

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2020-8878
(P2020-8878A)

(43) 公開日 令和2年1月16日(2020.1.16)

(51) Int.Cl.	F I	テーマコード (参考)
G10K 15/02 (2006.01)	G10K 15/02	5D162
H04S 1/00 (2006.01)	H04S 1/00	5D208
H03G 9/02 (2006.01)	H03G 9/02 025	5J030

審査請求 有 請求項の数 18 O L 外国語出願 (全 18 頁)

(21) 出願番号	特願2019-173808 (P2019-173808)	(71) 出願人	503260918 アップル インコーポレイテッド Apple Inc. アメリカ合衆国 95014 カリフォルニア州 クパチーノ アップル パーク ウェイ ワン One Apple Park Way, Cupertino, California 95014, U. S. A.
(22) 出願日	令和1年9月25日 (2019. 9. 25)	(74) 代理人	100094569 弁理士 田中 伸一郎
(62) 分割の表示	特願2018-511198 (P2018-511198) の分割	(74) 代理人	100103610 弁理士 ▲吉▼田 和彦
原出願日	平成28年9月26日 (2016. 9. 26)	(74) 代理人	100067013 弁理士 大塚 文昭
(31) 優先権主張番号	62/235, 293		
(32) 優先日	平成27年9月30日 (2015. 9. 30)		
(33) 優先権主張国・地域又は機関	米国 (US)		
(31) 優先権主張番号	15/275, 162		
(32) 優先日	平成28年9月23日 (2016. 9. 23)		
(33) 優先権主張国・地域又は機関	米国 (US)		

最終頁に続く

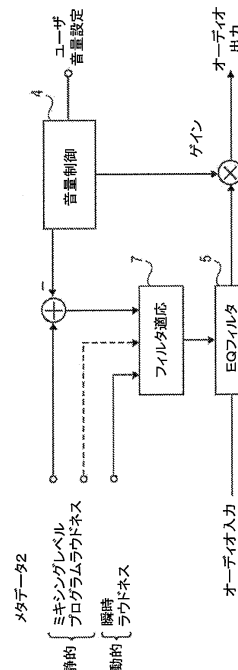
(54) 【発明の名称】 符号化されたオーディオメタデータに基づくラウドネス等化及びDRC中の動的等化

(57) 【要約】

【課題】再生デバイスにおいてメタデータベースのラウドネス等化(EQ)のために再生側の複雑さを軽減し、遅延を少なくし、品質を向上させる。

【解決手段】オーディオコンテンツの瞬時ラウドネス値を含むメタデータを使用した、再生システム内の受信オーディオコンテンツの動的ラウドネス等化。再生レベルは、再生システムのユーザ音量設定から導出され、オーディオコンテンツに割り当てられたミキシングレベルと比較される。瞬時ラウドネス値及び再生レベルの割り当てミキシングレベルとの比較に基づいて、フィルタリングされたオーディオコンテンツでスピーカを駆動する前に、オーディオコンテンツをフィルタリングしている等化フィルタを定義するパラメータが計算される。

【選択図】 図1



【特許請求の範囲】**【請求項 1】**

ダイナミックレンジ制御を適用している再生システムにおけるオーディオ等化のための方法であって、

a) オーディオコンテンツ、及び前記オーディオコンテンツのメタデータを受信するステップであって、

前記メタデータは、前記オーディオコンテンツについて計算された複数のダイナミックレンジ制御(DRC)ゲイン値を含む、前記受信するステップと、

b) 等化フィルタを定義する複数のパラメータを計算するステップであって、前記パラメータは、前記メタデータにおいて受信された前記複数のDRCゲイン値に基づいて計算される、ステップと、

c) 前記オーディオコンテンツを前記等化フィルタによってフィルタリングするステップと、

d) 前記オーディオコンテンツに前記複数のDRCゲイン値を適用することによって前記オーディオコンテンツのダイナミックレンジを調整するステップと、

e) 前記c)のステップにおいてフィルタリングされ、かつ前記d)のステップで調整され、かつダイナミックレンジとして前記オーディオコンテンツを提供して、前記再生システムにおけるスピーカを駆動するステップ、を含む方法。

【請求項 2】

前記等化フィルタを定義する前記パラメータは前記メタデータにおいて受信されたDRCゲイン値に基づいて直接計算される、請求項1に記載の方法。

【請求項 3】

前記メタデータにおいて受信された前記複数のDRCゲイン値を処理して複数のDRCゲイン調整値を計算するステップをさらに含み、前記複数のDRCゲイン値を前記オーディオコンテンツに適用するステップは、前記DRCゲイン調整値を前記オーディオコンテンツに適用するステップを含む、請求項2に記載の方法。

【請求項 4】

前記メタデータは、前記等化フィルタを定義するための以下の1つ以上を指定する静的フィルタ構成データを含み：a) タイプ、例えば低周波数カット若しくは低周波数ブースト、又は高周波数カット若しくは高周波数ブースト、b) フィルタ強度、及びc) 適応又は有効周波数帯域、

前記静的フィルタ構成データに従って構成された前記等化フィルタは、前記オーディオコンテンツがそれを通過している間に、前記パラメータによって動的に修正される、請求項2に記載の方法。

【請求項 5】

前記メタデータにおいて受信された前記複数のDRCゲイン値を処理して複数のDRCゲイン調整値を計算するステップをさらに含み、前記複数のDRCゲイン値を前記オーディオコンテンツに適用するステップは、前記DRCゲイン調整値を前記オーディオコンテンツに適用するステップを含む、請求項1に記載の方法。

【請求項 6】

前記メタデータは、静的フィルタ構成データは以下の1つ以上を指定する静的フィルタ構成データを含み：a) タイプ、例えば低周波数カット若しくは低周波数ブースト、又は高周波数カット若しくは高周波数ブースト、b) フィルタ強度、及びc) 適応又は有効周波数帯域、

前記静的フィルタ構成データに従って構成された前記等化フィルタは、前記オーディオコンテンツがそれを通過している間に、前記パラメータによって動的に修正される、請求項1に記載の方法。

【請求項 7】

前記等化フィルタを定義する前記複数のパラメータを計算するステップは、ミキシングレベルまたは再生レベルを使用しない、請求項1に記載の方法。

10

20

30

40

50

【請求項 8】

前記等下フィルタによって前記オーディオコンテンツをフィルタリングするステップは前記オーディオコンテンツのバックグラウンドノイズ音のバンドパスフィルタリングを行って前記バックグラウンドノイズ音の増幅を回避するステップを含む、請求項 1 に記載の方法。

【請求項 9】

前記複数の DRC ゲイン調整値を前記オーディオコンテンツに適用するステップは、前記オーディオコンテンツの同じゲインが全ての周波数成分に適用されるシングルバンド DRC を適用するステップを含む、請求項 8 に記載の方法。

【請求項 10】

オーディオ再生システムのプロセッサが実行する命令を記憶した非一時的機械可読媒体を備える製造品であって、

前記オーディオ再生システムのプロセッサが実行する命令が、

a) オーディオコンテンツ、及び前記オーディオコンテンツのメタデータを受信すること、であって、前記メタデータが、前記オーディオコンテンツについて計算された複数のダイナミックレンジ制御 (DRC) ゲイン値を含む、こと、

b) 等化フィルタを定義する複数のパラメータを計算することであって、前記パラメータが、前記メタデータにおいて受信された前記複数の DRC ゲイン値に基づいて計算されること、

c) 前記オーディオコンテンツを前記等化フィルタによってフィルタリングすること、

d) 前記オーディオコンテンツに前記複数の DRC ゲイン値を適用することによって前記オーディオコンテンツのダイナミックレンジを調整すること、及び、

e) 前記 c) においてフィルタリングされ、かつ前記 d) で調整され、かつダイナミックレンジとして前記オーディオコンテンツを提供して、前記再生システムにおけるスピーカを駆動すること、を含む製造品。

【請求項 11】

前記非一時的機械可読媒体は、その中に、前記プロセッサが、前記メタデータにおいて受信した前記 DRC ゲイン値に基づいて直接前記等下フィルタを定義する前記パラメータを計算することを実行する、命令を記憶している、請求項 10 に記載の製造品。

【請求項 12】

前記非一時的機械可読媒体は、その中に、前記プロセッサが、前記メタデータにおいて受信した前記複数の DRC ゲイン値を処理して、複数の DRC ゲイン調整値を計算することを実行する、命令を記憶しており、前記複数の DRC ゲイン値を前記オーディオコンテンツに適用することが、前記 DRC ゲイン調整値を前記オーディオコンテンツに適用することを含む、請求項 11 に記載の製造品。

【請求項 13】

前記メタデータは、前記等化フィルタを定義するための以下の 1 つ以上を指定する静的フィルタ構成データを含み： a) タイプ、例えば低周波数カット若しくは低周波数ブースト、又は高周波数カット若しくは高周波数ブースト、b) フィルタ強度、及び c) 適応又は有効周波数帯域、

前記静的フィルタ構成データに従って構成された前記等化フィルタは、前記オーディオコンテンツがそれを通過している間に、前記パラメータによって動的に修正される、請求項 11 に記載の製造品。

【請求項 14】

前記非一時的機械可読媒体は、その中に、前記プロセッサが、前記メタデータにおいて受信した前記複数の DRC ゲイン値を処理して、複数の DRC ゲイン調整値を計算することを実行する、命令を記憶しており、前記複数の DRC ゲイン値を前記オーディオコンテンツに適用することが、前記 DRC ゲイン調整値を前記オーディオコンテンツに適用することを含む、請求項 10 に記載の製造品。

【請求項 15】

前記メタデータは、前記等化フィルタを定義するための以下の1つ以上を指定する静的フィルタ構成データを含み：a) タイプ、例えば低周波数カット若しくは低周波数ブースト、又は高周波数カット若しくは高周波数ブースト、b) フィルタ強度、及びc) 適応又は有効周波数帯域、

前記静的フィルタ構成データに従って構成された前記等化フィルタは、前記オーディオコンテンツがそれを通過している間に、前記パラメータによって動的に修正される、請求項10に記載の製造品。

【請求項16】

前記プロセッサがミキシングレベル又は再生レベルを使用することなく、前記等下フィルタを定義する前記複数のパラメータを計算する、請求項10に記載の製造品。

10

【請求項17】

前記非一時的機械可読媒体は、その中に、前記プロセッサが、前記オーディオコンテンツの中の前記バックグラウンドノイズ音の増幅を回避するように前記オーディオコンテンツをフィルタリングするバンドパスフィルタとして前記等下フィルタを構成することを実行する、命令を記憶している、請求項10に記載の製造品。

【請求項18】

前記非一時的機械可読媒体は、その中に、前記プロセッサが、同じゲインが全ての周波数成分に適用されるシングルバンドDRCを適用することによって、前記複数のDRCゲイン値を前記オーディオコンテンツに適用することを実行する、命令を記憶している、請求項10に記載の製造品。

20

【発明の詳細な説明】

【技術分野】

【0001】

本出願は、2015年9月30日出願の、米国仮特許出願第62/235,293号の出願日遡及の特典を主張する。

【0002】

本発明の実施形態は、聴き手の経験を改善するために、メタデータに関連付けられたデジタルオーディオコンテンツの再生側でのデジタルオーディオ信号処理に関する。他の実施形態も記載される。

【背景技術】

30

【0003】

映画の音楽やサウンドトラックなどのオーディオコンテンツは、通常、特定の再生レベル（例えば、聴き手の位置で、オーディオコンテンツのプロデューサーが意図したものと同じ音圧レベルを得るために、その初期又は復号化された形式からスピーカによって音声に変換されるときまでの間の、再生中のオーディオ信号に適用される「全体ゲイン」。）を前提として生成される。異なる再生レベルが使用される場合、コンテンツはより大音量で又はより小音量で聞こえるだけでなく、異なるトーン特性を有するようにも感じられ得る。心理音響学から知られている効果は、低い周波数においては、再生レベルに対するラウドネス知覚は非線形的に増加することである。この効果は、知覚された等ラウドネス曲線によって、並びに再生レベル及び信号特性の関数としての知覚されたラウドネスの測定によって、定量化され得る。一般的に、コンテンツがプロデューサーが意図したレベルより低いレベルで再生されるとき、他の周波数と比較して低周波数成分の部分的な損失が報告されている。従来、ラウドネス等化は、再生音量の設定に応じて低周波数帯域をブーストする適応フィルタによって行われた。多くの古いオーディオレシーバは、そのように動作する「ラウドネス」ボタンを有する。

40

【発明の概要】

【0004】

メタデータベースのラウドネス等化(EQ)のためのいくつかのスキームを以下に記載する。そのいくつかは、例えば、再生側の複雑さが軽減され、遅延が少なくなり、品質が向上するなどの、1つ以上利点を有することができる。いくつかの品質改善は、符号化側

50

でのオフライン処理によるものであり、再生デバイスにおけるリアルタイム処理の制限及び低遅延要件によって限定されない。本明細書に記載されたメタデータベースのアプローチは、既存のMPEG-D DRC standard, ISO/IEC, 「Information technology - MPEG audio technologies - Part 4: Dynamic range control,」ISO/IEC 23003-4:2015、にシームレスに統合されてもよく、ダイナミックレンジ制御と共同する。

【0005】

DRCプロセス内に動的EQを提供するアプローチも記載される。それは、マルチバンドDRCと同様のEQを達成することができるが、より少数のバンド又は単一バンドのみのDRCである。動的EQはメタデータによって制御されてもよく、一般的なMPEG-D DRC規格に統合され得る。

10

【0006】

上述の概要は、本発明のすべての態様の網羅的なリストを含んでいない。本発明は、上でまとめた様々な態様のすべての適切な組み合わせによって実施できるすべてのシステム及び方法、並びに以下の「発明を実施するための形態」で開示されるもの、特に本出願と共に提出された請求項に指摘されるものを含むと考えられる。そのような組み合わせは、上記概要に具体的に記載されていない特定の利点を有する。

【0007】

本発明の実施形態は、同様の参照符号が同様の要素を示す添付の図面に例として限定としてではなく示されている。本開示における本発明の「一」実施形態又は「一つの」実施形態への言及は、必ずしも同じ実施形態に対するものではなく、それらは、少なくとも1つの実施形態を意味することに留意されたい。また、簡潔さ及び図の総数の低減のために、所与の図を使用して、本発明の複数の実施形態の特徴を例示することができ、図中のすべての要素が所与の実施形態に対して必要とされなくてもよい。

20

【図面の簡単な説明】

【0008】

【図1】オーディオコンテンツの受信メタデータから導出された瞬時ラウドネス値に基づく復号化側ラウドネス等化器のブロック図である。

【図2】瞬時ラウドネス値を含むメタデータを生成するための制作又は符号化側システムのブロック図である。

30

【図3】DRCゲイン値を計算するために符号化側で、及びその逆の形で復号化側で、使用され得るいくつかの例示的なDRC特性を示す。

【図4】ラウドネス等化フィルタの適応のために、復号化側で瞬時ラウドネス値を生成するために逆DRC特性がどのように使用されるかを示す図である。

【図5】ダイナミックレンジ圧縮及びラウドネス等化がオーディオコンテンツに適用されている復号化側を示す。

【図6】オーディオコンテンツにダイナミックレンジ圧縮と動的等化が適用されている復号化側を示す。

【図7】復号化側におけるラウドネス等化のための別のシステムのブロック図である。

40

【発明を実施するための形態】

【0009】

本発明のいくつかの実施形態を、添付図面を参照してここに説明する。実施形態に記載される部品の形状、相対位置、及び他の態様が明瞭には規定されない場合はいつでも、本発明の範囲は、示した部品のみ限定されず、その部品は、単に説明目的のためであることが意味される。多くの詳細が説明されるが、本発明のいくつかの実施形態は、これらの詳細なしに実施され得ることも理解される。他の事例では、本明細書の理解を妨げないように、周知の回路、構造、及び技術は詳細には示していない。

【0010】

背景技術の項で紹介した従来の「ラウドネス」ボタン機構は、聴き手によって報告され

50

た低周波数帯域損失量が、聴き手におけるその周波数帯域の音響レベルに依存し、それはオーディオコンテンツ自体に依存するという重要な問題を無視している。本発明の実施形態は、（スピーカを駆動する前に）オーディオ信号が通過させられる時間的に変化するフィルタを制御するために、オーディオコンテンツのスペクトル帯域の時間的に変化するレベルを考慮に入れたラウドネス等化スキームである。時間的に変化するフィルタ（本明細書では等化フィルタとも呼ばれるスペクトル整形フィルタ）は、再生レベル及び周波数帯域に対してラウドネス知覚が非線形的であることによって現れるスペクトルのゆがみを補償することを目的とする。

【0011】

再生されているオーディオコンテンツに関連付けられたメタデータに基づいて動作するラウドネス等化器の概念を示すブロック図を図1に示す。この図（及び本明細書の他の図面における全ての図）は、デジタル信号処理（DSP）動作又はDSP論理ハードウェアユニットを表しており（例えば、家庭用オーディオシステム、民生用電子機器スピーカデバイス、又は車両内のオーディオシステムにおけるローカル・ストレージ装置又はメモリなどの機械可読媒体に記憶された命令を実行しているプロセッサ）、例えば、デスクトップコンピュータ、家庭用オーディオエンターテイメントシステム、セットトップボックス、ラップトップコンピュータ、タブレットコンピュータ、スマートホンなどの、オーディオコンテンツを受信している復号化及び再生システム、又は得られたデジタルオーディオ出力信号がアナログ形式に変換されてから、スピーカ（例えば、ラウドスピーカ、イヤホン）を駆動しているオーディオ電力増幅器に供給される他の電子オーディオ再生システムとも呼ばれる。例えばインターネットストリーミング又はインターネットダウンロードを介して受信された当初のオーディオコンテンツは、符号化されメタデータと共に多重化されてビットストリームとされており、本明細書の図面に描かれた処理に到達するまでに、再生システムにおいて解凍され復号化されていてもよい。

【0012】

メタデータ2は、オーディオコンテンツのミキシングレベル及び任意選択でプログラムラウドネスを含む静的メタデータを含み、これは、例えば、完結したコンテンツ（本明細書ではオーディオプログラム又はオーディオアセットとも呼ばれる）毎の単一の値である。ミキシングレベルは、制作中に（又は符号化側で）確立された規格に従って、測定され得る。プログラムラウドネス値は、ITU, 「Algorithms to measure audio programme loudness and true-peak audio level,」ITU-R BS.1770-3で定義されているラウドネスモデルを使用して測定され得る。更に、瞬時ラウドネス値（例えば、音響レベル、音圧レベル、SPL）が、メタデータストリームを介する動的メタデータとして伝達され、そこでは、予想される聴き手の位置における音声コンテンツの音圧レベル（SPL）を（音声コンテンツ信号のフレームに同期した）フレーム又はブロック毎に記述する多数の瞬時ラウドネス値が時間の経過と共に順次受信される。言い換えると、瞬時ラウドネスは、オーディオコンテンツを定義するフレーム又はブロックの時系列によって変化する。メタデータは、復号化及び再生システムである、図示された再生又は復号化側に、オーディオコンテンツ（図では「オーディオ入力」として示されている）とともに、例えば、インターネットダウンロードを介して又はインターネットストリーミングを介して、転送されてもよい。復号化又は再生側では、瞬時ラウドネス値がメタデータ内にあるために追加の遅延は発生せず、したがって、再生側でのラウドネス推定プロセスは不要である。平滑度の改善、デコーダの複雑さの低減、及び追加的遅延のなさは、すべて、ラウドネス等化における最先端技術（メタデータの使用なしに再生側でもっぱら動作する）に対する本提案の利点である。

【0013】

再生側では、ユーザの音量設定（再生中のスピーカ又はイヤホンからの音量を手動で制御するための）が音量制御ブロック4に入力される。次いで、音量制御ブロック4は、デジタルオーディオ出力信号（図の「オーディオ出力」）に適用されることとなる適切なゲ

10

20

30

40

50

イン値（例えば、フルバンドスケーリングファクタ）を（例えば、おそらくテーブル参照を含む計算により）生成する。それは、ユーザの音量設定に基づいて、及び再生システムのレベル伝達特性（感度）に関する記憶された又は既定の知識に基づいて、再生レベルを導出する。後者は、所与のオーディオ出力信号が、どのように、聴き手の耳における音圧レベルを有する音としてレンダリングされるかを表現する（この感度は、ユーザ音量設定などの要因にも依存し得ることに留意を要する）。

【0014】

フィルタ適応ブロック7は、メタデータ2、及び静的ミキシングレベル（メタデータに示される）と再生レベルとの間の計算された差（例えば、ミキシングレベルと再生レベルの比較とも呼ばれる、2つのdB値間の減算として）を取り込み、その差に基づいて等化（EQ）フィルタ5を制御する（例えば、定義する）フィルタパラメータを生成する。フィルタ適応ブロック7は、まず、再生レベルがミキシングレベルより高いか低いかを判定してもよい。（ミキシングレベルがメタデータによって提供されない場合、平均ミキシングレベル（例えば、音声プログラム又はオーディオ録音制作環境で一般的に使用されるもの）を仮定することができる。再生レベルの方が低い場合、（メタデータ2によって）オーディオコンテンツについて報告された瞬時ラウドネスに応じて、低周波数帯域及び任意選択で高周波数帯域をある程度ブーストする必要がある。同様に、再生レベルの方が高い場合には、これらのスペクトル帯域をある程度減衰させる必要がある。EQフィルタ5は、こうしたことを行うように構成されており、（例えば、デジタルオーディオコンテンツ（オーディオ入力）の、メタデータ内の瞬時ラウドネス値に関連付けられたすべてのフレームについて更新されるようにするか、又は、EQフィルタ5を、オーディオコンテンツのすべてのフレームについては更新する必要がないように、いくつかのフレームをスキップすることによって）フレーム毎に更新され得る。

10

20

【0015】

再生レベルがミキシングレベルより低い場合、そして、再生レベルがミキシングレベルと比較して低く、瞬時ラウドネスが低いほど、EQフィルタ5によって与えられるブーストの量が大きくなることに留意されたい。これは、低い音圧レベルでは、レベルに対して、人間のラウドネス知覚の非線形性が増大するためである。また、再生レベルがミキシングレベルと十分には異なっていないことがわかっている一実施形態では、EQフィルタ5によるスペクトル整形は必要ない（例えば、そのレスポンスは0dBで平坦となるべきである）。

30

【0016】

一般に、オーディオスペクトルをいくつかの帯域に分割し、それらの帯域のそれぞれにおけるラウドネスを個別に推定することが有利である。特に本明細書が対象としているラウドネス等化の場合、符号化側でラウドネス測定を行う（メタデータ内の対の瞬時ラウドネス値のシーケンスとして現れる）ために、低周波数帯域及び高周波数での（重複しない）別の帯域を定義することができる。これは、それらの周波数帯域における人間の聴覚をモデル化する試みとして行われる。あるいは、瞬時ラウドネス値は、単一の周波数帯域だけ、例えば200Hz未満の低周波数に対して提供されてもよい。そしてまた、これらのラウドネス値は、その概念を上述した方法でEQフィルタ5を制御するのに適している。

40

【0017】

一実施形態では、EQフィルタ5を制御するために必要な情報は、ミキシング（制作）時（I）及び再生時（II）のオーディオコンテンツの特定のオーディオ帯域（スペクトル帯域）の瞬時SPLを含む。以下において、（I）は、 $L_{range,mix}(t)$ と呼ばれ、（II）は $L_{range,playback}(t)$ と呼ばれる。このような入力を使用し、そして従来の手法を使用して、特定の周波数帯域において必要とされるブーストゲイン又はカットゲインを計算することができる。

【0018】

オーディオコンテンツが与えられれば、オーディオ帯域内の瞬時ラウドネスレベルを、制作又は符号化側で推定できるが、再生中の絶対的なレベルは、再生システムの感度がわ

50

かっている場合にのみ決定できる。再生システムでは、感度 L_{playback} は音響レベル [音圧レベル] L_{playback} と、そのレベルをもたらすコンテンツの電気音響信号レベル [dBFS] との測定された差異を表す。制作システムでは、感度は、コンテンツの電気オーディオ信号レベル (L_{content}) と、得られた測定された SPL、例えば、SPL のミキシングレベルとの差として定義され得る。ミキシングシステムの感度 L_{mixing} は、メタデータ内に静的な値として含まれ得る。あるいは、それは、ミキシングレベル (例えば、ミキシングスタジオにおける測定された平均 SPL) と平均ラウドネスレベル (両方の値をメタデータとして伝達することができる) との間の差を計算することによって、再生システムにおいて推定され得る。平均ラウドネスレベルは、例えば、ITU-R BS. 1770-3 に記載された方法によって計算され得る。この推定はプログラムラウドネスと呼ばれる - 図 1 参照。典型的には、制作 / ミキシングシステムの感度は一定である。しかし、再生システムの場合、それは、ユーザが音量調整を行うとき、例えば、デバイスの音量ノブを回すことによって、変化し得る。

10

【0019】

ミキシングスタジオで瞬時ラウドネス値 [SPL] を計算又は測定できない場合、[dBFS] で表す絶対コンテンツレベル $L_{\text{range, content}}(t)$ に基づいて、再生側で推定瞬時ラウドネスレベルを計算できる。ここで (t) は、コンテンツ (オーディオ入力) のレベル変動によって経時的に変化することを示す。この推定値を使用すると、ミキシング側及び再生側でのスペクトル帯域の SPL を計算できる。

20

$$L_{\text{range, mixing}}(t) = L_{\text{range, content}}(t) - \text{delta } L_{\text{mixing}}$$

及び

$$L_{\text{range, playback}}(t) = L_{\text{range, content}}(t) + \text{delta } L_{\text{playback}}$$

【0020】

あるいは、生産側と再生側の間平均レベル差 L_{acoustic} は、平均ミキシングレベル L_{mixing} 及び再生レベル L_{playback} に基づいて直接的に計算され得る。

$$L_{\text{acoustic}} = L_{\text{playback}} - L_{\text{mixing}}$$

その結果に基づいて、再生側の瞬時 SPL は、次式で表される。

$$L_{\text{range, playback}}(t) = L_{\text{acoustic}} + L_{\text{range, mix}}(t)$$

【0021】

特定のスペクトル帯域で人間が知覚するラウドネスは、低い周波数帯域では、より低い音圧レベル SPL において非線形的に低下する。ミキシング時及び再生シナリオにおける低周波数及び中間周波数帯域に対する、実験室の環境で測定できる従来型の知覚ラウドネスカーブが存在する。

30

【0022】

様々な試験信号の周波数及びレベルに依存するラウドネスの心理音響的な測定の様々な公開物に基づいて、このようなブーストゲインの量を計算するために従来技術を使用することができる。例えば、T. Holman and F. Kampmann, 「Loudness Compensation: Use and Abuse」, Journal of the Audio Engineering Society, July / August 1978, Vol. 26, No. 2 / 8 参照。データの一般的な表現は、レベルに対するラウドネスの増加を示す等ラウドネス曲線又はグラフの形である。そのような心理音響データを用いて、ブーストゲインの量は、瞬時ラウドネス値の関数として (フィルタ適応ブロック 7 をプログラムすることによって)、かつ、上述の再生及びミキシングレベルに基づいて、容易に計算され得る。ブーストゲイン値と、周波数ブーストが適用されるべき帯域に基づいて、EQ フィルタ 5 の一部であって対象の周波数帯域において適切なブーストを生成するデジタルフィルタ要素のパラメータを導出することができる。

40

EQ フィルタ要素の実施例

【0023】

以下の実施例は、低周波数帯域及び高周波数帯域のカット及びブーストフィルタ要素の例を示しており、ラウドネス等化のための所望の周波数レスポンスを近似することができ

50

る。この実施例では、図1のEQフィルタ5の一部として、いくつかの異なるフィルタ要素が接続されてカスケードを形成しており、各要素は、オーディオスペクトルの残りの部分を変更することなく(0dBゲイン)、カット又はブースト周波数帯域を有することができる。これらは、低周波数整形フィルタ及び高周波数整形フィルタの実施例であり、低周波数整形フィルタ及び高周波数整形フィルタは、EQフィルタ5の一部としてカスケード接続されている。

【0024】

各低周波数整形フィルタ(EQフィルタ5の一部である)は、実係数を有する一次IIRフィルタであることができ、次の形を有する。

【数1】

$$H_{LF}(z) = \frac{1 - b_1 z^{-1}}{1 - a_1 z^{-1}}$$

10

低周波数カットフィルタは、所望のコーナ周波数に依存する固定係数 a_1 を有することができる。フィルタパラメータ b_1 は、上記で定義したブーストゲイン g_{boost} 又は L_{boost} に基づいて、以下のように動的に計算され得る。

【数2】

$$b_1 = \begin{cases} a_1; & L_{boost} \geq 0.0 \\ b_{1,bound}, a_1 + (b_{1,bound} - a_1) \frac{L_{boost}}{L_{boost,min}}; & L_{boost,min} < L_{boost} < 0.0 \\ b_{1,bound}; & \text{他} \end{cases}$$

20

【0025】

低周波数ブーストフィルタは、コーナ周波数に依存する固定係数 b_1 を有することができる。フィルタパラメータ a_1 は、線形ゲイン L_{boost} に基づいて動的に計算され得る。

【数3】

$$a_1 = \begin{cases} b_1 & L_{boost} \leq 0.0 \\ a_1 + (a_{1,bound} - b_1) \frac{L_{boost}}{L_{boost,max}} & 0.0 < L_{boost} < L_{boost,max} \\ a_{1,bound} & \text{他} \end{cases}$$

30

各高周波整形フィルタは、次の形を有する実係数を有する二次IIRフィルタであることができる。

【数4】

$$H_{HF}(z) = g_{norm} \frac{1 + b_1 z^{-1} + b_2 z^{-2}}{1 + a_1 z^{-1} + a_2 z^{-2}} \quad g_{norm} = \frac{1 + a_1 + a_2}{1 + b_1 + b_2}$$

40

フィルタのコーナ周波数は、オーディオサンプルレート及び正規化されたコーナ周波数に依存し得る。

$$f_c = f_{c,norm} f_s$$

$$f_{fklkl}$$

各高周波数カットフィルタは、 b_1 を除いて固定係数を有することができる。固定フィルタ係数は、コーナ周波数インデックス及び極/零半径パラメータに依存する。

$$r = 0.45$$

$$a_1 = -2r \cos(2\pi f_{c,norm})$$

$$a_2 = r^2$$

$$b_2 = a_2$$

フィルタパラメータ b_1 は、ブーストゲイン g_{boost} 又は L_{boost} に基づいて動的に計算

50

され得る。

【数 5】

$$b_1 = \begin{cases} a_1; & L_{boost} \geq 0.0 \\ a_1 + (b_{1, bound} - a_1) \frac{L_{boost}}{L_{boost, min}}; & L_{boost, min} < L_{boost} < 0.0 \\ b_{1, bound}; & \text{他} \end{cases} \quad \text{A}$$

【0026】

高周波数ブーストフィルタは、a係数がカットフィルタのb係数と同じ方法で計算され、b係数がカットフィルタのa係数と同じ方法で計算されること以外は同じ係数を有することができる。

$$b_1 = -2r \cos(2f_{c, norm})$$

$$b_2 = r^2$$

$$a_2 = b_2$$

フィルタパラメータ a_1 は、ブーストゲイン L_{boost} に基づいて動的に計算することができる。

【数 6】

$$a_1 = \begin{cases} b_1 & L_{boost} \leq 0.0 \\ b_1 + (a_{1, bound} - b_1) \frac{L_{boost}}{L_{boost, max}} & 0.0 < L_{boost} < L_{boost, max} \\ a_{1, bound} & \text{他} \end{cases}$$

【0027】

所与のオーディオプログラムの瞬時ラウドネス（例えば、SPL）値を含むメタデータを生成するための例示的な制作/符号化側システムが図2に示されている。符号化側で行われる瞬時ラウドネス又はSPL測定の精度を向上させるために、再生側のEQフィルタ5のより良いダウストリーム制御の目的で、オーディオ信号がラウドネス測定モジュール14に入る前に、オーディオ信号はまず、対象のスペクトル帯域（EQフィルタ5によって修正されるスペクトル帯域）外のすべての構成要素を除去するバンドパスフィルタ13によって処理されてもよい。このようにして、瞬時ラウドネスのより正確な推定を達成し、そのスペクトル帯域内のラウドネス等化のより良好な知覚品質を得ることができる。瞬時ラウドネスは、例えば、ラウドネス測定モジュール14がバンドパスフィルタ13の出力で短期エネルギーを計算することにより、次いで計算された短期エネルギーシーケンスを平滑化して（平滑化ブロック16）瞬時ラウドネス値の急速な変動を避けることにより、導出され得る。オーディオコンテンツのオフライン処理の場合（ライブストリーミングと比較して）、平滑化のルックアヘッドを増加させて平滑性を向上させることができ、さもなければEQフィルタ5が（瞬時ラウドネス値に応じて）あまりにも急速に又は不適切なときに適応される場合に生じるアーチファクトを回避することができる。

【0028】

本発明の別の実施形態によれば、以下のアプローチを使用することにより（復号化側/再生システムにおいてラウドネスEQを達成するために）、メタデータに（制作側又は符号化側で計算された）瞬時ラウドネス値のシーケンスを含める必要がなくなる。ISO/IEC, 「Information technology - MPEG audio technologies - Part 4: Dynamic range control」, ISO/IEC 23003-4:2015は、ラウドネス及びダイナミックレンジ制御（DRC）のための柔軟なスキームを定義している。それは、DRCゲイン値を復号化側に伝達し、DRCゲイン値を復号化されたオーディオ信号に適用することによって復号化側で圧縮効果を適用するために、メタデータ内のゲインシーケンスを使用する。符号

10

20

30

40

50

化に戻って、これらのDRCゲイン値は、通常、図3に示すようなDRC特性を平滑化された瞬時ラウドネス推定値に適用することによって、生成される。図3は、「ISO/IEC, 「Information technology - MPEG audio technologies - Part 4: Dynamic range control,」 ISO/IEC 23003-4:2015からのものである。図3のグラフにおけるDRC入力レベルは、平滑化された瞬時ラウドネスレベルである。

【0029】

本発明の一実施形態によれば、MPEG-D DRCで定義されるようなオーディオプログラムの圧縮のために意図された、同じメタデータソースDRCゲインシーケンスを（再生中の同じオーディオプログラムに対し）ラウドネスEQの目的で使用することが可能である。図4を参照すると、（復号化側で）DRCゲインシーケンスに逆DRC特性関数20を適用することによってこれを行うことができ、その平滑化された瞬時ラウドネス値を復元し、瞬時SPL値として再解釈し、次いで使用して上述のようにラウドネスEQフィルタ5を動的に更新することができる。この逆関数は、例えば、メタデータ内で受信された符号化されたDRCゲイン値のシーケンスを計算するために符号化側で適用されたDRC特性であるか又はそれを表す、いくつかのDRCゲイン曲線（例えば、図3に示すものなど）の1つである数学関数の入力変数及び出力変数を逆転させることによって得られ得る。言い換えると、逆DRC特性は、符号化側のオーディオコンテンツに適用されてDRCゲイン値を生成したDRC特性の逆数であってもよい。ここで、後者のシーケンスは数学的関数の「出力」に（又は、数学的関数の計算された逆関数への入力として）適用され、DRCフレーム毎に、瞬時ラウドネスレベルとして扱われる、対応するラウドネス値のシーケンスを生成する。オフセット調整されたシーケンスをフィルタ適応ブロック7に供給する前に、基準レベルオフセットを適用して、シーケンスのそのような瞬時値のそれぞれを調整することができることに留意されたい。例えば、ここで、オフセットは、基準音響レベルを表す例えばdBの固定値である。図4の他の全ての態様は、EQフィルタ5から出てきた（復号化されたオーディオプログラムの）フィルタリングされたオーディオコンテンツをスケーリング又は乗算して、現在のユーザ音量設定を反映させる前に、音量制御部4によって計算されたdB値を線形フォーマットに変換するのに必要であり得る線形ブロック22（図1には図示せず）への任意選択の変換を含め、図1と同じであってもよい。

【0030】

本発明の別の実施形態によれば、メタデータ内に、ラウドネスEQをもっぱら対象とする別個のDRCゲインシーケンスを有することも有用であることが認識される。その目的のために、MPEG-D DRC規格を、（メタデータに含まれる）いくつかのゲインシーケンスのどれが（EQフィルタ5を制御するため）ラウドネスEQに適しているか、及びどの周波数帯域を制御すべきかについての情報を運ぶ、追加的なメタデータ構文によって拡張することができる。メタデータには、それぞれ異なる周波数帯域でラウドネスEQを実行するように指示されている、そのような専用のDRCゲインシーケンスのいくつかが存在してもよい。更に、追加的なメタデータは、該当する場合、ラウドネスEQのために使用されるどのゲインシーケンスが特定のダウンミックス及びダイナミックレンジ制御にも適しているか、を指定することができる。この実施形態は、図5のブロック図を用いて説明される。図5と図4との類似点は明らかであるが、相違点には以下が挙げられる：図5では、乗算器においてダイナミックレンジ制御（DRCゲイン調整）が適用され、異なるDRCゲインシーケンス、シーケンス2（この場合任意選択のDRCゲイン修正ブロック25を使用する）から導出されるが、図4ではDRCゲイン調整は適用されない。また、EQフィルタ5（ラウドネスEQ）はここでは瞬時ラウドネス（SPL）値の関数として制御されており、それはもっぱらの目標とされたDRCゲインシーケンス1から導出され、動的制御のために乗算器で同時に適用されているDRCゲイン調整値（DRCゲインシーケンス2から導出される）によって（加算ユニットで）補正されている。

【0031】

10

20

30

40

50

図5においても、例えば固定された参照SPLである、瞬時ラウドネス値に対する静的オフセットの代わりに、DRCゲイン修正ブロック25の出力によって与えられ得る動的補正が行われている。しかしながら、瞬時ラウドネス値に対してなされる補正は代わりに、メタデータから供給されるDRCゲインシーケンス2によって直接与えられ得るので、ブロック25は任意選択である。DRCゲイン修正ブロック25は、再生中に適用されている圧縮プロファイル又はDRC特性を変更するために、制作/符号化側によって(メタデータソースDRCゲインシーケンスを計算するために)選択され使用されたものと比較して、任意選択で含まれ得る。DRCゲイン修正ブロック25は、この特定の再生システムにより好適であり得る、いわゆる「修正された」DRCゲイン(新しいDRCゲイン調整値)を生成する、米国特許出願公開第2014/0294200号(段落[0040] ~ [0045])の記述によることができる。いずれの場合でも、ここでは、フィルタ適応ブロック7に入力される瞬時ラウドネスシーケンスは、そのゲイン値が、例えば、図の乗算器によって示される、EQフィルタ5のダウンストリームにおけるオーディオコンテンツのスケージングにも適用されているDRCゲイン値シーケンスによって、ダイナミックレンジ制御の目的で補正される。したがって、このような技術では、メタデータがフレーム単位でDRCゲイン値を既に提供している場合、瞬時ラウドネス値の別個のシーケンス(図1に関連して上記参照)を(再生側でラウドネスEQを達成するために)メタデータに含める必要はない。

10

【0032】

更に別の実施形態では、図6を参照すると、図4のラウドネスEQスキームがDRCと組み合わせられ、ダイナミックレンジ調整されたオーディオコンテンツがEQフィルタ5によってフィルタリングされ、最終的にEQフィルタ及びダイナミックレンジ調整の両方がなされたオーディオコンテンツを生成する。しかしながら、これは、図5で行われた方法(ラウドネスEQとDRCも組み合わせる)とは異なる方法で行われる。図5との相違点には以下が含まれる: 所与のDRC利得シーケンス1(メタデータ内)に逆DRC特性関数20を適用することによって提供される瞬時ラウドネスシーケンスに基準レベル(例えば、固定値)を加えることにより、ミキシング時に瞬時音響レベルをフィルタ適応ブロック7に提供すること、及び同時に適用されているDRCゲインに従って、再生レベルとミキシングレベルとの間の静的な差を動的に調整することによって、フィルタ適応ブロック7への入力を動的に更新すること。他の相違点には、MPEG-D DRCで定義されるような、制御パラメータをDRCブロックに供給するためのフォーマットである、DRCインタフェースの追加、及びEQフィルタ5のアップストリーム(図5に見られるEQフィルタ5のダウンストリームに比較して)のオーディオコンテンツへのDRCゲインの適用が含まれる。

20

30

動的等化及びDRC

【0033】

上述のスキームは、MPEG-D DRCで提供されるようなDRCと組み合わせることができるラウドネスEQツールを提供する。しかし、いくつかのアプリケーションでは、例えば、再生レベルが不明なために、ラウドネスEQツールが複雑すぎることもあり得るか、又は適切な制御ができないこともあり得る。

40

【0034】

多くのアプリケーションでは、マルチバンドDRCが採用されてダイナミックレンジ圧縮を達成する。多くの場合、各DRCバンド内で個別に圧縮を制御することによって、「動的等化」も可能である。次のアプローチは、ラウドネスEQだけよりも一般的な目的のために、そのような動的EQスキームを提供する。

【0035】

以下では、上述のラウドネスEQといくつかの面で同様に動作するDRCの動的EQスキームについて説明する。相違点は、それは再生レベル(例えば、音量制御ブロック4によって生成される - 図1、図4、図5、図6参照)を考慮しないことである。それはむしろEQを、ダイナミックレンジ制御による色付け効果を補償するために適用しており、こ

50

の効果は、部分的に、DRCの適用に起因するレベルの変化及び関連するラウドネス知覚の心理音響特性から生じ得る。本明細書に記載される動的EQアプローチの他の有用な応用は、例えば、オーディオコンテンツにおけるノイズの多い低レベルバックグラウンド音のバンドパスフィルタリングを含み、さもなければ騒々しく聞こえ得るノイズの大きな増幅を回避する。

【0036】

以下に説明するスキームは、(メタデータに基づいて)MPEG-D DRCに統合され得る。しかし、それは一般的なリアルタイムダイナミックレンジ制御(メタデータなし)でも使用され得る。それは、以前はマルチバンドメタデータベースのDRCプロセスがサポートされている場合にのみ可能であったシングルバンドDRCの動的EQの利点を提供することができる。従来のメタデータベースのシングルバンドDRCプロセスは、すべての周波数成分に同じゲインを適用するので、例えば低周波数帯域のみでDRCゲインを選択的に低減することはできない。

10

【0037】

更に、以下に説明するスキームは、従来のメタデータに基づくマルチバンドDRCプロセスのサブバンド解像度に限定されず、したがって、より滑らかなスペクトル整形を提供することができ、より低い計算複雑性を有することができる。図7は、動的EQと組み合わせたDRC機能の実施例を示す。EQは、DRCゲインシーケンスによって間接的に動的に制御され、この態様は、逆DRC特性関数20、DRCゲイン値によって瞬時ラウドネスを補正する加算ユニット、及び任意選択のDRCゲイン修正ブロック25(ブロック25によって生成されたDRCゲイン値は、dBブロック26への変換によってラウドネスdB値に変換される)を含む点において、図5の一部に類似している。しかしながら、ラウドネス値は、逆DRC特性関数20を、DRCゲイン値を生成している同じメタデータソースDRCゲインシーケンスに適用することによって得られる。この実施形態では、EQフィルタ5は、以下を決定するビットストリーム内で伝達される静的メタデータに部分的に基づいて設定される：フィルタタイプ(例えば、低周波数カット/ブースト、高周波数カット/ブースト)、フィルタ強度、及び適応周波数帯域。ここで、図1、図4、図5の等化器の実施形態では、メタデータ内の静的フィルタ構成情報は必要ないことに留意されたい。

20

【0038】

図7とは代替的なスキームでは、逆DRC特性関数20は省略され、フィルタ適応ブロック7は、DRCゲイン値によって直接駆動されて複雑さを軽減する。言い換えると、EQフィルタ5を定義する動的パラメータは、メタデータ内で受信されたDRCゲイン値に直接基づいて計算される一方、それらのDRCゲイン値も、DRCゲイン調整値を計算するように処理され(例えば、任意選択のDRC修正ブロック25によって)、それは(ダイナミックレンジ制御を行うために)EQでフィルタリングされたオーディオコンテンツに適用される。

30

【0039】

ここで以下の発明の説明がなされる。製造品はその中に命令を記憶した機械可読媒体を含み、その命令は、オーディオ再生システムのプロセッサによって実行されるとき、以下のように、ダイナミックレンジ制御を適用しながら動的オーディオ等化を実行する。オーディオコンテンツが受信され、オーディオコンテンツのメタデータも受信され、メタデータは、オーディオコンテンツについて計算された、複数のダイナミックレンジ制御、DRC、ゲイン値を含む。逆DRC特性がメタデータ内で受信された複数のDRCゲイン値に適用され、オーディオコンテンツの複数の瞬時ラウドネス値を計算する。等化フィルタを定義する複数の動的パラメータが計算され、動的パラメータは計算された複数の瞬時ラウドネス値に基づいて計算される。オーディオコンテンツが等化フィルタによってフィルタリングされ、EQフィルタリングされたオーディオコンテンツを生成する。メタデータ内で受信された複数のDRCゲイン値が使用されて、複数のDRCゲイン調整値を計算する。複数のDRCゲイン調整値がEQフィルタリングされたオーディオコンテンツに適用さ

40

50

れ、ダイナミックレンジ制御を行う。動的等化の別の実施形態では、計算された複数の瞬時ラウドネス値が複数のDRCゲイン調整値に従って補正されて、補正された瞬時ラウドネス値を生成し、等化フィルタを定義する複数の動的パラメータは複数の補正された瞬時ラウドネス値に基づいて計算される。更に、計算された複数の瞬時ラウドネス値を補正することは、dBフォーマットで、計算された複数の瞬時ラウドネス値を複数のDRCゲイン調整値に加算することを含む。別の態様では、メタデータは、等化フィルタを定義するための以下の1つ以上を指定する静的フィルタ構成データを含む：a)タイプ、例えば低周波数カット若しくはブースト、又は高周波数カット若しくはブースト、b)フィルタ強度、及びc)適応又は有効周波数帯域。その場合、静的フィルタ構成データに従って構成された等化フィルタは、オーディオコンテンツがそれを通してに、動的パラメータによって動的に修正される。更に、更なる態様では等化フィルタを定義する前記複数の動的パラメータを計算することは、ミキシングレベル又は再生レベルを使用しない。

10

【0040】

製造品はその中に命令を記憶した機械可読媒体を含み、その命令は、オーディオ再生システムのプロセッサによって実行されるとき、以下のように、ダイナミックレンジ制御を適用しながら動的オーディオ等化を実行する。オーディオコンテンツ、及びオーディオコンテンツのメタデータを受信し、メタデータは、オーディオコンテンツについて計算された複数のダイナミックレンジ制御、DRC、ゲイン値を含む。等化フィルタを定義する複数の動的パラメータを計算し、動的パラメータを、メタデータで受信されたDRCゲイン値に基づいて計算する。等化フィルタを使用して受信されたオーディオコンテンツをフィルタリングし、EQフィルタリングされたオーディオコンテンツを生成する。メタデータ内で受信された複数のDRCゲイン値を処理して、複数のDRCゲイン調整値を計算する。計算された複数のDRCゲイン調整値をEQフィルタリングされたオーディオコンテンツに適用して、ダイナミックレンジ制御を行う。

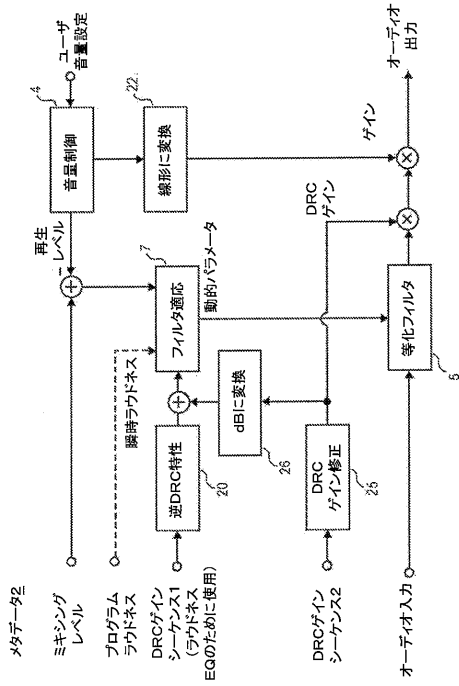
20

【0041】

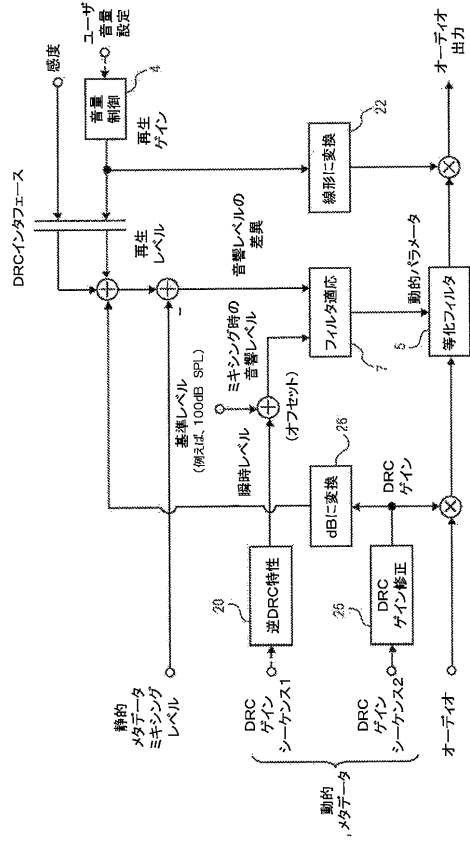
特定の実施形態を記載し添付の図面を示してきたが、そのような実施形態は、広範な発明を単に説明するものであって、限定するものではなく、当業者には様々な他の修正が生じ得るので、本発明は、図示又は記載された特定の構造及び配置に限定されないことが理解されるべきである。したがって、本明細書は、限定的ではなく例示的であると見なされるべきである。

30

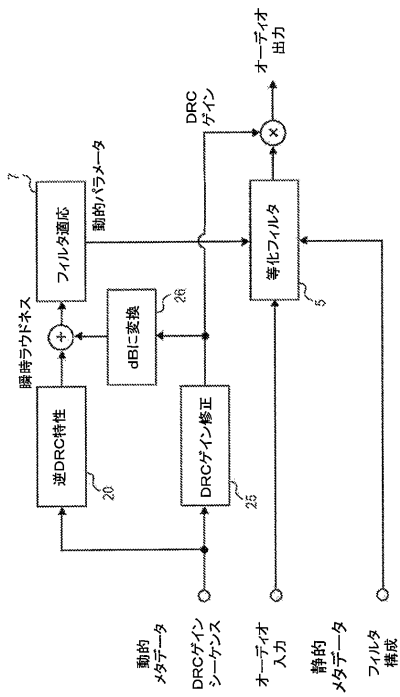
【 図 5 】



【 図 6 】



【 図 7 】



フロントページの続き

(74)代理人 100086771

弁理士 西島 孝喜

(74)代理人 100139712

弁理士 那須 威夫

(72)発明者 バウムガルテ フランク

アメリカ合衆国 9 5 0 1 4 カリフォルニア州 クパチーノ インフィニット ループ 1 メ
ール ストップ 3 8 - 2 アイエムジー

Fターム(参考) 5D162 AA09 BA02 BA03 BA07 CA05 EA01 EA07

5D208 BA01 BB01

5J030 AA05 AA10 AA14 AB01 AC10

【外国語明細書】
2020008878000001.pdf