

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第4987682号
(P4987682)

(45) 発行日 平成24年7月25日(2012.7.25)

(24) 登録日 平成24年5月11日(2012.5.11)

(51) Int.Cl.

F 1

G 1 O L 15/06 (2006.01)
G 1 O L 15/00 (2006.01)G 1 O L 15/06 300 J
G 1 O L 15/00 200 U
G 1 O L 15/00 200 A

請求項の数 19 (全 47 頁)

(21) 出願番号 特願2007-325274 (P2007-325274)
 (22) 出願日 平成19年12月17日 (2007.12.17)
 (65) 公開番号 特開2008-287210 (P2008-287210A)
 (43) 公開日 平成20年11月27日 (2008.11.27)
 審査請求日 平成22年12月1日 (2010.12.1)
 (31) 優先権主張番号 特願2007-107324 (P2007-107324)
 (32) 優先日 平成19年4月16日 (2007.4.16)
 (33) 優先権主張国 日本国 (JP)

(73) 特許権者 000002185
 ソニー株式会社
 東京都港区港南1丁目7番1号
 (73) 特許権者 397011373
 ソネットエンタテインメント株式会社
 東京都品川区大崎2丁目1番1号
 (74) 代理人 100095957
 弁理士 龟谷 美明
 (74) 代理人 100096389
 弁理士 金本 哲男
 (74) 代理人 100101557
 弁理士 萩原 康司
 (72) 発明者 倉田 宜典
 東京都港区港南1丁目7番1号 ソニー株式会社内

最終頁に続く

(54) 【発明の名称】音声チャットシステム、情報処理装置、音声認識方法およびプログラム

(57) 【特許請求の範囲】

【請求項 1】

音声認識を行ないつつ音声チャットを行なう複数の情報処理装置と、
 前記複数の情報処理装置と通信網を介して接続された検索サーバと、
 から構成され、

前記検索サーバは、当該検索サーバにおいて検索が行なわれた検索キーワードを記載した検索キーワードリストを、少なくとも1つの前記情報処理装置に対して開示し、
 少なくとも1つの前記情報処理装置は、

前記検索サーバから前記検索キーワードリストを取得して、前記音声認識を利用する単語が記載される認識単語辞書を生成する認識単語辞書生成部と、

前記音声チャットでの会話を音声データとし、前記認識単語辞書を含む認識用データベースを参照して、前記音声データを音声認識する音声認識部と、

他の情報処理装置との間における前記音声データの送受信の制御を行なう音声会話制御部と、

前記音声認識の結果から、前記音声チャットの会話内容に関するキーワードを検出するキーワード検出部と、

を備え、

前記認識単語辞書は、

所定の単語が予め登録された固定単語辞書と、

登録される単語を随時変更することが可能な可変単語辞書と、

10

20

を含み、

前記認識単語辞書生成部は、前記検索サーバにおける上位検索キーワードが記載された前記検索キーワードリストから前記可変単語辞書に登録する単語を選択して、前記可変単語辞書を生成し、

前記可変単語辞書に記載された単語には、前記固定単語辞書に記載された単語よりも高い重み付けがなされている、音声チャットシステム。

【請求項 2】

検索サーバから当該検索サーバで検索された検索キーワードが記載された検索キーワードリストを取得するとともに、他の情報処理装置との間で音声チャットを行う情報処理装置であって、

10

前記検索サーバから前記検索キーワードリストを取得して、音声認識に利用する単語が記載される認識単語辞書を生成する認識単語辞書生成部と、

前記音声チャットでの会話を音声データとし、前記認識単語辞書を含む認識用データベースを参照して、前記音声データを音声認識する音声認識部と、

前記他の情報処理装置との間ににおける前記音声データの送受信の制御を行う音声会話制御部と、

前記音声認識の結果から、前記音声チャットの会話内容に関するキーワードを検出するキーワード検出部と、

を備え、

前記認識単語辞書は、

20

所定の単語が予め登録された固定単語辞書と、

登録される単語を随時変更することが可能な可変単語辞書と、

を含み、

前記認識単語辞書生成部は、前記検索サーバにおける上位検索キーワードが記載された前記検索キーワードリストから前記可変単語辞書に登録する単語を選択して、前記可変単語辞書を生成し、

前記可変単語辞書に記載された単語には、前記固定単語辞書に記載された単語よりも高い重み付けがなされている、情報処理装置。

【請求項 3】

前記認識単語辞書生成部は、前記キーワード検出部が検出した前記キーワードの履歴情報に基づいて当該情報処理装置のユーザの嗜好を判断し、前記嗜好に基づいて、前記検索キーワードリストから前記単語を選択する、請求項2に記載の情報処理装置。

30

【請求項 4】

前記認識単語辞書生成部は、前記検索キーワードリストから選択された単語について前記検索サーバで検索を行い、前記検索によって得られた検索結果の中から、前記単語に関連するサブキーワードを取得し、

前記サブキーワードを、前記単語に関連付けて前記可変単語辞書に登録する、請求項2に記載の情報処理装置。

【請求項 5】

前記認識単語辞書生成部は、前記検索サーバにおける前記単語の検索頻度により決定される検索順位情報に基づいて、前記検索キーワードリストから選択した単語それぞれに対して、重み付けを行う、請求項2に記載の情報処理装置。

40

【請求項 6】

前記認識単語辞書生成部は、前記キーワード検出部が検索した前記キーワードの履歴情報に基づいて前記キーワードの頻度を算出し、前記可変単語辞書に登録されている前記キーワードに対応する単語の重み付けを、前記頻度に基づいて変更する、請求項5に記載の情報処理装置。

【請求項 7】

検索サーバから当該検索サーバで検索された検索キーワードが記載された検索キーワードリストを取得し、前記検索キーワードリストの中から単語を選択するとともに、選択し

50

た前記単語に対して音声認識に利用される重み付け情報を付加するものであり、前記重み付け情報には、選択した単語が優先的に音声認識されるような重み付けがなされる認識単語辞書生成装置から前記重み付け情報が付加された前記単語を取得し、他の情報処理装置との間で音声チャットを行う情報処理装置であって、

前記音声チャットでの会話を音声データとし、前記認識単語辞書生成装置から取得した前記単語を含む認識用データベースを参照して、前記音声データを音声認識する音声認識部と、

前記他の情報処理装置との間における前記音声データの送受信の制御を行う音声会話制御部と、

前記音声認識の結果から、前記音声チャットの会話内容に関するキーワードを検出するキーワード検出部と、

を備える、情報処理装置。

【請求項 8】

検索サーバから当該検索サーバで検索された検索キーワードが記載された検索キーワードリストを取得し、前記検索キーワードリストの中から単語を選択するとともに、選択した前記単語に対して音声認識に利用される重み付け情報を付加するものであり、前記重み付け情報には、選択した単語が優先的に音声認識されるような重み付けがなされる認識単語辞書生成装置から前記重み付け情報が付加された前記単語を取得し、他の情報処理装置との間で音声チャットを行う情報処理装置における音声認識方法であって、

前記認識単語辞書生成装置から前記重み付け情報が付加された単語を取得し、音声認識に用いられる認識用データベースに記憶するステップと、

前記音声チャットの会話内容を音声データとするステップと、

前記音声データを、当該音声データの音響的な特徴に基づいて解析し、前記会話内容に該当する単語列の候補を生成するステップと、

前記認識用データベースを参照しながら、生成した前記単語列の候補を、前記単語列を構成する単語間の言語的な連鎖の特徴に基づいて解析し、最も適した前記単語列を選択するステップと、

を含む、音声認識方法。

【請求項 9】

検索サーバから当該検索サーバで検索された検索キーワードが記載された検索キーワードリストを取得し、前記検索キーワードリストの中から単語を選択するとともに、選択した前記単語に対して音声認識に利用される重み付け情報を付加ものであり、前記重み付け情報には、選択した単語が優先的に音声認識されるような重み付けがなされる認識単語辞書生成装置から前記重み付け情報が付加された前記単語を取得し、他の情報処理装置との間で音声チャットを行う情報処理装置として機能するコンピュータに、

前記音声チャットでの会話を音声データとし、前記認識単語選択装置から取得した前記単語を含む認識用データベースを参照して、前記音声データを音声認識する音声認識機能と、

前記他の情報処理装置との間における前記音声データの送受信の制御を行う音声会話制御機能と、

前記音声認識の結果から、前記音声チャットの会話内容に関するキーワードを検出するキーワード検出機能と、
を実現させるためのプログラム。

【請求項 10】

検索サーバから当該検索サーバで検索された検索キーワードが記載された検索キーワードリストを取得して、言語認識に利用する単語が記載される認識単語辞書を生成する認識単語辞書生成部と、

放送局から送信された放送文字情報を受信する放送文字情報受信部と、

前記放送文字情報受信部が受信した前記放送文字情報に対応した単語列の中から抽出された所定のキーワードを管理するキーワード管理部と、

10

20

30

40

50

前記キーワード管理部により管理されている前記キーワードを参照して、前記放送文字情報から所定のキーワードを検出するキーワード検出部と、

外部表示装置との接続制御を行い、前記放送文字情報から検出された前記キーワードを、前記外部表示装置へと出力する外部表示装置接続制御部と、
を備え、

前記認識単語辞書は、

所定の単語が予め登録された固定単語辞書と、

登録される単語を随時変更することが可能な可変単語辞書と、
を含み、

前記認識単語辞書生成部は、前記検索サーバにおける上位検索キーワードが記載された前記検索キーワードリストから前記可変単語辞書に登録する単語を選択して、前記可変単語辞書を生成し、

前記可変単語辞書に記載された単語には、前記固定単語辞書に記載された単語よりも高い重み付けがなされている、情報処理装置。

【請求項 1 1】

前記放送文字情報受信部は、前記情報処理装置に接続された前記外部表示装置から、当該外部表示装置が受信している放送チャンネルに関する受信チャンネル情報と、前記放送チャンネルに対応する前記放送文字情報と、を取得する、請求項 1 0 に記載の情報処理装置。

【請求項 1 2】

前記情報処理装置は、前記外部表示装置での前記キーワードの参照履歴を記憶し、

前記外部表示接続制御部は、前記外部表示装置で選択されたキーワードを、前記検索サーバに通知する、請求項 1 0 に記載の情報処理装置。

【請求項 1 3】

検索サーバから当該検索サーバで検索された検索キーワードが記載された検索キーワードリストを取得して、音声認識に利用する単語が記載される認識単語辞書を生成する認識単語辞書生成部と、

放送局から送信された放送音声情報を受信する放送音声情報受信部と、

前記認識単語辞書を含む認識用データベースを参照して、前記放送音声情報を音声認識する音声認識部と、

前記放送音声情報の前記音声認識の結果から、所定のキーワードを検出するキーワード検出部と、

外部表示装置との接続制御を行い、前記放送文字情報から検出された前記キーワードを、前記外部表示装置へと出力する外部表示装置接続制御部と、
を備え、

前記認識単語辞書は、

所定の単語が予め登録された固定単語辞書と、

登録される単語を随時変更することが可能な可変単語辞書と、
を含み、

前記認識単語辞書生成部は、前記検索サーバにおける上位検索キーワードが記載された前記検索キーワードリストから前記可変単語辞書に登録する単語を選択して、前記可変単語辞書を生成し、

前記可変単語辞書に記載された単語には、前記固定単語辞書に記載された単語よりも高い重み付けがなされている、情報処理装置。

【請求項 1 4】

前記放送音声情報受信部は、前記情報処理装置に接続された前記外部表示装置から、当該外部表示装置が受信している放送チャンネルに関する受信チャンネル情報と、前記放送チャンネルに対応する前記放送音声情報と、を取得する、請求項 1 3 に記載の情報処理装置。

【請求項 1 5】

10

20

30

40

50

前記情報処理装置は、前記外部表示装置での前記キーワードの参照履歴を記憶し、前記外部表示装置接続制御部は、前記外部表示装置で選択されたキーワードを、前記検索サーバに通知する、請求項1_3に記載の情報処理装置。

【請求項 1_6】

検索サーバから当該検索サーバで検索された検索キーワードが記載された検索キーワードリストを取得して、言語認識に利用する単語が記載される認識単語辞書を生成するステップと、

放送局から送信された放送文字情報を受信するステップと、

前記認識単語辞書を参照して、前記放送文字情報から所定のキーワードを検出するステップと、

外部表示装置との接続制御を行い、前記放送文字情報から検出された前記キーワードを、前記外部表示装置へと出力するステップと、

を含み、

前記認識単語辞書は、

所定の単語が予め登録された固定単語辞書と、

登録される単語を随時変更することが可能な可変単語辞書と、

を含み、

前記認識単語辞書を生成するステップでは、前記検索サーバにおける上位検索キーワードが記載された前記検索キーワードリストから前記可変単語辞書に登録する単語を選択して、前記可変単語辞書が生成され、

前記可変単語辞書に記載された単語には、前記固定単語辞書に記載された単語よりも高い重み付けがなされている、キーワード検出方法。

【請求項 1_7】

検索サーバから当該検索サーバで検索された検索キーワードが記載された検索キーワードリストを取得して、音声認識に利用する単語が記載される認識単語辞書を生成するステップと、

放送局から送信された放送音声情報を受信するステップと、

前記認識単語辞書を含む認識用データベースを参照して、前記放送音声情報を音声認識するステップと、

前記放送音声情報の前記音声認識の結果から、所定のキーワードを検出するステップと、

を含み、

外部表示装置との接続制御を行い、前記放送文字情報から検出された前記キーワードを、前記外部表示装置へと出力するステップと、

を含み、

前記認識単語辞書は、

所定の単語が予め登録された固定単語辞書と、

登録される単語を随時変更することが可能な可変単語辞書と、

を含み、

前記認識単語辞書を生成するステップでは、前記検索サーバにおける上位検索キーワードが記載された前記検索キーワードリストから前記可変単語辞書に登録する単語を選択して、前記可変単語辞書が生成され、

前記可変単語辞書に記載された単語には、前記固定単語辞書に記載された単語よりも高い重み付けがなされている、キーワード検出方法。

【請求項 1_8】

コンピュータに、

検索サーバから当該検索サーバで検索された検索キーワードが記載された検索キーワードリストを取得して、言語認識に利用する単語が記載される認識単語辞書を生成する認識単語辞書生成機能と、

放送局から送信された放送文字情報を受信する放送文字情報受信機能と、

前記認識単語辞書を参照して、前記放送文字情報から所定のキーワードを検出するキー

10

20

30

40

50

ワード検出機能と、

外部表示装置との接続制御を行い、前記放送文字情報から検出された前記キーワードを、前記外部表示装置へと出力する外部表示装置接続制御機能と、
を実現させ、

前記認識単語辞書は、

所定の単語が予め登録された固定単語辞書と、

登録される単語を随時変更することが可能な可変単語辞書と、
を含み、

前記認識単語辞書生成機能は、前記検索サーバにおける上位検索キーワードが記載された前記検索キーワードリストから前記可変単語辞書に登録する単語を選択して、前記可変単語辞書を生成し、

前記可変単語辞書に記載された単語には、前記固定単語辞書に記載された単語よりも高い重み付けがなされるプログラム。

【請求項 19】

コンピュータに、

検索サーバから当該検索サーバで検索された検索キーワードが記載された検索キーワードリストを取得して、音声認識に利用する単語が記載される認識単語辞書を生成する認識単語辞書生成機能と、

放送局から送信された放送音声情報を受信する放送音声情報受信機能と、

前記認識単語辞書を含む認識用データベースを参照して、前記放送音声情報を音声認識する音声認識機能と、

前記放送音声情報の前記音声認識の結果から、所定のキーワードを検出するキーワード検出機能と、

外部表示装置との接続制御を行い、前記放送文字情報から検出された前記キーワードを、前記外部表示装置へと出力する外部表示装置接続制御機能と、
を実現させ、

前記認識単語辞書は、

所定の単語が予め登録された固定単語辞書と、

登録される単語を随時変更することが可能な可変単語辞書と、
を含み、

前記認識単語辞書生成機能は、前記検索サーバにおける上位検索キーワードが記載された前記検索キーワードリストから前記可変単語辞書に登録する単語を選択して、前記可変単語辞書を生成し、

前記可変単語辞書に記載された単語には、前記固定単語辞書に記載された単語よりも高い重み付けがなされるプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、音声チャットシステム、情報処理装置、音声認識方法およびプログラムに関する。

【背景技術】

【0002】

音声認識は古くから研究が行なわれており、読み上げられた音声に関しては、かなり高精度の認識率が得られるようになってきている。しかしながら、人間同士の自由な会話の認識は、未だに高い性能は得られていない。

【0003】

他方、音声からその話題を抽出する技術は、トピック検出技術とも呼ばれ、近年盛んに研究が行なわれている。このトピック検出技術を行なう場合には、音声からテキストの情報を取り出す音声認識部が重要な役割を果たす。

【0004】

10

20

30

40

50

音声からキーワードを抽出する際に、キーワードのみに注目して音声中からキーワードを発見する方法と、音声全体を大語彙音声認識などで全て認識し、認識結果からキーワードを抽出する方法がある。例えば前者では例えば音韻系列を認識し得られた音韻ラティスから単語列を抽出する方式が用いられ、後者では大語彙連続音声認識（Large Vocabulary Continuous Speech Recognition：LVCSR）が用いられる。キーワード数が多い場合後者の方が計算量の点で有効である。また、いずれの場合においても、認識すべき語彙の言語的知識が必要となるが、これは、検出すべき語彙の出現頻度情報等を利用することができます。

【0005】

音声認識には、単独の単語を認識する孤立単語認識と複数の単語からなる単語列を認識する連続単語認識がある。従来の連続単語認識では、言語モデルという「単語間のつながりやすさについてのデータベース」を持つことで、「音は似ているが滅茶苦茶な単語列」が認識結果として生成されることを防いでいる。

【0006】

しかしながら、言語モデルには、最初から認識できる単語（以下、適宜、既知語と称する。）についての情報のみ記述されるため、後で登録された単語（以下、適宜、登録単語と称する。）を正しく認識することが困難であった。なぜなら、孤立単語認識では、認識辞書に単語を登録すれば、以降その単語は認識されるようになるが、連続単語認識では辞書への登録だけでは不十分であり、登録単語を言語モデルにも反映させる必要があるが、言語モデルへの反映は一般的には困難であったからである。

10

20

【0007】

そこで、登録単語を、例えば「人名」、「地名」などのカテゴリに分類し、そのカテゴリに対応した言語モデルを用意し、登録単語をカテゴリに属させることで、新しい語彙であっても連続音声認識で扱える方法が開示されている（例えば、特許文献1参照。）。

【0008】

一方、登録単語の選択も大きな問題であった。特に固有名詞は重要なキーワードである場合が多く、固有名詞を認識することでユーザに対して有益な情報を提示することが可能になるからである。

【0009】

そこで、インターネット情報等から取得したキーワードを用いて、キーワードの抽出を行なう方法が提案されている（例えば、特許文献2参照。）。

30

【0010】

【特許文献1】特開2004-252121号公報

【特許文献2】特開2002-216026号公報

【0011】

ところが、固有名詞の数は極めて多く、ユーザが話すであろう単語をあらかじめ音声認識にすべて登録しておくことは事実上不可能である。

【発明の開示】

【発明が解決しようとする課題】

【0012】

このため、特許文献2に記載の方法を用いたとしても、音声チャットにおける会話の中から、例えば最近の話題に関するキーワードを高精度で認識することは、困難であるという問題があった。

40

【0013】

そこで、本発明は、このような問題に鑑みてなされたもので、その目的は、音声チャットにおける会話中に存在するキーワードを高精度で認識することが可能な、新規かつ改良された音声チャットシステム、情報処理装置、音声認識方法およびプログラムを提供することにある。

【0014】

さらに、本発明は、放送情報中に存在するキーワードを高精度に検出することが可能な

50

、新規かつ改良された情報処理装置、キーワード検出方法およびプログラムを提供する。

【課題を解決するための手段】

【0015】

上記課題を解決するために、本発明のある観点によれば、音声認識を行ないつつ音声チャットを行なう複数の情報処理装置と、前記複数の情報処理装置と通信網を介して接続された検索サーバと、から構成され、前記検索サーバは、当該検索サーバにおいて検索が行なわれた検索キーワードを記載した検索キーワードリストを、少なくとも1つの前記情報処理装置に対して開示し、少なくとも1つの前記情報処理装置は、前記検索サーバから前記検索キーワードリストを取得して、前記音声認識に利用する単語が記載される認識単語辞書を生成する認識単語辞書生成部と、前記音声チャットでの会話を音声データとし、前記認識単語辞書を含む認識用データベースを参照して、前記音声データを音声認識する音声認識部と、他の情報処理装置との間における前記音声データの送受信の制御を行なう音声会話制御部と、前記音声認識の結果から、前記音声チャットの会話内容に関するキーワードを検出するキーワード検出部と、を備える音声チャットシステムが提供される。

10

【0016】

上記課題を解決するために、本発明の別の観点によれば、検索サーバから当該検索サーバで検索された検索キーワードが記載された検索キーワードリストを取得するとともに、他の情報処理装置との間で音声チャットを行う情報処理装置であって、前記検索サーバから前記検索キーワードリストを取得して、音声認識に利用する単語が記載される認識単語辞書を生成する認識単語辞書生成部と、前記音声チャットでの会話を音声データとし、前記認識単語辞書を含む認識用データベースを参照して、前記音声データを音声認識する音声認識部と、前記他の情報処理装置との間における前記音声データの送受信の制御を行なう音声会話制御部と、前記音声認識の結果から、前記音声チャットの会話内容に関するキーワードを検出するキーワード検出部と、を備える情報処理装置が提供される。

20

【0017】

前記認識単語辞書は、所定の単語が登録された固定単語辞書と、登録される単語を随時変更することが可能な可変単語辞書と、を含み、前記認識単語辞書生成部は、前記検索サーバにおける上位検索キーワードが記載された前記検索キーワードリストから前記可変単語辞書に登録する単語を選択して、前記可変単語辞書を生成してもよい。

30

【0018】

前記認識単語辞書生成部は、前記キーワード検出部が検出した前記キーワードの履歴情報に基づいて当該情報処理装置のユーザの嗜好を判断し、前記嗜好に基づいて、前記検索キーワードリストから前記単語を選択してもよい。

【0019】

前記認識単語辞書生成部は、前記検索キーワードリストから選択された単語について前記検索サーバで検索を行い、前記検索によって得られた検索結果の中から、前記単語に関連するサブキーワードを取得し、前記サブキーワードを、前記単語に関連付けて前記可変単語辞書に登録してもよい。

【0020】

前記認識単語辞書生成部は、前記検索サーバにおける前記単語の検索頻度により決定される検索順位情報に基づいて、前記検索キーワードリストから選択した単語それぞれに対して、重み付けを行ってもよい。

40

【0021】

前記認識単語辞書生成部は、前記キーワード検出部が検索した前記キーワードの履歴情報に基づいて前記キーワードの頻度を算出し、前記可変単語辞書に登録されている前記キーワードに対応する単語の重み付けを、前記頻度に基づいて変更してもよい。

【0022】

上記課題を解決するために、本発明の更に別の観点によれば、検索サーバから当該検索サーバで検索された検索キーワードが記載された検索キーワードリストを取得し、前記検索キーワードリストの中から単語を選択するとともに、選択した前記単語に対して音声認

50

識に利用される重み付け情報を付加する認識単語辞書生成装置から前記重み付け情報が付加された前記単語を取得し、他の情報処理装置との間で音声チャットを行う情報処理装置であって、前記音声チャットでの会話を音声データとし、前記認識単語辞書生成装置から取得した前記単語を含む認識用データベースを参照して、前記音声データを音声認識する音声認識部と、前記他の情報処理装置との間における前記音声データの送受信の制御を行う音声会話制御部と、前記音声認識の結果から、前記音声チャットの会話内容に関するキーワードを検出するキーワード検出部と、を備える情報処理装置が提供される。

【 0 0 2 3 】

上記課題を解決するために、本発明の更に別の観点によれば、検索サーバから当該検索サーバで検索された検索キーワードが記載された検索キーワードリストを取得するとともに、他の情報処理装置との間で音声チャットを行う情報処理装置における音声認識方法であって、前記検索サーバから前記検索キーワードリストを取得して、音声認識を利用する単語が記載される認識単語辞書を生成するステップと、前記音声チャットの会話内容を音声データとするステップと、前記音声データを、当該音声データの音響的な特徴に基づいて解析し、前記会話内容に該当する単語列の候補を生成するステップと、音声認識に用いられ前記認識単語辞書を含む認識用データベースを参照しながら、生成した前記単語列の候補を、前記単語列を構成する単語間の言語的な連鎖の特徴に基づいて解析し、最も適した前記単語列を選択するステップと、を含む音声認識方法が提供される。10

【 0 0 2 4 】

上記課題を解決するために、本発明の更に別の観点によれば、検索サーバから当該検索サーバで検索された検索キーワードが記載された検索キーワードリストを取得し、前記検索キーワードリストの中から単語を選択するとともに、選択した前記単語に対して音声認識に利用される重み付け情報を付加する認識単語辞書生成装置から前記重み付け情報が付加された前記単語を取得し、他の情報処理装置との間で音声チャットを行う情報処理装置における音声認識方法であって、前記認識単語辞書生成装置から前記重み付け情報が付加された単語を取得し、音声認識に用いられる認識用データベースに記憶するステップと、前記音声チャットの会話内容を音声データとするステップと、前記音声データを、当該音声データの音響的な特徴に基づいて解析し、前記会話内容に該当する単語列の候補を生成するステップと、前記認識用データベースを参照しながら、生成した前記単語列の候補を、前記単語列を構成する単語間の言語的な連鎖の特徴に基づいて解析し、最も適した前記単語列を選択するステップと、を含む音声認識方法が提供される。20

【 0 0 2 5 】

上記課題を解決するために、本発明の更に別の観点によれば、検索サーバから当該検索サーバで検索された検索キーワードが記載された検索キーワードリストを取得するとともに、他の情報処理装置との間で音声チャットを行う情報処理装置を制御するコンピュータに、前記検索サーバから前記検索キーワードリストを取得して、音声認識を利用する単語が記載される認識単語辞書を生成する認識単語辞書生成機能と、前記音声チャットでの会話を音声データとし、前記認識単語辞書を含む認識用データベースを参照して、前記音声データを音声認識する音声認識機能と、前記他の情報処理装置との間における前記音声データの送受信の制御を行う音声会話制御機能と、前記音声認識の結果から、前記音声チャットの会話内容に関するキーワードを検出するキーワード検出機能と、を実現させるためのプログラムが提供される。30

【 0 0 2 6 】

かかる構成によれば、コンピュータプログラムは、コンピュータが備える記憶部に格納され、コンピュータが備えるCPUに読み込まれて実行されることにより、そのコンピュータを上記の情報処理装置として機能させる。また、コンピュータプログラムが記録された、コンピュータで読み取り可能な記録媒体も提供することができる。記録媒体は、例えば、磁気ディスク、光ディスク、光磁気ディスク、フラッシュメモリなどである。また、上記のコンピュータプログラムは、記録媒体を用いずに、例えばネットワークを介して配信してもよい。40

【 0 0 2 7 】

上記課題を解決するために、本発明の更に別の観点によれば、検索サーバから当該検索サーバで検索された検索キーワードが記載された検索キーワードリストを取得し、前記検索キーワードリストの中から単語を選択するとともに、選択した前記単語に対して音声認識に利用される重み付け情報を付加する認識単語辞書生成装置から前記重み付け情報が付加された前記単語を取得し、他の情報処理装置との間で音声チャットを行う情報処理装置を制御するコンピュータに、前記音声チャットでの会話を音声データとし、前記認識単語選択装置から取得した前記単語を含む認識用データベースを参照して、前記音声データを音声認識する音声認識機能と、前記他の情報処理装置との間における前記音声データの送受信の制御を行う音声会話制御機能と、前記音声認識の結果から、前記音声チャットの会話内容に関するキーワードを検出するキーワード検出機能と、を実現させるためのプログラムが提供される。

【 0 0 2 8 】

上記課題を解決するために、本発明の更に別の観点によれば、検索サーバから当該検索サーバで検索された検索キーワードが記載された検索キーワードリストを取得して、言語認識に利用する単語が記載される認識単語辞書を生成する認識単語辞書生成部と、放送局から送信された放送文字情報を受信する放送文字情報受信部と、前記認識単語辞書を参照して、前記放送文字情報から所定のキーワードを検出するキーワード検出部と、外部表示装置との接続制御を行い、前記放送文字情報から検出された前記キーワードを、前記外部表示装置へと出力する外部表示装置接続制御部と、を備える情報処理装置が提供される。

10

【 0 0 2 9 】

放送文字情報受信部は、情報処理装置に接続された外部表示装置から、当該外部表示装置が受信している放送チャンネルに関する受信チャンネル情報と、放送チャンネルに対応する放送文字情報と、を取得してもよい。また、放送文字情報受信部は、情報処理装置に接続された外部表示装置から当該外部表示装置が受信している放送チャンネルに関する受信チャンネル情報を取得するとともに、放送局から放送チャンネルに対応した放送文字情報を取得してもよい。

20

【 0 0 3 0 】

前記情報処理装置は、前記外部表示装置での前記キーワードの参照履歴を記憶し、前記外部表示装置接続制御部は、前記外部表示装置で選択されたキーワードを、前記検索サーバに通知してもよい。

30

【 0 0 3 1 】

上記課題を解決するために、本発明の更に別の観点によれば、検索サーバから当該検索サーバで検索された検索キーワードが記載された検索キーワードリストを取得して、音声認識に利用する単語が記載される認識単語辞書を生成する認識単語辞書生成部と、放送局から送信された放送音声情報を受信する放送音声情報受信部と、前記認識単語辞書を含む認識用データベースを参照して、前記放送音声情報を音声認識する音声認識部と、前記放送音声情報の前記音声認識の結果から、所定のキーワードを検出するキーワード検出部と、外部表示装置との接続制御を行い、前記放送文字情報から検出された前記キーワードを、前記外部表示装置へと出力する外部表示装置接続制御部と、を備える情報処理装置が提供される。

40

【 0 0 3 2 】

放送音声情報受信部は、情報処理装置に接続された外部表示装置から、当該外部表示装置が受信している放送チャンネルに関する受信チャンネル情報と、放送チャンネルに対応する放送音声情報と、を取得してもよい。また、放送音声情報受信部は、情報処理装置に接続された外部表示装置から当該外部表示装置が受信している放送チャンネルに関する受信チャンネル情報を取得するとともに、放送局から放送チャンネルに対応した放送音声情報を取得してもよい。

【 0 0 3 3 】

前記情報処理装置は、前記外部表示装置での前記キーワードの参照履歴を記憶し、前記

50

外部表示装置接続制御部は、前記外部表示装置で選択されたキーワードを、前記検索サーバに通知してもよい。

【0034】

上記課題を解決するために、本発明の更に別の観点によれば、検索サーバから当該検索サーバで検索された検索キーワードが記載された検索キーワードリストを取得して、言語認識に利用する単語が記載される認識単語辞書を生成するステップと、放送局から送信された放送文字情報を受信するステップと、前記認識単語辞書を参照して、前記放送文字情報から所定のキーワードを検出するステップと、外部表示装置との接続制御を行い、前記放送文字情報から検出された前記キーワードを、前記外部表示装置へと出力するステップと、を含むキーワード検出方法が提供される。

10

【0035】

上記課題を解決するために、本発明の更に別の観点によれば、検索サーバから当該検索サーバで検索された検索キーワードが記載された検索キーワードリストを取得して、音声認識に利用する単語が記載される認識単語辞書を生成するステップと、放送局から送信された放送音声情報を受信するステップと、前記認識単語辞書を含む認識用データベースを参照して、前記放送音声情報を音声認識するステップと、前記放送音声情報の前記音声認識の結果から、所定のキーワードを検出するステップと、外部表示装置との接続制御を行い、前記放送文字情報から検出された前記キーワードを、前記外部表示装置へと出力するステップと、を含むキーワード検出方法が提供される。

【0036】

20

上記課題を解決するために、本発明の更に別の観点によれば、コンピュータに、検索サーバから当該検索サーバで検索された検索キーワードが記載された検索キーワードリストを取得して、言語認識に利用する単語が記載される認識単語辞書を生成する認識単語辞書生成機能と、放送局から送信された放送文字情報を受信する放送文字情報受信機能と、前記認識単語辞書を参照して、前記放送文字情報から所定のキーワードを検出するキーワード検出機能と、外部表示装置との接続制御を行い、前記放送文字情報から検出された前記キーワードを、前記外部表示装置へと出力する外部表示装置接続制御機能と、を実現させるためのプログラムが提供される。

【0037】

30

かかる構成によれば、コンピュータプログラムは、コンピュータが備える記憶部に格納され、コンピュータが備えるCPUに読み込まれて実行されることにより、そのコンピュータを上記の情報処理装置として機能させる。また、コンピュータプログラムが記録された、コンピュータで読み取り可能な記録媒体も提供することができる。記録媒体は、例えば、磁気ディスク、光ディスク、光磁気ディスク、フラッシュメモリなどである。また、上記のコンピュータプログラムは、記録媒体を用いずに、例えばネットワークを介して配信してもよい。

【0038】

40

上記課題を解決するために、本発明の更に別の観点によれば、コンピュータに、検索サーバから当該検索サーバで検索された検索キーワードが記載された検索キーワードリストを取得して、音声認識に利用する単語が記載される認識単語辞書を生成する認識単語辞書生成機能と、放送局から送信された放送音声情報を受信する放送音声情報受信機能と、前記認識単語辞書を含む認識用データベースを参照して、前記放送音声情報を音声認識する音声認識機能と、前記放送音声情報の前記音声認識の結果から、所定のキーワードを検出するキーワード検出機能と、外部表示装置との接続制御を行い、前記放送文字情報から検出された前記キーワードを、前記外部表示装置へと出力する外部表示装置接続制御機能と、を実現させるためのプログラムが提供される。

【0039】

50

かかる構成によれば、コンピュータプログラムは、コンピュータが備える記憶部に格納され、コンピュータが備えるCPUに読み込まれて実行されることにより、そのコンピュータを上記の情報処理装置として機能させる。また、コンピュータプログラムが記録され

た、コンピュータで読み取り可能な記録媒体も提供することができる。記録媒体は、例えば、磁気ディスク、光ディスク、光磁気ディスク、フラッシュメモリなどである。また、上記のコンピュータプログラムは、記録媒体を用いずに、例えばネットワークを介して配信してもよい。

【発明の効果】

【0040】

本発明によれば、音声認識用の認識単語辞書を検索サーバから取得したデータベースを用いて作成し、この認識単語辞書を用いて音声認識を行なうため、音声チャットにおける会話中に存在するキーワードを高精度で認識することが可能である。また、本発明によれば、言語認識用の認識単語辞書を検索サーバから取得したデータベースを用いて作成し、この認識単語辞書を用いてキーワードの検出を行なうため、放送情報中に存在するキーワードを高精度で検出することが可能である。10

【発明を実施するための最良の形態】

【0041】

以下に添付図面を参照しながら、本発明の好適な実施の形態について詳細に説明する。なお、本明細書及び図面において、実質的に同一の機能構成を有する構成要素については、同一の符号を付することにより重複説明を省略する。

【0042】

(第1の実施形態)

以下に、本発明の第1の実施形態に係る音声チャットシステムについて、詳細に説明する。20

【0043】

図1は、本実施形態に係る音声チャットシステムを説明するための説明図である。音声チャットシステム10は、例えば、通信網12と、検索サーバ14と、情報処理装置20A、20Bと、を含む。

【0044】

通信網12は、情報処理装置20および検索サーバ14を双方向通信または一方向通信可能に接続する通信回線網である。この通信網12は、例えば、インターネット、電話回線網、衛星通信網、同報通信路等の公衆回線網や、WAN(Wide Area Network)、LAN(Local Area Network)、IP-VPN(Internet Protocol-Virtual Private Network)、Ethernet(登録商標)、ワイヤレスLAN等の専用回線網等で構成されており、有線/無線を問わない。30

【0045】

検索サーバ14は、インターネット等で公開されている情報についてキーワード等を使って検索できるWebサイトである。本実施形態に係る検索サーバ14は、キーワードによって蓄えられている情報の全文検索を行う全文検索型であってもよく、カテゴリ別に分類されている情報を検索するディレクトリ型であってもよい。

【0046】

情報処理装置20は、通信網12を介して接続された他の情報処理装置20と、音声によるチャット(音声チャット)を行う。また、情報処理装置20は、通信網12に接続されている検索サーバ14に対して、通信網12を介して情報の検索を要求することができる。また、音声チャットを行う他の情報処理装置20は、図示のように通信網12を介して接続されていてもよく、また、通信網12を介さずに、例えば、USB(Universal Serial Bus)ポートや、i.Link等のIEEE1394ポート、SCSI(Small Computer System Interface)ポート、RS-232Cポート等により直接接続されていてもよい。40

【0047】

なお、図示の例では、情報処理装置20として、デスクトップ型のPCを示しているが、本実施形態に係る情報処理装置20は、デスクトップ型PC、ノート型PCを問わない50

。また、本実施形態に係る情報処理装置20は、かかる例に限定されず、ネットワークを介した通信機能を有する機器であれば、例えば、テレビジョン受像器や家庭用ゲーム機等の情報家電、携帯電話、PDA(Personal Digital Assistant)等で構成することもできる。また、情報処理装置20は、契約者が持ち運びできるポータブルデバイス(Portable Device)、例えば、携帯型ゲーム機、PHS、携帯型映像／音声プレーヤなどであってもよい。

【0048】

さらに、図1では、通信網12に接続されている情報処理装置20は、2つのみであるが、本実施形態は、上記の場合に限定されるわけではなく、情報処理装置20は、通信網12上に複数接続されていてもよい。

10

【0049】

(情報処理装置20のハードウェア構成)

次に、本実施形態に係る情報処理装置20のハードウェア構成について、図2を参照しながら簡単に説明する。

【0050】

図2は、本実施形態にかかる情報処理装置20のハードウェア構成を示した説明図である。情報処理装置20は、主に、CPU(Central Processing Unit)201と、ROM(Read Only Memory)203と、RAM(Random Access Memory)205と、ホストバス207と、ブリッジ209と、外部バス211と、インターフェース213と、入力装置215と、出力装置217と、ストレージ装置219と、ドライブ221と、通信装置223とを備える。

20

【0051】

CPU201は、演算処理装置および制御装置として機能し、ROM203、RAM205、ストレージ装置219、またはリムーバブル記録媒体16に記録された各種プログラムに従って情報処理装置20内の動作全般またはその一部を制御する。ROM203は、CPU201が使用するプログラムや演算パラメータ等を記憶する。RAM205は、CPU201の実行において使用するプログラムや、その実行において適宜変化するパラメータ等を一次記憶する。これらはCPUバス等の内部バスにより構成されるホストバス207により相互に接続されている。

【0052】

30

ホストバス207は、ブリッジ209を介して、PCI(Peripheral Component Interconnect/Interface)バスなどの外部バス211に接続されている。

【0053】

入力装置215は、例えば、マウス、キーボード、タッチパネル、ボタン、スイッチおよびレバー等のユーザが操作する操作手段と、マイクロフォンやヘッドセット等の音声入力手段とを備える。また、入力装置215は、例えば、赤外線や他の電波を利用したリモートコントロール手段(いわゆる、リモコン)であってもよいし、情報処理装置20の操作に対応した携帯電話やPDA等の外部接続機器であってもよい。さらに、入力装置215は、例えば、上記の操作手段や音声入力手段を用いてユーザにより入力された情報に基づいて入力信号を生成し、CPU201に出力する入力制御回路などから構成されている。情報処理装置20のユーザは、この入力装置215を操作することにより、情報処理装置20に対して各種のデータを入力したり処理動作を指示したりすることができる。

40

【0054】

出力装置217は、例えば、CRT(Cathode Ray Tube)ディスプレイ装置、液晶ディスプレイ(Liquid Crystal Display:LCD)装置、プラズマディスプレイ(Plasma Display Panel:PDP)装置、EL(Electro-Luminescence)ディスプレイ装置およびランプなどの表示装置や、スピーカおよびヘッドホンなどの音声出力装置や、プリンタ装置、携帯電話、ファクシミリなど、取得した情報をユーザに対して視覚的または聴覚的に通知す

50

ることが可能な装置で構成される。出力装置 217 は、例えば、検索サーバを用いて検索した各種情報を出力する。具体的には、表示装置は、検索サーバによる各種情報の検索結果をテキストまたはイメージで表示する。他方、音声出力装置は、再生された音声データ等を音声に変換して出力する。

【0055】

ストレージ装置 219 は、本実施形態にかかる情報処理装置 20 の記憶部の一例として構成されたデータ格納用の装置であり、例えば、HDD (Hard Disk Drive) 等の磁気記憶部デバイス、半導体記憶デバイス、光記憶デバイス、または光磁気記憶デバイス等により構成される。このストレージ装置 219 は、CPU 201 が実行するプログラムや各種データ、および外部から取得した各種データなどを格納する。

10

【0056】

ドライブ 221 は、記憶媒体用リーダライタであり、情報処理装置 20 に内蔵、あるいは外付けされる。ドライブ 221 は、装着されている磁気ディスク、光ディスク、光磁気ディスク、または半導体メモリ等のリムーバブル記録媒体 16 に記録されている情報を読み出して、RAM 205 に出力する。また、ドライブ 221 は、装着されている磁気ディスク、光ディスク、光磁気ディスク、または半導体メモリ等のリムーバブル記録媒体 16 に記録を書き込むことも可能である。リムーバブル記録媒体 16 は、例えば、DVD メディア、HD-DVD メディア、Blu-ray メディア、コンパクトフラッシュ (Compact Flash : CF)、メモリースティック、または、SD メモリカード (Secure Digital memory card) 等である。また、リムーバブル記録媒体 16 は、例えば、非接触型 IC チップを搭載した IC カード (Integrated Circuit card) または電子機器等であってもよい。

20

【0057】

通信装置 223 は、例えば、通信網 12 に接続するための通信デバイス等で構成された通信インターフェースである。通信装置 223 は、例えば、有線または無線 LAN (Local Area Network)、Bluetooth、または WUSB (Wireless USB) 用の通信カード、光信用のルータ、ADSL (Asymmetric Digital Subscriber Line) 用のルータ、または、各種通信用のモデム等である。この通信装置 223 は、他の情報処理装置 20 との間で音声チャットに関する情報を送受信すると共に、例えば、インターネットや他の通信機器との間で各種の情報を送受信することができる。また、通信装置 223 に接続される通信網 12 は、有線または無線によって接続されたネットワーク等により構成され、例えば、インターネット、家庭内 LAN、赤外線通信、または衛星通信等であってもよい。

30

【0058】

以上説明した構成により、情報処理装置 20 は、当該情報処理装置 20 に直接接続された他の情報処理装置、または、通信網 12 に接続された他の情報処理装置と音声チャットを行うことが可能になると同時に、通信網 12 に接続された検索サーバ 14 等から、各種の情報を取得することが可能となる。さらに、情報処理装置 20 は、リムーバブル記録媒体 16 を用いて、当該情報処理装置 20 に蓄積されている情報を持ち出すことも可能である。

40

【0059】

以上、本実施形態に係る情報処理装置 20 の機能を実現可能なハードウェア構成の一例を示した。上記の各構成要素は、汎用的な部材を用いて構成されていてもよいし、各構成要素の機能に特化したハードウェアにより構成されていてもよい。従って、本実施形態を実施する時々の技術レベルに応じて、適宜、利用するハードウェア構成を変更することができる。また、上記のハードウェア構成は、あくまでも一例であり、これに限定されるものでないことは言うまでもない。また、利用形態によっては、ホストバス 207 や外部バス 211、またはインターフェース 213 等を省略する構成も可能である。

【0060】

(情報処理装置 20 の構成)

50

続いて、本実施形態に係る情報処理装置20の構成について詳細に説明を行うが、以下の説明においては、音声チャットを行う2つの情報処理装置20について、便宜的に第1情報処理装置20Aおよび第2情報処理装置20Bと称することとする。また、第1情報処理装置20Aおよび第2情報処理装置20Bは、各情報処理装置のユーザの声を音声データ化し、音声チャットの会話相手である他の情報処理装置に対して、音声データを送信するものとする。図3は、本実施形態に係る第1情報処理装置20Aの構成を説明するためのブロック図である。

【0061】

なお、以下の説明において、単語とは、音声を認識する処理において、1つのまとまりとして扱った方がよい単位のことを言い、言語学的な単語とは必ずしも一致しない。例えば、「タロウ君」は、それ全体を1単語として扱ってもよいし、「タロウ」、「君」という2単語として扱ってもよい。さらに、もっと大きな単位である「こんにちはタロウ君」等を1単語として扱ってもよい。

【0062】

また、音韻とは、音響的に1つの単位として扱った方が処理上都合のよいものについてい、音声学的な音韻や音素とは必ずしも一致しない。例えば、「東京」の「とう」の部分を“t / o / u”という3個の音韻記号で表すことも可能であり、または“o”の長音である“o :”という記号を用意してもよい。さらに、“t / o / o”と表してもよい。他にも、無音を表す記号を用意してもよく、さらに無音を表す記号を「発話前の無音」「発話に挟まれた短い無音区間」「「つ」」の部分の無音」のように細かく分類してもよい。

【0063】

図3に示したように、本実施形態に係る第1情報処理装置20Aは、例えば、音声入力部251と、音声認識部253と、キーワード検出部255と、キーワード管理部257と、音声会話制御部259と、表示部261と、音声出力部263と、通信部265と、記憶部267と、認識単語辞書生成部269と、を備える。

【0064】

音声入力部251は、第1情報処理装置20Aのユーザが発した音声である音声チャットでの会話を、第1情報処理装置20Aへと取り込むインターフェースであって、例えば、第1情報処理装置20Aに接続された入力装置であるマイクロフォンやヘッドセット等から構成される。マイクロフォンやヘッドセットに向かって発せられたユーザの会話は、自動的かつリアルタイムに第1情報処理装置20Aへと取り込まれて電気信号としての音声信号に変換され、音声入力部251を介して音声認識部253へと伝送される。

【0065】

音声認識部253は、例えば、CPU、ROM、RAM等で構成され、音声入力部251により自動的に取り込まれリアルタイムに伝送されてくるユーザの会話から音声データを生成して、後述する音声会話制御部259に音声データを伝送するとともに、生成した音声データに基づいて音声認識を行う。音声認識部253は、音声認識の結果として、生成した音声データに対応する単語列を生成して、後述するキーワード検出部255に対して、生成した単語列を伝送する。また、音声認識部253は、生成した単語列を後述する表示部261に伝送してもよい。なお、音声認識部253については、以下で詳細に説明する。

【0066】

キーワード検出部255は、音声認識部253から伝送された単語列の中に、キーワード管理部257が管理しているキーワードが存在するか否かを判断し、キーワードが存在している場合には、そのキーワードに該当する単語をキーワードとして出力する。キーワード検出部255は、一つの単語列の中に複数のキーワードが存在した場合には、該当する全てのキーワードを同時に検出してよい。また、キーワード検出部255は、単語列の中から検出したキーワードを、後述する表示部261に伝送してもよい。

【0067】

10

20

30

40

50

例えば、音声認識部 253 から「最近の構造改革は骨抜きだ」という認識結果が伝送され、キーワード管理部 257 が管理しているキーワードの中に「構造改革」というキーワードが存在する場合には、キーワード検出部 255 は、「構造改革」をキーワードとして出力する。

【0068】

キーワード管理部 257 は、音声認識部 253 が生成した単語列の中から抽出されるべきキーワードを管理する。キーワード管理部 257 は、第1情報処理装置 20A のユーザが検索サーバ 14 を利用した検索利用履歴情報や、本実施形態に係る音声チャットシステムの利用履歴情報等を記憶部 267 に記憶しておき、これらの利用履歴情報に含まれる検索キーワードや、検出されたキーワード等に基づいて、管理するキーワードを選択する。
キーワード管理部 257 で管理されているキーワードは、キーワード検出部 255 が自由に参照することが可能である。また、キーワード管理部 257 は、管理しているキーワードを音声認識部 253 に対して提供することも可能であり、音声認識部 253 は、キーワード管理部 257 から提供されたキーワードに基づいて、後述する認識用データベースを更新してもよい。

10

【0069】

音声会話制御部 259 は、例えば、CPU、ROM、RAM 等で構成され、通信網 12 を介して接続されている第2情報処理装置 20B との間で行われる音声チャットを制御する。音声会話制御部 259 は、音声認識部 253 で生成された音声データの伝送を受け、通信部 265 を介して音声データを第2情報処理装置 20B へと送信するとともに、第2情報処理装置 20B から送信される音声データを、通信部 265 を介して受信し、音声出力部 263 へと伝送する。また、音声会話制御部 259 は、音声チャットが行われる毎に、音声チャットの行われた日時、音声チャット自体を識別する識別子および音声チャットの相手を表す識別子を関連付けて、音声チャットの利用履歴情報として記憶部 267 に記憶してもよい。

20

【0070】

なお、音声会話制御部 259 は、音声会話制御に特化したハードウェアにより構成されていてもよく、音声チャットプログラム等のアプリケーションプログラムとして提供されてもよい。

【0071】

30

表示部 261 は、キーワード検出部 255 から伝送されたキーワードを、第1情報処理装置 20A のディスプレイ等の表示装置を介して、第1情報処理装置 20A のユーザに対して表示する。また、表示部 261 は、音声認識部 253 から伝送された音声認識結果である単語列そのものを、表示してもよい。

【0072】

音声出力部 263 は、第2情報処理装置 20B のユーザが発した音声の音声データを受信し、第1情報処理装置 20A へと取り込むインターフェースであって、例えば、第1情報処理装置 20A に接続された出力装置であるスピーカやイヤフォン等から構成される。通信部 265 を介して受信された第2情報処理装置 20B からの音声データは、音声出力部 263 を介して、第1情報処理装置 20A のユーザへと出力される。

40

【0073】

通信部 265 は、例えば第1情報処理装置 20A に設けられた通信装置であって、第1情報処理装置 20A の音声認識部 253 と、キーワード管理部 257 と、音声会話制御部 259 と、認識単語辞書生成部 269 とが、通信網 12 を介して第1情報処理装置 20A の外部の装置等である検索サーバ 14 や第2情報処理装置 20B 等と行う情報の送受信を、仲介する。なお、通信部 265 は、通信網 12 を介さずに、第1情報処理装置 20A に直接接続されているその他の情報処理装置等に対して、情報の送受信を行うことも可能である。

【0074】

記憶部 267 は、例えば第1情報処理装置 20A に設けられたストレージ装置であって

50

、キーワード検出部 255 が検出したキーワードやキーワード管理部 257 が管理しているキーワード情報等のデータを記憶する。また、これらのデータ以外にも、音声認識部 253 が生成した音声データや単語列等の認識結果や、各種のデータベース等を記憶することも可能である。更に、これらのデータ以外にも、第1情報処理装置 20A が、何らかの処理を行う際に保存する必要が生じた様々なパラメータや処理の途中経過等を、適宜記憶することが可能である。この記憶部 267 は、音声認識部 253 や、キーワード検出部 255 や、キーワード管理部 257 や、音声会話制御部 259 や、認識単語辞書生成部 269 等が、自由に読み書きを行うことが可能である。

【0075】

認識単語辞書生成部 269 は、検索サーバ 14 から取得する検索キーワードリスト等を取得して、取得した検索キーワードリストから適切なキーワードのみを選択し、音声認識を利用する認識単語辞書を生成する。また、認識単語辞書生成部 269 は、取得した検索キーワードリストからキーワードの選択を行なうだけでなく、選択したキーワードに対して、当該キーワードの属性情報や関連するサブキーワードを付加したり、音声認識を利用する際の認識重み情報を付加したりしてもよい。この認識単語辞書生成部 269 については、以下で詳細に説明する。

【0076】

<音声認識部 253 について>

続いて、図 4 を参照しながら、本実施形態に係る音声認識部 253 について、詳細に説明する。図 4 は、本実施形態に係る音声認識部 253 を説明するためのブロック図である。

【0077】

図 4 に示したように、本実施形態に係る音声認識部 253 は、例えば、AD 变換部 301 と、特徴パラメータ抽出部 303 と、マッチング部 305 と、共通辞書記憶部 307 と、認識用データベース 309 と、を備える。

【0078】

AD 变換部 301 は、音声入力部 251 から伝送されたアナログ信号である音声信号をサンプリングして量子化し、デジタル信号である音声データへと変換する。音声データへの変換が終了すると、AD 变換部 301 は、生成した音声データを、音声会話制御部 259 および特徴パラメータ抽出部 303 へと伝送する。また、AD 变換部 301 は、生成した音声データを、記憶部 267 に記憶してもよい。

【0079】

特徴パラメータ抽出部 303 は、AD 变換部 301 にて生成された音声データについて、適当なフレームごとに、例えば、メル周波数ケプストラム係数 (Mel Frequency Cepstrum Coefficients : MFCC) とその時間差分パラメータ、スペクトル、パワー線形予測係数、ケプストラム係数、線スペクトル対などの特徴パラメータを抽出する。特徴パラメータの抽出が終了すると、特徴パラメータ抽出部 303 は、抽出した特徴パラメータを、後述するマッチング部 305 へと伝送する。なお、特徴パラメータ抽出部 303 は、音声データから抽出した各種の特徴パラメータを、記憶部 267 に記憶してもよい。

【0080】

マッチング部 305 は、特徴パラメータ抽出部 303 から伝送された各種の特徴パラメータに基づき、後述する共通辞書記憶部 307 や認識用データベース 309 等を適宜参照しながら、音声入力部 251 に入力された音声（すなわち、音声チャットにおける会話内容）に最も近い単語列を、音声認識結果として生成する。単語列の生成方法については、以下で改めて説明する。単語列の生成が終了すると、マッチング部 305 は、生成した単語列を、キーワード検出部 255 へと伝送する。また、マッチング部 305 は、生成した単語列を、記憶部 267 に記憶してもよい。

【0081】

なお、上記のマッチング部 305 は、例えば、ビームサーチに基づくビタビ (Vite

10

20

30

40

50

r b i) デコーダや、A^{*}探索に基づくスタックデコーダなどの手法を用いてマッチングを行なうことも可能であり、例えば、いわゆるキーワードスポットティング等の手法を用いてマッチングを行なうことも可能である。また、マッチング部305が参照する各種の単語辞書に、後述する「認識重み」情報が付加されている場合には、後述する言語スコアに重みをつけた上で認識結果の順位付けを行なうことも可能である。

【0082】

共通辞書記憶部307は、音声認識で常に使用される単語の辞書である共通辞書を記憶する。共通辞書記憶部307に記憶されている共通辞書には、共通辞書に登録されている全ての単語について、発音情報とカテゴリ情報とがそれぞれ関連付けられて記述されている。例えば、固有名詞である「イチロー（人名）」が共通辞書に登録される場合には、「いちろう」という発音情報（音韻情報）と、“_人名_”というカテゴリとが、「イチロー」という固有名詞にそれぞれ関連付けられて登録される。なお、共通辞書の詳細については、以下で改めて説明する。

【0083】

認識用データベース309は、マッチング部305が単語列の生成に用いる各種のモデルや規則等を記憶しているデータベースである。この認識用データベース309については、以下で詳述する。

【0084】

<認識用データベース309について>

続いて、図5を参照しながら、本実施形態に係る認識用データベース309について、詳細に説明する。図5は、本実施形態に係る認識用データベース309を説明するための説明図である。

【0085】

図5に示したように、本実施形態に係る認識用データベース309は、例えば、認識単語辞書記憶部401と、音響モデル記憶部403と、言語モデル記憶部405と、音韻リスト407と、カナ音韻変換規則409と、を含む。

【0086】

認識単語辞書記憶部401は、マッチング部305が単語列を生成する際に用いる認識単語辞書を記憶する。認識単語辞書は、例えば、固定単語辞書と、可変単語辞書と、カテゴリテーブルとから構成されている。

【0087】

固定単語辞書には、単語登録および単語削除の対象外の単語、すなわち、予めシステムに設定されている単語（以下、適宜、固定単語と称する。）についての発音（音韻系列）と、音韻の連鎖関係を記述したモデル等の、各種の情報が記述されている。

【0088】

また、可変単語辞書には、キーワード単語についての発音や、音韻の連鎖関係を記述したモデル等の、各種の情報が記述されている。キーワード単語の登録・削除や発音の変更といった処理は、主に、この可変単語辞書に登録されている単語に対して行われる。また、可変単語辞書には、何も記憶されていなくともよい。

【0089】

例えば野球に関するキーワードを検出したい場合、共通辞書記憶部307には日常会話で通常使われる単語（例えば、国語辞典等の一般的な辞書に記載されているような単語等）を登録し、認識単語辞書記憶部401内の固定単語辞書には野球という分野で一般的に使われる「プレイ」や「ゲッター」「代打」「ベンチ入り」などの単語を登録する。また、認識単語辞書記憶部401内の可変単語辞書には「イチロー」などの時代とともに変化する固有名詞などを登録しておき、可変単語辞書を随時更新することで、最新の野球の話題のキーワードを含む認識が容易に実行できる。

【0090】

続いて、図6および図7を参照しながら、本実施形態に係る固定単語辞書について説明する。図6および図7は、本実施形態に係る固定単語辞書の一例を説明するための説明図

10

20

30

40

50

である。

【0091】

図6において、「シンボル」は単語を識別するための文字列であり、例えば、カナ表記などを用いて表すことができる。シンボルが同じエントリは、同じ単語のエントリであるとみなされる。また、本実施形態に係る言語モデルは、このシンボルを用いて表されている。また、「トランスクリプション」は、単語の表記を表し、認識結果として出力される文字列はこのトランスクリプションである。「音韻系列」は、単語の発音を音韻系列で表したものである。また、本実施形態に係る固定単語辞書は、図6に示したようなカナ表記のシンボルだけでなく、図7に示したような、漢字やひらがなの混ざった文字列についても記述されている。

10

【0092】

なお、図6および図7に記載されている「<先頭>」と「<終端>」とは特殊なシンボルであり、それぞれ「発話前の無音」と「発話後の無音」を表している。従って、トランスクリプションでは対応する表記は存在せず、「[]」(空欄)として表される。

【0093】

また、図6および図7に記載されている項目以外にも、本実施形態に係る固定単語辞書は、例えば、名詞や動詞といった単語の品詞やジャンル等を記述した「属性」欄や、以下で説明する言語スコアの算出の際に用いられる単語の重み付け情報を記述した「認識重み」欄等を備えても良い。

20

【0094】

続いて、図8を参照しながら、本実施形態に係る可変単語辞書について、説明する。図8は、本実施形態に係る可変単語辞書の一例を説明するための説明図である。

【0095】

図8に示したように、本実施形態に係る可変単語辞書には、例えば、「シンボル」欄と「カナ発音」欄が設けられる。また、図8に示した例では、「シンボル」として、カテゴリである“_人名_”と“_地名_”が記述されており、カテゴリ“_人名_”には、二つの単語の発音「イチロー」、「サニータロウ」がカナ発音で記述され、カテゴリ“_地名_”には、一つの単語の発音「キタシナガワ」がカナ発音で記述されている。また、本実施形態に係る可変単語辞書には、図8に示したようなカテゴリだけでなく、具体的な単語についても記述される。可変単語辞書に図示のようなカテゴリではなく具体的な単語が記載される場合には、「シンボル」欄に個々の単語が記載され、それらの単語に対応する「トランスクリプション」や「音韻系列」等の欄が設けられることとなる。

30

【0096】

また、本実施形態に係る可変単語辞書には、固定単語辞書の同様に、例えば、名詞や動詞といった単語の品詞やジャンル等を記述した「属性」欄や、以下で説明する言語スコアの算出の際に用いられる単語の重み付け情報を記述した「認識重み」欄等を備えても良い。可変単語辞書に記述される単語は、時代とともに変化する単語が主に記述されるため、「認識重み」欄に記載される値を、固定単語辞書に記載される「認識重み」よりも高い値に設定することが好ましい。このように、可変単語辞書に記載されている単語に高い値の「認識重み」を設定することで、時代に即した話題について会話することが多い音声チャットにおいて、特定のキーワードや当該特定のキーワードに関連した単語を優先的に選択することが可能となり、高い認識率で音声認識を行なうことが可能となる。

40

【0097】

なお、本実施形態に係る可変単語辞書の生成・更新処理については、以下で改めて詳細に説明する。

【0098】

カテゴリテーブルには、以下で説明する言語モデル記憶部405に記憶されている言語モデルに含まれるカテゴリと、そのカテゴリに含まれる単語の情報とが、対応づけて記述される。このカテゴリテーブルは、リスト状であってもよく、テーブルとして記載されていてもよい。なお、カテゴリテーブルは、言語モデルがカテゴリを含まない場合には、何

50

も記憶しなくてもよい。

【0099】

上記のカテゴリは、意味属性に基づく分類（例えば、“_人名_”、“_ユーザ名_”、“_地名_”、“_店名_”等）だけではなく、品詞に基づく分類（例えば、“_名詞_”、“_動詞_”、“_助詞_”等）であってもよい。なお、以下、“_・・・_”という表記は、カテゴリ名を表すものとする。

【0100】

図9は、カテゴリテーブルの例を示している。カテゴリテーブルは、以下で説明する言語モデルに使用されているカテゴリの種類と、そのカテゴリに属する単語の情報とが記述されている。例えば、言語モデルに、“_人名_”と“_地名_”の2個のカテゴリが使用されている場合には、カテゴリテーブルには、図9に示すように、“_人名_”と“_地名_”の2つのカテゴリがエントリされる。また、カテゴリテーブルに新たにカテゴリが追加される際には、追加されたカテゴリに属する単語は空欄でもよく、言語モデルの学習や適応化が進むにつれて、属する単語を追加していくことが可能である。なお、図9では、カテゴリ“_人名_”と“_地名_”に属する単語は、「[]」となっているが、この表記は、属する単語が何もないことを表している。

10

【0101】

図10は単語の情報がエントリされた認識単語辞書記憶部401のカテゴリテーブルの例を示している。図10において、第1行目のエントリは、カテゴリ“_人名_”に対して、単語「イチロー」および「サニータロウ」が属していることを表している。

20

【0102】

音響モデル記憶部403は、音声認識する音声の個々の音韻の音響的な特徴を表す音響モデルを記憶する。ここで、上記の音響モデルとは、母音や子音などの発音記号ごとに、音声の音響的な特徴を表現したモデルであり、入力音声の音声データと認識候補の単語とを、音としての類似性に着目して照合する際に用いられる。音響モデルとしては、例えば、隠れマルコフモデル(Hidden Markov Model: HMM)を用いることが可能であるが、本実施形態に係る音響モデルは、上記のHMMに規定されない。

【0103】

また、音響モデル記憶部403に記憶される音響モデルとして、例えば、静かな環境用（静かな環境で高い認識率が出る音響モデル）と雑音環境用（騒がしい環境でもそれなりの認識率が出る音響モデル）との2種類を用意し、環境に応じてどちらかを参照することも可能である。

30

【0104】

言語モデル記憶部405は、共通辞書記憶部307や認識単語辞書記憶部401中の各単語辞書に登録されている各単語が、それぞれどのように連鎖する（接続する）かを示す情報（以下、適宜、連鎖情報と称する。）を記述している言語モデルを記憶する。記述方法としては、例えば、統計的な単語連鎖確率(n-gramやclass n-gram)、生成文法、有限状態オートマトン(Finite State Automaton:FSA)等がある。

40

【0105】

言語モデル記憶部405に記憶されている言語モデルは、単語についての連鎖情報の他に、単語を特定の観点で分類したカテゴリについての連鎖情報も含んでいる。例えば、「人名を表す単語からなるカテゴリ」を“_人名_”というシンボルで表し、「食品を表す単語からなるカテゴリ」を“_食品_”というシンボルで表す場合、言語モデルは、“_人名_”、“_食品_”についての連鎖情報（すなわち、カテゴリ同士の連鎖、カテゴリと辞書に予め記憶されている単語との連鎖等）も記述している。

【0106】

したがって、言語モデルに含まれない単語についても、連鎖情報を取得することができる。例えば、「イチロー」と「は（助詞）」の連鎖情報を取得する場合、言語モデルに「イチロー」についての連鎖情報が記述されていなくても、「イチロー」が“_人名_”と

50

いうシンボルで表されるカテゴリに属していることがわかれば、代わりに“_人名_”と「は」との連鎖情報を取得することによって、「イチロー」と「は」の連鎖情報を取得することができる。

【0107】

<言語モデルについて>

続いて、図11を参照しながら、本実施形態に係る言語モデルの具体例について説明する。図11は、本実施形態に係る言語モデルの一例を説明するための説明図である。

【0108】

図11では、言語モデルとして、統計言語モデルが用いられている。統計言語モデルは、単語の連鎖情報を条件付確率で記述したモデルである。図11の言語モデルでは、3つの単語1, 2, 3の並び、すなわち単語の3連鎖の確率を表すtri-gramが、統計言語モデルとして用いられている。10

【0109】

図11において、「P(単語3 | 単語1 単語2)」は、単語列中に「単語1」、「単語2」という並びがあった場合に、その次に「単語3」が出現する確率を表す。例えば、「<先頭> “_人名_”」という並びがあった場合に、その次に「は」が出現する確率は、「0.012」である。なお、この確率は、大量の雑談を記述したテキストを解析することにより、予め求めることが可能である。また、言語モデルとしては、tri-gramの他に、bi-gram(2連鎖の確率)やuni-gram(単語の出現確率)等も、必要に応じて用いることが可能である。20

【0110】

また、図11の言語モデルにおいて、単語の他に、カテゴリを用いて文法が記述されている。すなわち、図11において、「_人名_」、「_地名_」は、カテゴリ“_人名_”、“_地名_”を意味するが、これらのカテゴリを用いてtri-gramを記述することによって、人名や地名を表す単語が可変単語辞書に登録された場合に、その単語を音声認識部253で認識することが可能となる。25

【0111】

音韻リスト407は、認識用データベース309で使用する音韻記号の一覧である。音韻リスト407は、例えば、図12に示したように、1つの音韻(に相当するもの)を1つの記号で表したものである。例えば、図12の音韻リストにおいて、母音+コロン(例えば、“a:”等)は、長音を表し、“N”は、撥音(「ん」)を表す。また、“sp”、“silB”、“silE”、“q”は、全て無音を表すが、それぞれ「発話中の無音」、「発話前の無音」、「発話後の無音」、「促音(「つ」)」を表す。30

【0112】

カナ音韻変換規則409は、カナ文字列を音韻系列に変換するための規則である。このように、カナ音韻変換規則409を記憶することによって、共通辞書記憶部307や認識単語辞書記憶部401は、発音情報として、音韻系列とは独立であるカナ文字列を保持することができる。カナ音韻変換規則409は、例えば、図13に示したように、カナによる表記と、この表記に対応づけられた音韻の記号とからなり、存在しうる全てのカナ表記に対応して、音韻への変換規則が記載される。図13のカナ音韻変換規則409によれば、例えば、「イチロー」というカナ文字列は、“i / ch / r / o:”という音韻系列に変換される。40

【0113】

以上、本実施形態に係る第1情報処理装置20Aの機能の一例を示した。上記の各構成要素は、汎用的な部材や回路を用いて構成されていてもよいし、各構成要素の機能に特化したハードウェアにより構成されていてもよい。また、各構成要素の機能を、CPU等が全て行ってもよい。従って、本実施形態を実施する時々の技術レベルに応じて、適宜、利用する構成を変更することが可能である。

【0114】

なお、第2情報処理装置20Bの構成は、第1情報処理装置20Aの構成と実質的に同50

一であるので、説明を省略する。

【0115】

(音声認識部253における音声認識処理について)

続いて、図14を参照しながら、本実施形態に係る音声認識部253における音声認識処理について、詳細に説明する。図14は、本実施形態に係る音声認識処理の一例を説明するための流れ図である。

【0116】

以下で説明する音声認識処理は、第1情報処理装置20Aのユーザからマイクロフォン等の音声入力部251に音声が入力されたとき、開始される。

【0117】

音声入力部251で生成された音声信号は、まず、AD変換部301により、デジタル信号である音声データに変換され、特徴パラメータ抽出部303に伝送される(ステップS101)。この際、AD変換部301は、生成した音声データを、記憶部267に記憶してもよい。次に、音声データが伝送された特徴パラメータ抽出部303は、伝送された音声データからメルケプストラム等の特徴量を抽出する(ステップS103)。特徴パラメータの抽出には、例えば、ケプストラム分析や、線形予測分析や、MFCC係数等を用いた聴覚フィルタに基づく分析等、各種の分析方法を用いることが可能である。また、特徴パラメータ抽出部303は、音声データから抽出した各種の特徴パラメータを、記憶部267に記憶してもよい。

【0118】

特徴パラメータ抽出部303において特徴量が抽出されると、マッチング部305は、共通辞書記憶部307に記憶されている共通辞書や、認識用データベース309の認識単語辞書記憶部401に記憶されている固定単語辞書と可変単語辞書を参照し、それぞれの辞書においてシンボルで表される単語のいくつかを連結し、単語列を生成する(ステップS105)。その後、マッチング部305は、生成した単語列について、認識用データベース309中の音響モデル記憶部403に記憶されている音響モデルに基づいて、音響スコアを計算する(ステップS105)。ここで、音響スコアとは、音声認識結果の候補である単語列と入力音声とが、音として(音響的に)どれだけ近いかを表すスコアであり、音響スコアが高いほど、生成した単語列が、入力音声に音響的に近いことを意味する。

【0119】

特徴パラメータに基づいて音響スコアが計算されると、マッチング部305は、得られた音響スコアに基づいて、音響スコアの高い単語列を所定の個数選択する(ステップS107)。なお、マッチング部305は、算出した音響スコアや、選択した単語列等を、記憶部267に記憶してもよい。

【0120】

音響スコアの算出が終了すると、マッチング部305は、ステップS107で選択した各単語列の言語スコアを、認識用データベース309中の言語モデル記憶部405に記憶されている言語モデルを用いて計算する(ステップS109)。例えば、言語モデル記憶部405に記憶されている各種の言語モデルの中から、文法や有限状態オートマンを使用している場合には、単語列がその言語モデルで受理することができるとき、言語スコアは「1」であり、受理することができないとき、言語スコアは「0」である。

【0121】

なお、マッチング部305は、生成された単語列を言語モデルが受理することができるとき、ステップS107で選択した単語列を残してもよく、生成された単語列を言語も出るが受理することができないとき、ステップS107で選択した単語列を削除してもよい。

【0122】

また、言語モデルとして、n-gramやclass n-gramのような統計言語モデルを使用している場合、その単語列の生成確率を言語スコアとする。この言語スコアを求める方法の詳細は、例えば、本出願人が先に提案した特願2001-382579号

10

20

30

40

50

に開示されている。なお、マッチング部305は、算出した言語スコア等を、記憶部267に記憶してもよい。

【0123】

なお、本実施形態に係る言語モデルは、共通辞書記憶部307に記憶されている共通辞書や、認識用データベース309内の認識単語辞書記憶部401に記憶されている固定単語辞書や可変単語辞書を用いて言語スコアの算出を行うが、可変単語辞書には、時代とともに変化する固有名詞などが随時更新されながら記述されているために、音声チャット等で話題になることが多い単語についても、言語スコアの算出対象とすることが可能となる。

【0124】

音響スコアと言語スコアの双方の算出が終了すると、マッチング部305は、ステップS105で計算された音響スコアと、ステップS109で計算された言語スコアを統合して各単語列をソートし、例えば、統合したスコアの一一番大きい単語列を認識結果として決定する（ステップS111）。すなわち、マッチング部305は、例えば、音響モデルから得られた音響スコアと言語モデルから得られた言語スコアとの積や、音響スコアの対数と言語スコアの対数との和等を最大とするような単語列を、認識結果として決定する。また、言語モデルで用いた各種の単語辞書に、単語の認識重み情報が付加されている場合には、認識重み情報を加味した言語スコアを利用してもよい。

【0125】

これにより、音響的にも言語的にも最もふさわしい単語列が認識結果として決定される。最もふさわしい単語列が認識結果として決定されると、マッチング部305は、決定した認識結果を、キーワード検出部255へと伝達する（ステップS113）。また、マッチング部305は、決定した認識結果である単語列を、記憶部267に記憶してもよい。

【0126】

このように、本実施形態に係る音声認識処理では、言語スコアの算出の際に、ユーザが音声チャットで話題にしやすい最近のトピック（特定のトピック）に関連した語彙を記載した可変単語辞書を参照することで、最近のトピックに関連した語彙を優先的に認識することができ、通常の音声認識処理では認識することが困難な最近のキーワードや当該キーワードに関連する語彙の認識率を向上させることができる。また、認識された結果表示される情報は、最近話題の、すなわちユーザにとっても興味深い可能性の高い情報であることが多いため、仮に提示された情報が音声チャットにおけるユーザの会話と多少離れていても（すなわち、音声認識結果に間違いが生じたとしても）、音声チャットのユーザに会話の広がりを与えることができる。これにより、音声チャットを行っているユーザも、音声認識処理の間違いをある程度許容することが考えられる。

【0127】

<言語スコアの算出方法について>

続いて、図15を参照しながら、本実施形態に係る言語スコアの算出方法について説明する。図15は、本実施形態に係る言語スコアの計算式の一例を説明するための説明図である。図15では、マッチング部305が、図14のステップS109で、例えば、単語列「<先頭>イチロー は 何時 に 起きた の<終端>」を選択した場合の言語スコアを求める式について、示している。

【0128】

言語スコア「Score(<先頭>イチロー は 何時 に 起きた の<終端>)」は、式(1)に示すように、単語列「<先頭>イチロー は 何時 に 起きた の<終端>」の生成確率である。

【0129】

言語スコア「Score(<先頭>イチロー は 何時 に 起きた の<終端>)」の値は、正確には、式(2)に示すように、「P(<先頭>)P(イチロー | <先頭>)P(は | <先頭>イチロー)P(何時 | <先頭>イチロー は)P(に | <先頭>イチロー は 何時)P(起きた | <先頭>イチロー は 何時 に)P(の | <先頭>イチロー は 何時 に 起きた)」

10

20

30

40

50

一は何時に起きた) P (<終端> | <先頭>イチロー は 何時に起きたの)で求められるが、図16に示すように、言語モデル112は、tri-gramを用いているので、条件部分「<先頭>イチロー は」、「<先頭>イチロー は 何時」、「<先頭>イチロー は 何時に」、「<先頭>イチロー は 何時に 起きた」、および「<先頭>イチロー は 何時に 起きた の」は、直前の最大2単語「イチロー は」、「は 何時」、「何時に」、「に 起きた」、および「起きた の」にそれぞれ限定した条件付確率で近似する(式(3))。

【0130】

この条件付確率は、図11に示したような言語モデルを参照することによって求められるが、言語モデルは、シンボル「イチロー」を含んでいないので、マッチング部305は、認識用データベース309中の認識単語辞書記憶部401に記憶されているカテゴリーテーブルを参照して、シンボル「イチロー」で表される単語のカテゴリが、“_人名_”であることを認識し、「イチロー」を“_人名_”に変換する。10

【0131】

即ち、式(4)に示すように、「P(イチロー | <先頭>)」は、「P(_人名_| <先頭>) P(イチロー | _人名)」に変更され、「P(_人名_| <先頭>) / N」で近似される。なお、Nは、カテゴリーテーブルの“_人名_”のカテゴリに属している単語の数を表す。

【0132】

即ち、確率を $P(X | Y)$ という形式で記述した場合、単語 X がカテゴリ C に属する単語である場合、言語モデルから $P(C | Y)$ を求め、その値に、 $P(X | C)$ (カテゴリ C から単語 X が生成される確率) を掛ける。カテゴリ C に属する単語が全て等確率で生成されると仮定すれば、カテゴリ C に属する単語が N 個ある場合、 $P(X | C)$ は、 $1/N$ と近似できる。20

【0133】

例えば、カテゴリ “_人名_” にシンボル「イチロー」で表される単語のみが属している場合には、上記の「N」は「1」となる。したがって、式(5)に示すように、「P(は | <先頭>イチロー)」は、「P(は | <先頭>_人名_)」となる。また、「P(何時 | イチロー は)」は、式(6)に示すように、「P(何時 | _人名_ は)」となる。30

【0134】

上記のような言語スコアの算出方法を用いることにより、可変単語を含む単語列に対しても、言語スコアを計算することができ、可変単語を認識結果に出現させることが可能となる。

【0135】

なお、上述の例では、システムの起動時に共通辞書記憶部307の共通辞書には何も記憶されていない状態であるとしたが、共通辞書に、いくつかの単語が予め記憶されていてもよい。

【0136】

図16は、システムの起動時に、キーワード「イチロー」がカテゴリ “_人名_” にエントリされている場合の共通辞書の例を示している。図16において、システムの起動時には、カテゴリ “_人名_” に、カナ発音「イチロー」がエントリされているので、キーワード登録を行わなくても、キーワードを検出できる。40

【0137】

また、上述の例では、固定単語辞書に記憶されている単語は、言語モデルに記述されている単語であり、可変単語辞書に記憶される単語は、カテゴリに属する単語であるとしたが、カテゴリに属する単語の一部を、固定単語辞書に記憶してもよい。

【0138】

図17は、固定単語辞書の例を示し、図18は、起動時のカテゴリーテーブルの例を示している。即ち、図16のカテゴリーテーブルには、カテゴリ “_人名_” と、そのカテゴリ “_人名_” に属する単語のシンボル「イチロー」が予め登録されている。また、図1750

の固定単語辞書 131 には、シンボル「イチロー」と、そのシンボル「イチロー」で表される単語のトランスクリプション「イチロー」、および音韻系列“ i / ch / r / o : ”が予め登録されている。

【 0 1 3 9 】

この場合、単語「イチロー」は、カテゴリ「_人名_」に属するものとして音声認識処理が行われる。即ち、単語「イチロー」は、最初から人名として扱われることになる。しかし、単語「イチロー」は固定単語辞書に記憶されているため、削除したり、変更したりすることはできない。

【 0 1 4 0 】

このように、想定される単語を予め固定単語辞書に記憶しておくことによって、登録を行わずに、キーワードを認識することができる。

【 0 1 4 1 】

(認識単語辞書生成部 269 について)

続いて、図19を参照しながら、本実施形態に係る認識単語辞書生成部269について、詳細に説明する。図19は、本実施形態に係る認識単語辞書生成部を説明するためのブロック図である。

【 0 1 4 2 】

本実施形態に係る認識単語辞書生成部269は、図19に示したように、例えば、検索キーワードリスト取得部501と、キーワード選択部503と、サブキーワード取得部505と、キーワード情報付加部507と、を備える。

[0 1 4 3]

検索キーワードリスト取得部 501 は、通信網 12 を介して接続されている検索サーバ 14 から、例えば、当該検索サーバで検索された上位検索キーワードのリストを取得する。取得する上位検索キーワードは、検索サーバ 14 全体の上位検索キーワードであってもよく、特定の分野における上位検索キーワードであってもよい。また、取得する上位検索キーワードは、任意の個数を取得することが可能である。

【 0 1 4 4 】

特定の分野における上位検索キーワードリストを取得する場合には、検索キーワードリスト取得部501は、例えば、記憶部267に記憶されている音声チャットの利用履歴情報や認識したキーワードの履歴情報等を参照して、第1情報処理装置20Aのユーザの嗜好等を判断し、ユーザの嗜好に合致した分野の上位検索キーワードを取得してもよい。例えば、検索キーワードリスト取得部501は、音声チャットの利用履歴情報や認識したキーワードの履歴情報を参照して、PLSA(Probabilistic Latent Semantic Analysis)等により所定の次元を有する話題ベクトルの形に変換しておく一方で、検索サーバの上位検索キーワードに関してもPLSA等により話題ベクトルの形に変換して、履歴情報に基づく話題ベクトルと上位検索キーワードに基づく話題ベクトルとの比較を行い、履歴情報を基にして生成された話題ベクトルに類似したベクトルを有する上位検索キーワードを取得してもよい。

[0 1 4 5]

検索キーワードリスト取得部 501 は、上記のようにして検索サーバ 14 から取得した検索キーワードリストを、キーワード選択部 503 へと伝送する。また、検索キーワードリスト取得部 501 は、取得した検索キーワードリストを、記憶部 267 に記憶してもよい。

[0 1 4 6]

キーワード選択部 503 は、検索キーワードリスト取得部 501 が検索サーバ 14 から取得した検索キーワードリストの中から、第 1 情報処理装置 20A のユーザの嗜好に基づいて、キーワードの選択を行なう。キーワードの選択は、音声チャットの利用履歴情報や、キーワード検出部 255 が検出したキーワードの履歴情報等を参照して、PLSA 等により所定の次元を有する話題ベクトルの形に変換しておく一方で、取得した検索キーワードリストに關しても PLSA 等により話題ベクトルの形に変換して、履歴情報に基づく話題ベクトルの形に変換する。

題ベクトルと上位検索キーワードに基づく話題ベクトルとの比較を行い、履歴情報を基にして生成された話題ベクトルに類似したベクトルを有する検索キーワードを選択することが可能である。

【0147】

また、音声チャットの相手である第2情報処理装置20Bから、音声チャットの利用履歴情報や認識したキーワードの履歴情報を取得できる場合には、第2情報処理装置20Bから取得した履歴情報と、第1情報処理装置20Aに記憶されている上記の履歴情報との整合をとり、キーワードの選択を行なってもよい。この場合、両者の履歴情報の整合は、例えば、両者の履歴情報の和集合や積集合やXOR集合をとってもよく、第1情報処理装置20Aのユーザの嗜好に統一をとってもよく、第2情報処理装置20Bにおけるユーザの嗜好に統一をとってもよい。また、取得したキーワードに対して、後述する認識重み情報が付加されている場合には、両者の認識重みの最大値や最小値や平均値を、新たな認識重み情報としてもよい。10

【0148】

また、キーワード選択部503は、検索キーワードリスト取得部501が検索サーバ14から取得した検索キーワードリストの中から、可変単語辞書に加えるのにふさわしくない単語を、適宜削除してもよい。ここで、可変単語辞書に加えるのにふさわしくない単語とは、例えば、i)一般的な単語や、ii)一般常識から鑑みて不適切な単語や、iii)1音韻や2音韻しかないように音声認識が困難な単語、等がある。

【0149】

キーワード選択部503は、検索サーバ14から取得した検索キーワードリストよりキーワードを選択すると、選択したキーワードを、後述するサブキーワード取得部505と、キーワード情報付加部507とに伝送するとともに、認識単語辞書記憶部401に記憶されている可変単語辞書に、選択したキーワードを記述する。また、キーワード選択部503は、選択したキーワードを、記憶部267に記憶してもよい。20

【0150】

サブキーワード取得部505は、キーワード選択部503が選択したキーワードを、検索サーバ14により検索し、キーワード選択部503が選択したキーワードに関連するサブキーワードを取得する。あるキーワードを検索サーバ14により検索すると、複数の文書が検索結果として得られるが、サブキーワード取得部505は、得られた複数の文書から、例えばTF・IDF(Term Frequency - Inverted Document Frequency)等のような重み付け方法を用いて、サブキーワードを取得する。なお、検索サーバ14による検索結果からサブキーワードを取得する方法は、上記の方法に規制されるわけではなく、公知のあらゆる方法を用いることが可能である。例えば、「サッカー日本代表」という単語がキーワードとしてキーワード選択部503から伝送された場合には、サブキーワード取得部505は「サッカー日本代表」というキーワードにより検索サーバ14にて検索を行い、得られた文書の中から、例えば、「オシム」や「中田」といったキーワードを取得する。30

【0151】

検索サーバ14からのサブキーワードの取得が終了すると、サブキーワード取得部505は、認識単語辞書記憶部401に記憶されている可変単語辞書に、取得したサブキーワードを、検索に利用したキーワードに関連付けて記述する。また、サブキーワード取得部505は、取得したサブキーワードを、後述するキーワード情報付加部507へと伝送する。また、サブキーワード取得部505は、取得したサブキーワードを記憶部267に記憶してもよい。40

【0152】

キーワード情報付加部507は、キーワード選択部503が選択したキーワードに対して、言語スコアの算出の際に用いられる認識重み情報や、キーワードの品詞やジャンル等に関する属性情報を含むキーワード情報を関連づけて付加し、可変単語辞書に記述する。また、キーワードに関連づけられているサブキーワードに対しても、上記のキーワード情50

報を付加してもよい。

【0153】

認識重み情報は、言語スコアの算出の際に利用される補正係数（重み付け係数）である認識重みが記載されている情報であって、認識重みが大きい値であるほど言語スコアの補正値は大きな値となり、認識結果として採用されやすくなる。例えば、認識重みが10であるキーワードは、認識重みが1であるキーワード（すなわち、重み付けがなされていないキーワード）よりも10倍高い確率で音声認識される。

【0154】

認識重みの決定においては、例えば、検索サーバ14から取得した検索キーワードリストでの順位情報や、音声認識結果の出力中に現れるキーワードの頻度等を利用することが可能である。検索キーワードリストは、検索サーバ14における上位いくつかの検索キーワードであるため、例えば、一番検索されているキーワードから順に順位付けを行い、順位付けの上位のものから順に、所定の認識重みを決定することが可能である。具体的には、一番検索されているキーワードから順に所定の係数を付加し、付加した係数に正規化したキーワードの頻度を乗じることで、認識重みとすることが可能である。

10

【0155】

また、認識重みの決定において、音声認識結果のキーワードに関して第1情報処理装置20Aのユーザが検索サーバ14を用いて検索を行ったかどうかという情報や、第1情報処理装置20Aのユーザの嗜好等も利用することが可能である。

【0156】

20

また、キーワードの品詞やジャンル等に関する属性情報は、例えば、インターネット上に設けられている掲示板やサイトなどの記載内容や、形態素解析ツール等を利用して付加することが可能である。また、インターネット上に設けられている百科事典やWikpediaや国語辞典等を適宜利用することも可能である。

【0157】

なお、上記の検索キーワードリストの取得、サブキーワードの取得およびキーワード情報の付加においては、検索サーバ14の代わりに、通信網12を介して接続されている任意のサーバ等に記憶されているシソーラス、オントロジーデータベース、百科事典、国語辞書および形態素解析ツール等を利用してよい。また、検索サーバ14と、上記のシソーラス、オントロジーデータベース、百科事典、国語辞書、形態素解析ツール等とを併用してもよい。

30

【0158】

<認識単語辞書の生成・更新処理について>

続いて、図20を参照しながら、本実施形態に係る認識単語辞書生成部の動作について、詳細に説明する。図20は、本実施形態に係る認識単語辞書生成・更新処理を説明するための説明図である。

【0159】

まず、認識単語辞書生成部269の検索キーワードリスト取得部501は、検索サーバ14から、ユーザの嗜好に沿った上位検索キーワードが記載された検索キーワードリストを取得する。この検索キーワードリストには、例えば図20に記載したように、「サッカー日本代表」、「ワールドベースボールクラシック」、「王監督」、「ジーコ」、「歌詞」、「27」、「ティラミス」、「万座ビーチ」、「飲み会」、「シユート」、「卑猥」といったキーワードが記載されている。検索キーワードリスト取得部501は、取得したキーワードを記憶部267に記憶するとともに、認識単語辞書生成部269のキーワード選択部503へと伝送する。

40

【0160】

キーワード選択部503は、伝送された検索キーワードリストを参照して、登録にふさわしくない単語を削除する。上記の例の場合では、一般的な単語である「飲み会」、「シユート」、「歌詞」、「27」と、一般常識から鑑みて不適切な単語である「卑猥」を、検索キーワードリストから削除することとなる。キーワード選択部503は、これらの單

50

語が削除された検索キーワードリストを、可変単語辞書に追加記載する。また、可変単語辞書への追加記載に当たっては、キーワード選択部 503 は、図 20 に示した項目以外に、例えば、音韻系列やトランスクリプション等の項目に関しても記載を行う。

【0161】

続いて、サブキーワード取得部 505 は、可変単語辞書に追加記載されたこれらのキーワードを取得し、取得したこれらのキーワードに関して、検索サーバ 14 等により、サブキーワードを取得する。例えば、図 20 に示したように、「サッカー日本代表」というキーワードを基に検索サーバ 14 を検索することにより、「オシム」、「中田」、「巻」、「ジーコ」といったサブキーワードを取得して、「サッカー日本代表」というキーワードに関連づけて可変単語辞書に追加記載する。同様に、上記の他のキーワードに関しても、サブキーワードを取得して可変単語辞書に追加記載を行う。10

【0162】

次に、キーワード情報付加部 507 は、不要な単語が削除された検索キーワードに対して、検索サーバ 14 での検索順位が上位なものから順にソートし、所定の係数を割り当てる。キーワード情報付加部 507 は、例えば、検索キーワードリストの一一番上に位置しているキーワードから順に、10、9、8 . . . と係数を割り当て、当該キーワードの頻度を一般的な頻度の期待値で割った値を、上記の割り当てられた係数にかけることで、認識重みを算出する。また、上記のようにして得られた認識重みに対して、例えばシグモイド関数のようなものを更に掛けてもよい。キーワード情報付加部 507 は、例えば上記のようにして算出した認識重み情報を、それぞれのキーワードに対して付加して、可変単語辞書に追加記載する。20

【0163】

なお、上記の認識重み情報の算出方法は、あくまでも一例であって、本実施形態に係る認識重み情報の算出方法は、上記の方法に規定されるわけではなく、上記以外の任意の方法を認識重み情報の算出に利用することが可能である。

【0164】

なお、図 20 に示したように、固定単語辞書には、一般的な日常会話に用いられるような単語である、「食べる」、「飲む」、「ヒット」、「シート」、「選手」、「飲み会」といったような単語が、予め登録されている。また、固定単語辞書に予め登録されている単語は、上述のように一般的な単語であり、音声認識において優先的に認識される必要性は低いと考えられるため、認識重み情報は、1 として登録される。30

【0165】

また、キーワード情報付加部 507 は、形態素解析ツールや検索サーバ 14 上の掲示板等の情報から、それぞれのキーワードの品詞やジャンルといった属性情報を取得して、可変単語辞書に追加記載する。例えば、図 20 に示した「サッカー日本代表」というキーワードには、属性情報として「名詞：スポーツ」という情報が付加される。

【0166】

上記のような処理を行うことで、認識単語辞書生成部 269 は、隨時可変単語辞書の生成・更新処理を行うことが可能となる。本実施形態に係る第 1 情報処理装置 20A を初めて起動する際には、可変単語辞書には何も情報が記載されていない場合がある。そのため、認識単語辞書生成部 269 は、上記の方法に基づいて、可変単語辞書の生成を行う。また、可変単語辞書の更新処理は、任意に行うことが可能である。例えば、ある期間毎に定期的に可変単語辞書の更新処理を行ってもよく、第 1 情報処理装置 20A のユーザからの更新命令に基づいて、可変単語辞書の更新処理を行っても良い。40

【0167】

また、可変単語辞書の更新に際して、サブキーワードとして記録されている単語が高頻度で音声認識されている場合には、高頻度で音声認識されているサブキーワードを、新たにキーワードとして可変単語辞書に記憶してもよい。

【0168】

マッチング部 305 は、音声認識処理を実行する際に、認識用データベース 309 中の50

言語モデル記憶部 405 に記憶されている、統計言語モデルや文法モデル等に基づいて、各種辞書に記載されている内容を利用して言語モデルの算出を行う。可変単語辞書に記憶されている単語の言語スコアを算出する場合には、マッチング部 305 は、言語モデルに基づいて通常の方法で言語スコアを算出し、更に、キーワードに付加されている認識重みを算出した言語スコアに掛けることで、実際に音声認識に用いる言語スコアとする。可変単語辞書に記憶されている単語は、図 20 に示したように 1 以上の認識重みが付加されているため、共通辞書や固定単語辞書に記憶されている単語に比べて、相対的に高い値の言語スコアが算出されることとなる。そのため、可変単語辞書に記憶されている、最近話題になることが多い特定の単語が、高い確率で認識されることとなる。

【0169】

10

なお、上記の方法では、通常の方法で算出された言語スコアに対して認識重みを掛けることとなり、場合によっては、算出される言語スコアが 1 以上の値を有することとなる。しかしながら、音声認識処理では、言語スコアの絶対値よりも、それぞれのキーワードに関する相対的な言語スコアの順位付けが重要であるため、言語スコアが 1 以上の値を有してもよい。

【0170】

上記の認識単語辞書生成部 269 は、図 21A に示したように、音声チャットを行っている第 1 情報処理装置 20A および第 2 情報処理装置 20B にそれぞれ実装されていてもよい。また、図 21B に示したように、認識単語辞書生成部は、音声チャットを行っている第 1 情報処理装置 20A および第 2 情報処理装置 20B には実装されずに、情報処理装置 20A および 20B を仲介しているサーバ 18 内に、実装されてもよい。ここで、図 21B におけるサーバ 図 21B に示したように、サーバ 18 内に実装される場合には、サーバ 18 が、検索サーバ 14 から検索キーワードリストを取得し、取得した検索キーワードリストから上記の方法でキーワードを選択するとともに、選択したキーワードに対して上記の重み付け情報を付加する、認識単語辞書生成装置として機能することとなる。

20

【0171】

サーバ 18 内の認識単語辞書生成部 183 は、音声会話制御部 181 から取得した音声会話の内容に基づいて、検索サーバ 14 から検索キーワードリストを取得し、取得した検索キーワードリスト用いて上記の処理方法で単語の選択と認識重み情報の付加を行なうことで、可変単語辞書を生成する。その後、サーバ 18 内の認識単語辞書生成部 183 は、第 1 情報処理装置 20A および第 2 情報処理装置 20B それぞれに、作成した可変単語辞書を伝送する。それぞれの情報処理装置は、伝送された可変単語辞書をそれぞれの認識用データベース 309 に記憶することで、音声認識処理に用いることが可能である。また、第 1 情報処理装置 20A および第 2 情報処理装置 20B は、サーバ 18 から認識重み情報が付加された単語のリストを取得し、各情報処理装置において、取得した単語のリストに基づいて可変単語辞書を作成してもよい。

30

【0172】

なお、図 21A および図 21B に記載されている第 1 情報処理装置 20A や第 2 情報処理装置 20B には、それぞれの情報処理装置が備える処理部の一部のみを記載しているが、記載されている処理部以外にも、例えば、図 3～図 5 に記載されているような処理部を有することは言うまでもない。

40

【0173】

以上説明したように、本実施形態に係る音声チャットシステムは、検索サーバ 14 のデータベースを音声認識に用いられる単語辞書の生成に用いることにより、最近話題にされることが多いキーワードを音声認識の言語知識として利用することができる。また、シソーラスなどを用いて話題のキーワードに関連した単語の出現確率も高く設定することができ、これらの関連キーワードを用いて、音声認識を行うことが出来る。これにより、本実施形態に係る音声チャットシステムは、ユーザが話題にしやすい最近のトピックに関連した語彙を優先的に認識することによって認識率を向上させることができる。また、認識された結果表示される情報は、最近話題の、すなわちユーザにとっても興味深い可能性の高

50

い情報が多いため、仮に提示された情報がユーザの会話と多少離れていても、ユーザには会話の広がりを与えるなどのメリットがある。

【0174】

(第2の実施形態)

以下に、本発明の第2の実施形態に係るキーワード検出システムについて、詳細に説明する。

【0175】

図22は、本実施形態に係るキーワード検出システムを説明するための説明図である。図22に示したように、本実施形態に係るキーワード検出システム11は、例えば、通信網12と、情報処理装置20A、20Bと、検索サーバ14と、を含む。また、情報処理装置20Aは、ホームネットワーク19内で、外部表示装置70と接続されている。さらに、情報処理装置20A、20Bと、外部表示装置70とは、放送局80から送信された放送情報を受信することが可能である。
10

【0176】

ここで、通信網12、検索サーバ14は、本発明の第1の実施形態に係る通信網12および検索サーバ14と同様の構成を有し、ほぼ同一の効果を奏するため、詳細な説明は省略する。

【0177】

情報処理装置20は、通信網12を介して接続された他の情報処理装置20と、音声によるチャット（音声チャット）を行う。また、情報処理装置20は、通信網12に接続されている検索サーバ14に対して、通信網12を介してデータベースの参照を要求したり、情報の検索を要求したりすることができる。また、音声チャットを行う他の情報処理装置20は、図示のように通信網12を介して接続されてもよく、また、通信網12を介さずに、例えば、USBポートや、i.LINK等のIEEE1394ポート、SCSIポート、RS-232Cポート等により直接接続されてもよい。
20

【0178】

また、情報処理装置20は、ホームネットワーク19内で、後述する外部表示装置70に接続されており、情報処理装置20と外部表示装置70との間で各種データの送受信を行うことが可能である。情報処理装置20と外部表示装置70との接続は、例えば、HDMI(High-Definition Multimedia Interface)-CEC(Consumer Electronics Control)等の接続ポートを介して行われる。また、情報処理装置20は、後述する放送局80から送信された放送情報を含む放送電波を、内部に設けられた受信機能を用いて受信することも可能である。
30

【0179】

なお、図示の例では、情報処理装置20として、デスクトップ型のPCを示しているが、本実施形態に係る情報処理装置20は、デスクトップ型PC、ノート型PCを問わない。また、本実施形態に係る情報処理装置20は、かかる例に限定されず、ネットワークを介した通信機能を有する機器であれば、例えば、テレビジョン受像器や家庭用ゲーム機等の情報家電、携帯電話、PDA等で構成することもできる。また、情報処理装置20は、契約者が持ち運びできるポータブルデバイス、例えば、携帯型ゲーム機、PHS、携帯型映像／音声プレーヤなどであってもよい。
40

【0180】

外部表示装置70は、CRTディスプレイ装置、液晶ディスプレイ装置、プラズマディスプレイ装置、ELディスプレイ装置等の表示装置であって、後述する放送局80から送信された放送情報を含む放送電波を受信して、当該表示装置の表示領域に表示する。ここで、放送局から送信された放送情報とは、放送文字情報や放送音声情報や画像情報等のデータを意味しており、放送情報を含む放送電波とは、ワンセグ（ワンセグメント放送）や12セグメントからなるディジタル放送（以下では、フルセグと略称する。）等のディジタル放送の電波を意味する。外部表示装置70は、HDMI-CEC等の接続ポートを介
50

して、受信した放送情報を情報処理装置20に送信することが可能である。また、外部表示装置70は、情報処理装置20から送信された各種データを受信して、当該データに対応した情報を表示することが可能である。

【0181】

放送局80は、放送文字情報や放送音声情報や画像情報等のデータから構成される放送情報を含む放送電波を送信する。外部表示装置70は、放送局80から送信された放送電波を受信し、放送電波に含まれる放送文字情報に基づいて字幕情報を表示したり、音声を出力したりする。また、情報処理装置20は、放送局80から送信された放送電波を受信して、各種の処理に利用することが可能である。

【0182】

(情報処理装置20の構成)

次に、本実施形態に係る情報処理装置20の構成について、詳細に説明する。なお、本実施形態に係る情報処理装置20のハードウェア構成は、本発明の第1の実施形態に係るハードウェア構成と実質的に同一であるため、詳細な説明は省略する。

【0183】

図23は、本実施形態に係る情報処理装置20の構成を説明するためのブロック図である。本実施形態に係る情報処理装置20は、例えば図23に示したように、キーワード管理部257と、表示部261と、通信部265と、記憶部267と、認識単語辞書生成部269と、放送文字情報受信部271と、キーワード検出部273と、外部表示装置接続制御部275と、を主に備える。

【0184】

キーワード管理部257は、後述する放送文字情報受信部271が受信した放送文字情報に対応した単語列の中から抽出されるべきキーワードを管理する。キーワード管理部257は、情報処理装置20のユーザが検索サーバ14を利用した検索利用履歴情報等を記憶部267に記憶しておき、これらの利用履歴情報に含まれる検索キーワードや、検出されたキーワード等に基づいて、管理するキーワードを選択する。キーワード管理部257で管理されているキーワードは、後述するキーワード検出部273が自由に参照することができる。また、キーワード管理部257は、管理しているキーワードを後述する認識単語辞書生成部269に対して提供することも可能であり、認識単語辞書生成部269は、キーワード管理部257から提供されたキーワードと、検索サーバ14から取得した検索キーワードリストと、を利用して、認識単語辞書の生成や更新を行うことができる。

【0185】

表示部261は、後述するキーワード検出部273から伝送されたキーワードを、情報処理装置20のディスプレイ等の表示装置を介して、情報処理装置20のユーザに対して表示する。また、表示部261は、放送文字情報受信部271が受信した放送文字情報に対応した単語列そのものを、表示してもよい。

【0186】

通信部265は、情報処理装置20に設けられた通信装置であって、情報処理装置20のキーワード管理部257と、認識単語辞書生成部269とが、通信網12を介して情報処理装置20の外部の装置等である検索サーバ14や他の情報処理装置20等と行う情報の送受信を、仲介する。なお、通信部265は、通信網12を介さずに、情報処理装置20に直接接続されているその他の情報処理装置等に対して、情報の送受信を行うことも可能である。

【0187】

記憶部267は、例えば情報処理装置20に設けられたストレージ装置であって、後述するキーワード検出部273が検出したキーワードやキーワード管理部257が管理しているキーワード情報等のデータを記憶する。また、これらのデータ以外にも、放送文字情報受信部271が受信した放送文字情報や、各種のデータベース等を記憶することも可能である。更に、これらのデータ以外にも、情報処理装置20が、何らかの処理を行う際に保存する必要が生じた様々なパラメータや処理の途中経過等を、適宜記憶することが可能

10

20

30

40

50

である。この記憶部 267 は、キーワード管理部 257、表示部 261、通信部 265、認識単語辞書生成部 269、放送文字情報受信部 271、キーワード検出部 273、外部表示装置接続制御部 275 等が、自由に読み書きを行うことが可能である。

【0188】

認識単語辞書生成部 269 は、検索サーバ 14 から取得する検索キーワードリスト等を取得して、取得した検索キーワードリストから適切なキーワードのみを選択し、音声認識を利用する認識単語辞書を生成する。また、認識単語辞書生成部 269 は、取得した検索キーワードリストからキーワードの選択を行なうだけでなく、選択したキーワードに対して、当該キーワードの属性情報や関連するサブキーワードを付加したり、音声認識を利用する際の認識重み情報を付加したりしてもよい。さらに、認識単語辞書生成部 269 は、キーワード管理部 257 から伝送されたキーワードや、後述する放送文字情報受信部 271 が受信した放送文字情報の中からキーワード検出部 273 が検出したキーワード等を用いて、認識単語辞書の生成や更新を行ってもよい。なお、本実施形態に係る認識単語辞書生成部 269 の詳細な構成や、認識単語辞書の生成方法等は、本発明の第 1 の実施形態に係る認識単語辞書生成部 269 の構成や生成方法と実質的に同一であるため、詳細な説明は省略する。10

【0189】

放送文字情報受信部 271 は、外部表示装置 70 が現在受信している放送チャンネルに関する受信チャンネル情報を、外部表示装置 70 から取得する。ここで、受信チャンネル情報とは、外部表示装置 70 が現在受信している放送チャンネルが何チャンネルかを表す情報である。放送文字情報受信部 271 は、この受信チャンネル情報を外部表示装置 70 から取得することにより、外部表示装置 70 が受信している放送チャンネルと、放送文字情報受信部 271 が受信する放送チャンネルとを同期させることができる。20

【0190】

また、放送文字情報受信部 271 は、取得した受信チャンネル情報に基づいて、外部表示装置 70 が受信している放送チャンネルに対応した放送電波を、放送局 80 から直接受信する。この際、放送文字情報受信部 271 は、放送局 80 が送信している放送電波のうちワンセグの電波を直接受信し、受信した電波の中から文字情報を取得してもよい。また、放送文字情報受信部 271 は、放送局 80 が送信しているフルセグの放送電波を直接受信してもよい。30

【0191】

また、放送文字情報受信部 271 は、取得した受信チャンネル情報に基づいて、外部表示装置 70 が受信し外部表示装置 70 から情報処理装置 20 に送信された放送文字情報を、受信してもよい。

【0192】

放送文字情報受信部 271 が受信した放送文字情報は、後述するキーワード検出部 273 へと出力される。また、放送文字情報受信部 271 は、受信した放送文字情報を記憶部 267 に記録してもよい。

【0193】

キーワード検出部 273 は、放送文字情報受信部 271 から伝送された放送文字情報に対応した単語列の中に、キーワード管理部 257 が管理しているキーワードが存在するか否かを判断し、キーワードが存在している場合には、そのキーワードに該当する単語をキーワードとして出力する。キーワード検出部 273 は、一つの単語列の中に複数のキーワードが存在した場合には、該当する全てのキーワードを同時に検出しててもよい。また、キーワード検出部 273 は、単語列の中から検出したキーワードを、表示部 261 に伝送してもよい。40

【0194】

例えば、放送文字情報受信部 271 から「最近の構造改革は骨抜きだ」という認識結果が伝送され、キーワード管理部 257 が管理しているキーワードの中に「構造改革」というキーワードが存在する場合には、キーワード検出部 273 は、「構造改革」をキーワー50

ドとして出力する。

【0195】

外部表示装置接続制御部275は、情報処理装置20に接続されている外部表示装置70と情報処理装置20との接続制御を行う。また、外部表示装置制御部275は、キーワード検出部273により抽出されたキーワード等を、外部表示装置70に送信する。また、外部表示装置70からある特定のキーワードやキーワードに関する記事について、外部表示装置70のユーザが参照したり選択したりしたことが通知されると、外部表示装置接続制御部267は、その参照履歴や選択履歴を記憶部267に記録するとともに、通信部265を介して検索サーバ14に通知してもよい。検索サーバ14は、これらの参照履歴や選択履歴を、検索サーバ14内で行われる各種の処理に利用することが可能である。なお、外部表示装置接続制御部275は、外部表示装置70に各種情報に対応したデータを送信する際に、送信予定のデータのデータ形式を、外部表示装置70が表示可能なデータ形式へと変換してもよい。10

【0196】

以上、本実施形態に係る情報処理装置20の機能の一例を示した。上記の各構成要素は、汎用的な部材や回路を用いて構成されていてもよいし、各構成要素の機能に特化したハードウェアにより構成されていてもよい。また、各構成要素の機能を、CPU等が全て行ってもよい。従って、本実施形態を実施する時々の技術レベルに応じて、適宜、利用する構成を変更することが可能である。20

【0197】

なお、本実施形態に係る情報処理装置20は、本発明の第1の実施形態に係る情報処理装置20が有する音声認識部や音声会話制御部を更に備えても良い。また、検索サーバ14に記録されている検索キーワードリストの内容を記憶するデータベース記憶部を更に備えてもよい。これにより、検索キーワードリストの内容が情報処理装置20に設けられたデータベース記憶部に格納されることとなり、情報処理装置20が検索サーバ14にアクセスするために要する時間を短縮することができる。また、データベース記憶部の内容を定期的に更新することで、検索サーバ14に記録されている最新の検索キーワードリストの内容を取得することが可能となる。20

【0198】

(外部表示装置70の構成)

30

続いて、図24および図25を参照しながら、本実施形態に係る外部表示装置70の構成について、詳細に説明する。図24は、本実施形態に係る外部表示装置の構成を説明するためのブロック図であり、図25は、本実施形態に係る外部表示装置の情報表示画面について説明するための説明図である。

【0199】

本実施形態に係る外部表示装置70のハードウェアは、CRTや、液晶パネル、プラスマディスプレイパネル、ELパネル等の表示素子と、CPU、ROM、RAM等から構成され、これらの表示素子を駆動制御するために用いられる駆動制御回路等から構成される。40

【0200】

また、本実施形態に係る外部表示装置70は、例えば図24に示したように、放送受信部701と、接続ポート制御部703と、放送情報表示制御部705と、キーワード情報表示制御部707と、表示部709と、を主に備える。

【0201】

放送受信部701は、放送局80から送信された放送電波を受信し、放送電波中に含まれる放送文字情報や放送音声情報や画像情報等のデータを、後述する接続ポート制御部703と、放送情報表示制御部705に出力する。また、放送受信部701は、外部表示装置70に備えられた記憶部(図示せず。)や、外部表示装置70に接続された記憶部(図示せず。)等に、受信した情報を記録してもよい。

【0202】

50

接続ポート制御部 703 は、情報処理装置 20 とディジタル通信が可能な HDMI - CEC 等の接続ポートを制御する。放送ポート制御部 703 は、放送受信部 701 が受信している放送チャンネルに関する受信チャンネル情報を、HDMI - CEC 等の接続ポートを介して情報処理装置 20 に送信する。また、放送受信部 701 が受信した放送文字情報や放送音声情報や画像情報等のデータは、接続ポート制御部 703 を介して情報処理装置 20 に出力されてもよい。また、情報処理装置 20 から送信された各種データは、接続ポート制御部 703 を介して外部表示装置 70 に入力される。情報処理装置 20 から送信された各種データは、後述するキーワード情報表示制御部 707 に伝送される。

【0203】

放送情報表示制御部 705 は、放送受信部 701 から伝送された放送文字情報、放送音声情報および画像情報を後述する表示部 709 に表示する際の表示制御を行う。 10

【0204】

キーワード情報表示制御部 707 は、情報処理装置 20 から送信されたキーワード情報を後述する表示部 709 に表示する際の表示制御を行う。ここで、情報処理装置 20 から送信されたキーワード情報とは、情報処理装置 20 により抽出されたキーワード、抽出されたキーワードの検索結果およびキーワードに関連する記事等の情報である。また、後述する表示部 709 に表示された情報を、外部表示装置 70 のユーザがマウスやキーボードやリモコン等の入力装置を用いて選択した場合に、選択された情報の詳細を取得するように表示部 709 から要請がなされると、キーワード情報表示制御部 707 は、この詳細情報取得要求を、接続ポート制御部 703 を介して情報処理装置 20 に送信する。 20

【0205】

表示部 709 は、放送情報表示制御部 705 により表示制御される放送文字情報、放送音声情報および画像情報と、キーワード情報表示制御部 707 により表示制御されるキーワード情報とを、外部表示装置 70 の情報表示画面に表示する。 30

【0206】

外部表示装置 70 の情報表示画面 751 は、通常は、放送文字情報、放送音声情報および画像情報等が表示される画像表示領域 753 が主に存在している。ここで、情報処理装置 20 からキーワード情報が伝送されると、情報表示画面 751 は、表示領域が分割され、キーワード情報が表示されるキーワード情報表示領域 755 が生成される。この画像表示領域 753 は、例えば放送情報表示制御部 705 により制御され、キーワード情報表示領域 755 は、例えばキーワード情報表示制御部 707 により制御される。 30

【0207】

なお、キーワード情報表示領域 755 の詳細については、本発明の第 1 の実施形態に係る情報処理装置における検索結果表示画面 50 と実質的に同一であるため、詳細な説明は省略する。 40

【0208】

以上、本実施形態に係る外部表示装置 70 の機能の一例を示した。上記の各構成要素は、汎用的な部材や回路を用いて構成されていてもよいし、各構成要素の機能に特化したハードウェアにより構成されていてもよい。また、各構成要素の機能を、CPU 等が全て行ってもよい。従って、本実施形態を実施する時々の技術レベルに応じて、適宜、利用する構成を変更することが可能である。 40

【0209】

(キーワード検出方法)

続いて、図 26 を参照しながら、本実施形態に係る情報処理装置 20 で行われるキーワード検出方法の一例について、詳細に説明する。図 26 は、本実施形態に係るキーワード検出方法について説明するための流れ図である。

【0210】

まず、情報処理装置 20 の放送文字情報受信部 271 は、外部表示装置 70 から取得した受信チャンネル情報に基づいて、外部表示装置 70 または放送局 80 から放送文字情報を受信し、放送字幕の文字列を文字情報データとして取得する（ステップ S201）。放 50

送文字情報受信部 271 は、取得した文字情報データを、キーワード検出部 273 に出力する。

【0211】

次に、キーワード検出部 273 は、伝送された文字情報データを形態素解析して文字列から単語を抽出し、情報処理装置 20 のメモリ上に記録する（ステップ S203）。

【0212】

次に、キーワード検出部 273 は、キーワード管理部 257 の内容を利用して、抽出した単語を検索する（ステップ S205）。

【0213】

続いて、キーワード検出部 273 は、キーワード管理部 257 の中に抽出単語が存在しているか否かを判定する（ステップ S207）。抽出単語がキーワード管理部 257 の中に存在している場合には、キーワード検出部 273 は、存在した抽出単語を、放送キーワードデータとしてメモリ上に記録する（ステップ S209）。また、抽出単語がキーワード管理部 257 の中に存在していない場合には、キーワード検出部 273 は、後述するステップ S211 を実行する。10

【0214】

次に、キーワード検出部 273 は、キーワード管理部 257 を用いて検索していない抽出単語がメモリ上に存在するか否かを判定する（ステップ S211）。メモリ上に未検索の単語が存在する場合には、ステップ S205 に戻って処理を実行する。また、メモリ上に未検索の単語が存在しない場合には、キーワード検出部 273 は、メモリ上に存在する放送キーワードデータを、抽出キーワードとして出力する（ステップ S213）。20

【0215】

以上説明したような方法で、本実施形態に係る情報処理装置 20 は、受信した放送文字情報の中からキーワードを抽出することが可能となる。

【0216】

以上、本実施形態に係る情報処理装置 20 では、放送局が送信している放送文字情報の中から、キーワード管理部 257 に存在している単語（キーワード）を自動的に抽出することが可能となり、抽出したキーワードを外部表示装置 70 に表示させることができる。かかる機能を用いることで、情報処理装置 20 や外部表示装置 70 でチャット等の双方向通信を実施しているユーザは、抽出されたキーワードをチャット等の話題として用いることが可能となる。30

【0217】

（本実施形態に係る情報処理装置の第 1 変形例）

続いて、図 27 および図 28 を参照しながら、本実施形態に係る情報処理装置 20 の第 1 変形例について、詳細に説明する。図 27 は、本変形例に係る情報処理装置 20 の構成を説明するためのブロック図であり、図 28 は、本変形例に係るキーワード検出方法を説明するための流れ図である。なお、本変形例に係る情報処理装置 20 のハードウェア構成は、本発明の第 2 の実施形態に係るハードウェア構成と実質的に同一であるため、詳細な説明は省略する。

【0218】

本変形例に係る情報処理装置 20 は、例えば図 27 に示したように、音声認識部 253 と、キーワード検出部 255 と、キーワード管理部 257 と、表示部 261 と、通信部 265 と、記憶部 267 と、認識単語辞書生成部 269 と、放送音声情報受信部 277 と、外部表示装置接続制御部 275 と、を主に備える。40

【0219】

本実施形態に係るキーワード管理部 257、表示部 261、通信部 265、記憶部 267 および外部表示装置接続制御部 275 については、本発明の第 2 の実施形態に係るキーワード管理部 257、表示部 261、通信部 265、記憶部 267 および外部表示装置接続制御部 275 と実質的に同一であり、それぞれ同様の効果を奏するため、詳細な説明は省略する。50

【0220】

音声認識部253は、後述する放送音声情報受信部277が受信した放送音声情報に基づいて、音声認識を行う。音声認識部253は、音声認識の結果として、放送音声情報に対応する単語列を生成し、後述するキーワード検出部255に対して、生成した単語列を伝送する。また、音声認識部253は、生成した単語列を表示部261に伝送してもよい。なお、音声認識部253の詳細な構成および音声認識方法は、本発明の第1の実施形態に係る音声認識部253の構成および音声認識方法と実質的に同一であるため、詳細な説明は省略する。

【0221】

キーワード検出部255は、音声認識部253から伝送された放送音声情報に対応した単語列の中に、キーワード管理部257が管理しているキーワードが存在するか否かを判断し、キーワードが存在している場合には、そのキーワードに該当する単語をキーワードとして出力する。キーワード検出部255は、一つの単語列の中に複数のキーワードが存在した場合には、該当する全てのキーワードを同時に検出してもよい。また、キーワード検出部255は、単語列の中から検出したキーワードを、表示部261に伝送してもよい。

10

【0222】

例えば、音声認識部253から「最近の構造改革は骨抜きだ」という認識結果が伝送され、キーワード管理部257が管理しているキーワードの中に「構造改革」というキーワードが存在する場合には、キーワード検出部255は、「構造改革」をキーワードとして出力する。

20

【0223】

認識単語辞書生成部269は、検索サーバ14から取得する検索キーワードリスト等を取得して、取得した検索キーワードリストから適切なキーワードのみを選択し、音声認識に利用する認識単語辞書を生成する。また、認識単語辞書生成部269は、取得した検索キーワードリストからキーワードの選択を行なうだけでなく、選択したキーワードに対して、当該キーワードの属性情報や関連するサブキーワードを付加したり、音声認識に利用する際の認識重み情報を付加したりしてもよい。さらに、認識単語辞書生成部269は、キーワード管理部257から伝送されたキーワードや、後述する放送音声情報受信部277が受信した放送音声情報に対応する単語列の中からキーワード検出部255が検出したキーワード等を用いて、認識単語辞書の生成や更新を行ってもよい。なお、本実施形態に係る認識単語辞書生成部269の詳細な構成や、認識単語辞書の生成方法等は、本発明の第1の実施形態に係る認識単語辞書生成部269の構成や生成方法と実質的に同一であるため、詳細な説明は省略する。

30

【0224】

放送音声情報受信部277は、外部表示装置70が現在受信している放送チャンネルに関する受信チャンネル情報を、外部表示装置70から取得する。放送音声情報受信部277は、この受信チャンネル情報を外部表示装置70から取得することにより、外部表示装置70が受信している放送チャンネルと、放送音声情報受信部277が受信する放送チャンネルとを同期させることができる。

40

【0225】

また、放送音声情報受信部277は、取得した受信チャンネル情報に基づいて、外部表示装置70が受信している放送チャンネルに対応した放送電波を、放送局80から直接受信する。この際、放送音声情報受信部277は、放送局80が送信している放送電波のうちワンセグの電波を直接受信し、受信した電波の中から音声情報を取得してもよい。また、放送音声情報受信部277は、放送局80が送信しているフルセグの放送電波を直接受信してもよい。

【0226】

また、放送音声情報受信部277は、取得した受信チャンネル情報に基づいて、外部表示装置70が受信し外部表示装置70から情報処理装置20に送信された放送音声情報を

50

、受信してもよい。

【0227】

放送音声情報受信部277が受信した放送音声情報は、音声認識部253へと出力される。また、放送音声情報受信部277は、受信した放送音声情報を記憶部267に記録してもよい。

【0228】

以上、本変形例に係る情報処理装置20の機能の一例を示した。上記の各構成要素は、汎用的な部材や回路を用いて構成されていてもよいし、各構成要素の機能に特化したハードウェアにより構成されていてもよい。また、各構成要素の機能を、CPU等が全て行つてもよい。従って、本実施形態を実施する時々の技術レベルに応じて、適宜、利用する構成を変更することが可能である。10

【0229】

なお、本変形例に係る情報処理装置20は、本発明の第1の実施形態に係る情報処理装置20が有する音声認識部や音声会話制御部を更に備えても良い。また、検索サーバ14に記録されている検索キーワードリストの内容を記憶するデータベース記憶部を更に備えてもよい。これにより、検索キーワードリストの内容が情報処理装置20に設けられたデータベース記憶部に格納されることとなり、情報処理装置20が検索サーバ14にアクセスするために要する時間を短縮することができる。また、データベース記憶部の内容を定期的に更新することで、検索サーバ14に記録されている最新の検索キーワードリストの内容を取得することが可能となる。20

【0230】

(キーワード検出方法)

続いて、図28を参照しながら、本変形例に係る情報処理装置20で行われるキーワード検出方法の一例について、詳細に説明する。

【0231】

まず、情報処理装置20の放送音声情報受信部277は、外部表示装置70から取得した受信チャンネル情報に基づいて、外部表示装置70または放送局80から放送音声情報を受信し、音声認識部253は、放送音声情報を音声認識し、音声認識結果を決定する(ステップS301)。30

【0232】

次に、キーワード検出部255は、伝送された文字情報データを形態素解析して文字列から単語を抽出し、情報処理装置20のメモリ上に記録する(ステップS303)。

【0233】

次に、キーワード検出部255は、キーワード管理部257の内容を利用して、抽出した単語を検索する(ステップS305)。

【0234】

続いて、キーワード検出部255は、キーワード管理部257の中に抽出単語が存在しているか否かを判定する(ステップS307)。抽出単語がキーワード管理部257の中に存在している場合には、キーワード検出部255は、存在した抽出単語を、放送キーワードデータとしてメモリ上に記録する(ステップS309)。また、抽出単語がキーワード管理部257の中に存在していない場合には、キーワード検出部255は、後述するステップS311を実行する。40

【0235】

次に、キーワード検出部255は、キーワード管理部257を用いて検索していない抽出単語がメモリ上に存在するか否かを判定する(ステップS311)。メモリ上に未検索の単語が存在する場合には、ステップS305に戻って処理を実行する。また、メモリ上に未検索の単語が存在しない場合には、キーワード検出部255は、メモリ上に存在する放送キーワードデータを、抽出キーワードとして出力する(ステップS313)。

【0236】

以上説明したような方法で、本実施形態に係る情報処理装置20は、受信した放送音声50

情報の中からキーワードを抽出することが可能となる。

【0237】

以上、本変形例に係る情報処理装置20では、放送局が送信している放送音声情報の中から、キーワード管理部257に存在している単語（キーワード）を自動的に抽出することが可能となり、抽出したキーワードを外部表示装置70に表示させることができる。かかる機能を用いることで、情報処理装置20や外部表示装置70でチャット等の双方向通信を実施しているユーザは、抽出されたキーワードをチャット等の話題として用いることが可能となる。

【0238】

以上、添付図面を参照しながら本発明の好適な実施形態について説明したが、本発明はかかる例に限定されることは言うまでもない。当業者であれば、特許請求の範囲に記載された範疇内において、各種の変更例または修正例に想到し得ることは明らかであり、それらについても当然に本発明の技術的範囲に属するものと了解される。10

【0239】

例えば、上述した実施形態においては、音声認識に用いられる各種の単語辞書が、共通辞書記憶部307と、認識用データベース309内の認識単語辞書記憶部401に記憶されている場合について説明したが、音声認識に用いられる各種単語辞書は、ある一つの記憶部内に記憶されていてもよい。

【0240】

また、上述した実施形態においては、認識単語辞書記憶部401に記憶される固定単語辞書、可変単語辞書、カテゴリーテーブルは、それぞれ1つずつである場合について説明しているが、固定単語辞書、可変単語辞書およびカテゴリーテーブルは、認識単語辞書記憶部401にそれぞれ複数記憶されていてもよい。認識単語辞書記憶部401にそれぞれ複数の固定単語辞書、可変単語辞書およびカテゴリーテーブルを記憶可能とすることで、例えば、ある特定の分野に特化した認識単語辞書を複数生成することが可能となる。また、ある特定のキーワードに関して、当該特定のキーワードに関連の深い単語の出現頻度を高くしたり、N-gramやclass-N-gramの確率を高くしたりすることで、特定のキーワードの認識率を高くしてもよい。20

【0241】

また、本明細書において説明した各種の処理方法は、必ずしも記載された順序に従って時系列的に行われる必要はなく、時系列的に処理されなくとも、または、並列的あるいは個別に実行されてもよい。30

【図面の簡単な説明】

【0242】

【図1】本発明の第1の実施形態に係る音声チャットシステムを説明するための説明図である。

【図2】同実施形態に係る情報処理装置のハードウェア構成を説明するためのブロック図である。

【図3】同実施形態に係る情報処理装置を説明するためのブロック図である。

【図4】同実施形態に係る音声認識部を説明するためのブロック図である。40

【図5】同実施形態に係る認識用データベースを説明するための説明図である。

【図6】同実施形態に係る固定単語辞書の一例を説明するための説明図である。

【図7】同実施形態に係る固定単語辞書の一例を説明するための説明図である。

【図8】同実施形態に係る可変単語辞書の一例を説明するための説明図である。

【図9】同実施形態に係るカテゴリーテーブルの一例を説明するための説明図である。

【図10】同実施形態に係るカテゴリーテーブルの一例を説明するための説明図である。

【図11】同実施形態に係る言語モデルの一例を説明するための説明図である。

【図12】同実施形態に係る音韻リストの一例を説明するための説明図である。

【図13】同実施形態に係るカナから音韻への変換規則を説明するための説明図である。

【図14】同実施形態に係る音声認識処理の一例を説明するための流れ図である。50

【図15】同実施形態に係る言語スコアの計算式の一例を説明するための説明図である。
 【図16】同実施形態に係る共通辞書の変形例を説明するための説明図である。
 【図17】同実施形態に係る固定単語辞書の変形例を説明するための説明図である。
 【図18】同実施形態に係るカテゴリテーブルの一例を説明するための説明図である。
 【図19】同実施形態に係る認識単語辞書生成部を説明するためのブロック図である。
 【図20】同実施形態に係る認識単語辞書生成・更新処理を説明するための説明図である。

。 【図21A】同実施形態に係る音声チャットシステムを説明するための説明図である。
 【図21B】同実施形態に係る音声チャットシステムの変形例を説明するための説明図である。

【図22】本発明の第2の実施形態に係るキーワード検出システムについて説明するための説明図である。

10

【図23】同実施形態に係る情報処理装置の構成を説明するためのブロック図である。
 【図24】同実施形態に係る外部表示装置の構成を説明するためのブロック図である。
 【図25】同実施形態に係る外部表示装置の情報表示画面について説明するための説明図である。

【図26】同実施形態に係るキーワード検出方法について説明するための流れ図である。
 【図27】同実施形態に係る情報処理装置の第1変形例の構成を説明するためのブロック図である。

【図28】同変形例に係るキーワード検出方法について説明するための流れ図である。

20

【符号の説明】

【0243】

1 0	音声チャットシステム
1 1	キーワード検出システム
1 2	通信網
1 4	検索サーバ
1 6	リムーバブル記録媒体
1 9	ホームネットワーク
2 0	情報処理装置
7 0	外部表示装置
8 0	放送局
2 5 1	音声入力部
2 5 3	音声認識部
2 5 5 , 2 7 3	キーワード検出部
2 5 7	キーワード管理部
2 5 9	音声会話制御部
2 6 1	表示部
2 6 3	音声出力部
2 6 5	通信部
2 6 7	記憶部
2 6 9	認識単語辞書生成部
2 7 1	放送文字情報受信部
2 7 3	キーワード検出部
2 7 5	外部表示装置接続制御部
2 7 7	放送音声情報受信部
3 0 1	A D 変換部
3 0 3	特徴パラメータ抽出部
3 0 5	マッチング部
3 0 7	共通辞書記憶部
3 0 9	認識用データベース

30

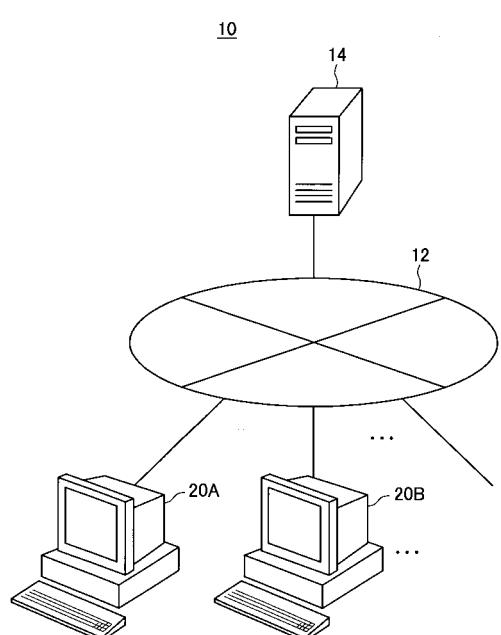
40

50

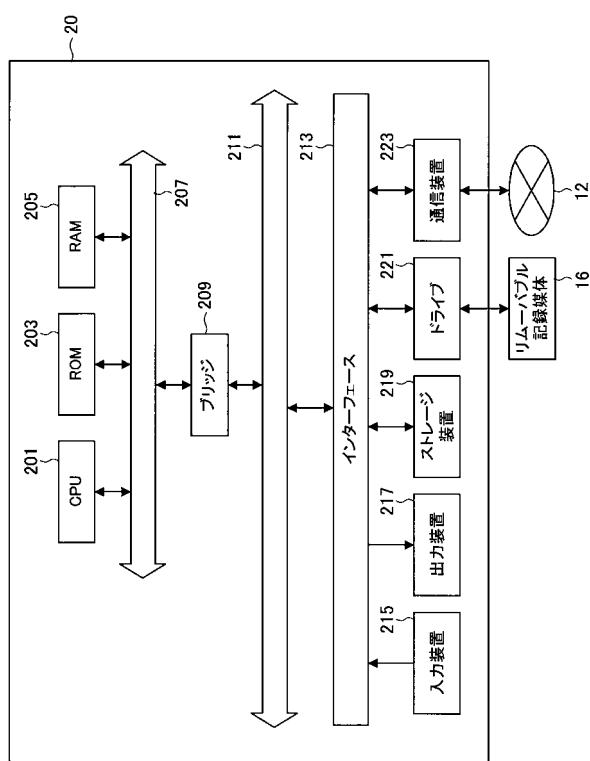
4 0 1	認識単語辞書記憶部
4 0 3	音響モデル記憶部
4 0 5	言語モデル記憶部
4 0 7	音韻リスト
4 0 9	カナ音韻変換規則
5 0 1	検索キーワードリスト取得部
5 0 3	キーワード選択部
5 0 5	サブキーワード取得部
5 0 7	キーワード情報付加部
7 0 1	放送受信部
7 0 3	接続ポート制御部
7 0 5	放送情報表示制御部
7 0 7	キーワード情報表示制御部
7 0 9	表示部
7 5 1	情報表示画面
7 5 3	画像表示領域
7 5 5	キーワード情報表示領域

10

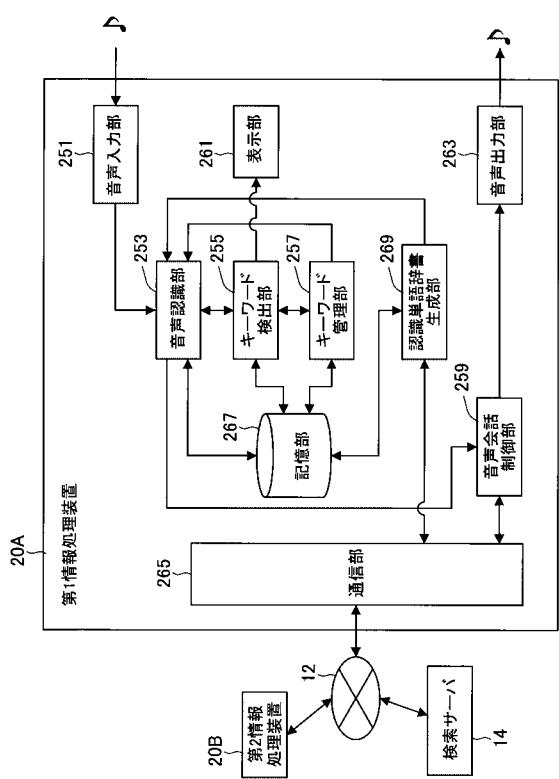
【図1】



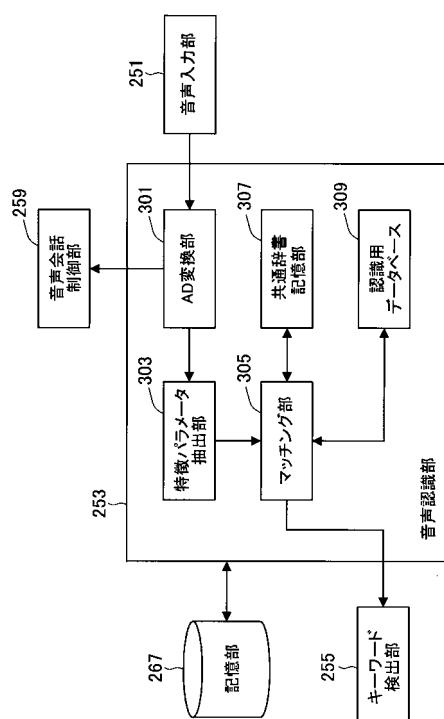
【図2】



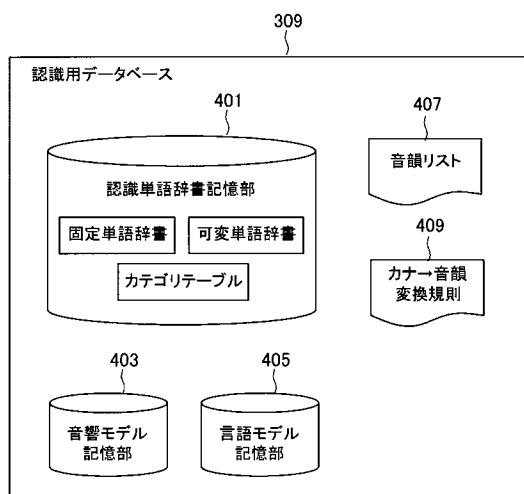
【図3】



【図4】



【図5】



【図6】

シンボル	トランスク립ション	音韻系列
<先頭>	[]	silB
<終端>	[]	silE
ア	[ア]	a
アー	[アー]	a:
イ	[イ]	i
イー	[イー]	i:
：		:
キヤ	[キヤ]	ky/a
キヤー	[キヤー]	ky/a:
：		:
ン	[ン]	N
ツ	[ツ]	t

【図7】

シンボル	トランスク립ション	音韻系列
<先頭>	[]	silB
<終端>	[]	silE
は	[は]	w/a
何時	[何時]	n/a/N/j/i
に	[ニ]	n/i
起きた	[起きた]	o/k/i/t/a
の	[の]	n/o
行きたい	[行きたい]	i/k/i/t/a/i
なあ	[なあ]	n/a
：		:

【図8】

カテゴリ	カナ発音
人名	イチロー
人名	サニータロウ
地名	キタシナガワ

【 図 9 】

カテゴリ	属する単語
人名	[]
地名	[]

【 义 1 0 】

カテゴリ	属する単語
人名	[イチロー、サニータロウ]
地名	キタシナガワ

【 図 1 1 】

単語1	単語2	単語3	P(単語3 単語1, 単語2)
<先頭>	<u>人名</u>	は	0.012
<u>人名</u>	は	何時	0.001
は	何時	に	0.001
何時	に	起きた	0.032
に	起きた	の	0.001
起きた	の	<終端>	0.004
<先頭>	<u>地名</u>	に	0.011
<u>地名</u>	に	行きたい	0.002
に	行きたい	なあ	0.010
行きたい	なあ	<終端>	0.009
:	:	:	:

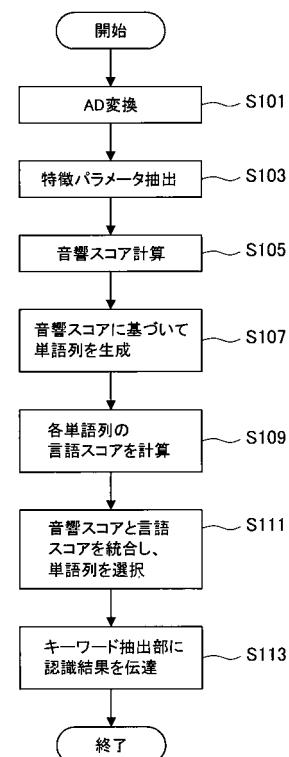
【 図 1 2 】

a,a:,i,i:,u,u;·,k,ky,g,gy,s,sh,z,j,t,ch,ts,·,y,w,N,sp,silB,silE,q

【 図 1 3 】

カナ	音韻
ア	a
アー	a:
イ	i
イー	i:
：	
エ	e
：	
キヤ	ky/a
キヤー	ky/a:
：	
ス	s/u
：	
ディ	d/i
ディー	d/i:
：	
ル	r/u
：	
ワ	w/a
ヲ	o
ン	N
ツ	t

【 図 1 4 】



【図15】

Score(<先頭> イチロー は 何時 に 起きた の <終端>)
= P(<先頭> イチロー は 何時 に 起きた の <終端>) ... (1)

= P(<先頭>)
P(イチロー | <先頭>)
P(は | <先頭> イチロー)
P(何時 | <先頭> イチロー は)
P(に | <先頭> イチロー は 何時)
P(起きた | <先頭> イチロー は 何時 に)
P(の | <先頭> イチロー は 何時 に 起きた)
P(<終端> | <先頭> イチロー は 何時 に 起きた の) ... (2)

= P(<先頭>)
P(イチロー | <先頭>)
P(は | <先頭> イチロー)
P(何時 | イチロー は)
P(に | は 何時)
P(起きた | 何時 | に)
P(の | に 起きた)
P(<終端> | 起きた の) ... (3)

P(イチロー | <先頭>)
=P(_人名_| <先頭>)P(イチロー | _人名_)
= P(_人名_| <先頭>)/N ... (4)

P(は | <先頭> イチロー)=P(は | <先頭> _人名_) ... (5)

P(何時 | イチロー | は)=P(何時 | _人名_ は) ... (6)

【図16】

カテゴリ	カナ発音
人名	イチロー
人名	マツイ

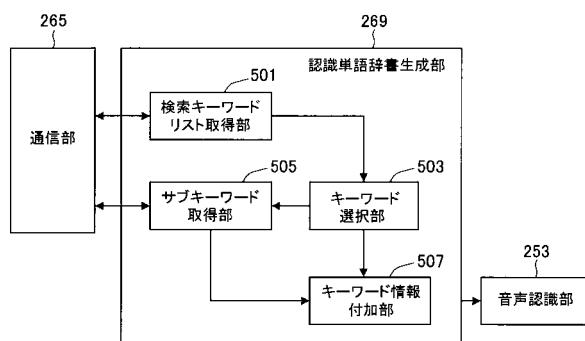
【図17】

シンボル	トランスクリプション	音韻系列
<先頭>	[]	silB
<終端>	[]	silE
何時	[何時]	n/a/N/j/i
に	[に]	n/i
起きた	[起きた]	o/k/i/t/a
の	[の]	n/o
イチロー	[イチロー]	i/ch/i/r/o:
松井	[松井]	m/a/ts/u/l

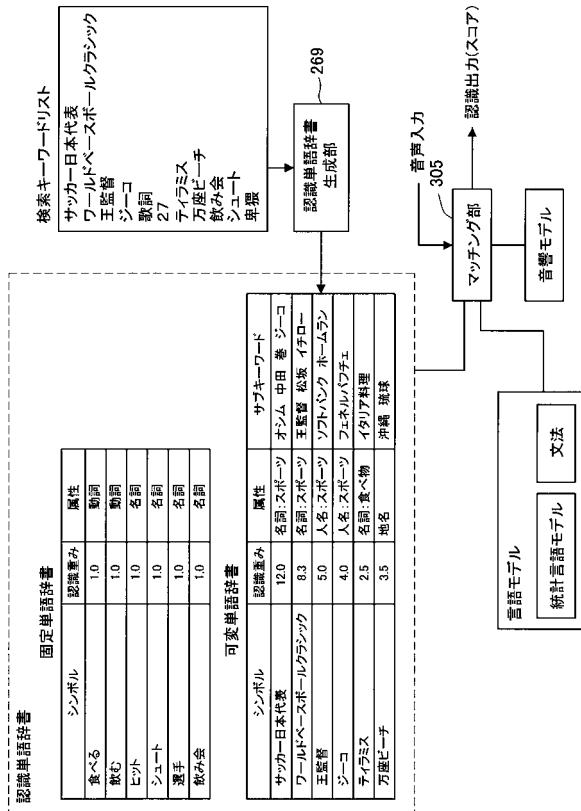
【図18】

カテゴリ	属する単語
人名	イチロー
人名	マツイ

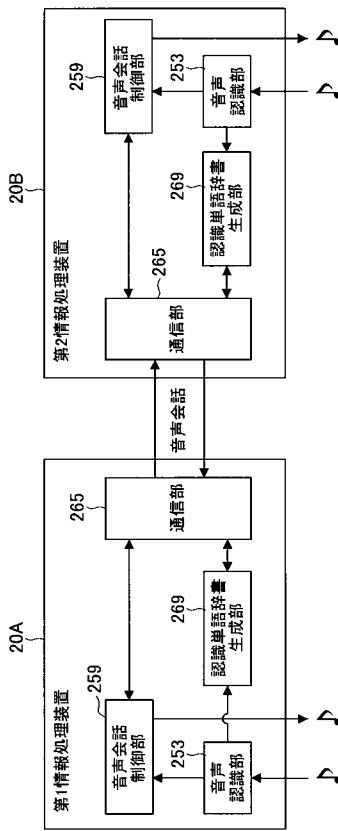
【図19】



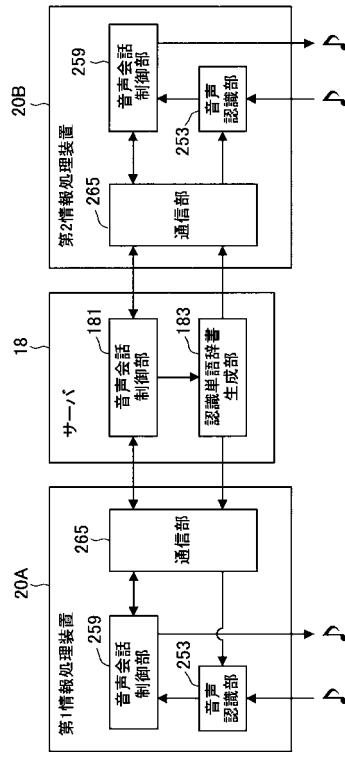
【図20】



【図 2 1 A】

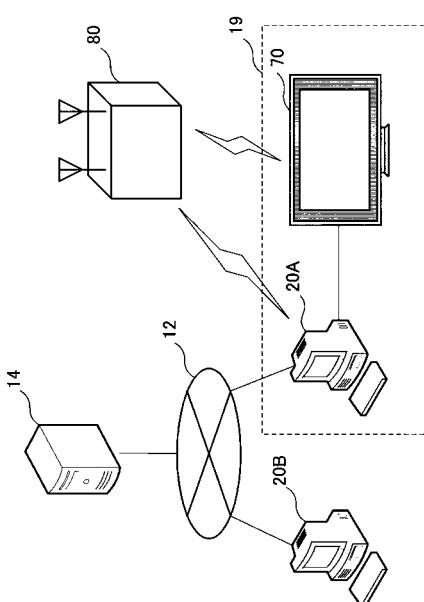


【図 2 1 B】

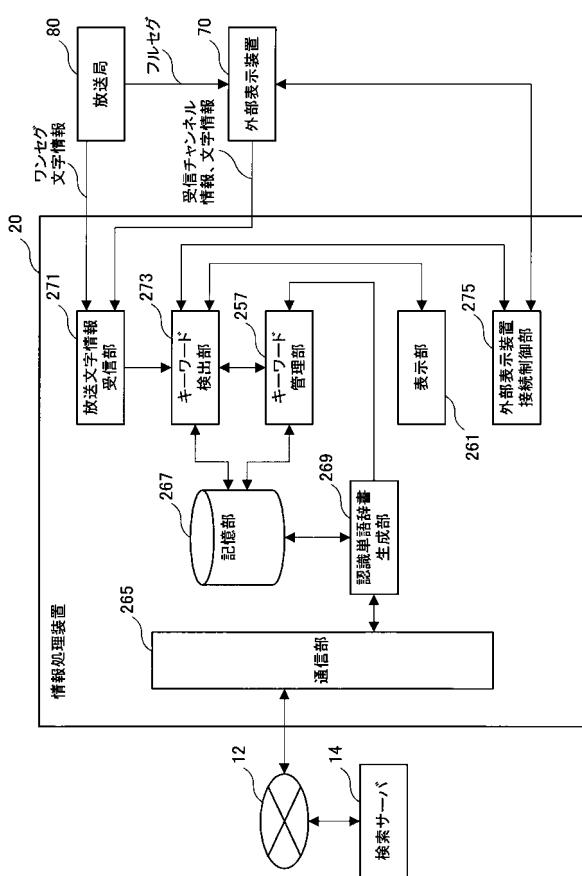


【図 2 2】

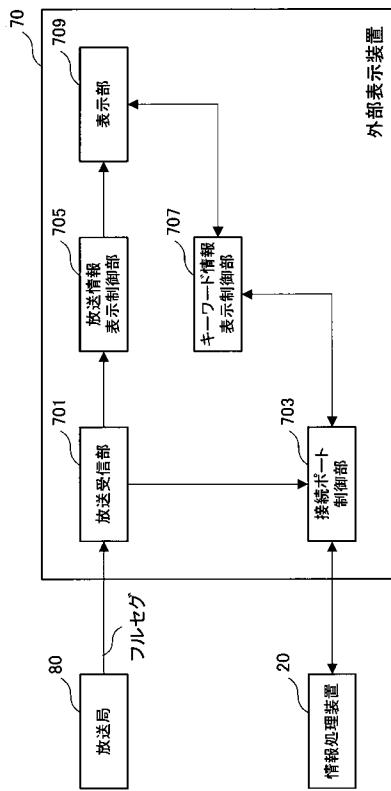
11



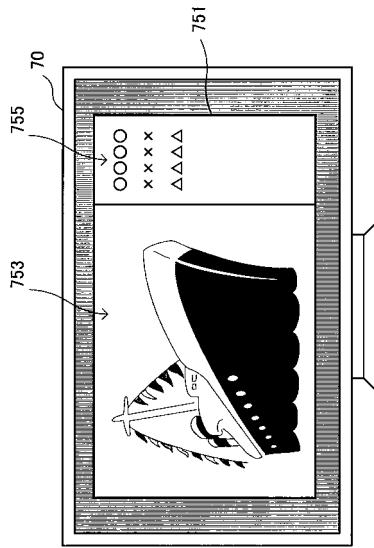
【図 2 3】



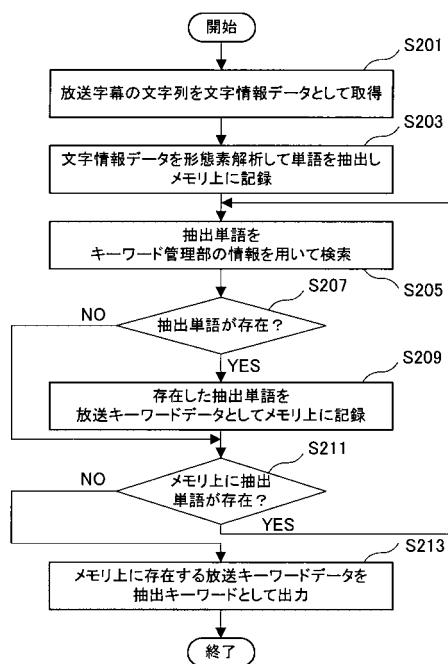
【図24】



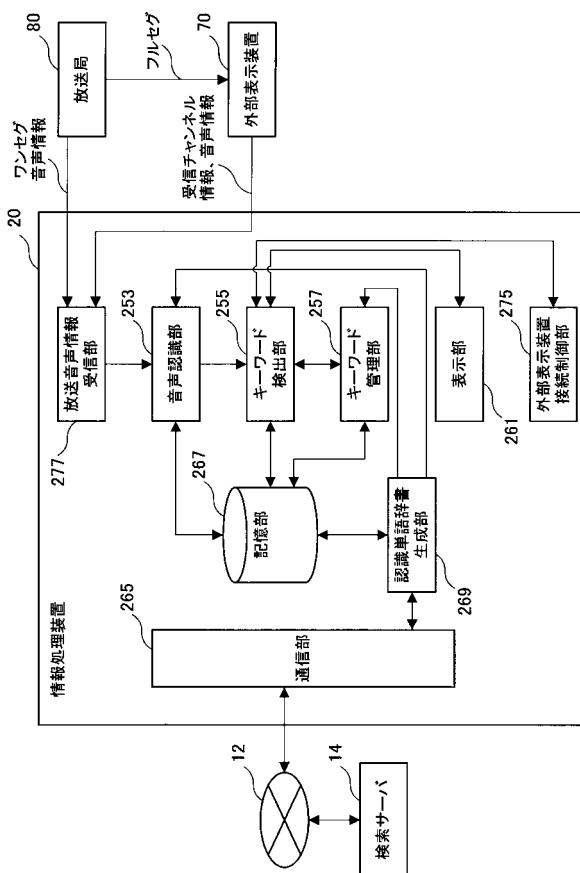
【図25】



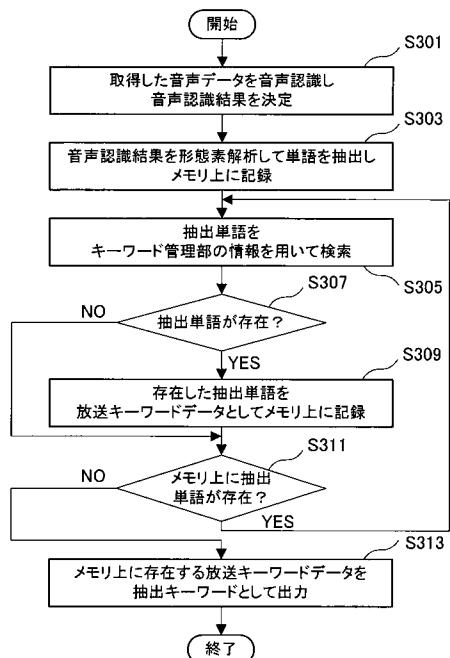
【図26】



【図27】



【図28】



フロントページの続き

(72)発明者 小川 浩明
東京都港区港南1丁目7番1号 ソニー株式会社内

(72)発明者 中出 元樹
東京都港区港南1丁目7番1号 ソニー株式会社内

(72)発明者 本田 等
東京都港区港南1丁目7番1号 ソニー株式会社内

(72)発明者 石塚 大介
東京都品川区大崎2丁目1番1号 ソネットエンタテインメント株式会社内

審査官 田部井 和彦

(56)参考文献 特開2005-237017(JP, A)
特開2002-245078(JP, A)
特開2003-323192(JP, A)
特開2004-295102(JP, A)
国際公開第2005/045806(WO, A1)
特開平09-186978(JP, A)
特開2006-319456(JP, A)
特開2008-225191(JP, A)
特開2004-267433(JP, A)
特表2007-514992(JP, A)
特開2005-309760(JP, A)
特開2003-186494(JP, A)
特開2005-241952(JP, A)
特開2004-252121(JP, A)

(58)調査した分野(Int.Cl., DB名)

G10L 15/00 - 17/00

G06F 17/30