

【公報種別】特許法第 17 条の 2 の規定による補正の掲載  
 【部門区分】第 6 部門第 2 区分  
 【発行日】平成 23 年 2 月 17 日 (2011.2.17)

【公表番号】特表 2010-515099 (P2010-515099A)  
 【公表日】平成 22 年 5 月 6 日 (2010.5.6)  
 【年通号数】公開・登録公報 2010-018  
 【出願番号】特願 2009-543949 (P2009-543949)  
 【国際特許分類】

G 1 0 L 19/02 (2006.01)

G 1 0 L 19/00 (2006.01)

【F I】

G 1 0 L 19/02 1 5 0

G 1 0 L 19/00 3 3 0 B

【手続補正書】

【提出日】平成 22 年 12 月 24 日 (2010.12.24)

【手続補正 1】

【補正対象書類名】明細書

【補正対象項目名】全文

【補正方法】変更

【補正の内容】

【発明の詳細な説明】

【発明の名称】情報ビットストリーム変換を含む多様なチャンネルで構成されたマルチオブジェクトオーディオ信号の符号化および復号化装置、並びに方法

【技術分野】

【0001】

本発明は、多様なチャンネルで構成されたマルチオブジェクトオーディオ信号の符号化および復号化装置、並びに方法に関し、より詳細には、付加情報ビットストリームを変換し、変換された付加情報ビットストリームを利用して所望の出力信号、すなわち多様なチャンネルを有するマルチオブジェクトオーディオ信号を復元する付加情報ビットストリーム変換を含む多様なチャンネルで構成されたマルチオブジェクトオーディオ信号の符号化および復号化装置、並びに方法に関する。

【0002】

ここで、多様なチャンネルを有するマルチオブジェクトオーディオ信号とは、マルチオブジェクトオーディオ信号であって、各々のオーディオオブジェクトが互いに異なるチャンネル（例えば、モノ、ステレオ、5.1 チャンネル）で構成されたオーディオ信号を意味する。

【0003】

本発明は、情報通信部および情報通信研究振興院の IT 次世代核心技術開発事業の一環として行われた研究から導出されたものである。[課題管理番号：2005-S-403-02、課題名：知能型統合情報放送 (SmaRTV) 技術 (Development of Super-intelligent Multimedia Anytime-anywhere Realistic TV (SmaRTV) Technology)]

【背景技術】

【0004】

従来のオーディオ符号化/復号化技術によれば、ユーザは、オーディオコンテンツを受動的に聴取するしかできなかった。従って、ユーザの必要に応じて互いに異なるチャンネルで構成された各オーディオオブジェクトを制御し、1つのオーディオコンテンツを多様な方法で組み合わせることによって、多様なオーディオオブジェクトを消費することができる

多様なチャンネルで構成された複数のオーディオオブジェクト別の符号化および復号化装置、並びに方法が要求される。

これと関連して、従来の S A C ( S p a t i a l A u d i o C o d i n g ) は、マルチチャンネルオーディオ信号をダウンミックスされたモノまたはステレオ信号で表現、伝送および復元する技術であって、低いビット率でも高品質のマルチチャンネルオーディオ信号を伝送することができる。

【 0 0 0 5 】

しかし、従来の S A C はマルチチャンネルの信号を、1つのオーディオオブジェクトに対してのみ符号化および復号化が可能な技術であるため、マルチチャンネル及びマルチオブジェクトオーディオ信号、例えば、モノ、ステレオ、および 5 . 1 チャンネルで構成された多様なオブジェクトのオーディオ信号を符号化 / 復号化することができないという問題がある。

【 0 0 0 6 】

また、従来のバイノーラルキューコーディング ( B i n a u r a l C u e C o d i n g 、 B C C ) は、マルチオブジェクトオーディオ信号を符号化 / 復号化することができる。しかし、オーディオオブジェクトのチャンネルは、モノチャンネルに限定されることから、モノチャンネルを含む多様なチャンネルで構成されたマルチオブジェクトオーディオ信号を符号化 / 復号化することができない。

【 0 0 0 7 】

整理すれば、従来技術は、単一チャンネルで構成されたマルチオブジェクトオーディオ信号またはマルチチャンネルの単一オブジェクトオーディオ信号に対してのみ符号化 / 復号化を行なうことができるため、多様なチャンネルで構成されたマルチオブジェクトオーディオ信号に対しては符号化 / 復号化を行なうことができない。したがって、従来のオーディオ符号化 / 復号化技術によれば、ユーザはオーディオコンテンツを受動的に聴取するほかなかった。

【 0 0 0 8 】

したがって、ユーザの必要に応じて、互いに異なるチャンネルで構成された複数のオーディオオブジェクト各々を制御し、1つのオーディオコンテンツを多様な方法で組み合わせることによって、多様なオーディオオブジェクトを消費できる多様なチャンネルで構成された複数のオーディオオブジェクト別の符号化および復号化装置、並びに方法が要求される。

【 0 0 0 9 】

また、マルチオブジェクトオーディオ符号化器から生成された付加情報ビットストリームが、従来の S A C 符号化 / 復号化器の付加情報ビットストリーム に対する 下位互換性を提供するために、マルチオブジェクトオーディオビットストリームと従来の S A C ビットストリームとの間で 変換が可能な 変換装置および方法が要求される。

【 0 0 1 0 】

前述したように、互いに異なるチャンネルで構成された複数のオーディオオブジェクト各々を制御し、1つのオーディオコンテンツを多様な方法で組み合わせることのできる多様なチャンネルのマルチオブジェクトオーディオ信号の符号化および復号化装置、並びに方法であって、前記マルチオブジェクトオーディオ符号化および復号化装置、並びに方法は、従来の S A C ビットストリームとの下位互換性を提供するためのビットストリーム変換が可能な多様なチャンネルのマルチオブジェクトオーディオ信号の符号化および復号化装置、並びに方法が要求される。

【 発明の概要 】

【 発明が解決しようとする課題 】

【 0 0 1 1 】

本発明は前述した要求に応じるために提案されたものであって、その目的は、多様なチャンネルを有するマルチオブジェクトオーディオ信号を符号化および復号化し、従来の S A C ビットストリーム に対する 下位互換性を提供する多様なチャンネルで構成されたマルチオブジェクトオーディオ信号符号化および復号化装置、並びに方法を提供することにある。

**【課題を解決するための手段】****【0012】**

前述した目的を達成するための本発明は、マルチオブジェクトオーディオ信号の符号化装置であって、前記符号化装置に入力されるオーディオオブジェクト信号を空間キュー基盤で符号化し、前記符号化されたオーディオオブジェクト信号に対するレンダリング情報を生成するオーディオオブジェクト符号化手段を含み、前記レンダリング情報は、前記オーディオオブジェクト信号に対する空間キュー情報、前記オーディオオブジェクト信号のチャンネル情報、および前記オーディオオブジェクト信号の識別情報を含む符号化装置を提供する。

**【0013】**

また、前述した目的を達成するための本発明は、マルチオブジェクトオーディオ信号を復号化するためのレンダリング情報を生成するトランスコーディング装置であって、符号化されたオーディオオブジェクト信号に対するオブジェクト制御情報および再生情報に基づいて、前記符号化されたオーディオオブジェクト信号に対するパワー利得情報および出力位置情報を含むレンダリング情報を生成する第1行列手段と、前記第1行列手段によって生成されたレンダリング情報および符号化装置から入力される前記符号化されたオーディオオブジェクト信号に対するレンダリング情報に基づいて、復号化装置から出力されるオーディオ信号に対する空間キュー情報を生成するレンダリング手段と、を備えるトランスコーディング装置を提供する。

**【0014】**

また、前述した目的を達成するための本発明は、マルチチャンネルオーディオ信号およびマルチオブジェクトオーディオ信号を復号化するためのレンダリング情報を生成するトランスコーディング装置であって、符号化装置から入力される符号化されたオーディオ信号に対するレンダリング情報から符号化されたオーディオオブジェクト信号に対するレンダリング情報、およびマルチチャンネルオーディオ信号に対するレンダリング情報を分離するパーシング手段と、符号化されたオーディオオブジェクト信号に対するオブジェクト制御情報および再生情報に基づいて、前記符号化されたオーディオオブジェクト信号に対するパワー利得情報および出力位置情報を含むレンダリング情報を生成する第1行列手段と、前記パーシング手段によって分離された符号化されたマルチチャンネルオーディオ信号に対するレンダリング情報に基づいて、前記マルチチャンネルオーディオ信号に対するチャンネル別のパワー利得情報を含むレンダリング情報を生成する第2行列手段と、前記第1行列手段によって生成されたレンダリング情報、前記第2行列手段によって生成されたレンダリング情報、および前記パーシング手段によって分離された符号化されたオーディオオブジェクト信号に対するレンダリング情報に基づいて、復号化装置から出力されるオーディオ信号に対する空間キュー情報を生成するレンダリング手段と、を備えるトランスコーディング装置を提供する。

**【0015】**

また、前述した目的を達成するための本発明は、マルチオブジェクトオーディオ信号の符号化方法であって、入力されるオーディオオブジェクト信号を空間キュー基盤で符号化し、前記符号化されたオーディオオブジェクト信号に対するレンダリング情報を生成するオーディオオブジェクト符号化ステップを含み、前記レンダリング情報は、前記オーディオオブジェクト信号に対する空間キュー情報、前記オーディオオブジェクト信号のチャンネル情報、および前記オーディオオブジェクト信号の識別情報を含む符号化方法を提供する。

**【0016】**

また、前述した目的を達成するための本発明は、マルチオブジェクトオーディオ信号を復号化するためのレンダリング情報を生成するトランスコーディング方法であって、符号化されたオーディオオブジェクト信号に対するオブジェクト制御情報および再生情報に基づいて、前記符号化されたオーディオオブジェクト信号に対するパワー利得情報および出力位置情報を含むレンダリング情報を生成する第1行列ステップと、前記第1行列ステッ

ブによって生成されたレンダリング情報および符号化ステップから入力される前記符号化されたオーディオオブジェクト信号に対するレンダリング情報に基づいて、復号化ステップから出力されるオーディオ信号に対する空間キュー情報を生成するレンダリングステップと、を含むトランスコーディング方法を提供する。

【 0 0 1 7 】

また、前述した目的を達成するための本発明は、マルチチャネルオーディオ信号およびマルチオブジェクトオーディオ信号を復号化するためのレンダリング情報を生成するトランスコーディング方法であって、符号化ステップから入力される符号化されたオーディオ信号に対するレンダリング情報から符号化されたオーディオオブジェクト信号に対するレンダリング情報、およびマルチチャネルオーディオ信号に対するレンダリング情報を分離するパーシングステップと、符号化されたオーディオオブジェクト信号に対するオブジェクト制御情報および再生情報に基づいて、前記符号化されたオーディオオブジェクト信号に対するパワー利得情報および出力位置情報を含むレンダリング情報を生成する第 1 行列ステップと、前記パーシングステップによって分離された符号化されたマルチチャネルオーディオ信号に対するレンダリング情報に基づいて、前記マルチチャネルオーディオ信号に対するチャンネル別のパワー利得情報を含むレンダリング情報を生成する第 2 行列ステップと、前記第 1 行列ステップによって生成されたレンダリング情報、前記第 2 行列ステップによって生成されたレンダリング情報、および前記パーシングステップによって分離された符号化されたオーディオオブジェクト信号に対するレンダリング情報に基づいて、復号化ステップから出力されるオーディオ信号に対する空間キュー情報を生成するレンダリングステップと、を含むトランスコーディング方法を提供する。

【 発明の効果 】

【 0 0 1 8 】

本発明は、付加情報ビットストリームの変換を含み、多様なチャネルを有するマルチオブジェクトオーディオ信号の符号化および復号化装置、並びに方法を提供することによって、多様なチャネルを有するマルチオブジェクトオーディオコンテンツを効率よく符号化および復号化し、ユーザの必要に応じて能動的にオーディオコンテンツを消費することができ、従来に利用されるビットストリームに対する下位互換性を提供することによって、従来における符号化および復号化装置との互換性を提供することができる。

【 図面の簡単な説明 】

【 0 0 1 9 】

【 図 1 】 本発明の実施形態に係るマルチオブジェクトオーディオ符号化器およびマルチオブジェクト復号化器を示すブロック図である。

【 図 2 】 本発明の実施形態に係るマルチオブジェクトオーディオ符号化器およびマルチオブジェクト復号化器を示すブロック図である。

【 図 3 】 本発明の実施形態に係る図 2 のトランスコーダ 1 0 3 を示すブロック図である。

【 図 4 】 本発明の実施形態に係る図 2 のビットストリームフォーマッタ 2 0 5 から生成される代表 S A O C ( s p a t i a l a u d i o o b j e c t c o d i n g ) ビットストリームを示す図である。

【 図 5 】 本発明の他の実施形態に係る図 2 の代表 S A O C ビットストリームを示す図である。

【 図 6 】 本発明の他の実施形態に係る図 2 のトランスコーダ 1 0 3 を示すブロック図である。

【 図 7 】 図 2 におけるマルチオブジェクトオーディオ符号化器および復号化器にオーディオオブジェクト除去部 7 0 1 が追加的に備えられた場合を示すブロック図である。

【 図 8 】 図 2 における S A C コーダ 2 0 1 および S A C デコーダ 1 0 5 を M P E G サラウンドコーダおよびデコーダに代替した場合を示すブロック図である。

【 発明を実施するための形態 】

【 0 0 2 0 】

前述した目的、特徴および長所は、添付の図面と関連した次の詳細な説明を介して更に

明確になるのであろう。以下、添付の図面に基づいて本発明に係る好ましい一実施形態を詳細に説明する。

【0021】

図1は、本発明に係るマルチオブジェクトオーディオ符号化器およびマルチオブジェクト復号化器を示すブロック図である。

【0022】

図1に示すように、本発明は、SAOC (Spatial Audio Object Coding) コーダ101、トランスコーダ103、およびSAC (Spatial Audio Coding) デコーダ105を備える。

【0023】

空間オーディオオブジェクトコーディング (SAOC: Spatial Audio Object Coding) 方式によれば、コーダに入力される信号はオーディオオブジェクトとして符号化される。ここで、各オーディオオブジェクトは、デコーダによって復元され、独立的に再生されない。しかし、所望のオーディオ場面を構成するためにオーディオオブジェクトに対する情報がレンダリングされ、多様なチャンネルを有するマルチオブジェクトオーディオ信号が出力される。したがって、SACデコーダは、所望のオーディオ場面を獲得するために入力されるオーディオオブジェクトに対する情報をレンダリングできる装置を要求する。

【0024】

前記SAOCコーダ101は、空間キュー基盤のコーダであって、入力オーディオ信号をオーディオオブジェクトとして符号化する。ここで、前記SAOCコーダ101に入力されるオーディオオブジェクトはモノまたはステレオ信号である。

【0025】

前記SAOCコーダ101は、入力される1つ以上のオーディオオブジェクトからダウンミックス信号を出力し、空間キューおよび付加情報 (side information) を抽出することによりSAOCビットストリームを生成する。ここで、出力されるダウンミックス信号はモノまたはステレオ信号である。前記SAOCコーダ101は「異質なレイアウトSAOC」または「Faller」技法に基づいて入力されるオーディオオブジェクト信号を分析する。

【0026】

前記抽出されたSAOCビットストリームは空間キューおよび付加情報を含み、前記付加情報は入力オーディオオブジェクトの空間情報を含む。前記空間キューは、一般的に周波数領域のサブバンド単位で分析されて抽出される。

【0027】

ここで、空間キュー (spatial cue) とは、オーディオ信号を符号化および復号化する過程で利用される情報である。空間キューは、周波数領域から抽出され、入力される2つの信号の大きさの差、遅延差、相関性などの情報を含む。例えば、空間キューは、オーディオ信号のパワー利得情報を表すオーディオ信号間のレベル差 (Channel Level Difference、CLD)、オーディオ信号間のエネルギー比 (Inter-Channel Level Difference、ICLD)、オーディオ信号間の時間差 (Inter Channel Time Difference、ICTD)、オーディオ信号間の相関性情報を表すオーディオ信号間の相関性 (Inter Channel Correlation、ICC)、および仮想音源位置情報 (Virtual Source Location Information) を含むが、これらの例に限定されない。

【0028】

また、前記付加情報は、オーディオ信号の復元、制御のための情報が含まれる。前記付加情報はヘッダー情報を含む。前記ヘッダー情報は、多様なチャンネルで構成されたマルチオブジェクトオーディオ信号の復元および再生のための情報を含み、オーディオオブジェクトに対するチャンネル情報および該当オーディオオブジェクトのIDを定義することによ

って、モノ、ステレオ、マルチチャネルのオーディオオブジェクトに対する復号化情報を提供することができる。例えば、符号化された特定オーディオオブジェクトがモノオーディオ信号であるかステレオオーディオ信号であるかを区分することができるよう各オブジェクトのIDおよび情報が定義され得る。前記ヘッダー情報は、一実施形態としてSAC(Spatial Audio Coding)ヘッダー情報、オーディオオブジェクト情報、およびプリセット(preset)情報を含むことができる。

【0029】

前記トランスコード103は、SAOCコード101に入力されるオーディオオブジェクトに対するレンダリングを行い、外部から入力される制御信号、すなわちオブジェクト別の音響情報および再生環境情報を利用して前記SAOCコード101から抽出されたSAOCビットストリームをSACビットストリームに変換する。

【0030】

すなわち、前記トランスコード103は、SAOCコード101に入力されるオーディオオブジェクトを多様なチャネルのマルチオブジェクトオーディオ信号に復元するために抽出されたSAOCビットストリームを利用してレンダリングする。付加情報を利用したレンダリングは、パラメータ領域で行なうことができる。

【0031】

また、前記トランスコード103は、SAOCビットストリームをSACビットストリームに変換する。まず、前記トランスコード103は、SAOCビットストリームから入力オーディオオブジェクトの情報を獲得し、入力オーディオオブジェクトの情報を所望のオーディオ場面に対応されるようレンダリングする。

【0032】

前記レンダリング過程において、前記トランスコード103は、所望のオーディオ場面对応する空間情報を予測し、前記予測された空間情報を変換してSACビットストリームとして出力する。

【0033】

前記トランスコード103に対しては、以下の図3で詳細に説明する。

【0034】

前記SACデコード105は、空間キュー基盤のマルチチャネルオーディオデコードであって、前記トランスコード103から出力されるSACビットストリームを利用し、前記SAOCコード101から出力されるダウンミックス信号をオブジェクト別のオーディオ信号に復元し、多様なチャネルのマルチオブジェクトオーディオ信号に復元する。前記SACデコード105は、MPEG(Moving Picture Experts Group)サラウンド復号化器、BCC(binaural cue coding)復号化器などで代替され得る。

【0035】

図2は、本発明に係るマルチオブジェクトオーディオ符号化器および復号化器を示す更なる一実施形態の構成図であって、入力信号が多様なチャネルを有する信号である場合を示す。

【0036】

同図に示したように、本発明はSAOCコード101、トランスコード103、SACデコード105、SACコード201、プリセットASI(audio scene information)部203、およびビットストリームフォーマッタ205を備える。

【0037】

前記SAOCコード101がモノまたはステレオのオーディオオブジェクトのみをサポートする場合、前記SACコード201は、入力されるマルチチャネルオーディオ信号から1つのオーディオオブジェクトを出力する。前記出力された1つのオーディオオブジェクトは、ダウンミックスされたモノまたはステレオ信号である。また前記SACコード201は、空間キューおよび付加情報を抽出してSACビットストリームを生成する。

## 【0038】

前記S A O Cコード101は、前記S A Cコード201から出力される1つのオーディオオブジェクトを備える1つ以上のオーディオオブジェクトから代表ダウンミックス信号を出力し、空間キューおよび付加情報を抽出してS A O Cビットストリームを生成する。

## 【0039】

前記プリセットA S I部203は、外部から入力される制御信号、すなわちオブジェクト別の音響情報および再生環境情報をプリセット - A S I ( P r e s e t - A u d i o S c e n e I n f o r m a t i o n ) 情報で構成し、前記プリセット - A S I 情報を含むプリセットA S I ビットストリームを生成する。前記プリセット - A S I 情報に対しては、以下図4で詳細に説明する。

## 【0040】

前記ビットストリームフォーマッタ205は、前記S A O Cコード101によって生成されたS A O Cビットストリーム、前記S A Cコード201によって生成されたS A Cビットストリーム、および前記プリセットA S I 部203によって生成されたプリセットA S I ビットストリームを利用し、代表S A O Cビットストリームを生成する。

## 【0041】

前記トランスコード103は、S A O Cコード101に入力されるオーディオオブジェクトに対するレンダリングを行い、外部から入力されるオブジェクト別の音響情報および再生環境情報を利用して前記ビットストリームフォーマッタ205によって生成された代表S A O Cビットストリームを代表S A Cビットストリームに変換する。前記トランスコード103は、前記S A Cデコード105に含まれて前記のような役割を行なうことができる。

## 【0042】

前記S A Cデコード105は、前記トランスコード103から出力されるS A Cビットストリームを利用し、前記S A O Cコード101から出力されるダウンミックス信号を多様なチャンネルのマルチオブジェクトオーディオ信号に復元する。前記S A Cデコード105は、M P E Gサラウンド復号化器、B C C 復号化器などで代替され得る。

## 【0043】

図3は、図2におけるトランスコード103を示す一実施形態の詳細な構造図である。

## 【0044】

同図に示すように、トランスコード103は、パーシング部301、レンダリング部303、第2行列部311、および第1行列部313を備え、代表S A O Cビットストリームを代表S A Cビットストリームに変換する。

## 【0045】

図1では、前記トランスコード103は、S A O CビットストリームをS A Cビットストリームに変換する。

## 【0046】

前記パーシング部301は、前記ビットストリームフォーマッタ205から生成された代表S A O Cビットストリーム、または図1のS A O Cコード101から生成されたS A O Cビットストリームをパーシングし、前記代表S A O Cビットストリームに含まれたS A O CビットストリームおよびS A Cビットストリームを分離する。また、前記パーシング部301は、前記分離したS A O CビットストリームからS A O Cコード101に入力されるオーディオオブジェクトの個数に対する情報を抽出する。ここで、図1のS A O Cコード101から生成されたS A O Cビットストリームをパーシングする場合、S A Cビットストリームが存在しないことから、分離する必要がない。

## 【0047】

前記第2行列部311は、前記パーシング部301によって分離したS A Cビットストリームを利用して第2行列を生成する。前記第2行列は、S A Cコード201に入力されるマルチチャンネルオーディオ信号に対する行列式である。もし、前記代表S A O CビットストリームにS A Cビットストリームが含まれていない場合、すなわち、図1のS A O C

コード 1 0 1 から生成された S A O C ビットストリームをパーシングする場合、前記第 2 行列部 3 1 1 は不要である。

【 0 0 4 8 】

第 2 行列は、S A C コード 2 0 1 に入力されるマルチチャネルオーディオ信号のパワー利得値に対するものであって、下記の数式 1 のとおりである。

【 数 1 】

$$\underbrace{\begin{bmatrix} \mathbf{w}_{ch\_1}^b \\ \mathbf{w}_{ch\_2}^b \\ \vdots \\ \mathbf{w}_{ch\_M}^b \end{bmatrix}}_{\text{Matrix II}} \left[ \mathbf{u}_{SAC}^b(k) \right] = \left[ \mathbf{Y}_{SAC}^b(k) \right] = \begin{bmatrix} y_{ch\_1}^b(k) \\ y_{ch\_2}^b(k) \\ \vdots \\ y_{ch\_M}^b(k) \end{bmatrix}$$

【 0 0 4 9 】

一般的に、1つのフレームをサブバンドに分割して解釈することが S A C の基本分析過程である。

【 0 0 5 0 】

ここで、 $\mathbf{u}_{SAC}^b(k)$  は S A C コード 2 0 1 から出力されるダウンミックス信号であって、 $k$  は周波数係数インデックスであり、 $b$  はサブバンドインデックスである。 $\mathbf{w}_{ch\_i}^b$  は前記 S A C ビットストリームから獲得できるマルチチャネル信号の空間キュー情報であって、 $i$  番目チャネル信号 ( $1 = i = M$ ) の周波数情報を復元するために利用される。したがって、 $\mathbf{w}_{ch\_i}^b$  は周波数係数の大きさ情報や位相情報として表現され得る。したがって、数式 1 の右側の項目を参照すれば、 $\mathbf{Y}_{SAC}^b(k)$  は数式 1 の結果であって、S A C デコード 1 0 5 から出力されるマルチチャネルオーディオ信号を表す。

【 0 0 5 1 】

前記  $\mathbf{u}_{SAC}^b(k)$  および  $\mathbf{w}_{ch\_i}^b$  はベクトルであり、 $\mathbf{u}_{SAC}^b(k)$  の前置行列次元 (Transpose Matrix Dimension) が  $\mathbf{w}_{ch\_i}^b$  の次元となる。例えば、下記の数式 2 のように定義され得る。ここで、S A C コード 2 0 1 から出力されるダウンミックス信号は、モノまたはステレオであるため、 $m$  は 1 または 2 である。

【 数 2 】

$$\mathbf{w}_{ch\_1}^b \times \mathbf{u}_{SAC}^b(k) = \begin{bmatrix} w_1^b & w_2^b & \cdots & w_m^b \end{bmatrix} \begin{bmatrix} u_1^b(k) \\ u_2^b(k) \\ \vdots \\ u_m^b(k) \end{bmatrix}$$

前記のように、 $\mathbf{w}_{ch\_i}^b$  は S A C ビットストリームに含まれた空間キュー情報である。

【 0 0 5 2 】

$\mathbf{w}_{ch\_i}^b$  が各チャネルのサブバンドでのパワー利得を表す場合、 $\mathbf{w}_{ch\_i}^b$  はチャネルレベル差空間キュー (Channel Level Difference、CLD) から予測されることができる。 $\mathbf{w}_{ch\_i}^b$  が周波数係数の位相差を補償するための係数として利用される場合、 $\mathbf{w}_{ch\_i}^b$  はチャネル時間差空間キュー (Channel Time Difference) または相関性空間キュー (Inter - Channel Coherence) から予測されることができる。

【 0 0 5 3 】

以下、例示的に  $\mathbf{w}_{ch\_i}^b$  が周波数係数の位相差を補償するための係数として利用さ

れる場合を基準にして説明する。

【 0 0 5 4 】

S A C コーダ 2 0 1 から出力されるダウンミックス信号との行列演算によって出力信号  $Y^b_{SAC}(k)$  が生成されるよう、数式 1 の第 2 行列は、各チャネル別のパワー利得値を表現し、前記ダウンミックス信号のベクトルの次元に逆にならなければならない。

【 0 0 5 5 】

前記第 2 行列部 3 1 1 が数式 1 および数式 2 を満足する第 2 行列を生成すれば、前記レンダリング部 3 0 3 は、前記生成された第 2 行列を前記第 1 行列部 3 1 3 の出力と結合する。

【 0 0 5 6 】

前記第 1 行列部 3 1 3 は、制御信号（例えば、オブジェクト制御情報、再生システム情報）を利用し、前記 S A O C コーダ 1 0 1 に入力される 1 つ以上のオーディオオブジェクトを所望の出力、すなわち多様なチャネルのマルチオブジェクトオーディオ信号でマッピングさせるための第 1 行列を生成する。

【 0 0 5 7 】

前記 S A O C コーダ 1 0 1 に入力されたオーディオオブジェクトの個数が  $N$  の場合、S A C コーダ 2 0 1 から出力されるダウンミックス信号も 1 つのオーディオオブジェクトとして見なされ、入力される  $N$  個のオーディオオブジェクトに含まれる。このような場合、S A C コーダ 2 0 1 から出力されるダウンミックス信号を除去した各オーディオオブジェクトは、第 1 行列を利用して S A C デコーダ 1 0 5 から出力されるチャネルでマッピングされ得る。

【 0 0 5 8 】

S A C デコーダ 1 0 5 から出力されるチャネルの個数を  $M$  とするとき、第 1 行列は下記の条件を満足することができる。

【 数 3 】

$$P \odot W_{oj}^b = \underbrace{\begin{bmatrix} p_{1,1}^b & p_{1,2}^b & \cdots & p_{1,N-1}^b \\ p_{2,1}^b & p_{2,2}^b & \cdots & p_{2,N-1}^b \\ \vdots & \vdots & \ddots & \vdots \\ p_{M,1}^b & p_{M,2}^b & \cdots & p_{M,N-1}^b \end{bmatrix}}_{\text{Matrix 1}} \odot \begin{bmatrix} w_{oj\_1}^b \\ w_{oj\_2}^b \\ \vdots \\ w_{oj\_N-1}^b \end{bmatrix} = \begin{bmatrix} w_{ch\_1}^b \\ w_{ch\_2}^b \\ \vdots \\ w_{ch\_M}^b \end{bmatrix}_{SAC}$$

ここで、 $w_{oj\_i}^b$  は、 $b$  番目サブバンドの  $i$  番目オーディオオブジェクト信号（ $1 = i = N - 1$ ）を表すベクトルであって、前記 S A O C ビットストリームから獲得できる空間キュー情報である。もし、オーディオオブジェクト  $i$  がステレオであれば  $w_{oj\_i}^b$  は  $2 \times 1$  行列のベクトルである。 $p_{i,j}^b$  は、 $j$  番目オーディオオブジェクトが  $i$  番目出力チャネルにマッピングされるためのパワー利得情報または位相情報を表す第 1 行列の要素ベクトルであって、外部から入力されたり、初期値で設定された制御情報（例えば、オブジェクト制御情報、再生システム情報）から獲得できる。

【 0 0 5 9 】

前記数式 3 の条件に適合する第 1 行列は、前記レンダリング部 3 0 3 に伝送され、前記数式 3 は前記レンダリング部 3 0 3 で演算される。

【 0 0 6 0 】

数式 3 の演算子

$\odot$

および演算過程は下記の数式 4 および数式 5 で詳細に説明する。

【数 4】

$$\begin{bmatrix} \mathbf{p}_{1,1}^b & \mathbf{p}_{1,2}^b & \cdots & \mathbf{p}_{1,N-1}^b \end{bmatrix} \odot \begin{bmatrix} \mathbf{w}_{oj\_1}^b \\ \mathbf{w}_{oj\_2}^b \\ \vdots \\ \mathbf{w}_{oj\_i(N-1)}^b \end{bmatrix} = \left[ \mathbf{p}_{1,1}^b \odot \mathbf{w}_{oj\_1}^b + \mathbf{p}_{1,2}^b \odot \mathbf{w}_{oj\_2}^b + \cdots + \mathbf{p}_{1,N-1}^b \odot \mathbf{w}_{oj\_i(N-1)}^b \right]$$

【数 5】

$$\begin{aligned} \mathbf{p}_{i,j}^b \odot \mathbf{w}_{oj\_i}^b &= \begin{bmatrix} p_{1,i,j}^b & p_{2,i,j}^b & \cdots & p_{m,i,j}^b \end{bmatrix} \odot \begin{bmatrix} w_{1,oj\_i}^b \\ w_{2,oj\_i}^b \\ \vdots \\ w_{m,oj\_i}^b \end{bmatrix} \\ &= \begin{bmatrix} p_{1,i,j}^b \times w_{1,oj\_i}^b & p_{2,i,j}^b \times w_{2,oj\_i}^b & \cdots & p_{m,i,j}^b \times w_{m,oj\_i}^b \end{bmatrix} \end{aligned}$$

ここで、入力されるオーディオオブジェクトがステレオである場合、 $m$ は2となる。

【0061】

例えば、入力されるオーディオオブジェクトの個数が $Y$ であり、 $m=2$ であり、出力されるチャンネルの個数が $M$ の場合、第1行列の次元は、 $M \times Y$ となり、 $Y$ 個の $\mathbf{p}_{i,j}^b$ は $2 \times 1$ 行列で構成される。ここで、SACコード201から出力されるオーディオオブジェクトが含まれる場合、 $Y=Y-1$ と見なされる。前記数式3の演算結果、出力されるチャンネルのパワー利得ベクトル $\mathbf{w}_{c,h,j}^b$ で構成された行列が表現されるべきである。前記表現されたベクトルの次元が $M \times 2$ の場合、出力されるチャンネル個数 $M$ および入力されるオーディオオブジェクトのレイアウトである2を反映する。

【0062】

再度、図3の説明に戻って、前記レンダリング部303は、前記第1行列部313および第2行列部311から第1行列および第2行列が伝送される。前記レンダリング部303は、前記パッシング部301によって分離されたSAOCビットストリームから獲得される各オーディオオブジェクトの空間キュー情報 $\mathbf{w}_{o,j,i}^b$ を獲得し、第1行列および第2行列を利用して算出された出力ベクトルを結合して所望の空間キュー情報を獲得し、前記所望の空間キュー情報を含む代表SAOCビットストリームを生成する。ここで、所望の空間キューとは、ユーザが前記SACデコーダ105からの出力を所望する出力マルチチャンネルオーディオ信号と関連した空間キューを意味する。

【0063】

前記第1行列および第2行列に基づいて、所望の空間キュー情報を獲得するための演算は下記の数式6のとおりである。

【数 6】

$$\text{pow}(\mathbf{p}_N) \begin{bmatrix} \mathbf{w}_{ch\_1}^b \\ \mathbf{w}_{ch\_2}^b \\ \vdots \\ \mathbf{w}_{ch\_M}^b \end{bmatrix}_{SAC} + (1 - \text{pow}(\mathbf{p}_N)) \begin{bmatrix} \mathbf{w}_{ch\_1}^b \\ \mathbf{w}_{ch\_2}^b \\ \vdots \\ \mathbf{w}_{ch\_M}^b \end{bmatrix}_{SAOC} = \begin{bmatrix} \mathbf{w}_{ch\_1}^b \\ \mathbf{w}_{ch\_2}^b \\ \vdots \\ \mathbf{w}_{ch\_M}^b \end{bmatrix} = \mathbf{W}_{\text{modified}}^b$$

ここで、 $\mathbf{p}_N$ は第1行列を生成するとき考慮しない事項であって、SACコード201から出力されるオーディオオブジェクトのパワーと、SAOCコード101に直接入力さ

れるオーディオオブジェクトのパワーとの和の比を表す。

【 0 0 6 4 】

前記  $P_N$  は数式 7 のように表すことができる。

【 数 7 】

$$p_N = \frac{\sum_{k=N-1} power(object\#k)}{power(object\#N)}$$

【 0 0 6 5 】

したがって、 $w_{ch\_j}^b$  が出力されるチャンネルのパワーであれば、オーディオオブジェクトがレンダリングされた後、各チャンネルのパワー比は  $w_{modified}^b$  として表す。 $w_{modified}^b$  から所望の空間キューパラメータを新たに抽出することができる。例えば、 $ch\_2$  と  $ch\_1$  との間の CLD パラメータを抽出すれば、下記の数式 8 のとおりである。

【 数 8 】

$$CLD_{ch1/ch2}^b = 20 \log_{10} \frac{w_{ch1}^b}{w_{ch2}^b} = \left[ 20 \log_{10} \frac{w_{ch1,1}^b}{w_{ch2,1}^b}, 20 \log_{10} \frac{w_{ch1,2}^b}{w_{ch2,2}^b} \right]_{m=2}$$

【 0 0 6 6 】

もし、伝送されるダウンミックス信号がモノ信号であれば、CLD パラメータは下記の数式 9 のとおりである。

【 数 9 】

$$CLD_{ch1/ch2}^b = 10 \log_{10} \frac{(w_{ch1,1}^b)^2 + (w_{ch1,2}^b)^2}{(w_{ch2,1}^b)^2 + (w_{ch2,2}^b)^2}$$

【 0 0 6 7 】

出力されるチャンネルのパワー比は、空間キューパラメータである CLD で表現することができ、隣接したチャンネル間の空間キューパラメータは与えられた  $w_{modified}^b$  情報から多様な組合せの形態で示すことができる。前記レンダリング部 303 は、 $w_{modified}^b$  から抽出された空間キュー（例えば、CLD パラメータ）を Huffman コーディング方式などを利用して SAC ビットストリームを生成する。

【 0 0 6 8 】

前記レンダリング部 303 により生成された SAC ビットストリームに含まれる空間キューは、デコーダの特性に応じて分析および抽出される方法が異なる。

例えば、BCC デコーダの場合、1 つのチャンネルを基準として前記数式 8 を利用して  $N-1$  個の CLD パラメータを抽出することができる。また、MP EG サラウンドデコーダの場合、MP EG サラウンドのチャンネル別の比較順に応じて CLD パラメータを抽出することができる。

【 0 0 6 9 】

整理すれば、前記パーシング部 301 は、SAC ビットストリームおよび SAOC ビットストリームを分離し、前記第 2 行列部 311 は、前記パーシング部 301 によって分離された SAC ビットストリームを利用して下記の数式 1 に基づいて第 2 行列を生成し、前記第 1 行列部 313 は、制御信号に対応される第 1 行列を生成する。前記パーシング部 3

01によって分離されたS A O Cビットストリームは前記レンダリング部303に伝送され、前記レンダリング部303は、伝送されたS A O Cビットストリームからオブジェクトの情報を獲得して第1行列と演算し、前記演算結果を第2行列と結合して前記 $W^b_{m o d i f i e d}$ を生成し、生成された $W^b_{m o d i f i e d}$ から空間キューを抽出して代表S A Cビットストリームを生成する。すなわち、生成された $W^b_{m o d i f i e d}$ から空間キューが所望の空間キューとなる。前記代表S A Cビットストリームは、M P E G サラウンドデコーダまたはB C Cデコーダの特性に応じて適切に変換されたビットストリームであって、多様なチャンネルを有するマルチオブジェクト信号に復元されることができる。

#### 【0070】

図4は、図2のビットストリームフォーマッタ205から生成される代表S A O Cビットストリームを示す一実施形態の構造図である。

#### 【0071】

同図に示すように、ビットストリームフォーマッタ205によって生成された代表S A O Cビットストリームは、前記S A O Cコーダ101によって生成されたS A O Cビットストリームと、前記S A Cコーダ201によって生成されたS A Cビットストリームとを結合して生成され、前記代表S A O Cビットストリームは、前記プリセットA S I部203によって生成されたプリセット - A S Iビットストリームを含む。前記プリセット - A S Iビットストリームは下記の図5で詳細に説明する。

#### 【0072】

前記S A O CビットストリームとS A Cビットストリームとを結合する方法は、第1に、各々のビットストリームを直接的にM U Xさせて1つのビットストリームを生成する方法であって、代表S A O CビットストリームでS A O CビットストリームおよびS A Cビットストリームが直列的に接続する401。

#### 【0073】

第2に、S A O C補助データ(ancillary data)領域が存在する場合、前記補助領域にS A Cビットストリーム情報を含ませて1つのビットストリームを生成する方法であって、代表S A O CビットストリームでS A O Cビットストリームおよび補助データ領域が直列的に接続され、前記補助データ領域はS A Cビットストリームを含む403。

#### 【0074】

第3に、前記S A O CビットストリームおよびS A Cビットストリームのうち、類似の空間キューを符号化した領域を同じビットストリームで表現する方法であって、例えば、代表S A O Cビットストリームのヘッダー情報領域には、S A O Cビットストリームヘッダー情報およびS A Cビットストリームヘッダー情報が含まれ、代表S A O Cビットストリームの任意の領域各々には、特定のC L Dと関連したS A O CビットストリームおよびS A Cビットストリームが含まれる405。

#### 【0075】

図5は、図2の代表S A O Cビットストリームを示す更なる一実施形態の構成図であって、前記代表S A O Cビットストリームが複数のプリセット - A S I情報を含む場合を示す。

#### 【0076】

同図に示すように、代表S A O Cビットストリームは、プリセット - A S I領域を含む。前記プリセット - A S I領域は複数のプリセット - A S Iを含み、前記プリセット - A S I情報は、オーディオオブジェクトの制御情報およびレイアウト情報などを含む。

#### 【0077】

前記トランスコーダ103を利用してオーディオオブジェクトをレンダリングする場合、各オーディオオブジェクトの位置情報および制御情報、出力される再生スピーカレイアウト情報が入力されなければならない。前記制御情報および再生スピーカレイアウト情報が入力されない場合、トランスコーダ103において各オーディオオブジェクトの制

御情報およびレイアウト情報は基本値として設定される。

【 0 0 7 8 】

基本値として設定された制御情報およびレイアウト情報を利用したり、入力されたオーディオオブジェクト制御情報およびレイアウト情報は、代表 S A O C ビットストリームまたは代表 S A C ビットストリームの付加情報またはヘッダー情報に含まれる。前記制御情報は 2 種類に表現され得る。第 1 に、各オーディオオブジェクトに対する制御情報（位置およびレベル）およびスピーカーのレイアウト情報を直接的に表現したり、第 2 に、制御情報およびスピーカーのレイアウト情報を第 1 行列の形態で表現し、前記第 1 行列部 3 1 3 の第 1 行列の代わりに利用され得る。

【 0 0 7 9 】

プリセット - A S I 情報とは、オーディオオブジェクト制御情報およびスピーカーのレイアウト情報を表す。すなわち、スピーカーのレイアウト情報およびスピーカーのレイアウト情報に適合するオーディオ場面を構成するための各オーディオオブジェクトの位置およびレベル情報などを含む。

【 0 0 8 0 】

前記パーシング部 3 0 1 によって抽出されたプリセット - A S I 情報を代表 S A C ビットストリームに伝送するために、前記したように、前記プリセット - A S I 情報は直接的に表現されたり、第 1 行列の形態で表現される。

【 0 0 8 1 】

プリセット - A S I 情報が直接的に表現される場合、プリセット - A S I 情報は、再生システムのレイアウト（モノ/ステレオ/マルチチャンネル）、オーディオオブジェクト ID、オーディオオブジェクトレイアウト（モノまたはステレオ）、オーディオオブジェクト位置、方位（Azimuth、0 degree ~ 360 degree）、ステレオ再生時高低（Elevation、- 50 degree ~ 90 degree）、オーディオオブジェクトレベル情報（- 50 dB ~ 50 dB）を含むことができる。

【 0 0 8 2 】

プリセット - A S I 情報が第 1 行列の形態で表現される場合、前記プリセット - A S I 情報を反映した数式 3 の P 行列を構成し、前記 P 行列をレンダリング部 3 0 3 に伝送する。前記 P 行列は、各オーディオオブジェクトが出力されるチャンネルにマッピングされるためのパワー利得情報または位相情報を要素ベクトルとして含む。

【 0 0 8 3 】

前記プリセット - A S I 情報は、入力される同一のオーディオオブジェクトに対して、所望の再生シナリオに対応される様々なオーディオ場면을定義することができる。例えば、ステレオまたはマルチチャンネル（5 . 1、7 . 1）再生システムに要求されるプリセット - A S I 情報がコンテンツ製作者および再生サービスの目的に適合するよう追加的に伝送され得る。

【 0 0 8 4 】

図 6 は、図 2 のトランスコード 1 0 3 を示す更なる一実施形態の構成図であって、外部から入力される制御信号がない場合にプリセット - A S I 情報を活用する一実施形態の構成図を示す。

【 0 0 8 5 】

同図に示すように、トランスコード 1 0 3 は、パーシング部 3 0 1 およびレンダリング部 3 0 3 を備える。前記トランスコード 1 0 3 は、第 2 行列部 3 1 1、第 1 行列部 3 1 3、プリセット - A S I 抽出部 6 0 1、および行列判断部 6 0 3 からの助けを受け得る。

【 0 0 8 6 】

前記パーシング部 3 0 1 は、前記代表 S A O C ビットストリームに含まれた S A O C ビットストリームと S A C ビットストリームとを分離し、前記代表 S A O C ビットストリームに含まれたプリセット - A S I ビットストリームをパーシングしてプリセット - A S I 抽出部 6 0 1 に伝送する。

【 0 0 8 7 】

前記プリセット - A S I 抽出部 6 0 1 は、パーシングされたプリセット - A S I ビットストリームからデフォルト ( d e f a u l t ) プリセット - A S I を出力するが、もし、プリセット - A S I に対する分離 ( s e l e c t i o n ) 要請がある場合、分離要請されたプリセット - A S I を出力する。

【 0 0 8 8 】

前記行列判断部 6 0 3 は、前記プリセット - A S I 抽出部 6 0 1 により出力されたプリセット - A S I が分離されたプリセット - A S I の場合、分離されたプリセット - A S I が第 1 行列の形態であるか判断する。分離したプリセット - A S I が情報を直接的に表現した場合、前記第 1 行列部 3 1 3 に伝送され、前記第 1 行列部 3 1 3 は、前記プリセット - A S I 情報を利用して第 1 行列を生成する。分離したプリセット - A S I が第 1 行列である場合、レンダリング部 3 0 3 に直接入力される信号として利用される。

【 0 0 8 9 】

図 7 は、図 2 のマルチオブジェクトオーディオ符号化器および復号化器にオーディオオブジェクト除去部 7 0 1 が追加的に備えられた場合を示す一実施形態の構造図である。

【 0 0 9 0 】

同図に示すように、オーディオオブジェクト除去部 7 0 1 は、前記 S A O C コーダ 1 0 1 によって生成された代表ダウンミックス信号から任意のオーディオオブジェクトを除去するために利用される。前記オーディオオブジェクト除去部 7 0 1 は、前記 S A O C コーダ 1 0 1 によって生成された代表ダウンミックス信号および前記トランスコーダ 1 0 3 から代表 S A O C ビットストリーム情報を受信し、任意のオーディオオブジェクトを除去する。前記オーディオオブジェクト除去部 7 0 1 に伝送される代表 S A O C ビットストリーム情報は、例えばレンダリング部 3 0 3 により提供され得る。

【 0 0 9 1 】

例えば、前記 S A C コーダ 2 0 1 のダウンミックス信号であるオーディオオブジェクト ( O B J E C T # N ) のみを前記 S A C デコーダ 1 0 5 の入力信号として利用しようとする場合を説明する。

【 0 0 9 2 】

前記 S A O C コーダ 1 0 1 は、入力される各々のオーディオオブジェクトのパワー大きさをサブバンド別に抽出して C L D 値で抽出し、前記 C L D 値を含む S A O C ビットストリームを生成する。任意のサブバンド m に対するパワー情報は、次のように獲得することができる。

【 0 0 9 3 】

$p_{m \text{ object} \# 1}, p_{m \text{ object} \# 2}, \dots, p_{m \text{ object} \# N}$

ここで、 $p_{m \text{ object} \# N}$  は、S A O C コーダ 1 0 1 によって出力される代表ダウンミックス信号の m 番目バンドのパワー大きさである。したがって、オーディオオブジェクト除去部 7 0 1 に入力される代表ダウンミックス信号を  $u(n)$  とし、前記代表ダウンミックス信号を周波数領域に変換したものを  $U(f)$  とし、オーディオオブジェクト除去部 7 0 1 の出力信号、すなわち前記 S A C デコーダ 1 0 5 の入力信号を  $U^{modified}(f)$  とすれば、 $U^{modified}(f)$  は、前記 S A C コーダ 2 0 1 のダウンミックス信号のオーディオオブジェクト ( O B J E C T # N ) と対応し、下記の数式 1 0 のように表すことができる。

【 数 1 0 】

$$U^{modified}(f) = U(f) \times \sqrt{\frac{p_{m \text{ object} \# N}}{\sum_{i=1}^N p_{m \text{ object} \# i}}} \times \delta, \quad A(m+1) \leq f \leq A(m+1)-1$$

ここで、 $A(m)$  は、m 番目のサブバンドの周波数領域で境界を示し、 $\delta$  は、レベルの

大きさを調整するための任意の定数値であり、 $U(f)$ は、モノまたはステレオである。

【0094】

以下、 $U(f)$ がモノである場合を説明する。ただ、 $U(f)$ がステレオである場合、 $U(f)$ が左/右チャンネルに分離して処理されること以外はモノの場合と同一である。

【0095】

前記  $U_{modified}(f)$  は、前記 SAC コーダ 201 のダウンミックス信号のオーディオオブジェクト (OBJECT # N) と同一に見られる。したがって、前記 SAC デコーダ 105 に入力される代表 SAC ビットストリームは、前記代表 SAOC ビットストリームから前記 SAOC ビットストリームが除外されたビットストリームであって、SAC コーダ 201 から出力される SAC ビットストリームと同一に利用することができる。すなわち、前記 SAC デコーダ 105 は、OBJECT # N を受信し、M 個のマルチチャンネル信号に復元できる。しかし、全体信号のレベルは、トランスコーダ 103 のレンダリング部 303 により調整されたり、または数式 10 に任意の定数値を乗算して OBJECT # N の信号レベルを調整することによって調整できる。

【0096】

一方、一実施形態において、前記 SAC コーダ 201 のダウンミックス信号のオーディオオブジェクト (OBJECT # N) のみを前記 SAC デコーダ 105 の入力信号から除去する場合について説明する。前記数式 10 は、下記の数式 11 のように示される。

【数 11】

$$U^{modified}(f) = U(f) \times \sqrt{\frac{\sum_{i=1}^{N-1} P_m^{object\#_i}}{\sum_{i=1}^N P_m^{object\#_i}}} \times \delta, \quad A(m+1) \leq f \leq A(m+1)-1$$

【0097】

したがって、前記 SAC デコーダ 105 に入力される代表 SAC ビットストリームは、前記代表 SAOC ビットストリームから SAC コーダ 201 の前記 SAC ビットストリームが除外されたビットストリームであって、レンダリング部 303 の第 2 行列に対する出力がないこととみられる。すなわち、前記トランスコーダ 103 は、代表 SAOC ビットストリーム部をパッシングし、OBJECT # N に関する情報を除去した残りのオーディオオブジェクト情報のみをレンダリングして代表 SAC ビットストリームを生成する。

【0098】

したがって、OBJECT # N に該当するパワー利得情報、相関性情報などを代表 SAC ビットストリームに含めない。前記数式 11 は数式 10 と同様に、 $\delta$  はレベルの大きさを調整するための任意の定数値であって、全体の出力信号レベルを調整することができる。

【0099】

前記オーディオオブジェクト除去部 701 は代表ダウンミックス信号からオーディオオブジェクトを除去し、除去命令はトランスコーダ 103 に入力される制御信号によって決定される。前記オーディオオブジェクト除去部 701 は、時間領域信号および周波数領域信号のすべてが適用可能である。また、前記代表ダウンミックス信号をサブバンドで分割するために DFT (Discrete Fourier Transform) または QMF (Quadrature Mirror Filter bank) を利用することができる。

【0100】

トランスコーダ 103 のレンダリング部 303 は、SAOC ビットストリームまたは SAC ビットストリームを除去して SAC デコーダ 105 に伝送し、前記オーディオオブジェクト除去部 701 は、前記 SAC デコーダ 105 に伝送されるビットストリームに対応

されるようオーディオオブジェクトを適切に除去する。

【0101】

前記トランスコード103がSACデコード105に含まれる場合、前記トランスコード103から出力される代表SACビットストリームは、付加の変換過程なしでSACデコード105に伝送され得る。ここで、付加の変換過程とは、量子化やHuffmanコーディング方法のような一般的なコーディング過程を意味する。

【0102】

また、SAOCコード101がSACコード201と接続しないものと見なし、SACコード201の出力オーディオオブジェクトを除去したSAOCコード101に入力されるオーディオオブジェクト、すなわちOBJECT#1～OBJECT#N-1のみを制御して復元することができる。

【0103】

図8は、図2のSACコード201およびSACデコード105をMP EGサラウンドコードおよびデコードに代替した場合を示す一実施形態の構造図である。

【0104】

同図に示すように、SACコード201はMP EGサラウンドコード(MPSコード)801に、SACデコード105はMP EGサラウンドデコード(MPSデコード)805に代替される。また、前記SAOCコード101から出力される代表ダウンミックス信号がステレオである場合、追加的に信号処理部803が要求される。

【0105】

前記MPSコード801は、図2のSACコード201と同一の役割を行なう。すなわち、前記MPSコード801は、入力されるマルチチャネルオーディオ信号から1つのオーディオオブジェクトを出力し、空間キューおよび付加情報を抽出してMPSビットストリームを生成する。前記出力された1つのオーディオオブジェクトは、ダウンミックスされたモノまたはステレオ信号である。

【0106】

また、前記MPSデコード805は、図2のSACデコード805と同一の役割を行なう。すなわち、前記MPSデコード805は、前記トランスコード103から出力されるSACビットストリームを利用し、前記SAOCコード101から出力されるダウンミックス信号または信号処理部803から出力される代表再ダウンミックス信号を多様なチャネルのマルチオブジェクトオーディオ信号に復元する。

【0107】

一方、前記信号処理部805は、前記SAOCコード101から出力されるダウンミックス信号がステレオである場合、すなわち前記MPSデコード805がステレオ信号を処理する場合、MPSデコード805がステレオ信号の左/右処理に制約されるために要求される。

【0108】

前記数式2は、一般的なSACデコードでダウンミックス信号がM個に一般化された場合を示す。ダウンミックス信号がステレオである場合、復元される出力チャネル1に対する数式2は下記の数式12のとおりである。

【数12】

$$\mathbf{w}_{ch\_1}^b \times \mathbf{u}_{SAC}^b(k) = \begin{bmatrix} w_{L,ch\_1}^b & w_{R,ch\_2}^b \end{bmatrix} \begin{bmatrix} u_L^b(k) \\ u_R^b(k) \end{bmatrix}$$

ここで、出力チャネルのベクトルは、すべてのダウンミックス信号に対して適用可能すべきであるが、現在MPSデコード805では不可能である。前記MPSデコード805は、下記の数式13のように、行列値が0に制限されるためである。

【数 1 3】

$$\mathbf{w}_{ch\_1}^b \times \mathbf{u}_{SAC}^b(k) = \begin{bmatrix} w_{L,ch\_1}^b & 0 \end{bmatrix} \begin{bmatrix} u_L^b(k) \\ u_R^b(k) \end{bmatrix}$$

【0 1 0 9】

すなわち、出力チャンネル 1 の復元において、 $u_R^b(k)$  要素が反映されないので、前記数式 3、数式 4、数式 5 によって生成された  $w_{ch\_2}^b$  が適用されない。したがって、ステレオ以上のレイアウトを有する信号に対して流動的な位置選定 (flexible positioning) が不可能である。すなわち、ステレオ信号の左側信号および右側信号の相互間に自由なレンダリングが不可能である。

【0 1 1 0】

しかし、前記信号処理部 803 を利用して前記 SAOC コーダ 101 から出力される代表ダウンミックス信号を再度ダウンミックスし、代表再ダウンミックス信号として出力する。前記信号処理部 803 の処理過程は、下記の数式 14 のとおりである。

【数 1 4】

$$\begin{bmatrix} \mathbf{w}_{ch\_1}^b \\ \mathbf{w}_{ch\_2}^b \\ \vdots \\ \mathbf{w}_{ch\_M}^b \end{bmatrix}_{\text{modified}} \times \begin{bmatrix} \mathbf{u}_{stereo}^b(k) \end{bmatrix} = \begin{bmatrix} y_{ch\_1}^b(k) \\ y_{ch\_2}^b(k) \\ \vdots \\ y_{ch\_M}^b(k) \end{bmatrix}$$

【0 1 1 1】

前記 SAOC コーダ 101 から出力される代表ダウンミックス信号がステレオである場合、前記信号処理部 803 の出力信号は、下記の数式 15 のとおりである。

【数 1 5】

$$\begin{bmatrix} \mathbf{w}_L^b \\ \mathbf{w}_R^b \end{bmatrix}_{\text{modified}} \times \begin{bmatrix} \mathbf{u}_{stereo}^b(k) \end{bmatrix} = \begin{bmatrix} y_{ch\_L}^b(k) \\ y_{ch\_R}^b(k) \end{bmatrix}$$

ここで、 $y_{ch\_L}^b(k)$ 、 $y_{ch\_R}^b(k)$  は、前記信号処理部 803 により出力される信号であって、前記 MPS デコーダ 805 に入力される。 $y_{ch\_L}^b(k)$  および  $y_{ch\_R}^b(k)$  は、数式 15 のように左側信号および右側信号のレンダリングがすべて反映された信号であるため、前記 MPS デコーダ 805 が前記数式 13 のように制限されても、前記 MPS デコーダ 805 は、左側信号および右側信号が自由にレンダリングされた信号を出力することができる。

【0 1 1 2】

例えば、前記  $w_L^b$ 、 $w_R^b$  が MPS デコーダ 805 により 5 チャンネルに復元される場合、前記数式 14 において  $w_L^b$ 、 $w_R^b$  は次のように表現され得る。

【数 1 6】

$$(e.g., \mathbf{w}_L^b = \mathbf{w}_{ch\_Lf}^b + \mathbf{w}_{ch\_Ls}^b + \mathbf{w}_{ch\_C}^b / \sqrt{2}, \mathbf{w}_R^b = \mathbf{w}_{ch\_Rf}^b + \mathbf{w}_{ch\_Rs}^b + \mathbf{w}_{ch\_C}^b / \sqrt{2})$$

## 【0113】

前記したように、MPEGサラウンドの制約によってMPSデコーダ805がステレオ信号の処理が困難な場合、前記信号処理部803は、前記トランスコーダ103から伝送されたオブジェクト位置情報を利用し、再度ダウンミックスして代表再ダウンミックス信号を出力する。前記信号処理部803に伝送されるオブジェクト位置情報は、例えばレンダリング部303により提供され得る。ここで、レンダリング部303は、前述で説明したものと類似した方式で前記代表SAOCビットストリームに基づいて前記SAOCコーダ101およびMPSコーダ801に入力されるオーディオ信号に対し、MPSデコーダ805により出力されるオーディオ信号の左側信号および右側信号の各々に対する空間キュー情報が含まれた代表MPSビットストリームを生成することができる。

## 【0114】

前記MPSデコーダ805は、前記信号処理部805とともに動作することによって、図2のSACデコーダ205と同一の役割を行なうことができる。

## 【0115】

前記MPSデコーダ805は、前記信号処理部803から出力された代表再ダウンミックス信号を所望の出力、すなわち多様なチャンネルを有するマルチオブジェクト信号に復元する。

## 【0116】

前記図2のSACデコーダ205または前記信号処理部805とともに動作する前記MPSデコーダ805の復号化方法は、マルチチャンネルマルチオブジェクトダウンミックス信号およびマルチチャンネルマルチオブジェクト付加情報(side information)信号を受信するステップと、前記マルチチャンネルマルチオブジェクトダウンミックス信号をマルチチャンネルダウンミックス信号に変換するステップと、前記マルチチャンネルマルチオブジェクト付加情報信号をマルチチャンネル付加情報信号に変換するステップと、前記変換されたマルチチャンネルダウンミックス信号およびマルチチャンネル付加情報信号を利用してオーディオ信号を合成するステップとを含む。

## 【0117】

前記マルチチャンネルダウンミックス信号の変換ステップは、前記マルチチャンネルマルチオブジェクト付加情報信号から獲得されるオブジェクト関連情報を利用し、前記マルチチャンネルマルチオブジェクトダウンミックス信号からオブジェクト情報を除去するステップを含む。前記マルチチャンネルダウンミックス信号の変換ステップは、前記マルチチャンネルマルチオブジェクト付加情報信号から獲得されるオブジェクト関連情報を利用し、前記マルチチャンネルマルチオブジェクトダウンミックス信号からオブジェクト情報を制御するステップを含む。

## 【0118】

ここで、前記マルチチャンネルダウンミックス信号の変換ステップで、前記オブジェクト関連情報はオブジェクト制御情報によって制御され得る。ここで、前記オブジェクト関連情報は、復号化システム情報によって制御され得る。

## 【0119】

以上で説明された本発明に係る符号化および復号化過程は装置の観点で説明したが、前記装置に備えられた各装置的な構成要素は、プロセス的構成要素に代替され得、この場合、本発明に係る符号化および復号化過程は、方法の観点で理解されることができるとは自明である。

## 【0120】

前述のような本発明の方法は、プログラムで具現され、コンピュータで読み出し可能な形態で記録媒体(CD-ROM、RAM、ROM、フロッピーディスク、ハードディスク、光磁気ディスクなど)に保存され得る。このような過程は、本発明が属する技術分野における通常の知識を有する者が容易に実施できるため、これ以上詳細に説明しない。

## 【0121】

以上で説明した本発明は、前述した実施形態および添付の図面によって限定されるので

なく、本発明の技術的な思想から脱離しない範囲内で様々な置換、変形、および変更が可能であることが、本発明が属する技術分野における通常の知識を有する者にとって明白であろう。

【手続補正 2】

【補正対象書類名】特許請求の範囲

【補正対象項目名】全文

【補正方法】変更

【補正の内容】

【特許請求の範囲】

【請求項 1】

マルチオブジェクトオーディオ信号を復号化するためのレンダリング情報を生成するトランスコーディング装置であって、

符号化されたオーディオオブジェクト信号に対するパワー利得情報および出力位置情報を含むレンダリングマトリックス情報を受信する受信手段と、

前記受信手段によって受信されたレンダリングマトリックス情報および符号化装置から入力される前記符号化されたオーディオオブジェクト信号に対するレンダリング情報に基づいて、復号化装置から出力されるオーディオ信号に対する空間キュー情報を生成するレンダリング手段と、

を備えることを特徴とするトランスコーディング装置。

【請求項 2】

前記レンダリング手段は、前記符号化されたオーディオオブジェクト信号のうちの所定オーディオオブジェクト信号に対する空間キュー情報を除いた、前記復号化装置から出力されるオーディオオブジェクト信号に対する空間キュー情報を生成し、

前記トランスコーディング装置は、前記符号化されたオーディオ信号のうちから前記所定オーディオオブジェクト信号を除去するオーディオオブジェクト除去手段をさらに含むことを特徴とする請求項 1 に記載のトランスコーディング装置。

【手続補正 3】

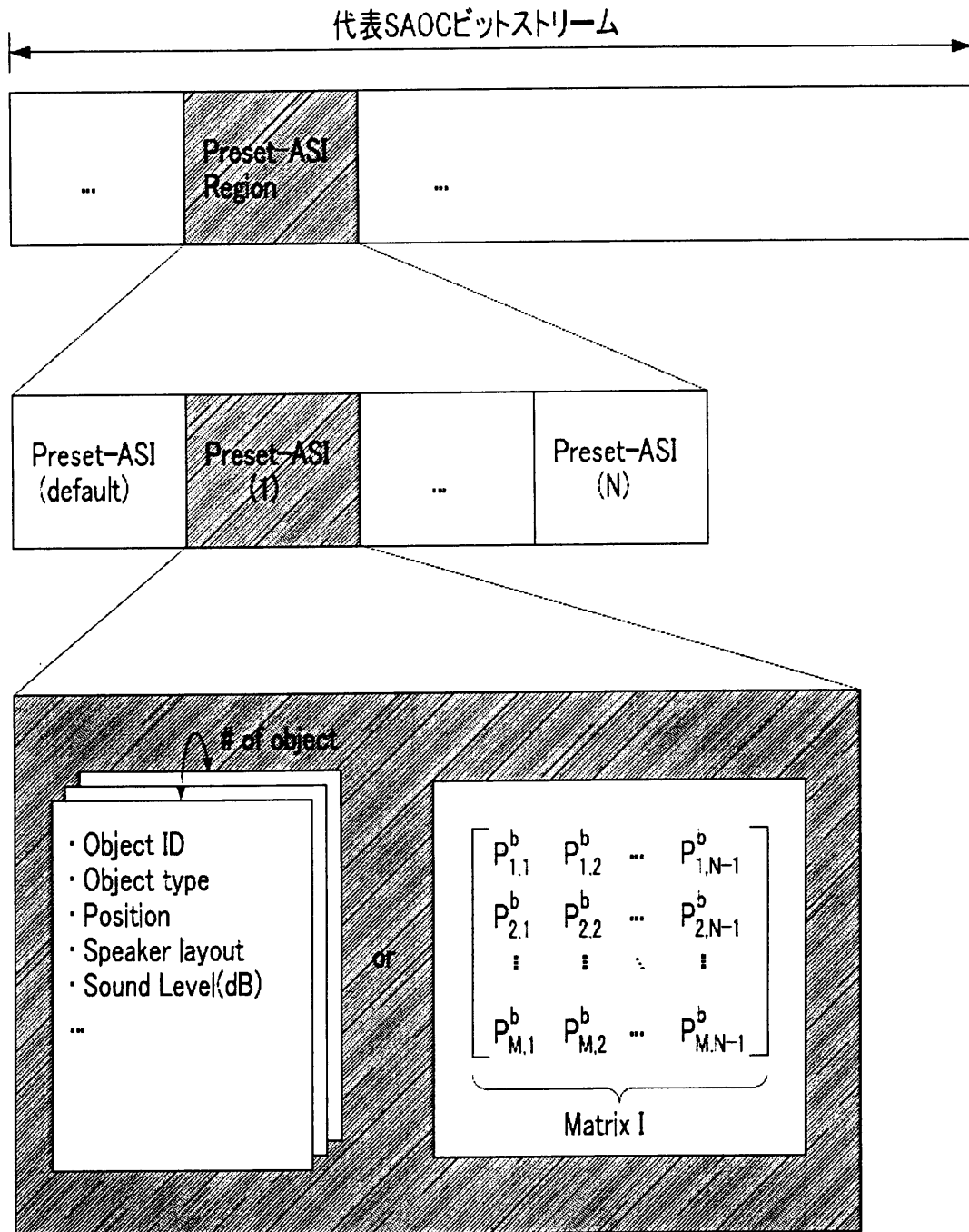
【補正対象書類名】図面

【補正対象項目名】図 5

【補正方法】変更

【補正の内容】

【 図 5 】



【 手続補正 4 】

【 補正対象書類名 】 図面

【 補正対象項目名 】 図 8

【 補正方法 】 変更

【 補正の内容 】

【図 8】

