

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7297266号
(P7297266)

(45)発行日 令和5年6月26日(2023.6.26)

(24)登録日 令和5年6月16日(2023.6.16)

(51)国際特許分類

F I

G 1 0 L	15/22	(2006.01)	G 1 0 L	15/22	2 0 0 H
G 1 0 L	15/00	(2013.01)	G 1 0 L	15/00	2 0 0 T
G 0 6 F	16/78	(2019.01)	G 1 0 L	15/22	4 5 3
G 0 6 F	3/16	(2006.01)	G 0 6 F	16/78	
			G 0 6 F	3/16	6 5 0

請求項の数 11 (全26頁) 最終頁に続く

(21)出願番号 特願2021-39268(P2021-39268)
 (22)出願日 令和3年3月11日(2021.3.11)
 (65)公開番号 特開2022-139053(P2022-139053
 A)
 (43)公開日 令和4年9月26日(2022.9.26)
 審査請求日 令和4年10月25日(2022.10.25)
 早期審査対象出願

(73)特許権者 398018021
 株式会社アドバンスト・メディア
 東京都豊島区東池袋3丁目1番1号 サ
 ンシャイン60 42F
 (74)代理人 100117260
 弁理士 福永 正也
 (72)発明者 鈴木 清幸
 東京都豊島区東池袋3丁目1番1号サン
 シャイン60 42階 株式会社アドバン
 スト・メディア内
 (72)発明者 大川 克利
 東京都豊島区東池袋3丁目1番1号サン
 シャイン60 42階 株式会社アドバン
 スト・メディア内
 (72)発明者 市川 竜平

最終頁に続く

(54)【発明の名称】 検索支援サーバ、検索支援方法及びコンピュータプログラム

(57)【特許請求の範囲】

【請求項1】

Webサイト上に公開される動画コンテンツの検索を支援する検索支援サーバであって、
 動画コンテンツの特定部分を示すタグ情報を選択するための発話を誘導するガイドデー
 タを選択的に表示するサイネージウィンドウと、

前記ガイドデータの中から、表示するべきガイドデータをユーザの発話により絞り込む
 ことが可能なスピーチウィンドウと

をデータ通信することが可能に接続されている端末装置に表示させ、

前記サイネージウィンドウ及び/又は前記スピーチウィンドウに表示された前記タグ情
 報のユーザによる発話を受け付けて、表示される前記タグ情報を絞り込み、絞り込まれた
 前記タグ情報の選択を受け付けることにより、選択を受け付けた前記タグ情報に対応する
 前記動画コンテンツの特定部分を前記端末装置において再生させることを特徴とする検索
 支援サーバ。

【請求項2】

前記タグ情報は、少なくとも前記動画コンテンツの特定部分を識別するタグ名称、前記
 タグ名称の内容を説明する単文又は単語群からなるテキストデータ及び前記タグ名称の属
 性を示す属性情報で構成されていることを特徴とする請求項1に記載の検索支援サーバ。

【請求項3】

前記サイネージウィンドウ及び/又は前記スピーチウィンドウに表示するガイドデータ
 の基礎となるサーチデータを生成するサーチデータ生成手段を備え、

該サーチデータ生成手段は、

前記動画コンテンツの内容に基づいて、内容を示す単文又は単語群を抽出して、抽出された単文又は単語群の選択を受け付けた場合に対応する前記動画コンテンツを再生する抽出・再生手段と、

再生される前記動画コンテンツに基づいて、前記タグ名称及び前記タグ名称の内容を示す単文又は単語群の入力を受け付けるタグ入力受付手段と、

再生される前記動画コンテンツに基づいて、前記動画コンテンツの特定部分の再生開始時間及び再生終了時間を含むポインタ情報を探索するポインタ探索手段と、

入力を受け付けた前記タグ名称ごとに、探索された前記ポインタ情報の割り付けを受け付けるポインタ割付受付手段と、

入力を受け付けた前記タグ名称ごとに、検索時に発話可能な前記属性情報の入力を受け付ける属性情報受付手段と

を備えることを特徴とする請求項 2 に記載の検索支援サーバ。

【請求項 4】

前記抽出・再生手段は、

前記動画コンテンツを精査して、含まれている音声を単文又は単語群として認識してテキストデータとして出力する音声文字起し手段と、

出力されたテキストデータの選択を受け付けた場合、選択を受け付けたテキストデータに対応する前記動画コンテンツを再生するコンテンツ再生手段と

を備えることを特徴とする請求項 3 に記載の検索支援サーバ。

【請求項 5】

前記ポインタ割付受付手段は、探索された前記ポインタ情報に含まれる、前記動画コンテンツの特定部分の再生開始時間及び再生終了時間について、前記動画コンテンツを再生しながら更新を受け付けるポインタ更新受付手段を備えることを特徴とする請求項 3 に記載の検索支援サーバ。

【請求項 6】

前記ガイドデータは、前記サーチデータに含まれる前記タグ情報を、前記サーチデータを識別するサーチデータ識別情報と対応付けて生成されることを特徴とする請求項 3 乃至 5 のいずれか一項に記載の検索支援サーバ。

【請求項 7】

前記スピーチウィンドウに表示され、ユーザによる選択を受け付けることが可能な前記タグ名称を絞り込むために、

ユーザにより発話された音声データの入力を受け付ける発話受付手段と、

入力を受け付けた音声データをテキストデータに変換し、変換したテキストデータの前記ガイドデータの前記タグ名称及び前記タグ名称の内容を示す単文又は単語群に対する一致度を算出し、算出された一致度が最大であるタグ名称を特定するとともに、前記一致度が所定値より大きい場合に前記タグ名称に対応する前記動画コンテンツの特定部分を表示するメタタグ音声認識手段と、

入力を受け付けた音声データに基づいて、前記ガイドデータの前記タグ情報のうち、前記属性情報に一致するものを抽出し、一致する一又は複数の属性情報の論理積による前記ガイドデータの絞り込みを行い、絞り込まれた前記ガイドデータの前記タグ名称及び前記属性情報を出力する絞り込み音声認識手段と

を備え、前記メタタグ音声認識手段及び前記絞り込み音声認識手段を並行して実行することを特徴とする請求項 3 乃至 6 のいずれか一項に記載の検索支援サーバ。

【請求項 8】

前記メタタグ音声認識手段は、変換されたテキストデータと、すべての前記タグ名称及び前記タグ名称の内容を示す単文又は単語群で形成された第一のデータセットとの一致度を算出し、

前記絞り込み音声認識手段は、すべてのタグ名称に対応付けられている前記属性情報で形成された第二のデータセットを音声認識フィルタとして用いた認識結果として抽出され

10

20

30

40

50

た前記属性情報の論理積により前記ガイドデータを絞り込むことを特徴とする請求項 7 に記載の検索支援サーバ。

【請求項 9】

前記絞り込み音声認識手段は、絞り込まれたガイドデータの前記属性情報により前記第二のデータセットを更新することを特徴とする請求項 8 に記載の検索支援サーバ。

【請求項 10】

Web サイト上に公開される動画コンテンツの検索を支援する検索支援サーバで実行することが可能な検索支援方法であって、

前記検索支援サーバが、

動画コンテンツの特定部分を示すタグ情報を選択するための発話を誘導するガイドデータを選択的に表示するサイネージウィンドウと、

前記ガイドデータの中から、表示すべきガイドデータをユーザの発話により絞り込むことが可能なスピーチウィンドウと

をデータ通信することが可能に接続されている端末装置に表示させる工程と、

前記サイネージウィンドウ及び/又は前記スピーチウィンドウに表示された前記タグ情報のユーザによる発話を受け付けて、表示される前記タグ情報を絞り込み、絞り込まれた前記タグ情報の選択を受け付けることにより、選択を受け付けた前記タグ情報に対応する前記動画コンテンツの特定部分を前記端末装置において再生させる工程と

を実行することを特徴とする検索支援方法。

【請求項 11】

Web サイト上に公開される動画コンテンツの検索を支援する検索支援サーバで実行することが可能なコンピュータプログラムであって、

前記検索支援サーバを、

動画コンテンツの特定部分を示すタグ情報を選択するための発話を誘導するガイドデータを選択的に表示するサイネージウィンドウと、

前記ガイドデータの中から、表示すべきガイドデータをユーザの発話により絞り込むことが可能なスピーチウィンドウと

をデータ通信することが可能に接続されている端末装置に表示させる手段、及び

前記サイネージウィンドウ及び/又は前記スピーチウィンドウに表示された前記タグ情報のユーザによる発話を受け付けて、表示される前記タグ情報を絞り込み、絞り込まれた前記タグ情報の選択を受け付けることにより、選択を受け付けた前記タグ情報に対応する前記動画コンテンツの特定部分を前記端末装置において再生させる手段

として機能させることを特徴とするコンピュータプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、Web サイトにおいてコンテンツ提供者が提供する多数の動画コンテンツの中から、検索主体であるユーザが、希望する動画コンテンツの特定のシーンを、発話により絞り込むことで確実に表示することが可能な検索支援サーバ、検索支援方法及びコンピュータプログラムに関する。

【背景技術】

【0002】

Web サイトには様々な情報が存在する。現状では、Web サイトのトップページのアイコンのクリックやクリック可能な言葉による説明、Google (R)、Yahoo (R) 等の検索エンジンを利用した検索窓を使うことで、ユーザが希望する内容を含むと考えられる Web ページへのアドレス、あるいは動画コンテンツそのものを抽出してユーザが視認可能な表示装置において再生している。

【0003】

また、ユーザの発話により検索キーの入力を受け付ける場合、Google (R)、Microsoft (R) 等が提供する音声認識エンジンを使用して、ユーザによる発話を

認識している。これらの商用の音声認識エンジンは、広汎に発話のデータを収集し、機械学習、あるいは深層学習により生成された汎用型の音声認識エンジンとして提供されている。

【 0 0 0 4 】

したがって、ユーザの発話を音声認識する場合に、コンテンツ提供者がユーザに対して提供を意図する検索キーを正しく認識することができない場合が生じるおそれがあった。そこで、例えば特許文献 1 には、複数の言語モデルを用いることで音声認識精度を高めている音声理解装置が開示されている。

【 0 0 0 5 】

また特許文献 2 には、ユーザに対して、番組やシーンを検索するためのキーとなる情報を画面上に表示しておき、ユーザの選択を受け付けるデジタル放送受信装置が開示されている。特許文献 2 では、コンテンツホルダが、検索情報を含むメタデータを作成して記憶しておき、該検索情報で検索される動画のシーン情報の一覧をユーザに表示することにより、ユーザが所望のシーンを選択して再生することが可能となっている。

10

【先行技術文献】

【特許文献】

【 0 0 0 6 】

【文献】特開 2 0 1 0 - 1 7 0 1 3 7 号公報
特開 2 0 0 9 - 0 5 5 6 4 3 号公報

【発明の概要】

20

【発明が解決しようとする課題】

【 0 0 0 7 】

特許文献 1 では、複数の音声認識エンジンを用いて並列に処理し、得られた複数の認識結果に基づく言語理解エンジンの処理結果を統合処理している。しかし、複数の認識結果及び言語理解処理結果の統合処理には相当の時間を要するだけではなく、統合処理後の認識結果の確度を保証することができないため現実的ではない。したがって、ユーザの発話によって、短時間で正しい検索キーを取得することができる保証がないという問題点があった。

【 0 0 0 8 】

また、特許文献 2 では、コンテンツホルダの作成する検索情報を含むメタデータが一覧表示される。しかし、特許文献 2 では表示されているメタデータの中から選択する必要があるのに対して、メタデータの表示領域には物理的な制約があることから表示可能なメタデータ数に制限がある。したがって、ユーザ所望のシーンに紐づいたメタデータが表示されないケースも想定され、ユーザが所望のシーンを再生するには、結局所望のシーンを再生しながら探し出す必要があり、相当の手間と時間を要するという問題点があった。また、一覧表示させるメタデータを作成するのも相当の手間と時間を要し、十分な数のメタデータを表示することができないので、メタデータの指示により再生されるシーンが必ずしもユーザの所望のシーンを再生しているという保証もない。

30

【 0 0 0 9 】

本発明は、上記事情に鑑みてなされたものであり、動画コンテンツの特定部分（シーン）を示すタグ情報を選択するための発話を誘導するガイドデータを選択的に表示するサイネージウィンドウと、全てのガイドデータを発話で絞り込んで表示することが可能なスピーチウィンドウを用いることで、ユーザが自己の目的に応じて絞り込んだガイドデータ通りに発話又は選択する限り、コンテンツ提供者がユーザに提供したい動画コンテンツの特定部分へ確実に誘導することが可能な検索支援サーバ、検索支援方法及びコンピュータプログラムを提供することを目的とする。

40

【課題を解決するための手段】

【 0 0 1 0 】

上記目的を達成するために本発明に係る検索支援サーバは、Web サイト上に公開される動画コンテンツの検索を支援する検索支援サーバであって、動画コンテンツの特定部分

50

を示すタグ情報を選択するための発話を誘導するガイドデータを選択的に表示するサインージウィンドウと、前記ガイドデータの中から、表示すべきガイドデータをユーザの発話により絞り込むことが可能なスピーチウィンドウとをデータ通信することが可能に接続されている端末装置に表示させ、前記サインージウィンドウ及び/又は前記スピーチウィンドウに表示された前記タグ情報のユーザによる発話を受け付けて、表示される前記タグ情報を絞り込み、絞り込まれた前記タグ情報の選択を受け付けることにより、選択を受け付けた前記タグ情報に対応する前記動画コンテンツの特定部分を前記端末装置において再生させることを特徴とする。

【0011】

また、本発明に係る検索支援サーバは、前記タグ情報は、少なくとも前記動画コンテンツの特定部分を識別するタグ名称、前記タグ名称の内容を説明する単文又は単語群からなるテキストデータ及び前記タグ名称の属性を示す属性情報で構成されていることが好ましい。

10

【0012】

また、本発明に係る検索支援サーバは、前記サインージウィンドウ及び/又は前記スピーチウィンドウに表示するガイドデータの基礎となるサーチデータを生成するサーチデータ生成手段を備え、該サーチデータ生成手段は、前記動画コンテンツの内容に基づいて、内容を示す単文又は単語群を抽出して、抽出された単文又は単語群の選択を受け付けた場合に対応する前記動画コンテンツを再生する抽出・再生手段と、再生される前記動画コンテンツに基づいて、前記タグ名称及び前記タグ名称の内容を示す単文又は単語群の入力を受け付けるタグ入力受付手段と、再生される前記動画コンテンツに基づいて、前記動画コンテンツの特定部分の再生開始時間及び再生終了時間を含むポイント情報を探索するポイント探索手段と、入力を受け付けた前記タグ名称ごとに、探索された前記ポイント情報の割り付けを受け付けるポイント割付受付手段と、入力を受け付けた前記タグ名称ごとに、検索時に発話可能な前記属性情報の入力を受け付ける属性情報受付手段とを備えることが好ましい。

20

【0013】

また、本発明に係る検索支援サーバは、前記抽出・再生手段は、前記動画コンテンツを精査して、含まれている音声を単文又は単語群として認識してテキストデータとして出力する音声文字起し手段と、出力されたテキストデータの選択を受け付けた場合、選択を受け付けたテキストデータに対応する前記動画コンテンツを再生するコンテンツ再生手段とを備えることが好ましい。

30

【0014】

また、本発明に係る検索支援サーバは、前記ポイント割付受付手段は、探索された前記ポイント情報に含まれる、前記動画コンテンツの特定部分の再生開始時間及び再生終了時間について、前記動画コンテンツを再生しながら更新を受け付けるポイント更新受付手段を備えることが好ましい。

【0015】

また、本発明に係る検索支援サーバは、前記ガイドデータは、前記サーチデータに含まれる前記タグ情報を、前記サーチデータを識別するサーチデータ識別情報と対応付けて生成されることが好ましい。

40

【0016】

また、本発明に係る検索支援サーバは、前記スピーチウィンドウに表示され、ユーザによる選択を受け付けることが可能な前記タグ名称を絞り込むために、ユーザにより発話された音声データの入力を受け付ける発話受付手段と、入力を受け付けた音声データをテキストデータに変換し、変換したテキストデータの前記ガイドデータの前記タグ名称及び前記タグ名称の内容を示す単文又は単語群に対する一致度を算出し、算出された一致度が最大であるタグ名称を特定するとともに、前記一致度が所定値より大きい場合に前記タグ名称に対応する前記動画コンテンツの特定部分を表示するメタタグ音声認識手段と、入力を受け付けた音声データに基づいて、前記ガイドデータの前記タグ情報のうち、前記属性情

50

報に一致するものを抽出し、一致する一又は複数の属性情報の論理積による前記ガイドデータの絞り込みを行い、絞り込まれた前記ガイドデータの前記タグ名称及び前記属性情報を出力する絞り込み音声認識手段とを備え、前記メタタグ音声認識手段及び前記絞り込み音声認識手段を並行して実行することが好ましい。

【0017】

また、本発明に係る検索支援サーバは、前記メタタグ音声認識手段は、変換されたテキストデータと、すべての前記タグ名称及び前記タグ名称の内容を示す単文又は単語群で形成された第一のデータセットとの一致度を算出し、前記絞り込み音声認識手段は、すべてのタグ名称に対応付けられている前記属性情報で形成された第二のデータセットを音声認識フィルタとして用いた認識結果として抽出された前記属性情報の論理積により前記ガイドデータを絞り込むことが好ましい。

10

【0018】

また、本発明に係る検索支援サーバは、前記絞り込み音声認識手段は、絞り込まれたガイドデータの前記属性情報により前記第二のデータセットを更新することが好ましい。

【0019】

次に、上記目的を達成するために本発明に係る検索支援方法は、Webサイト上に公開される動画コンテンツの検索を支援する検索支援サーバで実行することが可能な検索支援方法であって、前記検索支援サーバが、動画コンテンツの特定部分を示すタグ情報を選択するための発話を誘導するガイドデータを選択的に表示するサインージウインドウと、前記ガイドデータの中から、表示すべきガイドデータをユーザの発話により絞り込むことが可能なスピーチウインドウとをデータ通信することが可能に接続されている端末装置に表示させる工程と、前記サインージウインドウ及び/又は前記スピーチウインドウに表示された前記タグ情報のユーザによる発話を受け付けて、表示される前記タグ情報を絞り込み、絞り込まれた前記タグ情報の選択を受け付けることにより、選択を受け付けた前記タグ情報に対応する前記動画コンテンツの特定部分を前記端末装置において再生させる工程とを実行することを特徴とする。

20

【0020】

また、本発明に係る検索支援方法は、前記タグ情報は、少なくとも前記動画コンテンツの特定部分を識別するタグ名称、前記タグ名称の内容を説明する単文又は単語群からなるテキストデータ及び前記タグ名称の属性を示す属性情報で構成されていることが好ましい。

30

【0021】

また、本発明に係る検索支援方法は、前記検索支援サーバが、前記動画コンテンツの内容に基づいて、内容を示す単文又は単語群を抽出して、抽出された単文又は単語群の選択を受け付けた場合に対応する前記動画コンテンツを再生する工程と、再生される前記動画コンテンツに基づいて、前記タグ名称及び前記タグ名称の内容を示す単文又は単語群の入力を受け付ける工程と、再生される前記動画コンテンツに基づいて、前記動画コンテンツの特定部分の再生開始時間及び再生終了時間を含むポイント情報を探索する工程と、入力を受け付けた前記タグ名称ごとに、探索された前記ポイント情報の割り付けを受け付ける工程と、入力を受け付けた前記タグ名称ごとに、検索時に発話可能な前記属性情報の入力を受け付ける工程とを実行し、前記サインージウインドウ及び/又は前記スピーチウインドウに表示するガイドデータの基礎となるサーチデータを生成することが好ましい。

40

【0022】

また、本発明に係る検索支援方法は、前記検索支援サーバが、前記動画コンテンツを精査して、含まれている音声を単文又は単語群として認識してテキストデータとして出力する工程と、出力されたテキストデータの選択を受け付けた場合、選択を受け付けたテキストデータに対応する前記動画コンテンツを再生する工程とを実行することが好ましい。

【0023】

また、本発明に係る検索支援方法は、前記検索支援サーバが、探索された前記ポイント情報に含まれる、前記動画コンテンツの特定部分の再生開始時間及び再生終了時間について、前記動画コンテンツを再生しながら更新を受け付ける工程を実行することが好ましい。

50

【0024】

また、本発明に係る検索支援方法は、前記ガイドデータは、前記サーチデータに含まれる前記タグ情報を、前記サーチデータを識別するサーチデータ識別情報と対応付けて生成されることが好ましい。

【0025】

また、本発明に係る検索支援方法は、前記スピーチウィンドウに表示され、ユーザによる選択を受け付けることが可能な前記タグ名称を絞り込むために、前記検索支援サーバが、ユーザにより発話された音声データの入力を受け付ける工程を実行し、前記検索支援サーバが、入力を受け付けた音声データをテキストデータに変換し、変換したテキストデータの前記ガイドデータの前記タグ名称及び前記タグ名称の内容を示す単文又は単語群に対する一致度を算出し、算出された一致度が最大であるタグ名称を特定するとともに、前記一致度が所定値より大きい場合に前記タグ名称に対応する前記動画コンテンツの特定部分を表示する工程、及び入力を受け付けた音声データに基づいて、前記ガイドデータの前記タグ情報のうち、前記属性情報に一致するものを抽出し、一致する一又は複数の属性情報の論理積による前記ガイドデータの絞り込みを行い、絞り込まれた前記ガイドデータの前記タグ名称及び前記属性情報を出力する工程を、並行して実行することが好ましい。

10

【0026】

また、本発明に係る検索支援方法は、前記検索支援サーバが、変換されたテキストデータと、すべての前記タグ名称及び前記タグ名称の内容を示す単文又は単語群で形成された第一のデータセットとの一致度を算出する工程と、すべてのタグ名称に対応付けられている前記属性情報で形成された第二のデータセットを音声認識フィルタとして用いた認識結果として抽出された前記属性情報の論理積により前記ガイドデータを絞り込む工程とを実行することが好ましい。

20

【0027】

また、本発明に係る検索支援方法は、前記検索支援サーバが、絞り込まれたガイドデータの前記属性情報により前記第二のデータセットを更新する工程を実行することが好ましい。

【0028】

次に、上記目的を達成するために本発明に係るコンピュータプログラムは、Webサイト上に公開される動画コンテンツの検索を支援する検索支援サーバで実行することが可能なコンピュータプログラムであって、前記検索支援サーバを、動画コンテンツの特定部分を示すタグ情報を選択するための発話を誘導するガイドデータを選択的に表示するサイネージウィンドウと、前記ガイドデータの中から、表示すべきガイドデータをユーザの発話により絞り込むことが可能なスピーチウィンドウとをデータ通信することが可能に接続されている端末装置に表示させる手段、及び前記サイネージウィンドウ及び/又は前記スピーチウィンドウに表示された前記タグ情報のユーザによる発話を受け付けて、表示される前記タグ情報を絞り込み、絞り込まれた前記タグ情報の選択を受け付けることにより、選択を受け付けた前記タグ情報に対応する前記動画コンテンツの特定部分を前記端末装置において再生させる手段として機能させることを特徴とする。

30

【0029】

また、本発明に係るコンピュータプログラムは、前記タグ情報は、少なくとも前記動画コンテンツの特定部分を識別するタグ名称、前記タグ名称の内容を説明する単文又は単語群からなるテキストデータ及び前記タグ名称の属性を示す属性情報で構成されていることが好ましい。

40

【0030】

また、本発明に係るコンピュータプログラムは、前記検索支援サーバを、前記サイネージウィンドウ及び/又は前記スピーチウィンドウに表示するガイドデータの基礎となるサーチデータを生成するサーチデータ生成手段として機能させ、該サーチデータ生成手段を、前記動画コンテンツの内容に基づいて、内容を示す単文又は単語群を抽出して、抽出された単文又は単語群の選択を受け付けた場合に対応する前記動画コンテンツを再生する抽

50

出・再生手段、再生される前記動画コンテンツに基づいて、前記タグ名称及び前記タグ名称の内容を示す単文又は単語群の入力を受け付けるタグ入力受付手段、再生される前記動画コンテンツに基づいて、前記動画コンテンツの特定部分の再生開始時間及び再生終了時間を含むポイント情報を探索するポイント探索手段、入力を受け付けた前記タグ名称ごとに、探索された前記ポイント情報の割り付けを受け付けるポイント割付受付手段、及び入力を受け付けた前記タグ名称ごとに、検索時に発話可能な前記属性情報の入力を受け付ける属性情報受付手段として機能させることが好ましい。

【0031】

また、本発明に係るコンピュータプログラムは、前記抽出・再生手段を、前記動画コンテンツを精査して、含まれている音声を単文又は単語群として認識してテキストデータとして出力する音声文字起し手段、及び出力されたテキストデータの選択を受け付けた場合、選択を受け付けたテキストデータに対応する前記動画コンテンツを再生するコンテンツ再生手段として機能させることが好ましい。

10

【0032】

また、本発明に係るコンピュータプログラムは、前記ポイント割付受付手段を、探索された前記ポイント情報に含まれる、前記動画コンテンツの特定部分の再生開始時間及び再生終了時間について、前記動画コンテンツを再生しながら更新を受け付けるポイント更新受付手段として機能させることが好ましい。

【0033】

また、本発明に係るコンピュータプログラムは、前記ガイドデータは、前記サーチデータに含まれる前記タグ情報を、前記サーチデータを識別するサーチデータ識別情報と対応付けて生成されることが好ましい。

20

【0034】

また、本発明に係るコンピュータプログラムは、前記スピーチウィンドウに表示され、ユーザによる選択を受け付けることが可能な前記タグ名称を絞り込むために、前記検索支援サーバを、ユーザにより発話された音声データの入力を受け付ける発話受付手段として機能させ、前記検索支援サーバを、入力を受け付けた音声データをテキストデータに変換し、変換したテキストデータの前記ガイドデータの前記タグ名称及び前記タグ名称の内容を示す単文又は単語群に対する一致度を算出し、算出された一致度が最大であるタグ名称を特定するとともに、前記一致度が所定値より大きい場合に前記タグ名称に対応する前記動画コンテンツの特定部分を表示するメタタグ音声認識手段、及び入力を受け付けた音声データに基づいて、前記ガイドデータの前記タグ情報のうち、前記属性情報に一致するものを抽出し、一致する一又は複数の属性情報の論理積による前記ガイドデータの絞り込みを行い、絞り込まれた前記ガイドデータの前記タグ名称及び前記属性情報を出力する絞り込み音声認識手段として、並行して機能させることが好ましい。

30

【0035】

また、本発明に係るコンピュータプログラムは、前記メタタグ音声認識手段を、変換されたテキストデータと、すべての前記タグ名称及び前記タグ名称の内容を示す単文又は単語群で形成された第一のデータセットとの一致度を算出する手段として機能させ、前記絞り込み音声認識手段を、すべてのタグ名称に対応付けられている前記属性情報で形成された第二のデータセットを音声認識フィルタとして用いた認識結果として抽出された前記属性情報の論理積により前記ガイドデータを絞り込む手段として機能させることが好ましい。

40

【0036】

また、本発明に係るコンピュータプログラムは、前記絞り込み音声認識手段を、絞り込まれたガイドデータの前記属性情報により前記第二のデータセットを更新する手段として機能させることが好ましい。

【発明の効果】

【0037】

本発明によれば、Webサイト上に公開されている多数の動画コンテンツの中から、動画コンテンツを提供するコンテンツ提供者が、提供する動画コンテンツへユーザを誘導し

50

やすいガイドデータをスピーチウィンドウやサイネージウィンドウに表示することができるとともに、ユーザ自らの意思で選択対象となるガイドデータを絞り込むことができるので、個々のユーザが表示されているガイドデータ通りに発話あるいは選択することにより、簡便にしかも迅速にコンテンツ提供者が提供する動画コンテンツをユーザの意向に沿って表示・再生することが可能となる。

【図面の簡単な説明】

【0038】

【図1】本発明の実施の形態に係る音声検索システムの構成を模式的に示すブロック図である。

【図2】本発明の実施の形態に係る検索支援サーバの構成を模式的に示すブロック図である。

10

【図3】本発明の実施の形態に係る端末装置の構成を模式的に示すブロック図である。

【図4】本発明の実施の形態に係る検索支援サーバのサーチデータ生成処理の機能ブロック図である。

【図5】、本発明の実施の形態に係る検索支援サーバが、コンテンツ提供者が使用する端末装置に表示させる入力受付画面の例示図である。

【図6】本発明の実施の形態に係る検索支援サーバで生成されるサーチデータ及びガイドデータの例示図である。

【図7】本発明の実施の形態に係る検索支援サーバのサーチデータ及びガイドデータの例示図である。

20

【図8】本発明の実施の形態に係る検索支援サーバのサーチデータ及びガイドデータの他の例示図である。

【図9】本発明の実施の形態に係る検索支援サーバが、ユーザが使用する端末装置上で表示させる入力受付画面の例示図である。

【図10】本発明の実施の形態に係る検索支援サーバが、ユーザが使用する端末装置に表示させるスピーチウィンドウの例示図である。

【図11】本発明の実施の形態に係る検索支援サーバのCPUのコンテンツ提供者の設定処理手順を示すフローチャートである。

【図12】本発明の実施の形態に係る検索支援サーバのユーザによる発話の音声認識処理の機能ブロック図である。

30

【図13】本発明の実施の形態に係る検索支援サーバのCPUの音声認識処理の手順を示すフローチャートである。

【発明を実施するための形態】

【0039】

以下、本発明の実施の形態に係る検索支援サーバについて、図面を参照して説明する。以下の実施の形態は、特許請求の範囲に記載された発明を限定するものではなく、実施の形態の中で説明されている特徴的事項の組み合わせの全てが解決手段の必須事項であるとは限らないことは言うまでもない。

【0040】

また、本発明は多くの異なる態様にて実施することが可能であり、実施の形態の記載内容に限定して解釈されるべきものではない。実施の形態を通じて同じ要素には同一の符号を付している。

40

【0041】

以下の実施の形態では、コンピュータシステムにコンピュータプログラムを導入した音声検索システムとして説明するが、当業者であれば明らかな通り、本発明はその一部をコンピュータで実行することが可能なコンピュータプログラムとして実施することができる。したがって、本発明は、動画コンテンツの特定部分を示すタグ情報を選択するための発話を誘導するガイドデータを選択的に表示するサイネージウィンドウと、全てのガイドデータを発話で絞り込んで表示することが可能なスピーチウィンドウを用いることで、ユーザが自己の目的に応じて絞り込んだガイドデータ通りに発話又は選択する限り、コンテン

50

ツ提供者がユーザに提供したい動画コンテンツの特定部分へ確実に誘導することが可能な検索支援サーバというハードウェアとしての実施の形態、ソフトウェアとしての実施の形態、又はソフトウェアとハードウェアとの組み合わせの実施の形態をとることができる。コンピュータプログラムは、ハードディスク、DVD、CD、光記憶装置、磁気記憶装置等の任意のコンピュータで読み取ることが可能な記録媒体に記録することができる。

【0042】

本発明の実施の形態によれば、Webサイト上に公開されている多数の動画コンテンツの中から、動画コンテンツを提供するコンテンツ提供者が、提供する動画コンテンツへユーザを誘導しやすいガイドデータをスピーチウィンドウやサインウィンドウに表示することができるとともに、ユーザ自らの意思で選択対象となるガイドデータを絞り込むことができるので、個々のユーザが表示されているガイドデータ通りに発話あるいは選択することにより、簡便にしかも迅速にコンテンツ提供者が提供する動画コンテンツをユーザの意向に沿って表示・再生することが可能となる。

10

【0043】

図1は、本発明の実施の形態に係る音声検索システムの構成を模式的に示すブロック図である。本実施の形態に係る音声検索システムは、コンテンツ提供者が使用する端末装置1aと、コンテンツや動画コンテンツを検索するユーザが使用する端末装置1bと、端末装置1a及び1bとデータ通信することが可能にインターネット等のネットワーク網2を介して接続されている検索支援サーバ3とで構成されている。端末装置1a、1bは、マイクやスピーカを接続してある据え置き型のPCに限定されるものではなく、マイクやスピーカを内蔵しているスマートフォン、タブレット等の携帯端末であっても良い。

20

【0044】

図2は、本発明の実施の形態に係る検索支援サーバ3の構成を模式的に示すブロック図である。本実施の形態に係る検索支援サーバ3は、少なくともCPU(中央演算装置)31、メモリ32、記憶装置33、I/Oインタフェース34、ビデオインタフェース35、可搬型メモリドライブ36、通信インタフェース37及び上述したハードウェアを接続する内部バス38で構成されている。

【0045】

CPU31は、内部バス38を介して検索支援サーバ3の上述したようなハードウェア各部と接続されており、上述したハードウェア各部の動作を制御するとともに、記憶装置33に記憶されているコンピュータプログラム100に従って、種々のソフトウェア的機能を実行する。メモリ32は、SRAM、SDRAM等の揮発性メモリで構成され、コンピュータプログラム100の実行時にロードモジュールが展開され、コンピュータプログラム100の実行時に発生する一時的なデータ等を記憶する。

30

【0046】

記憶装置33は、内蔵される固定型記憶装置(ハードディスク)、ROM等で構成されている。記憶装置33に記憶されたコンピュータプログラム100は、プログラム及びデータ等の情報を記録したDVD、CD-ROM、USBメモリ、SDカード等の可搬型記録媒体90から、可搬型メモリドライブ36によりダウンロードされ、実行時には記憶装置33からメモリ32へ展開して実行される。もちろん、通信インタフェース37を介して接続されている外部コンピュータからダウンロードされたコンピュータプログラムであっても良い。

40

【0047】

記憶装置33は、サーチデータ記憶部331及びガイドデータ記憶部332とを備えている。サーチデータ記憶部331は、コンテンツ提供者が、提供するコンテンツに対してユーザがアクセス可能なサーチデータ(コンテンツの特定部分を識別するタグ名称、タグ名称の内容を説明する単文又は単語群からなるテキストデータ及びタグ名称の属性を示す属性情報)をポイント情報(動画コンテンツの場合には、加えてタイムスタンプ情報)に対応付けて記憶する。なお、ポイント情報とは、コンテンツにアクセスすることが可能なコンテンツの存在位置を示す情報を広く意味している。コンテンツがWebページである

50

場合にはURLがポインタ情報に相当し、動画コンテンツである場合には、再生可能なURLだけではなく、動画コンテンツの再生を開始するタイムスタンプ情報、再生を終了するタイムスタンプ情報もポインタ情報に含まれる。

【0048】

ガイドデータ記憶部332は、ユーザが検索するための発話あるいはクリック、タッチ等の選択操作を促すために、サーチデータに基づいて生成されるガイドデータを記憶する。ガイドデータは、サーチデータの中から、コンテンツの特定部分を識別するタグ名称、タグ名称の内容を説明する単文又は単語群からなるテキストデータ及びタグ名称の属性を示す属性情報を抽出して、対応するサーチデータの識別情報に対応付けて生成される。ガイドデータを表示するスピーチウィンドウに一覧表示された状態で、ユーザの発話により表示されるガイドデータが絞り込まれる。ユーザが絞り込まれて表示されているガイドデータの中からいずれかのタグ名称を選択することで、選択されたタグ名称に対応付けられたサーチデータで特定されるコンテンツ（動画コンテンツ含む）を表示（あるいは再生）することができる。

10

【0049】

通信インタフェース37は内部バス38に接続されており、インターネット、LAN、WAN等の外部のネットワーク網2に接続されることにより、外部コンピュータ等とデータ送受信を行うことが可能となっている。

【0050】

I/Oインタフェース34は、入力装置であるキーボード41、マウス42と接続され、データの入力を行う。本実施の形態では、実際に音声を入力するのは、コンテンツ提供者又はユーザが使用している端末装置1a、1b（スマートフォン、タブレット等）であり、入力された音声データを通信インタフェース37を介して受信する。もちろん、検索支援サーバ3にマイク、スピーカ等を直接接続していても良い。

20

【0051】

ビデオインタフェース35は、CRTディスプレイ、液晶ディスプレイ等の表示装置43と接続されている。本実施の形態では、実際に画像を出力表示するのは、コンテンツ提供者又はユーザが使用している端末装置1a、1b（スマートフォン、タブレット等）であり、検索支援サーバ3は、端末装置1a、1bへ（音声データを含む）画像データ等を通信インタフェース37を介して送信する。

30

【0052】

図3は、本発明の実施の形態に係る端末装置1（1a、1b共通）の構成を模式的に示すブロック図である。本実施の形態に係る端末装置1は、少なくともCPU（中央演算装置）11、メモリ12、記憶装置13、I/Oインタフェース14、ビデオインタフェース15、可搬型メモリドライブ16、通信インタフェース17及び上述したハードウェアを接続する内部バス18で構成されている。

【0053】

CPU11は、内部バス18を介して端末装置1の上述したようなハードウェア各部と接続されており、上述したハードウェア各部の動作を制御するとともに、記憶装置13に記憶されているコンピュータプログラム101に従って、種々のソフトウェア的機能を実行する。メモリ12は、SRAM、SDRAM等の揮発性メモリで構成され、コンピュータプログラム101の実行時にロードモジュールが展開され、コンピュータプログラム101の実行時に発生する一時的なデータ等を記憶する。

40

【0054】

記憶装置13は、内蔵される固定型記憶装置（ハードディスク）、ROM等で構成されている。記憶装置13に記憶されたコンピュータプログラム101は、通信インタフェース17を介して接続されている外部コンピュータからダウンロードされ、実行時には記憶装置13からメモリ12へ展開して実行される。もちろん、プログラム及びデータ等の情報を記録したSDカード等の可搬型記録媒体91から可搬型メモリドライブ16を介してダウンロードされたコンピュータプログラムであっても良い。

50

【 0 0 5 5 】

通信インタフェース 1 7 は内部バス 1 8 に接続されており、インターネット、LAN、WAN等の外部のネットワーク網 2 に接続されることにより、外部コンピュータ等とデータ送受信を行うことが可能となっている。

【 0 0 5 6 】

I/Oインタフェース 1 4 は、キーボード 2 0 3、マウス 2 0 4 等の入力装置の他、マイク 2 0 1 等の音声入力装置、スピーカ 2 0 2 等の音声出力装置と接続され、データの入出力を行う。通信インタフェース 1 7 を介してスマートフォン等を接続し、音声入力装置及び音声出力装置を代用しても良い。

【 0 0 5 7 】

ビデオインタフェース 1 5 は、表示装置 2 0 5 と接続されており、検索支援サーバ 3 から送信されてくる入出力用の画像をブラウザ等で表示する。検索されたコンテンツや動画コンテンツは、表示装置 2 0 5 に表示しても良いし、別途ネットワーク網を介してデータ通信することが可能に接続されている外部のコンピュータで表示しても良い。

【 0 0 5 8 】

以下、上述した構成の検索支援サーバ 3 の動作について説明する。

【 0 0 5 9 】

図 4 は、本発明の実施の形態に係る検索支援サーバ 3 のサーチデータ生成処理の機能ブロック図である。図 4 では、コンテンツ提供者がサーチデータを生成し、ガイドデータを生成して、サイネージウィンドウ及び/又はスピーチウィンドウに表示する手順について説明する。

【 0 0 6 0 】

図 4 において、サーチデータ生成部 4 0 1 は、サイネージウィンドウ及び/又はスピーチウィンドウに表示するガイドデータの基礎となるサーチデータを生成する。サーチデータ生成部 4 0 1 は、抽出・再生部 4 0 2、タグ入力受付部 4 0 3、ポイント探索部 4 0 4、ポイント割付受付部 4 0 5、及び属性情報受付部 4 0 6 を備えている。

【 0 0 6 1 】

抽出・再生部 4 0 2 は、動画コンテンツの内容に基づいて、内容を示す単文又は単語群を抽出して、抽出された単文又は単語群の選択を受け付けた場合に対応する動画コンテンツを再生する。より具体的には、抽出・再生部 4 0 2 は、音声文字起し部 4 0 2 1 と、動画コンテンツ表示部 4 0 2 2 とを備えている。

【 0 0 6 2 】

音声文字起し部 4 0 2 1 は、対象となる動画コンテンツの音声部分を抽出して書き起こし、適切な文節ごとのタイムスタンプ情報と対応付けたテキストデータとして出力する。出力されたテキストデータは、動画コンテンツに含まれている音声に基づいてタグ名称を生成する参考にすることができる。

【 0 0 6 3 】

動画コンテンツ再生部 4 0 2 2 は、出力されたテキストデータの選択を受け付けた場合、選択を受け付けたテキストデータに対応する動画コンテンツの特定部分を再生する。タグ名称の候補の選択を受け付けた場合に、正しい動画コンテンツであるか否か、そしてタイムスタンプ情報を更新するために再生される。

【 0 0 6 4 】

なお、選択を受け付けたテキストデータ、ガイドデータ及び外部から取得したテキストデータを教師データとして学習し、音声文字起し部 4 0 2 1 で使用する音声認識用の言語モデルを出力しても良い。学習方法としては、いわゆる AI を含み、周知の学習方法であれば特に限定されるものではない。

【 0 0 6 5 】

以下、動画コンテンツの特定部分に対応付けられるタグ情報の入力を受け付ける。本実施の形態では、タグ情報とは、少なくとも動画コンテンツの特定部分を識別するタグ名称、タグ名称の内容を説明する単文又は単語群からなるテキストデータ及びタグ名称の属性

10

20

30

40

50

を示す属性情報で構成されている。

【0066】

タグ入力受付部403は、表示出力されたテキストデータ、あるいは再生している動画コンテンツの特定部分に基づいて、タグ名称及びタグ名称の内容を示す単文又は単語群の入力を受け付ける。

【0067】

ポインタ探索部404は、再生されている動画コンテンツの特定部分に基づいて、動画コンテンツの特定部分を示すポインタ情報を探索する。探索されるポインタ情報は、既に対応付けられているURLだけではなく、再生開始タイミング及び再生終了タイミングに関するタイムスタンプ情報も探索される。

【0068】

動画コンテンツの場合、自動的に割り付けられた特定部分が、コンテンツ提供者が本当に提供したい部分であるとは限らない。不要な部分が含まれたり、必要な部分なのに分断されたりしている可能性がある。そこで、後述する画面操作によって、ポインタ情報の更新を行うことができる。これにより、過不足なく所望の動画コンテンツを提供することができる。

【0069】

ポインタ割付受付部405は、入力を受け付けたタグ名称ごとに、探索されたポインタ情報の割り付けを受け付ける。これにより、タグ名称の選択を受け付けることにより、割り付けられたポインタ情報に従って、動画コンテンツの特定部分を再生することができる。

【0070】

属性情報受付部406は、入力を受け付けたタグ名称ごとに、検索時に発話可能な属性情報の入力を受け付ける。後述するように属性情報をユーザが発話することにより、タグ名称を絞り込むことができる。

【0071】

ここで、属性情報としては、例えば「新着」、「動画」、「使い方」、「料金」、「医療」など、動画コンテンツを絞り込むのに役立つ情報であり、しかも発話しやすい言葉であれば特に限定されるものではない。

【0072】

このようにタグ名称及びタグ名称の内容を示す単文又は単語群、属性情報の入力を受け付け、タグ情報としてポインタ情報に対応付けてサーチデータを生成する。ガイドデータ生成部407は、生成されたサーチデータに含まれるタグ情報を、サーチデータを識別するサーチデータ識別情報と対応付けて、ガイドデータを生成する。

【0073】

図5は、本発明の実施の形態に係る検索支援サーバ3が、コンテンツ提供者が使用する端末装置1aに表示させる入力受付画面の例示図である。図5に示す初期画面50が、コンテンツ提供者の使用する端末装置1aにポップアップ表示される。コンテンツ提供者は、ユーザに提供したい動画コンテンツを再生領域51において再生しながら、「割付」ボタン52を選択することによりタグ情報に動画コンテンツのシーンを割り付けることができる。

【0074】

コンテンツ提供者による「割付」ボタン52の選択を受け付けた場合、選択を受け付けた動画コンテンツの音声部分を音声認識してすべて書き起こし、単文又は単語群として抽出して、文節ごとに再生開始点を示すタイムスタンプ情報を割り付けたトランスクリプションを作成する。作成されたトランスクリプションは、目次表示領域53に一覧表示、あるいは表示しきれない場合には、図5ではサーチデータ作成領域として使用されている共用ウィンドウ60に表示される。表示しきれない場合には、スクロール操作等ですべてのトランスクリプションを表示することができる。

【0075】

これにより、共用ウィンドウ60にトランスクリプションとして表示されている動画コ

10

20

30

40

50

コンテンツ中のシーンごとのタグ名称の候補の選択を受け付けた場合、対応付けられた動画コンテンツのシーンが再生される。

【0076】

コンテンツ提供者は、動画コンテンツを再生しながら、所望の再生開始タイミングにおいて「マーク入力」ボタン61を選択する。「マーク入力」ボタン61の選択を受け付けた場合、仮のタグ情報が再生開始タイムスタンプをポインタ情報として作成される。仮のタグ情報は、動画コンテンツの再生時に随時生成して記憶することができる。

【0077】

次に、コンテンツ提供者による「タグ入力」ボタン54の選択を受け付ける。「タグ入力」ボタン54の選択を受け付けた場合、一覧表示されている仮のタグ情報の選択を受け付けることで、共用ウインドウ60にタグ名称入力領域55、メタタグ入力領域56等を含むサーチデータ作成領域60が表示され、タグ情報の入力を受け付ける。

10

【0078】

仮のタグ情報の一覧表示は、「タグ入力終了」ボタン59の選択を受け付けることにより行われる。仮のタグ情報の選択を受け付けてタグ情報の入力を行うことで、すべての仮のタグ情報に対してサーチデータを生成することができる。なお、属性情報入力領域は図5には図示されていないが、キーインできる入力領域あるいは選択ボタンを配して表示させて、入力を受け付ければ良い。

【0079】

本実施の形態で特徴的なのは、コンテンツ提供者による発話でタグ情報の入力を受け付ける点と、「時間メータ」の存在である。例えばコンテンツ提供者が「時間メータ」と発話した場合、時間メータ57によりタグ情報と対応付けられている動画コンテンツの特定部分を更新することが可能なポインタ更新受付手段を備えている。

20

【0080】

具体的には、共用ウインドウ60に表示されたサーチデータ作成領域に更新するタグ情報の作成画面を表示し、「時間メータ」との発話を受け付けた場合、時間メータのスライダを該タグ情報に対応付けられているポインタ情報の再生開始タイムスタンプに位置付けて表示するとともに移動可能となる。再生開始タイムスタンプが再生時間表示領域58に表示された状態でスライダを移動させると、スライダの位置に対応した再生時点のタイムスタンプが再生時間表示領域58に表示されつつ、再生領域51の動画コンテンツが変化する。

30

【0081】

再生開始タイムスタンプ及び再生終了タイムスタンプ(以下、タイムスタンプ情報)も、コンテンツ提供者による「マーク入力」ボタン61の選択、あるいは「マーク」の発話で設定しても良い。例えば動画コンテンツを再生しながら、タイムスタンプ情報を設定したいシーンになった時点でコンテンツ提供者が「マーク入力」ボタン61を選択、あるいは「マーク」と発話することにより、タイムスタンプ情報を設定することができる。

【0082】

タグ情報、再生開始タイムスタンプ及び再生終了タイムスタンプの設定が終了した場合、コンテンツ提供者は「タグ入力」ボタン54を選択する。「タグ入力」ボタン54の選択を受け付けた時点で、検索支援サーバ3は端末装置1aのブラウザに表示されている動画コンテンツのURLをポインタ情報として探索し、設定されたタイムスタンプ情報とともに、入力されたタグ名称に対応付けてサーチデータとしてサーチデータ記憶部331に記憶する。

40

【0083】

ガイドデータは、サーチデータのタグ情報部分に、サーチデータを識別する識別情報を対応付けて生成する。図6は、本発明の実施の形態に係る検索支援サーバ3で生成されるサーチデータ及びガイドデータの例示図である。図6(a)は、生成されたガイドデータの例示図であり、図6(b)は、生成されたガイドデータの基礎となるサーチデータの例示図である。

50

【 0 0 8 4 】

図 6 (a) に示すように、ガイドデータは、サーチデータの何番目のタグ名称であるかを示す「番号」で対応付けられており、Webサイトを識別する識別情報である共通のサーチデータID (図 6 ではサーチデータID = X X X X X X) を有している。生成されたガイドデータはガイドデータ記憶部 3 3 2 に記憶される。

【 0 0 8 5 】

そして、「ポインタ情報」としてURL及びタイムスタンプ情報としての再生「開始」時刻及び再生「終了」時刻がタグ名称と対応付けて記憶されているのはサーチデータのみである。したがって、図 5 の「ガイドデータ」ボタン 6 2 の選択を受け付けることで、サーチデータのタグ情報をコピーし、サーチデータID (X X X X X X) と、サーチデータの何番目のタグ名称であるかを示す「番号」とを対応付けてガイドデータが生成される。図 6 (a) に示すガイドデータの「サーチデータID」欄が空白なのは、一のサーチデータのみに基づいて生成されたガイドデータであることを意味しており、他のサーチデータに基づいて生成されたガイドデータである場合には、他のサーチデータのサーチデータIDが記載される。

10

【 0 0 8 6 】

図 5 に示すサイネージウィンドウの共用ウィンドウ 6 0 には、生成されたガイドデータの中からコンテンツ提供者がユーザに表示したいデータとして選択したガイドデータのタグ名称が表示されることが好ましい。この場合、「ガイドデータ」ボタン 6 2 の選択を受け付けると「編集」ボタンと「拡張」ボタン (図示せず) が表示される。「編集」ボタンの選択を受け付けた場合、共用ウィンドウ 6 0 内にキーイン等することにより、サイネージウィンドウの共用ウィンドウ 6 0 に表示するタグ名称を設定することができる。「拡張」ボタンの選択を受け付けた場合には、他のサーチデータから生成されたガイドデータを追加することができる。追加されたガイドデータから目次データを取り出すことで、目次表示領域 5 3 に表示することができる。

20

【 0 0 8 7 】

すなわち、本実施の形態では、コンテンツ提供者ごとにサーチデータを生成しても良いし、同一コンテンツ提供者が複数立ち上げたWebサイトごとにサーチデータを生成しても良い。例えば、同一の会社のWebサイトであっても、事業部ごと、商品・サービスごとにWebサイトを立ち上げる機会が急増しており、コンテンツが日々増加する。これらのコンテンツを迅速にかつ確実に検索できるようにすることは、顧客サービスにおいても重要になる。

30

【 0 0 8 8 】

本実施の形態では、全てのガイドデータは、基礎となるサーチデータに基づいて生成されている。したがって、一のガイドデータに他のサーチデータに基づいて生成されたガイドデータさえ追加すれば、他のサーチデータに対応付けられているコンテンツ (動画コンテンツ含む) 、すなわち他のWebサイトのコンテンツであっても迅速かつ確実に表示・再生することができるようになる。

【 0 0 8 9 】

図 7 は、本発明の実施の形態に係る検索支援サーバ 3 のサーチデータ及びガイドデータの例示図である。図 7 (a) は、一のコンテンツ提供者のサーチデータに基づくガイドデータの例示図であり、図 7 (b) は、動画コンテンツを提供する他のコンテンツ提供者のサーチデータの例示図である。

40

【 0 0 9 0 】

図 6 に示すサーチデータ及びガイドデータとの違いは、共通のサーチデータIDではなく、異なるサーチデータIDを有するガイドデータが含まれている点にある。すなわち、サーチデータ及びガイドデータを生成した時点では、サーチデータIDは共通である。本実施の形態では、サーチデータとガイドデータとの二段構造にすることにより、サーチデータIDが異なるガイドデータを設定することができる。これにより、サーチデータIDが異なるガイドデータの基礎となるサーチデータを読み出すことができ、サーチデータI

50

Dが異なるサーチデータ、すなわち異なるコンテンツ提供者が提供する動画コンテンツの特定部分を再生することが可能となる。

【0091】

例えば共通のサーチデータIDを「YYYYYY」とする。そして、共通のサーチデータIDが「YYYYYY」を基礎として生成されたガイドデータには、他のサーチデータIDを基礎としたガイドデータを随時追加することができる。図7(a)の例では、サーチデータIDが「TTTTTT」であるガイドデータが追加されている。つまり、サーチデータIDが「TTTTTT」であるガイドデータの基礎となるサーチデータを読み出すことができるようになり、当該サーチデータに対応付けられている動画コンテンツを再生することができるようになる。

10

【0092】

つまり、図7(b)に示すサーチデータIDが「TTTTTT」であるサーチデータを参照することができ、所望のコンテンツのポインタ情報を取得することができる。したがって、ガイドデータを追加するだけで、どのWebサイトのコンテンツであっても所望のコンテンツを表示することができる。なお、図7(b)において、ガイドデータの「サーチデータID」欄の空欄は、追加されたガイドデータではなく、サーチデータIDが共通のサーチデータID「YYYYYY」であるガイドデータであることを意味している。

【0093】

図8は、本発明の実施の形態に係る検索支援サーバ3のサーチデータ及びガイドデータの他の例示図である。図8(a)は、一のコンテンツ提供者のサーチデータに基づくガイドデータの例示図であり、図8(b)は、他のコンテンツ提供者のサーチデータの例示図である。

20

【0094】

例えば共通のサーチデータIDを「YYYYYY」とする。そして、サーチデータIDが「YYYYYY」であるサーチデータを基礎として生成されたガイドデータには、他のサーチデータIDを基礎としたコンテンツのガイドデータを随時追加することができる。図8(a)では、サーチデータIDが「PPPPPP」であるガイドデータが追加されている。つまり、サーチデータIDが「PPPPPP」であるガイドデータの基礎となるサーチデータを読み出すことができるようになり、当該サーチデータに対応付けられているコンテンツを表示することができるようになる。

30

【0095】

つまり、図8(b)に示すサーチデータIDが「PPPPPP」であるサーチデータを参照することができ、所望のコンテンツのポインタ情報及びタイムスタンプ情報を取得することができる。したがって、ガイドデータを追加するだけで、どのWebサイトの動画コンテンツであっても所望の動画コンテンツの所望のシーンだけ再生することができる。なお、図8(b)において、ガイドデータの「サーチデータID」欄の空欄は、追加されたガイドデータではなく、サーチデータIDが「YYYYYY」であるガイドデータであることを意味している。

【0096】

図4に戻って、サイネージウィンドウ表示部408は、生成されたガイドデータのタグ名称等を、コンテンツ提供者が選択して、ユーザが使用する端末装置1b上でサイネージウィンドウに表示させる。これにより、ユーザに対して、動画コンテンツの特定部分を示すタグ情報を選択するための発話を誘導することができる。

40

【0097】

図9は、本発明の実施の形態に係る検索支援サーバ3が、ユーザが使用する端末装置1b上で表示する入力受付画面900の例示図である。図9に示すように、動画コンテンツを表示可能なコンテンツウィンドウ901と一体となって、サイネージウィンドウ902が表示される。サイネージウィンドウ902には、文字列が流れて表示されるような形態で表示されることが好ましい。

【0098】

50

具体的には、サイネージウィンドウ 902 に、選択されたガイドデータのタグ名称が表示される。図 9 において、矢印は、文字列がその方向に流れるように移動しながら表示することを意味する。サイネージウィンドウ 902 の表示を見たユーザは、コンテンツ提供者がどのような情報の提供を意図しているかを知ることができ、発話内容を工夫することができる。コンテンツ提供者は、ユーザに提供したい動画コンテンツの中のシーンを示すタグ名称あるいはユーザによる選択を誘導する文字列をサイネージウィンドウ 902 に表示することで、ユーザを提供したい動画コンテンツの中のシーンへと誘導することができる。

【0099】

また、入力受付画面 900 には、動画コンテンツ中のシーンごとのタグ名称が目次として一覧表示される目次表示領域 903 を含む。直接目次を選択して、動画コンテンツの中のシーンをコンテンツウィンドウ 901 に再生させることができる。目次表示領域 903 に表示されているタグ名称の横の「i」ボタンを選択してタグ名称の説明を表示させることもできる。

10

【0100】

図 4 に戻って、スピーチウィンドウ表示部 409 は、生成されたガイドデータのタグ名称をすべて、ユーザが使用する端末装置 1b 上でスピーチウィンドウに表示させる。目次表示領域 903 に表示しきれないガイドデータのタグ名称も、スピーチウィンドウには表示することができる。ユーザが発話することにより、ガイドデータのタグ名称を絞り込むことができる。絞り込むことで、スピーチウィンドウにスクロール等の操作をすることなくタグ名称を表示することができ、クリック、タッチ等の選択操作でタグ名称を選択することにより、対応付けられている動画コンテンツの特定部分を表示することができる。

20

【0101】

図 10 は、本発明の実施の形態に係る検索支援サーバ 3 が、ユーザが使用する端末装置 1b に表示させるスピーチウィンドウの例示図である。図 10 に示すように、生成されたガイドデータのタグ名称は、スピーチウィンドウ 1001 上で選択可能な状態でガイドデータ表示領域 1002 にすべて表示される。

【0102】

ガイドデータ表示領域 1002 に表示しきれない場合も、スクロールボタン 1003 を操作することですべてのガイドデータのタグ名称を閲覧することができる。ガイドデータ表示領域 1002 に表示されているガイドデータのタグ名称の中から、一のタグ名称の選択を受け付ける。これにより、選択を受け付けたタグ名称のガイドデータの基礎となるサーチデータを特定することができるので、サーチデータに対応付けられている動画コンテンツの特定部分を再生することができる。

30

【0103】

図 10 の例では、スピーチウィンドウ 1001 に表示されているタグ名称の選択を受け付けた場合、対応するコンテンツの特定部分を確認できるよう、端末装置 1b に表示する。ユーザが使用する端末装置 1b は、図 10 に示すようにデスクトップ型でも良いし、スマホ、タブレット等の携帯端末であっても良い。

【0104】

図 11 は、本発明の実施の形態に係る検索支援サーバ 3 の CPU 31 のコンテンツ提供者の設定処理手順を示すフローチャートである。図 11 において、検索支援サーバ 3 の CPU 31 は、動画コンテンツの内容に基づいて、内容を示す単文又は単語群を抽出して（ステップ S1101）、抽出された単文又は単語群の選択を受け付けた場合に対応する動画コンテンツを再生する（ステップ S1102）。

40

【0105】

CPU 31 は、表示出力されたテキストデータ、あるいは再生している動画コンテンツの特定部分に基づいて、タグ名称及びタグ名称の内容を示す単文又は単語群の入力を受け付ける（ステップ S1103）。

【0106】

50

C P U 3 1 は再生されている動画コンテンツの特定部分に基づいて、動画コンテンツの特定部分を示すポインタ情報を探索する（ステップ S 1 1 0 4）。C P U 3 1 は、入力を受け付けたタグ名称ごとに、探索されたポインタ情報の割り付けを受け付ける（ステップ S 1 1 0 5）。これにより、タグ名称の選択を受け付けることにより、割り付けられたポインタ情報に従って、動画コンテンツの特定部分を再生することができる。

【 0 1 0 7 】

C P U 3 1 は、入力を受け付けたタグ名称ごとに、検索時に発話可能な属性情報の入力を受け付ける（ステップ S 1 1 0 6）。後述するように、属性情報をユーザが発話することにより、タグ名称を絞り込むことができる。

【 0 1 0 8 】

C P U 3 1 は、入力を受け付けたタグ名称及びタグ名称の内容を示す単文又は単語群、属性情報を、ポインタ情報に対応付けてサーチデータを生成する（ステップ S 1 1 0 7）。C P U 3 1 は、生成されたサーチデータに含まれるタグ情報を、サーチデータを識別するサーチデータ識別情報と対応付けて、ガイドデータを生成する（ステップ S 1 1 0 8）。

【 0 1 0 9 】

C P U 3 1 は、生成されたガイドデータのタグ名称等を、コンテンツ提供者が選択して、ユーザが使用する端末装置 1 b 上でサイネージウィンドウに表示させ（ステップ S 1 1 0 9）、生成されたガイドデータのタグ名称をすべて、ユーザが使用する端末装置 1 b 上でスピーチウィンドウに表示させる（ステップ S 1 1 1 0）。

【 0 1 1 0 】

以下、ユーザが検索処理を実行する手順について説明する。本実施の形態では、コンテンツ提供者がユーザの検索を誘導するサイネージウィンドウ及びスピーチウィンドウを表示させている点に特徴を有している。ただし、それだけではなく、ユーザにとって所望のコンテンツを迅速にかつ確実に検索表示させる工夫として、表示されるガイドデータのタグ名称をユーザが発話することにより絞り込むことができる点が従来の検索システムとの大きな相違点である。

【 0 1 1 1 】

図 1 2 は、本発明の実施の形態に係る検索支援サーバ 3 のユーザによる発話の音声認識処理の機能ブロック図である。図 1 2 では、ユーザの使用する端末装置 1 b に、図 9 に示すサイネージウィンドウ及び図 1 0 に示すスピーチウィンドウが表示されている状態でユーザによる発話を入力として受け付け、正しく音声認識する手順について説明する。

【 0 1 1 2 】

図 1 2 に示すように、発話受付部 1 2 0 1 は、ユーザにより発話された音声データの入力を受け付ける。具体的には、端末装置 1 b においてユーザが発話した音声データを受信することで、音声データを取得する。

【 0 1 1 3 】

前処理部 1 2 0 2 は、入力を受け付けた音声データに対して雑音除去、発話区間の検出等を実行する。前処理された音声データは、メタタグ音声認識部 1 2 0 3 及び絞り込み音声認識部 1 2 0 8 へ渡され、メタタグ音声認識部 1 2 0 3 及び絞り込み音声認識部 1 2 0 8 を並行して実行する。

【 0 1 1 4 】

メタタグ音声認識部 1 2 0 3 は、前処理された音声データを認識してテキストデータに変換し、変換されたテキストデータに基づいて、ガイドデータのタグ名称及びタグ名称の内容を示す単文又は単語群に対する一致度を算出し、算出された一致度が最大であるタグ名称を特定するとともに、一致度が所定値より大きい場合にタグ名称に対応するコンテンツの特定部分を表示する。つまり、メタタグ音声認識部 1 2 0 3 で正しく認識できた場合には、他の音声認識処理と統合する処理等余分な処理を実行することなく認識結果を即座に出力することができる。

【 0 1 1 5 】

すなわち、メタタグ音声認識部 1 2 0 3 は、テキストデータ変換部 1 2 0 4、一致度算

10

20

30

40

50

出部 1 2 0 5、タグ名称特定部 1 2 0 6、コンテンツ表示・再生部 1 2 0 7を備えている。テキストデータ変換部 1 2 0 4は、前処理された音声データを、いわゆるディクテーションングラマーに基づいて認識してテキストデータに変換する。

【 0 1 1 6 】

一致度算出部 1 2 0 5は、すべてのタグ名称及びタグ名称の内容を示す単文又は単語群で形成された第一のデータセットを照会して、入力を受け付けた音声データとの一致度を算出する。

【 0 1 1 7 】

タグ名称特定部 1 2 0 6は、算出された一致度が最大であるタグ名称を特定する。一致度が最大であるタグ名称が、最も確からしい認識結果だからである。しかし、一致度が所定の閾値以下である場合には誤認識の可能性も高い。

10

【 0 1 1 8 】

コンテンツ表示・再生部 1 2 0 7は、算出された一致度が所定の閾値より大きいかなかを判断する。所定の閾値以下であると判断した場合には、認識精度が不十分であるとして再度の発話の待ち状態とする。所定の閾値より大きいと判断した場合には、特定されたタグ名称に対応付けられたコンテンツの特定部分を表示あるいは再生して、認識結果が正しいかなかを判断できる。

【 0 1 1 9 】

絞り込み音声認識部 1 2 0 8は、前処理された音声データに基づいて、いわゆるルールングラマーを用いて、ガイドデータのタグ情報のうち、属性情報に一致するものを抽出し、一致する一又は複数の属性情報の論理積によるガイドデータの絞り込みを行い、絞り込まれたガイドデータのタグ名称及び属性情報を入力する。これにより、メタタグ音声認識部 1 1 0 3で妥当な認識結果を得られない場合であっても、次にユーザが発話する内容をさらに絞り込むことができるので、音声認識の成功率を高めることができ、結果として短時間で正しい認識結果を得ることができる。具体的には、絞り込み音声認識部 1 2 0 8は、属性抽出部 1 2 0 9、絞り込み部 1 2 1 0を備えている。

20

【 0 1 2 0 】

属性抽出部 1 2 0 9は、すべてのタグ情報に対応付けられている属性情報で形成された第二のデータセットを音声認識フィルタとして用いた認識結果として属性情報を抽出する。絞り込み部 1 2 1 0は、抽出された属性情報の論理積としてタグ名称（ガイドデータ）を絞り込む。

30

【 0 1 2 1 】

絞り込み音声認識部 1 2 0 8は、絞り込まれたガイドデータの属性情報により、音声認識フィルタである第二のデータセットを更新する更新部 1 2 1 0を備えることが好ましい。第二のデータセットは、属性情報によりガイドデータが絞り込まれる都度更新されるので、ユーザによる次の発話をさらに制限することになり、正しい認識結果を得る確信度を高めることができ、より迅速に正しい認識結果を得ることが可能となる。

【 0 1 2 2 】

図 1 3は、本発明の実施の形態に係る検索支援サーバ 3のCPU 3 1の音声認識処理の手順を示すフローチャートである。検索支援サーバ 3のCPU 3 1は、ユーザにより発話された音声データの入力を受け付ける（ステップ S 1 3 0 1）。具体的には、端末装置 1 bにおいてユーザが発話した音声データを受信することで、音声データを取得する。

40

【 0 1 2 3 】

CPU 3 1は、入力を受け付けた音声データに対して雑音除去、発話区間の検出等を実行する（ステップ S 1 3 0 2）。前処理された音声データに基づいて、以下の2つの処理が並行して実行される。

【 0 1 2 4 】

まず、CPU 3 1は、前処理された音声データに基づいて、前処理された音声データを、いわゆるディクテーションングラマーに基づいて認識してテキストデータに変換する（ステップ S 1 3 0 3）。CPU 3 1は、すべてのタグ名称及びタグ名称の内容を示す単文又

50

は単語群で形成された第一のデータセットを照会して、入力を受け付けた音声データとの一致度を算出する（ステップS1304）。CPU31は、算出された一致度が最大であるタグ名称を特定する（ステップS1305）。

【0125】

CPU31は、算出された一致度が所定の閾値より大きいか否かを判断する（ステップS1306）。CPU31が、所定の閾値以下であると判断した場合（ステップS1306：NO）、CPU31は、処理をステップS1301へ戻して、再度の発話の待ち状態となる。

【0126】

CPU31が、所定の閾値より大きいと判断した場合（ステップS1306：YES）、CPU31は、特定されたタグ名称に対応付けられたコンテンツの特定部分を表示あるいは再生する（ステップS1307）。

10

【0127】

一方、CPU31は、前処理された音声データに基づいて、全てのタグ名称に対応付けられている属性情報で形成された第二のデータセットを音声認識フィルタとして用い、認識結果の属性情報を抽出する（ステップS1308）。CPU31は、抽出された属性情報の論理積としてタグ名称（ガイドデータ）を絞り込む（ステップS1309）。CPU31は、絞り込まれたガイドデータのタグ名称及び属性情報により、音声認識フィルタである第二のデータセットを更新し（ステップS1310）、処理をステップS1301へ戻して、再度の発話の待ち状態となる。

20

【0128】

以上のように本実施の形態によれば、Webサイト上に公開されている多数の動画コンテンツの中から、動画コンテンツを提供するコンテンツ提供者が、自分の動画コンテンツへユーザを誘導しやすいガイドデータをスピーチウィンドウやサイネージウィンドウに表示することができるのと同時に、ユーザ自らの意思で選択対象となるガイドデータを絞り込むことができるので、個々のユーザが表示されているガイドデータ通りに発話あるいは選択することにより、迅速にかつ確実にコンテンツ提供者が提供する動画コンテンツの特定部分をユーザに対して表示することが可能となる。

【0129】

また、本実施の形態によれば、メタタグ音声認識処理と絞り込み音声認識処理とを並行して実行することにより、メタタグ音声認識が正しい場合にはそのまま認識結果を出力することができ、一方で、認識結果の確からしさが足りない場合には、より絞り込んだガイドデータに基づくタグ名称をスピーチウィンドウに表示することができる。したがって、ユーザは、比較的短時間で所望のコンテンツを表示又は再生することが可能となる。

30

【0130】

なお、本発明は上記実施例に限定されるものではなく、本発明の趣旨の範囲内であれば多種の変更、改良等が可能である。例えば属性情報を含むタグ情報の入力方法は、上述した方法に限定されるものではなく、タグ名称を絞り込むことができる文字列を入力可能であれば特に限定されるものではない。

【0131】

また、ユーザによる発話を認識する場合に、メタタグ音声認識部1203と並行に処理を実行する認識処理は、絞り込み音声認識部1208に限定されるものではなく、ユーザによる次の発話を限定することができる処理であれば特に限定されるものではない。

40

【0132】

また、上記実施例では、スピーチウィンドウ81はユーザが使用する端末装置1b上に表示されているが、特にこれに限定されるものではなく、例えばデータ通信することが可能に接続されているスマートホン、タブレット等の外部の携帯端末装置上でリモート操作する形態であっても良い。

【0133】

また、音声認識に用いるディクテーショングラマー及びルールグラマーは、テキストデ

50

ータ、ガイドデータ及び外部から取得したテキストデータを教師データとして学習することにより生成することが望ましい。学習方法としては、いわゆる機械学習、深層学習のようにAIを用いても良いし、対応テーブルを拡充するような従来の方法であっても良く、特に限定されるものではない。

【符号の説明】

【0134】

1、1 a、1 b 端末装置

2 ネットワーク網

3 検索支援サーバ208

11、31 CPU

10

12、32 メモリ

13、33 記憶装置

14、34 I/Oインタフェース

15、35 ビデオインタフェース

16、36 可搬型ディスクドライブ

17、37 通信インタフェース

18、38 内部バス

90、91 記憶媒体

100、101 コンピュータプログラム

331 サーチデータ記憶部

20

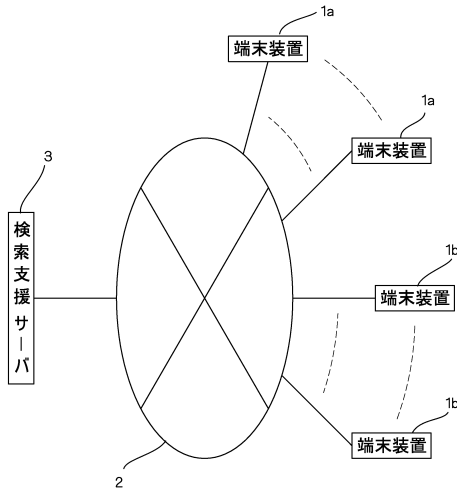
332 ガイドデータ記憶部

30

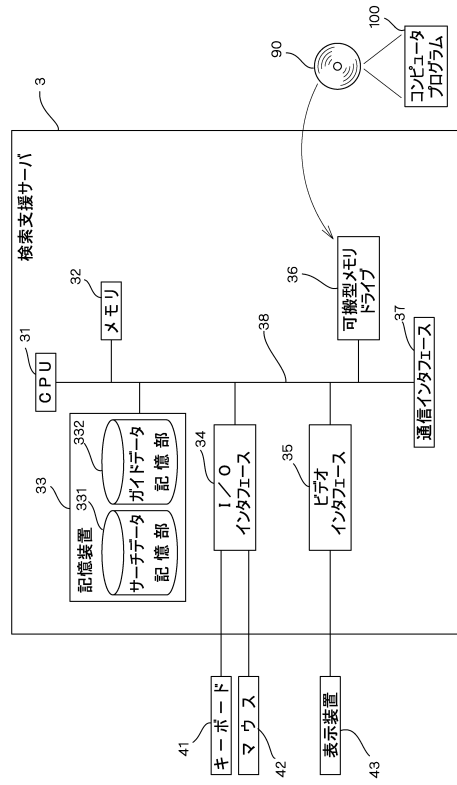
40

50

【図面】
【図 1】



【図 2】



10

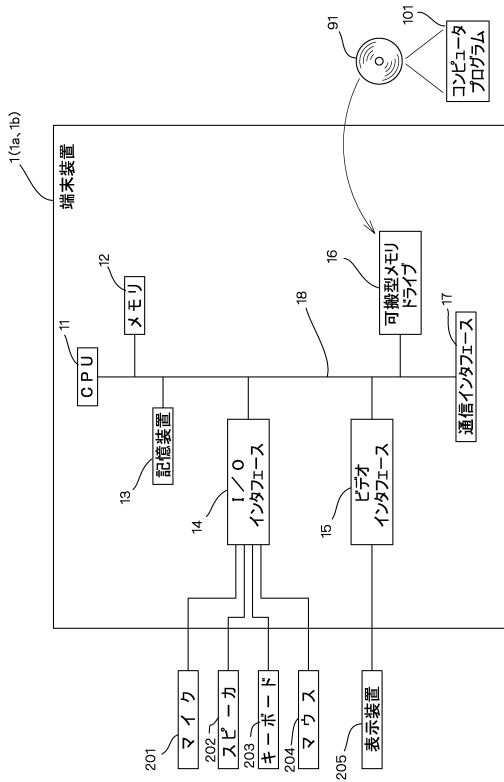
20

30

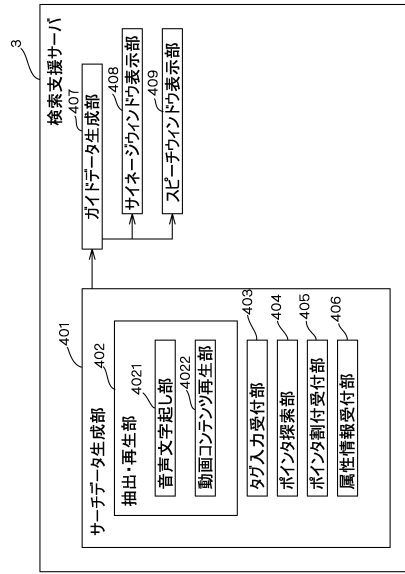
40

50

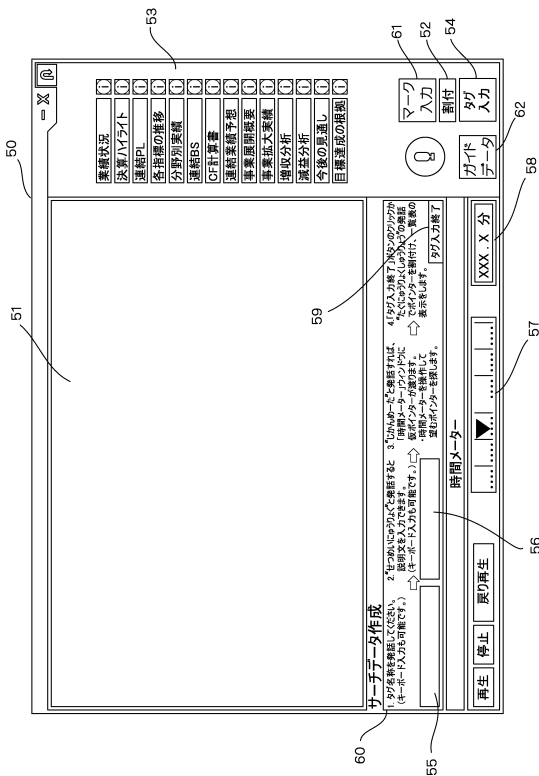
【図3】



【図4】



【図5】



【図6】

番号	タグ名称	説明	属性1	属性2	属性3	属性4	属性5	ユーザーID: XXXXXX	業号
1	2020年3月期決算説明	2020年3月期決算説明	AGP	AGS	AmiVoice	AGP	AGS	売上目録	...
2	決算ハイライト	11.5%増取、4%減益	AGP	AGS	AmiVoice	AGP	AGS	売上目録	...
3	新商品ロイヤリティ	各事業分野での対応などを説明しています	AGP	AGS	AmiVoice	AGP	AGS	売上目録	...
...
GN	今後の見通し	3年後、売上高80億円、営業利益2億円	AGP	AGS	AmiVoice	AGP	AGS	売上目録	...

番号	タグ名称	説明	属性1	属性2	属性3	属性4	属性5	ユーザーID: XXXXXX	業号
1	2020年3月期決算説明	2020年3月期決算説明	AGP	AGS	AmiVoice	AGP	AGS	売上目録	...
2	決算ハイライト	11.5%増取、4%減益	AGP	AGS	AmiVoice	AGP	AGS	売上目録	...
3	新商品ロイヤリティ	各事業分野での対応などを説明しています	AGP	AGS	AmiVoice	AGP	AGS	売上目録	...
...
SN	今後の見通し	3年後、売上高80億円、営業利益2億円	AGP	AGS	AmiVoice	AGP	AGS	売上目録	...

【図 7】

番号	タグ名称	説明	属性1	属性2	属性3	属性4	属性5	ユーザーID	ユーザーID
1
2	ACPとは？
...	オンライン診療
...	非対面営業
...	スクリーンアスタ
...	議事録作成支援システム
...	AmVoice EXの使い方
...

番号	タグ名称	説明	属性1	属性2	属性3	属性4	属性5	ユーザーID	ユーザーID
1
2
...	スクリーンアスタ
...	議事録作成支援システム
N-1
N

(a)

(b)

【図 8】

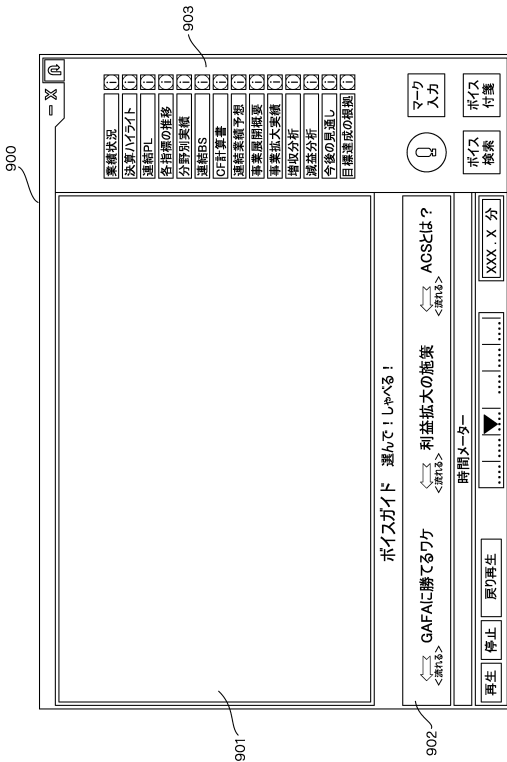
番号	タグ名称	説明	属性1	属性2	属性3	属性4	属性5	ユーザーID	ユーザーID
1
2	ACPとは？
...	オンライン診療
...	非対面営業
...	スクリーンアスタ
...	議事録作成支援システム
...	AmVoice EXの使い方
...

番号	タグ名称	説明	属性1	属性2	属性3	属性4	属性5	ユーザーID	ユーザーID
1
2
...	オンライン診療
...	非対面営業
...	COMPTON
G-1
G

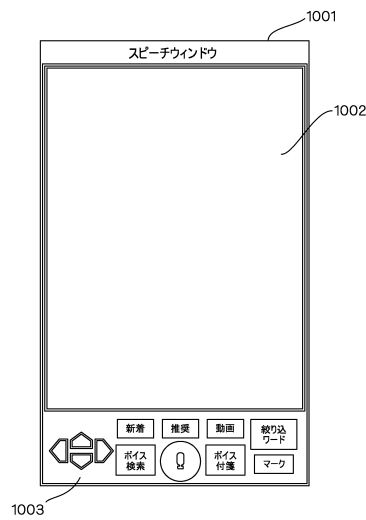
(a)

(b)

【図 9】



【図 10】



10

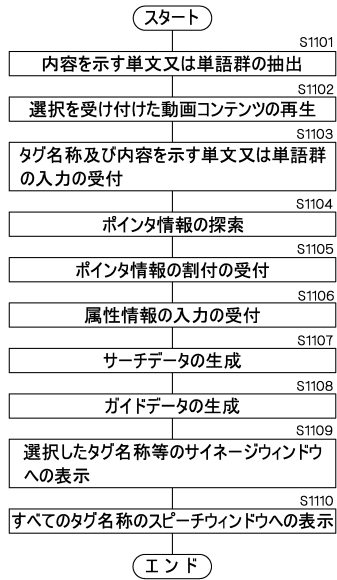
20

30

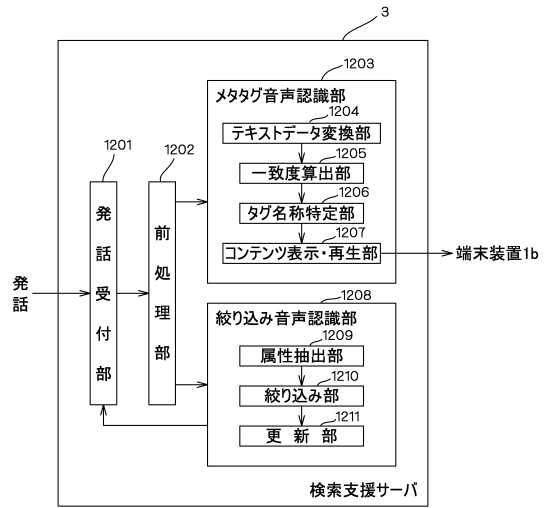
40

50

【 図 1 1 】



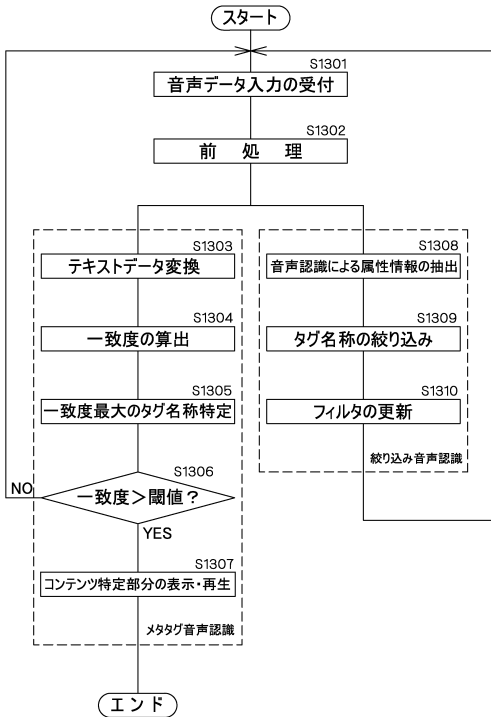
【 図 1 2 】



10

20

【 図 1 3 】



30

40

50

フロントページの続き

(51)国際特許分類

F I
G 0 6 F 3/16 6 2 0

東京都豊島区東池袋3丁目1番1号サンシャイン60 42階 株式会社アドバンスト・メディア内

審査官 三橋 竜太郎

(56)参考文献 特開2020-079982(JP,A)
特開2010-181461(JP,A)
特開2013-097578(JP,A)

(58)調査した分野 (Int.Cl., DB名)
G 0 6 F 1 6 / 0 0 - 1 6 / 9 5 8
G 1 0 L 1 5 / 0 0 - 1 5 / 2 2
G 0 6 F 3 / 1 6