

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2009-20291

(P2009-20291A)

(43) 公開日 平成21年1月29日(2009.1.29)

(51) Int.Cl.	F 1	テーマコード (参考)
G 1 O L 21/04 (2006.01)	G 1 O L 21/04 1 3 O Z	5 K O 2 7
H O 4 M 1/00 (2006.01)	H O 4 M 1/00 H	
G 1 O L 21/02 (2006.01)	G 1 O L 21/04 2 O O C	
H O 4 M 1/60 (2006.01)	G 1 O L 21/04 1 2 O Z	
	G 1 O L 21/02 4 O O	
審査請求 未請求 請求項の数 3 O L (全 9 頁) 最終頁に続く		

(21) 出願番号 特願2007-182458 (P2007-182458)
 (22) 出願日 平成19年7月11日 (2007.7.11)

(71) 出願人 000004075
 ヤマハ株式会社
 静岡県浜松市中区中沢町10番1号
 (74) 代理人 100064908
 弁理士 志賀 正武
 (74) 代理人 100089037
 弁理士 渡邊 隆
 (72) 発明者 川嶋 隆宏
 静岡県浜松市中区中沢町10番1号 ヤマ
 ハ株式会社内
 Fターム(参考) 5K027 BB03 DD16 HH20

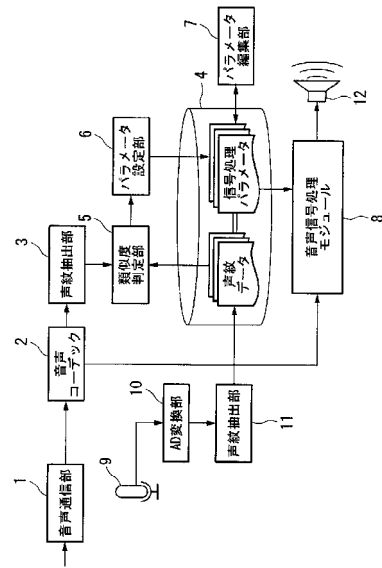
(54) 【発明の名称】 音声処理装置および通信端末装置

(57) 【要約】

【課題】 音声信号に対して、音声信号に応じた最適な処理を行うことができる音声処理装置および通信端末装置を提供する。

【解決手段】 声紋抽出部3は、音声信号の特徴を示す特徴データを生成する。音声信号処理モジュール8は、信号処理パラメータによって規定される処理内容に応じて音声信号を処理する。記憶部4は、特徴データと信号処理パラメータとを関連付けて記憶する。パラメータ設定部6は、記憶部4に記憶された特徴データの中から、声紋抽出部3によって生成された特徴データに類似するものを選択し、選択した特徴データと関連付けられている信号処理パラメータを音声信号処理モジュール8に設定する。

【選択図】 図1



【特許請求の範囲】**【請求項 1】**

音声信号に基づいて、当該音声信号の特徴を示す特徴データを生成するデータ生成手段と、

信号処理パラメータによって規定される処理内容に応じて前記音声信号を処理する信号処理手段と、

予め用意された複数の前記特徴データと、該複数の特徴データのそれぞれと関連付けられた前記信号処理パラメータとを記憶する記憶手段と、

前記記憶手段に記憶された前記特徴データの中から、前記データ生成手段によって生成された前記特徴データに類似するものを選択し、選択した前記特徴データと関連付けられている前記信号処理パラメータを前記信号処理手段に設定する設定手段と、

を備えたことを特徴とする音声処理装置。

【請求項 2】

外部から音声を入力し、音声信号を出力する音声入力手段と、

ユーザからの指示に応じて前記信号処理パラメータを編集するパラメータ編集手段と、をさらに備え、

前記データ生成手段は、前記音声入力手段から出力された前記音声信号の特徴を示す前記特徴データを生成し、

前記記憶手段は、前記音声入力手段から出力された前記音声信号の特徴を示す前記特徴データと、前記編集手段によって編集された前記信号処理パラメータとを関連付けて記憶する

ことを特徴とする請求項 1 に記載の音声処理装置。

【請求項 3】

請求項 1 または請求項 2 に記載の音声処理装置と、

外部の通信装置と通信を行い、前記音声信号を受信する受信手段と、

を備えたことを特徴とする通信端末装置。

【発明の詳細な説明】**【技術分野】****【0001】**

本発明は、音声信号を処理する音声処理装置および通信端末装置に関する。

【背景技術】**【0002】**

音声を聞き取り易くする技術として、例えば特許文献 1 には、通話相手の電話番号に応じて受話音質を自動的に切り替えることができるようにして、受話音声を聞き取り易い状態に調整することが可能な通信端末装置が記載されている。

【特許文献 1】特開 2005 - 136788 号公報

【発明の開示】**【発明が解決しようとする課題】****【0003】**

しかし、上記の通信端末装置では、電話番号に対応させて受話音の調整状態を登録しておく必要があるため、電話番号が登録されていない相手から電話が掛かってきた場合には、受話音を調整することができない。すなわち、従来技術では、受信された音声信号に対して、音声信号に応じた最適な処理を行うことができない状況が存在していた。

【0004】

本発明は、上述した課題に鑑みてなされたものであって、音声信号に対して、音声信号に応じた最適な処理を行うことができる音声処理装置および通信端末装置を提供することを目的とする。

【課題を解決するための手段】**【0005】**

本発明は、上記の課題を解決するためになされたもので、音声信号に基づいて、当該音

10

20

30

40

50

声信号の特徴を示す特徴データを生成するデータ生成手段と、信号処理パラメータによって規定される処理内容に応じて前記音声信号を処理する信号処理手段と、予め用意された複数の前記特徴データと、該複数の特徴データのそれぞれと関連付けられた前記信号処理パラメータとを記憶する記憶手段と、前記記憶手段に記憶された前記特徴データの中から、前記データ生成手段によって生成された前記特徴データに類似するものを選択し、選択した前記特徴データと関連付けられている前記信号処理パラメータを前記信号処理手段に設定する設定手段と、を備えたことを特徴とする音声処理装置である。

【0006】

また、本発明の音声処理装置は、外部から音声を入力し、音声信号を出力する音声入力手段と、ユーザからの指示に応じて前記信号処理パラメータを編集するパラメータ編集手段と、をさらに備え、前記データ生成手段は、前記音声入力手段から出力された前記音声信号の特徴を示す前記特徴データを生成し、前記記憶手段は、前記音声入力手段から出力された前記音声信号の特徴を示す前記特徴データと、前記編集手段によって編集された前記信号処理パラメータとを関連付けて記憶することを特徴とする。

10

【0007】

また、本発明は、上記の音声処理装置と、外部の通信装置と通信を行い、前記音声信号を受信する受信手段と、を備えたことを特徴とする通信端末装置である。

【発明の効果】

【0008】

本発明によれば、予め記憶された特徴データの中から、音声信号に基づいて生成された特徴データに類似するものが選択され、その特徴データと関連付けられている信号処理パラメータが設定されるので、音声信号に対して、音声信号に応じた最適な処理を行うことができるという効果が得られる。

20

【発明を実施するための最良の形態】

【0009】

以下、図面を参照し、本発明の実施形態を説明する。図1は、本発明の一実施形態による音声処理装置（例えば携帯電話等の通信端末装置）の構成を示している。なお、図1は、音声通信（特に音声信号の受信）に係る構成のみを示しており、他の構成は本実施形態の説明に不要であるので図示していない。

【0010】

音声通信部1は外部の通信装置と通信を行い、音声信号を受信する。音声コーデック2は、音声通信部1によって受信された、符号化されている音声信号をリニアの音声信号へ変換（デコード）するモジュールである。音声信号の符号化方式として、例えばQCELP（Qualcomm Code Excited Linear Prediction）やAMR（Advanced Multi Rate Codec）等がある。

30

【0011】

声紋抽出部3は、音声コーデック2から出力されたりニアの音声信号を分析し、音声信号の特徴を示す声紋データ（特徴データ）を抽出する。声紋データは長時間スペクトル法により求められる。つまり、音声信号に対して、FFT（Fast Fourier Transform）により周波数分析を逐次行い、周波数値を毎回累積加算していく。これをある区間まで続け、最後にその区間までの累積回数で周波数値を除算することにより、声紋データを生成することができる。

40

【0012】

記憶部4は、声紋データと、音声信号処理モジュール8による処理の内容を規定する信号処理パラメータとを関連付けて予め記憶している。類似度判定部5は、声紋抽出部3によって抽出された声紋データと、記憶部4に予め記憶されている声紋データとの類似度を判定する。類似度の判定方法にはどのような方法を用いてもよく、例えばメルケプストラム分析を用いて求めたメルケプストラムを時系列特徴ベクトルとし、そのベクトル距離を算出する方法等がある。

【0013】

50

パラメータ設定部 6 は、類似度判定部 5 による類似度判定の結果に基づいて、記憶部 4 に記憶された声紋データの中から、声紋抽出部 3 によって抽出された声紋データとの類似度が高い声紋データを選択し、その声紋データと関連付けられている信号処理パラメータを記憶部 4 から読み出して音声信号処理モジュール 8 に設定する。パラメータ編集部 7 は、例えば通信端末装置の操作キー（図示せず）の操作に応じてディスプレイ（図示せず）上に表示される GUI（Graphical User Interface）機能を利用して入力されたユーザからの指示に応じて信号処理パラメータを編集する。なお、パラメータ編集部 7 は、通信端末装置のインターフェース（図示せず）を介して接続された PC 等の外部機器により実現されてもよい。音声信号処理モジュール 8 は、音声コーデック 2 から出力された音声信号に対して、パラメータ設定部 6 によって設定された信号処理パラメータによって規定される処理内容に応じた処理を行うモジュールである。これによって、音質が改善され、音声が聞き取りやすくなる。

10

【 0 0 1 4 】

マイクロフォン 9 は音声を入力して音声信号に変換し、出力する。AD変換部 10 は、マイクロフォン 9 から出力されたアナログ音声信号をデジタル音声信号に変換する。声紋抽出部 11 は、AD変換部 10 から出力された音声信号を分析し、声紋抽出部 3 と同様に声紋データ（特徴データ）を抽出する。声紋抽出部 11 によって抽出された声紋データは、パラメータ編集部 7 によって編集された信号処理パラメータと共に記憶部 4 に格納される。スピーカ 12 は、音声信号処理モジュール 8 によって処理された音声信号に基づいた音声を発生する。

20

【 0 0 1 5 】

図 2 は音声信号処理モジュール 8 の構成を示している。高音域補填処理部 81 は、入力された音声信号に対して、音声コーデック 2 の帯域制限で失われた高域部分を補完すると共に、音声全体のざらつき感をなくす処理を行う。エンハンサー処理部 82 は、高音域補填処理部 81 から出力された音声信号に対して、高域倍音を強調することで、音にメリハリをつけ、音が前に出るような効果を出す処理を行う。

【 0 0 1 6 】

ダイナミックレンジコンプレッサー 83 は、エンハンサー処理部 82 から出力された音声信号に対して、特定のレベル（閾値）を超えた信号レベルを動的に減衰させる処理を行う。音量が大きい場合には、音量を押し下げ、その分、全体の音量レベルを持ち上げることで、トータルの音量差が均一になる。前段のエンハンサー処理部 82 によって、ピーク値が上がってしまう処理が行われた場合でも、音量を確保しつつ歪感のない音声を得ることができる。イコライザー 84 は、ダイナミックレンジコンプレッサー 83 から出力された音声信号の周波数帯域をバンド単位で補正する。上記の高音域補填処理部 81 ~ イコライザー 84 に対し、信号処理パラメータを設定し、その信号処理パラメータによって規定される処理を行わせることができるようになっている。

30

【 0 0 1 7 】

図 3 は、記憶部 4 に記憶される声紋データと信号処理パラメータの対応関係を示している。図 3 に示すように、声紋データ 300 に対して、高音域補填処理部 81 の処理内容を規定する信号処理パラメータ 310（図 4（a）参照）と、エンハンサー処理部 82 の処理内容を規定する信号処理パラメータ 320（図 4（b）参照）と、ダイナミックレンジコンプレッサー 83 の処理内容を規定する信号処理パラメータ 330（図 4（c）参照）と、イコライザー 84 の処理内容を規定する信号処理パラメータ 340（図 4（d）参照）とがそれぞれ関連付けられている。

40

【 0 0 1 8 】

例えば、「Type A」の声紋データ 300 に対して、図 4（a）に示す各パラメータで構成される「DB_set A」の信号処理パラメータ 310 と、図 4（b）に示す各パラメータで構成される「EH_set A」の信号処理パラメータ 320 と、図 4（c）に示す各パラメータで構成される「DR_set A」の信号処理パラメータ 330 と、図 4（d）に示す各パラメータで構成される「EQ_set A」の信号処理パラメータ 340 とが関連付けられている。

50

【 0 0 1 9 】

次に、図 5 を参照しながら、音声通信時（通話時）の音声処理装置の動作を説明する。音声処理装置を使用するユーザからのダイヤル発信または通話相手からのダイヤル発信を検知することで通話が開始される。音声通信部 1 は音声信号を受信し、音声コーデック 2 へ出力する。音声コーデック 2 は、符号化されている音声信号をリニアの音声信号へ変換する（ステップ S 1 0 0）。声紋抽出部 3 はこの音声信号から声紋データを抽出する（ステップ S 1 1 0）。

【 0 0 2 0 】

類似度判定部 5 は、声紋抽出部 3 によって抽出された声紋データと、記憶部 4 に予め記憶されている声紋データとの類似度を判定する。パラメータ設定部 6 は、この判定の結果に基づいて、記憶部 4 に記憶された声紋データの中から、声紋抽出部 3 によって抽出された声紋データとの類似度が高い（類似度が所定値以上の）声紋データを検索する（ステップ S 1 2 0）。

10

【 0 0 2 1 】

判定された類似度が所定値以上の声紋データが見つかった場合（ステップ S 1 3 0 で Y E S の場合）、処理はステップ S 1 4 0 に進む。また、判定された類似度が所定値以上の声紋データが見つからなかった場合（ステップ S 1 3 0 で N O の場合）、処理はステップ S 1 7 0 に進む。

【 0 0 2 2 】

ステップ S 1 4 0 では、パラメータ設定部 6 は、最も類似度が高い声紋データに関連付けられている信号処理パラメータを記憶部 4 から読み出す（ステップ S 1 4 0）。また、ステップ S 1 7 0 では、パラメータ設定部 6 は、予め用意されているデフォルトの信号処理パラメータを記憶部 4 から読み出す（ステップ S 1 7 0）。続いて、パラメータ設定部 6 は、読み出した信号処理パラメータを音声信号処理モジュール 8 に設定する（ステップ S 1 5 0）。

20

【 0 0 2 3 】

通話終了の合図があるまで、ステップ S 1 4 0 またはステップ S 1 7 0 で取得された信号処理パラメータが音声信号処理モジュール 8 に設定された状態が維持される。あるいは、通話の途中で通話相手が変わった場合でも音声の聞き取り易さを維持するため、所定時間毎にステップ S 1 0 0 へ戻って上記の処理を再度行うようにしてもよい。通話終了の合図があった場合、音声信号の受信が終了し（ステップ S 1 6 0）、通話に係る一連の処理が終了する。

30

【 0 0 2 4 】

以上のように、予め記憶部 4 に記憶された声紋データの中から、通話相手の音声信号から抽出した声紋データに類似するものが選択され、その声紋データと関連付けられている信号処理パラメータが音声信号処理モジュール 8 に設定されるので、受信された音声信号に対して、その音声信号に応じた最適な処理を行うことができる。特に、通話相手から初めて電話が掛かってきた場合でも、その通話相手の音声に係る音声信号から抽出される声紋データに類似したものが記憶部 4 に記憶されていれば、その音声信号に最適な信号処理パラメータに基づいた処理を行うことができる。

40

【 0 0 2 5 】

したがって、通話相手個人の声質に応じて、通話相手の声質に最適な信号処理パラメータを音声信号処理モジュール 8 に与え、音声を聞き取りやすくすることができる。例えば、静かに話す相手の声をより強調したり、濁声をより柔らかい音にしたりすることができる。

【 0 0 2 6 】

次に、図 6 を参照しながら、声紋データを記憶部 4 に登録するときの音声処理装置の動作を説明する。まず、ユーザからの指示により、声紋データを登録するためのモードに動作モードが変更される（ステップ S 2 0 0）。続いて、登録したい音声マイクロフォン 9 に入力される。A/D変換部 10 は、その音声に基づいたアナログ音声信号をデジタル

50

音声信号に変換する。声紋抽出部 11 はその音声信号を分析し、声紋データを抽出する（ステップ S 2 1 0）。抽出された声紋データは記憶部 4 に格納される。

【 0 0 2 7 】

続いて、信号処理パラメータを編集するための指示がユーザによって適宜入力される。このとき、信号処理パラメータが、入力した音声に対応した声紋データに最も適したものとなるように、GUI 機能を利用して編集の指示が入力される。パラメータ編集部 7 は、この指示に応じて信号処理パラメータを編集する（ステップ S 2 2 0）。パラメータ編集部 7 は、編集した信号処理パラメータを、声紋抽出部 11 によって抽出されて記憶部 4 に格納された声紋データと関連付けて記憶部 4 に格納する（ステップ S 2 3 0）。

【 0 0 2 8 】

声紋データの登録を続ける場合（ステップ S 2 4 0 で N O の場合）には、ステップ S 2 1 0 に戻って、音声の入力から処理が再度実行される。また、声紋データの登録を終了する場合（ステップ S 2 4 0 で Y E S の場合）には、声紋データの登録に係る一連の処理が終了する。

【 0 0 2 9 】

以上、図面を参照して本発明の実施形態について詳述してきたが、具体的な構成は上記の実施形態に限られるものではなく、本発明の要旨を逸脱しない範囲の設計変更等も含まれる。

【 図面の簡単な説明 】

【 0 0 3 0 】

【 図 1 】 本発明の一実施形態による音声処理装置の構成を示すブロック図である。

【 図 2 】 本発明の一実施形態による音声信号処理モジュールの構成を示すブロック図である。

【 図 3 】 本発明の一実施形態による記憶部に記憶される声紋データと信号処理パラメータの対応関係を示す図である。

【 図 4 】 本発明の一実施形態における信号処理パラメータの内容を示す図である。

【 図 5 】 本発明の一実施形態による音声処理装置の動作の手順を示すフローチャートである。

【 図 6 】 本発明の一実施形態による音声処理装置の動作の手順を示すフローチャートである。

【 符号の説明 】

【 0 0 3 1 】

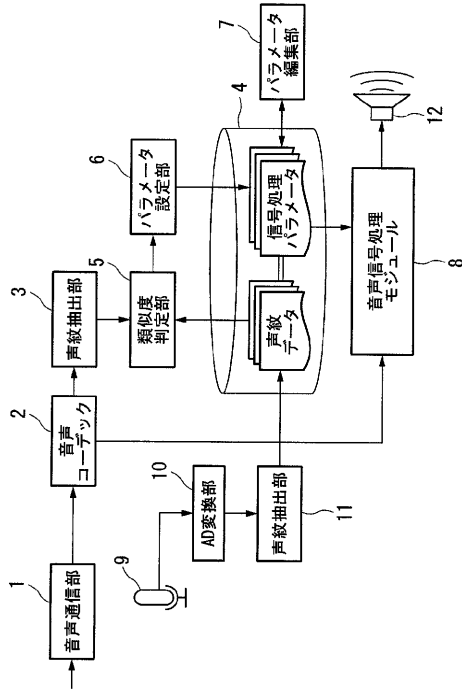
1・・・音声通信部、2・・・音声コーデック、3, 11・・・声紋抽出部（データ生成手段）、4・・・記憶部（記憶手段）、5・・・類似度判定部、6・・・パラメータ設定部（設定手段）、7・・・パラメータ編集部（パラメータ編集手段）、8・・・音声信号処理モジュール（信号処理手段）、9・・・マイクロフォン（音声入力手段）、10・・・A/D変換部、12・・・スピーカ、81・・・高音域補填処理部、82・・・エンハンサー処理部、83・・・ダイナミックレンジコンプレッサー、84・・・イコライザー

10

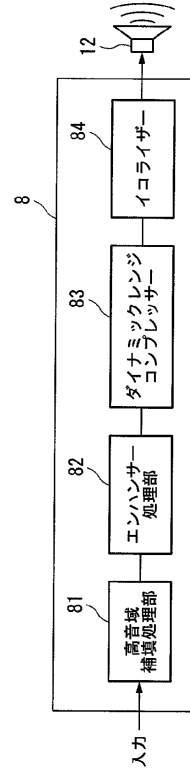
20

30

【図1】



【図2】



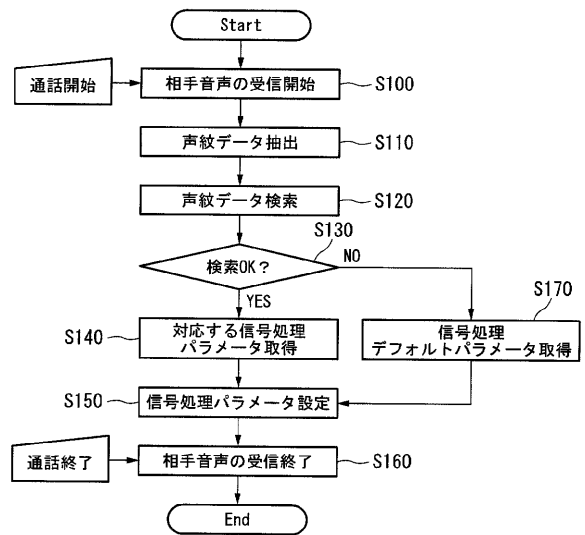
【図3】

声紋データ	信号処理パラメータ			
	高域補填	エンハンサ	DRC	EQ
Type A	DB_set A	EH_set A	DR_set A	EQ_set A
Type B	DB_set B	EH_set B	DR_set B	EQ_set B
⋮	⋮	⋮	⋮	⋮
Type Z	DB_set Z	EH_set Z	DR_set Z	EQ_set Z
デフォルト	DB_set 0	EH_set 0	DR_set 0	EQ_set 0
	300	310	320	330
				340

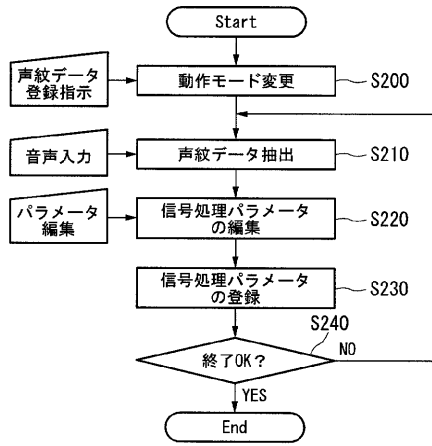
【図4】

- (a) DB_set A
DB gain=+3.4dB
Harmonics Type=3
Mid range Delay=9 sample
⋮
- (b) EH_set A
Algorithm=2
Scatter Gain=0.78
Wide Level=67%
⋮
- (c) DR_set A
Attack=2ms
Release=10ms
Threshold=-5dB
Ratio=1.0
⋮
- (d) EQ_set A
Band1 Filter Type=LPF
Band1 Gain=+3dB
Band1 Cutoff Freq=800Hz
Band1 Q=1.2
⋮

【図5】



【 図 6 】



フロントページの続き

(51)Int.Cl.

F I

テーマコード(参考)

G 1 0 L 21/02 3 0 1 Z

G 1 0 L 21/02 3 0 2 B

H 0 4 M 1/60 D