

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7619675号
(P7619675)

(45)発行日 令和7年1月22日(2025.1.22)

(24)登録日 令和7年1月14日(2025.1.14)

(51)国際特許分類

F I

G 1 0 L 19/00 (2013.01)

G 1 0 L 19/00 3 3 0 B

G 1 0 L 21/034 (2013.01)

G 1 0 L 21/034

請求項の数 4 外国語出願 (全39頁)

(21)出願番号	特願2023-182909(P2023-182909)	(73)特許権者	507236292
(22)出願日	令和5年10月25日(2023.10.25)		ドルビー ラボラトリーズ ライセンシン
(62)分割の表示	特願2022-136006(P2022-136006)		グ コーポレイション
	)の分割		アメリカ合衆国 9 4 1 0 3 カリフォル
原出願日	平成26年1月15日(2014.1.15)		ニア州 サンフランシスコ マーケット
(65)公開番号	特開2023-175019(P2023-175019)		ストリート 1 2 7 5
	A)	(73)特許権者	510185767
(43)公開日	令和5年12月8日(2023.12.8)		ドルビー・インターナショナル・アーベ
審査請求日	令和5年10月25日(2023.10.25)		ー
(31)優先権主張番号	61/754,882		アイルランド ディー 0 2 ヴィーケイ 6
(32)優先日	平成25年1月21日(2013.1.21)		0 ダブリン グランド カナル ドックラ
(33)優先権主張国・地域又は機関			ンズ サー ジョン ロジャーソンズ キー
	米国(US)		7 7 ブロック シー
(31)優先権主張番号	61/809,250	(74)代理人	100107766
(32)優先日	平成25年4月5日(2013.4.5)		弁理士 伊東 忠重
	最終頁に続く		最終頁に続く

(54)【発明の名称】 異なる再生装置を横断するラウドネスおよびダイナミックレンジの最適化

(57)【特許請求の範囲】

【請求項 1】

エンコードされたオーディオ・ビットストリームの一つまたは複数のフレームをデコードするためのオーディオ処理装置であって、前記エンコードされたオーディオ・ビットストリームはオーディオ・データと、複数のダイナミックレンジ制御（DRC）プロファイルについてのメタデータとを含み、当該オーディオ処理装置は：

前記エンコードされたオーディオ・ビットストリームをパースして、前記エンコードされたオーディオ・データと、前記DRCプロファイルのうちの一つまたは複数についてのメタデータとを抽出するよう構成されたビットストリーム・パーサと；

前記エンコードされたオーディオ・データをデコードし、デコードされたオーディオ・データにDRC利得を適用するよう構成されたオーディオ・デコーダとを有しており、

各DRCプロファイルは少なくとも一つの装置型または聴取環境のために好適であり、前記オーディオ・デコーダは当該オーディオ処理装置または前記聴取環境についての情報に応答して前記DRCプロファイルのうちの一つまたは複数を選択し；

前記デコードされたオーディオ・データに適用される前記DRC利得は、広帯域利得であり、前記一つまたは複数の選択されたDRCプロファイルに対応し、

前記一つまたは複数の選択されたDRCプロファイルに対応する前記DRC利得は、前記エンコードされたオーディオ・ビットストリームの前記メタデータに含まれている前記一つまたは複数の選択されたDRCプロファイルについてのDRCパラメータから決定される、オーディオ処理装置。

10

20

【請求項 2】

エンコードされたオーディオ・ビットストリームの一つまたは複数のフレームをデコードするための、オーディオ処理装置によって実行される方法であって、前記エンコードされたオーディオ・ビットストリームはオーディオ・データと、複数のダイナミックレンジ制御（DRC）プロファイルについてのメタデータとを含み、当該方法は：

前記エンコードされたオーディオ・ビットストリームをパースして、前記エンコードされたオーディオ・データと、前記DRCプロファイルのうちの一つまたは複数についてのメタデータとを抽出する段階と；

前記エンコードされたオーディオ・データをデコードし、デコードされたオーディオ・データにDRC利得を適用する段階とを含み、

各DRCプロファイルは少なくとも一つの装置型または聴取環境のために好適であり、

前記DRCプロファイルのうちの一つまたは複数は、前記オーディオ処理装置または前記聴取環境についての情報に応答して選択され、

前記デコードされたオーディオ・データに適用される前記DRC利得は、広帯域利得であり、前記一つまたは複数の選択されたDRCプロファイルに対応し、

前記一つまたは複数の選択されたDRCプロファイルに対応する前記DRC利得は、前記エンコードされたオーディオ・ビットストリームの前記メタデータに含まれている前記一つまたは複数の選択されたDRCプロファイルについてのDRCパラメータから決定される、方法。

【請求項 3】

オーディオ・デコード装置によって実行されたときに該オーディオ・デコード装置に請求項 2 記載の方法を実行させる命令のシーケンスを有する、非一時的なコンピュータ読み取り可能な記憶媒体。

【請求項 4】

コンピュータ上で実行されたときに請求項 2 に記載の方法を実行するための実行可能命令を有するコンピュータ・プログラム・プロダクト。

【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1月21日に出願された米国仮特許出願第61/754,882号、2013年4月5日に出願された米国仮特許出願第61/809,250号、2013年5月16日に出願された米国仮特許出願第61/824,010号の優先権を主張するものである。

【0002】

発明の分野

一つまたは複数の実施形態は概括的にはオーディオ信号処理に、より詳細には、再生環境および装置に基づいて、オーディオ・コンテンツのラウドネスおよびダイナミックレンジ特性を示すメタデータをもつオーディオ・データ・ビットストリームを処理することに関する。

【背景技術】**【0003】**

背景セクションで論じられる主題は、単に背景セクションで言及されている結果として従来技術であると想定されるべきではない。同様に、背景セクションで言及されるまたは背景セクションの主題に関連する問題は、従来技術において以前に認識されていたと想定されるべきではない。背景セクションにおける主題は単に種々のアプローチを表わすのであり、そうしたアプローチもそれ自身が発明であることがある。

【0004】

オーディオ信号のダイナミックレンジは一般に、信号中に具現される音の最大および最小の可能な値の間の比であり、通例、デシベル（底が10）値として測定される。多くのオーディオ処理システムでは、ダイナミックレンジ制御（またはダイナミックレンジ圧縮（

10

20

30

40

50

DRC: dynamic range compression)) が、大きな音のレベルを低下させるおよび / または静かな音のレベルを増幅することで、広いダイナミックレンジの源コンテンツを、電子設備を使ってより簡単に記憶および再生されることができるよう狭い記録されたダイナミックレンジに収まるようにするために使われる。オーディオ/ビジュアル (AV) コンテンツについて、DRC機構を通じた圧縮のための「ヌル」点を定義するために、ダイアログ参照レベルが使用されてもよい。DRCは、ダイアログ参照レベルより低いコンテンツをブーストし、参照レベルより上のコンテンツをカットするはたらきをする。

【 0 0 0 5 】

既知のオーディオ・エンコード・システムでは、コンテンツの型および意図される使用に基づいてDRCレベルを設定するために、オーディオ信号に関連付けられたメタデータが使われる。DRCモードは、オーディオ信号に適用される圧縮の量を設定し、デコードの出力参照レベルを定義する。そのようなシステムは、エンコーダ中にプログラムされ、ユーザーによって選択される二つのDRCレベルに制限されることがある。たとえば、-31dB (ライン) のdialnorm (dialog normalization [ダイアログ正規化]) 値が伝統的に、AVRまたはフル・ダイナミックレンジ対応装置上で再生されるコンテンツのために使われ、-20dBのdialnorm値がテレビジョン・セットまたは同様の装置上で再生されるコンテンツのために使われる。この型のシステムは、DRCメタデータの二つの異なるセットを使うことを通じて、単一のオーディオ・ビットストリームが、二つの一般的だが非常に異なる再生シナリオにおいて使われることを許容する。しかしながら、そのようなシステムは、事前設定されたdialnorm値に制限され、今日、デジタル・メディアおよびインターネット・ベースのストリーミング技術の到来を通じて可能になっている幅広い多様な異なる再生装置および聴取環境における再生のために最適化されていない。

【 0 0 0 6 】

現在のメタデータ・ベースのオーディオ・エンコード・システムにおいては、オーディオ・データのストリームは、オーディオ・コンテンツ (たとえばオーディオ・コンテンツの一つまたは複数のチャネル) およびオーディオ・コンテンツの少なくとも一つの特性を示すメタデータの両方を含みうる。たとえば、AC-3ビットストリームでは、聴取環境に送達されるプログラムの音を変えることに使うことが特に意図されているいくつかのオーディオ・メタデータ・パラメータがある。メタデータ・パラメータの一つは、dialnormパラメータであり、これはオーディオ・プログラムにおいて現われるダイアログの平均ラウドネス・レベル (またはコンテンツの平均ラウドネス) を指示し、オーディオ再生信号レベルを決定するために使われる。

【 0 0 0 7 】

(それぞれ異なるdialnormパラメータをもつ) 異なるオーディオ・プログラム・セグメントのシーケンスを含むビットストリームの再生の間に、AC-3デコーダは、各セグメントのdialnormパラメータを使って、セグメントのダイアログの知覚されるラウドネスが一貫したレベルになるよう、セグメントの再生レベルまたはラウドネスを修正する型のラウドネス処理を実行する。エンコードされたオーディオ項目のシーケンスにおける各エンコードされたオーディオ・セグメント (項目) は (一般に) 異なるdialnormパラメータをもち、デコーダは、各項目についてのダイアログの再生レベルまたはラウドネスが同じまたは非常に似通っているように各項目のレベルをスケーリングする。ただし、これは、再生の間に前記項目の異なるものに対して異なる量の利得を適用することを要求することがある。

【 0 0 0 8 】

いくつかの実施形態では、dialnormパラメータはユーザーによって設定され、自動的に生成されるのではない。ただし、ユーザーによって値が設定されない場合にはデフォルトのdialnorm値がある。たとえば、コンテンツ・クリエイターは、AC-3エンコーダの外部の装置を用いてラウドネス測定を行ない、次いでdialnorm値を設定するために (オーディオ・プログラムの話されたダイアログのラウドネスを示す) 結果をエンコーダに転送してもよい。こうして、dialnormパラメータを正しく設定するためにコンテンツ・クリエイターに依拠している。

10

20

30

40

50

【 0 0 0 9 】

AC-3ビットストリームにおけるdialnormパラメータが正しくないことがありうるいくつかの異なる理由がある。第一に、各AC-3エンコーダは、コンテンツ・クリエイターによってdialnorm値が設定されない場合にビットストリームの生成の間に使われるデフォルトのdialnorm値をもつ。このデフォルト値は、オーディオの実際のダイアログ・ラウドネス・レベルとは実質的に異なることがありうる。第二に、たとえコンテンツ・クリエイターがラウドネスを測定し、dialnorm値をしかるべく設定するとしても、推奨されるラウドネス測定方法に従わないラウドネス測定アルゴリズムまたはメーターが使用されたことがありえ、正しくないdialnorm値につながる。第三に、たとえAC-3ビットストリームがコンテンツ・クリエイターによって正しく測定され設定されたdialnorm値をもって生成されたとしても、ビットストリームの伝送および/または記憶の間に中間モジュールによって正しくない値に変更されたことがありうる。たとえば、テレビジョン放送アプリケーションでは、AC-3ビットストリームがデコードされ、修正され、次いで正しくないdialnormメタデータ情報を使って再エンコードされることはめずらしくない。このように、AC-3ビットストリームに含まれるdialnorm値は正しくないまたは不正確であることがあり、よって聴取経験の品質に対してマイナスの影響をもつことがある。

10

【 0 0 1 0 】

さらに、dialnormパラメータは、対応するオーディオ・データのラウドネス処理状態（たとえば、どんな型（単数または複数）のラウドネス処理がそのオーディオ・データに対して実行されたか）を示さない。さらに、ドルビー・デジタル（DD）およびドルビー・デジタル・プラス（DD+）システムにおけるシステムのような現在配備されているラウドネスおよびDRCシステムは、消費者の居間または映画館でAVコンテンツをレンダリングするために設計されている。そのようなコンテンツを他の環境および聴取設備（たとえばモバイル装置）における再生のために適応させるには、AVコンテンツをその聴取環境に適応させるために、再生装置において「盲目的に」後処理が適用される必要がある。換言すれば、後処理器（またはデコーダ）は受領されたコンテンツのラウドネス・レベルが特定のレベル（たとえば - 31または - 20dB）にあることを想定し、後処理器はレベルを、特定の装置に好適なあらかじめ決定された固定目標レベルに設定する。想定されるラウドネス・レベルまたはあらかじめ決定された目標レベルが正しくない場合には、後処理はその意図された効果と反対の効果をもつことがありうる。すなわち、後処理は、出力オーディオをユーザーにとってより望ましくないものにすることがありうる。

20

30

【 0 0 1 1 】

開示される実施形態は、AC-3ビットストリーム、E-AC-3ビットストリームまたはドルビーEビットストリームとの使用に限定されないが、便宜上、そのようなビットストリームが、ラウドネス処理状態メタデータを含むシステムとの関連で論じられる。ドルビー、ドルビー・デジタル、ドルビー・デジタル・プラスおよびドルビーEはドルビー・ラボラトリーズ・ライセンシング・コーポレイションの商標である。ドルビー・ラボラトリーズは、それぞれドルビー・デジタルおよびドルビー・デジタル・プラスとして知られる、AC-3およびE-AC-3の独自の実装を提供している。

【 発明の概要 】

40

【 課題を解決するための手段 】

【 0 0 1 2 】

諸実施形態は、オーディオ・データに関連付けられたメタデータを含むビットストリームを受領し、ビットストリーム中のメタデータを解析してオーディオ再生装置の第一の群のためのラウドネス・パラメータが前記ビットストリーム中で利用可能であるかどうかを判定することによって、オーディオ・データをデコードする方法に向けられる。第一の群のためにそれらのパラメータが存在すると判定するのに応答して、処理コンポーネントは、オーディオをレンダリングするためにそれらのパラメータおよびオーディオ・データを使う。第一の群のためにそれらのラウドネス・パラメータが存在しないと判定するのに応答して、処理コンポーネントは前記第一の群の一つまたは複数の特性を解析し、前記一つ

50

または複数の特性に基づいて前記パラメータを決定する。本方法はさらに、前記パラメータおよびオーディオ・データを再生のために前記オーディオをレンダリングする下流のモジュールに送信することによってオーディオをレンダリングするために前記パラメータおよびオーディオ・データを使ってもよい。前記パラメータおよびオーディオ・データは、前記パラメータおよびオーディオ・データに基づいて前記オーディオ・データをレンダリングすることによってオーディオをレンダリングするために使われてもよい。

【 0 0 1 3 】

ある実施形態では、本方法は、受領されたオーディオ・ストリームをレンダリングする出力装置を判別することおよび出力装置がオーディオ再生装置の前記第一の群に属するかどうかを判定することをも含む。ここで、ストリーム中のメタデータを解析してオーディオ再生装置の前記第一の群のためのラウドネス・パラメータが利用可能であるかどうかを判定する前記段階は、出力装置がオーディオ再生装置の前記第一の群に属することを判別する段階の後に実行される。ある実施形態では、出力装置がオーディオ再生装置の前記第一の群に属することを判別する前記段階は、前記出力装置の素性を指示するまたは前記出力装置を含む装置の群の素性を指示する前記出力装置に接続されたモジュールからの指標を受領し、受領された指標に基づいて、前記出力装置がオーディオ再生装置の前記第一の群に属することを判別することを含む。

10

【 0 0 1 4 】

諸実施形態はさらに、上記のエンコード方法実施形態において記述される工程を実行する処理コンポーネントを含む装置またはシステムにさらに向けられる。

20

【 0 0 1 5 】

諸実施形態はさらに、オーディオ・データおよび前記オーディオ・データに関連付けられたメタデータを受領し、ビットストリーム中のメタデータを解析してオーディオ装置の第一の群のためのラウドネス・パラメータに関連するラウドネス情報が前記ストリーム中で利用可能であるかどうかを判定し、第一の群のために前記ラウドネス情報が存在すると判定するのに応答して、前記ストリームからラウドネス情報を決定し、オーディオ・データおよびラウドネス情報をオーディオのレンダリングにおいて使うために伝送し、あるいは前記ラウドネス情報が前記第一の群のために存在しない場合には、出力プロファイルに関連付けられたラウドネス情報を決定し、前記出力プロファイルについての決定されたラウドネス情報をオーディオのレンダリングにおいて使うために伝送することによって、オーディオ・データをデコードする方法に向けられる。ある実施形態では、出力プロファイルに関連付けられたラウドネス情報を決定する段階は、出力プロファイルの特性を解析し、前記特性に基づいて前記パラメータを決定することを含んでいてもよく、決定されたラウドネス情報を伝送することは、決定されたパラメータを伝送することを含んでいてもよい。ラウドネス情報は出力プロファイルのためのラウドネス・パラメータまたは出力プロファイルの特性を含んでいてもよい。ある実施形態では、本方法はさらに、伝送されるべき低ビットレートのエンコードされたストリームを決定することを含み、ラウドネス情報は一つまたは複数の出力プロファイルについての特性を含む。

30

【 0 0 1 6 】

諸実施形態はさらに、上記のデコード方法実施形態において記述される工程を実行する処理コンポーネントを含む装置またはシステムに向けられる。

40

【図面の簡単な説明】

【 0 0 1 7 】

以下の図面において、同様の参照符号は同様の要素を指すために使われる。以下の図面はさまざまな例を描いているが、本稿に記述される実装は図面に描かれる例に限定されない。

【図 1】いくつかの実施形態のもとでの、ラウドネスおよびダイナミックレンジの最適化を実行するよう構成されたオーディオ処理システムの実施形態のブロック図である。

【図 2】いくつかの実施形態のもとでの、図 1 のシステムにおける使用のためのエンコーダのブロック図である。

50

【図 3】いくつかの実施形態のもとでの、図 1 のシステムにおける使用のためのデコーダのブロック図である。

【図 4】AC-3 フレームを、それが分割された諸セグメントを含めて示す図である。

【図 5】AC-3 フレームの同期情報 (SI: Synchronization Information) セグメントを、それが分割された諸セグメントを含めて示す図である。

【図 6】AC-3 フレームのビットストリーム情報 (BSI: Bitstream Information) セグメントを、それが分割された諸セグメントを含めて示す図である。

【図 7】E-AC-3 フレームを、それが分割された諸セグメントを含めて示す図である。

【図 8】いくつかの実施形態のもとでの、エンコードされたビットストリームのある種のフレームおよびメタデータのフォーマットを示す表である。

10

【図 9】いくつかの実施形態のもとでの、ラウドネス処理状態メタデータのフォーマットを示す表である。

【図 10】いくつかの実施形態のもとでの、ラウドネスおよびダイナミックレンジの最適化を実行するよう構成されうる図 1 のオーディオ処理システムのより詳細なブロック図である。

【図 11】例示的な使用事例における多様な再生装置および背景聴取環境についての種々のダイナミックレンジ要求を示す表である。

【図 12】ある実施形態のもとでの、ダイナミックレンジ最適化システムのブロック図である。

【図 13】ある実施形態のもとでの、多様な異なる再生装置クラスのための異なるプロファイルの間のインターフェースを示すブロック図である。

20

【図 14】ある実施形態のもとでの、複数の定義されたプロファイルについて長期ラウドネスと短期ダイナミックレンジの相関を示す表である。

【図 15】ある実施形態のもとでの、オーディオ・コンテンツの種々の型についてのラウドネス・プロファイルの例を示す図である。

【図 16】ある実施形態のもとでの、再生装置およびアプリケーションを横断してラウドネスおよびダイナミックレンジを最適化する方法を示すフローチャートである。

【発明を実施するための形態】

【0018】

定義および命名法

30

請求項を含む本開示を通じて、信号またはデータ「に対して」動作を実行する（たとえば信号またはデータをフィルタリングする、スケーリングする、変換するまたは利得を適用する）という表現は、信号またはデータに対して直接的に、または信号またはデータの処理されたバージョンに対して（たとえば、予備的なフィルタリングまたは前処理を該動作の実行に先立って受けている前記信号のバージョンに対して）該動作を実行することを表わすために広義で使用される。「システム」という表現は、装置、システムまたはサブシステムを表わす広義で使用される。たとえば、デコーダを実装するサブシステムは、デコーダ・システムと称されてもよく、そのようなサブシステムを含むシステム（たとえば、複数の入力にตอบสนองしてX個の出力信号を生成するシステムであって、前記サブシステムが入力のうちのM個を生成し、他のX - M個の入力は外部源から受領されるもの）もデコーダ・システムと称されることがある。「プロセッサ」という用語は、データ（たとえばオーディオまたはビデオまたは他の画像データ）に対して動作を実行するよう（たとえばソフトウェアまたはファームウェアを用いて）プログラム可能または他の仕方で構成可能であるシステムまたは装置を表わす広義で使用される。プロセッサの例は、フィールド・プログラム可能なゲート・アレイ（または他の構成可能な集積回路またはチップセット）、オーディオまたは他のサウンド・データに対してパイプライン化された処理を実行するようプログラムされたおよび/または他の仕方で構成されたデジタル信号プロセッサ、プログラム可能な汎用プロセッサもしくはコンピュータおよびプログラム可能なマイクロプロセッサ・チップまたはチップセットを含む。

40

【0019】

50

「オーディオ・プロセッサ」および「オーディオ処理ユニット」という用語は交換可能に、オーディオ・データを処理するよう構成されたシステムを表わす広義で使用される。オーディオ処理ユニットの例は、エンコーダ（たとえばトランスコーダ）、デコーダ、コーデック、前処理システム、後処理システムおよびビットストリーム処理システム（時にビットストリーム処理ツールと称される）を含むがこれに限られない。「処理状態メタデータ」という表現（たとえば「ラウドネス処理状態メタデータ」という表現におけるような）は、対応するオーディオ・データ（処理状態メタデータをも含むオーディオ・データ・ストリームのオーディオ・コンテンツ）とは別個の異なるデータを指す。処理状態メタデータは、オーディオ・データに関連付けられ、対応するオーディオ・データのラウドネス処理状態（たとえばどの型（単数または複数）の処理がそのオーディオ・データに対してすでに実行されているか）を示し、任意的にはそのオーディオ・データの少なくとも一つの特徴または特性をも示す。何らかの実施形態では、処理状態メタデータのオーディオ・データとの関連付けは、時間同期的である。このように、現在の（最も最近受領または更新された）処理状態メタデータは、対応するオーディオ・データが同時的に、示される型（単数または複数）のオーディオ・データ処理の結果を含むことを示す。場合によっては、処理状態メタデータは、処理履歴および／または示される型の処理において使われるおよび／または示される型の処理から導出されるパラメータの一部または全部を含んでもよい。さらに、処理状態メタデータは、オーディオ・データから計算されたまたは抽出された、対応するオーディオ・データの少なくとも一つの特徴または特性を含んでもよい。処理状態メタデータはまた、対応するオーディオ・データのいかなる処理にも関係せず対応するオーディオ・データのいかなる処理から導出されたのでもない他のメタデータを含んでもよい。たとえば、サードパーティー・データ、追跡情報、識別子、所有権があるか標準かの情報、ユーザー注釈データ、ユーザー選好データなどが、特定のオーディオ処理ユニットによって加えられて他のオーディオ処理ユニットに渡されてもよい。

【 0 0 2 0 】

「ラウドネス処理状態メタデータ（loudness processing state metadata）」（または「LPSM」）という表現は、対応するオーディオ・データのラウドネス処理状態（たとえばどの型（単数または複数）のラウドネス処理がそのオーディオ・データに対してすでに実行されているか）を、任意的にはまた対応するオーディオ・データの少なくとも一つの特徴または特性（たとえばラウドネス）をも示す処理状態メタデータを表わす。ラウドネス処理状態メタデータは、（単独で考えると）ラウドネス処理状態メタデータではないデータ（たとえば他のメタデータ）を含んでもよい。用語「結合する」または「結合される」という用語は、直接的または間接的な接続を意味するために使われる。

【 0 0 2 1 】

種々の目標ラウドネス値を要求または使用し、種々のダイナミックレンジ機能をもつさまざまな装置を横断してオーディオのラウドネスおよびダイナミックレンジを非破壊的に正規化するオーディオ・エンコーダ／デコーダのためのシステムおよび方法が記述される。いくつかの実施形態に基づく方法および機能コンポーネントは、一つまたは複数の装置プロファイルについて、オーディオ・コンテンツについての情報をエンコーダからデコーダに送る。装置プロファイルは、一つまたは複数の装置について所望される目標ラウドネスおよびダイナミックレンジを指定する。異なる「公称」ラウドネス目標をもつ新しい装置プロファイルがサポートされることができるよう、システムは拡張可能である。

【 0 0 2 2 】

ある実施形態では、システムは、エンコーダにおけるラウドネス制御およびダイナミックレンジ要求に基づいて適切な利得を生成し、データ・レートを低減するために、デコーダにおいて、もとの利得のパラメータ化を通じたエンコーダからの制御のもとで、前記利得を生成する。ダイナミックレンジ・システムは、ラウドネス制御を実装するための二つの機構を含む：オーディオがどのように再生されるかにおけるコンテンツ・クリエーターの制御を提供する芸術的ダイナミックレンジ・プロファイルと、さまざまな再生プロファイルについて過負荷が起こらないことを保証する別個の保護機構である。システムは、他

10

20

30

40

50

のメタデータ（内部または外部）パラメータがラウドネスおよびダイナミックレンジ利得および／またはプロファイルを適正に制御するために使われることを許容するようにも構成されている。デコーダは、デコーダ側ラウドネスおよびダイナミックレンジ設定／処理を利用するnチャンネル補助入力をサポートするよう構成されている。

【0023】

いくつかの実施形態では、ラウドネス処理状態メタデータ（LPSM）が、オーディオ・ビットストリームのメタデータ・セグメントの一つまたは複数のリザーブされたフィールド（またはスロット）に埋め込まれる。オーディオ・ビットストリームは他のセグメント（オーディオ・データ・セグメント）においてオーディオ・データをも含む。たとえば、ビットストリームの各フレームの少なくとも一つのセグメントがLPSMを含み、フレームの少なくとも一つの他のセグメントが対応するオーディオ・データ（すなわち、そのラウドネス処理状態およびラウドネスが前記LPSMによって示されているオーディオ・データ）を含む。いくつかの実施形態では、LPSMのデータ・ボリュームは、オーディオ・データを搬送するために割り当てられたビットレートに影響することなく搬送されるのに十分なほど小さくてもよい。

【0024】

オーディオ・データ処理チェーンにおいてラウドネス処理状態メタデータを通信することは、処理チェーン（またはコンテンツ・ライフサイクル）を通じて二つ以上のオーディオ処理ユニットが互いに縦続して機能する必要があるときに特に有用である。オーディオ・ビットストリーム中にラウドネス処理状態メタデータを含めることなくしては、チェーン内で二つ以上のオーディオ・コーデックが利用され、メディア消費装置（またはビットストリームのオーディオ・コンテンツのレンダリング点）までのビットストリームの行程の間でシングル・エンドのボリューム平準化が二回以上適用されるときに、品質、レベルおよび空間的劣化のようなメディア処理問題が起こりうる。

【0025】

ラウドネスおよびダイナミックレンジ・メタデータ処理システム

図1は、ある種のメディア処理（たとえば前処理および後処理）コンポーネントを使ういくつかの実施形態のもとでの、ラウドネスおよびダイナミックレンジの最適化を実行するよう構成されていてもよいオーディオ処理システムの実施形態のブロック図である。図1は、例示的なオーディオ処理チェーン（オーディオ・データ処理システム）を示しており、該システムの要素の一つまたは複数は、本発明のある実施形態に基づいて構成されていてもよい。図1のシステム10は、図のように一緒に結合された以下の要素を含む：前処理ユニット12、エンコーダ14、信号解析およびメタデータ補正ユニット16、トランスコーダ18、デコーダ20および後処理ユニット24。図示したシステムに対する諸変形では、要素の一つまたは複数が省略され、あるいは追加的なオーディオ・データ処理ユニットが含まれる。たとえば、ある実施形態では、後処理ユニット22は、別個のユニットではなくデコーダ20の一部である。

【0026】

いくつかの実装では、図1の前処理ユニットは、入力11としてオーディオ・コンテンツを含むPCM（時間領域）サンプルを受け入れ、処理されたPCMサンプルを出力するよう構成されている。エンコーダ14は、入力としてPCMサンプルを受け入れ、オーディオ・コンテンツを示すエンコードされた（たとえば圧縮された）オーディオ・ビットストリームを出力するよう構成されていてもよい。オーディオ・コンテンツを示すビットストリームのデータは本稿では時に「オーディオ・データ」と称される。ある実施形態では、エンコーダからのオーディオ・ビットストリーム出力は、オーディオ・データのほかにラウドネス処理状態メタデータを（および任意的には他のメタデータも）含む。

【0027】

信号解析およびメタデータ補正ユニット16は、入力として一つまたは複数のエンコードされたオーディオ・ビットストリームを受け入れて、信号解析を実行することにより、各エンコードされたオーディオ・ビットストリーム中の処理状態メタデータが正しいかど

10

20

30

40

50

うかを判定（たとえば有効確認）してもよい。いくつかの実施形態では、有効確認は、図 2 に示される要素 102 のような状態有効確認器コンポーネントによって実行されてもよく、一つのそのような有効確認技法はのちに状態有効確認器 102 のコンテキストにおいて述べる。いくつかの実施形態では、ユニット 16 はエンコーダに含められ、有効確認はユニット 16 または有効確認器 102 のいずれかによってなされる。信号解析およびメタデータ補正ユニットが含まれているメタデータが無効であることを見出す場合には、メタデータ補正ユニット 16 は、正しい値（単数または複数）を決定するための信号解析を実行し、正しくない値（単数または複数）を決定された正しい値で置き換える。このように、信号解析およびメタデータ補正ユニットから出力されるエンコードされた各オーディオ・ビットストリームは、エンコードされたオーディオ・データのほかに、補正された処理状態メタデータを含んでいてもよい。信号解析およびメタデータ補正ユニット 16 は、前処理ユニット 12、エンコーダ 14、トランスコーダ 18、デコーダ 20 または後処理ユニット 22 の一部であってもよい。あるいはまた、信号解析およびメタデータ補正ユニット 16 は、オーディオ処理チェーン内の別個のユニットまたは別のユニットの一部であってもよい。

10

【0028】

トランスコーダ 18 はエンコードされたオーディオ・ビットストリームを入力として受け入れ、修正された（たとえば異なる仕方でエンコードされた）オーディオ・ビットストリームを応答として出力してもよい（たとえば、入力ストリームをデコードし、デコードされたストリームを異なるエンコード・フォーマットにおいてエンコードし直すことにより）。トランスコーダから出力されたオーディオ・ビットストリームは、エンコードされたオーディオ・データのほかにラウドネス処理状態メタデータを（および任意的には他のメタデータも）含む。メタデータは、ビットストリーム中に含まれていてもよい。

20

【0029】

図 1 のデコーダ 20 は、入力としてエンコードされた（たとえば圧縮された）オーディオ・ビットストリームを受け入れ、（応答して）デコードされた PCM オーディオ・サンプルのストリームを出力してもよい。ある実施形態では、デコーダの出力は、以下のうちの任意のものであるまたはそれを含む：オーディオ・サンプルのストリームおよび入力のエンコードされたビットストリームから抽出されたラウドネス処理状態メタデータ（および任意的には他のメタデータも）の対応するストリーム；オーディオ・サンプルのストリームおよび入力のエンコードされたビットストリームから抽出されたラウドネス処理状態メタデータ（および任意的には他のメタデータも）から決定された制御ビットの対応するストリーム；または処理状態メタデータや処理状態メタデータから決定された制御ビットの対応するストリームのない、オーディオ・サンプルのストリーム。この最後の場合、デコーダは、抽出されたメタデータやそれから決定される制御ビットを出力しなくても、入力のエンコードされたビットストリームからラウドネス処理状態メタデータ（および/または他のメタデータ）を抽出し、抽出されたメタデータに対する少なくとも一つの動作（たとえば有効確認）を実行してもよい。

30

【0030】

本発明のある実施形態に従って図 1 の後処理ユニットを構成することにより、後処理ユニット 22 は、デコードされた PCM オーディオ・サンプルのストリームを受け入れ、該サンプルとともに受領されるラウドネス処理状態メタデータ（および任意的には他のメタデータも）または該サンプルとともに受領される制御ビット（ラウドネス処理状態メタデータおよび任意的には他のメタデータからデコーダによって決定される）を使って、それに対する後処理（たとえばオーディオ・コンテンツのボリューム平準化）を実行するよう構成されている。後処理ユニット 22 は任意的には、一つまたは複数のスピーカーによる再生のために後処理されたオーディオ・コンテンツをレンダリングするようにも構成されている。これらのスピーカーは、コンピュータ、テレビジョン、ステレオ・システム（家庭または映画館）、携帯電話および他のポータブル再生装置のような多様な異なる聴取装置または再生設備品目の任意のものにおいて具現されうる。スピーカーは任意の適切なサイ

40

50

ズおよび電力定格のものであってもよく、自立型のドライバ、スピーカー・エンクロージャー、サラウンドサウンド・システム、サウンドバー、ヘッドフォン、イヤープッドなどの形で提供されてもよい。

【0031】

いくつかの実施形態は、向上されたオーディオ処理チェーンであって、オーディオ処理ユニット（たとえばエンコーダ、デコーダ、トランスコーダおよび前処理および後処理ユニット）が、オーディオ・データに適用されるそれぞれの処理を、それぞれオーディオ処理ユニットによって受領されるラウドネス処理状態メタデータによって示されるメディア・データの同時の状態に従って適応させるものを提供する。システム100の任意のオーディオ処理ユニット（たとえば図1のエンコーダまたはトランスコーダ）へのオーディオ・データ入力11は、オーディオ・データ（たとえばエンコードされたオーディオ・データ）のほかにラウドネス処理状態メタデータ（および任意的には他のメタデータ）を含んでいてもよい。メタデータは、いくつかの実施形態によれば、別の要素または別の源によって入力オーディオに含められたものであってもよい。（メタデータをもつ）入力オーディオを受領する処理ユニットは、少なくとも一つの動作を、メタデータに対して（たとえば有効確認）、あるいはメタデータに応答して（たとえば入力オーディオの適応的な処理）実行し、また、その出力オーディオにおいて、前記メタデータ、前記メタデータの処理されたバージョンまたは前記メタデータから決定された制御ビットを含めるよう構成されていてもよい。

【0032】

オーディオ処理ユニット（またはオーディオ・プロセッサ）のある実施形態は、オーディオ・データに対応するラウドネス処理状態メタデータによって示されるオーディオ・データの状態に基づいてオーディオ・データの適応的な処理を実行するよう構成されている。いくつかの実施形態では、適応的な処理はラウドネス処理である（またはそれを含む）（メタデータが、そのオーディオ・データに対してラウドネス処理またはそれに類似した処理がすでに実行されているのではないことを示す場合）が、ラウドネス処理ではない（またはそれを含まない）（メタデータが、そのオーディオ・データに対してそのようなラウドネス処理またはそれに類似した処理がすでに実行されていることを示す場合）。いくつかの実施形態では、適応的な処理は、オーディオ処理ユニットがラウドネス処理状態メタデータによって示されるオーディオ・データの状態に基づいてオーディオ・データの他の適応的な処理を実行することを保証するための、メタデータ有効確認である（またはそれを含む）（たとえば、メタデータ有効確認サブユニットにおいて実行される）。いくつかの実施形態では、有効確認は、オーディオ・データに関連付けられた（たとえばオーディオ・データと一緒にビットストリームに含まれた）ラウドネス処理状態メタデータの信頼性を決定する。たとえば、メタデータが信頼できると有効確認される場合、ある型のすでに実行されたオーディオ処理からの結果が再使用されてもよく、同じ型のオーディオ処理の追加的な実行が回避されうる。他方、メタデータが細工されている（または他の事情で信頼できない）ことが見出される場合には、（信頼できないメタデータによって示されるところの）以前に実行されたと標榜される型のメディア処理は、オーディオ処理ユニットによって繰り返されてもよく、および/またはメタデータおよび/またはオーディオ・データに対してオーディオ処理ユニットによって他の処理が実行されてもよい。オーディオ処理ユニットは、該ユニットが（たとえば抽出された暗号学的な値および参照の暗号学的な値の一致に基づいて）処理状態メタデータが有効であると判定する場合、向上したメディア処理チェーンにおける下流の他のオーディオ処理ユニットに対して、（たとえばメディア・ビットストリーム中に存在する）ラウドネス処理状態メタデータが有効であることを信号伝達するよう構成されていてもよい。

【0033】

図1の実施形態について、前処理コンポーネント12はエンコーダ14の一部であってもよく、後処理コンポーネント22はデコーダ22の一部であってもよい。あるいはまた、前処理コンポーネント12はエンコーダ14とは別個である機能コンポーネントにおい

て具現されていてもよい。同様に、後処理コンポーネント 22 は、デコーダ 20 とは別個の機能コンポーネントにおいて具現されてもよい。

【0034】

図2は、図1のシステム10との関連で使用されうるエンコーダ100のブロック図である。エンコーダ100のコンポーネントまたは要素の任意のものは、ハードウェア、ソフトウェアまたはハードウェアとソフトウェアの組み合わせにおいて、一つまたは複数のプロセスおよび/または一つまたは複数の回路（たとえばASIC、FPGAまたは他の集積回路）として実装されうる。エンコーダ100は、図のように接続された、フレーム・バッファ110、パーサ111、デコーダ101、オーディオ状態有効確認器102、ラウドネス処理段103、オーディオ・ストリーム選択段104、エンコーダ105、詰め込み器（stuffer）/フォーマッタ段107、メタデータ生成段106、ダイアログ・ラウドネス測定サブシステム108およびフレーム・バッファ109を有する。任意的には、エンコーダ100は他の処理要素（図示せず）も含む。（トランスコーダである）エンコーダ100は、入力オーディオ・ビットストリーム（これはたとえばAC-3ビットストリーム、E-AC-3ビットストリームまたはドルビーEビットストリームのうちの一つであってもよい）をエンコードされた出力オーディオ・ビットストリーム（これはたとえばAC-3ビットストリーム、E-AC-3ビットストリームまたはドルビーEビットストリームのうちの別の一つであってもよい）に変換するよう構成されている。これは、入力ビットストリームに含まれるラウドネス処理状態メタデータを使って適応的および自動化されたラウドネス処理を実行することによることを含む。たとえば、エンコーダ100は、入力ドルビーEビットストリーム（製作および放送施設において典型的に使われるが、放送されたオーディオ・プログラムを受信する消費者装置においては典型的には使われないフォーマット）を、AC-3またはE-AC-3フォーマットの形のエンコードされた出力オーディオ・ビットストリーム（消費者装置への放送に好適）に変換するよう構成されていてもよい。

【0035】

図2のシステムはまた、エンコードされたオーディオの送達サブシステム150（これはエンコーダ100から出力されるエンコードされたビットストリームを記憶するおよび/または送達する）と、デコーダ152とを含む。エンコーダ100から出力されるエンコードされたオーディオ・ビットストリームは、サブシステム150によって（たとえばDVDまたはブルーレイ・ディスクの形で）記憶されても、あるいはサブシステム150（これは伝送リンクまたはネットワークを実装していてもよい）によって伝送されてもよく、あるいはサブシステム150によって記憶および伝送の両方をされてもよい。デコーダ152は、サブシステム150を介して受領する（エンコーダ100によって生成された）エンコードされたオーディオ・ビットストリームをデコードするよう構成されている。これは、ビットストリームの各フレームからラウドネス処理状態メタデータ（LPSM）を抽出し、デコードされたオーディオ・データを生成することによることを含む。ある実施形態では、デコーダ152は、デコードされたオーディオ・データに対してLPSMを使って適応ラウドネス処理を実行し、および/またはデコードされたオーディオ・データおよびLPSMを、デコードされたオーディオ・データに対してLPSMを使って適応ラウドネス処理を実行するよう構成されている後処理器に転送するよう構成される。任意的には、デコーダ152は、サブシステム150から受領されたエンコードされたオーディオ・ビットストリームを（たとえば非一時的な仕方）で記憶するバッファを含む。

【0036】

エンコーダ100およびデコーダ152のさまざまな実装が、本稿に記載される種々の実施形態を実行するよう構成される。フレーム・バッファ110は、エンコードされた入力オーディオ・ビットストリームを受領するよう結合されたバッファ・メモリである。動作では、バッファ110は、エンコードされたオーディオ・ビットストリームの少なくとも一つのフレームを（たとえば非一時的な仕方）で記憶し、エンコードされたオーディオ・ビットストリームのフレームのシーケンスがバッファ110からパーサ111に呈される。パーサ111は、ラウドネス処理メタデータ（LPSM）および他のメタデータを、エ

ンコードされた入力オーディオの各フレームから抽出し、少なくともLPSMをオーディオ状態有効確認器102、ラウドネス処理段103、段106およびサブシステム108に呈し、エンコードされた入力オーディオからオーディオ・データを抽出し、該オーディオ・データをデコーダ101に呈するよう結合され、構成されている。エンコーダ100のデコーダ101は、オーディオ・データをデコードしてデコードされたオーディオ・データを生成し、該デコードされたオーディオ・データをラウドネス処理段103、オーディオ・ストリーム選択段104、サブシステム108および任意的には状態有効確認器102にも呈するよう構成されている。

【0037】

状態有効確認器102は、それに対して呈されるLPSM（および任意的には他のメタデータ）を認証し、有効確認するよう構成される。いくつかの実施形態では、LPSMは、（たとえば本発明のある実施形態に従って）入力ビットストリームに含まれたデータ・ブロックである（または該データ・ブロックに含まれる）。該ブロックは、LPSM（および任意的には他のメタデータも）および/または基礎になるオーディオ・データを処理するための暗号学的ハッシュ（ハッシュ・ベースのメッセージ認証コードまたは「HMAC」）を含んでいてもよい（デコーダ101から有効確認器102に提供される）。該データ・ブロックは、これらの実施形態において、デジタル署名されてもよい。それにより、下流のオーディオ処理ユニットは比較的容易に、該処理状態メタデータを認証および有効確認しうる。

【0038】

たとえば、HMACは、ダイジェストを生成するために使われ、本発明のビットストリームに含まれる保護値（単数または複数）は該ダイジェストを含んでいてもよい。該ダイジェストは、AC-3フレームについては、以下のように生成されてもよい：（1）AC-3データおよびLPSMがエンコードされたのち、フレーム・データ・バイト（連結されたフレーム・データ#1およびフレーム・データ#2）およびLPSMデータ・バイトが、ハッシュ関数HMACのための入力として使われる。補助データ・フィールド内に存在していてもよい他のデータは、このダイジェストを計算するためには考慮に入れられない。そのような他のデータは、AC-3データにもLPSMデータにも属さないバイトであってもよい。LPSMに含まれる保護ビットは、HMACダイジェストを計算するためには考慮されなくてもよい。（2）ダイジェストが計算されたのち、該ダイジェストは保護ビットのためにリザーブされているフィールドにおいてビットストリームに書き込まれる。（3）完全なAC-3フレームの生成の最後の段階は、CRC検査の計算である。これは、フレームのいちばん最後に書かれ、LPSMビットを含む、このフレームに属するすべてのデータが考慮に入れられる。

【0039】

一つまたは複数のHMACでない暗号学的方法の任意のものを含むがそれに限定されない他の暗号学的方法が、LPSMおよび/または基礎になるオーディオ・データの安全な伝送および受領を保証するための（たとえば有効確認器102における）LPSMの有効確認のために使われてもよい。たとえば、（そのような暗号学的方法を使う）有効確認は、オーディオ・ビットストリームの実施形態を受領する各オーディオ処理ユニットにおいて実行され、ビットストリームに含まれるラウドネス処理状態メタデータおよび対応するオーディオ・データが（該メタデータによって示されるような）特定のラウドネス処理を受けている（および/または特定のラウドネス処理から帰結する）ものであり、そのような特定のラウドネス処理の実行後に修正されていないかどうかを判定することができる。

【0040】

状態有効確認器102は、有効確認動作の結果を示すために、オーディオ・ストリーム選択段104、メタデータ生成器106およびダイアログ・ラウドネス測定サブシステム108に制御データを呈する。該制御データに応答して、段104は次のいずれかを選択する（そしてエンコーダ105に伝える）ことができる：（1）（たとえば、LPSMがデコーダ101から出力されたオーディオ・データが特定の型のラウドネス処理を受けていないことを示し、有効確認器102からの制御ビットがLPSMが有効であることを示すと

10

20

30

40

50

き) ラウドネス処理段 103 の適応的に処理された出力; または (2) (たとえば、LPSM がデコーダ 101 から出力されたオーディオ・データが段 103 によって実行されるはずの特定の型のラウドネス処理をすでに受けていることを示し、有効確認器 102 からの制御ビットが LPSM が有効であることを示すとき) デコーダ 101 から出力された前記オーディオ・データ。

【0041】

エンコーダ 100 の段 103 は、デコーダ 101 から出力されたデコードされたオーディオ・データに対して、デコーダ 101 によって抽出された LPSM によって示される一つまたは複数のオーディオ・データ特性に基づいて、適応的なラウドネス処理を実行するよう構成されている。段 103 は、適応的な変換領域のリアルタイムのラウドネスおよびダイナミックレンジ制御プロセッサであってもよい。段 103 はユーザー入力 (たとえばユーザー目標ラウドネス / ダイナミックレンジ値または dialnorm 値) または他のメタデータ入力 (たとえば、一つまたは複数の型のサードパーティー・データ、追跡情報、識別子、所有権があるか標準かの情報、ユーザー注釈データ、ユーザー選好データなど) および / または (たとえばフィンガープリンティング・プロセスからの) 他の情報を受領して、そのような入力を、デコーダ 101 から出力されるデコードされたオーディオ・データを処理するために使ってもよい。

【0042】

ダイアログ・ラウドネス測定サブシステム 108 は、有効確認器 102 からの制御ビットが LPSM が無効であることを示す場合は、たとえばデコーダ 101 によって抽出された LPSM (および / または他のメタデータ) を使って、ダイアログ (または他の発話) を示す (デコーダ 101 からの) デコードされたオーディオの諸セグメントのラウドネスを決定するよう動作してもよい。有効確認器 102 からの制御ビットが LPSM が有効であることを示す場合には、LPSM が (デコーダ 101 からの) デコードされたオーディオのダイアログ (または他の発話) セグメントの以前に決定されたラウドネスを示しているときは、ダイアログ・ラウドネス測定サブシステム 108 の動作は無効にされてもよい。

【0043】

オーディオ・コンテンツにおけるダイアログのレベルを便利かつ簡単に測定するための有用なツール (たとえばドルビー LM100 ラウドネス・メーター) が存在している。APU (たとえばエンコーダ 100 の段 108) のいくつかの実施形態は、オーディオ・ビットストリーム (たとえば、エンコーダ 100 のデコーダ 101 から段 108 に呈されるデコードされた AC-3 ビットストリーム) のオーディオ・コンテンツの平均ダイアログ・ラウドネスを測定するためにそのようなツールを含むよう (またはそのようなツールの機能を実行するよう) 実装される。段 108 がオーディオ・データの真の平均ダイアログ・ラウドネスを測定するよう実装される場合、測定は、オーディオ・コンテンツの、主として発話を含んでいる諸セグメントを単離する段階を含んでいてもよい。主として発話であるオーディオ・セグメントは、次いで、ラウドネス測定アルゴリズムに従って処理される。AC-3 ビットストリームからデコードされるオーディオ・データについては、このアルゴリズムは、(国際規格 ITU-R BS.1770 に従う) 標準的な K 重み付けされたラウドネス指標 (K-weighted loudness measure) であってもよい。あるいはまた、他のラウドネス指標 (たとえばラウドネスの音響心理学的モデルに基づくもの) が使われてもよい。

【0044】

発話セグメントの単離は、オーディオ・データの平均ダイアログ・ラウドネスを測定するためには本質的ではないが、指標の精度を改善し、聴取者の観点からの、より満足のいく結果を与える。すべてのオーディオ・コンテンツがダイアログ (発話) を含むのではないので、オーディオ・コンテンツ全体のラウドネス指標は、発話が存在していたとした場合の、当該オーディオのダイアログ・レベルの十分な近似を提供しうる。

【0045】

メタデータ生成器 106 は、エンコーダ 100 から出力されるエンコードされたビットストリームに段 107 によって含まれるメタデータを生成する。メタデータ生成器 10

10

20

30

40

50

6 は、段 1 0 7 に、エンコーダ 1 0 1 によって抽出された LPSM（および / または他のメタデータも）を渡してもよい（たとえば、有効確認器 1 0 2 からの制御ビットが LPSM および / または他のメタデータが有効であることを示す場合）、あるいは新たな LPSM（および / または他のメタデータも）を生成して、該新たなメタデータを段 1 0 7 に呈してもよい（たとえば、有効確認器 1 0 2 からの制御ビットが、デコーダ 1 0 1 によって抽出された LPSM および / または他のメタデータが無効であることを示す場合）。あるいは、段 1 0 7 に対して、デコーダ 1 0 1 によって抽出されたメタデータと新たに生成されたメタデータとの組み合わせを呈してもよい。メタデータ生成器 1 0 6 は、サブシステム 1 0 8 によって生成されたラウドネス・データと、サブシステム 1 0 8 によって実行されたラウドネス処理の型を示す少なくとも一つの値とを、エンコーダ 1 0 0 から出力されるエンコードされたビットストリームに含めるために、段 1 0 7 に対して呈する LPSM 中に含めてもよい。メタデータ生成器 1 0 6 は、エンコードされたビットストリームに含めるべき LPSM（および任意的には他のメタデータも）および / またはエンコードされたビットストリームに含めるべき基礎になるオーディオ・データの解釈、認証または有効確認の少なくとも一つについて有用な保護ビット（これはハッシュ・ベースのメッセージ認証コードまたは「HMAC」からなっているてもよく、あるいはそれを含んでいてもよい）を生成してもよい。メタデータ生成器 1 0 6 はそのような保護ビットを、エンコードされたビットストリーム中に含めるために、段 1 0 7 に提供してもよい。

【 0 0 4 6 】

ある実施形態では、ダイアログ・ラウドネス測定サブシステム 1 0 8 は、デコーダ 1 0 1 から出力されたオーディオ・データを処理して、それに応答して、ラウドネス値（たとえば、ゲーティングされたおよびゲーティングされないダイアログ・ラウドネス値）およびダイナミックレンジ値を生成する。これらの値に応答して、メタデータ生成器 1 0 6 は、エンコーダ 1 0 0 から出力されるエンコードされたビットストリームに（詰め込み器 / フォーマッタ 1 0 7 によって）含めるためにラウドネス処理状態メタデータ（LPSM）を生成してもよい。ある実施形態では、ラウドネスは、ITU-R BS.1770-1 および ITU-R BS.1770-2 規格または他の同様のラウドネス測定規格によって規定される技法に基づいて計算されてもよい。ゲーティングされたラウドネスは、ダイアログ・ゲーテッド・ラウドネス（dialog-gated loudness）または相対ゲーテッド・ラウドネス（relative-gated loudness）またはこれらのゲーティングされたラウドネス型の組み合わせであることができ、システムは、用途の要求およびシステム制約条件に依存して適切なゲーティング・ブロックを用いることができる。

【 0 0 4 7 】

追加的、任意的または代替的に、エンコーダ 1 0 0 の 1 0 6 および / または 1 0 8 のサブシステムは、オーディオ・データの追加的な解析を実行して、段 1 0 7 から出力されるエンコードされたビットストリームに含めるための、オーディオ・データの少なくとも一つの特性を示すメタデータを生成してもよい。エンコーダ 1 0 5 は、選択段 1 0 4 から出力されたオーディオ・データを（たとえばそれに対して圧縮を実行することによって）エンコードし、段 1 0 7 から出力されるエンコードされたビットストリームに含めるために、エンコードされたオーディオを段 1 0 7 に呈する。

【 0 0 4 8 】

段 1 0 7 は、エンコーダ 1 0 5 からのエンコードされたオーディオと生成器 1 0 6 からのメタデータ（LPSM を含む）とを多重化して、段 1 0 7 から出力される、エンコードされたビットストリームを生成する。それにより、エンコードされたビットストリームは、ある実施形態によって指定されるフォーマットをもつようにされる。フレーム・バッファ 1 0 9 は、段 1 0 7 から出力されるエンコードされたオーディオ・ビットストリームの少なくとも一つのフレームを（たとえば非一時的な仕方でも）記憶するバッファ・メモリである。次いで、エンコードされたオーディオ・ビットストリームのそれらのフレームのシーケンスが、バッファ 1 0 9 から、エンコーダ 1 0 0 からの出力として、送達システム 1 5 0 に呈される。

10

20

30

40

50

【 0 0 4 9 】

メタデータ生成器 1 0 6 によって生成され、段 1 0 7 によって、エンコードされたビットストリームに含められたLPSMは、対応するオーディオ・データのラウドネス処理状態（たとえば、該オーディオ・データに対してどんな型（単数または複数）のラウドネス処理が実行されたか）および対応するオーディオ・データのラウドネス（たとえば、測定されたダイアログ・ラウドネス、ゲーティングされたおよび／またはゲーティングされないラウドネスおよび／またはダイナミックレンジ）を示す。ここで、ラウドネスおよび／またはオーディオ・データに対して実行されるレベル測定の「ゲーティング」とは、閾値を超える計算された値（単数または複数）が最終的な測定に含められる（たとえば、最終的な測定された値において - 60dBFSより低い短期的なラウドネス値を無視する）ような特定のレベルまたはラウドネスの閾値を参照する。絶対的な値に対するゲーティングは固定したレベルまたはラウドネスを参照し、相対値に対するゲーティングは現在の「ゲーティングされていない」測定値に依存する値を参照する。

10

【 0 0 5 0 】

エンコーダ 1 0 0 のいくつかの実装では、メモリ 1 0 9 にバッファリングされている（そして送達システム 1 5 0 に出力される）エンコードされたビットストリームは、AC-3ビットストリームまたはE-AC-3ビットストリームであり、オーディオ・データ・セグメント（たとえば、図 4 に示したフレームのAB0～AB5セグメント）およびメタデータ・セグメントを含む。ここで、オーディオ・データ・セグメントはオーディオ・データを示し、メタデータ・セグメントのうち少なくともいくつかのセグメントのそれぞれは、ラウドネス処理状態メタデータ（LPSM）を含む。段 1 0 7 はLPSMを次のフォーマットでビットストリーム中に挿入する。LPSMを含むメタデータ・セグメントのそれぞれは、ビットストリームのフレームのビットストリーム情報（「BSI」）セグメントの「addbsi」フィールドまたはビットストリームのフレームの末尾にある補助データ・フィールド（たとえば図 4 に示されるAUXセグメント）に含められる。

20

【 0 0 5 1 】

ビットストリームのフレームは、それぞれがLPSMを含む一つまたは二つのメタデータ・セグメントを含んでいてもよく、フレームが二つのメタデータ・セグメントを含む場合、一方がフレームのaddbsiフィールドに、他方がフレームのAUXフィールドに存在していてもよい。LPSMを含む各メタデータ・セグメントは、次のフォーマットをもつLPSMペイロード（またはコンテナ）・セグメントを含む：ヘッダ（たとえばLPSMペイロードの始まりを同定する同期語を含み、それに続いて少なくとも一つの識別情報値、たとえば下記の表 2 に示されるLPSMフォーマット・バージョン、長さ、期間（period）、カウントおよびサブストリーム関連付け値がくる）；ヘッダ後に、対応するオーディオ・データがダイアログを示すかダイアログを示さないか（たとえば、対応するオーディオ・データのどのチャンネルがダイアログを示すか）を示す少なくとも一つのダイアログ指示値（たとえば、表 2 のパラメータ「ダイアログ・チャンネル」）；対応するオーディオ・データがラウドネス規制の示されるセットに準拠しているかどうかを示す少なくとも一つのラウドネス規制準拠値（たとえば、表 2 のパラメータ「ラウドネス規制型」）；対応するオーディオ・データに対して実行されたラウドネス処理の少なくとも一つの型を示す少なくとも一つのラウドネス処理値（たとえば、表 2 のパラメータ「ダイアログ・ゲーテッド・ラウドネス補正フラグ」、「ラウドネス補正型」の一つまたは複数）；および対応するオーディオ・データに特徴的な少なくとも一つのラウドネス（たとえばピークまたは平均ラウドネス）を示す少なくとも一つのラウドネス値（たとえば、パラメータ「ITU相対ゲーテッド・ラウドネス」、「ITU発話ゲーテッド・ラウドネス」、「ITU（EBU3341）短時間3sラウドネス」および「真のピーク」の一つまたは複数）。

30

40

【 0 0 5 2 】

いくつかの実装では、当該ビットストリームのフレームの「addbsi」フィールドまたは補助データ・フィールドに段 1 0 7 によって挿入されるメタデータ・セグメントのそれぞれは次のフォーマットをもつ：コア・ヘッダ（たとえばメタデータ・セグメントの開始を

50

同定する同期語と、それに続く識別情報値、たとえば下記の表 1 に示されるコア要素バージョン、長さおよび期間 (period)、拡張要素カウントおよびサブストリーム関連付け値を含む) ; および コア・ヘッダ後に、ラウドネス処理状態メタデータまたは対応するオーディオ・データの少なくとも一方の解読、認証 (authentication) または有効確認 (validation) のうちの少なくとも一つのために有用な少なくとも一つの保護値 (たとえば、表 1 の HMAC ダイジェストおよびオーディオ・フィンガープリント値) ; および やはりコア・ヘッダ後に、当該メタデータ・セグメントが LPSM を含む場合、LPSM ペイロード識別情報 (「ID」) および LPSM ペイロード・サイズの値であって、後続のメタデータを LPSM ペイロードとして同定し、該 LPSM ペイロードのサイズを示すもの。

【 0 0 5 3 】

10

(たとえば上記のフォーマットをもつ) LPSM ペイロード (またはコンテナ) ・セグメントは、LPSM ペイロード ID および LPSM ペイロード・サイズの値に続く。

【 0 0 5 4 】

いくつかの実施形態では、フレームの補助データ・フィールド (または「addbsi」フィールド) 中の各メタデータ・セグメントは、三レベルの構造をもつ: 高レベル構造。これは、補助データ (または addbsi) フィールドがメタデータを含むかどうかを示すフラグと、どの型 (単数または複数) のメタデータが存在しているかを示す少なくとも一つの ID 値と、任意的にはまた (メタデータが存在する場合) (たとえば各型の) 何ビットのメタデータが存在するかを示す値とを含む。存在できるメタデータの一つの型は LPSM であり、存在できるメタデータのもう一つの型はメディア・リサーチ (research) ・メタデータ (たとえば、ニールセン・メディア・リサーチ (Nielsen Media Research) ・メタデータ) である; 中間レベル構造。これは、メタデータのそれぞれの同定される型についてのコア要素を含む (たとえば、メタデータのそれぞれの同定される型についての上述したようなコア・ヘッダ、保護値および LPSM ペイロード ID および LPSM ペイロード・サイズの値) ; 低レベル構造。これは、あるコア要素についての各ペイロード (たとえば、前記コア要素によって LPSM ペイロードが存在すると同定されている場合の LPSM ペイロード および / または前記コア要素によって別の型のメタデータ・ペイロードが存在すると同定されている場合の該別の型のメタデータ・ペイロード) 。

20

【 0 0 5 5 】

そのような三レベル構造におけるデータ値は、ネストされることができる。たとえば、コア要素によって同定される LPSM ペイロード および / または別のメタデータ・ペイロードについての保護値 (単数または複数) が、コア要素によって同定される各ペイロード後に (よって、コア要素のコア・ヘッダ後に) 含まれることができる。一例では、コア・ヘッダは、LPSM ペイロード および 別のメタデータ・ペイロードを同定することができ、第一のペイロード (たとえば LPSM ペイロード) についてのペイロード ID およびペイロード・サイズの値がコア・ヘッダに続くことができ、第一のペイロード自身が該 ID およびサイズの値に続くことができ、第二のペイロードについてのペイロード ID およびペイロード・サイズ値が第一のペイロードに続くことができ、第二のペイロード自身がこれらの ID およびサイズの値に続くことができ、両方のペイロードについての (またはコア要素値および両方のペイロードについての) 保護値が最後のペイロードに続くことができる。

30

40

【 0 0 5 6 】

いくつかの実施形態では、デコーダ 1 0 1 が、暗号的ハッシュをもつ本発明のある実施形態に従って生成されたオーディオ・ビットストリームを受領する場合、デコーダは、ビットストリームから決定されたデータ・ブロックからの該暗号的ハッシュをパースして取り出すよう構成されている。前記ブロックは、ラウドネス処理状態メタデータ (LPSM) を含む。有効確認器 1 0 2 は該暗号的ハッシュを使って、受領されたビットストリーム および / または関連付けられたメタデータを有効確認してもよい。たとえば、有効確認器 1 0 2 が、参照暗号的ハッシュと前記データ・ブロックから取り出された前記暗号的ハッシュとの間の一致に基づいて前記 LPSM が有効であると見出す場合、有効確認器 1 0 2 は、対応するオーディオ・データに対するプロセッサ 1 0 3 の動作を無効にしても

50

よく、選択段 104 にオーディオ・データを（変更なしに）素通りさせてもよい。追加的、任意的または代替的に、暗号的ハッシュに基づく方法の代わりに他の型の暗号技法が使用されてもよい。

【0057】

図2のエンコーダ100は、（デコーダ101によって抽出されたLPSMに応答して）後／前処理ユニットが、ある型のラウドネス処理を、（要素105、106および107において）エンコードされるべきオーディオ・データに対して実行したことを判別してもよく、よって前に実行されたラウドネス処理において使われたおよび／または前に実行されたラウドネス処理から導出された特定のパラメータを含むラウドネス処理状態メタデータを（生成器106において）生成してもよい。いくつかの実装では、エンコーダ100は、エンコーダがオーディオ・コンテンツに対して実行された処理の型を認識する限り、オーディオ・コンテンツに対する処理履歴を示す処理状態メタデータを生成して（そしてそれから出力されるエンコードされたビットストリームに含めて）もよい。

【0058】

図3は、図1のシステム10との関連で使用されうるデコーダのブロック図である。デコーダ200および後処理器300のコンポーネントまたは要素の任意のものは、ハードウェア、ソフトウェアまたはハードウェアとソフトウェアの組み合わせにおいて、一つまたは複数のプロセスおよび／または一つまたは複数の回路（たとえばASIC、FPGAまたは他の集積回路）として実装されうる。デコーダ200は、図のように接続された、フレーム・バッファ201、パーサ205、オーディオ・デコーダ202、オーディオ状態有効確認段（有効確認器）203および制御ビット生成段204を有する。デコーダ200は他の処理要素（図示せず）を含んでいてもよい。フレーム・バッファ201（バッファ・メモリ）は、デコーダ200によって受領されるエンコードされたオーディオ・ビットストリームの少なくとも一つのフレームを（たとえば非一時的な仕方）で記憶する。エンコードされたオーディオ・ビットストリームのフレームのシーケンスがバッファ201からパーサ205に呈される。パーサ205は、ラウドネス処理メタデータ（LPSM）および他のメタデータを、前記エンコードされた入力オーディオの各フレームから抽出し、少なくともLPSMをオーディオ状態有効確認器203および段204に呈し、LPSMを出力として（たとえば後処理器300に）呈し、エンコードされた入力オーディオからオーディオ・データを抽出し、抽出されたオーディオ・データをデコーダ202に呈するよう結合され、構成されている。デコーダ200に入力されるエンコードされたオーディオ・ビットストリームは、AC-3ビットストリーム、E-AC-3ビットストリームまたはドルビーEビットストリームのうちの一つであってもよい。

【0059】

図3のシステムは後処理器300をも含む。後処理器300は、フレーム・バッファ301と、バッファ301に結合された少なくとも一つの処理要素を含む他の処理要素（図示せず）とを有する。フレーム・バッファ301は、デコーダ200から後処理器300によって受領されるデコードされたオーディオ・ビットストリームの少なくとも一つのフレームを（たとえば非一時的な仕方）で記憶する。後処理器300の処理要素は、バッファ301から出力されるデコードされたオーディオ・ビットストリームのフレームのシーケンスを受領し、デコーダ202から出力される（LPSM値を含む）メタデータおよび／またはデコーダ200の段204から出力される制御ビットを使って適応的に処理するよう結合され、構成されている。ある実施形態では、後処理器300は、LPSM値を使って（たとえばLPSMによって示される、ラウドネス処理状態および／または一つまたは複数のオーディオ・データ特性に基づいて）デコードされたオーディオ・データに対して適応的なラウドネス処理を実行するよう構成されている。デコーダ200および後処理器300のさまざまな実装は、本稿に記載される実施形態に基づく方法の種々の実施形態を実行するよう構成されている。

【0060】

デコーダ200のオーディオ・デコーダ202は、パーサ205によって抽出されたオ

10

20

30

40

50

オーディオ・データをデコードして、デコードされたオーディオ・データを生成し、該デコードされたオーディオ・データを出力として（たとえば後処理器 300 に）呈するよう構成されている。状態有効確認器 203 は、それに対して呈される LPSM を（任意的には他のメタデータも）認証し、有効確認するよう構成されている。いくつかの実施形態では、LPSM は、（たとえば本発明のある実施形態に従って）入力ビットストリームに含められたデータ・ブロックである（または該データ・ブロックに含まれる）。該ブロックは、LPSM（および任意的には他のメタデータも）および/または基礎になるオーディオ・データ（パーサ 205 および/またはデコーダ 202 から有効確認器 203 に提供される）を処理するための暗号的ハッシュ（ハッシュ・ベースのメッセージ認証コードまたは「HMAC」）を含んでいてもよい。該データ・ブロックは、これらの実施形態において、デジタル署名されてもよい。それにより、下流のオーディオ処理ユニットは比較的容易に、該処理状態メタデータを認証および有効確認しうる。

10

【0061】

一つまたは複数の HMAC でない暗号学的方法の任意のものを含むがそれに限定されない他の暗号学的方法が、LPSM および/または基礎になるオーディオ・データの安全な送受信を保証するための（たとえば有効確認器 203 における）LPSM の有効確認のために使われてもよい。たとえば、（そのような暗号学的方法を使う）有効確認は、本発明のオーディオ・ビットストリームの実施形態を受領する各オーディオ処理ユニットにおいて実行され、ビットストリームに含まれるラウドネス処理状態メタデータおよび対応するオーディオ・データが（該メタデータによって示されるような）特定のラウドネス処理を受けている（および/または特定のラウドネス処理から帰結する）ものであり、そのような特定のラウドネス処理の実行後に修正されていないかどうかを判定することができる。

20

【0062】

状態有効確認器 203 は、有効確認動作の結果を示すために、ビット生成器 204 を制御する制御データを呈するおよび/または該制御データを出力として（たとえば後処理器 300 に）呈する。該制御データに（任意的には入力ビットストリームから抽出される他のメタデータにも）応答して、段 204 は次のいずれかを生成し（そして後処理器 300 に呈し）てもよい：

（たとえば、LPSM がデコーダ 202 から出力されたオーディオ・データが特定の型のラウドネス処理を受けていることを示し、有効確認器 203 からの制御ビットが LPSM が有効であることを示すとき）デコーダ 202 から出力されたデコードされたオーディオ・データが該特定の型のラウドネス処理を受けていることを示す制御ビット；または（たとえば、LPSM がデコーダ 202 から出力されたオーディオ・データが特定の型のラウドネス処理を受けていないことを示す、または LPSM がデコーダ 202 から出力されたオーディオ・データが特定の型のラウドネス処理を受けていることを示すが、有効確認器 203 からの制御ビットが LPSM が有効でないことを示すとき）デコーダ 202 から出力されたデコードされたオーディオ・データが該特定の型のラウドネス処理を受けるべきであることを示す制御ビット。

30

【0063】

あるいはまた、デコーダ 200 は、入力ビットストリームからデコーダ 202 によって抽出されたメタデータ（および他の任意のメタデータ）を後処理器 300 に呈し、後処理器 300 は LPSM を使って、デコードされたオーディオ・データに対してラウドネス処理を実行し、LPSM の有効確認を実行し、次いで有効確認が LPSM が有効であることを示す場合には、LPSM を使って、デコードされたオーディオ・データに対してラウドネス処理を実行する。

40

【0064】

いくつかの実施形態では、デコーダ 201 が、暗号的ハッシュをもつ本発明のある実施形態に従って生成されるオーディオ・ビットストリームを受領する場合、デコーダは、ビットストリームから決定されたデータ・ブロックからの該暗号的ハッシュをパースして取り出すよう構成されている。前記ブロックは、ラウドネス処理状態メタデータ（LPS

50

M)を含む。有効確認器203は該暗号的ハッシュを使って、受領されたビットストリームおよび/または関連付けられたメタデータを有効確認してもよい。たとえば、有効確認器203が、参照暗号的ハッシュと前記データ・ブロックから取り出された前記暗号的ハッシュとの間の一致に基づいて前記LPSMが有効であると見出す場合、有効確認器203は、下流のオーディオ処理ユニット(たとえば、ボリューム平準化ユニットであるまたはボリューム平準化ユニットを含んでいてもよい後処理器300)に対して、ビットストリームの該オーディオ・データを(変更なしに)素通りさせるよう信号伝達してもよい。追加的、任意的または代替的に、暗号的ハッシュに基づく方法の代わりに他の型の暗号技法が使用されてもよい。

【0065】

デコーダ100のいくつかの実装では、受領される(そしてメモリ201にバッファリングされる)エンコードされたビットストリームはAC-3ビットストリームまたはE-AC-3ビットストリームであり、オーディオ・データ・セグメント(たとえば図4に示されるフレームのAB0~AB5セグメント)およびメタデータ・セグメントを含む。ここで、オーディオ・データ・セグメントはオーディオ・データを示し、メタデータ・セグメントの少なくともいくつかの各セグメントはラウドネス処理状態メタデータ(LPSM)を含む。デコーダ段202は、ビットストリームから、以下のフォーマットをもつLPSMを抽出するよう構成されている。LPSMを含むメタデータ・セグメントのそれぞれは、ビットストリームのフレームのビットストリーム情報(「BSI」)セグメントの「addbsi」フィールド中に、あるいはビットストリームのフレームの末尾の補助データ・フィールド(たとえば図4に示されるAUXセグメント)中に含まれる。ビットストリームのフレームは、それぞれLPSMを含む一つまたは二つのメタデータ・セグメントを含んでいてもよく、フレームが二つのメタデータ・セグメントを含む場合、一方がフレームのaddbsiフィールドに存在し、他方がフレームのAUXフィールドに存在していてもよい。LPSMを含む各メタデータ・セグメントは、以下のフォーマットをもつLPSMペイロード(またはコンテナ)セグメントを含む:ヘッダ(たとえばLPSMペイロードの始まりを同定する同期語を含み、それに続いて識別情報値、たとえば、下記の表2に示される、LPSMフォーマット・バージョン、長さ、期間(period)、カウントおよびサブストリーム関連付け値を含む);ヘッダ後に、対応するオーディオ・データがダイアログを示すかダイアログを示さないか(たとえば、対応するオーディオ・データのどのチャンネルがダイアログを示すか)を示す少なくとも一つのダイアログ指示値(たとえば表2のパラメータ「ダイアログ・チャンネル」);対応するオーディオ・データがラウドネス規制の示されるセットに準拠しているかどうかを示す少なくとも一つのラウドネス規制準拠値(たとえば表2のパラメータ「ラウドネス規制型」);対応するオーディオ・データに対して実行されたラウドネス処理の少なくとも一つの型を示す少なくとも一つのラウドネス処理値(たとえば、表2のパラメータ「ダイアログ・ゲートッド・ラウドネス補正フラグ」「ラウドネス補正型」の一つまたは複数);および対応するオーディオ・データに特徴的な少なくとも一つのラウドネス(たとえばピークまたは平均ラウドネス)を示す少なくとも一つのラウドネス値(たとえば、表2のパラメータ「ITU相対ゲートッド・ラウドネス」「ITU発話ゲートッド・ラウドネス」「ITU(EBU3341)短時間3sラウドネス」および「真のピーク」の一つまたは複数)。

【0066】

いくつかの実装では、デコーダ段202は、ビットストリームのフレームの「addbsi」フィールドまたは補助データ・フィールドから、それぞれ以下のフォーマットをもつメタデータ・セグメントを抽出するよう構成されている:ヘッダ(たとえばメタデータ・セグメントの始まりを同定する同期語と、それに続く少なくとも一つの識別情報値、たとえば下記の表1に示されるコア要素バージョン、長さおよび期間(period)、拡張要素カウントおよびサブストリーム関連付け値を含む);およびコア・ヘッダ後に、ラウドネス処理状態メタデータまたは対応するオーディオ・データの少なくとも一方の解読、認証(authentication)または有効確認(validation)のうちの少なくとも一つのために有用な少

10

20

30

40

50

なくとも一つの保護値（たとえば、表 1 の HMAC ダイジェストおよびオーディオ・フィンガープリント値）；および やはりコア・ヘッダ後に、当該メタデータ・セグメントが LPSM を含む場合、LPSM ペイロード識別情報（「ID」）および LPSM ペイロード・サイズの値であって、後続のメタデータを LPSM ペイロードとして同定し、該 LPSM ペイロードのサイズを示すもの。（たとえば上記で指定したフォーマットをもつ）LPSM ペイロード（またはコンテナ）・セグメントは、LPSM ペイロード ID および LPSM ペイロード・サイズの値に続く。

【 0 0 6 7 】

より一般には、ある実施形態によって生成されるエンコードされたオーディオ・ビットストリームは、メタデータ要素およびサブ要素に、コア（必須）または拡張（任意的な要素）としてラベル付けする機構を提供する構造をもつ。これは、（メタデータも含めた）ビットストリームのデータ・レートを、多数の用途を横断してスケーリングすることを許容する。ビットストリーム・シンタックスのコア（必須）要素は、オーディオ・コンテンツに関連付けられた拡張（任意的）要素が（帯域内に（in-band））および/またはリモート位置に（帯域外に（out of band））存在することを信号伝達することもできるべきである。

【 0 0 6 8 】

いくつかの実施形態では、コア要素（単数または複数）は、ビットストリームの全フレームに存在することが要求される。コア要素のいくつかのサブ要素は任意的であり、任意の組み合わせにおいて存在していてもよい。拡張要素は全フレームに存在することは要求されない（ビットレート・オーバーヘッドを限定的にするため）。このように、拡張要素は、いくつかのフレームに存在していて、他のフレームには存在しなくてもよい。拡張要素のいくつかのサブ要素は任意的であり、任意の組み合わせにおいて存在していてもよいが、拡張要素のいくつかのサブ要素は必須であってもよい（つまり、その拡張要素がビットストリームのフレームに存在するならば必須）。

【 0 0 6 9 】

いくつかの実施形態では、オーディオ・データ・セグメントおよびメタデータ・セグメントのシーケンスを含むエンコードされたオーディオ・ビットストリームが（たとえば、本発明を具現するオーディオ処理ユニットによって）生成される。オーディオ・データ・セグメントはオーディオ・データを示し、メタデータ・セグメントのうち少なくともいくつかのセグメントのそれぞれは、ラウドネス処理状態メタデータ（LPSM）を含み、オーディオ・データ・セグメントはメタデータ・セグメントと時分割多重される。このクラスのいくつかの実施形態では、メタデータ・セグメントのそれぞれは、本稿に記載されるフォーマットをもつ。あるフォーマットでは、エンコードされたビットストリームは AC-3 ビットストリームまたは E-AC-3 ビットストリームであり、LPSM を含むメタデータ・セグメントのそれぞれは、追加的なビットストリーム情報として、ビットストリームのフレームのビットストリーム情報（「BSI」）セグメントの「addbsi」フィールド（図 6 に示される）に、またはビットストリームのフレームの補助データ・フィールドに（たとえばエンコーダ 100 の段 107 によって）含められる。各フレームは、図 8 の表 1 に示されるフォーマットをもつコア要素を、フレームの addbsi フィールドに含む。

【 0 0 7 0 】

あるフォーマットでは、addbsi（または補助データ）フィールドのうち、LPSM を含むそれぞれは、コア・ヘッダ（および任意的には追加的なコア要素）と、コア・ヘッダのあとの（またはコア・ヘッダおよび他のコア要素のあとの）次の LPSM 値（パラメータ）とを含む：ペイロード ID（該メタデータを LPSM として同定する）。これは（たとえば表 1 において指定されるような）コア要素値に続く； ペイロード・サイズ（LPSM ペイロードの大きさを示す）。これはペイロード ID に続く； LPSM データ（ペイロード ID およびペイロード・サイズ値に続く）。これは図 9 の表 2 に示されるフォーマットをもつ。

【 0 0 7 1 】

エンコードされたビットストリームの第二のフォーマットでは、ビットストリームは AC

10

20

30

40

50

-3ビットストリームまたはE-AC-3ビットストリームであり、メタデータ・セグメントのうちLPSMを含むそれぞれは：ビットストリームのフレームのビットストリーム情報（「BSI」）セグメントの「addbsi」フィールド（図6に示した）；またはビットストリームのフレームの末尾の補助データ・フィールド（たとえば図4に示されるAUXフィールド）のいずれかに（たとえばエンコーダ100の段107によって）含まれる。フレームは、それぞれがLPSMを含む一つまたは二つのメタデータ・セグメントを含んでいてもよく、フレームが二つのメタデータ・セグメントを含む場合、一方がフレームのaddbsiフィールドに存在し、他方がフレームのAUXフィールドに存在してもよい。LPSMを含む各メタデータ・セグメントは、上記の表1および表2を参照して上記で規定したフォーマットをもつ（すなわち、表1に指定されるコア要素を含み、それに続いて、上記で規定したペイロードID（当該メタデータをLPSMとして同定する）およびペイロード・サイズ値がきて、それにペイロード（表2に示されるフォーマットをもつLPSMデータ）が続く）。

10

【0072】

もう一つのフォーマットでは、エンコードされたビットストリームはドルビーEビットストリームであり、メタデータ・セグメントのうちLPSMを含むそれぞれは、ドルビーE保護帯域区間の最初のN個のサンプル位置である。LPSMを含むそのようなメタデータ・セグメントを含むドルビーEビットストリームは、たとえば、SMPTE 337MプリアンプルのPd語において信号伝達されるLPSMペイロード長を示す値を含む（SMPTE 337M Pa語反復レートは関連するビデオ・フレーム・レートと同じままであってもよい）。

【0073】

20

エンコードされたビットストリームがE-AC-3ビットストリームであるフォーマットでは、メタデータ・セグメントのうちLPSMを含むそれぞれは、ビットストリームのフレームのビットストリーム情報（「BSI」）セグメントの「addbsi」フィールドにおいて、追加的なビットストリーム情報として（たとえば、エンコーダ100の段107によって）含まれる。このフォーマットにおけるLPSMをもつE-AC-3ビットストリームのエンコードのさらなる諸側面について次の通り述べる。（1）E-AC-3ビットストリームの生成中において、（LPSM値をビットストリーム中に挿入する）E-AC-3エンコーダが「アクティブである」間は、ビットストリームは、生成されるすべてのフレーム（同期フレーム）について、フレームのaddbsiフィールドにおいて担持される（LPSMを含む）メタデータ・ブロックを含むべきである。該メタデータ・ブロックを担持するために必要とされるビットは、エンコーダ・ビットレート（フレーム長）を増大させるべきではない。（2）（LPSMを含む）すべてのメタデータ・ブロックは、以下の情報を含むべきである： loudness_correction_type_flag〔ラウドネス補正型フラグ〕：ここで、「1」は対応するオーディオ・データのラウドネスが当該エンコーダより上流で補正されたことを示し、「0」は該ラウドネスが当該エンコーダに組み込まれているラウドネス補正器（たとえば、図2のエンコーダ100のラウドネス処理器103）によって補正されたことを示す； speech_channel〔発話チャネル〕：どの源チャネル（単数または複数）が（それまでの0.5秒の間に）発話を含むかを示す。発話が検出されない場合、その旨が示される； speech_loudness〔発話ラウドネス〕：発話を含む各対応するオーディオ・チャネルの（それまでの0.5秒の間の）統合された発話ラウドネスを示す； ITU_loudness〔ITUラウドネス〕：各対応するオーディオ・チャネルの統合されたITU BS.1770-2ラウドネスを示す； 利得：（可逆性を実証するため）デコーダにおいて反転するためのラウドネス複合利得（単数または複数）。

30

40

【0074】

（LPSM値をビットストリーム中に挿入する）E-AC-3エンコーダが「アクティブ」であり、「信頼」フラグをもつAC-3フレームを受領している間は、当該エンコーダにおけるラウドネス・コントローラ（たとえば図2のエンコーダ100のラウドネス処理器103）はバイパスされる。「信頼される」源dialnorm〔ダイアログ正規化〕およびDRC値は（たとえばエンコーダ100の生成器106によって）E-AC-3エンコーダ・コンポーネント（たとえばエンコーダ100の段107）に渡されるべきである。LPSMブロック生成は

50

継続し、loudness_correction_type_flagは「1」に設定される。ラウドネス・コントローラ・バイパス・シーケンスは、「信頼」フラグが現われるデコードされたAC-3フレームの先頭に同期される必要がある。ラウドネス・コントローラ・バイパス・シーケンスは次のように実装される：leveler_amount〔平準化器量〕コントロールが、10オーディオ・ブロック期間（すなわち、53.3msec）にわたって値9から値0にデクリメントされ、leveler_back_end_meter〔平準化器バック・エンド・メーター〕コントロールがバイパス・モードにされる（この動作は、シームレスな遷移を与えるべきである）。平準化器の「信頼される」バイパスという用語は、源ビットストリームのdialnorm値が、エンコーダの出力においても再利用されることを含意する（たとえば、「信頼される」源ビットストリームが-30のdialnorm値をもつ場合、エンコーダの出力は出て行くdialnorm値について-30を利用するべきである）。

10

【0075】

（LPSM値をビットストリーム中に挿入する）E-AC-3エンコーダが「アクティブ」であり、「信頼」フラグなしのAC-3フレームを受領している間は、当該エンコーダに組み込まれたラウドネス・コントローラ（たとえば図2のエンコーダ100のラウドネス処理器103）はアクティブである。LPSMブロック生成は継続し、loudness_correction_type_flagは「0」に設定される。ラウドネス・コントローラ・アクティブ化シーケンスは、「信頼」フラグが消失するデコードされたAC-3フレームの先頭に同期される。ラウドネス・コントローラ・アクティブ化シーケンスは次のように実装される：leveler_amount〔平準化器量〕コントロールが、1オーディオ・ブロック期間（すなわち、5.3msec）にわたって値0から値9にインクリメントされ、leveler_back_end_meter〔平準化器バック・エンド・メーター〕コントロールが「アクティブ」モードにされる（この動作は、シームレスな遷移を与え、back_end_meter統合リセットを含む）。エンコード中、グラフィカル・ユーザー・インターフェース（GUI）はユーザーに対して以下のパラメータを示した：「入力オーディオ・プログラム〔信頼される／信頼されない〕」-このパラメータの状態は入力信号内の「信頼」フラグの存在に基づく；および「リアルタイム・ラウドネス補正：〔有効化／無効化〕」-このパラメータの状態は、エンコーダに組み込まれているこのラウドネス・コントローラがアクティブであるかどうかに基づく。

20

【0076】

ビットストリームの各フレームのビットストリーム情報（「BSI」）セグメントの「addbsi」フィールドに含まれる（記載されるフォーマットにおける）LPSMを有するAC-3またはE-AC-3ビットストリームをデコードするとき、デコーダは、（addbsiフィールド中の）LPSMブロック・データをパースして、抽出されたLPSM値をグラフィカル・ユーザー・インターフェース（GUI）に渡す。抽出されたLPSM値の組は、フレーム毎にリフレッシュされる。

30

【0077】

さらにもう一つのフォーマットでは、エンコードされたビットストリームはAC-3ビットストリームまたはE-AC-3ビットストリームであり、メタデータ・セグメントのうちLPSMを含むそれぞれは、（たとえばエンコーダ100の段107によって）ビットストリームのフレームのビットストリーム情報（「BSI」）セグメントの「addbsi」フィールド（図6に示した）における（またはAuxセグメントにおける）追加的なビットストリーム情報として、含められる。（表1および表2を参照して上述したフォーマットに対する変形である）このフォーマットでは、addbsi（またはAux）フィールドのうちLPSMを含むそれぞれは、以下のLPSM値を含む：表1に規定されるコア要素。それに続いてペイロードID（当該メタデータをLPSMとして同定する）およびペイロード・サイズ値、それに続いて次のフォーマット（上記の表2に示した要素と同様）をもつペイロード（LPSMデータ）：LPSMペイロードのバージョン：LPSMペイロードのバージョンを示す2ビット・フィールド；dialchan：対応するオーディオ・データの左、右および／または中央チャンネルが話されたダイアログを含んでいるかどうかを示す3ビット・フィールド。dialchanフィールドのビット割り当ては次のとおりであってもよい：左チャンネルにおけるダイアログの存

40

50

在を示すビット0はdialchanフィールドの最上位ビットに格納され、中央チャンネルにおけるダイアログの存在を示すビット2はdialchanフィールドの最下位ビットに格納される。対応するチャンネルがプログラムの先行する0.5秒の間に話されるダイアログを含んでいる場合には、dialchanフィールドの各ビットが「1」に設定される； loudregtyp：プログラム・ラウドネスがどの規制規格に準拠しているかを示す3ビット・フィールド。「loudregtyp」フィールドを「000」に設定することは、LPSMがラウドネス規制準拠を示さないことを示す。たとえば、このフィールドのある値（たとえば000）は、ラウドネス規制規格への準拠が示されないことを示してもよく、このフィールドの別の値（たとえば001）は当該プログラムのオーディオ・データがATSC A/85規格に準拠していることを示してもよく、このフィールドの別の値（たとえば010）は当該プログラムのオーディオ・データがEBU R128規格に準拠していることを示してもよい。この例において、このフィールドが「000」以外の何らかの値に設定される場合、loudcorrdialgatおよびloudcorrtpフィールドがペイロードのあとに続くべきである； loudcorrdialgat：ダイアログでゲーティングされたラウドネス補正が適用されたかどうかを示す1ビット・フィールド。プログラムのラウドネスがダイアログ・ゲーティングを使って補正されている場合には、loudcorrdialgatフィールドの値は「1」に設定される。そうでない場合にはその値は「0」に設定される； loudcorrtp：プログラムに適用されたラウドネス補正の型を示す1ビット・フィールド。プログラムのラウドネスが無限先読み（ファイル・ベース）のラウドネス補正プロセスで補正されている場合には、loudcorrtpフィールドの値は「0」に設定される。プログラムのラウドネスがリアルタイム・ラウドネス測定およびダイナミックレンジ制御の組み合わせを使って補正されている場合には、このフィールドの値は「1」に設定される； loudrelgate：相対ゲーティングされたラウドネス・データ（ITU）が存在するかどうかを示す1ビット・フィールド。loudrelgateフィールドが「1」に設定される場合、ペイロードにおいて、7ビットのituloudrelgatフィールドが後続するべきである； loudrelgat：相対ゲーティングされたプログラム・ラウドネス（ITU）を示す7ビット・フィールド。このフィールドは、dialnormおよびダイナミックレンジ圧縮に起因するいかなる利得調整も適用されることなく、ITU-R BS.1770-2に従って測定された、オーディオ・プログラムの統合されたラウドネスを示す。0ないし127の値は、0.5LKFSきざみで、-58LKFSから+5.5LKFSとして解釈される； loudspchgate：発話でゲーティングされたラウドネス・データ（ITU）が存在するかどうかを示す1ビット・フィールド。loudspchgateフィールドが「1」に設定される場合、ペイロードにおいて、7ビットのloudspchgatフィールドが後続するべきである； loudspchgat：発話ゲーティングされたプログラム・ラウドネスを示す7ビット・フィールド。このフィールドは、dialnormおよびダイナミックレンジ圧縮に起因するいかなる利得調整も適用されることなく、ITU-R BS.1770-3の公式(2)に従って測定された、対応するオーディオ・プログラム全体の統合されたラウドネスを示す。0ないし127の値は、0.5LKFSきざみで、-58から+5.5LKFSとして解釈される； loudstrm3se：短時間（3秒）ラウドネス・データが存在するかどうかを示す1ビット・フィールド。このフィールドが「1」に設定される場合、ペイロードにおいて7ビットのloudstrm3sフィールドが後続するべきである。 loudstrm3s：dialnormおよびダイナミックレンジ圧縮に起因するいかなる利得調整も適用されることなく、ITU-R BS.1771-1に従って測定された、対応するオーディオ・プログラムの先行する3秒のゲーティングされていないラウドネスを示す7ビット・フィールド。0ないし256の値は、0.5LKFSきざみで、-116LKFSから+11.5LKFSとして解釈される； truepke：真のピーク・ラウドネス・データが存在するかどうかを示す、1ビット・フィールド。truepkeフィールドが「1」に設定されていたら、ペイロードにおいて8ビットのtruepkフィールドが後続するべきである； truepk：dialnormおよびダイナミックレンジ圧縮に起因するいかなる利得調整も適用されることなく、ITU-R BS.1770-3の付属書2に従って測定された、プログラムの真のピーク・サンプル値を示す8ビット・フィールド。0ないし256の値は、0.5LKFSきざみで、-116LKFSから+11.5LKFSとして解釈される。

【0078】

10

20

30

40

50

いくつかの実施形態では、AC-3ビットストリームまたはE-AC-3ビットストリームのフレームの補助データ・フィールド（または「adbsi」フィールド）におけるメタデータ・セグメントのコア要素は、コア・ヘッダ（任意的に識別情報値、たとえばコア要素バージョンを含む）と、該コア・ヘッダ後に：メタデータ・セグメントのメタデータについてフィンガープリント・データが（または他の保護値が）含まれるかどうかを示す値と、（当該メタデータ・セグメントのメタデータに対応するオーディオ・データに係する）外部データが存在するかどうかを示す値と、コア要素によって同定される各型のメタデータ（たとえばLPSMおよび／またはLPSM以外の型のメタデータ）についてのペイロードIDおよびペイロード・サイズの値と、コア要素によって同定されるメタデータの少なくとも一つの型についての保護値とを含む。メタデータ・セグメントのメタデータ・ペイロード（単数または複数）は、コア・ヘッダに続き、（場合によっては）コア要素の値内にネストされる。

10

【0079】

最適化されたラウドネスおよびダイナミックレンジ・システム

上記の安全なメタデータ符号化および転送方式は、図1に示されるような、異なる再生装置、アプリケーションおよび聴取環境を横断してラウドネスおよびダイナミックレンジを最適化するためのスケーラブルかつ拡張可能なシステムとの関連で使われる。ある実施形態では、システム10は、異なる目標ラウドネス値を必要とし、異なるダイナミックレンジ機能をもつさまざまな装置を横断して入力オーディオ11のラウドネス・レベルおよびダイナミックレンジを正規化するように構成されている。ラウドネス・レベルおよびダイナミックレンジを正規化するために、システム10は、オーディオ・コンテンツとともに種々の装置プロファイルを含み、正規化はそれらのプロファイルに基づいてなされる。それらのプロファイルは、オーディオ処理チェーン内のオーディオ処理ユニットの一つによって含まれていてもよく、含まれるプロファイルは、オーディオ処理チェーン内の下流の処理ユニットによって、目標装置のための所望される目標ラウドネスおよびダイナミックレンジを決定するために使用されてもよい。追加的な処理コンポーネントが、（以下のパラメータ：ヌル帯域レンジ、真のピーク閾値、ラウドネス・レンジ、高速／低速時定数（係数）および最大ブーストを含む（がそれに限定されない））装置プロファイル管理、利得制御および広帯域および／またはマルチ帯域利得生成機能のための情報を提供または処理してもよい。

20

30

【0080】

図10は、いくつかの実施形態のもとでの、最適化されたラウドネスおよびダイナミックレンジ制御を提供するシステムのための、図1のシステムのより詳細な図を示している。図10のシステム321について、エンコーダ段は、オーディオ入力303をデコーダ312への送信のために好適なデジタル・フォーマットでエンコードするコア・エンコーダ・コンポーネント304を有している。オーディオは、それぞれ異なるラウドネスおよび／またはダイナミックレンジ目標設定を要求することがありうる多様な異なる聴取環境において再生されることができるよう処理される。こうして、図10に示されるように、デコーダはデジタル信号を出力し、該デジタル信号は、フル・レンジ・スピーカー320、ミニチュア・スピーカー322およびヘッドフォン324を含む多様な異なるドライバ型を通じた再生のために、デジタル・アナログ変換器316によってアナログ・フォーマットに変換される。これらのドライバは、可能な再生ドライバのほんのいくつかの例を示しており、任意の適切なサイズの任意のトランスデューサまたはドライバが使用される。さらに、図10のドライバ／トランスデューサ320～324は、任意の対応する聴取環境における使用のために任意の適切な再生装置において具現されてもよい。装置型はたとえば、AVR、テレビジョン、ステレオ設備、コンピュータ、携帯電話、タブレット・コンピュータ、MP3プレーヤーなどを含んでいてもよく、聴取環境はたとえば講堂、家、自動車、聴取ブースなどを含んでいてもよい。

40

【0081】

再生環境およびドライバ型の範囲が非常に小さな個人的なコンテキストから非常に大き

50

な公共会場まで多様でありうるので、可能および最適な再生ラウドネスおよびダイナミックレンジ構成のスペクトルは、コンテンツ型、背景雑音レベルなどに依存して大きく変わりうる。たとえば、家庭シアター環境では、サラウンドサウンド設備を通じて広いダイナミックレンジのコンテンツが再生されることができ、通常のテレビジョン・システム（フラットパネルLED/LCD型など）を通じてはより狭いダイナミックレンジのコンテンツが再生されることができ、一方、大きなレベル変動が望まれないある種の聴取条件のためには（たとえば夜間には、または厳しい音響出力パワー制限のある装置、たとえば携帯電話ノタブレットの内蔵スピーカーまたはヘッドフォン出力では）非常に狭いダイナミックレンジ・モードが使用されることがある。小型コンピュータまたはドック・スピーカーまたはヘッドフォンノイヤパッドを使うといったポータブルまたはモバイルな聴取コンテキストにおいては、再生の最適なダイナミックレンジは、環境に依存して変わりうる。たとえば、静かな環境では、最適なダイナミックレンジは騒がしい環境に比べてより大きいことがありうる。図10の適応的なオーディオ処理システムの諸実施形態は、聴取装置環境および再生装置型といったパラメータに依存してオーディオ・コンテンツをより了解可能にするために、ダイナミックレンジを変える。

【0082】

図11は、ある例示的な使用事例における多様な再生装置および背景聴取環境のための種々のダイナミックレンジ要求を示す表である。同様の要求がラウドネスについて導出されることができ、これら種々のダイナミックレンジおよびラウドネス要求は、最適化システム321によって使われる種々のプロファイルを生成する。システム321は、入力オーディオのラウドネスおよびダイナミックレンジを解析し、測定するラウドネスおよびダイナミックレンジ測定コンポーネント302を含む。ある実施形態では、システムは、システムは全体的なプログラム・コンテンツを解析して全体的なラウドネス・パラメータを決定する。このコンテキストにおいて、ラウドネスは、長期プログラム・ラウドネスまたはプログラムの平均ラウドネスを指す。ここで、プログラムは、映画、テレビ番組、コマーシャルまたは同様のプログラム・コンテンツといったオーディオ・コンテンツの単一の単位である。ラウドネスは、オーディオがどのように再生されるかを制御するためにコンテンツ・クリエイターによって使われる芸術的ダイナミックレンジ・プロファイルの指標を提供するために使われる。dialnormは単一のプログラム（たとえば映画、番組、コマーシャルなど）の平均ダイアログ・ラウドネスを表わすという点で、ラウドネスはdialnormメタデータ値に関係している。短期ダイナミックレンジは、プログラム・ラウドネスよりずっと短い時間期間にわたる信号の変動を定量化する。たとえば、短期ダイナミックレンジは数秒のオーダーで測定されてもよく、一方、プログラム・ラウドネスは数分またはさらには数時間のスペクトルにわたって測定されてもよい。短期ダイナミックレンジは、さまざまな再生プロファイルおよび装置型について過負荷が起こらないことを保証するための、プログラム・ラウドネスとは独立した保護機構を提供する。ある実施形態では、ラウドネス（長期プログラム・ラウドネス）目標はダイアログ・ラウドネスに基づき、短期ダイナミックレンジは相対ゲーティングされたおよびノまたはゲーティングされていないラウドネスに基づく。この場合、システム内のある種のDRCおよびラウドネス・コンポーネントはコンテンツ型およびノまたは目標装置型および特性に関して、コンテキストを意識する。このコンテキストを意識した機能の一部として、システムは、出力装置の一つまたは複数の特性を、該装置が、AVR型装置、テレビジョン、コンピュータ、ポータブル装置などといったある種のDRCおよびラウドネス再生条件のために最適化されている装置の特定の群の一員であるかどうかを判定するために解析するよう構成されている。

【0083】

前処理コンポーネントは、複数の異なるプロファイルの各プロファイルについて一意的なメタデータを生成するために、プログラム・コンテンツを解析して、ラウドネス、ピーク、真のピークおよび静穏期間を決定する。ある実施形態では、ラウドネスはダイアログ・ゲーテッド・ラウドネスおよびノまたは相対ゲーテッド・ラウドネスであってもよい。前記異なるプロファイルは、さまざまなDRC（Dynamic Range Control [ダイナミッ

10

20

30

40

50

クレンジ制御〕)および目標ラウドネス・モードを定義する。それらのモードでは、源オーディオ・コンテンツ、所望される目標ラウドネスおよび再生装置型および/または環境の特性に依存してエンコーダにおいて異なる利得値が生成される。デコーダは、(上記で言及したプロファイルによって可能にされる)種々のDRCおよび目標ラウドネス・モードをオファ―してもよく、以下を含みうる。DRCおよび目標ラウドネス・オフ/無効は、オーディオ信号の圧縮なしかつラウドネス正規化なしでフルのダイナミックレンジ・リスティングを許容する。DRCオフ/無効および家庭シアター・システム上での再生のための-31LKFSライン・モードの目標をもつラウドネス正規化は、-31LKFSの目標をもつラウドネス正規化をもつエンコーダにおいて(この再生モードおよび/または装置プロファイルのために特に)生成される利得値を通じて中程度のダイナミックレンジ圧縮を提供する。TVスピーカ―を通じた再生のためのRFモードは、-24、-23または-20LKFSの目標をもつラウドネス正規化による重度のダイナミックレンジ圧縮を提供する。コンピュータまたは同様の装置上での再生のための中間モードは、-14LKFSの目標でのラウドネス正規化により圧縮を提供する。ポータブル・モードは、-11LKFSのラウドネス正規化目標により非常に重度のダイナミックレンジ圧縮を提供する。-31、-23/-20、-14および-11のLKFSの目標ラウドネス値は、いくつかの実施形態のもとで本システムについて定義されうる異なる再生/装置プロファイルの例であることが意図されており、他のいかなる適切な目標ラウドネス値が使われてもよく、本システムは、特にこれらの再生モードおよび/または装置プロファイルのために適切な利得値を生成する。さらに、本システムは、エンコーダにおいて新しいプロファイルを定義するまたはエンコーダにロードされる新しいプロファイルを他のどこかで定義することにより、種々の再生装置および聴取環境に対応することができるよう、拡張可能かつ適応可能である。このようにして、将来の応用のために改善されたまたは異なる再生装置をサポートするために新しい、一意的な再生/装置プロファイルが生成されることができる。

【0084】

ある実施形態では、利得値は、システム321の任意の適切な処理コンポーネントにおいて、たとえばエンコーダ304、デコーダ312またはトランスコーダ308またはエンコーダに関連した任意の関連前処理コンポーネントまたはデコーダに関連した任意の後処理コンポーネントにおいて、計算されることができる。

【0085】

図13は、ある実施形態のもとでの、多様な異なる再生装置クラスについての異なるプロファイルの間のインターフェースを示すブロック図である。図13に示されるように、エンコーダ502はオーディオ入力501およびいくつかの異なるプロファイル506のうちの一つを受領する。エンコーダは、オーディオ・データを選択されたプロファイルと組み合わせて、出力ビットストリーム・ファイルを生成する。出力ビットストリーム・ファイルは、目標再生装置内のまたは目標再生装置に関連するデコーダ・コンポーネントにおいて処理されるものである。図13の例については、前記異なる再生装置は、コンピュータ510、携帯電話512、AVR 514およびテレビジョン516であってもよいが、他の多くの出力装置も可能である。装置510~516のそれぞれは、ドライバ320~324のようなスピーカ―(ドライバおよび/またはトランスデューサを含む)を含むまたはそれに結合されている。再生装置および関連するスピーカ―の処理、電力定格およびサイズの組み合わせが一般に、その特定の目標のためにどのプロファイルが最も最適であるかを指定する。このように、プロファイル506は、AVR、TV、モバイル・スピーカ―、モバイル・ヘッドフォンなどを通じた再生のために特に定義されてもよい。プロファイルは、静穏モード、夜間モード、戸外、屋内などといった特定の動作モードまたは条件のためにも定義されてもよい。図13に示されるプロファイルは、単に例示的なモードであり、特定の目標および環境のためのカスタム・プロファイルを含め、いかなる適切なプロファイルが定義されてもよい。

【0086】

図13は、エンコーダ502がプロファイル506を受領してラウドネスおよびDRC処

10

20

30

40

50

理のための適切なパラメータを生成する実施形態を示しているが、プロファイルおよびオーディオ・コンテンツに基づいて生成されたパラメータが、エンコーダ、デコーダ、トランスコーダ、前処理器、後処理器などといった任意の適切なオーディオ処理ユニット上で実行されることができることを注意しておくべきである。たとえば、図 13 の各出力装置 510 ~ 516 は、目標出力装置の装置または装置型にマッチするようラウドネスおよびダイナミックレンジの適応を可能にするためにエンコーダ 502 から送られたファイル 504 内のビットストリーム内のメタデータを処理するデコーダ・コンポーネントを有するまたはそれに結合されている。

【0087】

ある実施形態では、オーディオ・コンテンツのダイナミックレンジおよびラウドネスは、各可能な再生装置について最適化されている。これは、目標再生モードのそれぞれについて（信号ダイナミクス、サンプル・ピークおよび/または真のピークを制御することにより）オーディオ経験を最適化するように、長期ラウドネスを目標に維持し、短期ダイナミックレンジを制御することによって達成される。長期ラウドネスおよび短期ダイナミックレンジについて種々のメタデータ要素が定義される。図 10 に示されるように、コンポーネント 302 は入力オーディオ信号全体を（またはその一部、たとえばもし該当すれば発話成分を）解析して、これらの別個のDR成分の両方についての有意な特性を導出する。これは、芸術的な利得とクリップ（過負荷保護）利得値について、異なる利得値が定義されることを許容する。

【0088】

長期ラウドネスおよび短期ダイナミックレンジについてのこれらの利得値は、次いで、プロファイル 305 にマッピングされて、該ラウドネスおよびダイナミックレンジ制御利得値を記述するパラメータを提供する。これらのパラメータは、マルチプレクサ 306 または同様のコンポーネントにおいて、エンコーダ 304 からエンコードされたオーディオ信号と組み合わせられる。トランスコーダ 308 を通じてデコーダ段に伝送されるビットストリームの生成のためである。デコーダ段に入力されるビットストリームはデマルチプレクサ 310 において多重分離され、次いでデコーダ 312 においてデコードされる。利得コンポーネント 314 は、適切なプロファイルに対応する利得を適用して、デジタル・オーディオ・データを生成する。該デジタル・オーディオ・データは次いで、適切な再生装置およびドライバまたはトランスデューサ 320 ~ 324 を通じた再生のために、DAC Sユニット 416 を通じて処理される。

【0089】

図 14 は、ある実施形態のもとでの、複数の定義されたプロファイルについての長期ラウドネスと短期ダイナミックレンジの間の相関を示す表である。図 14 の表 4 に示されるように、各プロファイルは、システムのデコーダまたは各目標装置において適用されるダイナミックレンジ圧縮（DRC: dynamic range compression）の量を指定する利得値の組を含んでいる。プロファイル 1 ~ N と記される N 個のプロファイルのそれぞれは、デコーダ段において適用される対応する利得値を指定することによって、特定の長期ラウドネス・パラメータ（たとえば dialnorm）および過負荷（overload）圧縮パラメータを設定する。これらのプロファイルについてのDRC利得値は、エンコーダによって受け入れられる外部源によって定義されてもよく、あるいは外部値が提供されない場合にはデフォルトの利得値としてエンコーダ内で内部的に生成されてもよい。

【0090】

ある実施形態では、各プロファイルについての利得値は、ピーク、真のピーク、ダイアログの短期ラウドネスまたは全体的な短期ラウドネスまたは両方の組み合わせ（ハイブリッド）といったオーディオ信号のある種の特性の解析に基づいて計算されるDRC利得語において具現される。選ばれたプロファイルに基づく静的な利得（すなわち、伝達特性または曲線）および可能な各装置プロファイルおよび/または目標ラウドネスについての最終的なDRC利得の高速/低速アタックおよび高速/低速リリースを実装するために必要な時定数が計算される。上記のように、これらのプロファイルは、エンコーダ、デコーダにお

10

20

30

40

50

いて事前設定されていてもよく、または外部で生成されてコンテンツ・クリエイターから外部メタデータを介してエンコーダに搬送されてもよい。

【 0 0 9 1 】

ある実施形態では、利得値は、オーディオ・コンテンツのすべての周波数に対して同じ利得を適用する広帯域利得であってもよい。あるいはまた、利得は、マルチ帯域利得値からなり、オーディオ・コンテンツの異なる周波数または周波数帯域に異なる利得値が適用されてもよい。マルチチャネルの場合、各プロファイルは、単一の利得値の代わりに種々の周波数帯域についての利得を示す利得値のマトリクスを構成していてもよい。

【 0 0 9 2 】

図 1 0 を参照するに、ある実施形態では、聴取環境および / または再生装置の機能および構成の属性または特性に関する情報は、デコーダ段によって、フィードバック・リンク 3 3 0 によってエンコーダ段に提供される。プロファイル情報 3 3 2 もエンコーダ 3 0 4 に入力される。ある実施形態では、デコーダは、ビットストリーム中のメタデータを解析して、オーディオ再生装置の第一の群についてのラウドネス・パラメータがビットストリームにおいて利用可能であるかどうかを判定する。もし利用可能であれば、デコーダは、オーディオのレンダリングにおいて使うために、それらのパラメータを下流に伝送する。そうでなければ、エンコーダは前記装置のある種々の特性を解析してそれらのパラメータを導出する。次いで、これらのパラメータは再生のための下流のレンダリング・コンポーネントに送られる。エンコーダは、受領されたオーディオ・ストリームをレンダリングする出力装置（または該出力装置を含む出力装置の群）をも決定する。たとえば、出力装置は、携帯電話であるまたはポータブル装置のような群に属すると判別されてもよい。ある実施形態では、デコーダは、フィードバック・リンク 3 3 0 を使って、判別された出力装置または出力装置の群をエンコーダに示す。このフィードバックのために、出力装置に接続されたモジュール（たとえば、ヘッドセットに接続されたまたはラップトップ内のスピーカーに接続されたサウンドカード中のモジュール）がデコーダに対して出力装置の素性またはその出力装置を含む装置群の素性を示してもよい。デコーダはフィードバック・リンク 3 3 0 を通じてこの情報をエンコーダに送信する。ある実施形態では、デコーダは実行し、デコーダはラウドネスおよびDRCパラメータを決定する。ある実施形態では、デコーダはラウドネスおよびDRCパラメータを決定する。この実施形態では、フィードバック・リンク 3 3 0 を通じて前記情報を送信する代わりに、デコーダは、判別された装置または出力装置の群についての情報を使って、ラウドネスおよびDRCパラメータを決定する。もう一つの実施形態では、別のオーディオ処理ユニットがラウドネスおよびDRCパラメータを決定し、デコーダは前記情報をデコーダの代わりにそのオーディオ処理ユニットに送信する。

【 0 0 9 3 】

図 1 2 は、ある実施形態のもとの、ダイナミックレンジ最適化システムのブロック図である。図 1 2 に示されるように、エンコーダ 4 0 2 は入力オーディオ 4 0 1 を受領する。エンコードされたオーディオはマルチプレクサ 4 0 9 において選択された圧縮曲線 4 2 2 およびdialnorm値 4 2 4 から生成されたパラメータ 4 0 4 と組み合わせられる。結果として得られるビットストリームは、デマルチプレクサ 4 1 1 に伝送され、デマルチプレクサ 4 1 1 がデコーダ 4 0 6 によってデコードされるオーディオ信号を生成する。パラメータおよびdialnorm値は、デコーダ出力の増幅のために増幅器 4 1 0 を駆動する利得レベルを生成するために利得計算ユニット 4 0 8 によって使用される。図 1 2 は、ダイナミックレンジ制御がどのようにパラメータ化され、ビットストリーム中に挿入されるかを示している。ラウドネスも同様のコンポーネントを使ってパラメータ化され、ビットストリーム中に挿入されることができる。ある実施形態では、出力参照レベル制御（図示せず）がデコーダに提供されることもできる。図はラウドネスおよびダイナミックレンジ・パラメータがエンコーダにおいて決定され、挿入されるものとして示しているが、同様の決定は、前処理器、デコーダおよび後処理器のような他のオーディオ処理ユニットにおいて実行されることができる。

10

20

30

40

50

【 0 0 9 4 】

図 1 5 は、ある実施形態のもとでの、オーディオ・コンテンツの種々の型についてのラウドネス・プロファイルの例を示している。図 1 5 に示されるように、例示的な曲線 6 0 0 および 6 0 2 は、0 LKFSを中心とした利得に対する（LKFSでの）入力ラウドネスをプロットしている。種々の型のコンテンツは、図 1 5 に示されるような種々の曲線を示す。図では、曲線 6 0 0 は発話を表わしていてもよく、曲線 6 0 2 は標準的なフィルム・コンテンツを表わしていてもよい。図 1 5 に示されるように、発話コンテンツは、フィルム・コンテンツに比して、より大きな量の利得を受ける。図 1 5 は、ある種の型のオーディオ・コンテンツについての代表的なプロファイル曲線の例であることが意図されており、他のプロファイル曲線も使用されうる。図 1 5 に示されるようなプロファイル特性のある種の側面は、最適化システムのための有意なパラメータを導出するために使われる。ある実施形態では、これらのパラメータは：ヌル帯域幅、カット比、ブースト比、最大ブースト、FSアタック、FS減衰、ホールドオフ、ピーク制限および目標レベル・ラウドネスを含む。用途の要求およびシステム制約条件に依存してこれらのパラメータの少なくともいくつかに対して追加的または代替的に、他のパラメータが使われてもよい。

10

【 0 0 9 5 】

図 1 6 は、ある実施形態のもとでの、再生装置およびアプリケーションを横断してラウドネスおよびダイナミックレンジを最適化する方法を示している。図は、ラウドネスおよびダイナミックレンジ最適化をエンコーダにおいて実行されるものとして示しているが、同様の最適化は前処理器、デコーダおよび後処理器のような他のオーディオ処理ユニットにおいて実行されることができる。プロセス 6 2 0 において示されるように、本方法は、源からの入力信号を受領する（6 0 3）エンコーダ段をもって始まる。次いで、エンコーダまたは前処理コンポーネントが、源信号が、目標ラウドネスおよび/またはダイナミックレンジを達成するプロセスを受けたか否かを判定する（6 0 4）。目標ラウドネスは長期ラウドネスに対応し、外部または内部で定義されうる。源信号が目標ラウドネスおよび/またはダイナミックレンジを達成するためのプロセスを受けていない場合、本システムは、適切なラウドネスおよび/またはダイナミックレンジ制御動作を実行し（6 0 8）；そうでなく、源信号がこのラウドネスおよび/またはダイナミックレンジ制御動作を受けていた場合には、本システムは、バイパス・モードにはいってラウドネス制御および/またはダイナミックレンジ動作をスキップし、もとのプロセスが、適切な長期ラウドネスおよび/またはダイナミックレンジを指定することを許容する（6 0 6）。バイパス・モード 6 0 6 または実行されるモード 6 0 8 のいずれかについての適切な利得値（これは、単一の広帯域利得値であってもよく、あるいは周波数依存のマルチ帯域利得値であってもよい）が次いでデコーダにおいて適用される（6 1 2）。

20

30

【 0 0 9 6 】

ビットストリーム・フォーマット

先述したように、ラウドネスおよびダイナミックレンジを最適化するためのシステムは、エンコーダとデコーダの間または源とレンダリング/再生装置の間のビットストリームにおいて伝送されるメタデータおよびオーディオ・コンテンツが、互いから分離されたり、あるいは他の仕方でもネットワークもしくはサービス・プロバイダー・インターフェースなどのような他の独自の設備を通じた伝送中に損なわれたりしていないことを保証するために、安全な拡張可能なメタデータ・フォーマットを用いる。このビットストリームは、エンコーダおよび/またはデコーダ・コンポーネントが、適切なプロファイル情報を通じてオーディオ・コンテンツおよび出力装置特性に適合するようオーディオ信号のラウドネスおよびダイナミックレンジを適応させるための信号伝達のための機構を提供する。ある実施形態では、システムは、エンコーダとデコーダの間で伝送されるべき低ビットレートのエンコードされたビットストリームを決定するよう構成され、メタデータを通じてエンコードされるラウドネス情報は、一つまたは複数の出力プロファイルのための特性を含む。ある実施形態のもとでの、ラウドネスおよびダイナミックレンジ最適化システムとともに使うためのビットストリーム・フォーマットについて次に述べる。

40

50

【 0 0 9 7 】

AC-3のエンコードされたビットストリームは、メタデータおよび一ないし六個のチャンネルのオーディオ・コンテンツを有する。オーディオ・コンテンツは、知覚的オーディオ符号化を使って圧縮されたオーディオ・データである。メタデータは、聴取環境に送達されるプログラムの音を変える際に使うために意図されているいくつかのオーディオ・メタデータ・パラメータを含む。AC-3エンコードされたオーディオ・ビットストリームの各フレームは、デジタル・オーディオの1536サンプルについてのオーディオ・コンテンツおよびメタデータを含む。48kHzのサンプリング・レートについては、これは32ミリ秒のデジタル・オーディオまたはオーディオの31.25フレーム毎秒のレートを表わす。

【 0 0 9 8 】

E-AC-3エンコードされたオーディオ・ビットストリームの各フレームは、フレームに含まれるオーディオ・データが一、二、三または六ブロックのいずれであるかに依存して、それぞれデジタル・オーディオの256、512、768または1536サンプルについてのオーディオ・コンテンツおよびメタデータを含む。48kHzのサンプリング・レートについては、これはそれぞれ5.333、10.667、16または32ミリ秒のデジタル・オーディオまたはそれぞれオーディオの189.9、93.75、62.5または31.25フレーム毎秒のレートを表わす。

【 0 0 9 9 】

図4に示されるように、各AC-3フレームはセクション（セグメント）に分割される。セクションは、（図5に示されるように）同期語（SW）および二つの誤り訂正語のうち第一のもの（CRC1）を含む同期情報（SI）セクションと；メタデータの大半を含むビットストリーム情報（BSI）セクションと；データ圧縮されたオーディオ・コンテンツを含む（そしてメタデータも含むことができる）六つのオーディオ・ブロック（AB0からAB5）と；オーディオ・コンテンツが圧縮されたのちに残される未使用ビットがあればそれを含む余剰（waste）ビット・セグメント（W）と；さらなるメタデータを含んでいてもよい補助（AUX）情報セクションと；二つの誤り訂正語のうちの第二のもの（CRC2）とを含む。

【 0 1 0 0 】

図7に示されるように、各E-AC-3フレームはセクション（セグメント）に分割される。セクションは、（図5に示されるように）同期語（SW）を含む同期情報（SI）セクションと；メタデータの大半を含むビットストリーム情報（BSI）セクションと；データ圧縮されたオーディオ・コンテンツを含む（そしてメタデータも含むことができる）一から六個までの間のオーディオ・ブロック（AB0からAB5）と；オーディオ・コンテンツが圧縮されたのちに残される未使用ビットがあればそれを含む余剰（waste）ビット・セグメント（W）と；さらなるメタデータを含んでいてもよい補助（AUX）情報セクションと；誤り訂正語（CRC）とを含む。

【 0 1 0 1 】

AC-3（またはE-AC-3）ビットストリームでは、聴取環境に送達されるプログラムの音を変える際に使うよう特に意図されたいくつかのオーディオ・メタデータ・パラメータがある。そうしたメタデータ・パラメータの一つはdialnormパラメータであり、これはBSIセグメントに含まれる。

【 0 1 0 2 】

図6に示されるように、AC-3フレームのBSIセグメントは、当該プログラムについてのdialnorm値を示す五ビットのパラメータ（「dialnorm」）を含む。当該AC-3フレームのオーディオ符号化モード（「acmod」）が「0」であってデュアル・モノあるいは「1+1」チャンネル構成が使われていることを示す場合には、同じAC-3フレームにおいて担持される第二のオーディオ・プログラムについてのdialnorm値を示す五ビットのパラメータ（「dialnorm2」）が含まれる。

【 0 1 0 3 】

BSIセグメントは、フラグ（「addbsie」）であって、該「addbsie」ビットに続く追加的なビットストリーム情報の存在（または不在）を示すフラグと、パラメータ（「addbsil」）であって、該「addbsil」値に続く追加的なビットストリーム情報があればその長さ

10

20

30

40

50

を示すパラメータと、「addbsil」値に続く64ビットまでの追加的なビットストリーム情報（「addbsi」）とを含む。

【0104】

BSIセグメントは、図6に具体的に示されない他のメタデータ値を含んでいてもよい。

【0105】

本稿に記載される一つまたは複数の実施形態の諸側面は、ソフトウェア命令を実行する一つまたは複数のコンピュータまたは処理装置を含むネットワークを通じた伝送のためのオーディオ信号を処理するオーディオ・システムにおいて実装されてもよい。記載される実施形態のいずれも、単独で、または任意の組み合わせで互いと一緒に使用されてもよい。さまざまな実施形態が、本明細書の一つまたは複数の場所で論じられるまたは暗示されることがありうる従来技術でのさまざまな欠点によって動機付けられていることがありうるが、それらの実施形態は必ずしもこれらの欠点のいずれかに取り組むものではない。つまり、種々の実施形態は本明細書において論じられることがある種々の欠点に取り組むことがある。いくつかの実施形態は、本明細書において論じられることがあるいくつかの欠点または一つだけの欠点に部分的に取り組むだけであることがあり、いくつかの実施形態はこれらの欠点のどれにも取り組まないこともある。

10

【0106】

本稿に記述されるシステムの諸側面は、デジタルまたはデジタイズされたオーディオ・ファイルを処理するための適切なコンピュータ・ベースの音処理ネットワーク環境において実装されてもよい。適応オーディオ・システムの諸部分は、コンピュータ間で伝送されるデータをバッファリングおよびルーティングするはたらきをする一つまたは複数のルーター（図示せず）を含め、任意の所望される数の個々の機械を含む一つまたは複数のネットワークを含んでいてもよい。そのようなネットワークは、さまざまなネットワーク・プロトコル上で構築されてもよく、インターネット、広域ネットワーク（WAN）、ローカル・エリア・ネットワーク（LAN）またはその任意の組み合わせであってもよい。

20

【0107】

上記のコンポーネント、ブロック、プロセスまたは他の機能構成要素の一つまたは複数、システムのプロセッサ・ベースのコンピューティング装置の実行を制御するコンピュータ・プログラムを通じて実装されてもよい。本稿に開示されるさまざまな機能は、ハードウェア、ファームウェアのいくつもある組み合わせを使っておよび/またはさまざまな機械可読もしくはコンピュータ可読媒体において具現されたデータおよび/または命令として、挙動上の、レジスタ転送、論理コンポーネントおよび/または他の特性を用いて記載されることがあることも注意しておくべきである。そのようなフォーマットされたデータおよび/または命令が具現されうるコンピュータ可読媒体は、光学式、磁気式もしくは半導体記憶媒体のようなさまざまな形の物理的（非一時的）、不揮発性記憶媒体を含むがそれに限定されない。

30

【0108】

文脈がそうでないことを明確に要求するのではないが、本記述および請求項を通じて、単語「有する」「含む」などは、排他的もしくは網羅的な意味ではなく包含的な意味に解釈されるものとする。すなわち、「...を含むがそれに限定されない」の意味である。単数または複数を使った単語は、それぞれ複数または単数をも含む。さらに、「本稿で」「以下で」「上記で」「下記で」および類似の意味の単語は、全体としての本願を指すのであって、本願のいかなる特定の部分を指すものでもない。単語「または」が二つ以上の項目のリストを参照して使われるとき、その単語は該単語の以下の解釈のすべてをカバーする：リスト中の項目の任意のもの、リスト中の項目のすべておよびリスト中の項目の任意の組み合わせ。

40

【0109】

一つまたは複数の実装が、例として、個別的な実施形態を用いて記載されているが、一つまたは複数の実装は開示される実施形態に限定されないことは理解されるものとする。逆に、当業者に明白であろうさまざまな修正および類似の構成をカバーすることが意図さ

50

れている。したがって、付属の請求項の範囲は、そのようなすべての修正および類似の構成を包含するような最も広い解釈を与えられるべきである。

【 0 1 1 0 】

いくつかの態様を記載しておく。

〔 態 様 1 〕

ビットストリームにおいてオーディオ・データに関連付けられたメタデータを受領する段階と；

前記ビットストリーム中の前記メタデータを解析してオーディオ再生装置の第一の群のためのラウドネス・パラメータが前記ビットストリーム中で利用可能であるかどうかを判定する段階と；

前記第一の群のために前記パラメータが存在すると判定するのに応答して、オーディオをレンダリングするために前記パラメータおよびオーディオ・データを使う段階と；

前記第一の群のために前記ラウドネス・パラメータが存在しないと判定するのに応答して、前記第一の群の一つまたは複数の特性を解析し、前記一つまたは複数の特性に基づいて前記パラメータを決定する段階とを含む、方法。

〔 態 様 2 〕

前記一つまたは複数の特性は、前記オーディオ・データの異なるプロファイルについての利得レベルを含む、態様 1 記載の方法。

〔 態 様 3 〕

前記利得レベルは：前記オーディオ・データがある定義されたプログラムのためにどのように再生されるかを制御する芸術的ダイナミックレンジ・プロファイルと、前記定義されたプログラムの諸部分についての過負荷保護を提供する短期ダイナミックレンジ・プロファイルとのうちの少なくとも一つを定義する、態様 2 記載の方法。

〔 態 様 4 〕

オーディオをレンダリングするために前記パラメータおよびオーディオ・データを使う段階が、再生のために前記オーディオをレンダリングする下流のモジュールに前記パラメータおよびオーディオ・データを送信することを含む、態様 1 記載の方法。

〔 態 様 5 〕

オーディオをレンダリングするために前記パラメータおよびオーディオ・データを使う段階が、前記パラメータおよびオーディオ・データに基づいて前記オーディオ・データをレンダリングすることを含む、態様 1 記載の方法。

〔 態 様 6 〕

前記受領されたオーディオ・ストリームをレンダリングする出力装置を判別する段階と；

前記出力装置がオーディオ再生装置の前記第一の群に属するか否かを判定する段階とをさらに含み、ストリーム中のメタデータを解析してオーディオ再生装置の前記第一の群のためのラウドネス・パラメータが利用可能であるかどうかを判定する前記段階は、前記出力装置がオーディオ再生装置の前記第一の群に属することを判別する前記段階の後に実行される、

態様 1 記載の方法。

〔 態 様 7 〕

前記出力装置がオーディオ再生装置の前記第一の群に属することを判別する前記段階は：

前記出力装置の素性を指示するまたは前記出力装置を含む装置の群の素性を指示する前記出力装置に接続されたモジュールからの指標を受領し；

受領された指標に基づいて、前記出力装置がオーディオ再生装置の前記第一の群に属することを判別することを含む、

態様 6 記載の方法。

〔 態 様 8 〕

オーディオ・データに関連付けられたメタデータを含むビットストリームを受領するよう構成されたインターフェースと；

10

20

30

40

50

前記ビットストリーム中の前記メタデータを解析してオーディオ再生装置の第一の群のためのラウドネス・パラメータが前記ビットストリーム中で利用可能であるかどうかを判定するよう構成された、前記インターフェースに結合された解析器と；

前記解析器が前記第一の群のために前記パラメータが存在すると判定するのに応答して、オーディオをレンダリングするために前記パラメータおよびオーディオ・データを使うよう構成されたレンダリング・コンポーネントであって、前記解析器が前記第一の群のために前記ラウドネス・パラメータが存在しないと判定するのに応答して、前記第一の群の一つまたは複数の特性を解析し、前記一つまたは複数の特性に基づいて前記パラメータを決定するようさらに構成されている、レンダリング・コンポーネントとを有する、装置。

10

〔態様 9〕

前記レンダリング・コンポーネントがオーディオをレンダリングするために前記パラメータおよびオーディオ・データを使うことが、前記レンダリング・コンポーネントが、再生のために前記オーディオをレンダリングする下流のモジュールに、前記パラメータおよびオーディオ・データを送信することを含む、態様 8 記載の装置。

〔態様 10〕

前記レンダリング・コンポーネントがオーディオをレンダリングするために前記パラメータおよびオーディオ・データを使うことが、前記レンダリング・コンポーネントが、前記パラメータおよびオーディオ・データに基づいて前記オーディオ・データをレンダリングすることを含む、態様 9 記載の装置。

20

〔態様 11〕

前記受領されたオーディオ・ストリームをレンダリングする出力装置を判別し、前記出力装置がオーディオ再生装置の前記第一の群に属するか否かを判定するよう構成された第二のコンポーネントをさらに有し、

前記解析器は、ストリーム中のメタデータを解析してオーディオ再生装置の前記第一の群のためのラウドネス・パラメータが利用可能であるかどうかを判定することを、第二のコンポーネントが、前記出力装置がオーディオ再生装置の前記第一の群に属するか否かを判別した後に行なう、態様 10 記載の装置。

〔態様 12〕

30

前記出力装置の素性を指示するまたは前記出力装置を含む装置の群の素性を指示する前記出力装置に接続されたモジュールからの指標を受領し、受領された指標に基づいて、前記出力装置がオーディオ再生装置の前記第一の群に属するよう構成されたインターフェースをさらに有する、態様 11 記載の装置。

〔態様 13〕

オーディオ・データおよび前記オーディオ・データに関連付けられたメタデータを受領する段階と；

ビットストリーム中の前記メタデータを解析してオーディオ装置の第一の群のためのラウドネス・パラメータに関連するラウドネス情報が前記ストリーム中で利用可能であるかどうかを判定する段階と；

40

前記第一の群のために前記ラウドネス情報が存在すると判定するのに応答して、前記ストリームからラウドネス情報を決定し、前記オーディオ・データおよびラウドネス情報をオーディオのレンダリングにおいて使うために送信する段階と；

前記ラウドネス情報が前記第一の群のために存在しないと判定するのに応答して、出力プロファイルに関連付けられたラウドネス情報を決定し、前記出力プロファイルについての決定されたラウドネス情報をオーディオのレンダリングにおいて使うために送信する段階とを含む、方法。

〔態様 14〕

出力プロファイルに関連付けられたラウドネス情報を決定する段階はさらに：

50

前記出力プロファイルの特性を解析し；
前記特性に基づいて前記パラメータを決定することをさらに含み、
決定されたラウドネス情報を送信することは、決定されたパラメータを送信することを含む、
態様 1 3 記載の方法。

〔態様 1 5〕

前記特性は、前記オーディオ・データの異なるプロファイルについての利得レベルを含む、態様 1 4 記載の方法。

〔態様 1 6〕

前記利得レベルは：前記オーディオ・データがある定義されたプログラムのためにどのように再生されるかを制御する芸術的ダイナミックレンジ・プロファイルと、前記定義されたプログラムの諸部分についての過負荷保護を提供する短期ダイナミックレンジ・プロファイルとのうちの少なくとも一つを定義する、態様 1 5 記載の方法。

10

〔態様 1 7〕

前記ラウドネス情報が出力プロファイルのためのラウドネス・パラメータを含む、態様 1 3 記載の方法。

〔態様 1 8〕

前記ラウドネス情報が前記出力プロファイルの前記特性を含む、態様 1 3 記載の方法。

〔態様 1 9〕

送信されるべき低ビットレートのエンコードされたストリームを決定する段階をさらに含み、前記ラウドネス情報は一つまたは複数の出力プロファイルについての特性を含む、態様 1 3 記載の方法。

20

〔態様 2 0〕

一つまたは複数の出力プロファイルがプレミアム・コンテンツ・プロファイルを含まない、態様 1 7 記載の方法。

〔態様 2 1〕

オーディオ・データをデコードする装置であって：

前記オーディオ・データおよび前記オーディオ・データに関連付けられたメタデータを受領するインターフェースと；

ビットストリーム中のメタデータを解析してオーディオ装置の第一の群のためのラウドネス・パラメータに関連するラウドネス情報が前記ストリーム中で利用可能であるかどうかを判定する第一のコンポーネントとを有し、前記第一のコンポーネントは、前記第一の群のために前記ラウドネス情報が存在すると判定するのに応答して、前記ストリームからラウドネス情報を決定し、前記オーディオ・データおよびラウドネス情報をオーディオのレンダリングにおいて使うために送信し、前記ラウドネス情報が前記第一の群のために存在しないと判定するのに応答して、出力プロファイルに関連付けられたラウドネス情報を決定し、前記出力プロファイルについての決定されたラウドネス情報をオーディオのレンダリングにおいて使うために送信する、
装置。

30

〔態様 2 2〕

前記第一のコンポーネントが前記メタデータを解析することが、前記第一のコンポーネントが前記出力プロファイルの特性を解析し、前記特性に基づいて前記パラメータを決定することを含み、前記第一のコンポーネントが決定されたラウドネス情報を送信することが、決定されたパラメータを送信することを含む、
態様 2 1 記載の装置。

40

〔態様 2 3〕

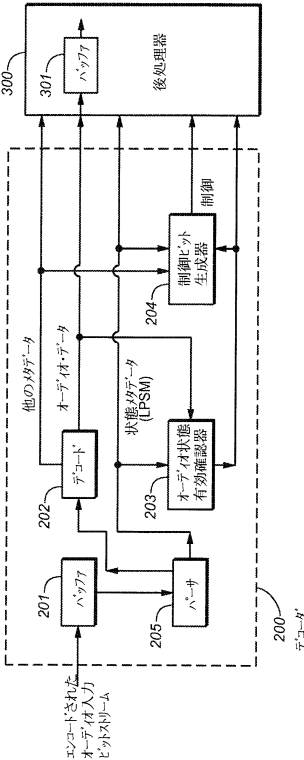
前記ラウドネス情報が出力プロファイルのためのラウドネス・パラメータを含む、態様 2 2 記載の装置。

〔態様 2 4〕

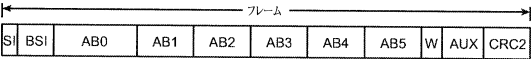
前記ラウドネス情報が前記出力プロファイルの前記特性を含む、態様 2 3 記載の装置。

50

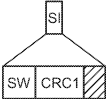
【図 3】



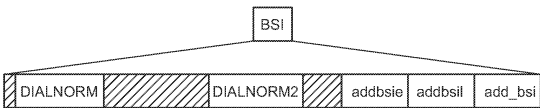
【図 4】



【図 5】



【図 6】



10

20

30

40

50

【 図 7 】

SI	BSI	AB0, AB1, AB2, AB3, AB4, AB5 のうちの1ないし6個	W	AUX	CRC2
----	-----	---	---	-----	------

【 図 8 】

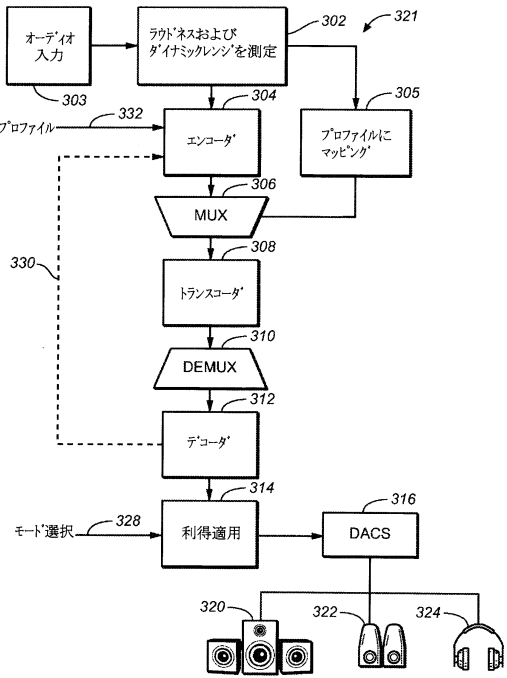
パラメータ	説明	必須 (M) / 任意的 (O)
SYNC [ID]		M
コア要素バージョン		M
コア要素長さ		M
コア要素期間 (xxx)		M
拡張要素カウント	当該コア要素に関連付けられた拡張メディア要素の数を示す。この値はビットストリームが生成から配信および最終的な配信へと渡される際にインクリメント/デクリメントしてもよい	M
サブストリーム関連付け	当該コア要素がどのサブストリーム (単数または複数) に関連付けられているかを記述	M
署名 (HMAC ダイジェスト)	フレーム全体のオーディオデータ、コア要素および全拡張要素に対して計算される 256 ビット HMAC ダイジェスト (SHA-2 アルゴリズムを使う)	M
PGM 境界カウントダウン	フィールドはオーディオプログラムファイル/ストリームの先頭または末尾のある数のフレームについてのみ現われる。よって、このパラメータが含まれることを合図するためには、コア要素バージョン変化を使うことができる	O
オーディオフィンガープリント	コア要素期間フィールドによって表わされるある数の PCM オーディオサンプルに対して取られるオーディオフィンガープリント	O
ビデオフィンガープリント	コア要素期間フィールドによって表わされるある数の圧縮されたビデオサンプル (もしあれば) に対して取られるビデオフィンガープリント	O
URL/UUID	このフィールドは、当該ビットストリームに関連する追加的なプログラムコンテンツ (エッセンス) および/またはメディアの外部の位置を参照する URL および/または UUID を担持するために定義される (フィンガープリントに対して冗長かもしれない)	O

表 1

【 図 9 】

表 2					
LFSM パラメータ [6 ビット/フィールド] 枠	説明	一意的な 状態の数	必須 (M)/ 任意的 (O)	挿入レートを 更新する 周期	
LFSM バージョン			M		
LFSM 期間 (xxx)	xxx フィールドにのみ該当		M		
LFSM カウント			M		
LFSM オブジェクト 関連付け			M		
ダイナミック (単数または 複数)	L, C, R オブジェクトのどの組み合わせが、それまでの 0.5 秒に発話を含んでいるかを示す。L, C, R のどの組み合わせにも発話が存在しないときは、このパラメータは「デフォルトなし」を示す	8	M	0.5 秒 (典型)	
ラウドネス規制型	関連付けられたオーディオストリームが特定の規制のセット (例: ATSC A/85 または EBU R128) に準拠していることを示す	8	M	フレーム	
ダイナミック ラウドネス補正	関連付けられたオーディオストリームが「ラウドネス」に基づいて補正されているかどうかを示す	2	O (ラウドネス規制型が対応するオーディオが未補正であることを示す場合にのみ存在)	フレーム	
ラウドネス補正型	関連付けられたオーディオストリームの補正が無制限読み取り (フィードバック) を用いてかリアルタイム (RT) ラウドネスおよび「デフォルト」の「デフォルト」を用いてかを示す	2	O (ラウドネス規制型が対応するオーディオが未補正であることを示す場合にのみ存在)	フレーム	
ITU 相対ラウドネス (INF)	適用されるラウドネスなしの関連付けられたオーディオストリームの ITU-R BS. 1770-2 統合ラウドネスを示す (例: 7 ビット: -58 → +5.5 LKFS 0.5 LKFS きざみ)	128	O	1 sec	
ITU 発話ラウドネス (INF)	適用されるラウドネスなしの関連付けられたオーディオストリームの ITU-R BS. 1770-2 統合ラウドネスを示す (例: 7 ビット: -58 → +5.5 LKFS 0.5 LKFS きざみ)	128	O	1 sec	
ITU (EBU 3341) 短時間 3s ラウドネス	約 10Hz の挿入レートでの (オーディオ) 密度適用されるラウドネスなしの関連付けられたオーディオストリームの 3 秒のラウドネスを示す (例: 8 ビット: 116 → +11.5 LKFS 0.5 LKFS きざみ)	256	O	0.1 sec	
真のピーク値	適用されるラウドネスなしでの関連付けられたオーディオストリームの ITU-R BS. 1770 付属書 2 の TruePeak 値 (dB TP) を示す (つまり、要素周期フィールドにおいて伝達されるフレーム期間にわたっての最大値) 116 → +11.5 LKFS 0.5 LKFS きざみ	256	O	0.5 sec	

【 図 10 】



10

20

30

40

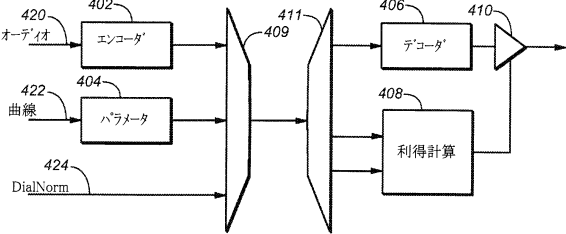
50

【図 1 1】

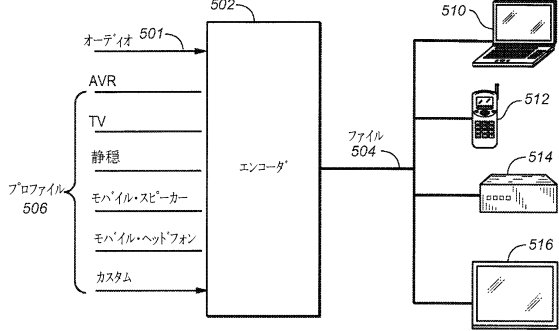
再生条件		
スピーカー	静穏	騒音あり
小/ミ	低ゲイミツクレンジ	低ゲイミツクレンジ
ヘツトフォン	中ゲイミツクレンジ	低ゲイミツクレンジ
大	高ゲイミツクレンジ	低ゲイミツクレンジ

表3

【図 1 2】



【図 1 3】

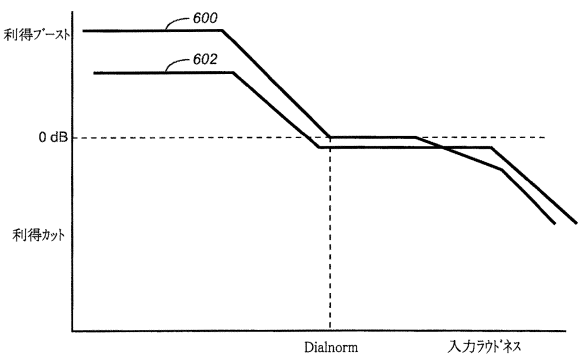


【図 1 4】

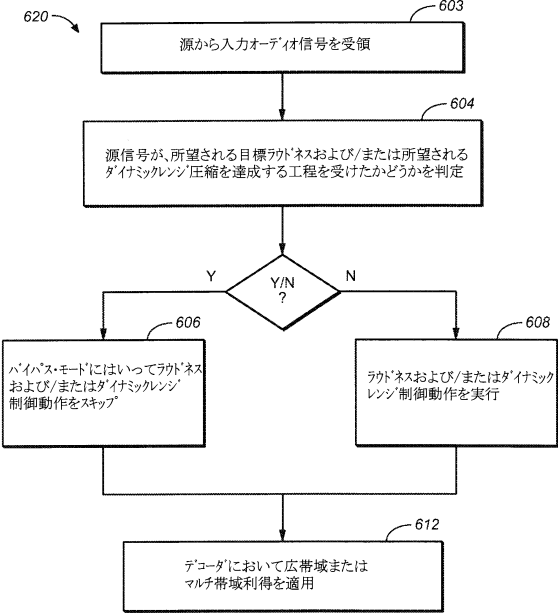
長期ラウドネス	DRC利得	過負荷圧縮
1	プロフィール 1	1
2	プロフィール 2	2
3	プロフィール 3	3
⋮	⋮	⋮
N	プロフィール N	N

表4

【図 1 5】



【図 1 6】



10

20

30

40

50

フロントページの続き

(33)優先権主張国・地域又は機関

米国(US)

(31)優先権主張番号 61/824,010

(32)優先日 平成25年5月16日(2013.5.16)

(33)優先権主張国・地域又は機関

米国(US)

(74)代理人 100070150

弁理士 伊東 忠彦

(74)代理人 100135079

弁理士 宮崎 修

(72)発明者 リードミラー, ジェフリー

アメリカ合衆国, カリフォルニア州 94103, サンフランシスコ, ポットレロ アヴェニュー 100, ドルビー ラボラトリーズ インコーポレイテッド内

(72)発明者 ノークロス, スコット グレゴリー

アメリカ合衆国, カリフォルニア州 94103, サンフランシスコ, ポットレロ アヴェニュー 100, ドルビー ラボラトリーズ インコーポレイテッド内

(72)発明者 ローエデン, カール ヨナス

スウェーデン王国, 11330 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 大野 弘

(56)参考文献 特表2012-504260(JP, A)

特表2008-505586(JP, A)

国際公開第2011/100155(WO, A1)

国際公開第2011/110525(WO, A1)

(58)調査した分野 (Int.Cl., DB名)

G10L 19/00

G10L 21/034

H04R 3/00