

(19)



(11)

**EP 3 149 969 B1**

(12)

**EUROPÄISCHE PATENTSCHRIFT**

(45) Veröffentlichungstag und Bekanntmachung des Hinweises auf die Patenterteilung:  
**18.09.2019 Patentblatt 2019/38**

(51) Int Cl.:  
**H04S 7/00 (2006.01)**

(21) Anmeldenummer: **15724972.3**

(86) Internationale Anmeldenummer:  
**PCT/EP2015/060792**

(22) Anmeldetag: **15.05.2015**

(87) Internationale Veröffentlichungsnummer:  
**WO 2015/180973 (03.12.2015 Gazette 2015/48)**

(54) **ERMITTLUNG UND NUTZUNG HÖRRAUMOPTIMIERTER ÜBERTRAGUNGSFUNKTIONEN**

DETERMINATION AND USE OF AUDITORY-SPACE-OPTIMIZED TRANSFER FUNCTIONS

DÉTERMINATION ET UTILISATION DE FONCTIONS DE TRANSFERT ACOUSTIQUEMENT OPTIMISÉES

(84) Benannte Vertragsstaaten:  
**AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM TR**

(72) Erfinder:  
• **BRANDENBURG, Karlheinz**  
98693 Ilmenau (DE)  
• **WERNER, Stephan**  
98693 Ilmenau (DE)  
• **SLADECZEK, Christoph**  
98693 Ilmenau (DE)

(30) Priorität: **28.05.2014 DE 102014210215**

(43) Veröffentlichungstag der Anmeldung:  
**05.04.2017 Patentblatt 2017/14**

(74) Vertreter: **Zinkler, Franz et al**  
**Schoppe, Zimmermann, Stöckeler**  
**Zinkler, Schenk & Partner mbB**  
Patentanwälte  
**Radtkoferstrasse 2**  
**81373 München (DE)**

(73) Patentinhaber:  
• **Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung e.V.**  
80686 München (DE)  
• **Technische Universität Ilmenau**  
98693 Ilmenau (DE)

(56) Entgegenhaltungen:  
**US-A1- 2006 045 294 US-A1- 2008 273 708**  
**US-A1- 2013 272 527**

**EP 3 149 969 B1**

Anmerkung: Innerhalb von neun Monaten nach Bekanntmachung des Hinweises auf die Erteilung des europäischen Patents im Europäischen Patentblatt kann jedermann nach Maßgabe der Ausführungsordnung beim Europäischen Patentamt gegen dieses Patent Einspruch einlegen. Der Einspruch gilt erst als eingelegt, wenn die Einspruchsgebühr entrichtet worden ist. (Art. 99(1) Europäisches Patentübereinkommen).

## Beschreibung

**[0001]** Ausführungsbeispiele der vorliegenden Erfindung beziehen sich auf eine Vorrichtung zur Ermittlung von für einen Abhörraum "hörraumoptimierten Übertragungsfunktionen", auf ein entsprechendes Verfahren und auf eine Vorrichtung zur räumlichen Wiedergabe eines Audiosignals mit entsprechenden Verfahren. Entsprechend bevorzugter Ausführungsbeispiele erfolgt die Wiedergabe mittels eines binauralen Nahbereichsschallwandlers, wie z.B. mittels eines Stereokopfhörers oder Stereo-In-Ear-Hörers. Weitere Ausführungsbeispiele beziehen sich auf ein System, umfassend die zwei Vorrichtungen, und auf ein Computerverfahren zur Durchführung der angesprochenen Verfahren.

**[0002]** Die perzeptive Qualität bei der Präsentation einer räumlichen auditiven Szene, z.B. auf Basis eines Mehrkanal-Audiosignals, hängt entscheidend von der akustisch künstlerischen Gestaltung des Inhalts der Präsentation, vom Wiedergabesystem und von der Raumakustik des Abhörraums bzw. Hörraums ab. Ein Hauptziel bei der Entwicklung von Audiowiedergabesystemen ist die Erzeugung von Hörereignissen, die vom Hörer als plausibel eingeschätzt werden. Dies spielt beispielsweise bei der Wiedergabe von Bild-Ton-Inhalten eine besondere Rolle. Bei von dem Nutzer als plausibel wahrgenommenen Inhalten müssen verschiedene perzeptive Qualitätsmerkmale, wie beispielsweise Lokalisierbarkeit, Distanzwahrnehmung, Räumlichkeitswahrnehmung und klangliche Aspekte der Abbildung den Erwartungen entsprechen. Im Idealfall stimmt somit die Wahrnehmung des wiedergegebenen mit der realen Situation im Raum überein.

**[0003]** Bei lautsprecherbasierten Audiowiedergabesystemen wird zwei- oder mehrkanaliges Audiomaterial im Abhörraum wiedergegeben. Dieses Audiomaterial kann einer kanalbasierten Mischung entstammen, bei der die fertigen Lautsprechersignale bereits vorliegen. Darüber hinaus können die Lautsprechersignale auch durch ein objektbasiertes Tonwiedergabeverfahren erzeugt werden. Hierbei werden basierend auf einer Beschreibung eines Tonobjekts (z.B.: Position, Lautstärke etc.) und der Kenntnis eines vorherrschenden Lautsprecheraufbaus, die Lautsprecherwiedergabesignale erzeugt. Dabei werden Phantomschallquellen erzeugt, die sich in der Regel auf den Verbindungsachsen zwischen den Lautsprechern befinden. Je nach gewähltem Lautsprecheraufbau und der vorherrschenden Raumakustik des Abhörraums können diese Phantomschallquellen vom Hörer in unterschiedlichen Richtungen und Distanzen wahrgenommen werden. Die Raumakustik selbst hat dabei einen entscheidenden Einfluss auf den Wohlklang der wiedergegebenen auditiven Szene.

**[0004]** Die Wiedergabe über Lautsprechersysteme ist allerdings nicht in allen Hörsituationen praktikabel. Weiterhin ist es nicht überall möglich Lautsprecher zu installieren. Als Beispiele solcher Situationen können das Musikhören auf mobilen Endgeräten, die Nutzung in sich

wechselnden Räumen, Nutzerakzeptanz oder auch die akustische Belästigung von Mitmenschen genannt werden. Als Alternative für Lautsprecher werden häufig Nahbereichsschallwandler, wie z.B. In-Ears oder Kopfhörer, die direkt an bzw. in unmittelbarer Umgebung zu dem Ohr "getragen" werden, eingesetzt.

**[0005]** Die klassische Stereowiedergabe über Schallwandler, die beispielsweise je mit einem akustischen Treiber pro Seite bzw. Ohr ausgestattet sind, erzeugen beim Hörer die Wahrnehmung, dass die sich abbildenden Phantomschallquellen im Kopf auf der Verbindungsachse zwischen den beiden Ohren befinden. Es kommt zur sogenannten "Im-Kopf-Lokalisierung". Eine plausibel wirkende externe Wahrnehmung (Externizität) der Phantomschallquellen kommt nicht zustande. Die so erzeugten Phantomschallquellen besitzen in der Regel weder ein für einen Nutzer dekodierbare Richtung(-sinformation) noch Distanz(-sinformation), die beispielsweise bei der Wiedergabe der gleichen akustischen Szene über ein Lautsprechersystem (z.B. 2.0 oder 5.1) im Abhörraum vorhanden wäre.

**[0006]** Zur Umgehung der Im-Kopf-Lokalisierung bei Kopfhörerwiedergabe werden Methoden der Binauralsynthese verwendet (ohne dass künstlerische Gestaltung und Mischung im Audiomaterial verloren geht). Bei der Binauralsynthese werden sogenannte "Außenohrübertragungsfunktionen" (head-related transfer function, HRTF, Kopfbezogene Übertragungsfunktion) für linkes und rechtes Ohr verwendet. Diese Außenohrübertragungsfunktionen umfassen für jedes Ohr eine Vielzahl an jeweiligen Richtungsvektoren für virtuelle Schallquellen zugeordnete Außenohrübertragungsfunktionen, entsprechend welchen eine Filterung der Audiosignale bei der Wiedergabe derselben erfolgt, so dass eine auditive Szene räumlich dargestellt bzw. die Räumlichkeit emuliert werden. Die Binauralsynthese macht es sich zunutze, dass interaurale Merkmale maßgeblich für das Zustandekommen einer Richtungswahrnehmung einer Schallquelle verantwortlich sind, wobei sich diese interauralen Merkmale in den Außenohrübertragungsfunktionen widerspiegeln. Soll also ein Audiosignal aus einer definierten Richtung wahrgenommen werden, wird dieses Signal mit den zu dieser Richtung gehörenden HRTFs des linken oder rechten Ohrs gefiltert. Mit Hilfe der Binauralsynthese ist es somit möglich, sowohl eine realistische Raumklangszene, z.B. gespeichert als Mehrkanalaudio über den Kopfhörer wiederzugeben. Um einen Lautsprecheraufbau virtuell zu simulieren, verwendet man die richtungsgebundenen HRTF-Paare für jeden zu simulierenden Lautsprecher. Für eine plausible Abbildung von Richtung und Distanz des Lautsprecheraufbaus müssen zusätzlich die richtungsabhängigen akustischen Übertragungsfunktionen des Abhörraums (room-related-transfer-function, RRTF, Raumbezogene Übertragungsfunktion) mit emuliert werden. Diese werden mit den HRTFs kombiniert und ergeben die binauralen Raumimpulsantworten (BRIRs, binaural room-impuls-respons). Die BRIRs können als Filter auf das akus-

tische Signal angewendet werden.

**[0007]** Aktuelle Forschungen und Untersuchungen machen allerdings deutlich, dass die Plausibilität einer Audiowiedergabe neben der physikalischen korrekten Synthese der Wiedergabesignale auch maßgeblich von kontextabhängigen Qualitätsparametern und insbesondere von dem Erwartungshorizont des Nutzers in Bezug auf die Raumakustik bestimmt wird. Deshalb besteht der Bedarf nach einem verbesserten Ansatz bei der binauralen Synthese.

**[0008]** Die US 2013/0272527 A1 beschreibt ein Audiosystem mit einem Empfänger zum Empfangen eines Audiosignals in einem sogenannten "Binauralschaltkreis zum Erzeugen eines Binauralsignals unter Zuhilfenahme welches eine virtuelle Schallquelle im Raum positioniert werden kann". Hierbei ist eine Anpassung der Binauraltransferfunktion in Abhängigkeit an die, akustischen Umweltparameter möglich, so dass der Klang sehr natürlich erscheint.

**[0009]** Die US 2008/0273708 A1 zeigt wie mittels HRTF-Bearbeitung Soundsignale genutzt werden können, um frühere Reflexionen zu simulieren.

**[0010]** Aufgabe der vorliegenden Erfindung liegt darin, eine verbesserte räumliche Wiedergabe mittels Nahbereichsschallwandlern, insbesondere in Bezug auf Übereinstimmung von akustischer Synthetisierung und Erwartungshorizont des Konsumenten, zu schaffen.

**[0011]** Die Aufgabe wird durch die unabhängigen Patentansprüche gelöst.

**[0012]** Ausführungsbeispiele der vorliegenden Erfindung schaffen eine (tragbare) Vorrichtung zur Ermittlung von für einen Abhörraum "hörraumoptimierten Übertragungsfunktionen" auf Basis einer Analyse der Raumakustik. Die hörraumoptimierten Übertragungsfunktionen dient zur hörraumoptimierten Nachbearbeitung von Audiosignalen bei der räumlichen Wiedergabe, wobei basierend auf den Außenohrübertragungsfunktionen (HRTFs) ein zu synthetisierender Raum emulierbar ist und wobei basierend auf den hörraumoptimierten Übertragungsfunktionen der Abhörraum emulierbar ist. Durch die Nutzung dieser zwei Übertragungsfunktionen, die in kombinierter Form auch als binaurale raumbezogene Raumimpulsantwort bezeichnet werden kann, kommt es zu einer realistischen Raumklangsimitation, die in Bezug auf die Räumlichkeit mit den durch das Mehrkanal(-Stereo)-Signal vorgegebenen Merkmalen entspricht, aber unter Berücksichtigung des Erwartungshorizonts, welcher insbesondere durch die Raumakustik vorweggenommen ist, verbessert ist.

**[0013]** Entsprechend weiteren Ausführungsbeispielen schafft die vorliegende Erfindung eine weitere (tragbare) Vorrichtung zur räumlichen Wiedergabe eines Audiosignals mittels eines binauralen Nahbereichsschallwandlers, bei der die räumliche Wiedergabe unter Zuhilfenahme von bekannten Außenohrübertragungsfunktionen und unter Zuhilfenahme von für einen hörraumoptimierten Übertragungsfunktionen emuliert wird, so dass bei der Wiedergabe von Audioinhalten den mittels des Nah-

bereichsschallwandlers ausgehenden akustischen Signalen die Abhörraumcharakteristik aufgeprägt wird.

**[0014]** Entsprechend dem Kerngedanken schafft die vorliegende Erfindung also die Voraussetzungen, um kognitive Effekte bei der Wiedergabe von Mehrkanal-Stereo zu berücksichtigen. Hierzu werden entsprechend einem ersten Aspekt hörraumoptimierte Übertragungsfunktionen für den jeweiligen Abhörraum, in welchem beispielsweise mittels eines Kopfhörers (allgemein mittels eines binauralen Nahbereichsschallwandlers) eine auditive Szene wiedergegeben werden soll, ermittelt. Das Ermitteln der hörraumoptimierten Übertragungsfunktion entspricht im Prinzip der Ableitung eines raumakustischen Filters auf Basis der ermittelten bzw. vermessenen Raumakustik mit der Zielsetzung, die akustischen Eigenschaften des Realraums synthetisch abzubilden. In einem zweiten Schritt kann dann die auditive Szene entsprechend einem zweiten Erfindungsaspekt sowohl unter Zuhilfenahme der HRTFs als auch unter Zuhilfenahme der hörraumoptimierten Übertragungsfunktionen als Raumklangsimitation wiedergegeben werden. Bei der Wiedergabe wird mittels HRTFs die Räumlichkeit erzeugt, während mittels hörraumoptimierter Übertragungsfunktionen die Anpassung der Räumlichkeit an die aktuelle Hörraumsituation erreicht wird. In anderen Worten heißt es, dass die hörraumoptimierten Übertragungsfunktionen eine Anpassung bzw. Nachbearbeitung der HRTFs oder der durch die HRTFs bearbeiteten Signale vornimmt. Im Ergebnis kann so bei der Wiedergabe von Audioinhalten die Divergenz zwischen dem wiederzugebenden Raum, definiert durch das Mehrkanalaudiomaterial, und dem Abhörraum, in dem sich der Hörer befindet, reduziert werden.

**[0015]** Für die Ermittlung der hörraumoptimierten Übertragungsfunktionen gibt es unterschiedliche Möglichkeiten, nämlich entsprechend einer ersten Variante die messtechnische Ermittlung unter Zuhilfenahme einer Testschallquelle und eines Mikrofons, so dass die Raumakustik über eine Teststrecke in dem Abhörraum analysiert werden kann, um ein akustisches Modell des Raums zu erhalten. Entsprechend einer zweiten Variante können auch natürlich vorkommende Geräusche, wie z.B. eine Stimme, als Testsignale genutzt werden. Die zweite Variante bietet insbesondere den Vorteil, dass praktisch jedes elektrische Endgerät mit einem Mikrofon, wie z.B. ein Mobiltelefon oder ein Smartphone, auf dem die oben beschriebene Funktionalität implementiert ist, ausreicht, um die Raumakustik zu ermitteln. Entsprechend einer dritten Variante kann die Analyse des Abhörraums bzw. die Ermittlung des akustischen Raummodells auf Basis von geometrischen Modellen erfolgen. Denkbar wäre es in diesem Zusammenhang auch, dass ein geometrisches Modell optisch, z.B. mit einer Kamera, die typischerweise auch in mobilen Endgeräten (wie Mobiltelefonen) integriert ist, erfasst wird, um so im Nachgang das akustische Modell des Abhörraums zu errechnen. Ausgehend von einem so ermittelten akustischen Raummodell können nun die hörraumoptimierten Übertragungs-

funktionen bestimmt werden.

**[0016]** Entsprechend weiteren Ausführungsbeispielen kann nicht nur dem Abhörraum alleine Rechnung getragen werden, sondern auch einer Positionierung des Hörers in dem Abhörraum. Hintergrund hierzu ist, dass sich die Raumakustik bzw. akustische Wahrnehmung entsprechend ändert, je nachdem ob die Hörposition sich näher an der Wand befindet oder in welche Richtung der Hörer schaut. Somit kann entsprechend weiteren Ausführungsbeispielen eine Vielzahl an richtungsabhängigen und/oder positionsabhängigen Übertragungsfunktionen (Übertragungsfunktionen-Scharen) innerhalb der hörraumoptimierten Übertragungsfunktionen hinterlegt sein, die hier beispielsweise in Abhängigkeit von der Position des Hörers in dem Abhörraum oder von dem Blickwinkel des Hörers ausgewählt werden.

**[0017]** Auch in Bezug auf die hörraumoptimierte Übertragungsfunktionen ist es vorteilhaft, wenn in der Vorrichtung zur räumlichen Wiedergabe oder in der an die Vorrichtung angekoppelten Datenbank eine Vielzahl von hörraumoptimierten Übertragungsfunktionen-Scharen für unterschiedliche Abhörräume hinterlegt sind, so dass diese in Abhängigkeit davon, in welchem Raum der Hörer sich gerade befindet, abrufbar sind. Hierzu kann beispielsweise die Vorrichtung zur räumlichen Wiedergabe auch eine Positionsbestimmungsvorrichtung, wie z.B. ein GPS umfassen.

**[0018]** Entsprechend weiteren Ausführungsbeispielen ist es auch möglich, neben der oder parallel zu der Abhörraumcharakteristik dem wiederzugebenden Audiomaterial die entsprechende Charakteristik eines virtuellen Lautsprecher setups aufzuprägen, welches beispielsweise dem realen Lautsprecher setup in dem Abhörraum entspricht oder frei konfiguriert ist.

**[0019]** Weitere Ausführungsbeispiele beziehen sich auf die entsprechenden Verfahren zur Ermittlung der hörraumoptimierten Übertragungsfunktionen und zur Wiedergabe von Mehrkanal-Stereo-Audiosignalen (oder objektbasierten Audiosignalen oder WFS-Audiosignalen) unter Nutzung der hörraumoptimierten Übertragungsfunktionen.

**[0020]** Nachfolgende Ausführungsbeispiele werden anhand der beiliegenden Figuren im Detail erläutert. Es zeigen:

Fig. 1a ein schematisches Blockschaltbild einer Vorrichtung zur Ermittlung von für einen Abhörraum hörraumoptimierten Übertragungsfunktionen;

Fig. 1b ein schematisches Flussdiagramm eines Verfahrens bei der Ermittlung hörraumoptimierter Übertragungsfunktionen;

Fig. 2a ein schematisches Blockschaltbild einer Vorrichtung zur räumlichen Wiedergabe von Mehrkanal-Stereo-Audiomaterial unter Berücksichtigung hörraumoptimierter Übertra-

gungsfunktionen;

Fig. 2b ein schematisches Flussdiagramm für ein Verfahren zur räumlichen Wiedergabe von Mehrkanal-Stereo-Audiomaterial unter Berücksichtigung hörraumoptimierter Übertragungsfunktionen; und

Fig. 3 ein schematisches Blockschaltbild eines Systems zur Ermittlung und Nutzung von hörraumoptimierten Übertragungsfunktionen.

**[0021]** Bevor nachfolgend Ausführungsbeispiele der vorliegenden Erfindung anhand der beiliegenden Zeichnungen näher erläutert werden, sei darauf hingewiesen, dass gleiche Elemente oder gleichwirkende Elemente mit gleichen Bezugszeichen versehen sind, so dass die Beschreibung derer aufeinander anwendbar bzw. austauschbar ist.

**[0022]** Im Vorfeld zu der Beschreibung der Erfindung wird nachfolgend auf die Motivation bei der Erfassung und Auralisation der Raumakustik eines Abhörraums zur standortabhängigen räumlichen Tonwiedergabe über Kopfhörer eingegangen. In diesem Zusammenhang wird auch kurz auf die Binauralsynthese eingegangen und ein Überblick über die für die Binauralsynthese genutzten Außenohrübertragungsfunktionen (HRTFs) und die in den Außenohrübertragungsfunktionen enthaltenen, manipulierbaren Variablen gegeben. Anhand dieses Überblicks wird ferner auch aufgezeigt, inwieweit die HRTFs durch die zu ermittelnden hörraumoptimierten Übertragungsfunktionen TF angepasst werden, um den Raumakustikbedingungen erfindungsgemäß Rechnung zu tragen.

**[0023]** Die Binauralsynthese basiert darauf, dass ein Audiosignal vor Ausgabe über einen Schallwandler (bevorzugt direkt an einem der Ohren) mit einer bestimmten Filterfunktion bzw. HRTF gefiltert wird, wobei die Filtercharakteristik sich je Richtungsvektor bzw. je virtueller Schallquelle unterscheidet, um so Raumklang, z.B. bei Nutzung eines Kopfhörers, zu emulieren. Die Filterfunktionen/HRTFs sind den natürliche Schalllokalisierungsmechanismen des menschlichen Gehörs nachempfunden. Hierdurch ist es möglich, das Audiosignal in der analogen oder digitalen Domäne so zu bearbeiten bzw. diesem eine akustische Charakteristik aufzuprägen, als ob dieses von einer beliebigen Position im Raum ausgesendet wird. Die Mechanismen bei der Lokalisierung von Schall sind:

- Erkennung der seitliche Einfallsrichtung;
- Erkennung der Einfallsrichtung in der medialen Ebene; und
- Erkennung der Entfernung.

**[0024]** Für die Lokalisierung in Bezug auf die seitliche Einfallsrichtung sind akustische Merkmale wie Laufzeitdifferenzen zwischen links/rechts und (frequenzabhän-

gige) Pegeldifferenzen zwischen links/rechts maßgeblich. Bei den Laufzeitunterschieden kann insbesondere zwischen Phasenlaufzeit bei niedrigen Frequenzen und Gruppenlaufzeit bei hohen Frequenzen unterschieden werden. Diese Laufzeitunterschiede können über einen beliebigen Stereo-Treiber via Signalverarbeitung nachgebildet werden. Die Bestimmung der Einfallrichtung in der medialen Ebene basiert insbesondere darauf, dass die Ohrmuschel und/oder dem Gehörgangseingang eine richtungsselektive Filterung des akustischen Signals durchführt. Diese Filterung ist frequenzselektiv, so dass ein Audiosignal vorab mit einem derartigen Frequenzfilter gefiltert werden kann, um eine bestimmte Einfallrichtung vorzutauschen bzw. eine Räumlichkeit zu emulieren. Die Bestimmung der Entfernung einer Schallquelle von dem Hörer basiert auf unterschiedlichen Mechanismen. Die Hauptmechanismen sind Lautstärke, frequenzselektive Filterung des zurückgelegten Schallwegs, Schallreflexion und Anfangszeitlücke. Ein Großteil der oben genannten Faktoren ist personenindividuell. Personenindividuelle Variable können z.B. der Ohrenabstand sowie die Form der Ohrmuschel sein, die sich insbesondere auf die seitliche und die mediale Lokalisierung auswirkt. Durch Manipulation eines Audiosignals in Hinblick auf die genannten Mechanismen erfolgt die Raumklangemulation, wobei die Manipulationsparameter (je Raumrichtung und Entfernung) in den HRTFs hinterlegt sind.

**[0025]** Diese HRTFs (Außenohrübertragungsfunktionen) sind in erster Linie für die Freifeldschallausbreitung gedacht. Hintergrund hierzu ist, dass die oben genannten drei Faktoren zur Lokalisierung für die Anwendung in geschlossenen Räumen insoweit verfälscht werden, dass der von einer Schallquelle ausgesandte Schall nicht nur direkt, sondern auch in reflektierter Form (z.B. über Wände) zu dem Hörer gelangt, was eine Veränderung der akustischen Wahrnehmung zu Folge hat. In Räumen kommt es also zu Direktschall und zu (später eintreffenden) reflektierten Schall, wobei diese Schallarten für den Hörer beispielsweise anhand von Laufzeit für bestimmte Frequenzgruppen und/oder Position der sekundären Schallquelle im Raum differenzierbar sind. Diese (Hall-)Parameter sind ferner von der Raumgröße und Beschaffenheit (z.B. Dämpfung, Form) abhängig, sodass ein Hörer die Raumgröße und Beschaffenheit abschätzen kann. Da diese Raumakustikparameter grundsätzlich über dieselben Mechanismen wahrgenommen werden wie die der Lokalisierung, ist die Raumakustik ebenfalls binaural emulierbar. Zur Emulation der Raumakustik wird die HRTF mittels der RRTF zu der binauralen Raumimpulsantwort (BRIR) erweitert, die dem Hörer im Falle der Kopfhörerwiedergabe bestimmte akustische Raumbedingungen simuliert. So erfolgt je nach virtueller Raumgröße eine Veränderung des Hallverhaltens, einer Verschiebung von Sekundärschallquellen, eine Veränderung der Lautheit der Sekundärschallquellen, insbesondere in Relation zu der Lautheit der Primärschallquellen.

**[0026]** Wie eingangs bereits erwähnt, spielen beim Hörer auch kognitive Effekte eine große Rolle. Untersuchungen zu derartigen kognitiven Effekten haben ergeben, dass die Relevanz von Parametern wie dem Übereinstimmungsgrad zwischen dem Abhörraum und dem zu synthetisierendem Raum, das Zustandekommen einer plausiblen auditiven Illusion hoch sind. Der Fachmann spricht im Falle einer geringen Divergenz zwischen Abhörraum und wiederzugebendem Raum von geringer Externizität des Hörereignisses.

**[0027]** Motiviert hierdurch soll nun die Binauralsynthese derart erweitert werden, dass die binaurale Simulation einer auditiven Szene an den Kontext der Nutzung angepasst werden kann. Im Detail erfolgt eine Anpassung der Simulation an die Abhörbedingungen, wie beispielsweise an eine aktuelle Raumakustik (Dämpfung) und die Geometrie des Abhörraums. Hierzu können die Distanzwahrnehmung, die Wahrnehmung von Räumlichkeit und die Richtungswahrnehmung so variiert werden, dass sie in Bezug auf den aktuellen Abhörraum plausibel erscheinen. Variationsparameter sind beispielsweise die HRTF- bzw. RRTF-Merkmale, wie z.B. Laufzeitdifferenzen, Pegeldifferenzen, frequenzselektive Filterung oder Anfangszeitlücke. Die Anpassung erfolgt beispielsweise in der Art, dass eine Raumgröße mit einem bestimmten Hallverhalten (Nachhallverhalten bzw. Reflexionsverhalten) emuliert wird oder Abstände z.B. zwischen Hörer und Schallquelle auf einen Maximalwert begrenzt werden. Weiterer Einflussfaktor auf das Raumklangverhalten sind die Position des Nutzers in dem Abhörraum, da es in Hinblick auf Hall und Reflexion entscheidend ist, ob der Nutzer zentral im Raum oder in der Nähe einer Wand steht. Auch dieses Verhalten ist durch Anpassung der HRTF- bzw. RRTF-Parameter emulierbar. Nachfolgend wird erläutert, wie bzw. mit welchen Mitteln die Anpassung der HRTF- bzw. RRTF-Parameter vorgenommen wird, um die Plausibilität der akustischen Simulation vor Ort zu verbessern.

**[0028]** Das Konzept zur Auralisation der Raumakustik umfasst in der grundlegenden Struktur zwei Bestandteile, die einerseits durch zwei unabhängige Vorrichtungen und andererseits durch zwei entsprechende Verfahren repräsentiert werden. Bezug nehmen auf Fig. 1a und 1b wird der erste Bestandteil, nämlich die Erfassung von hörraumoptimierten Übertragungsfunktionen TF erläutert, bevor Bezug nehmend auf Fig. 2a und 2b die Nutzung der hörraumoptimierten Übertragungsfunktionen TF erläutert werden.

**[0029]** Fig. 1a zeigt eine Vorrichtung 10 zur Ermittlung von für einen Abhörraum 12 optimierten Übertragungsfunktionen TF (Transfer Funktion). Zur Ermittlung der hörraumoptimierten Übertragungsfunktionen TF wird der Abhörraum 12 bzw. die Raumakustik desselben analysiert. Deshalb umfasst die Vorrichtung 10 eine Schnittstelle, z.B. wie hier illustriert eine Mikrofonschnittstelle (vgl. Bezugszeichen 14), zur Erfassung von hörraumbezogenen Daten. Da die hörraumoptimierte Übertragungsfunktionen TF, auf Basis welcher anschließend

mittels Binauralsynthese einem akustischen Material die Abhörraumcharakteristik aufgeprägt werden soll, typischerweise derart ausgelegt wird, dass bereits vorhandene HRTFs angepasst werden, kann die Vorrichtung 10 die Übertragungsfunktionen TF unter Berücksichtigung der zu nutzenden HRTFs ermitteln. Insofern umfasst die Vorrichtung 10 optionaler Weise eine weitere Schnittstelle zum Einlesen bzw. zum Weiterleiten von HRTFs.

**[0030]** Nachfolgend werden ausgehend von der Vorrichtung 10 unterschiedliche Vorgehensweisen für die Ermittlung der Raumakustik erläutert, auf Basis welcher dann in einem nachgelagerten Schritt die hörraumoptimierte Übertragungsfunktionen TF bestimmt werden. Entsprechend einer ersten Variante ist die Erfassung der vorherrschenden raumakustischen Bedingungen des Abhörraums messtechnisch möglich. Beispielsweise wird durch ein akustisches Messverfahren die Raumakustik des Abhörraums 12 unter Zuhilfenahme der Vorrichtung 10 vermessen. Dazu wird dann ein Testsignal, ausgesendet über einen optionalen Lautsprecher (nicht dargestellt), verwendet. Die Wiedergabe des Testsignals bzw. die Ansteuerung des Lautsprechers kann hierbei über die Vorrichtung 10 erfolgen, wenn die Vorrichtung 10 hierzu eine Lautsprecherschnittstelle (nicht dargestellt) oder den Lautsprecher selbst umfasst. Das über den Lautsprecher in den Raum 12 abgestrahlte Messsignal wird mittels des Mikrofons 14 aufgezeichnet, so dass ausgehend von der Signalveränderung über die Messstrecke (zwischen Lautsprechermikrofon) die Raumakustik bestimmt werden kann, so dass zumindest eine hörraumoptimierte Übertragungsfunktion TF z.B. für eine Raumrichtung oder eine Mehrzahl an hörraumoptimierten Übertragungsfunktionen TF ableitbar ist. Aus der gemessenen Übertragungsfunktion aus einer Richtung werden für den Hörraum relevante raumakustische Parameter abgeleitet. Diese werden genutzt, um die hörraumoptimierten Übertragungsfunktionen TF für die anderen benötigten Richtungen zu generieren. Hierfür können bspw. durch Stauchung und/oder Streckung von Bereichen der Impulsantwort (Übertragungsfunktion im Zeitbereich) die diskreten ersten Reflexionen an andere Raumrichtungen und Distanz der abzubildenden virtuellen Schallquellenpositionen angepasst werden. Die für die Richtungswahrnehmung relevanten Informationen liegen in den HRTFs vor. Um die hörraumoptimierten Übertragungsfunktionen TF für alle Raumrichtungen oder in einer sehr hohen Genauigkeit zu ermitteln, kann es entsprechend weiteren Ausführungsbeispielen vorteilhaft sein, die Analyse mittels des Testsignals für unterschiedliche Positionen von Mikrofon 14 und Lautsprecher in dem Abhörraum 12 zu wiederholen.

**[0031]** Entsprechend einer weiteren Variante kann die Ermittlung der Raumakustik unter Verwendung von akustischen Signalen, die bereits durch den Abhörraum 12 verhallt sind, geschätzt werden. Beispiele für derartige Signale sind die sowieso vorhandenen Umgebungsgeräusche, wie auch ein Sprachsignal eines Nutzers. Die

hierzu eingesetzten Algorithmen leiten sich aus Algorithmen zur Entfernung von Nachhall aus einem Sprachsignal ab. Hintergrund hierzu ist, dass typischerweise bei den Enthaltungsalgorithmen eine Schätzung über die auf dem zu enthaltenden Signal liegenden Raumübertragungsfunktion erfolgt. Bis dato werden diese Algorithmen dazu genutzt, um ein Filter zu bestimmen, welches angewendet auf das Originalsignal möglichst das unverhallte Signal zum Ergebnis hat. In der Anwendung bei der Analyse der Raumakustik wird die Filterfunktion nicht bestimmt, sondern nur ein Schätzverfahren dazu verwendet, die Eigenschaften des Abhörraums zu erkennen. Auch bei diesem Vorgehen kommt wiederum das Mikrofon 14, welches an die Vorrichtung 10 angekoppelt ist, zum Einsatz.

**[0032]** Entsprechend einer dritten Variante kann die Raumakustik basierend auf geometrischen Raumdaten simuliert werden. Dieses Vorgehen basiert darauf, dass geometrische Daten (z.B. Kantenabmessungen, freie Weglänge) eines Raumes 12 es ermöglichen, die Raumakustik abzuschätzen. Die Raumakustik des Raums 12 kann entweder direkt simuliert werden oder basierend auf raumakustischen Filterdatenbank, die akustische Vergleichsmodelle umfassen, näherungsweise bestimmt werden. In diesem Zusammenhang sind beispielsweise Verfahren wie das akustische Ray Tracing oder das Spiegelschallquellenverfahren in Verbindung mit einem Diffusschallmodell zu nennen. Die beiden genannten Verfahren setzen auf geometrischen Modellen des Abhörraums auf. Insofern muss die oben erläuterte Schnittstelle zur Erfassung von hörraumbezogenen Daten der Vorrichtung 10 nicht zwingendermaßen eine Mikrofonschnittstelle sein, sondern kann auch allgemein als Datenschnittstelle bezeichnet werden, die zum Einlesen von Geometriedaten dient. Ferner ist es auch möglich, dass weitere Daten über die Raumakustik hinaus mittels der Schnittstelle eingelesen werden, die beispielsweise Informationen über ein in dem Abhörraum vorhandenes Lautsprechersetup umfassen.

**[0033]** Zur Akquise der geometrischen Raumdaten sind mehrere Möglichkeiten denkbar: Entsprechend einer ersten Untervariante können die Daten aus einer Geometriedatenbank, wie z.B. Google Maps Inhouse entnommen werden. Diese Datenbanken umfassen typischerweise geometrische Modelle, wie z.B. Vektormodelle von Raumgeometrien, ausgehend von welchen in erster Linie die Abstände, aber auch Reflektionscharakteristika bestimmbar sind. Entsprechend einer weiteren Untervariante kann auch eine Bilddatenbank als Input genutzt werden, wobei dann im Nachgang mittels Bilderkennung die geometrischen Parameter in einem Zwischenschritt bestimmt werden. Entsprechend einer alternativen Untervariante wäre es auch möglich, anstatt Bildinformationen einer Bilddatenbank zu entnehmen, die Bildinformationen auch mittels einer Kamera oder allgemein eines optischen Sensors zu ermitteln, so dass direkt durch den Nutzer ein geometrisches Modell ermittelt werden kann. Ausgehend von der auf Basis von Bilddaten

ermittelten Raumgeometrie kann dann analog zum vorherigen Punkt die Raumakustik simuliert werden.

**[0034]** Mittels diesen so simulierten Raumakustikmodellen werden in einem nachgelagerten Schritt für zumindest einen, bevorzugterweise für eine Vielzahl von Räumen, die hörraumoptimierten Übertragungsfunktionen TF abgeleitet. Die Ableitung der hörraumoptimierten Übertragungsfunktionen TF, welche hinsichtlich ihrer Parameter mit den RRTFs vergleichbar ist, entspricht im Prinzip der Bestimmung einer Filterfunktion (je Raumrichtung), mittels welcher das akustische Verhalten im Raum, z.B. bei der Schallausbreitung in einer bestimmten Raumrichtung, nachbildbar ist. Die hörraumspezifischen Übertragungsfunktionen TF je Raum umfassen typischerweise eine Vielzahl an Übertragungsfunktionen, mittels welchen die Außenohrübertragungsfunktionen (zugeordnet zu einzelne Raumwinkeln) entsprechend (vergleichbar zu dem Vorgehen bei der Verarbeitung der Raumimpulsantwort) angepasst werden können. Die Anzahl an hörraumoptimierten Übertragungsfunktionen TF richtet sich deshalb typischerweise nach der Anzahl der Außenohrübertragungsfunktionen, die als Funktionscharakteristika vorkommen und eine Vielzahl, nämlich für links/rechts und für die relevanten Richtungen besteht. Die genaue Anzahl an Außenohrübertragungsfunktionen im HRTF-Modell richtet sich nach der gewünschten Raumaufauflösungsfähigkeit und kann aufgrund dessen, dass auch HRTF-Modelle existieren, bei denen eine Vielzahl der Richtungsvektoren mittels Interpolation ermittelt werden, erheblich variieren. Aus diesem Zusammenhang wird ersichtlich, warum es sinnvoll ist, dass die Vorrichtung zur Ermittlung der hörraumoptimierten Übertragungsfunktion TF das HRTF-Modell mit verwendet. In einem weiteren Schritt werden die ermittelten hörraumoptimierten Übertragungsfunktionen TF z.B. in einer raumakustischen Filterdatenbank, gespeichert.

**[0035]** Entsprechend einem weiteren Ausführungsbeispiel kann auch je Abhörraum auch eine Vielzahl an hörraumoptimierten Übertragungsfunktionen-Scharen (TF) ermittelt und gespeichert werden, wodurch Rechnung getragen wird, dass die Abhörraumfunktionen bzw. das akustische Verhalten im Abhörraum je nach Position des Hörers unterschiedlich ist. In anderen Worten ausgedrückt heißt es also, dass je (möglicher) Position des Nutzers in dem Abhörraum 12 eine eigene hörraumoptimierte Übertragungscharakteristik ermittelbar ist, wobei die Ermittlung derselben auf ein und demselben akustischen Modell des Abhörraums 12 basieren kann. Infolgedessen ist vorteilhafterweise die Analyse des Abhörraums nur einmal durchzuführen. Entsprechend einem weiteren Ausführungsbeispiel können auch je Raumrichtung in welche der Nutzer schaut unterschiedliche raumoptimierte Übertragungsfunktionen-Scharen (TF) ermittelt werden.

**[0036]** Die oben beschriebene Vorrichtung 10 kann unterschiedlich ausgeführt sein. Entsprechend bevorzugten Ausführungsbeispielen ist die Vorrichtung 10 als mobiles Gerät ausgeführt, wobei dann der Sensor 14, wie

z.B. das Mikrofon oder die Kamera, entsprechend integriert sein kann. D.h. also, dass sich weitere Ausführungsbeispiele auf eine Vorrichtung zur Bestimmung der hörraumoptimierten Übertragungsfunktion TF beziehen, die einerseits die Analyseeinheit 10 und andererseits ein Mikrofon und/oder eine Kamera umfassen. Hierbei kann die Analyseeinheit 10 beispielsweise als Hardware implementiert oder Software-basiert ausgeführt sein. Also umfassen Ausführungsbeispiele der Vorrichtung 10 eine interne oder via Cloudcomputing angebundene CPU oder eine andere Logik, die ausgebildet ist, um die Ermittlung von hörraumoptimierten Übertragungsfunktionen TF und/oder die Hörraumanalyse durchzuführen. Nachfolgend wird Bezug nehmend auf Fig. 1b das Verfahren bzw. insbesondere die Grundschriffe des Verfahrens, auf welchen der Algorithmus für Software-implementierte Ermittlung von hörraumoptimierten Übertragungsfunktionen TF basiert, erläutert.

**[0037]** Fig. 1b zeigt ein Flussdiagramm 100 des Verfahrens bei der Ermittlung der für einen hörraumoptimierten Übertragungsfunktionen TF. Das Verfahren 100 umfasst den zentralen Schritt 110 des Ermitteln der hörraumoptimierten Übertragungsfunktionen TF. Wie oben bereits erläutert, basiert der Schritt 110 auf der Analyse der Raumakustik 120 (vgl. Schritt 120 "Raumakustik analysieren") und optionaler Weise auch auf vorhandenen HRTF-Funktionen. Ausgehend von dem Schritt 110 kann ein weiterer optionaler Schritt, nämlich des Speicherns der Übertragungsfunktionen TF folgen. Dieser Schritt ist mit dem Bezugszeichen 130 versehen.

**[0038]** Entsprechend weiteren Ausführungsbeispielen wäre es bei dem im Rahmen der Fig. 1a und 1b erläuterten Ausführungsbeispielen auch denkbar, dass gleichsam mit dem Ermitteln der hörraumoptimierten Übertragungsfunktionen TF auch eine Bestimmung der Position des Abhörraums erfolgt, so dass der so erhaltene Datensatz über die Position direkt dem Abhörraum zugeordnet werden kann. Dies bietet den Vorteil, dass im Fall des späteren Abrufens der hörraumoptimierten Übertragungsfunktionen TF aus einer Datenbank eine Zuordnung des jeweiligen Datensatzes ausgehend von einer Positionsbestimmung möglich ist.

**[0039]** Nachfolgend wird Bezug nehmend auf Fig. 2a und 2b die Nutzung der eben ermittelten hörraumoptimierten Übertragungsfunktionen TF erläutert.

**[0040]** Fig. 2a zeigt eine Vorrichtung zur räumlichen Wiedergabe 20 unter Zuhilfenahme eines binauralen Nahbereichsschallwandlers 22. Die Funktionalität der Vorrichtung 20 wird unter anderem unter Zuhilfenahme des Flussdiagramms aus Fig. 2b, welches das Verfahren 200 der Wiedergabe illustriert, erläutert. Die Vorrichtung 20 ist dazu ausgebildet, das Audiosignal 24, wie z.B. ein Mehrkanal-Stereo-Audiosignal (oder ein objektbasiertes Audiosignal oder ein Audiosignal auf Basis eines Wellenfeldsynthese-Algorithmus (WFS)) wiederzugeben, und um gleichzeitig Raumklang zu emulieren (vgl. Schritt 210). Hierzu führt die Wiedergabevorrichtung 20 eine Bearbeitung des Audiosignals unter Zuhilfenahme von

HRTFs und unter Zuhilfenahme der hörraumoptimierten Übertragungsfunktionen TF durch.

**[0041]** Die Vorrichtung 20 kann einen HRTF-/TF-Speicher umfassen oder ist beispielsweise mit einer Datenbank verbunden, auf welche die HRTFs, wie auch die entsprechend obigen Verfahren ermittelten hörraumoptimierten Übertragungsfunktionen TF gespeichert sind. Entsprechend bevorzugten Ausführungsbeispielen erfolgt vor der Signalverarbeitung des Audiosignals ein Kombinieren (vgl. Schritt 220) der HRTF mit der TF bzw. ein Anpassen der HRTF auf Basis der TF. Das Ergebnis dieses Kombinierens ist eine mit der BRIR (Raumimpulsantwort) vergleichbare Übertragungsfunktion BRIR', mit welcher dann schlussendlich das Audiosignal 24 bearbeitet wird, um den Raumklang zu emulieren (vgl. Schritt 210). Diese Bearbeitung entspricht im Prinzip einem Anwenden eines BRIR'-basierten Filters auf das Audiosignal. Somit ist es also möglich, Binauralsynthese in Kombination mit der Verhallung der Audiosignale in Abhängigkeit von den in dem Abhörraum vorherrschenden akustischen Bedingungen durchzuführen, so dass bei der Wiedergabe ein hoher Grad an Übereinstimmung zwischen dem synthetisierten Raum und dem Abhörraum entsteht. Folglich stimmt der synthetisierte Raum (zumindest annähernd) mit dem Erwartungshorizont des Nutzers überein, was die Plausibilität der Szene erhöht.

**[0042]** Entsprechend Ausführungsbeispielen kann die Vorrichtung 20 auch die Positionsbestimmungseinheit, wie z.B. einem GPS-Empfänger, umfassen, mittels welcher die aktuelle Position des Hörers feststellbar ist. Ausgehend von der festgestellten Position kann nun der Abhörraum ermittelt werden und die dem Abhörraum zugeordnete hörraumoptimierte Übertragungsfunktionen TF geladen (und ggf. bei einem Raumwechsel aktualisiert) werden. Optional ist es auch möglich, mittels dieser Positionsbestimmungseinrichtung die Position des Hörers in dem Abhörraum zu ermitteln, um hier auch noch, sofern gespeichert, die Unterschiede in der Akustik in Abhängigkeit von der Position des Hörers in dem Raum darzustellen. Diese Positionsbestimmungseinheit kann entsprechend dritten Ausführungsbeispielen auch durch eine Orientierungsbestimmungseinheit erweitert werden, so dass auch die Blickrichtung des Hörers bestimmbar ist und die TFs in Abhängigkeit von der bestimmten Blickrichtung entsprechend geladen werden, um der richtungsabhängigen Abhörraumakustik gerecht zu werden.

**[0043]** Ausgehend von dieser Basisbetrachtung der zwei Einheiten 10 und 20 wird nun ein erweitertes Ausführungsbeispiel von Fig. 3 erläutert. Fig. 3 zeigt eine schematische Darstellung des Signalfusses beim Abhören von angepassten raumakustischen Simulationen zur Verwendung mit der Binauralsynthese ausgehend von einem System 10 + 20, welches die Vorrichtung zur Bestimmung der TFs und die Vorrichtung zur Wiedergabe der Audiosignale unter Nutzung der TFs umfasst.

**[0044]** Eine derartiges System 10 + 20 kann beispielsweise als mobiles Endgerät (z.B. als Smartphone) imp-

lementiert sein, auf dem auch die wiederzugebende Datei gespeichert ist. Das System 10 + 20 ist prinzipiell eine Kombination der Vorrichtung 10 aus Fig. 1a und der Vorrichtung 20 aus Fig. 1b, wobei die einzelnen Bestandteile zur funktionsorientierten Erläuterung anders untergliedert sind.

**[0045]** Das System 10 + 20 umfasst eine Funktionseinheit zur Auralisation des Abhörraums 20a und eine Funktionseinheit zur Binauralsynthese 20b. Ferner umfasst das System 10 + 20 einen Funktionsblock 10a zur Modellierung der Raumakustik und einen Funktionsblock 10b zur Modellierung des Übertragungsverhaltens. Die Modellierung der Raumakustik basiert wiederum auf einer Erfassung des Abhörraums, welche mit dem Funktionsblock 10c zur Erfassung der Raumakustik durchgeführt wird. Des Weiteren umfasst das System 10 + 20 in der dargestellten Ausführungsform zwei Speicher, nämlich einen zur Speicherung von Szenenpositionsdaten 30a und einen zur Speicherung von HRTF-Daten 30b. Nachfolgend wird ausgehend von dem Informationsfluss bei der Wiedergabe die Funktionalität des Systems 10 + 20 erläutert, wobei davon ausgegangen wird, dass der Abhörraum dem System 10 + 20 bekannt ist oder mittels eines Positionsbestimmungsverfahrens (vgl. oben) bereits bestimmt ist.

**[0046]** Bei der Wiedergabe von kanalbasierten bzw. objektbasierten Audiodaten 24 mit dem Kopfhörer 22 werden die Audiodaten in einem ersten Schritt der Signalverarbeitungseinheit 20a zugeführt, die die vorher modellierte Raumübertragungsfunktion TF auf das Signal 24 anwendet und dieses verhallt. Die Modellierung der Raumübertragungsfunktion TF erfolgt in einem Signalverarbeitungsblock 10a, wobei diese Modellierung durch das Modellierungsübertragungsverhalten (vgl. Funktionsblock 10b) überlagert sein kann, wie nachfolgend erläutert wird.

**[0047]** Dieser zweite (optionale) Funktionsblock 10b modelliert ein virtuelles Lautsprechersetup in dem jeweiligen Abhörraum. So kann also dem Nutzer ein akustisches Verhalten emuliert werden, als ob die abzuspielende Audiodatei auf einem bestimmten Lautsprecher-setup (2.0, 5.1, 9.2) wiedergegeben wird. Hierbei ist dann insbesondere die Lautsprecherposition fix mit dem Abhörraum verbunden und den jeweiligen Lautsprechern auch ein bestimmtes Übertragungsverhalten, z.B. definiert durch die Frequenzgang und Richtungscharakteristik oder unterschiedliche Pegelverhalten, zugeordnet. Hier ist es auch möglich, besondere Schallquellentypen, z.B. eine Spiegelschallquelle fix im Raum zu positionieren. Das Lautsprecher-setup wird auf Basis der Szenenpositionsdaten modelliert, die Information über die Position, den Abstand oder auch den Typ des virtuellen Lautsprechers umfassen. Diese Szenenpositionsdaten können einem real vorhandenen Lautsprecher-setup entsprechen, oder auf Basis eines virtuellen Lautsprecher-setups basieren und sind typischerweise durch den Nutzer individualisierbar.

**[0048]** Nach der Verhallung in der Auralisationsverar-

beitungseinheit 20a werden die verhallten Signale der Binauralsynthese 20b zugeführt, die durch einen Satz richtungsgebender HRTF-Filter (vgl. 30b) die Richtung der virtuellen Lautsprecher auf das dem Lautsprecher zugehörigen Audiomaterial aufprägen. Das Binauralsynthesystem kann, wie bereits oben erläutert, optionaler Weise die Kopfdrehung des Hörers auswerten. Das Resultat ist ein Kopfhörersignal, was durch eine entsprechende Entzerrung auf einen speziellen Kopfhörer angepasst werden kann, wobei sich das akustische Signal so verhält, als ob es mit einem spezifischen Lautsprecher-setup in dem jeweiligen Abhörraum abgegeben wurde.

**[0049]** Das System 10 + 20 kann beispielsweise als mobiles Endgerät oder aus Komponenten einer Heimkinoanlage ausgeführt sein. Im Allgemeinen sind Anwendungsgebiete die Wiedergabe von Musik und Entertainmentinhalten, wie z.B. Ton für Film oder Spielton über den binauralen Nahbereichsschallwandler.

**[0050]** An dieser Stelle sei angemerkt, dass entsprechend einem alternativen Ausführungsbeispiel die Vorrichtung 20 aus Fig. 2a ebenfalls so ausgelegt sein kann, auf Basis von Szenenpositionsdaten ein bestimmtes Lautsprecher-setup bzw. die Wiedergabe eines Audiosignals für ein bestimmtes Lautsprecher-setup zu emulieren. Entsprechend kann auch gemäß einem weiteren Ausführungsbeispiel die Vorrichtung 10 dazu ausgebildet sein, die Szenenpositionsdaten eines Lautsprecher-setups in dem Abhörraum 12 (z.B. über eine akustische Messung) zu ermitteln, so dass dieses Lautsprecher-setup mit der Vorrichtung 20 emuliert werden kann.

**[0051]** Obwohl manche Aspekte im Zusammenhang mit einer Vorrichtung beschrieben wurden, versteht es sich, dass diese Aspekte auch eine Beschreibung des entsprechenden Verfahrens darstellen, sodass ein Block oder ein Bauelement einer Vorrichtung auch als ein entsprechender Verfahrensschritt oder als ein Merkmal eines Verfahrensschrittes zu verstehen ist. Analog dazu stellen Aspekte, die im Zusammenhang mit einem oder als ein Verfahrensschritt beschrieben wurden, auch eine Beschreibung eines entsprechenden Blocks oder Details oder Merkmals einer entsprechenden Vorrichtung dar. Einige oder alle der Verfahrensschritte können durch einen Hardware-Apparat (oder unter Verwendung eines Hardware-Apparats), wie zum Beispiel einen Mikroprozessor, einen programmierbaren Computer oder eine elektronische Schaltung ausgeführt werden. Bei einigen Ausführungsbeispielen können einige oder mehrere der wichtigsten Verfahrensschritte durch einen solchen Apparat ausgeführt werden.

**[0052]** Ein erfindungsgemäß codiertes Signal, wie beispielsweise ein Audiosignal oder ein Videosignal oder ein Transportstromsignal, kann auf einem digitalen Speichermedium gespeichert sein oder kann auf einem Übertragungsmedium wie beispielsweise einem drahtlosen Übertragungsmedium oder einem verdrahteten Übertragungsmedium, z.B. dem Internet, übertragen werden

**[0053]** Das erfindungsgemäße kodierte Audiosignal

kann auf einem digitalen Speichermedium gespeichert sein, oder kann auf einem Übertragungsmedium, wie beispielsweise einem drahtlosen Übertragungsmedium oder einem drahtgebundenen Übertragungsmedium, wie beispielsweise dem Internet, übertragen werden.

**[0054]** Je nach bestimmten Implementierungsanforderungen können Ausführungsbeispiele der Erfindung in Hardware oder in Software implementiert sein. Die Implementierung kann unter Verwendung eines digitalen Speichermediums, beispielsweise einer Floppy-Disk, einer DVD, einer Blu-ray Disc, einer CD, eines ROM, eines PROM, eines EPROM, eines EEPROM oder eines FLASH-Speichers, einer Festplatte oder eines anderen magnetischen oder optischen Speichers durchgeführt werden, auf dem elektronisch lesbare Steuersignale gespeichert sind, die mit einem programmierbaren Computersystem derart zusammenwirken können oder zusammenwirken, dass das jeweilige Verfahren durchgeführt wird. Deshalb kann das digitale Speichermedium computerlesbar sein.

**[0055]** Manche Ausführungsbeispiele gemäß der Erfindung umfassen also einen Datenträger, der elektronisch lesbare Steuersignale aufweist, die in der Lage sind, mit einem programmierbaren Computersystem derart zusammenzuwirken, dass eines der hierin beschriebenen Verfahren durchgeführt wird.

**[0056]** Allgemein können Ausführungsbeispiele der vorliegenden Erfindung als Computerprogrammprodukt mit einem Programmcode implementiert sein, wobei der Programmcode dahin gehend wirksam ist, eines der Verfahren durchzuführen, wenn das Computerprogrammprodukt auf einem Computer abläuft.

**[0057]** Der Programmcode kann beispielsweise auch auf einem maschinenlesbaren Träger gespeichert sein.

**[0058]** Andere Ausführungsbeispiele umfassen das Computerprogramm zum Durchführen eines der hierin beschriebenen Verfahren, wobei das Computerprogramm auf einem maschinenlesbaren Träger gespeichert ist.

**[0059]** Mit anderen Worten ist ein Ausführungsbeispiel des erfindungsgemäßen Verfahrens somit ein Computerprogramm, das einen Programmcode zum Durchführen eines der hierin beschriebenen Verfahren aufweist, wenn das Computerprogramm auf einem Computer abläuft.

**[0060]** Ein weiteres Ausführungsbeispiel der erfindungsgemäßen Verfahren ist somit ein Datenträger (oder ein digitales Speichermedium oder ein computerlesbares Medium), auf dem das Computerprogramm zum Durchführen eines der hierin beschriebenen Verfahren aufgezeichnet ist.

**[0061]** Ein weiteres Ausführungsbeispiel des erfindungsgemäßen Verfahrens ist somit ein Datenstrom oder eine Sequenz von Signalen, der bzw. die das Computerprogramm zum Durchführen eines der hierin beschriebenen Verfahren darstellt bzw. darstellen. Der Datenstrom oder die Sequenz von Signalen kann bzw. können beispielsweise dahin gehend konfiguriert sein, über

eine Datenkommunikationsverbindung, beispielsweise über das Internet, transferiert zu werden.

**[0062]** Ein weiteres Ausführungsbeispiel umfasst eine Verarbeitungseinrichtung, beispielsweise einen Computer oder ein programmierbares Logikbauelement, die dahin gehend konfiguriert oder angepasst ist, eines der hierin beschriebenen Verfahren durchzuführen.

**[0063]** Ein weiteres Ausführungsbeispiel umfasst einen Computer, auf dem das Computerprogramm zum Durchführen eines der hierin beschriebenen Verfahren installiert ist.

**[0064]** Ein weiteres Ausführungsbeispiel gemäß der Erfindung umfasst eine Vorrichtung oder ein System, die bzw. das ausgelegt ist, um ein Computerprogramm zur Durchführung zumindest eines der hierin beschriebenen Verfahren zu einem Empfänger zu übertragen. Die Übertragung kann beispielsweise elektronisch oder optisch erfolgen. Der Empfänger kann beispielsweise ein Computer, ein Mobilgerät, ein Speichergerät oder eine ähnliche Vorrichtung sein. Die Vorrichtung oder das System kann beispielsweise einen Datei-Server zur Übertragung des Computerprogramms zu dem Empfänger umfassen.

**[0065]** Bei manchen Ausführungsbeispielen kann ein programmierbares Logikbauelement (beispielsweise ein feldprogrammierbares Gatterarray, ein FPGA) dazu verwendet werden, manche oder alle Funktionalitäten der hierin beschriebenen Verfahren durchzuführen. Bei manchen Ausführungsbeispielen kann ein feldprogrammierbares Gatterarray mit einem Mikroprozessor zusammenwirken, um eines der hierin beschriebenen Verfahren durchzuführen. Allgemein werden die Verfahren bei einigen Ausführungsbeispielen seitens einer beliebigen Hardwarevorrichtung durchgeführt. Diese kann eine universell einsetzbare Hardware wie ein Computerprozessor (CPU) sein oder für das Verfahren spezifische Hardware, wie beispielsweise ein ASIC.

**[0066]** Die oben beschriebenen Ausführungsbeispiele stellen lediglich eine Veranschaulichung der Prinzipien der vorliegenden Erfindung dar. Es versteht sich, dass Modifikationen und Variationen der hierin beschriebenen Anordnungen und Einzelheiten anderen Fachleuten einleuchten werden. Deshalb ist beabsichtigt, dass die Erfindung lediglich durch den Schutzzumfang der nachstehenden Patentansprüche und nicht durch die spezifischen Einzelheiten, die anhand der Beschreibung und der Erläuterung der Ausführungsbeispiele hierin präsentiert wurden, beschränkt sei.

## Patentansprüche

1. Vorrichtung (10) zur Ermittlung von für einen Abhörraum (12) hörraumoptimierten Übertragungsfunktionen (TF), die für den Abhörraum (12) abgeleitet sind und welche zur hörraumoptimierten Nachbearbeitung von Audiosignalen (24) bei der räumlichen Wiedergabe dienen, wobei die räumliche Wiedergabe der Audiosignale (24) mittels eines binauralen

Nahbereichsschallwandlers (22) unter Zuhilfenahme von bekannten Außenohrübertragungsfunktionen (HRTF) und unter Zuhilfenahme der hörraumoptimierten Übertragungsfunktionen (TF) emuliert wird,

wobei basierend auf den Außenohrübertragungsfunktionen (HRTF) ein zu synthetisierender Raum emulierbar ist und wobei basierend auf den hörraumoptimierten Übertragungsfunktionen (TF) der Abhörraum (12) emulierbar ist,

wobei die Vorrichtung (10) ausgebildet ist, um eine Raumakustik des Abhörraums (12) zu analysieren und um ausgehend von der Analyse der Raumakustik die hörraumoptimierten Übertragungsfunktionen (TF) für den Abhörraum (12), in welchem die räumliche Wiedergabe mittels des binauralen Nahbereichsschallwandlers (22) erfolgen soll, zu ermitteln, wobei die Vorrichtung (10) einen Speicher umfasst, in welchem eine Mehrzahl an hörraumoptimierten Übertragungsfunktionen-Scharen (TF) für eine Mehrzahl an Abhörräumen (12) hinterlegbar ist,

**dadurch gekennzeichnet, dass** die hörraumoptimierten Übertragungsfunktionen (TF) je Raum eine Vielzahl an Übertragungsfunktionen, zugeordnet zu einzelnen Raumwinkeln, umfassen, wobei jeder Raumwinkel eine Schallausbreitungsrichtung im Raum darstellt.

2. Vorrichtung (10) gemäß Anspruch 1, wobei die Vorrichtung (10) ein Mikrophon (14) einer tragbaren Vorrichtung zur akustischen Messung umfasst und/oder wobei die Analyse der Raumakustik des Abhörraums (12) mittels akustischer Messung in dem Abhörraum (12) unter Zuhilfenahme von Umgebungsgerauschen und/oder unter Zuhilfenahme von einem Testsignal erfolgt.

3. Vorrichtung (10) gemäß Anspruch 1, wobei die Analyse der Raumakustik des Abhörraums (12) auf Berechnung eines geometrischen Modells des Abhörraums (12) und/oder auf Modellierung des geometrischen Modells basierend auf einem kamerabasierten Modell des Abhörraums (12) basiert.

4. Vorrichtung (10) gemäß Anspruch 2 oder 3, wobei die hörraumoptimierten Übertragungsfunktionen (TF) so gewählt sind, dass auf Basis dieser eine Raumakustik des Abhörraums (12) emulierbar ist.

5. Vorrichtung (10) gemäß einem der Ansprüche 1 bis 4, wobei die Vorrichtung (10) ausgebildet ist, die hörraumoptimierten Übertragungsfunktionen (TF) unter Berücksichtigung eines virtuellen Lautsprechersextens zu ermitteln, entsprechend welchem eine Anzahl virtueller Lautsprecher in dem Abhörraum (12) positioniert ist.

6. Vorrichtung (10) gemäß einem der Ansprüche 1 bis

- 5, wobei die bekannten Außenohrübertragungsfunktionen (HRTF) eine Vielzahl an einzelnen Übertragungsfunktionen (TF) für das linke und das rechte Ohr umfassen, die richtungsgebundenen Vektoren für eine Vielzahl von virtuellen Schallquellen zugeordnet sind.
7. Vorrichtung (10) gemäß einem der Ansprüche 1 bis 6, wobei das Emulieren der räumlichen Wiedergabe auf interauralen Merkmalen, Balance-Merkmalen und Abstands-Merkmalen basiert, wobei die interauralen Merkmale einen Zusammenhang zwischen einer Einfallrichtung in der medialen Ebene und einer individuellen oder nicht-individuellen Außenohrfilterung umfassen, wobei die Balance-Merkmale einen Zusammenhang zwischen einer seitlichen Einfallrichtung und einer Lautstärkedifferenz und/oder einen Zusammenhang zwischen der seitlichen Einfallrichtung und einer Laufzeitdifferenz umfassen, wobei die Entfernungs-Merkmale einen Zusammenhang zwischen einer virtuellen Entfernung und einer frequenzabhängigen Filterung und/oder einen Zusammenhang zwischen der virtuellen Entfernung und einer Anfangszeitlücke und/oder einen Zusammenhang zwischen der virtuellen Entfernung und einem Reflexionsverhalten umfassen.
8. Vorrichtung (10) gemäß einem der Ansprüche 1 bis 7, wobei die binaurale Nahbereichsschallwandler (22) ein Kopfhörer ist, der ausgebildet ist, ein Mehrkanal-Stereo-Signal, ein objektbasiertes Audiosignal (24) und/oder ein Audiosignal (24) auf Basis eines Wellenfeldsynthese-Algorithmus als Audiosignal (24) auszugeben.
9. Verfahren (100) zur Ermittlung von für einen Abhörraum (12) hörraumoptimierten Übertragungsfunktionen (TF), die für den Abhörraum (12) abgeleitet sind und welche zur hörraumoptimierten Nachbearbeitung von Audiosignalen (24) bei der räumlichen Wiedergabe dienen können, wobei die räumliche Wiedergabe der Audiosignale (24) mittels eines binauralen Nahbereichsschallwandlers (22) unter Zuhilfenahme von bekannten Außenohrübertragungsfunktionen (HRTF) und unter Zuhilfenahme der hörraumoptimierten Übertragungsfunktionen (TF) emuliert wird, wobei basierend auf den Außenohrübertragungsfunktionen (HRTF) ein zu synthetisierender Raum emulierbar ist und wobei basierend auf den hörraumoptimierten Übertragungsfunktionen (TF) der Abhörraum (12) emulierbar ist, mit den Schritten:
- Analysieren (120) einer vorherrschenden Raumakustik des Abhörraums (12); und Ermitteln (110) der hörraumoptimierten Übertragungsfunktionen (TF) für den Abhörraum (12), in welchem die räumliche Wiedergabe mittels des binauralen Nahbereichsschallwandlers (22) erfolgen soll, auf Basis der Analyse der Raumakustik;
- Hinterlegen einer Mehrzahl an hörraumoptimierten Übertragungsfunktionen-Scharen (TF) für eine Mehrzahl an Abhörräumen (12) in einen Speicher, **dadurch gekennzeichnet, dass** die hörraumoptimierten Übertragungsfunktionen (TF) je Raum eine Vielzahl an Übertragungsfunktionen, zugeordnet zu einzelnen Raumwinkeln, umfassen, wobei jeder Raumwinkel eine Schallausbreitungsrichtung im Raum darstellt.
10. Vorrichtung (20) zur räumlichen Wiedergabe eines Audiosignals (24) mittels eines binauralen Nahbereichsschallwandlers (22), wobei die räumliche Wiedergabe unter Zuhilfenahme von bekannten Außenohrübertragungsfunktionen (HRTF) und unter Zuhilfenahme von für einen Abhörraum (12) hörraumoptimierten Übertragungsfunktionen (TF) emuliert wird, wobei basierend auf den Außenohrübertragungsfunktionen (HRTF) ein zu synthetisierender Raum emulierbar ist und wobei basierend auf den hörraumoptimierten Übertragungsfunktionen (TF) der Abhörraum (12) emulierbar ist, wobei die hörraumoptimierte Übertragungsfunktionen (TF) im Vorfeld für den jeweiligen Abhörraum (12) ermittelt sind; wobei die Vorrichtung (20) einen ersten Speicher, in welchem eine erste Vielzahl an Übertragungsfunktionen-Scharen (TF) für unterschiedliche Abhörräume (12) gespeichert ist, und eine Positionsbestimmungseinheit umfasst, wobei die Positionsbestimmungseinheit ausgebildet ist, um die Position zu bestimmen und anhand der bestimmten Position den Abhörraum (12) zu ermitteln; und wobei die Vorrichtung (20) ausgebildet ist, um für die Emulierung der räumlichen Wiedergabe die entsprechenden Übertragungsfunktionen (TF) für den jeweiligen Abhörraum (12) aus den Übertragungsfunktionen-Scharen auszuwählen, **dadurch gekennzeichnet, dass** die hörraumoptimierten Übertragungsfunktionen (TF) je Raum eine Vielzahl an Übertragungsfunktionen, zugeordnet zu einzelnen Raumwinkeln, umfassen, wobei jeder Raumwinkel eine Schallausbreitungsrichtung im Raum darstellt.
11. Vorrichtung (20) gemäß einem der Ansprüche 10, wobei die Vorrichtung (20) einen zweiten Speicher, in welchem eine zweite Vielzahl an Übertragungsfunktionen-Scharen (TF) für unterschiedliche Orientierungen gespeichert ist, und eine Orientierungsbestimmungseinheit umfasst, wobei die Orientierungsbestimmungseinheit ausge-

bildet ist, um eine Orientierung in dem Abhörraum (12) zu ermitteln, und  
 wobei die Vorrichtung (20) ausgebildet ist, um für die Emulierung der räumlichen Wiedergabe die entsprechenden Übertragungsfunktionen (TF) für die jeweilige Orientierung aus den Übertragungsfunktionen-Scharen auszuwählen; und/oder wobei die Vorrichtung (20) einen dritten Speicher, in welchem eine dritte Vielzahl an Übertragungsfunktionen (TF) für unterschiedliche Positionen in dem Abhörraum (12) gespeichert ist, und eine weitere Positionsbestimmungseinheit umfasst, wobei die weitere Positionsbestimmungseinheit ausgebildet ist, um eine Position in dem Abhörraum (12) zu ermitteln, und  
 wobei die Vorrichtung (20) ausgebildet ist, um für die Emulierung der räumlichen Wiedergabe die entsprechenden Übertragungsfunktionen (TF) für die jeweilige Position in dem Abhörraum (12) aus den Übertragungsfunktionen-Scharen auszuwählen; und/oder  
 wobei die Positionsbestimmungseinheit ausgebildet ist, um während der Wiedergabe die Positionen neu zu bestimmen, und wobei die Vorrichtung (20) ausgebildet ist, um basierend auf der aktualisierten Position die hörraumoptimierten Übertragungsfunktionen (TF) zu aktualisieren.

12. Verfahren (200) zur räumlichen Wiedergabe eines Audiosignals (24) mittels eines binauralen Nahbereichsschallwandlers (22), mit den Schritten:

Nachbearbeiten (210) des Audiosignals (24) unter Zuhilfenahme von bekannten Außenohrübertragungsfunktionen (HRTF) und unter Zuhilfenahme von für einen Abhörraum (12) hörraumoptimierten Übertragungsfunktionen (TF), welche für den Abhörraum (12), in welchem die Wiedergabe mittels des binauralen Nahbereichsschallwandlers (22) erfolgen soll, im Vorfeld ermittelt sind, wobei basierend auf den Außenohrübertragungsfunktionen (HRTF) ein zu synthetisierender Raum emulierbar ist und wobei basierend auf den hörraumoptimierten Übertragungsfunktionen (TF) der Abhörraum (12) emulierbar ist;  
 Speichern einer ersten Vielzahl an Übertragungsfunktionen-Scharen (TF) für unterschiedliche Abhörräume (12) in einem ersten Speicher;  
 Bestimmen einer Position; und  
 Ermitteln des Abhörraums (12) anhand der Position,  
 wobei die Vorrichtung (20) ausgebildet ist, um für die Emulierung der räumlichen Wiedergabe die entsprechenden Übertragungsfunktionen (TF) für den jeweiligen Abhörraum (12) aus den Übertragungsfunktionen-Scharen auszuwählen,

**dadurch gekennzeichnet, dass** die hörraumoptimierten Übertragungsfunktionen (TF) je Raum eine Vielzahl an Übertragungsfunktionen, zugeordnet zu einzelnen Raumwinkeln, umfassen, wobei jeder Raumwinkel eine Schallausbreitungsrichtung im Raum darstellt,.

13. Verfahren (200) gemäß Anspruch 12, wobei vor der Wiedergabe ein Kombinieren (220) der Außenohrübertragungsfunktionen (HRTF) und der hörraumoptimierten Übertragungsfunktionen (TF) zu einer Raumbezogenen Raumimpulsantwort (BRIR') erfolgt.
14. System (10 + 20), umfassend:  
 eine Vorrichtung (10) gemäß einem der Ansprüche 1 bis 8; und  
 eine Vorrichtung (20) gemäß einem der Ansprüche 10 bis 11.
15. Computerprogramm mit einem Programmcode, der die Durchführung des Verfahrens (100; 200) nach Anspruch 9 oder 12 bewirkt, wenn das Programm auf einem Computer, einer CPU oder einem mobilen Endgerät abläuft.

#### Claims

1. A device (10) for determining room-optimized transfer functions (TF) for a listening room (12) derived for the listening room (12) and serving for room-optimized post-processing of audio signals (24) in spatial reproduction, wherein the spatial reproduction of the audio signals (24) is emulated by means of a binaural close-range sound transducer (22) using known head-related transfer functions (HRTF) and using the room-optimized transfer functions (TF), wherein a room to be synthesized may be emulated based on the head-related transfer functions (HRTF), and wherein the listening room (12) may be emulated based on the room-optimized transfer functions (TF), wherein the device (10) is configured to analyze room acoustics of the listening room (12) and to determine, starting from analyzing the room acoustics, the room-optimized transfer functions (TF) for the listening room (12) where the spatial reproduction by means of the binaural close-range sound transducer (22) is to take place, wherein the device (10) comprises a storage in which may be deposited a plurality of room-optimized transfer function families (TF) for a plurality of listening rooms (12),  
**characterized in that**  
 the room-optimized transfer functions (TF) include per room a plurality of transfer functions assigned to

- individual solid angles, wherein each solid angle represents a sound propagation direction in the room.
2. The device (10) in accordance with claim 1, wherein the device (10) comprises a microphone (14) of a portable device for acoustic measurement and/or wherein analysis of the room acoustics of the listening room (12) takes place by means of an acoustic measurement in the listening room (12) using ambient noise and/or using a test signal.
  3. The device (10) in accordance with claim 1, wherein the analysis of the room acoustics of the listening room (12) is based on calculating a geometrical model of the listening room (12) and/or modeling the geometrical model based on a camera-based model of the listening room (12).
  4. The device (10) in accordance with claim 2 or 3, wherein the room-optimized transfer functions (TF) are selected such that room acoustics of the listening room (12) may be emulated on the basis thereof.
  5. The device (10) in accordance with any of claims 1 to 4, wherein the device (10) is configured to determine the room-optimized transfer functions (TF) considering a virtual loudspeaker setup in correspondence with which a number of virtual loudspeakers are positioned in the listening room (12).
  6. The device (10) in accordance with any of claims 1 to 5, wherein the known head-related transfer functions (HRTF) comprise a plurality of individual transfer functions (TF) for the left and right ears which are associated to directional vectors for a plurality of virtual sound sources.
  7. The device (10) in accordance with any of claims 1 to 7, wherein emulating the spatial reproduction is based on interaural features, balance features and distance features, wherein the interaural features comprise a connection between a direction of incidence in the medial planes and an individual or non-individual head-related filtering, wherein the balance features comprise a connection between a lateral direction of incidence and a difference in volume and/or a connection between the lateral direction of incidence and a run-time difference, wherein the distance features comprise a connection between a virtual distance and frequency-dependent filtering and/or a connection between the virtual distance and an initial time gap and/or a connection between the virtual distance and a reflection behavior.
  8. The device (10) in accordance with any of claims 1 to 7, wherein the binaural close-range sound transducer (22) is a headset configured to output as the audio signal (24) a multi-channel stereo signal, an object-based audio signal (24) and/or an audio signal (24) on the basis of a wave-field synthesis algorithm.
  9. A method (100) for determining room-optimized transfer functions (TF) for a listening room (12) which are derived for the listening room (12) and may serve for room-optimized post-processing of audio signals (24) in spatial reproduction, wherein the spatial reproduction of the audio signals (24) by means of a binaural close-range sound transducer (22) is emulated using known head-related transfer functions (HRTF) and using the room-optimized transfer functions (TF), wherein a room to be synthesized may be emulated based on the head-related transfer functions (HRTF), and wherein the listening room (12) may be emulated based on the room-optimized transfer functions (TF), comprising:
    - analyzing (120) prevailing room acoustics of the listening room (12); and
    - determining (110) the room-optimized transfer functions (TF) for the listening room (12) where spatial reproduction by means of the binaural close-range sound transducer (22) is to take place, on the basis of analyzing the room acoustics;
    - depositing a plurality of room-optimized transfer function families (TF) for a plurality of listening rooms (12) into a storage, **characterized in that** the room-optimized transfer functions (TF) include per room a plurality of transfer functions assigned to individual solid angles, wherein each solid angle represents a sound propagation direction in the room.
  10. A device (20) for spatial reproduction of an audio signal (24) by means of a binaural close-range sound transducer (22), wherein the spatial reproduction is emulated using known head-related transfer functions (HRTF) and using room-optimized transfer functions (TF) for a listening room (12), wherein a room to be synthesized may be emulated based on the head-related transfer functions (HRTF), and wherein the listening room (12) may be emulated based on the room-optimized transfer functions (TF), wherein the room-optimized transfer functions (TF) have been determined beforehand for the respective listening room (12); wherein the device (20) comprises a first storage in which are stored a first plurality of transfer function families (TF) for different listening rooms (12), and a position-determining unit, wherein the position-determining unit is configured to identify the position and determine the listening room (12) using the position identified; and

wherein the device (20) is configured to select, for emulating the spatial reproduction, the corresponding transfer functions (TF) for the respective listening room (12) from the transfer function families, **characterized in that** the room-optimized transfer functions (TF) include per room a plurality of transfer functions assigned to individual solid angles, wherein each solid angle represents a sound propagation direction in the room.

11. The device (20) in accordance with claim 10, wherein the device (20) comprises a second storage in which are stored a second plurality of transfer function families (TF) for different orientations, and an orientation-determining unit, wherein the orientation-determining unit is configured to determine an orientation in the listening room (12), and wherein the device (20) is configured to select, for emulating the spatial reproduction, the corresponding transfer functions (TF) for the respective orientation from the transfer function families, and/or wherein the device (20) comprises a third storage in which are stored a third plurality of transfer function families (TF) for different positions in the listening room (12), and another position-determining unit, wherein the other position-determining unit is configured to determine a position in the listening room (12), and wherein the device (20) is configured to select, for emulating the spatial reproduction, the corresponding transfer functions (TF) for the respective position in the listening room (12) from the transfer function families, and/or wherein the position-determining unit is configured to determine, while reproducing, the positions again, and wherein the device (20) is configured to update the room-optimized transfer functions (TF) based on the updated position.
12. A method (200) for spatially reproducing an audio signal (24) by means of a binaural close-range sound transducer (22), comprising:

post-processing (210) the audio signal (24) using known head-related transfer functions (HRTF) and using room-optimized transfer functions (TF) for a listening room (12) which have been determined beforehand for the listening room (12) where reproduction by means of the binaural close-range sound transducer (22) is to take place, wherein a room to be synthesized may be emulated based on the head-related transfer functions (HRTF), and wherein the listening room (12) may be emulated based on the room-optimized transfer functions (TF); storing a first plurality of transfer function families (TF) for different listening rooms (12) in a

first storage; identifying a position; and determining the listening room (12) using the position,

wherein the device (20) is configured to select, for emulating the spatial reproduction, the corresponding transfer functions (TF) for the respective listening room (12) from the transfer function families, **characterized in that** the room-optimized transfer functions (TF) include per room a plurality of transfer functions assigned to individual solid angles, wherein each solid angle represents a sound propagation direction in the room.

13. The method (200) in accordance with claim 12, wherein, before reproducing, combining (220) the head-related transfer functions (HRTF) and the room-optimized transfer functions (TF) to form a room-related room impulse response (BRIR') takes place.

14. A system (10 + 20) comprising:

a device (10) in accordance with any of claims 1 to 8; and  
a device (20) in accordance with any of claims 10 to 11.

15. A computer program having program code that causes performing the method (100; 200) in accordance with claim 9 or 12 when the program runs on a computer, CPU or mobile terminal.

## Revendications

1. Dispositif (10) de détermination de fonctions de transmission (TF) acoustiquement optimisées pour une salle d'écoute (12) qui sont dérivées pour la salle d'écoute (12) et qui servent pour le post-traitement acoustiquement optimisé de signaux audio (24) lors de la reproduction spatiale, où la reproduction spatiale des signaux audio (24) est émulée au moyen d'un transducteur acoustique binaural de proximité (22) à l'aide de fonctions de transmission de l'oreille externe (HRTF) connues et à l'aide des fonctions de transmission (TF) acoustiquement optimisées, dans lequel un local à synthétiser peut être émulé sur base des fonctions de transmission de l'oreille externe (HRTF) et dans lequel la salle d'écoute (12) peut être émulée sur base des fonctions de transmission acoustiquement optimisées (TF), dans lequel le dispositif (10) est conçu pour analyser une acoustique de la salle d'écoute (12) et pour déterminer, partant de l'analyse de l'acoustique de la salle, les fonctions de transmission (TF) acoustiquement optimisées pour la salle d'écoute (12) dans la-

quelle doit avoir lieu la reproduction spatiale au moyen du transducteur acoustique binaural de proximité (22),

dans lequel le dispositif (10) comporte une mémoire dans laquelle peuvent être mémorisés une pluralité de groupes de fonctions de transmission (TF) acoustiquement optimisées pour une pluralité de salles d'écoute (12),

**caractérisé par le fait que**

les fonctions de transmission (TF) acoustiquement optimisées comportent, par salle, une pluralité de fonctions de transmission associées à des angles spatiaux individuels, où chaque angle spatial représente une direction de propagation du son dans l'espace.

2. Dispositif (10) selon la revendication 1, dans lequel le dispositif (10) comporte un microphone (14) d'un dispositif portable pour la mesure acoustique et/ou dans lequel l'analyse de l'acoustique spatiale de la salle d'écoute (12) a lieu au moyen d'une mesure acoustique dans la salle d'écoute (12) à l'aide du bruit ambiant et/ou à l'aide d'un signal de test.
3. Dispositif (10) selon la revendication 1, dans lequel l'analyse acoustique spatiale de la salle d'écoute (12) est basée sur le calcul d'un modèle géométrique de la salle d'écoute (12) et/ou la modélisation du modèle géométrique sur base d'un modèle de la salle d'écoute (12) basé sur une caméra.
4. Dispositif (10) selon la revendication 2 ou 3, dans lequel les fonctions de transmission (TF) acoustiquement optimisées sont sélectionnées de sorte qu'une acoustique spatiale de la salle d'écoute (12) puisse être émulée sur base de ces dernières.
5. Dispositif (10) selon l'une des revendications 1 à 4, dans lequel le dispositif (10) est conçu pour déterminer les fonctions de transmission (TF) acoustiquement optimisées en tenant compte d'une configuration de haut-parleurs virtuelle selon laquelle un nombre de haut-parleurs virtuels sont positionnés dans la salle d'écoute (12).
6. Dispositif (10) selon l'une des revendications 1 à 5, dans lequel les fonctions de transmission de l'oreille externe (HRTF) connues comportent une pluralité de fonctions de transmission individuelles pour l'oreille gauche et l'oreille droite (TF) associées à des vecteurs liés à la direction pour une pluralité de sources de son virtuelles.
7. Dispositif (10) selon l'une des revendications 1 à 6, dans lequel l'émulation de la reproduction spatiale est basée sur des caractéristiques interaurales, des caractéristiques d'équilibre et des caractéristiques de distance,

dans lequel les caractéristiques interaurales comportent un rapport entre une direction d'incidence dans le plan médian et un filtrage individuel ou non individuel de l'oreille externe, dans lequel les caractéristiques d'équilibre comportent un rapport entre une direction d'incidence latérale et une différence de volume sonore et/ou un rapport entre la direction d'incidence latérale et une différence de durée, dans lequel les caractéristiques de distance comportent un rapport entre une distance virtuelle et un filtrage dépendant de la fréquence et/ou un rapport entre la distance virtuelle et un intervalle de temps initial et/ou un rapport entre la distance virtuelle et un comportement de réflexion.

8. Dispositif (10) selon l'une des revendications 1 à 7, dans lequel le transducteur acoustique binaural de proximité (22) est un casque qui est configuré pour sortir un signal stéréo multicanal, un signal audio à base d'objet (24) et/ou un signal audio (24) sur base d'un algorithme de synthèse de champ d'onde comme signal audio (24).
9. Procédé (100) pour déterminer des fonctions de transmission (TF) acoustiquement optimisées pour une salle d'écoute (12) qui sont dérivées pour la salle d'écoute (12) et qui peuvent servir pour le post-traitement acoustiquement optimisé de signaux audio (24) lors de la reproduction spatiale, dans lequel la reproduction spatiale des signaux audio (24) est émulée au moyen d'un transducteur acoustique binaural de proximité (22) à l'aide de fonctions de transmission de l'oreille externe (HRTF) connues et à l'aide des fonctions de transmission (TF) acoustiquement optimisées, dans lequel un local à synthétiser peut être émulé sur base des fonctions de transmission de l'oreille externe (HRTF) et dans lequel la salle d'écoute (12) peut être émulée sur base des fonctions de transmission acoustiquement optimisées (TF),  
aux étapes consistant à:

analyser (120) une acoustique spatiale prédominante de la salle d'écoute (12); et

déterminer (110) les fonctions de transmission (TF) acoustiquement optimisées pour la salle d'écoute (12) dans laquelle doit avoir lieu la reproduction spatiale au moyen du transducteur acoustique binaural de proximité (22), sur base de l'analyse de l'acoustique spatiale;  
mémoriser dans une mémoire une pluralité de groupes de fonctions de transmission (TF) acoustiquement optimisées pour une pluralité de salles d'écoute (12),

**caractérisé par le fait que**

les fonctions de transmission (TF) acoustiquement optimisées comportent, par salle, une pluralité de fonctions de transmission associées à

des angles spatiaux individuels, où chaque angle spatial représente une direction de propagation du son dans l'espace.

10. Dispositif (20) de reproduction spatiale d'un signal audio (24) au moyen d'un transducteur acoustique binaural de proximité (22), dans lequel la reproduction spatiale est émulée à l'aide de fonctions de transmission de l'oreille externe (HRTF) connues et à l'aide de fonctions de transmission (TF) acoustiquement optimisées pour une salle d'écoute (12), dans lequel un local à synthétiser peut être émulé sur base des fonctions de transmission de l'oreille externe (HRTF) et dans lequel la salle d'écoute (12) peut être émulée sur base des fonctions de transmission acoustiquement optimisées (TF), dans lequel les fonctions de transmission (TF) acoustiquement optimisées sont déterminées auparavant pour la salle d'écoute (12) respective; dans lequel le dispositif (20) comporte une première mémoire dans laquelle est mémorisée une première pluralité de groupes de fonctions de transmission (TF) pour différentes salles d'écoute (12), et une unité de détermination de position, dans lequel l'unité de détermination de position est conçue pour déterminer la position et pour déterminer la salle d'écoute (12) sur base de la position déterminée; et dans lequel le dispositif (20) est conçu pour sélectionner, pour l'émulation de la reproduction spatiale, les fonctions de transmission (TF) correspondantes pour la salle d'écoute respective (12) parmi les groupes de fonctions de transmission, **caractérisé par le fait que** les fonctions de transmission (TF) acoustiquement optimisées pour la salle d'écoute comportent, par salle, une pluralité de fonctions de transmission associées à des angles spatiaux individuels, où chaque angle spatial représente une direction de propagation de son dans l'espace.
11. Dispositif (20) selon la revendication 10, dans lequel le dispositif (20) comporte une deuxième mémoire dans laquelle est mémorisée une deuxième pluralité de groupes de fonctions de transmission (TF) pour différentes orientations, et une unité de détermination d'orientation, dans lequel l'unité de détermination d'orientation est conçue pour déterminer une orientation dans la salle d'écoute (12), et dans lequel le dispositif (20) est conçu pour sélectionner, pour l'émulation de la reproduction spatiale, les fonctions de transmission (TF) correspondantes pour l'orientation respective parmi les groupes de fonctions de transmission; et/ou dans lequel le dispositif (20) comporte une troisième mémoire dans laquelle est mémorisée une troisième pluralité de groupes de fonctions de transmission

(TF) pour différentes positions dans la salle d'écoute (12), et une autre unité de détermination de position, dans lequel l'autre unité de détermination de position est conçue pour déterminer une position dans la salle d'écoute (12), et

dans lequel le dispositif (20) est conçu pour sélectionner, pour l'émulation de la reproduction spatiale, les fonctions de transmission (TF) correspondantes pour la position respective dans la salle d'écoute (12) parmi les groupes de fonctions de transmission; et/ou

dans lequel l'unité de détermination de position est conçue pour déterminer à nouveau les positions pendant la reproduction, et dans lequel le dispositif (20) est conçu pour mettre à jour les fonctions de transmission (TF) acoustiquement optimisées sur base de la position mise à jour.

12. Procédé (200) de reproduction spatiale d'un signal audio (24) au moyen d'un transducteur acoustique binaural de proximité (22), aux étapes consistant à:

post-traiter (210) le signal audio (24) à l'aide de fonctions de transmission de l'oreille externe (HRTF) connues et à l'aide de fonctions de transmission (TF) acoustiquement optimisées pour une salle d'écoute (12) qui sont déterminées auparavant pour la salle d'écoute (12) dans laquelle doit avoir lieu la reproduction au moyen du transducteur acoustique binaural de proximité (22), où un espace à synthétiser peut être émulé sur base des fonctions de transmission de l'oreille externe (HRTF), et où la salle d'écoute (12) peut être émulée sur base des fonctions de transmission (TF) acoustiquement optimisées de la salle d'écoute; mémoriser une première pluralité de groupes de fonctions de transmission (TF) pour différentes salles d'écoute (12) dans une première mémoire;

déterminer une position; et déterminer la salle d'écoute (12) à l'aide de la position,

dans lequel le dispositif (20) est conçu pour sélectionner, pour l'émulation de la reproduction spatiale, les fonctions de transmission (TF) correspondantes pour la salle d'écoute (12) respective parmi les groupes de fonctions de transmission,

**caractérisé par le fait que**

les fonctions de transmission (TF) acoustiquement optimisées comportent, par salle, une pluralité de fonctions de transmission associées à des angles spatiaux individuels, où chaque angle spatial représente une direction de propagation de son dans l'espace.

13. Procédé (200) selon la revendication 12, dans lequel

a lieu, avant la reproduction, une combinaison (220) des fonctions de transmission de l'oreille externe (HRTF) et des fonctions de transmission (TF) acoustiquement optimisées pour une réponse impulsionnelle spatiale sur base de l'espace (BRIR'). 5

**14.** Système (10 + 20) comportant:

un dispositif (10) selon l'une des revendications 1 à 8; et 10  
un dispositif (20) selon l'une des revendications 10 à 11.

**15.** Programme d'ordinateur avec un code de programme qui provoque la réalisation du procédé (100; 200) selon la revendication 9 ou 12 lorsque le programme est exécuté sur un ordinateur, une unité centrale de traitement (CPU) ou un terminal mobile. 15

20

25

30

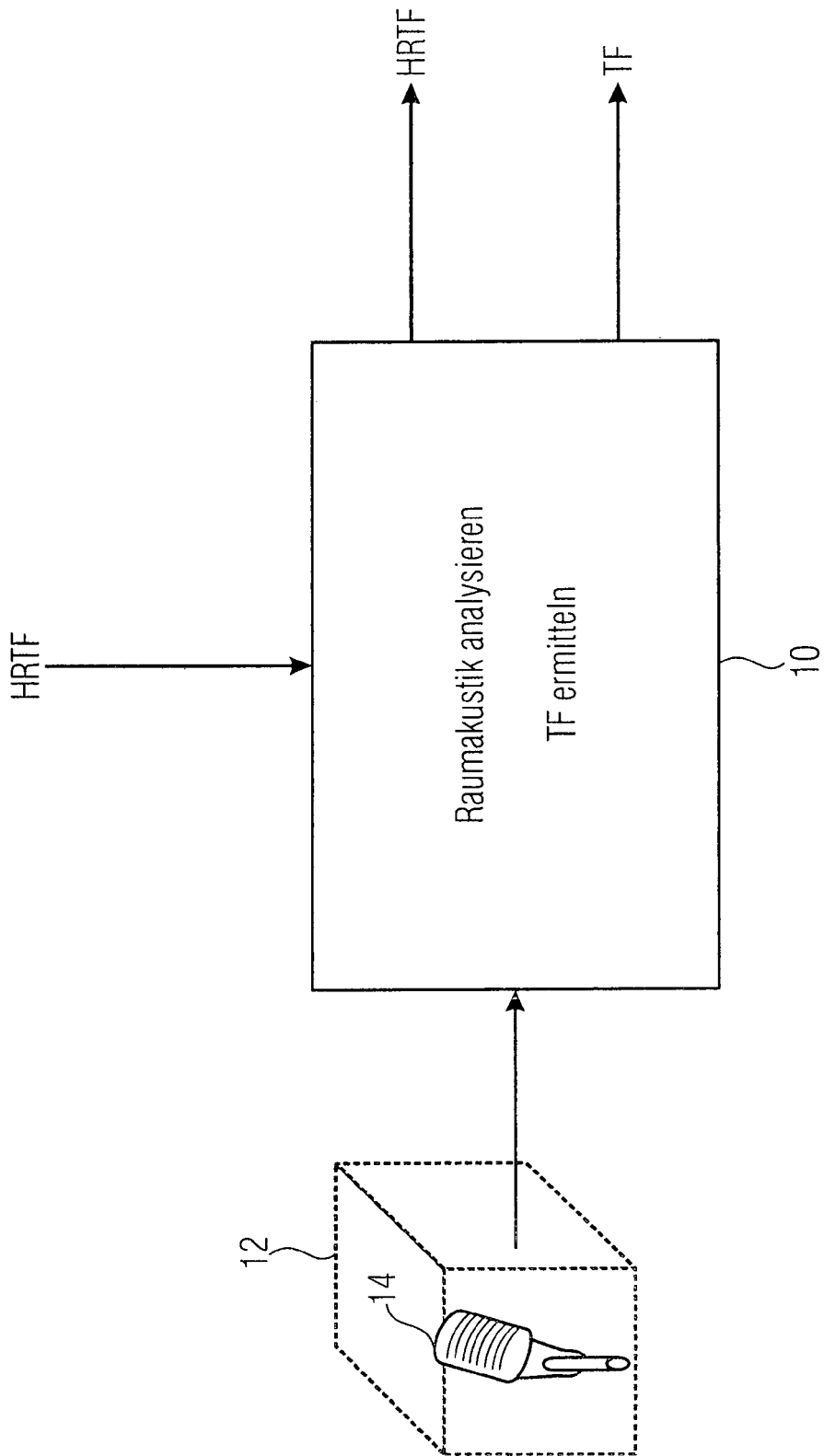
35

40

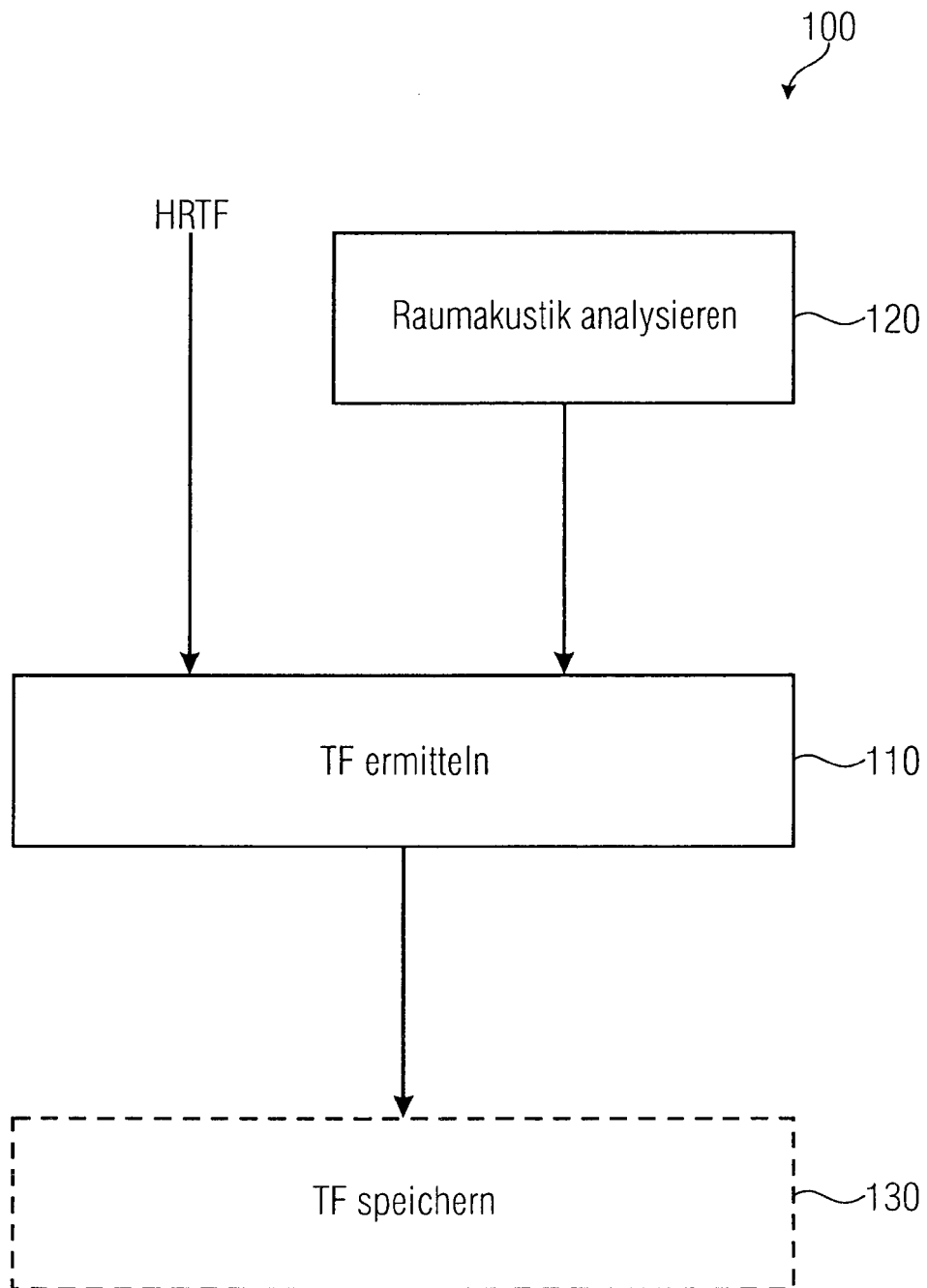
45

50

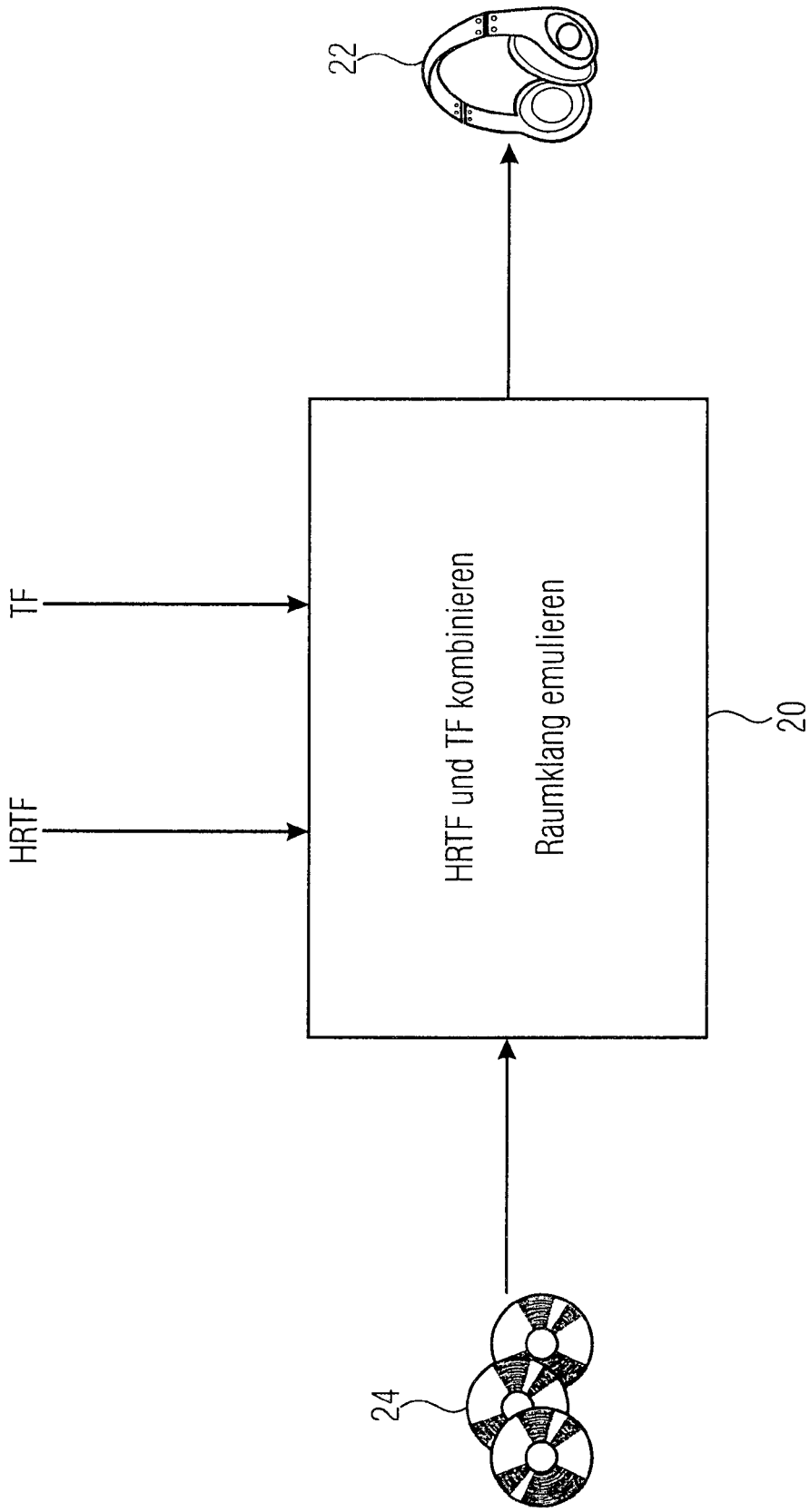
55



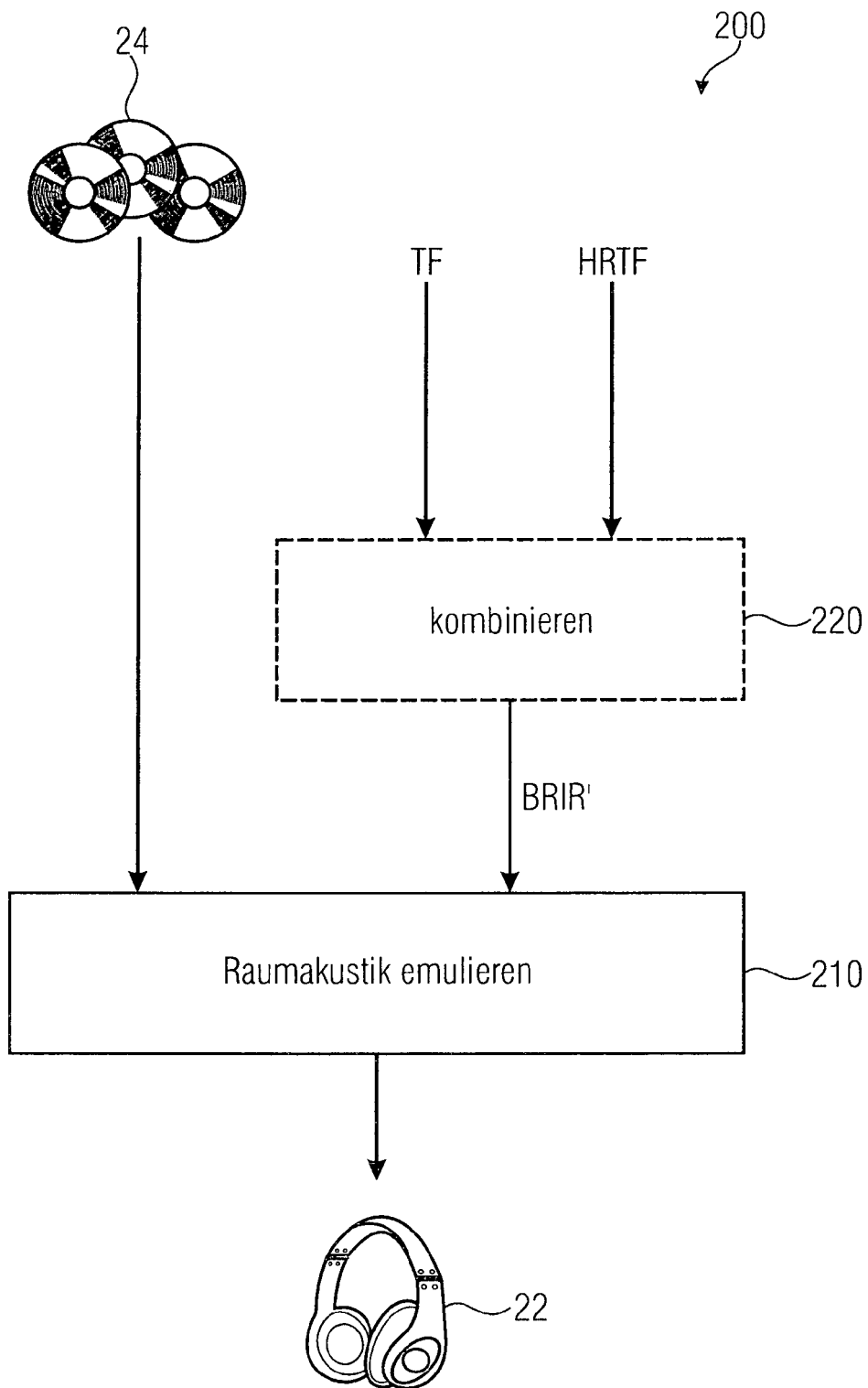
FIGUR 1A



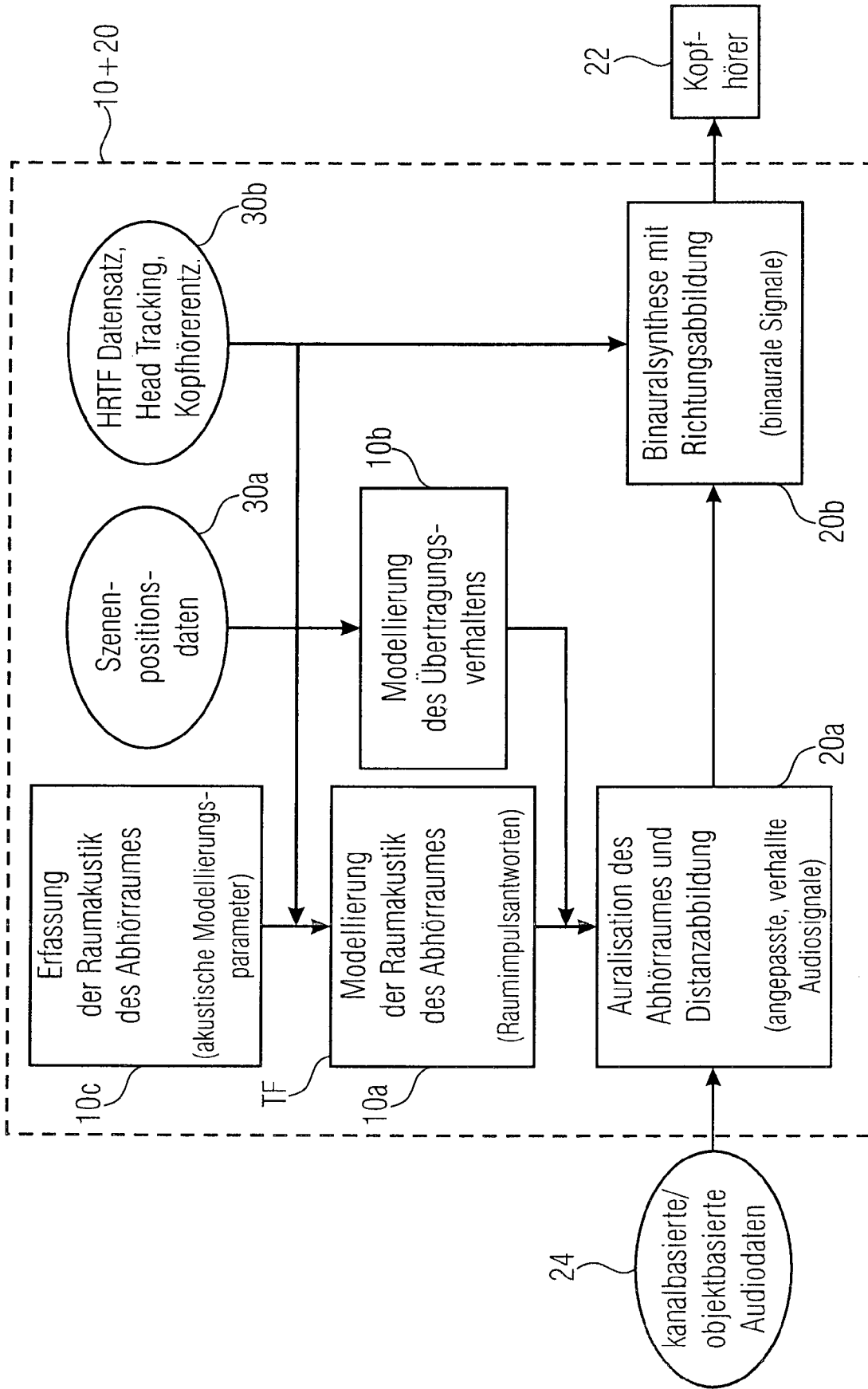
FIGUR 1B



FIGUR 2A



FIGUR 2B



FIGUR 3

**IN DER BESCHREIBUNG AUFGEFÜHRTE DOKUMENTE**

*Diese Liste der vom Anmelder aufgeführten Dokumente wurde ausschließlich zur Information des Lesers aufgenommen und ist nicht Bestandteil des europäischen Patentdokumentes. Sie wurde mit größter Sorgfalt zusammengestellt; das EPA übernimmt jedoch keinerlei Haftung für etwaige Fehler oder Auslassungen.*

**In der Beschreibung aufgeführte Patentdokumente**

- US 20130272527 A1 [0008]
- US 20080273708 A1 [0009]