

A ユーザ

目標歌唱 

歌詞 立ち止まる時 また ふと振り返る

合成歌唱 

G

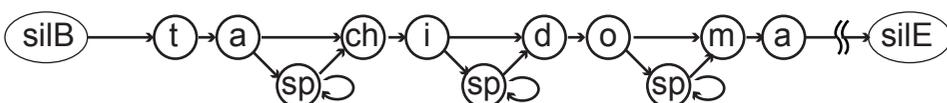
VocaListener-front-end
形態素解析、Viterbi アラインメント、音高推定、ビブラート検出、歌声合成システムとその音源データ

目標歌唱と歌詞(音素)の時間的対応付け (Viterbi アラインメント)

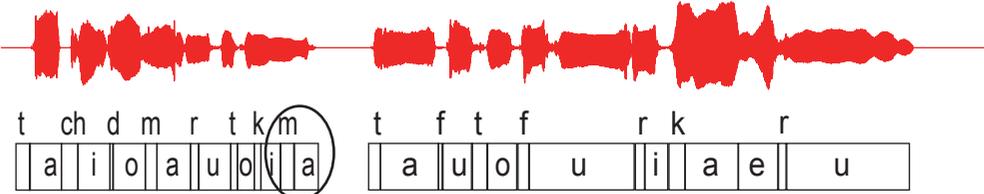
(1) かな文字列への変換(形態素解析 + ユーザによる訂正) **B**

たちどまるとき また ふとふりかえる
tachidomarutoki mata futofurikaeru

(2) アラインメント用文法の生成

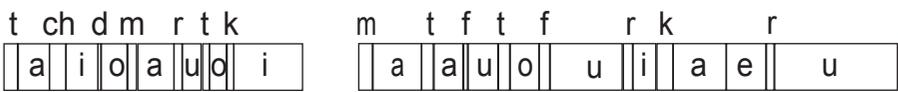


(3) Viterbi アラインメント (各音素の始端と終端を推定)

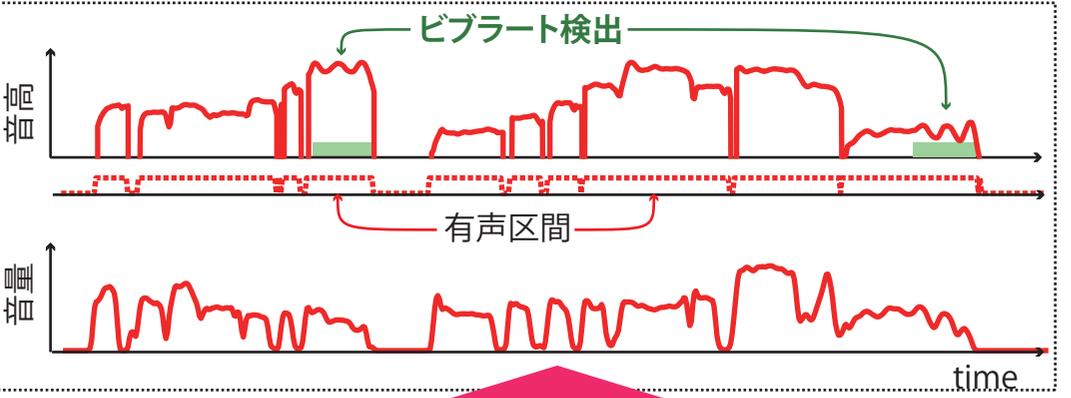


フレーズをまたがるなど、大きな誤りがあった場合、ユーザが訂正 **C**

(4) 誤りを訂正した場合、再度フレーズ内で Viterbi アラインメントを行う



目標歌唱の分析(音高・音量・ビブラート・有声/無声)



VocaListener-plus **D**

音高変更機能 (調子はずれ補正、音高トランスポーズ)
歌唱スタイル変更機能 (ビブラート深さの調節、音高・音量スムージング)

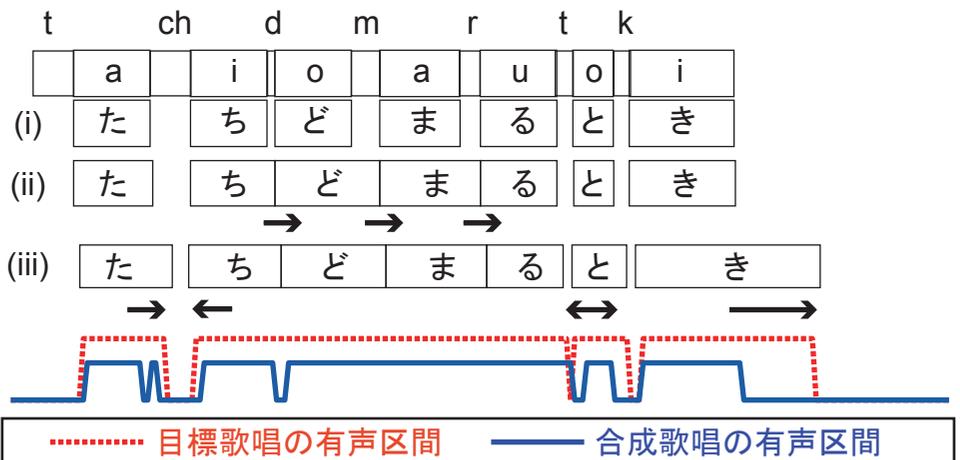
VocaListener-core
歌声分析、歌声合成、歌声合成パラメータの更新 **E**

歌詞アラインメント

(1) 有声区間のずれを補正

(i) 歌詞アラインメントの初期値 ⇒ Viterbiアラインメント結果の母音部分
(ii) 二つの音節をつなげる(連続する有声区間は、前の音節の終端を伸ばす)
(iii) 始端と終端を伸縮させる
「ノートナンバー決定 ⇒ 歌声合成 ⇒ 伸縮」の処理を繰り返す

t ch d m r t k
a i o a u o i
(i) た ち ど ま る と き
(ii) た ち ど ま る と き
(iii) た ち ど ま る と き



..... 目標歌唱の有声区間 — 合成歌唱の有声区間

(2) 音節境界の誤りを訂正 **F**

誤り箇所をユーザが指摘

た ち ど ま る と き

境界候補の選出:
- 目標歌唱の MFCC変動が大きい
- 過去に指摘されていない

ど ま る と
ど ま る と
ど ま る と

音高合わせ
歌声合成
エラー計算

⇒ エラー最小の候補をユーザへ提示 ⇒ 正解するまで繰り返す

音高パラメータの推定

・ノートナンバーの決定

た ち ど ま る と き
64 63 64 63 64 66 68

・ピッチバンド(PIT, PBS)の決定



音量パラメータの推定

・ダイナミクス(DYN)の決定

