

(19) 日本国特許庁(JP)

(12) 特許公報(B1)

(11) 特許番号

特許第6961906号
(P6961906)

(45) 発行日 令和3年11月5日(2021.11.5)

(24) 登録日 令和3年10月18日(2021.10.18)

(51) Int. Cl. F I
G 1 O L 15/10 (2006.01) G 1 O L 15/10 5 0 0 Z

請求項の数 5 (全 15 頁)

<p>(21) 出願番号 特願2021-67998 (P2021-67998)</p> <p>(22) 出願日 令和3年2月24日(2021.2.24)</p> <p>審査請求日 令和3年2月24日(2021.2.24)</p> <p>早期審査対象出願</p>	<p>(73) 特許権者 521158211 山口 真二郎 東京都新宿区中井1-2-4-1303</p> <p>(72) 発明者 山口 真二郎 東京都新宿区中井1-2-4-1303</p> <p>審査官 菊池 智紀</p>
--	---

最終頁に続く

(54) 【発明の名称】 外国人の国籍推定システム、外国人の母国語推定システム、外国人の国籍推定方法、外国人の母国語推定方法、及びプログラム

(57) 【特許請求の範囲】

【請求項1】

特定言語を母国語としない外国人の音声から外国人の国籍を推定する、外国人の国籍推定システムであって、

複数の言語発音データであって、前記特定言語の単位言語からなる単語情報と、外国人が前記単位言語を発音した音声からなる音声情報と、外国人の国籍からなる国籍情報と、を紐づけてなる、複数の言語発音データを記憶する、記憶手段と、

前記記憶手段内の前記複数の言語発音データを学習用データとする機械学習によって、特定の前記単語情報における前記音声情報と前記国籍情報を関連付けた、学習済みモデルを取得する、学習手段と、

外国人が前記特定言語を発音した音声からなる、生言語発音データを取得する、取得手段と、

前記取得手段が取得した前記生言語発音データを単位言語に分解した、複数の分解言語発音データを生成する、分解手段と、

前記複数の分解言語発音データに、前記学習済みモデルを適用することで、前記生言語発音データに係る外国人の国籍を推定する、推定手段と、を備え、

前記推定手段が、1の前記生言語発音データにおける複数の前記分解言語発音データの国籍を個別に推定する一次推定と、前記一次推定の結果を積算して前記1の生言語発音データの国籍を推定する二次推定と、前記二次推定の結果を積算して複数の前記生言語発音データに係る1の外国人の国籍を推定する三次推定と、を行うことを特徴とする、

国籍推定システム。

【請求項 2】

特定言語を母国語としない外国人の音声から外国人の母国語を推定する、外国人の母国語推定システムであって、

複数の言語発音データであって、前記特定言語の単位言語からなる単語情報と、外国人が前記単位言語を発音した音声からなる音声情報と、外国人の母国語からなる母国語情報と、を紐づけてなる、複数の言語発音データを記憶する、記憶手段と、

前記記憶手段内の前記複数の言語発音データを学習用データとする機械学習によって、特定の前記単語情報における前記音声情報と前記母国語情報を関連付けた、学習済みモデルを取得する、学習手段と、

外国人が前記特定言語を発音した音声からなる、生言語発音データを取得する、取得手段と、

前記取得手段が取得した前記生言語発音データを単位言語に分解した、複数の分解言語発音データを生成する、分解手段と、

前記複数の分解言語発音データに、前記学習済みモデルを適用することで、前記生言語発音データに係る外国人の母国語を推定する、推定手段と、を備え、

前記推定手段が、1の前記生言語発音データにおける複数の前記分解言語発音データの母国語を個別に推定する一次推定と、前記一次推定の結果を積算して前記1の生言語発音データの母国語を推定する二次推定と、前記二次推定の結果を積算して複数の前記生言語発音データに係る1の外国人の母国語を推定する三次推定と、を行うことを特徴とする、

母国語推定システム。

【請求項 3】

コンピュータを、請求項 1 に記載の国籍推定システム、又は請求項 2 に記載の母国語推定システムとして機能させるための、

プログラム。

【請求項 4】

特定言語を母国語としない外国人の音声から外国人の国籍を推定する、外国人の国籍推定方法であって、

複数の言語発音データであって、前記特定言語の単位言語からなる単語情報と、外国人が前記単位言語を発音した音声からなる音声情報と、外国人の国籍からなる国籍情報と、を紐づけてなる、複数の言語発音データを記憶する、記憶ステップと、

前記記憶ステップで記憶した前記複数の言語発音データを学習用データとする機械学習によって、特定の単語情報における前記音声情報と前記国籍情報を関連付けた、学習済みモデルを取得する、学習ステップと、

外国人が前記特定言語を発音した音声からなる、生言語発音データを取得する、取得ステップと、

前記取得ステップで取得した前記生言語発音データを単位言語に分解した、複数の分解言語発音データを生成する、分解ステップと、

前記複数の分解言語発音データに、前記学習済みモデルを適用することで、前記生言語発音データに係る外国人の国籍を推定する、推定ステップと、を備え、

前記推定ステップにおいて、1の前記生言語発音データにおける複数の前記分解言語発音データの国籍を個別に推定する一次推定と、前記一次推定の結果を積算して前記1の生言語発音データの国籍を推定する二次推定と、前記二次推定の結果を積算して複数の前記生言語発音データに係る1の外国人の国籍を推定する三次推定と、を行うことを特徴とする、

国籍推定方法。

【請求項 5】

特定言語を母国語としない外国人の音声から外国人の母国語を推定する、外国人の母国語推定方法であって、

複数の言語発音データであって、前記特定言語の単位言語からなる単語情報と、外国人が

前記単位言語を発音した音声からなる音声情報と、外国人の母国語からなる母国語情報と、を紐づけてなる、複数の言語発音データを記憶する、記憶ステップと、

前記記憶ステップで記憶した前記複数の言語発音データを学習用データとする機械学習によって、特定の単語情報における前記音声情報と前記母国語情報を関連付けた、学習済みモデルを取得する、学習ステップと、

外国人が前記特定言語を発音した音声からなる、生言語発音データを取得する、取得ステップと、

前記取得ステップで取得した前記生言語発音データを単位言語に分解した、複数の分解言語発音データを生成する、分解ステップと、

前記複数の分解言語発音データに、前記学習済みモデルを適用することで、前記生言語発音データに係る外国人の母国語を推定する、推定ステップと、を備え、

前記推定ステップにおいて、1の前記生言語発音データにおける複数の前記分解言語発音データの母国語を個別に推定する一次推定と、前記一次推定の結果を積算して前記1の生言語発音データの母国語を推定する二次推定と、前記二次推定の結果を積算して複数の前記生言語発音データに係る1の外国人の母国語を推定する三次推定と、を行うことを特徴とする、

母国語推定方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、外国人の発声による特定言語から当該外国人の国籍を推定するシステム、外国人の発声による特定言語から当該外国人の母国語を推定するシステム、及びこれらのシステムを機能させるプログラムに関する。

【背景技術】

【0002】

生産年齢人口の減少による労働力不足、政府の観光立国構想の推進、経済のグローバル化等の要因によって、近年、日本を訪問/滞在する外国人の人口が激増している。

日本国内に滞在する外国人の国籍は多様であり、営業サービスの質の向上、マーケティングの効率化、治安の維持、又は安全保障上の理由等の様々な目的のため、パスポート等の公的書類や本人の申告によらずに、当該外国人の国籍を判別する需要が存在する。

特許文献1には、個人の人種・形質のような生物学的特性、民族といった文化人類学的特性、発声した言語の属する母国語ないし方言といった言語学的特性その他の個人の外部的特徴によって、個人の国籍を自動的に判定する技術が開示されている。

【先行技術文献】

【特許文献】

【0003】

【特許文献1】特開2010-191530号公報

【非特許文献】

【0004】

【非特許文献1】

高木 直之 他, “日本語話者のための英語発音講座 6.

世界の英語発音”, [online], 東京海洋大学, [令和3年2月3日検索],

インターネット<<http://www2.kaiyodai.ac.jp/~takagi/pweb/world.htm>>

【発明の概要】

【発明が解決しようとする課題】

【0005】

武漢ウィルスの世界的蔓延によってマスクの着用が生活標準となった現在では、特許文

10

20

30

40

50

献 1 のような、顔認証によって国籍を判断する技術の実効性は著しく低下している。

また、特許文献 1 の内、言語学的特性を利用する技術は、外国人の発声する母国語に基づいて機械的に国籍を割り振るものであるが、例えば接客サービスや電話による問い合わせなど日本国内に居住する外国人が会話する場面では、会話の相手方が日本人であり、会話が日本語で行われることが通常であるため、そもそも対象の外国人の母国語を把握することが難しい。

【 0 0 0 6 】

近年では、受付ロボットを使用する無人ホテルが知られている。当該受付ロボットは、A I (人工知能) と音声認識を利用して客と会話を行うことで、客の要求を把握して、宿泊やこれに付随するサービスを提供する。

10

また、駅や空港の交通施設において、A I が音声認識する無人案内サービスが知られている。

これらのサービスにおいて、客が外国人旅行客である場合、外国人の話す日本語から当該外国人の国籍が推定できれば、会話の途中から受付ロボットの言語を当該外国人の母国語に切り替えることで、よりスムーズで快適な接客が可能となる。

【 0 0 0 7 】

犯罪捜査の現場では、外国人の容疑者が自らの氏名や国籍を黙秘する場合がある。

また、空港や港湾における入国管理の現場では、他人によるなりすましを防ぐため、外国人とパスポート記載の人物との照合が必要となる場合がある。

これらの場合、外国人の話す日本語から当該外国人の国籍が推定できれば、大使館への問い合わせなどによって、当該外国人の特定が容易になるなど、捜査や入国管理に資することが考えられる。

20

【 0 0 0 8 】

以上は全て日本語を中心にした例であるが、日本語に限らず、ある特定言語を母国語としない国の者(以下「外国人」と定義)の使用する特定言語から、当該外国人の国籍や母国語を推定する需要が広く存在する。

本発明は、上記のような課題を解決するための、外国人の国籍推定システム、外国人の母国語推定システム、及びプログラムを提供することを目的とするものである。

【課題を解決するための手段】

【 0 0 0 9 】

本発明の外国人の国籍推定システムは、特定言語を母国語としない外国人の音声から外国人の国籍を推定するシステムであって、複数の言語発音データであって、特定言語の単位言語からなる単語情報と、外国人が単位言語を発音した音声からなる音声情報と、外国人の国籍からなる国籍情報と、を紐づけてなる、複数の言語発音データを記憶する、記憶手段と、記憶手段内の複数の言語発音データを学習用データとする機械学習によって、特定の単語情報における音声情報と国籍情報を関連付けた、学習済みモデルを取得する、学習手段と、外国人が特定言語を発音した音声からなる、生言語発音データを取得する、取得手段と、取得手段が取得した生言語発音データを単位言語に分解した、複数の分解言語発音データを生成する、分解手段と、複数の分解言語発音データに、学習済みモデルを適用することで、生言語発音データに係る外国人の国籍を推定する、推定手段と、を備えることを特徴とする。

30

40

【 0 0 1 0 】

本発明の外国人の国籍推定システムは、推定手段が、1の生言語発音データにおける複数の分解言語発音データの国籍を個別に推定する一次推定と、一次推定の結果を積算して1の生言語発音データの国籍を推定する二次推定と、二次推定の結果を積算して複数の生言語発音データに係る1の外国人の国籍を推定する三次推定と、を行ってもよい。

【 0 0 1 1 】

本発明の外国人の国籍推定システムは、複数の言語発音データが、音声情報に紐づけた外国人の性別からなる性別情報を更に有し、学習手段が、特定の単語情報における音声情報と、国籍情報と、性別情報と、を関連付けた、学習済みモデルを取得してもよい。

50

【0012】

本発明の外国人の国籍推定システムは、複数の言語発音データが、音声情報に紐づけた外国人の性別からなる性別情報と、音声情報に紐づけた外国人の年齢又は年代からなる年齢情報と、を有し、学習手段が、特定の単語情報における音声情報と、国籍情報と、性別情報と、年齢情報と、を関連付けた、学習済みモデルを取得してもよい。

【0013】

本発明の外国人の母国語推定システムは、特定言語を母国語としない外国人の音声から外国人の母国語を推定するシステムであって、複数の言語発音データであって、特定言語の単位言語からなる単語情報と、外国人が単位言語を発音した音声からなる音声情報と、外国人の母国語からなる母国語情報と、を紐づけてなる、複数の言語発音データを記憶する、記憶手段と、記憶手段内の複数の言語発音データを学習用データとする機械学習によって、特定の単語情報における音声情報と母国語情報を関連付けた、学習済みモデルを取得する、学習手段と、外国人が特定言語を発音した音声からなる、生言語発音データを取得する、取得手段と、取得手段が取得した生言語発音データを単位言語に分解した、複数の分解言語発音データを生成する、分解手段と、複数の分解言語発音データに、学習済みモデルを適用することで、生言語発音データに係る外国人の母国語を推定する、推定手段と、を備えることを特徴とする。

【0014】

本発明の外国人の母国語推定システムは、推定手段が、1の生言語発音データにおける複数の分解言語発音データの母国語を個別に推定する一次推定と、一次推定の結果を積算して1の生言語発音データの母国語を推定する二次推定と、二次推定の結果を積算して複数の生言語発音データに係る1の外国人の母国語を推定する三次推定と、を行ってもよい。

【0015】

本発明の外国人の母国語推定システムは、複数の言語発音データが、音声情報に紐づけた外国人の性別からなる性別情報を更に有し、学習手段が、特定の単語情報における音声情報と、母国語情報と、性別情報と、を関連付けた、学習済みモデルを取得してもよい。

【0016】

本発明の外国人の母国語推定システムは、複数の言語発音データが、音声情報に紐づけた外国人の性別からなる性別情報と、音声情報に紐づけた外国人の年齢又は年代からなる年齢情報と、を有し、学習手段が、特定の単語情報における音声情報と、母国語情報と、性別情報と、年齢情報と、を関連付けた、学習済みモデルを取得してもよい。

【0017】

本発明のプログラムは、コンピュータを、本発明の国籍推定システム、又は本発明の母国語推定システムとして機能させることを特徴とする。

【0018】

本発明のコンピュータ読み取り可能な記録媒体は、本発明のプログラムを記録したことを特徴とする。

【0019】

本発明の外国人の国籍推定方法は、特定言語を母国語としない外国人の音声から外国人の国籍を推定する方法であって、複数の言語発音データであって、特定言語の単位言語からなる単語情報と、外国人が単位言語を発音した音声からなる音声情報と、外国人の国籍からなる国籍情報と、を紐づけてなる、複数の言語発音データを記憶する、記憶ステップと、記憶ステップで記憶した複数の言語発音データを学習用データとする機械学習によって、特定の単語情報における音声情報と国籍情報を関連付けた、学習済みモデルを取得する、学習ステップと、外国人が特定言語を発音した音声からなる、生言語発音データを取得する、取得ステップと、取得ステップで取得した生言語発音データを単位言語に分解した、複数の分解言語発音データを生成する、分解ステップと、複数の分解言語発音データに、学習済みモデルを適用することで、生言語発音データに係る外国人の国籍を推定する、推定ステップと、を備えることを特徴とする。

【 0 0 2 0 】

本発明の外国人の国籍推定方法は、推定ステップにおいて、1の生言語発音データにおける複数の分解言語発音データの国籍を個別に推定する一次推定と、一次推定の結果を積算して1の生言語発音データの国籍を推定する二次推定と、二次推定の結果を積算して複数の生言語発音データに係る1の外国人の国籍を推定する三次推定と、を行ってもよい。

【 0 0 2 1 】

本発明の外国人の国籍推定方法は、複数の言語発音データが、音声情報に紐づけた外国人の性別からなる性別情報を更に有し、学習ステップにおいて、特定の単語情報における音声情報と、国籍情報と、性別情報と、を関連付けた、学習済みモデルを取得してもよい。

10

【 0 0 2 2 】

本発明の外国人の国籍推定方法は、複数の言語発音データが、音声情報に紐づけた外国人の性別からなる性別情報と、音声情報に紐づけた外国人の年齢又は年代からなる年齢情報と、を有し、学習ステップにおいて、特定の単語情報における音声情報と、国籍情報と、性別情報と、年齢情報と、を関連付けた、学習済みモデルを取得してもよい。

【 0 0 2 3 】

本発明の外国人の母国語推定方法は、特定言語を母国語としない外国人の音声から外国人の母国語を推定する方法であって、複数の言語発音データであって、特定言語の単位言語からなる単語情報と、外国人が単位言語を発音した音声からなる音声情報と、外国人の母国語からなる母国語情報と、を紐づけてなる、複数の言語発音データを記憶する、記憶ステップと、記憶ステップで記憶した複数の言語発音データを学習用データとする機械学習によって、特定の単語情報における音声情報と母国語情報を関連付けた、学習済みモデルを取得する、学習ステップと、外国人が特定言語を発音した音声からなる、生言語発音データを取得する、取得ステップと、取得ステップで取得した生言語発音データを単位言語に分解した、複数の分解言語発音データを生成する、分解ステップと、複数の分解言語発音データに、学習済みモデルを適用することで、生言語発音データに係る外国人の母国語を推定する、推定ステップと、を備えることを特徴とする。

20

【 0 0 2 4 】

本発明の外国人の母国語推定方法は、推定ステップにおいて、1の生言語発音データにおける複数の分解言語発音データの母国語を個別に推定する一次推定と、一次推定の結果を積算して1の生言語発音データの母国語を推定する二次推定と、二次推定の結果を積算して複数の生言語発音データに係る1の外国人の母国語を推定する三次推定と、を行ってもよい。

30

【 0 0 2 5 】

本発明の外国人の母国語推定方法は、複数の言語発音データが、音声情報に紐づけた外国人の性別からなる性別情報を更に有し、学習ステップにおいて、特定の単語情報における音声情報と、母国語情報と、性別情報と、を関連付けた、学習済みモデルを取得してもよい。

【 0 0 2 6 】

本発明の外国人の母国語推定方法は、複数の言語発音データが、音声情報に紐づけた外国人の性別からなる性別情報と、音声情報に紐づけた外国人の年齢又は年代からなる年齢情報と、を有し、学習ステップにおいて、特定の単語情報における音声情報と、母国語情報と、性別情報と、年齢情報と、を関連付けた、学習済みモデルを取得してもよい。

40

【 発明の効果 】**【 0 0 2 7 】**

本発明によれば、機械学習によって外国人の国籍又は母国語を高い精度で推定することができる。また、推定結果のフィードバックにより、推定の精度を永続的に向上させ続けることができる。

【 図面の簡単な説明 】**【 0 0 2 8 】**

50

【図 1】本発明の国籍推定システムの説明図。

【図 2】本発明の母国語推定システムの説明図。

【図 3】実施例 2 の説明図。

【図 4】実施例 3 の説明図。

【発明を実施するための形態】

【0029】

以下、実施例を参照しながら本発明の外国人の国籍推定システム、母国語推定システム、及びプログラムについて詳細に説明する。

本発明において「特定言語」とは日本語、英語、スペイン語等の任意の 1 種類の言語を意味する。また、「外国人」とは特定言語を母国語としない者を意味し、必ずしも本発明を実施する国の国籍に基づくものではない。

【実施例 1】

【0030】

[国籍推定システム]

本発明の国籍推定システム 1 は、外国人の発声による特定言語から当該外国人の国籍を推定するシステムである。

非特許文献 1 には、英語を母国語としない者の英語には、母国語の影響による訛りがあり、経験のある船員であれば、VHF 無線から聞こえる英語の訛りによって相手の国籍が判断できることが開示されている。これは英語の例であるが、いずれの言語においても、その言語を母国語としない外国人の発音には程度の差こそあれ当該外国人の国籍に由来する特徴が表れる。すなわち、外国人による特定言語の発音は、当該外国人の国籍と相関関係を有する。

そこで、本発明の国籍推定システム 1 は、外国人が母国語以外の特定言語を話す際の発音の特徴（例えば訛り）に着目して、これを教師データとした機械学習によって学習モデルを取得し、未知の外国人が特定言語を話した音声に当該学習モデルを適用することで、未知の外国人の国籍を高精度に推定しようとする発明である。

なお、本実施例では「特定言語」が日本語である場合について説明するが、これは例示に過ぎず、本発明の国籍推定システム 1 は、いかなる言語にも適用が可能である。

【0031】

[国籍推定システムの構成]

国籍推定システム 1 は、少なくとも記憶手段 20、学習手段 30、取得手段 40、分解手段 50、及び推定手段 60 を備える（図 1）。本例では、更に生成手段 10 を備え、これらの各手段に情報処理を実行させるプログラムを、情報処理装置（コンピュータ）にインストールしてなる。

この情報処理装置は単体である必要はなく、複数の情報処理装置（演算装置、制御装置、記憶装置、入力装置、出力装置等を備えた装置）を、ネットワークを介して相互に接続して、1 つの国籍推定システム 1 を構成してもよい。

プログラムは、汎用のシステムに機能追加する形で導入するプラグイン形式のプログラムであってもよい。あるいはクラウドシステムを介して提供されてもよい。また、プログラムは、情報処理装置で読み取り可能な記憶媒体に格納することができる。

【0032】

（言語発音データ）

言語発音データ A は、機械学習の学習用データとなるデータである。

言語発音データ A は、少なくとも、単語情報 A 1 と、単語情報 A 1 と紐づけた音声情報 A 2 及び国籍情報 A 3 と、を含む。

単語情報 A 1 とは、特定言語の単位言語からなる。ここで「単位言語」とは、特定言語における意味のある文章を、単語、文節、文等の任意の単位に区切ったものであり、予め多数の単位言語をリスト化しておくことができる。

音声情報 A 2 は、外国人が単位言語を発音した音声からなる。

国籍情報 A 3 は、外国人の国籍からなる。なお、該当国が広域である場合、地域によ

て言語特性が大きく異なる場合等には、国籍情報 A 3 を同一国内の複数の地域に細分してもよい。

【 0 0 3 3 】

[生成手段]

生成手段 1 0 は、言語発音データ A を生成する手段である。

生成手段 1 0 は、外国人の発音による音声情報 A 2 に、対応する単語情報 A 1 を紐づけ、これに正解データである国籍情報 A 3 を付加して、言語発音データ A を生成する。

詳細には、例えば収音マイク等を用いて音声による空気の振動を波形データに変換し、環境音などのノイズを取り除くことで音声情報 A 2 を生成する。

続いて、音声情報 A 2 に公知の音声認識技術（後述）を適用して単位言語ごとに解析し、テキスト情報である単語情報 A 1 を生成する。

最後に、単語情報 A 1 に紐づいた音声情報 A 2 に、音声情報 A 2 の発音に係る外国人の国籍情報 A 3 を付加する。

言語発音データ A は後述する学習手段 3 0 の学習用データセットとし使用するため、多数の単位言語について相当量を生成する必要がある。

多数の言語発音データ A を取得する方法には、例えば外国人モニターに特定言語のリストを読み上げさせ、その音声を収音する方法、テレビやラジオ等の放送媒体から取得する方法、インターネット上でクロールした動画データ等から AI が音声情報 A 2 であると機械的に判断したデータを収集する方法、電話サービスの通話音声から収集する方法、等を用いることができる。

この場合、言語発音データ A から、当該外国人の氏名や電話番号を削除したり、動画であれば付随する画像データを削除するなどの匿名加工を行うことで、言語発音データ A を、個人を特定できない匿名加工情報とするのが望ましい。

また、言語発音データ A が欠測値や外れ値である場合には、当該言語発音データ A を除去することができる。

以上の手順により、言語発音データ A を生成して、記憶手段 2 0 へ送信する（生成ステップ S 1 ）。

【 0 0 3 4 】

[記憶手段]

記憶手段 2 0 は、複数の言語発音データ A を保存する手段である。

記憶手段 2 0 は、主記憶部及び補助記憶部を含むことができる。例えば、主記憶部は RAM などの揮発性メモリであり、補助記憶部は、ROM、EEPROM、フラッシュメモリ、又はハードディスクなどの不揮発性メモリである。

記憶手段 2 0 は、生成手段 1 0 によって入力された言語発音データ A を記憶し、多数の言語発音データ A を蓄積して、機械学習用の学習用データセットを構成する（記憶ステップ S 2 ）。

【 0 0 3 5 】

[学習手段]

学習手段 3 0 は、複数の言語発音データ A を学習データとした機械学習に基づいて、学習済みモデル X を取得する手段である。

学習手段 3 0 は、記憶手段 2 0 内に保存した大量の言語発音データ A からなる学習用データセットを、公知の機械学習アルゴリズムに基づいて機械学習して、学習済みパラメータが組み込まれた学習済みモデル X を取得する（学習ステップ S 3 ）。

この学習済みモデル X は、多数の単位言語における多数の音声情報 A 2 と国籍情報 A 3 の相関から導き出した一定の規則を反映したモデルである。

機械学習は、例えばディープラーニングの方法により、サポートベクターマシン（SVM）、ロジスティック回帰、ランダムフォレスト（決定木）等の公知の様々なアルゴリズムを利用できる。

【 0 0 3 6 】

（生言語発音データ）

10

20

30

40

50

生言語発音データ B は、推定の対象となる外国人による音声のデータである。

生言語発音データ B は、少なくとも外国人が特定言語を発音した音声データである生音声情報 B 1 を含む。

【 0 0 3 7 】

[取得手段]

取得手段 4 0 は、生言語発音データ B を取得する手段である。

取得手段 4 0 は、例えば收音マイクや電話回線である。

取得手段 4 0 により、推定の対象となる外国人の会話から生言語発音データ B を收音して、分解手段 5 0 へ送信する（取得ステップ S 4）。

【 0 0 3 8 】

(単位言語発音データ)

単位言語発音データ C は、生言語発音データ B を単位言語ごとに分解してなるデータである。

単位言語発音データ C は、分解音声情報 C 1 と、分解音声情報 C 1 に紐づいた分解テキスト情報 C 2 と、を有する。

単位言語発音データ C は、元になる生言語発音データ B ごとにグループ化される。

【 0 0 3 9 】

[分解手段]

分解手段 5 0 は、生言語発音データ B を単位言語毎に分解して、分解言語発音データ C を生成する手段である。

分解手段 5 0 は、最初に、取得手段 4 0 が取得した生言語発音データ B を、音声認識によってテキスト化する。

詳細には、例えば、生言語発音データ B を、音響分析、音響モデル、発音辞書、言語モデル等の要素を用いて解析し、音声情報をテキスト情報に変換する。

続いてテキスト情報を、自然言語処理 (Natural Language Processing) によって、当該テキスト情報を単位言語ごとに分節して、分解テキスト情報 C 2 を生成する。

自然言語処理とは、人間の言語 (自然言語) を機械で処理して内容を抽出する技術である。本例では、例えば特定言語の文法の性質を利用して言語間の境界を判断する文節構造モデルや、確率的言語モデルを適用することができる。

続いて、分解テキスト情報 C 2 の区分に基づいて、生言語発音データ B の生音声情報 B 1 を分解音声情報 C 1 に分解する。

こうして、生言語発音データ B ごとくの複数の分解音声情報 C 1 と、各分解音声情報 C 1 に紐づけた分解テキスト情報 C 2 を有する、単位言語発音データ C が得られる (分解ステップ S 5) 。

具体的には、例えば、生言語発音データ B が、外国人の発声による日本語「今日は頭が痛いので、アルバイトを休みます。」からなる場合、分解手段 5 0 によって、この生言語発音データ B を、「今日は」「頭が」「痛いので」「アルバイトを」「休みます」との文節による 5 つの単位言語に分解し、それぞれの単位言語に対応する分解音声情報 C 1 と分解テキスト情報 C 2 を得る。

この他、例えば「今日」「は」「頭」「が」「痛い」「ので」「アルバイト」「を」「休みます」と、単語による 9 つの単位言語に分解してもよいし、「今日は」「頭が痛いので」「アルバイトを休みます」と 3 つの単位言語に分解してもよい。

【 0 0 4 0 】

[推定手段]

推定手段 6 0 は、外国人の国籍を推定する手段である。

推定手段 6 0 は、分解手段 5 0 が生成した分解言語発音データ C に、学習手段 3 0 が取得した学習済みモデル X を適用して、分解言語発音データ C の発音に係る外国人の国籍を推定する (推定ステップ S 6) 。国籍の推定は、少なくとも一次推定及び二次推定を含む。

10

20

30

40

50

詳細には、推定手段60は、1の生言語発音データBに係る複数の分解言語発音データCを取得し、各分解言語発音データCの分解テキスト情報C2によって、各分解言語発音データCにかかる単位言語を特定する。

続いて、特定した単位言語に対応する学習済みモデルXを選択する。この際、対応する学習済みモデルXを有さない単位言語については、推定の対象から除くことができる。

各分解言語発音データCの分解音声情報C1に、単位言語に対応する学習済みモデルXを適用して、分解言語発音データCごとに国籍を推定する（一次推定）。一次推定の結果は、各分解言語発音データCにおける、複数の国別の推定率（％）として記録することができる。

上記の手順により、1の生言語発音データB内にグループ化された全ての分解言語発音データCに一次推定を行い、全ての一次推定の積算により、当該1の生言語発音データB全体の国籍を推定する（二次推定）。二次推定の結果は、当該1の生言語発音データBにおける、複数の国別の推定率（％）として記録することができる。

1の外国人から複数の生言語発音データBを取得できる場合には、一次推定及び二次推定に加えて、複数の二次推定結果を積算してなる三次推定を行うことができる。詳細な実施例について次項で説明する。

【0041】

[一次推定及び二次推定]

例えば、1つの生言語発音データB（AA00001）が「今日は」「頭が」「痛いので」「アルバイトを」「休みます」の5つの分解言語発音データCからなる場合、推定手段60は、以下の手順で一次推定及び二次推定を行う（表1）。

この5つの単位言語から、対応する学習済みモデルXを有さない「アルバイトを」を除いた4つの単位言語に、この単位言語に対応する学習済みモデルXをそれぞれ適用し、分解言語発音データC「今日は」に対応する国籍を[A国：71.1% B国：21.7% C国：7.2%]と推定する。

同様に、分解言語発音データC「頭が」に対応する国籍を[A国：55.6% B国：28.2% C国：9.0% D国：7.2%]と、分解言語発音データC「痛いので」に対応する国籍を[A国：45.2% B国：51.7% D国：3.1%]と、分解言語発音データC「休みます」に対応する国籍を[A国：38.9% B国：32.0% C国：29.1%]と、それぞれ一次推定する。

続いて、4つの単位言語それぞれの一次推定率を積算して、生言語発音データB（AA00001）の二次推定率を算出する。本例では二次推定率は[A国：52.7% B国：33.4% C国：11.3% D国：2.6%]と算出される。

なお、本例では説明の便宜上、複数の一次推定率を単純に合算して二次推定率を算出したが、実際には、特定言語の種類（発音の特徴が顕著に表れる特定言語とそうでない特定言語等）、特定言語の長さ、特定言語の発音順等によって、一次推定の結果に重みづけを行ったり、推定率の分布から信頼性の低い推定結果を除外するなど、統計処理上の各種の調整を行ってもよい。

【0042】

10

20

30

40

【表 1】

分解言語発音データ [AA00001]	A国	B国	C国	D国
00001 [今日は]	71.1%	21.7%	7.2%	-
00002 [頭が]	55.6%	28.2%	9.0%	7.2%
00003 [痛いので]	45.2%	51.7%	-	3.1%
00004 [アルバイトを]	-	-	-	-
00005 [休みます]	38.9%	32.0%	29.1%	-
合計	210.8	133.6	45.3	10.3
二次推定率	52.7%	33.4%	11.3%	2.6%

【0043】

[三次推定]

続いて推定手段60は、生言語発音データB(AA00001)を含む、同一の外国人に係る複数の生言語発音データB(AA00001、AA00002、AA00003・
20
・・・)の二次推定率を積算して、当該外国人に係る三次推定率を算出する(表2)。

本例では三次推定の結果が[A国:74.2% B国:17.4% C国:5.8%
D国:1.0% E国:1.3%・・・]と算出される。

以上の結果より、最終的に当該外国人の国籍を74.2%の推定率で、A国であると推定する。

なお、三次推定にも、二次推定と同様の各種の統計処理上の調整を行うことができる。

【0044】

【表 2】

生言語発音データ	A国	B国	C国	D国	E国	・・・
AA00001[今日は・・・ます]	52.7%	33.4%	11.3%	2.6%	-	
AA00002[・・・・・・・・]	72.2%	16.9%	8.2%	-	2.7%	
AA00003	42.5%	50.2%	3.0%	1.4%	2.9%	
AA00004	84.0%	8.2%	2.1%	4.1%	-	
・・・						
合計	816.1	191.1	63.5	11.0	13.8	
三次推定率	74.2%	17.4%	5.8%	1.0%	1.3%	・・・

【実施例 2】

【0045】

[母国語推定システム]

本発明の母国語推定システム1'は、外国人の発声による特定言語から当該外国人の母
50

国語を推定するシステムである（図２）。

本発明の母国語推定システム１'は、外国人が母国語以外の特定言語を話す際の発音の特徴（例えば訛り）に着目して、これを教師データとした機械学習によって学習モデルを取得し、未知の外国人が特定言語を話した音声に学習モデルを適用することで、未知の外国人の母国語を高精度に推定しようとする発明である。

母国語推定システム１'は、少なくとも記憶手段２０、学習手段３０、取得手段４０、分解手段５０、及び推定手段６０を備える。本例では、更に生成手段１０を備える。

母国語推定システム１'の具体的構成と母国語の推定方法は、（１）言語発音データＡが、国籍情報Ａ３ではなく、外国人の母国語の情報である母国語情報Ａ４を備える点、（２）学習手段３０が、国籍情報Ａ３ではなく母国語情報Ａ４を紐付けた機械学習に基づいて学習済みモデルＸを取得する点、及び（３）推定手段６０が、最終的に外国人の母国語を推定する点、以外は実施例１の国籍推定システム１と同一である。

【実施例３】

【００４６】

[言語発音データが性別情報を有する例]

本発明の国籍推定システム１及び母国語推定システム１'において、言語発音データＡに、性別情報Ａ５を備えてさせてもよい（図３）。

性別情報Ａ５とは、外国人の性別の情報である。

学習手段３０は、特定の単語情報Ａ１について、音声情報Ａ２と、国籍情報Ａ３（又は母国語情報Ａ４）と、性別情報Ａ５を関連付けて、学習済みモデルＸを取得する。

本例の場合、学習済みモデルＸに、性別による音声の違いを反映させることで、より精度の高い推定が可能となる。

【実施例４】

【００４７】

[言語発音データが年齢情報を有する例]

本発明の国籍推定システム１及び母国語推定システム１'において、言語発音データＡに、年齢情報Ａ６を備えてさせてもよい（図４）。

年齢情報Ａ６とは、外国人の年齢又は年代の情報である。

学習手段３０は、特定の単語情報Ａ１について、音声情報Ａ２と、国籍情報Ａ３（又は母国語情報Ａ４）と、年齢情報Ａ６を関連付けて、学習済みモデルＸを取得する。

本例の場合、学習済みモデルＸに、年齢や年代による音声の違いを反映させることで、より精度の高い推定が可能となる。

なお、実施例３と実施例４を掛け合わせ、性別情報Ａ５を年齢情報Ａ６に加える構成としてもよい。

【符号の説明】

【００４８】

- １ 国籍推定システム
- １' 母国語推定システム
- １０ 生成手段
- ２０ 記憶手段
- ３０ 学習手段
- ４０ 取得手段
- ５０ 分解手段
- ６０ 推定手段
- Ｓ１ 生成ステップ
- Ｓ２ 記憶ステップ
- Ｓ３ 学習ステップ
- Ｓ４ 取得ステップ
- Ｓ５ 分解ステップ
- Ｓ６ 推定ステップ

10

20

30

40

50

- A 言語発音データ
- A 1 単語情報
- A 2 音声情報
- A 3 国籍情報
- A 4 母国語情報
- A 5 性別情報
- A 6 年齢情報
- B 生言語発音データ
- B 1 生音声情報
- C 分解言語発音データ
- C 1 分解音声情報
- C 2 分解テキスト情報
- X 学習済みモデル

10

【要約】

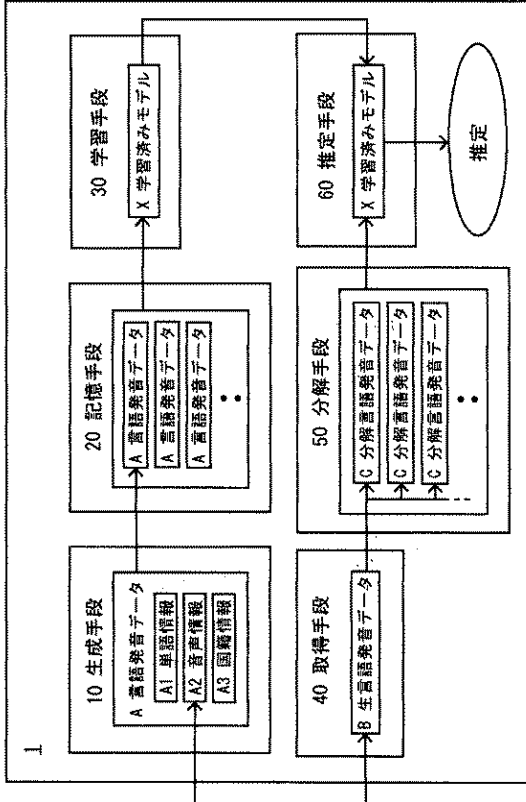
【課題】外国人の国籍又は母国語を高い精度で推定可能な、国籍推定システム、母国語推定システム、及びこれらのシステムを機能させるプログラムを提供すること。

【解決手段】本発明の外国人の国籍推定システム1は、少なくとも記憶手段20、学習手段30、取得手段40、分解手段50、及び推定手段60を備え、単語情報A1、音声情報A2、及び国籍情報A3、を紐づけてなる複数の言語発音データAを学習用データとする機械学習によって学習済みモデルXを取得し、生言語発音データBに学習済みモデルXを適用して、国籍を推定する。本発明の外国人の母国語推定システム1'は、少なくとも記憶手段20、学習手段30、取得手段40、分解手段50、及び推定手段60を備え、単語情報A1、音声情報A2、及び母国語情報A4、を紐づけてなる複数の言語発音データAを学習用データとする機械学習によって学習済みモデルXを取得し、生言語発音データBに学習済みモデルXを適用して、母国語を推定する。

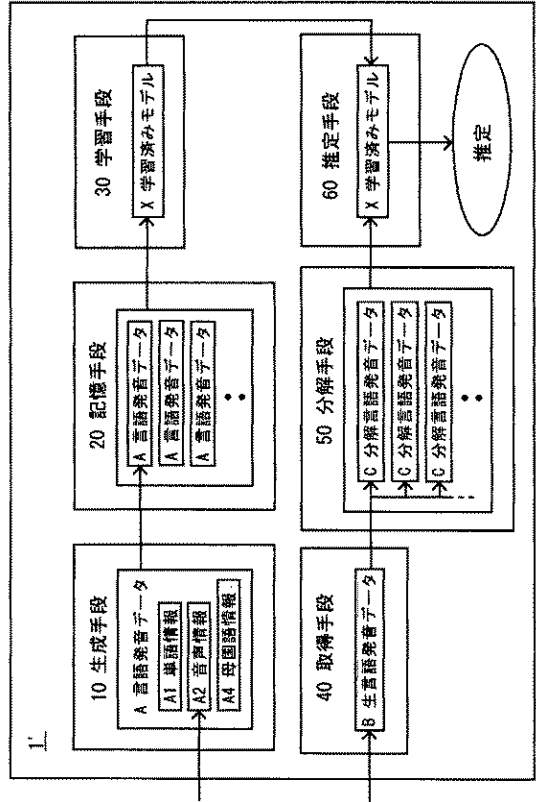
20

【選択図】図1

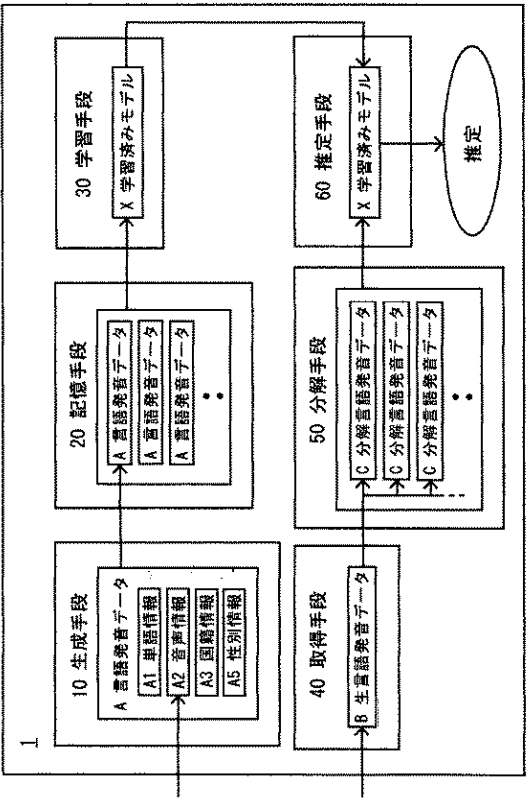
【図1】



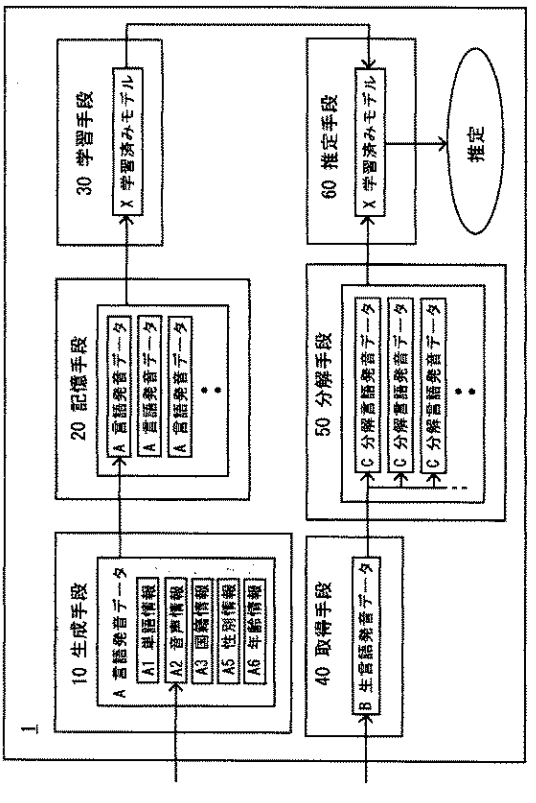
【図2】



【図3】



【図4】



フロントページの続き

(56)参考文献 特開2009-251388(JP,A)
特開2010-191530(JP,A)
特開2017-041120(JP,A)
特開2012-103554(JP,A)
米国特許出願公開第2004/0098259(US,A1)

(58)調査した分野(Int.Cl., DB名)
G10L 15/00 - 15/34