

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第6798312号
(P6798312)

(45) 発行日 令和2年12月9日(2020.12.9)

(24) 登録日 令和2年11月24日(2020.11.24)

(51) Int.Cl. F I
G 1 0 L 19/008 (2013.01) G 1 0 L 19/008 2 0 0
G 1 0 L 19/00 (2013.01) G 1 0 L 19/00 3 3 0 B

請求項の数 12 (全 24 頁)

(21) 出願番号	特願2016-547351 (P2016-547351)	(73) 特許権者	000002185 ソニー株式会社 東京都港区港南1丁目7番1号
(86) (22) 出願日	平成27年8月27日(2015.8.27)	(74) 代理人	100121131 弁理士 西川 孝
(86) 国際出願番号	PCT/JP2015/074132	(74) 代理人	100082131 弁理士 稲本 義雄
(87) 国際公開番号	W02016/039150	(72) 発明者	知念 徹 東京都港区港南1丁目7番1号 ソニー株式会社内
(87) 国際公開日	平成28年3月17日(2016.3.17)	(72) 発明者	島中 光行 東京都港区港南1丁目7番1号 ソニー株式会社内
審査請求日	平成30年8月14日(2018.8.14)		
(31) 優先権主張番号	特願2014-182548 (P2014-182548)		
(32) 優先日	平成26年9月8日(2014.9.8)		
(33) 優先権主張国・地域又は機関	日本国(JP)		

最終頁に続く

(54) 【発明の名称】 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

(57) 【特許請求の範囲】

【請求項1】

ダイアログ音源を含む複数のチャンネル音源のオーディオ信号を符号化し、符号化データを生成するチャンネル音源符号化部と、

前記ダイアログ音源とは異なる追加ダイアログ音源のオーディオ信号を符号化し、符号化データを生成する追加ダイアログ音源符号化部と、

複数の前記追加ダイアログ音源ごとに、それらの複数の前記追加ダイアログ音源のオーディオ信号をゲイン補正するための補正ゲイン情報と、マルチチャンネル構成である前記複数のチャンネル音源のオーディオ信号のチャンネルのうちの前記ダイアログ音源のチャンネルを示す情報とを含むダイアログ情報、前記複数のチャンネル音源の符号化データ、および前記追加ダイアログ音源の符号化データを多重化する多重化部と

を備え、

前記追加ダイアログ音源の符号化データおよび前記ダイアログ情報は、前記多重化により得られるビットストリームのユーザデータ領域に格納される

符号化装置。

【請求項2】

前記ユーザデータ領域は、MPEG AACまたはMPEG-D USACで規定されたDSEである

請求項1に記載の符号化装置。

【請求項3】

前記ダイアログ情報には、前記追加ダイアログ音源の種別を示す情報、および前記ダイ

アログ音源の種別を示す情報が含まれている

請求項 1 または請求項 2 に記載の符号化装置。

【請求項 4】

前記ダイアログ情報には、前記追加ダイアログ音源のオーディオ信号をゲイン補正するためのゲイン情報が含まれている

請求項 1 乃至請求項 3 の何れか一項に記載の符号化装置。

【請求項 5】

ダイアログ音源を含む複数のチャンネル音源のオーディオ信号を符号化して符号化データを生成し、

前記ダイアログ音源とは異なる追加ダイアログ音源のオーディオ信号を符号化して符号化データを生成し、

複数の前記追加ダイアログ音源ごとに、それらの複数の前記追加ダイアログ音源のオーディオ信号をゲイン補正するための補正ゲイン情報と、マルチチャンネル構成である前記複数のチャンネル音源のオーディオ信号のチャンネルのうちの前記ダイアログ音源のチャンネルを示す情報とを含むダイアログ情報、前記複数のチャンネル音源の符号化データ、および前記追加ダイアログ音源の符号化データを多重化する

ステップを含み、

前記追加ダイアログ音源の符号化データおよび前記ダイアログ情報は、前記多重化により得られるビットストリームのユーザデータ領域に格納される

符号化方法。

【請求項 6】

ダイアログ音源を含む複数のチャンネル音源のオーディオ信号を符号化して符号化データを生成し、

前記ダイアログ音源とは異なる追加ダイアログ音源のオーディオ信号を符号化して符号化データを生成し、

複数の前記追加ダイアログ音源ごとに、それらの複数の前記追加ダイアログ音源のオーディオ信号をゲイン補正するための補正ゲイン情報と、マルチチャンネル構成である前記複数のチャンネル音源のオーディオ信号のチャンネルのうちの前記ダイアログ音源のチャンネルを示す情報とを含むダイアログ情報、前記複数のチャンネル音源の符号化データ、および前記追加ダイアログ音源の符号化データを多重化する

ステップを含む処理をコンピュータに実行させ、

前記追加ダイアログ音源の符号化データおよび前記ダイアログ情報は、前記多重化により得られるビットストリームのユーザデータ領域に格納される

プログラム。

【請求項 7】

符号化ビットストリームを、ダイアログ音源を含む複数のチャンネル音源の符号化データ、前記符号化ビットストリームのユーザデータ領域に格納された、前記ダイアログ音源とは異なる 1 または複数の追加ダイアログ音源の符号化データ、および複数の前記追加ダイアログ音源ごとに、それらの複数の前記追加ダイアログ音源のオーディオ信号をゲイン補正するための補正ゲイン情報と、マルチチャンネル構成である前記複数のチャンネル音源のオーディオ信号のチャンネルのうちの前記ダイアログ音源のチャンネルを示す情報とを含む、前記ユーザデータ領域に格納されたダイアログ情報に分離させる分離部と、

前記 1 または複数の前記追加ダイアログ音源の符号化データのうちの選択された前記追加ダイアログ音源の符号化データを復号する追加ダイアログ音源復号部と、

前記補正ゲイン情報に基づいて、前記選択された前記追加ダイアログ音源の符号化データの復号により得られたオーディオ信号のゲイン補正を行うゲイン補正部と

を備える復号装置。

【請求項 8】

前記複数のチャンネル音源の符号化データを復号するチャンネル音源復号部をさらに備える請求項 7 に記載の復号装置。

10

20

30

40

50

【請求項 9】

前記ダイアログ音源の符号化データを復号して得られたオーディオ信号を、前記補正ゲイン情報に基づいてゲイン補正された、前記選択された前記追加ダイアログ音源のオーディオ信号に置き換えて出力するダイアログ選択部をさらに備える

請求項 7 または請求項 8 に記載の復号装置。

【請求項 10】

前記追加ダイアログ音源復号部は、前記ダイアログ情報の提示に応じて選択された前記追加ダイアログ音源の符号化データを復号する

請求項 7 乃至請求項 9 の何れか一項に記載の復号装置。

【請求項 11】

符号化ビットストリームを、ダイアログ音源を含む複数のチャンネル音源の符号化データ、前記符号化ビットストリームのユーザデータ領域に格納された、前記ダイアログ音源とは異なる 1 または複数の追加ダイアログ音源の符号化データ、および複数の前記追加ダイアログ音源ごとに、それらの複数の前記追加ダイアログ音源のオーディオ信号をゲイン補正するための補正ゲイン情報と、マルチチャンネル構成である前記複数のチャンネル音源のオーディオ信号のチャンネルのうちの前記ダイアログ音源のチャンネルを示す情報とを含む、前記ユーザデータ領域に格納されたダイアログ情報に分離させ、

前記 1 または複数の前記追加ダイアログ音源の符号化データのうちの選択された前記追加ダイアログ音源の符号化データを復号し、

前記補正ゲイン情報に基づいて、前記選択された前記追加ダイアログ音源の符号化データの復号により得られたオーディオ信号のゲイン補正を行う

ステップを含む復号方法。

【請求項 12】

符号化ビットストリームを、ダイアログ音源を含む複数のチャンネル音源の符号化データ、前記符号化ビットストリームのユーザデータ領域に格納された、前記ダイアログ音源とは異なる 1 または複数の追加ダイアログ音源の符号化データ、および複数の前記追加ダイアログ音源ごとに、それらの複数の前記追加ダイアログ音源のオーディオ信号をゲイン補正するための補正ゲイン情報と、マルチチャンネル構成である前記複数のチャンネル音源のオーディオ信号のチャンネルのうちの前記ダイアログ音源のチャンネルを示す情報とを含む、前記ユーザデータ領域に格納されたダイアログ情報に分離させ、

前記 1 または複数の前記追加ダイアログ音源の符号化データのうちの選択された前記追加ダイアログ音源の符号化データを復号し、

前記補正ゲイン情報に基づいて、前記選択された前記追加ダイアログ音源の符号化データの復号により得られたオーディオ信号のゲイン補正を行う

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。

【発明の詳細な説明】

【技術分野】

【0001】

本技術は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、オーディオ信号の復号の計算量を低減させることができるようにした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】

例えばオーディオ信号を符号化する方法として、国際標準規格であるMPEG(Moving Picture Experts Group)-2 AAC(Advanced Audio Coding)規格、MPEG-4 AAC規格やMPEG-D USAC規格のマルチチャンネル符号化が知られている(例えば、非特許文献1および非特許文献2参照)。

【先行技術文献】

【非特許文献】

【0003】

10

20

30

40

50

【非特許文献1】INTERNATIONAL STANDARD ISO/IEC 14496-3 Fourth edition 2009-09-01
Information technology-coding of audio-visual objects-part3:Audio

【非特許文献2】INTERNATIONAL STANDARD ISO/IEC 23003-3 First edition 2012-04-01
Information technology-coding of audio-visual objects-part3:Unified speech and audio coding

【発明の概要】

【発明が解決しようとする課題】

【0004】

ところで、従来の5.1チャンネルサラウンド再生を超える、より高臨場感な再生や、複数の音声（ダイアログ）を伝送するためには、より多くのオーディオチャンネルを用いた符号化技術が必要になる。

10

【0005】

例えば、22チャンネルのオーディオ信号と、英語、フランス語、ドイツ語などの複数のダイアログ音源のオーディオ信号とを符号化して伝送し、復号を行う場合を考える。

【0006】

このような場合、符号化で得られる符号化データは、22チャンネルの音源の符号化データのビット列と、複数のダイアログ音源の符号化データのビット列とから成り、それらのビット列が結合されて1つの符号化データのビット列として伝送される。

【0007】

また、復号側では1つの符号化データのビット列から22チャンネルの音源の符号化データの復号と、複数のダイアログ音源の符号化データの復号が行われる。これにより、22チャンネルの音源の復号データと、複数のダイアログ音源の復号データとが得られる。

20

【0008】

そして、復号側では、複数のダイアログ音源の復号データから所望のダイアログ音源の復号データのみが選択され、22チャンネルの音源におけるダイアログチャンネルの音源の復号データが、選択された所望のダイアログ音源の復号データに置き換えられる。

【0009】

一方、符号化データは1つのビット列として伝送されるため、復号側では、所望のダイアログ音源のみが最終的には必要であるにもかかわらず、複数あるダイアログ音源の符号化データの全てを復号することが必要となり、最終的に再生されないダイアログ音源の復号のための処理が無駄になってしまう。

30

【0010】

現状のAACやUSACなどのオーディオコーデックでは、全チャンネルおよび全ダイアログを復号する必要があるため、復号の計算量を低減させ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、復号の計算量を低減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】

【0012】

本技術の第1の側面の符号化装置は、ダイアログ音源を含む複数のチャンネル音源のオーディオ信号を符号化し、符号化データを生成するチャンネル音源符号化部と、前記ダイアログ音源とは異なる追加ダイアログ音源のオーディオ信号を符号化し、符号化データを生成する追加ダイアログ音源符号化部と、複数の前記追加ダイアログ音源ごとに、それらの複数の前記追加ダイアログ音源のオーディオ信号をゲイン補正するための補正ゲイン情報と、マルチチャンネル構成である前記複数のチャンネル音源のオーディオ信号のチャンネルのうちの前記ダイアログ音源のチャンネルを示す情報とを含むダイアログ情報、前記複数のチャンネル音源の符号化データ、および前記追加ダイアログ音源の符号化データを多重化する多重化部とを備え、前記追加ダイアログ音源の符号化データおよび前記ダイアログ情報は、前記多重化により得られるビットストリームのユーザデータ領域に格納される。

40

【0014】

50

前記ユーザデータ領域を、MPEG AACまたはMPEG-D USACで規定されたDSEとすることができる。

【0016】

前記ダイアログ情報には、前記追加ダイアログ音源の種別を示す情報、および前記ダイアログ音源の種別を示す情報が含まれているようにすることができる。

【0018】

前記ダイアログ情報には、前記追加ダイアログ音源のオーディオ信号をゲイン補正するためのゲイン情報が含まれているようにすることができる。

【0020】

本技術の第1の側面の符号化方法またはプログラムは、ダイアログ音源を含む複数のチャンネル音源のオーディオ信号を符号化して符号化データを生成し、前記ダイアログ音源とは異なる追加ダイアログ音源のオーディオ信号を符号化して符号化データを生成し、複数の前記追加ダイアログ音源ごとに、それらの複数の前記追加ダイアログ音源のオーディオ信号をゲイン補正するための補正ゲイン情報と、マルチチャンネル構成である前記複数のチャンネル音源のオーディオ信号のチャンネルのうちの前記ダイアログ音源のチャンネルを示す情報とを含むダイアログ情報、前記複数のチャンネル音源の符号化データ、および前記追加ダイアログ音源の符号化データを多重化するステップを含み、前記追加ダイアログ音源の符号化データおよび前記ダイアログ情報は、前記多重化により得られるビットストリームのユーザデータ領域に格納される。

10

【0021】

本技術の第1の側面においては、ダイアログ音源を含む複数のチャンネル音源のオーディオ信号が符号化されて符号化データが生成され、前記ダイアログ音源とは異なる追加ダイアログ音源のオーディオ信号が符号化されて符号化データが生成され、複数の前記追加ダイアログ音源ごとに、それらの複数の前記追加ダイアログ音源のオーディオ信号をゲイン補正するための補正ゲイン情報と、マルチチャンネル構成である前記複数のチャンネル音源のオーディオ信号のチャンネルのうちの前記ダイアログ音源のチャンネルを示す情報とを含むダイアログ情報、前記複数のチャンネル音源の符号化データ、および前記追加ダイアログ音源の符号化データが多重化される。また、前記追加ダイアログ音源の符号化データおよび前記ダイアログ情報は、前記多重化により得られるビットストリームのユーザデータ領域に格納される。

20

30

【0022】

本技術の第2の側面の復号装置は、符号化ビットストリームを、ダイアログ音源を含む複数のチャンネル音源の符号化データ、前記符号化ビットストリームのユーザデータ領域に格納された、前記ダイアログ音源とは異なる1または複数の追加ダイアログ音源の符号化データ、および複数の前記追加ダイアログ音源ごとに、それらの複数の前記追加ダイアログ音源のオーディオ信号をゲイン補正するための補正ゲイン情報と、マルチチャンネル構成である前記複数のチャンネル音源のオーディオ信号のチャンネルのうちの前記ダイアログ音源のチャンネルを示す情報とを含む、前記ユーザデータ領域に格納されたダイアログ情報に分離させる分離部と、前記1または複数の前記追加ダイアログ音源の符号化データのうちの選択された前記追加ダイアログ音源の符号化データを復号する追加ダイアログ音源復号部と、前記補正ゲイン情報に基づいて、前記選択された前記追加ダイアログ音源の符号化データの復号により得られたオーディオ信号のゲイン補正を行うゲイン補正部とを備える。

40

【0023】

復号装置には、前記複数のチャンネル音源の符号化データを復号するチャンネル音源復号部をさらに設けることができる。

【0024】

復号装置には、前記ダイアログ音源の符号化データを復号して得られたオーディオ信号を、前記補正ゲイン情報に基づいてゲイン補正された、前記選択された前記追加ダイアログ音源のオーディオ信号に置き換えて出力するダイアログ選択部をさらに設けることができる。

50

【0025】

前記追加ダイアログ音源復号部には、前記ダイアログ情報の提示に応じて選択された前記追加ダイアログ音源の符号化データを復号させることができる。

【0026】

本技術の第2の側面の復号方法またはプログラムは、符号化ビットストリームを、ダイアログ音源を含む複数のチャンネル音源の符号化データ、前記符号化ビットストリームのユーザデータ領域に格納された、前記ダイアログ音源とは異なる1または複数の追加ダイアログ音源の符号化データ、および複数の前記追加ダイアログ音源ごとに、それらの複数の前記追加ダイアログ音源のオーディオ信号をゲイン補正するための補正ゲイン情報と、マルチチャンネル構成である前記複数のチャンネル音源のオーディオ信号のチャンネルのうちの前記ダイアログ音源のチャンネルを示す情報とを含む、前記ユーザデータ領域に格納されたダイアログ情報に分離させ、前記1または複数の前記追加ダイアログ音源の符号化データのうちの選択された前記追加ダイアログ音源の符号化データを復号し、前記補正ゲイン情報に基づいて、前記選択された前記追加ダイアログ音源の符号化データの復号により得られたオーディオ信号のゲイン補正を行うステップを含む。

10

【0027】

本技術の第2の側面においては、符号化ビットストリームが、ダイアログ音源を含む複数のチャンネル音源の符号化データ、前記符号化ビットストリームのユーザデータ領域に格納された、前記ダイアログ音源とは異なる1または複数の追加ダイアログ音源の符号化データ、および複数の前記追加ダイアログ音源ごとに、それらの複数の前記追加ダイアログ音源のオーディオ信号をゲイン補正するための補正ゲイン情報と、マルチチャンネル構成である前記複数のチャンネル音源のオーディオ信号のチャンネルのうちの前記ダイアログ音源のチャンネルを示す情報とを含む、前記ユーザデータ領域に格納されたダイアログ情報に分離され、前記1または複数の前記追加ダイアログ音源の符号化データのうちの選択された前記追加ダイアログ音源の符号化データが復号され、前記補正ゲイン情報に基づいて、前記選択された前記追加ダイアログ音源の符号化データの復号により得られたオーディオ信号のゲイン補正が行われる。

20

【発明の効果】

【0028】

本技術の第1の側面および第2の側面によれば、復号の計算量を低減させることができる。

30

【0029】

なお、ここに記載された効果は必ずしも限定されるものではなく、本開示中に記載された何れかの効果であってもよい。

【図面の簡単な説明】

【0030】

【図1】22チャンネル音源対応再生システムについて説明する図である。

【図2】符号化装置の構成例を示す図である。

【図3】ダイアログ情報のビットストリームシンタックスの例を示す図である。

【図4】追加ダイアログ音源の符号化データのビットストリームシンタックスの例を示す図である。

40

【図5】符号化処理を説明するフローチャートである。

【図6】復号装置の構成例を示す図である。

【図7】復号処理を説明するフローチャートである。

【図8】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】

【0031】

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。

【0032】

第1の実施の形態

50

本技術の概要について

本技術は、複数チャンネルの音源の符号化データのビット列と、複数のダイアログ音源の符号化データのビット列とを結合して1つの符号化データのビット列として伝送する場合に、ビット列にダイアログの情報も加えて伝送するものである。

【0033】

これにより、ダイアログチャンネルとして再生されるダイアログ音源のみの復号を行うことができるようになり、復号の計算量を低減させることができる。

【0034】

このとき、復号側では、視聴者に追加ダイアログ音源の数、追加ダイアログ音源の情報を提示し、ダイアログチャンネルの音源を視聴者が選択したダイアログ音源に置き換えるか否かの指示を視聴者から受けるようにすれば、ダイアログチャンネルとして再生されるダイアログ音源のみの復号を行うことができるようになる。

【0035】

具体的には、このような本技術では符号化側において追加ダイアログ音源の数、追加ダイアログ音源の情報、ダイアログチャンネルの位置、および追加ダイアログ音源の利得（ゲイン）がダイアログ情報として符号化される。そして、ダイアログ情報の符号化データが、追加ダイアログ音源の符号化データおよび22チャンネル音源の符号化データと多重化されて、得られたビットストリームが出力される。

【0036】

また、復号側ではダイアログ情報の符号化データ、追加ダイアログ音源の符号化データ、および22チャンネル音源の符号化データが分離される。さらに、視聴者に対して追加ダイアログ音源の数や、追加ダイアログ音源の情報が提示され、視聴者から、ダイアログチャンネルの音源を視聴者が選択したダイアログ音源に置き換えるか否かの指示がなされる。

【0037】

このような視聴者からの指示がなされると、復号側では視聴者の指示に応じて、ダイアログチャンネルとして再生されるダイアログ音源のみの復号が行われてダイアログ音源の置き換えが行われ、22チャンネル音源の再生が行われる。

【0038】

22チャンネル音源対応再生システムについて

次に、本技術を適用したより具体的な実施の形態について説明する。なお、ここでは再生対象のオーディオ信号がマルチチャンネル構成である22チャンネルの信号である場合を例として説明を行うが、再生対象のオーディオ信号は何チャンネルの信号であってもよい。

【0039】

図1は、復号により得られたマルチチャンネル構成である22チャンネルのオーディオ信号を再生可能な、22チャンネル音源対応再生システムを示す図である。この例ではコンテンツの映像と、その映像に付随する22チャンネルのオーディオ信号とが再生される。

【0040】

図1では、視聴者US11の正面に映像が表示されるスクリーンSC11が設けられており、正面には9つのスピーカSP1乃至スピーカSP9が配置されている。

【0041】

すなわち、視聴者US11から見て正面中層には、中央にスピーカSP1が配置されており、そのスピーカSP1の左右にスピーカSP2およびスピーカSP3が配置されている。また、中層の図中、上側に位置する上層には、中央にスピーカSP4が配置されており、そのスピーカSP4の左右にスピーカSP5およびスピーカSP6が配置されている。

【0042】

同様に中層の図中、下側に位置する下層には、中央にスピーカSP7が配置されており、そのスピーカSP7の左右にスピーカSP8およびスピーカSP9が配置されている。また、視聴者US11の図中、上側、つまり天井にはスピーカSP10が配置されている。

【0043】

さらに視聴者US11から見て図中、左側に位置する左側面には6つのスピーカSP11乃至ス

10

20

30

40

50

ピーカSP16が配置されている。すなわち、左側面中層にはスピーカSP11およびスピーカSP12が配置されており、左側面上層にはスピーカSP13およびスピーカSP14が配置されており、左側面下層にはスピーカSP15およびスピーカSP16が配置されている。

【0044】

同様に視聴者US11から見て図中、右側に位置する右側面には6つのスピーカSP17乃至スピーカSP22が配置されている。すなわち、右側面中層にはスピーカSP17およびスピーカSP18が配置されており、右側面上層にはスピーカSP19およびスピーカSP20が配置されており、右側面下層にはスピーカSP21およびスピーカSP22が配置されている。

【0045】

このような22個のスピーカSP1乃至スピーカSP22と、図示せぬ復号装置とを有する22チャンネル音源対応再生システムでは、チャンネル音源の符号化データ、追加ダイアログ音源の符号化データ、およびダイアログ情報の符号化データが入力される。

10

【0046】

ここで、チャンネル音源の符号化データ、追加ダイアログ音源の符号化データ、およびダイアログ情報の符号化データについて説明する。

【0047】

チャンネル音源の符号化データは、22個のスピーカSP1乃至スピーカSP22のそれぞれに対応する、マルチチャンネルを構成する各チャンネルのオーディオ信号を符号化して得られたデータである。以下では、各チャンネルのオーディオ信号により再生される音声をチャンネル音源とも称することとする。

20

【0048】

また、各チャンネルに対応するチャンネル音源のなかには、ダイアログ音声であるダイアログ音源も含まれている。ここでは、チャンネル音源に含まれているダイアログ音源は1つである例について説明するが、ダイアログ音源の数はいくつであってもよい。さらに、以下では、ダイアログ音源を再生するためのオーディオ信号のチャンネルを、ダイアログチャンネルとも称することとする。

【0049】

追加ダイアログ音源の符号化データは、ダイアログ音源の置き換え候補となる追加されたダイアログ音声である追加ダイアログ音源のオーディオ信号を符号化して得られたデータである。22チャンネル音源対応再生システムでは、チャンネル音源の再生時に、チャンネル音源に含まれているダイアログ音源を、追加ダイアログ音源に置き換えて再生することが可能となっている。

30

【0050】

ダイアログ情報の符号化データは、ダイアログ音源や追加ダイアログ音源に関する情報が含まれているダイアログ情報を符号化して得られたデータである。

【0051】

22チャンネル音源対応再生システムでは、上述したチャンネル音源の符号化データ、追加ダイアログ音源の符号化データ、およびダイアログ情報の符号化データが含まれているビットストリームが供給されると、それらの符号化データが適宜復号され、チャンネル音源が再生される。

40

【0052】

例えば、まず、復号により得られたダイアログ情報が視聴者US11に対して提示される。すると、視聴者US11は、適宜、ダイアログ音源の置き換えを指示する。すなわち、ダイアログ音源と、1または複数の追加ダイアログ音源とのなかから、再生すべきダイアログの音源が選択される。

【0053】

ここで、例えば視聴者US11によりダイアログ音源が選択された場合には、ダイアログ音源を含むチャンネル音源の符号化データのみが復号される。そして、復号により得られた各チャンネル音源のオーディオ信号、つまり各チャンネルのオーディオ信号が、それらのチャンネルに対応するスピーカに供給され、チャンネル音源が再生される。

50

【 0 0 5 4 】

このとき、スピーカSP1乃至スピーカSP22のうちの1つのスピーカがダイアログ音源を再生するためのスピーカとして用いられる。この例では、視聴者US11の正面中層の中央（センター）位置にあるスピーカSP1において、ダイアログ音源が再生される。

【 0 0 5 5 】

なお、符号化側では、ダイアログ音源を再生するスピーカとして任意のスピーカを指定することが可能であるが、ここではダイアログ音源を再生するスピーカがスピーカSP1である例について説明することとする。

【 0 0 5 6 】

一方、再生するダイアログの音源として追加ダイアログ音源が選択された場合、チャンネル音源の符号化データが復号されるとともに、選択された追加ダイアログ音源の符号化データも復号される。そして、復号により得られたチャンネル音源のオーディオ信号のうち、ダイアログ音源のオーディオ信号が、復号により得られた、選択された追加ダイアログ音源のオーディオ信号に置き換えられてチャンネル音源が再生される。

10

【 0 0 5 7 】

したがって、この場合、スピーカSP1では選択された追加ダイアログ音源が再生され、残りのスピーカSP2乃至スピーカSP22では、それらのスピーカに対応するチャンネル音源が再生されることになる。また、この場合、選択されなかった他の追加ダイアログ音源の符号化データは復号されない。

【 0 0 5 8 】

なお、ここではチャンネル音源のうちの置き換えられる音源がダイアログ音源、すなわちダイアログ音声である場合を例として説明するが、その他、オブジェクト音源など、どのような音源が置き換えられるようにしてもよい。

20

【 0 0 5 9 】

符号化装置の構成例

続いて、以上において説明した22チャンネル音源対応再生システムに入力されるビットストリームを生成する符号化装置と、22チャンネル音源対応再生システムを構成し、符号化装置から出力されたビットストリームを復号する復号装置について説明する。

【 0 0 6 0 】

まず、符号化装置について説明する。図2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。

30

【 0 0 6 1 】

図2の符号化装置11は、チャンネル音源符号化部21、追加ダイアログ音源符号化部22、ダイアログ情報符号化部23、および多重化部24を有している。

【 0 0 6 2 】

チャンネル音源符号化部21は、供給された22チャンネル構成の各チャンネル音源のオーディオ信号を符号化し、その結果得られた符号化データを多重化部24に供給する。追加ダイアログ音源符号化部22は、供給された追加ダイアログ音源のオーディオ信号を符号化し、その結果得られた符号化データを多重化部24に供給する。

【 0 0 6 3 】

ダイアログ情報符号化部23は、供給されたダイアログ情報を符号化し、その結果得られた符号化データを多重化部24に供給する。多重化部24は、チャンネル音源符号化部21、追加ダイアログ音源符号化部22、およびダイアログ情報符号化部23のそれぞれから供給された符号化データを多重化してビットストリームを生成し、出力する。

40

【 0 0 6 4 】

符号化装置11においては、例えばAAC（MPEG AAC）規格やUSAC（MPEG-D USAC）規格に従ってチャンネル音源、追加ダイアログ音源、およびダイアログ情報が符号化される。

【 0 0 6 5 】

ダイアログ情報および追加ダイアログ音源について

ところで、多重化部24から出力されるビットストリームは、例えばAAC規格やUSAC規

50

格に準拠した符号化ビット列となっている。このビットストリームでは、例えばチャンネル音源の符号化データに続いてダイアログ情報の符号化データが配置され、さらにその後ろに各追加ダイアログ音源の符号化データが配置されている。

【 0 0 6 6 】

また、ダイアログ情報の符号化データは、例えばAAC規格やUSAC規格で規定された、ビットストリームのDSE (Data Stream Element) と呼ばれるユーザが自由に定義可能なユーザデータ領域に格納されている。同様に各追加ダイアログ音源の符号化データも、ダイアログ情報の符号化データが格納されたDSE (以下、適宜、DSE()) とともに記す) に続く、追加ダイアログ音源の個数分のDSEに格納される。すなわち、1つのDSEに1つの追加ダイアログ音源の符号化データが格納される。

10

【 0 0 6 7 】

ここで、ダイアログ情報の符号化データと追加ダイアログ音源の符号化データのビットストリームシンタックスの具体例について説明する。

【 0 0 6 8 】

図3は、ダイアログ情報の符号化データのビットストリームシンタックスを示す図である。

【 0 0 6 9 】

この例では、ダイアログ情報の符号化データとして、DSEの先頭に、符号化されたダイアログ情報の存在を示す同期用コードであるdialogue_info_syncが配置されている。

【 0 0 7 0 】

20

また、同期用コードdialogue_info_syncに続いて、チャンネル音源に含まれているダイアログ音源に関する情報 (以下、メインダイアログ情報と称する) であるmain_dialogue_infoが配置されている。例えばメインダイアログ情報main_dialogue_infoは日本語の主音声や、ナレーション音声等、ダイアログ音源がどのような種別 (種類) の音源であるかを示す情報などとされる。

【 0 0 7 1 】

dialogue_src_indexは、チャンネル音源に含まれているダイアログ音源のオーディオ信号のチャンネル、つまりダイアログ音源が再生されるスピーカ位置を示すダイアログ位置情報である。したがって、上述したスピーカSP1乃至スピーカSP22のうちのダイアログ位置情報dialogue_src_indexにより特定されるスピーカで、ダイアログ音源または追加ダイアログ音源が再生されることになる。

30

【 0 0 7 2 】

dialogue_gainは、ダイアログ置き換え時における追加ダイアログ音源のゲイン値、すなわち追加ダイアログ音源の再生時のゲイン補正を行うためのゲイン情報を示している。num_additional_dialogueは追加ダイアログ音源の数を示す追加ダイアログ個数情報である。

【 0 0 7 3 】

また、追加ダイアログ個数情報num_additional_dialogueの後には、追加ダイアログ個数情報num_additional_dialogueにより示される数だけadditional_dialogue_info[i]およびadditional_dialogue_gain[i]が配置されている。

40

【 0 0 7 4 】

additional_dialogue_info[i]は、i番目に配置されている追加ダイアログ音源に関する情報 (以下、追加ダイアログ情報と称する) である。

【 0 0 7 5 】

この追加ダイアログ情報additional_dialogue_info[i]は、メインダイアログ情報と同様に、フランス語の副音声や、コンテンツの解説の音声、低速度で朗読されたナレーション音声等、追加ダイアログ音源がどのような種別 (種類) の音源であるかを示す情報などとされる。

【 0 0 7 6 】

additional_dialogue_gain[i]は、i番目に配置されている追加ダイアログ音源のための

50

補正ゲイン値を示している。この補正ゲイン値additional_dialogue_gain[i]は、各追加ダイアログ音源間の音量のばらつきを補正するためのゲイン情報である。

【0077】

以上のようにダイアログ情報には、少なくともメインダイアログ情報、ダイアログ位置情報、追加ダイアログ音源のゲイン値、追加ダイアログ個数情報、追加ダイアログ情報、および追加ダイアログ音源の補正ゲイン値が含まれている。

【0078】

また、各追加ダイアログ音源の符号化データのビットストリームシンタックスは、例えば図4に示すようになる。

【0079】

この例では、追加ダイアログ音源の符号化データとして、DSEの先頭に、符号化された追加ダイアログ音源のオーディオ信号の存在を示す同期用コードであるadditional_dialogue_data_syncが配置されている。

【0080】

また、同期用コードadditional_dialogue_data_syncに続いて配置されているadditional_dialogue_indexは、このDSE内に格納されている追加ダイアログ音源のオーディオ信号（符号化データ）を識別するインデックスを示している。さらに、そのインデックスに続いて配置されているsingle_channel_element()は、AACやUSACで規定されるシングルチャネルの符号化データ、すなわちここでは追加ダイアログ音源の符号化されたオーディオ信号を示している。

【0081】

符号化処理の説明

次に、符号化装置11の動作について説明する。

【0082】

符号化装置11は、符号化対象とするマルチチャンネル構成の各フレームのオーディオ信号が供給されると、符号化処理を開始し、符号化により得られたビットストリームを出力する。以下、図5のフローチャートを参照して、符号化装置11による符号化処理について説明する。

【0083】

ステップS11において、符号化装置11はまだ符号化を行っていない1フレーム分のデータがあるか否かを判定する。

【0084】

ステップS11において1フレーム分の符号化データがあると判定された場合、ステップS12において、チャンネル音源符号化部21は処理対象とする処理対象フレームについて、ダイアログチャンネルを含むチャンネル音源のオーディオ信号の符号化を行う。すなわち、チャンネル音源符号化部21は、供給された22チャンネル構成の各チャンネル音源のオーディオ信号を符号化し、その結果得られた符号化データを多重化部24に供給する。

【0085】

ステップS13において、ダイアログ情報符号化部23は、ダイアログ情報があるか否かを判定する。例えば外部から処理対象フレームのダイアログ情報が供給された場合、ダイアログ情報があると判定される。

【0086】

ステップS13においてダイアログ情報がないと判定された場合、ステップS14乃至ステップS26の処理は行われず、処理はステップS27へと進む。この場合、追加ダイアログ音源がないので、チャンネル音源の符号化データのみがビットストリームに格納される。

【0087】

これに対してステップS13において、ダイアログ情報があると判定された場合、処理はステップS14に進む。

【0088】

10

20

30

40

50

ステップS 1 4において、ダイアログ情報符号化部 2 3 は、供給されたダイアログ情報に基づいて、ダイアログ情報の同期用コードdialogue_info_syncおよびメインダイアログ情報main_dialogue_infoを生成する。

【 0 0 8 9 】

ステップS 1 5において、ダイアログ情報符号化部 2 3 は、ダイアログ情報に基づいて、ダイアログ位置情報dialogue_src_indexを生成する。

【 0 0 9 0 】

ステップS 1 6において、ダイアログ情報符号化部 2 3 は、ダイアログ情報に基づいて、ダイアログ置き換え時における追加ダイアログ音源のゲイン値dialogue_gainを生成する。

【 0 0 9 1 】

ステップS 1 7において、ダイアログ情報符号化部 2 3 は、ダイアログ情報に基づいて、追加ダイアログ音源の数を示す追加ダイアログ個数情報num_additional_dialogueを生成する。

【 0 0 9 2 】

ステップS 1 8において、ダイアログ情報符号化部 2 3 は、ダイアログ情報に基づいて、追加ダイアログ音源のための追加ダイアログ情報additional_dialogue_info[i]および補正ゲイン値additional_dialogue_gain[i]を、追加ダイアログ個数情報num_additional_dialogueにより示される個数分だけ生成する。すなわち $i = 0$ 乃至 $\text{num_additional_dialogue} - 1$ について、追加ダイアログ情報additional_dialogue_info[i]および補正ゲイン値additional_dialogue_gain[i]が生成される。これにより、各追加ダイアログ音源の符号化された追加ダイアログ情報と補正ゲイン値が得られる。なお、追加ダイアログ音源の数は1つであってもよいし、複数であってもよい。

【 0 0 9 3 】

ステップS 1 9において、ダイアログ情報符号化部 2 3 は、ステップS 1 4乃至ステップS 1 8で生成されたデータを1つのDSE()の領域に格納する。

【 0 0 9 4 】

すなわち同期用コードdialogue_info_sync、メインダイアログ情報main_dialogue_info、ダイアログ位置情報dialogue_src_index、ゲイン値dialogue_gain、追加ダイアログ個数情報num_additional_dialogue、追加ダイアログ情報additional_dialogue_info[i]、および補正ゲイン値additional_dialogue_gain[i]がDSEに格納される。これにより、例えば図3に示したビットストリームシンタックスに従ってダイアログ情報の符号化データが生成されたことになる。

【 0 0 9 5 】

ダイアログ情報符号化部 2 3 は、このようにして得られたダイアログ情報の符号化データを多重化部 2 4 に供給する。

【 0 0 9 6 】

ステップS 2 0において、追加ダイアログ音源符号化部 2 2 は、処理対象の追加ダイアログ音源を示すインデックスiの値を $i=0$ とする。ここで、インデックスiにより特定される追加ダイアログ音源は、処理対象のi番目の追加ダイアログ音源であることを示している。ここでは、追加ダイアログ音源の数はnum_additional_dialogueだけあるので、インデックス $i = 0$ 乃至 $\text{num_additional_dialogue} - 1$ について追加ダイアログ音源が処理されればよい。

【 0 0 9 7 】

ステップS 2 1において、追加ダイアログ音源符号化部 2 2 は、インデックスiの値が $i < \text{num_additional_dialogue}$ であるか否かを判定する。

【 0 0 9 8 】

すなわち、 $i < \text{num_additional_dialogue}$ である場合、まだ全ての追加ダイアログ音源が処理されていないことになる。ステップS 2 1において $i < \text{num_additional_dialogue}$ であると判定された場合、処理はステップS 2 2に進む。

10

20

30

40

50

【 0 0 9 9 】

ステップ S 2 2 において、追加ダイアログ音源符号化部 2 2 は、追加ダイアログ音源の同期用コード additional_dialogue_data_sync を生成する。

【 0 1 0 0 】

ステップ S 2 3 において、追加ダイアログ音源符号化部 2 2 は、処理対象の追加ダイアログ音源の識別のためのインデックス additional_dialogue_index の値をインデックス i として生成する。すなわち、インデックス i の値が、そのままインデックス additional_dialogue_index の値とされる。

【 0 1 0 1 】

ステップ S 2 4 において、追加ダイアログ音源符号化部 2 2 は、 i 番目の追加ダイアログ音源の符号化データを、AAC や USAC で規定される single_channel_element() のビットストリームシンタックスに従い生成する。すなわち、追加ダイアログ音源符号化部 2 2 は、外部から供給された i 番目の追加ダイアログ音源のオーディオ信号を符号化する。

10

【 0 1 0 2 】

ステップ S 2 5 において、追加ダイアログ音源符号化部 2 2 は、ステップ S 2 2 乃至ステップ S 2 4 の処理で生成されたデータを 1 つの DSE() の領域に格納する。

【 0 1 0 3 】

すなわち同期用コード additional_dialogue_data_sync、インデックス additional_dialogue_index、および符号化された追加ダイアログ音源のオーディオ信号が DSE に格納される。これにより、例えば図 4 に示したビットストリームシンタックスに従って追加ダイアログ音源の符号化データが生成されたことになる。

20

【 0 1 0 4 】

ステップ S 2 6 において、追加ダイアログ音源符号化部 2 2 は、インデックス i の値を 1 だけインクリメントして処理はステップ S 2 1 に戻り、上述した処理が繰り返し行われる。

【 0 1 0 5 】

また、ステップ S 2 1 において、 $i < \text{num_additional_dialogue}$ でないと判定された場合、全ての追加ダイアログ音源の符号化が行われたので、追加ダイアログ音源符号化部 2 2 は、符号化により得られた各追加ダイアログ音源の符号化データを多重化部 2 4 に供給し、処理はステップ S 2 7 に進む。

30

【 0 1 0 6 】

ステップ S 2 1 において $i < \text{num_additional_dialogue}$ でないと判定されたか、またはステップ S 1 3 においてダイアログ情報がないと判定されると、ステップ S 2 7 において、多重化部 2 4 は 1 フレーム分の完了処理を行う。

【 0 1 0 7 】

すなわち、多重化部 2 4 は、チャンネル音源符号化部 2 1、追加ダイアログ音源符号化部 2 2、およびダイアログ情報符号化部 2 3 のそれぞれから供給された符号化データを多重化して 1 フレーム分のビットストリームを生成し、出力する。

【 0 1 0 8 】

ステップ S 2 7 において 1 フレーム分のビットストリームが出力されると、処理はステップ S 1 1 に戻り、上述した処理が繰り返し行われる。そして、ステップ S 1 1 において、次に処理対象として符号化すべき 1 フレーム分のデータがないと判定されると、符号化処理は終了する。

40

【 0 1 0 9 】

以上のようにして、符号化装置 1 1 は、チャンネル音源および追加ダイアログ音源のオーディオ信号だけでなく、チャンネル音源のうちのダイアログ音源と、追加ダイアログ音源についての情報を含むダイアログ情報も符号化し、それらの符号化データを多重化して出力する。

【 0 1 1 0 】

このようにダイアログ情報を符号化してビットストリームに含めて伝送させることで、

50

復号側において、必要な追加ダイアログ音源の符号化データのみを復号することができるようになり、復号時の計算量を低減させることができる。

【0111】

復号装置の構成例

次に、符号化装置11から出力されたビットストリームを入力とし、ビットストリームに含まれる符号化データを復号する復号装置について説明する。そのような復号装置は、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1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。

【0112】

図6に示す復号装置71は、分離部81、チャンネル音源復号部82、追加ダイアログ音源復号部83、ダイアログ情報復号部84、提示部85、入力部86、ゲイン補正部87、ダイアログ選択部88、ゲイン補正部89、および追加ダイアログ選択部90を有している。

10

【0113】

分離部81は、入力されたビットストリームをチャンネル音源の符号化データ、追加ダイアログ音源の符号化データ、およびダイアログ情報の符号化データに分離させる。分離部81は、チャンネル音源の符号化データをチャンネル音源復号部82に供給し、追加ダイアログ音源の符号化データを追加ダイアログ音源復号部83に供給し、ダイアログ情報の符号化データをダイアログ情報復号部84に供給する。

【0114】

チャンネル音源復号部82は、分離部81から供給された各チャンネル音源の符号化データを復号し、得られたダイアログチャンネルのオーディオ信号を追加ダイアログ選択部90に供給するとともに、他のチャンネル音源のオーディオ信号を、それらのチャンネル音源に対応するスピーカSP2乃至スピーカSP22に供給する。

20

【0115】

追加ダイアログ音源復号部83は、分離部81から供給された追加ダイアログ音源の符号化データのうち、入力部86から供給された選択指示により示される追加ダイアログ音源の符号化データのみを復号し、ゲイン補正部87に供給する。

【0116】

ダイアログ情報復号部84は、分離部81から供給されたダイアログ情報の符号化データを復号し、その結果得られたダイアログ情報を提示部85に供給する。また、ダイアログ情報復号部84は、復号により得られたダイアログ情報のゲイン値dialogue_gainをゲイン補正部89に供給するとともに、ダイアログ情報の補正ゲイン値additional_dialogue_gain[i]をゲイン補正部87に供給する。

30

【0117】

提示部84は、ダイアログ情報復号部84から供給されたダイアログ情報を視聴者US11に提示する。入力部86は、視聴者US11の入力操作に応じた信号、すなわち視聴者US11により選択されたダイアログ音源または追加ダイアログ音源を示す選択指示を、追加ダイアログ音源復号部83、ダイアログ選択部88、および追加ダイアログ選択部90に供給する。

40

【0118】

ゲイン補正部87は、ダイアログ情報復号部84から供給された補正ゲイン値additional_dialogue_gain[i]に基づいて、追加ダイアログ音源復号部83から供給された追加ダイアログ音源のオーディオ信号に対するゲイン補正を行い、ダイアログ選択部88に供給する。より詳細には、ゲイン補正部87は、各追加ダイアログ音源に対応する増幅器を有しており、各増幅器は追加ダイアログ音源復号部83から供給された追加ダイアログ音源のオーディオ信号に、それらの追加ダイアログ音源に対応する補正ゲイン値を乗じて出力する。

【0119】

ダイアログ選択部88は、入力部86から供給された選択指示に基づいて、ゲイン補正

50

部 8 7 から供給された各追加ダイアログ音源のオーディオ信号のうちの一つを選択し、ゲイン補正部 8 9 に供給する。より詳細には、ダイアログ選択部 8 8 は例えばスイッチからなり、入力部 8 6 からの選択指示に従って、スイッチの出力端をゲイン補正部 8 7 の各追加ダイアログ音源に対応する増幅器の出力端のうちの一つに接続する。これにより、選択指示により示される追加ダイアログ音源のオーディオ信号のみがスイッチの出力端からゲイン補正部 8 9 へと出力される。

【 0 1 2 0 】

ゲイン補正部 8 9 は、ダイアログ情報復号部 8 4 から供給されたゲイン値 `dialogue_gain` に基づいて、ダイアログ選択部 8 8 から供給された追加ダイアログ音源のオーディオ信号に対するゲイン補正を行い、追加ダイアログ選択部 9 0 に供給する。追加ダイアログ選択部 9 0 は例えばスイッチからなり、入力部 8 6 から供給された選択指示に基づいて、チャンネル音源復号部 8 2 から供給されたダイアログ音源のオーディオ信号、またはゲイン補正部 8 9 から供給された追加ダイアログ音源のオーディオ信号の何れか一方を選択し、スピーカ SP1 に供給する。

10

【 0 1 2 1 】

なお、ここでは復号装置 7 1 が分離部 8 1 乃至追加ダイアログ選択部 9 0 を有する例について説明するが、復号装置 7 1 に分離部 8 1 乃至ダイアログ情報復号部 8 4 のみが設けられ、他の提示部 8 5 乃至追加ダイアログ選択部 9 0 は復号装置 7 1 外に設けられてもよい。

【 0 1 2 2 】

復号処理の説明

続いて、復号装置 7 1 の動作について説明する。

20

【 0 1 2 3 】

復号装置 7 1 は、符号化装置 1 1 から伝送されてきたビットストリームが供給されると、復号処理を開始し、復号により得られたチャンネル音源を再生させる。以下、図 7 のフローチャートを参照して、復号装置 7 1 による復号処理について説明する。

【 0 1 2 4 】

ステップ S 6 1 において、分離部 8 1 は、供給されたビットストリームに、まだ処理されていない 1 フレーム分のデータがあるか否かを判定する。

【 0 1 2 5 】

ステップ S 6 1 において 1 フレーム分のデータがあると判定された場合、ステップ S 6 2 において、分離部 8 1 は 1 フレーム分のビットストリームにダイアログ情報が含まれているか否かを判定する。

30

【 0 1 2 6 】

例えばビットストリームを構成する `DSE()` のなかに、同期用コード `dialogue_info_sync` が格納されている `DSE()` があった場合、ダイアログ情報が含まれていると判定される。

【 0 1 2 7 】

ステップ S 6 2 においてダイアログ情報が含まれていないと判定された場合、ダイアログ情報はないので、ステップ S 6 3 およびステップ S 6 4 の処理は行われず、処理はステップ S 6 5 へと進む。このとき、分離部 8 1 は、ビットストリームからチャンネル音源の符号化データを分離させ、チャンネル音源復号部 8 2 へと供給する。

40

【 0 1 2 8 】

これに対して、ステップ S 6 2 においてダイアログ情報が含まれていると判定された場合、処理はステップ S 6 3 へと進む。このとき、分離部 8 1 はビットストリームからチャンネル音源の符号化データを分離させてチャンネル音源復号部 8 2 に供給するとともに、ビットストリームからダイアログ情報の符号化データを分離させてダイアログ情報復号部 8 4 に供給する。

【 0 1 2 9 】

ステップ S 6 3 において、ダイアログ情報復号部 8 4 は、分離部 8 1 から供給された符号化データからメインダイアログ情報 `main_dialogue_info`、ダイアログ位置情報 `dialogue`

50

_src_index、ゲイン値dialogue_gain、および追加ダイアログ個数情報num_additional_dialogueを取得する。また、ダイアログ情報復号部84は、取得したゲイン値dialogue_gainをゲイン補正部89に供給する。

【0130】

ステップS64において、ダイアログ情報復号部84は、分離部81から供給された符号化データから追加ダイアログ個数情報num_additional_dialogueにより示される数の分だけ、追加ダイアログ情報additional_dialogue_info[i]および補正ゲイン値additional_dialogue_gain[i]を取得する。すなわち、インデックスi=0乃至num_additional_dialogue-1について、追加ダイアログ情報additional_dialogue_info[i]および補正ゲイン値additional_dialogue_gain[i]が取得される。

10

【0131】

ダイアログ情報復号部84は、取得した各補正ゲイン値additional_dialogue_gain[i]をゲイン補正部87に供給し、処理はステップS65に進む。

【0132】

以上のステップS63およびステップS64の処理により、例えば図3に示したビットストリームシンタックスに従ってダイアログ情報が復号されたことになる。

【0133】

ステップS62においてダイアログ情報が含まれていないと判定されたか、またはステップS64において追加ダイアログ情報および補正ゲイン値が取得されると、ステップS65の処理が行われる。

20

【0134】

ステップS65において、チャンネル音源復号部82は、ダイアログチャンネルを含むチャンネル音源の各チャンネルの復号を行う。すなわち、チャンネル音源復号部82は、分離部81から供給されたチャンネル音源の符号化データを復号し、その結果得られた各チャンネル音源のオーディオ信号を出力する。このとき、ダイアログチャンネル、つまりダイアログ音源のオーディオ信号は追加ダイアログ選択部90に供給され、他のチャンネル、つまりダイアログ音源ではないチャンネル音源のオーディオ信号は、それらのチャンネル音源に対応するスピーカに供給される。

【0135】

ステップS66において、ダイアログ情報復号部84は、ダイアログ情報があるか否かを判定する。例えばビットストリームのなかに同期用コードdialogue_info_syncが格納されているDSE()があった場合、すなわちステップS63およびステップS64でダイアログ情報の復号が行われた場合、ダイアログ情報があると判定される。

30

【0136】

ステップS66においてダイアログ情報がないと判定された場合、ステップS67乃至ステップS69の処理は行われず、処理はステップS70に進む。

【0137】

これに対して、ステップS66においてダイアログ情報があると判定された場合、ダイアログ情報復号部84は、復号により得られたダイアログ情報を提示部85に供給し、処理はステップS67に進む。

40

【0138】

ステップS67において、提示部85は、ダイアログ情報復号部84から供給されたダイアログ情報を視聴者US11に対して提示する。

【0139】

例えばダイアログ情報のうち、少なくともメインダイアログ情報、追加ダイアログ個数情報、および追加ダイアログ情報が提示される。視聴者US11は、提示されたダイアログ情報を確認し、ダイアログ音源といくつかの追加ダイアログ音源のなかから再生対象とする1つのダイアログを選択する。そして、視聴者US11は自身の選択に応じて入力部86を操作し、ダイアログの選択指示を行う。

【0140】

50

ステップS 6 8において、入力部 8 6 は視聴者US11のダイアログの選択指示を取得する。そして入力部 8 6 は、視聴者US11の操作に応じた選択指示を追加ダイアログ音源復号部 8 3、ダイアログ選択部 8 8、および追加ダイアログ選択部 9 0 に供給する。

【 0 1 4 1 】

ステップS 6 9において、追加ダイアログ音源復号部 8 3、ダイアログ選択部 8 8、および追加ダイアログ選択部 9 0 は、入力部 8 6 から供給された選択指示に基づいて、視聴者US11の選択したダイアログはチャンネル音源に含まれるダイアログチャンネルのダイアログ音源であるか否かを判定する。

【 0 1 4 2 】

ステップS 6 9において、チャンネル音源に含まれるダイアログチャンネルのダイアログ音源であると判定された場合、つまり追加ダイアログ音源への切り替えが指示されなかった場合、処理はステップS 7 0へと進む。このとき、追加ダイアログ音源復号部 8 3では、追加ダイアログ音源の復号は行われない。

【 0 1 4 3 】

ステップS 6 9においてチャンネル音源に含まれるダイアログチャンネルのダイアログ音源であると判定されたか、またはステップS 6 6においてダイアログ情報がないと判定された場合、ステップS 7 0の処理が行われる。

【 0 1 4 4 】

ステップS 7 0において、復号装置 7 1 は、ダイアログチャンネルのダイアログ音源を含む各チャンネル音源を出力する。

【 0 1 4 5 】

すなわち、追加ダイアログ選択部 9 0 は、入力部 8 6 から供給された選択指示に基づいて、チャンネル音源復号部 8 2 から供給されたダイアログチャンネルのオーディオ信号をスピーカSP1に供給し、ダイアログ音源を再生させる。また、スピーカSP2乃至スピーカSP22は、チャンネル音源復号部 8 2 から供給されたオーディオ信号に基づいて、各チャンネル音源を再生する。

【 0 1 4 6 】

このようにして1フレーム分のチャンネル音源が再生されると、処理はステップS 6 1に戻り、上述した処理が繰り返し行われる。

【 0 1 4 7 】

これに対して、ステップS 6 9においてチャンネル音源に含まれるダイアログチャンネルのダイアログ音源でないと判定された場合、つまり追加ダイアログ音源であると判定された場合、処理はステップS 7 1に進む。

【 0 1 4 8 】

ステップS 7 1において、追加ダイアログ音源復号部 8 3 は、入力部 8 6 から供給された選択指示に基づいて、選択された追加ダイアログ音源の符号化データが格納されているDSE()を分離部 8 1 から取得し、例えば図 4 に示した追加ダイアログ音源の符号化データのビットストリームシンタックスに従って追加ダイアログ音源を復号する。

【 0 1 4 9 】

この処理により、選択指示により示される追加ダイアログ音源の符号化データのみが復号される。追加ダイアログ音源復号部 8 3 は、復号により得られた追加ダイアログ音源のオーディオ信号をゲイン補正部 8 7 に供給する。

【 0 1 5 0 】

ステップS 7 2において、ゲイン補正部 8 7 は、ダイアログ情報復号部 8 4 から供給された、選択指示により示される追加ダイアログ音源の補正ゲイン値additional_dialogue_gain[i]を、追加ダイアログ音源復号部 8 3 から供給された追加ダイアログ音源のオーディオ信号に乗算し、ゲイン補正を行う。

【 0 1 5 1 】

ゲイン補正部 8 7 はゲイン補正された追加ダイアログ音源のオーディオ信号をダイアログ選択部 8 8 に供給する。また、ダイアログ選択部 8 8 は、入力部 8 6 から供給された選

10

20

30

40

50

択指示に基づいてスイッチの接続先を切り替える。これにより、ゲイン補正部 87 から供給された、選択指示により示される追加ダイアログ音源のオーディオ信号が、ゲイン補正部 89 へと供給される。

【0152】

ステップ S73 において、ゲイン補正部 89 は、ダイアログ情報復号部 84 から供給されたゲイン値 `dialogue_gain` を、ダイアログ選択部 88 から供給された追加ダイアログ音源のオーディオ信号に乗じることでゲイン補正を行い、追加ダイアログ選択部 90 に供給する。

【0153】

ステップ S74 において、追加ダイアログ選択部 90 は、入力部 86 から供給された選択指示に基づいて、チャンネル音源復号部 82 から供給されたダイアログチャンネル、つまりダイアログ音源のオーディオ信号を、ゲイン補正部 89 から供給された追加ダイアログ音源のオーディオ信号に置き換えて出力する。

【0154】

この場合、選択指示は選択された追加ダイアログ音源を示しているので、追加ダイアログ選択部 90 はゲイン補正部 89 から供給された追加ダイアログ音源のオーディオ信号をスピーカ SP1 に供給し、追加ダイアログ音源を再生させる。また、スピーカ SP2 乃至スピーカ SP22 は、チャンネル音源復号部 82 から供給されたオーディオ信号に基づいて、各チャンネル音源を再生する。

【0155】

これにより、スピーカ SP2 乃至スピーカ SP22 でチャンネル音源が再生され、スピーカ SP1 ではチャンネル音源としてのダイアログ音源に代えて、追加ダイアログ音源が再生されることになる。

【0156】

このようにして1フレーム分のチャンネル音源が再生されると、処理はステップ S61 に戻り、上述した処理が繰り返し行われる。

【0157】

また、ステップ S70 またはステップ S74 においてチャンネル音源が再生された後、ステップ S61 において1フレーム分のデータがないと判定されると、復号処理は終了する。

【0158】

以上のようにして復号装置 71 は、ビットストリームからダイアログ情報の符号化データを取得して復号し、ダイアログ情報を提示するとともに、視聴者 US11 の選択指示に応じて、チャンネル音源と、必要な追加ダイアログ音源とを復号する。

【0159】

このようにダイアログ情報を復号することで、ダイアログ情報を提示することができ、またダイアログ情報の提示に応じてなされた選択指示に応じて、必要な追加ダイアログ音源のみを復号することで、復号の計算量を低減させることができる。

【0160】

なお、ここでは置き換え対象とされるチャンネル、すなわちダイアログチャンネルのオーディオ信号によりダイアログ音源が再生されるスピーカ位置がスピーカ SP1 である例について説明した。しかし、ダイアログ情報には、ダイアログ位置情報 `dialogue_src_index` が含まれているため、ダイアログ音源を再生するスピーカの位置として任意のスピーカ位置をフレームごとに指定することが可能である。

【0161】

そのような場合、ダイアログ位置情報 `dialogue_src_index` に基づいて、置き換え対象となるスピーカ位置、つまり置き換え対象となるチャンネル音源が特定される。そして、特定されたチャンネル音源について、視聴者 US11 の選択指示に応じて、チャンネル音源のオーディオ信号がそのまま出力されるか、またはそのチャンネル音源のオーディオ信号が、選択された追加音源のオーディオ信号に置き換えられるかが選択される。

10

20

30

40

50

【 0 1 6 2 】

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。

【 0 1 6 3 】

図 8 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。

10

【 0 1 6 4 】

コンピュータにおいて、CPU (Central Processing Unit) 5 0 1 , ROM (Read Only Memory) 5 0 2 , RAM (Random Access Memory) 5 0 3 は、バス 5 0 4 により相互に接続されている。

【 0 1 6 5 】

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。

【 0 1 6 6 】

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 5 0 7 は、ディスプレイ、スピーカなどよりなる。記録部 5 0 8 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 5 0 9 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 5 1 0 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 5 1 1 を駆動する。

20

【 0 1 6 7 】

以上のように構成されるコンピュータでは、CPU 5 0 1 が、例えば、記録部 5 0 8 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 5 0 5 及びバス 5 0 4 を介して、RAM 5 0 3 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。

【 0 1 6 8 】

コンピュータ (CPU 5 0 1) 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 5 1 1 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。

30

【 0 1 6 9 】

コンピュータでは、プログラムは、リムーバブルメディア 5 1 1 をドライブ 5 1 0 に装着することにより、入出力インターフェース 5 0 5 を介して、記録部 5 0 8 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 5 0 9 で受信し、記録部 5 0 8 にインストールすることができる。その他、プログラムは、ROM 5 0 2 や記録部 5 0 8 に、あらかじめインストールしておくことができる。

【 0 1 7 0 】

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。

40

【 0 1 7 1 】

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。

【 0 1 7 2 】

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。

【 0 1 7 3 】

50

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。

【0174】

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。

【0175】

さらに、本技術は、以下の構成とすることも可能である。

【0176】

[1]

ダイアログ音源を含む複数のチャンネル音源のオーディオ信号を符号化し、符号化データを生成するチャンネル音源符号化部と、

前記ダイアログ音源とは異なる追加ダイアログ音源のオーディオ信号を符号化し、符号化データを生成する追加ダイアログ音源符号化部と、

前記追加ダイアログ音源に関する情報を含むダイアログ情報、前記複数のチャンネル音源の符号化データ、および前記追加ダイアログ音源の符号化データを多重化する多重化部とを備える符号化装置。

[2]

前記追加ダイアログ音源の符号化データおよび前記ダイアログ情報は、前記多重化により得られるビットストリームのユーザデータ領域に格納される

[1]に記載の符号化装置。

[3]

前記ユーザデータ領域は、MPEG AACまたはMPEG-D USACで規定されたDSEである

[2]に記載の符号化装置。

[4]

前記ダイアログ情報には、前記ダイアログ音源に関する情報が含まれている

[1]乃至[3]の何れか一項に記載の符号化装置。

[5]

前記ダイアログ情報には、前記追加ダイアログ音源の種別を示す情報、および前記ダイアログ音源の種別を示す情報が含まれている

[4]に記載の符号化装置。

[6]

前記ダイアログ情報には、マルチチャンネル構成である前記複数のチャンネル音源のオーディオ信号のチャンネルのうちの前記ダイアログ音源のチャンネルを示す情報が含まれている

[4]または[5]に記載の符号化装置。

[7]

前記ダイアログ情報には、前記追加ダイアログ音源のオーディオ信号をゲイン補正するためのゲイン情報が含まれている

[1]乃至[6]の何れか一項に記載の符号化装置。

[8]

前記ダイアログ情報には、複数の前記追加ダイアログ音源ごとに、それらの複数の前記追加ダイアログ音源のオーディオ信号をゲイン補正するための補正ゲイン情報が含まれている

[1]乃至[7]の何れか一項に記載の符号化装置。

[9]

ダイアログ音源を含む複数のチャンネル音源のオーディオ信号を符号化して符号化データを生成し、

前記ダイアログ音源とは異なる追加ダイアログ音源のオーディオ信号を符号化して符号化データを生成し、

前記追加ダイアログ音源に関する情報を含むダイアログ情報、前記複数のチャンネル音源

10

20

30

40

50

の符号化データ、および前記追加ダイアログ音源の符号化データを多重化するステップを含む符号化方法。

[1 0]

ダイアログ音源を含む複数のチャンネル音源のオーディオ信号を符号化して符号化データを生成し、

前記ダイアログ音源とは異なる追加ダイアログ音源のオーディオ信号を符号化して符号化データを生成し、

前記追加ダイアログ音源に関する情報を含むダイアログ情報、前記複数のチャンネル音源の符号化データ、および前記追加ダイアログ音源の符号化データを多重化する

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。

10

[1 1]

符号化ビットストリームを、ダイアログ音源を含む複数のチャンネル音源の符号化データ、前記ダイアログ音源とは異なる1または複数の追加ダイアログ音源の符号化データ、および前記追加ダイアログ音源に関する情報を含むダイアログ情報に分離させる分離部と、

前記1または複数の前記追加ダイアログ音源の符号化データのうちの選択された前記追加ダイアログ音源の符号化データを復号する追加ダイアログ音源復号部と

を備える復号装置。

[1 2]

前記複数のチャンネル音源の符号化データを復号するチャンネル音源復号部をさらに備える

[1 1] に記載の復号装置。

20

[1 3]

前記ダイアログ音源の符号化データを復号して得られたオーディオ信号を、前記選択された前記追加ダイアログ音源の符号化データを復号して得られたオーディオ信号に置き換えて出力するダイアログ選択部をさらに備える

[1 1] または [1 2] に記載の復号装置。

[1 4]

前記追加ダイアログ音源復号部は、前記ダイアログ情報の提示に応じて選択された前記追加ダイアログ音源の符号化データを復号する

[1 1] 乃至 [1 3] の何れか一項に記載の復号装置。

[1 5]

符号化ビットストリームを、ダイアログ音源を含む複数のチャンネル音源の符号化データ、前記ダイアログ音源とは異なる1または複数の追加ダイアログ音源の符号化データ、および前記追加ダイアログ音源に関する情報を含むダイアログ情報に分離させ、

前記1または複数の前記追加ダイアログ音源の符号化データのうちの選択された前記追加ダイアログ音源の符号化データを復号する

ステップを含む復号方法。

30

[1 6]

符号化ビットストリームを、ダイアログ音源を含む複数のチャンネル音源の符号化データ、前記ダイアログ音源とは異なる1または複数の追加ダイアログ音源の符号化データ、および前記追加ダイアログ音源に関する情報を含むダイアログ情報に分離させ、

前記1または複数の前記追加ダイアログ音源の符号化データのうちの選択された前記追加ダイアログ音源の符号化データを復号する

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。

40

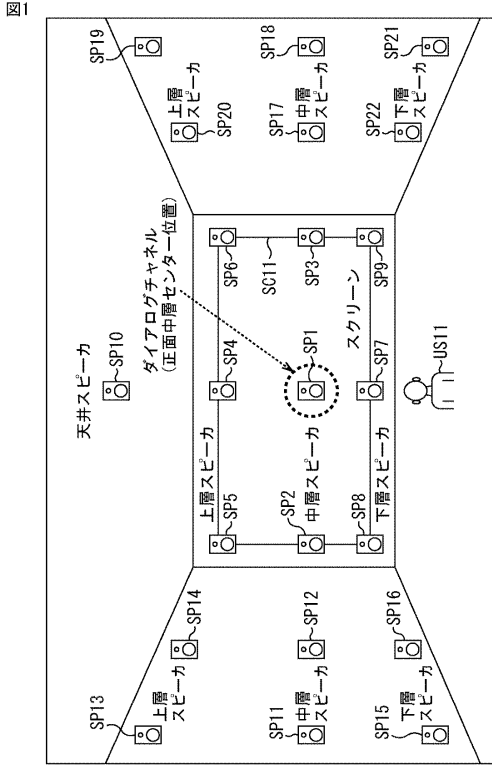
【符号の説明】

【 0 1 7 7 】

1 1 符号化装置, 2 1 チャンネル音源符号化部, 2 2 追加ダイアログ音源符号化部, 2 3 ダイアログ情報符号化部, 2 4 多重化部, 7 1 復号装置, 8 1 分離部, 8 2 チャンネル音源復号部, 8 3 追加ダイアログ音源復号部, 8 4 ダイアログ情報復号部, 8 5 提示部, 8 6 入力部, 8 7 ゲイン補正部, 8 8 ダイアログ選択部, 8 9 ゲイン補正部, 9 0 追加ダイアログ選択部

50

【図 1】



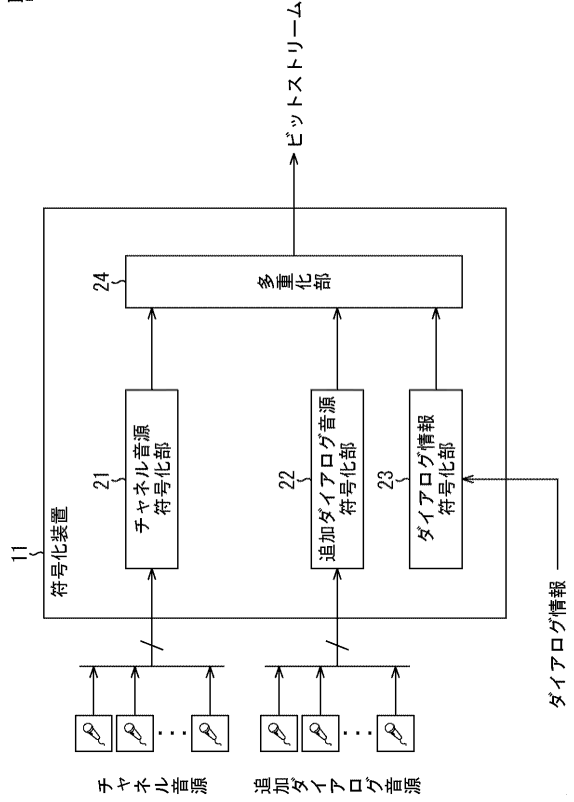
【図 3】

図3

Syntax	No. of Bits
<code>dialogue_info_sync;</code>	16
<code>main_dialogue_info;</code>	24
<code>dialogue_src_index;</code>	8
<code>dialogue_gain;</code>	32
<code>num_additional_dialogue;</code>	4
<code>for (i=0; i<<num_additional_dialogue; i++) {</code>	
<code>additional_dialogue_info[i];</code>	24
<code>additional_dialogue_gain[i];</code>	32
<code>}</code>	

【図 2】

図2

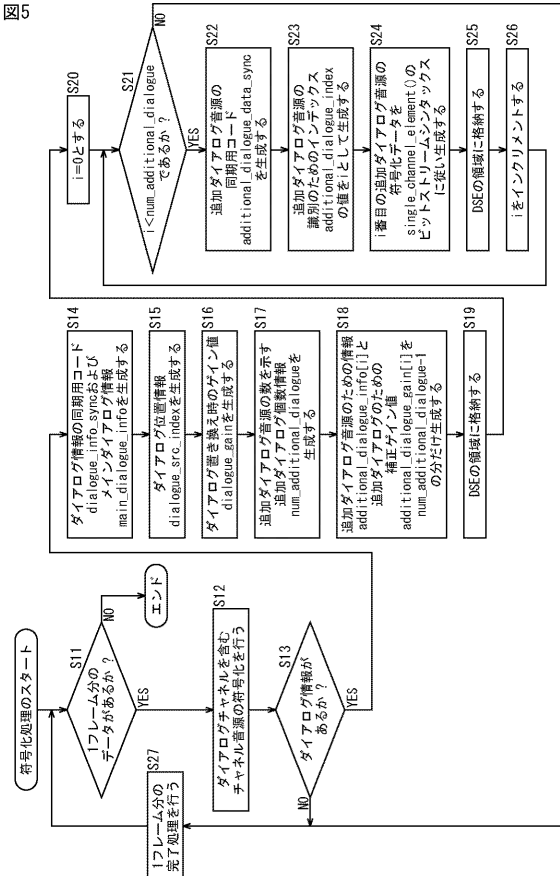


【図 4】

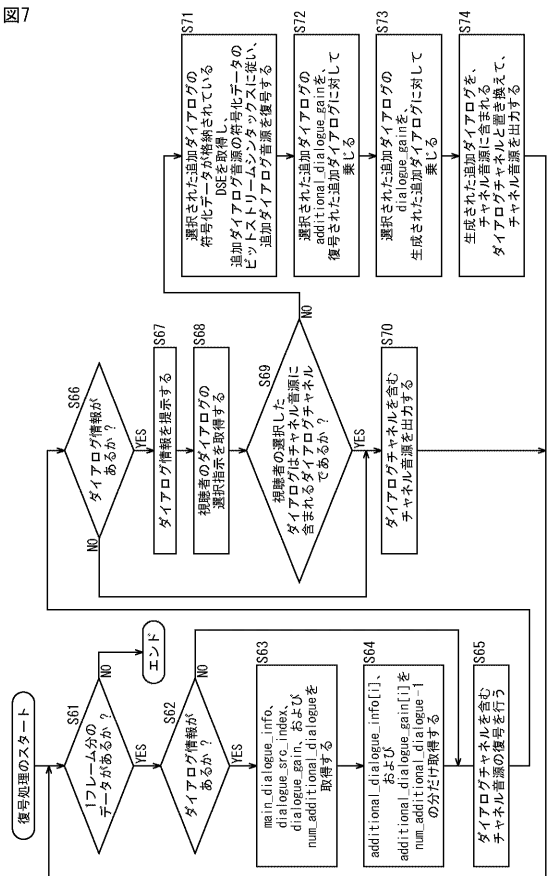
図4

Syntax	No. of Bits
<code>additional_dialogue_data_sync;</code>	16
<code>additional_dialogue_index;</code>	4
<code>single_channel_element();</code>	

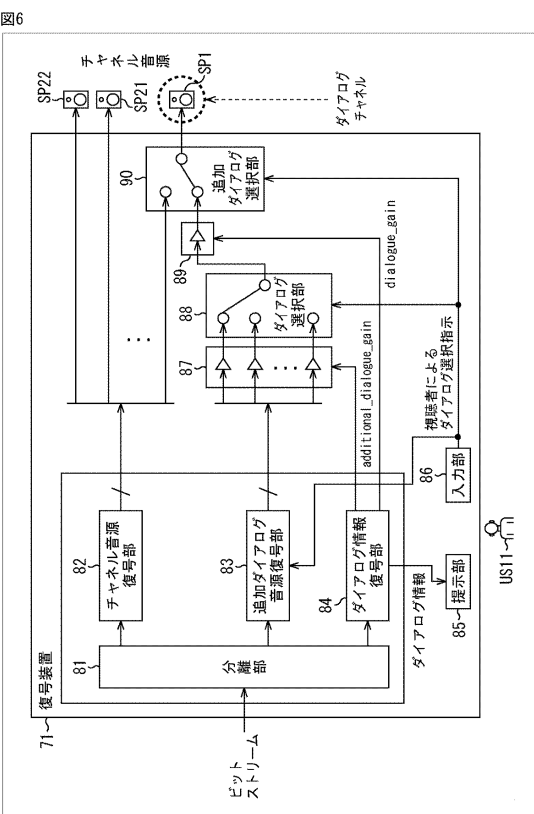
【図5】



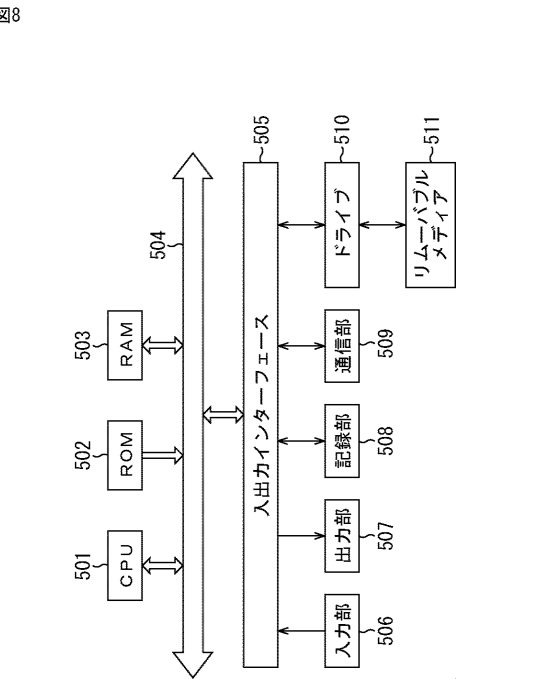
【図7】



【図6】



【図8】



フロントページの続き

- (72)発明者 辻 実
東京都港区港南1丁目7番1号 ソニー株式会社内
- (72)発明者 本間 弘幸
東京都港区港南1丁目7番1号 ソニー株式会社内

審査官 富澤 直樹

- (56)参考文献 特開2014-142475(JP,A)
特表2008-505586(JP,A)
国際公開第2014/114781(WO,A1)
特開2012-010311(JP,A)

- (58)調査した分野(Int.Cl., DB名)
G10L 19/00 - 19/26