(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局

(43) 国際公開日
2022年1月13日(13.01.2022)

WIPO | PCT

(10) 国際公開番号
WO 2022/009694 A1

(51) 国際特許分類:
G10L 19/00 (2013.01) *G10L 19/035* (2013.01)
G10L 19/008 (2013.01)

(21) 国際出願番号： PCT/JP2021/024098

(22) 国際出願日： 2021年6月25日(25.06.2021)

(25) 国際出願の言語： 日本語

(26) 国際公開の言語： 日本語

(30) 優先権データ：
特願 2020-118174 2020年7月9日(09.07.2020) JP
特願 2020-170985 2020年10月9日(09.10.2020) JP

(71) 出願人：ソニーグループ株式会社(**SONY GROUP CORPORATION**) [JP/JP]；〒1080075 東京都港区港南１丁目７番１号 Tokyo (JP).

(72) 発明者：河野 明文(**KONO Akifumi**)；〒1080075 東京都港区港南１丁目７番１号 ソニーグループ株式会社内 Tokyo (JP). 知念 徹(**CHINEN Toru**)；〒1080075 東京都港区港南１丁目７番１号 ソニーグループ株式会社内 Tokyo (JP). 本間 弘幸(**HONMA Hiroyuki**)；〒1080075 東京都港区港南１丁目７番１号 ソニーグループ株式会社内 Tokyo (JP). 辻 実(**TSUJI Minoru**)；〒1080075 東京都港区港南１丁目７番１号 ソニーグループ株式会社内 Tokyo (JP). 及川 芳明(**OIKAWA Yoshiaki**)；〒1080075 東京都港区港南１丁目７番１号 ソニーグループ株式会社内 Tokyo (JP).

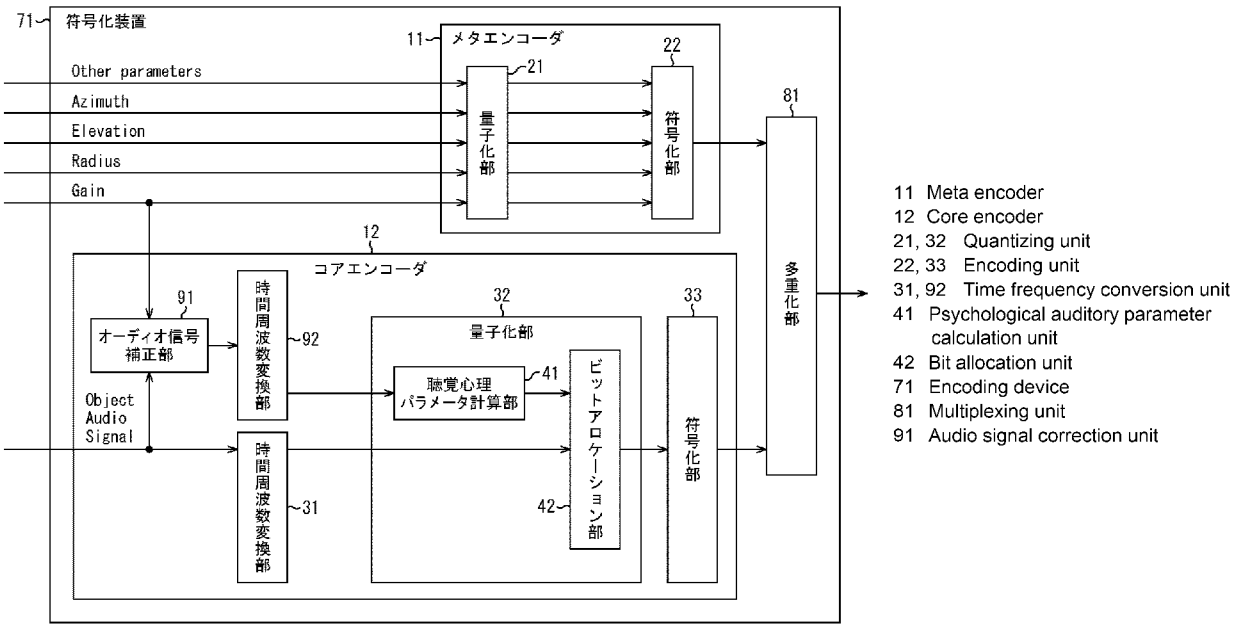
(74) 代理人：西川 孝，外 (**NISHIKAWA Takashi** et al.)；〒1700013 東京都豊島区東池袋３丁目９番１０号 池袋ＦＮビル４階 Tokyo (JP).

(81) 指定国(表示のない限り、全ての種類の国内保護が可能)：AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH,

(54) **Title:** SIGNAL PROCESSING DEVICE, METHOD, AND PROGRAM

(54) 発明の名称： 信号処理装置および方法、並びにプログラム

FIG. 4



(57) **Abstract:** This technology relates to a signal processing device, method, and program capable of improving encoding efficiency. This signal processing device comprises: a correction unit for correcting an audio signal of an audio object on the basis of a gain value included in metadata of the audio object; and a quantizing unit for calculating a psychological audio parameter on the basis of the corrected signal and quantizing the audio signal. This technology is applicable to an encoding device.

WO 2022/009694 A1

［続葉有］

CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) 指定国(表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, RU, TJ, TM), ヨーロッパ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

添付公開書類:

一　国際調査報告（条約第21条(3)）

(57) 要約：本技術は、符号化効率を向上させることができるようにする信号処理装置および方法、並びにプログラムに関する。 信号処理装置は、オーディオオブジェクトのメタデータに含まれるゲイン値に基づいて、オーディオオブジェクトのオーディオ信号に対する補正を行う補正部と、補正により得られた信号に基づいて聴覚心理パラメータを計算し、オーディオ信号を量子化する量子化部とを備える。本技術は符号化装置に適用することができる。

明　細　書

発明の名称 ： 信号処理装置および方法、並びにプログラム

技術分野

[0001] 本技術は、信号処理装置および方法、並びにプログラムに関し、特に、符号化効率を向上させることができるようにした信号処理装置および方法、並びにプログラムに関する。

背景技術

[0002] 従来、国際標準規格であるMPEG(Moving Picture Experts Group)-D USAC(Unified Speech and Audio Coding)規格や、MPEG-D USAC 規格をCore CoderにしたMPEG-H 3D Audio規格の符号化等が知られている（例えば、非特許文献１乃至非特許文献３参照）。

先行技術文献

非特許文献

[0003] 非特許文献1：ISO/IEC 23003-3, MPEG-D USAC

非特許文献2：ISO/IEC 23008-3, MPEG-H 3D Audio

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

発明の概要

発明が解決しようとする課題

[0004] MPEG-H 3D Audio規格等で扱われる3D Audioでは、音素材（オブジェクト）の位置を示す水平角度や垂直角度、距離、オブジェクトについてのゲインなどといったオブジェクトごとのメタデータを持ち、３次元的な音の方向や距離、拡がりなどを再現することができる。そのため、3D Audioでは、従来のステレオ再生に比べて、より臨場感のあるオーディオ再生が可能となる。

[0005] しかしながら、3D Audioで実現される多数のオブジェクトのデータを伝送するためには、より多くのオーディオチャネルを圧縮効率よく高速に復号可能な符号化技術が必要となる。すなわち、符号化効率の向上が望まれている。

[0006] 本技術は、このような状況に鑑みてなされたものであり、符号化効率を向上させることができるようにするものである。

課題を解決するための手段

[0007] 本技術の第１の側面の信号処理装置は、オーディオオブジェクトのメタデータに含まれるゲイン値に基づいて、前記オーディオオブジェクトのオーディオ信号に対する補正を行う補正部と、前記補正により得られた信号に基づいて聴覚心理パラメータを計算し、前記オーディオ信号を量子化する量子化部とを備える。

[0008] 本技術の第１の側面の信号処理方法またはプログラムは、オーディオオブジェクトのメタデータに含まれるゲイン値に基づいて、前記オーディオオブジェクトのオーディオ信号に対する補正を行い、前記補正により得られた信号に基づいて聴覚心理パラメータを計算し、前記オーディオ信号を量子化するステップを含む。

[0009] 本技術の第１の側面においては、オーディオオブジェクトのメタデータに含まれるゲイン値に基づいて、前記オーディオオブジェクトのオーディオ信号に対する補正が行われ、前記補正により得られた信号に基づいて聴覚心理パラメータが計算され、前記オーディオ信号が量子化される。

[0010] 本技術の第２の側面の信号処理装置は、オーディオオブジェクトのメタデータに含まれるゲイン値に基づいて、前記オーディオオブジェクトの前記ゲイン値およびオーディオ信号に対する修正を行う修正部と、前記修正により得られた修正オーディオ信号を量子化する量子化部とを備える。

[0011] 本技術の第２の側面の信号処理方法またはプログラムは、オーディオオブジェクトのメタデータに含まれるゲイン値に基づいて、前記オーディオオブジェクトの前記ゲイン値およびオーディオ信号に対する修正を行い、前記修正により得られた修正オーディオ信号を量子化するステップを含む。

[0012] 本技術の第２の側面においては、オーディオオブジェクトのメタデータに含まれるゲイン値に基づいて、前記オーディオオブジェクトの前記ゲイン値およびオーディオ信号に対する修正が行われ、前記修正により得られた修正

オーディオ信号が量子化される。

[0013] 本技術の第３の側面の信号処理装置は、オーディオオブジェクトのゲイン値と位置情報の少なくとも何れか一方を含むメタデータ、前記オーディオオブジェクトのオーディオ信号、および複数の前記オーディオオブジェクト間の聴覚マスキングに関する聴覚心理モデルに基づいて聴覚心理パラメータを計算し、前記聴覚心理パラメータに基づいて前記オーディオ信号を量子化する量子化部を備える。

[0014] 本技術の第３の側面の信号処理方法またはプログラムは、オーディオオブジェクトのゲイン値と位置情報の少なくとも何れか一方を含むメタデータ、前記オーディオオブジェクトのオーディオ信号、および複数の前記オーディオオブジェクト間の聴覚マスキングに関する聴覚心理モデルに基づいて聴覚心理パラメータを計算し、前記聴覚心理パラメータに基づいて前記オーディオ信号を量子化するステップを含む。

[0015] 本技術の第３の側面においては、オーディオオブジェクトのゲイン値と位置情報の少なくとも何れか一方を含むメタデータ、前記オーディオオブジェクトのオーディオ信号、および複数の前記オーディオオブジェクト間の聴覚マスキングに関する聴覚心理モデルに基づいて聴覚心理パラメータが計算され、前記聴覚心理パラメータに基づいて前記オーディオ信号が量子化される。

[0016] 本技術の第４の側面の信号処理装置は、オーディオオブジェクトのオーディオ信号、および前記オーディオオブジェクトの音源種別を示すラベル情報に基づいて、前記ラベル情報により示される前記音源種別に対して定められた調整パラメータとアルゴリズムの少なくとも一方により、前記オーディオ信号を量子化する量子化部を備える。

[0017] 本技術の第４の側面の信号処理方法またはプログラムは、オーディオオブジェクトのオーディオ信号、および前記オーディオオブジェクトの音源種別を示すラベル情報に基づいて、前記ラベル情報により示される前記音源種別に対して定められた調整パラメータとアルゴリズムの少なくとも一方により

、前記オーディオ信号を量子化するステップを含む。

[0018] 本技術の第４の側面においては、オーディオオブジェクトのオーディオ信号、および前記オーディオオブジェクトの音源種別を示すラベル情報に基づいて、前記ラベル情報により示される前記音源種別に対して定められた調整パラメータとアルゴリズムの少なくとも一方により、前記オーディオ信号が量子化される。

図面の簡単な説明

[0019] [図1]MPEG-H 3D Audioにおける符号化について説明する図である。

[図2]MPEG-H 3D Audioにおける符号化について説明する図である。

[図3]value rangeの例を示す図である。

[図4]符号化装置の構成例を示す図である。

[図5]符号化処理を説明するフローチャートである。

[図6]符号化装置の構成例を示す図である。

[図7]符号化処理を説明するフローチャートである。

[図8]符号化装置の構成例を示す図である。

[図9]ゲイン値の修正について説明する図である。

[図10]ゲイン値の修正に応じたオーディオ信号の修正について説明する図である。

[図11]ゲイン値の修正に応じたオーディオ信号の修正について説明する図である。

[図12]符号化処理を説明するフローチャートである。

[図13]ピンクノイズに対する聴感上の特性について説明する図である。

[図14]聴覚特性テーブルを用いたゲイン値の補正について説明する図である。

[図15]聴覚特性テーブルの例を示す図である。

[図16]聴覚特性テーブルの例を示す図である。

[図17]聴覚特性テーブルの例を示す図である。

[図18]ゲイン補正値の補間例を示す図である。

[図19]符号化装置の構成例を示す図である。

[図20]符号化処理を説明するフローチャートである。

[図21]符号化装置の構成例を示す図である。

[図22]符号化処理を説明するフローチャートである。

[図23]メタデータのConfigのシンタックス例を示す図である。

[図24]符号化装置の構成例を示す図である。

[図25]符号化処理を説明するフローチャートである。

[図26]符号化装置の構成例を示す図である。

[図27]符号化処理を説明するフローチャートである。

[図28]符号化装置の構成例を示す図である。

[図29]符号化処理を説明するフローチャートである。

[図30]コンピュータの構成例を示す図である。

発明を実施するための形態

[0020] 以下、図面を参照して、本技術を適用した実施の形態について説明する。

[0021] 〈第１の実施の形態〉

〈本技術について〉

本技術は、視聴時のレンダリングで適用されるメタデータのゲインを考慮し、実際の聴感に適した聴覚心理パラメータの計算とビットアロケーションを行うことで、符号化効率（圧縮効率）を向上させることができるようにするものである。

[0022] まず、MPEG-H 3D Audioにおけるオーディオオブジェクト（以下、単にオブジェクトと称する）のメタデータとオーディオ信号の符号化について説明する。

[0023] MPEG-H 3D Audioでは、図１に示すようにオブジェクトのメタデータはメタエンコーダによって符号化（エンコード）され、オブジェクトのオーディオ信号はコアエンコーダによって符号化される。

[0024] 具体的には、メタエンコーダは、メタデータを構成する各パラメータを量子化するとともに、その結果得られた量子化パラメータを符号化し、符号化

メタデータとする。

[0025] また、コアエンコーダは、オーディオ信号に対してMDCT（Modified Discrete Cosine Transform）（修正離散コサイン変換）を用いた時間周波数変換を行い、その結果得られたMDCT係数を量子化し、量子化MDCT係数を得る。MDCT係数の量子化時には、ビットアロケーションも行われる。さらに、コアエンコーダは、量子化MDCT係数を符号化し、符号化オーディオデータを得る。

[0026] そして、このようにして得られた符号化メタデータと符号化オーディオデータとが１つのビットストリームとしてまとめられ、出力される。

[0027] ここで、図２を参照して、MPEG-H 3D Audioでのメタデータとオーディオ信号の符号化について、より詳細に説明する。

[0028] この例では、メタデータとしての複数のパラメータがメタエンコーダ１１に入力され、オブジェクトの音を再生するための時間信号（波形信号）であるオーディオ信号がコアエンコーダ１２に入力される。

[0029] メタエンコーダ１１は、量子化部２１および符号化部２２を有しており、量子化部２１にメタデータが入力される。

[0030] メタエンコーダ１１でのメタデータの符号化の処理が開始されると、量子化部２１は、最初にメタデータとしての各パラメータの値に対して、必要に応じて上限値または下限値への置き換えを行い、その後、各パラメータを量子化し、量子化パラメータを得る。

[0031] この例では、メタデータを構成する各パラメータとして水平角度（Azimuth）、垂直角度（Elevation）、距離（Radius）、ゲイン値（Gain）、および他のパラメータ（Other parameters）が量子化部２１に入力される。

[0032] ここで、水平角度（Azimuth）および垂直角度（Elevation）は、３次元空間における基準となる聴取位置から見たオブジェクトの位置を示す水平方向および垂直方向の角度である。また、距離（Radius）は３次元空間におけるオブジェクトの位置を示す、基準となる聴取位置からオブジェクトまでの距離を示している。これらの水平角度、垂直角度、および距離からなる情報が、オブジェクトの位置を示す位置情報である。

[0033] さらに、ゲイン値（Gain）はオブジェクトのオーディオ信号のゲイン補正のためのゲインであり、他のパラメータ（Other parameters）は音像を広げるスプレッド処理のためのパラメータや、オブジェクトの優先度などである。

[0034] メタデータを構成する各パラメータは、図３に示す予め定められた所定の範囲であるvalue range内の値となるようにされる。

[0035] 図３の例では、メタデータを構成する各パラメータのvalue rangeが示されている。

[0036] なお、図３において「spread」、「spread width」、「spread height」、および「spread depth」は、スプレッド処理のためのパラメータであり、他のパラメータ（Other parameters）の例である。また、「dynamic object priority」はオブジェクトの優先度を示すパラメータであり、このパラメータも他のパラメータ（Other parameters）の例である。

[0037] 例えば、この例では水平角度（Azimuth）のvalue rangeは、下限値である-180度から上限値である180度までとなっている。

[0038] 量子化部２１に入力された水平角度が、このvalue rangeの範囲を超えている場合、つまり範囲外である場合には、その水平角度が下限値「-180」または上限値「180」に置き換えられてから量子化が行われる。すなわち、入力された水平角度が上限値よりも大きい値であれば、上限値「180」が制限（置き換え）後の水平角度とされ、水平角度が下限値よりも小さい値であれば、下限値「-180」が制限後の水平角度とされる。

[0039] また、例えばゲイン値（Gain）のvalue rangeは、下限値である0.004から上限値である5.957となっている。特に、ここではゲイン値はリニア値で記述されている。

[0040] 図２の説明に戻り、量子化部２１でメタデータを構成する各パラメータが量子化され、量子化パラメータが得られると、その量子化パラメータが符号化部２２により符号化され、その結果得られた符号化メタデータが出力される。例えば符号化部２２では、量子化パラメータに対して差分符号化が行わ

れ、符号化メタデータが生成される。

[0041] また、コアエンコーダ１２は、時間周波数変換部３１、量子化部３２、および符号化部３３を有しており、時間周波数変換部３１にオブジェクトのオーディオ信号が入力される。また、量子化部３２は、聴覚心理パラメータ計算部４１およびビットアロケーション部４２を有している。

[0042] コアエンコーダ１２では、オーディオ信号の符号化の処理が開始されると、最初に時間周波数変換部３１が、入力されたオーディオ信号に対してMDCT、すなわち時間周波数変換を行い、その結果として周波数スペクトル情報であるMDCT係数が得られる。

[0043] 次に、量子化部３２において、時間周波数変換（MDCT）により得られたMDCT係数がスケールファクターバンドごとに量子化され、その結果として量子化MDCT係数が得られる。

[0044] ここで、スケールファクターバンドとは、QMF（Quadrature Mirror Filter）分析フィルタの分解能である所定帯域幅のサブバンドを複数束ねて得られる帯域（周波数帯域）である。

[0045] 具体的には、量子化部３２での量子化では、聴覚心理パラメータ計算部４１により、MDCT係数に対して人間の聴覚特性（聴覚マスキング）を考慮するための聴覚心理パラメータが計算される。

[0046] また、ビットアロケーション部４２において、時間周波数変換により得られたMDCT係数と、聴覚心理パラメータ計算部４１で得られた聴覚心理パラメータとが用いられて、各スケールファクターバンドの量子化ビットと量子化ノイズの計算および評価を行う、聴覚心理モデルに基づいたビットアロケーションが行われる。

[0047] そしてビットアロケーション部４２は、そのビットアロケーションの結果に基づいてMDCT係数をスケールファクターバンドごとに量子化し、その結果得られた量子化MDCT係数を符号化部３３に供給する。

[0048] このようにすることで、MDCT係数の量子化で発生してしまう量子化ノイズがマスクされて知覚されないスケールファクターバンドの量子化ビットの一

部が、量子化ノイズが知覚されやすいスケールファクターバンドに割り当てられる（回される）。これにより、全体として音質の劣化を抑制し、効率的な量子化を行うことができる。すなわち、符号化効率を向上させることができる。

[0049] さらに、符号化部３３では、ビットアロケーション部４２から供給された量子化MDCT係数に対して、例えばコンテキストベースの算術符号化が行われ、その結果得られた符号化オーディオデータが、オーディオ信号の符号化データとして出力される。

[0050] 以上のようにして、オブジェクトのメタデータとオーディオ信号が、メタエンコーダ１１およびコアエンコーダ１２により符号化される。

[0051] ところで、聴覚心理パラメータの計算に使用されるMDCT係数は、入力されたオーディオ信号にMDCT、すなわち時間周波数変換を行って得られたものである。

[0052] しかしながら、実際に符号化されたオーディオ信号が復号（デコード）およびレンダリングされて視聴される際には、メタデータのゲイン値が適用されるため、聴覚心理パラメータ計算時と視聴時とで使用されるオーディオ信号に乖離が生じてしまう。

[0053] そのため、所定のスケールファクターバンドに対して、本来であれば聴感上聞こえることのない量子化ノイズの発生を防ぐために余分に多くのビットが使用されてしまうなど、符号化効率が低下してしまう場合がある。

[0054] そこで本技術では、メタデータのゲイン値を適用した補正MDCT係数を用いて聴覚心理パラメータを計算することで、より実際の聴感に即した聴覚心理パラメータを得て符号化効率を向上させることができるようにした。

[0055] 〈符号化装置の構成例〉

図４は、本技術を適用した符号化装置の一実施の形態の構成例を示す図である。なお、図４において図２における場合と対応する部分には同一の符号を付してあり、その説明は適宜省略する。

[0056] 図４に示す符号化装置７１は、例えばオーディオオブジェクトのコンテン

ツを配信するサーバ等の信号処理装置により実現され、メタエンコーダ１１、コアエンコーダ１２、および多重化部８１を有している。

[0057] また、メタエンコーダ１１は、量子化部２１および符号化部２２を有しており、コアエンコーダ１２は、オーディオ信号補正部９１、時間周波数変換部９２、時間周波数変換部３１、量子化部３２、および符号化部３３を有している。

[0058] さらに量子化部３２は、聴覚心理パラメータ計算部４１およびビットアロケーション部４２を有している。

[0059] 符号化装置７１の構成は、図２に示した構成に対して新たに多重化部８１、オーディオ信号補正部９１、および時間周波数変換部９２を設けた構成となっており、その他の点では図２に示した構成と同じとなっている。

[0060] 図４の例では、多重化部８１は、符号化部２２から供給された符号化メタデータと、符号化部３３から供給された符号化オーディオデータとを多重化してビットストリームを生成し、出力する。

[0061] また、オーディオ信号補正部９１にはオブジェクトのオーディオ信号と、メタデータを構成するオブジェクトのゲイン値とが供給される。

[0062] オーディオ信号補正部９１は、供給されたゲイン値に基づいて、供給されたオーディオ信号に対するゲイン補正を行い、ゲイン補正後のオーディオ信号を時間周波数変換部９２に供給する。例えばオーディオ信号補正部９１では、ゲイン値がオーディオ信号に乗算されて、オーディオ信号がゲイン補正される。すなわち、ここでは時間領域でオーディオ信号に対する補正が行われる。

[0063] 時間周波数変換部９２は、オーディオ信号補正部９１から供給されたオーディオ信号に対してMDCTを行い、その結果得られたMDCT係数を聴覚心理パラメータ計算部４１に供給する。

[0064] なお、以下では、オーディオ信号補正部９１でのゲイン補正により得られたオーディオ信号を、特に補正オーディオ信号とも称し、時間周波数変換部９２でのMDCTで得られたMDCT係数を、特に補正MDCT係数とも称することとす

る。

[0065] また、この例では時間周波数変換部３１で得られたMDCT係数は、聴覚心理パラメータ計算部４１には供給されず、聴覚心理パラメータ計算部４１では、時間周波数変換部９２から供給された補正MDCT係数に基づいて聴覚心理パラメータが計算される。

[0066] 符号化装置７１では、冒頭にあるオーディオ信号補正部９１において、入力されたオブジェクトのオーディオ信号に対して、レンダリング時と同様にメタデータに含まれるゲイン値を適用するゲイン補正が行われる。

[0067] その後、ゲイン補正により得られた補正オーディオ信号に対して、ビットアロケーション用とは別に時間周波数変換部９２によりMDCTが行われ、補正MDCT係数が求められる。

[0068] そして、最後に、補正MDCT係数に基づいて、聴覚心理パラメータ計算部４１により聴覚心理パラメータを計算することで、図２における場合よりも実際の聴感に即した聴覚心理パラメータが得られる。

[0069] これは、補正オーディオ信号に基づく音は、もとのオーディオ信号に基づく音よりも、復号側でのレンダリングにより得られる信号に基づく音により近いものとなるからである。このようにすることで、各スケールファクターバンドに対して、より適切に量子化ビットが割り当てられるようになり、符号化効率を向上させることができる。

[0070] なお、ここではオーディオ信号補正部９１でのゲイン補正に量子化前のメタデータのゲイン値を用いる例について説明したが、符号化後や量子化後のゲイン値をオーディオ信号補正部９１に供給し、ゲイン補正に利用するようにしてもよい。

[0071] そのような場合、オーディオ信号補正部９１において、符号化後や量子化後のゲイン値が復号されたり逆量子化されたりして、その結果得られたゲイン値に基づいてオーディオ信号のゲイン補正が行われ、補正オーディオ信号とされる。

[0072] 〈符号化処理の説明〉

次に、図４に示した符号化装置７１の動作について説明する。すなわち、以下、図５のフローチャートを参照して、符号化装置７１による符号化処理について説明する。

[0073] ステップＳ１１において量子化部２１は、供給されたメタデータとしての各パラメータを量子化し、その結果得られた量子化パラメータを符号化部２２に供給する。

[0074] このとき、量子化部２１は、予め定められたvalue rangeよりも大きいパラメータについてはそのvalue rangeの上限値への置き換えを行ってから量子化を行い、同様にvalue rangeよりも小さいパラメータについては下限値への置き換えを行ってから量子化を行う。

[0075] ステップＳ１２において符号化部２２は、量子化部２１から供給された量子化パラメータに対して差分符号化を行い、その結果得られた符号化メタデータを多重化部８１に供給する。

[0076] ステップＳ１３においてオーディオ信号補正部９１は、供給されたオブジェクトのオーディオ信号に対して、供給されたメタデータのゲイン値に基づくゲイン補正を行い、その結果得られた補正オーディオ信号を時間周波数変換部９２に供給する。

[0077] ステップＳ１４において時間周波数変換部９２は、オーディオ信号補正部９１から供給された補正オーディオ信号に対するMDCT（時間周波数変換）を行い、その結果得られた補正MDCT係数を聴覚心理パラメータ計算部４１に供給する。

[0078] ステップＳ１５において時間周波数変換部３１は、供給されたオブジェクトのオーディオ信号に対するMDCT（時間周波数変換）を行い、その結果得られたMDCT係数をビットアロケーション部４２に供給する。

[0079] ステップＳ１６において聴覚心理パラメータ計算部４１は、時間周波数変換部９２から供給された補正MDCT係数に基づいて聴覚心理パラメータを計算し、ビットアロケーション部４２に供給する。

[0080] ステップＳ１７においてビットアロケーション部４２は、聴覚心理パラメ

ータ計算部４１から供給された聴覚心理パラメータと、時間周波数変換部３１から供給されたMDCT係数とに基づいて聴覚心理モデルに基づくビットアロケーションを行い、その結果に基づきスケールファクターバンドごとにMDCT係数を量子化する。ビットアロケーション部４２は、量子化により得られた量子化MDCT係数を符号化部３３に供給する。

[0081] ステップＳ１８において符号化部３３は、ビットアロケーション部４２から供給された量子化MDCT係数に対して、コンテキストベースの算術符号化を行い、その結果得られた符号化オーディオデータを多重化部８１に供給する。

[0082] ステップＳ１９において多重化部８１は、符号化部２２から供給された符号化メタデータと、符号化部３３から供給された符号化オーディオデータとを多重化してビットストリームを生成し、出力する。

[0083] このようにしてビットストリームが出力されると、符号化処理は終了する。

[0084] 以上のようにして符号化装置７１は、符号化前のメタデータのゲイン値に基づいてオーディオ信号を補正し、その結果得られた補正オーディオ信号に基づいて聴覚心理パラメータを計算する。このようにすることで、より実際の聴感に即した聴覚心理パラメータを得ることができ、符号化効率を向上させることができる。

[0085] 〈第２の実施の形態〉

〈符号化装置の構成例〉

ところで、図４に示した符号化装置７１では、MDCTを２回行う必要があるため、演算負荷（演算量）が大きくなってしまう。そこで、周波数領域でMDCT係数（オーディオ信号）の補正を行うことで、演算量を低減させるようにしてもよい。

[0086] そのような場合、符号化装置７１は、例えば図６に示すように構成される。なお、図６において図４における場合と対応する部分には同一の符号を付してあり、その説明は適宜省略する。

[0087] 図６に示す符号化装置７１は、メタエンコーダ１１、コアエンコーダ１２、および多重化部８１を有している。

[0088] また、メタエンコーダ１１は、量子化部２１および符号化部２２を有しており、コアエンコーダ１２は、時間周波数変換部３１、MDCT係数補正部１３１、量子化部３２、および符号化部３３を有している。さらに量子化部３２は、聴覚心理パラメータ計算部４１およびビットアロケーション部４２を有している。

[0089] 図６に示す符号化装置７１の構成は、時間周波数変換部９２とオーディオ信号補正部９１に代えて、MDCT係数補正部１３１を設けた点で図４の符号化装置７１の構成と異なり、その他の点では図４の符号化装置７１と同じ構成となっている。

[0090] この例では、まず、オブジェクトのオーディオ信号に対して時間周波数変換部３１によりMDCTが行われ、その結果得られたMDCT係数がMDCT係数補正部１３１およびビットアロケーション部４２に供給される。

[0091] そして、MDCT係数補正部１３１では、時間周波数変換部３１から供給されたMDCT係数に対して、供給されたメタデータのゲイン値に基づいて補正が行われ、その結果得られた補正MDCT係数が聴覚心理パラメータ計算部４１へと供給される。

[0092] 例えばMDCT係数補正部１３１では、MDCT係数にゲイン値が乗算されて、MDCT係数が補正される。これにより、オーディオ信号が周波数領域でゲイン補正されたことになる。

[0093] このように周波数領域でゲイン補正を行う場合、時間領域で実際のレンダリングと同様にメタデータのゲイン値によるゲイン補正を行う第１の実施の形態の場合と比較して、ゲイン補正の再現度が多少低下する。すなわち、補正MDCT係数は、第１の実施の形態における場合ほど正確ではない。

[0094] しかし、補正MDCT係数に基づいて、聴覚心理パラメータ計算部４１により聴覚心理パラメータを計算することで、図２における場合と略同じ演算量で、図２における場合よりも実際の聴感に即した聴覚心理パラメータを得るこ

とができる。これにより、演算負荷を低く抑えつつ符号化効率を向上させることができる。

[0095] なお、図６では、MDCT係数の補正に量子化前のメタデータのゲイン値を用いる例について説明したが、符号化後や量子化後のゲイン値を利用するようにしてもよい。

[0096] そのような場合、MDCT係数補正部１３１において、符号化後や量子化後のゲイン値が復号されたり逆量子化されたりして、その結果得られたゲイン値に基づいてMDCT係数の補正が行われ、補正MDCT係数とされる。

[0097] 〈符号化処理の説明〉

次に、図６に示した符号化装置７１の動作について説明する。すなわち、以下、図７のフローチャートを参照して、図６の符号化装置７１による符号化処理について説明する。

[0098] なお、ステップＳ５１およびステップＳ５２の処理は、図５のステップＳ１１およびステップＳ１２の処理と同様であるので、その説明は省略する。

[0099] ステップＳ５３において時間周波数変換部３１は、供給されたオブジェクトのオーディオ信号に対するMDCTを行い、その結果得られたMDCT係数をMDCT係数補正部１３１およびビットアロケーション部４２に供給する。

[0100] ステップＳ５４においてMDCT係数補正部１３１は、時間周波数変換部３１から供給されたMDCT係数に対して、供給されたメタデータのゲイン値に基づいて補正を行い、その結果得られた補正MDCT係数を聴覚心理パラメータ計算部４１に供給する。

[0101] このようにして補正MDCT係数が得られると、その後、ステップＳ５５乃至ステップＳ５８の処理が行われて符号化処理は終了するが、これらの処理は図５のステップＳ１６乃至ステップＳ１９の処理と同様であるので、その説明は省略する。但し、ステップＳ５５では、聴覚心理パラメータ計算部４１は、MDCT係数補正部１３１から供給された補正MDCT係数に基づいて聴覚心理パラメータを計算する。

[0102] 以上のようにして符号化装置７１は、周波数領域でオーディオ信号（MDCT

係数）の補正を行い、得られた補正MDCT係数に基づいて聴覚心理パラメータを計算する。

[0103] このようにすることで、少ない演算量でも、より実際の聴感に即した聴覚心理パラメータを得ることができ、符号化効率を向上させることができる。

[0104] 〈第３の実施の形態〉

〈符号化装置の構成例〉

ところで、実際の3D Audioのコンテンツにおいては、符号化前のメタデータのゲイン値が必ずしもMPEG-Hの仕様範囲内の値であるとは限らない。

[0105] すなわち、例えばコンテンツ制作時に、波形レベルが極端に小さいオブジェクトの音量を他のオブジェクトの音量と揃えるために、メタデータのゲイン値が5.957（≒15.50dB）よりも大きい値とされることが考えられる。逆に、不要な音に対してメタデータのゲイン値が0.004（≒-49.76dB）より小さい値とされることもあり得る。

[0106] このようなコンテンツをMPEG-Hフォーマットで符号化および復号する場合、メタデータのゲイン値が図３に示したvalue rangeの上限値または下限値に制限されると、再生時に実際に聴取される音がコンテンツ制作者の意図とは異なるものとなってしまう。

[0107] そこで、メタデータのゲイン値がMPEG-Hの仕様の範囲外である場合に、MPEG-Hの仕様に準拠するようにメタデータのゲイン値とオーディオ信号を修正する前処理を行うことで、よりコンテンツ制作者の意図に近い音が再生されるようにしてもよい。

[0108] そのような場合、符号化装置７１は、例えば図８に示すように構成される。なお、図８において図６における場合と対応する部分には同一の符号を付してあり、その説明は適宜省略する。

[0109] 図８に示す符号化装置７１は、修正部１６１、メタエンコーダ１１、コアエンコーダ１２、および多重化部８１を有している。

[0110] また、メタエンコーダ１１は、量子化部２１および符号化部２２を有しており、コアエンコーダ１２は、時間周波数変換部３１、MDCT係数補正部１３

１、量子化部３２、および符号化部３３を有している。さらに量子化部３２は、聴覚心理パラメータ計算部４１およびビットアロケーション部４２を有している。

[0111] 図８に示す符号化装置７１の構成は、新たに修正部１６１が設けられている点で図６の符号化装置７１の構成と異なり、その他の点では図６の符号化装置７１と同じ構成となっている。

[0112] 図８に示す例では、コンテンツを構成するオブジェクトのメタデータおよびオーディオ信号が修正部１６１に供給される。

[0113] 修正部１６１は、符号化前に、供給されたメタデータのゲイン値のなかに、MPEG-Hの仕様範囲外、つまり上述のvalue range外のゲイン値があるかのチェック（確認）を行う。

[0114] そして、修正部１６１は、value range外のゲイン値がある場合には、そのゲイン値と、ゲイン値に対応するオーディオ信号に対して、MPEG-H 仕様に基づくゲイン値とオーディオ信号の修正処理を前処理として行う。

[0115] 具体的には、修正部１６１はvalue range（MPEG-Hの仕様範囲）外のゲイン値を、そのvalue rangeの上限値または下限値に修正し、修正ゲイン値とする。

[0116] すなわち、ゲイン値がvalue rangeの上限値よりも大きい場合には、その上限値が修正後のゲイン値である修正ゲイン値とされ、ゲイン値がvalue rangeの下限値よりも小さい場合には、その下限値が修正ゲイン値とされる。

[0117] なお、修正部１６１では、メタデータとしての複数のパラメータのうち、ゲイン値以外のパラメータについての修正（変更）は行われない。

[0118] また、修正部１６１は修正前のゲイン値と修正ゲイン値とに基づいて、供給されたオブジェクトのオーディオ信号に対するゲイン補正を行い、修正オーディオ信号とする。すなわち、修正前のゲイン値と修正ゲイン値の差分に基づき、オーディオ信号に対する修正（ゲイン補正）が行われる。

[0119] このとき、修正前のメタデータ（ゲイン値）およびオーディオ信号に基づくレンダリングの出力と、修正後のメタデータ（修正ゲイン値）および修正

オーディオ信号に基づくレンダリングの出力とが等しくなるようにオーディオ信号に対するゲイン補正が行われる。

[0120] 修正部１６１は、以上のゲイン値とオーディオ信号の修正を前処理として行い、必要に応じて修正されたゲイン値と、供給されたメタデータのゲイン値以外の他のパラメータとからなるデータを修正後のメタデータとして量子化部２１に供給するとともに、必要に応じて修正されたゲイン値をMDCT係数補正部１３１に供給する。

[0121] また、修正部１６１は、必要に応じて修正されたオーディオ信号を時間周波数変換部３１に供給する。

[0122] なお、以下では、説明を簡単にするため、修正が行われたか否かによらず、修正部１６１から出力されるメタデータおよびゲイン値を修正メタデータおよび修正ゲイン値とも称することとする。同様に、修正部１６１から出力されるオーディオ信号を修正オーディオ信号とも称することとする。

[0123] したがって、この例では修正メタデータがメタエンコーダ１１の入力となり、修正オーディオ信号および修正ゲイン値がコアエンコーダ１２の入力となる。

[0124] このようにすることで、実質的にMPEG-Hの仕様によってゲイン値が制限されることがなくなるため、コンテンツ制作者の意図通りのレンダリング結果を得ることができる。

[0125] メタエンコーダ１１やコアエンコーダ１２では、修正メタデータや修正オーディオ信号を入力として、図６に示した例と同様の処理が行われる。

[0126] すなわち、例えばコアエンコーダ１２では、時間周波数変換部３１により修正オーディオ信号に対してMDCTが行われ、その結果得られたMDCT係数がMDCT係数補正部１３１およびビットアロケーション部４２に供給される。

[0127] また、MDCT係数補正部１３１では、時間周波数変換部３１から供給されたMDCT係数に対して、修正部１６１から供給された修正ゲイン値に基づいて補正が行われ、その結果得られた補正MDCT係数が聴覚心理パラメータ計算部４１へと供給される。

[0128] なお、ここでは周波数領域でMDCT係数の補正が行われる例について説明したが、第１の実施の形態と同様に、時間領域において修正ゲイン値により修正オーディオ信号のゲイン補正が行われ、その後、MDCTにより補正MDCT係数が求められてもよい。

[0129] ここで、図９乃至図１１を参照して、ゲイン値とオーディオ信号の修正の具体的な例について説明する。

[0130] 図９は、所定のオブジェクトのメタデータのフレームごとのゲイン値を示している。なお、図９において横軸はフレームを示しており、縦軸はゲイン値を示している。

[0131] 特に、この例では折れ線L11は修正前における各フレームでのゲイン値を示しており、折れ線L12は各フレームでの修正後のゲイン値、つまり修正ゲイン値を示している。

[0132] また、直線L13はMPEG-Hの仕様範囲、つまり上述のvalue rangeの下限値（0.004（≒-49.76dB））を示しており、直線L14はMPEG-Hの仕様範囲の上限値（5.957（≒15.50dB））を示している。

[0133] ここでは、例えばフレーム「２」における修正前のゲイン値は、直線L13により示される下限値よりも小さい値であるので、ゲイン値が下限値へと置き換えられて修正ゲイン値とされている。また、例えばフレーム「４」における修正前のゲイン値は、直線L14により示される上限値よりも大きい値であるので、ゲイン値が上限値へと置き換えられて修正ゲイン値とされている。

[0134] このように、適宜、ゲイン値の修正を行うことで、各フレームにおける修正ゲイン値はMPEG-Hの仕様範囲（value range）内の値となっている。

[0135] また、図１０は修正部１６１での修正前のオーディオ信号を示しており、図１１は図１０に示したオーディオ信号を修正することで得られた修正オーディオ信号を示している。なお、図１０および図１１において横軸は時間を示しており、縦軸は信号レベルを示している。

[0136] 図１０に示されるように、修正前のオーディオ信号の信号レベルは時間によらず一定のレベルとなっている。

[0137] このようなオーディオ信号に対して修正部１６１でゲイン値および修正ゲイン値に基づくゲイン補正を行うと、図１１に示すように各時刻で信号レベルが異なる、つまり信号レベルが一定ではない修正オーディオ信号が得られる。

[0138] 特に図１１では、修正オーディオ信号における、メタデータのゲイン値が修正によって小さくなった、つまり上限値に置き換えられた影響が生じるサンプルでは、修正前よりも信号レベルが大きくなっていることが分かる。

[0139] これは、修正の前後でレンダリングの出力が同じとなるようにするためには、ゲイン値が小さくなった分だけ、オーディオ信号を大きくしておく必要があるからである。

[0140] 逆に、修正オーディオ信号における、メタデータのゲイン値が修正によって大きくなった、つまり下限値に置き換えられた影響が生じるサンプルでは、修正前よりも信号レベルが小さくなっていることが分かる。

[0141] 〈符号化処理の説明〉

次に、図８に示した符号化装置７１の動作について説明する。すなわち、以下、図１２のフローチャートを参照して、図８の符号化装置７１による符号化処理について説明する。

[0142] ステップＳ９１において修正部１６１は、供給されたオブジェクトのメタデータのゲイン値に応じて、必要に応じてメタデータ、より詳細にはメタデータのゲイン値と、供給されたオブジェクトのオーディオ信号とを修正する。

[0143] すなわち、修正部１６１は、メタデータのゲイン値がMPEG-Hの仕様範囲外、つまりvalue range外の値である場合、そのゲイン値をvalue rangeの上限値または下限値へと置き換える修正を行うとともに、修正前後のゲイン値に基づいてオーディオ信号を修正する。

[0144] 修正部１６１は、適宜修正を行うことにより得られた修正ゲイン値と、供給されたゲイン値以外のメタデータのパラメータとからなる修正メタデータを量子化部２１に供給するとともに、修正ゲイン値をMDCT係数補正部１３１

に供給する。

[0145] また、修正部１６１は、適宜修正を行うことにより得られた修正オーディオ信号を時間周波数変換部３１に供給する。

[0146] このようにして修正メタデータおよび修正オーディオ信号が得られると、その後、ステップＳ９２乃至ステップＳ９９の処理が行われて符号化処理は終了するが、これらの処理は図７のステップＳ５１乃至ステップＳ５８の処理と同様であるので、その説明は省略する。

[0147] 但し、ステップＳ９２およびステップＳ９３では、修正メタデータが量子化および符号化され、ステップＳ９４では修正オーディオ信号に対してMDCTが行われる。

[0148] また、ステップＳ９５では、ステップＳ９４で得られたMDCT係数と、修正部１６１から供給された修正ゲイン値とに基づいて、MDCT係数の補正が行われ、その結果得られた補正MDCT係数が聴覚心理パラメータ計算部４１へと供給される。

[0149] 以上のようにして符号化装置７１は、入力されたメタデータおよびオーディオ信号を、必要に応じて修正してから符号化を行う。

[0150] このようにすることで、実質的にMPEG-Hの仕様によってゲイン値が制限されてしまうことがなくなり、コンテンツ制作者の意図通りのレンダリング結果を得ることができる。

[0151] 〈第４の実施の形態〉

〈聴覚特性に応じたゲイン値の補正について〉

さらに、音源からの音の到来方向に関する聴覚特性に応じて、聴覚心理パラメータの計算に用いるオーディオ信号の補正を行うようにすることもできる。

[0152] 例えば聴覚の特性として、音の大きさの感じ方は、音源からの音の到来方向によって異なる。

[0153] つまり、同じオブジェクトであっても、聴取者に対して前方、側方、上方、および下方の各方向に音源がある場合で、それぞれ聴感上の音の大きさが

異なる。そのため、実際の聴感に即した聴覚心理パラメータの計算を行うためには、音源からの音の到来方向による音圧感度の違いを踏まえたゲイン補正が必要である。

[0154] ここで、音の到来方向に応じた音圧感度の違いと、その音圧感度に応じた補正について説明する。

[0155] 図１３は、あるピンクノイズが聴取者の真正面で再生されたときの聴感上の音の大きさを基準として、同じピンクノイズを異なる方向から再生した際に、聴感上の音の大きさが同じに感じるようにピンクノイズのゲイン補正を行ったときのゲイン補正量の一例を示している。

[0156] なお、図１３において縦軸はゲイン補正量を示しており、横軸は聴取者から見た音源位置を示す水平方向の角度であるAzimuth（水平角度）を示している。

[0157] 例えば、聴取者から見た真正面の方向を示すAzimuthは0度であり、聴取者から見た真横の方向、つまり側方を示すAzimuthは±90度であり、聴取者の後方、つまり真後ろの方向を示すAzimuthは180度である。特に、聴取者から見て左方向がAzimuthの正の方向となっている。

[0158] この例では、複数人の聴取者を対象として行われた実験の結果から得られた各Azimuthに対するゲイン補正量の平均値を示しており、特に各Azimuthにおいて点線で表される範囲は95％の信頼区間を示している。

[0159] 例えば側方（Azimuth＝±90度）でピンクノイズを再生するときには、ゲインを少し下げることで、聴取者は正面方向でピンクノイズを再生したときと同じ大きさで音が聞こえるように感じることが分かる。

[0160] また、例えば後方（Azimuth＝180度）でピンクノイズを再生するときには、ゲインを少し上げることで、聴取者は正面方向でピンクノイズを再生したときと同じ大きさで音が聞こえるように感じることが分かる。

[0161] すなわち、あるオブジェクト音源に対して、そのオブジェクト音源の定位位置が聴取者の側方にあるときにはオブジェクト音源の音のゲインを少し下げ、オブジェクト音源の定位位置が聴取者の後方にあるときにはオブジェク

ト音源の音のゲインを少し上げると、聴取者に同じ大きさで音が聞こえているように感じさせることができる。

[0162] そこで、オブジェクトの位置情報から、聴覚特性に基づき、そのオブジェクトに対するゲイン値の補正量を決定し、決定した補正量でゲイン値を補正すれば、聴覚特性を考慮した聴覚心理パラメータを得ることができる。

[0163] そのような場合、例えば図１４に示すようにゲイン補正部１９１および聴覚特性テーブル保持部１９２を設けるようにすればよい。

[0164] ゲイン補正部１９１には、オブジェクトのメタデータに含まれるゲイン値が供給されるとともに、オブジェクトのメタデータに含まれる位置情報としての水平角度（Azimuth）、垂直角度（Elevation）、および距離（Radius）が供給される。なお、ここでは説明を簡単にするため、ゲイン値は1.0であるものとする。

[0165] ゲイン補正部１９１は、供給されたメタデータとしての位置情報と、聴覚特性テーブル保持部１９２に保持されている聴覚特性テーブルとに基づいて、オブジェクトのゲイン値を補正するゲイン補正量を示すゲイン補正値を決定する。

[0166] また、ゲイン補正部１９１は、決定したゲイン補正値に基づいて、供給されたゲイン値を補正し、その結果得られたゲイン値を補正ゲイン値として出力する。

[0167] 換言すれば、ゲイン補正部１９１は位置情報により示される、聴取者から見たオブジェクトの方向（音の到来方向）に応じてゲイン補正値を決定することで、聴覚心理パラメータの計算に用いるオーディオ信号のゲイン補正のための補正ゲイン値を決定する。

[0168] 聴覚特性テーブル保持部１９２は、音源からの音の到来方向に関する聴覚特性を示す聴覚特性テーブルを保持しており、必要に応じて聴覚特性テーブルにより示されるゲイン補正値をゲイン補正部１９１に供給する。

[0169] ここで、聴覚特性テーブルは、音源であるオブジェクトから聴取者への音の到来方向、つまり聴取者から見た音源の方向（位置）と、その方向に応じ

たゲイン補正値とが対応付けられているテーブルである。換言すれば、聴覚特性テーブルは、音源からの音の到来方向に対して、聴感上の音の大きさを一定にするゲイン補正量を示す聴覚特性である。

[0170] 聴覚特性テーブルにより示されるゲイン補正値は、音の到来方向に対する人の聴覚特性に応じて定められたものであり、特に音の到来方向によらず聴感上の音の大きさが一定になるようなゲイン補正量となっている。換言すれば、ゲイン補正値は、音の到来方向に関する聴覚特性に基づくゲイン値の補正を行うための補正値である。

[0171] したがって、聴覚特性テーブルにより示されるゲイン補正値によりゲイン値を補正して得られた補正ゲイン値を用いてオブジェクトのオーディオ信号をゲイン補正すれば、オブジェクトの位置によらず、同じオブジェクトの音は同じ大きさで聞こえるようになる。

[0172] ここで、図１５に聴覚特性テーブルの例を示す。

[0173] 図１５に示す例では水平角度（Azimuth）、垂直角度（Elevation）、および距離（Radius）により定まるオブジェクトの位置、つまりオブジェクトの方向に対してゲイン補正値が対応付けられている。

[0174] 特に、この例では全ての垂直角度（Elevation）および距離（Radius）が0および1.0となっており、オブジェクトの垂直方向の位置は聴取者と同じ高さであり、かつ聴取者からオブジェクトまでの距離は常に一定であると想定されている。

[0175] 図１５の例では、例えば水平角度が180度である場合など、音源であるオブジェクトが聴取者の後方にある場合には、水平角度が0度や30度である場合など、オブジェクトが聴取者の前方にある場合よりもゲイン補正値が大きくなっている。

[0176] また、聴覚特性テーブル保持部１９２が図１５に示す聴覚特性テーブルを保持している場合における、ゲイン補正部１９１によるゲイン値の補正の具体的な例について説明する。

[0177] 例えばオブジェクトのメタデータのパラメータである水平角度、垂直角度

、および距離が90度、0度、および1.0mであるとすると、図１５からオブジェクトの位置に対応するゲイン補正値は-0.52dBとなる。

[0178] したがって、ゲイン補正部１９１は、聴覚特性テーブルから読み出したゲイン補正値「-0.52dB」と、ゲイン値「1.0」とに基づいて次式（１）の計算を行い、補正ゲイン値「0.94」を得る。

[0179] [数1]

$$1.0\times10^{-0.52/20}\cong0.94 \qquad \cdots(1)$$

[0180] 同様に、例えばオブジェクトの位置を示す水平角度、垂直角度、および距離が-150度、0度、および1.0mであるとすると、図１５からオブジェクトの位置に対応するゲイン補正値は0.51dBとなる。

[0181] したがって、ゲイン補正部１９１は、聴覚特性テーブルから読み出したゲイン補正値「0.51dB」と、ゲイン値「1.0」とに基づいて次式（２）の計算を行い、補正ゲイン値「1.06」を得る。

[0182] [数2]

$$1.0\times10^{0.51/20}\cong1.06 \qquad \cdots(2)$$

[0183] なお、図１５では水平方向のみが考慮された２次元の聴覚特性に基づいて決定されたゲイン補正値を利用する例について説明した。つまり、２次元の聴覚特性に基づいて生成された聴覚特性テーブル（以下、２次元聴覚特性テーブルとも称する）を利用する例について説明した。

[0184] しかし、水平方向だけでなく垂直方向の特性も考慮された３次元の聴覚特性に基づいて決定されたゲイン補正値を利用してゲイン値を補正するようにしてもよい。

[0185] そのような場合、例えば図１６に示す聴覚特性テーブルを利用することができる。

[0186] 図１６に示す例では、水平角度（Azimuth）、垂直角度（Elevation）、および距離（Radius）により定まるオブジェクトの位置、つまりオブジェクトの方向に対してゲイン補正値が対応付けられている。

[0187] 特に、この例では全ての水平角度と垂直角度の組み合わせにおいて、距離は1.0とされている。

[0188] 以下では、図１６に示すように音の到来方向に対する３次元の聴覚特性に基づいて生成された聴覚特性テーブルを、特に３次元聴覚特性テーブルとも称することとする。

[0189] ここで、聴覚特性テーブル保持部１９２が図１６に示す聴覚特性テーブルを保持している場合における、ゲイン補正部１９１によるゲイン値の補正の具体的な例について説明する。

[0190] 例えばオブジェクトの位置を示す水平角度、垂直角度、および距離が60度、30度、および1.0mであるとすると、図１６からオブジェクトの位置に対応するゲイン補正値は-0.07dBとなる。

[0191] したがって、ゲイン補正部１９１は、聴覚特性テーブルから読み出したゲイン補正値「-0.07dB」と、ゲイン値「1.0」とに基づいて次式（３）の計算を行い、補正ゲイン値「0.99」を得る。

[0192] [数3]

$$1.0\times10^{-0.07/20}\cong0.99 \qquad \cdots(3)$$

[0193] なお、以上において説明した補正ゲイン値算出の具体例では、オブジェクトの位置（方向）に対して定まる聴覚特性に基づくゲイン補正値が予め用意されていた。すなわち、オブジェクトの位置情報に対応するゲイン補正値が聴覚特性テーブルに格納されている例について説明した。

[0194] しかしながら、オブジェクトの位置は、聴覚特性テーブルにおいて、対応するゲイン補正値が格納されている位置にあるとは限らない。

[0195] 具体的には、例えば聴覚特性テーブル保持部１９２に図１６に示した聴覚特性テーブルが保持されており、位置情報としての水平角度、垂直角度、および距離が-120度、15度、および1.0mであるとする。

[0196] この場合、図１６の聴覚特性テーブルには、水平角度「-120」、垂直角度「15」、および距離「1.0」に対応するゲイン補正値は格納されていない。

[0197] そこで、聴覚特性テーブルに、位置情報により示される位置に対応するゲ

イン補正値がない場合には、その位置情報により示される位置に隣接する、対応するゲイン補正値が存在する複数の位置のゲイン補正値を用いて、ゲイン補正部１９１が補間処理等により所望の位置のゲイン補正値を算出するようにしてもよい。換言すれば、位置情報により示される位置の近傍にある、複数の各位置に対応付けられているゲイン補正値に基づいて補間処理等を行うことにより、位置情報により示される位置のゲイン補正値が求められる。

[0198] 例えば、ゲイン補正値の補間方法の１つとしてVBAP（Vector Base Amplitude Panning）を利用した方法がある。

[0199] VBAP（３点VBAP）は、３次元空間音響のレンダリングでよく用いられる振幅パンニング手法である。

[0200] VBAPでは、任意の仮想スピーカの近傍にある３つの実スピーカに対して、それぞれ重み付けのゲインを与えて音源信号を再生することで、仮想スピーカの位置を任意に変えることができる。

[0201] その際、聴取位置から各実スピーカまでの３方向のベクトルL1、ベクトルL2、およびベクトルL3を、実スピーカに与えるゲインで重み付け加算して得られる合成ベクトルの向きが、仮想スピーカの向き（Lp）と一致するように各実スピーカのゲインvg1、ゲインvg2、およびゲインvg3が求められる。具体的には、仮想スピーカの向き、すなわち聴取位置から仮想スピーカまでのベクトルをベクトルLpとすると、次式（４）を満たすゲインvg1乃至ゲインvg3が求められる。

[0202] [数4]

$$Lp=L1*vg1+L2*vg2+L3*vg3 \quad \cdots(4)$$

[0203] ここで、上述の３つの各実スピーカの位置を、聴覚特性テーブルに対応する３つのゲイン補正値CG1、ゲイン補正値CG2、およびゲイン補正値CG3が存在する位置とする。また、上述の仮想スピーカの位置を、聴覚特性テーブルに対応するゲイン補正値がない任意の位置とする。

[0204] このとき、次式（５）を計算することで、仮想スピーカの位置におけるゲイン補正値CGpを得ることができる。

[0205] [数5]

$$Ri = vgi / \sqrt{(vg1*vg1 + vg2*vg2 + vg3*vg3)} \qquad i = 1, 2, 3$$

$$CGp = R1*CG1 + R2*CG2 + R3*CG3 \qquad \cdots (5)$$

[0206] 式（５）では、まず、VBAPで求めた上述の重み付けのゲインvg1、ゲインvg2、およびゲインvg3に対して２乗和が１になるように正規化が行われて比率R1、比率R2、および比率R3が求められる。

[0207] そして、求められた比率R1、比率R2、および比率R3により、実スピーカ位置のゲイン補正値CG1、ゲイン補正値CG2、およびゲイン補正値CG3が重み付けされて加算された合成ゲインが、仮想スピーカの位置におけるゲイン補正値CGpとされる。

[0208] 具体的には、３次元空間内においてゲイン補正値が用意されている複数の位置でメッシュが区切られる。すなわち、例えば３次元空間内の３つの各位置のゲイン補正値が用意されているとすると、それらの３つの位置を頂点とする１つの３角形の領域が１つのメッシュとされる。

[0209] このようにして３次元空間が複数のメッシュに区切られると、ゲイン補正値を得ようとする所望の位置を注目位置として、その注目位置を内包するメッシュが特定される。

[0210] また、特定されたメッシュを構成する３つの頂点位置を示す位置ベクトルの乗加算により注目位置を示す位置ベクトルを表したときの３つの各頂点位置を示す位置ベクトルに乗算される係数がVBAPにより求められる。

[0211] そして、このようにして求められた３つの係数を２乗和が１になるように正規化したそれぞれが、注目位置を内包するメッシュの３つの各頂点位置のゲイン補正値のそれぞれに乗算され、係数が乗算されたゲイン補正値の和が注目位置のゲイン補正値として算出される。また、正規化は、和または３乗以上の和が１になるようにするなど、どのような方法で行われてもよい。

[0212] なお、ゲイン補正値の補間方法はVBAPを利用した補間に限らず、他のどのような方法であってもよい。

[0213] 例えば聴覚特性テーブルにおいてゲイン補正値が存在する位置のうち、注

目位置の近傍にあるN個（例えばN＝5）の位置など、複数個の位置のゲイン補正値の平均値を注目位置のゲイン補正値として用いてもよい。

[0214] また、例えば聴覚特性テーブルにおいてゲイン補正値が存在する位置のうち、注目位置に最も近い、ゲイン補正値が用意（格納）されている位置のゲイン補正値を注目位置のゲイン補正値として用いるようにしてもよい。

[0215] ところで、図１６に示した聴覚特性テーブルでは、各位置に対して１つのゲイン補正値が用意されている。言い換えると、全周波数でゲイン補正値が均一ということになる。

[0216] しかし、方向による主観的な音圧感度の違いは、周波数によって変わることも知られている。そこで、１つの位置に対して複数の周波数ごとにゲイン補正値が用意されるようにしてもよい。

[0217] ここで、１つの位置に対して３つの周波数におけるゲイン補正値がある場合における聴覚特性テーブルの例を図１７に示す。

[0218] 図１７に示す例では水平角度（Azimuth）、垂直角度（Elevation）、および距離（Radius）により定まる位置に対して、250Hz、1kHz、および8kHzの３つの各周波数におけるゲイン補正値が対応付けられている。なお、ここでは距離（Radius）は一定値であると想定されており、その値は聴覚特性テーブルには記されていない。

[0219] 例えば水平角度が-30度であり、垂直角度が0度である位置では、250Hzにおけるゲイン補正値は-0.91とされ、1kHzにおけるゲイン補正値は-1.34とされ、8kHzにおけるゲイン補正値は-0.92とされている。

[0220] なお、ここでは各位置について250Hz、1kHz、および8kHzという３つの周波数におけるゲイン補正値が用意された聴覚特性テーブルを例として示した。しかし、これに限らず、聴覚特性テーブルにおいて、位置ごとにゲイン補正値が用意される周波数の数や、ゲイン補正値を用意する周波数は任意の数や周波数とすることができる。

[0221] また、上述の例と同様に、聴覚特性テーブルにおいて、オブジェクトの位置に対して所望する周波数のゲイン補正値が格納されていないこともある。

[0222] そこで、ゲイン補正部１９１が、聴覚特性テーブルにおいてオブジェクトの位置またはその位置近傍の位置における、所望の周波数近傍の他の複数の周波数に対応付けられているゲイン補正値に基づいて補間処理等を行うことにより、オブジェクトの位置における所望の周波数のゲイン補正値を求めるようにしてもよい。

[0223] 例えば補間処理により所望の周波数のゲイン補正値を求める場合、０次補間や１次補間などの線形補間、スプライン補間などの非線形補間、任意の線形補間と非線形補間を組み合わせた補間処理など、どのような補間処理を行うようにしてもよい。

[0224] また、所望の位置について、周波数の最小値や最大値におけるゲイン補正値が存在しない（用意されていない）場合には、そのゲイン補正値を近傍の周波数のゲイン補正値に基づいて決定してもよいし、0dB等の固定値としてもよい。

[0225] ここで、聴覚特性テーブルにおいて、所定の位置の250Hz、1kHz、および8kHzの各周波数のゲイン補正値が存在しており、他の周波数のゲイン補正値がない場合に、他の各周波数のゲイン補正値を補間処理により求めたときの例を図１８に示す。なお、図１８において縦軸はゲイン補正値を示しており、横軸は周波数を示している。

[0226] この例では、250Hz、1kHz、および8kHzの各周波数におけるゲイン補正値に基づいて、線形補間や非線形補間等の補間処理が行われ、全周波数のゲイン補正値が求められている。

[0227] ところで、再生音圧によって等ラウドネス曲線が変わることが知られており、聴覚特性テーブルもオーディオ信号の再生音圧によって切り替える方がよいこともあり得る。

[0228] そこで、例えば聴覚特性テーブル保持部１９２に、複数の再生音圧ごとに聴覚特性テーブルを保持させておき、それらの聴覚特性テーブルのなかから、ゲイン補正部１９１がオブジェクトのオーディオ信号の音圧に基づいて適切なものを選択するようにしてもよい。つまり、ゲイン補正部１９１が、再

生音圧に応じて、ゲイン値の補正に用いる聴覚特性テーブルを切り替えるようにしてもよい。

[0229] この場合においても、上述した各位置や周波数のゲイン補正値の補間と同様に、対応する音圧の聴覚特性テーブルが聴覚特性テーブル保持部１９２にないときには、その聴覚特性テーブルのゲイン補正値を補間処理等により求めるようにしてもよい。

[0230] そのような場合、例えばゲイン補正部１９１は、オブジェクトのオーディオ信号の音圧に近い、つまりその音圧近傍の他の複数の再生音圧に対応付けられている聴覚特性テーブルの所定位置のゲイン補正値に基づいて補間処理等を行うことで、オブジェクトのオーディオ信号の音圧における所定位置のゲイン補正値を求める。このとき、例えば等ラウドネス曲線における曲線間の間隔等に応じた重みを付けて補間を行うようにしてもよい。

[0231] さらに、位置や周波数、再生音圧に応じて一律にオブジェクトのオーディオ信号（MDCT係数）のゲイン補正を行うと、かえって全体の音質が低下してしまう場合もある。

[0232] 具体的には、例えばもともと聴感上重要でない微小なノイズ音がオブジェクトのオーディオ信号とされている場合が考えられる。

[0233] この場合、微小なノイズ音のオブジェクトがゲイン補正値の大きい位置に配置されると、ビットアロケーション部４２において、そのオブジェクトのオーディオ信号に割り当てられるビット数が増えてしまう。そうすると、その分だけ重要な他のオブジェクトの音（オーディオ信号）に割り当てられるビット数が少なくなり、音質が低下してしまう可能性がある。

[0234] したがって、オブジェクトのオーディオ信号の特性に応じてゲイン補正の方法を変えるようにしてもよい。

[0235] 例えば上述の例であれば、ゲイン補正部１９１は、オーディオ信号のPE（Perceptual Entropy）や音圧がある閾値以下である、つまり重要でないオブジェクトであると判別できる場合には、ゲイン補正が行われないようにしたり、ゲイン補正の補正量を制限する、つまり補正ゲイン値が上限値以下となる

ように補正ゲイン値を制限したりするようにしてもよい。これにより、MDCT係数補正部１３１における、補正ゲイン値によるMDCT係数（オーディオ信号）の補正が制限されることになる。

[0236] その他、例えばゲイン補正部１９１は、オブジェクト音の周波数パワーに偏りがある場合には、主要な周波数帯域とそれ以外の周波数帯域とでゲイン補正に重み付けするようにしてもよい。そのような場合、例えば周波数帯域ごとに、周波数パワーに応じてゲイン補正値が補正される。

[0237] さらに、聴覚特性テーブルには、人によってその特性にばらつきがあることが知られている。よって、特定のユーザに最適化された聴覚特性テーブルを用いて、特定のユーザに最適化したエンコーダを構成することも可能である。

[0238] そのような場合、例えば聴覚特性テーブル保持部１９２にユーザごとに最適化された、複数のユーザごとの聴覚特性テーブルが保持されるようにすればよい。

[0239] なお、聴覚特性テーブルの最適化は、特定の人だけで聴覚特性を調べる実験を行った結果を使用してもよいし、他の別の方法で行ってもよい。

[0240] 〈符号化装置の構成例〉

以上のように聴覚特性に応じてゲイン値を補正する場合、符号化装置７１は、例えば図１９に示すように構成される。なお、図１９において図６または図１４における場合と対応する部分には同一の符号を付してあり、その説明は適宜省略する。

[0241] 図１９に示す符号化装置７１は、メタエンコーダ１１、コアエンコーダ１２、および多重化部８１を有している。

[0242] また、メタエンコーダ１１は、量子化部２１および符号化部２２を有しており、コアエンコーダ１２は、ゲイン補正部１９１、聴覚特性テーブル保持部１９２、時間周波数変換部３１、MDCT係数補正部１３１、量子化部３２、および符号化部３３を有している。さらに量子化部３２は、聴覚心理パラメータ計算部４１およびビットアロケーション部４２を有している。

[0243] 図１９に示す符号化装置７１の構成は、新たにゲイン補正部１９１および聴覚特性テーブル保持部１９２が設けられている点で図６の符号化装置７１の構成と異なり、その他の点では図６の符号化装置７１と同じ構成となっている。

[0244] 図１９の例では、聴覚特性テーブル保持部１９２には、例えば図１６に示した３次元聴覚特性テーブルが保持されている。

[0245] また、ゲイン補正部１９１には、オブジェクトのメタデータのゲイン値、水平角度、垂直角度、および距離が供給される。

[0246] ゲイン補正部１９１は、聴覚特性テーブル保持部１９２に保持されている３次元聴覚特性テーブルから、供給されたメタデータの位置情報としての水平角度、垂直角度、および距離に対応付けられているゲイン補正値を読み出す。

[0247] なお、メタデータの位置情報により示されるオブジェクトの位置に対応するゲイン補正値がない場合には、ゲイン補正部１９１は、適宜、補間処理等を行うことにより、位置情報により示されるオブジェクトの位置に対応するゲイン補正値を得る。

[0248] ゲイン補正部１９１は、このようにして得られたゲイン補正値により、供給されたオブジェクトのメタデータのゲイン値を補正し、その結果得られた補正ゲイン値をMDCT係数補正部１３１に供給する。

[0249] したがって、MDCT係数補正部１３１は、ゲイン補正部１９１から供給された補正ゲイン値に基づいて、時間周波数変換部３１から供給されたMDCT係数の補正を行い、その結果得られた補正MDCT係数を聴覚心理パラメータ計算部４１に供給する。

[0250] なお、図１９に示す例では、MDCT係数のゲイン補正に量子化前のメタデータを用いる例について説明したが、符号化後や量子化後のメタデータを利用するようにしてもよい。

[0251] そのような場合、ゲイン補正部１９１において、符号化後や量子化後のメタデータが復号されたり逆量子化されたりして、その結果得られたゲイン値

、水平角度、垂直角度、および距離に基づいて補正ゲイン値が求められる。

[0252] その他、図４や図８に示した構成に、ゲイン補正部１９１および聴覚特性テーブル保持部１９２を設けるようにしてもよい。

[0253] 〈符号化処理の説明〉

次に、図１９に示した符号化装置７１の動作について説明する。すなわち、以下、図２０のフローチャートを参照して、図１９の符号化装置７１による符号化処理について説明する。

[0254] なお、ステップＳ１３１およびステップＳ１３２の処理は、図７のステップＳ５１およびステップＳ５２の処理と同様であるので、その説明は省略する。

[0255] ステップＳ１３３においてゲイン補正部１９１は、供給されたメタデータのゲイン値、水平角度、垂直角度、および距離に基づいて補正ゲイン値を算出し、MDCT係数補正部１３１に供給する。

[0256] すなわち、ゲイン補正部１９１は、聴覚特性テーブル保持部１９２に保持されている３次元聴覚特性テーブルから、メタデータの水平角度、垂直角度、および距離に対応付けられているゲイン補正値を読み出して、そのゲイン補正値によりゲイン値を補正することで補正ゲイン値を算出する。このとき、適宜、補間処理等が行われて、水平角度、垂直角度、および距離により示されるオブジェクトの位置に対応するゲイン補正値が求められる。

[0257] このようにして補正ゲイン値が求められると、その後、ステップＳ１３４乃至ステップＳ１３９の処理が行われて符号化処理は終了するが、これらの処理は図７のステップＳ５３乃至ステップＳ５８の処理と同様であるので、その説明は省略する。

[0258] 但し、ステップＳ１３５では、時間周波数変換部３１で得られたMDCT係数が、ゲイン補正部１９１で得られた補正ゲイン値に基づいて補正され、補正MDCT係数とされる。

[0259] なお、上述したように最適化されたユーザごとの聴覚特性テーブルが聴覚特性テーブル保持部１９２に保持されているようにしてもよい。

[0260] また、聴覚特性テーブルにおいて、各位置について複数の周波数ごとにゲイン補正値が対応付けられているようにしてもよく、ゲイン補正部１９１が所望の周波数のゲイン補正値を、その周波数近傍の他の複数の周波数のゲイン補正値に基づく補間処理により求めるようにしてもよい。

[0261] 例えば聴覚特性テーブルにおいて、各位置に周波数ごとのゲイン補正値が対応付けられて格納されている場合、ゲイン補正部１９１では周波数ごとに補正ゲイン値が求められ、MDCT係数補正部１３１では、周波数ごとに補正ゲイン値によりMDCT係数が補正される。その他、聴覚特性テーブル保持部１９２には、再生音圧ごとに聴覚特性テーブルが保持されているようにしてもよい。

[0262] 以上のようにして符号化装置７１は、３次元聴覚特性テーブルを用いてメタデータのゲイン値を補正するとともに、その結果得られた補正ゲイン値を用いて得られた補正MDCT係数に基づいて聴覚心理パラメータを計算する。

[0263] このようにすることで、少ない演算量でも、実際の聴感に即した聴覚心理パラメータを得ることができ、符号化効率を向上させることができる。特に、３次元の聴覚特性に基づいてゲイン値が補正されるので、より実際の聴感に即した聴覚心理パラメータを得ることができる。

[0264] 〈第５の実施の形態〉

〈符号化装置の構成例〉

ところで、３次元の聴覚の特性として、音源からの音の到来方向による音圧感度の違いだけでなく、オブジェクト間の音にも聴覚マスキングがあり、さらにオブジェクト間の距離や音の周波数特性によってオブジェクト間のマスキング量が変わることが知られている。

[0265] しかし、一般的な聴覚心理パラメータの計算では、オブジェクトごとに個別に聴覚マスキングの計算を行っており、オブジェクト間の聴覚マスキングは考慮されていない。

[0266] そのため、複数のオブジェクトの音が同時に再生される場合に、実際にはオブジェクト間の聴覚マスキングによって、本来であれば量子化ノイズが知

覚されないにも関わらず量子化ビットを過剰に使用してしまう場合がある。

[0267] そこで、オブジェクトの位置や距離に応じた、複数のオブジェクト間の聴覚マスキングを考慮した３次元聴覚心理モデルを用いて聴覚心理パラメータの計算を行うことで、さらに符号化効率のよいビットアロケーションを行うようにしてもよい。

[0268] そのような場合、符号化装置７１は、例えば図２１に示すように構成される。なお、図２１において図４における場合と対応する部分には同一の符号を付してあり、その説明は適宜省略する。

[0269] 図２１に示す符号化装置７１は、メタエンコーダ１１、コアエンコーダ１２、および多重化部８１を有している。

[0270] また、メタエンコーダ１１は、量子化部２１および符号化部２２を有しており、コアエンコーダ１２は、時間周波数変換部３１、量子化部３２、および符号化部３３を有している。さらに量子化部３２は、聴覚心理モデル保持部２２１、聴覚心理パラメータ計算部２２２、およびビットアロケーション部４２を有している。

[0271] 図２１に示す符号化装置７１の構成は、オーディオ信号補正部９１と時間周波数変換部９２、および聴覚心理パラメータ計算部４１に代えて、聴覚心理モデル保持部２２１および聴覚心理パラメータ計算部２２２を設けた点で図４の符号化装置７１の構成と異なり、その他の点では図４の符号化装置７１と同じ構成となっている。

[0272] この例では、聴覚心理モデル保持部２２１は、予め用意された、複数のオブジェクト間の聴覚マスキングに関する３次元聴覚心理モデルを保持している。この３次元聴覚心理モデルは、オブジェクト単体の聴覚マスキングだけでなく、複数のオブジェクト間の聴覚マスキングも考慮された聴覚心理モデルである。

[0273] また、時間周波数変換部３１で得られたMDCT係数と、オブジェクトのメタデータの水平角度、垂直角度、距離、およびゲイン値とが聴覚心理パラメータ計算部２２２に供給される。

[0274] 聴覚心理パラメータ計算部２２２は、３次元聴覚特性に基づく聴覚心理パラメータ計算を行う。すなわち、聴覚心理パラメータ計算部２２２は、時間周波数変換部３１からのMDCT係数、供給されたメタデータの水平角度、垂直角度、距離、およびゲイン値、並びに聴覚心理モデル保持部２２１に保持されている３次元聴覚心理モデルに基づいて聴覚心理パラメータを計算し、ビットアロケーション部４２に供給する。

[0275] このような３次元聴覚特性に基づく聴覚心理パラメータ計算では、従来から考慮されていたオブジェクトごとの聴覚マスキングだけでなく、オブジェクト間の聴覚マスキングも考慮された聴覚心理パラメータを得ることができる。

[0276] これにより、３次元聴覚特性に基づく聴覚心理パラメータを用いてビットアロケーションを行うことができ、さらに符号化効率を向上させることができる。

[0277] 〈符号化処理の説明〉

次に、図２１に示した符号化装置７１の動作について説明する。すなわち、以下、図２２のフローチャートを参照して、図２１の符号化装置７１による符号化処理について説明する。

[0278] なお、ステップＳ１７１およびステップＳ１７２の処理は、図５のステップＳ１１およびステップＳ１２の処理と同様であるので、その説明は省略する。

[0279] ステップＳ１７３において時間周波数変換部３１は、供給されたオブジェクトのオーディオ信号に対するMDCT（時間周波数変換）を行い、その結果得られたMDCT係数を聴覚心理パラメータ計算部２２２およびビットアロケーション部４２に供給する。

[0280] ステップＳ１７４において聴覚心理パラメータ計算部２２２は、時間周波数変換部３１からのMDCT係数、供給されたメタデータの水平角度、垂直角度、距離、およびゲイン値、並びに聴覚心理モデル保持部２２１に保持されている３次元聴覚心理モデルに基づいて聴覚心理パラメータを計算し、ビット

アロケーション部４２に供給する。

[0281] このとき、聴覚心理パラメータ計算部２２２は、処理対象のオブジェクトのMDCT係数や、水平角度、垂直角度、距離、ゲイン値だけでなく、他のオブジェクトのMDCT係数や、水平角度、垂直角度、距離、ゲイン値も用いて聴覚心理パラメータを計算する。

[0282] 具体的な例として、例えば聴覚心理パラメータとしてのマスキング閾値を求める場合について説明する。

[0283] この場合、処理対象のオブジェクトのMDCT係数やゲイン値などに基づいてマスキング閾値が求められる。そして、処理対象のオブジェクトと他のオブジェクトのMDCT係数やゲイン値、位置情報と、３次元聴覚心理モデルとに基づいて、オブジェクト間の距離や相対的な位置関係、周波数のパワー（MDCT係数）の差などに応じたオフセット値（補正値）が求められる。さらに、求めたマスキング閾値がオフセット値により補正され、最終的なマスキング閾値とされる。

[0284] このようにすることで、オブジェクト間の聴覚マスキングも考慮された聴覚心理パラメータを得ることができる。

[0285] 聴覚心理パラメータが算出されると、その後、ステップＳ１７５乃至ステップＳ１７７の処理が行われて符号化処理は終了するが、これらの処理は図５のステップＳ１７乃至ステップＳ１９の処理と同様であるので、その説明は省略する。

[0286] 以上のようにして符号化装置７１は、３次元聴覚心理モデルに基づいて聴覚心理パラメータを計算する。このようにすることで、オブジェクト間の聴覚マスキングも考慮した３次元聴覚特性に基づく聴覚心理パラメータを用いてビットアロケーションを行うことができ、符号化効率を向上させることができる。

[0287] 〈第６の実施の形態〉

〈符号化装置の構成例〉

なお、以上において説明したオブジェクトのメタデータのゲイン値や位置

情報をビットアロケーションに利用する手法は、例えば、配信されたコンテンツの視聴の際に、オブジェクトのメタデータ、すなわち位置やゲインを、ユーザが変更なしにそのまま使用してレンダリングするようなサービスにおいては有効である。

[0288] 一方で、このような手法は、ユーザがレンダリング時のメタデータを編集できるようなサービスにおいては、符号化（エンコード）時とレンダリング時とでメタデータが異なる可能性があるため、そのままでは利用できない。

[0289] しかし、そのようなサービスであっても、コンテンツ制作者が全てのオブジェクトのメタデータの編集を許可するとは限らず、ユーザによるメタデータの編集を許可するオブジェクトと許可しないオブジェクトをコンテンツ制作者が指定することが考えられる。

[0290] ここで、コンテンツ制作者によるオブジェクトごとのメタデータの編集許可フラグ「editingPermissionFlag」を追加したメタデータのConfigのシンタックスを図２３に示す。編集許可フラグは、メタデータの編集を許可するか否かを示す編集許可情報の一例である。

[0291] この例では、メタデータのConfig（ObjectMetadataConfig）における矢印Q11に示す部分に編集許可フラグ「editingPermissionFlag」が含まれている。

[0292] ここで、「num_objects」はコンテンツを構成するオブジェクトの数を示しており、この例ではオブジェクトごとに編集許可フラグが格納されている。

[0293] 特に、編集許可フラグの値「１」はオブジェクトのメタデータの編集が許可されていることを示しており、編集許可フラグの値「０」はオブジェクトのメタデータの編集が許可されていないことを示している。コンテンツ制作者は、オブジェクトごとに編集許可フラグの値を指定（セット）する。

[0294] このような編集許可フラグをメタデータに含めるようにすれば、メタデータの編集が許可されていないオブジェクトについては、３次元聴覚心理モデルに基づいて聴覚心理パラメータを計算することができる。

[0295] そのような場合、符号化装置７１は、例えば図２４に示すように構成される。なお、図２４において図２１における場合と対応する部分には同一の符

号を付してあり、その説明は適宜省略する。

[0296] 図２４に示す符号化装置７１は、メタエンコーダ１１、コアエンコーダ１２、および多重化部８１を有している。

[0297] また、メタエンコーダ１１は、量子化部２１および符号化部２２を有しており、コアエンコーダ１２は、時間周波数変換部３１、量子化部３２、および符号化部３３を有している。さらに量子化部３２は、聴覚心理モデル保持部２２１、聴覚心理パラメータ計算部２２２、およびビットアロケーション部４２を有している。

[0298] 図２４に示す符号化装置７１は、基本的には図２１に示した符号化装置７１と同じであるが、図２４に示す符号化装置７１では、入力されるメタデータにオブジェクトごとの編集許可フラグが含まれている点で、図２１の符号化装置７１と異なっている。

[0299] この例では、メタデータの各パラメータとして、水平角度、垂直角度、距離、ゲイン値、編集許可フラグ、および他のパラメータ（Other parameters）が量子化部２１に入力される。また、メタデータのうちの水平角度、垂直角度、距離、ゲイン値、および編集許可フラグが聴覚心理パラメータ計算部２２２に供給される。

[0300] したがって、聴覚心理パラメータ計算部２２２は、供給された編集許可フラグに応じて、図４を参照して説明した聴覚心理パラメータ計算部４１と同様にして聴覚心理パラメータを計算するか、または図２１の例と同様にして聴覚心理パラメータを計算する。

[0301] 〈符号化処理の説明〉

次に、図２４に示した符号化装置７１の動作について説明する。すなわち、以下、図２５のフローチャートを参照して、図２４の符号化装置７１による符号化処理について説明する。

[0302] なお、ステップＳ２１１乃至ステップＳ２１３の処理は、図２２のステップＳ１７１乃至ステップＳ１７３の処理と同様であるので、その説明は省略する。

[0303] ステップＳ２１４において聴覚心理パラメータ計算部２２２は、供給されたオブジェクトのメタデータに含まれる編集許可フラグに応じて、聴覚心理パラメータを計算し、ビットアロケーション部４２に供給する。

[0304] 例えば聴覚心理パラメータ計算部２２２は、処理対象のオブジェクトの編集許可フラグが「１」であり、編集が許可されている場合、時間周波数変換部３１から供給された、処理対象のオブジェクトのMDCT係数に基づいて、聴覚心理パラメータを計算する。

[0305] このように、編集が許可されているオブジェクトについては、復号（再生）側でメタデータが編集される可能性があるので、オブジェクト間の聴覚マスキングは考慮されずに聴覚心理パラメータが算出される。

[0306] これに対して、例えば処理対象のオブジェクトの編集許可フラグが「０」であり、編集が許可されていない場合、聴覚心理パラメータ計算部２２２は、時間周波数変換部３１からのMDCT係数、供給されたメタデータの水平角度、垂直角度、距離、およびゲイン値、並びに聴覚心理モデル保持部２２１に保持されている３次元聴覚心理モデルに基づいて聴覚心理パラメータを計算する。

[0307] この場合、聴覚心理パラメータ計算部２２２は、図２２のステップＳ１７４における場合と同様にして聴覚心理パラメータを計算する。すなわち、処理対象のオブジェクトのMDCT係数や、水平角度、垂直角度、距離、ゲイン値だけでなく、他のオブジェクトのMDCT係数や、水平角度、垂直角度、距離、ゲイン値も用いられて聴覚心理パラメータが計算される。

[0308] このように、編集が許可されていないオブジェクトについては、復号（再生）側でメタデータが変わることがないので、オブジェクト間の聴覚マスキングが考慮されて聴覚心理パラメータが算出される。

[0309] 聴覚心理パラメータが算出されると、その後、ステップＳ２１５乃至ステップＳ２１７の処理が行われて符号化処理は終了するが、これらの処理は図２２のステップＳ１７５乃至ステップＳ１７７の処理と同様であるので、その説明は省略する。

[0310] 以上のようにして符号化装置７１は、編集許可フラグに応じて、適宜、３次元聴覚心理モデルを用いて聴覚心理パラメータを計算する。このようにすることで、編集が許可されていないオブジェクトについては、オブジェクト間の聴覚マスキングも考慮した３次元聴覚特性に基づく聴覚心理パラメータを用いてビットアロケーションを行うことができる。これにより、符号化効率を向上させることができる。

[0311] なお、ここでは、図２１に示した符号化装置７１の構成に対して、編集許可フラグを組み合わせて利用する例について説明した。しかし、これに限らず、例えば図１９に示した符号化装置７１の構成に対して、編集許可フラグを組み合わせて利用してもよい。

[0312] そのような場合、編集が許可されていないオブジェクトについては、３次元聴覚特性テーブルが用いられてオブジェクトのメタデータのゲイン値が補正されるようにすればよい。

[0313] これに対して、編集が許可されているオブジェクトについては、MDCT係数補正部１３１でMDCT係数の補正は行われず、聴覚心理パラメータ計算部４１において、時間周波数変換部３１で得られたMDCT係数がそのまま用いられて聴覚心理パラメータが計算される。

[0314] さらに、ここでは、１つの編集許可フラグ「editingPermissionFlag」によって、メタデータを構成する全パラメータの編集許可を一括管理する例について説明したが、メタデータのパラメータごとに編集許可フラグが用意されるようにしてもよい。そうすれば、編集許可フラグによって、メタデータに含まれる複数のパラメータのうちの一部または全部について選択的に編集を許可することができる。

[0315] そのような場合、例えばメタデータのパラメータのうちの編集許可フラグにより編集が許可されていないものだけが聴覚心理パラメータの計算に利用されるようにしてもよい。

[0316] 例えば図２４の例で、水平角度等からなる位置情報については編集が許可されており、ゲイン値の編集は許可されていない場合、位置情報は用いられ

ず、ゲイン値が用いられて、３次元聴覚心理モデルに基づく聴覚心理パラメータの計算が行われる。

[0317] 〈第７の実施の形態〉

〈符号化装置の構成例〉

ところで、2chや5.1ch、7.1ch等のチャンネルベースのオーディオ符号化では、様々な楽器のオーディオ信号がミキシングされた音が入力されることを前提としていた。

[0318] そのため、ビットアロケーションのアルゴリズムも、様々な楽器の信号に対して普遍的に安定動作するように調整する必要があった。

[0319] 一方で、オブジェクトベースの3D Audioの符号化では、オブジェクトとなる「Vocal」や「Guitar」、「Bass」といった個別の楽器のオーディオ信号が入力される。そのため、各楽器の信号用にビットアロケーション等のアルゴリズムやパラメータ（以下、調整パラメータとも称する）を最適化することで、符号化効率の向上や演算処理の高速化が可能になる。

[0320] そこで、例えば各オブジェクトの音源種別、つまり「Vocal」や「Guitar」などといった楽器を示すラベル情報を入力し、ラベル情報に応じたアルゴリズムや調整パラメータにより、聴覚心理パラメータを算出するようにしてもよい。換言すれば、ラベル情報に応じたビットアロケーションが行われるようにしてもよい。

[0321] そのような場合、符号化装置７１は、例えば図２６に示すように構成される。なお、図２６において図６における場合と対応する部分には同一の符号を付してあり、その説明は適宜省略する。

[0322] 図２６に示す符号化装置７１は、メタエンコーダ１１、コアエンコーダ１２、および多重化部８１を有している。

[0323] また、メタエンコーダ１１は、量子化部２１および符号化部２２を有しており、コアエンコーダ１２は、パラメータテーブル保持部２５１、時間周波数変換部３１、量子化部３２、および符号化部３３を有している。さらに量子化部３２は、聴覚心理パラメータ計算部４１およびビットアロケーション

部４２を有している。

[0324] 図２６に示す符号化装置７１の構成は、MDCT係数補正部１３１に代えてパラメータテーブル保持部２５１を設けた点で図６の符号化装置７１の構成と異なり、その他の点では図６の符号化装置７１と同じ構成となっている。

[0325] この例では、各オブジェクトの音源種別、すなわち例えばVocal、Chorus、Guitar、Bass、Drums、Kick、Snare、Hi-hat、Piano、Synth、String等のオブジェクトのオーディオ信号に基づく音の楽器種別を示すラベル情報が符号化装置７１に入力（供給）される。

[0326] このラベル情報は、例えば各オブジェクトのオブジェクト信号からなるコンテンツの編集等で利用されるものなどとすることができ、ラベル情報は楽器種別を示す文字列などとされてもよいし、楽器種別を示すID情報などとされてもよい。

[0327] パラメータテーブル保持部２５１には、ラベル情報により示される各楽器種別（音源種別）に対して、MDCTの計算や聴覚心理パラメータの計算、ビットアロケーションに用いるアルゴリズムや調整パラメータを示す情報が対応付けられているパラメータテーブルを保持している。なお、パラメータテーブルでは、楽器種別（音源種別）に対して、アルゴリズムを示す情報と調整パラメータのうちの少なくとも何れか一方が対応付けられていればよい。

[0328] 時間周波数変換部３１は、パラメータテーブル保持部２５１に保持されているパラメータテーブルを参照し、供給されたラベル情報により示される楽器種別に対して定められている調整パラメータやアルゴリズムにより、供給されたオーディオ信号に対するMDCTを行う。

[0329] 時間周波数変換部３１は、MDCTにより得られたMDCT係数を聴覚心理パラメータ計算部４１およびビットアロケーション部４２に供給する。

[0330] また、量子化部３２は、供給されたラベル情報およびMDCT係数に基づいて、ラベル情報により示される楽器種別に対して定められている調整パラメータやアルゴリズムにより、MDCT係数の量子化を行う。

[0331] すなわち、聴覚心理パラメータ計算部４１は、パラメータテーブル保持部

２５１に保持されているパラメータテーブルを参照し、供給されたラベル情報により示される楽器種別に対して定められている調整パラメータやアルゴリズムにより、時間周波数変換部３１からのMDCT係数に基づいて聴覚心理パラメータを計算し、ビットアロケーション部４２に供給する。

[0332] ビットアロケーション部４２は、パラメータテーブル保持部２５１に保持されているパラメータテーブルを参照し、時間周波数変換部３１からのMDCT係数、聴覚心理パラメータ計算部４１からの聴覚心理パラメータ、および供給されたラベル情報に基づいてビットアロケーションおよびMDCT係数の量子化を行う。

[0333] このとき、ビットアロケーション部４２では、MDCT係数および聴覚心理パラメータと、ラベル情報により示される楽器種別に対して定められている調整パラメータやアルゴリズムにより、ビットアロケーションが行われる。

[0334] なお、ラベル情報により示される楽器種別（音源種別）ごとのアルゴリズムや調整パラメータの最適化方法は様々であるが、以下、具体的な例について説明する。

[0335] 例えばMDCT（時間周波数変換）では、MDCTに用いる窓（変換窓）、すなわち窓関数の切り替えを行うことができる。

[0336] そこで、例えば楽器種別がHi-hatやGuitarなど、音の立ち上がりや立ち下がりが重要な楽器のオブジェクトに対しては、カイザー窓のような時間解像度の高い窓を用い、VocalやBassなどの量感が重要な楽器のオブジェクトにはサイン窓を用いるようにしてもよい。

[0337] このように、ラベル情報により示される楽器種別と、その楽器種別に対して定められた窓関数を示す情報とを対応付けてパラメータテーブルに格納しておけば、ラベル情報に応じた窓を用いたMDCTを行うことができる。

[0338] また、聴覚心理パラメータの計算やビットアロケーションにおいても、例えばラベル情報に応じた帯域制限が行われるようにすることができる。

[0339] すなわち、BassやKickといった低音域の楽器、Vocalのような中音域の楽器、Hi-hatのような高音域の楽器、Pianoのような全音域の楽器では、それぞれ

聴感上の重要な帯域や不要な帯域が異なる。そこで、ラベル情報を利用して、量子化ビットをそれぞれの不要な帯域から削減し、重要な帯域に多く割り当てるようにすることができる。

[0340] 具体的には、例えばBassやKickといった低音域の楽器のオブジェクト信号には、本来であれば、高域成分は殆ど含まれていない。しかし、そのような楽器のオブジェクト信号に高域のノイズが多く含まれていると、ビットアロケーションにおいて高域のスケールファクターバンドにも多くの量子化ビットが割り当てられてしまう。

[0341] そこで、例えばBassやKickといった低音域の楽器種別については、低域により多くの量子化ビットが割り当てられ、高域には少ない量子化ビットが割り当てられるように、聴覚心理パラメータの計算やビットアロケーションのための調整パラメータとアルゴリズムが定められる。

[0342] このようにすることで、目的とする信号成分が含まれていない高域の量子化ビット数を低く抑えてノイズを低減させるとともに、目的とする信号成分が含まれる低域の量子化ビット数を増加させ、音質と符号化効率を向上させることができる。

[0343] また、一方でマスキング閾値をはじめとした聴覚心理パラメータにおいても、トーン性の強い楽器、ノイズ性の高い楽器、信号の時間変動の多い楽器、信号の時間変動の少ない楽器などの楽器種別に応じて調整（調整パラメータ）を変えることで、楽器ごとの聴感に知覚されやすい音に多くの量子化ビットを割り当てることができる。

[0344] さらに、例えばAAC（Advanced Audio Coding）やUSACのような方式のエンコーダにおいては、周波数スペクトル情報（MDCT係数）がスケールファクターバンドごとに量子化される。

[0345] 各スケールファクターバンドの量子化値、すなわちスケールファクターバンドごとの割り当てのビット数は、所定の値を初期値としてスタートし、ビットアロケーションループを行うことで最終的な値が決められる。

[0346] 例えばビットアロケーションループでは、所定の条件が満たされるまで、

各スケールファクターバンドの量子化値を変化させながら、つまりビットアロケーションを行いながら、MDCT係数の量子化が繰り返し行われる。ここでいう所定の条件とは、例えば各スケールファクターバンドの量子化MDCT係数のビット数の和が所定の許容ビット数以下であるという条件、および量子化ノイズが十分に小さいという条件である。

[0347] リアルタイムエンコーダなど、符号化（量子化）にかかる時間を短くしたい場合も多く、そのような場合には多少の音質劣化を伴うが、上述のビットアロケーションループの回数（ループ数）に上限を設けることも行われている。

[0348] 当然、各スケールファクターバンドの量子化値の初期値が最終値に近いほど、ビットアロケーションループのループ数は少なくなり、エンコード時間が短くなる。また、ループ数の制限による音質劣化も減少することになる。

[0349] したがって、ラベル情報により示される楽器種別ごとに予め最適な初期値を求めておき、ラベル情報に従って初期値を切り替えることで、短い時間で高音質にオーディオ信号を符号化（量子化）することができる。この場合、例えばラベル情報を聴覚心理パラメータの１つとしたり、パラメータテーブルにおいて楽器種別ごとに調整パラメータとしての量子化値の初期値が定められていたりすればよい。

[0350] 以上のような楽器種別ごとの調整パラメータやアルゴリズムは、予め経験に基づく手調整や、統計的調整、機械学習などで求めることができる。

[0351] 図２６に示した構成の符号化装置７１では、これらの楽器種別ごとの調整パラメータやアルゴリズムが予めパラメータテーブルとして用意される。そして、ラベル情報に応じた調整パラメータやアルゴリズムにより、聴覚心理パラメータの計算やビットアロケーション、すなわち量子化が行われたり、MDCTが行われたりする。

[0352] なお、この例ではラベル情報を単独で使用しているが、他のメタデータの情報と組み合わせて使用してもよい。

[0353] 例えば、オブジェクトのメタデータの他のパラメータの例として、オブジ

ェクトの優先度（Priority）を示す優先度情報が含まれていることがある。

[0354] そこで、時間周波数変換部３１や聴覚心理パラメータ計算部４１、ビットアロケーション部４２において、オブジェクトの優先度情報により示される優先度の値によって、ラベル情報に対して定まる調整パラメータをさらに強弱調整してもよい。逆に、同じ優先度のオブジェクトでもラベル情報によって優先度に差をつけて処理してもよい。

[0355] また、ここではラベル情報を楽器種別に限定して説明したが、楽器種別以外にも、聴取環境を判別するラベル情報を利用することもできる。

[0356] 例えば、車内でコンテンツ等の音を聴取する場合には、エンジン音や走行ノイズにより低音域の量子化ノイズが知覚されにくくなる。また、静かな室内と雑踏のある屋外とでは、最小可聴限、つまり知覚できる音量が異なる。さらに、聴取環境自体も時間経過やユーザの移動に伴って変化する。

[0357] そこで、例えばユーザの聴取環境を示す聴取環境情報も含むラベル情報を符号化装置７１に入力し、ラベル情報に応じた調整パラメータやアルゴリズムにより、ユーザのいる聴取環境に最適な聴覚心理パラメータの計算等が行われるようにしてもよい。

[0358] この場合、例えばパラメータテーブルが参照されて、ラベル情報により示される聴取環境や楽器種別に対して定められた調整パラメータやアルゴリズムにより、MDCTや、聴覚心理パラメータの計算、ビットアロケーションが行われる。

[0359] このようにすることで、様々な聴取環境に対して、より高音質な量子化（符号化）を行うことができる。例えば車内であれば、MDCT係数の量子化時に、知覚されにくい低音域の量子化ノイズのマスキング閾値を大きくして、中高域にビットを多く割り当てることで、よりVocal等の楽器種別のオブジェクトの音質を向上させることができる。

[0360] 〈符号化処理の説明〉

次に、図２６に示した符号化装置７１の動作について説明する。すなわち、以下、図２７のフローチャートを参照して、図２６の符号化装置７１によ

る符号化処理について説明する。

[0361] なお、ステップＳ２５１およびステップＳ２５２の処理は、図７のステップＳ５１およびステップＳ５２の処理と同様であるので、その説明は省略する。

[0362] ステップＳ２５３において時間周波数変換部３１は、パラメータテーブル保持部２５１に保持されているパラメータテーブルと、供給されたラベル情報とに基づいて、供給されたオーディオ信号に対するMDCTを行い、その結果得られたMDCT係数を聴覚心理パラメータ計算部４１およびビットアロケーション部４２に供給する。

[0363] 例えばステップＳ２５３では、オブジェクトのラベル情報に対して定められた調整パラメータやアルゴリズムにより、そのオブジェクトのオーディオ信号に対するMDCTが行われる。

[0364] ステップＳ２５４において聴覚心理パラメータ計算部４１は、供給されたラベル情報に応じてパラメータテーブル保持部２５１に保持されているパラメータテーブルを参照し、時間周波数変換部３１から供給されたMDCT係数に基づいて聴覚心理パラメータを計算して、ビットアロケーション部４２に供給する。

[0365] 例えばステップＳ２５４では、オブジェクトのラベル情報に対して定められた調整パラメータやアルゴリズムにより、そのオブジェクトについての聴覚心理パラメータが算出される。

[0366] ステップＳ２５５においてビットアロケーション部４２は、供給されたラベル情報に応じてパラメータテーブル保持部２５１に保持されているパラメータテーブルを参照し、時間周波数変換部３１からのMDCT係数および聴覚心理パラメータ計算部４１からの聴覚心理パラメータに基づいてビットアロケーションを行い、MDCT係数を量子化する。

[0367] このようにしてMDCT係数が量子化されると、その後、ステップＳ２５６およびステップＳ２５７の処理が行われて符号化処理は終了するが、これらの処理は図７のステップＳ５７およびステップＳ５８の処理と同様であるので

、その説明は省略する。

[0368] 以上のようにして符号化装置７１は、ラベル情報に応じてMDCTや聴覚心理パラメータの計算、ビットアロケーションを行う。このようにすることで、符号化効率や量子化演算の処理速度を向上させることができるとともに、より高音質なオーディオ再生を実現することができる。

[0369] 〈第８の実施の形態〉

〈符号化装置の構成例〉

また、ラベル情報を用いて量子化（符号化）を行う符号化装置７１は、例えばMPEG-Iの自由視点のようなユーザの位置情報とオブジェクトの位置情報とを組み合わせて利用する場合にも適用可能である。

[0370] そのような場合、符号化装置７１は、例えば図２８に示すように構成される。なお、図２８において図２６における場合と対応する部分には同一の符号を付してあり、その説明は適宜省略する。

[0371] 図２８に示す符号化装置７１は、メタエンコーダ１１、コアエンコーダ１２、および多重化部８１を有している。

[0372] また、図示は省略されているが、メタエンコーダ１１は量子化部２１および符号化部２２を有している。

[0373] さらにコアエンコーダ１２は、パラメータテーブル保持部２５１、時間周波数変換部３１、量子化部３２、および符号化部３３を有しており、量子化部３２は聴覚心理パラメータ計算部４１およびビットアロケーション部４２を有している。

[0374] 図２８に示す符号化装置７１の構成は、基本的には図２６に示した符号化装置７１と同じであるが、図２８に示す符号化装置７１では、ユーザの位置、すなわちユーザによるコンテンツ等の音の聴取位置を示すユーザ位置情報がさらに入力される点で、図２６の符号化装置７１と異なっている。

[0375] メタエンコーダ１１は、オブジェクトの位置情報やゲイン値などの各パラメータが含まれるメタデータを符号化するが、ここではメタデータに含まれるオブジェクトの位置情報が図２６に示した例とは異なる。

[0376] 例えば、この例では、ユーザ位置情報と、供給されたオブジェクトの水平角度、垂直角度、および距離とに基づいて、ユーザ（聴取位置）から見たオブジェクトの相対的な位置を示す位置情報や、適宜、修正されたオブジェクトの絶対的な位置を示す位置情報などが、オブジェクトのメタデータを構成する位置情報として符号化される。

[0377] なお、ユーザ位置情報は、例えば符号化装置７１によるコンテンツ、つまり符号化メタデータと符号化オーディオデータとを含むビットストリームの配信先（送信先）である図示せぬクライアント装置から供給される。

[0378] また、聴覚心理パラメータ計算部４１では、ラベル情報だけでなく、供給されたオブジェクトの位置情報、つまりオブジェクトの位置を示す水平角度、垂直角度、および距離や、ユーザ位置情報も用いられて聴覚心理パラメータが計算される。

[0379] その他、ユーザ位置情報とオブジェクトの位置情報がビットアロケーション部４２にも供給され、それらのユーザ位置情報やオブジェクトの位置情報がビットアロケーションに用いられるようにしてもよい。

[0380] ここで、聴覚心理パラメータ計算部４１での聴覚心理パラメータの計算、およびビットアロケーション部４２でのビットアロケーションの例について説明する。特に、ここではコンテンツが音楽ライブコンテンツである例について説明する。

[0381] この場合、ユーザは仮想のライブ会場でコンテンツの音を聴取するが、ライブ会場の最前列と最後列とでは聞こえる音が大きく異なる。

[0382] そこで、例えば自由視点において、ユーザが最前列のオブジェクトと近い位置でコンテンツの音を聴取する場合には、複数のオブジェクトに同じラベル情報が付与されていても、均一にではなく、ユーザから近い位置にあるオブジェクトに優先的に量子化ビットが割り当てられるようにされる。このようにすることで、ユーザに対して、よりオブジェクトの近くにいるかのようなリアリティ、すなわち高い臨場感を与えることができる。

[0383] 逆に、ユーザが最後列のオブジェクトから遠い位置でコンテンツの音を聴

取する場合には、もとの楽器種別ごとの調整、つまりラベル情報に応じた調整パラメータやアルゴリズムに対して、さらに遠距離用の調整が行われるようにしてもよい。

[0384] 例えば、アタック音や持続音に多くのビットを割り当てた方がよい楽器の音でも、信号のディケイや残響、余韻の部分にビットを多く割り当てることで、空間感を向上させ、広い会場にいるかのような臨場感をユーザに対して与えることができる。

[0385] このように、ラベル情報だけでなく、３次元空間内におけるユーザの位置、つまりユーザ位置情報により示される聴取位置や、ユーザとオブジェクトとの間の距離に応じて、聴覚心理パラメータの計算やビットアロケーションを行うことで、さらに臨場感を向上させることができる。

[0386] 〈符号化処理の説明〉

次に、図２８に示した符号化装置７１の動作について説明する。すなわち、以下、図２９のフローチャートを参照して、図２８の符号化装置７１による符号化処理について説明する。

[0387] ステップＳ２８１においてメタエンコーダ１１の量子化部２１は、供給されたメタデータとしての各パラメータを量子化し、その結果得られた量子化パラメータを符号化部２２に供給する。

[0388] なお、ステップＳ２８１では、図２７のステップＳ２５１と同様の処理が行われるが、量子化部２１は、供給されたユーザ位置情報とオブジェクトの位置情報に基づいて、ユーザから見たオブジェクトの相対的な位置を示す位置情報、または、適宜修正されたオブジェクトの絶対的な位置を示す位置情報などを、オブジェクトのメタデータを構成する位置情報として量子化する。

[0389] ステップＳ２８１の処理が行われると、その後、ステップＳ２８２乃至ステップＳ２８７の処理が行われて符号化処理は終了するが、これらの処理は、図２７のステップＳ２５２乃至ステップＳ２５７の処理と同様であるので、その説明は省略する。

[0390] 但し、ステップＳ２８４では、上述したように、ラベル情報だけでなくユーザ位置情報やオブジェクトの位置情報も用いられて聴覚心理パラメータが計算される。また、ステップＳ２８５においてもユーザ位置情報やオブジェクトの位置情報が用いられてビットアロケーションが行われるようにしてもよい。

[0391] 以上のようにして符号化装置７１は、ラベル情報だけでなく、ユーザ位置情報やオブジェクトの位置情報も用いて聴覚心理パラメータの計算やビットアロケーションを行う。このようにすることで、符号化効率や量子化演算の処理速度を向上させることができるとともに、臨場感を向上させ、より高音質なオーディオ再生を実現することができる。

[0392] 以上のように、本技術においては視聴時のレンダリングで適用されるメタデータのゲイン値やオブジェクトの位置等を考慮しているため、実際の聴感に即した聴覚心理パラメータの計算とビットアロケーションを行うことができ、符号化効率を向上させることができる。

[0393] また、コンテンツ制作者が作成したメタデータのゲイン値がMPEG-Hの仕様の範囲外である場合でも、実質的にゲイン値が仕様範囲の上限値や下限値に制限されてしまうことがなく、量子化による音質劣化を除けば、製作者の意図通りのレンダリング音を再現することができる。

[0394] 例えば、あるオブジェクトのオーディオ信号が他のオブジェクトと同程度の利得をもっていて、かつメタデータのゲイン値が０（－∞dB）でノイズゲートを意図しているような場合がある。そのような場合、実際にレンダリングされて視聴するオーディオ信号はゼロデータにもかかわらず、一般的な符号化装置では他のオブジェクトと同様にビットが割り当てられてしまうところ、本技術では、ゼロデータとしてビットアロケーションが行われるため、量子化ビットを大幅に削減することができる。

[0395] 〈コンピュータの構成例〉

ところで、上述した一連の処理は、ハードウェアにより実行することもできるし、ソフトウェアにより実行することもできる。一連の処理をソフトウ

ェアにより実行する場合には、そのソフトウェアを構成するプログラムが、コンピュータにインストールされる。ここで、コンピュータには、専用のハードウェアに組み込まれているコンピュータや、各種のプログラムをインストールすることで、各種の機能を実行することが可能な、例えば汎用のパーソナルコンピュータなどが含まれる。

[0396] 図３０は、上述した一連の処理をプログラムにより実行するコンピュータのハードウェアの構成例を示すブロック図である。

[0397] コンピュータにおいて、CPU (Central Processing Unit) ５０１, ROM (Read Only Memory) ５０２, RAM (Random Access Memory) ５０３は、バス５０４により相互に接続されている。

[0398] バス５０４には、さらに、入出力インターフェース５０５が接続されている。入出力インターフェース５０５には、入力部５０６、出力部５０７、記録部５０８、通信部５０９、及びドライブ５１０が接続されている。

[0399] 入力部５０６は、キーボード、マウス、マイクロフォン、撮像素子などよりなる。出力部５０７は、ディスプレイ、スピーカなどよりなる。記録部５０８は、ハードディスクや不揮発性のメモリなどよりなる。通信部５０９は、ネットワークインターフェースなどよりなる。ドライブ５１０は、磁気ディスク、光ディスク、光磁気ディスク、又は半導体メモリなどのリムーバブル記録媒体５１１を駆動する。

[0400] 以上のように構成されるコンピュータでは、CPU５０１が、例えば、記録部５０８に記録されているプログラムを、入出力インターフェース５０５及びバス５０４を介して、RAM５０３にロードして実行することにより、上述した一連の処理が行われる。

[0401] コンピュータ（CPU５０１）が実行するプログラムは、例えば、パッケージメディア等としてのリムーバブル記録媒体５１１に記録して提供することができる。また、プログラムは、ローカルエリアネットワーク、インターネット、デジタル衛星放送といった、有線または無線の伝送媒体を介して提供することができる。

[0402] コンピュータでは、プログラムは、リムーバブル記録媒体５１１をドライブ５１０に装着することにより、入出力インターフェース５０５を介して、記録部５０８にインストールすることができる。また、プログラムは、有線または無線の伝送媒体を介して、通信部５０９で受信し、記録部５０８にインストールすることができる。その他、プログラムは、ROM５０２や記録部５０８に、あらかじめインストールしておくことができる。

[0403] なお、コンピュータが実行するプログラムは、本明細書で説明する順序に沿って時系列に処理が行われるプログラムであっても良いし、並列に、あるいは呼び出しが行われたとき等の必要なタイミングで処理が行われるプログラムであっても良い。

[0404] また、本技術の実施の形態は、上述した実施の形態に限定されるものではなく、本技術の要旨を逸脱しない範囲において種々の変更が可能である。

[0405] 例えば、本技術は、１つの機能をネットワークを介して複数の装置で分担、共同して処理するクラウドコンピューティングの構成をとることができる。

[0406] また、上述のフローチャートで説明した各ステップは、１つの装置で実行する他、複数の装置で分担して実行することができる。

[0407] さらに、１つのステップに複数の処理が含まれる場合には、その１つのステップに含まれる複数の処理は、１つの装置で実行する他、複数の装置で分担して実行することができる。

[0408] さらに、本技術は、以下の構成とすることも可能である。

[0409] （１）

オーディオオブジェクトのメタデータに含まれるゲイン値に基づいて、前記オーディオオブジェクトのオーディオ信号に対する補正を行う補正部と、

前記補正により得られた信号に基づいて聴覚心理パラメータを計算し、前記オーディオ信号を量子化する量子化部と

を備える信号処理装置。

（２）

前記補正部は、前記ゲイン値に基づいて時間領域で前記オーディオ信号を補正する

（１）に記載の信号処理装置。

（３）

前記補正部による補正により得られた補正オーディオ信号に対して時間周波数変換を行う時間周波数変換部をさらに備え、

前記量子化部は、前記時間周波数変換により得られた周波数スペクトル情報に基づいて前記聴覚心理パラメータを計算する

（２）に記載の信号処理装置。

（４）

前記オーディオ信号に対して時間周波数変換を行う時間周波数変換部をさらに備え、

前記補正部は、前記ゲイン値に基づいて、前記時間周波数変換により得られた周波数スペクトル情報を補正し、

前記量子化部は、前記補正部による補正により得られた補正周波数スペクトル情報に基づいて前記聴覚心理パラメータを計算する

（１）に記載の信号処理装置。

（５）

音の到来方向に関する聴覚特性に基づいて、前記ゲイン値を補正するゲイン補正部をさらに備え、

前記補正部は、補正された前記ゲイン値に基づいて前記オーディオ信号に対する補正を行う

（１）乃至（４）の何れか一項に記載の信号処理装置。

（６）

前記ゲイン補正部は、前記メタデータに含まれる位置情報により示される位置に対する前記聴覚特性に基づいて前記ゲイン値を補正する

（５）に記載の信号処理装置。

（７）

前記オーディオオブジェクトの位置と、前記オーディオオブジェクトの位置に対する、前記ゲイン値の前記聴覚特性に基づく補正を行うためのゲイン補正値とが対応付けられた聴覚特性テーブルを保持する聴覚特性テーブル保持部をさらに備える

（６）に記載の信号処理装置。

（８）

前記ゲイン補正部は、前記位置情報により示される位置に対応する前記ゲイン補正値が前記聴覚特性テーブルにない場合、前記聴覚特性テーブルにある複数の前記ゲイン補正値に基づいて補間処理を行い、前記位置情報により示される位置の前記ゲイン補正値を求める

（７）に記載の信号処理装置。

（９）

前記ゲイン補正部は、前記位置情報により示される位置近傍にある複数の各位置と対応付けられている前記ゲイン補正値に基づいて前記補間処理を行う

（８）に記載の信号処理装置。

（１０）

前記補間処理はVBAPを利用した補間処理である

（９）に記載の信号処理装置。

（１１）

前記聴覚特性テーブルにおいて、各位置に対して複数の周波数ごとの前記ゲイン補正値が対応付けられており、

前記ゲイン補正部は、前記位置情報により示される位置に対応する所定周波数の前記ゲイン補正値が前記聴覚特性テーブルにない場合、前記位置情報により示される位置に対応する、前記所定周波数近傍の他の複数の周波数の前記ゲイン補正値に基づいて前記補間処理を行い、前記位置情報により示される位置の前記所定周波数の前記ゲイン補正値を求める

（８）に記載の信号処理装置。

（１２）

前記聴覚特性テーブル保持部は、再生音圧ごとに前記聴覚特性テーブルを保持しており、

前記ゲイン補正部は、前記オーディオ信号の音圧に基づいて、前記ゲイン値の補正に用いる前記聴覚特性テーブルを切り替える

（８）に記載の信号処理装置。

（１３）

前記ゲイン補正部は、前記オーディオ信号の前記音圧に対応する前記聴覚特性テーブルが前記聴覚特性テーブル保持部に保持されていない場合、前記音圧近傍の他の複数の前記再生音圧の前記聴覚特性テーブルにおける、前記位置情報により示される位置に対応する前記ゲイン補正値に基づいて前記補間処理を行い、前記音圧に対応する前記位置情報により示される位置の前記ゲイン補正値を求める

（１２）に記載の信号処理装置。

（１４）

前記ゲイン補正部は、前記オーディオ信号の特性に応じて、前記ゲイン値を制限する

（７）乃至（１３）の何れか一項に記載の信号処理装置。

（１５）

前記ゲイン補正部は、前記位置情報により示される位置に対応する前記ゲイン補正値が前記聴覚特性テーブルにない場合、前記位置情報により示される位置に最も近い位置と対応付けられている前記ゲイン補正値により前記ゲイン値を補正する

（７）に記載の信号処理装置。

（１６）

前記ゲイン補正部は、前記位置情報により示される位置に対応する前記ゲイン補正値が前記聴覚特性テーブルにない場合、前記位置情報により示される位置近傍にある複数の各位置と対応付けられている前記ゲイン補正値の平

均値を、前記位置情報により示される位置の前記ゲイン補正値とする

（７）に記載の信号処理装置。

（１７）

信号処理装置が、

オーディオオブジェクトのメタデータに含まれるゲイン値に基づいて、前記オーディオオブジェクトのオーディオ信号に対する補正を行い、

前記補正により得られた信号に基づいて聴覚心理パラメータを計算し、前記オーディオ信号を量子化する

信号処理方法。

（１８）

オーディオオブジェクトのメタデータに含まれるゲイン値に基づいて、前記オーディオオブジェクトのオーディオ信号に対する補正を行い、

前記補正により得られた信号に基づいて聴覚心理パラメータを計算し、前記オーディオ信号を量子化する

ステップを含む処理をコンピュータに実行させるプログラム。

（１９）

オーディオオブジェクトのメタデータに含まれるゲイン値に基づいて、前記オーディオオブジェクトの前記ゲイン値およびオーディオ信号に対する修正を行う修正部と、

前記修正により得られた修正オーディオ信号を量子化する量子化部と

を備える信号処理装置。

（２０）

前記修正部は、前記ゲイン値が所定の範囲外の値である場合に前記修正を行う

（１９）に記載の信号処理装置。

（２１）

前記修正により得られた修正ゲイン値に基づいて、前記修正オーディオ信号に対する補正を行う補正部をさらに備え、

前記量子化部は、前記修正オーディオ信号に対する補正により得られた信号に基づいて前記修正オーディオ信号を量子化する

（１９）または（２０）に記載の信号処理装置。

（２２）

前記修正により得られた修正ゲイン値を含む前記メタデータを量子化および符号化するメタエンコーダと、

量子化された前記修正オーディオ信号を符号化する符号化部と、

符号化された前記メタデータと、符号化された前記修正オーディオ信号とを多重化する多重化部と

をさらに備える（１９）乃至（２１）の何れか一項に記載の信号処理装置。

（２３）

前記修正部は、前記ゲイン値と、前記修正により得られた修正ゲイン値との差分に基づいて前記オーディオ信号を修正する

（１９）乃至（２２）の何れか一項に記載の信号処理装置。

（２４）

信号処理装置が、

オーディオオブジェクトのメタデータに含まれるゲイン値に基づいて、前記オーディオオブジェクトの前記ゲイン値およびオーディオ信号に対する修正を行い、

前記修正により得られた修正オーディオ信号を量子化する

信号処理方法。

（２５）

オーディオオブジェクトのメタデータに含まれるゲイン値に基づいて、前記オーディオオブジェクトの前記ゲイン値およびオーディオ信号に対する修正を行い、

前記修正により得られた修正オーディオ信号を量子化する

ステップを含む処理をコンピュータに実行させるプログラム。

（２６）

オーディオオブジェクトのゲイン値と位置情報の少なくとも何れか一方を含むメタデータ、前記オーディオオブジェクトのオーディオ信号、および複数の前記オーディオオブジェクト間の聴覚マスキングに関する聴覚心理モデルに基づいて聴覚心理パラメータを計算し、前記聴覚心理パラメータに基づいて前記オーディオ信号を量子化する量子化部を備える

信号処理装置。

（２７）

前記オーディオ信号に対して時間周波数変換を行う時間周波数変換部をさらに備え、

前記量子化部は、前記時間周波数変換により得られた周波数スペクトル情報に基づいて前記聴覚心理パラメータを計算する

（２６）に記載の信号処理装置。

（２８）

前記量子化部は、処理対象の前記オーディオオブジェクトの前記メタデータおよび前記オーディオ信号と、他の前記オーディオオブジェクトの前記メタデータおよび前記オーディオ信号と、前記聴覚心理モデルとに基づいて前記聴覚心理パラメータを計算する

（２６）または（２７）に記載の信号処理装置。

（２９）

前記メタデータには、前記メタデータに含まれる、前記ゲイン値および前記位置情報を含む複数のパラメータの一部または全部の編集許可を示す編集許可情報が含まれており、

前記量子化部は、前記編集許可情報により編集が許可されていない前記パラメータ、前記オーディオ信号、および前記聴覚心理モデルに基づいて前記聴覚心理パラメータを計算する

（２６）乃至（２８）の何れか一項に記載の信号処理装置。

（３０）

信号処理装置が、

オーディオオブジェクトのゲイン値と位置情報の少なくとも何れか一方を含むメタデータ、前記オーディオオブジェクトのオーディオ信号、および複数の前記オーディオオブジェクト間の聴覚マスキングに関する聴覚心理モデルに基づいて聴覚心理パラメータを計算し、前記聴覚心理パラメータに基づいて前記オーディオ信号を量子化する

信号処理方法。

（３１）

オーディオオブジェクトのゲイン値と位置情報の少なくとも何れか一方を含むメタデータ、前記オーディオオブジェクトのオーディオ信号、および複数の前記オーディオオブジェクト間の聴覚マスキングに関する聴覚心理モデルに基づいて聴覚心理パラメータを計算し、前記聴覚心理パラメータに基づいて前記オーディオ信号を量子化する

ステップを含む処理をコンピュータに実行させるプログラム。

（３２）

オーディオオブジェクトのオーディオ信号、および前記オーディオオブジェクトの音源種別を示すラベル情報に基づいて、前記ラベル情報により示される前記音源種別に対して定められた調整パラメータとアルゴリズムの少なくとも一方により、前記オーディオ信号を量子化する量子化部を備える

信号処理装置。

（３３）

前記量子化部は、前記オーディオ信号および前記ラベル情報に基づいて聴覚心理パラメータを計算し、前記聴覚心理パラメータに基づいて前記オーディオ信号を量子化する

（３２）に記載の信号処理装置。

（３４）

前記量子化部は、前記ラベル情報に基づいて前記オーディオ信号のビットアロケーションおよび量子化を行う

（３２）または（３３）に記載の信号処理装置。

（３５）

前記ラベル情報に基づいて、前記ラベル情報により示される前記音源種別に対して定められた調整パラメータとアルゴリズムの少なくとも一方により、前記オーディオ信号に対して時間周波数変換を行う時間周波数変換部をさらに備え、

前記量子化部は、前記時間周波数変換により得られた周波数スペクトル情報に基づいて前記聴覚心理パラメータを計算し、前記周波数スペクトル情報を量子化する

（３２）乃至（３４）の何れか一項に記載の信号処理装置。

（３６）

前記ラベル情報には、前記オーディオ信号に基づく音の聴取環境を示す聴取環境情報がさらに含まれており、

前記量子化部は、前記ラベル情報により示される前記音源種別および前記聴取環境に対して定められた調整パラメータとアルゴリズムの少なくとも一方により、前記オーディオ信号を量子化する

（３２）乃至（３５）の何れか一項に記載の信号処理装置。

（３７）

前記量子化部は、前記オーディオオブジェクトの優先度に基づいて、前記ラベル情報により示される前記音源種別に対して定められた調整パラメータを調整する

（３２）乃至（３５）の何れか一項に記載の信号処理装置。

（３８）

前記量子化部は、ユーザの位置情報、前記オーディオオブジェクトの位置情報、前記オーディオ信号、および前記ラベル情報に基づいて前記オーディオ信号を量子化する

（３２）乃至（３５）の何れか一項に記載の信号処理装置。

（３９）

信号処理装置が、

オーディオオブジェクトのオーディオ信号、および前記オーディオオブジェクトの音源種別を示すラベル情報に基づいて、前記ラベル情報により示される前記音源種別に対して定められた調整パラメータとアルゴリズムの少なくとも一方により、前記オーディオ信号を量子化する

信号処理方法。

（４０）

オーディオオブジェクトのオーディオ信号、および前記オーディオオブジェクトの音源種別を示すラベル情報に基づいて、前記ラベル情報により示される前記音源種別に対して定められた調整パラメータとアルゴリズムの少なくとも一方により、前記オーディオ信号を量子化する

ステップを含む処理をコンピュータに実行させるプログラム。

符号の説明

[0410] １１　メタエンコーダ，　１２　コアエンコーダ，　３１　時間周波数変換部，　３２　量子化部，　３３　符号化部，　７１　符号化装置，　８１　多重化部，　９１　オーディオ信号補正部，　９２　時間周波数変換部，　１３１　MDCT係数補正部

請求の範囲

[請求項1] オーディオオブジェクトのメタデータに含まれるゲイン値に基づいて、前記オーディオオブジェクトのオーディオ信号に対する補正を行う補正部と、

前記補正により得られた信号に基づいて聴覚心理パラメータを計算し、前記オーディオ信号を量子化する量子化部と

を備える信号処理装置。

[請求項2] 前記補正部は、前記ゲイン値に基づいて時間領域で前記オーディオ信号を補正する

請求項１に記載の信号処理装置。

[請求項3] 前記補正部による補正により得られた補正オーディオ信号に対して時間周波数変換を行う時間周波数変換部をさらに備え、

前記量子化部は、前記時間周波数変換により得られた周波数スペクトル情報に基づいて前記聴覚心理パラメータを計算する

請求項２に記載の信号処理装置。

[請求項4] 前記オーディオ信号に対して時間周波数変換を行う時間周波数変換部をさらに備え、

前記補正部は、前記ゲイン値に基づいて、前記時間周波数変換により得られた周波数スペクトル情報を補正し、

前記量子化部は、前記補正部による補正により得られた補正周波数スペクトル情報に基づいて前記聴覚心理パラメータを計算する

請求項１に記載の信号処理装置。

[請求項5] 音の到来方向に関する聴覚特性に基づいて、前記ゲイン値を補正するゲイン補正部をさらに備え、

前記補正部は、補正された前記ゲイン値に基づいて前記オーディオ信号に対する補正を行う

請求項１に記載の信号処理装置。

[請求項6] 前記ゲイン補正部は、前記メタデータに含まれる位置情報により示

される位置に対する前記聴覚特性に基づいて前記ゲイン値を補正する

請求項５に記載の信号処理装置。

[請求項7] 前記オーディオオブジェクトの位置と、前記オーディオオブジェクトの位置に対する、前記ゲイン値の前記聴覚特性に基づく補正を行うためのゲイン補正値とが対応付けられた聴覚特性テーブルを保持する聴覚特性テーブル保持部をさらに備える

請求項６に記載の信号処理装置。

[請求項8] 前記ゲイン補正部は、前記位置情報により示される位置に対応する前記ゲイン補正値が前記聴覚特性テーブルにない場合、前記位置情報により示される位置近傍にある複数の各位置と対応付けられている前記ゲイン補正値に基づいて補間処理を行い、前記位置情報により示される位置の前記ゲイン補正値を求めるか、前記位置情報により示される位置に最も近い位置と対応付けられている前記ゲイン補正値を、前記位置情報により示される位置の前記ゲイン補正値とするか、または前記位置情報により示される位置近傍にある複数の各位置と対応付けられている前記ゲイン補正値の平均値を、前記位置情報により示される位置の前記ゲイン補正値とする

請求項７に記載の信号処理装置。

[請求項9] 前記補間処理はVBAPを利用した補間処理である

請求項８に記載の信号処理装置。

[請求項10] 信号処理装置が、

オーディオオブジェクトのメタデータに含まれるゲイン値に基づいて、前記オーディオオブジェクトのオーディオ信号に対する補正を行い、

前記補正により得られた信号に基づいて聴覚心理パラメータを計算し、前記オーディオ信号を量子化する

信号処理方法。

[請求項11] オーディオオブジェクトのメタデータに含まれるゲイン値に基づい

て、前記オーディオオブジェクトのオーディオ信号に対する補正を行い、

前記補正により得られた信号に基づいて聴覚心理パラメータを計算し、前記オーディオ信号を量子化する

ステップを含む処理をコンピュータに実行させるプログラム。

[請求項12] オーディオオブジェクトのメタデータに含まれるゲイン値に基づいて、前記オーディオオブジェクトの前記ゲイン値およびオーディオ信号に対する修正を行う修正部と、

前記修正により得られた修正オーディオ信号を量子化する量子化部と

を備える信号処理装置。

[請求項13] 前記修正部は、前記ゲイン値が所定の範囲外の値である場合に前記修正を行う

請求項１２に記載の信号処理装置。

[請求項14] 前記修正により得られた修正ゲイン値に基づいて、前記修正オーディオ信号に対する補正を行う補正部をさらに備え、

前記量子化部は、前記修正オーディオ信号に対する補正により得られた信号に基づいて前記修正オーディオ信号を量子化する

請求項１２に記載の信号処理装置。

[請求項15] 前記修正により得られた修正ゲイン値を含む前記メタデータを量子化および符号化するメタエンコーダと、

量子化された前記修正オーディオ信号を符号化する符号化部と、

符号化された前記メタデータと、符号化された前記修正オーディオ信号とを多重化する多重化部と

をさらに備える請求項１２に記載の信号処理装置。

[請求項16] 前記修正部は、前記ゲイン値と、前記修正により得られた修正ゲイン値との差分に基づいて前記オーディオ信号を修正する

請求項１２に記載の信号処理装置。

[請求項17] 信号処理装置が、

オーディオオブジェクトのメタデータに含まれるゲイン値に基づいて、前記オーディオオブジェクトの前記ゲイン値およびオーディオ信号に対する修正を行い、

前記修正により得られた修正オーディオ信号を量子化する

信号処理方法。

[請求項18] オーディオオブジェクトのメタデータに含まれるゲイン値に基づいて、前記オーディオオブジェクトの前記ゲイン値およびオーディオ信号に対する修正を行い、

前記修正により得られた修正オーディオ信号を量子化する

ステップを含む処理をコンピュータに実行させるプログラム。

[請求項19] オーディオオブジェクトのゲイン値と位置情報の少なくとも何れか一方を含むメタデータ、前記オーディオオブジェクトのオーディオ信号、および複数の前記オーディオオブジェクト間の聴覚マスキングに関する聴覚心理モデルに基づいて聴覚心理パラメータを計算し、前記聴覚心理パラメータに基づいて前記オーディオ信号を量子化する量子化部を備える

信号処理装置。

[請求項20] 前記オーディオ信号に対して時間周波数変換を行う時間周波数変換部をさらに備え、

前記量子化部は、前記時間周波数変換により得られた周波数スペクトル情報に基づいて前記聴覚心理パラメータを計算する

請求項１９に記載の信号処理装置。

[請求項21] 前記量子化部は、処理対象の前記オーディオオブジェクトの前記メタデータおよび前記オーディオ信号と、他の前記オーディオオブジェクトの前記メタデータおよび前記オーディオ信号と、前記聴覚心理モデルとに基づいて前記聴覚心理パラメータを計算する

請求項１９に記載の信号処理装置。

[請求項22] 前記メタデータには、前記メタデータに含まれる、前記ゲイン値および前記位置情報を含む複数のパラメータの一部または全部の編集許可を示す編集許可情報が含まれており、

前記量子化部は、前記編集許可情報により編集が許可されていない前記パラメータ、前記オーディオ信号、および前記聴覚心理モデルに基づいて前記聴覚心理パラメータを計算する

請求項１９に記載の信号処理装置。

[請求項23] 信号処理装置が、

オーディオオブジェクトのゲイン値と位置情報の少なくとも何れか一方を含むメタデータ、前記オーディオオブジェクトのオーディオ信号、および複数の前記オーディオオブジェクト間の聴覚マスキングに関する聴覚心理モデルに基づいて聴覚心理パラメータを計算し、前記聴覚心理パラメータに基づいて前記オーディオ信号を量子化する

信号処理方法。

[請求項24] オーディオオブジェクトのゲイン値と位置情報の少なくとも何れか一方を含むメタデータ、前記オーディオオブジェクトのオーディオ信号、および複数の前記オーディオオブジェクト間の聴覚マスキングに関する聴覚心理モデルに基づいて聴覚心理パラメータを計算し、前記聴覚心理パラメータに基づいて前記オーディオ信号を量子化する

ステップを含む処理をコンピュータに実行させるプログラム。

[請求項25] オーディオオブジェクトのオーディオ信号、および前記オーディオオブジェクトの音源種別を示すラベル情報に基づいて、前記ラベル情報により示される前記音源種別に対して定められた調整パラメータとアルゴリズムの少なくとも一方により、前記オーディオ信号を量子化する量子化部を備える

信号処理装置。

[請求項26] 前記量子化部は、前記オーディオ信号および前記ラベル情報に基づいて聴覚心理パラメータを計算し、前記聴覚心理パラメータに基づい

て前記オーディオ信号を量子化する

請求項２５に記載の信号処理装置。

[請求項27] 前記量子化部は、前記ラベル情報に基づいて前記オーディオ信号のビットアロケーションおよび量子化を行う

請求項２５に記載の信号処理装置。

[請求項28] 前記ラベル情報に基づいて、前記ラベル情報により示される前記音源種別に対して定められた調整パラメータとアルゴリズムの少なくとも一方により、前記オーディオ信号に対して時間周波数変換を行う時間周波数変換部をさらに備え、

前記量子化部は、前記時間周波数変換により得られた周波数スペクトル情報に基づいて前記聴覚心理パラメータを計算し、前記周波数スペクトル情報を量子化する

請求項２５に記載の信号処理装置。

[請求項29] 前記ラベル情報には、前記オーディオ信号に基づく音の聴取環境を示す聴取環境情報がさらに含まれており、

前記量子化部は、前記ラベル情報により示される前記音源種別および前記聴取環境に対して定められた調整パラメータとアルゴリズムの少なくとも一方により、前記オーディオ信号を量子化する

請求項２５に記載の信号処理装置。

[請求項30] 前記量子化部は、前記オーディオオブジェクトの優先度に基づいて、前記ラベル情報により示される前記音源種別に対して定められた調整パラメータを調整する

請求項２５に記載の信号処理装置。

[請求項31] 前記量子化部は、ユーザの位置情報、前記オーディオオブジェクトの位置情報、前記オーディオ信号、および前記ラベル情報に基づいて前記オーディオ信号を量子化する

請求項２５に記載の信号処理装置。

[請求項32] 信号処理装置が、

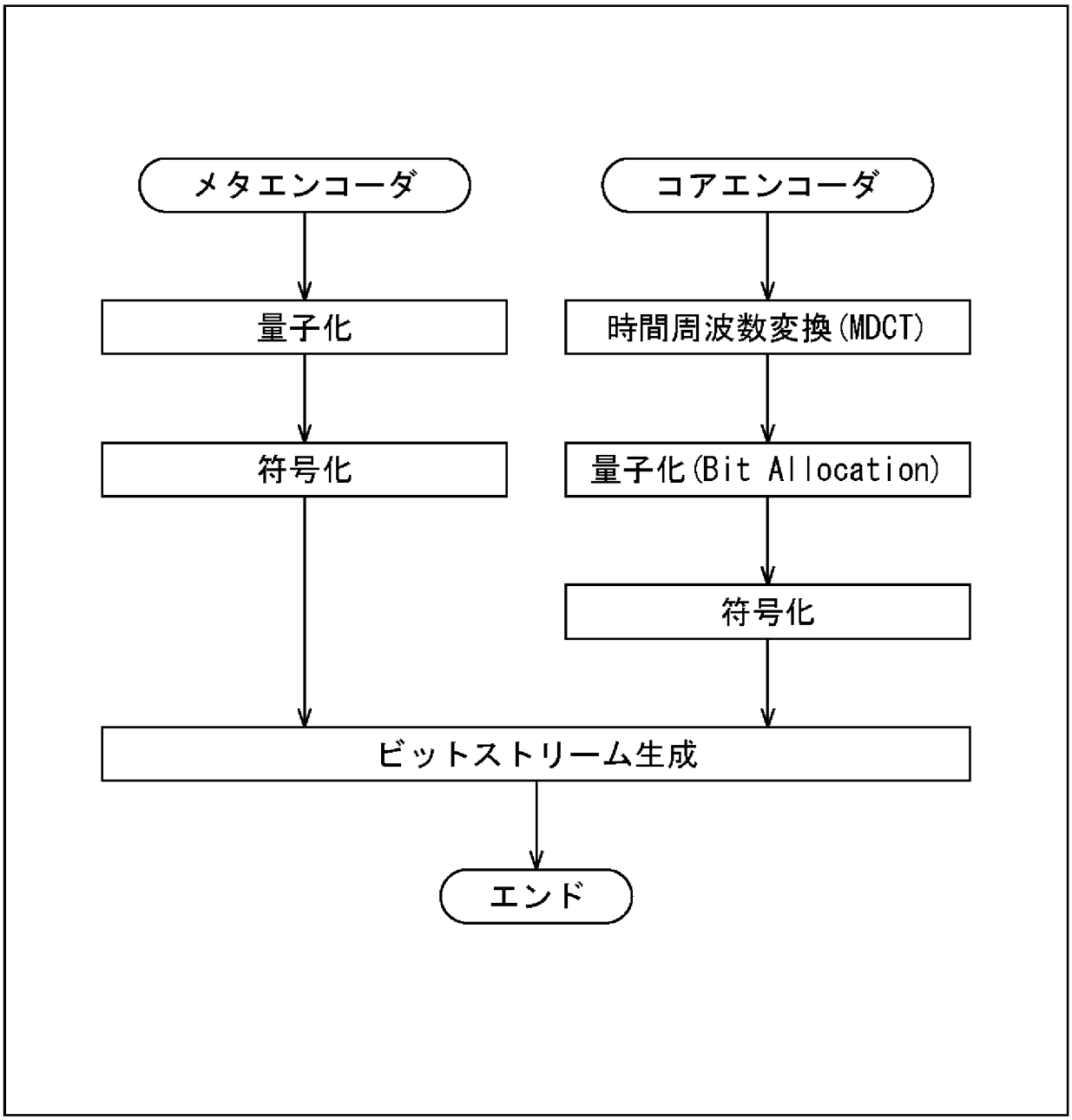
オーディオオブジェクトのオーディオ信号、および前記オーディオオブジェクトの音源種別を示すラベル情報に基づいて、前記ラベル情報により示される前記音源種別に対して定められた調整パラメータとアルゴリズムの少なくとも一方により、前記オーディオ信号を量子化する

信号処理方法。

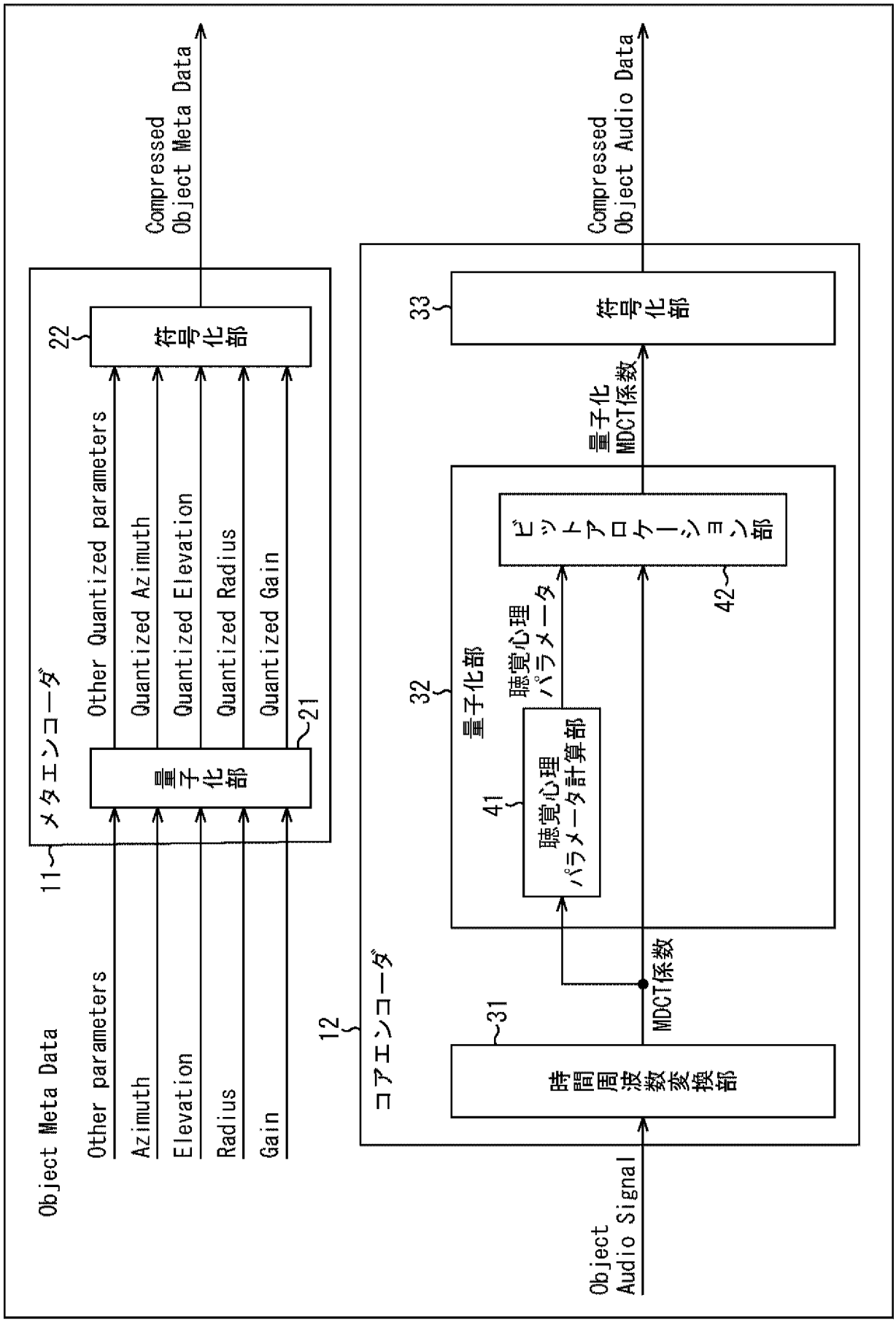
[請求項33] オーディオオブジェクトのオーディオ信号、および前記オーディオオブジェクトの音源種別を示すラベル情報に基づいて、前記ラベル情報により示される前記音源種別に対して定められた調整パラメータとアルゴリズムの少なくとも一方により、前記オーディオ信号を量子化する

ステップを含む処理をコンピュータに実行させるプログラム。

[図1]
FIG.1


メタエンコーダ
コアエンコーダ
量子化
時間周波数変換(MDCT)
符号化
量子化(Bit Allocation)
符号化
ビットストリーム生成
エンド

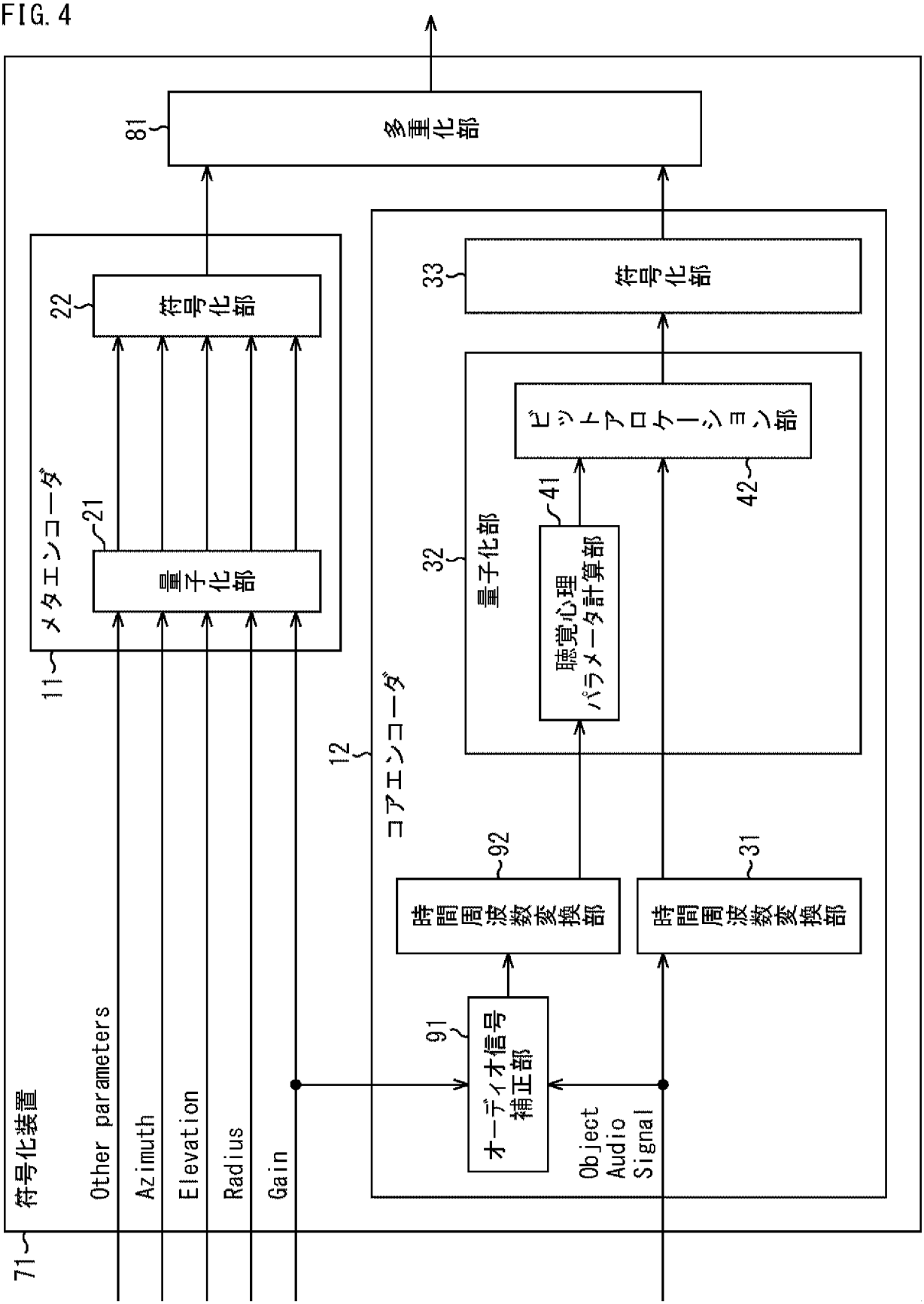
[図2]
FIG. 2


Object Meta Data
Other parameters
Azimuth
Elevation
Radius
Gain
11
メタエンコーダ
量子化部
21
Other Quantized parameters
Quantized Azimuth
Quantized Elevation
Quantized Radius
Quantized Gain
22
符号化部
Compressed Object Meta Data
Object Audio Signal
12
コアエンコーダ
時間周波数変換部
31
MDCT係数
32
量子化部
聴覚心理パラメータ
41
聴覚心理パラメータ計算部
ビットアロケーション部
42
量子化MDCT係数
33
符号化部
Compressed Object Audio Data

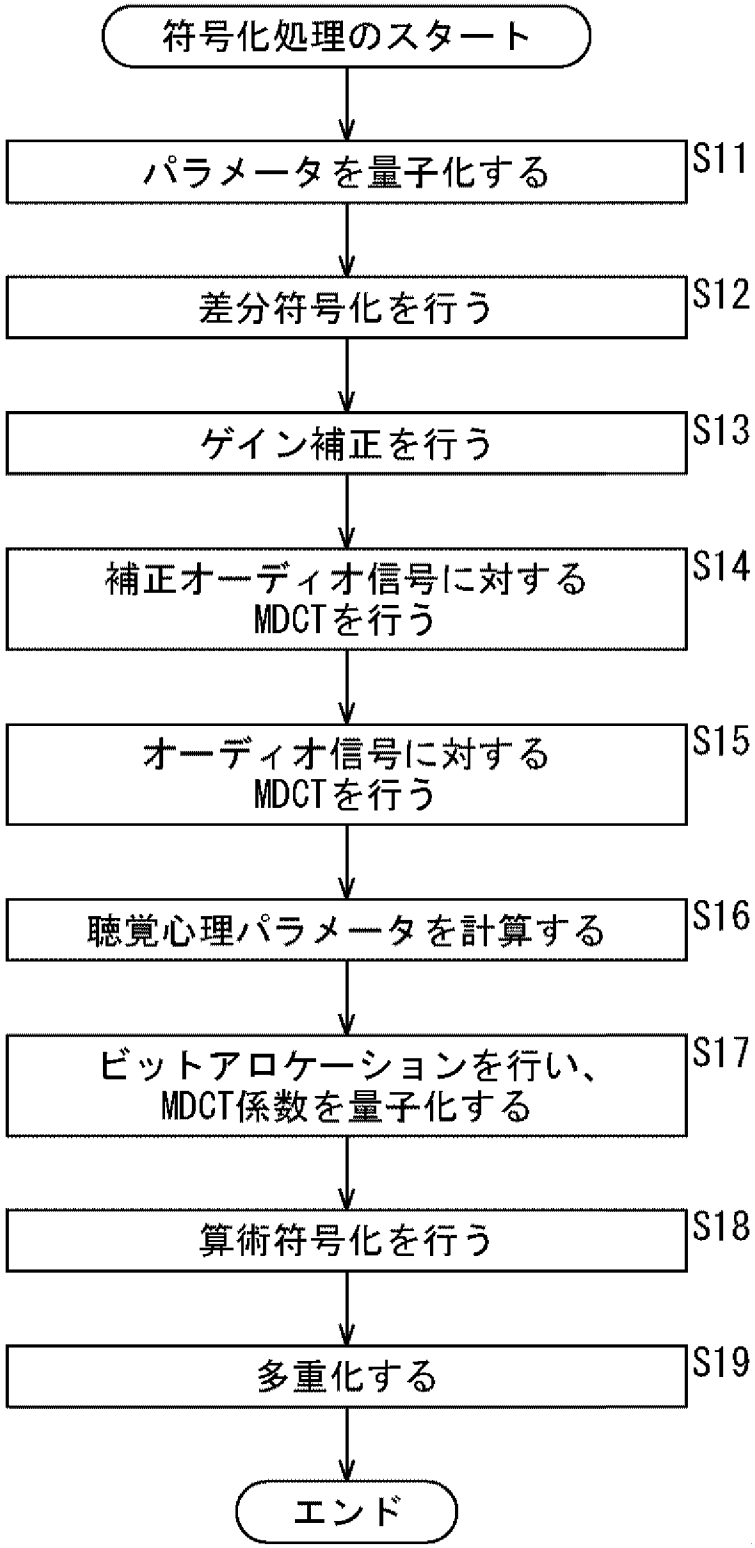
[図3]
FIG. 3

component	unit	value range
azimuth	° (degree)	-180:180
elevation	° (degree)	-90:90
radius	m(meter)	0.5:16
gain	none(linear)	0.004:5.957
spread(uniform)	° (degree)	0:180
spread width	° (degree)	0:180
spread height	° (degree)	0:90
spread depth	m(meter)	0:15.5
dynamic object priority	none	0:7

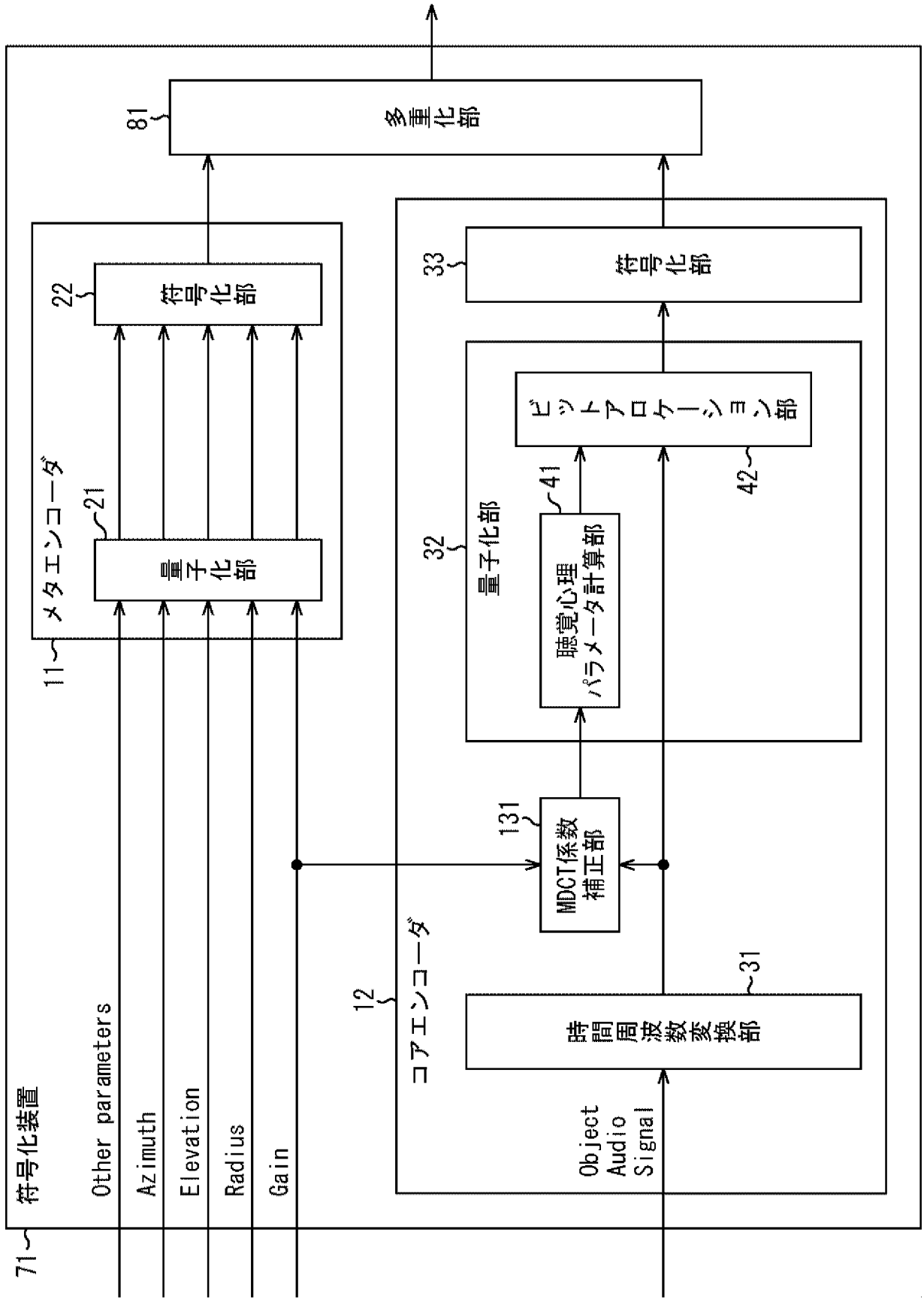
[図4]
FIG. 4


71
符号化装置
81
多重化部
11
メタエンコーダ
21
量子化部
22
符号化部
12
コアエンコーダ
32
量子化部
41
聴覚心理
パラメータ計算部
42
ビットアロケーション部
33
符号化部
91
オーディオ信号
補正部
92
時間周波数変換部
31
時間周波数変換部
Other parameters
Azimuth
Elevation
Radius
Gain
Object
Audio
Signal

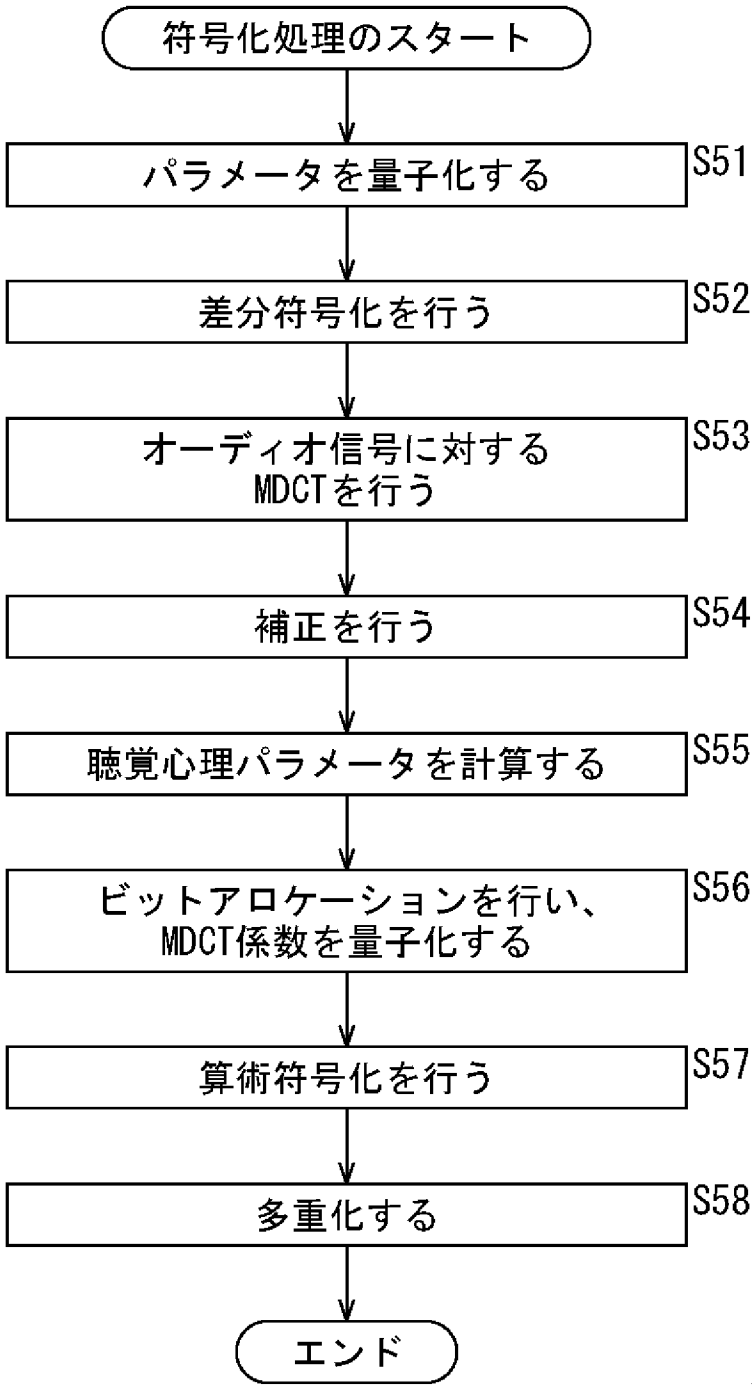
[図5]
FIG. 5


符号化処理のスタート
パラメータを量子化する
S11
差分符号化を行う
S12
ゲイン補正を行う
S13
補正オーディオ信号に対する
MDCTを行う
S14
オーディオ信号に対する
MDCTを行う
S15
聴覚心理パラメータを計算する
S16
ビットアロケーションを行い、
MDCT係数を量子化する
S17
算術符号化を行う
S18
多重化する
S19
エンド

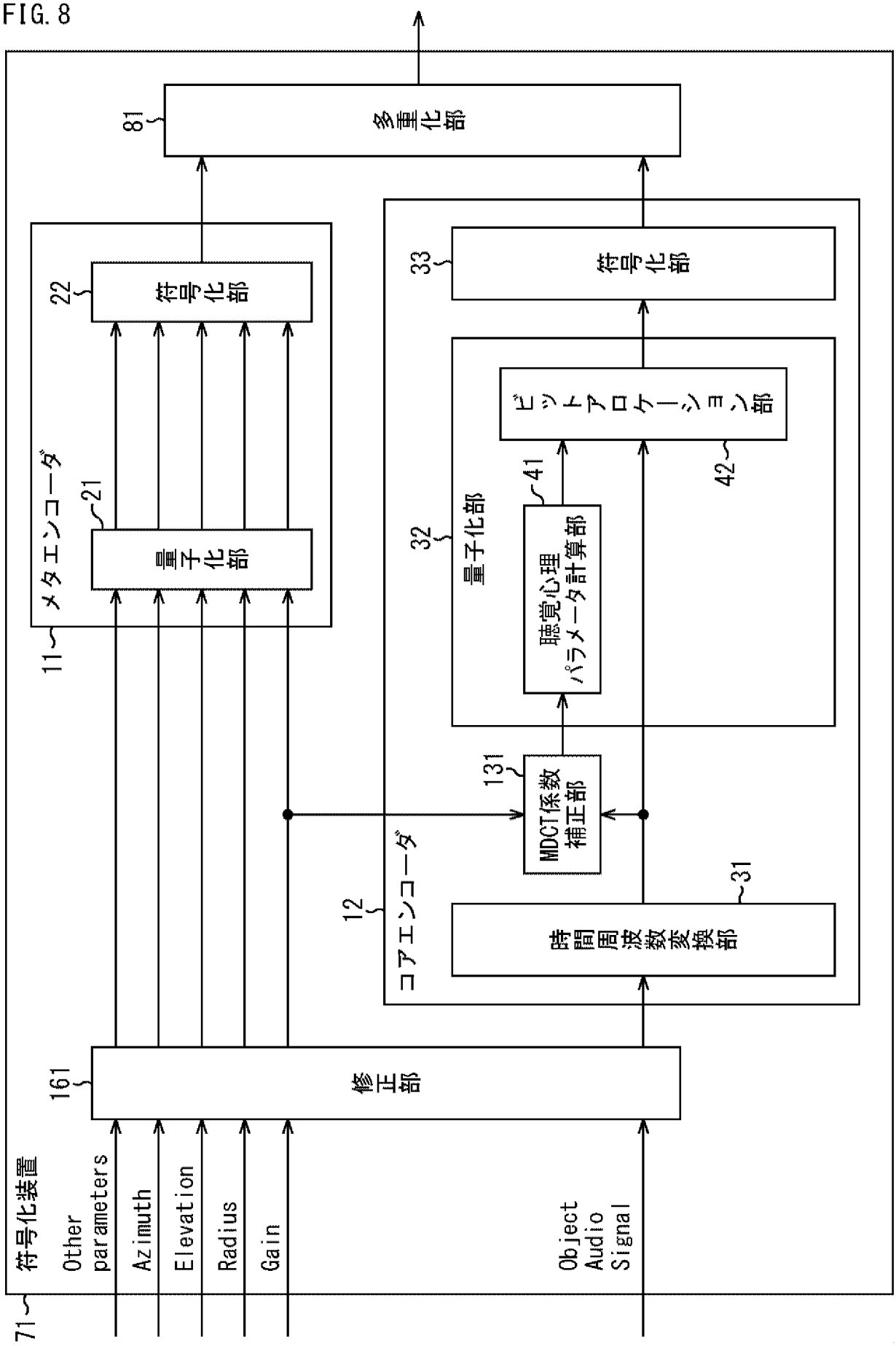
[図6]
FIG. 6


符号化装置
71
多重化部
81
メタエンコーダ
11
量子化部
21
符号化部
22
コアエンコーダ
12
時間周波数変換部
31
量子化部
32
聴覚心理パラメータ計算部
41
ビットアロケーション部
42
MDCT係数補正部
131
符号化部
33
Other parameters
Azimuth
Elevation
Radius
Gain
Object Audio Signal

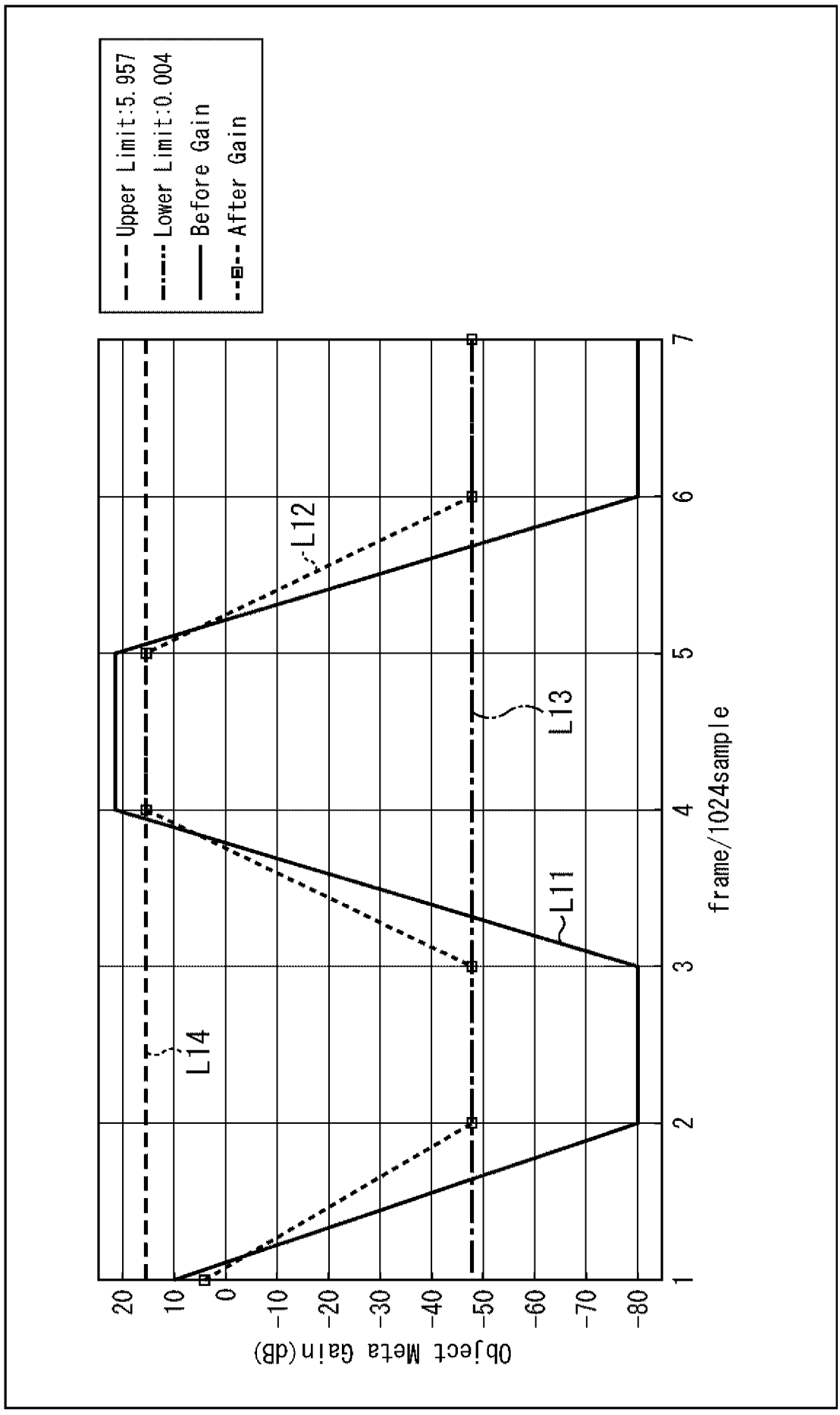
[図7]
FIG. 7


符号化処理のスタート
パラメータを量子化する
S51
差分符号化を行う
S52
オーディオ信号に対する
MDCTを行う
S53
補正を行う
S54
聴覚心理パラメータを計算する
S55
ビットアロケーションを行い、
MDCT係数を量子化する
S56
算術符号化を行う
S57
多重化する
S58
エンド

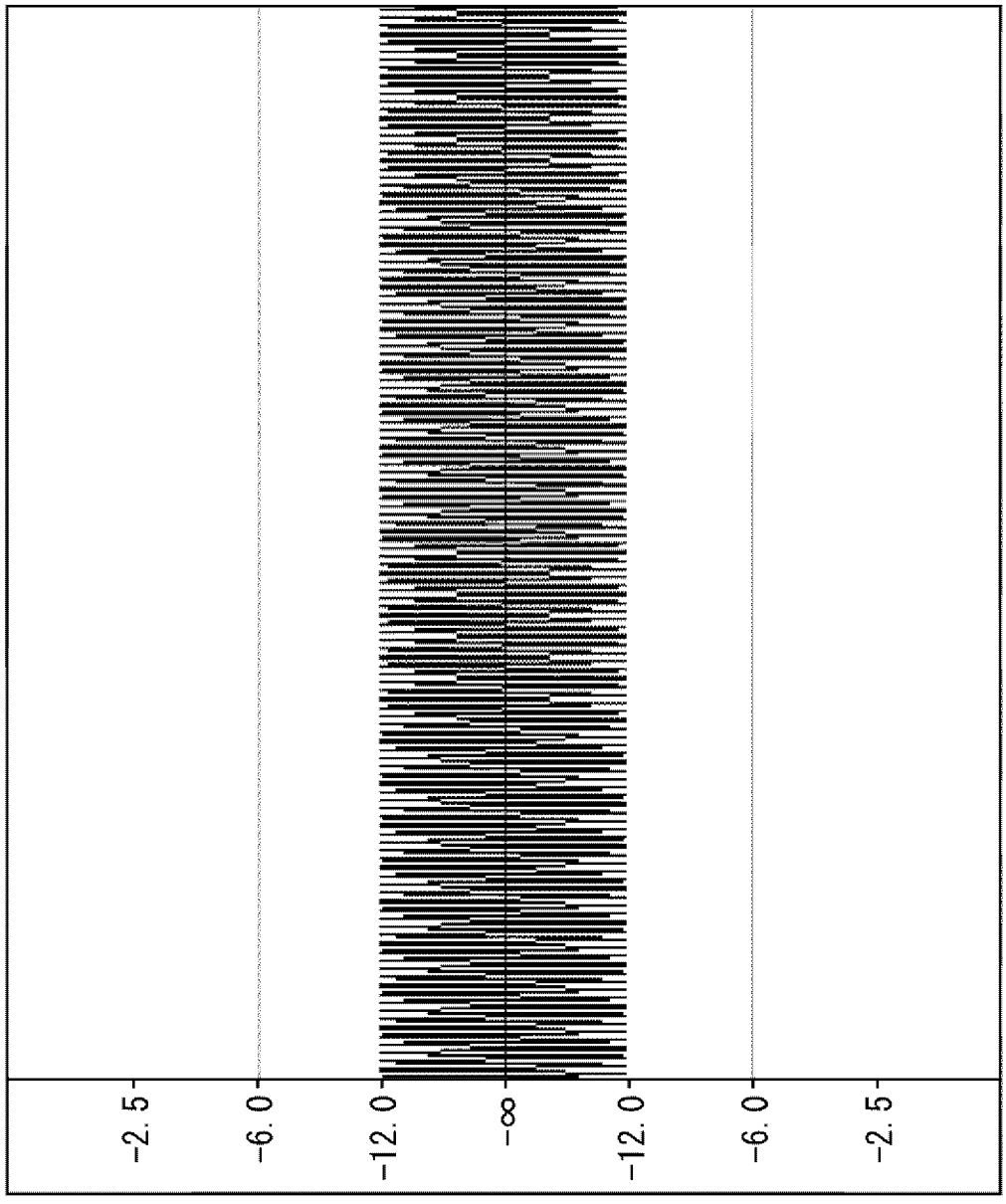
[図8]
FIG. 8


71
符号化装置
Other parameters
Azimuth
Elevation
Radius
Gain
Object Audio Signal
161
修正部
11
メタエンコーダ
21
量子化部
22
符号化部
12
コアエンコーダ
31
時間周波数変換部
131
MDCT係数補正部
32
量子化部
41
聴覚心理パラメータ計算部
42
ビットアロケーション部
33
符号化部
81
多重化部

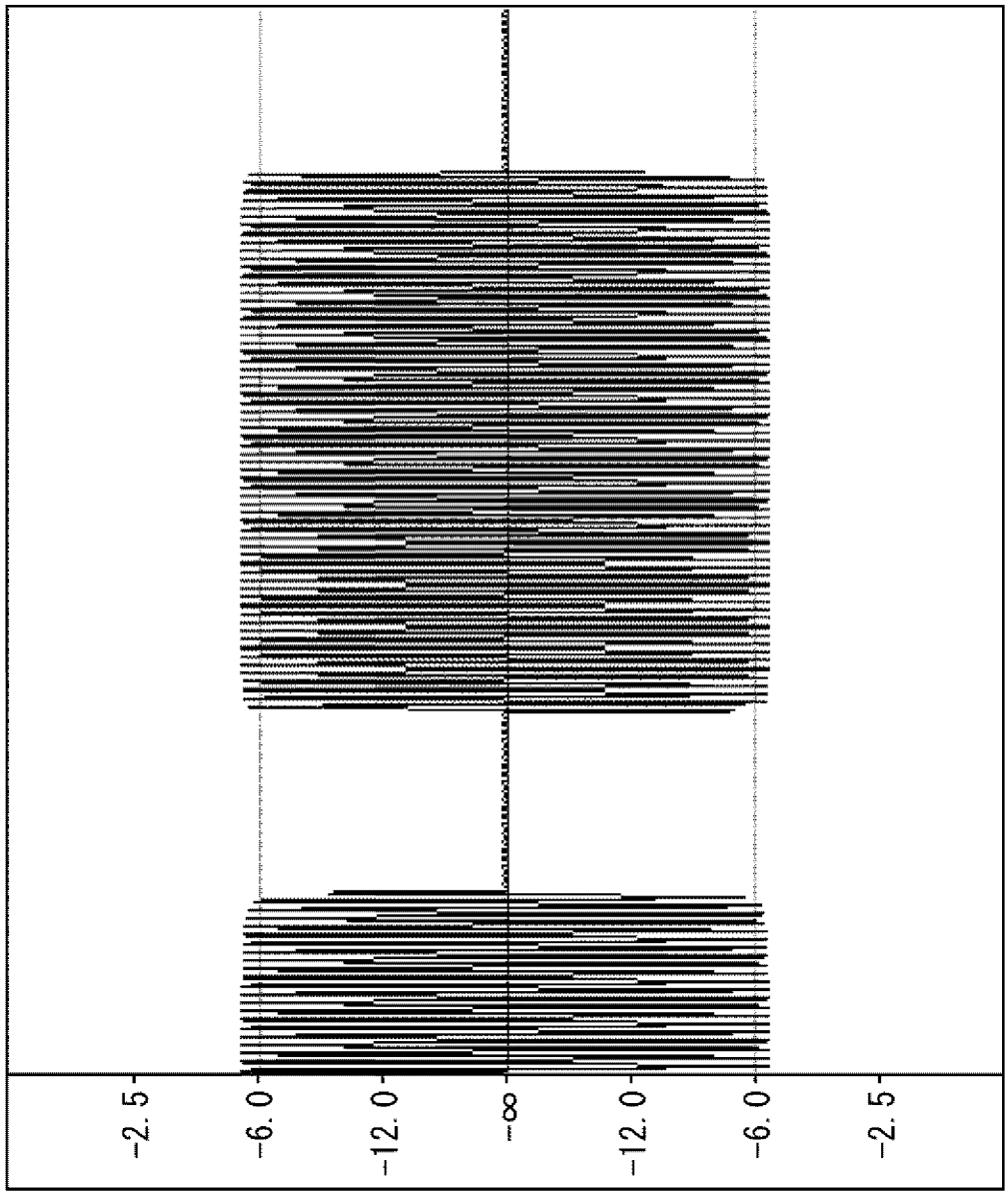
[図9]
FIG. 9


Upper Limit:5.957
Lower Limit:0.004
Before Gain
After Gain
L11
L12
L13
L14
Object Meta Gain(dB)
20
10
0
-10
-20
-30
-40
-50
-60
-70
-80
1
2
3
4
5
6
7
frame/1024sample

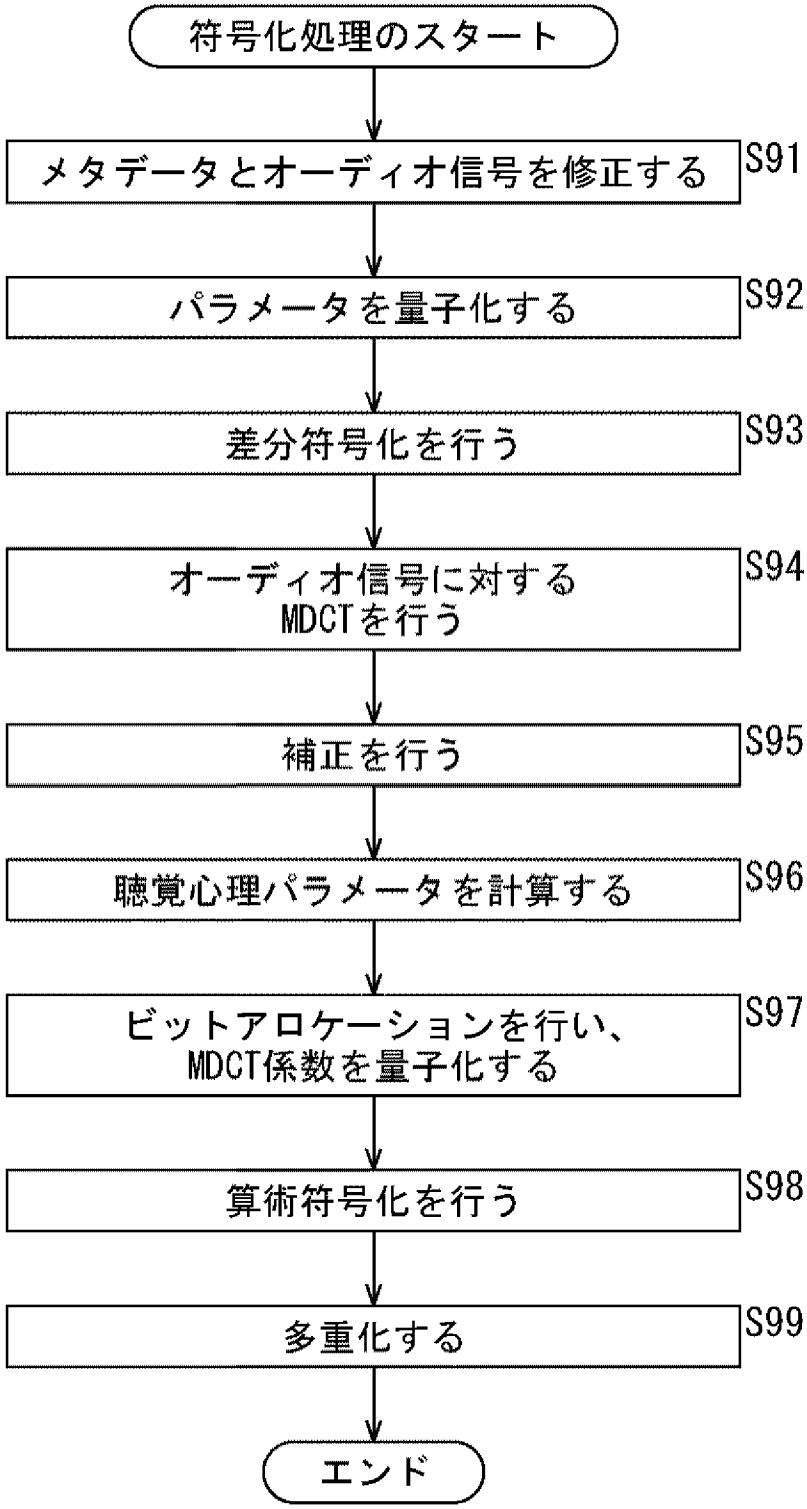
[図10]
FIG. 10



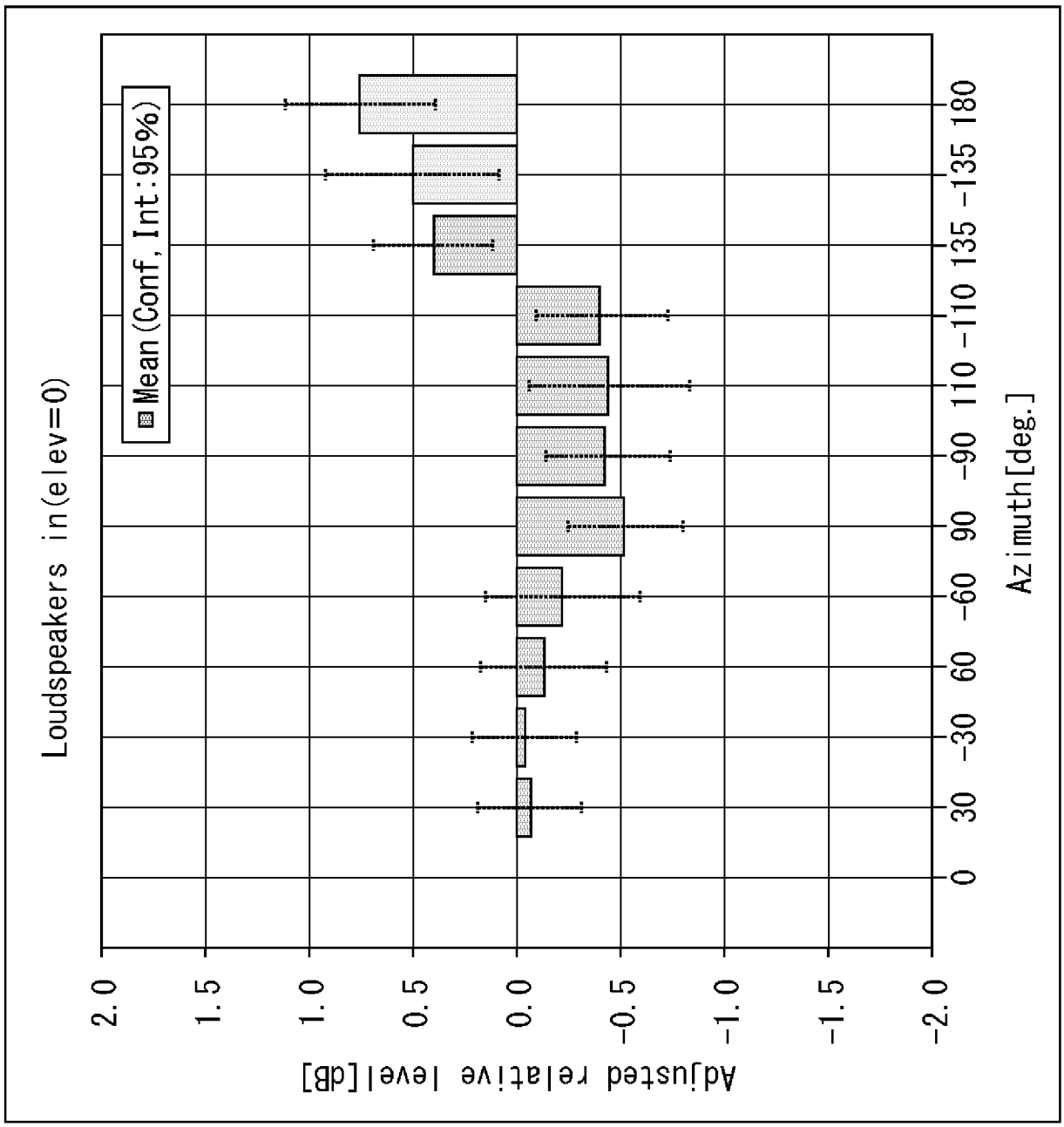
[図11]
FIG. 11



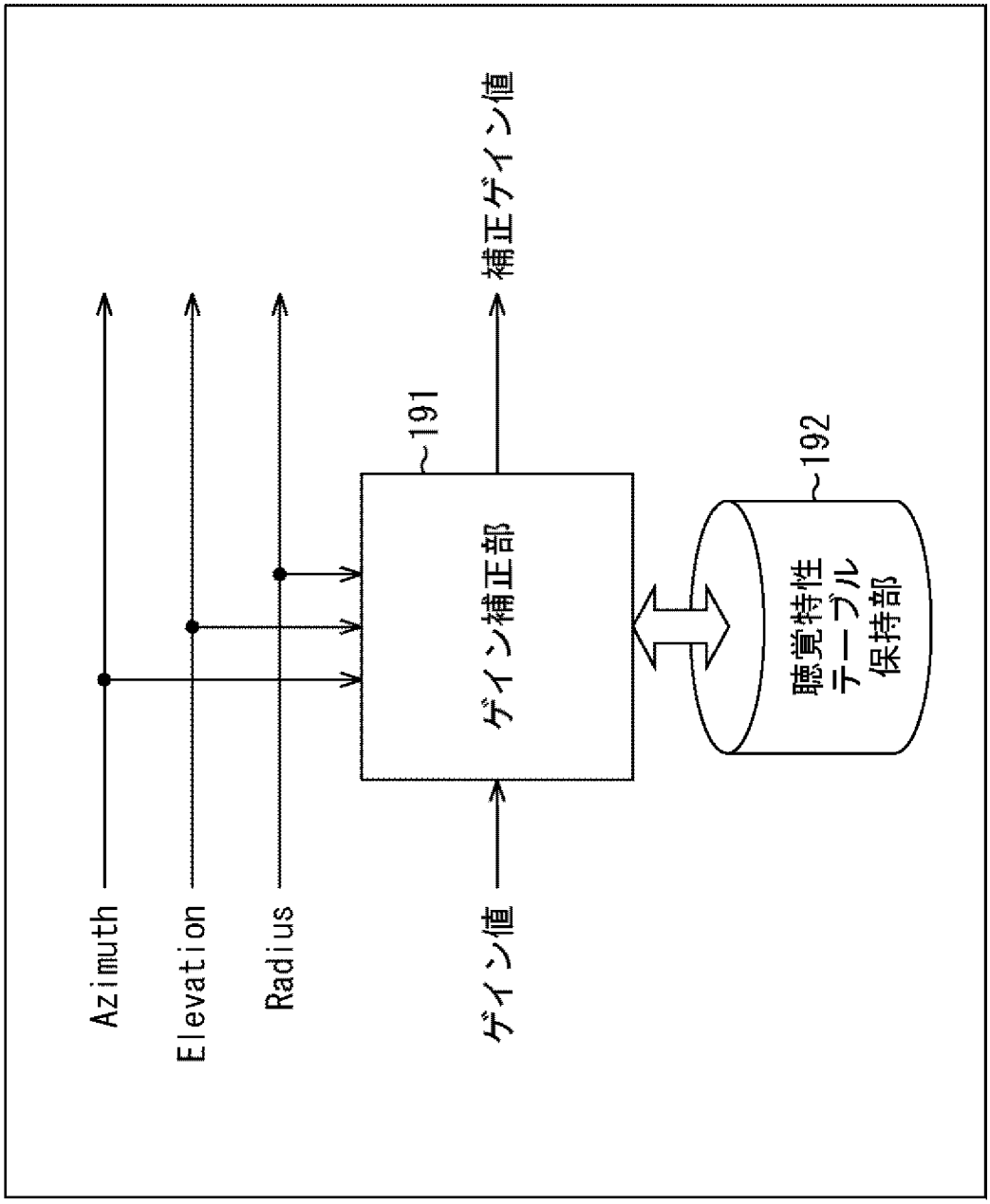
[図12]
FIG. 12


符号化処理のスタート
メタデータとオーディオ信号を修正する
S91
パラメータを量子化する
S92
差分符号化を行う
S93
オーディオ信号に対する
MDCTを行う
S94
補正を行う
S95
聴覚心理パラメータを計算する
S96
ビットアロケーションを行い、
MDCT係数を量子化する
S97
算術符号化を行う
S98
多重化する
S99
エンド

[図13]
FIG. 13


Loudspeakers in(elev=0)
Mean(Conf, Int:95%)
Adjusted relative level[dB]
2.0
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
-2.0
0
30
-30
60
-60
90
-90
110
-110
135
-135
180
Azimuth[deg.]

[図14]
FIG. 14


Azimuth
Elevation
Radius
ゲイン値
191
ゲイン補正部
補正ゲイン値
192
聴覚特性
テーブル
保持部

[図15]
FIG. 15

Azimuth [度]	Elevation [度]	Radius [m]	ゲイン補正値 [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

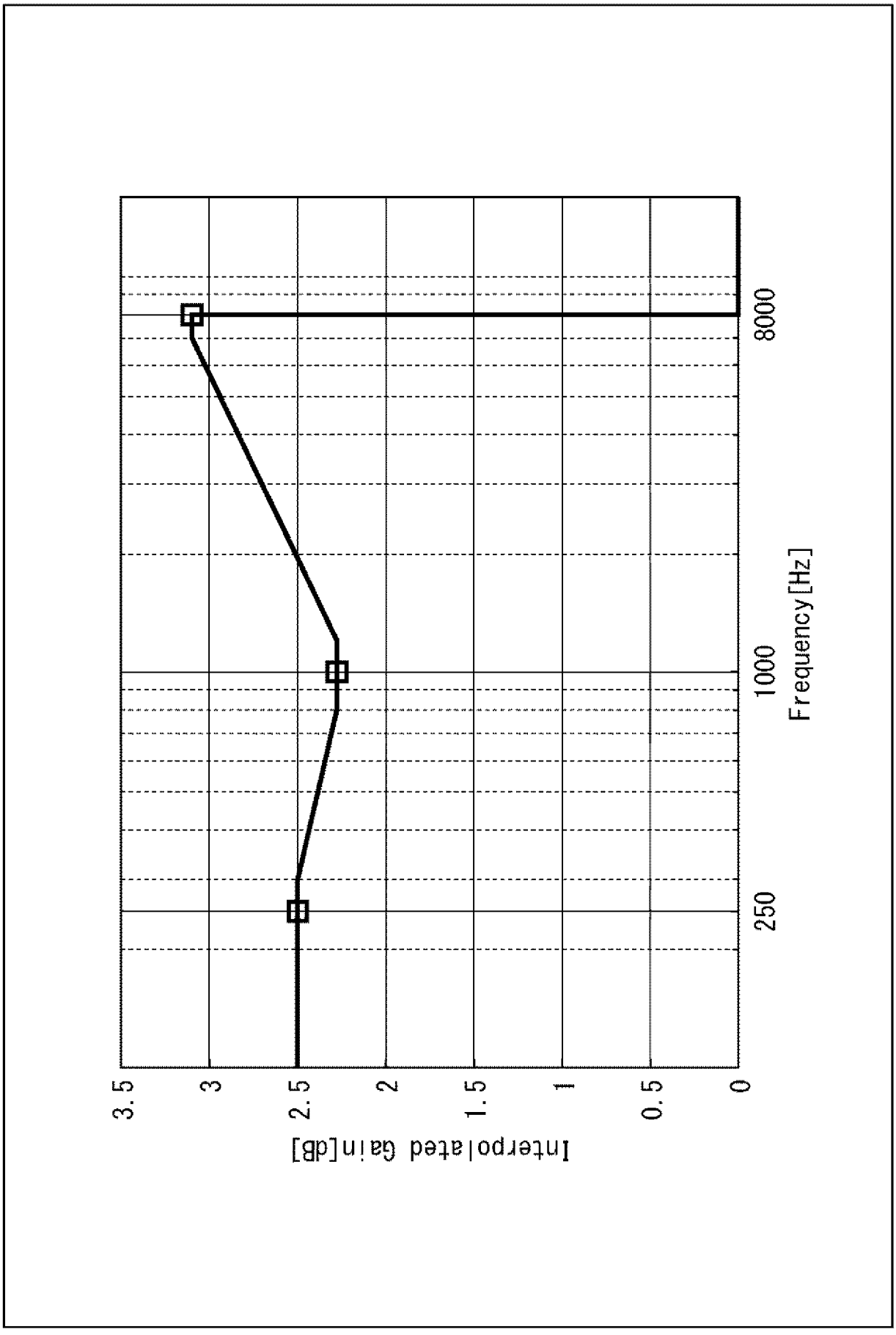
[図16]
FIG. 16

ゲイン補正値[dB]		Azimuth[度]											
		0	30	-30	60	-60	90	-90	120	-120	150	-150	180
Elevation[度]	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	

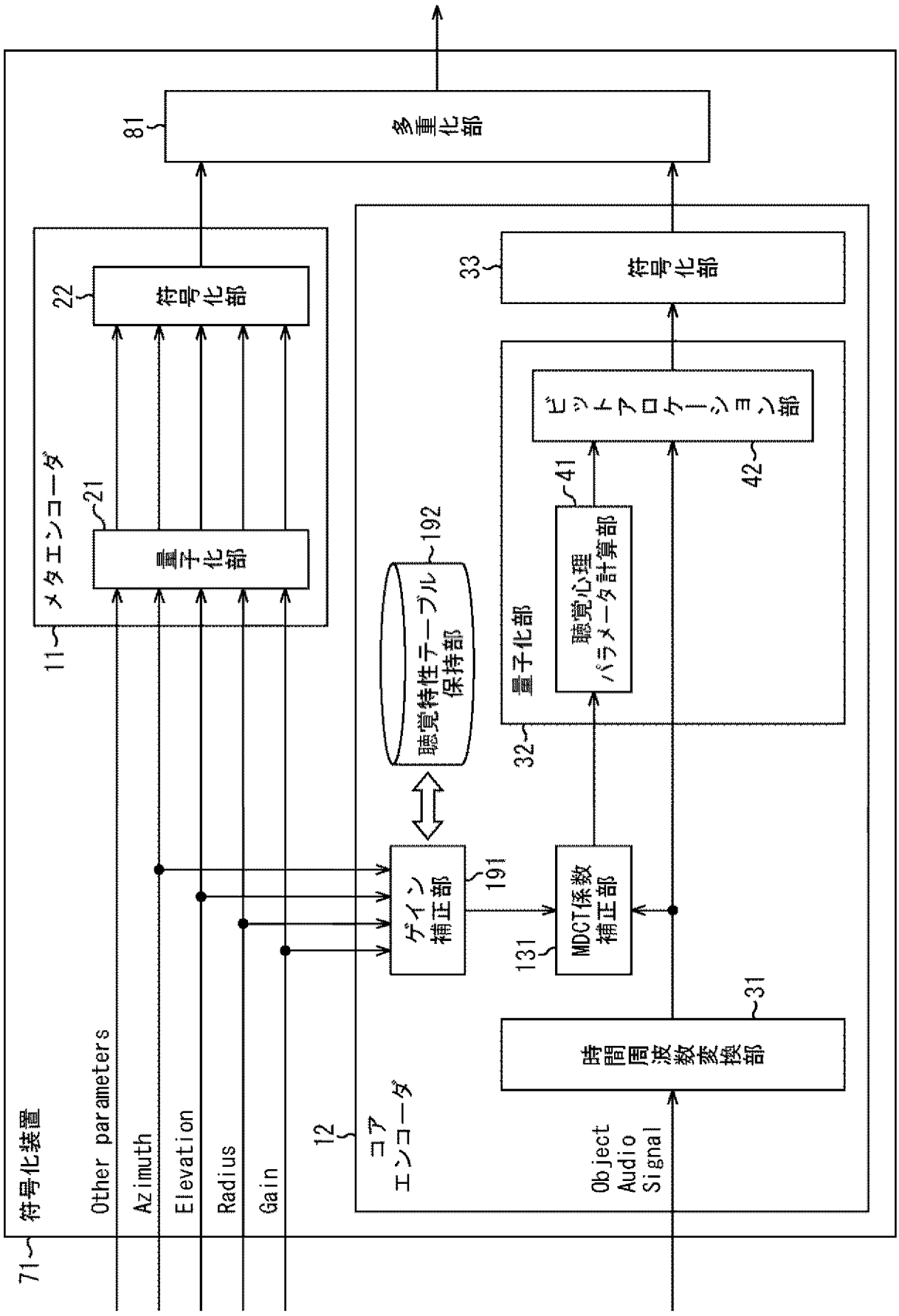
[図17]
FIG. 17

Azimuth	Elevation	ゲイン補正値		
		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

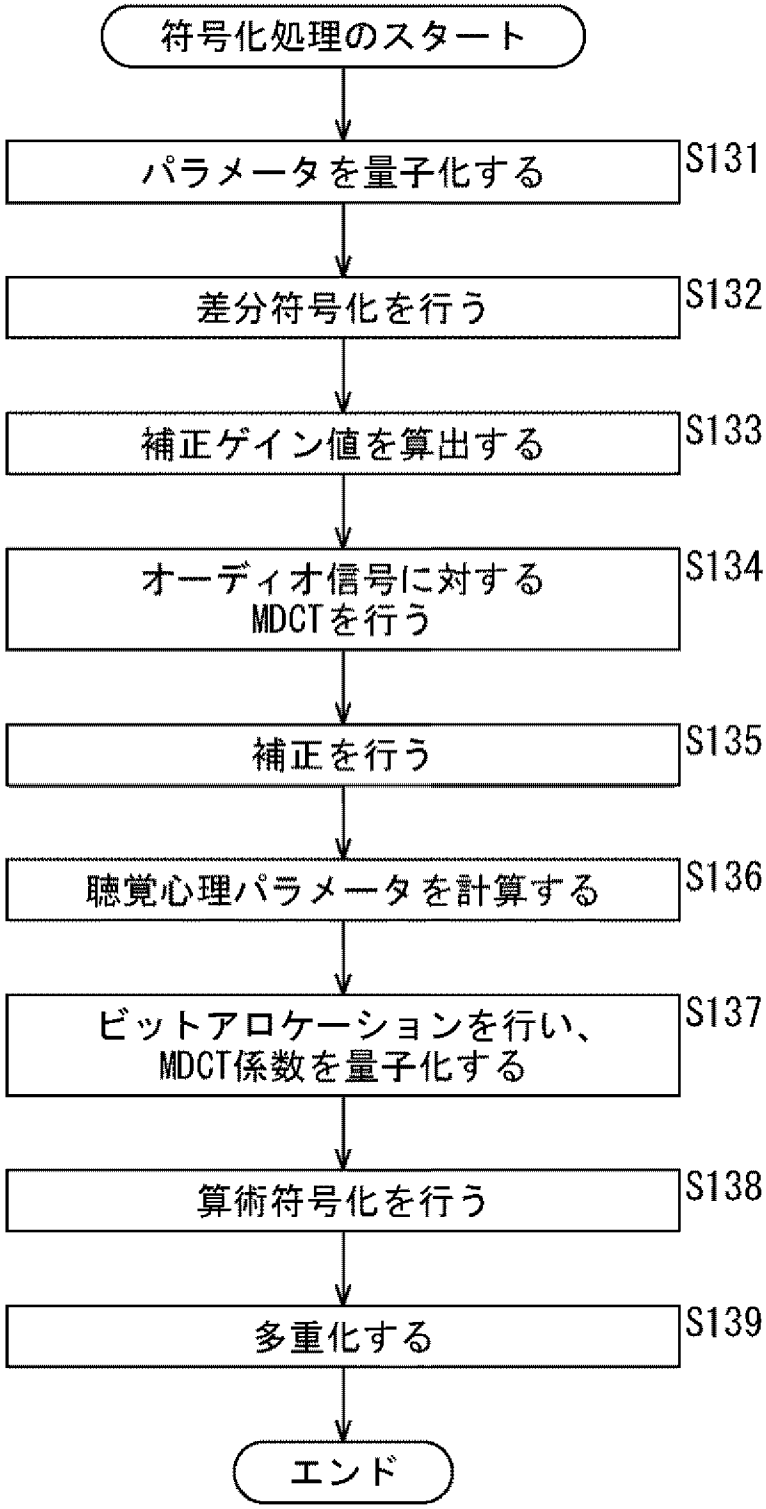
[図18]
FIG. 18


Interpolated Gain[dB]
3.5
3
2.5
2
1.5
1
0.5
0
Frequency[Hz]
250
1000
8000

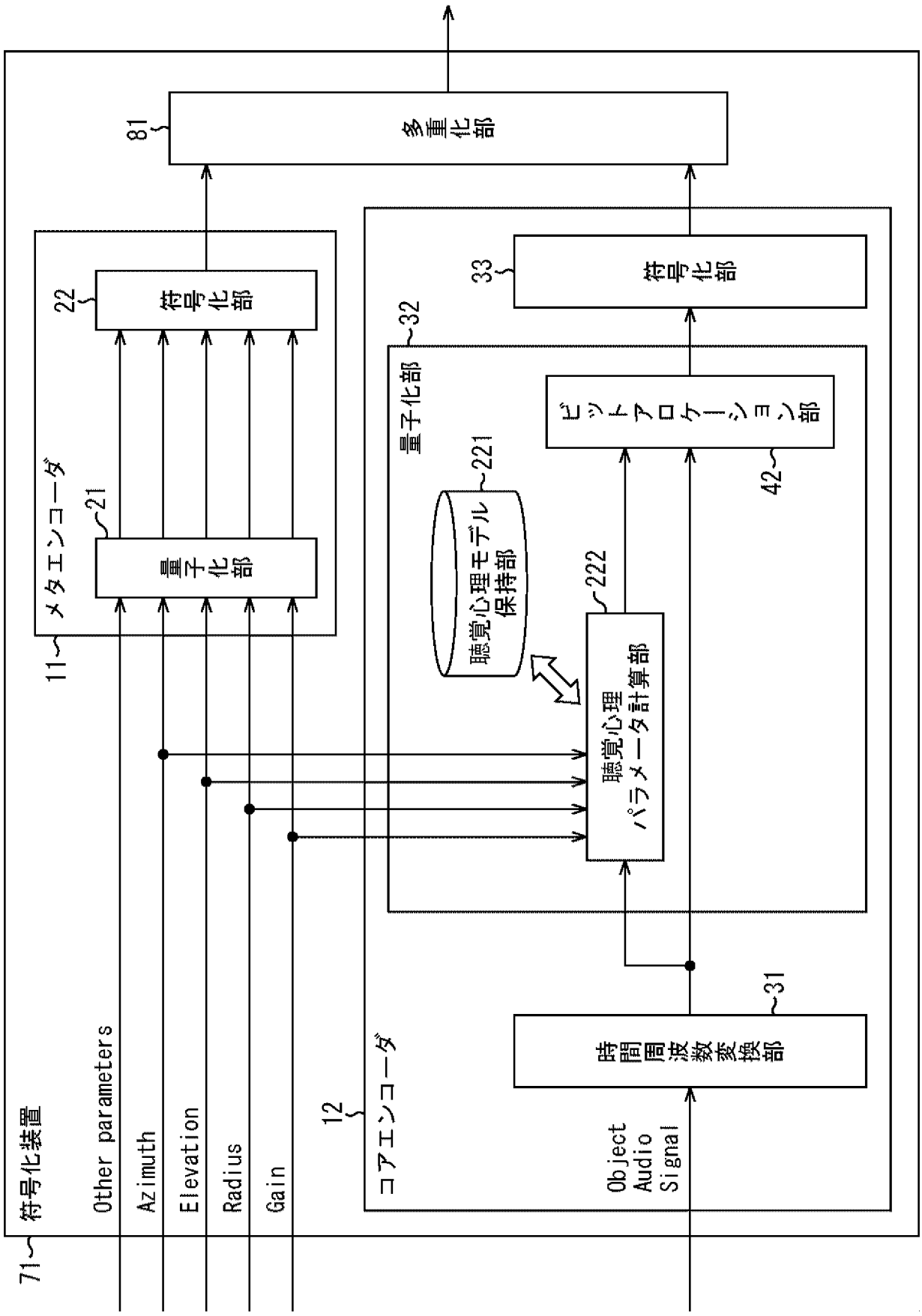
[図19]
FIG. 19


71 符号化装置
Other parameters
Azimuth
Elevation
Radius
Gain
Object Audio Signal
11 メタエンコーダ
21 量子化部
22 符号化部
12 コアエンコーダ
32 量子化部
聴覚特性テーブル保持部
192
ゲイン補正部
191
聴覚心理パラメータ計算部
41
MDCT係数補正部
131
ビットアロケーション部
42
33 符号化部
時間周波数変換部
31
81 多重化部

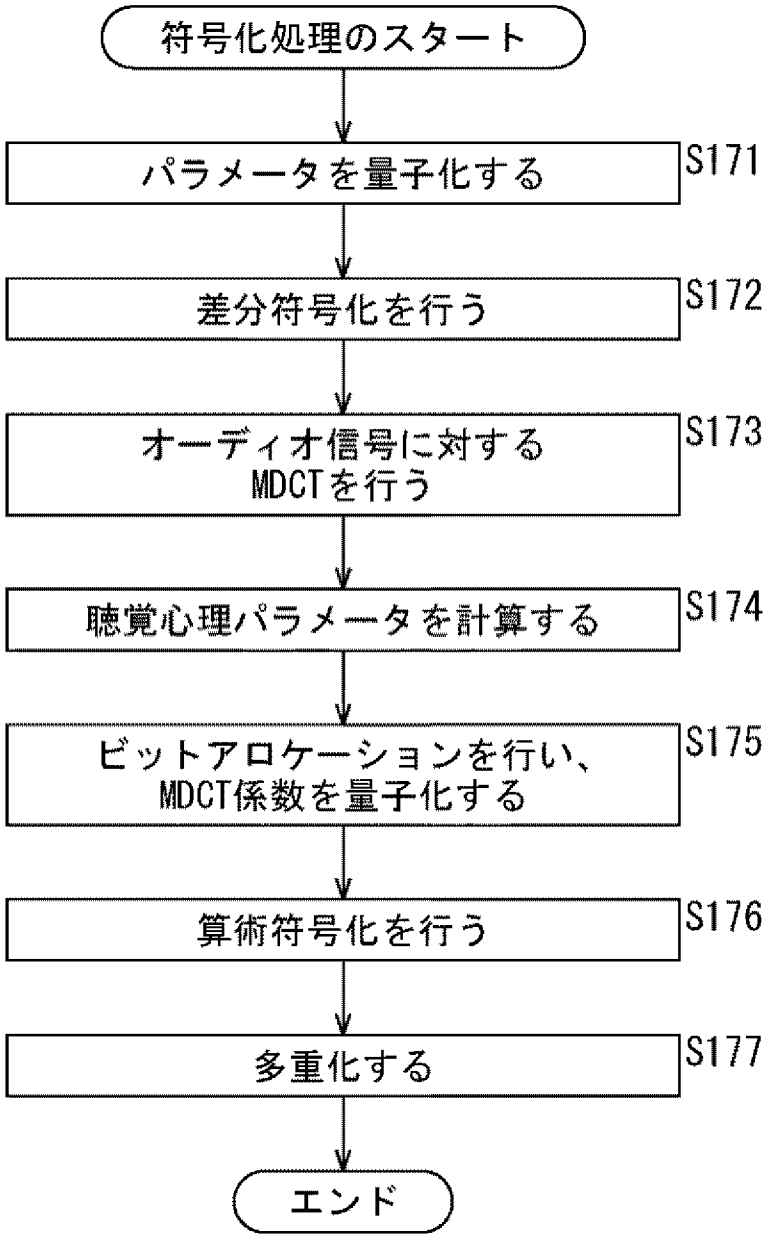
[図20]
FIG. 20


符号化処理のスタート
パラメータを量子化する
S131
差分符号化を行う
S132
補正ゲイン値を算出する
S133
オーディオ信号に対する
MDCTを行う
S134
補正を行う
S135
聴覚心理パラメータを計算する
S136
ビットアロケーションを行い、
MDCT係数を量子化する
S137
算術符号化を行う
S138
多重化する
S139
エンド

[図21]
FIG. 21


符号化装置
71
多重化部
81
メタエンコーダ
11
量子化部
21
符号化部
22
コアエンコーダ
12
量子化部
32
聴覚心理モデル保持部
221
聴覚心理パラメータ計算部
222
ビットアロケーション部
42
符号化部
33
時間周波数変換部
31
Other parameters
Azimuth
Elevation
Radius
Gain
Object Audio Signal

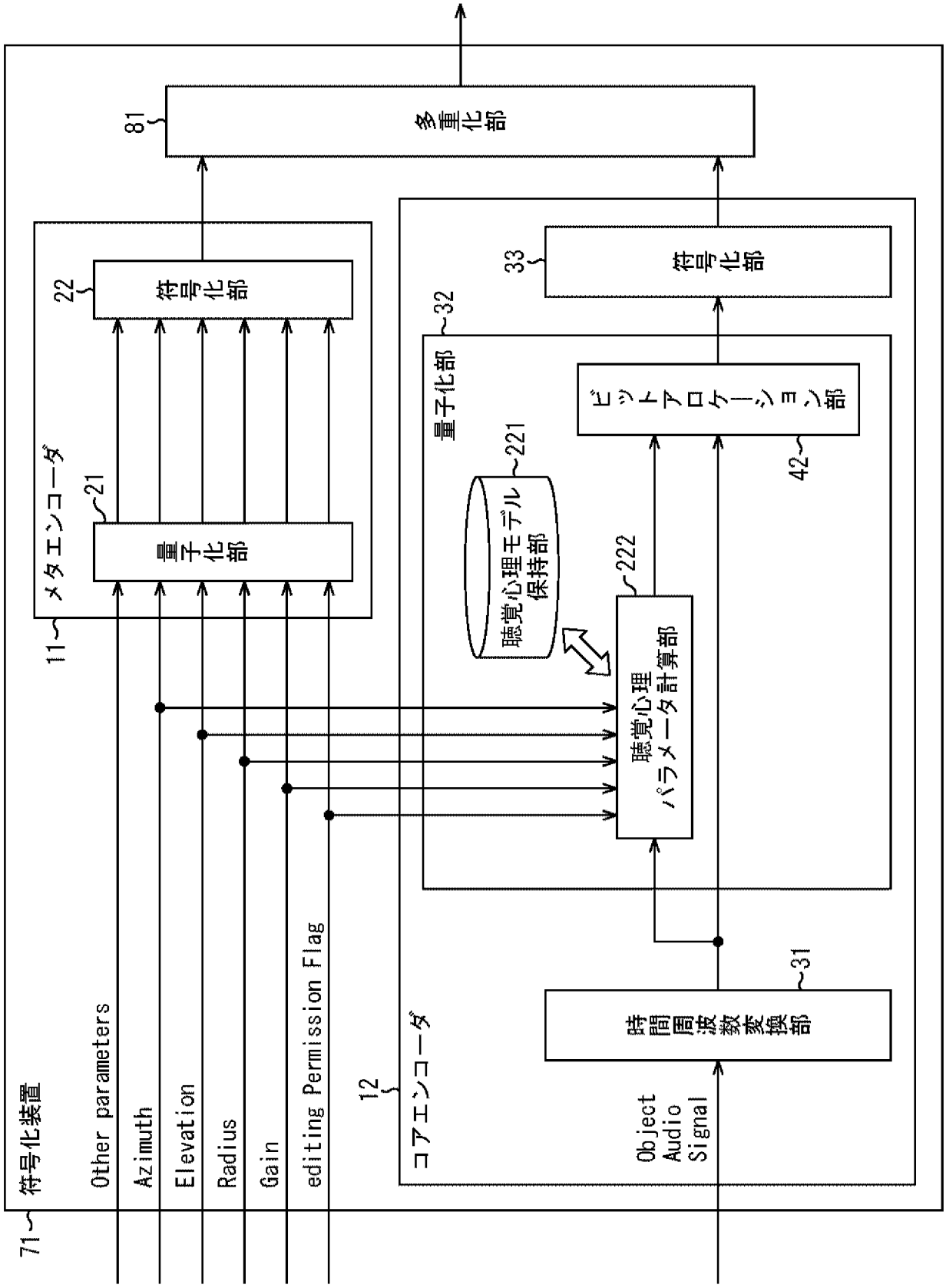
[図22]
FIG. 22


符号化処理のスタート
パラメータを量子化する
S171
差分符号化を行う
S172
オーディオ信号に対する
MDCTを行う
S173
聴覚心理パラメータを計算する
S174
ビットアロケーションを行い、
MDCT係数を量子化する
S175
算術符号化を行う
S176
多重化する
S177
エンド

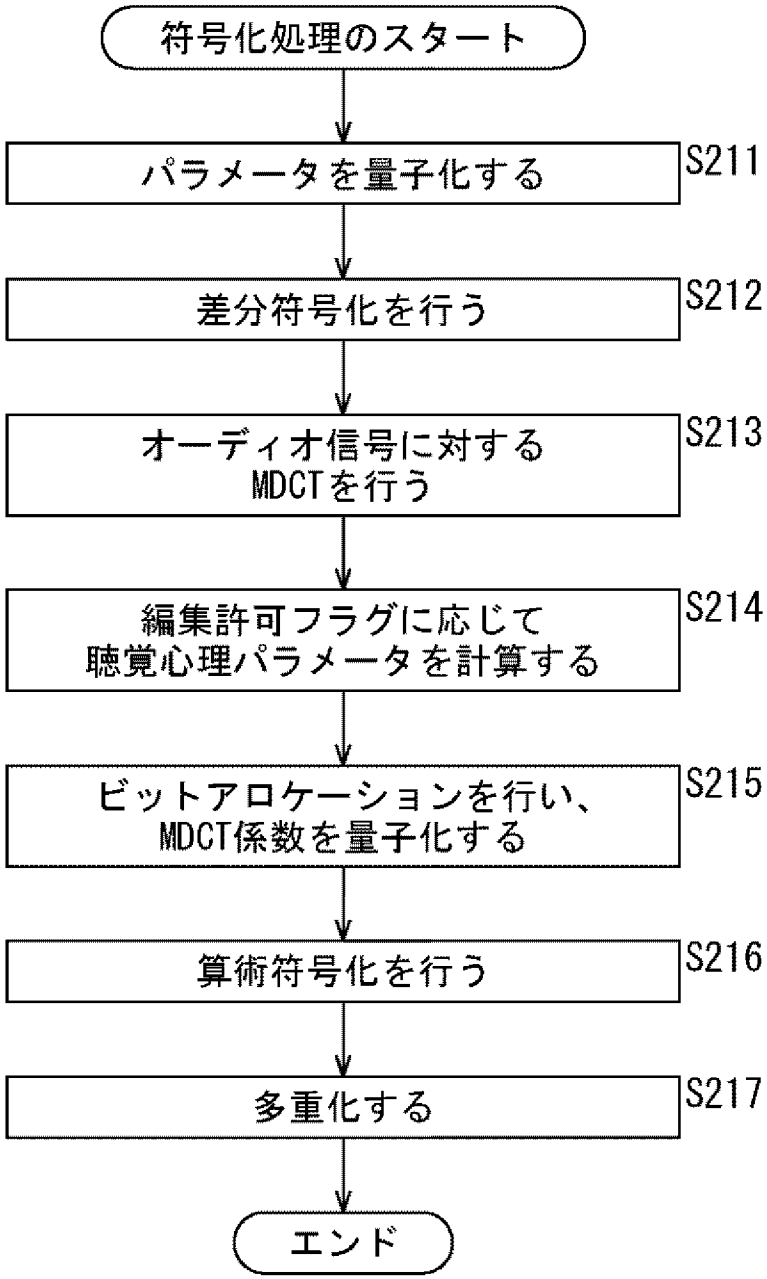
[図23]
FIG. 23

Syntax	No. of bits	Mnemonic
ObjectMetadataConfig()		
{		
lowDelayMetadataCoding;	1	bslbf
hasCoreLength;	1	bslbf
if(!hasCoreLength){		
frameLength;	6	uimsbf
OAMFrameLength = (frameLength+1)<<6;		
} else {		
OAMFrameLength = outputFrameLength;		
}		
hasScreenRelativeObjects;	1	bslbf
if(hasScreenRelativeObjects){		
for(o=0; o<num_objects; o++){		
isScreenRelativeObject[o];	1	bslbf
}		
}		
hasDynamicObjectPriority;	1	bslbf
hasUniformSpread;	1	bslbf
for(o=0; o<num_objects; o++){		
editingPermissionFlag[o]; ←Q11	1	bslbf
}		
}		

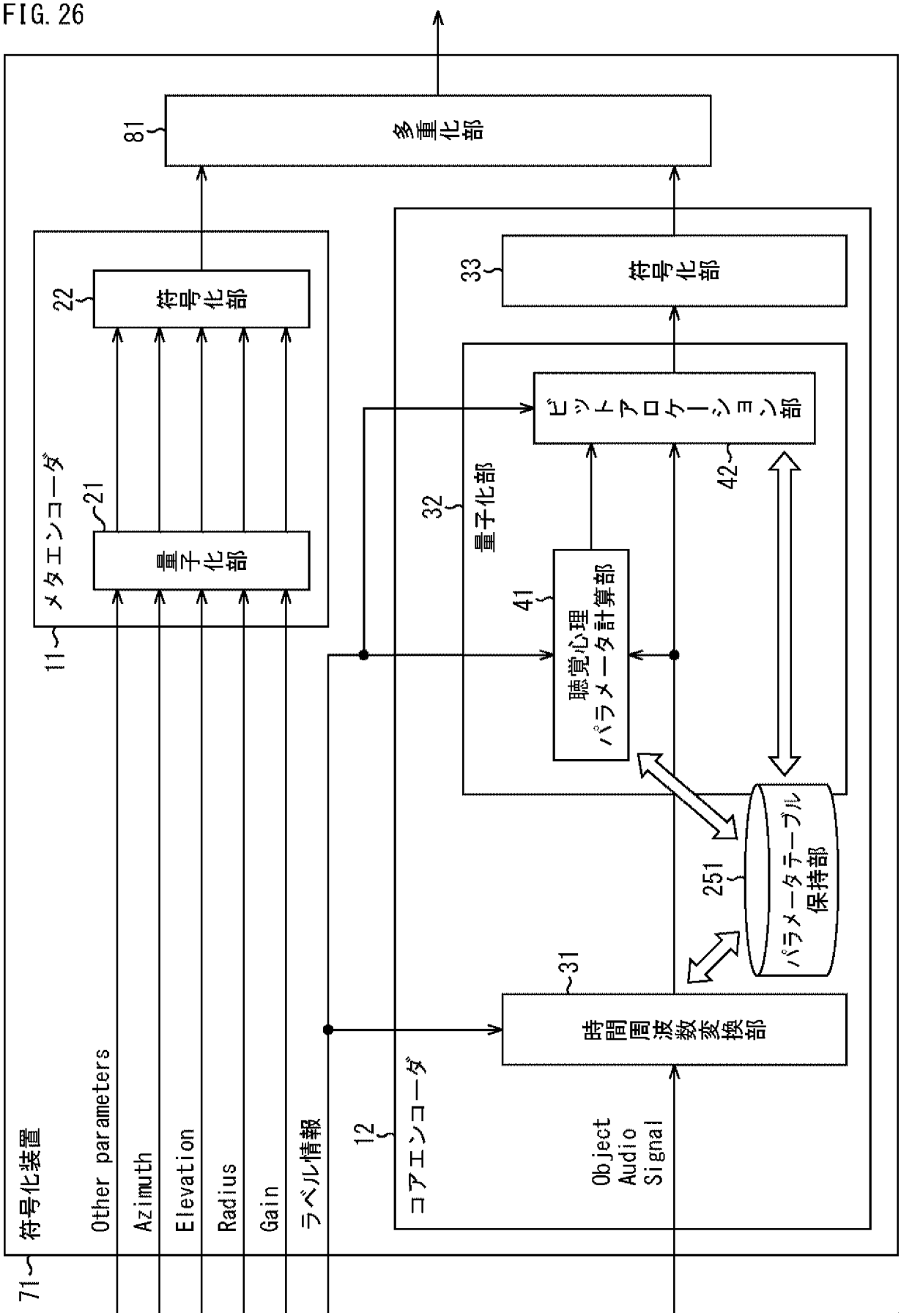
[図24]
FIG. 24


71
符号化装置
Other parameters
Azimuth
Elevation
Radius
Gain
editing Permission Flag
11
メタエンコーダ
21
量子化部
22
符号化部
12
コアエンコーダ
Object Audio Signal
31
時間周波数変換部
32
量子化部
聴覚心理モデル保持部
221
聴覚心理パラメータ計算部
222
ビットアロケーション部
42
33
符号化部
81
多重化部

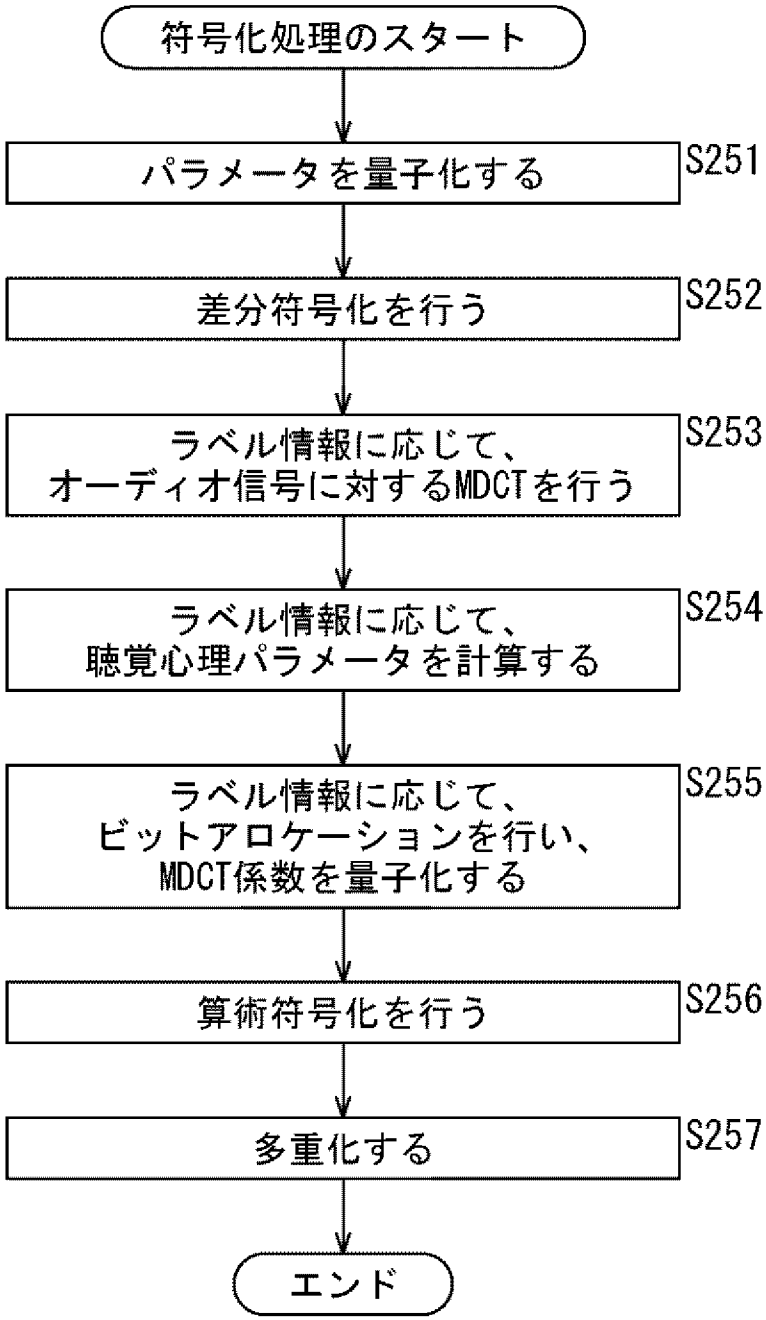
[図25]
FIG. 25


符号化処理のスタート
パラメータを量子化する
S211
差分符号化を行う
S212
オーディオ信号に対する
MDCTを行う
S213
編集許可フラグに応じて
聴覚心理パラメータを計算する
S214
ビットアロケーションを行い、
MDCT係数を量子化する
S215
算術符号化を行う
S216
多重化する
S217
エンド

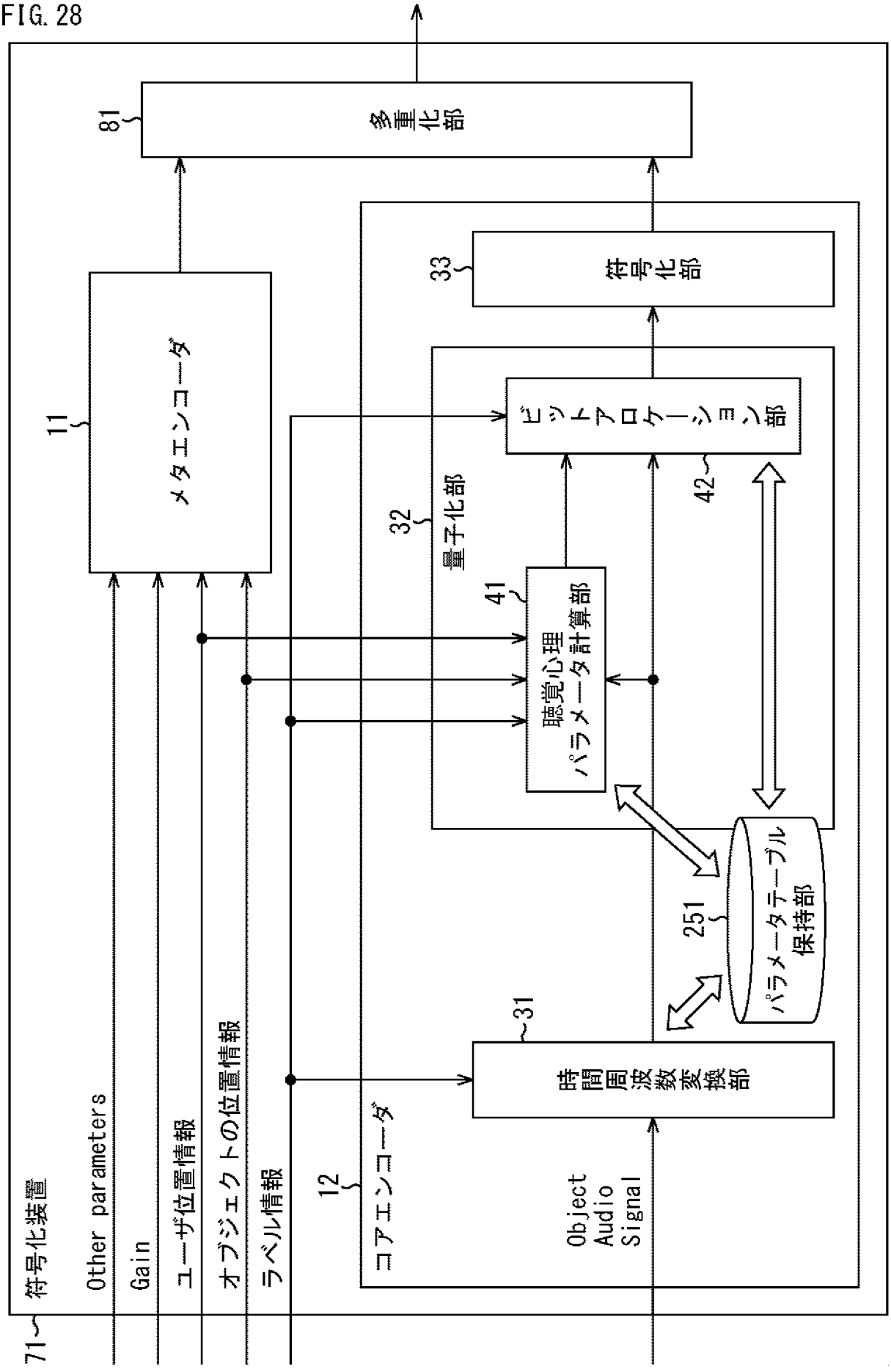
[図26]
FIG. 26


71
符号化装置
Other parameters
Azimuth
Elevation
Radius
Gain
ラベル情報
Object Audio Signal
12
コアエンコーダ
11
メタエンコーダ
21
量子化部
22
符号化部
31
時間周波数変換部
32
量子化部
33
符号化部
41
聴覚心理パラメータ計算部
42
ビットアロケーション部
251
パラメータテーブル保持部
81
多重化部

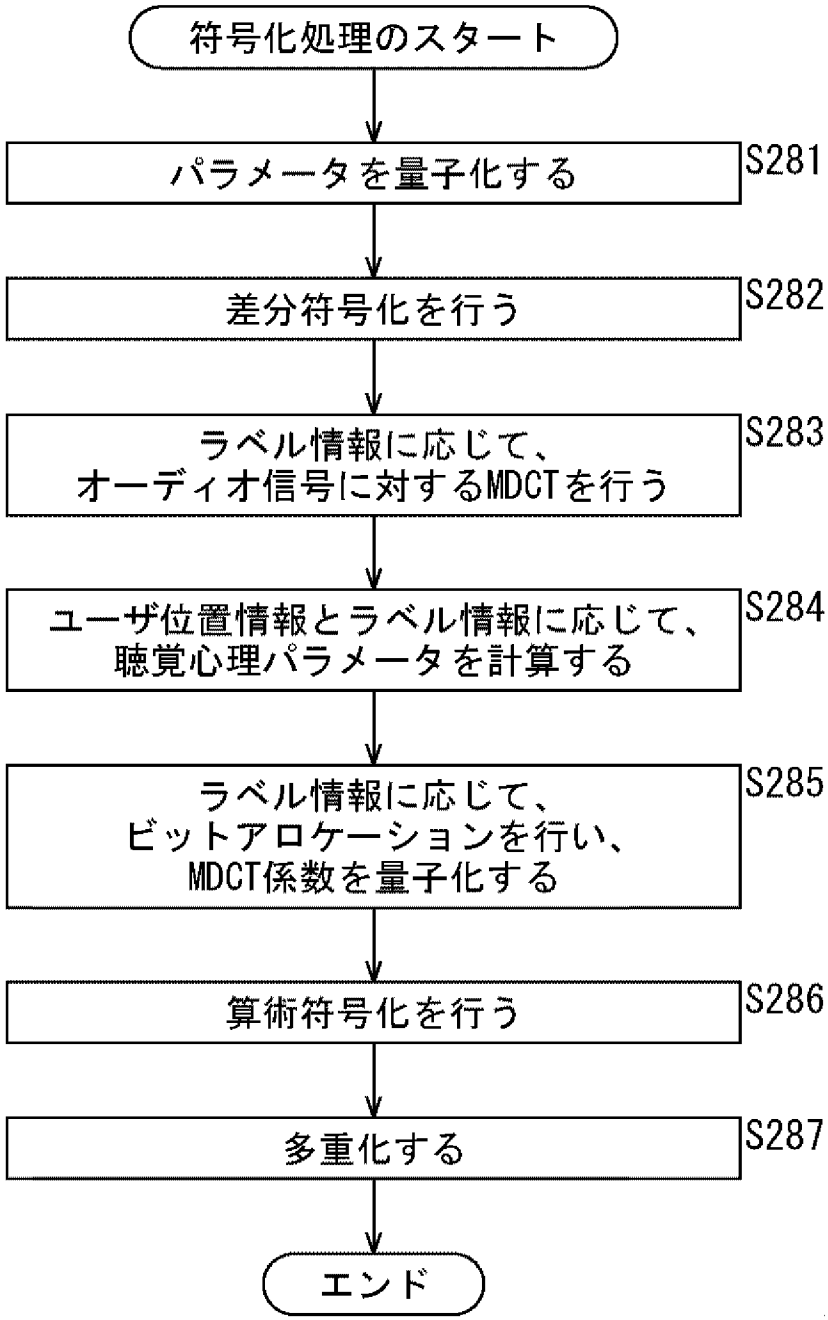
[図27]
FIG. 27


符号化処理のスタート
パラメータを量子化する
S251
差分符号化を行う
S252
ラベル情報に応じて、
オーディオ信号に対するMDCTを行う
S253
ラベル情報に応じて、
聴覚心理パラメータを計算する
S254
ラベル情報に応じて、
ビットアロケーションを行い、
MDCT係数を量子化する
S255
算術符号化を行う
S256
多重化する
S257
エンド

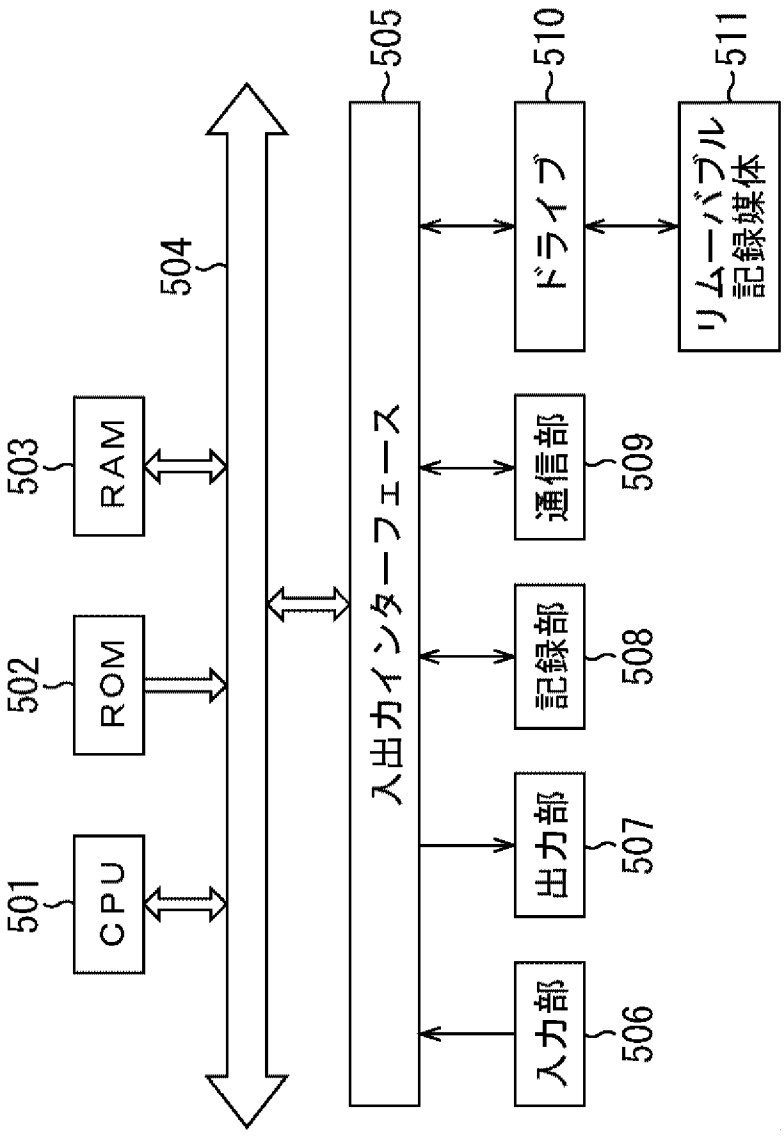
[図28]
FIG. 28


71
符号化装置
Other parameters
Gain
ユーザ位置情報
オブジェクトの位置情報
ラベル情報
12
コアエンコーダ
Object Audio Signal
31
時間周波数変換部
251
パラメータテーブル保持部
41
聴覚心理パラメータ計算部
32
量子化部
42
ビットアロケーション部
33
符号化部
11
メタエンコーダ
81
多重化部

[図29]
FIG. 29


符号化処理のスタート
パラメータを量子化する
S281
差分符号化を行う
S282
ラベル情報に応じて、
オーディオ信号に対するMDCTを行う
S283
ユーザ位置情報とラベル情報に応じて、
聴覚心理パラメータを計算する
S284
ラベル情報に応じて、
ビットアロケーションを行い、
MDCT係数を量子化する
S285
算術符号化を行う
S286
多重化する
S287
エンド

[図30]
FIG. 30


501
CPU
502
ROM
503
RAM
504
505
入出力インターフェース
506
入力部
507
出力部
508
記録部
509
通信部
510
ドライブ
511
リムーバブル記録媒体

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2021/024098

A. CLASSIFICATION OF SUBJECT MATTER
G10L 19/00(2013.01)i; G10L 19/008(2013.01)i; G10L 19/035(2013.01)i
FI: G10L19/008 200; G10L19/00 330B; G10L19/035 Z

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
G10L19/00-25/93

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Published examined utility model applications of Japan	1922-1996
Published unexamined utility model applications of Japan	1971-2021
Registered utility model specifications of Japan	1996-2021
Published registered utility model applications of Japan	1994-2021

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	JP 2018-522286 A (APPLE INC.) 09 August 2018 (2018-08-09) paragraphs [0013]-[0035]	1-4, 10-12, 14-15, 17-18
Y	JP 2001-154695 A (VICTOR COMPANY OF JAPAN, LTD.) 08 June 2001 (2001-06-08) paragraphs [0009]-[0018]	1-4, 10-12, 14-15, 17-18
A	JP 2016-530803 A (DOLBY LABORATORIES LICENSING CORPORATION) 29 September 2016 (2016-09-29) paragraph [0102]	25-33
A	WO 2016/203994 A1 (SONY CORP.) 22 December 2016 (2016-12-22) entire text	1-33
A	WO 2016/208406 A1 (SONY CORP.) 29 December 2016 (2016-12-29) entire text	1-33

☒ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
13 September 2021 (13.09.2021)	21 September 2021 (21.09.2021)

Name and mailing address of the ISA/	Authorized officer
Japan Patent Office 3-4-3, Kasumigaseki, Chiyoda-ku, Tokyo 100-8915, Japan	Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2021/024098

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 2006-139827 A (VICTOR COMPANY OF JAPAN, LTD.) 01 June 2006 (2006-06-01) entire text	1-33

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/JP2021/024098

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
JP 2018-522286 A	09 Aug. 2018	US 2017/0032793 A1 paragraphs [0020]-[0042] EP 3329487 A1 WO 2017/023601 A1 KR 10-2018-0019715 A CN 107851440 A	
JP 2001-154695 A	08 Jun. 2001	(Family: none)	
JP 2016-530803 A	29 Sep. 2016	US 2016/0192105 A1 paragraph [0117] WO 2015/017235 A1 EP 3028273 A1 KR 10-2016-0021892 A CN 105431900 A	
WO 2016/203994 A1	22 Dec. 2016	US 2018/0315436 A1 entire text EP 3316599 A1 KR 10-2017-0141276 A CN 107637097 A	
WO 2016/208406 A1	29 Dec. 2016	US 2018/0160250 A1 entire text EP 3319342 A1 KR 10-2018-0008609 A CN 107710790 A	
JP 2006-139827 A	01 Jun. 2006	(Family: none)	

国際調査報告

国際出願番号
PCT/JP2021/024098

A. 発明の属する分野の分類（国際特許分類（ＩＰＣ））

G10L 19/00(2013.01)i; G10L 19/008(2013.01)i; G10L 19/035(2013.01)i
FI: G10L19/008 200; G10L19/00 330B; G10L19/035 Z

B. 調査を行った分野

調査を行った最小限資料（国際特許分類（ＩＰＣ））

G10L19/00-25/93

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報	１９２２－１９９６年
日本国公開実用新案公報	１９７１－２０２１年
日本国実用新案登録公報	１９９６－２０２１年
日本国登録実用新案公報	１９９４－２０２１年

国際調査で使用した電子データベース（データベースの名称、調査に使用した用語）

C. 関連すると認められる文献

引用文献のカテゴリー*	引用文献名　及び一部の箇所が関連するときは、その関連する箇所の表示	関連する請求項の番号
Y	JP 2018-522286 A (アップル インコーポレイテッド) 09.08.2018 (2018-08-09) 段落[0013]-[0035]	1-4,10-12, 14-15,17-18
Y	JP 2001-154695 A (日本ビクター株式会社) 08.06.2001 (2001-06-08) 段落[0009]-[0018]	1-4,10-12, 14-15,17-18
A	JP 2016-530803 A (ドルビー ラボラトリーズ ライセンシング コーポレイション) 29.09.2016 (2016-09-29) 段落[0102]	25-33
A	WO 2016/203994 A1 (ソニー株式会社) 22.12.2016 (2016-12-22) 全文	1-33
A	WO 2016/208406 A1 (ソニー株式会社) 29.12.2016 (2016-12-29) 全文	1-33
A	JP 2006-139827 A (日本ビクター株式会社) 01.06.2006 (2006-06-01) 全文	1-33

☐ Ｃ欄の続きにも文献が列挙されている。

☑ パテントファミリーに関する別紙を参照。

＊ 引用文献のカテゴリー
“A” 特に関連のある文献ではなく、一般的技術水準を示すもの
“E” 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの
“L” 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）
“O” 口頭による開示、使用、展示等に言及する文献
“P” 国際出願日前で、かつ優先権の主張の基礎となる出願の日の後に公表された文献
“T” 国際出願日又は優先日後に公表された文献であって出願と抵触するものではなく、発明の原理又は理論の理解のために引用するもの
“X” 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの
“Y” 特に関連のある文献であって、当該文献と他の１以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの
“&” 同一パテントファミリー文献

国際調査を完了した日	国際調査報告の発送日
13.09.2021	21.09.2021

名称及びあて先
日本国特許庁(ISA/JP)
〒100-8915
日本国
東京都千代田区霞が関三丁目４番３号

権限のある職員（特許庁審査官）
山下　剛史 5Z 8946

電話番号 03-3581-1101 内線 3591

国際調査報告
パテントファミリーに関する情報

国際出願番号
PCT/JP2021/024098

引用文献	公表日	パテントファミリー文献	公表日
JP 2018-522286 A	09.08.2018	US 2017/0032793 A1 段落[0020]-[0042] EP 3329487 A1 WO 2017/023601 A1 KR 10-2018-0019715 A CN 107851440 A	
JP 2001-154695 A	08.06.2001	（ファミリーなし）	
JP 2016-530803 A	29.09.2016	US 2016/0192105 A1 段落[0117] WO 2015/017235 A1 EP 3028273 A1 KR 10-2016-0021892 A CN 105431900 A	
WO 2016/203994 A1	22.12.2016	US 2018/0315436 A1 全文 EP 3316599 A1 KR 10-2017-0141276 A CN 107637097 A	
WO 2016/208406 A1	29.12.2016	US 2018/0160250 A1 全文 EP 3319342 A1 KR 10-2018-0008609 A CN 107710790 A	
JP 2006-139827 A	01.06.2006	（ファミリーなし）	