

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第5111375号  
(P5111375)

(45) 発行日 平成25年1月9日(2013.1.9)

(24) 登録日 平成24年10月19日(2012.10.19)

(51) Int.Cl.		F I		
G 1 0 L 19/00	(2013.01)	G 1 0 L 19/00	3 3 0 B	
G 1 0 L 19/20	(2013.01)	G 1 0 L 19/00	2 1 3	
G 1 0 L 19/00B	(2013.01)	G 1 0 L 19/00	4 0 0 Z	

請求項の数 4 (全 43 頁)

(21) 出願番号	特願2008-528942 (P2008-528942)	(73) 特許権者	502032105
(86) (22) 出願日	平成18年8月30日 (2006. 8. 30)		エルジー エレクトロニクス インコーポ レイティド
(65) 公表番号	特表2009-506374 (P2009-506374A)		大韓民国, ソウル 150-721, ヨン ドンポーク, ヨイドードン, 20
(43) 公表日	平成21年2月12日 (2009. 2. 12)	(74) 代理人	100099759
(86) 国際出願番号	PCT/KR2006/003423		弁理士 青木 篤
(87) 国際公開番号	W02007/055462	(74) 代理人	100092624
(87) 国際公開日	平成19年5月18日 (2007. 5. 18)		弁理士 鶴田 準一
審査請求日	平成20年4月21日 (2008. 4. 21)	(74) 代理人	100102819
(31) 優先権主張番号	60/712, 119		弁理士 島田 哲郎
(32) 優先日	平成17年8月30日 (2005. 8. 30)	(74) 代理人	100151459
(33) 優先権主張国	米国 (US)		弁理士 中村 健一
(31) 優先権主張番号	60/719, 202	(74) 代理人	100108383
(32) 優先日	平成17年9月22日 (2005. 9. 22)		弁理士 下道 晶久
(33) 優先権主張国	米国 (US)		

最終頁に続く

(54) 【発明の名称】 オーディオ信号をエンコーディング及びデコーディングするための装置とその方法

(57) 【特許請求の範囲】

【請求項 1】

オーディオ信号をデコードする方法であって、  
ダウンミックス信号と、少なくとも一つのタイムスロット及び少なくとも一つのパラメータセットを有する少なくとも一つのフレームを含む空間情報とを含むオーディオ信号を受信するステップであって、前記パラメータセットは少なくとも一つのパラメータを含む、ステップと、

OTTボックスが低周波拡張(LFE)モードを示すかどうかを判定するステップであって、前記LFEモードは前記OTTボックスがLFEチャンネルに用いられていることを示す、ステップと、

パラメータセットが適用されるタイムスロットを示す、可変ビット長のタイムスロット情報を抽出するステップと、

固定ビット長のLFEパラメータ帯域情報を抽出するステップであって、該LFEパラメータ帯域情報は前記LFEチャンネルのパラメータ帯域の個数を示し、該パラメータ帯域はパラメータが適用される周波数帯域を示し、前記個数は前記OTTボックスが前記LFEモードを示すかどうかに応じて異なる、ステップと、

前記LFEパラメータ帯域情報及び前記タイムスロット情報に基づいて、前記タイムスロットのパラメータ帯域に前記パラメータセットのパラメータを適用することによって、前記オーディオ信号をデコードするステップと、を有し、

前記タイムスロット情報を抽出するステップは、

タイムスロット情報を特定するために、前記オーディオ信号からタイムスロットの個数及びパラメータセットの個数を抽出するステップと、

前記タイムスロット情報のビット長を判定するステップであって、該ビット長は、前記タイムスロットの個数と、前記パラメータセットの個数と、先行するパラメータセットと関係する先行するタイムスロット情報とによって可変である、ステップと、

前記ビット長に基づいて、前記タイムスロット情報を抽出するステップとを含み、

前記タイムスロット情報の個数は前記パラメータセットの個数に等しく、

前記タイムスロット情報は、第1パラメータセットが適用されるタイムスロットを示す絶対値、及び前記第1パラメータセットに続くパラメータセットが適用されるタイムスロットを示す差分値を含み、

10

前記差分値は、前記先行するタイムスロット情報と、前記続くパラメータセットが適用されるタイムスロットとの差を示し、前記先行するタイムスロット情報には前記続くパラメータセットの前記先行するパラメータセットが適用され、

前記第1パラメータセットに続くパラメータセットが適用されるタイムスロットは、前記先行するタイムスロット情報に前記差分値を加算することによって決定され、

前記絶対値は前記パラメータセットの個数及び前記タイムスロットの個数を用いて計算される第1最大範囲内で決定され、前記差分値は前記先行するタイムスロット情報によって計算される第2最大範囲内で決定される、方法。

#### 【請求項2】

前記タイムスロット情報は、パラメータセットが適用されるタイムスロットの位置を示す位置情報である、請求項1に記載の方法。

20

#### 【請求項3】

オーディオ信号をデコードする装置であって、該オーディオ信号は、ダウンミックス信号と、少なくとも一つのタイムスロット及び少なくとも一つのパラメータセットを有する少なくとも一つのフレームを含む空間情報とを含み、前記パラメータセットは少なくとも一つのパラメータを含み、前記装置は、

前記ダウンミックス信号をデコードするダウンミックス信号デコーディング部と、

OTTボックスが低周波拡張(LFE)モードを示すかどうかを判定し、前記LFEモードは前記OTTボックスがLFEチャンネルに用いられていることを示し、

パラメータセットが適用されるタイムスロットを示す、可変ビット長のタイムスロット情報を抽出し、

30

固定ビット長のLFEパラメータ帯域情報を抽出し、該LFEパラメータ帯域情報は前記LFEチャンネルのパラメータ帯域の個数を示し、該パラメータ帯域はパラメータが適用される周波数帯域を示し、前記個数は前記OTTボックスが前記LFEモードを示すかどうかに応じて異なり、

タイムスロット情報を特定するために、前記オーディオ信号からタイムスロットの個数及びパラメータセットの個数を抽出し、

前記タイムスロット情報のビット長を判定し、該ビット長は、前記タイムスロットの個数と、前記パラメータセットの個数と、先行するパラメータセットと関係する先行するタイムスロット情報とによって可変であり、

40

前記ビット長に基づいて、前記タイムスロット情報を抽出し、前記タイムスロット情報は、パラメータセットが適用されるタイムスロットを示す、空間情報デコーディング部と、

前記LFEパラメータ帯域情報及び前記タイムスロット情報に基づいて、前記タイムスロットのパラメータ帯域に前記パラメータセットのパラメータを適用することによって、前記オーディオ信号をデコードするアップミキシング部と、を備え、

前記タイムスロット情報の個数は前記パラメータセットの個数に等しく、前記タイムスロット情報は、第1パラメータセットが適用されるタイムスロットを示す絶対値、及び前記第1パラメータセットに続くパラメータセットが適用されるタイムスロットを示す差分値を含み、

50

前記差分値は、前記先行するタイムスロット情報と、前記続くパラメータセットが適用されるタイムスロットとの差を示し、前記先行するタイムスロット情報には前記続くパラメータセットの前記先行するパラメータセットが適用され、

前記第1パラメータセットに続くパラメータセットが適用されるタイムスロットは、前記先行するタイムスロット情報に前記差分値を加算することによって決定され、

前記絶対値は前記パラメータセットの個数及び前記タイムスロットの個数を用いて計算される第1最大範囲内で決定され、前記差分値は前記先行するタイムスロット情報によって計算される第2最大範囲内で決定される、装置。

【請求項4】

前記タイムスロット情報は、パラメータセットが適用されるタイムスロットの位置を示す位置情報である、請求項3に記載の装置。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は主としてオーディオ信号処理に関する。

【背景技術】

【0002】

通常、空間音声コーディング(SAC: Spatial Audio Coding)と呼ばれるマルチチャンネルオーディオのコーディングを認識するための新たなアプローチ方式が研究開発中にある。SACは、マルチチャンネルオーディオを低いビットレートにて送信可能にすることから、種々のオーディオ適用分野(例えば、インターネットストリーミング、音楽のダウンロードなど)に向いている。

【0003】

SACは、個別のオーディオ入力チャンネルを分散コーディングするよりも簡単なセットのパラメータでマルチチャンネルオーディオ信号の空間イメージを取得する。このようなパラメータは、デコーダーに送信されてオーディオ信号の空間特性を合成したり再構成するのに用いられる。

【0004】

一部のSAC適用分野においては、空間パラメータがビットストリームの一部としてデコーダーに送信される。このようなビットストリームは複数の空間フレームを含み、これらの空間フレームは空間パラメータが適用可能に並べられたタイムスロットセットを含む。また、ビットストリームは位置情報を含むが、デコーダーはこのような位置情報を用いて所定のパラメータセットが適用される正確なタイムスロットを識別することができる。

【0005】

一部のSAC適用分野においては、エンコーディング/デコーディング経路に概念的なエレメントを用いる。そのようなエレメントの一つとしては、通常OTT(One-Two)と呼ばれるものがあり、他のエレメントとしては、通常TTT(Two-Three)と呼ばれるものがあるが、このような名称は、それぞれ対応デコーダー要素の入力チャンネルと出力チャンネルの個数を意味している。OTTエンコーダーエレメントは、2つの空間パラメータを抽出してダウンミックス信号とレジデュアル信号を生成する。TTTエレメントは、3つのオーディオ信号を1つのダウンミックス信号と1つのレジデュアル信号にダウンミックスする。これらのエレメントが組み合わせられて種々の構成の空間オーディオ環境(例えば、サラウンドサウンドなど)を提供することができる。

【0006】

一部のSAC適用分野は、ノンガイド動作モードにて動作可能であるが、このような動作モードにおいては空間パラメータを送信する必要がなく、ステレオダウンミックス信号だけがエンコーダーからデコーダーに送信される。デコーダーは、ダウンミックス信号から空間パラメータを合成してマルチチャンネルオーディオ信号の生成に用いる。

【発明の開示】

10

20

30

40

50

## 【発明が解決しようとする課題】

## 【0007】

オーディオ信号に関連する空間情報は、デコーダーに送信されるか、または、記録媒体に記録可能なビットストリームにエンコーディングされる。このようなビットストリームは、時間、周波数及び空間エリアに関連する異なるシンタックスを含むことができる。一部の実施形態において、ビットストリームはパラメータが適用可能に並べられたセットのスロットを含む1以上のデータ構造（例えば、フレーム）を含むことができる。このようなデータ構造は、固定されたものであっても、可変的なものであってもよい。デコーダーがデータ構造タイプを判定して適切なデコーディング処理を呼び出せるようにするデータ構造タイプ表示子がビットストリームに含まれる。データ構造は位置情報を含むことができ、デコーダーはこのような位置情報を用いて所定のパラメータセットが適用される正確なスロットを識別することができる。スロット位置情報は、データ構造タイプ表示子が表示するデータ構造タイプに応じて、固定個数のビットまたは可変個数のビットでエンコーディング可能である。可変データ構造タイプについて、スロット位置情報は前記並べられたセットのスロットにおいて当該スロットの位置により可変個数のビットでエンコーディング可能である。

10

## 【課題を解決するための手段】

## 【0008】

一部の実施形態において、オーディオ信号をエンコーディングする方法は、オーディオ信号の第1の情報または第2の情報に対応するか、あるいは、前記第1の情報または第2の情報の範囲に対応するオーディオ信号のパラメータセットを生成するステップと、前記パラメータセット及び前記対応する第1の情報または第2の情報を前記オーディオ信号を示すビットストリームに挿入するステップと、を含み、前記第1の情報または前記第2の情報は可変個数のビットで表わされることを特徴とする。

20

## 【0009】

一部の実施形態において、オーディオ信号をデコーディングする方法は、オーディオ信号の第1の情報または第2の情報に対応するか、あるいは、前記第1の情報または第2の情報の範囲に対応するパラメータセットを含み、オーディオ信号を示すビットストリームを受信するステップと、前記パラメータセット及び前記第1の情報または第2の情報に基づき前記オーディオ信号をデコーディングするステップと、を含み、前記第1の情報または第2の情報は可変個数のビットで表わされることを特徴とする。

30

## 【0010】

システム、方法、装置、データ構造及びコンピュータにて読取り可能な媒体に関する多数フレームタイプのタイムスロット位置コーディングの他の実施形態も開示される。

## 【0011】

上述した通常の説明及び以下の実施形態の詳細な説明はいずれも例示的に説明するためのものであり、特許請求の範囲において請求される本発明の理解を容易にするためのものであるという点が理解されなければならない。

## 【発明を実施するための最良の形態】

## 【0012】

本発明への理解を容易にするために含まれる添付図面は本発明の実施形態を示すものであり、この明細書と一緒に本発明の原理を説明するためのものである。

40

## 【0013】

図1は、本発明の一実施形態により空間情報を生成する原理を示す図である。マルチチャンネルオーディオ信号に対するコーディング方式の概念は、人間がオーディオ信号を3次元的に認識するという点に基づく。オーディオ信号の3次元空間は空間情報を用いて表わすことができ、これは、チャンネルレベル差分（CLD；Channel Level Difference）と、チャンネル間相関／一貫性（ICC；Inter Channel Correlation/Coherence）と、チャンネル時間差分（CTD；Channel Time Difference）と、チャンネル予測係数（CPC

50

: Channel Prediction Coefficients)などを含むが、これらに制限されるものではない。CLDは2つのオーディオチャンネル間エネルギー(レベル)差分を意味し、ICCは2つのオーディオチャンネル間相関または一貫性の量を意味し、CTDは2チャンネル間の時間差分を意味する。

【0014】

図1にCTDとCLDパラメータの生成を示す。遠距離サウンドソース101から第1のダイレクトサウンド波103が人間の左耳107に達し、第2のダイレクトサウンド波102が人間の頭の周りにおいて回折された後、人間の右耳106に達する。2つのサウンド波102及び103は到達時間とエネルギーレベルにおいて互いに異なる。CTDパラメータとCLDパラメータはサウンド波102と103の到達時間及びエネルギーレベル差分に基づき生成される。また、反射されたサウンド波104及び105が両耳106及び107にそれぞれ達し、これらは互いに相関がない。ICCパラメータはサウンド波104及び105間相関に基づき生成可能である。

10

【0015】

エンコーダーにおいては、マルチチャンネルオーディオ信号において空間情報(例えば、空間パラメータなど)が抽出され、ダウンミックス信号が生成される。ダウンミックス信号と空間パラメータはデコーダーに転送される。これに制限されるものではないが、モノ信号、ステレオ信号またはマルチチャンネルオーディオ信号を含むダウンミックス信号に任意の個数のオーディオチャンネルが使用可能である。デコーダーにおいては、ダウンミックス信号と空間パラメータからマルチチャンネルアップミックス信号が生成される。

20

【0016】

図2は、本発明の一実施形態によりオーディオ信号をエンコーディングするエンコーダーのブロック図である。エンコーダーは、ダウンミキシング部202と、空間情報生成部203と、ダウンミックス信号エンコーディング部207とマルチプレクシング部209と、を備える。エンコーダーの他の構成も採用可能である。エンコーダーはハードウェアまたはソフトウェアにより実現されるか、あるいは、ハードウェアとソフトウェアとの組み合わせにより実現可能である。エンコーダーは集積回路チップ、チップセット、システムオンチップ(SoC: System on Chip)、デジタル信号プロセッサ、汎用プロセッサ及び種々のデジタル装置とアナログ装置により実現可能である。

30

【0017】

ダウンミキシング部202は、マルチチャンネルオーディオ信号201からダウンミックス信号204を生成する。図2において、 $x_1, \dots, x_n$ は入力オーディオチャンネルを示す。上述したように、ダウンミックス信号204は、モノ信号であっても、ステレオ信号であっても、オーディオ信号であってもよい。図示の例において、 $x'_1, \dots, x'_m$ はダウンミックス信号204のチャンネル番号を示す。一部の実施形態において、エンコーダーはダウンミックス信号204に代えて外部供給ダウンミックス信号205(例えば、精度よいダウンミックスなど)を処理する。

【0018】

空間情報生成部203は、マルチチャンネルオーディオ信号201から空間情報を抽出する。この場合、「空間情報」とは、デコーダーにおいてダウンミックス信号204をマルチチャンネルオーディオ信号にアップミキシングするのに用いられるオーディオ信号チャンネルに関連する情報を意味している。ダウンミックス信号204は、マルチチャンネルオーディオ信号をダウンミックスすることにより生成される。空間情報はエンコーディングされて、エンコーディングされた空間情報信号206を与える。

40

【0019】

ダウンミックス信号エンコーディング部207は、ダウンミキシング部202において生成されたダウンミックス信号204をエンコーディングして、エンコーディングされたダウンミックス信号208を生成する。

【0020】

マルチプレクシング部209は、エンコーディングされたダウンミックス信号208と

50

エンコーディングされた空間情報信号 206 とを含むビットストリーム 210 を生成する。ビットストリーム 210 はダウンストリームデコーダーに転送され、及び/または、記録媒体に記録される。

【0021】

図3は、本発明の一実施形態により、エンコーディングされたオーディオ信号をデコーディングするデコーダーのブロック図である。デコーダーは、デマルチプレクシング部302と、ダウンミックス信号デコーディング部305と、空間情報デコーディング部307と、アップミキシング部309と、を備える。デコーダーは、ハードウェアやソフトウェアにより、または、ハードウェアとソフトウェアとの組み合わせにより実現可能である。デコーダーは、集積回路チップ、チップセット、システムオンチップ(SoC)、デジタル信号プロセッサ、汎用プロセッサ及び種々のデジタル装置やデバイスにより実現可能である。

10

【0022】

一部の実施形態において、デマルチプレクシング部302は、オーディオ信号を示すビットストリーム301を受信して、このビットストリーム301から、エンコーディングされたダウンミックス信号303とエンコーディングされた空間情報信号304とを分離する。図3において、 $x'_1, \dots, x'_m$ はダウンミックス信号303のチャンネルを示す。ダウンミックス信号デコーディング部305は、エンコーディングされたダウンミックス信号303をデコーディングして、デコーディングされたダウンミックス信号306を出力する。デコーダーがマルチチャンネルオーディオ信号を出力することができない場合、ダウンミックス信号デコーディング部305はダウンミックス信号306を直接的に出力することができる。図3において、 $y'_1, \dots, y'_m$ はダウンミックス信号デコーディング部305の直接出力チャンネルを示す。

20

【0023】

空間情報信号デコーディング部307は、エンコーディングされた空間情報信号304から空間情報信号の構成情報を抽出し、抽出された構成情報を用いて空間情報信号304をデコーディングする。

【0024】

アップミキシング部309は、抽出された空間情報308を用いてダウンミックス信号306をマルチチャンネルオーディオ信号310にアップミックスすることができる。図3において、 $y_1, \dots, y_n$ はアップミキシング部309の出力チャンネル番号を示す。

30

【0025】

図4は、図3に示すデコーダーのアップミキシング部309に含まれるチャンネル変換モジュールのブロック図である。一部の実施形態において、アップミキシング部309は複数のチャンネル変換モジュールを含むことができる。チャンネル変換モジュールは、特定の情報を用いて入力チャンネルの個数と出力チャンネルの個数を区別可能な概念的な装置である。

【0026】

一部の実施形態において、チャンネル変換モジュールは、1チャンネルを2チャンネルに、及びその逆に変換するOTT(One-To-Two)ボックスと、2チャンネルを3チャンネルに、及びその逆に変換するTTT(Two-To-Three)ボックスを含む。OTTボックス及び/またはTTTボックスは種々の有用な構成にて配置可能である。例えば、図3に示すアップミキシング部309は、5-1-5構成、5-2-5構成、7-2-7構成、7-5-7構成などを含むことができる。5-1-5構成においては、5つのチャンネルを1つのチャンネルにダウンミキシングして1つのチャンネルを有するダウンミックス信号が生成されるが、これは、今後、5つのチャンネルにアップミックス可能である。OTTボックスとTTTボックスの種々の組み合わせを用いる他の構成も同様に生成可能である。

40

【0027】

図4には、アップミキシング部400の5-2-5構成例が示してある。5-2-5構

50

成においては、2チャンネルを有するダウンミックス信号401がアップミキシング部400に入力される。図示の例には、左側のチャンネルLと右側のチャンネルRがアップミキシング部400への入力として与えられる。この実施形態において、アップミキシング部400は、1つのTTTボックス402と3つのOTTボックス406、407及び408を備える。2チャンネルを有するダウンミックス信号401がTTTボックスTTT0に対する入力として与えられ、TTTボックスはダウンミックス信号401を処理して3つの出力チャンネル403、404及び405を与える。TTTボックス402に対する入力として、1以上の空間パラメータ（例えば、CPC、CLD、ICCなど）が与えられて、後述するように、ダウンミックス信号401を処理するのに使用可能である。この場合、CPCは2つのチャンネルから3つのチャンネルを生成する予測係数として説明可能である。

10

#### 【0028】

TTTボックス402からの出力として与えられるチャンネル403は、1以上の空間パラメータを用いて2つの出力チャンネルを生成するOTTボックス406に対する入力として与えられる。図示の例において、2つの出力チャンネルは、例えば、サラウンドサウンド環境における前左（FL；Front Left）スピーカー位置と後左（BL；Backward Left）スピーカー位置を示す。チャンネル404は1以上の空間パラメータを用いて2つの出力チャンネルを生成するOTTボックス407に対する入力として与えられる。図示の例において、2つの出力チャンネルは前右（FR；Front Right）のスピーカー位置と後右（BR；Backward Right）のスピーカー位置を示す。チャンネル405は2つの出力チャンネルを生成するOTTボックス408に対する入力として与えられる。図示の例において、2つの出力チャンネルはセンター（C；Center）スピーカー位置と低周波拡張（LFE；Low Frequency Enhancement）チャンネルを示す。この場合、空間情報（例えば、CLD、ICCなど）はOTTボックスのそれぞれに対する入力として与えられる。一部の実施形態においては、レジデュアル（残留）信号Res1、Res2がOTTボックス406及び407に対する入力として与えられうる。この実施形態において、レジデュアル信号はセンターチャンネルとLFEチャンネルを出力するOTTボックス408に対する出力として与えられないことがある。

20

#### 【0029】

図4に示す構成は、チャンネル変換モジュール用の構成の一例である。OTTボックスとTTTボックスの種々の組み合わせを含むチャンネル変換モジュール用の他の構成も採用可能である。チャンネル変換モジュールのそれぞれは周波数エリアにおいて動作可能であるため、チャンネル変換モジュールのそれぞれに適用されるパラメータ帯域の個数が定義可能である。パラメータ帯域は一つのパラメータに適用可能な少なくとも一つの周波数帯域を意味している。パラメータ帯域の個数については図6Bに基づき説明する。

30

#### 【0030】

図5は、本発明の一実施形態によりオーディオ信号のビットストリームを構成する方法を示す図である。図5の（a）は、空間情報信号だけを含むオーディオ信号のビットストリームを示し、図5の（b）及び（c）は、ダウンミックス信号と空間情報信号を含むオーディオ信号のビットストリームを示す。

40

#### 【0031】

図5の（a）を参照すると、オーディオ信号のビットストリームは、構成情報501とフレーム503を含むことができる。フレーム503は、ビットストリームにおいて繰り返し可能であり、一部の実施形態においては、空間オーディオ情報を含む1枚の空間フレーム502を含む。

#### 【0032】

一部の実施形態において、構成情報501は1枚の空間フレーム502内におけるタイムスロットの総数と、オーディオ信号の周波数範囲を拡張するパラメータ帯域の総数と、OTTボックスにおけるパラメータ帯域の個数と、TTTボックスにおけるパラメータ帯

50

域の個数と、レジデュアル信号におけるパラメータ帯域の個数を示す情報を含む。構成情報 501 には所望に応じて他の情報が含まれる。

【0033】

一部の実施形態において、空間フレーム 502 は、1 以上の空間パラメータ（例えば、CLD、ICC など）と、フレームタイプと、1 枚のフレーム内におけるパラメータセットの個数とパラメータセットが適用可能なタイムスロットと、を含む。所望に応じて、空間フレーム 502 には他の情報が含まれる。図 6 ~ 図 10 に基づき、構成情報 501 及び空間フレーム 502 に含まれる情報の意味と用途を説明する。

【0034】

図 5 の (b) を参照すると、オーディオ信号のビットストリームは、構成情報 504 と、ダウンミックス信号 505 と、空間フレーム 506 と、を含む。この場合、1 枚のフレーム 507 は、ダウンミックス信号 505 と空間フレーム 506 を含み、これらのフレーム 507 がビットストリームにおいて繰り返し可能である。

【0035】

図 5 の (c) を参照すると、オーディオ信号のビットストリームは、ダウンミックス信号 508 と、構成情報 509 と、空間フレーム 510 と、を含む。この場合、1 枚のフレーム 511 は、構成情報 509 と空間フレーム 510 を含み、フレーム 511 は、ビットストリームにおいて繰り返し可能である。各フレーム 511 に構成情報 509 が挿入される場合、オーディオ信号は再生装置により任意の位置において再生可能である。

【0036】

図 5 の (c) は、構成情報 509 がフレーム 511 ごとにビットストリームに挿入されることを示しているが、周期的にまたは非周期的に繰り返される複数のフレームごとに構成情報 509 がビットストリームに挿入可能であるということはいうまでもない。

【0037】

図 6 A と図 6 B は、本発明の一実施形態によるパラメータセット、タイムスロット及びパラメータ帯域間の関係を示す図である。パラメータセットとは、1 個のタイムスロットに適用される 1 以上の空間パラメータのことを言う。空間パラメータは、CLD、ICC、CPC などの空間情報を含むことができる。タイムスロットとは、空間パラメータが適用可能なオーディオ信号の時間間隔のことをいう。1 枚の空間フレームは 1 以上のタイムスロットを含むことができる。

【0038】

図 6 A を参照すると、多数のパラメータセット 1、...、P が空間フレームに使用可能であり、各パラメータセットは 1 以上のデータフィールド 1、...、Q - 1 を含むことができる。オーディオ信号の全体の周波数範囲に一つのパラメータセットが適用可能であり、このようなパラメータセットにおける各空間パラメータは当該周波数帯域の 1 以上の位置に適用可能である。例えば、パラメータセットが 20 個の空間パラメータを含む場合、オーディオ信号の全体の周波数帯域は 20 個のエリア（以下、「パラメータ帯域」と言う。）に分割可能であり、パラメータセットの 20 個の空間パラメータがこれらの 20 個のパラメータ帯域に適用可能である。パラメータは、所望に応じて、パラメータ帯域に適用可能である。例えば、低周波パラメータ帯域に空間パラメータが密に適用され、高周波パラメータ帯域には疎らに適用可能である。

【0039】

図 6 B には、パラメータセットとタイムスロットとの間の関係を示す時間 / 周波数グラフが示されている。図示の例においては、1 枚の空間フレームに 12 個のタイムスロットに並べられたセットに 3 つのパラメータセット（パラメータセット 1、パラメータセット 2、パラメータセット 3）が適用される。この場合、オーディオ信号の全体の周波数範囲は 9 個のパラメータ帯域に分割される。このため、水平軸はタイムスロットの個数を示し、垂直軸はパラメータ帯域の個数を示す。3 つのパラメータセットのそれぞれが特定のタイムスロットに適用される。例えば、最初のパラメータセット（パラメータセット 1）はタイムスロット # 1 に適用され、2 番目のパラメータセット（パラメータセット 2）はタ

10

20

30

40

50

タイムスロット# 5に適用され、3番目のパラメータセット(パラメータセット3)はタイムスロット# 9に適用される。タイムスロットにこれらのパラメータセットを補間及び/またはコピーすることにより、残りのタイムスロットにもこれらのパラメータセットが適用可能である。一般的に、パラメータセットの個数はタイムスロットの個数以下であってもよく、パラメータ帯域の個数はオーディオ信号の周波数帯域の個数以下であってもよい。オーディオ信号の全体の時間-周波数エリアの代わりに、オーディオ信号の一部の時間-周波数エリアに関する空間情報をエンコーディングすることにより、エンコーダーからデコーダーに送られる空間情報の量を低減することができる。このようなデータの低減が可能になる理由は、公知のオーディオコーディング認識の原理によれば、時間-周波数エリアにおける空間情報は、ほとんどの場合、人間の聴覚認識に十分であるためである。

10

**【0040】**

開示された実施形態の重要な特徴は、パラメータセットが適用されるタイムスロットの位置を、固定個数のビットまたは可変個数のビットを用いてエンコーディングしかつデコーディングする、というところにある。また、パラメータ帯域の個数も固定個数のビットまたは可変個数のビットで表わすことができる。これに制限されるものではないが、空間オーディオコーディングに用いられる他の情報として、時間エリア、空間エリア及び/または周波数エリアに関連する情報を含む情報にも可変コーディング方式が適用可能である(例えば、フィルタバンクから出力される多数の周波数副帯域に適用される)。

**【0041】**

図7Aは、本発明の一実施形態による空間情報の構成情報を示すシンタックスを示している。このような構成情報は多数のビットが割当て可能な複数のフィールド701~718を含む。

20

**【0042】**

「bsSamplingFrequencyIndex」フィールド701は、オーディオ信号のサンプリング処理から取得されるサンプリング周波数を示す。サンプリング周波数を示すために、「bsSamplingFrequencyIndex」フィールド701には4ビットが割り当てられる。「bsSamplingFrequencyIndex」フィールド701の値が15、すなわち、2進数「1111」であれば、サンプリング周波数を示すために、「bsSamplingFrequency」フィールド702が追加される。この場合、「bsSamplingFrequency」フィールド702には24ビットが割り当てられる。

30

**【0043】**

「bsFrameLength」フィールド703は、1枚の空間フレーム内のタイムスロットの総数(以下、「numSlots」という。)を示し、「numSlots」と「bsFrameLength」フィールド703との間には、「numSlots = bsFrameLength + 1」という関係が成り立つ。

**【0044】**

「bsFreqRes」フィールド704は、オーディオ信号の全体の周波数エリアを拡張するパラメータ帯域の総数を示す。「bsFreqRes」フィールド704については図7Bに基づき説明する。

40

**【0045】**

「bsTreeConfig」フィールド705は、図4に基づき説明したように、複数のチャンネル変換モジュールを含むツリー構成のための情報を示す。このようなツリー構成のための情報は、チャンネル変換モジュールのタイプ、チャンネル変換モジュールの個数、チャンネル変換モジュールに用いられた空間情報のタイプ、オーディオ信号の入力/出力チャンネルの個数などの情報を含む。

**【0046】**

ツリー構成は、チャンネル変換モジュールのタイプまたはチャンネルの個数に応じて、5-1-5構成、5-2-5構成、7-2-7構成、7-5-7構成などのうちいずれかであってもよい。ツリー構成のうち5-2-5構成が図4に示してある。

50

## 【0047】

「bsQuantMode」フィールド706は、空間情報の量子化モード情報を示す。

## 【0048】

「bsOneIcc」フィールド707は、1つのICCパラメータサブセットが全体のOTTボックスに対して用いられるか否かを示す。この場合、パラメータサブセットは、特定のタイムスロット及び特定のチャンネル変換モジュールに適用されるパラメータセットを意味する。

## 【0049】

「bsArbitraryDownmix」フィールド708は、任意のダウンミックスゲインの存否を示す。

10

## 【0050】

「bsFixedGainS(登録商標)ur」フィールド709は、LS(左側のサラウンド)及びRS(右側のサラウンド)などのサラウンドチャンネルに適用されるゲインを示す。

## 【0051】

「bsFixedGainLFE」は、LFEチャンネルに適用されるゲインを示す。

## 【0052】

「bsFixedGainDM」は、ダウンミックス信号に適用されるゲインを示す。

## 【0053】

「bsMatrixMode」フィールド712は、ステレオダウンミックス信号と互換可能な行列がエンコーダーから生成されるか否かを示す。

20

## 【0054】

「bsTempShapeConfig」フィールド713は、デコーダーにおける臨時の形態(例えば、TES(Temporal Envelope Shaping)及び/またはTP(Temporal Shaping))の動作モードを示す。

## 【0055】

「bsDecorConfig」フィールド714は、デコーダーの相関分離器の動作モードを示す。

## 【0056】

最後に、「bs3DaudioMode」フィールド715は、ダウンミックス信号が3D信号にエンコーディングされるか否かと、逆HRTF処理が用いられるか否かを示す。

30

## 【0057】

エンコーダー/デコーダーにおいて各フィールドの情報が決定/抽出された後、チャンネル変換モジュールに適用されるパラメータ帯域の個数に関する情報がエンコーダー/デコーダーにおいて決定/抽出される。先ず、OTTボックスに適用されるパラメータ帯域の個数が決定/抽出され(716)てから、TTTボックスに適用されるパラメータ帯域の個数が決定/抽出される(717)。OTTボックス及び/またはTTTボックスに対するパラメータ帯域の個数は、以下、図8A~図9Bに基づき詳述する。

40

## 【0058】

拡張フレームが存在する場合、「spatialExtensionConfig」ブロック718は、拡張フレームに関する構成情報を含む。「spatialExtensionConfig」ブロック718に含まれている情報について、以下、図10A~図10Dに基づき説明する。

## 【0059】

図7Bは、本発明の一実施形態による空間情報信号のパラメータ帯域の個数を示す表である。「numBands」は、オーディオ信号の全体の周波数エリアに対するパラメータ帯域の個数を示し、「bsFreqRes」は、パラメータ帯域の個数に関するインデックス情報を示す。例えば、オーディオ信号の全体の周波数エリアは、所望に応じて、パ

50

ラメータ帯域の個数（例えば、4、5、7、10、14、20、28など）に分割可能である。

【0060】

一部の実施形態においては、各パラメータ帯域に一つのパラメータが適用可能である。例えば、「numBands」が28である場合、オーディオ信号の全体の周波数エリアは28個のパラメータ帯域に分割され、これらの28個のパラメータ帯域のそれぞれに28個のパラメータがそれぞれ適用可能である。他の例において、「numBands」が4である場合、所定のオーディオ信号の全体の周波数エリアは4個のパラメータ帯域に分割され、これらの4個のパラメータ帯域のそれぞれには4個のパラメータがそれぞれ適用可能である。図7Bにおいて、「Reserve」は、所定のオーディオ信号の全体の周波数エリアに対するパラメータ帯域の個数が決定されていないことを意味する。

10

【0061】

人間の聴覚機関は、コーディング方式において用いられるパラメータ帯域の個数に敏感ではないということに留意する必要がある。このため、少数のパラメータ帯域を用いてもより多数のパラメータ帯域が用いられた場合に比べて、聴取者に類似する空間オーディオ効果を奏することができる。

【0062】

「numBands」とは異なり、図7Aに示す「bsFrameLength」フィールド703が示す「numSlots」は、全体の値を示すことができる。しかしながら、1枚の空間フレーム内のサンプルの個数が「numSlots」により明確に分割される場合、「numSlots」の値は制限される。このため、実質的に表わされるべき「numSlots」の最大値が「b」であれば、「bsFrameLength」フィールド703の全体の値は $\text{ceil}\{\log_2(b)\}$ ビットで表わすことができる。この場合、「 $\text{ceil}(x)$ 」は、値「x」以上の最大の整数を意味する。例えば、1枚の空間フレームが72個のタイムスロットを含む場合、 $\text{ceil}\{\log_2(72)\} = 7$ ビットが「bsFrameLength」フィールド703に割当て可能であり、チャンネル変換モジュールに適用されるパラメータ帯域の個数は「numBands」内において決定可能である。

20

【0063】

図8Aは、本発明の一実施形態により、OTTボックスに適用されるパラメータ帯域の個数を固定個数のビットで表わすシンタックスを示している。図7Aと図8Aを参照すると、「i」は「0」において「numOttBoxes - 1」の値を有し、「numOttBoxes」はOTTボックスの総数である。すなわち、「i」の値が各OTTボックスを示し、各OTTボックスに適用されるパラメータ帯域の個数は「i」の値により表わされる。OTTボックスがLFEチャンネルモードを有する場合、OTTボックスのLFEチャンネルに適用される帯域の個数（以下、「bsOttBands」という。）は固定個数のビットを用いて表わすことができる。図示の例においては、「bsOttBands」フィールド801に5ビットが割り当てられる。OTTボックスがLFEチャンネルモードを有さない場合、総数のパラメータ帯域numBandsがOTTボックスのチャンネルに割り当てられる。

30

40

【0064】

図8Bは、本発明の一実施の形態により、OTTボックスに適用されるパラメータ帯域の個数を可変個数のビットで表わすシンタックスを示している。図8Bは、図8Aとほとんど同様であるが、図8Bに示す「bsOttBands」フィールド802が可変個数のビットで表わすという点で図8Aとは異なる。具体的に、「numBands」以下の値を有する「bsOttBands」フィールド802は、「numBands」を用いる可変個数のビットで表わすことができる。

【0065】

「numBands」が $2^{(n-1)}$ 以上 $2^n$ 未満の範囲に収まると、「bsOttBands」フィールド802は可変nビットで表わすことができる。

50

## 【0066】

例えば、(a)「numBands」が40である場合、「bsOttBands」フィールド802は6ビットで表わされ、(b)「numBands」が28または20である場合、「bsOttBands」フィールド802は5ビットで表わされ、(c)「numBands」が14または10である場合、「bsOttBands」フィールド802は4ビットで表わされ、(d)「numBands」が7、5または4である場合、「bsOttBands」フィールド802は3ビットで表わされる。

## 【0067】

「numBands」が $2^{(n-1)}$ より大きく $2^n$ 以下の範囲に収まると、「bsOttBands」フィールド802は可変nビットで表わすことができる。

10

## 【0068】

例えば、(a)「numBands」が40である場合、「bsOttBands」フィールド802は6ビットで表わされ、(b)「numBands」が28または20である場合、「bsOttBands」フィールド802は5ビットで表わされ、(c)「numBands」が14または10である場合、「bsOttBands」フィールド802は4ビットで表わされ、(d)「numBands」が7または5である場合、「bsOttBands」フィールド802は3ビットで表わされ、(e)「numBands」が4である場合、「bsOttBands」フィールド802は2ビットとなる。

## 【0069】

「bsOttBands」フィールド802は、「numBands」を変数として取って最も近い整数に切り上げる関数(以下、「切り上げ関数」と言う。)により可変個数のビットで表わすことができる。

20

## 【0070】

具体的に、i)  $0 < \text{bsOttBands} \leq \text{numBands}$  または  $0 < \text{bsOttBands} < \text{numBands}$  である場合、「bsOttBands」フィールド802は  $\text{ceil}(\log_2 \text{numBands})$  の値に対応する数のビットで表わされるか、あるいは、ii)  $0 < \text{bsOttBands} \leq \text{numBands} + 1$  である場合、「bsOttBands」フィールド802は  $\text{ceil}(\log_2(\text{numBands} + 1))$  で表わすことができる。

## 【0071】

「numBands」(以下、「numberBands」という。)以下の値が任意に決定される場合、「bsBands」フィールド802は、「numberBands」を変数として取って切り上げ関数により可変個数のビットで表わすことができる。

30

## 【0072】

具体的に、i)  $0 < \text{bsOttBands} \leq \text{numberBands}$  または  $0 < \text{bsOttBands} < \text{numberBands}$  である場合、「bsOttBands」フィールド802は  $\text{ceil}(\log_2(\text{numberBands}))$  ビットで表わされるか、あるいは、ii)  $0 < \text{bsOttBands} \leq \text{numberBands} + 1$  である場合、「bsOttBands」フィールド802は  $\text{ceil}(\log_2(\text{numberBands} + 1))$  で表わすことができる。

40

## 【0073】

1以上のOTTボックスが用いられる場合、「bsOttBands」の組み合わせは以下の式1で表わすことができる。

## 【0074】

〔式1〕

## 【数 1】

$$\sum_{i=1}^N \text{numBands}^{i-1} \cdot \text{bsOttBands}_i, \quad 0 \leq \text{bsOttBands}_i < \text{numBands},$$

## 【0075】

ここで、 $\text{bsOttBands}_i$  は  $i$  番目の「 $\text{bsOttBands}$ 」を示す。例えば、OTTボックスが3個存在し、「 $\text{bsOttBands}$ 」フィールド802に対して3つの値 ( $N=3$ ) が存在するとする。この例において、3つのOTTボックスに適用される「 $\text{bsOttBands}$ 」フィールド802の3つの値 (以下、それぞれ  $a_1$ 、 $a_2$ 、 $a_3$  とする。) はそれぞれ2ビットで表わすことができる。このため、 $a_1$ 、 $a_2$ 、 $a_3$  の値を表わすために、合計で6ビットが必要となる。しかしながら、 $a_1$ 、 $a_2$ 、 $a_3$  の値がグループで表わされる場合、27個 ( $= 3 * 3 * 3$ ) の場合が発生可能であり、これは5ビットで表わすことができ、1ビットを節約することになる。「 $\text{numBands}$ 」が3であり、かつ、5ビットで表わされるグループ値が15である場合、グループ値は  $15 = 1 * (3^2) + 2 * (3^1) + 0 * (3^0)$  で表わすことができる。このため、デコーダーは式1の逆を適用してグループ値15から「 $\text{bsOttBands}$ 」フィールド802の3つの値  $a_1$ 、 $a_2$ 、 $a_3$  をそれぞれ1、2、0に決定することができる。

10

20

## 【0076】

多数のOTTボックスの場合、「 $\text{bsOttBands}$ 」の組み合わせは「 $\text{numberBands}$ 」を用いる式2～式4 (以下、後述する。) のうちいずれかとして表わすことができる。「 $\text{numberBands}$ 」を用いる「 $\text{bsOttBands}$ 」の表現は式1における「 $\text{numBands}$ 」を用いる表現とほとんど同様であるため、詳細な説明は省き、その式だけを以下に示す。

## 【0077】

〔式2〕

## 【数 2】

30

$$\sum_{i=1}^N (\text{numberBands} + 1)^{i-1} \cdot \text{bsOttBands}_i, \quad 0 \leq \text{bsOttBands}_i \leq \text{numberBands},$$

## 【0078】

〔式3〕

## 【数 3】

40

$$\sum_{i=1}^N \text{numberBands}^{i-1} \cdot \text{bsOttBands}_i, \quad 0 \leq \text{bsOttBands}_i < \text{numberBands},$$

## 【0079】

〔式4〕

## 【数 4】

$$\sum_{i=1}^N \text{numberBands}^{i-1} \cdot \text{bsOttBands}_i, \quad 0 < \text{bsOttBands}_i \leq \text{numberBands},$$

## 【0080】

図9Aは、本発明の一実施形態により、TTTボックスに適用されるパラメータ帯域の個数を固定個数のビットで示すシンタックスを示している。図7Aと図9Aを参照すると、「i」は「0」において「numTttBoxes - 1」の値を有し、「numTttBoxes」はTTTボックスの総数である。すなわち、「i」の値が各TTTボックスを示す。各TTTボックスに適用されるパラメータ帯域の個数は「i」の値により表わされる。一部の実施形態において、TTTボックスは低周波帯域範囲と高周波帯域範囲とに分割可能であり、これらの低周波帯域範囲と高周波帯域範囲には異なる処理が適用可能である。異なる分割も行うことが可能である。

10

## 【0081】

「bsTTTDualMode」フィールド901は、所定のTTTボックスが低周波帯域範囲と高周波帯域範囲に対してそれぞれ異なるモード（以下、「デュアルモード」という。）にて動作するか否かを示す。例えば、「bsTTTDualMode」フィールド901の値が「0」である場合、低周波帯域範囲と高周波帯域範囲を区別することなく、全体の帯域範囲に対して単一のモードが用いられる。「bsTTTDualMode」フィールド901の値が「1」である場合、低周波帯域範囲と高周波帯域範囲に対してそれぞれ異なるモードが用いられる。

20

## 【0082】

「bsTttModeLow」フィールド902は、所定のTTTボックスの動作モードを示すものであり、これは種々の動作モードを有することができる。例えば、TTTボックスは、CPCパラメータとICCパラメータなどを用いる予測モード、CLDパラメータを用いるエネルギー基盤のモードなどを有することができる。TTTボックスがデュアルモードを有する場合、高周波帯域範囲に関する追加情報が必要になる。

30

## 【0083】

「bsTttModeHigh」フィールド903は、TTTボックスがデュアルモードを有する場合に高周波帯域範囲の動作モードを示す。

## 【0084】

「bsTttBandsLow」フィールド904は、TTTボックスに適用されるパラメータ帯域の個数を示す。

## 【0085】

「bsTttBandsHigh」フィールド905は、「numBands」を有する。

## 【0086】

TTTボックスがデュアルモードを有する場合、低帯域範囲は「0」以上「bsTttBandsLow」未満であり、高帯域範囲は「bsTttBandsLow」以上「bsTttBandsHigh」未満である。

40

## 【0087】

TTTボックスがデュアルモードを有さない場合、TTTボックスに適用されるパラメータ帯域の個数は「0」以上「numBands」未満である（907）。

## 【0088】

「bsTttBandsLow」フィールド904は、固定個数のビットで表わすことができる。例えば、図9Aに示すように、「bsTttBandsLow」フィールド904を表わすために5ビットが割当て可能である。

50

## 【0089】

図9Bは、本発明の一実施形態により、TTTボックスに適用されるパラメータ帯域の個数を可変個数のビットで表わすシンタックスを示している。図9Bは、図9Aとほとんど同様であるが、図9Bにおいては、「bsTttBandsLow」フィールド907を可変個数のビットで表わし、図9Aにおいては、「bsTttBandsLow」フィールド904を固定個数のビットで表わすという点において図9Aと図9Bとは異なっている。具体的に、「bsTttBandsLow」フィールド907は「numBands」以下の値を有するため、「bsTttBandsLow」フィールド907は「numBands」を用いる可変個数のビットで表わすことができる。

## 【0090】

具体的に、「numBands」が $2^{(n-1)}$ 以上 $2^n$ 未満の範囲に収まると、「bsTttBandsLow」フィールド907はnビットで表わすことができる。

## 【0091】

例えば、(i)「numBands」が40である場合、「bsTttBandsLow」フィールド907は6ビットで表わされ、(ii)「numBands」が28または20である場合、「bsTttBandsLow」フィールド907は5ビットで表わされ、(iii)「numBands」が14または10である場合、「bsTttBandsLow」フィールド907は4ビットで表わされ、(iv)「numBands」が7、5または4である場合、「bsTttBandsLow」フィールド907は3ビットで表わされる。

## 【0092】

「numBands」が $2^{(n-1)}$ より大きく $2^n$ 以下の範囲に収まると、「bsTttBandsLow」フィールド907は可変nビットで表わすことができる。

## 【0093】

例えば、(i)「numBands」が40である場合、「bsTttBandsLow」フィールド907は6ビットで表わされ、(ii)「numBands」が28または20である場合、「bsTttBandsLow」フィールド907は5ビットで表わされ、(iii)「numBands」が14または10である場合、「bsTttBandsLow」フィールド907は4ビットで表わされ、(iv)「numBands」が7または5である場合、「bsTttBandsLow」フィールド907は3ビットで表わされ、(v)「numBands」が4である場合、「bsOttBands」フィールド802は2ビットで表わされる。

## 【0094】

「bsTttBandsLow」フィールド907は、「numBands」を変数として取って最も近い整数に切り上げて決定される個数のビットで表わすことができる。

## 【0095】

例えば、i)  $0 < \text{bsOttBandsLow} \text{ numBands}$  または  $0 \text{ bsOttBandsLow} < \text{numBands}$  である場合、「bsTttBandsLow」フィールド907は  $\text{ceil}(\log_2(\text{numBands}))$  の値に対応する数のビットで表わされるか、あるいは、ii)  $0 \text{ bsOttBandsLow} \text{ numBands}$  である場合、「bsTttBandsLow」フィールド907は  $\text{ceil}(\log_2(\text{numBands} + 1))$  で表わすことができる。

## 【0096】

「numBands」、すなわち、「numberBands」以下の値が任意に決定される場合、「bsTttBandsLow」フィールド907は、「numberBands」を用いて可変個数のビットで表わすことができる。

## 【0097】

具体的に、i)  $0 < \text{bsOttBandsLow} \text{ numberBands}$  または 0

10

20

30

40

50

$bsOttBandsLow < numberBands$ である場合、「 $bsTttBandsLow$ 」フィールド907は $ceil(\log_2(numberBands))$ の値に対応する数のビットで表わされるか、あるいは、 $ii) 0 \leq bsOttBandsLow < numberBands$ である場合、「 $bsTttBandsLow$ 」フィールド907は $ceil(\log_2(numberBands + 1))$ の値に対応する数のビットで表わすことができる。

【0098】

多数のOTTボックスが用いられる場合、「 $bsOttBandsLow$ 」の組み合わせは以下の式5で表わすことができる。

【0099】

〔式5〕

【数5】

$$\sum_{i=1}^N numBands^{i-1} \cdot bsTttBandsLow_i, \quad 0 \leq bsTttBandsLow_i < numBands,$$

【0100】

ここで、 $bsTttBandsLow_i$ はi番目の「 $bsTttBandsLow$ 」を示す。式5の意味は式1の意味と同様であるため、式5の詳細な説明は省く。

多数のOTTボックスの場合、「 $bsOttBandsLow$ 」の組み合わせは「 $numberBands$ 」を用いる式6～式8のうちいずれかとして表わすことができる。式6～式8の意味は式2～式4の意味と同様であるため、式6～式8の詳細な説明は省く。

【0101】

〔式6〕

【数6】

$$\sum_{i=1}^N (numberBands + 1)^{i-1} \cdot bsTttBandsLow_i, \quad 0 \leq bsTttBandsLow_i \leq numberBands,$$

【0102】

〔式7〕

【数7】

$$\sum_{i=1}^N numberBands^{i-1} \cdot bsTttBandsLow_i, \quad 0 \leq bsTttBandsLow_i < numberBands,$$

【0103】

〔式8〕

10

20

30

40

【数 8】

$$\sum_{i=1}^N \text{numberBands}^{i-1} \cdot \text{bsTttBandsLow}_i, \quad 0 < \text{bsTttBandsLow}_i \leq \text{numberBands},$$

【0104】

チャンネル変換モジュール（例えば、OTTボックス及び/またはTTTボックスなど）に適用されるパラメータ帯域の個数は「numBands」の分割値として表わすことができる。この場合、上述した分割値は、「numBands」の半値または特定の値で「numBands」を割った結果の値を用いる。

10

【0105】

OTT及び/またはTTボックスに適用されるパラメータ帯域の個数が一旦決定されると、前記パラメータ帯域の個数の範囲内において各OTTボックス及び/または各TTTボックスに適用可能なパラメータセットが決定される。各OTTボックス及び/または各TTTボックスには各パラメータセットがタイムスロット単位で適用可能である。すなわち、1個のタイムスロットに一つのパラメータセットが適用可能である。

【0106】

以上の述べたように、1枚の空間フレームは複数のタイムスロットを含むことができる。空間フレームが固定フレームタイプである場合、複数のタイムスロットには等間隔にてパラメータセットが適用可能である。空間フレームが可変フレームタイプである場合、パラメータが適用されるタイムスロットの位置情報が必要となる。これについては、図13A~13Cに基づき後述する。

20

【0107】

図10Aは、本発明の一実施形態により空間拡張フレームに関する空間拡張構成情報を示すシンタックスを示している。空間拡張構成情報は、「bsSacExtType」フィールド1001と、「bsSacExtLen」フィールド1002と、「bsSacExtLenAdd」フィールド1003と、「bsSacExtLenAddAdd」フィールド1004と、「bsFillBits」フィールド1007と、を含む。他の

30

【0108】

「bsSacExtType」フィールド1001は、空間拡張フレームのデータタイプを示す。例えば、空間拡張フレームは、「0」、レジデュアル信号データ、任意のダウンミックスレジデュアル信号データまたは任意のツリーデータで詰め込むことができる。

【0109】

「bsSacExtLen」フィールド1002は、空間拡張構成情報のバイト数を示す。

【0110】

「bsSacExtLenAdd」フィールド1003は、空間拡張構成情報のバイト数が、例えば15以上である場合、空間拡張構成情報の追加バイト数を示す。

40

【0111】

「bsSacLenAddAdd」フィールド1004は、空間拡張構成情報のバイト数が、例えば270以上である場合、空間拡張構成情報の追加バイト数を示す。

【0112】

エンコーダー/デコーダーにおいて各フィールドが決定/抽出された後、空間拡張フレームに含まれるデータタイプに関する構成情報が決定される(1005)。

【0113】

上述したように、空間拡張フレームにはレジデュアル信号データ、任意のダウンミックスレジデュアル信号データ、ツリー構成データなどが含まれる。

50

## 【0114】

続けて、空間拡張構成情報の長さのうち未使用のビットの個数が算出される(1006)。

## 【0115】

「bsFillBits」フィールド1007は、未使用のビットを詰め込むために見逃しうるデータのビット数を示す。

## 【0116】

図10Bと図10Cは、本発明の一実施形態により、空間拡張フレームにレジデュアル信号が含まれる場合、レジデュアル信号のための空間拡張情報を示すシンタックスを示している。

10

## 【0117】

図10Bを参照すると、「bsResidualSamplingFrequencyIndex」フィールド1008は、レジデュアル信号のサンプリング周波数を示す。

## 【0118】

「bsResidualFramesPerSpatialFrame」フィールド1009は、空間フレーム当たりのレジデュアルフレームの本数を示す。例えば、1枚の空間フレームに1枚、2枚、3枚または4枚のレジデュアルフレームが含まれる。

## 【0119】

「ResidualConfig」フィールド1010は、各OTT及び/またはTTTボックスに適用されるレジデュアル信号に対するパラメータ帯域の個数を示す。

20

## 【0120】

図10Cを参照すると、「bsResidualPresent」フィールド1011は、各OTT及び/またはTTTボックスにレジデュアル信号が適用されるか否かを示す。

## 【0121】

「bsResidualBands」フィールド1012は、各OTT及び/またはTTTボックスにレジデュアル信号が存在する場合、各OTT及び/またはTTTボックスに存在するレジデュアル信号のパラメータ帯域の個数を示す。レジデュアル信号のパラメータ帯域の個数は固定個数のビットまたは可変個数のビットで表わすことができる。パラメータ帯域の個数が固定個数のビットで表わされる場合、レジデュアル信号はオーディオ信号のパラメータ帯域の総数以下の値を有することができる。このため、全体のパラメータ帯域の個数を示すために必要となるビット数(例えば、図10Cにおける5ビットなど)が割当て可能である。

30

## 【0122】

図10Dは、本発明の一実施形態により、レジデュアル信号のパラメータ帯域の個数を可変個数のビットで表わすシンタックスを示している。「beResidualBands」フィールド1014は、「numBands」を用いる可変個数のビットで表わすことができる。

## 【0123】

「numBands」が $2^{(n-1)}$ 以上 $2^n$ 未満であれば、「beResidualBands」フィールド1014はnビットで表わすことができる。

40

## 【0124】

例えば、(i)「numBands」が40である場合、「beResidualBands」フィールド1014は6ビットで表わされ、(ii)「numBands」が28または20である場合、「beResidualBands」フィールド1014は5ビットで表わされ、(iii)「numBands」が14または10である場合、「beResidualBands」フィールド1014は4ビットで表わされ、(iv)「numBands」が7、5または4である場合、「beResidualBands」フィールド1014は3ビットで表わされる。

## 【0125】

50

「numBands」が $2^{(n-1)}$ より大きく $2^n$ 以下であれば、「beResidualBands」フィールド1014は可変nビットで表わすことができる。

【0126】

例えば、(i)「numBands」が40である場合、「beResidualBands」フィールド1014は6ビットで表わされ、(ii)「numBands」が28または20である場合、「beResidualBands」フィールド1014は5ビットで表わされ、(iii)「numBands」が14または10である場合、「beResidualBands」フィールド1014は4ビットで表わされ、(iv)「numBands」が7または5である場合、「beResidualBands」フィールド1014は3ビットで表わされ、(v)「numBands」が4である場合、「beResidualBands」フィールド1014は2ビットで表わされる。

10

【0127】

また、「beResidualBands」フィールド1014は、「numBands」を変数として取って最も近い整数に切り上げて決定される切り上げ関数により決定されるビット数で表わすことができる。

【0128】

具体的に、i)  $0 < \text{beResidualBands} \leq \text{numBands}$  または  $0 < \text{beResidualBands} < \text{numBands}$  である場合、「beResidualBands」フィールド1014は  $\text{ceil}(\log_2(\text{numBands}))$  ビットで表わされるか、あるいは、ii)  $0 < \text{beResidualBands} > \text{numBands}$  である場合、「beResidualBands」フィールド1014は  $\text{ceil}(\log_2(\text{numBands} + 1))$  ビットで表わすことができる。

20

【0129】

一部の実施形態においては、「numBands」以下の値「numberBands」を用いて「beResidualBands」フィールド1014を表わすことができる。

【0130】

具体的に、i)  $0 < \text{beResidualBands} \leq \text{numberBands}$  または  $0 < \text{beResidualBands} < \text{numberBands}$  である場合、「beResidualBands」フィールド1014は  $\text{ceil}(\log_2(\text{numberBands}))$  ビットで表わされるか、あるいは、ii)  $0 < \text{beResidualBands} > \text{numberBands}$  である場合、「beResidualBands」フィールド1014は  $\text{ceil}(\log_2(\text{numberBands} + 1))$  の値で表わすことができる。

30

【0131】

複数のレジデュアル信号Nが存在する場合、「beResidualBands」の組み合わせは、以下の式9で表わすことができる。

【0132】

〔式9〕

【数9】

40

$$\sum_{i=1}^N \text{numBands}^{i-1} \cdot \text{bsResidualBands}_i, \quad 0 \leq \text{bsResidualBands}_i < \text{numBands}$$

【0133】

この場合、 $\text{bsResidualBands}_i$  はi番目の「bsResidualBands」を示す。式9の意味は、式1の意味と同様であるため、式9の詳細な説明は省く。

50

## 【 0 1 3 4 】

多数のレジデュアル信号が存在する場合、「bsResidualBands」の組み合わせは「numberBands」を用いる式10～式12のうちいずれかとして表わすことができる。「numberBands」を用いて「bsResidualBands」を表わすことは式2～式4とほとんど同様であるため、詳細な説明を省く。

## 【 0 1 3 5 】

〔式10〕

## 【数10】

$$\sum_{i=1}^N (\text{numberBands} + 1)^{i-1} \cdot \text{bsResidualBands}_i, \quad 0 \leq \text{bsResidualBands}_i \leq \text{numberBands},$$

10

## 【 0 1 3 6 】

〔式11〕

## 【数11】

$$\sum_{i=1}^N \text{numberBands}^{i-1} \cdot \text{bsResidualBands}_i, \quad 0 \leq \text{bsResidualBands}_i < \text{numberBands},$$

20

## 【 0 1 3 7 】

〔式12〕

## 【数12】

$$\sum_{i=1}^N \text{numberBands}^{i-1} \cdot \text{bsResidualBands}_i, \quad 0 < \text{bsResidualBands}_i \leq \text{numberBands},$$

30

## 【 0 1 3 8 】

レジデュアル信号のパラメータ帯域の個数は「numBands」の分割値として表わすことができる。この場合、上述した分割値は、「numBands」の半値または特定の値で「numBands」を割った結果の値を用いる。

## 【 0 1 3 9 】

レジデュアル信号はダウンミックス信号及び空間情報信号と共にオーディオ信号のビットストリームに含まれてもよく、このようなビットストリームはデコーダーに転送可能である。デコーダーは、このようなビットストリームから前記ダウンミックス信号と、空間情報信号及びレジデュアル信号を抽出することができる。

40

## 【 0 1 4 0 】

続けて、ダウンミックス信号は空間情報を用いてアップミックスされる。一方、レジデュアル信号はアップミックスの過程においてダウンミックス信号に適用される。具体的に、ダウンミックス信号は空間情報を用いる複数のチャンネル変換モジュールにおいてアップミックスされる。このような過程において、レジデュアル信号がチャンネル変換モジュールに適用される。以上述べたように、チャンネル変換モジュールは複数のパラメータ帯域を有し、パラメータセットはタイムスロット単位でチャンネル変換モジュールに適用される。レジデュアル信号がチャンネル変換モジュールに適用される場合、レジデュアル信号が適用されるオーディオ信号のチャンネル間関連情報を更新するためにはレジデュアル

50

信号が必要となる。このような更新されたチャンネル間相関情報はアップミキシング処理に用いられる。

【0141】

図11Aは、本発明の一実施形態によるノンガイドコーディングのためのデコーダを示すブロック図である。ノンガイドコーディングは、オーディオ信号のビットストリームに空間情報が含まれていないことを意味する。

【0142】

一部の実施形態において、デコーダは、解析フィルタバンク1102と、解析部1104と、空間合成部1006と、合成フィルタバンク1108と、を備える。図11Aにはステレオ信号タイプのダウンミックス信号が示されているが、他のタイプのダウンミックス信号が使用可能である。

10

【0143】

動作に際し、デコーダはダウンミックス信号1101を受信し、解析フィルタバンク1102は前記受信されたダウンミックス信号1101を周波数エリア信号1103に変換する。解析部1104は、前記変換されたダウンミックス信号1103から空間情報を生成する。解析部1104がスロット単位で処理を行い、複数のスロットごとに空間情報1105を生成することができる。この場合、スロットはタイムスロットを含む。

【0144】

空間情報は2ステップで生成可能である。第一に、ダウンミックス信号からダウンミックスパラメータが生成される。第二に、前記ダウンミックスパラメータは空間パラメータなどの空間情報に変換される。一部の実施形態において、ダウンミックスパラメータはダウンミックス信号の行列演算により生成可能である。

20

【0145】

空間合成部1106は、前記生成された空間情報1105とダウンミックス信号1103とを合成してマルチチャンネルオーディオ信号1107を生成する。前記生成されたマルチチャンネルオーディオ信号1107は、合成フィルタバンク1108を通過して時間エリアオーディオ信号1109に変換される。

【0146】

空間情報は所定のスロット位置に生成可能である。このような位置間距離は同じであってもよい(すなわち、等距離)。例えば、空間情報は4個のスロットごとに生成可能である。また、空間情報は可変スロット位置に生成可能である。この場合、空間情報が生成される位置情報がビットストリームから抽出可能である。前記位置情報は可変個数のビットで表わすことができる。前記位置は以前のスロット位置情報からの絶対値及び差分値として表わすことができる。

30

【0147】

ノンガイドコーディングを用いる場合、オーディオ信号の各チャンネルに対するパラメータ帯域の個数(以下、「bsNumGuidedBlindBands」という。)は、固定個数のビットで表わすことができる。「bsNumGuidedBlindBands」は、「numBands」を用いる可変個数のビットで表わすことができる。例えば、「numBands」が $2^{(n-1)}$ 以上 $2^n$ 未満であれば、「bsNumGuidedBlindBands」は可変nビットで表わすことができる。

40

【0148】

具体的に、(a)「numBands」が40である場合、「bsNumGuidedBlindBands」は6ビットで表わされ、(b)「numBands」が28または20である場合、「bsNumGuidedBlindBands」は5ビットで表わされ、(c)「numBands」が14または10である場合、「bsNumGuidedBlindBands」は4ビットで表わされ、(d)「numBands」が7、5または4である場合、「bsNumGuidedBlindBands」は3ビットで表わされる。

【0149】

50

「numBands」が $2^{(n-1)}$ より大きく $2^n$ 以下であれば、「bsNumGuidedBlindBands」は可変nビットで表わすことができる。

【0150】

例えば、(a)「numBands」が40である場合、「bsNumGuidedBlindBands」は6ビットで表わされ、(b)「numBands」が28または20である場合、「bsNumGuidedBlindBands」は5ビットで表わされ、(c)「numBands」が14または10である場合、「bsNumGuidedBlindBands」は4ビットで表わされ、(d)「numBands」が7または5である場合、「bsNumGuidedBlindBands」は3ビットで表わされ、(e)「numBands」が4である場合、「bsNumGuidedBlindBands」は2ビットで表わされる。

10

【0151】

また、「bsNumGuidedBlindBands」は「numBands」を変数として取る切り上げ関数を用いて可変個数のビットで表わすことができる。

【0152】

例えば、i)  $0 < \text{bsNumGuidedBlindBands} \leq \text{numBands}$  または  $0 < \text{bsNumGuidedBlindBands} < \text{numBands}$  である場合、「bsNumGuidedBlindBands」は $\text{ceil}(\log_2(\text{numBands}))$ ビットで表わされるか、あるいは、ii)  $0 < \text{bsNumGuidedBlindBands} < \text{numBands}$  である場合、「bsNumGuidedBlindBands」は $\text{ceil}(\log_2(\text{numBands} + 1))$ ビットで表わすことができる。

20

【0153】

「numBands」以下の値、すなわち、「numberBands」が任意に決定される場合、「bsNumGuidedBlindBands」は以下のように表わすことができる。

【0154】

具体的に、i)  $0 < \text{bsNumGuidedBlindBands} \leq \text{numberBands}$  または  $0 < \text{bsNumGuidedBlindBands} < \text{numberBands}$  である場合、「bsNumGuidedBlindBands」は $\text{ceil}(\log_2(\text{numberBands}))$ ビットで表わされるか、あるいは、ii)  $0 < \text{bsNumGuidedBlindBands} < \text{numberBands}$  である場合、「bsNumGuidedBlindBands」は $\text{ceil}(\log_2(\text{numberBands} + 1))$ で表わすことができる。

30

【0155】

多数のチャンネルNが用いられる場合、「bsNumGuidedBlindBands」の組み合わせは以下の式13で表わすことができる。

【0156】

〔式13〕

【数13】

40

$$\sum_{i=1}^N \text{numBands}^{i-1} \cdot \text{bsNumGuidedBlindBands}_i, \quad 0 \leq \text{bsNumGuidedBlindBands}_i < \text{numBands},$$

【0157】

ここで、「bsNumGuidedBlindBands<sub>i</sub>」はi番目の「bsNumGuidedBlindBands」を示す。式13の意味は式1の意味と同様であるため、式13の詳細な説明は省く。

【0158】

50

多数のチャンネルが存在する場合、「bsNumGuidedBlindBands」は「numberBands」を用いる式14～式16のうちいずれかとして表わすことができる。「numberBands」を用いる「bsNumGuidedBlindBands」の表現は式2～式4の表現と同様であるため、式14～式16の詳細な説明は省く。

【0159】

〔式14〕

【数14】

$$\sum_{i=1}^N (\text{numberBands} + 1)^{i-1} \cdot \text{bsNumGuidedBlindBands}_i, 0 \leq \text{bsNumGuidedBlindBands}_i \leq \text{numberBands},$$

10

【0160】

〔式15〕

【数15】

$$\sum_{i=1}^N \text{numberBands}^{i-1} \cdot \text{bsNumGuidedBlindBands}_i, 0 \leq \text{bsNumGuidedBlindBands}_i < (\text{numberBands} + 1),$$

20

【0161】

〔式16〕

【数16】

$$\sum_{i=1}^N \text{numberBands}^{i-1} \cdot \text{bsNumGuidedBlindBands}_i, 0 < \text{bsNumGuidedBlindBands}_i \leq \text{numberBands},$$

30

【0162】

図11Bは、本発明の一実施形態によりパラメータ帯域の個数をグループとして表わす方法を示している。パラメータ帯域の個数は、チャンネル変換モジュールに適用されるパラメータ帯域の個数情報と、レジデュアル信号に適用されるパラメータ帯域の個数情報と、ノンガイドコーディングを用いる場合にオーディオ信号の各チャンネルに関するパラメータ帯域の個数情報と、を含む。パラメータ帯域の個数情報が複数存在する場合、複数の個数情報（例えば、「bsOttBands」、「bsTttBands」、「bsResidualBands」及び/または「bsNumGuidedBlindBands」など）は少なくとも一つのグループとして表わすことができる。

40

【0163】

図11Bを参照すると、パラメータ帯域の個数情報が $(kN + L)$ 個存在し、かつ、各パラメータ帯域の個数情報を表わすのに $Q$ ビットが必要となる場合、複数のパラメータ帯域の個数情報は以下のグループで表わすことができる。この場合、「 $k$ 」と「 $N$ 」は「0」ではない任意の整数であり、「 $L$ 」は $0 < L < N$ を満たす任意の整数である。

【0164】

グループ化方法は、パラメータ帯域の個数情報 $N$ 個を束ねて $k$ 個のグループを生成するステップと、最後のパラメータ帯域の個数情報 $L$ 個を束ねて最終のグループを生成するステップと、を含む。 $K$ 個のグループは $M$ ビットで表わすことができ、最終のグループは $p$ ビットで表わすことができる。この場合、 $M$ ビットが、好ましくは、パラメータ帯域の個

50

数情報のそれぞれをグループ化することなく表わす場合に用いられる  $N * Q$  ビットよりも小さい。P ビットが、好ましくは、パラメータ帯域の個数情報のそれぞれをグループ化することなく表わす場合に用いられる  $L * Q$  ビット以下である。

**【 0 1 6 5 】**

例えば、パラメータ帯域の個数情報の2つがそれぞれ  $b_1$  と  $b_2$  であるとする。 $b_1$  と  $b_2$  がそれぞれ5個の値を有する場合、 $b_1$  と  $b_2$  のそれぞれを表わすのに3ビットが必要である。この場合、たとえ3ビットは8個の値を表わすことができるとしても、実質的には5個の値が必要となる。このため、 $b_1$  と  $b_2$  のそれぞれは3個の余分を有する。しかしながら、 $b_1$  と  $b_2$  を束ねてグループとして表わす場合には、6ビット (= 3ビット + 3ビット) の代わりに5ビットが用いられる。具体的に、 $b_1$  と  $b_2$  との全ての組み合わせは25個 (=  $5 * 5$ ) のタイプを有するため、 $b_1$  と  $b_2$  のグループは5ビットで表わすことができる。5ビットは32個の値を表わすことができるため、グループ化表現の場合、7個の余分が生成される。しかしながら、 $b_1$  と  $b_2$  をグループ化して表わす場合、その余分は  $b_1$  と  $b_2$  をそれぞれ3ビットで表わす場合の余分よりも小さい。複数のパラメータ帯域の個数情報をグループとして表わす方法は、以下のような種々の方式により実現可能である。

10

**【 0 1 6 6 】**

複数のパラメータ帯域の個数情報がそれぞれ40種類の値を有する場合、Nとして2、3、4、5または6を用いてk個のグループが生成される。これらのk個のグループはそれぞれ11ビット、16ビット、22ビット、27ビット、32ビットとして表わすことができる。あるいは、これらのk個のグループは各場合を組み合わせる。

20

**【 0 1 6 7 】**

複数のパラメータ帯域の個数情報がそれぞれ28種類の値を有する場合、Nとして6を用いてk個のグループが生成され、kは29ビットで表わすことができる。

**【 0 1 6 8 】**

複数のパラメータ帯域の個数情報がそれぞれ20種類の値を有する場合、Nとして2、3、4、5、6または7を用いてk個のグループが生成される。これらのk個のグループはそれぞれ9ビット、13ビット、18ビット、22ビット、26ビット及び31ビットとして表わすことができる。あるいは、これらのk個のグループは各場合を組み合わせる。

30

**【 0 1 6 9 】**

複数のパラメータ帯域の個数情報がそれぞれ14種類の値を有する場合、Nとして6を用いてk個のグループが生成される。これらのk個のグループは23ビットで表わすことができる。

**【 0 1 7 0 】**

複数のパラメータ帯域の個数情報がそれぞれ10種類の値を有する場合、Nとして2、3、4、5、6、7、8または9を用いてk個のグループが生成される。これらのk個のグループはそれぞれ7、10、14、17、20、24、27及び30ビットで表わすことができる。あるいは、これらのk個のグループは各場合を組み合わせる。

40

**【 0 1 7 1 】**

複数のパラメータ帯域の個数情報がそれぞれ7種類の値を有する場合、Nとして6、7、8、9、10または11を用いてk個のグループが生成される。これらのk個のグループはそれぞれ17、20、23、26、29及び31ビットで表わすことができる。あるいは、これらのk個のグループは各場合を組み合わせる。

**【 0 1 7 2 】**

複数のパラメータ帯域の個数情報がそれぞれ5種類の値を有する場合、Nとして2、3、4、5、6、7、8、9、10、11、12または13を用いてk個のグループが生成可能である。これらのk個のグループはそれぞれ5、7、10、12、13、17、19、21、24、26、28及び31ビットで表わすことができる。あるいは、これらのk

50

個のグループは各場合を組み合わせることで表わすことができる。

【0173】

また、複数のパラメータ帯域の個数情報は、上述したグループとして表わされるように構成可能であるか、または、パラメータ帯域の個数情報のそれぞれを独立したビットシーケンスとして連続して表わされるように構成可能である。

【0174】

図12は、本発明の一実施形態により空間フレームの構成情報を示すシンタックスを示している。空間フレームは、「FramingInfo」ブロック1201と、「bsIndependenceFlag」ブロック1201と、「OttData」ブロック1203と、「TttData」ブロック1204と、「SmgData」ブロック1205と、「TempShapeData」ブロック1206と、を含む。

10

【0175】

「FramingInfo」ブロック1201は、パラメータセットの個数に関する情報と、各パラメータが適用されるタイムスロットに関する情報と、を含む。「FramingInfo」ブロック1201については、図13Aに基づき詳述する。

【0176】

「bsIndependenceFlag」フィールド1202は、現在のフレームが以前フレームに関する知識無しにデコーディング可能であるか否かを示す。

【0177】

「OttData」ブロック1203は、全体のOTTボックスに関する全体の空間パラメータ情報を含む。

20

【0178】

「TttData」ブロック1204は、全体のTTTボックスに関する全体の空間パラメータ情報を含む。

【0179】

「SmgData」ブロック1205は、非量子化された空間パラメータに適用される臨時平坦化に関する情報を含む。

【0180】

「TempShapeData」ブロック1206は、非相関信号に適用される臨時エンベロープ形状化に関する情報を含む。

30

【0181】

図13Aは、本発明の一実施形態により、パラメータセットが適用されるタイムスロット位置情報を示すシンタックスを示している。「bsFramingType」フィールド1301はオーディオ信号の空間フレームが固定フレームタイプであるか、あるいは、可変フレームタイプであることを示す。固定フレームとは、予め設定されたタイムスロットにパラメータセットが適用されるフレームのことを言う。例えば、等間隔にて予め設定されたタイムスロットにパラメータセットが適用される。可変フレームとは、パラメータセットが適用されるタイムスロットの位置情報を別途に受信するフレームのことを言う。

【0182】

「bsNumParamSets」フィールド1302は、1枚の空間フレーム内においてパラメータセットの個数を示し(以下、「numParamSets」という。)、 $\text{numParamSets}$ と「bsNumParamSets」との間には $\text{numParamSets} = \text{bsNumParamSets} + 1$ の関係が成り立つ。

40

【0183】

例えば、図13Aの「bsNumParamSets」フィールド1302に3ビットが割り当てられると、1枚の空間フレーム内には最大8個のパラメータセットが提供可能である。割り当てられるビットの個数については制限がないため、空間フレーム内により多くのパラメータセットが提供可能である。

【0184】

空間フレームが固定フレームタイプである場合、パラメータセットが適用されるタイム

50

スロットの位置情報は予め設定された規則により決定可能であり、パラメータセットが適用されるタイムスロットの追加的な位置情報は不要である。しかしながら、空間フレームが可変フレームタイプである場合、パラメータセットが適用される位置情報が必要となる。

【0185】

「bsParamSlot」フィールド1303は、パラメータセットが適用されるタイムスロットの位置情報を示す。「bsParamSlot」フィールド1303は、1枚の空間フレーム内におけるタイムスロットの個数、すなわち、「numSlots」を用いて可変個数のビットで表わすことができる。具体的に、「numSlots」が $2^{(n-1)}$ 以上 $2^n$ 未満の範囲に収まると、「bsParamSlot」フィールド1303はnビットで表わすことができる。

10

【0186】

例えば、(i)「numSlots」が64と127との間の範囲にある場合、「bsParamSlot」フィールド1303は7ビットで表わされ、(ii)「numSlots」が32と63との間の範囲にある場合、「bsParamSlot」フィールド1303は6ビットで表わされ、(iii)「numSlots」が16と31との間の範囲にある場合、「bsParamSlot」フィールド1303は5ビットで表わされ、(iv)「numSlots」が8と15との間の範囲にある場合、「bsParamSlot」フィールド1303は4ビットで表わされ、(v)「numSlots」が4と7との間の範囲にある場合、「bsParamSlot」フィールド1303は3ビットで表わされ、(vi)「numSlots」が2と3との間の範囲にある場合、「bsParamSlot」フィールド1303は2ビットで表わされ、(vii)「numSlots」が1である場合、「bsParamSlot」フィールド1303は1ビットで表わされ、(viii)「numSlots」が0である場合、「bsParamSlot」フィールド1303は0ビットで表わすことができる。これと同様に、「numSlots」が64と127との間の範囲にある場合、「bsParamSlot」フィールド1303は7ビットで表わすことができる。

20

【0187】

多数のパラメータセットNが存在する場合、「bsParamSlot」は式17により表わすことができる。

30

【0188】

〔式17〕

【数17】

$$\sum_{i=1}^N \text{numSlots}^{i-1} \cdot \text{bsParamSlot}_i, \quad 0 \leq \text{bsParamSlot}_i < \text{numSlots},$$

【0189】

この場合、「bsParamSlot<sub>i</sub>」はI番目のパラメータセットが適用されるタイムスロットを示す。例えば、「numSlots」が3であり、かつ、「bsParamSlot」フィールド1303が10個の値を有することができるとする。この場合、「bsParamSlot」フィールド1303に関する3つの情報(以下、それぞれ、c1、c2、c3という。)が必要となる。c1、c2、c3のそれぞれを表わすためには4ビットが必要であるため、合計で12ビットが必要となる。c1、c2、c3をグループに束ねて表わす場合、1、000個(=10\*10\*10)の場合が発生可能であり、これは10ビットで表わされ、2ビットを節約することになる。「numSlots」が3であり、かつ、5ビットで表わされるグループ値が31である場合、グループ値は $31 = 1 \times (3^2) + 5 \times (3^1) + 7 \times (3^0)$ で表わすことができる。このため

40

50

、デコーダ装置は式17の逆を適用してc1、c2、c3をそれぞれ1、5、7に決定することができる。

【0190】

図13Bは、本発明の一実施形態によりパラメータセットが絶対値及び差分値として適用されるタイムスロットの位置情報を示すシンタックスを示している。空間フレームが可変フレームタイプである場合、図13Aにおける「bsParamSlot」フィールド1301は「bsParamSlot」情報が単調増加するということを用いて絶対値及び差分値で表わすことができる。

【0191】

例えば、(i)最初のパラメータセットが適用されるタイムスロットの位置は絶対値、すなわち、「bsParamSlot[0]」として生成可能であり、(ii)2番目以上のパラメータセットが適用されるタイムスロットの位置は絶対値、すなわち、「bsParamSlot[ps]」と「bsParamSlot[ps-1]」との間の「difference value」または「difference value-1」として生成可能である(以下、「bsDiffParamSlot[ps]」という。)。この場合、「ps」はパラメータセットを意味する。

10

【0192】

「bsParamSlot[0]」フィールド1304は「numSlots」と「numParamSets」を用いて算出される個数のビットで表わすことができる(以下、「nBitsParamSlot(0)」という。)

20

【0193】

「bsDiffParamSlot[ps]」フィールド1305は、「numSlots」、「numParamSetst」及び以前のパラメータセットが適用されたタイムスロットの位置を用いて算出される個数のビットで表わすことができる(以下、「nBitParamSlot(ps)」という。)

【0194】

具体的に、最小数のビットで「bsParamSlot[ps]」を表わすために、「bsParamSlot[ps]」を表わすビットの個数は以下の規則により決定可能である：(i)複数の「bsParamSlot[ps]」は昇順に増加する( $bsParamSlot[ps] > bsParamSlot[ps-1]$ )、(ii)「bsParamSlot[0]」の最大値は「numSlots - NumParamSets」であり、(iii)  $0 < ps < numParamSets$ である場合、「bsParamSlot[ps]」は「bsParamSlot[ps-1] + 1」と「numSlots - numParamSets + ps」との間の値だけを有する。

30

【0195】

例えば、「numSlots」が10であり、かつ、「numParamSets」が3である場合、「bsParamSlot[ps]」は昇順に増加するため、「bsParamSlot[0]」の最大値は「 $10 - 3 = 7$ 」となる。すなわち、「bsParamSlot[0]」は0~7の値から選ばれる必要がある。これは、「bsParamSlot[0]」が7以上の値を有する場合、残りのパラメータセットに対するタイムスロットの個数が十分ではないためである。

40

【0196】

「bsParamSlot[0]」が5である場合、2番目のパラメータセットに対するタイムスロット位置bsParamSlot[1]は「 $5 + 1 = 6$ 」と「 $10 - 3 + 1 = 8$ 」との間の値から選ばれる必要がある。

【0197】

「bsParamSlot[1]」が7であれば、「bsParamSlot[2]」は8または9になりうる。「bsParamSlot[1]」が8であれば、「bsParamSlot[2]」は9になりうる。

【0198】

50

このため、「bsParamSlot[ps]」は、固定ビットとして表わされる代わりに、上記の特徴を用いて可変個数のビットで表わすことができる。

【0199】

「bsParamSlot[ps]」をビットストリームに構成するに当たって、「ps」が0である場合、「bsParamSlot[0]」は「nBitsParamSlot(0)」に対応する数のビットにより絶対値として表わすことができる。「ps」が0よりも大きな場合、「bsParamSlot[ps]」は「nBitsParamSlot(ps)」に対応する数により絶対値として表わすことができる。ビットストリームから上記のように構成された「bsParamSlot[ps]」を読み取るとき、各データに対するビットストリームの長さ、すなわち、「nBitsParamSlot[ps]」は式18を用いて表わすことができる。

10

【0200】

〔式18〕

【数18】

$$f_b(x) = \begin{cases} 0\text{bit}, & \text{if } x=1, \\ 1\text{bit}, & \text{if } x=2, \\ 2\text{bits}, & \text{if } 3 \leq x \leq 4, \\ 3\text{bits}, & \text{if } 5 \leq x \leq 8, \\ 4\text{bits}, & \text{if } 9 \leq x \leq 16, \\ 5\text{bits}, & \text{if } 17 \leq x \leq 32, \\ 6\text{bits}, & \text{if } 33 \leq x \leq 64, \end{cases}$$

20

30

【0201】

具体的に、「nBitsParamSlot[ps]」は  $n\text{BitsParamSlot}[0] = f_b(\text{numSlots} - \text{numParamSets} + 1)$  で表わすことができる。  $0 < ps < \text{numParamSets}$  であれば、「nBitsParamSlot[ps]」は  $n\text{BitsParamSlot}[ps] = f_b(\text{numSlots} - \text{numParamSets} + ps - \text{bsParamSlot}[ps - 1])$  で表わすことができる。「nBitsParamSlot[ps]」は式18を7ビットまで延長した式19を用いて決定可能である。

【0202】

〔式19〕

40

## 【数 19】

$$f_b(x) = \begin{cases} 0\text{bit}, & \text{if } x = 1, \\ 1\text{bit}, & \text{if } x = 2, \\ 2\text{bits}, & \text{if } 3 \leq x \leq 4, \\ 3\text{bits}, & \text{if } 5 \leq x \leq 8, \\ 4\text{bits}, & \text{if } 9 \leq x \leq 16, \\ 5\text{bits}, & \text{if } 17 \leq x \leq 32, \\ 6\text{bits}, & \text{if } 33 \leq x \leq 64, \\ 7\text{bits}, & \text{if } 65 \leq x \leq 128, \end{cases} \quad 10$$

20

## 【0203】

関数  $f_b(x)$  の例について後述する。「numSlots」が15であり、かつ、「numParamSets」が3である場合、上記の関数は  $nBitsParamSlot[1] = f_b(15 - 3 + 1 - 7) = 3$  ビットを求めることができる。この場合、「bsDiffParamSlot[1]」フィールド1305は3ビットで表わすことができる。

## 【0204】

3ビットで表わされる値が3である場合、「bsParamSlot[1]」は  $7 + 3 = 10$  となる。このため、 $nBitsParamSlot[2] = f_b(15 - 3 + 2 - 10) = 2$  ビットとなる。この場合、「bsDiffParamSlot[1]」フィールド1305は2ビットで表わすことができる。レジデュアルタイムスロットの個数がレジデュアルパラメータセットの個数と同数である場合、「bsDiffParamSlot[ps]」フィールドには0ビットが割り当てられる。換言すると、パラメータセットが適用されるタイムスロットの位置を表わすための追加情報が不要である。

30

## 【0205】

このため、「bsParamSlot[ps]」に対するビットの個数は可变的に決定可能である。「bsParamSlot[ps]」に対するビットの個数はデコーダにおいて関数  $f_b(x)$  を用いてビットストリームから読取り可能である。一部の実施形態において、関数  $f_b(x)$  は関数  $\text{ceil}(\log_2(x))$  を含むことができる。

## 【0206】

絶対値と差分値で表わされる「bsParamSlot[ps]」に関する情報をデコーダにおいてビットストリームから読み取るとき、まず、ビットストリームから「bsParamSlot[0]」が読み取られてから、 $0 < ps < \text{numParamSets}$  に対する「bsDiffParamSlot[ps]」が読み取られる。そして、 $0 < ps < \text{numParamSets}$  間隔に対する「bsParamSlot[ps]」は「bsParamSlot[0]」と「bsDiffParamSlot[ps]」を用いて求めることができる。例えば、図13Bに示すように、「bsParamSlot[ps]」は、「bsParamSlot[ps-1]」に「bsDiffParamSlot[ps]+1」を加えて求めることができる。

40

## 【0207】

50

図13Cは、本発明の一実施形態により、パラメータセットが適用されるタイムスロットの位置情報を示すシンタックスを示す図である。複数のパラメータセットが存在する場合、複数のパラメータセットに対する複数の「bsParamSlots」1307は少なくとも一つのグループとして表わすことができる。

【0208】

「bsParamSlots」1307の個数が $(kN + L)$ であり、かつ、「bsParamSlots」1307のそれぞれを表わすのに $Q$ ビットが必要となる場合、「bsParamSlots」1307は以下のグループとして表わすことができる。この場合、「 $k$ 」と「 $N$ 」は「0」ではない任意の整数であり、「 $L$ 」は $0 < L < N$ を満たす任意の整数である。

10

【0209】

グループ化方法は、 $N$ 個の「bsParamSlots」1307を束ねて $k$ 個のグループを生成するステップと、最後の $L$ 個の「bsParamSlots」1307を束ねて最終グループを生成するステップと、を含む。 $k$ 個のグループは $M$ ビットで表わすことができ、最終グループは $p$ ビットで表わすことができる。この場合、 $M$ ビットが、好ましくは、「bsParamSlots」1307のそれぞれをグループ化することなく表わす場合に用いられる $N * Q$ ビットよりも小さい。 $P$ ビットが、好ましくは、「bsParamSlots」1307のそれぞれをグループ化することなく表わす場合に用いられる $L * Q$ ビット以下である。

【0210】

20

例えば、2つのパラメータセットに対する1対の「bsParamSlots」1307がそれぞれ $d_1$ と $d_2$ であるとする。 $d_1$ と $d_2$ がそれぞれ5個の値を有する場合、 $d_1$ と $d_2$ のそれぞれを表わすのに3ビットが必要となる。この場合、たとえ3ビットは8個の値を表わすことができるとしても、実質的には5個の値が必要となる。このため、 $d_1$ と $d_2$ のそれぞれは3個の余分を有する。しかしながら、 $d_1$ と $d_2$ を束ねてグループとして表わす場合には、6ビット(=3ビット+3ビット)の代わりに5ビットが用いられる。具体的に、 $d_1$ と $d_2$ との全ての組み合わせは25個(=5\*5)のタイプを有するため、 $d_1$ と $d_2$ のグループは5ビットで表わすことができる。5ビットは32個の値を表わすことができるため、グループ化表現の場合、7個の余分が生成される。しかしながら、 $d_1$ と $d_2$ をグループ化して表わす場合、その余分は $d_1$ と $d_2$ をそれぞれ3ビットで表わす場合の余分よりも小さい。

30

【0211】

グループを構成するとき、グループに関するデータは、初期値に対する「bsParamSlot[0]」と、2番目以上の値に対する「bsParamSlot[ps]」との組間の差分値を用いて構成可能である。

【0212】

グループを構成するとき、パラメータセットの個数が1であれば、グループ化することなくビットが直接的に割当て可能であり、パラメータセットの個数が2以上であれば、グループ化を完了してからビットが割当て可能である。

【0213】

40

図14は、本発明の一実施形態によるエンコーディング方法のフローチャートである。以下、本発明によるオーディオ信号のエンコーディング方法及びエンコーダーの動作を説明する。

【0214】

まず、1枚の空間フレームにおいてタイムスロットの総数numSlotsとオーディオ信号のパラメータ帯域の総数numBandsが決定される(S1401)。

【0215】

そして、チャンネル変換モジュールに適用されるパラメータの個数及び/またはレジデュアル信号が決定される(S1402)。

【0216】

50

OTTボックスがLFEチャンネルモードを有する場合、OTTボックスに適用されるパラメータ帯域の個数は別途に決定される。

【0217】

OTTボックスがLFEチャンネルモードを有さない場合、OTTボックスに適用されるパラメータ帯域の個数として「numBands」が用いられる。

【0218】

続けて、空間フレームのタイプが判定される。この場合、空間フレームは固定フレームタイプと可変フレームタイプとに大別できる。

【0219】

空間フレームが可変フレームタイプである場合(S1403)、1枚の空間フレーム内において用いられるパラメータセットの個数が決定される(S1406)。この場合、パラメータセットはタイムスロット単位でチャンネル変換モジュールに適用可能である。

10

【0220】

続けて、パラメータセットが適用されるタイムスロットの位置が決定される(S1407)。この場合、パラメータセットが適用されるタイムスロットの位置は絶対値と差分値として表わすことができる。例えば、最初のパラメータセットが適用されるタイムスロットの位置は絶対値として表わすことができ、2番目以上のパラメータセットが適用されるタイムスロットの位置は以前のタイムスロットの位置からの差分値として表わすことができる。この場合、パラメータセットが適用されるタイムスロットの位置は可変個数のビットとして表わすことができる。

20

【0221】

具体的に、最初のパラメータセットが適用されるタイムスロットの位置は、タイムスロットの総数とパラメータセットの総数を用いて算出されるビットの個数で表わすことができる。2番目以上のパラメータセットが適用されるタイムスロットの位置は、タイムスロットの総数と、パラメータセットの総数と、以前のパラメータセットが適用されるタイムスロットの位置を用いて算出されるビットの個数で表わすことができる。

【0222】

空間フレームが固定フレームタイプである場合、1枚の空間フレームに用いられたパラメータセットの個数が決定される(S1404)。この場合、パラメータセットが適用されるタイムスロットの位置は予め設定された規則を用いて決定される。例えば、パラメータセットが適用されるタイムスロットの位置は、以前のパラメータセットが適用されるタイムスロットの位置から等間隔を有するように決定可能である(S1405)。

30

【0223】

続けて、ダウンミキシング部と空間生成部は、先に決定されたタイムスロットの総数と、パラメータ帯域の総数と、チャンネル変換部に適用されるべきパラメータ帯域の総数と、1枚の空間フレームにおけるパラメータセットの総数と、パラメータセットが適用されるタイムスロットの位置情報を用いて、ダウンミックス信号と空間情報をそれぞれ生成する(S1408)。

【0224】

最後に、マルチプレクシング部は、ダウンミックス信号と空間情報を含むビットストリームを生成して(S1409)、この生成されたビットストリームをデコーダーに転送する(S1409)。

40

【0225】

図15は、本発明の一実施形態によるデコーディング方法のフローチャートである。以下、本発明によるオーディオ信号のデコーディング方法及びデコーダーの動作を説明する。

【0226】

まず、デコーダーは、オーディオ信号のビットストリームを受信する(S1501)。デマルチプレクシング部は、受信されたビットストリームからダウンミックス信号と空間情報信号とを分離する(S1502)。続けて、空間情報信号デコーディング部は、空間

50

情報信号の構成情報から、1枚の空間フレームにおけるタイムスロットの総数、パラメータ帯域の総数及びチャンネル変換モジュールに適用されるパラメータ帯域の個数に関する情報を抽出する(S1503)。

【0227】

空間フレームが可変フレームタイプである場合(S1504)、1枚の空間フレームにおけるパラメータセットの個数とパラメータセットが適用されるタイムスロットの位置情報が空間フレームから抽出される(S1505)。タイムスロットの位置情報は、固定個数のビットまたは可変個数のビットで表わすことができる。この場合、最初のパラメータセットが適用されるタイムスロットの位置情報は絶対値として表わすことができ、2番目以上のパラメータセットが適用されるタイムスロットの位置情報は差分値として表わすことができる。2番目以上のパラメータセットが適用されるタイムスロットの実際の位置情報は、以前のパラメータセットが適用されたタイムスロットの位置情報に差分値を加えて求められる。

10

【0228】

最後に、抽出された情報を用いてダウンミックス信号がマルチチャンネルオーディオ信号に変換される(S1506)。

【0229】

この明細書に記載の実施形態は、従来のオーディオコーディング方式に比べて種々のメリットを与える。

【0230】

最初に、マルチチャンネルオーディオ信号のコーディングにおいて、パラメータセットが適用されるタイムスロットの位置を可変個数のビットで表わすことにより、転送データ量を低減することができる。

20

【0231】

第二に、最初のパラメータセットが適用されるタイムスロットの位置を絶対値で表わし、2番目以上のパラメータセットが適用されるタイムスロットの位置を差分値で表わすことにより、転送データ量を低減することができる。

【0232】

第三に、OTTボックス及び/またはTTTボックスなどに適用されるパラメータ帯域の個数を固定個数のビットまたは可変個数のビットで表わすことにより、転送データ量を低減することができる。この場合、パラメータセットが適用されるタイムスロットの位置は、上述した原理を用いて表わすことができ、ここで、パラメータセットはパラメータ帯域の個数範囲内に存在する。

30

【0233】

図16は、図1～図15に基づき説明されたオーディオエンコーダー/デコーダーを実現する装置構造1600の一例を示すブロック図である。この装置構造1600は、パソコン、サーバコンピュータ、家電装置、携帯電話、PDA、電子タブレット、テレビシステム、テレビセットトップボックス、ゲームコンソール、媒体再生器、音楽再生器、ナビゲーションシステム及びオーディオ信号をデコーディング可能な任意のその他の装置を含む種々の装置に適用可能であるが、これらに制限されるものではない。これらの装置の一部はハードウェアとソフトウェアとの組み合わせを用いて変更された構造を実現することができる。

40

【0234】

この構造1600は、1以上のプロセッサ1602(例えば、PowerPC(登録商標)、Intel Pentium(登録商標)4など)と、1以上のディスプレイ装置1604(例えば、CRT、LCDなど)と、オーディオサブシステム1606(例えば、オーディオハードウェア/ソフトウェア)と、1以上のネットワークインタフェース1608(例えば、Ethernet(登録商標)、FireWire(登録商標)、USBなど)と、入力装置1610(例えば、キーボード、マウスなど)と、1以上のコンピュータにて読取り可能な媒体1612(例えば、RAM、ROM、SDRAM、ハード

50

ディスク、光ディスク、フラッシュメモリなど)と、を備える。これらの構成要素は1以上のバス1614(例えば、EISA、PCI、PCI Expressなど)を介して通信し、データをやり取りすることができる。

【0235】

「コンピュータにて読取り可能な媒体」という用語は、プロセッサ1602に実行用命令語を与える任意の媒体をいい、不揮発性媒体(例えば、光ディスクや磁気ディスクなど)、揮発性媒体(例えば、メモリーなど)及び転送媒体を含むが、これらに制限されるものではない。転送媒体は、同軸ケーブル、銅配線及び光ファイバを含むが、これらに制限されるものではない。また、転送媒体は、音響波、光波または無線周波数波形の形態が採用可能である。

10

【0236】

コンピュータにて読取り可能な媒体1612は、オペレーティングシステム1616(例えば、Mac OS(登録商標)、Windows(登録商標)、Linux(登録商標)など)と、ネットワーク通信モジュール1618と、オーディオコーデック1620と、1以上のアプリケーション1622と、をさらに含む。

【0237】

オペレーティングシステム1616は、マルチユーザー、マルチプロセッシング、マルチタスキング、マルチスレディング、リアルタイムなどであってもよい。オペレーティングシステム1616は、入力装置1610から入力を認識し、ディスプレイ装置1604とサブシステム1606に出力を送り、コンピュータにて読取り可能な媒体1612(例えば、メモリーまたは記録装置など)の上にファイルとディレクトリを保持し、周辺装置(例えば、ディスクドライブ、プリンターなど)を制御し、1以上のバス1614上におけるトラフィックを管理する基本的なタスクを行うが、これらに制限されるものではない。

20

【0238】

ネットワーク通信モジュール1618は、ネットワーク接続を確立して保持する種々の構成要素(例えば、TCP/IP、HTTP、Ethernet(登録商標)などの通信プロトコルを実現するソフトウェアなど)を含む。ネットワーク通信モジュール1618は、装置構造1600のオペレーターに情報(例えば、オーディオコンテンツなど)を探索してネットワーク(例えば、インターネットなど)を検索可能にするブラウザを含むことができる。

30

【0239】

オーディオコーデック1620は、図1~図15に基づき説明したエンコーディング及び/またはデコーディングプロセスの全部または一部を実現する役割を果たす。一部の実施形態において、オーディオコーデックはハードウェア(例えば、プロセッサ1602、オーディオサブシステム1606など)と連動してこの明細書に記載の本発明によりオーディオ信号をエンコーディング及び/またはデコーディングすることをはじめとしてオーディオ信号を処理する。

【0240】

アプリケーション1622は、オーディオコンテンツに関連する任意のソフトウェアを含むことができ、オーディオコンテンツは、媒体再生器、音楽再生器(例えば、MP3再生器など)、移動電話アプリケーション、PDA、テレビシステム、セットトップボックスなどにおいてエンコーディング及び/またはデコーディングされるが、これらに制限されるものではない。一実施形態において、オーディオコーデックはアプリケーションサービスプロバイダーがネットワーク(例えば、インターネットなど)を介してエンコーディング/デコーディングサービスを提供するのに使用可能である。

40

【0241】

以上の説明においては、説明の目的から、本発明の完全な理解を提供するための種々の特定の詳細について開示している。しかしながら、当業者であれば、本発明がこれらの特定の詳細なしにも実行可能であるということが理解できるであろう。なお、本発明を不明

50

にすることを防ぐために、構造及び装置はブロック図として示してある。

【0242】

特に、当業者であれば、他の構造及びグラフィック環境が使用可能であり、上述とは異なるグラフィックツールと製品を用いて本発明が実現可能であるという点が理解できるであろう。特に、クライアント/サーバーアプローチは本発明のダッシュボード機能を与える構造の一例に過ぎず、当業者であれば、クライアント/サーバーアプローチではない他のものが使用可能であるという点が理解できるであろう。

【0243】

詳細な説明の一部は、コンピュータメモリ内におけるデータビットに対する演算のアルゴリズムとシンボル表現で実現されている。これらのアルゴリズム説明と表現は、データ処理分野の当業者が他の当業者に自分の作業の本質を最も有効に伝えるための手段である。一般的に、そしてこの明細書において、アルゴリズムは所望の結果に至るステップの一連のシーケンスとして認識される。これらのステップは物理量の操作を必要とする。一般に、必ずしもその限りではないが、このような量は貯蔵されたり、転送されたり、組み合わせられたり、比較されたり、他に操作可能な電気信号または磁気信号の形態をとる。主として、通常の使用の理由から、これらの信号をビット、値、エレメント、シンボル、キャラクター、述語、ナンバーなどと呼んだ方が便利である。

【産業上の利用可能性】

【0244】

しかしながら、これらの用語及びこれに類似する用語はあらゆる適切な物理量に関連するものであり、単にこれらの量に適用される便利な命名に過ぎない。論議から明らかなように、特に断わりのない限り、この明細書において、「処理」または「コンピューティング」または「算出」または「決定」または「表示」などの用語を用いるということは、コンピュータシステムのレジスターとメモリ内において物理（電気）量として表わされるデータを操作し、これをコンピュータシステムメモリまたはレジスターや他の情報の貯蔵、転送または表示装置において物理量で表わされるデータに変換するコンピュータシステムまたはこれに類似する電子計算装置のアクション及び処理を言う。

【0245】

本発明はこのような動作を行う装置に関する。この装置は、求められる目的により特別に構成可能であるか、あるいは、コンピュータに格納されたコンピュータプログラムにより選択的に活性化されるか、あるいは、再構成される汎用のコンピュータを含むことができる。このようなコンピュータプログラムは、フロッピー（登録商標）ディスク、光ディスク、CD-ROM、磁気光ディスクをはじめとする任意タイプのディスク、読み専用メモリー（ROM）、ランダムアクセスメモリー（RAM）、EPROM、EEPROM、磁気カードまたは光カード、または電子的な命令語を格納するのに適した任意タイプの媒体を含み、コンピュータシステムに接続されるコンピュータ記憶媒体に格納可能である。

【0246】

明細書に記載のアルゴリズムとモジュールが元々任意の特定のコンピュータや他の装置に関するものではない。この明細書に記載の技術によるプログラムに種々の汎用システムが使用可能であり、方法ステップを行う上でより特定化した装置を構成することが便利になることがある。このような種々のシステムに必要な構造については、後述する。この明細書に記載の本発明の教示事項を実現するために種々のプログラミング言語が使用可能である。また、当業者にとって自明であるように、本発明のモジュール、特徴、属性、方法論及びその他の態様は、ソフトウェア、ハードウェア、ファームウェアまたはこれらの3種類に対する任意の組み合わせにより実現可能である。もちろん、本発明の構成要素がソフトウェアとして実現される場合、これらの構成要素は独自のプログラムとして、より大きなプログラムの一部として、複数の個別プログラムとして、静的にまたは動的にリンクされるライブラリーとして、カーネルロード型モジュールとして、デバイスドライバとして実現されるか、あるいは、コンピュータプログラミングの分野における当業者に現在周知の、または、将来に周知になる予定のあらゆる任意の方式により実現可能である

10

20

30

40

50

。また、本発明が任意の特定のオペレーティングシステムや環境において実現されることに制限されることはない。

【0247】

当業者には本発明の思想または範囲から逸脱することなく、この明細書に記載の実施形態に種々の変形及び変更がなされうるということは自明であろう。よって、本発明は、このような変形及び変更が特許請求の範囲及びその等価物の範囲内にあるものである限り、開示された実施形態のこのような変形及び変更をいずれもカバーするものである。

【図面の簡単な説明】

【0248】

【図1】本発明の一実施形態により空間情報を生成する原理を示す図である。 10

【図2】本発明の一実施形態によりオーディオ信号をエンコーディングするエンコーダーのブロック図である。

【図3】本発明の一実施形態によりオーディオ信号をデコーディングするデコーダーのブロック図である。

【図4】本発明の一実施形態によるデコーダーのアップミキシング部に含まれるチャンネル変換部のブロック図である。

【図5】本発明の一実施形態によりオーディオ信号のビットストリームを構成する方法を説明するブロック図である。

【図6A】本発明の一実施形態による、パラメータセット、タイムスロット及びパラメータ帯域の関係を説明するための図面である。 20

【図6B】本発明の一実施形態による、パラメータセット、タイムスロット及びパラメータ帯域の関係を説明するための時間/周波数グラフである。

【図7A】本発明の一実施形態による空間情報信号の構成情報を表示するシンタックスを示す図である。

【図7B】本発明の一実施形態による空間情報信号のパラメータ帯域の個数を示す表である。

【図8A】本発明の一実施形態より、固定個数のビットとしてOTTボックスに適用されたパラメータ帯域の個数を示すシンタックスを示す図である。

【図8B】本発明の一実施形態により、可変個数のビットでOTTボックスに適用されたパラメータ帯域の個数を示すシンタックスを示す図である。 30

【図9A】本発明の一実施形態により、固定個数のビットでTTTボックスに適用されたパラメータ帯域の個数を示すシンタックスを示す図である。

【図9B】本発明の一実施形態により、可変個数のビットでTTTボックスに適用されたパラメータ帯域の個数を示すシンタックスを示す図である。

【図10A】本発明の一実施形態による空間拡張フレームのための空間拡張構成情報のシンタックスを示す図である。

【図10B】本発明の一実施形態により空間拡張フレームにレジデュアル信号が含まれる場合、レジデュアル信号のための空間拡張構成情報のシンタックスを示す図である。

【図10C】本発明の一実施形態により空間拡張フレームにレジデュアル信号が含まれる場合、レジデュアル信号のための空間拡張構成情報のシンタックスを示す図である。 40

【図10D】本発明の一実施形態によりレジデュアル信号のためのパラメータ帯域の個数を示す方法のためのシンタックスを示す図である。

【図11A】本発明の一実施形態よりノンガイドコーディングを用いるデコーディング装置のブロック図である。

【図11B】本発明の一実施形態によりパラメータ帯域の個数をグループで示す方法を示す図である。

【図12】本発明の一実施形態による空間フレームの構成情報のシンタックスを示す図である。

【図13A】本発明の一実施形態によりパラメータセットが適用されるタイムスロットの位置情報のシンタックスを示す図である。 50

【図13B】本発明の一実施形態により絶対値及び差分値としてパラメータセットが適用されるタイムスロットの位置情報を表わすためのシンタックスを示す図である。

【図13C】本発明の一実施形態によりパラメータセットがグループとして適用されるタイムスロットの複数の位置情報を示す図である。

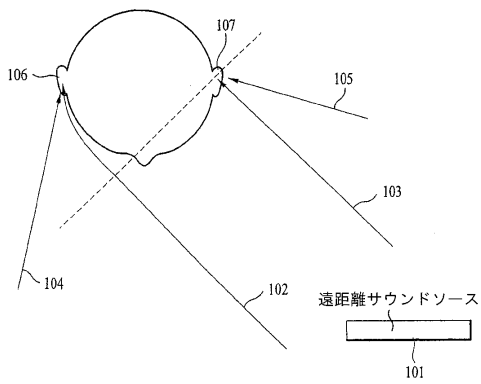
【図14】本発明の一実施形態によるエンコーディング方法のフローチャートである。

【図15】本発明の一実施形態によるデコーディング方法のフローチャートである。

【図16】図1～図15に基づき説明されるエンコーディング及びデコーディング処理を実現するための装置構造を示すブロック図である。

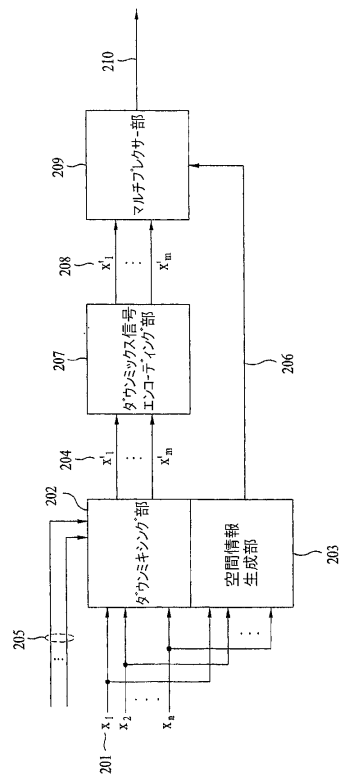
【図1】

FIG. 1



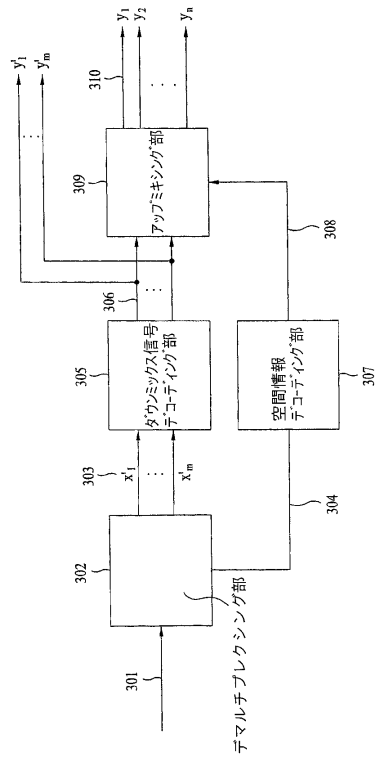
【図2】

FIG. 2



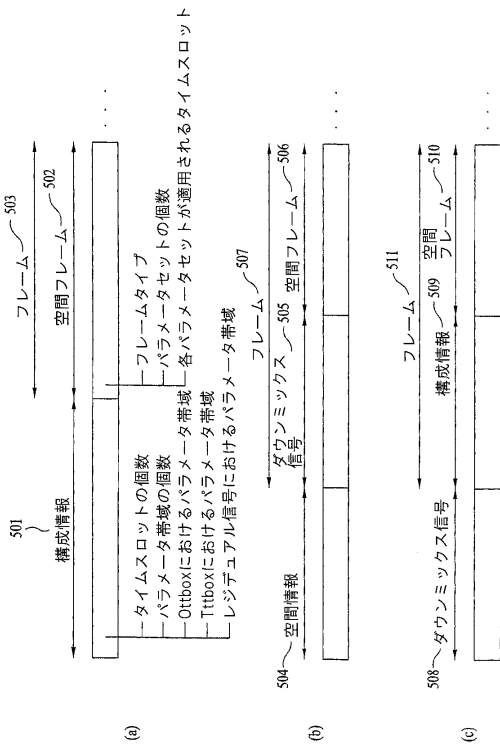
【 図 3 】

FIG. 3



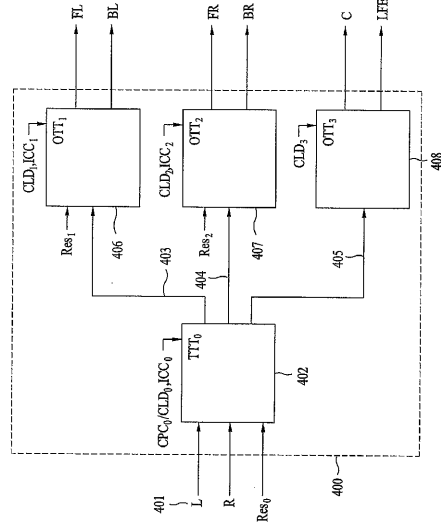
【 図 5 】

FIG. 5



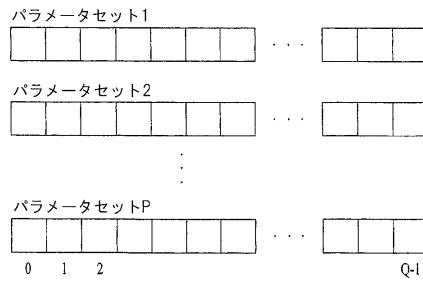
【 図 4 】

FIG. 4

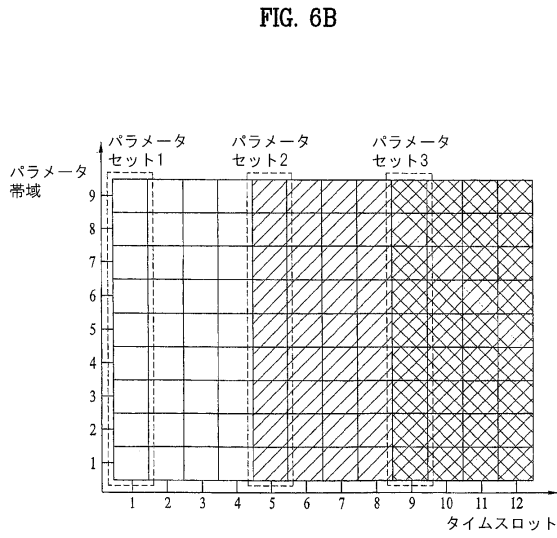


【 図 6 A 】

FIG. 6A



【 図 6 B 】



【 図 7 A 】

FIG. 7A

Syntax	No. of bits
SpatialSpecificConfig()	
{	
701 bsSamplingFrequencyIndex;	4
702 if ( bsSamplingFrequencyIndex == 0xf) {	
bsSamplingFrequency;	24
}	
703 bsFrameLength;	7
704 bsFreqRes;	3
705 bsTreeConfig;	4
706 bsQuantMode;	3
707 bsOneIcc;	1
708 bsArbitraryDownmix;	1
709 bsFixedGainSur;	3
710 bsFixedGainLFE;	3
711 bsFixedGainDMX;	3
712 bsMatrixMode;	1
713 bsTempShapeConfig;	4
714 bsDecorConfig;	4
715 bs3DAudioMode;	1
...	
716 for (i=0; i<numOttBoxes; i++) {	
OttConfig(i);	
}	
717 for (i=0; i<numTttBoxes; i++) {	
TttConfig(i);	
}	
...	
718 Spatial Extension Config()	
}	

【 図 7 B 】

FIG. 7B

bsFreqRes	numBands
0	Reserved
1	28
2	20
3	14
4	10
5	7
6	5
7	4

【 図 8 A 】

FIG. 8A

Syntax	No. of bits
OttConfig(i)	
{	
801 if (ottModelLfe[i]) {	
bsOttBands[i];	5
}	
else {	
bsOttBands[i] = numBands;	
}	
}	

【 図 8 B 】

FIG. 8B

Syntax	No. of bits
OttConfig(i)	
{	
802 if (ottModelLfe[i]) {	
bsOttBands[i];	Bitsnumbands
}	
else {	
bsOttBands[i] = numBands;	
}	
}	

numBandsの表示のための  
最小ビット数

【 図 9 A 】

FIG. 9A

Syntax	No. of bits
TttConfig(i)	
{	
901   bsTttDualMode[i];	1
902   bsTttModeLow[i];	3
if (bsTttDualMode[i]) {	
903      bsTttModeHigh[i];	3
904      bsTttBandsLow[i];	5
905      bsTttBandsHigh[i] = numBands;	
}	
else {	
906      bsTttBandsLow[i] = numBands;	
}	
}	

【 図 9 B 】

FIG. 9B

Syntax	No. of bits
TttConfig(i)	
{	
bsTttDualMode[i];	1
bsTttModeLow[i];	3
if (bsTttDualMode[i]) {	
bsTttModeHigh[i];	3
907      bsTttBandsLow[i];	BitsnumBands
bsTttBandsHigh[i] = numBands;	
}	
else {	
bsTttBandsLow[i] = numBands;	
}	
}	

numBandsの表示のための  
最小ビット数

【 図 1 0 A 】

FIG. 10A

Syntax	No. of bits
SpatialExtensionConfig()	
{	
while (bitsAvailble() >= 8) {	
bsSacExtType;	4
cnt = bsSacExtLen;	4
if (cnt == 15) {	
cnt += bsSacExtLenAdd;	8
}	
if (cnt == 15+255) {	
cnt += bsSacExtLenAddAdd;	16
}	
bitsRead = SpatialExtensionConfigData(bsSacExtType)	
nFillBits = 8*cnt-bitsRead;	
bsFillBits;	
}	
}	

【 図 1 0 C 】

FIG. 10C

Syntax	No. of bits
ResidualConfig(i)	
{	
1011   bsResidualPresent[i];	1
1012   if (bsResidualPresent[i]) {	
bsResidualBands[i];	5
}	
}	

【 図 1 0 B 】

FIG. 10B

Syntax	No. of bits
SpatialExtensionConfigData(1)	
{	
1008   bsResidualSamplingFrequencyIndex;	4
1009   bsResidualFramesPerSpatialFrame;	2
1010   for (i=0; i<numOutBoxes+numTttBoxes; i++) {	
ResidualConfig(i);	
}	
}	

【 図 1 0 D 】

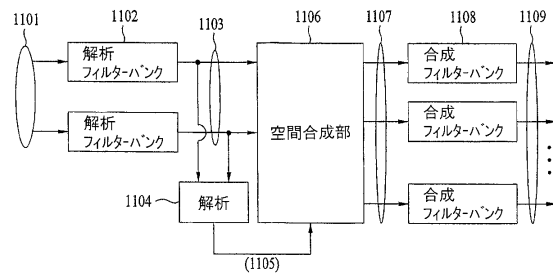
FIG. 10D

Syntax	No. of bits
ResidualConfig(i)	
{	
1013 bsResidualPresent[i];	1
1014 if (bsResidualPresent[i]) {	
bsResidualBands[i];	BitsnumBands
}	
}	

↓  
numBandsの表示のための  
最小ビット数

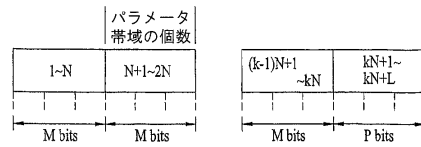
【 図 1 1 A 】

FIG. 11A



【 図 1 1 B 】

FIG. 11B



【 図 1 2 】

FIG. 12

Syntax	No. of bits
SpatialFrame()	
{	
1201 FramingInfo();	
1202 bsIndependencyFlag;	1
1203 OutData();	
1204 TutData();	
1205 SmgData();	
1206 TempShapeData();	
}	

【 図 1 3 A 】

FIG. 13A

Syntax	No. of bits
FramingInfo()	
{	
1301 bsFramingType;	1
1302 bsNumParamSets;	3
if (bsFramingType) {	
for (ps=0; ps<numParamSets; ps++) {	
1303 bsDiffParamSlot[ps];	BitsnumSlots
}	
}	
}	

↓  
numSlotsの表示のための  
最小ビット数

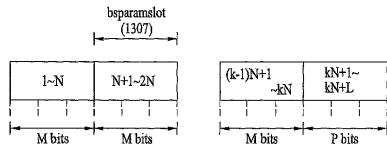
【図13B】

FIG. 13B

Syntax	No. of bits
FramingInfo()	
{	
bsFramingType;	
bsNumParamSets;	1
if (bsFramingType) {	3
for (ps=0; ps<numParamSets; ps++) {	
if (ps==0){	
bsParamSlot[0];	nBitsParamSlot(0)
} else {	
bsDiffParamSlot[ps];	nBitsParamSlot(ps)
bsParamSlot[ps] = bsParamSlot[ps-1]	
+ bsDiffParamSlot[ps] + 1;	
}	
}	
}	
}	

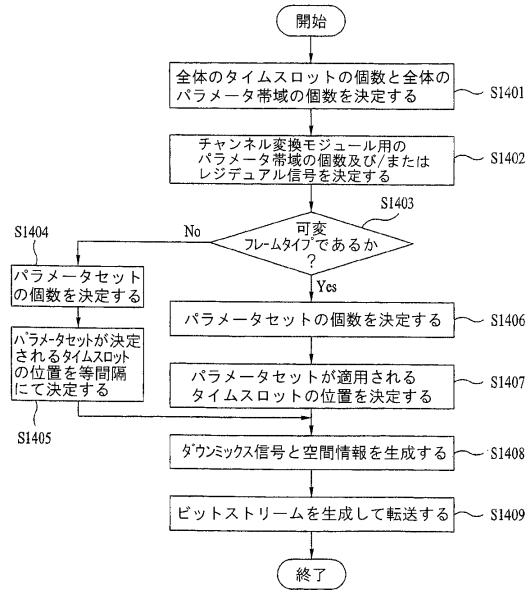
【図13C】

FIG. 13C



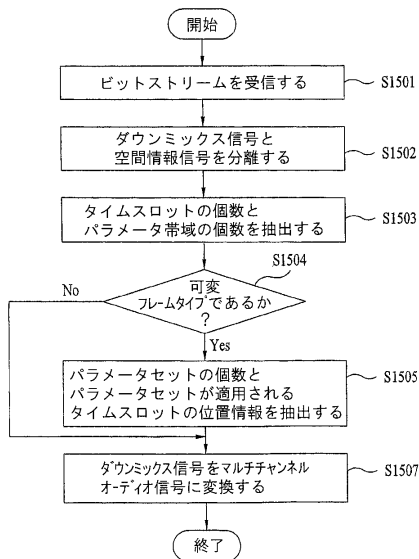
【図14】

FIG. 14



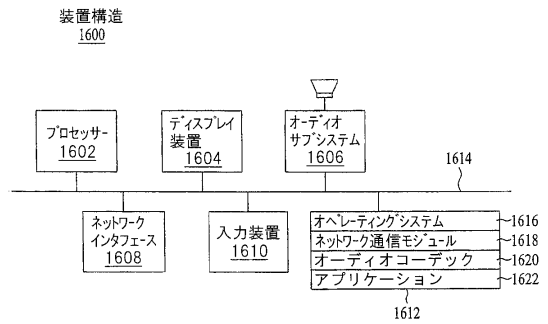
【図15】

FIG. 15



【図16】

FIG. 16



## フロントページの続き

- (31)優先権主張番号 60/723,007  
(32)優先日 平成17年10月4日(2005.10.4)  
(33)優先権主張国 米国(US)
- (31)優先権主張番号 60/726,228  
(32)優先日 平成17年10月14日(2005.10.14)  
(33)優先権主張国 米国(US)
- (31)優先権主張番号 60/729,225  
(32)優先日 平成17年10月24日(2005.10.24)  
(33)優先権主張国 米国(US)
- (31)優先権主張番号 10-2006-0004051  
(32)優先日 平成18年1月13日(2006.1.13)  
(33)優先権主張国 韓国(KR)
- (31)優先権主張番号 10-2006-0004057  
(32)優先日 平成18年1月13日(2006.1.13)  
(33)優先権主張国 韓国(KR)
- (31)優先権主張番号 10-2006-0004062  
(32)優先日 平成18年1月13日(2006.1.13)  
(33)優先権主張国 韓国(KR)
- (31)優先権主張番号 10-2006-0004063  
(32)優先日 平成18年1月13日(2006.1.13)  
(33)優先権主張国 韓国(KR)
- (31)優先権主張番号 10-2006-0004055  
(32)優先日 平成18年1月13日(2006.1.13)  
(33)優先権主張国 韓国(KR)
- (31)優先権主張番号 10-2006-0004065  
(32)優先日 平成18年1月13日(2006.1.13)  
(33)優先権主張国 韓国(KR)
- (31)優先権主張番号 60/762,536  
(32)優先日 平成18年1月27日(2006.1.27)  
(33)優先権主張国 米国(US)
- (74)代理人 100114018  
弁理士 南山 知広
- (72)発明者 パン, ヒー スク  
大韓民国, ソウル 137-130, ソチョ-グ, ヤンジエ-ドン, 101, #14-10
- (72)発明者 オー, ヒョン オ  
大韓民国, ギョンギ-ド 151-057, ゴヤン-シ, イルサン-グ, ジュヨブ 1(イル)-  
ドン, ガンソン メウル 3-ダンジ アパートメント, 306-403
- (72)発明者 キム, ドン スー  
大韓民国, ソウル 151-801, グワナク-グ, ナムヒョン-ドン, 602-265, ウーリ  
ム ビラ, 1502
- (72)発明者 リム, ジェ ヒュン  
大韓民国, ソウル 151-801, グワナク-グ, ナムヒョン-ドン, 1062-20, パーク  
ビル オフィステル, 609
- (72)発明者 ジュン, ヤン ウォン  
大韓民国, ソウル 135-270, カンナム-グ, ドゴク-ドン, ヨクサム ハンシン アパー  
トメント, 2-803

審査官 菊地 陽一

- (56)参考文献 国際公開第2004/097794(WO, A1)  
特表2008-501948(JP, A)  
国際公開第2004/008806(WO, A1)  
J.Herre, et al., The Reference Model Architecture for MPEG Spatial Audio Coding, Convention Paper of the Audio Engineering Society 118th Convention, 2005年 5月28日, Convention Paper 6447
- (58)調査した分野(Int.Cl., DB名)  
G10L 19/00