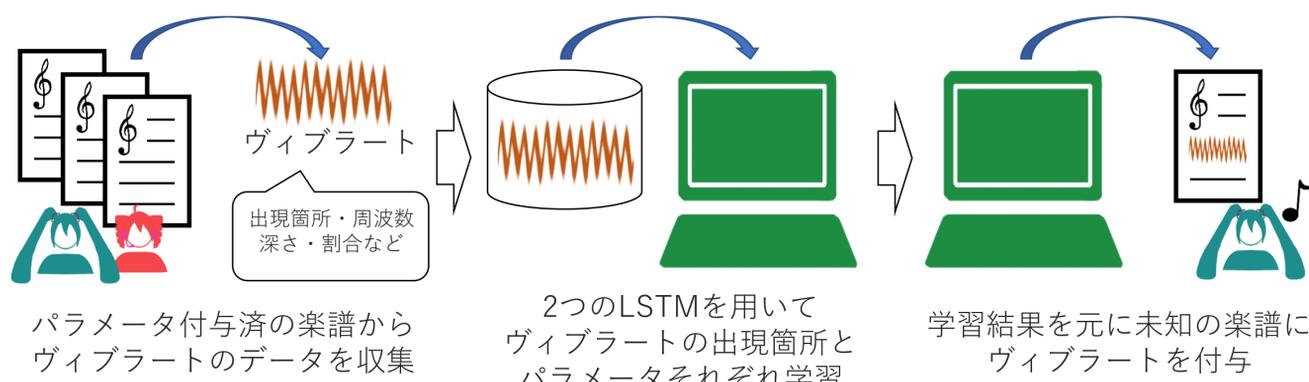


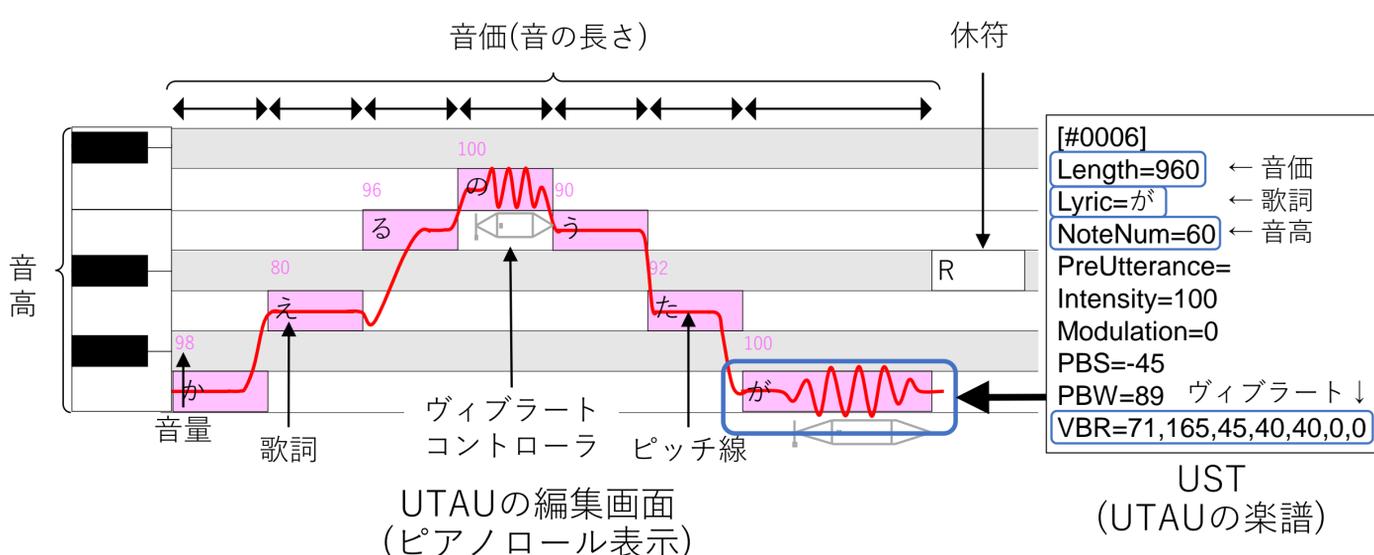
合成音声歌唱に固有なヴィブラートの楽譜レベルでの学習について

田中 瑞穂

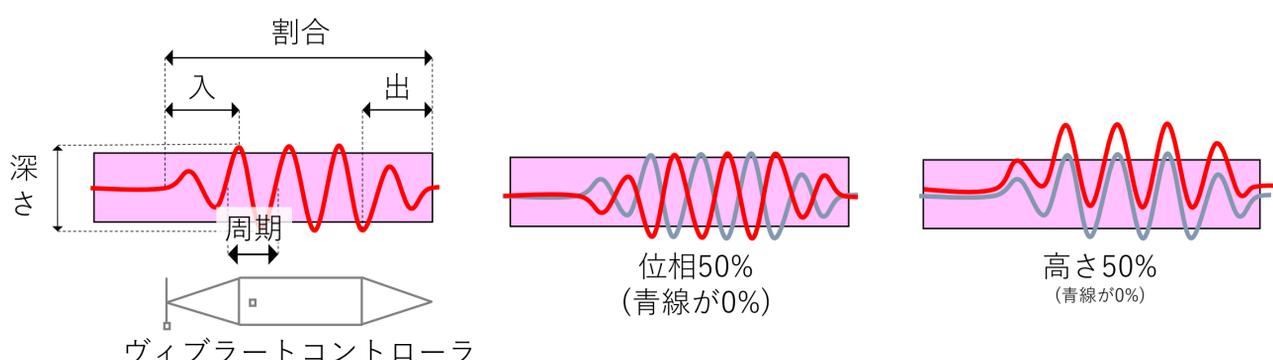
研究概要



音声合成ソフトUTAUについて



ヴィブラートコントローラについて



UTAUの歌唱にヴィブラートを付与する機能

割合	音価に対するヴィブラート長の割合	入	深さが最大になる地点
周期	波形の周期	出	深さが減少していく地点
深さ	波形の振幅	位相	位相, 100%で360度移動する
		高さ	音高, 100%で半音上がる

合成音声歌唱ヴィブラートの統計

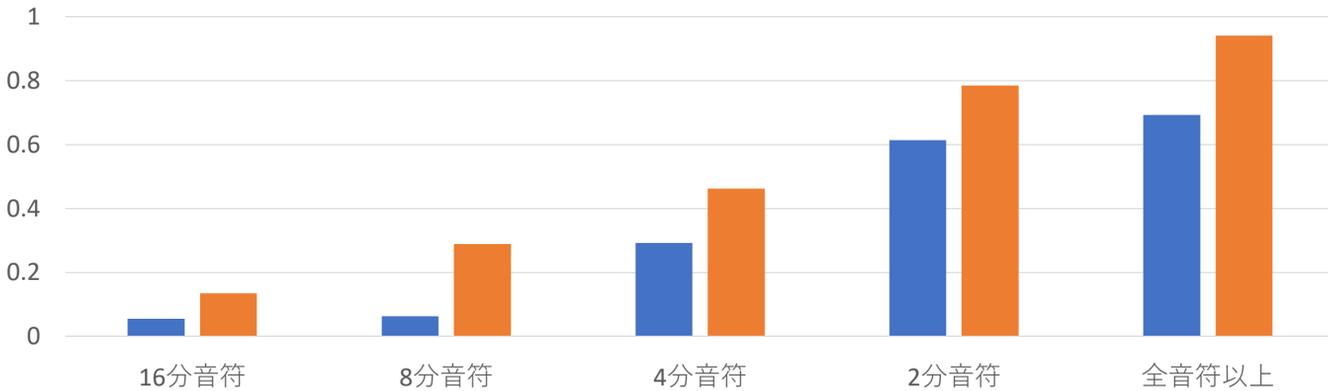
人間歌唱のヴィブラートにおける仮説

- 音価が大きいほどヴィブラートが出現する
- 音高が大きいほどヴィブラートが出現する(一定以上になると出現率が減少する)
- テンポが遅いほどヴィブラートが出現する

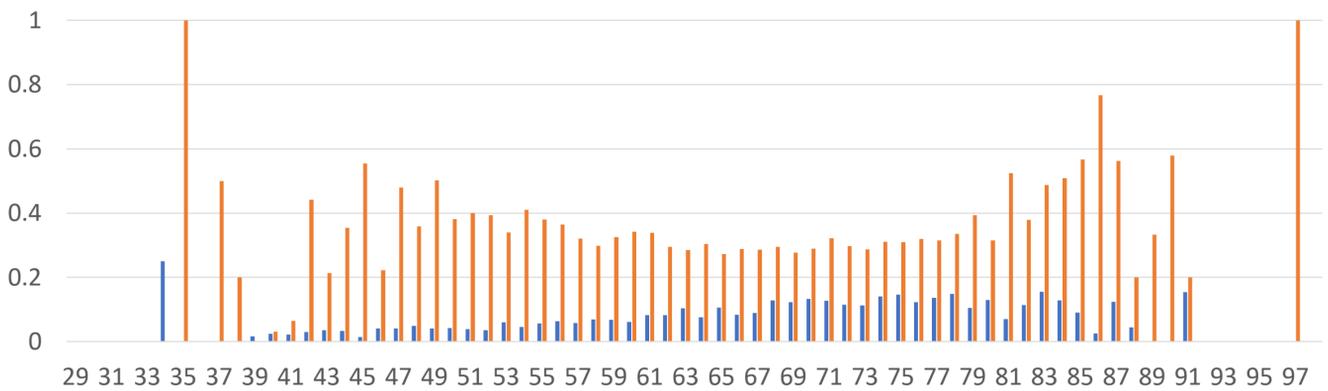
統計に用いた楽譜

- VSQX(VOCALOID楽譜, 青で表記): 144曲
- UST(UTAU楽譜, オレンジで表記): 773曲

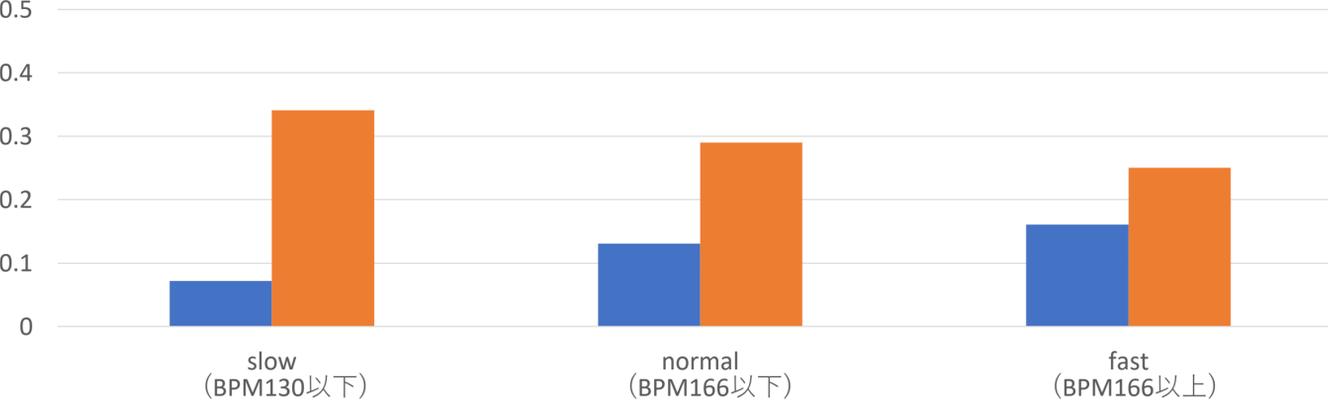
音価に対するヴィブラート出現率



音高に対するヴィブラート出現率



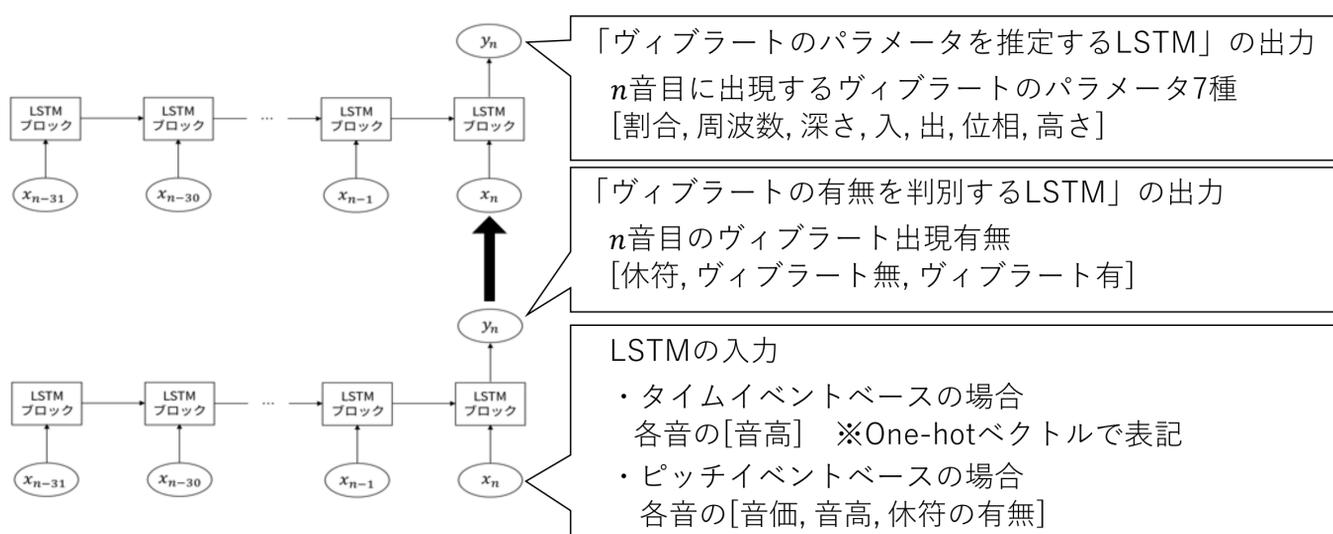
テンポに対するヴィブラート出現率



合成音声歌唱に固有なヴィブラートの楽譜レベルでの学習について

田中 瑞穂

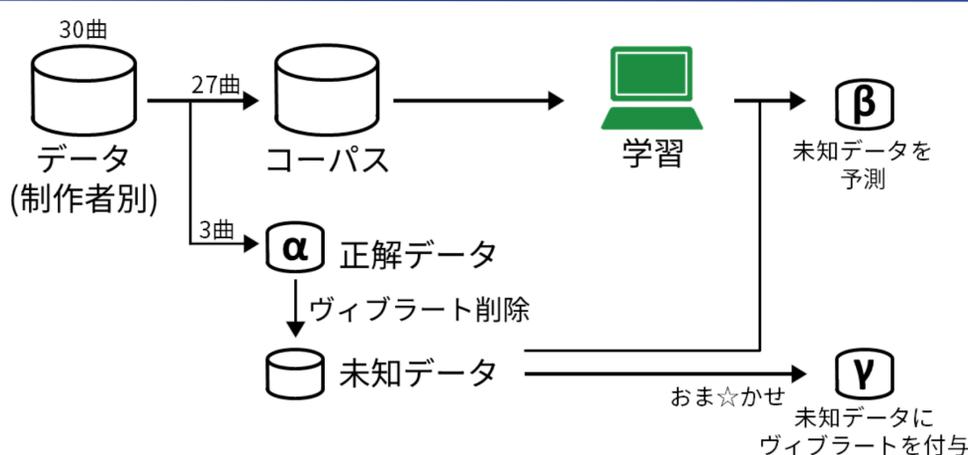
提案モデル



未知の楽譜へヴィブラートを付与する手順

- ①楽譜を「ヴィブラートの有無を判別するLSTM」に入力し、ヴィブラートの出現箇所を予測する
- ②ヴィブラートが出現すると判別された場合、「ヴィブラートのパラメータを推定するLSTM」に同じ楽譜を入力し、割合や周波数などの細かいパラメータの推定を行う

提案モデルの比較実験

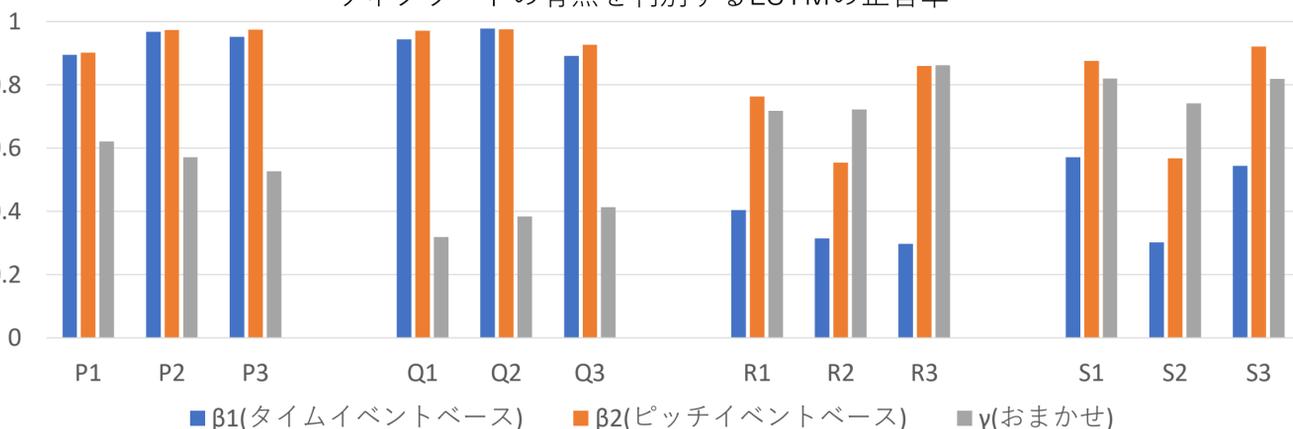


実験手順

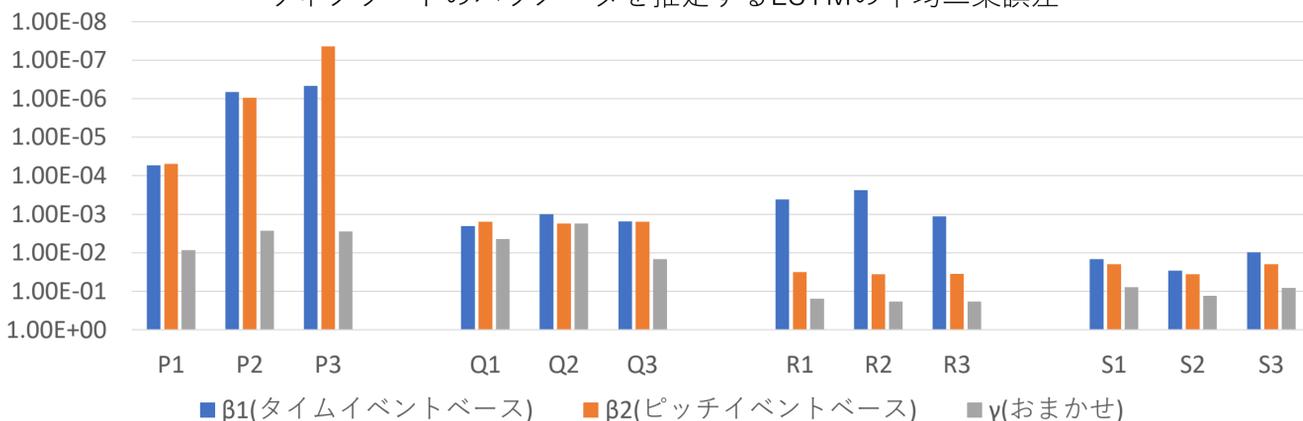
- ①制作者別に30曲のヴィブラートが付与された楽譜データを収集。
うち27曲をコーパス, 残り3曲を比較実験用の楽譜データとする。
 - ②比較実験用の楽譜データ(以下 α)からヴィブラートを削除した未知データを用意。
 - ③コーパスからヴィブラートの出現箇所とパラメータをそれぞれ学習。
学習結果から未知データのヴィブラートの出現箇所またはパラメータを付与。(以下 β)
 - ④従来手法「おま☆かせ」を用いて, 未知データにヴィブラートを付与。(以下 γ)
 - ⑤ α と β , α と γ を比較し, ヴィブラートの出現箇所の正答率, パラメータの平均二乗誤差を算出。
- ※今回は, P, Q, R, Sの4人のパラメータ制作者の楽曲を各3曲比較。(P1, P2, P3...と表記)

提案モデルの比較実験結果

ヴィブラートの有無を判別するLSTMの正答率



ヴィブラートのパラメータを推定するLSTMの平均二乗誤差



- ・制作者PとQの楽曲はヴィブラートの出現が少なく, RとSは多い傾向がある。
- ・おまかせは条件に当てはまる音全てにヴィブラートを付与するため, ヴィブラートの出現数が多いRとSの正答率が高くなっている。
- ・タイムイベントベースとピッチイベントベースを比較した際, ピッチイベントベースの方が位置推定の精度は高い。
- ・しかし, ピッチイベントベースの方が各パラメータの誤差は大きくなっている。

今後の展望

- ①入力データを変更した際の予測精度の比較を行う。
- ②Attention機構を用いて学習の精度を上昇させる。
- ③ポルタメント機能によって生成されたピッチ線を学習し, ピッチ線の付与を行う。
(下図の赤線)

