

人間の伴奏制御モデルに基づく伴奏システムの実装と評価

矢島 直人[†] 堀内 靖雄[†] 西田 昌史[†] 市川 薫[†]

[†] 千葉大学大学院自然科学研究科

〒263-8522 千葉市稲毛区弥生町 1-33

E-mail: [†] yajiman@graduate.chiba-u.jp, {hory, nishida, ichikawa}@faculty.chiba-u.jp

あらまし 本稿では、人間同士の合奏データから推定した演奏制御モデルを実装した伴奏システムについて述べる。制御モデルは「独奏者と伴奏者間の時間的ずれ」の履歴と「伴奏者のテンポ変化量」の履歴から伴奏者の未来の演奏時刻を決定する。しかしながら、このモデルは8分音符だけからなる楽曲から推定されているため、実際の曲に対して適用する場合には音が演奏されない部分についても考慮しなければならない。そこで、本研究では、人間が演奏時に自身の内部で利用しているであろう内的クロックを推定して、伴奏制御に用いる手法を提案する。人間の演奏者と提案システムによる合奏実験を行った結果、内的クロックの推定により、提案手法に基づく伴奏制御が可能であることが確認された。

Accompaniment System Based on the Model Estimated from Human Ensemble

Naoto YAJIMA[†], Yasuo HORIUCHI[†], Masafumi NISHIDA[†], and Akira ICHIKAWA[†]

[†] Graduate School of Science and Technology, Chiba University

1-33 Yayoi-cho, Inage-ku, Chiba, 263-8522 Japan

E-mail: [†] yajiman@graduate.chiba-u.jp, {hory, nishida, ichikawa}@faculty.chiba-u.jp

Abstract In this paper, we will introduce an accompaniment system that controls its own performance based on the model estimated from the data of human ensemble. The model can decide its future performance from the history of the time difference between the soloist and the accompanist and the history of the tempo modification of the accompanist. In order to apply this model to the actual scores, we have to estimate the inner clock of the soloist when the note does not exist in the score. So we introduced the estimation method of the inner clock using the notes around the missing position. We performed the experiment where human performers play with the introduced accompaniment system. As a result, we observed that the system can follow the tempo modification by the soloist. Moreover, it was confirmed that the estimation method of the inner clock worked effectively.

1. はじめに

現在、多くの分野において人間とコンピュータとのインタラクションを実現しようとする研究が盛んであり、我々は音楽を通じてリアルタイム・インタラクションを実現する手段として、伴奏システムに注目している。伴奏システムとは、人間を独奏者、コンピュータを伴奏者として合奏を行なうシステムである。従来のシステムは、主に独奏者の演奏ミスへの対応などを重視していたため、コンピュータが生成する伴奏の質が考慮されることはあまりなく、機械的で不自然な演奏となることが多かった。そこで我々は人間同士の合奏を分析し、その結果から伴奏者の演奏制御モデルを

推定する手法を提案した[4]。しかし、このモデルは8分音符のみの曲から推定されたものであるため、システムに実装する際、一般の曲にはそのままでは適用できない。そこで本研究では、上記モデルをコンピュータ上に実装するため、後述する内的クロックの概念を導入することによって、8分音符だけの曲ではなく、一般的な曲にも適用可能とする手法を提案する。また、提案手法を実装した伴奏システムと人間の独奏者との合奏を行ない、本手法の有効性を評価する。

2. 伴奏システム

伴奏システムは、あらかじめ与えられた楽譜情報を

表 1 伴奏者の演奏制御モデルにおけるパラメータの偏回帰係数, 重相関係数 R , 予測誤差 $E[\text{ms}]$ (8 分音符単位)

α_1	α_2	α_3	α_4	α_5	α_6	α_7	α_8		
0.50	—	-0.25	—	-0.04	—	—	-0.04		
β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8	R	E
-0.56	—	-0.02	—	-0.03	—	—	0.13	0.77	10.5

表 2 伴奏者の演奏制御モデルにおけるパラメータの偏回帰係数, 重相関係数 R , 予測誤差 $E[\text{ms}]$ (4 分音符単位)

α_1	α_2	α_3	α_4	α_5	α_6	α_7	α_8		
—	0.76	—	-0.05	—	-0.22	—	-0.21		
β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8	R	E
—	-0.45	—	-0.06	—	0.08	—	0.06	0.77	12.5

もとに, 入力された独奏者の演奏内容と協調するように伴奏の発音タイミングをリアルタイムで制御するインタラクティブ・システムである [1].

伴奏システムの研究は 1980 年代より始まったが, 従来のシステムでは主に独奏者の演奏ミスへの対応などを重視していた [2][3]. そのため, コンピュータが生成する伴奏の質はあまり考慮されず, 出力される演奏は機械的で不自然なものとなることが多かった.

そこで我々は人間同士の合奏は独奏者と伴奏者が互いに影響を与えあっていると考え, 人間らしい自然な演奏を生成するために, 実際に人間同士の合奏を収録, 分析することにより, 人間の合奏における伴奏者の演奏制御モデルを推定した [4]. その結果, モデルのパラメータとして, それぞれ過去 1 小節分 (8 分音符 8 つ分) の「独奏者と伴奏者の間の時間的ずれ」と「伴奏者のテンポ変化量」が有効であることがわかった. しかし, 推定に用いるパラメータを 8 分音符単位で取得しているため, 一般の楽譜のように, 独奏パートに休符や 8 分音符よりも長い時間長をもつ音符がある場合, その部分のパラメータが欠落してしまい, そのままではモデルを適用することができない.

本研究では, 拍を演奏制御の単位として, 拍に相当する箇所が発音しない状況においては, 過去の演奏履歴から発音時刻を推定する手法を提案する. 人間が演奏をする際には, 拍に相当する箇所において自身の内面にあるクロック情報 (内的クロック) を参照してテンポを保持していると考えられる [5][6][7]. また, 合奏においては, 独奏者と伴奏者が互いに自らの内的クロックにしたがって演奏を進めながら相手の内的クロックを推定し, 相手の演奏と協調するように自らの内的クロックを調整していると考えられる.

3. 伴奏者の演奏制御モデル

過去 1 小節分 (8 分音符 8 つ分) の「独奏者と伴奏者の間の時間的ずれ」と「伴奏者のテンポ変化量」を説明変数として, 次の「伴奏者のテンポ変化量」を予測するモデルは, 重回帰分析により, 式(1)のように表せる.

$$c_n = \sum_{i=1}^8 \alpha_i d_{n-i} \sum_{j=1}^8 \beta_j c_{n-j} \quad (1)$$

推定に用いるパラメータを 8 分音符単位で取得した場合, 表拍と裏拍に相当する音の間に高い相関が見られ, 多重共線性の影響が生じていたため, ステップワイズ法 (変数減増法) によって説明変数を選択している. 分析の結果得られた偏回帰係数と重相関係数, 予測誤差を表 1 に示す.

今回, 内的クロックを推定するにあたって, 制御の単位を拍単位で行なう必要がある. そこで, 推定に用いる説明変数を 4 分音符単位 (拍単位) とし, 重回帰分析した結果得られた偏回帰係数と重相関係数, 予測誤差を表 2 に示す. 4 分音符単位で選択した方が予測誤差は若干大きくなるが, 重相関係数の値は 8 分音符単位のとときとほとんど差はなかった. 拍単位の制御を行なう実装には 4 分音符単位の説明変数を用いる方が適しているため, 本稿で提案する伴奏制御モデルによる伴奏システムでは 4 分音符単位での重回帰分析に基いたモデルを用いることとする.

過去 1 小節分を 4 分音符 4 つ分とし, 拍単位で「独奏者と伴奏者の間の時間的ずれ」と「伴奏者のテンポ変化量」から次の「伴奏者のテンポ変化量」を予測するモデルは, 重回帰分析の結果, 式(2)のように表せる.

$$c_n = \sum_{i=1}^4 \gamma_i d_{n-i} \sum_{j=1}^4 \omega_j c_{n-j} \quad (2)$$

表 3 伴奏者の演奏制御モデル（拍単位の制御）におけるパラメータの偏回帰係数，重相関係数 R ，予測誤差 E [ms]

γ_1	γ_2	γ_3	γ_4		
0.76	-0.05	-0.22	-0.21		
ω_1	ω_2	ω_3	ω_4	R	E
-0.45	-0.06	0.08	0.06	0.77	12.5

各パラメータの偏回帰係数と重相関係数，予測誤差を表 3 に示す．ここで，「独奏者と伴奏者との間の時間的ずれ」と「伴奏者のテンポ変化量」は，それぞれ式(3)，式(4)のように表せる．

$$d_n = T_n - t_n \quad (3)$$

$$c_n = l_n - l_{n-1} = t_n - 2t_{n-1} + t_{n-2} \quad (4)$$

ただし， T_n は独奏者の次の拍での発音時刻， t_n は伴奏者の次の拍での発音時刻， t_{n-1} ， t_{n-2} ，...は伴奏者の過去の演奏時刻の履歴である．また， l_n は伴奏者の演奏時間間隔であり， $l_n = t_n - t_{n-1}$ と定義される．それぞれの添字は 1 拍（4 分音符）を単位とし，その値が小さいほど過去の時刻を表す．

式(4)を式(2)に代入すると，過去の演奏履歴から次の拍での伴奏の発音時刻を決定する式は，次のように表せる．

$$t_n = \sum_{i=1}^4 \gamma_i d_{n-i} + \sum_{j=1}^4 \omega_j c_{n-j} + 2t_{n-1} - t_{n-2} \quad (5)$$

式(2)，式(3)より，式(5)は次のように表せる．

$$t_n = \sum_{i=1}^4 A_i T_{n-i} + \sum_{j=1}^6 B_j t_{n-j} \quad (6)$$

すなわち，次の拍における伴奏音の発音時刻は，過去 4 拍分の独奏音の発音時刻と過去 6 拍分の伴奏音の発音時刻によって決定される．伴奏のテンポ決定は，独奏者の演奏のタイミングで行なうのではなく，伴奏の各拍ごとに上記のモデルを用いて決定する．このため，独奏音が伴奏の発音タイミング決定時までには発音されていなかった場合，発音タイミングは実際の発音を待たずに内的クロックの推定値を用いて決定される．この時点では独奏音が発音されていないものと見なされ，独奏音の発音時刻が反映されるのは，次の発音タイミングの決定時以降となる．

4. 内的クロック

独奏者の楽譜上で，拍上に音符がない場合，または拍上に音符があった場合でも，伴奏が次拍の演奏テン

ポを決定するタイミングにおいて，まだその音符が発音されていない場合は，伴奏者の演奏制御モデルを推定するためのパラメータが不足してしまうため，その拍に相当する独奏者の内的クロックを推定する必要が生じる．そこで本研究では，下記の考えに基づいて独奏者の内的クロックを推定することとした．

- (1) 内的クロックの推定には，独奏者の楽譜において，半拍単位（8 分音符単位）のタイミングに存在する音のみを用いる．8 分音符単位以外のタイミングの音符はテンポからの逸脱が起こることが多いため，内的クロックの推定に用いるべきではないと考えた．
- (2) 推定したい拍（内的クロック）の付近にある独奏者の音符（上記の条件を満たす）2 つから外挿（過去の 2 音から推定する場合）または内挿（過去と未来の 1 音ずつの合計 2 音から推定する場合）により推定する．ただし，その前後に独奏パートに長い音符，あるいは，休符が存在し，拍上に独奏の音符がない場合には，その時刻は伴奏システムの拍時刻が優先すると考え，伴奏システムが演奏した時刻を用いる．

上述のように，伴奏システムのテンポ決定には過去 4 拍の独奏者の演奏情報が必要である．過去の情報に関しては，現在までに入力されたすべての音が利用可能であり，その情報を積極的に用いて推定精度を向上させるべきと考えられる．したがって，未来の音符がまだ存在しない場合には外挿を使って推定し，未来の音符も利用可能な場合には内挿によって推定することとした．

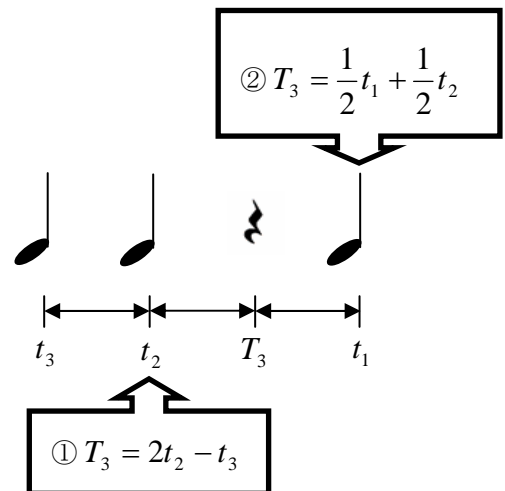


図 1 内的クロックの推定例

図 1 に、内的クロックを推定するプロセスの例を示す。この例では、独奏が 1 拍前の音を発音した時点 (t_2 の直後) で、 t_3 と t_2 から T_3 を外挿により推定する。推定結果は、時刻 T_3 以降の伴奏制御に利用される。続いて、独奏が内的クロック推定対象から 1 拍後の音を発音した時点 (t_1 の直後) で、 t_2 と t_1 から T_3 を内挿により推定し、その内的クロックの時刻を修正する。修正された結果は、時刻 t_1 以降の伴奏制御に利用される。このように、内的クロックの推定は、第 1 段階の推定 (外挿) と、第 2 段階の修正 (内挿) の 2 段階で行なう。

5. システムへの実装と合奏実験

内的クロックの推定を導入して伴奏者の演奏制御モデルを実装した伴奏システムが実際にどのようなふるまいをするか確認するため、以下の条件に基づいて人間の独奏者との合奏実験を行なった。

5.1. 演奏曲目

Hanon 作曲「ピアノの名手になる 60 練習曲」第 5 番を編曲し、収録対象とした。元来この曲はピアノの独奏曲であるが、互いの音を聴こえやすくするため、左手パートを 3 度下げ、10 度の並行移動としたうえで、左手パートを独奏パート、右手パートを伴奏パートとして、2 人の演奏者による合奏曲とし、被験者には独奏パートの楽譜を演奏するよう教示した。

今回の収録では、人間の独奏者のテンポ変化に対するコンピュータの伴奏者の追従性能を人間の演奏者に評価してもらうことを目的としたため、演奏は 1 つの鍵盤を打鍵するだけで楽譜上の音符が出力されるようにした。独奏者には 1 本の指で演奏を行なってもらい、その演奏 (テンポ変化) に対する伴奏システムの反応に意識を集中して評価してもらうように教示した。

5.2. 演奏者と収録条件

演奏者は音楽大学卒業生 2 名である。収録は実験 1 と実験 2 の 2 種類について実施した。なお、独奏者には伴奏を計算機が行なうことを教示している。以下にそれぞれの実験の概要を示す。

【実験 1】伴奏制御モデルの性能評価

従来モデル：過去 1 拍分の独奏者の演奏時刻と伴奏者の演奏時刻との間のずれのみを用いて演奏を制御するモデル。制御パラメータは本研究と同じデータを用い、単回帰分析により決定した。

提案モデル：提案手法に基づくモデル。ただし、内的クロックの推定を必要としない

次の 5 つの条件のそれぞれにおいて、従来モデルと提案モデルのどちらが良いかを合奏により評価してもらった。ただし、演奏者にはそれぞれがどちらのモデルであるかは教示せず、その順番もランダムに提示した。

- (条件 1)：ほぼ一定のテンポで演奏
- (条件 2)：徐々にテンポを上げて演奏
- (条件 3)：徐々にテンポを下げて演奏
- (条件 4)：テンポを突然上げて演奏
- (条件 5)：テンポを突然下げて演奏

ただし、あらかじめ以下の注意点を教示した。

- テンポ変化は小節の最初で行なう
- テンポを上げる際には 132 (bpm)、テンポを下げる際には 116 (bpm) をおおよその目安として演奏する
- 徐々にテンポを変える際には 1 小節程度にわたって行なう

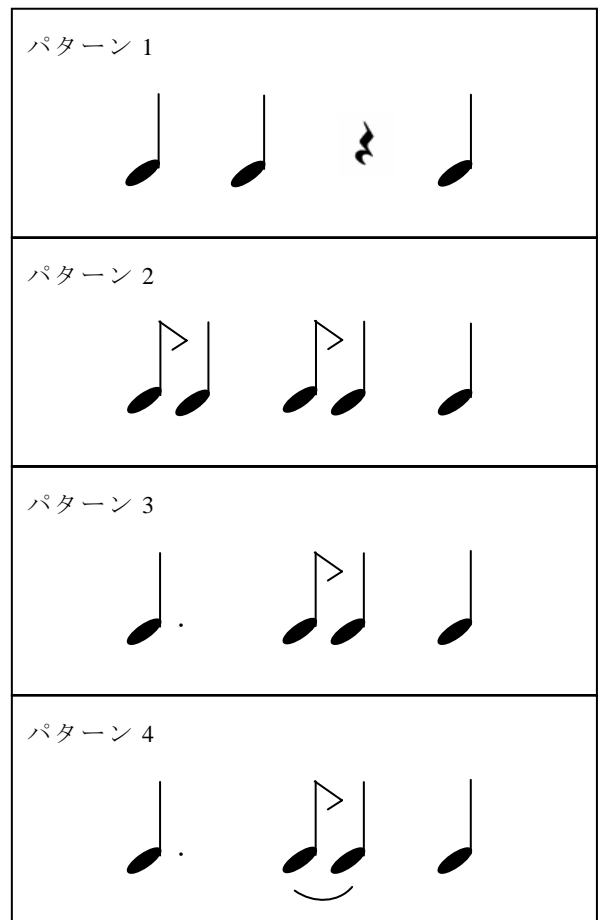


図 2 実験 2 で収録したリズムパターン

- テンポ変化を行なう回数は任意であるが、テンポを変化させた後は元のテンポに戻し（ただし、そのときの状況は評価しない）、テンポが安定した後に再びテンポ変化を行なう。

【実験 2】 内的クロック推定の性能評価

実験 1 の提案システムを用いて、実験 1 の楽譜と図 2 に示す 4 種類のリズムパターンの楽譜について、それぞれ合奏してもらい、その両者の伴奏の追従性能を評価してもらった。それぞれのリズムパターンに対し、まず実験 1 と同じ楽譜で合奏してもらい、続いて、各リズムパターンの楽譜で合奏をしてもらった。合奏終了後、後者の合奏（リズムパターン）を前者の合奏（拍ごと）と比較したときの追従性能の違いを 7 段階で評価してもらった。実験 2 ではテンポ変化は曲中で自由に行なってもらい、その総合評価で評価してもらうよう指示した。

5.3. 実験環境

独奏者は 88 鍵の MIDI ピアノで演奏を行ない、演奏者の演奏した打鍵時刻を記録すると同時に、MIDI 音源で該当する音を鳴らした。合奏の様子はビデオで撮影した。演奏の開始時には目や音による合図を送ることができないため、116 (bpm) で同じ音程の音を拍ごとに 8 拍分提示し、それに続けて演奏を開始するよう指示した。

5.4. 実験手順

システムとの合奏の練習として、収録前に提案モデルを実装したシステムとの合奏を 1 回行なった。練習終了後、上記の実験デザインに従って実験を行なった。

実験 1 については、システムとの合奏による慣れの影響も調べるため、実験 2 の終了後に再び収録を行なった。

6. 実験結果

実験後、演奏者から得たアンケートの回答結果を表 4、表 5 に示す。また、演奏者から得られたコメントを以下に挙げる。なお、演奏者にはランダムにシステムを提示したため、提案モデルや従来モデルの区別はわからない状態でのコメントであることに注意していただきたい。以下ではわかりやすくするため、それぞれに提案モデル、従来モデルと明記してある。

(被験者 A)

- 一定テンポでの演奏において、提案モデルは意図したテンポよりも先走っている（とくに上行形で）
- テンポを上げた場合の追従性能はよいが、テンポを下げた場合の追従は遅れぎみに感じる

表 4 実験 1 のアンケート回答結果 (Old: 従来モデルの方がよい, New: 提案モデルの方がよい, =: 甲乙つけがたい)

	被験者 A (1 回め)	被験者 B (1 回め)	被験者 A (2 回め)	被験者 B (2 回め)
条件 1	Old	Old	Old	Old
条件 2	Old	=	=	Old
条件 3	New	Old	New	Old
条件 4	Old	New	=	New
条件 5	New	Old	New	New

表 5 実験 2 のアンケート回答結果 (各パターンでの伴奏の実験 1 の曲での伴奏に対する比較評価)

	被験者 A	被験者 B
パターン 1	悪い	わずかに良い
パターン 2	わずかに悪い	わずかに悪い
パターン 3	わずかに悪い	わずかに良い
パターン 4	わずかに悪い	悪い

- 提案モデルは急激にテンポを下げた場合でも協調しやすい
- 実験 2 のパターン 1 では、テンポの上昇・下降によらず追従性能が悪い

(被験者 B)

- 提案モデルは全体的にバランスがとれていない、安定感がない、制御しにくい、という感がある、
- 急激にテンポを上げる場合、提案モデルの方が人間的には自然な追従をする（緩やかな変化ではあるが、従来モデルでは追従が急激すぎる）
- 急激にテンポを下げる場合は、急激に追従する方が人間らしく感じるため、従来モデルの方が自然
- 実験 2 のパターン 1 では、休符の間にテンポが変化するため、演奏していてストレスがない
- パターン 3 のように、8 分音符が入った場合の伴奏システムの演奏はとても人間的
- パターン 4 では、伴奏システムがリズムに惑わされている感じがする

表 4 や上記のコメントより、提案モデルが得た評価は、従来モデルの評価に比べて、必ずしも、高くなっているとはいえない。しかし、実験 1 では 2 回目の収録で提案モデルの評価が上がっているため、システムへの慣れが評価に影響しているとも考えられる。また、内的クロックの評価についてであるが、実験 2 ではわずかに良い、とするプラスの評価も見られ、内的クロックの推定が有効に機能しているといえる。しかし、

演奏者によって評価が反転している項目もあることから、音楽的嗜好による個人差が評価に影響を与えている可能性も考えられる。

7. おわりに

本研究では、独奏者の内的クロックを推定する手法を提案し、人間同士の合奏から得られた伴奏制御モデルに基づく伴奏システムを実現した。人間の演奏者との合奏実験を行ない、内的クロックの推定により伴奏制御が可能であることが明らかとなった。しかしながら、内的クロックを用いない場合での比較実験の結果から、人間同士の合奏から得られた推定モデルよりも過去1拍に敏感に反応する伴奏制御が好まれる場合もあることがわかり、必ずしも本提案手法が有効であったとは言いきれない。ただし、個人差による違いも大きく表れたため、より多くの演奏者との合奏実験を行なうことにより、今後とも引き続き検討していく必要があると考えられる。

今後の課題として、内的クロックの推定精度を向上するとともに、提案手法の伴奏制御モデルの問題点を明らかにし、より人間の演奏者に好まれる伴奏システムの実現を検討していきたいと考えている。また、現状ではモデルの推定に用いた4拍子の楽曲にしか対応していないが、合奏のデータを増やすことにより、3拍子の楽曲などにも対応できるようにしたいと考えている。

参 考 文 献

- [1] 堀内靖雄：自動伴奏，コンピュータと音楽の世界—基礎からフロンティアまで，長嶋洋一，橋本周司，平賀譲，平田圭二（編），pp.252-269，共立出版（1998）。
- [2] Dannenberg, R.B. : An On-Line Algorithm for Real Time Accompaniment, *Proc. International Computer Music Conference*, pp.193-198 (1984).
- [3] Vercoe, B. : The synthetic performer in the context of live performance, *Proc. International Computer Music Conference*, pp.199-200 (1984).
- [4] 堀内靖雄，坂本圭司，市川熹：合奏時の人間の演奏制御の分析・推定，情報処理学会論文誌，Vol.45, No.3, pp.690-697 (2004).
- [5] Gouyon, F and Dixon, S : A Review of Automatic Rhythm Description Systems, *Computer Music Journal*, Vol.29, Issue 1, pp.34-54 (2005).
- [6] 谷口高士：音は心の中で音楽になる[音楽心理学への招待]，北大路書房（2000）。
- [7] 大蔵康義：音と音楽の基礎知識，国書刊行会（1999）。
- [8] 田辺義和：Windows サウンドプログラミング—音の知識×プログラミングの知識，翔泳社（2001）。

- [9] Curtis Roads：コンピュータ音楽—歴史・テクノロジー・アート—，青柳龍也・小坂直敏・後藤真孝・引地孝文・平田圭二・平野砂峰旅・堀内靖雄・松島俊明（訳・監修），東京電機大学出版（2001）。