

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2017-199034

(P2017-199034A)

(43) 公開日 平成29年11月2日(2017.11.2)

(51) Int.Cl.		F I				テーマコード (参考)
G 1 O L	19/008	(2013.01)	G 1 O L	19/008	1 0 0	5 D 1 6 2
H O 4 S	7/00	(2006.01)	H O 4 S	7/00	3 0 0	
H O 4 S	5/00	(2006.01)	H O 4 S	5/00	5 0 0	

審査請求 有 請求項の数 15 O L 外国語出願 (全 50 頁)

(21) 出願番号	特願2017-152964 (P2017-152964)	(71) 出願人	510185767
(22) 出願日	平成29年8月8日 (2017.8.8)		ドルビー・インターナショナル・アーベー
(62) 分割の表示	特願2016-513406 (P2016-513406)		オランダ王国, セーエン アムステルダム
原出願日	平成26年5月23日 (2014.5.23)		ズイドースト 1101, ヘリケルベル
(31) 優先権主張番号	61/827, 246		グウェグ 1-35, アポロ ビルディン
(32) 優先日	平成25年5月24日 (2013.5.24)	(74) 代理人	グ 3エー
(33) 優先権主張国	米国 (US)		100107766
(31) 優先権主張番号	61/893, 770		弁理士 伊東 忠重
(32) 優先日	平成25年10月21日 (2013.10.21)	(74) 代理人	100070150
(33) 優先権主張国	米国 (US)		弁理士 伊東 忠彦
(31) 優先権主張番号	61/973, 625	(74) 代理人	100091214
(32) 優先日	平成26年4月1日 (2014.4.1)		弁理士 大貫 進介
(33) 優先権主張国	米国 (US)		

最終頁に続く

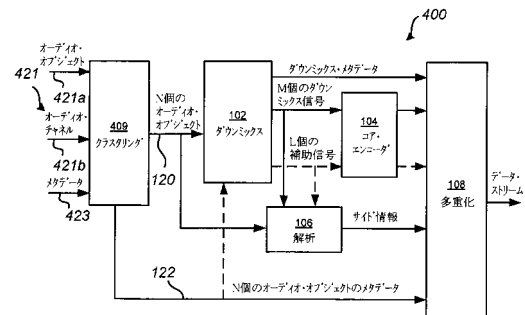
(54) 【発明の名称】 オーディオ・オブジェクトを含むオーディオ・シーンの効率的な符号化

(57) 【要約】

【課題】オブジェクト・ベース・オーディオのエンコードおよびデコードのためのエンコード方法およびデコード方法が提供される。

【解決手段】例示的なエンコード方法は、なかんづく、 M N であるとして、 N 個のオーディオ・オブジェクトの組み合わせを形成することによって M 個のダウンミックス信号を計算する段階と；前記 M 個のダウンミックス信号からの、前記 N 個のオーディオ・オブジェクトに基づいて形成された一組のオーディオ・オブジェクトの再構成を許容するパラメータを計算する段階とを含む。前記 M 個のダウンミックス信号の計算は、いかなるラウドスピーカー配位とも独立な基準に従ってなされる。

【選択図】図4



【特許請求の範囲】**【請求項 1】**

データ・ストリームに基づいてオーディオ・オブジェクトを再構成およびレンダリングする方法であって：

データ・ストリームを受領する段階であって、前記データ・ストリームは：

$N > 1$ および $M \geq N$ であるとして、 N 個のオーディオ・オブジェクトの組み合わせである M 個のダウンミックス信号を含む後方互換なダウンミックスと、

前記 M 個のダウンミックス信号からの、前記 N 個のオーディオ・オブジェクトの再構成を許容するパラメータを含む時間可変なサイド情報と、

前記 N 個のオーディオ・オブジェクトに関連付けられた複数のメタデータ・インスタンスであって、前記複数のメタデータ・インスタンスは前記 N 個のオーディオ・オブジェクトをレンダリングするためのそれぞれの所望されるレンダリング設定を指定する、メタデータ・インスタンス、および、各メタデータ・インスタンスについての、現在のレンダリング設定からそのメタデータ・インスタンスによって指定される所望されるレンダリング設定への補間の開始時刻および継続時間を含む遷移データとを含む、段階と；

前記後方互換なダウンミックスおよび前記サイド情報に基づいて、前記 N 個のオーディオ・オブジェクトを再構成する段階と；

前記 N 個のオーディオ・オブジェクトをあらかじめ定義されたチャネル配位の出力チャネルにレンダリングする段階とを含み、該レンダリングは：

現在のレンダリング設定に従ってレンダリングを実行し、

あるメタデータ・インスタンスについての前記遷移データによって定義される開始時刻において、前記現在のレンダリング設定から、そのメタデータ・インスタンスによって指定される所望されるレンダリング設定への補間を開始し、

そのメタデータ・インスタンスについての前記遷移データによって定義される継続時間後に前記所望されるレンダリング設定への前記補間を完了することによる、方法。

【請求項 2】

前記 N 個のオーディオ・オブジェクトに関連付けられた前記メタデータ・インスタンスは、前記オーディオ・オブジェクトの空間位置についての情報を含む、請求項 1 記載の方法。

【請求項 3】

前記 N 個のオーディオ・オブジェクトに関連付けられた前記メタデータ・インスタンスはさらに、オブジェクト・サイズ、オブジェクト・ラウドネス、オブジェクト重要性、オブジェクト・コンテンツ型およびゾーン・マスクのうちの一つまたは複数を含む、請求項 2 記載の方法。

【請求項 4】

前記複数のメタデータ・インスタンスに関連付けられた開始時刻は、フレーム境界のようなオーディオ・コンテンツに関係した時間イベントに対応する、請求項 1 ないし 3 のうちいずれか一項記載の方法。

【請求項 5】

前記現在のレンダリング設定から前記所望されるレンダリング設定への補間が線形補間である、請求項 1 ないし 4 のうちいずれか一項記載の方法。

【請求項 6】

前記データ・ストリームは、前記 N 個のオーディオ・オブジェクトを再構成するためのそれぞれの所望される再構成設定を指定する複数のサイド情報インスタンスと、各サイド情報インスタンスについての、現在の再構成設定から、そのサイド情報インスタンスによって指定される所望される再構成設定への補間を開始する時点および該補間を完了する時点を組み合わせて定義する、二つの独立に割り当て可能な部分を含む遷移データとを含み、前記 N 個のオーディオ・オブジェクトの再構成は：

現在の再構成設定に従って再構成を実行し、

あるサイド情報インスタンスについての前記遷移データによって定義される時点において、前記現在の再構成設定から、そのサイド情報インスタンスによって指定される所望される再構成設定への補間を開始し、

そのサイド情報インスタンスについての前記遷移データによって定義される時点において前記補間を完了することを含む、

請求項 1 ないし 5 のうちいずれか一項記載の方法。

【請求項 7】

データ・ストリームに基づいてオーディオ・オブジェクトを再構成およびレンダリングするシステムであって：

データ・ストリームを受領するよう構成された受領コンポーネントであって、前記データ・ストリームは：

$N > 1$ および $M \leq N$ であるとして、 N 個のオーディオ・オブジェクトの組み合わせである M 個のダウンミックス信号を含む後方互換なダウンミックスと、

前記 M 個のダウンミックス信号からの、前記 N 個のオーディオ・オブジェクトの再構成を許容するパラメータを含む時間可変なサイド情報と、

前記 N 個のオーディオ・オブジェクトに関連付けられた複数のメタデータ・インスタンスであって、前記複数のメタデータ・インスタンスは前記 N 個のオーディオ・オブジェクトをレンダリングするためのそれぞれの所望されるレンダリング設定を指定する、メタデータ・インスタンス、および、各メタデータ・インスタンスについての、現在のレンダリング設定からそのメタデータ・インスタンスによって指定される所望されるレンダリング設定への補間の開始時刻および継続時間を含む遷移データとを含む、受領コンポーネントと；

前記後方互換なダウンミックスおよび前記サイド情報に基づいて、前記 N 個のオーディオ・オブジェクトを再構成するよう構成された再構成コンポーネントと；

前記 N 個のオーディオ・オブジェクトをあらかじめ定義されたチャネル配位の出力チャネルにレンダリングするよう構成されたレンダラーとを有しており、該レンダリングは：現在のレンダリング設定に従ってレンダリングを実行し、

あるメタデータ・インスタンスについての前記遷移データによって定義される開始時刻において、前記現在のレンダリング設定から、そのメタデータ・インスタンスによって指定される所望されるレンダリング設定への補間を開始し、

そのメタデータ・インスタンスについての前記遷移データによって定義される継続時間後に前記所望されるレンダリング設定への前記補間を完了することによる、システム。

【請求項 8】

N 個のオーディオ・オブジェクトに関連付けられたメタデータを表現するデータ構造であって：

前記 N 個のオーディオ・オブジェクトをレンダリングするためのそれぞれの所望されるレンダリング設定を指定する複数のメタデータ・インスタンスと、

各メタデータ・インスタンスに関連付けられた遷移データであって、現在のレンダリング設定からそのメタデータ・インスタンスによって指定される所望されるレンダリング設定への補間の開始時刻および継続時間を含む遷移データとを含む、データ構造。

【請求項 9】

オーディオ・オブジェクトをデータ・ストリームとしてエンコードする方法であって：

$N > 1$ であるとして、 N 個のオーディオ・オブジェクトおよび前記 N 個のオーディオ・オブジェクトがデコーダ側での再生のためにどのようにレンダリングされるべきかを記述する前記 N 個のオーディオ・オブジェクトに関連付けられた時間可変のメタデータを受領する段階と；

$M \leq N$ であるとして、前記 N 個のオーディオ・オブジェクトの組み合わせを形成することによって、 M 個のダウンミックス信号を含む後方互換なダウンミックスを計算する段階と

;

前記M個のダウンミックス信号からの、前記N個のオーディオ・オブジェクトの再構成を許容するパラメータを含む時間可変のサイド情報を計算する段階と;

前記後方互換なダウンミックスおよび前記サイド情報を、デコーダに伝送するためのデータ・ストリームに含める段階と;

前記データ・ストリーム中にさらに:

前記N個のオーディオ・オブジェクトをレンダリングするためのそれぞれの所望されるレンダリング設定を指定する複数のメタデータ・インスタンスと、

各メタデータ・インスタンスについて、現在のレンダリング設定からそのメタデータ・インスタンスによって指定される所望されるレンダリング設定への補間の開始時刻および継続時間を含む遷移データとを含める段階とを含む、
方法。

10

【請求項 10】

前記N個のオーディオ・オブジェクトに関連付けられた前記メタデータは、前記オーディオ・オブジェクトの空間位置についての情報を含む、請求項9記載の方法。

【請求項 11】

前記N個のオーディオ・オブジェクトに関連付けられた前記メタデータはさらに、オブジェクト・サイズ、オブジェクト・ラウドネス、オブジェクト重要性、オブジェクト・コンテンツ型およびゾーン・マスクのうちの一つまたは複数を含む、請求項10記載の方法。

20

【請求項 12】

前記現在のレンダリング設定から前記所望されるレンダリング設定への補間が線形補間である、請求項9ないし11のうちいずれか一項記載の方法。

【請求項 13】

前記データ・ストリーム中にさらに:

前記N個のオーディオ・オブジェクトを再構成するためのそれぞれの所望される再構成設定を指定する複数のサイド情報インスタンスと、

各サイド情報インスタンスについての、現在の再構成設定から、そのサイド情報インスタンスによって指定される所望される再構成設定への補間を開始する時点および該補間を完了する時点を組み合わせて定義する、二つの独立に割り当て可能な部分を含む遷移データとを含める段階を含む、

30

請求項9ないし12のうちいずれか一項記載の方法。

【請求項 14】

オーディオ・オブジェクトをデータ・ストリームとしてエンコードするエンコーダであって:

$N > 1$ であるとして、N個のオーディオ・オブジェクトおよび前記N個のオーディオ・オブジェクトがデコーダ側での再生のためにどのようにレンダリングされるべきかを記述する前記N個のオーディオ・オブジェクトに関連付けられた時間可変のメタデータを受領するよう構成された受領器と;

M = Nであるとして、前記N個のオーディオ・オブジェクトの組み合わせを形成することによって、M個のダウンミックス信号を含む後方互換なダウンミックスを計算するよう構成されたダウンミックス・コンポーネントと;

40

前記M個のダウンミックス信号からの、前記N個のオーディオ・オブジェクトの再構成を許容するパラメータを含む時間可変のサイド情報を計算するよう構成された解析コンポーネントと;

前記後方互換なダウンミックスおよび前記サイド情報を、デコーダに伝送するためのデータ・ストリームに含めるよう構成された多重化コンポーネントとを有しており、

前記多重化コンポーネントは、前記データ・ストリーム中にさらに:

前記N個のオーディオ・オブジェクトをレンダリングするためのそれぞれの所望されるレンダリング設定を指定する複数のメタデータ・インスタンスと、

50

各メタデータ・インスタンスについて、現在のレンダリング設定からそのメタデータ・インスタンスによって指定される所望されるレンダリング設定への補間の開始時刻および継続時間を含む遷移データとを含めるよう構成されている、エンコーダ。

【請求項 15】

請求項 1 ないし 6 のうちいずれか一項記載の方法または請求項 9 ないし 13 のうちいずれか一項記載の方法をコンピュータに実行させるためのコンピュータ・プログラム。

【発明の詳細な説明】

【技術分野】

【0001】

10

関連出願への相互参照

本願は2013年5月24日に出願された米国仮特許出願第61/827,246号、2013年10月21日に出願された米国仮特許出願第61/827,246号、2014年4月1日に出願された米国仮特許出願第61/973,623号の出願日の利益を主張するものであり、そのそれぞれはここにその全体において参照によって組み込まれる。

【0002】

技術分野

本開示は概括的にはオーディオ・オブジェクトを含むオーディオ・シーンの符号化に関する。詳細には、オーディオ・オブジェクトのエンコードおよびデコードのためのエンコーダ、デコーダおよび関連する方法に関する。

20

【背景技術】

【0003】

オーディオ・シーンは一般にオーディオ・オブジェクトおよびオーディオ・チャンネルを含む。オーディオ・オブジェクトは、時間とともに変わりうる関連付けられた空間位置をもつオーディオ信号である。オーディオ・チャンネルは、三つの前方スピーカー、二つのサラウンド・スピーカーおよび低域効果スピーカーをもついわゆる5.1スピーカー配位のようなマルチチャンネル・スピーカー配位のチャンネルに直接対応するオーディオ信号である。

【0004】

オーディオ・オブジェクトの数は典型的には非常に多い、たとえば数百個のオーディオ・オブジェクトのオーダーであることがあるので、オーディオ・オブジェクトがデコーダ側で効率的に再構築されることを許容する符号化方法が必要とされている。オーディオ・オブジェクトをエンコーダ側でマルチチャンネル・ダウンミックスに（すなわち、5.1配位のようなある種のマルチチャンネル・スピーカー配位のチャンネルに対応する複数のオーディオ・チャンネルに）組み合わせて、デコーダ側で該マルチチャンネル・ダウンミックスからパラメトリックにオーディオ・オブジェクトを再構成する提案がされている。

30

【0005】

そのようなアプローチの利点は、オーディオ・オブジェクト再構成をサポートしないレガシー・デコーダがマルチチャンネル・ダウンミックスを直接、マルチチャンネル・スピーカー配位での再生のために使用しうるということである。例として、5.1ダウンミックスは、5.1配位のラウドスピーカーで直接再生されうる。

40

【発明の概要】

【発明が解決しようとする課題】

【0006】

だが、このアプローチの欠点は、マルチチャンネル・ダウンミックスがデコーダ側でオーディオ・オブジェクトの十分良好な再構成を与えないことがあるということである。たとえば、5.1配位の左前方スピーカーと同じ水平位置をもつが垂直位置は異なる二つのオーディオ・オブジェクトを考える。これらのオブジェクトは典型的には5.1ダウンミックスの同じチャンネルに組み合わされる。これは、同じダウンミックス・チャンネルから二つのオーディオ・オブジェクトの近似を再構成する必要があるデコーダ側でのオーディオ・オブジェクト再構成にとっては困難な状況をなす。このプロセスは完璧な再構成を保証できず

50

、時には耳に聞こえるアーチファクトにつながることもさへある。

【 0 0 0 7 】

このように、オーディオ・オブジェクトの効率的かつ改善された再構成を提供するエンコード/デコード方法が必要とされている。

【 0 0 0 8 】

たとえばダウンミックスからのオーディオ・オブジェクトの再構成の際、しばしばサイド情報またはメタデータが用いられる。そのようなサイド情報の形および内容はたとえば再構成されるオーディオ・オブジェクトの忠実さおよび/または再構成の実行の計算量に影響する。したがって、再構成されるオーディオ・オブジェクトの忠実さを高めることを許容するおよび/または再構成の計算量を軽減することを許容する新しい、代替的なサイド情報フォーマットをもつエンコード/デコード方法を提供することが望ましい。

10

【図面の簡単な説明】

【 0 0 0 9 】

ここで例示的な実施形態について付属の図面を参照して述べる。

【図 1】例示的な実施形態に基づくエンコーダの概略図である。

【図 2】例示的な実施形態に基づくオーディオ・オブジェクトの再構成をサポートするデコーダの概略図である。

【図 3】例示的な実施形態に基づくオーディオ・オブジェクトの再構成をサポートしない低計算量デコーダの概略図である。

【図 4】例示的な実施形態に基づくオーディオ・シーンの単純化のための逐次的に配列されたクラスタリング・コンポーネントを有するエンコーダの概略図である。

20

【図 5】例示的な実施形態に基づくオーディオ・シーンの単純化のための並列に配置されたクラスタリング・コンポーネントを有するエンコーダの概略図である。

【図 6】メタデータ・インスタンスのセットのためのレンダリング行列を計算する典型的な既知のプロセスを示す図である。

【図 7】オーディオ信号のレンダリングにおいて用いられる係数曲線の導出を示す図である。

【図 8】ある例示的な実施形態に基づくメタデータ・インスタンス補間方法を示す図である。

【図 9】例示的な実施形態に基づく追加的なメタデータ・インスタンスの導入の例を示す図である。

30

【図 10】例示的な実施形態に基づく追加的なメタデータ・インスタンスの導入の例を示す図である。

【図 11】ある例示的な実施形態に基づく、低域通過フィルタをもつサンプルホールド回路を使う補間方法を示す図である。すべての図面は概略的であり、一般に、本開示を明快にするために必要な部分を示すのみである。一方、他の部分は省略されたり示唆されるだけであつたりすることがある。特に断わりのない限り、同様の参照符号は異なる図面における同様の部分を指す。

【発明を実施するための形態】

【 0 0 1 0 】

40

上記に鑑み、オーディオ・オブジェクトの効率的な改善された再構成を許容するおよび/または再構成されるオーディオ・オブジェクトの忠実さを高めることを許容するおよび/または再構成の計算量を軽減することを許容するエンコーダ、デコーダおよび関連する方法を提供することが目的である。

【 0 0 1 1 】

I . 概観——エンコーダ

第一の側面によれば、オーディオ・オブジェクトをエンコードするためのエンコード方法、エンコーダおよびコンピュータ・プログラム・プロダクトが提供される。

【 0 0 1 2 】

例示的な実施形態によれば、オーディオ・オブジェクトをデータ・ストリームにエンコ

50

ードする方法であって：

$N > 1$ であるとして、 N 個のオーディオ・オブジェクトを受領する段階と；

$M = N$ であるとして、いかなるラウドスピーカー配位とも独立な基準に従って前記 N 個のオーディオ・オブジェクトの組み合わせを形成することによって、 M 個のダウンミックス信号を計算する段階と；

前記 M 個のダウンミックス信号からの、前記 N 個のオーディオ・オブジェクトに基づいて形成された一組のオーディオ・オブジェクトの再構成を許容するパラメータを含むサイド情報を計算する段階と；

前記 M 個のダウンミックス信号および前記サイド情報を、デコーダに伝送するためのデータ・ストリームに含める段階とを含む、方法が提供される。

10

【0013】

上記の構成では、 M 個のダウンミックス信号は、いかなるラウドスピーカー配位とも独立に N 個のオーディオ・オブジェクトから形成される。これは、 M 個のダウンミックス信号が、 M 個のチャンネルをもつスピーカー配位のチャンネルでの再生に好適なオーディオ信号に制約されないことを含意する。その代わり、 M 個のダウンミックス信号は、たとえば N 個のオーディオ・オブジェクトのダイナミクスに適応し、デコーダ側でのオーディオ・オブジェクトの再構成を改善するようある基準に従ってより自由に選択されうる。

【0014】

5.1配位の左前方スピーカーと同じ水平位置をもつが垂直位置は異なる二つのオーディオ・オブジェクトをもつ例に戻ると、提案される方法は、第一のオーディオ・オブジェクトを第一のダウンミックス信号に入れ、第二のオーディオ・オブジェクトを第二のダウンミックス信号に入れることを許容する。これは、デコーダにおけるオーディオ・オブジェクトの完璧な再構成を可能にする。一般に、そのような完璧な再構成は、アクティブなオーディオ・オブジェクトの数がダウンミックス信号の数を超えない限り、可能である。アクティブなオーディオ・オブジェクトの数のほうが多い場合には、提案される方法は、デコーダにおいて再構成されたオーディオ・オブジェクトにおいて生じる可能な近似誤差が、再構成されたオーディオ・シーンに対して全く知覚的影響を与えないまたは可能な最小の知覚的影響を与えるように、同じダウンミックス信号に混合される必要のあるオーディオ・オブジェクトの選択を許容する。

20

【0015】

M 個のダウンミックス信号が適応的であることの第二の利点は、ある種のオーディオ・オブジェクトが他のオーディオ・オブジェクトと厳密に別個のままにしておくことができることである。たとえば、ダイアログが空間的属性に関して正確にレンダリングされ、デコーダにおけるオブジェクト処理を許容することを保証するために、任意のダイアログ・オブジェクトをバックグラウンド・オブジェクトと別個にしておくことが有利であることがある。オブジェクト処理はたとえば、了解性を改善するための、ダイアログ向上またはダイアログ・ラウドネスの増大である。他の応用（たとえばカラオケ）では、一つまたは複数のオブジェクトの完全なミュートを許容することが有利でありうる。これも、そのようなオブジェクトが他のオブジェクトと混合されないことを要求する。特定のスピーカー配位に対応するマルチチャンネル・ダウンミックスを使う通常の方法は、他のオーディオ・オブジェクトとの混合中に存在するオーディオ・オブジェクトの完全なミュートは許容しない。

30

40

【0016】

ダウンミックス信号という語は、ダウンミックス信号が他の信号の混合、すなわち組み合わせであることを反映している。単語「ダウン」は、ダウンミックス信号の数 M が典型的にはオーディオ・オブジェクトの数 N より少ないことを示す。

【0017】

例示的な実施形態によれば、本方法は、さらに、各ダウンミックス信号を空間位置に関連付け、ダウンミックス信号の前記空間位置を前記データ・ストリーム中に前記ダウンミックス信号のためのメタデータとして含めることを含んでいてもよい。これは、レガシー

50

再生システムの場合に低計算量のデコードが使用されることを許容するので有利である。より精密には、ダウンミックス信号に関連付けられたメタデータは、デコーダ側で、ダウンミックス信号を、レガシー再生システムのチャンネルにレンダリングするために使われてもよい。

【0018】

例示的な実施形態によれば、N個のオーディオ・オブジェクトは、該N個のオーディオ・オブジェクトの空間位置を含むメタデータに関連付けられており、ダウンミックス信号に関連付けられる空間位置はN個のオーディオ・オブジェクトの空間位置に基づいて計算される。よって、ダウンミックス信号は、N個のオーディオ・オブジェクトの空間位置に依存する空間位置をもつオーディオ・オブジェクトとして解釈されてもよい。

10

【0019】

さらに、N個のオーディオ・オブジェクトの空間位置およびM個のダウンミックス信号に関連付けられた空間位置は時間変化してもよい。すなわち、オーディオ・フレームの時間フレームの間で変動してもよい。換言すれば、ダウンミックス信号は、時間フレームの間で変化する関連付けられた位置をもつ動的オーディオ・オブジェクトと解釈されてもよい。これは、ダウンミックス信号が固定した空間的なラウドスピーカー位置に対応する従来技術のシステムと対照的である。

【0020】

典型的には、サイド情報も時間変化し、それによりオーディオ・オブジェクトの再構成を支配するパラメータが時間的に変化することを許容する。

20

【0021】

エンコーダは、ダウンミックス信号の計算について異なる基準を適用してもよい。N個のオーディオ・オブジェクトがN個のオーディオ・オブジェクトの空間位置を含むメタデータに関連付けられる例示的な実施形態によれば、M個のダウンミックス信号を計算するための前記基準は、N個のオーディオ・オブジェクトの空間的近接性に基づいていてもよい。たとえば、互いに近接しているオーディオ・オブジェクトは同じダウンミックス信号に組み合わされてもよい。

【0022】

N個のオーディオ・オブジェクトに関連付けられたメタデータがさらに、N個のオブジェクトの互いとの関係での重要性を示す重要性値を含む例示的な実施形態によれば、M個のダウンミックス信号を計算するための前記基準はさらに、N個のオーディオ・オブジェクトの前記重要性値に基づいていてもよい。たとえば、N個のオーディオ・オブジェクトのうち最も重要なもの（単数または複数）はダウンミックス信号に直接マッピングされてもよく、一方、残りのオーディオ・オブジェクトは残りのダウンミックス信号を形成するために組み合わせられる。

30

【0023】

特に、例示的な実施形態によれば、M個のダウンミックス信号を計算する段階は、N個のオーディオ・オブジェクトの空間的近接性およびもし該当すれば重要性値に基づいてN個のオーディオ・オブジェクトをM個のクラスターと関連付け、各クラスターについてのダウンミックス信号を、そのクラスターに関連付けられたオーディオ・オブジェクトの組み合わせを形成することによって計算することを含む第一のクラスタリング手順を含む。いくつかの場合には、オーディオ・オブジェクトは高々一つのクラスターの一部をなしうる。他の場合には、オーディオ・オブジェクトはいくつかのクラスターの一部をなしてもよい。このようにして、オーディオ・オブジェクトから種々のグループ、すなわちクラスターが形成される。各クラスターは、ダウンミックス信号によって表わされることができ、該ダウンミックス信号はオーディオ・オブジェクトと考えられてもよい。このクラスタリング・アプローチは、各ダウンミックス信号を、そのダウンミックス信号に対応するクラスターに関連付けられたオーディオ・オブジェクトの空間位置に基づいて計算される空間位置に関連付けることを許容する。この解釈では、第一のクラスタリング手順は、このように、N個のオーディオ・オブジェクトの次元をM個のオーディオ・オブジェクトに、柔軟

40

50

な仕方で低減する。

【0024】

各ダウンミックス信号に関連付けられた空間位置はたとえば、そのダウンミックス信号に対応するクラスターに関連付けられたオーディオ・オブジェクトの空間位置の重心または重み付けされた重心として計算されてもよい。重みはたとえば、オーディオ・オブジェクトの重要性値に基づいていてもよい。

【0025】

例示的な実施形態によれば、N個のオーディオ・オブジェクトは、N個のオーディオ・オブジェクトの空間位置を入力としてもつK平均アルゴリズムを適用することによって、M個のクラスターに関連付けられる。

【0026】

オーディオ・シーンは、膨大な数のオーディオ・オブジェクトを有することがあるので、本方法は、オーディオ・シーンの次元を減らし、それによりオーディオ・オブジェクトを再構成するときのデコーダ側での計算量を低減するためのさらなる策を講じてもよい。特に、本方法はさらに、第一の複数のオーディオ・オブジェクトを第二の複数のオーディオ・オブジェクトに減らすための第二のクラスタリング手順を有していてもよい。

【0027】

ある実施形態によれば、第二のクラスタリング手順は、M個のダウンミックス信号の計算に先立って実行される。よって、その実施形態では、第一の複数のオーディオ・オブジェクトはオーディオ・シーンのもとのオーディオ・オブジェクトに対応し、第二の、低減した複数のオーディオ・オブジェクトは、M個のダウンミックス信号が計算されるベースとなるN個のオーディオ・オブジェクトに対応する。さらに、そのような実施形態では、N個のオーディオ・オブジェクトに基づいて形成された（デコーダにおいて再構成されるべき）前記一組のオーディオ・オブジェクトは、N個のオーディオ・オブジェクトに対応する、すなわち等しい。

【0028】

もう一つの実施形態によれば、第二のクラスタリング手順は、M個のダウンミックス信号の計算と並列に実行される。そのような実施形態では、M個のダウンミックス信号が計算されるベースとなるN個のオーディオ・オブジェクトも第二のクラスタリング手順に入力される前記第一の複数のオーディオ・オブジェクトも、オーディオ・シーンのもとのオーディオ・オブジェクトに対応する。さらに、そのような実施形態では、N個のオーディオ・オブジェクトに基づいて形成された前記一組のオーディオ・オブジェクト（デコーダにおいて再構成されるもの）は、前記第二の複数のオーディオ・オブジェクトに対応する。よって、このアプローチでは、M個のダウンミックス信号は、低減した数のオーディオ・オブジェクトに基づいてではなく、オーディオ・シーンのもとのオーディオ・オブジェクトに基づいて計算される。

【0029】

例示的な実施形態によれば、第二のクラスタリング手順は：

前記第一の複数のオーディオ・オブジェクトおよびそれらの関連付けられた空間位置を受領する段階と；

前記第一の複数のオーディオ・オブジェクトを、前記第一の複数のオーディオ・オブジェクトの空間的近接性に基づいて少なくとも一つのクラスターと関連付ける段階と；

前記少なくとも一つのクラスターのそれぞれを、そのクラスターに関連付けられたオーディオ・オブジェクトの組み合わせであるオーディオ・オブジェクトによって表わすことによって、前記第二の複数のオーディオ・オブジェクトを生成する段階と；

前記第二の複数のオーディオ・オブジェクトについての空間位置を含むメタデータを計算する段階であって、前記第二の複数のオーディオ・オブジェクトの各オーディオ・オブジェクトの空間位置が、対応するクラスターに関連付けられたオーディオ・オブジェクトの空間位置に基づいて計算される、段階と；

前記第二の複数のオーディオ・オブジェクトについてのメタデータを前記データ・スト

10

20

30

40

50

リーム中に含める段階とを含む。

【 0 0 3 0 】

換言すれば、第二のクラスタリング手順は、等しいまたは非常に近い位置をもつオブジェクトのような、オーディオ・シーンに存在する空間的な冗長性を活用する。さらに、前記第二の複数のオーディオ・オブジェクトを生成するときに、オーディオ・オブジェクトの重要性値が考慮に入れられてもよい。

【 0 0 3 1 】

上述したように、オーディオ・シーンはオーディオ・チャンネルを含んでいてもよい。そのようなオーディオ・チャンネルは、静的な位置、すなわちオーディオ・チャンネルに対応するラウドスピーカーの位置に関連付けられたオーディオ・オブジェクトと考えられてもよい。より詳細には、第二のクラスタリング手順はさらに：

少なくとも一つのオーディオ・チャンネルを受領する段階と；

前記少なくとも一つのオーディオ・チャンネルのそれぞれを、そのオーディオ・チャンネルのラウドスピーカー位置に対応する静的な空間位置をもつオーディオ・オブジェクトに変換する段階と；

変換された少なくとも一つのオーディオ・チャンネルを前記第一の複数のオーディオ・オブジェクトに含める段階とを含んでいてもよい。

【 0 0 3 2 】

このようにして、本方法は、オーディオ・チャンネルもオーディオ・オブジェクトも含むオーディオ・シーンのエンコードを許容する。

【 0 0 3 3 】

例示的な実施形態によれば、例示的な実施形態に基づくデコード方法を実行するための命令をもつコンピュータ可読媒体を有するコンピュータ・プログラム・プロダクトが提供される。

【 0 0 3 4 】

例示的な実施形態によれば、オーディオ・オブジェクトをデータ・ストリームにエンコードするためのエンコーダであって：

$N > 1$ であるとして、 N 個のオーディオ・オブジェクトを受領するよう構成された受領コンポーネントと；

$M \geq N$ であるとして、いかなるラウドスピーカー配位とも独立な基準に従って前記 N 個のオーディオ・オブジェクトの組み合わせを形成することによって、 M 個のダウンミックス信号を計算するよう構成されたダウンミックス・コンポーネントと；

前記 M 個のダウンミックス信号からの、前記 N 個のオーディオ・オブジェクトに基づいて形成された一組のオーディオ・オブジェクトの再構成を許容するパラメータを含むサイド情報を計算するよう構成された解析コンポーネントと；

前記 M 個のダウンミックス信号および前記サイド情報を、デコーダに伝送するためのデータ・ストリームに含めるよう構成された多重化コンポーネントとを有する、エンコーダが提供される。

【 0 0 3 5 】

II. 概観——デコーダ

第二の側面によれば、マルチチャンネル・オーディオ・コンテンツをデコードするためのデコード方法、デコーダおよびコンピュータ・プログラム・プロダクトが提供される。

【 0 0 3 6 】

第二の側面は、一般に第一の側面と同じ特徴および利点をもちうる。

【 0 0 3 7 】

例示的な実施形態によれば、エンコードされたオーディオ・オブジェクトを含むデータ・ストリームをデコードするためのデコーダにおける方法であって：

$M \geq N$ であるとして、いかなるラウドスピーカー配位とも独立な基準に従って計算された N 個のオーディオ・オブジェクトの組み合わせである M 個のダウンミックス信号と、前記 M 個のダウンミックス信号からの、前記 N 個のオーディオ・オブジェクトに基づいて形成さ

10

20

30

40

50

れた一組のオーディオ・オブジェクトの再構成を許容するパラメータを含むサイド情報とを含むデータ・ストリームを受領する段階と；

前記M個のダウンミックス信号および前記サイド情報から、前記N個のオーディオ・オブジェクトに基づいて形成された一組のオーディオ・オブジェクトを再構成する段階とを含む、方法が提供される。

【0038】

例示的な実施形態によれば、前記データ・ストリームはさらに、前記M個のダウンミックス信号に関連付けられた空間位置を含む前記M個のダウンミックス信号についてのメタデータを含み、当該方法はさらに：

前記デコーダがオーディオ・オブジェクト再構成をサポートするよう構成されている場合には、前記M個のダウンミックス信号および前記サイド情報から、前記N個のオーディオ・オブジェクトに基づいて形成された一組のオーディオ・オブジェクトを再構成する前記段階を実行し、

前記デコーダがオーディオ・オブジェクト再構成をサポートするよう構成されていない場合には、前記M個のダウンミックス信号についての前記メタデータを、前記M個のダウンミックス信号を再生システムの出力チャネルにレンダリングするために使うことを含む。

【0039】

例示的な実施形態によれば、前記M個のダウンミックス信号に関連付けられた空間位置は時間変化する。

【0040】

例示的な実施形態によれば、前記サイド情報は時間変化する。

【0041】

例示的な実施形態によれば、前記データ・ストリームはさらに、前記N個のオーディオ・オブジェクトに基づいて形成された一組のオーディオ・オブジェクトの空間位置を含む、前記N個のオーディオ・オブジェクトに基づいて形成された一組のオーディオ・オブジェクトについてのメタデータを含み、当該方法はさらに、

前記N個のオーディオ・オブジェクトに基づいて形成された前記一組のオーディオ・オブジェクトについての前記メタデータを、前記N個のオーディオ・オブジェクトに基づいて形成された再構成された一組のオーディオ・オブジェクトを、再生システムの出力チャネルにレンダリングするために使うことを含む。

【0042】

例示的な実施形態によれば、前記N個のオーディオ・オブジェクトに基づいて形成された前記一組のオーディオ・オブジェクトは、前記N個のオーディオ・オブジェクトに等しい。

【0043】

例示的な実施形態によれば、前記N個のオーディオ・オブジェクトに基づいて形成された前記一組のオーディオ・オブジェクトは、前記N個のオーディオ・オブジェクトの組み合わせである複数のオーディオ・オブジェクトを含み、その数はNより少ない。

【0044】

例示的な実施形態によれば、例示的な実施形態に基づくデコード方法を実行するための命令をもつコンピュータ可読媒体を有するコンピュータ・プログラム・プロダクトが提供される。

【0045】

例示的な実施形態によれば、エンコードされたオーディオ・オブジェクトを含むデータ・ストリームをデコードするためのデコーダであって：

M Nであるとして、いかなるラウドスピーカー配位とも独立な基準に従って計算されたN個のオーディオ・オブジェクトの組み合わせであるM個のダウンミックス信号と、前記M個のダウンミックス信号からの、前記N個のオーディオ・オブジェクトに基づいて形成された一組のオーディオ・オブジェクトの再構成を許容するパラメータを含むサイド情報とを含むデータ・ストリームを受領するよう構成された受領コンポーネントと；

前記M個のダウンミックス信号および前記サイド情報から、前記N個のオーディオ・オブジェクトに基づいて形成された一組のオーディオ・オブジェクトを再構成するよう構成された再構成コンポーネントとを有する、デコーダが提供される。

【0046】

III. 概観——サイド情報およびメタデータについてのフォーマット

第三の側面によれば、オーディオ・オブジェクトをエンコードするためのエンコード方法、エンコーダおよびコンピュータ・プログラム・プロダクトが提供される。

【0047】

第三の側面に基づく方法、エンコーダおよびコンピュータ・プログラム・プロダクトは一般に、第一の側面に基づく方法、エンコーダおよびコンピュータ・プログラム・プロダクトと共通する特徴および利点をもちうる。

10

【0048】

例示的な実施形態によれば、オーディオ・オブジェクトをデータ・ストリームとしてエンコードする方法であって：

N>1であるとして、N個のオーディオ・オブジェクトを受領する段階と；

M=Nであるとして、前記N個のオーディオ・オブジェクトの組み合わせを形成することによって、M個のダウンミックス信号を計算する段階と；

前記M個のダウンミックス信号からの、前記N個のオーディオ・オブジェクトに基づいて形成された一組のオーディオ・オブジェクトの再構成を許容するパラメータを含む時間可変のサイド情報を計算する段階と；

20

前記M個のダウンミックス信号および前記サイド情報を、デコーダに伝送するためのデータ・ストリームに含める段階とを含む、方法が提供される。

【0049】

本例示的な実施形態では、本方法はさらに、前記データ・ストリーム中に：

前記N個のオーディオ・オブジェクトに基づいて形成された前記一組のオーディオ・オブジェクトを再構成するためのそれぞれの所望される再構成設定を指定する複数のサイド情報インスタンスと；

各サイド情報インスタンスについての、現在の再構成設定から、そのサイド情報インスタンスによって指定される所望される再構成設定への遷移を開始する時点および該遷移を完了する時点を組み合わさって定義する、二つの独立に割り当て可能な部分を含む遷移データとを含めることを含む。

30

【0050】

本例示的な実施形態では、前記サイド情報は時間可変である、たとえば時間変化し、オーディオ・オブジェクトの前記再構成を支配するパラメータが時間的に変化することを許容する。このことは、前記の諸サイド情報インスタンスの存在によって反映される。現在の再構成設定からそれぞれの所望される再構成設定への遷移を開始する時点および完了する時点とを定義する遷移データを含むサイド情報フォーマットを用いることによって、現在の再構成設定と単一のサイド情報インスタンスによって指定される単一の所望される再構成設定とに基づいて、すなわち他のいかなるサイド情報インスタンスの知識もなしに、補間が実行されうるという意味で、サイド情報インスタンスどうしが互いにより独立にされる。

40

【0051】

したがって、提供されるサイド情報フォーマットは、既存のサイド情報インスタンスどうしの間の追加的なサイド情報インスタンスの計算／導入を容易にする。特に、提供されるサイド情報フォーマットは、再生品質に影響することなく、追加的なサイド情報インスタンスの計算／導入を許容する。本開示では、既存のサイド情報インスタンスどうしの間の新たなサイド情報インスタンスを計算／導入するプロセスは、サイド情報の「再サンプリング」と称される。サイド情報の再サンプリングはしばしば、ある種のオーディオ処理タスクの間に必要とされる。たとえば、オーディオ・コンテンツがたとえばカット／マージ／混合によって編集されるとき、そのような編集はサイド情報インスタンスの間で生起

50

することがある。この場合、サイド情報の再サンプリングが必要とされることがある。もう一つのそのような事例は、オーディオ信号および関連付けられたサイド情報がフレーム・ベースのオーディオ・コーデックを用いてエンコードされるときである。この場合、伝送の際のフレーム損失の耐性を改善するために、各オーディオ・コーデック・フレームについて、好ましくはそのコーデック・フレームの先頭のタイムスタンプをもつ、少なくとも一つのサイド情報インスタンスをもつことが望ましい。たとえば、オーディオ信号/オブジェクトは、ビデオ・コンテンツを含むオーディオビジュアル信号またはマルチメディア信号の一部であってもよい。そのような用途では、オーディオ・コンテンツのフレーム・レートを、ビデオ・コンテンツのフレーム・レートに一致するように修正することが望ましいことがある。それにより、サイド情報の対応する再サンプリングが望ましいことがありうる。

10

【0052】

ダウンミックス信号およびサイド情報が含まれるデータ・ストリームは、たとえば、ビットストリーム、特に記憶されるまたは伝送されるビットストリームであってもよい。

【0053】

前記N個のオーディオ・オブジェクトの組み合わせを形成することによって、M個のダウンミックス信号を計算するというのは、M個のダウンミックス信号のそれぞれが、前記N個のオーディオ・オブジェクトの一つまたは複数のオーディオ内容の組み合わせ、たとえば線形結合を形成することによって得られることを意味する。換言すれば、N個のオーディオ・オブジェクトのそれぞれが必ずM個のダウンミックス信号のそれぞれに寄与する必要はない。

20

【0054】

ダウンミックス信号という語は、ダウンミックス信号が他の信号の混合、すなわち組み合わせであることを反映している。ダウンミックス信号はたとえば、他の信号の加法的混合であってもよい。単語「ダウン」は、ダウンミックス信号の数Mが典型的にはオーディオ・オブジェクトの数Nより少ないことを示す。

【0055】

ダウンミックス信号は、たとえば、前記第一の側面内の例示的实施形態の任意のものに基づいて、いかなるラウドスピーカー配位とも独立な基準に従って前記N個のオーディオ信号の組み合わせを形成することによって計算されてもよい。あるいはまた、ダウンミックス信号は、ダウンミックス信号がM個のチャンネルをもつスピーカー配位のチャンネル上での再生のために好適であるように前記N個のオーディオ信号の組み合わせを形成することによって計算されてもよい。この後者は本稿では後方互換なダウンミックスと称される。

30

【0056】

二つの独立に割り当て可能な部分を含む遷移データというののは、それら二つの部分が互いに独立に割り当て可能である、すなわち互いとは独立に割り当てられてもよいことが意図されている。しかしながら、遷移データのそれらの部分がたとえばメタデータのサイド情報の他の型についての遷移データの部分と一致してもよいことは理解される。

【0057】

本例示的实施形態では、遷移データの二つの独立に割り当て可能な部分は、組み合わせられて、遷移を開始する時点および遷移を完了する時点を定義する。すなわち、これら二つの時点は、遷移データの前記二つの独立に割り当て可能な部分から導出可能である。

40

【0058】

ある例示的实施形態によれば、本方法はさらに、第一の複数のオーディオ・オブジェクトを第二の複数のオーディオ・オブジェクトに減らすためのクラスタリング手順を含んでもよい。ここで、前記N個のオーディオ・オブジェクトは、前記第一の複数のオーディオ・オブジェクトまたは前記第二の複数のオーディオ・オブジェクトをなす。前記N個のオーディオ・オブジェクトに基づいて形成された前記一組のオーディオ・オブジェクトは前記第二の複数のオーディオ・オブジェクトに一致する。本例示的实施形態では、クラスタリング手順は：

50

前記第二の複数のオーディオ・オブジェクトについての空間位置を含む時間可変なクラスター・メタデータを計算する段階と；

前記デコーダに伝送するために、前記データ・ストリーム中に；

第二の組のオーディオ・オブジェクトをレンダリングするためのそれぞれの所望されるレンダリング設定を指定する複数のクラスター・メタデータ・インスタンスと；

各クラスター・メタデータ・インスタンスについての、現在のレンダリング設定から、そのクラスター・メタデータ・インスタンスによって指定される所望されるレンダリング設定への遷移を開始する時点およびそのクラスター・メタデータ・インスタンスによって指定される前記所望されるレンダリング設定への前記遷移を完了する時点とを、組み合わせさせて定義する、二つの独立に割り当て可能な部分を含む遷移データとを含める段階とを含む。

10

【0059】

オーディオ・シーンは、膨大な数のオーディオ・オブジェクトを有することがあるので、本方法は、前記第一の複数のオーディオ・オブジェクトを第二の複数のオーディオ・オブジェクトに減らすことによってオーディオ・シーンの次元を減らすためのさらなる策を講じてよい。本例示的实施形態では、前記N個のオーディオ・オブジェクトに基づいて形成され、前記ダウンミックス信号および前記サイド情報に基づいてデコーダ側で再構成される前記一組のオーディオ・オブジェクトは、前記第二の複数のオーディオ・オブジェクトと一致する。該第二の複数のオーディオ・オブジェクトは、前記第一の複数のオーディオ信号によって表わされるオーディオ・シーンの簡略化および/またはより低次元の表現に対応し、デコーダ側での再構成のための計算量が軽減される。

20

【0060】

前記クラスター・メタデータを前記データ・ストリームに含めることは、たとえば前記ダウンミックス信号および前記サイド情報に基づいて前記第二の組のオーディオ信号が再構成されたあとに、デコーダ側での前記第二の組のオーディオ信号のレンダリングを許容する。

【0061】

前記サイド情報と同様に、本例示的实施形態における前記クラスター・メタデータは時間可変である、たとえば時間変化し、前記第二の複数のオーディオ・オブジェクトの前記レンダリングを支配するパラメータが時間的に変化することを許容する。前記ダウンミックス・メタデータについてのフォーマットは、前記サイド情報についてと同様であってもよく、同じまたは対応する利点を有していてもよい。特に、本例示的实施形態において提供されるクラスター・メタデータの形は、クラスター・メタデータの再サンプリングを容易にする。クラスター・メタデータの再サンプリングは、たとえば、前記クラスター・メタデータおよび前記サイド情報に関連するそれぞれの遷移を開始および完了する共通の時点を提供するために、および/または前記クラスター・メタデータを、関連付けられたオーディオ信号のフレーム・レートに調整するために用いられてもよい。

30

【0062】

ある例示的实施形態によれば、前記クラスタリング手順はさらに；

前記第一の複数のオーディオ・オブジェクトおよびそれらの関連付けられた空間位置を受領する段階と；

40

前記第一の複数のオーディオ・オブジェクトを、前記第一の複数のオーディオ・オブジェクトの空間的近接性に基づいて少なくとも一つのクラスターと関連付ける段階と；

前記少なくとも一つのクラスターのそれぞれを、そのクラスターに関連付けられたオーディオ・オブジェクトの組み合わせであるオーディオ・オブジェクトによって表わすことによって、前記第二の複数のオーディオ・オブジェクトを生成する段階と；

前記第二の複数のオーディオ・オブジェクトの各オーディオ・オブジェクトの空間位置を、それぞれのクラスターに関連付けられた、すなわちそのオーディオ・オブジェクトが代表するクラスターに関連付けられた諸オーディオ・オブジェクトの空間位置に基づいて計算する段階とを含む。

50

【 0 0 6 3 】

換言すれば、このクラスタリング手順は、等しいまたは非常に近い位置をもつオブジェクトのような、オーディオ・シーンに存在する空間的な冗長性を活用する。さらに、前記第二の複数のオーディオ・オブジェクトを生成するときに、オーディオ・オブジェクトの重要性値が考慮に入れられてもよい。これについては上記第一の側面における例示的实施形態との関連で述べた。

【 0 0 6 4 】

前記第一の複数のオーディオ・オブジェクトを少なくとも一つのクラスターと関連付けることは、前記第一の複数のオーディオ・オブジェクトのそれぞれを前記少なくとも一つのクラスターのうちの一つまたは複数と関連付けることを含む。いくつかの場合には、オーディオ・オブジェクトは高々一つのクラスターの一部をなしうる。他の場合には、オーディオ・オブジェクトはクラスタリング手順の一部としていくつかのクラスターの間で分割されてもよい。

10

【 0 0 6 5 】

前記第一の複数のオーディオ・オブジェクトの空間的近接性は、前記第一の複数のオーディオ・オブジェクトのうちのそれぞれのオーディオ・オブジェクトの間の距離および／またはそれらの相対位置に関係していてもよい。たとえば、互いに近いオーディオ・オブジェクトは、同じクラスターと関連付けられてもよい。

【 0 0 6 6 】

オーディオ・オブジェクトが、そのクラスターに関連付けられたオーディオ・オブジェクトの組み合わせであるとは、そのオーディオ・オブジェクトに関連付けられたオーディオ・コンテンツ／信号が、そのクラスターに関連付けられたそれぞれのオーディオ・オブジェクトに関連付けられたオーディオ・コンテンツ／信号の組み合わせとして形成されることを意味する。

20

【 0 0 6 7 】

ある例示的な実施形態によれば、それぞれのクラスター・メタデータ・インスタンスのために遷移データによって定義されるそれぞれの時点は、対応するサイド情報インスタンスのために遷移データによって定義されるそれぞれの時点と一致してもよい。

【 0 0 6 8 】

サイド情報およびクラスター・メタデータに関連付けられる遷移を開始および完了するのに同じ時点を用いることにより、サイド情報およびクラスター・メタデータの、合同再サンプリングのような合同処理が容易にされる。

30

【 0 0 6 9 】

さらに、サイド情報およびクラスター・メタデータに関連付けられる遷移を開始および完了するのに同じ時点を使うことは、デコーダ側での合同再構成およびレンダリングを容易にする。たとえば、再構成およびレンダリングがデコーダ側で合同動作として実行される場合、再構成およびレンダリングのための合同設定が、それぞれのサイド情報インスタンスおよびメタデータ・インスタンスについて決定されうるおよび／または再構成およびレンダリングのための合同設定の間の補間が、それぞれの設定について別個に補間を実行する代わりに用いられうる。そのような合同補間は、補間する必要のある係数／パラメータが少なくなるので、デコーダ側での計算量を減らしうる。

40

【 0 0 7 0 】

ある例示的な実施形態によれば、クラスタリング手順は、M個のダウンミックス信号の計算に先立って実行されてもよい。本例示的な実施形態では、前記第一の複数のオーディオ・オブジェクトはオーディオ・シーンのもとのオーディオ・オブジェクトに対応し、前記M個のダウンミックス信号が計算されるベースとなるN個のオーディオ・オブジェクトは前記第二の、低減した複数のオーディオ・オブジェクトをなす。よって、本例示的な実施形態では、N個のオーディオ・オブジェクトに基づいて形成された前記一組のオーディオ・オブジェクト（デコーダで再構成されるもの）は、前記N個のオーディオ・オブジェクトと一致する。

50

【 0 0 7 1 】

あるいはまた、クラスタリング手順は、M個のダウンミックス信号の計算と並列に実行されてもよい。本代替的实施形態によれば、M個のダウンミックス信号が計算されるベースとなるN個のオーディオ・オブジェクトは、オーディオ・シーンのもとのオーディオ・オブジェクトに対応する前記第一の複数のオーディオ・オブジェクトをなす。よって、このアプローチでは、M個のダウンミックス信号は、低減した数のオーディオ・オブジェクトに基づいてではなく、オーディオ・シーンのもとのオーディオ・オブジェクトに基づいて計算される。

【 0 0 7 2 】

ある例示的な実施形態によれば、本方法はさらに：

10

各ダウンミックス信号を、前記ダウンミックス信号をレンダリングするための時間可変な空間位置に関連付ける段階と；

前記ダウンミックス信号の前記空間位置を含むダウンミックス・メタデータを前記データ・ストリームにさらに含める段階とを含み、

当該方法はさらに、前記データ・ストリームに：

前記ダウンミックス信号をレンダリングするためのそれぞれの所望されるダウンミックス・レンダリング設定を指定する複数のダウンミックス・メタデータ・インスタンスと；

各ダウンミックス・メタデータ・インスタンスについての、現在のダウンミックス・レンダリング設定から、そのダウンミックス・メタデータ・インスタンスによって指定される所望されるダウンミックス・レンダリング設定への遷移を開始する時点および前記ダウンミックス・メタデータ・インスタンスによって指定される所望されるダウンミックス・レンダリング設定への遷移を完了する時点を組み合わせさせて定義する、二つの独立に割り当て可能な部分を含む遷移データとを含めることを含む。

20

【 0 0 7 3 】

ダウンミックス・メタデータを前記データ・ストリームに含めることは、レガシー再生設備の場合に低計算量のデコードが使用されることを許容する点で有利である。より正確には、ダウンミックス・メタデータは、ダウンミックス信号をレガシー再生システムのチャンネルにレンダリングするために、デコーダ側で使用されうる。すなわち、典型的には計算的により複雑な動作であるN個のオブジェクトに基づいて形成された前記複数のオーディオ・オブジェクトを再構成することがない。

30

【 0 0 7 4 】

本例示的な実施形態によれば、M個のダウンミックス信号に関連付けられた空間位置は時間可変である、たとえば時間変化してもよく、ダウンミックス信号は、時間フレームまたはダウンミックス・メタデータ・インスタンスの間で変化する関連付けられた位置をもつ動的オーディオ・オブジェクトと解釈されてもよい。これは、ダウンミックス信号が固定した空間的なラウドスピーカー位置に対応する従来技術のシステムと対照的である。より進化した機能をもつデコード・システムでは、同じデータ・ストリームが、オブジェクト指向の仕方で再生されることが想起される。

【 0 0 7 5 】

いくつかの例示的な実施形態では、N個のオーディオ・オブジェクトは、N個のオーディオ・オブジェクトの空間位置を含むメタデータに関連付けられていてもよく、ダウンミックス信号に関連付けられた空間位置はたとえば、N個のオーディオ・オブジェクトの空間位置に基づいて計算されてもよい。このように、ダウンミックス信号は、N個のオーディオ・オブジェクトの空間位置に依存する空間位置をもつオーディオ・オブジェクトと解釈されてもよい。

40

【 0 0 7 6 】

ある例示的な実施形態によれば、それぞれのダウンミックス・メタデータ・インスタンスについての遷移データによって定義されるそれぞれの時点は、対応するサイド情報インスタンスについての遷移データによって定義されるそれぞれの時点と一致してもよい。サイド情報およびダウンミックス・メタデータに関連付けられる遷移を開始および完了するの

50

に同じ時点を用いることにより、サイド情報およびダウンミックス・メタデータの合同処理、たとえば再サンプリングが容易になる。

【0077】

ある例示的实施形態によれば、それぞれのダウンミックス・メタデータ・インスタンスについての遷移データによって定義されるそれぞれの時点は、対応するクラスター・メタデータ・インスタンスについての遷移データによって定義されるそれぞれの時点と一致してもよい。クラスター・メタデータおよびダウンミックス・メタデータに関連付けられる遷移を開始および終了するのに同じ時点を用いることにより、クラスター・メタデータおよびダウンミックス・メタデータの合同処理、たとえば再サンプリングが容易になる。

【0078】

例示的な実施形態によれば、 $N > 1$ として、 N 個のオーディオ・オブジェクトをデータ・ストリームとしてエンコードするエンコーダが提供される。本エンコーダは：

M N であるとして、前記 N 個のオーディオ・オブジェクトの組み合わせを形成することによって、 M 個のダウンミックス信号を計算するよう構成されているダウンミックス・コンポーネントと；

前記 M 個のダウンミックス信号からの、前記 N 個のオーディオ・オブジェクトに基づいて形成された一組のオーディオ・オブジェクトの再構成を許容するパラメータを含む時間可変のサイド情報を計算するよう構成されている解析コンポーネントと；

前記 M 個のダウンミックス信号および前記サイド情報を、デコーダに伝送するためのデータ・ストリームに含めるよう構成された多重化コンポーネントとを有しており、

前記多重化コンポーネントは、前記デコーダへの伝送のために、前記データ・ストリーム中に：

前記 N 個のオーディオ・オブジェクトに基づいて形成された前記一組のオーディオ・オブジェクトを再構成するためのそれぞれの所望される再構成設定を指定する複数のサイド情報インスタンスと；

各サイド情報インスタンスについての、現在の再構成設定から、そのサイド情報インスタンスによって指定される所望される再構成設定への遷移を開始する時点および該遷移を完了する時点とを組み合わせることで定義する、二つの独立に割り当て可能な部分を含む遷移データとを含めるよう構成される。

【0079】

第四の側面によれば、マルチチャネル・オーディオ・コンテンツをデコードするためのデコード方法、デコーダおよびコンピュータ・プログラム・プロダクトが提供される。

【0080】

第四の側面に基づく方法、デコーダおよびコンピュータ・プログラム・プロダクトは、第三の側面に基づく方法、エンコーダおよびコンピュータ・プログラム・プロダクトと協働することが意図されており、対応する特徴および利点をもちうる。

【0081】

第四の側面に基づく方法、デコーダおよびコンピュータ・プログラム・プロダクトは、一般に、第二の側面に基づく方法、デコーダおよびコンピュータ・プログラム・プロダクトと共通の特徴および利点をもちうる。

【0082】

例示的な実施形態によれば、データ・ストリームに基づいてオーディオ・オブジェクトを再構成する方法であって：

$N > 1$ および M N であるとして、 N 個のオーディオ・オブジェクトの組み合わせである M 個のダウンミックス信号と、前記 M 個のダウンミックス信号からの、前記 N 個のオーディオ・オブジェクトに基づいて形成された一組のオーディオ・オブジェクトの再構成を許容するパラメータを含む時間可変なサイド情報とを含むデータ・ストリームを受領する段階と；

前記 M 個のダウンミックス信号および前記サイド情報に基づいて、前記 N 個のオーディオ・オブジェクトに基づいて形成された一組のオーディオ・オブジェクトを再構成する段階とを含み、

10

20

30

40

50

前記データ・ストリームは、複数のサイド情報インスタンスを含み、前記データ・ストリームはさらに、各サイド情報インスタンスについて、現在の再構成設定から、そのサイド情報インスタンスによって指定される所望される再構成設定への遷移を開始する時点および該遷移を完了する時点を組み合わせて定義する、二つの独立に割り当て可能な部分を含む遷移データを含み、

前記N個のオーディオ・オブジェクトに基づいて形成された一組のオーディオ・オブジェクトを再構成する段階は：

現在の再構成設定に従って再構成を実行し；

あるサイド情報インスタンスについての遷移データによって定義される時点において、前記現在の再構成設定から、そのサイド情報インスタンスによって指定される所望される再構成設定への遷移を開始し；

そのサイド情報インスタンスについての前記遷移データによって定義される時点において前記遷移を完了することを含む。

【0083】

上記のように、現在の再構成設定からそれぞれの所望される再構成設定への遷移を開始する時点および完了する時点を定義する遷移データを含むサイド情報フォーマットを用いることにより、たとえばサイド情報の再サンプリングが容易になる。

【0084】

データ・ストリームは、たとえば、エンコーダ側で生成されたビットストリームの形で受領されてもよい。

【0085】

前記M個のダウンミックス信号および前記サイド情報に基づいて、前記N個のオーディオ・オブジェクトに基づいて形成された一組のオーディオ・オブジェクトを再構成する段階は、たとえば、サイド情報に基づいて決定される係数を用いてダウンミックス信号の少なくとも一つの線形結合を形成することを含んでいてもよい。前記M個のダウンミックス信号および前記サイド情報に基づいて、前記N個のオーディオ・オブジェクトに基づいて形成された一組のオーディオ・オブジェクトを再構成する段階は、たとえば、サイド情報に基づいて決定される係数を用いて、ダウンミックス信号および任意的には該ダウンミックス信号から導出される一つまたは複数の追加的な（たとえば脱相関された）信号の少なくとも一つの線形結合を形成することを含んでいてもよい。

【0086】

ある例示的实施形態によれば、データ・ストリームはさらに、前記N個のオーディオ・オブジェクトに基づいて形成された前記一組のオーディオ・オブジェクトについての時間可変なクラスター・メタデータを含んでいてもよい。クラスター・メタデータは、前記N個のオーディオ・オブジェクトに基づいて形成された前記一組のオーディオ・オブジェクトについての空間位置を含む。データ・ストリームは、複数のクラスター・メタデータ・インスタンスを含んでいてもよく、データ・ストリームはさらに、各クラスター・メタデータ・インスタンスについて、現在のレンダリング設定から、そのクラスター・メタデータ・インスタンスによって指定される所望されるレンダリング設定への遷移を開始する時点およびそのクラスター・メタデータ・インスタンスによって指定される前記所望されるレンダリング設定への前記遷移を完了する時点と、組み合わせさせて定義する、二つの独立に割り当て可能な部分を含む遷移データを含んでいてもよい。

【0087】

本方法はさらに：

前記N個のオーディオ・オブジェクトに基づいて形成された再構成された一組のオーディオ・オブジェクトを、あらかじめ定義されたチャネル配位の出力チャネルにレンダリングするために前記クラスター・メタデータを使うことを含んでいてもよく、該レンダリングは：

現在のレンダリング設定に従ってレンダリングを実行し；

あるクラスター・メタデータ・インスタンスについての遷移データによって定義される

10

20

30

40

50

時点において、前記現在のレンダリング設定から、そのクラスター・メタデータ・インスタンスによって指定される所望されるレンダリング設定への遷移を開始し；

そのクラスター・メタデータ・インスタンスについての前記遷移データによって定義される時点において前記所望されるレンダリング設定への前記遷移を完了することを含む。

【0088】

前記あらかじめ定義されたチャンネル配位は、たとえば、特定の再生システムと互換な、すなわち特定の再生システムでの再生に好適な、出力チャンネルの配位に対応してもよい。

【0089】

前記N個のオーディオ・オブジェクトに基づいて形成された再構成された一組のオーディオ・オブジェクトを、あらかじめ定義されたチャンネル配位の出力チャンネルにレンダリングすることは、たとえば、レンダラーにおいて、前記N個のオーディオ・オブジェクトに基づいて形成された再構成された一組のオーディオ・オブジェクトを、クラスター・メタデータの制御のもとでレンダラーの出力チャンネル（のあらかじめ定義された配位）にマッピングすることを含んでいてもよい。

【0090】

前記N個のオーディオ・オブジェクトに基づいて形成された再構成された一組のオーディオ・オブジェクトをレンダリングすることは、たとえば、クラスター・メタデータに基づいて決定される係数を用いて、前記N個のオーディオ・オブジェクトに基づいて形成された再構成された一組のオーディオ・オブジェクトの線形結合を形成することを含んでいてもよい。

【0091】

ある例示的实施形態によれば、それぞれのクラスター・メタデータ・インスタンスについての遷移データによって定義されるそれぞれの時点は、対応するサイド情報インスタンスについて前記遷移データによって定義されるそれぞれの時点と一致してもよい。

【0092】

ある例示的实施形態によれば、本方法はさらに：

前記再構成の少なくとも一部および前記レンダリングの少なくとも一部を、それぞれ現在の再構成設定および現在のレンダリング設定に関連する再構成行列およびレンダリング行列の行列積として形成される第一の行列に対応する組み合わせられた動作として実行し；

あるサイド情報インスタンスおよびあるクラスター・メタデータ・インスタンスについての遷移データによって定義される時点において、前記現在の再構成およびレンダリング設定から、それぞれそのサイド情報インスタンスおよびそのクラスター・メタデータ・インスタンスによって指定される所望される再構成設定およびレンダリング設定への組み合わせられた遷移を開始し；

そのサイド情報インスタンスおよびそのクラスター・メタデータ・インスタンスについての前記遷移データによって定義される時点において前記組み合わせられた遷移を完了することを含み、前記組み合わせられた遷移は、前記第一の行列の行列要素と、それぞれ前記所望される再構成設定および前記所望されるレンダリング設定に関連する再構成行列およびレンダリング行列の行列積として形成される第二の行列の行列要素との間を補間することを含む。

【0093】

再構成設定およびレンダリング設定の別個の遷移ではなく、上記の意味での組み合わせられた遷移を実行することにより、より少数のパラメータ／係数を補間すればよくなり、計算量の低減が許容される。

【0094】

本例示的实施形態において言及されるような再構成行列またはレンダリング行列のような行列は、たとえば、単一行または単一系列からなってもよく、よってベクトルに対応していてもよいことは理解しておくものとする。

【0095】

ダウンミックス信号からのオーディオ・オブジェクトの再構成はしばしば、異なる周波

10

20

30

40

50

数帯域において異なる再構成行列を用いることによって実行される。一方、レンダリングはしばしば、すべての周波数について同じレンダリング行列を用いることによって実行される。そのような場合、再構成およびレンダリングの組み合わせられた動作に対応する行列、たとえば本例示的实施形態において言及した上記第一および第二の行列は、典型的には周波数依存であってもよい。すなわち、異なる周波数帯域については、行列要素の異なる値が典型的には用いられてもよい。

【0096】

ある例示的实施形態によれば、前記N個のオーディオ・オブジェクトに基づいて形成された前記一組のオーディオ・オブジェクトは、前記N個のオーディオ・オブジェクトと一致してもよい。すなわち、本方法は、前記M個のダウンミックス信号および前記サイド情報に基づいて前記N個のオーディオ・オブジェクトを再構成することを含んでいてもよい。

10

【0097】

あるいはまた、前記N個のオーディオ・オブジェクトに基づいて形成された前記一組のオーディオ・オブジェクトは、前記N個のオーディオ・オブジェクトの組み合わせである、N個より少ない複数のオーディオ・オブジェクトを含んでいてもよい。すなわち、本方法は、前記M個のダウンミックス信号および前記サイド情報に基づいて前記N個のオーディオ・オブジェクトのこれらの組み合わせを再構成することを含んでいてもよい。

【0098】

ある例示的实施形態によれば、データ・ストリームはさらに、前記M個のダウンミックス信号に関連付けられた時間可変な空間位置を含む前記M個のダウンミックス信号についてのダウンミックス・メタデータを含んでいてもよい。データ・ストリームは、複数のダウンミックス・メタデータ・インスタンスを含んでいてもよく、データ・ストリームはさらに、各ダウンミックス・メタデータ・インスタンスについて、現在のダウンミックス・レンダリング設定から、そのダウンミックス・メタデータ・インスタンスによって指定される所望されるダウンミックス・レンダリング設定への遷移を開始する時点およびそのダウンミックス・メタデータ・インスタンスによって指定される前記所望されるダウンミックス・レンダリング設定への前記遷移を完了する時点と、組み合わせさせて定義する、二つの独立に割り当て可能な部分を含む遷移データを含んでいてもよい。

20

【0099】

本方法はさらに：
デコーダがオーディオ・オブジェクト再構成をサポートするよう動作可能である（または構成されている）場合、再構成する段階を、前記M個のダウンミックス信号および前記サイド情報、前記N個のオーディオ・オブジェクトに基づいて形成された一組のオーディオ・オブジェクトに基づいて実行し；

30

デコーダがオーディオ・オブジェクト再構成をサポートするよう動作可能でない（または構成されていない）場合、前記ダウンミックス・メタデータおよび前記M個のダウンミックス信号を前記M個のダウンミックス信号のレンダリングのために出力することを含んでいてもよい。

【0100】

デコーダがオーディオ・オブジェクト再構成をサポートするよう動作可能であり、データ・ストリームがさらに、前記N個のオーディオ・オブジェクトに基づいて形成された前記一組のオーディオ・オブジェクトに関連付けられたクラスター・メタデータをさらに含む場合、デコーダはたとえば、再構成された一組のオーディオ・オブジェクトのレンダリングのために、再構成された一組のオーディオ・オブジェクト、クラスター・メタデータを出力してもよい。

40

【0101】

デコーダがオーディオ・オブジェクト再構成をサポートするよう動作可能でない場合には、たとえばサイド情報および該当するならクラスター・メタデータを破棄して、前記ダウンミックス・メタデータおよび前記M個のダウンミックス信号を出力として提供しても

50

よい。その場合、出力は、レンダラーによって、前記M個のダウンミックス信号をレンダラーの出力チャンネルにレンダリングするために用いられてもよい。

【0102】

任意的に、本方法はさらに、前記ダウンミックス・メタデータに基づいて、あらかじめ定義された出力配位の出力チャンネル、たとえばレンダラーの出力チャンネルに、あるいはデコーダの出力チャンネルに（デコーダがレンダリング機能をもつ場合）、前記M個のダウンミックス信号をレンダリングすることを含んでいてもよい。

【0103】

例示的实施形態によれば、データ・ストリームに基づいてオーディオ・オブジェクトを再構成するデコーダが提供される。本デコーダは：

N>1およびM Nであるとして、N個のオーディオ・オブジェクトの組み合わせであるM個のダウンミックス信号と、前記M個のダウンミックス信号からの、前記N個のオーディオ・オブジェクトに基づいて形成された一組のオーディオ・オブジェクトの再構成を許容するパラメータを含む時間可変なサイド情報とを含むデータ・ストリームを受領するよう構成されている受領コンポーネントと；

前記M個のダウンミックス信号および前記サイド情報に基づいて、前記N個のオーディオ・オブジェクトに基づいて形成された一組のオーディオ・オブジェクトを再構成するよう構成されている再構成コンポーネントとを有し、

前記データ・ストリームは、関連する複数のサイド情報インスタンスを含み、前記データ・ストリームはさらに、各サイド情報インスタンスについて、現在の再構成設定から、そのサイド情報インスタンスによって指定される所望される再構成設定への遷移を開始する時点および該遷移を完了する時点を組み合わさって定義する、二つの独立に割り当て可能な部分を含む遷移データを含む。

【0104】

再構成コンポーネントは、前記N個のオーディオ・オブジェクトに基づいて形成された一組のオーディオ・オブジェクトを再構成することを、少なくとも：

現在の再構成設定に従って再構成を実行し；

あるサイド情報インスタンスについての遷移データによって定義される時点において、前記現在の再構成設定から、そのサイド情報インスタンスによって指定される所望される再構成設定への遷移を開始し；

そのサイド情報インスタンスについての前記遷移データによって定義される時点において前記遷移を完了することによって行なうよう構成されている。

【0105】

ある例示的实施形態によれば、上記第三または第四の側面における方法はさらに、一つまたは複数の追加的なサイド情報インスタンスであって、該一つまたは複数の追加的なサイド情報インスタンスの直前または直後のサイド情報インスタンスと実質的に同じ再構成設定を指定するものを生成することを含んでいてもよい。同様の仕方で追加的なクラスター・メタデータ・インスタンスおよび/またはダウンミックス・メタデータ・インスタンスが生成される例示的实施形態も構想される。

【0106】

上記のように、より多くのサイド情報インスタンスを生成することによる前記サイド情報の再サンプリングは、いくつかの状況において有利でありうる。たとえばオーディオ信号/オブジェクトおよび関連するサイド情報がフレーム・ベースのオーディオ・コーデックを使ってエンコードされるとき、各オーディオ・コーデック・フレームについて少なくとも一つのサイド情報インスタンスをもつことが望ましいので、そのような構成が有利でありうる。エンコーダ側では、解析コンポーネントによって提供されるサイド情報インスタンスは、たとえば、ダウンミックス・コンポーネントによって提供されるダウンミックス信号のフレーム・レートと一致しないような仕方で時間的に分布していることがありうる。よって、サイド情報は、ダウンミックス信号の各フレームについて少なくとも一つのサイド情報インスタンスがあるよう新たなサイド情報インスタンスを導入することによ

10

20

30

40

50

て、有利に再サンプリングされうる。同様に、デコーダ側では、受領されたサイド情報インスタンスは、たとえば、受領されたダウンミックス信号のフレーム・レートと一致しないような仕方で時間的に分布していることがありうるので、サイド情報は、ダウンミックス信号の各フレームについて少なくとも一つのサイド情報インスタンスがあるような新たなサイド情報インスタンスを導入することによって、有利に再サンプリングされうる。

【0107】

追加的なサイド情報インスタンスは、たとえば、選択された時点について：該追加的なサイド情報インスタンスの直後のサイド情報インスタンスをコピーし、選択された時点および前記直後のサイド情報インスタンスについての遷移データによって定義される時点に基づいてその追加的なサイド情報インスタンスについての遷移データを決定することによって、生成されてもよい。

10

【0108】

第五の側面によれば、データ・ストリーム中でM個のオーディオ信号と一緒にエンコードされたサイド情報をトランスコードするための方法、装置およびコンピュータ・プログラム・プロダクトが提供される。

【0109】

第五の側面に基づく方法、装置およびコンピュータ・プログラム・プロダクトは、第三および第四の側面に基づく方法、エンコーダ、デコーダおよびコンピュータ・プログラム・プロダクトと協働することが意図されており、対応する特徴および利点をもちうる。

【0110】

20

例示的な実施形態によれば、データ・ストリームにおいてM個のオーディオ信号と一緒にエンコードされたサイド情報をトランスコードする方法が提供される。本方法は：

データ・ストリームを受領する段階と；

前記データ・ストリームから、M-1として、M個のオーディオ信号と、前記M個のオーディオ信号からの、一組のオーディオ・オブジェクトの再構成を許容するパラメータを含む、関連する時間可変なサイド情報とを抽出する段階であって、抽出されたサイド情報は：

前記オーディオ・オブジェクトを再構成するためのそれぞれの所望される再構成設定を指定する複数のサイド情報インスタンス；および

各サイド情報インスタンスについて、現在の再構成設定から、そのサイド情報インスタンスによって指定される所望される再構成設定への遷移を開始する時点および該遷移を完了する時点を組み合わせて定義する、二つの独立に割り当て可能な部分を含む遷移データを含む、段階と；

30

一つまたは複数の追加的なサイド情報インスタンスであって、該一つまたは複数の追加的なサイド情報インスタンスの直前または直後のサイド情報インスタンスと実質的に同じ再構成設定を指定する追加的なサイド情報インスタンスを生成する段階と；

前記M個のオーディオ信号および前記サイド情報をデータ・ストリームに含める段階とを含む。

【0111】

本例示的な実施形態では、前記一つまたは複数の追加的なサイド情報インスタンスは、前記サイド情報が受領されたデータ・ストリームから抽出された後に生成されてもよく、生成された一つまたは複数の追加的なサイド情報インスタンスは次いで、前記M個のオーディオ信号および他のサイド情報インスタンスと一緒にデータ・ストリーム中に含まれてもよい。

40

【0112】

第三の側面に関連して上記したように、より多くのサイド情報インスタンスを生成することによるサイド情報の再サンプリングは、いくつかの状況において有利でありうる。たとえばオーディオ信号/オブジェクトおよび関連するサイド情報がフレーム・ベースのオーディオ・コーデックを使ってエンコードされるとき、各オーディオ・コーデック・フレームについて少なくとも一つのサイド情報インスタンスをもつことが望ましいので、そのような構成が有利でありうる。

50

【0113】

データ・ストリームがさらに、第三および第四の側面との関係で述べたクラスター・メタデータおよび/またはダウンミックス・メタデータを含む実施形態も構想される。本方法はさらに、前記追加的なサイド情報インスタンスが生成される仕方と同様に、追加的なダウンミックス・メタデータ・インスタンスおよび/またはクラスター・メタデータ・インスタンスを生成することを含む。

【0114】

ある例示的实施形態によれば、前記M個のオーディオ信号は、第一のフレーム・レートに従って、受領されたデータ・ストリーム中で符号化されていてもよく、本方法はさらに：

10

前記M個のオーディオ信号を処理して、前記M個のダウンミックス信号が符号化されているフレーム・レートを前記第一のフレーム・レートとは異なる第二のフレーム・レートに変える段階と；

少なくとも、前記一つまたは複数の追加的なサイド情報インスタンスを生成することによって、前記第二のフレーム・レートに一致するおよび/または前記第二のフレーム・レートと互換になるよう、前記サイド情報を再サンプリングする段階とを含んでいてもよい。

【0115】

第三の側面との関係で上記したように、オーディオ信号を処理してオーディオ信号の符号化に使われるフレーム・レートを変える、たとえば修正されたフレーム・レートが、前記オーディオ信号が属するオーディオビジュアル信号のビデオ・コンテンツのフレーム・レートに一致するようにすることがいくつかの状況では有利でありうる。各サイド情報インスタンスについての遷移データの存在は、第三の側面との関係で上記したように、サイド情報の再サンプリングを容易にする。サイド情報は、たとえば、処理されたオーディオ信号の各フレームについて少なくとも一つのサイド情報インスタンスがあるよう、追加的なサイド情報インスタンスを生成することによって、新たなフレーム・レートに一致するよう再サンプリングされてもよい。

20

【0116】

例示的实施形態によれば、データ・ストリームにおいてM個のオーディオ信号と一緒にエンコードされたサイド情報をトランスコードする装置が提供される。

30

【0117】

本装置は：

データ・ストリームを受領し、前記データ・ストリームから、M-1として、M個のオーディオ信号と、前記M個のオーディオ信号からの、一組のオーディオ・オブジェクトの再構成を許容するパラメータを含む、関連する時間可変なサイド情報とを抽出するよう構成された受領コンポーネントを有しており、抽出されたサイド情報は：

前記オーディオ・オブジェクトを再構成するためのそれぞれの所望される再構成設定を指定する複数のサイド情報インスタンス；および

各サイド情報インスタンスについて、現在の再構成設定から、そのサイド情報インスタンスによって指定される所望される再構成設定への遷移を開始する時点および該遷移を完了する時点を組み合わせて定義する、二つの独立に割り当て可能な部分を含む遷移データを含む。

40

【0118】

本装置はさらに：

一つまたは複数の追加的なサイド情報インスタンスであって、該一つまたは複数の追加的なサイド情報インスタンスの直前または直後のサイド情報インスタンスと実質的に同じ再構成設定を指定する追加的なサイド情報インスタンスを生成するよう構成されている再サンプリング・コンポーネントと；

前記M個のオーディオ信号および前記サイド情報をデータ・ストリームに含めるよう構成されている多重化コンポーネントとを有する。

50

【 0 1 1 9 】

ある例示的实施形態によれば、第三、第四または第五の側面における方法は、さらに：第一のサイド情報インスタンスによって指定される第一の所望される再構成設定と、前記第一のサイド情報インスタンスの直後の一つまたは複数のサイド情報インスタンスによって指定される一つまたは複数の所望される再構成設定との間の差を計算し；計算された差があらかじめ定義された閾値より低いことに応答して前記一つまたは複数のサイド情報インスタンスを除去することを含んでいてもよい。クラスター・メタデータ・インスタンスおよび／またはダウンミックス・メタデータ・インスタンスが類似の仕方で除去される例示的实施形態も構想されている。

【 0 1 2 0 】

10

本例示的实施形態に基づいてサイド情報インスタンスを除去することにより、たとえばデコーダ側での再構成の際に、これらのサイド情報インスタンスに基づく無用な計算が回避されうる。あらかじめ定義された閾値を適切な（たとえば十分低い）レベルに設定することにより、再構成されたオーディオ信号の再生品質および／または忠実度は少なくとも近似的には維持されながらも、サイド情報インスタンスは除去されうる。

【 0 1 2 1 】

所望される再構成設定の間の差はたとえば、再構成の一部として用いられる一組の計数についてのそれぞれの値の間の差に基づいて計算されてもよい。

【 0 1 2 2 】

第三、第四または第五の側面における例示的实施形態によれば、各サイド情報インスタンスについての遷移データの前記二つの独立に割り当て可能な部分は：

20

所望される再構成設定への遷移を開始する時点を示すタイムスタンプと、所望される再構成設定への遷移を完了する時点を示すタイムスタンプ；

所望される再構成設定への遷移を開始する時点を示すタイムスタンプと、所望される再構成設定への遷移を開始する時点から所望される再構成設定に達するための継続時間を示す補間継続時間パラメータ；または

所望される再構成設定への遷移を完了する時点を示すタイムスタンプと、所望される再構成設定への遷移を開始する時点から所望される再構成設定に達するための継続時間を示す補間継続時間パラメータであってもよい。

【 0 1 2 3 】

30

換言すれば、遷移を開始および終了する時点は、それぞれの時点を示す二つのタイムスタンプによってまたはそれらのタイムスタンプの一方と遷移の継続時間を示す補間継続時間パラメータとの組み合わせによって、前記遷移データにおいて定義されうる。

【 0 1 2 4 】

それぞれのタイムスタンプはたとえば、M個のダウンミックス信号および／またはN個のオーディオ・オブジェクトを表わすために用いられる時間ベースを参照することによってそれぞれの時点を示してもよい。

【 0 1 2 5 】

第三、第四または第五の側面における例示的实施形態によれば、各クラスター・メタデータ・インスタンスについての遷移データの前記二つの独立に割り当て可能な部分は：

40

所望されるレンダリング設定への遷移を開始する時点を示すタイムスタンプと、所望されるレンダリング設定への遷移を完了する時点を示すタイムスタンプ；

所望されるレンダリング設定への遷移を開始する時点を示すタイムスタンプと、所望されるレンダリング設定への遷移を開始する時点から所望されるレンダリング設定に達するための継続時間を示す補間継続時間パラメータ；または

所望されるレンダリング設定への遷移を完了する時点を示すタイムスタンプと、所望されるレンダリング設定への遷移を開始する時点から所望されるレンダリング設定に達するための継続時間を示す補間継続時間パラメータであってもよい。

【 0 1 2 6 】

第三、第四または第五の側面における例示的实施形態によれば、各ダウンミックス・メ

50

タデータ・インスタンスについての遷移データの前記二つの独立に割り当て可能な部分は：

所望されるダウンミックス・レンダリング設定への遷移を開始する時点を示すタイムスタンプと、所望されるダウンミックス・レンダリング設定への遷移を完了する時点を示すタイムスタンプ；

所望されるダウンミックス・レンダリング設定への遷移を開始する時点を示すタイムスタンプと、所望されるダウンミックス・レンダリング設定への遷移を開始する時点から所望されるダウンミックス・レンダリング設定に達するための継続時間を示す補間継続時間パラメータ；または

所望されるダウンミックス・レンダリング設定への遷移を完了する時点を示すタイムスタンプと、所望されるダウンミックス・レンダリング設定への遷移を開始する時点から所望されるダウンミックス・レンダリング設定に達するための継続時間を示す補間継続時間パラメータであってもよい。

10

【0127】

例示的实施形態によれば、前記第三、第四または第五の側面における方法のいずれかの方法を実行するための命令をもつコンピュータ可読媒体を含むコンピュータ・プログラム・プロダクトが提供される。

【0128】

IV. 例示的实施形態

図1は、ある例示的实施形態に基づく、オーディオ・オブジェクト120をデータ・ストリーム140にエンコードするエンコーダ100を示している。エンコーダ100は、受領コンポーネント（図示せず）、ダウンミックス・コンポーネント102、エンコーダ・コンポーネント104、解析コンポーネント106および多重化コンポーネント108を有する。オーディオ・データの一つの時間フレームをエンコードするためのエンコーダ100の動作を下記で述べるが、下記の方法は時間フレーム・ベースで繰り返されることは理解される。同じことは図2～図5の説明にも当てはまる。

20

【0129】

受領コンポーネントは、複数のオーディオ・オブジェクト（N個のオーディオ・オブジェクト）120およびそれらのオーディオ・オブジェクト120に関連するメタデータ122を受領する。本稿での用法でのオーディオ・オブジェクトは、典型的には時間とともに（時間フレームと時間フレームの間で）変わる関連付けられた空間位置、すなわち動的な空間位置をもつオーディオ信号をいう。オーディオ・オブジェクト120に関連付けられたメタデータ122は典型的には、オーディオ・オブジェクト120が、デコーダ側で再生のためにどのようにレンダリングされるべきかを記述する情報を含む。特に、オーディオ・オブジェクト120に関連付けられたメタデータ122は、オーディオ・シーンの三次元空間内でのオーディオ・オブジェクト120の空間位置についての情報を含む。空間位置はデカルト座標において、あるいは方位角および高度のような方向角に任意的に距離で補足したものによって表現されることができる。オーディオ・オブジェクト120に関連付けられたメタデータ122はさらに、オブジェクト・サイズ、オブジェクト・ラウドネス、オブジェクト重要性、オブジェクト・コンテンツ型、ダイアログ向上の適用もしくははあるラウドスピーカーのレンダリングからの除外（いわゆるゾーン・マスク）などの個別的なレンダリング指示および/または他のオブジェクト属性を含んでいてもよい。

30

40

【0130】

図4を参照して述べるように、これらのオーディオ・オブジェクト120は、オーディオ・シーンの簡略化された表現に対応してもよい。

【0131】

N個のオーディオ・オブジェクト120はダウンミックス・コンポーネント102に入力される。ダウンミックス・コンポーネント102は、N個のオーディオ・オブジェクト120の組み合わせ、典型的には線形結合を形成することによって、M個のダウンミックス信号を計算する。たいていの場合、データ・ストリーム140に含まれるデータの量が

50

低減されるよう、ダウンミックス信号 1 2 4 の数はオーディオ・オブジェクト 1 2 0 の数より少ない。すなわち $M < N$ である。しかしながら、データ・ストリーム 1 4 0 の目標ビットレートが高い用途については、ダウンミックス信号 1 2 4 の数はオブジェクト 1 2 0 の数に等しくてもよい。すなわち $M = N$ となる。

【 0 1 3 2 】

ダウンミックス・コンポーネント 1 0 2 はさらに、ここで L 個の補助オーディオ信号 1 2 7 とラベル付けされる一つまたは複数の補助オーディオ信号 1 2 7 を計算してもよい。補助オーディオ信号 1 2 7 の役割は、デコーダ側での N 個のオーディオ・オブジェクト 1 2 0 の再構成を改善することである。補助オーディオ信号 1 2 7 は、 N 個のオーディオ・オブジェクト 1 2 0 の一つまたは複数に、直接またはこれらの組み合わせとして、対応してもよい。たとえば、補助オーディオ信号 1 2 7 は、ダイアログに対応するオーディオ・オブジェクト 1 2 0 など、 N 個のオーディオ・オブジェクト 1 2 0 のうちの特に重要なものに対応してもよい。重要性は、 N 個のオーディオ・オブジェクト 1 2 0 に関連付けられるメタデータ 1 2 2 によって反映されても、あるいはそれから導出されてもよい。

【 0 1 3 3 】

M 個のダウンミックス信号 1 2 4 および存在すれば L 個の補助信号 1 2 7 は、その後、ここでコア・エンコーダとラベル付けされるエンコーダ・コンポーネント 1 0 4 によってエンコードされて、 M 個のエンコードされたダウンミックス信号 1 2 6 および L 個のエンコードされた補助信号 1 2 9 を生成してもよい。エンコーダ・コンポーネント 1 0 4 は、当技術分野で知られている知覚的オーディオ・コーデックであってもよい。既知の知覚的オーディオ・コーデックの例はドルビー・デジタルおよび MPEG AAC を含む。

【 0 1 3 4 】

いくつかの実施形態では、ダウンミックス・コンポーネント 1 0 2 はさらに、 M 個のダウンミックス信号 1 2 4 をメタデータ 1 2 5 に関連付けてもよい。特に、ダウンミックス・コンポーネント 1 0 2 は、各ダウンミックス信号 1 2 4 を空間位置に関連付け、該空間位置をメタデータ 1 2 5 に含めてもよい。オーディオ・オブジェクト 1 2 0 に関連付けられたメタデータ 1 2 2 と同様に、ダウンミックス信号 1 2 4 に関連付けられたメタデータ 1 2 5 も、サイズ、ラウドネス、重要性および / または他の属性に関するパラメータを含んでいてもよい。

【 0 1 3 5 】

特に、ダウンミックス信号 1 2 4 に関連付けられた空間位置は、 N 個のオーディオ・オブジェクト 1 2 0 の空間位置に基づいて計算されてもよい。 N 個のオーディオ・オブジェクト 1 2 0 の空間位置は動的であってもよい、すなわち時間変化してもよく、 M 個のダウンミックス信号 1 2 4 に関連付けられた空間位置も動的であってもよい。換言すれば、 M 個のダウンミックス信号 1 2 4 はそれ自身がオーディオ・オブジェクトとして解釈されてもよい。

【 0 1 3 6 】

解析コンポーネント 1 0 6 は、 M 個のダウンミックス信号 1 2 4 およびもし存在すれば L 個の補助信号 1 2 9 からの N 個のオーディオ・オブジェクト 1 2 0 (または N 個のオーディオ・オブジェクト 1 2 0 の知覚的に好適な近似) の再構成を許容するパラメータを含むサイド情報 1 2 8 を計算する。サイド情報 1 2 8 も時間可変であってもよい。たとえば、解析コンポーネント 1 0 6 は、パラメトリック符号化のための任意の既知の技法に従って M 個のダウンミックス信号 1 2 4、もしあれば L 個の補助信号 1 2 7 および N 個のオーディオ・オブジェクト 1 2 0 を解析することによってサイド情報 1 2 8 を計算してもよい。あるいはまた、解析コンポーネント 1 0 6 は、 N 個のオーディオ・オブジェクトおよび (時間変化する) ダウンミックス行列を提供することによるなど該 N 個のオーディオ・オブジェクトから M 個のダウンミックス信号がどのように生成されたかについての情報を解析することによって、サイド情報 1 2 8 を計算してもよい。その場合、 M 個のダウンミックス信号 1 2 4 は、解析コンポーネント 1 0 6 への入力として厳密には要求されない。

【 0 1 3 7 】

M個のエンコードされたダウンミックス信号 1 2 6、L個のエンコードされた補助信号 1 2 9、サイド情報 1 2 8、N個のオーディオ・オブジェクトに関連付けられたメタデータ 1 2 2 およびダウンミックス信号に関連付けられたメタデータ 1 2 5 は次いで多重化コンポーネント 1 0 8 に入力され、該多重化コンポーネント 1 0 8 は、多重化技法を使って、その入力データを単一のデータ・ストリーム 1 4 0 に含める。このように、データ・ストリーム 1 4 0 は四つの型のデータを含みうる。

- a) M個のダウンミックス信号 1 2 6 (および任意的にL個の補助信号 1 2 9)
- b) M個のダウンミックス信号に関連付けられたメタデータ 1 2 5
- c) M個のダウンミックス信号からN個のオーディオ・オブジェクトを再構成するためのサイド情報 1 2 8
- d) N個のオーディオ・オブジェクトに関連付けられたメタデータ 1 2 2。

【0 1 3 8】

上述したように、オーディオ・オブジェクトを符号化するためのいくつかの従来技術のシステムは、M個のダウンミックス信号が、M個のチャンネルをもつスピーカー配位のチャンネルでの再生に好適なように選ばれることを要求する。これは本稿で後方互換なダウンミックスと称される。そのような従来技術の要求は、オーディオ・オブジェクトがあらかじめ定義された仕方でのみ組み合わせられうるという点で、ダウンミックス信号の計算を制約する。よって、従来技術によれば、ダウンミックス信号は、デコーダ側でのオーディオ・オブジェクトの再構成を最適にするという観点から選択されるのではない。

【0 1 3 9】

従来技術のシステムとは異なり、ダウンミックス・コンポーネント 1 0 2 は、N個のオーディオ・オブジェクトに関して信号適応的な仕方ではM個のダウンミックス信号 1 2 4 を計算する。特に、ダウンミックス・コンポーネント 1 0 2 は、各時間フレームについて、M個のダウンミックス信号 1 2 4 を、現在何らかの基準を最適化するオーディオ・オブジェクト 1 2 0 の組み合わせとして計算してもよい。基準は典型的には、5.1または他のラウドスピーカー配位のようないかなるラウドスピーカー配位とも独立であるよう定義される。これは、M個のダウンミックス信号 1 2 4 またはそのうちの少なくとも一つが、Mチャンネルをもつスピーカー配位のチャンネルでの再生に好適であるオーディオ信号に制約されないことを含意する。よって、ダウンミックス・コンポーネント 1 0 2 は、たとえばデコーダ側でのオーディオ・オブジェクト 1 2 0 の再構成を改善するために、M個のダウンミックス信号 1 2 4 を、N個のオーディオ・オブジェクト 1 2 0 の時間変動 (N個のオーディオ・オブジェクトの空間位置を含むメタデータ 1 2 2 の時間変動を含む) に適応させてもよい。

【0 1 4 0】

ダウンミックス・コンポーネント 1 0 2 は、M個のダウンミックス信号を計算するために異なる基準を適用してもよい。一例によれば、M個のダウンミックス信号は、該M個のダウンミックス信号に基づくN個のオーディオ・オブジェクトの再構成が最適化されるよう、計算されてもよい。たとえば、ダウンミックス・コンポーネント 1 0 2 は、N個のオーディオ・オブジェクト 1 2 0 およびM個のダウンミックス信号 1 2 4 に基づく該N個のオーディオ・オブジェクトの再構成から形成される再構成誤差を最小にしてもよい。

【0 1 4 1】

別の例によれば、基準はN個のオーディオ・オブジェクト 1 2 0 の空間位置に、特に空間的近接性に基づく。上記で論じたように、N個のオーディオ・オブジェクト 1 2 0 は、該N個のオーディオ・オブジェクト 1 2 0 の空間位置を含む関連付けられたメタデータ 1 2 2 を有する。該メタデータ 1 2 2 に基づいて、N個のオーディオ・オブジェクト 1 2 0 の空間的近接性が導出されてもよい。

【0 1 4 2】

より詳細には、ダウンミックス・コンポーネント 1 0 2 は、M個のダウンミックス信号 1 2 4 を決定するために、第一のクラスタリング手順を適用してもよい。第一のクラスタリング手順は、N個のオーディオ・オブジェクト 1 2 0 を、空間的近接性に基づいてM個の

10

20

30

40

50

クラスターに関連付けることを含んでいてもよい。関連付けられたメタデータ 1 2 2 によって表わされる、オブジェクト・サイズ、オブジェクト・ラウドネス、オブジェクト重要性を含む前記N個のオーディオ・オブジェクト 1 2 0 のさらなる属性も、オーディオ・オブジェクト 1 2 0 をM個のクラスターに関連付ける際に、考慮に入れられてもよい。

【 0 1 4 3 】

一例によれば、よく知られたK平均アルゴリズムが、N個のオーディオ・オブジェクトのメタデータ 1 2 2 (空間位置)を入力として、N個のオーディオ・オブジェクト 1 2 0 を空間的近接性に基づいてM個のクラスターに関連付けるために使われてもよい。N個のオーディオ・オブジェクト 1 2 0 の前記さらなる属性は、K平均アルゴリズムにおける重み付け因子として使われてもよい。

10

【 0 1 4 4 】

別の例によれば、第一のクラスタリング手順は、メタデータ 1 2 2 によって与えられるオーディオ・オブジェクトの重要性を選択基準として使う選択手順に基づいていてもよい。より詳細には、ダウンミックス・コンポーネント 1 0 2 は、最も重要な諸オーディオ・オブジェクト 1 2 0 をそのまま通過させ、M個のダウンミックス信号のうちの一つまたは複数がN個のオーディオ・オブジェクト 1 2 0 の一つまたは複数に対応するようにしてもよい。残りの、それほど重要でないオーディオ・オブジェクトは、上記で論じたように、空間的近接性に基づいてクラスターに関連付けられてもよい。

【 0 1 4 5 】

オーディオ・オブジェクトをクラスタリングするさらなる例は、米国仮特許出願第61/865,072号および同出願の優先権を主張するその後の出願において与えられている。

20

【 0 1 4 6 】

別の例によれば、第一のクラスタリング手順は、あるオーディオ・オブジェクト 1 2 0 をM個のクラスターのうちの二つ以上に関連付けてもよい。たとえば、オーディオ・オブジェクト 1 2 0 は、M個のクラスターにわたって分配されてもよい。ここで、分配はたとえば、オーディオ・オブジェクト 1 2 0 の空間位置に、また任意的にはオブジェクト・サイズ、オブジェクト・ラウドネス、オブジェクト重要性などを含むオーディオ・オブジェクトのさらなる属性にも依存する。分配は割合によって反映されてもよい。たとえば、オーディオ・オブジェクトは三つのクラスターに割合20%、30%、50%に従って分配される。

【 0 1 4 7 】

30

ひとたびN個のオーディオ・オブジェクト 1 2 0 がM個のクラスターに関連付けられたら、ダウンミックス・コンポーネント 1 0 2 は、各クラスターについてのダウンミックス信号 1 2 4 を、そのクラスターに関連付けられたオーディオ・オブジェクト 1 2 0 の組み合わせ、典型的には線形結合を形成することによって計算する。典型的には、ダウンミックス・コンポーネント 1 0 2 は、オーディオ・オブジェクト 1 2 0 に関連付けられたメタデータ 1 2 2 に含まれるパラメータを、組み合わせを形成するときの重みとして使ってもよい。例として、あるクラスターに関連付けられているオーディオ・オブジェクト 1 2 0 はオブジェクト・サイズ、オブジェクト・ラウドネス、オブジェクト重要性、オブジェクト位置、クラスターに関連付けられた空間位置に対するオブジェクトからの距離(詳細は後述)などに従って重み付けされてもよい。オーディオ・オブジェクト 1 2 0 がM個のクラスターにわたって分配される場合には、分配を反映する割合が、組み合わせを形成するときの重みとして使われてもよい。

40

【 0 1 4 8 】

第一のクラスタリング手順は、M個のダウンミックス信号 1 2 4 のそれぞれの空間位置との関連付けを簡単に許容するという点で有利である。たとえば、ダウンミックス・コンポーネント 1 0 2 は、クラスターに対応するダウンミックス信号 1 2 4 の空間位置を、そのクラスターに関連付けられたオーディオ・オブジェクト 1 2 0 の空間位置に基づいて計算してもよい。そのクラスターに関連付けられているオーディオ・オブジェクトの空間位置の重心または重み付けされた重心がこの目的のために使われてもよい。重み付けされた重心の場合には、クラスターに関連付けられたオーディオ・オブジェクト 1 2 0 の組み合

50

わせを形成するときと同じ重みが使われてもよい。

【0149】

図2は、図1のエンコーダ100に対応するデコーダ200を示している。デコーダ200は、オーディオ・オブジェクト再構成をサポートする型のものである。デコーダ200は、受領コンポーネント208、デコーダ・コンポーネント204および再構成コンポーネント206を有する。デコーダ200はさらに、レンダラー210を有していてもよい。あるいはまた、デコーダ200は再生システムの一部をなすレンダラー210に結合されてもよい。

【0150】

受領コンポーネント208は、エンコーダ100からデータ・ストリーム240を受領するよう構成されている。受領コンポーネント208は、受領されたデータ・ストリーム240をその成分に多重分離するよう構成された多重分離〔デマルチプレックス〕コンポーネントを有する。該成分は、この場合は、M個のエンコードされたダウンミックス信号226と、任意的にL個のエンコードされた補助信号229と、M個のダウンミックス信号およびL個の補助信号からのN個のオーディオ・オブジェクトの再構成のためのサイド情報228と、該N個のオーディオ・オブジェクトに関連付けられたメタデータ222である。

10

【0151】

デコーダ・コンポーネント204はM個のエンコードされたダウンミックス信号226および任意的にL個の補助信号227を処理してM個のダウンミックス信号224を生成する。上記でさらに論じたように、M個のダウンミックス信号224は、N個のオーディオ・オブジェクトからエンコーダ側で適応的に、すなわち、いかなるラウドスピーカー配位とも独立な基準に従ってN個のオーディオ・オブジェクトの組み合わせを形成することによって、形成されたものである。

20

【0152】

次いで、オブジェクト再構成コンポーネント206は、エンコーダ側で導出されたサイド情報228によって案内されてM個のダウンミックス信号224および任意的にL個の補助信号227に基づいてN個のオーディオ・オブジェクト220（またはこれらのオーディオ・オブジェクトの知覚的に好適な近似）を再構成する。オブジェクト再構成コンポーネント206は、オーディオ・オブジェクトのそのようなパラメトリックな再構成のための任意の既知の技法を適用しうる。

30

【0153】

次いで、再構成されたN個のオーディオ・オブジェクト220は、再生のために好適なマルチチャネル出力信号230を生成するために、オーディオ・オブジェクト222に関連付けられたメタデータ222および再生システムのチャネル配位についての知識を使ってレンダラーによって処理される。典型的なスピーカー再生配位は22.2および11.1を含む。サウンドバー・スピーカー・システムまたはヘッドフォン（バイノーラル呈示）での再生は、そのような再生システムのための専用レンダラーを用いても可能である。

【0154】

図3は、図1のエンコーダ100に対応する低計算量デコーダ300を示している。デコーダ300はオーディオ・オブジェクト再構成をサポートしない。デコーダ300は、受領コンポーネント308およびデコーダ・コンポーネント304を有する。デコーダ300はさらに、レンダラー310を有していてもよい。あるいはまた、該デコーダは再生システムの一部をなすレンダラー310に結合されてもよい。

40

【0155】

上記で論じたように、後方互換なダウンミックス（5.1ダウンミックスのような）、すなわちM個のチャネルをもつ再生システム上での直接再生のために好適なM個のダウンミックス信号を含むダウンミックスを使う従来技術のシステムは、（たとえば5.1マルチチャネル・ラウドスピーカー・セットアップしかサポートしない）レガシー再生システムのための低計算量のデコードを簡単に可能にする。そのような従来技術のシステムは典型的に

50

は、後方互換なダウンミックス信号自身をデコードし、サイド情報（図2の項目228参照）およびオーディオ・オブジェクトに関連付けられているメタデータ（図2の項目222参照）といったデータ・ストリームの追加的な部分を破棄する。しかしながら、ダウンミックス信号が上記のように適応的に形成されるとき、ダウンミックス信号は一般に、レガシー・システム上での直接再生のためには好適ではない。

【0156】

デコーダ300は、特定の再生配位のみをサポートするレガシー再生システム上での再生のために適応的に形成されるM個のダウンミックス信号の低計算量デコードを許容するデコーダの例である。

【0157】

受領コンポーネント308は、図1のエンコーダ100のようなエンコーダからビットストリーム340を受領する。受領コンポーネント308は、ビットストリーム340をその成分に多重分離する。この場合は、受領コンポーネント308は、エンコードされたM個のダウンミックス信号326およびM個のダウンミックス信号に関連付けられたメタデータ325を保持するだけである。L個の補助信号（図2の項目229参照）、N個のオーディオ・オブジェクトに関連付けられているメタデータ（図2の項目222参照）およびサイド情報（図2の項目228参照）といった、データ・ストリーム340の他の成分は破棄される。

【0158】

デコーダ・コンポーネント304はM個のエンコードされたダウンミックス信号326をデコードしてM個のダウンミックス信号324を生成する。次いで、M個のダウンミックス信号は、ダウンミックス・メタデータと一緒に、（典型的にはM個のチャンネルをもつ）M個のダウンミックス信号をレガシー再生フォーマットに対応するマルチチャンネル出力330にレンダリングするレンダラー310に入力される。ダウンミックス・メタデータ325はM個のダウンミックス信号324の空間位置を含むので、レンダラー310は典型的には図2のレンダラー210と同様であってもよい。唯一の違いは、レンダラー310は、オーディオ・オブジェクト220およびその関連付けられたメタデータ222の代わりに、今やM個のダウンミックス信号324および該M個のダウンミックス信号324に関連付けられたメタデータ325を入力として取るということである。

【0159】

図1との関連で上述したように、N個のオーディオ・オブジェクト120は、オーディオ・シーンの簡略化された表現に対応してもよい。

【0160】

一般に、オーディオ・シーンは、オーディオ・オブジェクトおよびオーディオ・チャンネルを含んでいてもよい。オーディオ・チャンネルとは、ここでは、マルチチャンネル・スピーカー配位のあるチャンネルに対応するオーディオ信号を意味する。そのようなマルチチャンネル・スピーカー配位の例は22.2配位、11.1配位などを含む。オーディオ・チャンネルは、そのチャンネルのスピーカー位置に対応する空間位置をもつ静的なオーディオ・オブジェクトとして解釈してもよい。

【0161】

いくつかの場合には、オーディオ・シーン内のオーディオ・オブジェクトおよびオーディオ・チャンネルの数は、100を超えるオーディオ・オブジェクトおよび1~24のオーディオ・チャンネルなど、膨大であることがある。これらのオーディオ・オブジェクト/チャンネルのすべてがデコーダ側で再構成される場合、大量の計算パワーが必要とされる。さらに、オブジェクト・メタデータおよびサイド情報に関連付けられた帰結するデータ・レートは、多数のオブジェクトが入力として与えられる場合には、一般に、非常に大きい。このため、デコーダ側で再構成されるべきオーディオ・オブジェクトの数を減らすために、オーディオ・シーンを簡略化することが有利である。この目的のために、エンコーダは、第二のクラスタリング手順に基づいてオーディオ・シーン内のオーディオ・オブジェクトの数を減らすクラスタリング・コンポーネントを有していてもよい。第二のクラスタリング手

10

20

30

40

50

順は、等しいまたは非常に近い位置をもつオーディオ・オブジェクトなど、オーディオ・シーン内に存在する空間的冗長性を活用することをねらいとする。追加的に、オーディオ・オブジェクトの知覚的重要性が考慮に入れられてもよい。一般に、そのようなクラスタリング・コンポーネントは、図1のダウンミックス・コンポーネント102と逐次的にまたは並列に配置されうる。逐次的な配置は図4を参照して述べる。並列な配置は図5を参照して述べる。

【0162】

図4はエンコーダ400を示している。図1を参照して述べたコンポーネントに加えて、エンコーダ400は、クラスタリング・コンポーネント409を有する。クラスタリング・コンポーネント409は、ダウンミックス・コンポーネント102と逐次的に配置されている。つまり、クラスタリング・コンポーネント409の出力がダウンミックス・コンポーネント102に入力される。

10

【0163】

クラスタリング・コンポーネント409は、オーディオ・オブジェクト421aおよび/またはオーディオ・チャンネル421bを、オーディオ・オブジェクト421aの空間位置を含む関連付けられたメタデータ423と一緒に、入力として受ける。クラスタリング・コンポーネント409は、各オーディオ・チャンネル421bをオーディオ・チャンネル421bに対応するスピーカー位置の空間位置に関連付けることによって、オーディオ・チャンネル421bを静的なオーディオ・オブジェクトに変換する。オーディオ・オブジェクト421aおよびオーディオ・チャンネル421bから形成された静的なオーディオ・オブジェクトは、第一の複数のオーディオ・オブジェクト421と見ることができる。

20

【0164】

クラスタリング・コンポーネント409は一般に、前記第一の複数のオーディオ・オブジェクト421を、ここでは図1のN個のオーディオ・オブジェクト120に対応する第二の複数のオーディオ・オブジェクトに減らす。この目的のため、クラスタリング・コンポーネント409は第二のクラスタリング手順を適用してもよい。

【0165】

第二のクラスタリング手順は一般に、ダウンミックス・コンポーネント102に関して上記した第一のクラスタリング手順と同様である。したがって、第一のクラスタリング手順の記述は第二のクラスタリング手順にも当てはまる。

30

【0166】

特に、第二のクラスタリング手順は、第一の複数のオーディオ・オブジェクト121を、第一の複数のオーディオ・オブジェクト121の空間的近接性に基づいて、少なくとも一つのクラスター、ここではN個のクラスターに関連付けることを含む。さらに上記で述べたように、クラスターとの関連付けは、メタデータ423によって表わされる、オーディオ・オブジェクトの他の属性にも基づいていてもよい。次いで、各クラスターは、そのクラスターに関連付けられたオーディオ・オブジェクトの組み合わせ（線形結合）であるオブジェクトによって表わされる。図示した例では、N個のクラスターがあるのでN個のオーディオ・オブジェクト120が生成される。クラスタリング・コンポーネント409はさらに、そのように生成されたN個のオーディオ・オブジェクト120についてのメタデータ122を計算する。メタデータ122は、N個のオーディオ・オブジェクト120の空間位置を含む。N個のオーディオ・オブジェクト120のそれぞれの空間位置は、対応するクラスターに関連付けられているオーディオ・オブジェクトの空間位置に基づいて計算されてもよい。例として、空間位置は、図1を参照して上記でさらに説明したように、クラスターに関連付けられているオーディオ・オブジェクトの空間位置の重み付けされた重心として計算されてもよい。

40

【0167】

クラスタリング・コンポーネント409によって生成されるN個のオーディオ・オブジェクト120は次いで、図1を参照してさらに述べたように、ダウンミックス・コンポーネント120に入力される。

50

【0168】

図5はエンコーダ500を示している。図1を参照して述べたコンポーネントに加えて、エンコーダ500は、クラスタリング・コンポーネント509を有する。クラスタリング・コンポーネント509は、ダウンミックス・コンポーネント102と並列に配置されている。つまり、ダウンミックス・コンポーネント102とクラスタリング・コンポーネント509は同じ入力をもつ。

【0169】

入力は、図1のN個のオーディオ・オブジェクト120に対応する第一の複数のオーディオ・オブジェクトを、該第一の複数のオーディオ・オブジェクトの空間位置を含むメタデータ122と一緒に含む。第一の複数のオーディオ・オブジェクト120は、図4の第一の複数のオーディオ・オブジェクト121と同様に、オーディオ・オブジェクトおよび静的なオーディオ・オブジェクトに変換されるオーディオ・チャンネルを含む。ダウンミックス・コンポーネント102がオーディオ・シーンの簡略化されたバージョンに対応する低減した数のオーディオ・オブジェクトに対して作用する図4の逐次的な配置とは対照的に、図5のダウンミックス・コンポーネント102は、M個のダウンミックス信号124を生成するためにオーディオ・シーンのフル・オーディオ・コンテンツに対して作用する。

【0170】

クラスタリング・コンポーネント509は機能において図4を参照して述べたクラスタリング・コンポーネント409と同様である。特に、クラスタリング・コンポーネント509は、上記の第二のクラスタリング手順を適用することによって、第一の複数のオーディオ・オブジェクト120を、ここではK個のオーディオ・オブジェクトによって示されている第二の複数のオーディオ・オブジェクト521に減らす。典型的には $M < K < N$ （高ビット用途については $M \approx K \approx N$ ）である。このように、第二の複数のオーディオ・オブジェクト521は、N個のオーディオ・オブジェクト126に基づいて形成された一組のオーディオ・オブジェクトである。さらに、クラスタリング・コンポーネント509は、第二の複数のオーディオ・オブジェクト521の空間位置を含む第二の複数のオーディオ・オブジェクト521（K個のオーディオ・オブジェクト）についてのメタデータ522を計算する。メタデータ522は、多重分離コンポーネント108によってデータ・ストリーム540に含められる。解析コンポーネント106は、M個のダウンミックス信号124からの第二の複数のオーディオ・オブジェクト521、すなわち前記N個のオーディオ・オブジェクトに基づいて形成された前記一組のオーディオ・オブジェクト（ここではK個のオーディオ・オブジェクト）の再構成を可能にするサイド情報528を計算する。サイド情報528は、多重化コンポーネント108によってデータ・ストリーム540に含められる。上記でさらに論じたように、解析コンポーネント106はたとえばサイド情報528を、第二の複数のオーディオ・オブジェクト521およびM個のダウンミックス信号124を解析することによって導出してもよい。

【0171】

エンコーダ500によって生成されるデータ・ストリーム540は、図2のデコーダまたは図3のデコーダ300によってデコードされうる。しかしながら、図2の再構成されたオーディオ・オブジェクト（N個のオーディオ・オブジェクトとラベル付けられる）は今や図5の第二の複数のオーディオ・オブジェクト521（K個のオーディオ・オブジェクトとラベル付けされる）に対応し、該オーディオ・オブジェクトに関連付けられたメタデータ222（N個のオーディオ・オブジェクトのメタデータとラベル付けられる）は今や図5の第二の複数のオーディオ・オブジェクトのメタデータ522（K個のオーディオ・オブジェクトのメタデータとラベル付けされる）に対応する。

【0172】

オブジェクト・ベースのオーディオ・エンコード/デコード・システムでは、オブジェクトに関連付けられたサイド情報またはメタデータは典型的には、付随するデータ・レートを制限するために、時間的に比較的低頻度で（まれに）更新される。オブジェクト位置

についての典型的な更新間隔は、オブジェクトの速度、要求される位置精度、メタデータの記憶もしくは伝送のための利用可能な帯域幅などに依存して、10から500ミリ秒の間の範囲でありうる。そのようなまれな、またさらには不規則なメタデータ更新は、二つの相続くメタデータ・インスタンスの中間のオーディオ・サンプルについて、メタデータおよび/またはレンダリング行列（すなわち、レンダリングにおいて用いられる行列）の補間を必要とする。補間なしでは、レンダリング行列における結果的なステップごとの変化が、ステップごとの行列更新によって導入されるスペクトル・スプラッター（splat）の結果として、望ましくないスイッチング・アーチファクト、クリック音、ジッパー・ノイズまたは他の望ましくないアーチファクトを引き起こすことがありうる。

【0173】

図6は、一組のメタデータ・インスタンスに基づく、オーディオ信号またはオーディオ・オブジェクトのレンダリングのためのレンダリング行列を計算するための典型的な既知のプロセスを示している。図6に示されるように、一組のメタデータ・インスタンス（m1ないしm4）610は、時間軸620に沿ってその位置によって示される一組の時点（t1ないしt4）に対応する。その後、各メタデータ・インスタンスは、該メタデータ・インスタンスと同じ時点において有効なそれぞれのレンダリング行列（c1ないしc4）630またはレンダリング設定に変換される。こうして、図のように、メタデータ・インスタンスm1は時点t1におけるレンダリング行列c1を生成し、メタデータ・インスタンスm2は時点t2におけるレンダリング行列c2を生成し、などとなる。簡単のため、図6は各メタデータ・インスタンスm1ないしm4について一つのレンダリング行列しか示していない。しかしながら、実際的なシステムでは、レンダリング行列c1は、出力信号 $y_j(t)$ を生成するためにそれぞれのオーディオ信号 $x_i(t)$ に適用されるべき一組のレンダリング行列係数または利得係数 $c_{1,i,j}$ を含んでいてもよい。

【0174】

$$y_j(t) = \sum_i x_i(t) c_{1,i,j}$$

レンダリング行列630は一般に、異なる時点における利得値を表わす係数を含む。メタデータ・インスタンスは、所定の離散的な時点において定義され、メタデータ時点の中間のオーディオ・サンプルについては、レンダリング行列630をつなぐ破線640によって示されるように、レンダリング行列は補間される。そのような補間は線形に実行することができるが、他の補間方法（帯域制限された補間、正弦/余弦補間など）が使われることもできる。メタデータ・インスタンス（および対応するレンダリング行列）の間の時間間隔は、「補間継続時間」と称される。そのような期間は一樣であってもよく、あるいは時刻t2とt3の間の補間継続時間に比べて時刻t3とt4の間の補間継続時間がより長いなど、異なっているもよい。

【0175】

多くの場合、メタデータ・インスタンスからのレンダリング行列係数の計算はよく定義されるが、（補間された）レンダリング行列が与えられてメタデータ・インスタンスを計算する逆のプロセスはしばしば困難であるか、またはさらには不可能である。これに関し、メタデータからレンダリング行列を生成するプロセスは時に暗号学的方向性関数と見なされることができる。既存のメタデータ・インスタンスの間の新しいメタデータ・インスタンスを計算するプロセスは、メタデータの「再サンプリング」と称される。メタデータの再サンプリングはしばしば、ある種のオーディオ処理タスクの間に必要となる。たとえば、オーディオ・コンテンツがカット/マージ/混合などにより編集されるとき、そのような編集はメタデータ・インスタンスの中間のところで生起することがある。この場合、メタデータの再サンプリングが必要とされる。もう一つのそのような事例はオーディオおよび関連付けられたメタデータがフレーム・ベースのオーディオ・コーデックを用いてエンコードされるときである。この場合、伝送の間のフレーム損失に対する耐性を改善するために、各オーディオ・コーデック・フレームについて、好ましくはそのコーデック・フレームの先頭にタイムスタンプをもつ少なくとも一つのメタデータ・インスタンスをもつことが望ましい。さらに、メタデータの補間は、二値のメタデータのような、標準的な

技法が二回目毎に多少なりとも正しくない値を導出することになるある種の型のメタデータについても効果がない。たとえば、0排除マスク (zero exclusion masks) のような二値フラグがある種の時点においてある種のオブジェクトをレンダリングから排除するために使われる場合、レンダリング行列係数からまたはメタデータの近隣のインスタンスからメタデータの有効な集合を推定することは事実上不可能である。これは図6では、時刻 t_3 と t_4 の間の補間継続時間におけるレンダリング行列係数からメタデータ・インスタンス m_{3a} を外挿または導出する失敗した試行として示されている。図6に示されるように、メタデータ・インスタンス m_x は、ある種の離散的な時点 t_x においてのみ確定的に定義され、それが関連付けられた一組の行列係数 c_x を生じる。これらの離散的な時刻 t_x の中間では、行列係数の諸セットが、過去または将来のメタデータ・インスタンスに基づいて補間される必要がある。しかしながら、上記のように、現在のメタデータ補間方式は、メタデータ補間プロセスにおける不可避免的な不正確さに起因する空間的オーディオ品質の損失を被る。例示的な実施形態に基づく代替的な補間方式について図7～図11を参照して下記に述べる。

10

20

30

40

50

【0176】

図1～図5を参照して述べた例示的な実施形態では、N個のオーディオ・オブジェクト120、220に関連付けられたメタデータ122、222およびK個のオブジェクト522に関連付けられたメタデータ522は、少なくともいくつかの例示的な実施形態では、クラスタリング・コンポーネント409および509に発し、クラスター・メタデータと称されてもよい。さらに、ダウンミックス信号124、324に関連付けられたメタデータ125、325はダウンミックス・メタデータと称されてもよい。

【0177】

図1、図4および図5を参照して述べたように、ダウンミックス・コンポーネント102は、N個のオーディオ・オブジェクト120の組み合わせを信号適応的な仕方で、すなわちいかなるラウドスピーカー配位とも独立な基準に従って、形成することによって、M個のダウンミックス信号124を計算してもよい。ダウンミックス・コンポーネント102のそのような動作は、第一の側面における例示的な実施形態の特徴である。他の側面における例示的な実施形態によれば、ダウンミックス・コンポーネント102はたとえば、信号適応的な仕方で、あるいはM個のダウンミックス信号がM個のチャンネルをもつスピーカー配位のチャンネルでの再生に好適なように、すなわち後方互換なダウンミックスとして、N個のオーディオ・オブジェクト120の組み合わせを形成することによって、M個のダウンミックス信号124を計算してもよい。

【0178】

ある例示的な実施形態では、図4を参照して述べたエンコーダ400は、再サンプリングのために、すなわち追加的なメタデータおよびサイド情報インスタンスを生成するために特に好適なメタデータおよびサイド情報フォーマットを用いる。本例示的な実施形態では、解析コンポーネント106は、N個のオーディオ・オブジェクト120を再構成するためのそれぞれの所望される再構成設定を指定する複数のサイド情報インスタンスと、各サイド情報インスタンスについて、現在の再構成設定から、そのサイド情報インスタンスによって指定される所望される再構成設定への遷移を開始する時点および該遷移を完了する時点を組み合わせて定義する、二つの独立に割り当て可能な部分を含む遷移データとを含む形で、サイド情報128を計算する。本例示的な実施形態では、各サイド情報インスタンスについての前記遷移データの前記二つの独立に割り当て可能な部分は：所望される再構成設定への遷移を開始する時点を示すタイムスタンプと、所望される再構成設定への遷移を開始する時点から所望される再構成設定に到達するための継続時間を示す補間継続時間パラメータとである。遷移が起こる期間は、本例示的な実施形態では、遷移が始まるべき時刻と遷移期間の長さによって一意的に定義される。サイド情報128のこの特定の形は、図7～図11を参照して後述する。この遷移期間を一意的に定義するいくつかの他の方法があることは理解される。たとえば、該期間の開始、終了または中間点の形の基準点に該期間の長さを付したものが前記遷移データにおいて、該期間を一意的に定義するために

用いられてもよい。あるいはまた、前記期間の始点および終点が、該期間を一意的に定義するために用いられてもよい。

【0179】

本例示的实施形態では、クラスタリング・コンポーネント409は第一の複数のオーディオ・オブジェクト421を、ここでは図1のN個のオーディオ・オブジェクト120に対応する第二の複数のオーディオ・オブジェクトに減らす。クラスタリング・コンポーネント409は、生成されたN個のオーディオ・オブジェクト120についてのクラスタ・メタデータ122を計算する。該メタデータは、デコーダ側のレンダラー210においてN個のオーディオ・オブジェクト122のレンダリングを可能にする。クラスタリング・コンポーネント409は、クラスタ・メタデータ122を、N個のオーディオ・オブジェクト120をレンダリングするためのそれぞれの所望されるレンダリング設定を指定する複数のクラスタ・メタデータ・インスタンスと、各クラスタ・メタデータ・インスタンスについて、現在のレンダリング設定から、そのクラスタ・メタデータ・インスタンスによって指定される所望されるレンダリング設定への遷移を開始する時点および所望されるレンダリング設定までの該遷移を完了する時点を組み合わせて定義する、二つの独立に割り当て可能な部分を含む遷移データとを含む形で、クラスタ・メタデータ122を提供する。本例示的实施形態では、各クラスタ・メタデータ・インスタンスについての前記遷移データの前記二つの独立に割り当て可能な部分は：所望されるレンダリング設定への遷移を開始する時点を示すタイムスタンプと、所望されるレンダリング設定への遷移を開始する時点から所望されるレンダリング設定に到達するための継続時間を示す補間継続時間パラメータとである。クラスタ・メタデータ122のこの特定の形は、図7～図11を参照して後述する。

10

20

【0180】

本例示的实施形態では、ダウンミックス・コンポーネント102は各ダウンミックス信号124を空間位置に関連付け、該空間位置を、デコーダ側でのレンダラー310においてM個のダウンミックス信号のレンダリングを許容するダウンミックス・メタデータ125に含める。ダウンミックス・コンポーネント102は、ダウンミックス信号をレンダリングするためのそれぞれの所望されるダウンミックス・レンダリング設定を指定する複数のダウンミックス・メタデータ・インスタンスと、各ダウンミックス・メタデータ・インスタンスについて、現在のダウンミックス・レンダリング設定から、そのダウンミックス・メタデータ・インスタンスによって指定される所望されるダウンミックス・レンダリング設定への遷移を開始する時点および所望されるダウンミックス・レンダリング設定までの該遷移を完了する時点を組み合わせて定義する、二つの独立に割り当て可能な部分を含む遷移データとを含む形で、ダウンミックス・メタデータ125を提供する。本例示的实施形態では、各ダウンミックス・メタデータ・インスタンスについての前記遷移データの前記二つの独立に割り当て可能な部分は：所望されるダウンミックス・レンダリング設定への遷移を開始する時点を示すタイムスタンプと、所望されるダウンミックス・レンダリング設定への遷移を開始する時点から所望されるダウンミックス・レンダリング設定に到達するための継続時間を示す補間継続時間パラメータとである。

30

【0181】

本例示的实施形態では、同じフォーマットがサイド情報128、クラスタ・メタデータ122およびダウンミックス・メタデータ125について用いられる。このフォーマットが、オーディオ信号のレンダリングのためのメタデータに関し、これから図7～図11を参照して記述される。しかしながら、図7～図11を参照して記述される以下の例において、「オーディオ信号のレンダリングのためのメタデータ」のような用語および表現は、「オーディオ・オブジェクトの再構成のためのサイド情報」、「オーディオ・オブジェクトのレンダリングのためのクラスタ・メタデータ」または「ダウンミックス信号のレンダリングのためのダウンミックス・メタデータ」のような対応する用語または表現によって置き換えられてもよいことは理解される。

40

【0182】

50

図 7 は、ある例示的な実施形態に基づく、オーディオ信号のレンダリングにおいて用いられる係数曲線の、メタデータに基づく導出を示している。図 7 に示されるように、たとえば一意的な時間サンプルに関連付けられた異なる時点 t_x において生成されるメタデータ・インスタンス m_x のセットは、変換器 710 によって対応する行列係数値 c_x の諸セットに変換される。係数のこれらのセットは、オーディオ・コンテンツがレンダリングされるべき再生システムにおけるさまざまなスピーカーおよびドライバにオーディオ信号をレンダリングするために用いられる、利得因子とも称される利得値を表わす。次いで、補間器 720 が利得因子 c_x を補間して、離散的な時刻 t_x の間の係数曲線を生成する。ある実施形態では、各メタデータ・インスタンス m_x に関連付けられたタイムスタンプ t_x は、ランダムな時点、クロック回路によって生成される同期的な諸時点、フレーム境界のようなオーディオ・コンテンツに関係した時間イベントまたは他の任意の適切な時間をもつイベントに対応してもよい。上記のように、図 7 を参照して与えられる説明は、オーディオ・オブジェクトの再構成のためのサイド情報にも同様に当てはまる。

【0183】

図 8 は、ある実施形態に基づくメタデータ・フォーマットを示している（そして、上記のように、以下の説明は対応するサイド情報フォーマットにも同様に当てはまる）。これは、上記のように、現在の方法に付随する補間問題の少なくともいくつかに対処する。それは、遷移もしくは補間の開始時刻としてのタイムスタンプを定義し、各メタデータ・インスタンスを、遷移継続時間または補間継続時間（「ライブ・サイズ」とも称される）を表わす補間継続時間パラメータをもって増強することによる。図 8 に示されるように、メタデータ・インスタンス m_2 ないし m_4 (810) のセットは、レンダリング行列 c_2 ないし c_4 (830) のセットを指定する。各メタデータ・インスタンスは、特定の時点 t_x において生成され、各メタデータ・インスタンスはそのタイムスタンプに関して定義される。 t_2 に対する m_2 、 t_3 に対する m_3 などである。関連付けられたレンダリング行列 830 は、それぞれの補間継続時間 d_2 、 d_3 、 d_4 (840) の間の遷移を実行した後に、各メタデータ・インスタンス 810 の関連付けられたタイムスタンプ (t_1 ないし t_4) から生成される。補間継続時間（またはランプ・サイズ）を示す補間継続時間パラメータは、各メタデータ・インスタンスと一緒に含まれる。すなわち、メタデータ・インスタンス m_2 は d_2 を含み、 m_3 は d_3 を含み、などとなる。概略的に、このことは $m_x = (\text{metadata}(t_x), d_x) \rightarrow c_x$ のように表現できる。このようにして、メタデータは本質的には、現在のレンダリング設定（たとえば前のメタデータから帰結する現在のレンダリング行列）から新しいレンダリング設定（たとえば現在のメタデータから帰結する新しいレンダリング行列）にどのように進むかの概略を提供する。各メタデータ・インスタンスは、該メタデータ・インスタンスが受領された瞬間に対する将来のある指定された時点において発効することが意図されており、係数曲線は係数の前の状態から導出される。よって、図 8 において、 m_2 は継続時間 d_2 後に c_2 を生成し、 m_3 は継続時間 d_3 後に c_3 を生成し、 m_4 は継続時間 d_4 後に c_4 を生成する。補間のためのこの方式では、前のメタデータが知られている必要はなく、前のレンダリング行列またはレンダリング状態だけが要求される。用いられる補間は、システム制約条件および構成に依存して線形または非線形でありうる。

【0184】

図 8 のメタデータ・フォーマットは、図 9 に示されるメタデータの無損失な再サンプリングを許容する。図 9 は、例示的な実施形態に基づく、メタデータの無損失な処理の第一の例を示している（そして、上記のように、以下の説明は対応するサイド情報フォーマットにも同様に当てはまる）。図 9 は、補間継続時間 d_2 ないし d_4 を含む、それぞれ将来のレンダリング行列 c_2 ないし c_4 を指すメタデータ・インスタンス m_2 ないし m_4 を示している。メタデータ・インスタンス m_2 ないし m_4 のタイムスタンプは t_2 ないし t_4 として与えられている。図 9 の例では、時刻 t_{4a} におけるメタデータ・インスタンス m_{4a} が加えられる。そのようなメタデータは、システムのエラー耐性を向上させるまたはメタデータ・インスタンスをオーディオ・フレームの先頭 / 末尾に同期させるなど、いくつかの理由で加えられる。たとえば、時刻 t_{4a} は、メタデータに関連付けられたオーディオ・コンテンツを符号化す

10

20

30

40

50

るために用いられるオーディオ・コーデックが新しいフレームを開始する時刻を表わしてもよい。無損失な動作のために、m4aのメタデータ値はm4のと同じである（すなわち、いずれも目標レンダリング行列c4を記述する）が、その点に到達するための時間d4aはd4 - d4aだけ短縮されている。換言すれば、メタデータ・インスタンスm4aは、前のメタデータ・インスタンスm4のと同じであり、よってc3とc4の間の補間曲線は変更されない。しかしながら、新たな補間継続時間d4aはもとの継続時間d4より短い。これは、事実上、メタデータ・インスタンスのデータ・レートを増大させる。これは、誤り訂正のようなある種の状況において有益でありうる。

【0185】

無損失なメタデータ補間の第二の例が図10に示されている（そして、上記のように、以下の説明は対応するサイド情報フォーマットにも同様に当てはまる）。この例では、この例では、目標は二つのメタデータ・インスタンスm3およびm4の間にメタデータm3aの新たなセットを含めることである。図10は、ある時間期間にわたってレンダリング行列が不変のままである場合を示している。したがって、この状況では、新たなセットのメタデータm3aの値は、補間継続時間m3aを除いて以前のメタデータm3のものと同じである。補間継続時間d3aの値は、t4～t3aに対応する値に、すなわち次のメタデータ・インスタンスm4に関連付けられた時刻t4と新たなセットのメタデータm3aに関連付けられた時刻t3aとの間の差に設定されるべきである。図10に示される場合は、たとえば、オーディオ・オブジェクトが静的であり、オーサリング・ツールがこの静的な性質のためそのオブジェクトについての新しいメタデータを送るのを停止するときに生じうる。そのような場合、たとえばメタデータをコーデック・フレームと同期させるために、新たなメタデータ・インスタンスm3aを挿入することが望ましいことがありうる。

【0186】

図8ないし図10に示した例では、現在から望まれるレンダリング行列またはレンダリング状態への補間は、線形補間によって実行された。他の例示的实施形態では、異なる補間方式が使われてもよい。あるそのような代替的な補間方式は、サンプルホールド回路をその後の低域通過フィルタと組み合わせたものを使う。図11は、ある例示的な実施形態に基づく、低域通過フィルタと一緒にサンプルホールド回路を使う補間方式を示している（そして、上記のように、以下の説明は対応するサイド情報フォーマットにも同様に当てはまる）。図11に示されるように、メタデータ・インスタンスm2ないしm4は、サンプルホールド・レンダリング行列係数c2およびc3に変換される。サンプルホールド・プロセスは、係数状態をすぐに所望される状態にジャンプさせ、その結果、図のようなステップ毎の曲線1110となる。この曲線1110はその後、低域通過フィルタリングされて、なめらかな、補間された曲線1120が得られる。補間フィルタ・パラメータ（たとえば、カットオフ周波数または時定数）は、時間サンプルおよび補間継続時間パラメータに加えて、メタデータの一部として信号伝達されることができる。システムの要求およびオーディオ信号の特性に依存して異なるパラメータが使用されてもよいことが理解される。

【0187】

ある例示的实施形態では、補間継続時間またはランブ・サイズは、0の値または実質的に0に近い値を含むいかなる実際的な値をもつこともできる。そのような小さな補間継続時間は、ファイルの最初のサンプルにおいてすぐにレンダリング行列を設定することまたはストリームの編集、スプライシングまたは連結を許容することを可能にするための初期化のような場合に特に有用である。この型の破壊的編集では、レンダリング行列を瞬間的に変更できることは、編集後のコンテンツの空間的属性を維持するために有益である。

【0188】

ある例示的实施形態では、本稿に記載される補間方式は、メタデータ・ビットレートを低下させる間引き方式におけるように、メタデータ・インスタンスの除去（および同様に上記のようにサイド情報インスタンスの除去）と互換である。メタデータ・インスタンスの除去は、システムが初期フレーム・レートより低いフレーム・レートで再サンプリングすることを許容する。この場合、エンコーダによって提供されるメタデータ・インスタン

スおよびその関連付けられた補間継続時間データは、ある種の特性に基づいて除去される。たとえば、エンコーダにおける解析コンポーネントは、信号の有意な停滞の期間があるかどうかを判定するためにオーディオ信号を解析し、そのような場合、デコーダ側へのデータの伝送の帯域幅要求を減らすために、すでに生成されたある種のメタデータ・インスタンスを除去してもよい。メタデータ・インスタンスの除去は、代替的または追加的に、デコーダにおいてまたはトランスコードにおいてなど、エンコーダとは別個のコンポーネントにおいて実行されてもよい。トランスコードは、エンコーダによって生成または追加されたメタデータ・インスタンスを除去してもよく、オーディオ信号を第一のレートから第二のレートに再サンプリングするデータ・レート変換器において用いられてもよい。ここで、第二のレートは第一のレートの整数倍であってもなくてもよい。どのメタデータ・インスタンスを除去するかを決定するためにオーディオ信号を解析する代わりに、エンコーダ、デコーダまたはトランスコードはメタデータを解析してもよい。たとえば、図10を参照するに、第一のメタデータ・インスタンスm3によって指定される第一の所望される再構成設定c3（または再構成行列）と、第一のメタデータ・インスタンスm3の直後のメタデータ・インスタンスm3aおよびm4によって指定される所望される再構成設定c3aおよびc4（または再構成行列）との間の差が計算されてもよい。この差はたとえば、それぞれのレンダリング行列に行列ノルムを用いることによって計算されてもよい。差が、たとえば再構成されるオーディオ信号の認容される歪みに対応する、あらかじめ定義された閾値より低ければ、第一のメタデータ・インスタンスm2に続くメタデータ・インスタンスm3aおよびm4は除去されてもよい。図10に示した例では、第一のメタデータ・インスタンスm3の直後のメタデータ・インスタンスm3aは第一のメタデータ・インスタンスm3aと同じレンダリング設定を指定し（ $c3 = c3a$ ）、よって除去される。一方、次のメタデータ設定m4は異なるレンダリング設定c4を指定し、用いられる閾値に依存して、メタデータとして保持されてもよい。

【0189】

図2を参照して記述されたデコーダ200では、オブジェクト再構成コンポーネント206は、M個のダウンミックス信号224およびサイド情報に基づいてN個のオーディオ・オブジェクト220を再構成することの一部として補間を用いてもよい。図7～図11を参照して述べた補間方式と同様に、N個のオーディオ・オブジェクト220の再構成はたとえば：現在の再構成設定に基づいて再構成を実行し；サイド情報インスタンスについて遷移データによって定義される時点において、現在の再構成設定から前記サイド情報インスタンスによって指定される所望される再構成設定への遷移を開始し；前記サイド情報インスタンスについて前記遷移データによって定義される時点において前記所望される再構成設定への前記遷移を完了することを含んでいてもよい。

【0190】

同様に、レンダラー210は、再生のために好適なマルチチャネル出力信号230を生成するために、再構成されたN個のオーディオ・オブジェクト220をレンダリングすることの一部として補間を用いてもよい。図7～図11を参照して述べた補間方式と同様に、レンダリングは：現在のレンダリング設定に基づいてレンダリングを実行し；クラスター・メタデータ・インスタンスについて遷移データによって定義される時点において、現在のレンダリング設定から前記クラスター・メタデータ・インスタンスによって指定される所望されるレンダリング設定への遷移を開始し；前記クラスター・メタデータ・インスタンスについて前記遷移データによって定義される時点において前記所望されるレンダリング設定への前記遷移を完了することを含んでいてもよい。

【0191】

いくつかの実施形態では、オブジェクト再構成部206およびレンダラー210は分離可能なユニットであってもよく、および/または別個のプロセスとして実行される動作に対応していてもよい。他の例示的实施形態では、オブジェクト再構成部206およびレンダラー210は、単一のユニットまたはプロセスとして具現され、再構成およびレンダリングが組み合わされた動作として実行されるのでもよい。そのような例示的实施形態では

、レンダリング行列および再構成行列に対して別個に補間を実行するのではなく、再構成およびレンダリングのために用いられる行列は単一の行列に組み合わされて、それが補間されてもよい。

【0192】

図3を参照して記述される低計算量デコーダ300では、レンダラー310が、M個のダウンミックス信号324をマルチチャネル出力330にレンダリングすることの一部として補間を実行してもよい。図7～図11を参照して述べた補間方式と同様に、レンダリングは：現在のダウンミックス・レンダリング設定に基づいてレンダリングを実行し；ダウンミックス・メタデータ・インスタンスについて遷移データによって定義される時点において、現在のダウンミックス・レンダリング設定から前記ダウンミックス・メタデータ・インスタンスによって指定される所望されるダウンミックス・レンダリング設定への遷移を開始し；前記ダウンミックス・メタデータ・インスタンスについて前記遷移データによって定義される時点において前記所望されるダウンミックス・レンダリング設定への前記遷移を完了することを含んでいてもよい。先述したように、レンダラー310は、デコーダ300に含まれてもよく、あるいは別個の装置／ユニットであってもよい。レンダラー310がデコーダ300とは別個である例示的实施形態では、デコーダは、レンダラー310におけるM個のダウンミックス信号のレンダリングのために、ダウンミックス・メタデータ325およびM個のダウンミックス信号324を出力してもよい。

【0193】

等価物、拡張、代替その他

上記の記述を吟味すれば、当業者には本開示のさらなる実施形態が明白になるであろう。本稿および図面は実施形態および例を開示しているが、本開示はこれらの個別的な例に制約されるものではない。付属の請求項によって定義される本開示の範囲から外れることなく数多くの修正および変形をなすことができる。請求項に現われる参照符号があったとしても、その範囲を限定するものと理解されるものではない。

【0194】

さらに、図面、本開示および付属の請求項の吟味から、本開示を実施する当業者によって、開示される実施形態に対する変形が理解され、実施されることができる。請求項において、「有する／含む」の語は他の要素またはステップを排除するものではなく、単数形の表現は複数を排除するものではない。ある種の施策が互いに異なる従属請求項に記載されているというだけの事実がこれらの施策の組み合わせが有利に使用できないことを示すものではない。

【0195】

上記で開示されたシステムおよび方法は、ソフトウェア、ファームウェア、ハードウェアまたはそれらの組み合わせとして実装されうる。ハードウェア実装では、上記の記述で言及された機能ユニットの間でのタスクの分割は必ずしも物理的なユニットへの分割に対応しない。むしろ、一つの物理的コンポーネントが複数の機能を有していてもよく、一つのタスクが協働していくつかの物理的コンポーネントによって実行されてもよい。ある種のコンポーネントまたはすべてのコンポーネントは、デジタル信号プロセッサまたはマイクロプロセッサによって実行されるソフトウェアとして実装されてもよく、あるいはハードウェアとしてまたは特定用途向け集積回路として実装されてもよい。そのようなソフトウェアは、コンピュータ記憶媒体（または非一時的な媒体）および通信媒体（または一時的な媒体）を含みうるコンピュータ可読媒体上で頒布されてもよい。当業者にはよく知られているように、コンピュータ記憶媒体という用語は、コンピュータ可読命令、データ構造、プログラム・モジュールまたは他のデータのような情報の記憶のための任意の方法または技術において実装される揮発性および不揮発性、リムーバブルおよび非リムーバブル媒体を含む。コンピュータ記憶媒体は、これに限られないが、RAM、ROM、EEPROM、フラッシュメモリまたは他のメモリ技術、CD-ROM、デジタル多用途ディスク（DVD）または他の光ディスク記憶、磁気カセット、磁気テープ、磁気ディスク記憶または他の磁気記憶デバイスまたは、所望される情報を記憶するために使用されることができ、コンピュータによ

ってアクセスされることができる他の任意の媒体を含む。さらに、通信媒体が典型的にはコンピュータ可読命令、データ構造、プログラム・モジュールまたは他のデータを、搬送波または他の転送機構のような変調されたデータ信号において具現し、任意の情報送達媒体を含むことは当業者にはよく知られている。

【0196】

すべての図面は概略的であり、一般に、本開示を明快にするために必要な部分を示すだけである。一方、他の部分は省略されたり、単に示唆するに留めたりすることもある。特に断わりのない限り、同様の参照符号は異なる図面において同様の部分を指す。

【0197】

いくつかの態様を記載しておく。

10

〔態様1〕

オーディオ・オブジェクトをデータ・ストリームとしてエンコードする方法であって：

N>1であるとして、N個のオーディオ・オブジェクトを受領する段階と；

M=Nであるとして、前記N個のオーディオ・オブジェクトの組み合わせを形成することによって、M個のダウンミックス信号を計算する段階と；

前記M個のダウンミックス信号からの、前記N個のオーディオ・オブジェクトに基づいて形成された一組のオーディオ・オブジェクトの再構成を許容するパラメータを含む時間可変のサイド情報を計算する段階と；

前記M個のダウンミックス信号および前記サイド情報を、デコーダに伝送するためのデータ・ストリームに含める段階とを含み、

20

本方法はさらに、前記データ・ストリーム中に：

前記N個のオーディオ・オブジェクトに基づいて形成された前記一組のオーディオ・オブジェクトを再構成するためのそれぞれの所望される再構成設定を指定する複数のサイド情報インスタンスと；

各サイド情報インスタンスについての、現在の再構成設定から、そのサイド情報インスタンスによって指定される所望される再構成設定への遷移を開始する時点および該遷移を完了する時点を組み合わせて定義する、二つの独立に割り当て可能な部分を含む遷移データとを含めることを含む、

方法。

〔態様2〕

30

第一の複数のオーディオ・オブジェクトを第二の複数のオーディオ・オブジェクトに減らすためのクラスタリング手順をさらに含み、前記N個のオーディオ・オブジェクトは、前記第一の複数のオーディオ・オブジェクトまたは前記第二の複数のオーディオ・オブジェクトをなし、前記N個のオーディオ・オブジェクトに基づいて形成された前記一組のオーディオ・オブジェクトは前記第二の複数のオーディオ・オブジェクトに一致し、前記クラスタリング手順は：

前記第二の複数のオーディオ・オブジェクトについての空間位置を含む時間可変なクラスター・メタデータを計算する段階と；

前記データ・ストリーム中に：

第二の組のオーディオ・オブジェクトをレンダリングするためのそれぞれの所望されるレンダリング設定を指定する複数のクラスター・メタデータ・インスタンスと；

40

各クラスター・メタデータ・インスタンスについての、現在のレンダリング設定から、そのクラスター・メタデータ・インスタンスによって指定される所望されるレンダリング設定への遷移を開始する時点およびそのクラスター・メタデータ・インスタンスによって指定される前記所望されるレンダリング設定への前記遷移を完了する時点とを、組み合わせて定義する、二つの独立に割り当て可能な部分を含む遷移データとを含める段階とを含む、

態様1記載の方法。

〔態様3〕

前記クラスタリング手順はさらに：

50

前記第一の複数のオーディオ・オブジェクトおよびそれらの関連付けられた空間位置を受領する段階と；

前記第一の複数のオーディオ・オブジェクトを、前記第一の複数のオーディオ・オブジェクトの空間的近接性に基づいて少なくとも一つのクラスターと関連付ける段階と；

前記少なくとも一つのクラスターのそれぞれを、そのクラスターに関連付けられたオーディオ・オブジェクトの組み合わせであるオーディオ・オブジェクトによって表わすことによって、前記第二の複数のオーディオ・オブジェクトを生成する段階と；

前記第二の複数のオーディオ・オブジェクトの各オーディオ・オブジェクトの空間位置を、そのオーディオ・オブジェクトが代表するクラスターに関連付けられた諸オーディオ・オブジェクトの空間位置に基づいて計算する段階とを含む、

10

態様 2 記載の方法。

〔態様 4〕

それぞれのクラスター・メタデータ・インスタンスについて前記遷移データによって定義されるそれぞれの時点は、対応するサイド情報インスタンスについて前記遷移データによって定義されるそれぞれの時点と一致する、態様 2 または 3 記載の方法。

〔態様 5〕

前記N個のオーディオ・オブジェクトが前記第二の複数のオーディオ・オブジェクトをなす、態様 2 ないし 4 のうちいずれか一項記載の方法。

〔態様 6〕

前記N個のオーディオ・オブジェクトが前記第一の複数のオーディオ・オブジェクトをなす、態様 2 ないし 4 のうちいずれか一項記載の方法。

20

〔態様 7〕

当該方法が、

各ダウンミックス信号を、前記ダウンミックス信号をレンダリングするための時間可変な空間位置に関連付ける段階と；

前記ダウンミックス信号の前記空間位置を含むダウンミックス・メタデータを前記データ・ストリームにさらに含める段階とをさらに含み、

当該方法がさらに、前記データ・ストリームに；

前記ダウンミックス信号をレンダリングするためのそれぞれの所望されるダウンミックス・レンダリング設定を指定する複数のダウンミックス・メタデータ・インスタンスと；

30

各ダウンミックス・メタデータ・インスタンスについての、現在のダウンミックス・レンダリング設定から、そのダウンミックス・メタデータ・インスタンスによって指定される所望されるダウンミックス・レンダリング設定への遷移を開始する時点および前記ダウンミックス・メタデータ・インスタンスによって指定される所望されるダウンミックス・レンダリング設定への遷移を完了する時点とを組み合わせさせて定義する、二つの独立に割り当て可能な部分を含む遷移データとを含めることを含む、

態様 1 ないし 6 のうちいずれか一項記載の方法。

〔態様 8〕

それぞれのダウンミックス・メタデータ・インスタンスについての前記遷移データによって定義されるそれぞれの時点は、対応するサイド情報インスタンスについての前記遷移データによって定義されるそれぞれの時点と一致する、態様 7 記載の方法。

40

〔態様 9〕

$N > 1$ として、N個のオーディオ・オブジェクトをデータ・ストリームとしてエンコードするエンコーダであって；

M Nであるとして、前記N個のオーディオ・オブジェクトの組み合わせを形成することによって、M個のダウンミックス信号を計算するよう構成されているダウンミックス・コンポーネントと；

前記M個のダウンミックス信号からの、前記N個のオーディオ・オブジェクトに基づいて形成された一組のオーディオ・オブジェクトの再構成を許容するパラメータを含む時間可変のサイド情報を計算するよう構成されている解析コンポーネントと；

50

前記M個のダウンミックス信号および前記サイド情報を、デコーダに伝送するためのデータ・ストリームに含めるよう構成された多重化コンポーネントとを有しており、

前記多重化コンポーネントは、前記データ・ストリーム中に：

前記N個のオーディオ・オブジェクトに基づいて形成された前記一組のオーディオ・オブジェクトを再構成するためのそれぞれの所望される再構成設定を指定する複数のサイド情報インスタンスと；

各サイド情報インスタンスについての、現在の再構成設定から、そのサイド情報インスタンスによって指定される所望される再構成設定への遷移を開始する時点および該遷移を完了する時点を組み合わさって定義する、二つの独立に割り当て可能な部分を含む遷移データとを含めるよう構成される、

エンコーダ。

〔態様 10〕

データ・ストリームに基づいてオーディオ・オブジェクトを再構成する方法であって：

N>1およびM Nであるとして、N個のオーディオ・オブジェクトの組み合わせであるM個のダウンミックス信号と、前記M個のダウンミックス信号からの、前記N個のオーディオ・オブジェクトに基づいて形成された一組のオーディオ・オブジェクトの再構成を許容するパラメータを含む時間可変なサイド情報とを含むデータ・ストリームを受領する段階と；

前記M個のダウンミックス信号および前記サイド情報に基づいて、前記N個のオーディオ・オブジェクトに基づいて形成された前記一組のオーディオ・オブジェクトを再構成する段階とを含み、

前記データ・ストリームは、複数のサイド情報インスタンスを含み、前記データ・ストリームはさらに、各サイド情報インスタンスについて、現在の再構成設定から、そのサイド情報インスタンスによって指定される所望される再構成設定への遷移を開始する時点および該遷移を完了する時点を組み合わさって定義する、二つの独立に割り当て可能な部分を含む遷移データを含み、

前記N個のオーディオ・オブジェクトに基づいて形成された前記一組のオーディオ・オブジェクトを再構成する段階は：

現在の再構成設定に従って再構成を実行し；

あるサイド情報インスタンスについての前記遷移データによって定義される時点において、前記現在の再構成設定から、そのサイド情報インスタンスによって指定される所望される再構成設定への遷移を開始し；

そのサイド情報インスタンスについての前記遷移データによって定義される時点において前記遷移を完了することを含む、方法。

〔態様 11〕

前記データ・ストリームはさらに、前記N個のオーディオ・オブジェクトに基づいて形成された前記一組のオーディオ・オブジェクトについての時間可変なクラスター・メタデータを含み、前記クラスター・メタデータは、前記N個のオーディオ・オブジェクトに基づいて形成された前記一組のオーディオ・オブジェクトについての空間位置を含み、前記データ・ストリームは、複数のクラスター・メタデータ・インスタンスを含み、前記データ・ストリームはさらに、各クラスター・メタデータ・インスタンスについて、現在のレンダリング設定から、そのクラスター・メタデータ・インスタンスによって指定される所望されるレンダリング設定への遷移を開始する時点およびそのクラスター・メタデータ・インスタンスによって指定される前記所望されるレンダリング設定への前記遷移を完了する時点と、組み合わさって定義する、二つの独立に割り当て可能な部分を含む遷移データを含み、

当該方法はさらに：

前記N個のオーディオ・オブジェクトに基づいて形成された再構成された一組のオーディオ・オブジェクトを、あらかじめ定義されたチャネル配位の出力チャネルにレンダリングするために前記クラスター・メタデータを使うことを含み、該レンダリングは：

10

20

30

40

50

現在のレンダリング設定に従ってレンダリングを実行し；

あるクラスター・メタデータ・インスタンスについての前記遷移データによって定義される時点において、前記現在のレンダリング設定から、そのクラスター・メタデータ・インスタンスによって指定される所望されるレンダリング設定への遷移を開始し；

そのクラスター・メタデータ・インスタンスについての前記遷移データによって定義される時点において前記所望されるレンダリング設定への前記遷移を完了することを含む、態様 10 記載の方法。

〔態様 12〕

それぞれのクラスター・メタデータ・インスタンスについての遷移データによって定義されるそれぞれの時点は、対応するサイド情報インスタンスについて前記遷移データによって定義されるそれぞれの時点と一致する、態様 11 記載の方法。

10

〔態様 13〕

当該方法はさらに：

前記再構成および前記レンダリングの少なくとも一部を、それぞれ現在の再構成設定および現在のレンダリング設定に関連する再構成行列およびレンダリング行列の行列積として形成される第一の行列に対応する組み合わせられた動作として実行し；

あるサイド情報インスタンスおよびあるクラスター・メタデータ・インスタンスについての前記遷移データによって定義される時点において、前記現在の再構成およびレンダリング設定から、それぞれそのサイド情報インスタンスおよびそのクラスター・メタデータ・インスタンスによって指定される所望される再構成およびレンダリング設定への組み合わせられた遷移を開始し；

20

そのサイド情報インスタンスおよびそのクラスター・メタデータ・インスタンスについての前記遷移データによって定義される時点において前記組み合わせられた遷移を完了することを含み、前記組み合わせられた遷移は、前記第一の行列の行列要素と、それぞれ前記所望される再構成設定および前記所望されるレンダリング設定に関連する再構成行列およびレンダリング行列の行列積として形成される第二の行列の行列要素との間を補間することを含む、

態様 12 記載の方法。

〔態様 14〕

前記N個のオーディオ・オブジェクトに基づいて形成された前記一組のオーディオ・オブジェクトは、前記N個のオーディオ・オブジェクトと一致する、態様 10 ないし 13 のうちいずれか一項記載の方法。

30

〔態様 15〕

前記N個のオーディオ・オブジェクトに基づいて形成された前記一組のオーディオ・オブジェクトは、前記N個のオーディオ・オブジェクトの組み合わせである、N個より少ない複数のオーディオ・オブジェクトを含む、態様 10 ないし 13 のうちいずれか一項記載の方法。

〔態様 16〕

デコーダにおいて実行される態様 10 ないし 15 のうちいずれか一項記載の方法であって、前記データ・ストリームはさらに、前記M個のダウンミックス信号に関連付けられた時間可変な空間位置を含む前記M個のダウンミックス信号についてのダウンミックス・メタデータを含み、前記データ・ストリームはさらに、複数のダウンミックス・メタデータ・インスタンスを含み、前記データ・ストリームはさらに、各ダウンミックス・メタデータ・インスタンスについて、現在のダウンミックス・レンダリング設定から、そのダウンミックス・メタデータ・インスタンスによって指定される所望されるダウンミックス・レンダリング設定への遷移を開始する時点およびそのダウンミックス・メタデータ・インスタンスによって指定される前記所望されるダウンミックス・レンダリング設定への前記遷移を完了する時点を、組み合わせさせて定義する、二つの独立に割り当て可能な部分を含む遷移データを含み、

40

当該方法はさらに：

50

前記デコーダがオーディオ・オブジェクト再構成をサポートするよう動作可能である場合、再構成する段階を、前記M個のダウンミックス信号および前記サイド情報、前記N個のオーディオ・オブジェクトに基づいて形成された前記一組のオーディオ・オブジェクトに基づいて実行し；

前記デコーダがオーディオ・オブジェクト再構成をサポートするよう動作可能でない場合、前記ダウンミックス・メタデータおよび前記M個のダウンミックス信号を前記M個のダウンミックス信号のレンダリングのために出力することを含む、方法。

〔態様 17〕

データ・ストリームに基づいてオーディオ・オブジェクトを再構成するデコーダであって；

N>1およびM Nであるとして、N個のオーディオ・オブジェクトの組み合わせであるM個のダウンミックス信号と、前記M個のダウンミックス信号からの、前記N個のオーディオ・オブジェクトに基づいて形成された一組のオーディオ・オブジェクトの再構成を許容するパラメータを含む時間可変なサイド情報とを含むデータ・ストリームを受領するよう構成されている受領コンポーネントと；

前記M個のダウンミックス信号および前記サイド情報に基づいて、前記N個のオーディオ・オブジェクトに基づいて形成された前記一組のオーディオ・オブジェクトを再構成するよう構成されている再構成コンポーネントとを有し、

前記データ・ストリームは、複数のサイド情報インスタンスを含み、前記データ・ストリームはさらに、各サイド情報インスタンスについて、現在の再構成設定から、そのサイド情報インスタンスによって指定される所望される再構成設定への遷移を開始する時点および該遷移を完了する時点を組み合わせて定義する、二つの独立に割り当て可能な部分を含む遷移データを含み、

前記再構成コンポーネントは、前記N個のオーディオ・オブジェクトに基づいて形成された前記一組のオーディオ・オブジェクトを再構成することを、少なくとも；

現在の再構成設定に従って再構成を実行し；

あるサイド情報インスタンスについての前記遷移データによって定義される時点において、前記現在の再構成設定から、そのサイド情報インスタンスによって指定される所望される再構成設定への遷移を開始し；

そのサイド情報インスタンスについての前記遷移データによって定義される時点において前記遷移を完了することによって行なうよう構成されている、デコーダ。

〔態様 18〕

態様 1 ないし 8 および 10 ないし 16 のうちいずれか一項記載の方法であって、さらに；

一つまたは複数の追加的なサイド情報インスタンスであって、該一つまたは複数の追加的なサイド情報インスタンスの直前または直後のサイド情報インスタンスと実質的に同じ再構成設定を指定するものを生成する段階を含む、

方法。

〔態様 19〕

データ・ストリーム中でM個のオーディオ信号と一緒にエンコードされたサイド情報をトランスコードする方法であって；

データ・ストリームを受領する段階と；

前記データ・ストリームから、M 1として、M個のオーディオ信号と、前記M個のオーディオ信号からの、一組のオーディオ・オブジェクトの再構成を許容するパラメータを含む、関連する時間可変なサイド情報とを抽出する段階であって、抽出されたサイド情報は；

前記オーディオ・オブジェクトを再構成するためのそれぞれの所望される再構成設定を指定する複数のサイド情報インスタンス；および

各サイド情報インスタンスについて、現在の再構成設定から、そのサイド情報インス

10

20

30

40

50

タンスによって指定される所望される再構成設定への遷移を開始する時点および該遷移を完了する時点を組み合わせて定義する、二つの独立に割り当て可能な部分を含む遷移データを含む、段階と；

一つまたは複数の追加的なサイド情報インスタンスであって、該一つまたは複数の追加的なサイド情報インスタンスの直前または直後のサイド情報インスタンスと実質的に同じ再構成設定を指定する追加的なサイド情報インスタンスを生成する段階と；

前記M個のオーディオ信号および前記サイド情報をデータ・ストリームに含める段階とを含む、方法。

〔態様 20〕

前記M個のオーディオ信号は、第一のフレーム・レートに従って、受領されたデータ・ストリーム中で符号化されており、当該方法はさらに：

前記M個のオーディオ信号を処理して、前記M個のダウンミックス信号が符号化されているフレーム・レートを前記第一のフレーム・レートとは異なる第二のフレーム・レートに変える段階と；

少なくとも、前記一つまたは複数の追加的なサイド情報インスタンスを生成することによって、前記第二のフレーム・レートに一致するよう、前記サイド情報を再サンプリングする段階とを含む、

態様 19 記載の方法。

〔態様 21〕

データ・ストリームにおいてM個のオーディオ信号と一緒にエンコードされたサイド情報をトランスコードする装置であって：

データ・ストリームを受領し、前記データ・ストリームから、M-1として、M個のオーディオ信号と、前記M個のオーディオ信号からの、一組のオーディオ・オブジェクトの再構成を許容するパラメータを含む、関連する時間可変なサイド情報とを抽出するよう構成された受領コンポーネントを有しており、抽出されるサイド情報は：

前記オーディオ・オブジェクトを再構成するためのそれぞれの所望される再構成設定を指定する複数のサイド情報インスタンス；および

各サイド情報インスタンスについて、現在の再構成設定から、そのサイド情報インスタンスによって指定される所望される再構成設定への遷移を開始する時点および該遷移を完了する時点を組み合わせて定義する、二つの独立に割り当て可能な部分を含む遷移データを含み、

当該装置はさらに：

一つまたは複数の追加的なサイド情報インスタンスであって、該一つまたは複数の追加的なサイド情報インスタンスの直前または直後のサイド情報インスタンスと実質的に同じ再構成設定を指定する追加的なサイド情報インスタンスを生成するよう構成されている再サンプリング・コンポーネントと；

前記M個のオーディオ信号および前記サイド情報をデータ・ストリームに含めるよう構成されている多重化コンポーネントとを有する、装置。

〔態様 22〕

態様 1 ないし 8、10 ないし 16 および 18 ないし 20 のうちいずれか一項記載の方法であって、さらに：

第一のサイド情報インスタンスによって指定される第一の所望される再構成設定と、前記第一のサイド情報インスタンスの直後の一つまたは複数のサイド情報インスタンスによって指定される一つまたは複数の所望される再構成設定との間の差を計算し；

計算された差があらかじめ定義された閾値より低いことに応答して前記一つまたは複数のサイド情報インスタンスを除去することを含む、方法。

〔態様 23〕

10

20

30

40

50

態様 1 ないし 8、10 ないし 16、18 ないし 20 および 22 のうちいずれか一項記載の方法、態様 9 記載のエンコーダ、態様 17 記載のデコーダまたは態様 21 記載の装置であって：

各サイド情報インスタンスについての前記遷移データの前記二つの独立に割り当て可能な部分は：

前記所望される再構成設定への遷移を開始する時点を示すタイムスタンプと、前記所望される再構成設定への遷移を完了する時点を示すタイムスタンプ；

前記所望される再構成設定への遷移を開始する時点を示すタイムスタンプと、前記所望される再構成設定への遷移を開始する時点から前記所望される再構成設定に達するための継続時間を示す補間継続時間パラメータ；または

前記所望される再構成設定への遷移を完了する時点を示すタイムスタンプと、前記所望される再構成設定への遷移を開始する時点から前記所望される再構成設定に達するための継続時間を示す補間継続時間パラメータである、

方法、エンコーダ、デコーダまたは装置。

〔態様 24〕

態様 2 ないし 8、11 ないし 16、18 および 22 ないし 23 のうちいずれか一項記載の方法であって：

各クラスター・メタデータ・インスタンスについての前記遷移データの前記二つの独立に割り当て可能な部分は：

前記所望されるレンダリング設定への遷移を開始する時点を示すタイムスタンプと、前記所望されるレンダリング設定への遷移を完了する時点を示すタイムスタンプ；

前記所望されるレンダリング設定への遷移を開始する時点を示すタイムスタンプと、前記所望されるレンダリング設定への遷移を開始する時点から前記所望されるレンダリング設定に達するための継続時間を示す補間継続時間パラメータ；または

前記所望されるレンダリング設定への遷移を完了する時点を示すタイムスタンプと、前記所望されるレンダリング設定への遷移を開始する時点から前記所望されるレンダリング設定に達するための継続時間を示す補間継続時間パラメータである、

方法。

〔態様 25〕

態様 7 ないし 8、16、18 および 22 ないし 24 のうちいずれか一項記載の方法であって、各ダウンミックス・メタデータ・インスタンスについての前記遷移データの前記二つの独立に割り当て可能な部分は：

前記所望されるダウンミックス・レンダリング設定への遷移を開始する時点を示すタイムスタンプと、前記所望されるダウンミックス・レンダリング設定への遷移を完了する時点を示すタイムスタンプ；

前記所望されるダウンミックス・レンダリング設定への遷移を開始する時点を示すタイムスタンプと、前記所望されるダウンミックス・レンダリング設定への遷移を開始する時点から前記所望されるダウンミックス・レンダリング設定に達するための継続時間を示す補間継続時間パラメータ；または

前記所望されるダウンミックス・レンダリング設定への遷移を完了する時点を示すタイムスタンプと、前記所望されるダウンミックス・レンダリング設定への遷移を開始する時点から前記所望されるダウンミックス・レンダリング設定に達するための継続時間を示す補間継続時間パラメータである、

方法。

〔態様 26〕

態様 1 ないし 8、10 ないし 16、18 ないし 20 および 22 ないし 25 のうちいずれか一項記載の方法を実行するための命令をもつコンピュータ可読媒体を含むコンピュータ・プログラム・プロダクト。

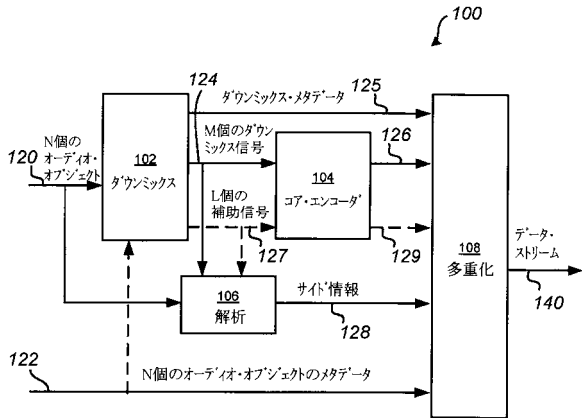
10

20

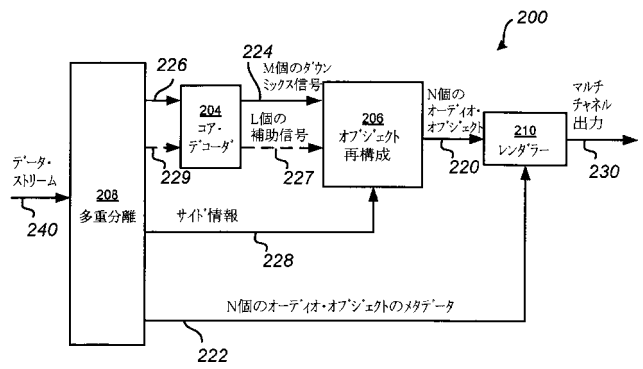
30

40

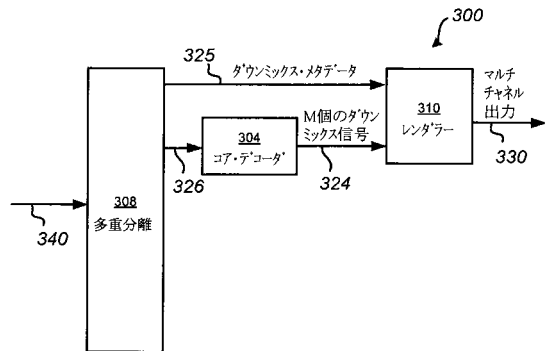
【 図 1 】



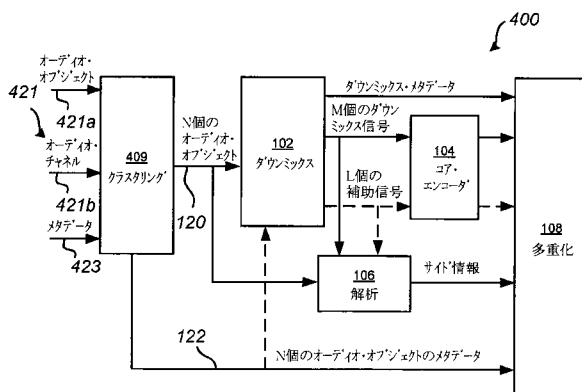
【 図 2 】



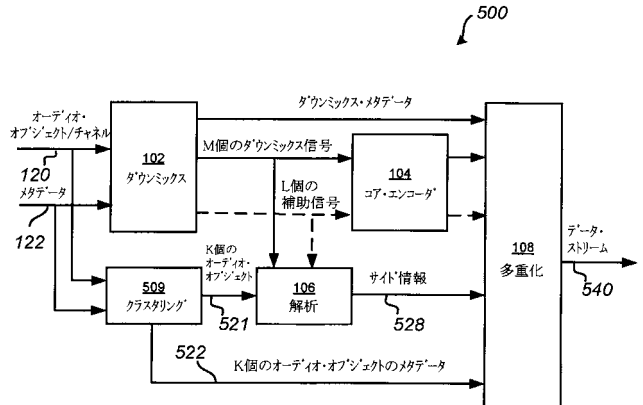
【 図 3 】



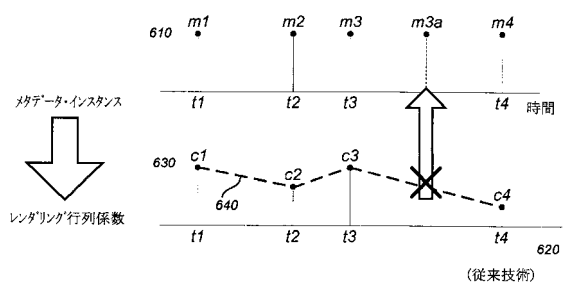
【 図 4 】



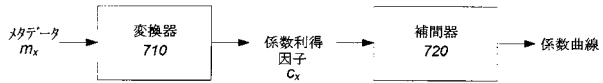
【 図 5 】



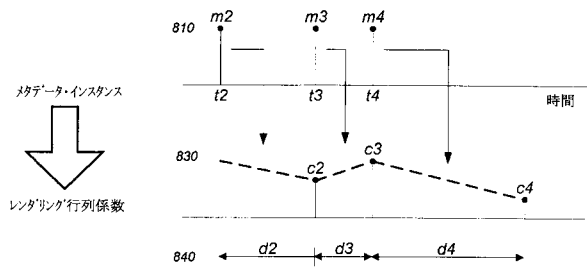
【 図 6 】



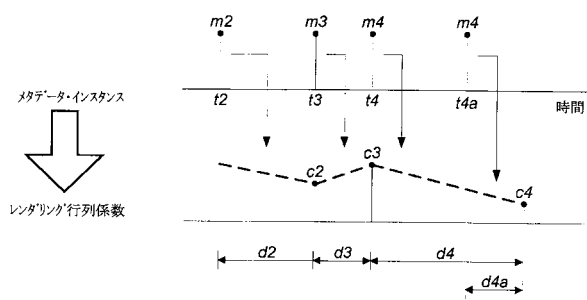
【図 7】



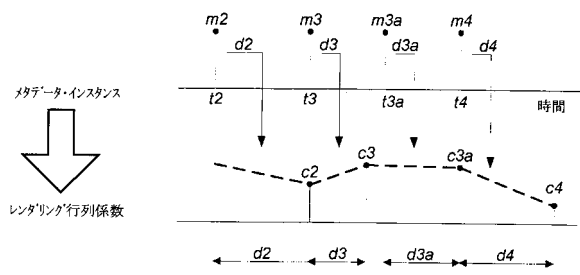
【図 8】



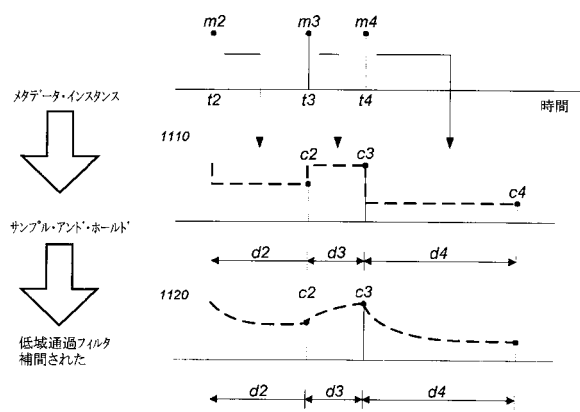
【図 9】



【図 10】



【図 11】



フロントページの続き

- (72)発明者 ブルンハーゲン, ヘイコ
スウェーデン王国, 1 1 3 3 0 ストックホルム, ガヴレガタン 1 2 アー ドルビー スウェーデン アーバー内
- (72)発明者 クヨーリング, クリストファー
スウェーデン王国, 1 1 3 3 0 ストックホルム, ガヴレガタン 1 2 アー ドルビー スウェーデン アーバー内
- (72)発明者 ヒルヴォーネン, トニ
スウェーデン国, 1 1 3 3 0 ストックホルム, ガヴレギヤタン 1 2 アー, ドルビー・スウェーデン・アーバー内
- (72)発明者 ヴィレモーズ, ラルス
スウェーデン王国, 1 1 3 3 0 ストックホルム, ガヴレガタン 1 2 アー ドルビー スウェーデン アーバー内
- (72)発明者 ブリーバート, ディルク ジェローエン
オーストラリア国, ニュー・サウス・ウェールズ 2 0 6 0, マクマホンズ・ポイント, ミッチェル・ストリート 3 5 レベル 3 ドルビー・オーストラリア・ピーティーワイ・リミテッド内
- F ターム(参考) 5D162 AA11 BA03 BA13 CB17 CD01 EE05 EG02

【外国語明細書】
2017199034000001.pdf