

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第6212565号  
(P6212565)

(45) 発行日 平成29年10月11日(2017.10.11)

(24) 登録日 平成29年9月22日(2017.9.22)

(51) Int.Cl. F I  
G 1 O L 19/00 (2013.01) G 1 O L 19/00 3 3 O B

請求項の数 24 (全 53 頁)

(21) 出願番号	特願2015-537032 (P2015-537032)	(73) 特許権者	507236292
(86) (22) 出願日	平成26年1月15日 (2014.1.15)		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
(65) 公表番号	特表2015-531498 (P2015-531498A)		アメリカ合衆国 9 4 1 0 3 カリフォル ニア州 サンフランシスコ マーケット ストリート 1 2 7 5
(43) 公表日	平成27年11月2日 (2015.11.2)		
(86) 国際出願番号	PCT/US2014/011672	(74) 代理人	100107766
(87) 国際公開番号	W02014/113465		弁理士 伊東 忠重
(87) 国際公開日	平成26年7月24日 (2014.7.24)	(74) 代理人	100070150
審査請求日	平成27年4月13日 (2015.4.13)		弁理士 伊東 忠彦
審査番号	不服2016-10128 (P2016-10128/J1)	(74) 代理人	100091214
審査請求日	平成28年7月5日 (2016.7.5)		弁理士 大貫 進介
(31) 優先権主張番号	61/754,882		
(32) 優先日	平成25年1月21日 (2013.1.21)		
(33) 優先権主張国	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	米国 (US)		

最終頁に続く

(54) 【発明の名称】 プログラム・ラウドネスおよび境界メタデータをもつオーディオ・エンコーダおよびデコーダ

(57) 【特許請求の範囲】

【請求項1】

オーディオ処理ユニットであって：

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記憶するバッファ・メモリであって、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・データおよびメタデータ・コンテナを含み、前記メタデータ・コンテナはヘッダ、一つまたは複数のメタデータ・ペイロードおよび保護データを含む、バッファ・メモリと；

前記バッファ・メモリに結合された、前記オーディオ・データをデコードするオーディオ・デコーダと；

前記オーディオ・デコーダに結合されるか前記オーディオ・デコーダと統合されている、前記オーディオ・データおよび前記メタデータ・コンテナを抽出する、パーサと；

前記保護データを使って、前記メタデータ・コンテナおよび該メタデータ・コンテナ内の前記一つまたは複数のペイロードの完全性を検証するよう構成されている検証手段とを有しており、

前記ヘッダは、前記メタデータ・コンテナの先頭を同定する同期語を含み、前記保護データは、前記一つまたは複数のメタデータ・ペイロードのあとに位置し、前記一つまたは複数のメタデータ・ペイロードは前記オーディオ・データに対して実行されたラウドネス処理の型を示すデータを含むプログラム・ラウドネス・ペイロードを含み、

前記プログラム・ラウドネス・ペイロードが、前記プログラム・ラウドネス・ペイロー

ドに含まれるラウドネス・データを生成するために使われたラウドネス測定方法を示すフィールドを含む、  
オーディオ処理ユニット。

【請求項 2】

前記メタデータ・コンテナが、スキップ・フィールド、補助データ・フィールド、adbsiフィールドおよびそれらの組み合わせからなる群から選択されるAC-3またはE-AC-3のリザーブされたデータ・スペースに格納される、請求項 1 記載のオーディオ処理ユニット。

【請求項 3】

前記一つまたは複数のメタデータ・ペイロードが、前記オーディオ・データの連続するオーディオ・プログラムの間の少なくとも一つの境界を示すメタデータを含む、請求項 1 または 2 記載のオーディオ処理ユニット。

10

【請求項 4】

前記プログラム・ラウドネス・ペイロードが、前記オーディオ・データのオーディオ・チャンネルが話されたダイアログを含むかどうかを示すフィールドを含む、請求項 1 記載のオーディオ処理ユニット。

【請求項 5】

前記プログラム・ラウドネス・ペイロードが、前記オーディオ・データのオーディオ・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かどうかを示すフィールドを含む、請求項 1 記載のオーディオ処理ユニット。

【請求項 6】

20

前記プログラム・ラウドネス・ペイロードが、前記オーディオ・データのオーディオ・プログラムのラウドネスが、無限先読みまたはファイル・ベースのラウドネス補正プロセスを使って補正されているかどうかを示すフィールドを含む、請求項 1 記載のオーディオ処理ユニット。

【請求項 7】

前記プログラム・ラウドネス・ペイロードが、ダイナミックレンジ圧縮に帰着できるいかなる利得調整もなしに前記オーディオ・データのオーディオ・プログラムの統合されたラウドネスを示すフィールドを含む、請求項 1 記載のオーディオ処理ユニット。

【請求項 8】

前記プログラム・ラウドネス・ペイロードが、ダイアログ正規化(dialog normalization)に帰着できるいかなる利得調整もなしに前記オーディオ・データのオーディオ・プログラムの統合されたラウドネスを示すフィールドを含む、請求項 1 記載のオーディオ処理ユニット。

30

【請求項 9】

前記プログラム・ラウドネス・ペイロードを使っ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、請求項 1 記載のオーディオ処理ユニット。

【請求項 10】

前記エンコードされたオーディオ・ビットストリームが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る、請求項 1 ないし 9 のうちいずれか一項記載のオーディオ処理ユニット。

40

【請求項 11】

前記エンコードされたオーディオ・ビットストリームから前記プログラム・ラウドネス・ペイロードを抽出し、前記プログラム・ラウドネス・ペイロードを認証または有効確認するよう構成されている、請求項 1 ないし 9 のうちいずれか一項記載のオーディオ処理ユニット。

【請求項 12】

前記一つまたは複数のメタデータ・ペイロードがそれぞれ、一意的なペイロード識別子を含み、前記一意的なペイロード識別子が各メタデータ・ペイロードの先頭に位置される、請求項 1 ないし 11 のうちいずれか一項記載のオーディオ処理ユニット。

【請求項 13】

50

前記同期語が値0x5838をもつ16ビット同期語である、請求項1ないし11のうちいずれか一項記載のオーディオ処理ユニット。

【請求項14】

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方法であって：

一つまたは複数のフレームにセグメント分割されている、エンコードされたオーディオ・ビットストリームを受領する段階と；

前記エンコードされたオーディオ・ビットストリームからオーディオ・データおよびメタデータのコンテナを抽出する段階であって、前記メタデータのコンテナは、ヘッダと、それに続く一つまたは複数のメタデータ・ペイロードと、それに続く保護データとを含む、段階と；

前記コンテナおよび前記一つまたは複数のメタデータ・ペイロードの完全性を、前記保護データの使用を通じて検証する段階とを含み、

前記一つまたは複数のメタデータ・ペイロードは、前記オーディオ・データに対して実行されたラウドネス処理の型を示すデータを含むプログラム・ラウドネス・ペイロードを含む、

前記プログラム・ラウドネス・ペイロードが、前記プログラム・ラウドネス・ペイロードに含まれるラウドネス・データを生成するために使われたラウドネス測定方法を示すフィールドを含む、

方法。

【請求項15】

前記エンコードされたオーディオ・ビットストリームが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る、請求項14記載の方法。

【請求項16】

前記エンコードされたオーディオ・ビットストリームから抽出されたオーディオ・データに対して、前記プログラム・ラウドネス・ペイロードを使って適応的なラウドネス処理を実行する段階をさらに含む、請求項14記載の方法。

【請求項17】

前記コンテナが、スキップ・フィールド、補助データ・フィールド、addbsiフィールドおよびそれらの組み合わせからなる群から選択されるAC-3またはE-AC-3のリザーブされたデータ・スペースに位置しておりそこから抽出される、請求項14記載の方法。

【請求項18】

前記プログラム・ラウドネス・ペイロードが、前記オーディオ・データのオーディオ・チャンネルが話されたダイアログを含むかどうかを示すフィールドを含む、請求項14記載の方法。

【請求項19】

前記プログラム・ラウドネス・ペイロードが、前記オーディオ・データのオーディオ・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かどうかを示すフィールドを含む、請求項14記載の方法。

【請求項20】

前記プログラム・ラウドネス・ペイロードが、前記オーディオ・データのオーディオ・プログラムのラウドネスが、無限先読みまたはファイル・ベースのラウドネス補正プロセスを使って補正されているかどうかを示すフィールドを含む、請求項14記載の方法。

【請求項21】

前記プログラム・ラウドネス・ペイロードが、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なしに前記オーディオ・データのオーディオ・プログラムの統合されたラウドネスを示すフィールドを含む、請求項14記載の方法。

【請求項22】

前記プログラム・ラウドネス・ペイロードが、ダイアログ正規化(dialog normalization)に帰着できるいかなる利得調整もなしに前記オーディオ・データのオーディオ・プログラムの統合されたラウドネスを示すフィールドを含む、請求項14記載の方法。

10

20

30

40

50

## 【請求項 2 3】

前記メタデータのコンテナが、前記オーディオ・データの連続するオーディオ・プログラムの間の少なくとも一つの境界を示すメタデータを含む、請求項 1 4 記載の方法。

## 【請求項 2 4】

前記メタデータのコンテナが、フレームの一つまたは複数のスキップ・フィールドまたは余剰ビット・セグメントに格納されている、請求項 1 4 記載の方法。

## 【発明の詳細な説明】

## 【技術分野】

## 【0001】

関連出願への相互参照

10

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号および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。各出願はここに引用によってその全体において組み込まれる。

## 【0002】

技術分野

本発明は、オーディオ信号処理に、より詳細にはオーディオ・コンテンツのラウドネス処理状態および当該ビットストリームによって示されるオーディオ・プログラム境界の位置を示すメタデータをもつ、オーディオ・データ・ビットストリームのエンコードおよびデコードに関する。本発明のいくつかの実施形態は、AC-3、向上AC-3またはE-AC-3またはドルビーEとして知られるフォーマットのの一つになっているオーディオ・データを生成またはデコードする。

20

## 【背景技術】

## 【0003】

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・コーポレー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。

## 【0004】

オーディオ・データ処理ユニットは典型的には、盲目的な仕方で動作し、データが受領される前に行なわれるオーディオ・データの処理履歴には注意を払わない。これは、単一エンティティが多様な目標メディア・レンダリング装置のためのすべてのオーディオ・データ処理およびエンコードを行ない、目標メディア・レンダリング装置がエンコードされたオーディオ・データのすべてのデコードおよびレンダリングを行なう処理フレームワークでは機能するかもしれない。しかしながら、この盲目的な処理は、複数のオーディオ処理ユニットが多様ネットワークを横断して分散しているまたは縦続的に（チェーンに）配置されていて、それぞれの型のオーディオ処理を最適に実行することが期待される状況では、うまく機能しない（または全く機能しない）。たとえば、一部のオーディオ・データは高性能メディア・システムのためにエンコードされていることがあり、メディア処理チェーンに沿ってモバイル装置に好適な軽減された形に変換される必要があることがある。よって、オーディオ処理ユニットは、すでに実行されている型の処理を該オーディオ・データに対して不必要に実行することがある。たとえば、ボリューム平準化（leveling）ユニットは、入力オーディオ・クリップに対して、以前に該入力オーディオ・クリップに対して同じまたは同様のボリューム平準化が実行されているか否かに関わりなく、処理を実行する。結果として、ボリューム平準化ユニットは、必要でないときにも平準化を実行することがある。この不必要な処理は、オーディオ・データ中のコンテンツをレンダリングする際に、特定の特徴の劣化および/または除去を引き起こすこともある。

30

40

## 【0005】

オーディオ・データの典型的なストリームは、オーディオ・コンテンツ（たとえば一つまたは複数のチャンネルのオーディオ・コンテンツ）および該オーディオ・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む。たとえば、AC-3ビットストリームで

50

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パラメータである。これはオーディオ・プログラムに現われるダイアログの平均レベルを示すことが意図されており、オーディオ再生信号レベルを決定するために使われる。

【0006】

種々のオーディオ・プログラム・セグメント（それぞれ異なるDIALNORMパラメータをもつ）のシーケンスを有するビットストリームの再生中、AC-3デコーダは各セグメントのDIALNORMパラメータを使って、該セグメントのシーケンス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あるよう再生レベルまたはラウドネスを修正するような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・アイテム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（アイテム）は、（一般に）異なるDIALNORMパラメータをもち、デコーダは、各アイテム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であるまたは非常に似通っているよう各アイテムのレベルをスケールリングする。ただし、これは再生中に、異なるアイテムに異なる量の利得を適用することを要求することがある。

【0007】

DIALNORMは典型的に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ていない場合には、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（オーディオ・プログラムの話されるダイアログのラウドネスを示す）結果を、DIALNORM値を設定するようエンコーダに転送してもよい。このように、DIALNORMパラメータを正しく設定するには、コンテンツ・クリエイターに頼っている。

【0008】

AC-3ビットストリーム中のDIALNORMパラメータがなぜ正しくないことがあるかについては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、DIALNORM値がコンテンツ・クリエイターによって設定されない場合にビットストリームの生成中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てDIALNORM値をしかるべく設定したとしても、推奨されるAC-3ラウドネス測定方法に準拠しないラウドネス測定アルゴリズムまたはメーターが使われ、その結果正しくないDIALNORM値を与えた可能性がある。第三に、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測定され正しく設定された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/または記憶の間に正しくない値に変更された可能性がある。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。

【0009】

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。本発明まで、オーディオ・ビットストリームは、オーディオ・ビットストリームのオーディオ・コンテンツのラウドネス処理状態（たとえば、それに適用されたラウドネス処理の型（単数または複数））またはビットストリームのオーディオ・コンテンツのラウドネス処理状態およびラウドネスを本開示に記述される型のフォーマットで示すメタデータを含まなかった。そのようなフォーマットにおけるラウドネス処理メタデータは、オーディオ・ビットストリームの適応ラウドネス処理および/またはオーディオ・コンテンツのラウドネス処理状態およびラウドネスの有効性の検証を、特に効率的な仕方で容易にするために有用である。

## 【 0 0 1 0 】

本発明は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一緒に使用に限定されるものではないが、便宜上、ラウドネス処理状態メタデータを含むそのようなビットストリームを生成、デコードまたは他の仕方処理する実施形態において記述される。

## 【 0 0 1 1 】

AC-3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・コンテンツを含む。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。

10

## 【 0 0 1 2 】

AC-3 (ドルビー・デジタルとしても知られる) 符号化の詳細はよく知られており、非特許文献1、特許文献1、2、3、4、5を含む多くの刊行物で記述されている。そのすべてはここに参照によってその全体において組み込まれる。

## 【 0 0 1 3 】

ドルビー・デジタル・プラス (E-AC-3) の詳細は、非特許文献2に記載されている。

## 【 0 0 1 4 】

ドルビーE符号化の詳細は、非特許文献3および非特許文献4に記載されている。

## 【 0 0 1 5 】

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。

20

## 【 0 0 1 6 】

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。

30

## 【 0 0 1 7 】

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第二のもの (CRC2) とを含む。余剰ビット・セグメント (W) は、「スキップ・フィールド」と称されることもある。

40

## 【 0 0 1 8 】

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) (一つの余剰ビット・セグメントしか示されていないが、典型的には各オーディオ・ブロックに続いて異なる余剰ビット・セグメントがある) と; さらなるメタデータを含んで

50

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。余剰ビット・セグメント (W) は、「スキップ・フィールド」と称されることもある。

【 0 0 1 9 】

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 0 2 0 】

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。

10

【 0 0 2 1 】

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。

20

【 0 0 2 2 】

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む。

【 先行技術文献 】

【 特許文献 】

【 0 0 2 3 】

【 特許文献 1 】 米国特許第5,583,962号

【 特許文献 2 】 米国特許第5,632,005号

【 特許文献 3 】 米国特許第5,633,981号

【 特許文献 4 】 米国特許第5,727,119号

【 特許文献 5 】 米国特許第6,021,386号

30

【 非特許文献 】

【 0 0 2 4 】

【 非特許文献 1 】 ATSC Standard A52/A: Digital Audio Compression Standard (AC-3), Revision A, Advanced Television Systems Committee, 20 Aug. 2001

【 非特許文献 2 】 Introduction to Dolby Digital Plus, an Enhancement to the Dolby Digital Coding System, AES Convention Paper 6196, 117th AES Convention, October 28, 2004

【 非特許文献 3 】 "Efficient Bit Allocation, Quantization, and Coding in an Audio Distribution System", AES Preprint 5068, 107th AES Conference, August 1999

40

【 非特許文献 4 】 "Professional Audio Coder Optimized for Use with Video", AES Preprint 5033, 107th AES Conference August 1999

【 発明の概要 】

【 課題を解決するための手段 】

【 0 0 2 5 】

あるクラスの諸実施形態において、本発明は、バッファ・メモリ、オーディオ・デコーダおよびパーサを含むオーディオ処理ユニットである。バッファ・メモリ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・ビットストリームはオーディオ・データおよびメタデータ・コンテナを含む。メタデータ・コンテナはヘッダと、一つまたは複数のメタデータ・ペイロードと

50

、保護データとを含む。ヘッダは、コンテナの先頭を同定する同期語を含む。前記一つまたは複数のメタデータ・ペイロードは、オーディオ・データに関連付けられたオーディオ・プログラムを記述する。保護データは、前記一つまたは複数のメタデータ・ペイロードのあとに位置する。保護データは、メタデータ・コンテナおよび該メタデータ・コンテナ内の前記一つまたは複数のペイロードの完全性を検証するために使われることもできる。前記オーディオ・デコーダは、前記バッファ・メモリに結合されており、前記オーディオ・データをデコードすることができる。前記パーサは、前記オーディオ・デコーダに結合されるかこれと統合されるかしており、前記メタデータ・コンテナをパースすることができる。

【0026】

典型的な実施形態では、本方法は、一つまたは複数のフレームにセグメント分割されている、エンコードされたオーディオ・ビットストリームを受領することを含む。前記オーディオ・データは、メタデータのコンテナとともに、前記エンコードされたオーディオ・ビットストリームから抽出される。前記メタデータのコンテナは、ヘッダと、それに続く一つまたは複数のメタデータ・ペイロードと、それに続く保護データとを含む。最後に、前記コンテナおよび前記一つまたは複数のメタデータ・ペイロードの完全性が、前記保護データの使用を通じて検証される。前記一つまたは複数のメタデータ・ペイロードは、前記オーディオ・データに関連付けられたオーディオ・プログラムの測定されたラウドネスを示すデータを含むプログラム・ラウドネス・ペイロードを含んでいてもよい。

【0027】

ラウドネス処理状態メタデータ(LPSM: loudness processing state metadata)と称される、本発明の典型的な実施形態に従ってオーディオ・ビットストリームに埋め込まれたプログラム・ラウドネス・メタデータのペイロードは、たとえばラウドネス規制エンティティが特定のプログラムのラウドネスがすでに指定された範囲内であるかどうかおよび対応するオーディオ・データ自身が修正されていないことを検証する(verify)(それにより該当する規制に準拠していることを保証する)ことができるようにするために、認証され(authenticated)有効確認され(validated)てもよい。これを検証するために、ラウドネスを再び計算する代わりに、ラウドネス処理状態メタデータを含むデータ・ブロックに含まれるラウドネス値が読み出されてもよい。LPSMに回答して、規制当局は、(LPSMによって示される場所により)対応するオーディオ・コンテンツがラウドネスの法制および/または規制上の要求(たとえば「CALM法」としても知られる商業広告ラウドネス緩和法(Commercial Advertisement Loudness Mitigation Act))に準拠していることを、オーディオ・コンテンツのラウドネスを計算する必要なしに、判別しうる。

【0028】

いくつかのラウドネス法制および/または規制要求(たとえばCALM法のもとで発布される規則)への準拠のために要求されるラウドネス測定は、統合された(integrated)プログラム・ラウドネスに基づく。統合されたプログラム・ラウドネスは、ダイアログ・レベルまたはフル・ミックス・レベルいずれかのラウドネス測定が、オーディオ・プログラム全体に対してなされることを要求する。このように、典型的な法律上の要求への準拠を検証するためにプログラム・ラウドネス測定を(たとえば放送チェーンにおけるさまざまな段階において)なすためには、どのオーディオ・データ(およびメタデータ)がオーディオ・プログラム全体を決定するかの知識をもって測定がなされることが本質的であり、これは典型的には、(たとえば諸オーディオ・プログラムのシーケンスを示すビットストリームの処理中に)当該プログラムの始まりと終わりの位置の知識を必要とする。

【0029】

本発明の典型的な実施形態によれば、エンコードされたオーディオ・ビットストリームは少なくとも一つのオーディオ・プログラム(たとえばオーディオ・プログラムのシーケンス)を示し、該ビットストリームに含まれるプログラム境界メタデータおよびLPSMが、プログラムの終わりにおけるプログラム・ラウドネス測定のリセットを可能にし、よって統合されたプログラム・ラウドネスを測定する自動化された方法を提供する。本発明の典

10

20

30

40

50



型的な実施形態は、エンコードされたオーディオ・ビットストリームにプログラム境界メタデータを効率的な仕方で含み、これはビットストリームによって示される連続するオーディオ・プログラムの間の少なくとも一つの境界の正確かつ堅牢な決定を許容する。典型的な実施形態は、異なるプログラムを示すビットストリームが、継ぎ合わされた (splice d) ビットストリームの一方または両方を打ち切る (よって継ぎ合わせ前のビットストリームの少なくとも一つに含まれていたプログラム境界メタデータを破棄する) 仕方で (本発明のビットストリームを生成するよう) 一緒に継ぎ合わされる場合でさえ、正確なプログラム境界決定を許容するという意味で、プログラム境界の正確かつ堅牢な決定を許容する。

#### 【 0 0 3 0 】

典型的な実施形態では、本発明のビットストリームのフレームにおけるプログラム境界メタデータは、フレーム・カウントを示すプログラム境界フラグである。典型的には、このフラグは現在フレーム (当該フラグを含んでいるフレーム) とプログラム境界 (現在のオーディオ・プログラムの始まりまたは終わり) との間のフレーム数を示す。いくつかの好ましい実施形態では、プログラム境界フラグは、単独のプログラムを示す各ビットストリーム・セグメントの始まりおよび終わりにおいて対称的で効率的な仕方で (すなわち、当該セグメントの始まりのあと何らかの所定数のフレーム内に生起するフレームにおいておよび当該セグメントの終わりの前の何らかの所定数のフレーム内に生起するフレームにおいて) 挿入される。それにより、二つのそのようなビットストリームが連結される (それにより二つのプログラムのシーケンスを示すようになる) とき、プログラム境界メタデータは、二つのプログラムの間の境界の両方の側に (たとえば対称的に) 存在することができる。

#### 【 0 0 3 1 】

(一つのオーディオ・プログラムまたはオーディオ・プログラムのシーケンスを示しうる) エンコードされたオーディオ・ビットストリームにプログラム境界メタデータを含めることから帰結するデータ・レート増を制限するために、典型的な実施形態では、プログラム境界フラグは、ビットストリームのフレームの部分集合にのみ挿入される。典型的には、境界フラグ挿入レートは、(フラグが挿入される) ビットストリームの各フレームの、該各フレームに最も近いプログラム境界からの増大する離間の非増加関数である。ここで、「境界フラグ挿入レート」とは、プログラム境界フラグを含む (プログラムを示す) フレームの数の、プログラム境界フラグを含まない (該プログラムを示す) フレームの数に対する、平均的な比を表わす。ここで、平均は、エンコードされたオーディオ・ビットストリームのある数 (たとえば比較的少数) の連続するフレームにわたる移動平均である。あるクラスの諸実施形態では、境界フラグ挿入レートは、最も近いプログラム境界からの (各フラグ挿入位置の) 増大する距離の対数的に減少する関数であり、フラグの一つを含む各フラグ含有フレームについて、該フラグ含有フレーム中のフラグのサイズは、該フラグ含有フレームよりも前記最も近いプログラム境界により近くに位置するフレームにおける各フラグのサイズ以上である (すなわち、各フラグ含有フレーム内のプログラム境界フラグのサイズは、当該フラグ含有フレームの、前記最も近いプログラム境界からの増大する離間の非減少関数である)。

#### 【 0 0 3 2 】

本発明のもう一つの側面は、本発明の方法のいずれかの実施形態を実行するよう構成されたオーディオ処理ユニット (APU: audio processing unit) である。もう一つのクラスの諸実施形態では、本発明は、本発明の方法のいずれかの実施形態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) 記憶しているバッファ・メモリ (バッファ) を含むAPUである。APU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 (前処理器)、後処理システム (後処理器)、オーディオ・ビットストリーム処理システムおよびそのような要素の組み合わせを含むがこれに限られない。

#### 【 0 0 3 3 】

もう一つのクラスの諸実施形態では、本発明は、オーディオ・データ・セグメントおよびメタデータ・セグメントを含むエンコードされたオーディオ・ビットストリームを生成するよう構成されたオーディオ処理ユニット（APU）である。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、前記メタデータ・セグメントのうち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、任意的にはプログラム境界メタデータをも、含む。典型的には、ビットストリームのフレーム中の少なくとも一つのそのようなメタデータ・セグメントが、そのフレームのオーディオ・データ（すなわち、そのフレームの少なくとも一つのオーディオ・データ・セグメント中のオーディオ・データ）に対して第一の型のラウドネス処理が実行されているかどうかを示すLPSMの少なくとも一つのセグメントと、そのフレームのオーディオ・データの少なくとも一部のラウドネス（たとえば、そのフレームのオーディオ・データの、ダイアログを示す少なくとも一部のデータのダイアログ・ラウドネス）を示すLPSMの少なくとも一つの他のセグメントとを含む。このクラスのある実施形態では、APUは、エンコードされたオーディオを生成するよう入力オーディオをエンコードするよう構成されているエンコーダであり、オーディオ・データ・セグメントはエンコードされたオーディオを含む。このクラスの典型的な実施形態では、メタデータ・セグメントのそれぞれは、本稿に記載される好ましいフォーマットをもつ。

#### 【0034】

いくつかの実施形態では、LPSM（たとえばLPSMおよびプログラム境界メタデータ）を含むエンコードされたビットストリーム（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）のメタデータ・セグメントのそれぞれは、そのビットストリームのフレームのスキップ・フィールド・セグメントの余剰ビット（たとえば、図4または図7に示される型の余剰ビット・セグメントW）に含まれる。他の実施形態では、LPSM（たとえばLPSMおよびプログラム境界メタデータ）を含むエンコードされたビットストリーム（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）のメタデータ・セグメントのそれぞれは、その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」フィールド中の追加的なビットストリーム情報として、あるいはその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または図7に示される型のAUXセグメント）中に、含まれる。LPSMを含む各メタデータ・セグメントは、下記の表1および表2を参照して本稿に記載されるフォーマットを有していてもよい（すなわち、表1に記載されるコア要素またはその変形と、それに続くペイロード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・サイズの値と、それに続くペイロード（表2に示されるようなフォーマットまたは本稿に記載される表2に対する変形で示されるようなフォーマットをもつLPSMデータ）とを含む）。いくつかの実施形態では、フレームは一つまたは複数のメタデータ・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、一方がそのフレームのaddbsiフィールドに存在していて他方がそのフレームのAUXフィールドに存在していてもよい。

#### 【0035】

あるクラスの諸実施形態では、本発明は、AC-3またはE-AC-3エンコードされたオーディオ・ストリームを生成するようオーディオ・データをエンコードする段階を含む方法である。該段階は、（ビットストリームの少なくとも一つのフレームの）メタデータ・セグメント中にLPSMおよびプログラム境界メタデータを、任意的にはそのフレームが属するオーディオ・プログラムについての他のメタデータをも、含むことによることを含む。いくつかの実施形態では、それぞれのそのようなメタデータ・セグメントは、フレームのaddbsiフィールドあるいはフレームの補助データ・フィールド中に含まれる。他の実施形態では、それぞれのそのようなメタデータは、フレームの余剰ビット・セグメント中に含まれる。いくつかの実施形態では、LPSMおよびプログラム境界メタデータを含む各メタデータ・セグメントは、コア・ヘッダ（および任意的には追加的なコア要素も）と、該コア・ヘッダ後に（または該コア・ヘッダおよび他のコア要素後に）下記のフォーマットをもつLP

10

20

30

40

50

SMペイロード（またはコンテナ）セグメントとを含む。

【0036】

ヘッダ。典型的には少なくとも一つの識別情報値を含む（たとえば、本稿の表2に記載されるような、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）、

ヘッダ後に、LPSMおよびプログラム境界メタデータ。プログラム境界メタデータは、プログラム境界フレーム・カウントと、そのフレームがプログラム境界フレーム・カウントのみを含むか、プログラム境界フレーム・カウントおよびオフセット値の両方を含むかを示す符号値（たとえば「offset\_exist」〔オフセット存在〕値）と、（場合によっては）オフセット値とを含んでいてもよい。

10

【0037】

LPSMは下記を含んでいてもよい：

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。ダイアログ指示値（単数または複数）は、対応するオーディオ・データのチャンネルの任意の組み合わせまたは全部にダイアログが存在しているかどうかを示してもよい；

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；

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；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。

20

【0038】

他の実施形態では、エンコードされる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はないビットストリームであり、LPSMを（および任意的にはプログラム境界メタデータをも）含むメタデータ・セグメントのそれぞれは、追加的なデータの記憶のためにリザーブされているビットストリームのセグメント（またはフィールドまたはスロット）に含まれる。LPSMを含む各メタデータ・セグメントは、下記の表1および表2を参照して本稿に記載されるものと同様または同一のフォーマットを有していてもよい（すなわち、表1に記載されるものと同様または同一のコア要素を、続いて（当該メタデータをLPSMとして同定する）ペイロードIDおよびペイロード・サイズ値、続いてペイロード（本稿に記載される表2または表2の変形に示されるフォーマットと同様または同一のフォーマットをもつLPSMデータ）を含む）。

30

【0039】

いくつかの実施形態では、エンコードされたビットストリームは、フレームのシーケンスを含み、各フレームは、「addbsi」フィールド（時にセグメントまたはスロットと称される）および補助データ・フィールドもしくはスロットを含むビットストリーム情報（「BSI」）セグメントを含む（たとえば、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る）。当該ビットストリームは、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）および任意的にはプログラム境界メタデータを含む。LPSMは、以下のフォーマットでビットストリームに存在している。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補助データ・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余剰ビット・セグメント中に含まれる。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。

40

50

## 【 0 0 4 0 】

ヘッダ。典型的には少なくとも一つの識別情報値、たとえば、下記の表 2 に示されるような、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、LPSMおよび任意的にはプログラム境界メタデータも。プログラム境界メタデータは、プログラム境界フレーム・カウントと、そのフレームがプログラム境界フレーム・カウントのみを含むか、プログラム境界フレーム・カウントおよびオフセット値の両方を含むかを示す符号値 (たとえば「offset\_exist」〔オフセット存在〕値) と、(場合によっては) オフセット値とを含んでいてもよい。

## 【 0 0 4 1 】

LPSMは下記を含んでいてもよい：

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」)。ダイアログ指示値 (単数または複数) は、対応するオーディオ・データのチャンネルの任意の組み合わせまたは全部にダイアログが存在しているかどうかを示してもよい；

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 (EBU 3341) 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) 。

## 【 0 0 4 2 】

対応するオーディオ・データを示す少なくとも一つのラウドネス値を考慮する、使用するまたは生成する本発明の任意の実施形態において、ラウドネス値 (単数または複数) は、オーディオ・データのラウドネスおよび/またはダイナミックレンジを処理するために利用される少なくとも一つのラウドネス測定特性を示してもよい。

## 【 0 0 4 3 】

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・フィールドまたは余剰ビット・セグメントにおける各メタデータ・セグメントは次のフォーマットをもつ。

## 【 0 0 4 4 】

コア・ヘッダ (典型的にはメタデータ・セグメントの開始を同定する同期語と、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；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、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。ここで、該HMACダイジェストは、表 1 に示されるように、フレーム全体のオーディオ・データ、コア要素およびすべての展開された (expanded) 要素に対して (SHA-2アルゴリズムを使って) 計算された256ビットHMACダイジェストであってもよい) ；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。(好ましくは上記のフォーマットをもつ) LPSMペイロード・セグメントは、LPSMペイロードIDおよびLPSM

10

20

30

40

50

ペイロード・サイズの値に続く。

【 0 0 4 5 】

前段に記載される型の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」フィールドまたは余剰ビット・セグメント）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、典型的にはまた（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プログラム境界メタデータであり、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）・メタデータである；

10

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たとえば上述した型のコア・ヘッダ、保護値およびペイロードIDおよびペイロード・サイズの値）；および

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別の型のメタデータ・ペイロード）を含む。

【 0 0 4 6 】

20

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、ペイロードの一方または両方についての（またはコア要素値ならびにペイロードの一方または両方についての）保護値（単数または複数）が最後のペイロードに続くことができる。

30

【 0 0 4 7 】

いくつかの実施形態では、フレームの補助フィールド（または「addbsi」フィールドまたは余剰ビット・セグメント）中のメタデータ・セグメントのコア要素はコア・ヘッダ（典型的には識別情報値、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・データに関係する）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。

40

【 0 0 4 8 】

もう一つの好ましいフォーマットでは、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、LPSM（および任意的にはプログラム境界メタデータも）を含むメタデータ・セグメントのそれぞれは、ドルビーE保護帯域区間の最初のN個のサンプル位置に含まれる。

【 0 0 4 9 】

50

もう一つのクラスの諸実施形態では、本発明は、オーディオ・データ・セグメントおよびメタデータ・セグメントを含むエンコードされたオーディオ・ビットストリームを受領するよう結合され、構成されたAPU（たとえばデコーダ）である。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メタデータ・セグメントは、ラウドネス処理状態メタデータ（LPSM）を、任意的にはプログラム境界メタデータをも含む。当該APUはまた、前記オーディオ・データに応答して、前記ビットストリームからLPSMを抽出してデコードされたオーディオ・データを生成し、前記LPSMを使って前記オーディオ・データに対して少なくとも一つの適応ラウドネス処理動作を実行するよう結合され、構成されている。このクラスのいくつかの実施形態はまた、当該APUに結合された後処理器をも含む。該後処理器は、前記LPSMを使って前記オーディオ・データに対して少なくとも一つの適応ラウドネス処理動作を実行するよう結合され、構成されている。

10

**【 0 0 5 0 】**

もう一つのクラスの諸実施形態では、本発明は、バッファ・メモリ（バッファ）および該バッファに結合された処理サブシステムを含むオーディオ処理ユニット（APU）である。当該APUは、オーディオ・データ・セグメントおよびメタデータ・セグメントを含むエンコードされたオーディオ・ビットストリームを受領するよう結合されている。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・セグメントの少なくともいくつかの各メタデータ・セグメントは、ラウドネス処理状態メタデータ（LPSM）を、任意的にはプログラム境界メタデータをも含む。前記バッファは、（たとえば非一時的な仕方で）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記憶し、前記処理サブシステムは、該ビットストリームから前記LPSMを抽出し、前記LPSMを使って前記オーディオ・データに対して少なくとも一つの適応ラウドネス処理動作を実行するよう構成されている。このクラスの典型的な実施形態では、前記APUはエンコーダ、デコーダおよび後処理器のうちの一つである。

20

**【 0 0 5 1 】**

本発明の方法のいくつかの実装では、生成されるオーディオ・ビットストリームはAC-3エンコードされたビットストリーム、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り、ラウドネス処理状態メタデータおよび他のメタデータ（たとえば、DIALNORMメタデータ・パラメータ、ダイナミックレンジ制御メタデータ・パラメータおよび他のメタデータ・パラメータ）を含む。本方法の他のいくつかの実装では、生成されるオーディオ・ビットストリームは別の型のエンコードされたビットストリームである。

30

**【 0 0 5 2 】**

本発明の諸側面は、本発明の方法の任意の実施形態を実行するよう構成された（たとえばプログラムされた）システムまたは装置ならびに本発明の方法またはそのステップの任意の実施形態を実装するためのコードを（たとえば非一時的な仕方で）記憶するコンピュータ可読媒体（たとえばディスク）を含む。たとえば、本発明のシステムは、プログラム可能な汎用プロセッサ、デジタル信号プロセッサまたはマイクロプロセッサであって、データに対して本発明の方法またはそのステップの実施形態を含む多様な動作のうち任意のものを実行するようソフトウェアまたはファームウェアによってプログラムされたおよび/または他の仕方で構成されたものであるまたはそれを含むことができる。そのような汎用プロセッサは、入力装置と、メモリと、呈されるデータに回答して本発明の方法（またはそのステップ）の実施形態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）処理回路とを含むコンピュータ・システムであるまたはそれを含むものであってもよい。

40

**【 図面の簡単な説明 】****【 0 0 5 3 】**

**【 図 1 】** 本発明の方法のある実施形態を実行するよう構成されていてもよいシステムのある実施形態のブロック図である。

50

【図2】本発明のオーディオ処理ユニットの実施形態であるエンコーダのブロック図である。

【図3】本発明のオーディオ処理ユニットの実施形態であるデコーダならびにそれに結合された、本発明のオーディオ処理ユニットのもう一つの実施形態である後処理器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描く図である。

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

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

10

【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描く図である。

【図8】本発明のある実施形態に基づくフォーマットをもつプログラム境界メタデータを含むエンコードされたオーディオ・ビットストリームの諸フレームの図である。

【図9】図9のエンコードされたオーディオ・ビットストリームの他の諸フレームの図である。これらのフレームのいくつかは、本発明のある実施形態に基づくフォーマットをもつプログラム境界メタデータを含む。

【図10】二つのエンコードされたオーディオ・ビットストリーム、すなわちプログラム境界(「境界」とラベル付けされている)がビットストリームの二つのフレームの間の遷移と整列されているビットストリーム(IEB)と、プログラム境界(「真の境界」とラベル付けされている)がビットストリームの二つのフレームの間の遷移から512サンプルだけオフセットされている別のビットストリーム(TB)とを描く図である。

20

【図11】四つのエンコードされたオーディオ・ビットストリームを示す図のセットである。図11のいちばん上のビットストリーム(「シナリオ1」とラベル付けされている)は、プログラム境界メタデータを含む第一のオーディオ・プログラム(P1)と、それに続く、やはりプログラム境界メタデータを含む第二のオーディオ・プログラム(P2)とを示しており；第二のビットストリーム(「シナリオ2」とラベル付けされている)は、プログラム境界メタデータを含む第一のオーディオ・プログラム(P1)と、それに続く、プログラム境界メタデータを含まない第二のオーディオ・プログラム(P2)とを示しており；第三のビットストリーム(「シナリオ3」とラベル付けされている)は、プログラム境界メタデータを含む打ち切りされた第一のオーディオ・プログラム(P1)であって、プログラム境界メタデータを含む第二のオーディオ・プログラム(P2)全体と継ぎ合わされているものを示しており；第四のビットストリーム(「シナリオ4」とラベル付けされている)は、プログラム境界メタデータを含む打ち切りされた第一のオーディオ・プログラム(P1)と、プログラム境界メタデータを含む打ち切りされた第二のオーディオ・プログラム(P2)であって、第一のオーディオ・プログラムの一部と継ぎ合わされているものを示している。

30

【発明を実施するための形態】

【0054】

#### 記法および命名法

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。

40

【0055】

請求項を含む本開示を通じて、「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

50

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。

【0056】

請求項を含む本開示を通じて、「プロセッサ」という表現は、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。

10

【0057】

請求項を含む本開示を通じて、「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。

20

【0058】

請求項を含む本開示を通じて、「処理状態メタデータ」(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)という表現は、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、典型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。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いてもよい。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いてもよい。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いてもよい。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。

30

40

【0059】

請求項を含む本開示を通じて、「ラウドネス処理状態メタデータ」(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、典型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)を含んでいてもよい。

【0060】

請求項を含む本開示を通じて、「チャンネル」(または「オーディオ・チャンネル」とい

50



う表現は、モノフォニック・オーディオ信号を表わす。

【0061】

請求項を含む本開示を通じて、「オーディオ・プログラム」という表現は、一つまたは複数のオーディオ・チャンネルおよび任意的には関連するメタデータ（たとえば、所望される空間的オーディオ呈示を記述するメタデータおよび/またはLPSMおよび/またはプログラム境界メタデータ）の集合を表わす。

【0062】

請求項を含む本開示を通じて、「プログラム境界メタデータ」という表現は、少なくとも一つのオーディオ・プログラム（たとえば二つ以上のオーディオ・プログラム）を示すエンコードされたオーディオ・ビットストリームのメタデータを表わし、プログラム境界メタデータは、少なくとも一つの前記オーディオ・プログラムの少なくとも一つの境界（始まりおよび/または終わり）のビットストリーム中の位置を示す。たとえば、（オーディオ・プログラムを示すエンコードされたオーディオ・ビットストリームの）プログラム境界メタデータは、プログラムの先頭の位置（たとえば、ビットストリームのN番目のフレームの始まりまたはビットストリームのN番目のフレームのM番目のサンプル位置）を示すメタデータと、プログラムの末尾の位置（たとえば、ビットストリームのJ番目のフレームの始まりまたはビットストリームのJ番目のフレームのK番目のサンプル位置）を示す追加的なメタデータとを含んでいてもよい。

【0063】

請求項を含む本開示を通じて、「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。よって、第一の装置が第二の装置に結合するとき、その接続は、直接接続を通じてであってもよいし、他の装置および接続を介した間接的な接続を通じてであってもよい。

【0064】

発明の実施形態の詳細な説明

本発明の典型的な実施形態によれば、ラウドネス処理状態メタデータ（「LPSM」）と称されるプログラム・ラウドネス・メタデータのペイロードおよび任意的にはまた、プログラム境界メタデータが、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。該オーディオ・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。典型的には、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、前記LPSMによってラウドネス処理状態およびラウドネスが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・データを担持するために割り当てられるビットレートに影響することなく担持されるのに十分小さくてもよい。

【0065】

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、二つ以上のオーディオ処理ユニットが処理チェーン（またはコンテンツ・ライフサイクル）を通じて互いに縦続的に機能する必要があるときに特に有用である。ラウドネス処理状態メタデータをオーディオ・ビットストリームに含めなければ、たとえばチェーンにおいて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）に至るビットストリームの行程の間に二回以上シングルエンドのボリューム平準化が適用されるときに、品質、レベルおよび空間的劣化といった深刻なメディア処理問題が起こりうる。

【0066】

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）のブロック図である。ここで、システムの要素の一つまたは複数が本発明の実施形態に基づいて構成されていてもよい。システムは、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、エンコーダ、信号解析およびメタデータ補正ユニット、トランスコーダ、

10

20

30

40

50

デコーダおよび前処理ユニット。図示したシステムの変形では、要素の一つまたは複数省略されたり、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たりする。

【0067】

いくつかの実装では、図1の前処理ユニットは、入力としてオーディオ・コンテンツを含むPCM（時間領域）サンプルを受け容れ、処理されたPCMサンプルを出力するよう構成されている。エンコーダは、入力として該PCMサンプルを受け容れ、前記オーディオ・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る。前記オーディオ・コンテンツを示す前記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。エンコーダが本発明の典型的な実施形態に従って構成されている場合、エンコーダからのオーディオ・ビット

10

【0068】

図1の信号解析およびメタデータ補正ユニットは、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容れ、（たとえばエンコードされたオーディオ・ビットストリーム中のプログラム境界メタデータを使って）信号解析を実行することによって、各エンコードされたオーディオ・ビットストリーム内の処理状態メタデータが正しいかどうかを判定（たとえば有効確認）してもよい。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、該ユニットは典型的には正しくない値（単数または複数）を信号解析から得られる正しい値（単数または複数）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各エンコードされたオーディオ・ビットストリームは、エンコードされたオーディオ・データのほかに訂正された（または訂正されていない）処理状態メタデータを含んでいてもよい。

20

【0069】

図1のトランスコーダは、入力としてエンコードされたオーディオ・ビットストリームを受け容れて、応答して（たとえば入力ストリームをデコードして、デコードされたストリームを異なるエンコード・フォーマットで再エンコードすることによって）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出力してもよい。トランスコーダが本発明の典型的な実施形態に基づいて構成されている場合、トランスコーダから出力されるオーディオ・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（典型的には他のメタデータも）含む。該メタデータはビットストリームに含められていてもよい。

30

【0070】

図1のデコーダは、入力としてエンコードされた（たとえば圧縮された）ビットストリームを受け容れ、（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する。デコーダが本発明の典型的な実施形態に基づいて構成される場合、典型的な動作におけるデコーダの出力は、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された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（および典型的には他のメタデータも）の対応するストリーム；または

40

オーディオ・サンプルのストリームおよび入力された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（および典型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から決定された処理状態メタデータや制御ビットの対応するストリームなしの、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、入力された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。

50

## 【 0 0 7 1 】

図1の後処理ユニットを本発明の典型的な実施形態に基づいて構成することによって、後処理ユニットは、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容れ、サンプルと一緒に受領されたラウドネス処理状態メタデータ（および典型的には他のメタデータも）または（デコーダによってラウドネス処理状態メタデータおよび典型的にはまた他のメタデータから決定される）制御ビットを使って、それに対して後処理（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る。後処理ユニットは典型的には、該後処理されたオーディオ・コンテンツを、一つまたは複数のスピーカによる再生のためにレンダリングするようにも構成される。

## 【 0 0 7 2 】

本発明の典型的な実施形態は、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）がそのそれぞれの処理を、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・データの同時的状态に従って適応させるものを提供する。

## 【 0 0 7 3 】

図1のシステムのいずれか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）に入力されるオーディオ・データは、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を（任意的には他のメタデータも）含んでいてもよい。本発明のある実施形態によれば、このメタデータは、図1のシステムの他の要素（または図1に示されない他の源）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。入力オーディオを（メタデータとともに）受領する本処理ユニットは、該メタデータに対してまたは該メタデータに回答して（たとえば入力オーディオの適応処理）少なくとも一つの動作（たとえば有効確認）を実行し、典型的にはまた、その出力オーディオ内に該メタデータ、該メタデータの処理されたバージョンまたは該メタデータから決定される制御ビットを含めるよう構成されていてもよい。

## 【 0 0 7 4 】

本発明のオーディオ処理ユニット（またはオーディオ・プロセッサ）の典型的な実施形態は、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・データの適応処理を実行するよう構成される。いくつかの実施形態では、適応処理は、（メタデータがラウドネス処理またはそれと同様の処理がすでにオーディオ・データに対して実行されているのでないことを示す場合は）ラウドネス処理である（またはラウドネス処理を含む）。だが、（メタデータがそのようなラウドネス処理またはそれと同様の処理がすでにオーディオ・データに対して実行されていることを示す場合は）ラウドネス処理ではない（またはラウドネス処理を含まない）。いくつかの実施形態では、適応処理は、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処理ユニットがオーディオ・データの他の適応処理を実行することを保証するための、（たとえばメタデータ検証サブユニットにおいて実行される）メタデータ有効確認であるまたはそれを含む。いくつかの実施形態では、該有効確認は、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ている）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、ある型の前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新たな実行は回避されてもよい。他方、メタデータが細工されている（または他の仕方で信頼できない）ことが見出される場合、（その信頼できないメタデータによって示される）前に実行されたとされる型のメディア処理がオーディオ処理ユニットによって反復されてもよく、および/またはオーディオ処理ユニットによって前記メタデータおよび/またはオーディオ・データに対し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

10

20

30

40

50

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。

【0075】

図2は、本発明のオーディオ処理ユニットの実施形態であるエンコーダ(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。典型的には、エンコーダ100は他の処理要素(図示せず)も含む。

10

【0076】

エンコーダ100(これはトランスコーダである)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。

20

【0077】

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、(任意的にはビットストリームからプログラム境界メタデータも抽出し、)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。典型的には、デコーダ152は、LPSM(および任意的にはプログラム境界メタデータも)を使ってデコードされたオーディオ・データに対し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、LPSM(および任意的にはプログラム境界メタデータも)を使ってデコードされたオーディオ・データに対し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。典型的には、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。

30

40

【0078】

エンコーダ100およびデコーダ152のさまざまな実装が、本発明の方法の種々の実施形態を実行するよう構成される。

50

## 【 0 0 7 9 】

フレーム・バッファ 1 1 0 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 1 1 0 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 1 1 0 からパーサ 1 1 1 に呈される。

## 【 0 0 8 0 】

パーサ 1 1 1 は、ラウドネス処理メタデータ (LPSM) を、任意的にはプログラム境界メタデータ (および/または他のメタデータをも、) そのようなメタデータが含まれている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 (任意的にはプログラム境界メタデータおよび/または他のメタデータをも) オーディオ状態有効確認器 1 0 2、ラウドネス処理段 1 0 3、段 1 0 6 およびサブシステム 1 0 8 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 1 0 1 に呈するよう結合され、構成されている。エンコーダ 1 0 0 のデコーダ 1 0 1 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 1 0 3、オーディオ・ストリーム選択段 1 0 4、サブシステム 1 0 8 および典型的には状態有効確認器 1 0 2 にも呈するよう構成されている。

## 【 0 0 8 1 】

状態有効確認器 1 0 2 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ていたデータ・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 (デコーダ 1 0 1 から有効確認器 1 0 2 に提供される) 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。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。

## 【 0 0 8 2 】

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。

## 【 0 0 8 3 】

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 1 0 2 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

10

20

30

40

50

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。

【0084】

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：

（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（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。

【0085】

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。段103は、（パーサ111によって抽出されるプログラム境界メタデータによって示される）単一のオーディオ・プログラムを示す（デコーダ101から出力される）デコードされたオーディオ・データに対して適応的なラウドネス処理を実行してもよく、パーサ111によって抽出されたプログラム境界メタデータによって示される異なるオーディオ・プログラムを示す（デコーダ101から出力される）デコードされたオーディオ・データを受領するのに応答して、ラウドネス処理をリセットしてもよい。

【0086】

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には、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。サブシステム108は、（パーサ111によって抽出されるプログラム境界メタデータによって示される）単一オーディオ・プログラムを示すデコードされたオーディオ・データに対してラウドネス測定を実行してもよく、そのようなプログラム境界メタデータによって示される異なるオーディオ・プログラムを示すデコードされたオーディオ・データを受領するのに応答して、前記測定をリセットしてもよい。

【0087】

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）が存在している。本発明の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。

#### 【0088】

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 B S.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。

10

#### 【0089】

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、典型的には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。

20

#### 【0090】

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（および/または段107に渡す）。メタデータ生成器106は、段107に、エンコーダ101および/またはパーサ111によって抽出されたLPSM（および任意的にはプログラム境界メタデータおよび/または他のメタデータも）を渡してもよいし（たとえば、有効確認器102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（および任意的にはプログラム境界メタデータおよび/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107に呈してもよい（たとえば、有効確認器102からの制御ビットが、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が無効であることを示す場合）。あるいは、段107に対して、デコーダ101および/またはパーサ111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106は、サブシステム108によって生成されたラウドネス・データと、サブシステム108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含めるために、段107に対して呈するLPSM中に含めてもよい。

30

#### 【0091】

メタデータ生成器106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「HM AC」からなっていないてもよく、あるいはそれを含んでいてもよい）を生成してもよい。

40

#### 【0092】

典型的な動作では、ダイアログ・ラウドネス測定サブシステム108は、デコーダ101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（詰め込み器/

50

フォーマッタ107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。

【0093】

追加的、任意的または代替的に、エンコーダ100の106および/または108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107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。

【0094】

エンコーダ105は、選択段104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107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107に呈する。

【0095】

段107は、エンコーダ105からのエンコードされたオーディオと生成器106からのメタデータ(LPSMを含む)とを多重化して、段107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。好ましくは、エンコードされたビットストリームは、本発明のある好ましい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。

【0096】

フレーム・バッファ109は、段107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これらのフレームのシーケンスが、バッファ109から、エンコーダ100からの出力として、送達システム150に呈される。

【0097】

メタデータ生成器106によって生成され、段107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。

【0098】

本稿において、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「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。

【0099】

エンコーダ100のいくつかの実装では、メモリ109にバッファリングされている(そして送達システム150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107はLPSMを(および任意的にはプログラム境界メタデータも)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(および任意的にはプログラム境界メタデータも)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余剰ビット・セグメント(たとえば、図4または図7に示される余剰ビット・セグメント「W」)また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

10

20

30

40

50



」)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または図7に示されるAUXセグメント)に含められる。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には、一方はフレームのaddbsiフィールドに、他方はフレームのAUXフィールドに存在していてもよい。いくつかの実施形態では、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(典型的には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)、

10

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

20

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)。

【0100】

いくつかの実施形態では、LPSMおよびプログラム境界メタデータを含む各メタデータ・セグメントは、コア・ヘッダを(任意的には追加的なコア要素も)含み、該コア・ヘッダのあとに(または該コア・ヘッダおよび他のコア要素のあとに)、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:

30

ヘッダ。典型的には少なくとも一つの識別情報値(たとえば、本稿に記載される表2に示されるような、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、LPSMおよびプログラム境界メタデータ。プログラム境界メタデータは、プログラム境界フレーム・カウントと、そのフレームがプログラム境界フレーム・カウントのみを含むか、プログラム境界フレーム・カウントおよびオフセット値の両方を含むかを示す符号値(たとえば「offset\_exist」[オフセット存在]値)と、(場合によっては)オフセット値とを含んでいてもよい。

【0101】

いくつかの実装では、段107によって余剰ビット・セグメントまた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・フィールドに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:

40

コア・ヘッダ(典型的にはメタデータ・セグメントの開始を同定する同期語と、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釈、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

50

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。

【0102】

（好ましくは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。

【0103】

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、典型的にはまた（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プログラム境界メタデータであり、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）；および

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別の型のメタデータ・ペイロード）を含む。

【0104】

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（またはコア要素値ならびに両方のペイロードについての）保護ビット（単数または複数）が最後のペイロードに続くことができる。

【0105】

いくつかの実施形態では、デコーダ101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）および任意的にはプログラム境界メタデータをも含む。有効確認器102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102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102は、対応するオーディオ・データに対するプロセッサ103の動作を無効にしてもよく、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。

【0106】

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、任意的にはプログラム境界メタデータにも応答して)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。

10

#### 【0107】

図3は、本発明のオーディオ処理ユニットのある実施形態であるデコーダ(200)およびそれに結合された後処理器(300)のブロック図である。後処理器(300)は、本発明のオーディオ処理ユニットの実施形態でもある。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。典型的には、デコーダ200は他の処理要素(図示せず)も含む。

20

#### 【0108】

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。

#### 【0109】

パーサ205は、ラウドネス処理メタデータ(LPSM)を、任意的にはプログラム境界メタデータおよび他のメタデータをも、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(プログラム境界メタデータが抽出される場合にはプログラム境界メタデータをも)オーディオ状態有効確認器203および段204に呈し、LPSMを(任意的にはプログラム境界メタデータをも)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。

30

#### 【0110】

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。

#### 【0111】

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。典型的には、後処理器300は、LPSM値および任意的にはプログラム境界メタデータも使って(たとえば、単一のオーディオ・プログラムを示すオーディオ・データについての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

40

50

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。

【0112】

デコーダ200および後処理器300のさまざまな実装は、本発明の方法の種々の実施形態を実行するよう構成されている。

【0113】

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。

10

【0114】

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(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

20

【0115】

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。

30

【0116】

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

40

(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203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。

【0117】

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入力ビットストリームからパーサ205によって抽出された

50

LPSM（および任意的にはプログラム境界メタデータも）を後処理器300に呈し、後処理器300はLPSM（および任意的にはプログラム境界メタデータも）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（および任意的にはプログラム境界メタデータも）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。

【0118】

いくつかの実施形態では、デコーダ200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（LPSM）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、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。

【0119】

デコーダ200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）および任意的にはプログラム境界メタデータをも含む。デコーダ段202（および/またはパーサ205）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（任意的にはプログラム境界メタデータも）抽出するよう構成されている。LPSMを（任意的にはプログラム境界メタデータも）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余剰ビット・セグメントまた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。いくつかの実施形態では、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。

【0120】

ヘッダ（典型的には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」）；

10

20

30

40

50

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（EBU 3341）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）。

#### 【0121】

いくつかの実施形態では、LPSMおよびプログラム境界メタデータを含む各メタデータ・セグメントは、コア・ヘッダ（任意的には追加的なコア要素も）を含み、該コア・ヘッダのあとに（または該コア・ヘッダおよび他のコア要素のあとに）、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。典型的には少なくとも一つの識別情報値（たとえば、下記の表2に示されるような、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、LPSMおよびプログラム境界メタデータ。プログラム境界メタデータは、プログラム境界フレーム・カウントと、そのフレームがプログラム境界フレーム・カウントのみを含むか、プログラム境界フレーム・カウントおよびオフセット値の両方を含むかを示す符号値（たとえば「offset\_exist」〔オフセット存在〕値）と、（場合によっては）オフセット値とを含んでいてもよい。

#### 【0122】

いくつかの実装では、パーサ205（および/またはデコーダ段202）は、ビットストリームのフレームの余剰ビット・セグメントまたは「addbsi」フィールドまたは補助データ・フィールドから、次のフォーマットをもつ各メタデータ・セグメントを抽出するよう構成される：

コア・ヘッダ（典型的にはメタデータ・セグメントの開始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。

#### 【0123】

（好ましくは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。

#### 【0124】

より一般には、本発明の好ましい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ルリングすることを許容する。好ましいビットストリーム・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。

#### 【0125】

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

10

20

30

40

50

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 1 2 6 】**

あるクラスの実施形態では、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)および任意的にはプログラム境界メタデータをも含み、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。このクラスの好ましい実施形態では、メタデータ・セグメントのそれぞれは、本稿に記載される好ましいフォーマットをもつ。

10

**【 0 1 2 7 】**

ある好ましいフォーマットでは、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、またはビットストリームのフレームの余剰ビット・セグメントに(たとえばエンコーダ100の好ましい実装の段107によって)含められる。

20

**【 0 1 2 8 】**

上記の好ましいフォーマットでは、各フレームは、下記の表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(または余剰ビット・セグメント)に含む。

**【 0 1 2 9 】**

【表 1】

表 1

パラメータ	説明	必須(M)/ 任意的(O)	
SYNC [ID]	同期語は 0x5838 の値に設定された 16 ビット値であつてもよい	M	10
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タデータ要素の数を示す。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発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	20
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	30
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/またはメタデータ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	40

該好ましいフォーマットでは、addbsi (または補助データ) フィールドまたは余剰ビット・セグメントのうち、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)。これは次の表 (表 2) に示されるフォーマットをもつ。

【 0 1 3 0 】

【表 2】

表 2

LPSM パラメータ [インテリジェント・ラウドネス]	説明	一意的な状態の数	必須(M)/ 任意的(O)	挿入レート (そのパラメータの更新の周期)	
LPSM バージョン			M		10
LPSM 期間 (xxx)	xxx フィールドにのみ該当		M		
LPSM カウント			M		
LPSM サブストリーム関連 付け			M		
ダイアログ・チャンネル(単数または複数)	L, C, R オーディオ・チャンネルのどの組み合わせが、それまでの0.5秒で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「ダイアログなし」を示す	8	M	~0.5 秒 (典型)	20
ラウドネス規制型	関連付けられたオーディオ・データ・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	30
ダイアログ・ゲートッド・ラウドネス補正フラグ	関連付けられたオーディオ・ストリームがダイアログ・ゲーティングに基づいて補正されているかどうかを示す	2	0 (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	40
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先読み(ファイル・ベース)を用いてリアルタイム(RT)ラウドネスおよびダイナミックレンジ・コントローラを用いてかを示す	2	0 (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 相対ゲートッド・ラウドネス (INF)	適用されるメタデータなしの関連付けられたオーディオ・ストリームの ITU-R BS. 1770-3 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	0	1 sec	
ITU 発話ゲートッド・ラウドネス (INF)	適用されるメタデータなしの関連付けられたオーディオ・ストリームの発話/ダイアログの ITU-R BS. 1770-1/3 統合ラウドネスを示す (例: 7 ビット: -58 → +5.5 LKFS 0.5LKFS きざみ)	128	0	1 sec	10
ITU (EBU 3341) 短時間 3s ラウドネス	約 10Hz の挿入レートでの (スライディング窓) 適用されるメタデータなしの関連付けられたオーディオ・ストリームの 3 秒のゲーティングされていない ITU (ITU-BS. 1771-1) ラウドネスを示す (例: 8 ビット: 116 → +11.5 LKFS 0.5 LKFS きざみ)	256	0	0.1 sec	20
真のピーク値	適用されるメタデータ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. 1770-3 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	0	0.5 sec	30
下方混合オフセット	下方混合ラウドネス・オフセットを示す				
プログラム境界	プログラム境界が生起するまたは生起したときを、フレーム単位で示す。プログラム境界がフレーム境界にないときには、任意的なサンプル・オフセットが、実際のプログラム境界がフレーム中でどのくらい遠くで生起するかを示す				40

フォーマットでは、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（および任意的にはプログラム境界メタデータも）を含むそれぞれは：ビットストリームのフレームの余剰ビット・セグメント；また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）のうちの任意のものに（たとえばエンコーダ100の好ましい実装の段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、一方はフレームのaddbsiフィールドに存在し、他方はフレームの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）。

#### 【0131】

もう一つの好ましいフォーマットでは、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（および任意的にはプログラム境界メタデータも）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、好ましくは、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語反復レートは好ましくは、関連するビデオ・フレーム・レートと同じまま）。

#### 【0132】

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ある好ましいフォーマットでは、メタデータ・セグメントのうちLPSM（および任意的にはプログラム境界メタデータも）を含むそれぞれは、ビットストリームのフレームの、余剰ビット・セグメントに、または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の好ましい実装の段107によって）含まれる。次に、この好ましい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述べる。

#### 【0133】

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、生成されるすべてのフレーム（同期フレーム）について、ビットストリームは、フレームのaddbsiフィールド（または余剰ビット・セグメント）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。

#### 【0134】

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の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〔発話チャンネル〕：どの源チャンネル（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・チャンネルの（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・チャンネルの統合されたITU BS.1770-3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

10

20

30

40

50

(単数または複数)。

【0135】

3. (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)はバイパスされるべきである。「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べきである。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)。(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べきである。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べきである。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べきである。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べきである)。

【0136】

5. 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すべきである:「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。

【0137】

(上記の好ましいフォーマットでは)ビットストリームの各フレームの余剰ビット・セグメントまたは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、(余剰ビット・セグメントまたは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値のすべてをグラフィカル・ユーザー・インターフェース(GUI)に渡すべきである。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。

【0138】

本発明に基づいて生成されるエンコードされたビットストリームのもう一つの好ましいフォーマットでは、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の好ましい実装の段107によって)余剰ビット・セグメントに、またはAuxセグメントに、またはビットストリームのフレームのビットストリーム情

10

20

30

40

50

報(「BSI」)セグメントの「addbsi」フィールド(図6に示した)における追加的なビットストリーム情報として、含まれる。(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または余剰ビット)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。

【0139】

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、それに続いてペイロード(LPSMデータ)。LPSMデータは次のフォーマット(上記の表2に示した必須要素と同様)をもつ。

【0140】

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。

10

【0141】

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。

【0142】

20

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4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0000)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0001)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、この値の別の値(たとえば0010)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「0000」以外の何らかの値に設定される場合、loudcorrldialgatおよびloudcorrtypフィールドがペイロードのあとに続くべきである。

【0143】

30

loudcorrldialgat：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、loudcorrldialgat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。

【0144】

loudcorrtyp：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、loudcorrtypフィールドは「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。

40

【0145】

loudrelgate：相対的な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。

【0146】

loudrelgat：相対的な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.1770-3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。

50

## 【 0 1 4 7 】

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。

## 【 0 1 4 8 】

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、-58LKFSから+5.5LKFSとして解釈される。

10

## 【 0 1 4 9 】

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。

## 【 0 1 5 0 】

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.1770-3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+5.5LKFSとして解釈される。

## 【 0 1 5 1 】

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。

20

## 【 0 1 5 2 】

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+5.5LKFSとして解釈される。

## 【 0 1 5 3 】

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余剰ビット・セグメントまたは補助データ (または「addbsi」) 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 (典型的には識別情報値、たとえばコア要素バージョン) 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・データに関する) 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。

30

40

## 【 0 1 5 4 】

本発明の典型的な実施形態は、ビットストリームによって示される連続するオーディオ・プログラム間の少なくとも一つの境界の正確かつ堅牢な判定を許容する効率的な仕方で、エンコードされたオーディオ・ビットストリーム内にプログラム境界メタデータを含める。典型的な実施形態は、異なるプログラムを示すビットストリームが、継ぎ合わされるビットストリーム的一方または両方を打ち切る (よって継ぎ合わせ前のビットストリームの少なくとも一方に含まれていたプログラム境界メタデータを破棄する) 仕方で一緒に (本発明のビットストリームを生成するよう) 継ぎ合わされている場合でも、正確なプログラム境界決定を許容するという意味で、プログラム境界の正確かつ堅牢な決定を許容す

50

る。

【0155】

典型的な実施形態では、本発明のビットストリームのフレームにおけるプログラム境界メタデータは、フレーム・カウントを示すプログラム境界フラグである。典型的には、このフラグは、現在フレーム（当該フラグを含んでいるフレーム）とプログラム境界（現在のオーディオ・プログラムの先頭または末尾）との間のフレーム数を示す。いくつかの好ましい実施形態では、プログラム境界フラグは対称的で効率的な仕方で、単一のプログラムを示す各ビットストリーム・セグメントの始まりおよび終わりに（すなわち、当該セグメントの始まりのあと何らかの所定数のフレーム内に生起するフレームにおいておよび当該セグメントの終わりの前の何らかの所定数のフレーム内に生起するフレームにおいて）挿入される。それにより、二つのそのようなビットストリームが連結される（それにより二つのプログラムのシーケンスを示すようになる）とき、プログラム境界メタデータは、二つのプログラムの間の境界の両方の側に（たとえば対称的に）存在することができる。

10

【0156】

プログラム境界フラグを、プログラムを示すビットストリームの全フレームに挿入することによって最大の堅牢性が達成できるが、これは典型的には、付随するデータ・レートの増大のため、実際的ではない。典型的な実施形態では、プログラム境界フラグは、エンコードされたオーディオ・ビットストリーム（これは一つのオーディオ・プログラムまたはオーディオ・プログラムのシーケンスを示しうる）のフレームの部分集合にのみ挿入され、境界フラグ挿入レートは、ビットストリームの（フラグが挿入される）各フレームの、該各フレームに最も近いプログラム境界からの離間の増大に対する非増加関数である。ここで、「境界フラグ挿入レート」とは、プログラム境界フラグを含む（プログラムを示す）フレームの数の、プログラム境界フラグを含まない（該プログラムを示す）フレームの数に対する、平均的な比を表わす。ここで、平均は、エンコードされたオーディオ・ビットストリームのある数（たとえば比較的少数）の連続するフレームにわたる移動平均である。

20

【0157】

（プログラム境界により近いビットストリーム中の位置において）境界フラグ挿入レートを増大させると、ビットストリームの送達のために必要とされるデータ・レートが増す。これを補償するために、挿入される各フラグのサイズ（ビット数）が、境界フラグ挿入レートが増すにつれて減少させられることが好ましい（それにより、Nは整数であるとして、ビットストリームのN番目のフレームにおけるプログラム境界フラグのサイズは、該N番目のフレームと最も近いプログラム境界との間の距離（フレーム数）の非増加関数である）。あるクラスの諸実施形態では、境界フラグ挿入レートは、最も近いプログラム境界からの（各フラグ挿入位置の）増大する距離の対数的に減少する関数であり、フラグの一つを含む各フラグ含有フレームについて、該フラグ含有フレーム中のフラグのサイズは、該フラグ含有フレームよりも前記最も近いプログラム境界により近くに位置するフレームにおける各フラグのサイズ以上である。典型的には、各フラグのサイズは、フラグの挿入位置から最も近いプログラム境界までのフレーム数の増加関数によって決定される。

30

【0158】

たとえば、図8および図9の実施形態を考える。ここで、フレーム番号（いちばん上の行）によって同定される各列が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を示す。ビットストリームは、図9の左側のフレーム番号「17」によって同定される列のすぐ左に現われる第一のプログラム境界（当該プログラムの始まりを示す）および図8の右側のフレーム番号「1」によって同定される列のすぐ右に現われる第二のプログラム境界（当該プログラムの終わりを示す）をもつオーディオ・プログラムを示す。図8に示される諸フレームに含まれるプログラム境界フラグは、現在フレームと第二のプログラム境界との間のフレーム数をカウントダウンする。図9に示される諸フレームに含まれるプログラム境界フラグは、現在フレームと第一のプログラム境界との間のフレーム数をカウントアップする。

40

50



## 【 0 1 5 9 】

図 8 および図 9 の実施形態では、プログラム境界フラグは、ビットストリームによって示されるオーディオ・プログラムの開始後、エンコードされたビットストリームの最初の X 個のフレームの  $2^N$  番目のフレームのそれぞれにおいて、および、ビットストリームによって示されるプログラムの終わりに最も近い（ビットストリームの最後の X 個のフレームの）  $2^N$  番目のフレームのそれぞれにおいてのみ示されている。ここで、プログラムは Y 個のフレームを含み、X は Y/2 以下の整数であり、N は 1 から  $\log_2 X$  までの範囲の正の整数である。このように、（図 8 および図 9 に示されるように）プログラム境界フラグがビットストリームの第二のフレーム（N=1）（プログラムの先頭に最も近いフラグ含有フレーム）、第四のフレーム（N=2）、第八のフレーム（N=3）などに挿入され、また、ビットストリームの終わりから八番目のフレーム、ビットストリームの終わりから四番目のフレームおよびビットストリームの終わりから二番目のフレーム（プログラムの末尾に最も近いフラグ含有フレーム）に挿入される。この例では、プログラムの先頭（または末尾）から  $2^N$  番目のフレームにおけるプログラム境界フラグは、図 8 および図 9 に示されるように、 $\log_2(2^{N+2})$  二進ビットを有する。こうして、プログラムの先頭（または末尾）から二番目のフレーム（N=1）におけるプログラム境界フラグは  $\log_2(2^{N+2}) = \log_2(2^3) = 3$  二進ビットを有し、プログラムの先頭（または末尾）から四番目のフレーム（N=2）におけるフラグは  $\log_2(2^{N+2}) = \log_2(2^4) = 4$  二進ビットを有する、などとなる。

## 【 0 1 6 0 】

図 8 および図 9 の例では、各プログラム境界フラグのフォーマットは次のとおり。各プログラム境界フラグは、先頭の「1」のビット、該先頭の「1」のビット後の「0」のビットのシーケンス（「0」のビットなしまたは一つまたは複数の連続する「0」のビット）および 2 ビットの後端符号（trailing code）からなる。ビットストリームの最後の X 個のフレーム（プログラム末尾に最も近い諸フレーム）におけるフラグについては、後端符号は、図 8 に示されるように、「11」である。ビットストリームの最初の X 個のフレーム（プログラム先頭に最も近い諸フレーム）におけるフラグについては、後端符号は、図 9 に示されるように、「10」である。このように、各フラグを読む（デコードする）ために、先頭の「1」のビットと後端符号との間の 0 の数が数えられる。後端符号が「11」であると識別される場合には、フラグは現在フレーム（そのフラグを含んでいるフレーム）とプログラムの終わりとの間に  $(2^{Z+1} - 1)$  個のフレームがあることを示す。ここで、Z は、このフラグの先頭の「1」のビットと後端符号との間の 0 の数である。このデコーダは、それぞれのそのようなフラグの最初と最後のビットを無視し、フラグの他の（中間の）諸ビットのシーケンスの逆を決定し（たとえば、中間ビットのシーケンスが「0001」であり、「1」のビットがシーケンスの最後のビットであれば、中間ビットの反転されたシーケンスは「1000」となり、「1」のビットが反転されたシーケンスの最初のビットになる）、中間ビットの反転されたシーケンスの二進値を、プログラムの終わりに対する現在フレーム（そのフラグが含まれているフレーム）のインデックスとして同定するよう効率的に実装されることができる。たとえば、中間ビットの反転されたシーケンスが「1000」であれば、この反転されたシーケンスは二進値  $2^4 = 16$  をもち、このフレームは、プログラムの終わりの前の 16 番目のフレームとして識別される（図 8 の、フレーム「0」を記述する列において示されるように）。

## 【 0 1 6 1 】

後端符号が「10」であると識別される場合には、フラグはプログラムの始まりと現在フレーム（そのフラグを含んでいるフレーム）との間に  $(2^{Z+1} - 1)$  個のフレームがあることを示す。ここで、Z は、このフラグの先頭の「1」のビットと後端符号との間の 0 の数である。このデコーダは、それぞれのそのようなフラグの最初と最後のビットを無視し、フラグの中間ビットのシーケンスの逆を決定し（たとえば、中間ビットのシーケンスが「0001」であり、「1」のビットがシーケンスの最後のビットであれば、中間ビットの反転されたシーケンスは「1000」となり、「1」のビットが反転されたシーケンスの最初のビットになる）、中間ビットの反転されたシーケンスの二進値を、プログラムの始まりに対する

10

20

30

40

50

現在フレーム（そのフラグが含まれているフレーム）のインデックスとして同定するよう効率的に実装されることができ。たとえば、中間ビットの反転されたシーケンスが「1000」であれば、この反転されたシーケンスは二進値 $2^4 = 16$ をもち、このフレームは、プログラムの始まりの後の16番目のフレームとして識別される（図9の、フレーム「32」を記述する列において示されるように）。

**【0162】**

図8および図9の例では、プログラム境界フラグは、ビットストリームによって示されるオーディオ・プログラムの開始後のエンコードされたビットストリームの最初のX個のフレームの $2^N$ 番目のフレームのそれぞれに、および、ビットストリームによって示されるオーディオ・プログラムの終了に最も近い（ビットストリームの最後のX個のフレームの） $2^N$ 番目のフレームのそれぞれに、存在するだけであり、プログラムはY個のフレームを有し、XはY/2以下の整数であり、Nは1から $\log_2 X$ までの範囲の正の整数である。プログラム境界フラグを含めることは、フラグなしでビットストリームを送信するのに必要されるビットレートに、1.875ビット/フレームの平均ビットレートを追加するだけである。

**【0163】**

ビットストリームがAC-3エンコードされたオーディオ・ビットストリームである図8および図9の実施形態の典型的な実装では、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。このように、そのような実施形態では、ある数のフレーム（「X」個のフレーム）だけプログラム境界から離間されるフレームにおけるプログラム境界フラグは、境界が、フラグ含有フレームの終了後32Xミリ秒（またはフラグ含有フレームの開始前32Xミリ秒）に現われることを示す。

**【0164】**

ビットストリームがE-AC-3エンコードされたオーディオ・ビットストリームである図8および図9の実施形態の典型的な実装では、ビットストリームの各フレームは、フレームが一ブロック、二ブロック、三ブロックまたは六ブロックのオーディオ・データのどれを含んでいるかに依存して、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。このように、そのような実施形態では、（各フレームが32ミリ秒のデジタル・オーディオを示すとすると）ある数のフレーム（「X」個のフレーム）だけプログラム境界から離間されるフレームにおけるプログラム境界フラグは、境界が、フラグ含有フレームの終了後32Xミリ秒（またはフラグ含有フレームの開始前32Xミリ秒）に現われることを示す。

**【0165】**

プログラム境界がオーディオ・ビットストリームのフレーム内に（すなわち、フレームの始まりまたは終わりに整列せずに）生起できるいくつかの実施形態では、ビットストリームのフレームに含まれるプログラム境界メタデータはプログラム境界フレーム・カウント（すなわち、フレーム・カウント含有フレームの先頭または末尾とプログラム境界との間の完全なフレームの数を示すメタデータ）と、オフセット値とを含む。オフセット値は、プログラム境界含有フレームの先頭または末尾と、プログラム境界含有フレーム内のプログラム境界の実際の位置との間のオフセット（典型的にはサンプル数）を示す。

**【0166】**

エンコードされたオーディオ・ビットストリームは、ビデオ・プログラムの対応するシーケンスのプログラム（サウンドトラック）のシーケンスを示していてもよく、そのようなオーディオ・プログラムの境界は、オーディオ・フレームのエッジではなくビデオ・フレームのエッジにおいて生起する傾向がある。また、いくつかのオーディオ・コーデック（たとえばE-AC-3の諸コーデック）は、ビデオ・フレームと整列していないオーディオ・

10

20

30

40

50

フレーム・サイズを使う。また、場合によっては、最初にエンコードされたオーディオ・ビットストリームがトランスコードを受けて、トランスコードされたビットストリームを生成し、最初にエンコードされたビットストリームはトランスコードされたビットストリームとは異なるフレーム・サイズをもつ。このため、プログラム境界（最初にエンコードされたビットストリームによって決定される）は、トランスコードされたビットストリームのフレーム境界に現われることは保証されない。たとえば、最初にエンコードされたビットストリーム（たとえば図10のビットストリーム「IEB」）が1536サンプル毎フレームのフレーム・サイズをもち、トランスコードされたビットストリーム（たとえば図10のビットストリーム「TB」）が1024サンプル毎フレームのフレーム・サイズをもつ場合、トランスコード・プロセスにより、異なるコーデックの異なるフレーム・サイズのため、実際のプログラム境界は、トランスコードされたビットストリームのフレーム境界にではなく、そのフレームのどこかに（たとえば図10に示されるように、トランスコードされたビットストリームのフレームに512サンプルはいったところに）現われることになりうる。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に含まれるプログラム境界メタデータがプログラム境界フレーム・カウントのほかにオフセット値を含む本発明の実施形態は、この段落に記された三つの場合（および他の場合）において有用である。

**【0167】**

図8および図9を参照して上述した実施形態は、エンコードされたビットストリームのフレームのいずれにもオフセット値（たとえばオフセット・フィールド）を含まない。この実施形態に対する諸変形では、オフセット値は、プログラム境界フラグを含む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に（たとえば、図8における0、8、12および14の番号を付されたフレームおよび図9における18、20、24および32の番号を付されたフレームに対応するフレームに）含まれる。

**【0168】**

あるクラスの諸実施形態では、（本発明のプログラム境界メタデータを含むエンコードされたビットストリームの各フレームにおける）データ構造が、当該フレームがプログラム境界フレーム・カウントのみを含んでいるか、プログラム境界フレーム・カウントおよびオフセット値の両方を含んでいるかを示す符号値を含む。たとえば、符号値は、単一ビット・フィールド（本稿では「offset\_exist〔オフセット存在〕」フィールドと称される）の値であってもよい。offset\_exist=0の値は、そのフレームにオフセット値が含まれないことを示してもよく、offset\_exist=1の値は、そのフレームにプログラム境界フレーム・カウントおよびオフセット値が含まれることを示してもよい。

**【0169】**

いくつかの実施形態では、AC-3または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少なくとも一つのフレームは、ビットストリームによって決定されるオーディオ・プログラムについてのLPSMおよびプログラム境界メタデータ（および任意的には他のメタデータも）を含むメタデータ・セグメントを含む。そのような各メタデータ・セグメント（これはビットストリームのaddbsiフィールドまたは補助データ・フィールドまたは余剰ビット・セグメントに含まれていてもよい）は、コア・ヘッダ（および任意的には追加的なコア要素も）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）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。

**【0170】**

ヘッダ（典型的には少なくとも一つの識別情報値、たとえば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、プログラム境界メタデータ（これは、プログラム境界フレーム・カウントと、そのフレームがプログラム境界フレーム・カウントのみを含むか、プログラム境界フレーム・カウントおよびオフセット値の両方を含むかを示す符号値（たとえば「offset\_exist」〔オフセット存在〕値）と、場合によっては、オフセット値とを含んでもよい）およびLPSM。該LPSMは下記を含んでもよい：

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。ダイアログ指示値（単数または複数）は、対応するオーディオ・データのチャンネルの任意の組み合わせまたは全部にダイアログが存在しているかどうかを示してもよい；

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；

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；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。

10

#### 【0171】

いくつかの実施形態では、LPSMペイロード・セグメントは、そのフレームがプログラム境界フレーム・カウントのみを含むか、プログラム境界フレーム・カウントおよびオフセット値の両方を含むかを示す符号値（たとえば「offset\_exist」〔オフセット存在〕値）を含む。たとえば、一つのそのような実施形態では、そのような符号値（たとえばoffset\_exist=1）が、そのフレームがプログラム境界フレーム・カウントおよびオフセット値を含むことを示すとき、LPSMペイロード・セグメントは、11ビット符号なし整数（すなわち、0から2048までの値をもつ）であり、合図されるフレーム境界（プログラム境界を含むフレームの境界）と実際のプログラム境界との間の追加的なオーディオ・サンプルの数を示すオフセット値を含んでいてもよい。プログラム境界フレーム・カウントが、プログラム境界含有フレームまでの（現在のフレーム・レートにおける）フレームの数を示す場合には、（LPSMペイロード・セグメントを含むフレームの始まりまたは終わりに対する）プログラム境界の（サンプル数単位での）精密な位置が：

20

$$S = (\text{frame\_counter} * \text{frame\_size}) + \text{offset}$$

として計算される。ここで、Sは（LPSMペイロード・セグメントを含んでいるフレームの始まりまたは終わりから）プログラム境界までのサンプル数であり、

「frame\_counter」は、プログラム境界フレーム・カウントによって示されるフレーム・カウントであり、「frame\_size」はフレーム当たりのサンプル数であり、「offset」は前記オフセット値によって示されるサンプル数である。

30

#### 【0172】

プログラム境界フラグの挿入レートが実際のプログラム境界の近くで増大するいくつかの実施形態は、フレームが、プログラム境界を含むフレームからある数（「Y」）のフレーム以下である場合には、そのフレームにはオフセット値は決して含まれないという規則を実装する。典型的には、Y=32である。この規則を（Y=32として）実装するE-AC-3エンコーダについては、エンコーダは、オーディオ・プログラムの最後の1秒には決してオフセット値を挿入しない。この場合、受領装置が、タイマーを維持し、それにより（プログラム境界含有フレームからYフレームより多く離れている当該エンコードされたビットストリームのフレームにおける、オフセット値を含むプログラム境界メタデータにตอบสนองして）自分自身のオフセット計算を実行することを受け持つ。

40

#### 【0173】

オーディオ・プログラムが対応するビデオ・プログラムのビデオ・フレームに「フレーム整列されている」ことがわかっているプログラム（たとえば、ドルビーEエンコードされたオーディオをもつ典型的な寄与フィード）については、オーディオ・プログラムを示すエンコードされたビットストリーム中のオフセット値を含めることは余計である。そこで、オフセット値は典型的にはそのようなエンコードされたビットストリームには含まれない。

#### 【0174】

図11を参照するに、次に、エンコードされたオーディオ・ビットストリームが本発明のオーディオ・ビットストリームの実施形態を生成するよう一緒に継ぎ合わされる場合を

50

考える。

【 0 1 7 5 】

図 1 1 のいちばん上のビットストリーム（「シナリオ 1」とラベル付けされている）は、プログラム境界メタデータ（プログラム境界フラグ F）を含む第一のオーディオ・プログラム（P1）全体と、それに続く、やはりプログラム境界メタデータ（プログラム境界フラグ F）を含む第二のオーディオ・プログラム（P2）全体とを示している。第一のプログラムの終わりの部分におけるプログラム境界フラグ（そのいくつかが図 1 1 に示されている）は、図 8 を参照して述べたものと同様であり、二つのプログラムの間の境界（すなわち、第二のプログラムの始まりにおける境界）の位置を決定する。第二のプログラムの始まりの部分におけるプログラム境界フラグ（そのいくつかが図 1 1 に示されている）は、図 9 を参照して述べたものと同様であり、やはり境界の位置を決定する。典型的な実施形態では、エンコーダまたはデコーダが、プログラム境界までカウントダウンするタイマー（第一のプログラム内のフラグによって較正される）を実装し、同じタイマー（第二のプログラム内のフラグによって較正される）が同じプログラム境界からカウントアップする。図 1 1 のシナリオ 1 における境界タイマー・グラフによって示されるように、そのようなタイマーのカウントダウン（第一のプログラム内のフラグによって較正される）は境界において 0 に達し、タイマーのカウントアップ（第二のプログラム内のフラグによって較正される）は境界の同じ位置を参照する。

10

【 0 1 7 6 】

図 1 1 の上から二番目のビットストリーム（「シナリオ 2」とラベル付けされている）は、プログラム境界メタデータ（プログラム境界フラグ F）を含む第一のオーディオ・プログラム（P1）全体と、それに続く、プログラム境界メタデータを含まない第二のオーディオ・プログラム（P2）全体とを示している。第一のプログラムの終わりの部分におけるプログラム境界フラグ（そのいくつかが図 1 1 に示されている）は、図 8 を参照して述べたものと同様であり、シナリオ 1 と同様に、二つのプログラムの間の境界（すなわち、第二のプログラムの始まりにおける境界）の位置を決定する。典型的な実施形態では、エンコーダまたはデコーダが、プログラム境界までカウントダウンするタイマー（第一のプログラム内のフラグによって較正される）を実装し、同じタイマーが（さらに較正されることなく）該プログラム境界からカウントアップすることを続ける（図 1 1 のシナリオ 2 における境界タイマー・グラフによって示されるように）。

20

30

【 0 1 7 7 】

図 1 1 の上から三番目のビットストリーム（「シナリオ 3」とラベル付けされている）は、プログラム境界メタデータ（プログラム境界フラグ F）を含む打ち切りされた第一のオーディオ・プログラム（P1）であって、やはりプログラム境界メタデータ（プログラム境界フラグ F）を含む第二のオーディオ・プログラム（P2）全体と継ぎ合わされているものを示している。継ぎ合わせは、第一のプログラムの最後の「N」個のフレームを除去している。第二のプログラムの始まりの部分におけるプログラム境界フラグ（そのいくつかが図 1 1 に示されている）は、図 9 を参照して述べたものと同様であり、打ち切りされた第一のプログラムと完全な第二のプログラムとの間の境界（継ぎ〔スプライス〕）の位置を決定する。典型的な実施形態では、エンコーダまたはデコーダが、打ち切りされていない第一のプログラムの終わりまでカウントダウンするタイマー（第一のプログラム内のフラグによって較正される）を実装し、同じタイマー（第二のプログラム内のフラグによって較正される）が第二のプログラムの先頭からカウントアップする。第二のプログラムの先頭が、シナリオ 3 におけるプログラム境界である。図 1 1 のシナリオ 3 における境界タイマー・グラフによって示されるように、そのようなタイマーのカウントダウン（第一のプログラム内のプログラム境界メタデータによって較正される）は、（第一のプログラム内のプログラム境界メタデータにตอบสนองして）0 に達する前に（第二のプログラム内のプログラム境界メタデータにตอบสนองして）リセットされる。このように、（継ぎ合わせによる）第一のプログラムの打ち切りはタイマーが、第一のプログラム内のプログラム境界メタデータだけにตอบสนองして（すなわち、それによる較正のもとで）、打ち切りされた

40

50

第一のプログラムと第二のプログラムの先頭との間のプログラム境界を同定することを妨げるが、第二のプログラムにおけるプログラム・メタデータがタイマーをリセットし、それによりリセットされたタイマーが、打ち切りされた第一のプログラムと第二のプログラムの先頭との間のプログラム境界の位置を（リセットされたタイマーの「0」カウントに対応する位置として）正しく示す。

【0178】

第四のビットストリーム（「シナリオ4」とラベル付けされている）は、プログラム境界メタデータ（プログラム境界フラグF）を含む打ち切りされた第一のオーディオ・プログラム（P1）と、プログラム境界メタデータ（プログラム境界フラグF）を含む打ち切りされた第二のオーディオ・プログラム（P2）であって、第一のオーディオ・プログラムの一部（打ち切りされていない部分）と継ぎ合わされているものとを示している。（打ち切り前の）第二のプログラム全体の始まりの部分におけるプログラム境界フラグ（そのいくつかが図11に示されている）は、図9を参照して述べたものと同様であり、（打ち切り前の）第一のプログラム全体の終わりの部分におけるプログラム境界フラグ（そのいくつかが図11に示されている）は、図8を参照して述べたものと同様である。継ぎ合わせにより、第一のプログラムの最後の「N」個のフレームが（よって継ぎの前に含まれていたプログラム境界フラグの一部が）ならびに第二のプログラムの最初の「M」個のフレームが（よって継ぎの前に含まれていたプログラム境界フラグの一部が）除去されている。典型的な実施形態では、エンコーダまたはデコーダが、打ち切りされていない第一のプログラムの終わりに向けてカウントダウンするタイマー（打ち切りされた第一のプログラム内のフラグによって較正される）を実装し、同じタイマー（打ち切りされた第二のプログラム内のフラグによって較正される）が打ち切りされていない第二のプログラムの先頭からカウントアップする。図11のシナリオ4における境界タイマー・グラフによって示されるように、そのようなタイマーのカウントダウン（第一のプログラム内のプログラム境界メタデータによって較正される）は、（第一のプログラム内のプログラム境界メタデータに回答して）0に達する前に（第二のプログラム内のプログラム境界メタデータに回答して）リセットされる。（継ぎ合わせによる）第一のプログラムの打ち切りは、タイマーが、第一のプログラム内のプログラム境界メタデータだけに回答して（すなわち、それによる較正のもとで）、打ち切りされた第一のプログラムと打ち切りされた第二のプログラムの先頭との間のプログラム境界を同定することを妨げる。しかしながら、リセットされたタイマーは、打ち切りされた第一のプログラムの終わりと打ち切りされた第二のプログラムの始まりとの間のプログラム境界の位置を正しく示さない。このように、継ぎ合わされるビットストリーム両方の打ち切りは、両者の間の境界の正確な決定を妨げることがある。

【0179】

本発明の実施形態は、ハードウェア、ファームウェアまたはソフトウェアまたは両者の組み合わせにおいて（たとえばプログラム可能な論理アレイとして）実装されてもよい。特に断わりのない限り、本発明の一部として含まれるアルゴリズムまたはプロセスは、いかなる特定のコンピュータまたは他の装置にも本来的に関係していない。特に、さまざまな汎用機械が、本願の教示に従って書かれたプログラムとともに使用されてもよく、あるいは必要とされる方法ステップを実行するためにより特化した装置（たとえば集積回路）を構築することがより便利であることがある。このように、本発明は、一つまたは複数のプログラム可能なコンピュータ・システム（たとえば、図1の諸要素または図2のエンコーダ100（またはその要素）または図3のデコーダ200（またはその要素）または図3の後処理器（またはその要素）のうちの任意のものの実装）上で実行される一つまたは複数のコンピュータ・プログラムにおいて実装されてもよい。各コンピュータ・システムは、少なくとも一つのプロセッサ、少なくとも一つのデータ記憶システム（揮発性および不揮発性メモリおよび/または記憶要素を含む）、少なくとも一つの入力装置またはポートおよび少なくとも一つの出力装置またはポートを有する。本稿に記載される機能を実行し、出力情報を生成するようプログラム・コードが入力データに適用される。出力情報は

10

20

30

40

50

、既知の仕方で一つまたは複数の出力装置に適用される。

【0180】

そのような各プログラムは、コンピュータ・システムと通信するためにいかなる所望されるコンピュータ言語（機械、アセンブリまたは高水準手続き型、論理的またはオブジェクト指向のプログラミング言語を含む）において実装されてもよい。いずれの場合にも、言語はコンパイルされる言語でもインタープリットされる言語でもよい。

【0181】

たとえば、コンピュータ・ソフトウェア命令のシーケンスによって実装されるとき、本発明の実施形態のさまざまな機能および段階は、好適なデジタル信号処理ハードウェアにおいて実行されるマルチスレッド式のソフトウェア命令シーケンスによって実装されてもよく、その場合、実施形態のさまざまな装置、段階および機能は、ソフトウェア命令の諸部分に対応してもよい。

【0182】

そのような各コンピュータ・プログラムは好ましくは、汎用または専用のプログラム可能なコンピュータによって読み取り可能な記憶媒体またはデバイス（たとえば半導体メモリまたはメディアまたは磁気式もしくは光学式メディア）に記憶されるまたはダウンロードされ、記憶媒体またはデバイスがコンピュータ・システムによって読まれたときに、本稿に記載される手順を実行するようコンピュータを構成するまたは動作させる。本発明のシステムは、コンピュータ・プログラムをもって構成された（すなわちコンピュータ・プログラムを記憶している）コンピュータ可読記憶媒体として実装されてもよく、そのように構成された記憶媒体はコンピュータ・システムに、本稿に記載される機能を実行するよう特定のあらかじめ定義された仕方で動作させる。

【0183】

本発明のいくつかの実施形態を記述してきたが、本発明の精神および範囲から外れることなくさまざまな修正がなしうことは理解されるであろう。上記の教示に照らして、本発明の数多くの修正および変形が可能である。付属の請求項の範囲内で、本発明が、本稿で具体的に記載される以外の仕方で実施されてもよいことは理解される。

いくつかの態様を記載しておく。

〔態様1〕

オーディオ処理ユニットであって：

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記憶するバッファ・メモリであって、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・データおよびメタデータ・コンテナを含み、前記メタデータ・コンテナはヘッダ、一つまたは複数のメタデータ・ペイロードおよび保護データを含む、バッファ・メモリと；

前記バッファ・メモリに結合された、前記オーディオ・データをデコードするオーディオ・デコーダと；

前記オーディオ・デコーダに結合されるか前記オーディオ・デコーダと統合されている、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する、パーサとを有しており、

前記ヘッダは、前記メタデータ・コンテナの先頭を同定する同期語を含み、前記一つまたは複数のメタデータ・ペイロードは、前記オーディオ・データに関連付けられたオーディオ・プログラムを記述し、前記保護データは、前記一つまたは複数のメタデータ・ペイロードのあとに位置し、前記保護データは、前記メタデータ・コンテナおよび該メタデータ・コンテナ内の前記一つまたは複数のペイロードの完全性を検証するために使用できる

オーディオ処理ユニット。

〔態様2〕

前記メタデータ・コンテナが、スキップ・フィールド、補助データ・フィールド、adbsiフィールドおよびそれらの組み合わせからなる群から選択されるAC-3またはE-AC-3のリ

10

20

30

40

50

ザーブされたデータ・スペースに格納される、態様 1 記載のオーディオ処理ユニット。

〔態様 3〕

前記一つまたは複数のメタデータ・ペイロードが、連続するオーディオ・プログラムの間の少なくとも一つの境界を示すメタデータを含む、態様 1 または 2 記載のオーディオ処理ユニット。

〔態様 4〕

前記一つまたは複数のメタデータ・ペイロードが、オーディオ・プログラムの測定されたラウドネスを示すデータを含むプログラム・ラウドネス・ペイロードを含む、態様 1 または 2 記載のオーディオ処理ユニット。

〔態様 5〕

前記プログラム・ラウドネス・ペイロードが、オーディオ・チャンネルが話されたダイアログを含むかどうかを示すフィールドを含む、態様 4 記載のオーディオ処理ユニット。

〔態様 6〕

前記プログラム・ラウドネス・ペイロードが、前記プログラム・ラウドネス・ペイロードに含まれるラウドネス・データを生成するために使われたラウドネス測定方法を示すフィールドを含む、態様 4 記載のオーディオ処理ユニット。

〔態様 7〕

前記プログラム・ラウドネス・ペイロードが、オーディオ・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かどうかを示すフィールドを含む、態様 4 記載のオーディオ処理ユニット。

〔態様 8〕

前記プログラム・ラウドネス・ペイロードが、オーディオ・プログラムのラウドネスが、無限先読みまたはファイル・ベースのラウドネス補正プロセスを使って補正されているかどうかを示すフィールドを含む、態様 4 記載のオーディオ処理ユニット。

〔態様 9〕

前記プログラム・ラウドネス・ペイロードが、ダイナミックレンジ圧縮に帰着できるいかなる利得調整もなしにオーディオ・プログラムの統合されたラウドネスを示すフィールドを含む、態様 4 記載のオーディオ処理ユニット。

〔態様 10〕

前記プログラム・ラウドネス・ペイロードが、ダイアログ正規化 (dialog normalization) に帰着できるいかなる利得調整もなしにオーディオ・プログラムの統合されたラウドネスを示すフィールドを含む、態様 4 記載のオーディオ処理ユニット。

〔態様 11〕

前記プログラム・ラウドネス・ペイロードを使っ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、態様 4 記載のオーディオ処理ユニット。

〔態様 12〕

前記エンコードされたオーディオ・ビットストリームが 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る、態様 1 ないし 11 のうちいずれか一項記載のオーディオ処理ユニット。

〔態様 13〕

前記エンコードされたオーディオ・ビットストリームから前記プログラム・ラウドネス・ペイロードを抽出し、前記プログラム・ラウドネス・ペイロードを認証または有効確認するよう構成されている、態様 4 ないし 11 のうちいずれか一項記載のオーディオ処理ユニット。

〔態様 14〕

前記一つまたは複数のメタデータ・ペイロードがそれぞれ、一意的なペイロード識別子を含み、前記一意的なペイロード識別子が各メタデータ・ペイロードの先頭に位置される、態様 1 ないし 13 のうちいずれか一項記載のオーディオ処理ユニット。

〔態様 15〕

前記同期語が値 0x5838 をもつ 16 ビット同期語である、態様 1 ないし 13 のうちいずれか

10

20

30

40

50



一項記載のオーディオ処理ユニット。

〔態様 16〕

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方法であって：

一つまたは複数のフレームにセグメント分割されている、エンコードされたオーディオ・ビットストリームを受領する段階と；

前記エンコードされたオーディオ・ビットストリームからオーディオ・データおよびメタデータのコンテナを抽出する段階であって、前記メタデータのコンテナは、ヘッダと、それに続く一つまたは複数のメタデータ・ペイロードと、それに続く保護データとを含む、段階と；

前記コンテナおよび前記一つまたは複数のメタデータ・ペイロードの完全性を、前記保護データの使用を通じて検証する段階であって、

前記一つまたは複数のメタデータ・ペイロードは、前記オーディオ・データに関連するオーディオ・プログラムの測定されたラウドネスを示すデータを含むプログラム・ラウドネス・ペイロードを含む、

方法。

〔態様 17〕

前記エンコードされたオーディオ・ビットストリームが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る、態様 16 記載の方法。

〔態様 18〕

前記エンコードされたオーディオ・ビットストリームから抽出されたオーディオ・データに対して、前記プログラム・ラウドネス・ペイロードを使って適応的なラウドネス処理を実行する段階をさらに含む、態様 16 記載の方法。

〔態様 19〕

前記コンテナが、スキップ・フィールド、補助データ・フィールド、addbsiフィールドおよびそれらの組み合わせからなる群から選択されるAC-3またはE-AC-3のリザーブされたデータ・スペースに位置しておりそこから抽出される、態様 16 記載の方法。

〔態様 20〕

前記プログラム・ラウドネス・ペイロードが、オーディオ・チャンネルが話されたダイアログを含むかどうかを示すフィールドを含む、態様 16 記載の方法。

〔態様 21〕

前記プログラム・ラウドネス・ペイロードが、前記プログラム・ラウドネス・ペイロードに含まれるラウドネス・データを生成するために使われたラウドネス測定方法を示すフィールドを含む、態様 16 記載の方法。

〔態様 22〕

前記プログラム・ラウドネス・ペイロードが、オーディオ・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かどうかを示すフィールドを含む、態様 16 記載の方法。

〔態様 23〕

前記プログラム・ラウドネス・ペイロードが、オーディオ・プログラムのラウドネスが、無限先読みまたはファイル・ベースのラウドネス補正プロセスを使って補正されているかどうかを示すフィールドを含む、態様 16 記載の方法。

〔態様 24〕

前記プログラム・ラウドネス・ペイロードが、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なしにオーディオ・プログラムの統合されたラウドネスを示すフィールドを含む、態様 16 記載の方法。

〔態様 25〕

前記プログラム・ラウドネス・ペイロードが、ダイアログ正規化(dialog normalization)に帰着できるいかなる利得調整もなしにオーディオ・プログラムの統合されたラウドネスを示すフィールドを含む、態様 16 記載の方法。

〔態様 26〕

10

20

30

40

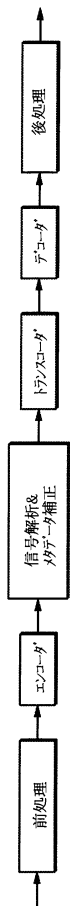
50

前記メタデータのコンテナが、連続するオーディオ・プログラムの間の少なくとも一つの境界を示すメタデータを含む、態様 16 記載の方法。

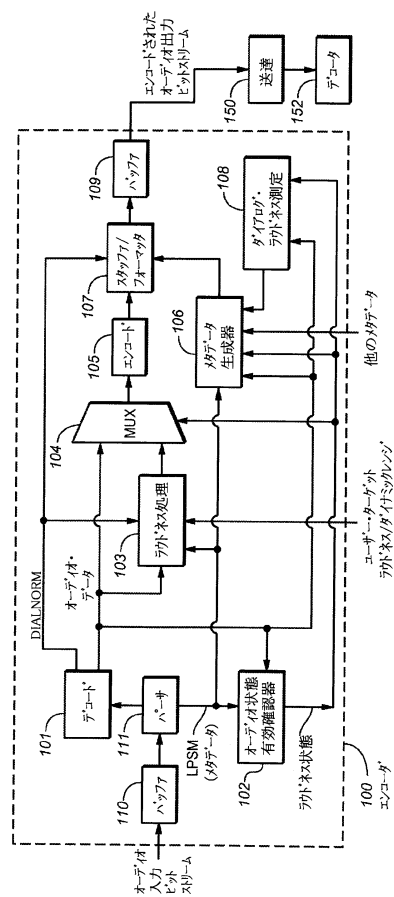
〔態様 27〕

前記メタデータのコンテナが、フレームの一つまたは複数のスキップ・フィールドまたは余剰ビット・セグメントに格納されている、態様 16 記載の方法。

【図 1】



【図 2】





【図 8】

フレーム#	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
境界までのフレーム数	-16	-15	-14	-13	-12	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1
プログラム境界 フラグ内容	100011	0	0	0	0	0	0	0	10011	0	0	0	1011	0	111	0
存在	1	0	0	0	0	0	0	0	1	0	0	0	1	0	1	0
フレームカウンタ	0001→16								001→8				01→4			1→2
カウンタアップ(+) ダウン(-)	1								1						1	

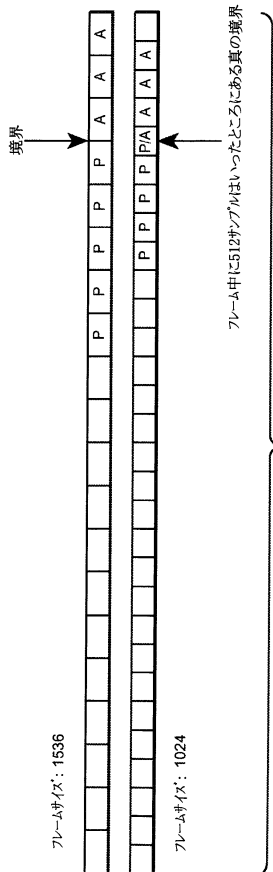
(プログラム境界までカウンタが増えるプログラム境界フラグのシンタックス)

【図 9】

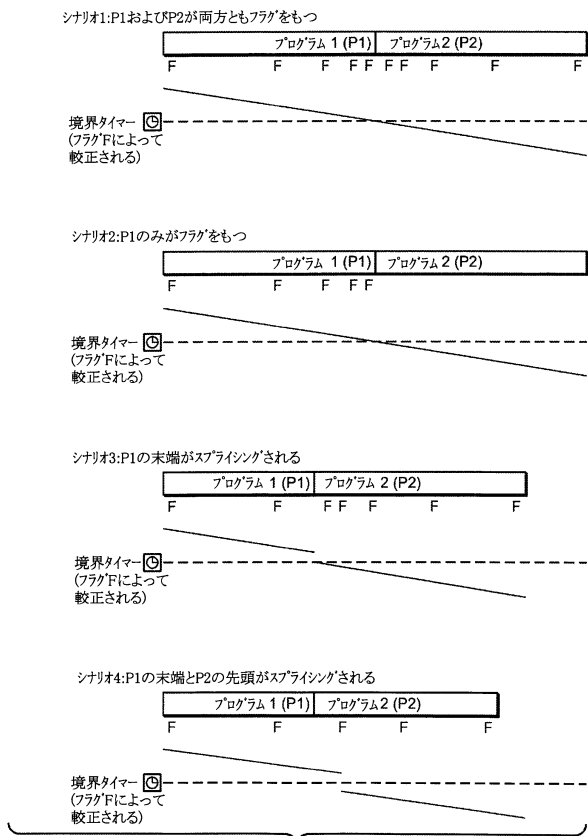
フレーム#	B	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
境界までのフレーム数	B	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
プログラム境界 フラグ内容	B	0	110	0	1010	0	0	0	10010	0	0	0	0	0	0	0	10010
存在	B	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	
フレームカウンタ	B	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0001→16
カウンタアップ(+) ダウン(-)	B	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0

(プログラムの先頭についてカウンタが増えるプログラム境界フラグのシンタックス)

【図 10】



【図 11】



---

フロントページの続き

早期審査対象出願

- (72)発明者 グラント, マイケル  
アメリカ合衆国, カリフォルニア州 94103 - 4813, 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 - 4813, 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 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 94103 - 4813, 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 ワード, マイケル  
アメリカ合衆国, カリフォルニア 94103 - 4813, 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

合議体

審判長 森川 幸俊

審判官 井上 信一

審判官 関谷 隆一

(56)参考文献 特開2008 - 536193 (JP, A)

(58)調査した分野(Int.Cl., DB名)

G10L 19/00