

【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 29 年 4 月 27 日 (2017.4.27)

【公開番号】特開 2016-142936 (P2016-142936A)
 【公開日】平成 28 年 8 月 8 日 (2016.8.8)
 【年通号数】公開・登録公報 2016-047
 【出願番号】特願 2015-19009 (P2015-19009)
 【国際特許分類】

G 1 0 L 13/10 (2013.01)

G 1 0 L 25/69 (2013.01)

【F I】

G 1 0 L 13/10 1 1 3 B

G 1 0 L 25/69

【手続補正書】
 【提出日】平成 29 年 3 月 22 日 (2017.3.22)
 【手続補正 1】
 【補正対象書類名】特許請求の範囲
 【補正対象項目名】全文
 【補正方法】変更
 【補正の内容】
 【特許請求の範囲】
 【請求項 1】

メモリから各種プログラムを読みだして実行し、音声合成処理で用いられる音声合成用データを作成するプロセッサが、第 1 種音声のコーパスデータの入力を受け付け、当該コーパスデータに含まれる音声データから第 1 韻律を抽出するステップと、

前記プロセッサが、予め用意されている第 2 種音声の韻律モデルを前記コーパスデータに含まれるテキストに対して適用し、前記テキストに対応する第 2 韻律を予測するステップと、

前記プロセッサが、前記第 1 韻律と前記第 2 韻律の差分値を算出するステップと、

前記プロセッサが、前記テキストに含まれる文字のうち、前記差分値が所定の閾値よりも大きい文字を判定するステップと、

前記プロセッサが、前記判定するステップの結果に基づいて、前記第 1 韻律と前記第 2 韻律との違いに起因する特徴テキスト部分に対応するデータを抽出するステップと、を含むことを特徴とする音声合成用データ作成方法。

【請求項 2】

請求項 1 において、

前記第 1 種音声は口語調音声であり、前記第 2 種音声は読み上げ調音声であり、

前記第 2 種音声の韻律モデルは、読み上げ韻律・音韻予測モデルであることを特徴とする音声合成用データ作成方法。

【請求項 3】

請求項 2 において、

前記抽出された特徴テキスト部分に対応するデータは、前記特徴テキスト部分の音声波形データ、韻律・音韻情報、及びテキストデータを含むことを特徴とする音声合成用データ作成方法。

【請求項 4】

請求項 3 において、さらに、

前記プロセッサが、前記特徴テキスト部分のテキストデータを用いて、与えられるテキストデータにおける口語調表現を抽出するためのルールを生成するステップを含むことを

特徴とする音声合成用データ作成方法。

【請求項 5】

請求項 4 において、

前記プロセッサは、前記特徴テキスト部分のテキストデータに加えて、当該テキストデータが含まれる口語調テキストの前後のコンテキスト情報を用いて前記ルールを生成することを特徴とする音声合成用データ作成方法。

【請求項 6】

メモリから各種プログラムを読みだして実行し、音声合成処理で用いられる音声合成用データを作成するプロセッサが、口語調音声のコーパスデータの入力を受け付け、当該コーパスデータに含まれる音声データから口語調韻律データを抽出するステップと、

前記プロセッサが、予め用意されている読み上げ調の韻律モデルを前記コーパスデータに含まれるテキストに対して適用し、前記テキストに対応する読み上げ調韻律データを予測するステップと、

前記プロセッサが、前記口語調韻律データと前記読み上げ調韻律データの差分値を算出するステップと、

前記プロセッサが、前記差分値に基づいて、前記テキストのセグメントに対して、当該セグメントの口語調の程度を示す口語調度を算出し、前記口語調韻律データに付与するステップと、

前記プロセッサが、前記口語調度が付与された前記口語調韻律データを用いて、前記音声合成用データを生成するステップと、

を含むことを特徴とする音声合成用データ作成方法。

【請求項 7】

請求項 6 において、

前記音声合成用データを生成するステップは、前記プロセッサが、前記口語調度が付与された前記口語調韻律データを用いて、入力テキストの口語調度を予測するための統計モデルである口語調度予測モデルを生成することを含むことを特徴とする音声合成用データ作成方法。

【請求項 8】

各種プログラムを格納するメモリと、

前記メモリから前記各種プログラムを読みだして実行し、音声合成処理で用いられる音声合成用データを作成するプロセッサと、を有し、

前記プロセッサは、

第 1 種音声のコーパスデータの入力を受け付け、当該コーパスデータに含まれる音声データから第 1 韻律を抽出する処理と、

予め用意されている第 2 種音声の韻律モデルを前記コーパスデータに含まれるテキストに対して適用し、前記テキストに対応する第 2 韻律を予測する処理と、

前記第 1 韻律と前記第 2 韻律の差分を算出する処理と、

前記テキストに含まれる文字のうち、前記差分が所定の閾値よりも大きい文字を判定する処理と、

前記第 1 韻律と前記第 2 韻律との違いに起因する特徴テキスト部分を抽出する処理と

、

を実行する音声合成用データ作成装置。

【請求項 9】

請求項 8 において、

前記第 1 種音声は口語調音声であり、前記第 2 種音声は読み上げ調音声であり、

前記第 2 種音声の韻律モデルは、読み上げ韻律・音韻予測モデルであることを特徴とする音声合成用データ作成装置。

【請求項 10】

請求項 9 において、

前記抽出された特徴テキスト部分に対応するデータは、前記特徴テキスト部分の音声波

形データ、韻律・音韻情報、及びテキストデータを含むことを特徴とする音声合成用データ作成装置。

【請求項 11】

請求項 10 において、

前記プロセッサは、さらに、前記特徴テキスト部分のテキストデータを用いて、与えられるテキストデータにおける口語調表現を抽出するためのルールを生成する処理を実行することを特徴とする音声合成用データ作成装置。

【請求項 12】

請求項 11 において、

前記プロセッサは、前記特徴テキスト部分のテキストデータに加えて、当該テキストデータが含まれる口語調テキストの前後のコンテキスト情報を用いて前記ルールを生成することを特徴とする音声合成用データ作成装置。