

ユーザ歌唱を真似て歌声合成する

中野 倫靖, 後藤 真孝 (産業技術総合研究所)

歌声合成を使用した楽曲制作支援

- ★ 高品質な (人間らしい) 歌声を手軽に合成したい
 - ★ 様々な歌声合成システムや音源データ (声質, 歌手) を自由に切り替えて利用したい
 - ★ 歌がうまくなくても使いたい
 - ★ 歌ってニュアンスを伝えたい
- 歌唱力補正
歌詞アラインメント
歌声合成パラメータ推定

VocaListener: ユーザ歌唱を真似る歌声合成パラメータ推定システム



従来 歌声合成パラメータ推定

大変

- ★ 手作業で調整
- ★ ユーザ歌唱から音高・音量を抽出しそのまま合成パラメータとする

ロボストでない

(原因) 合成結果は音源データ等に依存

CV01 (green line), CV02 (orange line)

全く同じ楽譜・歌詞・合成パラメータを与え異なる音源データ (CV01, CV02) で合成した結果

従来 歌詞アラインメント

ユーザ歌唱と歌詞の時間的対応付け

- ★ Viterbi アラインメントによるユーザ歌唱と歌詞の音素との時間的対応付け
- ✗ アラインメント結果に誤りが生じる (原因) Viterbi アラインメントの性能
- ✗ 発音開始時刻・音長と合成結果が一致しない (原因) 歌声合成システムの特長

本研究

手軽
ロボスト

- ★ ユーザ歌唱と合成歌唱の音高・音量を比較しながらパラメータを反復推定

反復により合成結果がユーザ歌唱へ近付いていく

合成 (音高) パラメータの反復推定の様子

本研究

Viterbi アラインメントに加えて

- ★ ユーザ歌唱と合成歌唱の有声区間のずれを補正

ユーザ歌唱と合成歌唱の有声区間

- ★ 音節境界の誤り箇所をユーザが指摘して訂正

誤り箇所をユーザが指摘

た ち ど ま る と き

ど ま る と ➔ 新しい候補をユーザへ提示

➔ 正解するまで繰り返す ➔ 直らなければ手作業で訂正

本研究 歌唱力補正

- ★ 歌唱スタイル変更機能
- ★ 音高変更機能

ビブラート区間とそれ以外の音高・音量を抑制・強調

ビブラート以外の区間 (抑制), ビブラート区間 (強調)

調子はずれ補正が効果的に機能した箇所

適用前 (blue line), 適用後 (red line)

歌唱力補正の適用

- 調子はずれ補正: 音高遷移が半音単位となるよう音高を (有声区間毎に) ずらす
- 音高トランスポーズ: 音高を全体的にもしくは部分的に上下させる

まとめ

- ★ ユーザ歌唱を真似る機能と歌唱力補正機能を持つ歌声合成パラメータ推定システム VocaListener を実現
 - ➔ 歌声合成パラメータの反復推定 ➔ 音源データの違いなどにロボストに推定できる
 - ➔ 歌詞アラインメントの誤り訂正 ➔ 音節境界の誤りをユーザが指摘することで訂正できる
 - ➔ 歌唱力補正機能 ➔ 歌唱力を補正して合成でき、ユーザ自身が歌唱できない表現も合成できる