

歌声の歌唱タイミングと同期する伴奏システム

野口真太郎¹ 堀内靖雄¹ 黒岩眞吾¹

概要：音楽情報処理の一分野として伴奏システムが研究されている。伴奏システムとは人間が行う伴奏のように、独奏者に協調した伴奏を計算機で実現しようとするシステムである。従来の伴奏システムは主に楽器の独奏に対し、ピアノなどで伴奏を行うものであったが、歌声を独奏とする伴奏システムも開発されてきている。フルートなどの楽器では各音の演奏タイミングはその発音開始時刻と考えられ、パワーや音高情報などから容易に検出可能であるが、歌詞を伴う歌声の場合その演奏タイミングは歌詞の音素により複雑に変化するため検出が困難である。先行研究において歌詞が子音+母音の場合は母音の開始時刻が歌唱者の意図する演奏タイミングと一致することが示されている。従来の楽器を対象とする伴奏システムをそのまま歌唱に適用した場合、子音の開始時刻をその発音開始時刻と判断してしまうため、独奏者の意図する演奏タイミングとずれてしまう。そこで本研究では母音の開始時刻をリアルタイムで検出する手法を提案し、その検出時刻に基づいて伴奏制御することにより歌声と正確に同期する伴奏システムの実現を目的とする。提案システムを実装し、従来システムとの比較評価実験を行った結果、提案手法の有用性が示された。

キーワード：伴奏システム, 歌声, 歌唱タイミング

1. はじめに

伴奏システムとは、独奏パートと伴奏パートの楽譜情報を事前に計算機に与えることで、人間同士で行われる合奏を人間と計算機によって実現するものであり、人間の独奏者の演奏に協調した伴奏を行うシステムである[1][2][3][4][5][6][7][8]。

従来の伴奏システムはピアノやフルート等の楽器を独奏の対象とするものが多かったが、歌声を独奏とする伴奏システムについても検討・開発されてきている。古くは[9][10]などで歌声に合わせる伴奏システムが提案されているが、当時の計算能力の限界から、リアルタイムで演奏追跡を行うためのピッチ検出や母音認識に焦点が当てられていた。最近の研究として[11]ではアカペラで欠員パートが出た場合を想定し、従来の伴奏システムでは独奏の入力が1人の歌声のところ、複数人の歌声を入力として用いることで欠員パートを補完するシステムを実現するための楽譜追跡手法が検討されている。[12]では既存の楽曲の音声ファイルからリアルタイムで独奏音と伴奏音を分離し、分離して得られた情報を用いてユーザの歌唱に追従するカラオケシステムが提案された。[13]ではアカペラにおけるハーモニーの美しさ(和音の響きの調和性)に着目し、独奏者の歌声をリアルタイムで認識し、その歌声に調和するように伴奏の音高を調整し出力するシステムが提案された。

ここで歌声に関する先行研究[14][15]において歌詞が子音+母音の音節を歌唱した場合、子音の開始時刻は音符の時刻より先行する傾向があることが示されている。これは人が歌唱をする際、一般的に母音を音符のタイミングに合わせて歌唱するためであると考えられる。そのため従来の楽器を対象とする伴奏システムをそのまま歌唱に適用した場合、子音の開始時刻をその発音開始時刻と判断してしまうため、独奏者の意図する演奏タイミングとずれてしまう

問題が生じる可能性がある。従来の歌声伴奏システム研究では子音+母音の歌唱において、母音の開始タイミングを検出して、そのタイミングに合わせて伴奏を制御する伴奏システムはまだ研究されていない。

そこで本研究では歌詞情報に依存せずに母音の開始時刻をリアルタイムで検出する手法を提案し、その検出時刻に基づいて伴奏制御をすることで歌声と正確に同期する伴奏システムの実現を目的とする。

2. 歌唱時のスペクトル包絡の分析

予備分析として歌唱時の音声の分析を行い、伴奏システムに利用可能なパラメータの検討を行う。

2.1 分析データ

分析データには合唱歴9年の男声歌唱者1名が歌唱した高野辰之作詞、岡野貞一作曲の「故郷」を用いた。分析データは予め録音したもので、サンプリング周波数44.1kHz、量子化ビット数16bitのモノラル音声である。

2.2 母音遷移時のスペクトル包絡の変化

まず始めに歌声の母音遷移時のスペクトル包絡を分析した。スペクトル包絡を分析した理由は歌唱時に音高の影響を受けずに安定した特徴を表す母音の特徴を調べるためである。スペクトル包絡の計算はケプストラム分析法を用いた。分析時のFFT窓幅は1024サンプル、スライド幅は441サンプル、リフタリング次数は60次元である。

2.2.1 歌唱時と話し声のスペクトル包絡の比較

まず始めに歌唱時と話し声のスペクトル包絡の比較を行う。

歌声のスペクトル包絡を図1に示す。図1では「うさぎ」と歌唱した際の音声のスペクトル包絡の時間変化を表している。また、話し声のスペクトル包絡を図2に示す。図2は歌唱を意識せずに「うさぎ」と発声した際の音声の

¹ 千葉大学

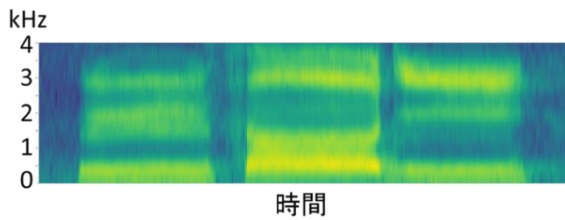


図 1 歌声のスペクトル包絡

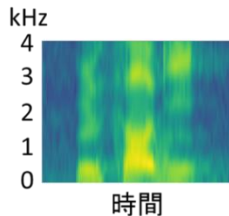


図 2 話し声のスペクトル包絡

スペクトル包絡の時間変化を表している。図 1, 図 2 のグラフの横軸は時間, 縦軸は周波数を表す。グラフは色が明るいほどその周波数のパワーが強いことを表す。

図 1, 図 2 を比較してわかる通り, 歌唱時の母音区間のフォルマントは話し声のフォルマントに比べ, 時間軸方向に長時間にわたって安定していることがわかる。また, 歌唱時のフォルマントは 0~4kHz の範囲だと第一フォルマント (F1), 第二フォルマント (F2), 歌唱フォルマントの 3 つが確認できるが, F1 のパワーが強く安定していることがわかる。そのため本研究では独奏者の母音検出に用いる特徴量として F1 付近のスペクトル包絡の安定性に着目することにした。

2.2.2 無声子音から母音への遷移

続いて無声子音から母音へ遷移する際のスペクトル包絡を分析した。歌唱中の無声子音から母音への遷移時のスペクトル包絡を図 3 に示す。図 3 では歌詞「ふるさと」の「るさ」のスペクトル包絡の時間変化を表している。

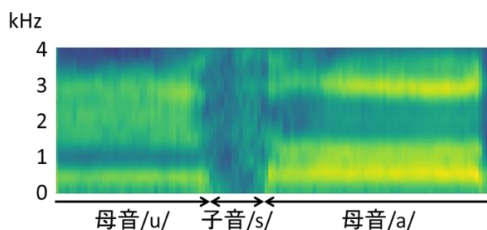


図 3 無声子音→母音のスペクトル包絡

図 3 より, 無声子音/s/の区間ではスペクトル包絡がノイズのようになっておりフォルマントが現れていないのに対し, 母音/a/の区間ではフォルマントが安定してはっきりと現れている。また, 現れているフォルマントの中では F1 が最もパワーが強く安定していることがわかる。したがって, 無声子音から母音への遷移は F1 が現れる周波数帯域のスペクトル包絡の特徴量で検出が可能であると考えられる。

2.2.3 有声子音から母音への遷移

最後に有声子音から母音へ遷移する際のスペクトル包絡を分析した。歌唱中の有声子音から母音への遷移時のスペクトル包絡を図 4 に示す。図 4 では歌詞「わすれがたき」の「れが」の区間のスペクトル包絡の時間変化を表している。

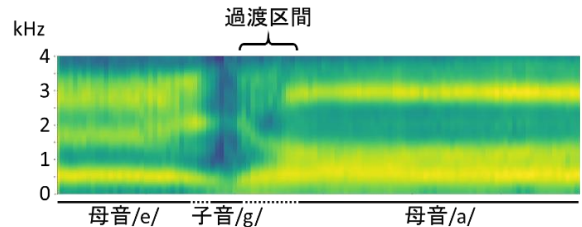


図 4 有声子音→母音のスペクトル包絡

図 4 より, 有声子音/g/の閉鎖区間で一時的にフォルマントが現れなくなるが, 直後に「バースト」が起こり, その後およそ十数フレームかけてフォルマントが母音/a/のフォルマントに向けて連続的に変化し, 最終的に時間軸方向に安定することが確認できる。このように有声子音から母音へ遷移する際は過渡区間が存在するため, 母音の開始時刻を厳密に決定することは難しい。また, 仮に母音の開始時刻が決定できたとしても, その母音開始時刻と歌唱者が音符の演奏タイミングと感じる時刻は完全に一致しているとは限らず, 歌唱者の個人差により変化する可能性が考えられる。そのため, 母音の開始時刻の決定には個人差を吸収する仕組みが必要であると考えられる。

3. 提案システム

本研究では独奏に歌声を用いることを想定し, 歌詞情報を用いずに母音の開始時刻を検出し, 得られた検出時刻を用いて歌声と同期するシステムを提案する。

3.1 音響特徴量の抽出

本研究では予備分析を基に, 母音遷移後のフォルマントの時間軸方向における安定性に着目し, 母音の開始時刻検出にスペクトル包絡の時間差分を用いることとした。

当研究室の伴奏システムはリアルタイムで独奏者の楽譜位置推定を行い, 伴奏を出力する。具体的には約 10ms 毎に独奏の最新 1024 サンプルの入力音声を用いて音響特徴量の計算等をリアルタイムで行うため, スペクトル包絡の計算にも計算量が少ない手法が求められる。そこで提案手法では比較的計算量が少ないケプストラム分析法を用いた。提案手法では 1024 サンプルの入力音声 (サンプリング周波数 44.1kHz, 量子化ビット数 16bit) に対し, リフタリングの次数は 60 とした。ケプストラム分析で得られるスペクトル包絡 $P_n(\omega)$ は対数値とする。

また, 音素等の違いによるパワー変動の影響を抑制す

るため、先ほど計算したスペクトル包絡 $P_n(\omega)$ (負の値)を式(1)で0~1になるように正規化を行った。

$$\hat{P}_n(\omega) = 1 - \frac{P_n(\omega)}{\sum_{a=1}^{1024} P_n(a)} \quad (1)$$

続いて正規化を行ったスペクトル包絡 $\hat{P}_n(\omega)$ を用いて、式(2)で前後2フレーム間のスペクトル包絡の時間差分を計算した。スペクトル包絡の差分は二乗平均平方根(RMS)と定義した。第2章で行った分析の結果より、母音のF1付近の変動を見ることで音素の遷移を検出できると判断したため、差分の計算にはF1が現れる範囲のスペクトル包絡の成分を用いることとする。先行研究[16][17]では日本語母音のF1は200~1000Hzの範囲で分布することが示されているため、差分の計算にはスペクトル包絡の0~1000Hzの成分を用いることとした。前述の通り入力音声のサンプリング周波数は44.1kHz、窓幅1024サンプルのため、計算されたスペクトル包絡 $\hat{P}_n(\omega)$ で0~1000Hzに対応する成分は先頭から25サンプル分になる。式(2)ではこの25サンプルを用いて計算を行う。

$$\Delta P_n = \sqrt{\frac{1}{25} (\hat{P}_n(m) - \hat{P}_{n-1}(m))^2} \quad (2)$$

提案手法ではフレーム毎に計算されるスペクトル包絡の時間差分 ΔP_n を母音開始時刻の検出に用いる。

3.2 母音の開始時刻の検出

続いて3.1で定義したスペクトル包絡の時間差分 ΔP_n を用いて母音の開始時刻を定義する。

提案手法では母音のスペクトル包絡が長時間にわたって安定している点に着目し、 ΔP_n が閾値 α を下回った時刻 t を起点に、一定時間連続して閾値 α 未満となった時点で母音の発音を確定し、起点の時刻 t を母音の開始時刻と定義した。伴奏システムの実装において閾値 α は0.00002とし、閾値を下回ってから開始時刻を確定させるまでの時間を50msとしている。

3.3 キャリブレーション

3.2において母音のスペクトル包絡の安定性を基準として母音の開始時刻を定義したが、実際には子音から母音への遷移には過渡区間が存在するため母音の開始時刻を厳密に定義することは難しく、また、母音の開始時刻と歌唱者の意図する演奏タイミングも個人差の影響を受ける可能性がある。そのため本手法ではキャリブレーションで個人差による誤差を吸収することにした。

キャリブレーションは伴奏システムの伴奏のテンポを固定した状態で行う。独奏には単純なメロディーの楽曲を用意し、メロディーに合わせて数種類の日本語音韻を発声しながらピアノ伴奏に合わせて合奏を行う。この合奏によって得られた伴奏システムのログを基に、独奏者の各音と同期する伴奏音に該当するMIDIのノートオン信号を送信した時刻と独奏音の検出時刻の誤差を計算する。この際、

独奏音の検出時刻は3.2で述べた手法によって得られた母音の開始時刻を用いる。次に各音の独奏の検出時刻と伴奏のノートオン時刻の誤差分布の中央値を求め、この値を母音の開始時刻の補正時間として用いる。

キャリブレーション後に行う実際の合奏時には、3.2で得られた母音遷移の時刻に対して、キャリブレーションで得られた値に基づいて補正した時刻を歌唱者の発音開始時刻としてテンポ計算を行う。テンポ計算については先行研究[18]の伴奏制御手法を用いた。

4. 実験

本章では提案手法を実装した伴奏システム(以下、提案システム)の有用性を評価するために行った実験について述べる。実験では比較のため、従来手法による独奏のパワーの立ち上がり時刻を発声開始時刻と定義する伴奏システム(以下、従来システム)を比較用システムとした。

4.1 実験環境

提案システムはIntel® Core™ i7-4500U CPU @ 1.80GHz 2.40GHz(8.00GB RAM)のPC上で実行した。

コンピュータ以外にMIDI音源としてRoland INTEGRA-7、独奏の音響情報を取得するマイクとしてAKG C414XLSコンデンサーマイク、オーディオインターフェースはRoland Quad-Capture(UA-55)を用いた。

4.2 キャリブレーション

キャリブレーションには図5のキャリブレーション用に独自に作曲した楽曲を用いた。上段が独奏パート、下段が伴奏パートである。歌唱者には独奏パートの各音符を指定された歌詞(子音+母音の音韻)で歌唱してもらった。反復記号は発声する音韻の回数に合わせて反復し、楽曲の途中で指定した音韻を全て発声し終わった場合はその時点で曲を終了させた。



図5 キャリブレーション用楽譜

4.3 定量評価実験

本研究で提案した検出手法が子音と母音を区別して正しく検出されるかどうかを確認するため、実験参加者に提案システムと従来システムでそれぞれ合奏をしてもらい、各システムより得られたログデータを用いてシステムの動作を定量的に評価した。本実験では提案手法の有効性を示すため、比較的子音長が長い音韻「な」に歌詞を固定して実験を行った。

4.3.1 実験手順

最初にキャリブレーションを行い、実験参加者ごとに補正量を決定した。発声した音韻は「な」のみで、音符に合わせて合計 40 回発声してもらった。

続いて、実験参加者には図 6 で示す実験用楽曲①を用いて伴奏システムとの合奏を 4 回 (=システム 2 種類×2 回) 行ってもらった。楽譜は上段が独奏、下段が伴奏であり、2 小節目以降の 1 拍目と 3 拍目を拍位置とし、これらの箇所伴奏のテンポを再計算するように設定した。また、楽譜上の反復箇所は 8 回演奏した。実験中、実験参加者には 2 種類のシステムを使用したことは伝えておらず、伴奏に合わせて歌唱するように教示を行った。歌唱誤りや認識誤りによって伴奏が独奏と比較して極端に加速もしくは減速した場合は妥当な評価が困難であるため、合奏のやり直しを行った。



図 6 実験用楽曲①

合奏を行う順番は「提案→従来→提案→従来」と「従来→提案→従来→提案」のどちらかで行った。前者をパターン 1、後者をパターン 2 とし、実験参加者の情報と演奏順パターンの組み合わせを表 1 に示す。

表 1 実験参加者の情報と演奏順パターン

実験参加者	性別	歌唱歴	演奏順パターン
A	女性	3 年 9 ヶ月	1
B	女性	3 年 9 ヶ月	2
C	男性	9 ヶ月	1
D	男性	9 ヶ月	2
E	男性	9 年 9 ヶ月	1

4.3.2 結果

実験参加者毎の補正量は A が 30.0ms, B が 88.0ms, C が 110.0ms, D が 130.0ms, E が 79.5ms であった。

続いて実験参加者毎に各拍のテンポ比率の平均値をシステム別に算出した。テンポ比率とは楽曲の初期テンポを 1.0 とした時のテンポ変化の比率を表した値である。図 7 に各実験参加者のテンポ比率の推移を示す。横軸は拍番号、縦軸はテンポ比率である。

4.3.3 考察

子音+母音の音韻の場合、提案手法は母音の開始時刻を検出する手法であるが、従来手法は子音の開始時刻を検出するため、従来手法の検出時刻は提案手法の検出時刻よりも早くなる。すなわち歌唱者が母音の開始時刻を伴奏に

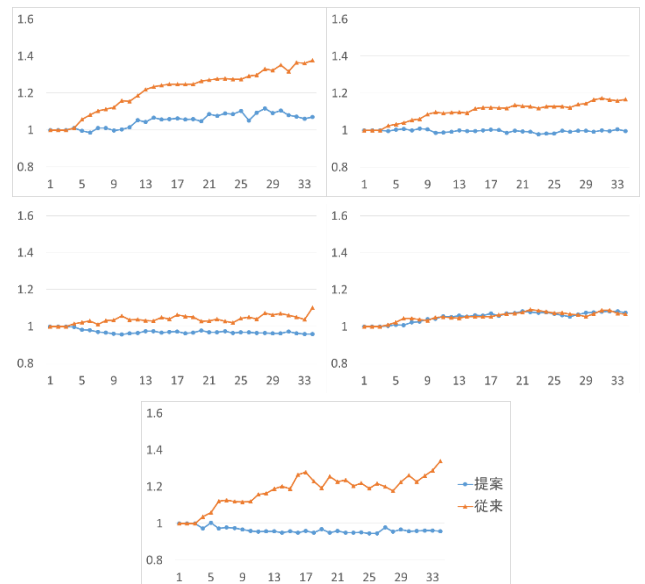


図 7 テンポ比率の推移 (上段左が A, 上段右が B, 中段左が C, 中段右が D, 下段が E)

合わせようと歌唱していると仮定した場合、従来手法の検出時刻を用いて計算されたテンポは伴奏に比べ相対的に速くなるはずである。A, B, E のテンポ比率の変化は上記の予想通りの結果となったが、C と D は何らかの要因によって差が現れなかったと考えられるため、実験時に得られたログデータから、実験結果の差について考察を行う。便宜上、A, B, E をグループ α 、C と D をグループ β と呼称する。

まずグループ α の分析として A の合奏時のログデータの考察を行う。図 8 に A が提案および従来システムで合奏した際の独奏のスペクトル包絡と各システムが独奏音を検出した時刻、独奏者の各音と同期する伴奏音に該当する MIDI のノートオン信号を送信した時刻を示す。スペクトル包絡は 0~4kHz の帯域を表示している。図 8 より、伴奏のノートオン時刻はどちらのシステムでも独奏音の子音から母音の過渡区間にあることがわかる。しかし独奏の検出時刻は、提案システムは伴奏のノートオン時刻と概ね一致するのに対し、従来システムでは子音/n/のフォルマントが現れ始めた時刻と一致することがわかる。同様の傾向は B と E でもみられたことから、この検出時刻の差がグループ α のテンポ比率の推移の差に表れたのではないかと考えられる。

次にグループ β の分析として C の合奏時のログデータの考察を行う。図 9 に C が提案および従来システムで合奏した際のログデータの一部を示す。図 9 より、C は有声子音/n/の子音長が短いため子音と母音の開始時刻がほぼ一致しており、各システムの独奏音の検出時刻もほとんど差がないことがわかる。このことからグループ β の両システムのテンポ比率の推移にあまり差が現れなかったと考えられる。

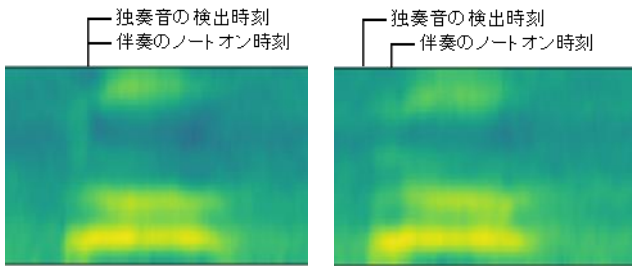


図 8 A のスペクトル包絡 (左: 提案, 右: 従来)

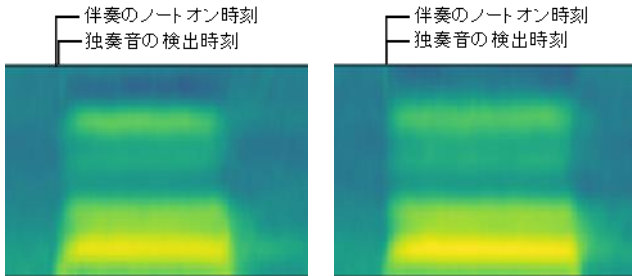


図 9 C のスペクトル包絡 (左: 提案, 右: 従来)

以上の考察より、提案手法は子音長が長い場合は母音の開始時刻の検出が有効であると考えられる。一方、子音長が短い場合は子音の開始時刻は母音の開始時刻とほぼ一致するため、従来手法でも提案手法とそれほど遜色のない合奏制御が行えると考えられる。

4.4 主観評価実験

実験参加者に実際の楽曲で提案システムと従来システムに対して主観評価実験を行った。

4.4.1 実験手順

始めにキャリブレーションを行い、実験参加者ごとに補正量を決定した。発声した音韻は子音の平均的な補正値を推定するため、母音単音や半母音を含まない日本語音韻 60 種類で、音符に合わせて合計 240 回 (=音韻 60 種類×4 回) 発声してもらった。

続いて、実験参加者には実験用楽曲②(高野辰之作詞、岡野貞一作曲「故郷」)を用いて伴奏システムとの合奏を 4 回 (=システム 2 種類×2 回) 行ってもらった。歌詞は「故郷」の一番を、独奏のテンポは一定で歌唱するように指示した。各合奏終了後、主観評価として以下の 5 段階評価の質問に回答してもらった。この際、実験参加者には 2 種類のシステムを使用したことは伝えていない。また、歌唱誤りや認識誤りによって伴奏が独奏と比較して極端に加速もしくは減速した場合は妥当な評価が困難であるため、合奏のやり直しを行った。全ての合奏が終了後、実験参加者から自由回答をもらった。なお、合奏を行う順番は定量評価実験の表 1 と同じ組み合わせを用いた。

また、B と E については歌詞を「な」に統一して、同様の実験を行った。この歌詞を「な」に統一した実験に限り、補正量は定量評価実験で得られた値を用いた。

Q. 自分の歌唱のテンポに伴奏は合っていたと思うか
(5 とてもそう思う 4 そう思う 3 どちらとも言えない 2 そう思わない 1 とてもそう思わない)

4.4.2 結果

検出時間の補正量と合奏後に得られた主観評価のシステム毎の平均値(通常歌詞と音韻統一)を表 2 に示す。

表 2 補正量と主観評価の平均値

	通常歌詞			音韻統一		
	補正 (ms)	主観評価		補正 (ms)	主観評価	
		提案	従来		提案	従来
A	110.0	2.5	4.5	/	/	/
B	98.0	3.0	4.0	88.0	5.0	3.0
C	121.0	3.0	5.0	/	/	/
D	137.0	1.5	3.5	/	/	/
E	92.0	3.5	3.5	79.5	4.5	2.5
全員	/	2.7	4.1	/	4.8	2.8

4.4.3 考察

本研究の提案手法は子音が拍より先行して発声される場合を想定して提案しており、定量評価実験の結果、グループαは想定された歌い方をしていたため、本項ではその中でも A の結果を中心に考察を行う。

A からは自由回答で曲の後半からずれていく感じがしたとコメントをもらった。そこで A が提案システムで「故郷」を合奏した際のテンポ比率の推移(2 回)を図 10 に示す。図 10 から、曲の折り返し地点付近の 32 拍目で急激にテンポが上昇していることがわかる。32 拍目は歌詞「かのかわ」の「の」の音符に対応する拍である。図 11 に A が該当箇所を提案システムで合奏した際の独奏のスペクトル包絡を示す。図 11 より、一つ前の音の母音/a/のフォルマントが現れているときに独奏音の検出が起きており、この検出タイミングのずれによってテンポ比率が大きくなった可能性が考えられる。ここで補正量に着目すると、主観評価実験で様々な種類の子音でキャリブレーションした時の補正量が 110.0ms であったのに対し、定量評価実験で子音/n/のみでキャリブレーションした時の補正量は 30.0ms であった。仮に図 11 の検出時刻の補正量が 30.0ms であった場合、検出時刻は 80.0ms 遅くなるため、概ね拍の時刻と一致する。そのため、この検出位置誤りは子音を区別せずに補正量を設定したことに問題があったと考えられる。

続いて歌詞を「な」に統一した時の主観評価について分析を行う。歌詞を統一した合奏において、B からは提案システムでの合奏は自分の歌に合っていたとコメントをもらった。ここで B が歌詞を「な」に統一して各システム

で「故郷」を合奏した際のテンポ比率（各2回）の平均値の推移を図12に示す。図12より、提案システムは従来システムと比較して終始テンポが安定していることがわかる。これは設定された補正量が最適であったためであると考えられる。これらの結果から、提案手法で推定された時刻と人間が演奏タイミングだと感じる時刻は子音によって大きなばらつきがあると推測され、子音毎に適切な補正量を設定する必要があると考えられる。

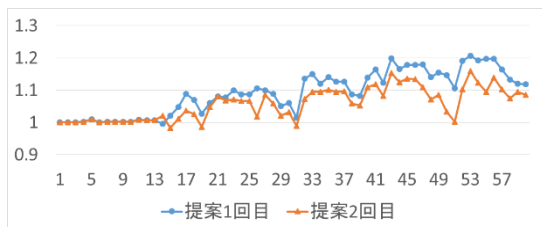


図10 Aのテンポ比率の推移（故郷，通常歌詞）

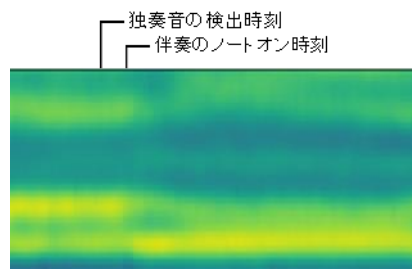


図11 Aのログとスペクトル包絡（提案）

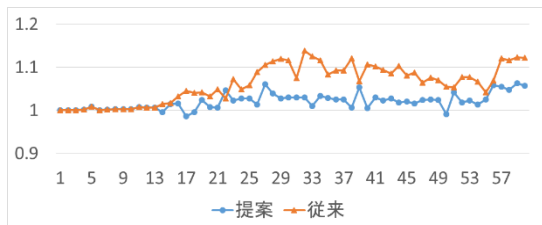


図12 Bのテンポ比率の推移（故郷，音韻統一）

5. まとめと今後の課題

本研究では楽器と歌声の音の演奏タイミングの違いに着目し、独奏のスペクトル包絡を利用してリアルタイムで独奏の母音開始時刻を検出する手法を提案し、歌声と同期する伴奏システムを実現した。定量評価実験の結果、提案手法の有効性を示すことができた。実際の楽曲を用いた主観評価実験では通常歌詞での歌唱は十分な結果を得ることができなかったが、音韻を統一歌唱した場合は提案システムの評価が高かった。これらの結果より補正量は子音毎に適切な値を設定する必要があることが示唆された。

今後の課題として、子音毎に適切な補正量を設定しなければならぬ。具体的には事前に歌詞情報をシステムに与えることにより、音韻ごとに補正量を調整する手法が考えられるが、演奏者の歌唱間違いに対しても適用できるよ

うにするため、伴奏システムに入力する歌声の音声認識を行い、システムが子音の種類に応じて補正量を切り替える手法を検討したい。また本研究では子音+母音の音韻に着目をしたが、今後は母音が連続する場合や半母音を伴う音韻などの場合、過渡区間の変動が緩やかであり、音の開始時刻の決定が困難であるが、そのような状況にも対応できる手法を検討したいと考えている。

謝辞 本研究は科研費17H00749の支援を受けています。

参考文献

- [1] Dannenberg, R.: An On-Line Algorithm for Real-Time Accompaniment, Proc. of ICMC, pp. 193-198 (1984).
- [2] Vercoe, B.: The Synthetic Performer in the Context of Live Performance, Proc. of ICMC, pp. 199-200 (1984).
- [3] Horiuchi, Y., Hozumi, T.: A Computer Accompaniment System With Independence, Proc. of ICMC, pp.418-420 (1993).
- [4] Cont, A.: ANTESCOFO: Anticipatory Synchronization and Control of Interactive Parameters in Computer Music, Proc. of ICMC (2008).
- [5] Raphael, C.: Music Plus One and Machine Learning, Proc. of ICML, pp. 21-28 (2010).
- [6] 中村栄太, 武田晴登, 山本龍一, 齋藤康之, 酒向慎司, 嵯峨山茂樹: 任意箇所への弾き直し・弾き飛ばしを含む演奏に追従可能な楽譜追跡と自動伴奏, 情処学論, Vol. 54, No. 4, pp. 1338-1349 (2013).
- [7] Wada, S., Horiuchi, Y., Kuroiwa, S.: Tempo prediction model for accompaniment system, Proc. of ICMC 2014, pp.1298-1303 (2014).
- [8] 前澤陽. 自動合奏のためのタイミング結合モデル. 研究報告音楽情報科学(MUS), 2016-MUS-112(19), pp. 1-8 (2016).
- [9] Katayose, H., Kanamori, K., Kamei, K., Nagashima, Y., Sato, K., Inokuchi, S., Simura, S.: Virtual Performer, Proc. of ICMC, pp.138-145 (1993).
- [10] 井上渉, 橋本周司, 大照完. 適応型歌声自動伴奏システム. 情処学論, 37(1), pp.31-38 (1996).
- [11] 森大毅, 上田新. アカペラ演奏支援のための歌声に対する楽譜追跡手法の検討. 研究報告音楽情報科学(MUS), 2015-MUS-107(62), pp.1-6 (2015).
- [12] 和田雄介, 坂東宜昭, 中村栄太, 糸山克寿, 吉井和佳. 楽曲中の歌声とユーザ歌唱のリアルタイムアラインメントに基づく伴奏追従型カラオケシステム. 研究報告音楽情報科学(MUS) 2017-MUS-116(3), pp.1-7 (2017).
- [13] 吉本侑平, 堀内靖雄, 黒岩眞吾. 和音の響きの調和性を考慮した歌声に対する伴奏システム. 研究報告音楽情報科学(MUS) 2019-MUS-122(23), pp.1-6 (2019).
- [14] 剣持秀紀. 音楽情報処理技術の最前線:3.歌声合成とその応用. 情報処理 50(8), 723-728 (2009).
- [15] 的場達矢, 馬場隆, 成山隆一, 松本秀一, 森勢将雅, 片寄晴弘. 歌唱のグルーブ感の構成要因の分析. 研究報告音楽情報科学(MUS) 2014-MUS-102(12), pp.1-5 (2014).
- [16] 益子幸江. 日本語の母音の音色とフォルマントについての一研究. 東京外国語大学論集(82), pp.105-121 (2011).
- [17] 遠藤希美, 川原繁人, 皆川泰代. 声楽的発声における母音知覚: 声楽経験および基本周波数の影響. 音声研究 21(2), pp.25-37 (2017).
- [18] 足立亜里紗, 堀内靖雄, 黒岩眞吾. 独奏認識誤りに頑健な音響入力伴奏システム. 研究報告音楽情報科学(MUS), 2017-MUS-114(1), pp.1-5 (2017).