

【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】 令和 6 年 9 月 10 日 (2024.9.10)

【公開番号】 特開 2023-164770 (P2023-164770A)
【公開日】 令和 5 年 11 月 13 日 (2023.11.13)
【年通号数】 公開公報 (特許) 2023-213
【出願番号】 特願 2023-151074 (P2023-151074)
【国際特許分類】

G 1 0 L 13/02 (2013.01)

G 1 0 L 25/63 (2013.01)

G 1 0 L 15/10 (2006.01)

G 1 0 L 13/033 (2013.01)

【F I】

G 1 0 L 13/02 1 3 0 Z

G 1 0 L 25/63

G 1 0 L 15/10 5 0 0 N

G 1 0 L 13/033 1 0 2 A

【手続補正書】

【提出日】 令和 6 年 8 月 30 日 (2024.8.30)

【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】

コールセンターの顧客対応に利用される音声処理装置であって、

第 1 のユーザの発話音声の信号である発話音声信号を取得する取得部と、

前記発話音声に含まれる前記第 1 のユーザの感情を抑制するように前記発話音声の音量
又は声質の少なくとも一方を変換した変換音声を作成する音声変換部と、

前記発話音声信号に対応する前記第 1 のユーザの感情情報を生成する感情認識部と、

第 2 のユーザに対して前記感情情報を表示する表示部と、

前記第 2 のユーザのストレス状況に関するストレス情報を生成するストレス認識部と、
を備え、

前記表示部は、生成された前記変換音声及び前記感情情報に基づいて、前記第 2 のユーザ
に対して出力される前記変換音声の出力タイミング毎に、前記変換音声に対応する前記
感情情報に基づく感情レベルを時系列に表示し、

前記表示部は、生成された前記ストレス情報に基づいて、前記第 2 のユーザの当該スト
レス情報を表示する、
音声処理装置。

【請求項 2】

前記感情認識部は、前記発話音声信号、当該発話音声信号から抽出した特徴量、当該発
話音声信号から生成したテキストデータ、当該テキストデータから抽出された特徴量、又
はこれらの少なくとも二つの組み合わせを入力とし、当該発話音声信号の発話者の感情情
報を出力するよう機械学習された感情認識モデルに、前記取得部が取得した発話音声信号
、当該発話音声信号から抽出した音声特徴量、当該発話音声信号から生成したテキストデ
ータ、当該テキストデータに対応するテキスト特徴量、又はこれらの少なくとも二つの組
み合わせを入力することにより、前記取得部が取得した発話音声信号に対応する第 1 のユ

10

20

30

40

50

ーザの感情情報を生成する、請求項 1 に記載の音声処理装置。

【請求項 3】

前記発話音声信号に基づいて抽出される特徴量を前記音声認識モデルに入力して、一以上の単語からなる単語列を含むテキストデータを生成する音声認識部と、

前記テキストデータに含まれる特定の単語列を検出し、前記特定の単語列を除去又は前記特定の単語列を他の単語列に置換したテキストデータを生成する除去部と、

をさらに備える、請求項 1 に記載の音声処理装置。

【請求項 4】

前記除去部は、前記特定の単語列が検出される場合、前記第 1 のユーザに対する警告に関する情報を生成し、

生成された当該情報は、前記第 1 のユーザが操作する情報処理装置において出力される、請求項 3 に記載の音声処理装置。

【請求項 5】

前記特定の単語列は、前記第 2 のユーザに心理的悪影響を与える一以上の単語を含む、請求項 3 又は 4 に記載の音声処理装置。

【請求項 6】

前記音声変換部は、前記発話音声をテキスト化せずに前記変換音声を生成する、請求項 1 に記載の音声処理装置。

【請求項 7】

コールセンターの顧客対応に利用される音声処理装置であって、

第 1 のユーザの発話音声の信号である発話音声信号を取得する取得部と、

第 2 のユーザに伝わる前記第 1 のユーザの感情を軽減するように前記発話音声の音量又は声質の少なくとも一方を変換した変換音声を生成する音声変換部と、

前記発話音声信号に対応する前記第 1 のユーザの感情情報を生成する感情認識部と、

前記第 2 のユーザに対して前記感情情報を表示する表示部と、

前記第 2 のユーザのストレス状況に関するストレス情報を生成するストレス認識部と、を備え、

前記表示部は、生成された前記変換音声及び前記感情情報に基づいて、前記第 2 のユーザに対して出力される前記変換音声の出力タイミング毎に、前記変換音声に対応する前記感情情報に基づく感情レベルを時系列に表示し、

前記表示部は、生成された前記ストレス情報に基づいて、前記第 2 のユーザの当該ストレス情報を表示する、音声処理装置。

【請求項 8】

前記第 1 のユーザの感情は、怒りの感情である、請求項 1 又は 7 に記載の音声処理装置。

【請求項 9】

音声を出力する出力部と、

前記ストレス情報に基づいて、前記発話音声又は前記変換音声のいずれを前記出力部から出力するかを切り替える制御部と、

をさらに備える、請求項 1 又は 7 に記載の音声処理装置。

【請求項 10】

前記ストレス認識部は、学習モデルに基づいて前記ストレス情報を生成し、

前記学習モデルは、前記発話音声と、前記発話音声を聞いた際の前記第 2 のユーザのストレスとが時間軸上で関連付けされた情報に基づいて学習される、

請求項 1 又は 7 に記載の音声処理装置。

【請求項 11】

コールセンターの顧客対応に利用される音声処理方法であって、

第 1 のユーザの発話音声の信号である発話音声信号を取得する取得工程と、

前記発話音声に含まれる前記第 1 のユーザの感情を抑制するように前記発話音声の音量

10

20

30

40

50

又は声質の少なくとも一方を変換した変換音声を作成する音声変換工程と、

前記発話音声信号に対応する第1のユーザの感情情報を生成する感情認識工程と、

第2のユーザに対して前記感情情報を表示する表示工程と、

を備え、

前記表示工程は、生成された前記変換音声及び前記感情情報に基づいて、前記第2のユーザに対して出力される前記変換音声の出力タイミング毎に、前記変換音声に対応する前記感情情報を表示するとともに、当該感情情報に基づく感情レベルを時系列に表示し、

前記表示工程は、生成された前記ストレス情報に基づいて、前記第2のユーザの当該ストレス情報を表示する、

音声処理方法。

10

【請求項12】

コンピュータを、請求項1から10のいずれか一項に記載の音声処理装置として機能させるためのプログラム。

【請求項13】

コールセンターの顧客対応に利用される音声処理装置であって、

第1のユーザの発話音声の信号である発話音声信号を取得する取得部と、

前記発話音声に含まれる前記第1のユーザの感情を抑制するように前記発話音声の音量又は声質の少なくとも一方を変換した変換音声を作成する音声変換部と、

前記発話音声信号に対応する前記第1のユーザの感情情報を生成する感情生成部と、

第2のユーザが感じるストレスの推定値を示すストレス情報を出力するストレス認識モデルを学習する学習部と、

20

前記第2のユーザから、前記第1のユーザの感情的発話に起因して感じたストレスの度合いの入力を受け付ける入力受付部と、

前記第2のユーザに対して前記感情情報を表示する表示部と、

を備え、

前記表示部は、前記第2のユーザに対して出力される前記変換音声の出力タイミングに合わせて、前記変換音声に対応する前記感情情報を表示し、

前記学習部は、前記入力受付部が受け付けた前記入力に基づいて生成された、発話音声信号と当該発話音声信号を聞いた際の前記第2のユーザのストレスとが時間軸上で関連付けされたデータセットを用いて、前記ストレス認識モデルを学習する、

30

音声処理装置。

40

50