



(11) **EP 3 069 530 B1**

(12) **EUROPÄISCHE PATENTSCHRIFT**

(45) Veröffentlichungstag und Bekanntmachung des Hinweises auf die Patenterteilung:
20.02.2019 Patentblatt 2019/08

(51) Int Cl.:
H04S 3/00 ^(2006.01) **G10L 19/008** ^(2013.01)

(21) Anmeldenummer: **14793183.6**

(86) Internationale Anmeldenummer:
PCT/EP2014/073808

(22) Anmeldetag: **05.11.2014**

(87) Internationale Veröffentlichungsnummer:
WO 2015/071148 (21.05.2015 Gazette 2015/20)

(54) **VERFAHREN UND VORRICHTUNG ZUM KOMPRIMIEREN UND DEKOMPRIMIEREN VON SCHALLFELDDATEN EINES GEBIETS**

METHOD AND DEVICE FOR COMPRESSING AND DECOMPRESSING SOUND FIELD DATA OF AN AREA

PROCÉDÉ ET DISPOSITIF POUR COMPRESSER ET DÉCOMPRESSER DES DONNÉES DE CHAMP SONORE D'UN DOMAINE

(84) Benannte Vertragsstaaten:
AL AT BE BG CH CY CZ 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

(56) Entgegenhaltungen:
WO-A2-2010/012478

(30) Priorität: **14.11.2013 DE 102013223201**

(43) Veröffentlichungstag der Anmeldung:
21.09.2016 Patentblatt 2016/38

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

- **Johannes Käsbach: "Development and evaluation of a mixed-order Ambisonics playback system", , 1. November 2010 (2010-11-01), Seiten 1-103, XP055161760, Danmarks Tekniske Universitet, 2800 Kgs. Lyngby, Denmark Gefunden im Internet: URL:http://johannes.kaesbach.de/Acoustics_files/Master%20Thesis%20-%20Mixed-Order%20Ambisonics%28small%29.pdf [gefunden am 2015-01-13]**
- **Chris Travis: "A NEW MIXED-ORDER SCHEME FOR AMBISONIC SIGNALS", Ambisonics Symposium 2009, 25. Juni 2009 (2009-06-25), Seiten 1-6, XP055161700, Graz Gefunden im Internet: URL:<http://ambisonics.iem.at/symposium2009/proceedings/ambisym09-travis-newmixedorder.pdf> [gefunden am 2015-01-13]**
- **IKHWANA ELFITRI ET AL: "Multichannel Audio Coding Based on Analysis by Synthesis", PROCEEDINGS OF THE IEEE, IEEE. NEW YORK, US, Bd. 99, Nr. 4, 1. April 2011 (2011-04-01), Seiten 657-670, XP011363629, ISSN: 0018-9219, DOI: 10.1109/JPROC.2010.2102310**

(72) Erfinder:
• **NOWAK, Johannes**
99084 Erfurt (DE)
• **SLADECZEK, Christoph**
98693 Ilmenau (DE)

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

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).

EP 3 069 530 B1

Beschreibung

[0001] Die vorliegende Erfindung bezieht sich auf die Audiotechnik und insbesondere auf die Komprimierung von räumlichen Schallfelddaten.

[0002] Die akustische Beschreibung von Räumen ist von hohem Interesse zur Ansteuerung von Wiedergabeanordnungen in Form z.B. eines Kopfhörers, einer Lautsprecheranordnung mit z.B. zwei bis zu einer mittleren Anzahl von Lautsprechern, wie beispielsweise 10 Lautsprechern oder auch für Lautsprecheranordnungen mit einer großen Anzahl von Lautsprechern, wie sie bei der Wellenfeldsynthese (WFS) zum Einsatz kommen.

[0003] Für die räumliche Audiocodierung allgemein existieren verschiedene Ansätze. Ein Ansatz besteht z.B. darin, verschiedene Kanäle für verschiedene Lautsprecher an vordefinierten Lautsprecherpositionen zu erzeugen, wie es beispielsweise bei MPEG-Surround der Fall ist. Dadurch erhält ein Hörer, der in dem Wiedergaberaum an einer bestimmten und optimalerweise der mittleren Position angeordnet ist, ein Raumgefühl für das wiedergegebene Schallfeld.

[0004] Eine alternative Raumbeschreibung besteht darin, einen Raum durch seine Impulsantwort zu beschreiben. Wird beispielsweise eine Schallquelle irgendwo in einem Raum oder Gebiet positioniert, so kann dieser Raum bzw. dieses Gebiet mit einem Kreisarray von Mikrofonen im Falle eines zweidimensionalen Gebiets oder mit einem Kugel-Mikrofon-Array im Falle eines dreidimensionalen Gebiets ausgemessen werden. Wenn beispielsweise ein Kugel-Mikrofon-Array mit einer hohen Anzahl von Mikrofonen betrachtet wird, wie beispielsweise 350 Mikrofonen, so wird eine Vermessung des Raumes folgendermaßen vonstatten gehen. An einer bestimmten Position innerhalb oder außerhalb des Mikrofon-Arrays wird ein Impuls erzeugt. Daraufhin wird von jedem Mikrofon die Antwort auf diesen Impuls, also die Impulsantwort gemessen. Abhängig davon, wie stark die Nachhalleigenschaften sind, wird dann eine längere oder kürzere Impulsantwort gemessen. So haben - bezüglich der Größenordnung - Messungen in großen Kirchen beispielsweise ergeben, dass Impulsantworten über 10 s dauern können.

[0005] Ein solcher Satz von z.B. 350 Impulsantworten beschreibt damit die Schallcharakteristik dieses Raumes für die spezielle Position einer Schallquelle, an der der Impuls erzeugt worden ist. Mit anderen Worten ausgedrückt stellt dieser Satz von Impulsantworten Schallfelddaten des Gebiets dar, und zwar für genau den einen Fall, bei dem eine Quelle an der Position positioniert ist, an der der Impuls erzeugt worden ist. Um den Raum weiter zu vermessen, also um die Schalleigenschaften des Raumes zu erfassen, wenn eine Quelle an einem anderen Raum positioniert wird, muss die dargestellte Prozedur für jede weitere Position z.B. außerhalb des Arrays (aber auch innerhalb des Arrays) wiederholt werden. Würde man daher z.B. einen Konzertsaal schallfeldmäßig erfassen, wenn z.B. ein Musiker-Quartett

spielt, bei dem die einzelnen Musiker an vier verschiedenen Positionen angeordnet sind, so werden bei dem genannten Beispiel für jede der vier Positionen 350 Impulsantworten gemessen, und diese $4 \times 350 = 1400$ Impulsantworten stellen dann die Schallfelddaten des Gebiets dar.

[0006] Da die zeitliche Länge der Impulsantworten durchaus beträchtliche Werte annehmen kann, und da womöglich noch eine detailliertere Darstellung der Schalleigenschaften des Raumes im Hinblick auf nicht nur vier sondern noch mehr Positionen gewünscht sein kann, ergibt sich eine riesige Menge an Impulsantwort-Daten, insbesondere wenn berücksichtigt wird, dass die Impulsantworten durchaus Längen über 10 s annehmen können.

[0007] Ansätze für räumliches Audiocoding ist z.B. spatial audio coding (SAC) [1] bzw. spatial audio object coding (SAOC) [2], die eine Bitraten-effiziente Kodierung von Mehrkanal-Audiosignalen bzw. objektbasierten räumlichen Audioszenen ermöglichen. Das spatial impulse response rendering (SIRR) [3] und die Weiterentwicklung directional audio coding (DirAc) [4] sind parametrische Kodierverfahren und basieren auf einer zeitabhängigen Schalleinfallrichtungsschätzung (direction of arrival - DOA), sowie einer Schätzung der Diffusität innerhalb von Frequenzbändern. Hier wird eine Trennung zwischen nichtdiffusem und diffusem Schallfeld vorgenommen. In [5] wird die verlustfreie Kompression von Kugelmikrofonarraydaten und die Kodierung von Higher-Order-Ambisonics-Signalen behandelt. Die Kompression wird durch Ausnutzen redundanter Daten zwischen den Kanälen (interchannel redundancy) erreicht.

[0008] Untersuchungen in [6] zeigen eine getrennte Betrachtung von frühem und spätem Schallfeld bei der binauralen Wiedergabe. Für dynamische Systeme, in denen Kopfbewegungen berücksichtigt werden wird die Filterlänge optimiert indem nur das frühe Schallfeld in Echtzeit gefaltet wird. Für das späte Schallfeld reicht lediglich ein Filter für alle Richtungen, ohne dabei die wahrgenommene Qualität zu reduzieren. In [7] werden kopfbezogene Übertragungsfunktionen (HRTF) auf einer Kugel im sphärischen harmonischen Bereich dargestellt. Der Einfluss verschiedener Genauigkeiten mittels unterschiedlicher Ordnungen sphärischer Harmonischer auf die Interaurale Kreuzkorrelation und die Raum-Zeit-Korrelation (spatio-temporal correlation) wird analytisch untersucht. Dies geschieht in Oktavbändern im diffusen Schallfeld.

[1] Herre, J et al (2004) Spatial Audio Coding: Next-generation efficient and compatible coding of multi-channel audio AES Convention Paper 6186 presented at the 117th Convention, San Francisco, USA

[2] Engdegard, J et al (2008) Spatial Audio Object Coding (SAOC) - The Upcoming MPEG Standard on Parametric Object Based Audio Coding, AES

Convention Paper 7377 presented at the 125th Convention, Amsterdam, Netherlands

[3] Merimaa J and Pulkki V (2003) Perceptually-based processing of directional room responses for multichannel loudspeaker reproduction, IEEE Workshop on Applications of Signal Processing to Audio and Acoustics 5

[4] Pulkki, V (2007) Spatial Sound Reproduction with Directional Audio Coding, J. Audio Eng. Soc., Vol. 55. No.6 10

[5] Hellerud E et al (2008) Encoding Higher Order Ambisonics with AAC AES Convention Paper 7366 presented at the 125th Convention, Amsterdam, Netherlands 15

[6] Liindau A, Kosanke L, Weinzierl S (2010) Perceptual evaluation of physical predictors of the mixing time in binaural room impulse responses AES Convention Paper presented at the 128th Convention, London, UK 20

[7] Avni, A and Rafaely B (2009) Interaural cross correlation and spatial correlation in a sound field represented by spherical harmonics in Ambisonics Symposium 2009, Graz, Austria 25

[0009] Ein Encoder-Dekoder Schema für niedrige Bitraten wird in [8] beschrieben. Der Encoder generiert ein Komposit-Audio-Informationen-Signal, das das zu reproduzierende Schallfeld beschreibt und einen Richtungsvektor oder Steering-Control-Signal. Das Spektrum wird in Subbänder zerlegt. Zur Steuerung wird in jedem Subband die dominante Richtung ausgewertet. Basierend auf der wahrgenommenen räumlichen Audioszene wird in [9] ein räumliches Audiokodier-Framework im Frequenzbereich beschrieben. Zeit-Frequenzabhängige Richtungsvektoren beschreiben die Eingangsaudioszene. 30 35 40

[10] beschreibt ein parametrisches, kanalbasiertes Audiokodierverfahren im Zeit- und Frequenzbereich. In [11] wird ein *binaural-cue-coding* (BCC) beschrieben, das ein oder mehr objektbasierte Cue-Codes verwendet. Diese beinhalten Richtung, Weite und Umhüllung einer auditorischen Szene. [12] bezieht sich auf die Verarbeitung von Kugelarraydaten für die Wiedergabe mittels Ambisonics. Dabei sollen die Verzerrungen des Systems durch Messfehler, wie z.B. Rauschen, equalisiert werden. In [13] wird ein kanalbasiertes Kodierverfahren beschrieben, dass sich auch auf Positionen der Lautsprecher, sowie einzelner Audio Objekte bezieht. In [14] wird ein Matrix-basiertes Kodierverfahren vorgestellt, das die Echtzeitübertragung von Higher Order Ambisonics Schallfeldern mit Ordnungen größer als 3 er- 45 50 55

möglichst.

In [15] wird eine Methode zur Kodierung von räumlichen Audiodaten beschrieben, das unabhängig vom Wiedergabesystem ist. Dabei wird das Eingangsmaterial in zwei Gruppen unterteilt, von denen die erste Gruppe das Audio beinhaltet, das hohe Lokalisierbarkeit benötigt, während die zweite Gruppe mit für die Lokalisation ausreichend niedrigen Ambisonics-Ordnungen beschrieben wird. In der ersten Gruppe wird das Signal in einen Satz aus Monokanälen mit Metadaten kodiert. Die Metadaten beinhalten Zeitinformationen, wann der entsprechende Kanal wiedergegeben werden soll und Richtungsinformationen zu jedem Moment. Bei der Wiedergabe werden die Audiokanäle für herkömmliche Panning-Algorithmen dekodiert, wobei das Wiedergabe-System bekannt sein muss. Das Audio in der zweiten Gruppe wird in Kanäle verschiedener Ambisonics-Ordnungen kodiert. Bei der Dekodierung werden dem Wiedergabesystem entsprechende Ambisonics-Ordnungen verwendet.

[8] Dolby R M (1999) Low-bit-rate spatial coding method and system, EP 1677576 A3

[9] Goodwin M and Jot J-M (2007) Spatial audio coding based on universal spatial cues, US 8,379,868 B2

[10] Seefeldt A and Vinton M (2006) Controlling spatial audio coding parameters as a function of auditory events, EP 2296142 A2

[11] Faller C (2005) Parametric coding of spatial audio with object-based side information, US 8340306 B2

[12] Kordon S, Batke J-M, Krüger A (2011) Method and apparatus for processing signals of a spherical microphone array on a rigid sphere used for generating an ambisonics representation of the sound field, EP 2592845 A1

[13] Corteel E and Rosenthal M (2011) Method and device for enhanced sound field reproduction of spatially encoded audio input signals, EP 2609759 A1

[14] Abeling S et al (2010) Method and apparatus for generating and for decoding sound field data including ambisonics sound field data of an order higher than three, EP 2451196 A1

[15] Arumi P and Sole A (2008) Method and apparatus for three-dimensional acoustic field encoding and optimal reconstruction, EP 2205007 A1

[0010] Die Fachveröffentlichung "Development and evaluation of a mixed-order Ambisonics playback sys-

tem", J. Käsbaach, Technical University of Denmark, November 2010 offenbart ein Ambisonic-Playback-System mit gemischter Ordnung, in dem die Zerlegung nach sphärischen harmonischen Komponenten des dreidimensionalen Schallfelds um zusätzliche horizontale Komponenten ergänzt worden ist. Unter Berücksichtigung der Orthonormalität-Eigenschaften des sphärischen Funktionen werden die maximalen zweidimensionalen und dreidimensionalen Ordnungen für ein gegebenes Lautsprecher-Array bestimmt. Basierend auf dieser Analyse wird eine alternative Implementation gemischter Ordnung vorgeschlagen, die eine abgeschnittene Ordnung der inherenten Legendre-Funktionen erforderlich macht.

[0011] Die Fachveröffentlichung "A New Mixed-Order Scheme for Ambisonic Signals", Chris Travis, Ambisonics Symposium 2009, 25. Juni 2009, Graz, Seiten 1-6 bezieht sich auf Systeme gemischter Ordnung, die eine höhere Auflösung in der horizontalen Ebene als an den Polen liefern. Ein Zwei-Parameter-Schema ($\#H\#V$) schneidet die sphärische harmonische Erweiterung in einer bestimmten Art und Weise ab. Sie resultiert in Auflösungsversus-Elevations-Kurven, die in und nahe der Horizontalebene flacher sind.

[0012] Die WO 2010/012478 A2 offenbart ein System zur Erzeugung von binauralen Signalen basierend auf einem Multikanalsignal, das eine Mehrzahl von Kanälen darstellt und zur Reproduktion durch eine Lautsprecherkonfiguration gedacht ist, die für jeden Kanal eine zugeordnete virtuelle Schallquellenposition aufweist.

[0013] Die Fachveröffentlichung "Multichannel Audio Coding Based on Analysis by Synthesis", I. Elfitri, u.a., Proceedings of the IEEE, New York, Band 99, Nr. 4, 1.4.2011, Seiten 657-670 beschreibt ein Closed-Loop-Codiersystem basierend auf einem Analyse-durch-Synthese-Prinzip, das auf die MPEG-Surround-Architektur angewendet wird.

[0014] Die Aufgabe der vorliegenden Erfindung besteht darin, ein effizienteres Konzept zum Handhaben wie z. B. Komprimieren oder Dekomprimieren von Schallfelddaten eines Gebiets zu schaffen.

[0015] Diese Aufgabe wird durch eine Vorrichtung zum Komprimieren von Schallfelddaten gemäß Anspruch 1, eine Vorrichtung zum Dekomprimieren von Schallfelddaten gemäß Anspruch 13, ein Verfahren zum Komprimieren von Schallfelddaten gemäß Anspruch 19, ein Verfahren zum Dekomprimieren von Schallfelddaten gemäß Anspruch 20 oder ein Computerprogramm gemäß Anspruch 21 gelöst.

[0016] Eine Vorrichtung zum Komprimieren von Schallfelddaten eines Gebiets umfasst einen Aufteiler zum Aufteilen der Schallfelddaten in einen ersten Anteil und in einen zweiten Anteil sowie einen nachgeordneten Umsetzer zum Umsetzen des ersten Anteils und des zweiten Anteils in harmonische Komponenten, wobei die Umsetzung so stattfindet, dass die zweite Anzahl in eine oder mehrere harmonische Komponenten mit einer zweiten Ordnung umgesetzt wird, und dass der erste Anteil

in harmonische Komponenten mit einer ersten Ordnung umgesetzt wird, wobei die erste Ordnung höher als die zweite Ordnung ist, um die komprimierten Schallfelddaten zu erhalten.

[0017] Damit wird eine Umsetzung der Schallfelddaten, wie beispielsweise der Menge an Impulsantworten in harmonische Komponenten durchgeführt, wobei bereits diese Umsetzung zu einer erheblichen Dateneinsparung führen kann. Harmonische Komponenten, wie sie beispielsweise mittels einer räumlichen Spektraltransformation erhaltbar sind, beschreiben ein Schallfeld wesentlich kompakter als Impulsantworten. Darüber hinaus ist die Ordnung der harmonischen Komponenten ohne Weiteres steuerbar. Die harmonische Komponente nullter Ordnung ist lediglich ein (ungerichtetes) Mono-Signal. Sie erlaubt noch keine Schallfeld-Richtungsbeschreibung. Dagegen erlauben die zusätzlichen harmonischen Komponenten erster Ordnung bereits eine relativ grobe Richtungsdarstellung analog zum Beamforming. Die harmonischen Komponenten zweiter Ordnung erlauben eine zusätzliche noch genauere Schallfeldbeschreibung mit noch mehr Richtungsinformation. Bei Ambisonics beispielsweise ist die Anzahl der Komponenten gleich $2n+1$, wobei n die Ordnung ist. Für die nullte Ordnung gibt es somit nur eine einzige harmonische Komponente. Für eine Umsetzung bis zur ersten Ordnung gibt es bereits drei harmonische Komponenten. Für eine Umsetzung mit fünfter Ordnung beispielsweise gibt es bereits 11 harmonische Komponenten und es hat sich herausgestellt, dass beispielsweise für 350 Impulsantworten eine Ordnung gleich 14 ausreichend ist. Dies bedeutet in anderen Worten, dass 29 harmonische Komponenten den Raum genauso gut beschreiben wie 350 Impulsantworten. Bereits diese Umsetzung von einem Wert von 350 Eingangskanälen auf 29 Ausgangskanäle bringt einen Kompressionsgewinn. Darüber hinaus wird noch eine Umsetzung verschiedener Anteile der Schallfelddaten, wie beispielsweise der Impulsantworten mit verschiedenen Ordnungen durchgeführt, da herausgefunden worden ist, dass nicht alle Anteile mit der gleichen Genauigkeit/Ordnung beschrieben werden müssen.

Ein Beispiel hierfür besteht darin, dass die Richtungswahrnehmung des menschlichen Gehörs hauptsächlich von den frühen Reflexionen abgeleitet wird, während die späten/diffusen Reflexionen in einer typischen Impulsantwort zur Richtungswahrnehmung nichts oder nur sehr wenig beitragen. Bei diesem Beispiel wird somit der erste Anteil der frühe Anteil der Impulsantworten sein, der mit einer höheren Ordnung in den Harmonischen-Komponenten-Bereich umgesetzt wird, während der späte diffuse Anteil mit einer geringen Ordnung und teilweise sogar mit einer Ordnung gleich null umgesetzt wird.

[0018] Ein anderes Beispiel besteht darin, dass die Richtungswahrnehmung des menschlichen Gehörs frequenzabhängig ist. Bei tiefen Frequenzen ist die Richtungswahrnehmung des menschlichen Gehörs relativ

schwach. Für die Kompression von Schallfelddaten genügt es daher, den niedrigen Spektralbereich der harmonischen Komponenten mit einer relativ geringen Ordnung in den Harmonische-Komponenten-Bereich umzusetzen, während die Frequenzbereiche der Schallfelddaten, in denen die Richtungswahrnehmung des menschlichen Gehörs sehr hoch ist, mit einer hohen und vorzugsweise sogar mit der maximalen Ordnung umgesetzt werden. Hierzu werden erfindungsgemäß die Schallfelddaten mittels einer Filterbank in einzelne Subband-Schallfelddaten zerlegt und diese Subband-Schallfelddaten werden dann mit unterschiedlichen Ordnungen zerlegt, wobei wiederum der erste Anteil Subband-Schallfelddaten bei höheren Frequenzen aufweist, während der zweite Anteil Subband-Schallfelddaten bei niedrigeren Frequenzen aufweist, wobei ganz niedrige Frequenzen ebenso wiederum sogar mit einer Ordnung gleich null, also nur mit einer einzigen harmonischen Komponente dargestellt werden können.

Bei einem weiteren Beispiel werden die vorteilhaften Eigenschaften der zeitlichen und frequenzmäßigen Verarbeitung kombiniert. So kann der frühe Anteil, der ohnehin mit höherer Ordnung umgesetzt wird, in Spektralanteile zerlegt werden, für die dann wieder für die einzelnen Bänder angepasste Ordnungen erhalten werden können. Insbesondere dann, wenn für die Subbandsignale eine dezimierende Filterbank, wie beispielsweise eine QMF-Filterbank (QMF = Quadrature Mirror Filterbank) eingesetzt wird, reduziert sich der Aufwand zur Umsetzung der Subband-Schallfelddaten in den Harmonische-Komponenten-Bereich zusätzlich. Darüber hinaus liefert die Differenzierung verschiedener Anteile der Schallfelddaten im Hinblick auf die zu berechnende Ordnung eine erhebliche Reduktion des Berechnungsaufwands, zumal die Berechnung der harmonischen Komponenten, wie beispielsweise der zylindrischen harmonischen Komponenten oder der sphärischen harmonischen Komponenten stark davon abhängt, bis zu welcher Ordnung die harmonischen Komponenten ausgerechnet werden sollen. Eine Berechnung der harmonischen Komponenten bis zur zweiten Ordnung beispielsweise benötigt wesentlich weniger Rechenaufwand und damit Rechenzeit bzw. Batterieleistung insbesondere bei mobilen Geräten als eine Berechnung der harmonischen Komponenten bis zur Ordnung 14 beispielsweise.

Bei den beschriebenen Ausführungsbeispielen ist der Umsetzer somit ausgebildet, um den Anteil, also den ersten Anteil der Schallfelddaten, der für eine Richtungswahrnehmung des menschlichen Gehörs wichtiger ist, mit einer höheren Ordnung umzusetzen als den zweiten Anteil, der für die Richtungswahrnehmung einer Schallquelle weniger wichtig als der erste Anteil ist.

[0019] Die vorliegende Erfindung kann nicht nur für eine zeitliche Zerlegung der Schallfelddaten in Anteile oder für eine spektrale Zerlegung der Schallfelddaten in Anteile eingesetzt werden, sondern auch für eine alternative, z. B. räumliche Zerlegung der Anteile, wenn beispielsweise berücksichtigt wird, dass die Richtungswahrnehmung

des menschlichen Gehörs für Schall in unterschiedlichen Azimuth- oder Elevations-Winkeln unterschiedlich ist. Wenn die Schallfelddaten beispielsweise als Impulsantworten oder andere Schallfeldbeschreibungen vorliegen, bei denen jeder einzelnen Beschreibung ein bestimmter Azimuth/Elevations-Winkel zugeordnet ist, so können die Schallfelddaten aus Azimuth/Elevations-Winkeln, bei denen die Richtungswahrnehmung des menschlichen Gehörs stärker ist, mit einer höheren Ordnung komprimiert werden als ein räumlicher Anteil der Schallfelddaten aus einer anderen Richtung.

[0020] Alternativ oder zusätzlich können die einzelnen Harmonischen "ausgedünnt" werden, also im Beispiel mit Ordnung 14, bei der es 29 Moden gibt. Es werden in Abhängigkeit der menschlichen Richtungswahrnehmung einzelne Moden eingespart, die das Schallfeld für unwichtige Schalleinfallsrichtungen abbilden. Im Falle von Mikrofonarraymessungen besteht hier eine Unsicherheit, weil man nicht weiss, in welche Richtung der Kopf bezogen zur Arraykugel ausgerichtet ist. Wenn man aber HRTFs mittels sphärischer Harmonischer darstellt, ist diese Unsicherheit behoben.

[0021] Weitere Zerlegungen der Schallfelddaten zusätzlich zu Zerlegungen in zeitlicher, spektraler oder räumlicher Richtung können ebenfalls verwendet werden, wie beispielsweise eine Zerlegung der Schallfelddaten in einen ersten und einen zweiten Anteil in Volumenklassen etc.

[0022] Bei Ausführungsbeispielen geschieht die Beschreibung akustischer Problemstellungen im zylindrischen oder sphärischen Koordinatensystem, also mittels vollständiger Sätze orthonormaler Eigenfunktionen, den sogenannte zylindrischen oder sphärischen harmonischen Komponenten. Mit höherer räumlicher Genauigkeit der Beschreibung des Schallfelds steigen das Datenaufkommen und die Rechenzeit bei der Verarbeitung bzw. Manipulation der Daten. Für hochqualitative Audioanwendungen sind hohe Genauigkeiten erforderlich, was zu den Problemen der langen Berechnungszeiten, die insbesondere nachteilig für Echtzeitsysteme sind, der großen Datenmengen, was die Übertragung räumlicher Schallfelddaten erschwert, und des hohen Energieverbrauchs durch intensiven Rechenaufwand insbesondere bei mobilen Geräten führt.

[0023] Alle diese Nachteile werden durch Ausführungsbeispiele der Erfindung dahin gehend gelindert bzw. eliminiert, weil aufgrund der Differenzierung der Ordnungen zur Berechnung der harmonischen Komponenten die Berechnungszeiten reduziert werden, und zwar im Vergleich zu einem Fall, bei dem sämtliche Anteile mit der höchsten Ordnung in harmonische Komponenten umgesetzt werden. Die großen Datenmengen werden dahin gehend erfindungsgemäß reduziert, dass die Darstellung durch harmonische Komponenten insbesondere bereits kompakter ist und dass zusätzlich noch unterschiedliche Anteile mit unterschiedlichen Ordnungen dargestellt werden, wobei die Datenmengenreduktion dadurch erreicht wird, dass eine niedrige Ordnung,

wie beispielsweise die erste Ordnung nur drei harmonische Komponenten hat, während die höchste Ordnung beispielsweise 29 harmonische Komponenten hat, und zwar am Beispiel einer Ordnung von 14.

[0024] Die reduzierte Rechenleistung und das reduzierte Speicheraufkommen verringern automatisch den Energieverbrauch, der insbesondere für einen Einsatz von Schallfelddaten in mobilen Geräten anfällt.

[0025] Bei Ausführungsbeispielen wird die räumliche Schallfeldbeschreibung im zylindrischen bzw. sphärischen harmonischen Bereich basierend auf der räumlichen Wahrnehmung des Menschen optimiert. Insbesondere eine Kombination aus zeit- und frequenzabhängiger Berechnung der Ordnung sphärischer Harmonischer in Abhängigkeit der räumlichen Wahrnehmung des menschlichen Gehörs führt zu einer erheblichen Aufwandsreduktion ohne die subjektive Qualität der Schallfeldwahrnehmung zu reduzieren. Selbstverständlich wird die objektive Qualität reduziert, da die vorliegende Erfindung eine verlustbehaftete Kompression darstellt. Diese verlustbehaftete Kompression ist jedoch unkritisch, zumal der letztendliche Empfänger das menschliche Gehör ist, und zumal es daher sogar für eine transparente Wiedergabe unerheblich ist, ob Schallfeldkomponenten, die vom menschlichen Gehör ohnehin nicht wahrgenommen werden, im wiedergegebenen Schallfeld vorhanden sind oder nicht.

[0026] In anderen Worten stellt daher bei der Wiedergabe/Auralisation entweder binaural, also mit Kopfhörern oder mit Lautsprechersystemen mit wenigen (z.B. Stereo) oder vielen Lautsprechern (z.B. WFS) das menschlichen Gehör das wichtigste Qualitätsmaß dar. Erfindungsgemäß wird die Genauigkeit der harmonischen Komponenten wie beispielsweise der zylindrischen oder sphärischen Harmonischen im Zeitbereich und/oder im Frequenzbereich und/oder in weiteren Bereichen gehörangepasst reduziert. Dadurch wird die Daten- und Rechenzeitreduktion erreicht.

[0027] Bevorzugte Ausführungsbeispiele der vorliegenden Erfindung werden nachfolgend Bezug nehmend auf die beiliegenden Zeichnungen detailliert erläutert. Es zeigen:

Fig. 1a ein Blockdiagramm einer Vorrichtung zum Komprimieren von Schallfelddaten gemäß einem Ausführungsbeispiel;

Fig. 1b ein Blockdiagramm einer Vorrichtung zum Dekomprimieren von komprimierten Schallfelddaten eines Gebiets;

Fig. 1c ein Blockdiagramm einer Vorrichtung zum Komprimieren mit zeitlicher Zerlegung;

Fig. 1d ein Blockdiagramm eines Ausführungsbeispiels einer Vorrichtung zum Dekomprimieren für den Fall einer zeitlichen Zerlegung;

Fig. 1e eine zu Fig. 1d alternative Vorrichtung zum Dekomprimieren;

5 Fig. 1f ein Beispiel für die Anwendung der Erfindung mit zeitlicher und spektraler Zerlegung am Beispiel von 350 gemessenen Impulsantworten als Schallfelddaten;

10 Fig. 2a ein Blockdiagramm einer Vorrichtung zum Komprimieren mit spektraler Zerlegung;

Fig. 2b ein Beispiel einer unterabgetasteten Filterbank und einer anschließenden Umsetzung der unterabgetasteten Subband-Schallfelddaten;

15 Fig. 2c eine Vorrichtung zum Dekomprimieren für das in Fig. 2a gezeigte Beispiel der spektralen Zerlegung;

20 Fig. 2d eine alternative Implementierung des Dekomprimierers für die spektrale Zerlegung;

25 Fig. 3a ein Übersichts-Blockdiagramm mit einem speziellen Analyse/Synthesecodierer gemäß einem weiteren Ausführungsbeispiel der vorliegenden Erfindung;

30 Fig. 3b eine detailliertere Darstellung eines Ausführungsbeispiels mit zeitlicher und spektraler Zerlegung;

Fig. 4 eine schematische Darstellung einer Impulsantwort;

35 Fig. 5 ein Blockschaltbild eines Umsetzers vom Zeit- oder Spektralbereich in den Harmonischen-Komponenten-Bereich mit variabler Ordnung; und

40 Fig. 6 eine Darstellung eines beispielhaften Umsetzers vom Harmonischen-Komponenten-Bereich in den Zeitbereich oder Spektralbereich mit anschließender Auralisation.

[0028] Fig. 1a zeigt ein Blockschaltbild einer Vorrichtung bzw. eines Verfahrens zum Komprimieren von Schallfelddaten eines Gebiets, wie sie an einem Eingang 10 in einen Aufteiler 100 eingegeben werden. Der Aufteiler 100 ist ausgebildet, um die Schallfelddaten in einen ersten Anteil 101 und einen zweiten Anteil 102 aufzuteilen. Darüber hinaus ist ein Umsetzer vorgesehen, der die zwei Funktionalitäten aufweist, die mit 140 oder 180 bezeichnet sind. Insbesondere ist der Umsetzer ausgebildet, um den ersten Anteil 101 umzusetzen, wie es bei 140 gezeigt ist, und um den zweiten Anteil 102 umzusetzen, wie es bei 180 gezeigt ist. Der Umsetzer setzt insbesondere den ersten Anteil 101 in eine oder mehrere

harmonische Komponenten 141 mit einer ersten Ordnung um, während der Umsetzer 180 den zweiten Anteil 102 in eine oder mehrere harmonische Komponenten 182 mit einer zweiten Ordnung umsetzt. Insbesondere ist die erste Ordnung, also die den harmonischen Komponenten 141 zugrunde liegende Ordnung höher als die zweite Ordnung, was, in anderen Worten ausgedrückt bedeutet, dass der Umsetzer 140 mit höherer Ordnung mehr harmonische Komponenten 141 ausgibt als der Umsetzer 180 mit niedrigerer Ordnung. Die Ordnung n_1 , durch die der Umsetzer 140 angesteuert wird, ist somit größer als die Ordnung n_2 , mit der der Umsetzer 180 angesteuert wird. Die Umsetzer 140, 180 können steuerbare Umsetzer sein. Alternativ kann die Ordnung jedoch festgelegt sein und damit fest einprogrammiert sein, so dass die Eingänge, die mit n_1 und n_2 bezeichnet sind, in diesem Ausführungsbeispiel nicht vorhanden sind.

[0029] Fig. 1b zeigt eine Vorrichtung zum Dekomprimieren von komprimierten Schallfelddaten 20, die erste harmonische Komponenten mit einer ersten Ordnung und eine oder mehrere zweite harmonische Komponenten mit einer zweiten Ordnung aufweisen, wie sie z.B. von Fig. 1a bei 141, 182 ausgegeben werden. Die dekomprimierten Schallfelddaten müssen jedoch nicht unbedingt die harmonischen Komponenten 141, 142 im "Rohformat" sein. Stattdessen könnte in Fig. 1a noch ein verlustloser Entropie-Codierer, wie beispielsweise ein Huffman-Codierer oder ein arithmetischer Codierer vorgesehen sein, um die Anzahl von Bits, die letztendlich zur Darstellung der harmonischen Komponenten benötigt werden, weiter zu reduzieren. Dann würde der Datenstrom 20, der in eine Eingangs-Schnittstelle 200 eingespeist wird, aus Entropie-codierten harmonischen Komponenten und gegebenenfalls Seiteninformationen bestehen, wie es noch anhand von Fig. 3a dargestellt wird. In diesem Fall wäre am Ausgang der Eingangs-Schnittstelle 200 ein jeweiliger Entropie-Decodierer vorgesehen, der an den Entropie-Encodierer auf Encodierer-Seite, also bezüglich Fig. 1a angepasst ist. So stellen somit die ersten harmonischen Komponenten der ersten Ordnung 201 und die zweiten harmonischen Komponenten der zweiten Ordnung 202, wie sie in Fig. 1b dargestellt sind, gegebenenfalls noch Entropiecodierte oder aber bereits Entropie-decodierte oder tatsächlich die harmonischen Komponenten in "Rohform", wie sie bei 141, 182 in Fig. 1a vorliegen, dar.

[0030] Beide Gruppen von harmonischen Komponenten werden in einen Decodierer bzw. Umsetzer/Kombinierer 240 eingespeist. Der Block 240 ist ausgebildet, um die komprimierten Schallfelddaten 201, 202 unter Verwendung einer Kombination des ersten Anteils und des zweiten Anteils und unter Verwendung einer Umsetzung von einer Harmonische-Komponenten-Darstellung in eine Zeitbereichsdarstellung zu dekomprimieren, um schließlich die dekomprimierte Darstellung des Schallfelds zu erhalten, wie es bei 240 dargestellt ist. Der Decodierer 240, der beispielsweise als Signalprozessor ausgebildet sein kann, ist somit ausgebildet, um zum ei-

nen eine Umsetzung in den Zeitbereich vom Sphärische-Harmonische-Komponentenbereich durchzuführen, und um zum anderen eine Kombination durchzuführen. Die Reihenfolge zwischen Umsetzung und Kombination kann jedoch unterschiedlich sein, wie es im Hinblick auf Fig. 1d, Fig. 1e oder Fig. 2c, 2d für verschiedene Beispiele dargestellt wird.

[0031] Fig. 1c zeigt eine Vorrichtung zum Komprimieren von Schallfelddaten eines Gebiets gemäß einem Ausführungsbeispiel, bei dem der Aufteiler 100 als zeitlicher Aufteiler 100a ausgebildet ist. Insbesondere ist der zeitliche Aufteiler 100a, der eine Implementierung des Aufteilers 100 von Fig. 1a ist, ausgebildet, um die Schallfelddaten in einen ersten Anteil, der erste Reflexionen in dem Gebiet umfasst, und in einen zweiten Anteil, der zweite Reflexionen in dem Gebiet umfasst, aufzuteilen, wobei die zweiten Reflexionen zeitlich später als die ersten Reflexionen auftreten. Anhand von Fig. 4 stellt der erste Anteil 101, der vom Block 100a ausgegeben wird, somit den Impulsantwort-Abschnitt 310 von Fig. 4 dar, während der zweite späte Anteil den Abschnitt 320 der Impulsantwort von Fig. 4 darstellt. Der Zeitpunkt der Aufteilung kann beispielsweise bei 100 ms liegen. Allerdings existieren auch andere Möglichkeiten der zeitlichen Aufteilung, wie beispielsweise früher oder später. Vorzugsweise wird die Aufteilung dort gelegt, wo die diskreten Reflexionen in diffuse Reflexionen übergehen. Dies kann je nach Raum ein unterschiedlicher Zeitpunkt sein, und es existieren Konzepte, um hier eine beste Aufteilung zu schaffen. Andererseits kann die Aufteilung in einen frühen und in einen späten Anteil auch abhängig von einer verfügbaren Datenrate durchgeführt werden, dahin gehend, dass die Aufteilungszeit immer kleiner gemacht wird, je weniger Bitrate vorhanden ist. Dies ist im Hinblick auf die Bitrate günstig, weil dann ein möglichst großer Anteil der Impulsantwort mit einer niedrigen Ordnung in den Harmonische-Komponenten-Bereich umgesetzt wird.

[0032] Der Umsetzer, der durch die Blöcke 140 und 180 in Fig. 1c dargestellt ist, ist somit ausgebildet, um den ersten Anteil 101 und den zweiten Anteil 102 in harmonische Komponenten umzusetzen, wobei der Umsetzer insbesondere den zweiten Anteil in eine oder mehrere harmonische Komponenten 182 mit einer zweiten Ordnung umsetzt und den ersten Anteil 101 in harmonische Komponenten 141 mit einer ersten Ordnung umsetzt, wobei die erste Ordnung höher als die zweite Ordnung ist, um schließlich das komprimierte Schallfeld zu erhalten, das von einer Ausgabeschnittstelle 190 schließlich zu Zwecken der Übertragung und/oder Speicherung ausgabbar ist.

[0033] Fig. 1d zeigt eine Implementierung des Dekomprimierers für das Beispiel der zeitlichen Aufteilung. Insbesondere ist der Dekomprimierer ausgebildet, um die komprimierten Schallfelddaten unter Verwendung einer Kombination des ersten Anteils 201 mit den ersten Reflexionen und des zweiten Anteils 202 mit den späten Reflexionen und einer Umsetzung von dem Harmoni-

sche-Komponenten-Bereich in den Zeitbereich durchzuführen. Fig. 1d zeigt eine Implementierung, bei der die Kombination nach der Umsetzung stattfindet. Fig. 1e zeigt eine alternative Implementierung, bei der die Kombination vor der Umsetzung stattfindet. Insbesondere ist der Umsetzer 241 ausgebildet, um harmonische Komponenten mit der hohen Ordnung in den Zeitbereich umzusetzen, während der Umsetzer 242 ausgebildet ist, um die harmonischen Komponenten mit der niedrigen Ordnung in den Zeitbereich umzusetzen. Im Hinblick auf Fig. 4 liefert somit der Ausgang des Umsetzers 241 etwas, das dem Bereich 210 entspricht, während der Umsetzer 242 etwas liefert, das dem Bereich 320 entspricht, wobei jedoch aufgrund der verlustbehafteten Kompression die Abschnitte am Ausgang der Brücke 241, 242 nicht identisch zu den Abschnitten 310, 320 sind. Insbesondere wird jedoch eine zumindest wahrnehmungsmäßige Ähnlichkeit oder Identität des Abschnitts am Ausgang des Blocks 241 zu dem Abschnitt 310 von Fig. 4 bestehen, während der Abschnitt am Ausgang des Blocks 242, der dem späten Anteil 320 der Impulsantwort entspricht, deutliche Unterschiede haben wird und somit den Verlauf der Impulsantwort lediglich näherungsweise darstellt. Diese Abweichungen sind jedoch für die menschliche Richtungswahrnehmung unkritisch, weil die menschliche Richtungswahrnehmung ohnehin kaum oder nicht auf dem späten Anteil bzw. den diffusen Reflexionen der Impulsantwort basiert.

[0034] Fig. 1e zeigt eine alternative Implementierung, bei der der Decodierer zunächst den Kombiner 245 und dann nachfolgend den Umsetzer 244 aufweist. Die einzelnen harmonischen Komponenten werden bei dem in Fig. 1e gezeigten Ausführungsbeispiel aufaddiert, woraufhin das Ergebnis der Aufaddition umgesetzt wird, um schließlich eine Zeitbereichsdarstellung zu erhalten. Im Unterschied hierzu wird bei der Ausführung in Fig. 1d eine Kombination nicht in einer Aufaddition bestehen, sondern in einer Serialisierung, dahin gehend, dass der Ausgang des Blocks 241 zeitlich früher in einer dekomprimierten Impulsantwort angeordnet sein wird als der Ausgang des Blocks 242, um wieder eine zu Fig. 4 entsprechende Impulsantwort zu erhalten, die dann für weitere Zwecke, wie beispielsweise eine Auralisation also in eine Aufbereitung von Tonsignalen mit dem gewünschten Raumeindruck verwendet werden kann.

[0035] Fig. 2a zeigt eine alternative Implementierung der vorliegenden Erfindung, bei der eine Aufteilung im Frequenzbereich vorgenommen wird. Insbesondere ist der Aufteiler 100 von Fig. 1a bei dem Ausführungsbeispiel von Fig. 2a als Filterbank implementiert, um zumindest einen Teil der Schallfelddaten zu filtern, um Schallfelddaten in verschiedenen Filterbankkanälen 101, 102 zu erhalten. Die Filterbank erhält bei einem Ausführungsbeispiel, bei dem die zeitliche Aufteilung von Fig. 1a nicht implementiert ist, sowohl den frühen als auch den späten Anteil, während bei einem alternativen Ausführungsbeispiel lediglich der frühe Anteil der Schallfelddaten in die Filterbank eingespeist wird, während der späte Anteil

nicht weiter spektral zerlegt wird.

[0036] Der Analyse-Filterbank 100b nachgeordnet ist der Umsetzer, der aus Teil-Umsetzern 140a, 140b, 140c ausgebildet sein kann. Der Umsetzer 140a, 140b, 140c ist ausgebildet, um die Schallfelddaten in verschiedenen Filterbankkanälen unter Verwendung verschiedener Ordnungen für verschiedene Filterbankkanäle umzusetzen, um für jeden Filterbankkanal einen oder mehrere harmonische Komponenten zu erhalten. Insbesondere ist der Umsetzer ausgebildet, um für einen ersten Filterbankkanal mit einer ersten Mittenfrequenz eine Umsetzung mit einer ersten Ordnung durchzuführen, und um für einen zweiten Filterbankkanal mit einer zweiten Mittenfrequenz eine Umsetzung mit einer zweiten Ordnung durchzuführen, wobei die erste Ordnung höher als die zweite Ordnung ist, und wobei die erste Mittenfrequenz, d.h. f_n , höher als die zweite Mittenfrequenz f_1 ist, um schließlich die komprimierte Schallfelddarstellung zu erhalten. Generell kann, je nach Ausführungsbeispiel, für das niedrigste Frequenzband eine niedrigere Ordnung als für ein mittleres Frequenzband verwendet werden. Allerdings muss je nach Implementierung das höchste Frequenzband, wie es bei dem in Fig. 2a gezeigten Ausführungsbeispiel der Filterbankkanal mit der Mittenfrequenz f_n ist, nicht unbedingt mit einer höheren Ordnung als z.B. ein mittlerer Kanal umgesetzt werden. Stattdessen kann in den Bereichen, in denen die Richtungswahrnehmung am höchsten ist, die höchste Ordnung verwendet werden, wie in den anderen Bereichen, zu denen auch z.B. ein bestimmter hoher Frequenzbereich gehören kann, die Ordnung niedriger ist, weil in diesen Bereichen auch die Richtungswahrnehmung des menschlichen Gehörs niedriger ist.

[0037] Fig. 2b zeigt eine detailliertere Implementierung der Analyse-Filterbank 100b. Diese umfasst bei dem in Fig. 2b gezeigten Ausführungsbeispiel einen Bandfilter und hat ferner nachgeschaltete Dezimierer 100c für jeden Filterbankkanal. Wenn beispielsweise eine Filterbank bestehend aus Bandfilter und Dezimierern eingesetzt wird, die 64 Kanäle hat, so kann jeder Dezimierer mit einem Faktor 1/64 dezimieren, so dass insgesamt die Anzahl der digitalen Abtastwerte am Ausgang der Dezimierer über alle Kanäle aufaddiert der Anzahl von Abtastwerten eines Blocks der Schallfelddaten im Zeitbereich entspricht, der durch die Filterbank zerlegt worden ist. Eine beispielhafte Filterbank kann eine reelle oder komplexe QMF-Filterbank sein. Jedes Subbandsignal vorzugsweise der frühen Anteile der Impulsantworten wird dann in harmonische Komponenten mittels der Umsetzer 140a bis 140c in Analogie zu Fig. 2a umgesetzt, um schließlich für verschiedene Subbandsignale der Schallfeldbeschreibung eine Beschreibung mit zylindrischen oder vorzugsweise sphärischen harmonischen Komponenten zu erhalten, die für unterschiedliche Subbandsignale unterschiedliche Ordnungen, also eine unterschiedlich große Anzahl von harmonischen Komponenten, aufweist.

[0038] Fig. 2c und Fig. 2d zeigen wieder verschiedene

Implementierungen des Dekomprimierers, wie er in Fig. 1b dargestellt ist, also eine unterschiedliche Reihenfolge der Kombination und anschließenden Umsetzung in Fig. 2c oder der zunächst durchgeführten Umsetzung und der anschließenden Kombination, wie es in Fig. 2d dargestellt ist. Insbesondere umfasst der Dekomprimierer 240 von Fig. 1b bei dem in Fig. 2c gezeigten Ausführungsbeispiel wieder einen Kombinierer 245, der eine Addition der unterschiedlichen harmonischen Komponenten aus den verschiedenen Subbändern durchführt, um dann eine Gesamtdarstellung der harmonischen Komponenten zu erhalten, die dann mit dem Umsetzer 244 in den Zeitbereich umgesetzt werden. Damit liegen die Eingangssignale in dem Kombinierer 245 im Harmonische-Komponenten-Spektralbereich vor, während der Ausgang des Kombinierers 345 eine Darstellung im Harmonische-Komponenten-Bereich darstellt, von dem dann durch den Umsetzer 244 eine Umsetzung in den Zeitbereich erhalten wird.

[0039] Bei dem in Fig. 2d gezeigten alternativen Ausführungsbeispiel werden die einzelnen harmonischen Komponenten für jedes Subband zunächst durch unterschiedliche Umsetzer 241a, 241b, 241c in den Spektralbereich umgesetzt, so dass die Ausgangssignale der Blöcke 241a, 241b, 241c den Ausgangssignalen der Blöcke 140a, 140b, 140c von Fig. 2a oder Fig. 2b entsprechen. Dann werden diese Subbandsignale in einer nachgeordneten Synthese-Filterbank, die im Falle eines Downsampling auf Encodierer-Seite (Block 100c von Fig. 2b) auch eine Hoch-Tast-Funktion, also eine Upsampling-Funktion aufweisen kann, verarbeitet. Die Synthese-Filterbank stellt dann die Kombinierer-Funktion des Decodierers 240 von Fig. 1b dar. Am Ausgang der Synthese-Filterbank liegt somit die dekomprimierte Schallfelddarstellung vor, die zur Auralisation verwendet werden kann, wie es noch dargestellt wird.

[0040] Fig. 1f zeigt ein Beispiel für die Zerlegung von Impulsantworten in harmonische Komponenten verschiedener Ordnungen. Die späten Abschnitte werden nicht spektral zerlegt sondern insgesamt mit der nullten Ordnung umgesetzt. Die frühen Abschnitte der Impulsantworten werden spektral zerlegt. Das niedrigste Band wird beispielsweise mit der ersten Ordnung verarbeitet, während das nächste Band bereits mit der fünften Ordnung verarbeitet wird und das letzte Band, weil es für die Richtungs/Raumwahrnehmung am wichtigsten ist, mit der höchsten Ordnung, also bei diesem Beispiel mit der Ordnung 14, verarbeitet wird.

[0041] Fig. 3a zeigt das gesamte Encodierer/Decodierer-Schema oder das gesamte Komprimierer/Dekomprimierer-Schema der vorliegenden Erfindung.

[0042] Insbesondere umfasst bei dem in Fig. 3a gezeigten Ausführungsbeispiel der Komprimierer nicht nur die Funktionalitäten von Fig. 1a, die mit 1 oder PENC bezeichnet sind, sondern auch einen Decodierer PDEC2, der wie in Fig. 1b ausgebildet sein kann. Darüber hinaus umfasst der Komprimierer noch eine Steuerung CTRL4, die ausgebildet ist, um vom Decodierer 2 erhal-

tene dekomprimierte Schallfelddaten mit ursprünglichen Schallfelddaten unter Berücksichtigung eines psychoakustischen Modells, wie beispielsweise dem Modell PEAQ, das von der ITU standardisiert worden ist, vergleichen.

[0043] Daraufhin erzeugt die Steuerung 4 optimierte Parameter für die Aufteilung, wie beispielsweise die zeitliche Aufteilung, die frequenzmäßige Aufteilung in der Filterbank oder optimierte Parameter für die Ordnungen in den einzelnen Umsetzern für die verschiedenen Anteile der Schallfelddaten, wenn diese Umsetzer steuerbar ausgebildet sind.

[0044] Steuerparameter, wie beispielsweise Aufteilungsinformationen, Filterbankparameter oder Ordnungen können dann zusammen mit einem Bitstrom, der die harmonischen Komponenten aufweist, zu einem Decodierer bzw. Dekomprimierer übertragen werden, der mit 2 in Fig. 3a dargestellt ist. Der Komprimierer 11 besteht somit aus dem Kontrollblock CTRL4 für die Codec-Steuerung sowie einen Parameter-Codierer PENC1 und dem Parameter-Decodierer PDEC2. Die Eingaben 10 sind Daten von Mikrofon-Array-Messungen. Der Steuerblock 4 initialisiert den Encodierer 1 und stellt sämtliche Parameter für die Encodierung der Array-Daten bereit. Im PENC-Block 1 werden die Daten gemäß der beschriebenen Methodik der gehörahängigen Aufteilung im Zeit- und im Frequenzbereich verarbeitet und für die Datenübertragung bereitgestellt.

[0045] Fig. 3b zeigt das Schema der Daten-En- und Decodierung. Die Eingangsdaten 10 werden zunächst durch den Aufteiler 100a in ein frühes 101 und ein spätes Schallfeld 102 zerlegt. Das frühe Schallfeld 101 wird mittels einer n-Band-Filterbank 100b in seine spektralen Anteile $f_1 \dots f_n$ zerlegt, die jeweils mit einer dem menschlichen Gehör angepassten Ordnung des sphärischen Harmonischen (x-Ordnung-SHD - SHD = Spherical Harmonics Decomposition) zerlegt werden. Diese Zerlegung in sphärische Harmonische stellt ein bevorzugtes Ausführungsbeispiel dar, wobei jedoch mittels jeder Schallfeldzerlegung (Sound Field Decomposition), die harmonische Komponenten erzeugt, gearbeitet werden kann. Da die Zerlegung in sphärische harmonische Komponenten in jedem Band je nach Ordnung unterschiedlich lange Berechnungszeiten braucht, wird es bevorzugt, die Zeitversätze in einer Verzögerungsleitung mit Verzögerungsblöcken 306, 304 zu korrigieren. Damit wird der Frequenzbereich im Rekonstruktionsblock 245, der auch als Kombinierer bezeichnet wird, rekonstruiert und mit dem späten Schallfeld im weiteren Kombinierer 243 wieder kombiniert, nachdem dieses mit einer gehörangepasst niedrigen Ordnung gerechnet wurde.

[0046] Der Steuerblock CTRL4 von Fig. 3a beinhaltet ein raumakustisches Analysemodul und ein Psychoakustikmodul. Der Steuerblock analysiert dabei sowohl die Eingangsdaten 10 als auch die Ausgangsdaten des Decodierers 2 von Fig. 3a, um die Codierparameter, die auch als Seiteninformationen 300 in Fig. 3a bezeichnet werden, bzw. die direkt im Komprimierer 11 dem Codie-

rer PENC1 bereitgestellt werden, adaptiv anzupassen. Aus den Eingangssignalen 10 werden raumakustische Parameter extrahiert, die mit den Parametern der verwendeten Array-Konfiguration die initialen Parameter der Codierung vorgeben. Diese beinhalten sowohl den Zeitpunkt der Trennung zwischen frühem und spätem Schallfeld, der auch als "mixing time" oder "Mischzeit" bezeichnet wird, als auch die Parameter für die Filterbank, wie beispielsweise entsprechende Ordnungen der sphärischen Harmonischen. Die Ausgabe, die z.B. in Form von binauralen Impulsantworten vorliegen kann, wie sie vom Kombinerer 243 ausgegeben wird, wird in ein psychoakustisches Modul mit einem auditorischen Modell geführt, das die Qualität evaluiert und die Codierparameter entsprechend anpasst. Alternativ kann das Konzept auch mit statischen Parametern arbeiten. Dann entfällt das Steuermodul CTRL4 sowie das PEDC-Modul 2 auf Encodierer- bzw. Komprimiererseite 11.

[0047] Die Erfindung ist dahin gehend vorteilhaft, dass Daten und Rechenaufwand bei der Verarbeitung und Übertragung von Kreis- und Kugelarraydaten in Abhängigkeit des menschlichen Gehörs reduziert werden. Es ist ferner vorteilhaft, dass die so verarbeiteten Daten in bestehende Kompressionsverfahren integriert werden können und somit eine zusätzliche Datenreduktion erlauben. Dies ist in bandbegrenzten Übertragungssystemen, wie für mobile Endgeräte von Vorteil. Ein weiterer Vorteil ist die mögliche Echtzeitverarbeitung der Daten im sphärischen harmonischen Bereich auch bei hohen Ordnungen. Die vorliegende Erfindung kann in vielen Bereichen Anwendung finden, und insbesondere in den Bereichen, wo das akustische Schallfeld mittels zylindrischer oder sphärischer Harmonischer dargestellt wird. Dies erfolgt z.B. bei der Schallfeldanalyse mittels Kreis- bzw. Kugelarrays. Wenn das analysierte Schallfeld auralisiert werden soll, kann das Konzept der vorliegenden Erfindung eingesetzt werden. Bei Geräten zur Simulation von Räumen werden Datenbanken zur Speicherung existierender Räume eingesetzt. Hier ermöglicht das erfindungsgemäße Konzept eine platzsparende und qualitativ hochwertige Speicherung. Es existieren Wiedergabeverfahren, die auf Kugelflächenfunktionen beruhen, wie beispielsweise Higher Order Ambisonics oder die binaurale Synthese. Hier liefert die vorliegende Erfindung eine Reduktion von Rechenzeit und Datenaufwand. Dies kann insbesondere im Hinblick auf die Datenübertragung z.B. bei Telekonferenz-Systemen von besonderem Vorteil sein.

[0048] Fig. 5 zeigt eine Implementierung eines Umsetzers 140 bzw. 180 mit einstellbarer Ordnung bzw. mit zumindest unterschiedlicher Ordnung, die auch fest eingestellt sein kann.

[0049] Der Umsetzer umfasst einen Zeit-Frequenz-Transformationsblock 502 und einen nachgeschalteten Raum-Transformationsblock 504. Der Raum-Transformationsblock 504 ist ausgebildet, um gemäß der Berechnungsvorschrift 508 zu arbeiten. In der Berechnungsvorschrift beträgt n die Ordnung. Die Berechnungsvorschrift

508 wird je nach Ordnung nur einmal gelöst, wenn die Ordnung gleich null ist, oder wird öfter gelöst, wenn die Ordnung z.B. bis zur Ordnung 5 oder bei dem vorstehend beschriebenen Beispiel bis zur Ordnung 14 geht. Insbesondere ist das Zeit-Frequenz-Transformationselement 502 ausgebildet, um die Impulsantworten auf den Eingangsleitungen 101, 102 in den Frequenzbereich zu transformieren, wobei vorzugsweise die schnelle Fourier-Transformation eingesetzt wird. Ferner wird dann nur das halbseitige Spektrum weitergeleitet, um den Rechenaufwand zu reduzieren. Dann wird eine räumliche Fourier-Transformation im Block Raumtransformation 504 durchgeführt, wie sie in dem Fachbuch Fourier Acoustics, Sound Radiation and Nearfield Acoustical Holography, Academic Press, 1999 von Earl G. Williams beschrieben ist. Vorzugsweise ist die Raumtransformation 504 optimiert für die Schallfeldanalyse und liefert gleichzeitig eine hohe numerische Genauigkeit und eine schnelle Berechnungsgeschwindigkeit.

[0050] Fig. 6 zeigt die bevorzugte Implementierung eines Umsetzers vom Harmonische-Komponenten-Bereich in den Zeitbereich, wo als Alternative ein Prozessor zur Zerlegung in ebene Wellen und Beamforming 602 dargestellt ist, und zwar alternativ zu einer inversen Raumtransformationsimplementierung 604. Die Ausgangssignale beider Blöcke 602, 604 können alternativ in einen Block 606 zur Erzeugung von Impulsantworten eingespeist werden. Die inverse Raumtransformation 604 ist ausgebildet, um die Hin-Transformation im Block 504 rückgängig zu machen. Alternativ hierzu führt die Zerlegung in ebene Wellen und das Beamforming in Block 606 dazu, dass eine große Menge an Zerlegungsrichtungen gleichmäßig verarbeitet werden können, was zu einer schnellen Verarbeitung insbesondere zur Visualisierung oder Auralisierung günstig ist. Vorzugsweise erhält der Block 602 radiale Filterkoeffizienten sowie, je nach Implementierung, zusätzliche Strahlformungs- bzw. Beamforming-Koeffizienten. Diese können entweder eine konstante Gerichtetheit haben oder frequenzabhängig sein. Alternative Eingangssignale in den Block 602 können modale Radialfilter sein, und insbesondere für sphärische Arrays oder unterschiedliche Konfigurationen, wie beispielsweise eine offene Kugel mit omnidirektionalen Mikrofonen, eine offene Kugel mit Cardioid-Mikrofonen und eine starre Kugel mit omnidirektionalen Mikrofonen. Der Block 606 zur Erzeugung von Impulsantworten erzeugt Impulsantworten oder Zeitbereichssignale aus Daten entweder vom Block 602 oder vom Block 604. Dieser Block rekombiniert insbesondere die vorher weggelassene negativen Anteile des Spektrums, führt eine schnelle inverse Fourier-Transformation durch und erlaubt ein Resampling bzw. Abtastumsetzung auf die ursprüngliche Abtast-Rate, falls das Eingangssignal an einer Stelle herunter-abgetastet wurde. Ferner kann eine Fensteroption eingesetzt werden.

[0051] Details zu der Funktionalität der Blöcke 502, 504, 602, 604, 606 sind in der Fachveröffentlichung "So-fia Sound Field Analysis Toolbox" von Bernschütz u.a.,

ICSA - International Conference on Spatial Audio, Detmold, 10. bis 13.11.2011" beschrieben, wobei diese Fachveröffentlichung hierin durch Bezugnahme in ihrer Gesamtheit aufgenommen ist.

[0052] Der Block 606 kann ferner ausgebildet sein, um den kompletten Satz von dekomprimierten Impulsantworten, also beispielsweise den verlustbehafteten Impulsantworten auszugeben, wobei dann der Block 608 wieder beispielsweise 350 Impulsantworten ausgeben würde. Je nach Auralisierung wird es jedoch bevorzugt, lediglich die letztendlich für die Wiedergabe nötigen Impulsantworten auszugeben, was durch einen Block 608 bewerkstelligt werden kann, der eine Auswahl oder eine Interpolation für ein bestimmtes Wiedergabeszenario liefert. Wird beispielsweise eine Stereowiedergabe angestrebt, wie es in Block 616 dargestellt ist, so wird abhängig von der Platzierung der beiden Stereo-Lautsprecher von den 350 beispielsweise wiedergewonnenen Impulsantworten die Impulsantwort ausgewählt, die jeweils der Raumrichtung des entsprechenden Stereo-Lautsprechers entspricht. Mit dieser Impulsantwort wird dann ein Vorfilter des entsprechenden Lautsprechers eingestellt, derart, dass das Vorfilter eine Filtercharakteristik hat, die dieser Impulsantwort entspricht. Dann wird ein wiederzugebendes Audiosignal zu den beiden Lautsprechern über die entsprechenden Vorfilter geführt und wiedergegeben, um schließlich den gewünschten Raumeindruck für eine Stereo-Auralisation zu erzeugen.

[0053] Existiert unter den zur Verfügung stehenden Impulsantworten eine Impulsantwort in einer bestimmten Richtung, in der im tatsächlichen Wiedergabeszenario ein Lautsprecher angeordnet ist, nicht, so werden die vorzugsweise zwei oder drei am nächsten benachbarten Impulsantworten verwendet und es wird eine Interpolation durchgeführt.

[0054] Bei einem alternativen Ausführungsbeispiel, bei dem die Wiedergabe bzw. Auralisation durch eine Wellenfeldsynthese 612 stattfindet, wird es bevorzugt, eine Wiedergabe von frühen und späten Reflexionen über virtuelle Quellen durchzuführen, wie es in der PH. D.-Arbeit "Spatial Sound Design based on Measured Room Impulse Responses" von Frank Melchior an der TU Delft aus dem Jahr 2011 detailliert dargestellt ist, wobei diese Fachveröffentlichung ebenfalls durch Bezugnahme in ihrer Gesamtheit hierin aufgenommen ist.

[0055] Insbesondere werden die Reflexionen einer Quelle bei der Wellefeldsynthese-Wiedergabe 612 durch vier Impulsantworten an bestimmten Positionen für die frühen Reflexionen und 8 Impulsantworten an bestimmten Positionen für die späten Reflexionen wiedergegeben. Der Auswahlblock 608 wählt dann die 12 Impulsantworten für die 12 virtuellen Positionen aus. Hierauf werden diese Impulsantworten zusammen mit den zugehörigen Positionen in einem Wellenfeld-Synthese-Renderer, der in Block 612 angeordnet sein kann, zugeführt, und der Wellenfeld-Synthese-Renderer berechnet dann unter Verwendung dieser Impulsantworten die Lautsprecher signale für die tatsächlich vorhandenen

Lautsprecher, damit diese dann die entsprechenden virtuellen Quellen abbilden. Damit wird für jeden Lautsprecher im Wellefeld-Synthese-Wiedergabesystem ein eigenes Vorfilter berechnet, über das dann ein letztendlich wiederzugebendes Audiosignal gefiltert wird, bevor es vom Lautsprecher ausgegeben wird, um eine entsprechende Wiedergabe mit hohen qualitativen Raumeffekten zu erreichen.

[0056] Eine alternative Implementierung der vorliegenden Erfindung besteht in einer Erzeugung eines Kopfhörersignals, also in einer Binaural-Anwendung, bei der der Raumeindruck des Gebiets über die Kopfhörerwiedergabe erzeugt werden soll.

[0057] Obgleich im Vorstehenden hauptsächlich Impulsantworten als Schallfelddaten dargestellt worden sind, können beliebige andere Schallfelddaten, beispielsweise Schallfelddaten nach Betrag und Vektor also im Hinblick auf z.B. Schalldruck und Schallschnelle an bestimmten Positionen im Raum ebenfalls eingesetzt werden. Auch diese Schallfelddaten können in wichtigere und weniger wichtigere Anteile im Hinblick auf die menschliche Richtungswahrnehmung aufgeteilt und in harmonische Komponenten umgesetzt werden. Die Schallfelddaten können auch jede Art von Impulsantworten, wie beispielsweise Head-Related Transfer Functions (HRTF-) Funktionen oder Binaural Room Impulse Responses (BRIR-) Funktionen oder Impulsantworten von jeweils einem diskreten Punkt zu einer vorbestimmten Position in dem Gebiet umfassen.

[0058] Vorzugsweise wird ein Raum mit einem Kugelarray abgetastet. Dann liegt das Schallfeld als Satz von Impulsantworten vor. Im Zeitbereich wird das Schallfeld in seine frühen und späten Anteile zerlegt. Anschließend werden beide Teile in ihre sphärischen oder zylindrischen harmonischen Komponenten zerlegt. Da die relativen Richtungsinformationen im frühen Schallfeld vorhanden sind, wird hier eine höhere Ordnung der sphärischen Harmonischen gerechnet als im späten Schallfeld, das für eine niedrige Ordnung ausreichend ist. Der frühe Teil ist verhältnismäßig kurz, beispielsweise 100 ms und wird genau, also mit vielen harmonischen Komponenten dargestellt, während der späte Teil, beispielsweise 100 ms bis 2 s oder 10 s lang ist. Dieser späte Teil wird jedoch mit weniger oder nur einer einzigen harmonischen Komponente dargestellt.

[0059] Eine weitere Datenreduktion ergibt sich durch die Aufspaltung des frühen Schallfelds in einzelne Bänder vor der Darstellung als sphärische Harmonische. Dazu wird nach der Trennung im Zeitbereich in frühes und spätes Schallfeld das frühe Schallfeld mittels einer Filterbank in seine spektralen Anteile zerlegt. Durch Unterabtastung der einzelnen Frequenzbänder wird eine Datenreduktion erreicht, die die Berechnung der harmonischen Komponenten deutlich beschleunigt. Zusätzlich wird für jedes Frequenzband eine in Abhängigkeit der menschlichen Richtungswahrnehmung perzeptiv ausreichend frühe Ordnung verwendet. So sind für niedrige Frequenzbänder, in denen die menschliche Richtungs-

wahrnehmung niedrig ist, niedrige Ordnungen oder sogar für das niedrigste Frequenzband die Ordnung null ausreichen, während bei hohen Bändern höhere Ordnungen bis zur maximal sinnvollen Ordnung im Hinblick auf die Genauigkeit des gemessenen Schallfeldes benötigt werden. Auf der Decodierer- bzw. Dekomprimierer-Seite wird das komplette Spektrum rekonstruiert. Anschließend werden frühes oder spätes Schallfeld wieder kombiniert. Die Daten stehen nun zur Auralisation bereit.

[0060] 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.

[0061] 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.

[0062] 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.

[0063] 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.

[0064] Der Programmcode kann beispielsweise auch auf einem maschinenlesbaren Träger gespeichert sein.

[0065] Andere Ausführungsbeispiele umfassen das

Computerprogramm zum Durchführen eines der hierin beschriebenen Verfahren, wobei das Computerprogramm auf einem maschinenlesbaren Träger gespeichert ist.

5 **[0066]** 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.

10 **[0067]** 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.

15 **[0068]** 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.

20 **[0069]** 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.

30 **[0070]** Ein weiteres Ausführungsbeispiel umfasst einen Computer, auf dem das Computerprogramm zum Durchführen eines der hierin beschriebenen Verfahren installiert ist.

35 **[0071]** 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.

40 **[0072]** 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
45
50
55 einigen Ausführungsbeispielen seitens einer beliebigen Hardwarevorrichtung durchgeführt. Diese kann eine universell einsetzbare Hardware wie ein Computerprozessor (CPU) sein oder für das Verfahren spezifische Hard-

ware, wie beispielsweise ein ASIC.

[0073] 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 zum Komprimieren von Schallfelddaten (10) eines Gebiets, mit folgenden Merkmalen:

einem Aufteiler (100) zum Aufteilen der Schallfelddaten in einen ersten Anteil (101) und in einen zweiten Anteil (102); und
 einem Umsetzer (140, 180) zum Umsetzen des ersten Anteils (101) und des zweiten Anteils (102) in harmonische Komponenten (141, 182) einer Schallfeldbeschreibung, wobei der Umsetzer (140, 180) ausgebildet ist, um den zweiten Anteil (102) in eine oder mehrere harmonische Komponenten (141) mit einer zweiten Ordnung umzusetzen, und um den ersten Anteil (101) in harmonische Komponenten mit einer ersten Ordnung umzusetzen, wobei die erste Ordnung höher als die zweite Ordnung ist, um die komprimierten Schallfelddaten zu erhalten, wobei der Aufteiler (100) ausgebildet ist, um eine spektrale Aufteilung durchzuführen und eine Filterbank (100b) zum Filtern zumindest eines Teils der Schallfelddaten (10) aufweist, um Schallfelddaten in verschiedenen Filterbankkanälen (140a, 140b, 140c) zu erhalten, und wobei der Umsetzer ausgebildet ist, um für ein Subbandsignal aus einem ersten Filterbankkanal (140c), das den ersten Anteil (101) darstellt, der verschiedenen Filterbankkanäle (140a, 140b, 140c) die harmonischen Komponenten mit der ersten Ordnung zu berechnen, und um für ein Subbandsignal aus einem zweiten Filterbankkanal (140a), das den zweiten Anteil (102) darstellt, der verschiedenen Filterbankkanäle (140a, 140b, 140c) die harmonischen Komponenten mit der zweiten Ordnung zu berechnen, wobei eine Mittenfrequenz (f_n) des ersten Filterbankkanals (140a) höher als eine Mittenfrequenz (f_1) des zweiten Filterbankkanals (140c) ist.

2. Vorrichtung nach Anspruch 1, bei der der Umsetzer (140, 180) ausgebildet ist, um

für den ersten Anteil, der für eine Richtungswahrnehmung des menschlichen Gehörs wichtiger als der zweite Anteil ist, die harmonischen Komponenten mit der ersten Ordnung zu berechnen, die höher als die zweite Ordnung ist.

3. Vorrichtung nach Anspruch 1 oder 2, bei der der Aufteiler (100) ausgebildet ist, um die Schallfelddaten (10) in den ersten Anteil, der erste Reflexionen in dem Gebiet umfasst, und in den zweiten Anteil, der zweite Reflexionen in dem Gebiet umfasst, aufzuspalten, wobei die zweiten Reflexionen zeitlich später als die ersten Reflexionen auftreten.

4. Vorrichtung nach einem der vorhergehenden Ansprüche, bei der der Aufteiler (100) ausgebildet ist, um die Schallfelddaten in den ersten Anteil, der erste Reflexionen in dem Gebiet umfasst, und in den zweiten Anteil, der zweite Reflexionen in dem Gebiet umfasst, aufzuteilen, wobei die zweiten Reflexionen zeitlich später als die ersten Reflexionen auftreten, und bei dem der Aufteiler (100) ferner ausgebildet ist, um den ersten Anteil in spektrale Anteile (101, 102) zu zerlegen, und um die spektralen Anteile jeweils in eine oder mehrere harmonische Komponenten mit unterschiedlicher Ordnung umzusetzen, wobei eine Ordnung für einen spektralen Anteil mit einem höheren Frequenzband höher ist als eine Ordnung für einen spektralen Anteil in einem niedrigeren Frequenzband.

5. Vorrichtung nach einem der vorhergehenden Ansprüche, die ferner eine AusgabeSchnittstelle (190) aufweist, um die eine oder die mehreren harmonischen Komponenten (182) mit der zweiten Ordnung und die harmonischen Komponenten mit der ersten Ordnung (141) zusammen mit Seiteninformationen (300), die einen Hinweis auf die erste Ordnung oder die zweite Ordnung aufweisen, zur Übertragung und Speicherung zu liefern.

6. Vorrichtung nach einem der vorhergehenden Ansprüche, bei der die Schallfelddaten ein dreidimensionales Gebiet beschreiben und der Umsetzer ausgebildet ist, um als die harmonischen Komponenten zylindrische harmonische Komponenten zu berechnen, oder bei der die Schallfelddaten (10) ein dreidimensionales Gebiet beschreiben und der Umsetzer (140, 180) ausgebildet ist, um als die harmonischen Komponenten sphärische harmonische Komponenten zu berechnen.

7. Vorrichtung nach einem der vorhergehenden Ansprüche, bei der die Schallfelddaten als eine erste Anzahl von

diskreten Signalen vorliegen,
bei der der Umsetzer (140, 180) für den ersten Anteil
(101) und den zweiten Anteil (102) eine zweite ge-
samte Anzahl von harmonischen Komponenten lie-
fert, und
bei der die zweite gesamte Anzahl von harmoni-
schen Komponenten kleiner als die erste Anzahl von
diskreten Signalen ist.

8. Vorrichtung nach einem der vorhergehenden An-
sprüche,
bei der der Aufteiler (100) ausgebildet ist, um als
Schallfelddaten (10) eine Vielzahl verschiedener Im-
pulsantworten zu verwenden, die unterschiedlichen
Positionen in dem Gebiet zugeordnet sind.

9. Vorrichtung nach Anspruch 8,
bei der die Impulsantworten Head-Related Transfer
Functions (HRTF-) Funktionen oder Binaural Room
Impulse Responses (BRIR-) Funktionen oder Im-
pulsantworten von jeweils einem diskreten Punkt in
dem Gebiet zu einer vorbestimmten Position in dem
Gebiet sind.

10. Vorrichtung nach einem der vorhergehenden An-
sprüche, die ferner folgendes Merkmal aufweist:

einen Decodierer (2) zum Dekomprimieren der
komprimierten Schallfelddaten unter Verwen-
dung einer Kombination des ersten und des
zweiten Anteils und unter Verwendung einer
Umsetzung von einer Harmonische-Kompo-
nenten-Darstellung in eine Zeitbereichsdarstel-
lung, um eine dekomprimierte Darstellung zu er-
reichen; und
eine Steuerung (4) zum Steuern des Aufteilers
(100) oder des Umsetzers (140, 180) bezüglich
der ersten oder zweiten Ordnung, wobei die
Steuerung (4) ausgebildet ist, um unter Verwen-
dung eines psychoakustischen Moduls die de-
komprimierten Schallfelddaten mit den Schall-
felddaten (10) zu vergleichen und um den Auf-
teiler (100) oder den Umsetzer (140, 180) unter
Verwendung des Vergleichs anzusteuern.

11. Vorrichtung nach Anspruch 10,
bei der der Decodierer ausgebildet ist, um die har-
monischen Komponenten der zweiten Ordnung und
die harmonischen Komponenten der ersten Ord-
nung umzusetzen (241, 242), und dann eine Kom-
bination der umgesetzten harmonischen Kompo-
nenten durchzuführen, oder
wobei der Decodierer (2) ausgebildet ist, um die har-
monischen Komponenten mit der zweiten Ordnung
und die harmonischen Komponenten mit der ersten
Ordnung zu kombinieren (245), und um ein Ergebnis
der Kombination in dem Kombinierer (245) von ei-
nem Harmonische-Komponenten-Bereich in den

Zeitbereich umzusetzen (244).

12. Vorrichtung nach Anspruch 10,
bei der der Decodierer ausgebildet ist, um harmoni-
sche Komponenten verschiedener spektraler Antei-
le mit unterschiedlichen Ordnungen umzusetzen
(140a, 140b),
um unterschiedliche Verarbeitungszeiten für unter-
schiedliche spektrale Anteile auszugleichen (304,
306), und
um in einen Zeitbereich umgesetzte spektrale Antei-
le des ersten Anteils mit in den Zeitbereich umge-
setzten spektralen Komponenten des zweiten An-
teils durch zueinander serielle Anordnung zu kom-
binieren.

13. Vorrichtung zum Dekomprimieren von komprimier-
ten Schallfelddaten, die erste harmonische Kompo-
nenten (141) einer Schallfeldbeschreibung bis zu ei-
ner ersten Ordnung und eine oder mehrere zweite
harmonische Komponenten (182) einer Schallfeld-
beschreibung bis zu einer zweiten Ordnung aufwei-
sen, wobei die erste Ordnung größer als die zweite
Ordnung ist, mit folgenden Merkmalen:

einer Eingangsschnittstelle (200) zum Erhalten
der komprimierten Schallfelddaten; und
einem Prozessor (240) zum Verarbeiten der ers-
ten harmonischen Komponenten (201) und der
zweiten harmonischen Komponenten (202) un-
ter Verwendung einer Kombination des ersten
und des zweiten Anteils und unter Verwendung
einer Umsetzung von einer Harmonische-Kom-
ponenten-Darstellung in eine Zeitbereichsdar-
stellung, um eine dekomprimierte Darstellung
zu erreichen, wobei der erste Anteil durch die
ersten harmonischen Komponenten und der
zweite Anteil durch die zweiten harmonischen
Komponenten dargestellt sind,
wobei die ersten harmonischen Komponenten
(HK_n) der ersten Ordnung einen ersten Spek-
tralbereich darstellen, und die eine oder die
mehreren harmonischen Komponenten (HK₁)
der zweiten Ordnung einen unterschiedlichen
Spektralbereich darstellen,
wobei der Prozessor (240) ausgebildet ist, um
die ersten harmonischen Komponenten (HK_n)
der ersten Ordnung in den Spektralbereich um-
zusetzen (241a), und um die eine oder die meh-
reren zweiten harmonischen Komponenten
(HK₁) der zweiten Ordnung in den Spektralbe-
reich umzusetzen (241c), und um die umgesetz-
ten harmonischen Komponenten mittels einer
Synthese-Filterbank (245) zu kombinieren, um
eine Darstellung von Schallfelddaten in dem
Zeitbereich zu erhalten.

14. Vorrichtung nach Anspruch 13, bei der der Prozes-

sor (240) folgende Merkmale aufweist:

einen Kombinierer (245) zum Kombinieren der ersten harmonischen Komponenten und der zweiten harmonischen Komponenten, um kombinierte harmonische Komponenten zu erhalten; und
einen Umsetzer (244) zum Umsetzen der kombinierten harmonischen Komponenten in den Zeitbereich.

15. Vorrichtung nach Anspruch 13, bei der der Prozessor folgende Merkmale aufweist:

einen Umsetzer (241, 242) zum Umsetzen der ersten harmonischen Komponenten und der zweiten harmonischen Komponenten in den Zeitbereich; und
einen Kombinierer (243, 245) zum Zusammen-
setzen der in den Zeitbereich umgesetzten harmonischen Komponenten, um die dekomprimierten Schallfelddaten zu erhalten.

16. Vorrichtung nach einem der Ansprüche 13 bis 15, bei der der Prozessor (240) ausgebildet ist, um Informationen über eine Wiedergabeordnung (610, 612, 614) zu erhalten, und
bei der der Prozessor (240) ausgebildet ist, um die dekomprimierten Schallfelddaten (602, 604, 606) zu berechnen und aufgrund der Informationen über die Wiedergabeordnung einen Teil der Schallfelddaten der dekomprimierten Schallfelddaten zu Zwecken einer Wiedergabe auszuwählen (608), oder bei der der Prozessor ausgebildet ist, um nur einen Teil der für die Wiedergabeordnung nötigen dekomprimierten Schallfelddaten zu berechnen.

17. Vorrichtung nach einem der Ansprüche 13 bis 16, bei der die ersten harmonischen Komponenten mit der ersten Ordnung frühe Reflexionen des Gebiets darstellen und die zweiten harmonischen Komponenten mit der zweiten Ordnung späte Reflexionen des Gebiets darstellen, und
bei der der Prozessor (240) ausgebildet ist, um die ersten harmonischen Komponenten und die zweiten harmonischen Komponenten zu addieren und um ein Ergebnis der Addition in den Zeitbereich umzusetzen, um die dekomprimierten Schallfelddaten zu erhalten.

18. Vorrichtung nach einem der Ansprüche 13 bis 17, bei der der Prozessor ausgebildet ist, um zum Umsetzen eine inverse Raumtransformation (604) und eine inverse Fouriertransformation (606) durchzuführen.

19. Verfahren zum Komprimieren von Schallfelddaten (10) eines Gebiets, mit folgenden Schritten:

Aufteilen (100) der Schallfelddaten in einen ersten Anteil (101) und in einen zweiten Anteil (102); und

Umsetzen (140, 180) des ersten Anteils (101) und des zweiten Anteils (102) in harmonische Komponenten (141, 182) einer Schallfeldbeschreibung, wobei der zweite Anteil (102) in eine oder mehrere harmonische Komponenten (141) mit einer zweiten Ordnung umgesetzt wird, und wobei der erste Anteil (101) in harmonische Komponenten mit einer ersten Ordnung umgesetzt wird, wobei die erste Ordnung höher als die zweite Ordnung ist, um die komprimierten Schallfelddaten zu erhalten,

wobei das Aufteilen (100) eine spektrale Aufteilung durch ein Filtern mit einer Filterbank (100b) zum Filtern zumindest eines Teils der Schallfelddaten (10) aufweist, um Schallfelddaten in verschiedenen Filterbankkanälen (140a, 140b, 140c) zu erhalten, und

wobei das Umsetzen ein Berechnen der harmonischen Komponenten mit der ersten Ordnung für ein Subbandsignal aus einem ersten Filterbankkanal (140c), das den ersten Anteil (101) darstellt, der verschiedenen Filterbankkanäle (140a, 140b, 140c), und ein Berechnen der harmonischen Komponenten mit der zweiten Ordnung für ein Subbandsignal aus einem zweiten Filterbankkanal (140a), das den zweiten Anteil (102) darstellt, der verschiedenen Filterbankkanäle (140a, 140b, 140c) aufweist, wobei eine Mittenfrequenz (f_n) des ersten Filterbankkanals (140a) höher als eine Mittenfrequenz (f_1) des zweiten Filterbankkanals (140c) ist.

20. Verfahren zum Dekomprimieren von komprimierten Schallfelddaten, die erste harmonischen Komponenten (141) einer Schallfeldbeschreibung bis zu einer ersten Ordnung und eine oder mehrere zweite harmonische Komponenten (182) einer Schallfeldbeschreibung bis zu einer zweiten Ordnung aufweisen, wobei die erste Ordnung größer als die zweite Ordnung ist, mit folgenden Schritten:

Erhalten (200) der komprimierten Schallfelddaten; und

Verarbeiten (240) der ersten harmonischen Komponenten (201) und der zweiten harmonischen Komponenten (202) unter Verwendung einer Kombination des ersten und des zweiten Anteils und unter Verwendung einer Umsetzung von einer Harmonische-Komponenten-Darstellung in eine Zeitbereichsdarstellung, um eine dekomprimierte Darstellung zu erreichen, wobei der erste Anteil durch die ersten harmonischen Komponenten und der zweite Anteil durch die zweiten harmonischen Komponenten dargestellt sind,

wobei die ersten harmonischen Komponenten (HK_n) der ersten Ordnung einen ersten Spektralbereich darstellen, und die eine oder die mehreren harmonischen Komponenten (HK_1) der zweiten Ordnung einen unterschiedlichen Spektralbereich darstellen, wobei das Verarbeiten (240) ein Umsetzen der ersten harmonischen Komponenten (HK_n) der ersten Ordnung in den Spektralbereich und ein Umsetzen der einen oder der mehreren zweiten harmonischen Komponenten (HK_1) der zweiten Ordnung in den Spektralbereich und ein Kombinieren der umgesetzten harmonischen Komponenten mittels einer Synthese-Filterbank (245) aufweist, um eine Darstellung von Schallfelddaten in dem Zeitbereich zu erhalten.

21. Computerprogramm zum Durchführen eines Verfahrens nach einem der Ansprüche 19 bis 20, wenn das Verfahren auf einem Computer abläuft.

Claims

1. Apparatus for compressing sound field data (10) of an area, comprising:

a divider (100) for dividing the sound field data into a first portion (101) and into a second portion (102); and

a converter (140, 180) for converting the first portion (101) and the second portion (102) into harmonic components (141, 182) of a sound field description, wherein the converter (140, 180) is configured to convert the second portion (102) into one or several harmonic components (141) of a second order, and to convert the first portion (101) into harmonic components of a first order, wherein the first order is higher than the second order, to obtain the compressed sound field data,

wherein the divider (100) is configured to perform spectral division and comprises a filterbank (100b) for filtering at least part of the sound field data (10) for obtaining sound field data in different filterbank channels (140a, 140b, 140c), and wherein the converter is configured to compute, for a subband signal from a first filterbank channel (140c), which represents the first portion (101), of the different filterbank channels (140a, 140b, 140c), the harmonic components of the first order, and to compute, for a subband signal from a second filterbank channel (140a), which represents the second portion (102), of the different filterbank channels (140a, 140b, 140c), the harmonic components of the second order, wherein a center frequency (f_n) of the first filterbank channel (140a) is higher than a center fre-

quency (f_1) of the second filterbank channel (140c).

2. Apparatus according to claim 1, wherein the converter (140, 180) is configured to compute the harmonic components of the first order, which is higher than the second order, for the first portion, which is more important for directional perception of the human hearing than the second portion.
3. Apparatus according to claims 1 or 2, wherein the divider (100) is configured to divide the sound field data (10) into the first portion including first reflections in the area and into the second portion including second reflections in the area, wherein the second reflections occur later in time than the first reflections.
4. Apparatus according to one of the previous claims, wherein the divider (100) is configured to divide the sound field data into the first portion including first reflections in the area and into the second portion including second reflections in the area, wherein the second reflections occur later in time than the first reflections, and wherein the divider (100) is further configured to decompose the first portion into spectral portions (101, 102) and to convert the spectral portions each into one or several harmonic components of different orders, wherein an order for a spectral portion with a higher frequency band is higher than an order for a spectral portion in a lower frequency band.
5. Apparatus according to one of the previous claims, further comprising an output interface (190) for providing the one or several harmonic components (182) of the second order and the harmonic components of the first order (141) together with side information (300) comprising an indication on the first order or the second order for transmission and storage.
6. Apparatus according to one of the previous claims, wherein the sound field data describe a three-dimensional area and the converter is configured to compute cylindrical harmonic components as the harmonic components, or wherein the sound field data (10) describe a three-dimensional area and the converter (140, 180) is configured to compute spherical harmonic components as the harmonic components.
7. Apparatus according to one of the previous claims, wherein the sound field data exist as a first number of discrete signals, wherein the converter (140, 180) for the first portion (101) and the second portion (102) provides a second total number of harmonic components, and

wherein the second total number of harmonic components is smaller than the first number of discrete signals.

8. Apparatus according to one of the previous claims, wherein the divider (100) is configured to use, as sound field data (10), a plurality of different impulse responses that are allocated to different positions in the area.
9. Apparatus according to claim 8, wherein the impulse responses are head-related transfer functions (HRTF) or binaural room impulse responses (BRIR) functions or impulse responses of a respective discrete point in the area to a predetermined position in the area.
10. Apparatus according to one of the previous claims, further comprising:
 - a decoder (2) for decompressing the compressed sound field data by using a combination of the first and second portions and by using a conversion from a harmonic component representation into a time domain representation for obtaining a decompressed representation; and
 - a control (4) for controlling the divider (100) or the converter (140, 180) with respect to the first or second order, wherein the control (4) is configured to compare, by using a psychoacoustic module, the decompressed sound field data with the sound field data (10) and to control the divider (100) or the converter (140, 180) by using the comparison.
11. Apparatus according to claim 10, wherein the decoder is configured to convert the harmonic components of the second order and the harmonic components of the first order (241, 242) and to then perform a combination of the converted harmonic components, or wherein the decoder (2) is configured to combine the harmonic components of the second order and the harmonic components of the first order (245) and to convert a result of the combination in the combiner (245) from a harmonic component domain into the time domain (244).
12. Apparatus according to claim 10, wherein the decoder is configured to convert harmonic components of different spectral portions with different orders (140a, 140b), to compensate different processing times for different spectral portions (304, 306), and to combine spectral portions of the first portion converted into a time domain with the spectral components of the second portion converted into the time domain by serially arranging the same.

13. Apparatus for decompressing compressed sound field data comprising first harmonic components (141) of a sound field description up to a first order and one or several second harmonic components (182) of a sound field description up to a second order, wherein the first order is higher than the second order, comprising:

an input interface (200) for obtaining the compressed sound field data; and
a processor (240) for processing the first harmonic components (201) and the second harmonic components (202) by using a combination of the first and the second portion and by using a conversion of a harmonic component representation into a time domain representation to obtain a decompressed illustration, wherein the first portion is represented by the first harmonic components and the second portion by the second harmonic components, wherein the first harmonic components (HK_n) of the first order represent a first spectral domain, and the one or the several harmonic components (HK_1) of the second order represent a different spectral domain, wherein the processor (240) is configured to convert the harmonic components (HK_n) of the first order into the spectral domain (241a) and to convert the one or the several second harmonic components (HK_1) of the second order into the spectral domain (241c), and to combine the converted harmonic components by means of a synthesis filterbank (245) to obtain a representation of sound field data in the time domain.

14. Apparatus according to claim 13, wherein the processor (240) comprises:

a combiner (245) for combining the first harmonic components and the second harmonic components to obtain combined harmonic components; and
a converter (244) for converting the combined harmonic components into the time domain.

15. Apparatus according to claim 13, wherein the processor comprises:

a converter (241, 242) for converting the first harmonic components and the second harmonic components into the time domain; and
a combiner (243, 245) for combining the harmonic components converted into the time domain for obtaining the decompressed sound field data.

16. Apparatus according to one of claims 13 to 15, wherein the processor (240) is configured to obtain

information on a reproduction arrangement (610, 612, 614), and

wherein the processor (240) is configured to compute the decompressed sound field data (602, 604, 606) and to select, based on the information on the reproduction arrangement, part of the sound field data of the decompressed sound field data for reproduction purposes (608), or
 wherein the processor is configured to compute only a part of the decompressed sound field data necessitated for the reproduction arrangement.

17. Apparatus according to one of claims 13 to 16, wherein the first harmonic components of the first order represent early reflections of the area and the second harmonic components of the second order represent late reflections of the area, and wherein the processor (240) is configured to add the first harmonic components and the second harmonic components and to convert a result of the addition into the time domain for obtaining the decompressed sound field data.

18. Apparatus according to one of claims 13 to 17, wherein the processor is configured to perform, for the conversion, an inverse room transformation (604) and an inverse Fourier transformation (606).

19. Method for compressing sound field data (10) of an area, comprising the steps of:

dividing (100) the sound field data into a first portion (101) and into a second portion (102), and

converting (140, 180) the first portion (101) and the second portion (102) into harmonic components (141, 182) of a sound field description, wherein the second portion (102) is converted into one or several harmonic components (141) of a second order, and wherein the first portion (101) is converted into harmonic components of a first order, wherein the first order is higher than the second order, to obtain the compressed sound field data,

wherein dividing (100) comprises spectral division by filtering with a filterbank (100b) for filtering at least part of the sound field data (10) for obtaining sound field data in different filterbank channels (140a, 140b, 140c), and

wherein converting represents a computation of the harmonic components of the first order for a subband signal from a first filterbank channel (140c), which represents the first portion (101), of the different filterbank channels (140a, 140b, 140c), and a computation of the harmonic components of the second order for a subband signal from a second filterbank channel (140a), which represents the second portion (102), of the dif-

ferent filterbank channels (140a, 140b, 140c), wherein a center frequency (f_n) of the first filterbank channel (140a) is higher than a center frequency (f_1) of the second filterbank channel (140c).

20. Method for decompressing compressed sound field data comprising first harmonic components (141) of a sound field description up to a first order and one or several second harmonic components (182) of a sound field description up to a second order, wherein the first order is higher than the second order, comprising steps of:

obtaining (200) the compressed sound field data; and

processing (240) the first harmonic components (201) and the second harmonic components (202) by using a combination of the first and second portions and by using a conversion from a harmonic component representation into a time domain representation to obtain a decompressed representation, wherein the first portion is represented by the first harmonic components and the second portion by the second harmonic components,

wherein the first harmonic components (HK_n) of the first order represent a first spectral domain, and the one or the several harmonic components (HK_1) of the second order represent a different spectral domain,

wherein processing (240) comprises converting the first harmonic components (HK_n) of the first order into the spectral domain and converting the one or the several second harmonic components (HK_1) of the second order into the spectral domain and combining the converted harmonic components by means of a synthesis filterbank (245) to obtain a representation of sound field data in the time domain.

21. Computer program for performing a method according to one of claims 19 to 20 when the method runs on a computer.

Revendications

1. Dispositif de compression de données de champ sonore (10) d'une région, aux caractéristiques suivantes:

un diviseur (100) destiné à diviser les données de champ sonore en une première part (101) et une deuxième part (102); et

un convertisseur (140, 180) destiné à convertir la première part (101) et la deuxième part (102) en composantes harmoniques (141, 182) d'une

ponses (BRIR) ou réponses impulsionnelles de chaque fois un point discret dans la région à une position prédéterminée dans la région.

10. Dispositif selon l'une des revendications précédentes, qui présente par ailleurs la caractéristique suivante:

un décodeur (2) destiné à décompresser les données de champ sonore comprimées à l'aide d'une combinaison de la première et de la deuxième part et à l'aide d'une conversion d'une représentation de composantes harmoniques en une représentation dans le domaine temporel, pour obtenir une représentation décompressée; et
un moyen de commande (4) destiné à commander le diviseur (100) ou le convertisseur (140, 180) en ce qui concerne le premier ou le deuxième ordre, où le moyen de commande (4) est conçu pour comparer, à l'aide d'un module psycho-acoustique, les données de champ sonore décompressées avec les données de champ sonore (10) et pour commander le diviseur (100) ou le convertisseur (140, 180) à l'aide de la comparaison.

11. Dispositif selon la revendication 10, dans lequel le décodeur est conçu pour convertir les composantes harmoniques du deuxième ordre et les composantes harmoniques du premier ordre (241, 242), et pour effectuer alors une combinaison des composantes harmoniques converties, ou dans lequel le décodeur (2) est conçu pour combiner les composantes harmoniques présentant le deuxième ordre et les composantes harmoniques présentant le premier ordre (245), et pour convertir un résultat de la combinaison dans le combineur (245) d'un domaine de composantes harmoniques au domaine temporel (244).

12. Dispositif selon la revendication 10, dans lequel le décodeur est conçu pour convertir (140a, 140b) des composantes harmoniques de différentes parts spectrales présentant des ordres différents, pour compenser (304, 306) différents temps de traitement pour différentes parts spectrales, et pour combiner les parts spectrales de la première partie converties à un domaine temporel avec les composantes spectrales de la deuxième part converties au domaine temporel par disposition en série l'une par rapport à l'autre.

13. Dispositif pour décompresser des données de champ sonore comprimées qui présentent des premières composantes harmoniques (141) d'une description de champ sonore jusqu'à un premier ordre et une

ou plusieurs deuxièmes composantes harmoniques (182) d'une description de champ sonore jusqu'à un deuxième ordre, où le premier ordre est supérieur au deuxième ordre, aux caractéristiques suivantes:

une interface d'entrée (200) destinée à obtenir les données de champ sonore comprimées; et un processeur (240) destiné à traiter les premières composantes harmoniques (201) et les deuxièmes composantes harmoniques (202) à l'aide d'une combinaison de la première et de la deuxième part et à l'aide d'une conversion d'une représentation de composantes harmoniques en une représentation dans le domaine temporel, pour obtenir une représentation décompressée, où la première part est représentée par les premières composantes harmoniques et la deuxième part est représentée par les deuxièmes composantes harmoniques, dans lequel les premières composantes harmoniques (HK_n) du premier ordre représentent un premier domaine spectral et les une ou plusieurs composantes harmoniques du deuxième ordre (HK_1) représentent un domaine spectral différent, dans lequel le processeur (240) est conçu pour convertir les premières composantes harmoniques (HK_n) du premier ordre au domaine spectral (241a), et pour convertir les une ou plusieurs deuxièmes composantes harmoniques (HK_1) du deuxième ordre au domaine spectral (241c), et pour combiner les composantes harmoniques converties au moyen d'un banc de filtres de synthèse (245) pour obtenir une représentation des données de champ sonore dans le domaine temporel.

14. Dispositif selon la revendication 13, dans lequel le processeur (240) présente les caractéristiques suivantes:

un combineur (245) destiné à combiner les premières composantes harmoniques et les deuxièmes composantes harmoniques pour obtenir des composantes harmoniques combinées; et
un convertisseur (244) destiné à convertir les composantes harmoniques combinées au domaine temporel.

15. Dispositif selon la revendication 13, dans lequel le processeur présente les caractéristiques suivantes:

un convertisseur (241, 242) destiné à convertir les premières composantes harmoniques et les deuxièmes composantes harmoniques au domaine temporel; et
un combineur (243, 245) destiné à assembler

les composantes harmoniques converties au domaine temporel pour obtenir les données de champ sonore décompressées.

16. Dispositif selon l'une des revendications 13 à 15, dans lequel le processeur (240) est conçu pour obtenir des informations sur un aménagement d'affichage (610, 612, 614), et dans lequel le processeur (240) est conçu pour calculer les données de champ sonore décompressées (602, 604, 606) et pour sélectionner, sur base des informations relatives à l'aménagement d'affichage, une partie des données de champ sonore des données de champ sonore décompressées à des fins d'une reproduction (608), ou dans lequel le processeur est conçu pour calculer uniquement une partie des données de champ sonore décompressées nécessaires pour l'aménagement d'affichage.
17. Dispositif selon l'une des revendications 13 à 16, dans lequel les premières composantes harmoniques présentant le premier ordre représentent les réflexions précoces de la région et les deuxièmes composantes harmoniques présentant le deuxième ordre représentent les réflexions tardives de la région, et dans lequel le processeur (240) est conçu pour additionner les premières composantes harmoniques et les deuxièmes composantes harmoniques et pour convertir un résultat de l'addition au domaine temporel pour obtenir les données de champ sonore décompressées.
18. Dispositif selon l'une des revendications 13 à 17, dans lequel le processeur est conçu pour effectuer, pour la conversion, une transformation spatiale inverse (604) et une transformation de Fourier inverse (606).
19. Procédé de compression de données de champ sonore (10) d'une région, aux étapes suivantes consistant à:
 - diviser (100) les données de champ sonore en une première part (101) et une deuxième part (102); et
 - convertir (140, 180) la première part (101) et la deuxième part (102) en composantes harmoniques (141, 182) d'une description de champ sonore, où la deuxième part (102) est convertie en une ou plusieurs composantes harmoniques (141) présentant un deuxième ordre et dans lequel la première partie (101) est convertie en composantes harmoniques présentant un premier ordre, où le premier ordre est supérieur au deuxième ordre, pour obtenir les données de champ sonore comprimées,

dans lequel la division (100) présente une division spectrale par un filtrage par un banc de filtres (100b) pour filtrer au moins une partie des données de champ sonore (10) pour obtenir des données de champ sonore dans différents canaux de banc de filtres (140a, 140b, 140c), et dans lequel la conversion présente un calcul des composantes harmoniques présentant le premier ordre pour un signal de sous-bande d'un premier canal de banc de filtres (140c) représentant la première partie (101) des différents canaux de banc de filtres (140a, 140b, 140c) et un calcul des composantes harmoniques présentant le deuxième ordre pour un signal de sous-bande d'un deuxième canal de banc de filtres (140a) représentant la deuxième partie (102) des différents canaux de banc de filtres (140a, 140b, 140c), où une fréquence centrale (f_n) du premier canal de banc de filtres (140a) est supérieure à une fréquence centrale (f_1) du deuxième canal de banc de filtres (140c).

20. Procédé de décompression de données de champ sonore comprimées qui présentent des premières composantes harmoniques (141) d'une description de champ sonore jusqu'à un premier ordre et une ou plusieurs deuxièmes composantes harmoniques (182) d'une description de champ sonore jusqu'à un deuxième ordre, où le premier ordre est supérieur au deuxième ordre, aux étapes suivantes consistant à:

obtenir (200) les données de champ sonore comprimées; et
traiter (240) les premières composantes harmoniques (201) et les deuxièmes composantes harmoniques (202) à l'aide d'une combinaison de la première et de la deuxième part et à l'aide d'une conversion d'une représentation de composantes harmoniques en d'une représentation dans le domaine temporel, pour obtenir une représentation décompressée, où la première part est représentée par les premières composantes harmoniques et la deuxième part est représentée par les deuxièmes composantes harmoniques,
dans lequel les premières composantes harmoniques (HK_n) du premier ordre représentent un premier domaine spectral et les une ou plusieurs composantes harmoniques (HK_1) du deuxième ordre représentent un domaine spectral différent,
dans lequel le traitement (240) présente une conversion des premières composantes harmoniques (HK_n) du premier ordre au domaine spectral et une conversion des une ou plusieurs deuxièmes composantes harmoniques (HK_1) du deuxième ordre au domaine spectral et une

combinaison des composantes harmoniques converties au moyen d'un banc de filtres de synthèse (245), pour obtenir une représentation des données de champ sonore dans le domaine temporel.

5

21. Programme d'ordinateur pour la mise en oeuvre d'un procédé selon l'une des revendications 19 à 20 lorsque le procédé est exécuté sur un ordinateur.

10

15

20

25

30

35

40

45

50

55

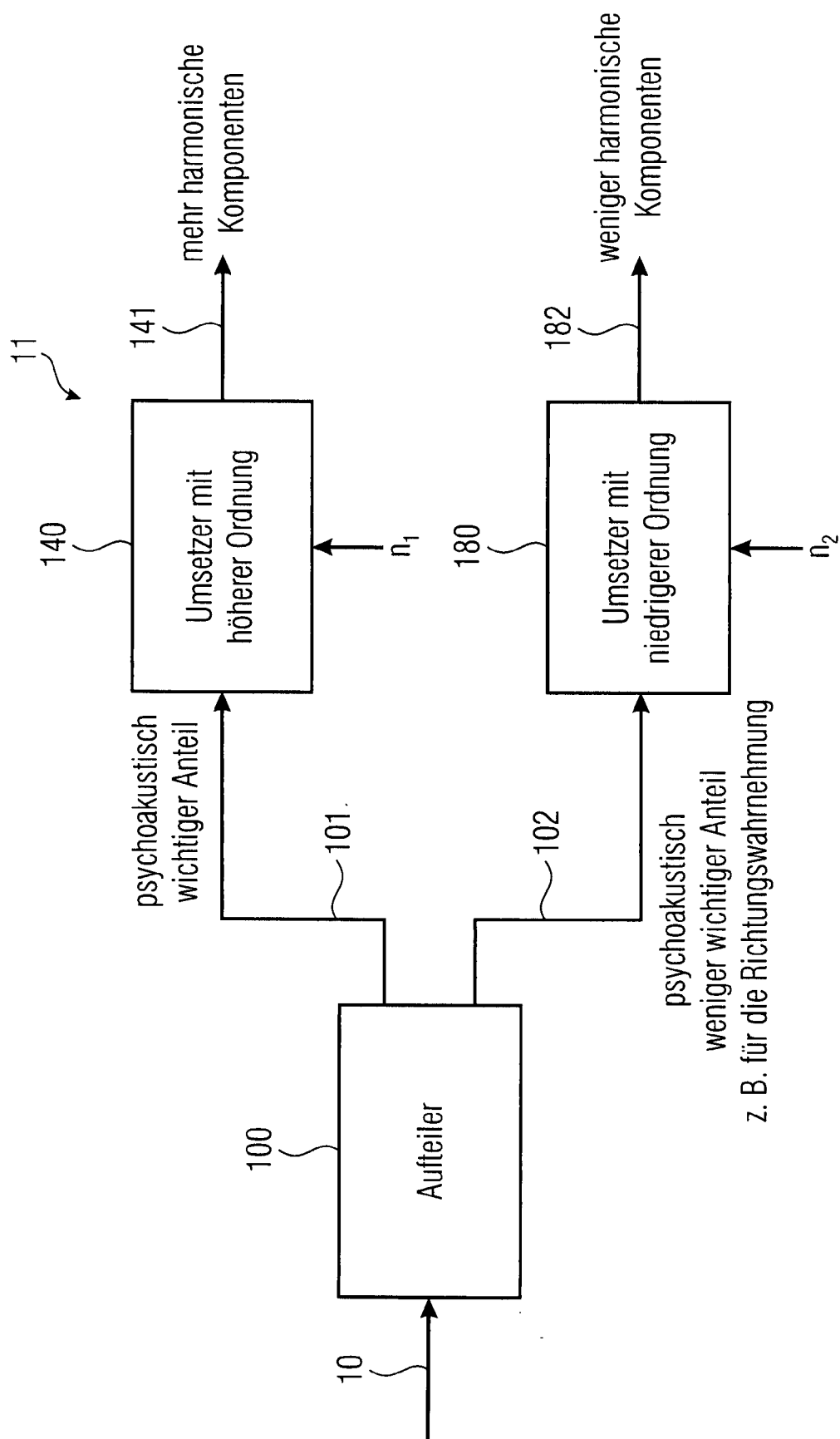


FIG 1A

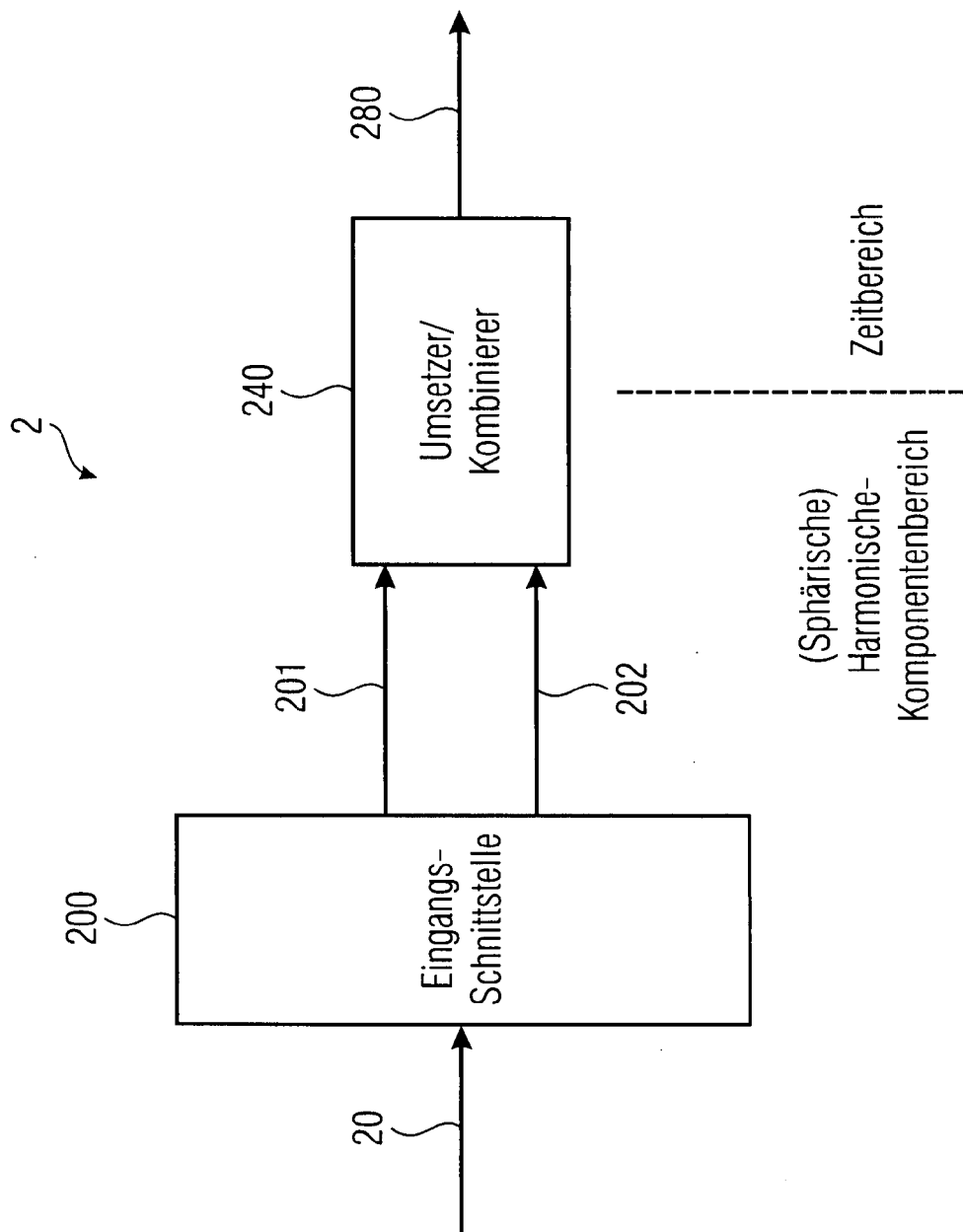


FIG 1B

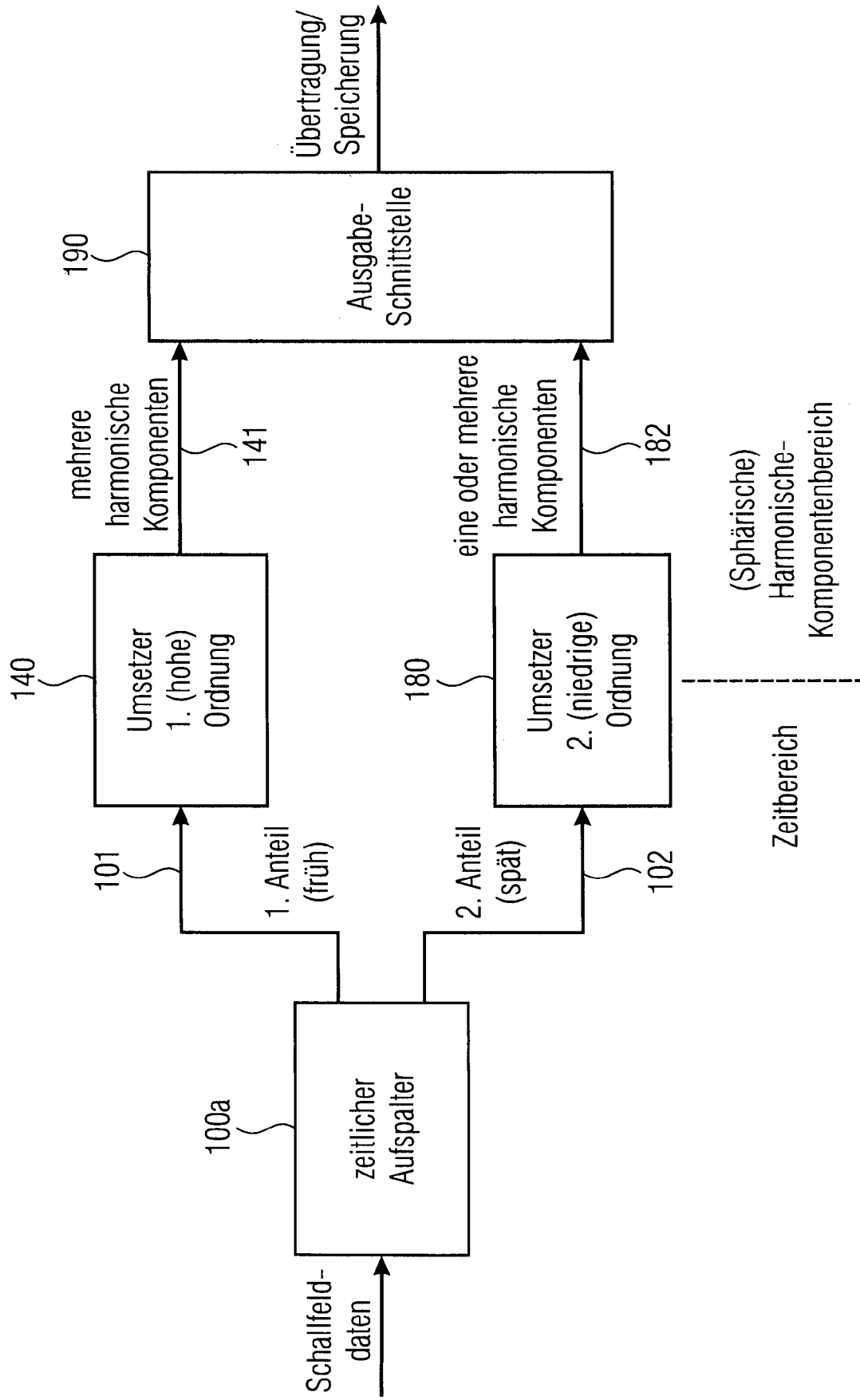


FIG 1C

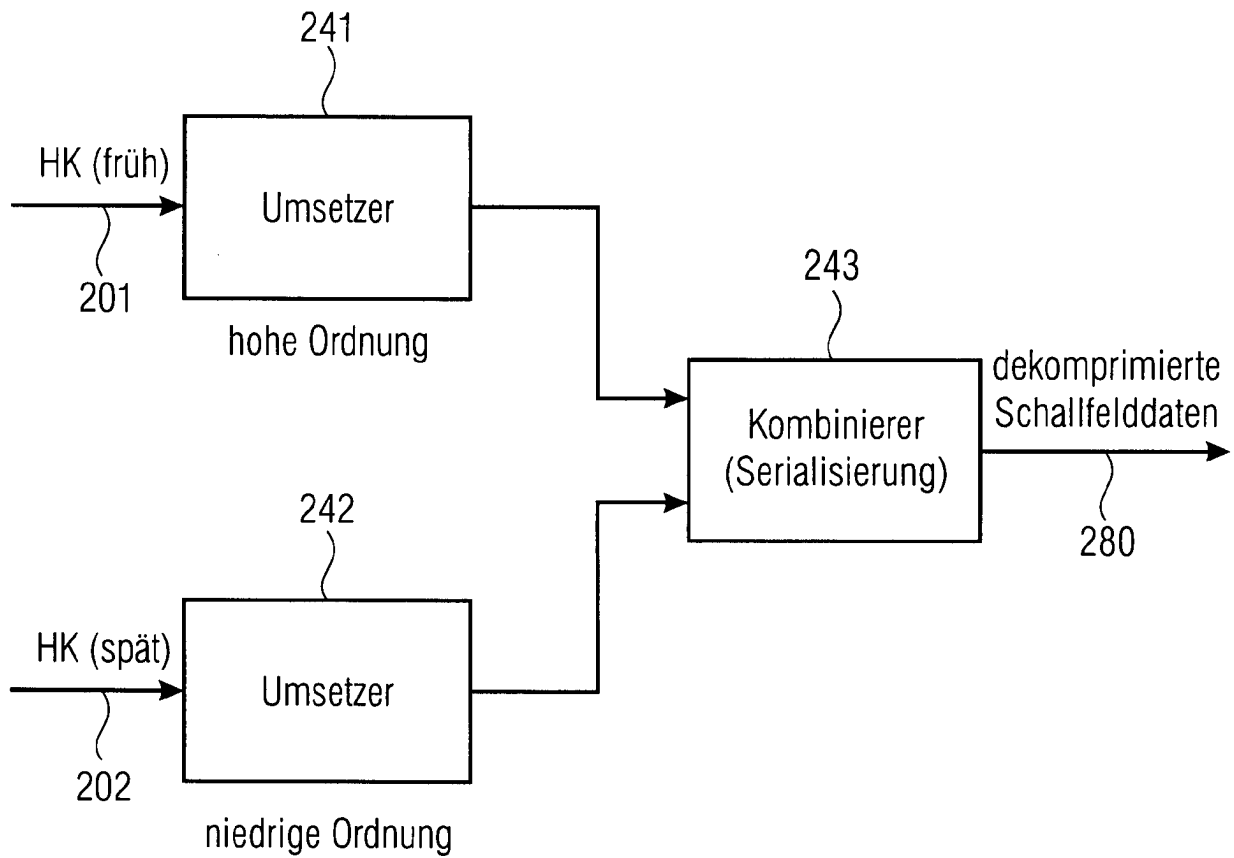


FIG 1D

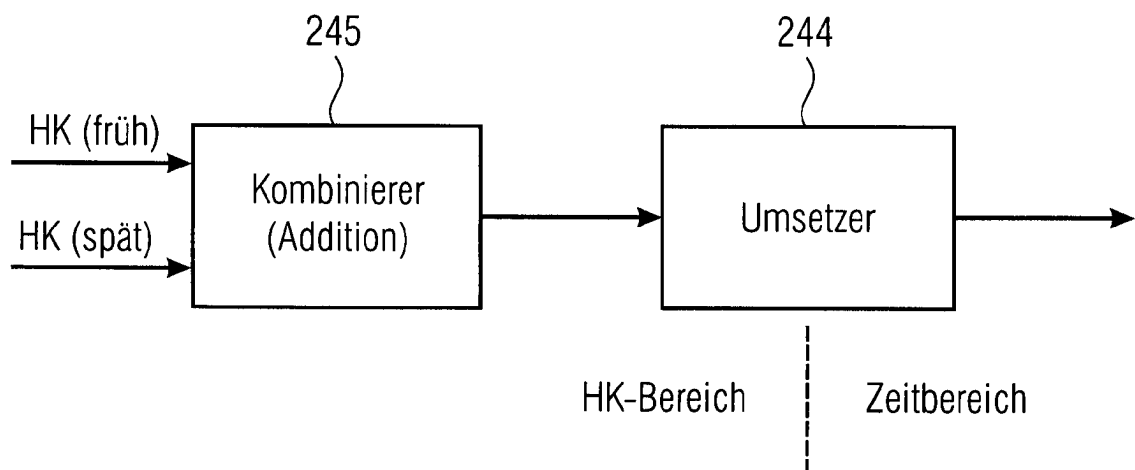


FIG 1E

Anteil	Ordnung	Anzahl der SHK
später Teil (volles Band)	0	1
früher Teil (erstes Band)	1	3
früher Teil (zweites Band)	5	11
⋮	⋮	⋮
früher Teil (letztes Band)	14	29

(für eine Raumbeschreibung mit 181 Impulsantworten)

FIG 1F

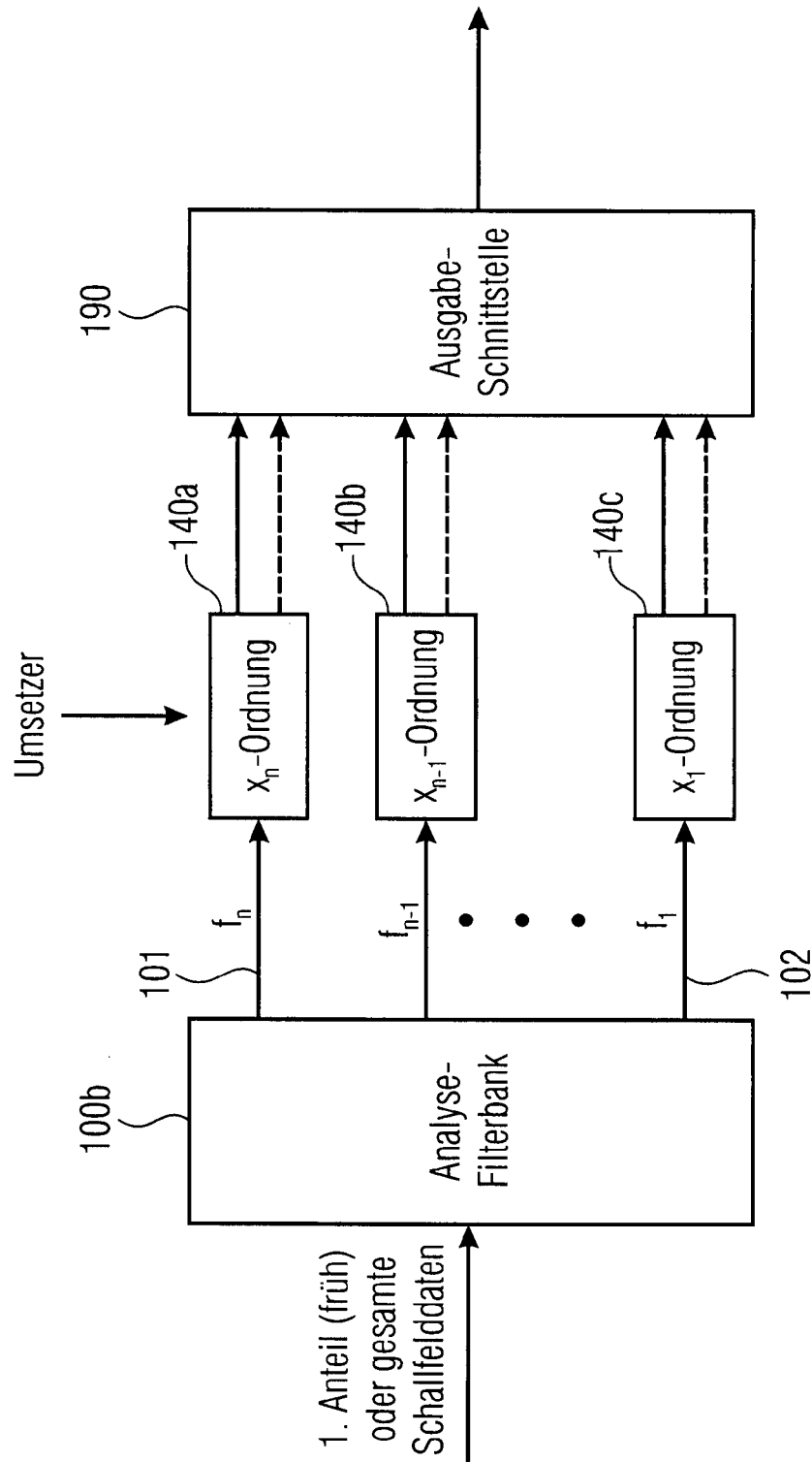


FIG 2A

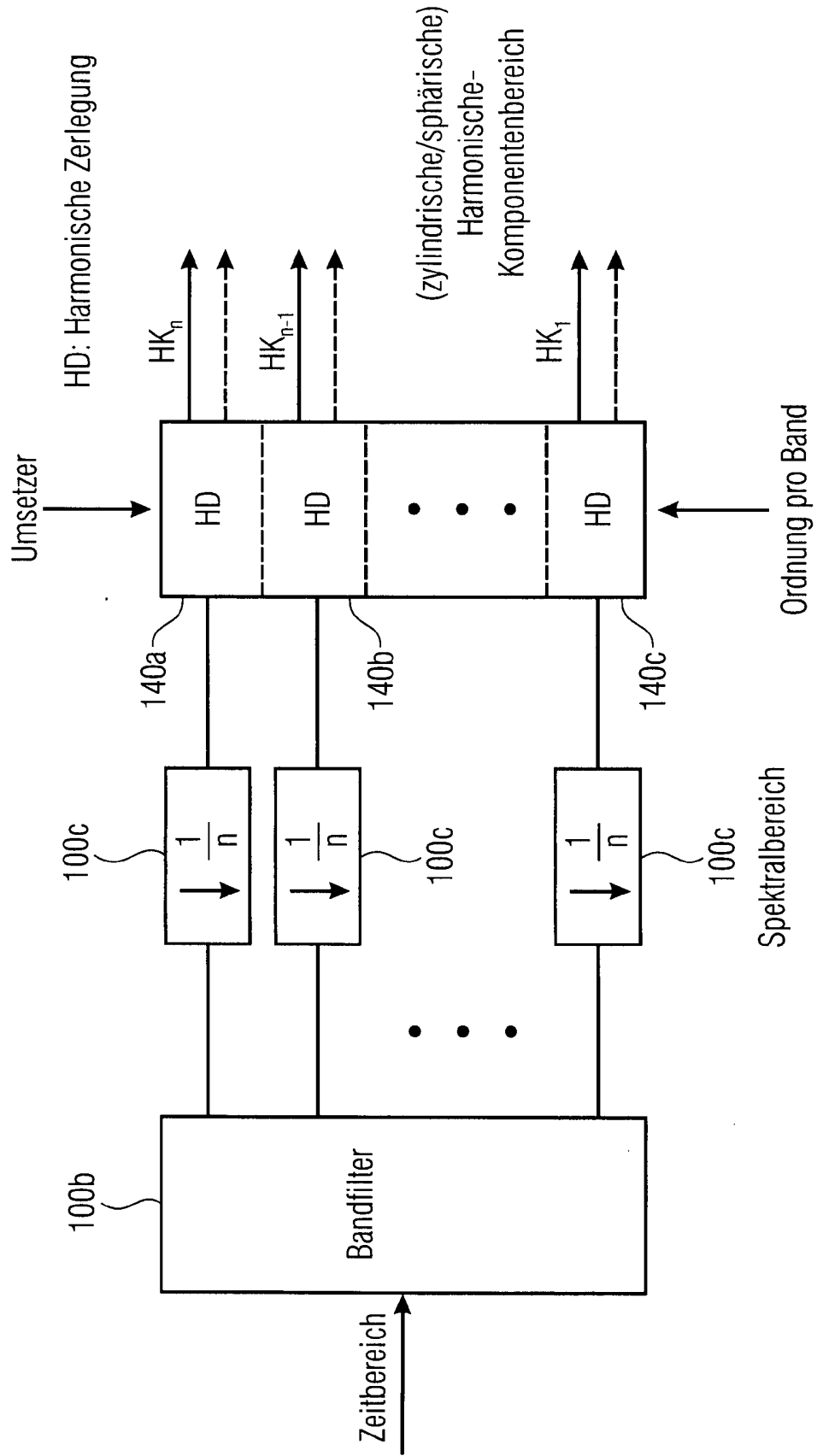


FIG 2B

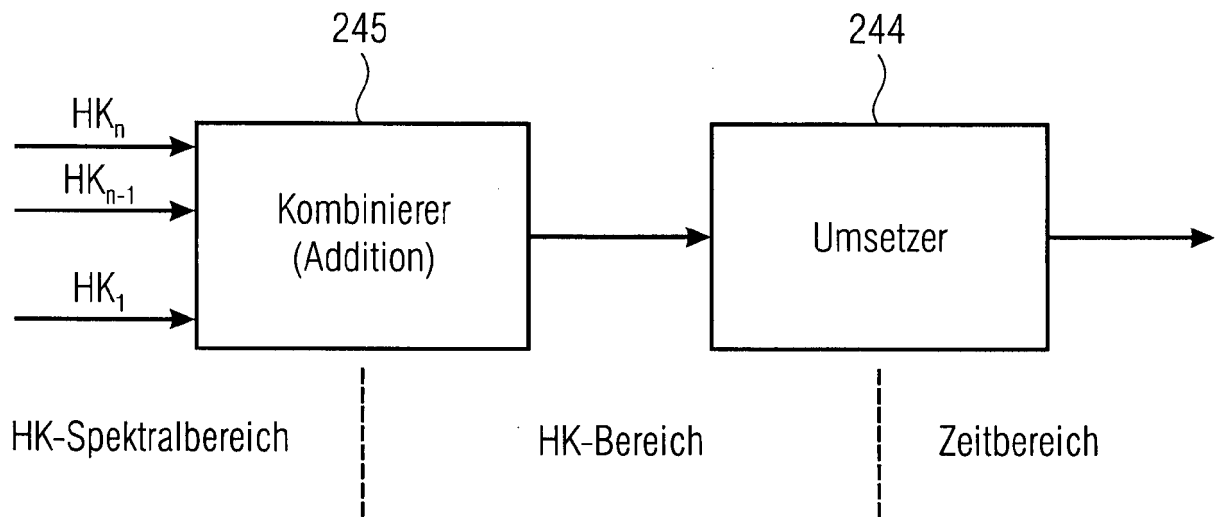


FIG 2C

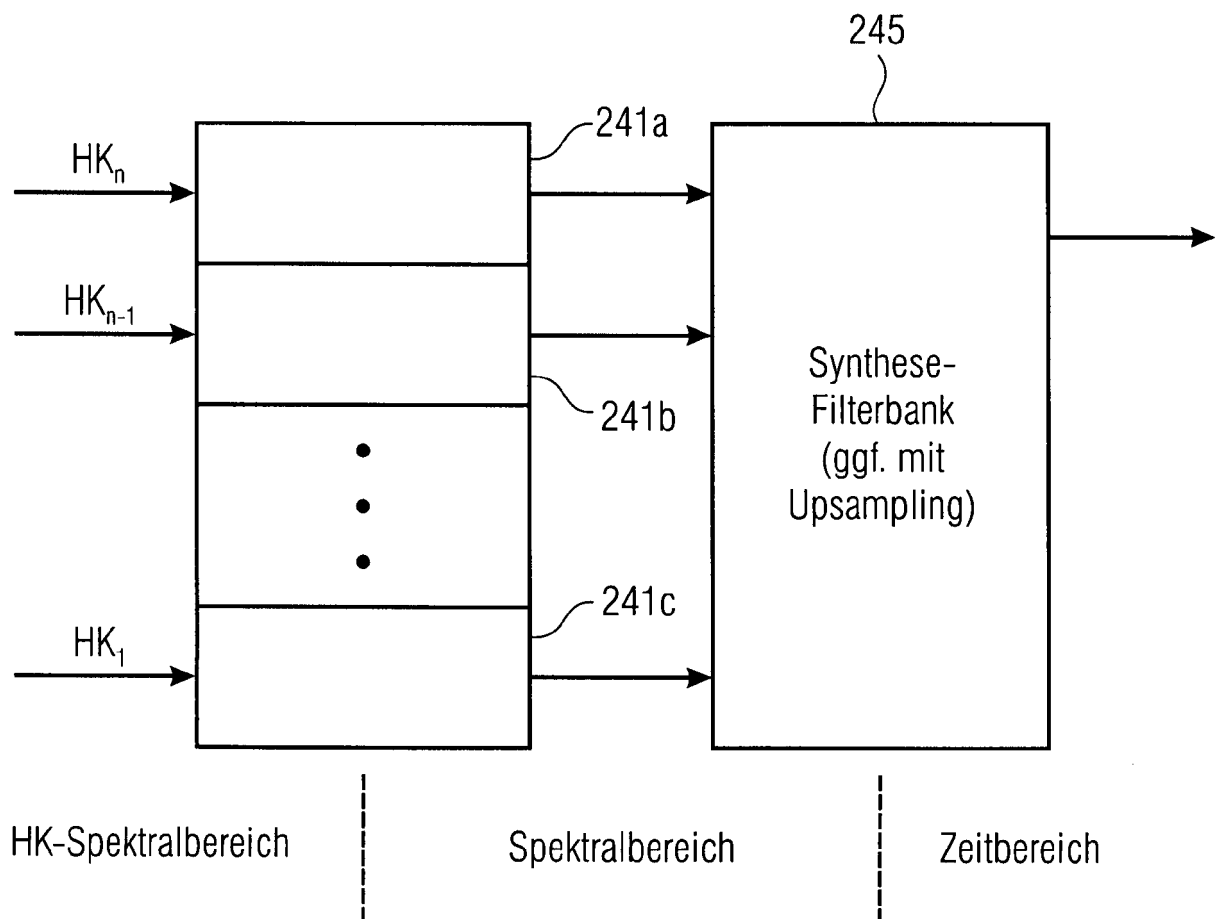
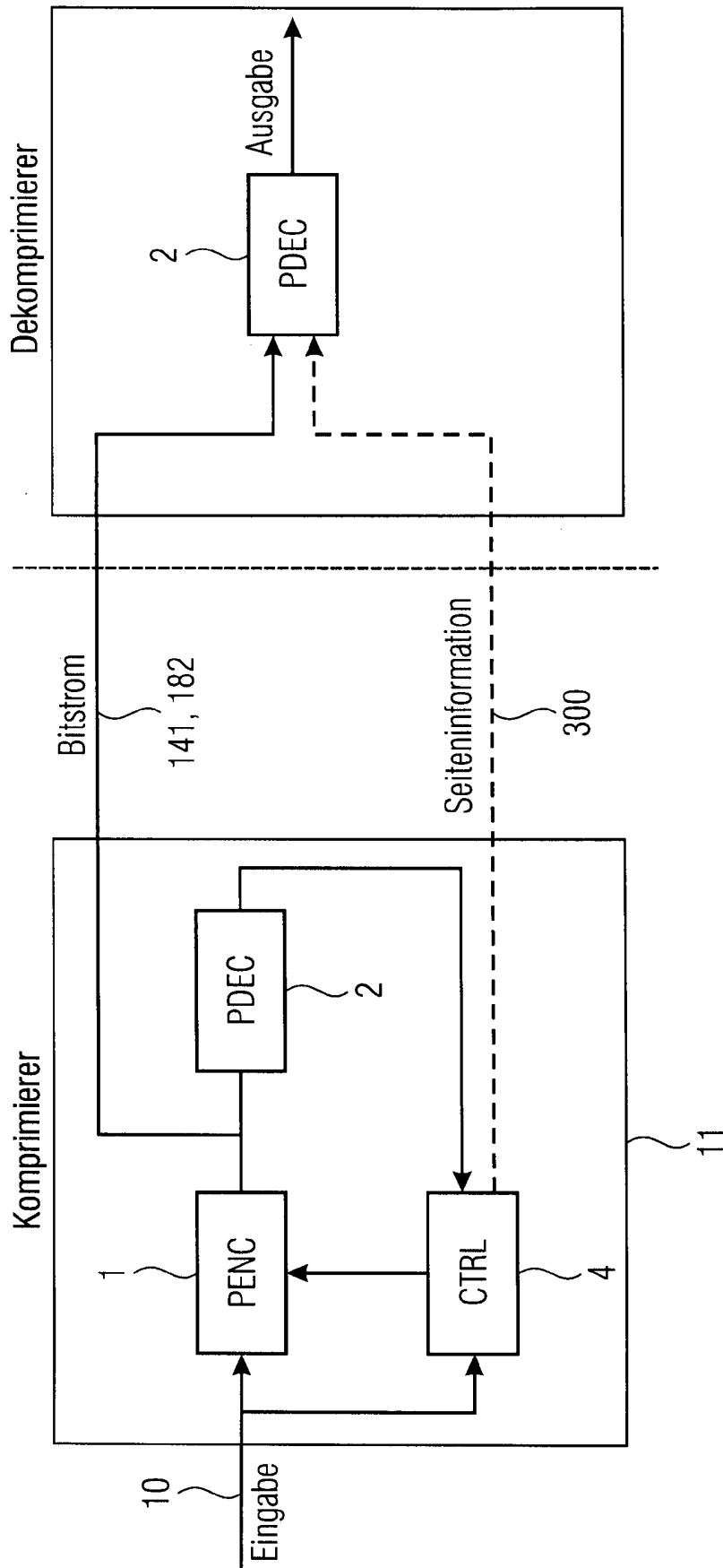


FIG 2D



Steuerung (CTRL)

- vergleicht dekomprimierte Schallfelddaten mit ursprünglichen Schallfelddaten
- unter Berücksichtigung eines psychoakustischen Modells (wie z. B. PEAQ der ITU)
- generiert optimierte Parameter für Aufteilung, Filterbank, Ordnungen
- steuert den Encodierer-Kern PENC

FIG 3A

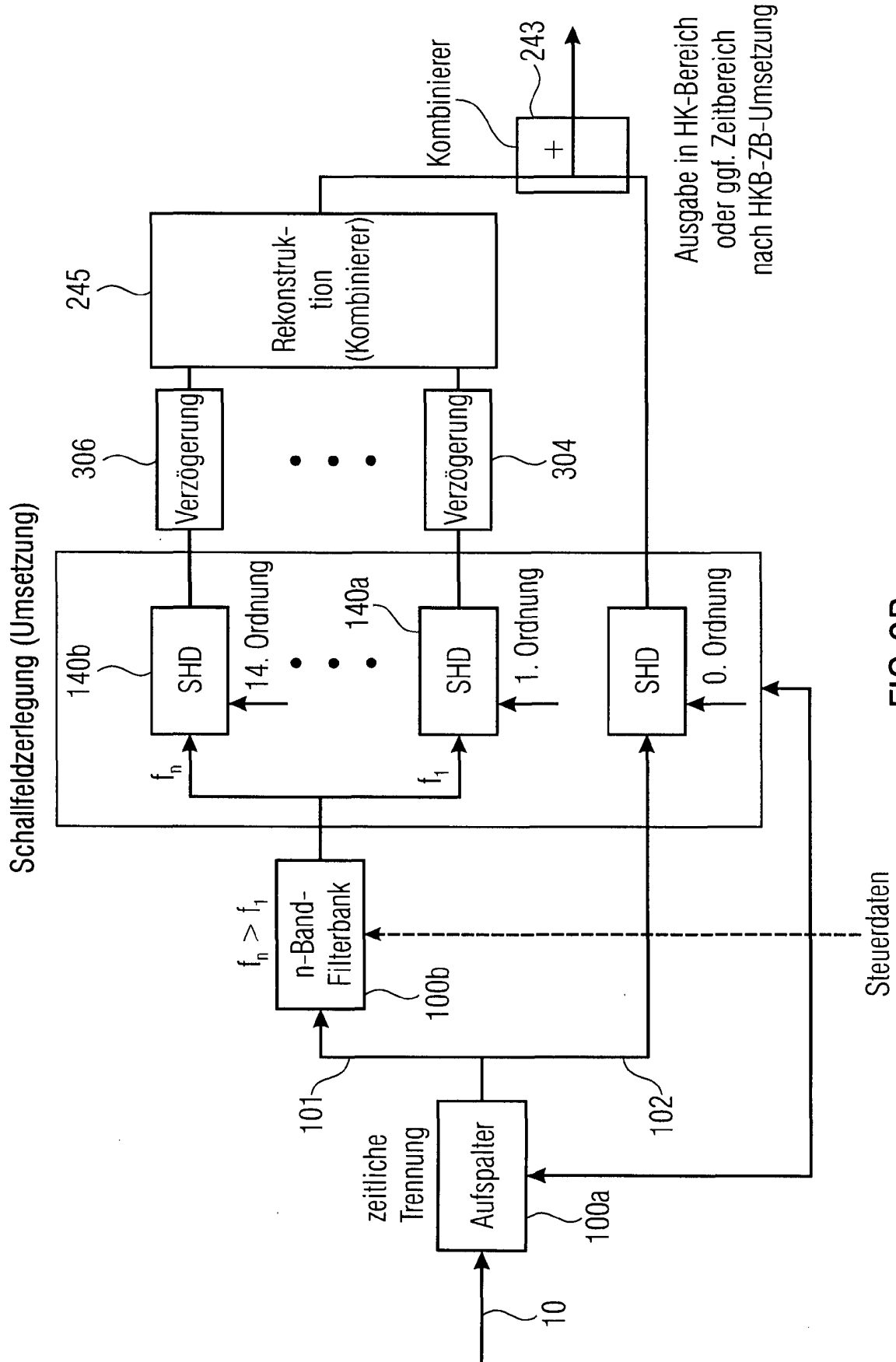


FIG 3B

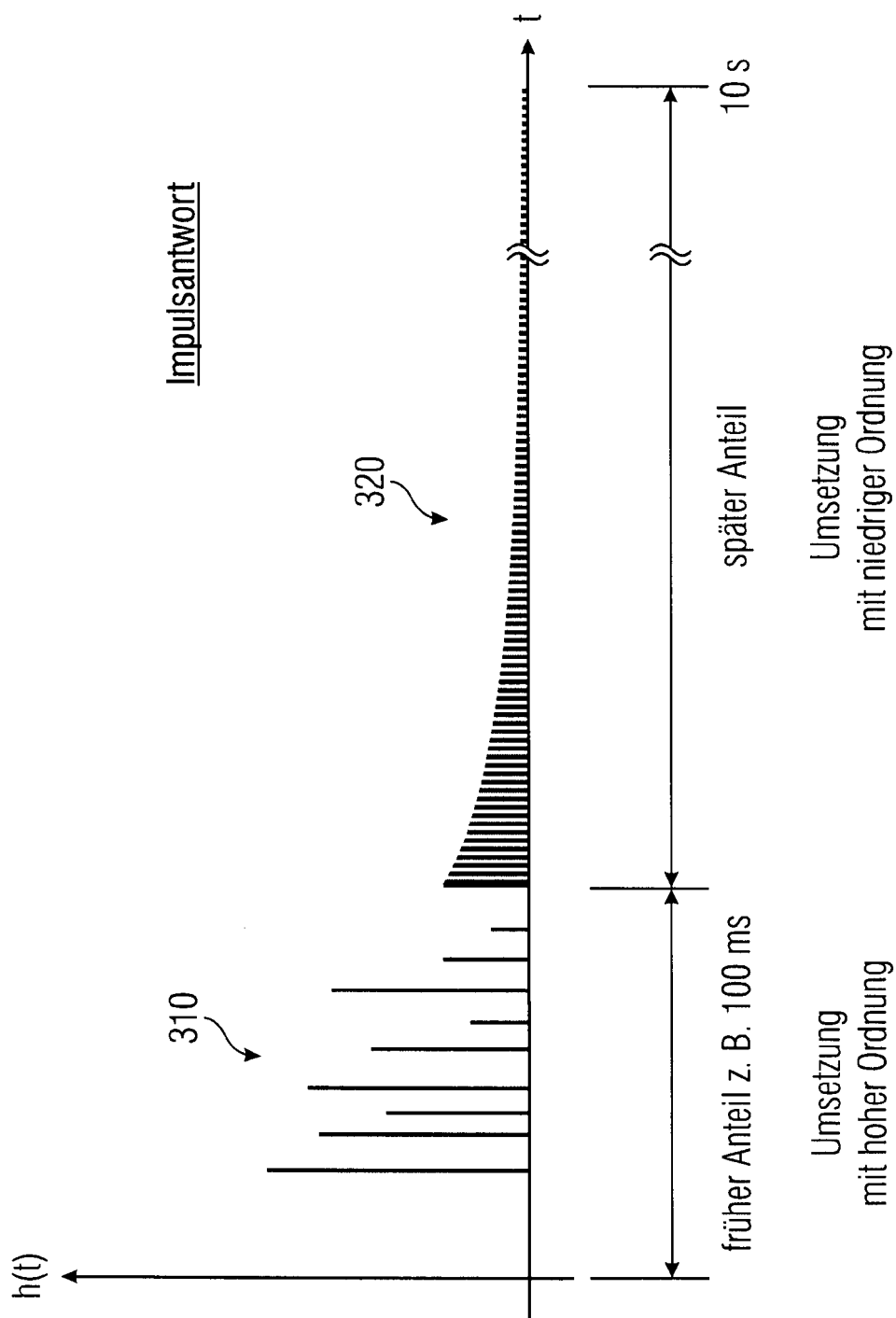
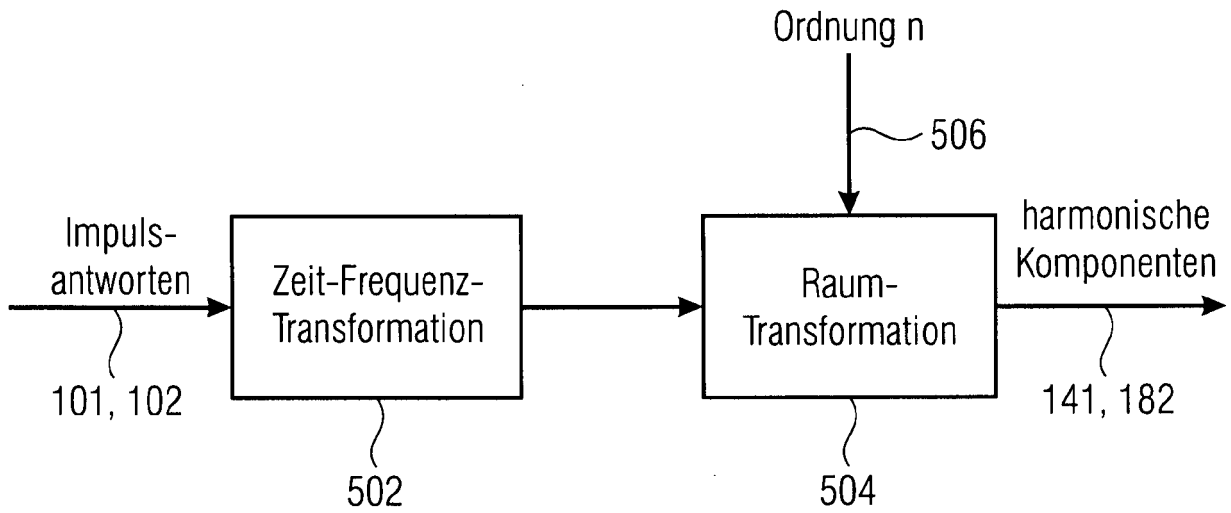


FIG 4



$$Y_n^m(\theta, \phi) = \sqrt{\frac{(2n+1)}{4\pi} \frac{(n-m)!}{(n+m)!}} P_n^m(\cos\theta) e^{im\phi}.$$

508

Y : harmonische Komponenten

n : Ordnung

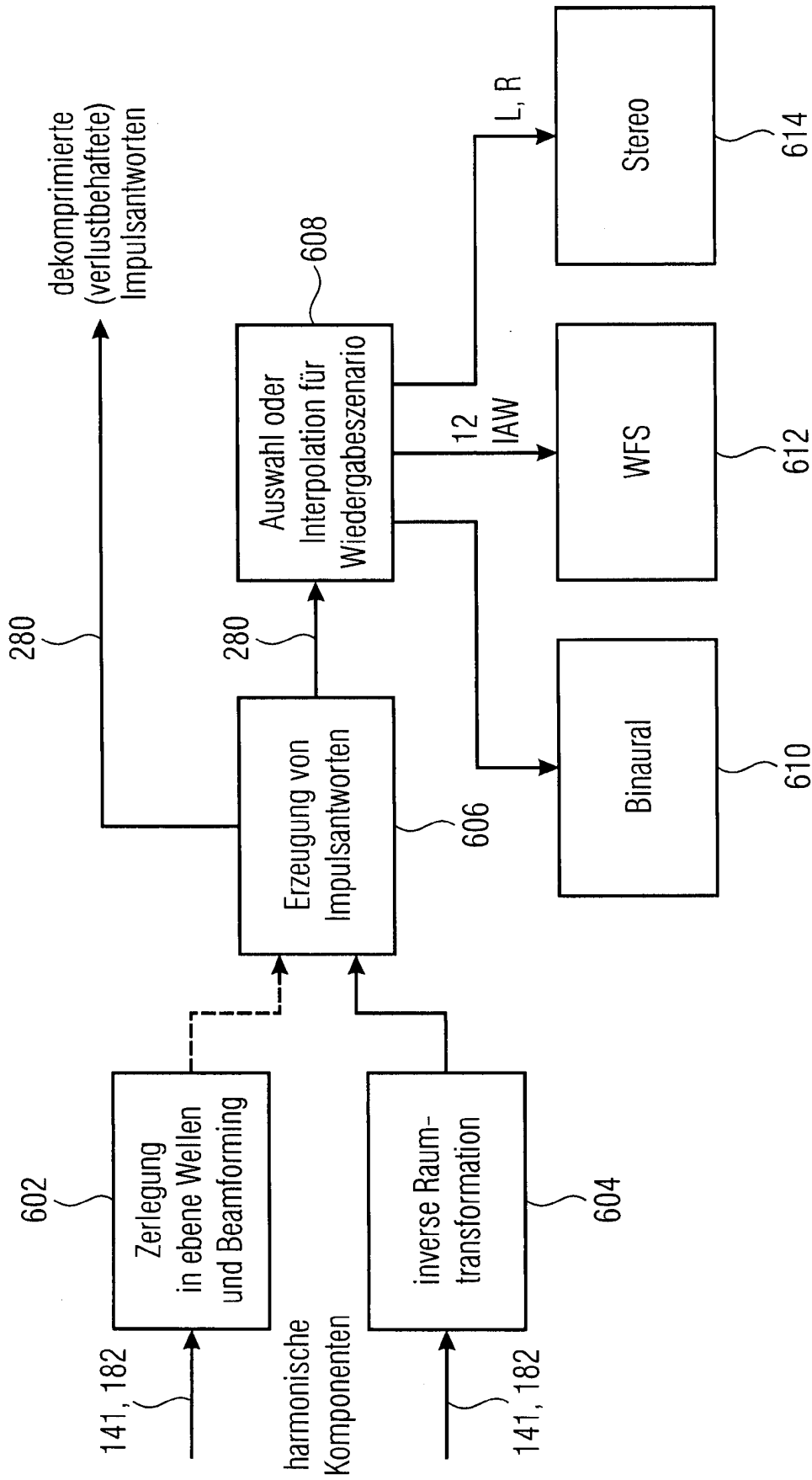
m : Mode

θ : Elevationswinkel

ϕ : Azimuthwinkel

P_n^m : Legendre-Funktionen

FIG 5



4 IAW für späte Reflexionen
 8 IAW für frühe Reflexionen
 (Auswahl nach Position der
 jeweiligen virtuellen Quelle)

FIG 6

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

- EP 1677576 A3 [0009]
- US 8379868 B2 [0009]
- EP 2296142 A2 [0009]
- US 8340306 B2 [0009]
- EP 2592845 A1 [0009]
- EP 2609759 A1 [0009]
- EP 2451196 A1 [0009]
- EP 2205007 A1 [0009]
- WO 2010012478 A2 [0012]

In der Beschreibung aufgeführte Nicht-Patentliteratur

- **HERRE, J et al.** Spatial Audio Coding: Next-generation efficient and compatible coding of multi-channel audio. *AES Convention Paper 6186 presented at the 117th Convention, 2004* [0008]
- **ENGDEGARD, J et al.** Spatial Audio Object Coding (SAOC) - The Upcoming MPEG Standard on Parametric Object Based Audio Coding. *AES Convention Paper 7377 presented at the 125th Convention, 2008* [0008]
- **MERIMAA J ; PULKKI V.** Perceptually-based processing of directional room responses for multichannel loudspeaker reproduction. *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, 2003* [0008]
- **PULKKI, V.** Spatial Sound Reproduction with Directional Audio Coding. *J. Audio Eng. Soc., 2007, vol. 55 (6* [0008]
- **HELLERUD E et al.** Encoding Higher Order Ambisonics with AAC. *AES Convention Paper 7366 presented at the 125th Convention, 2008* [0008]
- **LIINDAU A ; KOSANKE L ; WEINZIERL S.** Perceptual evaluation of physical predictors of the mixing time in binaural room impulse responses. *AES Convention Paper presented at the 128th Convention, 2010* [0008]
- **AVNI, A ; RAFAELY B.** Interaural cross correlation and spatial correlation in a sound field represented by spherical harmonics. *Ambisonics Symposium 2009, 2009* [0008]
- Development and evaluation of a mixed-order Ambisonics playback system. J. Käsbach. Technical University of Denmark, November 2010 [0010]
- **CHRIS TRAVIS.** A New Mixed-Order Scheme for Ambisonic Signals. *Ambisonics Symposium 2009, 25. Juni 2009, 1-6* [0011]
- **I. ELFITRI.** Multichannel Audio Coding Based on Analysis by Synthesis. *Proceedings of the IEEE, 01. April 2011, vol. 99 (4), 657-670* [0013]
- Sound Radiation and Nearfield Acoustical Holography. **EARL G. WILLIAMS.** Fachbuch Fourier Acoustics. Academic Press, 1999 [0049]
- SofiA Sound Field Analysis Toolbox. **BERNSCHÜTZ U.A.** Detmold. ICOSA - International Conference on Spatial Audio, 10. November 2011 [0051]
- Spatial Sound Design based on Measured Room Impulse Responses. **PH. D.-ARBEIT.** TU Delft. 2011 [0054]