

【公報種別】特許法第17条の2の規定による補正の掲載
 【部門区分】第6部門第2区分
 【発行日】平成17年10月27日(2005.10.27)

【公開番号】特開2002-91467(P2002-91467A)
 【公開日】平成14年3月27日(2002.3.27)
 【出願番号】特願2000-277024(P2000-277024)
 【国際特許分類第7版】

G 1 0 L 11/02
 G 1 0 L 15/04
 G 1 0 L 15/02
 // G 1 0 L 101:12

【F I】

G 1 0 L 3/00 5 1 3 C
 G 1 0 L 3/00 5 1 3 A
 G 1 0 L 9/14 3 0 1 A
 G 1 0 L 101:12

【手続補正書】

【提出日】平成17年7月21日(2005.7.21)

【手続補正1】

【補正対象書類名】明細書

【補正対象項目名】特許請求の範囲

【補正方法】変更

【補正の内容】

【特許請求の範囲】

【請求項1】

音声認識の対象とする音声の区間を検出する音声区間検出手段を備えた音声認識システムであって、

前記音声区間検出手段は、

予め音声の特徴を学習ベクトルとして生成する学習ベクトル生成手段と、

発話によって生じる入力音声の特徴ベクトルと前記学習ベクトルの内積を求め、その内積値が所定値以上となることを音声区間と判定する内積値判定手段とを備え、

前記内積値判定手段で判定される音声区間における前記入力音声を音声認識の対象とすることを特徴とする音声認識システム。

【請求項2】

音声認識の対象とする音声の区間を検出する音声区間検出手段を備えた音声認識システムであって、

前記音声区間検出手段は、

予め音声の特徴を学習ベクトルとして生成する学習ベクトル生成手段と、

非音声期間内に生じる入力音声の線形予測残差パワーに基づいて、音声と雑音とを識別するための閾値を生成する閾値生成手段と、

発話によって生じる入力音声の特徴ベクトルと前記学習ベクトルの内積を求め、その内積値が所定値以上となることを音声区間と判定する内積値判定手段と、

前記発話によって生じる入力音声の線形予測残差パワーが前記閾値生成手段で生成された閾値より大きくなることを音声区間と判定する線形予測残差パワー判定手段とを備え、

前記内積値判定手段と線形予測残差パワー判定手段で判定される音声区間における前記入力音声を音声認識の対象とすることを特徴とする音声認識システム。

【請求項3】

前記非音声期間内に生じる入力音声の特徴ベクトルと前記学習ベクトルの内積を求め、

その内積値が所定値以上になるときは、前記内積値判定手段の判定処理を停止させる誤判定制御手段を備えることを特徴とする請求項2に記載の音声認識システム。

【請求項4】

前記非音声期間内によって生じる入力音声の線形予測残差パワーを求める演算手段を備えると共に、

前記演算手段が求める前記線形予測残差パワーが所定値以下になるときは、前記内積値判定手段の判定処理を停止させる誤判定制御手段を備えることを特徴とする請求項2に記載の音声認識システム。

【請求項5】

前記非音声期間内によって生じる入力音声の線形予測残差パワーを求める演算手段を備えると共に、

前記非音声期間内に生じる入力音声の特徴ベクトルと前記学習ベクトルの内積を求めその内積値が所定値以上になるとき、または、前記非音声期間内に生じる入力音声の線形予測残差パワーが所定値以下になるときは、前記内積値判定手段の判定処理を停止させる誤判定制御手段を備えることを特徴とする請求項2に記載の音声認識システム。

【手続補正2】

【補正対象書類名】明細書

【補正対象項目名】0049

【補正方法】変更

【補正の内容】

【0049】

更に第1の検出部100には、実際に発話された入力音声のデータ S_a を上記同様のフレーム単位に区分けして出力するフレーム化部16と、フレーム単位の入力音声データ S_{af} をLPC分析することにより、ケプストラム領域でのM次元の特徴ベクトル A と予測残差パワーとを求めるLPCケプストラム演算部17と、学習ベクトル V と特徴ベクトル A の内積 VTA を求める内積演算部18と、内積 VTA を所定の閾値と比較し、 VTA のときに音声区間と判定する第1の閾値判定部19が備えられており、第1の閾値判定部19の判定結果 $D1$ を音声区間決定部300に供給するようになっている。

【手続補正3】

【補正対象書類名】明細書

【補正対象項目名】0096

【補正方法】変更

【補正の内容】

【0096】

また、非音声期間内に生じる背景音の特徴ベクトルと学習ベクトルの内積の値が所定値以上になる場合、又は、非音声期間内に生じる入力音声の線形予測残差パワーが所定値以下になる場合、又は、これら両者の場合に、入力音声の特徴ベクトルとの内積の値に基づいて音声区間の検出を行わず、実際の発話によって生じる入力音声の予測残差パワーが所定閾値より大きくなるときを音声区間とすることとしたので、SN比が高くなって背景音のスペクトルが高域で高くなるような背景下での音声区間の検出精度を向上させることができる。