

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第5184234号
(P5184234)

(45) 発行日 平成25年4月17日(2013.4.17)

(24) 登録日 平成25年1月25日(2013.1.25)

(51) Int.Cl.	F I	
G 1 O K 15/04 (2006.01)	G 1 O K	15/04 3 0 2 D
G 1 O L 15/00 (2013.01)	G 1 O L	15/00 2 0 0 G
G 1 O L 15/22 (2006.01)	G 1 O L	15/22 4 5 3
H O 4 N 5/445 (2011.01)	H O 4 N	5/445 Z
H O 4 N 7/173 (2011.01)	H O 4 N	7/173 6 3 0

請求項の数 18 (全 18 頁)

(21) 出願番号	特願2008-176844 (P2008-176844)	(73) 特許権者	000005821
(22) 出願日	平成20年7月7日(2008.7.7)		パナソニック株式会社
(65) 公開番号	特開2010-15088 (P2010-15088A)		大阪府門真市大字門真1006番地
(43) 公開日	平成22年1月21日(2010.1.21)	(74) 代理人	100105647
審査請求日	平成23年7月5日(2011.7.5)		弁理士 小栗 昌平
		(74) 代理人	100108589
			弁理士 市川 利光
		(74) 代理人	100119552
			弁理士 橋本 公秀
		(72) 発明者	塚本 有一
			東京都中央区京橋2丁目13番10号 B
			rothers & Co. 株式会社内
		(72) 発明者	進藤 功
			東京都中央区京橋2丁目13番10号 B
			rothers & Co. 株式会社内
			最終頁に続く

(54) 【発明の名称】 データ生成装置及びデータ生成プログラム、並びに、再生装置

(57) 【特許請求の範囲】

【請求項1】

音声の再生に合わせてテキストを所定のタイミングで表示するための連動表示用データを生成するデータ生成装置であって、

音源データから特徴パラメータを抽出し、当該抽出した特徴パラメータを所定の言語の特徴パラメータと比較する言語列生成部と、

前記テキストを1つ以上の文字列に区分けして、各文字列に読みを付与するテキスト処理部と、

読みが付与された各文字列と前記音源データから抽出した特徴パラメータと前記所定の言語の特徴パラメータとに基づいて、対象とされている文字列が再生されるタイミングを前記音源データの再生経過時間によって示すタイムスタンプ情報を含む連動表示用データを生成する第1データ処理部と、

読みが付与されなかった文字列の前後の文字列のタイムスタンプ情報に基づいて、前記読みが付与されなかった文字列に所定長の時間帯を割り当てて、当該読みが付与されなかった文字列と、この文字列に割り当てた時間帯を示すタイムスタンプ情報とを前記連動表示用データに含める第2データ処理部と、

を備えたことを特徴とするデータ生成装置。

【請求項2】

請求項1に記載のデータ生成装置であって、

前記テキスト処理部によって読みが付与されなかった文字列と、当該文字列のテキスト

内位置を示す情報とを記憶する位置管理部を備え、

前記第2データ処理部は、前記第1データ処理部が、前記テキストに含まれる読みが付与された文字列の全てと、読みが付与された各文字列のタイムスタンプ情報とを含む連動表示用データを生成した後、前記位置管理部に記録された各文字列に所定長の時間帯を割り当ててことを特徴とするデータ生成装置。

【請求項3】

請求項2に記載のデータ生成装置であって、

前記第2データ処理部は、

前記位置管理部が記憶する対象とされた文字列の直前の文字列のタイムスタンプ情報が示す終了時間から直後の文字列のタイムスタンプ情報が示す開始時間までの時間長を算出し、

10

前記時間長が所定時間以上であれば、前記終了時間から前記開始時間までの時間内の所定長の時間帯を、前記対象とされた文字列に割り当て、

前記時間長が所定時間未満であれば、前記直前の文字列に割り当てられた時間帯又は前記直後の文字列に割り当てられた時間帯の少なくともいずれか一方を短縮し、当該短縮した時間帯を、前記対象とされた文字列に割り当ててことを特徴とするデータ生成装置。

【請求項4】

請求項3に記載のデータ生成装置であって、

前記第2データ処理部は、前記直前の文字列に割り当てられた時間帯又は前記直後の文字列に割り当てられた時間帯のいずれか一方を前記所定長の時間、短縮することを特徴とするデータ生成装置。

20

【請求項5】

請求項2に記載のデータ生成装置であって、

前記第2データ処理部は、読みの付与できない文字列の特徴に応じて、前記位置管理部が記憶する当該文字列の直前の文字列に割り当てられた時間帯を前記所定長の時間、短縮するか、前記直後の文字列に割り当てられた時間帯を前記所定長の時間、短縮するかを判断することを特徴とするデータ生成装置。

【請求項6】

請求項2に記載のデータ生成装置であって、

前記位置管理部に記録される文字列のテキスト内位置を示す情報は、当該文字列の前記テキストの先頭からの文字数を含むことを特徴とするデータ生成装置。

30

【請求項7】

請求項2に記載のデータ生成装置であって、

前記位置管理部に記録される文字列のテキスト内位置を示す情報は、前記テキストを所定の表示形態で表示したときの表示領域における行番号及び列番号を含むことを特徴とするデータ生成装置。

【請求項8】

請求項2に記載のデータ生成装置であって、

前記位置管理部に記録される文字列のテキスト内位置を示す情報は、当該文字列の前記テキストの先頭の文字列からの文字列番号を含むことを特徴とするデータ生成装置。

40

【請求項9】

請求項2に記載のデータ生成装置であって、

前記位置管理部に記録される文字列のテキスト内位置を示す情報は、前記テキストに含まれる前記テキスト処理部で分けられた各文字列同士の順序関係を含むことを特徴とするデータ生成装置。

【請求項10】

音声の再生に合わせてテキストを所定のタイミングで表示するための連動表示用データを生成するデータ生成装置であって、

音源データから特徴パラメータを抽出し、当該抽出した特徴パラメータを所定の言語の特徴パラメータと比較する言語列生成部と、

50

前記テキストを1つ以上の文字列に区分けして、各文字列に読みを付与するテキスト処理部と、

読みが付与された各文字列と前記音源データから抽出した特徴パラメータと前記所定の言語の特徴パラメータとに基づいて、対象とされている文字列が再生されるタイミングを前記音源データの再生経過時間によって示すタイムスタンプ情報を含む連動表示用データを生成するデータ処理部と、

読みが付与されなかった文字列に特別な読みを付与する特別読み付与部と、

前記データ処理部は、前記特別な読みが付与された文字列の直前の文字列のタイムスタンプ情報に基づいて、前記特別な読みが付与された文字列と前記音源データから抽出した特徴パラメータと前記所定の言語の特徴パラメータとに基づいて得られる前記再生される
10
タイミングを、当該特別な読みが付与された文字列に割り当てて、当該特別な読みが付与された文字列と、この文字列に割り当てた時間帯を示すタイムスタンプ情報とを前記連動表示用データに含めることを特徴とするデータ生成装置。

【請求項11】

請求項1又は10に記載のデータ生成装置であって、

読みが付与されなかった文字列に含まれるいずれかの文字に対し、読みを推測する読み推測部を備えることを特徴とするデータ生成装置。

【請求項12】

請求項11に記載のデータ生成装置であって、

前記特別読み付与部は、前記読み推測部が読みを推測できなかった文字列に前記特別な読みを付与することを特徴とするデータ生成装置。
20

【請求項13】

請求項10又は12に記載のデータ生成装置であって、

前記データ処理部は、前記特別な読みが付与された文字列の直前の文字列のタイムスタンプ情報が示す終了時間から所定長の時間帯を、前記特別な読みが付与された文字列に割り当てることを特徴とするデータ生成装置。

【請求項14】

請求項1～13のいずれか1項に記載のデータ生成装置であって、

前記データ処理部は、前記所定長の時間帯が割り当てられた文字列を前後の文字列に含め、当該文字列のタイムスタンプ情報も結合することを特徴とするデータ生成装置。
30

【請求項15】

コンピュータを、請求項1～14のいずれか1項に記載のデータ生成装置が備える各部として機能させるためのデータ生成プログラム。

【請求項16】

テキストを表示させながら音声を再生する再生装置であって、

請求項1～14のいずれか1項に記載のデータ生成装置又は請求項15に記載のデータ生成プログラムにより作成した連動表示用データに基づき、音声の再生に合わせてテキスト内の読みを付与できない文字列を所定のタイミングで表示することを特徴とする再生装置。

【請求項17】
40

テキストを表示させながら音声を再生する再生装置であって、

請求項1～14のいずれか1項に記載のデータ生成装置又は請求項15に記載のデータ生成プログラムにより作成した連動表示用データに基づき、音声の再生に合わせてテキスト内の読みを付与できない文字列を所定のタイミングで表示し、その表示部分を利用してユーザの選択箇所を取得する事により、読みの付与されなかった文字列から再生を開始することを特徴とする再生装置。

【請求項18】

請求項1～14のいずれか1項に記載のデータ生成装置又は請求項15に記載のデータ生成プログラムにより作成した連動表示用データに基づき、音声の再生に合わせてテキスト内の読みを付与できない文字列を所定のタイミングで表示し、その表示部分を利用して
50

ユーザの選択箇所を取得する事により、ユーザの選択した所定単位の文字列の先頭から再生を開始する再生装置であって、

ユーザが読みの付与されなかった文字列を選択した際に、前記読みの付与されなかった文字列の前後の読みの付与された文字列から再生を開始することを特徴とする再生装置。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、音声の再生に合わせてテキストを所定のタイミングで表示するための連動表示用データを生成するデータ生成装置及びデータ生成プログラム、並びに、再生装置に関する。

10

【背景技術】

【0002】

音楽に合わせて画面に表示される歌詞を見ながら歌を歌うためのカラオケ機器が広く普及している。多くのカラオケ機器は、音楽の再生に合わせて歌詞を画面に表示し、歌われる部分の歌詞の表示色を適当なタイミングで滑らかに変えていく。例えば、グレーから白に歌詞の表示色を変えていく。このようなカラオケ機器は、音楽データ及び歌詞データの他に、音楽の再生に合わせて歌詞の表示色を変えるための連動表示用データを用いる。特許文献1及び2は、連動表示用データを作成するための技術について説明する。

【0003】

特許文献1には、音声コンテンツと歌詞等のテキストデータとの間の同期タイミングを設定するリンクテーブルを作成するファイル作成装置が開示されている。当該ファイル作成装置はリンク作成部を備え、リンク作成部は、図18に示すように、周波数弁別部201と、特徴抽出部202と、言語作成部203と、言語モデルデータベース204と、言語モデルメモリ部205と、フレームカウント部206と、経過時間算出部207と、テキスト分割部208と、テキストメモリ部209と、テーブル作成部211とを有する。

20

【0004】

周波数弁別部201は、フレームデータとして供給された音楽データの周波数スペクトルを解析し、人の音声帯域のスペクトル成分を抽出する。特徴抽出部202は、抽出されたスペクトル成分を音響分析して特徴パラメータを抽出する。言語作成部203は、特徴抽出部202で抽出された特徴パラメータと、音声モデルデータベース204に格納された各基準言語の特徴パラメータとを比較して言語列を生成する。音声モデルデータベース204は、五十音や濁音などの各言語モデルの特徴パラメータを言語モデル毎に記憶している。言語モデルメモリ部205は、言語作成部203で生成された言語列をテキストデータに変換して記憶する。

30

【0005】

フレームカウント部206は、供給されたフレーム数を供給開始から累積し、その累積フレーム数を経過時間算出部207に供給する。経過時間算出部207は、累積フレーム数に基づいて、再生開始時からの経過時間を算出する。テキスト分割部208は、テキストデータを複数のブロックに区分けする。テキストメモリ部209は、ブロック毎に区分けされたテキストデータを記憶する。

40

【0006】

マッチング部210は、テキストメモリ部209に記憶されたブロック毎のテキストデータと、言語モデルメモリ部205に記憶された言語列のテキストデータとを比較し、両者が整合したタイミングで識別信号をテーブル作成部211に送る。リンク作成部211は、マッチング部210から識別信号が供給されたタイミングで経過時間算出部207から経過時間データを取得し、この経過時間データを当該テキストブロックの再生タイミングとしてリンクテーブルに設定する。

【0007】

また、特許文献2には、図19に示すように、漢字仮名混じりの文字列で構成された歌詞情報を変換辞書部16を用いて読みに変換し、当該変換した読みを音符等の楽譜情報に

50

割り当てる歌詞割当装置が開示されている。変換辞書部 16 は、所定の単語とその読み情報を対応して記憶する。

【0008】

【特許文献1】特開2003-280670号公報

【特許文献2】特開2002-82665号公報

【発明の開示】

【発明が解決しようとする課題】

【0009】

特許文献1のファイル作成装置は、上述のように、人の音声帯域のスペクトル成分を音響分析して特徴パラメータを抽出し、この特徴パラメータを音声モデルデータベース204に格納された各基準言語の特徴パラメータと比較して言語列を生成している。このため、言語列には句読点やクエスチョンマーク等の無音記号が含まれない。その結果、テキストメモリ部209に記憶されたテキストデータに無音記号が含まれていても、無音記号は言語作成部203によって生成された言語列とは整合しないため、当該無音記号を含む区分けされたテキストブロックには適した同期タイミングが設定されない可能性がある。また、特許文献1のファイル作成装置では、テキストブロックの先頭文字に対応する言語情報が抽出言語列の言語情報にマッチングした時点を比較開始時点とする事により、前記無音記号が含まれる場合の不一致をテキストブロック単位で抑制・回避しているが、テキストブロックが無音記号だけで構成されている場合、又は著しく多くの無音記号で構成されている場合には、適した同期タイミングが設定されない可能性がある。

【0010】

また、特許文献2の歌詞割当装置は、上述のように、変換辞書部16を用いて歌詞情報を読みに変換する。しかし、当該歌詞割当装置は、読みを持たない句読点やクエスチョンマーク等の無音記号は変換辞書部16には登録されていないため、変換できない。このため、無音記号を楽譜情報に割り当てることができない。

【0011】

また、当該歌詞割当装置は、変換辞書部16に記録されていない固有名詞等の文字列を正確な読みに変換できない可能性がある。本来とは異なる読みを楽譜情報に割り当てると、適切なタイミングで歌詞が表示されない可能性があるため好ましくない。なお、ユーザは、変換辞書の内容を追加及び削除できる。このため、上述の固有名詞等の文字列とその読みを新たに追加できるが、この作業は、変換辞書の内容を変更するための操作方法を知るユーザが時間と手間をかけて行わなければならない。

【0012】

本発明の目的は、読みが付与されない文字列又は正確な読みが付与されなかった文字列を含むテキストを音声の再生に合わせた所定のタイミングで表示するための連動表示用データを生成するデータ生成装置及びデータ生成プログラム、並びに、再生装置を提供することである。

【課題を解決するための手段】

【0013】

本発明は、音声の再生に合わせてテキストを所定のタイミングで表示するための連動表示用データを生成するデータ生成装置であって、音源データから特徴パラメータを抽出し、当該抽出した特徴パラメータを所定の言語の特徴パラメータと比較する言語列生成部と、前記テキストを1つ以上の文字列に区分けして、各文字列に読みを付与するテキスト処理部と、読みが付与された各文字列と前記音源データから抽出した特徴パラメータと前記所定の言語の特徴パラメータとに基づいて、対象とされている文字列が再生されるタイミングを前記音源データの再生経過時間によって示すタイムスタンプ情報を含む連動表示用データを生成する第1データ処理部と、読みが付与されなかった文字列の前後の文字列のタイムスタンプ情報に基づいて、前記読みが付与されなかった文字列に所定長の時間帯を割り当てて、当該読みが付与されなかった文字列と、この文字列に割り当てた時間帯を示すタイムスタンプ情報とを前記連動表示用データに含める第2データ処理部と、を備えた

10

20

30

40

50

データ生成装置を提供する。

【0014】

上記データ生成装置は、前記テキスト処理部によって読みが付与されなかった文字列と、当該文字列のテキスト内位置を示す情報とを記憶する位置管理部を備え、前記第2データ処理部は、前記第1データ処理部が、前記テキストに含まれる読みが付与された文字列の全てと、読みが付与された各文字列のタイムスタンプ情報とを含む連動表示用データを生成した後、前記位置管理部に記録された各文字列に所定長の時間帯を割り当てる。

【0015】

上記データ生成装置では、前記第2データ処理部は、前記位置管理部が記憶する対象とされた文字列の直前の文字列のタイムスタンプ情報が示す終了時間から直後の文字列のタイムスタンプ情報が示す開始時間までの時間長を算出し、前記時間長が所定時間以上であれば、前記終了時間から前記開始時間までの時間内の所定長の時間帯を、前記対象とされた文字列に割り当て、前記時間長が所定時間未満であれば、前記直前の文字列に割り当てられた時間帯又は前記直後の文字列に割り当てられた時間帯の少なくともいずれか一方を短縮し、当該短縮した時間帯を、前記対象とされた文字列に割り当てる。

10

【0016】

上記データ生成装置では、前記第2データ処理部は、前記直前の文字列に割り当てられた時間帯又は前記直後の文字列に割り当てられた時間帯のいずれか一方を前記所定長の時間、短縮する。

【0017】

上記データ生成装置では、前記第2データ処理部は、前記終了時間から前記開始時間までの時間帯及び前記短縮した時間帯を、前記読みが付与されなかった文字列に割り当てる。

20

【0018】

上記データ生成装置では、前記位置管理部に記録される文字列のテキスト内位置を示す情報は、当該文字列の前記テキストの先頭からの文字数を含む。

【0019】

上記データ生成装置では、前記位置管理部に記録される文字列のテキスト内位置を示す情報は、前記テキストを所定の表示形態で表示したときの表示領域における行番号及び列番号を含む。

30

【0020】

上記データ生成装置では、前記位置管理部に記録される文字列のテキスト内位置を示す情報は、当該文字列の前記テキストの先頭の文字列からの文字列番号を含む。

【0021】

上記データ生成装置では、前記位置管理部に記録される文字列のテキスト内位置を示す情報は、前記テキストに含まれる前記テキスト処理部で分けられた各文字列同士の順序関係を含む。

【0022】

本発明は、音声の再生に合わせてテキストを所定のタイミングで表示するための連動表示用データを生成するデータ生成装置であって、音源データから特徴パラメータを抽出し、当該抽出した特徴パラメータを所定の言語の特徴パラメータと比較する言語列生成部と、前記テキストを1つ以上の文字列に区分けして、各文字列に読みを付与するテキスト処理部と、読みが付与された各文字列と前記音源データから抽出した特徴パラメータと前記所定の言語の特徴パラメータとに基づいて、対象とされている文字列が再生されるタイミングを前記音源データの再生経過時間によって示すタイムスタンプ情報を含む連動表示用データを生成するデータ処理部と、読みが付与されなかった文字列に特別な読みを付与する特別読み付与部と、前記データ処理部は、前記特別な読みが付与された文字列の直前の文字列のタイムスタンプ情報に基づいて、前記特別な読みが付与された文字列と前記音源データから抽出した特徴パラメータと前記所定の言語の特徴パラメータとに基づいて得られる前記再生されるタイミングを、当該特別な読みが付与された文字列に割り当てて、当

40

50

該特別な読みが付与された文字列と、この文字列に割り当てた時間帯を示すタイムスタンプ情報とを前記連動表示用データに含めるデータ生成装置を提供する。

【0023】

上記データ生成装置では、読みが付与されなかった文字列に含まれるいずれかの文字に対し、読みを推測する読み推測部を備える。

また、上記データ生成装置では、前記特別読み付与部は、前記読み推測部が読みを推測できなかった文字列に前記特別な読みを付与する。

【0024】

上記データ生成装置では、前記データ処理部は、前記特別な読みが付与された文字列の直前の文字列のタイムスタンプ情報が示す終了時間から所定長の時間帯を、前記特別な読みが付与された文字列に割り当てる。

10

【0025】

本発明は、コンピュータを、上記データ生成装置が備える各部として機能させるためのデータ生成プログラムを提供する。

【0026】

本発明は、上記データ生成装置又は上記データ生成プログラムで作成した連動表示用データに基づき、音声の再生に合わせてテキスト内の読みを付与できない文字列を所定のタイミングで表示する再生装置を提供する。

【0027】

本発明は、上記データ生成装置又は上記データ生成プログラムで作成した連動表示用データに基づき、音声の再生に合わせてテキスト内の読みを付与できない文字列を所定のタイミングで表示し、その表示部分を利用してユーザの選択箇所を取得する事により、読みの付与できない文字列から再生を開始する再生装置を提供する。

20

【発明の効果】

【0028】

本発明に係るデータ生成装置及びデータ生成プログラム、並びに、再生装置によれば、読みが付与されない文字列又は正確な読みが付与されなかった文字列を含むテキストを音声の再生に合わせた所定のタイミングで表示するための連動表示用データを生成することができる。

【発明を実施するための最良の形態】

30

【0029】

以下、本発明の実施形態について、図面を参照して説明する。以下説明する実施形態のデータ生成装置は、音源データ及びテキストデータに基づいて、音源データの再生に合わせた所定のタイミングでテキストを表示するための連動表示用データを生成する。なお、テキストは、句読点や特殊文字を含む日本語又は外国語の文章、歌詞等である。また、テキストを表示するとは、例えば、テキストの表示色の透過率を50%から0%に変更したり、テキストの表示色を別の見やすい色に変更することや、テキストが全く表示されていない状態から表示することである。

【0030】

データ生成装置が連動表示用データを生成する際、テキストは1文字以上の文字で構成される文字列に分けられる。この文字列は、形態素（言語で意味を持つ最小単位）でも文字1つでも音素（言語の持つ音の最小単位）でも構わない。生成された連動表示用データは、テキストデータと、テキストを細分化した各文字列に割り当てられた固有の時間情報であるタイムスタンプ情報とを含む。タイムスタンプ情報は、各文字列を表示する開始時間と、その文字列を表示し終えるまでの時間長又は終了時間とを含む。これら開始時間及び終了時間は、音声の所定の再生経過時間によって示される。

40

【0031】

（第1の実施形態）

図1は、第1の実施形態のデータ生成装置を示すブロック図である。図1に示すように、第1の実施形態のデータ生成装置100は、入力部101と、データ生成部103と、

50

位置管理部 105 と、データ加工部 107 と、データ記憶部 109 とを備える。なお、データ生成部 103、位置管理部 105 及びデータ加工部 107 はプログラムを実行することによって動作する。

【0032】

入力部 101 は、音源データ及びテキストデータをデータ生成装置 100 に入力するためのインタフェースである。データ生成部 103 は、入力部 101 を介して入力された音源データとテキストデータのマッチングを行い、連動表示用データを生成する。なお、データ生成部 103 による音源データとテキストデータのマッチングの詳細については後述する。位置管理部 105 は、データ生成部 103 における処理の過程で読みが付与されなかった文字列や音源データに基づくテキストとマッチングしない文字列とそのテキスト内位置を示す情報を記憶する。なお、音源データに基づくテキストとマッチングしない文字列は、データ生成部 103 による処理の過程で、当該文字列に正確な読みが付与されなかった結果、音源データに基づくテキストとマッチングしないことが考えられる。

10

【0033】

データ加工部 107 は、読みが付与されなかった文字列や音源データに基づくテキストとマッチングしない文字列にタイムスタンプ情報を割り当てて、データ生成部 103 によって生成された連動表示用データを更新する。データ記憶部 109 は、データ生成部 103 によって生成され、データ加工部 107 によって更新された連動表示用データを記憶する。

【0034】

20

なお、位置管理部 105 に記録されるテキスト内位置を示す情報は、複数の形態の中のいずれか 1 つの形態によって示される。複数の形態とは、(1) テキストデータが示すテキストの先頭からの文字数、(2) テキストデータが示すテキストを所定の表示形態で表示したときの表示領域における行番号及び列番号、(3) テキストデータが示すテキストを形態素解析処理等を行うことによって複数の文字列に区分けした際の先頭の文字列からの文字列番号、(4) テキストデータが示すテキストを形態素解析処理等を行うことによって複数の文字列に区分けした際の区分けされた各文字列同士の順序関係(例えば音声対話記述言語である VoiceXML 等のデータ構造が理解しやすい形式で記述されるのが好ましい。)に基づく記述情報である。

【0035】

30

図 2 は、第 1 の実施形態のデータ生成装置 100 の動作を示すフローチャートである。図 2 に示すように、ステップ S101 では、入力部 101 を介して入力された音源データ及びテキストデータがデータ生成部 103 に送られる。次に、ステップ S103 では、データ生成部 103 が音源データとテキストデータのマッチングを行い、連動表示用データを生成する。次に、ステップ S105 では、データ加工部 107 が、読みを付与できなかった文字列にタイムスタンプ情報を割り当てて、ステップ S103 で生成された連動表示用データを更新する。次に、ステップ S107 では、データ加工部 107 が、ステップ S105 で更新した連動表示用データをデータ記憶部 109 に格納する。

【0036】

図 3 及び図 4 は、第 1 の実施形態のデータ生成装置 100 が行うステップ S103 の詳細を示すフローチャートである。図 3 に示すように、データ生成部 103 は、図 2 のステップ S101 で入力された音源データの周波数スペクトルを解析し、人の音声帯域のスペクトル成分を抽出する(ステップ S201)。次に、データ生成部 103 は、ステップ S201 で抽出したスペクトル成分を音響分析して特徴パラメータを抽出する(ステップ S203)。次に、データ生成部 103 は、ステップ S203 で抽出した特徴パラメータを所定の言語の特徴パラメータと比較して言語列を生成する(ステップ S205)。次に、データ生成部 103 は、ステップ S205 で生成した言語列をテキストデータに変換する(ステップ S207)。なお、ステップ S205 で用いられる所定の言語の特徴パラメータが日本語の特徴パラメータの場合、ステップ S207 で得られる言語列のテキストデータはひらがな又はカタカナである。

40

50

【 0 0 3 7 】

次に、データ生成部 1 0 3 は、図 2 のステップ S 1 0 1 で入力されたテキストデータが示すテキストを、形態素解析等の構造解析処理を行うことにより複数の文字列に区分けする（ステップ S 2 0 9）。図 2 のステップ S 1 0 1 で入力されたテキストデータが日本語の場合、当該テキストには漢字や句読点等が含まれている可能性が高い。一方、ステップ S 2 0 7 で得られた言語列のテキストデータはひらがな又はカタカナである。このため、図 4 に示すように、データ生成部 1 0 3 は、図示しない読み変換辞書を用いて、ステップ S 2 0 9 で区分けされた各文字列に読みを付与する（ステップ S 2 1 1）。

【 0 0 3 8 】

次に、データ生成部 1 0 3 は、ステップ S 2 1 1 で読みが付与されたか否かを文字列毎に判断する（ステップ S 2 1 3）。データ生成部 1 0 3 は、ステップ S 2 1 3 で読みが付与されなかったと判断された文字列とそのテキスト内位置を示す情報を位置管理部 1 0 5 に登録する（ステップ S 2 1 5）。一方、データ生成部 1 0 3 は、ステップ S 2 1 3 で読みが付与されたと判断された文字列をステップ S 2 0 7 で得られた言語列のテキストデータと比較して（ステップ S 2 1 7）、言語列のテキストデータが示すテキスト内に対象とされている文字列と読み的一致する文字列があるか否かを判断する（ステップ S 2 1 9）。

【 0 0 3 9 】

ステップ S 2 1 9 で、対象とされている文字列が言語列のテキスト内であればステップ S 2 2 1 に進み、言語列のテキスト内になればステップ S 2 1 5 に進み、当該対象とされている文字列とそのテキスト内位置を示す情報を位置管理部 1 0 5 に記録する。ステップ S 2 2 1 では、データ生成部 1 0 3 が、対象とされている文字列及びこの文字列に対応するタイムスタンプ情報を連動表示用データに書き出す。なお、各文字列のタイムスタンプ情報は、音源データを再生した際に、対象とされている文字列が発声されるタイミングに応じて決定される。

【 0 0 4 0 】

ステップ S 2 2 1 を行った後、データ生成部 1 0 3 は、ステップ S 2 1 7 で文字列と比較する言語列のテキストデータの解析対象位置を時間情報に基づき更新する（ステップ S 2 2 3）。データ生成部 1 0 3 は、ステップ S 2 0 9 で区分けされた文字列の全てに対して、ステップ S 2 1 5 における位置管理部 1 0 5 への記録又はステップ S 2 2 1 における連動表示用データへの書き出しが行われたかを判断し（ステップ S 2 2 5）、全ての文字列に対して当該処理が行われた場合はこのサブファンクションを終了し、そうでなければステップ S 2 1 1 に戻る。

【 0 0 4 1 】

図 5 は、第 1 の実施形態のデータ生成装置 1 0 0 が行うステップ S 1 0 5 の詳細を示すフローチャートである。図 5 に示すように、データ加工部 1 0 7 は、位置管理部 1 0 5 から文字列とそのテキスト内位置を示す情報を読み込む（ステップ S 3 0 1）。次に、データ加工部 1 0 7 は、ステップ S 1 0 3 で生成した連動表示用データに基づいて、ステップ S 3 0 1 で読み込んだ文字列の前後の文字列を特定し、前の文字列の終了時間から後の文字列の開始時間までの時間長を計算する（ステップ S 3 0 3）。

【 0 0 4 2 】

次に、データ加工部 1 0 7 は、ステップ S 3 0 3 で算出した時間長が例えば 1 0 m 秒といった所定時間以上であるかを判断し（ステップ S 3 0 5）、時間長が所定時間以上であればステップ S 3 0 7 に進み、所定時間未満であればステップ S 3 0 9 に進む。ステップ S 3 0 7 で、データ加工部 1 0 7 は、ステップ S 3 0 1 で読み込んだ文字列に、前の文字列の終了時間から後の文字列の開始時間までの時間内の所定長の時間帯を割り当てて、連動表示用データを更新する。すなわち、ステップ S 3 0 7 では、ステップ S 3 0 1 で読み込んだ文字列及びこの文字列に割り当てた時間帯を示すタイムスタンプ情報を連動表示用データに書き出す。

【 0 0 4 3 】

一方、ステップ S 3 0 9 で、データ加工部 1 0 7 は、ステップ S 3 0 1 で読み込んだ文

10

20

30

40

50

字列に所定長の時間帯を割り当てるよう、当該文字列の前後の文字列に割り当てられた2つの時間帯の少なくともいずれか一方を短縮し、その短縮した時間帯を当該文字列に割り当てた上で、連動表示用データを更新する。データ加工部107は、当該文字列の前の文字列に割り当てられた時間帯を短縮する場合、当該文字列の前の文字列の時間帯の終了時間を所定時間長早め、当該設定変更によって空いた時間帯を当該文字列に割り当てる。また、データ加工部107は、当該文字列の後の文字列に割り当てられた時間帯を短縮する場合、当該文字列の後の文字列の時間帯の開始時間を所定時間長遅らせ、当該設定変更によって空いた時間帯を当該文字列に割り当てる。

尚、当該文字列の前後の文字列に割り当てられた時間帯を短縮する場合、当該文字列の特徴（開始文字、終了文字等）に応じて前後のどちらの文字列に割り当てられた時間帯を短縮するか判断してもよい。

10

【0044】

なお、データ加工部107は、当該文字列の前の文字列に割り当てられた時間帯の終了時間を早め、かつ、当該文字列の後の文字列に割り当てられた時間帯の開始時間を遅らせることによって、当該設定変更によって空いた時間帯を当該文字列に割り当てても良い。さらに、当該文字列の前後の文字列から短縮する時間長は、ステップS303で算出した時間長に応じて調整しても良い。すなわち、ステップS303で算出した時間長と、当該文字列の前後の文字列から短縮して得られた時間長との和が所定長となるよう、当該文字列の前後の文字列の時間帯を短縮しても良い。

【0045】

20

ステップS307又はステップS309が行われた後、データ加工部107は、位置管理部105に登録された全ての文字列に対して上記処理が行われたかを判断し（ステップS311）、全ての文字列に対して当該処理が行われた場合はこのサブファンクションを終了し、そうでなければステップS301に戻る。

【0046】

本実施形態のデータ生成装置100には、図6に示すように、ユーザからの指示に応じて音源データの再生や連動表示用データに基づくテキストの表示等を行う再生装置50を接続しても良い。再生装置50には、操作受付部51、再生制御部53、再生部55及び表示部57が設けられている。操作受付部51はユーザによる再生装置50の操作を受け付け、再生制御部53は操作内容に応じた処理を行う。当該処理によって、再生部55は音源データを再生し、表示部57はテキストを表示する。このとき、再生制御部53は、データ生成装置100のデータ記憶部109に格納されている連動表示用データに基づいて、音源データの再生に合わせて読みが付与できないまたは読みが正確でない文字列を含むテキストを所定のタイミングで表示するよう処理する。図7は、表示部57に表示されるテキストの一例を示す図である。

30

【0047】

また、操作受付部51の一形態として、表示部57上にタッチパネルが設けられている場合、ユーザは表示部57に表示された読みが付与できないまたは読みが正確でない文字列を含むテキスト中の所望の文字列を選択することができる。再生制御部53は、当該操作に応じて、選択された文字列に対応する再生時間から音源データを再生し、かつ、テキストを表示するよう処理する。尚、選択対象となる文字列の単位は形態素で区切られた各文字列でも、表示部57における表示上の行でも問題ない。また、選択対象となる各文字列に読みが付与されない文字列のみで構成されている文字列がある場合、前後の読みが付与されている文字列に含める事で、選択対象となる文字列には必ず読みを持つ文字列が含まれるようにしても良い。

40

【0048】

以上説明したように、本実施形態のデータ生成装置100によれば、読みが付与されない文字列又は正確な読みが付与されなかった文字列に、所定長の時間帯が割り当てられる。したがって、本実施形態のデータ生成装置100によって生成及び更新された連動表示用データに基づくテキストの表示では、句読点やクエスチョンマーク等の無音記号で表さ

50

れる文字列、又は固有名詞や略字等の正確な読みが付与されにくい文字列を含むテキストが、音源の再生に合わせて実際に近いタイミングで表示される。さらに、ユーザは、このような文字列を再生開始点として指定することができる。

【 0 0 4 9 】

(第2の実施形態)

図8は、第2の実施形態のデータ生成装置を示すブロック図である。図8に示すように、第2の実施形態のデータ生成装置110は、入力部111と、データ生成部113と、特別読み付与部115と、データ記憶部117とを備える。なお、データ生成部113及び特別読み付与部115はプログラムを実行することによって動作する。

【 0 0 5 0 】

入力部111は、音源データ及びテキストデータをデータ生成装置110に入力するためのインタフェースである。データ生成部113は、入力部111を介して入力された音源データとテキストデータのマッチングを第1の実施形態と同様に行い、連動表示用データを生成する。特別読み付与部115は、データ生成部113における処理の過程で読みが付与されなかった文字列に特別な読みを付与する。なお、「特別な読み」とは、全ての前記言語列に相当する「*(Linux(登録商標)等のオペレーティングシステムにおけるワイルドカードに相当する)」等である。データ記憶部117は、データ生成部113によって生成された連動表示用データを記憶する。

【 0 0 5 1 】

図9は、第2の実施形態のデータ生成装置110の動作を示すフローチャートである。図9に示すように、ステップS111では、入力部111を介して入力された音源データ及びテキストデータがデータ生成部113に送られる。次に、ステップS113では、データ生成部113が音源データとテキストデータのマッチングを行い、連動表示用データを生成する。次に、ステップS115では、データ生成部113が、ステップS113で生成した連動表示用データをデータ記憶部117に格納する。

【 0 0 5 2 】

図10及び図11は、第2の実施形態のデータ生成装置110が行うステップS113の詳細を示すフローチャートである。図10及び図11に示すように、データ生成部113は、図3及び図4を参照して説明したステップS201～S213及びステップS217～S225を第1の実施形態と同様に行う。本実施形態では、図11に示すステップS213が行われた後、特別読み付与部115は、読みが付与されなかった文字列に特別な読みを付与する(ステップS251)。次に、データ生成部113は、特別な読みが付与された文字列に、当該文字列の直前の文字列の終了時間以降の該当する言語列と特別な読みを比較して(ステップS253)、読みの一致する言語列が存在するか否かを判断する(ステップS255)。その後、該当する言語列と特別な読みが一致した場合はステップS257へ進み、一致しなかった場合はステップS259へ進む。ステップS257では、特別な読みが付与された文字列へ該当する言語列の時間情報を割り当てて、ステップS221へ進む。一方、ステップS259では、特別な読みが付与された文字列の直前の文字列の終了時間から始まる所定長の時間帯を割り当てた後、ステップS221に進む。また、本実施形態では、ステップS219で、対象とされている文字列が言語列のテキスト内ないと判断されたときにもステップS253に進み、データ生成部113は同様の処理を行う。

【 0 0 5 3 】

以上説明したように、本実施形態のデータ生成装置110によれば、読みが付与されない文字列には特別な読みが付与されるため、連動表示用データの更新を行わずに、句読点やクエスチョンマーク等の無音記号で表される文字列を含むテキストの連動表示用データを生成することができる。なお、本実施形態のデータ生成装置110にも、第1の実施形態と同様に、再生装置50を接続しても良い。

【 0 0 5 4 】

(第3の実施形態)

10

20

30

40

50

図12は、第3の実施形態のデータ生成装置を示すブロック図である。図12に示すように、第3の実施形態のデータ生成装置120は、入力部121と、データ生成部123と、読み推測部125と、特別読み付与部127と、データ記憶部129とを備える。なお、データ生成部123及び読み推測部125はプログラムを実行することによって動作する。

【0055】

入力部121は、音源データ及びテキストデータをデータ生成装置120に入力するためのインタフェースである。データ生成部123は、入力部121を介して入力された音源データとテキストデータのマッチングを第1の実施形態と同様に行い、連動表示用データを生成する。読み推測部125は、データ生成部123における処理の過程で読みが付与されなかった文字列の読みを推測する。特別読み付与部127は、読み推測部125で読みを推測できなかった文字列に特別な読みを付与する。なお、「特別な読み」とは、第2の実施形態と同様、全ての前記言語列に相当する「*(Linux(登録商標)等のオペレーティングシステムにおけるワイルドカードに相当する)」等である。データ記憶部129は、データ生成部123によって生成された連動表示用データを記憶する。

10

【0056】

図13は、第3の実施形態のデータ生成装置120の動作を示すフローチャートである。図13に示すように、ステップS121では、入力部121を介して入力された音源データ及びテキストデータがデータ生成部123に送られる。次に、ステップS123では、データ生成部123が音源データとテキストデータのマッチングを行い、連動表示用データを生成する。次に、ステップS125では、データ生成部123が、ステップS123で生成した連動表示用データをデータ記憶部129に格納する。

20

【0057】

図14及び図15は、第3の実施形態のデータ生成装置120が行うステップS123の詳細を示すフローチャートである。図14及び図15に示すように、データ生成部123は、図3及び図4を参照して説明したステップS201～S213及びステップS217～S225を第1の実施形態と同様に行う。本実施形態では、図15に示すステップS213が行われた後、読み推測部125は、読みが付与されなかった文字列の読みを推測する(ステップS271)。ステップS273では、ステップS271で読み推測部125が読みを推測できたか否かを判断し、読みを推測できた場合はステップS217に進み、推測できなかった場合はステップS275に進む。

30

【0058】

ステップS275では、特別読み付与部127は、読み推測部125が読みを推測できなかった文字列に特別な読みを付与する。次に、データ生成部123は、特別な読みが付与された文字列に、当該文字列の直前の文字列の終了時間から始まる所定長の時間帯を割り当て(ステップS253)た後、ステップS221に進む。また、本実施形態では、第2の実施形態と同様、ステップS219で、対象とされている文字列が言語列のテキスト内ないと判断されたときにもステップS253に進み、データ生成部123は同様の処理を行う。

【0059】

図16は、第3の実施形態で読み推測部125が行うステップS271の詳細を示すフローチャートである。図16に示すように、読み推測部125は、読みが付与されなかった文字列に含まれる文字毎に読みとして登録されている音素が含まれているかを解析する(ステップS401)。なお、読み推測部125は、図17に一例が示される音素リストを用いて前記解析を行う。次に、読み推測部125は、文字列に含まれる文字毎に、各文字に対応する音素が音素リスト中に含まれるかを判断し(ステップS403)、含まれる場合はステップS405に進み、含まれない場合はステップS407に進む。

40

【0060】

ステップS405では、読み推測部125は、対象とされた文字に対応する音素を読みとして設定する。一方、ステップS407では、読み推測部125は、対象とされた文字

50

に読みを設定しない。ステップS 4 0 5又はステップS 4 0 7が行われた後、読み推測部1 2 5は、対象とされている文字列に含まれる全ての文字に対して上記処理を行ったかを判断し(ステップS 4 0 9)、全ての文字に対して当該処理が行われた場合はこのサブアクションを終了し、そうでなければステップS 4 0 1に戻る。

【0 0 6 1】

以上説明したように、本実施形態のデータ生成装置1 2 0によれば、読みが付与されない文字列に対しては読みを推測した上で言語列のテキストデータと比較するため、新語や固有名詞等の文字列を含むテキストの連動表示用データをより適切に生成することができる。また、読みを推測できない文字列には特別な読みが付与されるため、第2の実施形態と同様、句読点やクエスチョンマーク等の無音記号で表される文字列を含むテキストの連動表示用データを生成することができる。なお、本実施形態のデータ生成装置1 2 0にも、第1の実施形態と同様に、再生装置5 0を接続しても良い。

【産業上の利用可能性】

【0 0 6 2】

本発明に係るデータ生成装置は、読みが付与されない文字列又は正確な読みが付与されなかった文字列を含むテキストを音声の再生に合わせた所定のタイミングで表示するための連動表示用データを生成する装置等として有用である。

【図面の簡単な説明】

【0 0 6 3】

【図1】第1の実施形態のデータ生成装置を示すブロック図

【図2】第1の実施形態のデータ生成装置の動作を示すフローチャート

【図3】第1の実施形態のデータ生成装置が行うステップS 1 0 3の詳細を示すフローチャート

【図4】第1の実施形態のデータ生成装置が行うステップS 1 0 3の詳細を示すフローチャート

【図5】第1の実施形態のデータ生成装置が行うステップS 1 0 5の詳細を示すフローチャート

【図6】第1の実施形態のデータ生成装置に再生装置が接続された構成を示すブロック図

【図7】再生装置の表示部に表示されるテキストの一例を示す図

【図8】第2の実施形態のデータ生成装置を示すブロック図

【図9】第2の実施形態のデータ生成装置の動作を示すフローチャート

【図10】第2の実施形態のデータ生成装置が行うステップS 1 1 3の詳細を示すフローチャート

【図11】第2の実施形態のデータ生成装置が行うステップS 1 1 3の詳細を示すフローチャート

【図12】第3の実施形態のデータ生成装置を示すブロック図

【図13】第3の実施形態のデータ生成装置の動作を示すフローチャート

【図14】第3の実施形態のデータ生成装置が行うステップS 1 2 3の詳細を示すフローチャート

【図15】第3の実施形態のデータ生成装置が行うステップS 1 2 3の詳細を示すフローチャート

【図16】第3の実施形態で読み推測部が行うステップS 2 7 1の詳細を示すフローチャート

【図17】音素リストの一例を示す図

【図18】特許文献1に開示されたファイル作成装置が備えるリンク作成部の内部構成を示すブロック図

【図19】特許文献2に開示された歌詞割当装置の内部構成を示すブロック図

【符号の説明】

【0 0 6 4】

1 0 0 , 1 1 0 , 1 2 0 データ生成装置

10

20

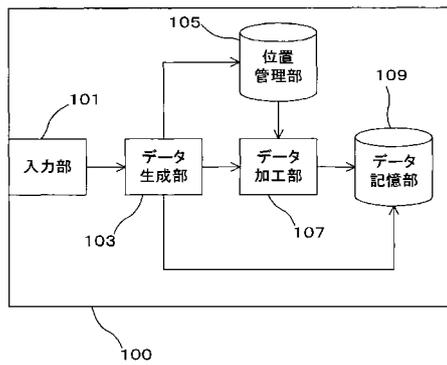
30

40

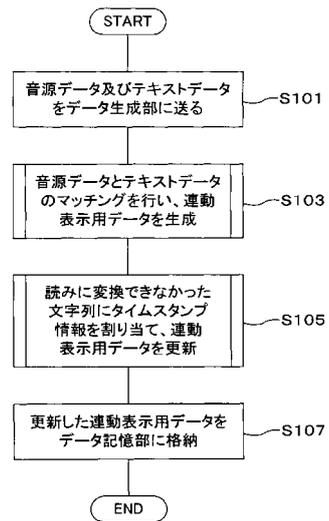
50

- 1 0 1 , 1 1 1 , 1 2 1 入力部
- 1 0 3 , 1 1 3 , 1 2 3 データ生成部
- 1 0 5 位置管理部
- 1 0 7 データ加工部
- 1 0 9 , 1 1 7 , 1 2 9 データ記憶部
- 1 1 5 , 1 2 7 特別読み付与部
- 1 2 5 読み推測部
- 5 0 再生装置
- 5 1 操作受付部
- 5 3 再生制御部
- 5 5 再生部
- 5 7 表示部

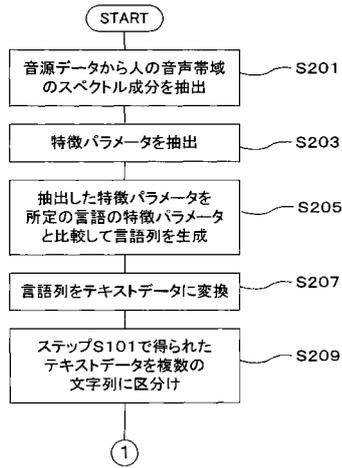
【図1】



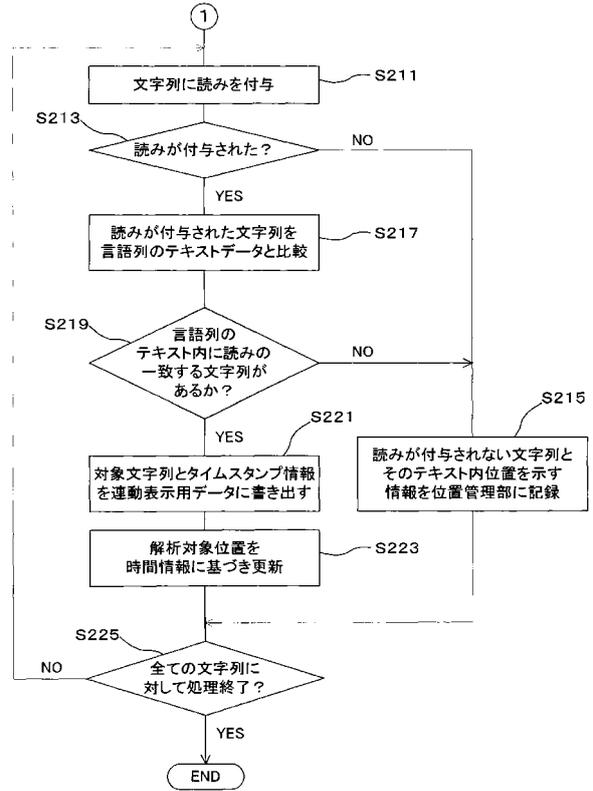
【図2】



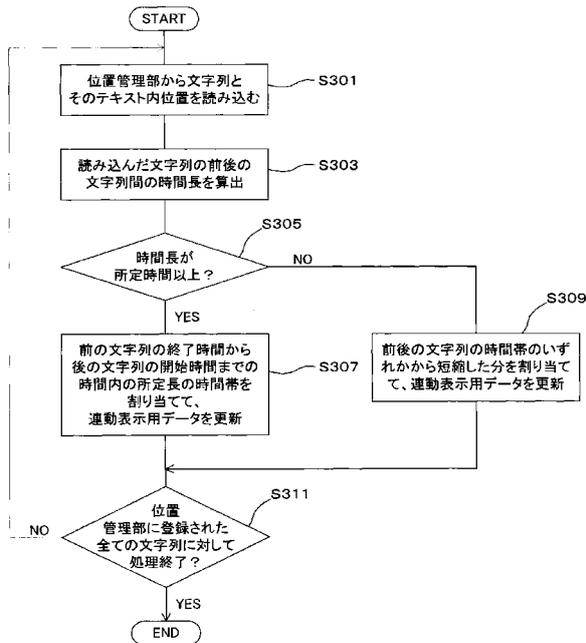
【図3】



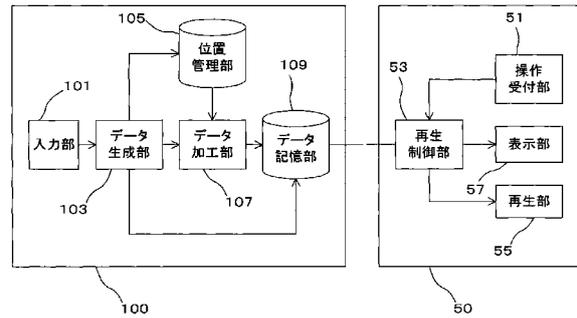
【図4】



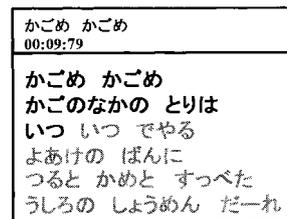
【図5】



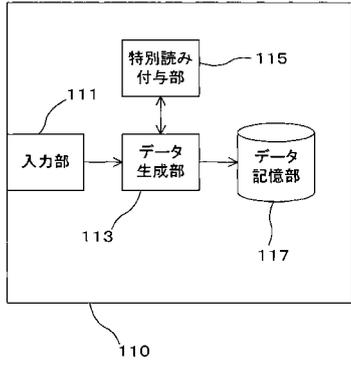
【図6】



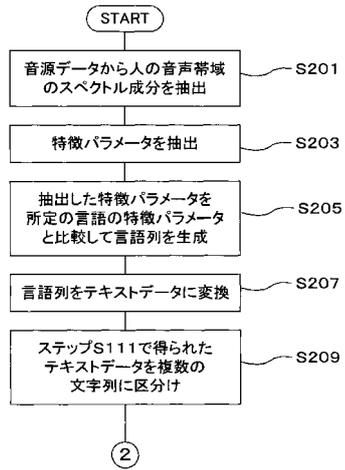
【図7】



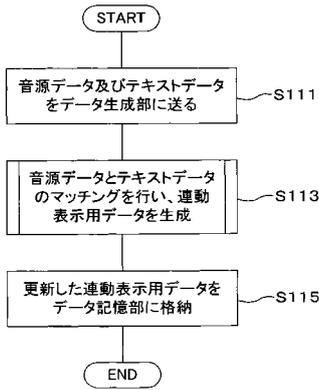
【図 8】



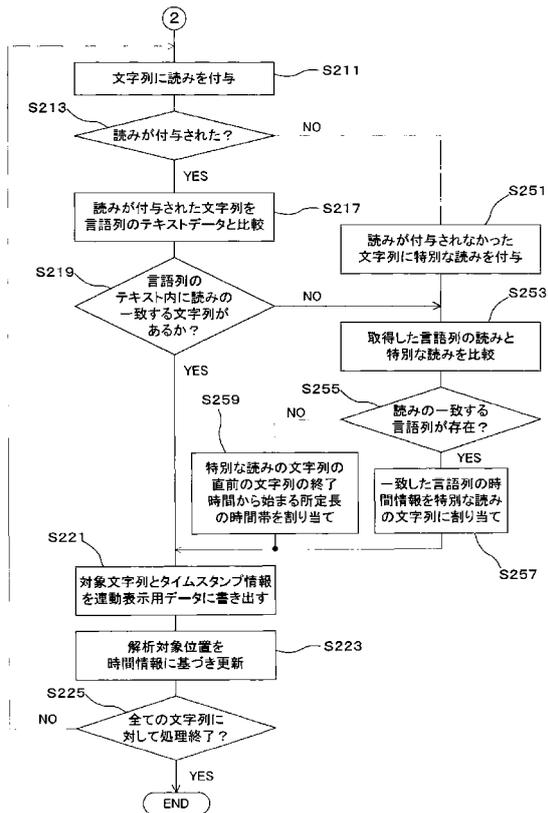
【図 10】



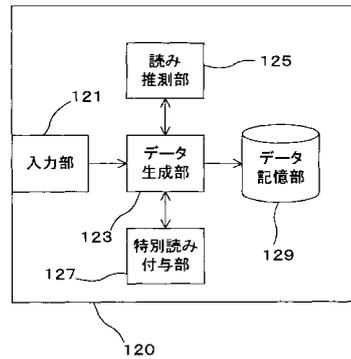
【図 9】



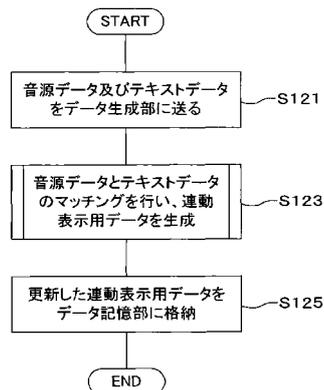
【図 11】



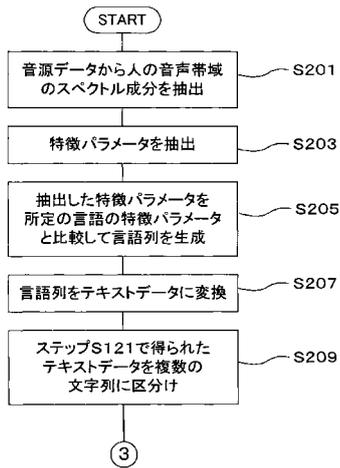
【図 12】



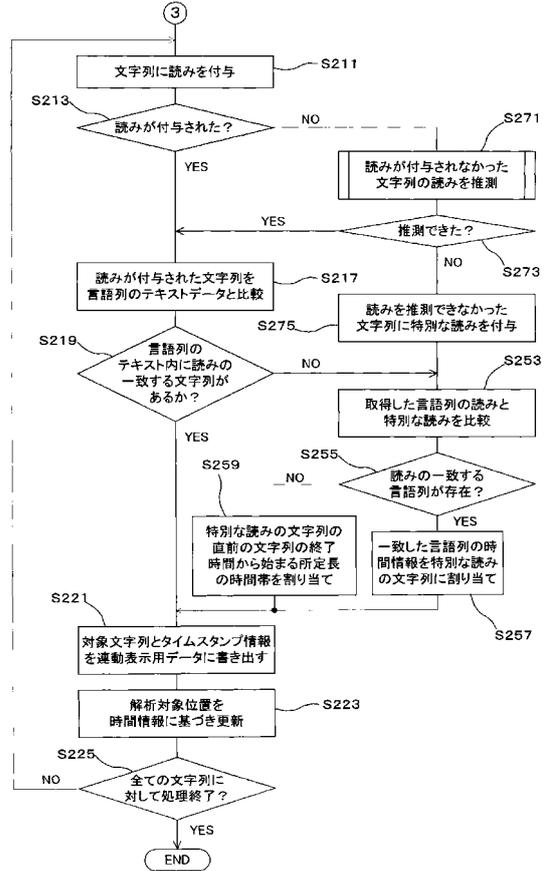
【図 13】



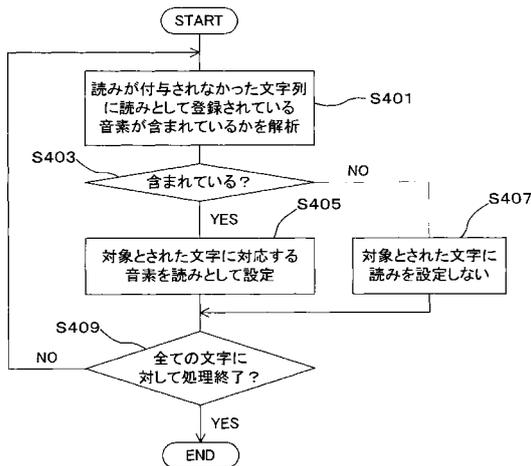
【図14】



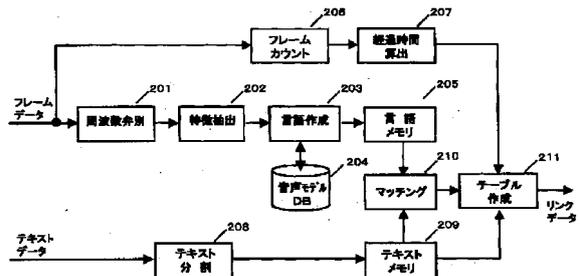
【図15】



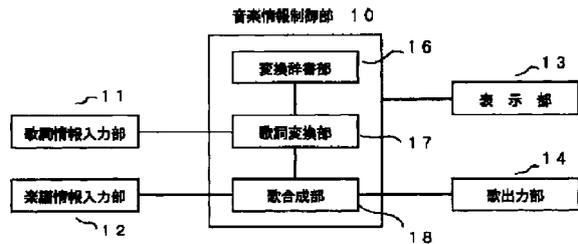
【図16】



【図18】



【図19】



【図17】

文字	音素
a	a
b	b
c	c
d	d
:	:

フロントページの続き

審査官 柴垣 俊男

- (56)参考文献 特開2003-186491(JP,A)
特開平10-254478(JP,A)
特開平8-212190(JP,A)
特開2003-280670(JP,A)

(58)調査した分野(Int.Cl., DB名)

G10K 15/04
G10L 15/00
G10L 15/22
H04N 5/445
H04N 7/173