

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4883750号
(P4883750)

(45) 発行日 平成24年2月22日(2012.2.22)

(24) 登録日 平成23年12月16日(2011.12.16)

(51) Int.Cl.

F I

G 1 0 L 11/00 (2006.01)

G 0 9 B 15/00 (2006.01)

G 0 9 B 19/06 (2006.01)

G 1 0 L 15/00 (2006.01)

G 1 0 L 15/10 (2006.01)

G 1 0 L 11/00 4 0 2 F

G 1 0 L 11/00 1 0 1 F

G 1 0 L 11/00 1 0 1 Z

G 0 9 B 15/00 Z

G 0 9 B 19/06

請求項の数 10 (全 22 頁) 最終頁に続く

(21) 出願番号 特願2005-167467 (P2005-167467)
 (22) 出願日 平成17年6月7日(2005.6.7)
 (65) 公開番号 特開2006-227564 (P2006-227564A)
 (43) 公開日 平成18年8月31日(2006.8.31)
 審査請求日 平成20年3月26日(2008.3.26)
 (31) 優先権主張番号 特願2005-13159 (P2005-13159)
 (32) 優先日 平成17年1月20日(2005.1.20)
 (33) 優先権主張国 日本国(JP)

(出願人による申告)平成17年度独立行政法人情報通信研究機構、研究テーマ「人間情報コミュニケーションの研究開発」に関する委託研究、産業活力再生特別措置法第30条の適用を受ける特許出願

(73) 特許権者 393031586
 株式会社国際電気通信基礎技術研究所
 京都府相楽郡精華町光台二丁目2番地2
 (74) 代理人 100115749
 弁理士 谷川 英和
 (72) 発明者 加藤 宏明
 京都府相楽郡精華町光台二丁目2番地2
 株式会社国際電気通信基礎技術研究所内
 (72) 発明者 山田 玲子
 京都府相楽郡精華町光台二丁目2番地2
 株式会社国際電気通信基礎技術研究所内

審査官 毛利 太郎

最終頁に続く

(54) 【発明の名称】 音響評定装置、およびプログラム

(57) 【特許請求の範囲】

【請求項1】

音響の入力を受け付ける音響受付部と、
 前記音響受付部が受け付けた音響から韻律的特徴を示す2種類以上の韻律的特徴情報を抽出する韻律的特徴情報抽出部と、
 前記2種類以上の各韻律的特徴情報に、前記音響の言語の特性に応じた重要度であり、各韻律的特徴情報の重み付けを示す重要度が付与されており、当該2種類以上の各韻律的特徴情報と各韻律的特徴情報の重要度とを用いて、前記音響受付部が受け付けた音響の良し悪しを評定する評定部と、
 前記評定部が評定した結果である評定結果を出力する処理部を具備する音響評定装置。

10

【請求項2】

前記音響は、音声であり、
 前記評定部は、
 言語ごとに良し悪しを評定するための情報であり、言語ごとに前記2種類以上の各韻律的特徴情報の重要度であり、前記音響の言語の特性に応じた重要度を有する言語別評定情報を保持している言語別評定情報格納手段と、
 前記音響受付部が受け付けた音声の言語に対応する言語別評定情報を、前記言語別評定情報格納手段から取得する言語別評定情報取得手段と、
 前記言語別評定情報取得手段が取得した言語別評定情報と、前記韻律的特徴情報に基づいて、前記音響受付部が受け付けた音声の良し悪しを評定する評定手段を具備する請求項1

20

記載の音響評価装置。

【請求項 3】

前記音響受付部が受け付けた音声の言語を判定する種別判定部をさらに具備し、
前記言語別評価情報取得手段は、
前記種別判定部が判定した言語に対応する言語別評価情報を取得する請求項 2 記載の音響
評価装置。

【請求項 4】

前記評価部は、
前記 2 種類以上の各韻律的特徴情報を正規化する正規化手段と、
前記 2 種類以上の韻律的特徴情報ごとに、音響の良し悪しを評価するための情報である模
範評価情報を格納している模範評価情報格納手段と、
前記正規化手段が正規化した 2 種類以上の各韻律的特徴情報と、前記 2 以上の模範評価情
報に基づいて、前記音響受付部が受け付けた音響の良し悪しを評価する評価手段を具備す
る請求項 1 から請求項 3 いずれか記載の音響評価装置。

10

【請求項 5】

前記処理部は、
前記評価結果に基づいて、前記 2 種類以上の各韻律的特徴情報を補正する韻律的特徴情報
補正手段と、
前記韻律的特徴情報補正手段が補正した 2 種類以上の韻律的特徴情報と前記音響受付部が
受け付けた音響に基づいて、音響を合成する音響合成手段と、
前記音響合成手段が合成した音響を出力する音響出力手段を具備する請求項 1 から請求項
4 いずれか記載の音響評価装置。

20

【請求項 6】

前記評価部は、
前記 2 種類以上の各韻律的特徴情報に、前記音響の言語の特性に応じた重要度であり、各
韻律的特徴情報の重み付けを示す重要度が付与されており、当該 2 種類以上の各韻律的特
徴情報と各韻律的特徴情報の重要度とを用いて、かつ、音韻の特性ごとに重み付けして、
前記音響受付部が受け付けた音響の良し悪しを評価する請求項 1 から請求項 5 いずれか記
載の音響評価装置。

【請求項 7】

音響の入力を受け付ける音響受付部と、
前記音響受付部が受け付けた音響から韻律的特徴を示す 2 種類以上の韻律的特徴情報を抽
出する韻律的特徴情報抽出部と、
前記 2 種類以上の韻律的特徴情報ごとに、音響の良し悪しを評価するための情報である模
範評価情報を格納している模範評価情報格納手段と、
前記音響受付部が受け付けた音響を、前記韻律的特徴情報抽出部が取得した 2 種類以上の
韻律的特徴情報と、前記模範評価情報に基づいて補正し、出力する処理部を具備する音響
評価装置。

30

【請求項 8】

前記処理部は、
前記模範評価情報に基づいて、前記 2 種類以上の各韻律的特徴情報を補正する韻律的特徴
情報補正手段と、
前記韻律的特徴情報補正手段が補正した 2 種類以上の韻律的特徴情報と前記音響受付部が
受け付けた音響に基づいて、音響を合成する音響合成手段と、
前記音響合成手段が合成した音響を出力する音響出力手段を具備する請求項 7 記載の音響
評価装置。

40

【請求項 9】

前記韻律的特徴情報は、
音響の時間構造に関する情報である時間構造情報、音響の強さに関する情報である強弱情
報、音響の抑揚に関する情報である抑揚情報のうちの 2 以上の情報である請求項 1 から請

50

求項 8 いずれか記載の音響評定装置。

【請求項 10】

コンピュータに、

音響の入力を受け付ける音響受付ステップと、

前記音響受付ステップで受け付けた音響から韻律的特徴を示す 2 種類以上の韻律的特徴情報を抽出する韻律的特徴情報抽出ステップと、

前記 2 種類以上の各韻律的特徴情報に、前記音響の言語の特性に応じた重要度であり、各韻律的特徴情報の重み付けを示す重要度が付与されており、当該 2 種類以上の各韻律的特徴情報と各韻律的特徴情報の重要度とを用いて前記音響受付ステップで受け付けた音響の良し悪しを評定する評定ステップと、

10

前記評定ステップにおける評定結果を出力する処理ステップを実行させるためのプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、入力された音声や楽音などの音響の良し悪しを評価する装置等に関し、特に、語学学習や音楽演奏の学習等に利用できる音響評定装置等に関するものである。

【背景技術】

【0002】

従来の技術として、以下の語学学習装置がある（特許文献 1 参照）。本語学学習装置は、学習者が選択した役割の発音をレファランデータと比較して一致度によって点数化して表示し、点数によって適当な次の画面を自動に表示することにより、学習能率を向上させる装置である。本従来の語学学習装置は、入力された音声信号は音声認識技術により分析された後、標準音データと一致度が比較されて点数が与えられ、学習者発音のスペクトルと抑揚とが学習者発音表示ボックスに表れるという構成になっている。

20

【特許文献 1】特開 2003 - 228279（第 1 頁、第 1 図等）

【発明の開示】

【発明が解決しようとする課題】

【0003】

しかし、入力された音響の自然性などの音響の良し悪しの評価は、特に、語学学習等において極めて重要であるにも関わらず、従来の語学学習装置は、当該音響の良し悪しを評価する機能を有しなかった。

30

【課題を解決するための手段】

【0004】

本第一の発明の音響評定装置は、音響の入力を受け付ける音響受付部と、前記音響受付部が受け付けた音響から韻律的特徴を示す韻律的特徴情報を抽出する韻律的特徴情報抽出部と、前記韻律的特徴情報に基づいて、前記音響受付部が受け付けた音響の良し悪しを評定する評定部と、前記評定部における評定結果に基づいて、処理を行う処理部を具備する音響評定装置である。

かかる構成により、入力された音響の自然性などの音響の良し悪しの評価ができ、語学等の学習の効果が向上する。

40

【0005】

また、本第二の発明の音響評定装置は、第一の発明の音響評定装置に対して、前記音響は、音声であり、前記評定部は、音声の自然さを示す自然性を評定し、言語ごとに自然性評定のための情報である言語別評定情報を保持している言語別評定情報格納手段と、前記音響受付部が受け付けた音声の言語に対応する言語別評定情報を、前記言語別評定情報格納手段から取得する言語別評定情報取得手段と、前記言語別評定情報取得手段が取得した言語別評定情報と、前記韻律的特徴情報に基づいて、前記音響受付部が受け付けた音声の自然性を評定する評定手段を具備する音響評定装置である。

かかる構成により、各言語に適した評定方法で精度高く、音声の自然性の評定ができる

50

。

【 0 0 0 6 】

また、本第三の発明の音響評定装置は、音響の入力を受け付ける音響受付部と、前記音響受付部が受け付けた音響から韻律的特徴を示す韻律的特徴情報を抽出する韻律的特徴情報抽出部と、音響の良し悪しを評定するための情報である模範評定情報を格納している模範評定情報格納手段と、前記音響受付部が受け付けた音響を、前記模範評定情報に基づいて補正し、出力する処理部を具備する音響評定装置である。

かかる構成により、入力された音響の特徴を残しながら、模範的な音響を出力でき、語学等の学習の効果が大幅に向上する。

【 発明の効果 】

10

【 0 0 0 7 】

本発明による音響評定装置によれば、入力された音響の良し悪しの評価ができたり、または、模範的な音響を出力できたりする機能を有し、語学等の学習の効果が向上する。

【 発明を実施するための最良の形態 】

【 0 0 0 8 】

以下、音響評定装置等の実施形態について図面を参照して説明する。なお、実施の形態において同じ符号を付した構成要素は同様の動作を行うので、再度の説明を省略する場合がある。

（ 実施の形態 1 ）

【 0 0 0 9 】

20

本実施の形態における音響評定装置は、たとえば、英語や中国語などの語学学習等に利用される装置であり、入力された音響の良し悪しを評価し、当該評価結果を出力する装置である。なお、本実施の形態において、主として、音響は音声であり、音響の良し悪しは音声の自然性である。しかし、音響は、音声以外の楽音等であってもよく、音響が楽音の場合は、評価対象は模範の楽音との類似度となる。

図 1 は、本実施の形態における音響評定装置のブロック図である。本音響評定装置は、音響受付部 1 0 1、種別判定部 1 0 2、韻律的特徴情報抽出部 1 0 3、評定部 1 0 4、処理部 1 0 5 を具備する。

評定部 1 0 4 は、模範評定情報格納手段 1 0 4 1、言語別評定情報格納手段 1 0 4 2、言語別評定情報取得手段 1 0 4 3、正規化手段 1 0 4 4、評定手段 1 0 4 5 を具備する。

30

【 0 0 1 0 】

音響受付部 1 0 1 は、音響の入力を受け付ける。音響とは、音声や楽音などである。楽音とは、楽器の演奏により出力される音である。音響受付部 1 0 1 は、例えば、マイクとそのドライバーソフト、またはマイクのドライバーソフト等により実現され得る。また、音響は、マイクだけではなく、磁気テープや C D - R O M などの記録媒体から読み出されても良い。なお、ここでは、音響は主として、音声として説明する。

【 0 0 1 1 】

種別判定部 1 0 2 は、音響受付部 1 0 1 が受け付けた音声の言語を判定する。言語とは、例えば、日本語、英語、中国語、韓国語等である。音声の言語を判定する技術は、公知技術であるので、詳細な説明は省略する。種別判定部 1 0 2 は、通常、M P U やメモリ等から実現され得る。種別判定部 1 0 2 の処理手順は、通常、ソフトウェアで実現され、当該ソフトウェアは R O M 等の記録媒体に記録されている。但し、ハードウェア（専用回路）で実現しても良い。

40

【 0 0 1 2 】

韻律的特徴情報抽出部 1 0 3 は、音響受付部 1 0 1 が受け付けた音声から韻律的特徴を示す韻律的特徴情報を抽出する。韻律的特徴情報は、音声の時間構造に関する情報である時間構造情報、音声の強さに関する情報である強弱情報、音声の抑揚に関する情報である抑揚情報のうちの 1 以上の情報である。韻律的特徴情報抽出部 1 0 3 が抽出する韻律的特徴情報は、音声のどの単位（音韻、単語など）についての情報であっても良い。韻律的特徴情報抽出部 1 0 3 は、通常、M P U やメモリ等から実現され得る。韻律的特徴情報抽出

50

部 1 0 3 の処理手順は、通常、ソフトウェアで実現され、当該ソフトウェアは R O M 等の記録媒体に記録されている。但し、ハードウェア（専用回路）で実現しても良い。

【 0 0 1 3 】

評定部 1 0 4 は、韻律的特徴情報抽出部 1 0 3 が抽出した韻律的特徴情報に基づいて、音響受付部 1 0 1 が受け付けた音声の良し悪しを評定する。評定部 1 0 4 は、韻律的特徴情報を構成する 2 以上の情報の、それぞれについて良し悪しを評定しても良いし、一つの総合点を算出しても良い。ここで、音声の良し悪しとは、例えば、音声の自然性である。また、評定部 1 0 4 は、総合点の算出のために、通常、1 以上の韻律的特徴情報の評定結果を使用する。評定部 1 0 4 は、通常、M P U やメモリ等から実現され得る。評定部 1 0 4 の処理手順は、通常、ソフトウェアで実現され、当該ソフトウェアは R O M 等の記録媒体に記録されている。但し、ハードウェア（専用回路）で実現しても良い。

10

【 0 0 1 4 】

処理部 1 0 5 は、評定部 1 0 4 における評定結果に基づいて、処理を行う。ここでは、処理部 1 0 5 は、評定結果を出力する。出力する評定結果は、時間構造情報、強弱情報、および抑揚情報のそれぞれについての評定結果でも良いし、それらの合計点でも良いし、良いか悪いかを示す情報等でも良い。ここで、出力とは、ディスプレイへの表示、プリンタへの印字、音出力、外部の装置への送信、記録媒体への蓄積等を含む概念である。処理部 1 0 5 は、例えば、ディスプレイやスピーカー等の出力デバイスを含むと考えると含まないとも考えても良い。処理部 1 0 5 は、出力デバイスのドライバソフトまたは、出力デバイスのドライバソフトと出力デバイス等で実現され得る。

20

【 0 0 1 5 】

模範評定情報格納手段 1 0 4 1 は、音声の良し悪しを評定するための情報である模範評定情報を格納している。模範評定情報は、例えば、模範となる発声者（以下、「模範者」ともいう。）が学習対象の文章などを発声して、当該音声から抽出した情報でも良いし、複数人の模範者の音声から学習した情報でも良いし、コンピュータ処理により作り出したモデルデータでも良い。模範評定情報格納手段 1 0 4 1 は、ハードディスクや R O M 等の不揮発性の記録媒体が好適であるが、R A M 等の揮発性の記録媒体でも実現可能である。

【 0 0 1 6 】

言語別評定情報格納手段 1 0 4 2 は、言語ごとに自然性を評定するための情報である言語別評定情報を保持している。言語別評定情報は、例えば、言語別の韻律的特徴情報別の重み付けの情報である。その具体例は後述する。なお、言語別評定情報の構造は問わない。言語別評定情報格納手段 1 0 4 2 は、ハードディスクや R O M 等の不揮発性の記録媒体が好適であるが、R A M 等の揮発性の記録媒体でも実現可能である。

30

言語別評定情報取得手段 1 0 4 3 は、音響受付部 1 0 1 が受け付けた音響の言語に対応する言語別評定情報を取得する。

正規化手段 1 0 4 4 は、韻律的特徴情報を正規化する。正規化のアルゴリズム例は後述する。

【 0 0 1 7 】

評定手段 1 0 4 5 は、言語別評定情報取得手段 1 0 4 3 が取得した言語別評定情報と、韻律的特徴情報抽出部 1 0 3 が抽出した韻律的特徴情報に基づいて、音響受付部 1 0 1 が受け付けた音声の良し悪しを評定する。正規化手段 1 0 4 4 が正規化した韻律的特徴情報と、模範評定情報に基づいて、音響受付部 1 0 1 が受け付けた音声の良し悪しを評定する。

40

【 0 0 1 8 】

言語別評定情報取得手段 1 0 4 3、正規化手段 1 0 4 4、および評定手段 1 0 4 5 は、通常、M P U やメモリ等から実現され得る。言語別評定情報取得手段 1 0 4 3 等の処理手順は、通常、ソフトウェアで実現され、当該ソフトウェアは R O M 等の記録媒体に記録されている。但し、ハードウェア（専用回路）で実現しても良い。

以下、本音響評定装置の動作について図 2 のフローチャートを用いて説明する。

（ステップ S 2 0 1）音響受付部 1 0 1 は、評定する対象の音響（ここでは、音声）を

50

受け付けたか否かを判断する。音声を受け付ければステップ S 2 0 2 に行き、音声を受け付けなければステップ S 2 0 1 に戻る。

(ステップ S 2 0 2) 種別判定部 1 0 2 は、ステップ S 2 0 1 で受け付けた音声の言語を判別する。

【 0 0 1 9 】

(ステップ S 2 0 3) 韻律的特徴情報抽出部 1 0 3 は、ステップ S 2 0 1 で受け付けた音声から韻律的特徴を示す韻律的特徴情報を抽出する。韻律的特徴とは、ここでは、時間構造情報、強弱情報、抑揚情報である。ただし、韻律的特徴は、他の情報を含んでも良いし、時間構造情報、強弱情報、抑揚情報のうちの 2 以下の情報でも良い。

【 0 0 2 0 】

(ステップ S 2 0 4) 正規化手段 1 0 4 4 は、ステップ S 2 0 3 で取得した韻律的特徴情報のうちの時間構造情報を正規化する。時間構造情報の正規化とは、ステップ S 2 0 1 で受け付けた音声と、模範となる音声の発話速度を揃えるための処理である。つまり、発話速度は、一般に、発話の自然性など、発話の良し悪しとは無関係である、という考えに基づく。時間構造情報の正規化において、例えば、正規化手段 1 0 4 4 は、ステップ S 2 0 3 で取得した時間構造情報が示す発話全体の時間長と、模範となる時間構造情報(以下、適宜「模範時間構造情報」という。)が示す発話全体の時間長を同じにするために、時間構造情報の時間情報を短縮、または伸長する。なお、正規化の方法として、発話の全体長を揃える方法以外に、各音韻の時間長を平均した長さを揃える方法や、母音のみに着目して、その開始時間間隔の平均値を揃える方法などでも良い。母音のみに着目して、その開始時間間隔の平均値を揃える方法は、母音の開始点の間隔が、人が知覚する発話速度をよく反映している点が、優れている。

(ステップ S 2 0 5) 評定手段 1 0 4 5 は、模範評定情報格納手段 1 0 4 1 から模範時間構造情報を取得する。

【 0 0 2 1 】

(ステップ S 2 0 6) 評定手段 1 0 4 5 は、ステップ S 2 0 4 で正規化した時間構造情報(以下、適宜「正規化時間構造情報」という。)と、ステップ S 2 0 5 で取得した模範時間構造情報に基づいて、受け付けた音声の時間構造情報について評定する。かかる評定アルゴリズムは、問わない。例えば、評定手段 1 0 4 5 は、正規化時間構造情報と模範時間構造情報の音韻ごとの差(絶対値)の合計をパラメータとして、時間構造情報についての評定値(以下、適宜「時間構造情報評定値」という。)を算出する。なお、通常、正規化時間構造情報と模範時間構造情報の音韻ごとの差(絶対値)の合計が低いほど、模範の音声に近い、という結果であり、時間構造情報評定値は高得点となる。なお、正規化時間構造情報と模範時間構造情報の音韻ごとの差(例えば、「t o k k u m i a i」の中の同じ「t」の時間の情報の差)を取得するのではなく、評定手段 1 0 4 5 は、複数の音韻(例えば、「t o k k u m i a i」の中の「t o」)の正規化時間構造情報と、対応する複数の音韻(例えば、「t o」)の模範時間構造情報における差を取得して、時間構造情報評定値を算出しても良い。かかる算出方法の方が、知覚に則した評価となり好適である場合も多い。

【 0 0 2 2 】

(ステップ S 2 0 7) 正規化手段 1 0 4 4 は、ステップ S 2 0 3 で取得した韻律的特徴情報のうちの強弱情報を正規化する。強弱情報の正規化とは、ステップ S 2 0 1 で受け付けた音声と、模範となる音声の大きさを揃えるための処理である。つまり、一般に、声の大きさと発声の良し悪しとは無関係である、という考えに基づく。特に、録音された声の大きさは、話された時点でのもともとの声の大きさに加えて、話者とマイクの距離、設定された録音レベルなどが影響する。これらは自然性などの音声の良し悪しには関係しない。強弱情報の正規化には、例えば、発話全体の平均的な強さあるいは音韻毎の強さを使用する。なお、音圧レベルに基づく尺度に加えて、人間の耳の感度を考慮したいいくつかの方法が国際規格として確立されている(例えば、A特性補正, ラウドネスなど。)。強弱情報の正規化の方法として、発話全体の大きさを揃える方法の他に、各音韻の大きさを平均

10

20

30

40

50

したものを揃える方法、母音のみに着目して、その平均的大きさを揃える方法等がある。この母音のみに着目して、その平均的大きさを揃える方法は、音声コミュニケーションにおいて母音は子音が担う情報の拡声器の役割を果たしており、人が知覚する声の大きさの印象はほぼ母音の大きさによって決まるという考え方に添っており良好なものである。

(ステップS208) 評価手段1045は、模範評価情報格納手段1041から模範となる強弱情報(以下、適宜「模範強弱情報」という。)を取得する。

【0023】

(ステップS209) 評価手段1045は、ステップS207で正規化した強弱情報(以下、適宜「正規化強弱情報」という。)と、ステップS206で取得した模範強弱情報に基づいて、受け付けた音声の強弱情報について評価する。かかる評価アルゴリズムは、問わない。例えば、評価手段1045は、正規化強弱情報と模範強弱情報の音韻ごとの差(絶対値)の合計をパラメータとして、強弱情報についての評価値(以下、適宜「強弱情報評価値」という。)を算出する。なお、通常、正規化強弱情報と模範強弱情報の音韻ごとの差(絶対値)の合計が低いほど、模範の音声に近い、という結果であり、強弱情報評価値は高得点となる。

【0024】

(ステップS210) 正規化手段1044は、ステップS203で取得した韻律的特徴情報のうちの抑揚情報を正規化する。抑揚情報の正規化とは、ステップS201で受け付けた音声と、模範となる音声の高さを揃えるための処理である。一般に、声の平均的な高さや自然性などの音声の良し悪しは無関係であるという考えに基づく。平均的な声の高さは、体の大きさと性別に大きく関係しており、通常、子供が最も声が高く、次いで大人の女性、そして大人の男性が最も声が高い。また、それぞれのグループの中でも人によって、声の高低は異なる。つまり、平均的な声の高さは人によって異なるが、一般にはこの声の高低が発話の自然性評価などの発話の良し悪しの評価に無関係である、と考えられる。抑揚情報の正規化には、例えば、発話全体の平均的な声の高さあるいは音韻毎の高さを使用する。ここで、声の高さは、例えば、基本周波数(音声・音響分野では“F0(エフゼロ)”)と略称することが多い。)、または、ピッチ感と相関のあるその他の尺度(例えば、ささやき声のスペクトル重心周波数)等を尺度とする。

(ステップS211) 評価手段1045は、模範評価情報格納手段1041から模範となる抑揚情報(以下、適宜「模範抑揚情報」という。)を取得する。

【0025】

(ステップS212) 評価手段1045は、ステップS207で正規化した抑揚情報(以下、適宜「正規化抑揚情報」という。)と、ステップS206で取得した模範抑揚情報に基づいて、受け付けた音声の抑揚情報について評価する。かかる評価アルゴリズムは、問わない。例えば、評価手段1045は、正規化抑揚情報が有する振れ幅と模範抑揚情報が有する振れ幅の差をパラメータとして、抑揚情報についての評価値(以下、適宜「抑揚情報評価値」という。)を算出する。なお、通常、正規化抑揚情報が有する振れ幅と模範抑揚情報が有する振れ幅の差が小さいほど、模範の音声に近い、という結果であり、抑揚情報評価値は高得点となる。

(ステップS213) 言語別評価情報取得手段1043は、ステップS202で判定された言語に対応する言語別評価情報を、言語別評価情報格納手段1042から取得する。

【0026】

(ステップS214) 評価手段1045は、ステップS206で評価した時間構造情報評価値、ステップS209で評価した強弱情報評価値、ステップS212で評価した抑揚情報評価値、およびステップS213で取得した言語別評価情報に基づいて、評価結果を算出する。評価結果の具体的な算出アルゴリズムの例は後述する。

【0027】

(ステップS215) 処理部105は、ステップS214で算出した評価結果を出力する。なお、評価結果の出力態様は、問わない。つまり、評価の最終結果(総合点)のみを出力しても良いし、音韻毎の点数を出力しても良いし、韻律的特徴情報ごと(時間構造情

10

20

30

40

50

報、強弱情報、および抑揚情報)に、その評定値を出力しても良い。ステップS201に戻る。

【0028】

また、上記において、正規化の処理(ステップS204、ステップS207、ステップS210の処理)は、評定の精度を上げるために重要であるが、音声の自然性等を評定する場合に、正規化の処理は行わなくても良い。

なお、図2のフローチャートにおいて、電源オフや処理終了の割り込みにより処理は終了する。

以下、本実施の形態における音響評定装置の具体的な動作について説明する。

【0029】

音響評定装置の言語別評定情報格納手段1042は、例えば、図3に示す言語別評定情報を格納している。言語別評定情報は、「言語」、「時間構造」、「強弱」、「抑揚」の情報を有するレコードを1以上、保持している。「言語」は、言語を示す。「時間構造」「強弱」「抑揚」は、3つの韻律的特徴情報の重要度を示す比率を示す。つまり、この言語別評定情報は、評定値を算出する場合に使用する情報であって、言語別の各韻律的特徴情報の重み付けを示す情報である。具体的には、図3は、日本語の音声を評定する場合に、時間構造情報と抑揚情報を同等に重視し、強弱情報は、あまり重視しないことを示す。

【0030】

図4は、模範評定情報格納手段1041に格納されている模範時間構造情報を示す。模範時間構造情報は、「音韻」「時間(ms)」を有するレコードを1以上保持している。「音韻」は、評定する発話を音韻に区切ったものである。なお、本具体例において、評定対象の発話は、「とっくみあいはい、ゆうがたまでつづいた」である。「時間(ms)」は、各音韻の長さ(ms)を示す。なお、図4において、模範時間構造情報は、音韻ごとの時間の情報であるが、単語ごと等の時間の情報でも良い。つまり、模範時間構造情報は、音声の時間構造に関する情報であり、模範となる情報であれば良い。

【0031】

図5は、模範評定情報格納手段1041に格納されている模範強弱情報を示す。模範強弱情報は、「音韻」「強弱情報」を有するレコードを1以上保持している。「強弱情報」は、ここでは、音韻毎の音圧レベルの相対値である。なお、図5において、模範強弱情報は、音韻ごとの強弱情報であるが、単語ごと等の強弱情報でも良い。つまり、模範強弱情報は、音声の強さに関する情報であり、模範となる情報であれば良い。

【0032】

図6は、模範評定情報格納手段1041に格納されている模範抑揚情報を示す。模範抑揚情報は、「文節(句)」「代表値(Hz)」「振れ幅」を有するレコードを1以上有する。模範抑揚情報は、文節毎の「代表値(Hz)」「振れ幅」を有する。「代表値(Hz)」とは、例えば、中間値である。また、振れ幅は、文節内の最大の基本周波数と、最小の基本周波数の差である。

かかる場合、例えば、日本語を学習する外国人は、「とっくみあいはい、ゆうがたまでつづいた」と、音響評定装置のマイクに向かって発話する。

次に、マイク(音響受付部101の一部)は、当該外国人の音声を受け付ける。そして、種別判定部102は、所定のアルゴリズムにより、言語を「日本語」と判別する。

【0033】

次に、韻律的特徴情報抽出部103は、音響受付部101が受け付けた音声を分析し、韻律的特徴情報を抽出する。具体的には、韻律的特徴情報抽出部103は、例えば、図7に示す時間構造情報、および強弱情報を得る。また、図7に示す情報から、例えば、韻律的特徴情報抽出部103は、音韻毎の時間(ms)である時間構造情報を得る(図9の属性「(1)時間(ms)」参照)。また、例えば、韻律的特徴情報抽出部103は、図7に示す情報から、音韻毎の平均の音の強弱を示す情報である強弱情報を得る(図10の属性「(1)強弱情報」参照)。

【0034】

10

20

30

40

50

また、韻律的特徴情報抽出部 103 は、例えば、図 8 に示す抑揚情報を得る。さらに、韻律的特徴情報抽出部 103 は、図 8 に示す抑揚情報から文節毎の代表値の基本周波数、および振れ幅の情報を取得する、とする（図 11 参照）。なお、図 11 は、韻律的特徴情報抽出部 103 が、最終的に抽出した抑揚情報であり、模範抑揚情報（図 6 参照）と同様の構造である。

【0035】

次に、正規化手段 1044 は、時間構造情報を正規化する。つまり、模範時間構造情報が示す全体発話長「2320 (ms)」に対して、韻律的特徴情報抽出部 103 が抽出した時間構造情報が示す全体発話長は「2500 (ms)」である。そこで、韻律的特徴情報抽出部 103 が抽出した時間構造情報が示す全体発話長が「2320 (ms)」になるように、図 9 の属性値「(1) 時間 (ms)」を短縮する。そして、正規化手段 1044 は、図 9 の属性値「(2) 正規化後」の時間構造情報を得る。

10

【0036】

次に、評価手段 1045 は、図 9 の属性値「(2) 正規化後」の時間構造情報と、模範時間構造情報（図 9 の属性値「(3) 模範」）との差異を算出する。かかる差異（絶対値）は、図 9 の属性値「(4) 差異（絶対値）」である。そして、評価手段 1045 は、図 9 の属性値「(4) 差異（絶対値）」の合計「363」を得る。かかる値が、評価対象の発話と模範音声との、時間構造情報に関する差異となる。そして、時間構造情報評価値は、この差異に基づいて算出される。評価手段 1045 は、例えば、「時間構造情報評価値 = f_1 （差異（絶対値））」により、時間構造情報評価値を算出する。ここで、 $f_1(x)$ は、 x の値が大きくなればなるほど、時間構造情報評価値が小さくなる関数である。

20

【0037】

次に、正規化手段 1044 は、強弱情報を正規化する。つまり、模範強弱情報が示す音韻毎の強弱情報の全体平均値が示す強弱情報「11.6」に対して、韻律的特徴情報抽出部 103 が抽出した音韻毎の強弱情報が示す平均の強弱情報は「7.66」である。そこで、正規化手段 1044 は、韻律的特徴情報抽出部 103 が抽出した音韻毎の強弱情報を「11.6 / 7.66」倍し、正規化強弱情報を得る。正規化強弱情報は、図 10 の属性値「(2) 正規化後」である。

【0038】

次に、評価手段 1045 は、図 10 の属性値「(2) 正規化後」の正規化強弱情報と、模範時間構造情報（図 10 の属性値「(3) 模範」）との差異を算出する。かかる差異（絶対値）は、図 10 の属性値「(4) 差異（絶対値）」である。そして、評価手段 1045 は、図 10 の属性値「(4) 差異（絶対値）」の合計「85.43」を得る。かかる値が、評価対象の発話と模範音声との、強弱情報に関する差異となる。そして、強弱情報評価値は、この差異に基づいて算出される。評価手段 1045 は、例えば、「強弱情報評価値 = f_2 （差異（絶対値））」により、強弱情報評価値を算出する。ここで、 $f_2(x)$ は、 x の値が大きくなればなるほど、強弱情報評価値が小さくなる関数である。

30

【0039】

次に、正規化手段 1044 は、抑揚情報を正規化する。つまり、例えば、抑揚情報は、文節内での代表値の基本周波数、および発話全体での振れ幅である、とする。かかる場合、正規化手段 1044 は、例えば、韻律的特徴情報抽出部 103 が取得した図 8 のグラフの基本周波数の代表値が、模範抑揚情報の基本周波数の代表値と等しくなるように、図 8 のグラフ（抑揚情報を構成する情報）を上または下に移動させる。かかる処理が正規化である。そして、正規化手段 1044 は、正規化後のグラフの代表値の基本周波数を得る。なお、発話全体での振れ幅は、ここでは正規化の影響を受けない、とする。

40

【0040】

次に、評価手段 1045 は、正規化手段 1044 が取得した代表値の基本周波数と模範抑揚情報が有する代表値の基本周波数との差、および、韻律的特徴情報抽出部 103 が取得した振れ幅と模範抑揚情報が有する振れ幅との差の 2 種類の情報をパラメータとして、抑揚情報評価値を算出する。なお、上記 2 種類の差の情報は、文節毎に取得する。つまり

50

、評定手段 1045 は、中間的に、図 11 の表を得る、とする。そして、評定手段 1045 は、図 11 の表（取得した抑揚情報）と図 6 の表（模範抑揚情報）に基づいて、それぞれの差を算出し、当該差から抑揚情報評定値を算出する。具体的には、例えば、評定手段 1045 は、「 $(1880 - 1720) + (1630 - 1590) + (410 - 250) + (628 - 420) = 568$ 」を得る。そして、評定手段 1045 は、として、「抑揚情報評定値 = $f_3(568)$ 」により、抑揚情報評定値を算出する。ここで、 $f_3(x)$ は、 x の値が大きくなればなるほど、抑揚情報評定値が小さくなる関数である。

次に、言語別評定情報取得手段 1043 は、判定された言語「日本語」に対応する言語別評定情報「時間構造：0.4、強弱：0.2、抑揚：0.4」を、言語別評定情報格納手段 1042 から取得する。

10

次に、評定手段 1045 は、例えば、「 $f = 0.4 \times \text{時間構造情報評定値} + 0.2 \times \text{強弱情報評定値} + 0.4 \times \text{抑揚情報評定値}$ 」により、総合的な評定値を算出する。

【0041】

さらに、処理部 105 は、上記算出した評定結果「 f の演算結果」を出力する。出力の態様は、「76点」などの点数でも良いし、評価対象の音声（音響）のどこがどう悪いのかを出力しても良い。かかる場合、処理部 105 は、例えば、図 9 や図 10 の表を、そのまま出力し、例えば、差異が所定の値より大きい音韻の文字色や背景色を、他の音韻の文字色や背景色と区別して、目立つように出力することは好適である。かかる出力態様により、韻律的特徴情報毎に、どの部分で模範音声と大きく食い違っていたのかが一目瞭然に分かり、好適である。

20

以上、本実施の形態によれば、入力された音響の自然性などの音響の良し悪しの評価ができ、語学等の学習の効果が向上する。また、各言語に適した評定方法で精度高く、音声の自然性の評定ができる。

【0042】

なお、本実施の形態によれば、主として、音響は音声であるとして説明した。しかし、音響は、音声以外の楽音等であってもよく、音響が楽音の場合は、評価対象は模範の楽音との類似度となる。かかる場合、本音響評定装置は、楽器の演奏教育に利用され得る。その他、本音響評定装置は、模範となる音響との類似度を評定する装置であれば、そのアプリケーションは問わない。

【0043】

30

また、本実施の形態によれば、音響評定装置は、言語別評定情報を保持しており、言語に適した評定を行ったが、当該構成は必須ではない。本音響評定装置は、音響の入力を受け付ける音響受付部と、前記音響受付部が受け付けた音響から韻律的特徴を示す韻律的特徴情報を抽出する韻律的特徴情報抽出部と、前記韻律的特徴情報に基づいて、前記音響受付部が受け付けた音響の良し悪しを評定する評定部と、前記評定部における評定結果に基づいて、処理を行う処理部を具備すれば良い。

また、本実施の形態において、母音や、強さの値が大きい音韻について、比重を大きくして評価するなど、全体の評価の際に、音韻の特性ごとに重み付けして評価することは好適である。母音や、強さの値が大きい音韻が模範となる音響に近い場合は、模範の音響と、より類似している、と評価する方が、人間の知覚に則した評価となる。一方、人間の知覚にとって、一般に、子音や強さの弱い音韻の影響度合いは少ない。

40

また、本実施の形態において、正規化手段は必須ではない。

【0044】

さらに、本実施の形態における処理は、ソフトウェアで実現しても良い。そして、このソフトウェアをソフトウェアダウンロード等により配布しても良い。また、このソフトウェアを CD-ROM などの記録媒体に記録して流布しても良い。なお、このことは、本明細書における他の実施の形態においても該当する。なお、本実施の形態における情報処理装置を実現するソフトウェアは、以下のようなプログラムである。つまり、このプログラムは、コンピュータに、音響の入力を受け付ける音響受付ステップと、前記音響受付ステップで受け付けた音響から韻律的特徴を示す韻律的特徴情報を抽出する韻律的特徴情報抽

50

出ステップと、前記韻律的特徴情報に基づいて、前記音響受付ステップで受け付けた音響の良し悪しを評定する評定ステップと、前記評定ステップにおける評定結果に基づいて、処理を行う処理ステップを実行させるためのプログラム、である。なお、前記処理ステップにおいて、前記評定結果を出力する、ことは好適である。

また、上記プログラムにおいて、前記音響は、音声であり、前記評定ステップにおいて、前記音響受付ステップで受け付けた音声の言語に対応する言語別評定情報を、取得する言語別評定情報取得サブステップと、前記言語別評定情報取得サブステップで取得した言語別評定情報と、前記韻律的特徴情報に基づいて、前記音響受付ステップで受け付けた音声の自然性を評定する評定サブステップを具備することは好適である。

10

【0045】

また、上記プログラムにおいて、コンピュータに、前記音響受付ステップで受け付けた音声の言語を判定する種別判定ステップをさらに実行させ、前記言語別評定情報取得サブステップは、前記種別判定ステップで判定した言語に対応する言語別評定情報を取得することは好適である。

【0046】

さらに、上記プログラムにおいて、評定ステップは、前記韻律的特徴情報を正規化する正規化サブステップと、前記正規化サブステップで正規化した韻律的特徴情報と、格納している模範評定情報に基づいて、前記音響受付ステップで受け付けた音響の良し悪しを評定する評定サブステップを具備しても良い。

20

(実施の形態2)

【0047】

本実施の形態における音響評定装置は、たとえば、英語や中国語などの語学学習等に利用される装置であり、入力された音響を好適な音響に補正し、模範的な音響にして出力する装置である。なお、本実施の形態において、主として、音響は音声である。しかし、音響は、音声以外の楽音等であってもよく、音響が楽音の場合は、評価対象は模範の楽音との類似度となる。

図12は、本実施の形態における音響評定装置のブロック図である。本音響評定装置は、音響受付部101、種別判定部102、韻律的特徴情報抽出部103、評定部104、処理部1205を具備する。

30

処理部1205は、韻律的特徴情報補正手段12051、音響合成手段12052、音響出力手段12053を具備する。

処理部1205は、評定部104における評定結果に基づいて、音響受付部101が受け付けた音声を補正して、出力する。

【0048】

韻律的特徴情報補正手段12051は、評定部104における評定結果に基づいて、韻律的特徴情報抽出部103が抽出した韻律的特徴情報を補正する。また、韻律的特徴情報補正手段12051は、格納している模範評定情報に基づいて、韻律的特徴情報抽出部103が抽出した韻律的特徴情報を補正しても良い。

音響合成手段12052は、韻律的特徴情報補正手段12051が補正した韻律的特徴情報と音響受付部101が受け付けた音響に基づいて、音響を合成する。

40

【0049】

韻律的特徴情報補正手段12051、音響合成手段12052は、通常、MPUやメモリ等から実現され得る。韻律的特徴情報補正手段12051等の処理手順は、通常、ソフトウェアで実現され、当該ソフトウェアはROM等の記録媒体に記録されている。但し、ハードウェア(専用回路)で実現しても良い。

【0050】

音響出力手段12053は、音響合成手段12052が合成した音響を出力する。出力とは、スピーカーを用いた音出力、外部の装置への送信、記録媒体への蓄積等を含む概念である。音響出力手段12053は、スピーカー等の出力デバイスを含むと考えることも含ま

50

ないと考えても良い。音響出力手段 1 2 0 5 3 は、出力デバイスのドライバーソフトまたは、出力デバイスのドライバーソフトと出力デバイス等で実現され得る。

次に、音響評定装置の動作について図 1 3 のフローチャートを用いて説明する。なお、図 1 3 のフローチャートにおいて、図 2 と異なるステップについてのみ説明する。

【 0 0 5 1 】

(ステップ S 1 3 0 1) 韻律的特徴情報補正手段 1 2 0 5 1 は、ステップ S 2 0 6 で算出した時間構造情報評定値に基づいて、受け付けた音響の時間構造情報を補正する必要があるか否かを判断する。時間構造情報を補正する必要がある場合はステップ S 1 3 0 2 に行き、時間構造情報を補正する必要があるなければステップ S 1 3 0 3 に行く。なお、韻律的特徴情報補正手段 1 2 0 5 1 は、例えば、時間構造情報評定値の補正の閾値を持っており、時間構造情報評定値が当該閾値を超える場合に、補正する必要があると判断する。

10

【 0 0 5 2 】

(ステップ S 1 3 0 2) 韻律的特徴情報補正手段 1 2 0 5 1 は、模範時間構造情報に基づいて、時間構造情報を修正する。なお、韻律的特徴情報補正手段 1 2 0 5 1 は、例えば、単純に、各音韻の時間長が、模範時間構造情報と同一になるようにしても良いし、母音の時間長のみが模範時間構造情報と同一になるようにしても良い。その他、補正のアルゴリズムは問わない。

【 0 0 5 3 】

(ステップ S 1 3 0 3) 韻律的特徴情報補正手段 1 2 0 5 1 は、ステップ S 2 0 9 で算出した強弱情報評定値に基づいて、受け付けた音響の強弱情報を補正する必要があるか否かを判断する。補正する必要がある場合はステップ S 1 3 0 4 に行き、補正する必要があるなければステップ S 1 3 0 5 に行く。なお、韻律的特徴情報補正手段 1 2 0 5 1 は、例えば、強弱情報評定値の補正の閾値を持っており、強弱情報評定値が当該閾値を超える場合に、補正する必要があると判断する。

20

【 0 0 5 4 】

(ステップ S 1 3 0 4) 韻律的特徴情報補正手段 1 2 0 5 1 は、模範強弱情報に基づいて、強弱情報を修正する。なお、韻律的特徴情報補正手段 1 2 0 5 1 は、例えば、単純に、各音韻の強弱情報が、模範強弱情報と同一になるようにしても良いし、母音の強弱情報のみを模範強弱情報と同一になるようにしても良い。その他、補正のアルゴリズムは問わない。

30

【 0 0 5 5 】

(ステップ S 1 3 0 5) 韻律的特徴情報補正手段 1 2 0 5 1 は、ステップ S 2 1 2 で算出した抑揚情報評定値に基づいて、受け付けた音響の抑揚情報を補正する必要があるか否かを判断する。補正する必要がある場合はステップ S 1 3 0 6 に行き、補正する必要があるなければステップ S 1 3 0 7 に行く。なお、韻律的特徴情報補正手段 1 2 0 5 1 は、例えば、抑揚情報評定値の補正の閾値を持っており、抑揚情報評定値が当該閾値を超える場合に、補正する必要があると判断する。

【 0 0 5 6 】

(ステップ S 1 3 0 6) 韻律的特徴情報補正手段 1 2 0 5 1 は、模範抑揚情報に基づいて、抑揚情報を修正する。なお、韻律的特徴情報補正手段 1 2 0 5 1 は、例えば、文節毎の振れ幅を、模範抑揚情報と同一になるようにしても良い。その他、補正のアルゴリズムは問わない。なお、抑揚情報の補正は、入力された音声から抽出された抑揚情報(例えば、図 8 の情報)の一点(例えば、始点や代表値の点)を固定し、抑揚情報が示す形状(例えば、図 8 のグラフが示す形状)を、模範抑揚情報が示す形状と同じになるように修正しても良い。つまり、かかる補正は、入力された音声から抽出された抑揚情報の一点を基点として、当該抑揚情報の他の点との差(相対値)を、模範抑揚情報の点であり、抽出された抑揚情報の一点に対応する一点からの相対値と同一にする修正である。

40

(ステップ S 1 3 0 7) 音響合成手段 1 2 0 5 2 は、ステップ S 2 0 1 で受け付けた音響、上記ステップで補正した韻律的特徴情報(時間構造情報、強弱情報、抑揚情報)に基づいて、音響を合成する。

50

(ステップS 1 3 0 8) 音響出力手段 1 2 0 5 3 は、ステップS 1 3 0 7 で合成した音響を出力する。

【 0 0 5 7 】

なお、図 1 3 のフローチャートにおいて、各韻律的特徴情報を補正する必要があるか否かを判断したのち、補正する必要があると判断した場合のみ補正したが、上記判断をすることなしに、各韻律的特徴情報を補正するようにしても良い。また、補正は、模範となる音声情報の各韻律的特徴情報をそのまま使用する修正でも良い。

また、補正する各韻律的特徴情報は、上記の時間構造情報、強弱情報、抑揚情報のうち、2 以下の韻律的特徴情報であっても良い。

なお、図 1 3 のフローチャートにおいて、電源オフや処理終了の割り込みにより処理は終了する。

10

以下、本実施の形態における音響評定装置の具体的な動作について説明する。ここでは、実施の形態 1 における処理と異なる処理である、韻律的特徴情報を補正して、音響を合成して出力する処理について説明する。

【 0 0 5 8 】

今、実施の形態 1 において説明した処理と同様の処理により、本音響評定装置は、受け付けた音声进行分析し、評定し、図 9 に示す時間構造情報に関する情報、図 1 0 に示す強弱情報に関する情報、図 1 1 に示す抑揚情報に関する情報を得たとする。

【 0 0 5 9 】

次に、韻律的特徴情報補正手段 1 2 0 5 1 は、正規化時間構造情報と模範時間構造情報との差異が 1 0 以上である音韻について、模範時間構造情報の値を適用する。つまり、韻律的特徴情報補正手段 1 2 0 5 1 は、図 1 4 の属性「補正後(1)」を得る。属性「補正後(1)」の網掛けの属性値が、模範の時間構造情報に置き換わった属性値である。次に、韻律的特徴情報補正手段 1 2 0 5 1 は、属性「補正後(1)」の各属性値を、正規化前の元の音声の長さになるように、伸長、または短縮する。なお、ここでは、伸長であり、属性「補正後(1)」の各属性値に「2 5 0 0 / 2 3 2 0」を乗じて、小数点以下を 4 捨五入して、整数化している。その属性値を属性「補正後(2)」に示す。以上の処理により、韻律的特徴情報補正手段 1 2 0 5 1 は、補正した時間構造情報を得る。

20

【 0 0 6 0 】

次に、韻律的特徴情報補正手段 1 2 0 5 1 は、正規化強弱情報と模範強弱情報との差異が 5 以上である音韻について、模範強弱情報の値を適用する。つまり、韻律的特徴情報補正手段 1 2 0 5 1 は、図 1 5 の属性「補正後(1)」を得る。属性「補正後(1)」の網掛けの属性値が、模範の強弱情報に置き換わった属性値である。次に、韻律的特徴情報補正手段 1 2 0 5 1 は、属性「補正後(1)」の各属性値を、正規化前の元の音声の強さになるように属性値に対して正規化した際の補正值の逆数(0 . 6 5 8)を乗じる。そして、受け付けた元の音声の強さに近づける。

30

次に、韻律的特徴情報補正手段 1 2 0 5 1 は、抑揚情報を模範抑揚情報に変更する。なお、抑揚情報についても、上述した補正の可否を判断する等の処理を行ってから補正をしても良い、ことは言うまでもない。

【 0 0 6 1 】

40

以上の処理により、韻律的特徴情報補正手段 1 2 0 5 1 は、韻律的特徴情報の補正を完了する。なお、上記した各韻律的特徴情報の補正アルゴリズムは、一例であることは言うまでもない。例えば、韻律的特徴情報補正手段 1 2 0 5 1 は、時間構造情報や強弱情報も、模範の各情報に修正しても良い。

【 0 0 6 2 】

次に、音響合成手段 1 2 0 5 2 は、上記の補正した韻律的特徴情報と音響受付部 1 0 1 が受け付けた音響に基づいて、音響を合成する。かかる場合、音響合成手段 1 2 0 5 2 は、音響受付部 1 0 1 が受け付けた音響进行分析し、上記の補正した韻律的特徴情報を除いて、受け付けた音響が有する情報をそのまま用いて音響を合成する。

次に、音響出力手段 1 2 0 5 3 は、音響合成手段 1 2 0 5 2 が合成した音響を出力する

50

。

以上、本実施の形態によれば、入力された音響の特徴を残しながら、模範的な音響を出力でき、語学等の学習の効果が大幅に向上する。

【 0 0 6 3 】

なお、本実施の形態によれば、主として、音響は音声であるとして説明した。しかし、音響は、音声以外の楽音等であってもよく、音響が楽音の場合は、出力される音響は、模範の楽音、または模範の楽音に近い楽音となる。かかる場合、本音響評定装置は、楽器の演奏教育に利用され得る。

【 0 0 6 4 】

また、本実施の形態によれば、音響評定装置において、音響の評定処理は必須ではない。本音響評定装置は、音響の入力を受け付ける音響受付部と、前記音響受付部が受け付けた音響から韻律的特徴を示す韻律的特徴情報を抽出する韻律的特徴情報抽出部と、音響の良し悪しを評定するための情報である模範評定情報を格納している模範評定情報格納手段と、前記音響受付部が受け付けた音響を、前記韻律的特徴情報抽出部が取得した韻律的特徴情報と、前記模範評定情報に基づいて補正し、出力する処理部を具備する構成であれば良い。

10

また、本実施の形態によれば、音声合成する前に、所定の要件を満たす韻律的特徴情報（不適切な韻律的特徴情報）を補正したり、模範の韻律的特徴情報に差し替える補正をしたりするアルゴリズムについて説明した。しかし、ユーザ（学習者や演奏者など）が補正したい部分（例えば、始めの単語だけ、あるいは抑揚情報のみ）を指示し、音響評定装置はかかる部分に関する情報を格納しており、かかる補正する部分のみを補正することは好適である。かかる場合、出力された評定結果を見たユーザが、例えば、自分の弱点を補強するために、補正したい部分（例えば、始めの単語だけ、あるいは抑揚情報のみ）を指示し、音響評定装置が当該指示に基づいて、部分を特定する情報を蓄積する。

20

また、本実施の形態によれば、時間構造情報、強弱情報、抑揚情報は、主として一音韻ごとに評価され、また補正された。しかし、2以上の音韻に対して評価し、または補正するようにしても良い。2以上の音韻に対して評価、補正することは、知覚に則しており、好適である場合も多い。

【 0 0 6 5 】

さらに、本実施の形態における処理は、ソフトウェアで実現しても良い。そして、このソフトウェアをソフトウェアダウンロード等により配布しても良い。また、このソフトウェアをCD-ROMなどの記録媒体に記録して流布しても良い。なお、このことは、本明細書における他の実施の形態においても該当する。なお、本実施の形態における情報処理装置を実現するソフトウェアは、以下のようなプログラムである。つまり、このプログラムは、コンピュータに、音響の入力を受け付ける音響受付ステップと、前記音響受付ステップで受け付けた音響から韻律的特徴を示す韻律的特徴情報を抽出する韻律的特徴情報抽出ステップと、前記音響受付ステップで受け付けた音響を、前記韻律的特徴情報抽出ステップで取得した韻律的特徴情報と、格納している模範評定情報に基づいて補正し、出力する処理ステップを実行させるためのプログラム、である。

30

【 0 0 6 6 】

また、上記処理ステップは、前記模範評定情報に基づいて、前記韻律的特徴情報を補正する韻律的特徴情報補正サブステップと、前記韻律的特徴情報補正サブステップで補正した韻律的特徴情報と前記音響受付ステップで受け付けた音響に基づいて、音響を合成する音響合成サブステップと、前記音響合成サブステップで合成した音響を出力する音響出力サブステップを具備する構成でも良い。

40

【 0 0 6 7 】

本明細書で述べた前記韻律的特徴情報は、音響の時間構造に関する情報である時間構造情報、音響の強さに関する情報である強弱情報、音響の抑揚に関する情報である抑揚情報のうちの1以上の情報であることは好適である。

【 0 0 6 8 】

50

また、図 16 は、本明細書で述べたプログラムを実行して、上述した種々の実施の形態の音響評定装置を実現するコンピュータの外観を示す。上述の実施の形態は、コンピュータハードウェア及びその上で実行されるコンピュータプログラムで実現され得る。図 16 は、このコンピュータシステム 160 の概観図であり、図 17 は、システム 160 のブロック図である。

【0069】

図 16 において、コンピュータシステム 160 は、FD (Flexible Disk) ドライブ、CD-ROM (Compact Disk Read Only Memory) ドライブを含むコンピュータ 161 と、キーボード 162 と、マウス 163 と、モニタ 164 と、マイク 165 と、スピーカー 166 とを含む。

10

【0070】

図 17 において、コンピュータ 161 は、FD ドライブ 1611、CD-ROM ドライブ 1612 に加えて、CPU (Central Processing Unit) 1613 と、CPU 1613、CD-ROM ドライブ 1612 及び FD ドライブ 1611 に接続されたバス 1614 と、ブートアッププログラム等のプログラムを記憶するための ROM (Read-Only Memory) 1615 と、CPU 1613 に接続され、アプリケーションプログラムの命令を一時的に記憶するとともに一時記憶空間を提供するための RAM (Random Access Memory) 1616 と、アプリケーションプログラム、システムプログラム、及びデータを記憶するためのハードディスク 1617 とを含む。ここでは、図示しないが、コンピュータ 161 は、さらに、LAN への接続を提供するネットワークカードを含んでも良い。

20

【0071】

コンピュータシステム 160 に、上述した実施の形態の音響評定装置の機能を実行させるプログラムは、CD-ROM 1701、または FD 1702 に記憶されて、CD-ROM ドライブ 1612 または FD ドライブ 1611 に挿入され、さらにハードディスク 1617 に転送されても良い。これに代えて、プログラムは、図示しないネットワークを介してコンピュータ 161 に送信され、ハードディスク 1617 に記憶されても良い。プログラムは実行の際に RAM 1616 にロードされる。プログラムは、CD-ROM 1701、FD 1702 またはネットワークから直接、ロードされても良い。

【0072】

30

プログラムは、コンピュータ 161 に、上述した実施の形態の音響評定装置の機能を実行させるオペレーティングシステム (OS)、またはサードパーティプログラム等は、必ずしも含まなくても良い。プログラムは、制御された態様で適切な機能 (モジュール) を呼び出し、所望の結果が得られるようにする命令の部分のみを含んでいれば良い。コンピュータシステム 160 がどのように動作するかは周知であり、詳細な説明は省略する。

【0073】

なお、上記プログラムにおいて、情報を入力するステップなどでは、ハードウェアによって行われる処理、例えば、出力するステップにおけるモニタなどで行われる処理 (ハードウェアでしか行われない処理) は含まれない。

また、このプログラムを実行するコンピュータは、単数であってもよく、複数であってもよい。すなわち、集中処理を行ってもよく、あるいは分散処理を行ってもよい。

40

【0074】

また、上記各実施の形態において、各処理 (各機能) は、単一の装置 (システム) によって集中処理されることによって実現されてもよく、あるいは、複数の装置によって分散処理されることによって実現されてもよい。

本発明は、以上の実施の形態に限定されることなく、種々の変更が可能であり、それらも本発明の範囲内に包含されるものであることは言うまでもない。

【産業上の利用可能性】

【0075】

以上のように、本発明にかかる音響評定装置は、入力された音響の良し悪しの評価がで

50

きたり、または、模範的な音響を出力できたりする機能を有し、語学学習装置等として有用である。

【図面の簡単な説明】

【 0 0 7 6 】

【図 1】実施の形態 1 における音響評定装置のブロック図

【図 2】同音響評定装置の動作について説明するフローチャート

【図 3】同言語別評定情報の例を示す図

【図 4】同模範時間構造情報の例を示す図

【図 5】同模範強弱情報の例を示す図

【図 6】同模範抑揚情報の例を示す図

10

【図 7】同音響評定装置が取得する時間構造情報、および強弱情報の例を示す図

【図 8】同音響評定装置が取得する抑揚情報の例を示す図

【図 9】同時間構造情報について説明する図

【図 10】同強弱情報について説明する図

【図 11】同抑揚情報について説明する図

【図 12】実施の形態 2 における音響評定装置のブロック図

【図 13】同音響評定装置の動作について説明するフローチャート

【図 14】同時間構造情報の補正について説明する図

【図 15】同強弱情報の補正について説明する図

【図 16】同音響評定装置を構成するコンピュータシステムの概観図

20

【図 17】同音響評定装置を構成するコンピュータのブロック図

【符号の説明】

【 0 0 7 7 】

1 0 1 音響受付部

1 0 2 種別判定部

1 0 3 韻律的特徴情報抽出部

1 0 4 評定部

1 0 5、1 2 0 5 処理部

1 0 4 1 模範評定情報格納手段

1 0 4 2 言語別評定情報格納手段

30

1 0 4 3 言語別評定情報取得手段

1 0 4 4 正規化手段

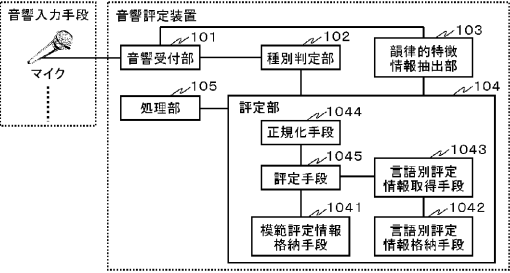
1 0 4 5 評定手段

1 2 0 5 1 韻律的特徴情報補正手段

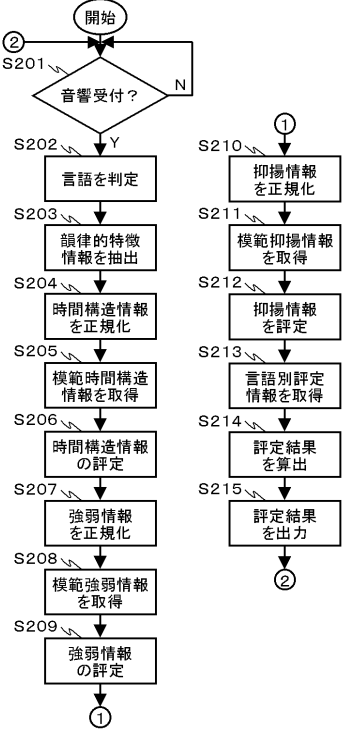
1 2 0 5 2 音響合成手段

1 2 0 5 3 音響出力手段

【図 1】



【図 2】



【図 3】

	時間構造	強弱	抑揚
日本語	0.4	0.2	0.4
英語	0.3	0.35	0.35
中国語	0.25	0.25	0.5

【図 4】

音韻	時間 (ms)
t	30
o	80
k	140
u	70
m	70
i	90
a	110
i	80
w	60
a	150
休止	180
y	70
u	120
g	60
a	90
t	30
a	70
m	40
a	80
d	40
e	80
ts	130
u	50
z	60
u	50
i	90
t	110
a	90

【図 5】

音韻	強弱情報
t	6.36E+00
o	1.57E+01
k	-5.68E+00
u	1.67E+01
m	1.47E+01
i	1.04E+01
a	1.57E+01
i	1.81E+01
w	1.70E+01
a	1.47E+01
休止	-5.82E+00
y	1.61E+01
u	2.10E+01
g	2.31E+01
a	1.77E+01
t	8.43E-01
a	1.56E+01
m	1.74E+01
a	1.66E+01
d	1.07E+01
e	1.54E+01
ts	-3.98E+01
u	9.95E+00
z	4.88E+00
u	1.53E+01
i	1.20E+01
t	-4.40E+00
a	1.06E+01

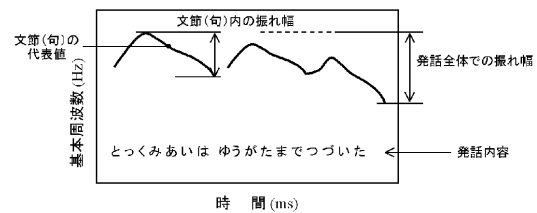
【図 6】

文節(句)	代表値(Hz)	振れ幅
とっくみあいはい	1880	250
ゆうがたまでつづいた	1590	420

【図 7】

時間(ms)	音韻ラベル	強弱情報
10	休止	-1.64E+01
20	t	5.39E+00
30	t	6.08E+00
40	t	7.61E+00
50	o	1.57E+01
60	o	8.78E+00
70	o	1.57E+01
80	o	1.60E+01
90	o	1.70E+01
100	o	1.90E+01
110	o	1.72E+01
120	o	1.59E+01
2390	t	-7.07E+00
2400	t	-3.58E+00
2410	t	2.10E+00
2420	a	3.66E+00
2430	a	1.34E+01
2440	a	1.39E+01
2450	a	1.18E+01
2460	a	1.28E+01
2470	a	1.26E+01
2480	a	1.21E+01
2490	a	6.77E+00
2500	a	8.12E+00

【図 8】



【図 9】

音韻	(1)時間 (ms)	(2)正規化後	(3)模範	(4)差異 (絶対値)
t	35	32	30	2
o	90	83	80	3
k	150	138	140	1
u	103	95	70	25
m	88	81	70	11
i	70	64	90	26
a	89	82	110	28
i	97	90	80	10
w	40	37	60	23
a	120	111	150	39
休止	130	120	180	60
y	70	64	70	6
u	129	119	120	1
g	70	64	80	4
a	102	94	90	4
t	46	42	30	12
a	99	91	70	21
m	47	43	40	3
a	81	75	80	5
d	45	41	40	1
e	102	94	80	14
ts	173	160	130	30
u	66	63	50	13
z	67	62	60	2
u	56	53	50	3
i	102	94	90	4
t	132	122	110	12
a	97	80	90	0

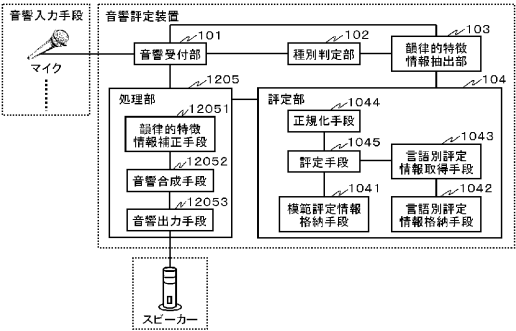
【図 10】

音韻	(1)強弱情報	(2)正規化後	(3)模範	(4)差異 (絶対値)
t	2.36E+00	3.58E+00	6.36E+00	2.78
o	1.17E+01	1.78E+01	1.57E+01	2.07
k	-1.27E+01	-1.93E+01	-5.68E+00	13.58
u	1.17E+01	1.78E+01	1.67E+01	1.07
m	4.70E+00	7.14E+00	1.47E+01	7.56
i	8.40E+00	1.28E+01	1.04E+01	2.36
a	1.17E+01	1.78E+01	1.57E+01	2.07
i	8.10E+00	1.23E+01	1.81E+01	5.80
w	1.20E+01	1.82E+01	1.70E+01	1.22
a	1.17E+01	1.78E+01	1.47E+01	3.07
休止	-8.82E+00	-1.34E+01	-5.82E+00	7.57
y	1.01E+01	1.53E+01	1.61E+01	0.76
u	1.60E+01	2.43E+01	2.10E+01	3.30
g	2.01E+01	3.05E+01	2.31E+01	7.43
a	1.17E+01	1.78E+01	1.77E+01	0.07
t	7.20E-01	1.09E+00	8.43E-01	0.25
a	1.06E+01	1.61E+01	1.56E+01	0.50
m	1.21E+01	1.84E+01	1.74E+01	0.98
a	1.12E+01	1.70E+01	1.66E+01	0.41
d	8.70E+00	1.32E+01	1.07E+01	2.51
e	8.40E+00	1.28E+01	1.54E+01	2.64
ts	-3.90E+00	-5.92E+00	-3.98E+00	1.94
u	9.00E+00	1.37E+01	9.95E+00	3.72
z	4.81E+00	7.30E+00	4.88E+00	2.42
u	1.13E+01	1.72E+01	1.53E+01	1.86
i	1.00E+01	1.52E+01	1.20E+01	3.19
t	4.11E+00	6.24E+00	4.40E+00	1.84
a	8.60E+00	1.31E+01	1.06E+01	2.46

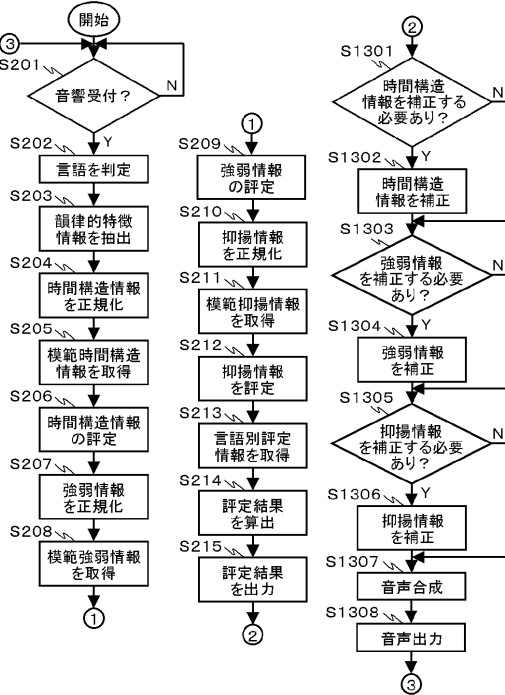
【図 11】

文節 (句)	代表値 (Hz)	振れ幅
とっくみあい	1720	410
ゆうがたまでつづいた	1630	628

【図 12】



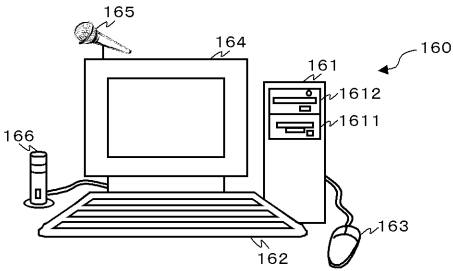
【図 13】



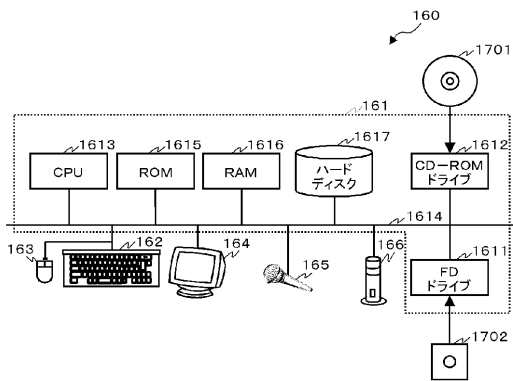
【図 14】

音韻	時間 (ms)	正規化後	極値	差異 (絶対値)	補正後 (1)	補正後 (2)
t	35	32	30	2	32	34
o	90	83	80	3	83	89
k	150	139	140	1	139	149
u	103	95	70	25	79	75
m	88	81	70	11	70	75
i	70	64	90	26	90	96
a	89	82	110	28	110	118
i	97	90	80	10	89	86
w	40	37	60	23	60	64
a	120	111	150	39	150	161
休止	130	120	190	80	190	193
y	70	64	70	6	64	68
u	128	119	120	1	119	128
g	70	64	60	4	64	68
a	102	94	90	4	94	101
t	46	42	30	12	30	32
a	99	91	70	21	70	75
m	47	43	40	3	43	46
a	81	75	80	5	75	80
d	45	41	40	1	41	44
e	102	94	90	14	90	86
ts	173	160	130	30	130	140
u	68	63	50	13	50	53
z	67	62	60	2	62	66
u	58	53	50	3	53	57
i	102	94	90	4	94	101
t	132	122	110	12	118	118
a	97	90	90	0	90	96

【図 16】



【図 17】



【図 15】

音韻	(1)強弱情報	(2)正規化後	(3)模範	(4)差異(絶対値)	補正後(1)	補正後(2)
t	2.36E+00	3.58E+00	6.36E+00	2.78	3.58E+00	2.36E+00
o	1.17E+01	1.78E+01	1.57E+01	2.07	1.78E+01	1.17E+01
k	-1.27E+01	-1.93E+01	-5.68E+00	13.58	-5.68E+00	-3.74E+00
u	1.17E+01	1.78E+01	1.67E+01	1.07	1.78E+01	1.17E+01
m	4.70E+00	7.14E+00	1.47E+01	7.56	1.47E+01	9.68E+00
i	8.40E+00	1.28E+01	1.04E+01	2.36	1.28E+01	8.40E+00
a	1.17E+01	1.78E+01	1.57E+01	2.07	1.78E+01	1.17E+01
i	8.10E+00	1.23E+01	1.81E+01	5.80	1.81E+01	1.19E+01
w	1.20E+01	1.82E+01	1.70E+01	1.22	1.82E+01	1.20E+01
a	1.17E+01	1.78E+01	1.47E+01	3.07	1.78E+01	1.17E+01
休止	-8.82E+00	-1.34E+01	-5.82E+00	7.57	-1.34E+01	-8.82E+00
y	1.01E+01	1.53E+01	1.61E+01	0.76	1.53E+01	1.01E+01
u	1.60E+01	2.43E+01	2.10E+01	3.30	2.43E+01	1.60E+01
ε	2.01E+01	3.05E+01	2.31E+01	7.43	3.05E+01	2.01E+01
a	1.17E+01	1.78E+01	1.77E+01	0.07	1.78E+01	1.17E+01
t	7.20E-01	1.09E+00	8.43E-01	0.25	1.09E+00	7.20E-01
a	1.06E+01	1.61E+01	1.56E+01	0.50	1.61E+01	1.06E+01
m	1.21E+01	1.84E+01	1.74E+01	0.98	1.84E+01	1.21E+01
s	1.12E+01	1.70E+01	1.66E+01	0.41	1.70E+01	1.12E+01
d	8.70E+00	1.32E+01	1.07E+01	2.51	1.32E+01	8.70E+00
e	8.40E+00	1.28E+01	1.54E+01	2.64	1.28E+01	8.40E+00
ts	-3.90E+00	-5.92E+00	-3.98E+00	1.94	-5.92E+00	-3.90E+00
u	9.00E+00	1.37E+01	9.95E+00	3.72	1.37E+01	9.00E+00
z	4.81E+00	7.30E+00	4.88E+00	2.42	7.30E+00	4.81E+00
u	1.13E+01	1.72E+01	1.53E+01	1.86	1.72E+01	1.13E+01
i	1.00E+01	1.52E+01	1.20E+01	3.19	1.52E+01	1.00E+01
t	4.11E+00	6.24E+00	4.40E+00	1.84	6.24E+00	4.11E+00
a	8.60E+00	1.31E+01	1.06E+01	2.46	1.31E+01	8.60E+00

フロントページの続き

(51)Int.Cl. F I
G 1 0 L 15/02 (2006.01) G 1 0 L 15/00 2 0 0 E
G 1 0 L 15/10 5 0 0 Z
G 1 0 L 15/02 3 0 0 A

(56)参考文献 特開 2 0 0 4 - 2 7 1 8 9 5 (J P , A)
特開 2 0 0 4 - 0 5 3 6 5 2 (J P , A)
特開平 1 1 - 1 4 3 3 4 6 (J P , A)
特開平 0 3 - 2 5 2 7 0 0 (J P , A)
特開 2 0 0 1 - 2 9 6 8 8 2 (J P , A)
特開平 0 5 - 3 4 1 7 0 5 (J P , A)

(58)調査した分野(Int.Cl. , D B 名)
G 1 0 L 1 1 / 0 0 - 2 1 / 0 6
G 0 9 B 1 5 / 0 0
G 0 9 B 1 9 / 0 6