

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第5224552号
(P5224552)

(45) 発行日 平成25年7月3日(2013.7.3)

(24) 登録日 平成25年3月22日(2013.3.22)

(51) Int.Cl.		F I		
G 1 O L 13/02	(2013.01)	G 1 O L 13/02	1 1 O Z	
G 1 O L 13/047	(2013.01)	G 1 O L 13/04	B	
G 1 O L 13/06	(2013.01)	G 1 O L 13/06	1 2 O A	

請求項の数 5 (全 15 頁)

(21) 出願番号	特願2010-183923 (P2010-183923)	(73) 特許権者	599008654 伊福部 達 東京都杉並区和泉3-59-15-205
(22) 出願日	平成22年8月19日(2010.8.19)	(73) 特許権者	310010575 地方独立行政法人北海道立総合研究機構 北海道札幌市北区北19条西11丁目1番地9
(65) 公開番号	特開2012-42722 (P2012-42722A)	(73) 特許権者	391026106 株式会社電制 北海道江別市工栄町8番地の13
(43) 公開日	平成24年3月1日(2012.3.1)	(74) 代理人	110000121 アイアット国際特許業務法人
審査請求日	平成23年6月8日(2011.6.8)	(72) 発明者	伊福部 達 東京都杉並区和泉3丁目59番15号 シ ティハウス永福町205号室 最終頁に続く

(54) 【発明の名称】 音声生成装置およびその制御プログラム

(57) 【特許請求の範囲】

【請求項1】

基本周波数の音声データを生成する音源生成手段と、

入力手段の操作に基づいて第1ホルマント周波数と第2ホルマント周波数の2次元平面上におけるX座標値とY座標値および第3ホルマント周波数と第4ホルマント周波数の2次元平面上におけるX座標値とY座標値を検出する座標値検出手段と、

前記音源生成手段で生成された前記基本周波数の音声データを、前記第1ホルマント周波数と第2ホルマント周波数の2次元平面上における前記X座標値に対応する前記第1ホルマント周波数で共振させる第1の共振手段と、

前記第1の共振手段により共振された前記音声データを、前記第1ホルマント周波数と前記第2ホルマント周波数の2次元平面上における前記Y座標値に対応する前記第2ホルマント周波数で共振させる第2の共振手段と、

前記第2の共振手段により共振された前記音声データを、前記第3ホルマント周波数と前記第4ホルマント周波数の2次元平面上における前記X座標値に対応する前記第3ホルマント周波数で共振させる第3の共振手段と、

前記第3の共振手段により共振された前記音声データを、前記第3ホルマント周波数と前記第4ホルマント周波数の2次元平面上における前記Y座標値に対応する前記第4ホルマント周波数で共振させる第4の共振手段と、

前記第4の共振手段により共振された前記音声データを出力する出力手段と、
備えることを特徴とする音声生成装置。

10

20

【請求項 2】

請求項 1 に記載の音声生成装置であって、

前記座標値検出手段は、前記入力手段の操作に基づいて鼻音の有無を判定し、その判定結果を出力するものであり、

前記音源生成手段からの基本周波数の音声データのうち、高周波を通過させ、遮断周波数より低い周波数の帯域を減衰させて前記第 1 の共振手段に出力するハイパスフィルタと

前記座標値検出手段から鼻音があるという判定結果が入力された場合に、前記ハイパスフィルタから出力された低周波成分が除去された音声データを、所定の共振周波数で共振させて鼻音となる音声データを生成して出力する鼻音生成手段と、

前記第 4 の共振手段により共振された音声データと前記鼻音生成手段により生成された音声データを加算して出力する第 1 の加算手段と、を有し、

前記出力手段は、前記第 1 の加算手段により出力された音声データを出力することを特徴とする音声生成装置。

【請求項 3】

請求項 1 又は 2 に記載の音声生成装置であって、

前記座標値検出手段は、前記入力手段の操作に基づいて乱流音の有無を判定し、その判定結果を出力するものであり、

前記座標値検出手段から乱流音があるという判定結果が入力された場合に、子音を合成するための音源を、第 1 ～ 第 4 のホルマント周波数それぞれに対して個別に設けられた疑似乱数発生器によってそれぞれ生成すると共に、前記座標値検出手段からの入力情報に対応するホルマント周波数でそれぞれ共振させた音声データを出力する乱流音生成手段と、

前記乱流音生成手段から出力される音声データと前記第 1 の加算手段により出力される音声データを加算する第 2 の加算手段とを有し、

前記出力手段は、前記第 2 の加算手段により出力される音声データを出力することを特徴とする音声生成装置。

【請求項 4】

請求項 1 から 3 のいずれか 1 項に記載の音声生成装置であって、

前記音声生成装置は、ジャイロセンサを内蔵した操作バーから情報を受信できるように構成されると共に前記操作バーの空間上の動作に応じて、それぞれのホルマント周波数の 2 次元平面上における分布に対応させた情報を記憶しており、前記操作バーのジャイロセンサが検出した X 軸、Y 軸、Z 軸方向の加速度に基づいて音声データを生成することを特徴とする音声生成装置。

【請求項 5】

基本周波数の音声データを生成する音源生成ステップと、

入力手段の操作に基づいて第 1 ホルマント周波数と第 2 ホルマント周波数の 2 次元平面上における X 座標値と Y 座標値および第 3 ホルマント周波数と第 4 ホルマント周波数の 2 次元平面上における X 座標値と Y 座標値を検出する座標値検出ステップと、

前記音源生成ステップで生成された前記基本周波数の音声データを、前記第 1 ホルマント周波数と第 2 ホルマント周波数の 2 次元平面上における前記 X 座標値に対応する前記第 1 ホルマント周波数で共振させる第 1 の共振ステップと、

前記第 1 の共振ステップにより共振された前記音声データを、前記第 1 ホルマント周波数と前記第 2 ホルマント周波数の 2 次元平面上における前記 Y 座標値に対応する前記第 2 ホルマント周波数で共振させる第 2 の共振ステップと、

前記第 2 の共振ステップにより共振された前記音声データを、前記第 3 ホルマント周波数と前記第 4 ホルマント周波数の 2 次元平面上における前記 X 座標値に対応する前記第 3 ホルマント周波数で共振させる第 3 の共振ステップと、

前記第 3 の共振ステップにより共振された前記音声データを、前記第 3 ホルマント周波数と前記第 4 ホルマント周波数の 2 次元平面上における前記 Y 座標値に対応する前記第 4 ホルマント周波数で共振させる第 4 の共振ステップと、

10

20

30

40

50

前記第4の共振ステップにより共振された前記音声データを出力する出力ステップと、を含む処理をコンピュータに実行させることを特徴とするプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、音声生成装置およびその制御プログラムに関し、特に、簡単な操作で、リアルタイムに音声を生成することが可能な音声生成装置およびその制御プログラムに関する。

【背景技術】

【0002】

人間が互いに意思や感情を伝達し合うコミュニケーションの手段には、言語、文字、その他の視覚や聴覚に訴える身振り、表情、声などの手段があるが、日常生活においては、音声による会話が果たす役割は非常に大きい。音声は、コミュニケーションの手段として言語情報を伝達するだけでなく、その音質によって、話し手が誰であるかという情報や、音楽的情報をも表現することができる。

【0003】

人が声を出す際には、肺の呼吸運動によって与えられる肺呼気（空気の流れ）を喉頭の中央にある声帯において振動エネルギーに変換し、この振動によって音声の基本となる音（声）を生成している。これを喉頭原音あるいは声帯原音と呼ぶ。

【0004】

一方、人は、この喉頭原音（原音の基本周波数は、成人男性で約120Hz、女性で約240Hz程度）を、声道と呼ばれる咽頭、口腔、鼻腔などで共鳴させることによって修飾し、さらに唇、舌、顎などの助力によって音色に変化を与えることで所望の音声波形を生成している。これを構音と呼ぶ。

【0005】

しかし、唇、舌、顎などの欠損や変形、脳性麻痺や脳血管障害、筋ジストロフィーやパーキンソン病等の筋・神経系難病などにより、唇、舌、顎などを使った構音機能に何らかの異常が生じると、音声会話に必要な音色の変化を十分に生成することができないという発声障害を引き起こしてしまう。

【0006】

そこで、近年、このような発声障害を支援するいくつかの機器が提案されている。例えば、ユーザがスイッチを操作することによって予め決められた言葉を発する装置、発話内容を第三者が予め録音しておき、それを再生する装置、あるいは、キー操作によって発話内容を入力すると、その発話内容を音声合成して発する装置がある。

【先行技術文献】

【特許文献】

【0007】

【特許文献1】特開2005-241744号公報

【特許文献2】特開平11-237946号公報

【発明の概要】

【発明が解決しようとする課題】

【0008】

しかしながら、従来の発声障害支援装置は、発話内容が限定され、リアルタイムな発話が困難であり、さらに感情（抑揚）を表現することが困難であった。

【0009】

本発明はこのような状況に鑑みてなされたものであり、その目的は、簡単な操作で、リアルタイムに音声を生成することが可能な音声生成装置およびその制御プログラムを提供することである。

【課題を解決するための手段】

【0010】

10

20

30

40

50

本発明の一側面は、音声生成装置に関するものである。すなわち、本発明の音声生成装置は、基本周波数の音声データを生成する音源生成手段と、入力手段の操作に基づいて第1ホルマント周波数と第2ホルマント周波数の2次元平面上におけるX座標値とY座標値および第3ホルマント周波数と第4ホルマント周波数の2次元平面上におけるX座標値とY座標値を検出する座標値検出手段と、音源生成手段で生成された基本周波数の音声データを、第1ホルマント周波数と第2ホルマント周波数の2次元平面上におけるX座標値に対応する第1ホルマント周波数で共振させる第1の共振手段と、第1の共振手段により共振された音声データを、第1ホルマント周波数と第2ホルマント周波数の2次元平面上におけるY座標値に対応する第2ホルマント周波数で共振させる第2の共振手段と、第2の共振手段により共振された音声データを、第3ホルマント周波数と第4ホルマント周波数の2次元平面上におけるX座標値に対応する第3ホルマント周波数で共振させる第3の共振手段と、第3の共振手段により共振された音声データを、第3ホルマント周波数と第4ホルマント周波数の2次元平面上におけるY座標値に対応する第4ホルマント周波数で共振させる第4の共振手段と、第4の共振手段により共振された音声データを出力する出力手段と、備えることを特徴とする。

10

【0011】

本発明の一側面は、プログラムに関するものである。すなわち、基本周波数の音声データを生成する音源生成ステップと、入力手段の操作に基づいて第1ホルマント周波数と第2ホルマント周波数の2次元平面上におけるX座標値とY座標値および第3ホルマント周波数と第4ホルマント周波数の2次元平面上におけるX座標値とY座標値を検出する座標値検出ステップと、前記音源生成ステップで生成された前記基本周波数の音声データを、前記第1ホルマント周波数と第2ホルマント周波数の2次元平面上における前記X座標値に対応する前記第1ホルマント周波数で共振させる第1の共振ステップと、前記第1の共振ステップにより共振された前記音声データを、前記第1ホルマント周波数と前記第2ホルマント周波数の2次元平面上における前記Y座標値に対応する前記第2ホルマント周波数で共振させる第2の共振ステップと、前記第2の共振ステップにより共振された前記音声データを、前記第3ホルマント周波数と前記第4ホルマント周波数の2次元平面上における前記X座標値に対応する前記第3ホルマント周波数で共振させる第3の共振ステップと、前記第3の共振ステップにより共振された前記音声データを、前記第3ホルマント周波数と前記第4ホルマント周波数の2次元平面上における前記Y座標値に対応する前記第4ホルマント周波数で共振させる第4の共振ステップと、前記第4の共振ステップにより共振された前記音声データを出力する出力ステップと、を含む処理をコンピュータに実行させることを特徴とする。

20

30

【発明の効果】

【0012】

本発明によれば、簡単な操作で、リアルタイムに音声を生成することが可能な音声生成装置およびその制御プログラムを提供することができる。

【図面の簡単な説明】

【0013】

【図1】母音の発生の仕組みについて説明するための図である。

40

【図2】第1の実施の形態に係る音声生成装置の構成例を示す図である。

【図3】音声生成GUIの表示例を示す図である。

【図4】音声生成装置の機能構成例を示すブロック図である。

【図5】音声生成処理を説明するフローチャートである。

【図6】第2の実施の形態に係る操作バーと携帯型の音声生成装置の接続例を示す図である。

【図7】操作バーと携帯型の音声生成装置の内部の構成例を示すブロック図である。

【図8】音声生成装置の他の機能構成例を示すブロック図である

【図9】本発明を適用したコンピュータの構成例を示すブロック図である。

【発明を実施するための形態】

50

【 0 0 1 4 】

以下、本発明の実施の形態について図面を参照して詳細に説明する。

【 0 0 1 5 】

[母音の発声の仕組み]

図 1 は、母音の発声の仕組みについて説明するための図である。

【 0 0 1 6 】

人間から発声される音声は、肺から押し出される呼気が唇から放射されるまでに通過する声道(喉頭、咽頭、口腔、鼻腔からなる共鳴腔)の形によって作り出される。つまり、声帯から唇までの声道を 1 つの音響管と考えると、声道で共鳴現象が生じることで発声される。この共鳴によって強められた共振周波数をホルマントと呼ぶ。ホルマントは、複数個発生し、周波数の低い方から第 1 ホルマント (F 1)、第 2 ホルマント (F 2)、・・・と呼ぶ。この複数のホルマントによって、声の種類 (音色) が決まる。

10

【 0 0 1 7 】

また、声道は、母音によって極めて複雑な形状を示し、声道の形は、舌や唇を使って変えられている。図 1 (A) に示すように、舌の最も盛り上がっているところを舌の調音位置と呼び、この調音位置が母音 a , i , u , e , o の種類によって特徴的に推移している。この推移しているところを線で結ぶと、五角形になる。母音の種類によって推移する顎の開閉具合と舌の調音位置 (舌によって声道が狭められる位置) と、第 1 ホルマントの周波数および第 2 ホルマントの周波数の間には、図 1 (B) に示すような密接な対応関係がある。図 1 (B) において、横軸は第 1 ホルマント周波数 (F 1) を示し、縦軸は第 2 ホルマント周波数 (F 2) を示している。

20

【 0 0 1 8 】

図 1 (B) には、分かりやすい例として、代表的な男性の声 (図中実線で示す) と代表的な女性の声 (図中点線で示す) のパターンの例を示している。点 M a , 点 M i , 点 M u , 点 M e , 点 M o は、それぞれ、男性が母音 a , i , u , e , o を発したときの第 1 ホルマント周波数と第 2 ホルマント周波数を示し、点 F a , 点 F i , 点 F u , 点 F e , 点 F o は、それぞれ、女性が母音 a , i , u , e , o を発したときの第 1 ホルマント周波数と第 2 ホルマント周波数を示している。

【 0 0 1 9 】

図 1 (B) に示すように、男性の声と女性の声とでは、同じ母音であっても、第 1 ホルマント周波数と第 2 ホルマント周波数の組み合わせが異なる。また、5 つの母音全てについての組み合わせは、男性と女性とで異なる五角形を描くことができる。図 1 (B) では、男性と女性の例により示したが、実際には、この五角形は話者によって異なる。

30

【 0 0 2 0 】

以上のように、第 1 ホルマント周波数と第 2 ホルマント周波数の組み合わせ (合成) によって、母音を模倣することができ、音声を疑似的に生成することが可能となる。

【 0 0 2 1 】

[本発明の第 1 の実施の形態]

図 2 は、本発明の第 1 の実施の形態としての音声生成装置 1 の構成例を示す図である。

【 0 0 2 2 】

音声生成装置 1 は、CPU (Central Processing Unit)、ROM (Read Only Memory)、および RAM (Random Access Memory)、HDD (Hard Disk Drive) などを実装した汎用のコンピュータシステムで構成され、入力デバイス 1 1 と表示部 1 2 を有している。なお、図 2 に示す入力デバイス 1 1 と表示部 1 2 は、音声生成装置 1 と一体に構成されているが、別体で構成するようにしても良い。

40

【 0 0 2 3 】

音声生成装置 1 は、入力デバイス 1 1 からの入力信号に基づいて、ROM などに記憶されている音声生成ソフトウェアを読み出し、読み出した音声生成ソフトウェアを実行する。音声生成装置 1 は、音声生成ソフトウェアの実行により、表示部 1 2 に音声生成 GUI (Graphical User Interface) を表示させ、その GUI への入力信号に基づいて、所定の音声を生

50

成し、スピーカ 1 3 を介して再生（発声）させる。

【 0 0 2 4 】

例えば、表示部 1 2 には、図 2 に示すように、第 1 ホルマント周波数と第 2 ホルマント周波数の 2 次元平面上における分布が音声生成 GUI として表示されており、その GUI 上には、母音 a, i, u, e, o を発したときの第 1 ホルマント周波数と第 2 ホルマント周波数が、「a」、「i」、「u」、「e」、「o」としてそれぞれ示されている。

【 0 0 2 5 】

ユーザは、スピーカ 1 3 から発声させたい発話内容を、マウス 1 1 A（あるいはタッチパッド 1 1 B）を用いて音声生成 GUI 上に軌跡を描く。軌跡の描き方は、例えば、マウス 1 1 A を押下したまま、発話内容に準ずる位置を辿りながら、所望の位置でマウス 1 1 A の押下を解除する。音声生成装置 1 は、マウス 1 1 A の動作に追従するポインタ P が描いた軌跡から、第 1 ホルマント周波数と第 2 ホルマント周波数の 2 次元平面上における X Y 座標位置を検出し、検出した X 座標値で規定されている第 1 ホルマント周波数の音声と、Y 座標値で規定されている第 2 ホルマント周波数の音声とを合成し、合成した疑似的な音声をスピーカ 1 3 から発声させる。

【 0 0 2 6 】

入力デバイス 1 1 は、マウス 1 1 A、タッチパッド 1 1 B、およびキーボード 1 1 C などからなり、ユーザによって入力された入力信号を音声生成装置 1 に供給する。

【 0 0 2 7 】

表示部 1 2 は、例えば、液晶ディスプレイであり、ユーザによって起動された音声生成ソフトウェアに応じた音声生成 GUI を表示する。

【 0 0 2 8 】

図 3 は、音声生成ソフトウェアが実行されることに応じて表示部 1 2 に表示される、音声生成 GUI の表示例を示す図である。

【 0 0 2 9 】

図 3 に示す表示例では、軌跡 L 1（図中実線で示す）と軌跡 L 2（図中点線で示す）が示されている。ユーザによってマウス 1 1 A を用いて軌跡 L 1 が描かれると、音声生成装置 1 は、検出した X Y 座標値から、「おはよう」に聞こえる疑似的な音声を生成し、スピーカ 1 3 から発声させる。また、ユーザによってマウス 1 1 A を用いて軌跡 L 2 が描かれると、音声生成装置 1 は、検出した X Y 座標値から、「あおいうみ」に聞こえる疑似的な音声を生成し、スピーカ 1 3 から発声させる。

【 0 0 3 0 】

図 4 は、音声生成装置 1 の機能構成例を示すブロック図である。図 4 に示す機能部のうちの少なくとも一部は、音声生成装置 1 の CPU により音声生成ソフトウェアが実行されることによって実現される。

【 0 0 3 1 】

音声生成装置 1 は、音源生成部 2 1、音声生成部 2 2、D/A (Digital to Analog) 変換器 2 3、および増幅器 2 4 から構成される。

【 0 0 3 2 】

音源生成部 2 1 は、ユーザの操作によって図示せぬ ON/OFF スイッチからオン信号が供給されると、基本周波数の音声データ（基本音声データ）を生成し、それを音声生成部 2 2 に出力する。

【 0 0 3 3 】

音声生成部 2 2 は、座標値検出部 3 1、第 1 ホルマント共振器 3 2、および第 2 ホルマント共振器 3 3 を有する。

【 0 0 3 4 】

座標値検出部 3 1 は、入力デバイス 1 1 からの入力信号に基づいて、第 1 ホルマント周波数と第 2 ホルマント周波数の 2 次元平面上における X Y 座標値を検出し、検出した X 座標値を第 1 ホルマント共振器 3 2 に出力し、検出した Y 座標値を第 2 ホルマント共振器 3 3 に出力する。

10

20

30

40

50

【 0 0 3 5 】

第1ホルマント共振器32は、音源生成部21から入力された基本音声データを、座標値検出部31からの入力情報（第1ホルマント周波数と第2ホルマント周波数の2次元平面上におけるX座標値）に対応する第1ホルマント周波数で共振させた後、第2ホルマント共振器33に出力する。

【 0 0 3 6 】

第2ホルマント共振器33は、第1ホルマント共振器32から入力された、第1ホルマント周波数で共振された音声データを、座標値検出部31からの入力情報（第1ホルマント周波数と第2ホルマント周波数の2次元平面上におけるY座標値）に対応する第2ホルマント周波数でさらに共振させた後、D/A変換器23に出力する。

10

【 0 0 3 7 】

D/A変換器23は、第2ホルマント共振器32から入力された、第1ホルマント周波数および第2ホルマント周波数で共振された音声データをD/A変換し、増幅器24に出力する。

【 0 0 3 8 】

増幅器24は、D/A変換器23の出力信号（疑似的な音声）を増幅し、スピーカ13に出力する。

【 0 0 3 9 】

次に、図5のフローチャートを参照して、音声生成ソフトウェアが実行する音声生成処理について説明する。

20

【 0 0 4 0 】

この処理を開始するにあたり、音声生成ソフトウェアの起動に伴って、表示部12には、図2に示したような音声生成GUIが表示されている。

【 0 0 4 1 】

ステップS1において、座標値検出部31は、音声生成GUI上で操作が開始されたか否かを判定し、音声生成GUI上で操作が開始されるまで待機する。操作の開始とは、例えば、ユーザによりマウス11Aが押下されることである。また後述する、操作の終了とは、マウス11Aが押下されたままドラッグされた後（軌跡が描かれた後）、押下が解除されることである。

【 0 0 4 2 】

ステップS1において、座標値検出部31は、音声生成GUI上で操作が開始された、すなわち、マウス11Aが押下されたと判定した場合、ステップS2に進み、第1ホルマント周波数と第2ホルマント周波数の2次元平面上におけるXY座標値を検出する。またこのとき、ユーザの操作によって図示せぬON/OFFスイッチからオン信号が供給され、音源生成部21から基本周波数の音声データ（基本音声データ）が出力される。

30

【 0 0 4 3 】

ステップS3において、第1ホルマント共振器32は、音源生成部21から入力された基本音声データを、ステップS2の処理によって検出された、第1ホルマント周波数と第2ホルマント周波数の2次元平面上におけるX座標値に対応する第1ホルマント周波数で共振させる。

40

【 0 0 4 4 】

ステップS4において、第2ホルマント共振器33は、ステップS3の処理によって第1ホルマント周波数で共振された音声データを、ステップS2の処理によって検出された、第1ホルマント周波数と第2ホルマント周波数の2次元平面上におけるY座標値に対応する第2ホルマント周波数で共振させる。

【 0 0 4 5 】

ステップS5において、D/A変換器23は、ステップS4の処理によって第2ホルマント共振器33で共振された音声データをD/A変換する。増幅器24は、D/A変換された出力信号を増幅し、疑似的な音声、すなわち、第1ホルマント周波数と第2ホルマント周波数の変化に応じて模倣された母音をスピーカ13から発声させる。

50

【 0 0 4 6 】

ステップ S 6 において、座標値検出部 3 1 は、音声生成 GUI 上で操作が終了されたか否か、すなわち、マウス 1 1 A の押下が解除されたか否かを判定し、まだ操作が終了していないと判定した場合、ステップ S 2 に戻り、上述した処理を繰り返し実行する。そして、ステップ S 6 において、座標値検出部 3 1 は、音声生成 GUI 上で操作が終了したと判定した場合、音声生成処理を終了する。

【 0 0 4 7 】

[発明の第 1 の実施の形態における効果]

以上のように、第 1 の実施の形態によれば、マウス 1 1 A やタッチパッド 1 1 B などを用いて、直感的な操作で、疑似的な音声をリアルタイムに生成することが可能となる。

10

【 0 0 4 8 】

また、音声生成部 2 2 は、第 1 ホルマント周波数と第 2 ホルマント周波数の値を通じて、発声時の顎の開閉具合や舌の位置による調音位置をシミュレートしているため、生成できる音声は、日本語 5 母音に限らず、外国語の各種母音や、日本語として意味をなさない音声を生成させることも可能となる。

【 0 0 4 9 】

さらに、マウス 1 1 A やタッチパッド 1 1 B による操作軌跡と操作速度を適当に選択することによって、半母音や鼻音に似た音声を生成することも可能である。

【 0 0 5 0 】

[本発明の第 2 の実施の形態]

次に、本発明の第 2 の実施の形態について、図 6 および図 7 を参照して説明する。

20

【 0 0 5 1 】

図 6 は、第 2 の実施の形態としての操作バー 5 1 と携帯型の音声生成装置 5 2 の接続例を示し、図 7 は、操作バー 5 1 と携帯型の音声生成装置 5 2 の内部の構成例を示すブロック図である。

【 0 0 5 2 】

操作バー 5 1 には、図 7 に示すように、回転素子または振動素子などを内蔵したジャイロセンサ 6 1 が搭載されている。ジャイロセンサ 6 1 は、X 軸方向、Y 軸方向、Z 軸方向の加速度をそれぞれ検出し、検出結果を音声生成装置 5 2 に出力する。音声生成装置 5 2 は、予め、空間上における操作バー 5 1 の動作（移動方向と移動量）に応じて、第 1 ホルマント周波数と第 2 ホルマント周波数の 2 次元平面上における分布を対応させた情報を記憶しており、操作バーの動作に応じて疑似的な音声を生成する。

30

【 0 0 5 3 】

音声生成装置 5 2 は、電源部 7 1、音声生成部 7 2、およびスピーカ 7 3 を有している。

【 0 0 5 4 】

電源部 7 1 は、例えば、電池あるいはバッテリーであり、音声生成部 7 2 やスピーカ 7 3 などへ電力を供給する。

【 0 0 5 5 】

音声生成部 7 2 は、第 1 の実施の形態において図 4 に示した機能を有しており、操作バー 5 1 の動作（検出結果）から、第 1 ホルマント周波数と第 2 ホルマント周波数の 2 次元平面上における X Y 座標値を検出し、検出した X 座標値で規定されている第 1 ホルマント周波数の音声と、Y 座標値で規定されている第 2 ホルマント周波数の音声とを合成し、合成した疑似的な音声をスピーカ 7 3 から発声させる。

40

【 0 0 5 6 】

[発明の第 2 の実施の形態における効果]

以上のように、第 2 の実施の形態によれば、操作バー 5 1 を用いて、直感的な操作で、疑似的な音声をリアルタイムに生成することが可能となる。

【 0 0 5 7 】

[変形例]

50

1. 以上においては、入力デバイスとして、マウス 1 1 A、タッチパッド 1 1 B、および操作バー 5 1 を用いる場合を例に説明したが、他にも、タッチペンやジョイスティックなどを利用することも勿論可能である。つまり、ユーザの症例に合わせて入力デバイスを切り替えるようにすることが好ましい。

【 0 0 5 8 】

2. また、操作バー 5 1 にジャイロセンサ 6 1 を搭載するようにしたが、加圧センサをさらに搭載することにより、操作バー 5 1 を握る量に応じて周波数を変化させ、生成する音声に抑揚を持たせることも可能である。

【 0 0 5 9 】

3. さらに、音源生成部 2 1 で生成する基本周波数の音声データの種類を変更することにより、男性の声、女性の声など様々な音声を生成することが可能となる。

10

【 0 0 6 0 】

4. なお、上述において、第 1 ホルマント周波数と第 2 ホルマント周波数を組み合わせることにより、母音を疑似的に生成することができるが、本実施の形態はこれに限定されず、第 3 ホルマント周波数と第 4 ホルマント周波数をさらに組み合わせることにより、声の性質や特徴をも加味することが可能である。

【 0 0 6 1 】

5. また、以上においては、母音の組み合わせで疑似的な音声を生成するようにしたが、さらに子音を組み合わせることにより、より自然な音声を生成することが可能となる。

【 0 0 6 2 】

20

図 8 は、より高品質な音声を生成する音声生成装置 1 の機能構成例を示すブロック図である。なお、図 4 に示した構成要素と同一の構成要素には同一の符号を付してあり、重複する説明は適宜省略する。

【 0 0 6 3 】

図 8 に示す音声生成装置 1 には、音源生成部 2 1、音声生成部 2 2、D/A変換器 2 3、増幅器 2 4 の他、加算器 2 5、乱流音生成部 2 6、加算器 2 7、鼻音生成部 2 8、およびハイパスフィルタ (HPF) 3 4 が新たに設けられている。

【 0 0 6 4 】

音声生成部 2 2 には、座標値検出部 3 1、第 1 ホルマント共振器 3 2、第 2 ホルマント共振器 3 3 の他、第 3 ホルマント共振器 3 5 および第 4 ホルマント共振器 3 6 が新たに設けられている。また、鼻音生成部 2 8 には、鼻音生成用共振器 3 7 が設けられている。

30

【 0 0 6 5 】

座標値検出部 3 1 は、入力デバイス 1 1 からの入力信号に基づいて、第 1 ホルマント周波数と第 2 ホルマント周波数の 2 次元平面上における X Y 座標値を検出し、検出した X 座標値を第 1 ホルマント共振器 3 2 および乱流音生成部 2 6 の第 1 ホルマント共振器 8 6 に出力し、検出した Y 座標値を第 2 ホルマント共振器 3 3 および乱流音生成部 2 6 の第 2 ホルマント共振器 8 7 に出力する。

【 0 0 6 6 】

また座標値検出部 3 1 は、入力デバイス 1 1 からの入力信号に基づいて、第 3 ホルマント周波数と第 4 ホルマント周波数の 2 次元平面上における X Y 座標値を検出し、検出した X 座標値を第 3 ホルマント共振器 3 5 および乱流音生成部 2 6 の第 3 ホルマント共振器 8 8 に出力し、検出した Y 座標値を第 4 ホルマント共振器 3 6 および乱流音生成部 2 6 の第 4 ホルマント共振器 8 9 に出力する。さらに座標値検出部 3 1 は、入力デバイス 1 1 からの入力信号に基づいて、鼻音の有無を判断し、その判断結果を鼻音生成部 2 8 の鼻音生成用共振器 3 7 を通知する。たとえば、図 3 の例において、「u」のやや左側から「i」付近の位置に向かう軌跡を描くようにマウス 1 1 A が操作されたとき、「み」に近い音 (鼻音を含む音) が発生されるので、入力信号からそのような軌跡が検出された場合は、鼻音の有と判断され、その判断結果が、鼻音生成部 2 8 の鼻音生成用共振器 3 7 に通知される。

40

【 0 0 6 7 】

50

ハイパスフィルタ 34 は、音源生成部 21 からの基本周波数の音声データのうち、高周波を通過させ、遮断周波数より低い周波数の帯域を減衰させた後、第 1 ホルマント共振器 32 および鼻音生成用共振器 37 に出力する。

【0068】

第 1 ホルマント共振器 32、第 2 ホルマント共振器 33、第 3 ホルマント共振器 35、および第 4 ホルマント共振器 36 は、音源生成部 21 で生成されハイパスフィルタ 34 で低周波成分が除去された音声データを、座標値検出部 31 からの入力情報に対応してそれぞれの共振周波数で共振させる。鼻音生成用共振器 37 は、鼻音の有の旨が、座標値検出部 31 から通知されると、音源生成部 21 で生成された音声データを所定の共振周波数で共振させて鼻音となる音声データを生成する。

10

【0069】

加算器 25 は、第 1 ホルマント共振器 32、第 2 ホルマント共振器 33、第 3 ホルマント共振器 35、および第 4 ホルマント共振器 36 のそれぞれのホルマント周波数で共振された音声データと、鼻音生成用共振器 37 の共振周波数で共振された音声データを加算する。

【0070】

乱流音生成部 26 は、疑似乱数発生器 81 乃至 85、第 1 ホルマント共振器 86、第 2 ホルマント共振器 87、第 3 ホルマント共振器 88、および第 4 ホルマント共振器 89 を有する。

【0071】

疑似乱数発生器 81 乃至 84 は、摩擦音などの子音を合成するための音源を疑似乱数によって生成し、第 1 ホルマント共振器 86、第 2 ホルマント共振器 87、第 3 ホルマント共振器 88、第 4 ホルマント共振器 89 にそれぞれ供給する。疑似乱数発生器 85 は、声道での共鳴を伴わない子音を合成するための音源を疑似乱数によって生成し、第 4 ホルマント共振器 89 の出力後段に供給している。

20

【0072】

第 1 ホルマント共振器 86 は、疑似乱数発生器 81 で生成された子音用の音源データを、座標検出部 31 からの入力情報に対応する第 1 ホルマント周波数で共振させる。

【0073】

第 2 ホルマント共振器 87 は、第 1 ホルマント共振器 86 から出力された音声データ、及び、疑似乱数発生器 82 で生成された子音用の音源データを、座標検出部 31 からの入力情報に対応する第 2 ホルマント周波数で共振させる。

30

【0074】

第 3 ホルマント共振器 88 は、第 2 ホルマント共振器 87 から出力された音声データ、及び、疑似乱数発生器 83 で生成された子音用の音源データを、座標検出部 31 からの入力情報に対応する第 3 ホルマント周波数で共振させる。

【0075】

第 4 ホルマント共振器 89 は、第 3 ホルマント共振器 88 から出力された音声データ、及び、疑似乱数発生器 84 で生成された子音用の音源データを、座標検出部 31 からの入力情報に対応する第 4 ホルマント周波数で共振させる。

40

【0076】

加算器 27 は、加算器 25 から出力された音声データと乱流音生成部 26 から出力された音声データを加算する。D/A変換器 23 は、加算器 27 から入力された音声データをD/A変換し、増幅器 24 を介してスピーカ 13 に出力する。

【0077】

以上のような構成によって、第 1 ホルマント周波数および第 2 ホルマント周波数だけでなく、第 3 ホルマント周波数および第 4 ホルマント周波数を組み合わせた音声を生じることが可能となる。また、鼻音と乱流音の音声データを組み合わせることにより、鼻音化された母音や、摩擦音などの子音を含んだより自然な音声を生じることが可能となる。

【0078】

50

6. 上述した一連の処理は、ハードウェアにより実行することもできるし、ソフトウェアにより実行することもできる。一連の処理をソフトウェアにより実行する場合には、そのソフトウェアを構成するプログラムが、専用のハードウェアに組み込まれているコンピュータ、または、各種のプログラムをインストールすることで、各種の機能を実行することが可能な、例えば汎用のパーソナルコンピュータなどに、プログラム記録媒体からインストールされる。

【0079】

図9は、上述した一連の処理をプログラムにより実行するコンピュータのハードウェアの構成例を示すブロック図である。

【0080】

コンピュータにおいて、CPU101、ROM102、RAM103、および入出力インターフェース104は、バス105により相互に接続されている。

【0081】

入出力インターフェース104には、さらに、キーボード、マウス、タッチパッド、操作バー、マイクロホンなどよりなる入力部106、ディスプレイ、スピーカなどよりなる出力部107、ハードディスクや不揮発性のメモリなどよりなる記憶部108、ネットワークインタフェースなどよりなる通信部109、磁気ディスク、光ディスク、光磁気ディスク、或いは半導体メモリなどのリムーバブルメディア111を駆動するドライブ110が接続されている。

【0082】

以上のように構成されるコンピュータでは、CPU101が、例えば、記憶部108に記憶されているプログラムを、入出力インターフェース104およびバス105を介して、RAM103にロードして実行することにより、上述した一連の処理が行われる。

【0083】

コンピュータ(CPU101)が実行するプログラムは、例えば、磁気ディスク(フレキシブルディスクを含む)、光ディスク(CD-ROM(Compact Disc-Read Only Memory)、DVD(Digital Versatile Disc)等)、光磁気ディスク、もしくは半導体メモリなどよりなるパッケージメディアであるリムーバブルメディア111に記録して、あるいは、ローカルエリアネットワーク、インターネット、デジタル衛星放送といった、有線または無線の伝送媒体を介して提供される。

【0084】

そして、プログラムは、リムーバブルメディア111をドライブ110に装着することにより、入出力インターフェース104を介して、記憶部108にインストールすることができる。また、プログラムは、有線または無線の伝送媒体を介して、通信部109で受信し、記憶部108にインストールすることができる。その他、プログラムは、ROM102や記憶部108に、あらかじめインストールしておくことができる。

【0085】

なお、コンピュータが実行するプログラムは、本明細書で説明する順序に沿って時系列に処理が行われるプログラムであっても良いし、並列に、あるいは呼び出しが行われたとき等の必要なタイミングで処理が行われるプログラムであっても良い。さらに、プログラムは、遠方のコンピュータに転送されて実行されるものであっても良い。また、プログラムを実行するハードウェアとして、汎用コンピュータの他に、携帯電話、ゲーム端末、電子音楽プレーヤ、電子書籍リーダーなどを利用して良い。

【0086】

7. この発明は、上記実施の形態そのままに限定されるものではなく、実施段階ではその要旨を逸脱しない範囲で構成要素を変形して具体化したり、上記実施の形態に開示されている複数の構成要素を適宜組み合わせたりすることにより種々の発明を形成できる。例えば、実施の形態に示される全構成要素から幾つかの構成要素を削除してもよい。さらに、異なる実施の形態に亘る構成要素を適宜組み合わせてもよい。

【符号の説明】

10

20

30

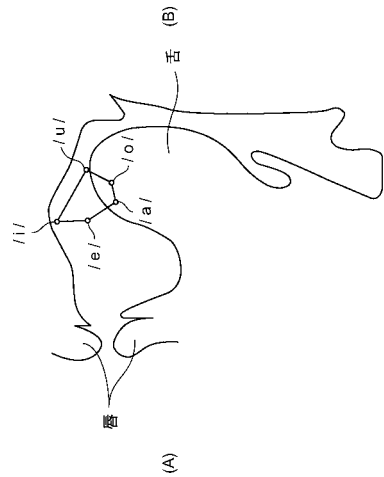
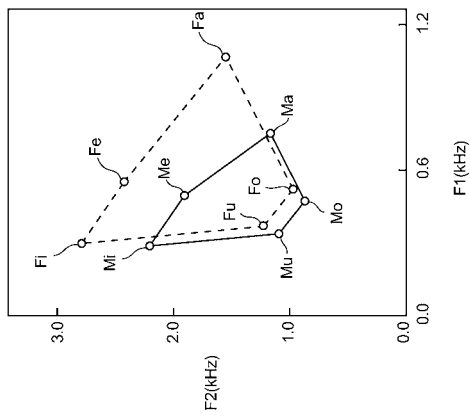
40

50

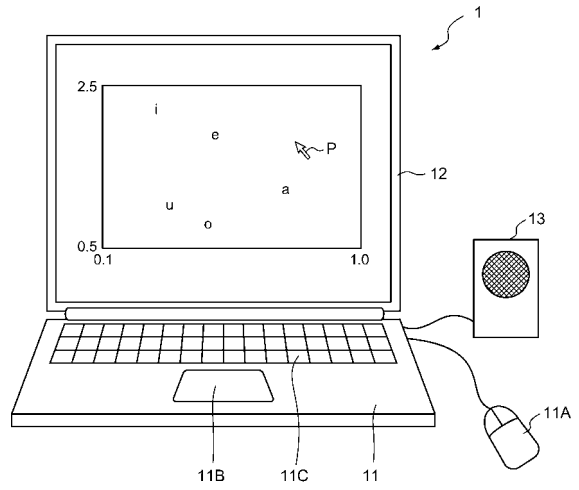
【 0 0 8 7 】

- 1 音声生成装置
- 1 1 入力デバイス
- 1 2 表示部
- 1 3 スピーカ
- 2 1 音源生成部
- 2 2 音声生成部

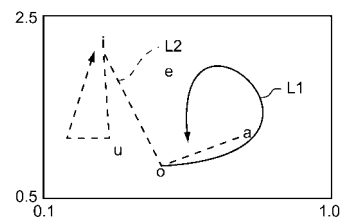
【 図 1 】



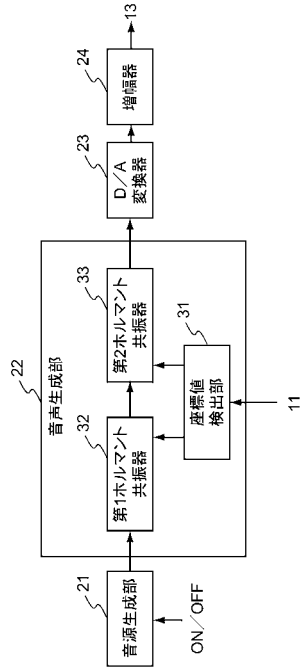
【 図 2 】



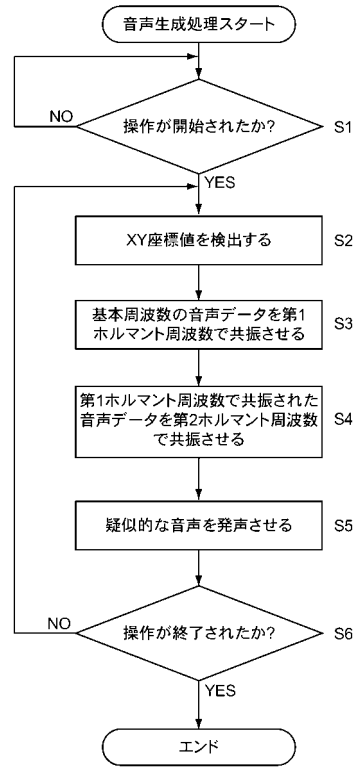
【 図 3 】



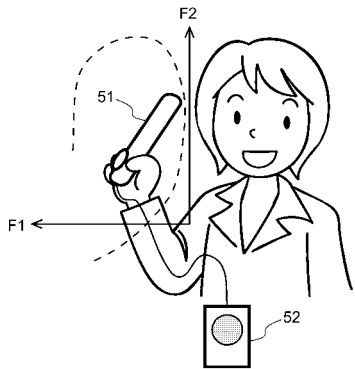
【 図 4 】



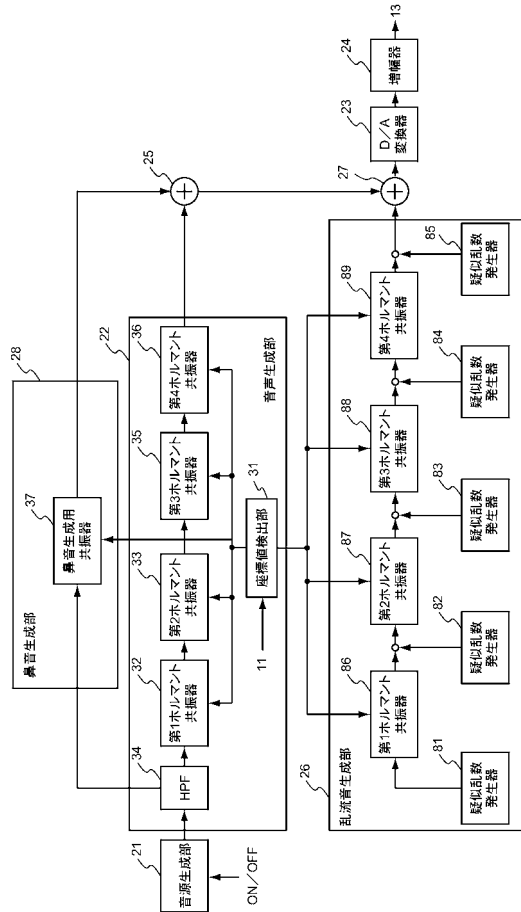
【 図 5 】



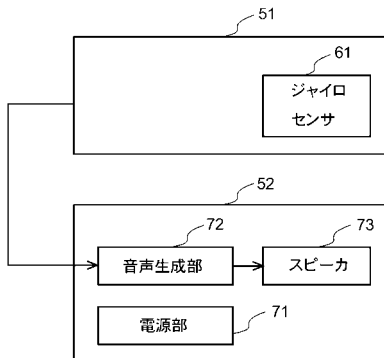
【 図 6 】



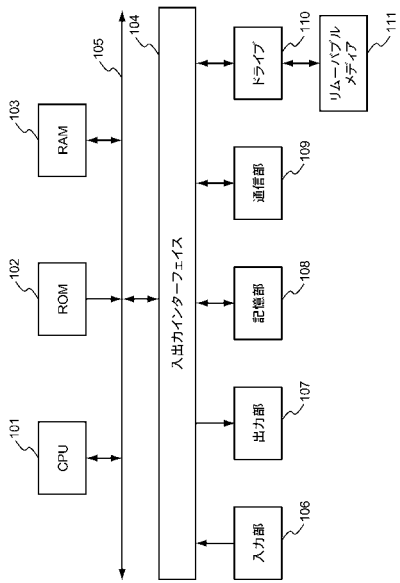
【 図 8 】



【 図 7 】



【図9】



フロントページの続き

(72)発明者 橋場 参生

北海道札幌市北区北19条西1丁目1番地9 地方独立行政法人北海道立総合研究機構内

(72)発明者 須貝 保徳

北海道江別市工栄町8番地の13 株式会社電制 江別工場内

審査官 山下 剛史

(56)参考文献 特開2009-258366(JP,A)

特表2000-503412(JP,A)

藪謙一郎他, "発話障害者支援のための音声合成器の基礎的設計", 電子情報通信学会技術研究報告, 2006年 3月, Vol.105, No.686, pp.59-64

(58)調査した分野(Int.Cl., DB名)

G10L 13/00 - 13/10