

(19)日本国特許庁(JP)

(12)公開特許公報(A)

(11)公開番号
特開2022-17568
(P2022-17568A)

(43)公開日 令和4年1月25日(2022.1.25)

(51)国際特許分類

F I

G 1 0 L 15/04 (2013.01) G 1 0 L 15/04 2 0 0

G 1 0 L 15/28 (2013.01) G 1 0 L 15/28 5 0 0

審査請求 有 請求項の数 7 O L (全20頁)			
(21)出願番号	特願2021-183893(P2021-183893)	(71)出願人	000004075
(22)出願日	令和3年11月11日(2021.11.11)		ヤマハ株式会社
(62)分割の表示	特願2021-4411(P2021-4411)の分割	(74)代理人	静岡県浜松市中区中沢町 1 0 番 1 号
原出願日	平成29年3月15日(2017.3.15)		110003177
			特許業務法人旺知国際特許事務所
		(72)発明者	岩田 貴裕
			静岡県浜松市中区中沢町 1 0 番 1 号 ヤマハ株式会社内
		(72)発明者	瀬戸 優樹
			静岡県浜松市中区中沢町 1 0 番 1 号 ヤマハ株式会社内
		(72)発明者	大野 康法
			静岡県浜松市中区中沢町 1 0 番 1 号 ヤマハ株式会社内
		(72)発明者	森口 翔太
最終頁に続く			

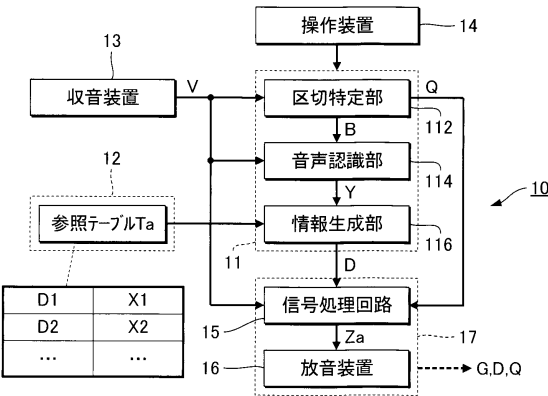
(54)【発明の名称】 情報提供方法、情報提供システムおよびプログラム

(57)【要約】

【課題】音声に対する情報提供の遅延を低減する。

【解決手段】情報提供システム 1 0 は、案内音声 G の収音に並行して当該案内音声 G の時間軸上の区切 B を順次に特定する区切特定部 1 1 2 と、区切特定部 1 1 2 による区切 B の特定毎に、記憶装置 1 2 に記憶された複数の登録文字列 X のうち、当該区切 B の直前の区間に対する音声認識により特定された認識文字列 Y に類似する登録文字列 X に対応する配信情報 D を生成する情報生成部 1 1 6 と、配信情報 D を送信する情報送信部 1 7 とを具備する。

【選択図】図 3



【特許請求の範囲】**【請求項 1】**

コンピュータシステムが、
音声の収音に並行して当該音声の時間軸上の区切を特定し、
前記区切の特定毎に、記憶装置に記憶された複数の登録文字列のうち、当該区切の直前の区間に対する音声認識により特定された認識文字列に類似する登録文字列に対応する配信情報を生成し、
前記配信情報を送信する
情報提供方法。

【請求項 2】

前記区切の特定毎に、当該区切が前記音声の途中の時点である場合に、当該途中の時点であることを示す区切情報を送信する
請求項 1 の情報提供方法。

【請求項 3】

前記区切の特定毎に、当該区切が前記音声の末尾の時点である場合に、当該末尾の時点であることを示す区切情報を送信する
請求項 1 または請求項 2 の情報提供方法。

【請求項 4】

前記配信情報の送信においては、前記配信情報の生成毎に当該配信情報を送信する
請求項 1 から請求項 3 の何れかの情報提供方法。

【請求項 5】

前記配信情報の送信においては、前記配信情報の生成毎に、当該配信情報と、前記音声について過去に生成した配信情報とを送信する
請求項 4 の情報提供方法。

【請求項 6】

音声の収音に並行して当該音声の時間軸上の区切を特定する区切特定部と、
前記区切特定部による前記区切の特定毎に、記憶装置に記憶された複数の登録文字列のうち、当該区切の直前の区間に対する音声認識により特定された認識文字列に類似する登録文字列に対応する配信情報を生成する情報生成部と、
前記配信情報を送信する情報送信部と
を具備する情報提供システム。

【請求項 7】

音声の収音に並行して当該音声の時間軸上の区切を特定する区切特定部、および、
前記区切特定部による前記区切の特定毎に、記憶装置に記憶された複数の登録文字列のうち、当該区切の直前の区間に対する音声認識により特定された認識文字列に類似する登録文字列に対応する配信情報を生成する情報生成部、
としてコンピュータを機能させるプログラム。

【発明の詳細な説明】**【技術分野】****【0001】**

本発明は、情報を提供する技術に関する。

【背景技術】**【0002】**

音声が表示する文字列を特定する音声認識技術が従来から提案されている。例えば特許文献 1 には、音節または単語等の認識単位の境界点を音声の特徴パラメータから検出する音声認識装置が開示されている。また、特許文献 2 には、音声から抽出された特徴ベクトルの時系列から単語列を特定する連続音声認識装置が開示されている。

【先行技術文献】**【特許文献】****【0003】**

10

20

30

40

50

【特許文献１】特開平 9 - 2 7 4 4 9 6 号公報

【特許文献２】特開 2 0 0 4 - 1 2 6 1 5 号公報

【発明の概要】

【発明が解決しようとする課題】

【 0 0 0 4 】

ところで、例えば交通機関または商業施設等の各種の施設で再生される音声について音声認識を実行し、認識結果に応じた情報を利用者の端末装置に提供するサービスが想定される。しかし、音声の終了により認識範囲が確定した段階で音声認識を開始し、認識結果に応じた情報を提供する構成では、情報の提供が音声の再生に対して遅延するという問題がある。以上の事情を考慮して、本発明の好適な態様は、音声に対する情報提供の遅延を低減することを目的とする。

【課題を解決するための手段】

【 0 0 0 5 】

以上の課題を解決するために、本発明の好適な態様に係る情報提供方法は、コンピュータシステムが、音声の收音に並行して当該音声の時間軸上の区切を特定し、前記区切の特定毎に、記憶装置に記憶された複数の登録文字列のうち、当該区切の直前の区間に対する音声認識により特定された認識文字列に類似する登録文字列に対応する配信情報を生成し、前記配信情報を送信する。

本発明の好適な態様に係る情報提供システムは、音声の收音に並行して当該音声の時間軸上の区切を特定する区切特定部と、前記区切特定部による区切の特定毎に、記憶装置に記憶された複数の登録文字列のうち、当該区切の直前の区間に対する音声認識により特定された認識文字列に類似する登録文字列に対応する配信情報を生成する情報生成部と、前記配信情報を送信する情報送信部とを具備する。

本発明の好適な態様に係るプログラムは、音声の收音に並行して当該音声の時間軸上の区切を特定する区切特定部、および、前記区切特定部による区切の特定毎に、記憶装置に記憶された複数の登録文字列のうち、当該区切の直前の区間に対する音声認識により特定された認識文字列に類似する登録文字列に対応する配信情報を生成する情報生成部、としてコンピュータを機能させる。

【図面の簡単な説明】

【 0 0 0 6 】

【図 1】第 1 実施形態に係る情報提供システムが使用される状況の説明図である。

【図 2】案内音声の説明図である。

【図 3】情報提供システムの構成図である。

【図 4】情報生成部の動作を示すフローチャートである。

【図 5】信号処理回路の構成図である。

【図 6】情報提供システムの動作を示すフローチャートである。

【図 7】端末装置の構成図である。

【図 8】端末装置の動作を示すフローチャートである。

【図 9】第 2 実施形態の動作の説明図である。

【図 10】第 3 実施形態における参照テーブルの模式図である。

【図 11】第 3 実施形態における配信情報および区間情報の説明図である。

【図 12】第 3 実施形態の動作の説明図である。

【図 13】変形例における関連情報の表示例である。

【発明を実施するための形態】

【 0 0 0 7 】

< 第 1 実施形態 >

図 1 は、本発明の第 1 実施形態に係る情報提供システム 10 が使用される状況の説明図である。第 1 実施形態の情報提供システム 10 は、店舗またはショッピングモール等の商業施設内の利用者 U b に情報を提供するコンピュータシステムである。利用者 U b は、端末装置 20 を携帯する。端末装置 20 は、例えば携帯電話機またはスマートフォン等の可搬

10

20

30

40

50

型の情報端末である。例えば、商業施設内に設置される電光掲示板または電子看板（例えばデジタルサイネージ）等の案内用の表示端末を端末装置 20 として利用することも可能である。なお、実際には多数の利用者 U b が情報提供システム 10 のサービスを利用し得るが、以下の説明では便宜的に 1 個の端末装置 20 に着目する。

【0008】

商業施設の従業者等の案内者 U a は、商業施設内の利用者 U b を案内するための音声（以下「案内音声」という）G を発声する。図 2 に例示される通り、案内音声 G は、複数の区間 S（S 1, S 2, ...）を含んで構成される。複数の区間 S は、例えば案内者 U a が一連の案内として順次に発声する複数の文にそれぞれ対応する。図 2 に例示された案内音声 G は、「お客様にご案内申し上げます」という区間 S 1 と、「当店は間もなく閉店いたします」という区間 S 2 と、「ご来店ありがとうございました」という区間 S 3 とで構成される。第 1 実施形態では、案内者 U a は、事前に用意された複数の文字列（以下「登録文字列」という）の何れかを選択的に発声する。各登録文字列は、案内音声 G の 1 個の区間 S に対応する。案内者 U a は、複数の登録文字列から選択された 2 個以上の登録文字列の時系列を案内音声 G として発声する。

10

【0009】

図 3 は、情報提供システム 10 の機能に着目した構成図である。図 3 に例示される通り、第 1 実施形態の情報提供システム 10 は、制御装置 11 と記憶装置 12 と收音装置 13 と操作装置 14 と信号処理回路 15 と放音装置 16 とを具備する。例えばタブレット端末またはパーソナルコンピュータ等の情報端末が情報提供システム 10 として利用され得る。なお、例えば商業施設内に設置される電光掲示板または電子看板（例えばデジタルサイネージ）等の案内用の表示端末を情報提供システム 10 として利用することも可能である。また、情報提供システム 10 は、単体の装置で実現されるほか、相互に別体で構成された複数の装置でも実現され得る。

20

【0010】

收音装置 13 は、周囲の音響を收音する音響機器（マイクロホン）である。案内者 U a は、收音装置 13 に対して案内音声 G を発声する。收音装置 13 は、案内者 U a が発声した案内音声 G を收音し、当該案内音声 G を表す音声信号 V を生成する。なお、收音装置 13 が生成した音声信号 V をアナログからデジタルに変換する A/D 変換器の図示は便宜的に省略した。操作装置 14 は、利用者 U b からの操作を受け付ける入力機器である。放音装置 16 は、制御装置 11 から指示された音を再生する。

30

【0011】

制御装置 11 は、例えば CPU（Central Processing Unit）等の処理回路で構成され、情報提供システム 10 の各要素を統括的に制御する。記憶装置 12 は、制御装置 11 が実行するプログラムと制御装置 11 が使用する各種のデータとを記憶する。例えば半導体記録媒体および磁気記録媒体等の公知の記録媒体、または複数種の記録媒体の組合せが、記憶装置 12 として任意に採用され得る。なお、情報提供システム 10 とは別体で記憶装置 12（例えばクラウドストレージ）を用意し、移動体通信網またはインターネット等の通信網を介して制御装置 11 が記憶装置 12 に対する読出および書込を実行することも可能である。

40

【0012】

第 1 実施形態の記憶装置 12 は参照テーブル T a を記憶する。図 3 に例示される通り、参照テーブル T a は、複数の登録文字列 X（X 1, X 2, ...）の各々について配信情報 D（D 1, D 2, ...）を対応させたデータテーブルである。各配信情報 D は、登録文字列 X を識別するための識別情報である。

【0013】

第 1 実施形態の制御装置 11 は、記憶装置 12 に記憶されたプログラムを実行することで、商業施設の利用者 U b に情報を提供するための複数の機能（区切特定部 112, 音声認識部 114 および情報生成部 116）を実現する。なお、制御装置 11 の一部の機能を専用の電子回路で実現した構成、または、制御装置 11 の機能を複数の装置に分散した構成

50

も採用され得る。

【 0 0 1 4 】

区切特定部 1 1 2 は、收音装置 1 3 が生成した音声信号 V が表す案内音声 G について時間軸上の区切 B を特定する。区切 B は、図 2 に例示される通り、案内音声 G を構成する各区間 S の終点（すなわち案内音声 G が表す各文の境界点）であり、相前後する 2 個の区間 S の境界点とも換言され得る。区切特定部 1 1 2 による区切 B の特定は、收音装置 1 3 による案内音声 G の收音に並行して（すなわち案内音声 G の終了前に）順次に行われる。

【 0 0 1 5 】

一般的な発話においては、相前後する 2 個の文の間に、低音量または無音の区間が存在するという傾向がある。以上の傾向を考慮して、第 1 実施形態の区切特定部 1 1 2 は、音声信号 V が表す案内音声 G の音量が低下した時点を区切 B として特定する。例えば、区切特定部 1 1 2 は、各区間 S に想定される時間長と比較して十分に短い周期で音声信号 V の音量を順次に算定する。そして、区切特定部 1 1 2 は、音量が閾値を下回る期間が所定の時間にわたり継続した場合に、当該期間内の時点（例えば始点または終点）を区切 B として特定する。

【 0 0 1 6 】

また、第 1 実施形態の区切特定部 1 1 2 は、区切 B の特定毎に区切情報 Q を生成する。任意の 1 個の区切 B に対応する区切情報 Q は、当該区切 B が案内音声 G の途中の時点であるか案内音声 G の末尾の時点であることを示す情報（例えばフラグ）である。具体的には、区切特定部 1 1 2 は、1 個の区切 B から所定の時間内に音声信号 V の音量が増加した場合に、当該区切 B が案内音声 G の途中の時点であることを示す区切情報 Q を生成する。他方、特定済の 1 個の区切 B から所定の時間内に音声信号 V の音量が増加しない場合には、当該区切 B が案内音声 G の末尾の時点であることを示す区切情報 Q を生成する。

【 0 0 1 7 】

例えば図 2 に例示される通り、区切特定部 1 1 2 は、区間 S 1 と区間 S 2 との間の区切 B 1 を特定した段階では、当該区切 B 1 が案内音声 G の途中の時点であることを示す区切情報 Q 1 を生成する。同様に、区間 S 2 と区間 S 3 との間の区切 B 2 を特定した段階では、当該区切 B 2 が案内音声 G の途中の時点であることを示す区切情報 Q 2 が生成される。他方、区間 S 3 の区切 B 3 を特定した段階では、区切特定部 1 1 2 は、当該区切 B 3 が案内音声 G の末尾の時点であることを示す区切情報 Q 3 を生成する。

【 0 0 1 8 】

図 3 の音声認識部 1 1 4 は、音声信号 V に対する音声認識により認識文字列 Y を特定する。認識文字列 Y は、案内者 U a が発声した内容を表す文字列（すなわち音声認識の結果）である。第 1 実施形態の音声認識部 1 1 4 は、收音装置 1 3 による案内音声 G の收音に並行して認識文字列 Y を順次に特定する。具体的には、音声認識部 1 1 4 は、区切特定部 1 1 2 による区切 B の特定毎に、音声信号 V のうち当該区切 B の直前の区間 S に対する音声認識で認識文字列 Y を特定する。すなわち、案内音声 G の 1 個の区間 S を案内者 U a が発声するたびに、当該区間 S の発声の内容を表す認識文字列 Y が特定される。なお、音声信号 V の音声認識には、例えば H M M (Hidden Markov Model) 等の音響モデルと言語的な制約を示す言語モデルとを利用した公知の認識処理が任意に採用され得る。

【 0 0 1 9 】

なお、案内者 U a は、事前に用意された複数の登録文字列 X の何れかを発声するから、理想的には、音声認識部 1 1 4 が特定する認識文字列 Y は何れかの登録文字列 X と一致する。しかし、実際には、案内者 U a 毎の発話の特徴または施設内の背景雑音等の種々の事情に起因して、音声認識部 1 1 4 による解析には誤認識が発生し得る。したがって、認識文字列 Y と登録文字列 X とは、相互に類似するけれども一致しない場合がある。

【 0 0 2 0 】

図 3 の情報生成部 1 1 6 は、音声認識部 1 1 4 による音声認識の結果（すなわち認識文字列 Y）に応じた配信情報 D を生成する。情報生成部 1 1 6 による配信情報 D の生成は、区切特定部 1 1 2 による区切 B の特定毎（音声認識部 1 1 4 による認識文字列 Y の認識毎）

10

20

30

40

50

に順次に実行される。すなわち、收音装置 1 3 による案内音声 G の收音に並行して区間 S の発声毎に配信情報 D が順次に生成される。配信情報 D の生成には、記憶装置 1 2 に記憶された参照テーブル T a が利用される。

【 0 0 2 1 】

図 4 は、情報生成部 1 1 6 が配信情報 D を生成する処理のフローチャートである。音声認識部 1 1 4 が 1 個の区間 S について認識文字列 Y を特定するたびに図 4 の処理が開始される。図 4 の処理を開始すると、情報生成部 1 1 6 は、参照テーブル T a に登録された複数の登録文字列 X のうち認識文字列 Y に類似する登録文字列 X を検索する (S a 5 1)。具体的には、情報生成部 1 1 6 は、複数の登録文字列 X の各々について認識文字列 Y との類似度を算定し、複数の登録文字列 X のうち類似度が最大となる 1 個の登録文字列 X (すなわち認識文字列 Y に最も類似する登録文字列 X) を特定する。以上の説明から理解される通り、第 1 実施形態では、認識文字列 Y に類似する登録文字列 X が特定されるから、音声認識における誤認識の影響を含まない登録文字列 X が特定される。

10

【 0 0 2 2 】

以上の手順で登録文字列 X を特定すると、情報生成部 1 1 6 は、当該登録文字列 X に対応する配信情報 D を参照テーブル T a から検索する (S a 5 2)。すなわち、案内音声 G を構成する複数の区間 S の各々について、案内者 U a が当該区間 S について発声した登録文字列 X に対応する配信情報 D が生成される。

【 0 0 2 3 】

図 3 の信号処理回路 1 5 は、区切特定部 1 1 2 が生成した区切情報 Q と情報生成部 1 1 6 が生成した配信情報 D とを表す通知信号 Z a を生成する。図 5 は、信号処理回路 1 5 の構成図である。図 5 に例示される通り、第 1 実施形態の信号処理回路 1 5 は、変調処理部 1 5 2 と混合処理部 1 5 4 とを含んで構成される。なお、信号処理回路 1 5 の少なくとも一部の機能を制御装置 1 1 が実現することも可能である。

20

【 0 0 2 4 】

変調処理部 1 5 2 は、情報生成部 1 1 6 が生成した配信情報 D と区切特定部 1 1 2 が生成した区切情報 Q とを示す音響成分を表す変調信号 M を生成する。具体的には、変調処理部 1 5 2 は、例えば所定の周波数の搬送波を配信情報 D および区切情報 Q により変調する周波数変調、または、拡散符号を利用した配信情報 D および区切情報 Q の拡散変調等の変調処理により変調信号 M を生成する。配信情報 D および区切情報 Q を表す音響成分の周波数帯域は、例えば、放音装置 1 6 による再生が可能な周波数帯域であり、かつ、利用者 U b が通常的环境中で聴取する音の周波数帯域を上回る範囲 (例えば 1 8 k H z 以上かつ 2 0 k H z 以下) である。

30

【 0 0 2 5 】

図 5 の混合処理部 1 5 4 は、收音装置 1 3 から供給される音声信号 V と変調処理部 1 5 2 が生成した変調信号 M とを混合 (例えば加算) することで通知信号 Z a を生成する。混合処理部 1 5 4 が生成した通知信号 Z a が放音装置 1 6 に供給される。放音装置 1 6 は、通知信号 Z a が表す音を放音する。したがって、音声信号 V が表す案内音声 G と変調信号 M が表す配信情報 D および区切情報 Q の音響成分とが放音装置 1 6 から再生される。すなわち、第 1 実施形態の放音装置 1 6 は、案内者 U a が発声した案内音声 G を再生する音響機器として機能するほか、空気振動としての音波を伝送媒体とした音響通信で配信情報 D および区切情報 Q を送信する送信機としても機能する。なお、通知信号 Z a をデジタルからアナログに変換する D / A 変換器の図示は便宜的に省略した。

40

【 0 0 2 6 】

前述の通り、配信情報 D および区切情報 Q は、区切特定部 1 1 2 による区切 B の特定毎に生成される。したがって、配信情報 D と区切情報 Q とは、收音装置 1 3 による案内音声 G の收音に並行して区切 B の特定毎に送信される。以上の説明から理解される通り、信号処理回路 1 5 および放音装置 1 6 は、配信情報 D の生成毎 (すなわち区切 B の特定毎) に配信情報 D と区切情報 Q とを送信する情報送信部 1 7 として機能する。

【 0 0 2 7 】

50

図 6 は、情報提供システム 10 が実行する動作（以下「情報提供処理」という）を例示するフローチャートである。情報提供処理（情報提供方法の例示）は、例えば操作装置 14 に対する案内者 Ua からの指示を契機として開始され、案内者 Ua による案内音声 G の発声に並行して実行される。

【0028】

区切特定部 112 は、案内音声 G の区切 B を特定するための処理（以下「区切特定処理」という）を実行する（Sa1）。区切特定部 112 が実際に区切 B を特定するまで区切特定処理 Sa1 は反復される（Sa2：NO）。区切特定処理 Sa1 により区切 B を特定すると（Sa2：YES）、区切特定部 112 は、当該区切 B について区切情報 Q を生成する（Sa3）。

10

【0029】

音声認識部 114 は、音声信号 V のうち区切特定部 112 が特定した区切 B の直前の区間 S に対する音声認識で認識文字列 Y を特定する（Sa4）。情報生成部 116 は、図 4 を参照して説明した通り、音声認識部 114 が特定した認識文字列 Y に応じた配信情報 D を生成する（Sa5）。すなわち、情報生成部 116 は、認識文字列 Y に類似する登録文字列 X を参照テーブル Ta から検索し（Sa51）、当該登録文字列 X に対応する配信情報 D を参照テーブル Ta から特定する（Sa52）。そして、情報送信部 17 は、情報生成部 116 が生成した配信情報 D と区切特定部 112 が生成した区切情報 Q とを送信する（Sa6）。具体的には、配信情報 D および区切情報 Q を音響成分として含む通知信号 Za を信号処理回路 15 が生成し、通知信号 Za が表す音を放音装置 16 が放音する。

20

【0030】

制御装置 11 は、情報提供処理を終了すべきか否かを判定する（Sa7）。例えば案内音声 G が終了した場合、または、操作装置 14 に対する操作で案内者 Ua が終了を指示した場合、制御装置 11 は情報提供処理を終了する（Sa7：YES）。他方、情報提供処理を継続すべき場合（Sa7：NO）、制御装置 11 は処理をステップ Sa1 に移行して、同様の処理（Sa1 - Sa7）を反復する。すなわち、收音装置 13 による案内音声 G の收音に並行して、区切情報 Q の生成（Sa3）と区間 S の音声認識（Sa4）と配信情報 D の生成（Sa5）と配信情報 D および区切情報 Q の送信（Sa6）とが、区切 B の特定毎に順次に反復される。以上の手順で情報提供システム 10 から送信された配信情報 D および区切情報 Q は、商業施設内の利用者 Ub が携帯する端末装置 20 により受信される。

30

【0031】

図 7 は、端末装置 20 の構成図である。図 7 に例示される通り、端末装置 20 は、制御装置 21 と記憶装置 22 と收音装置 23 と表示装置 24 とを具備する。收音装置 23 は、周囲の音を收音する音響機器（マイクロホン）である。具体的には、收音装置 23 は、情報提供システム 10 の放音装置 16 による再生音を收音して通知信号 Zb を生成する。通知信号 Zb は、配信情報 D および区切情報 Q の音響成分を含み得る。以上の説明から理解される通り、收音装置 23 は、端末装置 20 の相互間の音声通話または動画撮影時の音声収録に利用されるほか、空気振動としての音波を伝送媒体とする音響通信で配信情報 D および区切情報 Q を受信する受信機としても機能する。表示装置 24（例えば液晶表示パネル）は、制御装置 21 による制御のもとで各種の画像を表示する。

40

【0032】

制御装置 21 は、例えば CPU 等の処理回路で構成され、端末装置 20 の各要素を統括的に制御する。記憶装置 22 は、制御装置 21 が実行するプログラムと制御装置 21 が使用する各種のデータとを記憶する。例えば半導体記録媒体および磁気記録媒体等の公知の記録媒体、または複数種の記録媒体の組合せが、記憶装置 22 として任意に採用され得る。なお、端末装置 20 とは別体で記憶装置 22（例えばクラウドストレージ）を用意し、制御装置 21 が通信網を介して記憶装置 22 に対する読出および書込を実行することも可能である。

【0033】

第 1 実施形態の記憶装置 22 は、図 7 に例示された参照テーブル Tb を記憶する。参照テ

50

ーブル T b は、情報提供システム 10 から送信され得る複数の配信情報 D (D 1 , D 2 , ...) の各々について関連情報 C (C 1 , C 2 , ...) を対応させたデータテーブルである。任意の 1 個の配信情報 D に対応する関連情報 C は、当該配信情報 D が示す登録文字列 X に関連する情報である。例えば登録文字列 X 自体、または登録文字列 X を他言語に翻訳した文字列が、関連情報 C の好適例である。登録文字列 X と関連情報 C とは 1 対 1 に対応するから、配信情報 D は、関連情報 C を示す情報 (例えば関連情報 C の識別情報) とも換言され得る。

【 0 0 3 4 】

制御装置 21 は、記憶装置 22 に記憶されたプログラムを実行することで、情報提供システム 10 が送信した配信情報 D および区切情報 Q に関する処理を実行するための複数の機能 (情報抽出部 212 および提示制御部 214) を実現する。なお、制御装置 21 の一部の機能を専用の電子回路で実現した構成、または、制御装置 21 の機能を複数の装置に分散した構成も採用され得る。

10

【 0 0 3 5 】

情報抽出部 212 は、收音装置 23 が生成した通知信号 Z b から配信情報 D および区切情報 Q を抽出する。具体的には、情報抽出部 212 が通知信号 Z b から配信情報 D を抽出するための処理 (以下「抽出処理」という) は、例えば、通知信号 Z b のうち配信情報 D および区切情報 Q の音響成分を含む周波数帯域を強調するフィルタ処理と、前述の変調処理部 152 が実行する変調処理に対応した復調処理とを包含する。各区間 S に想定される時間長と比較して十分に短い周期で抽出処理が反復される。したがって、案内者 U a による案内音声 G の発声 (または放音装置 16 による案内音声 G の再生) に並行して、当該案内音声 G の区間 S 毎に、配信情報 D と区切情報 Q とが順次に抽出される。

20

【 0 0 3 6 】

提示制御部 214 は、表示装置 24 による情報の表示を制御する。第 1 実施形態の提示制御部 214 は、情報抽出部 212 が抽出した配信情報 D に対応する関連情報 C を表示装置 24 に表示させる。具体的には、提示制御部 214 は、参照テーブル T b に登録された複数の関連情報 C のうち情報抽出部 212 が抽出した配信情報 D に対応する関連情報 C を検索し、当該関連情報 C を表示装置 24 に表示させる。したがって、放音装置 16 による案内音声 G の再生に並行して、当該案内音声 G の各区間 S に対応する関連情報 C が表示装置 24 に表示される。

30

【 0 0 3 7 】

第 1 実施形態の提示制御部 214 は、情報抽出部 212 が抽出した区切情報 Q に応じて関連情報 C の表示を制御する。具体的には、区切 B が案内音声 G の末尾の時点であることを示す区切情報 Q の抽出までは、提示制御部 214 は、複数の関連情報 C を一連の情報として表示装置 24 の 1 画面内に並列に表示させる。例えば、図 2 の例示を想定すると、区切 B (B 1 , B 2) が案内音声 G の途中の時点であることを区切情報 Q 1 および区切情報 Q 2 の各々が示し、区切 B 3 が案内音声 G の末尾の時点であることを区切情報 Q 3 が示す。したがって、提示制御部 214 は、配信情報 D 1 に対応する関連情報 C 1 と配信情報 D 2 に対応する関連情報 C 2 と配信情報 D 3 に対応する関連情報 C 3 とを、案内音声 G に対応する一連の情報として表示装置 24 の 1 画面内に表示させる。他方、区切 B が案内音声 G の末尾の時点であることを区切情報 Q が示す場合、以後に抽出される配信情報 D に対応する関連情報 C は、表示済の関連情報 C とは別個の情報として、表示済の関連情報 C とは別画面に表示される。

40

【 0 0 3 8 】

図 8 は、端末装置 20 が実行する動作 (以下「情報取得処理」という) を例示するフローチャートである。記憶装置 22 に記憶されたプログラムの起動を契機として情報取得処理が開始される。

【 0 0 3 9 】

端末装置 20 の情報抽出部 212 は、收音装置 23 が生成した通知信号 Z b に対する抽出処理を実行する (S b 1)。情報抽出部 212 が実際に配信情報 D および区切情報 Q を抽

50

出するまで抽出処理は反復される（S b 2：NO）。抽出処理 S b 1により配信情報 D および区切情報 Q が抽出されると（S b 2：YES）、提示制御部 2 1 4 は、配信情報 D に対応する関連情報 C を表示装置 2 4 に表示させる（S b 3）。前述の通り、提示制御部 2 1 4 は、抽出処理で抽出された区切情報 Q に応じて、配信情報 D に対応する関連情報 C を、表示済の関連情報 C と同画面で表示するか別画面で表示するかを制御する。

【0040】

制御装置 2 1 は、情報取得処理を終了すべきか否かを判定する（S b 4）。例えば端末装置 2 0 の利用者 U b が終了を指示した場合（S b 4：YES）には、情報取得処理は終了する。他方、情報取得処理を継続すべき場合（S b 4：NO）、制御装置 2 1 は処理をステップ S b 1に移行して同様の処理（S b 1 - S b 4）を反復する。以上の説明から理解される通り、案内者 U a による案内音声 G の発声に並行して、配信情報 D および区切情報 Q の抽出処理（S b 1, S b 2）と関連情報 C の表示（S b 3）とが順次に反復される。したがって、図 2 に例示される通り、案内音声 G の区間 S 毎に関連情報 C が表示装置 2 4 に順次に表示される。

10

【0041】

以上に説明した通り、第 1 実施形態では、案内音声 G の收音に並行した区切 B の特定毎に、当該区切 B の直前の区間 S に対する音声認識の結果に応じた配信情報 D が順次に生成および送信される。すなわち、案内音声 G の終了前に配信情報 D の送信が開始される。したがって、例えば案内音声 G の終了を契機として音声信号 V に対する音声認識と配信情報 D の生成および送信とを開始する構成と比較して、案内音声 G に対する配信情報 D の送信（さらには関連情報 C の提示）の遅延を低減することが可能である。災害等の緊急事態の発生時には案内音声 G により利用者 U b に迅速に通知する必要がある。配信情報 D の送信の遅延を低減できる第 1 実施形態は、緊急事態の発生を案内音声 G により利用者 U b に通知する場合に特に好適である。

20

【0042】

また、第 1 実施形態では、区切 B が案内音声 G の途中の時点であるか末尾の時点であることを示す区切情報 Q が配信情報 D とともに情報提供システム 1 0 から送信される。したがって、情報提供システム 1 0 から順次に送信される複数の配信情報 D が相互に関連しているのか否かを端末装置 2 0 において把握する（さらには関連情報 C の表示を制御する）ことが可能である。

30

【0043】

< 第 2 実施形態 >

本発明の第 2 実施形態について説明する。なお、以下に例示する各構成において作用または機能が第 1 実施形態と同様である要素については、第 1 実施形態の説明で使用した符号を流用して各々の詳細な説明を適宜に省略する。

【0044】

第 1 実施形態（図 2）では、情報生成部 1 1 6 による配信情報 D の生成毎に当該配信情報 D を端末装置 2 0 に送信する構成を例示した。第 2 実施形態では、情報生成部 1 1 6 による配信情報 D の生成毎に、当該配信情報 D と、案内音声 G について過去に特定された配信情報 D とを、情報提供システム 1 0 から端末装置 2 0 に送信する。

40

【0045】

図 9 は、第 2 実施形態における情報提供システム 1 0 の動作の説明図である。案内音声 G の最初の区間 S 1 について区切特定部 1 1 2 が区切 B 1 を特定すると、情報生成部 1 1 6 は、第 1 実施形態と同様に、区間 S 1 の認識文字列 Y に応じた配信情報 D 1 を生成する。情報送信部 1 7 は、配信情報 D 1 と区切 B 1 の区切情報 Q 1 とを送信する。配信情報 D 1 は記憶装置 1 2 に格納される。

【0046】

案内音声 G の区間 S 2 について区切特定部 1 1 2 が区切 B 2 を特定すると、情報生成部 1 1 6 は、第 1 実施形態と同様に、区間 S 2 の認識文字列 Y に応じた配信情報 D 2 を生成する。第 2 実施形態の情報送信部 1 7 は、情報生成部 1 1 6 が新規に生成した配信情報 D 2

50

と、直前の区間 S 1 について生成済の配信情報 D 1 とを、区間 S 2 の区切情報 Q 2 とともに端末装置 2 0 に送信する。配信情報 D 2 は配信情報 D 1 とともに記憶装置 1 2 に格納される。

【 0 0 4 7 】

案内音声 G の区間 S 3 について区切特定部 1 1 2 が区切 B 3 を特定すると、情報生成部 1 1 6 は、区間 S 3 の認識文字列 Y に応じた配信情報 D 3 を生成する。第 2 実施形態の情報送信部 1 7 は、情報生成部 1 1 6 が新規に生成した配信情報 D 3 と、案内音声 G について過去に生成した配信情報 D 1 および配信情報 D 2 とを、区間 S 3 の区切情報 Q 3 とともに端末装置 2 0 に送信する。以上の説明から理解される通り、第 2 実施形態では、配信情報 D の生成毎に、当該配信情報 D と過去の配信情報 D とが端末装置 2 0 に送信される。

10

【 0 0 4 8 】

他方、端末装置 2 0 の情報抽出部 2 1 2 は、收音装置 2 3 が生成した通知信号 Z b から 1 個以上の配信情報 D と区切情報 Q とを抽出する。情報抽出部 2 1 2 が 1 個の配信情報 D を抽出した場合、提示制御部 2 1 4 は、当該配信情報 D に対応する関連情報 C を参照テーブル T b から特定して表示装置 2 4 に表示させる。また、情報抽出部 2 1 2 が複数の配信情報 D を抽出した場合、提示制御部 2 1 4 は、複数の配信情報 D の各々について関連情報 C を参照テーブル T b から特定し、複数の関連情報 C を表示装置 2 4 に表示させる。

【 0 0 4 9 】

第 2 実施形態においても第 1 実施形態と同様の効果が実現される。また、第 2 実施形態では、情報生成部 1 1 6 が新規に生成した配信情報 D とともに過去の配信情報 D が送信される。したがって、案内音声 G の 1 個の区間 S について受信に失敗した配信情報 D を、当該案内音声 G における以降の区間 S に対応する配信情報 D とともに受信し直すことが可能である。すなわち、案内音声 G に関する特定の関連情報 C を端末装置 2 0 において表示できない可能性が低減されるという利点がある。また、案内音声 G の発声の開始前から商業施設内にいる利用者 U b の端末装置 2 0 だけでなく、案内音声 G の発声の途中で商業施設に到着した利用者 U b の端末装置 2 0 にも、配信情報 D を配信して関連情報 C を表示させることが可能である。

20

【 0 0 5 0 】

< 第 3 実施形態 >

図 1 0 は、第 3 実施形態において端末装置 2 0 の記憶装置 2 2 に記憶される参照テーブル T b の模式図である。図 1 0 に例示される通り、第 3 実施形態の参照テーブル T b は、案内者 U a が発声し得る複数の案内音声 G の各々について配信情報 E (E 1 , E 2 , ...) と関連情報 R (R 1 , R 2 , ...) とを相互に対応させたデータテーブルである。任意の 1 個の案内音声 G に対応する関連情報 R は、例えば当該案内音声 G を任意の言語 (案内音声 G と同言語または他言語) で表現した文字列であり、時系列に配列された複数の部分 r (r 1 , r 2 , ...) で構成される。複数の部分 r は、案内音声 G の複数の区間 S にそれぞれ対応する。例えば、案内音声 G の任意の 1 個の区間 S に対応する部分 r は、当該区間 S を表す文字列自体、または当該文字列を他言語に翻訳した文字列である。また、任意の 1 個の案内音声 G に対応する配信情報 E は、案内音声 G (または当該案内音声 G に対応する関連情報 R) を識別するための識別情報である。配信情報 E は、関連情報 R を構成する複数の部分 r の組合せを識別するための情報とも換言され得る。

30

40

【 0 0 5 1 】

情報提供システム 1 0 の情報生成部 1 1 6 は、図 1 1 に例示される通り、区切特定部 1 1 2 による区切 B の特定毎 (音声認識部 1 1 4 による認識文字列 Y の認識毎) に配信情報 E と区間情報 F とを生成する。具体的には、情報生成部 1 1 6 は、複数の関連情報 R のうち認識文字列 Y に類似する登録文字列 X を含む関連情報 R の配信情報 E を生成する。例えば、登録文字列 X が図 2 の区間 S 1 の「お客様にご案内申し上げます」である場合、当該登録文字列 X を含む図 2 の案内音声 G の関連情報 R に対応した配信情報 E が生成される。

【 0 0 5 2 】

区間情報 F は、図 1 1 に例示される通り、関連情報 R を構成する複数の部分 r (または案

50

内音声 G を構成する複数の区間 S) にそれぞれ対応する複数の単位データ f (f 1 , f 2 , f 3) を含んで構成される。区間情報 F の複数の単位データ f のうち関連情報 R の 1 個の部分 r に対応する単位データ f は、関連情報 R の当該部分 r を端末装置 2 0 の利用者 U b に提示するか否かを示す情報 (例えばフラグ) である。例えば、区間情報 F の複数の単位データ f のうち、登録文字列 X が特定された区間 S に対応する単位データ f は数値 a 1 (許可) に設定され、登録文字列 X の未特定の区間 S に対応する単位データ f は数値 a 0 (禁止) に設定される。すなわち、区間情報 F は、案内音声 G の複数の区間 S の各々について登録文字列 X が特定されたか否か (案内者 U a が当該区間 S を発音したか否か、または、音声認識部 1 1 4 が当該区間の認識文字列 Y を特定したか否か) を示す情報である。図 1 1 に例示される通り、情報送信部 1 7 は、配信情報 E の生成毎 (すなわち区切 B の特定毎) に、配信情報 E と区間情報 F とを周囲に送信する。配信情報 E および区間情報 F の送信には、第 1 実施形態における配信情報 D および区切情報 Q の送信と同様に、音響通信が好適に利用される。

【 0 0 5 3 】

図 2 の例示と同様に、「お客様にご案内申し上げます」という区間 S 1 と「当店は間もなく閉店いたします」という区間 S 2 と「ご来店ありがとうございました」という区間 S 3 とで構成される案内音声 G を、案内者 U a が発声する場合を想定する。

【 0 0 5 4 】

図 1 2 に例示される通り、案内音声 G の最初の区間 S 1 の登録文字列 X が特定された段階 t 1 において、情報生成部 1 1 6 は、当該登録文字列 X を含む案内音声 G に対応した配信情報 E と、複数の単位データ f のうち区間 S 1 に対応する単位データ f 1 が数値 a 1 (許可) に設定された区間情報 F とを生成する。区間情報 F のうち区間 S 1 以外の各区間 S (S 2 , S 3) に対応する単位データ f は数値 a 0 (禁止) に設定される。他方、配信情報 E と区間情報 F とを情報送信部 1 7 から受信した端末装置 2 0 の提示制御部 2 1 4 は、配信情報 E が示す関連情報 R を参照テーブル T b から検索し、当該関連情報 R を構成する複数の部分 r のうち数値 a 1 の単位データ f に対応する部分 r を表示装置 2 4 に表示させる。すなわち、関連情報 R のうち「お客様にご案内申し上げます」という区間 S 1 に対応する部分 r 1 が表示装置 2 4 に表示される。関連情報 R のうち数値 a 0 の単位データ f に対応する部分 r 2 および部分 r 3 は表示されない。

【 0 0 5 5 】

案内音声 G の第 2 番目の区間 S 2 の登録文字列 X が特定された段階 t 2 では、情報生成部 1 1 6 は、区間 S 1 および区間 S 2 について特定された登録文字列 X を含む案内音声 G の配信情報 E と、区間情報 F とを生成する。区間情報 F の複数の単位データ f のうち、区間 S 1 に対応する単位データ f 1 と区間 S 2 に対応する単位データ f 2 とは数値 a 1 (許可) に設定され、区間 S 3 に対応する単位データ f 3 は数値 a 0 (禁止) に設定される。他方、配信情報 E と区間情報 F とを情報送信部 1 7 から受信した端末装置 2 0 の提示制御部 2 1 4 は、配信情報 E が示す関連情報 R の複数の部分 r のうち、数値 a 1 の単位データ f に対応する部分 r を表示装置 2 4 に表示させる。すなわち、関連情報 R のうち、「お客様にご案内申し上げます」という区間 S 1 に対応する部分 r 1 と、「当店は間もなく閉店いたします」という区間 S 2 に対応する部分 r 2 とが、表示装置 2 4 に表示される。関連情報 R のうち数値 a 0 の単位データ f 3 に対応する部分 r 3 は表示されない。

【 0 0 5 6 】

案内音声 G の最後の区間 S 3 の登録文字列 X が特定された段階 t 3 では、情報生成部 1 1 6 は、区間 S 1 から区間 S 3 について特定された登録文字列 X を含む案内音声 G の配信情報 E と、区間情報 F とを生成する。区間情報 F においては、区間 S 1 から区間 S 3 にそれぞれ対応する全部の単位データ f (f 1 ~ f 3) が数値 a 1 (許可) に設定される。他方、配信情報 E と区間情報 F とを情報送信部 1 7 から受信した端末装置 2 0 の提示制御部 2 1 4 は、配信情報 E が示す関連情報 R のうち数値 a 1 の単位データ f に対応する部分 r を表示装置 2 4 に表示させる。すなわち、関連情報 R のうち「お客様にご案内申し上げます」という区間 S 1 に対応する部分 r 1 と、「当店は間もなく閉店いたします」という区間

S 2 に対応する部分 r 2 と、「ご来店ありがとうございました」という区間 S 3 に対応する部分 r 3 と（すなわち関連情報 R の全部）が、表示装置 2 4 に表示される。

【 0 0 5 7 】

第 3 実施形態においても第 1 実施形態と同様の効果が実現される。また、第 3 実施形態では、関連情報 R を構成する複数の部分 r の各々について利用者 U b に対する提示の可否を示す区間情報 F が端末装置 2 0 に送信される。したがって、案内音声 G の発声の開始から商業施設内にいる利用者 U b だけでなく、案内音声 G の発声の途中で商業施設に到着した利用者 U b にも、案内音声 G のうち利用者 U b の到着前に発音された区間 S に対応する部分 r を含む関連情報 R を提示することが可能である。

【 0 0 5 8 】

なお、複数の部分 r の組合せが相違する複数の関連情報 R のなかには、例えば最初の部分 r 1 が共通するものがある。最初の部分 r 1 が複数の関連情報 R にわたり共通する場合、例えば案内音声 G の最初の区間 S 1 の登録文字列 X が特定された段階 t 1 では、関連情報 R を一意に特定できない可能性がある。情報生成部 1 1 6 は、登録文字列 X を含む複数の案内音声 G のうちの何れかに対応する配信情報 E を端末装置 2 0 に送信する。以上の説明から理解される通り、端末装置 2 0 に送信される配信情報 E が示す案内音声 G は、案内者 U a が実際に発声する案内音声 G とは相違する可能性がある。しかし、第 3 実施形態では、区切 B の特定毎に端末装置 2 0 に送信される区間情報 F により、関連情報 R の各部分 r の提示の可否が指定されるから、関連情報 R のうち案内者 U a が実際には発声しない区間 S に対応する部分 r は端末装置 2 0 において表示されない。配信情報 E は、区切 B の特定毎に更新される。したがって、最初の区間 S が発声された段階では、案内者 U a が実際に発声する予定の内容とは相違する案内音声 G を示す配信情報 E が生成されたとしても、案内音声 G の全部の区間 S が発声された段階では、当該案内音声 G に対応した適正な配信情報 E が生成される。

【 0 0 5 9 】

< 変形例 >

以上に例示した各態様は多様に変形され得る。具体的な変形の態様を以下に例示する。以下の例示から任意に選択された 2 個以上の態様は、相互に矛盾しない範囲で適宜に併合され得る。

【 0 0 6 0 】

(1) 前述の各形態では、案内音声 G の音量が低下した時点を区切 B として特定したが、区切特定部 1 1 2 が案内音声 G の区切 B を特定する方法は以上の例示に限定されない。例えば、操作装置 1 4 に対する操作で案内者 U a が指示した時点を区切特定部 1 1 2 が案内音声 G の区切 B として特定することも可能である。例えば操作装置 1 4 に所定の操作が付与された時点が区切 B として特定される。以上の構成によれば、案内者 U a が意図した時点を区切 B として設定できるという利点がある。

【 0 0 6 1 】

(2) 前述の各形態では、認識文字列 Y の認識毎に配信情報 D を生成および送信したが、配信情報 D の送信の契機は以上の例示に限定されない。例えば、認識文字列 Y が特定の語句（以下「登録語句」という）を含むことを契機として、情報送信部 1 7 が配信情報 D を送信することも可能である。例えば、記憶装置 1 2 には登録語句が事前に記憶される。音声認識部 1 1 4 は、認識結果である認識文字列 Y に登録語句が含まれるか否かを判定する。認識文字列 Y に登録語句が含まれると音声認識部 1 1 4 が判定した場合に、情報送信部 1 7 は、情報生成部 1 1 6 が生成した配信情報 D を端末装置 2 0 に送信する。以上の構成によれば、案内者 U a による登録語句の発声を契機として配信情報 D が端末装置 2 0 に送信されるから、配信情報 D の送信のタイミングまたは頻度を案内者 U a が適宜に調整できるという利点がある。

【 0 0 6 2 】

(3) 前述の各形態では、区切 B が案内音声 G の途中の時点であるか末尾の時点であるかを示す区切情報 Q を端末装置 2 0 に送信したが、区切 B が案内音声 G の途中の時点である

10

20

30

40

50

場合に限定して区切情報 Q を端末装置 20 に送信することも可能である。すなわち、区切 B が案内音声 G の末尾の時点である場合には区切情報 Q は送信されない。端末装置 20 の提示制御部 214 は、区切情報 Q を受信できた場合には区切 B が案内音声 G の途中の時点であると判定し、区切情報 Q を受信できない場合には区切 B が案内音声の末尾の時点であると判定する。

【0063】

同様に、区切 B が案内音声 G の末尾の時点である場合に限定して区切情報 Q を端末装置 20 に送信することも可能である。すなわち、区切 B が案内音声 G の途中の時点である場合には区切情報 Q は送信されない。端末装置 20 の提示制御部 214 は、区切情報 Q を受信できた場合には区切 B が案内音声 G の末尾の時点であると判定し、区切情報 Q を受信できない場合には区切 B が案内音声の途中の時点であると判定する。

10

【0064】

(4) 前述の各形態では、配信情報 D と区切情報 Q とを情報提供システム 10 から端末装置 20 に送信したが、区切情報 Q の送信は省略され得る。端末装置 20 の提示制御部 214 は、配信情報 D に対応した区切 B が案内音声 G の途中の時点であるか末尾の時点であるかを認識することなく、配信情報 D が示す関連情報 C を表示装置 24 に表示させる。例えば、図 13 に例示される通り、配信情報 D の受信毎に、当該配信情報 D が示す関連情報 C (C1, C2, C3) を順次に別画面で表示装置 24 に表示させることも可能である。

【0065】

(5) 前述の各形態では、情報提供システム 10 が音声認識部 114 を具備する構成を例示したが、情報提供システム 10 が通信可能な音声認識装置 (例えばウェブサーバ) に、音声信号 V の各区間 S に対する音声認識を実行させることも可能である。例えば、区切特定部 112 が区切 B を特定すると、音声信号 V のうち当該区切 B の直前の区間 S が音声認識装置に送信され、音声認識の結果である認識文字列 Y が音声認識装置から情報提供システム 10 に送信される。以上の説明から理解される通り、音声認識部 114 は情報提供システム 10 から省略され得る。

20

【0066】

(6) 前述の各形態では、登録文字列 X の識別情報 (または関連情報 C の識別情報) を配信情報 D として例示したが、配信情報 D の内容は以上の例示に限定されない。例えば、関連情報 C 自体を配信情報 D として情報提供システム 10 から端末装置 20 に送信することも可能である。以上の構成において、端末装置 20 の提示制御部 214 は、情報抽出部 212 が抽出した配信情報 D (すなわち関連情報 C) を表示装置 24 に表示させる。したがって、参照テーブル T b を端末装置 20 に保持する必要はない。

30

【0067】

音声認識部 114 が特定した認識文字列 Y を示す配信情報 D を情報生成部 116 が生成することも可能である。例えば、認識文字列 Y 自体が配信情報 D として端末装置 20 に送信される。また、認識文字列 Y を識別するための識別情報を配信情報 D として端末装置 20 に送信する構成も想定される。

【0068】

(7) 前述の各形態では、登録文字列 X が参照テーブル T a に事前に登録された構成を例示したが、定型句と 1 個以上の挿入句とから構成される文字列を登録文字列 X とすることも可能である。定型句は、複数の登録文字列 X について共用される定型的な文字列であり、挿入句は、定型句に挿入され得る文字列である。相異なる複数の挿入句の何れかが定型句に挿入されることで複数の登録文字列 X が構成される。情報生成部 116 は、音声認識部 114 による認識結果 (認識文字列 Y) から定型句と挿入句とを特定し、定型句および挿入句の組合せを示す配信情報 D を生成する。

40

【0069】

(8) 前述の各形態では、関連情報 C を表示装置 24 に表示したが、関連情報 C を端末装置 20 の利用者 U b に提示する方法は以上の例示に限定されない。例えば、関連情報 C が示す音声を再生することで関連情報 C を利用者 U b に提示することも可能である。例えば

50

、関連情報 C が音声を表す構成、または、関連情報 C が表す文字列に対する音声合成により音声を生成する構成が想定される。

【 0 0 7 0 】

(9) 前述の各形態では、端末装置 2 0 の記憶装置 2 2 に参照テーブル T b を保持したが、例えば移動体通信網またはインターネット等の通信網を介して端末装置 2 0 と通信する配信装置に参照テーブル T b を保持することも可能である。端末装置 2 0 は、配信情報 D を含む情報要求を配信装置に送信する。配信装置は、情報要求内の配信情報 D に対応する関連情報 C を参照テーブル T b から検索して要求元の端末装置 2 0 に送信する。端末装置 2 0 の提示制御部 2 1 4 は、配信装置から受信した関連情報 C を表示装置 2 4 に表示させる。以上の説明から理解される通り、参照テーブル T b を端末装置 2 0 の記憶装置 2 2 に記憶することは必須ではない。

10

【 0 0 7 1 】

(1 0) 前述の各形態では、情報提供システム 1 0 を商業施設で利用する場面を例示したが、情報提供システム 1 0 が利用される場面は以上の例示に限定されない。情報提供システム 1 0 が設置される施設としては、例えば、電車またはバス等の交通施設、旅館またはホテル等の宿泊施設、博物館または美術館等の展示施設、史跡または名所等の観光施設、および、競技場または体育館等の運動施設が例示され得る。また、前述の各形態では、利用者 U b に施設を案内するための案内音声 G を例示したが、区切 B が検出される音声は案内音声 G に限定されず、任意の音声について情報提供システム 1 0 を利用することが可能である。

20

【 0 0 7 2 】

(1 1) 前述の各形態では、音響を伝送媒体とする音響通信で配信情報 D および区切情報 Q を端末装置 2 0 に送信したが、端末装置 2 0 に配信情報 D を送信するための通信方式は音響通信に限定されない。例えば、電波または赤外線等の電磁波を伝送媒体とした無線通信で情報提供システム 1 0 から端末装置 2 0 に配信情報 D および区切情報 Q を送信することも可能である。すなわち、無線通信用の通信機器が情報送信部 1 7 として利用される。具体的には、Bluetooth (登録商標) またはWiFi (登録商標) 等の無線通信が配信情報 D の送信に好適である。

【 0 0 7 3 】

(1 2) 前述の各形態に係る情報提供システム 1 0 は、各形態での例示の通り、制御装置 1 1 とプログラムとの協働により実現される。前述の各形態に係るプログラムは、制御装置 1 1 (コンピュータの例示) に、案内音声 G の收音に並行して当該案内音声 G の時間軸上の区切 B を順次に特定する区切特定処理 (S a 1) と、区切特定処理による区切 B の特定毎に、当該区切 B の直前の区間 S に対する音声認識の結果 (認識文字列 Y) に応じた配信情報 D を順次に生成する情報生成処理 (S a 5) とを実行させる。

30

【 0 0 7 4 】

以上に例示したプログラムは、コンピュータが読取可能な記録媒体に格納された形態で提供されてコンピュータにインストールされ得る。記録媒体は、例えば非一過性 (non-transitory) の記録媒体であり、CD-ROM 等の光学式記録媒体 (光ディスク) が好例であるが、半導体記録媒体または磁気記録媒体等の公知の任意の形式の記録媒体を包含し得る。なお、非一過性の記録媒体とは、一過性の伝搬信号 (transitory, propagating signal) を除く任意の記録媒体を含み、揮発性の記録媒体を除外するものではない。また、通信網を介した配信の形態でプログラムをコンピュータに提供することも可能である。

40

【 0 0 7 5 】

(1 3) 以上に例示した形態から、例えば以下の構成が把握される。

< 態様 1 >

本発明の好適な態様 (態様 1) に係る情報提供方法は、コンピュータシステムが、音声の收音に並行して当該音声の時間軸上の区切を順次に特定し、前記区切の特定毎に、当該区切の直前の区間に対する音声認識の結果に応じた配信情報を順次に生成し、前記配信情報を順次に送信する。以上の態様では、音声の收音に並行した区切の特定毎に、当該区間に

50

対する音声認識の結果に応じた配信情報が生成される。したがって、音声の全区間の終了を契機として音声認識と配信情報の生成とを開始する構成と比較して、音声に対する配信情報の送信の遅延を低減することが可能である。

< 態様 2 >

態様 1 の好適例（態様 2）において、前記区切の特定では、前記音声の音量が低下した時点を前記区切として特定する。以上の態様では、音声の音量が低下した時点が区切として特定されるから、例えば利用者が手動で音声の区切を指示する手間を必要とせずに、音声の適切な時点を区切として特定できるという利点がある。

< 態様 3 >

態様 1 の好適例（態様 3）において、前記区切の特定では、操作装置に対する操作で指示された時点を前記区切として特定する。以上の態様によれば、利用者（例えば案内者）が意図した任意の時点を区切として設定できるという利点がある。

< 態様 4 >

態様 1 から態様 3 の何れかの好適例（態様 4）において、前記配信情報の送信では、前記配信情報の生成毎に当該配信情報を送信する。以上の態様では、配信情報の生成毎に当該配信情報が送信されるから、音声に対する配信情報の送信の遅延を低減することが可能である。

< 態様 5 >

態様 1 から態様 3 の何れかの好適例（態様 5）において、前記配信情報の送信では、前記音声認識により特定された文字列が特定の語句を含む場合に、前記配信情報を送信する。以上の態様では、特定の語句の発声を契機として配信情報が送信されるから、配信情報の送信のタイミングを発声者が適宜に調整できるという利点がある。

< 態様 6 >

態様 1 から態様 5 の何れかの好適例（態様 6）において、前記配信情報の送信では、前記特定した区切が前記音声の途中の時点である場合に、当該途中の時点であることを示す区切情報と前記配信情報とを送信する。以上の態様では、区切が音声の途中の時点であることを示す区切情報が配信情報とともに送信される。したがって、順次に送信される複数の配信情報が、相互に関連しているのか否かを受信側において把握することが可能である。

< 態様 7 >

態様 1 から態様 6 の何れかの好適例（態様 7）において、前記配信情報の送信では、前記特定した区切が前記音声の末尾の時点である場合に、当該末尾の時点であることを示す区切情報と前記配信情報とを送信する。以上の態様では、区切が音声の末尾の時点であることを示す区切情報が配信情報とともに送信される。したがって、順次に送信される複数の配信情報が、相互に関連しているのか否かを受信側において把握することが可能である。

< 態様 8 >

態様 1 から態様 7 の何れかの好適例（態様 8）において、前記配信情報の送信では、前記配信情報の生成毎に、当該配信情報と、前記音声について過去に生成した配信情報とを送信する。以上の態様では、新規に生成された配信情報とともに過去の配信情報が送信される。したがって、配信情報の受信側では、音声の 1 個の区間について受信に失敗した配信情報を、後続の区間に対応する配信情報とともに受信し直すことが可能である。

< 態様 9 >

態様 1 から態様 4 の何れかの好適例（態様 9）において、前記配信情報の生成では、前記区切の特定毎に、当該区切の直前の区間を含む複数の区間に対応する関連情報を示す配信情報を生成し、前記配信情報の送信においては、前記配信情報の生成毎に、当該配信情報と、前記関連情報の複数の部分の各々について利用者に対する提示の可否を示す区間情報とを送信する。以上の態様では、関連情報の複数の部分の各々について提示の可否を示す区間情報が送信される。したがって、音声の発声の開始から配信情報を取得できた利用者だけでなく、音声の発声の途中で配信情報を取得可能になった利用者にも、音声の最初の区間に対応する部分を含む関連情報を提示することが可能である。

< 態様 10 >

10

20

30

40

50

本発明の好適な態様（態様１０）に係る情報提供システムは、音声の收音に並行して当該音声の時間軸上の区切を順次に特定する区切特定部と、前記区切特定部による区切の特定毎に、当該区切の直前の区間に対する音声認識の結果に応じた配信情報を順次に生成する情報生成部と、前記配信情報を順次に送信する情報送信部とを具備する。以上の態様では、音声の收音に並行した区切の特定毎に、当該区間に対する音声認識の結果に応じた配信情報が生成される。したがって、音声の全区間の終了を契機として音声認識と配信情報の生成とを開始する構成と比較して、音声に対する配信情報の送信の遅延を低減することが可能である。

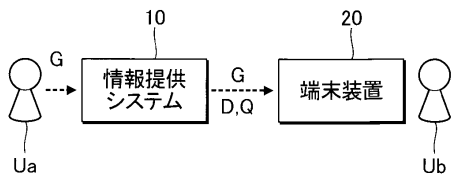
【符号の説明】

【００７６】

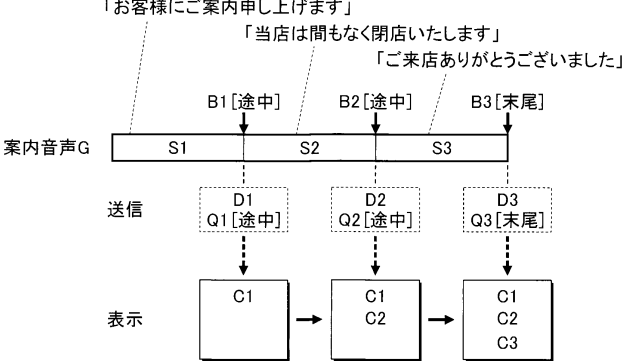
１０…情報提供システム、１１…制御装置、１１２…区切特定部、１１４…音声認識部、１１６…情報生成部、１２…記憶装置、１３…收音装置、１４…操作装置、１５…信号処理回路、１５２…変調処理部、１５４…混合処理部、１６…放音装置、１７…情報送信部、２０…端末装置、２１…制御装置、２１２…情報抽出部、２１４…提示制御部、２２…記憶装置、２３…收音装置、２４…表示装置。

【図面】

【図１】



【図２】



10

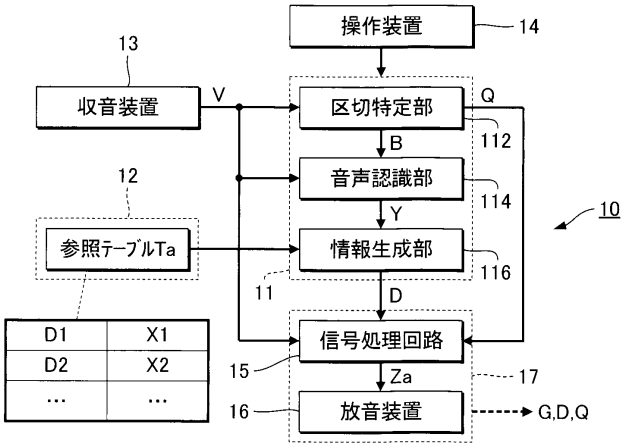
20

30

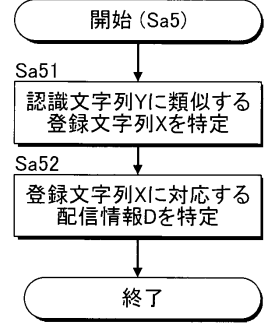
40

50

【図 3】

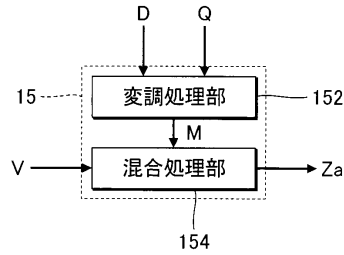


【図 4】

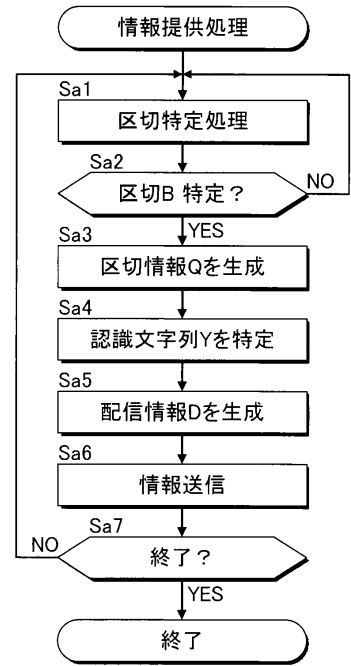


10

【図 5】



【図 6】



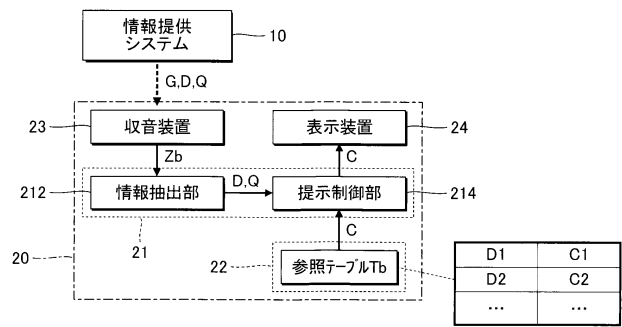
20

30

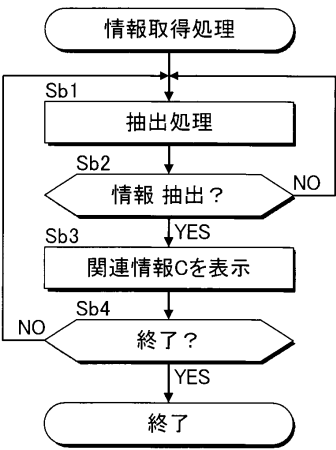
40

50

【図 7】

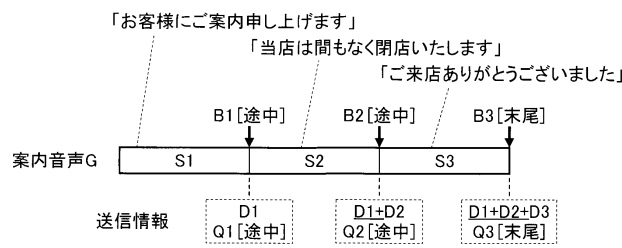


【図 8】

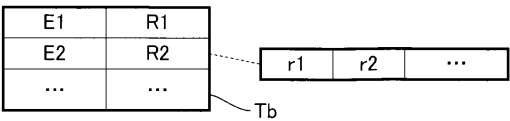


10

【図 9】

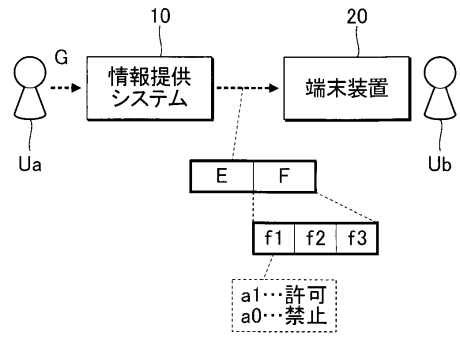


【図 10】

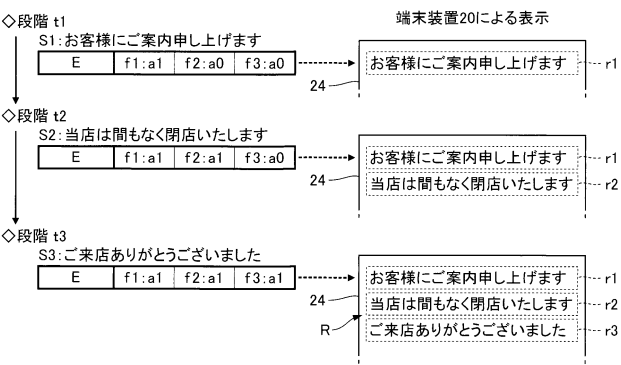


20

【図 11】



【図 12】



30

40

50

【 図 1 3 】



10

20

30

40

50

フロントページの続き

(72)発明者 静岡県浜松市中区中沢町 1 0 番 1 号 ヤマハ株式会社内
石田 哲朗
静岡県浜松市中区中沢町 1 0 番 1 号 ヤマハ株式会社内