

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第5101292号  
(P5101292)

(45) 発行日 平成24年12月19日 (2012.12.19)

(24) 登録日 平成24年10月5日 (2012.10.5)

|                      |                 |
|----------------------|-----------------|
| (51) Int. Cl.        | F I             |
| H03G 9/00 (2006.01)  | H03G 9/00 A     |
| G10L 11/00 (2006.01) | G10L 11/00 101F |
| G10L 21/04 (2006.01) | G10L 11/00 101E |
| H03G 3/32 (2006.01)  | G10L 21/04 120D |
| H03G 3/30 (2006.01)  | H03G 3/32       |

請求項の数 36 (全 47 頁) 最終頁に続く

|               |                               |           |                     |
|---------------|-------------------------------|-----------|---------------------|
| (21) 出願番号     | 特願2007-539070 (P2007-539070)  | (73) 特許権者 | 507236292           |
| (86) (22) 出願日 | 平成17年10月25日 (2005.10.25)      |           | ドルビー ラボラトリーズ ライセンシン |
| (65) 公表番号     | 特表2008-518565 (P2008-518565A) |           | グ コーポレイション          |
| (43) 公表日      | 平成20年5月29日 (2008.5.29)        |           | アメリカ合衆国 カリフォルニア州 94 |
| (86) 国際出願番号   | PCT/US2005/038579             |           | 103-4813 サンフランシスコ ポ |
| (87) 国際公開番号   | W02006/047600                 |           | トレロ アベニュー 100       |
| (87) 国際公開日    | 平成18年5月4日 (2006.5.4)          | (74) 代理人  | 100068755           |
| 審査請求日         | 平成20年10月24日 (2008.10.24)      |           | 弁理士 恩田 博宣           |
| (31) 優先権主張番号  | 60/622, 458                   | (74) 代理人  | 100105957           |
| (32) 優先日      | 平成16年10月26日 (2004.10.26)      |           | 弁理士 恩田 誠            |
| (33) 優先権主張国   | 米国 (US)                       | (74) 代理人  | 100142907           |
| (31) 優先権主張番号  | 60/638, 607                   |           | 弁理士 本田 淳            |
| (32) 優先日      | 平成16年12月21日 (2004.12.21)      | (74) 代理人  | 100071010           |
| (33) 優先権主張国   | 米国 (US)                       |           | 弁理士 山崎 行造           |

最終頁に続く

(54) 【発明の名称】 オーディオ信号の感知音量及び／又は感知スペクトルバランスの計算と調整

(57) 【特許請求の範囲】

【請求項 1】

オーディオ信号における特段のラウドネス特性を制御する方法であって、  
 該特段のラウドネス特性とは、周波数と時間の関数としての感知音量の指標である特定ラウドネス、又は、2 次的な干渉性のあるサウンドの存在下で周波数と時間の関数としての信号の感知音量の指標である部分特定ラウドネスであり、該方法は、  
目標特定ラウドネスを取得するステップであって、前記目標特定ラウドネスは、前記オーディオ信号の特定ラウドネスから導き出される、前記取得するステップと、  
 前記オーディオ信号の特段のラウドネス特性と前記目標特定ラウドネスとの差を縮めるために、前記オーディオ信号を修正するのに有用な周波数に依存する修正パラメータを導き出すステップと、

a) 前記オーディオ信号の特段のラウドネス特性と前記目標特定ラウドネスとの差を縮めるために、前記修正パラメータを該オーディオ信号に適用するステップと、  
 b) 前記オーディオ信号の特段のラウドネス特性と前記目標特定ラウドネスとの差を縮めるために、時間的に及び／又は空間的に分離して前記修正パラメータを前記オーディオ信号に適用するように、前記修正パラメータと前記オーディオ信号を伝達又は保存するステップと、  
 を具備することを特徴とする方法。

【請求項 2】

前記取得するステップでは、前記オーディオ信号及び前記目標特定ラウドネス又は該目標

10

20

特定ラウドネスを代理するものを、伝達系から受け取り又は保存媒体から複製し、前記導き出すステップでは、前記前記目標特定ラウドネス又はそれを代理するものから前記修正パラメータを導き出すことを特徴とする請求項 1 に記載の方法。

【請求項 3】

オーディオ信号の特定ラウドネス特性を制御する方法であって、  
該特定ラウドネス特性とは、周波数と時間の関数としての感知音量の指標である特定ラウドネス、又は、2 次的な干渉性のあるサウンドの存在下で周波数と時間の関数としてのオーディオ信号の感知音量の指標である部分特定ラウドネスであり、該方法は、  
前記オーディオ信号を修正するための周波数に依存する修正パラメータであって、前記修正パラメータは、目標特定ラウドネスから導き出され、前記目標特定ラウドネスは、前記オーディオ信号の特定ラウドネスから導き出される、前記修正パラメータと、前記オーディオ信号とを、伝達系から受け取り又は保存媒体から複製するステップと、  
前記オーディオ信号の特定ラウドネス特性と前記目標特定ラウドネスとの差を縮めるために、受け取った前記修正パラメータに応答して、前記オーディオ信号を修正するステップと、  
を具備することを特徴とする方法。

【請求項 4】

前記目標特定ラウドネスは、前記オーディオ信号の特定ラウドネスの 1 以上の圧伸を含むことを特徴とする請求項 1 乃至 3 のいずれか 1 項に記載の方法。

【請求項 5】

【数 1】

前記 1 以上の圧伸は、以下のような関係

で前記特定ラウドネスの圧伸を行う、時間的に変化する周波数的に変化するスケールファクタ  $\Xi[b, t]$  を含み、

$$\hat{N}[b, t] = \Xi[b, t] N[b, t]$$

ここで、 $(\hat{N}[b, t])$  は前記オーディオ信号の目標特定ラウドネス、 $(N[b, t])$  は前記オーディオ信号の特定ラウドネスであることを特徴とする請求項 4 に記載の方法。

【請求項 6】

前記圧伸は、少なくとも一部は、好ましい複数帯域音量と前記オーディオ信号の複数帯域音量との比により定められることを特徴とする請求項 5 に記載の方法。

【請求項 7】

前記圧伸は、以下のような関係で、 $L_o[b, t] / L_i[b, t]$  として表現でき、

【数 2】

$$\hat{N}[b, t] = \frac{L_o[b, t]}{L_i[b, t]} N[b, t],$$

ここで、 $N[b, t]$  は、前記オーディオ信号の特定ラウドネスであり、 $L_o[b, t]$  は、好ましい複数帯域音量であり、 $L_i[b, t]$  は、前記オーディオ信号の複数帯域音量であり、 $\hat{N}[b, t]$  は、前記目標特定ラウドネスであることを特徴とする請求項 6 に記載の方法。

【請求項 8】

$L_o[b, t]$  は、 $L_i[b, t]$  の関数であることを特徴とする請求項 7 に記載の方法。

【請求項 9】

$L_i[b, t]$  の関数としての  $L_o[b, t]$  は、下記のように表現することができ、

$$L_o[b, t] = DRC\{L_i[b, t]\}$$

ここで、 $DRC\{\}$  は、 $L_i[b, t]$  から  $L_o[b, t]$  への写像するダイナミックレンジ関数であることを特徴とする請求項 8 に記載の方法。

【請求項 10】

$L_i[b, t]$  は、前記オーディオ信号の前記特定ラウドネスの時間的に平滑化したもの及び / 又は周波数的に平滑化したものであることを特徴とする請求項 7 に記載の方法。

10

【請求項 11】

前記方法は、前記修正するステップ又は、前記修正パラメータを生成する応用例又は、目標特定ラウドネスが、前記感知されたオーディオスペクトル又は干渉性信号の存在下で感知されたオーディオスペクトルが異なった特定ラウドネスの圧伸の値に対して異なる値となることができるオーディオ信号に対応する、ダイナミックレンジ制御として有用であることを特徴とする請求項 6 乃至 10 のいずれか 1 項に記載の方法。

【請求項 12】

前記ダイナミックレンジ関数は、各帯域に適用した平均変化がすべての帯域で実質的に等しいとき、各帯域に適用した短期変化が帯域間で独立に変化するように各帯域の音量を制御することを特徴とする請求項 11 に記載の方法。

20

【請求項 13】

$L_i[b, t]$  の関数としての  $L_o[b, t]$  は、下記のように表現することができ、

【数 3】

$$L_o[b, t] = \frac{\bar{L}_i[b, t]}{\bar{L}_i[t]} DRC\left\{\frac{\bar{L}_i[t]}{\bar{L}_i[b, t]} L_i[b, t]\right\}$$

ここで、 $L_o[t] = DRC\{L_i[t]\}$  は、前記オーディオ信号のトータル音量の

好ましいトータル音量への写像を表し、 $\bar{L}_i[t]$  は、前記オーディオ信号の広帯域音量  $L_i[t]$

30

の時間平均したものを表し、 $\bar{L}_i[b, t]$  は、前記オーディオ信号の複数帯域ラウドネス

$L_i[b, t]$  の時間平均したものを表すことを特徴とする請求項 12 に記載の方法。

【請求項 14】

前記方法は、前記修正するステップ又は、前記修正パラメータを生成する応用例又は、目標特定ラウドネスが、前記感知されたオーディオスペクトル又は干渉性信号の存在下で感知されたオーディオスペクトルが異なった特定ラウドネスの圧伸の値に対して、前記オーディオ信号の感知されたオーディオスペクトルの圧伸と実質的に同じ値となることができるオーディオ信号に対応する、ダイナミックレンジ制御として有用であることを特徴とする請求項 12 又は請求項 13 に記載の方法。

40

【請求項 15】

前記オーディオ信号の複数帯域音量は、前記オーディオ信号をクリティカル帯域に分割し前記クリティカル帯域のいくつかについて周波数平滑化することにより、概算することを特徴とする請求項 5 乃至請求項 14 のいずれか 1 項に記載の方法。

【請求項 16】

特定の帯域  $b$  で複数帯域音量の帯域平滑化したもの、すなわち  $L[b, t]$  は、全帯域  $c$  での畳み込み積算として以下のように表現され、

50

【数 4】

$$L[b, t] = \sum_c Q(b - c) N[c, t],$$

ここで、 $N[b, t]$  は、前記オーディオ信号の特定ラウドネスであり、 $Q(b - c)$  は前記平滑フィルターの帯域シフトした応答であることを特徴とする請求項 15 に記載の方法。

【請求項 17】

前記 1 以上の平滑化は、下式の関係で特定ラウドネスを圧伸する時間的に変化する周波数的に変化しないスケールファクタを含み

10

【数 5】

$$\hat{N}[b, t] = \Phi[t] N[b, t],$$

ここで、 $(\hat{N}[b, t])$  は、目標特定ラウドネスであり、 $(N[b, t])$  は、オーディオ信号の特定ラウドネスであり、 $b$  は周波数の指標を示し、 $t$  は時間の指標を示すことを特徴とする請求項 4 に記載の方法。

20

【請求項 18】

前記圧伸は、少なくとも一部は、好ましい広帯域音量と前記オーディオ信号の広帯域音量との比により定められることを特徴とする請求項 17 に記載の方法。

【請求項 19】

前記オーディオ信号の特定ラウドネスの関数における前記圧伸は、下式のような関係で、 $L_o[b, t] / L_i[b, t]$  として表現でき、

【数 6】

$$\hat{N}[b, t] = \frac{L_o[t]}{L_i[t]} N[b, t],$$

30

ここで、 $N[b, t]$  は、前記オーディオ信号の特定ラウドネスであり、 $L_o[t]$  は、好ましい広帯域音量であり、 $L_i[t]$  は、前記オーディオ信号の広帯域音量であり、 $\hat{N}[b, t]$  は、前記目標特定ラウドネスであることを特徴とする請求項 17 又は請求項 18 に記載の方法。

【請求項 20】

$L_o[t]$  は、 $L_i[t]$  の関数であることを特徴とする請求項 19 に記載の方法。

40

【請求項 21】

$L_i[t]$  の関数としての  $L_o[t]$  は、下式のように表現することができ、

$$L_o[t] = DRC\{L_i[t]\}$$

ここで、 $DRC\{\}$  は、 $L_i[t]$  から  $L_o[t]$  への写像するダイナミックレンジ関数であることを特徴とする請求項 20 に記載の方法。

【請求項 22】

$L_i[t]$  は、前記オーディオ信号の前記特定ラウドネスの時間的に平滑化したものであることを特徴とする請求項 19 に記載の方法。

【請求項 23】

50

$L_i[t]$  は、前記オーディオ信号の長期音量の指標であることを特徴とする請求項 1 9 に記載の方法。

【請求項 2 4】

$L_i[t]$  は、前記オーディオ信号の短期音量の指標であることを特徴とする請求項 1 9 に記載の方法。

【請求項 2 5】

前記方法は、前記修正するステップ又は、前記修正パラメータを生成する応用例又は、前記目標特定ラウドネスが、感知されたオーディオスペクトル又は干渉性信号の存在下での感知されたオーディオスペクトルが前記オーディオ信号の感知されたオーディオスペクトルの圧伸の値として異なった特定ラウドネスの圧伸の値又は異なった部分特定ラウドネスの圧伸の値に対して実質的に同じ値となることができるオーディオ信号に対応する、自動ゲイン制御又はダイナミックレンジ制御として有用であることを特徴とする請求項 1 7 乃至請求項 2 4 のいずれか 1 項に記載の方法。

10

【請求項 2 6】

前記修正するステップ又は、前記導き出すステップは、( 1 ) 特定ラウドネス、及び / 又は、( 2 ) 部分特定ラウドネス、及び / 又は、( 3 ) 目標特定ラウドネスを明示的に計算することを特徴とする請求項 1 乃至請求項 2 5 のいずれか 1 項に記載の方法。

【請求項 2 7】

前記修正するステップ又は、前記導き出すステップは、( 1 ) 特定ラウドネス、及び / 又は、( 2 ) 部分特定ラウドネス、及び / 又は、( 3 ) 目標特定ラウドネスを黙示的に計算

20

【請求項 2 8】

前記修正するステップ又は、前記導き出すステップ又は、前記生成するステップは、本質的に ( 1 ) 特定ラウドネス、及び / 又は、( 2 ) 部分特定ラウドネス、及び / 又は、( 3 ) 目標特定ラウドネスを決定するルックアップテーブルを採用することを特徴とする請求項 2 7 に記載の方法。

【請求項 2 9】

前記修正するステップ又は、前記導き出すステップ又は、前記生成するステップは、本質的に ( 1 ) 特定ラウドネス、及び / 又は、( 2 ) 部分特定ラウドネス、及び / 又は、( 3 ) 目標特定ラウドネスを決定する閉形式の数学表現を採用することを特徴とする請求項 2 8 に記載の方法。

30

【請求項 3 0】

前記修正パラメータは時間的に平滑化されていることを特徴とする、請求項 1 乃至請求項 2 9 のいずれか 1 項に記載の方法。

【請求項 3 1】

前記修正パラメータは、前記オーディオ信号の周波数帯域に関する複数の振幅スケールファクタを具備することを特徴とする請求項 3 0 に記載の方法。

【請求項 3 2】

前記複数の振幅スケールファクタの少なくともいくつかは、時間的に変化することを特徴とする請求項 3 1 に記載の方法。

40

【請求項 3 3】

前記修正パラメータは、1 以上のフィルターを制御するためのフィルター係数を含むことを特徴とする請求項 3 0 に記載の方法。

【請求項 3 4】

前記 1 以上のフィルターの少なくともいくつかは時間的に変化し、前記フィルター係数の少なくともいくつかは時間的に変化することを特徴とする請求項 3 3 に記載の方法。

【請求項 3 5】

請求項 1 乃至請求項 3 4 のいずれか 1 項に記載の方法のすべてのステップを実施するための装置。

【請求項 3 6】

50

請求項 1 乃至請求項 3 4 のいずれか 1 項に記載の方法のすべてのステップをコンピュータに実行させるための、コンピュータ読取可能な媒体に保存されたコンピュータプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明はオーディオ信号処理に関する。さらに詳細には、本発明はオーディオ信号の感知音量及び／又は感知スペクトルバランスの計測と制御に関する。本発明は、例えば、音量補償ボリューム制御、自動ゲイン制御、ダイナミックレンジ制御（例えば、リミッター、圧縮器、伸張器、等を含む）、ダイナミックイコライゼーション、及び、オーディオ再生環境における暗騒音の相殺、のうちの 1 つ以上で有用である。本発明には方法が含まれるのみならず対応するコンピュータプログラム及び装置が含まれる。

10

【0002】

なお、本願は米国暫定特許出願 S . N . 6 0 / 6 2 2 , 4 5 8 : 2 0 0 4 年 1 0 月 2 6 日出願、及び S . N . 6 0 / 6 3 8 , 6 0 7 : 2 0 0 4 年 1 2 月 2 1 日出願、に基づく優先権を主張する。

【背景技術】

【0003】

音量を計測する十分客観的な方法を開発する試みは何度もなされてきた。フレッチャーとマンソンは、人間の聴覚は中間周波（又は音声周波）より低周波又は高周波において感度が鈍くなることを 1 9 3 3 年に明らかにした。彼らはまた、サウンドのレベルが上がってゆくにつれて感度が相対的に下がりゆくことも発見した。初期の音量メータは、マイクロフォン、増幅器、及び、低サウンドレベル、中間サウンドレベル、及び高サウンドレベルにおける聴覚の周波数応答を大雑把に模倣するよう組み合わせたフィルターにより構成されていた。

20

【0004】

このような装置が単一で孤立したトーンの一定レベルの音量の計測を行なうが、もっと複雑なサウンドの計測においては、音量の主観的な印象とうまく合致しなかった。このタイプのサウンドレベルメータは標準化されているが、工業的なノイズの監視と制御のような特定の目的にのみ使われる。

30

【0005】

1 9 5 0 年代前半、中でもツビッカーとスチーブンスが、フレッチャーとマンソンの研究を拡張して音量認識プロセスのより現実的なモデルを開発した。スチーブンスは、1 9 5 6 年に、Journal of the Acoustical Society of Americaに、「複合ノイズの音量の測定」方法を発表し、ツビッカーは、1 9 5 8 年に、Acoustical に、彼の「音量の心理的理論」の論文を発表した。1 9 5 9 年にツビッカーは、音量計算の図式的な手順を発表し、そのすぐ後いくつかの類似の論文を発表した。スチーブンスとツビッカーの方法は、I S O 5 3 2 の（それぞれ）パート A 及びパート B として規格化された。両方法は、同様の方法を必要とする。

【0006】

40

まず、エキサイテーションと言われる、内耳基底膜に沿う時間的に変化するエネルギーの分布が、クリティカル帯域レートスケール上で均一な間隔を持つ中央周波数の通過帯域聴覚フィルターのバンクにオーディオを通すことによりシミュレートされる。各聴覚フィルターは、フィルターのその位置に対応する中央周波数を持ち、内耳基底膜に沿う特定の位置に対応する周波数応答をシミュレートするよう設計される。クリティカル帯域幅は、このようなフィルターの帯域幅として定義される。ヘルツ（Hz）で計測されて、これらの聴覚フィルターのクリティカル帯域幅は中央周波数の増大と共に増加する。したがって、全ての聴覚フィルターにおいてこのゆがんだスケールで計測したクリティカル帯域幅が一定になるようなゆがんだ周波数スケールを定義することは有益である。このようなゆがんだスケールは、クリティカル帯域レートスケールと称され、広い範囲の音響心理学的現

50

象を理解しシミュレートする上で有益である。例えば、E・ツビッカーとH・ファステルによる「音響心理学 実際とモデル」1990年ベルリン、を参照のこと。スチーブンスとツビッカーの方法は、バークスケールと称されるクリティカル帯域レートスケールを用い、そこでは、クリティカル帯域幅は500Hz以下では一定であり、500Hzを越えると増大する。最近ではムーアとグラスバーグは、クリティカル帯域レートスケールを定義し、彼等はそれを等価直角帯域幅(Equivalent Rectangular Bandwidth)(ERB)スケールと名づけた。(B. C. J. Moore, B. Glasberg, T. Baer, 「閾値音量と部分音量の予想のためのモデル」Journal of the Audio Engineering Society, Vol. 45, No. 4, 1997年4月, 224~240ページ)。ノッチ雑音マスキングを用いた音響心理学的実験を通して、ムーアとグラスバーグは、バークスケールにおいてクリティカル帯域幅が一定であるのに対して、クリティカル帯域幅が500Hz以下で減少し続けることを示した。

10

#### 【0007】

エキサイテーションの計算に続くのが、「特定ラウドネス」と称される量を生成する、非線形圧縮機能である。特定ラウドネスは、周波数及び時間の関数としての感知音量の指標であり、上述のバークスケール又はERBスケールのような、クリティカル帯域レートスケールに沿って単位周波数毎の感知音量を単位として測定することができる。最終的に時間的に変化する「トータル音量」は周波数横断的に特定ラウドネスを積算することで計算される。特定ラウドネスは、クリティカル帯域レートスケールに沿って一様に配分された有限個の聴覚フィルターから推定され、トータル音量は、各フィルターからの特定ラウドネスを単純に積算することにより計算することができる。

20

#### 【0008】

音量はフォンの単位で計測される。フォン単位の所定のサウンドの音量は、そのサウンドの音量に等しい主観的音量を持つ1kHz毎の音圧レベル(SPL)である。普通、SPLで0dBは、 $2 \times 10^{-5}$ パスカルの2乗平均平方根圧力であり、したがって0フォンでもある。1kHzにおける音量を持つ1kHz以外の周波数のトーンの音量を比較するためにこの定義を用いて、等しい音量の曲線を所定のフォンレベルで定めることができる。図11は、20Hzと12.5kHzとの間の周波数、及び、4.2フォン(聴覚の閾値を考慮した)から120フォン(ISO226:1087(E)、「音質 標準等音量レベル曲線」)までのフォンレベルにおける等音量レベルの曲線を示す。フォンの計測は、周波数の変化に対する人間の聴覚の感度の変化を考慮しているが、SPLにおける音量の増大の非直線性を補正する試み、すなわち曲線の間隔が変化するということに対する補正の試みがなされていないので、計測結果で、レベルが変化したときにサウンドの主観的相対音量を評価することはできない。

30

#### 【0009】

音量は、「ソーン」の単位で計測してもよい。図11に示すように、フォン単位とソーン単位とは1対1対応が成り立つ。1ソーンは、1kHzの純粋なサイン波で40dB(SPL)の音量で定義され、40フォンと等価である。ソーン単位は感知する音量が2倍になったときソーン単位で2倍になるような単位である。例えば、4ソーンは2ソーンの2倍の大きさと感知される。このようにソーンで音量レベルを表現することはより参考になる。特定ラウドネスを感知音量の指標として、周波数と時間の関数として定義すると、特定ラウドネスは、単位周波数当たりのソーン単位で計測してもよい。したがって、バークスケールを用いたとき、特定ラウドネスは、ソーン毎バークとなり、同様にERBスケールを用いたとき、単位はソーン毎ERBとなる。

40

#### 【0010】

上述のように、人間の耳の感度は周波数とレベルの両方で変化し、この事実は音響心理学の文献によく記載されている。1つの結果によれば、感知された所定のサウンドのスペクトル又は音質は、そのサウンドを聞いたときの音量レベルと共に変化する。例えば、低周波数、中間周波数、及び高周波数を含むサウンドについて、このような周波数成分に相対的に比例する感知成分は、そのサウンドの全体的な音量と共に変化する。静かなときの

50

低周波数成分のサウンドと高周波数成分のサウンドは、うるさいときの中間周波数より静かである。この現象はよく知られており、サウンドを再生する機器においていわゆる音量制御により緩和される。音量制御は、ボリュームを下げるとき、低周波、ときには高周波も強めるボリューム制御である。したがって、周波数の極限で耳の感度が低くなるのをそれらの周波数で人工的に増大させることにより補償する。このような制御は完全に受動的であり、補償の程度は、ボリューム制御の設定値又は他のユーザが操作する制御の設定値の関数であり、オーディオ信号内容の関数ではない。

#### 【 0 0 1 1 】

実際には、低周波数、中間周波数、及び高周波数における感知される相対的スペクトルのバランスに変化は信号、特に、その実際のスペクトル及び意図しているのが大きな音が穏やかな音かに依存する。交響楽団のレコーディングを考える。聴衆の一員がコンサートホールで聞いたときと同じレベルで再生されたとすると、この交響楽団が大きな音で又は穏やかな音のどちらで演奏しているかによってスペクトル横断的なバランスを修正することができる。もし、例えばこの音楽が 10 dB 静かに再生されたならば、感知されたスペクトル横断的なバランスを、大きな音の部分に対してある方法で、静かな音の部分に対して別の方法で変化させる。従来の受動的音量制御では、音楽の関数として異なった補償を適用することはない。

#### 【 0 0 1 2 】

シーフェルト他による、2004年5月27日に出願された国際特許出願 No. PCT / US 2004 / 016964 であって、2004年12月23日に WO 2004 / 111994 A2 として国際公開されたものには、特に、オーディオ信号の感知音量を計測し調整するシステムが開示されている。米国を指定国とする前記 PCT 出願は、その全てを参照として本出願に組み込む。前記出願において、音響心理学的モデルにより、知覚単位でオーディオ信号の音量が計算される。加えて、この出願は、オーディオに適用したとき、ゲインにより補正されたオーディオが実質的に参照音量と同じ音量となる、広帯域に乘算できるゲインを計算する技術が導入されている。しかしながら、このような広帯域ゲインを適用することにより、感知されるオーディオのバランスが変化する。

#### 【 発明の開示 】

#### 【 0 0 1 3 】

1つの特徴として、本発明は、その特定ラウドネスと目標特定ラウドネスとの差を縮めるために、オーディオ信号を修正することで、オーディオ信号の特定ラウドネスを制御するのに有用な情報を導き出すものである。特定ラウドネスとは、周波数と時間の関数としての感知音量の指標である。実際の実施の形態において、修正されたオーディオ信号の特定ラウドネスを、目標特定ラウドネスに近似させる。この近似は、以下に既述するように、通常の信号処理の検討に影響されるだけでなく、この修正に用いた時間平滑及び/又は周波数平滑にも影響される。

#### 【 0 0 1 4 】

特定ラウドネスは周波数と時間の関数としての感知音量の指標なので、オーディオ信号の特定ラウドネスと目標特定ラウドネスとの差を縮めるために、前記修正では、周波数の関数としてこのオーディオ信号を修正してもよい。場合によって目標特定ラウドネスは時間不変であるかもしれないが、オーディオ信号自身は定常的な時間不変信号かもしれないが、一般に、この修正では、時間の関数としてオーディオ信号を修正してもよい。

#### 【 0 0 1 5 】

本発明の特徴は、オーディオ再生環境における暗騒音を相殺するために採用することができることである。暗騒音がある状態でオーディオを聞くと、オーディオのレベルとスペクトル及びノイズのレベルとスペクトルの両方に依存するような形で部分的に又は完全にノイズがオーディオをマスクすることがある。その結果、感知されるオーディオスペクトルの変化があらわれる。音響心理学的研究によれば、(B. C. J. Moore, B. Glasberg, T. Baer, 「閾値音量と部分音量の予想のためのモデル」 Journal of the Audio Engineering Society, Vol. 45, No. 4, 1997年4月参照)ノイズのよう

10

20

30

40

50



な２次的な干渉性のあるサウンド信号の存在下でオーディオの感知音量として、オーディオの「部分的特定ラウドネス」を定義することができる。

【００１６】

このように、他の特徴として、本発明は、その特定ラウドネスと目標特定ラウドネスとの差を縮めるために、オーディオ信号を修正することで、オーディオ信号の部分的特定ラウドネスを制御するのに有用な情報を導き出すものである。このようにすることで、知覚的に正確な方法でノイズの影響を緩和する。干渉性のあるノイズ信号を考慮に入れる本発明のこの特徴又は他の特徴において、オーディオ信号自身への及び２次的な干渉信号自身へのアクセスを仮定している。

【００１７】

他の特徴として、本発明は、その特定ラウドネスと目標特定ラウドネスとの差を縮めるために、オーディオ信号を修正することで、オーディオ信号の特定ラウドネスを制御するものである。

【００１８】

他の特徴として、本発明は、その部分的特定ラウドネスと目標特定ラウドネスとの差を縮めるために、オーディオ信号を修正することで、オーディオ信号の部分的特定ラウドネスを制御するものである。

【００１９】

目標特定ラウドネスがオーディオ信号の関数ではない場合、目標特定ラウドネスを保存又は受信してもよい。目標特定ラウドネスがオーディオ信号の関数ではない場合、前記修正又は前記導き出しにより、明示的又は黙示的に特定ラウドネス又は部分的特定ラウドネスを計算する。黙示的な計算の例には、特定ラウドネス及び／又は部分的特定ラウドネスを本来的に決定するルックアップテーブル又は「クローズドフォーム」の数学的表現が含まれる（クローズドフォームの用語は、指数関数や正弦関数のように、有限個の標準的な数学的演算及び関数を用いて正確に表現することのできる数学的表現の記述を意味する）。さらに、目標特定ラウドネスがオーディオ信号の関数ではない場合、目標特定ラウドネスは、時間及び周波数的に不変であっても、時間に対してのみ不変であってもよい。

【００２０】

さらに他の特徴として、本発明は、目標特定ラウドネスを生成するための１以上のプロセスパラメータ及び１以上のプロセス制御パラメータに従い、オーディオ信号を処理すること又はオーディオ信号を計測することによりオーディオ信号を処理するものである。目標特定ラウドネスが時間的に不変（「固定」）であっても、目標特定ラウドネスは、都合よく、オーディオ信号の特定ラウドネスの関数とすることができる。それが静的で、時間及び周波数的に不変であるかもしれないが、一般に、オーディオ信号自身は時間及び周波数的に変化し、したがって、目標特定ラウドネスがオーディオ信号の関数であるとき、目標特定ラウドネスは時間及び周波数的に変化する。

【００２１】

オーディオ及び目標特定ラウドネス又は目標特定ラウドネスを代理するものは、伝達により受け取り、又は記憶媒体から再生することができる。目標特定ラウドネスを代理するものは、オーディオ信号又はオーディオ信号の尺度を圧伸する１以上のスケールファクタとすることができる。

【００２２】

本発明の上記特徴を持つ目標特定ラウドネスは、オーディオ信号の関数又はオーディオ信号の尺度とすることができる。適切なオーディオ信号の尺度の１つが、オーディオ信号の目標特定ラウドネスである。オーディオ信号の関数又はオーディオ信号の尺度の関数は、オーディオ信号の圧伸又はオーディオ信号の尺度の圧伸とすることができる。例えば、この圧伸は、以下の圧伸のうちの１つ又はこれらの組合せとすることができる。

10

20

30

40

## 【数 1】

(a) 時間及び周波数に対して変化するスケールファクタ  $\Xi[b, t]$  は、以下のように特定ラウドネスを圧伸する。

$$\hat{N}[b, t] = \Xi[b, t] N[b, t]$$

## 【0023】

(b) 時間的に変化する、周波数的に不変であるスケールファクタ  $\Phi[t]$  は、以下のように特定ラウドネスを圧伸する。

10

## 【数 2】

$$\hat{N}[b, t] = \Phi[t] N[b, t];$$

## 【0024】

(c) 時間的に不変であり、周波数的に変化するスケールファクタ  $\Theta[b]$  は、以下のように特定ラウドネスを圧伸する。

## 【数 3】

$$\hat{N}[b, t] = \Theta[b] N[b, t]; \text{ and}$$

20

## 【0025】

(d) 時間的に不変であり、周波数的に不変であるスケールファクタ  $\alpha$  は、以下のように特定ラウドネスを圧伸する。

## 【数 4】

$$\hat{N}[b, t] = \alpha N[b, t],$$

## 【0026】

30

ここで、 $\hat{N}[b, t]$  は、目標特定ラウドネスであり、 $N[b, t]$  は、オーディオ信号の特定ラウドネスであり、 $b$  は周波数の指標を示し、 $t$  は時間の指標を示す。

## 【0027】

(a) の時間及び周波数的に変化するスケールファクタの場合、圧伸は、少なくとも部分的には、好ましい複数帯域音量のそのオーディオ信号の複数帯域音量に対する比により決定することができる。このような圧伸はダイナミックレンジ制御として利用することができる。ダイナミックレンジ制御として本発明の特徴を採用することについての詳細を以下に説明する。

40

## 【0028】

(a) の時間及び周波数的に変化するスケールファクタの場合はまた、特定ラウドネスは、好ましいスペクトル形状の指標とそのオーディオ信号のスペクトル形状の指標との比により圧伸することができる。このような圧伸は、時間的に変化する感知されたスペクトルから実質的に時間的に不変である感知されたスペクトルへと、感知されたオーディオ信号のスペクトルを変換するために採用される。特定ラウドネスが、好ましいスペクトル形状と指標のそのオーディオ信号のスペクトル形状の指標との比により圧伸された場合、このような圧伸は、ダイナミックイコライザーとして利用することができる。ダイナミックイコライザーとして本発明の特徴を採用することについての詳細を以下に説明する。

50

## 【 0 0 2 9 】

( b ) の時間的に変化し、周波数的に不変であるスケールファクタの場合、圧伸は、少なくとも部分的には、好ましい広帯域音量とそのオーディオ信号の広帯域音量との比により決定することができる。このような圧伸は、自動ゲイン制御又はダイナミックレンジ制御として利用することができる。自動ゲイン制御又はダイナミックレンジ制御として本発明の特徴を採用することについての詳細を以下に説明する。

## 【 0 0 3 0 】

( a ) の場合 ( 時間及び周波数的に変化するスケールファクタ ) 、又は、( b ) の場合 ( 時間的に変化し、周波数的に不変であるスケールファクタ ) 、スケールファクタは、オーディオ信号の関数又はオーディオ信号の指標の関数とすることができる。

10

## 【 0 0 3 1 】

( c ) の時間的に不変であり、周波数的に変化するスケールファクタの場合、又は ( d ) の時間的に不変であり、周波数的に不変であるスケールファクタの場合、両方の場合において、前記修正または前記導き出しには、スケールファクタの保存が含まれることがあり、又は、スケールファクタは外部から受信することができる。

## 【 0 0 3 2 】

( c ) 及び ( d ) のどちらも、スケールファクタはオーディオ信号の関数又はオーディオ信号の指標の関数でなくてもよい。

## 【 0 0 3 3 】

本発明の全ての特徴及びその修正したものにおいて、修正、導き出し、又は生成には、明示的に又は黙示的に ( 1 ) 特定ラウドネス、及び / 又は、( 2 ) 部分的特定ラウドネス、及び / 又は、( 3 ) 目標特定ラウドネスの計算を行う。黙示的な計算の例には、ルックアップテーブル又はクローズドフォームの数学的表現が含まれる。

20

## 【 0 0 3 4 】

修正パラメータは時間的に平滑化されてもよい。修正パラメータは、マルチタップの FIR フィルター又はマルチポールの HR フィルターのような、例えば、( 1 ) オーディオ信号の周波数帯域に関する複数の振幅スケールファクタ、又は ( 2 ) 1 以上のフィルターを制御する複数のフィルター係数とすることができる。このスケールファクタ又はフィルター係数 ( 及び、これらが用いられるフィルター ) は時間的に変化するものであってもよい。

30

## 【 0 0 3 5 】

目標特定ラウドネスを定義するオーディオ信号の特定ラウドネスの関数又はその逆演算の計算において、このような計算処理は、感知 ( 音響心理学的 ) 音量領域として特徴づけられるような処理、すなわち、計算の入力及び出力が特定ラウドネスとなるような処理で行われる。一方、オーディオ信号の周波数帯域に振幅スケールファクタを適用することにより、あるいは、オーディオ信号の制御可能なフィルタリング処理にフィルター係数を適用することにより、修正パラメータが作用して、電気信号領域として特徴づけられる感知 ( 音響心理学的 ) 音量領域の外のオーディオ信号を修正する。オーディオ信号の修正が電気信号領域で行われるかもしれないが、電気信号領域でのこのような変化は、修正されたオーディオ信号が好ましい目標特定ラウドネスに近似する特定ラウドネスような、感知 ( 音響心理学的 ) 音量領域における計算結果から導き出される。

40

## 【 0 0 3 6 】

音量領域での計算から修正パラメータを導き出すことにより、このような修正パラメータを電気信号領域で導き出すのに比べて、感知音量と感知スペクトルバランスに対する卓越した制御を行うことができる。加えて、音量領域での計算を行うに当たって基底膜をシミュレートする音響心理学的フィルターバンク又はその等価物を用いることにより、電気信号領域で修正パラメータを導き出す構成より、感知スペクトルの詳細な制御を行うことができる。

## 【 0 0 3 7 】

修正、導出、及び生成の各々は、干渉性のあるオーディオ信号の指標、目標特定ラウド

50

ネス、特定ラウドネスから導き出された修正されていないオーディオ信号又は修正されたオーディオ信号の部分特定ラウドネスの推定値、修正されていないオーディオ信号の特定ラウドネス、及び特定ラウドネス又は修正されたオーディオ信号の部分特定ラウドネスから導き出された目標特定ラウドネスの概算値のうちの1つ以上に依存する。

【0038】

この修正又は導出により、少なくとも部分的には、干渉性のあるオーディオ信号の指標、目標特定ラウドネス、特定ラウドネスから導き出された修正されていないオーディオ信号又は修正されたオーディオ信号の部分特定ラウドネスの推定値、修正されていないオーディオ信号の特定ラウドネス、及び特定ラウドネス又は修正されたオーディオ信号の部分特定ラウドネスから導き出された目標特定ラウドネスの概算値のうちの1つ以上から修正パラメータを導くことができる。

10

【0039】

もっと具体的には、この修正又は導出により、少なくとも部分的には、

(1) 目標特定ラウドネスと、

修正されたオーディオ信号の特定ラウドネスから受け取った修正されていないオーディオ信号の特定ラウドネスの推定値と、

のいずれか1つと、

(2) 修正されていないオーディオ信号の特定ラウドネスと、

修正されたオーディオ信号の特定ラウドネスから導き出された目標特定ラウドネスの概算値と、

20

のいずれか1つと、

から修正パラメータを導くことができ、

又は、

干渉性のあるオーディオ信号を考慮に入れるときは、この修正又は導出により、少なくとも部分的には、

(1) 干渉性のあるオーディオ信号の指標と、

(2) 目標特定ラウドネスと、

修正されたオーディオ信号の特定ラウドネスから受け取った修正されていないオーディオ信号の特定ラウドネスの推定値と、

のいずれか1つと、

30

(3) 修正されていないオーディオ信号の特定ラウドネスと、

修正されたオーディオ信号の特定ラウドネスから導き出された目標特定ラウドネスの概算値と、

のいずれか1つと、

から修正パラメータを導くことができる。

【0040】

オーディオ信号から特定ラウドネスから導き出され、目標特定ラウドネスがこの方法以外の音源から受け取られるか、又は、修正又は導出に目標特定ラウドネスの保存が含まれているときには保存したものから受け取られるようなフィードフォワード構成を採用することができる。代替的に、フィードフォワード/フィードバック構成を採用することもでき、ここでは、目標特定ラウドネスの推定が修正されたオーディオ信号から導き出され、目標特定ラウドネスがこの方法以外の音源から受け取られるか、又は、修正又は導出に目標特定ラウドネスの保存が含まれているときには保存したものから受け取られる。

40

【0041】

この修正又は導出には、明示的又は黙示的に、目標特定ラウドネスを取得する1つ以上の処理が含まれることがあり、この1つ以上の処理では、明示的又は黙示的に、オーディオ信号の関数又はオーディオ信号の指標の計算を行う。1つの代替的形態では、フィードフォワード構成を採用することができ、ここでは、特定ラウドネスと目標特定ラウドネスとは、オーディオ信号から導き出され、目標特定ラウドネスの導出ではオーディオ信号の関数又はオーディオ信号の指標を採用する。他の代替的形態では、フィードフォワード/

50

フィードバックのハイブリッド構成を採用することができ、ここでは、目標特定ラウドネスの概算値は修正されたオーディオ信号から導き出され、目標特定ラウドネスはオーディオ信号から導き出され、目標特定ラウドネスの導出ではオーディオ信号の関数又はオーディオ信号の指標を採用する。

【 0 0 4 2 】

この修正又は導出には、明示的又は黙示的に、修正されたオーディオ信号に応答して修正されていないオーディオ信号の特定ラウドネスの推定値を取得する1つ以上の処理が含まれることがあり、この1つ以上の処理では、明示的又は黙示的に、オーディオ信号の関数の逆演算又はオーディオ信号の指標の逆演算を行う。1つの代替的形態では、フィードバック構成を採用することができ、ここでは、修正されていないオーディオ信号の特定ラウドネスの推定値と目標特定ラウドネスの概算値は、修正されたオーディオ信号から導き出され、特定ラウドネスの推定値は、オーディオ信号の関数の逆演算又はオーディオ信号の指標の逆演算を用いて計算される。他の代替的形態では、フィードフォワード/フィードバックのハイブリッド構成を採用することができ、ここでは、特定ラウドネスは、オーディオ信号から導き出され、修正されていないオーディオ信号の特定ラウドネスの推定値は、修正されたオーディオ信号から導き出され、この推定値の導出はオーディオ信号の関数の逆演算又はオーディオ信号の指標の逆演算を用いて行われる。

【 0 0 4 3 】

修正されたオーディオ信号を生成させるために修正パラメータをオーディオ信号に適用することができる。

【 0 0 4 4 】

本発明の他の特徴は、実質的なエンコーダ又はエンコーディング処理及びデコーダ又はデコーディング処理となるように、処理又は装置の時間的及び/又は空間的分割ができるという点である。例えば、エンコーディング/デコーディングシステムがあり、そこでは、修正又は導出では、伝達及び受信又は保存を行い、さらに、オーディオ信号と(1)修正パラメータ又は(2)目標特定ラウドネス又は目標特定ラウドネスを代理するものを複製する。代替的に、実質的にエンコーダ又はエンコーディング処理のみがあり、そこでは、オーディオ信号及び(1)修正パラメータ又は(2)目標特定ラウドネス又は目標特定ラウドネスを代理するものの伝達又は保存を行う。代替的に、上述のように、デコーダ又はデコーディング処理のみがあり、そこでは、オーディオ信号及び(1)修正パラメータ又は(2)目標特定ラウドネス又は目標特定ラウドネスを代理するものの受信及び複製を行う。

【 発明を実施するための最良の形態 】

【 0 0 4 5 】

図1から4は、本発明の特徴による実行可能なフィードフォワード構成、フィードバック構成、及びフィードフォワード/フィードバックの2つのハイブリッド構成の実施例を示す機能ブロックダイアグラムである。

【 0 0 4 6 】

図1のフィードフォワードトポロジーの例を参照すると、オーディオ信号は2つの経路に用いられる。すなわち、(1)修正パラメータに応答してオーディオを修正することのできる処理又は装置2(「オーディオ信号の修正」)を有する信号経路、と(2)このような修正パラメータを生成することのできる処理又は装置4(「修正パラメータの生成」)を有する信号経路、である。図1のフィードフォワードトポロジーの例と図2~4の例における修正パラメータの生成2は、修正パラメータの生成4(図2~4の例ではそれぞれ、コンピュータ処理又はコンピュータ4'、4''、4''')から受け取った修正パラメータMに応じて周波数及び/又は時間を変更する方法で、オーディオ信号、例えば、その振幅を修正する処理又は装置とすることができる。修正パラメータの生成4とそれに相当する図2~4のコンピュータは、それぞれ少なくとも部分的には感知音量領域で動作する。オーディオ信号の修正2は、電気信号領域で動作し、図1~4の各例における修正されたオーディオ信号を生成する。図1~4の例においてもオーディオ信号の修正2と修正

パラメータの生成 4（又はそれに相当するコンピュータ）はオーディオ信号を修正し、その特定ラウドネスと目標特定ラウドネスとの差を減少させる。

【0047】

図1のフィードフォワードの例では、処理又は装置4には、いくつかの処理及び/又装置が含まれる。すなわち、オーディオ信号又はオーディオ信号の特定ラウドネスのようなオーディオ信号の指標に応答して目標特定ラウドネスを計算する「目標特定ラウドネスの計算」処理又は装置と、オーディオ信号又はそのエキサイテーションのようなオーディオ信号の指標に応答して特定ラウドネスを計算する「特定ラウドネスの計算」処理又は装置8と、特定ラウドネスと目標特定ラウドネスとに応答して修正パラメータを計算する「修正パラメータの計算」処理又は装置10と、である。目標特定ラウドネスの計算6は、それぞれが関数パラメータを持つことのできる1以上の関数「F」を実行することができる。例えば、オーディオ信号の特定ラウドネスを計算してそれに1以上の関数Fを適用して目標特定ラウドネスを算出する。これは図1に図解的に示したように、「関数Fと関数パラメータの選択」として処理又は装置6に入力される。処理又は装置6で計算する代わりに、目標特定ラウドネスは、修正パラメータの生成4に含まれるかあるいはこれに関連づけられた（図解的に「保存」として処理又は装置10に入力する）保存処理又は保存装置、又は全体的な処理又は装置の外部の供給源（図解的に「外部」として処理又は装置10に入力する）から提供されてもよい。このように、修正パラメータは、少なくとも部分的には、感知（音響心理学的）音量領域での計算（すなわち、少なくとも特定ラウドネスと、場合によって、目標特定ラウドネスの計算）に基づく。

【0048】

処理又は装置6、8及び10（及び、図2の例では処理又は装置12、14、10'、図3の例では処理又は装置6、14、10''、及び、図4の例では処理又は装置8、12、10'''）によりなされる計算は、明示的及び/又は黙示的に行うことができる。黙示的に行う例には、（1）項目が特定ラウドネス及び/又は目標特定ラウドネス及び/又は修正パラメータの全体又は部分的に基づいているルックアップテーブルと、（2）特定ラウドネス及び/又は目標特定ラウドネス及び/又は修正パラメータの全体又は部分的に本質的に基づいている閉形式の数学的表現と、が含まれる。

【0049】

図1の例の計算処理又は装置6、8及び10（及び、図2の例では処理又は装置12、14、10'、図3の例では処理又は装置6、14、10''、及び、図4の例では処理又は装置8、12、10'''）は、図解的に示され、ここに説明されているが、これは説明のためだけのものである。これらの処理又は装置のいくつか又は全てを、単一の処理又は装置あるいは、複数の処理又は装置の様々な結合に統合してもよいことは了解されよう。例えば、以下の図9の構成、すなわち、図1の例におけるフィードフォワードトポロジーにおいて、修正パラメータを計算する処理又は装置では、オーディオ信号と目標特定ラウドネスとから導き出された平滑化されたエキサイテーションに応答してそれを行う。図9の例では、修正パラメータを計算する処理又は装置では、オーディオ信号の特定ラウドネスを黙示的に計算する。

【数5】

本発明の特徴として、図1の例及びここに記載した他の実施の形態の例において、目標特定ラウドネス（ $\hat{N}[b, t]$ ）は、1以上のスケールファクタにより圧伸することにより計算される。この圧伸は、時間及び周波数に対して変化するスケールファクタ $\Xi[b, t]$ は、以下の関係を持つ特定ラウドネスの圧伸とすることができる。

$$\hat{N}[b, t] = \Xi[b, t] N[b, t]$$

【0050】

時間的に変化する、周波数的に不変であるスケールファクタ  $[t]$  は、以下の関係を持つ特定ラウドネスの圧伸となる、

【数 6】

$$\hat{N}[b, t] = \Phi[t]N[b, t],$$

【0051】

時間的に不変で、周波数的に変化するスケールファクタ  $[b]$  は、以下の関係を持つ特定ラウドネスの圧伸となる、

【数 7】

$$\hat{N}[b, t] = \Theta[b]N[b, t], \text{ or}$$

【0052】

スケールファクタ  $\alpha$  は、以下の関係を持つオーディオ信号の特定ラウドネスの圧伸となる、

【数 8】

$$\hat{N}[b, t] = \alpha N[b, t],$$

【0053】

ここで、 $b$  は周波数の尺度であり（例えば帯域数）、 $t$  は時間の尺度である（例えば、ブロック数）。特定の圧伸及び / 又は特定の圧伸の組み合わせの事例を用いて複数の圧伸を採用することができる。このような複数の圧伸の例を以下に示す。場合によって、以下にさらに説明するように、この圧伸はオーディオ信号又はオーディオ信号の指標の関数とすることができる。他の場合では、以下にさらに説明するように、圧伸がオーディオ信号の指標の関数でないとき、圧伸は、別の方法で定められるか又は提供される。例えば、時間と周波数的に不変なスケールファクタ  $\alpha$ 、又は、時間的に不変で、周波数的に変化するスケールファクタ  $[b]$  での圧伸を選択又は適用することができる。

【0054】

このようにして、目標特定ラウドネスはオーディオ信号の 1 以上の関数又はオーディオ信号の指標を表現することができる（特定ラウドネスはオーディオ信号の 1 つの指標である）。

【数 9】

$$\hat{N}[b, t] = F(N[b, t])$$

関数  $F$  が可逆であることを条件に、特定ラウドネス（修正されていないオーディオ信号の  $N[b, t]$ ）を目標特定ラウドネス（ $\hat{N}[b, t]$ ）の逆関数すなわち関数  $F^{-1}$  として計算することができる。

$$N[b, t] = F^{-1}(\hat{N}[b, t])$$

【0055】

以下に示すように、逆関数すなわち関数  $F^{-1}$  は、図 2 及び 4 の例におけるフィードバック構成及びフィードフォワード / フィードバックのハイブリッド構成において計算される。

【0056】

目標特定ラウドネスの計算 6 へ入力する「関数と関数パラメータの選択」は、1 以上の関数パラメータに従い 1 以上の関数を適用することにより装置又は処理 6 は特定ラウドネ

10

20

30

40

50

スの計算を行うことができることを示すために記載されている。例えば、特定ラウドネスの計算 8 は、目標特定ラウドネスを定義するためにオーディオ信号の特定ラウドネスの関数「F」を計算することができる。例えば、「関数と関数パラメータの選択」入力は、その関数に関連する定数（例えばスケールファクタ）のような、1 以上の関数パラメータと共に、1 以上の上記形式の圧伸に該当する 1 以上の特定の関数を選択することができる。

【0057】

上述の通り、圧伸に係するスケールファクタは、目標特定ラウドネスが特定ラウドネスの圧伸として計算される限り、目標特定ラウドネスを代理するものとして機能を果たす。このように、図 9 の例において、以下に説明し上記したように、特定ラウドネスと目標特定ラウドネスがテーブル内で固有であるルックアップテーブルは、スケールファクタ及びエキサイテーションにより索引付けることができる。

【0058】

ルックアップテーブル、閉形式の数学的表現、あるいは他の技術を採用したとしても、修正パラメータ 4（及び各々図 2～4 の例におけるそれに対応する処理又は装置 4'，4''，及び 4'''）は、特定ラウドネス及び目標特定ラウドネスが明示的に計算されないとしても、感知（音響心理学的）音量領域で計算される。そこには明示的な特定ラウドネス又は、概念的黙示的な特定ラウドネスが存在する。同様に、そこには明示的な目標特定ラウドネス又は、概念的黙示的な目標特定ラウドネスが存在する。いずれにせよ、修正パラメータの計算は、修正パラメータを生成し、特定ラウドネスと目標特定ラウドネスとの差を減少させるようオーディオ信号を修正することを目的とする。

【0059】

ノイズのような 2 次的な干渉性のあるオーディオ信号を有する再生環境において、修正パラメータの計算 10（及び各々図 2～4 の例におけるそれに対応する処理又は装置 10'，10''，及び 10'''）は、付加的な入力として、2 次的な干渉性のあるオーディオ信号の指標又は 2 次的な干渉信号それ自身もその入力として受け取ることができる。このような付加的な入力とは図 1（及び図 2～4）に破線で示されている。2 次的な干渉性のあるオーディオ信号の指標は、以下に説明する、図 17 の例におけるそのエキサイテーションでもよい。修正パラメータの計算 10（及び各々図 2～4 の例におけるそれぞれに対応する処理又は装置 10'，10''，及び 10'''）に対する干渉性のある信号の指標又は信号それ自身（この干渉性のある信号は別に処理することが可能であると仮定する）の処理は、以下に表題「ノイズ補償」でさらに説明するように干渉性信号を考慮に入れた修正パラメータを計算するよう適切に構成した処理又は装置を許容する。図 2～4 の例において、部分特定ラウドネスの計算では、適切な干渉性信号の指標をそれぞれ修正パラメータの処理又は装置 10'，10''，又は 10''' に適用するのみならず、この処理又は装置によって、部分特定ラウドネスの計算を容易にするために「修正されていないオーディオの特定ラウドネスの近似値の計算」処理又は装置 12 及び／又は「目標特定ラウドネスの近似値の計算」処理又は装置 14 にも用いられる。図 1 のフィードフォワード例において、部分特定ラウドネスは明示的に計算されない。すなわち、図 1 の修正パラメータの計算は、修正されたオーディオの部分特定ラウドネスを目標特定ラウドネスに近似させるために、修正パラメータの近似値を計算する。これについては、先に言及した表題「ノイズ補償」でさらに以下に説明する。

【0060】

上述のように、各図 1～4 の例において、修正パラメータ M は、信号修正演算 2 をオーディオ信号に適用するとき、オーディオ修正の結果生じた特定ラウドネス又は部分特定ラウドネスと、目標特定ラウドネスとの差を縮小させる。修正されたオーディオ信号の特定ラウドネスが目標特定ラウドネスに極めて近似するか又は同じとなることが理想である。修正パラメータ M は、例えば、フィルタバンクから導き出された周波数帯域又は時間的に変動するフィルタの係数に適用する時間的に変動するゲインの形をとる。したがって、図 1～4 の全ての例において、オーディオ信号の修正 2 は、例えば、それぞれ周波数低域で動作する複数の振幅圧伸器又は時間的に変化するフィルタ（例えば、複数タップ



のFIRフィルタ又は複数極のIIRフィルタ)として実施してもよい。

【0061】

ここで、又は本明細書の他の部分で、同じ参照番号の使用しているものは、その装置又は処理が実質的に同じ参照番号を持つものと同じであることを示す。主要な番号(例えば「10」)の参照番号は、その装置又は処理が構成又は機能において類似するが、同一の基本番号又はその最上位桁が同一の番号を持つ他のものを修正していることを示す。

【0062】

ある構成では、図1のフィードフォワード例とほぼ等しいフィードバック構成を実現することができる。図2は、信号経路中のオーディオ信号の修正装置又は処理2にもオーディオ信号が用いられるような例を示す。この装置又は処理2は、フィードバック構成において修正パラメータの生成装置又は処理4'がその入力として、オーディオ信号の修正装置又は処理2の出力から修正されたオーディオ信号も受け取る。このようにして、図2の例では、修正されていないオーディオ信号ではなく修正されたオーディオ信号が制御経路に用いられる。オーディオ信号の修正装置又は処理2と修正パラメータの生成装置又は処理4'とは、オーディオ信号を修正し、その特定ラウドネスと目標特定ラウドネスとの差を減少させる。処理又は装置4'には、いくつかの機能又は装置を含むことができる。すなわち、「修正されていないオーディオの特定ラウドネスの近似値の計算」処理又は装置12、「目標特定ラウドネスの近似値の計算」処理又は装置14、及び修正パラメータを計算する「修正パラメータの計算」処理又は装置10'である。

【0063】

関数Fが可逆であるという条件で、処理又は装置12は、修正されたオーディオ信号の特定ラウドネス又は部分特定ラウドネスに逆関数 $F^{-1}$ を適用することにより、修正されていないオーディオ信号の特定ラウドネスを推定する。処理又は装置12は、上述のように、逆関数 $F^{-1}$ を計算することができる。これは、図2に処理又は装置12への入力「逆関数 $F^{-1}$ と逆関数パラメータの選択」として概略的に示されている。「目標特定ラウドネスの近似値の計算」14は、修正されたオーディオ信号の、特定ラウドネス又は部分特定ラウドネスの計算により動作する。このような特定ラウドネス又は部分特定ラウドネスは、目標特定ラウドネスの近似値となる。修正されていないオーディオ信号の特定ラウドネス及び目標特定ラウドネスの近似値は、修正パラメータの計算10'で使われ、オーディオ信号の修正2に適用された場合は、修正されたオーディオ信号の特定ラウドネス又は部分特定ラウドネスと目標特定ラウドネスとの差を減少させる。上述の通り、修正パラメータMは、例えば、フィルタバンクの周波数帯域に用いられる時間的に変化するゲイン又は時間的に変化するフィルタの係数の形を取る。修正パラメータの計算10'の実際の形態において、フィードバックループでは、修正パラメータの計算と適用の間に時間遅れを導入することもできる。

【0064】

上述したように、雑音のような2次的な干渉性のあるオーディオ信号を有する再生環境において、修正パラメータの計算10'、目標特定ラウドネスの近似値の計算12、及び目標特定ラウドネスの近似値の計算14ではそれぞれ、付加的な入力として、2次的な干渉性のあるオーディオ信号又はその入力の1つとしての2次的な干渉性のある信号自身のような指標を受け取ることができる。処理又は装置12及び処理又は装置14は、それぞれ修正されたオーディオ信号の部分特定ラウドネスの計算を行うことができる。このような付加的な入力図2に破線で示されている。

【0065】

上述したように、本発明のフィードフォワード/フィードバックのハイブリッド構成の実施例が可能である。図3及び4は、このような実施の形態の2つの例を示す。図3及び4の例において、図1及び2の例のように、オーディオ信号はやはり信号経路中のオーディオ信号の修正処理又は装置2に用いられるが、それぞれの制御経路中の修正パラメータの生成(図3における4'、図4における4'')では修正されていないオーディオ信号と修正されたオーディオ信号の両方を受信する。図3及び4の両方において、オーデ

ィオ信号の修正 2 と修正パラメータの生成（それぞれ 4' と 4' '）では、黙示的であることもある特定ラウドネスと、これも黙示的であることもある目標特定ラウドネスとの差を減少させるためにオーディオ信号を修正する。

【 0 0 6 6 】

図 3 の例では、修正パラメータの生成処理又は装置 4' には、いくつかの機能と装置を含むことができる。すなわち、図 1 の例のような目標特定ラウドネスの計算 6、図 2 のフィードバック例のような、目標特定ラウドネスの近似値の計算 1 4、及び「修正パラメータの計算」処理又は装置 1 0' ' である。図 1 の例のように、このフィードフォワード/フィードバックのハイブリッド構成例のフィードフォワード部分において、目標特定ラウドネスの計算 6 では、それぞれが関数パラメータを有することができる 1 以上の関数「F」を実行する。これは図 3 に処理又は装置 6 に入力する「関数 F と関数パラメータの選択」として図解的に示されている。このフィードフォワード/フィードバックのハイブリッド構成例のフィードバック部分において、図 2 のフィードバックの例のように、修正されたオーディオ信号が、目標特定ラウドネスの近似値の計算 1 4 に用いられる。処理又は装置 1 4 は、修正されたオーディオ信号の特定ラウドネス又は部分特定ラウドネスを計算することにより、図 3 の例において、図 2 の例で行ったような動作を行う。このような特定ラウドネス又は部分特定ラウドネスは、目標特定ラウドネスの 1 つの近似値となる。（処理又は装置 6 からの）目標特定ラウドネスと（処理又は装置 1 4 からの）目標特定ラウドネスの近似値とは、修正パラメータの計算 1 0' ' に適用され、修正パラメータ M を導き出す。修正パラメータは、オーディオ信号の修正 2 によりオーディオ信号に用いられると、修正されないオーディオ信号の特定ラウドネスと目標特定ラウドネスとの差を減少させる。上述のように、これらの修正パラメータ M は、例えば、フィルタバンクの周波数帯域に用いられる時間的に変化するゲイン又は時間的に変化するフィルタの係数の形を取る。実際の実施の形態において、フィードバックループでは、修正パラメータ M の、計算と適用との間に時間遅れを導入することもできる。上述したように、ノイズのような 2 次的な干渉性のあるオーディオ信号を有する再生環境において、修正パラメータの計算 1 0' ' 及び目標特定ラウドネスの近似値の計算 1 4 ではそれぞれまた、2 次的な干渉性のあるオーディオ信号又はその入力の 1 つとしての 2 次的な干渉性のある信号自身のような指標を受け取ることができ、処理又は装置 1 4 では、修正されたオーディオ信号の部分特定ラウドネスを計算することができる。このような付加的な入力図 3 において破線で示されている。

【 0 0 6 7 】

修正パラメータの計算 1 0' ' では、目標特定ラウドネスと目標特定ラウドネス近似値入力との差が、目標特定ラウドネスの近似値と「実際の」目標特定ラウドネスとの差を減少させるように修正パラメータを調整するように、誤差検出装置又は機能を採用することができる。このような調整は、黙示的かもしれないが、修正されていないオーディオ信号の特定ラウドネスと目標特定ラウドネスとの差を減少させる。このようにして、修正パラメータ M は、関数 F を用いて元のオーディオの特定ラウドネスからのフィードフォワード経路で計算された目標特定ラウドネスと、修正されたオーディオの特定ラウドネス又は部分特定ラウドネスからのフィードバック経路で計算された目標特定ラウドネスの近似値との誤差に基づいて修正してもよい。

【 0 0 6 8 】

図 4 の例では、代替的なフィードフォワード/フィードバック構成例が示されている。この代替案は、フィードフォワード経路で関数 F が計算されるのではなく、フィードバック経路で逆関数  $F^{-1}$  が計算される点で図 3 の例とは異なる。図 4 の例では、修正パラメータの生成処理又は装置 4' には、いくつかの関数と装置とを含むことができる。すなわち、図 1 のフィードフォワード例のような特定ラウドネスの計算 8、図 2 のフィードバック例のような修正されていないオーディオの特定ラウドネスの近似値の計算 1 2、及び修正パラメータの計算 1 0' ' ' である。図 1 のフィードフォワード例のような特定ラウドネスの計算 8 では、修正パラメータの計算 1 0' ' ' の入力として、修正されないオーデ

ィオ信号の特定ラウドネスを与える。図2のフィードバック例のように、関数 $F$ が可逆であることを条件として、処理又は装置12は、修正されたオーディオ信号に特定ラウドネス又は部分特定ラウドネス逆関数 $F^{-1}$ を適用することにより、修正されていないオーディオ信号の特定ラウドネスを推定する。修正されていないオーディオの特定ラウドネスの近似値の計算処理又は装置12に入力される「逆関数及び逆関数パラメータの選択」は、処理又は装置12が上述のように逆関数 $F^{-1}$ を計算することのできることを示すために示されている。これは、図4において、処理又は装置12への「逆関数 $F^{-1}$ 逆関数パラメータの選択」入力として図解的に示されている。このように、処理又は装置12は、もう1つの修正パラメータの計算10' ' 'の入力として、修正されていないオーディオ信号の特定ラウドネスの近似値を提供する。

10

#### 【0069】

図1~3の例のように、修正パラメータの計算10' ' 'では、この例では黙示的となるが、オーディオ信号の修正2に適用すると、修正されていないオーディオ信号の特定ラウドネスと目標特定ラウドネスとの差を減少させる修正パラメータ $M$ を導き出す。上述のように、修正パラメータ $M$ は、例えば、フィルタバンクの周波数帯域に用いられる時間的に変化するゲイン又は時間的に変化するフィルタの係数の形を取る。実際の実施の形態において、フィードバックループでは、修正パラメータ $M$ の、計算と適用との間に時間遅れを導入することもできる。上述のように、ノイズのような2次的な干渉性のあるオーディオ信号を有する再生環境において、修正パラメータの計算10' ' '及び修正されていないオーディオの特定ラウドネスの近似値の計算12ではそれぞれまた、付加的な入力として、2次的な干渉性のあるオーディオ信号又はその入力の1つとしての2次的な干渉性のある信号自身のような指標を受け取ることができ、処理又は装置12では、修正されたオーディオ信号の部分特定ラウドネスを計算することができる。このような付加的な入力は図4において破線で示されている。

20

#### 【0070】

修正パラメータの計算10' ' 'では、特定ラウドネスと特定ラウドネス近似値入力との差が、特定ラウドネスの近似値と「実際の」特定ラウドネスとの差を減少させるように修正パラメータを調整する出力を生じさせるように、誤差検出装置又は機能を採用することができる。目標特定ラウドネスの近似値と考えることのできる、特定ラウドネスの近似値は、修正されたオーディオの特定ラウドネス又は部分特定ラウドネスから導き出されるので、本質的に関数 $F^{-1}$ であるこのような調整により、修正されたオーディオの特定ラウドネスと目標特定ラウドネスとの差を減少させる。このようにして、修正パラメータ $M$ は、特定ラウドネス同士の差に基づいて更新され、元のオーディオからのフィードフォワード経路で計算され、そして、特定ラウドネスの近似値は、逆関数 $F^{-1}$ を用いて修正されたオーディオの特定ラウドネス又は部分特定ラウドネスからのフィードバック経路で計算される。フィードバック経路により、現実的な実施例では修正パラメータの、更新と適用との間に時間遅れを導入する。

30

#### 【0071】

図1~4の例における修正パラメータ $M$ は、オーディオ信号の修正処理又は装置2に適用されたとき、オーディオ信号の特定ラウドネスと目標特定ラウドネスとの差を減少せるとはいえ、現実的な実施例では、同じオーディオ信号に応答して生じた対応する修正パラメータはお互いに同じである必要はない。

40

#### 【0072】

本発明の重要で本質的な特徴ではないが、オーディオ信号又は修正されたオーディオ信号の特定ラウドネスの計算では、前記国際特許出願No. PCT/US2004/016964であって、WO2004/111964A2として国際公開された技術を採用することが有利であり、ここでは、計算において、2以上の特定ラウドネスモデル関数のグループ、2以上の特定ラウドネスモデル関数の1つ又はその組み合わせ、から選択し、その選択は、入力オーディオ信号の特性の指標により制御される。図1の特定ラウドネスの記述104は、以下のように、このような構成について説明する。

50

## 【0073】

本発明のさらなる特徴によれば、修正されていないオーディオ信号及び、（１）修正パラメータ、又は、（２）目標特定ラウドネス又は目標特定ラウドネスを代理するもの（例えば、目標特定ラウドネスを、明示的又は黙示的に計算するのに使うことのできるスケールファクタ）のどちらか一方を、例えば時間的に及び／又は空間的に分離した装置又は処理において使用するために、保存又は伝達することができる。修正パラメータ又は、目標特定ラウドネス又は、目標特定ラウドネスを代理するものは、例えば、上述の図１～４のフィードフォワード構成、フィードバック構成、及びフィードフォワードとフィードバックのハイブリッド構成のうちの１つような、適切な方法で決定することができる。実際には、図１の例のようなフィードフォワード構成が、修正されたオーディオ信号に基づく計算を避ける限りにおいて、最も簡単で速い。修正されていないオーディオと修正パラメータの伝達と保存の例を図５に示し、修正されていないオーディオと目標特定ラウドネス又は目標特定ラウドネスを代理するものの伝達と保存の例を図６に示す。

10

## 【0074】

図５の例のような構成は、オーディオ信号への修正パラメータの適用とこのような修正パラメータの生成とを時間的に及び／又は空間的に分離するために用いることができる。図６の例のような構成は、修正パラメータの生成と適用の両方と、目標特定ラウドネス又は目標特定ラウドネスを代理するものとを時間的に及び／又は空間的に分離するために用いることができる。両方の構成形式により、修正パラメータの生成や目標特定ラウドネスの生成における複雑さを避けた単純な低価格の再現又は受信構成が可能となる。図５の形式の構成のほうが図６の形式の構成より簡単であるが、図６の構成は、保存又は伝達に必要とされる情報がより少なく、特に、１以上のスケールファクタのような目標特定ラウドネスを代理するものが保存又は伝達されるとき、有利となる。このような、情報の保存と伝達の削減は、低ビットレートのオーディオ環境で特に有用である。

20

## 【0075】

したがって、本発明のさらなる特徴によれば、（１）保存又は伝達装置又は処理から修正パラメータＭを受け取り、それを、受け取ったオーディオ信号に適用する装置又は処理、又は、（２）保存又は伝達装置又は処理から、目標特定ラウドネス又は目標特定ラウドネスを代理するものを受け取り、受け取ったオーディオ信号に（又は、このオーディオ信号～導き出される特定ラウドネスのようなオーディオ信号の指標に）この目標特定ラウドネス又はそれを代理するものを適用することにより、修正パラメータＭを生成し、この修正パラメータＭをオーディオ信号に適用する装置又は処理を提供するものである。このような装置又は処理は、デコーディング処理又はデコーダとして特徴づけることができ、保存又は伝達された情報を生成するために必要な装置又は処理は、エンコーディング処理又はエンコーダとして特徴づけることができる。このようなエンコーディング処理又はエンコーダは、それぞれのデコーディング処理又はデコーダで必要とされる情報を生成するのに有効な、図１～４の構成例の一部となる。このようなデコーディング処理又はデコーダは、サウンドを処理及び／又は再生する事実上すべての装置又は処理に関係付けること又は動作させることができる。

30

## 【0076】

本発明の１つの特徴によれば、図５の例において、例えば、図１の修正パラメータの生成４、図２の４'、図３の４''、図４の４'''のような、修正パラメータの生成処理又は生成装置により作られた修正されていないオーディオ信号及び修正パラメータＭは、適切な保存装置又は伝達装置（「保存又は伝達」）１６に適用することができる。エンコーディング処理又はエンコーダとして図１のフィードフォワード例を用いる場合、オーディオ信号の修正２は、修正されたオーディオを生成するためには必要なく、エンコーダ又はエンコーディング処理の時間的又は空間的位置にこの修正されたオーディオを供給する必要がなければ削除することができる。保存又は伝達１６には、例えば、適切な磁気記憶再生装置、光学的記憶装置再生装置、又は、半導体記憶再生装置、あるいは、適切な電線による伝達受信装置、ワイヤレス伝達受信装置を含めることができるが、これらの選択は

40

50

本発明に重要ではない。再生した修正パラメータ又は受信した修正パラメータは、次いで、その特定ラウドネスが、修正パラメータが導き出された構成における目標特定ラウドネス又はこの構成において本質的な目標特定ラウドネスに近似するように、再生した修正パラメータ又は受信した修正パラメータを修正するために、図 1 ~ 4 の例において採用されたオーディオ信号の修正 2 に用いられる。この修正パラメータはさまざまな方法で保存又は伝達することができる。例えば、これらはオーディオ信号に伴うメタデータとして保存又は伝達することができ、これらは、別の経路又はチャンネルで送ってもよく、オーディオ中に電子迷彩技術によりエンコードしてもよく、マルチプレックスしてもよい。オーディオ信号を修正する修正パラメータの使用は任意としてもよく、もし任意とするならば、例えばユーザによりその使用を選択できるようにしてもよい。例えば、オーディオ信号に修正パラメータが用いられる場合は、オーディオ信号のダイナミックレンジを下げるかもしれない。このようなダイナミックレンジを採用するかどうかはユーザにより選択可能にすることができる。

【 0 0 7 7 】

図 6 の例のように、本発明の他の特徴によれば、修正されていないオーディオ信号と目標特定ラウドネス又は目標特定ラウドネスを代理するものは、適切な保存又は伝達装置又は機能（「伝達又は保存」）16 に適用することができる。図 1 の例のように、エンコーディング処理又はエンコーダとしてフィードフォワード構成を用いる場合は、修正パラメータの計算型の処理又は装置もオーディオ信号の修正型の処理又は装置も必要ではなく、エンコーダ又はエンコーディング処理の時間的又は空間的位置に修正パラメータ又は修正されたオーディオを供給することを必要としないならば、削除してもよい。図 5 の例の場合のように、保存又は伝達 16 には、例えば、適切な磁気記憶再生装置、光学的記憶装置再生装置、又は、半導体記憶再生装置、あるいは、適切な電線による伝達受信装置、ワイヤレス伝達受信装置を含めることができるが、これらの選択は本発明に重要ではない。再生した目標特定ラウドネス又は目標特定ラウドネスを代理するもの又は受信した目標特定ラウドネス又は目標特定ラウドネスを代理するものは、次いで、図 1 ~ 4 の例において採用された形式のオーディオ信号の修正 2 に適用する修正パラメータ M を供給するために、再生したオーディオ信号又は受信したオーディオ信号の特定ラウドネスが、修正パラメータを導き出す構成において目標特定ラウドネスに近似するか又は本質的となるよう、再生したオーディオ信号又は受信したオーディオ信号を修正するために、図 1 の例で採用された形式の、修正パラメータの計算 10 に、又は図 3 の例で採用された形式の、修正パラメータの計算 10' ' に修正されていないオーディオと共に用いられる。目標特定ラウドネス又はその代理をするものが図 1 の例の形式のエンコーディング処理又はエンコーダで極簡単に取得できるかもしれないが、目標特定ラウドネス又はその代理をするもの又は目標特定ラウドネス又はその代理をするものの近似値は、図 2 から 4 までの例の形式のエンコーディング処理又はエンコーダで（近似値は図 2 及び 3 の処理又は装置 14、及び図 4 の処理又は装置 12 で）取得することができる。目標特定ラウドネス又はその代理をするものは、種々の方法で保存又は伝達される。例えば、これらはオーディオ信号に伴うメタデータとして保存又は伝達することができ、これらは、別の経路又はチャンネルで送ってもよく、オーディオ中に電子迷彩技術によりエンコードしてもよく、マルチプレックスしてもよい。オーディオ信号を修正するために、保存又は伝達された目標特定ラウドネス又はその代理をするものから導き出された修正パラメータの使用は任意としてもよく、もし任意とするならば、例えばユーザによりその使用を選択できるようにしてもよい。例えば、オーディオ信号に修正パラメータが用いられる場合は、オーディオ信号のダイナミックレンジを下げるかもしれない。このようなダイナミックレンジを採用するかどうかはユーザにより選択可能にすることができる。

【 0 0 7 8 】

開示された発明をデジタルシステムとして実施する場合は、フィードフォワード構成が最も実用的なので、そのような構成の例を以下に詳細に説明するが、これは本発明をそれに限定するものではないことは了解されよう。

## 【 0 0 7 9 】

本明細書を通じて、ここで「フィルター」又は「フィルターバンク」のような用語には、本質的にＩＩＲフィルター又は変換のようなあらゆる再帰的及び非再帰的なフィルター処理が含まれ、「フィルターされた」情報は、このようなフィルターを適用した結果得られたものである。以下に記載の実施の形態では、変換により実行されたフィルターバンクを採用する。

## 【 0 0 8 0 】

図 7 は、フィードフォワード構成で実施した本発明の特徴の模範的な実施の形態の詳細を示したものである。オーディオは、まず分析フィルターバンク機能又は装置（「分析フィルターバンク」）１００を通過し、そこでオーディオ信号を複数の周波数帯域に分割する（したがって、図 5 は、各出力が周波数低域を表す、分析フィルターバンク１００からの複数の出力を示し、その出力が種々の機能又は装置を通して合成フィルターバンクに伝わり、以下に説明するように、その帯域を加算して結合された広帯域の信号となる）。分析フィルターバンク１００における各周波数帯域に関連するフィルターの応答は、内耳における基底膜の特定の位置の応答をシミュレートするよう設計している。分析フィルターバンク１００の各フィルターの出力は、次に、外耳又は中耳を通るオーディオの伝達における効果のフィルター処理をシミュレートする伝達フィルター又は伝達フィルター機能（「伝達フィルター」）１０１に入る。このオーディオのラウドネスが計測されさえすれば、伝達フィルターを分析フィルターバンクに先立って適用することができるが、分析フィルターバンクの出力は、修正されたオーディオの合成に用いられるので、このフィルターバンクの後に伝達フィルターを適用する方が好都合である。伝達フィルター１０１の出力は、次に、エキサイテーション機能又は装置（「エキサイテーション」）１０２に入り、その出力は基底膜に沿ってのエネルギーの分布をシミュレートする。エキサイテーションエネルギー値は、平滑化機能又は装置（「平滑化」）１０３により時間的に平滑化される。平滑化機能の時定数は、必要なアプリケーションでの要求にしたがって設定される。平滑化されたエキサイテーション信号は、続いて、特定ラウドネス機能又は装置（「特定ラウドネス（ＳＬ）」）１０４の特定ラウドネスに変換される。特定ラウドネスは、ゾーン毎単位周波数の単位で表現される。各帯域に関連づけられたこの特定ラウドネスの成分は、特定ラウドネス修正機能又は装置（「ＳＬ修正」）１０５に入る。ＳＬ修正は元の特定ラウドネスをその入力とし、本発明の特徴に従い、元の特定ラウドネスの関数であることが好ましい（次の「目標特定ラウドネス」の見出しの表題を参照のこと）、好ましい又は「目標」特定ラウドネスを出力する。ＳＬ修正１０５は、各帯域で独立に動作することができ、好ましい効果によっては、帯域間で相互依存が存在していてもよい（図 7 の相互接続ラインで示唆したような周波数平滑化）。入力としてエキサイテーション１０２から平滑化されたエキサイテーション周波数帯域と、ＳＬ修正１０５から目標特定ラウドネスとを取り込み、ゲインソルバー機能又は装置（「ゲインソルバー」）１０６は、計測した特定ラウドネスを目標特定ラウドネスに変換するために、分析フィルターバンク１００の出力の各大域に適用する必要があるゲインを決定する。ゲインソルバーはさまざまな方法で実行することができる。例えば、ゲインソルバーには、前記国際特許出願 No. PCT / US 2004 / 016964 であって、WO 2004 / 111964 A2 として国際公開されたものに開示された方法のような反復的な方法、又は、代替的にルックアップテーブルを見る方法が含まれる。ゲインソルバー１０６により作られた帯域毎のゲインを、感知できるアーティファクトを最小限にするために任意的な平滑化機能又は装置（「平滑化」）１０７により時間的にさらに平滑化することができるが、他で説明したように、処理又は装置全体の他の場所で時間的平滑化を適用することが好ましい。最後に、それぞれの乗算結合機能又は結合器１０８を介して分析フィルターバンク１００のそれぞれの帯域にゲインが適用され、処理又は「修正された」オーディオは、合成フィルターバンク機能又は装置（「合成フィルターバンク」）１１０において、ゲイン修正された帯域から合成される。加えて、分析フィルターバンクからの出力は、ゲインの計算における待ち時間を埋め合わせるために、ゲインを適用する前に時間遅れ機能又は装置（「時間遅れ」）１０９に

10

20

30

40

50

より時間遅れが加えられる。代替的に、周波数帯域でのゲイン修正の適用に用いるためにゲインを計算する代わりに、ゲインソルバー 106 は、複数次ップの F I R フィルターまたは複数次極の I I R フィルターのような時間的に変化するフィルターを制御するフィルター係数を計算してもよい。説明を簡単にするために、本発明の特徴は主に周波数帯域に用いられるゲイン要素を採用するような記載となっているが、このフィルター係数と時間的に変化するフィルターを実際的な実施の形態に採用することもできる。

【0081】

実際的な実施の形態において、オーディオの処理は、デジタル領域で実行するかもしれない。したがって、オーディオ入力信号は、オーディオ源からサンプリング周波数  $f_s$  でサンプルされた離散的な時間シーケンス  $x[n]$  で表される。下式で与えられる  $x[n]$  の  $r$  m s パワーが人間のリスナに聴取される d B 単位の音圧レベルに等しくなるように、シーケンス  $x[n]$  は適切に圧伸されていると仮定する。

【数10】

$$RMS_{dB} = 10 \log_{10} \left( \frac{1}{L} \sum_{n=0}^L x^2[n] \right)$$

【0082】

加えて、オーディオ信号は説明を簡単にするためにモノラルであると仮定する。

【0083】

分析フィルターバンク 100、伝達フィルター 101、エキサイテーション 102、特定ラウドネス 104、特定ラウドネス修正 105、ゲインソルバー 106、及びフィルターバンク 110 については以下に詳細を説明する。

【0084】

[分析フィルターバンク 100]

オーディオ入力信号は分析フィルターバンク又は機能(「分析フィルターバンク」) 100 に用いられる。分析フィルターバンク 100 の各フィルターは、内耳の基底膜に沿う特定の位置での周波数応答をシミュレートするよう設計されている。分析フィルターバンク 100 には、Moore と Glasberg と Baer が定義したような (B. C. J. Moore, B. Glasberg, T. Baer, 「閾値音量と部分音量の予想のためのモデル」参照)、等価矩形帯域幅 (ERB) で帯域幅と間隔が一定の線形フィルターが含まれる。

【0085】

ERB 周波数スケールは、人間の知覚にしっかりと合致しており、より主観的な音量に合致することとなる客観的な音量計測を行う上で性能の改善を示しているが、性能を下げたバーク周波数スケールを適用してもよい。

【0086】

ヘルツ単位の中心周波数  $f$ 、ヘルツ単位の ERB の 1 つの帯域幅は以下のように近似できる。

【数11】

$$ERB(f) = 24.7(4.37f/1000 + 1) \quad (1)$$

【0087】

この関係から、歪んだ周波数スケールが、歪んだスケールのどの点でも歪んだスケールの 1 つにおける対応する ERB が等しくなるように定義される。ヘルツ単位の線形周波数からこの ERB 周波数スケールへの変換のための関数は式 (1) の逆数を積分することにより得られる。

【数12】

$$HzToERB(f) = \int \frac{1}{24.7(4.37f/1000 + 1)} df = 21.4 \log_{10}(4.37f/1000 + 1) \quad (2a)$$

10

20

30

40

50

【 0 0 8 8 】

f について式 ( 2 a ) を解くことにより E R B スケールから線形周波数スケールに戻す変換を表現することもまた有益である。

【 数 1 3 】

$$ERBToHz(e) = f = \frac{1000}{4.37} 10^{(e/21.4-1)}, \quad (2b)$$

【 0 0 8 9 】

ここで、e は、E R B スケールの単位である。図 9 は、E R B スケールとヘルツ単位の周波数との関係を示す。

10

【 0 0 9 0 】

分析フィルタバンク 1 0 0 には、E R B スケールに沿って一定間隔の中央周波数  $f_c$  [ I ] . . .  $f_c$  [ B ] で表される B 聴覚フィルタを含むことができる。より具体的には

【 数 1 4 】

$$f_c[1] = f_{\min} \quad (3a)$$

$$f_c[b] = f_c[b-1] + ERBToHz(HzToERB(f_c[b-1]) + \Delta) \quad b = 2 \dots B \quad (3b)$$

$$f_c[B] < f_{\max}, \quad (3c)$$

20

【 0 0 9 1 】

ここで、 $\Delta$  は、分析フィルタバンク 1 0 0 の好ましい E R B 間隔であり、 $f_{\min}$  と  $f_{\max}$  とはそれぞれ好ましい最大中央周波数と最小中央周波数である。 $\Delta = 1$  を選び、人間の耳が感知できる周波数範囲を考慮に入れて、 $f_{\min} = 50 \text{ Hz}$  そして  $f_{\max} = 20000 \text{ Hz}$  に設定することができる。このようなパラメータにより、例えば、式 ( 3 a ) ~ ( 3 c ) に適用して  $B = 40$  の聴覚フィルタが得られる。

【 0 0 9 2 】

各聴覚フィルタの振幅周波数応答は、Moore と Glasberg が示したように、端数を切り捨てた指数関数により特徴づけることができる。具体的には、中央周波数  $f_c$  [ b ] を持つフィルタの振幅応答は、以下のように計算される、

30

【 数 1 5 】

$$H_b(f) = (1 + pg)e^{-pg} \quad (4a)$$

【 0 0 9 3 】

ここで、

【 数 1 6 】

$$g = \left| \frac{f - f_c[b]}{f_c[b]} \right|, \quad (4b)$$

40

$$p = \frac{4f_c[b]}{ERB(f_c[b])} \quad (4c)$$

【 0 0 9 4 】

このような B 聴覚フィルタの振幅応答は、E R B スケールでのクリティカル帯域を近似し、図 1 0 に示される。

【 0 0 9 5 】

分析フィルタバンク 1 0 0 のフィルタ処理は、一般に短時間離散フーリエ変換 ( S T D F T ) と称される有限長の離散フーリエ変換を用いて適切に近似される。なぜなら、

50



フルレート実施と称される、オーディオ信号のサンプリングレートでこのフィルターを実施することは、正確な音量計測に必要とする以上の時間分解能が得られると信じられているからである。フルレート実施の代わりに S T D F T を用いることにより、効率を改善し計算の複雑さを緩和することができる。

【 0 0 9 6 】

入力オーディオ信号  $x[n]$  の S T D F T は以下のように定義される、  
【 数 1 7 】

$$X[k,t] = \sum_{n=0}^{N-1} w[n]x[n+tT]e^{-j\frac{2\pi kt}{N}}, \quad (5a)$$

10

【 0 0 9 7 】

ここで、 $k$  は周波数インデックスであり、 $t$  は時間ブロックインデックスであり、 $N$  は D F T のサイズであり、 $T$  はホップサイズであり、 $w[n]$  は、以下のように規格化した長さ  $N$  の窓である。

【 数 1 8 】

$$\sum_{n=0}^{N-1} w^2[n] = 1 \quad (5b)$$

【 0 0 9 8 】

20

ここで、式 ( 5 a ) の変数  $t$  は、秒での時間尺度に対立する S T D F T の時間ブロックを表す離散インデックスであることに留意すべきである。 $t$  における増加は、信号  $x[n]$  に沿って  $T$  サンプルのホップを表す。それに続くインデックス  $t$  の参照はこの定義を想定する。実施の形態の詳細に応じて、異なったパラメータの設定と窓の形を用いることができるが、 $f_s = 44100 \text{ Hz}$ 、 $N = 2048$ 、 $T = 1024$ 、そして  $w[n]$  をハニング窓にすることで、適切な時間分解能と周波数分解能とのバランスが得られる。上述の S T D F T は、高速フーリエ変換 ( F F T ) を用いるとより効率的になる。

【 0 0 9 9 】

S T D F T の代わりに、変形離散コサイン変換 ( M D C T ) を分析フィルターバンクを実施するために用いることができる。M D C T は、ドルビー A C - 3 のような、知覚オーディオコードに一般的に用いられる変換である。開示したシステムがこのような知覚的にコード化されたオーディオを実施する場合、開示した音量計測と修正により、コード化されたオーディオの既存の M D C T 係数をより効率的に処理することができ、それにより、分析フィルターバンク変換が削除される。入力信号  $x[n]$  の M D C T は下式で得られる。

30

【 数 1 9 】

$$\text{where } n_0 = \frac{(N/2) + 1}{2} \quad (6)$$

【 0 1 0 0 】

40

一般にホップサイズ  $T$  は、信号  $x[n]$  の完全な再構成が可能となるよう変換長  $N$  のちょうど 2 分の 1 となるよう選ばれる。

【 0 1 0 1 】

[ 伝達フィルター 1 0 1 ]

分析フィルターバンク 1 0 0 の出力は、フィルターバンクの各帯域が外耳及び中耳を通るオーディオの伝達に合致するようにフィルター処理する、伝達フィルター又は伝達フィルター機能 ( 「伝達フィルター」 ) 1 0 1 に用いられる。図 8 は、可聴周波数レンジでの伝達フィルター  $P(f)$  の適切な振幅周波数応答を示す。この応答は 1 k H z 以下で一定であり、1 k H z 以上では、I S O 2 2 6 標準で定義された聴覚閾値の逆数に従い、この閾値は、1 k H z で一致するよう正規化される。

50

## 【 0 1 0 2 】

## [ エキサイテーション 1 0 2 ]

入力オーディオ信号の音量を計算するためには、伝達フィルター 1 0 1 を適用した後の分析フィルターバンク 1 0 0 の各フィルターにおけるオーディオ信号の短時間エネルギーの指標が必要である。この時間的にまた空間的に変化する指標はエキサイテーションと称される。分析フィルターバンク 1 0 0 の各フィルターにおける短時間エネルギー出力は、周波数領域でのフィルター応答に入力信号のパワースペクトルを乗算することにより、エキサイテーション関数 1 0 2 に近似させることができる。

## 【 数 2 0 】

$$E[b,t] = \frac{1}{N} \sum_{k=0}^{N-1} |H_b[k]|^2 |P[k]|^2 |X[k,t]|^2, \quad (7)$$

10

## 【 0 1 0 3 】

ここで、 $b$  は帯域数、 $t$  はブロック数、 $H_b[k]$  と  $P[k]$  は、STDFT ビンインデックス  $k$  又は MDCT ビンインデックス  $k$  に対応する周波数でサンプルされた、それぞれ聴覚フィルター及び聴覚伝達フィルターの周波数応答である。ここで、同様の結果を得るために、式 (4 a) ~ (4 c) で説明したものとは違う聴覚フィルターの振幅応答の形を式 (7) で用いることに留意すべきである。例えば、前記国際特許出願 No. PCT/US 2004/016964 であって、国際公開番号 WO 2004/111964 A2 として国際公開されたものでは、2つの代案が記載されている。すなわち、12次のIIR変換で特徴づけられる聴覚フィルターと、低価格の「レンガ壁 (brick-wall)」帯域通過近似値である。

20

## 【 0 1 0 4 】

要約すると、エキサイテーション関数 1 0 2 の出力は、時間間隔  $t$  毎のそれぞれの E R B 帯域におけるエネルギー  $E$  の周波数領域での表現である。

## 【 0 1 0 5 】

## [ 時間平均化 (「平滑化」) 1 0 3 ]

開示された発明の応用において、以下に説明するように、それを特定ラウドネスに変換する前に、エキサイテーション  $E[b, t]$  を平滑化することが好ましい。例えば、下式により、平滑化関数 1 0 3 を再帰的に行うことで平滑化を実行することができる。

30

## 【 数 2 1 】

$$\bar{E}[b,t] = \lambda_b \bar{E}[b,t] + (1 - \lambda_b) E[b,t], \quad (8)$$

## 【 0 1 0 6 】

ここで、帯域  $b$  での時定数  $\lambda_b$  は、望みの応用例に応じて選定する。多くの場合帯域  $b$  内で人の音量感知の時間積分に比例して選定することが好都合である。WatsonとGengelが、この時間積分が低周波数 (125 ~ 200 Hz) で 150 ~ 175 ms 以内であり、高周波数で 40 ~ 60 ms であることを実験的に示した (Charles S. Watson 及び Roy W. Gengel の「聴覚感度に関連する信号継続と信号周波数」, Journal of the Acoustical Society of America, Vol.46, No.4 (Part2), 1969, pp.989-997)。

40

## 【 0 1 0 7 】

## [ 特定ラウドネス 1 0 4 ]

特定ラウドネス変換又は変換関数 (「特定ラウドネス」) 1 0 4 において、エキサイテーションの各周波数帯域は、ゾーン毎 E R B で計測される、特定ラウドネスの成分値に変換される。

## 【 0 1 0 8 】

最初に、特定ラウドネスの計算において、 $\bar{E}[b, t]$ の各帯域におけるエキサイテーションレベルは、伝達フィルター $P(z)$ (図12)により正規化されたISO226の等音量曲線(図11)により定義されるような、1kHzにおける等価なエキサイテーションレベルに変換することができる。

【数22】

$$\bar{E}_{1kHz}[b, t] = T_{1kHz}(\bar{E}[b, t], f_c[b]), \quad (9)$$

10

【0109】

ここで、 $T_{1kHz}(E, f)$ は、周波数 $f$ ではレベル $E$ の音量に等しい1kHzでのレベルを生成する関数である。実際には $T_{1kHz}(E, f)$ は、伝達フィルターにより正規化された等音量曲線のルックアップテーブルの補間として実行される。1kHzでの等価なレベルに変換することで、以下の特定ラウドネスの計算が簡単になる。

【0110】

次に、各帯域の特定ラウドネスは以下のように計算される。

【数23】

$$N[b, t] = \alpha[b, t]N_{NB}[b, t] + (1 - \alpha[b, t])N_{WB}[b, t], \quad (10)$$

20

【0111】

ここで、 $N_{NB}[b, t]$ と $N_{WB}[b, t]$ は、それぞれ狭帯域信号モデルと広帯域信号モデルに基づく特定ラウドネス値である。値 $[b, t]$ は、オーディオ信号から計算される0と1との間の補間係数である。前記国際特許出願No. PCT/US2004/016964であって、国際公開番号WO2004/111964A2で国際公開されたものには、エキサイテーションのスペクトルの平坦性から $[b, t]$ を計算する技術が記載されている。また、「狭帯域信号」信号モデルと「広帯域」信号モデルについても詳細に記載されている。

【0112】

30

狭帯域特定ラウドネス値と広帯域特定ラウドネス値、 $N_{NB}[b, t]$ と $N_{WB}[b, t]$ は、指数関数を用いて変換されたエキサイテーションから推定することができる。

【数24】

$$N_{NB}[b, t] = \begin{cases} G_{NB} \left( \left( \frac{\bar{E}_{1kHz}[b, t]}{TQ_{1kHz}} \right)^{\beta_{NB}} - 1 \right), & \bar{E}_{1kHz}[b, t] > TQ_{1kHz} \\ 0, & otherwise \end{cases} \quad (11a)$$

$$N_{WB}[b, t] = \begin{cases} G_{WB} \left( \left( \frac{\bar{E}_{1kHz}[b, t]}{TQ_{1kHz}} \right)^{\beta_{WB}} - 1 \right), & \bar{E}_{1kHz}[b, t] > TQ_{1kHz}, \\ 0, & otherwise \end{cases} \quad (11b)$$

40

【0113】

ここで $TQ_{1kHz}$ は、1kHzのトーンに対する静音の閾値でのエキサイテーションレベルである。等音量曲線(図11と12) $TQ_{1kHz}$ からは4.2dBである。これらの特定ラウドネス関数は両方とも、エキサイテーションが静音の閾値に等しいとき、ゼロに等しいことに留意すべきである。静音の閾値より大きなエキサイテーションに対して、両方の関数は、感覚の強さについてのスチーブンスの法則に従うパワーの法則により単調に増大する。狭帯域関数の指数は広帯域関数より大きくなるよう選定し、狭帯域関数を広

50

帯域関数より速く増大させる。狭帯域と広帯域の場合の指数 とゲイン G の具体的な選定は、トーンに対する音量とノイズの増大に対する経験的なデータと適合するように選択される。

【 0 1 1 4 】

MooreとGlasbergが示したように、特定ラウドネスはエキサイテーションが聴覚の閾値にあるとき、ゼロの代わりに小さな値にしておくべきである。エキサイテーションがゼロに向かって減少するにつれて、特定ラウドネスは単調にゼロに向かって減少してゆく。この理由は、聴覚の閾値が蓋然的な閾値であり（トーンが50%の回数で検出される点）、一緒に現れた、各々閾値である複数のトーンが、個々のトーンより可聴であるサウンドの和であるかもしれないからである。開示した本出願において、この特性を有する特定ラウドネスの増大により、以下に説明するように、エキサイテーションが閾値に近いときゲインソルバーに、より適切な振る舞いをさせるという利点が加えられる。エキサイテーションが閾値又は閾値以下にあるとき、特定ラウドネスをゼロに定めると、ゲインソルバーの一意的な解は存在しない。一方、MooreとGlasbergが示したように、特定ラウドネスをゼロ以上のすべてのエキサイテーションの値に対して単調増加するように定めた場合、一意的な解が存在する。音量の圧伸が一定以上であればゲインは一定以上となり、その逆も成り立つ。式(11a)と(11b)の特定ラウドネス関数は、下記の好ましい特性を持つよう修正することができる。

【 数 2 5 】

$$N_{NB}[b, t] = \begin{cases} G_{NB} \left( \left( \frac{\bar{E}_{1kHz}[b, t]}{TQ_{1kHz}} \right)^{\rho_{NB}} - 1 \right), & \bar{E}_{1kHz}[b, t] > \lambda TQ_{1kHz} \\ \exp \left\{ K_{NB} \left( -\log \left( \frac{\bar{E}_{1kHz}[b, t]}{TQ_{1kHz}} \right) + C_{NB} \right)^{\eta_{NB}} \right\}, & otherwise \end{cases} \quad (11c)$$

【 数 2 6 】

$$N_{WB}[m, t] = \begin{cases} G_{WB} \left( \left( \frac{\bar{E}_{1kHz}[b, t]}{TQ_{1kHz}} \right)^{\rho_{WB}} - 1 \right), & \bar{E}_{1kHz}[b, t] > \lambda TQ_{1kHz} \\ \exp \left\{ K_{WB} \left( -\log \left( \frac{\bar{E}_{1kHz}[b, t]}{TQ_{1kHz}} \right) + C_{WB} \right)^{\eta_{WB}} \right\}, & otherwise \end{cases} \quad (11d)$$

【 0 1 1 5 】

ここで、定数 $\lambda$ は1より大きく、指数 $\eta$ は1より小さい。そして、定数KとCとは、特定ラウドネス関数とその第1導関数とが、 $\bar{E}_{1kHz}[b, t] = \lambda TQ_{1kHz}$ で連続となるよう選定される。

【 0 1 1 6 】

特定ラウドネスから、すべての帯域に亘って特定ラウドネスを積算することにより、全体の又は「トータル」音量 $L[t]$ が得られる。

【 数 2 7 】

$$L[t] = \sum_b N[b, t] \quad (12)$$

10

20

30

40

50

## 【 0 1 1 7 】

## [特定ラウドネス修正 1 0 5]

特定ラウドネス修正機能（「特定ラウドネス修正」） 1 0 5において、 $\hat{N}[b,t]$ のよう  
に表される目標特定ラウドネスは、装置又は処理全体の望みのアプリケーションに応じ  
てさまざまな方法で特定ラウドネス S L 1 0 4（図 7）から計算することができる。

## 【 0 1 1 8 】

10

以下に詳細に説明するように、例えば音量制御の場合、スケールファクタ を用いて目標  
特定ラウドネスを計算することができる。以下の式（ 1 6 ）及びそれについての説明を参  
照のこと。自動ゲイン制御及びダイナミックレンジ制御の場合、目標特定ラウドネスは、  
好ましい出力音量と入力音量との比を用いて計算することができる。以下の式（ 1 7 ）と  
（ 1 8 ）及びそれらについての説明を参照のこと。ダイナミックイコライゼーションの場  
合は、式（ 2 3 ）及びそれについての説明に示した関係を用いて目標特定ラウドネスを計  
算する。

## 【 0 1 1 9 】

## [ゲインソルバー 1 0 6]

20

この例では、各帯域  $b$  及び各時間間隔  $t$  に対してゲインソルバー 1 0 6 は入力として  
平滑化されたエキサイテーション  $\bar{E}[b,t]$  と目標特定ラウドネス  $\hat{N}[b,t]$  とを取り、引き  
続いてオーディオの修正に用いるゲイン  $G[b,t]$  を生成する。関数  $\Psi\{\}$  を下式のように  
エキサイテーションから特定ラウドネスに変換する非線形変換とする。

## 【 数 2 8 】

$$N[b,t] = \Psi \{ \bar{E}[b,t] \}, \quad (13)$$

30

## 【 0 1 2 0 】

ゲインソルバーは以下となるように  $G [ b , t ]$  を求める。

## 【 数 2 9 】

$$\hat{N}[b,t] = \Psi \{ G^2[b,t] \bar{E}[b,t] \}. \quad (14a)'$$

## 【 0 1 2 1 】

ゲインソルバー 1 0 6 は、元のエキサイテーションに適用したとき、特定ラウドネスが  
理想的には目標特定ラウドネスに等しくなるような特定ラウドネスが結果的に得られるよ  
うな周波数的に変化し時間的に変化するゲインを定める。実際には、ゲインソルバーは、  
周波数領域でのオーディオ信号に適用したとき、特定ラウドネスと目標特定ラウドネスと  
の差を減少させるようにオーディオ信号を修正する結果となるような周波数的に変化し時  
間的に変化するゲインを定める。理想的には、修正されたオーディオ信号が目標特定ラウ  
ドネスに非常に近似した特定ラウドネスを持つように、修正を行う。式（ 1 4 a ）の解は  
、さまざまな方法で実行できる。例えば、 $\Psi^{-1} \{ \cdot \}$  で表される、特定ラウドネスの逆  
数の閉形式数学的表現が存在するなら、式（ 1 4 a ）を再構成することにより直接的にゲ  
インを計算することができる。

40

【数 3 0】

$$G[b, t] = \sqrt{\frac{\Psi^{-1}(\hat{N}[b, t])}{\bar{E}[b, t]}} \quad (14b)$$

【 0 1 2 2】

代替的に、もし、 $\Psi^{-1}\{\cdot\}$  の閉形式の解が存在しないならば、式 (14 a) の繰り返しを、そのときのゲインの推定値を用いて評価する反復的なアプローチを採用することができる。結果得られた特定ラウドネスを好ましい目標と比較し、ゲインを誤差に基づいて更新する。もしゲインが適切に更新されたなら、好ましい解に収束する。もう 1 つの方法は、ルックアップテーブルを作るために各帯域のエキサイテーション値の範囲で関数  $\Psi^{-1}\{\cdot\}$  をあらかじめ計算する。このルックアップテーブルから、逆関数  $\Psi^{-1}\{\cdot\}$  の近似が得られ、そして式 (14 b) からゲインが計算される。先に説明したように、目標特定ラウドネスは、特定ラウドネスを圧伸することで表現することができるから、

10

【数 3 1】

$$\hat{N}[b, t] = \Xi[b, t] N[b, t] \quad (14c)$$

【 0 1 2 3】

式 (13) を (14 c) に代入し、(14 c) を (14 b) に代入してゲインの他の表現が得られる。

【数 3 2】

20

$$G[b, t] = \sqrt{\frac{\Psi^{-1}(\Xi[b, t] \Psi(\bar{E}[b, t]))}{\bar{E}[b, t]}} \quad (14d)$$

【 0 1 2 4】

ゲインは、エキサイテーション  $\bar{E}[b, t]$  の関数と特定ラウドネスの圧伸  $\Xi[b, t]$  によりきれいに表現できることが分かる。

【 0 1 2 5】

したがって、ゲインは、中間値として特定ラウドネスや目標特定ラウドネスを明示的に計算することなく、(14 d) の評価又は等価なルックアップテーブルから計算することができる。しかし、これらの値は式 (14 d) の使用を通じて黙示的に計算される。特定ラウドネスや目標特定ラウドネスを明示的又は黙示的に計算することで修正パラメータを計算する等価な他の方法も工夫することができ、この発明はこのような方法のすべてを包含するものである。

30

【 0 1 2 6】

[ 合成フィルタバンク 1 1 0 ]

上述したように、分析フィルタバンク 1 0 0 は、短時間離散フーリエ変換 (STDF T) 又は変形離散コサイン変換を用いることにより効率的に実施できるが、STDF T 又は MDC T は同様に合成フィルタバンク 1 1 0 の実施に用いることができる。具体的には、 $X[k, t]$  に入力オーディオの STDF T 又は MDC T を表現させると、先に説明したように、合成フィルタバンク 1 1 0 で処理した (修正された) オーディオ STDF T 又は MDC T は以下のように計算することができる。

40

【数 3 3】

$$\hat{X}[k, t] = \sum_b G[b, t] S_b[k] X[k, t - d], \quad (15)$$

【 0 1 2 7】

ここで  $S_b[k]$  は、帯域  $b$  に関する合成フィルタの応答であり、 $d$  は図 7 の時間遅れブロック 1 0 9 に関する時間遅れである。合成フィルタ  $S_b[k]$  の形は、分析フィル

50

ターバンクで用いたフィルター、 $H_b[k]$ 、と同じに選んでもよく、又はこれらを、ゲイン修正なしで（すなわち  $G[b, t] = 1$  で）完全な再構成ができるように修正してもよい。

【0128】

次いで、 $\hat{X}[k, t]$  の逆フーリエ変換又は変形離散コサイン変換により最終的に処理されたオーディオを生成し、当業者によく知られているような方法で重複加算して合成する。

〔目標特定ラウドネス〕

10

図1～7に例示したような本発明の特徴を実施する形態の振る舞いは、目標特定ラウドネス  $\hat{N}[b, t]$  を計算する方法により指図される。

【0129】

本発明は、目標特定ラウドネスを計算するための特定の関数又は逆関数に制限されるものではないが、いくつかのそのような関数と適切な応用例をここで説明する。

【0130】

〔音量制御に適した時間的に変化せず、周波数的に変化しない関数〕

標準の音量制御ではオーディオに広帯域ゲインを適用することによりオーディオ信号の音量を調整する。一般に、ゲインは、オーディオ音量が望みのレベルになるまでユーザが調整するノブ又はスライダーと結合している。本発明の特徴によれば、このような制御をより音響心理学的に満足させるようにすることができる。本発明のこの特徴によれば、感知するスペクトルの変化をもたらすかもしれない、全周波数帯域に亘って同じ量だけゲインを変化させる音量制御と結合した広帯域ゲインを持つのではなく、理想的には感知するスペクトルに変化がないよう複数の周波数帯域のそれぞれのゲインが人の聴覚モデルを考慮に入れた量だけ変化するように、特定ラウドネススケールファクタが代わって音量制御の調整と結合している。本発明の特徴及びその応用例の記載において、「一定」又は「時間的に変化しない」は、例えば人により、音量制御スケールファクタの設定を時間と共に変化させることを許容するものである。このような、「時間的に変化しない」は、しばしば「ある程度時間的に変化しない」、「ある程度一定である」、「区分的に時間的に変化しない」、「区分的に一定である」、「階段状に時間的に変化しない」、及び「階段状に一定である」というように称される。このようなスケールファクタが与えられると、で乗算され、計測された特定ラウドネスとして目標特定ラウドネスを計算することができる。

20

30

【数34】

$$\hat{N}[b, t] = \alpha N[b, t]. \quad (16)$$

【0131】

トータル音量  $L[t]$  が特定ラウドネス  $N[b, t]$  の全帯域  $b$  に亘っての和なので、上記修正ではファクタによりトータル音量を圧伸するが、音量制御の調整を変更させるときに感知したスペクトルを同じに維持するような方法で行う。言い換えれば、どんなときでも、音量制御の調整を変更することにより、感知した音量が変化することになるが、修正されたオーディオの感知したスペクトルと修正されていないオーディオの感知したスペクトルとにおいて変化はない。図13aは、女性の音声を含むオーディオ信号に対して  $\alpha = 0.25$  のとき、時間「 $t$ 」で帯域「 $b$ 」に亘って得られた複数帯域ゲイン  $G[b, t]$  を示す。比較のために、標準の音量制御として、元のトータル音量を  $0.25$  で圧伸するのに必要な広帯域ゲインを点線で示す（水平線）。この複数帯域ゲイン  $G[b, t]$  は、中間周波数帯域に比べて高周波数帯域と低周波数帯域とで増加している。これ

40

50

は、人の耳が高周波数と低周波数とで感度が低くなることを示す等音量曲線を満足している。

【 0 1 3 2 】

図 1 3 b は、元のオーディオ信号の特定ラウドネスと、先行技術の音量制御により修正されたものとしての広帯域ゲイン修正された信号の特定ラウドネスと、本発明の特徴により修正されたものとしての複数帯域ゲイン修正された信号の特定ラウドネスとを示したものである。複数帯域ゲイン修正された信号の特定ラウドネスは、オリジナルを 0 . 2 5 で圧伸したものである。広帯域ゲイン修正された信号の特定ラウドネスは、元の修正されていない信号の特定ラウドネスに対してそのスペクトルの形が変わっている。この場合、特定ラウドネスは相対的に、低周波数と高周波数の両方で音量が失われている。これは、オーディオの音量を下げたときのオーディオの曇りとして感知され、この問題は、音量が感知音量領域で導き出されたゲインで制御されている、複数帯域で修正された信号では起こらない。

10

【 0 1 3 3 】

従来の音量制御に関する感知されたスペクトルバランスの歪みに加えて、第 2 の問題が存在する。式 ( 1 1 a ) ~ ( 1 1 d ) に反映された音量モデルに反映されている音量検知の特性は、信号レベルが聴覚の閾値に近づくにつれて、あらゆる周波数において信号の音量が速く減少して行くことである。その結果、音量の小さな信号に同じ音量減衰を与えるのに必要な電氣的減衰は、音量の大きな信号に必要なものより少ない。従来の音量制御では、信号レベルを考慮しないで一定の減衰を与えているので、音量を下げたとき、音量の大きな信号に対して音量の小さな信号は、「音量が小さすぎ」になる。多くの場合この結果オーディオの詳細が失われる結果となる。カスタネットを残響室でレコーディングすることを考える。このようなレコーディングにおいて、カスタネットの主となる「一打」は反射エコーに比べて非常に大音量であるが、部屋の大きさを伝達するのは反射エコーである。従来の音量制御で音量を下げるにつれて、反射エコーは、主となる一打に比べて音量が小さくなり、最終的に聴覚の閾値以下に消えてしまい、「乾いた」カスタネット音を残す。音量に基づく音量制御により、音量の大きい部分と音量の小さい部分との相対音量を一定に保つために、音量の大きい部分に比較して音量の小さい残響部分を押し上げることで、レコーディングにおける音量の小さい部分の消失を防止する。このような効果を達成するために、複数帯域ゲイン  $G [ b , t ]$  は、人の音量に対する時間分解能とつりあった速さで時間的に変化しなければならない。

20

30

【 0 1 3 4 】

複数帯域ゲイン  $G [ b , t ]$  が、平滑化されたエキサイテーション  $\bar{E}[b,t]$  の関数として計算されるので、式 ( 8 ) の時定数  $\lambda_b$  の選択は、各帯域  $b$  でゲインが時間的にどれだけ速く変化するかで示される。

【 0 1 3 5 】

先に述べたように、これらの時定数は帯域  $b$  内での人の音量感知の時間の積分に比例して選定することができ、そして時間について適切な  $G [ b , t ]$  の変化を生み出す。もし時定数が不適切 ( 速すぎるか遅すぎる ) に選定されるならば、感知できる不快なアーティファクトが処理したオーディオに持ち込まれる。

40

【 0 1 3 6 】

[ 固定イコライゼーションに適した時間的に不変で周波数的に変化する関数 ]

いくつかの応用例において、オーディオに感知可能な固定のイコライゼーションの適用を望むかもしれない、この場合、目標特定ラウドネスは、下式のように、時間的に不変であるが周波数的に変化するスケールファクタ  $[ b ]$  をその関連として用いて計算することができる。



【数 3 5】

$$\hat{N}[b, t] = \Theta[b] N[b, t],$$

【 0 1 3 7】

ここで、 $\hat{N}[b, t]$ は、目標特定ラウドネスであり、 $N[b, t]$ はオーディオ信号の特定ラウドネスであり、 $b$ は周波数の指標であり、 $t$ は時間の指標である。

【 0 1 3 8】

この場合、圧伸は帯域から帯域で変化してもよい。このような応用例は、例えば理解しやすさを増すために音声周波数で目立つスペクトルを部分的に強調するのに有効である。

【 0 1 3 9】

[自動ゲイン制御及び自動ダイナミックレンジ制御に適した周波数的に不変で時間的に変化する関数]

自動ゲイン制御及び自動ダイナミックレンジ制御（AGC及びDRC）の技法は、オーディオ処理の分野でよく知られている。抽象的な意味では、両方ともある方法でオーディオ信号のレベルを計測し、そして、計測したレベルの関数である量で信号のゲインを修正する。AGCの場合では、計測したレベルが、ユーザが選択した参照レベルに近似するように信号のゲインが修正される。DRCでは、信号の計測したレベルが好ましいレンジに変換されるように信号のゲインが修正される。例えば、オーディオの静かな部分を大きく、音量の大きな部分を静かにすることを望むことができる。RobinsonとGundryにより記載されている（Charles RobinsonとKenneth Gundry、「メタデータによるダイナミックレンジ制御」、107th Convention of the AES, Preprint 5028, September 24-27, 1999, New

York）。従来のAGCとDRCの実施例は、一般に、平滑化されたピーク又は2乗平均平方根（rms）振幅のようなオーディオ信号レベルの簡単な計測を用いて、ゲイン修正をおこなう。このような簡単な計測は、オーディオの感知音量にある程度まで関連しているが、本発明の特徴によれば、音響心理学的モデルに基づく音量の計測によりゲイン修正を行うことにより、感覚的に適切なAGCとDRCが可能となる。さらに、多くの従来のAGCシステムとDRCシステムでは、広帯域ゲインのゲイン修正を適用し、したがって、処理したオーディオ中に先に説明したような音質の（スペクトルの）歪みを受ける。一方、本発明の特徴によれば、そのような歪みを除去又は最低限にするような方法で特定ラウドネスを形成するために複数帯域のゲインを用いる。

【 0 1 4 0】

本発明の特徴を採用するAGCの応用例とDRCの応用例とは両方とも、入力広帯域音量 $L_i[t]$ を好ましい出力広帯域音量 $L_o[t]$ に変換又は写像し、ここで音量は、ゾーンのような感知音量単位で計測される。入力広帯域音量 $L_i[t]$ は、入力オーディオ信号の具体的な入力特定ラウドネス $N[b, t]$ の関数となる。それは入力オーディオ信号のトータル音量と同じかもしれないが、オーディオ信号のトータル音量を時間的に平滑化したものとしてもよい。

【 0 1 4 1】

図14aと14bは、それぞれAGCとDRCに一般的な写像の例を示す。 $L_o[t]$ が $L_i[t]$ の関数であるような写像を仮定すると目標特定ラウドネスは以下のように計算される。

【数 3 6】

$$\hat{N}[b, t] = \frac{L_o[t]}{L_i[t]} N[b, t]. \quad (17)$$

【 0 1 4 2】

10

20

30

40

50

このオーディオ信号の元の特定ラウドネス  $N[b, t]$  を好ましい出力広帯域音量の入力広帯域音量に対する比で単純に圧伸することで、出力目標特定ラウドネス  $\hat{N}[b, t]$  を生成する。

#### 【 0 1 4 3 】

A G C システムに対して、入力広帯域音量  $L_i [t]$  は、原則として、オーディオの長時間トータル音量の計測値であるべきである。これは、トータル音量  $L [t]$  を時間で平滑化して  $L_i [t]$  を作るにより達成できる。

10

#### 【 0 1 4 4 】

A G C と比較して、D R C システムは、信号の音量短時間変化に反応し、したがって、 $L_i [t]$  は単純に  $L [t]$  と等しくなる。結果として、 $L_o [t] / L_i [t]$  で与えられる特定ラウドネスの圧伸は、急速に動揺し、処理したオーディオに不必要なアーティファクトを導く。1つの典型的なアーティファクトは、周波数スペクトルの一部が、他の比較的關係のないスペクトル部分で可聴に変調されたものである。例えば、クラシック音楽を選択した場合、低周波数には大きく響くティンパニーが含まれる一方、一様な弦楽器の響きが目立つ高周波数が含まれる。このティンパニーを叩いたときはいつでも、全体的な音量  $L_i [t]$  が増大し、D R C システムはこの特定ラウドネス全体に減衰を適用する。そして、弦楽器は「ポンプ」ダウンし、ティンパニーは「ポンプ」アップして聞こえる。このようなスペクトル中のクロスポンピングは、今までの広帯域 D R C システムでも問題であり、一般的な解決手段は、D R C を異なった周波数帯域に独立に適用することである。ここに開示したシステムは、フィルタバンクと感知音量モデルを用いる特定ラウドネスの計算に起因して、本質的に複数帯域であり、したがって、本発明の特徴に従うように複数帯域で D R C システムを動作させるように修正することは、比較的簡単であり、次に説明する。

20

#### 【 0 1 4 5 】

[ ダイナミックレンジ制御に適した周波数的に変化し、時間的に変化する関数 ]

D R C システムは、入力音量と出力音量を帯域  $b$  とは独立に変化させることにより、複数帯域又は周波数的に変化する方法で動作させるよう拡張することができる。これらの複数帯域音量値は  $L_i [b, t]$  と  $L_o [b, t]$  として参照され、目標特定ラウドネスは、下式で得られる。

30

#### 【 数 3 7 】

$$\hat{N}[b, t] = \frac{L_o[b, t]}{L_i[b, t]} N[b, t], \quad (18)$$

#### 【 0 1 4 6 】

ここで、 $L_o [b, t]$  は、図 1 4 b で示したように  $L_i [b, t]$  から各帯域  $b$  とは独立に計算又は写像されたものである。入力複数帯域音量  $L_i [b, t]$  は、入力オーディオ信号の特定ラウドネス  $N [b, t]$  の関数である。これは、入力オーディオ信号の特定ラウドネスと同じかもしれないが、オーディオ信号の特定ラウドネスを時間的に平滑化し及び/又は周波数的に平滑化したものであってもよい。

40

#### 【 0 1 4 7 】

$L_i [b, t]$  を計算する最も簡単な方法は、それを特定ラウドネス  $N [b, t]$  に等しく設定することである。この場合、D R C は、上記表題「自動ゲイン制御及び自動ダイナミックレンジ制御に適した周波数的に不変で時間的に変化する関数」で説明したような、全帯域で同じ入力音量対出力音量比によるのではなく、感知音量モデルの聴覚フィルタバンク内の各帯域で独立に行われる。40 帯域を用いた実際の実施の形態において、周

50

波数軸に沿うこれらの帯域間の間隔は、音量の正確な計測とを提供するために比較的細かい。しかし、各帯域に独立に DRC スケールファクタを適用することにより、処理したオーディオを「引き裂かれた」サウンドにするかもしれない。この問題を避けるため、DRC の量を 1 つの帯域から次の帯域に適用するとき大幅に変化することのないように特定ラウドネス  $N[b, t]$  を帯域を通して平滑化して  $L_i[b, t]$  を計算するような選択ができる。これは、帯域平滑化フィルタ  $Q(b)$  を定義し、標準的な畳み込み積算により特定ラウドネスを全帯域  $c$  で平滑化することで達成できる。

【数 38】

$$L_i[b, t] = \sum_c Q(b-c) N[c, t]. \quad (19)$$

10

【0148】

ここで、 $N[c, t]$  は、オーディオ信号の特定ラウドネスであり、 $Q(b-c)$  は、平滑化フィルタの帯域シフトした応答である。図 15 は、このような帯域平滑化フィルタの 1 例を示す。

【0149】

もし、 $L_o[b, t]$  の関数として  $L_i[b, t]$  を計算する DRC 関数がすべての帯域  $b$  で固定されているならば、特定ラウドネス  $N[b, t]$  の各帯域で被る変化の形式は、信号の全体的な音量が同じに維持されるとしても、処理されるオーディオのスペクトルにより変化する。例えば、大きな音量のバスと静かなトレブルを有するオーディオ信号は、バスをカットしトレブルをブーストすることができる。正味の効果はオーディオの音質又は感知されるスペクトルの変化であり、これは特定の応用例では好ましいかもしれない。

20

【0150】

しかし、オーディオの平均感知スペクトルを修正しないで複数帯域の DRC を行うことを望むかもしれない。修正の短時間変動を帯域間で独立に動作させる一方、各帯域における平均修正をほぼ同じにすることを望むかもしれない。各帯域におけるこの DRC の平均的な作用を基準となる作用と同じにさせることにより好ましい効果が達成できるかもしれない。広帯域入力音量  $L_i[b, t]$  の好ましい DRC としてこの基準となる作用を選ぶことができる。関数  $L_o[t] = DRC\{L_i[t]\}$  が広帯域音量の好ましい DRC 写像を表すとする。

30

【0151】

そして、 $\bar{L}_i[t]$  を広帯域入力音量の時間平均したものを表すとし、 $\bar{L}_i[b, t]$  を複数帯域入力音量  $L_i[b, t]$  の時間平均したものを表すとする。そうすると、複数帯域出力音量は下式で計算される。

【数 39】

$$L_o[b, t] = \frac{\bar{L}_i[b, t]}{\bar{L}_i[t]} DRC \left\{ \frac{\bar{L}_i[t]}{\bar{L}_i[b, t]} L_i[b, t] \right\}. \quad (20)$$

40

【0152】

ここで、複数帯域入力音量は広帯域入力音量と同じ平均レンジに最初に圧伸されることに留意すべきである。広帯域音量に設計された DRC 関数が次いで適用される。最後に、その結果が圧伸されて複数帯域音量の平均レンジに戻される。複数帯域 DRC をこのように公式化することにより、縮小したスペクトルのポンピングの利点が保たれ、同時にオーディオの平均感知スペクトルが保たれる。

【0153】

50

〔ダイナミックイコライゼーションに適した周波数的に変化し時間的に変化する関数〕

本発明の特徴の他の応用例は、オーディオの元のダイナミックレンジを保持しながら、時間的に変化するオーディオの感知スペクトルを目標となる時間的に変化する感知スペクトルに意図的に変換することである。この処理をダイナミックイコライゼーション（DEQ）と称してもよい。従来の静的なイコライゼーションでは、単純な固定のフィルター処理がそのスペクトルを変化させるためにオーディオに適用されている。例えば、固定したバス又はトレブルブーストに適用できる。このような処理は、オーディオの現在のスペクトルを考慮に入れないので、同じ信号すなわち、すでに比較的大量のバス又はトレブルを含む信号には適さない。DEQにより、信号のスペクトルが計測され、この計測されたスペクトルを本質的に静的な好ましい形に変換するためにその信号をダイナミックに修正する。本発明の特徴として、このような好ましい形はフィルタバンクの帯域を通じて仕様を定められ、EQ [b] と称される。実際の実施の形態では、計測されたスペクトルは、特定ラウドネス  $N[b, t]$  を時間で平滑化することにより生成することができるオーディオの平均スペクトルの形を表す。

【0154】

この平滑化された特定ラウドネスを  $\bar{N}[b, t]$  と称してもよい。複数帯域DRCを用いたとき、1つの帯域から次の帯域へ移るときにDEQが急激に変化することを望まないかもしれないので、帯域平滑化関数を帯域平滑化されたスペクトル  $\bar{L}_i[b, t]$  を生成するために用いてもよい。

【数40】

$$\bar{L}[b, t] = \sum_c Q(b-c) \bar{N}[c, t]. \quad (21)$$

【0155】

オーディオの元のダイナミックレンジを保存するために、好ましいスペクトルEQ [b] は、 $\bar{L}[b, t]$  で与えられる計測されたスペクトル形状の全体的な音量と同じとなるよう正規化される。この正規化されたスペクトル形状を  $\bar{L}_{EQ}[b, t]$  と称することができる。

【数41】

$$\bar{L}_{EQ}[b, t] = \left( \frac{\sum_c \bar{L}[c, t]}{\sum_c EQ[c]} \right) EQ[b]. \quad (22)$$

【0156】

最終的に目標特定ラウドネスは以下のように計算される。

【数42】

$$\hat{N}[b, t] = \left( \frac{\bar{L}[b, t]}{\bar{L}_{EQ}[b, t]} \right)^{1-\beta} \frac{\bar{L}_{EQ}[b, t]}{\bar{L}[b, t]} N[b, t], \quad (23)$$

【0157】

ここで  $\beta$  は、ユーザが指定する0から1までの間のパラメータであり、適用されるDEQ

10

20

30

40

50

の程度を示す。式(23)を参照して、 $\alpha = 0$ のとき、もとの特定ラウドネスは修正されず、 $\alpha = 1$ のとき、特定ラウドネスは好ましいスペクトル形状の計測されたスペクトル形状に対する比で圧伸される。

【0158】

好ましいスペクトル形状EQ[b]を生成する便利な方法は、ユーザが満足するスペクトルバランスのオーディオについて計測したものと、ユーザがそれを $\bar{L}[b,t]$ に等しくなるよう設定することである。実際の実施の形態において、例えば図16に示すように、ユーザはボタン又は他の適切な駆動部507により、それが駆動されたとき、現時点のオーディオのスペクトル形状 $\bar{L}[b,t]$ の計測値を捕捉(キャプチャー)させ、そしてこの計測値を、DEBが可能になったとき後で(プリセット選択508により)EQ[b]に取り込むことのできるプリセットとして(目標特定ラウドネス、プリセット、キャプチャー、及び保存506に)保存する。

10

【0159】

図16は、図7を簡単にしたもので、分析フィルターバンク100から合成フィルターバンク110までの複数帯域を示す1つのラインのみを示している。図17の例では、上述したようなダイナミックイコライゼーションに従い機能又は装置104で計測された特定ラウドネスに修正を施す、ダイナミックEQ特定ラウドネス(SL)修正505も提供する。

20

【0160】

[結合された処理]

音量制御(VC)、AGC、DRC、及びDEQを含む、先に説明したすべての処理を1つのシステムに結合することを望むかもしれない。これらの各処理は特定ラウドネスの圧伸として表すことができるから、以下のようにこれらを結合することができる。

【数43】

$$\hat{N}[b,t] = (\Xi_{VC}[b,t]\Xi_{AGC}[b,t]\Xi_{DRC}[b,t]\Xi_{DEQ}[b,t])N[b,t], \quad (24)$$

30

【0161】

ここで、 $\Xi_{*}[b,t]$ は、処理「\*」と関連するスケールファクタを表す。

【0162】

次いで、結合された処理を表す目標特定ラウドネスのゲインG[b,t]の単一のセットを計算することができる。

【0163】

場合によっては、このラウドネスの修正処理の1つ又は組み合わせについてのスケールファクタが時間的に速く変動しすぎ、処理したオーディオ中にアーティファクトを生じることがある。したがって、これらのスケールファクタの一部を平滑化させることが望ましい。一般に、VCとDEQからのスケールファクタは時間的に平滑に変化するが、AGCスケールファクタとDRCスケールファクタの結合を平滑化することは必要であろう。これらのスケールファクタの結合を以下で表すこととする。

40

【数44】

$$\Xi_C[b,t] = \Xi_{AGC}[b,t]\Xi_{DRC}[b,t] \quad (25)$$

【0164】

50

平滑化の背後にある概念は、特定ラウドネスが増加するとき結合されたスケールファクタがすばやく反応することであり、特定ラウドネスが減少するときスケールファクタがより重く平滑化されることである。この概念は、オーディオ圧縮器の設計におけるファーストアタック (fast attack) 及びスローリリース (slow release) の使用としてよく知られているものに相当する。スケールファクタを平滑化するための適切な時定数は、特定ラウドネスの帯域平滑化したものを時間的に平滑化することにより計算することができる。最初に、特定ラウドネスの帯域平滑化したものは下式で計算される。

【数 4 5】

$$L[b, t] = \sum_c Q(b - c) N[c, t]. \quad (26)$$

10

【0 1 6 5】

ここで  $N[c, t]$  は、オーディオ信号の特定ラウドネスであり、 $Q[b - c]$  は、前記式 (19) のような平滑フィルターの帯域シフトした応答である。

【0 1 6 6】

この帯域平滑化した特定ラウドネスの時間平滑化したものは下式で計算される。

【数 4 6】

$$\bar{L}[b, t] = \lambda[b, t] L[b, t] + (1 - \lambda[b, t]) \bar{L}[b, t - 1] \quad (27)$$

20

【0 1 6 7】

ここで、帯域に依存する平滑化係数  $\lambda[b, t]$  は下式で与えられる。

【数 4 7】

$$\lambda[b, t] = \begin{cases} \lambda_{fast}, & L[b, t] > \bar{L}[b, t] \\ \lambda_{slow}, & L[b, t] \leq \bar{L}[b, t] \end{cases}. \quad (28)$$

【0 1 6 8】

平滑化した結合されたスケールファクタは、次に、下式で計算される。

【数 4 8】

$$\bar{\Xi}_c[b, t] = \lambda_M[b, t] \Xi_c[b, t] + (1 - \lambda_M[b, t]) \bar{\Xi}_c[b, t - 1], \quad (29)$$

30

【0 1 6 9】

ここで、 $\lambda_M[b, t]$  は、 $\lambda[b, t]$  の帯域平滑化したものである。

【数 4 9】

$$\lambda_M[b, t] = \left( \frac{1}{\sum_c Q(c)} \right) \sum_c Q(b - c) \lambda[b, t]. \quad (30)$$

40

【0 1 7 0】

平滑化係数の帯域平滑化により、時間平滑化したスケールファクタが帯域をまたがって急激に変化することを防ぐ。時間平滑化と帯域平滑化の記載のスケールファクタにより、処理したオーディオに、耳障りな感知できるアーティファクトをほとんど含有させない。

【0 1 7 1】

[ ノイズ補償 ]

多くのオーディオ再生環境において、リスナが聞こうとするオーディオと干渉する暗騒音が存在する。例えば、移動中の乗用車内のリスナは備え付けのステレオシステムで音楽を再生することができ、エンジン及び道路からのノイズにより、音楽が顕著に異なって感

50

知される。特に、ノイズのエネルギーが音楽のエネルギーに強く関連するスペクトルの部分では、音楽の感知音量が減少する。もしノイズが十分大きい場合は、音楽は完全にマスクされる。

【 0 1 7 2 】

現発明の特徴に関して、干渉性のあるノイズの存在下での処理したオーディオの特定ラウドネスが目標特定ラウドネス  $\bar{N}[b, t]$  に等しくなるように、ゲイン  $G[b, t]$  を選ぼうとするだろう。

10

【 0 1 7 3 】

この効果を達成するため、MooreとGlasbergによって定義された部分ラウドネスの概念を用いてもよい。ノイズ単独の計測値とオーディオ単独の計測値とが得られると仮定する。 $E_N[b, t]$  はノイズからのエキサイテーションを表し、 $E_A[b, t]$  はオーディオからのエキサイテーションを表すものとする。オーディオとノイズの結合した特定ラウドネスは、下式で与えられる。

【数 5 0】

$$N_{TOT}[b, t] = \Psi\{E_A[b, t] + E_N[b, t]\}, \quad (31)$$

【 0 1 7 4 】

ここで、 $\{\cdot\}$  は、エキサイテーションから特定ラウドネスへの非線形変換を表す。結合した特定ラウドネスを保持するような方法で、リスナーの聴覚が結合した特定ラウドネスを、オーディオの部分特定ラウドネスとノイズの部分特定ラウドネスとを区別すると仮定する。すなわち、

20

【数 5 1】

$$N_{TOT}[b, t] = N_A[b, t] + N_N[b, t]. \quad (32)$$

【 0 1 7 5 】

オーディオの部分特定ラウドネス、 $N_A[b, t]$  は、制御したい値であり、したがってこの値を求めなければならない。ノイズの部分特定ラウドネスは、以下のように近似することができる。

30

【数 5 2】

$$N_N[b, t] = \left( \frac{E_{TN}[b, t]}{E_A[b, t]} \right)^k \left( \Psi\{E_N[b, t] + E_{TN}[b, t]\} - \Psi\{E_{TQ}[b]\} \right) \quad (33)$$

【 0 1 7 6 】

ここで  $E_{TN}[b, t]$  は、ノイズの存在によりマスクされる閾値であり、 $E_{TQ}[b]$  は帯域  $b$  での静音閾値であり、 $k$  は 0 と 1 との間の指数である。式 (31) ~ (33) を結合して、オーディオの部分特定ラウドネスの表現に到達する。

40

【数 5 3】

$$N_A[b, t] = \Psi\{E_A[b, t] + E_N[b, t]\} - \left( \frac{E_{TN}[b, t]}{E_A[b, t]} \right)^k \left( \Psi\{E_N[b, t] + E_{TN}[b, t]\} - \Psi\{E_{TQ}[b]\} \right) \quad (34)$$

【 0 1 7 7 】

オーディオのエキサイテーションがノイズによりマスクされる閾値 ( $E_A[b, t] = E_{TN}[b, t]$ ) であるとき、オーディオの部分特定ラウドネスは、静音閾値での信号のラウドネスに等しく、これは好ましい結果であることに留意すべきである。オーディオのエキサイテーションがノイズよりはるかに大きいとき、式 (34) の 2 番目の項が消え

50

、オーディオの特定ラウドネスはノイズが無い場合の値にほぼ等しくなる。言い換えると、オーディオがノイズよりはるかに大きいので、ノイズがオーディオにマスクされる。指数は、信号対ノイズ比の関数として、ノイズ中のトーンの音量のデータにうまく適合するように経験的に選択される。MooreとGlasbergは  $\alpha = 0.3$  が適当であることを見つけている。マスクされるノイズの閾値は、ノイズのエキサイテーション自身の関数として近似される。

【数54】

$$E_{TN}[b,t] = K[b]E_N[b,t] + E_{TQ}[b] \quad (35)$$

【0178】

ここで、 $K[b]$  は低周波数帯域で増大する定数である。したがって、式(34)で与えられるオーディオの部分特定ラウドネスは、抽象的にはオーディオのエキサイテーションとノイズのエキサイテーションの関数として表現することができる。すなわち、

【数55】

$$N_A[b,t] = \Phi\{E_A[b,t], E_N[b,t]\}. \quad (36)$$

【0179】

修正されたゲインソルバーは、次いで、ノイズの存在下で処理したオーディオの部分特定ラウドネスが目標特定ラウドネスに等しくなるようゲイン  $G[b, t]$  の計算に用いられる。

【数56】

$$\hat{N}[b,t] = \Phi\{G^2[b,t]E_A[b,t], E_N[b,t]\} \quad (37)$$

【0180】

図17は、元のゲインソルバー106を記載したノイズ補償ゲインソルバー206に置き換えた図7のシステムを示す（フルターバンクの複数帯域を示すブロック間の複数の垂直線は、図を単純化するために1本の線に置き換えてあることに留意すべきである）。加えて、図は、（平滑化103からの）オーディオのエキサイテーションと（SL修正からの）目標特定ラウドネスと共に新しいゲインソルバー206に（ブロック100, 101, 102および103の動作に相当する方法で、分析フルターバンク200, 伝達フィルタ201, エキサイテーション202及び平滑化203により）供給するノイズエキサイテーションの測定も示している。

【0181】

最も基本的な動作では、図17のSL修正105は、単に目標特定ラウドネス  $\hat{N}[b,t]$

をオーディオの元の特定ラウドネス  $N[b,t]$  に等しく設定することとしてもよい。

【0182】

言い換えれば、SL修正は、周波数的に変化しない、 $\alpha = 1$  である、オーディオ信号の特定ラウドネスを圧伸するスケールファクタを提供する。図17のような構成により、ノイズの存在下で処理したオーディオの感知されるラウドネススペクトラムがノイズが存在しない状態でのラウドネススペクトラムと等しくなるよう計算される。さらに、VC, AGC, DRC, 及びDEQを含む元の関数として目標特定ラウドネスを計算するための先に説明した技術の1つ又はその組み合わせを、ノイズ補償音量修正システムと共に用いてもよい。

【0183】

現実的な実施の形態において、ノイズの計測はオーディオが再生される環境の近傍にお

10

20

30

40

50



かれたマイクロフォンから得ることもできる。あるいは、さまざまな状況の下での予想ノイズスペクトルに近似する所定のノイズエキサイテーションのテンプレートのセットを用いてもよい。例えば、自動車の室内のノイズは種々の運転速度であらかじめ分析しておき、それをノイズエキサイテーション対速度のルックアップテーブルとして保存することができる。次いで、自動車の速度が変わるときにこのルックアップテーブルから、図17のゲインソルバー206に送られるノイズエキサイテーションを近似する。

【0184】

〔実施の形態〕

本発明は、ハードウェア又はソフトウェア又は両者の組み合わせ（例えば、プログラムブルロジックアレー）により実施することができる。特記なき場合は、本発明の一部として含まれるアルゴリズムは、特定のコンピュータ又は他の装置に本質的に関連するものではない。特に、ここでの開示にしたがって書かれたプログラムによりさまざまな汎用マシンを使用してもよく、また、要求される方法を実行するために特別に作られた装置（例えば、集積回路）を作ることはもっと便利かもしれない。このように、各々が少なくとも1つのプロセッサ、少なくとも1つのデータ記憶システム（揮発性メモリー、不揮発性メモリー、及び/又は記憶素子を含む）、少なくとも1つの入力装置又は入力ポート、及び少なくとも1つの出力装置又は出力ポートからなる1以上のプログラム可能なコンピュータシステム上で動作する1以上のコンピュータプログラムにより本発明を実施することができる。プログラムコードは、本明細書に記載した機能を実行し出力情報を生成するために入力データにおいて適用される。出力情報は、既知の方法で、1以上の出力装置において適用される。

【0185】

このようなプログラムの各々は、コンピュータシステムと通信するために好ましいどんなコンピュータ言語（機械語、アセンブリ言語、又は高級手続言語、ロジック言語、又はオブジェクト指向プログラム言語を含む）で実行してもよい。どの場合でも、言語はコンパイラでもインタプリタでもよい。

【0186】

ここに記載した手続を実行させるために記憶媒体又は装置をコンピュータが読み出したとき、コンピュータを設定し動作させるために、このようなコンピュータプログラムは、汎用コンピュータ又は特定目的のプログラム可能なコンピュータで読み出し可能な記憶媒体又は装置（例えば、半導体メモリー又は媒体、又は次期媒体又は光学媒体）に記憶しておくことが好ましい。本発明のシステムは、コンピュータプログラムとして構成され、この記憶媒体がここに記載した機能を実行するために具体的かつあらかじめ定めた方法でコンピュータを動作させるよう構成されている、コンピュータで読み出し可能な記憶媒体として実施することも考えられる。

【0187】

本発明の種々の実施の形態をここで説明した。しかしながら、本発明の思想や範囲から逸脱することなしにさまざまな変更を加えることが可能であることは了解されよう。例えば、上述の工程のいくつかはその順序に関係がなく、したがって、ここに説明したものと異なる順序で実施することができる。したがって、他の実施の形態も以下の特許請求の範囲に含まれる。

【図面の簡単な説明】

【0188】

【図1】本発明の特徴によるフィードフォワード構成の実施形態の1例を示す機能ブロックダイアグラムである。

【図2】本発明の特徴によるフィードバック構成の実施形態の1例を示す機能ブロックダイアグラムである。

【図3】本発明の特徴によるフィードフォワード/フィードバックのハイブリッド構成の実施形態の1例を示す機能ブロックダイアグラムである。

【図4】本発明の特徴によるもう1つのフィードフォワード/フィードバックのハイブリ

ッド構成の実施形態の 1 例を示す機能ブロックダイアグラムである。

【図 5】フィードフォワード構成、フィードバック構成、及びフィードフォワード／フィードバックのハイブリッド構成のいずれか 1 つにより決定された修正されていないオーディオ信号と修正パラメータを、例えば、時間的に又は空間的に分離した装置又は処理において用いるために保存又は伝達する様子を示す機能ブロックダイアグラムである。

【図 6】フィードフォワード構成、フィードバック構成、及びフィードフォワード／フィードバックのハイブリッド構成のいずれか 1 つにより決定された修正されていないオーディオ信号と目標特定ラウドネス又はそれを代理するものを、例えば、時間的に又は空間的に分離した装置又は処理において用いるために保存又は伝達する様子を示す機能ブロックダイアグラムである。

10

【図 7】本発明の特徴を概観する概略機能ブロックダイアグラム又は概略フローチャートである。

【図 8】本発明の実施の形態における伝達フィルタに適した線形フィルタ  $P(z)$  の理想的な特性応答を示し、ここで、垂直軸はデシベル (dB) 単位の減衰であり、水平軸は 10 を底とする対数で表したヘルツ (Hz) である。

【図 9】ERB 周波数スケール (垂直軸) とヘルツ (Hz) 単位の周波数 (水平軸) との関係を示す。

【図 10】ERB スケールにおけるクリティカル帯域を近似した応答特性を持つ理想的な聴覚フィルタを示す。水平軸はヘルツ単位で示した周波数で、垂直軸はデシベル単位で示したレベルである。

20

【図 11】ISO 226 の等音量曲線を示す。水平軸はヘルツ (Hz) 単位で示した周波数であり (底 10 の対数)、垂直軸はデシベル単位で示した音圧レベルである。

【図 12】伝達フィルタ  $P(z)$  により正規化された ISO 226 の等音量曲線を示す。水平軸はヘルツ単位で示した周波数であり (底 10 の対数)、垂直軸はデシベル単位で示した音圧レベルである。

【図 13a】女性の音声セグメント上の 0.25 で音量縮小させる広帯域及び複数帯域ゲインを示す理想的な図である。水平軸の尺度は ERB 単位であり、垂直軸の尺度はデシベル (dB) 単位の相対ゲインである。

【図 13b】それぞれ、元の信号、広帯域でゲイン修正された信号、及び複数帯域でゲイン修正された信号の特定ラウドネスを示す理想的な図である。水平軸の尺度は ERB 単位であり、垂直軸の尺度はデシベル (dB) 単位の相対ゲインである。

30

【図 14a】典型的な AGC に対する  $L_i[t]$  の関数としての  $L_o[t]$  を示す理想的な図である。水平軸は  $\log(L_i[t])$  であり垂直軸は  $\log(L_o[t])$  である。

【図 14b】典型的な DRC に対する  $L_i[t]$  の関数としての  $L_o[t]$  を示す理想的な図である。水平軸は  $\log(L_i[t])$  であり垂直軸は  $\log(L_o[t])$  である。

【図 15】複数帯域 DRC に対する典型的な帯域平滑関数を示す理想的な図である。水平軸の尺度は帯域数であり、垂直軸の尺度は帯域  $b$  に対するゲイン出力である。

【図 16】本発明の特徴を概観する概略機能ブロックダイアグラム又は概略フローチャートである。

40

【図 17】再生環境における騒音の相殺を含む、図 1 に類似の概略機能ブロックダイアグラム又は概略フローチャートである。

【図 1】

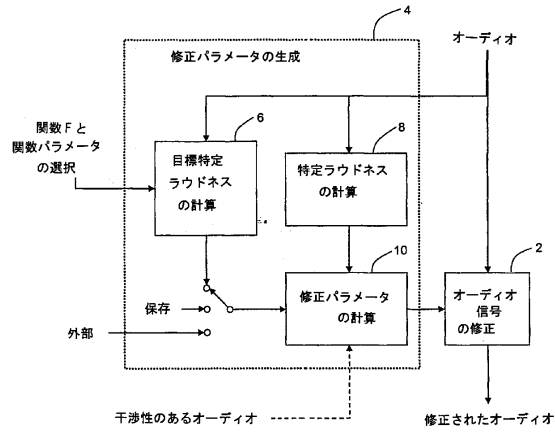


図 1

【図 2】

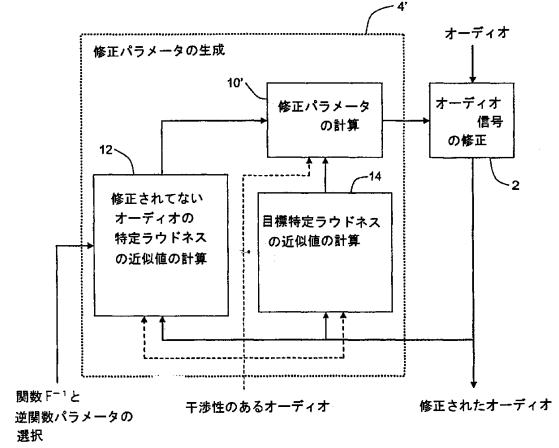


図 2

【図 3】

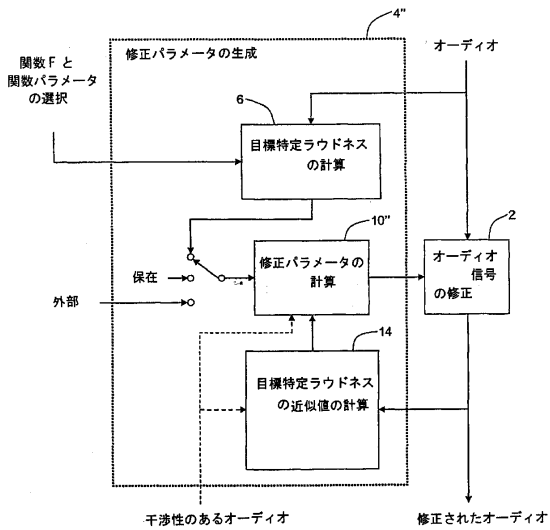


図 3

【図 4】

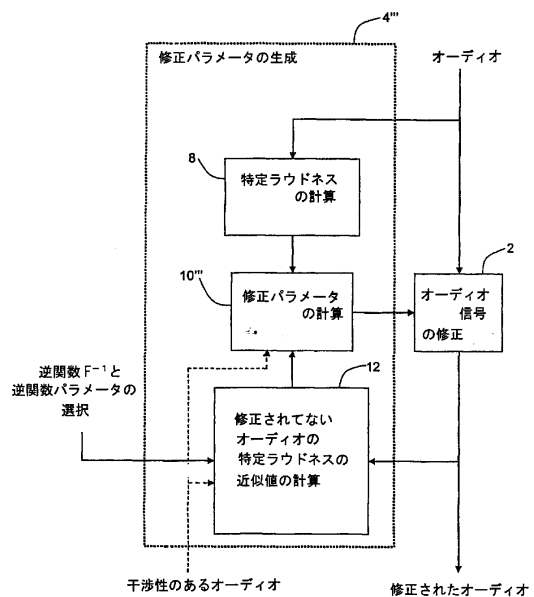


図 4

【図 5】

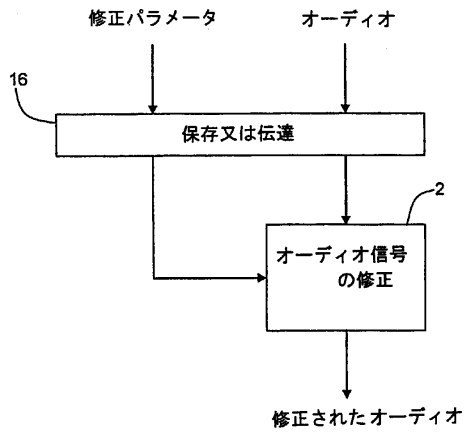


図 5

【図 6】

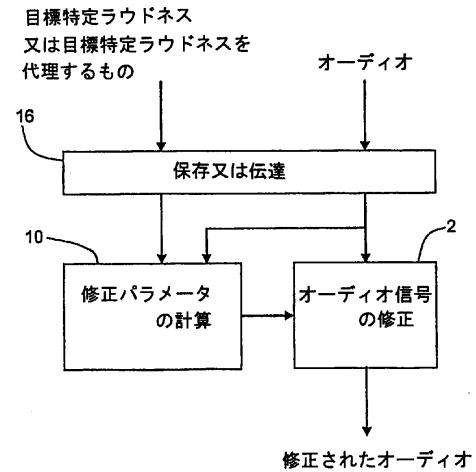


図 6

【図 7】

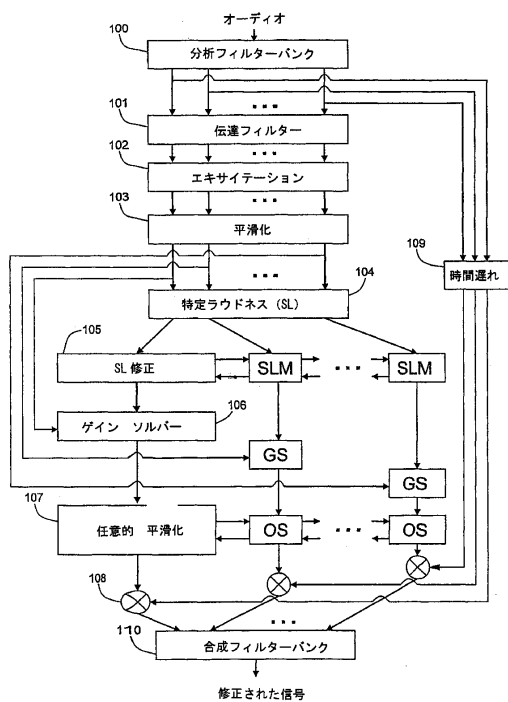


図 7

【図 8】

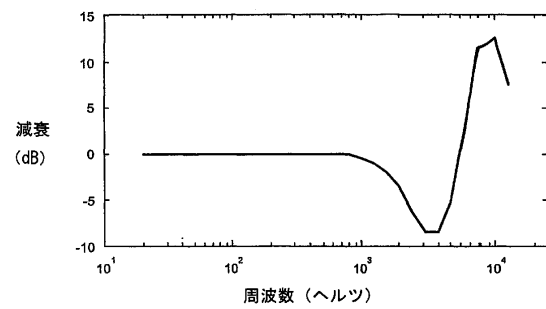


図 8

【図 9】

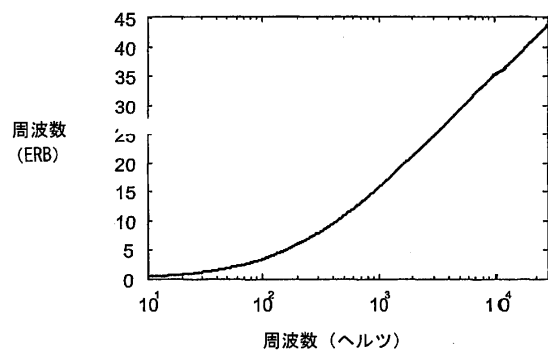


図 9

【図 10】

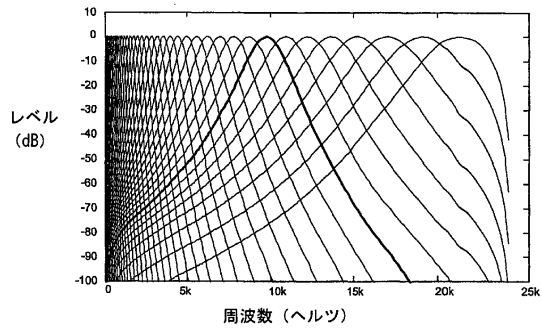


図 10

【図 12】

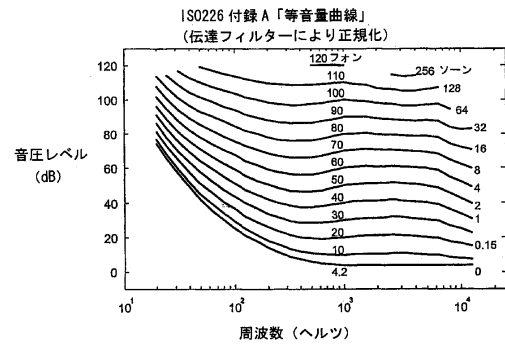


図 12

【図 11】

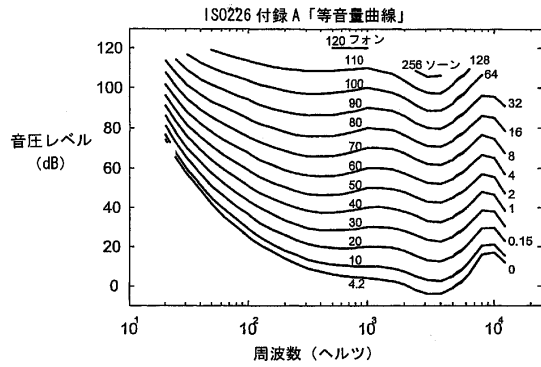


図 11

【図 13 a】

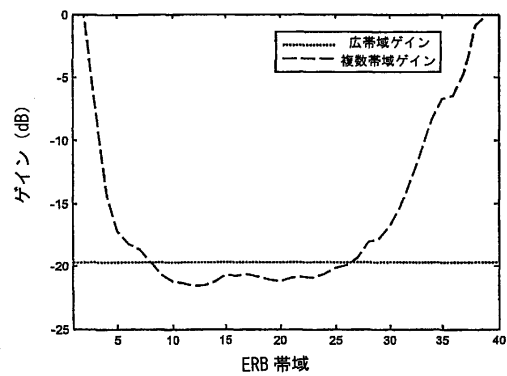


図 13a

【図 13 b】

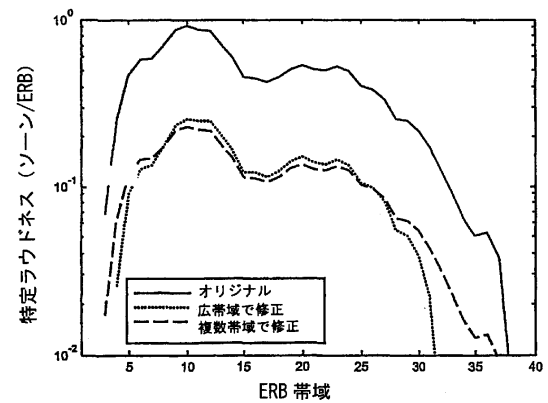


図 13b

【図 14 a】

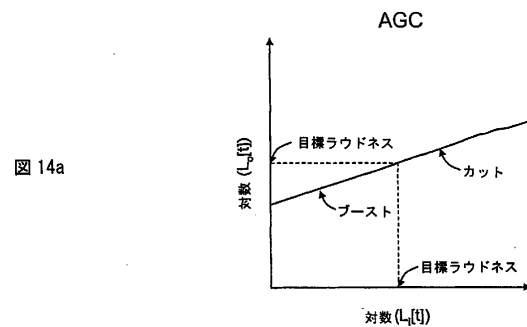
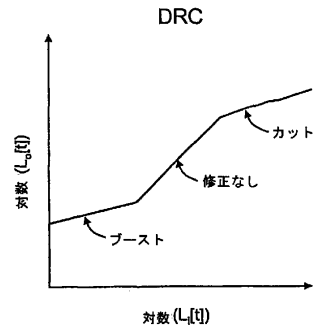


図 14a

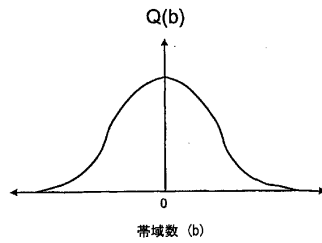
【図 14b】

図 14b



【図 15】

図 15



【図 16】

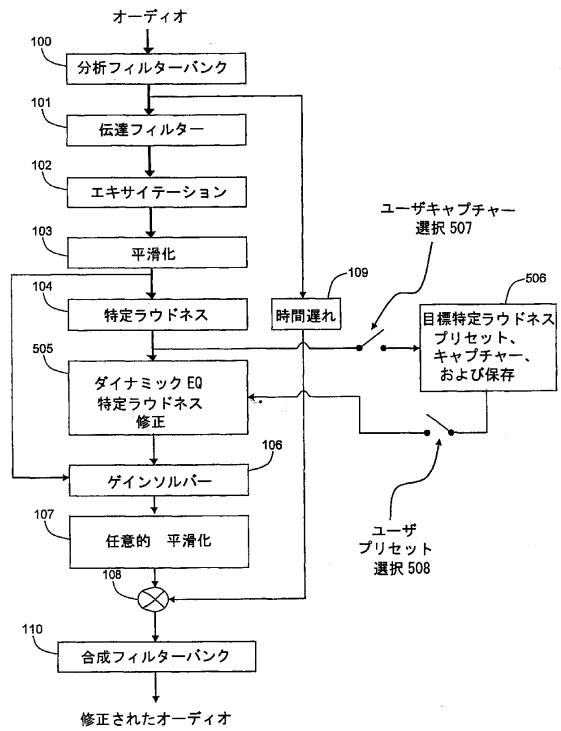


図 16

【図 17】

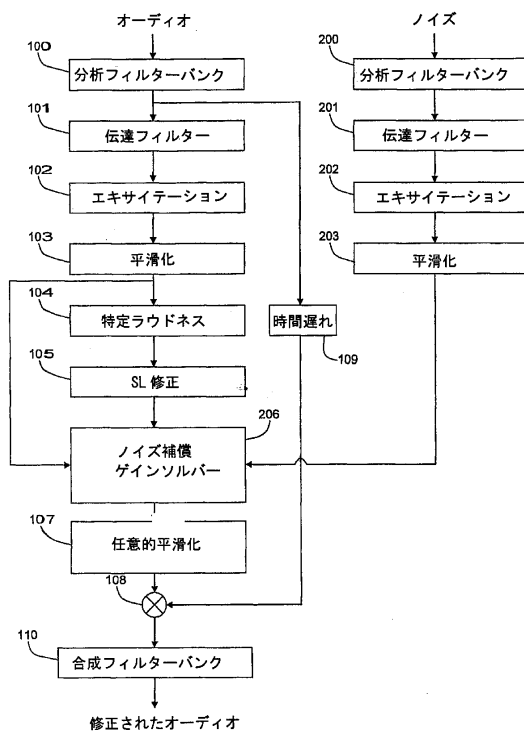


図 17

---

 フロントページの続き

(51)Int.Cl.

F I

H 0 3 G 3/30

B

H 0 3 G 3/30

C

(74)代理人 100121762

弁理士 杉山 直人

(74)代理人 100126767

弁理士 白銀 博

(74)代理人 100118647

弁理士 赤松 利昭

(74)代理人 100138519

弁理士 奥谷 雅子

(74)代理人 100120145

弁理士 田坂 一朗

(74)代理人 100138438

弁理士 尾首 亘聰

(74)代理人 100147740

弁理士 保坂 俊

(72)発明者 シーフェルト、アラン・ジェフリー

アメリカ合衆国、カリフォルニア州 9 4 1 0 3、サンフランシスコ、ポトレロ・アベニュー 1  
0 0

審査官 高橋 義昭

(56)参考文献 特開 2 0 0 3 - 2 6 4 8 9 2 ( J P , A )

特開平 5 - 1 7 5 7 7 2 ( J P , A )

特開平 1 1 - 2 2 5 3 9 8 ( J P , A )

特表 2 0 0 2 - 5 4 3 7 0 3 ( J P , A )

特開平 1 0 - 3 4 1 1 2 3 ( J P , A )

特開 2 0 0 2 - 7 8 0 9 6 ( J P , A )

(58)調査した分野(Int.Cl. , D B 名)

H03G 3/30 3/32 5/02 5/16 9/00

G10L 11/00 21/04