

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第5969727号
(P5969727)

(45) 発行日 平成28年8月17日 (2016. 8. 17)

(24) 登録日 平成28年7月15日 (2016. 7. 15)

(51) Int. Cl.	F I
H03G 9/02 (2006.01)	H03G 9/02 O25
G10L 21/034 (2013.01)	G10L 21/034
H03G 7/00 (2006.01)	H03G 7/00 O07
H03G 9/00 (2006.01)	H03G 9/00 O05

請求項の数 20 (全 16 頁)

(21) 出願番号	特願2016-510698 (P2016-510698)	(73) 特許権者	507236292
(86) (22) 出願日	平成26年4月14日 (2014. 4. 14)		ドルビー ラボラトリーズ ライセンシン
(65) 公表番号	特表2016-518788 (P2016-518788A)		グ コーポレイション
(43) 公表日	平成28年6月23日 (2016. 6. 23)		アメリカ合衆国 94103 カリフォル
(86) 国際出願番号	PCT/US2014/034024		ニア州 サンフランシスコ マーケット
(87) 国際公開番号	W02014/179021		ストリート 1275
(87) 国際公開日	平成26年11月6日 (2014. 11. 6)	(74) 代理人	100107766
審査請求日	平成27年10月22日 (2015. 10. 22)		弁理士 伊東 忠重
(31) 優先権主張番号	61/817, 175	(74) 代理人	100070150
(32) 優先日	平成25年4月29日 (2013. 4. 29)		弁理士 伊東 忠彦
(33) 優先権主張国	米国 (US)	(74) 代理人	100091214
早期審査対象出願			弁理士 大貫 進介

最終頁に続く

(54) 【発明の名称】 動的閾値を用いた周波数帯域圧縮

(57) 【特許請求の範囲】

【請求項 1】

入力オーディオ信号に応じて圧縮器の閾値を動的に調整する方法であって：

複数の周波数帯域成分をもつ入力オーディオ信号を受領する段階と；

前記複数の周波数帯域成分に応じて、複数の時間変化する閾値を決定する段階であって、各時間変化する閾値はそれぞれの周波数帯域成分に対応する、段階と；

圧縮器によって、各周波数帯域成分に対して、前記対応する時間変化する閾値をもつ圧縮動作を実行して、複数の利得を生成する段階であって、各利得はそれぞれの周波数帯域成分に対応する、段階と；

各利得を、遅延された対応する周波数帯域成分に適用して、複数の処理された周波数帯域成分を生成する段階と；

前記処理された周波数帯域成分を加算して出力信号を生成する段階と、

前記出力信号の知覚されるスペクトルの時間変化する推定値を決定する段階と；

前記知覚されるスペクトル推定値によって誘起される歪みスペクトルの時間変化する推定値を決定する段階と；

前記知覚されるスペクトル推定値の存在における前記歪みスペクトル推定値の可聴性指標を予測する段階と；

前記時間変化する閾値の一つまたは複数を、予測された可聴性指標に従って調整する段階とを含む、

方法。

10

20

【請求項 2】

前記歪みスペクトル推定値は、前記知覚されるスペクトル推定値に対する歪みモデルの応答に従って決定される、請求項 1 記載の方法。

【請求項 3】

前記歪みスペクトル推定値は、第一の周波数帯域成分の第一の推定された歪みを含み、前記第一の推定された歪みは、前記第一の周波数帯域成分におよび前記第一の周波数帯域成分より高い周波数の周波数帯域成分の少なくとも一部に誘起される歪みの最大として決定される、請求項 2 記載の方法。

【請求項 4】

前記知覚されるスペクトル推定値の決定は：

前記処理された周波数帯域成分に平滑化処理を適用することを含む、
請求項 1 ないし 3 のうちいずれか一項記載の方法。

【請求項 5】

前記知覚されるスペクトル推定値の存在における前記歪みスペクトル推定値の可聴性指標を予測する段階が、

前記知覚されるスペクトル推定値からマスキング閾値を計算し；

前記歪みスペクトル推定値と前記マスキング閾値との間の差を決定し；

決定された差の正の値を合計して前記の予測された可聴性指標を生成することを含む、
請求項 1 ないし 4 のうちいずれか一項記載の方法。

【請求項 6】

前記マスキング閾値は、前記知覚されるスペクトル推定値に基づくトーン性スペクトルを参照して計算され、前記トーン性スペクトルは、トーン的な周波数帯域成分からノイズ的な周波数帯域成分を区別するトーン性値を含む、請求項 5 記載の方法。

【請求項 7】

合計される、決定された差の正の値は、一つまたは複数の上の周波数帯域成分および一つまたは複数の下の周波数帯域成分が、該上および下の帯域成分の間の周波数帯域成分より低い重みをもつよう重み付けされる、請求項 5 または 6 記載の方法。

【請求項 8】

前記時間変化する閾値はさらに、複数の固定閾値に従って決定される、請求項 1 ないし 7 のうちいずれか一項記載の方法。

【請求項 9】

各時間変化する閾値は、周波数帯域成分に従ってかつ前記複数の固定閾値に従って決定される、請求項 8 記載の方法。

【請求項 10】

各時間変化する閾値は、前記対応する周波数帯域成分に従ってかつ対応する固定閾値に従って決定される、請求項 9 記載の方法。

【請求項 11】

歪みの可聴性指標を予測し；

予測された可聴性指標を規格化し；

前記固定閾値の一つまたは複数を参照して、かつオフセット値に適用される前記の規格化された可聴性指標に従って、前記時間変化する閾値の一つまたは複数を上げるまたは下げることを含む、
請求項 8 記載の方法。

【請求項 12】

前記出力信号のデータを記憶媒体上に記憶する段階をさらに含む、
請求項 1 ないし 11 のうちいずれか一項記載の方法。

【請求項 13】

請求項 1 記載の動作を実行するよう構成された装置。

【請求項 14】

請求項 2 ないし 12 のうち一つまたは複数の記載の動作を実行するよう構成された装置

10

20

30

40

50

。

【請求項 1 5】

前記入力オーディオ信号をフィルタリングして前記複数の周波数帯域成分を生成することができるフィルタリング・モジュールをさらに有する、
請求項 1 3 または 1 4 記載の装置。

【請求項 1 6】

前記フィルタリング・モジュールが、複数の帯域通過フィルタを有するマルチ帯域フィルタを有し、各帯域通過フィルタはそれぞれの周波数帯域成分に対応する、請求項 1 5 記載の装置。

【請求項 1 7】

前記出力信号を受領するよう結合された一つまたは複数の増幅器をさらに有し、前記一つまたは複数の増幅器は、前記出力信号を増幅して、増幅された出力信号を生成することができる、請求項 1 3 ないし 1 6 のうちいずれか一項記載の装置。

【請求項 1 8】

前記出力信号または前記増幅された出力信号を受領するよう結合されたディスプレイ装置をさらに有し、前記ディスプレイ装置は、前記受領された信号に関連するグラフィック・データを表示することができる、請求項 1 7 記載の装置。

【請求項 1 9】

入力オーディオ信号に応じて圧縮器の閾値を動的に調整するために方法を実行させるようコンピューティング装置によって実行可能な命令を記憶しており、前記方法が請求項 1 記載の動作を含む、非一時的なコンピュータ可読記憶媒体。

【請求項 2 0】

前記方法がさらに、請求項 2 ないし 1 2 のうち一つまたは複数の記載の一つまたは複数の動作を含む、請求項 1 9 記載の非一時的なコンピュータ可読記憶媒体。

【発明の詳細な説明】

【技術分野】

【0 0 0 1】

関連出願への相互参照

本願は2013年4月29日に出願されたAlan J. Seefedtによる「歪み可聴性モデルに基づく動的閾値を用いた、歪み低減マルチ帯域圧縮器」という名称の本願と同時係属中であり同一被譲渡者の米国仮特許出願第61/817,175号（代理人整理番号D13064USP1）の優先権を主張するものである。同出願の内容は参照により、あらゆる目的についてその全体において組み込まれる。

【0 0 0 2】

技術分野

本特許開示は概括的にはオーディオ呈示に、詳細には呈示の際の歪み低減に関する。

【背景技術】

【0 0 0 3】

スマートフォンおよびタブレットのようないくつかのオーディオ再生装置は限られた出力能力をもつ増幅器およびスピーカーを含む。そのような装置では、オーディオ再生は知覚的に歪められることがあり、呈示中に再生レベルが上げられるとしばしば鋭く歪められることがある。歪みの特性はしばしば再生装置にとって周波数依存である。たとえば、テレビジョン・キャビネットは特定の周波数で共鳴応答を示すことがある。この共鳴応答は、その特定の周波数における出力信号が特定のレベルに達するかそれを超えるとときに励起され、わずらわしいガタガタ音を生じる。

【0 0 0 4】

歪みを低減し、再生装置での再生レベルを最大化しようとするために、マルチ帯域圧縮が再生前のオーディオ信号に適用されることができる。信号の各周波数帯域について歪み閾値が指定される。圧縮器は、各帯域における信号レベルが対応する歪み閾値を超えないことを保証するよう、各帯域に独立な利得を適用する。

【発明の概要】

【課題を解決するための手段】

【0005】

入力オーディオ信号に応じて圧縮器の閾値を動的に調整する技法を実装するシステム、装置、方法およびコンピュータ可読記憶媒体のいくつかの例が開示される。いくつかの実装では、いくつかの周波数帯域成分をもつ入力信号が処理される。それらの周波数帯域成分に応じて、時間変化する閾値が決定されることができる。各時間変化する閾値は一般に、それぞれの周波数帯域成分に対応する。圧縮器が、各周波数帯域成分に対して、前記対応する時間変化する閾値をもつ圧縮動作を実行して、いくつかの利得を生成する。ここで、各利得はそれぞれの周波数帯域成分に対応する。各利得は、遅延された対応する周波数帯域成分に適用されて、いくつかの処理された帯域成分を生成し、それらの処理された帯域成分は加算されて出力信号を生成する。

10

【0006】

いくつかの実装では、出力信号の知覚されるスペクトルの時間変化する推定値が決定される。また、知覚されるスペクトル推定値によって誘起される歪みスペクトルの時間変化する推定値が、たとえば歪みモデルの応答に従って決定される。知覚されるスペクトル推定値の存在における歪みスペクトル推定値の可聴性指標が予測される。時間変化する閾値は、予測された可聴性指標に従って調整されることができる。

【0007】

いくつかの実装では、可聴性指標を予測することは、知覚されるスペクトル推定値からマスキング閾値を計算し；歪みスペクトル推定値とマスキング閾値との間の差を決定し；決定された差の正の値を合計して予測された可聴性指標を生成することを含む。マスキング閾値は、知覚されるスペクトル推定値に基づくトーン性スペクトルを参照して計算されることができる。トーン性スペクトルは、トーン的な周波数帯域成分からノイズ的な周波数帯域成分を区別するトーン性値を含む。合計される、決定された差の正の値は、一つまたは複数の上の周波数帯域成分および一つまたは複数の下の周波数帯域成分が、該上および下の帯域成分の間の周波数帯域成分より低い重みをもつよう重み付けされることができる。

20

【0008】

いくつかの実装では、時間変化する閾値は、少なくとも部分的には、固定された名目閾値に従って決定される。このように、各時間変化する閾値は、対応する周波数帯域成分に従ってかつそれぞれの固定した閾値に従って決定されることができる。これらの実装において、歪みの予測される可聴性指標は規格化されることができる。時間変化する閾値は、一つまたは複数の固定した閾値を参照して、かつオフセット値に適用される規格化された可聴性指標に従って上げられたり下げられたりすることができる。

30

【0009】

いくつかの実装では、入力オーディオ信号にตอบสนองして圧縮閾値を動的に調整する装置は、上記の動作の一部または全部を実行させるよう動作可能な一つまたは複数のコントローラを含む。さらに、そのような装置は、前記入力信号をフィルタリングして前記周波数帯域成分を生成することができるフィルタリング・モジュールおよびいくつかの帯域通過フィルタを含むマルチ帯域フィルタを含むことができる。ここで、各帯域通過フィルタはそれぞれの周波数帯域成分に対応する。本装置は、前記出力信号を受領し、増幅して増幅された出力信号を生成するよう結合された一つまたは複数の増幅器と、増幅された出力信号を受領して再生するよう結合された一つまたは複数のスピーカーとを含むこともできる。

40

【0010】

いくつかの実装では、上記の動作の一部または全部を実行させるようコンピューティング装置によって実行可能な命令を記憶する非一時的なコンピュータ可読記憶媒体。コンピューティング装置の限定しない例は、サーバーおよびデスクトップ・コンピュータならびにスマートフォン、タブレット、携帯型音楽プレーヤーなどといった携帯型ハンドヘルド装置を含む。いくつかの事例では、一つまたは複数のサーバーが、クラウド・ベースのサ

50

ービスの一部として、開示される技法の一つまたは複数を使って入力オーディオ信号を処理し、処理された出力信号をインターネットを通じてユーザーの装置にストリーミングするよう構成されることができる。

【図面の簡単な説明】

【0011】

含まれる図面は例解目的のためであり、開示される発明のシステム、装置、方法およびコンピュータ可読記憶媒体についての可能な構造および動作の例を提供するだけのものである。これらの図面はいかなる意味でも、開示される実装の精神および範囲から外れることなく当業者によってなされうる形および細部におけるいかなる変化をも制限するものではない。

10

【図1】いくつかの実装に基づく、入力オーディオ信号に応じて圧縮器の閾値を動的に調整する歪み可聴性モデル(DAM: distortion audibility model)を組み込む圧縮器100の例を示している。

【図2】いくつかの実装に基づいて実行される、入力オーディオ信号に応じて圧縮器の閾値を動的に調整する方法200の例を示す図である。

【図3】いくつかの実装に基づいて実行される、入力オーディオ信号に応じて圧縮器の閾値を動的に調整する方法300の別の例を示す図である。

【図4】いくつかの実装に基づいて実行される、知覚されるスペクトル推定値の存在において歪みスペクトル推定値の可聴性指標を予測する方法400の例を示す図である。

【図5】Aは、いくつかの実装に基づく、表示装置に表示される入力オーディオ信号の周波数帯域成分のエネルギー・レベルのグラフィック表現500Aの例を示す。Bは、いくつかの実装に基づく、表示装置に表示される、聴取者によって知覚される推定された出力信号スペクトルの対応するエネルギー・レベルのグラフィック表現500Bの例を示す。

20

【図6】いくつかの実装に基づく、オーディオ信号を呈示するための圧縮器100を組み込む再生システム600の例を示す図である。

【発明を実施するための形態】

【0012】

入力オーディオ信号に応じて圧縮器の閾値を動的に調整する技法を実装するシステム、装置、方法およびコンピュータ可読記憶媒体のいくつかの例が開示される。開示される技法のいくつかは、マルチ帯域圧縮器によって適用されることのできる動的閾値を決定するための歪み可聴性モデルを組み込む。歪み可聴性モデルは、動的な信号依存の仕方で、入力信号の存在におけるその入力信号によって誘起される歪みの知覚される可聴性を予測するよう構成される。この予測される可聴性は、圧縮器の閾値を動的に修正するために使用されることができる。こうして、開示される技法を組み込むいくつかの装置およびシステムは、最小限の知覚される歪みをもって増大した再生レベルに対応できる。

30

【0013】

いくつかの実装では、歪み可聴性モデルは、聴取者に聞こえる信号スペクトルの時間変化する推定値およびその信号によって誘起される歪みスペクトルの時間変化する推定値を予測するよう構成される。その際、歪み可聴性モデルは、信号スペクトル推定値の存在における、この歪みスペクトル推定値の可聴性を予測する。このようにして、圧縮器の一つまたは複数の時間変化する閾値がしかるべく動的に変調されることができる。

40

【0014】

圧縮閾値の動的な調整のための開示される技法は、他の圧縮プロセスとの関連で使われてもよく、スマートフォン、タブレット、ラップトップ・コンピュータ、携帯型音楽プレーヤー、テレビジョン、モニタおよびサーバー・ベースのシステムといったさまざまな装置およびシステムにおいて実装されることができる。

【0015】

開示される技法を実装するいくつかの装置およびシステムは、音色、つまり等しいラウドネスおよびピッチの二つの音が異なるように知覚されうる聴取者知覚の属性を知覚できるほどに変えてしまう通常のマルチ帯域圧縮器に対して改善する。ある種の周波数が歪み

50

閾値に達し、他の周波数が達しないとき、いくつかの通常の圧縮器は、これらの周波数の間の相対バランスを変更することによって不都合を導入してしまう。結果として得られる音は、異常なものとして現われ、不自然な聴取体験につながる。

【 0 0 1 6 】

さらに、利得が積極的すぎる場合には、再生レベルは不必要に低減されることがある。各帯域における閾値が、その帯域を中心とする狭帯域信号についての知覚される歪みをなくすよう設定されれば、広帯域信号が圧縮器を通過することから帰結する減衰はしばしば、任意の誘起される歪みを知覚的に消去するために必要とされるよりも強い。これは、広帯域信号は、該広帯域信号が誘起する歪みの一部を有意にマスクすることがありうる一方、狭帯域信号はその誘起される歪みをマスクする効果がずっと弱いことがありうるという事実のためである。

【 0 0 1 7 】

図 1 は、入力オーディオ信号に応じて圧縮器の閾値を動的に調整するための歪み可聴性モデル (DAM) を組み込む圧縮器 1 0 0 の例を示している。図 1 では、フィルタバンク 1 0 4 の形のフィルタリング・モジュールが入力オーディオ信号 $x[n]$ を受領する。フィルタバンク 1 0 4 は、入力信号 $x[n]$ をフィルタリングして入力信号 $x[n]$ をいくつかの周波数帯域成分 $x_1[n] \sim x_B[n]$ に分離するよう構成されている。いくつかの実装では、フィルタバンク 1 0 4 は、B 個の帯域通過フィルタとして実装されるマルチ帯域フィルタとして構成される。ここで、各帯域通過フィルタがそれぞれの周波数帯域成分に対応する。たとえば、フィルタバンク 1 0 4 の各帯域 b は、式 (1)

$$x_b[n] = h_b[n] * x[n] \quad b = 1 \dots B \quad (1)$$

に表わされるように入力信号 $x[n]$ を帯域通過フィルタ応答 $h_b[n]$ と畳み込みしたものとして計算される周波数帯域成分 $x_b[n]$ を出力するよう構成されることができる。

【 0 0 1 8 】

図 1 において、DAM 1 0 8 はフィルタバンク 1 0 4 から出力される周波数帯域成分 $x_1[n] \sim x_B[n]$ を受領するよう結合された入力をもつ。DAM 1 0 8 に提供される各周波数帯域成分 $x_b[n]$ は、図 1 に示されるように、それぞれの固定した圧縮閾値 L_b と対にされる。DAM 1 0 8 は、時間変化する閾値 $D_1[n] \sim D_B[n]$ を計算して出力するよう構成される。いくつかの実装では、時間変化する閾値 $D_b[n]$ は、それぞれの周波数帯域成分 $x_b[n]$ の関数として計算される。いくつかの他の実装では、各帯域について独立に計算されるのではなく、各時間変化する閾値 $D_b[n]$ は、式 (2)

$$D_b[n] = \text{DAM}(\{x_i[n], L_i \mid i = 1 \dots B\}) \quad (2)$$

に表わされるように、諸帯域 $b = 1 \dots B$ にわたるすべての周波数帯域成分 $x_1[n] \sim x_B[n]$ および固定した閾値 L_b の関数として計算される。

【 0 0 1 9 】

各周波数帯域成分 $x_b[n]$ は、それぞれの時間変化する閾値 $D_b[n]$ とともに圧縮関数 (CF: compression function) 1 1 2 に入力として与えられる。該時間変化する閾値は、それより上でその帯域 b における信号が歪みを生じはじめるレベルを表わす。各圧縮関数 1 1 2 b は、式 (3)

$$g_b[n] = \text{CF}(x_b[n], D_b[n]) \quad (3)$$

に表わされるように、周波数帯域成分 $x_b[n]$ および時間変化する閾値 $D_b[n]$ を処理して時間変化する利得 $g_b[n]$ を生成するよう構成される。これは、帯域 b をその限界閾値 L_b より下に保つための利得を表わす。

【 0 0 2 0 】

周波数帯域成分 $x_1[n] \sim x_B[n]$ 全部の遅延されたバージョンにそれらの対応する利得信号 $g_1[n] \sim g_B[n]$ を乗算したものを合計することによって、処理された出力信号 $y[n]$ が計算される。図 1 では、乗算器ユニット 1 1 6 は利得信号を遅延された周波数帯域成分に乘算して ($y_b[n] = g_b[n] x_b[n - d]$) 処理された帯域成分 $y_1[n] \sim y_B[n]$ を生成するよう構成されている。これらの処理された帯域成分が次いで加算ユニット 1 2 0 において合計されて出力信号 $y[n]$ を生じる。たとえば、遅延 d は、利得信号の計算に関連する任意の処理遅延を考

10

20

30

40

50

慮に入れるよう設計されることができる。式(4)は、処理された信号 $y[n]$ の生成の表現を表わしている。

【 0 0 2 1 】

【 数 1 】

$$y[n] = \sum_{b=1}^B g_b[n] x_b[n-d] \quad (4)$$

図2は、いくつかの実装に基づいて実行される、入力オーディオ信号に応じて圧縮器の閾値を動的に調整する方法200の例を示している。図2は、図1の例を参照して記述される。図2の204において、上記で説明したように、周波数帯域成分 $x_1[n] \sim x_B[n]$ がDAM 108への入力として受領される。208では、DAM 108が周波数帯域成分 $x_1[n] \sim x_B[n]$ に従って時間変化する閾値 $D_1[n] \sim D_B[n]$ を決定する。212では、各CF 112bは、対応する時間変化する閾値 $D_b[n]$ を使って対応する周波数帯域成分 $x_b[n]$ に対して圧縮動作を実行して利得 $g_1[n] \sim g_B[n]$ を生成するよう構成される。216では、各利得 $g_b[n]$ は、たとえば乗算器ユニット116を使って、対応する周波数帯域成分 $x_b[n]$ の遅延されたバージョンに適用され、処理された帯域成分 $y_1[n] \sim y_B[n]$ を生成する。220では、処理された帯域成分 $y_1[n] \sim y_B[n]$ は加算ユニット120において加算されて出力信号 $y[n]$ を生成する。次いでこれがメモリ装置に記憶されるおよび/または追加的な処理モジュールに提供されることができる。

【 0 0 2 2 】

図3は、いくつかの実装に基づいて実行される、入力オーディオ信号に応じて圧縮器の閾値を動的に調整する方法300のもう一つの例を示している。方法300の動作304～320は上記の方法200の動作204～220と実質的に同じである。図3では、出力信号 $y[n]$ の生成に続いて、324において、聴取者に聞こえる出力信号 $y[n]$ のパワースペクトルの時間変化する推定値が推定されることができる。この目的に向け、平滑化動作が処理された帯域成分 $y_1[n] \sim y_B[n]$ に適用されることができる。たとえば、各処理された帯域成分 $y_b[n] = g_b[n] x_b[n-d]$ が高速アタック/低速リリース一極平滑化器 (fast-attack/slow-release one pole smoother) を用いて平滑化されることができる。 $s_b[n]$ は、聴取者によって知覚される出力信号スペクトルの推定値を表わし、式(5)において計算される。

【 0 0 2 3 】

【 数 2 】

$$s_b[n] = \begin{cases} \alpha_A s_b[n-1] + (1 - \alpha_A) y_b^2[n], & y_b^2[n] \geq s_b[n-1] \\ \alpha_R s_b[n-1] + (1 - \alpha_R) y_b^2[n], & \text{他の場合} \end{cases} \quad (5)$$

圧縮閾値のその後の変調から生じるアーチファクトを軽減するために、いくつかの事例では、利得 $g_1[n] \sim g_B[n]$ のアタックおよびリリースを支配するために式(5)において使われるものよりやや速いアタックおよびやや遅いリリースを利用することが望ましいことがある。そのような事例では、聴取者によって知覚される推定された出力信号スペクトルは、式(6)

$$S_b[n] = 10 \log_{10}(s_b[n]) \quad (6)$$

に示されるように、デシベル (dB) で表わされることができる。

【 0 0 2 4 】

図3では、328において、知覚されるスペクトル推定値 $s_b[n]$ によって誘起される歪みスペクトルの時間変化する推定値が決定される。たとえば、 $s_b[n]$ によって再生システムにおいて誘起される歪みは、 $s_b[n]$ を歪み生成モデルに入力することによって推定でき

10

20

30

40

50

る。そのようなモデルは、個々の再生システムを測定し、特徴付けることによって生成されてもよい。いくつかの場合には、小型スピーカをもつ装置についての一般化されたモデルが十分であり、帯域 b における信号が、 b を含め b 以上のすべての帯域に、帯域 b における信号レベルより何らかの固定オフセットだけ下のレベルで歪みを誘起すると想定する。

【 0 0 2 5 】

いくつかの実装では、任意の所与の帯域における歪みスペクトル推定値は、その帯域中に生成される歪みの、すべての帯域にわたる最大によって与えられる。こうして、第一の周波数帯域成分の第一の推定された歪みは、第一の周波数帯域成分中におよび前記第一の周波数帯域成分より高い周波数の周波数帯域成分の少なくとも一部に誘起される歪みの最大として決定されることができる。これは、任意の単一の帯域が一般に、当該帯域を含む当該帯域以上の諸帯域に歪みを生成するからである。図 1 および図 2 との関係で上記したように時間変化する閾値のはたらきをする歪みスペクトル推定値 $D_1[n] \sim D_B[n]$ は、最低帯域から始まって最高帯域まで逐次的な仕方では計算される。これは式(7)に例解される。ここで、固定オフセット D_{offset} についての値の例は40dBである。

【 0 0 2 6 】

【 数 3 】

$$D_1[n] = S_1[n] - D_{offset} \quad (7)$$

$$D_b[n] = \max\{D_{b-1}[n], S_b[n] - D_{offset}\} \quad b = 2 \dots B$$

図 3 では、3 3 2 において、知覚されるスペクトル推定値 $s_b[n]$ の存在における歪みスペクトル推定値 $D_b[n]$ の可聴性指標が、たとえば図 1 の DAM 1 0 8 を使って予測される。後述する図 4 は、この可聴性指標を予測するための技法の例を示す。

【 0 0 2 7 】

図 3 では、3 3 6 において、時間変化する閾値の一つまたは複数が、予測された可聴性指標 $A[n]$ に従って調整される。いくつかの実装では、予測された可聴性指標 $A[n]$ は規格化され、時間変化する閾値 $D_b[n]$ の一つまたは複数が、固定した諸閾値 L_b の一つまたは複数を参照して、かつオフセット値に適用される規格化された可聴性指標に従って、上げられたり下げられたりする。たとえば、予測された可聴性指標は、時間変化する閾値 $D_1[n] \sim D_B[n]$ を変調するために利用される前に、式(8)に示されるように、0と1の間に規格化されることができる。

【 0 0 2 8 】

【 数 4 】

$$A_{norm}[n] = \begin{cases} 0 & A[n] < A_{low} \\ \frac{A[n] - A_{low}}{A_{high} - A_{low}} & A_{low} < A[n] < A_{high} \\ 1 & A_{high} < A[n] \end{cases} \quad (8)$$

式(8)の例において、規格化限界は、規格化された予測された可聴性指標 $A_{norm}[n]$ が0に等しいときは誘起される歪みが出力信号によってよくマスクされ、 $A_{norm}[n]$ が1に等しいときは歪みが可聴性の縁にあるよう、選ばれる。したがって、 $A_{norm}[n]$ が0に等しいときは、時間変化する閾値 $D_1[n] \sim D_B[n]$ はよりラウドな再生を許容するよう上げられることができるが、 $A_{norm}[n]$ が1に等しいときは、閾値 $D_1[n] \sim D_B[n]$ はそれらの名目値のままとなる。よって、閾値 $D_1[n] \sim D_B[n]$ は

$$D_b[n] = L_b + (1 - A_{norm}[n])L_{offset} \quad (9)$$

に従って固定した閾値 L_b から計算できる。

【 0 0 2 9 】

式(9)において、 $A_{norm}[n]$ が0に等しいときは、閾値 $D_b[n]$ は L_{offset} dBだけその名目値より上に上げられる。ある事例では、 L_{offset} を6dBの範囲に設定すると、歪みの知覚される増大はなしに、広帯域の信号についての知覚されるラウドネスにおける知覚的な実質的な増大を与えた。他の事例では、 L_{offset} は特定の再生装置に合わせて調整された。

【 0 0 3 0 】

図4は、いくつかの実施形態に基づいて実行される、知覚されるスペクトル推定値の存在において歪みスペクトル推定値の可聴性指標を予測する方法400の例を示している。404では、マスキング閾値 $M_b[n]$ が知覚されるスペクトル推定値 $s_b[n]$ から計算される。マスキング閾値 $M_b[n]$ についての一つのモデルは、知覚されるスペクトル推定値より固定オフセット M_{offset} 低いというものである。これは式(10)

$$M_b[n] = S_b[n] - M_{offset} \quad (10)$$

において示される。ここで、 M_{offset} についての値の例は10dBである。

【 0 0 3 1 】

あるいはまた、マスキング信号のトーン性の関数としてのマスキングの可変性を考慮に入れるマスキング・モデルが使用されてもよい。トーン的な信号のマスキング能力はノイズ的な信号より有意に低いことが一般に知られている。よって、マスキング閾値 $M_b[n]$ は、 $s_b[n]$ に基づいてトーン性スペクトルを参照して計算されることができる。トーン性スペクトルは、ノイズ的な周波数帯域成分をトーン的な周波数帯域成分から区別するトーン性値を含む。各帯域における $s_b[n]$ のトーン性を、トーン性スペクトル $T_b[n]$ を生成するための既知の技法を使って特徴付けてもよい。ここで、 $T_b[n]$ は0から1まで変わる。0はノイズ的な信号を示し、1はトーン的な信号を示す。このトーン性スペクトルを利用して、マスキング閾値は式(11)

$$M_b[n] = S_b[n] - (T_b[n]M_{tone} + (1 - T_b[n])M_{noise}) \quad (11)$$

に表わされるように計算されてもよい。

【 0 0 3 2 】

ある試験事例において、 $M_{tone} = 30\text{dB}$ および $M_{noise} = 10\text{dB}$ が適切な値の例であった。これはノイズ的な信号よりトーン性信号について20dB低いマスキングを与える。

【 0 0 3 3 】

図4では、408において、諸帯域にわたって歪みスペクトル推定値 $D_b[n]$ がマスキング閾値 $M_b[n]$ と比較され、 $D_b[n]$ と $M_b[n]$ との間の差があればそれを決定する。412では、決定された差のあらゆる正の値が加算されて、予測された可聴性指標 $A[n]$ を生成する。いくつかの実装では、 $A[n]$ は、式(12)に示されるように、 $D_b[n]$ と $M_b[n]$ との間の差の正の値の重み付けされた和として計算される。

【 0 0 3 4 】

【 数 5 】

$$A[n] = \sum_b W_b \max\{D_b[n] - M_b[n], 0\} \quad (12)$$

式(12)において、いくつかの実装では、重み付け W_b は知覚的に動機付けられて、高周波数帯域および低周波数帯域が中間周波数帯域より低く重み付けされてもよい。

【 0 0 3 5 】

いくつかの他の実装では、明示的な歪み生成およびマスキング・モデルを利用するのではなく、歪み可聴性の指標が信号スペクトル $S_b[n]$ の関数から推定されてもよい。一つのそのような例は、式(13)に示されるような、諸帯域にわたるこのスペクトルの標準偏差である。

【 0 0 3 6 】

【数 6】

$$A[n] = \sqrt{\frac{1}{B} \sum_b \left(S_b[n] - \frac{1}{B} \sum_i S_i[n] \right)^2} \quad (13)$$

標準偏差が小さいときは、すべての帯域の値はほぼ同じである。つまり、 $S_b[n]$ は大まかに広帯域である。この場合、 $S_b[n]$ は歪みをそこそよくマスクするはずである。標準偏差が比較的大きい場合には、 $S_b[n]$ の値は有意に変動するということであり、スペクトルにおける、歪みが可聴となる「穴」の可能性を示す。結果として、式(13)における値 $A[n]$ はごく大まかに、式(12)での値の振る舞いに一致する。式(13)からの可聴性値は次いで、歪み生成およびマスキング・モデルで使われたのとは異なる規格化限界をもって式(8)に従って規格化され、次いで式(9)におけるように、閾値 $D_b[n]$ を変調するために利用されてもよい。

【0037】

図5のAは、いくつかの実装に基づく、表示装置に表示される、入力オーディオ信号の周波数帯域成分のエネルギー・レベルのグラフィック表現500Aの例を示す。図5のBは、いくつかの実装に基づく、表示装置に表示される、聴取者によって知覚される推定された出力信号スペクトルの対応するエネルギー・レベルのグラフィック表現500Bの例を示す。

【0038】

図5のAおよびBの例では、図1の圧縮器100は上記のように入力オーディオ信号 $x[n]$ を処理する。図5のAでは、入力信号の20個の周波数帯域成分 $x_1[n] \sim x_{20}[n]$ が図1のフィルタバンク104によって生成され、人間の聴覚の知覚的な分解能を模倣するよう離間される。図1の圧縮関数112の入力となる各周波数帯域成分 $x_b[n]$ の信号エネルギーは、図5のAにおけるバー504によって表わされる。上記のように計算された各周波数帯域成分についての時間変化する閾値 $D_1[n] \sim D_{20}[n]$ はクリアなセグメント508によって表わされる。圧縮関数112から出力される利得 $g_1[n] \sim g_{20}[n]$ はセグメント512によって描かれる。

【0039】

図5のBでは知覚されるスペクトル推定値 $S_1[n] \sim S_{20}[n]$ のエネルギー・レベルがバー514によって表わされる。上記のように知覚されるスペクトル推定値 $s_b[n]$ から計算されたマスキング閾値 $M_1[n] \sim M_{20}[n]$ は図5のBではセグメント516によって表わされる。図1のDAM 108によって生成された時間変化する閾値 $D_1[n] \sim D_{20}[n]$ は図5のBにおけるクリアなセグメント520によって表わされる。図1では、DAM 108は図5のBの時間変化する閾値 $D_1[n] \sim D_{20}[n]$ 520を、圧縮器100を通じて再生されるオーディオ信号の存在において聴取者によって知覚される何らかの誘起された歪みが生じる前に、可能な最大レベルまで増大させる。一般に、広帯域信号については、時間変化する閾値 $D_1[n] \sim D_{20}[n]$ は、それらの名目上設定されたレベル L_b より上まで増大させられる。誘起される歪みがあったとしてもそのかなりの部分を信号自身がマスクするからである。狭帯域信号については、時間変化する閾値はそれらの名目レベルの近くに留まる可能性が高い。何らかの誘起される歪みに対して信号がほとんどマスキングをもたらさないからである。

【0040】

図6は、いくつかの実装に基づく、オーディオ信号を呈示するための圧縮器100を組み込む再生システム600の例を示す図である。図6では、コントローラ604が上記の圧縮器100を実装するよう構成される。たとえば、図2～図4を参照して上記した動作の一つまたは複数がコントローラ604によって実行されることができるとはまたコントローラ604によって実行されるようにされることができるとは。コントローラ604は、下記

でより詳細に説明するように幅広い多様な異なる構成をもつことができる。さらに、コントローラ 604 は、図 6 の場合におけるようにユーザー装置上に実装されることができるし、あるいは場合によっては一つまたは複数のサーバー上に実装されることができる。

【0041】

図 6 に示したものに対する代替例として、コントローラ 604 を実装するサーバーが、オーディオをユーザーの装置に、好適なネットワークを通じてストリーミングすることができる。当業者は、そのようなネットワークがサーバーと、スマートフォンまたはタブレットのようなユーザー装置との間の通信を提供することを理解するはずである。ネットワークは、たとえば、TCP/IP ベースのネットワーク、テレコミュニケーション・ネットワーク、無線ネットワーク、ケーブル・ネットワーク、公共ネットワーク、私設ネットワーク、広域ネットワーク、ローカル・エリア・ネットワーク、インターネット、ワールドワイドウェブ、イントラネット、エクストラネットなどを含む幅広い多様なネットワーク環境の任意の部分集合または組み合わせであることができる。

【0042】

図 6 に戻ると、再生システム 600 は、圧縮器 100 から出力信号 $y[n]$ を受領して $y[n]$ をアナログ信号に変換するよう結合されたデジタル アナログ変換器 (DAC) 608 を含む。DAC 608 はこの例ではコントローラ 604 の一つの処理モジュールとして実装されているが、当業者は DAC 608 がコントローラ 604 とは別個のユニットとして、場合によってはコントローラ 604 が位置しているサーバーまたはユーザーの装置から別個の装置として構築されることができることを理解するはずである。

【0043】

図 6 では、前置増幅器 612 が、DAC 608 によって出力されるアナログ信号の信号強度を受領し、ブーストするよう結合されている。電流感応前置増幅器、寄生容量前置増幅器、電荷感応前置増幅器またはそれらの何らかの組み合わせといったさまざまな型の前置増幅器が使用されることができる。当業者は、いくつかの代替例では前置増幅器 612 がシステム 600 から省略できることを理解するはずである。

【0044】

図 6 では、電力増幅器 616 は前置増幅器 612 からのブーストされた信号を受領し、増幅して、一つまたは複数のスピーカ 620 を駆動するのに好適な増幅された出力信号を生成するよう結合されている。前置増幅器 612 がシステム 600 から省略される状況では、電力増幅器 616 が、統合された増幅器で置換されてもよい。システム 600 の増幅器およびスピーカは、上記のようなユーザーの装置のコンポーネントであることができる。いくつかの実装では、ユーザーの装置と一体のまたはユーザーの装置に接続された別個の装置の形のディスプレイが制御されて、図 5 の A および B に呈示した情報のような、前記出力信号の諸周波数帯域および前記信号を特徴付ける他の情報を示すグラフィック・データを表示することができる。他のいくつかのシナリオでは、図 6 の増幅器 612 およびスピーカ 620 は公共アナウンス (PA: public address) システム、劇場もしくはコンサート・サウンド・システムまたはステレオもしくはホームシアター・システムのような家庭システムの一部である。

【0045】

本稿に記載した技法は、一つまたは複数のコンピューティング装置によって実装されることができる。たとえば、特殊目的コンピューティング装置のコントローラが、開示される動作を実行するようまたはそのような動作を実行させるよう固定結線されてもよく、動作を実行するまたは動作を実行させるよう持続的にプログラムされた一つまたは複数の特定用途向け集積回路 (ASIC) またはフィールド・プログラマブル・ゲート・アレイ (FPGA) のようなデジタル電子回路を含んでいてもよい。いくつかの実装では、カスタムの固定結線論理、ASIC および / または FPGA でカスタム・プログラミングをもつものが組み合わされて上記技法を達成する。

【0046】

他のいくつかの実装では、汎用コンピューティング装置が、ファームウェア、メモリ、

10

20

30

40

50

他の記憶またはその組み合わせの中のプログラム命令に従って開示される動作の一つまたは複数を実行させるようプログラムされたコントローラを含むことができる。汎用コンピューティング装置の例は、サーバー、ネットワーク装置およびスマートフォン、タブレット、ラップトップ、デスクトップ・コンピュータ、携帯型メディアプレーヤー、他のさまざまな携帯型ハンドヘルド装置のようなユーザー装置ならびに開示される動作を実装するまたは該動作を実装および実行させるデータ処理ハードウェアおよび/またはプログラム論理を組み込む他の任意の装置を含む。コンピューティング装置は、モニタ、プリンタまたは本稿で言及される結果の任意のものをユーザーに提供するための他の好適なディスプレイを含んでいてもよい。

【0047】

本稿での用法では、用語「記憶媒体」および「記憶メディア」は、コンピュータまたは機械の型に特定の仕方で動作させるデータおよび/または命令を記憶する任意の媒体を指す。本稿に記載されるモデル、モジュール、ユニット、エンジンおよび動作の任意のものは、任意の好適なコンピュータ言語を使ってコントローラのプロセッサによって実行可能なソフトウェア・コードとして実装されてもよいし、そのようなソフトウェア・コードによって実装されるようにされてもよい。ソフトウェア・コードは、記憶および/または伝送のためのコンピュータ可読媒体上の一連の命令またはコマンドとして記憶されてもよい。好適なコンピュータ可読媒体の例は、ランダム・アクセス・メモリ (RAM)、読み出し専用メモリ (ROM)、ハードドライブまたはフロッピーディスクのような磁気媒体、コンパクトディスク (CD) または DVD (デジタル多用途ディスク) のような光学式媒体、半導体ドライブ、フラッシュ・メモリおよび他の任意のメモリ・チップまたはカートリッジを含む。コンピュータ可読媒体は、そのような記憶デバイスの任意の組み合わせであってもよい。ソフトウェア/プログラム・コードをエンコードされたコンピュータ可読媒体は、上記のようなユーザー装置またはサーバーのような互換な装置と一緒にパッケージングされてもよいし、あるいは他の装置とは別個に提供されてもよい。任意のそのようなコンピュータ可読媒体は、単一のコンピューティング装置の上もしくは中にまたはコンピュータ・システム全体の上または中に存在してもよく、システムまたはネットワーク内の他にもあるコンピュータ可読媒体のうちのものであってもよい。

【0048】

記憶媒体は、伝送媒体とは異なるが、伝送媒体との関連で使われてもよい。伝送媒体は記憶媒体の間で情報を転送することに参加する。たとえば、伝送媒体は同軸ケーブル、銅線および光ファイバーを含む。伝送媒体は、音響波または光波、たとえば電波および赤外線データ通信の際に生成されるものの形を取ることもできる。

【0049】

本稿において具体的なコンピューティング・パラダイムおよびソフトウェア・ツールを参照したにもかかわらず、開示される技法はハードウェアおよびソフトウェアのいかなる特定の組み合わせにも、コンピューティング装置またはデータ処理装置によって実行される命令のいかなる特定の源にも限定されない。さまざまな実装のベースとなるプログラム命令は幅広い多様なプログラミング言語、ソフトウェア・ツールおよびデータ・フォーマットの任意のものに対応してもよく、任意の型の非一時的なコンピュータ可読記憶媒体もしくはメモリ・デバイス (単数または複数) に記憶されてもよく、多様なコンピューティング・モデルに従って実行されてもよい。コンピューティング・モデルは、たとえば、クライアント/サーバー・モデル、ピアツーピア・モデル、スタンドアローン・コンピューティング装置上またはさまざまな機能が異なる位置で実施されるまたは用いられる分散コンピューティング・モデルに従うことを含む。さらに、本稿における具体的なプロトコルへの言及は、単に例である。当業者に既知の好適な代替が用いられてもよい。

【0050】

本稿での用法での用語「スピーカー」は、単に例として、エンクロージャー内に設けられた直接放射型エレクトロダイナミック・ドライバを組み込むラウドスピーカー、ホーン・ラウドスピーカー、圧電型スピーカー、磁歪スピーカー、静電ラウドスピーカー、リボ

10

20

30

40

50

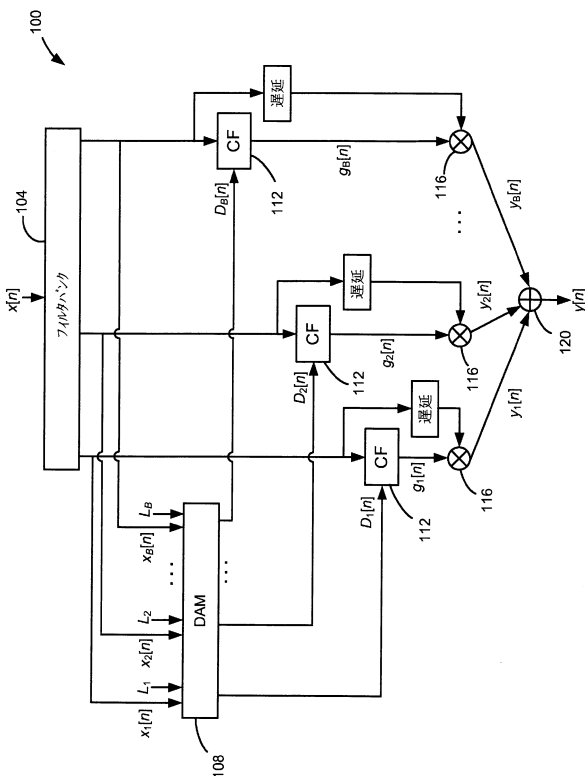
ン型および平面型ラウドスピーカー、屈曲波ラウドスピーカー、フラット・パネル・ラウドスピーカー、分散モード・ラウドスピーカー、ハイル（Heil）空気動トランスデューサ、プラズマ・アーク・スピーカー、デジタル・スピーカーおよびそれらの任意の組み合わせを含むことができる。

【 0 0 5 1 】

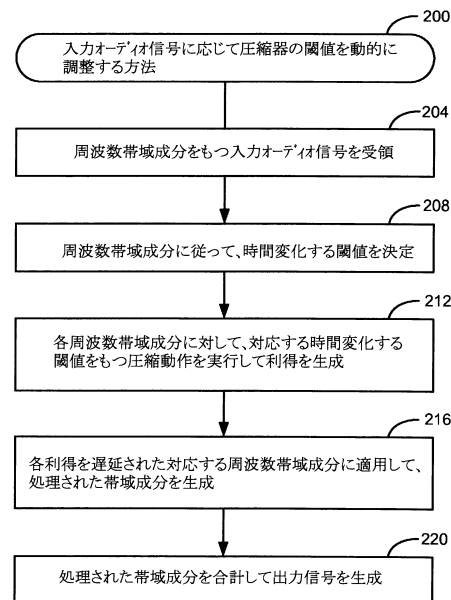
本願の主題はその個別的な実装を参照して具体的に示され、記述されてきたが、当業者は、本開示の精神および範囲から外れることなく開示される実装の形および詳細における変更ができることを理解するであろう。これらの実装のいくつかの例は付属の図面に示されており、その十全な理解を与えるために個別的な詳細が記載されている。こうした個別的な詳細の一部または全部なしでも実装が実施されうことは注意しておくべきである。さらに、よく知られた特徴は明確さを増進するために詳細には記述されなかったことがある。最後に、本稿でいくつかの実装を参照して利点が論じられているが、範囲はそのような利点への言及によって限定されるべきではないことは理解されるであろう。むしろ、範囲は、付属の請求項を参照して決定されるべきである。

10

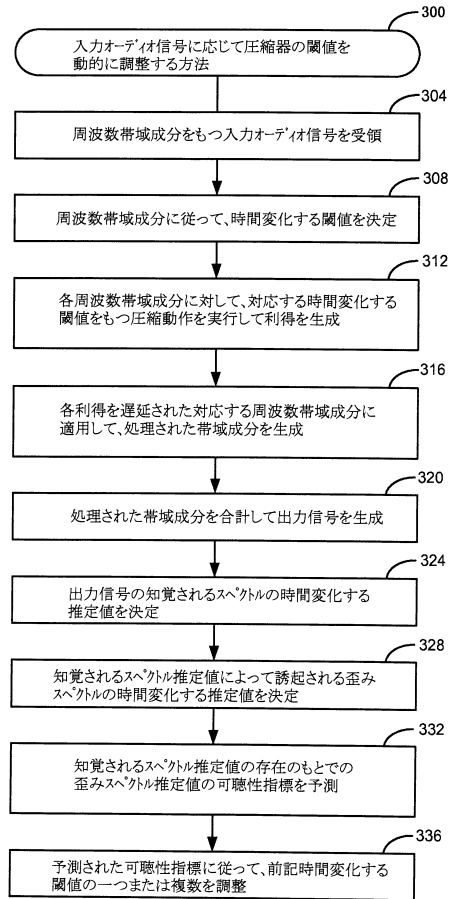
【 図 1 】



【 図 2 】



【図 3】



【図 5 A】

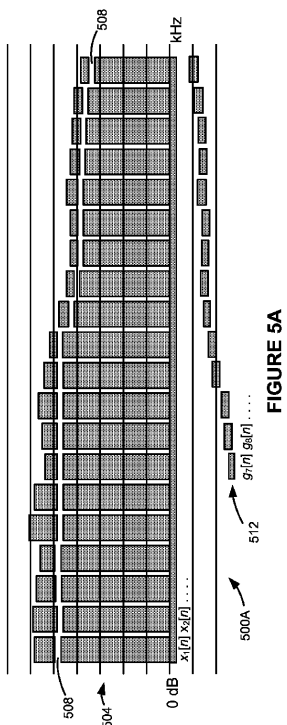
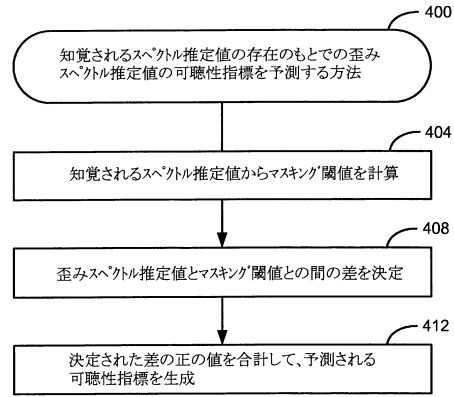


FIGURE 5A

【図 4】



【図 5 B】

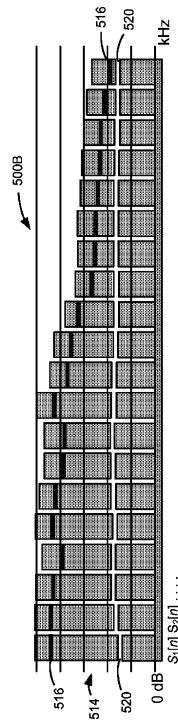
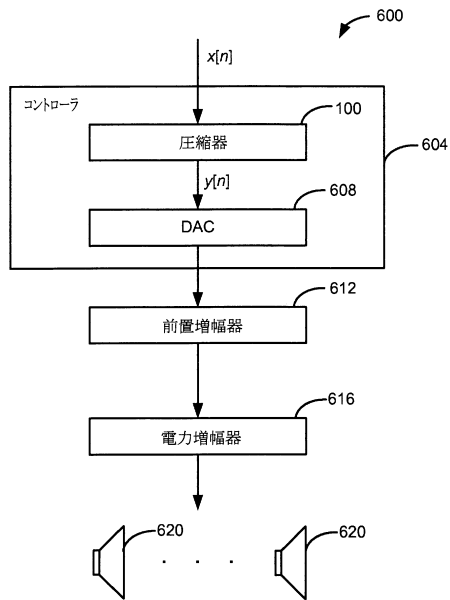


FIGURE 5B

【図 6】



フロントページの続き

(72)発明者 シーフェルドット, アラン ジェイ.
アメリカ合衆国, カリフォルニア州 94103, サンフランシスコ, ポットレロ アヴェニュー
100, ドルビー ラボラトリーズ, インコーポレイテッド内

審査官 高 橋 義昭

(56)参考文献 米国特許出願公開第2013/013096(US, A1)
米国特許出願公開第2010/215193(US, A1)
米国特許出願公開第2005/0123153(US, A1)
米国特許出願公開第2009/0254339(US, A1)
米国特許第5832444(US, A)
米国特許第4249042(US, A)
特開2005-160038(JP, A)
特表2010-521715(JP, A)
特表2003-520469(JP, A)

(58)調査した分野(Int.Cl., DB名)

H03G	9/02
G10L	21/034
H03G	7/00
H03G	9/00