

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第6605995号
(P6605995)

(45) 発行日 令和1年11月13日(2019.11.13)

(24) 登録日 令和1年10月25日(2019.10.25)

(51) Int.Cl. F I
G 1 O L 15/22 (2006.01) G 1 O L 15/22 4 7 O Z
 G 1 O L 15/22 4 6 O Z

請求項の数 12 (全 20 頁)

(21) 出願番号	特願2016-52983 (P2016-52983)	(73) 特許権者	000003078 株式会社東芝 東京都港区芝浦一丁目1番1号
(22) 出願日	平成28年3月16日(2016.3.16)	(74) 代理人	110002147 特許業務法人酒井国際特許事務所
(65) 公開番号	特開2017-167368 (P2017-167368A)	(72) 発明者	布目 光生 東京都港区芝浦一丁目1番1号 株式会社東芝内
(43) 公開日	平成29年9月21日(2017.9.21)	(72) 発明者	芦川 平 東京都港区芝浦一丁目1番1号 株式会社東芝内
審査請求日	平成30年2月22日(2018.2.22)	(72) 発明者	芦川 将之 東京都港区芝浦一丁目1番1号 株式会社東芝内

最終頁に続く

(54) 【発明の名称】 音声認識誤り修正装置、方法及びプログラム

(57) 【特許請求の範囲】

【請求項1】

第1発話者の第1発話音声を取得する第1取得部と、
 前記第1発話音声を認識し、第1テキストに変換する第1認識部と、
 前記第1テキストを示す第1字幕画像を出力する第1出力部と、
 前記第1テキストに含まれる修正対象文字列を修正するための第2発話者の第2発話音声を取得する第2取得部と、
 前記第2発話音声を認識し、第2テキストに変換する第2認識部と、
 前記第1テキストと前記第2テキストとのテキストマッチングに基づいて、前記第1テキストに含まれる文字列のうち、前記第2テキストとの類似度が閾値以上の文字列を前記修正対象文字列として推定する推定部と、
 前記修正対象文字列が前記第2テキストに置き換えられることを示す第2字幕画像を出力する第2出力部と、
 を備える音声認識誤り修正装置。

【請求項2】

前記推定部は、前記第1発話音声と前記第2発話音声との音響マッチングに更に基づいて、前記修正対象文字列を推定する請求項1に記載の音声認識誤り修正装置。

【請求項3】

前記推定部は、前記テキストマッチングと前記音響マッチングとに基づいて、前記第1テキストに含まれる文字列のうち、前記第2テキストとの類似度が閾値以上の文字列を前

記修正対象文字列に推定する請求項 2 に記載の音声認識誤り修正装置。

【請求項 4】

前記推定部は、前記第 1 テキストに前記第 2 テキストとの類似度が閾値以上となる文字列が複数含まれる場合、当該複数の文字列それぞれを修正候補文字列に推定し、

前記第 2 出力部は、前記第 1 テキストにおいて前記複数の修正候補文字列の中から前記第 2 テキストに置き換えられる修正候補文字列を前記第 2 発話者に選択させるための選択画像を出力し、

前記第 2 テキストに置き換えられる修正候補文字列を前記選択画像上で選択する選択入力を受け付ける受付部を更に備え、

前記推定部は、前記複数の修正候補文字列のうち、前記選択入力により選択された修正候補文字列を前記修正対象文字列に推定し、

前記第 2 出力部は、前記第 1 テキストにおいて前記修正対象文字列が前記第 2 テキストに置き換えられることが示された前記第 2 字幕画像を出力する請求項 3 に記載の音声認識誤り修正装置。

【請求項 5】

前記選択入力は、前記第 2 テキストに置き換えられる修正候補文字列上の位置又は当該修正候補文字列の近傍位置を前記選択画像上でタッチする入力であり、

前記推定部は、前記複数の修正候補文字列のうち、前記選択入力によりタッチされた位置に近い方の修正候補文字列を前記修正対象文字列に推定する請求項 4 に記載の音声認識誤り修正装置。

【請求項 6】

前記選択入力は、前記第 2 テキストに置き換えられる修正候補文字列上の位置又は当該修正候補文字列の近傍位置を前記選択画像上でタッチすることで、当該修正候補文字列を選択する入力である請求項 4 に記載の音声認識誤り修正装置。

【請求項 7】

前記選択入力は、前記第 2 テキストを、当該第 2 テキストに置き換えられる修正候補文字列上の位置又は当該修正候補文字列の近傍位置に移動させる入力であり、

前記推定部は、前記複数の修正候補文字列のうち、前記選択入力により前記第 2 テキストが移動された位置に近い方の修正候補文字列を前記修正対象文字列に推定する請求項 4 に記載の音声認識誤り修正装置。

【請求項 8】

前記第 2 テキストに置き換えられる文字列を前記第 1 字幕画像上で選択する選択入力を受け付ける受付部を更に備え、

前記推定部は、前記第 1 テキストに前記第 2 テキストとの類似度が閾値以上となる文字列が複数含まれる場合、前記選択入力により選択された文字列を前記修正対象文字列に推定する請求項 3 に記載の音声認識誤り修正装置。

【請求項 9】

前記選択入力は、前記第 2 テキストに置き換えられる文字列上の位置又は当該文字列の近傍位置を前記第 1 字幕画像上でタッチする入力であり、

前記推定部は、前記第 2 テキストとの類似度が閾値以上となる複数の文字列のうち、前記選択入力によりタッチされた位置に近い方の文字列を前記修正対象文字列に推定する請求項 7 に記載の音声認識誤り修正装置。

【請求項 10】

前記選択入力は、前記第 2 テキストに置き換えられる文字列上の位置又は当該文字列の近傍位置を前記第 1 字幕画像上でタッチすることで、当該文字列を選択する入力である請求項 7 に記載の音声認識誤り修正装置。

【請求項 11】

第 1 発話者の第 1 発話音声を取得する第 1 取得ステップと、

前記第 1 発話音声を認識し、第 1 テキストに変換する第 1 認識ステップと、

前記第 1 テキストを示す第 1 字幕画像を出力する第 1 出力ステップと、

10

20

30

40

50

前記第1テキストに含まれる修正対象文字列を修正するための第2発話者の第2発話音声を取得する第2取得ステップと、

前記第2発話音声を認識し、第2テキストに変換する第2認識ステップと、

前記第1テキストと前記第2テキストとのテキストマッチングに基づいて、前記第1テキストに含まれる文字列のうち、前記第2テキストとの類似度が閾値以上の文字列を前記修正対象文字列として推定する推定ステップと、

前記修正対象文字列が前記第2テキストに置き換えられることを示す第2字幕画像を出力する第2出力ステップと、

を含む音声認識誤り修正方法。

【請求項12】

第1発話者の第1発話音声を取得する第1取得ステップと、

前記第1発話音声を認識し、第1テキストに変換する第1認識ステップと、

前記第1テキストを示す第1字幕画像を出力する第1出力ステップと、

前記第1テキストに含まれる修正対象文字列を修正するための第2発話者の第2発話音声を取得する第2取得ステップと、

前記第2発話音声を認識し、第2テキストに変換する第2認識ステップと、

前記第1テキストと前記第2テキストとのテキストマッチングに基づいて、前記第1テキストに含まれる文字列のうち、前記第2テキストとの類似度が閾値以上の文字列を前記修正対象文字列として推定する推定ステップと、

前記修正対象文字列が前記第2テキストに置き換えられることを示す第2字幕画像を出力する第2出力ステップと、

をコンピュータに実行させるためのプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明の実施形態は、音声認識誤り修正装置、方法及びプログラムに関する。

【背景技術】

【0002】

従来から、発話者が発話した音声を音声認識処理でテキストに変換し、変換したテキストを字幕として出力する技術が知られている。

【0003】

またこのような技術において、音声認識処理による認識誤りを修正するため、修正者が、音声から変換されたテキスト上で認識誤り箇所を手動で選択し、選択した認識誤り箇所の文字列を修正するための正しい文字列をキーボードなどから手入力することで、認識誤り箇所の文字列を正しい文字列に修正した字幕を出力する技術が知られている。

【先行技術文献】

【特許文献】

【0004】

【特許文献1】特開2004-151614号公報

【発明の概要】

【発明が解決しようとする課題】

【0005】

しかしながら、上述したような従来技術では、認識誤りを修正するために、認識誤り箇所の選択かつ認識誤り箇所の文字列を修正するための手入力が必要であり、認識誤りの修正に手間を要していた。

【0006】

本発明が解決しようとする課題は、発話者が発話した音声を認識して字幕として出力する場合の認識誤りの修正を簡易化することができる音声認識誤り修正装置、方法及びプロ

10

20

30

40

50

グラムを提供することである。

【課題を解決するための手段】

【0007】

実施形態の音声認識誤り修正装置は、第1取得部と、第1認識部と、第1出力部と、第2取得部と、第2認識部と、推定部と、第2出力部と、を備える。第1取得部は、第1発話者の第1発話音声を取得する。第1認識部は、前記第1発話音声を認識し、第1テキストに変換する。第1出力部は、前記第1テキストを示す第1字幕画像を出力する。第2取得部は、前記第1テキストに含まれる修正対象文字列を修正するための第2発話者の第2発話音声を取得する。第2認識部は、前記第2発話音声を認識し、第2テキストに変換する。推定部は、前記第1テキストと前記第2テキストとのテキストマッチングに基づいて、前記修正対象文字列を推定する。第2出力部は、前記修正対象文字列が前記第2テキストに置き換えられることを示す第2字幕画像を出力する。

10

【図面の簡単な説明】

【0008】

【図1】第1実施形態の音声認識誤り修正システムの例を示す構成図。

【図2】第1実施形態の音声認識誤り修正装置の例を示す構成図。

【図3】第1実施形態の第1字幕画像の例を示す図。

【図4】第1実施形態の第2テキストの例を示す図。

【図5】第1実施形態の第2字幕画像の例を示す図。

【図6】第1実施形態の処理例を示すフローチャート。

20

【図7】第1実施形態の推定処理の例を示すフローチャート。

【図8】第1実施形態のテキストマッチング処理例を示すフローチャート。

【図9】第1実施形態の修正対象文字列の範囲の推定例の説明図。

【図10】第1実施形態の修正対象文字列の範囲の推定例の説明図。

【図11】第1実施形態の修正対象文字列の範囲の推定例の説明図。

【図12】第2実施形態の音声認識誤り修正装置の例を示す構成図。

【図13】第2実施形態の選択画像の例を示す図。

【図14】第2実施形態の選択入力の例の説明図。

【図15】第2実施形態の第2字幕画像の例を示す図。

【図16】第2実施形態の処理例を示すフローチャート。

30

【図17】変形例2の選択入力の例の説明図。

【図18】変形例2の選択入力の例の説明図。

【図19】変形例2の選択入力の例の説明図。

【図20】変形例3の選択入力の例の説明図。

【図21】各実施形態及び各変形例の音声認識誤り修正装置のハードウェア構成例を示す構成図。

【発明を実施するための形態】

【0009】

以下、添付図面を参照しながら、実施形態を詳細に説明する。

【0010】

40

(第1実施形態)

図1は、第1実施形態の音声認識誤り修正システム1の一例を示す構成図である。図1に示すように、音声認識誤り修正システム1は、音声認識誤り修正装置10と、端末装置20-1~20-Nと、を備える。

【0011】

音声認識誤り修正装置10と端末装置20-1~20-N(Nは、3以上の自然数)とは、ネットワーク2を介して接続されている。ネットワーク2は、例えば、LAN(Local Area Network)やインターネットなどにより実現できる。また、ネットワーク2は、有線のネットワークであっても無線のネットワークであってもよい。

【0012】

50

なお、以下の説明では、端末装置 20 - 1 ~ 20 - N を各々区別する必要がない場合は、単に端末装置 20 と称する場合がある。

【 0 0 1 3 】

音声認識誤り修正装置 10 は、例えば、音声認識用のサーバ装置などのコンピュータにより実現できる。端末装置 20 は、音声入力装置として機能するマイクや、操作装置及び表示装置として機能するタッチパネルディスプレイなどを備えたコンピュータであり、例えば、スマートフォン、タブレット端末、及びノート P C (Personal Computer) などにより実現できる。

【 0 0 1 4 】

第 1 実施形態では、音声認識誤り修正システム 1 が、授業やセミナーなどの講演内容をリアルタイムに字幕として提示するシステムである場合を例に取り説明するが、これに限定されるものではない。例えば、音声認識誤り修正システム 1 を、会議などの発言内容をリアルタイムに字幕として提示するシステムに適用してもよい。

10

【 0 0 1 5 】

以下では、講演者（第 1 発話者の一例）が端末装置 20 - 1 を使用し、講演内容の字幕を修正する修正者（第 2 発話者の一例）が端末装置 20 - 2 を使用し、講演内容を聴講する聴講者が端末装置 20 - N を使用する場合を例にとり説明するが、これに限定されるものではない。なお第 1 実施形態では、端末装置 20 - N を使用する聴講者は、音声の聞き取りが困難な聴覚障害者を想定しているが、これに限定されるものではない。

【 0 0 1 6 】

20

第 1 実施形態では、音声認識誤り修正システム 1 が、以下のような利用形態で利用される場合を例に取り説明する。

【 0 0 1 7 】

まず、講演者は、講演内容である第 1 発話音声を、マイクを介して端末装置 20 - 1 に音声入力し、端末装置 20 - 1 は、音声入力された第 1 発話音声を音声認識誤り修正装置 10 に送信する。音声認識誤り修正装置 10 は、端末装置 20 - 1 から受信した第 1 発話音声を音声認識して第 1 テキストに変換し、変換した第 1 テキストを示す第 1 字幕画像を端末装置 20 - 1 ~ 20 - N それぞれに配信する。各端末装置 20 は、音声認識誤り修正装置 10 から受信した第 1 字幕画像をタッチパネルディスプレイに表示する。

【 0 0 1 8 】

30

続いて、修正者が端末装置 20 - 2 上で第 1 字幕画像を確認し、第 1 テキストに認識誤りがある場合は、認識誤りを修正するための正しい内容である第 2 発話音声を、マイクを介して端末装置 20 - 2 に音声入力し（リスピークし）、端末装置 20 - 2 は、音声入力された第 2 発話音声を音声認識誤り修正装置 10 に送信する。音声認識誤り修正装置 10 は、端末装置 20 - 2 から受信した第 2 発話音声を音声認識して第 2 テキストに変換し、第 1 テキストにおける認識誤り箇所の修正対象文字列を第 2 テキストに置き換える第 2 字幕画像を端末装置 20 - 1 ~ 20 - N それぞれに配信する。各端末装置 20 は、音声認識誤り修正装置 10 から受信した第 2 字幕画像をタッチパネルディスプレイに表示する。

【 0 0 1 9 】

このように第 1 実施形態では、第 1 字幕画像が示す第 1 テキストに認識誤りがある場合には、当該認識誤りを修正した第 2 字幕画像が各端末装置 20 に配信されるため、聴講者は、字幕画像から講演内容を正しく理解することができる。

40

【 0 0 2 0 】

図 2 は、第 1 実施形態の音声認識誤り修正装置 10 の一例を示す構成図である。図 2 に示すように、音声認識誤り修正装置 10 は、第 1 取得部 11 と、第 1 認識部 13 と、生成部 15 と、第 1 出力部 17 と、第 2 取得部 19 と、第 2 認識部 21 と、推定部 23 と、更新部 25 と、第 2 出力部 27 と、を備える。

【 0 0 2 1 】

第 1 取得部 11、第 1 認識部 13、生成部 15、第 1 出力部 17、第 2 取得部 19、第 2 認識部 21、推定部 23、更新部 25、及び第 2 出力部 27 は、C P U (Central Pro

50

cessing Unit)などの処理装置にプログラムを実行させること、即ち、ソフトウェアにより実現してもよいし、IC(Integrated Circuit)などのハードウェアにより実現してもよいし、ソフトウェア及びハードウェアを併用して実現してもよい。

【0022】

第1取得部11は、第1発話者の第1発話音声を取得する。第1実施形態では、第1取得部11は、前述の通り、端末装置20-1から、第1発話者の第1発話音声として、講演者の講演内容の発話音声を取得する。

【0023】

第1認識部13は、第1取得部11により取得された第1発話音声を認識し、第1テキストに変換する。具体的には、第1認識部13は、第1取得部11により取得された第1発話音声を音声認識処理により認識し、第1発話音声の内容を第1テキストに変換する。なお、音声認識処理は、公知の技術を用いればよい。

10

【0024】

生成部15は、第1認識部13により変換された第1テキストを示す第1字幕画像を生成する。第1字幕画像は、第1テキストを示す画像であればどのような画像であってもよい。

【0025】

例えば、第1字幕画像としては、講演者、修正者、及び聴講者などの利用者が見やすいように、第1テキストのフォントサイズや背景色を変更した画像が挙げられる。このような画像として、いわゆる情報保障でよく見られるサブディスプレイで黒背景に第1テキストを白文字で表示するような形態が挙げられる。なお、一般に、聴覚障害者向けには、黒背景かつ白文字、横幅が20文字、縦が10行などの表示形式が知られている。

20

【0026】

また例えば、第1字幕画像として、講演者の映像など何らかの映像上に第1テキストをオーバーレイ表示する形態や、修正者向けの編集を前提として第1テキストを表示した形態なども挙げられる。

【0027】

但し、第1字幕画像は、これらに限定されるものではない。

【0028】

図3は、第1実施形態の第1字幕画像の一例を示す図である。図3に示す例では、第1字幕画像上に第1テキスト51が表示されている。なお、第1テキスト51には、音声認識処理による認識誤りも生じており、後述する修正対象の修正対象文字列52も含まれている。

30

【0029】

第1出力部17は、生成部15により生成された第1字幕画像を出力する。第1実施形態では、第1出力部17は、各端末装置20に第1字幕画像を出力する。

【0030】

これにより、各端末装置20では、第1字幕画像が表示される。ここで、修正者は、端末装置20-2上で第1字幕画像を確認し、第1テキストに認識誤りがあることを確認したとする。例えば、第1字幕画像が図3に示す画像である場合、修正者は、修正対象文字列52が認識誤り箇所であることを確認する。

40

【0031】

この場合、第2取得部19は、第1テキストに含まれる修正対象文字列を修正するための第2発話者の第2発話音声を取得する。第1実施形態では、第2取得部19は、前述の通り、端末装置20-2から、第2発話者の第2発話音声として、修正対象文字列を修正するための修正者の発話音声を取得する。

【0032】

第2認識部21は、第2取得部19により取得された第2発話音声を認識し、第2テキストに変換する。具体的には、第2認識部21は、第2取得部19により取得された第2発話音声を音声認識処理により認識し、第2発話音声の内容を第2テキストに変換する。

50

なお、音声認識処理は、第1認識部13同様、公知の技術を用いればよい。

【0033】

図4は、第1実施形態の第2テキストの一例を示す図である。図4に示す例では、図3に示す修正対象文字列52を修正するための第2テキスト53が示されている。

【0034】

推定部23は、第1認識部13により変換された第1テキストと第2認識部21により変換された第2テキストとのテキストマッチング(テキストベースマッチング)を行い、当該テキストマッチングの結果に基づいて、第1テキスト内の修正対象文字列を推定する。具体的には、推定部23は、第1認識部13により第1テキストに変換される第1発話音声と第2認識部21により第2テキストに変換される第2発話音声との音響マッチング(音響データマッチング)を更に行い、テキストマッチングの結果と音響マッチングの結果とに基づいて、第1テキスト内の修正対象文字列を推定する。

10

【0035】

第1実施形態では、推定部23は、以下のように、第1テキストと第2テキストとのテキストマッチングを行う。

【0036】

まず、推定部23は、第1テキスト及び第2テキストそれぞれに対し、形態素解析処理、及び読み推定処理を施し、第1テキスト及び第2テキストそれぞれから形態素情報及び読み情報を取得する。読み推定処理とは、テキストを揺らぎのない状態に修正する処理であり、例えば日本語であれば、漢字をひらがなやカタカナに修正することが挙げられる。

20

【0037】

次に、推定部23は、第1テキストの形態素情報及び読み情報を用いて、第1テキストを表層文字、形態素(品詞)の組み合わせ、及び読みの文字列の各値に展開するとともに、第2テキストの形態素情報及び読み情報を用いて、第2テキストを表層文字、形態素(品詞)の組み合わせ、及び読みの文字列の各値に展開する。

【0038】

次に、推定部23は、文字列(表層文字)単位、形態素単位、及び読み単位それぞれで、第1テキストに対する第2テキストのスキャンを行い、第1テキスト内から第2テキストに一致又は類似する文字列(類似ブロック)を探索する。

【0039】

そして推定部23は、類似ブロックが探索された場合、当該類似ブロックでの第2テキストとの文字列(表層文字)単位、形態素単位、及び読み単位それぞれでの類似度を算出し、算出した各類似度を統合し、テキスト類似度とする。各類似度の統合には、例えば、数式(1)を用いればよい。

30

【0040】

$$\text{TextSim}(a, b) = \text{Sim}_S(a, b) + \text{Sim}_M(a, b) + \text{Sim}_R(a, b) \quad \dots (1)$$

【0041】

ここで、aは、第2テキストを示す。bは、第1テキスト内の類似ブロックを示す。 $\text{Sim}_S(a, b)$ は、aとbとの文字列(表層文字)単位での類似度を示す。 $\text{Sim}_M(a, b)$ は、aとbとの形態素単位での類似度を示す。 $\text{Sim}_R(a, b)$ は、aとbとの読み単位での類似度を示す。 $\text{TextSim}(a, b)$ は、テキスト類似度を示す。

40

【0042】

また第1実施形態では、推定部23は、音声波形などの音響情報を用いて音声間のマッチング位置を認識するボイストリガなどの公知技術を用いて、第1発話音声と第2発話音声との音響マッチングを行い、音響類似度を算出する。なお、推定部23は、第1発話音声の比較区間(第2発話音声との比較が行われる第1発話音声の区間)をずらしながら第2発話音声との音響マッチングを行い、各比較区間での音響類似度を算出する。

【0043】

50

そして推定部 2 3 は、テキストマッチングの結果と音響マッチングの結果とに基づいて、第 1 テキストに含まれる文字列のうち、第 2 テキストとの類似度が閾値以上の文字列を修正対象文字列に推定する。

【 0 0 4 4 】

具体的には、推定部 2 3 は、算出したテキスト類似度毎に、当該テキスト類似度と、当該テキスト類似度の類似ブロックに対応する比較区間での音響類似度と、を統合し、推定類似度とする。テキスト類似度と音響類似度との統合には、例えば、数式 (2) を用いればよい。

【 0 0 4 5 】

$MergeSim(a, b) = TextSim(a, b) + VoiceSim(a', b') \dots (2)$ 10

【 0 0 4 6 】

ここで、 a' は、第 2 発話音声を示す。 b' は、類似ブロックに対応する第 1 発話音声内の比較区間を示す。 $VoiceSim(a', b')$ は、 a' と b' との音響類似度を示す。 $MergeSim(a, b)$ は、推定類似度を示す。

【 0 0 4 7 】

例えば推定部 2 3 は、推定類似度が閾値以上となる類似ブロックの文字列を修正対象文字列に推定する。なお第 1 実施形態では、推定類似度が閾値以上となる類似ブロックが 1 つである場合を想定して説明する。

【 0 0 4 8 】

なお、第 1 テキストが図 3 に示す第 1 テキスト 5 1 であり、第 2 テキストが図 4 に示す第 2 テキスト 5 3 である場合、推定部 2 3 は、このように修正対象文字列を推定することで、第 1 テキスト 5 1 から修正対象文字列 5 2 を推定する。 20

【 0 0 4 9 】

更新部 2 5 は、生成部 1 5 により生成された第 1 字幕画像を、第 1 テキストにおいて修正対象文字列が第 2 テキストに置き換えられることが示された第 2 字幕画像に更新する。

【 0 0 5 0 】

図 5 は、第 1 実施形態の第 2 字幕画像の一例を示す図である。図 5 に示す例では、第 2 字幕画像上に第 1 テキスト 5 1 が表示されている。また、図 5 に示す例では、第 1 テキスト 5 1 内の修正対象文字列 5 2 がハイライトされて表示されている。また、図 5 に示す例では、リスピーク表示として第 2 テキスト 5 3 が表示されるとともに、修正対象文字列 5 2 が第 2 テキスト 5 3 に置き換えられることが示されている。 30

【 0 0 5 1 】

但し、第 2 字幕画像の表示態様は、これに限定されるものではなく、例えば、修正対象文字列 5 2 を第 2 テキスト 5 3 に予め置き換えておいてもよいし、図 5 に示すような表示を一定時間行った後に修正対象文字列 5 2 を第 2 テキスト 5 3 に置き換えるようにしてもよい。

【 0 0 5 2 】

第 2 出力部 2 7 は、更新部 2 5 により更新された第 2 字幕画像を出力する。第 1 実施形態では、第 2 出力部 2 7 は、各端末装置 2 0 に第 2 字幕画像を出力する。 40

【 0 0 5 3 】

これにより、各端末装置 2 0 では、第 2 字幕画像が表示される。このように第 1 実施形態では、第 1 字幕画像が示す第 1 テキストに認識誤りがある場合には、当該認識誤りを修正した第 2 字幕画像が各端末装置 2 0 に配信されるため、聴講者は、字幕画像から講演内容を正しく理解することができる。

【 0 0 5 4 】

図 6 は、第 1 実施形態の音声認識誤り修正装置 1 0 で行われる処理の流れの一例を示すフローチャートである。

【 0 0 5 5 】

まず、第 1 取得部 1 1 は、端末装置 2 0 - 1 から、第 1 発話者の第 1 発話音声として、 50

講演者の講演内容の発話音声を取得する（ステップS101）。

【0056】

続いて、第1認識部13は、第1取得部11により取得された第1発話音声を音声認識処理により認識し、第1発話音声の内容を第1テキストに変換する（ステップS103）。

【0057】

続いて、生成部15は、第1認識部13により変換された第1テキストを示す第1字幕画像を生成する（ステップS105）。

【0058】

続いて、第1出力部17は、生成部15により生成された第1字幕画像を各端末装置20に出力する（ステップS107）。

10

【0059】

続いて、第2取得部19が、端末装置20-2から、第2発話者の第2発話音声として、修正対象文字列を修正するための修正者の発話音声を取得すると（ステップS109でYes）、第2認識部21は、第2取得部19により取得された第2発話音声を音声認識処理により認識し、第2発話音声の内容を第2テキストに変換する（ステップS111）。

【0060】

続いて、推定部23は、第1テキスト内の修正対象文字列を推定する推定処理を行う（ステップS113）。なお、推定処理の詳細については、後述する。

20

【0061】

続いて、更新部25は、生成部15により生成された第1字幕画像を、第1テキストにおいて修正対象文字列が第2テキストに置き換えられることが示された第2字幕画像に更新する（ステップS115）。

【0062】

続いて、第2出力部27は、更新部25により更新された第2字幕画像を各端末装置20に出力する（ステップS117）。

【0063】

なお、第2取得部19が、端末装置20-2から、第2発話者の第2発話音声を取得しない場合（ステップS109でNo）、ステップS111以降の処理は行われない。

30

【0064】

図7は、図6に示すフローチャートのステップS113の推定処理の流れの一例を示すフローチャートである。

【0065】

まず、推定部23は、第1認識部13から第1テキストを取得する（ステップS121）。なお、推定部23は、生成部15から第1テキストを取得してもよい。

【0066】

続いて、推定部23は、第2認識部21から第2テキストを取得する（ステップS123）。

【0067】

続いて、推定部23は、第1テキスト及び第2テキストそれぞれに対し、形態素解析処理を施し、第1テキスト及び第2テキストそれぞれから形態素情報を取得する（ステップS125）。

40

【0068】

続いて、推定部23は、第1テキスト及び第2テキストそれぞれに対し、読み推定処理を施し、第1テキスト及び第2テキストそれぞれから読み情報を取得する（ステップS127）。

【0069】

続いて、推定部23は、第1テキストの形態素情報及び読み情報、並びに第2テキストの形態素情報及び読み情報を用いて、第1テキストと第2テキストとのテキストマッチン

50

グ処理を行う（ステップS 1 2 9）。

【0070】

ここで、図8を参照しながら、テキストマッチング処理を説明する。図8は、図7に示すフローチャートのステップS 1 2 9のテキストマッチング処理の流れの一例を示すフローチャートである。

【0071】

まず、推定部23は、第1テキストの形態素情報及び読み情報を用いて、第1テキストを表層文字、形態素（品詞）の組み合わせ、及び読みの文字列の各値に展開するとともに、第2テキストの形態素情報及び読み情報を用いて、第2テキストを表層文字、形態素（品詞）の組み合わせ、及び読みの文字列の各値に展開しておく。

10

【0072】

続いて、推定部23は、文字列（表層文字）単位で、第1テキストに対する第2テキストのスキャンを行い、第1テキスト内から第2テキストに一致又は類似する文字列（類似ブロック）を探索する（ステップS 1 4 1）。

【0073】

続いて、推定部23は、形態素単位で、第1テキストに対する第2テキストのスキャンを行い、第1テキスト内から第2テキストに一致又は類似する文字列（類似ブロック）を探索する（ステップS 1 4 3）。

【0074】

続いて、推定部23は、読み単位で、第1テキストに対する第2テキストのスキャンを行い、第1テキスト内から第2テキストに一致又は類似する文字列（類似ブロック）を探索する（ステップS 1 4 5）。

20

【0075】

続いて、推定部23は、探索された類似ブロック毎に、当該類似ブロックでの第2テキストとの文字列（表層文字）単位、形態素単位、及び読み単位それぞれでの類似度を算出し、算出した各類似度を統合することで、テキスト類似度を算出する（ステップS 1 4 7）。

【0076】

続いて、図7に戻り、推定部23は、第1認識部13から第1発話音声を取得するとともに、第2認識部21から第2発話音声を取得し、音響マッチング処理を行う（ステップS 1 3 1）。具体的には、推定部23は、第1発話音声の比較区間（第2発話音声との比較が行われる第1発話音声の区間）をずらしながら第2発話音声との音響マッチングを行い、各比較区間での音響類似度を算出する。

30

【0077】

続いて、推定部23は、算出したテキスト類似度毎に、当該テキスト類似度と、当該テキスト類似度の類似ブロックに対応する比較区間での音響類似度と、を統合することで両マッチング処理結果を統合し、推定類似度とする（ステップS 1 3 3）。

【0078】

続いて、推定部23は、統合結果である推定類似度に基づいて、第1テキスト内の修正対象文字列を推定する（ステップS 1 3 5）。例えば推定部23は、推定類似度が閾値以上となる類似ブロックの文字列を修正対象文字列に推定する。

40

【0079】

以上のように第1実施形態によれば、発話者が発話した音声を認識して字幕として出力する場合の認識誤りを、認識誤りを修正するための正しい内容を復唱（リスピーク）するだけで修正できるため、認識誤りの修正を簡易化でき、作業の手間やコストが削減できる。

【0080】

また第1実施形態の手法では、テキストマッチングや音響マッチングにより言語的な特徴や音響的な特徴を利用して認識誤り箇所を自動で推定するため、認識誤り箇所の選択かつ認識誤り箇所の文字列を修正するための手入力を省略することが可能となる。

50

【 0 0 8 1 】

特に第 1 実施形態の手法では、文字列（表層文字）単位、形態素単位、及び読み単位それぞれで、テキストマッチングを行うため、固定された単位ではなく様々な単位で修正対象文字列（の範囲）を推定することができる。

【 0 0 8 2 】

例えば、図 9 に示す例では、第 2 テキスト 6 3 が名詞であるため、第 1 テキスト 6 1 の中から名詞に相当する修正対象文字列 6 2 を推定しており、品詞の出現類似性で修正対象文字列（の範囲）を推定することができる。

【 0 0 8 3 】

また例えば、図 1 0 に示す例では、第 2 テキスト 6 6 に対し、第 1 テキスト 6 4 の中から修正対象文字列 6 5 を推定しており、形態素単位での読みの文字列の類似性で修正対象文字列（の範囲）を推定することができる。

10

【 0 0 8 4 】

また例えば、図 1 1 に示す例では、第 2 テキスト 6 9 に対し、第 1 テキスト 6 7 の中から修正対象文字列 6 8 を推定しており、文字列の距離と当該文字列に含まれる品詞内容の類似性で修正対象文字列（の範囲）を推定することができる。

【 0 0 8 5 】

このように第 1 実施形態の手法では、認識誤り箇所の選択かつ認識誤り箇所の文字列を修正するための手入力が不要であるため、ある程度のスキルを要する修正者でなくても認識誤りの修正を行うことが可能であり、万人が認識誤りの修正を行うことが可能となる。

20

【 0 0 8 6 】

このため、これまで特別な会議や講演、授業などでしか情報保障をサポートできなかった状況でも、より多くの人々が簡単にサポートに回ることができ、字幕の精度を高めることが可能となる。

【 0 0 8 7 】

（第 2 実施形態）

第 2 実施形態では、第 1 テキスト内に第 2 テキストへの置き換え候補となる文字列が複数存在する場合について説明する。以下では、第 1 実施形態との相違点の説明を主に言い、第 1 実施形態と同様の機能を有する構成要素については、第 1 実施形態と同様の名称・符号を付し、その説明を省略する。

30

【 0 0 8 8 】

図 1 2 は、第 2 実施形態の音声認識誤り修正装置 1 1 0 の一例を示す構成図である。図 1 2 に示すように、第 2 実施形態の音声認識誤り修正装置 1 1 0 では、推定部 1 2 3、更新部 1 2 5、第 2 出力部 1 2 7、及び受付部 1 2 9 が、第 1 実施形態と相違する。

【 0 0 8 9 】

なお、受付部 1 2 9 は、CPU などの処理装置にプログラムを実行させること、即ち、ソフトウェアにより実現してもよいし、IC などのハードウェアにより実現してもよいし、ソフトウェア及びハードウェアを併用して実現してもよい。

【 0 0 9 0 】

推定部 1 2 3 は、第 1 認識部 1 3 により変換された第 1 テキストに第 2 認識部 2 1 により変換された第 2 テキストとの類似度が閾値以上となる文字列が複数含まれる場合、当該複数の文字列それぞれを修正候補文字列に推定する。

40

【 0 0 9 1 】

更新部 1 2 5 は、生成部 1 5 により生成された第 1 字幕画像を、第 1 テキストにおいて複数の修正候補文字列の中から第 2 テキストに置き換えられる修正候補文字列を修正者に選択させるための選択画像に更新する。

【 0 0 9 2 】

図 1 3 は、第 2 実施形態の選択画像の一例を示す図である。図 1 3 に示す例では、選択画像上に第 1 テキスト 1 5 1 が表示されている。また、図 1 3 に示す例では、リスピーク表示として第 2 テキスト 1 5 4 が表示されている。また、図 1 3 に示す例では、第 2 テキ

50

スト 1 5 4 に置き換えられる候補となる修正候補文字列 1 5 2、1 5 3 が含まれている。
 なお、修正候補文字列 1 5 2、1 5 3 をハイライトして表示するようにしてもよい。

【 0 0 9 3 】

第 2 出力部 1 2 7 は、更新部 1 2 5 により更新された選択画像を出力する。第 2 実施形態では、第 2 出力部 1 2 7 は、修正者が使用する端末装置 2 0 - 2 に選択画像を出力する。

【 0 0 9 4 】

受付部 1 2 9 は、第 2 テキストに置き換えられる修正候補文字列を選択画像上で選択する選択入力を受け付ける。第 2 実施形態では、受付部 1 2 9 は、端末装置 2 0 - 2 から、選択入力として、第 2 テキストに置き換えられる修正候補文字列上の位置又は当該修正候補文字列の近傍位置を選択画像上でタッチする入力（選択画像上のタッチされた位置の位置座標の入力）を受け付ける。

10

【 0 0 9 5 】

例えば、受付部 1 2 9 は、図 1 4 に示すように、修正者がタッチペン 2 0 1 などで修正候補文字列 1 5 2 の近傍位置 1 5 5 を選択画像上でタッチ（ポインティング、又はタップ）する入力を受け付ける。

【 0 0 9 6 】

推定部 1 2 3 は、複数の修正候補文字列のうち、受付部 1 2 9 により受け付けられた選択入力により選択された修正候補文字列を修正対象文字列に推定する。第 2 実施形態では、推定部 1 2 3 は、複数の修正候補文字列のうち、選択入力によりタッチされた位置に近い方の修正候補文字列を修正対象文字列に推定する。修正候補文字列と選択入力によりタッチされた位置との距離に応じたスコアは、例えば、数式（3）により求められる。このため、このスコアを用いることで、複数の修正候補文字列のうち、選択入力によりタッチされた位置に近い方の修正候補文字列を特定できる。

20

【 0 0 9 7 】

$$Score(x) = 1(\sqrt{2\pi}) \exp(-x^2/2) \quad \dots (3)$$

【 0 0 9 8 】

なお、x の値は、数式（4）に示す通りである。

【 0 0 9 9 】

$x = \text{ポインティングされた位置から修正候補文字列までの距離} (px) / \text{修正候補文字列のフォントサイズ} (px) \quad \dots (4)$

30

【 0 1 0 0 】

従って、図 1 3 に示す例の場合であれば、近傍位置 1 5 5 は、修正候補文字列 1 5 3 よりも修正候補文字列 1 5 2 の方が近いため、推定部 1 2 3 は、修正候補文字列 1 5 2 を修正対象文字列に推定する。

【 0 1 0 1 】

更新部 1 2 5 は、選択画像を、第 1 テキストにおいて修正対象文字列が第 2 テキストに置き換えられることが示された第 2 字幕画像に更新する。

【 0 1 0 2 】

図 1 5 は、第 2 実施形態の第 2 字幕画像の一例を示す図である。図 1 5 に示す例では、第 2 字幕画像上に、第 1 テキスト 1 5 1 に含まれる修正候補文字列 1 5 2 が、第 2 テキスト 1 5 4 を漢字に変換した第 2 テキスト 1 5 7 に置き換えられた第 1 テキスト 1 5 6 が表示されている。

40

【 0 1 0 3 】

第 2 出力部 1 2 7 は、更新部 1 2 5 により更新された第 2 字幕画像を出力する。第 2 実施形態では、第 2 出力部 1 2 7 は、各端末装置 2 0 に第 2 字幕画像を出力する。

【 0 1 0 4 】

図 1 6 は、第 2 実施形態の音声認識誤り修正装置 1 1 0 で行われる処理の流れの一例を示すフローチャートである。

【 0 1 0 5 】

50

まず、ステップS 2 0 1 ~ S 2 1 3までの処理は、図6に示すフローチャートのステップS 1 0 1 ~ S 1 1 3までの処理と同様である。

【0106】

続いて、推定部123は、第1認識部13により変換された第1テキストに第2認識部21により変換された第2テキストとの類似度が閾値以上となる文字列が複数含まれる場合(ステップS 2 1 5でYes)、当該複数の文字列それぞれを修正候補文字列に推定する。そして更新部125は、生成部15により生成された第1字幕画像を、第1テキストにおいて複数の修正候補文字列の中から第2テキストに置き換えられる修正候補文字列を修正者に選択させるための選択画像に更新する(ステップS 2 1 7)。

【0107】

続いて、第2出力部127は、更新部125により更新された選択画像を端末装置20-2に出力する(ステップS 2 1 9)。

【0108】

続いて、受付部129は、端末装置20-2から、選択入力として、第2テキストに置き換えられる修正候補文字列上の位置又は当該修正候補文字列の近傍位置を選択画像上でタッチする入力(選択画像上のタッチされた位置の位置座標の入力)を受け付ける(ステップS 2 2 1)。

【0109】

続いて、推定部123は、複数の修正候補文字列のうち、選択入力によりタッチされた位置に近い方の修正候補文字列を修正対象文字列に推定する(ステップS 2 2 3)。

【0110】

以降のステップS 2 2 5 ~ S 2 2 7までの処理は、図6に示すフローチャートのステップS 1 1 5 ~ S 1 1 7までの処理と同様である。

【0111】

なおステップS 2 1 5において、推定部123は、第1認識部13により変換された第1テキストに第2認識部21により変換された第2テキストとの類似度が閾値以上となる文字列が単数である場合(ステップS 2 1 5でNo)、当該文字列を修正対象文字列に推定し、ステップS 2 2 5へ進む。

【0112】

以上のように第2実施形態によれば、第2テキストにより置き換えられる候補となる文字列が第1テキスト内に複数存在する場合であっても、修正者が意図した文字列を第2テキストに置き換えることができる。

【0113】

特に第2実施形態によれば、修正者は、修正対象となる認識誤り箇所を正確に選択する必要はなく(認識誤り箇所の始点及び終点を厳密に指定する必要はなく)、修正対象となる認識誤り箇所の近傍をタッチ(タップ)するなどの曖昧な指定で足りるため、認識誤り箇所を選択する場合であっても、認識誤り箇所の選択を簡易化することができ、作業の時間やコストが削減できる。

【0114】

なお第2実施形態では、タッチ(ポインティング、又はタップ)により第2テキストに置き換えられる修正候補文字列を選択する場合を例に取り説明したが、これに限定されず、例えば、タップ及びホールドにより第2テキストに置き換えられる修正候補文字列を選択するようにしてもよい。

【0115】

(変形例1)

上記第2実施形態では、修正者による第2発話音声の入力の後に、修正者による第2テキストに置き換えられる修正候補文字列の選択が行われる場合を例に取り説明したが、第2発話音声の入力と第2テキストに置き換えられる文字列の選択とを、略同時に行うようにしてもよい。

【0116】

10

20

30

40

50

この場合、修正者は、第2テキストに置き換えられる文字列を第1字幕画像上で選択しながら、第2発話音声を入力するようにすればよい。つまり、第2取得部19による第2発話音声の取得と略同時に、受付部129は、第2テキストに置き換えられる文字列を第1字幕画像上で選択する選択入力を受け付ける。第2実施形態では、受付部129は、端末装置20-2から、選択入力として、第2テキストに置き換えられる文字列上の位置又は当該文字列の近傍位置を第1字幕画像上でタッチする入力（第1字幕画像上のタッチされた位置の位置座標の入力）を受け付ける。

【0117】

そして推定部123は、第1テキストに第2テキストとの類似度が閾値以上となる文字列が複数含まれる場合、選択入力により選択された文字列を修正対象文字列に推定する。第2実施形態では、推定部123は、第2テキストとの類似度が閾値以上となる複数の文字列のうち、選択入力によりタッチされた位置に近い方の文字列を修正対象文字列に推定する。

【0118】

以上のように変形例1においても、第2実施形態と同様の効果を奏することができる。

【0119】

（変形例2）

上記第2実施形態において、選択入力を、第2テキストに置き換えられる修正候補文字列上の位置又は当該修正候補文字列の近傍位置を選択画像上でタッチすることで、当該修正候補文字列を選択する入力としてもよい。

【0120】

例えば、図17に示すように、修正者がタッチペン201などで修正候補文字列152の近傍位置を選択画像上で最初にタッチ（ポインティング、又はタップ）すると、範囲161が選択される。続いて、図18に示すように、修正者がタッチペン201などで修正候補文字列152の近傍位置を選択画像上でもう一度タッチ（ポインティング、又はタップ）すると、範囲162が選択される。続いて、図19に示すように、修正者がタッチペン201などで修正候補文字列152の近傍位置を選択画像上でもう一度タッチ（ポインティング、又はタップ）すると、範囲163が選択され、修正候補文字列152が選択される。

【0121】

このように、修正候補文字列152の近傍位置を繰り返しタッチすることで、修正候補文字列152を含む文字列を、範囲を変えながら選択するようにしてもよい。なお、範囲の変更は、形態素単位で縮小させるようにしても拡張させるようにしてもよい。また、タッチで選択範囲を変更し、その後、ダブルタッチ（ダブルタップ）を行うことで、選択範囲に含まれる修正候補文字列を第2テキストに置き換えるようにしてもよい。

【0122】

同様に変形例1において、選択入力を、第2テキストに置き換えられる文字列上の位置又は当該文字列の近傍位置を第1字幕画像上でタッチすることで、当該文字列を選択する入力としてもよい。

【0123】

（変形例3）

上記第2実施形態において、選択入力を、第2テキストを、当該第2テキストに置き換えられる修正候補文字列上の位置又は当該修正候補文字列の近傍位置に移動させる入力とし、推定部123は、複数の修正候補文字列のうち、選択入力により第2テキストが移動された位置に近い方の修正候補文字列を修正対象文字列に推定するようにしてもよい。

【0124】

例えば、受付部129は、図20に示すように、修正者がタッチペン201などで第2テキスト154を修正候補文字列152の方向171へドラッグ&ドロップする入力を受け付け、推定部123は、複数の修正候補文字列のうち、ドラッグ&ドロップにより第2テキスト154が移動された位置に近い方の修正候補文字列152を修正対象文字列に推

10

20

30

40

50

定するようにしてもよい。

【0125】

(ハードウェア構成)

図21は、上記各実施形態及び各変形例の音声認識誤り修正装置のハードウェア構成の一例を示す構成図である。図21に示すように、上記各実施形態及び各変形例の音声認識誤り修正装置は、CPUなどの制御装置901と、ROMやRAMなどの主記憶装置902と、HDDやSSDなどの補助記憶装置903と、ディスプレイなどの表示装置904と、マウスやキーボードなどの入力装置905と、通信装置906とを、備えており、通常のコンピュータを利用したハードウェア構成で実現できる。

【0126】

上記各実施形態及び各変形例の音声認識誤り修正装置で実行されるプログラムは、ROM等に予め組み込んで提供される。

【0127】

また、上記各実施形態及び各変形例の音声認識誤り修正装置で実行されるプログラムを、インストール可能な形式又は実行可能な形式のファイルでCD-ROM、CD-R、メモリカード、DVD、フレキシブルディスク(FD)等のコンピュータで読み取り可能な記憶媒体に記憶されて提供するようにしてもよい。

【0128】

また、上記各実施形態及び各変形例の音声認識誤り修正装置で実行されるプログラムを、インターネット等のネットワークに接続されたコンピュータ上に格納し、ネットワーク経由でダウンロードさせることにより提供するようにしてもよい。また、上記各実施形態及び各変形例の音声認識誤り修正装置で実行されるプログラムを、インターネット等のネットワーク経由で提供または配布するようにしてもよい。

【0129】

上記各実施形態及び各変形例の音声認識誤り修正装置で実行されるプログラムは、上述した各部をコンピュータ上で実現させるためのモジュール構成となっている。実際のハードウェアとしては、例えば、制御装置901が補助記憶装置903からプログラムを主記憶装置902上に読み出して実行することにより、上記各部がコンピュータ上で実現されるようになっている。

【0130】

以上説明したとおり、上記各実施形態及び各変形例によれば、発話者が発話した音声認識して字幕として出力する場合の認識誤りの修正を簡易化することができる。

【0131】

なお本発明は、上記各実施形態そのままに限定されるものではなく、実施段階ではその要旨を逸脱しない範囲で構成要素を変形して具体化することができる。また上記各実施形態に開示されている複数の構成要素の適宜な組み合わせにより、種々の発明を形成することができる。例えば、実施形態に示される全構成要素からいくつかの構成要素を削除してもよい。さらに、異なる実施形態にわたる構成要素を適宜組み合わせても良い。

【0132】

例えば、上記実施形態のフローチャートにおける各ステップを、その性質に反しない限り、実行順序を変更し、複数同時に実施し、あるいは実施毎に異なった順序で実施してもよい。

【符号の説明】

【0133】

- 1、110 音声認識誤り修正システム
- 2 ネットワーク
- 10 音声認識誤り修正装置
- 20 - 1 ~ 20 - N (20) 端末装置
- 11 第1取得部
- 13 第1認識部

10

20

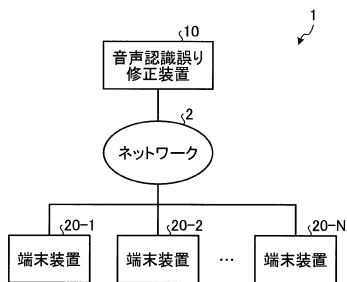
30

40

50

- 1 5 生成部
- 1 7 第1出力部
- 1 9 第2取得部
- 2 1 第2認識部
- 2 3、1 2 3 推定部
- 2 5、1 2 5 更新部
- 2 7、1 2 7 第2出力部
- 1 2 9 受付部
- 9 0 1 制御装置
- 9 0 2 主記憶装置
- 9 0 3 外部記憶装置
- 9 0 4 表示装置
- 9 0 5 入力装置
- 9 0 6 通信装置

【図1】

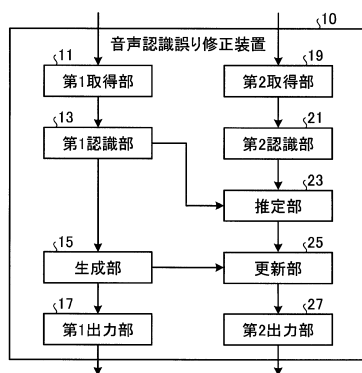


【図3】

51
 えとまずしそのあの一、地震の時のリスクという観点から言うと、
 まず一番遠い地を安全衛生推進員としては一番気になっているのが諸侯の本の学科です。
 あの一、まあ、大人がほとんどなので致命的なダメージはないんですけども
 あの一、一撃で当たったりする等、

52

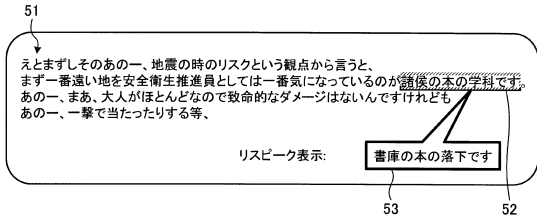
【図2】



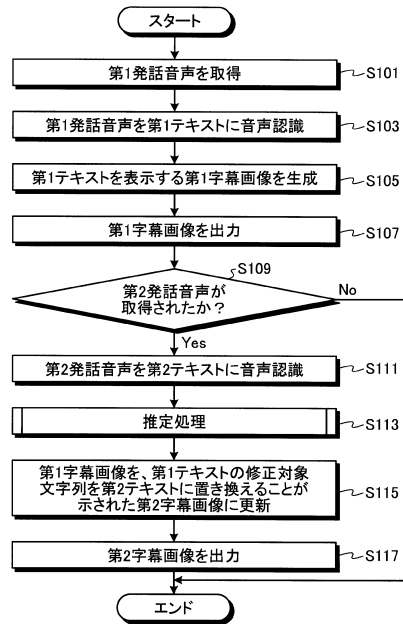
【図4】

53
 書庫の本の落下です

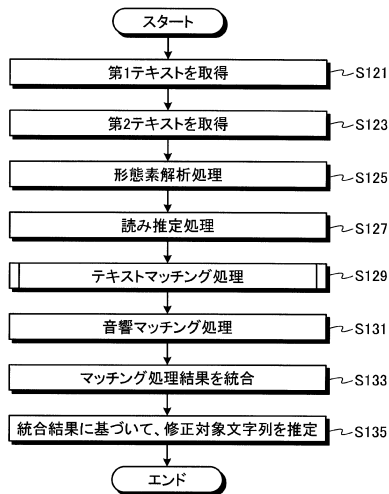
【図5】



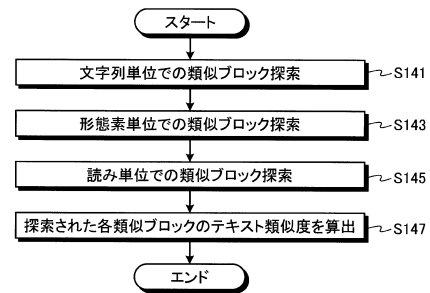
【図6】



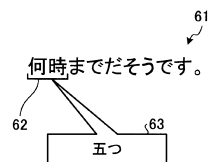
【図7】



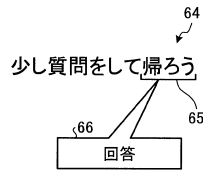
【図8】



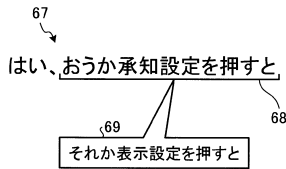
【図9】



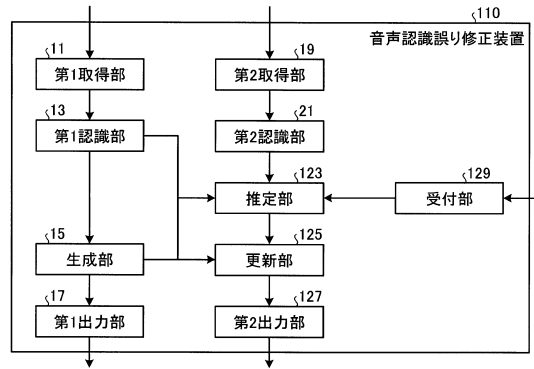
【図10】



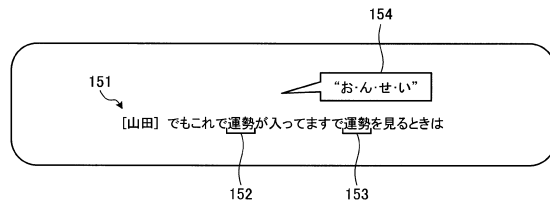
【図11】



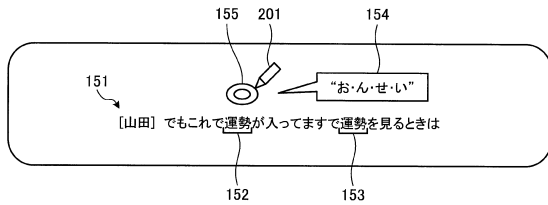
【図12】



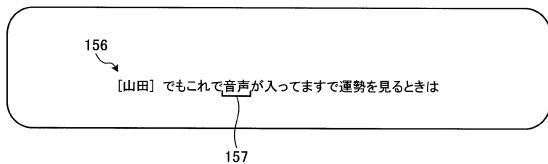
【図13】



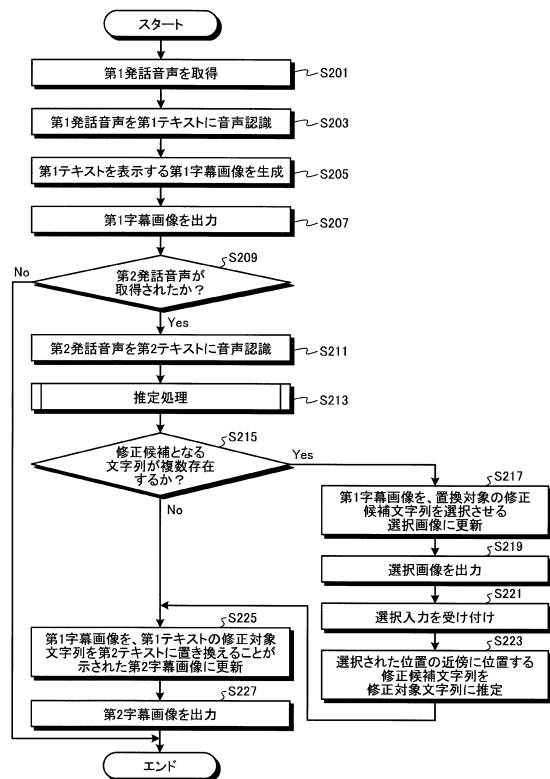
【図14】



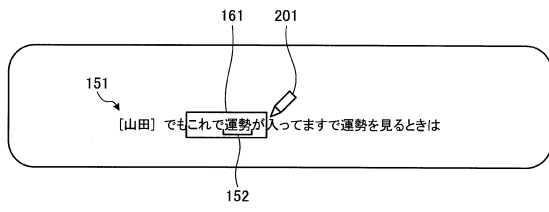
【図15】



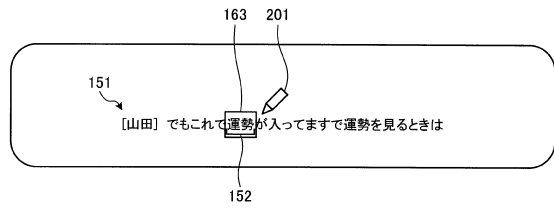
【図16】



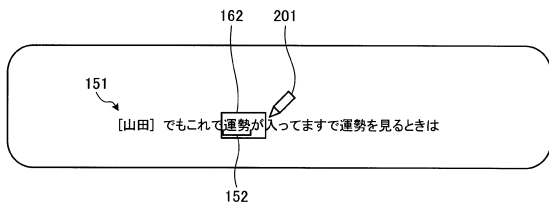
【図17】



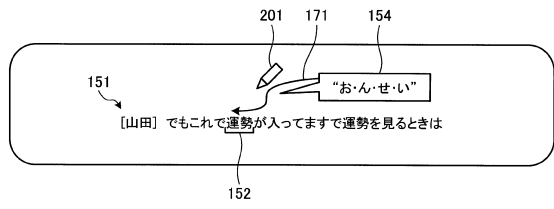
【図19】



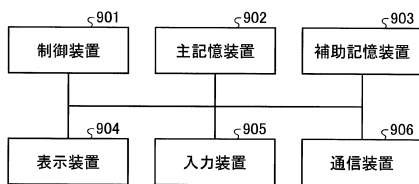
【図18】



【図20】



【図21】



フロントページの続き

審査官 千本 潤介

- (56)参考文献 特開2002-287792(JP,A)
特開平04-254896(JP,A)
特開2001-060192(JP,A)
特開2000-010586(JP,A)
特開2007-093789(JP,A)
特開平05-143593(JP,A)
特開2001-042996(JP,A)
特表2004-508594(JP,A)
米国特許出願公開第2016/0063994(US,A1)

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