

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第6542269号
(P6542269)

(45) 発行日 令和1年7月10日 (2019.7.10)

(24) 登録日 令和1年6月21日 (2019.6.21)

(51) Int. Cl.

F I

G 1 0 L 19/008 (2013.01)

G 1 0 L 19/008 1 0 0

G 1 0 L 19/02 (2013.01)

G 1 0 L 19/02 1 5 0

請求項の数 24 (全 61 頁)

(21) 出願番号 特願2016-573839 (P2016-573839)
 (86) (22) 出願日 平成27年7月2日 (2015.7.2)
 (65) 公表番号 特表2017-523451 (P2017-523451A)
 (43) 公表日 平成29年8月17日 (2017.8.17)
 (86) 国際出願番号 PCT/EP2015/065086
 (87) 国際公開番号 WO2016/001356
 (87) 国際公開日 平成28年1月7日 (2016.1.7)
 審査請求日 平成30年6月26日 (2018.6.26)
 (31) 優先権主張番号 14306080.4
 (32) 優先日 平成26年7月2日 (2014.7.2)
 (33) 優先権主張国 欧州特許庁 (EP)
 (31) 優先権主張番号 14194186.4
 (32) 優先日 平成26年11月20日 (2014.11.20)
 (33) 優先権主張国 欧州特許庁 (EP)

(73) 特許権者 510185767
 ドルビー・インターナショナル・アーベー
 オランダ王国、セーエン アムステルダム
 ズイドースト 1101, ヘリケルベル
 グウェグ 1-35, アポロ ビルディン
 グ 3エー
 (74) 代理人 100107766
 弁理士 伊東 忠重
 (74) 代理人 100070150
 弁理士 伊東 忠彦
 (74) 代理人 100091214
 弁理士 大貫 進介

最終頁に続く

(54) 【発明の名称】 圧縮HOA表現をデコードする方法および装置ならびに圧縮HOA表現をエンコードする方法および装置

(57) 【特許請求の範囲】

【請求項 1】

圧縮されたHOA表現をデコードする本方法であって、

・前記圧縮されたHOA表現から、複数の打ち切りされたHOA係数シーケンス、前記打ち切りされたHOA係数シーケンスのシーケンス・インデックスを示すまたは含む割り当てベクトル、サブバンドに関係した方向情報、複数の予測行列および利得制御サイド情報を抽出する段階と；

・前記複数の打ち切りされたHOA係数シーケンス、前記利得制御サイド情報および前記割り当てベクトルから打ち切りされたHOA表現を再構成する段階と；

・分解フィルタバンクにおいて、前記再構成された打ち切りされたHOA表現を複数F個の周波数サブバンドについての周波数サブバンド表現に分解する段階と；

・各周波数サブバンド表現について方向性サブバンド合成ブロックにおいて、前記再構成された打ち切りされたHOA表現のそれぞれの周波数サブバンド表現と、前記サブバンドに関係した方向情報と、前記予測行列とから、予測された方向性HOA表現を合成する段階と；

・前記F個の周波数サブバンドのそれぞれについてサブバンド組成ブロックにおいて、係数シーケンスをもつデコードされたサブバンドHOA表現を組成する段階であって、該係数シーケンスは、該係数シーケンスが前記割り当てベクトルに含まれるインデックスnをもつ場合には前記打ち切りされたHOA表現の係数シーケンスから得られ、それ以外の場合には前記方向性サブバンド合成ブロックの一つによって与えられる前記予測された方向性HO

10

20

A成分の係数シーケンスから得られる、段階と；

・合成フィルタバンクにおいて、前記デコードされたサブバンドHOA表現を合成して、デコードされたHOA表現を得る段階とを含む、方法。

【請求項 2】

前記抽出する段階が、エンコードされた打ち切りされたHOA係数シーケンスを含む知覚的に符号化された部分を取得することを含み、知覚的デコーダにおいて、前記エンコードされた打ち切りされたHOA係数シーケンスを知覚的にデコードして、前記打ち切りされたHOA係数シーケンスを取得することをさらに含む、請求項 1 記載の方法。

【請求項 3】

前記抽出する段階が、エンコードされたサイド情報部分を取得することを含み、サイド情報源デコーダにおいて、前記エンコードされたサイド情報部分をデコードして、前記サブバンドに関係した方向情報、予測行列、利得制御サイド情報および割り当てベクトルを取得することをさらに含む、請求項 1 記載の方法。

【請求項 4】

前記サブバンドに関係した方向情報は、アクティブな方向の集合 ($M_{DIR}(k)$) と、第一および第二のインデックスをもつインデックスのタプルを含むタプル集合 ($M_{DIR}(k, f_1), \dots, M_{DIR}(k, f_F)$) とを含み、前記第二のインデックスは、現在の周波数サブバンドについてのアクティブな方向の前記集合内のアクティブな方向のインデックスであり、前記第一のインデックスはそのアクティブな方向の軌跡インデックスであり、軌跡とは特定の音源の方向の時間的シーケンスである、請求項 1 記載の方法。

【請求項 5】

少なくとも一つの周波数サブバンド表現が二つ以上の周波数サブバンドのサブバンド・グループを含む、請求項 1 記載の方法。

【請求項 6】

サブバンド・グループ構成情報が前記圧縮されたHOA表現から受領または抽出され、前記サブバンド・グループ構成情報が前記合成フィルタバンクをセットアップするために使われる、請求項 5 記載の方法。

【請求項 7】

所与の数の係数シーケンスを有する入力HOA信号のフレームをエンコードする方法であって、各係数シーケンスはインデックスをもち、当該方法が、

・打ち切りされたHOA表現に含まれるべきアクティブな係数シーケンスのインデックスの集合を決定する段階と；

・低減された数の0でない係数シーケンスもつ前記打ち切りされたHOA表現を計算する段階と；

・前記入力HOA信号から、候補方向の第一の集合を推定する段階と；

・前記入力HOA信号を複数の周波数サブバンドに分割する段階であって、それらの周波数サブバンドの係数シーケンスが得られる、段階と；

・各周波数サブバンドについて、方向の第二の集合を推定する段階であって、方向の前記第二の集合の各要素は第一および第二のインデックスをもつインデックスのタプルであり、前記第二のインデックスは現在の周波数サブバンドについてのアクティブな方向のインデックスであり、前記第一のインデックスは該アクティブな方向の軌跡インデックスであり、各アクティブな方向は前記入力HOA信号の候補方向の前記第一の集合にも含まれるものであり、

・各周波数サブバンドについて、その周波数サブバンドの係数シーケンスからそれぞれの周波数サブバンドの方向の前記第二の集合に従って、方向性サブバンド信号を計算する段階と；

・各周波数サブバンドについて、それぞれの周波数サブバンドのアクティブな係数シーケンスのインデックスの前記集合を使ってその周波数サブバンドの係数シーケンスから方向性サブバンド信号を予測するよう適応された予測行列を計算する段階と；

10

20

30

40

50

・候補方向の前記第一の集合、方向の前記第二の集合、前記予測行列および前記打ち切りされたHOA表現をエンコードする段階とを含む、方法。

【請求項 8】

二つ以上のサブバンドの少なくとも一つのグループが生成され、前記少なくとも一つのグループが単一のサブバンドの代わりに使われ、単一のサブバンドと同じように扱われる、請求項 7 記載の方法。

【請求項 9】

前記打ち切りされたHOA表現のエンコードは、

- ・前記打ち切りされたHOAチャンネル・シーケンスの部分的な脱相関と；
 - ・前記打ち切りされたHOAチャンネル・シーケンスをトランスポート・チャンネルに割り当てるためのチャンネル割り当てと；
 - ・前記トランスポート・チャンネルのそれぞれに対して利得制御を実行する段階であって、各トランスポート・チャンネルについての利得制御サイド情報が生成される、段階と；
 - ・知覚的エンコーダにおいて、利得制御された打ち切りされたHOAチャンネル・シーケンスをエンコードする段階と；
 - ・前記利得制御サイド情報、候補方向の前記第一の集合、方向の前記第二の集合および前記予測行列をサイド情報源符号化器においてエンコードする段階と；
 - ・前記知覚的エンコーダおよび前記サイド情報源符号化器の出力を多重化して、エンコードされたHOA信号フレームを得る段階とを含む、
- 請求項 7 記載の方法。

【請求項 10】

各周波数サブバンドについて方向の第二の集合を推定する前記段階において、周波数サブバンドの方向は、フルバンドHOA信号の方向 ($M_{DIR}(k)$) のうちでのみ探索される、請求項 7 記載の方法。

【請求項 11】

アクティブな方向の軌跡を決定する段階をさらに含み、アクティブな方向とは音源の方向であり、軌跡とは特定の音源の方向の時間的なシーケンスである、請求項 7 記載の方法。

【請求項 12】

打ち切りされたHOA表現とは、一つまたは複数の係数シーケンスが0に設定されているHOA信号である、請求項 7 記載の方法。

【請求項 13】

HOA信号をデコードする装置であって、

- ・前記圧縮されたHOA表現から、複数の打ち切りされたHOA係数シーケンス、前記打ち切りされたHOA係数シーケンスのシーケンス・インデックスを示すまたは含む割り当てベクトル、サブバンドに関係した方向情報、複数の予測行列および利得制御サイド情報を抽出するよう構成された抽出モジュールと；
 - ・前記複数の打ち切りされたHOA係数シーケンス、前記利得制御サイド情報および前記割り当てベクトルから、打ち切りされたHOA表現を再構成するよう構成された再構成モジュールと；
 - ・前記再構成された打ち切りされたHOA表現を複数F個の周波数サブバンドについての周波数サブバンド表現に分解するよう構成された分解フィルタバンク・モジュールと；
 - ・各周波数サブバンドについて、前記再構成された打ち切りされたHOA表現のそれぞれの周波数サブバンド表現と、前記サブバンドに関係した方向情報と、前記予測行列とから、予測された方向性HOA表現を合成するよう構成された少なくとも一つの方向性サブバンド合成モジュールと；
 - ・前記F個の周波数サブバンドのそれぞれについて、係数シーケンスをもつデコードされたサブバンドHOA表現を組成するよう構成された少なくとも一つのサブバンド組成モジュール
- 55であって、該係数シーケンスは、該係数シーケンスが前記割り当てベクトルに含

まれるインデックス n をもつ場合には前記打ち切りされたHOA表現の係数シーケンスから得られ、あるいはそれ以外の場合には前記方向性サブバンド合成モジュールの一つによって与えられる前記予測された方向性HOA成分の係数シーケンスから得られる、モジュールと；

・前記デコードされたサブバンドHOA表現を合成して、デコードされたHOA表現を得るよう構成された合成フィルタバンク・モジュールとを有する、装置。

【請求項 1 4】

前記抽出モジュールが、少なくとも、

- ・エンコードされたサイド情報部分と、エンコードされた打ち切りされたHOA係数シーケンスを含む知覚的に符号化された部分とを取得するためのデマルチプレクサと；
- ・前記エンコードされた打ち切りされたHOA係数シーケンスを知覚的にデコードして前記打ち切りされたHOA係数シーケンスを得るよう構成された知覚的デコーダと；
- ・前記エンコードされたサイド情報部分をデコードして、前記サブバンドに関係した方向情報、予測行列、利得制御サイド情報および割り当てベクトルを得るよう構成されたサイド情報源デコーダとを有する、

請求項 1 3 記載の装置。

【請求項 1 5】

前記抽出モジュールが、エンコードされたサイド情報部分を取得し、さらに、前記エンコードされたサイド情報部分をデコードして、前記サブバンドに関係した方向情報、予測行列、利得制御サイド情報および割り当てベクトルを取得するよう構成されたサイド情報源デコーダを有する、請求項 1 3 記載の装置。

【請求項 1 6】

前記サブバンドに関係した方向情報は、アクティブな方向の集合 ($M_{DIR}(k)$) と、第一および第二のインデックスをもつインデックスのタプルを含むタプル集合 ($M_{DIR}(k, f_1), \dots, M_{DIR}(k, f_F)$) とを含み、前記第二のインデックスは、現在の周波数サブバンドについてのアクティブな方向の前記集合内のアクティブな方向のインデックスであり、前記第一のインデックスはそのアクティブな方向の軌跡インデックスであり、軌跡とは特定の音源の方向の時間的シーケンスである、請求項 1 3 記載の装置。

【請求項 1 7】

少なくとも一つの周波数サブバンド表現が二つ以上の周波数サブバンドのサブバンド・グループを含む、請求項 1 3 記載の装置。

【請求項 1 8】

サブバンド・グループ構成情報が前記圧縮されたHOA表現から受領または抽出され、前記サブバンド・グループ構成情報が前記合成フィルタバンクをセットアップするために使われる、請求項 1 7 記載の装置。

【請求項 1 9】

所与の数の係数シーケンスをもつ入力HOA信号のフレームをエンコードするための装置であって、各係数シーケンスはインデックスをもち、当該装置が、

- ・低減された数の0でない係数シーケンスをもつ打ち切りされたHOA表現を計算するよう構成され、前記打ち切りされたHOA表現に含まれているアクティブな係数シーケンスのインデックスの集合を決定するさらに構成された計算および決定モジュールと；
- ・前記入力HOA信号を複数の周波数サブバンドに分割するよう構成された分解フィルタバンク・モジュールであって、それらの周波数サブバンドの係数シーケンスが得られる、分解フィルタバンク・モジュールと；

・前記入力HOA信号から、候補方向の第一の集合を推定するよう構成され、前記周波数サブバンドのそれぞれについて、方向の第二の集合を推定するようさらに構成された方向推定モジュールであって、方向の前記第二の集合の各要素は第一および第二のインデックスをもつインデックスのタプルであり、前記第二のインデックスは現在の周波数サブバンドについてアクティブな方向のインデックスであり、前記第一のインデックスは該アクティ

10

20

30

40

50

ブな方向の軌跡インデックスであり、各アクティブな方向は前記入力HOA信号の候補方向の前記第一の集合にも含まれる、方向推定モジュールと；

・前記周波数サブバンドのそれぞれについて、それぞれの周波数サブバンドの方向の前記第二の集合に応じてその周波数サブバンドの係数シーケンスから方向性サブバンド信号を計算するよう構成された少なくとも一つの方向性サブバンド計算モジュールと；

・前記周波数サブバンドのそれぞれについて、それぞれの周波数サブバンドのアクティブな係数シーケンスのインデックスの前記集合を使ってその周波数サブバンドの係数シーケンスから方向性サブバンド信号を予測するよう適応された予測行列を計算するよう構成された少なくとも一つの方向性サブバンド予測モジュールと；

・候補方向の前記第一の集合、方向の前記第二の集合、前記予測行列および前記打ち切りされたHOA表現をエンコードするよう構成されたエンコード・モジュールとを有する、装置。

10

【請求項 2 0】

二つ以上のサブバンドの少なくとも一つのグループが生成され、前記少なくとも一つのグループが単一のサブバンドの代わりに使われ、単一のサブバンドと同じように扱われる、請求項 1 9 記載の装置。

【請求項 2 1】

・前記打ち切りされたHOAチャンネル・シーケンスを部分的に脱相関させるよう構成された部分的脱相関器と；

・前記打ち切りされたHOAチャンネル・シーケンスをトランスポート・チャンネルに割り当てるよう構成されたチャンネル割り当てモジュールと；

・前記トランスポート・チャンネルに対して利得制御を実行するよう構成された少なくとも一つの利得制御ユニットであって、各トランスポート・チャンネルについての利得制御サイド情報が生成される、ユニットとをさらに有しており、

前記エンコード・モジュールが、

・利得制御された打ち切りされたHOAチャンネル・シーケンスをエンコードするよう構成された知覚的エンコーダと；

・前記利得制御サイド情報、候補方向の前記第一の集合、方向の前記第二の集合および前記予測行列をエンコードするよう構成されたサイド情報源符号化器と；

・前記知覚的エンコーダおよび前記サイド情報源符号化器の出力を多重化して、エンコードされたHOA信号フレームを得るよう構成されたマルチプレクサとを有する、

請求項 1 9 記載の装置。

20

30

【請求項 2 2】

前記方向推定モジュールが、各周波数サブバンドについて方向の第二の集合を推定するときに、周波数サブバンドの方向を、フルバンドHOA信号の方向 ($M_{DIR}(k)$) のうちでのみ探索する、請求項 1 9 記載の装置。

【請求項 2 3】

アクティブな方向の軌跡を決定するよう構成された軌跡決定モジュールをさらに有しており、アクティブな方向とは音源の方向であり、軌跡とは特定の音源の方向の時間的なシーケンスである、請求項 1 9 記載の装置。

【請求項 2 4】

打ち切りされたHOA表現とは、一つまたは複数の係数シーケンスが0に設定されているHOA信号である、請求項 1 9 記載の装置。

【発明の詳細な説明】

【技術分野】

【0 0 0 1】

本発明は、所与の数の係数シーケンスを有する入力HOA信号のフレームをエンコードする方法、HOA信号をデコードする方法、所与の数の係数シーケンスを有する入力HOA信号のフレームをエンコードする装置、HOA信号をデコードする装置に関する。

【背景技術】

40

50

【 0 0 0 2 】

高次アンビソニックス (HOA: Higher Order Ambisonics) は、波面合成 (WFS: wave field synthesis) または22.2として知られるもののようなチャンネル・ベースのアプローチのような他の技法もあるうちの、三次元音を表現するための一つの可能性を提供する。チャンネル・ベースの方法とは対照的に、HOA表現は特定のスピーカー・セットアップとは独立であるという利点をもたらす。この柔軟性は、特定のスピーカー・セットアップでのHOA表現の再生のために必要とされるデコード・プロセスの代償を伴う。必要とされるスピーカーの数が通例非常に多いWFSアプローチに比べ、HOAは少数のスピーカーのみからなるセットアップにレンダリングされてもよい。HOAのさらなる利点は、同じ表現を、いかなる修正もなしでヘッドフォンへのバイノーラル・レンダリングのために用いることもできるということである。

10

【 0 0 0 3 】

HOAは、複素調和平面波振幅のいわゆる空間密度の、打ち切りされた球面調和関数 (SH) 展開による表現に基づく。各展開係数は角周波数の関数であり、これは時間領域関数によって等価に表現できる。よって、一般性を失うことなく、完全なHOA音場表現は、実際に、0個の時間領域関数からなると理解できる。ここで、0は展開係数の数を表わす。これらの時間領域関数は、以下では、等価だが、HOA係数シーケンスまたはHOAチャンネルと称される。

【 0 0 0 4 】

HOA表現の空間分解能は、展開の最大次数Nの増大とともに改善する。残念ながら、展開係数の数0は次数Nとともに二次で、特に $0 = (N+1)^2$ の形で増大する。たとえば、次数N=4を使う典型的なHOA表現は0=25個のHOA (展開) 係数を必要とする。上記の考察によれば、HOA表現の伝送のための全ビットレートは、所望される単一チャンネル・サンプリング・レート f_s およびサンプル当たりのビット数 N_b を与えられて、 $0 \cdot f_s \cdot N_b$ によって決定される。結果として、たとえば次数N=4のHOA表現を、 $f_s = 48\text{kHz}$ のサンプリング・レートで、サンプル当たり $N_b = 16$ ビットを用いて伝送することは、19.2Mbits/sのビットレートにつながる。これは、たとえばストリーミングのような多くの実的な用途にとって非常に高い。このように、HOA表現の圧縮がきわめて望ましい。

20

【 0 0 0 5 】

HOA音場表現の圧縮のためのさまざまな手法が特許文献4、5、6において提案されている。これらの手法は、音場解析を実行し、与えられたHOA表現を方向性成分 (directional component) と残差周囲成分 (residual ambient component) に分解することで共通している。一方では、最終的な圧縮された表現は、いくつかの量子化された信号を含み、該量子化された信号は、いわゆる方向性およびベクトル・ベースの信号と周囲HOA成分 (ambient HOA component) の関連する係数シーケンスとの知覚的符号化から帰結する。他方では、最終的な圧縮された表現は、量子化された信号に関係する追加的なサイド情報を含む。このサイド情報は、HOA表現の、その圧縮されたバージョンからの再構成のために必要である。

30

【 先行技術文献 】

【 特許文献 】

40

【 0 0 0 6 】

【特許文献4】欧州特許出願第13305558.2号。Alexander Krueger and Sven Kordon. Intelligent signal extraction and packing for compression of HOA sound field representations. (Technicolor Internal Reference: PD130015), filed 29. April 2013

【特許文献5】欧州特許出願公開第2743922号。A. Krueger, S. Kordon, and J. Boehm. HOA compression by decomposition into directional and ambient components. (Technicolor Internal Reference: PD120055), December 2012

【特許文献6】欧州特許出願公開第2665208号。Alexander Krueger, Sven Kordon, Johannes Boehm, and Jan-Mark Batke. Method and apparatus for compressing

50

and decompressing a higher order ambisonics signal representation. (Technicolor Internal Reference: PD120015), May 2012

【特許文献7】欧州特許出願公開第2738962号。Alexander Krueger. Method and apparatus for robust sound source direction tracking based on Higher Order Ambisonics. (Technicolor Internal Reference: PD120049), November 2012

【非特許文献】

【0007】

【非特許文献1】Jérôme Daniel. Représentation de champs acoustiques, application à la transmission et à la reproduction de scènes sonores complexes dans un contexte multimédia. PhD thesis, Université Paris 6, 2001 10

【非特許文献2】Jörg Fliege and Ulrike Maier. A two-stage approach for computing cubature formulae for the sphere. Technical report, Fachbereich Mathematik, Universität Dortmund, 1999. ノード数は<http://www.mathematik.uni-dortmund.de/lx/research/projects/fliege/nodes/nodes.html>にある。

【非特許文献3】Sven Kordon and Alexander Krueger. Adaptive value range control for HOA signals. Patent application (Technicolor Internal Reference: PD130016), July 2013

【非特許文献8】Daniel D. Lee and H. Sebastian Seung. Learning the parts of objects by nonnegative matrix factorization. Nature, 401:788-791, 1999 20

【非特許文献9】ISO/IEC JTC 1/SC 29 N. Text of ISO/IEC 23008-3/CD, MPEG-H 3d audio, April 2014

【非特許文献10】Boaz Rafaely. Plane-wave decomposition of the sound field on a sphere by spherical convolution. J. Acoust. Soc. Am., 4(116): 2149-2157, October 2004

【非特許文献11】Earl G. Williams. Fourier Acoustics, volume 93 of Applied Mathematical Sciences. Academic Press, 1999

【発明の概要】

【発明が解決しようとする課題】 30

【0008】

特許文献4、5、6についての量子化された信号の合理的な最小数は8である。よって、これらの方法の一つでのデータ・レートは、個々の知覚的符号化器について32kbit/sのデータ・レートを想定すると、典型的には256kbit/sより低くはない。たとえばモバイル装置へのオーディオ・ストリーミングのようなある種の用途のためには、この全データ・レートは高すぎることもありうる。よって、たとえば128kbit/sのような顕著により低いデータ・レートに対処するHOA圧縮方法に対する需要がある。

【課題を解決するための手段】

【0009】

音場の高次アンビソニックス(HOA)表現の低ビットレート圧縮のための新たな方法および装置が開示される。 40

【0010】

音場のHOA表現のための低ビットレート圧縮方法の一つの主要な側面は、HOA表現を複数の周波数サブバンドに分解し、各周波数サブバンド(すなわちサブバンド)内の係数を、打ち切りされたHOA表現と、いくつかの予測された方向性サブバンド信号に基づく表現との組み合わせによって近似することである。

【0011】

打ち切りされたHOA表現は少数の選択された係数シーケンスを含む。ここで、選択は時間とともに変わることが許容される。たとえば、フレーム毎に新たな選択がなされる。打ち切りされたHOA表現を表わす選択された係数シーケンスは知覚的に符号化され、最終的 50

な圧縮されたHOA表現の一部となる。ある実施形態では、符号化効率を高め、レンダリングにおけるノイズ・マスキング解除 (noise unmasking) の効果を低減するために、それらの選択された係数シーケンスは知覚的符号化の前に脱相関される。それらの選択されたHOA係数シーケンスのうちの所定数に対して空間的変換を適用することによって、部分的脱相関が達成される。圧縮解除のためには、脱相関は、再相関によって反転される。そのような部分的脱相関の大いなる利点は、圧縮解除に際して脱相関を反転させるために追加のサイド情報が必要とされないことである。

【 0 0 1 2 】

近似されたHOA表現の他方の成分は、対応する方向をもついくつかの方向性サブバンド信号によって表現される。これらは、打ち切りされたHOA表現の係数シーケンスからの予測を含むパラメトリック表現によって符号化される。ある実施形態では、それぞれの方向性サブバンド信号が、打ち切りされたHOA表現の係数シーケンスのスケーリングされた和によって予測 (または表現) される。ここで、スケーリングは、一般的には複素数値である。圧縮解除のために方向性サブバンド信号のHOA表現を再合成することができるために、圧縮された表現は、方向の量子化されたバージョンとともに複素数値の予測スケーリング因子の量子化されたバージョンを含む。

【 0 0 1 3 】

ある実施形態では、所与の数の係数シーケンスを有する入力HOA信号のフレームをエンコードする (そしてそれにより圧縮する) 方法であって、各係数シーケンスはインデックスをもつ、方法が、

打ち切りされたHOA表現に含まれるべきアクティブな係数シーケンスのインデックスの集合 $I_{C,ACT}(k)$ を決定する段階と；

低減された数の0でない係数シーケンス (すなわち、入力HOA信号より少ない0でない係数シーケンス、よってより多くの0の係数シーケンス) をもつ打ち切りされたHOA表現 $C_T(k)$ を計算する段階と；

前記入力HOA信号から、候補方向の第一の集合 $M_{DIR}(k)$ を推定する段階と；

前記入力HOA信号を複数の周波数サブバンドに分割する段階であって、それらの周波数サブバンドの係数シーケンス

【 数 1 】

$$\tilde{c}(k-1, k, f_{1,...,F})$$

が得られる、段階と；

各周波数サブバンドについて、方向の第二の集合 $M_{DIR}(k, f_1), \dots, M_{DIR}(k, f_F)$ を推定する段階であって、方向の前記第二の集合の各要素は第一および第二のインデックスをもつインデックスのタプルであり、前記第二のインデックスは現在の周波数サブバンドについてのアクティブな方向のインデックスであり、前記第一のインデックスは該アクティブな方向の軌跡インデックスであり、各アクティブな方向は前記入力HOA信号の候補方向の前記第一の集合 $M_{DIR}(k)$ にも含まれるものであり (すなわち、方向の前記第二の集合におけるアクティブなサブバンド方向は、フルバンド方向の前記第一の集合のサブセット)、

各周波数サブバンドについて、その周波数サブバンドの係数シーケンス

【 数 2 】

$$\tilde{c}(k-1, k, f_{1,...,F})$$

からそれぞれの周波数サブバンドの方向の前記第二の集合 $M_{DIR}(k, f_1), \dots, M_{DIR}(k, f_F)$ に従って、方向性サブバンド信号

【数 3】

$$\tilde{X}(k-1, k, f_1), \dots, \tilde{X}(k-1, k, f_F)$$

を計算する段階と；

各周波数サブバンドについて、それぞれの周波数サブバンドのアクティブな係数シーケンスのインデックスの前記集合 $I_{C,ACT}(k)$ を使ってその周波数サブバンドの係数シーケンス

【数 4】

$$\tilde{C}(k-1, k, f_{1,\dots,F})$$

から方向性サブバンド信号

【数 5】

$$\tilde{X}(k-1, k, f_{1,\dots,F})$$

を予測するよう適応された予測行列 $A(k, f_1), \dots, A(k, f_F)$ を計算する段階と；

候補方向の前記第一の集合 $M_{DIR}(k)$ 、方向の前記第二の集合 $M_{DIR}(k, f_1), \dots, M_{DIR}(k, f_F)$ 、前記予測行列 $A(k, f_1), \dots, A(k, f_F)$ および前記打ち切りされたHOA表現 $C_T(k)$ をエンコードする段階とを含む。

【0014】

方向の前記第二の集合は周波数サブバンドに関係する。候補方向の前記第一の集合はフル周波数帯域に関係する。有利なことに、サブバンド方向の前記第二の集合はフルバンド方向の前記第一の集合の部分集合なので、各周波数サブバンドについて方向の第二の集合を推定する段階において、周波数サブバンドの方向 $M_{DIR}(k, f_1), \dots, M_{DIR}(k, f_F)$ は、フルバンドHOA信号の方向 $M_{DIR}(k)$ のうちでのみ探索すればよい。ある実施形態では、各タブル内での前記第一および第二のインデックスの逐次順は入れ替えられてもよい。すなわち、前記第一のインデックスが現在の周波数サブバンドについてのアクティブな方向のインデックスとなり、前記第二のインデックスが前記アクティブな方向についての軌跡インデックスとなる。

【0015】

完備なHOA信号は、複数の係数シーケンスまたは係数チャネルを含む。これらの係数シーケンスのうちの一つまたは複数の0に設定されているHOA信号は、本稿では打ち切りされたHOA表現と呼ばれる。打ち切りされたHOA表現を計算または生成することは、一般には、0に設定されるまたはされない係数シーケンスの選択を含む。この選択は、さまざまな基準に従って、たとえば、0に設定されない係数シーケンスとして、最大エネルギーを含む係数シーケンスまたは知覚的に最も有意である係数シーケンスを選択することによって、あるいは任意に係数シーケンスを選択することなどによって、できる。HOA信号を周波数サブバンドに分割することは、たとえば直交ミラー・フィルタ(QMF)を含む分解フィルタバンクによって実行されることができる。

【0016】

ある実施形態では、打ち切りされたHOA表現 $C_T(k)$ のエンコードは、打ち切りされたHOAチャネル・シーケンスの部分的な脱相関と、(相関したまたは脱相関された)打ち切りされたHOAチャネル・シーケンス $y_1(k), \dots, y_I(k)$ をトランスポート・チャネルに割り当てるためのチャネル割り当てと、前記トランスポート・チャネルのそれぞれに対して利得制御を実行する段階であって、各トランスポート・チャネルについての利得制御サイド情報 $e_i(k-1)$ 、 $i(k-1)$ が生成される、段階と、知覚的エンコードにおいて、利得制御された打ち切りされたHOAチャネル・シーケンス $z_1(k), \dots, z_I(k)$ をエンコードする段階と、前記利得制御サイド情報 $e_i(k-1)$ 、 $i(k-1)$ 、候補方向の前記第一の集合 $M_{DIR}(k)$ 、方向の前

10

20

30

40

50

記第二の集合 $M_{DIR}(k, f_1)$ 、...、 $M_{DIR}(k, f_F)$ および前記予測行列 $A(k, f_1)$ 、...、 $A(k, f_F)$ をサイド情報源符号化器においてエンコードする段階と、前記知覚的エンコーダおよび前記サイド情報源符号化器の出力を多重化して、エンコードされたHOA信号フレーム

【数 6】

$$\tilde{B}(k-1)$$

を得る段階とを含む。

【0017】

ある実施形態では、コンピュータ可読媒体が、入力HOA信号のフレームをエンコードまたは圧縮するための前記方法をコンピュータに実行させるための実行可能命令を記憶している。

10

【0018】

ある実施形態では、所与の数の係数シーケンスを有する入力HOA信号のフレームをフレームごとにエンコードする（そしてそれにより圧縮する）ための装置であって、各係数シーケンスがインデックスをもつ、装置が、プロセッサと、該プロセッサ上で実行されたときに、入力HOA信号のフレームをエンコードまたは圧縮するための上述した方法の段階を実行するソフトウェア・プログラムのためのメモリとを有する。

【0019】

さらに、ある実施形態では、圧縮されたHOA表現をデコードする（そしてそれにより圧縮解除する）ための方法が、

20

前記圧縮されたHOA表現から、複数の打ち切りされたHOA係数シーケンス

【数 7】

$$\hat{z}_1(k), \dots, \hat{z}_I(k),$$

前記打ち切りされたHOA係数シーケンスのシーケンス・インデックスを示す（または含む）割り当てベクトル $v_{AMB, ASSIGN}(k)$ 、サブバンドに関係した方向情報 $M_{DIR}(k+1, f_1)$ 、...、 $M_{DIR}(k+1, f_F)$ 、複数の予測行列 $A(k+1, f_1)$ 、...、 $A(k+1, f_F)$ および利得制御サイド情報 $e_1(k)$ 、...、 $e_I(k)$ 、 $\hat{z}_1(k)$ を抽出する段階と；

30

前記複数の打ち切りされたHOA係数シーケンス

【数 8】

$$\hat{z}_1(k), \dots, \hat{z}_I(k),$$

前記利得制御サイド情報 $e_1(k)$ 、...、 $e_I(k)$ 、 $\hat{z}_1(k)$ および前記割り当てベクトル $v_{AMB, ASSIGN}(k)$ から打ち切りされたHOA表現

【数 9】

$$\hat{c}_T(k)$$

40

を再構成する段階と；

分解フィルタバンクにおいて、前記再構成された打ち切りされたHOA表現

【数 10】

$$\hat{c}_T(k)$$

を複数F個の周波数サブバンドについての周波数サブバンド表現

50

【数 1 1】

$$\tilde{\mathcal{C}}_T(k, f_1), \dots, \tilde{\mathcal{C}}_T(k, f_F)$$

に分解する段階と；

各周波数サブバンド表現について方向性サブバンド合成ブロックにおいて、前記再構成された打ち切りされたHOA表現の前記それぞれの周波数サブバンド表現

【数 1 2】

$$\tilde{\mathcal{C}}_T(k, f_1), \dots, \tilde{\mathcal{C}}_T(k, f_F)$$

10

と、前記サブバンドに関係した方向情報 $M_{\text{DIR}}(k+1, f_1), \dots, M_{\text{DIR}}(k+1, f_F)$ と、前記予測行列 $A(k+1, f_1), \dots, A(k+1, f_F)$ とから、予測された方向性HOA表現

【数 1 3】

$$\tilde{\mathcal{C}}_D(k, f_1), \dots, \tilde{\mathcal{C}}_D(k, f_F)$$

を合成する段階と；

前記F個の周波数サブバンドのそれぞれについてサブバンド組成ブロックにおいて、係数シーケンス

【数 1 4】

$$\tilde{\mathcal{C}}_n(k, f_j), n = 1, \dots, O$$

をもつデコードされたサブバンドHOA表現

【数 1 5】

$$\tilde{\mathcal{C}}(k, f_1), \dots, \tilde{\mathcal{C}}(k, f_F)$$

30

を組成する段階であって、該係数シーケンスは、該係数シーケンスが前記割り当てベクトル $V_{\text{AMB, ASSIGN}}(k)$ に含まれる（すなわちその要素である）インデックスnをもつ場合には前記打ち切りされたHOA表現

【数 1 6】

$$\tilde{\mathcal{C}}_T(k, f_j)$$

の係数シーケンスから得られ、それ以外の場合には前記方向性サブバンド合成ブロックの一つによって与えられる前記予測された方向性HOA成分

40

【数 1 7】

$$\tilde{\mathcal{C}}_D(k, f_j)$$

の係数シーケンスから得られる、段階と；

合成フィルタバンクにおいて、前記デコードされたサブバンドHOA表現

【数 1 8】

$$\tilde{\mathcal{C}}(k, f_1), \dots, \tilde{\mathcal{C}}(k, f_F)$$

50

を合成して、デコードされたHOA表現

【数 1 9】

$$\hat{c}(k)$$

を得る段階とを含む。

【0 0 2 0】

ある実施形態では、前記抽出する段階は、前記圧縮されたHOA表現を多重分離して、知覚的に符号化された部分とエンコードされたサイド情報部分とを取得することを含む。ある実施形態では、知覚的に符号化された部分は、知覚的にエンコードされた打ち切りされたHOA係数シーケンス

10

【数 2 0】

$$\check{z}_1(k), \dots, \check{z}_I(k)$$

を含み、前記抽出する段階は、知覚的デコードにおいて、知覚的にエンコードされた打ち切りされたHOA係数シーケンス

【数 2 1】

$$\check{z}_1(k), \dots, \check{z}_I(k)$$

20

をデコードして前記打ち切りされたHOA係数シーケンス

【数 2 2】

$$\hat{z}_1(k), \dots, \hat{z}_I(k)$$

を得ることを含む。ある実施形態では、前記抽出する段階は、サイド情報源デコードにおいて、前記エンコードされたサイド情報部分をデコードして、前記サブバンド関係方向 $M_{DIR}(k+1, f_1), \dots, M_{DIR}(k+1, f_F)$ 、予測行列 $A(k+1, f_1), \dots, A(k+1, f_F)$ 、利得制御サイド情報 $e_1(k), \dots, e_I(k)$ および割り当てベクトル $v_{AMB, ASSIGN}(k)$ を得ることを含む。

30

【0 0 2 1】

ある実施形態では、コンピュータ可読媒体が、優勢な方向性信号の方向のデコードのための前記方法をコンピュータに実行させるための実行可能命令を記憶している。

【0 0 2 2】

ある実施形態では、圧縮されたHOA表現をフレームごとにデコードする（そしてそれにより圧縮解除する）ための装置が、プロセッサと、該プロセッサ上で実行されたときに、入力HOA信号のフレームをデコードまたは圧縮解除するための上述した方法の段階を実行するソフトウェア・プログラムのためのメモリとを有する。

【0 0 2 3】

40

ある実施形態では、HOA信号をデコードする装置が、デコードされるべきHOA信号表現についての最大数の方向Dのインデックスを受領するよう構成された第一モジュールと、デコードされるべきHOA信号表現の最大数の方向Dの方向を再構成するよう構成された第二モジュールと、サブバンド毎にアクティブな方向性信号のインデックスを受領するよう構成された第三モジュールと、デコードされるべきHOA信号表現のそれらの再構成された方向Dからサブバンド毎にアクティブな方向性信号を再構成するよう構成された第四モジュールと、諸サブバンドの諸方向性信号を予測するよう構成された第五モジュールとを有し、あるサブバンドの現在フレームにおける方向性信号の前記予測は、そのサブバンドの先行フレームの諸方向性信号を決定することを含み、その方向性信号のインデックスが前記先行フレームにおいて0であり、現在フレームにおいて0でなければ、新たな方向性信号が生成

50

され、その方向性信号のインデックスが前記先行フレームにおいて0でなく、現在フレームにおいて0であれば、前の方向性信号が取り消され、その方向性信号のインデックスが第一の方向から第二の方向に変わる場合には、方向性信号の方向は前記第一の方向から前記第二の方向に動かされる。

【0024】

サブバンドは一般に、複素数値のフィルタバンクから得られる。前記割り当てベクトルの一つの目的は、送信/受信される、よって打ち切りされたHOA表現に含まれる係数シーケンスのシーケンス・インデックスを示し、それによりこれらの係数シーケンスの最終的なHOA信号への割り当てを可能にすることである。換言すれば、割り当てベクトルは、打ち切りされたHOA表現の各係数シーケンスについて、それが最終的なHOA信号におけるどの係数シーケンスに対応するかを示す。たとえば、打ち切りされたHOA表現が四つの係数シーケンスを含み、最終的なHOA信号が九つの係数シーケンスをもつ場合、割り当てベクトルは(原理的に)[1,2,5,7]であってもよい。これにより、打ち切りされたHOA表現の第一、第二、第三および第四の係数シーケンスが実際には最終的なHOA信号における第一、第二、第五および第七の係数シーケンスであることを示す。

【0025】

本発明のさらなる目的、特徴および利点は、添付の図面との関連で参酌されるとき以下の記述および付属の請求項を考慮することから明白となるであろう。

【図面の簡単な説明】

【0026】

本発明の例示的实施形態が、付属の図面を参照して記述される。

【図1】空間的HOAエンコードのアーキテクチャを示す図である。

【図2】方向推定ブロックのアーキテクチャを示す図である。

【図3】知覚的サイド情報源エンコードを示す図である。

【図4】知覚的サイド情報源デコードを示す図である。

【図5】空間的HOAデコードのアーキテクチャを示す図である。

【図6】球面座標系を示す図である。

【図7】方向推定処理ブロックを示す図である。

【図8】打ち切りされたHOA表現の方向、軌跡インデックス集合および係数を示す図である。

【図9】MPEGにおいて使われる通常のオーディオ・エンコードである。

【図10】MPEGにおいて使用できる改善されたオーディオ・エンコードである。

【図11】MPEGにおいて使われる通常のオーディオ・デコードである。

【図12】MPEGにおいて使用できる改善されたオーディオ・デコードである。

【図13】エンコード方法のフローチャートである。

【図14】デコード方法のフローチャートである。

【発明を実施するための形態】

【0027】

音場のHOA表現のための提案される低ビットレート圧縮方法の一つの主要な発想は、もとのHOA表現をフレームごとおよび周波数サブバンドごとに、すなわち各HOAフレームの個々の周波数サブバンド内で、二つの部分の組み合わせによって近似するということである。二つの部分とは、打ち切りされたHOA表現と、いくつかの予測された方向性サブバンド信号に基づく表現である。HOAの基本の概要を下記でさらに与えておく。

【0028】

近似されたHOA表現の第一の部分は、少数の選択された係数シーケンスからなる打ち切りされたHOAバージョンであり、該選択は時間とともに(たとえばフレーム毎に)変わることが許容される。打ち切りされたHOAバージョンを表わすための選択されたそれらの係数シーケンスは、次いで、知覚的に符号化され、最終的な圧縮されたHOA表現の一部となる。符号化効率を高め、レンダリングにおけるノイズ・マスキング解除(noise unmasking)の効果を低減するために、それらの選択された係数シーケンスを、知覚的符号化の前

に脱相関することが有利である。それらの選択されたHOA係数シーケンスのうちの所定数に対して空間的変換、つまり所与の数の仮想スピーカ信号へのレンダリングを適用することによって、部分的脱相関が達成される。かかる部分的脱相関の大いなる利点は、圧縮解除に際して脱相関を反転させるために追加のサイド情報が必要とされないことである。

【0029】

近似されたHOA表現の第二の部分は、対応する方向をもついくつかの方向性サブバンド信号によって表現される。しかしながら、これらは従来は符号化されない。その代わり、これらは前記第一の部分、すなわち打ち切りされたHOA表現の係数シーケンスからの予測によってパラメトリック表現として符号化される。具体的には、各方向性サブバンド信号は、打ち切りされたHOA表現の係数シーケンスのスケーリングされた和によって予測される。ここで、スケーリングは一般的には複素数値である。両方の部分が合わさって、HOA信号の圧縮された表現をなし、よって低ビットレートを達成する。圧縮解除のために方向性サブバンド信号のHOA表現を再合成することができるために、圧縮された表現は、方向の量子化されたバージョンとともに複素数値の予測スケーリング因子の量子化されたバージョンを含む。

【0030】

このコンテキストにおいて特に重要な側面は、方向および複素数値の予測スケーリング因子の計算ならびにいかにしてそれらを効率的に符号化するかである。

【0031】

低ビットレートHOA圧縮

提案される低ビットレートHOA圧縮のために、低ビットレートHOA圧縮器は空間的HOAエンコード部と、知覚的および源エンコード部とに細分されることができる。空間的HOAエンコード部の例示的なアーキテクチャが図1に示されており、知覚的および源エンコード部の例示的なアーキテクチャが図3に示されている。空間的HOAエンコーダ10は、1個の信号を含む第一の圧縮されたHOA表現を、そのHOA表現をどのように生成するかを記述するサイド情報とともに提供する。知覚的およびサイド情報源符号化器30では、これら1個の信号が知覚的符号化器31において知覚的にエンコードされ、サイド情報はサイド情報源符号化器32における源エンコードにかけられる。サイド情報源符号化器32は、符号化されたサイド情報

【数23】

\tilde{r}

を提供する。次いで、知覚的符号化器31およびサイド情報源符号化器32によって提供された二つの符号化された表現は、マルチプレクサ33において多重化されて、低ビットレートの圧縮されたHOAデータ・ストリーム

【数24】

\tilde{b}

が得られる。

【0032】

空間的HOAエンコード

図1に示される空間的HOAエンコーダは、フレームごとの処理を実行する。フレームは0個の時間連続的なHOA係数シーケンスの諸部分として定義される。たとえば、エンコードされるべき入力HOA表現のk番目のフレームC(k)は、時間連続的なHOA係数シーケンスのベクトルc(t) (式(46)参照) に関して次のように定義される。

【0033】

10

20

30

40

【数 2 5】

$$\mathbf{C}(k) := [\mathbf{c}((kL+1)T_s) \quad \mathbf{c}((kL+2)T_s) \quad \dots \quad \mathbf{c}((k+1)L T_s)] \in \mathbb{R}^{O \times L} \quad (1)$$

ここで、 k はフレーム・インデックスを表わし、 L はフレーム長（サンプル単位）を表わし、 $O = (N+1)^2$ はHOA係数シーケンスの数を表わし、 T_s はサンプリング周期を表わす。

【0 0 3 4】

打ち切りされたHOA表現の計算

図1に示されるように、打ち切りされたHOA表現の計算における第一段階は、もとのHOAフレーム $\mathbf{C}(k)$ から、打ち切りされたバージョン $\mathbf{C}_T(k)$ を計算する（11）ことを含む。このコンテキストにおける打ち切りとは、入力HOA表現の O 個の係数シーケンスのうち1個の特定の係数シーケンスを選択し、他のすべての係数シーケンスを0に設定することを意味する。係数シーケンスの選択のためのさまざまな解決策が、特許文献4、5、6から知られている。たとえば、最大のパワーまたは人間の知覚に関して最高の有意性をもつものの選択である。選択された係数シーケンスは、打ち切りされたHOAバージョンを表わす。選択された係数シーケンスのインデックスを含むデータ・セット $I_{C,ACT}(k)$ が生成される。次いで、のちにさらに述べるように、打ち切りされたHOAバージョン $\mathbf{C}_T(k)$ は部分的に脱相関され（12）、部分的に脱相関された打ち切りされたHOAバージョン $\mathbf{C}_I(k)$ はチャンネル割り当て13を受ける。ここで、選ばれた係数シーケンスが、利用可能な1個のトランスポート・チャンネルに割り当てられる。のちにさらに述べるように、これらの係数シーケンスは、次いで知覚的にエンコードされ（30）、最終的に、圧縮された表現の一部になる。チャンネル割り当て後に知覚的エンコードのためのなめらかな信号を得るために、 k 番目のフレームでは選択されているが $k+1$ 番目のフレームでは選択されていない係数シーケンスが判別される。あるフレームでは選択されるが次のフレームでは選択されなくなる係数シーケンスは、フェードアウトされる。それらのインデックスは、データ・セット $I_{C,ACT}(k)$ の部分集合であるデータ・セット $I_{C,ACT,OUT}(k)$ に含まれる。同様に、 k 番目のフレームでは選択されるが $k-1$ 番目のフレームでは選択されていなかった係数シーケンスは、フェードインされる。それらのインデックスは、やはりデータ・セット $I_{C,ACT}(k)$ の部分集合であるデータ・セット $I_{C,ACT,IN}(k)$ に含まれる。フェードのためには、窓関数 $w_{OA}(l)$ 、 $l = 1, \dots, 2L$ （のちにたとえば式(39)で導入されるようなもの）が使われてもよい。

【0 0 3 5】

ひっくりめると、打ち切りされたバージョン $\mathbf{C}_T(k)$ のHOAフレーム k が O 個の個々の係数シーケンス・フレームの L 個のサンプルから

【数 2 6】

$$\mathbf{C}_T(k) = \begin{bmatrix} c_{T,1}(k,1) & \dots & c_{T,1}(k,L) \\ c_{T,2}(k,1) & \dots & c_{T,2}(k,L) \\ \vdots & \ddots & \vdots \\ c_{T,O}(k,1) & \dots & c_{T,O}(k,L) \end{bmatrix} \quad (2)$$

によって構成されるとすると、打ち切りは、係数シーケンス・インデックス $n = 1, \dots, O$ およびサンプル・インデックス $l = 1, \dots, L$ について、

10

20

30

40

【数 2 7】

$$c_{T,n}(k) = \begin{cases} c_n(k,l) \cdot w_{OA}(l) & \text{if } n \in J_{CACTIN}(k) \\ c_n(k,l) \cdot w_{OA}(L+1) & \text{if } n \in J_{CACTOUT}(k) \\ c_n(k,l) & \text{if } n \in J_{CACT}(k) \setminus (J_{CACTIN}(k) \cup J_{CACTOUT}(k)) \\ 0 & \text{else} \end{cases} \quad (3)$$

によって表現されることができる。

10

【0036】

係数シーケンスの選択のための基準についてはいくつかの可能性がある。たとえば、一つの有利な解決策は、信号パワーの大半を表わす係数シーケンスを選択することである。もう一つの有利な解決策は、人間の知覚に関して最も有意な係数シーケンスを選択することである。後者の場合、有意性は、たとえば、異なる仕方で打ち切りされた表現を仮想スピーカー信号にレンダリングし、それらの信号とものHOA表現に対応する仮想スピーカー信号との間の誤差を決定し、最後に、音マスキング効果を考慮しつつ誤差の有意性を解釈することによって、決定されてもよい。

【0037】

集合 $I_{CACT}(k)$ 内のインデックスを選択するための合理的な戦略は、ある実施形態では、常に最初の 0_{MIN} 個のインデックス $1, \dots, 0_{MIN}$ を選択するというものである。ここで、 $0_{MIN} = (N_{MIN} + 1)^2 - 1$ であり、 N_{MIN} は、打ち切りされたHOA表現の所与の最小フル次数を表わす。次いで、上述した基準のうちの一つに従って集合 $\{0_{MIN} + 1, \dots, 0_{MAX}\}$ の残り $l - 0_{MIN}$ 個のインデックスを選択する。ここで、 $0_{MAX} = (N_{MAX} + 1)^2 - 0$ であり、 N_{MAX} は、選択のために考慮されるHOA係数シーケンスの最大次数を表わす。 0_{MAX} はサンプル当たりの転送可能な係数の最大数であり、これは係数の総数 0 以下であることを注意しておく。この戦略によれば、打ち切り処理ブロック11は、いわゆる割り当てベクトル

20

【数 2 8】

$$v_A(k) \in \mathbb{N}^{l-0_{MIN}}$$

30

をも提供する。該割り当てベクトルの要素 $v_{A,i}(k)$ 、 $i = 1, \dots, l - 0_{MIN}$ は

【数 2 9】

$$v_{A,i}(k) = n \quad (4)$$

に従って設定される。ここで、 n (ただし $n \geq 0_{MIN} + 1$) は、のちに i 番目のトランスポート信号 $y_i(k)$ に割り当てられる、 $C(k)$ の追加的に選択されたHOA係数シーケンスのHOA係数シーケンス・インデックスを表わす。 $y_i(k)$ の定義は下記の式(10)において与えられる。よって、 $C_T(k)$ の最初の 0_{MIN} 行は、デフォルトにより、HOA係数シーケンス $1, \dots, 0_{MIN}$ を含み、 $C_T(k)$ の続く $0 - 0_{MIN}$ 個の (または $0 = 0_{MAX}$ なら $0_{MAX} - 0_{MIN}$ 個の) 行のうち、割り当てベクトル $v_A(k)$ にインデックスが格納されている、フレームごとに変化するHOA係数シーケンスを含む $l - 0_{MIN}$ 個の行がある。最後に、 $C_T(k)$ の残りの行は 0 を含む。結果として、後述するように、利用可能な l 個のトランスポート信号のうち最初の (または式(10)のように最後の) 0_{MIN} 個はデフォルトによりHOA係数シーケンス $1, \dots, 0_{MIN}$ に割り当てられ、残りの $l - 0_{MIN}$ 個のトランスポート信号は、割り当てベクトル $v_A(k)$ にインデックスが格納されている、フレームごとに変化するHOA係数シーケンスに割り当てられる。

40

【0038】

50

部分的脱相関

第二段階では、その後の知覚的エンコードの効率を高めるとともに、レンダリングにおける選択されたHOA係数シーケンスの行列処理後に生じる符号化ノイズ・マスキング解除を避けるために、選択されたHOA係数シーケンスの部分的脱相関12が実行される。例示的な部分的脱相関12は、最初の O_{MIN} 個の選択されたHOA係数シーケンスに空間的変換を適用することによって達成される。これは、 O_{MIN} 個の仮想スピーカ信号へのレンダリングを意味する。それぞれの仮想スピーカ位置は、図6に示される球面座標系によって表わされる。ここで、各位置は単位球上にある、すなわち動径1をもつと想定される。よって、これらの位置は等価に方向 $\mathbf{j} = (\theta_j, \phi_j)$ によって表わされることが出来る。ここで、 $1 \leq j \leq O_{MIN}$ であり、 θ_j および ϕ_j はそれぞれ傾斜角および方位角を表わす（球面座標系の定義については下記参照）。これらの方向は、単位球上にできるだけ一様に分散されるべきである（個別的な方法の計算についてはたとえば非特許文献2を参照）。HOAは一般に N_{MIN} に依存して方向を定義するので、本稿で \mathbf{j} が書かれているところでは実際には $\mathbf{j}^{(N_{MIN})}$ が意図されることを注意しておく。

10

【0039】

以下では、すべての仮想スピーカ信号のフレームは
【数30】

$$\mathbf{W}(k) = \begin{bmatrix} w_1(k) \\ w_2(k) \\ \vdots \\ w_{O_{MIN}}(k) \end{bmatrix} \quad (5)$$

20

によって表わされる。ここで、 $w_j(k)$ は j 番目の仮想スピーカ信号の k 番目のフレームを表わす。さらに、 \mathbf{U}_{MIN} は仮想スピーカ方向 \mathbf{j} に関するモード行列を表わす。ここで、 $1 \leq j \leq O_{MIN}$ である。モード行列は次式によって定義される。

【0040】

【数31】

30

$$\mathbf{\Psi}_{MIN} = [\mathbf{S}_{MIN,1} \quad \dots \quad \mathbf{S}_{MIN,O_{MIN}}] \in \mathbb{R}^{O_{MIN} \times O_{MIN}} \quad (6)$$

ここで、

【数32】

$$\mathbf{S}_{MIN,i} := [S_0^0(\boldsymbol{\Omega}_i) \quad S_1^{-1}(\boldsymbol{\Omega}_i) \quad S_1^0(\boldsymbol{\Omega}_i) \quad S_1^1(\boldsymbol{\Omega}_i) \quad \dots \quad S_N^{N-1}(\boldsymbol{\Omega}_i) \quad S_N^N(\boldsymbol{\Omega}_i)] \in \mathbb{R}^{O_{MIN}} \quad (7)$$

は仮想方向 \mathbf{j} に関するモード・ベクトルを示す。その要素 $S_n^m(\cdot)$ のそれぞれは、のちに定義する（式(48)参照）実数値の球面調和関数を表わす。この記法を使うと、レンダリング・プロセスは次の行列乗算によって定式化できる。

40

【0041】

【数 3 3】

$$W(k) = (\Psi_{\text{MIN}})^{-1} \cdot \begin{bmatrix} c_1(k) \\ \vdots \\ c_{O_{\text{MIN}}}(k) \end{bmatrix} \quad (8)$$

よって、部分的脱相関 1 2 の出力である中間表現 $C_l(k)$ の信号は
【数 3 4】

10

$$c_{l,n}(k) = \begin{cases} w_n(k) & \text{if } 1 \leq n \leq O_{\text{MIN}} \\ c_{T,n}(k) & O_{\text{MIN}} + 1 \leq n \leq O \end{cases} \quad (9)$$

によって与えられる。

【0 0 4 2】

チャンネル割り当て

中間表現 $C_l(k)$ のフレームを計算したのち、その個々の信号 $c_{l,n}(k)$ (ここで $n = 1, \dots, O_{\text{ACT}}(k)$) は、利用可能な l 個のチャンネルに割り当てられて (1 3)、知覚的エンコードのため
20 のトランスポート信号 $y_i(k)$ 、 $i = 1, \dots, l$ を与える。割り当て 1 3 の一つの目的は、相続くフレームの間で選択が変化する場合に起こりうる、知覚的にエンコードされるべき信号の不連続を避けることである。割り当ては、

【数 3 5】

$$y_i(k) = \begin{cases} c_{L,n,i}(k) & \text{if } 1 \leq i \leq l - O_{\text{MIN}} \\ c_{L,i-(l-O_{\text{MIN}})}(k) & \text{if } l - O_{\text{MIN}} < i \leq l \end{cases} \quad (10)$$

30

によって表わすことができる。

【0 0 4 3】

利得制御

トランスポート信号 $y_i(k)$ のそれぞれは、利得制御ユニット 1 4 によって最終的に処理される。ここで、信号利得は、知覚的エンコードに好適な値範囲を達成するためになめらかに修正される。利得修正は、相続くブロックの間での深刻な利得変化を避けるために、一種の先読みを必要とし、よって一フレームの遅延を導入する。各トランスポート信号 $y_i(k)$ について、利得制御ユニット 1 4 は、遅延されたフレーム $y_i(k-1)$ 、 $i = 1, \dots, l$ を受領または生成する。利得制御後の修正された信号フレームは、 $z_i(k-1)$ 、 $i = 1, \dots, l$ と表わされる。さらに、何らかの修正がなされた場合に空間的デコードにおいて反転できるように、利得制御サイド情報が提供される。利得制御サイド情報は、指数 $e_i(k-1)$ および例外フラグ $f_i(k-1)$ 、 $i = 1, \dots, l$ を含む。利得制御のより詳細な説明については、たとえば非特許文献 9、C.5.2.5 節または非特許文献 3 で得られる。このように、打ち切りされた HOA バージョン 1 9 は、利得制御された信号フレーム $z_i(k-1)$ および利得制御サイド情報 $e_i(k-1)$ 、 $f_i(k-1)$ 、 $i = 1, \dots, l$ を含む。

40

【0 0 4 4】

分解フィルタバンク

上述したように、近似された HOA 表現は二つの部分から構成される。すなわち、打ち切りされた HOA バージョン 1 9 と、打ち切りされた HOA 表現の係数シーケンスから予測される、対応する方向をもつ方向性のサブバンド信号によって表わされる成分とである。よって

50

、第二の部分のパラメトリック表現を計算するために、もとのHOA表現の個々の係数シーケンスの各フレーム $c_n(k)$ 、 $n = 1, \dots, N$ がまず個々のサブバンド信号

【数 3 6】

$$\tilde{c}_n(k, f_1), \dots, \tilde{c}_n(k, f_F)$$

のフレームに分解される。これは、一つまたは複数の分解フィルタバンク 1 5 においてなされる。各サブバンド f_j 、 $j = 1, \dots, F$ について、個々のHOA係数シーケンスのサブバンド信号のフレームがサブバンドHOA表現

【数 3 7】

$$\tilde{c}(k, f_j) = \begin{bmatrix} \tilde{c}_1(k, f_j) \\ \tilde{c}_2(k, f_j) \\ \vdots \\ \tilde{c}_O(k, f_j) \end{bmatrix} \quad \text{for } j = 1, \dots, F \quad (11)$$

にまとめられてもよい。分解フィルタバンク 1 5 は、方向性サブバンド信号計算のために、このサブバンドHOA表現を方向推定処理ブロック 1 6 および一つまたは複数の計算ブロック 1 7 に与える。

【0 0 4 5】

原理的には、任意の型のフィルタ（すなわち、任意の複素数値のフィルタバンク、たとえばQMF、FFT）が分解フィルタバンク 1 5 において使用されうる。分解フィルタバンクと対応する合成フィルタバンクの相続く適用が遅延された恒等関数を与えることは要求されない。そのような恒等関数を与える性質は完璧再構成属性として知られるものである。HOA係数シーケンス $c_n(k)$ とは対照的に、それらのサブバンド表現

【数 3 8】

$$\tilde{c}_n(k, f_j)$$

は一般には複素数値である。さらに、これらのサブバンド信号

【数 3 9】

$$\tilde{c}_n(k, f_j)$$

は一般に、もとの時間領域信号に比べて時間的に間引きされている。結果として、フレーム

【数 4 0】

$$\tilde{c}_n(k, f_j)$$

内のサンプル数は通例、時間領域信号フレーム $c_n(k)$ 内のサンプル数（これは L ）より顕著に少ない。

【0 0 4 6】

ある実施形態では、処理を人間の聴覚系の属性によりよく適合させるために、二つ以上

10

20

30

40

50

のサブバンド信号がサブバンド信号グループに組み合わせられる。各グループの帯域幅は、そのサブバンド信号の数によって、たとえばよく知られたバーク尺度に適合されることができる。すなわち、特に高周波では、二つ以上のグループが一つに組み合わせられることができる。この場合、各サブバンド・グループが一つのHOA係数シーケンス【数41】

$$\tilde{c}(k, f_j)$$

からなり、抽出されるパラメータの数は単一のサブバンドについてと同じであることを注意しておく。ある実施形態では、グループ化は、分解フィルタバンク・ブロック15に組み込まれてもよい一つまたは複数のサブバンド信号グループ化ユニット（明示的には図示していない）において実行される。

【0047】

方向推定

方向推定処理ブロック16は、入力HOA表現を解析し、各周波数サブバンド f_j 、 $j=1, \dots, F$ について、音場に主要な寄与を加えるサブバンド一般平面波関数の方向の集合 $M_{DIR}(k, f_j)$ を計算する。このコンテキストにおいて、「主要な寄与」という用語は、信号パワーが他の方向から入射するサブバンド一般平面波の信号パワーより高いことを指しうる。また、人間の知覚に関して高い有意性を指すこともある。サブバンド・グループ化が使われる場合には、 $M_{DIR}(k, f_j)$ の計算のために、単一のサブバンドの代わりに、サブバンド・グループが使われることもできることを注意しておく。

【0048】

圧縮解除の間、相続くフレームの間での推定される方向および予測係数の変化のため、予測される方向性サブバンド信号におけるアーチファクトが生じることがある。そのようなアーチファクトを避けるためには、エンコードの間の方向推定および方向性サブバンド信号の予測は、連結された長フレームに対して実行される。連結された長フレームは、現在フレームおよびその先行フレームからなる。圧縮解除のためには、これらの長フレームに対して推定された量は、予測された方向性のサブバンド信号を用いた重複加算処理を実行するために使われる。

【0049】

方向推定のための素直なアプローチは、各サブバンドを別個に扱うことであろう。方向探索のために、ある実施形態では、たとえば特許文献7において提案される技法が適用されてもよい。このアプローチは、個々の各サブバンドについて、方向推定値のなめらかな時間的軌跡を提供し、急峻な方向変化または開始を捕捉できる。しかしながら、この既知のアプローチには二つの欠点がある。第一に、各サブバンドにおける独立した方向推定は、フルバンドの一般平面波（たとえばある方向からの過渡的なドラム・ビート）が存在する場合に、個々のサブバンド方向における推定誤差が、足し合わせても単一の方向からの所望されるフルバンド・バージョンにならない種々の方向からのサブバンド一般平面波につながりうるという、望まれない効果につながることもありうる。特に、ある諸方向からの過渡信号がぼかされる。

【0050】

第二に、低ビットレート圧縮を得るという意図を考えると、サイド情報から帰結する全ビットレートを念頭に置く必要がある。以下では、そのような素朴なアプローチのためのビットレートがかなり高くなることを例によって示す。例示的に、サブバンドの数 F は10であると想定され、各サブバンドについての方向の数（これは各集合 $M_{DIR}(k, f_j)$ における要素の数に対応する）は4であると想定される。さらに、非特許文献9で提案されているように、各サブバンドについて、 $Q=900$ 個の潜在的な方向候補の格子上で探索を実行するとする。これは、単一の方向の単純な符号化のために $\log_2(Q) = 10$ ビットを要求する。毎秒約50フレームのフレーム・レートを想定すると、結果として得られる全体的なデー

10

20

30

40

50

タ・レートは、方向の符号化された表現だけのために、
 (10ビット/方向)・(4方向/帯域)・(10帯域/フレーム)・(50フレーム/秒)=20kbit/s

となる。たとえ毎秒25フレームのフレーム・レートを想定するとしても、結果として得られるデータ・レート10kbit/sはまだかなり高い。

【0051】

改善として、ある実施形態では、方向推定ブロック20において、方向推定のための以下の方法が使われる。一般的な発想は図2に示されている。第一段階では、フルバンド方向推定ブロック21が、Q個の試験方向 $T_{EST,q}$ 、 $q=1, \dots, Q$ からなる方向格子上で、連結された長フレーム

【数42】

$$\bar{C}(k-1;k) = [C(k-1) \quad C(k)] \quad (12)$$

を使って、予備的なフルバンド方向推定または探索を実行する。ここで、 $C(k)$ および $C(k-1)$ はフルバンドのもののHOA表現の現在および前の入力フレームである。この方向探索は、 $D(k)$ 個の方向候補 $CAND_d(k)$ 、 $d=1, \dots, D(k)$ を提供し、これらが集合 $M_{DIR}(k)$ に含まれる。すなわち、

$$M_{DIR}(k) = \{ CAND_1(k), \dots, CAND_{D(k)}(k) \} \quad (13)$$

となる。

【0052】

フレーム毎の方向候補の最大数についての典型的な値は $D=16$ である。方向推定は、たとえば特許文献7において提案されている方法によって達成されることができる。発想は、入力HOA表現の方向性パワー分布から得られる情報を、方向のベジアン推定についての単純な源動きモデルと組み合わせるというものである。

【0053】

第二段階では、方向探索が、サブバンド（またはサブバンド・グループ）毎に、サブバンド方向推定ブロック22によって個々の各サブバンドについて実行される。しかしながら、サブバンドについてのこの方向探索は、Q個の試験方向からなる初期のフルの方向格子を考える必要はなく、各サブバンドについて $D(k)$ 個の方向のみを含む候補集合 $M_{DIR}(k)$ のみを考えればよい。 f_j 番目のサブバンド、 $j=1, \dots, F$ についての方向の数 $D_{SB}(k, f_j)$ は、 D_{SB} より小さくなく、 D_{SB} は典型的には D より顕著に小さく、たとえば $D_{SB}=4$ である。フルバンド方向探索と同様に、サブバンドに関係した方向探索も、前のフレームと現在フレームからなる、サブバンド信号の長い連結されたフレーム

【数43】

$$\bar{C}(k-1;k;f_j) = [\tilde{C}(k-1,f_j) \quad \tilde{C}(k,f_j)] \quad j=1, \dots, F \quad (14)$$

に対して実行される。原理的には、フルバンドに関係した方向探索についてと同じベジアン推定方法が、サブバンド関係の方向探索のために適用されてもよい。

【0054】

特定の音源の方向は時間とともに変化してもよい（だが変化する必要はない）。特定の音源の方向の時間的シーケンスは本稿では「軌跡」と呼ばれる。各サブバンド関係の方向または軌跡は曖昧さのないインデックスを得て、それにより異なる軌跡の混同が防止され、連続的な方向性サブバンド信号が与えられる。このことは、方向性サブバンド信号の後述する予測のために重要である。これは特に、さらに下記で定義される相続く予測係数行列 $A(k, f_j)$ の間の時間的依存性を活用することを許容する。したがって、 f_j 番目のサブバンドのための方向推定は、タプルの集合 $M_{DIR}(k, f_j)$ を提供する。各タプルは、個々の（ア

10

20

30

40

50

クティブな) 方向軌跡を同定するインデックス $d \in I_{DIR}(k, f_j) = \{1, \dots, D_{SB}\}$ と、それぞれの推定される方向 $\Omega_{SB,d}(k, f_j)$ とからなる。すなわち、
【数 4 4】

$$\mathcal{M}_{DIR}(k, f_j) = \left\{ \left(d, \Omega_{SB,d}(k, f_j) \right) \mid d \in I_{DIR}(k, f_j) \right\}. \quad (15)$$

サブバンド方向探索は上述したように現在フレームの方向候補 $CAND_{d,d}(k)$ 、 $d = 1, \dots, D(k)$ の間でのみ実行されるので、定義により、集合 $\{(\Omega_{SB,d}(k, f_j) \mid d \in I_{DIR}(k, f_j))\}$ は各 $j = 1, \dots, F$ についての $M_{DIR}(k)$ の部分集合である。各インデックスは Q 個ではなく $D(k)$ 個の候補方向のうちの一つの方向を定義し、 $D(k) = Q$ であるので、これは、方向に関するサイド情報のより効率的な符号化を許容する。インデックス d は、軌跡を生成するためにその後のフレームにおいて方向を追跡するために使われる。図 2 に示され、上記したように、方向推定処理ブロック 16 はある実施形態では、フルバンド方向推定ブロック 21 と、各サブバンドもしくはサブバンド・グループについてのサブバンド方向推定ブロック 22 とを有する方向推定ブロック 20 を備える。これはさらに、図 7 に示されるように上述した長フレームを方向推定ブロック 20 に提供する長フレーム生成ブロック 23 を有していてもよい。長フレーム生成ブロック 23 は、たとえば一つまたは複数のメモリを使って、それぞれ L サンプルの長さをもつ二つの相続く入力フレームから長フレームを生成する。長フレームは、本稿では L によって、また $k-1$ および k の二つのインデックスをもつことによって示される。他の実施形態では、長フレーム生成ブロック 23 は、図 1 に示されるエンコーダ内の別個のブロックであってもよく、あるいは他のブロックに組み込まれてもよい。

【0055】

方向性サブバンド信号の計算

図 1 に戻ると、分解フィルタバンク 15 によって提供されるサブバンド HOA 表現フレーム $\tilde{\mathbf{c}}(k, f_j)$ 、 $j = 1, \dots, F$

【数 4 5】

$$\tilde{\mathbf{c}}(k, f_j), j = 1, \dots, F$$

は、一つまたは複数の方向性サブバンド信号計算ブロック 17 にも入力される。方向性サブバンド信号計算ブロック 17 では、すべての D_{SB} 個の潜在的な方向性サブバンド信号の長フレーム

【数 4 6】

$$\tilde{\mathbf{x}}_d(k-1; k; f_j), d = 1, \dots, D_{SB}$$

が

【数 4 7】

$$\tilde{\mathbf{X}}(k-1; k; f_j) = \begin{bmatrix} \tilde{\mathbf{x}}_1(k-1; k; f_j) \\ \tilde{\mathbf{x}}_2(k-1; k; f_j) \\ \vdots \\ \tilde{\mathbf{x}}_{D_{SB}}(k-1; k; f_j) \end{bmatrix} \in \mathbb{C}^{D_{SB} \times 2L}. \quad (16)$$

10

20

30

40

50

のように行列に配列される。

【 0 0 5 6 】

さらに、非アクティブな方向性サブバンド信号のフレーム、すなわちインデックスdが集合 $I_{DIR}(k, f_j)$ 内に含まれない長い信号フレーム

【 数 4 8 】

$$\tilde{\mathbf{x}}_d(k-1; k; f_j)$$

は0に設定される。

【 0 0 5 7 】

残りの長い信号フレーム

【 数 4 9 】

$$\tilde{\mathbf{x}}_d(k-1; k; f_j)$$

すなわち、インデックスd $I_{DIR}(k, f_j)$ をもつものは、行列

【 数 5 0 】

$$\tilde{\mathbf{x}}_{ACT}(k-1; k; f_j) \in \mathbb{C}^{D_{SB}(k, f_j) \times 2L}$$

20

内に集められる。それに含まれるアクティブな方向性サブバンド信号を計算するための一つの可能性は、そのHOA表現ともの入力サブバンドHOA表現との間の誤差を最小にすることである。解は

【 数 5 1 】

$$\tilde{\mathbf{x}}_{ACT}(k-1; k; f_j) = \left(\boldsymbol{\Psi}_{SB}(k, f_j) \right)^+ \tilde{\mathbf{c}}(k-1; k; f_j) \quad (17)$$

30

によって与えられる。ここで、 $(\cdot)^+$ はムーア・ペンローズの擬似逆行列を表わし、

【 数 5 2 】

$$\boldsymbol{\Psi}_{SB}(k, f_j) \in \mathbb{R}^{O \times D_{SB}(k, f_j)}$$

は、集合 $\{(\boldsymbol{\Psi}_{SB, d}(k, f_j) \mid d \in I_{DIR}(k, f_j))\}$ 内の方向推定値に関するモード行列を表わす。サブバンド・グループの場合、方向性サブバンド信号

【 数 5 3 】

$$\tilde{\mathbf{x}}_{ACT}(k-1; k; f_j)$$

40

の集合は、グループのすべてのHOA表現

【 数 5 4 】

$$\tilde{\mathbf{c}}(k-1; k; f_j)$$

に一つの列 $(\boldsymbol{\Psi}_{SB}(k, f_j))^+$ を乗算することから計算されることを注意しておく。長フレ

50

ームは、上記したものと同様の一つまたは複数のさらなる長フレーム生成ブロックによって生成されることができることを注意しておく。同様に、長フレームは、長フレーム分解ブロックにおいて、正規長さのフレームに分解されることができる。ある実施形態では、方向性サブバンドの計算のための諸ブロック 17 は、それらの出力において、長フレーム【数 5 5】

$$\bar{\tilde{x}}_{ACT}(k-1; k; f_j), j = 1, \dots, F$$

を、方向性サブバンド予測ブロック 18 に提供する。

10

【0058】

方向性サブバンド信号の予測

上述したように、近似的なHOA表現は、アクティブな方向性サブバンド信号によって部分的に表現される。だが、これは従来は符号化されないものである。従来と違って、現在記載される実施形態では、符号化された表現の伝送のための全データ・レートを低く保つために、パラメトリック表現が使われる。パラメトリック表現では、それぞれのアクティブな方向性サブバンド信号

【数 5 6】

$$\tilde{x}_d(k-1; k; f_j)$$

20

すなわちインデックス $d = I_{DIR}(k, f_j)$ をもつものが、打ち切りされたサブバンドHOA表現【数 5 7】

$$\tilde{c}_n(k-1, f_j) \text{ and } \tilde{c}_n(k, f_j)$$

の係数シーケンスの重み付けされた和によって予測される。ここで、 $n = I_{C, ACT}(k-1)$ であり、重みは一般に複素数値である。

30

【0059】

よって、

【数 5 8】

$$\bar{\tilde{x}}_p(k-1; k; f_j)$$

が

【数 5 9】

40

$$\bar{\tilde{x}}(k-1; k; f_j)$$

の予測されたバージョンを表わすとする、予測は【数 6 0】

$$\bar{\tilde{x}}_p(k-1; k; f_j) = A(k, f_j) \bar{\tilde{c}}_T(k-1; k; f_j), \quad (18)$$

として行列乗算によって表わされる。ここで、

50

【数 6 1】

$$A(k, f_j) \in \mathbb{C}^{O \times D_{SB}}$$

はサブバンド f_j についてのすべての重み付け因子（あるいは等価だが予測係数）をもつ行列である。予測行列 $A(k, f_j)$ の計算は、一つまたは複数の方向性サブバンド予測ブロック 18 において実行される。ある実施形態では、図 1 に示されるように、サブバンド毎に一つの方向性サブバンド予測ブロック 18 が使われる。別の実施形態では、複数またはすべてのサブバンドのために単一の方向性サブバンド予測ブロック 18 が使われる。サブバンド・グループの場合には、各グループについて一つの行列 $A(k, f_j)$ が計算されるが、それはグループのそれぞれの HOA 表現

10

【数 6 2】

$$\bar{\tilde{c}}_T(k-1; k; f_j)$$

を乗算され、グループ毎に一組の行列

【数 6 3】

20

$$\bar{\tilde{x}}_p(k-1; k; f_j)$$

を生成する。構築法により、インデックス $d_{DIR}(k, f_j)$ をもつもの以外の $A(k, f_j)$ のすべての行は 0 であることを注意しておく。つまり、アクティブな方向性サブバンド信号のみが予測される。さらに、インデックス $n_{C, ACT}(k-1)$ をもつもの以外の $A(k, f_j)$ のすべての列も 0 である。つまり、予測のためには、伝送され、HOA 圧縮解除の間に予測のために利用可能である HOA 係数シーケンスのみが考慮される。

【0 0 6 0】

予測行列 $A(k, f_j)$ の計算のために、以下の諸側面が考慮される必要がある。

30

【0 0 6 1】

第一に、もとの打ち切りされたサブバンド HOA 表現

【数 6 4】

$$\tilde{c}_T(k, f_j)$$

は一般に、HOA 圧縮解除において利用可能ではない。その代わり、その知覚的にデコードされたバージョン

【数 6 5】

40

$$\hat{\tilde{c}}_T(k, f_j)$$

が利用可能であり、方向性サブバンド信号の予測のために使われる。低ビットレートでは、典型的なオーディオ・コーデック（たとえば AAC または USAC）は、スペクトルの低周波数および中程度の周波数は通常通り符号化されるが、高周波数内容（たとえば 5kHz 以降）は高周波数包絡についての追加のサイド情報を使って低周波数および中程度の周波数から複製されるスペクトル帯域複製（SBR）を使う。そのため、知覚的デコード後の打ち切りされた HOA 成分

50

【数 6 6】

$$\tilde{\mathbf{c}}_{\mathbf{T}}(k, f_j)$$

の再構成されたサブバンド係数シーケンスの大きさは、もとの

【数 6 7】

$$\tilde{\mathbf{c}}_{\mathbf{T}}(k, f_j)$$

10

のものに似る。しかしながら、これは位相については成り立たない。よって、高周波数サブバンドについては、複素数値の予測係数を使うことによる予測のために何らかの位相関係を活用することは意味がない。その代わり、実数値の予測係数のみを使うほうが合理的である。特に、 f_j 番目のサブバンドがSBRの開始周波数を含むようインデックス j_{SBR} を定義して、予測係数の型を次のように設定することが有利である。

【数 6 8】

$$\mathbf{A}(k, f_j) \in \begin{cases} \mathbb{C}^{O \times D_{\text{SB}}} & \text{for } 1 \leq j < j_{\text{SBR}} \\ \mathbb{R}^{O \times D_{\text{SB}}} & \text{for } j_{\text{SBR}} \leq j \leq F \end{cases} \quad (19)$$

20

換言すれば、ある実施形態では、より低いサブバンドについての予測係数は複素数値であり、一方、より高いサブバンドについての予測係数は実数値である。

【0 0 6 2】

第二に、ある実施形態では、行列 $\mathbf{A}(k, f_j)$ の計算の戦略は、その型に適合される。特に、SBRによって影響されない低周波数サブバンド f_j 、 $1 \leq j < j_{\text{SBR}}$ については、

【数 6 9】

$$\bar{\mathbf{x}}(k-1; k; f_j)$$

30

とその予測されたバージョン

【数 7 0】

$$\bar{\mathbf{x}}_{\mathbf{p}}(k-1; k; f_j)$$

との間の誤差のユークリッド・ノルムを最小化することによって、 $\mathbf{A}(k, f_j)$ の0でない要素を決定することが可能である。知覚的符号化器 3 1 は、 j_{SBR} (図示せず) を定義し、提供する。このようにして、関係する信号の位相関係が予測のために明示的に活用される。サブバンド・グループについては、グループのすべての方向性信号に対する予測誤差のユークリッド・ノルムが最小化されるべきである(すなわち、最小二乗予測誤差)。SBRによって影響される高周波数サブバンド f_j 、 $j_{\text{SBR}} \leq j \leq F$ については、打ち切りされたHOA成分

40

【数 7 1】

$$\tilde{\mathbf{c}}_{\mathbf{T}}(k, f_j)$$

の再構成されたサブバンド係数シーケンスの位相は、もとのサブバンド係数シーケンスの位相にいささかも似ていると想定することができないため、上述した基準は合理的ではな

50

い。

【 0 0 6 3 】

この場合、一つの解決策は、位相を無視し、その代わりに予測のためには信号パワーのみに集中することである。予測係数の決定のための合理的な基準は次の誤差を最小化することである。

【 数 7 2 】

$$\left| \bar{\mathbf{x}}(k-1; k; f_j) \right|^2 - \left| \mathbf{A}(k, f_j) \right|^2 \left| \bar{\mathbf{c}}_T(k-1; k; f_j) \right|^2 \quad (20)$$

10

ここで、演算 $|\cdot|^2$ は要素ごとに行列に適用されると想定される。換言すれば、予測係数は、打ち切りされたHOA成分のすべての重み付けされたサブバンドまたはサブバンド・グループの係数シーケンスのパワーの和が、方向性サブバンド信号のパワーを最もよく近似するように、選ばれる。この場合、この最適化問題を解いて予測行列 $\mathbf{A}(k, f_j)$ 、 $j = 1, \dots, F$ の予測係数を得るために、非負行列因子分解 (NMF: Nonnegative Matrix Factorization) 技法 (たとえば非特許文献 8 参照) を使うことができる。これらの行列は、次いで、知覚的および源エンコード段 30 に提供される。

【 0 0 6 4 】

知覚的および源エンコード

上記の空間的HOA符号化後、 $k-1$ 番目のフレームについての結果として得られる利得適応されたトランスポート信号 $z_i(k-1)$ 、 $i = 1, \dots, I$ は符号化されて、それらの符号化された表現

20

【 数 7 3 】

$$\tilde{z}_i(k-1)$$

が得られる。これは、図 3 に示される知覚的および源エンコード段 30 における知覚的符号化器 31 によって実行される。さらに、集合 $M_{DIR}(k)$ 、 $M_{DIR}(k, f_j)$ 、 $f = 1, \dots, F$ 、予測係数行列 $\mathbf{A}(k, f_j) \in \mathbb{C}^{O \times DSB}$ 、 $j = 1, \dots, F$ 、利得制御パラメータ $e_i(k-1)$ および $\tilde{z}_i(k-1)$ 、 $i = 1, \dots, I$ および割り当てベクトル $v_A(k-1)$ に含まれる情報が、効率的な記憶または伝送のために冗長性を除去するために源エンコードにかけられる。これは、サイド情報源符号化器 32 によって実行される。結果として得られる符号化された表現

30

【 数 7 4 】

$$\tilde{\mathbf{r}}(k-1)$$

はマルチプレクサ 33 において、符号化されたトランスポート信号表現

【 数 7 5 】

40

$$\tilde{z}_i(k-1), i = 1, \dots, I$$

と一緒に多重化されて、最終的な符号化されたフレーム

【 数 7 6 】

$$\tilde{\mathbf{B}}(k-1)$$

を生成する。

50

【 0 0 6 5 】

原理的には、利得制御パラメータの源符号化および割り当ては非特許文献 9 と同様に実行できるので、本稿は方向および予測パラメータの符号化のみに焦点を当て、それについて以下で詳述する。

【 0 0 6 6 】

方向の符号化

個々のサブバンド方向の符号化のために、上記に基づく非有意性削減を活用して、選ばれるべき個々のサブバンド方向を制約することができる。前述したように、これら個々のサブバンド方向は、あらゆる可能な試験方向 TEST_q , $q = 1, \dots, Q$ から選ばれるのではなく、フルバンド HOA 表現の各フレームについて決定される少数の候補から選ばれる。例示的に、サブバンド方向の源符号化のための可能な仕方は、下記のアルゴリズム 1 においてまとめられる。

【 0 0 6 7 】

【 表 1 】

アルゴリズム 1 サブバンド方向の符号化

NoOfGlobalDirs(k) ($\lceil \log_2(D) \rceil$ ビットで符号化される)

{Fill GlobalDirGridIndices(k) (NoOfGlobalDirs(k) 個の要素をもつ配列。各要素は $\lceil \log_2(D) \rceil$ ビットで符号化される)}

for $d = 1$ to NoOfGlobalDirs(k) do

GlobalDirGridIndices(k) [d] = q such that $\Omega_{FB,d}(k) = \Omega_{\text{TEST},q}$

end for

// グローバル方向

for $j = 1$ to F do

{Fill bSubBandDirIsActive(k, f_j) (D_{SB} 個の要素をもつビット配列)}

for $d = 1$ to D_{SB} do

if $d \in \mathcal{I}_{\text{DIR}}(k, f_j)$ then

bSubBandDirIsActive(k, f_j) [d] = 1

else

bSubBandDirIsActive(k, f_j) [d] = 0

end if

end for

// サブバンド毎の
// アクティブな方向

{Fill RelDirIndices(k, f_j)

($D_{SB}(k, f_j)$ 個の要素をもつ配列。各要素は $\lceil \log_2(\text{NoOfGlobalDirs}(k)) \rceil$ ビットで符号化される)}

for $d = 1$ to D_{SB} do

$d_1 = 1$

if bSubBandDirIsActive(k, f_j) [d] = 1 then

RelDirIndices(k, f_j) [d_1] = i such that $\Omega_{SB,d}(k, f_j) = \Omega_{FB,i}(k)$

$d_1 = d_1 + 1$

end if

end for

end for

// フルバンドの
// 方向インデックス

アルゴリズム 1 の第一段階では、サブバンド方向として実際に生起するすべてのフルバンド方向候補の集合 $\mathcal{M}_{FB}(k)$ が決定される。すなわち、

【 数 7 7 】

$$\mathcal{M}_{FB}(k) := \left\{ \Omega_{\text{CAND},d}(k) \mid \exists j \in \{1, \dots, F\} \text{ and } d \in \mathcal{I}_{\text{DIR}}(k, f_j) \right. \\ \left. \text{such that } \Omega_{\text{CAND},d}(k) = \Omega_{SB,d}(k, f_j) \right\} \quad (21)$$

NoOfGlobalDirs(k) によって表わされるこの集合の要素数は、方向の符号化された表現の第一の部分である。 $\mathcal{M}_{FB}(k)$ は定義により $\mathcal{M}_{\text{DIR}}(k)$ の部分集合なので、NoOfGlobalDirs(k) は

$\lceil \log_2(D) \rceil$ ビットで符号化できる。さらなる記述を明確にするために、集合 $\mathcal{M}_{FB}(k)$ 内の方向は $\Omega_{FB,d}(k)$, $d = 1, \dots, \text{NoOfGlobalDirs}(k)$ によって表わされる。すなわち、

10

20

30

40

50

【数 7 8】

$$\mathcal{M}_{FB}(k) := \{\Omega_{FB,d}(k) | d = 1, \dots, \text{NoOfGlobalDirs}(k)\} \quad (22)$$

第二段階では、集合 $\mathcal{M}_{FB}(k)$ 内の方向が、ここで格子 (grid) と称される可能な試験方向 $\text{TEST}_{q,q}$ のインデックス $q = 1, \dots, Q$ によって符号化される。各方向 $\Omega_{FB,d}(k)$ 、 $d = 1, \dots, \text{NoOfGlobalDirs}(k)$ について、それぞれの格子インデックスは、 $\log_2(Q)$ ビットのサイズをもつ配列要素 $\text{GlobalDirGridIndices}(k)[d]$ において符号化される。すべての符号化されたフルバンド方向を表わす全配列 $\text{GlobalDirGridIndices}(k)$ は $\text{NoOfGlobalDirs}(k)$ 個の要素からなる。

10

【0 0 6 8】

第三段階では、各サブバンドまたはサブバンド・グループ f_j 、 $j = 1, \dots, F$ について、 d 番目の方向性サブバンド信号 ($d = 1, \dots, D_{SB}$) がアクティブであるか否か、すなわち $d \text{ I}_{DIR}(k, f_j)$ であるかどうかの情報が、配列要素 $\text{bSubBandDirIsActive}(k, f_j)[d]$ において符号化される。全配列 $\text{bSubBandDirIsActive}(k, f_j)$ は D_{SB} 個の要素からなる。 $d \text{ I}_{DIR}(k, f_j)$ であれば、それぞれのサブバンド方向 $\text{SB}_{d,d}(k, f_j)$ はそれぞれのフルバンド方向 $\text{FB}_{i,i}(k)$ のインデックス i によって、 $D_{SB}(k, f_j)$ 個の要素からなる配列 $\text{RelDirIndices}(k, f_j)$ に符号化される。

【0 0 6 9】

20

この方向エンコード方法の効率を示すために、上記の例に基づく方向の符号化された表現についての最大データ・レートが計算される： $F = 10$ サブバンド、 $D_{SB}(k, f_j) = D_{SB} =$ サブバンド毎に4方向、 $Q = 900$ 個の潜在的な試験方向および毎秒25フレームのフレーム・レートが想定される。通常の符号化方法では、必要とされるデータ・レートは10kbit/sであった。ある実施形態に基づくこの改善された符号化方法では、フルバンド方向の数が $\text{NoOfGlobalDirs}(k) = D = 8$ であるとすれば、 $\text{GlobalDirGridIndices}(k)$ を符号化するためにフレーム当たり $D \cdot \log_2(Q) = 80$ ビットが、 $\text{bSubBandDirIsActive}(k, f_j)$ を符号化するために $D_{SB} \cdot F = 40$ ビットが、 $\text{RelDirIndices}(k, f_j)$ を符号化するために $D_{SB} \cdot F \cdot \log_2(\text{NoOfGlobalDirs}(k)) = 120$ ビットが必要とされる。その結果、 $(240 \text{ ビット / フレーム}) \cdot (25 \text{ フレーム / 秒}) = 6 \text{ kbit/s}$ のデータ・レートとなり、これは10kbit/sより顕著に小さい。フルバンド方向のより多くの数 $\text{NoOfGlobalDirs}(k) = D = 16$ についてでさえも、たった7kbit/sのデータ・レートで十分である。

30

【0 0 7 0】

予測係数行列の符号化

予測係数行列の符号化のためには、方向軌跡、よって方向性サブバンド信号のなめらかさのために相続くフレームの予測係数の間には高い相関があるという事実を活用できる。さらに、各予測係数行列 $A(k, f_j)$ についてフレーム毎に $(D_{SB}(k, f_j) \cdot M_{C,ACT}(k-1))$ 個の潜在的な0でない要素の比較的多くの数がある。ここで、 $M_{C,ACT}(k-1)$ は集合 $I_{C,ACT}(k-1)$ 内の要素数を表わす。サブバンド・グループが使われなければ、全部で、フレーム当たり符号化されるべき F 個の行列がある。サブバンド・グループが使われる場合には、フレーム当たり符号化されるべき行列は相応して F 個より少なくなる。

40

【0 0 7 1】

ある実施形態では、各予測係数のためのビット数を低く保つために、各複素数値の予測係数はその絶対値とその偏角によって表現され、次いで偏角および絶対値が相続くフレームの間で異なる仕方で、行列 $A(k, f_j)$ のそれぞれの特定の要素について独立に、符号化される。絶対値が区間 $[0, 1]$ 内にあると想定される場合、絶対値の差は区間 $[-1, 1]$ 内にある。複素数の偏角の差は区間 $[-\pi, \pi]$ 内にあると想定されてもよい。絶対値および偏角の差の両方の量子化のために、それぞれの区間は等しいサイズのたとえば 2^{N_Q} 個の部分区間に細分されることができる。その場合、それぞれの絶対値および偏角の差について、素直な符号化だと N_Q ビットを必要とする。さらに、相続くフレームの予測係数の間の上述した

50

相関のため、個々の差の生起確率はきわめて非一様に分布していることが実験的に見出されている。特に、絶対値および偏角における小さな差は、より大きな差よりも有意に高頻度で生起する。よって、たとえばハフマン符号化のような、符号化されるべき個々の値の事前確率に基づく符号化方法が、予測係数毎の平均ビット数を有意に減らすために活用できる。換言すれば、予測行列 $A(k, f_j)$ の実部および虚部の代わりに、予測行列 $A(k, f_j)$ における値の絶対値および位相を異なる仕方でエンコードすることが通例は有利である。しかしながら、実部および虚部の使用が受け入れ可能な状況も生じるかもしれない。

【 0 0 7 2 】

ある実施形態では、非差分的に符号化された行列係数を含む特別なアクセス・フレームが、ある間隔（用途固有、たとえば毎秒一回）において送られる。これは、デコーダがこれらの特殊なアクセス・フレームから差分デコードを再開することを許容し、よってデコードのためのランダムな入場を可能にする。

10

【 0 0 7 3 】

以下では、上記のように構築された低ビットレート圧縮されたHOA表現の圧縮解除を記述する。圧縮解除もフレーム毎に機能する。

【 0 0 7 4 】

原理的には、ある実施形態に基づく低ビットレートHOAデコーダは、上記の低ビットレートHOAエンコーダ・コンポーネントの、逆順に配列された対応物を有する。特に、低ビットレートHOAデコーダは、図4に描かれるような知覚的および源デコード部と、図6に示されるような空間的HOAデコード部とに細分できる。

20

【 0 0 7 5 】

知覚的および源デコード

図4は、ある実施形態における、知覚的およびサイド情報源デコード40を示している。知覚的およびサイド情報源デコード40において、低ビットレート圧縮されたHOAビットストリーム

【 数 7 9 】

\tilde{B}

はまず多重分離されて（41）、1個の信号の知覚的に符号化された表現

30

【 数 8 0 】

$\tilde{z}_i, i = 1, \dots, I$

およびどのようにしてHOA表現を生成するかを記述する符号化されたサイド情報

【 数 8 1 】

\tilde{r}

を生じる。次いで、1個の信号の知覚的デコードおよびサイド情報のデコードが実行される。

40

【 0 0 7 6 】

知覚的デコーダ42は1個の信号

【 数 8 2 】

$\tilde{z}_i(k), i = 1, \dots, I$

を知覚的にデコードされた信号

【数 8 3】

$$\hat{\mathbf{z}}_i(k), i = 1, \dots, I$$

にデコードする。

【0 0 7 7】

サイド情報源デコーダ 4 3 は、符号化されたサイド情報

【数 8 4】

 $\tilde{\mathbf{r}}$

10

をタプル集合 $M_{\text{DIR}}(k+1, f_j)$ 、 $j = 1, \dots, F$ 、各サブバンドまたはサブバンド・グループ f_j ($j = 1, \dots, F$) についての予測係数行列 $A(k+1, f_j)$ 、利得補正指数 $e_i(k)$ および利得補正例外フラグ $i(k)$ ならびに割り当てベクトル $v_{\text{AMB, ASSIGN}}(k)$ にデコードする。

【0 0 7 8】

アルゴリズム 2 は、符号化されたサイド情報

【数 8 5】

 $\tilde{\mathbf{r}}$

20

からどのようにしてタプル集合 $M_{\text{DIR}}(k, f_j)$ 、 $j = 1, \dots, F$ を生成するかを例示的にまとめている。サブバンド方向のデコードは下記で詳細に記述される。

【0 0 7 9】

【表 2】

アルゴリズム 2 サブバンド方向のデコード

Read NoOfGlobalDirs (k) ($\lceil \log_2(D) \rceil$ ビットで符号化される)

{Read GlobalDirGridIndices (k) (NoOfGlobalDirs (k) 個の要素をもつ配列。各要素は $\lceil \log_2(Q) \rceil$ ビットで符号化される)}

{Compute $M_{\text{FB}}(k)$ }

for $d = 1$ to NoOfGlobalDirs (k) do

$\Omega_{\text{FB},d}(k) = \Omega_{\text{FBST,GlobalDirGridIndices}(k)[d]}$

end for

30

for $j = 1$ to F do

{Read bSubBandDirIsActive (k, f_j) (D_{SB} 個の要素をもつビット配列)}

{Compute $D_{\text{SB}}(k, f_j)$ }

$D_{\text{SB}}(k, f_j) = 0$

for $d = 1$ to $D_{\text{SB}}(k, f_j)$ do

if bSubBandDirIsActive (k, f_j) [d] = 1 then

$D_{\text{SB}}(k, f_j) = D_{\text{SB}}(k, f_j) + 1$

end if

end for

40

{Read ReIDirIndices (k, f_j) ($D_{\text{SB}}(k, f_j)$ 個の要素をもつ配列。各要素は $\lceil \log_2(\text{NoOfGlobalDirs}(k)) \rceil$ ビットで符号化される)}

{Compute $M_{\text{DIR}}(k, f_j)$ }

for $d = 1$ to $D_{\text{SB}}(k, f_j)$ do

$d_1 = 1$

if bSubBandDirIsActive (k, f_j) [d] = 1 then

$\Omega_{\text{SB},d}(k, f_j) = \Omega_{\text{FB,ReIDirIndices}(k,f_j)[d_1]}(k)$

$M_{\text{DIR}}(k, f_j) = M_{\text{DIR}}(k, f_j) \cup \{d, \Omega_{\text{SB},d}(k, f_j)\}$

$d_1 = d_1 + 1$

end if

end for

end for

第一に、符号化されたサイド情報

50

【数 8 6】

 \tilde{f}

からフルバンド方向の数NoOfGlobalDirs(k)が抽出される。上記のように、これらはサブバンド方向としても使われる。それは $\log_2(D)$ ビットで符号化される。

【0080】

第二段階では、NoOfGlobalDirs(k)個の要素からなる配列GlobalDirGridIndices(k)が抽出される。各要素は $\log_2(Q)$ ビットによって符号化される。この配列は、フルバンド方向 $_{FB,d}(k)$ 、 $d = 1, \dots, \text{NoOfGlobalDirs}(k)$ を表わす格子インデックスを含み、

$$_{FB,d}(k) = \text{TEST_GlobalDirGridIndices}(k)[d] \quad (23)$$

となる。

【0081】

次いで、各サブバンドまたはサブバンド・グループ f_j 、 $j = 1, \dots, F$ について、 D_{SB} 個の要素からなる配列bSubBandDirIsActive(k, f_j)が抽出される。ここで、 d 番目の要素bSubBandDirIsActive(k, f_j)[d]は、 d 番目のサブバンド方向がアクティブであるか否かを示す。さらに、アクティブなサブバンド方向の総数 $D_{SB}(k, f_j)$ が計算される。

【0082】

最後に、各サブバンドまたはサブバンド・グループ f_j 、 $j = 1, \dots, F$ について、タプルの集合 $M_{DIR}(k, f_j)$ が計算される。これは、個々の（アクティブな）サブバンド方向軌跡を同定するインデックス $d \in \{1, D_{SB}\}$ およびそれぞれの推定される方向 $_{SB}(k, f_j)$ からなる。

【0083】

次に、各サブバンドまたはサブバンド・グループ f_j 、 $j = 1, \dots, F$ について、予測係数行列 $A(k+1, f_j)$ が、符号化されたフレーム

【数 8 7】

 $\tilde{B}(k)$

から再構成される。ある実施形態では、再構成は、サブバンドまたはサブバンド・グループ f_j 毎に以下のステップを含む：第一に、各行列係数の偏角および絶対値の差がエントロピー復号によって得られる。次いで、エントロピー復号された偏角および絶対値の差が、それらの符号化のために使われたビット数 N_Q に従って実際の値範囲に再スケーリングされる。最後に、現在の予測係数行列 $A(k+1, f_j)$ が、再構成された偏角および絶対値の差を、最新の係数行列 $A(k, f_j)$ 、すなわち直前フレームの係数行列の係数に加えることによって、構築される。

【0084】

このように、現在の行列 $A(k+1, f_j)$ のデコードのためには直前の行列 $A(k, f_j)$ が知られている必要がある。ある実施形態では、ランダム・アクセスを可能にするために、特別なアクセス・フレームがある間隔で受領される。該特別なアクセス・フレームは、これらのフレームから差分デコードを再開するために、非差分的に符号化された行列係数を含む。

【0085】

知覚的およびサイド情報源デコーダ40は、知覚的にデコードされた信号

【数 8 8】

$$\hat{\mathbf{z}}_i(k), i = 1, \dots, I$$

と、タプル集合 $M_{DIR}(k+1, f_j)$ 、 $j = 1, \dots, F$ と、予測係数行列 $A(k+1, f_j)$ と、利得補正指数 $e_i(k)$ と、利得補正例外フラグ $_{i}(k)$ と、割り当てベクトル $v_{AMB, ASSIGN}(k)$ とをその後の空間的HOAデコーダ50に出力する。

10

20

30

40

50

【 0 0 8 6 】

空間的HOAデコード

図 5 は、ある実施形態における例示的な空間的HOAデコード 5 0 を示している。空間的HOAデコード 5 0 は、I 個の信号

【数 8 9】

$$\hat{\mathbf{z}}_i(k), i = 1, \dots, I$$

およびサイド情報デコード 4 3 によって与えられる上記のサイド情報から、再構成されたHOA表現を生成する。空間的HOAデコード 5 0 内の個々の処理ユニットについて下記で詳細に述べる。

10

【 0 0 8 7 】

逆利得制御

空間的HOAデコード 5 0 において、知覚的デコードされた信号

【数 9 0】

$$\hat{\mathbf{z}}_i(k), i = 1, \dots, I$$

は、関連付けられた利得補正指数 $e_i(k)$ および利得補正例外フラグ $e_{ex,i}(k)$ と一緒に、まず一つまたは複数の逆利得制御処理ブロック 5 1 に入力される。逆利得制御処理ブロックは、利得補正された信号フレーム

20

【数 9 1】

$$\hat{\mathbf{y}}_i(k), i = 1, \dots, I$$

を与える。ある実施形態では、I 個の信号

【数 9 2】

$$\hat{\mathbf{z}}_i(k)$$

30

のそれぞれは、図 5 と同様に別個の逆利得制御処理ブロック 5 1 に入力され、i 番目の逆利得制御処理ブロックが利得補正された信号フレーム

【数 9 3】

$$\hat{\mathbf{y}}_i(k)$$

を与える。逆利得制御のより詳細な記述は、たとえば非特許文献 9 の節 11.4.2.1 から知られている。

【 0 0 8 8 】

打ち切りされたHOA再構成

打ち切りされたHOA再構成ブロック 5 2 では、I 個の利得補正された信号フレーム

40

【数 9 4】

$$\hat{\mathbf{y}}_i(k), i = 1, \dots, I$$

が、割り当てベクトル $V_{\text{AMB, ASSIGN}}(k)$ によって与えられる情報に従ってHOA係数シーケンス行列に再分配（すなわち、再割り当て）される。それにより、打ち切りされたHOA表現

【数 9 5】

$$\hat{\mathbf{c}}_T(k)$$

が再構成される。割り当てベクトル $\mathbf{v}_{\text{AMB, ASSIGN}}(k)$ は、各伝送チャネルについて、もとの HOA成分のどの係数シーケンスを含んでいるかを示す I 個の成分を有する。さらに、割り当てベクトルの要素は、 k 番目のフレームについてのすべての受領された係数シーケンスのインデックス（もとのHOA成分を指すインデックス）の集合 $I_{\text{C, ACT}}(k)$

$$I_{\text{C, ACT}}(k) = \{v_{\text{AMB, ASSIGN}, i}(k) \mid i = 1, \dots, I\} \quad (24)$$

をなす。

10

【0 0 8 9】

打ち切りされたHOA表現

【数 9 6】

$$\hat{\mathbf{c}}_T(k)$$

の再構成は以下のステップを含む。

【0 0 9 0】

第一に、デコードされた中間表現

【数 9 7】

20

$$\hat{\mathbf{c}}_i(k) = \begin{bmatrix} \hat{c}_{i,1}(k) \\ \vdots \\ \hat{c}_{i,o}(k) \end{bmatrix} \quad (25)$$

の個々の成分

【数 9 8】

$$\hat{c}_{i,n}(k), n = 1, \dots, O$$

30

は、割り当てベクトル内の情報に依存して、0に設定されるか、利得補正された信号フレーム

【数 9 9】

$$\hat{\mathbf{y}}_i(k)$$

の対応する成分によって置き換えられる。すなわち、

【数 1 0 0】

40

$$\hat{c}_{i,n}(k) = \begin{cases} \hat{\mathbf{y}}_i(k) & \exists i \in \{1, \dots, I\} \text{ such that } v_{\text{AMB, ASSIGN}, i}(k) = n \text{ の場合} \\ 0 & \text{それ以外の場合} \end{cases} \quad (26)$$

つまり、上述したように、割り当てベクトルの i 番目の要素（式(26)では n ）は、 i 番目の係数

【数 1 0 1】

$$\hat{y}_i(k)$$

が、デコードされた中間表現行列（式(25)）のn番目の行における

【数 1 0 2】

$$\hat{c}_{l,n}(k)$$

を置換することを示す。

10

【0 0 9 1】

第二に、

【数 1 0 3】

$$\hat{c}_1(k)$$

内の最初の 0_{MIN} 個の信号の再相関が、逆空間的変換を適用することによって実行され、フレーム

【数 1 0 4】

20

$$\hat{c}_{T,\text{MIN}}(k) = \Psi_{\text{MIN}} \begin{bmatrix} \hat{c}_{1,1}(k) \\ \hat{c}_{1,2}(k) \\ \vdots \\ \hat{c}_{1,0_{\text{MIN}}}(k) \end{bmatrix} \quad (27)$$

を与える。ここで、モード行列 Ψ_{MIN} は式(6)において定義されたとおりである。モード行列は、各 0_{MIN} または N_{MIN} についてあらかじめ定義されている所与の諸方向に依存し、よってエンコーダおよびデコーダの両方において独立して構築されることができる。 0_{MIN} （または N_{MIN} ）も慣習上あらかじめ定義される。

30

【0 0 9 2】

最後に、再構成された打ち切りされたHOA表現

【数 1 0 5】

$$\hat{c}_T(k)$$

が再相関された信号

【数 1 0 6】

40

$$\hat{c}_{T,\text{MIN}}(k)$$

および中間表現の信号

【数 1 0 7】

$$\hat{c}_{l,n}(k), n = 0_{\text{MIN}} + 1, \dots, 0$$

50

から

【数 1 0 8】

$$\hat{\mathbf{c}}_T(k) = \begin{bmatrix} \hat{\mathbf{c}}_{T,\text{MIN}}(k) \\ \hat{\mathbf{c}}_{L,\text{MIN}+1}(k) \\ \vdots \\ \hat{\mathbf{c}}_{L,0}(k) \end{bmatrix} \in \mathbb{R}^{O \times L} \quad (28)$$

に従って組成される。

10

【 0 0 9 3】

分解フィルタバンク

予測された方向性サブバンド信号によって表わされる第二のHOA成分をさらに計算するために、圧縮解除された打ち切りされたHOA表現

【数 1 0 9】

$$\hat{\mathbf{c}}_T(k)$$

の個々の係数シーケンスnの各フレーム

20

【数 1 1 0】

$$\hat{\mathbf{c}}_{T,n}(k), \quad n = 1, \dots, O$$

がまず、一つまたは複数の分解フィルタバンク 5 3 において、個々のサブバンド信号のフレーム

【数 1 1 1】

$$\hat{\tilde{\mathbf{c}}}_{T,n}(k, f_j), \quad j = 1, \dots, F$$

30

に分解される。各サブバンド f_j 、 $j = 1, \dots, F$ について、個々のHOA係数シーケンスのサブバンド信号のフレームは、サブバンドHOA表現

【数 1 1 2】

$$\hat{\tilde{\mathbf{c}}}_T(k, f_j) = \begin{bmatrix} \hat{\tilde{\mathbf{c}}}_{T,1}(k, f_j) \\ \hat{\tilde{\mathbf{c}}}_{T,2}(k, f_j) \\ \vdots \\ \hat{\tilde{\mathbf{c}}}_{T,O}(k, f_j) \end{bmatrix} \quad \text{for } j = 1, \dots, F \quad (29)$$

40

にまとめられてもよい。

【 0 0 9 4】

HOA空間的デコード段において適用される前記一つまたは複数の分解フィルタバンク 5 3 は、HOA空間的エンコード段における一つまたは複数の分解フィルタバンク 1 5 と同じであり、サブバンド・グループについては、HOA空間的エンコード段からのグループ化が適用される。このように、ある実施形態では、グループ化情報がエンコードされた信号に含められる。グループ化情報についてのさらなる詳細は後述する。

【 0 0 9 5】

50

ある実施形態では、HOA圧縮段における打ち切りされたHOA表現の計算のために最大次数 N_{MAX} が考慮され（上記の式(4)付近を参照）、HOA圧縮器および圧縮解除器の分解フィルタバンク 1 5、5 3 の適用は、インデックス $n = 1, \dots, 0_{MAX}$ をもつHOA係数シーケンス【数 1 1 3】

$$\hat{c}_{T,\pi}(k)$$

のみに制約される。インデックス $n = 0_{MAX} + 1, \dots, 0$ をもつサブバンド信号フレーム【数 1 1 4】

10

$$\hat{c}_{T,\pi}(k, f_j)$$

は0に設定されることができる。

【0 0 9 6】

方向性サブバンドHOA表現の合成

各サブバンドまたはサブバンド・グループについて、方向性サブバンドもしくはサブバンド・グループHOA表現

【数 1 1 5】

20

$$\tilde{c}_D(k, f_j), j = 1, \dots, F$$

が、一つまたは複数の方向性サブバンド合成ブロック 5 4 において合成される。ある実施形態では、相続くフレームの間での方向および予測係数の変化に起因するアーチファクトを避けるために、方向性サブバンドHOA表現の計算は、重複加算の概念に基づく。よって、ある実施形態では、 f_j 番目のサブバンド、 $j = 1, \dots, F$ に関係したアクティブな方向性サブバンド信号のHOA表現

【数 1 1 6】

30

$$\tilde{c}_D(k, f_j)$$

は、フェードアウトされる成分とフェードインされる成分との和として計算される：

【数 1 1 7】

$$\tilde{c}_D(k, f_j) = \tilde{c}_{D,OUT}(k, f_j) + \tilde{c}_{D,IN}(k, f_j). \quad (30)$$

40

第一段階では、二つの個々の成分を計算するために、フレーム $k_1 \in \{k, k+1\}$ についての予測係数行列 $A(k_1, f_j)$ と k 番目のフレームについての打ち切りされたサブバンドHOA表現

【数 1 1 8】

$$\tilde{c}_T(k, f_j)$$

とに関係したすべての方向性サブバンド信号

【数 1 1 9】

$$\hat{\tilde{\mathbf{x}}}_1(k_1; k; f_j)$$

の瞬時フレームが、

【数 1 2 0】

$$\hat{\tilde{\mathbf{x}}}_1(k_1; k; f_j) = \mathbf{A}(k_1, f_j) \hat{\tilde{\mathbf{c}}}_T(k, f_j) \quad \text{for } k_1 \in \{k, k+1\} \quad (31)$$

10

によって計算される。

【0 0 9 7】

サブバンド・グループについては、各グループについてのHOA表現

【数 1 2 1】

$$\hat{\tilde{\mathbf{c}}}_T(k, f_j)$$

が固定した行列 $\mathbf{A}(k_1, f_j)$ を乗算されて、そのグループのサブバンド信号

20

【数 1 2 2】

$$\hat{\tilde{\mathbf{x}}}_1(k_1; k; f_j)$$

を生成する。

【0 0 9 8】

第二段階では、方向 $\mathbf{s}_{B,d}(k, f_j)$ に関する方向性サブバンド信号

【数 1 2 3】

30

$$\hat{\tilde{\mathbf{x}}}_{1,d}(k_1; k; f_j)$$

の瞬時サブバンドHOA表現

【数 1 2 4】

$$\hat{\tilde{\mathbf{c}}}_{D,d}^{(d)}(k_1; k; f_j), \quad d \in \mathcal{M}_{\text{DIR}}(k, f_j), \quad j = 1, \dots, F$$

が

【数 1 2 5】

40

$$\hat{\tilde{\mathbf{c}}}_{D,I}^{(d)}(k_1; k; f_j) = \boldsymbol{\psi}(\mathbf{n}_{\text{SB},d}(k, f_j)) \hat{\tilde{\mathbf{x}}}_{1,d}(k_1; k; f_j) \quad (32)$$

として得られる。ここで、 $(\mathbf{n}_{\text{SB},d}(k, f_j))^\top \mathbf{R}^0$ は、方向 $\mathbf{s}_{B,d}(k, f_j)$ に関する（式(7)におけるモード・ベクトルのような）モード・ベクトルを表わす。サブバンド・グループについては、式(32)はグループのすべての信号について実行され、行列 $(\mathbf{n}_{\text{SB},d}(k, f_j))$ は各グループについて固定されている。

【0 0 9 9】

行列

50

【数 1 2 6】

$$\hat{\tilde{c}}_{D,OUT}(k, f_j), \hat{\tilde{c}}_{D,IN}(k, f_j), \text{ および } \hat{\tilde{c}}_{D,I}^{(d)}(k_1; k; f_j)$$

がそのサンプルから

【数 1 2 7】

$$\hat{\tilde{c}}_{D,OUT}(k, f_j) = \begin{bmatrix} \hat{\tilde{c}}_{D,OUT,1}(k, f_j; 1) & \dots & \hat{\tilde{c}}_{D,OUT,1}(k, f_j; L) \\ \vdots & \ddots & \vdots \\ \hat{\tilde{c}}_{D,OUT,O}(k, f_j; 1) & \dots & \hat{\tilde{c}}_{D,OUT,O}(k, f_j; L) \end{bmatrix} \in \mathbb{R}^{O \times L} \quad (33) \quad 10$$

$$\hat{\tilde{c}}_{D,IN}(k, f_j) = \begin{bmatrix} \hat{\tilde{c}}_{D,IN,1}(k, f_j; 1) & \dots & \hat{\tilde{c}}_{D,IN,1}(k, f_j; L) \\ \vdots & \ddots & \vdots \\ \hat{\tilde{c}}_{D,IN,O}(k, f_j; 1) & \dots & \hat{\tilde{c}}_{D,IN,O}(k, f_j; L) \end{bmatrix} \in \mathbb{R}^{O \times L} \quad (34)$$

$$\hat{\tilde{c}}_{D,I}^{(d)}(k_1; k; f_j) = \begin{bmatrix} \hat{\tilde{c}}_{D,I,1}^{(d)}(k-1; k; f_j; 1) & \dots & \hat{\tilde{c}}_{D,I,1}^{(d)}(k-1; k; f_j; L) \\ \vdots & \ddots & \vdots \\ \hat{\tilde{c}}_{D,I,O}^{(d)}(k-1; k; f_j; 1) & \dots & \hat{\tilde{c}}_{D,I,O}^{(d)}(k-1; k; f_j; L) \end{bmatrix} \in \mathbb{R}^{O \times L} \quad (35) \quad 20$$

によって構成されたとすると、アクティブな方向性サブバンド信号のHOA表現のフェードアウトされる成分およびフェードインされる成分のサンプル値は最終的に

【数 1 2 8】

$$\hat{\tilde{c}}_{D,OUT,n}(k, f_j; l) = \sum_{d \in \mathcal{D}_{DIR}(k, f_j)} \hat{\tilde{c}}_{D,I,n}^{(d)}(k; k; f_j; l) \cdot w_{OA}(L+l) \quad (36)$$

$$\hat{\tilde{c}}_{D,IN,n}(k, f_j; l) = \sum_{d \in \mathcal{D}_{DIR}(k+1, f_j)} \hat{\tilde{c}}_{D,I,n}^{(f_j)}(k+1; k; d; l) \cdot w_{OA}(l) \quad (37) \quad 30$$

によって決定される。ここで、ベクトル

【数 1 2 9】

$$\mathbf{w}_{OA} = [w_{OA}(1) \ w_{OA}(2) \ \dots \ w_{OA}(2L)]^T \in \mathbb{R}^{2L} \quad (38)$$

は重複加算の窓関数を表わす。窓関数の例は、周期的ハン窓によって与えられ、その要素は

【数 1 3 0】

$$w_{OA}(l) = \frac{1}{2} \left[1 - \cos \left(2\pi \frac{l-1}{2L} \right) \right] \quad (39)$$

によって定義される。

【 0 1 0 0 】

サブバンドHOA組成 (composition)

各サブバンドまたはサブバンド・グループ f_j 、 $j = 1, \dots, F$ について、デコードされたサブバンドHOA表現

【数 1 3 1】

$$\tilde{\mathcal{E}}(k, f_j)$$

の係数シーケンス

【数 1 3 2】

10

$$\hat{\mathcal{E}}_n(k, f_j), n = 1, \dots, O$$

は、以前に伝送されていれば打ち切りされたHOA表現

【数 1 3 3】

$$\tilde{\mathcal{E}}_T(k, f_j)$$

20

の係数シーケンスに設定され、そうでなければ方向性サブバンド合成ブロック 5 4 の一つによって与えられる方向性HOA成分

【数 1 3 4】

$$\tilde{\mathcal{E}}_D(k, f_j)$$

の係数シーケンスに設定される。すなわち、

【数 1 3 5】

30

$$\hat{\mathcal{E}}_n(k, f_j) = \begin{cases} \tilde{\mathcal{E}}_{T,n}(k, f_j) & n \in \mathcal{I}_{\text{CACT}}(k) \text{ の場合} \\ \tilde{\mathcal{E}}_{D,n}(k, f_j) & \text{それ以外の場合} \end{cases} \quad (40)$$

このサブバンド組成は、一つまたは複数のサブバンド組成ブロック 5 5 によって実行される。ある実施形態では、各サブバンドまたはサブバンド・グループについて、よって前記一つまたは複数の方向性サブバンド合成ブロック 5 4 のそれぞれについて、別個のサブバンド組成ブロック 5 5 が使われる。ある実施形態では、方向性サブバンド合成 (Directional Subband Synthesis) ブロック 5 4 およびその対応するサブバンド組成ブロック (Subband Composition Block) 5 5 は、単一のブロックに統合される。

40

【0 1 0 1】

合成フィルタバンク

最終段階では、デコードされたHOA表現が、すべてのデコードされたサブバンドHOA表現

【数 1 3 6】

$$\hat{\mathcal{E}}(k, f_j), j = 1, \dots, F$$

から合成される。圧縮解除されたHOA表現

50

【数 1 3 7】

$$\hat{c}(k)$$

の個々の時間領域係数シーケンス

【数 1 3 8】

$$\hat{c}_n(k), n = 1, \dots, 0$$

10

は、一つまたは複数の合成フィルタバンク 5 6 によって対応するサブバンド係数シーケンス

【数 1 3 9】

$$\hat{c}_n(k, f_j), j = 1, \dots, F$$

から合成される。該一つまたは複数の合成フィルタバンク 5 6 は、圧縮解除されたHOA表現

20

【数 1 4 0】

$$\hat{c}(k)$$

を最終的に出力する。

【0 1 0 2】

合成された時間領域係数シーケンスは通例、分解および合成フィルタバンク 5 3、5 6 の相続く適用のため、遅延をもつことを注意しておく。

【0 1 0 3】

30

図 8 は、単一の周波数サブバンド f_1 について、一組のアクティブな方向の候補、それらの選ばれた軌跡および対応するタプル集合を例示的に示している。フレーム k では、四つの方向が周波数サブバンド f_1 においてアクティブである。これらの方向はそれぞれの軌跡 T_1 、 T_2 、 T_3 および T_5 に属する。前のフレーム $k-2$ および $k-1$ では、異なる方向、すなわちそれぞれ T_1 、 T_2 、 T_6 および $T_1 \sim T_4$ がアクティブであった。フレーム k におけるアクティブな方向の集合 $M_{DIR}(k)$ は、フルバンドに関係し、いくつかのアクティブな方向の候補を含む。たとえば、 $M_{DIR}(k) = \{ \text{ }_3, \text{ }_8, \text{ }_{52}, \text{ }_{101}, \text{ }_{229}, \text{ }_{446}, \text{ }_{581} \}$ である。各方向は任意の仕方、たとえば二つの角度によってまたはあらかじめ定義されたテーブルのインデックスとして、表現できる。アクティブなフルバンドの方向の集合から、サブバンドおよびその対応する軌跡において実際にアクティブである方向が、各周波数サブバンドにつ

40

【0 1 0 4】

下記は、例示的な集合 $I_{C,ACT}(k) = \{1, 2, 4, 6\}$ における係数シーケンスに対応する例示的な打ち切りされたHOA表現 $C_T(k)$ の係数行列の一部である。

【0 1 0 5】

【数 1 4 1】

$$C_T(k) = \begin{bmatrix} c_{T,1}(k,1) & c_{T,1}(k,2) & c_{T,1}(k,3) & \dots \\ c_{T,2}(k,1) & c_{T,2}(k,2) & c_{T,2}(k,3) & \dots \\ 0 & 0 & 0 & \dots \\ c_{T,4}(k,1) & c_{T,4}(k,2) & c_{T,4}(k,3) & \dots \\ 0 & 0 & 0 & \dots \\ c_{T,6}(k,1) & c_{T,6}(k,2) & c_{T,6}(k,3) & \dots \\ \dots & \dots & \dots & \dots \end{bmatrix}$$

10

$I_{C,ACT}(k)$ によれば、行1、2、4、6の係数のみが0に設定されない（とはいえ、信号によっては0になることもある）。行列 $C_T(k)$ の各列は、サンプルを指し、該行列の各行は係数シーケンスである。圧縮は、すべての係数シーケンスではなく、いくつかの選択された係数シーケンス、すなわちインデックスがそれぞれ $I_{C,ACT}(k)$ および割り当てベクトル $v_A(k)$ に含まれている係数シーケンスのみがエンコードされ伝送されることを含む。デコーダでは、それらの係数は圧縮解除されて、再構成された打ち切りされたHOA表現の正しい行列に配置される。それらの行についての情報は、割り当てベクトル $v_{AMB,ASSIGN}(k)$ から得られる。 $v_{AMB,ASSIGN}(k)$ は、それぞれの伝送される係数シーケンスについて使われるトランスポート・チャネルをさらに与える。残りの係数シーケンスは0で満たされ、のちに、受領された（通例は0でない）係数から、受領されたサイド情報、たとえばサブバンドもしくはサブバンド・グループに関係した予測行列および方向に従って予測される。

20

【0 1 0 6】

サブバンド・グループ化

ある実施形態では、使用される諸サブバンドは、人間の聴覚の音響心理学的属性に適合された異なる帯域幅をもつ。あるいはまた、分解フィルタバンク53からのいくつかのサブバンドが組み合わせられて、異なる帯域幅をもつサブバンドをもつ適応されたフィルタバンクをなす。分解フィルタバンク53からの隣接するサブバンドのグループが同じパラメータを使って処理される。組み合わせられたサブバンドのグループが使われる場合、エンコーダ側で適用される対応するサブバンド構成がデコーダ側に知られていなければならない。ある実施形態では、構成情報が伝送され、デコーダによって合成フィルタバンクをセットアップするために使われる。ある実施形態では、構成情報は、（たとえばリスト内の）複数のあらかじめ定義された既知の構成のうちの一つについての識別子を含む。

30

【0 1 0 7】

もう一つの実施形態では、サブバンド構成を定義するための必要とされるビット数を減らす次の柔軟な解決策が使われる。サブバンド構成の効率的なエンコードのために、最初、最後から二番目および最後のサブバンド・グループのデータが、他のサブバンド・グループとは異なる仕方で扱われる。さらに、サブバンド・グループ帯域幅差分値がエンコードにおいて使われる。原理的には、サブバンド・グループ化情報符号化方法は、オーディオ信号の一つまたは複数のフレームについて有効なサブバンド・グループについてサブバンド構成データを符号化するために好適である。ここで、各サブバンド・グループは一つまたは複数の隣接するもとのサブバンドの組み合わせであり、もとのサブバンドの数はあらかじめ定義されている。ある実施形態では、次のサブバンド・グループの帯域幅は、現在サブバンド・グループの帯域幅以上である。この方法は、 N_{SB} 個のサブバンド・グループを、 $N_{SB} - 1$ を表わす固定数のビットを用いて符号化することを含み、 $N_{SB} > 1$ であれば、最初のサブバンド・グループ g_1 について、帯域幅値 $B_{SB}[1]$ を、 $B_{SB}[1] - 1$ を表わす単進符号を用いて符号化することを含む。 $N_{SB} = 3$ であれば、固定数のビットをもつ帯域幅差分値 $B_{SB}[2] = B_{SB}[2] - B_{SB}[1]$ が、第二のサブバンド・グループ g_2 について符号化される。 $N_{SB} > 3$ であれば、対応する数の帯域幅差分値 $B_{SB}[g] = B_{SB}[g] - B_{SB}[g - 1]$ が、サブバンド・グループ $g_2, \dots, g_{N_{SB}-2}$ について単進符号を用いて符号化され、帯域幅差分値 $B_{SB}[N_{SB} - 1] = B_{SB}[N_{SB} - 1] - B_{SB}[N_{SB} - 2]$ が最後のサブバンド・グループ $g_{N_{SB}-1}$ について固定数の

40

50

ビットを用いて符号化される。サブバンド・グループについての帯域幅値は、隣接するものとのサブバンドの数として表現される。最後のサブバンド・グループ g_{SB} については、符号化されるサブバンド構成データに対応する値が含まれる必要はない。

【0108】

図9は、通常のMPEG-H 3Dオーディオ・エンコーダのHOAエンコード経路の一般化されたブロック図を示している。二つの型の優勢音信号が抽出される：方向性音抽出（Directional Sound Extraction）ブロックDSEにおける方向性信号と、VVec音抽出（VVec Sound Extraction）ブロックVSEにおけるベクトル・ベースの信号VVecである。ベクトル・ベースの信号VVec（Vベクトル）に属するベクトルは、対応するベクトル・ベースの信号についての音場の空間的分布を表わす。さらに、残差／周囲音用計算器（Calculator for Residuuum/Ambience）CRAにおいて周囲成分もエンコードされる。CRAにより、方向性音抽出ブロックDSEおよびVVec音抽出ブロックVSEからの出力データのいずれか一方または両方が使われてもよく、あるいはどちらも使われなくてもよい。周囲信号は空間分解能低減（Spatial Resolution Reduction）ブロックSRR、部分的脱相関（partial decorrelation）PDおよび利得制御 GC_A にかけられる。四角内のブロックは音シーン解析（Sound Scene Analysis）SSAによって制御される。統合音声音響符号化器USAC3Dに入力される前に、優勢音信号もそれぞれの利得制御ブロック GC_D 、 GC_V によって処理される。最後に、USAC3DエンコーダENCc&HEPcがHOA空間的サイド情報をHOA拡張ペイロード中にパッキングする。

【0109】

図10は、ある実施形態に基づく、MPEGにおいて使用可能な改善されたオーディオ・エンコーダを示している。開示される技術は、現行のMPEG-H 3Dオーディオのシステムを、低帯域幅のためのビットストリームが既知のMPEG-H 3Dオーディオのフォーマットの真の上位集合であるような仕方で修正する。図9に比べると、音シーン解析SSAにおいて、二つの新たなブロックを含む経路が追加されている。これらは、周囲信号に適用されるQMF分解（QMF Analysis）フィルタバンクQAcと、方向性サブバンド信号のパラメータの計算のための方向性サブバンド計算（Directional Subband Calculation）ブロックDSCcである。これらのパラメータは、伝送された周囲信号に基づいて方向性信号を合成することを許容する。さらに、欠けている周囲信号を再生することを許容するパラメータが計算される。合成プロセスについてのサイド情報パラメータはUSAC3DエンコーダENC&HEPに渡され、それにより圧縮された出力信号HOAc,oのHOA拡張ペイロード中にパッキングされる。有利なことに、圧縮は、図9の構成で達成される通常の圧縮より効率的である。

【0110】

図11は、通常のMPEG-H 3Dオーディオ・デコーダの一般化されたブロック図を示している。まず、圧縮された入力ビットストリームHOAc,iからHOAサイド情報が抽出され、USAC3DおよびHOA拡張ペイロード・デコーダDECc&HEPcが伝送チャネル波形信号を再生する。これらに対応する逆利得制御ブロック IGC_D 、 IGC_V 、 IGC_A に入力される。ここで、エンコーダにおいて適用された正規化が打ち消される。対応する伝送チャネルは、サイド情報とともに、それぞれHOA方向性音合成（Directional Sound Synthesis）ブロックDSSおよび／またはVVec音合成（VVec Sound Synthesis）ブロックVSSにおいて優勢音信号（方向性および／またはベクトル・ベース）を合成するために使われる。第三の経路では、周囲成分が逆部分的脱相関（Inverse Partial Decorrelation）IPDブロックおよびHOA周囲音合成（HOA Ambience Synthesis）HASブロックによって再生される。続くHOA組成（HOA Composition）ブロックHCcは、優勢音成分と周囲音を組み合わせ、デコードされたHOA信号を構築する。これはHOAレンダラー（HOA renderer）HRに入力されて、出力信号HOA'd,o、すなわち最終的なスピーカー・フィールドを生成する。

【0111】

図12は、ある実施形態に基づく、MPEGにおいて使用可能な改善されたオーディオ・デコーダを示している。エンコーダと同様に、経路が加わっている。それは、サブバンド信号の計算のためのデコーダ側QMF分解（QMF Analysis）ブロックQA_Dと、パラメトリック

にエンコードされた方向性サブバンド信号の合成のための方向性サブバンド信号合成 (Directional Subband signal Synthesis) ブロックDSC_Dとを有する。計算されたサブバンド信号は、対応する伝送されたサイド情報とともに、方向性信号のHOA表現を合成するために使われる。その後、合成された信号成分はQMF合成 (QMF synthesis) フィルタバンクQSを使って時間領域に変換される。その出力信号はさらに、向上されたHOA組成ブロックHCに入力される。デコードされたHOA出力信号HOA_{D, O}を提供するための、続くHOAレンダリング・ブロックHRは変更なしとされる。

【 0 1 1 2 】

以下では、高次アンビソニックスのいくつかの基本的な特徴が説明される。

【 0 1 1 3 】

10

高次アンビソニックス (HOA) は、音源がないと想定されるコンパクトな関心領域内の音場の記述に基づく。その場合、関心領域内の位置 \mathbf{x} および時刻 t における音圧の空間時間的挙動 $p(t, \mathbf{x})$ は、斉次の波の式 (homogeneous wave equation) によって物理的に完全に決定される。以下では、図 6 に示される球面座標系を想定する。この座標系では、 x 軸は前方位置を向き、 y 軸は左を向き、 z 軸は上を向く。空間内の位置 $\mathbf{x} = (r, \theta, \phi)^T$ は動径 $r > 0$ (すなわち、座標原点までの距離)、極軸 z (!) から測った傾斜角 $\theta \in [0, \pi]$ および xy 平面において x 軸から反時計回りに測った方位角 $\phi \in [0, 2\pi]$ によって表現される。さらに、 $(\cdot)^T$ は転置を表わす。

【 0 1 1 4 】

すると、 ω が角周波数を表わし、 i は虚数単位を示すものとして、 $F_t(\cdot)$ によって表わされる時間に関する音圧のフーリエ変換、すなわち

20

【 数 1 4 2 】

$$P(\omega, \mathbf{x}) = F_t(p(t, \mathbf{x})) = \int_{-\infty}^{\infty} p(t, \mathbf{x}) e^{-i\omega t} dt \quad (41)$$

は、

【 数 1 4 3 】

$$P(\omega = kc_s, r, \theta, \phi) = \sum_{n=0}^N \sum_{m=-n}^n A_n^m(k) j_n(kr) S_n^m(\theta, \phi) \quad (42)$$

30

に従って球面調和関数級数に展開されうることが示せる (非特許文献 1 1)。式 (42) において、 c_s は音速を表わし、 k は角波数を表わす。角波数は角周波数 ω に、 $k = \omega / c_s$ によって関係付けられる。さらに、 $j_n(\cdot)$ は第一種の球面ベッセル関数を表わし、 $S_n^m(\cdot, \cdot)$ は次数 (order) n 、陪数 (degree) m の実数値の球面調和関数を表わす。展開係数 $A_n^m(k)$ は角波数 k のみに依存する。音圧が空間的に帯域制限されていることが暗黙的に想定されていることを注意しておく。よって、級数は次数インデックス n に関して上限 N で打ち切られる。この N はHOA符号化表現の次数と呼ばれる。

【 0 1 1 5 】

40

音場が異なる角周波数 ω の無限個の調和平面波の重ね合わせによって表現され、角タプル (θ, ϕ) によって指定されるすべての可能な方向から到来するとすると、それぞれの平面波複素振幅関数 $C(\theta, \phi)$ は次の球面調和関数展開によって表わせることを示せる (非特許文献 1 0)。

【 0 1 1 6 】

【数 1 4 4】

$$C(\omega = kc_s, \theta, \phi) = \sum_{n=0}^N \sum_{m=-n}^n C_n^m(k) S_n^m(\theta, \phi) \quad (43)$$

ここで、展開係数 $C_n^m(k)$ は展開係数 $A_n^m(k)$ に、

$$A_n^m(k) = i^n C_n^m(k) \quad (44)$$

によって関係付けられる。個々の係数 $C_n^m(k = \omega/c_s)$ が角周波数 ω の関数であるとする
て、逆フーリエ変換 ($F^{-1}(\cdot)$ によって表わされる) の適用は、各次数 n および陪数 m につい

10

【数 1 4 5】

$$c_n^m(t) = F_t^{-1}(C_n^m(\omega/c_s)) = \frac{1}{2\pi} \int_{-\infty}^{\infty} C_n^m\left(\frac{\omega}{c_s}\right) e^{i\omega t} d\omega \quad (45)$$

を与える。これらの時間領域関数はここでは連続時間HOA係数シーケンスと称され、これ

【数 1 4 6】

20

 $\mathbf{c}(t)$

$$= \left[c_0^0(t) \ c_1^{-1}(t) \ c_1^0(t) \ c_1^1(t) \ c_2^{-2}(t) \ c_2^{-1}(t) \ c_2^0(t) \ c_2^1(t) \ c_2^2(t) \ \dots \ c_N^{N-1}(t) \ c_N^N(t) \right]^T$$

(46)

によって単一のベクトル $\mathbf{c}(t)$ にまとめることができる。

ベクトル $\mathbf{c}(t)$ 内のHOA係数シーケンス $c_n^m(t)$ の位置インデックスは

30

$n(n+1) + 1 + m$

によって与えられる。

【0 1 1 7】

ベクトル $\mathbf{c}(t)$ 内の全体的な要素数は $O = (N+1)^2$ によって与えられる。

最終的なアンビソニックス・フォーマットは、サンプリング周波数 f_s を使って、 $\mathbf{c}(t)$ のサンプリングされたバージョンを、

【数 1 4 7】

$$\{\mathbf{c}(lT_s)\}_{l \in \mathbf{N}} = \{\mathbf{c}(T_s), \mathbf{c}(2T_s), \mathbf{c}(3T_s), \mathbf{c}(4T_s), \dots\} \quad (47)$$

40

として与える。ここで、 $T_s = 1/f_s$ はサンプリング周期を表わす。 $\mathbf{c}(lT_s)$ の要素はここでは離散時間HOA係数シーケンスと称される。これは常に実数値であることが示せる。この属性は、連続時間バージョン $c_n^m(t)$ についても成り立つ。

【0 1 1 8】

実数値の球面調和関数の定義

実数値の球面調和関数 $S_n^m(\cdot, \cdot)$ (SN3D規格化 (非特許文献 1、3.1章) を想定) は次式によって与えられる。

【0 1 1 9】

【数 1 4 8】

$$S_n^m(\theta, \phi) = \sqrt{(2n+1) \frac{(n-|m|)!}{(n+|m|)!}} P_{n,|m|}(\cos\theta) \text{trg}_m(\phi) \quad (48)$$

ここで

$$\text{trg}_m(\phi) = \begin{cases} \sqrt{2}\cos(m\phi) & m > 0 \\ 1 & m = 0 \\ -\sqrt{2}\sin(m\phi) & m < 0 \end{cases} \quad (49)$$

10

ルジャンドル陪関数 $P_{n,m}(x)$ は次式によって定義される。

【0 1 2 0】

【数 1 4 9】

$$P_{n,m}(x) = (1-x^2)^{m/2} \frac{d^m}{dx^m} P_n(x), m \geq 0 \quad (50)$$

20

ここで、ルジャンドル多項式 $P_n(x)$ を用いているが、非特許文献 1 1 とは異なり、コンドン・ショートリー (Condon-Shortley) 位相項 $(-1)^m$ は使っていない。

【0 1 2 1】

ある実施形態では、(複素数値のフィルタバンクから得られた) HOA 信号表現のサブバンドまたはサブバンド・グループ内の優勢な方向性信号の方向をフレームごとに決定し、効率的にエンコードする方法は、

各現在フレーム k について：HOA 信号におけるフルバンド方向候補の集合 $M_{DIR}(k)$ 、集合 $M_{DIR}(k)$ における要素数 $NoOfGlobalDirs$ および該要素数をエンコードするために必要とされる数 $D(k) = \log_2(NoOfGlobalDirs)$ を決定する段階であって、各フルバンド方向候補は、 Q 個の可能な方向のあらかじめ定義されたフル集合に関するグローバル・インデックス q (q

30

$[1, \dots, Q]$) をもつ、段階と；

現在フレーム k の各サブバンドまたはサブバンド・グループ j について、集合 $M_{DIR}(k)$ 内のフルバンド方向候補のうちどの方向がアクティブなサブバンド方向として現われるかを判別する段階と；

前記サブバンドまたはサブバンド・グループのいずれかにおいてアクティブなサブバンド方向として現われる、使用されるフルバンド方向候補の集合 $M_{FB}(k)$ (みな HOA 信号中のフルバンド方向候補の集合 $M_{DIR}(k)$ に含まれる) および使用されるフルバンド方向候補の集合 $M_{FB}(k)$ 内の要素数 $NoOfGlobalDirs(k)$ を決定する段階と；

現在フレーム k の各サブバンドまたはサブバンド・グループ j について：集合 $M_{DIR}(k)$ 内のフルバンド方向候補のうち d 個 ($d \in [1, \dots, D]$) までの方向のどの方向がアクティブなサブバンド方向であるかを判別し；アクティブなサブバンド方向のそれぞれについて軌跡および軌跡インデックスを決定し；前記軌跡インデックスをそれぞれのアクティブなサブバンド方向に割り当てる段階と；

40

現在のサブバンドまたはサブバンド・グループ j 内のアクティブなサブバンド方向のそれぞれを $D(k)$ ビットをもつ相対インデックスによってエンコードする段階とを含む。

【0 1 2 2】

ある実施形態では、コンピュータ可読媒体が、コンピュータに、優勢な方向性信号の方向をフレームごとに決定し、効率的にエンコードするためのこの方法を実行させる実行可能命令を記憶している。

【0 1 2 3】

50

さらに、ある実施形態では、HOA信号表現のサブバンド内の優勢な方向性信号の方向をデコードする方法が、
 デコードされるべきHOA信号表現について最大数の方向Dの諸インデックスを受領する段階と；デコードされるべきHOA信号表現の最大数の方向Dの諸方向を再構成する段階と；サブバンド毎のアクティブな諸方向信号の諸インデックスを受領する段階と；デコードされるべきHOA信号表現の再構成された諸方向Dおよびサブバンド毎のアクティブな方向信号の諸インデックスからサブバンド毎のアクティブな諸方向を再構成する段階と；諸サブバンドの方向性信号を予測する段階とを含み、サブバンドの現在フレームにおける方向性信号の前記予測は、そのサブバンドの先行フレームの方向性信号を決定することを含み、前記方向性信号のインデックスが前記先行フレームにおいて0であり現在フレームにおいて0でない場合には新たな方向性信号が生成され、前記方向性信号のインデックスが前記先行フレームにおいて0でなく現在フレームにおいて0である場合には前の方向性信号がキャンセルされ、方向性信号のインデックスが第一の方向から第二の方向に変わる場合には、該方向性信号の方向が該第一の方向から該第二の方向に動かされる。

10

【 0 1 2 4 】

ある実施形態では、図 1 および図 3 に示され、上記で論じたように、所与の数の係数シーケンスをもつ入力HOA信号のフレームをエンコードするための装置であって、各係数シーケンスはインデックスをもつ、装置が、少なくとも一つのハードウェア・プロセッサと、少なくとも一つのソフトウェア・コンポーネントを有体に具現する非一時的な有体なコンピュータ可読記憶媒体とを有する。前記ソフトウェア・コンポーネントは、前記少なくとも一つのハードウェア・プロセッサ上で実行されるときに、前記ハードウェア・プロセッサに、

20

低減された数の0でない係数シーケンスをもつ打ち切りされたHOA表現 $C_T(k)$ を計算する段階 1 1 と、

打ち切りされたHOA表現に含まれているアクティブな係数シーケンスのインデックスの集合 $I_{C,ACT}(k)$ を決定する段階 1 1 と、

前記入力HOA信号から、候補方向の第一の集合 $M_{DIR}(k)$ を推定する段階 1 6 と；

前記入力HOA信号を複数の周波数サブバンド f_1, \dots, f_F に分割する段階 1 5 であって、それらの周波数サブバンドの係数シーケンス

【 数 1 5 0 】

30

$$\tilde{C}(k-1, k, f_1), \dots, \tilde{C}(k-1, k, f_F)$$

が得られる、段階と、

それらの周波数サブバンドのそれぞれについて、方向の第二の集合 $M_{DIR}(k, f_1), \dots, M_{DIR}(k, f_F)$ を推定する段階 1 6 であって、方向の前記第二の集合の各要素は第一および第二のインデックスをもつインデックスのタプルであり、前記第二のインデックスは現在の周波数サブバンドについてアクティブな方向のインデックスであり、前記第一のインデックスは該アクティブな方向の軌跡インデックスであり、各アクティブな方向は前記入力HOA信号の候補方向の前記第一の集合 $M_{DIR}(k)$ にも含まれる、段階と、

40

前記周波数サブバンドのそれぞれについて、それぞれの周波数サブバンドの方向の前記第二の集合 $M_{DIR}(k, f_1), \dots, M_{DIR}(k, f_F)$ に応じてその周波数サブバンドの係数シーケンス

【 数 1 5 1 】

$$\tilde{C}(k-1, k, f_1), \dots, \tilde{C}(k-1, k, f_F)$$

から方向性サブバンド信号

【数 1 5 2】

$$\bar{X}(k-1, k, f_1), \dots, \bar{X}(k-1, k, f_F)$$

を計算する段階 1 7 と、

前記周波数サブバンドのそれぞれについて、それぞれの周波数サブバンドのアクティブな係数シーケンスのインデックスの前記集合 $I_{C, ACT}(k)$ を使ってその周波数サブバンドの係数シーケンス

【数 1 5 3】

10

$$\bar{C}(k-1, k, f_1), \dots, \bar{C}(k-1, k, f_F)$$

から方向性サブバンド信号

【数 1 5 4】

$$\bar{X}(k-1, k, f_1), \dots, \bar{X}(k-1, k, f_F)$$

を予測するよう適応された予測行列 $A(k, f_1), \dots, A(k, f_F)$ を計算する段階 1 8 と、

20

候補方向の前記第一の集合 $M_{DIR}(k)$ 、方向の前記第二の集合 $M_{DIR}(k, f_1), \dots, M_{DIR}(k, f_F)$ 、前記予測行列 $A(k, f_1), \dots, A(k, f_F)$ および前記打ち切りされたHOA表現 $C_T(k)$ をエンコードする段階とを実行させる。

【0 1 2 5】

ある実施形態では、図 4 および図 5 に示され、上記で論じたように、圧縮されたHOA表現をデコードするための装置が、少なくとも一つのハードウェア・プロセッサと、少なくとも一つのソフトウェア・コンポーネントを有体に具現する非一時的な有体なコンピュータ可読記憶媒体とを有する。前記ソフトウェア・コンポーネントは、前記少なくとも一つのハードウェア・プロセッサ上で実行されるときに、前記ハードウェア・プロセッサに、前記圧縮されたHOA表現から、複数の打ち切りされたHOA係数シーケンス

30

【数 1 5 5】

$$\hat{z}_1(k), \dots, \hat{z}_I(k),$$

前記打ち切りされたHOA係数シーケンスのシーケンス・インデックスを示すまたは含む割り当てベクトル $v_{AMB, ASSIGN}(k)$ 、サブバンドに関係した方向情報 $M_{DIR}(k+1, f_1), \dots, M_{DIR}(k+1, f_F)$ 、複数の予測行列 $A(k+1, f_1), \dots, A(k+1, f_F)$ および利得制御サイド情報 $e_1(k), \dots, e_I(k)$ 、 $\hat{z}_I(k)$ を抽出する段階 4 1、4 2、4 3 と；

前記複数の打ち切りされたHOA係数シーケンス

40

【数 1 5 6】

$$\hat{z}_1(k), \dots, \hat{z}_I(k),$$

前記利得制御サイド情報 $e_1(k), \dots, e_I(k)$ 、 $\hat{z}_I(k)$ および前記割り当てベクトル $v_{AMB, ASSIGN}(k)$ から、打ち切りされたHOA表現

【数 1 5 7】

$$\hat{C}_T(k)$$

50

を再構成する段階 5 1、5 2 と；
 分解フィルタバンク 5 3 において、前記再構成された打ち切りされたHOA表現
 【数 1 5 8】

$$\hat{c}_T(k)$$

を複数F個の周波数サブバンドについての周波数サブバンド表現
 【数 1 5 9】

10

$$\hat{c}_T(k, f_1), \dots, \hat{c}_T(k, f_F)$$

に分解する段階と；
 各周波数サブバンド表現について方向性サブバンド合成ブロック 5 4 において、前記再構成された打ち切りされたHOA表現の前記それぞれの周波数サブバンド表現
 【数 1 6 0】

$$\hat{c}_T(k, f_1), \dots, \hat{c}_T(k, f_F)$$

20

と、前記サブバンドに関係した方向情報 $M_{DIR}(k+1, f_1), \dots, M_{DIR}(k+1, f_F)$ と、前記予測行列 $A(k+1, f_1), \dots, A(k+1, f_F)$ とから、予測された方向性HOA表現
 【数 1 6 1】

$$\hat{c}_D(k, f_1), \dots, \hat{c}_D(k, f_F)$$

を合成する段階 5 4 と；
 前記F個の周波数サブバンドのそれぞれについてサブバンド組成ブロック 5 5 において、
 係数シーケンス
 【数 1 6 2】

30

$$\hat{c}_n(k, f_j), n = 1, \dots, 0$$

をもつデコードされたサブバンドHOA表現
 【数 1 6 3】

$$\hat{c}(k, f_1), \dots, \hat{c}(k, f_F)$$

40

を組成する段階 5 5 であって、該係数シーケンスは、該係数シーケンスが前記割り当てベクトル $v_{AMB, ASSIGN}(k)$ に含まれるインデックスnをもつ場合には前記打ち切りされたHOA表現
 【数 1 6 4】

$$\hat{c}_T(k, f_j)$$

50

の係数シーケンスから得られ、あるいはそれ以外の場合には前記方向性サブバンド合成ブロック 5 4 の一つによって与えられる前記予測された方向性HOA成分

【数 1 6 5】

$$\tilde{\mathbf{c}}_D(k, f_j)$$

の係数シーケンスから得られる、段階と；

一つまたは複数の合成フィルタバンク 5 6 において、前記デコードされたサブバンドHOA表現

【数 1 6 6】

10

$$\tilde{\mathbf{c}}(k, f_1), \dots, \tilde{\mathbf{c}}(k, f_F)$$

を合成して、デコードされたHOA表現

【数 1 6 7】

$$\hat{\mathbf{c}}(k)$$

を得る段階 5 6 とを実行させる。

20

【0 1 2 6】

ある実施形態では、所与の数の係数シーケンスをもつ入力HOA信号のフレームをエンコードするための装置 1 0 であって、各係数シーケンスはインデックスをもつ、装置が、低減された数の0でない係数シーケンスをもつ打ち切りされたHOA表現 $\mathbf{C}_T(k)$ を計算するよう構成され、打ち切りされたHOA表現に含まれているアクティブな係数シーケンスのインデックスの集合 $I_{C,ACT}(k)$ を決定するさらに構成された計算および決定モジュール 1 1 と；

前記入力HOA信号を複数の周波数サブバンド f_1, \dots, f_F に分割するよう構成された分解フィルタバンク・モジュール 1 5 であって、それらの周波数サブバンドの係数シーケンス

【数 1 6 8】

30

$$\bar{\mathbf{c}}(k-1, k, f_1), \dots, \bar{\mathbf{c}}(k-1, k, f_F)$$

が得られる、分解フィルタバンク・モジュールと；

前記入力HOA信号から、候補方向の第一の集合 $M_{DIR}(k)$ を推定するよう構成され、前記周波数サブバンドのそれぞれについて、方向の第二の集合 $M_{DIR}(k, f_1), \dots, M_{DIR}(k, f_F)$ を推定するようさらに構成された方向推定モジュール 1 6 であって、方向の前記第二の集合の各要素は第一および第二のインデックスをもつインデックスのタプルであり、前記第二のインデックスは現在の周波数サブバンドについてアクティブな方向のインデックスであり、前記第一のインデックスは該アクティブな方向の軌跡インデックスであり、各アクティブな方向は前記入力HOA信号の候補方向の前記第一の集合 $M_{DIR}(k)$ にも含まれる、方向推定モジュールと；

40

前記周波数サブバンドのそれぞれについて、それぞれの周波数サブバンドの方向の前記第二の集合 $M_{DIR}(k, f_1), \dots, M_{DIR}(k, f_F)$ に応じてその周波数サブバンドの係数シーケンス

【数 1 6 9】

$$\bar{\mathbf{c}}(k-1, k, f_1), \dots, \bar{\mathbf{c}}(k-1, k, f_F)$$

から方向性サブバンド信号

50

【数 1 7 0】

$$\bar{\bar{X}}(k-1, k, f_1), \dots, \bar{\bar{X}}(k-1, k, f_F)$$

を計算するよう構成された少なくとも一つの方向性サブバンド計算モジュール 1 7 と；
前記周波数サブバンドのそれぞれについて、それぞれの周波数サブバンドのアクティブな
係数シーケンスのインデックスの前記集合 $I_{C, ACT}(k)$ を使ってその周波数サブバンドの係
数シーケンス

【数 1 7 1】

10

$$\bar{\bar{C}}(k-1, k, f_1), \dots, \bar{\bar{C}}(k-1, k, f_F)$$

から方向性サブバンド信号

【数 1 7 2】

$$\bar{\bar{X}}(k-1, k, f_1), \dots, \bar{\bar{X}}(k-1, k, f_F)$$

20

を予測するよう適応された予測行列 $A(k, f_1), \dots, A(k, f_F)$ を計算するよう構成された少なく
とも一つの方向性サブバンド予測モジュール 1 8 と；

候補方向の前記第一の集合 $M_{DIR}(k)$ 、方向の前記第二の集合 $M_{DIR}(k, f_1), \dots, M_{DIR}(k, f_F)$
、前記予測行列 $A(k, f_1), \dots, A(k, f_F)$ および前記打ち切りされたHOA表現 $C_T(k)$ をエンコード
するよう構成されたエンコード・モジュール 3 0 とを有する。

【0 1 2 7】

ある実施形態では、前記装置はさらに、前記打ち切りされたHOAチャネル・シーケンス
を部分的に脱相関するよう構成された部分的脱相関器 1 2 と；前記打ち切りされたHOAチ
ャネル・シーケンス $y_1(k), \dots, y_I(k)$ をトランスポート・チャネルに割り当てるよう構成さ
れたチャネル割り当てモジュール 1 3 と；前記トランスポート・チャネルに対して利得制
御を実行するよう構成された少なくとも一つの利得制御ユニット 1 4 であって、各トラン
スポート・チャネルについての利得制御サイド情報 $e_i(k-1)$ 、 $i(k-1)$ が生成される、
ユニットとをさらに有する。

30

【0 1 2 8】

ある実施形態では、エンコード・モジュール 3 0 は、利得制御された打ち切りされたHO
Aチャネル・シーケンス $z_1(k), \dots, z_I(k)$ をエンコードするよう構成された知覚的エンコー
ダと；前記利得制御サイド情報 $e_i(k-1)$ 、 $i(k-1)$ 、候補方向の前記第一の集合 $M_{DIR}(k)$
、方向の前記第二の集合 $M_{DIR}(k, f_1), \dots, M_{DIR}(k, f_F)$ および前記予測行列 $A(k, f_1), \dots, A(k, f_F)$
をエンコードするよう構成されたサイド情報源符号化器 3 2 と；前記知覚的エンコー
ダ 3 1 および前記サイド情報源符号化器 3 2 の出力を多重化して、エンコードされたHOA
信号フレーム

40

【数 1 7 3】

$$\bar{\bar{B}}(k-1)$$

を得るよう構成されたマルチプレクサ 3 3 とを有する。

【0 1 2 9】

ある実施形態では、HOA信号をデコードする装置 5 0 が、
前記圧縮されたHOA表現から、複数の打ち切りされたHOA係数シーケンス

【数 1 7 4】

$$\hat{\mathbf{z}}_1(k), \dots, \hat{\mathbf{z}}_I(k),$$

前記打ち切りされたHOA係数シーケンスのシーケンス・インデックスを示すまたは含む割り当てベクトル $\mathbf{v}_{\text{AMB}, \text{ASSIGN}}(k)$ 、サブバンドに関係した方向情報 $M_{\text{DIR}}(k+1, f_1), \dots, M_{\text{DIR}}(k+1, f_F)$ 、複数の予測行列 $A(k+1, f_1), \dots, A(k+1, f_F)$ および利得制御サイド情報 $e_1(k), \dots, e_I(k), \dots, e_I(k)$ を抽出するよう構成された抽出モジュール40と；
前記複数の打ち切りされたHOA係数シーケンス

【数 1 7 5】

10

$$\hat{\mathbf{z}}_1(k), \dots, \hat{\mathbf{z}}_I(k),$$

前記利得制御サイド情報 $e_1(k), \dots, e_I(k), \dots, e_I(k)$ および前記割り当てベクトル $\mathbf{v}_{\text{AMB}, \text{ASSIGN}}(k)$ から、打ち切りされたHOA表現
【数 1 7 6】

$$\hat{\mathbf{c}}_T(k)$$

20

を再構成するよう構成された再構成モジュール51、52と；
前記再構成された打ち切りされたHOA表現
【数 1 7 7】

$$\hat{\mathbf{c}}_T(k)$$

を複数F個の周波数サブバンドについての周波数サブバンド表現
【数 1 7 8】

30

$$\tilde{\mathbf{c}}_T(k, f_1), \dots, \tilde{\mathbf{c}}_T(k, f_F)$$

に分解するよう構成された分解フィルタバンク・モジュール53と；
各周波数サブバンドについて、前記再構成された打ち切りされたHOA表現の前記それぞれの周波数サブバンド表現
【数 1 7 9】

$$\tilde{\mathbf{c}}_T(k, f_1), \dots, \tilde{\mathbf{c}}_T(k, f_F)$$

40

と、前記サブバンドに関係した方向情報 $M_{\text{DIR}}(k+1, f_1), \dots, M_{\text{DIR}}(k+1, f_F)$ と、前記予測行列 $A(k+1, f_1), \dots, A(k+1, f_F)$ とから、予測された方向性HOA表現
【数 1 8 0】

$$\tilde{\mathbf{c}}_D(k, f_1), \dots, \tilde{\mathbf{c}}_D(k, f_F)$$

を合成するよう構成された少なくとも一つの方向性サブバンド合成モジュール54と；
前記F個の周波数サブバンドのそれぞれについて、係数シーケンス

50

【数 1 8 1】

$$\hat{\tilde{c}}_n(k, f_j), n = 1, \dots, O$$

をもつデコードされたサブバンドHOA表現

【数 1 8 2】

$$\hat{\tilde{c}}(k, f_1), \dots, \hat{\tilde{c}}(k, f_R)$$

10

を組成するよう構成された少なくとも一つのサブバンド組成モジュール 5 5 であって、該係数シーケンスは、該係数シーケンスが前記割り当てベクトル $v_{\text{AMB, ASSIGN}}(k)$ に含まれるインデックス n をもつ場合には前記打ち切りされたHOA表現

【数 1 8 3】

$$\hat{\tilde{c}}_T(k, f_j)$$

の係数シーケンスから得られ、あるいはそれ以外の場合には前記方向性サブバンド合成モジュール 5 4 の一つによって与えられる前記予測された方向性HOA成分

20

【数 1 8 4】

$$\hat{\tilde{c}}_D(k, f_j)$$

の係数シーケンスから得られる、モジュールと；
前記デコードされたサブバンドHOA表現

【数 1 8 5】

30

$$\hat{\tilde{c}}(k, f_1), \dots, \hat{\tilde{c}}(k, f_R)$$

を合成して、デコードされたHOA表現

【数 1 8 6】

$$\hat{\tilde{c}}(k)$$

を得るよう構成された合成フィルタバンク・モジュール 5 6 とを有する。

【0 1 3 0】

40

前記抽出モジュール 4 0 は少なくとも、エンコードされたサイド情報部分と知覚的に符号化された部分とを取得するためのデマルチプレクサ 4 1 であって、該知覚的に符号化された部分は、エンコードされた打ち切りされたHOA係数シーケンス

【数 1 8 7】

$$\check{z}_1(k), \dots, \check{z}_I(k)$$

を含む、デマルチプレクサと；

前記エンコードされた打ち切りされたHOA係数シーケンス

50

【数 1 8 8】

$$\check{z}_1(k), \dots, \check{z}_I(k)$$

を知覚的にデコードして前記打ち切りされたHOA係数シーケンス

【数 1 8 9】

$$\hat{z}_1(k), \dots, \hat{z}_I(k)$$

10

を得るよう構成された知覚的デコーダ42と；前記エンコードされたサイド情報部分をデコードして(s43)、前記サブバンド関係方向情報 $M_{DIR}(k+1, f_1), \dots, M_{DIR}(k+1, f_F)$ 、予測行列 $A(k+1, f_1), \dots, A(k+1, f_F)$ 、利得制御サイド情報 $e_1(k), \dots, e_I(k), \dots, e_I(k), \dots, e_I(k)$ および割り当てベクトル $v_{AMB, ASSIGN}(k)$ を得るよう構成されたサイド情報源デコーダ43とを有する。

【0 1 3 1】

図13は、ある実施形態における、低ビットレートのエンコード方法のフローチャートを示している。所与の数の係数シーケンスを有する入力HOA信号のフレームを低ビットレート・エンコードする方法であって、各係数シーケンスはインデックスをもつ、方法が、打ち切りされたHOA表現に含まれるアクティブな係数シーケンスのインデックスの集合 $I_{C, ACT}(k)$ を決定する段階s11と；

20

前記入力HOA信号から、候補方向の第一の集合 $M_{DIR}(k)$ を推定する段階s16と；

前記入力HOA信号を複数の周波数サブバンド f_1, \dots, f_F に分割する段階s15であって、それらの周波数サブバンドの係数シーケンス

【数 1 9 0】

$$\bar{\bar{c}}(k-1, k, f_1), \dots, \bar{\bar{c}}(k-1, k, f_F)$$

が得られる、段階と；

30

各周波数サブバンドについて、方向の第二の集合 $M_{DIR}(k, f_1), \dots, M_{DIR}(k, f_F)$ を推定する段階s161であって、方向の前記第二の集合の各要素は第一および第二のインデックスをもつインデックスのタプルであり、前記第二のインデックスは現在の周波数サブバンドについてのアクティブな方向のインデックスであり、前記第一のインデックスは該アクティブな方向の軌跡インデックスであり、各アクティブな方向は前記入力HOA信号の候補方向の前記第一の集合 $M_{DIR}(k)$ にも含まれるものであり、

各周波数サブバンドについて、その周波数サブバンドの係数シーケンス

【数 1 9 1】

$$\bar{\bar{c}}(k-1, k, f_1), \dots, \bar{\bar{c}}(k-1, k, f_F)$$

40

からそれぞれの周波数サブバンドの方向の前記第二の集合 $M_{DIR}(k, f_1), \dots, M_{DIR}(k, f_F)$ に従って、方向性サブバンド信号

【数 1 9 2】

$$\bar{\bar{x}}(k-1, k, f_1), \dots, \bar{\bar{x}}(k-1, k, f_F)$$

を計算する段階s17と；

50

各周波数サブバンドについて、それぞれの周波数サブバンドのアクティブな係数シーケンスのインデックスの前記集合 $I_{C,ACT}(k)$ を使ってその周波数サブバンドの係数シーケンス

【数 1 9 3】

$$\tilde{C}(k-1, k, f_1), \dots, \tilde{C}(k-1, k, f_F)$$

から方向性サブバンド信号

【数 1 9 4】

10

$$\tilde{X}(k-1, k, f_1), \dots, \tilde{X}(k-1, k, f_F)$$

を予測するよう適応された予測行列 $A(k, f_1), \dots, A(k, f_F)$ を計算する段階s18と；
候補方向の前記第一の集合 $M_{DIR}(k)$ 、方向の前記第二の集合 $M_{DIR}(k, f_1), \dots, M_{DIR}(k, f_F)$ 、
前記予測行列 $A(k, f_1), \dots, A(k, f_F)$ および前記打ち切りされたHOA表現 $C_T(k)$ をエンコードする
段階s19とを含む。

【0 1 3 2】

ある実施形態では、前記打ち切りされたHOA表現 $C_T(k)$ のエンコードは、前記打ち切りされたHOAチャンネル・シーケンスの部分的な脱相関s12と、前記打ち切りされたHOAチャンネル・シーケンス $y_1(k), \dots, y_I(k)$ をトランスポート・チャンネルに割り当てるためのチャンネル割り当てs13と、前記トランスポート・チャンネルのそれぞれに対して利得制御を実行する段階s14であって、各トランスポート・チャンネルについての利得制御サイド情報 $e_i(k-1)$ 、

20

$i(k-1)$ が生成される、段階と、知覚的エンコーダ31において、利得制御された打ち切りされたHOAチャンネル・シーケンス $z_1(k), \dots, z_I(k)$ をエンコードする段階s31と、前記利得制御サイド情報 $e_i(k-1)$ 、 $i(k-1)$ 、候補方向の前記第一の集合 $M_{DIR}(k)$ 、方向の前記第二の集合 $M_{DIR}(k, f_1), \dots, M_{DIR}(k, f_F)$ および前記予測行列 $A(k, f_1), \dots, A(k, f_F)$ をサイド情報源符号化器32においてエンコードする段階s32と、前記知覚的エンコーダ31および前記サイド情報源符号化器32の出力を多重化s33して、エンコードされたHOA信号フレーム

30

【数 1 9 5】

$$\tilde{B}(k-1)$$

を得る段階とを含む。

【0 1 3 3】

ある実施形態では、所与の数の係数シーケンスを有する入力HOA信号のフレームをフレームごとにエンコードする装置であって、各係数シーケンスがインデックスをもつ、装置が、プロセッサと、該プロセッサによって実行されたときに、該プロセッサに請求項8の段階を実行させる命令を記憶しているメモリとを有する。

40

【0 1 3 4】

図14は、ある実施形態におけるデコード方法のフローチャートを示している。低ビットレート圧縮されたHOA表現をデコードするための本方法は、前記圧縮されたHOA表現から、複数の打ち切りされたHOA係数シーケンス

【数 1 9 6】

$$\hat{z}_1(k), \dots, \hat{z}_I(k),$$

前記打ち切りされたHOA係数シーケンスのシーケンス・インデックスを示すまたは含む割

50

り当てベクトル $v_{\text{AMB}, \text{ASSIGN}}(k)$ 、サブバンドに関係した方向情報 $M_{\text{DIR}}(k+1, f_1), \dots, M_{\text{DIR}}(k+1, f_F)$ 、複数の予測行列 $A(k+1, f_1), \dots, A(k+1, f_F)$ および利得制御サイド情報 $e_1(k), \dots, e_I(k)$ 、 $\hat{c}_I(k)$ を抽出する段階と；
前記複数の打ち切りされたHOA係数シーケンス
【数 1 9 7】

$$\hat{z}_1(k), \dots, \hat{z}_I(k),$$

前記利得制御サイド情報 $e_1(k), \dots, e_I(k)$ 、 $\hat{c}_I(k)$ および前記割り当てベクトル $v_{\text{AMB}, \text{ASSIGN}}(k)$ から打ち切りされたHOA表現
【数 1 9 8】

$$\hat{c}_T(k)$$

を再構成する段階s51、s52と；
分解フィルタバンク 5 3 において、前記再構成された打ち切りされたHOA表現
【数 1 9 9】

$$\hat{c}_T(k)$$

を複数F個の周波数サブバンドについての周波数サブバンド表現
【数 2 0 0】

$$\hat{\tilde{c}}_T(k, f_1), \dots, \hat{\tilde{c}}_T(k, f_F)$$

に分解する段階s53と；
各周波数サブバンド表現について方向性サブバンド合成ブロック 5 4 において、前記再構成された打ち切りされたHOA表現の前記それぞれの周波数サブバンド表現
【数 2 0 1】

$$\hat{\tilde{c}}_T(k, f_1), \dots, \hat{\tilde{c}}_T(k, f_F)$$

と、前記サブバンドに関係した方向情報 $M_{\text{DIR}}(k+1, f_1), \dots, M_{\text{DIR}}(k+1, f_F)$ と、前記予測行列 $A(k+1, f_1), \dots, A(k+1, f_F)$ とから、予測された方向性HOA表現
【数 2 0 2】

$$\hat{\tilde{c}}_D(k, f_1), \dots, \hat{\tilde{c}}_D(k, f_F)$$

を合成する段階s54と；
前記F個の周波数サブバンドのそれぞれについてサブバンド組成ブロック 5 5 において、
係数シーケンス
【数 2 0 3】

$$\hat{\tilde{c}}_n(k, f_j), n = 1, \dots, O$$

をもつデコードされたサブバンドHOA表現

10

20

30

40

【数 2 0 4】

$$\tilde{\mathcal{C}}(k, f_1), \dots, \tilde{\mathcal{C}}(k, f_F)$$

を組成する段階s55であって、該係数シーケンスは、該係数シーケンスが前記割り当てベクトル $v_{AMB, ASSIGN}(k)$ に含まれるインデックス n をもつ場合には前記打ち切りされたHOA表現

【数 2 0 5】

$$\tilde{\mathcal{C}}_T(k, f_j)$$

10

の係数シーケンスから得られ、それ以外の場合には前記方向性サブバンド合成ブロック 54 の一つによって与えられる前記予測された方向性HOA成分

【数 2 0 6】

$$\tilde{\mathcal{C}}_D(k, f_j)$$

20

の係数シーケンスから得られる、段階と；

合成フィルタバンク 56 において、前記デコードされたサブバンドHOA表現

【数 2 0 7】

$$\tilde{\mathcal{C}}(k, f_1), \dots, \tilde{\mathcal{C}}(k, f_F)$$

を合成して、デコードされたHOA表現

【数 2 0 8】

$$\hat{\mathcal{C}}(k)$$

30

を得る段階とを含む。

【0 1 3 5】

ある実施形態では、前記抽出する段階は、前記圧縮されたHOA表現を多重分離s41して、知覚的に符号化された部分とエンコードされたサイド情報部分とを取得し、前記エンコードされた打ち切りされたHOA係数シーケンスを知覚的にデコードs42し、サイド情報源デコーダ 43 において、前記エンコードされたサイド情報をデコードs43することを含む。ある実施形態では、前記複数の打ち切りされたHOAシーケンスから打ち切りされたHOA表現

【数 2 0 9】

40

$$\hat{\mathcal{C}}_T(k)$$

を再構成することは、逆利得制御s51を実行することおよび前記打ち切りされたHOA表現

【数 2 1 0】

$$\hat{\mathcal{C}}_T(k)$$

50

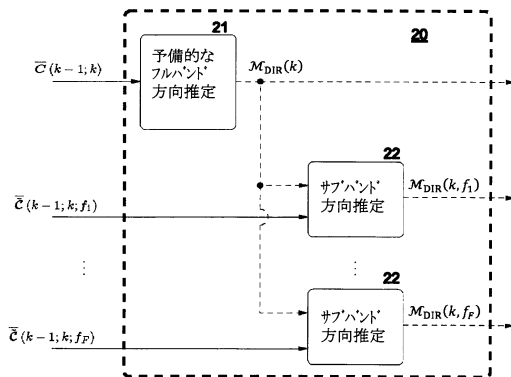
【 0 1 3 6 】

【 0 1 3 7 】

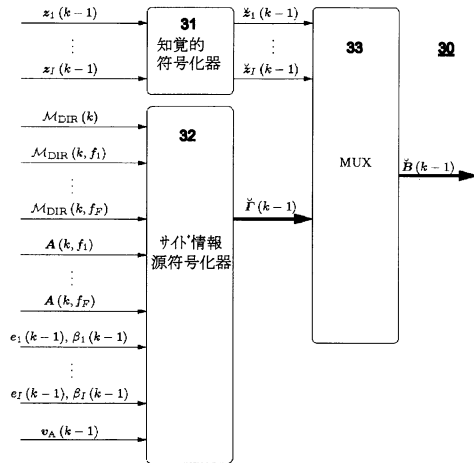
【 0 1 3 8 】

10

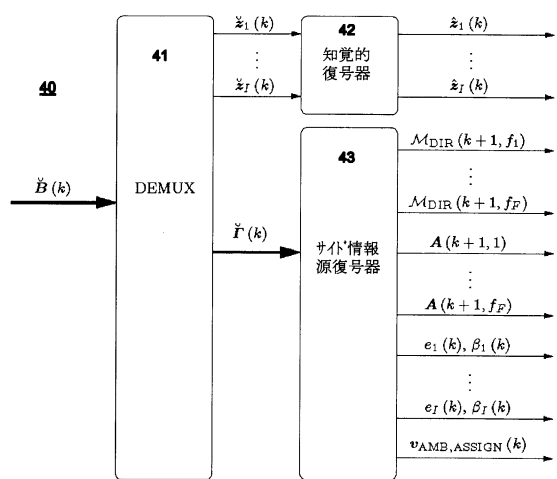
【圖 2】



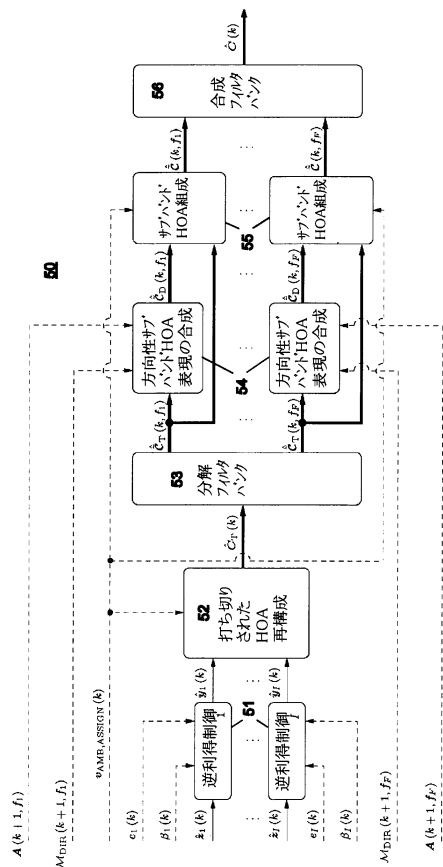
【図 3】



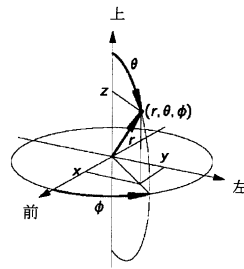
【図 4】



【図 5】



【図 6】



【図 7】

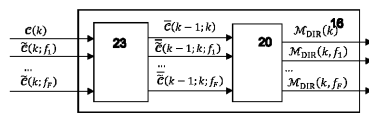


Fig.7

【図 8】

$$\mathbf{M}_{\text{mm}}(\mathbf{k}) = \{\Omega_{23}, \Omega_{28}, \Omega_{52}, \Omega_{101}, \Omega_{229}, \Omega_{446}, \Omega_{581}\}$$

Γ_1	x	x	x	x
Γ_2	x	x	x	x
Γ_3		x	x	
Γ_4		x	x	
Γ_5		x	x	
Γ_6	x			
...				
	k-2	k-1	k	k+1

7V-ム

$$\mathbf{M}_{\text{mm}}(\mathbf{k}, \mathbf{f}_1) = \{(1, \Omega_{23}), (2, \Omega_{229}), (3, \Omega_{28}), (5, \Omega_{581})\}$$

$$\mathbf{M}_{\text{mm}}(\mathbf{k}, \mathbf{f}_2) = \{(1, \Omega_{23}), (2, \Omega_{229})\}$$

【図 9】

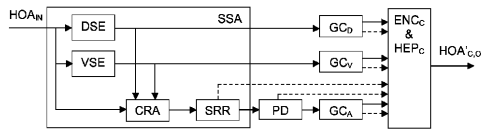


Fig.9

【図 10】

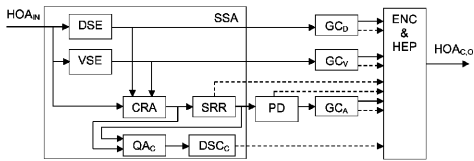


Fig.10

【図 11】

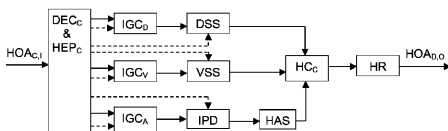


Fig.11

【図 12】

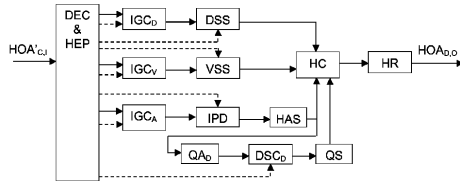


Fig.12

【図 13】

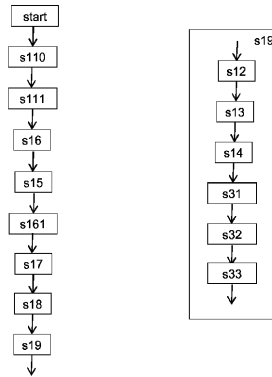


Fig.13

【図 14】

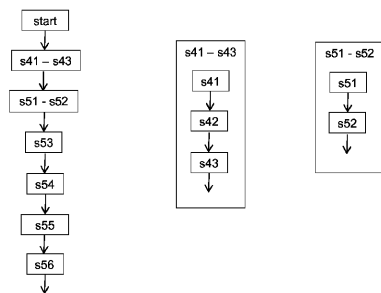


Fig.14

フロントページの続き

- (72)発明者 クルーガー, アレクサンダー
ドイツ連邦共和国, 3 0 6 5 5 ハノーヴァー, フォイエルバッハシュトラッセ 1 6
- (72)発明者 コルドン, スヴェン
ドイツ連邦共和国, 3 1 5 1 5 ヴンストルフ, ミューレンカンブシュトラッセ 5 0 アー

審査官 山下 剛史

- (56)参考文献 欧州特許出願公開第2 6 6 5 2 0 8 (E P , A 1)
欧州特許出願公開第2 7 4 3 9 2 2 (E P , A 1)
岩谷幸雄他, 球面調和解析による音場表現, 日本音響学会誌, 2 0 1 1 年 1 1 月, 67巻, 11号,
p.544-549
Erik HELLERUD, et al., SPATIAL REDUNDANCY IN HIGHER ORDER AMBISONICS AND ITS USE FOR L
OW DELAY LOSSLESS COMPRESSION, ICASSP 2009, IEEE, 2 0 0 9 年 4 月, p.269-272
- (58)調査した分野(Int.Cl., D B 名)
G 1 0 L 1 9 / 0 0 - 1 9 / 2 6
H 0 4 S 1 / 0 0 - 7 / 0 0