

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7677432号
(P7677432)

(45)発行日 令和7年5月15日(2025.5.15)

(24)登録日 令和7年5月7日(2025.5.7)

(51)国際特許分類

G 1 0 L 17/00 (2013.01)

G 1 0 L 15/10 (2006.01)

G 1 0 L 17/24 (2013.01)

F I

G 1 0 L 17/00 2 0 0 D

G 1 0 L 15/10 2 0 0 W

G 1 0 L 17/00 2 0 0 Z

G 1 0 L 17/24

請求項の数 7 (全26頁)

(21)出願番号	特願2023-539570(P2023-539570)	(73)特許権者	000004237 日本電気株式会社 東京都港区芝五丁目7番1号
(86)(22)出願日	令和3年8月6日(2021.8.6)	(74)代理人	100104765 弁理士 江上 達夫
(86)国際出願番号	PCT/JP2021/029412	(74)代理人	100107331 弁理士 中村 聡延
(87)国際公開番号	WO2023/013060	(74)代理人	100131015 弁理士 三輪 浩誉
(87)国際公開日	令和5年2月9日(2023.2.9)	(72)発明者	幸田 芳紀 東京都港区芝五丁目7番1号 日本電気株式会社内
審査請求日	令和6年1月9日(2024.1.9)	審査官	大野 弘

最終頁に続く

(54)【発明の名称】 情報処理システム、情報処理装置、情報処理方法、及び記録媒体

(57)【特許請求の範囲】

【請求項1】

複数人の音声情報を含む会話データを取得する取得手段と、
前記音声情報から所定回数以上発話された単語をキーワードとして抽出するキーワード抽出手段と、
前記音声情報から前記キーワードを発話した際の音声に関する特徴量である第1特徴量を抽出する特徴量抽出手段と、
前記キーワードと前記第1特徴量とを関連付けた照合用情報を生成する生成手段と、
を備える情報処理システム。

【請求項2】

前記複数人のうち少なくとも1人の音声に関する特徴量である第2特徴量を取得する特徴量取得手段と、
前記第1特徴量と前記第2特徴量とを比較することで、前記第1特徴量から前記キーワードを発話した話者を特定できるか否かを判定する判定手段と、
を更に備える請求項1に記載の情報処理システム。

【請求項3】

前記会話データに対する所定処理を要求するユーザに対して、前記照合用情報が生成された前記キーワードの発話を促す情報を提示する提示手段と、
前記ユーザの発話内容から、前記ユーザの音声に関する特徴量である第3特徴量を抽出する認証用特徴量抽出手段と、

発話を促した前記キーワードに関連付いた前記第 1 特徴量と、前記第 3 特徴量との比較結果に基づいて、前記ユーザによる前記所定処理の実行を許可するか否かを判定する許可判定手段と、

を更に備える請求項 1 又は 2 に記載の情報処理システム。

【請求項 4】

前記照合用情報は複数の前記キーワードについて生成されており、

前記提示手段は、一部の前記キーワードの発話を促す情報を提示して、前記ユーザによる前記所定処理の実行を許可しないと判定された場合に、他の前記キーワードの発話を促す情報を提示する、

請求項 3 に記載の情報処理システム。

10

【請求項 5】

複数人の音声情報を含む会話データを取得する取得手段と、

前記音声情報から所定回数以上発話された単語をキーワードとして抽出するキーワード抽出手段と、

前記音声情報から前記キーワードを発話した際の音声に関する特徴量である第 1 特徴量を抽出する特徴量抽出手段と、

前記キーワードと前記第 1 特徴量とを関連付けた照合用情報を生成する生成手段と、
を備える情報処理装置。

【請求項 6】

少なくとも 1 つのコンピュータが実行する情報処理方法であって、

複数人の音声情報を含む会話データを取得し、

前記音声情報から所定回数以上発話された単語をキーワードとして抽出し、

前記音声情報から前記キーワードを発話した際の音声に関する特徴量である第 1 特徴量を抽出し、

前記キーワードと前記第 1 特徴量とを関連付けた照合用情報を生成する、
情報処理方法。

20

【請求項 7】

少なくとも 1 つのコンピュータに、

複数人の音声情報を含む会話データを取得し、

前記音声情報から所定回数以上発話された単語をキーワードとして抽出し、

前記音声情報から前記キーワードを発話した際の音声に関する特徴量である第 1 特徴量を抽出し、

前記キーワードと前記第 1 特徴量とを関連付けた照合用情報を生成する、
情報処理方法を実行させるコンピュータプログラム。

30

【発明の詳細な説明】

【技術分野】

【0001】

この開示は、情報処理システム、情報処理装置、情報処理方法、及び記録媒体の技術分野に関する。

【背景技術】

40

【0002】

この種のシステムとして、音声認識技術にキーワードを利用するものが知られている。例えば特許文献 1 では、入力音声から所定のキーワードを発話した音声であるキーワード音声を検出する技術が開示されている。特許文献 2 では、キーワードリストを作成して、音声情報から重要語を抽出する技術が開示されている。特許文献 3 では、音声認識された入力内容からユーザの興味特定に使用するキーワードを抽出する技術が開示されている。特許文献 4 では、音声認識によって生成した文字情報からキーワードを生成する技術が開示されている。

【0003】

その他の関連する技術として、特許文献 5 では、ユーザの声道及びユーザの話し方のパ

50

ターンの挙動に関する情報に基づいて、ユーザのボイスプリントを生成する技術が開示されている。

【先行技術文献】

【特許文献】

【0004】

【文献】特開2020-086011号公報

【文献】特開2015-099290号公報

【文献】特開2009-294790号公報

【文献】特開2007-257134号公報

【文献】特表2014-517366号公報

【発明の概要】

【発明が解決しようとする課題】

【0005】

この開示は、先行技術文献に開示された技術を改善することを目的とする。

【課題を解決するための手段】

【0006】

この開示の情報処理システムの一の態様は、複数人の音声情報を含む会話データを取得する取得手段と、前記音声情報からキーワードを抽出するキーワード抽出手段と、前記音声情報から前記キーワードを発話した際の音声に関する特徴量である第1特徴量を抽出する特徴量抽出手段と、前記キーワードと前記第1特徴量とを関連付けた照合用情報を生成する生成手段と、を備える。

【0007】

この開示の情報処理装置の一の態様は、複数人の音声情報を含む会話データを取得する取得手段と、前記音声情報からキーワードを抽出するキーワード抽出手段と、前記音声情報から前記キーワードを発話した際の音声に関する特徴量である第1特徴量を抽出する特徴量抽出手段と、前記キーワードと前記第1特徴量とを関連付けた照合用情報を生成する生成手段と、を備える。

【0008】

この開示の情報処理方法の一の態様は、少なくとも1つのコンピュータが実行する情報処理方法であって、複数人の音声情報を含む会話データを取得し、前記音声情報からキーワードを抽出し、前記音声情報から前記キーワードを発話した際の音声に関する特徴量である第1特徴量を抽出し、前記キーワードと前記第1特徴量とを関連付けた照合用情報を生成する。

【0009】

この開示の記録媒体の一の態様は、少なくとも1つのコンピュータに、複数人の音声情報を含む会話データを取得し、前記音声情報からキーワードを抽出し、前記音声情報から前記キーワードを発話した際の音声に関する特徴量である第1特徴量を抽出し、前記キーワードと前記第1特徴量とを関連付けた照合用情報を生成する、情報処理方法を実行させるコンピュータプログラムが記録されている。

【図面の簡単な説明】

【0010】

【図1】第1実施形態に係る情報処理システムのハードウェア構成を示すブロック図である。

【図2】第1実施形態に係る情報処理システムの機能的構成を示すブロック図である。

【図3】第1実施形態に係る情報処理システムによる情報生成動作の流れを示すフローチャートである。

【図4】第2実施形態に係る情報処理システムの機能的構成を示すブロック図である。

【図5】第2実施形態に係る情報処理システムによる情報生成動作の流れを示すフローチャートである。

【図6】第3実施形態に係る情報処理システムによる話者分類の具体例を示す概念図であ

10

20

30

40

50

るである。

【図 7】第 3 実施形態に係る情報処理システムによる話者集約の具体例を示す概念図である。

【図 8】第 3 実施形態に係る情報処理システムによるキーワード抽出の具体例を示す概念図である。

【図 9】第 3 実施形態に係る情報処理システムにおけるキーワードの記憶態様の一例を示す表である。

【図 10】第 4 実施形態に係る情報処理システムの機能的構成を示すブロック図である。

【図 11】第 4 実施形態に係る情報処理システムによる許可判定動作の流れを示すフローチャートである。

【図 12】第 4 実施形態に係る情報処理システムによる提示例を示す平面図である。

【図 13】第 4 実施形態に係る情報処理システムが扱うファイルの表示例を示す平面図である。

【図 14】第 5 実施形態に係る情報処理システムの機能的構成を示すブロック図である。

【図 15】第 5 実施形態に係る情報処理システムの許可判定動作の流れを示すフローチャートである。

【図 16】第 5 実施形態に係る情報処理システムによるキーワード表示変更の一例を示す平面図である。

【図 17】第 6 実施形態に係る情報処理システムのアプリケーション適用例を示すブロック図（その 1）である。

【図 18】第 6 実施形態に係る情報処理システムのアプリケーション適用例を示すブロック図（その 2）である。

【図 19】第 6 実施形態に係る情報処理システムのアプリケーション適用例を示すブロック図（その 3）である。

【図 20】第 7 実施形態に係る情報処理システム 10 による表示例を示す平面図である。

【発明を実施するための形態】

【0011】

以下、図面を参照しながら、情報処理システム、情報処理方法、及び記録媒体の実施形態について説明する。

【0012】

<第 1 実施形態>

第 1 実施形態に係る情報処理システムについて、図 1 から図 3 を参照して説明する。

【0013】

（ハードウェア構成）

まず、図 1 を参照しながら、第 1 実施形態に係る情報処理システムのハードウェア構成について説明する。図 1 は、第 1 実施形態に係る情報処理システムのハードウェア構成を示すブロック図である。

【0014】

図 1 に示すように、第 1 実施形態に係る情報処理システム 10 は、プロセッサ 11 と、RAM (Random Access Memory) 12 と、ROM (Read Only Memory) 13 と、記憶装置 14 とを備えている。情報処理システム 10 は更に、入力装置 15 と、出力装置 16 と、を備えていてもよい。上述したプロセッサ 11 と、RAM 12 と、ROM 13 と、記憶装置 14 と、入力装置 15 と、出力装置 16 とは、データバス 17 を介して接続されている。

【0015】

プロセッサ 11 は、コンピュータプログラムを読み込む。例えば、プロセッサ 11 は、RAM 12、ROM 13 及び記憶装置 14 のうちの少なくとも一つが記憶しているコンピュータプログラムを読み込むように構成されている。或いは、プロセッサ 11 は、コンピュータで読み取り可能な記録媒体が記憶しているコンピュータプログラムを、図示しない記録媒体読み取り装置を用いて読み込んでもよい。プロセッサ 11 は、ネットワークイン

10

20

30

40

50

タフェースを介して、情報処理システム 10 の外部に配置される不図示の装置からコンピュータプログラムを取得してもよい（つまり、読み込んでもよい）。プロセッサ 11 は、読み込んだコンピュータプログラムを実行することで、RAM 12、記憶装置 14、入力装置 15 及び出力装置 16 を制御する。本実施形態では特に、プロセッサ 11 が読み込んだコンピュータプログラムを実行すると、プロセッサ 11 内には、会話データからキーワードを抽出して情報を生成するための機能ブロックが実現される。

【0016】

プロセッサ 11 は、例えば CPU (Central Processing Unit)、GPU (Graphics Processing Unit)、FPGA (field-programmable gate array)、DSP (Demand-Side Platform)、ASIC (Application Specific Integrated Circuit) として構成されてよい。プロセッサ 11 は、これらのうち一つで構成されてもよいし、複数を並列で用いるように構成されてもよい。

10

【0017】

RAM 12 は、プロセッサ 11 が実行するコンピュータプログラムを一時的に記憶する。RAM 12 は、プロセッサ 11 がコンピュータプログラムを実行している際にプロセッサ 11 が一時的に使用するデータを一時的に記憶する。RAM 12 は、例えば、DRAM (Dynamic RAM) であってもよい。

【0018】

ROM 13 は、プロセッサ 11 が実行するコンピュータプログラムを記憶する。ROM 13 は、その他に固定的なデータを記憶していてもよい。ROM 13 は、例えば、PROM (Programmable ROM) であってもよい。

20

【0019】

記憶装置 14 は、情報処理システム 10 が長期的に保存するデータを記憶する。記憶装置 14 は、プロセッサ 11 の一時記憶装置として動作してもよい。記憶装置 14 は、例えば、ハードディスク装置、光磁気ディスク装置、SSD (Solid State Drive) 及びディスクアレイ装置のうちの少なくとも一つを含んでいてもよい。

【0020】

入力装置 15 は、情報処理システム 10 のユーザからの入力指示を受け取る装置である。入力装置 15 は、例えば、キーボード、マウス及びタッチパネルのうちの少なくとも一つを含んでいてもよい。入力装置 15 は、スマートフォンやタブレット等の携帯端末として構成されていてもよい。

30

【0021】

出力装置 16 は、情報処理システム 10 に関する情報を外部に対して出力する装置である。例えば、出力装置 16 は、情報処理システム 10 に関する情報を表示可能な表示装置（例えば、ディスプレイ）であってもよい。また、出力装置 16 は、情報処理システム 10 に関する情報を音声出力可能なスピーカ等であってもよい。出力装置 16 は、スマートフォンやタブレット等の携帯端末として構成されていてもよい。

【0022】

なお、図 1 では、複数の装置を含んで構成される情報処理システム 10 の例を挙げたが、これらの全部又は一部の機能を、1 つの装置（情報処理装置）で実現してもよい。この情報処理装置は、例えば、上述したプロセッサ 11、RAM 12、ROM 13 のみを備えて構成され、その他の構成要素（即ち、記憶装置 14、入力装置 15、出力装置 16）については、例えば情報処理装置に接続される外部の装置が備えるようにしてもよい。また、情報処理装置は、一部の演算機能を外部の装置（例えば、外部サーバやクラウド等）によって実現するものであってもよい。

40

【0023】

（機能的構成）

次に、図 2 を参照しながら、第 1 実施形態に係る情報処理システム 10 の機能的構成について説明する。図 2 は、第 1 実施形態に係る情報処理システムの機能的構成を示すプロ

50

ック図である。

【0024】

図2に示すように、第1実施形態に係る情報処理システム10は、その機能を実現するための構成要素として、会話データ取得部110と、キーワード抽出部120と、特徴量抽出部130と、照合用情報生成部140と、を備えて構成されている。会話データ取得部110、キーワード抽出部120、特徴量抽出部130、及び照合用情報生成部140の各々は、例えば上述したプロセッサ11（図1参照）によって実現される処理ブロックであってよい。

【0025】

会話データ取得部110は、複数人の音声情報を含む会話データを取得する。会話データ取得部110は、例えばマイク等から直接会話データを取得してもよいし、他の装置等で生成された会話データを取得してもよい。会話データの一例としては、会議の音声を録音した会議データ等が挙げられる。また、会話データ取得部110は、取得した会話データに対して各種処理を実行可能に構成されてよい。例えば、会話データ取得部110は、会話データにおいて話者が発話している区間を検出する処理、会話データを音声認識してテキスト化する処理、及び発話している話者を分類する処理等を実行可能に構成されてよい。

【0026】

キーワード抽出部120は、会話データ取得部110で取得された会話データの音声情報から、発話内容に含まれているキーワードを抽出する。キーワード抽出部120は、音声情報に含まれている単語の中からランダムにキーワードを抽出するようにしてもよいし、予め定められた単語をキーワードとして抽出するようにしてもよい。また、キーワード抽出部120は、会話データの内容に応じて抽出するキーワードを決定してもよい。例えば、キーワード抽出部120は、会話データにおいて出現頻度の高い単語（例えば、所定回数以上発話された単語）をキーワードとして抽出するようにしてもよい。キーワード抽出部120は、1つの会話データから複数のキーワードを抽出するようにしてもよい。キーワード抽出部120は、複数人の各々について少なくとも1つのキーワードを抽出するようにしてもよい。

【0027】

特徴量抽出部130は、キーワード抽出部120において抽出されたキーワードを発話した際の音声に関する特徴量（以下、適宜「第1特徴量」と称する）を抽出可能に構成されている。キーワード抽出部120において複数のキーワードが抽出されている場合、特徴量抽出部130は、全てのキーワードについて特徴量を抽出してもよいし、一部のキーワードについてのみ特徴量を抽出してもよい。なお、音声に関する特徴量の抽出手法については、既存の技術を適宜採用することができるため、ここでの詳細な説明は省略する。

【0028】

照合用情報生成部140は、キーワード抽出部120で抽出されたキーワードと、特徴量抽出部130で抽出された第1特徴量とを関連付けることで、照合用情報を生成可能に構成されている。例えば、照合用情報生成部140は、第1のキーワードと、第1のキーワードを発話した際の音声に関する特徴量とを互いに関連付け、第2のキーワードと、第2のキーワードを発話した際の音声に関する特徴量とを互いに関連付けてよい。照合用情報生成部140で生成された照合用情報は、会話に参加した複数人の音声照合に用いられる。照合用情報の具体的な利用方法については、後述する他の実施形態において詳しく説明する。

【0029】

（情報生成動作）

次に、図3を参照しながら、第1実施形態に係る情報処理システム10による照合用情報を生成する際の動作（以下、適宜「情報生成動作」と称する）の流れについて説明する。図3は、第1実施形態に係る情報処理システムによる情報生成動作の流れを示すフローチャートである。

【0030】

図3に示すように、第1実施形態に係る情報処理システム10による情報生成動作では、まず会話データ取得部110が、複数人の音声情報を含む会話データを取得する（ステップS101）。そして、会話データ取得部110は、会話データにおいて話者が発話している区間を検出する処理（以下、適宜「区間検出処理」と称する）を実行する（ステップS102）。区間検出処理は、例えば無音区間を検出してトリミングする処理であってよい。

【0031】

続いて、会話データ取得部110は、区間検出処理が実行された会話データ（即ち、発話している区間の音声情報）から、話者を分類する処理（以下、適宜「話者分類処理」と称する）を実行する（ステップS103）。話者分類処理は、例えば会話データの各区間に話者に応じたラベルを付与する処理であってよい。

10

【0032】

他方で、会話データ取得部110は、区間検出処理が実行された会話データに対して音声認識を行ってテキスト化する処理（以下、適宜「音声認識処理」と称する）を実行する（ステップS104）。音声認識処理の具体的な手法については、既存の技術を適宜採用できるため、ここでの詳細な説明は省略する。なお、音声認識処理と、上述した話者分類処理とは、並行して同時に実行されてもよいし、相前後して順次実行されてもよい。

【0033】

続いて、キーワード抽出部120が、音声認識処理が実行された会話データ（即ち、テキストデータ）からキーワードを抽出する（ステップS105）。この際、キーワード抽出部120は、話者分類処理の結果を用いて（例えば、話者を区別して）キーワードを抽出してよい。なお、キーワード抽出部120は、同じ漢字の単語であっても異なる読み方をするものについては、それらを区別して抽出してよい。例えば、「一」という漢字の場合、「いち」と読むものと、「ひとつ」と読むものとで別々に抽出してよい。

20

【0034】

続いて、特徴量抽出部130が、キーワード抽出部120で抽出されたキーワードを発話した際の音声に関する特徴量（即ち、第1特徴量）を抽出する（ステップS106）。そして、照合用情報生成部140は、キーワード抽出部120で抽出されたキーワードと、特徴量抽出部130で抽出された第1特徴量とを関連付けて、照合用情報を生成する（ステップS107）。

30

【0035】

（技術的効果）

次に、第1実施形態に係る情報処理システム10によって得られる技術的効果について説明する。

【0036】

図1から図3で説明したように、第1実施形態に係る情報処理システム10では、会話データから抽出されたキーワードと、その音声に関する特徴量（即ち、第1特徴量）とを関連付けて照合用情報が生成される。このようにすれば、複数人の音声情報を含む会話データから、照合用情報を適切に生成することができる。よって、会話に参加した複数人に対して、キーワードを用いた音声照合処理を適切に実行することが可能となる。また、本実施形態では、会話データからキーワードが抽出されるため、音声照合処理に用いるキーワードを別途用意する必要がない。よって、照合用情報を生成するのに要する手間を削減することが可能である。

40

【0037】

事前に決めた音声のキーワードを使い回す場合、悪意により録音された音声や音声合成で対応されてしまうおそれがある。しかしながら本実施形態では、事前に決めたキーワードを用いない（会話データからキーワードを生成できる）ため、悪意への堅牢性を高めることが可能である。また、会話データから自動的にキーワードが生成されるため、事前登録が不要となり、ユーザに意識してキーワードを用意させる必要もない。更に、キーワー

50

ドを失念してしまうことを回避できる。例えば、複数の会議で異なるキーワードを用意しておけば、精度を高めることができる一方で、キーワードを失念する可能性も高くなってしまう。しかしながら本実施形態では、複数キーワードを用意する場合と同様の精度を実現しつつ、キーワードを失念してしまうような状況も回避できる。

【0038】

＜第2実施形態＞

第2実施形態に係る情報処理システム10について、図4及び図5を参照して説明する。なお、第2実施形態は、上述した第1実施形態と一部の構成及び動作が異なるのみであり、その他の部分については第1実施形態と同一であってよい。このため、以下では、すでに説明した第1実施形態と異なる部分について詳細に説明し、その他の重複する部分については適宜説明を省略するものとする。

10

【0039】

(機能的構成)

まず、図4を参照しながら、第2実施形態に係る情報処理システム10の機能的構成について説明する。図4は、第2実施形態に係る情報処理システムの機能的構成を示すブロック図である。なお、図4では、図2で示した構成要素と同様の要素に同一の符号を付している。

【0040】

図4に示すように、第2実施形態に係る情報処理システム10は、その機能を実現するための構成要素として、会話データ取得部110と、キーワード抽出部120と、特徴量抽出部130と、照合用情報生成部140と、特徴量取得部150と、利用可否判定部160と、を備えて構成されている。即ち、第2実施形態に係る情報処理システム10は、第1実施形態の構成(図2参照)に加えて、特徴量取得部150と、利用可否判定部160と、を更に備えて構成されている。なお、特徴量取得部150及び利用可否判定部160の各々は、例えば上述したプロセッサ11(図1参照)によって実現される処理ブロックであってよい。

20

【0041】

特徴量取得部150は、会話に参加した複数人の少なくとも1人の音声に関する特徴量(以下、適宜「第2特徴量」と称する)を取得可能に構成されている。特徴量取得部150は、会話データ取得部110が取得した会話データから第2特徴量を取得してよい。例えば、特徴量取得部150は、話者分類処理が実行された会話データから第2特徴量を抽出してよい。或いは、特徴量取得部150は、予め用意された第2特徴量を取得してもよい。例えば、会話に参加した複数人の各々の個人IDや保有端末と紐付けて記憶されている第2特徴量を取得してもよい。

30

【0042】

利用可否判定部160は、特徴量抽出部130で抽出された第1特徴量と、特徴量取得部150で取得された第2特徴量とを比較することで、第1特徴量からキーワードを発話した話者を特定できるか否かを判定可能に構成されている。即ち、利用可否判定部160は、キーワードに対応する第1特徴量が、音声照合に利用可能であるかを判定可能に構成されている。利用可否判定部160は、同一の話者から抽出した第1特徴量と第2特徴量とを照合して、それらの話者が同一人物であると判定できた場合に、その第1特徴量は音声照合に利用可能であると判定してよい。また、利用可否判定部160は、同一の話者から抽出した第1特徴量と第2特徴量とを照合して、それらの話者が同一人物でないと判定された場合に、その第1特徴量は音声照合に利用可能でないと判定してよい。

40

【0043】

(情報生成動作)

次に、図5を参照しながら、第2実施形態に係る情報処理システム10による情報生成動作の流れについて説明する。図5は、第2実施形態に係る情報処理システムによる情報生成動作の流れを示すフローチャートである。なお、図5では、図3で説明した処理と同様の処理に同一の符号を付している。

50

【0044】

図5に示すように、第2実施形態に係る情報処理システム10による情報生成動作では、まず会話データ取得部110が、複数人の音声情報を含む会話データを取得する（ステップS101）。そして、会話データ取得部110は、区間検出処理を実行する（ステップS102）。

【0045】

続いて、会話データ取得部110は、区間検出処理が実行された会話データに対して、話者分類処理を実行する（ステップS103）。そして、第2実施形態では、特徴量取得部150が、話者分類処理を実行した会話データから第2特徴量を取得する（ステップS201）。なお、既に説明したように、特徴量取得部150は、会話データ以外から第2特徴量を取得してもよい。

10

【0046】

他方、会話データ取得部110は、区間検出処理が実行された会話データに対して、音声認識処理を実行する（ステップS104）。そして、キーワード抽出部120が、音声認識処理が実行された会話データからキーワードを抽出する（ステップS105）。この際、キーワード抽出部120は、話者分類処理の結果を用いて（例えば、話者を区別して）キーワードを抽出してよい。その後、特徴量抽出部130が、キーワード抽出部120で抽出されたキーワードに対応する第1特徴量を抽出する（ステップS106）。

【0047】

なお、上述したステップS103及びS201の処理（即ち、フロー左側の処理）と、ステップS104、S105及びS106の処理（即ち、フロー右側の処理）とは、並行して同時に実行されてもよいし、相前後して順次実行されてもよい。

20

【0048】

続いて、第2実施形態では、利用可否判定部160が、特徴量抽出部130で抽出された第1特徴量と、特徴量取得部150で取得された第2特徴量とを比較して、第1特徴量からキーワードを発話した話者を特定できるか否かを判定する（ステップS202）。ここで、第1特徴量からキーワードを発話した話者を特定できると判定された場合（ステップS202：YES）、照合用情報生成部140は、キーワード抽出部120で抽出されたキーワードと、特徴量抽出部130で抽出された第1特徴量とを関連付けて、照合用情報を生成する（ステップS107）。一方、第1特徴量からキーワードを発話した話者を特定できないと判定された場合（ステップS202：NO）、上述したステップS107の処理は省略される。即ち、話者を特定できないと判定されたキーワードについては、照合用情報は生成されない。

30

【0049】

（技術的効果）

次に、第2実施形態に係る情報処理システム10によって得られる技術的効果について説明する。

【0050】

図4及び図5で説明したように、第2実施形態に係る情報処理システム10では、第1特徴量と第2特徴量を比較して、キーワードによる音声照合が可能であるか否かが判定される。このようにすれば、音声照合に適さないキーワードについて照合用情報が生成されてしまうことを防止できる。よって、照合用情報を用いた音声照合の精度を高めることが可能である。

40

【0051】

<第3実施形態>

第3実施形態に係る情報処理システム10について、図6から図9を参照して説明する。なお、第3実施形態は、上述した第1及び第2実施形態で実行される処理の具体例等を説明するものであり、その構成や動作については第1及び第2実施形態と同一であってよい。このため、以下では、すでに説明した各実施形態と異なる部分について詳細に説明し、その他の重複する部分については適宜説明を省略するものとする。

50

【0052】

(話者分類処理)

まず、図6を参照しながら、第3実施形態に係る情報処理システム10が実行する話者分類処理(即ち、図3及び図5のステップS103の処理)の具体例について説明する。図6は、第3実施形態に係る情報処理システムによる話者分類の具体例を示す概念図である。

【0053】

図6に示すような音声認識データ(即ち、会話データをテキスト化したデータ)が、第3実施形態に係る情報処理システム10で取得されているとする。この場合、話者分類処理では、音声認識データの各区間に、話者に対応するラベルが付与されてよい。図6に示す例では、音声認識データの各区間に対して、話者A、話者B、及び話者Cに対応するラベルが付与されている。これにより、どの区間を、どの話者が発話したものが認識できるようになる。

10

【0054】

(話者集約処理)

次に、図7を参照しながら、第3実施形態に係る情報処理システム10が実行する話者集約処理(即ち、話者分類データから話者を絞り込む処理)の具体例について説明する。図7は、第3実施形態に係る情報処理システムによる話者集約の具体例を示す概念図である。

【0055】

図7に示すような話者分類データ(即ち、話者分類されたデータ)が、第3実施形態に係る情報処理システム10で取得されているとする。この場合、話者集約する処理では、話者分類データから、いずれか1人の話者が発話した区間が抽出されてよい。なお、図7に示す例では、話者Aの発話した区間を抽出する例を挙げているが、これらに加えて又は代えて、他の話者が発話した区間を抽出する処理が実行されてよい。

20

【0056】

(キーワード抽出処理)

次に、図8を参照しながら、第3実施形態に係る情報処理システム10が実行するキーワード抽出処理(即ち、話者集約データからキーワードを抽出する処理)の具体例について説明する。図8は、第3実施形態に係る情報処理システムによるキーワード抽出の具体例を示す概念図である。

30

【0057】

図8に示すような話者集約データが、第3実施形態に係る情報処理システム10で取得されているとする。この場合、キーワード抽出処理では、話者集約データにおいて複数回発話されている単語がキーワードとして抽出される。図8に示す例では、太字で示す「本日」、「会議」、「保存」の3つの単語が複数回発話されている。このため、これら3つの単語がキーワードとして抽出される。なお、話者集約データが複数の話者に対して取得されている場合(例えば、話者Bや話者Cについても話者集約データが取得されている場合)、複数の話者について、それぞれキーワードを抽出する処理が実行されてよい。

【0058】

(キーワード記憶)

次に、図9を参照しながら、第3実施形態に係る情報処理システム10におけるキーワードの記憶態様の具体例について説明する。図9は、第3実施形態に係る情報処理システムにおけるキーワードの記憶態様の一例を示す表である。

40

【0059】

図9に示すように、キーワード抽出処理で抽出されたキーワードは、話者ごとに別々に記憶されてよい。例えば、話者A、話者B、話者C、話者Dが存在する場合、話者Aの発話区間から抽出されたキーワードは話者Aに対応するキーワードとして記憶される。話者Bの発話区間から抽出されたキーワードは話者Bに対応するキーワードとして記憶される。話者Cの発話区間から抽出されたキーワードは話者Cに対応するキーワードとして記憶

50

される。話者Dの発話区間から抽出されたキーワードは話者Dに対応するキーワードとして記憶される。なお、これらのキーワードから照合用情報が生成された場合、照合用情報についても話者ごとに記憶されてよい。

【0060】

(技術的效果)

次に、第3実施形態に係る情報処理システム10によって得られる技術的效果について説明する。

【0061】

図6から図9で説明したように、第3実施形態に係る情報処理システム10によれば、照合用情報を生成する各種処理を適切な態様で実行することが可能である。ただし、上述した各種処理は上記態様に限定されるものではなく、各種処理をここで説明した態様とは異なる態様で実行するようにしてもよい。

【0062】

<第4実施形態>

第4実施形態に係る情報処理システム10について、図10から図13を参照して説明する。なお、第4実施形態は、上述した第1から第3実施形態と一部の構成及び動作が異なるのみであり、その他の部分については第1から第3実施形態と同一であってよい。このため、以下では、すでに説明した各実施形態と異なる部分について詳細に説明し、その他の重複する部分については適宜説明を省略するものとする。

【0063】

(機能的構成)

まず、図10を参照しながら、第4実施形態に係る情報処理システム10の機能的構成について説明する。図10は、第4実施形態に係る情報処理システムの機能的構成を示すブロック図である。なお、図10では、図2で示した構成要素と同様の要素に同一の符号を付している。

【0064】

図10に示すように、第4実施形態に係る情報処理システム10は、その機能を実現するための構成要素として、会話データ取得部110と、キーワード抽出部120と、特徴量抽出部130と、照合用情報生成部140と、照合用情報記憶部210と、キーワード提示部220と、認証用特徴量抽出部230と、許可判定部240と、を備えて構成されている。即ち、第4実施形態に係る情報処理システム10は、第1実施形態の構成(図2参照)に加えて、照合用情報記憶部210と、キーワード提示部220と、認証用特徴量抽出部230と、許可判定部240と、を更に備えて構成されている。照合用情報記憶部210は、例えば上述した記憶装置14によって実現されてよい。また、キーワード提示部220、認証用特徴量抽出部230、許可判定部240の各々は、例えば上述したプロセッサ11(図1参照)によって実現される処理ブロックであってよい。

【0065】

照合用情報記憶部210は、照合用情報生成部140で生成された照合用情報を記憶可能に構成されている。照合用情報記憶部210は、既に説明したように、会話に参加していた話者ごとに照合用情報を記憶可能に構成されてよい(図9参照)。照合用情報記憶部210が記憶している照合用情報は、キーワード提示部220によって適宜読み出し可能とされている。

【0066】

キーワード提示部220は、会話データに対する所定処理を要求するユーザに対して、照合用情報記憶部210に記憶されている照合用情報に含まれるキーワードを提示可能に構成されている。キーワード提示部220は、例えば出力装置16(図1参照)を用いてキーワードを提示してよい。キーワード提示部220は、ユーザが所定処理を実行するための操作(例えば、右クリックやダブルクリック等)を行ったタイミングで、キーワードを提示するようにしてもよい。なお、所定処理の一例としては、会話データのファイルを開く処理、暗号化された会話データのファイルを復号化する処理、及び会話データのファ

10

20

30

40

50

イルを編集する処理等が挙げられる。

【0067】

照合用情報が話者ごとに記憶されている場合、キーワード提示部220は、ユーザがどの話者であるかを判別してから、その話者に対応するキーワードを提示するようにしてもよい。キーワード提示部220は、例えばユーザの入力（例えば、氏名や個人ID等の入力）から話者を判別し、その話者に対応するキーワードを提示してよい。或いは、キーワード提示部220は、顔認証等を用いてどの話者であるかを判別し、その話者に対応するキーワードを提示してよい。

【0068】

また、照合用情報記憶部210が複数のキーワードを記憶している場合、キーワード提示部220は、記憶されている複数のキーワードの中から、提示するキーワードを選択して提示するようにしてもよい。また、キーワード提示部220は、複数のキーワードを接合して提示してもよい。この場合、キーワード提示部220は、予め定められた個数のキーワードを接合して提示してもよい。或いは、キーワード提示部220は、接合したキーワードの長さが話者を特定するために十分な長さとなるように（即ち、適切な音声照合が行える長さとなるように）、キーワードを選択してもよい。例えば、話者を特定するのに1.5秒の発話が必要となる場合、0.5秒に相当するワードを3つ接合したものを選択して提示すればよい。

【0069】

認証用特徴量抽出部230は、キーワードを提示した後にユーザが発話した内容（即ち、提示したキーワードに対応する発話内容）から、その音声に関する特徴量（以下、適宜「第3特徴量」と称する）を抽出可能に構成されている。第3特徴量は、第1特徴量（即ち、照合用情報としてキーワードと関連付けて記憶されている特徴量）と照合可能な特徴量である。

【0070】

許可判定部240は、キーワード提示部220で提示したキーワードに関連付いた第1特徴量と、認証用特徴量抽出部230で抽出した第3特徴量を比較して、ユーザによる所定処理の実行を許可するか否かを判定する。具体的には、許可判定部240は、第1特徴量と第3特徴量の照合の結果、会話データにおいてキーワードを発話している人物と、会話データに対する所定処理を要求しているユーザとが同一人物であると判定された場合に、ユーザによる所定処理の実行を許可してよい。また、会話データにおいてキーワードを発話している人物と、会話データに対する所定処理を要求しているユーザとが同一人物でないと判定された場合に、ユーザによる所定処理の実行を禁止してよい。

【0071】

（許可判定動作）

次に、図11を参照しながら、第4実施形態に係る情報処理システム10による所定処理を許可するか否かを判定する動作（以下、適宜「許可判定動作」と称する）の流れについて説明する。図11は、第4実施形態に係る情報処理システムによる許可判定動作の流れを示すフローチャートである。なお、図11で示す許可判定動作は、第1及び第2実施形態で説明した情報生成動作が実行された後に（言い換えれば、照合用情報が生成された状態で）行われるものとする。

【0072】

図11に示すように、第4実施形態に係る情報処理システム10による許可判定動作では、まずキーワード提示部220が、照合用情報記憶部210が記憶している照合用情報を読み出して、ユーザに提示するキーワードを生成する（ステップS401）。そして、キーワード提示部220は、生成したキーワードをユーザに対して提示する（ステップS402）。

【0073】

なお、ユーザに提示するキーワードが1つである場合、キーワード提示部220は、読み出した照合用情報に含まれるキーワードをそのまま提示すればよい。また、ユーザに提

10

20

30

40

50

示するキーワードが複数である場合、キーワード提示部 220 は、読み出した照合用情報に含まれるキーワードを接合して提示すればよい。なお、キーワードの具体的な提示例については後に詳しく説明する。

【0074】

続いて、認証用特微量抽出部 230 が、ユーザの発話データ（具体的には、キーワードの提示を受けたユーザの発話によって取得された音声情報）を取得する（ステップ S403）。そして、認証用特微量抽出部 230 は、取得した発話データから、第 3 特微量を抽出する（ステップ S404）。

【0075】

続いて、許可判定部 240 が、提示したキーワードに対応する第 1 特微量と、認証用特微量抽出部 230 で抽出した第 3 特微量と、を照合して認証処理を実行する（ステップ S405）。ここで、認証が成功した場合（ステップ S405：YES）、許可判定部 240 は、ユーザに対して所定処理の実行を許可する（ステップ S406）。一方、認証が成功しなかった場合（ステップ S405：NO）、許可判定部 240 は、ユーザに対して所定処理の実行を許可しない（ステップ S407）。

【0076】

（キーワードの提示例）

次に、図 12 を参照しながら、第 4 実施形態に係るキーワード提示部 220 によるキーワードの提示例について説明する。図 12 は、第 4 実施形態に係る情報処理システムによる提示例を示す平面図である。

【0077】

図 12 に示すように、キーワード提示部 220 は、ディスプレイにキーワードを表示することで、ユーザにキーワードを提示してよい。この例では、「本日」、「会議」、「保存」の 3 つのキーワードがユーザに提示されている。また、キーワードに加えて、「以下の言葉を発話してください。」のようなメッセージを表示することで、ユーザがキーワードを発話することを促してもよい。なお、キーワードの提示は、音声で行われてもよい。具体的には、図 12 で表示されているキーワードやメッセージを、スピーカ等を用いて音声出力してもよい。

【0078】

なお、ここでは提示した 3 つのキーワードをすべて発話するように促しているが、提示した複数のキーワードの中から一部のキーワードを選択して発話させるように促してもよい。この場合、「下記のキーワードから 1 つのキーワードを選択して発話してください。」のようなメッセージを表示してもよい。また、複数のキーワードを発話させる場合、その順序を固定してもよいし、固定しなくてもよい。具体的には、「本日」、「会議」、「保存」の 3 つのキーワードをユーザに提示した場合、「本日」、「会議」、「保存」の順番（即ち、表示された順番）で発話した場合にのみ認証が成功するようにしてもよいし、「会議」、「保存」、「本日」の順番（即ち、表示された順番とは異なる順番）で発話した場合でも認証が成功するようにしてもよい。

【0079】

（ファイルの表示例）

次に、図 13 を参照しながら、第 4 実施形態に係る情報処理システム 10 で扱われるデータファイル（即ち、所定処理の対象となるファイル）の表示例について説明する。図 13 は、第 4 実施形態に係る情報処理システムが扱うファイルの表示例を示す平面図である。

【0080】

図 13 に示すように、第 4 実施形態に係る情報処理システム 10 で扱われるデータファイルは、音声アイコンを付して表示されてもよい。このようにすれば、会話データに対して所定処理を要求するユーザが、直感的に認証方法を把握することができる。即ち、キーワードの発話によって認証可能なデータファイルであることを、ユーザに対して視覚的に知らせることができる。

【0081】

(技術的効果)

次に、第４実施形態に係る情報処理システム１０によって得られる技術的効果について説明する。

【００８２】

図１０から図１３で説明したように、第４実施形態に係る情報処理システム１０では、キーワードを提示した際のユーザの発話内容に基づいて、会話データに対する所定処理の実行可否が判定される。このようにすれば、所定処理を要求するユーザについて、所定処理を実行する権限があるか否かを適切に判定することができる。言い換えれば、ユーザが会話に参加していた人物か否かを適切に判定できる。よって、会話に参加していない第三者によって所定処理が実行されてしまうことを防止できる。なお、発話によって所定処理を許可する手法として、例えば予め定型句を用意しておく方法も考えられるが、発話する際に盗聴されるおそれがある。また、キーワードを毎回変えてもよいが、手間がかかり、キーワードを忘れてしまうおそれもある。しかるに本実施形態に係る情報処理システム１０によれば、会話データから抽出されるキーワードを提示し、そのキーワードの発話によって所定処理を許可できるため、上述した問題点をすべて解消することが可能である。

【００８３】

<第５実施形態>

第５実施形態に係る情報処理システム１０について、図１４から図１６を参照して説明する。なお、第５実施形態は、上述した第４実施形態と一部の構成及び動作が異なるのみであり、その他の部分については第１から第４実施形態と同一であってよい。このため、以下では、すでに説明した各実施形態と異なる部分について詳細に説明し、その他の重複する部分については適宜説明を省略するものとする。

【００８４】

(機能的構成)

まず、図１４を参照しながら、第５実施形態に係る情報処理システム１０の機能的構成について説明する。図１４は、第５実施形態に係る情報処理システムの機能的構成を示すブロック図である。なお、図１４では、図１０で示した構成要素と同様の要素に同一の符号を付している。

【００８５】

図１４に示すように、第５実施形態に係る情報処理システム１０は、その機能を実現するための構成要素として、会話データ取得部１１０と、キーワード抽出部１２０と、特徴量抽出部１３０と、照合用情報生成部１４０と、照合用情報記憶部２１０と、キーワード提示部２２０と、認証用特徴量抽出部２３０と、許可判定部２４０と、キーワード変更部２５０と、を備えて構成されている。即ち、第５実施形態に係る情報処理システム１０は、第４実施形態の構成（図１０参照）に加えて、キーワード変更部２５０を更に備えて構成されている。なお、キーワード変更部２５０は、例えば上述したプロセッサ１１（図１参照）によって実現される処理ブロックであってよい。

【００８６】

キーワード変更部２５０は、キーワード提示部２２０が提示するキーワードを変更可能に構成されている。具体的には、キーワード変更部２５０は、許可判定部２４０において会話データに対する所定処理の実行が許可されなかった場合に、キーワード提示部２２０が提示するキーワードを変更可能に構成されている。

【００８７】

(許可判定動作)

次に、図１５を参照しながら、第５実施形態に係る情報処理システム１０による許可判定動作の流れについて説明する。図１５は、第５実施形態に係る情報処理システムによる許可判定動作の流れを示すフローチャートである。なお、図１５では、図１１で示した処理と同様の処理に同一の符号を付している。

【００８８】

図１５に示すように、第５実施形態に係る情報処理システム１０による許可判定動作で

は、まずキーワード提示部 220 が、照合用情報記憶部 210 が記憶している照合用情報を読み出して、ユーザに提示するキーワードを生成する（ステップ S401）。そして、キーワード提示部 220 は、生成したキーワードをユーザに対して提示する（ステップ S402）。

【0089】

続いて、認証用特微量抽出部 230 が、ユーザの発話データ（即ち、ユーザの発話に応じた音声情報）を取得する（ステップ S403）。そして、認証用特微量抽出部 230 は、取得した発話データから、第 3 特微量を抽出する（ステップ S404）。

【0090】

続いて、許可判定部 240 が、提示したキーワードに対応する第 1 特微量と、認証用特微量抽出部 230 で抽出した第 3 特微量とを照合して認証処理を実行する（ステップ S405）。ここで、認証が成功した場合（ステップ S405：YES）、許可判定部 240 は、ユーザに対して所定処理の実行を許可する（ステップ S406）。一方、認証が成功しなかった場合（ステップ S405：NO）、許可判定部 240 は、ユーザに対して所定処理の実行を許可しない（ステップ S407）。

【0091】

そして本実施形態では特に、ユーザに対して所定処理の実行が許可されなかった場合に、キーワード変更部 250 が、他のキーワード（即ち、まだ提示していない別のキーワード）が残っているか否かを判定する（ステップ S501）。そして、他のキーワードが残っている場合（ステップ S501：YES）、キーワード変更部 250 は、キーワード提示部 220 が提示するキーワードを他のキーワードに変更する（ステップ S502）。この場合、ステップ S402 から処理が再開される。即ち、変更されたキーワードの発話に基づいて、再び同様の判定が行われることになる。なお、他のキーワードが残っていない場合（ステップ S501：NO）、所定処理の実行が許可されないまま一連の処理が終了する。

【0092】

（キーワードの変更例）

次に、図 16 を参照しながら、第 5 実施形態に係るキーワード変更部 250 によるキーワードの変更例について説明する。図 16 は、第 5 実施形態に係る情報処理システムによるキーワード表示変更の一例を示す平面図である。

【0093】

図 16 に示すように、最初に「本日」、「会議」、「保存」の 3 つのキーワードが提示されていたとする。そして、所定処理の実行を許可しないと判定された場合、キーワード変更部 250 は、提示するキーワードを「会議」、「予算」、「機能」の 3 つに変更する。このように、キーワード変更部 250 は、一部のキーワードのみを変更してもよい。即ち、複数のキーワードを接合して提示する場合には、変更前後において一部重複したキーワードが提示されてもよい。また、キーワード変更部 250 は、すべてのキーワードを変更するようにしてもよい。更に、キーワード変更部 250 は、表示するキーワードの数を変更するようにしてもよい。

【0094】

なお、キーワード提示部 220 は、キーワードを変更した際に、キーワードと共に表示していたメッセージを変更してもよい。例えば、図 16 で示すように「認証に失敗しました。再認証する場合は以下の言葉を発話してください。」というメッセージを表示するようにしてもよい。このようにすれば、ユーザに対して再度キーワードの発話を促すことができる。

【0095】

（技術的効果）

次に、第 5 実施形態に係る情報処理システム 10 によって得られる技術的効果について説明する。

【0096】

図 1 4 から図 1 6 で説明したように、第 5 実施形態に係る情報処理システム 1 0 では、キーワードを用いた認証処理が失敗した場合に、ユーザに提示されるキーワードが変更される。通常の生体認証やパスワード照合では、照合用の情報を変更することはできないが、第 5 実施形態に係る複数のキーワードは本人性を示すもののため、変更が可能である。このようにすれば、認証処理によって本人拒否が発生した場合でも、再び認証処理を行うことが可能である。本実施形態は特に、再認証の際にキーワードが変更されるため、キーワードが照合に不適切なものであった場合であっても、変更後に適切な認証処理が行われることになる。

【 0 0 9 7 】

＜第 6 実施形態＞

第 6 実施形態に係る情報処理システム 1 0 について、図 1 7 から図 1 9 を参照して説明する。なお、第 6 実施形態は、上述した第 1 から第 5 実施形態に係る情報処理システムの具体的な適用例を説明するものであり、その構成や動作については第 1 から第 5 実施形態と同一であってよい。このため、以下では、すでに説明した各実施形態と異なる部分について詳細に説明し、その他の重複する部分については適宜説明を省略するものとする。

【 0 0 9 8 】

（会議アプリと共通のアプリに適用）

まず、図 1 7 を参照しながら、第 6 実施形態に係る情報処理システム 1 0 が、会話データを生成する会議アプリと共通のアプリに適用される例について説明する。図 1 7 は、第 6 実施形態に係る情報処理システムのアプリケーション適用例を示すブロック図（その 1）である。なお、図 1 7 では、説明の便宜上、第 6 実施形態に係る情報処理システム 1 0 が備える構成要素として、会話データ取得部 1 1 0、キーワード抽出部 1 2 0、特徴量抽出部 1 3 0、及び照合用情報生成部 1 4 0（即ち、第 1 実施形態（図 2 参照）の構成要素）のみを図示しているが、第 6 実施形態に係る情報処理システム 1 0 は、第 2 から第 5 実施形態で説明した構成要素を備えていてもよい。

【 0 0 9 9 】

図 1 7 に示すように、第 6 実施形態に係る情報処理システム 1 0 は、端末 5 0 0 にインストールされている会議アプリ A p p 1 の一部の機能として実現されてもよい。この場合、会話データ取得部 1 1 0 は、会議アプリ A p p 1 が有する会話データ生成部 5 0 において生成される会話データを取得するように構成されればよい。

【 0 1 0 0 】

（会議アプリと別のアプリに適用）

次に、図 1 8 を参照しながら、第 6 実施形態に係る情報処理システム 1 0 が、会話データを生成する会議アプリと別のアプリに適用される例について説明する。図 1 8 は、第 6 実施形態に係る情報処理システムのアプリケーション適用例を示すブロック図（その 2）である。なお、図 1 8 では、図 1 7 で示した構成要素と同様の要素に同一の符号を付している。

【 0 1 0 1 】

図 1 8 に示すように、第 6 実施形態に係る情報処理システム 1 0 は、端末 5 0 0 にインストールされている会議アプリ A p p 2 とは別のアプリ（情報生成アプリ A p p 3）の機能として実現されてもよい。この場合、会議アプリ A p p 2 と、情報生成アプリ A p p 3 を連携させることで、会話データ生成部 5 0 において生成される会話データが、会話データ取得部 1 1 0 によって取得される。

【 0 1 0 2 】

（会議アプリと別端末のアプリに適用）

次に、図 1 9 を参照しながら、第 6 実施形態に係る情報処理システム 1 0 が、会話データを生成する会議アプリと別端末のアプリに適用される例について説明する。図 1 9 は、第 6 実施形態に係る情報処理システムのアプリケーション適用例を示すブロック図（その 3）である。なお、図 1 9 では、図 1 8 で示した構成要素と同様の要素に同一の符号を付している。

10

20

30

40

50

【0103】

図19に示すように、第6実施形態に係る情報処理システム10は、会議アプリApp2がインストールされている端末501とは別の端末（即ち、端末502）にインストールされた情報生成アプリApp3の機能として実現されてもよい。この場合、会議アプリApp2がインストールされた端末501と、情報生成アプリApp3がインストールされた端末502とがデータ通信を行うことで、会話データ生成部50において生成される会話データが、会話データ取得部110によって取得される。

【0104】

なお、上述したアプリApp1～App3で使用する各種情報（例えば、会話データ、キーワード及び特徴量等）等については、端末500、501及び端末502のストレージではなく、外部サーバの記憶装置等に記憶されていてもよい。この場合、端末500、501及び502は、必要に応じて外部サーバと通信を行い、利用する情報を適宜送受信すればよい。

【0105】

（技術的效果）

次に、第6実施形態に係る情報処理システム10によって得られる技術的效果について説明する。

【0106】

図16から図19で説明したように、第6実施形態に係る情報処理システム10によれば、上述した第1から第5実施形態の各種機能を、適切な態様で実現することができる。なお、ここで挙げた適用例はあくまで一例であり、本実施形態に係る情報処理システム10の機能は、ここで挙げていない態様によって実現することもできる。また、第6実施形態では、会話データを生成するアプリの一例として会議アプリ（会議の録画や録音を行うアプリ）を挙げているが、会議アプリを他のアプリに置き換えても同様に適用可能である。

【0107】

＜第7実施形態＞

第7実施形態に係る情報処理システム10について、図20を参照して説明する。なお、第7実施形態は、上述した第1から第6実施形態と一部の構成及び動作が異なるのみであり、その他の部分については第1から第6実施形態と同一であってよい。このため、以下では、すでに説明した各実施形態と異なる部分について詳細に説明し、その他の重複する部分については適宜説明を省略するものとする。

【0108】

（管理画面の表示）

まず、図20を参照しながら、第7実施形態に係る情報処理システム10による表示例（特に、管理画面の表示例）について説明する。図20は、第7実施形態に係る情報処理システム10による表示例を示す平面図である。

【0109】

図20に示すように、第7実施形態に係る情報処理システム10は、管理画面（例えば、システム管理者等が見る画面）に、会話データのファイル名と、会話データから生成されたキーワード（即ち、照合用情報として紐付けられているキーワード）と、が一覧表示される。なお、管理画面は、例えば上述した出力装置16を用いて表示されるものであってよい。

【0110】

図20の例では、1つ目の「20210115_meeting.txt」というファイルに、「会議」、「予算」、「新規」というキーワードが紐付けられている。2つ目の「20210303_meeting.txt」というファイルに、「来季」、「年度」、「実行」というキーワードが紐付けられている。3つ目の「20210310_meeting.txt」というファイルに、「指示」、「予算」、「決定」というキーワードが紐付けられている。なお、図20では、3つのファイルについて一覧表示する例を挙げているが、より多くのファイルについて一覧表示されてもよい。また、すべてのファイルが画面に収まらない場合には、スクロール可能な態様で表

10

20

30

40

50

示されてもよいし、複数のページに分けて表示されてもよい。

【0111】

(技術的効果)

次に、第7実施形態に係る情報処理システム10によって得られる技術的効果について説明する。

【0112】

図20で説明したように、第7実施形態に係る情報処理システム10によれば、管理画面において、ファイル名とキーワードが一覧形式で表示される。このようにすれば、システム管理者等に対して、どの会話データにどのようなキーワードが紐付けられているかを分かりやすく提示することが可能である。

【0113】

上述した各実施形態の機能を実現するように該実施形態の構成を動作させるプログラムを記録媒体に記録させ、該記録媒体に記録されたプログラムをコードとして読み出し、コンピュータにおいて実行する処理方法も各実施形態の範疇に含まれる。すなわち、コンピュータ読取可能な記録媒体も各実施形態の範囲に含まれる。また、上述のプログラムが記録された記録媒体はもちろん、そのプログラム自体も各実施形態に含まれる。

【0114】

記録媒体としては例えばフロッピー（登録商標）ディスク、ハードディスク、光ディスク、光磁気ディスク、CD-ROM、磁気テープ、不揮発性メモリカード、ROMを用いることができる。また該記録媒体に記録されたプログラム単体で処理を実行しているものに限らず、他のソフトウェア、拡張ボードの機能と共同して、OS上で動作して処理を実行するものも各実施形態の範疇に含まれる。更に、プログラム自体がサーバに記憶され、ユーザ端末にサーバからプログラムの一部または全てをダウンロード可能なようにしてもよい。

【0115】

<付記>

以上説明した実施形態に関して、更に以下の付記のようにも記載されうるが、以下には限られない。

【0116】

(付記1)

付記1に記載の情報処理システムは、複数人の音声情報を含む会話データを取得する取得手段と、前記音声情報からキーワードを抽出するキーワード抽出手段と、前記音声情報から前記キーワードを発話した際の音声に関する特徴量である第1特徴量を抽出する特徴量抽出手段と、前記キーワードと前記第1特徴量とを関連付けた照合用情報を生成する生成手段と、を備える情報処理システムである。

【0117】

(付記2)

付記2に記載の情報処理システムは、前記複数人のうち少なくとも1人の音声に関する特徴量である第2特徴量を取得する特徴量取得手段と、前記第1特徴量と前記第2特徴量とを比較することで、前記第1特徴量から前記キーワードを発話した話者を特定できるか否かを判定する判定手段と、を更に備える付記1に記載の情報処理システムである。

【0118】

(付記3)

付記3に記載の情報処理システムは、前記会話データに対する所定処理を要求するユーザに対して、前記照合用情報が生成された前記キーワードの発話を促す情報を提示する提示手段と、前記ユーザの発話内容から、前記ユーザの音声に関する特徴量である第3特徴量を抽出する認証用特徴量抽出手段と、発話を促した前記キーワードに関連付いた前記第1特徴量と、前記第3特徴量との比較結果に基づいて、前記ユーザによる前記所定処理の実行を許可するか否かを判定する許可判定手段と、を更に備える付記1又は2に記載の情報処理システムである。

【0119】

(付記4)

付記4に記載の情報処理システムは、前記照合用情報は複数の前記キーワードについて生成されており、前記提示手段は、一部の前記キーワードの発話を促す情報を提示して、前記ユーザによる前記所定処理の実行を許可しないと判定された場合に、他の前記キーワードの発話を促す情報を提示する、付記3に記載の情報処理システムである。

【0120】

(付記5)

付記5に記載の情報処理装置は、複数人の音声情報を含む会話データを取得する取得手段と、前記音声情報からキーワードを抽出するキーワード抽出手段と、前記音声情報から前記キーワードを発話した際の音声に関する特徴量である第1特徴量を抽出する特徴量抽出手段と、前記キーワードと前記第1特徴量とを関連付けた照合用情報を生成する生成手段と、を備える情報処理装置である。

【0121】

(付記6)

付記6に記載の情報処理方法は、少なくとも1つのコンピュータが実行する情報処理方法であって、複数人の音声情報を含む会話データを取得し、前記音声情報からキーワードを抽出し、前記音声情報から前記キーワードを発話した際の音声に関する特徴量である第1特徴量を抽出し、前記キーワードと前記第1特徴量とを関連付けた照合用情報を生成する、情報処理方法である。

【0122】

(付記7)

付記7に記載の記録媒体は、少なくとも1つのコンピュータに、複数人の音声情報を含む会話データを取得し、前記音声情報からキーワードを抽出し、前記音声情報から前記キーワードを発話した際の音声に関する特徴量である第1特徴量を抽出し、前記キーワードと前記第1特徴量とを関連付けた照合用情報を生成する、情報処理方法を実行させるコンピュータプログラムが記録された記録媒体である。

【0123】

(付記8)

付記8に記載のコンピュータプログラムは、少なくとも1つのコンピュータに、複数人の音声情報を含む会話データを取得し、前記音声情報からキーワードを抽出し、前記音声情報から前記キーワードを発話した際の音声に関する特徴量である第1特徴量を抽出し、前記キーワードと前記第1特徴量とを関連付けた照合用情報を生成する、情報処理方法を実行させるコンピュータプログラムである。

【0124】

この開示は、請求の範囲及び明細書全体から読み取ることのできる発明の要旨又は思想に反しない範囲で適宜変更可能であり、そのような変更を伴う情報処理システム、情報処理装置、情報処理方法、及び記録媒体もまたこの開示の技術思想に含まれる。

【符号の説明】

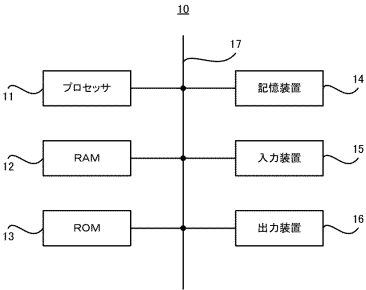
【0125】

- 10 情報処理システム
- 11 プロセッサ
- 110 会話データ取得部
- 120 キーワード抽出部
- 130 特徴量抽出部
- 140 照合用情報生成部
- 150 特徴量取得部
- 160 利用可否判定部
- 210 照合用情報記憶部
- 220 キーワード提示部

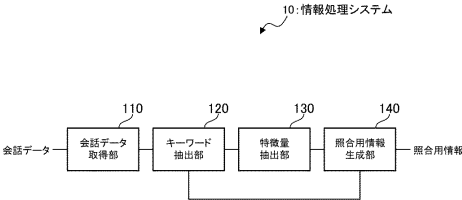
- 2 3 0 認証用特微量抽出部
- 2 4 0 許可判定部
- 2 5 0 キーワード変更部
- 5 0 0 端末

【図面】

【図 1】



【図 2】



10

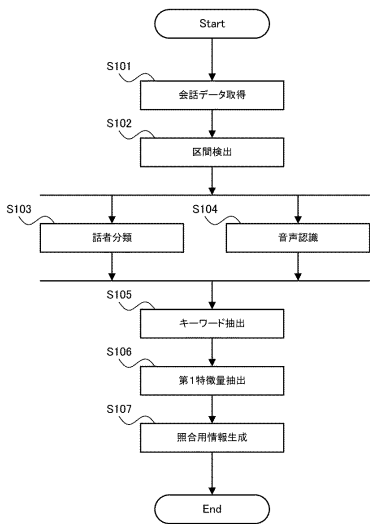
20

30

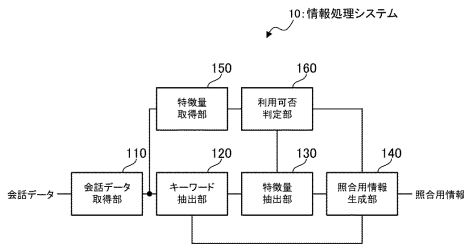
40

50

【図 3】



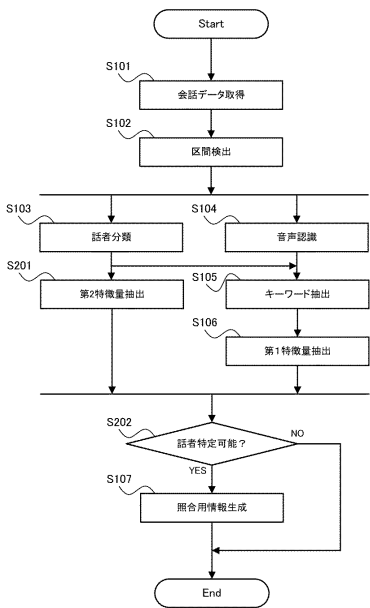
【図 4】



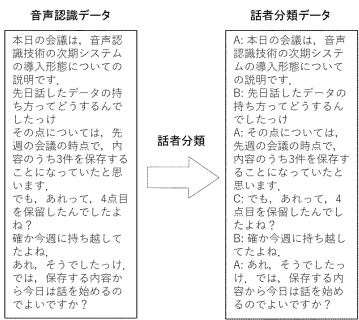
10

20

【図 5】



【図 6】

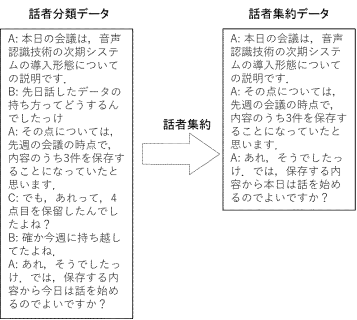


30

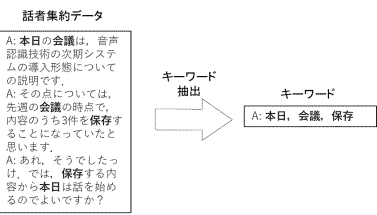
40

50

【図 7】



【図 8】



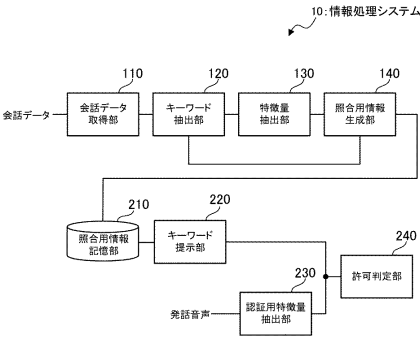
10

20

【図 9】

ユーザ	キーワード
A	本日、会議、保存
B	〇〇、●●
C	△△、▲▲
D	□□、■

【図 10】

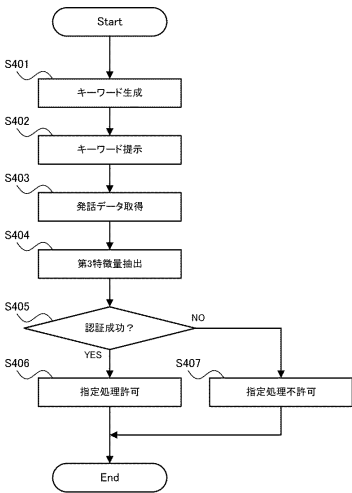


30

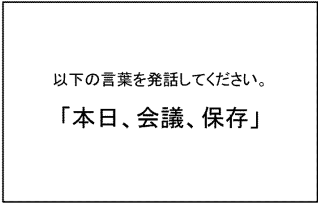
40

50

【図 1 1】



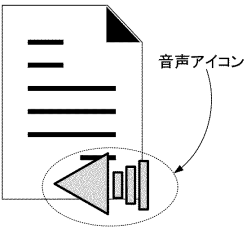
【図 1 2】



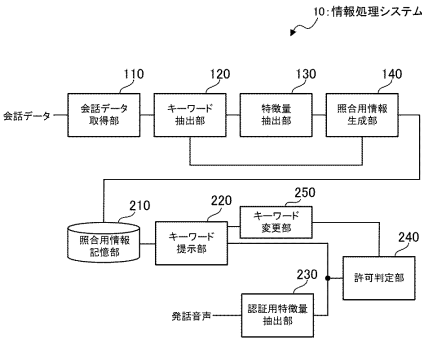
10

20

【図 1 3】



【図 1 4】

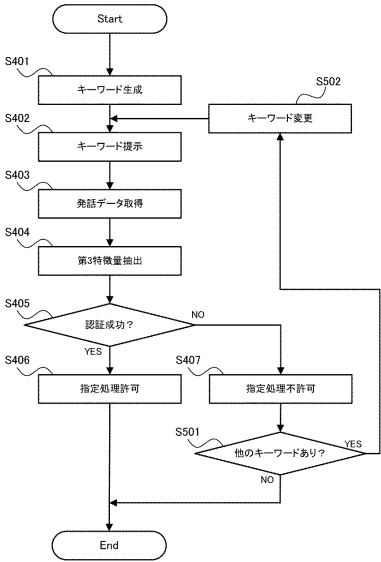


30

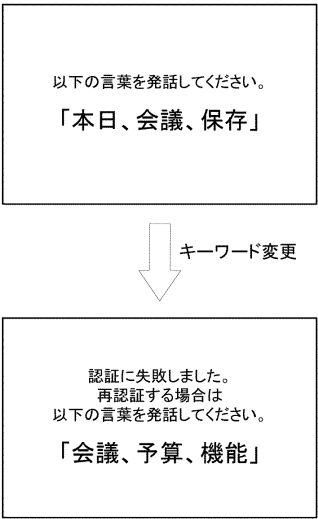
40

50

【図 15】



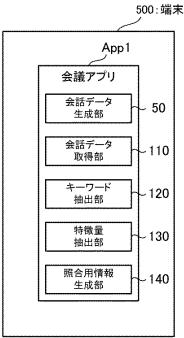
【図 16】



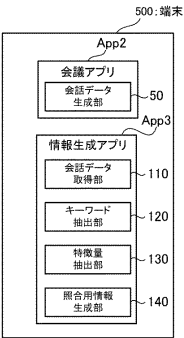
10

20

【図 17】



【図 18】

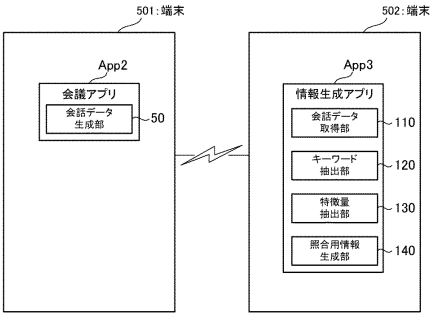


30

40

50

【図 19】



【図 20】

ファイル名	キーワード
20210115_meeting.txt	会議、予算、新規
20210303_meeting.txt	来季、年度、実行
202103010_meeting.txt	指示、予算、決定

10

20

30

40

50

フロントページの続き

(56)参考文献 韓国登録特許第2113879 (KR, B1)
特開2016-206428 (JP, A)
米国特許第08543834 (US, B1)
特開2004-294755 (JP, A)

(58)調査した分野 (Int.Cl., DB名)
G10L 17/02
G10L 15/10
G10L 17/00
G10L 17/24
G10L 17/08