

(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
2014年12月4日(04.12.2014)



(10) 国際公開番号

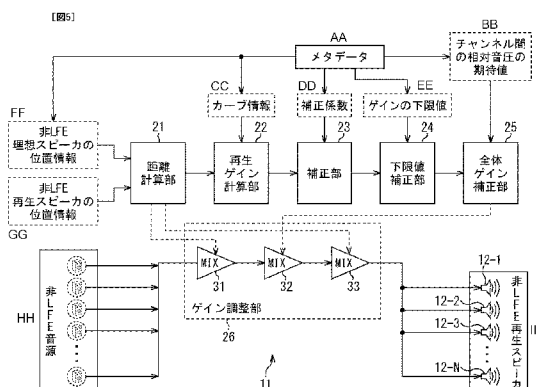
WO 2014/192603 A1

- (51) 国際特許分類:
H04S 5/02 (2006.01) G10L 21/034 (2013.01)
- (21) 国際出願番号: PCT/JP2014/063410
- (22) 国際出願日: 2014年5月21日(21.05.2014)
- (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 2013-115725 2013年5月31日(31.05.2013) JP
- (71) 出願人: ソニー株式会社 (SONY CORPORATION)
[JP/JP]; 〒1080075 東京都港区港南1丁目7番1号 Tokyo (JP).
- (72) 発明者: 史 潤宇 (SHI Runyu); 〒1080075 東京都港区港南1丁目7番1号 ソニー株式会社内 Tokyo (JP). 知念 徹 (CHINEN Toru); 〒1080075 東京都港区港南1丁目7番1号 ソニー株式会社内 Tokyo (JP). 山本 優樹 (YAMAMOTO Yuki); 〒1080075 東京都港区港南1丁目7番1号 ソニー株式会社内 Tokyo (JP). 畠中 光行 (HATANAKA Mitsuyuki); 〒1080075 東京都港区港南1丁目7番1号 ソニー株式会社内 Tokyo (JP).
- (74) 代理人: 西川 孝, 外 (NISHIKAWA Takashi et al.); 〒1600023 東京都新宿区西新宿7丁目5番25号 西新宿木村屋ビルディング9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, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, 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, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, 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, 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).

[続葉有]

(54) Title: AUDIO SIGNAL OUTPUT DEVICE AND METHOD, ENCODING DEVICE AND METHOD, DECODING DEVICE AND METHOD, AND PROGRAM

(54) 発明の名称: オーディオ信号出力装置および方法、符号化装置および方法、復号装置および方法、並びにプログラム



- 21 Distance calculation unit
22 Playback gain calculation unit
23 Correction unit
24 Lower limit correction unit
25 Overall gain correction unit
26 Gain adjustment unit
31, 32, 33 Mix
AA Metadata
BB Expected value for relative sound pressure between channels
CC Curve information
DD Correction coefficient
EE Lower limit for gain
FF Position information for non-LFE ideal speakers
GG Position information for non-LFE playback speakers
HH Non-LFE sound sources
II Non-LFE playback speakers

(57) Abstract: The present invention relates to an audio signal output device and method capable of more realistic audio reproduction, and also to an encoding device and method, a decoding device and method, and a program. Given input of an audio signal that is generated to be outputted as sound by a virtual speaker, i.e. a hypothetical speaker arranged in the ideal position, the distance between the position of the virtual speaker and the position of the real reproduction speaker is calculated. Then, the gain of the audio signal is adjusted on the basis of the gain corresponding to the calculated distance, and the gain-adjusted audio signal is reproduced by the reproduction speaker. By this means, even if there is positional deviation between the virtual speaker and the reproduction speaker, more realistic audio reproduction can be achieved. The present invention can be applied to a reproduction device.

(57) 要約: 本技術は、より臨場感のあるオーディオ再生を行なうことができるようにするオーディオ信号出力装置および方法、符号化装置および方法、復号装置および方法、並びにプログラムに関する。理想的な位置に配置された仮想のスピーカである理想スピーカから音声出力されるものとして生成されたオーディオ信号が入力された場合、理想スピーカの位置と、実際の再生スピーカの位置との距離が求められる。そして、求められた距離に応じたゲインによりオーディオ信号のゲイン調整が行なわれ、ゲイン調整後のオーディオ信号が再生スピーカで再生される。これにより、理想スピーカと再生スピーカとの配置位置にずれがある場合であっても、より臨場感のあるオーディオ再生を実現することができる。本技術は、再生装置に適用するこ

とができる。



添付公開書類:

— 国際調査報告 (条約第 21 条(3))

明 細 書

発明の名称：

オーディオ信号出力装置および方法、符号化装置および方法、復号装置および方法、並びにプログラム

技術分野

[0001] 本技術はオーディオ信号出力装置および方法、符号化装置および方法、復号装置および方法、並びにプログラムに関し、特に、より臨場感のあるオーディオ再生を行なうことができるようにしたオーディオ信号出力装置および方法、符号化装置および方法、復号装置および方法、並びにプログラムに関する。

背景技術

[0002] マルチチャンネルのオーディオ再生において、再生側のスピーカ配置は、音源の位置と完全に一致することが望ましいが、現実では再生側のスピーカ位置と音源の位置とが異なることが多い。

[0003] 再生側のスピーカの配置位置と音源の位置が異なると、スピーカの位置にない音源が生じるので、このような音源の音声をどのように再生するかは重要な事項である。

[0004] また、任意の位置にある音源の音声を、任意の位置のスピーカから再生する方法としてVBAP (Vector Base Amplitude Panning) と呼ばれる技術が提案されている (例えば、非特許文献 1 参照)。

[0005] VBAPでは、目標となる音像の定位位置が、その定位位置の周囲にある 2 つまたは 3 つのスピーカの方向を向くベクトルの線形和で表現される。そして、その線形和において各ベクトルに乗算されている係数が、各スピーカから出力されるオーディオ信号のゲインとして用いられてゲイン調整が行なわれ、目標となる位置に音像が定位するようになされる。

先行技術文献

非特許文献

- [0006] 非特許文献1: Ville Pulkki, “Virtual Sound Source Positioning Using Vector Base Amplitude Panning”, Journal of AES, vol.45, no.6, pp.456-466, 1997

発明の概要

発明が解決しようとする課題

- [0007] ところで、従来では音源のチャンネル数およびスピーカ配置と、再生側のスピーカのチャンネル数およびスピーカ配置が、それぞれ予め決められたチャンネル数とスピーカ配置となる場合、例えばいくつかの国際標準会議で推奨された7.1チャンネル配置と5.1チャンネル配置、5.1チャンネル配置と2.1チャンネル配置、または22.2チャンネル配置と5.1チャンネル配置のような場合については、音声の再生方法が提案されている。このような場合、ダウンミックス処理により各スピーカから適切なゲインで音声が出力され、臨場感のあるオーディオ再生を実現することができる。
- [0008] しかしながら、上記以外の場合や、音源位置またはスピーカ配置位置が、予め定められた位置からずれた位置にある場合には、提案されている再生方法では、音声を再生することができなかったり、再生はできるものの音質と音像定義が大きく劣化してしまったりすることになる。
- [0009] また、上述したVBAPではチャンネルベースの音源を再生する場合、チャンネルベースの音源の音像と音源が再生される理想のスピーカの位置が異なることがほとんどであるので、音像定義が大きく劣化することになる。
- [0010] 以上のように上述した技術では、臨場感のあるオーディオ再生を実現することが困難であった。
- [0011] 本技術は、このような状況に鑑みてなされたものであり、より臨場感のあるオーディオ再生を行なう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

- [0012] 本技術の第1の側面のオーディオ信号出力装置は、オーディオ信号を再生する理想的なスピーカの位置と、前記オーディオ信号を再生する実際のスピーカの位置との距離を計算する距離計算部と、前記距離に基づいて前記オー

ディオ信号の再生ゲインを計算するゲイン計算部と、前記再生ゲインに基づいて前記オーディオ信号のゲイン調整を行なうゲイン調整部とを備える。

[0013] 前記ゲイン計算部には、各前記距離に対する前記再生ゲインを得るためのカーブ情報に基づいて、前記再生ゲインを計算させることができる。

[0014] 前記カーブ情報を、折れ線カーブまたは関数カーブを示す情報とすることができる。

[0015] 前記ゲイン調整部には、前記理想的なスピーカが所定の基準点を中心とする単位円上にない場合、前記基準点から前記理想的なスピーカまでの距離と前記単位円の半径とに基づいて定められたゲインにより前記オーディオ信号のゲイン調整をさらに行なわせることができる。

[0016] 前記ゲイン調整部には、前記基準点から前記理想的なスピーカまでの距離と前記単位円の半径とに基づいて定められた遅延時間に基づいて前記オーディオ信号を遅延させるようにすることができる。

[0017] 前記ゲイン調整部には、前記実際のスピーカが所定の基準点を中心とする単位円上にない場合、前記基準点から前記実際のスピーカまでの距離と前記単位円の半径とに基づいて定められたゲインにより前記オーディオ信号のゲイン調整をさらに行なわせることができる。

[0018] 前記ゲイン調整部には、前記基準点から前記実際のスピーカまでの距離と前記単位円の半径とに基づいて定められた遅延時間に基づいて前記オーディオ信号を遅延させるようにすることができる。

[0019] オーディオ信号出力装置には、理想的なセンタースピーカの位置と、前記実際のスピーカの位置との距離に基づいて、前記再生ゲインを補正するゲイン補正部をさらに設けることができる。

[0020] オーディオ信号出力装置には、前記再生ゲインが予め定められた下限値よりも小さい場合、前記再生ゲインを補正する下限値補正部をさらに設けることができる。

[0021] オーディオ信号出力装置には、入力された前記オーディオ信号に基づく入力音声の音圧の期待値、および前記再生ゲインに基づいて、前記再生ゲイン

によりゲイン調整された前記オーディオ信号に基づく出力音声の全体のパワーと、前記入力音声の全体のパワーとの比を算出し、前記比に基づいて前記再生ゲインを補正する全体ゲイン補正部をさらに設けることができる。

[0022] 本技術の第1の側面のオーディオ信号出力方法またはプログラムは、オーディオ信号を再生する理想的なスピーカの位置と、前記オーディオ信号を再生する実際のスピーカの位置との距離を計算し、前記距離に基づいて前記オーディオ信号の再生ゲインを計算し、前記再生ゲインに基づいて前記オーディオ信号のゲイン調整を行なうステップを含む。

[0023] 本技術の第1の側面においては、オーディオ信号を再生する理想的なスピーカの位置と、前記オーディオ信号を再生する実際のスピーカの位置との距離が計算され、前記距離に基づいて前記オーディオ信号の再生ゲインが計算され、前記再生ゲインに基づいて前記オーディオ信号のゲイン調整が行なわれる。

[0024] 本技術の第2の側面の符号化装置は、オーディオ信号を再生する理想的なスピーカの位置と、前記オーディオ信号を再生する実際のスピーカの位置との距離に応じて前記オーディオ信号のゲインを補正するための補正情報を生成する補正情報生成部と、前記オーディオ信号を符号化する符号化部と、前記補正情報および符号化された前記オーディオ信号を含むビットストリームを出力する出力部とを備える。

[0025] 本技術の第2の側面の符号化方法は、オーディオ信号を再生する理想的なスピーカの位置と、前記オーディオ信号を再生する実際のスピーカの位置との距離に応じて前記オーディオ信号のゲインを補正するための補正情報を生成し、前記オーディオ信号を符号化し、前記補正情報および符号化された前記オーディオ信号を含むビットストリームを出力するステップを含む。

[0026] 本技術の第2の側面においては、オーディオ信号を再生する理想的なスピーカの位置と、前記オーディオ信号を再生する実際のスピーカの位置との距離に応じて前記オーディオ信号のゲインを補正するための補正情報が生成され、前記オーディオ信号が符号化され、前記補正情報および符号化された前

記オーディオ信号を含むビットストリームが出力される。

[0027] 本技術の第3の側面の復号装置は、オーディオ信号を再生する理想的なスピーカの位置と、前記オーディオ信号を再生する実際のスピーカの位置との距離に応じて前記オーディオ信号のゲインを補正するための補正情報、および符号化された前記オーディオ信号をビットストリームから抽出する抽出部と、前記符号化された前記オーディオ信号を復号する復号部と、復号された前記オーディオ信号および前記補正情報を出力する出力部とを備える。

[0028] 前記補正情報を、前記理想的なスピーカの位置情報とすることができる。

[0029] 前記補正情報を、各前記距離に対するゲインを得るためのカーブ情報とすることができる。

[0030] 前記カーブ情報を、折れ線カーブまたは関数カーブを示す情報とすることができる。

[0031] 本技術の第3の側面の復号方法は、オーディオ信号を再生する理想的なスピーカの位置と、前記オーディオ信号を再生する実際のスピーカの位置との距離に応じて前記オーディオ信号のゲインを補正するための補正情報、および符号化された前記オーディオ信号をビットストリームから抽出し、前記符号化された前記オーディオ信号を復号し、復号された前記オーディオ信号および前記補正情報を出力するステップを含む。

[0032] 本技術の第3の側面においては、オーディオ信号を再生する理想的なスピーカの位置と、前記オーディオ信号を再生する実際のスピーカの位置との距離に応じて前記オーディオ信号のゲインを補正するための補正情報、および符号化された前記オーディオ信号がビットストリームから抽出され、前記符号化された前記オーディオ信号が復号され、復号された前記オーディオ信号および前記補正情報が出力される。

発明の効果

[0033] 本技術の第1の側面乃至第3の側面によれば、より臨場感のあるオーディオ再生を行なうことができる。

図面の簡単な説明

[0034] [図1]本技術の概要について説明する図である。

[図2]折れ線カーブについて説明する図である。

[図3]関数カーブについて説明する図である。

[図4]再生ゲインについて説明する図である。

[図5]再生装置の構成例を示す図である。

[図6]ダウンミックス処理を説明するフローチャートである。

[図7]オーディオシステムの構成例を示す図である。

[図8]メタデータについて説明する図である。

[図9]符号化処理を説明するフローチャートである。

[図10]復号処理を説明するフローチャートである。

[図11]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

[0035] 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。

[0036] 〈第1の実施の形態〉

〈本技術の概要について〉

本技術は、任意のチャンネルの音源を任意の数のスピーカで再生する再生方法と、その再生方法の実現に必要な情報（メタデータ）の符号化および復号技術に関するものである。

[0037] まず、本技術の概要について説明する。

[0038] 例えば、複数の各チャンネルのオーディオ信号と、それらのオーディオ信号のメタデータとが再生装置に供給され、再生装置において、メタデータとオーディオ信号に基づいて、音声の再生を制御するものとする。

[0039] ここで、各チャンネルのオーディオ信号は、メタデータにより示される理想的な位置に配置されたスピーカにより再生されるものとして生成された信号であるものとする。以下では、メタデータにより示される位置にある、各チャンネルのオーディオ信号を再生する仮想のスピーカを、理想スピーカと称することとする。また、再生装置から出力されたオーディオ信号に基づいて音声を出力する実際のスピーカを再生スピーカと称することとする。

[0040] 本技術では、全チャンネルのオーディオ信号が、LFE (Low Frequency Effect) 用のオーディオ信号と、LFE用でないオーディオ信号とに分類される。すなわち、全理想スピーカがLFE用のスピーカと、LFE用でないスピーカとに分類される。同様に、再生スピーカも、LFE用のスピーカと、LFE用でないスピーカとに分類される。

[0041] まず、LFE用ではないチャンネルのオーディオ信号の再生について説明する。

[0042] LFE用ではないチャンネルのオーディオ信号の再生では、例えば図1に示すように、理想スピーカと再生スピーカとの距離に基づいて、オーディオ信号のゲイン調整が行なわれる。

[0043] 図1では、視聴者であるユーザU11の位置を中心とする半径 r_u の球PH11の表面上に理想スピーカVSP1と、再生スピーカRSP11-1乃至再生スピーカRSP11-3とが配置されている。理想スピーカVSP1および再生スピーカRSP11-1乃至再生スピーカRSP11-3はLFE用ではないスピーカである。

[0044] なお、以下、再生スピーカRSP11-1乃至再生スピーカRSP11-3を特に区別する必要のない場合、単に再生スピーカRSP11とも称する。また、この例では、1つの理想スピーカおよび3つの再生スピーカだけが図示されているが、実際には他の理想スピーカや再生スピーカも存在する。

[0045] 例えば、理想スピーカVSP1に対応するチャンネルのオーディオ信号に基づく音声は、理想スピーカVSP1の位置に音像を定位させることが理想的である。

[0046] そこで、本技術では、理想スピーカVSP1と再生スピーカRSP11との距離に応じて各再生スピーカRSP11の再生ゲインを定め、それらの再生ゲインでオーディオ信号に基づく音声を各再生スピーカRSP11から出力させることで、理想スピーカVSP1の位置に音像を定位させる。

[0047] 具体的には、理想スピーカVSP1と再生スピーカRSP11の距離は、ユーザU11を始点とし、理想スピーカVSP1方向を向くベクトルと、ユ

ーザU 1 1を始点とし、再生スピーカR S P 1 1方向を向くベクトルとのなす角度とされる。

[0048] 換言すれば、球P H 1 1の表面上における理想スピーカV S P 1と再生スピーカR S P 1 1との距離、つまり2つのスピーカを結ぶ弧の長さが、理想スピーカV S P 1と再生スピーカR S P 1 1の距離とされる。

[0049] 図1の例では、矢印A 1 1と矢印A 1 2とがなす角度が、理想スピーカV S P 1と再生スピーカR S P 1 1-1との距離DistM1とされている。同様に、矢印A 1 1と矢印A 1 3とがなす角度が、理想スピーカV S P 1と再生スピーカR S P 1 1-2との距離DistM2とされ、矢印A 1 1と矢印A 1 4とがなす角度が、理想スピーカV S P 1と再生スピーカR S P 1 1-3との距離DistM3とされている。

[0050] そして、例えば理想スピーカV S P 1のチャンネルのオーディオ信号が、距離DistM1に基づいてゲイン調整されて再生スピーカR S P 1 1-1で再生される。また、理想スピーカV S P 1のチャンネルのオーディオ信号は、それぞれ距離DistM2および距離DistM3に基づいてゲイン調整され、再生スピーカR S P 1 1-2および再生スピーカR S P 1 1-3で再生される。

[0051] これにより、理想スピーカV S P 1と再生スピーカR S P 1 1との位置にずれがある場合であっても、それによって生じた音像のずれを軽減することができ、より臨場感のあるオーディオ再生を実現することができる。

[0052] 次に、LFE用でないチャンネルのオーディオ信号の再生について、さらに詳細に説明していく。

[0053] ここでは、例えば具体的に、LFE用でないM個の理想スピーカ、つまりMチャンネルのオーディオ信号をダウンミックスしてNチャンネルのオーディオ信号とし、それらのNチャンネルのオーディオ信号をLFE用でないN個の再生スピーカにより再生する例について説明する。

[0054] ダウンミックス処理では、主に以下に示す6つの処理S T E 1乃至処理S T E 6が順番に行なわれる。

[0055] (処理S T E 1) : 理想スピーカと再生スピーカの距離を求める

(処理 S T E 2) : 求めた距離と予め定めた減衰カーブに基づいて、理想スピーカごとに各再生スピーカの再生ゲインを求める

(処理 S T E 3) : 再生スピーカの配置位置に応じて再生ゲインを補正する

(処理 S T E 4) : 下限値に基づいて再生ゲインを補正する

(処理 S T E 5) : 出力音声全体のエネルギーが、入力音声全体のエネルギーと近くなるように再生ゲインを補正する

(処理 S T E 6) : 再生ゲインをオーディオ信号に適用し、ゲイン調整を行なう

[0056] 続いて、これらの処理 S T E 1 乃至処理 S T E 6 について、さらに説明する。

[0057] 〈処理 S T E 1 について〉

まず、処理 S T E 1 では、スピーカ間の距離が求められるが、各スピーカの位置は、水平方向角度 θ ($-180^\circ \leq \theta \leq +180^\circ$)、垂直方向角度 γ ($-90^\circ \leq \gamma \leq +90^\circ$)、およびユーザからスピーカまでの距離 r ($0 \leq r \leq +\infty$) により表現されるものとする。

[0058] 例えば図 1 において、ユーザ U 1 1 の位置を原点とし、 x 軸、 y 軸、および z 軸からなる 3 次元座標系を考えるとする。

[0059] ここで、図中、奥行き方向の直線と、図中、横方向の直線とを含む平面を $x y$ 平面とすると、 $x y$ 平面において基準となる方向の直線、例えば y 軸と、ユーザ U 1 1 を始点とするスピーカ方向のベクトルとが $x y$ 平面上においてなす角度が水平方向角度 θ とされる。つまり、水平方向角度 θ は、図 1 中、水平方向の角度である。

[0060] また、ユーザ U 1 1 を始点とするスピーカ方向のベクトルと、 $x y$ 平面とがなす角度が垂直方向角度 γ とされ、ユーザ U 1 1 とスピーカとを結ぶ直線の長さが距離 r とされる。

[0061] 各理想スピーカの位置を示す水平方向角度 θ 、垂直方向角度 γ 、および距離 r は、オーディオ信号のメタデータとして再生装置に供給される。また、

再生装置には、各再生スピーカの位置を示す水平方向角度 θ 、垂直方向角度 γ 、および距離 r も供給される。

[0062] なお、以下では、特にM個の理想スピーカのうちの m 番目の理想スピーカの水平方向角度 θ 、垂直方向角度 γ 、および距離 r を、それぞれ θ_{im} 、 γ_{im} 、および r_{im} で表すこととする。同様に、以下、N個の再生スピーカのうちの n 番目の再生スピーカの水平方向角度 θ 、垂直方向角度 γ 、および距離 r を、それぞれ θ_{on} 、 γ_{on} 、および r_{on} で表すこととする。

[0063] 再生装置では、M個の理想スピーカごとに、それらの理想スピーカと、N個の各再生スピーカとの距離が求められる。

[0064] 例えば、 m 番目の理想スピーカと、 n 番目の再生スピーカとの距離 $\text{Dist}(m, n)$ は、次式(1)により求められる。

[0065] [数1]

$$\text{Dist}(m, n) = \arccos[\cos \theta_{im} \times \cos \theta_{on} \times \cos(\gamma_{im} - \gamma_{on}) + \sin \theta_{im} \times \sin \theta_{on}]$$
$$(0^\circ \leq \text{Dist}(m, n) \leq 180^\circ) \quad \dots (1)$$

[0066] 再生装置では、M個の理想スピーカとN個の再生スピーカの組み合わせごとに式(1)の計算が行なわれ、合計で $M \times N$ 個の距離 $\text{Dist}(m, n)$ が計算される。

[0067] ところで、各理想スピーカや再生スピーカが半径 r_u の単位円上、つまり図1に示した球PH11上に配置されていれば、各スピーカから出力された音声は、同時にユーザU11に到達する。しかし、一部のスピーカが球PH11上にはない場合には、そのスピーカからの音声は、他のスピーカからの音声と比べてユーザU11に早くまたは遅く到達することになるだけでなく、ユーザに聞こえる音声の音圧も変化してしまう。

[0068] そこで、再生装置では、距離 $r_{im} \neq r_u$ である理想スピーカのオーディオ信号に対して、補正值 $\text{SoundPressureCorrection}_{im}$ により音圧の補正が行なわれ、遅延時間 Delay_{im} により遅延処理が施される。

[0069] これにより、理想スピーカが球PH11上に位置するものとして扱うことができるようになる。

[0070] 具体的には、距離 r_{im} と半径 r_u に基づいて次式（２）の計算が行なわれて、補正值 $\text{SoundPressureCorrection}_{im}$ が算出される。

[0071] [数2]

$$\text{SoundPressureCorrection}_{im} = -10 \times \log_{10} \left[\left(\frac{r_{im}}{r_u} \right)^2 \right] \text{ (dB)} \quad \dots (2)$$

[0072] 式（２）により求まる補正值 $\text{SoundPressureCorrection}_{im}$ は、理想スピーカ側、つまり再生装置に入力されたチャンネル m のオーディオ信号に対する補正に用いられる。以下、特に、再生装置に入力されたオーディオ信号を入力オーディオ信号とも称し、再生装置から出力されるオーディオ信号を出力オーディオ信号とも称することとする。

[0073] また、理想スピーカの入力オーディオ信号に対する遅延処理のための遅延時間 Delay_{im} は、距離 r_{im} と半径 r_u に基づいて次式（３）により算出される。なお、 $r_{im} > r_u$ の場合には遅延時間 Delay_{im} は負の値となり、遅延処理ではオーディオ信号が負の方向に遅延される、つまりオーディオ信号が時間的に前の方向にシフトされることになる。

[0074] [数3]

$$\text{Delay}_{im} = (r_u - r_{im}) \times \text{音速 (s)} \quad \dots (3)$$

[0075] これらの補正值 $\text{SoundPressureCorrection}_{im}$ と遅延時間 Delay_{im} は、距離 $r_{im} \neq r_u$ である理想スピーカについて算出される。同様に、距離 $r_{on} \neq r_u$ である再生スピーカについても補正值 $\text{SoundPressureCorrection}_{on}$ と遅延時間 Delay_{on} が算出される。

[0076] すなわち、次式（４）により補正值 $\text{SoundPressureCorrection}_{on}$ が算出され、式（５）により遅延時間 Delay_{on} が算出される。

[0077] [数4]

$$\text{SoundPressureCorrection}_{on} = -10 \times \log_{10} \left[\left(\frac{r_{on}}{r_u} \right)^2 \right] \text{ (dB)} \quad \dots (4)$$

[0078] [数5]

$$\text{Delay}_{on} = (r_u - r_{on}) \times \text{音速} (s) \quad \cdot \cdot \cdot (5)$$

[0079] このようにして求められた補正值SoundPressureCorrection_{on}と遅延時間Delay_{on}は、再生スピーカ側、つまり出力オーディオ信号に対する音圧の補正值と遅延時間である。再生装置では、距離 $r_{on} \neq r_u$ である再生スピーカに供給されるオーディオ信号に対して、補正值SoundPressureCorrection_{on}により音圧の補正が行なわれ、遅延時間Delay_{on}により遅延処理が施される。

[0080] 〈処理STE 2について〉

続いて、処理STE 2では、理想スピーカごとに各再生スピーカの再生ゲインが求められる。

[0081] まずは、M個の各理想スピーカについて、その理想スピーカとの距離Dist(m, n)が「0」となる再生スピーカが存在するか否かが特定され、各理想スピーカが、再生スピーカ位置にあるスピーカ、または再生スピーカ位置にないスピーカの何れかに分類される。

[0082] そして、再生スピーカ位置にあるスピーカとされたm番目の理想スピーカについては、そのm番目の理想スピーカに対応するチャンネルmのオーディオ信号についてのn番目の再生スピーカの再生ゲインMixGain(m, n)が次式(6)により算出される。

[0083] [数6]

$$\text{MixGain}(m, n) = \begin{cases} 0\text{dB}, & \text{Dist}(m, n) = 0 \\ -\infty\text{dB}, & \text{Dist}(m, n) > 0 \end{cases} \quad \cdot \cdot \cdot (6)$$

[0084] 式(6)では、距離Dist(m, n)が「0」である再生スピーカ、つまりm番目の理想スピーカと同じ位置にある再生スピーカの再生ゲインMixGain(m, n)は0dBとされる。また、距離Dist(m, n)が「0」でない再生スピーカ、つまりm番目の理想スピーカと異なる位置にある再生スピーカの再生ゲインMixGain(m, n)は $-\infty\text{dB}$ とされる。

[0085] これにより、m番目の理想スピーカに対応するチャンネルmのオーディオ信号は、その理想スピーカと同じ位置にある再生スピーカにおいて再生され

ることになる。つまり、他の再生スピーカからは、チャンネル m の音声成分は出力されない。

[0086] これに対して、再生スピーカ位置にないスピーカとされた m 番目の理想スピーカについては、折れ線カーブか、または関数カーブの何れかの減衰カーブが用いられて、その理想スピーカについての各再生スピーカの再生ゲイン $MixGain(m, n)$ が求められる。

[0087] 具体的には、再生装置に供給されるメタデータには、折れ線カーブ、または関数カーブの何れのカーブを用いて再生ゲインを求めるかを示すカーブ情報が含まれており、再生装置はメタデータに含まれるカーブ情報に示される種類のカーブを用いて再生ゲインを算出する。

[0088] また、メタデータにはカーブ情報に示される種類のカーブのうちの具体的にどのカーブを用いるかを示すカーブインデックスも含まれている。カーブインデックスは、再生装置には記録されていない新たなカーブを示す情報であることもある。

[0089] 再生装置は、カーブインデックスが予め定められたカーブを示す情報である場合には、予め記録されている、係数等のカーブを得るための情報を用いて再生ゲインの算出を行なう。一方、カーブインデックスが新たなカーブを示す情報である場合には、再生装置は、メタデータから、新たなカーブを得るための情報を読み出し、その情報から得られるカーブを用いて再生ゲインを算出する。

[0090] 例えば、再生ゲインの算出に用いられる折れ線カーブは、各距離 $Dist(m, n)$ に対する再生ゲインの値から構成される数列により表現される。

[0091] 具体的には、再生ゲインの値の数列として、 $[0, -1.5, -4.5, -6, -9, -10.5, -12, -13.5, -15, -15, -16.5, -16.5, -18, -18, -18, -19.5, -19.5, -21, -21, -21, -\infty, -\infty, -\infty, -\infty, -\infty]$ (dB) が、再生ゲインを得るための情報とされているとする。

[0092] そのような場合、数列の始点の値が距離 $Dist(m, n) = 0^\circ$ のときの再生ゲインとされ、数列の終点の値が距離 $Dist(m, n) = 180^\circ$ のときの再生ゲインとさ

れる。また、数列の k 番目の点の値が、次式 (7) に示す距離 $\text{Dist}(m, n)$ のときの再生ゲインとされる。

[0093] [数7]

$$\text{Dist}(m, n) = (k-1) \times \frac{180^\circ}{\text{数列の長さ}-1} \quad \dots (7)$$

[0094] また、数列の隣接する点の間は、距離 $\text{Dist}(m, n)$ によって再生ゲインが線形に変化するとされる。このような数列により得られる折れ線カーブは、再生ゲイン $\text{MixGain}(m, n)$ と、距離 $\text{Dist}(m, n)$ のマッピングを表すカーブである。

[0095] 例えば、上述した数列によって図2に示す折れ線カーブが得られる。

[0096] 図2では、縦軸は再生ゲインの値を示しており、横軸は理想スピーカと再生スピーカとの間の距離を示している。また、折れ線 $CV11$ が折れ線カーブを表しており、折れ線カーブ上の四角形は、再生ゲインの値の数列を構成する1つの数値を示している。

[0097] この例では、 n 番目の再生スピーカと m 番目の理想スピーカとの距離 $\text{Dist}(m, n)$ が DistM1 である場合には、その n 番目の再生スピーカの再生ゲイン $\text{MixGain}(m, n)$ は、折れ線カーブ上の DistM1 におけるゲインの値である -3.5dB とされる。

[0098] また、距離 $\text{Dist}(m, n)$ が DistM2 である再生スピーカの再生ゲイン $\text{MixGain}(m, n)$ は、折れ線カーブ上の DistM2 におけるゲインの値である -8dB とされ、距離 $\text{Dist}(m, n)$ が DistM3 である再生スピーカの再生ゲイン $\text{MixGain}(m, n)$ は、折れ線カーブ上の DistM3 におけるゲインの値である -16.5dB とされる。

[0099] 一方、再生ゲインの算出に用いられる関数カーブは、3つの係数 coef1 、係数 coef2 、および係数 coef3 と、予め定められた下限となるゲイン値 MinGain により表現される。

[0100] この場合、再生装置は係数 coef1 乃至係数 coef3 、ゲイン値 MinGain 、および距離 $\text{Dist}(m, n)$ により表現される、次式 (8) に示す関数 $f(\text{Dist}(m, n))$ を用いて、以下の式 (9) を計算し、 m 番目の理想スピーカについての各再生スピーカの再生ゲイン $\text{MixGain}(m, n)$ を算出する。

[0101] [数8]

$$f(\text{Dist}(m, n)) = \text{MinGain} \times \left(\text{Coef1} \times \left(\frac{\text{Dist}(m, n)}{180^\circ} \right)^3 + \text{Coef2} \times \left(\frac{\text{Dist}(m, n)}{180^\circ} \right)^2 + \text{Coef3} \times \left(\frac{\text{Dist}(m, n)}{180^\circ} \right) \right) \quad \dots (8)$$

[0102] [数9]

$$\text{MixGain}(m, n) = \begin{cases} 0\text{dB}, & f(\text{Dist}(m, n)) > 0\text{dB} \\ f(\text{Dist}(m, n)), & \text{otherwise} \\ -\infty\text{dB}, & \text{Dist}(m, n) > \text{Cut_thre} \end{cases} \quad \dots (9)$$

[0103] なお、式(9)において、Cut_threは、次式(10)を満たす最小値である。

[0104] [数10]

$$f(\text{Cut_thre}) = \text{MinGain} = -21\text{dB}, \quad f'(\text{Cut_thre}) < 0 \quad \dots (10)$$

[0105] このような関数 $f(\text{Dist}(m, n))$ 等により表される関数カーブは、例えば図3に示すカーブとなる。なお、図3において縦軸は再生ゲインの値を示しており、横軸は理想スピーカと再生スピーカとの間の距離を示している。また、曲線C V 2 1が関数カーブを表している。

[0106] 図3に示す関数カーブでは、関数 $f(\text{Dist}(m, n))$ により示される再生ゲインの値が始めて下限となるゲイン値MinGainより小さくなると、それ以降の各距離 $\text{Dist}(m, n)$ における再生ゲインの値は「 $-\infty$ 」とされている。なお、図中の点線は、各距離 $\text{Dist}(m, n)$ におけるもとの関数 $f(\text{Dist}(m, n))$ の値を示している。

[0107] この例では、 n 番目の再生スピーカと m 番目の理想スピーカとの距離 $\text{Dist}(m, n)$ が DistM1 である場合には、その n 番目の再生スピーカの再生ゲイン $\text{MixGain}(m, n)$ は、関数カーブ上の DistM1 におけるゲインの値である -6dB とされる。

[0108] また、距離 $\text{Dist}(m, n)$ が DistM2 である再生スピーカの再生ゲイン $\text{MixGain}(m,$

n)は、関数カーブ上のDistM2におけるゲインの値である-12dBとされ、距離Dist(m,n)がDistM3である再生スピーカの再生ゲインMixGain(m,n)は、関数カーブ上のDistM3におけるゲインの値である-18dBとされる。

[0109] なお、関数カーブにより再生ゲインMixGain(m,n)を求める場合、係数coef1乃至係数coef3の組み合わせ [coef1, coef2, coef3] は、例えば [8, -12, 6] や、 [1, -3, 3] 、 [2, -5.3, 4.2] などとされる。

[0110] 以上の処理によって、M個の理想スピーカごとに、N個の各再生スピーカの再生ゲインMixGain(m,n)が得られることになる。これらの再生スピーカの再生ゲインの値は、理想スピーカとの距離Dist(m,n)に近いほど大きい値となり、その音声の音量も大きくなる。なお、再生ゲインMixGain(m,n)は、 $M > N$ となる場合にはミックスゲインとなる。

[0111] 〈処理STE3について〉

さらに、処理STE3では、処理STE2で得られた $M \times N$ 個の再生ゲインMixGain(m,n)に対して、n番目の再生スピーカの配置位置に応じて補正が行なわれる。

[0112] 例えば、ユーザの前方に位置する音源からの音声、ユーザの後方から聞こえると違和感が生じてしまう。これに対して、ユーザの後方に位置する音源からの音声、ユーザの前方から聞こえても大きな違和感が生じない。

[0113] そこで、N個の各再生スピーカがユーザに対して前方や後方など、どこに位置しているかに応じて、各再生スピーカの再生ゲインを補正し、再生スピーカの位置によって出力される音声に違和感が生じないようになされる。すなわち、理想スピーカとの距離Dist(m,n)が同じである、ユーザの前方と後方にある2つの再生スピーカで、この理想スピーカのオーディオ信号を再生する場合、ユーザの後方にある再生スピーカの再生ゲインは、前方の再生スピーカの再生ゲインより小さくなるように補正が行なわれる。

[0114] 具体的には、まず再生装置は、メタデータから再生スピーカの配置位置に応じた再生ゲインの補正が必要か否かを示す情報を取得し、取得した情報が、再生ゲインの補正が不要である旨の情報である場合には、処理STE3を

行なわない。すなわち、処理STE 2の後、処理STE 3はスキップされて処理STE 4が行なわれる。

[0115] 一方、メタデータから取得した情報が、再生ゲインの補正が必要である旨の情報である場合には、再生装置は、式(1)と同様の計算を行なってN個の各再生スピーカと、空間原点Cとの距離 $\text{Dist}(n, C)$ を求める。

[0116] ここで、空間原点Cとは、再生スピーカが配置される空間上の基準となる位置であり、例えば空間原点Cの位置は、水平方向角度 $\theta = 0$ 、垂直方向角度 $\gamma = 0$ 、および距離 $r = r_0$ により表現される位置とされる。この場合、空間原点Cの位置は単位円、つまり図1の球PH 1 1上にあり、かつユーザU 1 1の正面に位置することになる。このような空間原点Cの位置は、理想的なセンタースピーカの位置である。

[0117] N個の再生スピーカについて空間原点Cとの距離 $\text{Dist}(n, C)$ が求められると、次式(11)の計算によりN個の各再生スピーカの補正係数 $\text{spkr_pos_correction_coefficient}(n)$ が求められる。

[0118] [数11]

$$\begin{aligned} & \text{spkr_pos_correction_coefficient}(n) \\ &= \text{Max_spkr_pos_correction_coefficient} \times \frac{\text{Dist}(n, c)}{180^\circ} \\ & \quad \dots (11) \end{aligned}$$

[0119] なお、式(11)において、Max_spkr_pos_correction_coefficientは、距離 $\text{Dist}(n, C)$ が最大(180°)となるときの補正係数を示している。

[0120] さらに、求められた補正係数 $\text{spkr_pos_correction_coefficient}(n)$ が、m番目の理想スピーカについてのn番目の再生スピーカの再生ゲイン $\text{MixGain}(m, n)$ に乗算されて、補正後の再生ゲイン $\text{MixGain_pos_corr}(m, n)$ とされる。すなわち、次式(12)の計算が行なわれる。

[0121]

[数12]

$$\begin{aligned} & \text{MixGain_pos_corr}(m, n) \\ &= \text{MixGain}(m, n) \times \text{spkr_pos_correction_coeffcient}(n) \\ & \quad - \text{MaxMixGain}(n) \times (\text{Max_spkr_pos_correction_coeffcient} - 1) \\ & \quad \times \frac{\text{Dist}(n, c)}{180^\circ} \quad \cdot \cdot \cdot (12) \end{aligned}$$

[0122] なお、式(12)において、MaxMixGain(n)は、n番目の再生スピーカについてのM個の再生ゲイン、つまりnの値が同じである再生ゲインMixGain(m, n)の最大値である。式(12)において、MaxMixGain(n)が含まれる項は、spkr_pos_correction_coeffcient(n)による補正が行なわれ過ぎないようにするための逆補正の項である。

[0123] 以上の処理により、再生スピーカの配置位置に応じて適宜補正された、M×N個の再生ゲインMixGain_pos_corr(m, n)が得られる。

[0124] なお、再生スピーカの配置位置に応じた再生ゲインの補正が行なわれない場合には、再生ゲインMixGain(m, n)がそのまま再生ゲインMixGain_pos_corr(m, n)とされる。

[0125] 〈処理STE4について〉

また、処理STE3の後に行なわれる処理STE4では、全ての再生スピーカの再生ゲインが小さい値となっている理想スピーカのオーディオ信号に対して、少なくとも1つの再生スピーカにおいて、予め定められた再生ゲインの下限值でオーディオ信号が再生されるように再生ゲインが補正される。

[0126] すなわち、処理STE3で得られた理想スピーカごとの再生ゲイン、つまりmの値が同じであるN個の再生ゲインMixGain_pos_corr(m, n)の最大値MaxMixGain_i(m)が求められ、その最大値MaxMixGain_i(m)と下限値MixGain_{MinThre}とが比較される。

[0127] そして、所定のm番目の理想スピーカについて、最大値MaxMixGain_i(m)が下限値MixGain_{MinThre}より小さい場合、そのm番目の理想スピーカについてのN個

の再生ゲイン $\text{MixGain_pos_corr}(m, n)$ に対して、補正值 $\text{MinGain}_{\text{correction}_i}(m)$ が加算される。ここで、補正值 $\text{MinGain}_{\text{correction}_i}(m)$ は、次式(13)に示すように最大値 $\text{MaxMixGain}_i(m)$ と下限値 $\text{MixGain}_{\text{MinThre}}$ との差分である。

[0128] [数13]

$$\text{MinGain}_{\text{correction}_i}(m) = \text{MaxMixGain}_i(m) - \text{MixGain}_{\text{MinThre}} \quad \cdot \cdot \cdot (13)$$

[0129] このような補正によって、少なくとも1つの再生スピーカによりチャンネル m のオーディオ信号が所定の最低再生ゲインで再生されることになり、特定チャンネルの音声聞こえなくなってしまうことを防止することができる。

[0130] 〈処理S T E 5について〉

また、処理S T E 5では、出力音声全体のエネルギーが、入力音声全体のエネルギーと近くなるように再生ゲイン $\text{MixGain_pos_corr}(m, n)$ が補正される。

[0131] まず、再生装置は、メタデータから理想スピーカの各チャンネル間の相対音圧の期待値 $\text{SPR}_i(m)$ を読み出し、そのうちの一番音圧の大きい理想スピーカの絶対音圧を0dBFSと仮定し、理想スピーカごとの各期待値 $\text{SPR}_i(m)$ から各チャンネルのオーディオ信号の音声の音圧を計算し、入力オーディオ信号の音声全体のパワー値 pow_i を求める。

[0132] ここで、パワー値 pow_i は、M個の各チャンネルのオーディオ信号を再生することで理想スピーカから出力される音声（以下、入力音声とも称する）の全体のパワーである。また、以下、N個の各チャンネルのオーディオ信号を再生することで再生スピーカから出力される音声を出力音声とも称することとする。

[0133] 次に、再生装置は処理S T E 4で得られた再生ゲイン $\text{MixGain_pos_corr}(m, n)$ に対して、期待値 $\text{SPR}_i(m)$ を乗算することで、各再生スピーカの出力音声の音圧の期待値 $\text{SPR}_o(n)$ を求め、期待値 $\text{SPR}_o(n)$ から出力音声全体のパワー値 pow_o を求める。

[0134] そして、再生装置は処理S T E 4で得られた全ての再生ゲイン MixGain_pos

$_corr(m,n)$ に、入力音声と出力音声のパワー値の比 (pow_o/pow_i) を乗算して、出力音声全体の音圧補正を行なう。このようにして得られた再生ゲインが、理想スピーカごとの各再生スピーカの最終的な再生ゲインとなる。

[0135] ここで、一番音圧の大きい理想スピーカの絶対音圧を0dBと仮定して、入力音声と出力音声のパワー値の比 (pow_o/pow_i) を求めたが、その値は実際の絶対音圧を用いて求めた入力音声と出力音声のパワー値の比 (pow_o/pow_i) と同じ値となる。このように入力音声の絶対音圧を仮定すれば、実際の入力音声の絶対音圧が分からない場合でも、入力音声と出力音声のパワー値の比 (pow_o/pow_i) を求めることが可能となる。なお、仮定した音圧値を0dBではなく、他の任意の値としても、結果として得られるパワー値の比は同じ値となる。

[0136] 〈LFE用のスピーカについて〉

さらに、LFE用のチャンネルのオーディオ信号の再生について説明する。

[0137] 例えば、LFE用の理想スピーカの数、0個、1個、または2個のうちの何れかとなり、同様にLFE用の再生スピーカの数も0個、1個、または2個のうちの何れかとなる。

[0138] LFE用の理想スピーカの数、またはLFE用の再生スピーカの数、何れかが0個である場合には、LFE用のチャンネルのオーディオ信号は再生不可能であり、そのオーディオ信号のゲインは $-\infty$ とされる。

[0139] これに対して、LFE用の理想スピーカや再生スピーカの数、1個または2個である場合には、再生装置では例えば図4に示す再生ゲインにより、LFE用の各チャンネルのオーディオ信号が生成される。

[0140] すなわち、LFE用の理想スピーカとLFE用の再生スピーカがともに1つまたは2つである場合には、LFE用の理想スピーカのオーディオ信号が、そのままLFE用の再生スピーカのオーディオ信号とされて再生される。

[0141] また、LFE用の理想スピーカが1つであり、LFE用の再生スピーカが2つである場合、またはLFE用の理想スピーカが2つであり、LFE用の再生スピーカが1つである場合には、各チャンネルのオーディオ信号が均等に分配される。

。

[0142] すなわち、LFE用の理想スピーカ1つに対して、LFE用の再生スピーカが2つである場合、理想スピーカのオーディオ信号が同じ再生ゲインでゲイン調整され、それぞれ2つの再生スピーカで再生される。また、LFE用の理想スピーカ2つに対して、LFE用の再生スピーカが1つである場合、理想スピーカのオーディオ信号が同じ再生ゲインで足し合わされて1つのオーディオ信号とされ、再生スピーカで再生される。

[0143] 〈再生装置の構成例〉

続いて、以上において説明した再生装置の具体的な実施の形態について説明する。

[0144] 再生装置は、例えば図5に示すように構成される。

[0145] 図5に示す再生装置11は、図示せぬデコーダ等からメタデータとオーディオ信号の供給を受けて、メタデータに基づいてオーディオ信号のゲイン調整を行い、その結果得られたオーディオ信号をスピーカ12-1乃至スピーカ12-Nに供給する。

[0146] なお、図5では再生装置11のうちのLFE用でないチャンネルのオーディオ信号を再生するための機能ブロックのみが図示されており、LFE用のチャンネルのオーディオ信号を再生するための機能ブロックの図示は省略されている。

[0147] また、図5では、LFE用でないM個の各理想スピーカに対応するMチャンネルのオーディオ信号が供給され、それらのMチャンネルのオーディオ信号がNチャンネルのオーディオ信号に変換されて出力されるものとする。さらに、スピーカ12-1乃至スピーカ12-Nは、以上において説明したLFE用でない再生スピーカに対応する。

[0148] 以下、スピーカ12-1乃至スピーカ12-Nを特に区別する必要のない場合、単にスピーカ12とも称することとする。また、各スピーカ12は、以上において説明した再生スピーカRSP11に対応するスピーカであるので、スピーカ12を再生スピーカ12とも称することとする。

- [0149] 図5に示す再生装置11には、距離計算部21、再生ゲイン計算部22、補正部23、下限値補正部24、全体ゲイン補正部25、およびゲイン調整部26が設けられている。また、ゲイン調整部26は、増幅部31、増幅部32、および増幅部33を備えている。
- [0150] 距離計算部21には、メタデータに含まれているLFE用ではない各理想スピーカの位置情報と、各再生スピーカ12の位置情報とが供給される。距離計算部21は、理想スピーカの位置情報と再生スピーカ12の位置情報に基づいて距離 $\text{Dist}(m, n)$ を算出し、再生ゲイン計算部22に供給する。
- [0151] ここで、各スピーカの位置情報とは水平方向角度 θ 、垂直方向角度 γ 、および距離 r からなる情報である。
- [0152] また、距離計算部21は、必要に応じて理想スピーカ側の補正值 $\text{SoundPressureCorrection}_{im}$ および遅延時間 Delay_{im} を算出して増幅部31に供給するとともに、再生スピーカ12側の補正值 $\text{SoundPressureCorrection}_{on}$ および遅延時間 Delay_{on} を算出して増幅部33に供給する。つまり、距離計算部21では処理STE1が行なわれる。
- [0153] 再生ゲイン計算部22には、メタデータに含まれているカーブ情報とカーブインデックスが供給され、再生ゲイン計算部22は、カーブ情報やカーブインデックスと、距離計算部21から供給された距離とを用いて再生ゲイン $\text{MixGain}(m, n)$ を算出し、補正部23に供給する。すなわち、再生ゲイン計算部22では処理STE2が行なわれる。
- [0154] 補正部23には、再生スピーカ12の位置情報と、メタデータに含まれている、再生スピーカ12の配置位置に応じた再生ゲインの補正が必要か否かを示す情報、および補正係数 $\text{Max_spkr_pos_correction_coeffcient}$ とが供給される。
- [0155] 補正部23は、これらの供給された情報に基づいて、再生ゲイン計算部22から供給された再生ゲインに対して再生スピーカ12の配置位置に応じた補正を行い、その結果得られた再生ゲイン $\text{MixGain_pos_corr}(m, n)$ を下限値補正部24に供給する。すなわち、補正部23では処理STE3が行なわれる。

。

[0156] 下限値補正部 24 には、メタデータに含まれている、再生ゲインの下限值 $MixGain_{MinThre}$ が供給される。下限値補正部 24 は、下限値 $MixGain_{MinThre}$ に基づいて補正部 23 から供給された再生ゲインの補正を行い、全体ゲイン補正部 25 に供給する。すなわち、下限値補正部 24 では、処理 STE 4 が行なわれる。

[0157] 全体ゲイン補正部 25 には、メタデータに含まれている理想スピーカの各チャンネル間の相対音圧の期待値 $SPR_i(m)$ が供給される。全体ゲイン補正部 25 は、期待値 $SPR_i(m)$ に基づいて、下限値補正部 24 から供給された再生ゲインの補正を行い、その結果得られた最終的な再生ゲインを増幅部 32 に供給する。全体ゲイン補正部 25 では、処理 STE 5 が行なわれる。

[0158] ゲイン調整部 26 は、図示せぬデコーダから供給された M 個の理想スピーカのオーディオ信号に対してゲイン調整を行なって N チャンネルのオーディオ信号を生成し、各チャンネルのオーディオ信号を再生スピーカ 12 に供給して音声を再生させる。ゲイン調整部 26 では、処理 STE 6 が行なわれる。

。

[0159] すなわち、増幅部 31 は、距離計算部 21 から供給された補正值および遅延時間に基づいて、供給された M チャンネルのオーディオ信号に対して、適宜、ゲイン補正および遅延処理を施して増幅部 32 に供給する。

[0160] 増幅部 32 は、増幅部 31 から供給された M チャンネルのオーディオ信号に、全体ゲイン補正部 25 から供給された再生ゲインを乗算する。また、増幅部 32 は、再生ゲインが乗算された各理想スピーカのオーディオ信号を加算することで N チャンネルのオーディオ信号を生成し、増幅部 33 に供給する。

[0161] 増幅部 33 は、距離計算部 21 から供給された補正值および遅延時間に基づいて、増幅部 32 から供給された N チャンネルのオーディオ信号に対して、適宜、ゲイン補正および遅延処理を施して再生スピーカ 12 に供給する。

[0162] 〈ダウンミックス処理の説明〉

続いて、再生装置 11 の動作について説明する。

[0163] 再生装置 11 に各理想スピーカのオーディオ信号とメタデータが供給されると、再生装置 11 は、LFE用のオーディオ信号と、LFE用ではないオーディオ信号とに対してそれぞれ再生スピーカに供給するオーディオ信号を生成し、出力する。

[0164] 以下、図 6 のフローチャートを参照して、再生装置 11 がLFE用ではないオーディオ信号に対して行なうダウンミックス処理について説明する。

[0165] ステップ S 11 において、距離計算部 21 は、メタデータに含まれているLFE用でない理想スピーカの位置情報と、LFE用でない再生スピーカ 12 の位置情報に基づいて、理想スピーカと再生スピーカ 12 の距離 $\text{Dist}(m, n)$ を求め、再生ゲイン計算部 22 に供給する。具体的には、理想スピーカと再生スピーカ 12 の組み合わせごとに式 (1) の計算が行なわれ、 $M \times N$ 個の距離 $\text{Dist}(m, n)$ が求められる。

[0166] ステップ S 12 において、距離計算部 21 は必要に応じて、理想スピーカ側と再生スピーカ 12 側の補正值および遅延時間を求める。

[0167] 具体的には距離計算部 21 は、距離 $r_{im} \neq r_u$ である理想スピーカについて、理想スピーカの位置情報としての距離 r_{im} に基づいて式 (2) および式 (3) の計算を行い、補正值 $\text{SoundPressureCorrection}_{im}$ および遅延時間 Delay_{im} を算出して増幅部 31 に供給する。

[0168] また、距離計算部 21 は、距離 $r_{on} \neq r_u$ である再生スピーカについて、再生スピーカ 12 の位置情報としての距離 r_{on} に基づいて式 (4) および式 (5) の計算を行い、補正值 $\text{SoundPressureCorrection}_{on}$ および遅延時間 Delay_{on} を算出して増幅部 33 に供給する。

[0169] ステップ S 13 において、再生ゲイン計算部 22 は、距離計算部 21 から供給された距離 $\text{Dist}(m, n)$ に基づいて、理想スピーカごとに各再生スピーカ 12 の再生ゲインを求める。

[0170] 例えば、再生ゲイン計算部 22 は、理想スピーカと再生スピーカ 12 との距離 $\text{Dist}(m, n)$ が「0」となる再生スピーカ 12 が存在する理想スピーカにつ

いては、式（６）の計算を行なって、その理想スピーカについての各再生スピーカ 1 2 の再生ゲイン $MixGain(m, n)$ を算出する。

[0171] また、再生ゲイン計算部 2 2 は、距離 $Dist(m, n) = 0$ となる再生スピーカ 1 2 が存在しない理想スピーカについては、メタデータに含まれているカーブ情報に示されるカーブ、すなわち折れ線カーブまたは関数カーブを得る。このとき、再生ゲイン計算部 2 2 は、カーブインデックスを参照し、必要に応じてメタデータから折れ線カーブまたは関数カーブを読み出す。

[0172] 再生ゲイン計算部 2 2 は、折れ線カーブまたは関数カーブが得られると、得られたカーブに基づいて距離 $Dist(m, n)$ に対応するゲイン値を求め、求めたゲイン値を理想スピーカの再生スピーカ 1 2 についての再生ゲイン $MixGain(m, n)$ とする。このとき、必要に応じて式（７）や式（９）の計算が行なわれる。

[0173] 再生ゲイン計算部 2 2 は、各理想スピーカについて、再生スピーカ 1 2 ごとの再生ゲイン $MixGain(m, n)$ が得られると、再生ゲイン $MixGain(m, n)$ を補正部 2 3 に供給する。

[0174] ステップ S 1 4 において、補正部 2 3 は、メタデータに含まれている再生ゲインの補正が必要か否かを示す情報に基づいて、必要に応じて再生ゲイン計算部 2 2 から供給された再生ゲインを再生スピーカ 1 2 の配置位置に応じて補正し、下限値補正部 2 4 に供給する。

[0175] 具体的には、補正部 2 3 は、各再生スピーカ 1 2 の位置情報と、メタデータに含まれている補正係数 $Max_spkr_pos_correction_coeffcient$ とを用いて、式（１１）および式（１２）を計算することで、再生ゲイン $MixGain_pos_corr(m, n)$ を算出する。

[0176] ステップ S 1 5 において、下限値補正部 2 4 は必要に応じて、メタデータに含まれている下限値 $MixGain_{MinThre}$ に基づいて、補正部 2 3 から供給された再生ゲインを補正し、全体ゲイン補正部 2 5 に供給する。具体的には、必要に応じて式（１３）が計算され、再生ゲイン $MixGain_pos_corr(m, n)$ に補正值 $MinGain_{correctioni}(m)$ が加算される。

- [0177] ステップS 1 6において、全体ゲイン補正部 2 5は出力音声全体の音圧補正を行なう。
- [0178] すなわち、全体ゲイン補正部 2 5は、メタデータに含まれている期待値 $SPR_i(m)$ と、下限値補正部 2 4から供給された再生ゲイン $MixGain_pos_corr(m, n)$ とに基づいて、入力音声と出力音声の全体のパワー値の比 (pow_o/pow_i) を算出する。そして、全体ゲイン補正部 2 5は、パワー値の比 (pow_o/pow_i) を、再生ゲイン $MixGain_pos_corr(m, n)$ に乗算して最終的な再生ゲインとし、増幅部 3 2に供給する。
- [0179] ステップS 1 7において、増幅部 3 1は、距離計算部 2 1から供給された理想スピーカ側の補正值および遅延値に基づいて、オーディオ信号のゲイン調整を行なう。
- [0180] 具体的には増幅部 3 1は、補正值と遅延値が供給されたチャンネル m のオーディオ信号について、オーディオ信号に補正值 $SoundPressureCorrection_{im}$ を乗算し、その結果得られたオーディオ信号を遅延時間 $Delay_{im}$ だけ時間方向に遅延させて増幅部 3 2に供給する。
- [0181] ステップS 1 8において、増幅部 3 2は、全体ゲイン補正部 2 5から供給された再生ゲインと、増幅部 3 1から供給されたオーディオ信号とに基づいて、各再生スピーカ 1 2のオーディオ信号を生成し、増幅部 3 3に供給する。
- [0182] 具体的には、増幅部 3 2は、再生スピーカ 1 2に対応する N 個のチャンネルのうちの1つを注目チャンネル nc とすると、注目チャンネル nc についての各理想スピーカの再生ゲインを、それらの理想スピーカのオーディオ信号に乗算する。そして、増幅部 3 2は、再生ゲインが乗算された各理想スピーカのオーディオ信号、つまり M 個のオーディオ信号を足し合わせて得られる1つのオーディオ信号を、注目チャンネル nc のオーディオ信号とする。 N 個の各チャンネルを注目チャンネルとして同様の処理を行なうことで、 M 個の各理想スピーカのオーディオ信号が、 N 個の各再生スピーカ 1 2のオーディオ信号へと変換される。

- [0183] ステップS 19において、増幅部33は、距離計算部21から供給された再生スピーカ12側の補正值および遅延値に基づいて、増幅部32から供給されたオーディオ信号のゲイン調整を行なう。
- [0184] 具体的には増幅部33は、補正值と遅延値が供給されたチャンネルnのオーディオ信号について、オーディオ信号に補正值 $\text{SoundPressureCorrection}_{on}$ を乗算し、その結果得られたオーディオ信号を遅延時間 Delay_{on} だけ時間方向に遅延させて再生スピーカ12に供給する。
- [0185] 各チャンネルのオーディオ信号が再生スピーカ12に出力されると、ダウンミックス処理は終了する。また、再生スピーカ12では、再生装置11から供給されたオーディオ信号に基づいて音声再生される。
- [0186] 以上のようにして、再生装置11は、理想スピーカの位置と実際の再生スピーカ12の配置位置との距離に応じてオーディオ信号のゲイン調整（ゲイン補正）を行なう。これにより、理想スピーカと再生スピーカ12との位置にずれがある場合であっても、出力音声の音質と音像定義の劣化を抑制することができ、より臨場感のあるオーディオ再生を行なうことができる。
- [0187] 以上において説明した処理によって、入力された任意の1以上のチャンネルのオーディオ信号を、1以上の任意の個数の任意の位置に配置された再生スピーカで再生することができるようになる。また、入力される各チャンネルのオーディオ信号が各オブジェクトを音源とするオーディオ信号である場合であっても、同様のダウンミックス処理により、正しい音像位置でのオーディオ再生を行なうことができる。
- [0188] 〈エンコーダとデコーダについて〉
- 次に、再生装置11に供給されるメタデータを符号化するエンコーダと、符号化されたメタデータを復号するデコーダについて説明する。
- [0189] 例えば図7に示すように、本技術を適用したオーディオシステムでは、エンコーダ61からデコーダ62にメタデータが供給され、さらにデコーダ62から再生装置11へとメタデータが供給される。
- [0190] エンコーダ61は、外部からメタデータを得るために必要な情報と、M個

の各理想スピーカのオーディオ信号とを取得して、符号化されたメタデータとオーディオ信号とからなるビットストリームを生成する。

[0191] エンコーダ 6 1 は、メタデータ生成部 7 1、オーディオ信号符号化部 7 2、および出力部 7 3 を備えている。

[0192] メタデータ生成部 7 1 は、外部から必要な情報を取得するとともに、取得した情報を必要に応じて符号化し、符号化されたメタデータを生成する。

[0193] 例えば、メタデータには、各理想スピーカの位置情報、理想スピーカのうちのLFE用の理想スピーカの数（チャンネル数）、およびカーブ情報、カーブインデックスが含まれている。また、メタデータには、再生スピーカ 1 2 の配置位置に応じた再生ゲインの補正が必要か否かを示す情報、再生スピーカ 1 2 の配置による補正係数Max_sprk_pos_correction_coefficient、ゲインの下限值MixGain_{MinThre}、およびチャンネル間の相対音圧の期待値SPR_{i(m)}も含まれている。

[0194] オーディオ信号符号化部 7 2 は、外部から供給されたオーディオ信号を符号化する。出力部 7 3 は、符号化されたメタデータと、符号化されたオーディオ信号とが含まれるビットストリームを生成し、デコーダ 6 2 に出力する。

[0195] また、デコーダ 6 2 は、抽出部 8 1、オーディオ信号復号部 8 2、および出力部 8 3 を備えている。デコーダ 6 2 は、エンコーダ 6 1 から送信されてきたビットストリームを受信し、抽出部 8 1 は、受信されたビットストリームからメタデータとオーディオ信号を抽出する。このとき、抽出部 8 1 は必要に応じてメタデータの復号を行なう。

[0196] オーディオ信号復号部 8 2 は、抽出部 8 1 により抽出されたオーディオ信号を復号する。出力部 8 3 は、抽出部 8 1 により抽出されたメタデータと、オーディオ信号復号部 8 2 により復号されたオーディオ信号とを再生装置 1 1 に供給する。

[0197] さらに、エンコーダ 6 1 からデコーダ 6 2 に出力されるビットストリームに記述されるメタデータの一部は、例えば図 8 に示すようになる。すなわち

、図8はメタデータの一部のシンタックスを示している。

[0198] 図8の例では、ヘッダの先頭には、ダウンミックスに必要な情報がメタデータに含まれているか否かを示す情報として「down mix coef exist flag」が配置されている。

[0199] また、メタデータにはカーブ情報として「down mix coef mode」が配置されており、その下には、カーブインデックスとしての「polyline curve idx」または「function curve idx」が配置される。

[0200] 「polyline curve idx」は、折れ線カーブを示しており、この値が2進数「111」である場合には、新たな折れ線カーブであることを示している。この場合、新たな折れ線カーブを得るための情報として「polyline curve coefficient[j]」が記述されている。

[0201] 新たな折れ線カーブを得るための情報は、例えば図2に示した折れ線C V 1 1上の四角形の各点（以下、記述点と称する）、つまり数列を構成する各値を特定する情報とされる。

[0202] 具体的には、例えば再生ゲイン軸（縦軸）が16分割されるとし、16本の分割線が定義される。各記述点は、縦軸の各分割線上に順次配置されるとする。

[0203] メタデータ内では、記述点が「0」で表され、各記述点がどの分割線上に配置されるかを示す情報は「1」で表されることとする。

[0204] 図2中、左から順番に記述点を記述することとし、まず、左から1番目の記述点が上から何本下の分割線上に位置するかを示す情報が「1」の数で記述され、その後、記述点を表す「0」が記述される。ここでは、左から1番目の記述点は最上位の分割線上に位置するため、記述点を表す「0」だけが記述される。

[0205] また、それ以降においては、その記述点が、直前の記述点が位置する分割線からQ本下の分割線上に位置することを示す情報がQ個の「1」で記述され、その後に記述点を表す「0」が記述される。

[0206] 例えば、左から3番目の記述点は2番目の記述点より2本下の分割線上に

位置するため、2個の「1」が記述され、その後、1個の「0」が記述される。また、左から10番目の記述点は9番目の記述点と同じ分割線上に位置する、すなわち0本下の分割線上に位置するので、「1」が記述されず、1個の「0」だけが記述される。

[0207] 上記の方法により記述が行なわれ、全ての記述点の記述が完了した場合、1個の「1」が記述され、折れ線カーブの情報の記述が終了したことが表される。但し、記述点の数が多く、合わせて64個の「1」と「0」を用いても記述しきれない場合には、「1」と「0」の数が64に達するまで記述が行なわれ、その後は記述終了とされる。

[0208] したがって、折れ線カーブを得るための情報をメタデータから読み出す場合には、16個の「1」、または合計で64個の「1」と「0」（すなわち、「1」と「0」が合計で64個）が読み出されるまで、順番に各記述点を得るための情報が読み出され、折れ線カーブが生成される。

[0209] また、「function curve idx」は、関数カーブを示しており、この値が2進数の「111」である場合には、新たな関数カーブであることを示している。この場合、新たな関数カーブの係数として「function_curve_coefficient[i]」が記述されている。

[0210] また、メタデータに記述されている「minimun_gain_threshold_idx」は、ゲインの下限值 $MixGain_{MinThre}$ を示すインデックスである。さらに、メタデータに記述されている「gain_correction_coefficient」は再生スピーカ12の配置位置に応じた再生ゲインの補正に必要な補正係数 $Max_spkr_pos_correction_coefficient$ である。ここで、 $Max_spkr_pos_correction_coefficient$ の値が「1」である場合には、再生スピーカ12の配置位置に応じた再生ゲインの補正は必要ないことを示していることになる。

[0211] さらに、メタデータには、チャンネル間の相対音圧の期待値 $SPR_i(m)$ がメタデータ内に記述されているか否かを示す情報である「sound_level_exist_flag」が記述されており、この「sound_level_exist_flag」の値に応じて「channel sound level[i]」が記述される。「channel sound level[i]」は、期

待値SPR_i(m)である。

[0212] 〈符号化処理の説明〉

さらに、エンコーダ61とデコーダ62の動作について説明する。

[0213] まず、図9のフローチャートを参照して、エンコーダ61による符号化処理について説明する。

[0214] ステップS41において、メタデータ生成部71は、外部から必要な情報を取得するとともに、取得した情報の符号化を行なうことで、符号化されたメタデータを生成する。例えば、メタデータ生成部71は、図8に示したシンタックスに対応するメタデータを生成する。

[0215] ステップS42において、オーディオ信号符号化部72は、外部から供給されたオーディオ信号を符号化する。

[0216] ステップS43において、出力部73は、符号化されたメタデータと、符号化されたオーディオ信号とが含まれるビットストリームを生成し、デコーダ62に出力する。そして、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。

[0217] 以上のようにしてエンコーダ61は、理想スピーカの位置情報や、カーブ情報などが含まれるメタデータを生成し、出力する。このように理想スピーカの位置情報やカーブ情報などからなる情報をメタデータとして生成することで、再生装置11において、理想スピーカの位置と実際の再生スピーカ12の配置位置との距離に応じたゲイン補正など、適切なゲイン補正を行なうことができるようになる。その結果、より臨場感のあるオーディオ再生を行なうことができる。

[0218] 〈復号処理の説明〉

続いて、図10のフローチャートを参照して、デコーダ62により行なわれる復号処理について説明する。

[0219] ステップS71において、デコーダ62はエンコーダ61から送信されてきたビットストリームを受信し、抽出部81は、受信されたビットストリームからメタデータとオーディオ信号を抽出する。また、抽出部81は、メタ

データの復号も行なう。

[0220] ステップS 7 2において、オーディオ信号復号部 8 2は、抽出部 8 1により抽出されたオーディオ信号を復号する。

[0221] ステップS 7 3において、出力部 8 3は、復号されたメタデータと、復号されたオーディオ信号とを再生装置 1 1に出力し、復号処理は終了する。

[0222] 以上のようにしてデコーダ 6 2は、メタデータとオーディオ信号を復号し、理想スピーカの位置情報や、カーブ情報などが含まれるメタデータとオーディオ信号を再生装置 1 1に出力する。このように理想スピーカの位置情報やカーブ情報などからなる情報をメタデータとして出力することで、再生装置 1 1において、理想スピーカの位置と実際の再生スピーカ 1 2の配置位置との距離に応じたゲイン補正など、適切なゲイン補正を行なうことができるようになる。その結果、より臨場感のあるオーディオ再生を行なうことができる。

[0223] 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、例えば汎用のコンピュータなどが含まれる。

[0224] 図 1 1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。

[0225] コンピュータにおいて、CPU 5 0 1、ROM 5 0 2、RAM 5 0 3は、バス 5 0 4により相互に接続されている。

[0226] バス 5 0 4には、さらに、入出力インターフェース 5 0 5が接続されている。入出力インターフェース 5 0 5には、入力部 5 0 6、出力部 5 0 7、記録部 5 0 8、通信部 5 0 9、及びドライブ 5 1 0が接続されている。

[0227] 入力部 5 0 6は、キーボード、マウス、マイクロホン、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メディア511を駆動する。

[0228] 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。

[0229] 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メディア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。

[0230] コンピュータでは、プログラムは、リムーバブルメディア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。

[0231] 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。

[0232] 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。

[0233] 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。

。

[0234] 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。

[0235] 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。

[0236] さらに、本技術は、以下の構成とすることも可能である。

[0237] [1]

オーディオ信号を再生する理想的なスピーカの位置と、前記オーディオ信号を再生する実際のスピーカの位置との距離を計算する距離計算部と、

前記距離に基づいて前記オーディオ信号の再生ゲインを計算するゲイン計算部と、

前記再生ゲインに基づいて前記オーディオ信号のゲイン調整を行なうゲイン調整部と

を備えるオーディオ信号出力装置。

[2]

前記ゲイン計算部は、各前記距離に対する前記再生ゲインを得るためのカーブ情報に基づいて、前記再生ゲインを計算する

[1]に記載のオーディオ信号出力装置。

[3]

前記カーブ情報は、折れ線カーブまたは関数カーブを示す情報である

[2]に記載のオーディオ信号出力装置。

[4]

前記ゲイン調整部は、前記理想的なスピーカが所定の基準点を中心とする単位円上にない場合、前記基準点から前記理想的なスピーカまでの距離と前記単位円の半径とに基づいて定められたゲインにより前記オーディオ信号のゲイン調整をさらに行う

[1]または[2]に記載のオーディオ信号出力装置。

[5]

前記ゲイン調整部は、前記基準点から前記理想的なスピーカまでの距離と前記単位円の半径とに基づいて定められた遅延時間に基づいて前記オーディオ信号を遅延させる

[4] に記載のオーディオ信号出力装置。

[6]

前記ゲイン調整部は、前記実際のスピーカが所定の基準点を中心とする単位円上にない場合、前記基準点から前記実際のスピーカまでの距離と前記単位円の半径とに基づいて定められたゲインにより前記オーディオ信号のゲイン調整をさらに行う

[1] または [2] に記載のオーディオ信号出力装置。

[7]

前記ゲイン調整部は、前記基準点から前記実際のスピーカまでの距離と前記単位円の半径とに基づいて定められた遅延時間に基づいて前記オーディオ信号を遅延させる

[6] に記載のオーディオ信号出力装置。

[8]

理想的なセンタースピーカの位置と、前記実際のスピーカの位置との距離に基づいて、前記再生ゲインを補正するゲイン補正部をさらに備える

[1] 乃至 [7] の何れかに記載のオーディオ信号出力装置。

[9]

前記再生ゲインが予め定められた下限値よりも小さい場合、前記再生ゲインを補正する下限値補正部をさらに備える

[1] 乃至 [8] の何れかに記載のオーディオ信号出力装置。

[1 0]

入力された前記オーディオ信号に基づく入力音声の音圧の期待値、および前記再生ゲインに基づいて、前記再生ゲインによりゲイン調整された前記オーディオ信号に基づく出力音声の全体のパワーと、前記入力音声の全体のパ

ワーとの比を算出し、前記比に基づいて前記再生ゲインを補正する全体ゲイン補正部をさらに備える

[1] 乃至 [9] の何れかに記載のオーディオ信号出力装置。

[1 1]

オーディオ信号を再生する理想的なスピーカの位置と、前記オーディオ信号を再生する実際のスピーカの位置との距離を計算し、

前記距離に基づいて前記オーディオ信号の再生ゲインを計算し、

前記再生ゲインに基づいて前記オーディオ信号のゲイン調整を行なう

ステップを含むオーディオ信号出力方法。

[1 2]

オーディオ信号を再生する理想的なスピーカの位置と、前記オーディオ信号を再生する実際のスピーカの位置との距離を計算し、

前記距離に基づいて前記オーディオ信号の再生ゲインを計算し、

前記再生ゲインに基づいて前記オーディオ信号のゲイン調整を行なう

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。

[1 3]

オーディオ信号を再生する理想的なスピーカの位置と、前記オーディオ信号を再生する実際のスピーカの位置との距離に応じて前記オーディオ信号のゲインを補正するための補正情報を生成する補正情報生成部と、

前記オーディオ信号を符号化する符号化部と、

前記補正情報および符号化された前記オーディオ信号を含むビットストリームを出力する出力部と

を備える符号化装置。

[1 4]

オーディオ信号を再生する理想的なスピーカの位置と、前記オーディオ信号を再生する実際のスピーカの位置との距離に応じて前記オーディオ信号のゲインを補正するための補正情報を生成し、

前記オーディオ信号を符号化し、

前記補正情報および符号化された前記オーディオ信号を含むビットストリームを出力する

ステップを含む符号化方法。

[1 5]

オーディオ信号を再生する理想的なスピーカの位置と、前記オーディオ信号を再生する実際のスピーカの位置との距離に応じて前記オーディオ信号のゲインを補正するための補正情報、および符号化された前記オーディオ信号をビットストリームから抽出する抽出部と、

前記符号化された前記オーディオ信号を復号する復号部と、

復号された前記オーディオ信号および前記補正情報を出力する出力部とを備える復号装置。

[1 6]

前記補正情報は、前記理想的なスピーカの位置情報である

[1 5] に記載の復号装置。

[1 7]

前記補正情報は、各前記距離に対するゲインを得るためのカーブ情報である

[1 5] または [1 6] に記載の復号装置。

[1 8]

前記カーブ情報は、折れ線カーブまたは関数カーブを示す情報である

[1 7] に記載の復号装置。

[1 9]

オーディオ信号を再生する理想的なスピーカの位置と、前記オーディオ信号を再生する実際のスピーカの位置との距離に応じて前記オーディオ信号のゲインを補正するための補正情報、および符号化された前記オーディオ信号をビットストリームから抽出し、

前記符号化された前記オーディオ信号を復号し、

復号された前記オーディオ信号および前記補正情報を出力する

ステップを含む復号方法。

符号の説明

[0238] 1 1 再生装置, 2 1 距離計算部, 2 2 再生ゲイン計算部, 2 3 補正部, 2 4 下限値補正部, 2 5 全体ゲイン補正部, 2 6 ゲイン調整部, 6 1 エンコーダ, 6 2 デコーダ, 7 1 メタデータ生成部, 7 2 オーディオ信号符号化部, 7 3 出力部, 8 1 抽出部, 8 2 オーディオ信号復号部, 8 3 出力部

請求の範囲

- [請求項1] オーディオ信号を再生する理想的なスピーカの位置と、前記オーディオ信号を再生する実際のスピーカの位置との距離を計算する距離計算部と、
- 前記距離に基づいて前記オーディオ信号の再生ゲインを計算するゲイン計算部と、
- 前記再生ゲインに基づいて前記オーディオ信号のゲイン調整を行なうゲイン調整部と
- を備えるオーディオ信号出力装置。
- [請求項2] 前記ゲイン計算部は、各前記距離に対する前記再生ゲインを得るためのカーブ情報に基づいて、前記再生ゲインを計算する
- 請求項1に記載のオーディオ信号出力装置。
- [請求項3] 前記カーブ情報は、折れ線カーブまたは関数カーブを示す情報である
- 請求項2に記載のオーディオ信号出力装置。
- [請求項4] 前記ゲイン調整部は、前記理想的なスピーカが所定の基準点を中心とする単位円上にない場合、前記基準点から前記理想的なスピーカまでの距離と前記単位円の半径とに基づいて定められたゲインにより前記オーディオ信号のゲイン調整をさらに行う
- 請求項1に記載のオーディオ信号出力装置。
- [請求項5] 前記ゲイン調整部は、前記基準点から前記理想的なスピーカまでの距離と前記単位円の半径とに基づいて定められた遅延時間に基づいて前記オーディオ信号を遅延させる
- 請求項4に記載のオーディオ信号出力装置。
- [請求項6] 前記ゲイン調整部は、前記実際のスピーカが所定の基準点を中心とする単位円上にない場合、前記基準点から前記実際のスピーカまでの距離と前記単位円の半径とに基づいて定められたゲインにより前記オーディオ信号のゲイン調整をさらに行う

請求項 1 に記載のオーディオ信号出力装置。

[請求項7] 前記ゲイン調整部は、前記基準点から前記実際のスピーカまでの距離と前記単位円の半径とに基づいて定められた遅延時間に基づいて前記オーディオ信号を遅延させる

請求項 6 に記載のオーディオ信号出力装置。

[請求項8] 理想的なセンタースピーカの位置と、前記実際のスピーカの位置との距離に基づいて、前記再生ゲインを補正するゲイン補正部をさらに備える

請求項 1 に記載のオーディオ信号出力装置。

[請求項9] 前記再生ゲインが予め定められた下限値よりも小さい場合、前記再生ゲインを補正する下限値補正部をさらに備える

請求項 1 に記載のオーディオ信号出力装置。

[請求項10] 入力された前記オーディオ信号に基づく入力音声の音圧の期待値、および前記再生ゲインに基づいて、前記再生ゲインによりゲイン調整された前記オーディオ信号に基づく出力音声の全体のパワーと、前記入力音声の全体のパワーとの比を算出し、前記比に基づいて前記再生ゲインを補正する全体ゲイン補正部をさらに備える

請求項 1 に記載のオーディオ信号出力装置。

[請求項11] オーディオ信号を再生する理想的なスピーカの位置と、前記オーディオ信号を再生する実際のスピーカの位置との距離を計算し、
前記距離に基づいて前記オーディオ信号の再生ゲインを計算し、
前記再生ゲインに基づいて前記オーディオ信号のゲイン調整を行なう

ステップを含むオーディオ信号出力方法。

[請求項12] オーディオ信号を再生する理想的なスピーカの位置と、前記オーディオ信号を再生する実際のスピーカの位置との距離を計算し、
前記距離に基づいて前記オーディオ信号の再生ゲインを計算し、
前記再生ゲインに基づいて前記オーディオ信号のゲイン調整を行な

う

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13] オーディオ信号を再生する理想的なスピーカの位置と、前記オーディオ信号を再生する実際のスピーカの位置との距離に応じて前記オーディオ信号のゲインを補正するための補正情報を生成する補正情報生成部と、

前記オーディオ信号を符号化する符号化部と、

前記補正情報および符号化された前記オーディオ信号を含むビットストリームを出力する出力部と

を備える符号化装置。

[請求項14] オーディオ信号を再生する理想的なスピーカの位置と、前記オーディオ信号を再生する実際のスピーカの位置との距離に応じて前記オーディオ信号のゲインを補正するための補正情報を生成し、

前記オーディオ信号を符号化し、

前記補正情報および符号化された前記オーディオ信号を含むビットストリームを出力する

ステップを含む符号化方法。

[請求項15] オーディオ信号を再生する理想的なスピーカの位置と、前記オーディオ信号を再生する実際のスピーカの位置との距離に応じて前記オーディオ信号のゲインを補正するための補正情報、および符号化された前記オーディオ信号をビットストリームから抽出する抽出部と、

前記符号化された前記オーディオ信号を復号する復号部と、

復号された前記オーディオ信号および前記補正情報を出力する出力部と

を備える復号装置。

[請求項16] 前記補正情報は、前記理想的なスピーカの位置情報である
請求項15に記載の復号装置。

[請求項17] 前記補正情報は、各前記距離に対するゲインを得るためのカーブ情

報である

請求項 1 5 に記載の復号装置。

[請求項18] 前記カーブ情報は、折れ線カーブまたは関数カーブを示す情報である

請求項 1 7 に記載の復号装置。

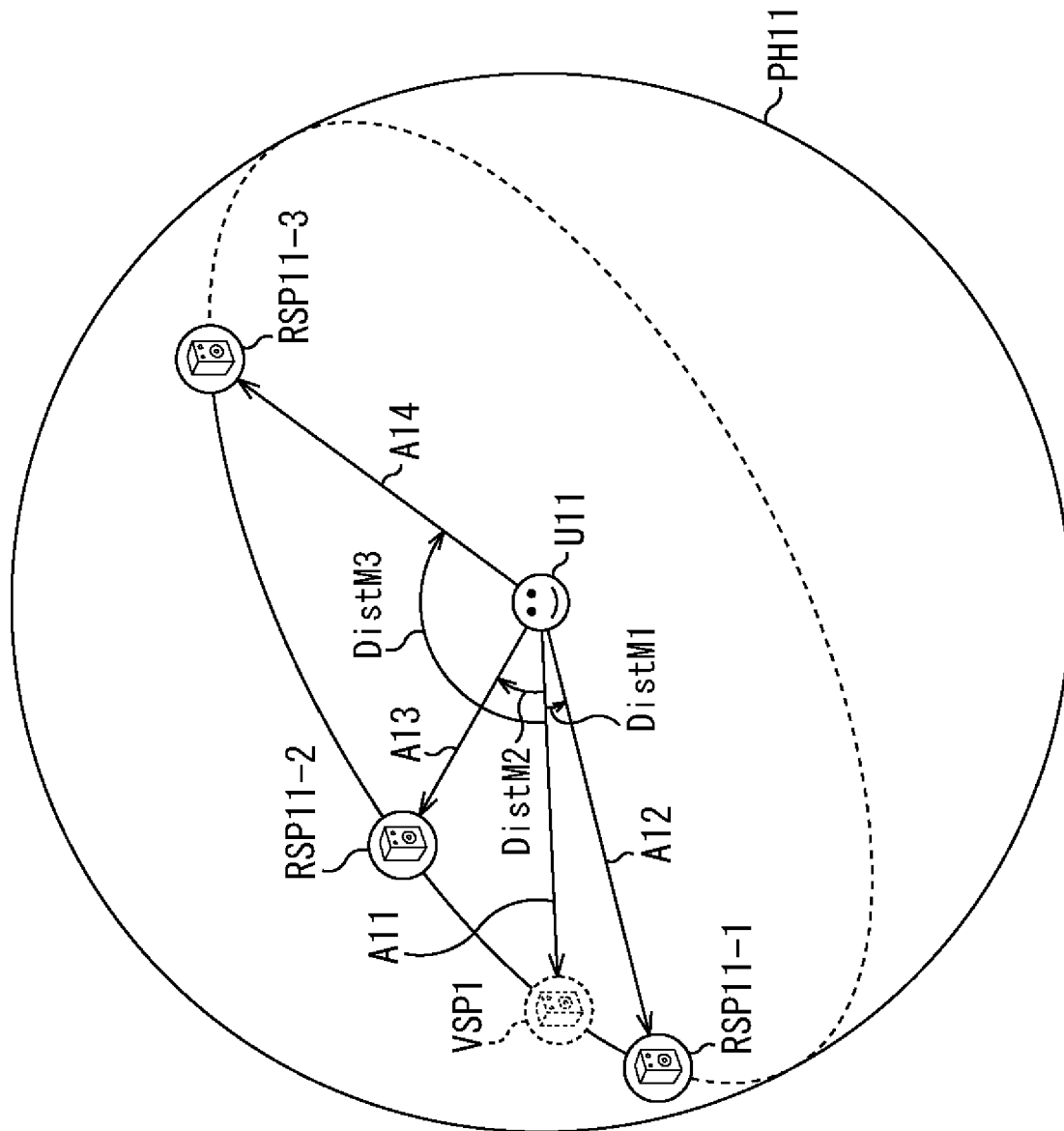
[請求項19] オーディオ信号を再生する理想的なスピーカの位置と、前記オーディオ信号を再生する実際のスピーカの位置との距離に応じて前記オーディオ信号のゲインを補正するための補正情報、および符号化された前記オーディオ信号をビットストリームから抽出し、

前記符号化された前記オーディオ信号を復号し、

復号された前記オーディオ信号および前記補正情報を入力するステップを含む復号方法。

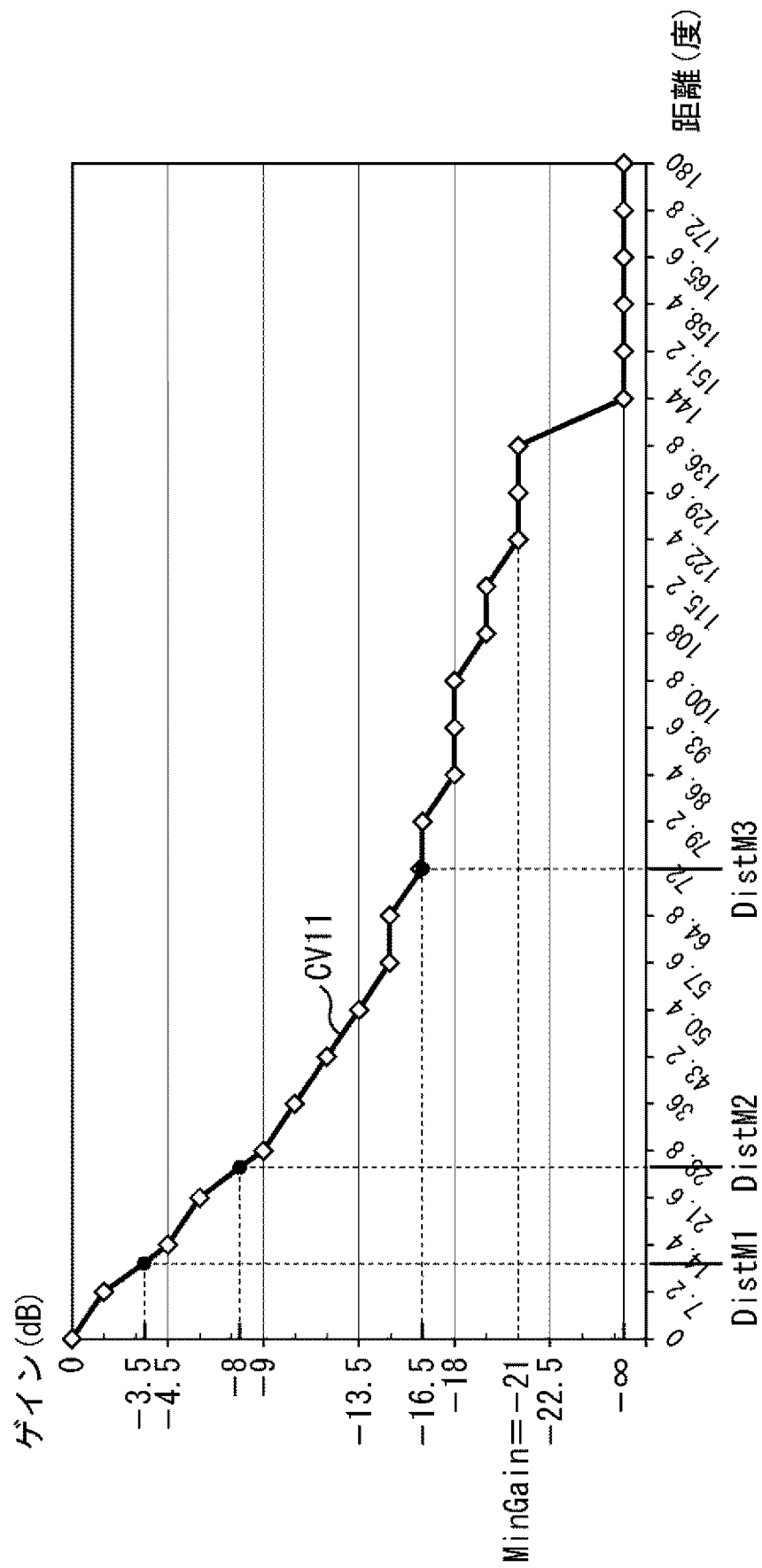
[図1]

図1



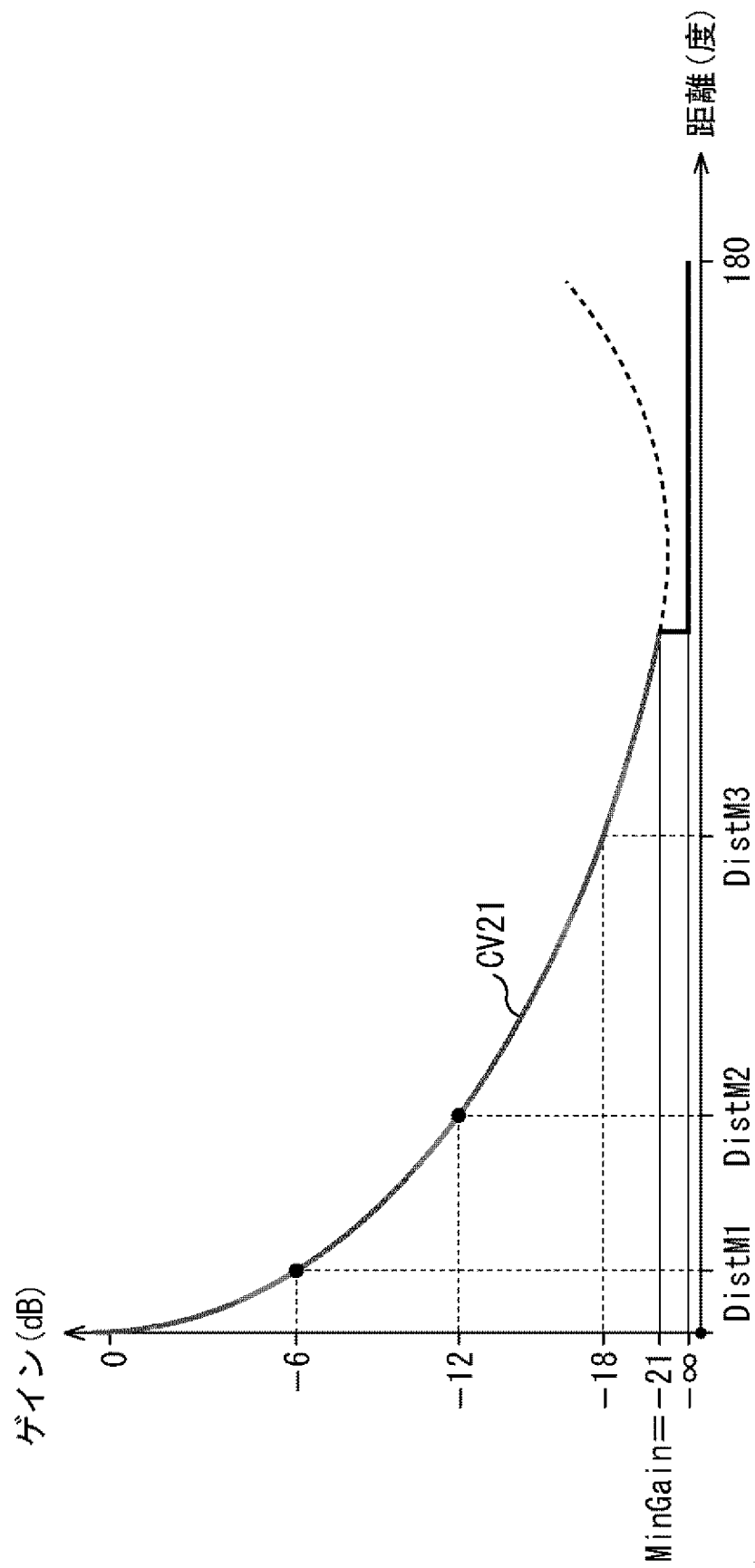
[図2]

図2



[図3]

図3

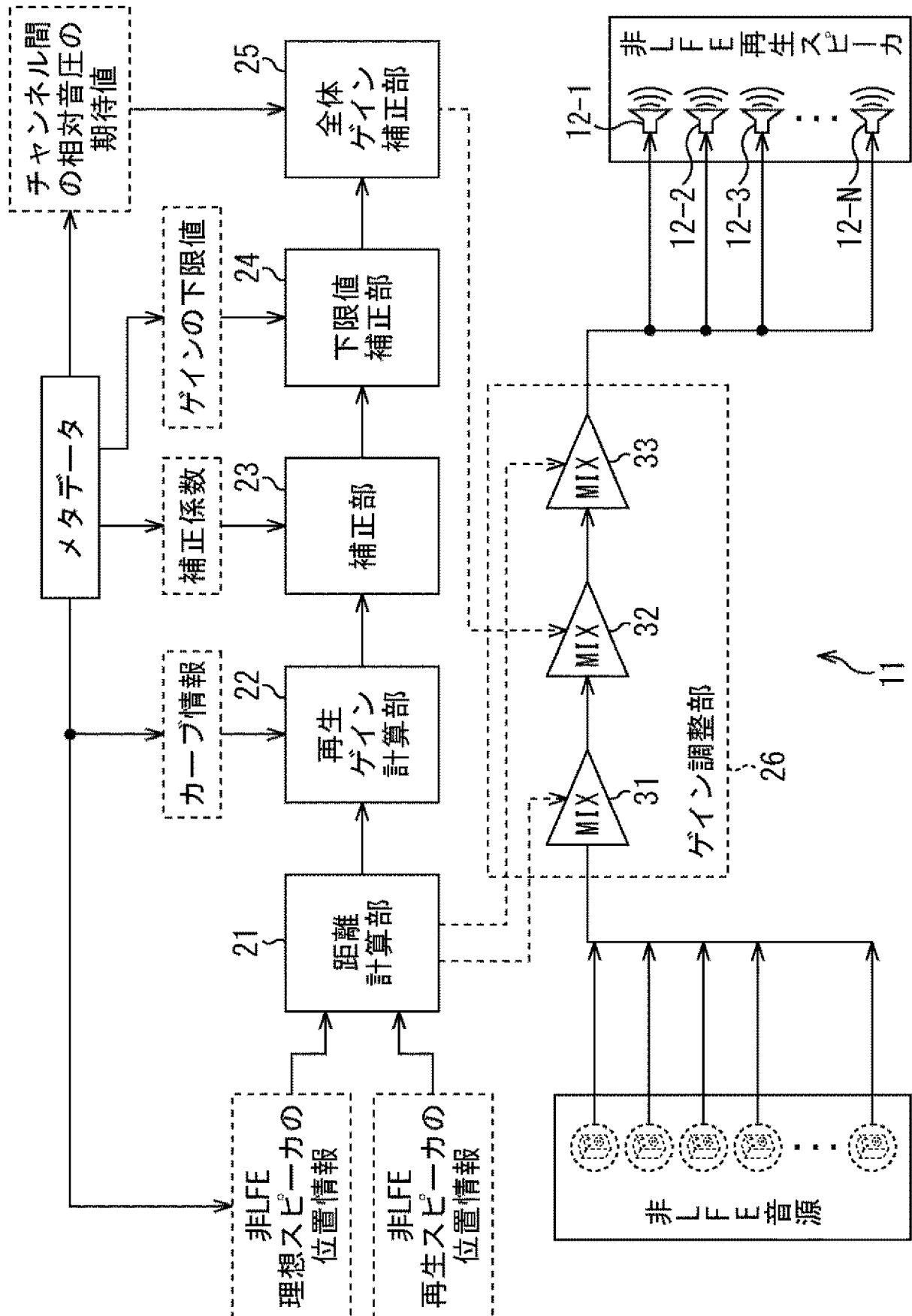


[図4]
図4

LFE理想スピーカの数	LFE再生スピーカの数	ゲインのマッピング関数 MixGain _{io}
1	1	MixGain ₁₁ =MixGain ₂₂ =0dB
2	2	MixGain ₁₂ =MixGain ₂₁ =-∞dB
1	2	MixGain ₁₁ =MixGain ₁₂ =-3dB
2	1	MixGain ₁₁ =MixGain ₂₁ =-3dB

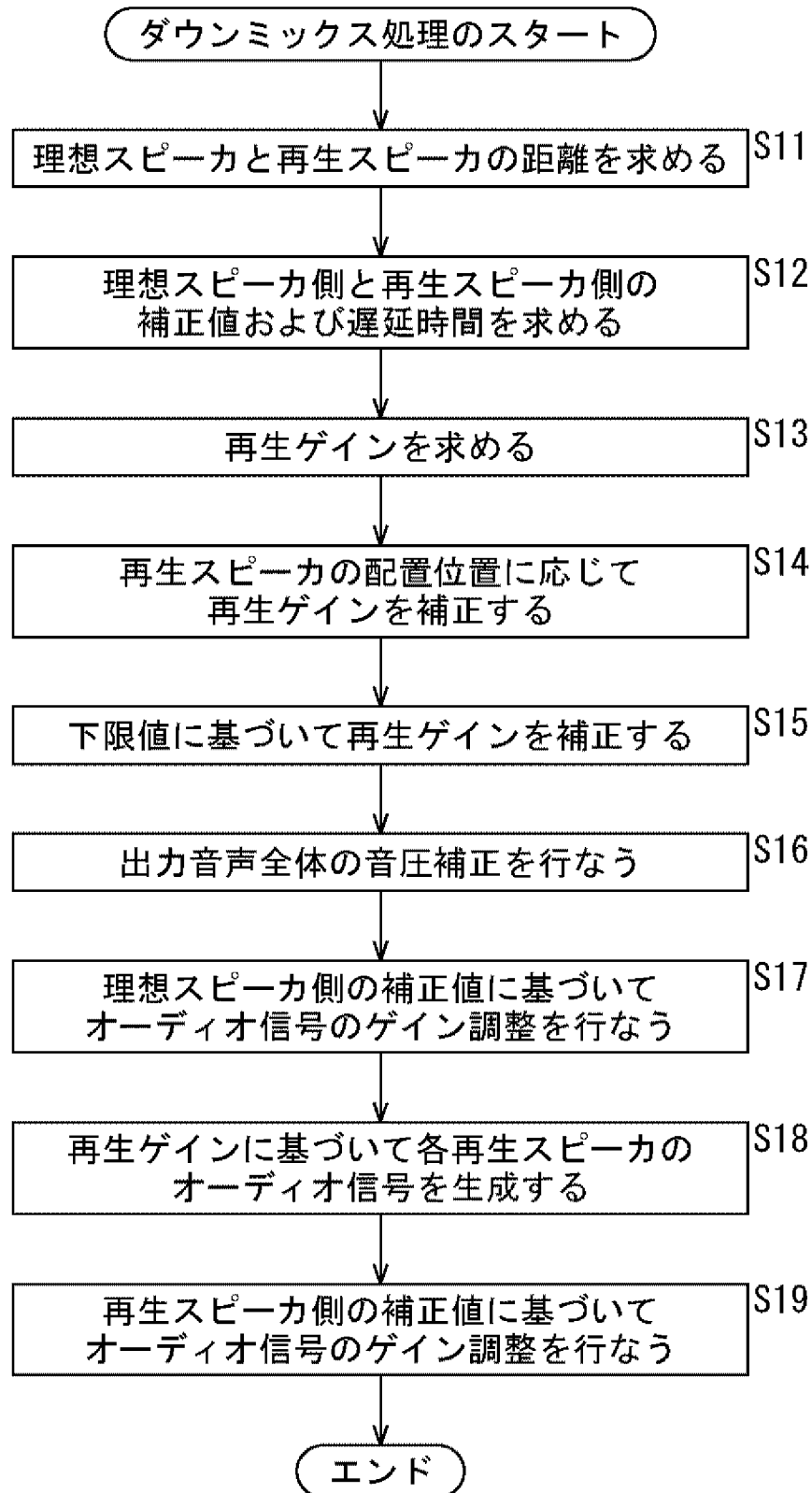
[図5]

図5



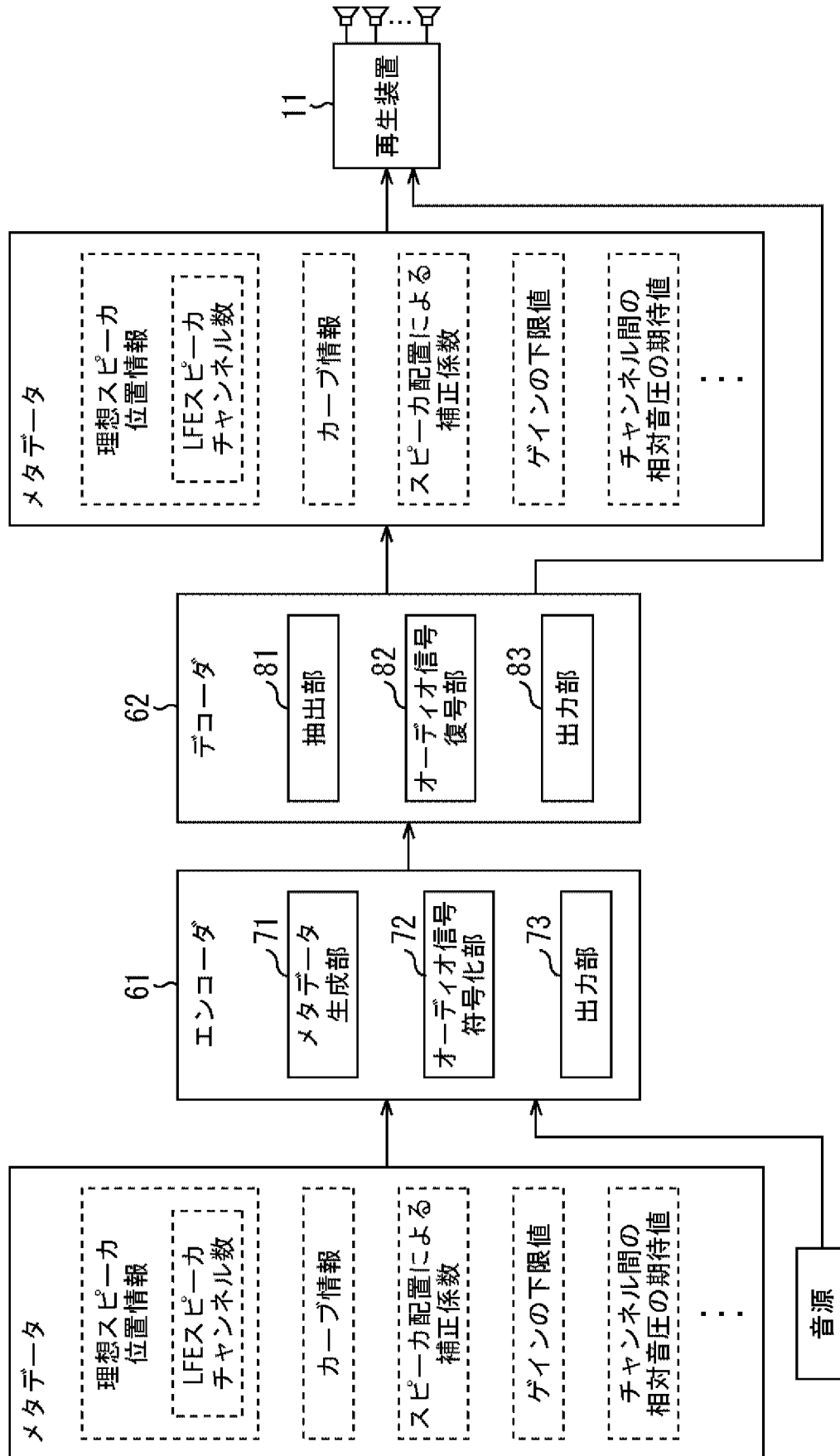
[図6]

図6



[図7]

図7

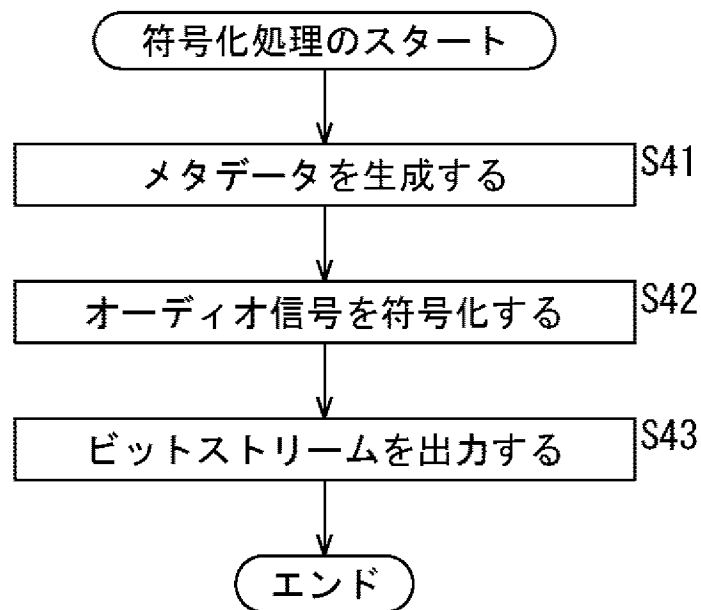


[図8]
8

Syntax	No. of bits
<pre>scene_description_header() { if(down mix coef exist flag == "1") { if(down mix coef mode) { if(polyline curve idx == "111") { i=0; j=0; while(i<16 j<64) { if(polyline curve coefficient[j] == "1") i++; j++; } } } else { if (function curve idx == "111") { for (i=0; i<3; i++) { function_curve_coefficient[i]; } } } minimum_gain_threshold_idx gain_correction_coefficient if(sound_level_exist_flag == "1") { for (i=0; i<number of channels; i++) { channel sound level[i]; } } } }</pre>	1 1 3 1 3 8 4 7 1 10

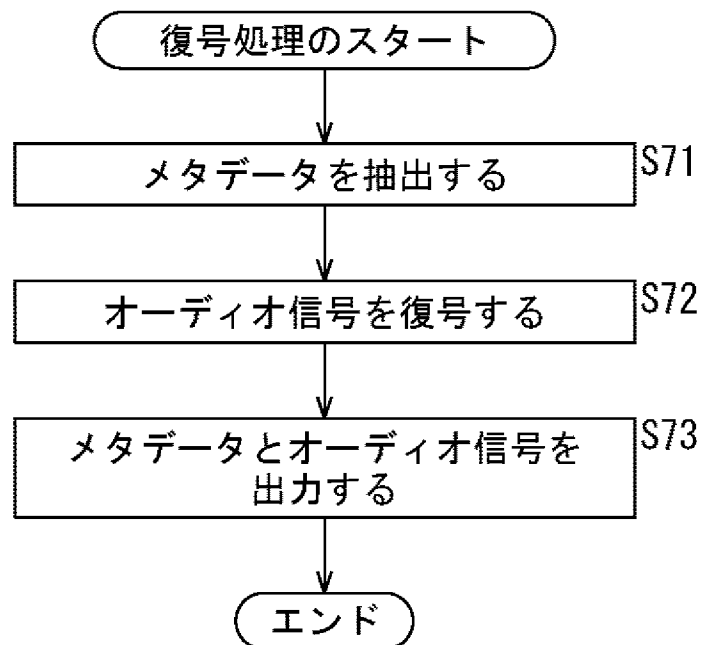
[図9]

図9



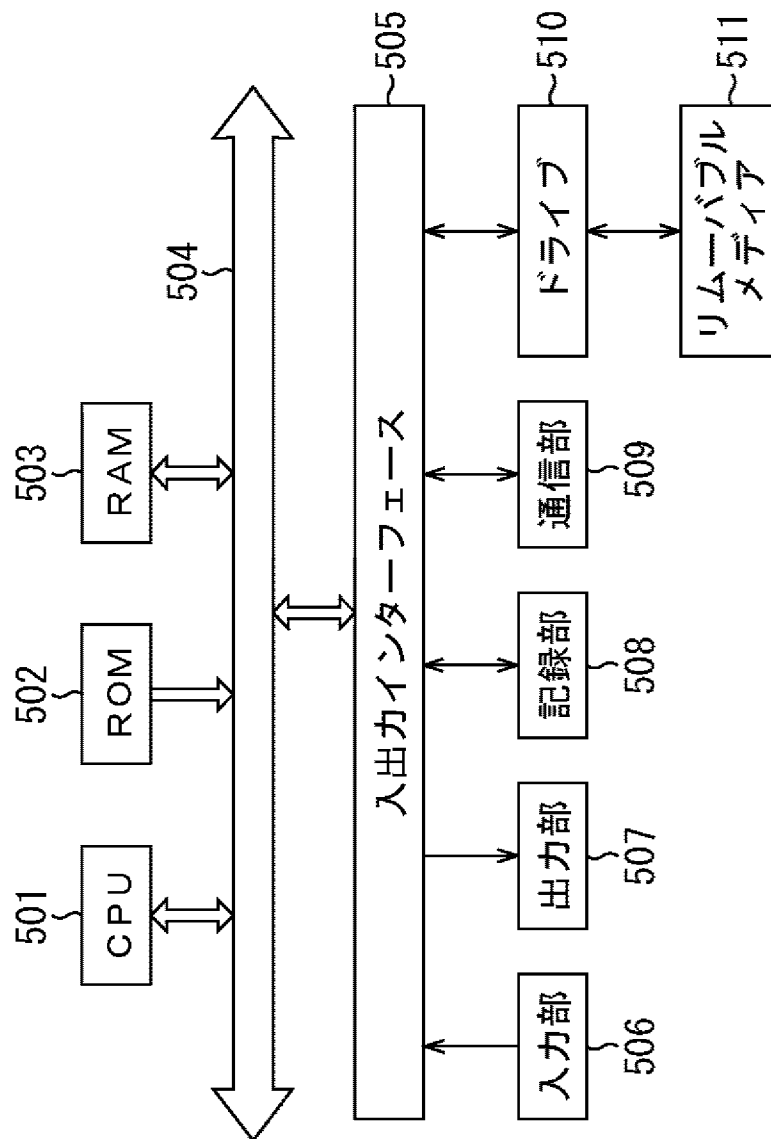
[図10]

図10



[図11]

図11



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2014/063410

A. CLASSIFICATION OF SUBJECT MATTER

H04S5/02(2006.01)i, G10L21/034(2013.01)i

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)

H04S5/02, G10L21/034

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

Jitsuyo Shinan Koho	1922-1996	Jitsuyo Shinan Toroku Koho	1996-2014
Kokai Jitsuyo Shinan Koho	1971-2014	Toroku Jitsuyo Shinan Koho	1994-2014

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 A	JP 2006-101248 A (Victor Company of Japan, Ltd.), 13 April 2006 (13.04.2006), paragraphs [0008] to [0021]; fig. 1 to 6 (Family: none)	1-3, 11-19 4-10
Y A	JP 2010-41190 A (Yamaha Corp.), 18 February 2010 (18.02.2010), paragraphs [0019] to [0027]; fig. 1 (Family: none)	1-3, 11-19 4-10
Y A	JP 8-65169 A (Sony Corp.), 08 March 1996 (08.03.1996), paragraphs [0035] to [0040], [0063] to [0066]; fig. 1, 6 & US 5859826 A & EP 688113 A2	13-19 1-12

☒ 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
05 June, 2014 (05.06.14)

Date of mailing of the international search report
17 June, 2014 (17.06.14)

Name and mailing address of the ISA/
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2014/063410

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2007/0253561 A1 (TSP Systems, Inc.), 01 November 2007 (01.11.2007), paragraphs [0055] to [0063]; fig. 3A & WO 2007/127077 A2	1-19

A. 発明の属する分野の分類（国際特許分類（I P C））

Int.Cl. H04S5/02(2006.01)i, G10L21/034(2013.01)i

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（I P C））

Int.Cl. H04S5/02, G10L21/034

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	1 9 2 2 - 1 9 9 6 年
日本国公開実用新案公報	1 9 7 1 - 2 0 1 4 年
日本国実用新案登録公報	1 9 9 6 - 2 0 1 4 年
日本国登録実用新案公報	1 9 9 4 - 2 0 1 4 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 A	JP 2006-101248 A（日本ビクター株式会社）2006.04.13, 段落[0008]-[0021], [図 1]-[図 6]（ファミリーなし）	1-3, 11-19 4-10
Y A	JP 2010-41190 A（ヤマハ株式会社）2010.02.18, 段落[0019]-[0027], [図 1]（ファミリーなし）	1-3, 11-19 4-10
Y A	JP 8-65169 A（ソニー株式会社）1996.03.08, 段落[0035]-[0040], [0063]-[0066], [図 1], [図 6] & US 5859826 A & EP 688113 A2	13-19 1-12

☒ 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矛盾するものではなく、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

0 5 . 0 6 . 2 0 1 4

国際調査報告の発送日

1 7 . 0 6 . 2 0 1 4

国際調査機関の名称及びあて先

日本国特許庁（I S A / J P）
郵便番号 1 0 0 - 8 9 1 5
東京都千代田区霞が関三丁目 4 番 3 号

特許庁審査官（権限のある職員）

菊池 充

5 Z

4 5 4 5

電話番号 0 3 - 3 5 8 1 - 1 1 0 1 内線 3 5 8 9

C (続き) . 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
A	US 2007/0253561 A1 (TSP Systems, Inc.) 2007.11.01, 段落[0055]-[0063], [FIG. 3A] & WO 2007/127077 A2	1-19