

(19) 日本国特許庁(JP)

## (12) 特許公報(B2)

(11) 特許番号

特許第6045696号  
(P6045696)

(45) 発行日 平成28年12月14日(2016.12.14)

(24) 登録日 平成28年11月25日(2016.11.25)

(51) Int.Cl.

G 1 O L 19/008 (2013.01)  
H 0 4 S 3/00 (2006.01)

F 1

G 1 O L 19/008 200  
H 0 4 S 3/00 Z

請求項の数 7 (全 27 頁)

(21) 出願番号	特願2015-523022 (P2015-523022)
(86) (22) 出願日	平成25年7月26日 (2013.7.26)
(65) 公表番号	特表2015-531078 (P2015-531078A)
(43) 公表日	平成27年10月29日 (2015.10.29)
(86) 國際出願番号	PCT/KR2013/006732
(87) 國際公開番号	W02014/021588
(87) 國際公開日	平成26年2月6日 (2014.2.6)
審査請求日	平成27年1月15日 (2015.1.15)
(31) 優先権主張番号	10-2012-0084229
(32) 優先日	平成24年7月31日 (2012.7.31)
(33) 優先権主張国	韓国 (KR)
(31) 優先権主張番号	10-2012-0084230
(32) 優先日	平成24年7月31日 (2012.7.31)
(33) 優先権主張国	韓国 (KR)

(73) 特許権者	513105524 インテレクチュアル ディスカバリー シ ーオー エルティディ I N T E L L E C T U A L D I S C O V E R Y C O., LTD. 大韓民国 135-745ソウル、カンナム グ、サムソンロ 511番 (サムソンドン、 ゴールデン・タワー) 10エフ
(74) 代理人	100101454 弁理士 山田 阜二
(74) 代理人	100081422 弁理士 田中 光雄
(74) 代理人	100125874 弁理士 川端 純市

最終頁に続く

(54) 【発明の名称】オーディオ信号処理方法および装置

## (57) 【特許請求の範囲】

## 【請求項 1】

複数のオブジェクト信号を含む第1オブジェクト信号グループに対する第1信号と、複数のオブジェクト信号を含む第2オブジェクト信号グループに対する第2信号とを受信するステップと、

前記第1オブジェクト信号グループに対する第1メタデータと、前記第2オブジェクト信号グループに対する第2メタデータとを受信するステップと、

前記第1信号と前記第1メタデータを用いて、前記第1オブジェクト信号グループに属する複数のオブジェクト信号を生成するステップと、

前記第2信号と前記第2メタデータを用いて、前記第2のオブジェクト信号グループに属する複数のオブジェクト信号を生成するステップとを含むオーディオ信号処理方法であつて、

前記メタデータの各々は、対応するオブジェクト信号グループの各々に属するオブジェクト信号に対応するオブジェクトの位置情報を含み、

前記オブジェクトが当該オブジェクトの位置が時間に応じて変化するダイナミックオブジェクトであるときに、前記オブジェクトの位置情報は前記オブジェクトの前の位置情報に対する位置値を表すことを特徴とする、オーディオ信号処理方法。

## 【請求項 2】

前記第1オブジェクト信号グループに属するオブジェクト信号のうちの少なくとも1つのオブジェクト信号と、前記第2オブジェクト信号グループに属する少なくとも1つのオ

10

20

プロジェクト信号を用いて、出力オーディオ信号を生成するステップをさらに含むことを特徴とする、請求項 1 に記載のオーディオ信号処理方法。

【請求項 3】

前記第 1 メタデータと前記第 2 メタデータは、1 つのビット列から受信されることを特徴とする、請求項 1 に記載のオーディオ信号処理方法。

【請求項 4】

前記第 1 オブジェクト信号グループに属した少なくとも 1 つのオブジェクト信号に対するダウンミックスゲイン情報は前記第 1 メタデータから得られ、

前記ダウンミックスゲイン情報を用いて前記少なくとも 1 つのオブジェクト信号を生成することを特徴とする、請求項 1 に記載のオーディオ信号処理方法。 10

【請求項 5】

グローバルゲイン情報を受信するステップをさらに含み、前記グローバルゲイン情報は、前記第 1 オブジェクト信号グループと前記第 2 オブジェクト信号グループにすべて適用されるゲイン値であることを特徴とする、請求項 1 に記載のオーディオ信号処理方法。

【請求項 6】

前記第 1 オブジェクト信号グループに属するオブジェクト信号のうちの少なくとも 1 つのオブジェクト信号と、前記第 2 オブジェクト信号グループに属する少なくとも 1 つのオブジェクト信号は、同じ時間帯に再生されることを特徴とする、請求項 1 に記載のオーディオ信号処理方法。 20

【請求項 7】

前記メタデータは、前記オブジェクトの位置情報が前記オブジェクトの前の位置値に対する位置値であることを示す情報をさらに含むことを特徴とする、請求項 1 に記載のオーディオ信号処理方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、オブジェクトオーディオ信号処理方法および装置に関するものであって、より詳細には、オブジェクトオーディオ信号の符号化および復号化を行ったり、3 次元空間にレンダリングするための方法および装置に関するものである。

【背景技術】 30

【0002】

3 D オーディオとは、既存のサラウンドオーディオで提供する水平面上のサウンド場面(2 D)に、高さ方向にさらに他の軸(dimension)を提供することにより、言葉通り3次元空間での臨場感のあるサウンドを提供するための一連の信号処理、伝送、符号化、再生技術などを通称する。特に、3 D オーディオを提供するためには、従来より多い数のスピーカーを用いたり、あるいは少ない数のスピーカーを用いても、スピーカーが存在しない仮想の位置で音像が結ばれるようにするレンダリング技術が広く要求される。

【0003】

3 D オーディオは、今後発売される超高解像度テレビ(UHDTV)に対応するオーディオソリューションになることが予想され、高品質のインフォテインメント空間に進化している車両でのサウンドを含めて、その他、劇場サウンド、個人用3 D テレビ、タブレット、スマートフォン、クラウドゲームなど多様に応用されることが予想される。 40

【発明の概要】

【発明が解決しようとする課題】

【0004】

3 D オーディオは、まず、最大22.2チャンネルまで、従来より多いチャンネルの信号を伝送することが必要であるが、このためには、これに適した圧縮伝送技術が要求される。従来のMP3、AAC、DTS、AC3などの高音質符号化の場合、主に5.1チャンネル未満のチャンネルのみを伝送するのに最適化されていた。

【0005】 50

また、22.2チャンネル信号を再生するためには、24個のスピーカシステムを設置した聴取空間に対するインフラが必要であるが、市場へ短期間の拡散が容易でないので、22.2チャンネル信号を、それより小さい数のスピーカを有する空間で効果的に再生するための技術、逆に、既存のステレオ、あるいは5.1チャンネルの音源を、それより多い数のスピーカである10.1チャンネル、22.2チャンネル環境で再生できるようにする技術、ひいては、規定されたスピーカ位置と規定された聴取室環境でない所でも元の音源が提供するサウンド場面を提供できるようにする技術、そして、ヘッドホン聴取環境においても3Dサウンドを楽しめるようにする技術などが要求される。これらの技術を、本願では、通称、レンダリング(rendering)といい、詳細には、それぞれダウンミックス、アップミックス、柔軟なレンダリング(flexible rendering)、バイノーラルレンダリング(binaural rendering)などと呼ぶ。10

#### 【0006】

一方、このようなサウンド場面を効果的に伝送するための代案として、オブジェクトベースの信号伝送方策が必要である。音源によって、チャンネルベースで伝送するより、オブジェクトベースで伝送することがより有利な場合があるだけでなく、オブジェクトベースで伝送する場合、ユーザが任意にオブジェクトの再生の大きさと位置を制御することができるなど、インタラクティブな音源の聴取を可能にする。これによって、オブジェクト信号を高伝送率で圧縮することのできる効果的な伝送方法が必要である。

#### 【0007】

また、前記チャンネルベースの信号とオブジェクトベースの信号とが混合された形態の音源も存在することができ、これによって新たな形態の聴取経験を提供することもできる。したがって、チャンネル信号とオブジェクト信号をともに効果的に伝送し、これを効果的にレンダリングするための技術も必要である。

#### 【課題を解決するための手段】

#### 【0008】

本発明の一態様によれば、複数のオブジェクト信号を定められた方法によって分類した第1オブジェクト信号グループと第2オブジェクト信号グループを生成するステップと、第1オブジェクト信号グループに対して第1ダウンミックス信号を生成するステップと、第2オブジェクト信号グループに対して第2ダウンミックス信号を生成するステップと、第1オブジェクト信号グループに含まれたオブジェクト信号に対して、第1ダウンミックス信号に対応して第1オブジェクト抽出情報を生成するステップと、第2オブジェクト信号グループに含まれたオブジェクト信号に対して、第2ダウンミックス信号に対応して第2オブジェクト抽出情報を生成するステップとを含むオーディオ信号処理方法が提供できる。30

#### 【0009】

本発明の他の態様によれば、第1ダウンミックス信号と第2ダウンミックス信号を含む複数のダウンミックス信号を受信するステップと、第1ダウンミックス信号に対応する、第1オブジェクト信号グループに対する第1オブジェクト抽出情報を受信するステップと、第2ダウンミックス信号に対応する、第2オブジェクト信号グループに対する第2オブジェクト抽出情報を受信するステップと、第1ダウンミックス信号と第1オブジェクト抽出情報を用いて、第1オブジェクト信号グループに属するオブジェクト信号を生成するステップと、第2ダウンミックス信号と第2オブジェクト抽出情報を用いて、第2オブジェクト信号グループに属するオブジェクト信号を生成するステップとを含むオーディオ信号処理方法が提供できる。40

#### 【発明の効果】

#### 【0010】

本発明によれば、オーディオ信号を効果的に表現し、符号化し、伝送および格納することができ、多様な再生環境および機器を介して高品質のオーディオ信号を再生することができる。

10

20

30

40

50

**【0011】**

本発明の効果が前記効果に制限されるわけではなく、言及されていない効果は、本明細書および添付した図面から本発明の属する技術分野における通常の知識を有する者に明確に理解できる。

**【図面の簡単な説明】****【0012】**

【図1】同じ視聴距離で映像の大きさに応じた視聴角度を説明するための図である。

【図2】マルチチャンネルの一例として22.2chのスピーカ配置構成図である。

【図3】聞き手が3Dオーディオを聴取する聴取空間上での各サウンドオブジェクトの位置を示す概念図である。 10

【図4】図3に示されたオブジェクトに対して、本発明によるグルーピング方法を用いてオブジェクト信号グループを形成した例示的構成図である。

【図5】本発明によるオブジェクトオーディオ信号符号化器の一実施例に対する構成図である。

【図6】本発明の一実施例による復号化装置の例示的な構成図である。

【図7】本発明による符号化方法によって符号化して生成したビット列の一実施例である。

【図8】本発明によるオブジェクトおよびチャンネル信号復号化システムをブロック図に示す一実施例である。

【図9】本発明による他の形態のオブジェクトおよびチャンネル信号復号化システムのブロック図である。 20

【図10】本発明による復号化システムの一実施例である。

【図11】本発明による複数のオブジェクト信号に対するマスキング閾値を説明するための図である。

【図12】本発明による複数のオブジェクト信号に対するマスキング閾値を算出する符号化器の一実施例である。

【図13】5.1チャンネルセットアップに対してITU-R勧告案による配置と、任意の位置に配置された場合を説明するための図である。

【図14】本発明によるオブジェクトのビット列に対する復号化器とこれを用いたフレキシブルレンダリングシステムが接続された一実施例の構造である。 30

【図15】本発明によるオブジェクトのビット列に対する復号化とレンダリングを実現した他の実施例の構造である。

【図16】復号化器とレンダラとの間の伝送計画を決定して伝送する構造を示す図である。

【図17】22.2チャンネルシステムにおいて、全面配置スピーカのうち、ディスプレイによって不在のスピーカをその周辺チャンネルを用いて再生する概念を説明するための概念図である。

【図18】本発明による不在スピーカ位置への音源配置のための処理方法の一実施例である。

【図19】各バンドで生成された信号をテレビの周辺に配置されたスピーカとマッピングさせる一実施例である。 40

【図20】本発明の一実施例によるオーディオ信号処理装置が実現された製品の関係を示す図である。

**【発明を実施するための形態】****【0013】**

本発明の一態様によれば、複数のオブジェクト信号を定められた方法によって分類した第1オブジェクト信号グループと第2オブジェクト信号グループを生成するステップと、第1オブジェクト信号グループに対して第1ダウンミックス信号を生成するステップと、第2オブジェクト信号グループに対して第2ダウンミックス信号を生成するステップと、第1オブジェクト信号グループに含まれたオブジェクト信号に対して、第1ダウンミック 50

ス信号に対応して第1オブジェクト抽出情報を生成するステップと、第2オブジェクト信号グループに含まれたオブジェクト信号に対して、第2ダウンミックス信号に対応して第2オブジェクト抽出情報を生成するステップとを含むオーディオ信号処理方法が提供できる。

【0014】

ここで、前記オーディオ信号処理方法は、前記第1オブジェクト信号グループと第2オブジェクト信号グループは、混合されて1つのサウンド場面をなす信号をさらに含むことができる。

【0015】

また、前記オーディオ信号処理方法は、前記第1オブジェクト信号グループと第2オブジェクト信号グループは、同じ時間に再生される信号で構成できる。 10

【0016】

本発明において、第1オブジェクト信号グループと第2オブジェクト信号グループは、1つのオブジェクト信号のビット列に符号化できる。

【0017】

ここで、前記第1ダウンミックス信号を生成するステップは、前記第1オブジェクト信号グループに含まれたオブジェクト信号に対して各オブジェクト別ダウンミックスゲイン情報を適用して得られ、前記オブジェクト別ダウンミックスゲイン情報は、前記第1オブジェクト抽出情報に含まれる。

【0018】

ここで、前記オーディオ信号処理方法は、前記第1オブジェクト抽出情報と前記第2オブジェクト抽出情報を符号化するステップをさらに含むことができる。 20

【0019】

本発明において、前記オーディオ信号処理方法は、前記第1オブジェクト信号グループと第2オブジェクト信号グループを含む前記オブジェクト信号全体に対するグローバルゲイン情報を生成するステップをさらに含み、前記グローバルゲイン情報は、前記オブジェクト信号のビット列に符号化できる。

【0020】

本発明の他の態様によれば、第1ダウンミックス信号と第2ダウンミックス信号を含む複数のダウンミックス信号を受信するステップと、第1ダウンミックス信号に対応する、第1オブジェクト信号グループに対する第1オブジェクト抽出情報を受信するステップと、第2ダウンミックス信号に対する第2オブジェクト信号グループに対する第2オブジェクト抽出情報を受信するステップと、第1ダウンミックス信号と第1オブジェクト抽出情報を用いて、第1オブジェクト信号グループに属するオブジェクト信号を生成するステップと、第2ダウンミックス信号と第2オブジェクト抽出情報を用いて、第2オブジェクト信号グループに属するオブジェクト信号を生成するステップとを含むオーディオ信号処理方法が提供できる。 30

【0021】

ここで、前記オーディオ信号処理方法は、前記第1オブジェクト信号グループに属するオブジェクト信号のうちの少なくとも1つのオブジェクト信号と、前記第2オブジェクト信号グループに属する少なくとも1つのオブジェクト信号を用いて出力オーディオ信号を生成するステップをさらに含むことができる。 40

【0022】

ここで、前記第1オブジェクト抽出情報と第2オブジェクト抽出情報は、1つのビット列から受信することができる。

【0023】

また、前記オーディオ信号処理方法は、第1オブジェクト抽出情報から第1オブジェクト信号グループに属した少なくとも1つのオブジェクト信号に対するダウンミックスゲイン情報が得られ、前記ダウンミックスゲイン情報を用いて前記少なくとも1つのオブジェクト信号を生成することができる。 50

**【 0 0 2 4 】**

さらに、前記オーディオ信号処理方法は、グローバルゲイン情報を受信するステップをさらに含み、前記グローバルゲイン情報は、前記第1オブジェクト信号グループと前記第2オブジェクト信号グループにすべて適用されるゲイン値であってよい。

**【 0 0 2 5 】**

また、前記第1オブジェクト信号グループに属するオブジェクト信号のうちの少なくとも1つのオブジェクト信号と、前記第2オブジェクト信号グループに属する少なくとも1つのオブジェクト信号は、同じ時間帯に再生できる。

**【 0 0 2 6 】**

本明細書に記載された実施例は、本発明の属する技術分野における通常の知識を有する者に本発明の思想を明確に説明するためのものであるので、本発明が本明細書に記載された実施例によって限定されるものではなく、本発明の範囲は、本発明の思想を逸脱しない修正例または変形例を含むと解釈されなければならない。10

**【 0 0 2 7 】**

本明細書で使用される用語と添付した図面は、本発明を容易に説明するためのものであり、図面に示された形状は、必要に応じて本発明の理解のために誇張されて表示されたものであるので、本発明が本明細書で使用される用語と添付した図面によって限定されるものではない。

**【 0 0 2 8 】**

本明細書において、本発明にかかる公知の構成または機能に対する具体的な説明が本発明の要旨をあいまいにし得ると判断された場合に、これに関する詳細な説明は必要に応じて省略する。20

**【 0 0 2 9 】**

本発明において、次の用語は次のような基準により解釈され得、記載されていない用語でも下記の趣旨によって解釈され得る。コーディングは、場合によって、エンコーディングまたはデコーディングと解釈され得、情報(information)は、値(values)、パラメータ(parameter)、係数(coefficients)、成分(elements)などをすべて網羅する用語であって、場合によって、意味は異なつて解釈できるが、本発明はこれに限定されない。

**【 0 0 3 0 】**

以下、本発明の実施例によるオブジェクトオーディオ信号の処理方法および装置について説明する。

**【 0 0 3 1 】**

図1は、同じ視聴距離上で映像の大きさ(例:UHDTVおよびHDTV)に応じた視聴角度を説明するための図である。ディスプレイの作製技術が発展し、消費者の要求に応じて映像の大きさが大型化する傾向にある。図1に示されているように、HDTV(1920\*1080ピクセルの映像)の場合に比べて、UHDTV(7680\*4320ピクセルの映像)は、約16倍大きくなった映像である。HDTVが居間の壁面に設置され、視聴者が一定の視聴距離をおいて居間のソファーに座った場合、視聴角度が約30度となり得る。しかし、同じ視聴距離でUHDTVが設置された場合、視聴角度は約100度に達する。このように高画質高解像度の大型スクリーンが設置された場合、この大型コンテンツにふさわしいような高い実在感と臨場感を有するサウンドが提供されることが好ましい。視聴者がまるで現場にいるのとほぼ同じ環境を提供するためには、1-2個のサラウンドチャンネルスピーカーが存在するだけでは足りないことがある。したがって、より多いスピーカおよびチャンネル数を有するマルチチャンネルオーディオ環境が要求されることがある。40

**【 0 0 3 2 】**

前記説明のように、ホームシアター環境のほか、個人3Dテレビ(personal 3DTV)、スマートフォンテレビ、22.2チャンネルのオーディオプログラム、自動車、3D video、テレプレゼンスルーム(telopresence room)50

、クラウドベースのゲーム( c l o u d - b a s e d g a m i n g )などがあり得る。

#### 【0033】

図2は、マルチチャンネルの一例として22.2chのスピーカ配置を示す図である。22.2chは音場感を高めるためのマルチチャンネル環境の一例であってよいし、本発明は、特定のチャンネル数または特定のスピーカ配置に限定されない。図2を参照すれば、最も高いレイヤ(t o p l a y e r)1010に計9個のチャンネルが提供できる。前面に3個、中間位置に3個、サラウンド位置に3個の計9個のスピーカが配置されていることが分かる。中間レイヤ(m i d d l e l a y e r)1020には、前面に5個、中間位置に2個、サラウンド位置に計3個のスピーカが配置できる。前面の5個のスピーカのうち、中央位置の3個は、テレビスクリーン内に含まれてよい。底(b o t t o m l a y e r)1030には、前面に計3個のチャンネルおよび2個のLFEチャンネル1040が設けられてよい。  
10

#### 【0034】

このように、最大数十個のチャンネルに達するマルチチャンネル信号を伝送し再生するにあたり、高い演算量が必要であり得る。また、通信環境などを考慮する時、高い圧縮率が要求されることがある。それだけでなく、一般家庭では、マルチチャンネル(例: 22.2ch)のスピーカ環境を備える場合は多くなく、2chまたは5.1chのセットアップを有する聴取者が多いため、すべてのユーザに共通に伝送する信号がマルチチャンネルをそれぞれエンコーディングして送る場合には、そのマルチチャンネルを2chおよび5.1chに再変換して再生しなければならない場合、通信的な非効率が発生するだけでなく、22.2chのPCM信号を格納しなければならないため、メモリ管理における非効率が発生することがある。  
20

#### 【0035】

図3は、聞き手110が3Dオーディオを聴取する聴取空間上130で3次元のサウンド場面を構成する各サウンドオブジェクト120の位置を示す概念図である。図3を参照すれば、図式化の便宜上、各オブジェクト120が点ソース(p o i n t s o u r c e)であると示したが、点ソースのほか、平面波(p l a i n w a v e)形態の音源や、アンビエント(a m b i e n t)音源(サウンド場面の空間を認識できる全方位に行き渡っている余音)などもあり得る。

#### 【0036】

図4は、図3の図式化されたオブジェクトに対して、本発明によるグルーピング方法を用いてオブジェクト信号グループ410、420を形成したことを示す。本発明によれば、オブジェクト信号に対する符号化あるいは処理をする上で、オブジェクト信号グループを形成して、グルーピングされたオブジェクトを単位で符号化したり処理することが特徴である。この時、符号化の場合、オブジェクトを個別信号として独立符号化(d i s c r e t e c o d i n g)する場合や、オブジェクト信号に対するパラメトリック符号化をする場合を含む。特に、本発明によれば、オブジェクト信号に対するパラメータ符号化のためのダウンミックス信号の生成と、ダウンミックスに対応するオブジェクトのパラメータ情報を生成するにあたり、グルーピングされたオブジェクトを単位として生成することが特徴である。すなわち、従来の例として、S A O C 符号化技術の場合、サウンド場面を構成するすべてのオブジェクトを1つのダウンミックス信号(この時、ダウンミックス信号は、モノ(1チャンネル)、あるいはステレオ(2チャンネル)であってよいが、便宜上、1つのダウンミックス信号と表現する)とそれに対応するオブジェクトパラメータ情報と表現したが、このような方法を、本発明で考慮するシナリオのように、20個のオブジェクト以上、多くて200個、500個を、1つのダウンミックスとそれに対応するパラメータで表現する場合、所望の水準の音質を提供するアップミックスおよびレンダリングが事実上不可能である。これによって、本発明では、符号化対象になるオブジェクトをグループ化し、グループ単位でダウンミックスを生成する方法を利用する。グループ単位でダウンミックスされる過程で、各オブジェクトがダウンミックスされる時、ダウンミックスゲインが適用可能であり、適用されたオブジェクト別ダウンミックスゲインは、付加  
40  
50

情報として各グループに対するビット列に含まれる。一方、符号化の効率性あるいは全体ゲインに対する効果的な制御のために、各グループに共通に適用されるグローバルゲインと各グループ別オブジェクトに限定して適用されるオブジェクトグループゲインが使用可能であり、これらは符号化されてビット列に含まれて受信段に伝送される。

#### 【 0 0 3 7 】

グループを形成する第一の方法は、サウンド場面上で、各オブジェクトの位置を考慮して、近いオブジェクト同士でグループを形成する方法である。図4のオブジェクトグループ410、420は、このような方法で形成した一例である。これは、パラメータ符号化的不完全性により各オブジェクトの間に発生するクロストーク歪みや、オブジェクトを第3の位置に移動したり大きさを変更するレンダリングを行う時に発生する歪みが、聞き手110になるべく聞こえないようにするための方法である。同じ位置にあるオブジェクトに発生した歪みは、相対的にマスキングによって聞き手に聞こえない可能性が高い。同じ理由で個別符号化をする場合も、空間的に類似の位置にあるオブジェクト間のグルーピングにより付加情報を共有するなどの効果を期待することができる。10

#### 【 0 0 3 8 】

図5は、本発明の一実施例によるオブジェクトオーディオ信号符号化器500を示すブロック図である。図示のように、オブジェクトオーディオ信号符号化器500は、オブジェクトグルーピング部550と、ダウンミキサおよびパラメータ符号化器520、540とを含むことができる。オブジェクトグルーピング部550は、本発明の実施例により、複数のオブジェクトをグルーピングして、少なくとも1つのオブジェクト信号グループを生成する。図5の実施例では、第1オブジェクト信号グループ510および第2オブジェクト信号グループ530が生成されたことを示したが、本発明の実施例において、オブジェクト信号グループの数はこれに限定しない。この時、各オブジェクト信号グループは、図4の例で説明した方法のように空間的な類似度を考慮して生成されてもよく、音色、周波数分布、音圧などの信号特性によって区分して生成されてもよい。ダウンミキサおよびパラメータ符号化器520、540は、生成された各グループごとにダウンミックスを行い、この過程でダウンミックスされたオブジェクトを復元するのに必要なパラメータを生成する。各グループごとに生成されたダウンミックス信号は、AAC、MP3のようなチャンネル別ウェーブフォーム(waveform)を符号化するウェーブフォーム符号化器560を介して追加的に符号化される。これを一般にコアコードック(Core codec)と呼ぶ。また、各ダウンミックス信号間のカップリングなどによる符号化が行われてよい。各符号化器520、540、560を介して生成された信号は、MUX570を介して1つのビット列として形成されて伝送される。したがって、ダウンミキサおよびパラメータ符号化器520、540とウェーブフォーム符号化器560を介して生成されたビット列は、すべて1つのサウンド場面をなす構成オブジェクトを符号化した信号と見なすことができる。また、生成されたビット列内の互いに異なるオブジェクトグループに属したオブジェクト信号は、同じ時間フレームを有して符号化され、したがって、同じ時間帯に再生される特徴を有したりする。一方、オブジェクトグルーピング部550で生成したグルーピング情報は、符号化されて受信段に伝達されることが可能である。2030

#### 【 0 0 3 9 】

図6は、本発明の一実施例によるオブジェクトオーディオ信号復号化器600を示すブロック図である。オブジェクトオーディオ信号復号化器600は、図5の実施例により符号化されて伝送された信号を復号化することができる。復号化過程は、符号化の逆過程であって、DEMUX610は、符号化器からビット列を受信し、ビット列から少なくとも1つのオブジェクトパラメータセットとウェーブフォーム符号化された信号を抽出する。仮に、図5のオブジェクトグルーピング部550が生成したグルーピング情報がビット列に含まれる場合、DEMUX610は、ビット列から当該グルーピング情報を抽出することができる。ウェーブフォーム復号化器620は、ウェーブフォーム復号化を行って複数のダウンミックス信号を生成し、生成された複数のダウンミックス信号は、それぞれ対応するオブジェクトパラメータセットとともに、アップミキサおよびパラメータ復号化器64050

30、650に入力される。アップミキサおよびパラメータ復号化器630、650は、入力されたダウンミックス信号をそれぞれアップミキシングして、少なくとも1つのオブジェクト信号グループ640、660に復号化する。この時、各オブジェクト信号グループ640、660の復元には、ダウンミックス信号およびこれに対応するオブジェクトパラメータセットが用いられる。図6の実施例において、複数のダウンミックス信号が存在するため、複数のパラメータ復号化が必要である。図6では、第1ダウンミックス信号および第2ダウンミックス信号がそれぞれ第1オブジェクト信号グループ640および第2オブジェクト信号グループ660に復号化されることを示したが、本発明の実施例で抽出されたダウンミックス信号の個数および対応するオブジェクト信号グループの数はこれに限定しない。一方、オブジェクトデグルーピング部670は、グルーピング情報を用いて、各オブジェクト信号グループを個別オブジェクト信号にデグルーピングすることができる。10

#### 【0040】

本発明の実施例によれば、伝送されたビット列にグローバルゲインおよびオブジェクトグループゲインが含まれている場合、これらを適用して正常なオブジェクト信号の大きさを復元することができる。一方、レンダリングあるいはトランスコーディング過程でこのゲイン値は制御が可能であり、グローバルゲインの調整により全体信号の大きさを、オブジェクトグループゲインの調整によりグループ別信号の大きさを調整することができる。例えば、再生スピーカ単位でオブジェクトグルーピングが行われた場合、後述の柔軟なレンダリングを実現するためにゲインを調整する時、オブジェクトグループゲインの調整により容易に実現することができる。20

#### 【0041】

図5および図6において、複数のパラメータ符号化器あるいは復号化器は、説明の便宜上、並列に処理されるように示されたが、1つのシステムを介して順次に複数のオブジェクトグループに対する符号化あるいは復号化を行うことも可能である。

#### 【0042】

オブジェクトグループを形成する他の方法は、互いに相関度の低いオブジェクト同士で1つのグループにグルーピングする方法である。これは、パラメータ符号化の特徴で、相関度の高いオブジェクトはダウンミックスからそれを分離しにくい特徴を考慮したものである。この時、ダウンミックス時、ダウンミックスゲインなどのパラメータを調整して、グルーピングされた各オブジェクトがより相関性が遠くなるようにする符号化方法も可能である。この時、使用されたパラメータは、復号化時、信号の復元に使用できるように伝送されることが好ましい。30

#### 【0043】

オブジェクトグループを形成するさらに他の方法は、互いに相関度の高いオブジェクトを1つのグループにグルーピングする方法である。これは、相関度の高いオブジェクトの場合、パラメータを用いた分離に困難があるが、そのような活用度が高くない応用で圧縮効率を高めるための方法である。多様なスペクトルを有する複雑な信号の場合、コアコーデックでの信号処理にそれだけビットを多く要するため、相関度の高いオブジェクトをまとめて1つのコアコーデックを活用すれば符号化効率が高い。40

#### 【0044】

オブジェクトグループを形成するさらに他の方法は、オブジェクト間のマスキングの有無を判断して符号化することである。例えば、オブジェクトAがオブジェクトBをマスキングする関係にある場合、2つの信号を1つのダウンミックスに含めてコアコーデックで符号化すると、オブジェクトBは符号化過程で省略されてよい。この場合、復号化段でパラメータを用いてオブジェクトBを得る場合、歪みが大きい。したがって、このような関係を有するオブジェクトAとオブジェクトBは、別のダウンミックスに含めることができない。反面、オブジェクトAとオブジェクトBがマスキングの関係にあるものの、2つのオブジェクトを分離してレンダリングする必要がない応用や、少なくともマスキングされたオブジェクトに対する別途処理の必要がない場合は、逆にオブジェクトAとBを1つの50

ダウンミックスに含ませることが好ましい。したがって、応用に応じて選択方法が異なり得る。例えば、符号化過程において、好ましいサウンド場面上で特定のオブジェクトがマスキングされて無くなったり、少なくとも微弱な場合であれば、これをオブジェクトリストから除外し、マスカーになるオブジェクトに含ませたり、2つのオブジェクトを合わせて1つのオブジェクトとして表現する方式で実現することができる。

#### 【0045】

オブジェクトグループを形成するさらに他の方法は、平面波ソースオブジェクトやアンビエントソースオブジェクトなどの、点ソースオブジェクトでないものを分離して別途にグループ化することである。このようなソースは、点ソースと異なる特性によって、他の形態の圧縮符号化方法やパラメータが必要であり、したがって、別途に分離して処理することが好ましい。10

#### 【0046】

本発明の実施例によれば、グルーピング情報は、前述のオブジェクトグループが形成された方法に関する情報を含むことができる。オーディオ信号復号化器は、伝送されたグルーピング情報を参照して、復号化されたオブジェクト信号グループを元のオブジェクトに還元するオブジェクトデグルーピングを行うことができる。

#### 【0047】

図7は、本発明による符号化方法によって符号化して生成したビット列の一実施例である。図7を参照すれば、符号化されたチャンネルあるいはオブジェクトデータが伝送される主ビット列700が、チャンネルグループ720、730、740、あるいはオブジェクトグループ750、760、770の順に整列されていることが分かる。各チャンネルグループには、チャンネルグループに属した個別チャンネルが設定された順序によって整列されて配置される。図面符号721、731、751は、それぞれチャンネル1、チャンネル8、チャンネル92の信号を示す例である。また、ヘッダ710に各グループのビット列内での位置情報であるチャンネルグループポジション情報C H G \_ P O S \_ I N F O 711、オブジェクトグループポジション情報O B J \_ P O S \_ I N F O 712を含んでいるため、これを参照すれば、ビット列を順次に復号化しなくても、所望のグループのデータのみを優先復号化することができる。したがって、復号化器は、一般的にグループ単位で先に到着したデータから復号化を行うが、他の政策や理由によって復号化する順序を任意に変更することができる。また、図7は、主ビット列700のほか、別途に、主な復号化関連情報とともに、各チャンネルあるいはオブジェクトに対するメタデータ703、704を盛り込んでいる副ビット列701を例示する。副ビット列は、主ビット列が伝送される途中に間欠的に伝送されたり、別の伝送チャンネルを介して伝送されてよい。一方、チャンネルとオブジェクト信号に統いて、選択的にA N C ( A n c i l l a r y d a t a ) 780が含まれてよい。2030

#### 【0048】

##### (オブジェクトグループ別にビット割当する方法)

複数のグループ別にダウンミックスを生成し、各グループ別に独立したパラメトリックオブジェクト符号化を行うにあたり、各グループで使用されるビット数は互いに異なり得る。グループ別ビットを割り当てる基準は、グループ内に含まれたオブジェクトの数、グループ内のオブジェクト間のマスキング効果を考慮した有効オブジェクト数、人の空間解像度を考慮した位置に応じた加重値、オブジェクトの音圧の大きさ、オブジェクト間相関度、サウンド場面上のオブジェクトの重要度などを考慮することができる。例えば、A、B、Cの3つの空間的オブジェクトグループを有する場合、それぞれグループのo b j e c t信号が3、2、1個ずつ含まれていれば、割り当てられたビットは3 a 1 ( n - x )、2 a 2 ( n - y )、a 3 n に割り当てられてよい。ここで、x, yは、各グループ内でオブジェクト間、そしてオブジェクト内でマスキング効果によってビットをより少なく割り当ててもよい程度を示し、a 1、a 2、a 3は、グループ別に前記言及した多様な要素によって決定できる。40

#### 【0049】

10

20

30

40

50

(オブジェクトグループ内の主オブジェクト、副オブジェクト位置情報の符号化)

一方、オブジェクト情報の場合、プロデューサーが生成した意図によって勧告したり、他のユーザが提案するミックス情報などを、オブジェクトの位置および大きさ情報としてメタデータを介して伝達する手段を有することが好ましい。本発明では、これを、便宜上、プリセット情報と呼ぶ。オブジェクトが時間に応じて位置可変するダイナミックオブジェクトの場合、プリセット情報を介して伝送されるべき位置情報の情報量が少なくない。例えば、1000個のオブジェクトに対して毎フレーム可変する位置情報を伝送すれば、非常に大きいデータ量になる。したがって、オブジェクトの位置情報も効果的に伝送することが好ましい。そこで、本発明では、主オブジェクトと副オブジェクトという定義を利用して、位置情報の効果的な符号化方法を用いる。

10

#### 【0050】

主オブジェクトは、オブジェクトの位置情報を3次元空間上の絶対的な座標値で表現するオブジェクトを意味する。副オブジェクトは、3次元空間上の位置を主オブジェクトに対する相対的な値で表現して、位置情報を有するオブジェクトを意味する。したがって、副オブジェクトの位置情報を知るためには、対応する主オブジェクトが何であるかを知らなければならない。本発明の実施例によれば、グルーピングを行う場合、特に空間上の位置を基準としてグルーピングをする場合、同一グループ内に1つのオブジェクトを主オブジェクト、残りのオブジェクトを副オブジェクトとして位置情報を表現する方法で実現可能である。符号化のためのグルーピングがないか、これを用いることが副オブジェクト位置情報の符号化に有利でない場合、位置情報符号化のための別の集合を形成することができる。副オブジェクト位置情報を相対的に表現することを、絶対値で表現するより有利にするためには、グループあるいは集合内に属するオブジェクトは空間上で一定範囲内に位置することが好ましい。

20

#### 【0051】

本発明の他の位置情報の符号化方法は、主オブジェクトに対する相対的な表現の代わりに、固定されたスピーカ位置に関する相対情報として各オブジェクトの位置情報を表現することである。例えば、22チャンネルスピーカの指定された位置値を基準として、オブジェクトの相対的位置情報を表現する。この時、基準として使用するスピーカの個数と位置値などは、現在のコンテンツで設定した値を参照することができる。

30

#### 【0052】

本発明の他の実施例によれば、位置情報を絶対値あるいは相対値で表現した後、量子化を行うが、量子化ステップは、絶対位置を基準として可変的であることを特徴とする。例えば、聞き手の正面付近は、側面あるいは後面に比べて、位置に対する区別能力がはるかに高いことが知られているため、正面領域に対する解像度は、側面領域に対する解像度より高いように量子化ステップを設定することが好ましい。同様に、人は、方位に対する解像度が、高低に対する解像度より高いため、方位角に対する量子化を、高度に対する解像度より高くすることが好ましい。

#### 【0053】

本発明のさらに他の実施例では、位置が時変するダイナミックオブジェクトの場合、主オブジェクトあるいは他の基準点に対する相対的な位置値を表現する代わりに、当該オブジェクトの前の位置値に対する相対的な値で表現することが可能である。したがって、ダイナミックオブジェクトに対する位置情報は、時間的に前、空間的に隣りの基準点のうちのいずれを基準としたかを区別するためのフラグ情報をともに伝送することが好ましい。

40

#### 【0054】

(復号化器全体のアーキテクチャ)

図8は、本発明によるオブジェクトおよびチャンネル信号復号化システム800をブロック図に示す一実施例である。システム800は、オブジェクト信号801、チャンネル信号802、あるいはオブジェクト信号とチャンネル信号との組み合わせを受けることができる。また、オブジェクト信号あるいはチャンネル信号は、それぞれウェーブフォーム符号化(801、802)されたり、パラメトリック符号化(803、804)されてい

50

てよい。復号化システム 800 は、大きく、3DA復号化部 860 と、3DAレンダリング部 870 とに区分され、3DAレンダリング部 870 は、任意の外部システムあるいはソリューションが使用されてもよい。したがって、3DA復号化部 860 と 3DAレンダリング部 870 は、外部と容易に互換される標準化されたインターフェースを提供することが好ましい。

#### 【0055】

図9は、本発明のさらに他の形態のオブジェクトおよびチャンネル信号復号化システム 900 のブロック図である。同様に、本システム 900 は、オブジェクト信号 901 、チャンネル信号 902 、あるいはオブジェクト信号とチャンネル信号との組み合わせを受けることができる。また、オブジェクト信号あるいはチャンネル信号は、それぞれウェーブフォーム符号化 (901、902) されたり、パラメトリック符号化 (903、904) されていてよい。図8のシステム 800 と比較する時、相違点は、図9の復号化システム 900 では、それぞれ分離されていた個別オブジェクト復号化器 810 と個別チャンネル復号化器 820 、そしてパラメトリックチャンネル復号化器 840 とパラメトリックオブジェクト復号化器 830 が、それぞれ1つの個別復号化器 910 とパラメトリック復号化器 920 に統合された点である。また、図9の復号化システム 900 には、3DAレンダリング部 940 と、便利かつ標準化されたインターフェースのためのレンダリングインターフェース部 930 が追加された。レンダリングインターフェース部 930 は、内部あるいは外部に存在する3DAレンダラ 940 からユーザ環境情報、レンダラバージョンなどを受信し、これに互換される形態のチャンネル信号あるいはオブジェクト信号を生成して、3DA Renderer 940 に伝達する役割を果たす。また、チャンネル数、オブジェクト別の名前などのように再生に必要な付加情報をユーザに提供するために必要なメタデータを標準化された形式で生成して、3DA Renderer 940 に伝達することができる。レンダリングインターフェース部 930 は、後述の順序制御部 1630 を含むことができる。

#### 【0056】

パラメトリック復号化器 920 は、オブジェクト信号あるいはチャンネル信号を生成するためにダウンミックス信号が必要であるが、必要なダウンミックス信号は、個別復号化器 910 を介して復号化されて入力される。オブジェクトおよびチャンネル信号復号化システムに対応する符号化器は、様々なタイプになっていてよいし、図8および図9に表現された形態のビット列 801、802、803、804、901、902、903、904 のうちの少なくとも1つを生成できれば、互換される符号化器と見なすことができる。また、本発明によれば、図8および図9に提示された復号化システムは、過去のシステムあるいはビット列との互換性を保障するようにデザインされた。例えば、AACで符号化された個別チャンネルのビット列が入力された場合、当該ビット列は、個別(チャンネル)復号化器を介して復号化されて、3DAレンダラに送られてよい。MPS(MPEG Surround)ビット列の場合、ダウンミックス信号とともに送られるが、ダウンミックスされた後、AACで符号化された信号は、個別(チャンネル)復号化器を介して復号化されてパラメトリックチャンネル復号化器に伝達され、パラメトリックチャンネル復号化器は、まるでMPEG Surround復号化器のように動作する。SAOC(Spatial Audio Object Coding)で符号化されたビット列の場合も同様に動作する。図8のシステム 800 において、SAOCビット列は、従来のように、SAOCトランスコーダ 830 でトランスコーディングされた後、MPEG Surroundデコーダ 840 を介して個別チャンネルにレンダリングされる構造を有する。このためには、SAOCトランスコーダ 830 は、再生チャンネル環境情報を受けて、これに合うように最適化されたチャンネル信号を生成して伝送することが好ましい。したがって、本発明のオブジェクトおよびチャンネル信号復号化システムは、従来のSAOCビット列を受けて復号化するが、ユーザあるいは再生環境に特化したレンダリングを行うことができる。図9のシステム 900 では、SAOCビット列が入力される場合、MPSビット列に変換するトランスコーディング動作の代わりに、直ちにチャンネルあるいはレンダリングに適した個別オブジェクト形態に変換する方法で実現される。したがって、トラ

ンスコーディングする構造に比べて演算量が低く、音質の面においても有利である。図9において、オブジェクト復号化器の出力を channel でのみ表示したが、個別オブジェクト信号としてレンダラインタフェース 930 に伝達されてもよい。また、図9でのみ表記されたが、図8の場合を含めて、パラメトリックビット列上にレジデュアル信号が含まれた場合、これに対する復号化は個別復号化器を介して復号化されることが特徴である。

#### 【0057】

(チャンネルに対する個別、パラメータ組み合わせ、レジデュアル)

図10は、本発明の他の実施例によるエンコーダおよびデコーダの構成を示す図である。

10

#### 【0058】

図10は、デコーダのスピーカセットアップがそれぞれ異なる場合に、スケーラブルなコーディングのための構造を示す。

#### 【0059】

エンコーダは、ダウンミキシング部 210 を含み、デコーダは、第1デコーディング部 230～第3デコーディング部 250 のうちの1つ以上と、デマルチプレキシング部 220 を含む。

#### 【0060】

ダウンミキシング部 210 は、マルチチャンネルに相当する入力信号 (CH\_N) をダウンミキシングすることにより、ダウンミックス信号 (DMX) を生成する。この過程で、アップミックスパラメータ (UP) およびアップミックスレジデュアル (UR) のうちの1つ以上を生成する。その後、ダウンミックス信号 (DMX)、アップミックスパラメータ (UP) (およびアップミックスレジデュアル (UR)) をマルチプレキシングすることにより、1以上のビットストリームを生成し、デコーダに伝送する。

20

#### 【0061】

ここで、アップミックスパラメータ (UP) は、1以上のチャンネルを2以上のチャンネルにアップミキシングするために必要なパラメータであって、空間パラメータおよびチャンネル間位相差 (IPD) などが含まれてよい。

#### 【0062】

そして、アップミックスレジデュアル (UR) は、原信号の入力信号 (CH\_N) と復元された信号との差であるレジデュアル信号に相当する。ここで、復元された信号は、ダウンミックス信号 (DMX) にアップミックスパラメータ (UP) を適用してアップミキシングされた信号であってもよいし、ダウンミキシング部 210 によってダウンミキシングされていないチャンネルが discrete な方式でエンコーディングされた信号であってよい。

30

#### 【0063】

デコーダのデマルチプレキシング部 220 は、1以上のビットストリームからダウンミックス信号 (DMX) およびアップミックスパラメータ (UP) を抽出し、アップミックスレジデュアル (UR) をさらに抽出することができる。ここで、レジデュアル信号は、ダウンミックス信号に対する個別符号化と類似の方法で符号化できる。したがって、レジデュアル信号の復号化は、図8あるいは図9に提示されたシステムでは、個別 (チャンネル) 復号化器を介して行われることが特徴である。

40

#### 【0064】

デコーダは、スピーカセットアップ環境に応じて、第1デコーディング部 230～第3デコーディング部 250 のうちの1つ (または1つ以上) を選択的に含むことができる。デバイスの種類 (スマートフォン、ステレオテレビ、5.1ch ホームシアター、22.2ch ホームシアターなど) に応じて、ラウドスピーカのセットアップ環境が多様であり得る。このように多様な環境にもかかわらず、22.2chなどのマルチチャンネル信号を生成するためのビットストリームおよびデコーダが選択的でなければ、22.2ch の信号をすべて復元した後に、スピーカの再生環境に応じて、再びダウンミックスしなけれ

50

ばならない。この場合、復元およびダウンミックスに必要な演算量が非常に高いだけでなく、遅延が発生することもある。

#### 【0065】

しかし、本発明の他の実施例によれば、各デバイスのセットアップ環境に応じて、デコーダが第1デコーディング部～第3デコーディング部のうちの1つ（または1つ以上）を選択的に備えることにより、前記のような不都合を解消することができる。

#### 【0066】

第1デコーディング部230は、ダウンミックス信号（DMX）のみをデコーディングする構成であって、チャンネル数の増加を伴わない。すなわち、第1デコーディング部230は、ダウンミックス信号がモノの場合、モノチャンネル信号を出力し、ステレオの場合、ステレオ信号を出力するものである。第1デコーディング部230は、スピーカチャンネル数が1つまたは2つの、ヘッドホンが備えられた装置、スマートフォン、テレビなどに好適であり得る。10

#### 【0067】

一方、第2デコーディング部240は、ダウンミックス信号（DMX）およびアップミックスパラメータ（UP）を受信し、これに基づいて、パラメトリックMチャンネル（PM）を生成する。第2デコーディング部240は、第1デコーディング部230に比べて出力チャンネル数が増加する。しかし、アップミックスパラメータ（UP）が総Mチャンネルまでのアップミックスに相当するパラメータのみが存在する場合、第2デコーディング部240は、原チャンネル数（N）に及ばないMチャンネル数の信号を出力することができる。例えば、エンコーダの入力信号である原信号が22.2ch信号であり、Mチャンネルは5.1ch、7.1chチャンネルなどであってよい。20

#### 【0068】

第3デコーディング部250は、ダウンミックス信号（DMX）およびアップミックスパラメータ（UP）だけでなく、アップミックスレジデュアル（UR）まで受信する。第2デコーディング部240は、Mチャンネルのパラメトリックチャンネルを生成するのに対し、第3デコーディング部250は、これにアップミックスレジデュアル信号（UR）まで追加的に適用することにより、N個のチャンネルの復元された信号を出力することができる。30

#### 【0069】

各デバイスは、第1デコーディング部～第3デコーディング部のうちの1つ以上を選択的に備え、ビットストリームの中からアップミックスパラメータ（UP）およびアップミックスレジデュアル（UR）を選択的にパーシングすることにより、各スピーカセットアップ環境に合わせた信号を直ちに生成することにより、複雑度および演算量を低減することができる。

#### 【0070】

（マスキングを考慮したオブジェクトのウェーブフォーム符号化）

本発明によるオブジェクトのウェーブフォーム符号化器（以下、ウェーブフォーム（waveform）符号化器は、チャンネルオーディオ信号あるいはオブジェクトオーディオ信号を各チャンネルあるいはオブジェクト別に独立して復号化可能に符号化する場合をいい、パラメトリック符号化／復号化に相対する概念で、さらに個別（discrete）符号化／復号化とも呼ぶ）は、オブジェクトのサウンド場面上の位置を考慮してビット割当する。これは、心理音響のBMLD（Binaural Masking Level Difference）現象とオブジェクト信号符号化の特徴を利用したものである。40

#### 【0071】

BMLD現象を説明するために、既存のオーディオ符号化方法で使用していたMS（Mid-Side）ステレオ符号化を説明すれば、次の通りである。すなわち、心理音響におけるマスキング現象は、マスキングを発生させるマスカー（masker）と、マスキングになるマスキーラー（maskee）とが空間的に同一方向にある時に可能であるのがB50

M L D である。ステレオオーディオ信号の 2 チャンネルのオーディオ信号の間の相関性が非常に高く、その大きさが等しい場合、その音に対する像（音像）が 2 つのスピーカの間の中央に結ばれ、相関性がない場合、各スピーカから独立した音が出て、その像がそれぞれスピーカに結ばれる。仮に、相関性が最大である入力信号に対して各チャネルを独立して符号化 (d u a l m o n o) する場合、オーディオ信号の音像は中央に、量子化雑音の音像は各スピーカに別に結ばれることとなる。すなわち、各チャネルにおける量子化雑音は互いに相関性がないため、その像が各スピーカに別に結ばれるのである。したがって、マスキーになるべき量子化雑音が空間的不一致によってマスキングされず、結局、人に歪みとして聞こえる問題が発生する。和差符号化は、このような問題を解決すべく、2 つのチャネル信号を加えた信号 (M i d 信号) と引いた信号 (D i f f e r e n c e) を生成した後、これを用いて心理音響モデルを行い、これを用いて量子化する。このような方法によれば、発生した量子化雑音の音像がオーディオ信号の音像と同じ位置に結ばれる。

#### 【 0 0 7 2 】

従来のチャネル符号化の場合、各チャネルは、再生されるスピーカにマッピングされ、当該スピーカの位置は固定されて互いに離れているため、チャネル間のマスキングは考慮できなかった。しかし、各オブジェクトを独立して符号化する場合は、当該オブジェクトのサウンド場面上の位置に応じてマスキングされるか否かが異なり得る。したがって、他のオブジェクトによって、現在符号化されるオブジェクトがマスキングされるか否かを判断して、それによってビットを割り当てて符号化することが好ましい。

#### 【 0 0 7 3 】

図 1 1 は、オブジェクト 1 とオブジェクト 2 に対するそれぞれの信号と、これら信号から取得できるマスキング閾値 1 1 1 0 、 1 1 2 0 と、オブジェクト 1 とオブジェクト 2 を合わせた信号に対するマスキング閾値 1 1 3 0 を示す。オブジェクト 1 とオブジェクト 2 が少なくとも聞き手の位置を基準として同一の位置、あるいは B M L D の問題が発生しないだけの範囲内に位置すると見なすと、聞き手に当該信号によってマスキングされる領域は 1 1 3 0 のようになるので、オブジェクト 1 に含まれた S 2 信号は完全にマスキングされて聞こえない信号となるはずである。そのため、オブジェクト 1 を符号化する過程において、オブジェクト 2 に対するマスキング閾値を考慮して符号化することが好ましい。マスキングス閾値は互いに加算的に合わされる性質があるので、結局、オブジェクト 1 とオブジェクト 2 に対するそれぞれのマスキング閾値を加える方法で求めることができる。あるいは、マスキング閾値を計算する過程自体も演算量が非常に高いので、オブジェクト 1 とオブジェクト 2 とを予め合わせて生成した信号を用いて、1 つのマスキング閾値を計算してオブジェクト 1 とオブジェクト 2 をそれぞれ符号化することも好ましい。

#### 【 0 0 7 4 】

図 1 2 は、図 1 1 のような例示の内容を実現するために、本発明による複数のオブジェクト信号に対するマスキング閾値を算出する符号化器 1 2 0 0 の一実施例である。2 つのオブジェクト信号が入力されると、これに対する S U M 1 2 1 0 で和信号を生成する。和信号を入力として、心理音響モデル演算部 1 2 3 0 では、オブジェクト 1 、オブジェクト 2 に対応するマスキング閾値をそれぞれ算出する。この時、図 1 2 に示さないが、心理音響モデル演算部 1 2 3 0 の入力として、和信号のほか、オブジェクト 1 とオブジェクト 2 の信号も追加的に提供できる。生成されたマスキング閾値 1 を用いてオブジェクト信号 1 に対するウェーブフォーム符号化 1 2 2 0 が行われ、符号化されたオブジェクト信号 1 を出力し、マスキング閾値 2 を用いてオブジェクト信号 2 に対するウェーブフォーム符号化 1 2 4 0 が行われ、符号化されたオブジェクト信号 2 を出力する。

#### 【 0 0 7 5 】

本発明の他のマスキング閾値算出方法は、2 つのオブジェクト信号の位置が聴音覚を基準として完全に一致しない場合、2 つのオブジェクトに対するマスキング閾値を加える代わりに、2 つのオブジェクトが空間上に離れた程度を考慮してマスキングレベルを減衰して反映することも可能である。すなわち、オブジェクト 1 に対するマスキング閾値を M 1

( f )、オブジェクト 2 に対するマスキング閾値を M 2 ( f )とする時、各オブジェクトを符号化するのに使用する最終ジョイントマスキング閾値 M 1 ' ( f )、M 2 ' ( f )は、次のような関係を有するように生成される。

【 0 0 7 6 】

【 数 1 】

$$M1'(f) = M1(f) + A(f)M2(f)$$

【 数 2 】

$$M2'(f) = A(f)M1(f) + M2(f)$$

10

【 0 0 7 7 】

この時、A ( f )は、2つのオブジェクトの間の空間上の位置と距離、および2つのオブジェクトの属性などにより生成される減衰ファクターであって、0 . 0 = < A ( f ) = < 1 . 0 の範囲を有する。

【 0 0 7 8 】

人の方向に対する解像度は、正面を基準として左右へいくほど悪くなり、後方へいく時にさらに悪くなる特性を有するが、したがって、オブジェクトの絶対的位置は、A ( f )を決定するさらに他の要素として作用することができる。

【 0 0 7 9 】

本発明の他の実施例では、2つのオブジェクトのうちの1つのオブジェクトに対しては自身のマスキング閾値のみを用い、他のオブジェクトに対してのみ相手オブジェクトに対するマスキング閾値を持ってくる方法で実現することができる。これを、それぞれ独立オブジェクト、依存オブジェクトという。自分自身のマスキング閾値のみを利用するオブジェクトは、相手オブジェクトに関係なく高音質符号化されるため、当該オブジェクトから空間的に分離されるレンダリングが行われても音質が保存される利点を有することができる。オブジェクト 1 を独立オブジェクト、オブジェクト 2 を依存オブジェクトとすれば、次のような式でマスキング閾値が表現される。

20

【 0 0 8 0 】

【 数 3 】

$$M1'(f) = M1(f)$$

30

【 数 4 】

$$M2'(f) = A(f)M1(f) + M2(f)$$

【 0 0 8 1 】

独立オブジェクトと依存オブジェクトの有無は、各オブジェクトに対する付加情報として復号化およびレンダラに伝達することが好ましい。

【 0 0 8 2 】

本発明のさらに他の実施例では、2つのオブジェクトが空間上で一定程度類似する場合、マスキング閾値のみを合わせて生成するのではなく、信号自体を1つのオブジェクトに合わせて処理することも可能である。

40

【 0 0 8 3 】

本発明のさらに他の実施例では、特にパラメータ符号化を行う場合、2つの信号の相関度と2つの信号の空間上の位置を考慮して、1つのオブジェクトに合わせて処理することが好ましい。

【 0 0 8 4 】

(トランスクーディングの特徴)

本発明のさらに他の実施例では、カップリングされたオブジェクトを含むビット列をより低いビット率にトランスクーディングするために、データの大きさを低減するために、オブジェクトの数字を低減しなければならない場合（すなわち、複数のオブジェクトを1

50

つにダウンミックスして1つのオブジェクトで表現する場合)、カップリングされたオブジェクトに対して1つのオブジェクトで表現することが好ましい。

#### 【0085】

以上のオブジェクト間のカップリングによる符号化を説明するにあたり、説明の便宜のために、2つのオブジェクトのみをカップリングする場合のみを例に挙げたが、2つ以上の多数のオブジェクトに対するカップリングも類似の方法で実現可能である。

#### 【0086】

##### (柔軟なレンダリング必要)

3Dオーディオのために必要な技術のうち、柔軟なレンダリングは、3Dオーディオの品質を最上に引き上げるために解決すべき重要な課題の一つである。居間の構造、家具の配置に応じて、5.1チャンネルスピーカーの位置が非常に不定形的なのは周知の事実である。このような不定形的位置にスピーカーが存在しても、コンテンツ制作者の意図したサウンド場面を提供できるようにしなければならない。このためには、ユーザごとにまちまちの再生環境でのスピーカー環境を知らなければならないとともに、規格に応じた位置対比の差を補正するためのレンダリング技術が必要である。すなわち、伝送されたビット列をデコーディングすることでコードックの役割が終わるのではなく、これをユーザの再生環境に合わせて最適化変形する過程に対する一連の技術が要求される。

10

#### 【0087】

図13は、5.1チャンネルセットアップに対して、ITU-R勧告案によって配置されたスピーカー(灰色)1310と、任意の位置に配置されたスピーカー(白色)1320を示す。実際の居間環境では、このようにITU-R勧告案対比、スピーカーの方向角と距離がすべて異なる問題が発生することがある(図に示さないが、スピーカーの高さにも差があり得る。)。このように異なるスピーカー位置で元のチャンネル信号をそのまま再生する場合、理想的な3Dサウンド場面を提供しにくい。

20

#### 【0088】

##### (フレキシブルレンダリング)

信号の大きさを基準として、2つのスピーカーの間の音源の方向情報を決定するAmplitude Panningや、3次元空間上で3つのスピーカーを用いて音源の方向を決定するのに広く使用されるVBAP(Vector-Based Amplitude Panning)を用いると、オブジェクト別に伝送されたオブジェクト信号に対しては相対的に便利にフレキシブルレンダリングを実現できることが分かる。チャンネルの代わりにオブジェクト信号を伝送することの利点の一つである。

30

#### 【0089】

##### (オブジェクトの復号化とレンダリング構造)

図14は、本発明によるオブジェクトのビット列に対する復号化器とこれを用いたフレキシブルレンダリングシステムが接続された2つの実施例の構造1400、1401を示す。前述のように、オブジェクトの場合、所望のサウンド場面に合わせてオブジェクトを音源に位置させることができるのである利点があり、ここでは、ミックス(Mix)部1420でミキシング行列で表現された位置情報を受信して、優先チャンネル信号に変更する。すなわち、サウンド場面に対する位置情報を、出力チャンネルに対応するスピーカーからの相対的な情報として表現されるのである。この時、実際にスピーカーの個数と位置が定められた位置に存在しない場合、当該位置情報(Speaker Config)を用いて再びレンダリングする過程が必要である。下記に記述するように、チャンネル信号を再び異なる形態のチャンネル信号にレンダリングすることは、オブジェクトを最終チャンネルに直接レンダリングする場合より実現しにくい。

40

#### 【0090】

図15は、本発明によるオブジェクトのビット列に対する復号化とレンダリングを実現した他の実施例の構造1500を示す。図14の場合と比較すると、ビット列から復号化とともに最終スピーカー環境に合わせたフレキシブルレンダリング1510を直接実現することである。すなわち、ミキシング行列に基づいて定形のチャンネルで行うミキシングと

50

、該生成された定形チャンネルからフレキシブルスピーカにレンダリングする過程の2つのステップを経る代わりに、ミキシング行列とスピーカ位置情報1520を用いて1つのレンダリング行列あるいはレンダリングパラメータを生成し、これを用いてオブジェクト信号を対象スピーカに直ちにレンダリングするのである。

#### 【0091】

(チャンネルで付けてフレキシブルレンダリング)

一方、チャンネル信号が入力として伝送され、当該チャンネルに対応するスピーカの位置が任意の位置に変更された場合、オブジェクト信号に対するパンニング手法のような方法が適用されにくく、別のチャンネルマッピングプロセスが必要である。より大きい問題は、このようにオブジェクト信号とチャンネル信号に対してレンダリングのために必要な過程と解決方法が異なるため、オブジェクト信号とチャンネル信号が同時に伝送され、2つの信号をミックスした形態のサウンド場面を演出しようとする場合は、空間の不整合による歪みが発生しやすいというのである。このような問題を解決するために、本発明の他の実施例では、オブジェクトに対するフレキシブルレンダリングを別途に行わず、チャンネル信号にミックスを先に行った後、チャンネル信号に対するフレキシブルレンダリングを行うようとする。HRTFを用いたレンダリングなども同様の方法で実現されることが好ましい。

10

#### 【0092】

(復号化段のダウンミックス：パラメータ伝送あるいは自動生成)

ダウンミックスレンダリングにおいて、マルチチャンネルコンテンツをそれより少ない数の出力チャンネルを介して再生する場合、今までにはM-Nダウンミックスマトリクス(Mは入力チャンネル数、Nは出力チャンネル数)を用いて実現することが一般的であった。すなわち、5.1チャンネルコンテンツをステレオで再生する時、与えられた数式によってダウンミックスを行う方式で実現される。しかし、このようなダウンミックス実現方法は、まず、ユーザの再生スピーカ環境が5.1チャンネルだけであるにもかかわらず、伝送された22.2チャンネルに相当するすべてのビット列を復号化しなければならない演算量の問題が発生する。携帯機器での再生のためのステレオ信号生成のためにも、22.2チャンネル信号をすべて復号化しなければならぬとすれば、その演算量の負担が非常に高いだけでなく、おびただしい量のメモリの無駄使い(22.2チャンネル復号化されたオーディオ信号の格納)が発生する。

20

#### 【0093】

(ダウンミックスの代案としてのトランスコーディング)

これに対する代案として、巨大な22.2チャンネルの原ビット列から効果的なトランスクーディングにより、目標機器あるいは目標の再生空間に適した数のビット列に切り替える方法を考えることができる。例えば、クラウドサーバに格納された22.2チャンネルコンテンツであれば、クライアント端末から再生環境情報を受信し、これに合わせて変換して伝送するシナリオが実現可能である。

30

#### 【0094】

(復号化順序あるいはダウンミックス順序；順序制御部)

一方、復号化器とレンダリングが分離されているシナリオの場合、例えば、22.2チャンネルのオーディオ信号とともに、50個のオブジェクト信号を復号化して、これをレンダラに伝達しなければならない場合が発生することがある。この時、伝送されるオーディオ信号は、復号化が完了した高データ率の信号であるので、復号化器とレンダラとの間に非常に大きい帯域幅を要求する問題がある。したがって、一度にこのように多いデータを同時に伝送することは好ましくなく、効果的な伝送計画を立てることが好ましい。そして、これに合わせて復号化器が復号化順序を決定して伝送することが好ましい。図16は、このように復号化器とレンダラとの間の伝送計画を決定して伝送する構造1600を示すブロック図である。

40

#### 【0095】

順序制御部1630は、ビット列に対する復号化により付加情報を取得し、メタデータ

50

とレンダラ 1620 から再生環境、レンダリング情報などを受信する。次に、順序制御部 1630 は、前記受信された情報を用いて、復号化順序、復号化された信号をレンダラ 1620 に伝送する伝送順序および単位などの統制情報を決定し、決定された統制情報を復号化器 1610 とレンダラ 1620 に再び伝達する。例えば、レンダラ 1620 で特定のオブジェクトを完全に除去するように命令した場合、このオブジェクトは、レンダラ 1620 への伝送が不要な上に、復号化もする必要がない。あるいは、他の例として特定のオブジェクトを特定のチャンネルにのみレンダリングする状況の場合、当該オブジェクトを別途に伝送する代わりに、伝送される当該チャンネルに予めダウンミックスして伝送すれば伝送帯域が減少するはずである。他の実施例として、サウンド場面を空間的にグルーピングして、各グループごとにレンダリングに必要な信号をともに伝送すれば、レンダラの内部のバッファで不要に待機する信号の量を最小化することができる。一方、レンダラ 1620 によって一度に収容可能なデータの大きさが異なり得るが、このような情報も順序制御部 1630 に通知して、これに合わせて復号化器 1610 が復号化タイミングおよび伝送量を決定することができる。

#### 【0096】

一方、順序制御部 1630 による復号化の統制は、符号化段に伝達され、符号化過程まで統制することができる。すなわち、符号化器で不要な信号を符号化時に除外したり、オブジェクト、チャンネルに対するグルーピングを決定することなどが可能である。

#### 【0097】

(音声高速道路)

10

一方、ビット列のうち、双方向通信の音声に相当するオブジェクトが含まれてよい。双方向通信は、他のコンテンツと異なり、時間遅延に非常に敏感であるため、これに相当するオブジェクトあるいはチャンネル信号が受信された場合、これを優先してレンダラに伝送しなければならない。これに相当するオブジェクトあるいはチャンネル信号は、別のフラグなどで表示することができる。まず、伝送オブジェクトは、他のオブジェクト / チャンネルと異なり、同じフレームに入っている他のオブジェクト / チャンネル信号と再生時間 (presentation time) において独立した特性を有する。

#### 【0098】

(A V 整合および Phantom Center )

20

U H D T V 、すなわち超高解像度テレビを考慮する時、発生する新たな問題の一つとして、一般に Near Field と呼ぶ状況がある。すなわち、一般的なユーザ環境（居間）の視聴距離を考慮する時、再生されるスピーカからの聞き手までの距離が各スピーカの間の距離より短くなることによって、各スピーカが点音源として動作する点と、広くて大きいスクリーンによって中央部にスピーカが不在する状況でビデオに同期化した音オブジェクトの空間解像度が非常に高くてはじめて、高品質の 3D オーディオサービスが可能であるという点である。

#### 【0099】

従来の 30 度程度の視聴角度では、左右に配置されたステレオスピーカが Near Field 状況に置かれず、画面上のオブジェクトの移動（例えば、左から右へ移動する自動車）に合わせたサウンド場面を提供するに十分である。しかし、視聴角度が 100 度に達する U H D T V 環境では、左右の解像度だけでなく、画面の上下を構成する追加の解像度が必要である。例えば、画面上の 2 人の登場人物がいる場合、現在の H D T V では、2 人の声がすべて真ん中から発話されるかのように聞こえても、現実感において大きな問題に感じられないが、U H D T V の大きさにおいては、画面とそれに対応する声の不一致が新たな形態の歪みとして認識されるはずである。

30

#### 【0100】

これに対する解決方策の一つとして、22.2 チャンネルスピーカ configuration の形態が挙げられる。図 2 は、22.2 チャンネル配置の一例である。図 2 によれば、前面部に計 11 個のスピーカを配置して、前面の左右および上下の空間解像度を大きく高めている。従来 3 個のスピーカが担当していた中間層に 5 個のスピーカを配置する

40

50

。そして、上位階層3個、下位階層に3個を追加することで音の高低も十分に対応できるようにした。このような配置を利用すれば、従来に比べて前面の空間解像度が高くなるため、それだけビデオ信号との整合に有利になる。しかし、LCD、OLEDなどのディスプレイ素子を用いる現在のテレビにおいて、スピーカーが存在すべき位置をディスプレイが占める問題がある。すなわち、ディスプレイ自体が音を提供したり、あるいは音を貫通する素子の性格を有しない限り、ディスプレイ領域の外部に存在するスピーカーを用いて、画面内の各オブジェクト位置に整合された音を提供しなければならない問題が存在する。図2において、少なくともFLc、FC、FRcに相当するスピーカーは、ディスプレイと重複した位置に配置される。

## 【0101】

10

図17は、22.2チャンネルシステムにおいて、全面配置スピーカーのうち、ディスプレイによって不在のスピーカーをその周辺チャンネルを用いて再生する概念を説明するための概念図である。FLc、FC、FRcの不在に対応するために、点線で表示した円のように追加のスピーカーをディスプレイの上下周辺部に配置する場合も考慮することができる。図17によれば、FLcを生成するのに使用可能な周辺チャンネルは7個あり得る。この7個のスピーカーを用いて仮想ソースを生成する原理で不在のスピーカー位置に相当する音を再生することができる。

## 【0102】

周辺スピーカーを用いて仮想ソースを生成する方法として、V B A PやH A A S E f f e c t（先行効果）のような技術および性質を利用することができる。あるいは、周波数帯域に応じて互いに異なるパニング手法を適用することができる。ひいては、H R T Fを用いた方位角の変更および高さ調整などを考慮することができる。例えば、B t F Cを用いてFCを代替する場合、上昇性質を有するH R T Fを適用してFCチャンネル信号をB t F Cに加える方法で実現することができる。H R T Fの観察を通じて把握できる性質は、音の高さを調整するためには、高周波数帯域の特定N u l lの位置（これは、人によって異なる）を制御しなければならないということである。しかし、人によって異なるN u l lを一般化して実現するためには、高周波数帯域を広く拡大したり縮小する方法で高さ調整を実現することができる。このような方法を用いると、代わりにフィルタの影響で信号に歪みが発生する欠点がある。

20

## 【0103】

30

本発明による不在スピーカ位置への音源配置のための処理方法は、図18に示されている通りである。図18によれば、ファンタムスピーカ位置に対応するチャンネル信号が入力信号として使用され、入力信号は、3つのバンドに分割するサブバンドフィルタ部1810を経る。スピーカアレイがない方法で実現されてもよいが、この場合、3つのバンドの代わりに、2つのバンドに区分したり、3つのバンドに分割した代わりに、上位2つのバンドに対してそれぞれ異なる処理を経る方法で実現されてもよい。1番目バンド（S L、S 1）は、低周波帯域で相対的に位置に鈍い代わりに、大きさの大きいスピーカを介して再生することが好ましいため、ウーファあるいはサブウーファを介して再生できる信号である。この時、先行効果を利用するため、1番目バンドの信号は、時間遅延フィルタ部1820によって遅延されることがある。この時、時間遅延は、他のバンドでの処理過程で発生するフィルタの時間遅延を補償するためではなく、他のバンド信号対比より遅く再生されるようにするために、すなわち、先行効果を提供するための追加的な時間遅延を提供する。

40

## 【0104】

2番目バンド（S M、S 2～S 5）は、ファンタムスピーカ周辺の（テレビのディスプレイのベゼルおよびその周辺に配置されるスピーカ）スピーカを介して再生されるのに使用される信号であって、少なくとも2つのスピーカに分割されて再生され、V B A Pなどのパニングアルゴリズム1830を適用するための係数が生成されて適用される。したがって、2番目バンドの出力が再生されるスピーカの個数と位置（ファンタムスピーカに対して相対的な）を正確に提供してはじめて、これを通したパニング効果が向上することが

50

できる。この時、V B A P パニングのほか、H R T F を考慮したフィルタの適用や、時間パニング効果を提供するために、互いに異なる位相フィルタあるいは時間遅延フィルタを適用することも可能である。このようにバンドを分けてH R T F を適用する時に得られるさらに他の利点は、H R T F によって発生する信号歪みの範囲を処理する帯域内に制限することができる点である。

#### 【 0 1 0 5 】

3番目バンド(S H、S 6 ~ S \_ N )は、スピーカアレイが存在する場合、これを用いて再生される信号を生成するためであり、スピーカアレイ制御部 1 8 4 0 で少なくとも 3 つのスピーカを介した音源仮想化のためのアレイ信号処理技術を適用することができる。あるいは、W F S (Wave Field Synthesis)により生成される係数を適用することができる。この時、3番目バンドと2番目バンドは実際に同じバンドであってもよい。10

#### 【 0 1 0 6 】

図 1 9 は、各バンドで生成された信号をテレビ周辺に配置されたスピーカとマッピングさせる一実施例を示す。図 1 9 によれば、2番目バンド(S 2 ~ S 5 )および3番目バンド(S 6 ~ S \_ N )に対応するスピーカの個数および位置情報は相対的に正確に定義された位置になければならず、その位置情報は、図 1 8 の処理システムに提供されることが好みしい。

#### 【 0 1 0 7 】

図 2 0 は、本発明の一実施例によるオーディオ信号処理装置が実現された製品の関係を示す図である。まず、図 2 0 を参照すれば、有無線通信部 3 1 0 は、有無線通信方式によりビットストリームを受信する。具体的には、有無線通信部 3 1 0 は、有線通信部 3 1 0 A、赤外線通信部 3 1 0 B、ブルートゥース部 3 1 0 C、無線R A N 通信部 3 1 0 D のうちの 1 つ以上を含むことができる。20

ユーザ認証部 3 2 0 は、ユーザ情報を受信してユーザ認証を行うものであって、指紋認識部 3 2 0 A、虹彩認識部 3 2 0 B、顔認識部 3 2 0 C、および音声認識部 3 2 0 D のうちの 1 つ以上を含むことができるが、それぞれ指紋、虹彩情報、顔輪郭情報、音声情報を受信して、ユーザ情報に変換し、ユーザ情報および既に登録されているユーザデータと一致するか否かを判断して、ユーザ認証を行うことができる。

#### 【 0 1 0 8 】

入力部 3 3 0 は、ユーザが様々な種類の命令を入力するための入力装置であって、キーパッド部 3 3 0 A、タッチパッド部 3 3 0 B、リモコン部 3 3 0 C のうちの 1 つ以上を含むことができるが、本発明はこれに限定されない。30

#### 【 0 1 0 9 】

信号コーディングユニット 3 4 0 は、有無線通信部 3 1 0 を介して受信されたオーディオ信号および／またはビデオ信号に対してエンコーディングまたはデコーディングを行い、時間ドメインのオーディオ信号を出力する。前記信号コーディングユニット 3 4 0 は、オーディオ信号処理装置 3 4 5 を含むことができる。この時、オーディオ信号処理装置 3 4 5 は、先に説明した本発明の実施例（すなわち、一実施例によるデコーダ 6 0 0 、および他の実施例によるエンコーダおよびデコーダ 1 4 0 0 ）に相当するものであって、このようにオーディオ処理装置 3 4 5 およびこれを含む信号コーディングユニット 3 4 0 は、1 つ以上のプロセッサによって実現できる。40

#### 【 0 1 1 0 】

制御部 3 5 0 は、入力装置から入力信号を受信し、信号コーディングユニット 3 4 0 および出力部 3 6 0 のすべてのプロセスを制御する。出力部 3 6 0 は、信号デコーディングユニット 3 4 0 によって生成された出力信号などが出力される構成要素であって、スピーカ部 3 6 0 A と、ディスプレイ部 3 6 0 B とを含むことができる。出力信号がオーディオ信号の時、出力信号はスピーカに出力され、ビデオ信号の時、出力信号はディスプレイを介して出力される。

#### 【 0 1 1 1 】

1020304050

本発明によるオーディオ信号処理方法は、コンピュータで実行されるためのプログラムに作成され、コンピュータ読み取り可能な記録媒体に格納されてよいし、本発明によるデータ構造を有するマルチメディアデータもコンピュータ読み取り可能な記録媒体に格納されてよい。前記コンピュータ読み取り可能な記録媒体は、コンピュータシステムによって読み込まれるデータが格納されるすべての種類の格納装置を含む。コンピュータ読み取り可能な記録媒体の例には、ROM、RAM、CD-ROM、磁気テープ、フロッピー（登録商標）ディスク、光データ記憶装置などがあり、また、キャリアウェーブ（例えば、インターネットを介した伝送）の形態で実現されるものも含む。さらに、前記エンコーディング方法によって生成されたビットストリームは、コンピュータ読み取り可能な記録媒体に格納されたり、有／無線通信網を用いて伝送可能である。

10

### 【0112】

以上、本発明は、限定された実施例と図面によって説明されたが、本発明はこれによって限定されず、本発明の属する技術分野における通常の知識を有する者によって本発明の技術思想と以下に記載される特許請求の範囲の均等範囲内で多様な修正および変形が可能であることは当然である。

### 【実施例】

#### 【0113】

上述のように、発明を実施するための形態において、関連事項を記述した。

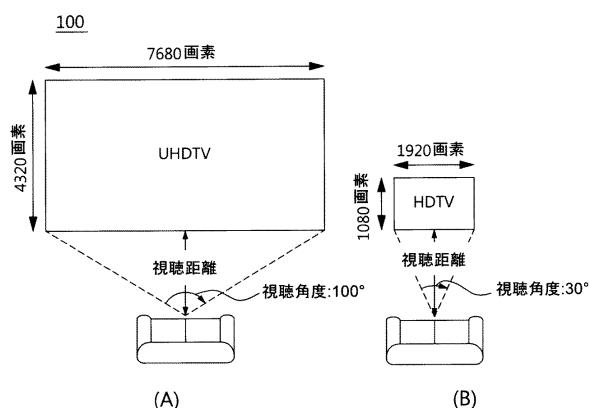
#### 【産業上の利用可能性】

#### 【0114】

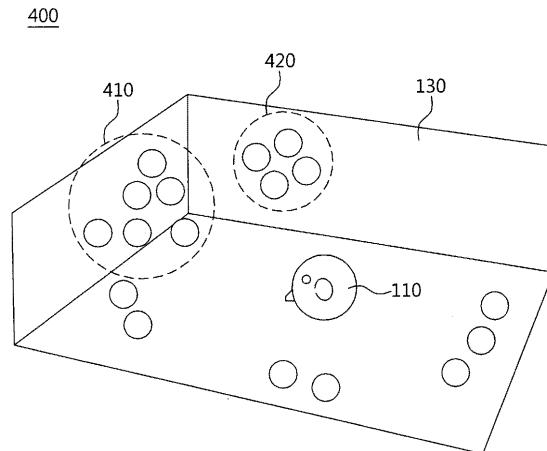
20

本発明は、オーディオ信号をエンコーディングおよびデコーディングしたり、オーディオ信号に多様な処理をする過程に適用可能である。

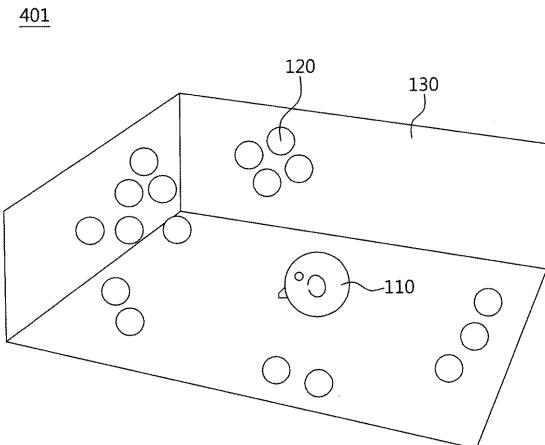
【図1】



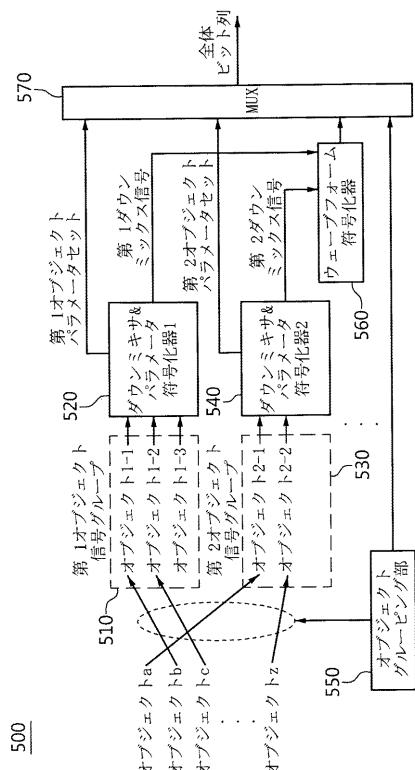
【図4】



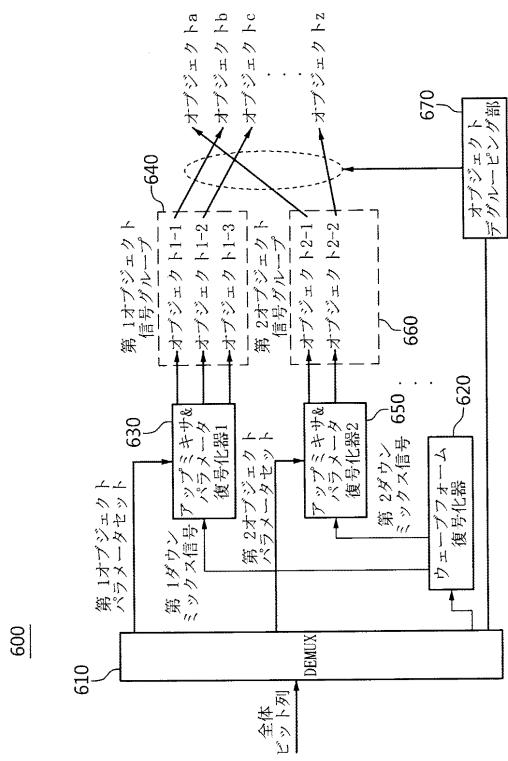
【図3】



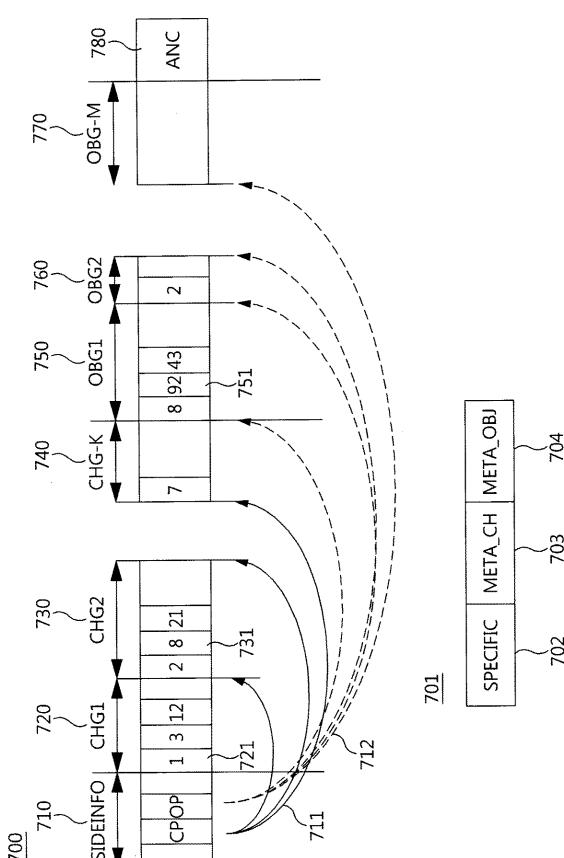
【図5】



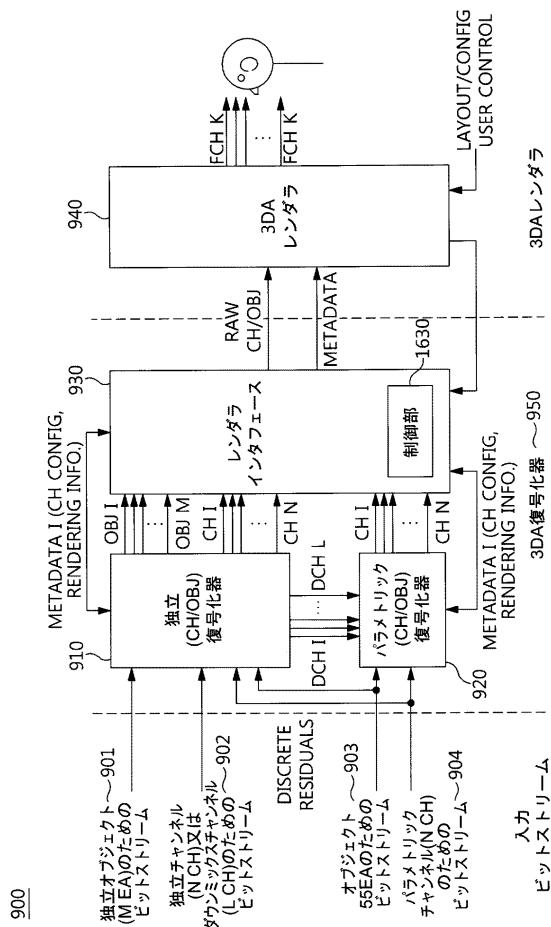
【図6】



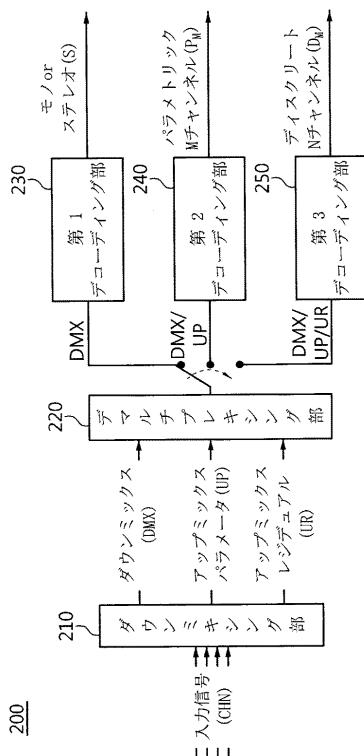
【図7】



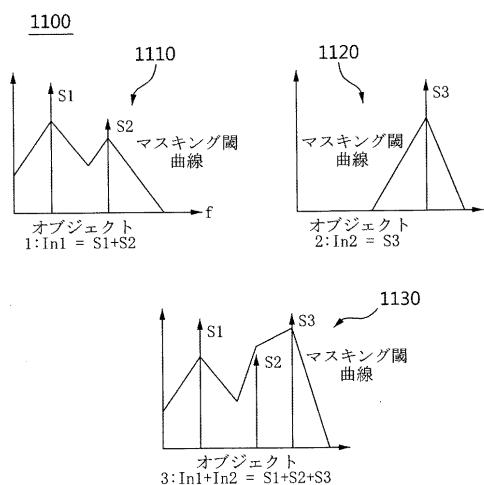
【図9】



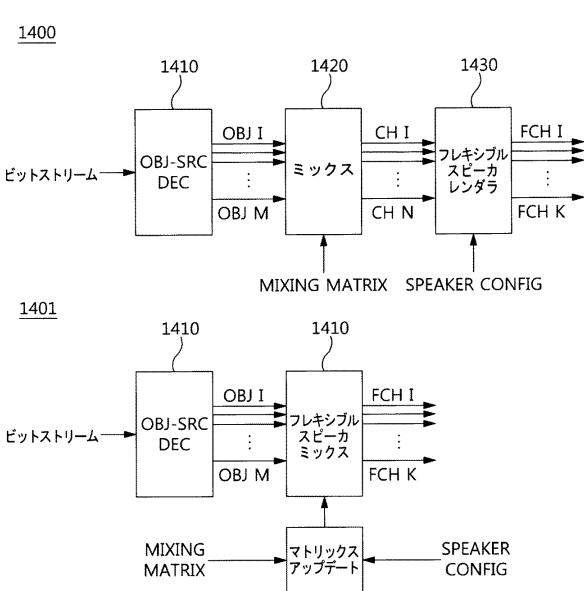
【図10】



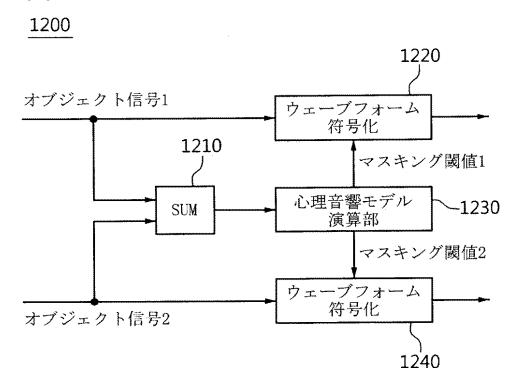
【図11】



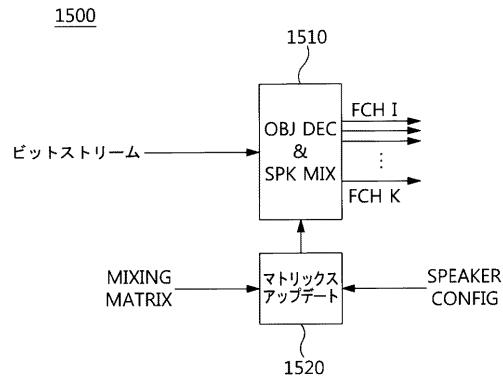
【図14】



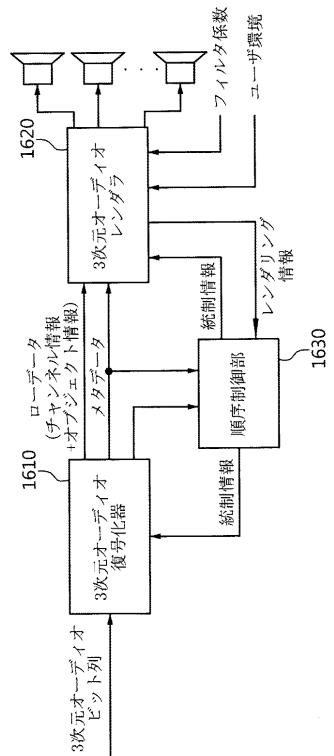
【図12】



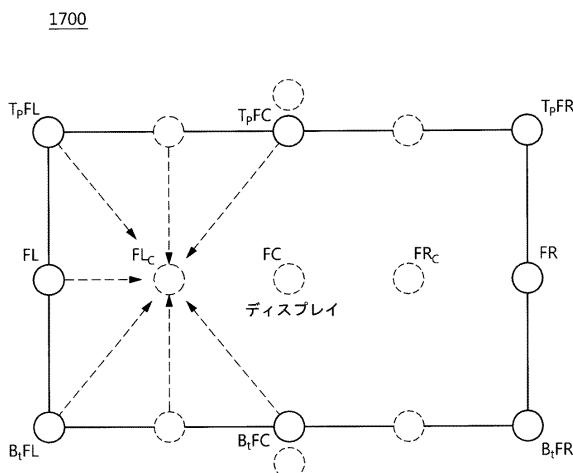
【図15】



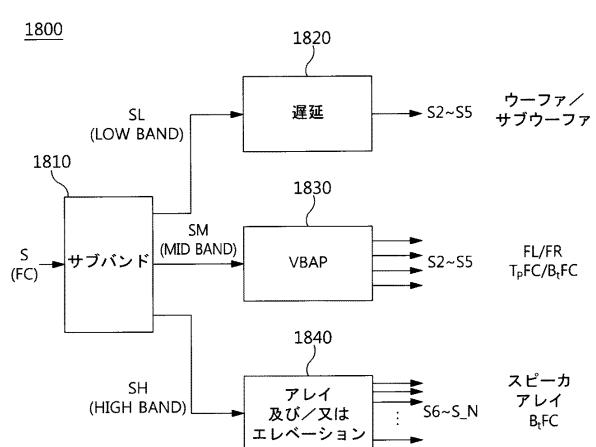
【図16】



【図17】

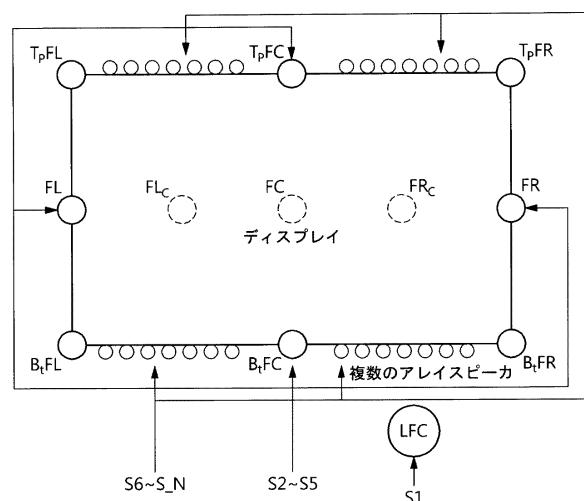


【図18】



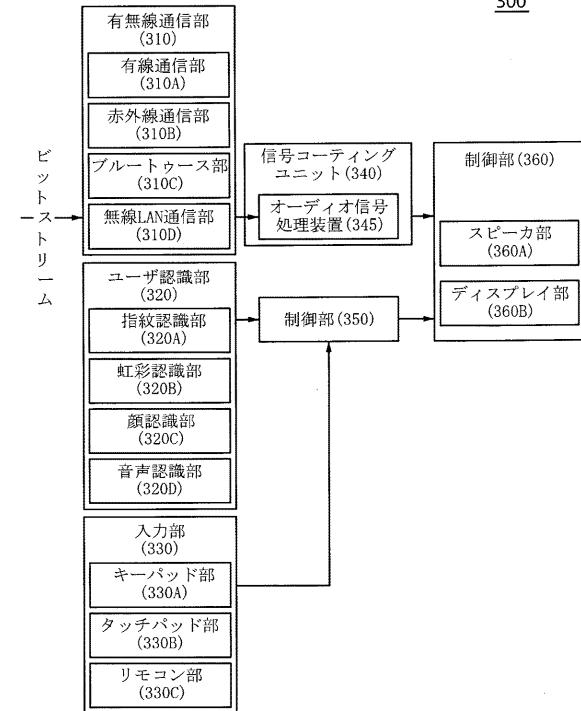
【図19】

1900



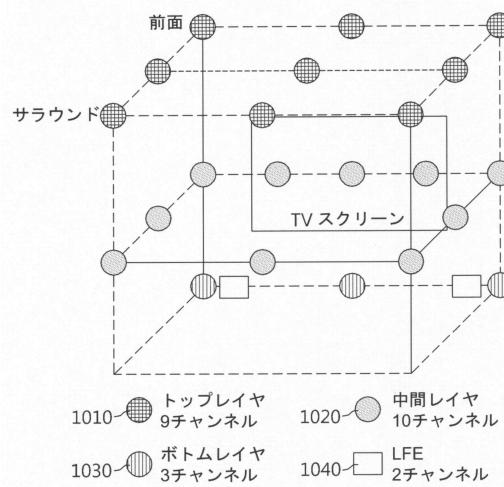
【図20】

300



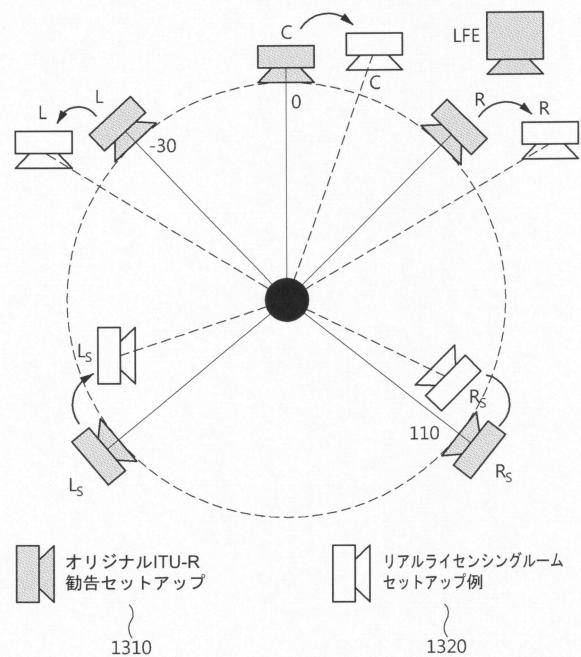
【図2】

1000



【図13】

1300



---

フロントページの続き

(31)優先権主張番号 10-2012-0083944  
(32)優先日 平成24年7月31日(2012.7.31)  
(33)優先権主張国 韓国(KR)  
(31)優先権主張番号 10-2012-0084231  
(32)優先日 平成24年7月31日(2012.7.31)  
(33)優先権主張国 韓国(KR)

(72)発明者 オ・ヒョンオ  
大韓民国 4 6 3 - 8 8 1 キョンギド、ソンナムシ、ブンダング、トンパンギヨロ 1 2 3 番、ブルジ  
オ - グランブルー・アパートメント 1 1 0 - 1 2 0 3  
(72)発明者 ソン・チョンオク  
大韓民国 1 2 0 - 7 4 9 ソウル、ソデムング、ヨンセイロ 5 0 番、ヨンセイ・ユニバーシティ、デ  
パートメント・オブ・エレクトリカル・アンド・エレクトロニック・エンジニアリング、スクール  
・オブ・エンジニアリング、ビ-6 0 1  
(72)発明者 ソン・ミョンソク  
大韓民国 1 2 0 - 7 4 9 ソウル、ソデムング、ヨンセイロ 5 0 番、ヨンセイ・ユニバーシティ、デ  
パートメント・オブ・エレクトリカル・アンド・エレクトロニック・エンジニアリング、スクール  
・オブ・エンジニアリング、ビ-6 0 1  
(72)発明者 チョン・セウォン  
大韓民国 1 2 0 - 7 4 9 ソウル、ソデムング、ヨンセイロ 5 0 番、ヨンセイ・ユニバーシティ、デ  
パートメント・オブ・エレクトリカル・アンド・エレクトロニック・エンジニアリング、スクール  
・オブ・エンジニアリング、ビ-6 0 1  
(72)発明者 イ・テギュ  
大韓民国 1 2 0 - 7 4 9 ソウル、ソデムング、ヨンセイロ 5 0 番、ヨンセイ・ユニバーシティ、デ  
パートメント・オブ・エレクトリカル・アンド・エレクトロニック・エンジニアリング、スクール  
・オブ・エンジニアリング、ビ-6 0 1

審査官 菊池 智紀

(56)参考文献 特表 2 0 1 0 - 5 0 7 1 1 4 (JP, A)  
米国特許出願公開第 2 0 1 2 / 0 1 8 3 1 4 8 (US, A1)  
特表 2 0 1 0 - 5 1 1 9 1 2 (JP, A)  
特表 2 0 1 0 - 5 2 9 5 0 0 (JP, A)  
特表 2 0 1 1 - 5 0 1 8 2 3 (JP, A)  
国際公開第 2 0 1 0 / 1 0 9 9 1 8 (WO, A1)

(58)調査した分野(Int.Cl., DB名)  
G 10 L 19 / 00 - 19 / 26  
H 04 S 3 / 00