche Zusatzinformationen sind zum Beispiel die Positionen einer sich bewegenden Schallquelle. Um solche Informationen mit dem Audiosignal zu verbinden, nutzt er das ARA-Wave Format. Dieses Audioformat wurde bei der 35. international Conference der AES in London im Februar 2009 vorgestellt.

Erkenntnis aus der Masterarbeit von Mikko Peltola ist ein funktionierendes MARA-System mit der Einschränkung, dass der AR-Ansatz vernachlässigt wurde. Ebenso stellt er das ARA-Wave Format vor, welches gegebenenfalls auch in meinem Projekt eingesetzt werden kann. Des Weiteren wird auf die Dissertation von Mikka Tikander hingewiesen, welche das Themengebiet der ARA behandelt.

2.4 Urban Sound Gardens

Dieses Projekt ist eine Zusammenarbeit der Uni Glasgow und Nokia [VAOB11]. Dabei wurde ein Auditory Display mittels eines Nokia N95, Head-Tracking Sensor und GPS aufgebaut. Die Audioverarbeitung wurde mittels der JSR 234 Library (draft) realisiert. Diese Library ist für die Symbian-Plattform entwickelt worden. In der in diesem Projekt verwendeten Version kann diese Library ein 16 bit Mono-Signal mit 16kHz mittels HRTF verarbeiten. Eine Einbindung der realen Audioumgebung findet nicht statt.

Das System wurde im Municipal Gardens in Funchal, Madeira eingesetzt (Abbildung 4). Dabei wurde getestet, welche Auditory Display Verfahren unter dem Schwerpunkt, der überlappenden Audioquellen, für den Benutzer am besten geeignet ist. Jede Audioquelle hat eine aktive Zone (5m Radius) und einen Nahbereich (12,5m Radius).



Abbildung 4: Urban Sound Gardens

Bei dem Testaufbau wurden vier verschiedene Auditory Display Verfahren auf die Informationsvermittlung und Benutzerwahrnehmung verglichen.

- **Baseline:** Wenn der Benutzer in die aktive Zone kommt, wird der dazugehörige Audio-Clip abgespielt.
- **Earcons:** In der aktiven Zone wurde die Audioquelle entsprechend der Position abgespielt.

- Spatial: Wenn der Benutzer in den Nahbereich kommt, wird das Audiosignal (z.B. Tierstimme) entsprechend der Position ausgelöst. Eine Entfernungssimulation wird über die Lautstärke simuliert.
- Spatial 3D: Wie Spatial aber mit voller Verräumlichung.

Das Ergebnis dieses Tests zeigt, dass es für den Menschen eindeutig möglich ist auch mehrere simulierte Audioquellen voneinander zu unterscheiden und die Position zu bestimmen. Von den Benutzern wurde die 3D-Audio-Umgebung als angenehmste und informativste wahrgenommen.

2.5 3D Audio Augmented Reality: Implementation and Experiments

Dieses Paper wurde von der US Firma Rockwell Scientific (Militär) vorgestellt [SWC⁺03]. Es wurde ein tragbares 3D-Audiosystem mittels eines tragbaren PC sowie Headtracking-Sensoren und GPS-Empfänger aufgebaut. Ein Augmented Reality Audio Ansatz wurde nicht verfolgt.

Das entwickelte System wurde für eine Testreihe benutzt, in der getestet wurde, wie gut sich ein Mensch an das 3D-Audio-System anpassen kann. Dazu mussten die Testpersonen mehrere Tage hintereinander das System benutzen und virtuelle Schallquellen orten. Es konnte gezeigt werden, dass die Testpersonen die Position der Schallquellen schon nach 4 Tagen deutlich besser lokalisieren konnten. Durch die tägliche Benutzung des Systems konnte ebenfalls eine deutliche Verbesserung der Front-Back Lokalisierung festgestellt werden.

In diesem Projekt wurde durch die Testreihe gezeigt, dass sich der Mensch sehr schnell an ein solches System anpassen kann. Diese Erkenntnis lässt für meine Weiterentwicklung des Systems die Hoffnung aufkommen, dass sich durch eine ungefähre Abbildung der Realität eine gute Simulation von virtuellen Schallquellen erzeugen lässt.

3 Zusammenfassung und Ausblick

Die hier vorgestellten Arbeiten geben einen hervorragenden Überblick über die Schwerpunkte von “Augmented Reality Audio“. Die Dissertation von Jyri Huopaniemi zeigt zugleich verschiedene Ansatzmöglichkeiten von “Augmented Reality Audio“ (Kapitel: 2.1). Des Weiteren geht er in seiner Arbeit davon aus, dass sich eine Vielzahl von Audio-Filterungen von ARA mittels DSPs bearbeiten lassen. Dieses spricht dafür, dass die von uns gewählte Entwicklungsplattform eines SoCs auf einem FPGA ein guter Weg ist.

Der Bericht aus dem Journal der AES von der Helsinki University of Technology und dem Nokia Research Center (Media Technologies Lab) (Kapitel: 2.2) zeigt welche Eigenschaften für ein Mobile Augmented Reality Audio System wichtig sind. Ebenso zeigt die durchgeführte Testreihe, dass es möglich ist den Benutzern virtuelle Schallquellen vorzutäuschen.

Die Masterarbeit von Mikko Peltola (Kapitel: 2.3) ist eines der ersten mobilen ARA-Systeme. In seiner Arbeit stellt er das ARA-Wave Format vor, welches gegebenenfalls auch für das Audiostreaming in meinem System eingesetzt werden kann.

Die Erkenntnisse aus der Zusammenarbeit der Uni Glasgow mit Nokia (Urban Sound Garden) (Kapitel: 2.4) sowie die der US Firma Rockwell Scientific (Kapitel: 2.5) zeigen, dass der Mensch die Vorteile eines MARA Systems nutzen kann. Ebenso wird gezeigt, dass Menschen in der Lage sind sich schnell an ein solches System anzupassen.

Die Erkenntnisse zeigen, dass es sich lohnt ein solches System aufzubauen. Ebenso bestärkten die Erkenntnisse mich darin, dass der eingeschlagene Weg ein MARA System mittels eines SoCs zu lösen ein guter Weg ist. Eine Vielzahl von Anforderungen eines solchen MARA Systems gut mittels eines FPGAs bearbeitet werden können. Dazu zählt die Audiofilterung in parallelen DSPs und die daraus folgende Echtzeitfähigkeit.

4 Anhang

4.1 Texas Instruments - Spatial Audio Technology

Im vierten Quartal 2011 hat Texas Instruments den Baustein LM28901 [Ins11] vorgestellt. Dieser Baustein ist für kleine mobile Geräte gedacht. Er erzeugt über kleine Lautsprecher ein räumliches Klangerlebnis. Jeder Baustein kann ein Lautsprecherarray von vier Lautsprechern verwalten. Zudem können bis zu vier dieser Bausteine miteinander verbunden werden. Somit können bis zu 16 Lautsprecher im Lautsprecherarray zusammengeschaltet werden. Um das räumliche Klangerlebnis zu erzeugen, arbeitet der Baustein mit dem Verfahren Beamforming, Crosstalk Cancellation und Head Related Transfer Function (HRTF). (siehe Abbildung 5)

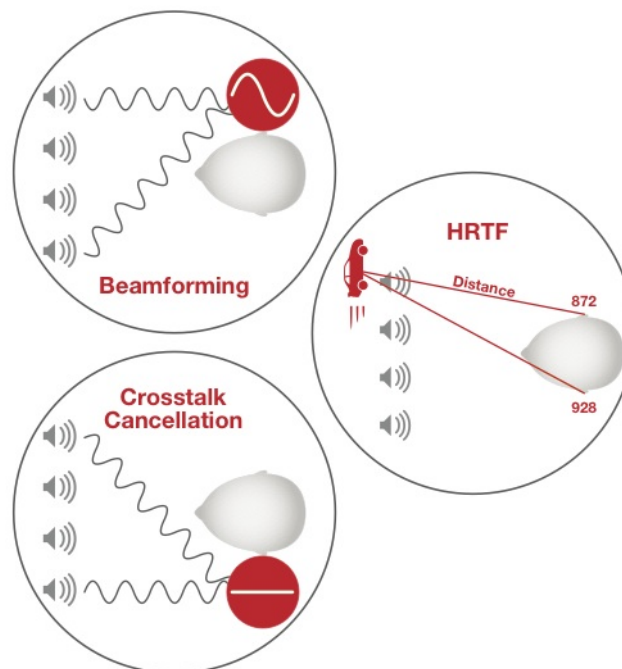


Abbildung 5: Texas Instruments - Spatial Audio Technology

- **Beamforming:** Durch dieses Verfahren werden Soundbereiche erzeugt, indem durch Zusatz von konstruktiv-störenden Schallwellen andere Schallwellen verstärkt werden und daraus eine reichere Klangpalette für den Zuhörer erzeugt wird.
- **Cross Cancellation:** Erzeugt zerstörende Wellen für ein Soundbereich. Diese Anti-Wellen sind dafür gedacht, dass das rechte Ohr nur die Signale des rechten Ohres hört und nicht die Signale für das linke Ohr und umgekehrt.

- HRTF: Erzeugt die Wahrnehmung von Klang aus verschiedenen Richtungen. Dazu arbeitet dieses Verfahren mit Zeitlicher-Differenz und Intensität-Differenz zwischen den beiden Ohren. Hieraus wird ein realistischer 3D-Raum erzeugt.

Abbildungsverzeichnis

| | | |
|---|--|----|
| 1 | Signalflussgraph: links: binaural und rechts: crosstalk canceled binaural audio processing | 3 |
| 2 | Zuhörer Test Aufbau. Vergleich virtueller Audioquellen und realer Audioquellen über AR-Audio Kopfhörer | 6 |
| 3 | Entwickeltes MARA-System | 7 |
| 4 | Urban Sound Gardens | 8 |
| 5 | Texas Instruments - Spatial Audio Technology | 11 |

Abkürzungsverzeichnis

| | |
|---------|--|
| AES | Audio Engineering Society |
| AR | Augmented Reality |
| ARA | Augmented Reality Audio |
| DSP | Digital Signal Processor |
| FPGA | Field Programmable Gate Array |
| GPS | Global Positioning System |
| HAW | Hochschule für Angewandte Wissenschaften |
| HRIR | Head Related Impulse Response (kopfbezogene Impulsantwort) |
| HRTF | Head Related Transfer Function (kopfbezogene Transferfunktion) |
| ILD | Interaural-Level-Difference (Pegeldifferenz) |
| IP-Core | Intellectual Property Core |
| ITD | Interaural-Time-Difference (Laufzeitdifferenz) |
| JAES | Journal of the Audio Engineering Society |
| MARA | Mobile Augmented Reality Audio |
| OSC | Open Sound Control |
| RTP | Real-Time Protocol |
| SoC | System on Chip |
| VR | Virtual Reality |
| WLAN | Wireless Local Area Network |

Literatur

- [HJT⁺04] HÄRMÄ, Aki ; JAKKA, Julia ; TIKANDER, Miikka ; KARJALAINEN, Matti ; LOKKI, Tapio ; HIIPAKKA, Jarmo ; LORHO, Gaetan: *Augmented Reality Audio for Mobile and Wearable Appliances**. J. Audio Eng. Soc., Vol. 52, No. 6, 2004
- [Huo99] HUOPANIEMI, Jyri: *Virtual Acoustics and 3-D Sound in Multimedia Signal Processing*. Dissertation, Helsinki University of Technology, 1999
- [Ins11] INSTRUMENTS, Texas: *Spatial Audio Technology*. 2011
- [Pel09] PELTOLA, Mikko: *Augmented Reality Audio Applications in Outdoor Use*. Master's Thesis, Helsinki University of Technology, 2009
- [SWC⁺03] SUNDARESWARAN, V. ; WANG, Kenneth ; CHEN, Steven ; BEHRINGER, Reinhold ; MCGEE, Joshua ; TAM, Clement ; ZAHORIK, Pavel: *3D Audio Augmented Reality: Implementation and Experiments*. ACM International Symposium on Mixed and Augmented Realty (ISMAR '03), 2003
- [Tik09] TIKANDER, Miikka: *DEVELOPMENT AND EVALUATION OF AUGMENTED REALITY AUDIO SYSTEMS*. Helsinki University of Technology Dissertation, 2009
- [VAOB11] VAZQUEZ-ALVAREZ, Yolanda ; OAKLEY, Ian ; BREWSTER, Stephen A.: *Auditory display design for exploration in mobile audio-augmented reality*. Springer-Verlag, 2011