

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第5585401号
(P5585401)

(45) 発行日 平成26年9月10日 (2014.9.10)

(24) 登録日 平成26年8月1日 (2014.8.1)

(51) Int.Cl. F 1
H04R 3/00 (2006.01) H04R 3/00 310

請求項の数 7 (全 31 頁)

(21) 出願番号	特願2010-250714 (P2010-250714)	(73) 特許権者	000002185
(22) 出願日	平成22年11月9日 (2010.11.9)		ソニー株式会社
(65) 公開番号	特開2012-104992 (P2012-104992A)		東京都港区港南1丁目7番1号
(43) 公開日	平成24年5月31日 (2012.5.31)	(74) 代理人	100095957
審査請求日	平成25年9月20日 (2013.9.20)		弁理士 亀谷 美明
		(74) 代理人	100096389
			弁理士 金本 哲男
		(74) 代理人	100101557
			弁理士 萩原 康司
		(74) 代理人	100128587
			弁理士 松本 一騎
		(72) 発明者	山下 功誠
			東京都港区港南1丁目7番1号 ソニー株式会社内

最終頁に続く

(54) 【発明の名称】 再生装置、再生方法、提供装置および再生システム

(57) 【特許請求の範囲】

【請求項1】

音響データを再生する再生部と、

前記再生部により再生された前記音響データの音量を、複数の音響データの各々の音量特性値と前記音響データの音量特性値との関係に基づいて補正する音量補正部と、

前記複数の音響データの各々の音量特性値と前記音響データの音量特性値との関係に基づいて音量補正のためのパラメータを算出するパラメータ算出部と、を備え、

前記音量特性値は、

前記音響データの音量平均値を示し、

前記パラメータ算出部は、

前記複数の音響データの前記音量平均値の分布に基づいて決定される基準平均値より前記音響データの前記音量平均値が小さい場合に前記音響データの音量を増幅するためのパラメータを算出し、前記基準平均値よりも前記音響データの前記音量平均値が大きい場合に前記音響データの音量を減衰するためのパラメータを算出し、

前記音量補正部は、

非線形補正部および第1の線形補正部を有し、前記パラメータ算出部により算出されたパラメータに従って前記音響データの音量を補正し、

前記パラメータ算出部は、

前記第1の線形補正部に設定される線形補正用パラメータ、および前記非線形補正部に設定される非線形補正用パラメータを算出する、

再生装置。

【請求項 2】

前記パラメータ算出部は、

前記基準平均値より前記音響データの前記音量平均値が小さい場合、前記音響データの音量ピーク値に対する音量上限値の比率を前記線形補正用パラメータとして算出し、前記音響データの前記音量平均値に対する前記基準平均値の比率に基づいて前記非線形補正用パラメータを算出する、

請求項 1 に記載の再生装置。

【請求項 3】

前記パラメータ算出部は、

前記基準平均値より前記音響データの前記音量平均値が大きい場合、前記音響データの前記音量平均値に対する前記基準平均値の比率を前記線形補正用パラメータとして算出する、

請求項 2 に記載の再生装置。

【請求項 4】

前記パラメータ算出部は、

前記基準平均値よりも小さい下側基準値より前記音響データの前記音量平均値が小さい場合に前記音響データの音量を増幅するためのパラメータを算出し、

前記基準平均値よりも大きい上側基準値より前記音響データの前記音量平均値が大きい場合に前記音響データの音量を減衰するためのパラメータを算出し、

前記下側基準値と前記上側基準値との間に前記音響データの前記音量平均値が含まれる場合、前記音響データをスルー出力させるためのパラメータを前記音量補正部に設定する、

請求項 3 に記載の再生装置。

【請求項 5】

前記音量補正部は、

前記第 1 の線形補正部の後段に前記非線形補正部と並列に接続される第 2 の線形補正部を有し、

前記パラメータ算出部は、

前記音響データをスルー出力させるためのパラメータとして、前記第 1 の線形補正部および前記第 2 の線形補正部に増幅率 1 を設定し、前記非線形補正用パラメータとして増幅率 0 を設定する、

請求項 4 に記載の再生装置。

【請求項 6】

音響データを再生するステップと；

再生された前記音響データの音量を、複数の音響データの各々の音量特性値と前記音響データの音量特性値の関係とに基づいて補正するステップと；

を含み、

前記複数の音響データの各々の音量特性値と前記音響データの音量特性値との関係に基づいて音量補正のためのパラメータを算出するステップを含み、

前記音量特性値は、前記音響データの音量平均値を示し、

前記パラメータの算出において、前記複数の音響データの前記音量平均値の分布に基づいて決定される基準平均値より前記音響データの前記音量平均値が小さい場合に前記音響データの音量を増幅するためのパラメータを算出し、前記基準平均値よりも前記音響データの前記音量平均値が大きい場合に前記音響データの音量を減衰するためのパラメータを算出し、

前記音響データの音量補正において、算出されたパラメータに従って前記音響データの音量を補正し、

前記パラメータの算出において、線形補正用パラメータ、および非線形補正用パラメータを算出する、

10

20

30

40

50

再生方法。

【請求項7】

音響データおよび複数の音響データの各々の音量特性値を記憶する記憶部、
前記音響データおよび前記複数の音響データの各々の音量特性値を取得する取得部、
前記音響データを解析するとともに前記複数の音響データの各々の音量特性値の分布を
解析する解析部、および

前記解析部による解析結果を送信する送信部、

を有する提供装置と；

前記提供装置から、前記音響データ、および前記解析結果を受信する受信部、

前記音響データを再生する再生部、

前記再生部により再生された前記音響データの音量を前記解析結果に基づいて補正する
音量補正部、および

前記複数の音響データの各々の音量特性値と前記音響データの音量特性値との関係に基
づいて音量補正のためのパラメータを算出するパラメータ算出部を有し、

前記音量特性値は、

前記音響データの音量平均値を示し、

前記パラメータ算出部は、

前記複数の音響データの前記音量平均値の分布に基づいて決定される基準平均値より前
記音響データの前記音量平均値が小さい場合に前記音響データの音量を増幅するためのパ
ラメータを算出し、前記基準平均値よりも前記音響データの前記音量平均値が大きい場合
に前記音響データの音量を減衰するためのパラメータを算出し、

前記音量補正部は、

非線形補正部および第1の線形補正部を有し、前記パラメータ算出部により算出された
パラメータに従って前記音響データの音量を補正し、

前記パラメータ算出部は、

前記第1の線形補正部に設定される線形補正用パラメータ、および前記非線形補正部に
設定される非線形補正用パラメータを算出する再生装置と、

を備える、再生システム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、再生装置、再生方法、提供装置および再生システムに関する。

【背景技術】

【0002】

近年、HDD (Hard Disk Drive) などの記憶装置の大容量化に伴い、
ユーザが携帯機器などを使用して楽曲を聴くなどといった一般的な環境などにおいても携
帯機器などにおいて容易に大量の楽曲を保持することができる環境が整ってきている。し
かし、そのような大量の楽曲を順に再生する場合、楽曲ごとの録音状態が異なっているな
どといった理由から、楽曲ごとの音量が所望の範囲内に収まらず、ある楽曲の音量は大き
過ぎる一方、別の楽曲の音量は小さ過ぎるなどといった状況が起こり得る。

【0003】

さらに、最近では、ADSL (Asymmetric Digital Subscriber Line) や光ファイバーなどの高速ネットワークも普及してきており、一般
家庭に存在するPC (Personal Computer) などにおいても、ネットワ
ーク上の楽曲サーバに常時接続して、連続的に楽曲をストリーミング再生できる環境が整
ってきている。このような環境においても同様に、ある楽曲の音量は大き過ぎる一方、別
の楽曲の音量は小さ過ぎるなどといった状況が起こり得る。

【0004】

かかる状況においては、ユーザは楽曲を心地よく聴くのが困難であり、楽曲の音量を調

10

20

30

40

50

整する操作はユーザにとって煩わしいものである。このような状況を改善するために、ノーマライザーという自動音量補正装置が提案されている。自動音量補正装置は、各楽曲の再生時の音量を自動的に補正するものである（例えば、特許文献1参照。）。

【先行技術文献】

【特許文献】

【0005】

【特許文献1】特開2001-320793号公報

【発明の概要】

【発明が解決しようとする課題】

【0006】

しかしながら、上記したノーマライザーは、一般的には楽曲の音量が所定の目標レベルに収まるように動的に楽曲の音量を制御するため、再生中の楽曲に関して原曲の音量カーブを忠実に再現できないという問題があった。例えば、楽曲の先頭付近は音量が小さいが、再生の途中から急激に音量が上がる楽曲であっても、常に所定の音量レベルで再生されてしまう。このような楽曲の再生では、楽曲の作曲者や録音者が意図した音量カーブを得ることはできないのは明らかである。楽曲以外の音響データに関しても同様の問題があった。

【0007】

また、このような楽曲の再生では、再生中の楽曲の音量が必要以上に増幅されることがあり、増幅による音声の歪みが発生し易く、楽曲の音質を損ねてしまうという問題が起こっている。楽曲以外の音響データに関しても同様の問題が起こっている。

【0008】

そこで、本発明は、上記問題に鑑みてなされたものであり、本発明の目的とするところは、元の音響データの音量変化に比較的近い音量変化を維持しつつ各音響データ間の音量の違いを低減させ、音量増幅による音声の歪みを低減させながら音響データを再生することが可能な、新規かつ改良された技術を提供することにある。

【課題を解決するための手段】

【0009】

上記課題を解決するために、本発明のある観点によれば、音響データを再生する再生部と、上記再生部により再生された上記音響データの音量を、複数の音響データの各々の音量特性値と上記音響データの音量特性値との関係に基づいて補正する音量補正部と、を備える、再生装置が提供される。

【0010】

上記再生装置は、上記複数の音響データの各々の音量特性値と上記音響データの音量特性値との関係に基づいて音量補正のためのパラメータを算出するパラメータ算出部をさらに備え、上記音量補正部は、上記パラメータ算出部により算出されたパラメータに従って上記音響データの音量を補正することとしてもよい。

【0011】

上記音量特性値は、上記音響データの音量平均値を示し、上記パラメータ算出部は、上記複数の音響データの上記音量平均値の分布に基づいて決定される基準平均値より上記音響データの上記音量平均値が小さい場合に上記音響データの音量を増幅するためのパラメータを算出し、上記基準平均値よりも上記音響データの上記音量平均値が大きい場合に上記音響データの音量を減衰するためのパラメータを算出することとしてもよい。

【0012】

上記音量補正部は、非線形補正部および第1の線形補正部を有し、上記パラメータ算出部は、上記第1の線形補正部に設定される線形補正用パラメータ、および上記非線形補正部に設定される非線形補正用パラメータを算出することとしてもよい。

【0013】

上記パラメータ算出部は、上記基準平均値より上記音響データの上記音量平均値が小さい場合、上記音響データの音量ピーク値に対する音量上限値の比率を上記線形補正用パラ

10

20

30

40

50

メータとして算出し、上記音響データの上記音量平均値に対する上記基準平均値の比率に基づいて上記非線形補正用パラメータを算出することとしてもよい。

【0014】

上記パラメータ算出部は、上記基準平均値より上記音響データの上記音量平均値が大きい場合、上記音響データの上記音量平均値に対する上記基準平均値の比率を上記線形補正用パラメータとして算出することとしてもよい。

【0015】

上記パラメータ算出部は、上記基準平均値よりも小さい下側基準値より上記音響データの上記音量平均値が小さい場合に上記音響データの音量を増幅するためのパラメータを算出し、上記基準平均値よりも大きい上側基準値より上記音響データの上記音量平均値が大きい場合に上記音響データの音量を減衰するためのパラメータを算出し、上記下側基準値と上記上側基準値との間に上記音響データの上記音量平均値が含まれる場合、上記音響データをスルー出力させるためのパラメータを上記音量補正部に設定することとしてもよい。

【0016】

上記音量補正部は、上記第1の線形補正部の後段に上記非線形補正部と並列に接続される第2の線形補正部を有し、上記パラメータ算出部は、上記音響データをスルー出力させるためのパラメータとして、上記第1の線形補正部および上記第2の線形補正部に増幅率1を設定し、上記非線形補正用パラメータとして増幅率0を設定することとしてもよい。

【発明の効果】

【0017】

以上説明したように、本発明によれば、元の音響データの音量変化に比較的近い音量変化を維持しつつ各音響データ間の音量の違いを低減させ、音量増幅による音声の歪みを低減させながら音響データを再生することが可能である。

【図面の簡単な説明】

【0018】

【図1】音響データを再生して得られる音声の波形の一例を示す図である。

【図2】一般的な音量補正技術による音量補正を音響データに適用した場合において音響データを再生して得られる音声の波形の一例を示す図である。

【図3】本発明の実施形態に係る再生装置による音量補正を音響データに適用した場合において音響データを再生して得られる音声の波形の一例を示す図である。

【図4】本発明の実施形態に係る音量特性値の一例としてのPeak値およびRMS値について説明するための図である。

【図5】本発明の実施形態に係る音量特性値の一例としてのPeak値の計算例を示す図である。

【図6】本発明の実施形態に係る音量特性値の一例としてのRMS値の計算例を示す図である。

【図7】本発明の実施形態に係る再生装置のハードウェア構成を示す図である。

【図8】本発明の実施形態に係る再生システムの機能構成を示す図である。

【図9】本発明の実施形態に係る属性情報の管理手法の一例について説明するための図である。

【図10】本発明の実施形態に係る属性情報の管理手法の他の一例について説明するための図である。

【図11】本発明の実施形態に係る再生システムにより実行される処理の流れを示すフローチャートである。

【図12】本発明の実施形態に係る再生装置による音響データの音量補正に適用し得るコンプレッサー回路を示す図である。

【図13】本発明の実施形態に係る再生装置に適用し得るコンプレッサー回路が有する静特性（音響データの音量を増幅する場合）の一例を示す図である。

【図14】本発明の実施形態に係る再生装置に適用し得るコンプレッサー回路が有する静

10

20

30

40

50

特性（音響データの音量を減衰する場合）の一例を示す図である。

【図15】本発明の実施形態に係る再生装置に適用し得るコンプレッサー回路が有する静特性（音響データの音量を増幅する場合）の他の一例を示す図である。

【図16】本発明の実施形態に係る再生装置に適用し得るコンプレッサー回路が有する動特性の一例を示す図である。

【図17】本発明の実施形態に係る再生装置に適用し得るコンプレッサー回路における制御信号（エンベロープ信号）の生成手法の一例を示す図である。

【図18】本発明の実施形態に係る再生装置による音響データの音量補正の処理の流れの詳細を示すフローチャートである。

【図19】本発明の実施形態に係る再生装置に適用し得るパラメータ算出部によるパラメータの算出例（音響データの音量を増幅する場合）を示す図である。 10

【図20】本発明の実施形態に係る再生装置に適用し得るパラメータ算出部によるパラメータの算出例（音響データの音量を減衰する場合）を示す図である。

【図21】本発明の実施形態に係る再生装置に適用し得るパラメータ算出部によるパラメータの算出例（音響データをスルー出力する場合）を示す図である。

【図22】本発明の実施形態に係る音量特性値の一例としてのPeak値の統計分布について説明するための図である。

【図23】本発明の実施形態に係る音量特性値の一例としてのRMS値の統計分布について説明するための図である。

【図24】本発明の実施形態に係る再生装置による音響データの音量補正の処理に使用され得るRMS基準値の決定手法の一例について説明するための図である。 20

【図25】本発明の実施形態に係る再生装置による音響データの音量補正の処理に使用され得るRMS基準値の決定手法の他の一例について説明するための図である。

【発明を実施するための形態】

【0019】

以下に添付図面を参照しながら、本発明の好適な実施の形態について詳細に説明する。なお、本明細書及び図面において、実質的に同一の機能構成を有する構成要素については、同一の符号を付すことにより重複説明を省略する。

【0020】

また、以下の順序にしたがって当該「発明を実施するための形態」を説明する。 30

1. 実施形態

- 1-1. 音響データを再生して得られる音声の波形
- 1-2. 一般的な音量補正技術による音量補正
- 1-3. 本発明の実施形態に係る再生装置による音量補正
- 1-4. 音量特性値の一例としてのPeak値およびRMS値
- 1-5. 音量特性値の一例としてのPeak値の計算例
- 1-6. 音量特性値の一例としてのRMS値の計算例
- 1-7. 再生装置のハードウェア構成
- 1-8. 再生システムの機能構成
- 1-9. 属性情報の管理手法の一例 40
- 1-10. 属性情報の管理手法の他の一例
- 1-11. 再生システムにより実行される処理
- 1-12. 音響データの音量補正に適用し得るコンプレッサー回路
- 1-13. コンプレッサー回路が有する静特性（音量増幅時）の一例
- 1-14. コンプレッサー回路が有する静特性（音量減衰時）の一例
- 1-15. コンプレッサー回路が有する静特性（音量増幅時）の他の一例
- 1-16. コンプレッサー回路が有する動特性の一例
- 1-17. コンプレッサー回路における制御信号の生成手法の一例
- 1-18. 音響データの音量補正の処理の流れの詳細
- 1-19. パラメータ算出部によるパラメータ（音量増幅時）の算出例 50

- 1 - 2 0 . パラメータ算出部によるパラメータ (音量減衰時) の算出例
- 1 - 2 1 . パラメータ算出部によるパラメータ (スルー出力時) の算出例
- 1 - 2 2 . 音量特性値の一例としての P e a k 値の統計分布
- 1 - 2 3 . 音量特性値の一例としての R M S 値の統計分布
- 1 - 2 4 . R M S 基準値の決定手法の一例
- 1 - 2 5 . R M S 基準値の決定手法の他の一例

- 2 . 使用例
- 3 . 変形例
- 4 . まとめ

【 0 0 2 1 】

10

< 1 . 実施形態 >

[1 - 1 . 音響データを再生して得られる音声の波形]

図 1 は、音響データを再生して得られる音声の波形の一例を示す図である。図 1 を参照しながら、音響データを再生して得られる音声の波形の一例について説明する。なお、図 1 には、音声波形の P e a k 値および R M S (R o o t M e a n S q u a r e) 値のラインが示されているが、P e a k 値および R M S 値は、本発明の実施形態において、音量特性の一例として用いられ得るものである。また、R M S 値は、音量平均値の一例として使用されるものである。

【 0 0 2 2 】

20

ある音響データを再生して得られる音声の波形は、図 1 に示すように示される。ここで、図 1 に示された波形のうちの上側に示した波形は、音響データを再生して得られる音声を左 (L c h) および右 (R c h) のスピーカの双方から出力する方式 (以下、「ステレオ音声方式」とも言う。) により L c h から出力された音声の波形を示したものである。また、図 1 に示された波形のうちの下側に示した波形は、ステレオ音声方式により R c h から出力された音声の波形を示したものである。音響データに対する補正は行われていない。

【 0 0 2 3 】

30

図 1 に示された音声波形は、C D - D A 形式 (C o m p a c t D i s c D i g i t a l A u d i o) により記録された音響データと同様の記録方式により記録された音響データを再生して得られたものである。C D - D A 形式は、サンプリング周波数 4 4 . 1 k H z 、量子化ビット数 1 6 b i t 、ステレオ P C M (P u l s e C o d e M o d u l a t i o n) 形式により音響データを記録するものである。

【 0 0 2 4 】

音響データが量子化ビット数 1 6 b i t 、P C M 形式により記録されたものである場合、その音響データは、符号 1 b i t およびデータ 1 5 b i t により構成された 2 の補数コードにより表現される。この音響データにより表現できる音声の音量 (振幅値) の最大値 (あるいは最小値) は、正側で + 3 2 7 6 7 、負側で - 3 2 7 6 8 である。最小値から最大値までの範囲を超える音量の音声については最大値 (あるいは最小値) に丸められることとなるが、その結果としてクリップノイズと呼ばれる不快な歪みが発生する可能性がある。

40

【 0 0 2 5 】

他方、音量が小さ過ぎる音声音響データとして記録されている場合には、1 6 b i t の情報量を使い切れずに、音響データが再生された結果として良好な S / N の音声を得られないこととなる。それゆえ、音響データの一例として楽曲を使用する場合には、楽曲の制作者は 1 曲を通して音量が 1 6 b i t により表現される範囲内に収まるようにミキサーなどの機器を用いて音量を調整しながら楽曲を録音する。

【 0 0 2 6 】

しかしながら、図 1 に示した音声波形を有する楽曲のように、市販されている C D もしくは楽曲ファイルの中には、全体を通して 1 6 b i t により表現できる値の最大値に音量が達しない楽曲が録音されていることもある。そのような楽曲が録音されている理由とし

50

ては様々なものが想定されるが、録音技術が発達していなかった古い年代に録音された楽曲には、一般的にそのような傾向がある。

【 0 0 2 7 】

図 1 に示した例は、他の楽曲と比べて比較的音量が小さいクラシック曲の音声波形である。このようなクラシック曲を単独で再生する場合には音量はさほど気にならないが、音量が比較的大きい楽曲の前後にこのようなクラシック曲を再生する場合は、このようなクラシック曲の音量が相対的に小さくなるため、音量の小ささが気になる可能性が高い。それゆえ、このクラシック曲の音量のみを上げたくならないという状況となり得る。

【 0 0 2 8 】

図 1 に示した音声波形を音声波形部分 M 1 1 ~ M 1 4 に分けた場合、音声波形部分 M 1 1 は静かな部分であり、音声波形部分 M 1 2 は徐々に盛り上がる部分であり、音声波形部分 M 1 3 は静かな部分であり、音声波形部分 M 1 4 は盛り上がる部分である。上記したように、図 1 に示した音声波形はクラシック曲を再生して得られるものであり、聴く人に対して全体的に音量が小さいという印象を与え易いと想定される。そこで、このクラシック曲を再生する場合には、音声の音量を全体的に上げる必要があると考えられる。

【 0 0 2 9 】

[1 - 2 . 一般的な音量補正技術による音量補正]

図 2 は、一般的な音量補正技術による音量補正を音響データに適用した場合において音響データを再生して得られる音声の波形の一例を示す図である。図 2 を参照しながら、一般的な音量補正技術による音量補正を音響データに適用した場合において音響データを再生して得られる音声の波形の一例について説明する。

【 0 0 3 0 】

図 2 に示した音声波形を音声波形部分 M 2 1、M 2 2 に分けた場合、音声波形部分 M 2 1 は一気に盛り上がる部分であり、音声波形部分 M 2 2 は盛り上がりっぱなしによりメリハリがない部分である。図 2 に示すように、一般的な音量補正技術による音量補正により、上記したクラシック曲の音量は大きくなっているが、原曲が有する音量変化（以下、「音量カーブとも言う。」）を維持することができておらず、聴く人に不自然な感じを与える結果となっている。

【 0 0 3 1 】

また、一般的な音量補正技術による音量補正により、図 2 に示された音声波形からは把握できないような音声の歪みも発生している。聴く人に不自然な感じを与えることとなった理由の 1 つは、音量増幅が音声波形全体を考慮してなされたものではなく、音声波形の一部を考慮して部分的に適切な音量増幅を行っていることが挙げられる。

【 0 0 3 2 】

[1 - 3 . 本発明の実施形態に係る再生装置による音量補正]

図 3 は、本発明の実施形態に係る再生装置による音量補正を音響データに適用した場合において音響データを再生して得られる音声の波形の一例を示す図である。図 3 を参照しながら、本発明の実施形態に係る再生装置による音量補正を音響データに適用した場合において音響データを再生して得られる音声の波形の一例について説明する。

【 0 0 3 3 】

図 3 に示した音声波形を音声波形部分 M 3 1、M 3 2、M 3 3、M 3 4 に分けた場合、音声波形部分 M 3 1 は静かな部分であり、音声波形部分 M 3 2 は徐々に盛り上がる部分であり、音声波形部分 M 3 3 は静かな部分であり、音声波形部分 M 3 4 は徐々に盛り上がる部分である。図 3 に示すように、本発明の実施形態に係る再生装置による音量補正により、上記したクラシック曲の音量は大きくなっており、原曲が有する音量カーブを維持することができている。

【 0 0 3 4 】

また、本発明の実施形態に係る再生装置による音量補正により、音声の歪みも軽減することができる。聴く人に自然な感じを与えるような音量補正を行うことが可能となった理由の 1 つは、音量増幅が音声波形全体を考慮してなされたものであることが挙げられる。

音声波形全体を考慮して音量増幅を行うために、例えば、上記した Peak 値および RMS 値などといったメタデータを使用することができる。Peak 値および RMS 値などのメタデータは、例えば、サーバに保持させておき、再生装置にダウンロードして使用することができる。

【0035】

[1 - 4 . 音量特性値の一例としての Peak 値および RMS 値]

図4は、本発明の実施形態に係る音量特性値の一例としての Peak 値および RMS 値について説明するための図である。図4を参照しながら、本発明の実施形態に係る音量特性値の一例としての Peak 値および RMS 値について説明する。

【0036】

図4に示すように、1箇所だけ突発的な衝撃音のような大きな音があり、それ以外は歌声のみが含まれている楽曲が録音されていたとする。このような楽曲は、一般的にはダイナミックレンジの広い音と呼ばれる。衝撃音の音量が最大音量に相当し、16bitにより表現できる値の最大値に相当するため、衝撃音以外の音声については相対的に音量を下げた録音せざるを得ない。

【0037】

例えば、衝撃音の Peak 値が +32767 であった場合は、衝撃音以外の音声に相当する歌声などの音量は Peak 値を +1000 くらいまで下げて録音しないと、楽曲全体が 16bit に収まらない。このようなケースでは、衝撃音以外の音声に相当する歌声の音量が、相対的にかなり小さく感じられてしまう。

【0038】

このような背景から、本発明の実施形態に係る再生システムは、例えば、楽曲の音量を表す特徴量として Peak 値および RMS 値の2つを保持することとしている。Peak 値は、楽曲1曲の中に出現する突発的な音量変化が生じたときの音量レベルを表す特徴量として用いるのに適していると言える。図4に示した例では、Peak 値は、衝撃音の音量レベルを表現するのに適していると言える。また、RMS 値は、楽曲の音声の有するパワーに対する積分値の時間平均を表すものであり、楽曲全体の平均的な音量レベルを表す特徴量として用いるのに適していると言える。図4に示した例においては、RMS 値は、主に衝撃音以外の音声に相当する歌声の音量レベルを表す特徴量となっている。

【0039】

[1 - 5 . 音量特性値の一例としての Peak 値の計算例]

図5は、本発明の実施形態に係る音量特性値の一例としての Peak 値の計算例を示す図である。図5を参照しながら、本発明の実施形態に係る音量特性値の一例としての Peak 値の計算例について説明する。

【0040】

音響データの一例としての楽曲1曲をモノラル音声により再生した場合、例えば、図5に示した式1のような関数により計算を行うことにより、変数 Peak に Peak 値が格納される。また、音響データの一例としての楽曲1曲をステレオ音声により再生した場合、例えば、図5に示した式2のような関数により計算を行うことにより、変数 Peak に Peak 値が格納される。本発明の実施形態に係る再生システムは、例えば、このような計算により変数 Peak に Peak 値を格納し、変数 Peak に格納した Peak 値を音量特性値の一例として使用することができる。

【0041】

[1 - 6 . 音量特性値の一例としての RMS 値の計算例]

図6は、本発明の実施形態に係る音量特性値の一例としての RMS 値の計算例を示す図である。図6を参照しながら、本発明の実施形態に係る音量特性値の一例としての RMS 値の計算例について説明する。

【0042】

音響データの一例としての楽曲1曲をモノラル音声により再生した場合、例えば、図6に示した式3のような関数により計算を行うことにより、変数 RMS に RMS 値が格納さ

10

20

30

40

50

れる。また、音響データの一例としての楽曲1曲をステレオ音声により再生した場合、例えば、図6に示した式5のような関数により計算を行うことにより、変数RMSにRMS値が格納される。本発明の実施形態に係る再生システムは、例えば、このような計算により変数RMSにRMS値を格納し、変数RMSに格納したRMS値を音量特性値の一例として使用することができる。

【0043】

[1-7.再生装置のハードウェア構成]

図7は、本発明の実施形態に係る再生装置のハードウェア構成例を示す図である。図7を参照しながら、本発明の実施形態に係る再生装置のハードウェア構成例について説明する。

10

【0044】

図7に示すように、本発明の実施形態に係る再生装置100は、例えば、CPU(Central Processing Unit)910と、ROM(Read Only Memory)920と、RAM(Random Access Memory)930と、バス940と、ユーザインタフェース950と、記憶装置960と、通信装置970と、デコード回路980と、を備えるものである。

【0045】

CPU910は、演算処理装置および制御装置として機能し、例えば、ROM920に記録された各種プログラムを読み出して記憶装置960に記憶させておき、各種プログラムを実行するに際して、記憶装置960により記憶された各種プログラムをRAM930に展開し、RAM930に展開した各種プログラムを実行することができる。かかる各種プログラムの実行により、再生装置100内の動作全般またはその一部を制御することが可能である。

20

【0046】

ROM920は、読み出し専用のメモリであり、例えば、CPU910により読み出される各種プログラムやその実行に使用されるパラメータなどを記憶することができる。

【0047】

RAM930は、記憶装置960からCPU910により読み出された各種プログラムやその実行に使用されるパラメータなどを一時的に記憶することが可能である。また、CPU910による各種プログラムの実行に際して一時的に記憶しておく必要が生じた各種データなどを記憶することも可能である。

30

【0048】

バス940は、再生装置100の内部に設けられている各ハードウェアを接続し、各ハードウェア間における信号の送受信を実現するために機能する。

【0049】

ユーザインタフェース950は、入力装置により構成されるものである。例えば、マウス、キーボード、タッチパネル、ボタン、スイッチおよびレバーなどにより構成され、ユーザからの操作を受け付ける機能を有する。入力装置は、例えば、受け付けた操作に基づいて入力信号を生成し、CPU910に出力する。音響データを再生して得られる音声を聴いている人は、例えば、入力装置を操作することにより、再生装置100に対して各種のデータを入力したり処理動作を指示したりすることができる。

40

【0050】

記憶装置960は、データ格納用の装置であり、例えば、HDDなどの磁気記憶デバイス、半導体記憶デバイス、光記憶デバイス、または光磁気記憶デバイス等により構成される。記憶装置960は、外部から取得した音響データや画像データなどを格納することができる。記憶装置960は、例えば、CPU910によりROM920から読み出された各種プログラムを記憶しておく機能などを有するものである。

【0051】

50

通信装置 970 は、例えば、ネットワークに接続するための通信デバイス等で構成された通信インタフェースである。通信装置 970 は、例えば、有線または無線 LAN (Local Area Network) 用の通信カード、ADSL (Asymmetric Digital Subscriber Line) 用のルータ、各種通信用のモデムなどにより構成される。この通信装置 970 は、例えば、他の通信機器との間で音響データなどのデータを送受信することができる。また、通信装置 970 に接続されるネットワークは、有線または無線によって接続されたネットワークなどにより構成され、例えば、インターネット、家庭内 LAN などであってもよい。

【0052】

デコード回路 980 は、符号化された音響データを復号し、再生する機能を有するものである。また、デコード回路 980 は、符号化された映像データを復号し、再生する機能を有していてもよい。

10

【0053】

以上、本発明の実施形態に係る再生装置 100 の機能を実現可能なハードウェア構成の一例を示した。上記の各構成要素は、汎用的な部材を用いて構成されていてもよいし、各構成要素の機能に特化したハードウェアにより構成されていてもよい。従って、本発明の実施形態に係る再生装置 100 を実施する時の技術レベルに応じて、適宜、利用するハードウェア構成を変更することが可能である。

【0054】

[1- 8 . 再生システムの機能構成]

20

図 8 は、本発明の実施形態に係る再生システムの機能構成を示す図である。図 8 を参照しながら、本発明の実施形態に係る再生システムの機能構成について説明する。

【0055】

図 8 に示すように、本発明の実施形態に係る再生システム 10 は、再生装置 100 と、提供装置 200 と、出力装置 300 とを備えるものである。再生装置 100 と提供装置 200 とは、例えば、ネットワークを介して接続されており、互いに通信可能とされている。再生装置 100 と出力装置 300 とは、例えば、専用線により接続されており、出力装置 300 は、再生装置 100 から出力された音声信号を音声として出力することが可能である。また、出力装置 300 は、再生装置 100 から出力された画像信号を画像として出力することが可能である。

30

【0056】

再生装置 100 は、少なくとも、再生部の一例としての映像音響データデコード回路 130 (以下、単に「映像音響データデコード回路 130」とも言う。)と、音量補正部の一例としてのコンプレッサー回路 160 (以下、単に「コンプレッサー回路 160」とも言う。)を備えるものである。ここでは、再生装置 100 は、映像音響データデコード回路 130 を備えることとしており、映像音響データデコード回路 130 は、符号化された映像データと符号化された音響データとの双方を復号して再生することとする。しかしながら、再生装置 100 は、再生部は、符号化された映像データを復号して再生する機能は有していなくてもよい。映像音響データデコード回路 130 は、例えば、上記したデコード回路 980 により構成される。

40

【0057】

コンプレッサー回路 160 は、映像音響データデコード回路 130 により再生された音響データ (以下、「再生対象音響データ」とも言う。)の音量を、複数の音響データの各々の音量特性値と再生対象音響データの音量特性値との関係に基づいて補正する機能を有するものである。複数の音響データは、特に限定されるものではなく、再生対象音響データを含むものであってもよいし、含まないものであってもよい。音量特性値としては、Peak 値や RMS 値などを例として挙げることができる。

【0058】

再生装置 100 は、映像音響データデコード回路 130 と、コンプレッサー回路 160 とを備えることにより、元の音響データの音量変化に比較的近い音量変化を維持しつつ各

50

音響データ間の音量の違いを低減させ、音量増幅による音声の歪みを低減させながら音響データを再生することが可能となる。

【0059】

再生装置100は、パラメータ算出部150を備えることとしてもよい。パラメータ算出部150は、複数の音響データの各々の音量特性値と音響データの音量特性値との関係に基づいて音量補正のためのパラメータを算出する機能を有するものである。音量補正のためのパラメータは、コンプレッサー回路160において音響データの音量補正に使用される。すなわち、コンプレッサー回路160は、パラメータ算出部150により算出されたパラメータに従って再生対象音響データの音量を補正する。

【0060】

再生装置100は、受信部110、映像音響データ取得部120、映像音響データデコード回路130、属性情報取得部140、ミキサー装置170、記憶部180などを備えることとしてもよい。受信部110は、提供装置200から、上記した複数の音響データの各々の音量特性値、再生対象音響データの音量特性値、再生対象音響データ自体などを受信することが可能である。しかしながら、上記した複数の音響データの各々の音量特性値、再生対象音響データの音量特性値、再生対象音響データ自体などは、提供装置200から受信せずに、再生装置100が備える記憶部180により記憶されているものを再生装置100において使用することもできる。受信部110は、提供装置200から映像データを受信することとしてもよい。受信部110は、例えば、通信装置970により構成されるものである。

【0061】

映像音響データ取得部120は、受信部110により受信された再生対象音響データを受信部110から取得する機能を有するものである。映像音響データ取得部120は、取得した再生対象音響データを映像音響データデコード回路130に出力する機能をさらに有している。映像音響データ取得部120は、受信部110により映像データが受信された場合には、その映像データを受信部110から取得する機能を有していてもよい。その場合には、映像音響データ取得部120は、受信部110から取得した映像データを映像音響データデコード回路130に出力する機能をさらに有する。映像音響データデコード回路130に出力された再生対象音響データや映像データは、映像音響データデコード回路130において使用される。

【0062】

属性情報取得部140は、受信部110により受信された再生対象音響データの音量特性値を受信部110から取得する機能を有するものである。属性情報取得部140は、受信部110により受信された再生対象音響データの音量特性値をパラメータ算出部150に出力する機能をさらに有している。パラメータ算出部150に出力された再生対象音響データの音量特性値は、パラメータ算出部150において使用される。パラメータ算出部150は、再生対象音響データの音量特性値を再生対象音響データの属性情報(「メタデータ」とも言う。以下同じ。)として属性情報取得部140から受信することができる。

【0063】

ミキサー装置170は、コンプレッサー回路160から出力された再生対象音響データおよび映像データを合成する機能を有するものである。コンプレッサー回路160から出力された再生対象音響データは、必要に応じて音量が補正されている。ミキサー装置170は、再生対象音響データおよび映像データが合成されてなる合成信号を出力装置300に出力する。

【0064】

再生装置100は、記憶部180をさらに備えることとしてもよい。記憶部180は、例えば、記憶装置960によって構成されるものである。記憶部180は、再生装置100を構成する各機能ブロックの動作を制御するための各種プログラムを記憶する機能や、そのプログラムが実行される際に使用される各種データなどを記憶する機能を有するもの

10

20

30

40

50

である。

【 0 0 6 5 】

映像音響データ取得部 1 2 0 や、属性情報取得部 1 4 0、パラメータ算出部 1 5 0 など
は、例えば、CPU 9 1 0、RAM 9 3 0 などにより構成され、CPU 9 1 0 が記憶部 1
8 0 によって記憶されているプログラムを RAM 9 3 0 に展開して実行することによりそ
の機能が実現されるものである。しかし、このような構成に限らず、映像音響データ取得
部 1 2 0 や、属性情報取得部 1 4 0、パラメータ算出部 1 5 0 の中には、専用のハードウ
ェアにより構成されるものが存在してもよい。

【 0 0 6 6 】

提供装置 2 0 0 は、少なくとも、記憶部 2 1 0 と、取得部 2 2 0 と、解析部 2 3 0 と、
送信部 2 4 0 とを備えるものである。記憶部 2 1 0 は、例えば、記憶装置によって構成さ
れるものである。記憶部 1 8 0 は、提供装置 2 0 0 を構成する各機能ブロックの動作を制
御するための各種プログラムを記憶する機能や、そのプログラムが実行される際に使用さ
れる各種データなどを記憶する機能を有するものである。記憶部 1 8 0 は、例えば、上記
した複数の音響データの各々の音量特性値を記憶する機能を有していてもよい。また、記
憶部 1 8 0 は、再生対象音響データを記憶する機能を有していてもよい。

10

【 0 0 6 7 】

取得部 2 2 0 は、複数の音響データの各々の音量特性値を取得する機能を有するもの
である。取得部 2 2 0 は、例えば、記憶部 1 8 0 により記憶されている複数の音響データ
の各々の音量特性値を取得する。複数の音響データの各々の音量特性値としては、例え
ば、複数の音響データの各々の RMS 値を使用することができる。また、取得部 2 2 0 は、再
生対象音響データを取得する機能を有するものである。取得部 2 2 0 は、例えば、記憶部
1 8 0 により記憶されている再生対象音響データを取得する。

20

【 0 0 6 8 】

解析部 2 3 0 は、取得部 2 2 0 により取得された複数の音響データの各々の音量特性値
の分布を解析する機能を有するものである。複数の音響データの各々の音量特性値の分布
としては、例えば、複数の音響データの各々の RMS 値の分布を使用することができる。
解析部 2 3 0 は、複数の音響データの各々の音量特性値の分布を解析することにより、R
M S 基準値を解析結果として算出してもよい。RMS 基準値の算出手法については、後に
詳細に説明する。解析部 2 3 0 により複数の音響データの各々の音量特性値の分布が解析
されて得られた結果は、解析結果の一例として送信部 2 4 0 に出力される。RMS 基準値
は、基準平均値の一例である。

30

【 0 0 6 9 】

また、解析部 2 3 0 は、取得部 2 2 0 により取得された再生対象音響データを解析する
機能をさらに有するものである。解析部 2 3 0 は、再生対象音響データを解析すること
により、再生対象音響データの音量特性値を算出してもよい。また、解析部 2 3 0 は、再生
対象音響データの音量特性値の一例として、再生対象音響データの RMS 値を算出しても
よい。解析部 2 3 0 は、このように解析して得られた再生対象音響データの音量特性値を
、例えば、再生対象音響データの属性情報として記憶部 2 1 0 に記憶させておき、必要に
応じて記憶部 2 1 0 から取得できるようにしておくこともできる。解析部 2 3 0 により再
生対象音響データが解析されて得られた結果は、解析結果の一例として送信部 2 4 0 に出
力される。

40

【 0 0 7 0 】

ここでは、解析部 2 3 0 は、複数の音響データの各々の音量特性値の分布を解析して得
た結果と、再生対象音響データを解析して得た結果とを、解析結果として送信部 2 4 0 に
出力することとしている。しかしながら、解析部 2 3 0 は、これらの解析結果を用いてさ
らに解析を進め、その結果として得られた解析結果を送信部 2 4 0 に出力するようにして
もよい。例えば、解析部 2 3 0 は、解析をさらに進めた結果として、音量補正のためのパ
ラメータを算出することとしてもよい。かかる場合には、当該パラメータを解析結果の一
例として送信部 2 4 0 に出力することとしてもよい。音量補正のためのパラメータについ

50

ては、後に詳細に説明する。

【 0 0 7 1 】

送信部 2 4 0 は、再生対象音響データを再生する再生装置 1 0 0 に対して、再生対象音響データ、および解析部 2 3 0 による解析結果を送信する機能を有するものである。送信部 2 4 0 は、例えば、取得部 2 2 0 により記憶部 2 1 0 から取得された再生対象音響データを再生装置 1 0 0 に送信することができる。送信部 2 4 0 は、例えば、通信装置により構成されるものである。送信部 2 4 0 は、記憶部 2 1 0 により映像データが記憶されている場合には、記憶部 2 1 0 により記憶されている映像データを再生装置 1 0 0 に送信し、再生させることとしてもよい。

【 0 0 7 2 】

取得部 2 2 0 や、解析部 2 3 0 などは、例えば、CPU、RAM などにより構成され、CPU が記憶部 2 1 0 によって記憶されているプログラムを RAM に展開して実行することによりその機能が実現されるものである。しかし、このような構成に限らず、取得部 2 2 0 や、解析部 2 3 0 の中には、専用のハードウェアにより構成されるものが存在してもよい。

【 0 0 7 3 】

出力装置 3 0 0 は、少なくとも、音声出力部 3 1 0 を備えるものである。音声出力部 3 1 0 は、例えば、スピーカまたはヘッドホンによって構成されるものである。音声出力部 3 1 0 は、映像音響データデコード回路 1 3 0 により再生された音響データをアナログ信号に変換して出力する機能を有している。出力装置 3 0 0 は、画像出力部 3 2 0 を有して

【 0 0 7 4 】

[1 - 9 . 属性情報の管理手法の一例]

図 9 は、本発明の実施形態に係る属性情報の管理手法の一例について説明するための図である。図 9 を参照しながら、本発明の実施形態に係る属性情報の管理手法の一例について説明する。

【 0 0 7 5 】

図 9 に示すように、再生対象音響データの属性情報は、提供装置 2 0 0 が備える記憶部 2 1 0 に記憶されている再生対象音響データと同一の楽曲ファイルにヘッダデータとして埋め込まれていることとしてもよい。解析部 2 3 0 は、例えば、上記した解析結果を属性情報として楽曲ファイルにヘッダデータとして埋め込んでおけばよい。そのようにすれば、提供装置 2 0 0 は、再生装置 1 0 0 からの再生対象音響データの送信要求に応じて、取得部 2 2 0 により属性情報が埋め込まれた楽曲ファイルを取得し、送信部 2 4 0 により再生装置 1 0 0 に送信すればよい。

【 0 0 7 6 】

[1 - 1 0 . 属性情報の管理手法の他の一例]

図 1 0 は、本発明の実施形態に係る属性情報の管理手法の他の一例について説明するための図である。図 1 0 を参照しながら、本発明の実施形態に係る属性情報の管理手法の他の例について説明する。

【 0 0 7 7 】

図 1 0 に示すように、再生対象音響データの属性情報は、提供装置 2 0 0 が備える記憶部 2 1 0 に記憶されている再生対象音響データと同一の楽曲ファイルに埋め込まれていなくてもよい。解析部 2 3 0 は、例えば、上記した解析結果を属性情報として再生対象音響データに紐付けをして記憶部 2 1 0 に記憶させておけばよい。提供装置 2 0 0 は、例えば、再生対象音響データを識別するための情報と属性情報とを対応付けて記憶部 2 1 0 に記憶しておく。

【 0 0 7 8 】

一方、再生装置 1 0 0 は、必要に応じて、再生対象音響データを識別するための情報を

10

20

30

40

50

含む要求を提供装置 200 に送信し、その要求に対する応答として再生対象音響データを識別するための情報に対応する属性情報を取得することができる。再生対象音響データを識別するための情報としては、例えば、再生対象音響データの一部を解析して得られる情報 (Finger print) を使用することができる。

【0079】

[1-11. 再生システムにより実行される処理]

図 11 は、本発明の実施形態に係る再生システムにより実行される処理の流れを示すフローチャートである。図 11 を参照しながら、本発明の実施形態に係る再生システム 10 により実行される処理の流れについて説明する。

【0080】

図 11 に示すように、提供装置 200 の解析部 230 は、再生対象音響データの Peak 値および RMS 値を予め計算しておくことができる (ステップ S101)。再生対象音響データの Peak 値および RMS 値は、再生対象音響データの音量特性値の例に相当する。続いて、記憶部 210 は、解析部 230 により計算された再生対象音響データの Peak 値および RMS 値を再生対象音響データと対応付けて属性情報として保持しておくことができる (ステップ S102)。計算結果と再生対象音響データとの対応付けについては、上記したように様々な手法が想定される。ステップ S101 およびステップ S102 は、再生システム 10 において 1 パス目の処理として実行され得るものである。記憶部 210 がこのように属性情報を保持しておき、必要に応じて取得されるようにしておけば、同一の再生対象音響データを次回以降に再生する際に再度 1 パス目の処理を実行する必要はなくなるという利点がある。

【0081】

続いて、再生装置 100 の受信部 110 は、映像音響データデコード回路 130 による再生対象音響データの再生前に再生対象音響データの属性情報をダウンロードする (ステップ S201)。再生装置 100 のパラメータ算出部 150 は、この属性情報に基づいて音量補正パラメータを算出し (ステップ S202)、算出した音量補正パラメータをコンプレッサー回路 160 に設定する (ステップ S203)。映像音響データデコード回路 130 は、再生対象音響データをデコードしながらコンプレッサー回路 160 に通す (ステップ S204)。ステップ S201 ~ ステップ S204 は、再生システム 10 において 2 パス目の処理として実行され得るものである。2 パス目の処理を実行すると、音量が補正された後の再生対象音響データが出力される。

【0082】

[1-12. 音響データの音量補正に適用し得るコンプレッサー回路]

図 12 は、本発明の実施形態に係る再生装置 100 による音響データの音量補正に適用し得るコンプレッサー回路を示す図である。図 12 を参照しながら、本発明の実施形態に係る再生装置 100 による音響データの音量補正に適用し得るコンプレッサー回路 160 について説明する。

【0083】

図 12 に示すように、コンプレッサー回路 160 は、映像音響データデコード回路 130 から入力された信号の大きさが出力目標レベルになるように信号を調整し、調整後の信号を出力信号として出力する。本発明の実施形態に係るコンプレッサー回路 160 は、既存のコンプレッサー回路とは異なり、例えば、線形増幅部 161 および非線形増幅部 163 を有している。また、コンプレッサー回路 160 は、例えば、線形増幅部 161 の後段に非線形増幅部 163 と並列に接続される線形増幅部 162 を有しており、かかるコンプレッサー回路 160 を有することは、本発明の実施形態に係る再生装置 100 の特徴の 1 つである。これらの増幅部をメタデータにより適切に制御することで、制作者の意図する元の音響データの音量変化カーブを維持しつつ各音響データ間の音量の違いを低減させ、音量増幅による音声の歪みを低減させながら音響データを再生することが可能となる。以下、映像音響データデコード回路 130 から入力される信号が音声信号 (例えば、再生対象音響データの音声信号) であるとして説明する。

10

20

30

40

50

【0084】

線形増幅部161は、第1の線形補正部の一例として機能し、図12においては「S gain」と表されている。また、線形増幅部162は、第2の線形補正部の一例として機能し、図12においては「T gain」と表されている。また、非線形増幅部163は、非線形補正部の一例として機能し、図12においては「D gain」と表されている。「S gain」「T gain」「D gain」は、各増幅部における増幅率を示している。

【0085】

線形増幅部161は、音声信号に対して線形の増幅処理を施すものである。線形増幅部162は、線形増幅部161から出力された音声信号に対して線形の増幅処理を施すものである。非線形増幅部163は、線形増幅部161から出力された音声信号に対して非線形の増幅処理を施すものであり、主に音声レベルの高い音声波形への振幅圧縮処理を行う。その振幅圧縮処理は、非線形処理により行われるため、音声に歪みが生じ、振幅圧縮率が高まるにつれ、その歪み率が大きくなってしまふ。

10

【0086】

[1-13. コンプレッサー回路が有する静特性(音量増幅時)の一例]

図13は、本発明の実施形態に係る再生装置100に適用し得るコンプレッサー回路160内の非線形増幅部163が有する静特性(音響データの音量を増幅する場合)の一例を示す図である。図13を参照しながら、本発明の実施形態に係る再生装置100に適用し得るコンプレッサー回路160内の非線形増幅部163が有する静特性(音響データの音量を増幅する場合)の一例について説明する。

20

【0087】

図13に示したグラフは、特に、再生対象音響データの音量を増幅する場合において、非線形増幅部163(増幅率D gainを有する増幅部)が有する静特性を示したものである。図13には、非線形増幅部163に対して入力される信号の大きさと非線形増幅部163から出力される信号の大きさとの関係が示されている。非線形増幅部163が有する動特性を考慮する場合には、図13に示された非線形増幅部163から出力される信号の大きさは、当該大きさを有する信号になるように非線形増幅部163により制御される出力目標レベルとしての意味を持つ。

【0088】

図13に示した例では、非線形増幅部163への入力信号の大きさがTに達するまでは、入力信号の大きさの増分に比例して非線形増幅部163からの出力信号の大きさの増分が変化する。また、非線形増幅部163への入力信号の大きさがTを超えると、非線形増幅部163からの出力信号の大きさの増分が0dBに飽和する。非線形増幅部163への入力信号の大きさがTに達するまでは、非線形増幅部163からの出力信号の大きさは、非線形増幅部163への入力信号の大きさに増幅率D gainを乗じた分だけ大きい値となっている。非線形増幅部163への入力信号の大きさがTに達するまでは、音声信号に対して非線形な増幅が行われていることになる。

30

【0089】

[1-14. コンプレッサー回路が有する静特性(音量減衰時)の一例]

図14は、本発明の実施形態に係る再生装置100に適用し得るコンプレッサー回路160内の非線形増幅部163が有する静特性(音響データの音量を減衰する場合)の一例を示す図である。図14を参照しながら、本発明の実施形態に係る再生装置100に適用し得るコンプレッサー回路160内の非線形増幅部163が有する静特性(音響データの音量を減衰する場合)の一例について説明する。

40

【0090】

図14に示したグラフは、特に、再生対象音響データの音量を減衰する場合において、非線形増幅部163(増幅率D gainを有する増幅部)が有する静特性を示したものである。図14には、非線形増幅部163に対して入力される信号の大きさと非線形増幅部163から出力される信号の大きさとの関係が示されている。非線形増幅部163が有する動特性を考慮する場合には、図14に示された非線形増幅部163から出力される信号

50

の大きさは、当該大きさを有する信号になるように非線形増幅部 163 により制御される出力目標レベルとしての意味を持つ。

【0091】

図 14 に示した例では、非線形増幅部 163 からの出力信号の大きさは、非線形増幅部 163 への入力信号の大きさに増幅率 $Dgain$ を乗じた分だけ小さい値となっている。非線形増幅部 163 への入力信号の大きさが T に達するまでは、音声信号に対して非線形な増幅が行われていることになる。図 14 に示した場合のように、音声信号の音量を減衰する場合には、音声信号に対して線形減衰を施すこととなるため、音声に歪みが発生することは考慮しなくてよい。

【0092】

[1-15. コンプレッサー回路が有する静特性(音量増幅時)の他の一例]

図 15 は、本発明の実施形態に係る再生装置 100 に適用し得るコンプレッサー回路 160 内の非線形増幅部 163 が有する静特性(音響データの音量を増幅する場合)の他の一例を示す図である。図 15 を参照しながら、本発明の実施形態に係る再生装置 100 に適用し得るコンプレッサー回路 160 内の非線形増幅部 163 が有する静特性(音響データの音量を増幅する場合)の他の一例について説明する。

【0093】

図 15 に示したグラフは、特に、再生対象音響データの音量を増幅する場合において、非線形増幅部 163 (増幅率 $Dgain$ を有する増幅部) が有する静特性を示したものであり、図 13 を参照しながら説明した静特性に代わって使用される例であり、図 13 に示した静特性と図 14 に示した静特性との複合型に相当するものである。図 15 には、非線形増幅部 163 に対して入力される信号の大きさと非線形増幅部 163 から出力される信号の大きさとの関係が示されている。非線形増幅部 163 が有する動特性を考慮する場合には、図 15 に示された非線形増幅部 163 から出力される信号の大きさは、当該大きさを有する信号になるように非線形増幅部 163 により制御される出力目標レベルとしての意味を持つ。

【0094】

図 15 に示した例では、非線形増幅部 163 への入力信号の大きさが T_1 に達するまでは、比較的大きな増幅率により入力信号の大きさの増分に比例して非線形増幅部 163 からの出力信号の大きさの増分が変化する。また、非線形増幅部 163 への入力信号の大きさが T_1 を超えると、非線形増幅部 163 からの出力信号の大きさの増分が低下していく。非線形増幅部 163 への入力信号の大きさが T_2 を超えてからは、非線形増幅部 163 からの出力信号の大きさは、非線形増幅部 163 への入力信号の大きさよりも低くなり、音声信号に対して非線形な増幅が行われていることになる。かかる増幅処理も振幅圧縮処理に相当し、音声に歪みが生じ得る。

【0095】

[1-16. コンプレッサー回路が有する動特性の一例]

図 16 は、本発明の実施形態に係る再生装置に適用し得るコンプレッサー回路が有する動特性の一例を示す図である。図 16 を参照しながら、本発明の実施形態に係る再生装置 100 に適用し得るコンプレッサー回路 160 内の非線形増幅部 163 が有する動特性の一例について説明する。

【0096】

一般的に、コンプレッサーは聴感上自然な音声振幅圧縮を実現するため、時間的に増幅率の変動する動特性を有している。図 16 (A) に示した入力信号は、トーンバースト信号に相当する。また、図 16 (B) に示した信号は、コンプレッサーが生成したエンベロープ信号に相当する。コンプレッサー回路 160 内の非線形増幅部 163 は、図 16 (A) に示した入力信号に基づいて図 16 (B) に示したエンベロープ信号を生成することができる。エンベロープ信号の生成については、図 17 を参照しながら後に詳細に説明する。

【0097】

10

20

30

40

50

非線形増幅部 163 は、図 16 (B) に示したエンベロープ信号の大きさと出力目標レベルとの比率から、増幅ゲイン率を求める。具体的には、非線形増幅部 163 は、出力目標レベルをエンベロープ信号の大きさとで除算して増幅ゲイン率とする。非線形増幅部 163 は、この増幅ゲイン率を、時々刻々、図 16 (A) に示した入力信号に乗算することで、図 16 (C) に示した出力信号を得ることができる。図 16 (C) に示した出力信号は、トーンバースト信号のエンベロープを表している。

【0098】

[1 - 17 . コンプレッサー回路における制御信号の生成手法の一例]

図 17 は、本発明の実施形態に係る再生装置 100 に適用し得るコンプレッサー回路 160 内の非線形増幅部 163 における制御信号 (エンベロープ信号) の生成手法の一例を示す図である。図 17 を参照しながら、本発明の実施形態に係る再生装置 100 に適用し得るコンプレッサー回路 160 における制御信号 (エンベロープ信号) の生成手法の一例について説明する。

10

【0099】

非線形増幅部 163 は、図 16 (A) に示した入力信号に対して絶対値をとり、時定数を持たせる。時定数には 3 種類あり、立ち上がり率を表すアタックタイム、ピークを保持するホールドタイム、立ち下がり率を表すリリースタイムである。非線形増幅部 163 は、これらのパラメータを変えることで、コンプレッサー回路 160 内の非線形増幅部 163 の挙動を変えることができる。具体的には、非線形増幅部 163 は、これらのパラメータを変えることでコンプレッサー回路 160 に音響データを通して得られる音声の音色が変化する。一般的には、聴感上において適切な値をパラメータの固定値として設定しておく。

20

【0100】

[1 - 18 . 音響データの音量補正の処理の流れの詳細]

図 18 は、本発明の実施形態に係る再生装置による音響データの音量補正の処理の流れの詳細を示すフローチャートである。図 18 を参照しながら、本発明の実施形態に係る再生装置 100 による音響データの音量補正の処理の流れの詳細について説明する。なお、図 18 に示した処理は、上記した 2 パス目の処理に相当する。

【0101】

図 18 に示すように、再生装置 100 の属性情報取得部 140 は、受信部 110 を介して再生対象音響データの RMS 値を取得する (ステップ S301)。パラメータ算出部 150 は、属性情報取得部 140 により取得された RMS 値とあらかじめ定められた RMS 基準値との大小を比較する (ステップ S302)。RMS 基準値は、複数の音響データの RMS 値の分布に基づいて決定されるものであり、その決定手法については後に詳細に説明する。

30

【0102】

パラメータ算出部 150 は、RMS 値が基準値より小さい場合は、音量増幅 (ゲイン増幅) を行うためのパラメータを算出し、RMS 値が基準値より大きい場合は、音量減衰 (ゲイン減衰) を行うためのパラメータを算出する。RMS 値が基準値と同じ場合は、音量補正を行わない (スルー出力する) ようにするためのパラメータを算出する。パラメータ算出部 150 は、例えば、コンプレッサー回路 160 内の線形増幅部 161 に設定される線形補正用パラメータ (例えば、増幅率 S_{gain})、およびコンプレッサー回路 160 内の非線形増幅部 163 に設定される非線形補正用パラメータ (例えば、増幅率 D_{gain}) を算出することができる。コンプレッサー回路 160 は、パラメータ算出部 150 により算出されたパラメータに基づいて、音量補正を行う。

40

【 0 1 0 3 】

[1 - 1 9 . パラメータ算出部によるパラメータ（音量増幅時）の算出例]

図 1 9 は、本発明の実施形態に係る再生装置 1 0 0 に適用し得るパラメータ算出部 1 5 0 によるパラメータの算出例（音響データの音量を増幅する場合）を示す図である。図 1 9 を参照しながら、本発明の実施形態に係る再生装置 1 0 0 に適用し得るパラメータ算出部 1 5 0 によるパラメータの算出例（音響データの音量を増幅する場合）について説明する。

【 0 1 0 4 】

図 1 9 に示すように、音響データの音量を増幅する場合、パラメータ算出部 1 5 0 は、再生対称音響データの音量が全体で（RMS 基準値 / RMS 値）倍に増幅するよう、線形増幅部 1 6 1 に設定される増幅率 S g a i n（線形補正用パラメータの一例）および非線形増幅部 1 6 3 に設定される増幅率 D g a i n（非線形補正用パラメータの一例）を算出する。

10

【 0 1 0 5 】

まず、線形増幅部 1 6 1 に設定される増幅率 S g a i n の算出方法について述べる。ここでは、再生対象音響データのメタデータの 1 つである P e a k 値を利用する。パラメータ算出部 1 5 0 は、例えば、メタデータの 1 つである P e a k 値と音量上限値（例えば、「3 2 7 6 8」）との比率を線形増幅部 1 6 1 に設定される増幅率 S g a i n として算出する。パラメータ算出部 1 5 0 により算出された増幅率 S g a i n は、パラメータ算出部 1 5 0 により線形増幅部 1 6 1 に設定される。線形増幅部 1 6 1 では線形増幅を行い、メタデータの 1 つである P e a k 値を考慮した増幅のため、飽和しないことが保証されている。このため、線形増幅部 1 6 1 における音量増幅に際して音声の歪みは発生しない。但し、線形増幅部 1 6 1 では音量増幅が足りない可能性があり、音量増幅の不足分については非線形増幅部 1 6 3 にて補うことができる。

20

【 0 1 0 6 】

次に、非線形増幅部 1 6 3 に設定される増幅率 D g a i n の算出方法について述べる。ここでは、再生対象音響データのメタデータの 1 つである R M S 値を利用する。パラメータ算出部 1 5 0 は、再生対象音響データの R M S 値に対する R M S 基準値の比率に基づいて増幅率 D g a i n を算出する。パラメータ算出部 1 5 0 により算出された増幅率 D g a i n は、パラメータ算出部 1 5 0 により非線形増幅部 1 6 3 に設定される。パラメータ算出部 1 5 0 は、例えば、（RMS 基準値 / RMS 値） / S g a i n を増幅率 D g a i n として算出する。かかる値が増幅率 D g a i n としてパラメータ算出部 1 5 0 により設定されれば、非線形増幅部 1 6 3 における増幅を、増幅率 S g a i n を考慮して低く抑えられる。すなわち、音声の歪みを低く抑えることができる。増幅率 D g a i n は、図 1 3 に示した静特性における増幅率に相当する。

30

【 0 1 0 7 】

[1 - 2 0 . パラメータ算出部によるパラメータ（音量減衰時）の算出例]

図 2 0 は、本発明の実施形態に係る再生装置 1 0 0 に適用し得るパラメータ算出部 1 5 0 によるパラメータの算出例（音響データの音量を減衰する場合）を示す図である。図 2 0 を参照しながら、本発明の実施形態に係る再生装置 1 0 0 に適用し得るパラメータ算出部 1 5 0 によるパラメータの算出例（音響データの音量を減衰する場合）について説明する。

40

【 0 1 0 8 】

図 2 0 に示すように、音響データの音量を減衰する場合、パラメータ算出部 1 5 0 は、再生対象音響データの R M S 値に対する R M S 基準値の比率を線形増幅部 1 6 1 に設定される増幅率 S g a i n（線形補正用パラメータの一例）として算出する。パラメータ算出部 1 5 0 により算出された増幅率 S g a i n は、パラメータ算出部 1 5 0 により線形増幅部 1 6 1 に設定される。

【 0 1 0 9 】

パラメータ算出部 1 5 0 は、「0 . 0」を増幅率 D g a i n（非線形補正用パラメータ

50

の一例)として算出する。「1.0」を増幅率 $T g a i n$ として算出する。パラメータ算出部 150 により算出された増幅率 $D g a i n$ および増幅率 $T g a i n$ は、パラメータ算出部 150 により非線形増幅部 163 および線形増幅部 162 に設定される。かかる値がパラメータ算出部 150 により設定されれば、線形増幅部 162 により線形減衰を行うため、音声の歪み無く減衰することができる。線形増幅部 162 はスルー出力を行い、非線形増幅部 163 は無音を出力し、線形増幅部 162 および非線形増幅部 163 から出力された両出力が加算され、最終段の出力信号となる。

【0110】

[1-21. パラメータ算出部によるパラメータ(スルー出力時)の算出例]

図 21 は、本発明の実施形態に係る再生装置 100 に適用し得るパラメータ算出部 150 によるパラメータの算出例(音響データをスルー出力する場合)を示す図である。図 21 を参照しながら、本発明の実施形態に係る再生装置 100 に適用し得るパラメータ算出部 150 によるパラメータの算出例(音響データをスルー出力する場合)について説明する。

10

【0111】

図 21 に示すように、音響データをスルー出力する場合、パラメータ算出部 150 は、「1.0」を線形増幅部 161 に設定される増幅率 $S g a i n$ (線形補正用パラメータの一例)として算出する。また、パラメータ算出部 150 は、「0.0」を増幅率 $D g a i n$ (非線形補正用パラメータの一例)として算出する。また、パラメータ算出部 150 は、「1.0」を増幅率 $T g a i n$ として算出する。

20

【0112】

パラメータ算出部 150 により算出された増幅率 $S g a i n$ 、増幅率 $D g a i n$ および増幅率 $T g a i n$ は、パラメータ算出部 150 により線形増幅部 161、非線形増幅部 163 および線形増幅部 162 に設定される。かかる値がパラメータ算出部 150 により設定されれば、非線形増幅部 163 による処理が行われないため、歪みのない音声を出力することができる。

【0113】

次に、パラメータ算出部 150 において算出に用いられている $R M S$ 基準値の求め方について述べる。 $R M S$ 基準値は、例えば、再生対象音響データの $P e a k$ 値及び $R M S$ 値の統計分布から算出する。

30

[1-22. 音量特性値の一例としての $P e a k$ 値の統計分布]

図 22 は、本発明の実施形態に係る音量特性値の一例としての $P e a k$ 値の統計分布について説明するための図である。図 22 を参照しながら、本発明の実施形態に係る音量特性値の一例としての $P e a k$ 値の統計分布について説明する。

【0114】

図 22 は、一般的な楽曲(音響データの例)群の $P e a k$ 値の統計分布(ヒストグラム)を示している。多くの楽曲の $P e a k$ 値は音量の上限値である 32768 の付近となっているが、図 22 に示されたヒストグラムを参照する限りは、 $P e a k$ 値が音量の上限値付近にない曲も多く存在しているので、そのような楽曲に対しては、図 19 に示したような、線形増幅のための各増幅率の組み合わせを適用することで、非線形増幅による音声の歪みを適切に抑えることができる。本例では、 $R M S$ 基準値の算出に $P e a k$ 値の統計分布を直接利用していないが、例えば、 $P e a k$ 値の統計分布の平均値と後述する $R M S$ 値の統計分布の平均値との差分とから、最適な増幅率となるよう非線形増幅部 163 が有する静特性と動特性を変更してもよい。

40

【0115】

[1-23. 音量特性値の一例としての $R M S$ 値の統計分布]

図 23 は、本発明の実施形態に係る音量特性値の一例としての $R M S$ 値の統計分布について説明するための図である。図 23 を参照しながら、本発明の実施形態に係る音量特性値の一例としての $R M S$ 値の統計分布について説明する。

【0116】

50

図 2 3 は、一般的な楽曲（音響データの例）群の R M S 値の統計分布（ヒストグラム）を示している。図 2 3 を参照すると、R M S 値の平均値を中心にほぼ正規分布で広範囲に広がっているのがわかる。数十 d B の音量差で、さまざまな楽曲が分布していることがわかる。リスナーが再生しようとしている楽曲群に相当する音楽ライブラリは既知であるため、それらの楽曲群の統計分布を予め求めておくことは可能である。ネットワークを介したストリーミング再生のユースケースに関しても同様である。本発明の実施形態においては、再生する楽曲群の統計分布を、図 2 2、図 2 3 などに示したようにあらかじめ計算して求めておくことができる。

【 0 1 1 7 】

[1 - 2 4 . R M S 基準値の決定手法の一例]

図 2 4 は、本発明の実施形態に係る再生装置 1 0 0 による音響データの音量補正の処理に使用され得る R M S 基準値の決定手法の一例について説明するための図である。図 2 4 を参照しながら、本発明の実施形態に係る再生装置 1 0 0 による音響データの音量補正の処理に使用され得る R M S 基準値の決定手法の一例について説明する。

【 0 1 1 8 】

図 2 4 に示すように、解析部 2 3 0 は、統計分布から R M S 値の平均値を求めることができる。解析部 2 3 0 は、R M S 値の平均値を R M S 基準値としてもよいが、図 2 4 に示した例では、R M S 値の平均値 + 1 . 5 シグマ（シグマは標準偏差）を R M S 基準値としている。また、解析部 2 3 0 は、R M S 値の平均値 + 3 シグマ - 数 d B（例えば、6 d B）を R M S 基準値としてもよい。解析部 2 3 0 は、R M S 値の平均値を R M S 基準値としない理由は、より多くの楽曲に対して、音量補正をゲイン増幅モードで動作させるためであり、量子化ビット数 1 6 ビットを有効に利用した S / N の良い音にすることができる。

【 0 1 1 9 】

ノーマライザーを 3 つの動作モード（ゲイン増幅、ゲイン減衰、スルー出力）のいずれにより動作させるのが好ましいかはユーザに嗜好によるので、この R M S 基準値の決め方は一通りではなく、ユーザが統計分布の R M S 平均値を基準に好きな音量レベルに決めてよい。

【 0 1 2 0 】

[1 - 2 5 . R M S 基準値の決定手法の他の一例]

図 2 5 は、本発明の実施形態に係る再生装置 1 0 0 による音響データの音量補正の処理に使用され得る R M S 基準値の決定手法の他の一例について説明するための図である。図 2 5 を参照しながら、本発明の実施形態に係る再生装置 1 0 0 による音響データの音量補正の処理に使用され得る R M S 基準値の決定手法の他の一例について説明するための図である。

パラ

【 0 1 2 1 】

図 2 4 に示した例では、R M S 基準値を一つとしたが、図 2 5 に示した例では、R M S 基準値の上限値と下限値の 2 つを設けている。このとき、パラメータ算出部 1 5 0 は、R M S 基準値よりも小さい下側基準値（下限の R M S 基準値）より再生対象音響データの R M S 値が小さい場合に再生対象音響データの音量を増幅するためのパラメータを算出する。また、パラメータ算出部 1 5 0 は、R M S 基準値よりも大きい上側基準値（上限の R M S 基準値）より再生対象音響データの R M S 値が大きい場合に再生対象音響データの音量を減衰するためのパラメータを算出する。パラメータ算出部 1 5 0 は、下限の R M S 基準値と上限の R M S 基準値との間に再生対象音響データの R M S 値が含まれる場合、再生対象音響データをスルー出力させるためのパラメータを算出する。

【 0 1 2 2 】

R M S 基準値と下側基準値との差分および R M S 基準値と上側基準値との差分については、特に限定されるものではなく、例えば、3 d B とすることができる。すなわち、R M S 基準値 + 3 d B を上限の R M S 基準値とし、R M S 基準値 - 3 d B レベルを下限の R M S 基準値とすることができる。コンプレッサー回路 1 6 0 は、上限の R M S 基準値より大き

10

20

30

40

50

いRMS値を持つ楽曲に対してはゲイン減衰の音量補正動作を行い、下限のRMS基準値より小さいRMS値を持つ楽曲に対してはゲイン増幅の音量補正動作を行うことができる。上限のRMS基準値と下側のRMS基準値との間のRMS値を持つ楽曲に対しては何もせず、楽曲をスルー出力する。スルー出力することで、再生対象音響データに必要な以上の音量制御処理をする必要がなくなり、音質の劣化を最小限にすることができる。これらの動作モードの設定に関しても、ユーザ側で統計分布のRMS平均値を基準に自由に決めることができる。

【0123】

< 2 . 使用例 >

以上に説明した本発明の実施形態に係る再生システム10の使用例と一般的な再生システムの使用例とについて説明する。以下に示すように、一般的な再生システムの使用例についてシナリオ1、2に記載し、本発明の実施形態に係る再生システム10の使用例についてシナリオ例3に記載する。

【0124】

シナリオ例1)

Aさんはこれまでに購入したCDが100枚くらいあり、これらの楽曲は全てPCなどにリップングして保存している。楽曲数として1000曲くらいがPCに保存されている。その中にはクラシックの曲やロックの曲が混じっている。AさんはPC上の音楽ファイルを専用のソフトウェアを使ってランダム再生してみた。

【0125】

1曲目はロックが再生された。音量が大きめだったので音量ボリュームを10段階中3にした。次に2曲目にクラシックが再生された。音量がとても小さく感じたので、Aさんは音量ボリュームを3から7に上げた。ところがこのクラシックの曲は後半にとっても盛り上がり、うるさく感じられてきた。そこでAさんはボリュームを5に下げた。3曲目は再び、別のロックの曲が再生された、音量がすごく大きく感じられたので、ボリュームを5から3に下げた。Aさんは音量調整をととても煩わしく感じた。

【0126】

シナリオ例2)

Aさんは、既存の「ダイナミック型」のノーマライザー処理をONにした。AさんはPC上の音楽ファイルを専用のソフトウェアを使ってランダム再生してみた。1曲目はロックが再生された。音量が大きめだったので音量ボリュームを10段階中3にした。次に2曲目にクラシックが再生された。音量はロックの曲とそんなに差を感じなかったので3のまま再生した。ところがクラシックの曲のフルートやピアノの音が原曲と比べて、なぜか歪みっぽく感じられた。

【0127】

そして、このクラシック曲は前半が静かで後半がすごく盛り上がるはずなのに、前半からフルートとピアノの音量がすごく上がり過ぎているのが気になった。前半の音量感の勢いそのまま後半の盛り上がり部分に突入したので、後半のドラマチックな盛り上がり方を感じ取ることができなかった。3曲目は再び、別のロックの曲が再生された。音量調整は全体を通して3のままで変える必要がなかったが、クラシックの曲の音量の不自然さや歪みによる音質の劣化が気になった。

【0128】

シナリオ例3)

Aさんは本発明の実施形態に係る再生システム10の音量補正機能をONにした。AさんはPC上の音楽ファイルを専用のソフトウェアを使ってランダム再生してみた。1曲目はロックが再生された。音量が大きめだったので音量ボリュームを10段階中3にした。次に2曲目にクラシックが再生された。静かな感じのフルートとピアノの演奏から始まったが、音量を上げようと思う程ではなかった。むしろ、静寂の中のフルートとピアノの演奏が心地良かった。音も透き通ったようなきれいな音色が響いていた。後半は打楽器を含む大人数でのオーケストラの演奏が始まり、すごく盛り上がり感を感じとることができた

10

20

30

40

50

。この場合でも音量は3のまま最適音量感を感じとることができた。

【0129】

クラシック曲の再生が終わり、3曲目は再び、別のロックの曲が再生された。クラシックの後にロック調の気持の良いドラムやギター音が再生されたが、音量感はちょうど良かった。音量調整は全体を通して3のまま変える必要がなかった。全体を通して、クラシックのやさしい音色とロックの激しい音色の対比が気持ちよかった。音量感は全体を通して3でちょうど良かった。

【0130】

<3.変形例>

以上、添付図面を参照しながら本発明の好適な実施形態について詳細に説明したが、本発明はかかる例に限定されない。本発明の属する技術の分野における通常の知識を有する者であれば、特許請求の範囲に記載された技術的思想の範疇内において、各種の変更例または修正例に想到し得ることは明らかであり、これらについても、当然に本発明の技術的範囲に属するものと了解される。

10

【0131】

例えば、上記実施形態では、提供装置200により保持されている楽曲に対してRMS基準値を算出するための統計処理を行うこととしたが、本発明はかかる例に限定されない。例えば、RMS基準値を楽曲のジャンルごとに行うようにしてもよい。例えば、クラシック曲の楽曲群のRMS基準値とロック曲の楽曲群のRMS基準値とを別個に算出し、再生装置100による音量補正に使用することも可能である。また、各ユーザのプレイリストごとに算出してもよい。また、上記実施形態では、RMS基準値を算出するための統計処理を提供装置200が行うこととしたが、再生装置100が行うこととすることもできる。

20

【0132】

<4.まとめ>

本実施形態によれば、元の音響データの音量変化に比較的近い音量変化を維持しつつ各音響データ間の音量の違いを低減させ、音量増幅による音声の歪みを低減させながら音響データを再生することが可能である。

【符号の説明】

【0133】

- 10 再生システム
- 100 再生装置
- 110 受信部
- 120 映像音響データ取得部
- 130 映像音響データデコード回路(再生部)
- 140 属性情報取得部
- 150 パラメータ算出部
- 160 コンプレッサー回路(音量補正部)
- 161 線形増幅部
- 162 線形増幅部
- 163 非線形増幅部
- 170 ミキサー装置
- 180 記憶部
- 200 提供装置
- 210 記憶部
- 220 取得部
- 230 解析部
- 240 送信部
- 300 出力装置
- 310 音声出力部

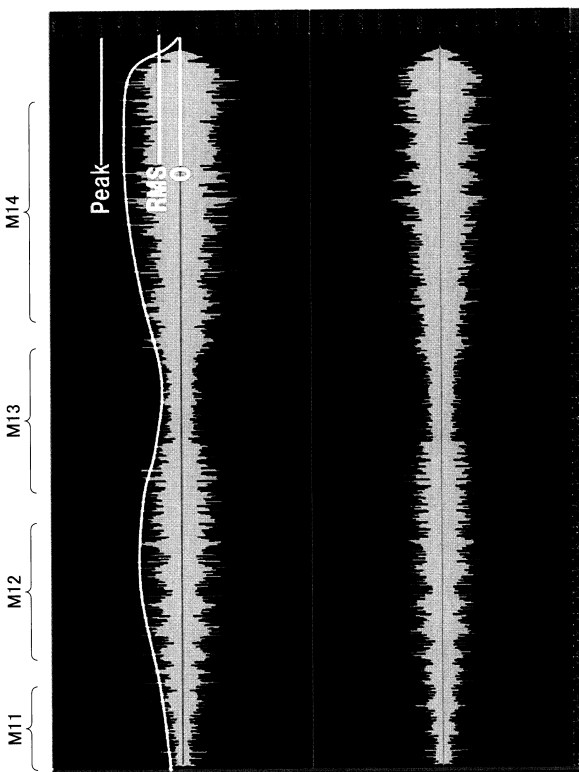
30

40

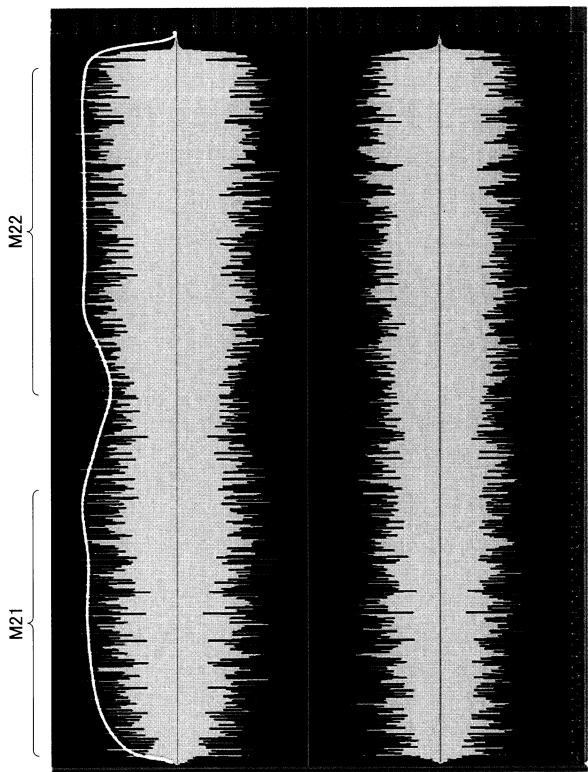
50

3 2 0 画像出力部

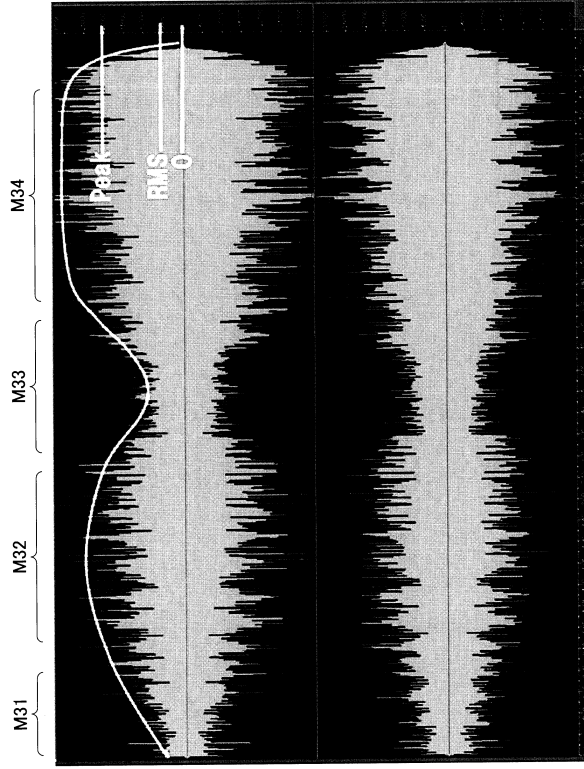
【 図 1 】



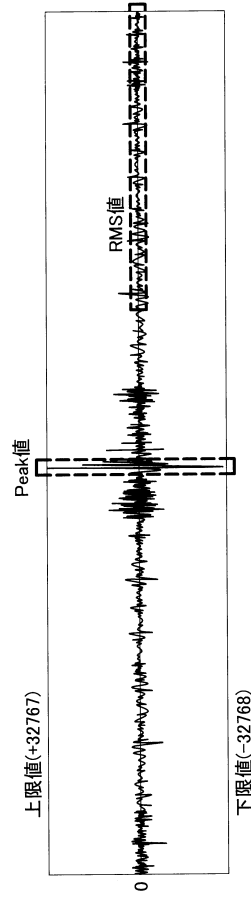
【 図 2 】



【 図 3 】



【 図 4 】



【 図 5 】

Peak値計算式

- モノラル音声の場合
モノラルPCMの各サンプル値をXi、1曲の総サンプル数をnとする。
max(a,b)は大きい方の値を返す関数。
abs(a)は絶対値を返す関数。

式1)
繰り返し返し(i=1からi=nまで)
{
 Peak = max(Peak, abs(Xi))
}

- ステレオ音声の場合
ステレオPCMの場合、Lチャンネルのサンプル値をXiL、Rチャンネルのサンプル値をXiRとすると、下記計算式となる。

式2)
繰り返し返し(i=1からi=nまで)
{
 Peak = max(Peak, abs(XiL))
 Peak = max(Peak, abs(XiR))
}

【 図 6 】

RMS値計算式

- モノラル音声の場合
モノラルPCMの各サンプル値をXi、1曲の総サンプル数をnとする。
sqrt()は平方根を返す関数。

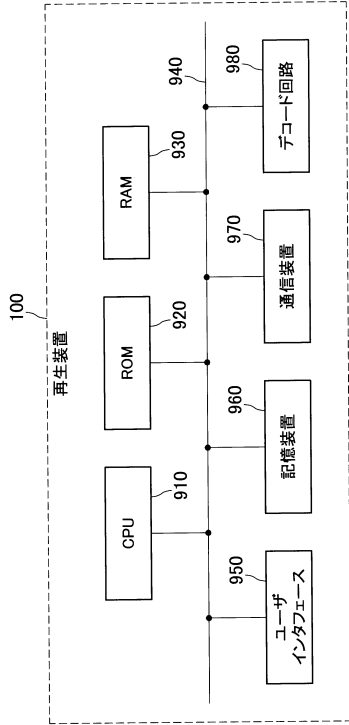
式3)
$$RMS = \sqrt{\left(\sum_{i=1}^n (Xi * Xi)\right) / n}$$

- ステレオ音声の場合
ステレオPCMの場合、Lチャンネルのサンプル値をXiL、Rチャンネルのサンプル値をXiRとすると、下記計算式となる。
sqrt()は平方根を返す関数。

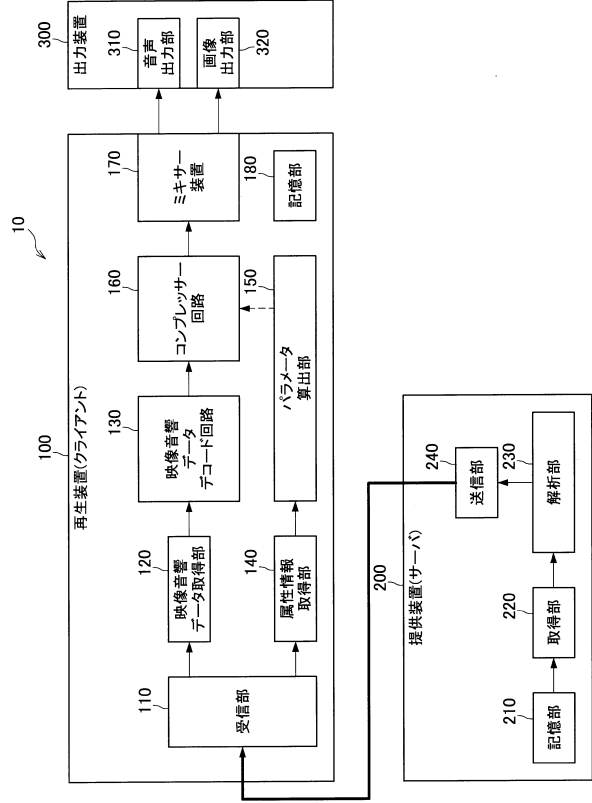
式4)
繰り返し返し(i=1からi=nまで)
{
 Xi = (XiL + XiR) / 2

式5)
$$RMS = \sqrt{\left(\sum_{i=1}^n (Xi * Xi)\right) / n}$$

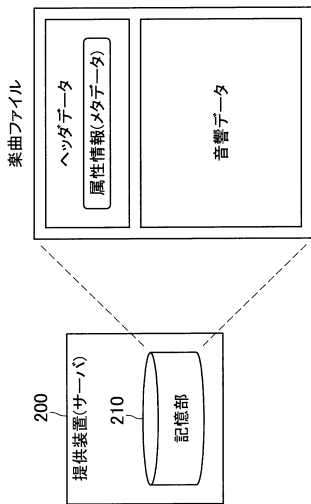
【図7】



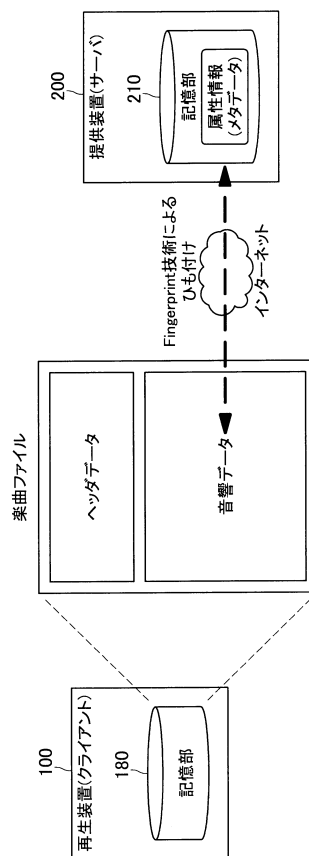
【図8】



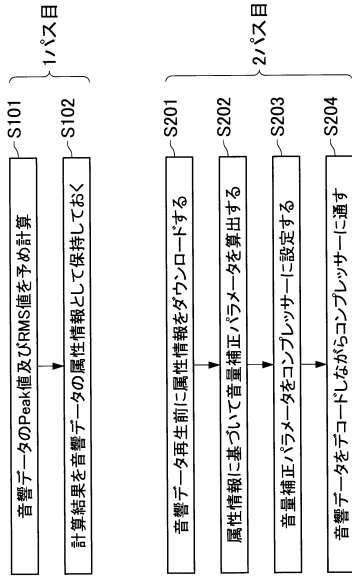
【図9】



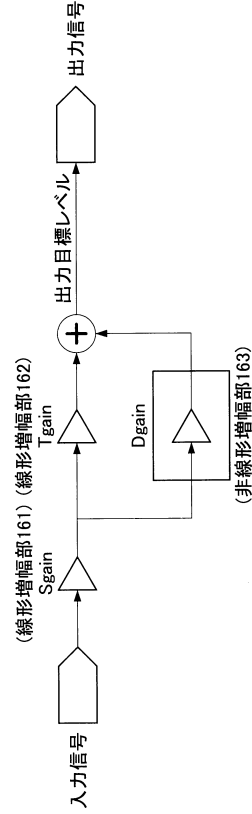
【図10】



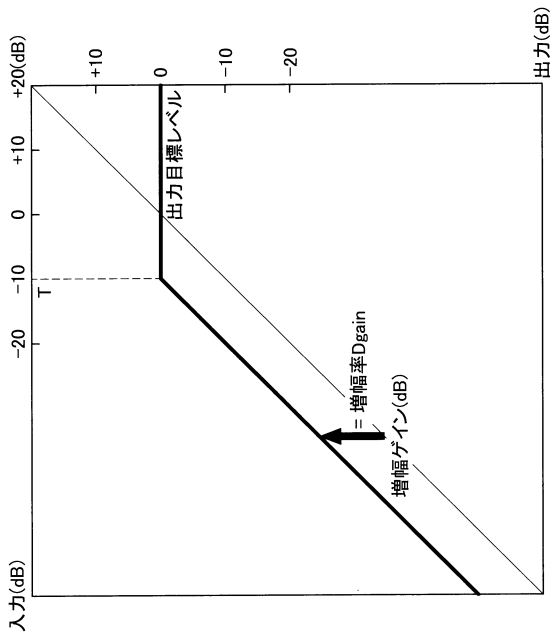
【 図 1 1 】



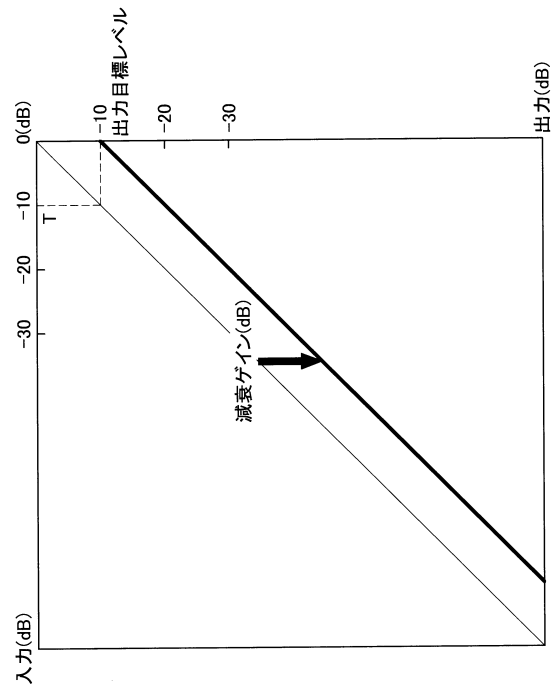
【 図 1 2 】



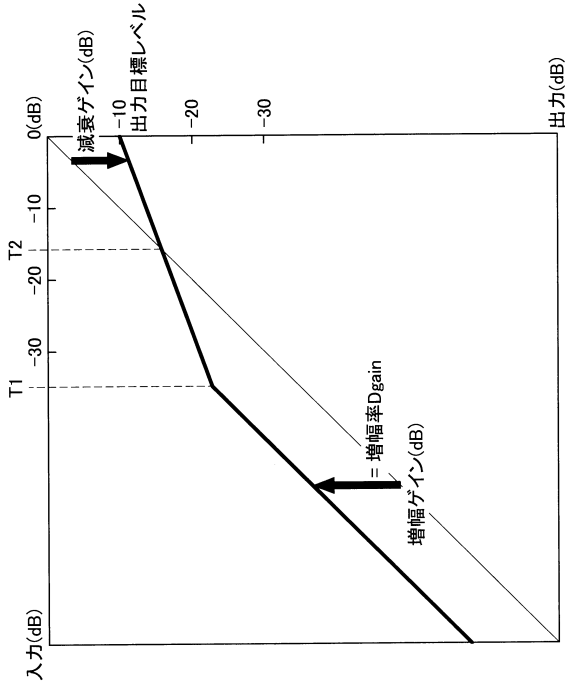
【 図 1 3 】



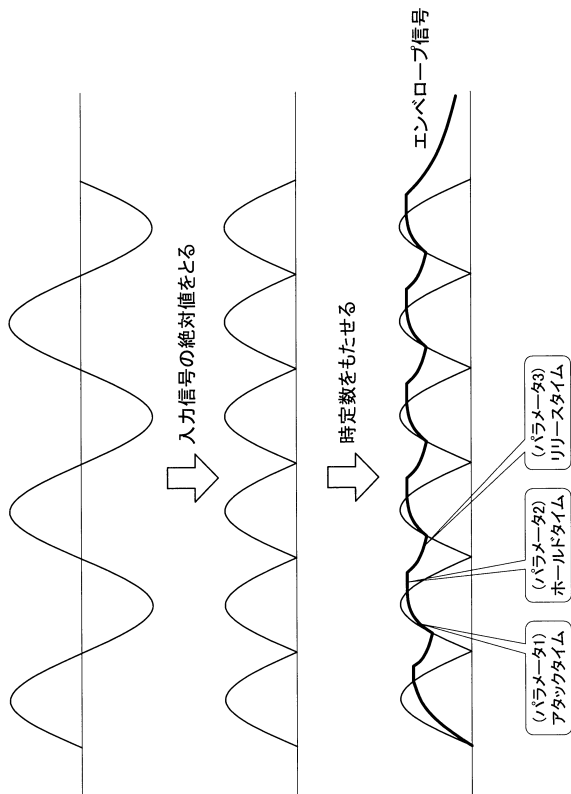
【 図 1 4 】



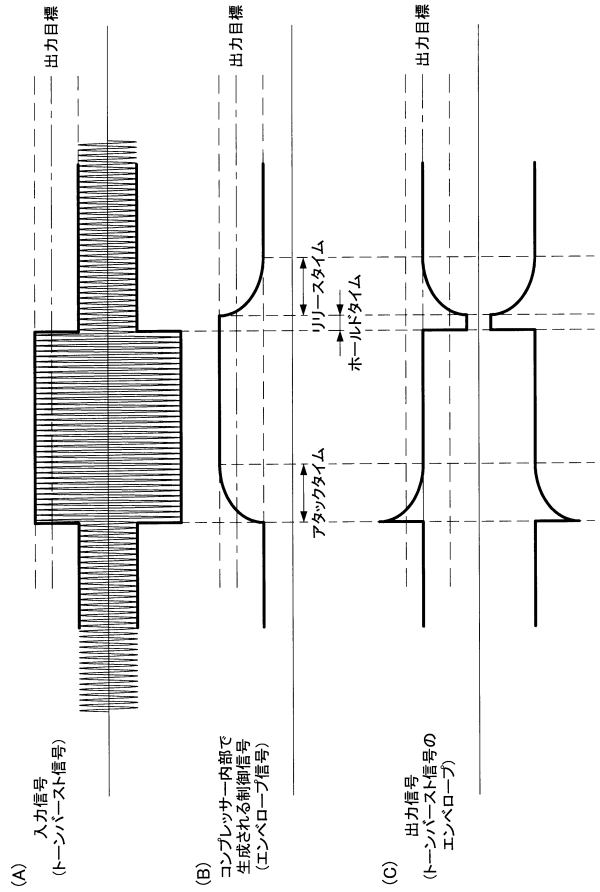
【図 15】



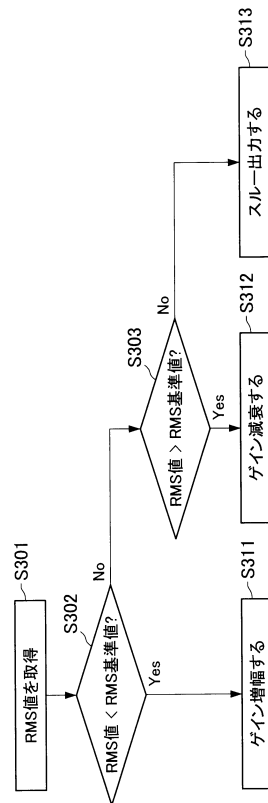
【図 17】



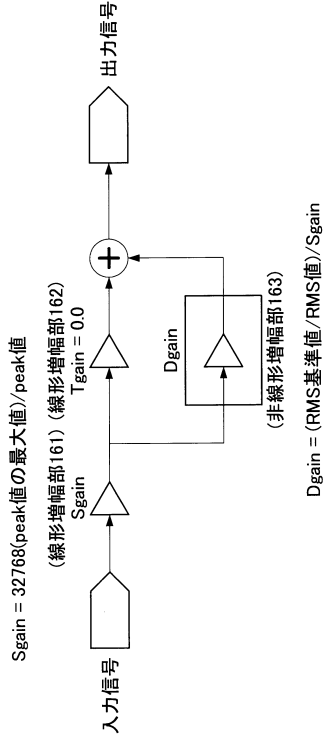
【図 16】



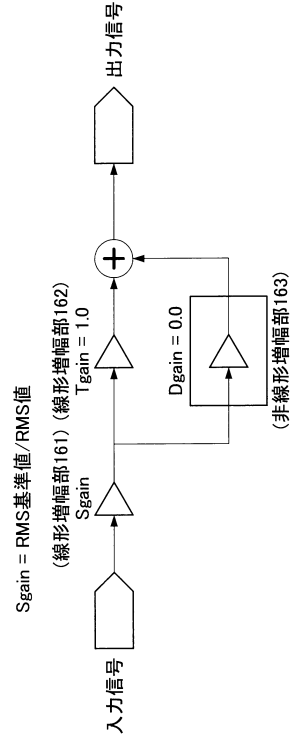
【図 18】



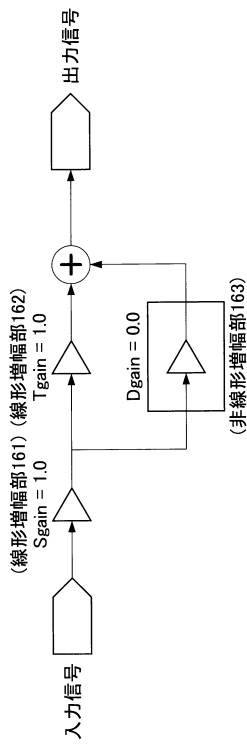
【 図 19 】



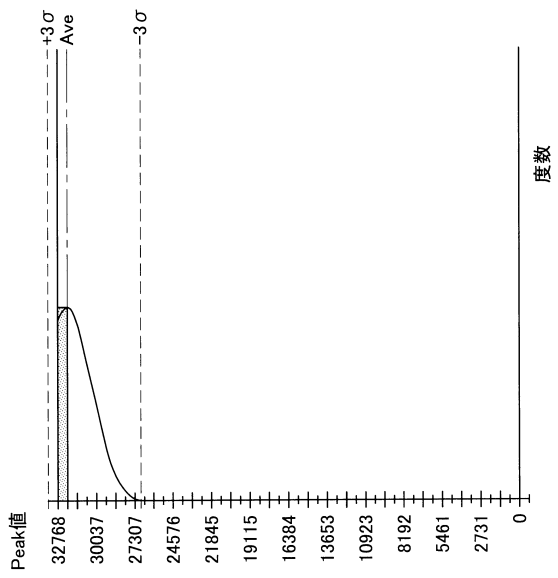
【 図 20 】



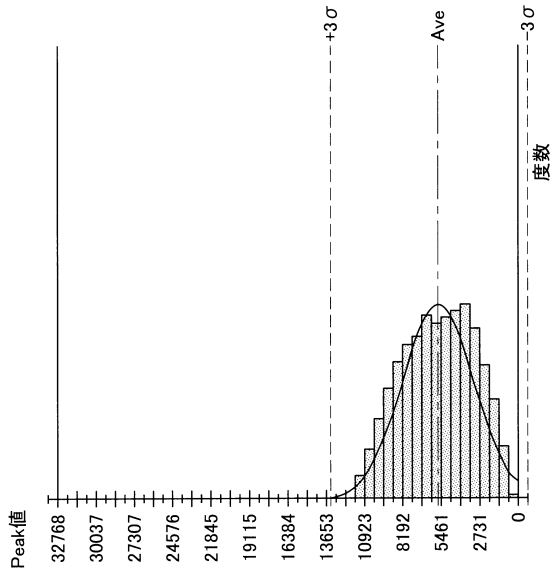
【 図 21 】



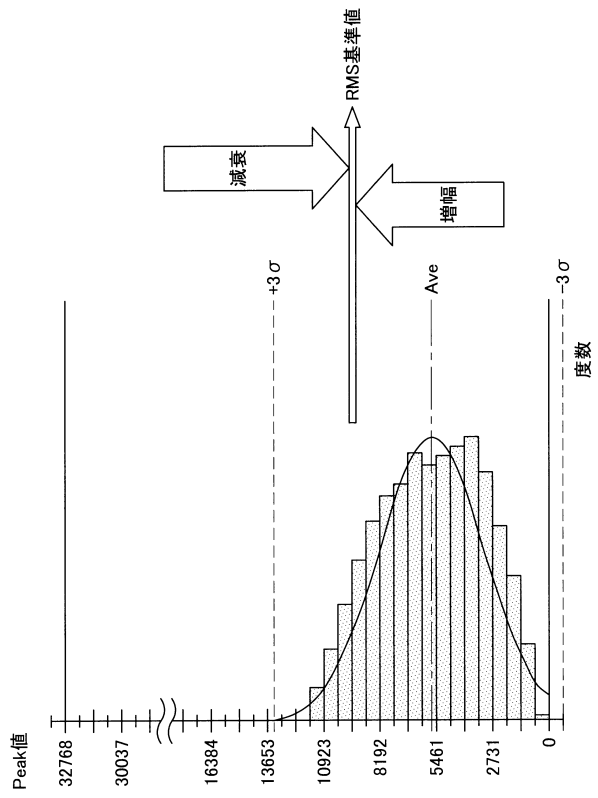
【 図 22 】



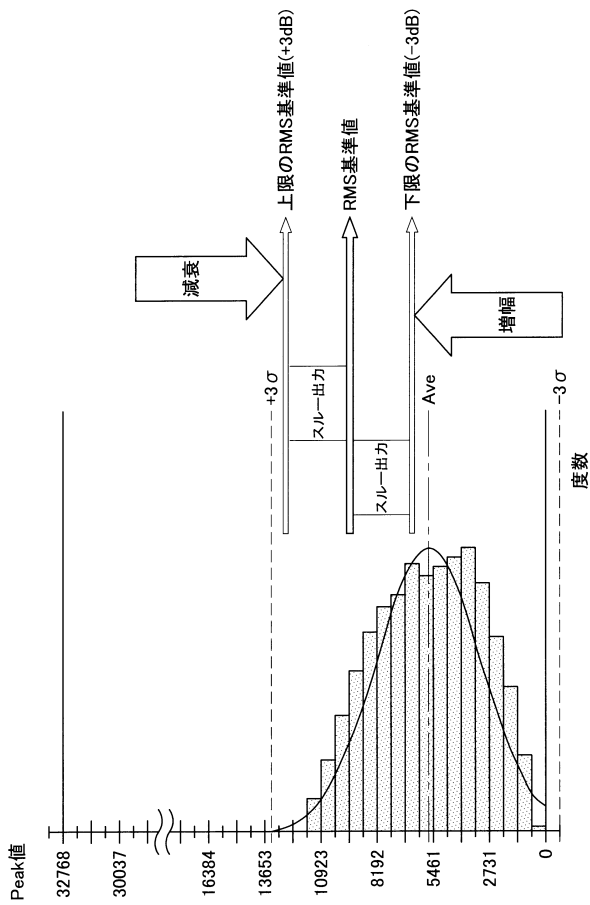
【 図 2 3 】



【 図 2 4 】



【 図 2 5 】



フロントページの続き

(72)発明者 板橋 徹徳
東京都港区港南1丁目7番1号 ソニー株式会社内

審査官 吉 澤 雅博

(56)参考文献 特開2008-065904(JP,A)
特表2008-521028(JP,A)
特開平11-352967(JP,A)
特開2007-003845(JP,A)

(58)調査した分野(Int.Cl., DB名)
H04R 3/00