



(12) **Veröffentlichung**

der internationalen Anmeldung mit der
 (87) Veröffentlichungs-Nr.: **WO 2022/009694**
 in der deutschen Übersetzung (Art. III § 8 Abs. 2
 IntPatÜbkG)
 (21) Deutsches Aktenzeichen: **11 2021 003 663.7**
 (86) PCT-Aktenzeichen: **PCT/JP2021/024098**
 (86) PCT-Anmeldetag: **25.06.2021**
 (87) PCT-Veröffentlichungstag: **13.01.2022**
 (43) Veröffentlichungstag der PCT Anmeldung
 in deutscher Übersetzung: **27.04.2023**

(51) Int Cl.: **G10L 19/00 (2013.01)**
G10L 19/008 (2013.01)
G10L 19/002 (2013.01)
G10L 19/035 (2013.01)

(30) Unionspriorität:
2020-118174 **09.07.2020** **JP**
2020-170985 **09.10.2020** **JP**

(72) Erfinder:
Kono, Akifumi, Tokyo, JP; Chinen, Toru, Tokyo, JP; Honma, Hiroyuki, Tokyo, JP; Tsuji, Minoru, Tokyo, JP; Oikawa, Yoshiaki, Tokyo, JP

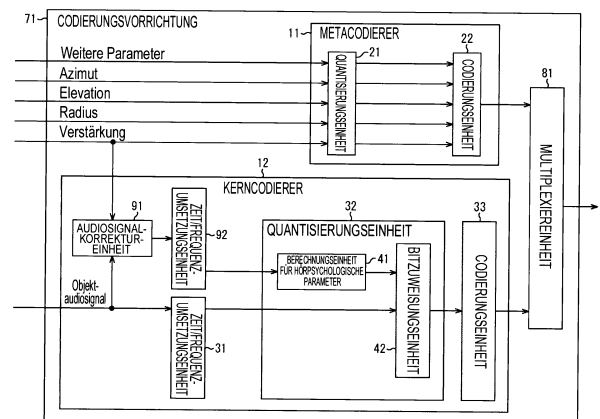
(71) Anmelder:
Sony Group Corporation, Tokyo, JP

(74) Vertreter:
**2SPL Patentanwälte PartG mbB Schuler Schacht
 Platzer Lehmann, 81373 München, DE**

Die folgenden Angaben sind den vom Anmelder eingereichten Unterlagen entnommen.

(54) Bezeichnung: **Signalverarbeitungsvorrichtung, Verfahren und Programm**

(57) Zusammenfassung: Diese Technologie bezieht sich auf eine Signalverarbeitungsvorrichtung, ein Verfahren und ein Programm, die in der Lage sind, eine Codierungseffizienz zu verbessern. Die Signalverarbeitungsvorrichtung weist Folgendes auf: eine Korrektureinheit zum Korrigieren eines Audiosignals eines Audioobjekts auf der Grundlage eines Verstärkungswerts, der in Metadaten des Audioobjekts enthalten ist; und eine Quantisierungseinheit zum Berechnen eines psychologischen Audioparameters auf der Grundlage des korrigierten Signals und Quantisieren des Audiosignals. Diese Technologie ist auf eine Codierungsvorrichtung anwendbar.



Beschreibung

[Technisches Gebiet]

[0001] Die vorliegende Technologie bezieht sich auf eine Signalverarbeitungsvorrichtung, ein Signalverarbeitungsverfahren und ein Programm und insbesondere auf eine Signalverarbeitungsvorrichtung, ein Signalverarbeitungsverfahren und ein Programm, die in der Lage sind, eine Codierungseffizienz zu verbessern.

[Technischer Hintergrund]

[0002] Im verwandten Gebiet waren ein Codieren der vereinheitlichten Sprach- und Audio-Codierungsnorm (USAC-Norm) der Bewegbildeexpertengruppe (MPEG)-D, die eine internationale Norm ist, oder die MPEG-H-3D-Audio-Norm unter Verwendung der MPEG-D-USAC-Norm als ein Kerncodierer und dergleichen bekannt (siehe z. B. NPL 1 bis NPL 3).

[Entgegenhaltungsliste]

[Nichtpatentliteratur]

[NPL 1]

ISO/IEC 23003-3, MPEG-D USAC

[NPL 2]

ISO/IEC 23008-3, MPEG-H 3D Audio

[NPL 3]

ISO/IEC 23008-3:2015/AMENDMENT3,
MPEG-H 3D Audio Phase 2

[Zusammenfassung]

[Technisches Problem]

[0003] In 3D-Audio, das in der MPEG-H-3D-Audio-Norm und dergleichen behandelt wird, ist es möglich, die Richtung, die Entfernung, die Ausbreitung und dergleichen eines dreidimensionalen Schalls mit Metadaten für jedes Objekt wie z. B. horizontale und vertikale Winkel, die die Position eines Schallmaterials (eines Objekts) angeben, eine Entfernung und eine Verstärkung für das Objekt, wiederzugeben. Aus diesem Grund ist es in 3D-Audio möglich, Audio mit einem stärkeren Präsenzgefühl im Vergleich zur Stereowiedergabe des verwandten Gebiets wiederzugeben.

[0004] Allerdings besteht, um Daten einer großen Anzahl von Objekten, die durch 3D-Audio realisiert werden, zu übertragen, die Notwendigkeit einer Codierungstechnologie, die eine größere Anzahl von Audiokanälen mit höherem Kompressionswirkungsgrad mit einer hohen Geschwindigkeit decodieren kann. Das heißt, es besteht Bedarf an einer Verbesserung der Codierungseffizienz.

[0005] Die vorliegende Technologie wird im Hinblick auf derartige Umstände entwickelt und ermöglicht, die Codierungseffizienz zu verbessern.

[Lösung des Problems]

[0006] Eine Signalverarbeitungsvorrichtung gemäß einem ersten Aspekt der vorliegenden Technologie weist eine Korrekturereinheit, die konfiguriert ist, ein Audiosignal eines Audioobjekts auf der Grundlage eines Verstärkungswerts, der in Metadaten des Audioobjekts enthalten ist, zu korrigieren, und eine Quantisierungseinheit, die konfiguriert ist, Hörpsychologieparameter auf der Grundlage eines Signals, das durch die Korrektur erhalten wird, zu berechnen und das Audiosignal zu quantisieren, auf.

[0007] Ein Signalverarbeitungsverfahren oder ein Programm gemäß dem ersten Aspekt der vorliegenden Technologie umfasst ein Korrigieren eines Audiosignals eines Audio-Objekts auf der Grundlage eines Verstärkungswerts, der in Metadaten des Audio-Objekts enthalten ist, ein Berechnen von Hörpsychologieparametern auf der Grundlage eines Signals, das durch die Korrektur erhalten wird, und ein Quantisieren des Audiosignals.

[0008] Im ersten Aspekt der vorliegenden Technologie wird ein Audiosignal eines Audio-Objekts auf der Grundlage eines Verstärkungswerts, der in Metadaten des Audio-Objekts enthalten ist, korrigiert, werden Hörpsychologieparameter auf der Grundlage eines Signals, das durch die Korrektur erhalten wird, berechnet und wird das Audiosignal quantisiert.

[0009] Eine Signalverarbeitungsvorrichtung gemäß einem zweiten Aspekt der vorliegenden Technologie weist eine Änderungseinheit, die konfiguriert ist, einen Verstärkungswert eines Audio-Objekts und ein Audiosignal auf der Grundlage des Verstärkungswerts, der in Metadaten des Audio-Objekts enthalten ist, zu ändern, und eine Quantisierungseinheit, die konfiguriert ist, das geänderte Audiosignal, das durch die Änderung erhalten wird, zu quantisieren, auf.

[0010] Ein Signalverarbeitungsverfahren oder ein Programm gemäß dem zweiten Aspekt der vorliegenden Technologie umfasst ein Modifizieren eines Verstärkungswerts eines Audio-Objekts und eines Audiosignals auf der Grundlage des Verstärkungswerts, der in Metadaten des Audio-Objekts enthalten ist, und ein Quantisieren des geänderten Audiosignals, das durch die Änderung erhalten wird.

[0011] Im zweiten Aspekt der vorliegenden Technologie werden ein Verstärkungswert eines Audio-Objekts und ein Audiosignal auf der Grundlage des Verstärkungswerts, der in Metadaten des Audio-Objekts enthalten ist, geändert und wird das geän-

derte Audiosignal, das durch die Änderung erhalten wird, quantisiert.

[0012] Eine Signalverarbeitungsvorrichtung gemäß einem dritten Aspekt der vorliegenden Technologie weist eine Quantisierungseinheit auf, die konfiguriert ist, Hörpsychologieparameter auf der Grundlage von Metadaten, die einen Verstärkungswert und/oder Positionsinformationen eines Audio-Objekts aufweisen, eines Audiosignals des Audio-Objekts und eines Hörpsychologiemodells, das mit Hörverdeckung zwischen mehreren Audio-Objekten in Beziehung steht, zu berechnen und das Audiosignal auf der Grundlage der Hörpsychologieparameter zu quantisieren.

[0013] Ein Signalverarbeitungsverfahren oder ein Programm gemäß dem dritten Aspekt der vorliegenden Technologie umfasst ein Berechnen von Hörpsychologieparametern auf der Grundlage von Metadaten, die einen Verstärkungswert und/oder Positionsinformationen eines Audio-Objekts aufweisen, eines Audiosignals des Audio-Objekts und eines Hörpsychologiemodells, das mit der Hörverdeckung zwischen mehreren Audio-Objekten in Beziehung steht, und ein Quantisieren des Audiosignals auf der Grundlage der Hörpsychologieparameter.

[0014] Im dritten Aspekt der vorliegenden Technologie werden Hörpsychologieparameter auf der Grundlage von Metadaten, die einen Verstärkungswert und/oder Positionsinformationen eines Audio-Objekts aufweisen, eines Audiosignals des Audio-Objekts und eines Hörpsychologiemodells, das mit der Hörverdeckung zwischen mehreren Audio-Objekten in Beziehung steht, berechnet und wird das Audiosignal auf der Grundlage der Hörpsychologieparameter quantisiert.

[0015] Eine Signalverarbeitungsvorrichtung gemäß einem vierten Aspekt der vorliegenden Technologie weist eine Quantisierungseinheit auf, die konfiguriert ist, ein Audiosignal eines Audio-Objekts unter Verwendung mindestens eines Einstellungsparameters und eines Algorithmus, der für den Typ der Schallquelle, der durch Etiketteninformationen angegeben ist, die den Typ der Schallquelle des Audio-Objekts angeben, bestimmt wird, auf der Grundlage des Audiosignals des Audio-Objekts und der Etiketteninformationen zu quantisieren.

[0016] Ein Signalverarbeitungsverfahren oder ein Programm gemäß dem vierten Aspekt der vorliegenden Technologie umfasst ein Quantisieren eines Audiosignals eines Audio-Objekts unter Verwendung mindestens eines Einstellungsparameters und eines Algorithmus, der für den Typ der Schallquelle, der durch Etiketteninformationen angegeben ist, die den Typ der Schallquelle des Audio-Objekts angeben, bestimmt wird, auf der Grundlage des Audiosig-

nals des Audio-Objekts und der Etiketteninformationen.

[0017] Im vierten Aspekt der vorliegenden Technologie wird ein Audiosignal eines Audio-Objekts unter Verwendung mindestens eines Einstellungsparameters und eines Algorithmus, der für den Typ der Schallquelle, der durch Etiketteninformationen angegeben ist, die den Typ der Schallquelle des Audio-Objekts angeben, bestimmt wird, auf der Grundlage des Audiosignals des Audio-Objekts und der Etiketteninformationen quantisiert.

Figurenliste

[Fig. 1] Fig. 1 ist ein Diagramm, das das Codieren in MPEG-H-3D-Audio veranschaulicht.

[Fig. 2] Fig. 2 ist ein Diagramm, das das Codieren in MPEG-H-3D-Audio veranschaulicht.

[Fig. 3] Fig. 3 ist ein Diagramm, das ein Beispiel eines Wertebereichs veranschaulicht.

[Fig. 4] Fig. 4 ist ein Diagramm, das ein Konfigurationsbeispiel einer Codierungsvorrichtung veranschaulicht.

[Fig. 5] Fig. 5 ist ein Ablaufplan, der eine Codierungsverarbeitung veranschaulicht.

[Fig. 6] Fig. 6 ist ein Diagramm, das ein Konfigurationsbeispiel der Codierungsvorrichtung veranschaulicht.

[Fig. 7] Fig. 7 ist ein Ablaufplan, der eine Codierungsverarbeitung veranschaulicht.

[Fig. 8] Fig. 8 ist ein Diagramm, das ein Konfigurationsbeispiel der Codierungsvorrichtung veranschaulicht.

[Fig. 9] Fig. 9 ist ein Diagramm, das eine Änderung von Verstärkungswerten veranschaulicht.

[Fig. 10] Fig. 10 ist ein Diagramm, das eine Änderung eines Audiosignals gemäß der Änderung eines Verstärkungswerts veranschaulicht.

[Fig. 11] Fig. 11 ist ein Diagramm, das eine Änderung eines Audiosignals gemäß der Änderung eines Verstärkungswerts veranschaulicht.

[Fig. 12] Fig. 12 ist ein Ablaufplan, der eine Codierungsverarbeitung veranschaulicht.

[Fig. 13] Fig. 13 ist ein Diagramm, das Hörcharakteristiken von rosa Rauschen veranschaulicht.

[Fig. 14] Fig. 14 ist ein Diagramm, das eine Korrektur eines Verstärkungswerts unter Verwendung einer Hörcharakteristiktabelle veranschaulicht.

[Fig. 15] Fig. 15 ist ein Diagramm, das ein Beispiel einer Hörcharakteristiktafel veranschaulicht.

[Fig. 16] Fig. 16 ist ein Diagramm, das ein Beispiel einer Hörcharakteristiktafel veranschaulicht.

[Fig. 17] Fig. 17 ist ein Diagramm, das ein Beispiel einer Hörcharakteristiktafel veranschaulicht.

[Fig. 18] Fig. 18 ist ein Diagramm, das ein Beispiel einer Interpolation von Verstärkungskorrekturwerten veranschaulicht.

[Fig. 19] Fig. 19 ist ein Diagramm, das ein Konfigurationsbeispiel einer Codierungsvorrichtung veranschaulicht.

[Fig. 20] Fig. 20 ist ein Ablaufplan, der eine Codierungsverarbeitung veranschaulicht.

[Fig. 21] Fig. 21 ist ein Diagramm, das ein Konfigurationsbeispiel einer Codierungsvorrichtung veranschaulicht.

[Fig. 22] Fig. 22 ist ein Ablaufplan, der eine Codierungsverarbeitung veranschaulicht.

[Fig. 23] Fig. 23 ist ein Diagramm, das ein Syntaxbeispiel der Konfiguration von Metadaten veranschaulicht.

[Fig. 24] Fig. 24 ist ein Diagramm, das ein Konfigurationsbeispiel einer Codierungsvorrichtung veranschaulicht.

[Fig. 25] Fig. 25 ist ein Ablaufplan, der eine Codierungsverarbeitung veranschaulicht.

[Fig. 26] Fig. 26 ist ein Diagramm, das ein Konfigurationsbeispiel einer Codierungsvorrichtung veranschaulicht.

[Fig. 27] Fig. 27 ist ein Ablaufplan, der eine Codierungsverarbeitung veranschaulicht.

[Fig. 28] Fig. 28 ist ein Diagramm, das ein Konfigurationsbeispiel einer Codierungsvorrichtung veranschaulicht.

[Fig. 29] Fig. 29 ist ein Ablaufplan, der eine Codierungsverarbeitung veranschaulicht.

[Fig. 30] Fig. 30 ist ein Diagramm, das ein Konfigurationsbeispiel eines Computers veranschaulicht.

[Beschreibung der Ausführungsformen]

[0018] Im Folgenden werden Ausführungsformen, auf die die vorliegende Technologie angewendet wird, unter Bezugnahme auf die Zeichnungen beschrieben.

<Erste Ausführungsform>

<Vorliegende Technologie>

[0019] Die vorliegende Technologie kann eine Codierungseffizienz (eine Komprimierungseffizienz) durch Berechnen von Hörpsychologieparametern, die für eine tatsächliche Hörempfindung geeignet sind, und Durchführen einer Bitzuweisung unter Berücksichtigung einer Verstärkung von Metadaten, die beim Rendern während des Betrachtens angewendet werden, verbessern.

[0020] Zunächst wird ein Codieren von Metadaten und eines Audiosignals eines Audio-Objekts (das im Folgenden einfach als ein Objekt bezeichnet wird) in MPEG-H-3D-Audio beschrieben.

[0021] In MPEG-H-3D-Audio werden Metadaten eines Objekts durch einen Metacodierer codiert und wird ein Audiosignal des Objekts durch einen Kerncodierer codiert, wie in Fig. 1 veranschaulicht ist.

[0022] Speziell quantisiert der Metacodierer Parameter, die Metadaten bilden, und codiert die resultierenden quantisierten Parameter, um codierte Metadaten zu erhalten.

[0023] Zusätzlich führt der Kerncodierer eine Zeit-/Frequenz-Umsetzung unter Verwendung einer modifizierten diskreten Kosinustransformation (MDCT) am Audiosignal durch und quantisiert den resultierenden MDCT-Koeffizienten, um den quantisierten MDCT-Koeffizienten zu erhalten. Eine Bitzuweisung wird während der Quantisierung des MDCT-Koeffizienten auch durchgeführt. Ferner codiert der Kerncodierer den quantisierten MDCT-Koeffizienten, um codierte Audiodaten zu erhalten.

[0024] Dann werden die codierten Metadaten und die codierten Audiodaten, die auf diese Weise erhalten werden, als ein einzelner Bitstrom zusammengesetzt und ausgegeben.

[0025] Hier wird das Codieren von Metadaten und eines Audiosignals in MPEG-H-3D-Audio unter Bezugnahme auf Fig. 2 ausführlicher beschrieben.

[0026] In diesem Beispiel werden mehrere Parameter in den Metacodierer 11 als Metadaten eingegeben und wird ein Audiosignal, das ein Zeitsignal (ein Wellenformsignal) zum Wiedergeben eines Schalls eines Objekts ist, in den Kerncodierer 12 eingegeben.

[0027] Der Metacodierer 11 weist eine Quantisierungseinheit 21 und eine Codierungseinheit 22 auf und Metadaten werden in die Quantisierungseinheit 21 eingegeben.

[0028] Wenn im Metacodierer 11 eine Metadaten-codierungsverarbeitung gestartet wird, ersetzt die Quantisierungseinheit 21 zuerst nach Bedarf den Wert jedes Metadatenparameters mit einem oberen Grenzwert oder einem unteren Grenzwert und quantisiert dann die Parameter, um quantisierte Parameter zu erhalten.

[0029] In diesem Beispiel werden ein horizontaler Winkel (ein Azimut) ein vertikaler Winkel (eine Elevation), eine Entfernung (ein Radius), ein Verstärkungswert (eine Verstärkung) und weitere Parameter in die Quantisierungseinheit 21 als Parameter, die Metadaten bilden, eingegeben.

[0030] Hier sind der horizontale Winkel (der Azimut) und der vertikale Winkel (die Elevation) Winkel in der horizontalen Richtung und der vertikalen Richtung, die die Position des Objekts aus einer Bezugshörposition in einem dreidimensionalen Raum betrachtet angeben. Ferner gibt die Entfernung (der Radius) die Position des Objekts im dreidimensionalen Raum an und gibt eine Entfernung von der Bezugshörposition zum Objekt an. Informationen, die aus dem horizontalen Winkel, dem vertikale Winkel und der Entfernung bestehen, sind Positionsinformationen, die die Position des Objekts angeben.

[0031] Ferner ist der Verstärkungswert (die Verstärkung) eine Verstärkung zur Verstärkungskorrektur eines Audiosignals des Objekts und sind die weiteren Parameter Parameter zur Ausbreitungsverbreitung zum Erweitern eines Schallbilds, der Priorität des Objekts und dergleichen.

[0032] Jeder Parameter, der Metadaten bildet, wird zu einem Wert in einem Wertebereich gesetzt, der ein vorgegebener Bereich ist, der in **Fig. 3** veranschaulicht ist.

[0033] Im Beispiel in **Fig. 3** ist ein Wertebereich jedes Parameters, der Metadaten bildet, veranschaulicht.

[0034] Es ist festzuhalten, dass in **Fig. 3**, „Ausbreitung“, „Ausbreitungsbreite“, „Ausbreitungshöhe“ und „Ausbreitungstiefe“ Parameter zur Ausbreitungsverarbeitung sind und Beispiele von weiteren Parametern sind. Zusätzlich ist „dynamische Objektpriorität“ ein Parameter, der die Priorität eines Objekts angibt und ist dieser Parameter auch ein Beispiel weiterer Parameter.

[0035] Zum Beispiel ist in diesem Beispiel der Wertebereich des horizontalen Winkels (des Azimuts) von einem unteren Grenzwert von -180 Grad zu einem oberen Grenzwert von 180 Grad.

[0036] Falls die horizontale Winkeleingabe in die Quantisierungseinheit 21 den Wertebereich über-

schreitet, d. h. falls der horizontale Winkel aus dem Bereich fällt, wird der horizontale Winkel durch den unteren Grenzwert „-180“ oder den oberen Grenzwert „180“ ersetzt und dann quantisiert. Das heißt, wenn der eingegebene horizontale Winkel ein Wert größer als der obere Grenzwert ist, wird der obere Grenzwert „180“ als der horizontale Winkel nach einer Beschränkung (einem Austausch) eingestellt, und wenn der horizontale Winkel ein kleinerer Wert als der untere Grenzwert ist, wird der untere Grenzwert „-180“ als der horizontale Winkel nach einer Beschränkung eingestellt.

[0037] Zusätzlich ist z. B. der Wertebereich des Verstärkungswerts (der Verstärkung) von einem unteren Grenzwert von 0,004 zu einem oberen Grenzwert von 5,957. Insbesondere ist hier der Verstärkungswert als ein geradliniger Wert beschrieben.

[0038] Zurückkommend zur Beschreibung von **Fig. 2** werden dann, wenn Parameter, die Metadaten bilden, durch die Quantisierungseinheit 21 quantisiert werden und die quantisierten Parameter erhalten werden, die quantisierten Parameter durch die Codierungseinheit 22 codiert und werden die resultierenden codierten Metadaten ausgegeben. Zum Beispiel führt die Codierungseinheit 22 ein Differenzcodieren an den quantisierten Parametern durch, um codierte Metadaten zu erzeugen.

[0039] Zusätzlich weist der Kerncodierer 12 eine Zeitfrequenzumsetzungseinheit 31, eine Quantisierungseinheit 32 und eine Codierungseinheit 33 auf und wird ein Audiosignal eines Objekts in die Zeitfrequenzumsetzungseinheit 31 eingegeben. Zusätzlich weist die Quantisierungseinheit 32 eine Hörpsychologieparameterberechnungseinheit 41 und eine Bit-zuweisungseinheit 42 auf.

[0040] Im Kerncodierer 12 führt, wenn die Codierungsverarbeitung für das Audiosignal gestartet wird, die Zeitfrequenzumsetzungseinheit 31 zuerst eine MDCT, d. h. eine Zeit/Frequenz-Umsetzung am Eingangsaudiosignal durch und wird folglich ein MDCT-Koeffizient, was Frequenzspektruminformationen sind, erhalten.

[0041] Dann wird in der Quantisierungseinheit 32 der MDCT-Koeffizient, der durch die Zeit/Frequenz-Umsetzung (die MDCT) erhalten wird, für jedes Skalierungsfaktorband quantisiert und folglich wird ein quantisierter MDCT-Koeffizient erhalten.

[0042] Hier ist das Skalierungsfaktorband ein Band (ein Frequenzband), das durch Bündelung mehrerer Unterbänder mit einer vorgegebenen Bandbreite, die die Auflösung eines Quadraturspiegelfilter-Analysefilters (QMF-Analysefilter) ist, erhalten wird.

[0043] Speziell berechnet in der Quantisierung, die durch die Quantisierungseinheit 32 durchgeführt wird, die Hörpsychologieparameter-Berechnungseinheit 41 Hörpsychologieparameter zum Berücksichtigen von menschlichen Hörcharakteristiken (der Hörverdeckung) für den MDCT-Koeffizienten.

[0044] Ferner wird in der Bitzuweisungseinheit 42 der MDCT-Koeffizient, der durch die Zeit/Frequenz-Umsetzung erhalten wird, und die Hörpsychologieparameter, die durch die Hörpsychologieparameter-Berechnungseinheit 41 erhalten werden, verwendet, um eine Bitzuweisung auf der Grundlage eines Hörpsychologiemodells zum Berechnen und Bewerten von quantisierten Bits und quantisiertem Rauschen jedes Skalierungsfaktorbands durchzuführen.

[0045] Dann quantisiert die Bitzuweisungseinheit 42 den MDCT-Koeffizienten für jedes Skalierungsfaktorband auf der Grundlage eines Ergebnisses der Bitzuweisung und liefert den resultierenden quantisierten MDCT-Koeffizienten zur Codierungseinheit 33.

[0046] Auf diese Weise wird ein Teil der Quantisierungsbits des Skalierungsfaktorbands, in dem das Quantisierungsrauschen, das durch die Quantisierung des MDCT-Koeffizienten erzeugt wird, verdeckt ist und nicht wahrgenommen wird, dem Skalierungsfaktorband, in dem das Quantisierungsrauschen einfach wahrgenommen wird, zugewiesen (gewandelt). Dadurch ist es möglich, eine Verschlechterung der Klangqualität insgesamt zu unterdrücken und eine effiziente Quantisierung durchzuführen. Das heißt, es ist möglich, die Codierungseffizienz zu verbessern.

[0047] Ferner wird in der Codierungseinheit 33 z. B. eine kontextbasierte Arithmetikcodierung am quantisierten MDCT-Koeffizienten, der von der Bitzuweisungseinheit 42 geliefert wird, durchgeführt und werden die resultierenden codierten Audiodaten als codierte Daten eines Audiosignals ausgegeben.

[0048] Wie oben beschrieben ist, werden Metadaten eines Objekts und ein Audiosignal durch den Metacodierer 11 und den Kerncodierer 12 codiert.

[0049] Im Übrigen wird der MDCT-Koeffizient, der verwendet wird, um Hörpsychologieparameter zu berechnen, durch Durchführen einer MDCT, d. h. einer Zeit/Frequenz-Umsetzung, am Eingangsaudiosignal erhalten.

[0050] Allerdings werden, wenn das tatsächlich codierte Audiosignal decodiert, gerendert und betrachtet wird, Verstärkungswerte von Metadaten angewendet und tritt somit eine Diskrepanz zwischen Audiosignalen, die während des Berechnens von Hörpsychologieparametern und während des Betrachtens verwendet werden, auf.

[0051] Aus diesem Grund kann eine Verringerung der Codierungseffizienz wie z. B. unter Verwendung von zusätzlichen Bits, um die Erzeugung eines Quantisierungsrauschens, das ursprünglich in einer Hörempfindung für ein vorgegebenes Skalierungsfaktorband nicht hörbar ist, zu verhindern, auftreten.

[0052] Folglich werden in der vorliegenden Technologie, Hörpsychologieparameter unter Verwendung eines korrigierten MDCT-Koeffizienten berechnet, auf den Verstärkungswerte von Metadaten angewendet werden, und somit ist es möglich, Hörpsychologieparameter, die stärker an die tatsächliche Hörempfindung angepasst sind, zu erhalten und die Codierungseffizienz zu verbessern.

<Konfigurationsbeispiel der Codierungsvorrichtung>

[0053] Fig. 4 ist ein Diagramm, das ein Konfigurationsbeispiel einer Ausführungsform einer Codierungsvorrichtung veranschaulicht, auf die die vorliegende Technologie angewendet wird. Es ist festzuhalten, dass in Fig. 4, Abschnitte, die denen in Fig. 2 entsprechen, durch dieselben Bezugsnummern und -zeichen bezeichnet werden, und ihre Beschreibung geeignet ausgelassen wird.

[0054] Eine Codierungsvorrichtung 71, die in Fig. 4 veranschaulicht ist, ist durch eine Signalverarbeitungsvorrichtung wie z. B. einen Server implementiert, die den Inhalt eines Audio-Objekts verteilt und weist einen Metacodierer 11, einen Kerncodierer 12 und eine Multiplexiereinheit 81 auf.

[0055] Ferner weist der Metacodierer 11 eine Quantisierungseinheit 21 und eine Codierungseinheit 22 auf und weist der Kerncodierer 12 eine Audiosignalkorrekturereinheit 91, eine Zeitfrequenzumsetzungseinheit 92, eine Zeitfrequenzumsetzungseinheit 31, eine Quantisierungseinheit 32 und eine Codierungseinheit 33 auf.

[0056] Ferner weist die Quantisierungseinheit 32 eine Hörpsychologieparameter-Berechnungseinheit 41 und eine Bitzuweisungseinheit 42 auf.

[0057] Die Codierungsvorrichtung 71 ist derart konfiguriert, dass eine Multiplexiereinheit 81, eine Audiosignalkorrekturereinheit 91 und eine Zeitfrequenzumsetzungseinheit 92 zu der Konfiguration, die in Fig. 2 veranschaulicht ist, erneut hinzugefügt werden und besitzt sonst dieselbe Konfiguration wie die, die in Fig. 2 veranschaulicht ist.

[0058] Im Beispiel von Fig. 4 multiplexiert die Multiplexiereinheit 81 codierte Metadaten, die von der Codierungseinheit 22 geliefert werden, und codierte Audiodaten, die von der Codierungseinheit 33 geliefert werden, um einen Bitstrom zu erzeugen und auszugeben.

[0059] Zusätzlich werden ein Audiosignal eines Objekts und Verstärkungswerte des Objekts, die Metadaten bilden, zur Audiosignalkorrekturereinheit 91 geliefert.

[0060] Die Audiosignalkorrekturereinheit 91 führt eine Verstärkungskorrektur am gelieferten Audiosignal auf der Grundlage des gelieferten Verstärkungswerts durch und liefert das Audiosignal, das der Verstärkungskorrektur unterworfen wurde, zur Zeitfrequenzumsetzungseinheit 92. Zum Beispiel multipliziert die Audiosignalkorrekturereinheit 91 das Audiosignal mit dem Verstärkungswert, um eine Verstärkungskorrektur des Audiosignals durchzuführen. Das heißt, hier wird die Korrektur am Audiosignal in einer Zeitdomäne durchgeführt.

[0061] Die Zeitfrequenzumsetzungseinheit 92 führt eine MDCT am Audiosignal, das von der Audiosignalkorrekturereinheit 91 geliefert wird, durch und liefert den resultierenden MDCT-Koeffizienten zur Hörpsychologieparameterberechnungseinheit 41.

[0062] Es ist festzuhalten, dass im Folgenden, das Audiosignal, das durch die Verstärkungskorrektur in der Audiosignalkorrekturereinheit 91 erhalten wird, speziell auch als ein korrigiertes Audiosignal bezeichnet wird und der MDCT-Koeffizient, der durch die MDCT in der Zeitfrequenzumsetzungseinheit 92 erhalten wird, speziell als ein korrigierter MDCT-Koeffizient bezeichnet wird.

[0063] Ferner wird in diesem Beispiel der MDCT-Koeffizient, der durch die Zeitfrequenzumsetzungseinheit 31 erhalten wird, nicht zur Hörpsychologieparameterberechnungseinheit 41 geliefert und werden in der Hörpsychologieparameterberechnungseinheit 41 Hörpsychologieparameter auf der Grundlage des korrigierten MDCT-Koeffizienten, der von der Zeitfrequenzumsetzungseinheit 92 geliefert wird, berechnet.

[0064] In der Codierungsvorrichtung 71 führt die Audiosignalkorrekturereinheit 91 am Kopf eine Verstärkungskorrektur an einem Eingangsaudiosignal eines Objekts durch Anwenden von Verstärkungswerten, die in Metadaten enthalten sind, genau so wie während des Renderns durch.

[0065] Danach führt die Zeitfrequenzumsetzungseinheit 92 eine MDCT am korrigierten Audiosignal, das durch die Verstärkungskorrektur erhalten wird, getrennt von der zur Bitzuweisung durch, um einen korrigierten MDCT-Koeffizienten zu erhalten.

[0066] Dann werden schließlich Hörpsychologieparameter durch die Hörpsychologieparameterberechnungseinheit 41 auf der Grundlage des korrigierten MDCT-Koeffizienten berechnet, wodurch Hörpsychologieparameter erhalten werden, die stär-

ker an die tatsächliche Hörempfindung angepasst sind als im Falle von **Fig. 2**.

[0067] Dies ist darauf zurückzuführen, dass ein Schall auf der Grundlage des korrigierten Audiosignals näher bei einem Schall auf der Grundlage eines Signals, das durch Rendern auf der Decodierungsseite erhalten wird, liegt als ein Schall auf der Grundlage des ursprünglichen Audiosignals. Auf diese Weise werden Quantisierungsbits jedem Skalierungsfaktorband geeigneter zugewiesen und kann die Codierungseffizienz verbessert werden.

[0068] Es ist festzuhalten, dass, obwohl hier ein Beispiel beschrieben wurde, in dem Verstärkungswerte von Metadaten vor der Quantisierung zur Verstärkungskorrektur in der Audiosignalkorrekturereinheit 91 verwendet werden, Verstärkungswerte nach dem Codieren oder der Quantisierung zur Audiosignalkorrekturereinheit 91 geliefert und zur Verstärkungskorrektur verwendet werden können.

[0069] In einem derartigen Fall werden die Verstärkungswerte nach dem Codieren oder der Quantisierung in der Audiosignalkorrekturereinheit 91 decodiert oder umgekehrt quantisiert und wird eine Verstärkungskorrektur eines Audiosignals auf der Grundlage von Verstärkungswerten durchgeführt, die als Ergebnis des Decodierens oder der Quantisierung erhalten werden, um ein korrigiertes Audiosignal zu erhalten.

<Beschreibung der Codierungsverarbeitung>

[0070] Als nächstes wird der Betrieb der Codierungsvorrichtung 71, die in **Fig. 4** veranschaulicht ist, beschrieben. Das heißt, die Codierungsverarbeitung, die durch die Codierungsvorrichtung 71 durchgeführt wird, wird unten unter Bezugnahme auf einen Ablaufplan von **Fig. 5** beschrieben.

[0071] In Schritt S11 quantisiert die Quantisierungseinheit 21 Parameter als gelieferte Metadaten und liefert die resultierenden quantisierten Parameter zur Codierungseinheit 22.

[0072] Zum jetzigen Zeitpunkt führt die Quantisierungseinheit 21 eine Quantisierung durch, nachdem Parameter größer als ein vorgegebener Wertebereich durch einen oberen Grenzwert des Wertebereichs ersetzt worden sind, und führt ähnlich eine Quantisierung nach dem Ersetzen von Parametern, die kleiner als der Wertebereich sind, durch einen unteren Grenzwert durch.

[0073] In Schritt S12 führt die Codierungseinheit 22 ein Differenzcodieren an den quantisierten Parametern, die aus der Quantisierungseinheit 21 geliefert werden, durch und liefert die resultierenden codierten Metadaten zur Multiplexiereinheit 81.

[0074] In Schritt S13 führt die Audiosignalkorrektur-einheit 91 eine Verstärkungskorrektur auf der Grundlage von Verstärkungswerten der gelieferten Metadaten an einem gelieferten Audiosignal eines Objekts durch und liefert das resultierende korrigierte Audiosignal zur Zeitfrequenzumsetzungseinheit 92.

[0075] In Schritt S14 führt die Zeitfrequenzumsetzungseinheit 92 eine MDCT (eine Zeit/Frequenz-Umsetzung) an dem korrigierten Audiosignal, das von der Audiosignalkorrektur-einheit 91 geliefert wird, durch und liefert den resultierenden korrigierten MDCT-Koeffizienten zur Hörpsychologieparameter-berechnungseinheit 41.

[0076] In Schritt S15 führt die Zeitfrequenzumsetzungseinheit 31 eine MDCT (eine Zeit/Frequenz-Umsetzung) am gelieferten Audiosignal des Objekts durch und liefert den resultierenden MDCT-Koeffizienten zur Bitzuweisungseinheit 42.

[0077] In Schritt S16 berechnet die Hörpsychologieparameter-Berechnungseinheit 41 Hörpsychologieparameter auf der Grundlage des korrigierten MDCT-Koeffizienten, der von der Zeitfrequenzumsetzungseinheit 92 geliefert wird, und liefert die berechneten Hörpsychologieparameter zur Bitzuweisungseinheit 42.

[0078] In Schritt S17 führt die Bitzuweisungseinheit 42 eine Bitzuweisung auf der Grundlage eines Hörpsychologiemodells auf der Grundlage der Hörpsychologieparameter, die von der Hörpsychologieparameterberechnungseinheit 41 geliefert werden, und des MDCT-Koeffizienten, der von der Zeitfrequenzumsetzungseinheit 31 geliefert wird, durch und quantisiert den MDCT-Koeffizienten für jedes Skalierungsfaktorband auf der Grundlage der Ergebnisse. Die Bitzuweisungseinheit 42 liefert den quantisierten MDCT-Koeffizient, der durch die Quantisierung erhalten wird, zur Codierungseinheit 33.

[0079] In Schritt S18 führt die Codierungseinheit 33 ein kontextbasiertes Arithmetikcodieren an dem quantisierten MDCT-Koeffizienten, der von der Bitzuweisungseinheit 42 geliefert wird, durch und liefert die resultierenden codierten Audiodaten zur Multiplexiereinheit 81.

[0080] In Schritt S19 multiplexiert die Multiplexiereinheit 81 die codierten Metadaten, die von der Codierungseinheit 22 geliefert werden, und die codierten Audiodaten, die von der Codierungseinheit 33 geliefert werden, um einen Bitstrom zu erzeugen und auszugeben.

[0081] Wenn der Bitstrom auf diese Weise ausgegeben wird, ist die Codierungsverarbeitung beendet.

[0082] Wie oben beschrieben ist, korrigiert die Codierungsvorrichtung 71 das Audiosignal auf der Grundlage der Verstärkungswerte der Metadaten vor dem Codieren und berechnet Hörpsychologieparameter auf der Grundlage des resultierenden korrigierten Audiosignals. Auf diese Weise ist es möglich, Hörpsychologieparameter zu erhalten, die stärker an die tatsächliche Hörempfindung angepasst sind, und die Codierungseffizienz zu verbessern.

<Zweite Ausführungsform>

<Konfigurationsbeispiel der Codierungsvorrichtung>

[0083] Im Übrigen muss die Codierungsvorrichtung 71, die in **Fig. 4** veranschaulicht ist, eine MDCT zweimal durchführen und erhöht sich somit eine Rechenlast (der Berechnungsaufwand). Folglich kann der Berechnungsaufwand durch Korrigieren eines MDCT-Koeffizienten (von Audiosignalen) in einer Frequenzdomäne verringert werden.

[0084] In einem derartigen Fall ist die Codierungsvorrichtung 71 z. B. konfiguriert, wie in **Fig. 6** veranschaulicht ist. Es ist festzuhalten, dass in **Fig. 6** Abschnitte, die denen in **Fig. 4** entsprechen, durch dieselben Bezugsnummern und -zeichen bezeichnet werden, und ihre Beschreibung geeignet ausgelassen wird.

[0085] Die Codierungsvorrichtung 71, die in **Fig. 6** veranschaulicht ist, weist einen Metacodierer 11, einen Kerncodierer 12 und eine Multiplexiereinheit 81 auf.

[0086] Ferner weist der Metacodierer 11 eine Quantisierungseinheit 21 und eine Codierungseinheit 22 auf und weist der Kerncodierer 12 eine Zeitfrequenzumsetzungseinheit 31, eine MDCT-Koeffizientenkorrektur-einheit 131, eine Quantisierungseinheit 32 und eine Codierungseinheit 33 auf. Ferner weist die Quantisierungseinheit 32 eine Hörpsychologieparameter-Berechnungseinheit 41 und eine Bitzuweisungseinheit 42 auf.

[0087] Die Konfiguration der Codierungsvorrichtung 71, die in **Fig. 6** veranschaulicht ist, unterscheidet sich von der Konfiguration der Codierungsvorrichtung 71 in **Fig. 4** dahingehend, dass die MDCT-Koeffizientenkorrektur-einheit 131 statt der Zeitfrequenzumsetzungseinheit 92 und der Audiosignalkorrektur-einheit 91 vorgesehen ist, und ist sonst gleich der Konfiguration der Codierungsvorrichtung 71 in **Fig. 4**.

[0088] In diesem Beispiel führt erst die Zeitfrequenzumsetzungseinheit 31 eine MDCT an einem Audiosignal eines Objekts durch und wird der resultierende MDCT-Koeffizient zu der MDCT-Koeffizientenkorrek-

tureinheit 131 und der Bitzuweisungseinheit 42 geliefert.

[0089] Dann korrigiert die MDCT-Koeffizientenkorrekturereinheit 131 den MDCT-Koeffizienten, der von der Zeitfrequenzumsetzungseinheit 31 geliefert wird, auf der Grundlage von Verstärkungswerten der gelieferten Metadaten und wird der resultierende korrigierte MDCT-Koeffizient zur Hörpsychologieparameterberechnungseinheit 41 geliefert.

[0090] Zum Beispiel multipliziert die MDCT-Koeffizientenkorrekturereinheit 131 den MDCT-Koeffizienten mit den Verstärkungswerten, um den MDCT-Koeffizienten zu korrigieren. Dadurch wird die Verstärkungskorrektur des Audiosignals in einer Frequenzdomäne durchgeführt.

[0091] Falls die Verstärkungskorrektur in der Frequenzdomäne auf diese Weise durchgeführt wird, ist die Reproduzierbarkeit der Verstärkungskorrektur etwas kleiner als im Falle der ersten Ausführungsform, in dem die Verstärkungskorrektur auf der Grundlage von Verstärkungswerten von Metadaten auf dieselbe Weise wie das tatsächliche Rendern in einer Zeitdomäne durchgeführt wird. Das heißt, der korrigierte MDCT-Koeffizient ist nicht so genau wie in der ersten Ausführungsform.

[0092] Allerdings ist es durch Berechnen der Hörpsychologieparameter durch die Hörpsychologieparameter-Berechnungseinheit 41 auf der Grundlage des korrigierten MDCT-Koeffizienten möglich, Hörpsychologieparameter, die stärker an die tatsächliche Hörempfindung angepasst sind als im Falle von **Fig. 2**, mit im Wesentlichen im Wesentlichen demselben Berechnungsaufwand wie im Fall von **Fig. 2** zu erhalten. Dadurch ist es möglich, die Codierungseffizienz zu verbessern, während die Rechenlast niedrig gehalten wird.

[0093] Es ist festzuhalten, dass, obwohl in **Fig. 6** ein Beispiel beschrieben wurde, in dem Verstärkungswerte von Metadaten vor der Quantisierung für die Korrektur eines MDCT-Koeffizienten verwendet werden, Verstärkungswerte nach dem Codieren oder der Quantisierung verwendet werden können.

[0094] In einem derartigen Fall korrigiert die MDCT-Koeffizientenkorrekturereinheit 131 einen MDCT-Koeffizienten auf der Grundlage von Verstärkungswerten, die als Ergebnis eines Decodierens oder einer inversen Quantisierung, die an Verstärkungswerten nach dem Codieren oder der Quantisierung durchgeführt wird, erhalten werden, um einen korrigierten MDCT-Koeffizienten zu erhalten.

<Beschreibung der Codierungsverarbeitung>

[0095] Als nächstes wird der Betrieb der Codierungsvorrichtung 71, die in **Fig. 6** veranschaulicht ist, beschrieben. Das heißt, die Codierungsverarbeitung, die durch die Codierungsvorrichtung 71 in **Fig. 6** durchgeführt wird, wird unten unter Bezugnahme auf einen Ablaufplan von **Fig. 7** beschrieben.

[0096] Es ist festzuhalten, dass die Prozesse der Schritte S51 und S52 gleich den Prozessen der Schritte S11 und S12 in **Fig. 5** sind und deshalb ihre Beschreibung ausgelassen wird.

[0097] In Schritt S53 führt die Zeitfrequenzumsetzungseinheit 31 eine MDCT am gelieferten Audiosignal eines Objekts durch und liefert den resultierenden MDCT-Koeffizienten zu der MDCT-Koeffizientenkorrekturereinheit 131 und der Bitzuweisungseinheit 42.

[0098] In Schritt S54 korrigiert die MDCT-Koeffizientenkorrekturereinheit 131 den MDCT-Koeffizienten, der von der Zeitfrequenzumsetzungseinheit 31 geliefert wird, auf der Grundlage der Verstärkungswerte der gelieferten Metadaten und liefert den resultierenden korrigierten MDCT-Koeffizienten zur Hörpsychologieparameterberechnungseinheit 41.

[0099] Wenn der korrigierte MDCT-Koeffizient auf diese Weise erhalten wird, werden die Prozesse der Schritte S55 bis S58 danach durchgeführt und wird die Codierungsverarbeitung beendet. Allerdings sind diese Prozesse gleich den Prozessen der Schritte S16 bis S19 in **Fig. 5** und wird deshalb ihre Beschreibung ausgelassen. Allerdings berechnet in Schritt S55 die Hörpsychologieparameter-Berechnungseinheit 41 Hörpsychologieparameter auf der Grundlage des korrigierten MDCT-Koeffizienten, der von der MDCT-Koeffizientenkorrekturereinheit 131 geliefert wird.

[0100] Wie oben beschrieben ist, korrigiert die Codierungsvorrichtung 71 das Audiosignal (den MDCT-Koeffizienten) in einer Frequenzdomäne und berechnet Hörpsychologieparameter auf der Grundlage des erhaltenen korrigierten MDCT-Koeffizienten.

[0101] Auf diese Weise ist es möglich, Hörpsychologieparameter, die stärker an die tatsächliche Hörempfindung angepasst sind, selbst mit einem geringen Berechnungsaufwand zu erhalten und die Codierungseffizienz zu verbessern.

<Dritte Ausführungsform>

<Konfigurationsbeispiel der Codierungsvorrichtung>

[0102] Im Übrigen sind im tatsächlichen 3D - Audioinhalt Verstärkungswerte von Metadaten vor dem Codieren nicht notwendigerweise in einem Spezifikationsumfang von MPEG-H.

[0103] Das heißt, es ist z. B. dann, wenn ein Inhalt erstellt wird, denkbar, dass Verstärkungswerte von Metadaten zu Werten gesetzt werden, die größer als 5,957 ($\approx 15,50$ dB) sind, um die Lautstärke eines Objekts, dessen Wellenformpegel extrem niedrig ist, an die Lautstärke weiterer Objekte anzupassen. Dagegen können die Verstärkungswerte der Metadaten für einen unnötigen Schall Werte sein, die kleiner als 0,004 ($\approx 49,76$ dB) sind.

[0104] Wenn die Verstärkungswerte der Metadaten auf einen oberen Grenzwert oder einen unteren Grenzwert des Wertebereichs, der in **Fig. 3** veranschaulicht ist, beschränkt sind, ist, falls ein derartiger Inhalt in an MPEG-H Format codiert und decodiert wird, ein Schall, der während der Wiedergabe tatsächlich gehört wird, von der Intention eines Inhaltserstellers verschieden.

[0105] Folglich kann, falls die Verstärkungswerte der Metadaten aus dem Umfang der MPEG-H-Spezifikationen fallen, eine Vorverarbeitung zum Ändern der Verstärkungswerte der Metadaten und des Audiosignals, um den MPEG-H-Spezifikationen zu entsprechen, durchgeführt werden, um einen Schall nahe an der Intention eines Inhaltserstellers wiederzugeben.

[0106] In einem derartigen Fall ist die Codierungsvorrichtung 71 z. B. konfiguriert, wie in **Fig. 8** veranschaulicht ist. Es ist festzuhalten, dass in **Fig. 8** Abschnitte, die denen in **Fig. 6** entsprechen, durch dieselben Bezugsnummern und -zeichen bezeichnet werden, und ihre Beschreibung geeignet ausgelassen wird.

[0107] Die Codierungsvorrichtung 71, die in **Fig. 8** veranschaulicht ist, weist eine Änderungseinheit 161, einen Metacodierer 11, einen Kerncodierer 12 und eine Multiplexiereinheit 81 auf.

[0108] Ferner weist der Metacodierer 11 eine Quantisierungseinheit 21 und eine Codierungseinheit 22 auf und weist der Kerncodierer 12 eine Zeitfrequenzumsetzungseinheit 31, eine MDCT-Koeffizientenkorrekturereinheit 131, eine Quantisierungseinheit 32 und eine Codierungseinheit 33 auf. Ferner weist die Quantisierungseinheit 32 eine Hörpsychologieparameter-Berechnungseinheit 41 und eine Bitzuweisungseinheit 42 auf.

[0109] Die Konfiguration der Codierungsvorrichtung 71, die in **Fig. 8** veranschaulicht ist, unterscheidet sich von der Konfiguration der Codierungsvorrichtung 71 in **Fig. 6** dahingehend, dass die Änderungseinheit 161 erneut vorgesehen ist, und ist sonst gleich der Konfiguration der Codierungsvorrichtung 71 in **Fig. 6**.

[0110] In dem Beispiel, das in **Fig. 8** veranschaulicht ist, werden Metadaten und Audiosignale von Objekten, die einen Inhalt bilden, zur Änderungseinheit 161 geliefert.

[0111] Vor dem Codieren prüft (bestätigt) die Änderungseinheit 161, ob ein Verstärkungswert, der aus dem Spezifikationsumfang von MPEG-H, d. h. aus dem Wertebereich, der oben beschrieben ist, fällt, unter den Verstärkungswerten der gelieferten Metadaten vorliegt.

[0112] Dann führt, falls ein Verstärkungswert vorliegt, der aus dem Wertebereich fällt, die Änderungseinheit 161 eine Änderungsverarbeitung eines Verstärkungswerts und eines Audiosignals auf der Grundlage der MPEG-H-Spezifikation als Vorverarbeitung in Bezug auf den Verstärkungswert und das Audiosignal, das dem Verstärkungswert entspricht, durch.

[0113] Speziell ändert die Änderungseinheit 161 den Verstärkungswert, der aus dem Wertebereich (dem Spezifikationsumfang von MPEG-H) fällt, zu dem oberen Grenzwert oder dem unteren Grenzwert des Wertebereichs, um einen geänderten Verstärkungswert zu erhalten.

[0114] Mit anderen Worten wird, falls der Verstärkungswert größer als der obere Grenzwert des Wertebereichs ist, der obere Grenzwert derart eingestellt, dass er ein geänderter Verstärkungswert ist, der ein Verstärkungswert nach einer Änderung ist, und wird, falls der Verstärkungswert kleiner als der untere Grenzwert des Wertebereichs ist, der untere Grenzwert derart eingestellt, dass er ein geänderter Verstärkungswert ist.

[0115] Es ist festzuhalten, dass die Änderungseinheit 161 Parameter außer dem Verstärkungswert unter den mehreren Parametern als Metadaten nicht ändert (abwandelt).

[0116] Ferner führt die Änderungseinheit 161 eine Verstärkungskorrektur am gelieferten Audiosignal des Objekts auf der Grundlage des Verstärkungswerts vor der Änderung und des geänderten Verstärkungswert durch, um ein geändertes Audiosignal zu erhalten. Das heißt, das Audiosignal wird auf der Grundlage einer Differenz zwischen dem Verstärkungswert vor der Änderung und dem geänderten Verstärkungswert geändert (Verstärkungskorrektur).

[0117] Zum jetzigen Zeitpunkt wird eine Verstärkungskorrektur am Audiosignal derart durchgeführt, dass eine Ausgabe des Renderns auf der Grundlage der Metadaten (des Verstärkungswerts) und des Audiosignals vor der Änderung und eine Ausgabe des Renderns auf der Grundlage der Metadaten (des geänderten Verstärkungswerts) und des geänderten Audiosignals nach der Änderung zueinander gleich sind.

[0118] Die Änderungseinheit 161 führt die oben beschriebene Änderung des Verstärkungswerts und des Audiosignals als Vorverarbeitung durch, liefert Daten, die durch einen Verstärkungswert, der nach Bedarf geändert ist, und Parametern außer dem Verstärkungswert der gelieferten Metadaten gebildet sind, als Metadaten nach der Änderung zur Quantisierungseinheit 21 und liefert den Verstärkungswert, der nach Bedarf geändert wurde, zur MDCT-Koeffizientenkorrekturereinheit 131.

[0119] Ferner liefert die Änderungseinheit 161 das Audiosignal, das nach Bedarf geändert wurde, zur Zeitfrequenzumsetzungseinheit 31.

[0120] Es ist festzuhalten, dass im Folgenden, um die Beschreibung zu vereinfachen, Metadaten und eine Verstärkungswert, die von der Änderungseinheit 161 ausgegeben werden, auch als geänderte Metadaten und ein geänderter Verstärkungswert bezeichnet werden, ungeachtet davon, ob eine Änderung durchgeführt wurde oder nicht. Ähnlich wird ein Audiosignal, das von der Änderungseinheit 161 ausgegeben wird, auch als ein geändertes Audiosignal bezeichnet.

[0121] Somit sind in diesem Beispiel geänderte Metadaten eine Eingabe des Metacodierers 11 und sind ein geändertes Audiosignal und ein geänderter Verstärkungswert Eingaben des Kerncodierers 12.

[0122] Auf diese Weise ist ein Verstärkungswert durch die MPEG-H-Spezifikationen nicht im Wesentlichen beschränkt und somit ist es möglich, ein Rendering-Ergebnis gemäß der Intention des Inhaltserstellers zu erhalten.

[0123] Der Metacodierer 11 und der Kerncodierer 12 führen eine Verarbeitung ähnlich dem Beispiel, das in **Fig. 6** veranschaulicht ist, unter Verwendung geänderter Metadaten und eines geänderten Audiosignals als Eingaben durch.

[0124] Das heißt, z. B. führt im Kerncodierer 12 die Zeitfrequenzumsetzungseinheit 31 eine MDCT am geänderten Audiosignal durch und wird der resultierende MDCT-Koeffizient zu der MDCT-Koeffizientenkorrekturereinheit 131 und der Bitzuweisungseinheit 42 geliefert.

[0125] Ferner führt die MDCT-Koeffizientenkorrekturereinheit 131 eine Korrektur an dem MDCT-Koeffizienten, der von der Zeitfrequenzumsetzungseinheit 31 geliefert wird, auf der Grundlage des geänderten Verstärkungswerts, der von der Änderungseinheit 161 geliefert wird, durch und wird der korrigierte MDCT-Koeffizient zur Hörpsychologieparameterberechnungseinheit 41 geliefert.

[0126] Es ist festzuhalten, dass, obwohl hier ein Beispiel beschrieben wurde, in dem ein MDCT-Koeffizient in einer Frequenzdomäne korrigiert wird, eine Verstärkungskorrektur eines geänderten Audiosignals unter Verwendung eines geänderten Verstärkungswerts in einer Zeitdomäne wie in der ersten Ausführungsform durchgeführt wird und dann der korrigierte MDCT-Koeffizient durch MDCT erhalten werden kann.

[0127] Hier wird ein konkretes Beispiel einer Änderung eines Verstärkungswerts eines Audiosignals unter Bezugnahme auf **Fig. 9** bis **Fig. 11** beschrieben.

[0128] **Fig. 9** veranschaulicht Verstärkungswerte für jeden Rahmen von Metadaten eines vorgegebenen Objekts. Es ist festzuhalten, dass in **Fig. 9** die horizontale Achse einen Rahmen angibt und die vertikale Achse einen Verstärkungswert angibt.

[0129] Insbesondere gibt in diesem Beispiel eine Polygonlinie L11 einen Verstärkungswert in jedem Rahmen vor einer Änderung an und gibt eine Polygonlinie L12 einen Verstärkungswert in jedem Rahmen nach einer Änderung, d. h. einen geänderten Verstärkungswert, an.

[0130] Zusätzlich gibt eine Gerade L13 einen Spezifikationsumfang von MPEG-H, d. h. einen unteren Grenzwert (0,004 ($\approx -49,76$ dB)) des oben erwähnten Wertebereichs, an und gibt eine Gerade L14 einen oberen Grenzwert des Spezifikationsumfangs von MPEG-H (5,957 ($\approx 15,50$ dB)) an.

[0131] Hier ist z. B. ein Verstärkungswert vor einer Änderung in einem Rahmen „2“ ein kleinerer Wert als der untere Grenzwert, der durch die Gerade L13 angegeben ist, und wird somit der Verstärkungswert durch den unteren Grenzwert ersetzt, um einen geänderten Verstärkungswert zu erhalten. Zusätzlich ist z. B. ein Verstärkungswert vor einer Änderung in einem Rahmen „4“ ein Größerer Wert als der obere Grenzwert, der durch die Gerade L14 angegeben ist, und somit wird der Verstärkungswert durch den oberen Grenzwert ersetzt, um einen geänderten Verstärkungswert zu erhalten.

[0132] Auf diese Weise wird die Änderung eines Verstärkungswerts geeignet durchgeführt und wird somit ein geänderter Verstärkungswert in jedem

Rahmen zu einem Wert im Spezifikationsumfang (Wertebereich) von MPEG-H gesetzt.

[0133] Zusätzlich veranschaulicht **Fig. 10** ein Audiosignal vor einer Änderung, die durch die Änderungseinheit 161 durchgeführt wird, und veranschaulicht **Fig. 11** ein geändertes Audiosignal, das durch Ändern des Audiosignals, das in **Fig. 10** veranschaulicht ist, erhalten wird. Es ist festzuhalten, dass in **Fig. 10** und **Fig. 11** die horizontale Achse eine Zeit angibt und die vertikale Achse einen Signalpegel angibt.

[0134] Wie in **Fig. 10** veranschaulicht ist, ist der Signalpegel eines Audiosignals vor einer Änderung ungeachtet einer Zeit ein fester Pegel.

[0135] Wenn die Änderungseinheit 161 eine Verstärkungskorrektur auf der Grundlage eines Verstärkungswerts und eines geänderten Verstärkungswerts an einem derartigen Audiosignal durchführt, wird ein geändertes Audiosignal erhalten, das einen Signalpegel aufweist, der jedes Mal variiert, wie in **Fig. 11** veranschaulicht ist, d. h., das einen Signalpegel aufweist, der nicht fest ist.

[0136] Insbesondere kann aus **Fig. 11** verstanden werden, dass der Signalpegel eines geänderten Audiosignals in einem Abtastwert, der durch eine Verringerung des Verstärkungswerts von Metadaten aufgrund der Änderung, d. h. durch Austausch durch einen oberen Grenzwert, beeinflusst wird, stärker erhöht ist als vor einer Änderung.

[0137] Dies ist darauf zurückzuführen, dass es nötig ist, das Audiosignal um einen Betrag zu erhöhen, der der Abnahme des Verstärkungswerts entspricht, um Ausgaben des Renderns vor und nach der Änderung gleich zu gestalten.

[0138] Dagegen ist ersichtlich, dass der Signalpegel eines geänderten Audiosignals in einem Abtastwert, der durch eine Erhöhung des Verstärkungswerts von Metadaten aufgrund der Änderung, d. h. durch Austausch durch einen unteren Grenzwert beeinflusst wird, stärker verringert ist als vor einer Änderung.

<Beschreibung der Codierungsverarbeitung>

[0139] Als nächstes wird der Betrieb der Codierungsvorrichtung 71, die in **Fig. 8** veranschaulicht ist, beschrieben. Das heißt, die Codierungsverarbeitung, die durch die Codierungsvorrichtung 71 in **Fig. 8** durchgeführt wird, wird unten unter Bezugnahme auf einen Ablaufplan von **Fig. 12** beschrieben.

[0140] In Schritt S91 ändert die Änderungseinheit 161 Metadaten, insbesondere einen Verstärkungswert der Metadaten und eines gelieferten Audiosig-

nals eines Objekts nach Bedarf in Übereinstimmung mit dem gelieferten Verstärkungswert der Metadaten des Objekts.

[0141] Das heißt, falls der Verstärkungswert der Metadaten aus dem Spezifikationsumfang von MPEG-H fällt, d. h. ein Wert ist, der aus dem Wertebereich fällt, führt die Änderungseinheit 161 eine Änderung zum Ersetzen des Verstärkungswerts durch den oberen Grenzwert oder den unteren Grenzwert des Wertebereichs durch und ändert das Audiosignal auf der Grundlage der Verstärkungswerte vor und nach der Änderung.

[0142] Die Änderungseinheit 161 liefert die geänderten Metadaten, die durch den geänderten Verstärkungswert, der durch geeignetes Durchführen einer Änderung erhalten wird, und Parameter der Metadaten außer den gelieferten Verstärkungswerten gebildet sind, zur Quantisierungseinheit 21 und liefert die geänderten Verstärkungswerte zur MDCT-Koeffizientenkorrekturereinheit 131.

[0143] Ferner liefert die Änderungseinheit 161 das geänderte Audiosignal, das durch geeignetes Durchführen einer Änderung erhalten wird, zur Zeitfrequenzumsetzungseinheit 31.

[0144] Wenn die geänderten Metadaten und das geänderte Audiosignal auf diese Weise erhalten werden, werden die Prozesse der Schritte S92 bis S99 danach durchgeführt und wird die Codierungsverarbeitung beendet. Allerdings sind diese Prozesse gleich den Prozessen der Schritte S51 bis S58 in **Fig. 7** und wird deshalb ihre Beschreibung ausgelassen.

[0145] Allerdings werden in den Schritten S92 und S93 die geänderten Metadaten quantisiert und codiert und wird in Schritt S94 eine MDCT am geänderten Audiosignal durchgeführt.

[0146] Ferner wird in Schritt S95 der MDCT-Koeffizient auf der Grundlage des MDCT-Koeffizienten, der in Schritt S94 erhalten wird, und der geänderten Verstärkungswerte, die von der Änderungseinheit 161 geliefert werden, korrigiert und wird der resultierende korrigierte MDCT-Koeffizient zur Hörpsychologieparameterberechnungseinheit 41 geliefert.

[0147] Wie oben beschrieben ist, ändert die Codierungsvorrichtung 71 die Eingangsmetadaten und das Audiosignal nach Bedarf und codiert sie dann.

[0148] Auf diese Weise sind die Verstärkungswerte im Wesentlichen nicht durch die Spezifikationen von MPEG-H beschränkt und können Rendering-Ergebnisse erhalten werden, wie durch den Inhaltsersteller vorgesehen ist.

<Vierte Ausführungsform>

<Hinsichtlich der Korrektur des Verstärkungswerts gemäß Hörcharakteristiken>

[0149] Ferner ist es auch möglich, ein Audiosignal zu korrigieren, das zur Berechnung von Hörpsychologieparametern in Übereinstimmung mit Hörcharakteristiken, die mit der Einfallrichtung eines Schalls von einer Schallquelle in Beziehung stehen, verwendet wird.

[0150] Zum Beispiel variiert als Eigenschaft des Hörens die Wahrnehmung der Lautheit eines Schalls abhängig von der Einfallrichtung eines Schalls von einer Schallquelle.

[0151] Das heißt, selbst für dasselbe Objekt variiert eine Hörlautstärke, falls Schallquellen in jeweiligen Richtungen, d. h. auf Seiten vorne, seitlich, oben und unten eines Hörers angeordnet sind. Aus diesem Grund ist es, um die Hörpsychologieparameter, die an die tatsächliche Hörempfindung angepasst sind, zu berechnen, nötig, eine Verstärkungskorrektur auf der Grundlage einer Differenz der Schalldruckempfindlichkeit abhängig von der Einfallrichtung eines Schalls von einer Schallquelle durchzuführen.

[0152] Hier werden die Differenz der Schalldruckempfindlichkeit abhängig von der Einfallrichtung eines Schalls und die Korrektur gemäß der Schalldruckempfindlichkeit beschrieben.

[0153] **Fig. 13** veranschaulicht ein Beispiel des Betrags der Verstärkungskorrektur, wenn eine Verstärkungskorrektur von rosa Rauschen derart durchgeführt wird, dass eine Hörlautstärke während des Wiedergebens desselben rosa Rauschens aus verschiedenen Richtungen auf der Grundlage einer Hörlautstärke, wenn ein bestimmtes rosa Rauschen direkt vor dem Hörer wiedergegeben wird, gleich empfunden wird.

[0154] Es ist festzuhalten, dass in **Fig. 13** die vertikale Achse den Betrag der Verstärkungskorrektur angibt und die horizontale Achse einen Azimut (einen horizontalen Winkel) angibt, der ein Winkel in der horizontalen Richtung ist, der die Position einer Schallquelle vom Hörer gesehen angibt.

[0155] Zum Beispiel ist der Azimut, der vom Hörer gesehen die Richtung der rechten vorderen Seite angibt, 0 Grad, ist der Azimut, der vom gesehen Hörer die rechte seitliche Richtung, d. h. die laterale Seite, angibt, ± 90 Grad und ist der Azimut, der die Rückseite, d. h. die Richtung genau hinter dem Hörer, angibt, 180 Grad. Insbesondere ist die Linksrichtung vom Hörer gesehen die positive Richtung des Azimuts.

[0156] Dieses Beispiel zeigt einen Durchschnittswert des Betrags der Verstärkungskorrektur für jeden Azimut, der aus Ergebnissen von Experimenten erhalten wird, die an mehreren Hörern durchgeführt wurden, und gibt insbesondere gibt ein Bereich, der durch eine gestrichelte Linie in jedem Azimut repräsentiert wird, ein 95 %-Konfidenzintervall an.

[0157] Zum Beispiel lässt sich erkennen, dass, wenn rosa Rauschen auf der lateralen Seite (Azimut = ± 90 Grad) wiedergegeben wird, ein Hörer dieselbe Schalllautstärke empfindet, als wenn das rosa Rauschen, das in Richtung der Frontseite des Hörers wiedergegeben wird, mit einer etwas verringerten Verstärkung gehört wird.

[0158] Zusätzlich lässt sich z. B. dann, wenn rosa Rauschen auf der lateralen Seite (Azimut = 180 Grad) wiedergegeben wird, erkennen, dass der Hörer dieselbe Schalllautstärke empfindet, als wenn das rosa Rauschen, das in Richtung der Frontseite des Hörers wiedergegeben wird, mit einer etwas erhöhten Verstärkung gehört wird.

[0159] Das heißt, dass es für eine bestimmte Zielschallquelle, falls eine Verstärkung eines Schalls der Zielschallquelle etwas verringert wird, wenn die Lokalisierungsposition der Zielschallquelle auf der lateralen Seite des Hörers ist, und falls die Verstärkung des Schalls der Zielschallquelle etwas erhöht wird, wenn die Lokalisierungsposition der Zielschallquelle auf der lateralen Seite des Hörers ist, möglich ist, den Hörer empfinden zu lassen, dass dieselbe Schalllautstärke gehört wird.

[0160] Folglich ist es, wenn der Korrekturbetrag eines Verstärkungswerts für ein Objekt auf der Grundlage von Hörcharakteristiken aus Positionsinformationen des Objekts bestimmt wird, und der Verstärkungswert mit dem bestimmten Korrekturbetrag korrigiert wird, möglich, Hörpsychologieparameter zu erhalten, die die Hörcharakteristiken berücksichtigen.

[0161] In einem derartigen Fall können z. B., wie in **Fig. 14** veranschaulicht ist, eine Verstärkungskorrektureinheit 191 und eine Hörcharakteristiktabelleinheit 192 vorgesehen sein.

[0162] Verstärkungswerte, die in Metadaten eines Objekts enthalten sind, werden zur Verstärkungskorrektureinheit 191 geliefert und der horizontale Winkel (der Azimut), der vertikale Winkel (die Elevation) und die Entfernung (der Radius), die als Positionsinformationen in den Metadaten des Objekts enthalten sind, werden dazu geliefert. Es ist festzuhalten, dass ein Verstärkungswert der Einfachheit der Beschreibung halber hier zu 1,0 angenommen wird.

[0163] Die Verstärkungskorrektureinheit 191 bestimmt einen Verstärkungskorrekturwert, der den Betrag der Verstärkungskorrektur zum Korrigieren eines Verstärkungswerts eines Objekts angibt, auf der Grundlage der Positionsinformationen als die gelieferten Metadaten und eine Hörcharakteristiktabelle, die in der Hörcharakteristiktabelle-Halteinheit 192 gehalten wird.

[0164] Zusätzlich korrigiert die Verstärkungskorrektureinheit 191 den gelieferten Verstärkungswert auf der Grundlage des bestimmten Verstärkungskorrekturwerts und gibt den resultierenden Verstärkungswert als einen korrigierten Verstärkungswert aus.

[0165] Mit anderen Worten bestimmt die Verstärkungskorrektureinheit 191 einen Verstärkungskorrekturwert in Übereinstimmung mit der Richtung eines Objekts von einem Hörer gesehen (der Einfallrichtung eines Schalls), die durch die Positionsinformationen angegeben ist, um dadurch einen korrigierten Verstärkungswert zur Verstärkungskorrektur eines Audiosignals bestimmen, der zur Berechnung von Hörpsychologieparametern verwendet wird.

[0166] Die Hörcharakteristiktabelle-Halteinheit 192 hält Hörcharakteristiktabelle, die Hörcharakteristiken angibt, die mit der Einfallrichtung eines Schalls von einer Schallquelle in Beziehung stehen, und liefert einen Verstärkungskorrekturwert, der durch die Hörcharakteristiktabelle angegeben wird, nach Bedarf zur Verstärkungskorrektureinheit 191.

[0167] Hier ist die Hörcharakteristiktabelle eine Tabelle, in der die Einfallrichtung eines Schalls von einem Objekt, das eine Schallquelle ist, zum Hörer, d. h. die Richtung (die Position) der Schallquelle gesehen vom Hörer, und ein Verstärkungskorrekturwert, der der Richtung entspricht, einander zugeordnet sind. Mit anderen Worten ist die Hörcharakteristiktabelle eine Hörcharakteristik, die den Betrag der Verstärkungskorrektur angibt, die eine akustische Lautstärke in Bezug auf die Einfallrichtung des Schalls von der Schallquelle konstant gestaltet.

[0168] Ein Verstärkungskorrekturwert, der durch die Hörcharakteristiktabelle angegeben ist, wird in Übereinstimmung mit menschlichen Hörcharakteristiken in Bezug auf die Einfallrichtung eines Schalls bestimmt und ist insbesondere die Verstärkungsbetragskorrektur, die eine akustische Lautstärke ungeachtet der Einfallrichtung des Schalls konstant gestaltet. Mit anderen Worten ist der Verstärkungskorrekturwert ein Korrekturwert zum Korrigieren eines Verstärkungswerts auf der Grundlage von Hörcharakteristiken, die mit der Einfallrichtung des Schalls in Beziehung stehen.

[0169] Somit werden, wenn ein Audiosignal eines Objekts einer Verstärkungskorrektur unter Verwendung eines korrigierten Verstärkungswerts unterworfen wird, der durch Korrigieren eines Verstärkungswerts unter Verwendung des Verstärkungskorrekturwerts, der durch die Hörcharakteristiktabelle angegeben ist, erhalten wird, Schalle desselben Objekts ungeachtet der Position des Objekts bei derselben Lautstärke gehört.

[0170] Hier veranschaulicht **Fig. 15** ein Beispiel der Hörcharakteristiktabelle.

[0171] In dem Beispiel, das in **Fig. 15** veranschaulicht ist, ist ein Verstärkungskorrekturwert der Position eines Objekts zugeordnet, die durch den horizontalen Winkel (den Azimut), den vertikalen Winkel (die Elevation) und die Entfernung (den Radius), d. h. die Richtung des Objekts, bestimmt ist.

[0172] Speziell sind in diesem Beispiel alle vertikalen Winkel (Elevationen) und Entfernungen (Radien) 0 und 1,0, ist die Position des Objekts in der vertikalen Richtung dieselbe Höhe wie ein Hörer und wird eine Entfernung vom Hörer zum Objekt jederzeit als konstant angenommen.

[0173] Im Beispiel von **Fig. 15** ist dann, wenn ein Objekt, das eine Schallquelle ist, hinter dem Hörer ist, wie z. B. dann, wenn der horizontale Winkel 180 Grad ist, ein Verstärkungskorrekturwert größer, als wenn das Objekt vor dem Hörer ist, wie z. B. dann, wenn der horizontale Winkel 0 Grad oder 30 Grad ist.

[0174] Ferner wird ein konkretes Beispiel einer Verstärkungswertkorrektur, die durch die Verstärkungskorrektureinheit 191 durchgeführt wird, wenn die Hörcharakteristiktabelle-Halteinheit 192 die Hörcharakteristiktabelle hält, die in **Fig. 15** veranschaulicht ist, beschrieben.

[0175] Zum Beispiel ist, wenn angenommen wird, dass der horizontale Winkel, der vertikale Winkel und die Entfernung, die Parameter von Metadaten des Objekts sind, gleich 90 Grad, 0 Grad bzw. 1,0 m sind, ein Verstärkungskorrekturwert, der der Position des Objekts entspricht, -0,52 dB, wie in **Fig. 15** veranschaulicht ist.

[0176] Somit berechnet die Verstärkungskorrektureinheit 191 die folgende Gleichung (1) auf der Grundlage des Verstärkungskorrekturwerts „-0,52 dB“, der aus der Hörcharakteristiktabelle gelesen wird, und eines Verstärkungswerts „1,0“, um einen korrigierten Verstärkungswert „0,94“ zu erhalten.

[Math. 1]

$$1.0 \times 10^{-0.52/20} \cong 0.94 \quad (1)$$

[0177] Ähnlich ist z. B. dann, wenn angenommen wird, dass der horizontale Winkel, der vertikale Winkel und die Entfernung, die die Position des Objekts angeben, gleich -150 Grad, 0 Grad bzw. 1,0 m sind, ein Verstärkungskorrekturwert, der der Position des Objekts entspricht, 0,51 dB, wie in **Fig. 15** veranschaulicht ist.

[0178] Somit berechnet die Verstärkungskorrektur-einheit 191 die folgende Gleichung (2) auf der Grundlage des Verstärkungskorrekturwerts „0,51 dB“, der aus der Hörcharakteristiktabelle gelesen wird, und eines Verstärkungswerts „1,0“, um einen korrigierten Verstärkungswert „1,06“ zu erhalten.

[Math. 2]

$$1.0 \times 10^{0.51/20} \cong 1.06 \quad (2)$$

[0179] Es ist festzuhalten, dass in **Fig. 15** ein Beispiel beschrieben wurde, in dem ein Verstärkungskorrekturwert verwendet wird, der auf der Grundlage von zweidimensionalen Hörcharakteristiken bestimmt wird, die lediglich die horizontale Richtung berücksichtigen. Das heißt, es wurde ein Beispiel beschrieben, in dem eine Hörcharakteristiktabelle (die im Folgenden auch als eine zweidimensionale Hörcharakteristiktabelle bezeichnet wird), die auf der Grundlage der zweidimensionalen Hörcharakteristiken erzeugt wird, verwendet wird.

[0180] Allerdings kann ein Verstärkungswert unter Verwendung eines Verstärkungskorrekturwerts korrigiert werden, der auf der Grundlage von dreidimensionalen Hörcharakteristiken bestimmt wird, die nicht nur die horizontale Richtung, sondern auch Eigenschaften in der vertikalen Richtung berücksichtigen.

[0181] In einem derartigen Fall kann z. B. eine Hörcharakteristiktabelle, die in **Fig. 16** veranschaulicht ist, verwendet werden.

[0182] In dem Beispiel, das in **Fig. 16** veranschaulicht ist, ist ein Verstärkungskorrekturwert der Position eines Objekts zugeordnet, das durch den horizontalen Winkel (den Azimut), den vertikalen Winkel (die Elevation) und die Entfernung (den Radius), d. h. die Richtung, des Objekts bestimmt ist.

[0183] Speziell ist in diesem Beispiel eine Entfernung für alle Kombinationen von horizontalen Winkeln und vertikalen Winkel gleich 1,0.

[0184] Im Folgenden wird eine Hörcharakteristiktabelle, die auf der Grundlage von dreidimensionalen Hörcharakteristiken in Bezug auf die Einfallrichtung eines Schalls erzeugt wird, wie in **Fig. 16** veranschaulicht ist, auch insbesondere als eine dreidimensionale Hörcharakteristiktabelle bezeichnet werden.

[0185] Hier wird ein konkretes Beispiel einer Korrektur eines Verstärkungswerts durch die Verstärkungskorrektur-einheit 191 beschrieben, falls die Hörcharakteristiktabelle-Halteinheit 192 die Hörcharakteristiktabelle, die in **Fig. 16** veranschaulicht ist, hält.

[0186] Zum Beispiel ist, wenn angenommen wird, dass ein horizontaler Winkel, ein vertikaler Winkel und eine Entfernung, die die Position eines Objekts angeben, gleich 60 Grad, 30 Grad bzw. 1,0 m sind, ein Verstärkungskorrekturwert, der der Position des Objekts entspricht, gleich -0,07 dB, wie in **Fig. 16** veranschaulicht ist.

[0187] Somit berechnet die Verstärkungskorrektur-einheit 191 die folgende Gleichung (3) auf der Grundlage eines Verstärkungskorrekturwerts „-0,07 dB“, der aus der Hörcharakteristiktabelle gelesen wird, und eines Verstärkungswerts „1,0“, um einen korrigierten Verstärkungswert „0,99“ zu erhalten.

[Math. 3]

$$1.0 \times 10^{-0.07/20} \cong 0.99 \quad (3)$$

[0188] Es ist festzuhalten, dass in dem konkreten Beispiel einer Berechnung eines korrigierten Verstärkungswerts, das oben beschrieben ist, der Verstärkungskorrekturwert auf der Grundlage der Hörcharakteristiken, die in Bezug auf die Position (die Richtung) des Objekts bestimmt werden, im Voraus vorbereitet wird. Das heißt, es wurde ein Beispiel beschrieben, in dem ein Verstärkungskorrekturwert, der den Positionsinformationen des Objekts entspricht, in der Hörcharakteristiktabelle gespeichert ist.

[0189] Allerdings ist die Position des Objekts nicht notwendigerweise eine Position, bei der der entsprechende Verstärkungskorrekturwert in der Hörcharakteristiktabelle gespeichert ist.

[0190] Speziell wird z. B. angenommen, dass die Hörcharakteristiktabelle, die in **Fig. 16** gezeigt ist, in der Hörcharakteristiktabelle-Halteinheit 192 gehalten wird und ein horizontaler Winkel, ein vertikaler Winkel und eine Entfernung als Positionsinformationen gleich -120 Grad, 15 Grad bzw. 1,0 m sind.

[0191] In diesem Fall speichert die Hörcharakteristiktabelle von **Fig. 16** keinen Verstärkungskorrekturwert, der einem horizontalen Winkel „-120“, einem vertikalen Winkel „15“ und einer Entfernung „1,0“ entspricht.

[0192] Folglich kann, falls kein Verstärkungskorrekturwert vorliegt, der einer Position, die durch Positionsinformationen in der Hörcharakteristiktabelle angegeben ist, entspricht, die Verstärkungskorrektur-einheit 191 einen Verstärkungskorrekturwert für

eine gewünschte Position durch Interpolationsverarbeitung oder dergleichen unter Verwendung von Verstärkungskorrekturwerten für mehrere Positionen, die entsprechende Verstärkungskorrekturwerte aufweisen, berechnen, wobei die mehreren Positionen zu den Positionen, die durch die Positionsinformationen angegeben sind, benachbart sind. Mit anderen Worten wird eine Interpolationsverarbeitung oder dergleichen auf der Grundlage von Verstärkungskorrekturwerten, die mehreren Positionen in der Nähe der Position, die durch die Positionsinformationen angegeben ist, zugeordnet sind, durchgeführt und wird somit ein Verstärkungskorrekturwert für die Position, die durch die Positionsinformationen angegeben ist, erhalten.

[0193] Zum Beispiel existiert ein Verfahren unter Verwendung von vektorbasiertem Amplitudenschwenken (VBAP) als eines von Verstärkungskorrekturwert-Interpolationsverfahren.

[0194] VBAP (3-point VBAP) ist eine Amplitudenschwenktechnik, die häufig in dreidimensionalem räumlichem Audio-Rendering verwendet wird.

[0195] In VBAP kann die Position eines virtuellen Lautsprechers durch Geben einer gewichteten Verstärkung auf jeden von drei realen Lautsprechern in der Nähe eines beliebigen virtuellen Lautsprechers beliebig geändert werden, um ein Schallquellsignal wiederzugeben.

[0196] Zum jetzigen Zeitpunkt werden eine Verstärkung vg_1 , eine Verstärkung vg_2 und eine Verstärkung vg_3 der realen Lautsprecher derart erhalten, dass die Orientierung eines Zusammensetzungsvektors, der durch Gewichten und Addieren von Vektoren L_1 , L_2 und L_3 in drei Richtungen von einer Hörposition zu den realen Lautsprechern mit den Verstärkungen, die den realen Lautsprechern gegeben werden, erhalten wird, mit der Orientierung (L_p) des virtuellen Lautsprechers übereinstimmt. Speziell werden dann, wenn die Orientierung des virtuellen Lautsprechers, d. h. ein Vektor von der Hörposition zum virtuellen Lautsprecher derart gesetzt wird, dass er ein Vektor L_p ist, die Verstärkungen vg_1 bis vg_3 erhalten, die die folgende Gleichung (4) erfüllen.

[Math. 4]

$$L_p = L_1 * vg_1 + L_2 * vg_2 + L_3 * vg_3 \quad (4)$$

[0197] Hier wird angenommen, dass die Positionen der drei realen Lautsprecher, die oben beschrieben sind, Positionen sind, bei denen drei Verstärkungskorrekturwerte CG_1 , CG_2 und CG_3 vorliegen, die der Hörcharakteristiktabelle entsprechen. Zusätzlich wird angenommen, dass die Position des virtuellen Lautsprechers, die oben beschrieben ist, eine beliebige Position ist, bei der kein Verstärkungskorrektur-

wert vorliegt, der der Hörcharakteristiktabelle entspricht.

[0198] Zum jetzigen Zeitpunkt ist es möglich, einen Verstärkungskorrekturwert CG_p bei der Position des virtuellen Lautsprechers durch Berechnen der folgenden Gleichung (5) zu erhalten.

[Math. 5]

$$R_i = vg_i / \sqrt{(vg_1 * vg_1 + vg_2 * vg_2 + vg_3 * vg_3)} \quad i = 1, 2, 3$$

$$CG_p = R_1 * CG_1 + R_2 * CG_2 + R_3 * CG_3 \quad (5)$$

[0199] In Gleichung (5) werden zunächst die oben beschriebenen gewichteten Verstärkungen vg_1 , vg_2 und vg_3 , die durch VBAP erhalten werden, derart normiert, dass die Summe von Quadraten zu 1 gesetzt ist, wodurch die Verhältnisse R_1 , R_2 und R_3 erhalten werden.

[0200] Dann wird eine Zusammensetzungsverstärkung, die durch Gewichten und Addieren der Verstärkungskorrekturwerte CG_1 , CG_2 und CG_3 für die Position des realen Lautsprechers auf der Grundlage der erhaltenen Verhältnisse R_1 , R_2 und R_3 erhalten wird, als der Verstärkungskorrekturwert CG_p bei der Position des virtuellen Lautsprechers gesetzt.

[0201] Speziell wird eine Masche bei mehreren Positionen abgeteilt, für die Verstärkungskorrekturwerte in einem dreidimensionalen Raum vorbereitet sind. Das heißt, wenn z. B. angenommen wird, dass Verstärkungskorrekturwerte für drei Positionen im dreidimensionalen Raum vorbereitet sind, wird ein Dreiecksbereich mit den drei Positionen als Scheitelpunkte zu einer Masche gesetzt.

[0202] Wenn der dreidimensionale Raum auf diese Weise in mehrere Maschen unterteilt ist, wird eine gewünschte Position zum Erhalten eines Verstärkungskorrekturwerts als eine Zielposition gesetzt und wird eine Masche, die die Zielposition aufweist, festgelegt.

[0203] Zusätzlich wird ein Koeffizient, der mit Positionsvektoren, die drei Scheitelpunktpositionen angeben, die die festgelegte Masche bilden, multipliziert wird, während ein Positionsvektor, der eine Zielposition angibt, durch Multiplikation und Addition der Positionsvektoren, die die drei Scheitelpunktpositionen angeben, repräsentiert wird, durch VBAP erhalten.

[0204] Dann werden die drei Koeffizienten, die auf diese Weise erhalten werden, derart normiert, dass die Summe von Quadraten zu 1 gesetzt ist, wird jeder der normierten Koeffizienten mit jedem der Verstärkungskorrekturwerte für die drei Scheitelpunktpositionen der Masche, die die Zielposition aufweist, multipliziert und wird die Summe der

Verstärkungskorrekturwerte, die mit den Koeffizienten multipliziert werden, als ein Verstärkungskorrekturwert für die Zielposition berechnet. Zusätzlich kann die Normierung durch ein beliebiges Verfahren wie z. B. Gestalten der Summe oder der Summe von Kubikzahlen oder mehr gleich eins durchgeführt werden.

[0205] Es ist festzuhalten, dass das Verstärkungskorrekturwert-Interpolationsverfahren nicht auf die Interpolation unter Verwendung von VBAP beschränkt ist und beliebige weitere Verfahren verwendet werden können.

[0206] Zum Beispiel kann ein Durchschnittswert von Verstärkungskorrekturwerten für mehrere Positionen wie z. B. N Positionen (z. B. $N = 5$) in der Nähe der Zielposition unter den Positionen, bei denen Verstärkungskorrekturwerte in der Hörcharakteristiktabelle vorliegen, als der Verstärkungskorrekturwert für die Zielposition verwendet werden.

[0207] Ferner kann z. B. ein Verstärkungskorrekturwert für eine Position, bei der ein Verstärkungskorrekturwert vorbereitet (gespeichert) ist, die unter den Positionen, bei denen Verstärkungskorrekturwerte in der Hörcharakteristiktabelle vorliegen, am nächsten bei der Zielposition liegt, als der Verstärkungskorrekturwert für die Zielposition verwendet werden.

[0208] Im Übrigen ist in der Hörcharakteristiktabelle, die in **Fig. 16** veranschaulicht ist, ein Verstärkungskorrekturwert für jede Position vorbereitet. Mit anderen Worten sind Verstärkungskorrekturwerte gleichförmig bei allen Frequenzen.

[0209] Allerdings ist auch bekannt, dass sich eine subjektive Differenz der Schalldruckempfindlichkeit, die von einer Richtung abhängt, abhängig von einer Frequenz ändert. Folglich kann ein Verstärkungskorrekturwert für alle mehreren Frequenzen für eine Position vorbereitet werden.

[0210] Hier veranschaulicht **Fig. 17** ein Beispiel einer Hörcharakteristiktabelle, falls Verstärkungskorrekturwerte bei drei Frequenzen für eine Position vorhanden sind.

[0211] In dem Beispiel, das in **Fig. 17** veranschaulicht ist, sind Verstärkungskorrekturwerte bei drei Frequenzen von 250 Hz, 1 kHz und 8 kHz einer Position zugeordnet, die durch den horizontalen Winkel (den Azimut), den vertikalen Winkel (die Elevation) und die Entfernung (den Radius) bestimmt ist. Es ist festzuhalten, dass die Entfernung (der Radius) als fester Wert angenommen wird und der Wert nicht in der Hörcharakteristiktabelle aufgezeichnet ist.

[0212] Zum Beispiel sind bei einer Position, bei der ein horizontaler Winkel -30 Grad ist und ein vertikaler Winkel 0 Grad ist, ein Verstärkungskorrekturwert bei 250 Hz -0,91, ein Verstärkungskorrekturwert bei 1 kHz -1,34 und ein Verstärkungskorrekturwert bei 8 kHz -0,92.

[0213] Es ist festzuhalten, dass eine Hörcharakteristiktabelle, in der Verstärkungskorrekturwerte bei drei Frequenzen von 250 Hz, 1 kHz und 8 kHz für jede Position vorbereitet sind, hier als Beispiel gezeigt ist. Allerdings ist die vorliegende Technologie (nicht darauf beschränkt und kann die Anzahl von Frequenzen, bei denen Verstärkungskorrekturwerte für jede Position und Frequenz, für die Verstärkungskorrekturwerte vorbereitet sind, vorbereitet sind, zu einer beliebigen Anzahl und Frequenzen in der Hörcharakteristiktabelle gesetzt werden.

[0214] Zusätzlich muss ähnlich zum oben beschriebenen Beispiel ein Verstärkungskorrekturwert bei einer gewünschten Frequenz für eine Position eines Objekts nicht in der Hörcharakteristiktabelle gespeichert sein.

[0215] Folglich kann die Verstärkungskorrekturwert-191 eine Interpolationsverarbeitung oder dergleichen auf der Grundlage von Verstärkungskorrekturwerten, die weiteren mehreren Frequenzen in der Nähe der gewünschten Frequenz bei der Position des Objekts oder einer Position in der Nähe der Position in der Hörcharakteristiktabelle zugeordnet sind, durchführen, um einen Verstärkungskorrekturwert bei der gewünschten Frequenz bei der Position des Objekts zu erhalten.

[0216] Zum Beispiel kann, falls ein Verstärkungskorrekturwert bei einer gewünschten Frequenz durch Interpolationsverarbeitung erhalten wird, eine beliebige Interpolationsverarbeitung, z. B. lineare Interpolation wie z. B. Interpolation nullter Ordnung oder Interpolation erster Ordnung, nichtlineare Interpolation wie z. B. Spline-Interpolation oder eine Interpolationsverarbeitung, in der lineare Interpolation und nichtlineare Interpolation beliebig kombiniert werden, durchgeführt werden.

[0217] Ferner kann, falls ein Verstärkungskorrekturwert bei einer minimalen oder einer maximalen Frequenz für eine gewünschte Position nicht vorhanden ist (nicht vorbereitet ist), der Verstärkungskorrekturwert auf der Grundlage von Verstärkungskorrekturwerten bei den Umgebungsfrequenzen bestimmt werden oder kann zu einem festen Wert wie z. B. 0 dB gesetzt werden.

[0218] Hier veranschaulicht **Fig. 18** ein Beispiel, in dem Verstärkungskorrekturwerte bei weiteren Frequenzen durch eine Interpolationsverarbeitung erhalten werden, falls Verstärkungskorrekturwerte

bei Frequenzen von 250 Hz, 1 kHz und 8 kHz für eine vorgegebene Position in der Hörcharakteristiktabelle vorliegen und keine Verstärkungskorrekturwerte bei weiteren Frequenzen vorliegen. Es ist festzuhalten, dass in **Fig. 18** die vertikale Achse einen Verstärkungskorrekturwert angibt und die horizontale Achse eine Frequenz angibt.

[0219] In diesem Beispiel wird eine Interpolationsverarbeitung wie z. B. eine lineare Interpolation oder eine nichtlineare Interpolation auf der Grundlage von Verstärkungskorrekturwerten bei Frequenzen von 250 Hz, 1 kHz und 8 kHz durchgeführt, um Verstärkungskorrekturwerte bei allen Frequenzen zu erhalten.

[0220] Im Übrigen ist bekannt, dass sich eine Kurve gleicher Lautheit abhängig von einem Wiedergabeschalldruck ändert und es besser sein kann, die Hörcharakteristiktabelle gemäß einem Wiedergabeschalldruck eines Audiosignals zu wechseln.

[0221] Folglich hält z. B. die Hörcharakteristiktabelle-Halteinheit 192 eine Hörcharakteristiktabelle für jeden von mehreren Wiedergabeschalldrücken und kann die Verstärkungskorrekturereinheit 191 eine geeignete aus den Hörcharakteristiktabelle auf der Grundlage des Schalldrucks eines Audiosignals eines Objekts wählen. Das heißt, die Verstärkungskorrekturereinheit 191 kann die Hörcharakteristiktabelle, die zur Korrektur eines Verstärkungswerts verwendet werden soll, in Übereinstimmung mit einem Wiedergabeschalldruck wechseln.

[0222] Selbst in diesem Fall können ähnlich zur oben beschriebenen Interpolation von Verstärkungskorrekturwerten für jede Position und Frequenz, wenn keine Hörcharakteristiktabelle eines entsprechenden Schalldrucks in der Hörcharakteristiktabelle-Halteinheit 192 vorliegt, Verstärkungskorrekturwerte der Hörcharakteristiktabelle durch Interpolationsverarbeitung oder dergleichen erhalten werden.

[0223] In einem derartigen Fall führt z. B. die Verstärkungskorrekturereinheit 191 die Interpolationsverarbeitung oder dergleichen auf der Grundlage von Verstärkungskorrekturwerten für eine vorgegebene Position in der Hörcharakteristiktabelle, die mehreren weiteren Wiedergabeschalldrücken in der Nähe des Schalldrucks des Audiosignals des Objekts zugeordnet ist, d. h. in der Nähe des Schalldrucks, um Verstärkungskorrekturwerte für eine vorgegebene Position bei dem Schalldruck des Audiosignals des Objekts zu erhalten, durch. Zum jetzigen Zeitpunkt kann z. B. eine Interpolation durch Addieren von Gewichten gemäß Intervallen zwischen Kurven in einer Kurve gleicher Lautheit durchgeführt werden.

[0224] Ferner kann sich, wenn eine Verstärkungskorrektur eines Audiosignals (eines MDCT-Koeffizienten) eines Objekts gemäß der Position, der Frequenz und dem Wiedergabeschalldruck gleichförmig durchgeführt wird, die Gesamtklangqualität eher verschlechtern.

[0225] Speziell ist z. B. ein Fall denkbar, in dem eine Minute eines Rauschsalls, der ursprünglich für eine Hörempfindung unwichtig ist, als das Audiosignal des Objekts verwendet wird.

[0226] In diesem Fall ist, wenn ein Objekt einer Minute Rauschschall bei einer Position mit einem großen Verstärkungskorrekturwert angeordnet ist, die Bitanzahl, die dem Audiosignal des Objekts zugewiesen ist, in der Bitzuweisungseinheit 42 erhöht. Dann ist die Bitanzahl, die Schallen (Audiosignale) von weiteren wichtigen Objekten zugewiesen ist, entsprechend verringert, was zu einer Möglichkeit führt, dass die Klangqualität verschlechtert ist.

[0227] Somit kann ein Verstärkungskorrekturverfahren gemäß den Eigenschaften des Audiosignals des Objekts geändert werden.

[0228] Zum Beispiel muss die Verstärkungskorrekturereinheit 191 im oben beschriebenen Beispiel, falls bestimmt werden kann, dass die Wahrnehmungsentropie (PE) oder der Schalldruck eines Audiosignals gleich oder kleiner als ein Schwellenwert ist, d. h. das Objekt kein unwichtiges Objekt ist, keine Verstärkungskorrektur durchführen oder kann die Verstärkungsbetragskorrektur begrenzen, d. h. kann einen korrigierten Verstärkungswert derart begrenzen, dass der korrigierte Verstärkungswert gleich oder kleiner als ein oberer Grenzwert ist. Dadurch ist die Korrektur des MDCT-Koeffizienten (des Audiosignals) unter Verwendung des korrigierten Verstärkungswerts in der MDCT-Koeffizientenkorrekturereinheit 131 beschränkt.

[0229] Zusätzlich kann z. B. dann, wenn die Frequenzleistung eines Objektschalls ungleichförmig ist, die Verstärkungskorrekturereinheit 191 die Verstärkungskorrektur in einem Hauptfrequenzband und den weiteren Frequenzbändern gewichten. In einem derartigen Fall wird z. B. ein Verstärkungskorrekturwert gemäß der Frequenzleistung für jedes Frequenzband korrigiert.

[0230] Ferner ist bekannt, dass eine Hörcharakteristiktabelle Schwankungen der Eigenschaften abhängig von einer Person besitzt. Somit ist es auch möglich, einen Codierer, der für einen bestimmten Anwender optimiert ist, unter Verwendung einer Hörcharakteristiktabelle, die für einen bestimmten Anwender optimiert ist, zu konfigurieren.

[0231] In einem derartigen Fall kann z. B. die Hörcharakteristiktabelle-Halteinheit 192 eine Hörcharakteristiktabelle für alle mehreren Anwender halten, wobei die Hörcharakteristiktabelle für jeden Anwender optimiert ist.

[0232] Es ist festzuhalten, dass die Optimierung der Hörcharakteristiktabelle unter Verwendung von Ergebnissen eines Experiments durchgeführt werden kann, das durchgeführt wird, um Hörcharakteristiken lediglich einer bestimmten Person zu prüfen, oder durch ein weiteres Verfahren durchgeführt werden kann.

<Konfigurationsbeispiel der Codierungsvorrichtung>

[0233] Falls ein Verstärkungswert in Übereinstimmung mit Hörcharakteristiken, die oben beschrieben sind, korrigiert wird, ist die Codierungsvorrichtung 71 z. B. konfiguriert, wie in **Fig. 19** veranschaulicht ist. Es ist festzuhalten, dass in **Fig. 19** Abschnitte, die denen in **Fig. 6** oder **Fig. 14** entsprechen, durch dieselben Bezugsnummern und Bezugszeichen bezeichnet sind und ihre Beschreibung geeignet ausgelassen wird.

[0234] Die Codierungsvorrichtung 71, die in **Fig. 19** veranschaulicht ist, weist einen Metacodierer 11, einen Kerncodierer 12 und eine Multiplexiereinheit 81 auf.

[0235] Ferner weist der Metacodierer 11 eine Quantisierungseinheit 21 und eine Codierungseinheit 22 auf und weist der Kerncodierer 12 eine Verstärkungskorrekturereinheit 191, eine Hörcharakteristiktabelle-Halteinheit 192, eine Zeitfrequenzumsetzungseinheit 31, eine MDCT-Koeffizientenkorrekturereinheit 131, eine Quantisierungseinheit 32 und eine Codierungseinheit 33 auf. Ferner weist die Quantisierungseinheit 32 eine Hörpsychologieparameter-Berechnungseinheit 41 und eine Bitzuweisungseinheit 42 auf.

[0236] Die Konfiguration der Codierungsvorrichtung 71, die in **Fig. 19** veranschaulicht ist, unterscheidet sich von der Konfiguration der Codierungsvorrichtung 71 in **Fig. 6** dahingehend, dass die Verstärkungskorrekturereinheit 191 und die Hörcharakteristiktabelle-Halteinheit 192 erneut vorgesehen sind, und ist sonst gleich der Konfiguration der Codierungsvorrichtung 71 in **Fig. 6**.

[0237] Im Beispiel von **Fig. 19** hält die Hörcharakteristiktabelle-Halteinheit 192 z. B. eine dreidimensionale Hörcharakteristiktabelle, die in **Fig. 16** veranschaulicht ist.

[0238] Zusätzlich werden ein Verstärkungswert, ein horizontaler Winkel, ein vertikaler Winkel und eine

Entfernung von Metadaten eines Objekts zur Verstärkungskorrekturereinheit 191 geliefert.

[0239] Die Verstärkungskorrekturereinheit 191 liest Verstärkungskorrekturwerte, die dem horizontalen Winkel, dem vertikalen Winkel und der Entfernung zugeordnet sind, als Positionsinformationen der gelieferten Metadaten aus der dreidimensionalen Hörcharakteristiktabelle, die in der Hörcharakteristiktabelle-Halteinheit 192 gehalten wird.

[0240] Es ist festzuhalten, dass, falls kein Verstärkungskorrekturwert vorliegt, der der Position des Objekts entspricht, das durch die Positionsinformationen der Metadaten angegeben ist, die Verstärkungskorrekturereinheit 191 eine Interpolationsverarbeitung oder dergleichen geeignet durchführt, um einen Verstärkungskorrekturwert zu erhalten, der der Position des Objekts entspricht, das durch die Positionsinformationen angegeben ist.

[0241] Die Verstärkungskorrekturereinheit 191 korrigiert einen Verstärkungswert der gelieferten Metadaten des Objekts unter Verwendung des Verstärkungskorrekturwerts, der auf diese Weise erhalten wird, und liefert den resultierenden korrigierten Verstärkungswert zur MDCT-Koeffizientenkorrekturereinheit 131.

[0242] Somit korrigiert die MDCT-Koeffizientenkorrekturereinheit 131 den MDCT-Koeffizienten, der von der Zeitfrequenzumsetzungseinheit 31 geliefert wird, auf der Grundlage des korrigierten Verstärkungswerts der von der Verstärkungskorrekturereinheit 191 geliefert wird, und liefert den resultierenden korrigierten MDCT-Koeffizienten zur Hörpsychologieparameterberechnungseinheit 41.

[0243] Es ist festzuhalten, dass in dem Beispiel, das in **Fig. 19** veranschaulicht ist, ein Beispiel, in dem Metadaten vor der Quantisierung zur Verstärkungskorrektur eines MDCT-Koeffizienten verwendet werden, beschrieben wurde, jedoch Metadaten nach dem Codieren oder der Quantisierung verwendet werden können.

[0244] In einem derartigen Fall decodiert die Verstärkungskorrekturereinheit 191 die codierten oder die quantisierten Metadaten oder quantisiert sie invers, um einen korrigierten Verstärkungswert auf der Grundlage des resultierenden Verstärkungswerts, des resultierenden horizontalen Winkels, des resultierenden vertikalen Winkels und der resultierenden Entfernung zu erhalten.

[0245] Zusätzlich können die Verstärkungskorrekturereinheit 191 und die Hörcharakteristiktabelle-Halteinheit 192 in den Konfigurationen vorgesehen sein, die in **Fig. 4** und **Fig. 8** veranschaulicht sind.

<Beschreibung der Codierungsverarbeitung>

[0246] Als nächstes wird der Betrieb der Codierungsvorrichtung 71, die in **Fig. 19** veranschaulicht ist, beschrieben. Das heißt, die Codierungsverarbeitung, die durch die Codierungsvorrichtung 71 in **Fig. 19** durchgeführt wird, wird unten unter Bezugnahme auf einen Ablaufplan von **Fig. 20** beschrieben.

[0247] Es ist festzuhalten, dass die Prozesse der Schritte S131 und S 132 gleich den Prozessen der Schritte S51 und S52 in **Fig. 7** sind und somit ihre Beschreibung ausgelassen wird.

[0248] In Schritt S133 berechnet die Verstärkungskorrekturereinheit 191 einen korrigierten Verstärkungswert auf der Grundlage des Verstärkungswerts, des horizontalen Winkels, des vertikalen Winkels und der Entfernung der gelieferten Metadaten und liefert den korrigierten Verstärkungswert zur MDCT-Koeffizientenkorrekturereinheit 131.

[0249] Das heißt, die Verstärkungskorrekturereinheit 191 liest einen Verstärkungskorrekturwert, der dem horizontalen Winkel, dem vertikalen Winkel und der Entfernung der Metadaten aus der dreidimensionalen Hörcharakteristiktabelle, die in der Hörcharakteristiktabelle-Halteinheit 192 gehalten werden, zugeordnet ist, und korrigiert den Verstärkungswert unter Verwendung des Verstärkungskorrekturwerts, um einen korrigierten Verstärkungswert zu berechnen. Zum jetzigen Zeitpunkt wird eine Interpolationsverarbeitung oder dergleichen geeignet durchgeführt und somit wird ein Verstärkungskorrekturwert, der der Position des Objekts, das durch den horizontalen Winkel, den vertikalen Winkel und die Entfernung angegeben ist, entspricht, erhalten.

[0250] Wenn der korrigierte Verstärkungswert auf diese Weise erhalten wird, werden die Prozesse der Schritte S134 bis S 139 danach durchgeführt und wird die Codierungsverarbeitung beendet. Allerdings sind diese Prozesse gleich den Prozessen der Schritte S53 bis S58 in **Fig. 7** und wird deshalb ihre Beschreibung ausgelassen.

[0251] Allerdings wird in Schritt S135 der MDCT-Koeffizient, der durch die Zeitfrequenzumsetzungseinheit 31 erhalten wird, auf der Grundlage des korrigierten Verstärkungswerts, der durch die Verstärkungskorrekturereinheit 191 erhalten wird, korrigiert, um einen korrigierten MDCT-Koeffizienten zu erhalten.

[0252] Es ist festzuhalten, dass eine Hörcharakteristiktabelle für jeden Anwender, die wie oben beschrieben optimiert ist, in der Hörcharakteristiktabelle-Halteinheit 192 gehalten werden kann.

[0253] Ferner kann in der Hörcharakteristiktabelle ein Verstärkungskorrekturwert jeder von mehreren Frequenzen in Bezug auf jede Position zugeordnet sein und kann die Verstärkungskorrekturereinheit 191 einen Verstärkungskorrekturwert für eine gewünschte Frequenz durch Interpolationsverarbeitung auf der Grundlage der Verstärkungskorrekturwerte für mehrere weitere Frequenzen in der Nähe der Frequenz erhalten.

[0254] Zum Beispiel erhält in der Hörcharakteristiktabelle, falls ein Verstärkungskorrekturwert für jede Frequenz jeder Position zugeordnet und gespeichert ist, die Verstärkungskorrekturereinheit 191 einen korrigierten Verstärkungswert für jede Frequenz und korrigiert die MDCT-Koeffizientenkorrekturereinheit 131 einen MDCT-Koeffizienten unter Verwendung des korrigierten Verstärkungswerts für jede Frequenz. Zusätzlich kann eine Hörcharakteristiktabelle für jeden Wiedergabeschalldruck in der Hörcharakteristiktabelle-Halteinheit 192 gehalten werden.

[0255] Wie oben beschrieben ist, korrigiert die Codierungsvorrichtung 71 einen Verstärkungswert von Metadaten unter Verwendung einer dreidimensionalen Hörcharakteristiktabelle und berechnet Hörpsychologieparameter auf der Grundlage eines korrigierten MDCT-Koeffizienten, der unter Verwendung des resultierenden korrigierten Verstärkungswerts erhalten wird.

[0256] Auf diese Weise ist es möglich, Hörpsychologieparameter zu erhalten, die selbst mit einem geringen Berechnungsaufwand an die tatsächliche Hörempfindung angepasst sind, und die Codierungseffizienz zu verbessern. Insbesondere wird ein Verstärkungswert auf der Grundlage von dreidimensionalen Hörcharakteristiken korrigiert und ist es somit möglich, Hörpsychologieparameter zu erhalten, die stärker an die tatsächliche Hörempfindung angepasst sind.

<Fünfte Ausführungsform>

<Konfigurationsbeispiel der Codierungsvorrichtung>

[0257] Im Übrigen weisen dreidimensionale Hörcharakteristiken nicht nur eine Differenz der Schalldruckempfindlichkeit abhängig von der Einfallrichtung eines Schalls von einer Schallquelle auf, sondern auch eine Hörverdeckung von Schall zwischen Objekten, und es ist bekannt, dass der Betrag der Verdeckung zwischen Objekten abhängig von einer Entfernung zwischen den Objekten und Schallfrequenzeigenschaften variiert.

[0258] Allerdings wird in einer allgemeinen Berechnung von Hörpsychologieparametern die Hörverdeckung für jedes Objekt einzeln berechnet und wird

eine Hörverdeckung zwischen Objekten nicht betrachtet.

[0259] Aus diesem Grund können, falls Schalle von mehreren Objekten gleichzeitig wiedergegeben werden, Quantisierungsbits durch Hörverdeckung zwischen Objekten ungeachtet von Quantisierungsrauschen, das tatsächlich ursprünglich nicht wahrnehmbar ist, übermäßig verwendet werden.

[0260] Folglich kann eine Bitzuweisung mit höherer Codierungseffizienz durch Berechnen von Hörpsychologieparametern unter Verwendung eines dreidimensionalen Hörpsychologiemodells, das eine Hörverdeckung zwischen mehreren Objekten gemäß den Positionen und Entfernungen der Objekte berücksichtigt, durchgeführt werden.

[0261] In einem derartigen Fall ist die Codierungsvorrichtung 71 z. B. konfiguriert, wie in **Fig. 21** veranschaulicht ist. In **Fig. 21** werden Abschnitte, die denen in **Fig. 4** entsprechen, durch dieselben Bezugsnummern und -zeichen bezeichnet und ihre Beschreibung wird weggelassen.

[0262] Die Codierungsvorrichtung 71, die in **Fig. 21** veranschaulicht ist, weist einen Metacodierer 11, einen Kerncodierer 12 und eine Multiplexiereinheit 81 auf.

[0263] Zusätzlich weist der Metacodierer 11 eine Quantisierungseinheit 21 und eine Codierungseinheit 22 auf und weist der Kerncodierer 12 eine Zeitfrequenzumsetzungseinheit 31, eine Quantisierungseinheit 32 und eine Codierungseinheit 33 auf. Ferner weist die Quantisierungseinheit 32 eine Hörpsychologiemodell-Halteeinheit 221, eine Hörpsychologieparameter-Berechnungseinheit 222 und eine Bitzuweisungseinheit 42 auf.

[0264] Die Konfiguration der Codierungsvorrichtung 71, die in **Fig. 21** veranschaulicht ist, unterscheidet sich von der Konfiguration der Codierungsvorrichtung 71 in **Fig. 4** dahingehend, dass die Hörpsychologiemodell-Halteeinheit 221 und die Hörpsychologieparameter-Berechnungseinheit 222 statt der Audiosignalkorrekturereinheit 91, der Zeitfrequenzumsetzungseinheit 92 und der Hörpsychologieparameter-Berechnungseinheit 41 vorgesehen sind, und sie sonst gleich der Konfiguration der Codierungsvorrichtung 71 in **Fig. 4** ist.

[0265] In diesem Beispiel hält die Hörpsychologiemodell-Halteeinheit 221 ein dreidimensionales Hörpsychologiemodell, das im Voraus und hinsichtlich einer Hörverdeckung zwischen mehreren Objekten vorbereitet wird. Dieses dreidimensionale Hörpsychologiemodell ist ein Hörpsychologiemodell, das nicht nur eine Hörverdeckung eines einzelnen

Objekts, sondern auch eine Hörverdeckung zwischen mehreren Objekten berücksichtigt.

[0266] Zusätzlich werden ein MDCT-Koeffizient, der durch die Zeitfrequenzumsetzungseinheit 31 erhalten wird, und ein horizontaler Winkel, ein vertikaler Winkel, eine Entfernung und ein Verstärkungswert von Metadaten eines Objekts, zur Hörpsychologieparameterberechnungseinheit 222 geliefert.

[0267] Die Hörpsychologieparameter-Berechnungseinheit 222 berechnet Hörpsychologieparameter auf der Grundlage von dreidimensionalen Hörcharakteristiken. Das heißt, die Hörpsychologieparameter-Berechnungseinheit 222 berechnet die Hörpsychologieparameter auf der Grundlage des MDCT-Koeffizienten, der von der Zeitfrequenzumsetzungseinheit 31 empfangen wurde, des horizontalen Winkels, des vertikalen Winkels, der Entfernung und des Verstärkungswerts der gelieferten Metadaten und des dreidimensionalen Hörpsychologiemodells, das in der Hörpsychologiemodell-Halteeinheit 221 gehalten wird, und liefert die berechneten Hörpsychologieparameter zur Bitzuweisungseinheit 42.

[0268] In der Hörpsychologieparameterberechnung auf der Grundlage von derartigen dreidimensionalen Hörcharakteristiken ist es möglich, Hörpsychologieparameter zu erhalten, die nicht nur eine Hörverdeckung für jedes Objekt, das bisher betrachtet worden ist, sondern auch eine Hörverdeckung zwischen Objekten berücksichtigt.

[0269] Dadurch ist es möglich, eine Bitzuweisung unter Verwendung von Hörpsychologieparametern auf der Grundlage von dreidimensionalen Hörcharakteristiken durchzuführen und die Codierungseffizienz zu verbessern.

<Beschreibung der Codierungsverarbeitung>

[0270] Als nächstes wird der Betrieb der Codierungsvorrichtung 71, die in **Fig. 21** veranschaulicht ist, beschrieben. Das heißt, die Codierungsverarbeitung, die durch die Codierungsvorrichtung 71 in **Fig. 21** durchgeführt wird, wird unten unter Bezugnahme auf einen Ablaufplan von **Fig. 22** beschrieben.

[0271] Es ist festzuhalten, dass die Prozesse der Schritte S171 und S 172 gleich den Prozessen der Schritte S11 und S12 in **Fig. 5** sind und deshalb ihre Beschreibung ausgelassen wird.

[0272] In Schritt S173 führt die Zeitfrequenzumsetzungseinheit 31 eine MDCT (eine Zeit/Frequenzumsetzung) am gelieferten Audiosignal des Objekts durch und liefert den resultierenden MDCT-Koeffizienten zu der Hörpsychologieparameterberechnungseinheit 222 und der Bitzuweisungseinheit 42.

[0273] In Schritt S174 berechnet die Hörpsychologieparameter-Berechnungseinheit 222 Hörpsychologieparameter auf der Grundlage des MDCT-Koeffizienten, der von der Zeitfrequenzumsetzungseinheit 31 empfangen wurde, des horizontalen Winkels, des vertikalen Winkels, der Entfernung und des Verstärkungswerts der gelieferten Metadaten und des dreidimensionalen Hörpsychologiemodells, das in der Hörpsychologiemodell-Halteeinheit 221 gehalten wird, und liefert die berechneten Hörpsychologieparameter zur Bitzuweisungseinheit 42.

[0274] Zum jetzigen Zeitpunkt berechnet die Hörpsychologieparameter-Berechnungseinheit 222 Hörpsychologieparameter unter Verwendung nicht nur des MDCT-Koeffizienten, des horizontalen Winkels, des vertikalen Winkels, der Entfernung und des Verstärkungswerts des Objekts, das verarbeitet werden soll, sondern auch von MDCT-Koeffizienten, horizontalen Winkeln, vertikalen Winkeln, Entfernungen und Verstärkungswerten von weiteren Objekten.

[0275] Als ein konkretes Beispiel wird z. B. ein Fall beschrieben, in dem ein Mithörschwellenwert als ein Hörpsychologieparameter erhalten wird.

[0276] In diesem Fall wird der Mithörschwellenwert auf der Grundlage eines MDCT-Koeffizienten, eines Verstärkungswerts und dergleichen eines Objekts, das verarbeitet werden soll, erhalten. Zusätzlich werden ein Versatzwert (ein Korrekturwert), der einer Entfernung und einer relativen Positionsbeziehung zwischen Objekten entspricht, eine Differenz der Frequenzleistung (des MDCT-Koeffizienten) und dergleichen auf der Grundlage von MDCT-Koeffizienten, Verstärkungswerten und Positionsinformationen eines Objekts, das verarbeitet werden soll, und weiterer Objekte und eines dreidimensionalen Hörpsychologiemodells erhalten. Ferner wird der erhaltene Mithörschwellenwert unter Verwendung des Versatzwerts korrigiert und wird zu einem endgültigen Mithörschwellenwert gesetzt.

[0277] Auf diese Weise ist es möglich, Hörpsychologieparameter zu erhalten, die auch eine Hörverdeckung zwischen Objekten unter Berücksichtigung.

[0278] Wenn die Hörpsychologieparameter berechnet werden, werden die Prozesse der Schritte S175 bis S177 danach durchgeführt und wird die Codierungsverarbeitung beendet. Allerdings sind diese Prozesse gleich den Prozessen der Schritte S17 bis S19 in **Fig. 5** und wird deshalb ihre Beschreibung ausgelassen.

[0279] Wie oben beschrieben ist, berechnet die Codierungsvorrichtung 71 Hörpsychologieparameter auf der Grundlage eines dreidimensionalen Hörpsychologiemodells. Auf diese Weise ist es möglich, eine Bitzuweisung unter Verwendung von Hörpsy-

chologieparametern auf der Grundlage von dreidimensionalen Hörcharakteristiken, die auch eine Hörverdeckung zwischen Objekten berücksichtigen, durchzuführen und die Codierungseffizienz zu verbessern.

<Sechste Ausführungsform>

<Konfigurationsbeispiel der Codierungsvorrichtung>

[0280] Es ist festzuhalten, dass das oben beschriebene Verfahren des Verwendens eines Verstärkungswerts und von Positionsinformationen von Metadaten eines Objekts zur Bitzuweisung z. B. für einen Dienst wirksam ist, in dem ein Anwender ein Rendern unter Verwendung von Metadaten eines Objekts, d. h. Positionen und Verstärkungen, wie sie sind, ohne Änderung während des Betrachtens eines verteilten Inhalts durchführt.

[0281] Andererseits kann ein derartiges Verfahren nicht unverändert verwendet werden, weil die Möglichkeit besteht, dass sich in einem Dienst in dem ein Anwender während des Renderns Metadaten bearbeiten kann, Metadaten zwischen dem Fall des Codierens und dem Fall des Renderns unterscheiden.

[0282] Allerdings erlauben selbst mit einem derartigen Dienst Inhaltsersteller nicht notwendigerweise ein Bearbeiten von Metadaten von allen Objekten und es ist denkbar, dass Inhaltsersteller Objekte, für die Anwender Metadaten bearbeiten dürfen, und Objekte, für die sie es nicht dürfen, festlegen.

[0283] Hier veranschaulicht **Fig. 23** die Syntax des Konfig von Metadaten, zu denen ein Bearbeitungs-genehmigungsmerker „BearbeitungsGenehmigungsMerker“ von Metadaten für jedes Objekt durch einen Inhaltsersteller hinzugefügt wird. Der Bearbeitungs-genehmigungsmerker ist ein Beispiel von Bearbeitungsgenehmigungsinformationen, die angeben, ob ein Bearbeiten von Metadaten erlaubt ist oder nicht.

[0284] In diesem Beispiel weist ein Abschnitt, der durch einen Pfeil Q 11 in Konfig (ObjektMetadaten-Konfig) der Metadaten angegeben ist, einen Bearbeitungs-genehmigungsmerker „BearbeitungsGenehmigungsMerker“ auf.

[0285] Hier gibt „Anz_Objekte“ die Anzahl von Objekten an, die einen Inhalt bilden, und in diesem Beispiel ist ein Bearbeitungsgenehmigungsmerker für jedes Objekt gespeichert.

[0286] Insbesondere gibt ein Wert „1“ eines Bearbeitungs-genehmigungsmerklers an, dass ein Bearbeiten von Metadaten eines Objekts erlaubt ist, und gibt ein Wert „0“ eines Bearbeitungs-genehmigungsmerklers

an, dass ein Bearbeiten von Metadaten eines Objekts nicht erlaubt ist. Der Inhaltsersteller legt den Wert eines Bearbeitungsgenehmigungsmerkers für jedes Objekt fest (stellt ihn ein).

[0287] Wenn ein derartiger Bearbeitungsgenehmigungsmerker in den Metadaten enthalten ist, ist es möglich, Hörpsychologieparameter auf der Grundlage eines dreidimensionalen Hörpsychologiemodells für ein Objekt zu berechnen, für das es nicht erlaubt ist, Metadaten zu bearbeiten.

[0288] In einem derartigen Fall ist die Codierungsvorrichtung 71 z. B. konfiguriert, wie in **Fig. 24** veranschaulicht ist. Es ist festzuhalten, dass in **Fig. 24** Abschnitte, die denen in **Fig. 21** entsprechen, durch dieselben Bezugsnummern und -zeichen bezeichnet sind und ihre Beschreibung geeignet ausgelassen wird.

[0289] Die Codierungsvorrichtung 71, die in **Fig. 24** veranschaulicht ist, weist einen Metacodierer 11, einen Kerncodierer 12 und eine Multiplexiereinheit 81 auf.

[0290] Zusätzlich weist der Metacodierer 11 eine Quantisierungseinheit 21 und eine Codierungseinheit 22 auf und weist der Kerncodierer 12 eine Zeitfrequenzumsetzungseinheit 31, eine Quantisierungseinheit 32 und eine Codierungseinheit 33 auf. Ferner weist die Quantisierungseinheit 32 eine Hörpsychologiemodell-Halteeinheit 221, eine Hörpsychologieparameter-Berechnungseinheit 222 und eine Bitzuweisungseinheit 42 auf.

[0291] Die Codierungsvorrichtung 71, die in **Fig. 24** veranschaulicht ist, ist im Wesentlichen gleich der Codierungsvorrichtung 71, die in **Fig. 21** veranschaulicht ist, jedoch ist die Codierungsvorrichtung 71, die in **Fig. 24** veranschaulicht ist, von der Codierungsvorrichtung 71 in **Fig. 21** dahingehend verschieden, dass ein Bearbeitungsgenehmigungsmerker für jedes Objekt in Metadaten, die eingegeben werden sollen, enthalten ist.

[0292] In diesem Beispiel werden ein horizontaler Winkel, ein vertikaler Winkel, eine Entfernung, ein Verstärkungswert, ein Bearbeitungsgenehmigungsmerker und weitere Parameter in die Quantisierungseinheit 21 als Metadatenparameter eingegeben. Zusätzlich werden der horizontale Winkel, der vertikale Winkel, die Entfernung, der Verstärkungswert und der Bearbeitungsgenehmigungsmerker unter den Metadaten zur Hörpsychologieparameterberechnungseinheit 222 geliefert.

[0293] Somit berechnet die Hörpsychologieparameter-Berechnungseinheit 222 Hörpsychologieparameter genau so wie die Hörpsychologieparameter-Berechnungseinheit 41, die unter Bezugnahme auf

Fig. 4 beschrieben ist, in Übereinstimmung mit dem gelieferten Bearbeitungsgenehmigungsmerker oder berechnet Hörpsychologieparameter auf dieselbe Weise wie im Beispiel von **Fig. 21**.

<Beschreibung der Codierungsverarbeitung>

[0294] Als nächstes wird der Betrieb der Codierungsvorrichtung 71, die in **Fig. 24** veranschaulicht ist, beschrieben. Das heißt, die Codierungsverarbeitung, die durch die Codierungsvorrichtung 71 in **Fig. 24** durchgeführt wird, wird unten unter Bezugnahme auf einen Ablaufplan von **Fig. 25** beschrieben.

[0295] Es ist festzuhalten, dass die Prozesse der Schritte S211 bis S213 gleich den Prozessen der Schritte S171 bis S173 in **Fig. 22** sind und deshalb ihre Beschreibung ausgelassen wird.

[0296] In Schritt S214 berechnet die Hörpsychologieparameter-Berechnungseinheit 222 Hörpsychologieparameter in Übereinstimmung mit dem Bearbeitungsgenehmigungsmerker, der in den gelieferten Metadaten des Objekts enthalten ist, und liefert die berechneten Hörpsychologieparameter zur Bitzuweisungseinheit 42.

[0297] Zum Beispiel berechnet, falls ein Bearbeitungsgenehmigungsmerker eines Objekts, das verarbeitet werden soll, „1“ ist und ein Bearbeiten erlaubt ist, die Hörpsychologieparameter-Berechnungseinheit 222 Hörpsychologieparameter auf der Grundlage eines MDCT-Koeffizienten des Objekts, das verarbeitet werden soll, wobei der MDCT-Koeffizient, der von der Zeitfrequenzumsetzungseinheit 31 geliefert wird.

[0298] Auf diese Weise besteht für ein Objekt, für das Bearbeiten erlaubt ist, die Möglichkeit, dass Metadaten auf einer Decodierungsseite (einer Wiedergabeseite) bearbeitet werden, und deshalb werden Hörpsychologieparameter berechnet, ohne eine Hörverdeckung zwischen Objekten zu berücksichtigen.

[0299] Andererseits berechnet z. B. dann, wenn ein Bearbeitungsgenehmigungsmerker eines Objekts, das verarbeitet werden soll, „0“ ist und ein Bearbeiten nicht erlaubt ist, die Hörpsychologieparameter-Berechnungseinheit 222 Hörpsychologieparameter auf der Grundlage der MDCT-Koeffizienten, die von der Zeitfrequenzumsetzungseinheit 31 empfangen wurden, des horizontalen Winkels, des vertikalen Winkels, der Entfernung und des Verstärkungswerts der gelieferten Metadaten, und des dreidimensionalen Hörpsychologiemodells, das in der Hörpsychologiemodell-Halteeinheit 221 gehalten wird.

[0300] In diesem Fall berechnet die Hörpsychologieparameter-Berechnungseinheit 222 Hörpsychologieparameter genau so wie im Fall eines Schritts S174 in **Fig. 22**. Das heißt, die Hörpsychologieparameter werden nicht nur unter Verwendung des MDCT-Koeffizienten, des horizontalen Winkels, des vertikalen Winkels, der Entfernung und des Verstärkungswerts des Objekts, das verarbeitet werden soll, berechnet, sondern auch von MDCT-Koeffizienten, horizontalen Winkeln, vertikalen Winkeln, Entfernungen und Verstärkungswerten von weiteren Objekten.

[0301] Auf diese Weise werden für ein Objekt, für das ein Bearbeiten nicht erlaubt ist, Hörpsychologieparameter unter Berücksichtigung einer Hörverdeckung zwischen Objekten berechnet, weil Metadaten nicht auf der Decodierungsseite (der Wiedergabe-seite) geändert werden.

[0302] Wenn die Hörpsychologieparameter berechnet werden, werden die Prozesse der Schritte S215 bis S217 danach durchgeführt und wird die Codierungsverarbeitung beendet. Allerdings sind diese Prozesse gleich den Prozessen der Schritte S175 bis S177 in **Fig. 22** und wird deshalb ihre Beschreibung ausgelassen.

[0303] Wie oben beschrieben ist, berechnet die Codierungsvorrichtung 71 Hörpsychologieparameter geeignet unter Verwendung eines dreidimensionalen Hörpsychologiemodells in Übereinstimmung mit einem Bearbeitungsgenehmigungsmerker. Auf diese Weise ist es möglich, für ein Objekt, für das ein Bearbeiten nicht erlaubt ist, eine Bitzuweisung unter Verwendung von Hörpsychologieparametern auf der Grundlage von dreidimensionalen Hörcharakteristiken, die auch eine Hörverdeckung zwischen Objekten berücksichtigen, durchzuführen. Dadurch ist es möglich, die Codierungseffizienz zu verbessern.

[0304] Es ist festzuhalten, dass in Bezug auf die Konfiguration der Codierungsvorrichtung 71, die in **Fig. 21** veranschaulicht ist, ein Beispiel, in dem ein Bearbeitungsgenehmigungsmerker in Kombination verwendet wird, beschrieben wurde. Allerdings ist die vorliegende Erfindung nicht darauf beschränkt und es kann z. B. ein Bearbeitungsgenehmigungsmerker in Kombination in Bezug auf die Konfiguration der Codierungsvorrichtung 71, die in **Fig. 19** veranschaulicht ist, verwendet werden.

[0305] In einem derartigen Fall ist es für ein Objekt, für das ein Bearbeiten nicht erlaubt ist, lediglich erforderlich, dass eine dreidimensionale Hörcharakteristiktabelle verwendet wird, um einen Verstärkungswert von Metadaten des Objekts zu korrigieren.

[0306] Andererseits korrigiert für ein Objekt, für das ein Bearbeiten erlaubt ist, die MDCT-Koeffizienten-

korrekturereinheit 131 einen MDCT-Koeffizienten nicht und die Hörpsychologieparameter-Berechnungseinheit 41 berechnet Hörpsychologieparameter unter Verwendung des unveränderten MDCT-Koeffizienten, der durch die Zeitfrequenzumsetzungseinheit 31 erhalten wird.

[0307] Ferner kann, obwohl hier ein Beispiel beschrieben wurde, in dem Bearbeitungsgenehmigungen für alle Parameter, die Metadaten bilden, durch einen Bearbeitungsgenehmigungsmerker „BearbeitungsGenehmigungsMerker“ gemeinsam gemanagt werden, ein Bearbeitungsgenehmigungsmerker für jeden Parameter von Metadaten vorbereitet werden. Auf diese Weise ist es möglich, ein Bearbeiten einiger oder aller der mehreren Parameter, die in den Metadaten enthalten sind, durch den Bearbeitungsgenehmigungsmerker wahlweise zu erlauben.

[0308] In einem derartigen Fall können z. B. lediglich Parameter der Metadaten, für die ein Bearbeiten durch den Bearbeitungsgenehmigungsmerker nicht erlaubt ist, für die Berechnung der Hörpsychologieparameter verwendet werden.

[0309] Zum Beispiel wird im Beispiel von **Fig. 24**, falls ein Bearbeiten von Positionsinformationen, die einen horizontalen Winkel und dergleichen aufweisen, erlaubt ist, jedoch ein Bearbeiten eines Verstärkungswerts nicht erlaubt ist, der Verstärkungswert ohne Verwendung der Positionsinformationen verwendet und werden Hörpsychologieparameter auf der Grundlage eines dreidimensionalen Hörpsychologiemodells berechnet.

<Siebte Ausführungsform>

<Konfigurationsbeispiel der Codierungsvorrichtung>

[0310] Im Übrigen verwendet eine kanalbasierte Audiocodierung wie z. B. 2-Kanal, 5.1-Kanal und 7.1-Kanal die Annahme als Grundlage, dass Schalle, die durch Mischen von Audiosignalen von verschiedenen musikalischen Mitteln erhalten werden, eingegeben werden.

[0311] Aus diesem Grund ist es auch nötig, einen Bitzuweisungsalgorithmus derart einzustellen, dass für Signale von den verschiedenen Musikinstrumenten im Allgemeinen ein stabiler Betrieb erreicht wird.

[0312] Andererseits werden in objektbasierter 3D-Audiocodierung Audiosignale einzelner Musikinstrumente wie z. B. „Gesang“, „Gitarre“ und „Bass“, die als Objekte dienen, eingegeben. Aus diesem Grund ist es möglich, die Codierungseffizienz zu verbessern und die Geschwindigkeit einer Arithmetikverarbeitung durch Optimieren von Algorithmen wie z. B. Bitzuweisung und Parameter (die im Folgenden auch

als Einstellungsparameter bezeichnet werden) für Signale der musikalischen Mittel zu erhöhen.

[0313] Folglich können z. B. die Typen von Schallquellen von Objekten, d. h. Etiketteninformationen, die Musikinstrumente angeben, wie z. B. „Gesang“ und „Gitarre“ eingegeben werden und können Hörpsychologieparameter unter Verwendung eines Algorithmus oder von Einstellungsparametern, die den Etiketteninformationen entsprechen, berechnet werden. Mit anderen Worten kann eine Bitzuweisung, die Etiketteninformationen entspricht, durchgeführt werden.

[0314] In einem derartigen Fall ist die Codierungsvorrichtung 71 z. B. konfiguriert, wie in **Fig. 26** veranschaulicht ist. Es ist festzuhalten, dass in **Fig. 26** Abschnitte, die denen in **Fig. 6** entsprechen, durch dieselben Bezugsnummern und -zeichen bezeichnet sind und ihre Beschreibung geeignet ausgelassen wird.

[0315] Die Codierungsvorrichtung 71, die in **Fig. 26** veranschaulicht ist, weist einen Metacodierer 11, einen Kerncodierer 12 und eine Multiplexiereinheit 81 auf.

[0316] Ferner weist der Metacodierer 11 eine Quantisierungseinheit 21 und eine Codierungseinheit 22 auf und weist der Kerncodierer 12 eine Parametertabellenhalteeinheit 251, eine Zeitfrequenzumsetzungseinheit 31, eine Quantisierungseinheit 32 und eine Codierungseinheit 33 auf. Ferner weist die Quantisierungseinheit 32 eine Hörpsychologieparameter-Berechnungseinheit 41 und eine Bitzuweisungseinheit 42 auf.

[0317] Die Konfiguration der Codierungsvorrichtung 71, die in **Fig. 26** veranschaulicht ist, unterscheidet sich von der Konfiguration der Codierungsvorrichtung 71 in **Fig. 6** dahingehend, dass eine Parametertabellenhalteeinheit 251 statt der MDCT-Koeffizientenkorrekturereinheit 131 vorgesehen ist, und ist sonst gleich der Konfiguration der Codierungsvorrichtung 71 in **Fig. 6**.

[0318] In diesem Beispiel werden Etiketteninformationen, die die Typen von Schallquellen von Objekten, d. h. die Typen von Musikinstrumenten von Schallen angeben, auf der Grundlage von Audiosignalen von Objekten wie z. B. eines Gesangs, eines Chorus, einer Gitarre, eines Basses, eines Schlagzeugs, einer großen Trommel, einer kleinen Trommel, einer Hi-Hat, eines Pianos, eines Synthesizers und von Streichern in die Codierungsvorrichtung 71 eingegeben (ihr zugeführt).

[0319] Zum Beispiel können die Etiketteninformationen zum Bearbeiten oder dergleichen von Inhalten verwendet werden, die durch Objektsignale von

Objekten gebildet sind, und können die Etiketteninformationen eine Zeichenfolge oder dergleichen sein, die den Typ des Musikinstruments angibt, oder können Kennungsinformationen oder dergleichen sein, die den Typ des Musikinstruments angeben.

[0320] Die Parametertabellenhalteeinheit 251 hält eine Parametertabelle, in der Informationsanzeigialgorithmen und Einstellungsparameter, die zur MDCT-Berechnung, Berechnung von Hörpsychologieparametern und Bitzuweisung verwendet werden, jeweils dem Typ eines Musikinstruments (dem Typ der Schallquelle), das durch die Etiketteninformationen angegeben wird, zugeordnet sind. Es ist festzuhalten, dass in der Parametertabelle Informationsanzeigialgorithmen und/oder Einstellungsparameter dem Typ eines Musikinstruments (dem Typ der Schallquelle) zugeordnet sein können.

[0321] Die Zeitfrequenzumsetzungseinheit 31 führt eine MDCT an einem gelieferten Audiosignal unter Verwendung von Einstellungsparametern und Algorithmen durch, die für den Typ des Musikinstruments, der durch gelieferte Etiketteninformationen angegeben ist, unter Bezugnahme auf die Parametertabelle, die in der Parametertabellenhalteeinheit 251 gehalten wird, bestimmt werden.

[0322] Die Zeitfrequenzumsetzungseinheit 31 liefert einen MDCT-Koeffizienten, der durch die MDCT erhalten wird, zu der Hörpsychologieparameterberechnungseinheit 41 und der Bitzuweisungseinheit 42.

[0323] Zusätzlich quantisiert die Quantisierungseinheit 32 den MDCT-Koeffizienten auf der Grundlage der Einstellungsparameter und der Algorithmen, die für den Typ des Musikinstruments, der durch die Etiketteninformationen angegeben ist, auf der Grundlage der gelieferten Etiketteninformationen und MDCT-Koeffizienten bestimmt werden.

[0324] Das heißt, die Hörpsychologieparameter-Berechnungseinheit 41 berechnet Hörpsychologieparameter auf der Grundlage des MDCT-Koeffizienten, der von der Zeitfrequenzumsetzungseinheit 31 unter Verwendung der Einstellungsparameter und Algorithmen empfangen wurde, die für den Typ des Musikinstruments, der durch die gelieferten Etiketteninformationen angegeben ist, unter Bezugnahme auf die Parametertabelle, die in der Parametertabellenhalteeinheit 251 gehalten wird, bestimmt werden, und liefert die berechneten Hörpsychologieparameter zur Bitzuweisungseinheit 42.

[0325] Die Bitzuweisungseinheit 42 führt eine Bitzuweisung und eine Quantisierung des MDCT-Koeffizienten auf der Grundlage des MDCT-Koeffizienten, der von der Zeitfrequenzumsetzungseinheit 31 empfangen wurde, der Hörpsychologieparameter, die

von der Hörpsychologieparameterberechnungseinheit 41 empfangen wurden, und der gelieferten Etiketteninformationen unter Bezugnahme auf die Parametertabelle, die in der Parametertabellenhalteeinheit 251 gehalten wird, durch.

[0326] Zum jetzigen Zeitpunkt führt die Bitzuweisungseinheit 42 eine Bitzuweisung unter Verwendung des MDCT-Koeffizienten, von Hörpsychologieparametern und der Einstellungsparameter und Algorithmen, die für den Typ des Musikinstruments, der durch die Etiketteninformationen angegeben ist, bestimmt werden, durch.

[0327] Es ist festzuhalten, dass verschiedene Verfahren zum Optimieren von Algorithmen und Einstellungsparametern für jeden Typ eines Musikinstruments (Typ einer Schallquelle), der durch Etiketteninformationen angegeben ist, vorliegen und konkrete Beispiele unten beschrieben werden.

[0328] Zum Beispiel ist es in der MDCT (der Zeit-/Frequenz-Umsetzung) möglich, Fenster, die zur MDCT verwendet werden (Transformationsfenster), d. h. Fensterfunktionen, zu wechseln.

[0329] Folglich kann z. B. ein Fenster mit einer hohen Zeitauflösung wie z. B. das Kaiser-Fenster für Musikinstrumentenobjekte wie z. B. die Typen eines Musikinstruments einer Hi-Hat und einer Gitarre verwendet werden, bei denen ein Anstieg und ein Abfall von Schallen wichtig sind, und kann ein Sinusfenster für Musikinstrumentenobjekte wie z. B. Gesang und Bass verwendet werden, bei denen eine Voluminösität wichtig ist.

[0330] Auf diese Weise kann, wenn der Typ des Musikinstruments, das durch die Etiketteninformationen angegeben ist, und Informationen, die eine Fensterfunktion angeben, die für den Typ des Musikinstruments bestimmt wurde, in der Parametertabelle in Verbindung miteinander gespeichert sind, eine MDCT unter Verwendung eines Fensters, das den Etiketteninformationen entspricht, durchgeführt werden.

[0331] Ferner kann auch in der Berechnung von Hörpsychologieparametern und einer Bitzuweisung z. B. eine Bandbegrenzung gemäß Etiketteninformationen durchgeführt werden.

[0332] Das heißt, Musikinstrumente in einem unteren Register wie z. B. ein Bass und eine große Trommel, Musikinstrumente in einem mittleren Register wie z. B. ein Gesang, Musikinstrumente in einem hohen Register wie z. B. eine Hi-Hat, und Musikinstrumenten in einem gesamten Register wie z. B. ein Piano unterscheiden sich in wichtigen und unnötigen Bändern in einer Hörempfindung. Folglich ist es möglich, Quantisierungsbits von jedem von unnötigen

Bändern unter Verwendung der Etiketteninformationen zu verringern und einem wichtigen Band viele Quantisierungsbits zuzuweisen.

[0333] Speziell weist ein Objektsignal eines Musikinstruments in einem unteren Register wie z. B. ein Bass oder eine große Trommel ursprünglich nahezu keine Komponenten eines hohen Bereichs auf. Allerdings werden, wenn ein Objektsignal eines derartigen Musikinstruments viel Rauschen eines hohen Bereichs aufweist, in einer Bitzuweisung auch viele Quantisierungsbits eines Skalierungsfaktorbands eines hohen Bereichs zugewiesen.

[0334] Folglich werden für den Typ des Musikinstruments in einem unteren Register wie z. B. einen Bass oder eine große Trommel Einstellungsparameter und Algorithmen für die Berechnung von Hörpsychologieparametern und eine Bitzuweisung derart bestimmt, dass viele Quantisierungsbits aufgrund eines niedrigen Bereichs zugewiesen sind und weniger Quantisierungsbits einem hohen Bereich zugewiesen sind.

[0335] Auf diese Weise ist es möglich, ein Rauschen durch Verringern der Anzahl von Quantisierungsbits eines hohen Bereichs, der keine Zielsignalkomponenten aufweist, zu verringern, die Anzahl von Quantisierungsbits eines niedrigen Bereichs, der Zielsignalkomponenten aufweist, zu erhöhen und eine Klangqualität und eine Codierungseffizienz zu verbessern.

[0336] Andererseits können auch in Hörpsychologieparametern wie z. B. einem Mithörschwellenwert Schallen, die durch eine Hörempfindung einfach wahrgenommen werden, für jedes Musikinstrument viele Quantisierungsbits zugewiesen sein, indem eine Einstellung (Einstellungsparameter) in Übereinstimmung mit dem Typ des Musikinstruments wie z. B. ein Musikinstrument, das eine starke Tonalität aufweist, ein Musikinstrument, das eine hohe Rauscheneigenschaft aufweist, ein Musikinstrument, das eine starke zeitliche Schwankung eines Signals aufweist, und ein Musikinstrument, das eine geringe zeitliche Schwankung eines Signals aufweist, geändert wird.

[0337] Ferner werden z. B. in Codierern wie z. B. fortschrittlicher Audiocodierung (AAC) und USAC Frequenzspektruminformationen (ein MDCT-Koeffizient) für jedes Skalierungsfaktorband quantisiert.

[0338] Der quantisierte Wert jedes Skalierungsfaktorbands, d. h. die Bitanzahl, die für jedes Skalierungsfaktorband zugewiesen werden soll, startet mit einem vorgegebenen Wert als ein Anfangswert und ein endgültiger Wert wird durch Durchführen einer Bitzuweisungsschleife bestimmt.

[0339] Zum Beispiel wird in der Bitzuweisungsschleife eine Quantisierung eines MDCT-Koeffizien-

ten wiederholt durchgeführt, während der quantisierte Wert jedes Skalierungsfaktorbands geändert wird, d. h. während des Durchführens einer Bitzuweisung, bis vorgegebene Bedingungen erfüllt sind. Die vorgegebenen Bedingungen, die hier erwähnt werden, sind z. B. eine Bedingung, dass die Summe der Bitanzahl des quantisierten MDCT-Koeffizienten jedes Skalierungsfaktorbands gleich oder kleiner als eine vorgegebene zulässige Bitanzahl ist, und ein Bedingung, dass das Quantisierungsrauschen ausreichend klein ist.

[0340] In vielen Fällen ist es wünschenswert, die Zeit zu verkürzen, die zum Codieren (zur Quantisierung) wie z. B. unter Verwendung eines Echtzeitcodierers erforderlich ist, und derartige Fälle werden durch eine geringe Schwächung einer Klangqualität begleitet. Allerdings ist auch eine Obergrenze für die Anzahl von Bitzuweisungsschleifen (die Anzahl von Schleifen), die oben beschrieben sind, eingestellt.

[0341] Natürlich ist die Anzahl von Bitzuweisungsschleifen umso geringer und ist die Codierungszeit umso kürzer, je näher der Anfangswert des quantisierten Werts jedes Skalierungsfaktorbands am endgültigen Wert liegt. Zusätzlich ist die Verschlechterung der Klangqualität aufgrund der Begrenzung der Anzahl von Schleifen auch verringert.

[0342] Somit ist es möglich, ein Audiosignal mit hoher Klangqualität in einem kurzen Zeitraum durch Erhalten eines optimalen Anfangswerts im Voraus für jeden Typ eines Musikinstruments, das durch die Etiketteninformationen angegeben ist, und Schalten des Anfangswerts in Übereinstimmung mit den Etiketteninformationen zu codieren (zu quantisieren). In diesem Fall können z. B. die Etiketteninformationen als einer von Hörpsychologieparametern gesetzt werden oder kann ein Anfangswert eines quantisierten Werts als ein Einstellungsparameter für jeden Typ eines Musikinstruments in einer Parametertabelle bestimmt sein.

[0343] Die oben beschriebenen Einstellungsparameter und Algorithmen für jeden Typ eines Musikinstruments können im Voraus durch manuelle Einstellung auf der Grundlage einer Erfahrung, statistische Einstellung, maschinelles Lernen oder dergleichen erhalten werden.

[0344] In der Codierungsvorrichtung 71 die die Konfiguration, die in **Fig. 26** veranschaulicht ist, aufweist, werden Einstellungsparameter und Algorithmen für jeden Typ eines Musikinstruments im Voraus als eine Parametertabelle vorbereitet. Zusätzlich werden eine Berechnung von Hörpsychologieparametern, eine Bitzuweisung, d. h. eine Quantisierung, und eine MDCT gemäß Einstellungsparametern und Algorithmen durchgeführt, die den Etiketteninformationen entsprechen.

[0345] Es ist festzuhalten, dass, obwohl die Etiketteninformationen in diesem Beispiel allein verwendet werden, sie in Kombination mit weiteren Metadateninformationen verwendet werden können.

[0346] Zum Beispiel können weitere Parameter von Metadaten eines Objekts Prioritätsinformationen aufweisen, die die Priorität des Objekts angeben.

[0347] Folglich können in der Zeitfrequenzumsetzungseinheit 31, der Hörpsychologieparameter-Berechnungseinheit 41 und der Bitzuweisungseinheit 42 die Die Stärke und die Schwäche von Einstellungsparametern, die für die Etiketteninformationen bestimmt werden, ferner unter Verwendung des Werts der Priorität, die durch die Prioritätsinformationen des Objekts angegeben ist, durchgeführt werden. Dagegen können Objekte mit derselben Priorität mit verschiedenen Prioritäten unter Verwendung der Etiketteninformationen verarbeitet werden.

[0348] Zusätzlich ist es auch möglich, obwohl hier eine Beschreibung durch Beschränken der Etiketteninformationen auf den Typ des Musikinstruments gegeben wurde, Etiketteninformationen zum Bestimmen einer Hörumgebung außer dem Typ des Musikinstruments zu verwenden.

[0349] Zum Beispiel wird, falls ein Schall wie z. B. ein Inhalt in einem Fahrzeug gehört wird, Quantisierungsrauschen in einem unteren Register aufgrund eines Kraftmaschinenschalls und eines Fahrgeräuschs weniger wahrscheinlich wahrgenommen. Zusätzlich unterscheidet sich ein minimal hörbarer Bereich, d. h. eine wahrnehmbare Lautstärke zwischen einem ruhigen Raum und eine überfüllten Außenbereich. Ferner ändert sich die Hörumgebung selbst auch im Zeitablauf und mit der Bewegung eines Anwenders.

[0350] Folglich können z. B. Etiketteninformationen, die Hörumgebungsinformationen aufweisen, die die Hörumgebung des Anwenders angeben, in die Codierungsvorrichtung 71 eingegeben werden und können Hörpsychologieparameter, die für die Hörumgebung des Anwenders optimal sind, unter Verwendung von Einstellungsparametern und Algorithmen, die den Etiketteninformationen entsprechen, berechnet werden und dergleichen.

[0351] In diesem Fall werden die MDCT, die Berechnung von Hörpsychologieparametern und die Bitzuweisung unter Verwendung von Einstellungsparametern und Algorithmen durchgeführt, die für die Hörumgebung und den Typ des Musikinstruments, die durch die Etiketteninformationen angegeben werden, z. B. unter Bezugnahme auf eine Parametertabelle bestimmt werden.

[0352] Auf diese Weise ist es möglich, eine Quantisierung (eine Codierung) mit höherer Klangqualität für verschiedene Hörumgebungen durchzuführen. Zum Beispiel werden in einem Fahrzeug einem mittleren/hohen Bereich durch Erhöhen eines Mithörschwellenwerts eines Quantisierungsrauschens in einem unteren Register, das weniger wahrscheinlich wahrgenommen wird, während des Quantisierens eines MDCT-Koeffizienten viele Bits zugewiesen, und somit ist es möglich, die Klangqualität eines Objekts, das vom Typ eines Musikinstruments wie z. B. von Gesang ist, zu verbessern.

<Beschreibung der Codierungsverarbeitung>

[0353] Als nächstes wird der Betrieb der Codierungsvorrichtung 71, die in **Fig. 26** veranschaulicht ist, beschrieben. Das heißt, die Codierungsverarbeitung, die durch die Codierungsvorrichtung 71 in **Fig. 26** durchgeführt wird, wird unten unter Bezugnahme auf einen Ablaufplan von **Fig. 27** beschrieben.

[0354] Es ist festzuhalten, dass die Prozesse der Schritte S251 und S252 gleich den Prozessen der Schritte S51 und S52 in **Fig. 7** sind und deshalb ihre Beschreibung ausgelassen wird.

[0355] In Schritt S253 führt die Zeitfrequenzumsetzungseinheit 31 eine MDCT am gelieferten Audiosignal auf der Grundlage der Parametertabelle, die in der Parametertabellenhalteeinheit 251 gehalten wird, und der gelieferten Etiketteninformationen durch und liefert den resultierenden MDCT-Koeffizienten zu der Hörpsychologieparameterberechnungseinheit 41 und der Bitzuweisungseinheit 42.

[0356] Zum Beispiel wird in Schritt S253 eine MDCT am Audiosignal des Objekts unter Verwendung von Einstellungsparametern und Algorithmen, die für die Etiketteninformationen des Objekts bestimmt werden, durchgeführt.

[0357] In Schritt S254 berechnet die Hörpsychologieparameter-Berechnungseinheit 41 Hörpsychologieparameter auf der Grundlage des MDCT-Koeffizienten, der von der Zeitfrequenzumsetzungseinheit 31 geliefert wird, unter Bezugnahme auf die Parametertabelle, die in der Parametertabellenhalteeinheit 251 gehalten wird, gemäß den gelieferten Etiketteninformationen und liefert die berechneten Hörpsychologieparameter zur Bitzuweisungseinheit 42.

[0358] Zum Beispiel werden in Schritt S254 die Hörpsychologieparameter für das Objekt unter Verwendung der Einstellungsparameter und Algorithmen berechnet, die für die Etiketteninformationen des Objekts bestimmt werden.

[0359] In Schritt S255 führt die Bitzuweisungseinheit 42 eine Bitzuweisung auf der Grundlage des MDCT-Koeffizienten, der von der Zeitfrequenzumsetzungseinheit 31 geliefert wird, und der Hörpsychologieparameter, die von der Hörpsychologieparameterberechnungseinheit 41 geliefert werden, unter Bezugnahme auf die Parametertabelle, die in der Parametertabellenhalteeinheit 251 gehalten wird, gemäß den gelieferten Etiketteninformationen durch und quantisiert den MDCT-Koeffizienten.

[0360] Wenn der MDCT-Koeffizient auf diese Weise quantisiert wird, werden die Prozesse der Schritte S256 und S257 danach durchgeführt und wird die Codierungsverarbeitung beendet. Allerdings sind diese Prozesse gleich den Prozessen der Schritte S57 und S58 in **Fig. 7** und deshalb wird ihre Beschreibung ausgelassen.

[0361] Wie oben beschrieben ist, führt die Codierungsvorrichtung 71 eine MDCT, eine Berechnung von Hörpsychologieparametern und eine Bitzuweisung in Übereinstimmung mit den Etiketteninformationen durch. Auf diese Weise ist es möglich, die Codierungseffizienz und die Verarbeitungsgeschwindigkeit einer Quantisierungsberechnung zu verbessern und eine Audiowiedergabe mit höherer Klangqualität zu realisieren.

<Achte Ausführungsform>

<Konfigurationsbeispiel der Codierungsvorrichtung>

[0362] Zusätzlich ist die Codierungsvorrichtung 71, die eine Quantisierung (ein Codieren) unter Verwendung von Etiketteninformationen durchführt, auch anwendbar, falls Positionsinformationen eines Anwenders und Positionsinformationen eines Objekts in Kombination verwendet werden, wie z. B. bei MPEG-I, freier Standpunkt.

[0363] In einem derartigen Fall ist die Codierungsvorrichtung 71 z. B. konfiguriert, wie in **Fig. 28** veranschaulicht ist. In **Fig. 28** sind Abschnitte, die denen in **Fig. 26** entsprechen, durch dieselben Bezugsnummern und -zeichen bezeichnet und wird ihre Beschreibung geeignet ausgelassen.

[0364] Die Codierungsvorrichtung 71, die in **Fig. 28** veranschaulicht ist, weist einen Metacodierer 11, einen Kerncodierer 12 und eine Multiplexiereinheit 81 auf.

[0365] Obwohl es in der Zeichnung nicht veranschaulicht ist, weist der Metacodierer 11 eine Quantisierungseinheit 21 und eine Codierungseinheit 22 auf.

[0366] Ferner weist der Kerncodierer 12 eine Parametertabellenhalteeinheit 251, eine Zeitfrequenzum-

setzungseinheit 31, eine Quantisierungseinheit 32 und eine Codierungseinheit 33 auf und weist die Quantisierungseinheit 32 eine Hörpsychologieparameter-Berechnungseinheit 41 und eine Bitzuweisungseinheit 42 auf.

[0367] Die Konfiguration der Codierungsvorrichtung 71, die in **Fig. 28** veranschaulicht ist, ist im Wesentlichen dieselbe wie die der Codierungsvorrichtung 71, die in **Fig. 26** veranschaulicht ist, unterscheidet sich jedoch von der Konfiguration der Codierungsvorrichtung 71, die in **Fig. 26** veranschaulicht ist, dahingehend, dass die Position eines Anwenders, d. h. Anwenderpositionsinformationen, die eine Hörposition eines Schalls wie z. B. eines Inhalts angeben, ferner durch den Anwender in die Codierungsvorrichtung 71 eingegeben werden, die in **Fig. 28** veranschaulicht ist.

[0368] Der Metacodierer 11 codiert Metadaten, die Parameter wie z. B. Positionsinformationen eines Objekts und Verstärkungswerte aufweisen, jedoch sind die Positionsinformationen des Objekts, die in den Metadaten enthalten sind, von denen im Beispiel, das in **Fig. 26** veranschaulicht ist, verschieden.

[0369] Zum Beispiel werden in diesem Beispiel Positionsinformationen, die die relative Position des Objekts vom Anwender (der Hörposition) gesehen angeben, Positionsinformationen, die die absolute Position des Objekts, die geeignet geändert wurde, angeben, und dergleichen als Positionsinformationen, die Metadaten des Objekts bilden, auf der Grundlage der Anwenderpositionsinformationen und des gelieferten horizontalen Winkels, des gelieferten vertikalen Winkels und der gelieferten Entfernung des Objekts codiert.

[0370] Es ist festzuhalten, dass z. B. die Anwenderpositionsinformationen von einer Client-Vorrichtung (die nicht dargestellt ist) geliefert werden, die ein Verteilungsziel (ein Übertragungsziel) eines Bitstroms ist, der einen Inhalt, der durch die Codierungsvorrichtung 71 erzeugt wird, d. h. codierte Metadaten und codierte Audiodaten enthält.

[0371] Ferner berechnet die Hörpsychologieparameter-Berechnungseinheit 41 Hörpsychologieparameter unter Verwendung nicht nur der Etiketteninformationen, sondern auch der gelieferten Positionsinformationen des Objekts, d. h. des horizontalen Winkels, des vertikalen Winkels und der Entfernung, die die Position des Objekts angeben, und der Anwenderpositionsinformationen.

[0372] Zusätzlich können die Anwenderpositionsinformationen und die Objektpositionsinformationen auch zur Bitzuweisungseinheit 42 geliefert werden und können die Anwenderpositionsinformationen

und die Objektpositionsinformationen zur Bitzuweisung verwendet werden.

[0373] Hier wird ein Beispiel einer Berechnung von Hörpsychologieparametern, die durch die Hörpsychologieparameter-Berechnungseinheit 41 durchgeführt wird, und einer Bitzuweisung, die durch die Bitzuweisungseinheit 42 durchgeführt wird, beschrieben. Insbesondere wird hier ein Beispiel beschrieben, in dem ein Inhalt eine Live-Musik-Inhalt ist.

[0374] In diesem Fall hört ein Anwender einem Schall eines Inhalts in einer virtuellen Live-Halle zu, jedoch unterscheiden sich Schalle, die in einer vorderen Reihe und einer letzten Reihe der Live-Halle gehört werden, wesentlich.

[0375] Folglich werden z. B., falls der Anwender dem Schall des Inhalts bei einer Position in der Nähe eines Objekts in der vorderen Reihe an einem freien Standpunkt zuhört, bevorzugt Quantisierungsbits einem Objekt zugewiesen, das bei einer Position in der Nähe des Anwenders angeordnet ist, statt gleichförmig zugewiesen zu werden, selbst dann, wenn dieselben Etiketteninformationen mehreren Objekten zugewiesen werden. Auf diese Weise ist es möglich, dem Anwender eine Empfindung einer Realität, als ob der Anwender näher bei dem Objekt ist, d. h. ein stärkeres Präsenzgefühl zu geben.

[0376] Dagegen kann, falls der Anwender dem Schall des Inhalts bei einer Position entfernt vom Objekt in der letzten Reihe zuhört, die ursprüngliche Einstellung für jeden Typ eines Musikinstruments, d. h. die Einstellung für eine größere Entfernung an den Einstellungsparametern und Algorithmen durchgeführt werden, die den Etiketteninformationen entsprechen.

[0377] Zum Beispiel werden selbst mit dem Schall eines Musikinstruments, wobei es besser ist, mehrere Bits einem Anstiegsschall und einem Verbindungsschall zuzuweisen, viele Bits dem Abklingen eines Signals, Echos und einem Nachhallabschnitt zugewiesen und ist es somit möglich, eine Wahrnehmung eines Raums zu verbessern und dem Anwender ein Präsenzgefühl zu vermitteln, als ob der Anwender in einer großen Halle ist.

[0378] Auf diese Weise ist es möglich, ein Präsenzgefühl durch Durchführen einer Berechnung von Hörpsychologieparametern und eine Bitzuweisung in Übereinstimmung nicht nur mit den Etiketteninformationen, sondern auch der Position des Anwenders in einem dreidimensionalen Raum, d. h. einer Hörposition, die durch die Anwenderpositionsinformationen angegeben ist, und einer Entfernung zwischen dem Anwender und dem Objekt weiter zu verbessern.

<Beschreibung der Codierungsverarbeitung>

[0379] Als nächstes wird der Betrieb der Codierungsvorrichtung 71, die in **Fig. 28** veranschaulicht ist, beschrieben. Das heißt, die Codierungsverarbeitung, die durch die Codierungsvorrichtung 71 in **Fig. 28** durchgeführt wird, wird unten unter Bezugnahme auf einen Ablaufplan von **Fig. 29** beschrieben.

[0380] In Schritt S281 quantisiert die Quantisierungseinheit 21 des Metacodierers 11 Parameter als gelieferte Metadaten und liefert die resultierenden quantisierten Parameter zur Codierungseinheit 22.

[0381] Es ist festzuhalten, dass in Schritt S281 dieselbe Verarbeitung wie in Schritt S251 von **Fig. 27** durchgeführt wird, jedoch die Quantisierungseinheit 21 Positionsinformationen, die die relative Position des Objekts vom Anwender gesehen angeben, Positionsinformationen, die die geeignet geänderte absolute Position des Objekts angeben oder dergleichen als Positionsinformationen, die die Metadaten des Objekts bilden, auf der Grundlage der gelieferten Anwenderpositionsinformationen und Objektpositionsinformationen quantisiert.

[0382] Wenn der Prozess eines Schritts S281 durchgeführt wird, werden die Prozesse der Schritte S282 bis S287 danach durchgeführt und wird die Codierungsverarbeitung beendet. Allerdings sind diese Prozesse gleich den Prozessen der Schritte S252 bis S257 in **Fig. 27** und deshalb wird ihre Beschreibung ausgelassen.

[0383] Allerdings werden in Schritt S284 die Hörpsychologieparameter unter Verwendung nicht nur der Etiketteninformationen, sondern auch der Anwenderpositionsinformationen und der Objektpositionsinformationen, die oben beschrieben sind, berechnet. Ferner kann in Schritt S285 eine Bitzuweisung unter Verwendung der Anwenderpositionsinformationen oder der Objektpositionsinformationen durchgeführt werden.

[0384] Wie oben beschrieben ist, führt die Codierungsvorrichtung 71 eine Berechnung von Hörpsychologieparametern und eine Bitzuweisung unter Verwendung nicht nur der Etiketteninformationen, sondern auch der Anwenderpositionsinformationen und der Objektpositionsinformationen durch. Auf diese Weise ist es möglich, die Codierungseffizienz und die Verarbeitungsgeschwindigkeit einer Quantisierungsberechnung zu verbessern, ein Präsenzgefühl verbessern und eine Audiowiedergabe mit höherer Klangqualität zu realisieren.

[0385] Wie oben beschrieben ist, berücksichtigt die vorliegende Technologie einen Verstärkungswert

von Metadaten, die beim Rendern angewendet werden, während des Betrachtens, die Position von Objekten, und dergleichen und ist es dadurch möglich, eine Berechnung von Hörpsychologieparametern und eine Bitzuweisung, die an die tatsächliche Hörempfindung angepasst sind, durchzuführen und die Codierungseffizienz zu verbessern.

[0386] Zusätzlich ist selbst dann, wenn ein Verstärkungswert von Metadaten, die durch einen Inhaltsersteller erstellt werden, aus dem Umfang der MPEG-H-Spezifikationen fällt, der Verstärkungswert tatsächlich nicht auf obere und untere Grenzwerte im Spezifikationsbereich beschränkt und ist es mit Ausnahme einer Klangqualitätsverschlechterung aufgrund einer Quantisierung möglich, einen Rendering-Schall wiederzugeben, wie durch den Urheber vorgesehen wurde.

[0387] Zum Beispiel liegt ein Fall vor, in dem ein Audiosignal eines bestimmten Objekts denselben Verstärkungspegel wie ein weiteres Objekt aufweist und ein Verstärkungswert in Metadaten 0 ($-\infty$ dB) ist, was für eine Rauschsperrung beabsichtigt ist. In einem derartigen Fall werden, obwohl ein Audiosignal, das tatsächlich gerendert und betrachtet wird, aus Nulldaten besteht, Bits auf dieselbe Weise zugewiesen, wie in weiteren Objekten in einer allgemeinen Codierungsvorrichtung. Allerdings wird in der vorliegenden Technologie eine Bitzuweisung als Nulldaten durchgeführt und ist es somit möglich, die Anzahl von Quantisierungsbits wesentlich zu verringern.

<Konfigurationsbeispiel eines Computers>

[0388] Im Übrigen kann die oben beschriebene Reihe von Prozessen auch durch Hardware oder Software ausgeführt werden. Falls die Reihe von Prozessen durch Software ausgeführt wird, ist ein Programm, das die Software konfiguriert, in einem Computer installiert. Hier weist der Computer z. B. einen Computer, der in fest zugeordneter Hardware gebildet ist, einen allgemein verwendbaren Personal Computer, in dem verschiedene Programme installiert sind, um in der Lage zu sein, verschiedene Funktionen auszuführen, und dergleichen auf.

[0389] **Fig. 30** ist ein Blockdiagramm, das ein Konfigurationsbeispiel einer Computer-Hardware veranschaulicht, die die oben beschriebene Reihe von Prozessen unter Verwendung eines Programms ausführt.

[0390] Im Computer sind eine Zentraleinheit (CPU) 501, ein Festwertspeicher (ROM) 502 und ein Schreib-/Lese-Speicher (RAM) 503 miteinander durch einen Bus 504 verbunden.

[0391] Ferner ist eine Eingabe/Ausgabe-Schnittstelle 505 mit dem Bus 504 verbunden. Eine Eingabe-

beeinheit 506, eine Ausgabeinheit 507, eine Aufzeichnungseinheit 508, eine Kommunikationseinheit 509 und ein Laufwerk 510 sind mit der Eingabe/Ausgabe-Schnittstelle 505 verbunden.

[0392] Die Eingabeeinheit 506 ist eine Tastatur, eine Maus, ein Mikrofon, ein Bildgebungselement oder dergleichen. Die Ausgabeinheit 507 ist eine Anzeigevorrichtung, ein Lautsprecher oder dergleichen. Die Aufzeichnungseinheit 508 ist aus einer Festplatte, einem nichtflüchtigen Speicher oder dergleichen gebildet. Die Kommunikationseinheit 509 ist eine Netzchnittstelle oder dergleichen. Das Laufwerk 510 steuert eine entnehmbares Aufzeichnungsmedium 511 wie z. B. einen Magnetdatenträger, einen optischen Datenträger, eine magneto-optische Platte oder einen Halbleiterspeicher.

[0393] Im Computer, der die oben beschriebene Konfiguration aufweist, führt z. B. die CPU 501 die oben beschriebene Reihe von Prozessen durch Laden eines Programms, das in der Aufzeichnungseinheit 508 gespeichert ist, in den RAM 503 mittels der Eingabe/Ausgabe-Schnittstelle 505 und des Buses 504 und Ausführen des Programms durch.

[0394] Das Programm, das durch den Computer (die CPU 501) ausgeführt wird, kann z. B. im entnehmbaren Aufzeichnungsmedium 511, das als ein Verpackungsmedium zur Lieferung dient, aufgezeichnet sein. Das Programm kann mittels eines drahtgebundenen oder drahtlosen Übertragungsmediums wie z. B. eines lokalen Netzes, des Internet oder einer digitalen Satellitenübertragung geliefert werden.

[0395] Im Computer ist es durch Installieren des entnehmbaren Aufzeichnungsmediums 511 im Laufwerk 510 möglich, das Programm mittels der Eingabe/Ausgabe-Schnittstelle 505 in der Aufzeichnungseinheit 508 zu installieren. Ferner kann das Programm durch die Kommunikationseinheit 509 mittels eines drahtgebundenen oder drahtlosen Übertragungsmediums, das in der Aufzeichnungseinheit 508 installiert werden soll, empfangen werden. Zusätzlich kann dieses Programm in dem ROM 502 oder der Aufzeichnungseinheit 508 im Voraus installiert werden.

[0396] Es ist festzuhalten, dass das Programm, das durch einen Computer ausgeführt wird, ein Programm sein kann, das eine Verarbeitung chronologisch in der Reihenfolge durchführt, die in der vorliegenden Spezifikation beschrieben ist, oder ein Programm sein kann, das eine Verarbeitung parallel oder zu einem erforderlichen Zeitpunkt wie z. B. einer Aufrufzeit durchführt.

[0397] Ausführungsformen der vorliegenden Technologie sind nicht auf die oben beschriebenen Aus-

führungsformen beschränkt und können im Umfang der vorliegenden Technologie unterschiedlich geändert werden, ohne vom Hauptinhalt der vorliegenden Technologie abzuweichen.

[0398] Zum Beispiel kann die vorliegende Technologie als Cloud-Berechnung konfiguriert sein, wobei mehrere Vorrichtungen eine Funktion mittels eines Netzes gemeinsam verwenden und sie zusammenwirkend verarbeiten.

[0399] Zusätzlich kann jeder Schritt, der in dem Ablaufplan oben beschrieben ist, durch eine Vorrichtung ausgeführt werden oder in einer gemeinsam verwendeten Weise durch mehrere Vorrichtungen ausgeführt werden.

[0400] Ferner können, falls ein Schritt mehrere Prozesse aufweist, die mehreren Prozessen, die in dem einen Schritt enthalten sind, durch eine Vorrichtung ausgeführt werden oder in einer gemeinsam verwendeten Weise durch mehrere Vorrichtungen ausgeführt werden.

[0401] Ferner kann die vorliegende Technologie konfiguriert sein, wie folgt.

(1) Eine Signalverarbeitungsvorrichtung, die Folgendes aufweist:

eine Korrekturereinheit, die konfiguriert ist, ein Audiosignal eines Audio-Objekts auf der Grundlage eines Verstärkungswerts, der in Metadaten des Audio-Objekts enthalten ist, zu korrigieren; und

eine Quantisierungseinheit, die konfiguriert ist, Hörpsychologieparameter auf der Grundlage eines Signals, das durch die Korrektur erhalten wird, zu berechnen und das Audiosignal zu quantisieren.

(2) Die Signalverarbeitungsvorrichtung nach (1), wobei die Korrekturereinheit das Audiosignal in einer Zeitdomäne auf der Grundlage des Verstärkungswerts korrigiert.

(3) Die Signalverarbeitungsvorrichtung nach (2), die ferner Folgendes aufweist:

eine Zeitfrequenzumsetzungseinheit, die konfiguriert ist, eine Zeit/Frequenzumsetzung an dem korrigierten Audiosignal, das durch die Korrektur durch die Korrekturereinheit erhalten wird, durchzuführen, wobei

die Quantisierungseinheit die Hörpsychologieparameter auf der Grundlage von Frequenzspektruminformationen, die durch die Zeit/Frequenzumsetzung erhalten werden, berechnet.

(4) Die Signalverarbeitungsvorrichtung nach (1), die ferner Folgendes aufweist:

eine Zeitfrequenzumsetzungseinheit, die konfiguriert ist, eine Zeit/Frequenzumsetzung am Audiosignal durchzuführen, wobei

die Korrekturereinheit Frequenzspektruminformationen, die durch die Zeit/Frequenz-Umsetzung auf der Grundlage des Verstärkungswerts erhalten werden, korrigiert, und

die Quantisierungseinheit die Hörpsychologieparameter auf der Grundlage der korrigierten Frequenzspektruminformationen, die durch die Korrektur der Korrekturereinheit erhalten werden, berechnet.

(5) Die Signalverarbeitungsvorrichtung nach einem von (1) bis (4), die ferner Folgendes aufweist:

eine Verstärkungskorrekturereinheit, die konfiguriert ist, den Verstärkungswert auf der Grundlage von Hörcharakteristiken, die mit einer Einfallrichtung eines Schalls in Beziehung stehen, zu korrigieren, wobei

die Korrekturereinheit das Audiosignal auf der Grundlage des korrigierten Verstärkungswerts korrigiert.

(6) Die Signalverarbeitungsvorrichtung nach (5), wobei die Verstärkungskorrekturereinheit den Verstärkungswert auf der Grundlage der Hörcharakteristiken in Bezug auf eine Position, die durch Positionsinformationen, die in den Metadaten enthalten sind, angegeben ist, korrigiert.

(7) Die Signalverarbeitungsvorrichtung nach (6), die ferner Folgendes aufweist:

eine Hörcharakteristiktabellen-Halteeinheit, die konfiguriert ist, eine Hörcharakteristiktafel zu halten, in der die Position des Audio-Objekts und ein Verstärkungskorrekturwert zum Durchführen einer Korrektur auf der Grundlage der Hörcharakteristik des Verstärkungswerts für die Position des Audio-Objekts einander zugeordnet sind.

(8) Die Signalverarbeitungsvorrichtung nach (7), wobei, falls der Verstärkungskorrekturwert, der der Position entspricht, die durch die Positionsinformationen angegeben ist, nicht in der Hörcharakteristiktafel ist, die Verstärkungskorrekturereinheit eine Interpolationsverarbeitung auf der Grundlage von mehreren Verstärkungskorrekturwerten in der Hörcharakteristiktafel durchführt, um den Verstärkungskorrekturwert für eine Position, die durch die Positionsinformationen angegeben ist, zu erhalten.

(9) Die Signalverarbeitungsvorrichtung nach (8), wobei die Verstärkungskorrekturereinheit die Interpolationsverarbeitung auf der Grundlage der Verstärkungskorrekturwerte durchführt, die den mehreren Positionen in der Nähe der Posi-

tion, die durch die Positionsinformationen angegeben ist, zugeordnet sind.

(10) Die Signalverarbeitungsvorrichtung nach (9), wobei die Interpolationsverarbeitung eine Interpolationsverarbeitung unter Verwendung von VBAP ist.

(11) Die Signalverarbeitungsvorrichtung nach (8), wobei der Verstärkungskorrekturwert jeder von mehreren Frequenzen für jede Position in der Hörcharakteristiktafel zugeordnet ist, und falls die Hörcharakteristiktafel den Verstärkungskorrekturwert für eine vorgegebene Frequenz, der der Position, die durch die Positionsinformationen angegeben ist, entspricht, nicht aufweist, die Verstärkungskorrekturereinheit die Interpolationsverarbeitung auf der Grundlage der Verstärkungskorrekturwerte von mehreren weiteren Frequenzen in der Nähe der vorgegebenen Frequenz durchführt, um den Verstärkungskorrekturwert für die vorgegebene Frequenz für die Position, die durch die Positionsinformationen angegeben ist, zu erhalten wobei, die mehreren weiteren Frequenzen der Position entsprechen, die durch die Positionsinformationen angegeben ist.

(12) Die Signalverarbeitungsvorrichtung nach (8), wobei die Hörcharakteristiktabellen-Halteeinheit die Hörcharakteristiktafel für jeden Wiedergabeschalldruck hält, und die Verstärkungskorrekturereinheit die Hörcharakteristiktafel, die verwendet wird, um den Verstärkungswert zu korrigieren, auf der Grundlage eines Schalldrucks des Audiosignals wechselt.

(13) Die Signalverarbeitungsvorrichtung nach (12), wobei, falls die Hörcharakteristiktafel, die dem Schalldruck des Audiosignals entspricht, nicht in der Hörcharakteristiktabellen-Halteeinheit gehalten wird, die Verstärkungskorrekturereinheit die Interpolationsverarbeitung auf der Grundlage des Verstärkungskorrekturwerts, der der Position, die durch die Positionsinformationen angegeben ist, in der Hörcharakteristiktafel von mehreren weiteren Wiedergabeschalldrücken in der Nähe des Schalldrucks entspricht, durchführt, um den Verstärkungskorrekturwert für die Position, die durch die Positionsinformationen, die dem Schalldruck entsprechen, angegeben ist, zu erhalten.

(14) Die Signalverarbeitungsvorrichtung nach einem von (7) bis (13), wobei die Verstärkungskorrekturereinheit den Verstärkungswert in Übereinstimmung mit Eigenschaften des Audiosignals begrenzt.

(15) Die Signalverarbeitungsvorrichtung nach (7), wobei, falls der Verstärkungskorrekturwert, der der Position entspricht, die durch die Positionsinformationen angegeben ist, nicht in der

Hörcharakteristiktabelle ist, die Verstärkungskorrektureinheit den Verstärkungswert unter Verwendung des Verstärkungskorrekturwerts, der einer Position zugeordnet ist, die am nächsten zu der Position liegt, die durch die Positionsinformationen angegeben ist, korrigiert.

(16) Die Signalverarbeitungsvorrichtung nach (7), wobei, falls der Verstärkungskorrekturwert, der der Position entspricht, die durch die Positionsinformationen angegeben ist, nicht in der Hörcharakteristiktabelle ist, die Verstärkungskorrektureinheit einen Durchschnittswert der Verstärkungskorrekturwerte, die den mehreren Positionen in der Nähe der Position, die durch die Positionsinformationen angegeben ist, zugeordnet sind, als den Verstärkungskorrekturwert der Position, die durch die Positionsinformationen angegeben ist, einstellt.

(17) Ein Signalverarbeitungsverfahren, das Folgendes aufweist:

Verursachen, dass eine Signalverarbeitungsvorrichtung ein Audiosignal eines Audio-Objekts auf der Grundlage eines Verstärkungswerts, der in Metadaten des Audio-Objekts enthalten ist, korrigiert und Hörpsychologieparameter auf der Grundlage eines Signals, das durch die Korrektur erhalten wird, berechnet und das Audiosignal quantisiert.

(18) Ein Programm, das verursacht, dass ein Computer eine Verarbeitung ausführt, die Schritte aufweist, die Folgendes umfassen:

Korrigieren eines Audiosignals eines Audio-Objekts auf der Grundlage eines Verstärkungswerts, der in Metadaten des Audio-Objekts enthalten ist; und

Berechnen von Hörpsychologieparametern auf der Grundlage eines Signals, das durch die Korrektur und ein Quantisieren des Audiosignals erhalten wird.

(19) Eine Signalverarbeitungsvorrichtung, die Folgendes aufweist:

eine Änderungseinheit, die konfiguriert ist, einen Verstärkungswert eines Audio-Objekts und ein Audiosignal auf der Grundlage des Verstärkungswerts, der in Metadaten des Audio-Objekts enthalten ist, zu ändern; und

eine Quantisierungseinheit, die konfiguriert ist, das geänderte Audiosignal, das durch die Änderung erhalten wird, zu quantisieren.

(20) Die Signalverarbeitungsvorrichtung nach (19), wobei die Änderungseinheit die Änderung durchführt, falls der Verstärkungswert ein Wert ist, der aus einem vorgegebenen Bereich fällt.

(21) Die Signalverarbeitungsvorrichtung nach (19) oder (20), die ferner Folgendes aufweist:

eine Korrektureinheit, die konfiguriert ist, das geänderte Audiosignal auf der Grundlage des geänderten Verstärkungswerts, der durch die Änderung erhalten wird, zu korrigieren, wobei

die Quantisierungseinheit das geänderte Audiosignal auf der Grundlage eines Signals, das durch das Korrigieren des geänderten Audiosignals erhalten wird, quantisiert.

(22) Die Signalverarbeitungsvorrichtung nach einem von (19) bis (21), die ferner Folgendes aufweist:

einen Metacodierer, der konfiguriert ist, die Metadaten, die den geänderten Verstärkungswert aufweisen, der durch die Änderung erhalten wird, zu quantisieren und zu codieren;

eine Codierungseinheit, die konfiguriert ist, das quantisierte geänderte Audiosignal zu codieren; und

eine Multiplexiereinheit, die konfiguriert ist, die codierten Metadaten und das codierte geänderte Audiosignal zu multiplexieren.

(23) Die Signalverarbeitungsvorrichtung nach einem von (19) bis (22), wobei die Änderungseinheit das Audiosignal auf der Grundlage einer Differenz zwischen dem Verstärkungswert und dem geänderten Verstärkungswert, der durch die Änderung erhalten wird, ändert.

(24) Ein Signalverarbeitungsverfahren, das Folgendes aufweist: Verursachen, dass eine Signalverarbeitungsvorrichtung einen Verstärkungswert eines Audio-Objekts und ein Audiosignal auf der Grundlage des Verstärkungswerts, der in Metadaten des Audio-Objekts enthalten ist, ändert und das geänderte Audiosignal, das durch die Änderung erhalten wird, zu quantisieren.

(25) Ein Programm, das verursacht, dass ein Computer eine Verarbeitung ausführt, die Schritte aufweist, die Folgendes umfassen:

Ändern eines Verstärkungswerts eines Audio-Objekts und eines Audiosignals auf der Grundlage des Verstärkungswerts, der in Metadaten des Audio-Objekts enthalten ist; und

Quantisieren des geänderten Audiosignals, das durch die Änderung erhalten wird.

(26) Eine Signalverarbeitungsvorrichtung, die Folgendes aufweist:

eine Quantisierungseinheit, die konfiguriert ist, Hörpsychologieparameter auf der Grundlage von Metadaten, die einen Verstärkungswert und/oder Positionsinformationen eines Audio-

Objekts aufweisen, eines Audiosignals des Audio-Objekts und eines Hörpsychologiemodells, das mit der Hörverdeckung zwischen mehreren Audio-Objekten in Beziehung steht, zu berechnen und das Audiosignal auf der Grundlage der Hörpsychologieparameter zu quantisieren.

(27) Die Signalverarbeitungsvorrichtung nach (26), die ferner Folgendes aufweist:

eine Zeitfrequenzumsetzungseinheit, die konfiguriert ist, eine Zeit/FrequenzUmsetzung am Audiosignal durchzuführen, wobei

die Quantisierungseinheit die Hörpsychologieparameter auf der Grundlage von Frequenzspektruminformationen, die durch die Zeit/Frequenz-Umsetzung erhalten werden, berechnet.

(28) Die Signalverarbeitungsvorrichtung nach (26) oder (27), wobei die Quantisierungseinheit die Hörpsychologieparameter auf der Grundlage der Metadaten und des Audiosignals des Audio-Objekts, das verarbeitet werden soll, der Metadaten und der Audiosignale der weiteren Audio-Objekte und des Hörpsychologiemodells berechnet.

(29) Die Signalverarbeitungsvorrichtung nach einem von (26) bis (28), wobei die Metadaten Bearbeitungsgenehmigungsinformationen aufweisen, die eine Genehmigung angeben, einige oder alle von mehreren Parametern, die den Verstärkungswert und die Positionsinformationen, die in den Metadaten enthalten sind, aufweisen, zu bearbeiten, und die Quantisierungseinheit die Hörpsychologieparameter auf der Grundlage der Parameter, für die ein Bearbeiten durch die Bearbeitungsgenehmigungsinformationen nicht erlaubt ist, der Audiosignale und des Hörpsychologiemodells berechnet.

(30) Ein Signalbehandlungsverfahren, das Folgendes aufweist:

Verursachen, dass eine Signalverarbeitungsvorrichtung Hörpsychologieparameter auf der Grundlage von Metadaten, die einen Verstärkungswert und/oder Positionsinformationen eines Audio-Objekts aufweisen, eines Audiosignals des Audio-Objekts und eines Hörpsychologiemodells, das mit der Hörverdeckung zwischen mehreren Audio-Objekten in Beziehung steht, berechnet und das Audiosignal auf der Grundlage der Hörpsychologieparameter quantisiert.

(31) Ein Programm, das verursacht, dass ein Computer eine Verarbeitung ausführt, die Schritte aufweist, die Folgendes umfassen: Berechnen von Hörpsychologieparametern auf der Grundlage von Metadaten, die einen Ver-

stärkungswert und/oder Positionsinformationen eines Audio-Objekts aufweisen, eines Audiosignals des Audio-Objekts und eines Hörpsychologiemodells, das mit der Hörverdeckung zwischen mehreren Audio-Objekten in Beziehung steht, und Quantisieren des Audiosignals auf der Grundlage der Hörpsychologieparameter.

(32) Eine Signalverarbeitungsvorrichtung, die Folgendes aufweist:

eine Quantisierungseinheit, die konfiguriert ist, ein Audiosignal eines Audio-Objekts unter Verwendung mindestens eines Einstellungsparameters und eines Algorithmus, der für den Typ der Schallquelle, der durch Etiketteninformationen angegeben ist, die den Typ der Schallquelle des Audio-Objekts angeben, bestimmt wird, auf der Grundlage des Audiosignals des Audio-Objekts und der Etiketteninformationen zu quantisieren.

(33) Die Signalverarbeitungsvorrichtung nach (32), wobei die Quantisierungseinheit Hörpsychologieparameter auf der Grundlage des Audiosignals und der Etiketteninformationen berechnet und das Audiosignal auf der Grundlage der Hörpsychologieparameter quantisiert.

(34) Die Signalverarbeitungsvorrichtung nach (32) oder (33), wobei die Quantisierungseinheit eine Bitzuweisung und eine Quantisierung des Audiosignals auf der Grundlage der Etiketteninformationen durchführt.

(35) Die Signalverarbeitungsvorrichtung nach einem von (32) bis (34), die ferner Folgendes aufweist:

eine Zeitfrequenzumsetzungseinheit, die konfiguriert ist, eine Zeit/FrequenzUmsetzung am Audiosignal unter Verwendung mindestens eines der Einstellungsparameter und des Algorithmus, der für den Typ der Schallquelle,

der durch die Etiketteninformationen auf der Grundlage der Etiketteninformationen angegeben ist, bestimmt wird, durchzuführen, wobei die Quantisierungseinheit die Hörpsychologieparameter auf der Grundlage von Frequenzspektruminformationen, die durch die Zeit/Frequenz-Umsetzung erhalten werden, berechnet und die Frequenzspektruminformationen quantisiert.

(36) Die Signalverarbeitungsvorrichtung nach einem von (32) bis (35), wobei die Etiketteninformationen ferner Hörumgebungsinformationen aufweisen, die eine Schallhörumgebung auf der Grundlage des Audiosignals angeben, und die Quantisierungseinheit das Audiosignal unter Verwendung mindestens eines Einstellungsparameters und eines Algorithmus, der für den Typ der Schallquelle bestimmt wird, und der

Hörumgebung, die durch die Etiketteninformationen angegeben ist, quantisiert.

(37) Die Signalverarbeitungsvorrichtung nach einem von (32) bis (35), wobei die Quantisierungseinheit einen Einstellungsparameter, der für den Typ der Schallquelle, der durch die Etiketteninformationen angegeben ist, bestimmt wird, auf der Grundlage der Priorität des Audio-Objekts einstellt.

(38) Die Signalverarbeitungsvorrichtung nach einem von (32) bis (35), wobei die Quantisierungseinheit das Audiosignal auf der Grundlage von Positionsinformationen eines Anwenders, Positionsinformationen des Audio-Objekts, des Audiosignals und der Etiketteninformationen quantisiert.

(39) Ein Signalverarbeitungsverfahren, das Folgendes aufweist:

Verursachen, dass eine Signalverarbeitungsvorrichtung ein Audiosignal eines Audio-Objekts unter Verwendung mindestens eines Einstellungsparameters und eines Algorithmus, der für den Typ der Schallquelle, der durch Etiketteninformationen angegeben ist, die den Typ der Schallquelle des Audio-Objekts angeben, bestimmt wird auf der Grundlage des Audiosignals des Audio-Objekts und der Etiketteninformationen quantisiert.

(40) Ein Programm, das verursacht, dass ein Computer eine Verarbeitung ausführt, die Schritte aufweist, die Folgendes umfassen:

Quantisieren eines Audiosignals eines Audio-Objekts unter Verwendung mindestens eines Einstellungsparameters und eines Algorithmus, der für den Typ der Schallquelle, der durch Etiketteninformationen angegeben ist, die den Typ der Schallquelle des Audio-Objekts angeben, bestimmt wird, auf der Grundlage des Audiosignals des Audio-Objekts und der Etiketteninformationen.

Bezugszeichenliste

11	Metacodierer
12	Kerncodierer
31	Zeitfrequenzumsetzungseinheit
32	Quantisierungseinheit
33	Codierungseinheit
71	Codierungsvorrichtung
81	Multiplexiereinheit
91	Audiosignalkorrekturereinheit
92	Zeitfrequenzumsetzungseinheit
131	MDCT-Koeffizientenkorrekturereinheit

ZITATE ENTHALTEN IN DER BESCHREIBUNG

Diese Liste der vom Anmelder aufgeführten Dokumente wurde automatisiert erzeugt und ist ausschließlich zur besseren Information des Lesers aufgenommen. Die Liste ist nicht Bestandteil der deutschen Patent- bzw. Gebrauchsmusteranmeldung. Das DPMA übernimmt keinerlei Haftung für etwaige Fehler oder Auslassungen.

Zitierte Nicht-Patentliteratur

- ISO/IEC 23003-3, MPEG-D USAC [0002]
- ISO/IEC 23008-3, MPEG-H 3D Audio [0002]

Patentansprüche

1. Signalverarbeitungsvorrichtung, die Folgendes umfasst:

eine Korrekturereinheit, die konfiguriert ist, ein Audiosignal eines Audio-Objekts auf der Grundlage eines Verstärkungswerts, der in Metadaten des Audio-Objekts enthalten ist, zu korrigieren; und
eine Quantisierungseinheit, die konfiguriert ist, Hörpsychologieparameter auf der Grundlage eines Signals, das durch die Korrektur erhalten wird, zu berechnen und das Audiosignal zu quantisieren.

2. Signalverarbeitungsvorrichtung nach Anspruch 1, wobei die Korrekturereinheit das Audiosignal in einer Zeitdomäne auf der Grundlage des Verstärkungswerts korrigiert.

3. Signalverarbeitungsvorrichtung nach Anspruch 2, die ferner Folgendes umfasst:
eine Zeitfrequenzumsetzungseinheit, die konfiguriert ist, eine Zeit/Frequenzumsetzung an dem korrigierten Audiosignal, das durch die Korrektur durch die Korrekturereinheit erhalten wird, durchzuführen, wobei die Quantisierungseinheit die Hörpsychologieparameter auf der Grundlage von Frequenzspektruminformationen, die durch die Zeit/Frequenz-Umsetzung erhalten werden, berechnet.

4. Signalverarbeitungsvorrichtung nach Anspruch 1, die ferner Folgendes umfasst:
eine Zeitfrequenzumsetzungseinheit, die konfiguriert ist, eine Zeit/Frequenzumsetzung am Audiosignal durchzuführen, wobei die Korrekturereinheit Frequenzspektruminformationen, die durch die Zeit/Frequenz-Umsetzung auf der Grundlage des Verstärkungswerts erhalten werden, korrigiert, und
die Quantisierungseinheit die Hörpsychologieparameter auf der Grundlage der korrigierten Frequenzspektruminformationen, die durch die Korrektur der Korrekturereinheit erhalten werden, berechnet.

5. Signalverarbeitungsvorrichtung nach Anspruch 1, die ferner Folgendes umfasst:
eine Verstärkungskorrekturereinheit, die konfiguriert ist, den Verstärkungswert auf der Grundlage von Hörcharakteristiken, die mit einer Einfallsrichtung eines Schalls in Beziehung stehen, zu korrigieren, wobei die Korrekturereinheit das Audiosignal auf der Grundlage des korrigierten Verstärkungswerts korrigiert.

6. Signalverarbeitungsvorrichtung nach Anspruch 5, wobei die Verstärkungskorrekturereinheit den Verstärkungswert auf der Grundlage der Hörcharakteristiken in Bezug auf eine Position, die durch Positionsinformationen, die in den Metadaten enthalten sind, angegeben ist, korrigiert.

7. Signalverarbeitungsvorrichtung nach Anspruch 6, die ferner Folgendes umfasst:
eine Hörcharakteristiktabelle-Halteeinheit, die konfiguriert ist, eine Hörcharakteristiktabelle zu halten, in der die Position des Audio-Objekts und ein Verstärkungskorrekturwert zum Durchführen einer Korrektur auf der Grundlage der Hörcharakteristik des Verstärkungswerts für die Position des Audio-Objekts einander zugeordnet sind.

8. Signalverarbeitungsvorrichtung nach Anspruch 7, wobei, falls der Verstärkungskorrekturwert, der der Position entspricht, die durch die Positionsinformationen angegeben ist, nicht in der Hörcharakteristiktabelle ist, die Verstärkungskorrekturereinheit eine Interpolationsverarbeitung auf der Grundlage der Verstärkungskorrekturwerte, die mehreren Positionen in der Nähe der Position, die durch die Positionsinformationen angegeben ist, zugeordnet sind, durchführt, um den Verstärkungskorrekturwert der Position, die durch die Positionsinformationen angegeben ist, zu erhalten, um den Verstärkungskorrekturwert, der einer Position, die am nächsten zu der Position liegt, die durch die Positionsinformationen angegeben ist, zugeordnet ist, als den Verstärkungskorrekturwert der Position, die durch die Positionsinformationen angegeben ist, zu setzen oder um einen Durchschnittswert der Verstärkungskorrekturwerte, die den mehreren Positionen in der Nähe der Position, die durch die Positionsinformationen angegeben ist, zugeordnet sind, als den Verstärkungskorrekturwert der Position, die durch die Positionsinformationen angegeben ist, zu setzen.

9. Signalverarbeitungsvorrichtung nach Anspruch 8, wobei die Interpolationsverarbeitung eine Interpolationsverarbeitung unter Verwendung von VBAP ist.

10. Signalverarbeitungsverfahren, das Folgendes umfasst:
Verursachen, dass eine Signalverarbeitungsvorrichtung ein Audiosignal eines Audio-Objekts auf der Grundlage eines Verstärkungswerts, der in Metadaten des Audio-Objekts enthalten ist, korrigiert und Hörpsychologieparameter auf der Grundlage eines Signals, das durch die Korrektur erhalten wird, berechnet und das Audiosignal quantisiert.

11. Programm, das verursacht, dass ein Computer eine Verarbeitung ausführt, die Schritte aufweist, die Folgendes umfassen:
Korrigieren eines Audiosignals eines Audio-Objekts auf der Grundlage eines Verstärkungswerts, der in Metadaten des Audio-Objekts enthalten ist; und Berechnen von Hörpsychologieparametern auf der Grundlage eines Signals, das durch die Korrektur erhalten wird, und Quantisieren des Audiosignals.

12. Signalverarbeitungsvorrichtung, die Folgendes umfasst:

eine Änderungseinheit, die konfiguriert ist, einen Verstärkungswert eines Audio-Objekts und ein Audiosignal auf der Grundlage des Verstärkungswerts, der in Metadaten des Audio-Objekts enthalten ist, zu ändern; und
eine Quantisierungseinheit, die konfiguriert ist, das geänderte Audiosignal, das durch die Änderung erhalten wird, zu quantisieren.

13. Signalverarbeitungsvorrichtung nach Anspruch 12, wobei die Änderungseinheit die Änderung durchführt, falls der Verstärkungswert ein Wert ist, der aus einem vorgegebenen Bereich fällt.

14. Signalverarbeitungsvorrichtung nach Anspruch 12, die ferner Folgendes umfasst:

eine Korrekturereinheit, die konfiguriert ist, das geänderte Audiosignal auf der Grundlage des geänderten Verstärkungswerts, der durch die Änderung erhalten wird, zu korrigieren, wobei die Quantisierungseinheit das geänderte Audiosignal auf der Grundlage eines Signals, das durch das Korrigieren des geänderten Audiosignals erhalten wird, quantisiert.

15. Signalverarbeitungsvorrichtung nach Anspruch 12, die ferner Folgendes umfasst:

einen Metacodierer, der konfiguriert ist, die Metadaten, die den geänderten Verstärkungswert aufweisen, der durch die Änderung erhalten wird, zu quantisieren und zu codieren;
eine Codierungseinheit, die konfiguriert ist, das quantisierte geänderte Audiosignal zu codieren; und eine Multiplexiereinheit, die konfiguriert ist, die codierten Metadaten und das codierte geänderte Audiosignal zu multiplexieren.

16. Signalverarbeitungsvorrichtung nach Anspruch 12, wobei die Änderungseinheit das Audiosignal auf der Grundlage einer Differenz zwischen dem Verstärkungswert und dem geänderten Verstärkungswert, der durch die Änderung erhalten wird, ändert.

17. Signalbehandlungsverfahren, das Folgendes umfasst: Verursachen, dass eine Signalverarbeitungsvorrichtung einen Verstärkungswert eines Audio-Objekts und ein Audiosignal auf der Grundlage des Verstärkungswerts, der in Metadaten des Audio-Objekts enthalten ist, ändert und das geänderte Audiosignal, das durch die Änderung erhalten wird, quantisiert.

18. Programm, das verursacht, dass ein Computer eine Verarbeitung ausführt, die Schritte aufweist, die Folgendes umfassen:

Ändern eines Verstärkungswerts eines Audio-Objekts und eines Audiosignals auf der Grundlage

des Verstärkungswerts, der in Metadaten des Audio-Objekts enthalten ist; und
Quantisieren des geänderten Audiosignals, das durch die Änderung erhalten wird.

19. Signalverarbeitungsvorrichtung, die Folgendes umfasst:

eine Quantisierungseinheit, die konfiguriert ist, Hörpsychologieparameter auf der Grundlage von Metadaten, die einen Verstärkungswert und/oder Positionsinformationen eines Audio-Objekts aufweisen, eines Audiosignals des Audio-Objekts und eines Hörpsychologiemodells, das mit der Hörmaskierung zwischen mehreren Audio-Objekten in Beziehung steht, zu berechnen und das Audiosignal auf der Grundlage der Hörpsychologieparameter zu quantisieren.

20. Signalverarbeitungsvorrichtung nach Anspruch 19, die ferner Folgendes umfasst:

eine Zeitfrequenzumsetzungseinheit, die konfiguriert ist, eine Zeit/Frequenzumsetzung am Audiosignal durchzuführen, wobei die Quantisierungseinheit die Hörpsychologieparameter auf der Grundlage von Frequenzspektruminformationen, die durch die Zeit/Frequenzumsetzung erhalten werden, berechnet.

21. Signalverarbeitungsvorrichtung nach

Anspruch 19, wobei die Quantisierungseinheit die Hörpsychologieparameter auf der Grundlage der Metadaten und des Audiosignals des Audio-Objekts, das verarbeitet werden soll, der Metadaten und der Audiosignale der weiteren Audio-Objekte, und des Hörpsychologiemodells berechnet.

22. Signalverarbeitungsvorrichtung nach

Anspruch 19, wobei die Metadaten Bearbeitungsgenehmigungsinformationen aufweisen, die eine Genehmigung angeben, einige oder alle von mehreren Parametern, die den Verstärkungswert und die Positionsinformationen, die in den Metadaten enthalten sind, aufweisen, zu bearbeiten, und die Quantisierungseinheit die Hörpsychologieparameter auf der Grundlage der Parameter, für die ein Bearbeiten durch die Bearbeitungsgenehmigungsinformationen nicht erlaubt ist, der Audiosignale und des Hörpsychologiemodells berechnet.

23. Signalbehandlungsverfahren, das Folgendes umfasst:

Verursachen, dass eine Signalverarbeitungsvorrichtung Hörpsychologieparameter auf der Grundlage von Metadaten, die einen Verstärkungswert und/oder Positionsinformationen eines Audio-Objekts aufweisen, eines Audiosignals des Audio-Objekts und eines Hörpsychologiemodells, das mit der Hörmaskierung zwischen mehreren Audio-Objekten in Beziehung steht, berechnet und das Audiosignal

auf der Grundlage der Hörpsychologieparameter quantisiert.

24. Programm, das verursacht, dass ein Computer eine Verarbeitung ausführt, die Schritte aufweist, die Folgendes umfassen:

Berechnen von Hörpsychologieparametern auf der Grundlage von Metadaten, die einen Verstärkungswert und/oder Positionsinformationen eines Audio-Objekts aufweisen, eines Audiosignals des Audio-Objekts und eines Hörpsychologiemodells, das mit der Hörmaskierung zwischen mehreren Audio-Objekten in Beziehung steht, und Quantisieren des Audiosignals auf der Grundlage der Hörpsychologieparameter.

25. Signalverarbeitungsvorrichtung, die Folgendes umfasst:

eine Quantisierungseinheit, die konfiguriert ist, ein Audiosignal eines Audio-Objekts unter Verwendung mindestens eines Einstellungsparameters und eines Algorithmus, der für den Typ der Schallquelle, der durch Etiketteninformationen angegeben ist, die den Typ der Schallquelle des Audio-Objekts angeben, bestimmt wird, auf der Grundlage des Audiosignals und der Etiketteninformationen zu quantisieren.

26. Signalverarbeitungsvorrichtung nach Anspruch 25, wobei die Quantisierungseinheit Hörpsychologieparameter auf der Grundlage des Audiosignals und der Etiketteninformationen berechnet und das Audiosignal auf der Grundlage der Hörpsychologieparameter quantisiert.

27. Signalverarbeitungsvorrichtung nach Anspruch 25, wobei die Quantisierungseinheit eine Bitzuweisung und eine Quantisierung des Audiosignals auf der Grundlage der Etiketteninformationen durchführt.

28. Signalverarbeitungsvorrichtung nach Anspruch 25, die ferner Folgendes umfasst:

eine Zeitfrequenzumsetzungseinheit, die konfiguriert ist, eine Zeit/Frequenzumsetzung am Audiosignal unter Verwendung mindestens eines der Einstellungsparameter und des Algorithmus, der für den Typ der Schallquelle, der durch die Etiketteninformationen angegeben ist, bestimmt wird, auf der Grundlage der Etiketteninformationen durchzuführen, wobei

die Quantisierungseinheit die Hörpsychologieparameter auf der Grundlage von Frequenzspektruminformationen, die durch die Zeit/Frequenzumsetzung erhalten werden, berechnet und die Frequenzspektruminformationen quantisiert.

29.

Signalverarbeitungsvorrichtung nach Anspruch 25, wobei die Etiketteninformationen ferner Hörumge-

bungsinformationen aufweisen, die eine Schallhörumgebung auf der Grundlage des Audiosignals angeben, und

die Quantisierungseinheit das Audiosignal unter Verwendung mindestens eines Einstellungsparameters und eines Algorithmus, der für den Typ der Schallquelle bestimmt wird, und der Hörumgebung, die durch die Etiketteninformationen angegeben ist, quantisiert.

30. Signalverarbeitungsvorrichtung nach Anspruch 25, wobei die Quantisierungseinheit einen Einstellungsparameter, der für den Typ der Schallquelle, der durch die Etiketteninformationen angegeben ist, bestimmt wird, auf der Grundlage der Priorität des Audio-Objekts einstellt.

31. Signalverarbeitungsvorrichtung nach Anspruch 25, wobei die Quantisierungseinheit das Audiosignal auf der Grundlage von Positionsinformationen eines Anwenders, Positionsinformationen des Audio-Objekts, des Audiosignals und der Etiketteninformationen quantisiert.

32. Signalverarbeitungsverfahren, das Folgendes umfasst:

Verursachen, dass eine Signalverarbeitungsvorrichtung ein Audiosignal eines Audio-Objekts unter Verwendung mindestens eines Einstellungsparameters und eines Algorithmus, der für den Typ der Schallquelle, der durch Etiketteninformationen angegeben ist, die den Typ der Schallquelle des Audio-Objekts angeben, bestimmt wird, auf der Grundlage des Audiosignals des Audio-Objekts und der Etiketteninformationen quantisiert.

33. Programm, das verursacht, dass ein Computer eine Verarbeitung ausführt, die Schritte aufweist, die Folgendes umfassen:

Quantisieren eines Audiosignals eines Audio-Objekts unter Verwendung mindestens eines Einstellungsparameters und eines Algorithmus, der für den Typ der Schallquelle, der durch Etiketteninformationen angegeben ist, die den Typ der Schallquelle des Audio-Objekts angeben, bestimmt wird, auf der Grundlage des Audiosignals des Audio-Objekts und der Etiketteninformationen.

Es folgen 30 Seiten Zeichnungen

Anhängende Zeichnungen

Fig. 1

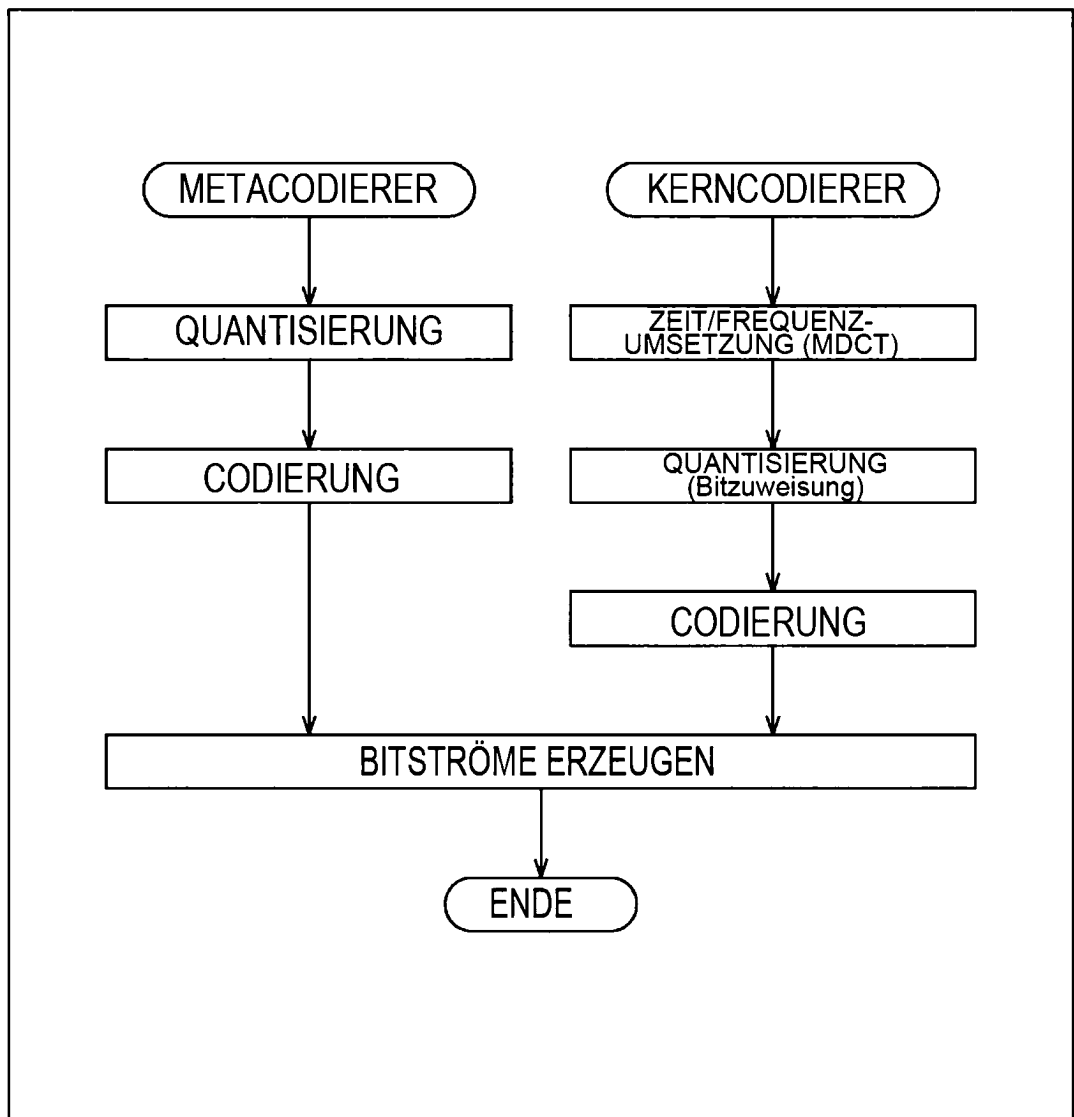


Fig. 2

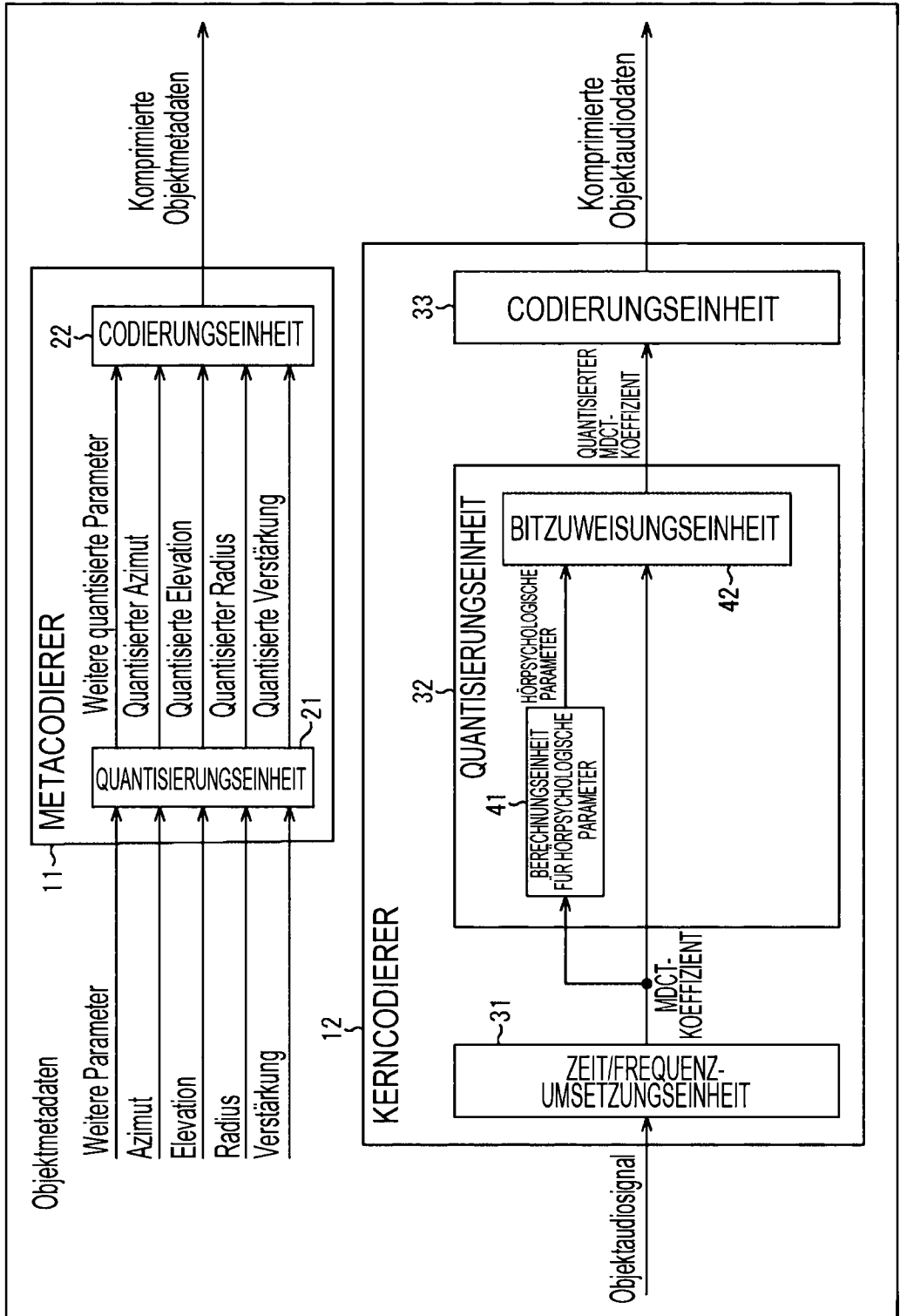


Fig. 3

Komponente	Einheit	Wertebereich
Azimut	° (Grad)	-180;180
Elevation	° (Grad)	-90;90
Radius	m (Meter)	0.5;16
Verstärkung	keine (linear)	0,004;5,957
Ausbreitung (gleichförmig)	° (Grad)	0;180
Ausbreitungsbreite	° (Grad)	0;180
Ausbreitungshöhe	° (Grad)	0;90
Ausbreitungstiefe	m (Meter)	0;15,5
dynamische Objektpriorität	keine	0;7

Fig. 4

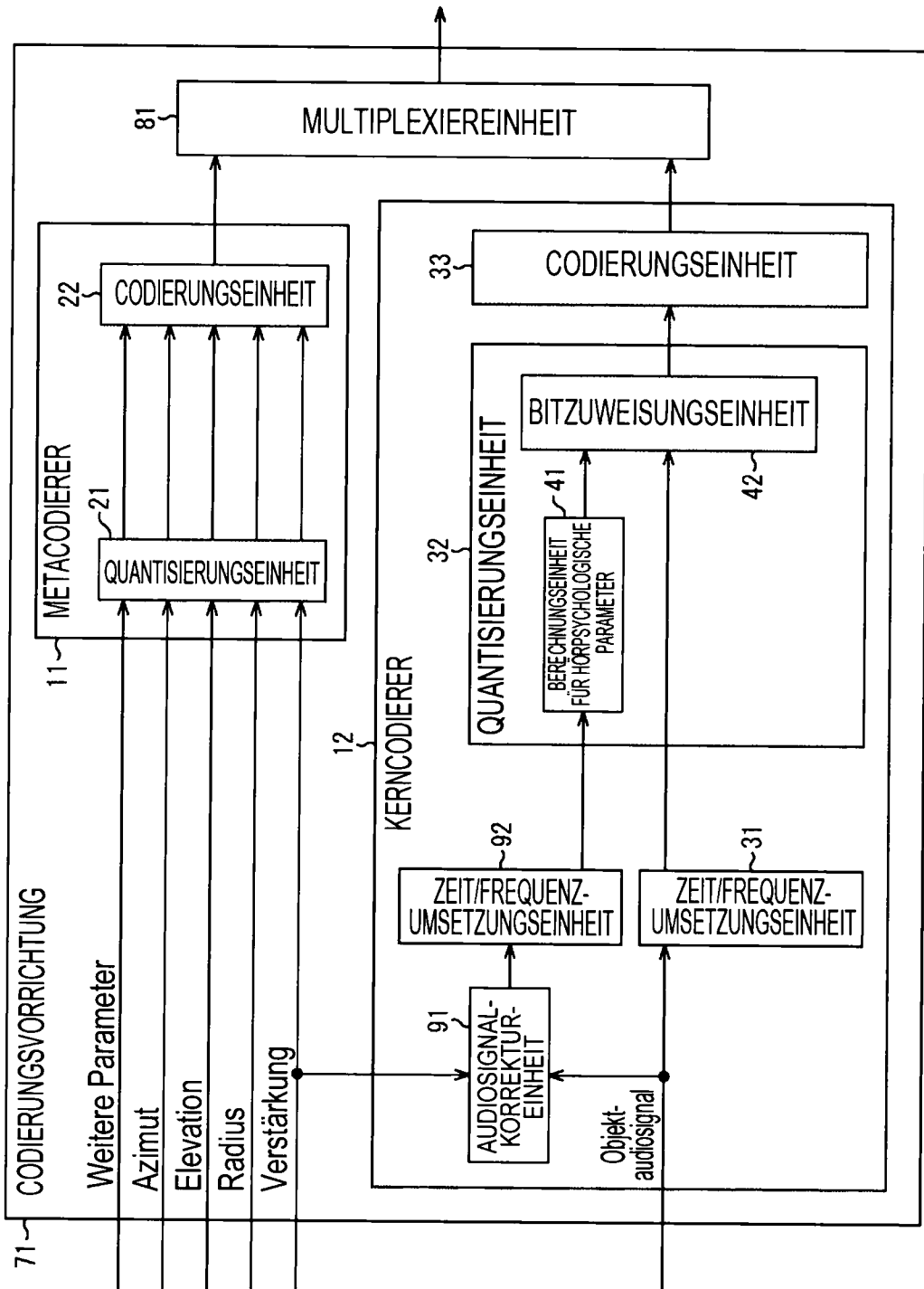


Fig. 5

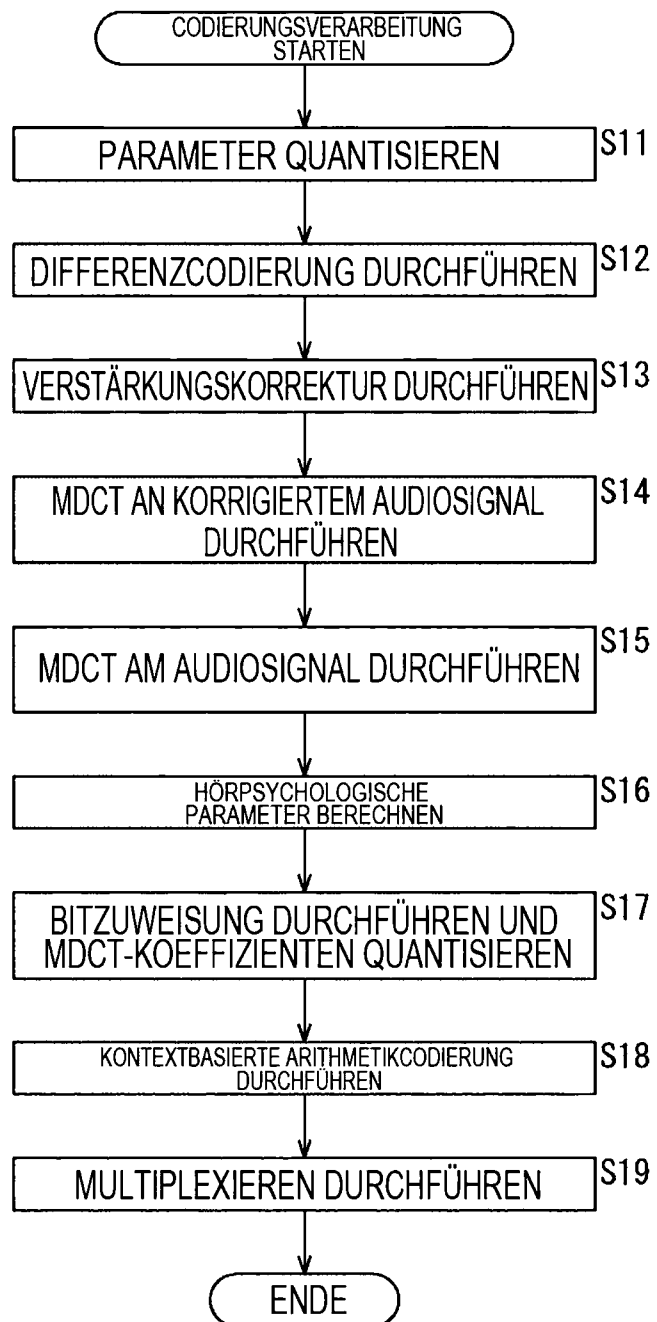


Fig. 6

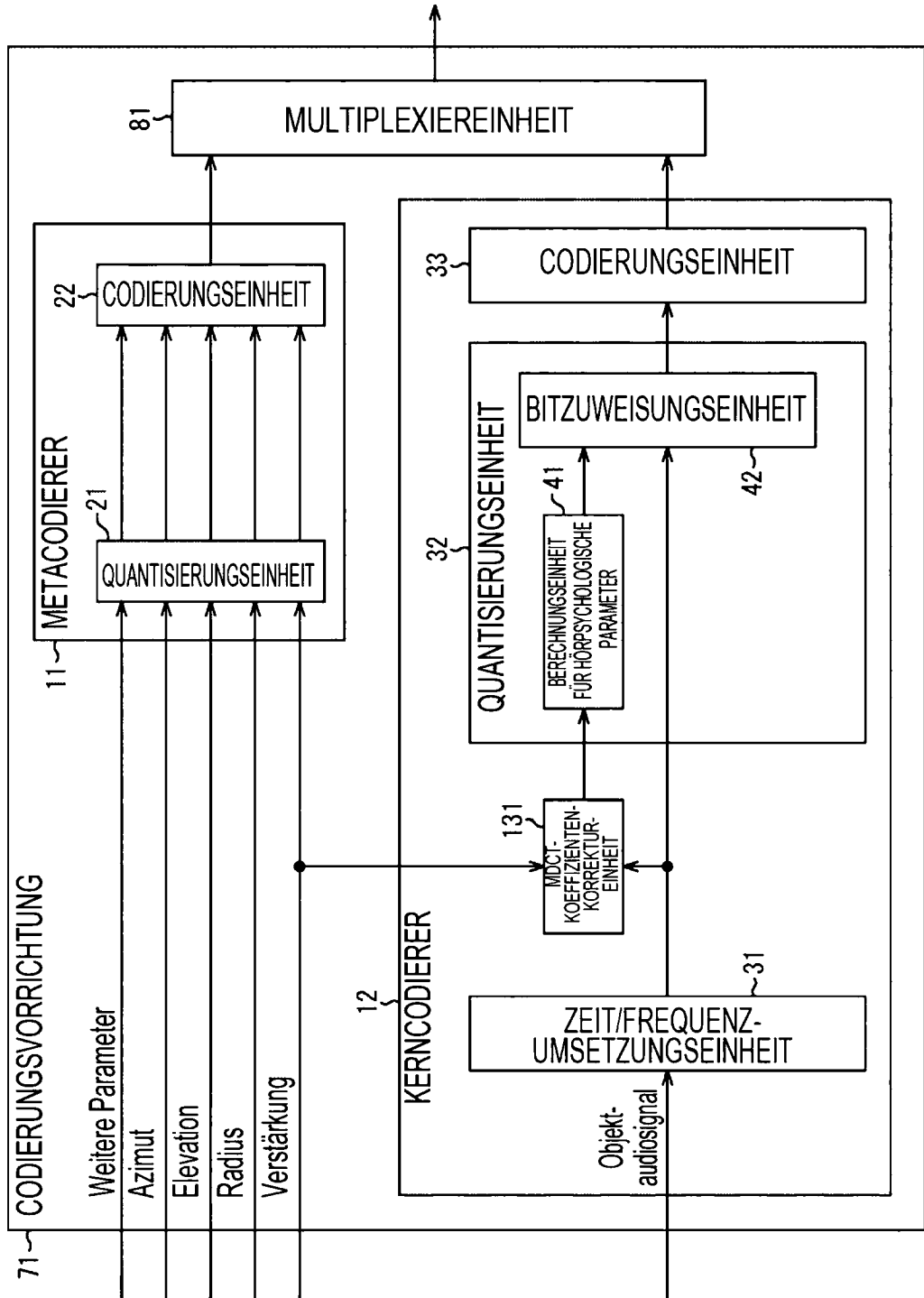


Fig. 7

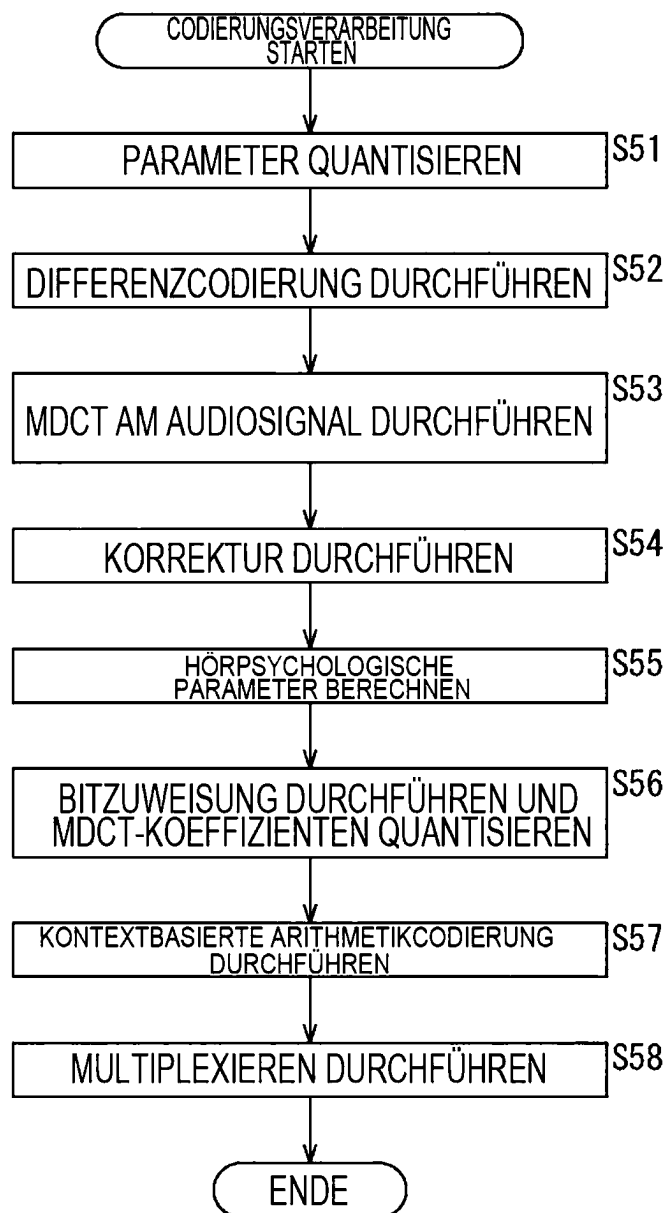


Fig. 8

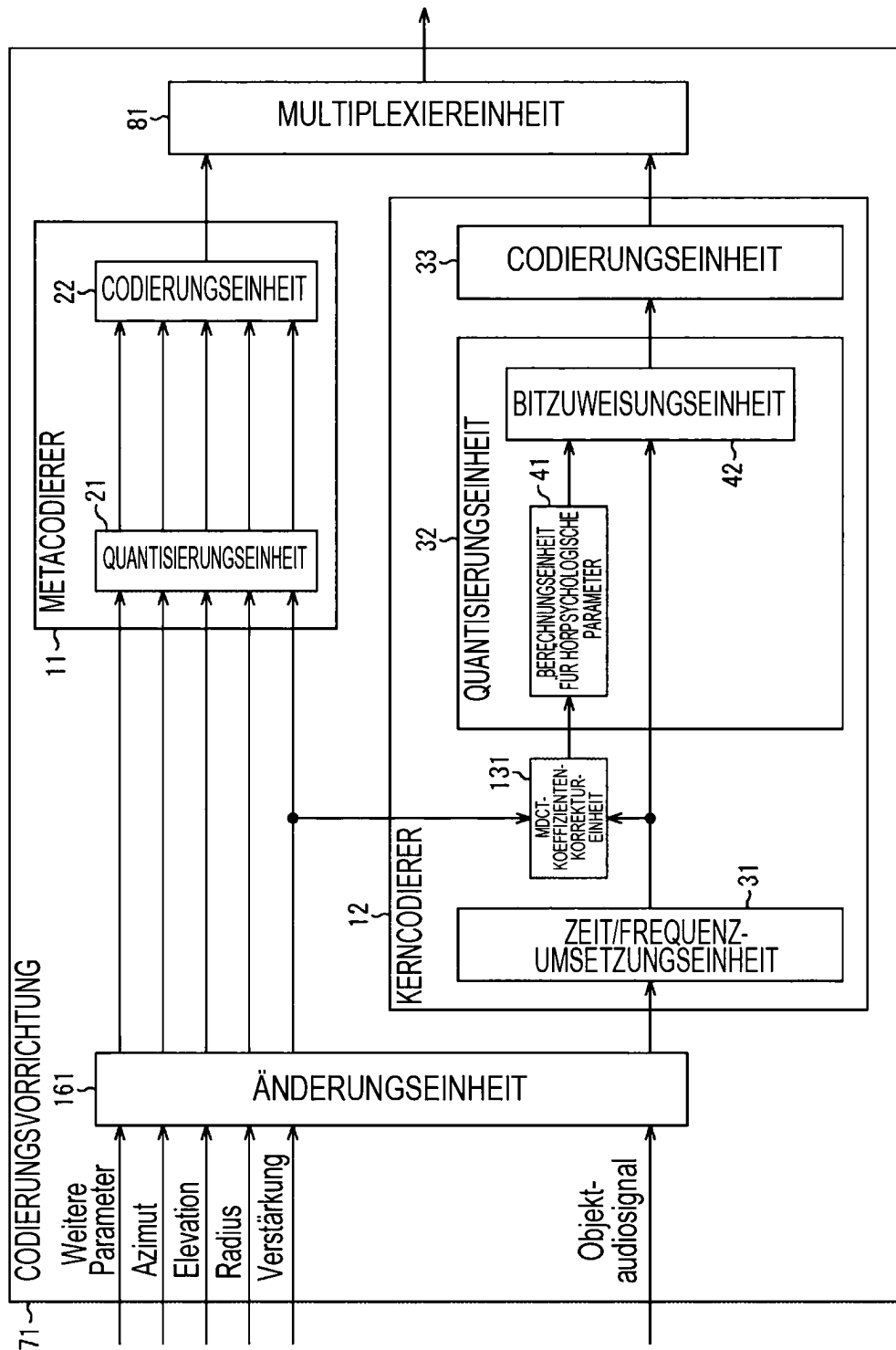


Fig. 9

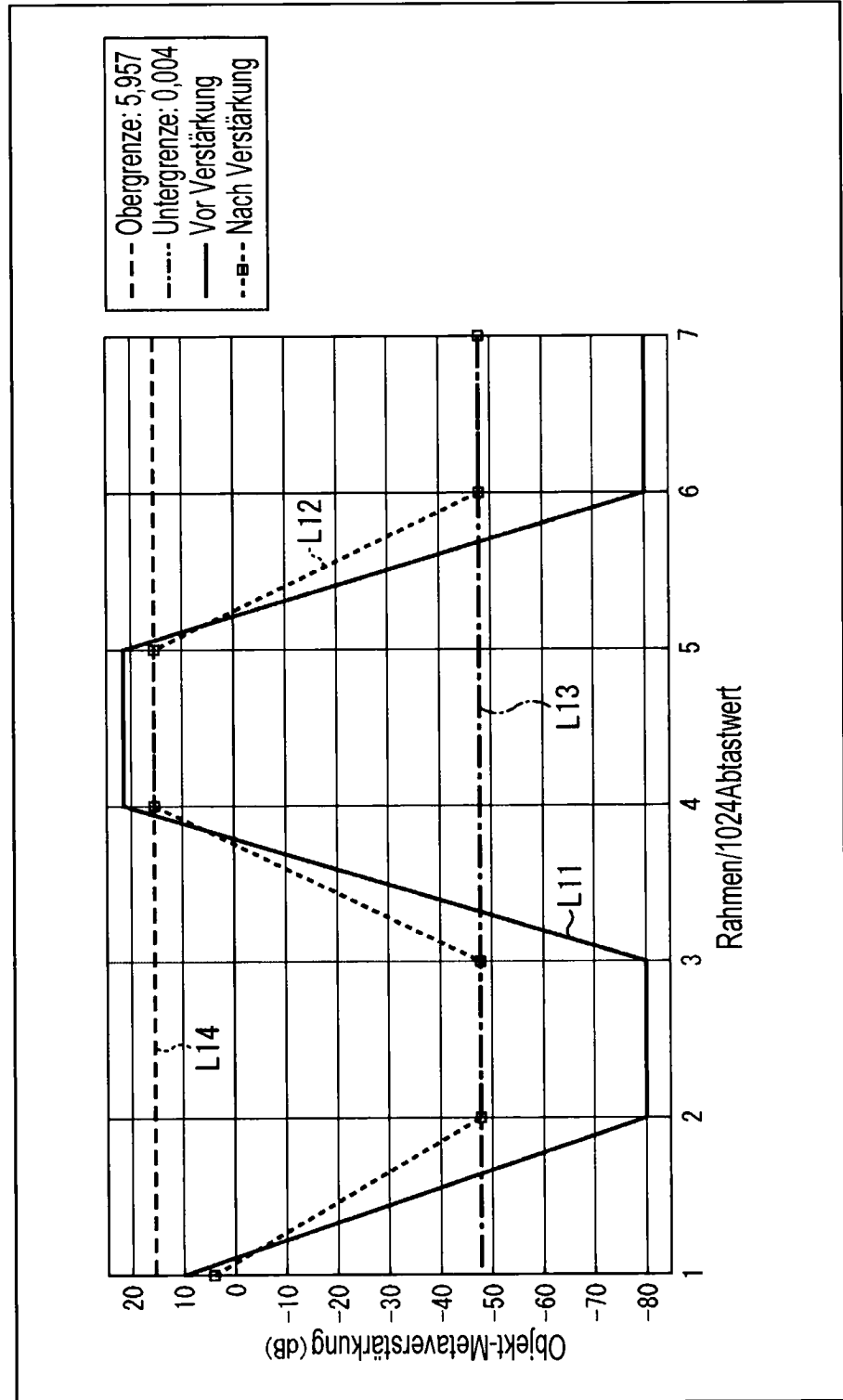


Fig. 10

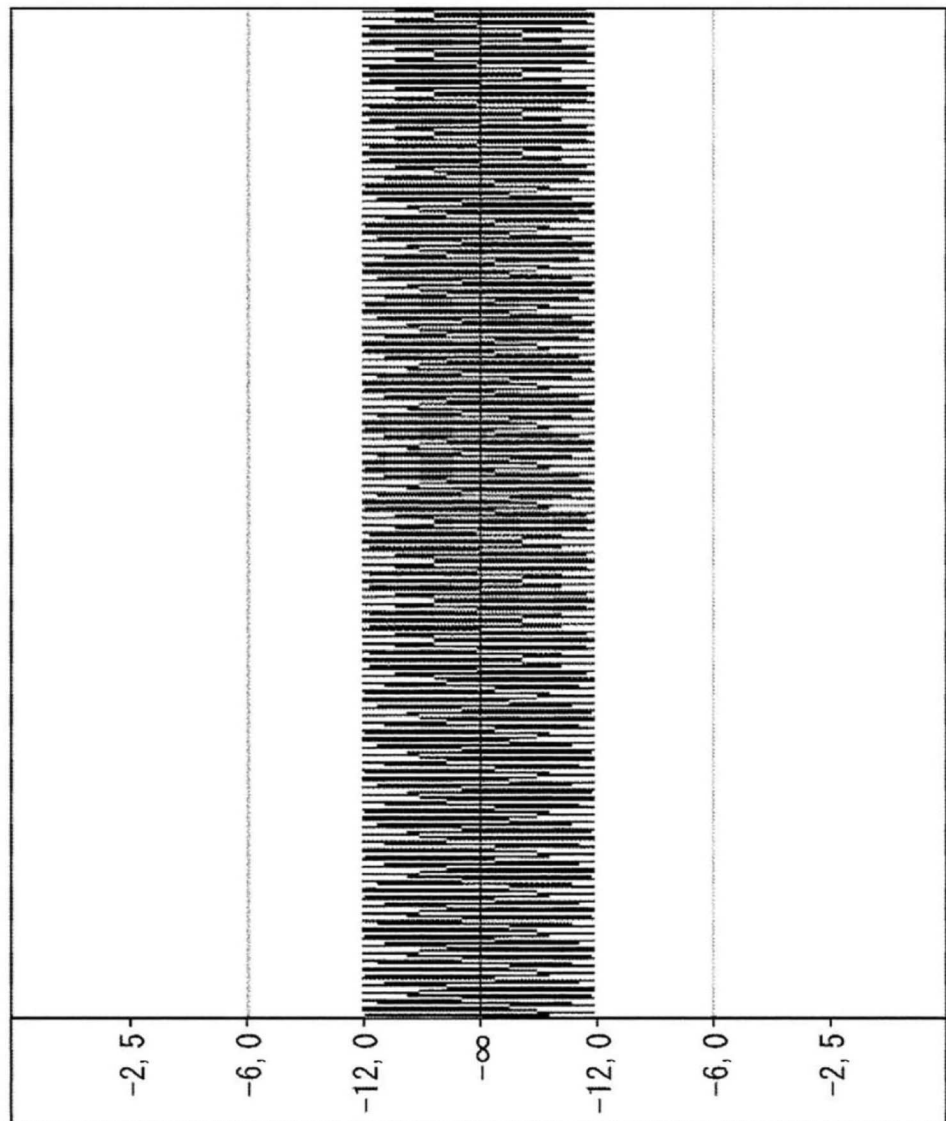


Fig. 11

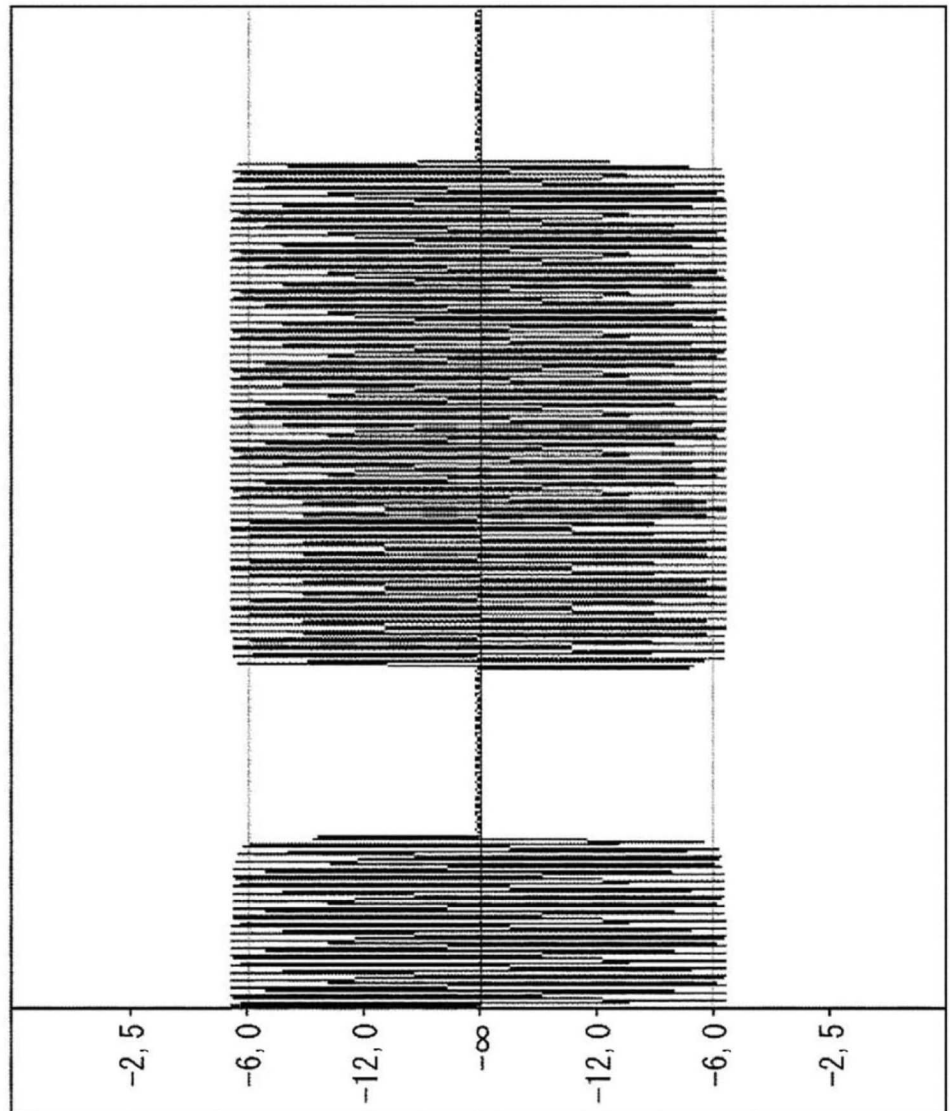


Fig. 12

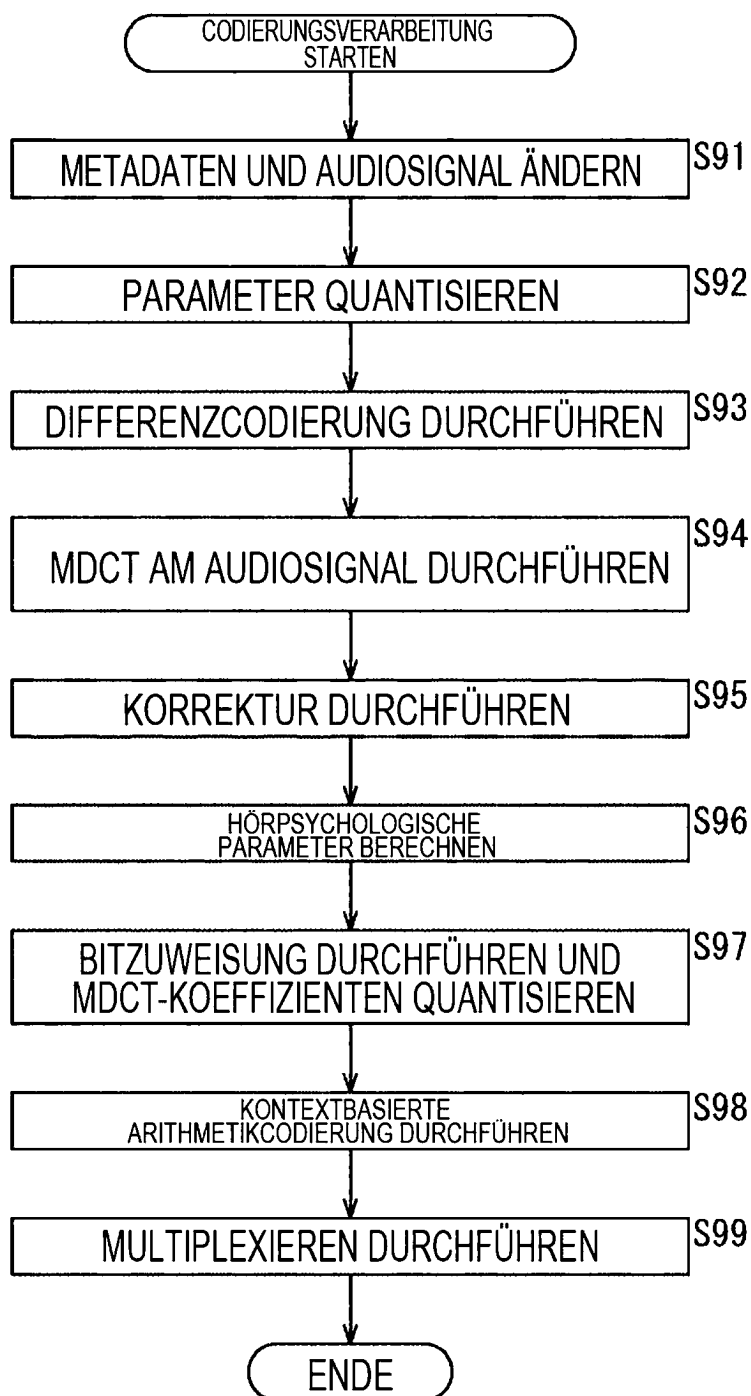


Fig. 13

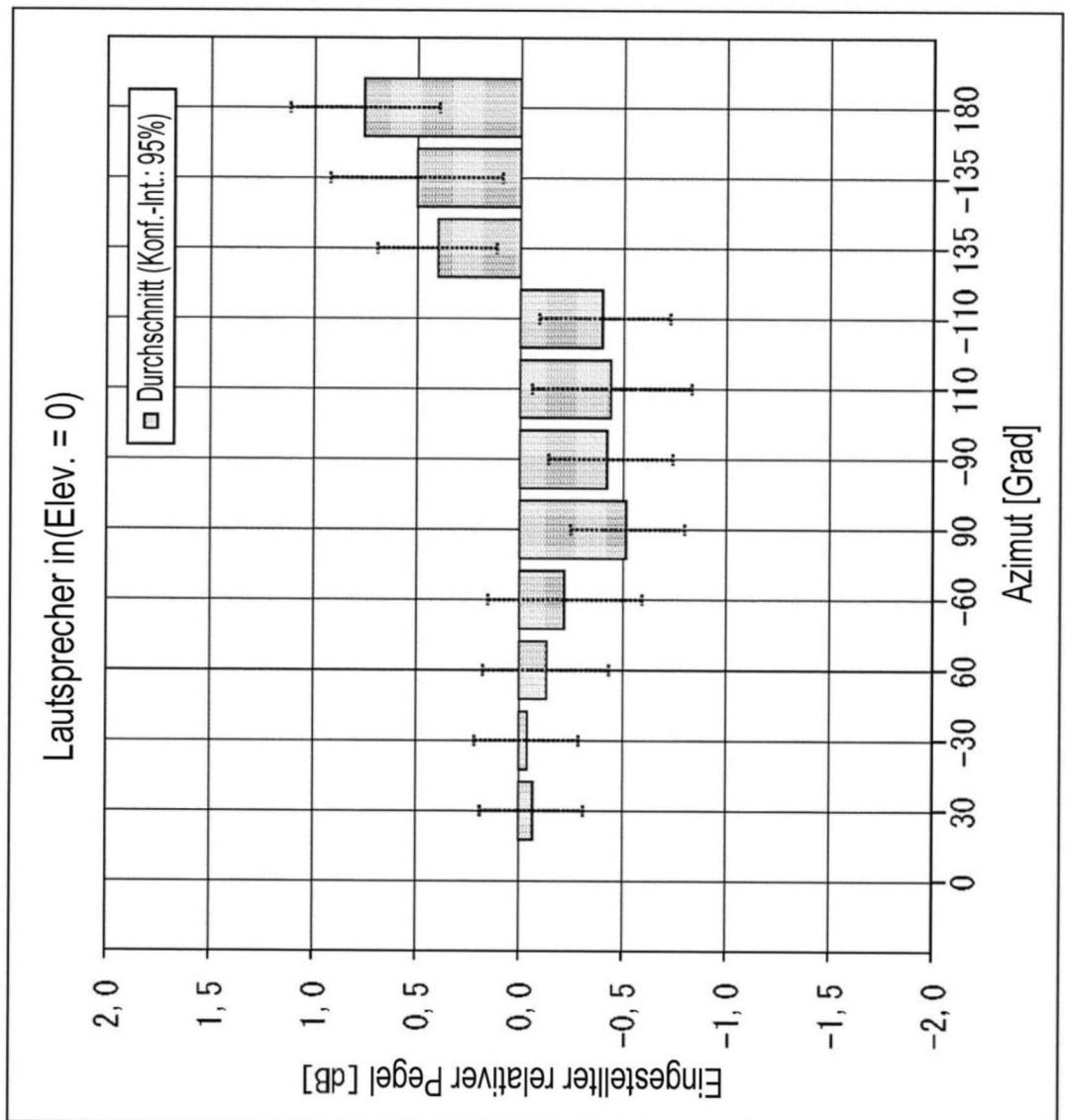


Fig. 14

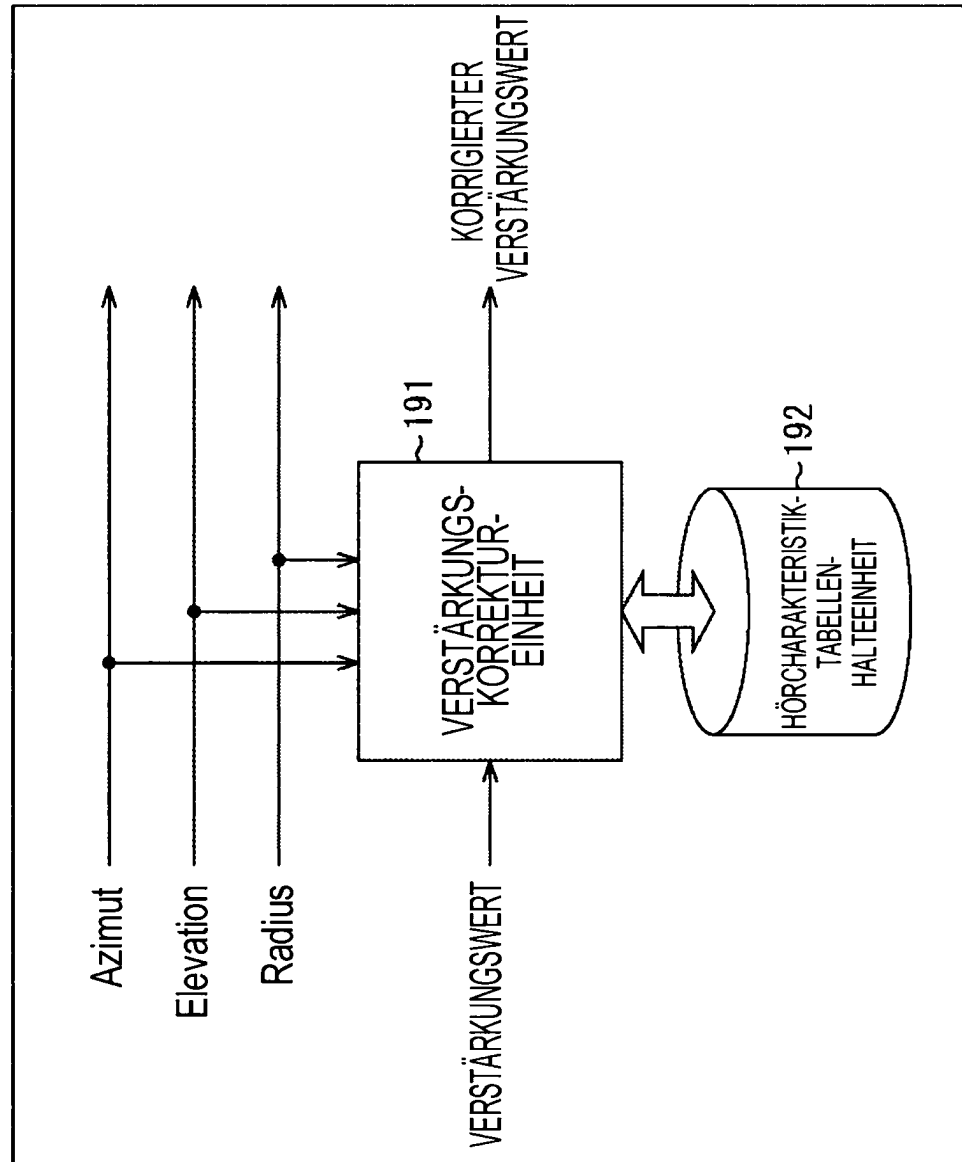


Fig. 15

AZIMUT [GRAD]	ELEVATION [GRAD]	Radius [m]	VERSTÄRKUNGSKORREKTURWERT [dB]
0	0	1,0	0,0
30	0	1,0	-0,02
-30	0	1,0	-0,07
60	0	1,0	-0,06
-60	0	1,0	-0,32
90	0	1,0	-0,52
-90	0	1,0	-0,44
120	0	1,0	-0,43
-120	0	1,0	-0,41
150	0	1,0	0,33
-150	0	1,0	0,51
180	0	1,0	0,63

Fig. 16

VERSTÄRKUNGSKORREKTURWERT [dB]		AZIMUT [GRAD]											
		0	30	-30	60	-60	90	-90	120	-120	150	-150	180
30		0,33	0,12	0,03	-0,07	-0,02	0,61	0,17	0,21	0,46	0,43	0,81	0,66
0		0,00	-0,02	-0,07	-0,06	-0,32	-0,52	-0,44	-0,43	-0,41	0,33	0,51	0,63
-30		-0,44			-0,28	-0,32					-0,15	-0,01	

Fig. 17

Azimut	Elevation	VERSTÄRKUNGSKORREKTURWERT		
		250Hz	1kHz	8kHz
0	0	0	0	0
-30	0	-0,91	-1,34	-0,92
-60	0	-1,45	-1,74	-1,52
-90	0	-1,43	-2,69	-3,04
-110	0	-1,23	-2,17	-3,14
-135	0	0,04	-2,28	-1,50
180	0	2,25	-1,67	1,67
135	0	0,54	-2,42	-1,62
110	0	-0,39	-2,39	-3,42
90	0	-0,51	-2,79	-2,89
60	0	-2,51	-2,28	-3,09
30	0	-1,01	-0,91	-0,57
0	30	0,75	1,39	-1,00
-30	30	-1,19	-0,61	-1,99
-45	30	-0,47	-0,06	-3,42
-90	30	0,20	-1,40	-4,14
-110	30	0,60	-0,38	-2,98
-135	30	2,19	-0,83	-1,54
180	30	0,93	-0,12	-0,53
135	30	2,63	-0,89	-1,06
110	30	0,11	-0,38	-3,45
90	30	0,39	-1,68	-4,24
45	30	-0,15	0,26	-2,92
30	30	-0,69	-0,14	-2,32
0	-30	2,36	-1,86	-0,45
-30	-30	1,00	-1,58	-0,94
-45	-30	-0,17	-1,83	-1,01
-135	-20	-0,18	-1,86	-1,14
180	-20	0,62	-0,52	1,16
135	-20	0,11	-1,42	-1,63
45	-30	-0,09	-1,94	-2,06
30	-30	0,66	-1,69	-1,18

Fig. 18

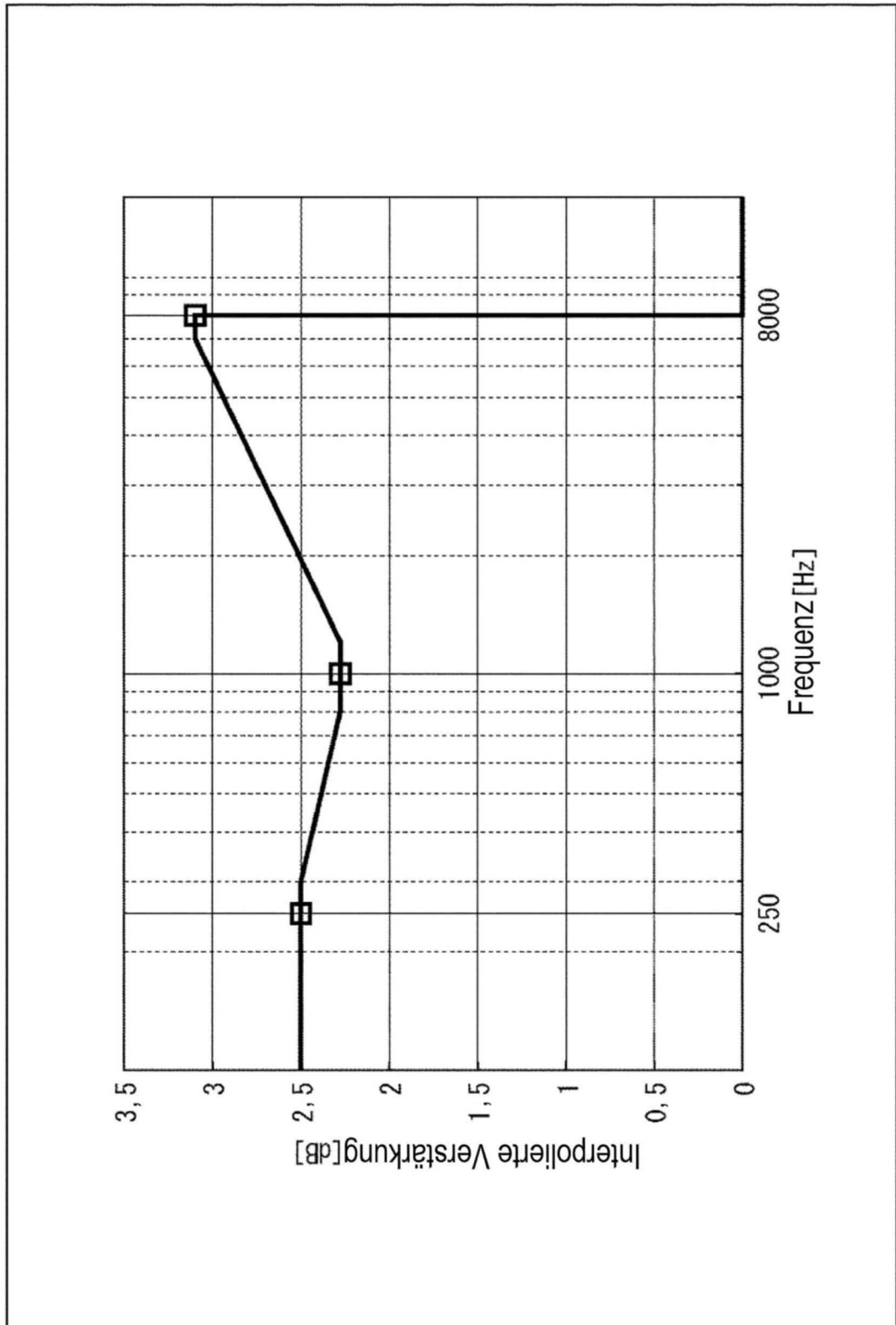


Fig. 19

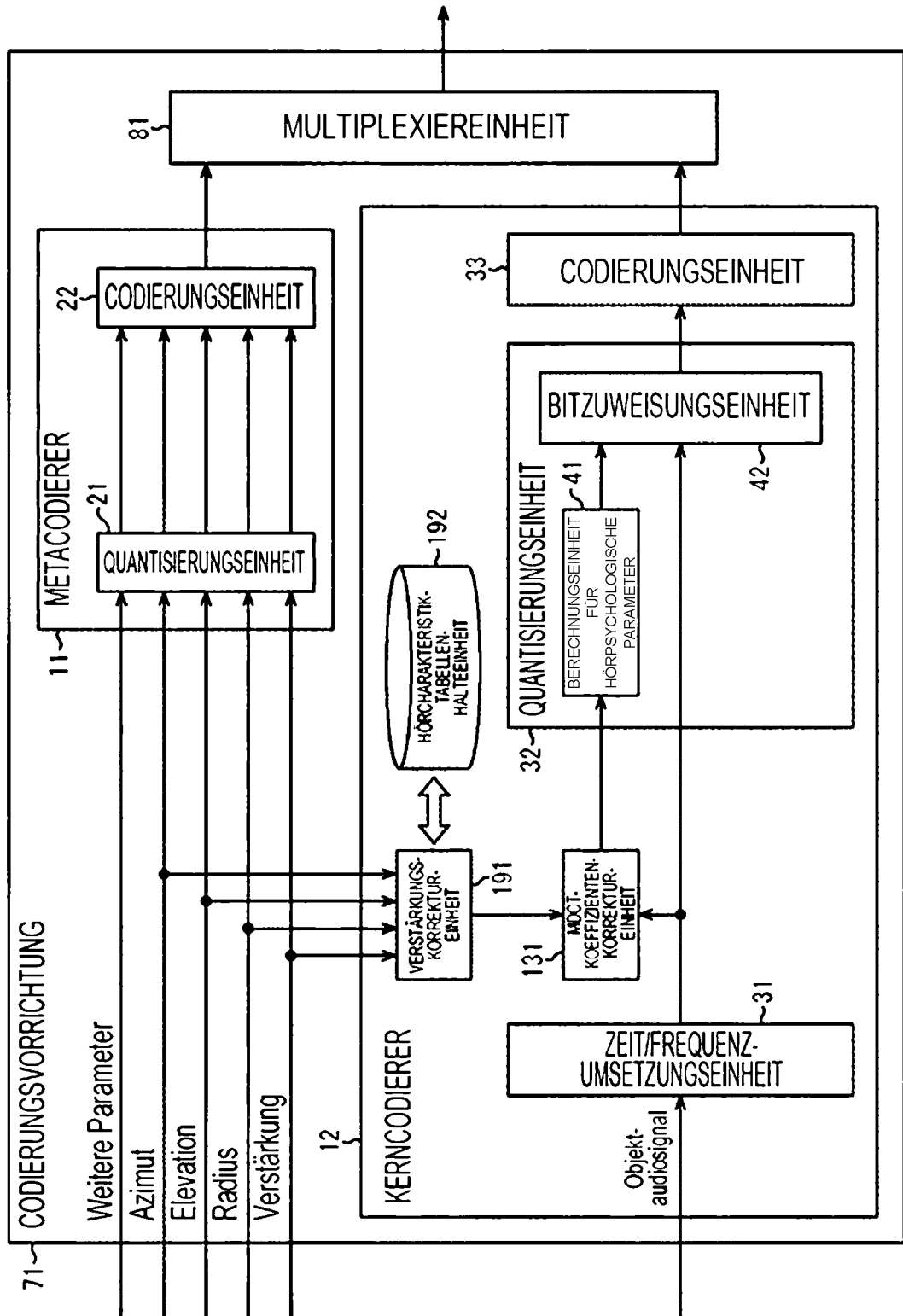


Fig. 20

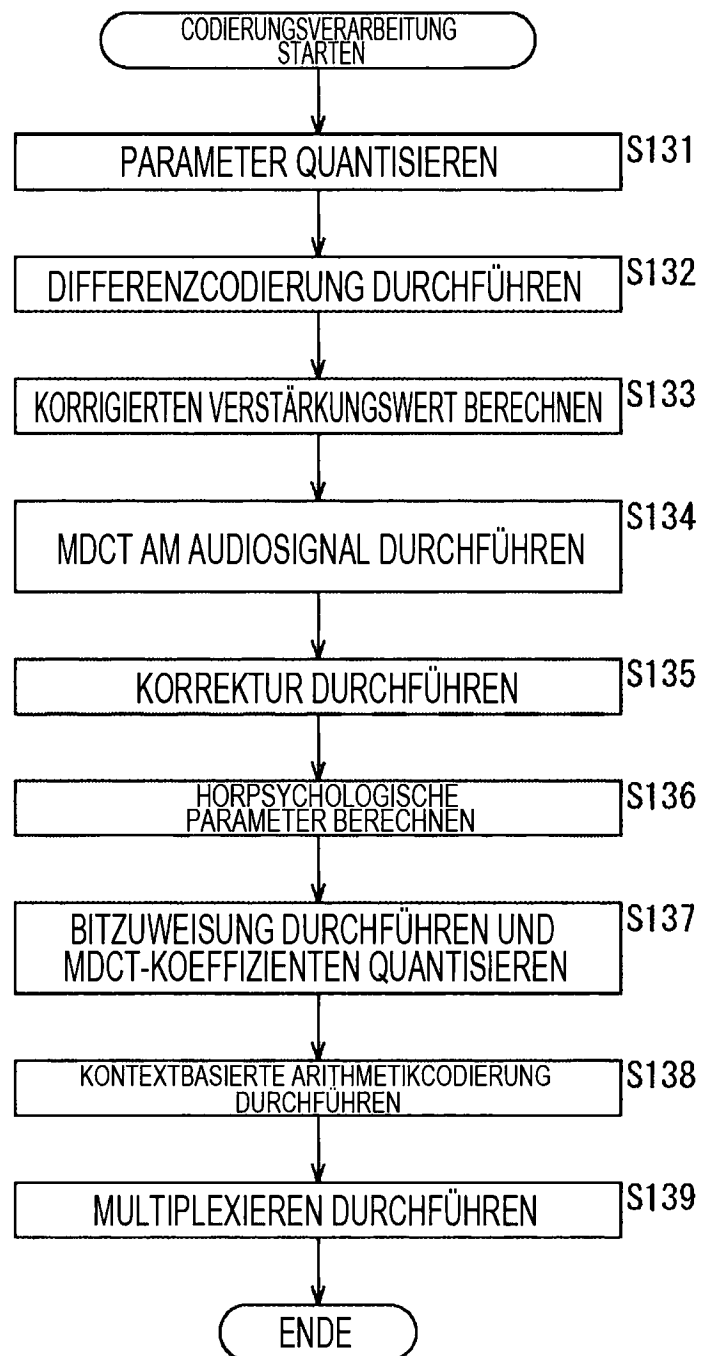


Fig. 21

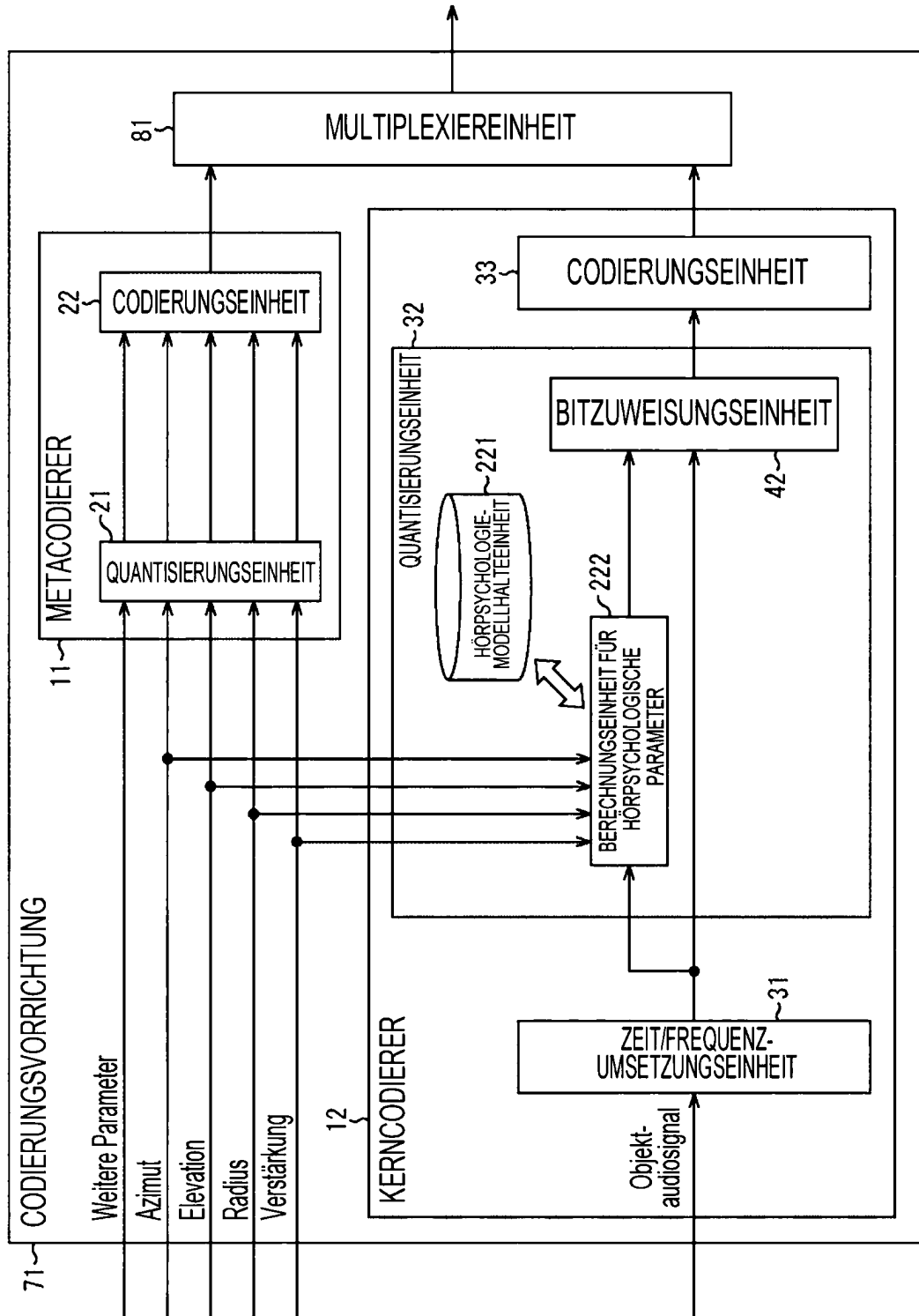


Fig. 22

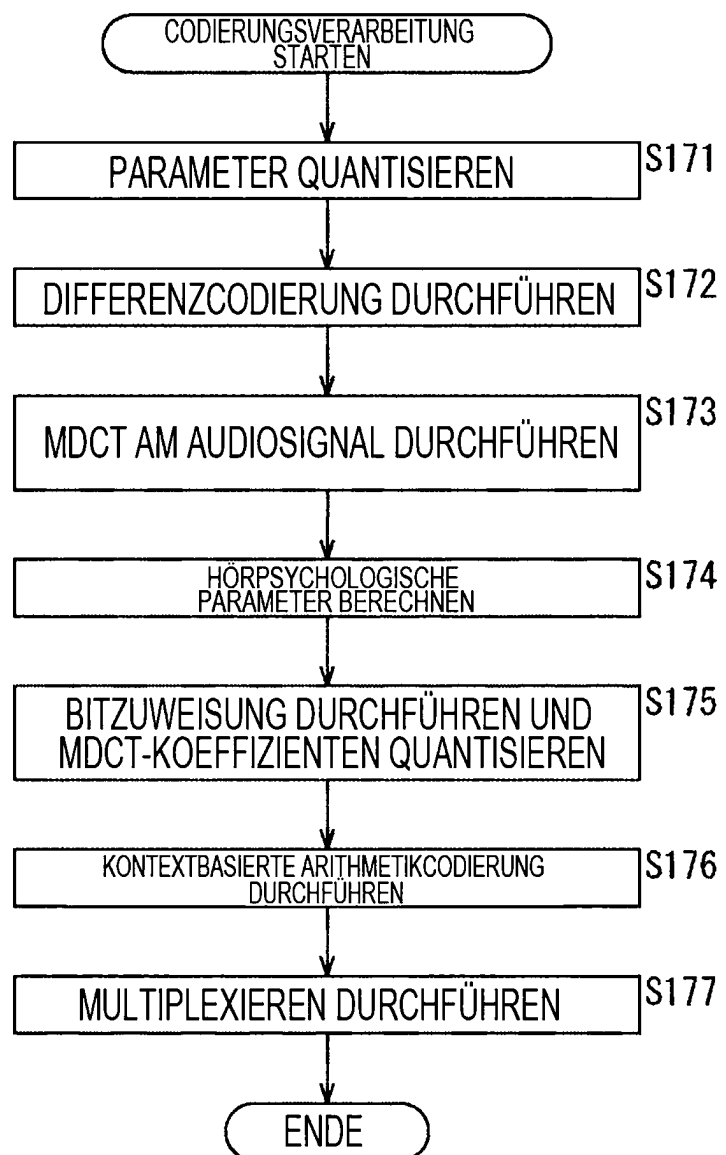


Fig. 23

Syntax	Bitanzahl	Gedächtnishilfe
ObjektMetadatenKonfig ()		
{		
geringeVerzögerungMetadatenCodierung :	1	bslbf
besitztKernLänge :	1	bslbf
falls(!besitztKernLänge) {		
RahmenLänge :	6	uimsbf
OAMRahmenLänge = (RahmenLänge+1)<<6:		
} sonst {		
OAMRahmenLänge = AusgabeRahmenLänge:		
}		
besitztBildschirmBezogeneObjekte :	1	bslbf
falls(besitztBildschirmBezogeneObjekte) {		
für (o = 0 : o < Anz_Objekte : o++) {		
istBildschirmBezogenesObjekt [o] :	1	bslbf
}		
}		
besitztDynamischeObjektPriorität :	1	bslbf
besitztGleichförmigeAusbreitung :	1	bslbf
für (o = 0 : o < Anz_Objekte : o++) {		
BearbeitungsGenehmigungsMerker [o] : ← Q11	1	bslbf
}		
}		

Fig. 24

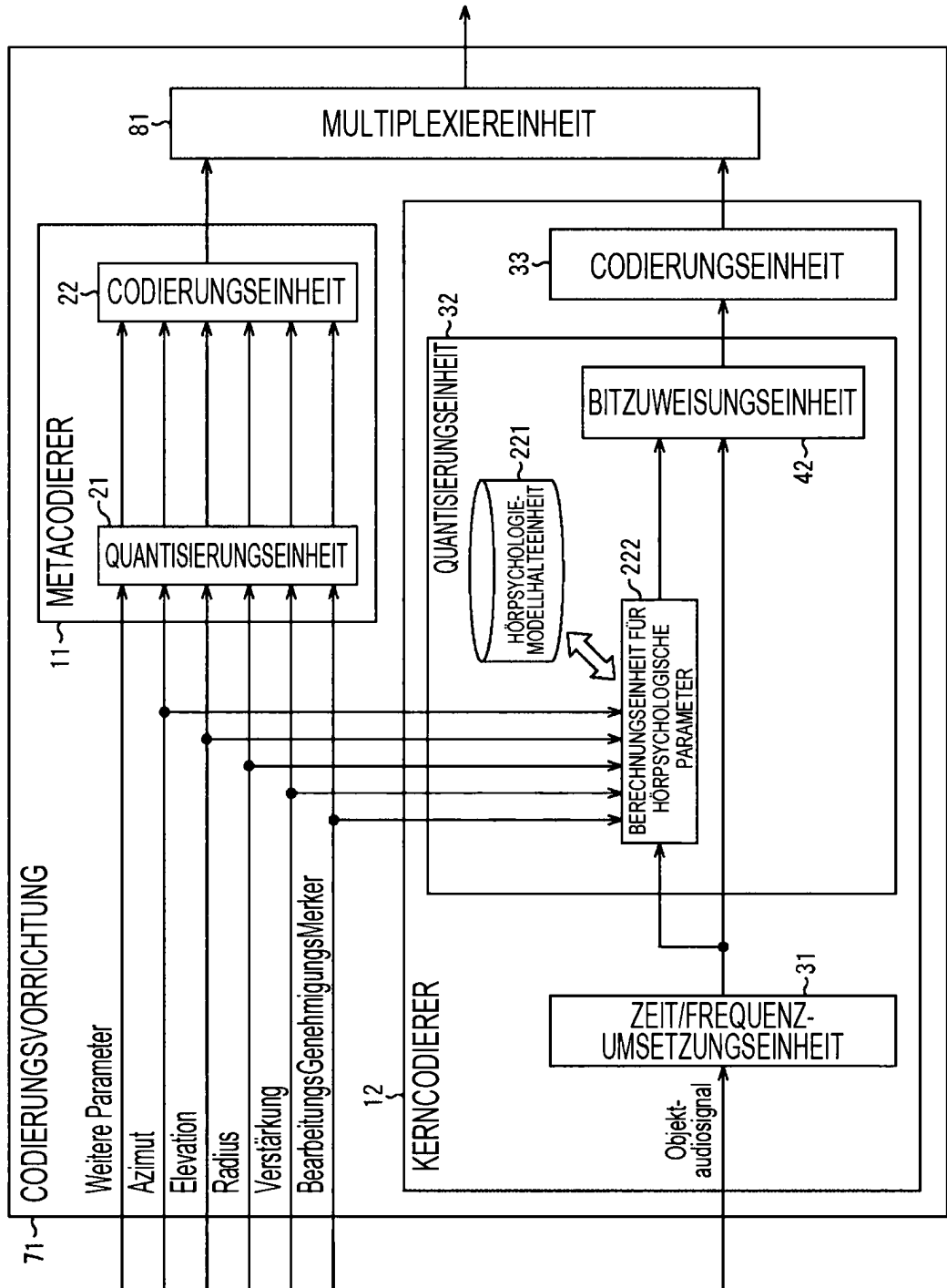


Fig. 25

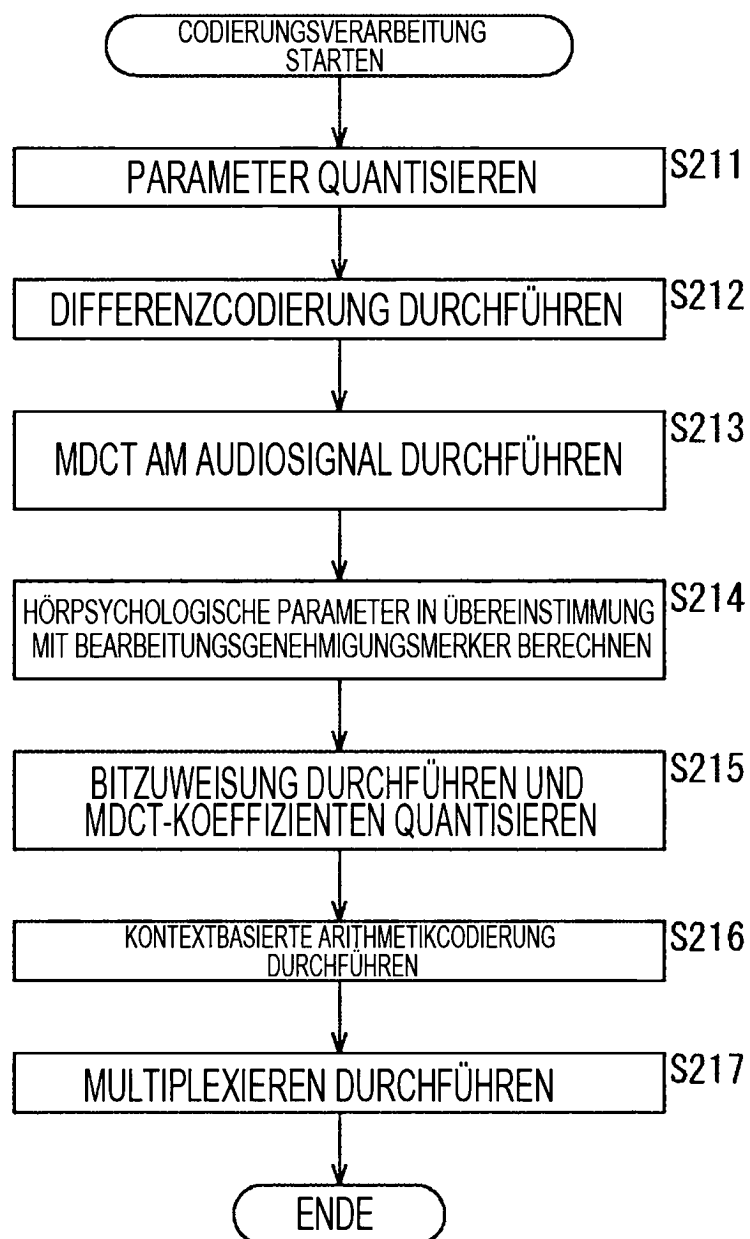


Fig. 26

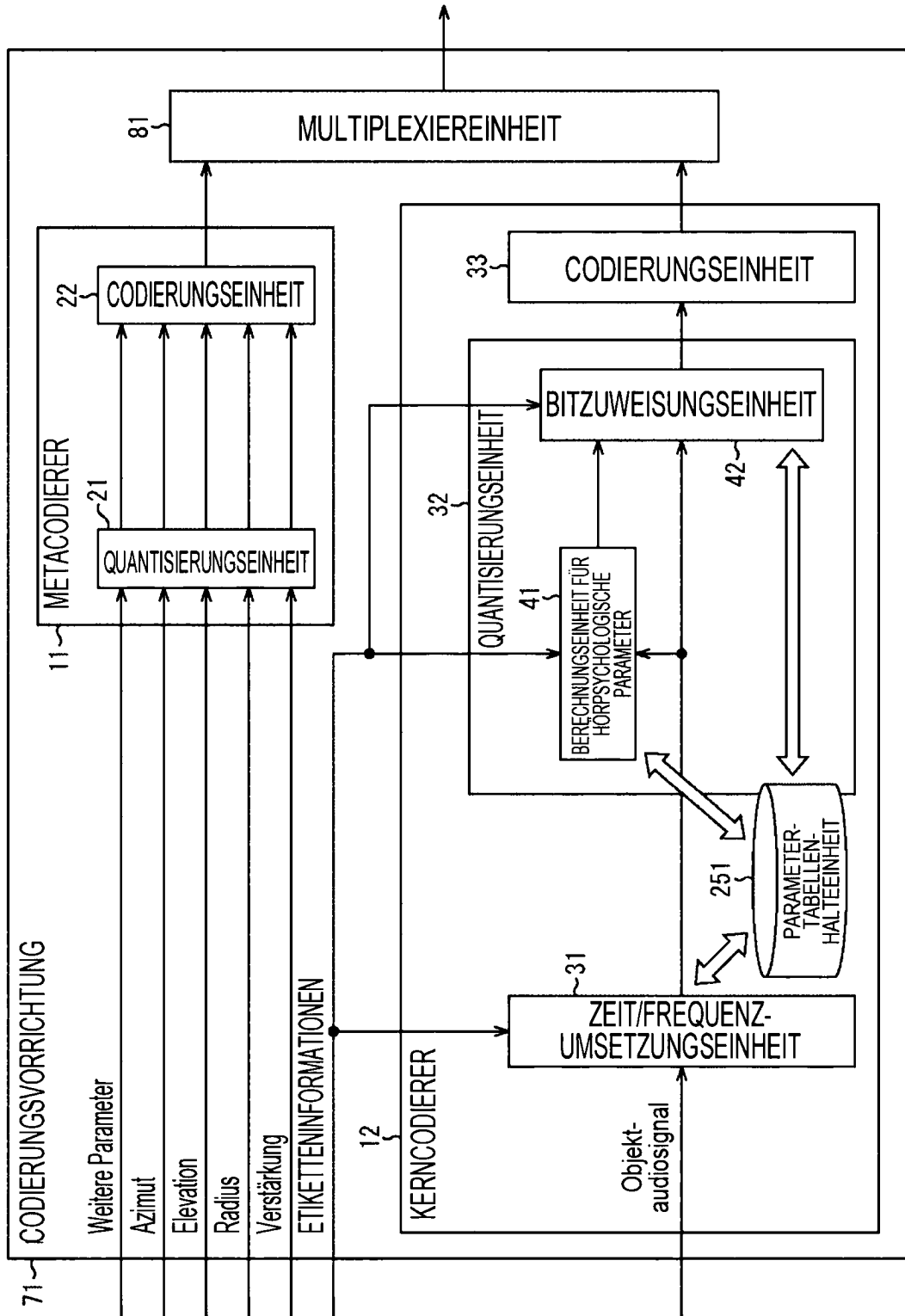


Fig. 27

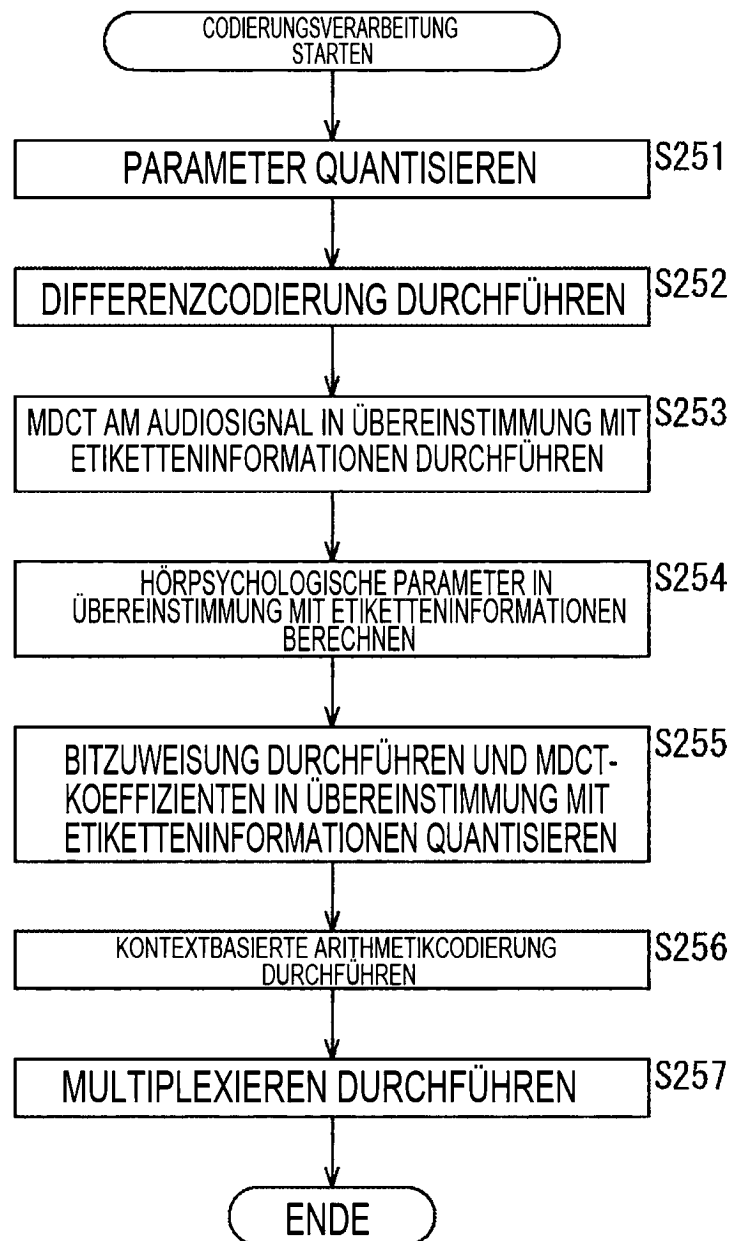


Fig. 28

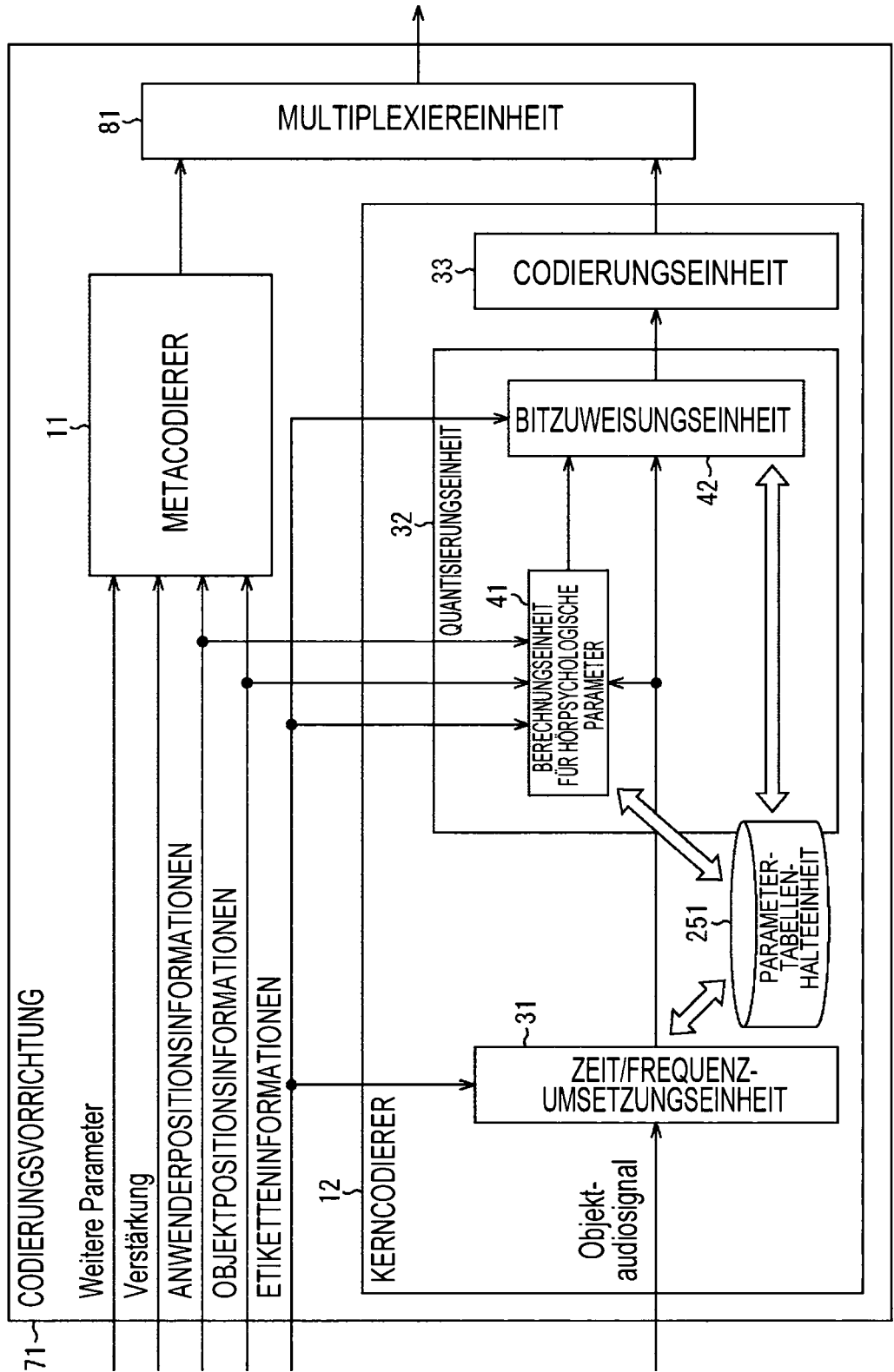


Fig. 29

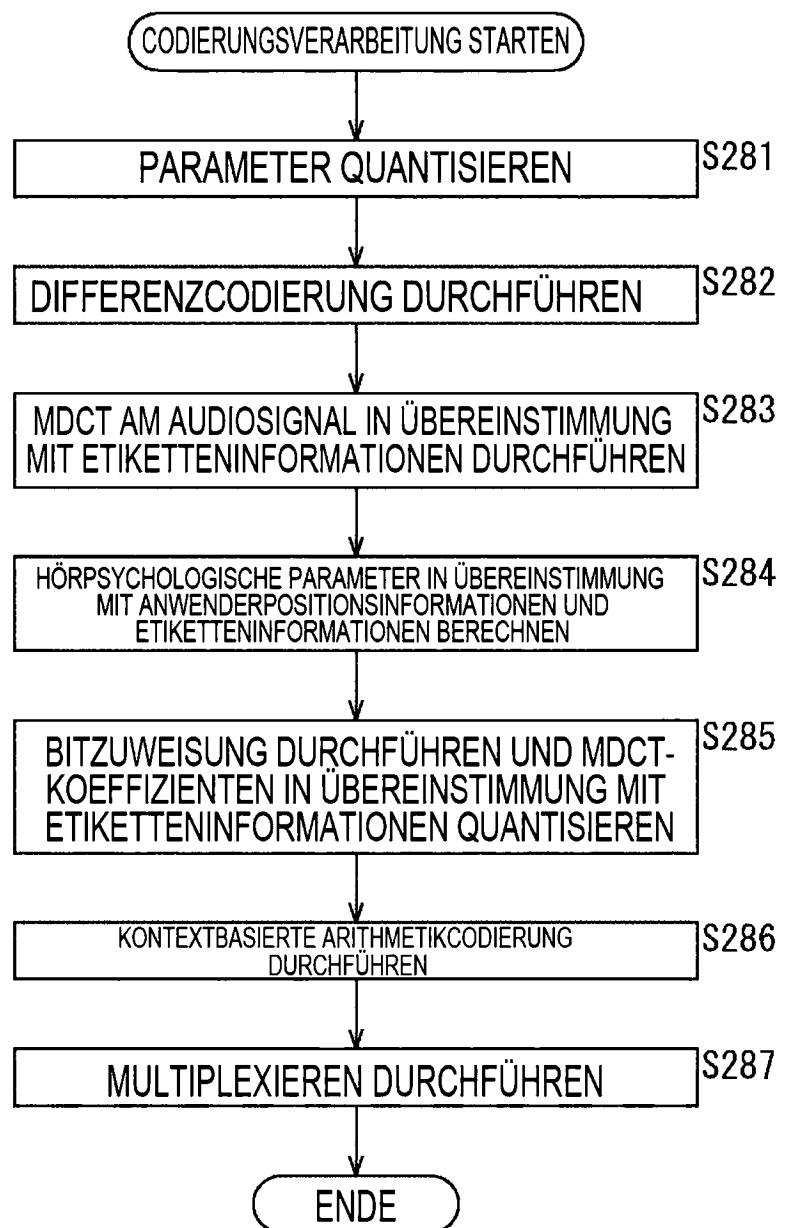


Fig. 30

