

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第6190942号  
(P6190942)

(45) 発行日 平成29年8月30日(2017.8.30)

(24) 登録日 平成29年8月10日(2017.8.10)

(51) Int.Cl. F I  
**G 1 O L 19/02 (2013.01)** G 1 O L 19/02 1 7 O Z  
**G 1 O L 19/022 (2013.01)** G 1 O L 19/022

請求項の数 12 (全 23 頁)

(21) 出願番号	特願2016-505840 (P2016-505840)	(73) 特許権者	510185767
(86) (22) 出願日	平成26年4月4日(2014.4.4)		ドルビー・インターナショナル・アーベー
(65) 公表番号	特表2016-515722 (P2016-515722A)		オランダ王国, セーエン アムステルダム
(43) 公表日	平成28年5月30日(2016.5.30)		ズイドーフト 1101, ヘリケルベル
(86) 国際出願番号	PCT/EP2014/056848		グウェグ 1-35, アポロ ビルディング
(87) 国際公開番号	W02014/161990		グ 3エー
(87) 国際公開日	平成26年10月9日(2014.10.9)	(74) 代理人	100107766
審査請求日	平成27年9月30日(2015.9.30)		弁理士 伊東 忠重
(31) 優先権主張番号	61/808,701	(74) 代理人	100070150
(32) 優先日	平成25年4月5日(2013.4.5)		弁理士 伊東 忠彦
(33) 優先権主張国	米国 (US)	(74) 代理人	100091214
			弁理士 大貫 進介

最終頁に続く

(54) 【発明の名称】 オーディオ・エンコーダおよびデコーダ

(57) 【特許請求の範囲】

【請求項1】

オーディオ・データを担持するオーディオ・フレームにセグメント分割されたオーディオ・ビットストリームを受け入れるよう構成されたオーディオ処理システムであって、当該オーディオ処理システムは：

フロントエンド・コンポーネントであって：

前記オーディオ・ビットストリーム中のあるオーディオ・フレームに対応する量子化されたスペクトル係数を繰り返し受領し、中間信号の第一の周波数領域表現を出力するよう適応された量子化解除段；および

前記中間信号の前記第一の周波数領域表現を受領し、それに基づいて前記中間信号の時間領域表現を合成する逆変換段とを含むフロントエンド・コンポーネントと；

処理段であって：

前記中間信号の前記時間領域表現を受領し、前記中間信号の第二の周波数領域表現を出力する分解フィルタバンク；

前記中間信号の前記第二の周波数領域表現を受領し、処理されたオーディオ信号の周波数領域表現を出力する少なくとも一つの処理コンポーネント；および

前記処理されたオーディオ信号の前記周波数領域表現を受領し、前記処理されたオーディオ信号の時間領域表現を出力する合成フィルタバンクとを含む処理段と；

前記処理されたオーディオ信号の前記時間領域表現を受領し、ターゲット・サンプリング周波数でサンプリングされた再構成されたオーディオ信号を出力するサンプル・レート

10

20

変換器とを有しており、

当該オーディオ処理システムは、少なくとも二つの異なるあらかじめ定義されたオーディオ・フレーム・レートのうちの任意のオーディオ・フレーム・レートをもつオーディオ・ビットストリームを受け入れるよう構成されており、前記フロントエンド・コンポーネントは可変の最大フレーム長をもち、前記異なるあらかじめ定義されたオーディオ・フレーム・レートに対応する少なくとも二つのモードで動作可能であり、前記あらかじめ定義されたオーディオ・フレーム・レートは、5%以下の変動範囲のオーディオ・フレーム・レートのサブグループを含み、前記フロントエンド・コンポーネントは、前記サブグループ中のすべてのフレーム・レートについて同じフレーム長を使うよう構成されている、ことを特徴とする、

10

オーディオ処理システム。

【請求項 2】

前記量子化解除段は、量子化インデックスを、各フレーム長に関連付けられた、非一様な周波数依存の量子化レベルにマッピングし、各周波数について人間の聴覚に対して知覚不能な量子化ノイズ割り当てを達成するよう構成されている、請求項 1 記載のオーディオ処理システム。

【請求項 3】

前記フロントエンド・コンポーネントが、各最大フレーム長について、該最大フレーム長の同サイズのブロックへの細分である代替的なフレーム長で動作可能である、請求項 1 または 2 記載のオーディオ処理システム。

20

【請求項 4】

前記オーディオ・ビットストリームの全ビットレートが、二チャンネル・オーディオ・データを担持するとき、200kbps未<sup>下</sup>満である、請求項 1 ないし 3 のうちいずれか一項記載のオーディオ処理システム。

【請求項 5】

前記分解フィルタバンクおよび合成フィルタバンクが固定数の周波数帯域をもつフィルタバンクである、請求項 1 ないし 4 のうちいずれか一項記載のオーディオ処理システム。

【請求項 6】

前記量子化解除段が、臨界サンプリングされた量子化解除されたスペクトル係数を含む前記中間オーディオ信号の量子化解除された周波数領域表現を出力するよう構成されており、

30

前記逆変換段が重複加算合成フィルタバンクを有する、請求項 1 ないし 5 のうちいずれか一項記載のオーディオ処理システム。

【請求項 7】

前記オーディオ・ビットストリームがマルチチャンネル・オーディオ信号をパラメトリック・エンコードするものである、請求項 1 ないし 6 のうちいずれか一項記載のオーディオ処理システム。

【請求項 8】

請求項 1 ないし 7 のうちいずれか一項記載のオーディオ処理システムと；

前記少なくとも二つの異なるあらかじめ定義されたオーディオ・フレーム・レートに対応する少なくとも二つの異なるあらかじめ定義されたビデオ・フレーム・レートのうちの任意のビデオ・フレーム・レートをもつ入来ビデオ・フレームのストリームを受け入れ、処理されたビデオ・フレームのストリームを出力するよう構成されたビデオ処理システムとを有するマルチメディア処理システムであって、前記あらかじめ定義されたオーディオ・フレーム・レートのそれぞれについて、一つまたは複数のオーディオ・フレームが一緒になって、対応するあらかじめ定義されたビデオ・フレーム・レートについての一つまたは複数のビデオ・フレームと等しい継続時間のオーディオ・データを担持し、

40

当該マルチメディア処理システムが、前記あらかじめ定義されたオーディオ・フレーム・レートのうちの任意のオーディオ・フレーム・レートをもつオーディオ・ビットストリームおよび対応するあらかじめ定義されたビデオ・フレーム・レートをもつ入来ビデオ・

50

フレームのストリームについて：

一つまたは複数のビデオ・フレームの複製を実行し、複製される一つまたは複数のビデオ・フレームと等しい継続時間のオーディオ・データを一緒になって担持する一つまたは複数のオーディオ・フレームの複製を実行する；および/または

一つまたは複数のビデオ・フレームの削除を実行し、削除される一つまたは複数のビデオ・フレームと等しい継続時間のオーディオ・データを一緒になって担持する一つまたは複数のオーディオ・フレームの削除を実行する

よう構成されている、

マルチメディア処理システム。

【請求項 9】

オーディオ・データを担持するオーディオ・フレームにセグメント分割されたオーディオ・ビットストリームを処理する方法であって、前記オーディオ・ビットストリームはフレームにセグメント分割されており、各ビデオ・フレームは、一つまたは複数のビデオ・フレームと等しい継続時間の関連付けられたオーディオ・データを担持する、前記オーディオ・ビットストリーム中のフレームに時間的に関係しており、当該方法は：

・前記オーディオ・ビットストリーム中のあるオーディオ・フレームに対応する量子化されたスペクトル係数を繰り返し受領し、周波数から時間への変換が後続する逆量子化を実行する段階であって、それにより中間オーディオ信号の表現が得られる、段階と；

・前記中間オーディオ信号に対して周波数領域で少なくとも一つの処理段階を実行する段階と；

・処理されたオーディオ信号のサンプリング・レートをターゲット・サンプリング周波数に変更する段階であって、それにより再構成されたオーディオ信号の時間領域表現が得られる、段階とを含み、

前記オーディオ・ビットストリームの現在フレーム・レートを確立するさらなる段階および前記逆量子化および周波数から時間への変換が、異なる最大フレーム長をもつ少なくとも二つのモードで動作可能な機能コンポーネント(112)において実行され、現在モードは、前記オーディオ・ビットストリームの現在フレーム・レートに応答して選択され、前記オーディオ・ビットストリームの確立された現在フレーム・レートが、いくつかの所定のフレーム・レートの一つに対してマッチングされ、前記所定のフレーム・レートは、5%以下の変動範囲のサブグループを含み、量子化されたスペクトル係数を繰り返し受領する前記の段階は、同じフレーム長をもつ前記機能コンポーネントの動作中に実行されること、を特徴とする、

方法。

【請求項 10】

前記逆量子化が、静的な第二の機能コンポーネント(114)において実行される、請求項9記載の方法。

【請求項 11】

前記オーディオ・ビットストリームがマルチチャンネル・オーディオ信号をパラメトリック・エンコードするものである、請求項9または10記載の方法。

【請求項 12】

コンピュータに請求項9ないし11のうちいずれか一項記載の方法を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】

関連出願への相互参照

本願は2013年4月5日に出願された米国仮特許出願第61/808,701号の優先権を主張するものである。同出願の内容はここに参照によってその全体において組み込まれる。

【0002】

10

20

30

40

50

## 技術分野

本稿に開示される発明は概括的にはオーディオビジュアル・メディア頒布に関する。特に、複数オーディオ・フレーム長を受け入れるよう適応されており、よってフレーム同期的オーディオビジュアル・メディア・フォーマットを扱うメディア・デコーダの一部をなすのに好適なオーディオ・デコーダに関する。

### 【背景技術】

#### 【0003】

今日利用可能なたいていの商業用途において使われるオーディオおよびビデオのフレーム・レート（またはフレーム周波数）は、記録および再生ソフトウェア・プロダクト、ハードウェア・コンポーネントならびに通信当事者間でオーディオおよびビデオを伝送するための合意されたフォーマットの両方において現われる別個の確立された業界標準に従う。オーディオ・フレーム・レートは典型的には種々の符号化アルゴリズムに固有であり、44.1および48kHzのようなオーディオ・サンプリング周波数に関連付けられる。これらはそれぞれの地理的領域におけるビデオ・フレーム・レート29.97fps（NTSC）および25fps（PAL）と同じくらい有名である。さらなる標準的なビデオ・フレーム・レートは23.98、24および30fpsまたはより一般化された形では24、25、30fpsおよび(24,25,30) × 1000/1001fpsを含む。オーディオ・フレーム・レートを統一または調和させる試みは、アナログからデジタル頒布へのシフトにもかかわらず、まだうまくいっていない。このことは、オーディオ・フレーム（たとえばネットワークを通じた伝送のために好適なパケットまたは符号化単位）が一般には整数個のビデオ・フレームに対応しないことを含意する。

#### 【0004】

オーディオビジュアル・データ・ストリームを同期させる必要性は、クロック・ドリフトの結果として、あるいはいくつかのストリームが異なる現からサーバーにおける共通の処理、編集またはスプライシングのために受領されるとき、絶えず生じている。こうした状況は放送局では頻繁に遭遇される。図3に示した状況では、オーディオ・フレーム（ストリームS1内のA11、A12、...およびストリームS2内のA21、A22、...）およびビデオ・フレーム（ストリームS1内のV11、V12、...およびストリームS2内のV21、V22、...）は一致しない。（たとえば両ストリームをスプライシングする試みにおいて）ストリームの一方におけるビデオ・フレームを複製または除去することによってストリーム間でビデオからビデオへの同期を改善する試みは、典型的にはそのストリーム内でのオーディオからビデオの非同期につながる。一般に、たとえ対応するオーディオ・フレームが削除または複製されたとしても、少なくともある程度は非同期は持続する。

#### 【0005】

さらなる処理を代償として、同期の間にオーディオを時間的にデコードしてフレーム分割とは独立な低レベル・フォーマット、たとえばもとのサンプリング周波数の分解能のベースバンド・フォーマットまたはパルス符号変調（PCM）にすることによって、より大きな工夫の余地を作り出すことができる。しかしながら、そのようなデコードは、メタデータの、特定のオーディオ・セグメントへの厳密なアンカー付けをばかし、「完璧な」中間フォーマットにデコードすることによっては是正できない情報損失を生じる。一例として、ダイナミックレンジ制御（DRC）は典型的にはモード依存かつ設備依存であり、したがって実際の再生の時点でのみ消費されることができる。オーディオ・パケットを通じてDRCの特性を支配するデータ構造は、同期が行なわれた後で忠実に復元するのは難しい。よって、相続くデコード、同期およびエンコード段を過ぎてこの型のメタデータを保存するタスクは、複雑さの制約条件を受ける場合には、簡単なタスクではない。

#### 【0006】

一層深刻な困難が、二チャンネルPCM信号を担持するよう設計されており、よって符号化された形でしかマルチチャンネル・コンテンツを扱えないレガシー・インフラストラクチャーとの関連で生じうる。

#### 【0007】

所与のフレームにおけるデータが記録され、符号化されたオーディオビジュアル信号に

10

20

30

40

50

おける同じ時間セグメントに正確に対応するという意味でフレーム同期的にオーディオおよびビデオ・データをエンコードすることがより便利であることは確かである。これはオーディオビジュアル・ストリームのフレームごとの操作、すなわち、ストリーム内の一つまたは複数の独立した符号化単位全体の複製または除去のもとで、オーディオからビデオへの同期を保存する。ドルビーE(商標)オーディオ・フォーマットにおいて利用可能なフレーム長はビデオ・フレーム長に一致する。典型的なビットレート448kbpsでは、このフォーマットは、デジタル・ビデオ・カセットのようなハード・メディアを好ましい記憶態様として、主として業務用の制作のために設計された。

【発明の概要】

【発明が解決しようとする課題】

10

【0008】

フレーム同期的オーディオビジュアル・フォーマットの一部としての頒布目的に好適な代替的なオーディオ・フォーマットならびにそれとともに使うのに好適な符号化および復号設備が必要とされている。

【図面の簡単な説明】

【0009】

下記において、例示的な実施形態が付属の図面を参照して、より詳細に記述される。

【図1】ある例示的な実施形態に基づくオーディオ処理システムの構造および該システム内のコンポーネントの内部機構を示す図である。

【図2】メディア頒布システムを示す図である。

20

【図3】従来技術に基づく、独立したフレーム長をもつオーディオ・ビットストリームおよびビデオ・フレームの関連したフレームをそれぞれ含む二つのマルチメディア・ストリームを示す図である。

【図4】ある例示的な実施形態に基づく、調整されたフレーム長をもつオーディオ・ビットストリームおよびビデオ・フレームの関連したフレームをそれぞれ含む二つのマルチメディア・ストリームを示す図である。

【図5】ある例示的な実施形態に基づくエンコーダ・システムの構造を示す図である。すべての図面は概略的であり、一般に本発明を明快にするのに必要な部分を示すのみである。他の部分は省略されたり、あるいは単に示唆され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。

30

【発明を実施するための形態】

【0010】

#### 1. 概観

オーディオ処理システムは、オーディオ・データを担持するフレームにセグメント分割されたオーディオ・ビットストリームを受け入れる。オーディオ・データは、音波をサンプリングし、こうして得られた電子的な時間サンプルをスペクトル係数に変換し、それが次いで量子化され、伝送または記憶に好適なフォーマットで符号化されることによって用意されたものでもよい。オーディオ処理システムは、単一チャンネル、ステレオまたはマルチチャンネルのフォーマットにおいてサンプリングされた音波を再構成するよう適応されている。フレーム中のオーディオ・データは、当該オーディオ・データが表わす音波の追加的な時間区間を再構成するための基礎として十分であるという意味で自己完結的であってもよい。再構成は、重複変換などのために先行するオーディオ・データの知識を必要としなくてもよい。

40

【0011】

オーディオ処理システムは、少なくともフロントエンド・コンポーネントと、処理段と、所望されるターゲット・サンプリング周波数で処理されたオーディオ信号の時間領域表現を提供するためのサンプル・レート変換器とを有する。ターゲット・サンプリング周波数はあらかじめ定義された量であり、入来オーディオ・ビットストリームの属性(たとえばオーディオ・フレーム・レート)とは独立にユーザーまたはシステム設計者によって構成設定可能であることを強調しておく。一つのオプションとして、ターゲット・サンプリ

50

ング周波数はオーディオ・フレーム・レートの関数であってもよい。もう一つのオプションとして、ターゲット・サンプリング周波数は一定および/またはオーディオ・フレーム・レートに関して非適応的であってもよい。

【 0 0 1 2 】

フロントエンド・コンポーネントにおいて、量子化解除段がオーディオ・ビットストリームからのオーディオ・フレームぶんの数の量子化されたスペクトル係数を受領し、該係数を合同して処理してフレーム中のオーディオ・データの逆量子化を制御する情報を生成し、逆量子化を実行する。フロントエンド・コンポーネントにおいて量子化解除段の下流で、逆変換段が前記出力 中間信号の第一の周波数領域表現をなすスペクトル係数

を受け、該中間信号の時間領域表現を合成する。次いで、フロントエンド・コンポーネントは、オーディオ・ビットストリームから同数のその後の量子化されたスペクトル係数（ただしこれはいくつかの変換にわたって分散していてもよい）を受領し、処理し、これらの係数を処理することに進む。量子化解除段は、オーディオ・データを所定の量子化レベル（または再構成レベルまたは再構成点）にデコードすることによって前記スペクトル係数を生成するよう構成されている。量子化レベルは、音響心理学的考察に基づいてエンコードによって選ばれている。たとえば、所与の周波数（または周波数帯域）についての量子化雑音がマスキング閾値を超えないような仕方で行なわれている。マスキング閾値は周波数依存なので、経済的な観点からは、エンコードに、周波数に関して非一様な諸量子化レベルを選択させることが好ましい。結果として、量子化および量子化解除は典型的には、最適な出力が生成される特定の物理的サンプリング周波数を念頭に行なわれる。

【 0 0 1 3 】

オーディオ処理システムにおける処理段は、周波数領域で処理を実行するよう適応されていてもよい。この目的のために、処理段は、前記中間信号の周波数領域表現を提供する初期分解フィルタバンクと、一つまたは複数の処理コンポーネントとを有する。該処理コンポーネントに続いて、処理されたオーディオ信号を時間領域に戻す変換をする合成フィルタバンクがくる。

【 0 0 1 4 】

ある例示の実施形態では、オーディオ処理システムは、少なくとも二つの異なるあらかじめ定義されたオーディオ・フレーム・レートのうちの任意のものをもつオーディオ・ビットストリームを受け入れるよう構成されている。このことは、オーディオ処理システムが、各ビデオ・フレームが一つまたは複数のビデオ・フレーム、好ましくは一つまたは複数のフル・ビデオ・フレームの継続時間に等しい継続時間のオーディオ・データを担持するオーディオ・フレームに時間的に関係しているという意味でビデオ・フレームのストリームと関連付けられているオーディオ・ストリームに対して作用することを許容する。それにより、二つ以上の異なるビデオ・フレーム・レートが、オーディオ処理システムにおけるオーディオからビデオへの同期を犠牲にすることなくマッチされることができる。本稿での用法では、ビデオ・フレームの「継続時間」（秒単位）は、ビデオ・フレーム・レートの逆数として理解されてもよい。同様に、オーディオ・フレームの「継続時間」（秒単位）は、オーディオ・フレーム・レートの逆数として定義されてもよい。フレーム中のビデオ・データは、関係するオーディオ・データがサンプリングされた期間の初期時点、中間時点または最終時点におけるサンプリングから帰結しうる。あるいはまた、ビデオ・データは、オーディオ・サンプリング期間と少なくとも部分的に重なる（たとえばローリング・シャッター・プロセスによる）長い期間にわたってサンプリングされている。フロントエンド・コンポーネントは（サンプル数で測られる）可変最大フレーム長をもち、前記あらかじめ定義されたオーディオ・フレーム・レートに対応する少なくとも二つのモードで動作可能である。

【 0 0 1 5 】

オーディオ処理システムは、相対的により低いオーディオ・フレーム・レートについて相対的により大きなフレーム長（または最大フレーム長、可能性のある細分は考慮する、下記参照）を選択することができるので、所望されるフレーム・レート適応性を達成する



ア・ビットストリームは、オーディオ・ビットストリームと、各ビデオ・フレームがオーディオ・ビットストリーム中のあるフレームに時間的に関係しているという意味でオーディオ・ビットストリームに関連付けられているビデオ・フレームのストリームとを含んでもよい。オーディオおよびビデオ・フレームの間の時間的關係は、一対一、一対多または多対一でありうる。これは、マルチメディア処理システムがフレーム同期的な仕方ですなわちマルチメディア・ストリームにおけるオーディオからビデオの同期性を乱すことなくオーディオおよびビデオ・フレームの同時の複製または削除を実行することを許容する。すでに述べたように、これは、二つのマルチメディア・ストリームのスプライシングのような動作を容易にする。フレーム同期的フォーマットの二つのマルチメディア・ストリームがスプライシングされる場合、フレーム境界をスプライシング点として使うことが常に安全である。それでも、本発明の範囲内で、マルチメディア処理システムは、異なる源、たとえばライブ・レコーディング・サイト、電波またはネットワーク受信機または記憶媒体からの二つ以上のマルチメディア・ストリームを受け入れてもよい。

10

**【 0 0 2 1 】**

ある例示的实施形態では、オーディオ処理方法は、ビデオ・フレームのストリームと関連付けられたオーディオ・ビットストリームを入力する。ここで、オーディオ・ビットストリームはフレームにセグメント分割されており、各ビデオ・フレームは、一つまたは複数のビデオ・フレームと同じ継続時間の関連付けられたオーディオ・データを担持する、オーディオ・ビットストリーム中のフレームに時間的に関係している。そしてオーディオ処理方法は、再構成されたオーディオ信号を出力する。本方法は：

20

- ・オーディオ・ビットストリームの現在フレーム・レートを確認する段階と；
- ・オーディオ・ビットストリーム中のあるオーディオ・フレームに対応する量子化されたスペクトル係数を繰り返し受領し、周波数から時間への変換が後続する逆量子化を実行する段階であって、それにより中間オーディオ信号の表現が得られる、段階と；
- ・前記中間オーディオ信号に対して周波数領域で少なくとも一つの処理段階を実行する段階と；
- ・処理されたオーディオ信号のサンプリング・レートをターゲット・サンプリング周波数に変更する段階であって、それにより再構成されたオーディオ信号の時間領域表現が得られる、段階とを含む。

**【 0 0 2 2 】**

30

この例示的实施形態では、第二の段階（逆量子化および周波数から時間への変換）は、それぞれ異なる最大フレーム長をもつ少なくとも二つのモードで動作可能な、ソフトウェア・モジュールまたは量子化解除および合成回路内のコンポーネントといった機能コンポーネントにおいて実行される。機能コンポーネントを動作させるモードは、オーディオ・ビットストリームの現在フレーム・レートに応答して選択される。ここで、二つの異なるフレーム・レート値が機能コンポーネントの異なる動作モードを生じさせる。

**【 0 0 2 3 】**

上記の特性をもつオーディオ処理方法は、多様なオーディオ・フレーム・レートに適用可能である。逆量子化を先述した同じ機能コンポーネントで実行する必要も、そもそもオーディオ・フレーム・レート変化に応答する機能コンポーネントで実行する必要もない。実際、逆量子化は、オーディオ・フレーム・レート変化に特に適応しないという意味で静的な（または単一モード）型の機能コンポーネントにおいて実行されてもよい。

40

**【 0 0 2 4 】**

ある例示的实施形態では、上記のオーディオ処理方法を実行するための命令をもつコンピュータ可読な一時的または非一時的な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 2 5 】**

第二の側面では、本発明は、いくつかの受け入れ可能なフレーム・レートのうちの一つをもつオーディオ・ビットストリームを処理する問題に対する代替的な解決策を提供する。ある例示的实施形態では、オーディオ処理システムは、ビデオ・フレームのストリーム

50

に関連付けられたオーディオ・ビットストリームを受け入れるよう構成されている。ここで、オーディオ・ビットストリームはフレームにセグメント分割されており、各ビデオ・フレームは、一つまたは複数のビデオ・フレームと同じ継続時間の関連付けられたオーディオ・データを担持する、オーディオ・ビットストリーム中のフレームに時間的に関係している。本オーディオ処理システムは：

・フロントエンド・コンポーネントであって：

オーディオ・ビットストリーム中のあるオーディオ・フレームに対応する量子化されたスペクトル係数を繰り返し受領し、中間信号の第一の周波数領域表現を出力するよう適応された量子化解除段；および

前記中間信号の前記第一の周波数領域表現を受領し、それに基づいて前記中間信号の時間領域表現を合成する逆変換段とを含むフロントエンド・コンポーネントと；

・処理段であって：

前記中間信号の前記時間領域表現を受領し、前記中間信号の第二の周波数領域表現を出力する分解フィルタバンク；

前記中間信号の前記周波数領域表現を受領し、処理されたオーディオ信号の周波数領域表現を出力する少なくとも一つの処理コンポーネント；および

前記処理されたオーディオ信号の前記周波数領域表現を受領し、前記処理されたオーディオ信号の時間領域表現を出力する合成フィルタバンクとを含む処理段と；

・前記処理されたオーディオ信号の前記時間領域表現を受領し、ターゲット・サンプリング周波数でサンプリングされた再構成されたオーディオ信号を出力するサンプル・レート変換器とを有する。

【0026】

この例示的实施形態によれば、フロントエンド・コンポーネントは固定フレーム長を有していてもよい。フロントエンド・コンポーネントが異なるフレーム継続時間（または等価だが異なるフレーム・レート）だが等しいフレーム長に対応する量子化されたスペクトル係数を処理するとき、サンプル・レート変換器がオーディオ処理システムの出力の物理的なサンプリング周波数を所望されるターゲット物理的なサンプリング周波数に復元してもよい。先の例示的实施形態と同様に、異なるフレーム・レートへの適応可能性は、フロントエンド・コンポーネントが前提とする（あるいはより正確にはそれが受領する制御データを準備するエンティティが前提とする）物理的なサンプリング周波数と再構成されたオーディオ信号の物理的なサンプリング周波数との間の限られたミスマッチを受け入れることによって達成される。

【0027】

本発明のあるさらなる側面は、上記で概説したオーディオ処理システムのいずれかによってデコードされるオーディオ・ビットストリームを準備するのに好適なエンコーダ・システムおよびエンコード方法を提供する。

【0028】

特に断りのない限り、本発明は、たとえ互いに異なる請求項に記載されていたとしても、特徴のあらゆる組み合わせに関する。

【0029】

## II. 例示的实施形態

図1は、エンコードされたオーディオ・ビットストリームPを受領し、図1においてステレオ・ベースバンド信号L、Rの対として示される再構成されたオーディオ信号をその最終出力とする、オーディオ処理システム100の一般化されたブロック図である。この例では、ビットストリームPが量子化された変換符号化された二チャンネル・オーディオ・データを含むことが想定される。オーディオ処理システム100は通信ネットワーク、無線受信機またはメモリ（図示せず）からオーディオ・ビットストリームPを受領しうる。システム100の出力は、再生のためにラウドスピーカに供給されてもよく、あるいは通信ネットワークまたは無線リンクを通じたさらなる伝送のためまたはメモリ内での記憶のために同じまたは異なるフォーマットで再エンコードされてもよい。

## 【 0 0 3 0 】

オーディオ処理システム 1 0 0 は、ビットストリーム P を量子化されたスペクトル係数および制御データにデコードするためのデコーダ 1 0 8 を有する。その構造についてのちにより詳細に論じるフロントエンド・コンポーネント 1 1 0 は、これらのスペクトル係数を量子化解除し、処理段 1 2 0 によって処理されるべき中間オーディオ信号の時間領域表現を供給する。中間オーディオ信号は分解フィルタバンク 1 2 2 L、1 2 2 R によって、先述した符号化変換に関連付けられたものとは異なる第二の周波数領域に変換される。第二の周波数領域表現は、直交ミラー・フィルタ (QMF) 表現であってもよく、その場合、分解フィルタバンク 1 2 2 L、1 2 2 R は QMF フィルタバンクとして提供されてもよい。分解フィルタバンク 1 2 2 L、1 2 2 R の下流では、高周波数再構成を受け持つスペクトル帯域複製 (SBR) モジュール 1 2 4 およびダイナミックレンジ制御 (DRC) モジュール 1 2 6 が、中間オーディオ信号の第二の周波数領域表現を処理する。その下流では、合成フィルタバンク 1 2 8 L、1 2 8 R がこのように処理されたオーディオ信号の時間領域表現を生成する。本開示を吟味した当業者は理解するであろうが、スペクトル帯域複製モジュール 1 2 4 もダイナミックレンジ制御モジュール 1 2 6 も本発明の必須要素ではない。逆に、異なる例示的实施形態に基づくオーディオ処理システムは、処理段 1 2 0 内の追加的または代替的なモジュールを含んでいてもよい。処理段 1 2 0 の下流では、サンプル・レート変換器 1 3 0 が、処理されたオーディオ信号のサンプリング・レートを調整して、意図された再生設備 (図示せず) が指定されている所望されるオーディオ・サンプリング・レート、たとえば 44.1kHz または 48kHz にするよう動作可能である。出力におけるアーチファクトが少ないサンプル・レート変換器 1 3 0 をどのように設計するか自身は当技術分野において既知である。サンプル・レート変換器 1 3 0 は、これが必要とされない時、すなわち処理段 1 2 0 が供給する処理されたオーディオ信号がすでにターゲット・サンプリング周波数をもつ時には非アクティブ化されてもよい。サンプル・レート変換器 1 3 0 の下流に配置される任意的な信号制限モジュール 1 4 0 が、クリップなし (no-clip) 条件に従って必要に応じてベースバンド信号値を制限するよう構成される。クリップなし条件はやはり特定の意図される再生設備に鑑みて選ばれてもよい。

## 【 0 0 3 1 】

図 1 の下部に示されるように、フロントエンド・コンポーネント 1 1 0 は、異なるブロック・サイズをもついくつかのモードの一つで動作できる量子化解除段 1 1 4 と、やはり異なるブロック・サイズに対して動作できる逆変換段 1 1 8 L、1 1 8 R とを有する。好ましくは、量子化解除段 1 1 4 および逆変換段 1 1 8 L、1 1 8 R のモード変化は同期的であり、よってブロック・サイズはすべての時点においてマッチする。これらのコンポーネントの上流では、フロントエンド・コンポーネント 1 1 0 は、制御データから量子化されたスペクトル係数を分離するためのデマルチプレクサ 1 1 2 を有する。典型的には、それは制御データを逆変換段 1 1 8 L、1 1 8 R に転送し、量子化されたスペクトル係数 (および任意的には前記制御データ) を量子化解除段 1 1 4 に転送する。量子化解除段 1 1 4 は、(典型的には整数として表わされる) 量子化インデックスの一つのフレームから (典型的には浮動小数点数として表わされる) スペクトル係数の一つのフレームへのマッピングを実行する。各量子化インデックスは、量子化レベル (または再構成点) に関連付けられている。オーディオ・ビットストリームが上記で論じたように非一様量子化を使って準備されたとすると、量子化インデックスがどの周波数帯域を指すかが指定されない限り、その関連付けは一意的ではない。別の言い方をすれば、量子化解除プロセスは、各周波数帯域についての異なるコードブックに従ってもよく、コードブックの集合がフレーム長および/またはビットレートの関数として変わってもよい。図 1 では、これは概略的に示されている。ここで、縦軸は周波数を表わし、横軸は単位周波数当たり割り当てられた符号化ビットの量を表わす。周波数帯域は典型的にはより高い周波数についてはより広くなり、内部サンプリング周波数  $f_i$  の半分で終わることを注意しておく。内部サンプリング周波数は、サンプル・レート変換器 1 3 0 における再サンプリングの結果として、数値的に異なる物理的なサンプリング周波数にマッピングされてもよい。たとえば、4.3% のアップサン

10

20

30

40

50

プリングは  $f_i = 46.034\text{kHz}$  を近似的な物理的周波数  $48\text{kHz}$  にマッピングし、低いほうの周波数帯域境界を同じ因子だけ増大させる。図 1 がさらに示唆するように、オーディオ・ビットストリームを準備するエンコーダは典型的には、異なる周波数帯域には、符号化される信号の複雑さおよび人間の聴覚の期待される感度変動に従って、異なる量の符号化ビットを割り当てる。

【 0 0 3 2 】

オーディオ処理システム 1 0 0、特にフロントエンド・コンポーネント 1 1 0 の動作モードを特徴付ける定量的データが表 1 に与えられている。

【 0 0 3 3 】

【表 1】

表 1: オーディオ処理システムの例示的な動作モード a-m									
モード	フレームレート [Hz]	フレーム 継続時間 [ms]	フロントエンド・ コンポーネント における フレーム長 [サンプル]	フロントエンド・ コンポーネント における ビット幅 [Hz]	内部 サンプリング 周波数 [kHz]	分解フィルタ バンク [帯域]	分解 周波数 帯域の幅 [Hz]	SRC 因子	外部サンプリ ング周波数 [kHz]
a	23.976	41.708	1920	11.988	46.034	64	359.640	0.9590	48.000
b	24.000	41.667	1920	12.000	46.080	64	360.000	0.9600	48.000
c	24.975	40.040	1920	12.488	47.952	64	374.625	0.9990	48.000
d	25.000	40.000	1920	12.500	48.000	64	375.000	1.0000	48.000
e	29.970	33.367	1536	14.985	46.034	64	359.640	0.9590	48.000
f	30.000	33.333	1536	15.000	46.080	64	360.000	0.9600	48.000
g	47.952	20.854	960	23.976	46.034	64	359.640	0.9590	48.000
h	48.000	20.833	960	24.000	46.080	64	360.000	0.9600	48.000
i	50.000	20.000	960	25.000	48.000	64	375.000	1.0000	48.000
j	59.940	16.683	768	29.970	46.034	64	359.640	0.9590	48.000
k	60.000	16.667	768	30.000	46.080	64	360.000	0.9600	48.000
l	120.000	8.333	384	60.000	46.080	64	360.000	0.9600	48.000
m	25.000	40.000	3840	12.500	96.000	128	375.000	1.0000	96.000

表 1 における強調された列は、制御可能な量の値を含んでいる。残りの量はこれらに依

10

20

30

40

50

存するものと見なされてよい。さらに、再サンプリング (SRC) 因子の理想的な値が  $(24/25) \times (1000/1001) = 0.9560$ 、 $24/25 = 0.96$  および  $1000/1001 = 0.9990$  であることを注意しておく。表 1 に挙げた SRC 因子の値は丸められている。フレーム・レート値も同様である。再サンプリング因子 1.000 は厳密であり、SRC 130 が非アクティブ化されているまたは完全に存在しないことに対応する。例示的实施形態では、オーディオ処理システム 100 は、そのうちの一つまたは複数が表 1 のエントリーに一致してもよい異なるフレーム長をもつ少なくとも二つのモードで動作可能である。

#### 【0034】

フロントエンド・コンポーネントのフレーム長が 1920 サンプルに設定されるモード a-d は、普及している符号化フォーマットのビデオ・フレーム・レートに厳密に一致するよう  
10  
選択された (オーディオ) フレーム・レート 23.976、24.000、24.975 および 25.000 Hz を扱うために使われる。異なるフレーム長のため、モード a~d では、内部サンプリング周波数 (フレーム・レート  $\times$  フレーム長) は約 46.034 kHz から 48.000 kHz まで変わる。臨界サンプリングおよび均等に離間した周波数ピンを想定すると、これは 11.988 Hz から 12.500 Hz の範囲内のピン幅値 (内部サンプリング周波数の半分 / フレーム長) に対応する。内部サンプリング周波数の変動は制限されているので (フレーム・レートの変動範囲が約 5% である結果、それは約 5%)、オーディオ処理システム 100 は、はいつてくるオーディオ・ビットストリームが準備された物理的なサンプリング周波数に厳密にマッチしないにもかかわらず、四つのモード a~d のすべてにおいて合理的な出力品質を与える。

#### 【0035】

フロントエンド・コンポーネント 110 の下流に続けると、分解 (QMF) フィルタバンク 122 は、すべてのモード a~d において 64 個の帯域または QMF フレーム当たり 30 個のサンプルをもつ。物理的には、これは各分解周波数帯域のわずかに変動する幅に対応するが、変動はやはり無視できるほど限られている。特に、SBR および DRC 処理モジュール 124、126 は、出力品質に不都合なく、現在モードについて不可知であってもよい。しかしながら、SRC 130 はモード依存であり、処理されたオーディオ信号の各フレームが物理的単位において 48 kHz のターゲット外部サンプリング周波数に対応する数のサンプルを含むことを保証するために、ターゲット外部サンプリング周波数と内部サンプリング周波数の商に一致するよう選ばれる。特定の再サンプリング因子を使う。

#### 【0036】

モード a~d のそれぞれにおいて、オーディオ処理システム 100 は、ビデオ・フレーム・レートと外部サンプリング周波数両方に厳密にマッチする。すると、オーディオ処理システム 100 は図 4 のマルチメディア・ビットストリーム T1 および T2 のオーディオ部分を扱う。ここで、オーディオ・フレーム A11、A12、A13、... ; A22、A23、A24、... およびビデオ・フレーム V11、V12、V13、... ; V22、V23、V24 は各ストリーム内において時間的に一致する。図 4 に示唆されるように、このとき、進んでいるストリームにおけるオーディオ・フレームおよび関連付けられたビデオ・フレームを削除することによって、ストリーム T1、T2 の同期を改善することができる。あるいはまた、遅れているストリームにおけるオーディオ・フレームおよび関連付けられたビデオ・フレームが複製され、もとの位置の次に挿入される。この際、可能性としては知覚可能なアーチファクトを軽減するために補間措置が組み合わされる。

#### 【0037】

フレーム・レート 29.97 Hz および 30.00 Hz を扱うよう意図されたモード e および f は、第二のサブグループとして見分けられる。すでに説明したように、オーディオ・データの量子化は、約 48 kHz の内部サンプリング周波数のために適応される (または最適化される)。よって、各フレームがより短いので、フロントエンド・コンポーネント 110 のフレーム長はより小さな値、1536 サンプルに設定され、それにより約 46.034 および 46.080 kHz の内部サンプリング周波数となる。分解フィルタバンク 122 が 64 個の周波数帯域をもちモード独立であれば、各 QMF フレームは 24 サンプルを含む。

#### 【0038】

10

20

30

40

50

同様に、50Hzおよび60Hz（標準化されたテレビジョン・フォーマットにおけるリフレッシュ・レートの二倍に対応）および120Hzまたはその付近のフレーム・レートは、それぞれモードg~i（フレーム長960サンプル）、モードj~k（フレーム長768サンプル）およびモードl（フレーム長384サンプル）によってカバーされる。内部サンプリング周波数はそれぞれの場合において48kHzに近いままであり、よってオーディオ・ビットストリームが生成されたときの量子化プロセスのいかなる音響心理学的チューニングも少なくとも近似的には有効なままであることを注意しておく。64帯域フィルタバンクにおけるそれぞれのQMFフレーム長は15、12および6サンプルである。

【0039】

上述したように、オーディオ処理システム100は、オーディオ・フレームをより短いサブフレームに細分するよう動作可能であってもよい。これを行なう理由は、オーディオ過渡成分をより効率的に捕捉することであってもよい。48kHzサンプリング周波数および表1で与えられた設定については、下記の表2~表4が2、4、8（図4におけるオーディオ・フレームA29参照）および16サブフレームへの細分から帰結するビン幅およびフレーム長を示している。表1に基づく設定は時間および周波数分解能の有利なバランスを達成すると思われる。

【0040】

【表2】

表 2:フレーム長 2048 サンプルでの時間/周波数分解能					
	サブフレーム数				
	1	2	4	8	16
ビン数	2048	1024	512	256	128
ビン幅[Hz]	11.72	23.44	46.88	93.75	187.50
フレーム継続時間[ms]	42.67	21.33	10.67	5.33	2.67

【0041】

【表3】

表 3:フレーム長 1920 サンプルでの時間/周波数分解能					
	サブフレーム数				
	1	2	4	8	16
ビン数	1920	960	480	240	120
ビン幅[Hz]	12.50	25.00	50.00	100.00	200.00
フレーム継続時間[ms]	40.00	20.00	10.00	5.00	2.50

【0042】

【表 4】

表 4:フレーム長 1536 サンプルでの時間/周波数分解能					
	サブフレーム数				
	1	2	4	8	16
ビン数	1536	768	384	192	96
ビン幅[Hz]	15.63	31.25	62.50	125.00	250.00
フレーム継続時間[ms]	32.00	16.00	8.00	4.00	2.00

フレームの細分に関する決定は、オーディオ・エンコード・システム（図示せず）におけるようなオーディオ・ビットストリームを準備するプロセスの一部として行なわれてもよい。

## 【 0 0 4 3 】

表 1 においてモード m によって示されるように、オーディオ処理システム 1 0 0 はさらに、96kHz の増大した外部サンプリング周波数および 128QMF 帯域で動作することを可能にされてもよい。これは QMF フレーム当たり 30 サンプルに対応する。外部サンプリング周波数はたまたま内部サンプリング周波数と一致するので、SRC 因子は 1 である。これは再サンプリングが必要ないことに相当する。

## 【 0 0 4 4 】

図 2 は、受信サイト 2 1 0 を有するメディア頒布システムを示している。受信サイト 2 1 0 は衛星受信機 2 1 1 および / または電波周波数受信機 2 1 2 を有していてもよく、これらは統合受信機兼復号器（IRD: integrated receiver and decoder）におけるコンポーネントとして設けられてもよい。オーディオおよびビデオ・データを搬送するマルチメディア・ビットストリームは受信サイト 2 1 0 において受信される。マルチメディア・ビットストリームは、ビットストリーム・コンテンツのローカル時間 2 1 9、たとえばサンプリング装置におけるローカル時間を定義するタイムスタンプを有する。受信され、可能性としては復号されたマルチメディア・ビットストリームは次いで、コンテンツ・ルーティングおよび管理が行なわれてもよいターンアラウンド・サイト 2 2 0 に伝送される。いくつかのターンアラウンド・サイト 2 2 0 は復号を全く実行しなくてもよい。マルチメディア・ビットストリームのオーディオ部分はたとえば、映画テレビ技術者協会によって定義された業界標準 SMPTE337 のあるバージョンに従ってフォーマットされてもよい。このフォーマットは、データが他のインターフェースを通じて送達される情報コンテンツと同期されることを許容する時間整列情報を含むので、使うのが有利である。そのような属性をもつ他のフォーマットがこの使用において SMPTE337 を置き換えてもよい。

## 【 0 0 4 5 】

ターンアラウンド・サイト 2 2 0 におけるローカル時間 2 2 9 は、ターンアラウンド・サイト 2 2 0 および下流のエンティティに対して共通のローカル参照クロックまたはネットワーク参照時間によって定義されてもよい。理想的には、コンテンツのローカル時間 2 1 9 およびターンアラウンド・サイト 2 2 0 におけるローカル時間 2 2 9 は等しいが、実際上は、クロック・ドリフトの結果として違いがあることがある。その点で、受信されたマルチメディア・ビットストリームを同期させる必要性が生じる。より正確には、ターンアラウンド処理が HD-SDI (high definition serial digital interface [ 高精細度シリアル・デジタル・インターフェース ]) レベルでなされる実装では、圧縮解除されたビデオは、従来は未使用の HANC および VANC (horizontal and vertical ancillary data [ 水平および垂直補助データ ]) スペースをいまだ有するフレームにおいて搬送される。HANC および VANC スペースは、エンコードされたオーディオ・データを輸送するために使われる。この実装では、オーディオおよびビデオ・データの間の唯一の相対的なタイミング

10

20

30

40

50

情報は、HD-SDI信号の特定の諸フレームにおけるオーディオ・データの配置によって暗黙的に与えられる。オーディオが同期的な仕方でフレーミングされないとすると、HD-SDIフレームの反復または脱落は、そのビデオ・フレームと一緒に反復/脱落されるオーディオ・データの継続時間（または時間的な局在化）についていかなる情報も返さないであろう。反復/脱落が隣接する諸HD-SDIフレームが、次のまたは直前のHD-SDIフレームに依拠することによって完全なオーディオ・フレームに復元されることのできないオーディオ・フレームの断片を含むようになる場合、該隣接する諸HD-SDIフレームも破損することがありうる。フレーム同期的なオーディオ符号化は、この障害を回避する。等しい物理的継続時間のオーディオおよびビデオ・データおよび実際にはその完全なフレームを反復する/脱落させることしか許容しないからである。

10

## 【0046】

複数のマルチメディア・ビットストリームをたとえば異なるコンテンツ・オブジェクトの間でまたはコンテンツと広告素材の間で接合することは、フレーム同期性が非常に有用であるさらなる例である。実際、これは各フレームがオーディオおよびビデオ・データの対応するおよび時間同期的な部分を符号化することを保証する。よって、二つの連続するフレームの間でビットストリームを中断または開始することが常に安全である。したがって、フレーム境界が接合点として使われてもよい。

## 【0047】

ターンアラウンド・サイト220の下流では、シリアル・デジタル・インターフェース(SDI)ルーター230、次いでビデオ・エンコーダが配置されていてもよい。ビデオ・エンコーダにはオーディオ・パススルー・コンポーネントが並列接続されていて、まとめて参照符号240で示される。SDIルーター230はたとえば、それぞれSMPTE259MおよびSMPTE292Mとして標準化された標準精細度および高精細度フォーマットに準拠していてもよい。ビデオ・エンコーダおよびオーディオ・パススルー・コンポーネント240から供給されるトランスコードされたオーディオビジュアル・データは次いで、有線または無線通信ネットワーク290を通じて消費者291、292、293に伝送される。

20

## 【0048】

図3および図4はすでに論じた。

## 【0049】

本発明の第二の側面に関し、図1に示したオーディオ処理システム100と同じ一般的な特性をもつが、フロントエンド・コンポーネントのマルチモード機能を全く必要としない代替的なオーディオ処理システムを提供することが構想される。よって、そのフロントエンド・コンポーネントの(恒久的な)設定に依存して、代替的なオーディオ処理システムは、モードa~dまたはモードe~fまたはモードg~iまたはモードj~kのいずれかで動作可能である。代替的なオーディオ処理システムの適応可能性は主としてSRCに起因する。フロントエンド・コンポーネントは、変わりうる物理的継続時間を必ずしも知ることなく(あるいは少なくともそれに適応することなく)各フレーム内のサンプルを処理する。よって、物理的なサンプリング周波数は変わりうるが、その変動が限られている限り、著しい不都合な音響心理学的な副作用はない。

30

## 【0050】

図5では、ある例示的实施形態に基づくエンコーダ・システム501が概略的に描かれている。入力信号5011はサンプリング・レートFsをもち、これは48kHzであってもよい。さらに、ビデオ・フレーム・レート5010がエンコーダに入力され、これにオーディオ・フレーム・レートが整列されるべきである。「適応」モジュール502は、要求されるフレーム・レートを扱うようシステムをベース変換サイズ(または最大フレーム長)に適応させ、このベース変換サイズを、(表2~4に基づく)入力信号特性の関数として、より小さな変換に細分する。所望されるフレーム・レートはさらにサンプル・レート変換器(SRC)モジュール508に供給される。システムが、単一のベース変換サイズをもって種々のフレーム・レートの倍数を扱うことができるようにする内部サンプリング・レート(内部Fs)を与えるためである。ベース変換サイズは、オーディオ符号化の観点から

40

50

最適または少なくとも有利であるよう選ばれる。すなわち、それは静的な信号については（合理的な限界の範囲内で）できるだけ長いべきであり、たとえば48kHzのサンプリング・レートについて2048、1920、1536MDCTラインであり、好ましくは過渡的なパッセージについてはより小さな変換に、たとえば8個または16個の変換に細分可能である。本例示的实施形態によれば、SRCモジュール508は、外部サンプリング・レートからの制限された逸脱をもつ内部サンプリング・レートを提供するように制限されるよう設計される。それにより、選択されたMDCTベース変換サイズの物理的な時間/周波数属性を有意に変えない。

#### 【0051】

MDCTモジュール504は、サンプリング・レート「内部FS」の入力時間領域信号を周波数領域に変換する。「量子化&エンコード」モジュール503はMDCTラインを、実際の（物理的な）周波数の関数としての人間の聴覚系の感度および周波数分解能の変動を考慮に入れる音響心理学的モデルに従って量子化する。「量子化およびエンコード」モジュール503は、内部サンプリング周波数または外部サンプリング周波数を使って量子化および音響心理学を実行するよう適応されることができる。後者が選択される場合、サンプリング・レート変換器が考慮に入れられないので、実際の物理的周波数からの逸脱が存在する。しかしながら、本システムの設計を与えられると、この逸脱は無視できるほど十分に小さい。別の言い方をすれば、音響心理学は、MDCTの特定のサイズを与えられてサンプリング・レートの異なる範囲に適応する能力があるが、サンプリング・レート変換器に起因する信号の外部サンプリング・レートからの内部サンプリング・レートの逸脱は、外部サンプリング・レートについてのチューニング・パラメータの範囲外にはならないほど十分に小さい。

#### 【0052】

デコーダ側で高周波数再構成が有効にされていることに頼る「HFRエンコーダ」506は、たとえば64サブバンドの固定したQMF変換507のサブバンドに対して作用する。HFRエンコードおよびデコードのためのQMFフィルタバンクは常に固定した変換サイズなので（HFR処理に一般に使われる擬似QMFはMDCTのように変化する窓シーケンスおよび変換サイズをもつことができない）、関連するビデオ/オーディオ・フレーム・レートの大半について64サブバンドに固定されたQMF変換サイズを維持することが可能であるということは、本例示的实施形態の恩恵である。

#### 【0053】

HFRエンコーダ506は、QMFサンプルのグループ、たとえば48kHzにおける2048サンプル・フレームについての32サンプルに対して作用する（64サブバンドかける32QMFサンプルは2048時間領域サンプルに等しい）。表1で概説された提案された変換サイズおよびSRC因子を与えられて、HFRエンコーダ506は、ビデオ・フレーム・レートに従ってフレーム・サイズを適応させる必要があるだけである。これは上記で概説した32サンプルの例とはわずかに異なる仕方で時間的にQMFサブバンド・サンプルをグループ化することによる。SRC因子の選択は、フレームに対応するQMFサンプルの数が整数（やはり表1で概説している）であることを保証する。

#### 【0054】

さらに、HFRエンコーダ506は典型的には、バーク・スケールに関係した近似に従って、スペクトルを広い諸帯域にグループ分けする。この観点から、サブバンドのグループ分けがビデオ・フレーム・レートの関数として変化しないので、QMFサブバンドの数を一定に保つことは、本例示的实施形態の利点である。

#### 【0055】

MDCTからの量子化されたスペクトルは制御データとともにマルチプレクサ505によって多重化されて出力オーディオ・ビットストリームを形成する。

#### 【0056】

まとめると、図5に示した例示的实施形態は、次の性質をもつエンコーダ・システム501を（実際は、デコーダとして作用する対応するオーディオ処理システム100も）提

10

20

30

40

50

供する：

- ・内部および外部サンプリング・レートの間の差を最小にするために、SRC因子を1に近く保つ；
- ・オーディオ符号化のために機能することが知られており、過渡成分を扱うためにサブ変換にうまく細分できるMDCTベース変換サイズを使う；
- ・実装、チューニングを簡単にし、エンコーダからデコーダへの信号伝達オーバーヘッドを最小にするために種々のMDCTベース変換サイズの最小のものを使い、
- ・高周波数再構成および同様のアルゴリズムについてうまく機能することが知られており、整数個のQMFサンプルをフレームにグループ化することを許容するサブバンドの数（64）をもつ単一の固定サイズのQMFフィルタバンクを含む。

10

【0057】

エンコーダ・システム501のあるさらなる発展では、システムは帯域幅制限コンポーネント（図示せず）を含んでいてもよい。実際、入力サンプリング・レート48kHzの信号については、そのような帯域幅制限コンポーネントはオーディオ帯域幅をたとえば20kHz、すなわちナイキスト周波数より4kHz下制限してもよい。これは、デコーダ側でSRC 130の設計を軽減する。サンプリング・レート変換プロセスにおいてより急峻でない低域通過フィルタが使用できるからである。

【0058】

III. 等価物、拡張、代替その他

上記の記述を吟味すれば、当業者には本発明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、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発明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。

20

【0059】

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。

30

40

いくつかの態様を記載しておく。

〔態様1〕

ビデオ・フレームのストリームに関連付けられたオーディオ・ビットストリームを受け入れるよう構成されたオーディオ処理システムであって、前記オーディオ・ビットストリ

50

ームはフレームにセグメント分割されており、各ビデオ・フレームは、一つまたは複数のビデオ・フレームと等しい継続時間の関連付けられたオーディオ・データを担持する前記オーディオ・ビットストリーム中のフレームに時間的に関係しており、当該オーディオ処理システムは：

フロントエンド・コンポーネントであって：

前記オーディオ・ビットストリーム中のあるオーディオ・フレームに対応する量子化されたスペクトル係数を繰り返し受領し、中間信号の第一の周波数領域表現を出力するよう適応された量子化解除段；および

前記中間信号の前記第一の周波数領域表現を受領し、それに基づいて前記中間信号の時間領域表現を合成する逆変換段とを含むフロントエンド・コンポーネントと；

処理段であって：

前記中間信号の前記時間領域表現を受領し、前記中間信号の第二の周波数領域表現を出力する分解フィルタバンク；

前記中間信号の前記第二の周波数領域表現を受領し、処理されたオーディオ信号の周波数領域表現を出力する少なくとも一つの処理コンポーネント；および

前記処理されたオーディオ信号の前記周波数領域表現を受領し、前記処理されたオーディオ信号の時間領域表現を出力する合成フィルタバンクとを含む処理段と；

前記処理されたオーディオ信号の前記時間領域表現を受領し、ターゲット・サンプリング周波数でサンプリングされた再構成されたオーディオ信号を出力するサンプル・レート変換器とを有しており、

当該オーディオ処理システムは、少なくとも二つの異なるあらかじめ定義されたオーディオ・フレーム・レートのうちの任意のオーディオ・フレーム・レートをもつオーディオ・ビットストリームを受け入れるよう構成されており、前記フロントエンド・コンポーネントは可変の最大フレーム長をもち、前記異なるあらかじめ定義されたオーディオ・フレーム・レートに対応する少なくとも二つのモードで動作可能であることを特徴とする、オーディオ処理システム。

〔態様 2〕

前記あらかじめ定義されたオーディオ・フレーム・レートは、5%以下の変動範囲のオーディオ・フレーム・レートのサブグループを含み；

前記フロントエンド・コンポーネントは、前記サブグループ中のすべてのフレーム・レートについて同じフレーム長を使うよう構成されている、

態様 1 記載のオーディオ処理システム。

〔態様 3〕

前記量子化解除段は、量子化インデックスを、各フレーム長に関連付けられた、非一様な周波数依存の量子化レベルにマッピングし、各周波数について人間の聴覚に対して知覚不能な量子化ノイズ割り当てを達成するよう構成されている、態様 1 または 2 記載のオーディオ処理システム。

〔態様 4〕

前記フロントエンド・コンポーネントが、各最大フレーム長について、該最大フレーム長の同サイズのブロックへの細分である代替的なフレーム長で動作可能である、態様 1 ないし 3 のうちいずれか一項記載のオーディオ処理システム。

〔態様 5〕

前記オーディオ・ビットストリームの全ビットレートが、二チャンネル・オーディオ・データを担持するとき、200kbps未滿、好ましくは100kbps未滿である、態様 1 ないし 4 のうちいずれか一項記載のオーディオ処理システム。

〔態様 6〕

前記分解フィルタバンクおよび合成フィルタバンクが固定数の周波数帯域をもつフィルタバンクである、態様 1 ないし 5 のうちいずれか一項記載のオーディオ処理システム。

〔態様 7〕

前記量子化解除段が、臨界サンプリングされた量子化解除されたスペクトル係数を含む

10

20

30

40

50

前記中間オーディオ信号の量子化解除された周波数領域表現を出力するよう構成されており、

前記逆変換段が重複加算合成フィルタバンクを有する、  
態様 1 ないし 6 のうちいずれか一項記載のオーディオ処理システム。

〔態様 8〕

前記オーディオ・ビットストリームがマルチチャネル・オーディオ信号をパラメトリック・エンコードするものである、態様 1 ないし 7 のうちいずれか一項記載のオーディオ処理システム。

〔態様 9〕

態様 1 ないし 8 のうちいずれか一項記載のオーディオ処理システムと；

前記オーディオ・ビットストリームに関連付けられたビデオ・フレームのストリームを受け入れ、処理されたビデオ・フレームのストリームを出力するよう構成されたビデオ処理システムとを有するマルチメディア処理システムであって、

当該マルチメディア処理システムが、一つまたは複数のオーディオ・フレームおよびそれに時間的に関係した一つまたは複数のビデオ・フレームの複製および/または削除を実行するよう動作可能である、

マルチメディア処理システム。

〔態様 10〕

ビデオ・フレームのストリームと関連付けられたオーディオ・ビットストリームを処理する方法であって、前記オーディオ・ビットストリームはフレームにセグメント分割されており、各ビデオ・フレームは、一つまたは複数のビデオ・フレームと等しい継続時間の関連付けられたオーディオ・データを担持する、前記オーディオ・ビットストリーム中のフレームに時間的に関係しており、当該方法は：

・前記オーディオ・ビットストリーム中のあるオーディオ・フレームに対応する量子化されたスペクトル係数を繰り返し受領し、周波数から時間への変換が後続する逆量子化を実行する段階であって、それにより中間オーディオ信号の表現が得られる、段階と；

・前記中間オーディオ信号に対して周波数領域で少なくとも一つの処理段階を実行する段階と；

・処理されたオーディオ信号のサンプリング・レートをターゲット・サンプリング周波数に変更する段階であって、それにより再構成されたオーディオ信号の時間領域表現が得られる、段階とを含み、

前記オーディオ・ビットストリームの現在フレーム・レートを確立するさらなる段階および前記逆量子化および周波数から時間への変換が、異なる最大フレーム長をもつ少なくとも二つのモードで動作可能な機能コンポーネント（112）において実行され、現在モードは、前記オーディオ・ビットストリームの現在フレーム・レートに応答して選択される、

方法。

〔態様 11〕

前記逆量子化が、静的な第二の機能コンポーネント（114）において実行される、態様 10 記載の方法。

〔態様 12〕

前記オーディオ・ビットストリームがマルチチャネル・オーディオ信号をパラメトリック・エンコードするものである、態様 10 または 11 記載の方法。

〔態様 13〕

前記オーディオ・ビットストリームの確立された現在フレーム・レートが、いくつかの所定のフレーム・レートの一つに対してマッチングされ、前記所定のフレーム・レートは、5%以下の変動範囲のサブグループを含み、量子化されたスペクトル係数を繰り返し受領する前記の段階は、同じフレーム長をもつ前記機能コンポーネントの動作中に実行される、態様 10 ないし 12 のうちいずれか一項記載の方法。

〔態様 14〕

10

20

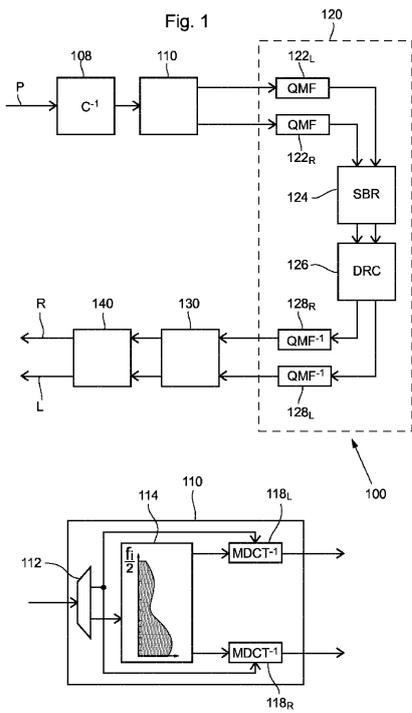
30

40

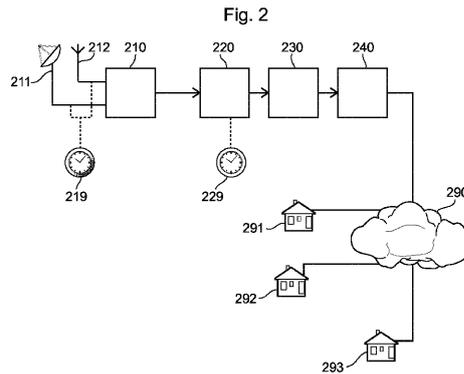
50

態様 10 ないし 13 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有するコンピュータ・プログラム・プロダクト。

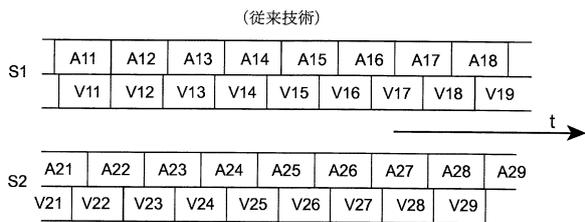
【 図 1 】



【 図 2 】



【 図 3 】



【 図 4 】

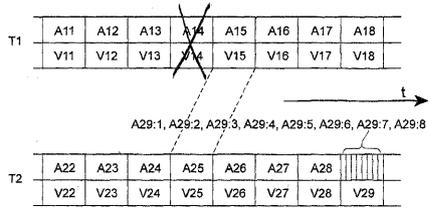
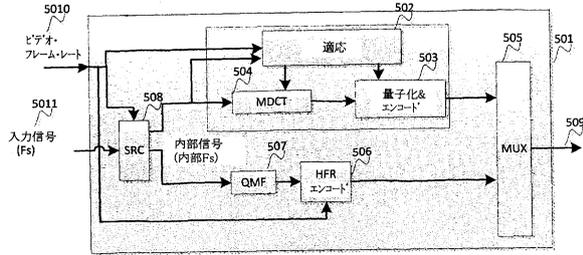


Fig. 4

【 図 5 】



## フロントページの続き

- (72)発明者 ブルンハーゲン, ハイコ  
スウェーデン王国, 113 30 ストックホルム, ガヴレガタン 12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, 113 30 ストックホルム, ガヴレガタン 12アー ドルビー スウェーデン アーバー内
- (72)発明者 グロッシェル, アレクサンダー  
ドイツ連邦共和国, 90429 ニュルンベルグ, ドイッチヘルンシュトラッセ 15 - 19, ドルビー・ジャーマニー・ゲーエムベーハー内
- (72)発明者 ポップ, イェンス  
ドイツ連邦共和国, 90429 ニュルンベルグ, ドイッチヘルンシュトラッセ 15 - 19, ドルビー・ジャーマニー・ゲーエムベーハー内
- (72)発明者 ローエデン, カール ヨナス  
スウェーデン王国, 113 30 ストックホルム, ガヴレガタン 12アー ドルビー スウェーデン アーバー内

審査官 上田 雄

- (56)参考文献 特表2002-536681(JP, A)  
米国特許出願公開第2006/0140591(US, A1)  
STEFAN MELTZER, MPEG-4 HE-AAC V2 - AUDIO CODING FOR TODAY'S DIGITAL MEDIA WORLD, [ONLINE], 2006年 1月31日, P1-12, URL, <http://tech.ebu.ch/Jahia/site/tech/cache/bypass/publications>
- (58)調査した分野(Int.Cl., DB名)  
G10L 19/00 - 19/26