

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第4543919号  
(P4543919)

(45) 発行日 平成22年9月15日(2010.9.15)

(24) 登録日 平成22年7月9日(2010.7.9)

(51) Int. Cl. F I  
**G09B 19/04 (2006.01)** G09B 19/04  
**G09B 5/06 (2006.01)** G09B 5/06  
**G09B 19/06 (2006.01)** G09B 19/06

請求項の数 4 (全 12 頁)

(21) 出願番号 特願2004-371875 (P2004-371875)  
 (22) 出願日 平成16年12月22日(2004.12.22)  
 (65) 公開番号 特開2006-178214 (P2006-178214A)  
 (43) 公開日 平成18年7月6日(2006.7.6)  
 審査請求日 平成19年10月23日(2007.10.23)

(73) 特許権者 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 100098084  
 弁理士 川▲崎▼ 研二  
 (72) 発明者 畑 紀行  
 静岡県浜松市中沢町10番1号 ヤマハ株  
 式会社内  
 審査官 古川 直樹

最終頁に続く

(54) 【発明の名称】 語学学習装置

(57) 【特許請求の範囲】

【請求項1】

模範音声を記憶する記憶手段と、  
 ユーザ音声を入力する入力手段と、  
 前記模範音声および前記ユーザ音声をそれぞれ、子音部分と母音部分とに分離する分離手段と、  
 前記模範音声および前記ユーザ音声の長さが同一となるように、前記模範音声あるいは前記ユーザ音声の母音部分の長さを変化させる処理手段と、  
 前記処理手段により長さが揃えられた模範音声とユーザ音声とを比較し、差異点を抽出する比較手段と、  
 前記比較手段により抽出された差異点を強調する差異点強調手段と、  
 前記差異点強調手段により差異点が強調された音声を出力する出力手段とを有し、  
 前記差異点強調手段が、前記比較手段により差異点が抽出された部分については模範音声の音量をユーザ音声の音量より大きくし、前記比較手段により差異点が抽出されなかった部分については模範音声の音量をユーザ音声の音量より小さくすることを特徴とする語学学習装置。

【請求項2】

前記模範音声および前記ユーザ音声のいずれか一方の子音部分を、他方の子音部分と時間軸上の位置が同じとなるように再配置する再配置手段をさらに有し、

前記処理手段が、前記再配置手段により再配置された子音部分と、時間軸上でその子音部分の次に現れる子音部分との間に位置する母音部分の長さをそれぞれ変化させることにより、前記模範音声および前記ユーザ音声の長さを同一とする

ことを特徴とする請求項 1 に記載の語学学習装置。

【請求項 3】

前記ユーザ音声および前記模範音声に対し所定のパラメータを抽出し、該パラメータの変化量に応じた図形を示す画像データを生成する画像生成手段と、

前記画像生成手段により生成された画像データを、前記比較手段により抽出された差異点で異なる表示態様で表示を行う表示手段と

をさらに有する請求項 1 に記載の語学学習装置。

10

【請求項 4】

前記差異点強調手段が、前記比較手段により差異点が抽出された部分については模範音声のうち特定の周波数領域を増幅することを特徴とする請求項 1 に記載の語学学習装置。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、語学学習を支援する語学学習システムに関する。

【背景技術】

【0002】

外国語あるいは母国語の語学学習、特に、発音あるいは発話の独習においては、CD (Compact Disk) 等の記録媒体に記録された模範音声を再生し、その模範音声の真似をして発音あるいは発話するという学習方法が広く用いられている。これは模範音声の真似をすることで正しい発音を身につけることを目的とするものである。ここで、学習をより効果的に進めるためには、模範音声と自分の音声との差を客観的に評価する必要がある。しかし、CD に記録された模範音声を聞いてその真似をするだけでは、自分の発した音声と模範音声との差を具体的に把握することが困難であるという問題があった。

20

【0003】

このような問題を解決する技術として、例えば特許文献 1 ~ 3 に記載の技術がある。特許文献 1 には、模範音声とユーザ音声とを同時に再生する技術が開示されている。特許文献 2 には、模範音声の波形（模範波形）とユーザ音声の波形（自声波形）とを同時に出力する技術が開示されている。特許文献 3 には、模範音声とユーザ音声との比較を行う際に両者の頭を揃えたり、両者の長さを同一にするために一方の音声を一律に引き伸ばす技術が開示されている。

30

【特許文献 1】特開平 7 - 219418 号公報

【特許文献 2】特開 2002 - 23613 号公報

【特許文献 3】特開平 11 - 143496 号公報

【発明の開示】

【発明が解決しようとする課題】

【0004】

特許文献 1 に記載の技術においては、模範音声とユーザ音声とが同時に再生されるのみで、両者の差異点が分かりにくいという問題があった。また、特許文献 2 に記載の技術においても、模範音声の波形とユーザ音声の波形とが同時に出力されるのみで、両者の差異点が分かりにくいという問題があった。さらに、特許文献 3 に記載の技術においては、模範音声とユーザ音声の長さを揃えるために、一方の音声の長さを一律に引き伸ばすのみであり、時間軸に対して両者の音韻が必ずしも一致しないという問題があった。

40

【0005】

本発明は上述の事情に鑑みてなされたものであり、模範音声とユーザ音声との差異点を提示することができる語学学習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】

【0006】

50

上述の課題を解決するため、本発明は、模範音声を記憶する記憶手段と、ユーザ音声を入力する入力手段と、前記模範音声および前記ユーザ音声をそれぞれ、子音部分と母音部分とに分離する分離手段と、前記模範音声および前記ユーザ音声の長さが同一となるように、前記模範音声あるいは前記ユーザ音声の母音部分の長さを変化させる処理手段と、前記処理手段により長さが揃えられた模範音声とユーザ音声とを比較し、差異点を抽出する比較手段とを有する語学学習装置を提供する。

【0007】

好ましい態様において、この語学学習装置は、前記模範音声および前記ユーザ音声のいずれか一方の子音部分を、他方の子音部分と時間軸上の位置が同じとなるように再配置する再配置手段をさらに有し、前記処理手段が、前記再配置手段により再配置された子音部分と、時間軸上でその子音部分の次に現れる子音部分との間に位置する母音部分の長さをそれぞれ変化させることにより、前記模範音声および前記ユーザ音声の長さを同一としてもよい。

10

別の好ましい態様において、この語学学習装置は、前記ユーザ音声および前記模範音声に対し所定のパラメータを抽出し、該パラメータの変化量に応じた図形を示す画像データを生成する画像生成手段と、前記画像生成手段により生成された画像データを、前記比較手段により抽出された差異点で異なる表示態様で表示を行う表示手段とをさらに有してもよい。

【0008】

さらに別の好ましい態様において、この語学学習装置は、前記比較手段により抽出された差異点を強調する差異点強調手段と、前記差異点強調手段により差異点が強調された音声を入力する出力手段とをさらに有してもよい。

20

この態様において、前記差異点強調手段が、前記比較手段により差異点が抽出された部分については模範音声の音量をユーザ音声の音量より大きくし、前記比較手段により差異点が抽出されなかった部分については模範音声の音量をユーザ音声の音量より小さくすることとしてもよい。

あるいは、前記差異点強調手段が、前記比較手段により差異点が抽出された部分については模範音声のうち特定の周波数領域を増幅することとしてもよい。

【発明の効果】

【0009】

本発明によれば、ユーザは自分の音声と模範音声との差異がある部分を具体的に特定することができる。

30

【発明を実施するための最良の形態】

【0010】

以下、図面を参照して本発明の実施形態について説明する。

<第1実施形態>

図1は、本発明の第1実施形態に係る語学学習装置1の機能構成を示すブロック図である。記憶部10は、語学学習においてお手本となる音声を示す模範音声データを記憶する。入力部20は、ユーザ（学習者あるいは生徒）の音声を取得し、ユーザ音声データを入力する。データ処理部30は、ユーザ音声と模範音声とが同一の長さとなるようにユーザ音声データの処理を行う。差異点抽出部40は同一の長さに揃えられたユーザ音声データと模範音声データとを比較して、両者の差異点を抽出する。差異点強調部50は、模範音声データおよびユーザ音声データに対して差異点抽出部40で抽出された差異点を強調する処理を行う。音声出力部60は、差異点が強調された模範音声およびユーザ音声を再生する。各構成要素の機能の詳細については後述する。

40

【0011】

図2は、語学学習装置1のハードウェア構成を示すブロック図である。CPU（Central Processing Unit）101は、RAM（Random Access Memory）102を作業エリアとして、ROM（Read Only Memory）103あるいはHDD（Hard Disk Drive）104に記憶されているプログラムを読み出して実行する。HDD104は、各種アプリケーション

50

ンプログラムやデータを記憶する記憶装置である。本実施形態に関して、特に、HDD 104は、語学学習プログラム、この語学学習プログラムで使用する模範音声データを記録した模範音声データベースDB1を記憶している。

#### 【0012】

ディスプレイ105は、CRT (Cathode Ray Tube) やLCD (Liquid Crystal Display) 等、CPU 101の制御下で文字や画像を表示する表示装置である。マイク106は、ユーザの音声を取得するための集音装置であり、ユーザの発した音声に対応する音声信号を出力する。音声処理部107は、マイク106により出力されたアナログ音声信号をデジタル音声データに変換する機能や、HDD 104に記憶された音声データを変換してスピーカ108に出力する機能を有する。また、ユーザはキーボード109を操作することにより、語学学習装置1に対して指示入力を行うことができる。以上で説明した各構成要素は、バス110を介して相互に接続されている。

10

#### 【0013】

図3は、模範音声データベースDB1の内容を示す図である。模範音声データベースDB1には、語学学習に用いる例文(例えば、英語の学習において「Good to see you again. How are you?」等の例文)のテキストデータと、例文単位の音声波形をデジタル化した音声波形データと、その例文を特定する識別子とが複数記憶されている。例文テキストデータ、音声波形データ、識別子はそれぞれ対応付けられている。

#### 【0014】

続いて、語学学習装置1の動作について説明する。本実施形態においては、CPU 101がHDD 104に記憶された語学学習プログラムを実行することにより、語学学習装置1において図1に示される各機能構成要素に相当する機能が実現される。

20

#### 【0015】

図4は、本実施形態に係る語学学習装置1の動作を示すフローチャートである。語学学習プログラムを実行すると、CPU 101は、ディスプレイ105上に例文の選択を促すメッセージを表示する。ユーザはディスプレイ105上に表示されたメッセージに従い、模範音声データベースDB1に記録された例文から1の例文を選択する。CPU 101は選択された例文の音声を再生する(ステップS101)。具体的には次のとおりである。CPU 101は、模範音声データベースDB1から、選択された例文に対応する模範音声データを抽出する。CPU 101は、抽出した模範音声データを音声処理部107に出力する。音声処理部107は入力された模範音声データをデジタル/アナログ変換してアナログ音声信号としてスピーカ108に出力する。こうして模範音声再生される。

30

#### 【0016】

ユーザはスピーカ108から再生された模範音声を聞き、マイク106に向かって模範音声を真似して例文を発声する。すなわち、ユーザ音声の入力が行われる(ステップS102)。具体的には次のとおりである。模範音声の再生が終了すると、CPU 101は、「次はあなたの番です。例文を発音してください」等、ユーザに例文の発生を促すメッセージをディスプレイ105に表示する。さらにCPU 101は、「スペースキーを押してから発音し、発音が終わったらもう一度スペースキーを押してください」等、ユーザ音声の入力を行うための操作を指示するメッセージをディスプレイ105に表示する。ユーザは、ディスプレイ105に表示されたメッセージに従ってキーボード109を操作し、ユーザ音声の入力を行う。すなわち、キーボード109のスペースキーを押した後に、マイク106に向かって例文を発声する。発声が終了したら、ユーザはもう一度スペースキーを押す。

40

#### 【0017】

ユーザの音声はマイク106により電気信号に変換される。マイク106は、ユーザ音声信号を出力する。ユーザ音声信号は、音声処理部107によりデジタル音声データに変換され、ユーザ音声データとしてHDD 104に記録される。CPU 101は、模範音声の再生が完了した後、スペースキーの押下をトリガとしてユーザ音声データの記録を開始し、再度のスペースキーの押下をトリガとしてユーザ音声データの記録を終了する。すな

50

わち、ユーザが最初にスペースキーを押してから、もう一度スペースキーを押すまでの間のユーザ音声はHDD104に記録される。

【0018】

続いてCPU101は、ユーザ音声と模範音声の長さが同一となるようにユーザ音声データを処理する(ステップS103)。具体的には次のとおりである。図5は、模範音声(図5(A))およびユーザ音声(図5(B))の波形を例示する図である。図5に示される例では、模範音声およびユーザ音声はともに同一の例文を発声したものであるが、発話の速度が異なっているため、長さが異なっている。すなわち、ユーザ音声の方が発話速度が遅いため、音声の長さが長くなっている。CPU101は、以下のようにしてユーザ音声の長さを模範音声と同一にする。

10

【0019】

図6は、ステップS103における、ユーザ音声と模範音声の長さを同一にする処理をより詳細に示すフローチャートである。CPU101は、まず、データサイズを計測する等の方法により、模範音声およびユーザ音声の長さを算出する(ステップS103-1)。CPU101は、この算出結果から、模範音声とユーザ音声の長さの差(図5のt)をさらに算出する。

【0020】

続いてCPU101は、ユーザ音声データのうち、子音に係る部分および母音に係る部分のそれぞれに識別子を付加する(ステップS103-2)。これは次のような目的による。すなわち、一般に子音の長さは話者によらずほぼ同一であるのに対し、母音の長さは話者によって大きく異なる。したがって、ユーザ音声の長さを変更する際に、子音の長さはそのまま母音の長さのみ変化させれば、聴感上の不自然さを生じさせずに音声の長さを変更することができる。従来技術においては、子音であるか母音であるかにかかわらず様に音声の長さを変更されるので、不自然な聴感を与えてしまうという問題があった。しかし、本実施形態によればこのような問題は起こらない。

20

【0021】

母音部分と子音部分の分離は例えば次のように行う。CPU101は、選択された例文のテキストデータから、その例文に含まれる母音を抽出する。例えばテキストデータの先頭から順に「a」「u」「i」「e」という母音が抽出された場合を考える。CPU101は、音声データをあらかじめ決められた時間(フレーム)毎に分割する。CPU101は、フレームに分解された模範音声データが示す波形およびユーザ音声信号が示す波形をフーリエ変換して得られた振幅スペクトルの対数を求め、それをフーリエ逆変換してフレームごとのスペクトル包絡を得る。CPU101は、こうして得られたスペクトル包絡から第1フォルマントおよび第2、第3フォルマントのフォルマント周波数を抽出する。一般に母音は第1および第2、第3フォルマントの分布により特徴付けられる。CPU101は、音声データの先頭からまず母音「a」のフォルマント周波数分布とマッチングを行う。マッチングによりそのフレームが母音「a」に相当するものであると判断された場合、CPU101は、検出された母音の種類および音声データ上の位置を示すデータCを生成する。既にデータCが生成されているときは、データCに新たな情報を追加する。CPU101は、後続するフレームについても母音「a」とのマッチングを行い、マッチしなくなったら続いて母音「u」とのマッチングを行う。このようにして先頭から母音を検索し、データCを生成する。なお、母音の種類および位置を示すデータを生成する代わりに、子音の位置を示すデータを生成してもよいし、母音と子音両方の位置を示すデータを生成してもよい。また、音と音のリエゾン区間(中間的な音)の区間位置データや、無音区間データを生成してもよい。

30

40

【0022】

続いて、CPU101は、模範音声データのうち、子音に係る部分について、データの先頭から順に番号を付し、子音に係る部分それぞれの先頭位置を示す情報(例えば、データの先頭からの時間)と共にテーブルTB1としてRAM102に記憶する。ここで、模範音声データにおける母音と子音の分離は、上述のユーザ音声データの場合と同様に行っ

50

てもよいし、あらかじめ模範音声データベースDB1に母音の位置または子音の位置を示す情報を記憶しておき、その情報に基づいて子音部分を特定してもよい。続いてCPU101は、ユーザ音声データから子音に係る部分を切り出し、テーブルTB1を参照して、模範音声と子音の位置が一致するように、切り出した子音部分を再配置する(ステップS103-3)。さらにCPU101は、ユーザ音声データのうち、子音と子音の間に位置する母音に係る部分について、模範音声と同じ長さとなるようにデータの加工を行う(ステップS103-3)。これは例えば、ユーザ音声の母音部分の方が模範音声の母音部分よりも長い場合には、その長い部分のデータを削除することにより実現できる。あるいは、ユーザ音声の母音部分の方が模範音声の母音部分よりも短い場合には、所望の長さになるまで母音部分の波形を繰り返し足しつけていけばよい。このようにして、ユーザ音声の長さは模範音声と同一となり、また、時間軸上の子音および母音の位置も一致することとなる。

10

#### 【0023】

再び図4を参照して説明する。CPU101は、模範音声とユーザ音声との差異点を抽出する(ステップS104)。この処理は例えば次のように行われる。CPU101は、前述のように模範音声データが示す波形をあらかじめ決められた時間(フレーム)ごとに分割する。また、CPU101は、ユーザ音声データが示す波形についてもフレームごとに分割する。CPU101は、フレームに分解された模範音声データが示す波形およびユーザ音声信号が示す波形をフーリエ変換して得られた振幅スペクトルの対数を求め、それをフーリエ逆変換してフレームごとのスペクトル包絡を得る。

20

#### 【0024】

図7は、模範音声(上)およびユーザ音声(下)のスペクトル包絡を例示する図である。図7に示されるスペクトル包絡は、フレームI~フレームIIIの3つのフレームから構成されている。CPU101は、得られたスペクトル包絡をフレームごとに比較する。CPU101は、模範音声のスペクトル包絡とユーザ音声のスペクトル包絡との差異が、あらかじめ決められたしきい値を超えた場合は、そのフレームにおいて模範音声とユーザ音声との差異があるものと判断する。模範音声とユーザ音声との差異は、例えば、特徴的なフォルマントの周波数とスペクトル密度とをスペクトル密度-周波数図に表したときの2点間の距離によって求めてもよいし、特定の周波数においてスペクトル密度を比較することによって求めてもよい。あるいは、1以上の特定のフォルマントのフォルマント周波数を比較することにより模範音声とユーザ音声との差異を求めてもよい。図5に示される例では、CPU101はフレームIIについて差異があるものと判断する。CPU101は、模範音声とユーザ音声に差異があったことを示すフラグを記録したデータFを生成し、RAM102に記憶する。模範音声とユーザ音声とに差異が無い場合は、CPU101は、そのことを示すデータFを生成し、RAM102に記憶する。さらに、CPU101は、そのフレームにおける模範音声とユーザ音声との差異の有無を示すフラグを記録したデータDを生成し、RAM102に記憶する。すなわち、データDは、フレームごとにユーザの発音の良否(模範音声との差異の有無)を示している。CPU101はこのようにしてすべてのフレームについて模範音声のスペクトルとユーザ音声のスペクトルを比較する。RAM102には、模範音声と差異があると判断されたフレームを特定するデータDが記憶されている。

30

40

#### 【0025】

再び図4を参照して説明する。CPU101は、データFに基づいてユーザ音声に模範音声と異なっている部分が存在するか否か判断する(ステップS105)。ユーザ音声に模範音声と異なっている部分がある場合(ステップS105: YES)、CPU101は、以下で説明する差異点強調処理を行う(ステップS106)。CPU101は、ステップS105、S106の処理を全フレームに渡って行う(ステップS108)。これにより、差異点が強調された模範音声が再生される。ユーザの発音に悪い部分が無い場合(ステップS105: NO)、CPU101は「良好です」等のメッセージをディスプレイ105に表示し、処理を終了する。

50

## 【0026】

ステップS106における差異点強調処理は、例えば以下のように行われる。ユーザが自分の発音を確認するという目的から、後述するステップS107において、基本的にはユーザ音声再生される。しかし、模範音声との差異があった部分については、強調処理として、ユーザ音声ではなく模範音声を再生する。これにより、ユーザが模範音声との差異があった部分を具体的に特定することができるという効果、および差異があった部分について正しい発音をユーザに示すことができるという効果が奏される。差異点強調処理は具体的には、次のように行われる。CPU101は、ステップS103において長さを揃えられた模範音声およびユーザ音声に対し、それぞれ音量係数を乗じて加算する。音量係数は、再生される音声の音量を示すパラメータである。例えばユーザ音声の音量係数が1で模範音声の音量係数が0である場合は、スピーカ108からはユーザ音声のみが再生される。逆にユーザ音声の音量係数が0で模範音声の音量係数が1である場合にはスピーカ108からは模範音声のみが再生される。

10

## 【0027】

CPU101は、データDを参照してフレーム毎に音量係数を決定する。すなわち、ユーザ音声データにおいて、データDが差異点ありを示している場合には、CPU101はそのフレームの音量係数を0に設定する。逆に、データDが模範音声との差異点なしを示している場合には、CPU101はそのフレームの音量係数を1に設定する。一方、模範音声データにおいて、データDが差異点ありを示している場合には、CPU101はそのフレームの音量係数を1に設定する。逆に、データDが模範音声との差異点なしを示している場合には、CPU101はそのフレームの音量係数を0に設定する。CPU101は、このようにして求められた音量係数をユーザ音声データおよび模範音声データに乗じて、ユーザ音声データと模範音声データとを混合する。CPU101は、こうして得られた混合音声データを音声処理部107に出力する。

20

## 【0028】

続いて音声処理部107は、入力された混合音声データをデジタル/アナログ変換し、音声信号としてスピーカ108に出力する。スピーカ108からは、強調処理を施された音声再生される(ステップS107)。この音声を聞くことにより、ユーザは自分の音声と模範音声との差異がある部分を具体的に特定することができ、また、その差異がある部分については正しい発音を知ることができる。

30

## 【0029】

<第2実施形態>

続いて、本発明の第2実施形態について説明する。なお、以下の説明において第1実施形態と共通の要素には共通の参照符号を付与し、その説明を省略する。

図8は、本発明の第2実施形態に係る語学学習装置2の機能構成を示すブロック図である。第1実施形態に係る語学学習装置1と異なる部分についてのみ説明すると、差異点強調部51は、模範音声データおよびユーザ音声データに対して差異点抽出部40で抽出された差異点に基づいて、模範音声およびユーザ音声を視覚化した画像であって、両者の差異点強調された画像を示す画像データを生成する。差異点表示部61は、差異点強調部51により生成された画像データに基づいて画像表示を行う。なお、語学学習装置2のハードウェア構成は図2に示される語学学習装置1のハードウェア構成と同一であるのでその説明を省略する。

40

## 【0030】

図9は、本発明の第2実施形態に係る語学学習装置2の動作を示すフローチャートである。ステップS101~S105の処理は第1実施形態と同じであるのでその説明を省略する。ユーザ音声に模範音声と異なっている部分がある場合(ステップS105: YES)、CPU101は、差異点強調処理を行う(ステップS206)。本実施形態において、模範音声とユーザ音声との差異点は画像で視覚的に表される。

## 【0031】

図10は、本実施形態において出力される画像を例示する図である。模範音声画像Aお

50

よびユーザ音声画像Bが、縦に並べられて表示される。各音声画像は、音量および音程を示す図(A-1、A-2)と、周波数特性を示す図(B-1、B-2)と、主として2つの図から構成される。これらの図はいずれも、水平方向が時間軸となっている。ステップS206において、CPU101は、模範音声データおよびユーザ音声データに基づいて、図10に示されるような画像を示す画像データを生成する。すなわち、CPU101は各音声データから、フレーム毎に音程(ピッチ)および音量を算出する。CPU101は、図10に示されるように、音量に対応させて図形の幅を、音程に対応させて図形の上下方向の位置(座標)を決定する。CPU101は、処理対象となっているフレームに相当する時間軸上の位置に、決定された位置(上下方向の座標)に決定された幅を有する図形を表示させる画像データを生成する。また、CPU101は、音声波形に対しフレーム毎

10

【0032】

20

CPU101は生成した画像データをディスプレイ105に出力する。ディスプレイ105は、画像データに従って図10に示されるような画像を表示する(ステップS207)。

【0033】

<変形例>

本発明は上述の実施形態に限定されるものではなく、種々の変形実施が可能である。

上述の各実施形態では、CPU101がHDD104に記憶された語学学習プログラムを実行することにより語学学習装置1あるいは語学学習装置2としての機能が実現されたが、図1あるいは図8に示される機能構成要素に相当する電子回路等を用いて、語学学習装置としての機能をハードウェア的に実現してもよい。

30

【0034】

また、上述の各実施形態では、模範音声およびユーザ音声の長さを揃える処理(ステップS103)において、ユーザ音声の長さを模範音声に合わせる態様について説明したが、模範音声の長さをユーザ音声の長さに合わせるようにしてもよい。

【0035】

また、第1実施形態では、模範音声とユーザ音声との差異点を強調する処理は、音量の調整により行う態様について説明したが、さらに付加的な処理を行ってもよい。例えば、模範音声とユーザ音声とで差異があった部分が子音である場合について、模範音声を再生する際に、より模範音声を明確に聞かせるために特定の周波数領域を増幅する構成としてもよい。あるいは、例えば一般に英語におけるlとrの発音においては第2フォルマントおよび第3フォルマントの位置が異なることが知られているため、特定のフォルマントを増幅する強調処理を行ってもよい。この強調処理は、語学学習プログラムに従ってCPU101が行ってもよいし、バンドパスフィルタと増幅器とを用いてハードウェア的に実現してもよい。

40

また、音量の調整によって差異点の強調を行う態様においても、音量は「0」「1」の2段階だけでなく、中間の音量値を用いてもよい。

【0036】

また、第2実施形態では、模範音声画像およびユーザ音声画像を縦に並べて表示する態様について説明したが、両画像の表示の態様はこれに限定されない。例えば、両画像を横に並べても良いし、重ねて表示してもよい。重ねて表示する場合には、模範音声とユーザ

50



音声とで表示色を変えることが望ましい。

また、模範音声画像およびユーザ音声画像を並べて表示する態様において、図10の例では、音程および音量を示す画像と、周波数特性を示す画像とをそれぞれ別個に表示する態様について説明したが、これらを1つの画像で表現してもよい。例えば、第2実施形態では、音程および音量を示す画像は単一の色で表示されたが、これに周波数特性を示す色をつけて表示してもよい。

また、音声画像は図10で例示したものに限定されない。例えば、音声波形を直接表示してもよい。

また、ステップS206における周波数分析の手法は、第2実施形態で説明したものに限定されない。例えば、フレーム毎に音声波形のスペクトル包絡を求め、フォルマント周波数を求めることとしてもよい。フォルマント周波数としては、例えば、第1～第3フォルマントのいずれかを用いることができる。

【0037】

また、第1実施形態で説明した差異点強調処理を行う機能と、第2実施形態で説明した画像生成・表示機能とを同時に具備する語学学習装置を提供することもできる。

【図面の簡単な説明】

【0038】

【図1】本発明の第1実施形態に係る語学学習装置1の機能構成を示すブロック図である。

。

【図2】語学学習装置1のハードウェア構成を示すブロック図である。

【図3】模範音声データベースDB1の内容を示す図である。

【図4】同実施形態に係る語学学習装置1の動作を示すフローチャートである。

【図5】模範音声(A)およびユーザ音声(B)の波形を例示する図である。

【図6】ステップS103の処理をより詳細に示すフローチャートである。

【図7】模範音声(上)およびユーザ音声(下)のスペクトル包絡を例示する図である。

【図8】本発明の第2実施形態に係る語学学習装置2の機能構成を示すブロック図である。

。

【図9】同実施形態に係る語学学習装置2の動作を示すフローチャートである。

【図10】同実施形態において出力される画像を例示する図である。

【符号の説明】

【0039】

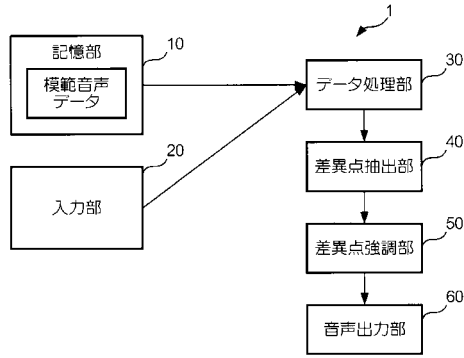
1...語学学習装置、2...語学学習装置、10...記憶部、20...入力部、30...音声処理部、40...差異点抽出部、50...差異点強調部、51...差異点強調部、60...音声出力部、61...差異点表示部、101...CPU、102...RAM、103...ROM、104...HDD、105...ディスプレイ、106...マイク、107...音声処理部、108...スピーカ、109...キーボード、110...バス

10

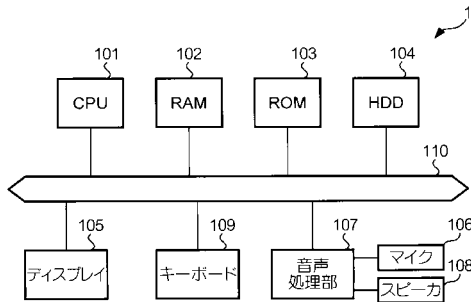
20

30

【図1】



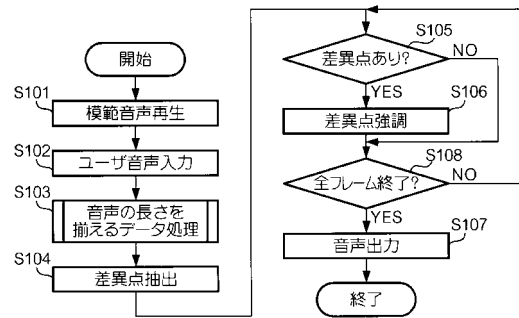
【図2】



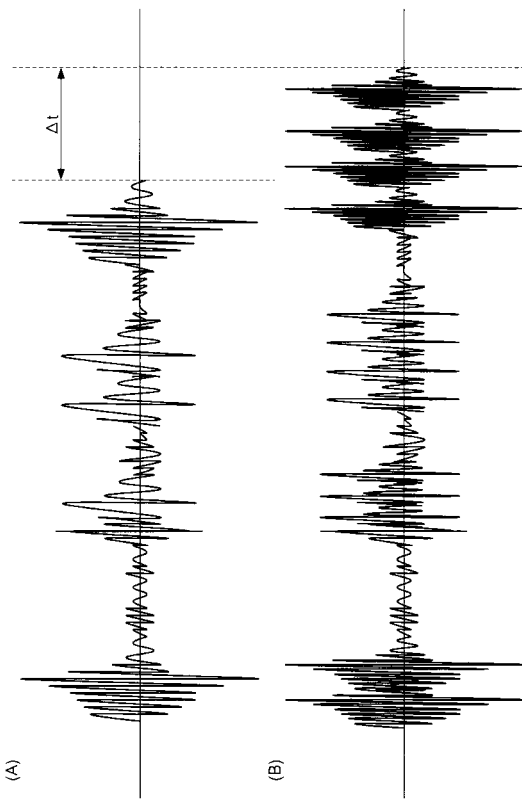
【図3】

識別子	テキストデータ	音声波形データ
001	Good to see you again. How are you?	
002	Good morning. How are you?	
003	Fine thank you. And you?	
004	When did you get up this morning?	
005	I am sorry to hear that.	
006	It is fine today.	
...	...	...

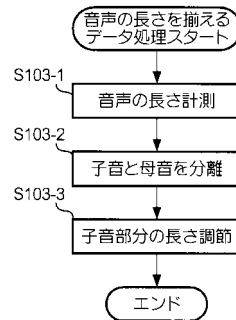
【図4】



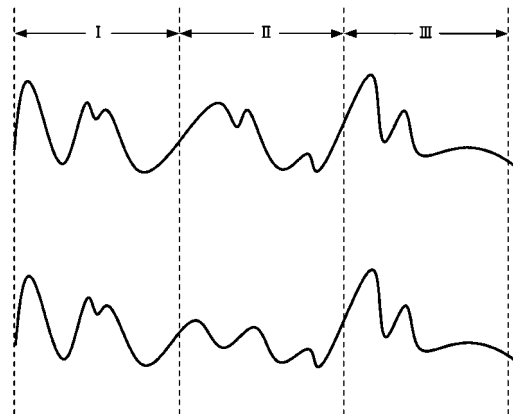
【図5】



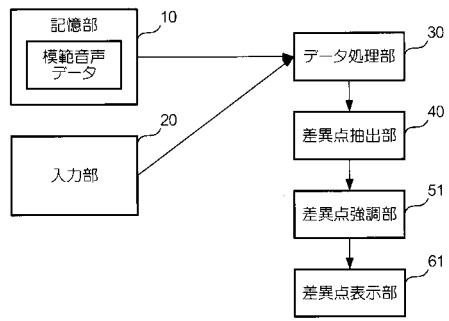
【図6】



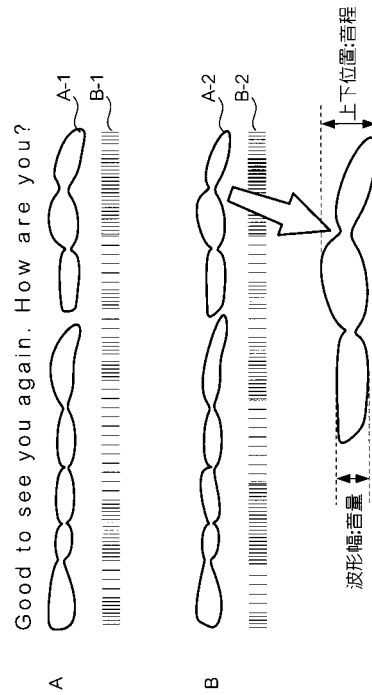
【図7】



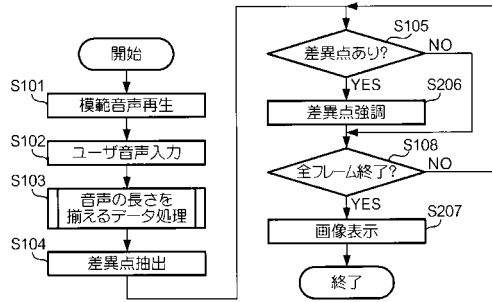
【図8】



【図10】



【図9】



---

フロントページの続き

- (56)参考文献 特開平03 - 273280 (JP, A)  
特開平05 - 232856 (JP, A)  
特開平11 - 143496 (JP, A)  
特開2000 - 075778 (JP, A)  
特開平07 - 295465 (JP, A)  
特開2000 - 162954 (JP, A)  
登録実用新案第3003950 (JP, U)  
特開2003 - 162291 (JP, A)  
特開平02 - 082287 (JP, A)  
特開2000 - 250401 (JP, A)

(58)調査した分野(Int.Cl., DB名)

G09B 1/00 - 19/26