

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第6526602号  
(P6526602)

(45) 発行日 令和1年6月5日(2019.6.5)

(24) 登録日 令和1年5月17日(2019.5.17)

(51) Int.Cl. F I  
G 1 0 L 1 5 / 2 0 ( 2 0 0 6 . 0 1 ) G 1 0 L 1 5 / 2 0 2 0 0 Z

請求項の数 8 (全 10 頁)

(21) 出願番号	特願2016-104880 (P2016-104880)	(73) 特許権者	000004226 日本電信電話株式会社 東京都千代田区大手町一丁目5番1号
(22) 出願日	平成28年5月26日(2016.5.26)	(74) 代理人	100121706 弁理士 中尾 直樹
(65) 公開番号	特開2017-211513 (P2017-211513A)	(74) 代理人	100128705 弁理士 中村 幸雄
(43) 公開日	平成29年11月30日(2017.11.30)	(74) 代理人	100147773 弁理士 義村 宗洋
審査請求日	平成30年6月26日(2018.6.26)	(72) 発明者	福富 隆明 東京都千代田区大手町一丁目5番1号 日 本電信電話株式会社内
		(72) 発明者	岡本 学 東京都千代田区大手町一丁目5番1号 日 本電信電話株式会社内

最終頁に続く

(54) 【発明の名称】 音声認識装置、その方法、及びプログラム

(57) 【特許請求の範囲】

【請求項1】

所定の閾値  $\tau_1$  未満の非音声区間を音声区間の一部とみなして、音声データの音声区間である第一音声区間  $s_1, s_2, \dots, s_L$  を検出する第一発話区間検出部と、

$\tau_1 < \tau_2$  とし、所定の閾値  $\tau_2$  未満の非音声区間を音声区間の一部とみなして、前記音声データの音声区間である第二音声区間  $t_1, t_2, \dots, t_J$  を検出する第二発話区間検出部と、

前記第一音声区間  $s_1, s_2, \dots, s_L$  に対応する前記音声データに対して音声認識を行い第一認識結果  $sr_1, sr_2, \dots, sr_L$  を得、前記第二音声区間  $t_1, t_2, \dots, t_J$  に対応する前記音声データに対して音声認識を行い第二認識結果  $tr_1, tr_2, \dots, tr_J$  を得る音声認識部と、

$i=1, \dots, L-1$  とし、 $m=1, 2, \dots, M$  とし、 $M$  は1以上  $L-1$  以下の何れかの整数とし、 $L > J$  のとき、第一認識結果  $sr_i$  と第一認識結果  $sr_{i+m}$  との類似度  $sim_{i,m}$  を計算し、前記類似度  $sim_{i,m}$  が所定の閾値より大きい場合、第一認識結果  $sr_i$  を削除し、残った第一認識結果のみを認識結果とする認識結果修正部とを含む、

音声認識装置。

【請求項2】

請求項1の音声認識装置であって、

前記類似度  $sim_{i,m}$  は、前記第一認識結果  $sr_i$  の末尾から  $n$  文字と、前記第一認識結果  $sr_{i+m}$  の先頭から  $n$  文字との間の類似度である、

音声認識装置。

【請求項3】

10

20

請求項 1 または請求項 2 の音声認識装置であって、  
M=1である、  
音声認識装置。

【請求項 4】

請求項 1 から請求項 3 の何れかの音声認識装置であって、  
第一認識結果 $sr_i$ の発話開始時刻と第一認識結果 $sr_{i+m}$ の発話開始時刻との差分を  $t_{i+m}$ とし、Mは2以上i未満の何れかの整数であり、前記認識結果修正部において、前記類似度 $i_{i,m}$ を次式により更新し、更新後の類似度 $i_{i,m}$ が所定の閾値より大きい場合、第一認識結果 $sr_i$ を削除し、残った第一認識結果のみを認識結果とする、

$$i_{i,m} = i_{i,m} / t_{i+m}$$

音声認識装置。

10

【請求項 5】

請求項 1 から請求項 4 の何れかの音声認識装置であって、  
前記認識結果修正部において、前記類似度 $i_{i,m}$ を第一認識結果 $sr_i$ に対する音素系列と第一認識結果 $sr_{i+m}$ に対する音素系列とから計算する、  
音声認識装置。

【請求項 6】

請求項 1 から請求項 5 の何れかの音声認識装置であって、  
前記認識結果修正部は、L=Jのとき、類似度 $i_{i,m}$ の計算及び第一認識結果 $sr_i$ の削除を省略し、第二認識結果を認識結果とする、  
音声認識装置。

20

【請求項 7】

所定の閾値 $t_1$ 未満の非音声区間を音声区間の一部とみなして、音声データの音声区間である第一音声区間 $s_1, s_2, \dots, s_L$ を検出する第一発話区間検出ステップと、  
 $t_1 < t_2$ とし、所定の閾値 $t_2$ 未満の非音声区間を音声区間の一部とみなして、前記音声データの音声区間である第二音声区間 $t_1, t_2, \dots, t_J$ を検出する第二発話区間検出ステップと、

前記第一音声区間 $s_1, s_2, \dots, s_L$ に対応する前記音声データに対して音声認識を行い第一認識結果 $sr_1, sr_2, \dots, sr_L$ を得、前記第二音声区間 $t_1, t_2, \dots, t_J$ に対応する前記音声データに対して音声認識を行い第二認識結果 $tr_1, tr_2, \dots, tr_J$ を得る音声認識ステップと、

30

$i=1, \dots, L-1$ とし、 $m=1, 2, \dots, M$ とし、Mは1以上L-1以下の何れかの整数とし、L>Jのとき、第一認識結果 $sr_i$ と第一認識結果 $sr_{i+m}$ との類似度 $i_{i,m}$ を計算し、前記類似度 $i_{i,m}$ が所定の閾値より大きい場合、第一認識結果 $sr_i$ を削除し、残った第一認識結果のみを認識結果とする認識結果修正ステップとを含む、

音声認識方法。

【請求項 8】

請求項 1 から請求項 6 の何れかの音声認識装置としてコンピュータを機能させるためのプログラム。

【発明の詳細な説明】

【技術分野】

40

【0001】

本発明は、人間の発話などをコンピュータに認識させ、発話を文字列に変換する音声認識技術に関する。

【背景技術】

【0002】

音声認識技術においては、発話区間検出技術(非特許文献1参照)を用いて、利用者の発話のみを切り出し、音声認識を行う。利用者の発話のみを精度よく切り出すことで、発話していない雑音区間を除外し、精度よく音声認識ができる。

【0003】

発話区間検出においては、入力音声の各フレームに対して音声尤度、非音声尤度を算出

50

し、音声区間を決定する手法がとられることが多い。

【0004】

ひとつの発話として音声認識に渡す発話は、尤度計算から音声区間と判定された区間のみを渡すのではなく、音声区間の前後にマージンとしての非音声区間を付与したり、音声区間と音声区間の間（以下、「ギャップ」ともいう）に存在する短い非音声区間も含まれる。

【0005】

どの程度のマージンを付与するか、及び、どの程度のギャップであればひとつの発話とするかは、一般に利用環境に応じて人手で調整する。

【先行技術文献】

10

【非特許文献】

【0006】

【非特許文献1】藤本雅清，“音声区間検出の基礎と最近の研究動向”，IEICE Technical Report., SP2010-23 (2010-06)，pp.7-12.

【発明の概要】

【発明が解決しようとする課題】

【0007】

しかしながら、音声対話システムなどにおいて、利用者が目的の発話を言い直すケースがあり、その場合、発話の言い直しは、自然な連続した発話と比較して言語的なつながりが低下することから音声認識精度が劣化する。また、音声対話システムとしてもその意味解釈が難しくなり、音声対話システムとしての信頼度が低下する。

20

【0008】

言い直し箇所も発話区間であることから発話区間検出処理では除外することができない。

【0009】

発話区間検出において結合する音声区間ギャップを短く設定することで、音声認識に渡される発話区間は短くなる。言い直しをする際には、短い非音声区間を挟むことから、ギャップの調整により言い直しとそれ以降の発話に分割することができるが、短く分割されやすくなると文脈としてのつながりが利用しづらくなるため認識精度が劣化する。

【0010】

30

本発明は、言い直しを検出することで言い直しによる認識精度の劣化を回避することができる音声認識装置、その方法及びプログラムを提供することを目的とする。

【課題を解決するための手段】

【0011】

上記の課題を解決するために、本発明の一態様によれば、音声認識装置は、所定の閾値  $\tau_1$  未満の非音声区間を音声区間の一部とみなして、音声データの音声区間である第一音声区間  $s_1, s_2, \dots, s_L$  を検出する第一発話区間検出部と、  $\tau_1 < \tau_2$  とし、所定の閾値  $\tau_2$  未満の非音声区間を音声区間の一部とみなして、音声データの音声区間である第二音声区間  $t_1, t_2, \dots, t_J$  を検出する第二発話区間検出部と、第一音声区間  $s_1, s_2, \dots, s_L$  に対応する音声データに対して音声認識を行い第一認識結果  $sr_1, sr_2, \dots, sr_L$  を得、第二音声区間  $t_1, t_2, \dots, t_J$  に対応する音声データに対して音声認識を行い第二認識結果  $tr_1, tr_2, \dots, tr_J$  を得る音声認識部と、  $i=1, \dots, L-1$  とし、  $m=1, 2, \dots, M$  とし、  $M$  は1以上  $L-1$  以下の何れかの整数とし、  $L > J$  のとき、第一認識結果  $sr_i$  と第一認識結果  $sr_{i+m}$  との類似度  $sim_{i,m}$  を計算し、類似度  $sim_{i,m}$  が所定の閾値より大きい場合、第一認識結果  $sr_i$  を削除し、残った第一認識結果のみを認識結果とする認識結果修正部とを含む。

40

【0012】

上記の課題を解決するために、本発明の他の態様によれば、音声認識方法は、所定の閾値  $\tau_1$  未満の非音声区間を音声区間の一部とみなして、音声データの音声区間である第一音声区間  $s_1, s_2, \dots, s_L$  を検出する第一発話区間検出ステップと、  $\tau_1 < \tau_2$  とし、所定の閾値  $\tau_2$  未満の非音声区間を音声区間の一部とみなして、音声データの音声区間である第二音

50

声区間 $t_1, t_2, \dots, t_j$ を検出する第二発話区間検出ステップと、第一音声区間 $s_1, s_2, \dots, s_L$ に対応する音声データに対して音声認識を行い第一認識結果 $sr_1, sr_2, \dots, sr_L$ を得、第二音声区間 $t_1, t_2, \dots, t_j$ に対応する音声データに対して音声認識を行い第二認識結果 $tr_1, tr_2, \dots, tr_j$ を得る音声認識ステップと、 $i=1, \dots, L-1$ とし、 $m=1, 2, \dots, M$ とし、 $M$ は1以上 $L-1$ 以下の何れかの整数とし、 $L>J$ のとき、第一認識結果 $sr_i$ と第一認識結果 $sr_{i+m}$ との類似度 $\rho_{i,m}$ を計算し、類似度 $\rho_{i,m}$ が所定の閾値より大きい場合、第一認識結果 $sr_i$ を削除し、残った第一認識結果のみを認識結果とする認識結果修正ステップとを含む。

【発明の効果】

【0013】

本発明によれば、言い直しを検出することで言い直しによる認識精度の劣化を回避することができ、音声認識を用いるシステムの信頼性を向上させるという効果を奏する。

10

【図面の簡単な説明】

【0014】

【図1】第一実施形態に係る音声認識装置の機能ブロック図。

【図2】第一実施形態に係る音声認識装置の処理フローの例を示す図。

【図3】入力音声データ、第一音声区間及び第二音声区間の例を示す図。

【図4】入力音声データ、第一認識結果及び第二認識結果の例を示す図。

【図5】認識結果修正部103の処理フローの例を示す図。

【発明を実施するための形態】

【0015】

20

以下、本発明の実施形態について、説明する。なお、以下の説明に用いる図面では、同じ機能を持つ構成部や同じ処理を行うステップには同一の符号を記し、重複説明を省略する。

【0016】

<第一実施形態>

本実施形態では、感度の異なる(ギャップの長さの異なる)2つの発話区間検出を動作させ、それぞれに対する音声認識結果を参照することで、言い直しを検出する。

【0017】

図1は第一実施形態に係る音声認識装置100の機能ブロック図を、図2はその処理フローを示す。

30

【0018】

音声認識装置100は、第一発話区間検出部101-1、第二発話区間検出部101-2、音声認識部102及び認識結果修正部103を含む。

【0019】

音声認識装置100は、マイクロホン等の音声デバイスで収音された音声データを入力とし、その音声データを音声認識し、音声認識結果を出力する。

【0020】

<第一発話区間検出部101-1及び第二発話区間検出部101-2>

第一発話区間検出部101-1は、音声データを入力とし、所定の閾値 $\theta_1$ 未満の非音声区間を音声区間の一部とみなして、音声データの音声区間である第一音声区間 $s_1, s_2, \dots, s_L$ を検出し(S101-1)、音声認識部102に出力する。

40

【0021】

第二発話区間検出部101-2は、音声データを入力とし、所定の閾値 $\theta_2$ 未満の非音声区間を音声区間の一部とみなして、音声データの音声区間である第二音声区間 $t_1, t_2, \dots, t_j$ を検出し(S101-2)、音声認識部102に出力する。ただし、 $\theta_1 < \theta_2$ とする。

【0022】

なお、本実施形態では、第一音声区間 $s_1, s_2, \dots, s_L$ 及び第二音声区間 $t_1, t_2, \dots, t_j$ として、音声区間を切り出した音声データを出力する。ただし、音声区間と、その音声区間に対応する音声データとを取得できればよいため、切り出す前の音声データと音声区間を示す情報(例えば(i)音声区間の開始時刻と終了時刻、(ii)音声区間の開始時刻と音声区間の長

50

さ等)を出力してもよい。

【0023】

例えば、第一発話区間検出部101-1及び第二発話区間検出部101-2は、音声デバイスからの入力音声データを解析し、音声区間の検出を行う。音声区間検出方法は、既存のいかなる音声区間検出方法を用いてもよく、利用環境等に合わせて最適なものを適宜選択すればよい。例えば、非特許文献1の音声区間検出方法を用いればよい。ただし、本実施形態では、検出感度の異なる2つの設定で音声区間検出処理を行う。図3は、入力音声データ、第一音声区間 $s_1, s_2, \dots, s_L$ 及び第二音声区間 $t_1, t_2, \dots, t_J$ の例を示す。

【0024】

第一発話区間検出部101-1は、言い直しや考えながら話す際に生じるような発話内での短い非音声区間にも敏感に発話を分割するように設定する(図3参照)。つまり、閾値 $\theta_1$ を小さな値に設定する。他方、第二発話区間検出部101-2は、そのような短い非音声区間は音声区間の一部とみなすように設定する(図3参照)。つまり、閾値 $\theta_2$ を閾値 $\theta_1$ よりも大きな値に設定する。例えば、第一発話区間検出部101-1及び第二発話区間検出部101-2は、内部で音声尤度と非音声尤度との比(以下、尤度比ともいう)を算出し、尤度比がある閾値 $\theta$ を下回る区間(非音声区間)が閾値 $\theta$ を上回る区間(音声区間)に挟まれている場合において、当該非音声区間がある時間長 $\tau_1$ または $\tau_2$ よりも短い場合はその非音声区間を音声区間として処理する。ここでの設定の違いは、この時間長 $\tau_1$ を非常に短く、たとえば0.1秒などに設定し、他方、時間長 $\tau_2$ を、通常の発話内における一般的な間を考慮し、0.5秒程度に設定することにある。 $\tau_1$ を非常に短くすることで発話区間(音声区間)が細かく分割されやすくなる。第一発話区間検出部101-1及び第二発話区間検出部101-2では、異なる設定で動かした複数の区間検出結果(第一音声区間 $s_1, s_2, \dots, s_L$ 及び第二音声区間 $t_1, t_2, \dots, t_J$ 、ただし、 $\theta_1 < \theta_2$ なので、第一音声区間の個数Lは必ず第二音声区間の個数J以上となる、つまり、 $L \geq J$ となる)が出力される。

【0025】

<音声認識部102>

音声認識部102は、音声データ、第一音声区間 $s_1, s_2, \dots, s_L$ 及び第二音声区間 $t_1, t_2, \dots, t_J$ (音声区間を切り出した音声データ)を入力とし、第一音声区間 $s_1, s_2, \dots, s_L$ に対応する音声データに対して音声認識を行い第一認識結果 $sr_1, sr_2, \dots, sr_L$ を得、第二音声区間 $t_1, t_2, \dots, t_J$ に対応する音声データに対して音声認識を行い第二認識結果 $tr_1, tr_2, \dots, tr_J$ を得(S102)、認識結果修正部103に出力する。音声認識方法は、既存のいかなる音声認識方法を用いてもよく、利用環境等に合わせて最適なものを適宜選択すればよい。図4は、第一認識結果 $sr_1, sr_2, \dots, sr_L$ 及び第二認識結果 $tr_1, tr_2, \dots, tr_J$ の例を示す。

【0026】

<認識結果修正部103>

認識結果修正部103は、第一認識結果 $sr_1, sr_2, \dots, sr_L$ と第二認識結果 $tr_1, tr_2, \dots, tr_J$ とを受け取り、 $L > J$ のとき、第一認識結果 $sr_i$ と第一認識結果 $sr_{i+m}$ との類似度 $\rho_{i,m}$ を計算し、類似度 $\rho_{i,m}$ が所定の閾値より大きい場合、第一認識結果 $sr_i$ を削除し、残った第一認識結果のみを認識結果とし(S103)、音声認識装置100の出力値として出力する。ただし、 $i=1, 2, \dots, L-1$ とし、 $m=1$ とする。なお、 $L > J$ のとき、Jは第二認識結果に含まれる発話の個数を示すため少なくとも1以上の値となり、Lは1より大きな値となる。

【0027】

図5は、認識結果修正部103の処理フローの例を示す。認識結果修正部103は、第一発話区間検出部101-1及び第二発話区間検出部101-2から入力された発話区間の数L, Jが異なる場合( $L \neq J$ 、ただし、 $L \geq J$ なので $L > J$ となる、図5のS103-Aのyesの場合)、発話区間数が多い方の認識結果(つまり、第一認識結果)に対して言い直しの検出を行う(S103-B~S103-H)。なお、この例では、flagが0のとき言い直しが行われていないと判断し、flagが1のとき言い直しが行われたと判断する。そこで、まず、flagを0に設定する(S103-B)。発話区間の数L, Jが同一である場合( $L=J$ 、図5のS103-Aのnoの場合)には処理は行わず、発話区間長が長い方の認識結果(第二認識結果

10

20

30

40

50

) $tr_1, tr_2, \dots, tr_J$ をそのまま出力する(S 1 0 3 - J)。このような処理により、以降の処理S 1 0 3 - B ~ S 1 0 3 - Iを省略することができる。言い直しの検出は、複数発話区間のうち時間軸上で前後するもの(本実施形態では、 $sr_i$ と $sr_{i+1}$ )の認識結果文字列の類似度 $\rho_{i,m}$ を算出し(S 1 0 3 - D)、その類似度がある閾値 $\rho_{th}$ を上回っている際に(S 1 0 3 - Eのyesの場合)、 $sr_{i+1}$ は $sr_i$ の言い直し箇所だとして判断し、flagを1に設定する(S 1 0 3 - F)。ただし、類似度の計算には認識結果文字列すべてを用いるのではなく、過去の認識結果 $sr_i$ の末尾からn文字、 $sr_{i+1}$ の先頭からn文字のみを用いるものとし、類似度計算自体は公知の技術を用いて算出する。ここでnは例えば5とする。類似度の計算は、2つの文字系列を比較し、スコアを出力できるものであればよく、たとえばDPマッチングなどを用いるとよい。類似度のスコアを0から1に正規化した場合は0.85程度にしておくともよい。計算した類似度計算結果から、近接する発話で類似度が $\rho_{th}$ を上回っている場合には、近接する過去の認識結果 $sr_i$ を削除する(S 1 0 3 - G)。この処理をすべての認識結果 $sr_i$ に対して行い(S 1 0 3 - C、なお、図5のS 1 0 3 - Cの上端内の数値は、変数iの初期値、終値、増分値を表す。)、言い直しが行われていた場合(S 1 0 3 - Hのyesの場合)、近接する発話で類似度が $\rho_{th}$ 以下であった認識結果(削除されていない第一音声認識結果 $sr_i$ )のみを出力する(S 1 0 3 - I)。なお、第一音声認識結果 $sr_L$ ( $i=L$ のとき)は、 $i=L+1$ の比較が行われることがないため、必ず出力される。なお、第一認識結果 $sr_1, sr_2, \dots, sr_L$ は、第二認識結果 $tr_1, tr_2, \dots, tr_J$ に比べると、短く分割される。そのため、第一認識結果 $sr_1, sr_2, \dots, sr_L$ は、第二認識結果 $tr_1, tr_2, \dots, tr_J$ に比べると、文脈としてのつながりが利用しづらくなり、認識精度が劣化すると考えられる。そのため、近接する

10

20

【0028】

&lt;効果&gt;

以上の構成により、言い直しを検出することで言い直しによる認識精度の劣化を回避することができ、音声認識を用いるシステムの信頼性を向上させるという効果を奏する。

【0029】

&lt;第二実施形態&gt;

第一実施形態と異なる部分を中心に説明する。

【0030】

認識結果修正部103は、第一認識結果 $sr_1, sr_2, \dots, sr_L$ と第二認識結果 $tr_1, tr_2, \dots, tr_J$ とを受け取り、 $L > J$ のとき、第一認識結果 $sr_i$ と第一認識結果 $sr_{i+m}$ との類似度 $\rho_{i,m}$ を計算し、類似度 $\rho_{i,m}$ が所定の閾値より大きい場合、第一認識結果 $sr_i$ を削除し、残った第一認識結果のみを認識結果とし(S 1 0 3)、音声認識装置100の出力値として出力する。ただし、 $i=1, 2, \dots, L-1$ とし、 $m=1, 2, \dots, M$ とし、Mは1以上L-1以下の何れかの整数とする。

30

【0031】

認識結果修正部103において、直前の発話区間とのみ認識結果の類似度比較を行うのではなく、過去の発話区間すべて、または過去M個分の発話区間(認識結果)との類似度を算出する形態もある。この場合、直近にある発話区間の方が言い直しである可能性が高いことから比較対象との時間的距離に応じた重みを導入するのがよい。たとえば、DPマッチング等で算出した類似度のスコアを $\rho_{i,m}$ とし、発話区間同士の開始時刻のずれ(第一認識結果 $sr_i$ の発話開始時刻と第一認識結果 $sr_{i+m}$ の発話開始時刻との差分)を $t_{i+m}$ とした場合、距離重みを考慮した類似度は、 $\rho_{i,m} / t_{i+m}$ と定義するとよい。つまり、

40

$$\rho_{i,m} = \rho_{i,m} / t_{i+m}$$
 として、類似度 $\rho_{i,m}$ を更新する。また、Mは3,4個程度に設定するのがよい。あるいは、過去T秒以内の発話区間のように時間的な制約をかけてもよい。Tはたとえば5秒などにするとよい。

【0032】

&lt;効果&gt;

このような構成とすることで、第一実施形態と同様の効果を得ることができる。さらに

50

、言い直しが行われているか否かの判定範囲を広げることで、より精度の高い音声認識が可能となる。なお、M=1のとき、第一実施形態と同様の構成となるため、第一実施形態は第二実施形態の一例とも言える。

【0033】

<第三実施形態>

第一実施形態と異なる部分を中心に説明する。

【0034】

認識結果修正部103において、認識結果文字列の類似度を計算するが、文字列の表記での類似度ではなく、一度音素に変換し、音素系列の類似度を計算してもよい。なお、音声認識部102で音声認識処理を行う際に、音声データから音声認識結果(文字列)に変換する過程で、音素系列を取得するため、その音素系列を利用して類似度を計算してもよい。また、本実施形態では、音素系列の類似度を計算しているが、音声データの類似度を計算してもよいし、音声データから得られる特徴量(例えばMFCC(メル周波数ケプストラム係数)等)の類似度を計算してもよい。ただし、文字列や音素系列で類似度を計算したほうが類似度の精度は良い。

10

【0035】

<その他の変形例>

本発明は上記の実施形態及び変形例に限定されるものではない。例えば、上述の各種の処理は、記載に従って時系列に実行されるのみならず、処理を実行する装置の処理能力あるいは必要に応じて並列的あるいは個別に実行されてもよい。その他、本発明の趣旨を逸脱しない範囲で適宜変更が可能である。

20

【0036】

<プログラム及び記録媒体>

また、上記の実施形態及び変形例で説明した各装置における各種の処理機能をコンピュータによって実現してもよい。その場合、各装置が有すべき機能の処理内容はプログラムによって記述される。そして、このプログラムをコンピュータで実行することにより、上記各装置における各種の処理機能がコンピュータ上で実現される。

【0037】

この処理内容を記述したプログラムは、コンピュータで読み取り可能な記録媒体に記録しておくことができる。コンピュータで読み取り可能な記録媒体としては、例えば、磁気記録装置、光ディスク、光磁気記録媒体、半導体メモリ等のようなものでもよい。

30

【0038】

また、このプログラムの流通は、例えば、そのプログラムを記録したDVD、CD-ROM等の可搬型記録媒体を販売、譲渡、貸与等することによって行う。さらに、このプログラムをサーバコンピュータの記憶装置に格納しておき、ネットワークを介して、サーバコンピュータから他のコンピュータにそのプログラムを転送することにより、このプログラムを流通させてもよい。

【0039】

このようなプログラムを実行するコンピュータは、例えば、まず、可搬型記録媒体に記録されたプログラムもしくはサーバコンピュータから転送されたプログラムを、一旦、自己の記憶部に格納する。そして、処理の実行時、このコンピュータは、自己の記憶部に格納されたプログラムを読み取り、読み取ったプログラムに従った処理を実行する。また、このプログラムの別の実施形態として、コンピュータが可搬型記録媒体から直接プログラムを読み取り、そのプログラムに従った処理を実行することとしてもよい。さらに、このコンピュータにサーバコンピュータからプログラムが転送されるたびに、逐次、受け取ったプログラムに従った処理を実行することとしてもよい。また、サーバコンピュータから、このコンピュータへのプログラムの転送は行わず、その実行指示と結果取得のみによって処理機能を実現する、いわゆるASP(Application Service Provider)型のサービスによって、上述の処理を実行する構成としてもよい。なお、プログラムには、電子計算機による処理の用に供する情報であってプログラムに準ずるもの(コンピュータに対する直

40

50

接の指令ではないがコンピュータの処理を規定する性質を有するデータ等)を含むものとする。

【0040】

また、コンピュータ上で所定のプログラムを実行させることにより、各装置を構成することとしたが、これらの処理内容の少なくとも一部をハードウェア的に実現することとしてもよい。

【図1】

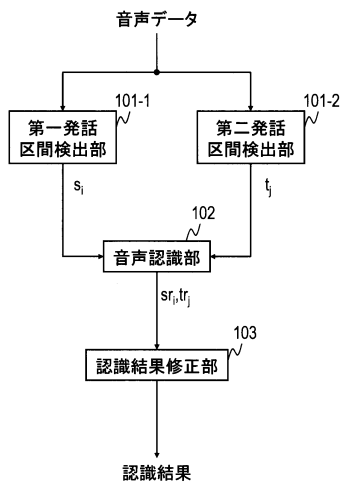


図1

【図2】

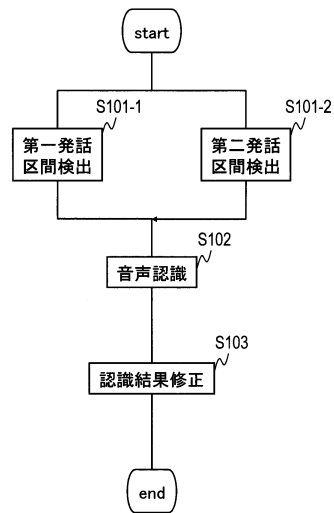


図2



【 図 3 】

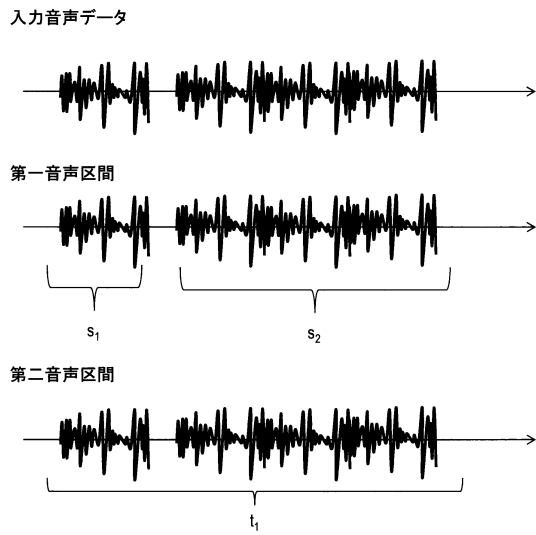


図3

【 図 4 】

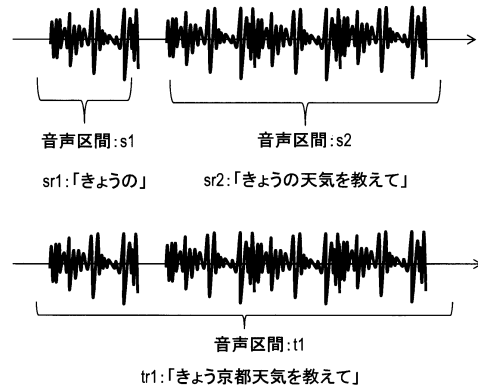


図4

【 図 5 】

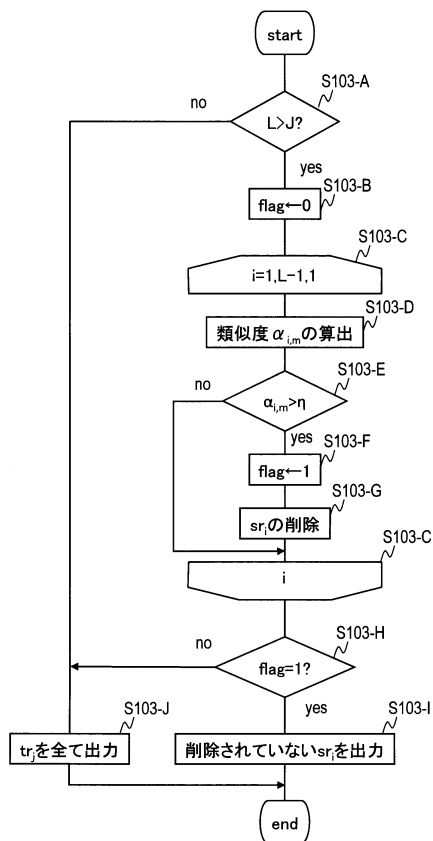


図5

---

フロントページの続き

(72)発明者 松井 清彰

東京都千代田区大手町一丁目5番1号 日本電信電話株式会社内

審査官 千本 潤介

(56)参考文献 国際公開第2012/093661(WO, A1)

特開2007-057844(JP, A)

(58)調査した分野(Int.Cl., DB名)

G10L 15/20

IEEE Xplore