

声に対する印象を用いた合成音声ライブラリ探索システムの提案

情報 太郎 情報 花子

情報大学情報学部

1 はじめに

人の歌声や喋り声を人工的に再現する音声合成ソフトは数多く存在しており、それらソフトのほとんどが複数種類の声を切り替えて使用できる。また、その中でもいくつかのソフトでは個人が声の元となる合成音声ライブラリを作成し、第三者による利用を前提とした配布を行える。例えば、喋り声を対象とした合成音声ソフト COEIROINK ではユーザの作成した音声合成モデルが 350 キャラクタ分以上配布されているほか [1]、歌声を対象とした合成音声ソフト UTAU では同ソフト上で使用できる UTAU 音源が 7000 キャラクタ分以上存在する [2]。このように、今や合成音声ソフトの利用者は使える声に対し非常に多くの選択肢を持っており、その全てを把握することは現実的ではない。合成音声を利用するシーンにおいて、声を持つイメージや印象は声を選ぶ上で考慮すべき要素であり、例えば喋らせるアナウンスの内容や、歌わせる曲調など用途に合った声質を持つライブラリを探すことは重要なプロセスである。しかし、現状声を持つ印象を知るには実際に聴いてみるのが最も有力な手段であり、数多あるライブラリの生み出す声を十分な数聴き比べ適切な声を選択するには多大な手間と時間を要する。さらにその結果として、多くのユーザがライブラリを選ぶ際、普段の生活の中で聞いたことのある声の中から声を選択し、結果としてユーザ全体の中で使われる声に大きな偏りが生じる問題も発生する。万に近い数存在するライブラリのうち実際にユーザに用いられる声は一握りであり、ほとんどのライブラリはユーザに用いられることなく埋もれてしまう。

そこで本研究では、ライブラリごとの声に対する印象を事前に数値化し、それを用いてユーザの求める声に近いライブラリを探索するシステムを提案する。探索対象とする合成音声ソフトは特に利用できるライブラリの多い UTAU を対象とし、一般に UTAU 音源と呼ばれるライブラリを扱うが、将来的に他の合成音声ソフトにも対応できるよう拡張性を備えるものとする。本システムでは声に対する印象を複数の印象軸ごとに評価スコアとして数値化し、ユーザは理想としてイメージする声の評価スコアを入力することで、目的に合った音源を探索することができる。評価スコアの軸には、例えばスコアの高低を女性らしい声・男性らしい声に対応させた”性別感”など、ユーザが声からスコアを、あるいはスコアから声のある程度想定できるような直感的な軸を用いることが望ましい。各音源に対する評価スコアは、アンケート調査によって集められたデータをもとに音源ファイルから各評価スコアを推定できる機械学習モデルを作成し、それを用いて付与する。また、本システムは開発する上で実際にユーザに利用されることを想定し、多くの人が手軽に利用できるよう Web アプリケーションとして実装する。

2 関連研究

2.1 *声質とは

*声の性別感などは音声波形などのうちの部分によって決まるのか？先行研究を*探して*引用する

2.2 *声から声質を

- (1) ランダムフォレスト法を用いた歌声が聞き手に与える印象の単語推定に関する研究
- (2) 重回帰混合正規分布モデルに基づく声質制御における制御パラメータの設計
- (3) スペクトル包絡情報を入力とした Deep Neural Network に基づく歌声のための声質評価

2.3 *探索システム

- (1) 声質の可視化を用いた所望音声検索システムの提案
- (2) 歌声合成システムの音源データ検索のための声質評価値推定法

3 *声質に対する評価スコアの推定

3.1 *推定の流れ

- (1) UTAU 音源の形式について
- (2) UTAU 音源アンケートについて
- (3) データの前処理・理由と目的
- (4) モデルの構築・理由と目的
- (5) 結果

3.2 *結果の評価

- (1) モデルの評価
- (2) 評価スコアの自動推定

4 *ここに探索システムの名前を入力

4.1 *システムの概要

- (1) システムの機能
- (2) システムの画面イメージ
- (3) システムの利用方法
- (4) システムの実装

4.2 *システムの評価

- (1) システムの有用性
- (2) システムの拡張性

5 おわりに

参考文献

- [1] COEIROINK, MYCOEIROINK, <https://coeiroink.com/mycoeiroink/list>
- [2] Vocaloid Database, <https://vocaldb.net/Search?searchType=Artist&artistType=UTAU>