

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第5527827号
(P5527827)

(45) 発行日 平成26年6月25日(2014.6.25)

(24) 登録日 平成26年4月25日(2014.4.25)

(51) Int.Cl. F I
H03G 9/00 (2006.01) H03G 9/00 A
H03G 7/00 (2006.01) H03G 7/00 A

請求項の数 10 (全 23 頁)

(21) 出願番号 特願2012-94004(P2012-94004)
 (22) 出願日 平成24年4月17日(2012.4.17)
 (65) 公開番号 特開2013-223130(P2013-223130A)
 (43) 公開日 平成25年10月28日(2013.10.28)
 審査請求日 平成24年4月17日(2012.4.17)

(73) 特許権者 303013763
 NECエンジニアリング株式会社
 神奈川県川崎市中原区下沼部1753番地
 (74) 代理人 100103894
 弁理士 冢入 健
 (72) 発明者 濱島 力
 東京都品川区東品川四丁目10番27号
 NECエンジニアリング株式会社内
 審査官 高橋 義昭

最終頁に続く

(54) 【発明の名称】 ラウドネス調整装置、ラウドネス調整方法、及びプログラム

(57) 【特許請求の範囲】

【請求項1】

入力デジタル音声信号のラウドネス調整を行うラウドネス調整装置であって、
 前記入力デジタル音声信号から抽出された音声データに対し、ゲイン調整を行うゲイン調整部と、

前記ゲイン調整部によるゲイン調整済みの音声データを特性フィルタ処理したフィルタ済み音声データを生成する第1フィルタ部と、

前記第1フィルタ部の生成したフィルタ済み音声データが第1閾値以下の場合、前記ゲイン調整部によるゲイン調整済みの音声データを増幅して出力するブースト処理部と、

前記ブースト処理部が出力した音声データを特性フィルタ処理したフィルタ済み音声データを生成する第2フィルタ部と、

前記第2フィルタ部の生成したフィルタ済み音声データが第2閾値以上の場合、前記ブースト処理部が増幅した音声データに対してリミッタコンプレッサ処理を行って出力するリミッタコンプレッサ部と、を備えるラウドネス調整装置。

【請求項2】

前記リミッタコンプレッサ部が出力した音声データを特性フィルタ処理したフィルタ済み音声データを生成する第3フィルタ部と、

前記第3フィルタ部が生成したフィルタ済み音声データに対し、第1の時間区間を用いた移動二乗平均により第1二乗平均値を算出する第1二乗平均算出部と、

前記第3フィルタ部が生成したフィルタ済み音声データに対し、第1の時間区間より長

10

20

い第 2 の時間区間を用いた移動二乗平均により第 2 二乗平均値を算出する第 2 二乗平均算出部と、

前記第 1 及び第 2 二乗平均値、及びターゲットラウドネス値に基づいて前記ゲイン調整に用いるゲイン調整値を算出する制御部と、

を備える請求項 1 に記載のラウドネス調整装置。

【請求項 3】

前記制御部は、前記第 1 二乗平均値が前記ターゲットラウドネス値よりも大きく、前記第 2 二乗平均値が増加傾向にある場合、前記入力デジタル音声信号から抽出された前記音声データを抑圧するように前記ゲイン調整値を調整する抑圧処理を行い、

前記第 1 二乗平均値が前記ターゲットラウドネス値よりも大きく、前記第 2 二乗平均値が減少傾向にある場合、前記抑圧処理を中断し、

上記以外の場合には、前記抑圧処理は行わない、

ことを特徴とする請求項 2 に記載のラウドネス調整装置。

【請求項 4】

前記制御部は、

外部からの制御情報の入力無く、前記入力デジタル音声信号の音声モードの変更がなく、前記抑圧処理の実行中であり、かつ前記第 1 二乗平均値が開放閾値以下である場合、所定時間で前記入力デジタル音声信号から抽出された前記音声データの開放が完了するように前記ゲイン調整値を調整する開放処理を行い、

外部からの制御情報の入力無く、前記入力デジタル音声信号の音声モードの変更がなく、前記抑圧処理の実行中であり、かつ前記第 1 二乗平均値が前記開放閾値以下ではない場合、前記開放処理を行わない、

ことを特徴とする請求項 3 に記載のラウドネス調整装置。

【請求項 5】

前記制御部は、

前記抑圧処理の実行中であり、かつ外部からの制御情報の入力及び前記入力デジタル音声信号の音声モードの変更のいずれか一方が生じた場合、前記所定時間よりも短い時間で前記入力デジタル音声信号から抽出された前記音声データの開放が完了するように前記ゲイン調整値を調整する、

ことを特徴とする請求項 4 に記載のラウドネス調整装置。

【請求項 6】

前記入力デジタル音声信号から前記音声データを抽出するデコーダを更に備え、前記デコーダは、前記入力デジタル音声信号から付加情報を抽出し、前記付加情報内にゲイン調整を行わないことを指示する情報が含まれている場合、抽出した前記音声データを前記ゲイン調整部に供給せず他の任意の出力装置に供給し、

それ以外の場合には前記ゲイン調整部に抽出した前記音声データを供給する、

ことを特徴とする請求項 4 または請求項 5 に記載のラウドネス調整装置。

【請求項 7】

前記デコーダが抽出した前記音声データを解析し、無音状態が一定時間続いた場合に前記制御部に通知を行う無音検出部を更に備え、

前記制御部は、前記抑圧処理の実行中であり、かつ前記無音検出部からの通知を受信した場合、前記所定時間よりも短い時間で前記音声データの開放が完了するように前記ゲイン調整値を調整する、

ことを特徴とする請求項 6 に記載のラウドネス調整装置。

【請求項 8】

請求項 1 乃至請求項 7 のいずれか 1 項に記載のラウドネス調整装置を内蔵した映像 / 音声処理装置。

【請求項 9】

入力デジタル音声信号のラウドネス調整を行うラウドネス調整方法であって、

前記入力デジタル音声信号から抽出された音声データに対し、ゲイン調整を行うゲイン

10

20

30

40

50

調整ステップと、

前記ゲイン調整ステップにおけるゲイン調整済みの音声データを特性フィルタ処理したフィルタ済み音声データを生成する第1フィルタステップと、

前記第1フィルタステップにて生成したフィルタ済み音声データが第1閾値以下の場合、前記ゲイン調整ステップにおいてゲイン調整済みの音声データを増幅して出力するブースト処理ステップと、

前記ブースト処理ステップにて出力された音声データを特性フィルタ処理したフィルタ済み音声データを生成する第2フィルタステップと、

前記第2フィルタステップにて生成したフィルタ済み音声データが第2閾値以上の場合、前記ブースト処理ステップにて増幅した音声データに対してリミッタコンプレッサ処理を行って出力するリミッタコンプレッサステップと、を備えるラウドネス調整方法。

10

【請求項10】

入力デジタル音声信号のラウドネス調整をコンピュータに実行させるプログラムであって、

コンピュータに、

前記入力デジタル音声信号から抽出された音声データに対し、ゲイン調整を行うゲイン調整ステップと、

前記ゲイン調整ステップにおけるゲイン調整済みの音声データを特性フィルタ処理したフィルタ済み音声データを生成する第1フィルタステップと、

前記第1フィルタステップにて生成したフィルタ済み音声データが第1閾値以下の場合、前記ゲイン調整ステップにおいてゲイン調整済みの音声データを増幅して出力するブースト処理ステップと、

20

前記ブースト処理ステップにて出力された音声データを特性フィルタ処理したフィルタ済み音声データを生成する第2フィルタステップと、

前記第2フィルタステップにて生成したフィルタ済み音声データが第2閾値以上の場合、前記ブースト処理ステップにて増幅した音声データに対してリミッタコンプレッサ処理を行って出力するリミッタコンプレッサステップと、

を実行させるプログラム。

【発明の詳細な説明】

【技術分野】

30

【0001】

本発明はラウドネス調整装置、ラウドネス調整方法、及びプログラムに関する。

【背景技術】

【0002】

地上波デジタルテレビ放送の本格運用が開始されている。デジタルテレビ放送は、コンテンツの作成者が生成した音声をそのまま視聴者の視聴する音声とすることができる。しかしその反面、作成者がダイナミックレンジの広いコンテンツを作成した場合、コンテンツ内での音量差が大きくなる。視聴者が生活雑音の中でこのようなコンテンツをデジタルテレビ等で視聴すると、小さな音声が聞きづらいと感じる場合がある。そこで、視聴者がデジタルテレビ等の音量をリモートコントローラ等で大きくした場合、シーンチェンジ等が生じた場合に急に大きな音声が出力されてしまい、不快に感じる場合がある。また、このような場合にはユーザが主導で音量を再調整する必要が生じる。

40

【0003】

このような状況を鑑みて、音声信号の音量に関する世界的な標準規格の策定が進められている。この標準規格は、番組・素材単位での平均ラウドネス値を計算し、ターゲットラウドネス値に合わせた素材を生成・運用することを目的としている。なお、ラウドネスとは、人間が感じる音の大きさ（音の感覚量）である。ラウドネス値とは、デジタル録音レベルからラウドネス測定アルゴリズムに基づいて算出したラウドネスの計算値である。平均ラウドネス値とは、任意の測定区間のラウドネス値である。ターゲットラウドネス値とは、番組の聴取レベルを適切に保つために目標とする番組の平均のラウドネス値である。

50

詳細は、非特許文献 1 を参照されたい。

【 0 0 0 4 】

デジタルテレビ放送には、スポーツ中継や記者会見中継等の生放送番組があり、これらの番組に対しては基本的に編集処理を行うことが出来ない。そのため、前述の標準規格に規定された音量感の調整処理をリアルタイムで行う装置を提供する必要がある。

【先行技術文献】

【特許文献】

【 0 0 0 5 】

【特許文献 1】特開 2 0 0 6 - 9 3 9 0 8 号公報

【非特許文献】

10

【 0 0 0 6 】

【非特許文献 1】社団法人 電波産業会、デジタルテレビ放送番組におけるラウドネス運用規定 技術資料 (A R I B T R - B 3 2 1 . 0 版)、平成 2 3 年 3 月 2 8 日策定

【発明の概要】

【発明が解決しようとする課題】

【 0 0 0 7 】

上述したように、生放送番組等を考慮した場合、音量感の調整処理をリアルタイムで行うことが望ましい。しかしながら、前述の標準規格 (非特許文献 1) に記載された方式では、ラウドネス値が移動二乗平均値であるため、すなわち任意の区間 (時間) を待った後に値を算出するため、音量感の調整処理に一定の遅れが生じてしまう。生放送番組等ではない場合 (リアルタイム性が求められない場合)、平均ラウドネス値の算出時間を考慮して音声信号を遅延させる手法を用いることができる。しかしながら、生放送番組等である場合 (リアルタイム性が求められる場合)、このような音声信号を遅延させる手法を用いることはできない。

20

【 0 0 0 8 】

すなわち、前述の標準規格 (非特許文献 1) に記載された一般的な方式では、リアルタイム性を担保しつつ、音量感の調整処理を行うことが困難であるという問題が生じていた。

【 0 0 0 9 】

本発明は、上述した問題点を鑑みてなされたものであり、リアルタイム性を担保しつつ、音量感の調整処理を行うことができるラウドネス調整装置、ラウドネス調整方法、及びプログラムを提供することを主たる目的とする。

30

【課題を解決するための手段】

【 0 0 1 0 】

本発明にかかるラウドネス調整装置の一態様は、
入力デジタル音声信号のラウドネス調整を行うラウドネス調整装置であって、
前記入力デジタル音声信号から抽出された音声データに対し、ゲイン調整を行うゲイン調整部と、

前記ゲイン調整部によるゲイン調整済みの音声データを特性フィルタ処理したフィルタ済み音声データを生成する第 1 フィルタ部と、

40

前記第 1 フィルタ部の生成したフィルタ済み音声データが第 1 閾値以下の場合、前記ゲイン調整部によるゲイン調整済みの音声データを増幅して出力するブースト処理部と、

前記ブースト処理部が出力した音声データを特性フィルタ処理したフィルタ済み音声データを生成する第 2 フィルタ部と、

前記第 2 フィルタ部の生成したフィルタ済み音声データが第 2 閾値以上の場合、前記ブースト処理部が増幅した音声データに対してリミッタコンプレッサ処理を行って出力するリミッタコンプレッサ部と、を備えるものである。

【 0 0 1 1 】

本発明にかかるラウドネス調整方法の一態様は、

入力デジタル音声信号のラウドネス調整を行うラウドネス調整方法であって、

50

前記入力デジタル音声信号から抽出された音声データに対し、ゲイン調整を行うゲイン調整ステップと、

前記ゲイン調整ステップにおけるゲイン調整済みの音声データを特性フィルタ処理したフィルタ済み音声データを生成する第1フィルタステップと、

前記第1フィルタステップにて生成したフィルタ済み音声データが第1閾値以下の場合、前記ゲイン調整ステップにおいてゲイン調整済みの音声データを増幅して出力するブースト処理ステップと、

前記ブースト処理ステップにて出力された音声データを特性フィルタ処理したフィルタ済み音声データを生成する第2フィルタステップと、

前記第2フィルタステップにて生成したフィルタ済み音声データが第2閾値以上の場合、前記ブースト処理部が増幅した音声データに対してリミッタコンプレッサ処理を行って出力するリミッタコンプレッサステップと、を備えるものである。

10

【0012】

本発明にかかるプログラムの一態様は、

入力デジタル音声信号のラウドネス調整をコンピュータに実行させるプログラムであって、

コンピュータに、

前記入力デジタル音声信号から抽出された音声データに対し、ゲイン調整を行うゲイン調整ステップと、

前記ゲイン調整ステップにおけるゲイン調整済みの音声データを特性フィルタ処理したフィルタ済み音声データを生成する第1フィルタステップと、

20

前記第1フィルタステップにて生成したフィルタ済み音声データが第1閾値以下の場合、前記ゲイン調整ステップにおいてゲイン調整済みの音声データを増幅して出力するブースト処理ステップと、

前記ブースト処理ステップにて出力された音声データを特性フィルタ処理したフィルタ済み音声データを生成する第2フィルタステップと、

前記第2フィルタステップにて生成したフィルタ済み音声データが第2閾値以上の場合、前記ブースト処理部が増幅した音声データに対してリミッタコンプレッサ処理を行って出力するリミッタコンプレッサステップと、

を実行させる、ものである。

30

【発明の効果】

【0013】

本発明によれば、リアルタイム性を担保しつつ、音量感の調整処理を行うことができるラウドネス調整装置、ラウドネス調整方法、及びプログラムを提供することができる。

【図面の簡単な説明】

【0014】

【図1】実施の形態1にかかるラウドネス調整装置1の全体構成を示すブロック図である。

【図2】一般的な抑圧処理（ATK処理）を示す概念図である。

【図3】一般的な開放処理（REL処理）を示す概念図である。

40

【図4】実施の形態1にかかる制御部19の抑圧処理（ATK処理）を示すフローチャートである。

【図5】実施の形態1にかかる制御部19の開放処理（REL処理）を示すフローチャートである。

【図6】実施の形態1にかかるラウドネス調整装置1のラウドネス調整を示す概念図である。

【図7】実施の形態1にかかる制御部19によるゲイン調整を示す概念図である。

【図8】実施の形態1にかかる制御部19によるゲイン調整を示す概念図である。

【図9】実施の形態1にかかる制御部19による開放処理（REL処理）を示す概念図である。

50

【図 1 0】実施の形態 2 にかかるラウドネス調整装置 1 の全体構成を示すブロック図である。

【図 1 1】実施の形態 3 にかかるラウドネス調整装置 1 の全体構成を示すブロック図である。

【図 1 2】本発明にかかるラウドネス調整装置 1 の全体構成を示すブロック図である。

【発明を実施するための形態】

【0015】

<実施の形態 1>

以下、図面を参照して本発明の実施の形態について説明する。図 1 は、本実施の形態にかかるラウドネス調整装置 1 の構成を示すブロック図である。ラウドネス調整装置 1 は、例えばデジタルテレビ受像機に内蔵される装置である。また、ラウドネス調整装置 1 は、インターネット放送やラジオ放送等の音声を扱う装置（例えばコンピュータ、ラジオ装置、映像/音声処理装置とも記載する。）に内蔵されても良い。

10

【0016】

ラウドネス調整装置 1 は、デコーダ 1 1 と、ゲイン調整部 1 2 と、音声調整部 1 3 と、セクタ 1 8 と、制御部 1 9 と、移動二乗平均算出部 2 0 と、移動二乗平均算出部 2 1 と、K フィルタ 2 2 と、エンコーダ 2 3 と、を備える。音声調整部 1 3 は、K フィルタ 1 4 と、音声データブースト部 1 5 と、K フィルタ 1 6 と、リミッタコンプレッサ部 1 7 と、を備える。

【0017】

デコーダ 1 1 には、入力デジタル音声信号が供給される。入力デジタル音声信号は、デジタルテレビ受像機のアンテナ等から入力されたデータである。入力音声データは、例えば AES / EBU フォーマット信号、エンベディッドオーディオ信号等である。デコーダ 1 1 は、入力デジタル音声信号をデコードし、デコード済みの音声データをゲイン調整部 1 2 に供給し、デコードにより得られた付加情報（音声モード情報、切替情報、ラウドネス運用情報等）をセクタ 1 8 に供給する。

20

【0018】

なお、デコーダ 1 1 は、デコードにより得られた付加情報内にラウドネス調整を行わないことを指示する情報が含まれる場合、デコード済みの音声データをエンコーダ 2 3 または任意の処理部にそのまま供給し、ゲイン調整部 1 2 への音声データの供給は行わない。

30

【0019】

ゲイン調整部 1 2 は、制御部 1 9 から入力されるゲイン調整値を用いて、デコーダ 1 1 から入力された音声データのゲインを調整する。ゲイン調整部 1 2 は、ゲイン調整済みの音声データを K フィルタ 1 4、及び音声データブースト部 1 5 に供給する。

【0020】

音声調整部 1 3 は、後述する移動二乗平均の算出処理を待つことなく、入力された音声データのブースト処理及びリミッタコンプレッサ処理を行う。当該ブースト処理は、ラウドネス運用基準の周波数特性に適應した処理である。以下、音声調整部 1 3 内の各処理部の処理について説明する。

【0021】

K フィルタ 1 4 は、入力された音声データに対して K 特性（ITU-R BS. 1770 で規定されている聴感補正フィルタの特性）フィルタリング処理を行い、フィルタリング済みの音声データを音声データブースト部 1 5 に供給する。

40

【0022】

音声データブースト部 1 5 は、K フィルタ 1 4 から入力された音声データが予め定められた閾値以下の音声データであるか否かを判定する。当該判定は、例えばデシベル値（dB）の比較により行う。閾値以下であると判定した場合、音声データブースト部 1 5 は、ゲイン調整部 1 2 から入力された音声データを増幅する（ブースト処理を行う）。音声データブースト部 1 5 は、増幅済みの音声データを K フィルタ 1 6 及びリミッタコンプレッサ部 1 7 に供給する。なお、閾値より大きいと判定した場合、音声データブースト部 1 5

50

は、ゲイン調整部 12 から入力された音声データをそのまま K フィルタ 16 及びリミッタコンプレッサ部 17 に供給する。

【0023】

K フィルタ 16 は、音声データブースト部 15 から入力された音声データに対して K 特性 (ITU-R BS.1770 で規定されている聴感補正フィルタの特性) フィルタリング処理を行い、フィルタリング済みの音声データをリミッタコンプレッサ部 17 に供給する。

【0024】

リミッタコンプレッサ部 17 は、K フィルタ 16 から入力された音声データが予め定められた閾値以上の音声データであるか否かを判定する。当該判定は、例えばデシベル値 (dB) の比較により行う。閾値以上であると判定した場合、リミッタコンプレッサ部 17 は、音声データブースト部 15 から入力された音声データに対してリミッタコンプレッサ処理を行い、処理後の音声データを K フィルタ 22 及びエンコーダ 23 に供給する。ここで、リミッタコンプレッサ処理は、一般的な放送音声の処理について行われる任意の音声レベルの適正化処理であれば良い。なお、閾値以上ではないと判定した場合、リミッタコンプレッサ部 17 は、音声データブースト部 15 から入力された音声データをそのまま K フィルタ 22 及びエンコーダ 23 に供給する。

【0025】

エンコーダ 23 は、リミッタコンプレッサ部 17 またはデコーダ 11 から入力された音声データを所定のフォーマットにエンコードし、エンコード済みの音声データを任意の処理部 (例えばデジタルテレビ受像機のスピーカ等) に供給する。

【0026】

K フィルタ 22 は、リミッタコンプレッサ部 17 から入力された音声データに対して K 特性 (ITU-R BS.1770 で規定されている聴感補正フィルタの特性) フィルタリング処理を行い、フィルタリング済みの音声データを移動二乗平均算出部 20 及び移動二乗平均算出部 21 に供給する。

【0027】

移動二乗平均算出部 20 及び移動二乗平均算出部 21 は、K フィルタ 22 から入力された音声データの信号パワーを求めるための二乗平均値を算出する。ここで、移動二乗平均算出部 20 と移動二乗平均算出部 21 は、異なる時間の移動二乗平均処理を行う。以下の説明では、移動二乗平均算出部 20 は、非特許文献 1 に規定された最小ブロックである 100ms (0.1秒) 単位で移動二乗平均値を算出し、算出値 S (短い時間での移動二乗平均値) を制御部 19 に供給する。一方、移動二乗平均算出部 21 は、EBU R-128 に規定された時間である 3s (3秒) 単位で移動二乗平均を算出し、算出値 L (長い時間での移動二乗平均値) を制御部 19 に供給する。なお、移動二乗平均の算出方法の詳細は、非特許文献 1 を参照されたい。また、上述の移動二乗平均の単位となる時間 (0.1秒、3秒) はあくまで一例であり、これに必ずしも限られない。

【0028】

セレクタ 18 には、外部からの制御情報と、デコーダ 11 からの付加情報と、が入力される。制御情報とは、後述する高速な開放処理 (REL 処理) を指示する指示情報を含むことができる情報である。付加情報には、上述のように音声モードの切替 (例えば、CM (コマーシャル) 音声と番組音声の切替) に関する情報が含まれる。セレクタ 18 は、両入力のうち制御部 19 の動作を変更する情報を適宜選択し、選択した情報を制御部 19 に供給する。セレクタ 18 の動作詳細は、図 5 を参照して後述する。

【0029】

制御部 19 には、セレクタ 18 の選択した情報と、ターゲットラウドネス値と、移動二乗平均算出部 20 の算出値 S と、移動二乗平均算出部 21 の算出値 L と、が入力される。制御部 19 は、これらの情報を基にゲイン調整値を算出し、算出したゲイン調整値をゲイン調整部 12 に供給する。

【0030】

10

20

30

40

50

ターゲットラウドネス値とは、前述のように出力音量の目的値となる平均ラウドネス値である。制御部 19 は、ゲイン調整値に関する抑圧処理（ATK（アタック）処理）、または開放処理（REL（リリース）処理）を行うことにより、ゲイン調整部 12 における音声データのゲイン調整を制御する。抑圧処理（ATK 処理）とは、規定値レベル（dB）以上の音声データが入力された場合には、音声データを抑圧するようなゲイン調整値を定めることを意味する。開放処理（REL 処理）とは、規定値レベル（dB）以下の音声データが入力された場合には、音声データを開放するようなゲイン調整値を定めることを意味する。制御部 19 は、設定したゲイン調整値をゲイン調整部 12 に供給する。

【0031】

図 2 及び図 3 を参照し、一般的な抑圧処理（ATK 処理）及び開放処理（REL 処理）の概要を説明する。ある音声データが規定値レベル（ t_h ）である場合、図 2 に示すように音声データを抑圧する処理（ATK 処理）を行う。例えば音声データのデシベル値（dB）が規定レベル（ t_h ）より大きい場合、音声データのレベルを弱めるようなゲイン調整値を設定する。この時、抑圧量がある一定量以上（通常は約 90% 以上）になるまでの時間を抑圧時間（アタック時間）という。

【0032】

また、図 3 に示すように、抑圧状態から規定値レベル（ t_h ）以下の音声データとなった場合、抑圧状態から開放する処理（REL 処理）を行う。この時、開放される量がある一定量以上（通常は開放量が約 90% 以上）になるまでの時間を開放時間（リリース時間）という。なお、抑圧処理（ATK 処理）及び開放処理（REL 処理）の詳細は、例えば特許文献 1 を参照されたい。

【0033】

以下、制御部 19 の詳細な処理を図 4 及び図 5 を参照して説明する。はじめに、図 4 を参照して制御部 19 による抑圧処理（ATK 処理）について説明する。制御部 19 は、算出値 S （移動二乗平均算出部 20 の出力値）と、ターゲットラウドネス値（ L_{t_h} ）と、を比較する（ST11）。算出値 S がターゲットラウドネス値（ L_{t_h} ）よりも大きくない場合（ST11: No）、制御部 19 は、処理を行わない。算出値 S がターゲットラウドネス値（ L_{t_h} ）よりも大きい場合（ST11: Yes）、制御部 19 は、抑圧処理（ATK 処理）の実行中であるか否かを判定する（ST12）。

【0034】

抑圧処理（ATK 処理）の実行中ではない場合（ST12: No）、制御部 19 は、算出値 L （移動二乗平均算出部 21 の出力値）が増加傾向にあるか否かを判定する（ST13）。算出値 L が増加傾向にはない場合（ST13: No）、制御部 19 は、処理を行わない。一方、算出値 L が増加傾向にある場合（ST13: Yes）、制御部 19 は、抑圧処理（ATK 処理）を実行する（ST14）。ゲイン調整値は、所望の抑圧時間と、入力音声データの dB 値に応じて調整される

【0035】

抑圧処理（ATK 処理）の実行中である場合（ST12: Yes）、制御部 19 は、算出値 L （移動二乗平均算出部 21 の出力値）が増加傾向にあるか否かを判定する（ST15）。算出値 L が増加傾向にはない場合（ST15: No）、制御部 19 は、抑圧処理（ATK 処理）を中断する（ST16）。一方、算出値 L が増加傾向にある場合（ST15: Yes）、制御部 19 は、抑圧処理（ATK 処理）を実行する（ST14）。

【0036】

続いて、抑圧処理（ATK 処理）からの開放処理（REL）処理の流れを図 5 を参照して説明する。セレクタ 18 は、音声モードが変更された場合に、変更されたことを示す信号を出力する。制御部 19 は、セレクタ 18 からの出力信号により音声モードが切り替わったか否かを判定する（ST21）。音声モードが切り替わっていない場合（ST21: No）、制御部 19 は、開放用の閾値（開放閾値（ REL_{start_th} ）とも記載する。）と、算出値 S と、を比較する（ST22）。開放閾値は、ターゲットラウドネス値よりも低い値を持つ。

10

20

30

40

50

【 0 0 3 7 】

算出値 S が開放閾値以下である場合 (S T 2 2 : Y e s)、制御部 1 9 は、通常の開放処理となるようにゲイン調整値を設定する (S T 2 3)。すなわち、制御部 1 9 は、所定の開放時間 (T 1) となるようにゲイン調整値を設定する (S T 2 3)。一方、算出値 S が開放閾値以下ではない場合 (S T 2 2 : N o)、制御部 1 9 は、処理を行わない。

【 0 0 3 8 】

抑圧処理 (A T K 処理) の実行中であって音声モードの変更が生じた場合 (S T 2 1 : Y e s)、制御部 1 9 は、高速の開放処理となるようにゲイン調整値を設定する (S T 2 4)。すなわち、制御部 1 9 は、所定の開放時間 (T 1) よりも短い開放時間 (T 2) となるようにゲイン調整値を設定する (S T 2 4)。換言すると、制御部 1 9 は、通常の開放処理 (R E L) よりも開放量が大きくなるようにゲイン調整値を設定する。ゲイン調整値は、所望の開放時間 (T 1 または T 2) と、入力音声データの d B 値に応じて調整される。

10

【 0 0 3 9 】

なお、上述の判定において、高速な開放処理 (R E L) 処理を指示する制御情報がセクタ 1 8 から入力された場合、制御部 1 9 は、音声モードの変更時と同様の処理 (S T 2 1 : Y e s) を行う。

【 0 0 4 0 】

続いて、各処理部の動作と出力音声データとの関係を図 6 ~ 図 9 を参照して説明する。図 6 は、音声調整部 1 3 (K フィルタ 1 4、音声データブースト部 1 5、K フィルタ 1 6、リミッタコンプレッサ 1 7) の制御を示すグラフである。

20

【 0 0 4 1 】

前述のように非特許文献 1 に記載されたラウドネス値の調整手法では、移動二乗平均を用いるため、リアルタイムでのラウドネス調整が困難である。そのため、図 6 に示すように、瞬間的に大きな音量の音声データが入力された場合、一般的なラウドネス値の調整手法では大きな音量差 (図中の (1)) が生じてしまう。これにより、視聴者に不快感を与えてしまう。

【 0 0 4 2 】

一方、本実施の形態にかかるラウドネス調整装置 1 は、音声調整部 1 3 を備える。そして、当該音声調整部 1 3 は、移動二乗平均値を用いることなく、入力音声データの抑圧処理 (リミッタコンプレッサ処理) を行う。そのため、瞬間的に大きな音量の音声データが入力された場合であっても、音量差を抑えることが出来る (図中の (2))。これにより、視聴者にとって違和感の無い、すなわち不快感を与えることの無い音声データの提供を実現することができる。

30

【 0 0 4 3 】

図 7 及び図 8 は、制御部 1 9 のゲイン調整値の調整処理を示すグラフである。図 7 は、各タイミング (T M 1 ~ T M 5) における算出値 S と各閾値 (ターゲットラウドネス値及び開放閾値)、及び算出値 L の増減を示す図である。図 8 は、制御部 1 9 によるゲイン調整値の調整を示す図である。図 7 及び図 8 には、同一の入力音声データ、算出値 S、及び算出値 L が表示されている。

40

【 0 0 4 4 】

タイミング T M 1 において、算出値 S は、ターゲットラウドネス値以上であり、かつ算出値 L は増加傾向にある (図 7 T M 1)。そのため、制御部 1 9 は、タイミング T M 1 から抑圧処理 (A T K 処理) を開始する (図 8 T M 1)。タイミング T M 1 ~ T M 2 の間、制御部 1 9 は、抑圧処理 (A T K 処理) を実行する。

【 0 0 4 5 】

タイミング T M 2 において、算出値 L は減少傾向にある (図 7 T M 2)。そのため、制御部 1 9 は、タイミング T M 2 から抑圧処理 (A T K 処理) の実行を保留する (図 8 T M 2)。タイミング T M 2 ~ T M 3 の間、制御部 1 9 は、抑圧処理 (A T K 処理) を行わない。

50

【 0 0 4 6 】

タイミング T M 3 において、算出値 S は、ターゲットラウドネス値以上であり、かつ算出値 L は増加傾向となる（図 7 T M 3）。そのため、制御部 1 9 は、タイミング T M 1 から抑圧処理（A T K 処理）を再開する（図 8 T M 3）。

【 0 0 4 7 】

タイミング T M 4 において、算出値 S は、開放閾値以下となる（図 7 T M 4）。そのため、制御部 1 9 は、タイミング T M 4 から開放処理（R E L 処理）を開始する（図 8 T M 4）。

【 0 0 4 8 】

タイミング T M 1 において、算出値 S は、ターゲットラウドネス値以上となり、かつ算出値 L は増加傾向にある（図 7 T M 5）。そのため、制御部 1 9 は、タイミング T M 5 から抑圧処理（A T K 処理）を開始する（図 8 T M 5）。以上のように、制御部 1 9 は、算出値 S 及び算出値 L を用いてゲイン調整値の調整を行う。

【 0 0 4 9 】

図 9 は、制御部 1 9 による開放処理を示す図である。図 9 は、一般的な開放処理（ターゲットラウドネス値との比較）を用いて生成した音声データと、開放閾値を用いた開放処理を用いて生成した音声データと、を示すグラフである。図示するように、一般的な開放処理では、音声データがターゲットラウドネス値を下回った場合、即座に開放処理を行う。そのため、ターゲットラウドネス値付近での抑圧、及び開放が交互に繰り返される。これにより、元の音声データに対して振幅の変化が揺らいだ状態になる。換言すると、元の音声データの抑揚を正確に表した音声データを生成することができない。

【 0 0 5 0 】

一方、ターゲットラウドネス値よりも小さい値である開放閾値を用いた開放処理を用いて生成した音声データは、ターゲットラウドネス値付近において抑圧と開放が交互に繰り返される事象が無くなる。これにより、元の音声データに対して振幅の変化が揺らいだ状態を回避することができる。よって、図 9 に示すように、元の音声データの抑揚を正確に表した音声データを生成することができる。

【 0 0 5 1 】

続いて、本実施の形態にかかるラウドネス調整装置 1 の効果について改めて説明する。上述のように、音声調整部 1 3 は、移動二乗平均の算出値を用いることなく、特性フィルタ処理に従ったブースト処理及びリミッタコンプレッサ処理を行う。これらの処理は、移動二乗平均の算出とは異なり、即時的に処理が可能である。当該構成により、極端に大きな音声データを抑圧することができ、かつ、極端に大きな音声データに後続する通常の音量の音声データを不要に抑圧することが無くなる。すなわち、リアルタイム性を担保しつつ、音量感の調整を適切に行うことができる。換言すると、リアルタイム性を保ちつつ、音質を一定に保つ（違和感の無い音声を生成する）ラウドネス調整を実現することが可能となる。

【 0 0 5 2 】

さらに、制御部 1 9 は、図 4 に示すように、算出値 L の増加時にのみ抑圧処理（A T K 処理）を行い、算出値 L の減少時には抑圧処理（A T K 処理）の保留を行うことにより、過度の抑圧処理（A T K 処理）を行わない。これにより、抑圧が行われすぎた音声データを出力することを回避できる。

【 0 0 5 3 】

さらにまた、抑圧処理（A T K 処理）の実行中から開放処理（R E L 処理）に移行する際にも、音声データのレベルがターゲットラウドネス値よりも小さい値である開放閾値以下となった場合に開放処理を行う。これにより、不要な開放処理を行うことなく、生成する音声データの抑揚を基の音声データの抑揚と近づけることができる。

【 0 0 5 4 】

制御部 1 9 は、抑圧処理（A T K 処理）の実行中に音声モードの変更や外部からの指示が生じた場合に、通常の開放処理（R E L 処理）よりも開放時間が短くなるようにゲイン

10

20

30

40

50

調整値を設定する（高速開放処理を実行する）。デジタルテレビに関する音声信号は、番組コンテンツとコマーシャルが混在し、両者の音量感が全く異なる場合がある。一般的なラウドネス調整の手法では、番組コンテンツからコマーシャルに切り替わる場合、番組コンテンツにかかる音声データの解析（移動二乗平均の算出）に基づく抑圧処理（ATK処理）が行われてしまう恐れがある。すなわち、番組コンテンツに適した抑圧処理（ATK処理）をコマーシャルに対しても行ってしまふ。しかし、本実施形態にかかるラウドネス調整装置では、上述の高速開放処理を行うことにより、抑圧量を即座にキャンセルすることができ、コンテンツに応じたラウドネス調整を瞬時に行うことが可能になる。

【0055】

さらにまた、デコーダ11は、デコードにより得られた付加情報内にラウドネス調整を行わないことを指示する情報が含まれる場合、デコード済みの音声データをエンコーダ23または任意の処理部にそのまま供給する。デコーダ11に入力される入力デジタル音声信号は、生放送等のような番組にかかるものの他に、予め非特許文献1に規定のラウドネス調整が行われたコンテンツにかかるものがある。後者の場合、既に制作者の意図通りのラウドネス調整が行われているため、改めてラウドネス調整を行う必要がない。デコーダ11が後者の信号の場合にラウドネス調整を行うことなくそのまま出力を行うため、制作者の意図を変えることなく音声データを供給することができる。

【0056】

<実施の形態2>

本実施の形態にかかるラウドネス調整装置1は、多チャンネル（サラウンドまたはそれ以上のチャンネル数）の音声データを処理できることを特徴とする。本実施の形態にかかるラウドネス調整装置1について、実施の形態1と異なる点を以下に説明する。

【0057】

図10は、本実施の形態にかかるラウドネス調整装置1の構成を示すブロック図である。ラウドネス調整装置1には、多チャンネル（サラウンドまたはそれ以上のチャンネル数）に関する入力デジタル音声信号が入力される。

【0058】

ラウドネス調整装置1は、入力デジタル音声信号内の各チャンネルに応じた数の音声調整部13（13-1～13-n、nは2以上の整数であり以下の記載でも同様である。）、移動二乗平均算出部20（20-1～20-n）、移動二乗平均算出部21（21-1～21-n）、Kフィルタ22（22-1～22-n）、ゲイン乗算部24（24-1～24-n）、二乗平均積算部25-1、及び二乗平均積算部25-2を備える。

【0059】

デコーダ11は、入力デジタル音声信号をデコードし、各チャンネルに対応する音声データを音声調整部13-1～13-nに供給する。実施の形態1と同様の符号を付した処理部の処理は、チャンネル毎に実施の形態1と同様の処理を行うため、その詳細な説明は省略する。

【0060】

ゲイン乗算部24-1は、算出値S及び算出値Lに対して、処理対象のチャンネルに対応付けられたゲインを乗算する。そして、ゲイン乗算部24-1は、算出値Sの乗算結果を二乗平均積算部25-1に供給し、算出値Lの乗算結果を二乗平均積算部25-2に供給する。

【0061】

二乗平均積算部25-1は、各チャンネルの算出値S（最小ブロックである100ms（0.1秒）単位で移動二乗平均値）を積算（累積加算）する。二乗平均積算部25-1は、積算結果を制御部19に供給する。

【0062】

二乗平均積算部25-2は、各チャンネルの算出値L（（3秒）単位で移動二乗平均）を積算（累積加算）する。二乗平均積算部25-1は、積算結果を制御部19に供給する。

【0063】

10

20

30

40

50

制御部 19 は、二乗平均積算部 25 - 1 及び二乗平均積算部 25 - 2 から入力された積算値を基に、各ゲイン調整部 12 - 1 ~ 12 - n に供給するゲイン調整部を算出し、供給する。

【0064】

続いて、本実施の形態にかかるラウドネス調整装置 1 の効果について説明する。ラウドネス調整装置 1 は、上述の構成により、多チャンネルに関する入力デジタル音声信号が入力された場合であってもリアルタイム性を担保しつつ、音量感の調整処理を行うことができる。

【0065】

<実施の形態 3>

本実施の形態にかかるラウドネス調整装置 1 は、音声データの無音判定を行う処理部を有することを特徴とする。本実施の形態にかかるラウドネス調整装置 1 について、実施の形態 1 と異なる点を以下に説明する。

【0066】

図 11 は、本実施の形態にかかるラウドネス調整装置 1 の構成を示すブロック図である。ラウドネス調整装置 1 は、実施の形態 1 の構成（図 1）に加えて無音検出部 26 を更に備える。

【0067】

デコーダ 11 は、デコード済みの音声データをゲイン調整部 12 と、無音検出部 26 とに適宜供給する。無音検出部 26 は、音声データのレベル変動、すなわちデシベル（dB）変動の有無を検知し、レベル変動が一定時間生じない場合、あるいは、素材切替時に音声データをミュート（無音状態）とする時に生成される音声データが全て "0" となる場合、あるいは特定の音声データとなる状態が一定時間経過する場合、のいずれかを検出した際に制御部 19 に通知を行う。

【0068】

制御部 19 は、抑圧処理（ATK 処理）を行っている際に無音検出部 26 からの通知を受け付けた場合、上述の高速開放処理（REL 処理、図 5ST24）を実行する。

【0069】

続いて、本実施の形態にかかるラウドネス調整装置 1 の効果について説明する。一般的に、テレビ番組コンテンツからコマーシャルに切り替わる場合、一定時間の無音区間が存在する。ここで、入力音声デジタル信号内に抽出した付加情報内に、コンテンツの切り替わりにかかる情報が含まれない場合がある。本実施の形態にかかるラウドネス調整装置 1 は、無音検出部 26 により無音区間を検出している。そして、制御部 19 は、付加情報に加えて無音区間の有無に応じて高速開放処理（REL 処理）を実行する。これにより、コンテンツが切り替わる前に行われていた抑圧処理（ATK 処理）を即座にキャンセルすることができ、コンテンツに応じたラウドネス調整を瞬時に行うことが可能になる。

【0070】

ここで、改めて本発明の概略を図 12 を参照して説明する。図 12 は、ラウドネス調整装置 1 の概略を示したブロック図である。ラウドネス調整装置 1 は、ゲイン調整部 12 と、音声調整部 13 と、を備える。音声調整部 13 は、K フィルタ 14 と、音声データブースト部 15 と、K フィルタ 16 と、リミッタコンプレッサ部 17 と、を備える。

【0071】

ゲイン調整部 12 には、音声データが入力される。ここで、音声データとは、例えば AES / EBU フォーマット信号、エンベディッドオーディオ信号等をデコードすることにより得られるデータである。ゲイン調整部 12 は、所定のゲイン調整値を用いて入力された音声データのゲインを調整し、ゲイン調整済みの音声データを K フィルタ 14 及び音声データブースト部 15 に供給する。

【0072】

K フィルタ 14 は、入力された音声データに対して K 特性（フィルタリング処理を行い、フィルタリング済みの音声データを音声データブースト部 15 に供給する。

10

20

30

40

50

【 0 0 7 3 】

音声データブースト部 1 5 は、K フィルタ 1 4 から入力された音声データが予め定められた閾値以下の音声データであるか否かを判定する。閾値以下であると判定した場合、音声データブースト部 1 5 は、ゲイン調整部 1 2 から入力された音声データを増幅する（ブースト処理を行う）。音声データブースト部 1 5 は、増幅済みの音声データを K フィルタ 1 6 及びリミッタコンプレッサ部 1 7 に供給する。

【 0 0 7 4 】

K フィルタ 1 6 は、音声データブースト部 1 5 から入力された音声データに対して K 特性フィルタリング処理を行い、フィルタリング済みの音声データをリミッタコンプレッサ部 1 7 に供給する。

10

【 0 0 7 5 】

リミッタコンプレッサ部 1 7 は、K フィルタ 1 6 から入力された音声データが予め定められた閾値以上の音声データであるか否かを判定する。閾値以上であると判定した場合、リミッタコンプレッサ部 1 7 は、音声データブースト部 1 5 から入力された音声データに対してリミッタコンプレッサ処理を行い、処理後の音声データを出力する。

【 0 0 7 6 】

上記のとおり、ラウドネス調整装置 1 は、移動二乗平均の算出値を用いることなく、特性フィルタ処理に従ったブースト処理及びリミッタコンプレッサ処理を行う。これらの処理は、移動二乗平均の算出とは異なり、即時的に処理が可能である。当該構成により、例えば図 6 に示すように極端に大きな音声データを抑圧することができ、かつ、極端に大きな音声データに後続する通常の音量の音声データを不要に抑圧することが無くなる。すなわち、リアルタイム性を担保しつつ、音量感の調整を適切に行うことができる。

20

【 0 0 7 7 】

以上、本発明者によってなされた発明を実施の形態に基づき具体的に説明したが、本発明は前記実施の形態に限定されるものではなく、その要旨を逸脱しない範囲で種々変更可能であることはいうまでもない。

【 0 0 7 8 】

上述のデコーダ 1 1、ゲイン調整部 1 2、音声調整部 1 3、制御部 1 9、移動二乗平均算出部 2 0、移動二乗平均算出部 2 1、K フィルタ 2 2、エンコーダ 2 3 の各処理は、任意のコンピュータ内で動作するプログラムとして実現することが可能である。

30

【 0 0 7 9 】

プログラムは、様々なタイプの非一時的なコンピュータ可読媒体（non-transitory computer readable medium）を用いて格納され、コンピュータに供給することができる。非一時的なコンピュータ可読媒体は、様々なタイプの実体のある記録媒体（tangible storage medium）を含む。非一時的なコンピュータ可読媒体の例は、磁気記録媒体（例えばフレキシブルディスク、磁気テープ、ハードディスクドライブ）、光磁気記録媒体（例えば光磁気ディスク）、CD-ROM（Read Only Memory）、CD-R、CD-R/W、半導体メモリ（例えば、マスク ROM、PROM（Programmable ROM）、EPROM（Erasable PROM）、フラッシュ ROM、RAM（random access memory））を含む。また、プログラムは、様々なタイプの一時的なコンピュータ可読媒体（transitory computer readable medium）によってコンピュータに供給されてもよい。一時的なコンピュータ可読媒体の例は、電気信号、光信号、及び電磁波を含む。一時的なコンピュータ可読媒体は、電線及び光ファイバ等の有線通信路、又は無線通信路を介して、プログラムをコンピュータに供給できる。

40

【 0 0 8 0 】

上記実施形態の一部又は全部は、以下の付記のようにも記載されうるが、以下には限られない。

【 0 0 8 1 】

< 付記 1 >

50

入力デジタル音声信号のラウドネス調整を行うラウドネス調整装置であって、
前記入力デジタル音声信号から抽出された音声データに対し、ゲイン調整を行うゲイン調整部と、

前記ゲイン調整部によるゲイン調整済みの音声データを特性フィルタ処理したフィルタ済み音声データを生成する第1フィルタ部と、

前記第1フィルタ部の生成したフィルタ済み音声データが第1閾値以下の場合、前記ゲイン調整部によるゲイン調整済みの音声データを増幅して出力するブースト処理部と、

前記ブースト処理部が出力した音声データを特性フィルタ処理したフィルタ済み音声データを生成する第2フィルタ部と、

前記第2フィルタ部の生成したフィルタ済み音声データが第2閾値以上の場合、前記ブースト処理部が増幅した音声データに対してリミッタコンプレッサ処理を行って出力するリミッタコンプレッサ部と、を備えるラウドネス調整装置。

10

【0082】

<付記2>

前記リミッタコンプレッサ部が出力した音声データを特性フィルタ処理したフィルタ済み音声データを生成する第3フィルタ部と、

前記第3フィルタ部が生成したフィルタ済み音声データに対し、第1の時間区間を用いた移動二乗平均により第1二乗平均値を算出する第1二乗平均算出部と、

前記第3フィルタ部が生成したフィルタ済み音声データに対し、第1の時間区間より長い第2の時間区間を用いた移動二乗平均により第2二乗平均値を算出する第2二乗平均算出部と、

20

前記第1及び第2二乗平均値、及びターゲットラウドネス値に基づいて前記ゲイン調整に用いるゲイン調整値を算出する制御部と、

を備える付記1に記載のラウドネス調整装置。

【0083】

<付記3>

前記リミッタコンプレッサ部が出力した音声データを特性フィルタ処理したフィルタ済み音声データを生成する第3フィルタ部と、

前記第3フィルタ部が生成したフィルタ済み音声データに対し、第1の時間区間を用いた移動二乗平均により第1二乗平均値を算出する第1二乗平均算出部と、

30

前記第3フィルタ部が生成したフィルタ済み音声データに対し、第1の時間区間より長い第2の時間区間を用いた移動二乗平均により第2二乗平均値を算出する第2二乗平均算出部と、

前記第1及び第2二乗平均値、及びターゲットラウドネス値に基づいて前記ゲイン調整に用いるゲイン調整値を算出する制御部と、

を備える付記1に記載のラウドネス調整装置。

【0084】

<付記4>

前記制御部は、

外部からの制御情報の入力無く、前記入力デジタル音声信号の音声モードの変更がなく、前記抑圧処理の実行中であり、かつ前記第1二乗平均値が開放閾値以下である場合、所定時間で前記入力デジタル音声信号から抽出された前記音声データの開放が完了するように前記ゲイン調整値を調整する開放処理を行い、

40

外部からの制御情報の入力無く、前記入力デジタル音声信号の音声モードの変更がなく、前記抑圧処理の実行中であり、かつ前記第1二乗平均値が前記開放閾値以下ではない場合、前記開放処理を行わない、

ことを特徴とする付記3に記載のラウドネス調整装置。

【0085】

<付記5>

前記制御部は、

50

前記抑圧処理の実行中であり、かつ外部からの制御情報の入力及び前記入力デジタル音声信号の音声モードの変更のいずれか一方が生じた場合、前記所定時間よりも短い時間で前記入力デジタル音声信号から抽出された前記音声データの開放が完了するように前記ゲイン調整値を調整する、

ことを特徴とする付記 3 または付記 4 に記載のラウドネス調整装置。

【 0 0 8 6 】

< 付記 6 >

前記入力デジタル音声信号から前記音声データを抽出するデコーダを更に備え、前記デコーダは、前記入力デジタル音声信号から付加情報を抽出し、前記付加情報内にゲイン調整を行わないことを指示する情報が含まれている場合、抽出した前記音声データを前記ゲイン調整部に供給せず、他の任意の出力装置に供給し、

それ以外の場合には前記ゲイン調整部に抽出した前記音声データを供給する、

ことを特徴とする付記 1 乃至付記 5 のいずれか 1 項に記載のラウドネス調整装置。

【 0 0 8 7 】

< 付記 7 >

前記デコーダが抽出した前記音声データを解析し、無音状態が一定時間続いた場合に前記制御部に通知を行う無音検出部を更に備え、

前記制御部は、前記抑圧処理の実行中であり、かつ前記無音検出部からの通知を受信した場合、前記所定時間よりも短い時間で前記音声データの開放が完了するように前記ゲイン調整値を調整する、

ことを特徴とする付記 6 に記載のラウドネス調整装置。

【 0 0 8 8 】

< 付記 8 >

入力デジタル音声信号のラウドネス調整を行うラウドネス調整装置であって、

前記入力デジタル音声信号から第 1 の音声データ及び第 2 の音声データを抽出するエンコーダと、

前記エンコーダが抽出した前記第 1 の音声データに対し、第 1 ゲイン調整値を用いてゲイン調整を行う第 1 ゲイン調整部と、

前記エンコーダが抽出した前記第 2 の音声データに対し、第 2 ゲイン調整値を用いてゲイン調整を行う第 2 ゲイン調整部と、

前記第 1 ゲイン調整部によるゲイン調整済みの音声データを特性フィルタ処理したフィルタ済み音声データを生成する第 1 フィルタ部と、

前記第 2 ゲイン調整部によるゲイン調整済みの音声データを特性フィルタ処理したフィルタ済み音声データを生成する第 2 フィルタ部と、

前記第 1 フィルタ部の生成したフィルタ済み音声データが第 1 閾値以下の場合、前記第 1 ゲイン調整部によるゲイン調整済みの音声データを増幅して出力する第 1 ブースト処理部と、

前記第 2 フィルタ部の生成したフィルタ済み音声データが前記第 1 閾値以下の場合、前記第 2 ゲイン調整部によるゲイン調整済みの音声データを増幅して出力する第 2 ブースト処理部と、

前記第 1 ブースト処理部が出力した音声データを特性フィルタ処理したフィルタ済み音声データを生成する第 3 フィルタ部と、

前記第 2 ブースト処理部が出力した音声データを特性フィルタ処理したフィルタ済み音声データを生成する第 4 フィルタ部と、

前記第 3 フィルタ部の生成したフィルタ済み音声データが第 2 閾値以上の場合、前記ブースト処理部が増幅した音声データに対してリミッタコンプレッサ処理を行って出力する第 1 リミッタコンプレッサ部と、

前記第 4 フィルタ部の生成したフィルタ済み音声データが前記第 2 閾値以上の場合、前記ブースト処理部が増幅した音声データに対してリミッタコンプレッサ処理を行って出力する第 2 リミッタコンプレッサ部と、

を備えるラウドネス調整装置。

【0089】

<付記9>

前記第1リミッタコンプレッサ部が出力した音声データを特性フィルタ処理したフィルタ済み音声データを生成する第5フィルタ部と、

前記第2リミッタコンプレッサ部が出力した音声データを特性フィルタ処理したフィルタ済み音声データを生成する第6フィルタ部と、

前記第5フィルタ部が処理したフィルタ済み音声データに対し、第1の時間区間を用いた移動二乗平均により算出した第1二乗平均値を算出する第1二乗平均算出部と、

前記第5フィルタ部が処理したフィルタ済み音声データに対し、第1の時間区間より長い第2の時間区間を用いた移動二乗平均により算出した第2二乗平均値を算出する第2二乗平均算出部と、

前記第6フィルタ部が処理したフィルタ済み音声データに対し、前記第1の時間区間を用いた移動二乗平均により算出した第3二乗平均値を算出する第3二乗平均算出部と、

前記第6フィルタ部が処理したフィルタ済み音声データに対し、前記第2の時間区間を用いた移動二乗平均により算出した第4二乗平均値を算出する第4二乗平均算出部と、

前記第1及び第2二乗平均値、及びターゲットラウドネス値に基づいて前記ゲイン調整値を算出する制御部と、

前記第1及び第3二乗平均値の積算値、前記第2及び第4二乗平均値の積算値、及びターゲットラウドネス値に基づいて、前記第1ゲイン調整値及び前記第2ゲイン調整値を算出する制御部と、を備えることを特徴とする付記8に記載のラウドネス調整装置。

【0090】

<付記10>

付記1乃至付記9のいずれか1項に記載のラウドネス調整装置を内蔵した映像/音声処理装置。

【0091】

<付記11>

入力デジタル音声信号のラウドネス調整を行うラウドネス調整方法であって、

前記入力デジタル音声信号から抽出された音声データに対し、ゲイン調整を行うゲイン調整ステップと、

前記ゲイン調整ステップにおけるゲイン調整済みの音声データを特性フィルタ処理したフィルタ済み音声データを生成する第1フィルタステップと、

前記第1フィルタステップにて生成したフィルタ済み音声データが第1閾値以下の場合、前記ゲイン調整ステップにおいてゲイン調整済みの音声データを増幅して出力するブースト処理ステップと、

前記ブースト処理ステップにて出力された音声データを特性フィルタ処理したフィルタ済み音声データを生成する第2フィルタステップと、

前記第2フィルタステップにて生成したフィルタ済み音声データが第2閾値以上の場合、前記ブースト処理部が増幅した音声データに対してリミッタコンプレッサ処理を行って出力するリミッタコンプレッサステップと、を備えるラウドネス調整方法。

【0092】

<付記12>

前記リミッタコンプレッサステップに手出力した音声データを特性フィルタ処理したフィルタ済み音声データを生成する第3フィルタステップと、

前記第3フィルタステップにて生成したフィルタ済み音声データに対し、第1の時間区間を用いた移動二乗平均により第1二乗平均値を算出する第1二乗平均算出ステップと、

前記第3フィルタステップにて生成したフィルタ済み音声データに対し、第1の時間区間より長い第2の時間区間を用いた移動二乗平均により第2二乗平均値を算出する第2二乗平均算出ステップと、

前記第1及び第2二乗平均値、及びターゲットラウドネス値に基づいて前記ゲイン調整

10

20

30

40

50

に用いるゲイン調整値を算出する制御ステップと、
を備える付記 1 1 に記載のラウドネス調整方法。

【 0 0 9 3 】

< 付記 1 3 >

前記制御ステップでは、前記第 1 二乗平均値が前記ターゲットラウドネス値よりも大きく、前記第 2 二乗平均値が増加傾向にある場合、前記入力デジタル音声信号から抽出された前記音声データを抑圧するように前記ゲイン調整値を調整する抑圧処理を行い、

前記第 1 二乗平均値が前記ターゲットラウドネス値よりも大きく、前記第 2 二乗平均値が減少傾向にある場合、前記抑圧処理を中断し、

上記以外の場合には、前記抑圧処理は行わない、
ことを特徴とする付記 1 2 に記載のラウドネス調整方法。

10

【 0 0 9 4 】

< 付記 1 4 >

前記制御ステップでは、

外部からの制御情報の入力無く、前記入力デジタル音声信号の音声モードの変更がなく、前記抑圧処理の実行中であり、かつ前記第 1 二乗平均値が開放閾値以下である場合、所定時間で前記入力デジタル音声信号から抽出された前記音声データの開放が完了するように前記ゲイン調整値を調整する開放処理を行い、

外部からの制御情報の入力無く、前記入力デジタル音声信号の音声モードの変更がなく、前記抑圧処理の実行中であり、かつ前記第 1 二乗平均値が前記開放閾値以下ではない場合、前記開放処理を行わない、

ことを特徴とする付記 1 3 に記載のラウドネス調整方法。

20

【 0 0 9 5 】

< 付記 1 5 >

前記制御ステップでは、

前記抑圧処理の実行中であり、かつ外部からの制御情報の入力及び前記入力デジタル音声信号の音声モードの変更のいずれか一方が生じた場合、前記所定時間よりも短い時間で前記入力デジタル音声信号から抽出された前記音声データの開放が完了するように前記ゲイン調整値を調整する、

ことを特徴とする付記 1 3 または付記 1 4 に記載のラウドネス調整方法。

30

【 0 0 9 6 】

< 付記 1 6 >

入力デジタル音声信号のラウドネス調整をコンピュータに実行させるプログラムであって、

コンピュータに、

前記入力デジタル音声信号から抽出された音声データに対し、ゲイン調整を行うゲイン調整ステップと、

前記ゲイン調整ステップにおけるゲイン調整済みの音声データを特性フィルタ処理したフィルタ済み音声データを生成する第 1 フィルタステップと、

前記第 1 フィルタステップにて生成したフィルタ済み音声データが第 1 閾値以下の場合、前記ゲイン調整ステップにおいてゲイン調整済みの音声データを増幅して出力するブースト処理ステップと、

40

前記ブースト処理ステップにて出力された音声データを特性フィルタ処理したフィルタ済み音声データを生成する第 2 フィルタステップと、

前記第 2 フィルタステップにて生成したフィルタ済み音声データが第 2 閾値以上の場合、前記ブースト処理部が増幅した音声データに対してリミッタコンプレッサ処理を行って出力するリミッタコンプレッサステップと、

を実行させるプログラム。

【 0 0 9 7 】

< 付記 1 7 >

50

前記リミッタコンプレッサステップに手出力した音声データを特性フィルタ処理したフィルタ済み音声データを生成する第3フィルタステップと、

前記第3フィルタステップにて生成したフィルタ済み音声データに対し、第1の時間区間を用いた移動二乗平均により第1二乗平均値を算出する第1二乗平均算出ステップと、

前記第3フィルタステップにて生成したフィルタ済み音声データに対し、第1の時間区間より長い第2の時間区間を用いた移動二乗平均により第2二乗平均値を算出する第2二乗平均算出ステップと、

前記第1及び第2二乗平均値、及びターゲットラウドネス値に基づいて前記ゲイン調整に用いるゲイン調整値を算出する制御ステップと、

を備える付記16に記載のプログラム。

10

【0098】

<付記18>

前記制御ステップでは、前記第1二乗平均値が前記ターゲットラウドネス値よりも大きく、前記第2二乗平均値が増加傾向にある場合、前記入力デジタル音声信号から抽出された前記音声データを抑圧するように前記ゲイン調整値を調整する抑圧処理を行い、

前記第1二乗平均値が前記ターゲットラウドネス値よりも大きく、前記第2二乗平均値が減少傾向にある場合、前記抑圧処理を中断し、

上記以外の場合には、前記抑圧処理は行わない、ことを特徴とする付記17に記載のプログラム。

20

【0099】

<付記19>

前記制御ステップでは、

外部からの制御情報の入力無く、前記入力デジタル音声信号の音声モードの変更がなく、前記抑圧処理の実行中であり、かつ前記第1二乗平均値が開放閾値以下である場合、所定時間で前記入力デジタル音声信号から抽出された前記音声データの開放が完了するように前記ゲイン調整値を調整する開放処理を行い、

外部からの制御情報の入力無く、前記入力デジタル音声信号の音声モードの変更がなく、前記抑圧処理の実行中であり、かつ前記第1二乗平均値が前記開放閾値以下ではない場合、前記開放処理を行わない、

ことを特徴とする付記18に記載のプログラム。

30

【0100】

<付記20>

前記制御ステップでは、

前記抑圧処理の実行中であり、かつ外部からの制御情報の入力及び前記入力デジタル音声信号の音声モードの変更のいずれか一方が生じた場合、前記所定時間よりも短い時間で前記入力デジタル音声信号から抽出された前記音声データの開放が完了するように前記ゲイン調整値を調整する、

ことを特徴とする付記18または付記19に記載のプログラム。

40

【符号の説明】

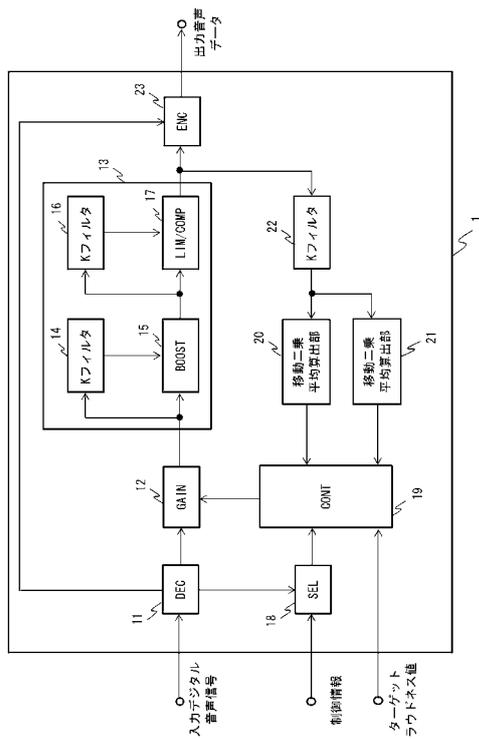
【0101】

- 1 ラウドネス調整装置
- 11 デコーダ
- 12 ゲイン調整部
- 13 音声調整部
- 14 Kフィルタ
- 15 音声データブースト部
- 16 Kフィルタ
- 17 リミッタコンプレッサ部
- 18 セレクタ

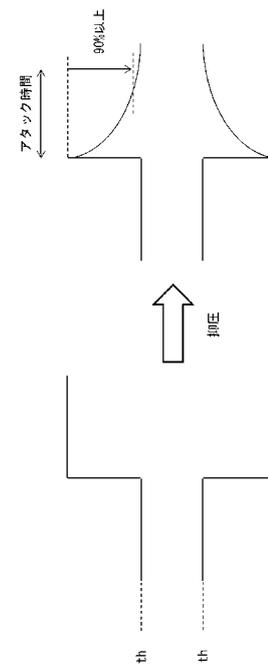
50

- 19 制御部
- 20 移動二乗平均算出部
- 21 移動二乗平均算出部
- 22 Kフィルタ
- 23 エンコーダ
- 24 ゲイン乗算部
- 25 二乗平均積算部
- 26 無音検出部

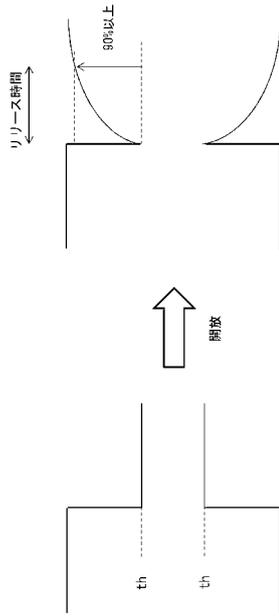
【図1】



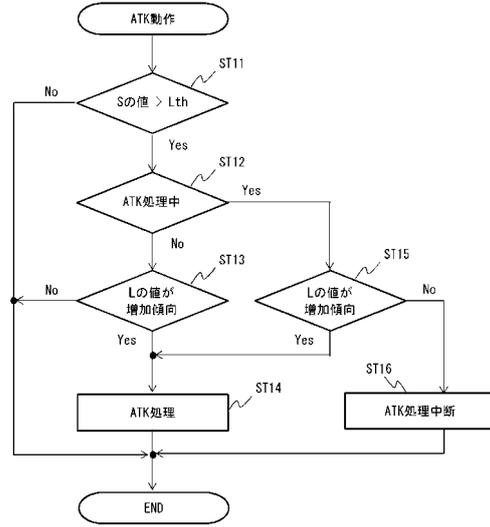
【図2】



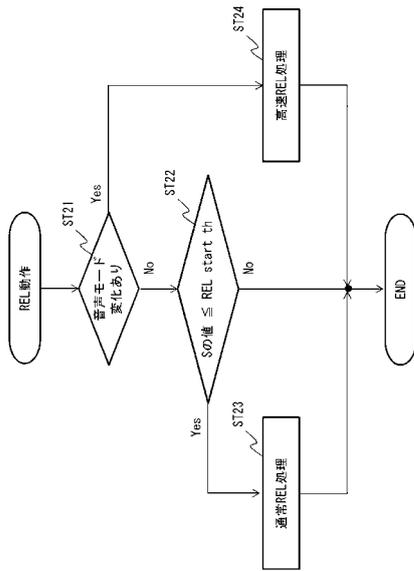
【図3】



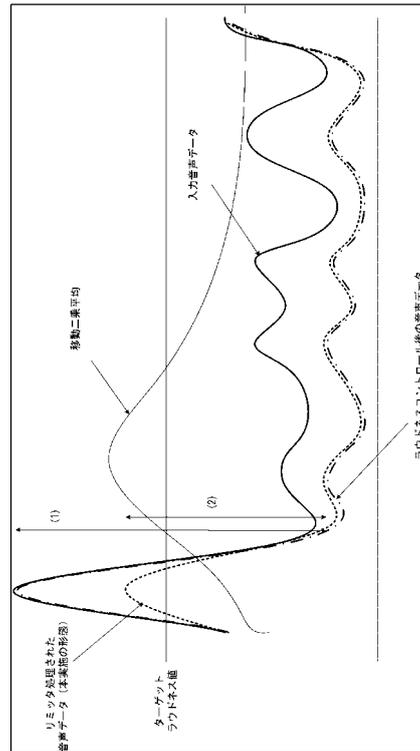
【図4】



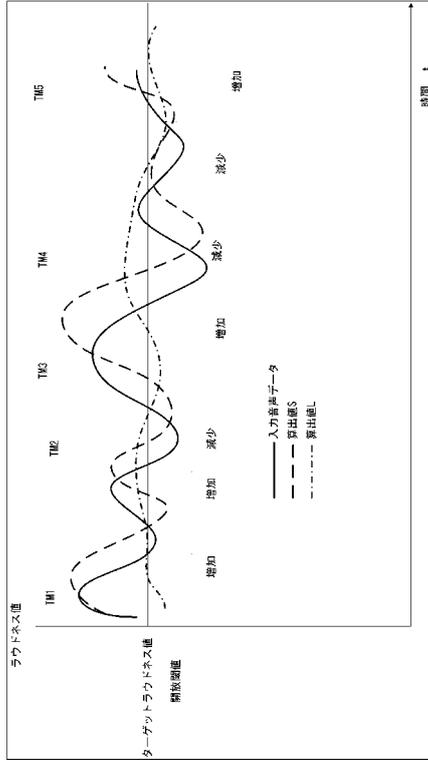
【図5】



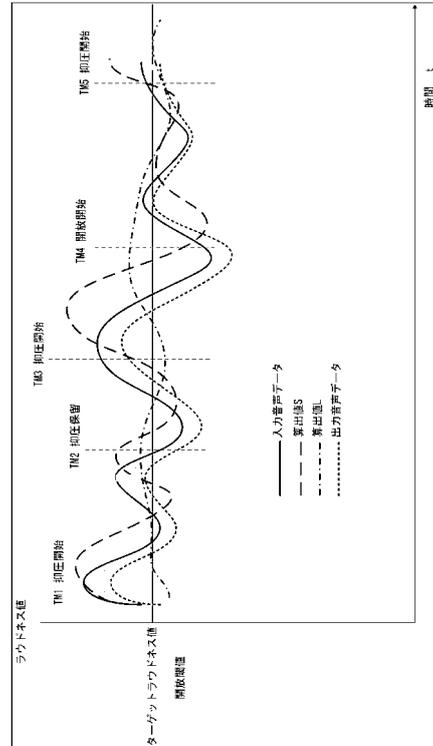
【図6】



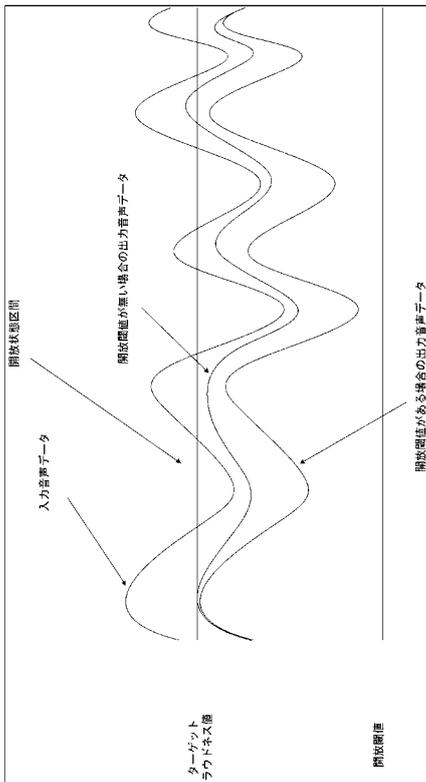
【図7】



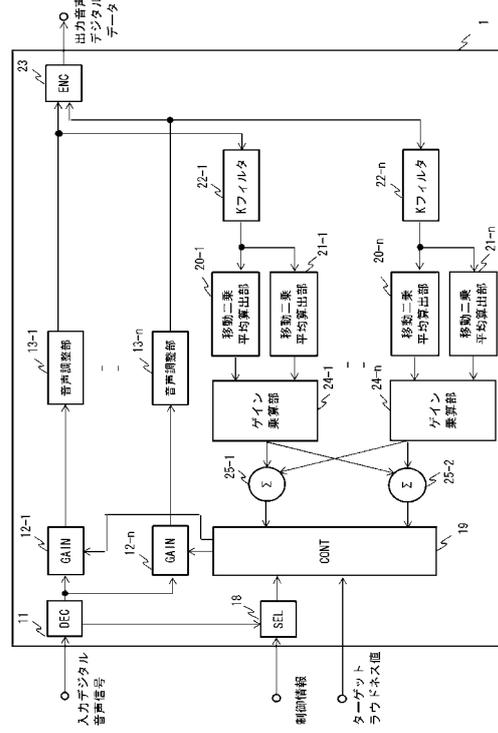
【図8】



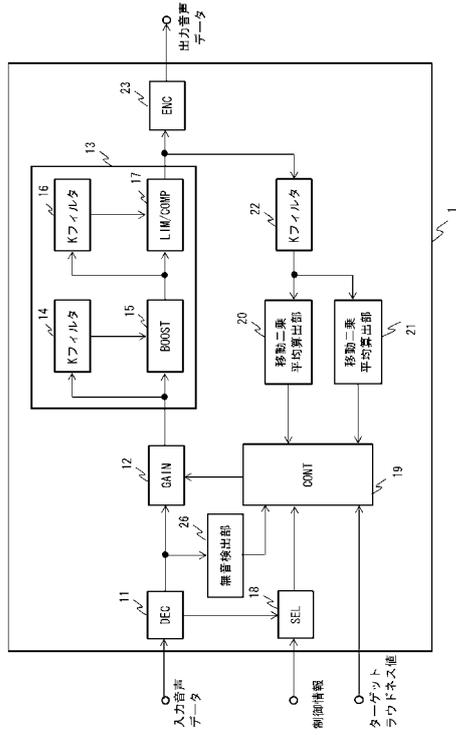
【図9】



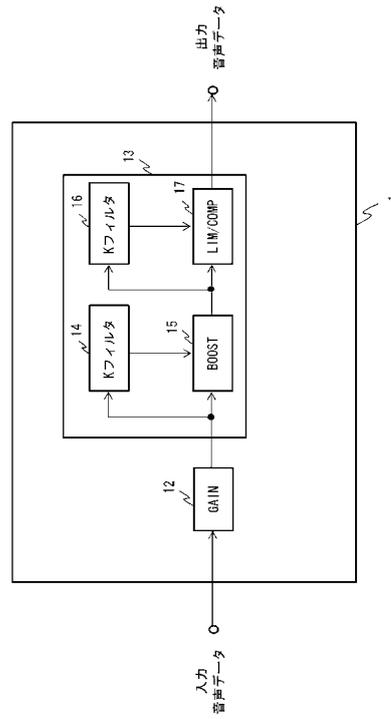
【図10】



【図 11】



【図 12】



フロントページの続き

- (56)参考文献 特開2011-151811(JP,A)
特開2009-171092(JP,A)
国際公開第2010/044439(WO,A1)
特開2007-060041(JP,A)
特開平07-312530(JP,A)

(58)調査した分野(Int.Cl., DB名)

H03G 9/00

H03G 7/00