

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第6306071号  
(P6306071)

(45) 発行日 平成30年4月4日(2018.4.4)

(24) 登録日 平成30年3月16日(2018.3.16)

(51) Int.Cl. F I  
**A 6 1 B 5/16 (2006.01)** A 6 1 B 5/16  
**G 1 0 L 25/66 (2013.01)** G 1 0 L 25/66

請求項の数 6 (全 14 頁)

(21) 出願番号	特願2016-22895 (P2016-22895)	(73) 特許権者	512286484 P S T株式会社
(22) 出願日	平成28年2月9日(2016.2.9)		神奈川県横浜市中区山下町2 産業貿易セ ンター905
(65) 公開番号	特開2017-140170 (P2017-140170A)	(73) 特許権者	504119583 光吉 俊二
(43) 公開日	平成29年8月17日(2017.8.17)		東京都港区高輪四丁目10番31号532 号室
審査請求日	平成28年9月13日(2016.9.13)	(74) 代理人	100072718 弁理士 古谷 史旺
		(74) 代理人	100116001 弁理士 森 俊秀
		(72) 発明者	篠原 修二 東京都港区高輪四丁目10番31号 53 2号室 P S T株式会社内

最終頁に続く

(54) 【発明の名称】 推定装置、推定プログラム、推定装置の作動方法および推定システム

(57) 【特許請求の範囲】

【請求項1】

被験者が発話した音声データを用いて複数の特徴量を算出して、算出した前記複数の特徴量のうち、他の前記特徴量との相関が無い、あるいは他の前記特徴量との相関係数の絶対値が所定値より小さい少なくとも1つの前記特徴量を決定する第1算出部と、

前記第1算出部が決定した前記特徴量に基づいて前記被験者における心身状態を示す度合いを算出する第2算出部と、

前記第2算出部が算出した前記度合いに基づいて前記被験者における心身状態を推定する推定部と

を備えることを特徴とする推定装置。

10

【請求項2】

請求項1に記載の推定装置において、

前記第2算出部は、決定された前記特徴量を重み付けして前記度合いを算出することを特徴とする推定装置。

【請求項3】

請求項1または請求項2に記載の推定装置において、

前記第1算出部は、算出した前記複数の特徴量のうち、他の前記特徴量との相関が無い、あるいは他の前記特徴量との相関係数の絶対値が所定値より小さく、かつROC (Receiver Operating Characteristic) 曲線におけるAUC (Area Under the Curve) を0.7以上有する少なくとも1つの前記特徴量を決定することを特徴とする推定装置。

20

## 【請求項 4】

被験者が発話した音声データを用いて複数の特徴量を算出して、算出した前記複数の特徴量のうち、他の前記特徴量との相関が無い、あるいは他の前記特徴量との相関係数の絶対値が所定値より小さい少なくとも1つの前記特徴量を決定し、

決定した前記特徴量に基づいて前記被験者における心身状態を示す度合いを算出し、算出した前記度合いに基づいて前記被験者における心身状態を推定する処理をコンピュータに実行させる推定プログラム。

## 【請求項 5】

被験者の心身状態を推定する推定装置の作動方法において、

前記被験者が発話した音声データを用いて複数の特徴量を算出して、算出した前記複数の特徴量のうち、他の前記特徴量との相関が無い、あるいは他の前記特徴量との相関係数の絶対値が所定値より小さい少なくとも1つの前記特徴量を決定し、

決定した前記特徴量に基づいて前記被験者における心身状態を示す度合いを算出し、算出した前記度合いに基づいて前記被験者における心身状態を推定する

ことを特徴とする推定装置の作動方法。

## 【請求項 6】

被験者が発話した音声データを取得する取得装置と、

前記取得装置が取得した前記音声データを用いて複数の特徴量を算出して、算出した前記複数の特徴量のうち、他の前記特徴量との相関が無い、あるいは他の前記特徴量との相関係数の絶対値が所定値より小さい少なくとも1つの前記特徴量を決定する第1算出部と、前記第1算出部が決定した特徴量に基づいて前記被験者における心身状態を示す度合いを算出する第2算出部と、前記第2算出部が算出した前記度合いに基づいて前記被験者における心身状態を推定する推定部とを含む推定装置と

を備えることを特徴とする推定システム。

## 【発明の詳細な説明】

## 【技術分野】

## 【0001】

本発明は、被験者の健康状態を推定する推定装置、推定プログラム、推定装置の作動方法および推定システムに関する。

## 【背景技術】

## 【0002】

人の脳の活動（すなわち、人の感情状態あるいは心身状態）は、人が発話する音声に現れることが知られている。例えば、人が発話した音声の信号からピッチ周波数等を算出し、ピッチ周波数等と感情状態との対応関係を示す情報と、算出したピッチ周波数等のパラメータとに基づいて、人の感情状態（あるいは心身状態）を推定する技術が提案されている（例えば、引用文献1参照）。

## 【先行技術文献】

## 【特許文献】

## 【0003】

【特許文献1】国際公開第2006/132159号

## 【発明の概要】

## 【発明が解決しようとする課題】

## 【0004】

ピッチ周波数等のパラメータと感情状態との対応関係は、様々な状況や場面において発話された音声のサンプルデータを用いて、複数の人の各々に、各サンプルデータの音声を発話した人の感情状態（例えば、喜びや怒り等）を判定させることで生成される。すなわち、ピッチ周波数等のパラメータと感情状態との対応関係の生成には、可能な限り多くのサンプルデータを予め用意し、サンプルデータ毎に発話した人の感情状態を、複数の人それぞれに判定させることから、時間が掛かってしまう。また、複数の人それぞれが人の感情状態を主観的に判定することにより、ピッチ周波数等のパラメータと感情状態との対応

10

20

30

40

50

関係に基づいて推定される感情状態あるいは心身状態は、客観性に欠けることがある。

【0005】

また、生成された対応関係には、パラメータから感情状態を推定するための様々な閾値が設定されている。しかしながら、設定される閾値は、音声のデータに含まれるノイズや、音声のデータに対するダウンサンプリング等の処理による音質の劣化に対して、影響を受けやすいという問題がある。

【0006】

一つの側面では、本件開示の推定方法、推定プログラム、推定装置および推定システムは、ピッチ周波数等のパラメータと感情状態との対応関係を示す情報を予め用意することなく、従来と比べて容易に被験者の心身状態を推定できる技術を提供することを目的とする。

10

【課題を解決するための手段】

【0007】

一つの観点による推定装置は、被験者が発話した音声データを用いて複数の特徴量を算出して、算出した複数の特徴量のうち、他の特徴量との相関が無い、あるいは他の特徴量との相関係数の絶対値が所定値より小さい少なくとも一つの特徴量を決定する第1算出部と、第1算出部が決定した特徴量に基づいて被験者における心身状態を示す度合いを算出する第2算出部と、第2算出部が算出した度合いに基づいて被験者における心身状態を推定する推定部とを有する。

【0008】

20

別の観点による推定プログラムは、被験者が発話した音声データを用いて複数の特徴量を算出して、算出した複数の特徴量のうち、他の特徴量との相関が無い、あるいは他の特徴量との相関係数の絶対値が所定値より小さい少なくとも一つの特徴量を決定し、決定した特徴量に基づいて被験者における心身状態を示す度合いを算出し、算出した度合いに基づいて被験者における心身状態を推定する処理をコンピュータに実行させる。

【0009】

別の観点による推定装置の作動方法は、被験者の心身状態を推定する推定装置の作動方法において、被験者が発話した音声データを用いて複数の特徴量を算出して、算出した複数の特徴量のうち、他の特徴量との相関が無い、あるいは他の特徴量との相関係数の絶対値が所定値より小さい少なくとも一つの特徴量を決定し、決定した特徴量に基づいて被験者における心身状態を示す度合いを算出し、算出した度合いに基づいて被験者における心身状態を推定する。

30

【0010】

別の観点による推定システムは、被験者が発話した音声データを取得する取得装置と、取得装置が取得した音声データを用いて複数の特徴量を算出して、算出した複数の特徴量のうち、他の特徴量との相関が無い、あるいは他の特徴量との相関係数の絶対値が所定値より小さい少なくとも一つの特徴量を決定する第1算出部と、第1算出部が決定した特徴量に基づいて被験者における心身状態を示す度合いを算出する第2算出部と、第2算出部が算出した度合いに基づいて被験者における心身状態を推定する推定部とを含む推定装置とを有する。

40

【発明の効果】

【0011】

本件開示の推定装置、推定プログラム、推定装置の作動方法および推定システムは、ピッチ周波数等のパラメータと感情状態との対応関係を示す情報を予め用意することなく、従来と比べて容易に被験者の心身状態を推定できる。

【図面の簡単な説明】

【0012】

【図1】推定装置の一実施形態を示す図である。

【図2】推定装置の別の実施形態を示す図である。

【図3】図2に示した携帯通信端末を介して取得した音声データの一例を示す図である。

50

【図4】図2に示した第1算出部10aが算出した特徴量間における相関性を示す図である。

【図5】図4に示した特徴量間における相関性の続きを示す図である。

【図6】図2に示した推定システムにおける推定処理の一例を示す図である。

【発明を実施するための形態】

【0013】

以下、図面を用いて実施形態について説明する。

【0014】

図1は、推定装置および推定方法の一実施形態を示す。

【0015】

図1に示した推定装置100は、CPU (Central Processing Unit) 等の演算処理装置と、ハードディスク装置等の記憶装置とを有するコンピュータ装置等である。推定装置100は、例えば、演算処理装置が記憶装置に記憶された推定プログラムを実行することにより、第1算出部10、第2算出部20および推定部30として機能する。なお、第1算出部10、第2算出部20および推定部30は、ハードウェアにより実現されてもよい。

【0016】

第1算出部10は、推定装置100の記憶装置に記憶された被験者が発話した音声データ、あるいはスマートフォン等の携帯通信端末を介して取得した被験者の音声データを用いて、音声のピッチ周波数や音声の強度等を算出する。そして、第1算出部10は、ピッチ周波数や音声の強度等の時間変化に基づいて、被験者の発話におけるピッチ周波数の検出頻度、および音声の強度の増加（または減少率）等を算出する。ピッチ周波数の検出頻度および音声の強度の増加率等は、音声の特徴量の一例である。

【0017】

なお、第1算出部10は、ピッチ周波数の検出頻度および音声強度の増加率（または減少率）の少なくとも1つを算出してもよい。

【0018】

第2算出部20は、算出したピッチ周波数の検出頻度等の特徴量に基づいて、被験者における心身状態を示す度合いを算出する。例えば、第2算出部20は、算出したピッチ周波数の検出頻度および音声強度の増加率等を加算した値を、被験者の心身状態を示す度合い（以下、元気度とも称される）として算出する。なお、第2算出部20は、ピッチ周波数の検出頻度あるいは強度の増加率等の少なくとも1つを、被験者の元気度としてもよい。また、第2算出部20は、ピッチ周波数の検出頻度あるいは強度の増加率等を、重み付け加算した値を被験者の元気度としてもよい。

【0019】

推定部30は、算出した元気度に基づいて被験者における心身状態（例えば、うつ状態等か否か）を推定する。そして、推定装置100は、推定部30により推定された心身状態を示す情報を、外部の有機EL (Organic Electro-Luminescence) や液晶等のディスプレイに出力する。

【0020】

なお、推定装置100の構成および動作は、図1に示した例に限定されない。例えば、推定装置100と、スマートフォン等の携帯通信端末と、有機EL等のディスプレイとを含めることで、推定システムとしてもよい。

【0021】

以上、図1に示した実施形態では、推定装置100は、被験者の音声データを用いて、被験者の発話におけるピッチ周波数の検出頻度、および音声の強度の増加率等の特徴量を算出し、算出した特徴量に基づいて被験者の心身状態を示す元気度を算出する。そして、推定装置100は、算出した元気度に基づいて被験者の心身状態を推定する。これにより、推定装置100は、ピッチ周波数等のパラメータと感情状態との対応関係を示す情報を予め用意することなく、従来と比べて容易に被験者の心身状態を推定できる。また、元気

10

20

30

40

50

度は算出した特徴量に基づいて算出されることから、推定装置 100 は、被験者の心身状態を客観的に評価することができる。また、推定装置 100 は、ピッチ周波数等のパラメータと感情状態との対応関係を示す情報を用いないため、ノイズ等の音質劣化に対する耐性を有する。

【0022】

図 2 は、推定方法、推定装置および推定システムの別の実施形態を示す。

【0023】

図 2 に示した推定システム S Y S は、推定装置 100 a および携帯通信端末 200 を有する。推定装置 100 a と携帯通信端末 200 とは、有線または無線を介して接続される。なお、推定装置 100 a と携帯通信端末 200 とは、ネットワークを介して接続されてもよい。

10

【0024】

携帯通信端末 200 は、携帯電話、スマートフォンあるいはタブレット型端末である。携帯通信端末 200 は、携帯通信端末 200 に含まれるマイクロホンを通じて被験者 P A が発話する音声の信号を取得し、取得した信号を所定のサンプリング周波数（例えば、11 キロヘルツ等）でサンプリングすることでデジタル信号の音声データを生成する。そして、携帯通信端末 200 は、生成した音声データを推定装置 100 a に送信する。また、携帯通信端末 200 は、推定装置 100 a により推定された結果を、携帯通信端末 200 に含まれる有機 E L 等のディスプレイに表示する。携帯通信端末 200 は、取得装置の一例である。

20

【0025】

なお、携帯通信端末 200 の代わりに、マイクロホンが、有線または無線を介して、推定装置 100 a に接続されてもよい。この場合、推定装置 100 a は、マイクロホンからの音声の信号を、所定のサンプリング周波数でサンプリングし、デジタル信号の音声データを生成してもよい。

【0026】

推定装置 100 a は、C P U 等の演算処理装置と、ハードディスク装置等の記憶装置とを有するコンピュータ装置等である。推定装置 100 a は、例えば、演算処理装置が記憶装置に記憶された推定プログラムを実行することにより、第 1 算出部 10 a、第 2 算出部 20 a および推定部 30 a として機能する。なお、第 1 算出部 10 a、第 2 算出部 20 a および推定部 30 a は、ハードウェアにより実現されてもよい。

30

【0027】

なお、推定装置 100 a は、携帯通信端末 200 において実現されてもよい。すなわち、携帯通信端末 200 に含まれる C P U が、携帯通信端末 200 に含まれるメモリ等の記憶部に記憶された推定プログラムを実行することにより、携帯通信端末 200 は、第 1 算出部 10 a、第 2 算出部 20 a および推定部 30 a として機能してもよい。

【0028】

第 1 算出部 10 a は、携帯通信端末 200 から取得した被験者 P A の音声データを用いて、音声のピッチ周波数や音声の強度等を算出する。そして、第 1 算出部 10 a は、算出したピッチ周波数や強度等の時間変化に基づいて、被験者 P A の発話におけるピッチ周波数の検出頻度、および音声強度の増加率等の音声の特徴量を算出する。なお、第 1 算出部 10 a の動作については、図 3 で説明する。

40

【0029】

第 2 算出部 20 a は、第 1 算出部 10 a が算出したピッチ周波数の検出頻度等の特徴量に基づいて、被験者 P A における心身状態を示す度合い（元気度）を算出する。第 2 算出部 20 a の動作については、図 4 で説明する。

【0030】

推定部 30 a は、第 2 算出部 20 a が算出した度合いに基づいて被験者 P A における心身状態を推定する。そして、推定装置 100 a は、推定部 30 a により推定された心身状態を示す情報を携帯通信端末 200 へ出力し、携帯通信端末 200 に含まれる有機 E L 等

50

のディスプレイに表示させる。

【 0 0 3 1 】

図 3 は、図 2 に示した携帯通信端末 2 0 0 を介して取得した音声データの一例を示す。図 3 ( a ) は、携帯通信端末 2 0 0 を介して取得した被験者 P A が発話した音声の音圧の時間変化を示し、図 3 ( b ) は、被験者 P A が発話した音声の強度の時間変化を示す。なお、図 3 の横軸は、時刻を示し、図 3 ( a ) の縦軸は、音声の音圧を示し、図 3 ( b ) の縦軸は、音声の強度 P W を示す。音声の強度 P W は、音圧の二乗のことである。

【 0 0 3 2 】

図 3 は、被験者 P A による発話の音声データのうち、“ありがとう”と発話した発話単位のデータを示す。時刻  $t_0$ 、 $t_1$ 、 $t_2$ 、 $t_3$ 、 $t_4$  は、発話単位に含まれる“あ”、“り”、“が”、“と”、“う”の各語の発話の開始時刻を示す。なお、“ありがとう”の発話単位のうち、“り”の語を発話した音声データに対する第 1 算出部 1 0 a の算出処理について説明するが、第 1 算出部 1 0 a は、“ありがとう”の他の語および他の発話単位に対しても、同一または同様に、算出処理を実行する。

10

【 0 0 3 3 】

第 1 算出部 1 0 a は、例えば、携帯通信端末 2 0 0 から取得した音声データを用いて、ピッチ周波数、強度およびゼロ点交差数等を、ウィンドウ W D 毎に算出する。例えば、第 1 算出部 1 0 a は、取得した音声データを用いて、“ありがとう”等の発話単位のデータ毎に、ウィンドウ W D の幅で F F T (Fast Fourier Transform) 等のスペクトラム解析を実行し、パワースペクトルを算出する。第 1 算出部 1 0 a は、算出される音声の強度 P W の値が大きくなばらつきを示す場合があるため、音声の強度 P W をウィンドウ W D 毎に平均をとって算出する。すなわち、図 3 ( b ) は、移動平均した音声の強度 P W の時間変化を示す。

20

【 0 0 3 4 】

また、第 1 算出部 1 0 a は、算出したパワースペクトルに対して自己相関の処理を実行し、算出した自己相関係数の分布において隣接する極大値(または極小値)間の周波数の間隔に基づいて、ピッチ周波数を算出する。さらに、第 1 算出部 1 0 a は、各ウィンドウ W D の音声データにおいて、音声の波形の音圧が基準圧力(例えば、“0”とする)を横切る回数を、ゼロ点交差数として算出する。

【 0 0 3 5 】

なお、ウィンドウ W D の幅は、例えば、5 1 2 等のサンプル数を有し、第 1 算出部 1 0 a は、ウィンドウ W D の幅の 4 分の 1 等の所定の間隔でウィンドウ W D を移動させて、ピッチ周波数等を、各ウィンドウ W D で算出する。すなわち、図 3 ( b ) に示した時刻  $t_1$ 、 $t_{11}$ 、 $t_{12}$ 、 $t_{13}$ 、 $t_{14}$ 、 $t_{15}$ 、 $t_{16}$ 、 $t_{17}$ 、 $t_{18}$ 、 $t_{19}$ 、 $t_2$  における強度 P W は、“り”の語が発話された音声データを用いて、第 1 算出部 1 0 a が算出した強度を示す。そして、時刻  $t_1$ 、 $t_{11}$ 、 $t_{12}$ 、 $t_{13}$ 、 $t_{14}$ 、 $t_{15}$ 、 $t_{16}$ 、 $t_{17}$ 、 $t_{18}$ 、 $t_{19}$ 、 $t_2$  の各々の時間間隔は、ウィンドウ W D を移動させた所定の間隔と等しい。

30

【 0 0 3 6 】

また、第 1 算出部 1 0 a は、ピッチ周波数 F 0 やテンポ等のパラメータを、音声データから算出してもよい。

40

【 0 0 3 7 】

次に、第 1 算出部 1 0 a は、“ありがとう”等の発話単位毎に算出したピッチ周波数、強度、ゼロ点交差数等に基づいて、被験者 P A の発話におけるピッチ周波数の検出頻度、および強度 P W の増加率等の特徴量を算出する。例えば、第 1 算出部 1 0 a は、発話単位の各々において、全てのウィンドウ W D のうちピッチ周波数が算出されたウィンドウ W D の割合を、ピッチ周波数の検出率を示す P I T C H \_ R A T E として算出する。すなわち、P I T C H \_ R A T E は、被験者 P A の音声において有声音(母音)が発話された割合を示す。第 1 算出部 1 0 a は、各発話単位の P I T C H \_ R A T E を、各発話単位の開始時刻(例えば、“ありがとう”の場合の時刻  $t_0$ )や終了時刻(例えば、“ありがとう”

50

の場合の時刻  $t_5$  ) 等の時刻に対応付ける。そして、第 1 算出部 10 a は、被験者 P A の発話における PITCH\_RATE の時間変化を取得する。

【0038】

また、第 1 算出部 10 a は、ウィンドウ W D 毎に算出したゼロ点交差数を用いて、隣接するウィンドウ W D のゼロ点交差数との差分であるゼロ点交差数の各発話単位における、ばらつきの度合いを示す DELTA\_ZERO\_DIV を算出する。例えば、第 1 算出部 10 a は、互いに隣接するウィンドウ W D 間でゼロ点交差数の差分をそれぞれ求め、求めたゼロ点交差数の差分の標準偏差を DELTA\_ZERO\_DIV として算出する。なお、第 1 算出部 10 a は、各ウィンドウ W D 間で求めたゼロ点交差数の差分の分散値を、DELTA\_ZERO\_DIV として算出してもよい。あるいは、第 1 算出部 10 a は、各ウィンドウ W D 間で求めたゼロ点交差数の差分の平均値と、各ウィンドウ W D 間のゼロ点交差数の差分との差の絶対値を加算した値を、DELTA\_ZERO\_DIV として算出してもよい。第 1 算出部 10 a は、各発話単位の DELTA\_ZERO\_DIV を、各発話単位の開始時刻や終了時刻等の時刻に対応付けて、被験者 P A の発話における DELTA\_ZERO\_DIV の時間変化を取得する。

10

【0039】

また、図 3 ( b ) に示すように、例えば、“ありがとう”の発話単位において、“あ”、“り”、“が”、“と”、“う”の各語の強度 P W の時間変化は、強度が増加する Attack の領域と、強度が一定に保たれる Keep の領域と、強度が減少する Decay の領域とを有する。そこで、第 1 算出部 10 a は、Attack の領域および Decay の領域における強度 P W の傾きを算出する。例えば、第 1 算出部 10 a は、“り”の語で算出した強度 P W のうち、Attack の領域に含まれる時刻  $t_1$  の強度 P W ( $t_1$ ) と、時刻  $t_{12}$  の強度 P W ( $t_{12}$ ) と、式 ( 1 ) とを用いて、Attack の領域における強度 P W の傾き A S を算出する。

20

$$A S = ( P W ( t_{12} ) - P W ( t_1 ) ) / ( t_{12} - t_1 ) \dots ( 1 )$$

そして、第 1 算出部 10 a は、“ありがとう”の各語で算出した傾き A S の平均値を、“ありがとう”の発話単位の Attack の領域における強度 P W の傾き ATTACK\_SLOPE として算出する。そして、第 1 算出部 10 a は、各発話単位の ATTACK\_SLOPE を、各発話単位の開始時刻や終了時刻等の時刻に対応付けて、被験者 P A の発話における ATTACK\_SLOPE の時間変化を取得する。

30

【0040】

なお、第 1 算出部 10 a は、時刻  $t_1$  の強度 P W ( $t_1$ )、時刻  $t_{11}$  の強度 P W ( $t_{11}$ ) および時刻  $t_{12}$  の強度 P W ( $t_{12}$ ) を用いて、“り”の語の Attack の領域における強度 P W の傾き A S を算出してもよい。例えば、第 1 算出部 10 a は、時刻  $t_1$  の強度 P W ( $t_1$ ) と時刻  $t_{11}$  の強度 P W ( $t_{11}$ ) との間の強度の傾きと、時刻  $t_{11}$  の強度 P W ( $t_{11}$ ) と時刻  $t_{12}$  の強度 P W ( $t_{12}$ ) との間の強度の傾きとをそれぞれ算出する。そして、第 1 算出部 10 a は、算出した各強度の傾きの平均値を、“り”の語の Attack の領域における強度 P W の傾き A S として算出する。

【0041】

一方、例えば、第 1 算出部 10 a は、“り”の語で算出した強度 P W のうち、Decay の領域に含まれる時刻  $t_{18}$  の強度 P W ( $t_{18}$ ) と、時刻  $t_2$  の強度 P W ( $t_2$ ) と、式 ( 2 ) とを用いて、Decay の領域における強度 P W の傾き D S を算出する。

40

$$D S = ( P W ( t_2 ) - P W ( t_{18} ) ) / ( t_2 - t_{18} ) \dots ( 2 )$$

そして、第 1 算出部 10 a は、“ありがとう”の各語で算出した傾き D S の平均値を、“ありがとう”の発話単位の Decay の領域における強度 P W の傾き DECAY\_SLOPE として算出する。そして、第 1 算出部 10 a は、各発話単位の DECAY\_SLOPE を、各発話単位の開始時刻や終了時刻等の時刻に対応付けて、被験者 P A の発話における DECAY\_SLOPE の時間変化を取得する。

【0042】

なお、第 1 算出部 10 a は、時刻  $t_{18}$  の強度 P W ( $t_{18}$ )、時刻  $t_{19}$  の強度 P W

50

( t 1 9 ) および時刻 t 2 の強度 P W ( t 2 ) を用いて、“ り ” の語の D e c a y の領域における強度 P W の傾き D S を算出してもよい。例えば、第 1 算出部 1 0 a は、時刻 t 1 8 の強度 P W ( t 1 8 ) と時刻 t 1 9 の強度 P W ( t 1 9 ) との間の強度の傾きと、時刻 t 1 9 の強度 P W ( t 1 9 ) と時刻 t 2 の強度 P W ( t 2 ) との間の強度の傾きとをそれぞれ算出する。そして、第 1 算出部 1 0 a は、算出した各強度の傾きの平均値を、“ り ” の語の D e c a y の領域における強度 P W の傾き D S として算出する。

【 0 0 4 3 】

また、第 1 算出部 1 0 a は、図 3 ( b ) に示すように、発話単位 “ ありがとう ” の各語の発話の終了時刻 t 1、t 2、t 3、t 4、t 5 における強度 P W の平均値を、D E C A Y \_ P O W E R として算出する。第 1 算出部 1 0 a は、各発話単位の D E C A Y \_ P O W E R を、各発話単位の開始時刻や終了時刻等の時刻に対応付けて、被験者 P A の発話における D E C A Y \_ P O W E R の時間変化を取得する。

10

【 0 0 4 4 】

なお、第 1 算出部 1 0 a は、D E C A Y \_ S L O P E \_ D I V、D E L T A \_ Z E R O \_ M A X \_ A B S、D E L T A \_ Z E R O \_ D I V \_ A B S、D E C A Y \_ C O U N T、および P O W E R \_ P E A K \_ C O U N T 等の特徴量を算出してもよい。また、第 1 算出部 1 0 a は、D E C A Y \_ P O W E R \_ D I V、A T T A C K \_ S L O P E \_ D I V、A T T A C K \_ C O U N T、および P I T C H \_ T I M E \_ C O R R E 等の特徴量を算出してもよい。

【 0 0 4 5 】

20

なお、D E C A Y \_ S L O P E \_ D I V は、各発話単位における D E C A Y \_ S L O P E のばらつきの度合いを示す標準偏差や分散値等である。D E L T A \_ Z E R O \_ M A X \_ A B S は、各発話単位において算出された複数の D E L T A \_ Z E R O \_ D I V のうち、所定の強度以上の強度 P W を有する音声データを用いて算出された最大の D E L T A \_ Z E R O \_ D I V の絶対値である。D E L T A \_ Z E R O \_ D I V \_ A B S は、D E L T A \_ Z E R O \_ D I V の絶対値である。D E C A Y \_ C O U N T は、図 3 ( b ) に示すように、各発話単位の強度 P W において、D e c a y の領域でサンプリングされたデータ数である。

【 0 0 4 6 】

また、P O W E R \_ P E A K \_ C O U N T は、各発話単位において、例えば、互いに隣接する 3 つのウィンドウ W D において算出された強度 P W の時間変化が凸形状となる、1 秒等の単位時間当たりの数である。なお、P O W E R \_ P E A K \_ C O U N T を算出するにあたり、隣接するウィンドウ W D の数は、3 以上の複数でもよい。また、各ウィンドウ W D の強度 P W は、ノイズレベル以上であることが好ましい。

30

【 0 0 4 7 】

また、D E C A Y \_ P O W E R \_ D I V は、例えば、各発話単位における D E C A Y \_ P O W E R のばらつきの度合いを示す標準偏差や分散値等である。A T T A C K \_ S L O P E \_ D I V は、各発話単位における A T T A C K \_ S L O P E のばらつきを示す標準偏差や分散値等である。A T T A C K \_ C O U N T は、図 3 ( b ) に示すように、各発話単位の強度 P W において、A t t a c k の領域でサンプリングされたデータ数である。P I T C H \_ T I M E \_ C O R R E は、各発話単位において、ウィンドウ W D を番号付けした場合に、ウィンドウ W D の順番 ( すなわち時間経過 ) とピッチ周波数の時間変化との間の相関係数である。

40

【 0 0 4 8 】

図 4 および図 5 は、図 2 に示した第 1 算出部 1 0 a が算出した特徴量間における相関性を示す。特徴量間の相関性を示す相関テーブル C T は、L A B E L、A U C ( Area Under the Curve ) および D E C A Y \_ P O W E R 等の複数の特徴量のそれぞれを記憶する領域を有する。L A B E L の領域には、D E C A Y \_ P O W E R 等の特徴量を示す名称が格納される。

【 0 0 4 9 】

50

AUCの領域には、例えば、医師によってラベル（例えば、うつ状態か否か、脳梗塞か否か等）を付与された複数の被験者音声データを、LABEL領域の各特徴量を用いて分類した場合のROC曲線に対するAUCが格納される。すなわち、格納されるAUCの値は、各特徴量有する被験者PAの心身状態を判定する能力の度合いを示す。なお、ROCは、Receiver Operating Characteristicの略である。

【0050】

すなわち、例えば、AUCの値が0.7以上の特徴量は、単独で用いられた場合でも、被験者PAの心身状態を正しく判定でき、AUCの値が0.7より低い特徴量は、単独で用いられた場合、被験者PAの心身状態を正しく判定できないことを示す。関連テーブルCTでは、AUCの値を0.7以上有する特徴量を示す。

10

【0051】

複数の特徴量の領域（以下、特徴量領域とも称される）の各々には、被験者PAの音声データを用いて算出された、各特徴量領域の特徴量が示す時間変化と、LABELの領域の各特徴量が示す時間変化との間の相互相関係数が格納される。なお、図4および図5に示した関連テーブルCTでは、相互相関係数の絶対値が所定の係数值、例えば、0.65以上を示す特徴量領域を網掛けで示している。これは、相互相関係数の絶対値が所定の係数值以上を示す、特徴量領域の特徴量とLABELの領域の特徴量とは、推定装置100aが、被験者PAの音声データを用いて、いずれか一方の特徴量を算出した場合、他方の特徴量を推定できることを示す。すなわち、推定装置100aは、図4および図5に示した関連テーブルCTのLABELの領域の特徴量のうち、一部の特徴量を算出することで、LABELの領域の全ての特徴量を算出するのと同等であることを示す。

20

【0052】

そこで、推定装置100aは、図4および図5に示した関連テーブルCTに基づいて、LABELの領域の特徴量のうち、AUCの値が0.7以上と大きく、かつ他の特徴量との相関性が無い、あるいは他の特徴量との相関性が所定の係数值より小さい特徴量を選択する。例えば、推定装置100aは、LABELの領域において網掛けで示すDECAY\_\_POWER、DECAY\_\_SLOPE、PITCH\_\_RATE、およびDELTA\_\_ZERO\_\_DIVの4つの特徴量を選択する。

なお、推定装置100aは、主成分分析法、あるいは自己符号化器等のニューラルネットを用いて、AUCの値が0.7以上と大きく、かつ他の特徴量との関係性が無い、あるいは関係性が低い特徴量を選択してもよい。

30

【0053】

第1算出部10aは、被験者PAの音声データを用いて、選択されたDECAY\_\_POWER、DECAY\_\_SLOPE、PITCH\_\_RATE、およびDELTA\_\_ZERO\_\_DIVの特徴量を算出する。第2算出部20aは、式(1)を用いて、算出されたDECAY\_\_POWER、DECAY\_\_SLOPE、PITCH\_\_RATE、およびDELTA\_\_ZERO\_\_DIVの特徴量を重み付け加算して、被験者PAの心身状態を示す度合い（元気度）を算出する。

$$= - \text{DECAY\_POWER} + \text{DECAY\_SLOPE} + \text{PITCH\_RATE} + 0.5 \times \text{DELTA\_ZERO\_DIV} \dots (1)$$

40

なお、DECAY\_\_POWER、DECAY\_\_SLOPE、PITCH\_\_RATE、およびDELTA\_\_ZERO\_\_DIVの特徴量の重み付け係数は、式(1)の場合に限定されない。例えば、各特徴量を重み付け係数は、対象となる被験者PAの職業、家族構成あるいは生活環境等に応じて適宜設定されることが好ましい。例えば、第1算出部10aは、式(2)を用いて、被験者PAの元気度を算出してもよい。なお、式(2)におけるDECAY\_\_SLOPEの係数は、“0”である。

$$= - 0.5 \times \text{DECAY\_POWER} + \text{PITCH\_RATE} + 0.5 \times \text{DELTA\_ZERO\_DIV} \dots (2)$$

なお、DECAY\_\_POWER、DECAY\_\_SLOPE、PITCH\_\_RATE、およびDELTA\_\_ZERO\_\_DIVの特徴量の各々は、所定の係数值、例えば、0.65

50

以上の相互相関係数を示す特徴量に代えられてもよい。例えば、DECAY\_SLOPEは、ATTACK\_SLOPEとの間で相互相関係数が0.79であることから、第1算出部10aは、DECAY\_SLOPEの代わりに、ATTACK\_SLOPEを算出してもよい。そして、第2算出部20aは、DECAY\_POWER、PITCH\_RATE、DELTA\_ZERO\_DIV、式(1)とともに、ATTACK\_SLOPEを用いて、元気度を算出する。なお、ATTACK\_SLOPEの重み付け係数は、適宜設定されることが好ましい。

【0054】

また、元気度は、AUCの値が高いDECAY\_POWER、DECAY\_SLOPE、PITCH\_RATE、およびDELTA\_ZERO\_DIV等の特徴量のいずれか1つを用いて算出されてもよい。

10

【0055】

なお、DECAY\_POWER\_DIVおよびPITCH\_TIME\_CORREは、他の特徴量との相関性が低い、他の特徴量と比べてAUCの値が最も小さいことから、元気度を算出する式(1)には含まれていない。しかしながら、DECAY\_POWER\_DIVおよびPITCH\_TIME\_CORREは、第1算出部10aにより算出され、式(1)に含まれてもよい。

【0056】

推定部30aは、第2算出部20aが算出した元気度と閾値との比較に基づいて、被験者PAにおける心身状態、例えば、被験者PAがうつ状態か否かを推定する。例えば、推定部30aは、元気度が閾値より小さい(すなわち、被験者PAが発話する音声不明瞭で、歯切れが悪い)場合、被験者PAの心身状態が良くない(例えば、うつ状態)と推定する。一方、推定部30aは、元気度が閾値以上(すなわち、被験者PAが発話する音声明瞭で、歯切れが良い)場合、被験者PAの心身状態が良く、健康であると推定する。

20

【0057】

なお、推定部30aが用いる閾値は、例えば、敏感度が“1”で、偽陽性率(1-特異度)が“0”の点からの距離が最小となる元気度のROC曲線上の点に基づいて設定される。あるいは、閾値は、元気度のROC曲線と、AUCが0.5の場合のROC曲線との間の距離(敏感度+特異点-1)を示すYouden Index等に基づいて設定されてもよい。

30

【0058】

図6は、図2に示した推定装置100aによる推定処理の一例を示す。ステップS100からステップS130は、推定装置100aに搭載される演算処理装置が、推定装置100aの記憶装置に記憶される推定プログラムを実行することにより実現される。すなわち、図6は、推定プログラムおよび推定方法の別の実施形態を示す。この場合、図2に示した第1算出部10a、第2算出部20aおよび推定部30aは、推定プログラムの実行により実現される。なお、図6に示した処理は、推定装置100aに搭載されるハードウェアにより実現されてもよい。この場合、図2に示した第1算出部10a、第2算出部20aおよび推定部30aは、推定装置100a内に配置される回路により実現される。

40

【0059】

なお、推定プログラムは、例えば、DVD(Digital Versatile Disc)等のリムーバブルディスクに記録して頒布することができる。また、推定プログラムは、USB(Universal Serial Bus)メモリ等の可搬型記憶媒体に記録して頒布してもよい。あるいは、推定装置100aは、推定装置100aに含まれるネットワークインタフェースを介して、ネットワークを通じて推定プログラムをダウンロードし、メモリ等の記憶部に格納してもよい。

【0060】

ステップS100では、第1算出部10aは、携帯通信端末200を介して取得した被験者PAが発話した音声データを用いて、ピッチ周波数、強度およびゼロ点交差数等のパ

50

ラメータを、ウィンドウWD毎に算出する。

【0061】

ステップS110では、第1算出部10aは、ステップS100で算出したピッチ周波数、強度およびゼロ点交差数等のパラメータを用いて、DECAY\_\_POWER、DECAY\_\_SLOPE、PITCH\_\_RATE、およびDELTA\_\_ZERO\_\_DIVの特徴量を算出する。

【0062】

ステップS120では、第2算出部20aは、ステップS110で算出した特徴量と、式(1)とを用いて、被験者PAの元気度 を算出する。

【0063】

ステップS130では、推定部30aは、ステップS120で算出された元気度 と、閾との比較に基づいて、被験者PAの心身状態(例えば、うつ状態か否か)を推定する。

【0064】

推定装置100aは、推定部30aにより推定された心身状態を示す情報を携帯通信端末200に出力し、携帯通信端末200のディスプレイに表示させる。そして、推定装置100aは、推定処理を終了する。図6に示した処理は、被験者PAが携帯通信端末200に向かって発話する度に繰り返し実行される。

【0065】

以上、図2から図6に示した実施形態では、推定装置100aは、被験者PAの音声データを用いて、被験者の発話におけるDECAY\_\_POWER、DECAY\_\_SLOPE、PITCH\_\_RATE、およびDELTA\_\_ZERO\_\_DIVの特徴量を算出する。そして、推定装置100aは、算出した特徴量と式(1)とを用いて、被験者PAの心身状態を示す元気度 を算出する。そして、推定装置100aは、算出した元気度 と閾値との比較に基づいて被験者の心身状態を推定する。これにより、推定装置100aは、ピッチ周波数等のパラメータと感情状態との対応関係を示す情報を予め用意することなく、従来と比べて容易に被験者の心身状態を推定できる。また、元気度 は算出した特徴量に基づいて算出されることから、推定装置100aは、被験者PAの心身状態を客観的に評価することができる。また、推定装置100aは、ピッチ周波数等のパラメータと感情状態との対応関係を示す情報を用いないため、ノイズ等の音質劣化に対する耐性を有する。

【0066】

なお、推定装置100(100a)は、精神分析、行動予測、行動分析等の心理カウンセリング、精神医療、一般医療における面接や処方へ適用した場合を示したが、これに限定されない。例えば、推定装置100は、ロボット、人工知能や自動車、あるいはコールセンター、エンターテイメント、インターネット、スマートフォンやタブレット型端末等の携帯端末装置アプリケーションやサービス、検索システムへ応用されてもよい。また、推定装置100は、診断装置、自動問診装置、災害トリアージ等に応用されてもよい。また、推定装置100は、金融与信管理システムや行動予測、企業、学校、行政機関、警察や軍事、情報収集活動等での情報分析、虚偽発見に繋がる心理分析、組織グループ管理へ応用されてもよい。また、推定装置100は、組織の構成員、研究者や従業員、管理者等の心の健康や行動予測を管理するシステム、住居やオフィス、飛行機や宇宙船といった環境を制御するシステム、あるいは家族や友人の心の状態や行動予測を知るための手段に適用されてもよい。また、推定装置100は、音楽や映画配信、一般的な情報検索、情報分析管理や情報処理、あるいは顧客感性嗜好マーケット分析等やこれらをネットワークやスタンドアロンで管理するシステム等へ適用されてもよい。

【0067】

以上の詳細な説明により、実施形態の特徴点および利点は明らかになるであろう。これは、特許請求の範囲がその精神および権利範囲を逸脱しない範囲で前述のような実施形態の特徴点および利点にまで及ぶことを意図するものである。また、当該技術分野において通常の知識を有する者であれば、あらゆる改良および変更に容易に想到できるはずである。したがって、発明性を有する実施形態の範囲を前述したものに限定する意図はなく、実

10

20

30

40

50

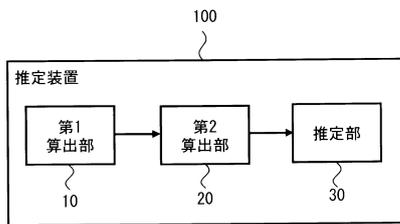
施形態に開示された範囲に含まれる適当な改良物および均等物に拠ることも可能である。

【符号の説明】

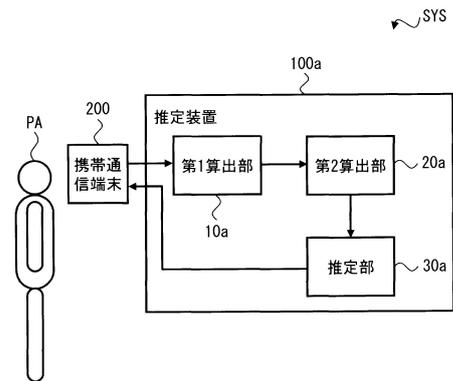
【0068】

10, 10a...第1算出部; 20, 20a...第2算出部; 30, 30a...推定部; 100, 100a...推定装置; 200...携帯通信端末; CT...相関テーブル; SYS...推定システム

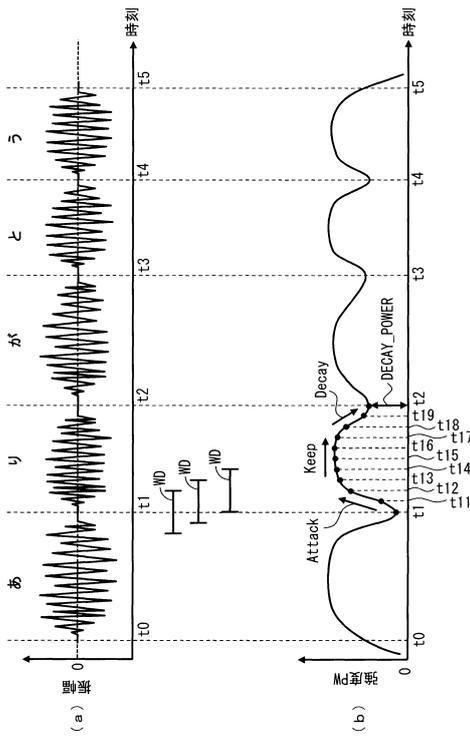
【図1】



【図2】



【図3】



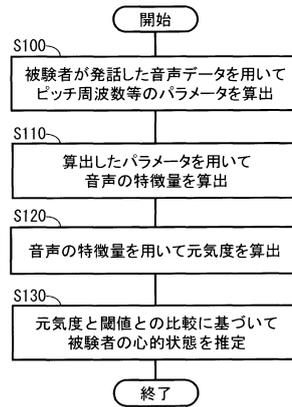
【図4】

	AUC	DECAY_POWER	DECAY_SLOPE_DIV	DECAY_SLOPE	DELTA_ZERO_ABS	DELTA_ZERO_DIV_ABS	PITCH_RATE	DELTA_ZERO_DIV
DECAY_POWER	0.95							
DECAY_SLOPE_DIV	0.95	-0.22						
DECAY_SLOPE	0.93	-0.19	0.69					
DELTA_ZERO_MAX_ABS	0.91	-0.30	-0.20	-0.19				
DELTA_ZERO_DIV_ABS	0.91	-0.35	-0.10	-0.08	0.74			
PITCH_RATE	0.89	-0.21	0.04	0.11	0.16	0.17		
DELTA_ZERO_DIV	0.89	-0.34	-0.12	-0.09	0.73	0.99	0.13	
DELTA_ZERO	0.89	0.19	-0.72	-0.58	0.11	0.04	0.02	0.04
POWER_PEAK_COUNT	0.88	-0.24	0.72	0.62	-0.16	-0.09	-0.05	-0.05
DECAY_POWER_DIV	0.84	0.51	-0.38	-0.33	-0.09	-0.15	-0.05	-0.14
ATTACK_SLOPE_DIV	0.80	-0.22	0.91	0.65	-0.21	-0.10	-0.02	-0.12
ATTACK_SLOPE	0.79	-0.16	0.67	0.79	-0.23	-0.06	0.03	-0.07
ATTACK_COUNT	0.77	0.24	-0.49	-0.49	0.08	-0.01	0.10	-0.01
PITCH_TIME_CORRE	0.71	0.15	0.00	0.00	-0.12	-0.11	-0.20	-0.11

【図5】

	DECAY_COUNT	POWER_PEAK_COUNT	DECAY_POWER_DIV	ATTACK_SLOPE_DIV	ATTACK_SLOPE	ATTACK_COUNT	PITCH_TIME_CORRE
DECAY_POWER							
DECAY_SLOPE_DIV							
DECAY_SLOPE							
DELTA_ZERO_MAX_ABS							
DELTA_ZERO_DIV_ABS							
PITCH_RATE							
DELTA_ZERO_DIV							
DECAY_COUNT	-0.83						
POWER_PEAK_COUNT	0.33	-0.38					
DECAY_POWER_DIV	-0.67	0.75	-0.37				
ATTACK_SLOPE_DIV	-0.49	0.63	-0.27	0.71			
ATTACK_SLOPE	0.68	-0.61	0.34	-0.73	-0.58		
ATTACK_COUNT	-0.05	-0.04	0.03	0.02	-0.04	0.07	
PITCH_TIME_CORRE							

【図6】



---

フロントページの続き

(72)発明者 光吉 俊二

東京都港区高輪四丁目10番31号 532号室

審査官 関根 裕

(56)参考文献 特開2013-072979(JP,A)

特開2007-000280(JP,A)

特開2011-045520(JP,A)

国際公開第2006/132159(WO,A1)

特開2003-228391(JP,A)

特開2012-000049(JP,A)

特開2015-128579(JP,A)

特開2008-76905(JP,A)

第23回 診断精度の指標とROC, 2017年 7月31日, URL, <http://web.archive.org/web/20100402220305/http://www.asca-co.com/nuclear/2010/03/post-25.html>

TOKUNO, S, Stress Evaluation by Voice: From Prevention to Treatment in Mental Health Care, (ESMSJ) Econophysics, Sociophysics & other Multidisciplinary Sciences Journal, 2015年, Vol.5, No.1, p.30-35, URL, [http://www.esmsj.upit.ro/ESMSJ%20vol%205\(1\)%20pentru%20Denis%20pe%20site/ARTICOL%203.pdf](http://www.esmsj.upit.ro/ESMSJ%20vol%205(1)%20pentru%20Denis%20pe%20site/ARTICOL%203.pdf)

(58)調査した分野(Int.Cl., DB名)

A61B 5/16 - 5/18

G10L 25/63 - 25/66

G10L 15/00 - 17/26

A61B 5/00, A61B 10/00