

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第5267115号
(P5267115)

(45) 発行日 平成25年8月21日(2013.8.21)

(24) 登録日 平成25年5月17日(2013.5.17)

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

請求項の数 6 (全 23 頁)

(21) 出願番号	特願2008-332031 (P2008-332031)	(73) 特許権者	000002185 ソニー株式会社
(22) 出願日	平成20年12月26日(2008.12.26)		東京都港区港南1丁目7番1号
(65) 公開番号	特開2010-154388 (P2010-154388A)	(74) 代理人	100112955 弁理士 丸島 敏一
(43) 公開日	平成22年7月8日(2010.7.8)	(72) 発明者	渡辺 秀明 東京都港区港南1丁目7番1号 ソニー株式会社内
審査請求日	平成23年11月4日(2011.11.4)	(72) 発明者	藤田 式曜 東京都港区港南1丁目7番1号 ソニー株式会社内
		(72) 発明者	松本 淳 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】 信号処理装置、その処理方法およびプログラム

(57) 【特許請求の範囲】

【請求項1】

コンテンツにおける音声信号の周波数特性と人間の聴覚特性とに基づいて第1の音声調整情報を生成する第1の音声調整情報生成部と、

音声入力部により入力された騒音信号のうち音声出力部により出力された前記音声信号と他の騒音信号とを分離する音声分離部と、

前記音声分離部により分離された前記他の騒音信号の信号レベルに基づいて第2の音声調整情報を生成する第2の音声調整情報生成部と、

前記音声出力部に出力される前記音声信号の音量を前記第1および第2の音声調整情報に基づいて前記第1の音声調整情報が大きいほど前記音声信号の音量を小さくし、前記第2の音声調整情報が大きいほど前記音声信号の音量を大きくするように調整する音声調整部と

を具備し、

前記第1の音声調整情報生成部は、前記音声信号の周期性と前記音声信号の信号レベルとに基づいて前記音声信号が有声音であるか無声音であるかを判定する音声判定部を備え

、前記音声調整部は、前記音声判定部により有声音と判定された場合には前記音声判定部により無声音と判定されたときに比べて前記音声信号の音量を大きくする信号処理装置。

【請求項2】

前記第1の音声調整情報生成部は、前記音声信号の信号レベルに基づいて前記音声信号が無音であるか否かを判定する無音判定部をさらに備え、

前記音声調整部は、前記無音判定部により無音であると判定された場合には前記音声信号の音量を大きくしない

請求項1記載の信号処理装置。

【請求項3】

前記音声分離部は、前記音声信号と前記騒音信号に含まれる音声信号との類似度を前記音声信号に基づいて算出して前記算出された類似度に基づいて前記他の騒音信号を推定する請求項1記載の信号処理装置。

【請求項4】

前記音声分離部は、エコーキャンセラーを備える請求項3記載の信号処理装置。

【請求項5】

騒音信号を集音する音声入力部と、コンテンツにおける音声信号を出力する音声出力部とを備える信号処理装置における音量調整方法であって、

前記音声信号の周波数特性と人間の聴覚特性とに基づいて第1の音声調整情報を生成する第1の音声調整情報生成手順と、

前記音声入力部により入力された騒音信号のうち前記音声出力部により出力された前記音声信号と他の騒音信号とを分離する音声分離手順と、

前記音声分離手順により分離された前記他の騒音信号の信号レベルに基づいて第2の音声調整情報を生成する第2の音声調整情報生手順と、

前記音声出力部に出力される前記音声信号の音量を前記第1および第2の音声調整情報に基づいて前記第1の音声調整情報が大きいほど前記音声信号の音量を小さくし、前記第2の音声調整情報が大きいほど前記音声信号の音量を大きくするように調整する音声調整手順と

を具備し、

前記第1の音声調整情報生成手順において、前記音声信号の周期性と前記音声信号の信号レベルとに基づいて前記音声信号が有声音であるか無声音であるかを判定する音声判定手順を備え、

前記音声調整手順において、前記音声判定手順により有声音と判定された場合には前記音声判定手順により無声音と判定されたときに比べて前記音声信号の音量を大きくする音量調整方法。

【請求項6】

騒音信号を集音する音声入力部と、コンテンツにおける音声信号を出力する音声出力部とを備える信号処理装置において、

前記音声信号の周波数特性と人間の聴覚特性とに基づいて第1の音声調整情報を生成する第1の音声調整情報生成手順と、

前記音声入力部により入力された騒音信号のうち前記音声出力部により出力された前記音声信号と他の騒音信号とを分離する音声分離手順と、

前記音声分離手順により分離された前記他の騒音信号の信号レベルに基づいて第2の音声調整情報を生成する第2の音声調整情報生手順と、

前記音声出力部に出力される前記音声信号の音量を前記第1および第2の音声調整情報に基づいて前記第1の音声調整情報が大きいほど前記音声信号の音量を小さくし、前記第2の音声調整情報が大きいほど前記音声信号の音量を大きくするように調整する音声調整手順と

をコンピュータに実行させるプログラムであって、

前記第1の音声調整情報生成手順において、前記音声信号の周期性と前記音声信号の信号レベルとに基づいて前記音声信号が有声音であるか無声音であるかを判定する音声判定手順を備え、

前記音声調整手順において、前記音声判定手順により有声音と判定された場合には前記音声判定手順により無声音と判定されたときに比べて前記音声信号の音量を大きくする

10

20

30

40

50

プログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、信号処理装置に関し、特に音声信号の音量を自動的に制御する信号処理装置、および、その処理方法ならびに当該方法をコンピュータに実行させるプログラムに関する。

【背景技術】

【0002】

近年、テレビ放送の多チャンネル化やDVD (Digital Versatile Disk) の普及などにより、多様なAV (Audio Visual) コンテンツが記録再生装置において再生される状況にある。このような状況においては、コンテンツごとの音声信号のレベルにばらつきが多いため、視聴者は、コンテンツごとに音声信号の出力レベルを手動により調整する必要がある。このため、コンテンツごとの音声の出力レベルが均一になるよう調整する手法が考案されている。例えば、コンテンツに含まれるシーンの種別を検出することによって、シーンごとの音量差を自動的に調整する記録再生装置が提案されている (例えば、特許文献1参照。)。

10

【0003】

一方、視聴者の視聴環境において騒音ノイズが変動する場合もある。このような場合にも、視聴者は、その騒音ノイズの変動に応じて、手動により記録再生装置から出力される音声の音量を調整しなければならない。このため、マイクにより集音された信号からノイズレベルを検出して、その検出されたノイズレベルに基づいて出力音声の音量を調整する音声出力装置が提案されている (例えば、特許文献2参照。)。

20

【特許文献1】特開2007-53510号公報 (図1)

【特許文献2】特許第3286981号 (図1)

【発明の開示】

【発明が解決しようとする課題】

【0004】

上述の2つの従来技術において、前者では、コンテンツに含まれるシーンの種別を判別することによって、出力音声の音量を自動的に調整することができる。しかしながら、このときには、視聴者の視聴環境における騒音ノイズが変動する場合には、その変動に応じて、その都度手動による音声調整が必要となる。一方、後者では、視聴環境における騒音レベルに応じて出力音声の音量を調整することができる。しかしながら、コンテンツの音声信号レベルにばらつきがある場合には、コンテンツごとに手動による音声調整が必要となる。

30

【0005】

このように、従来技術では、騒音ノイズの変動またはコンテンツごとの音量レベルの違いによって、手動による音声調整が必要となってしまう。

【0006】

本発明はこのような状況に鑑みてなされたものであり、音声信号を最適な出力レベルに調整することを目的とする。

40

【課題を解決するための手段】

【0007】

本発明は、上記課題を解決するためになされたものであり、その第1の側面は、コンテンツにおける音声信号に基づいて第1の音声調整情報を生成する第1の音声調整情報生成部と、音声入力部により入力された騒音信号のうち音声出力部により出力された上記音声信号と他の騒音信号とを分離する音声分離部と、上記音声分離部により分離された上記他の騒音信号に基づいて第2の音声調整情報を生成する第2の音声調整情報生成部と、上記音声出力部に出力される上記音声信号の音量を上記第1および第2の音声調整情報に基づいて調整する音声調整部とを具備する信号処理装置およびその処理方法ならびに当該方法

50

をコンピュータに実行させるプログラムである。これにより、音声信号に基づいて生成される第1の音声調整情報と、他の騒音信号に基づいて生成される第2の音声調整情報とに基づいて音声信号の音量を調整させるという作用をもたらす。

【0008】

また、この第1の側面において、上記第1の音声調整情報生成部は、上記音声信号の周波数特性と人間の聴覚特性とに基づいて上記第1の音声調整情報を生成し、上記第2の音声調整情報生成部は、上記他の騒音信号の信号レベルに基づいて上記第2の音声調整情報を生成し、上記音声調整部は、上記第1の音声調整情報が大きいほど上記音声信号の音量を小さくし、上記第2の音声調整情報が大きいほど上記音声信号の音量を大きくするようにしてもよい。これにより、音声調整部は、音声信号の周波数特性と人間の聴覚特性とに基づいて生成された第1の音声調整情報が大きいほど音声信号の音量を小さくさせ、他の騒音信号の信号レベルに基づいて生成された第2の音声調整情報が大きいほど音声信号の音量を大きくさせるという作用をもたらす。この場合において、上記第1の音声調整情報生成部は、上記音声信号の周期性と上記音声信号の信号レベルとに基づいて上記音声信号が有声音であるか無声音であるかを判定する音声判定部を備え、上記音声調整部は、上記音声判定部により有声音と判定された場合には上記音声判定部により無声音と判定されたときに比べて上記音声信号の音量を大きくするようにしてもよい。これにより、音声調整部は、音声信号の周期性および音声信号の信号レベルとに基づいて音声信号が有声音であるか否かが判定され、有声音であると判定された場合には無声音と判定されたときに比べて音声信号の音量を大きくさせるという作用をもたらす。

10

20

【0009】

また、上記第1の音声調整情報生成部は、上記音声信号の周波数特性と人間の聴覚特性とに基づいて上記第1の音声調整情報を生成し、上記第2の音声調整情報生成部は、上記他の騒音信号の信号レベルに基づいて上記第2の音声調整情報を生成し、上記音声調整部は、上記第1の音声調整情報が大きいほど上記音声信号の音量を小さくし、上記第2の音声調整情報が大きいほど上記音声信号の音量を大きくする場合において、上記第1の音声調整情報生成部は、上記音声信号の信号レベルに基づいて上記音声信号が無音であるか否かを判定する無音判定部をさらに備え、上記音声調整部は、上記無音判定部により無音であると判定された場合には上記音声信号の音量を大きくしないようにしてもよい。これにより、音声信号の信号レベルに基づいて音声信号が無音であると判定された場合には音声信号の音量を大きくさせないという作用をもたらす。

30

【0010】

また、この第1の側面において、上記音声分離部は、上記音声信号と上記騒音信号に含まれる音声信号との類似度を上記音声信号に基づいて算出して上記算出された類似度に基づいて上記他の騒音信号を推定するようにしてもよい。これにより、音声分離部は、音声信号に基づいて推定される騒音信号に含まれる音声信号を騒音信号から除去させるという作用をもたらす。この場合において、上記音声分離部は、エコーキャンセラーを備えるようにしてもよい。これにより、音声分離部は、エコーキャンセラーを備えることによって騒音信号のうち騒音信号に含まれる音声信号を除去させるという作用をもたらす。

40

【発明の効果】

【0011】

本発明によれば、音声信号を最適な出力レベルに調整させるという優れた効果を奏し得る。

【発明を実施するための最良の形態】

【0012】

以下、本発明を実施するための最良の形態（以下、実施の形態と称する）について説明する。説明は以下の順序により行う。

1. 第1の実施の形態（音声信号の音量制御：記録再生装置の構成例）
2. 第2の実施の形態（音声信号の音量制御：制御情報のデータ形式の例）
3. 第3の実施の形態（音声信号の音量制御：利得算出手法の例）

50

4. 第4の実施の形態（音声信号の音量制御：音量制御手順の例）

【0013】

< 1. 第1の実施の形態 >

[記録再生装置の構成例]

図1は、本発明の第1の実施の形態における記録再生装置の一構成例を示すブロック図である。記録再生装置100は、アンテナ110と、チューナ部120と、コンテンツ記録部130と、コンテンツ再生部140と、コンテンツ解析情報生成部150と、スピーカ160とを備える。また、記録再生装置100は、マイク170と、環境騒音分離部180と、環境騒音解析情報生成部190と、音声調整部200とを備える。

【0014】

アンテナ110は、放送信号を受信するために用いられるものである。このアンテナ110は、例えば、地上波放送、BS放送やCS放送などの放送信号を受信する。

【0015】

チューナ部120は、アンテナ110により受信された信号を復調するものである。このチューナ部120は、その復調された受信データであるコンテンツデータをコンテンツ再生部140に供給する。ここにいうコンテンツデータには、例えば、地上波放送、BS放送、CS放送などの放送コンテンツや放送コンテンツに付随するEPGなどのメタデータが含まれる。

【0016】

コンテンツ記録部130は、チューナ部120から供給されるコンテンツデータを所定のフォーマットに変換して記録するものである。また、このコンテンツ記録部130は、コンテンツ記録部130において記録されているコンテンツデータをコンテンツ再生部140に供給する。ここでは、チューナ部120からのコンテンツデータを記録する例について説明したが、記録再生装置100にAV入力端子を設けて外部装置から供給されるコンテンツデータを記録するようにしてもよい。

【0017】

コンテンツ再生部140は、チューナ部120から供給されるコンテンツデータまたはコンテンツ記録部130から供給されるコンテンツデータを再生するものである。このコンテンツ再生部140は、例えば、チューナ部120から供給される音声データを復号することによって音声信号を生成する。また、このコンテンツ再生部140は、コンテンツ記録部130からのAVコンテンツデータに含まれる画像データと、これに対応する音声データとを復号することによって画像信号および音声信号を生成する。

【0018】

また、コンテンツ再生部140は、信号線149および201を介してコンテンツ解析情報生成部150および音声調整部200に復号された音声信号をそれぞれ供給する。このコンテンツ再生部140は、例えば、復号された音声信号に加えて、画像データやメタデータなどを供給する。また、コンテンツ再生部140は、チューナ部120から供給されたコンテンツデータを復号してコンテンツ記録部130に供給する。

【0019】

コンテンツ解析情報生成部150は、コンテンツ再生部140から供給されるコンテンツデータを解析することによって、コンテンツ再生部140において再生される音声信号に関するコンテンツ解析情報を生成するものである。このコンテンツ解析情報生成部150は、コンテンツ再生部140から供給される音声信号に基づいて、フレームごとにコンテンツ解析情報を生成する。ここにいうフレームとは、音声信号における標本化された一定数のサンプルのことをいう。このコンテンツ解析情報生成部150は、例えば、コンテンツ再生部140から供給された音声信号の周波数特性に基づいてコンテンツ解析情報を生成する。

【0020】

また、このコンテンツ解析情報生成部150は、例えば、コンテンツにおけるシーンの種別（CM（Commercial Message）/ニュース番組など）を判定して、その判定結果をコ

10

20

30

40

50

コンテンツ解析情報として音声調整部 200 に供給する。この例において、コンテンツ解析情報生成部 150 は、コンテンツにおける映像信号の輝度情報などの時間的変化からシーンを検出する。そして、コンテンツ解析情報生成部 150 は、その検出結果と EPG (Electric Program Guide) データなどの情報とを組み合わせることでシーンの種別を判定する。

【0021】

また、コンテンツ解析情報生成部 150 は、生成されたコンテンツ解析情報を音声調整部 200 に供給する。なお、コンテンツ解析情報生成部 150 は、特許請求の範囲に記載の第 1 の音声調整情報生成部の一例である。また、コンテンツ解析情報は、特許請求の範囲の第 1 の音声調整情報の一例である。

【0022】

スピーカ 160 は、音声調整部 200 から供給される音声信号を出力音声として出力する拡音器である。なお、スピーカ 160 は、特許請求の範囲に記載の音声出力部の一例である。

【0023】

マイク 170 は、記録再生装置 100 の周辺の環境音を集音するマイクロフォンである。このマイク 170 は、その集音された環境音を電気信号に変換して、騒音信号として環境騒音分離部 180 に供給する。この騒音信号には、スピーカ 160 から出力された出力音声と、それ以外の環境騒音とが含まれる。なお、マイク 170 は、特許請求の範囲に記載の音声入力部の一例である。

【0024】

環境騒音分離部 180 は、マイク 170 から供給される騒音信号と、音声調整部 200 から供給される音声信号とに基づいて、騒音信号に含まれるスピーカ 160 からの出力音声除去するものである。すなわち、環境騒音分離部 180 は、マイク 170 からの騒音信号のうち、スピーカ 160 から出力された音声信号の成分と、他の騒音信号である環境騒音信号の成分とを分離する。

【0025】

この環境騒音分離部 180 は、音声調整部 200 から供給される音声信号と、騒音信号に含まれる出力音声との類似度を音声調整部 200 からの音声信号に基づいて算出して、その算出された類似度に基づいて環境騒音信号を推定する。この環境騒音分離部 180 は、例えば、エコーキャンセラーにより実現される。また、環境騒音分離部 180 は、分離された環境騒音信号を、信号線 189 を介して環境騒音解析情報生成部 190 に供給する。なお、環境騒音分離部 180 は、特許請求の範囲に記載の音声分離部の一例である。

【0026】

環境騒音解析情報生成部 190 は、環境騒音分離部 180 から供給される環境騒音信号を解析することによって、環境騒音信号に関する環境騒音解析情報を生成するものである。この環境騒音解析情報生成部 190 は、環境騒音分離部 180 から供給される環境騒音信号に基づいて環境騒音解析情報を生成する。この環境騒音解析情報生成部 190 は、例えば、環境騒音分離部 180 から供給される環境騒音信号の信号レベルに基づいて環境騒音解析情報を生成する。また、環境騒音解析情報生成部 190 は、その生成された環境騒音解析情報を音声調整部 200 に供給する。なお、環境騒音解析情報生成部 190 は、特許請求の範囲に記載の第 2 の音声調整情報生成部の一例である。また、環境騒音解析情報は、特許請求の範囲の第 2 の音声調整情報の一例である。

【0027】

音声調整部 200 は、コンテンツ解析情報生成部 150 からのコンテンツ解析情報および環境騒音解析情報生成部 190 からの環境騒音解析情報に基づいて、コンテンツ再生部 140 から供給された音声信号の音量を調整するものである。この音声調整部 200 は、信号線 209 を介して、その調整された音声信号をスピーカ 160 および環境騒音分離部 180 に供給する。なお、音声調整部 200 は、特許請求の範囲に記載の音声調整部の一例である。

【0028】

10

20

30

40

50

[コンテンツ解析情報生成部および環境騒音解析情報生成部の構成例]

図2は、本発明の第1の実施の形態におけるコンテンツ解析情報生成部150および環境騒音解析情報生成部190の一構成例を示すブロック図である。ここでは、音声調整部200は、図1と同一のものであるため、同一符号を付してここでの説明を省略する。

【0029】

コンテンツ解析情報生成部150は、音声レベル算出部151と、無音判定部152と、ピッチゲイン算出部153と、音声判定部154と、パワースペクトル算出部155と、ラウドネスレベル算出部156とを備える。また、環境騒音解析情報生成部190は、騒音レベル算出部191およびパワースペクトル算出部192を備える。

【0030】

音声レベル算出部151は、信号線149を介して供給される音声信号の信号レベルをフレームごとに算出するものである。この音声レベル算出部151は、音声信号の各フレームにおけるサンプル値の二乗平均平方根(パワー値)である信号レベルを音声レベルとして算出する。この音声レベル算出部151は、その算出された音声レベルを無音判定部152、音声判定部154および音声調整部200に供給する。

【0031】

無音判定部152は、音声レベル算出部151から供給される音声信号の音声レベルに基づいて、その音声信号に関して無音であるか否かを判定するものである。この無音判定部152は、音声レベル算出部151からの音声レベルおよび予め定められた閾値 T_s (音声レベル閾値)に基づいて音声信号が無音であるか否かを判定して、その判定結果に基づいて無音判定情報を生成する。この無音判定部152は、例えば、音声レベルが閾値未満である場合には無音である旨を示す無音判定情報(Silence Flag = True)を生成する。一方、この無音判定部152は、音声レベルが閾値以上である場合には無音でない旨を示す無音判定情報(Silence Flag = False)を生成する。また、この無音判定部152は、その生成された無音判定情報をコンテンツ解析情報として音声調整部200に供給する。なお、無音判定部152は、特許請求の範囲に記載の無音判定部の一例である。

【0032】

ピッチゲイン算出部153は、信号線149を介して供給される音声信号を分析することによってピッチゲインを算出するものである。ここにいうピッチゲインとは、人間の発する音声の特徴を表わすピッチ成分の強度を示す指標である。このピッチゲイン算出部153は、信号線149から供給される音声信号の周期性に基づいてピッチゲインをフレームごとに算出する。このピッチゲイン算出部153は、その算出されたピッチゲインを音声判定部154に供給する。

【0033】

音声判定部154は、音声レベル算出部151から供給される音声レベルと、ピッチゲイン算出部153から供給されるピッチゲインとに基づいて、コンテンツ再生部140からの音声信号が有音声であるか無音声であるかを判定するものである。すなわち、音声判定部154は、時間領域における音声信号の周期性と音声信号の信号レベルとに基づいて有音声であるか否かを判定する。そして、この音声判定部154は、その判定結果に基づいて音声判定情報を生成する。

【0034】

この音声判定部154は、例えば、音声レベルが上述の閾値 T_s (音声レベル閾値)以上であり、かつ、ピッチゲインが一定の閾値 T_p (ピッチゲイン閾値)以上である場合には有音声である旨を示す音声判定情報を生成する。それ以外の場合には、音声判定部154は、無音声である旨を示す音声判定情報を生成する。また、この音声判定部154は、例えば、音声レベルおよびピッチゲインに対応付けられた音声信号の特徴量を予め保持しておき、音声レベル算出部151からの音声レベルおよびピッチゲイン算出部153からのピッチゲインに対応する音声信号の特徴量を取得する。そして、この音声判定部154は、その音声信号の特徴量が予め定められた閾値 T_f (特徴量閾値)より大きい場合には

10

20

30

40

50

有音声である旨を示す音声判定情報を生成し、特徴量が閾値未満である場合には無音声である旨を示す音声判定情報を生成する。また、音声判定部 154 は、その生成された音声判定情報をコンテンツ解析情報として音声調整部 200 に供給する。なお、音声判定部 154 は、特許請求の範囲に記載の音声判定部の一例である。

【0035】

パワースペクトル算出部 155 は、信号線 149 を介して供給される音声信号の周波数特性に基づいて、その音声信号のパワースペクトルを算出するものである。このパワースペクトル算出部 155 は、その算出されたパワースペクトルをラウドネスレベル算出部 156 に供給する。

【0036】

ラウドネスレベル算出部 156 は、パワースペクトル算出部 155 から供給されるパワースペクトルに基づいてラウドネスレベルを算出するものである。ここにいうラウドネスレベルとは、人間の聴覚特性を考慮した音の大きさを示す指標のことをいう。すなわち、このラウドネスレベル算出部 156 は、音声信号の周波数特性と人間の聴覚特性とに基づいてラウドネスレベルを算出する。

【0037】

このラウドネスレベル算出部 156 は、例えば、ISO (International Organization for Standardization) 532B の規定に基づいてラウドネスレベルを算出する。この例において、ラウドネスレベル算出部 156 は、臨界帯域ごとの音声信号におけるパワーに対するマスキングカーブをそれぞれ生成する。そして、ラウドネスレベル算出部 156 は、その生成された複数のマスキングカーブが重ね合わせられた領域の面積を算出することによって、ラウドネスレベルを算出する。また、ラウドネスレベル算出部 156 は、その算出されたラウドネスレベルをコンテンツ解析情報として音声調整部 200 に供給する。なお、ラウドネスレベル算出部 156 は、特許請求の範囲に記載の第 1 の音声調整情報生成部の一例である。

【0038】

騒音レベル算出部 191 は、信号線 189 を介して供給される環境騒音信号の信号レベルをフレームごとに算出するものである。この騒音レベル算出部 191 は、環境騒音信号の各フレームにおけるサンプル値の二乗平均平方根 (パワー値) である信号レベルを騒音レベルとして算出する。この騒音レベル算出部 191 は、その算出された騒音レベルを環境騒音解析情報として音声調整部 200 に供給する。なお、騒音レベル算出部 191 は、特許請求の範囲に記載の第 2 の音声調整情報生成部の一例である。

【0039】

パワースペクトル算出部 192 は、信号線 189 を介して供給される環境騒音信号の周波数特性に基づいて、当該環境騒音信号のパワースペクトルを算出するものである。このパワースペクトル算出部 192 は、その算出されたパワースペクトルを環境騒音解析情報として音声調整部 200 に供給する。

【0040】

[音声調整部の構成例]

図 3 は、本発明の第 1 の実施の形態における音声調整部 200 の一構成例を示すブロック図である。

【0041】

音声調整部 200 は、ゲイン特性決定部 210 と、目標利得算出部 220 と、調整利得算出部 230 と、利得設定部 240 と、コンプレッサ処理部 251 と、イコライジング処理部 252 と、全体音量増幅部 253 と、調整帯域設定部 260 とを備える。

【0042】

ゲイン特性決定部 210 は、コンテンツ解析情報および環境騒音解析情報に基づいて、音声信号の音量の増加量を算出するためのゲイン特性を決定するものである。このゲイン特性決定部 210 は、ラウドネスレベル算出部 156 からのラウドネスレベルと、音声判定部 154 からの音声判定情報と、騒音レベル算出部 191 からの騒音レベルとに基づい

10

20

30

40

50

てゲイン特性を決定する。ゲイン特性決定部 2 1 0 は、最大利得テーブル 2 1 1 と、最大利得取得部 2 1 2 と、ゲイン特性傾き決定部 2 1 3 と、最小騒音レベル抽出部 2 1 4 とを備える。

【 0 0 4 3 】

最大利得テーブル 2 1 1 は、音声信号のラウドネスレベルおよび騒音レベルに対応付けられたゲイン特性における最大利得を保持するものである。このゲイン特性における最大利得は、ゲイン特性における上限値であり、音声信号を過剰に増幅させることを防止するために設けられている。また、最大利得テーブル 2 1 1 は、最大利得取得部 2 1 2 からの音声信号のラウドネスレベルおよび騒音レベルに基づいて対応付けられた最大利得を最大利得取得部 2 1 2 に出力する。

10

【 0 0 4 4 】

この最大利得テーブル 2 1 1 は、音声信号のラウドネスレベルが大きい程、視聴者にとって聞き取り易い音声であるため、小さい値の最大利得を出力し、ラウドネスレベルが小さい程、聞き取り難い音声であるため、大きい値の最大利得を出力する。一方、この最大利得テーブル 2 1 1 は、騒音レベルが大きい程、環境騒音が大きいため、大きい値の最大利得を出力し、騒音レベルが小さい程、環境騒音が小さいため、小さい値の最大利得を出力する。すなわち、この最大利得テーブル 2 1 1 は、音声信号のラウドネスレベルが大きい程、小さい値の最大利得を出力し、騒音レベルが大きい程、大きい値の最大利得を出力する。

【 0 0 4 5 】

最大利得取得部 2 1 2 は、ラウドネスレベル算出部 1 5 6 から供給されるラウドネスレベルと、騒音レベル算出部 1 9 1 から供給される騒音レベルとに基づいてゲイン特性における最大利得を取得するものである。この最大利得取得部 2 1 2 は、ラウドネスレベル算出部 1 5 6 からのラウドネスレベルおよび騒音レベル算出部 1 9 1 からの騒音レベルをフレームごとに最大利得テーブル 2 1 1 に供給する。そして、この最大利得取得部 2 1 2 は、ラウドネスレベル算出部 1 5 6 からのラウドネスレベル、および、騒音レベル算出部 1 9 1 からの騒音レベルに対応付けられた最大利得を最大利得テーブル 2 1 1 から取得する。また、この最大利得取得部 2 1 2 は、その取得された最大利得を目標利得算出部 2 2 0 に供給する。なお、最大利得取得部 2 1 2 は、特許請求の範囲に記載の音声調整部の一例である。

20

30

【 0 0 4 6 】

なお、ここでは、音声信号のラウドネスレベルに基づいて最大利得を取得する例について説明したが、音声信号のラウドネスレベルに代えて音声信号の音声レベルを用いるようにしてもよい。また、環境騒音信号の騒音レベルに基づいて最大利得を取得する例について説明したが、騒音レベルに代えて環境騒音信号のラウドネスレベルを環境騒音解析情報生成部 1 9 0 において生成し、その生成されたラウドネスレベルを用いるようにしてもよい。

【 0 0 4 7 】

ゲイン特性傾き決定部 2 1 3 は、音声判定部 1 5 4 から供給される音声判定情報に基づいてゲイン特性における傾きを決定するものである。このゲイン特性傾き決定部 2 1 3 は、音声判定情報が有音声である旨を示す場合には、音声判定情報が無音声である旨を示すときに比べてゲイン特性の傾きを大きくする。

40

【 0 0 4 8 】

このゲイン特性傾き決定部 2 1 3 は、例えば、ゲイン特性の傾きを予め保持しておき、音声判定情報が有音声を示す場合には、音声判定情報が無音声を示すときに比べて大きい値のゲイン特性の傾きを選択する。一方、このゲイン特性傾き決定部 2 1 3 は、音声判定情報が無音声を示す場合には、音声判定情報が有音声を示すときに比べて小さい値のゲイン特性の傾きを選択する。また、ゲイン特性傾き決定部 2 1 3 は、選択されたゲイン特性における傾きを目標利得算出部 2 2 0 に供給する。なお、ゲイン特性傾き決定部 2 1 3 は、特許請求の範囲に記載の音声調整部の一例である。また、ここでは、音声判定情報に基づい

50

てゲイン特性の傾きを決定する例について説明したが、再生されるコンテンツにおけるシーンの種別に応じてゲイン特性の傾きを決定するようにしてもよい。

【 0 0 4 9 】

最小騒音レベル抽出部 2 1 4 は、騒音レベル算出部 1 9 1 から供給されるフレームごとの騒音レベルのうち最小の騒音レベルを抽出するものである。この最小騒音レベル抽出部 2 1 4 は、例えば、一定期間において最も小さい騒音レベルである最小騒音レベルを抽出して、その抽出された最小騒音レベルが過去の最小騒音レベルより小さい場合には、その最小騒音レベルを新たな最小騒音レベルとして保持する。また、最小騒音レベル抽出部 2 1 4 は、その抽出された最小騒音レベルを暗騒音レベルとして目標利得算出部 2 2 0 に供給する。

10

【 0 0 5 0 】

目標利得算出部 2 2 0 は、ゲイン特性決定部 2 1 0 から供給されるゲイン特性における最大利得、傾きおよび暗騒音レベルを用いることによって、騒音レベル算出部 1 9 1 からの騒音レベルに基づいて目標利得を算出するものである。この目標利得算出部 2 2 0 は、最大利得取得部 2 1 2 からの最大利得と、ゲイン特性傾き決定部 2 1 3 からの傾きと、最小騒音レベル抽出部 2 1 4 からの暗騒音レベルとを用いてゲイン特性を生成する。そして、この目標利得算出部 2 2 0 は、その生成されたゲイン特性において、騒音レベル算出部 1 9 1 からの騒音レベルに対応する目標利得を算出する。また、目標利得算出部 2 2 0 は、その算出された目標利得を調整利得算出部 2 3 0 に供給する。

【 0 0 5 1 】

20

調整利得算出部 2 3 0 は、音声信号に対する不自然な音量の増減を抑制するために、目標利得に基づいて調整利得を算出するものである。この調整利得算出部 2 3 0 は、目標利得算出部 2 2 0 から供給される目標利得と、無音判定部 1 5 2 から供給される無音判定情報とに基づいて調整利得を算出する。調整利得算出部 2 3 0 は、例えば、無音判定情報が無音でない旨を示し、かつ、目標利得 ($t a r g e t _ g a i n [m]$) が前のフレームの調整利得 ($e q _ g a i n [m - 1]$) より大きい場合には式 1 に基づいて調整利得 ($e q _ g a i n [m]$) を算出する。それ以外の場合には、調整利得算出部 2 3 0 は、式 2 に基づいて調整利得 ($e q _ g a i n [m]$) を算出する。

$$e q _ g a i n [m] = t 1 \cdot t a r g e t _ g a i n [m] + (1 - t 1) \cdot e q _ g a i n [m - 1] \quad \dots \text{式 1}$$

30

$$e q _ g a i n [m] = t 2 \cdot e q _ g a i n [m - 1] \quad \dots \text{式 2}$$

【 0 0 5 2 】

ここで、 $t 1$ および $t 2$ は定数である。また、 $t 1$ は「 0 . 0 」より大きく設定され、 $t 2$ は「 1 . 0 」未満に設定される。

【 0 0 5 3 】

式 1 により、現在のフレームの音声信号が無音状態でなく、かつ、1 つ前のフレームの音声信号の音量からさらに音量を増加させる場合には、1 つ前のフレームの音量からの急激な音量増加を抑制することができる。また、式 2 により、現在のフレームの音声信号が無音と判定された場合には、目標利得に関係なく、1 つ前のフレームの音量に基づいて音量調整するため、無音状態における不自然な音量増加を防止することができる。また、1 つ前のフレームの音量から低下させる場合にも、急激な音量低下を抑制することができる。また、調整利得算出部 2 3 0 は、その算出された調整利得を利得設定部 2 4 0 に供給する。なお、調整利得算出部 2 3 0 は、特許請求の範囲に記載の音声調整部の一例である。

40

【 0 0 5 4 】

利得設定部 2 4 0 は、調整利得算出部 2 3 0 から供給された調整利得に基づいて、コンプレッサ処理部 2 5 1、イコライジング処理部 2 5 2 および全体音量増幅部 2 5 3 の利得をそれぞれ設定するものである。

【 0 0 5 5 】

この利得設定部 2 4 0 は、例えば、調整利得算出部 2 3 0 から供給される調整利得が一

50

定の閾値 T_a (コンプレッサ処理閾値) 以下である場合には、コンプレッサ処理部 251 のみに音声信号を増幅させるように利得を設定する。また、この利得設定部 240 は、調整利得が閾値 T_a より大きく、かつ、一定の閾値 T_b (イコライジング処理閾値) 以下である場合には、コンプレッサ処理部 251 およびイコライジング処理部 252 によって音声信号を増幅させるように利得を設定する。この利得設定部 240 は、調整利得が閾値 T_b より大きい場合には、コンプレッサ処理部 251、イコライジング処理部 252 および全体音量増幅部 253 によって音声信号を増幅させるように利得を設定する。

【0056】

コンプレッサ処理部 251 は、音声レベル算出部 151 から供給される音声レベルに応じて、音声信号の音圧を補正するものである。このコンプレッサ処理部 251 は、利得設定部 240 により設定される利得と、音声レベル算出部 151 から供給される音声レベルとに基づいて、信号線 201 を介して供給される音声信号を増幅させる。このコンプレッサ処理部 251 は、例えば、音声レベル算出部 151 から供給される音声レベルに応じて、音声信号の音量の増幅率を変更する。また、このコンプレッサ処理部 251 は、その増幅された音声信号をイコライジング処理部 252 に供給する。

10

【0057】

イコライジング処理部 252 は、環境騒音信号の周波数帯域に基づいて音声信号における周波数成分を増幅させるものである。このイコライジング処理部 252 は、利得設定部 240 により設定される利得と、調整帯域設定部 260 により設定される最大周波数とに基づいて、コンプレッサ処理部 251 によって増幅された音声信号をさらに増幅する。また、イコライジング処理部 252 は、増幅された音声信号を全体音量増幅部 253 に供給する。

20

【0058】

全体音量増幅部 253 は、利得設定部 240 により設定された利得に基づいて、イコライジング処理部 252 により増幅された音声信号をさらに増幅するものである。この全体音量増幅部 253 は、その増幅された音声信号を信号線 209 に供給する。

【0059】

調整帯域設定部 260 は、パワースペクトル算出部 192 から供給される環境騒音信号のパワースペクトルに基づいて、イコライジング処理部 252 において音量調整すべき音声信号の周波数帯域を設定するものである。この調整帯域設定部 260 は、例えば、パワースペクトルに基づいてスペクトルセントロイドを算出する。そして、この調整帯域設定部 260 は、その算出されたスペクトルセントロイドに、予め定められた値を乗算することによって、音声信号を増幅させる帯域の上限である最大周波数を算出する。ここにいうスペクトルセントロイドとは、環境騒音信号におけるパワースペクトルの重心に相当する周波数である。また、調整帯域設定部 260 は、その設定された最大周波数をイコライジング処理部 252 に設定する。

30

【0060】

このように、最大利得取得部 212 を設けることによって、音声信号のラウドネスレベルと環境騒音信号の騒音レベルとに基づいて最大利得を設定することができる。また、ゲイン特性傾き決定部 213 を設けることによって音声判定情報に基づいてゲイン特性の傾きの大きさを設定することができる。また、調整利得算出部 230 を設けることによって、無音状態における不自然な音量増加を防止することができるとともに、急激な音量の増減を抑制することができる。また、最小騒音レベル抽出部 214 を設けることによって、暗騒音レベルの異なる環境に応じた適切なゲイン特性を生成することができる。

40

【0061】

[環境騒音分離部の構成例]

図 4 は、本発明の第 1 の実施の形態における環境騒音分離部 180 の一構成例を示すブロック図である。ここでは、スピーカ 160 と、マイク 170 と、環境騒音分離部 180 とが示されている。また、ここでは、信号線 209 から供給される再生音声のうちの 1 つのサンプル値を $x[n]$ とし、スピーカ 160 から出力されたサンプル値 $x[n]$ の出力

50

音声を $y'[n]$ とし、出力音声 $y'[n]$ 以外の環境騒音を $s[n]$ とする。これにより、マイク 170 から供給される騒音信号は $y'[n] + s[n]$ となる。なお、スピーカ 160 およびマイク 170 は、図 1 に示したものと同様であるため、同一符号を付してここでの説明を省略する。

【0062】

環境騒音分離部 180 は、適応フィルタ 181 および減算器 182 を備える。適応フィルタ 181 は、信号線 209 からの再生音声 $x[n]$ に基づいて、マイク 170 からの騒音信号に含まれる出力音声成分 $y[n]$ を推定するものである。この適応フィルタ 181 は、減算器 182 からフィードバックされた信号に基づいて推定される室内伝達系のインパルス応答を再生音声 $x[n]$ に畳み込むことによって出力音声成分 $y[n]$ を推定する。

10

【0063】

減算器 182 は、マイク 170 から供給される騒音信号 ($y'[n] + s[n]$) と、適応フィルタ 181 により推定される出力音声成分 $y[n]$ との差分を算出する減算器である。この減算器 182 が、マイク 170 から供給される騒音信号 ($y'[n] + s[n]$) に対して適応フィルタ 181 により推定された出力音声成分 $y[n]$ を減算することにより、環境騒音信号 $e[n]$ を生成する。この減算器 182 は、その生成された環境騒音信号 $e[n]$ を、適応フィルタ 181 に供給するとともに、信号線 189 を介して環境騒音解析情報生成部 190 に供給する。

【0064】

20

このように、適応フィルタ 181 および減算器 182 を設けることによって、マイク 170 から供給された騒音信号に含まれる出力音声成分を除去して、環境騒音信号 $e[n]$ を抽出することができる。

【0065】

< 2 . 第 2 の実施の形態 >

[コンテンツ解析情報のデータ形式例]

図 5 は、本発明の第 2 の実施の形態におけるコンテンツ解析情報生成部 150 により生成されるコンテンツ解析情報のデータ形式を例示する図である。ここでは、再生音声信号 310 およびコンテンツ解析情報 320 のデータ形式が示されている。ここでは、横軸を時間軸としている。

30

【0066】

再生音声信号 310 は、コンテンツ再生部 140 により再生された音声信号における振幅の変動を示している。この再生音声信号 310 は、標準化されたサンプルのうち連続する N 個のサンプルを 1 フレームとして構成される。この再生音声信号 310 は、1 フレームごとにコンテンツ解析情報生成部 150 によって解析される。また、再生音声 $x[n]$ は、1 フレームにおける 1 サンプルの振幅の値である。

【0067】

コンテンツ解析情報 320 は、コンテンツ解析情報生成部 150 においてフレームごとに生成される再生音声信号 310 に対するコンテンツ解析情報のデータ形式を示す概念図である。このコンテンツ解析情報 320 には、フレーム番号 321、音声レベル 322、無音判定情報 323、ラウドネスレベル 324 および音声判定情報 325 が含まれる。

40

【0068】

フレーム番号 321 には、再生音声信号 310 のフレームを識別するための番号が格納される。音声レベル 322 には、音声レベル算出部 151 において算出される再生音声信号 310 の 1 フレームにおける二乗平均平方根の値 ($RMS[m]$) が格納される。

【0069】

無音判定情報 323 には、無音判定部 152 において、再生音声信号 310 が無音であるか否かが判定されて、その判定結果 ($Silence\ Flag[m]$) が格納される。この無音判定情報 323 には、例えば、無音判定部 152 により無音であると判定された場合には "True" が格納され、無音でないと判定された場合には "False" が格納

50

される。

【 0 0 7 0 】

ラウドネスレベル 3 2 4 には、ラウドネスレベル算出部 1 5 6 において算出されたラウドネスレベルの値 ($L [m]$) が格納される。このラウドネスレベル 3 2 4 は、人間の聴覚特性を考慮した音の大きさを表わすものである。そのため、視聴者には、ラウドネスレベル 3 2 4 の値が大きい程、スピーカ 1 6 0 から出力される再生音声信号 3 1 0 は聞こえ易く、ラウドネスレベル 3 2 4 の値が小さい程、スピーカ 1 6 0 から出力される再生音声信号 3 1 0 は聞こえ難い。

【 0 0 7 1 】

音声判定情報 3 2 5 には、音声判定部 1 5 4 において、再生音声信号 3 1 0 が有音声であるか無音声であるかが判定され、その判定結果 ($S p e e c h F l a g [m]$) が格納される。

10

【 0 0 7 2 】

このように、コンテンツ解析情報生成部 1 5 0 により生成されたコンテンツ解析情報 3 2 0 が音声調整部 2 0 0 に供給される。

【 0 0 7 3 】

[環境騒音解析情報のデータ形式例]

図 6 は、本発明の第 2 の実施の形態における環境騒音解析情報生成部 1 9 0 により生成される環境騒音情報のデータ形式を例示する図である。ここでは、環境騒音信号 4 1 0 および環境騒音解析情報 4 2 0 のデータ形式が示されている。ここでは、横軸を時間軸として

20

【 0 0 7 4 】

環境騒音信号 4 1 0 は、環境騒音分離部 1 8 0 により分離された環境騒音信号における振幅の変動を示している。この環境騒音信号 4 1 0 は、標本化された N サンプルを 1 フレームとして構成される。この環境騒音信号 4 1 0 は、1 フレームごとに環境騒音解析情報生成部 1 9 0 により解析される。また、環境騒音信号 $e [n]$ は、1 フレームにおける 1 サンプルの振幅の値である。

【 0 0 7 5 】

環境騒音解析情報 4 2 0 は、環境騒音解析情報生成部 1 9 0 においてフレームごとに生成される環境騒音信号 4 1 0 に対する環境騒音解析情報のデータ形式を示す概念図である。この環境騒音解析情報 4 2 0 には、フレーム番号 4 2 1、騒音レベル 4 2 2 およびパワースペクトル 4 2 3 が含まれる。

30

【 0 0 7 6 】

フレーム番号 4 2 1 には、環境騒音信号 4 1 0 のフレームを識別するための番号が格納される。騒音レベル 4 2 2 には、騒音レベル算出部 1 9 1 において算出される環境騒音信号 4 1 0 の二乗平均平方根の値 ($R M S _ e [m]$) が格納される。パワースペクトル 4 2 3 には、パワースペクトル算出部 1 9 2 において算出された k 個のパワースペクトルの値 ($s p _ e [m] [1] \sim [m] [k]$) が格納される。なお、k は、N サンプルの半分の値である。

【 0 0 7 7 】

このように、環境騒音解析情報生成部 1 9 0 により生成された環境騒音解析情報 4 2 0 が音声調整部 2 0 0 に供給される。次に、環境騒音解析情報およびコンテンツ解析情報に基づいて算出される目標利得の算出手法について図面を参照して説明する。

40

【 0 0 7 8 】

< 3 . 第 3 の実施の形態 >

[目標利得の算出手法の例]

図 7 は、本発明の第 3 の実施の形態の音声調整部 2 0 0 における目標利得の算出手法を例示する図である。ここでは、ゲイン特性 5 1 0 および 5 2 0 が示されている。また、ここでは、縦軸を音声信号に対する音量の利得とし、横軸を騒音レベルとする。

【 0 0 7 9 】

50

最大利得 (g a i n _ s u p) は、最大利得取得部 2 1 2 において取得されたゲイン特性における最大利得である。この最大利得 (g a i n _ s u p) は、最大利得取得部 2 1 2 において音声信号のラウドネスレベル (L) と環境騒音信号の騒音レベル (R M S _ e) とに基づいて決定される。この最大利得 (g a i n _ s u p) は、騒音レベル (R M S _ e) が大きいほど最大利得が大きくなり、騒音レベル (R M S _ e) が小さいほど小さくなる。一方、音声信号に対するラウドネスレベル (L) が大きいほど、視聴者にとって聞き取り易い音であるため、最大利得 (g a i n _ s u p) は小さくなる。これに対し、ラウドネスレベル (L) が小さいほど、視聴者にとって聞き取り難い音であるため、最大利得 (g a i n _ s u p) は大きくなる。

【 0 0 8 0 】

10

暗騒音レベル (R M S _ e _ i n f) は、最小騒音レベル抽出部 2 1 4 により抽出された最小の騒音レベルである。この暗騒音レベル (R M S _ e _ i n f) は、最小騒音レベル抽出部 2 1 4 において各フレームにおける騒音レベル (R M S _ e) のうち、最小の騒音レベルを抽出することによって設定される。これにより、暗騒音レベル (R M S _ e _ i n f) の異なる環境に応じたゲイン特性が生成される。

【 0 0 8 1 】

ゲイン特性 5 1 0 および 5 2 0 の傾きは、ゲイン特性傾き決定部 2 1 3 により音声判定情報 (S p e e c h F l a g) に基づいて、予め定められたゲイン特性における傾きが決定される。

【 0 0 8 2 】

20

このように、最大利得 (g a i n _ s u p) 、暗騒音レベル (R M S _ e _ i n f) およびゲイン特性 5 1 0 および 5 2 0 の傾きを定めることによって、ゲイン特性 5 1 0 および 5 2 0 が決定される。

【 0 0 8 3 】

ゲイン特性 5 1 0 は、音声判定情報 (S p e e c h F l a g) が有音声を示す場合におけるゲイン特性である。このゲイン特性 5 1 0 は、ゲイン特性 5 2 0 に比べて傾きの大きい特性を示す。これにより、音声信号が有音声である場合には、視聴者に対して音声信号を聞き取り易くすることができる。

【 0 0 8 4 】

ゲイン特性 5 2 0 は、音声判定情報 (S p e e c h F l a g) が無音声を示す場合におけるゲイン特性である。例えば、音声判定情報 (S p e e c h F l a g) が無音声を示す場合には、ゲイン特性 5 2 0 に基づいて、騒音レベル (R M S _ e) に対応する目標利得 (t a r g e t _ g a i n) が算出される。

30

【 0 0 8 5 】

このように、音声信号のラウドネスレベル (L) および環境騒音信号の騒音レベル (R M S _ e) に基づいて最大利得が決定されるため、目標利得は、ラウドネスレベル (L) が大きいほど小さくなり、騒音レベル (R M S _ e) が大きいほど大きくなる。すなわち、記録再生装置 1 0 0 は、スピーカ 1 6 0 から出力される音声信号が聞き取り易い特性であれば出力音声レベルの増加量を抑え、マイク 1 7 0 からの環境騒音が大きければ出力音声レベルの増加量を大きくする。

40

【 0 0 8 6 】

また、音声判定情報に基づいてゲイン特性の傾き選択するため、目標利得は、音声判定情報が有音声である旨を示す場合には大きくなり、無音声である旨を示す場合には小さくなる。すなわち、記録再生装置 1 0 0 は、スピーカ 1 6 0 から出力される音声信号が有音声である場合には、その音声信号を視聴者に対して聞き取り易くするために、出力音声レベルを無音声である場合に比べて大きくする。

【 0 0 8 7 】

[コンプレッサ処理部による音量調整手法の例]

図 8 は、本発明の第 3 の実施の形態におけるコンプレッサ処理部 2 5 1 による音量調整手法の例に関する図である。ここでは、利得補正特性 6 1 0 が示されている。また、こ

50

では、横軸をコンテンツ解析情報生成部 150 により算出される音声レベル (RMS) とし、縦軸をコンプレッサ処理部 251 によって増幅された音声信号の音声出力レベルとする。

【0088】

利得補正特性 610 は、コンテンツ解析情報生成部 150 によって算出された音声レベル (RMS) に応じて、コンテンツ再生部 140 により再生された音声信号の音量の増加率を補正するための利得特性の一例である。この利得補正特性 610 は、区間 1 乃至 3 の区間ごとに増加率が異なっている。

【0089】

この場合において、コンプレッサ処理部 251 では、音声信号の音声レベル (RMS) が閾値 Th_comp1 (増加率増大閾値) 未満 (区間 1) である場合には、音声レベル (RMS) が微小であるため、利得の補正を行わない。また、音声レベル (RMS) が閾値 Th_comp1 以上であり、かつ、閾値 Th_comp2 (増加率抑制閾値) 未満 (区間 2) である場合には、音声信号の音圧を効果的に増加させるために、区間 1 に比べて音声信号の音量の増加率を大きくする。さらに、音声レベル (RMS) が閾値 Th_comp2 以上である場合 (区間 3) には、音声信号における振幅の増加を抑制するため、区間 1 に比べて音声信号の音量の増加率を小さくする。

【0090】

このように、利得補正特性 610 を用いることによって、音声信号の最大振幅を抑制しつつ、音声信号の音圧を効果的に増加させることができる。次に、コンプレッサ処理部 251 により増幅された音声信号を、イコライジング処理部 252 においてさらに増幅する場合における音量調整手法について次図を参照して説明する。

【0091】

[イコライジング処理部による音声調整手法の例]

図 9 は、本発明の第 3 の実施の形態におけるイコライジング処理部 252 による音量調整手法の例に関する概念図である。ここでは、スペクトルセントロイド $C1$ および $C2$ と、これらに対応する音量調整領域 711 および 712 が示されている。ここでは、横軸を周波数とし、縦軸を音声信号の音量の利得とする。

【0092】

スペクトルセントロイド $C1$ および $C2$ は、調整帯域設定部 260 において環境騒音信号のパワースペクトル (sp_e) に基づいて算出されたスペクトル重心周波数である。このスペクトルセントロイド $C1$ および $C2$ を算出することにより、環境騒音信号におけるレベルの高い周波数成分を特定することができる。この例では、スペクトルセントロイド $C1$ は、環境騒音信号における 1 番目のフレームに対するスペクトル重心周波数であり、スペクトルセントロイド $C2$ は、2 番目のフレームに対するスペクトル重心周波数である。

【0093】

音量調整周波数 $f1$ および $f2$ は、イコライジング処理部 252 によって増幅させる音声信号の最大周波数である。この音量調整周波数 $f1$ および $f2$ は、スペクトルセントロイド $C1$ および $C2$ に一定の値をそれぞれ乗算することによって求められる最大周波数である。

【0094】

設定利得 eq_gain1' および eq_gain2' は、利得設定部 240 により設定された利得である。設定利得 eq_gain1' は、音声信号における 1 番目のフレームに対する設定利得であり、設定利得 eq_gain2' は、2 番目のフレームに対する設定利得である。

【0095】

音量調整領域 711 および 712 は、イコライジング処理部 252 において音声信号を増幅させる領域を示す概念図である。音量調整領域 711 は、音声信号における 1 番目のフレームに対する音量の増幅領域である。音量調整領域 712 は、音声信号における 2 番

10

20

30

40

50

目のフレームに対する音量の増幅領域である。

【 0 0 9 6 】

このように、イコライジング処理部 2 5 2 により音声信号を増幅させる周波数帯域を、環境騒音信号の周波数特性に基づいて算出することによって、適切な音質調整を行うことができる。

【 0 0 9 7 】

< 4 . 第 4 の実施の形態 >

[記録再生装置の動作例]

次に本発明の第 4 の実施の形態における記録再生装置 1 0 0 の動作について図面を参照して説明する。

10

【 0 0 9 8 】

図 1 0 は、本発明の第 4 の実施の形態における記録再生装置 1 0 0 の音声調整方法の処理手順例を示すフローチャートである。

【 0 0 9 9 】

まず、コンテンツ再生部 1 4 0 により、コンテンツデータが再生されることによって、音声信号が生成される (ステップ S 9 1 0)。次に、コンテンツ解析情報生成部 1 5 0 により、コンテンツ再生部 1 4 0 からの音声信号に基づいてコンテンツ解析情報が生成される (ステップ S 9 2 0)。なお、ステップ S 9 2 0 は、特許請求の範囲に記載の第 1 の音声調整情報生成手順の一例である。

20

【 0 1 0 0 】

次に、環境騒音分離部 1 8 0 により、音声調整部 2 0 0 から供給される音声信号に基づいて、マイク 1 7 0 から供給された騒音信号のうち、スピーカ 1 6 0 により出力された音声信号と環境騒音信号とが分離される (ステップ S 9 3 0)。なお、ステップ S 9 3 0 は、特許請求の範囲に記載の音声分離手順の一例である。次に、環境騒音解析情報生成部 1 9 0 により、環境騒音分離部 1 8 0 において分離された環境騒音信号に基づいて環境騒音解析情報が生成される (ステップ S 9 4 0)。なお、ステップ S 9 4 0 は、特許請求の範囲に記載の第 2 の音声調整情報生成手順の一例である。

【 0 1 0 1 】

そして、音声調整部 2 0 0 において、コンテンツ解析情報および環境騒音解析情報に基づいて音声信号の音量を調整する音声調整処理が実行される (ステップ S 9 5 0)。なお、ステップ S 9 5 0 は、特許請求の範囲に記載の音声調整手順の一例である。次に、スピーカ 1 6 0 により、音声調整部 2 0 0 において増幅された音声信号が出力される (ステップ S 9 6 0)。次に、次の音声信号のフレームがあるか否かが判断される (ステップ S 9 7 0)。次のフレームがあれば、最後のフレームまで音声処理を繰り返し、次のフレームが無ければ音声処理を終了する。

30

【 0 1 0 2 】

[音声調整部の動作例]

図 1 1 は、本発明の第 4 の実施の形態における音声調整部 2 0 0 の音声調整処理 (ステップ S 9 5 0) の処理手順例を示すフローチャートである。

【 0 1 0 3 】

まず、コンテンツ解析情報生成部 1 5 0 からのコンテンツ解析情報と、環境騒音解析情報生成部 1 9 0 からの環境騒音解析情報とが取得される (ステップ S 9 5 1)。次に、最大利得取得部 2 1 2 により、ラウドネスレベル算出部 1 5 6 からの音声信号のラウドネスレベル (L)、および、騒音レベル算出部 1 9 1 からの騒音レベル (R M S _ e) に対応する最大利得 (g a i n _ s u p) が取得される。それとともに、ゲイン特性傾き決定部 2 1 3 により、音声判定情報 (S p e e c h _ F l a g) に基づいてゲイン特性の傾きが決定される。さらに、最小騒音レベル抽出部 2 1 4 により、現在のフレームまでの騒音レベル (R M S _ e) のうち最小の騒音レベルである暗騒音レベル (R M S _ e _ i n f) が抽出される (ステップ S 9 5 2)。これにより、目標利得 (t a r g e t _ g a i n) を算出するためのゲイン特性が生成される。

40

50

【0104】

次に、目標利得算出部220により、ゲイン特性における最大利得、傾きおよび暗騒音レベルを用いることによって、現在のフレームにおける騒音レベル(RMS_e)に基づいて目標利得(target_{gain})が算出される(ステップS953)。そして、調整利得算出部230により、目標利得(target_{gain})および無音判定情報(Silence Flag)に基づいて調整利得(eq_{gain})が算出される(ステップS954)。

【0105】

次に、利得設定部240により、調整利得(eq_{gain})に基づいてコンプレッサ処理部251に利得が設定されるとともに、音声レベル算出部151により、音声レベル(RMS)がコンプレッサ処理部251に供給される。そして、コンプレッサ処理部251により、利得設定部240によって設定された利得と、音声レベル算出部151からの音声レベルとに基づいて、コンテンツ再生部140からの音声信号が増幅される(ステップS955)。

【0106】

次に、利得設定部240により、調整利得(eq_{gain})が閾値Th_{gain1}以下であるか否かが判断される(ステップS956)。そして、調整利得(eq_{gain})が閾値Th_{gain1}以下である場合には、音声調整処理が終了する。一方、調整利得(eq_{gain})が閾値Th_{gain1}より大きい場合には、利得設定部240により、調整利得(eq_{gain})に基づいてイコライジング処理部252に利得が設定される。それとともに、調整帯域設定部260により、環境騒音信号のパワースペクトルに基づいて音声信号を増幅させる周波数帯域が算出される。そして、イコライジング処理部252により、利得設定部240によって設定された利得と、調整帯域設定部260によって算出された周波数帯域とに基づいて、コンプレッサ処理部251からの音声信号が増幅される(ステップS957)。

【0107】

次に、利得設定部240により、調整利得(eq_{gain})が閾値Th_{gain2}以下であるか否かが判断される(ステップS958)。そして、調整利得(eq_{gain})が閾値Th_{gain2}以下である場合には、音声調整処理が終了する。一方、調整利得(eq_{gain})が閾値Th_{gain2}より大きい場合には、利得設定部240により、調整利得(eq_{gain})に基づいて全体音量増幅部253に利得が設定される。そして、全体音量増幅部253により、利得設定部240によって設定された利得に基づいて、イコライジング処理部252からの音声信号が増幅され(ステップS959)、音声調整処理が終了し、ステップS960の処理に進む。

【0108】

このように、本発明の実施の形態によれば、再生されるコンテンツに基づいて生成されるコンテンツ解析情報と、環境騒音信号に基づいて生成される環境騒音解析情報とに基づいて、再生されるコンテンツの音声信号を適切に調整することができる。

【0109】

また、最大利得取得部212を設けることによって、音声信号のラウドネスレベルが大きい程、最大利得が小さくなり、目標利得も小さくなることから、音声信号の音量を小さくすることができる。さらに、騒音レベルが大きい程、最大利得が大きくなり、目標利得も小さくなることから、音声信号の音量を大きくすることができる。

【0110】

また、ゲイン特性傾き決定部213を設けることによって、音声信号が有音声と判定された場合には、ゲイン特性の傾きが大きくなるため、目標利得が大きくなることから、出力音声を増大させることができる。これにより、有音声と判定された音声信号の音量を増大させることにより、出力音声を聞き取り易くすることができる。

10

20

30

40

50

【 0 1 1 1 】

なお、本発明の実施の形態は本発明を具現化するための一例を示したものであり、上述のように特許請求の範囲における発明特定事項とそれぞれ対応関係を有する。ただし、本発明は実施の形態に限定されるものではなく、本発明の要旨を逸脱しない範囲において種々の変形を施すことができる。

【 0 1 1 2 】

また、本発明の実施の形態において説明した処理手順は、これら一連の手順を有する方法として捉えてもよく、また、これら一連の手順をコンピュータに実行させるためのプログラム乃至そのプログラムを記憶する記録媒体として捉えてもよい。この記録媒体として、例えば、CD (Compact Disc)、MD (MiniDisc)、DVD、メモリカード、ブルーレイディスク (Blu-ray Disc (登録商標)) 等を用いることができる。

10

【 図面の簡単な説明 】

【 0 1 1 3 】

【 図 1 】 本発明の第 1 の実施の形態における記録再生装置の一構成例を示すブロック図である。

【 図 2 】 本発明の第 1 の実施の形態におけるコンテンツ解析情報生成部 1 5 0 および環境騒音解析情報生成部 1 9 0 の一構成例を示すブロック図である。

【 図 3 】 本発明の第 1 の実施の形態における音声調整部 2 0 0 の一構成例を示すブロック図である。

【 図 4 】 本発明の第 1 の実施の形態における環境騒音分離部 1 8 0 の一構成例を示すブロック図である。

20

【 図 5 】 本発明の第 2 の実施の形態におけるコンテンツ解析情報生成部 1 5 0 により生成されるコンテンツ解析情報のデータ形式を例示する図である。

【 図 6 】 本発明の第 2 の実施の形態における環境騒音解析情報生成部 1 9 0 により生成される環境騒音情報のデータ形式を例示する図である。

【 図 7 】 本発明の第 3 の実施の形態の音声調整部 2 0 0 における目標利得の算出手法を例示する図である。

【 図 8 】 本発明の第 3 の実施の形態におけるコンプレッサ処理部 2 5 1 による音量調整手法の例に関する図である。

【 図 9 】 本発明の第 3 の実施の形態におけるイコライジング処理部 2 5 2 による音量調整手法の例に関する概念図である。

30

【 図 1 0 】 本発明の第 4 の実施の形態における記録再生装置 1 0 0 の音声調整方法の処理手順例を示すフローチャートである。

【 図 1 1 】 本発明の第 4 の実施の形態における音声調整部 2 0 0 の音声調整処理 (ステップ S 9 5 0) の処理手順例を示すフローチャートである。

【 符号の説明 】

【 0 1 1 4 】

- 1 0 0 記録再生装置
- 1 1 0 アンテナ
- 1 2 0 チューナ部
- 1 3 0 コンテンツ記録部
- 1 4 0 コンテンツ再生部
- 1 4 9、1 8 9、2 0 1、2 0 9 信号線
- 1 5 0 コンテンツ解析情報生成部
- 1 5 1 音声レベル算出部
- 1 5 2 無音判定部
- 1 5 3 ピッチゲイン算出部
- 1 5 4 音声判定部
- 1 5 5 パワースペクトル算出部
- 1 5 6 ラウドネスレベル算出部

40

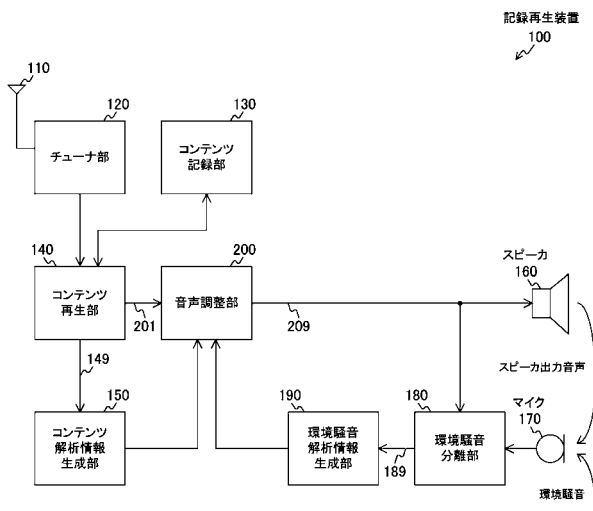
50

- 160 スピーカ
- 170 マイク
- 180 環境騒音分離部
- 181 適応フィルタ
- 182 減算器
- 190 環境騒音解析情報生成部
- 191 騒音レベル算出部
- 192 パワースペクトル算出部
- 200 音声調整部
- 210 ゲイン特性決定部
- 211 最大利得テーブル
- 212 最大利得取得部
- 213 ゲイン特性傾き決定部
- 214 最小騒音レベル抽出部
- 220 目標利得算出部
- 230 調整利得算出部
- 240 利得設定部
- 251 コンプレッサ処理部
- 252 イコライジング処理部
- 253 全体音量増幅部
- 260 調整帯域設定部

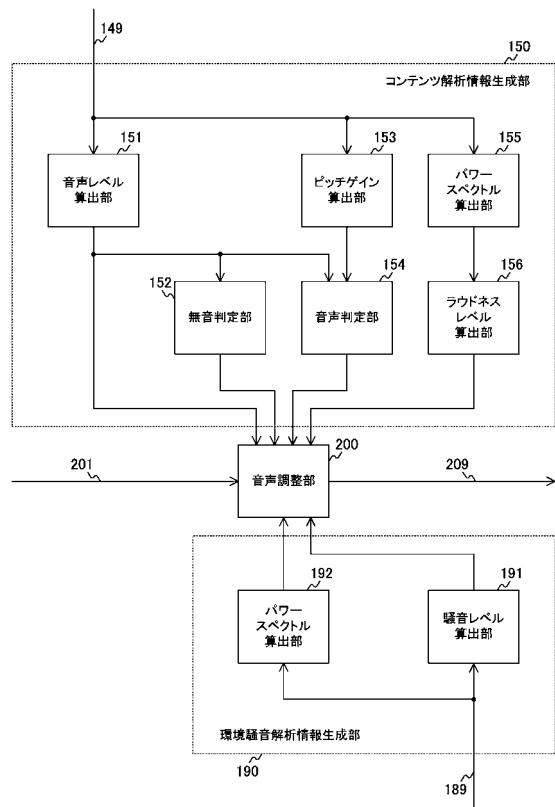
10

20

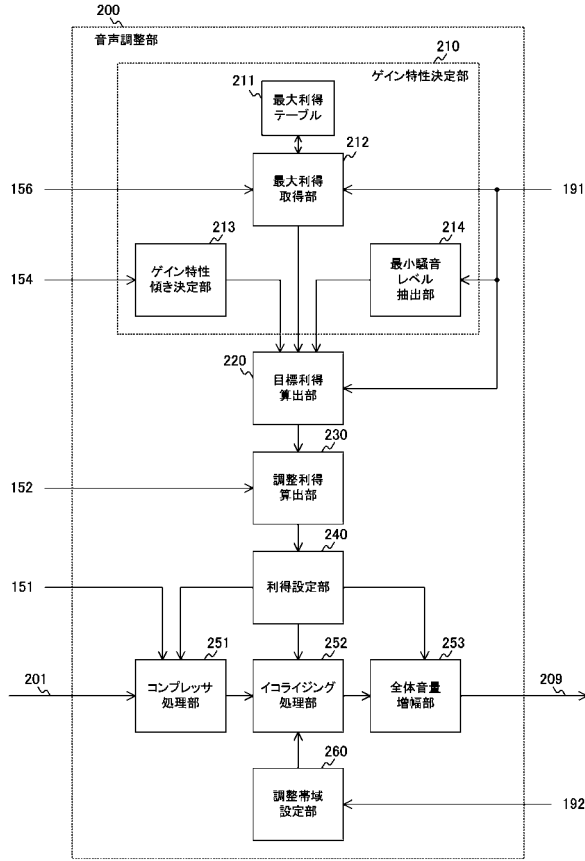
【図1】



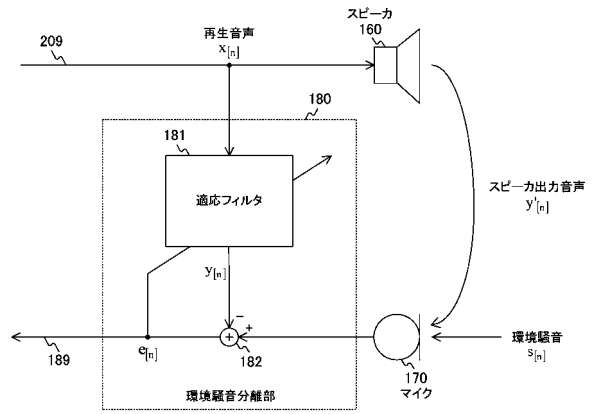
【図2】



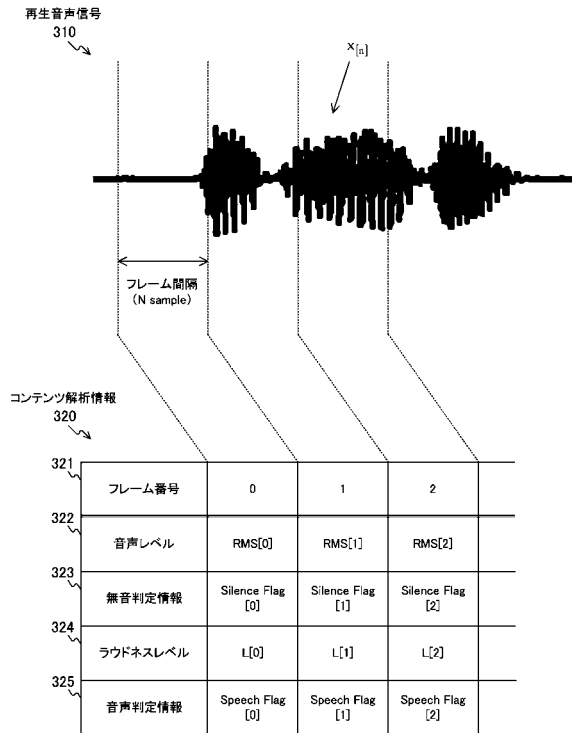
【図3】



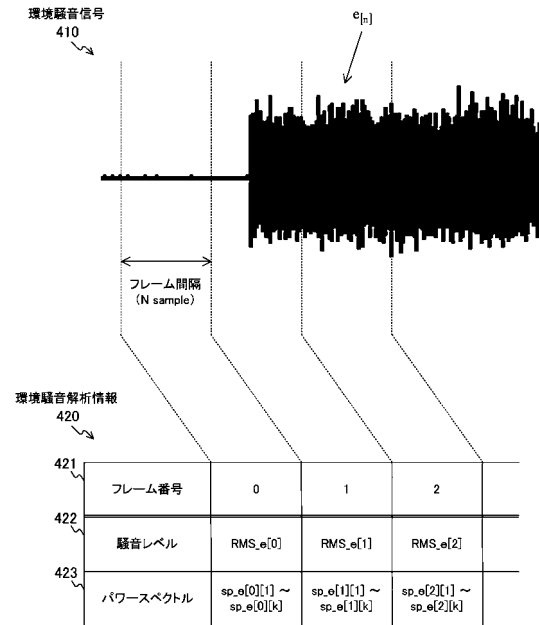
【図4】



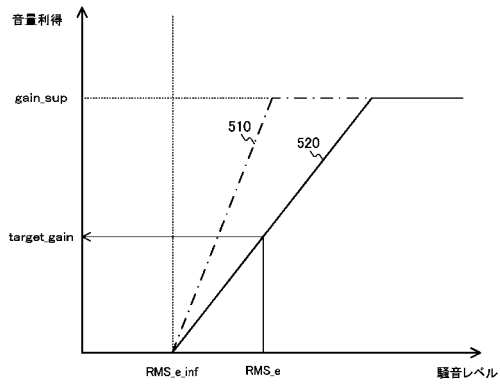
【図5】



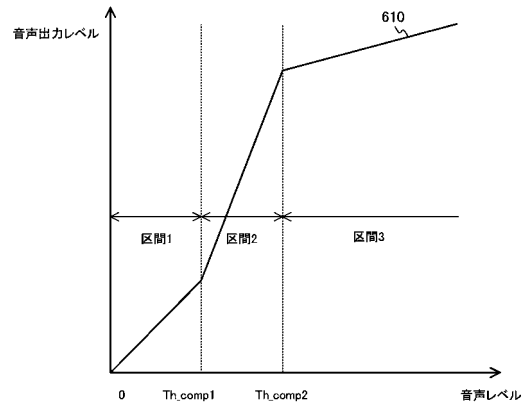
【図6】



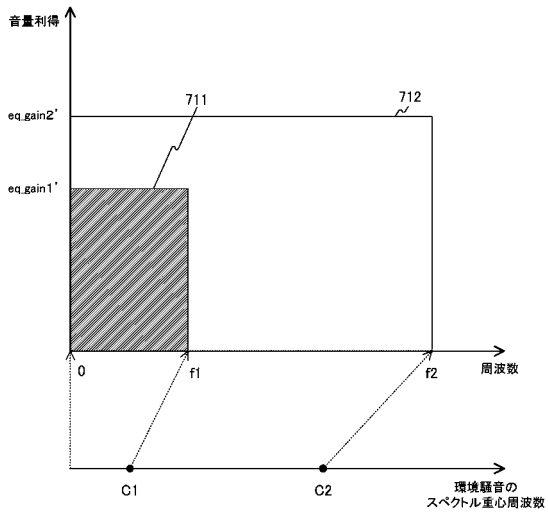
【図7】



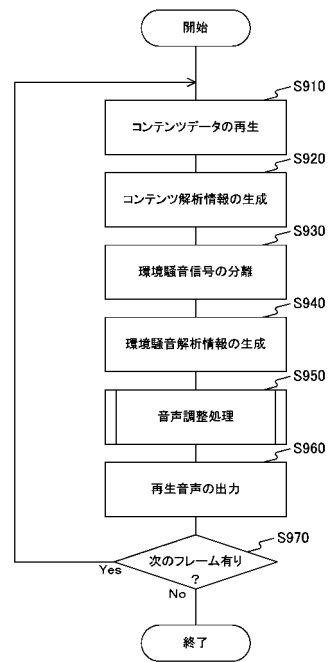
【図8】



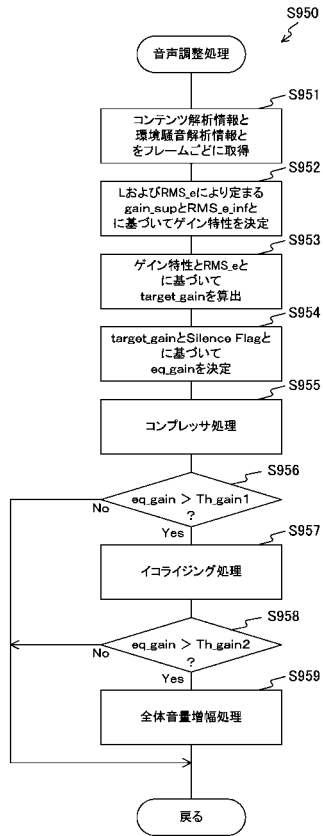
【図9】



【図10】



【図 11】



フロントページの続き

審査官 大野 弘

- (56)参考文献 特開平11-239310(JP,A)
特開平06-310962(JP,A)
特開平06-319192(JP,A)
特開2007-053510(JP,A)

- (58)調査した分野(Int.Cl., DB名)
H04R 3/00