

伴奏システムのための演奏時刻予測モデル

和田静花^{†1} 堀内靖雄^{†1} 黒岩眞吾^{†1}

概要: 伴奏システムが人間らしい伴奏を行うには、人間の協調演奏法を分析しモデル化する必要がある。我々の先行研究では「独奏者と伴奏者の演奏タイミングのずれ」と「伴奏者の時間長変化」という2つの演奏情報から次拍の時間長変化を予測するモデルを提案した。しかし本研究で名演奏家の合奏録音を分析した結果、直前の時間長変化よりも過去数拍分の平均テンポからのずれの方が予測に有効であることがわかった。そこで新たに線形予測モデルを構築し、予測誤差による評価を行ったところ、従来手法よりも高い精度での予測が可能となった。さらに、本研究では一般的な楽曲における長い音符や休符などにも対応可能な予測モデルを提案する。

1. はじめに

伴奏システムとは人間の独奏者の伴奏をコンピュータが行うもので、人間とコンピュータとのリアルタイム・インタラクティブ・システムの一つとして、音楽情報処理研究のみならず、人間とのリアルタイム・インタラクションを考える上で重要な研究テーマとなっている[1][2][3][4][5]。伴奏システムでは楽譜は独奏・伴奏の両方がシステムに与えられており、独奏者の演奏と協調するようにシステムの演奏タイミングを制御する。従来の研究ではおもに演奏ミス等、独奏者の追跡方法に重点が置かれていたため、伴奏システム自身の演奏は機械的で不自然なものとなることが多かった。この原因として、人間の伴奏者がどのような制御方法により独奏者と合わせているのかが解明されていないことが挙げられる。我々の先行研究[6]では二人の演奏者が二台の MIDI ピアノで合奏したデータを分析することにより、伴奏者が独奏者の演奏に対し、どのように演奏しているかを表現する伴奏制御モデルを提案した。しかしながら、分析に利用した楽曲は音価一定の単純な練習曲であったため、現実的な合奏における有効性は不明であった。先行研究[7]では[6]の伴奏制御モデルで使用したパラメータについて、市販されている著名な演奏家による録音 CD を分析することにより、一般的な楽曲のテンポ安定部における有効性を示した。[8]ではさらに、休符や長い音符などで予測に必要なパラメータが揃わない場合の伴奏制御法を提案した。本研究では[6]で用いたパラメータに対して、別の予測パラメータを提案し、さらに[8]のアイデアも利用し、一般的な楽曲での全ての拍における演奏予測可能な予測モデルを提案する。

2. 先行研究の伴奏制御モデル

[6]では「独奏者と伴奏者の演奏タイミングのずれ」と「伴奏者の時間長変化」の過去の履歴から伴奏者の次の演奏タイミングを予測する伴奏制御モデルが提案されている。

「独奏者と伴奏者の演奏タイミングのずれ」とは各拍における両者の演奏時刻のずれであり、 n 拍目のずれ d_n は式(1)で定義される。

$$d_n = T_n - t_n \quad (1)$$

ただし、 T_n は n 拍目の独奏者の演奏時刻、 t_n は n 拍目の伴奏者の演奏時刻である。

「伴奏者の時間長変化」とは各拍における伴奏者の演奏時間間隔(Inter onset interval, 以下 IOI)の変化を表し、 n 拍目の時間長変化 c_n は式(2)で定義される。

$$c_n = l_n - l_{n-1} \quad (2)$$

ここで l_n は伴奏者の IOI であり、式(3)で定義される。

$$l_n = t_n - t_{n-1} \quad (3)$$

以上 2 種類のパラメータにより未来の伴奏者の演奏を予測する推定モデルとして、式(4)を用いた。

$$c_n = \sum_{i=1}^p \alpha_i d_{n-i} + \sum_{j=1}^q \beta_j c_{n-j} \quad (4)$$

式(4)における各パラメータの係数 α_i および β_j は、人間の合奏データに対して重回帰分析を適用し、結果として求められる偏回帰係数を用いている。 p, q は履歴の長さを表す数であるが、一般的な楽曲においては休符や長い音符の影響により、過去数拍分の連続したデータの履歴を取得することが困難である。したがって、[7]では直前 1 拍前のずれ d_{n-1} と時間長変化 c_{n-1} について、次拍の時間長変化 c_n の予測に対する有効性を評価した。

3. 分析データ

3.1 分析対象曲ならびに分析 CD

本研究では、C. Franck 作曲「ピアノとヴァイオリンのためのソナタ イ長調 第 4 楽章」において分析を行った。この楽曲はピアノとヴァイオリンによる合奏であるが、今回はピアノ伴奏・フルート独奏の合奏用に編曲されたものを用いた。この楽曲から、3-36, 53-64, 89-98, 187-218, 222-235 小節を使用した。実際に使用した楽曲の一部分を図 1 に示す。これらの箇所はテンポ変化が少なく、独奏・伴奏ともに拍上に存在する音が多いために採用した。なお、本研究ではテンポ安定部における予測モデルを想定しているため、楽譜上に大きなテンポ変化の指示がある箇所は対象外と考え、分析から除外した。本研究では 4 組の名演奏家の CD を使用したが、表 1 にそれぞれの演奏家情報を示す。また、表 2 に各演奏家における伴奏者の IOI お

^{†1} 千葉大学

よびテンポ, IOI の標準偏差の平均を示す.

3.2 演奏時刻の計測手法

先行研究[9]では, 様々な楽器による合奏において指向性マイクを使用して収録することにより各楽器の音を分離して分析することを可能にしているが, この手法では大量にデータ収集をすることが困難である. そのため本研究では, 楽譜から MIDI データを作成し, 文献[10]のツールを用いて市販 CD の音響信号に MIDI データをフィッティングさせ, 各音符の立ち上がり時刻を測定した. しかしながら, フィッティングエラーや微妙な誤差が生じるため, フィッティング結果に対し, 手修正を行った. その際, 文献[11]の P. 67 のツールを利用すると同時に人間の耳による聴取も行い, 各拍上の音符の立ち上がり時刻を厳密に決定した[12].

4. 提案モデル

4.1 従来モデルのパラメータ検証

従来モデルで用いた 2 つのパラメータについて, 本研究使用曲において予測値との相関を調査する.

4.1.1 独奏者と伴奏者の演奏タイミングのずれ

ここでは直前の二者間の演奏タイミングのずれと伴奏者の次拍の時間長変化との関係を調査する. 図 2 に横軸を d_{n-1} , 縦軸を c_n とした散布図を演奏家ごとに示す.

図 2 より, 全演奏家においてある程度の正の相関関係が見られることから, 例えば伴奏者が独奏者に対して先行してしまったために両者のタイミングがずれた場合, 伴奏者は次の時間長を直前の時間長よりも長くする (テンポを遅くする) という傾向が表れている. つまり, 演奏者間の直前のずれを補正するようにテンポ変化を行っているのではないかと考えられる.

4.1.2 伴奏者の時間長変化

次に, 伴奏者の時間長変化と次拍の時間長変化との関係を調査する. 図 3 に, 横軸を c_{n-1} , 縦軸を c_n とした散布図を, 演奏家ごとに示す.

図 3 より, 全演奏家においてある程度の負の相関関係が見られることから, 例えば伴奏者が直前の時間長をその前の拍の時間長よりも長くしていた場合, 次拍の時間長は直前の時間長よりも短くするという傾向が表れている. つまり, 伴奏者は直前の時間長の変化を相殺するように次拍の時間長を調整して演奏を行っているのではないかと考えられる.

4.2 新たなパラメータの導入

図 3 を観察すると, 人間の伴奏者の演奏では直前の時間長変化に対して, その変動を打ち消すように時間長が変化していることから, 伴奏者は自身のテンポを一定に保とうと演奏しているのではないかと考えた. テンポを一定に保つということは, 人間が心の中で一定周期の拍を刻んでおり, その拍時刻に合わせようと調整しているのではないかと考えられる.



(a) 1-3 小節



(b) 87-90 小節

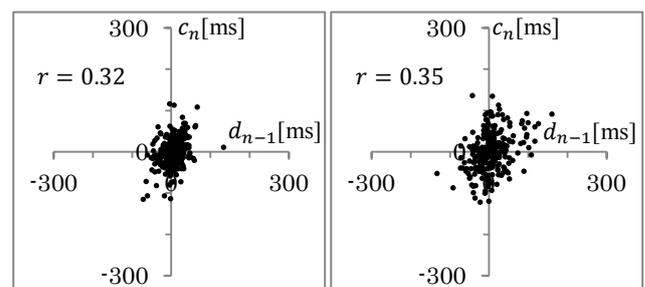
図 1 本研究使用楽譜抜粋

表 1 分析に使用した CD

	独奏者 (フルート)	伴奏者 (ピアノ)
Duo A	Emmanuel Pahud	Eric Le Sage
Duo B	James Galway	Martha Argerich
Duo C	Jean-Pierre Rampal	Pierre Barbizet
Duo D	William Bennett	Clifford Benson

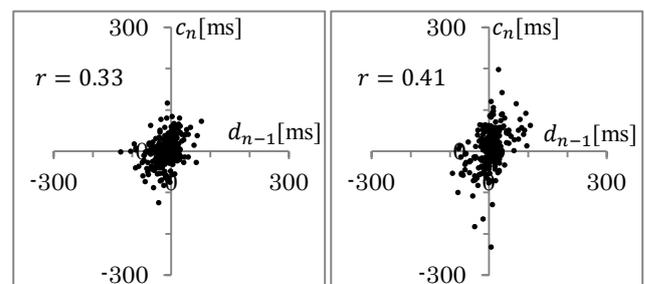
表 2 伴奏者の IOI・テンポ・標準偏差の平均

	Duo A	Duo B	Duo C	Duo D
IOI [ms]	362	371	314	403
テンポ [bpm]	166	162	191	149
IOI の標準偏差 [ms]	33.4	57.2	31.0	41.7



(a) Duo A

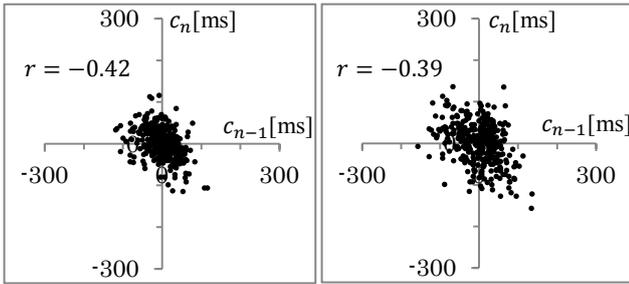
(b) Duo B



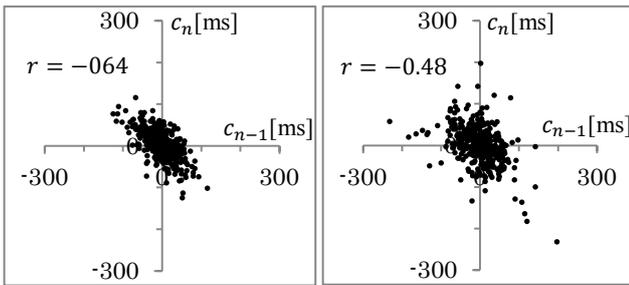
(c) Duo C

(d) Duo D

図 2 d_{n-1} と c_n との関係

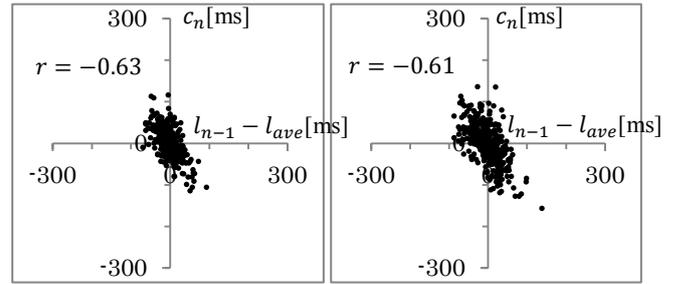


(a) Duo A (b) Duo B

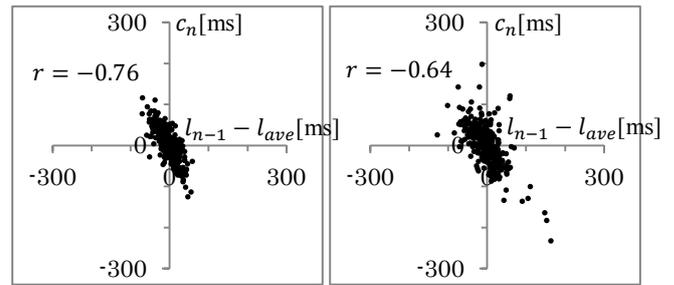


(c) Duo C (d) Duo D

図3 c_{n-1} と c_n との関係



(a) Duo A (b) Duo B



(c) Duo C (d) Duo D

図4 $l_{n-1} - l_{ave}$ と c_n との関係

そこで、本研究では「心の中で感じている」一定の拍時間長を伴奏者の平均時間長 l_{ave} とみなし、その平均時間長からのずれ $l_{n-1} - l_{ave}$ と予測値 c_n との関係を調査する. 平均時間長からのずれ $l_{n-1} - l_{ave}$ を横軸に、それに対する次拍の時間長変化 c_n を縦軸に取った散佈図を図4に示す. なお、ここでの平均時間長とは便宜的に、伴奏者の直前8拍の平均時間長を取ったものとする.

図4から、 $l_{n-1} - l_{ave}$ と c_n の間には、非常に高い負の相関が見られた. 図3と比較すると、 c_{n-1} とよりも $l_{n-1} - l_{ave}$ との相関の方が高く、 $l_{n-1} - l_{ave}$ を用いることにより、予測性能が向上すると考えられる. したがって、本研究では従来モデルで使用された説明変数 c_{n-1} の代わりに $l_{n-1} - l_{ave}$ を用いることとし、式(5)のように定義する.

$$v_{n-1} = l_{n-1} - l_{ave} \quad (5)$$

ここでの平均時間長 l_{ave} は直前の時間長 l_{n-1} までの平均時間長であるが、何拍分の平均を取るのが適当であるのかを検討する必要がある.

そこで、式(6)において l_{ave} を直前1拍の平均時間長から12拍平均まで変化させ、それぞれ t_n を予測するモデルを構築した. 算出した演奏時刻について、実測値との誤差を平均二乗誤差により求めた. 誤差の値を図5に示す.

$$t_n = t_{n-1} + l_{ave} \quad (6)$$

図5より、Duo A, Duo B, Duo Cは8拍平均までは誤差が減少しており、それ以上の拍で平均を取った場合はほぼ横ばいとなっている. Duo Dでは8拍平均を取った場合が最も誤差が小さくなっている. この結果から、本研究では平均時間長 l_{ave} として伴奏者の直前8拍の平均を使用することが妥当であると考え、モデルを構築する.

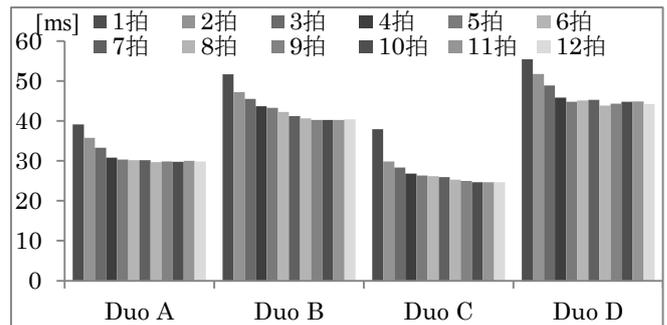


図5 1~12拍平均を用いた場合の予測誤差

以上の結果から、本研究での提案モデルでは直前の二者間の演奏タイミングのずれ d_{n-1} と伴奏者の8拍の平均時間長からのずれ v_{n-1} から、次拍の時間長変化 c_n を予測し、式(7)により次拍の伴奏者の演奏時刻 t_n を導出する.

$$t_n = t_{n-1} + (l_{n-1} + c_n) \quad (7)$$

4.3 パターン別予測パラメータの分類

本研究では一般的な楽曲を分析対象としているため、長い音符や休符などにより、予測に必要な上記二つのパラメータが揃わない拍が存在する. そのため、[8]で用いた分類法に従い、適用する楽譜の状況により次のように使用するパラメータを変更することで、どのような場面でも適用可能なモデルを構築する.

① 直前1拍前に独奏・伴奏の音符が共に存在する場合

直前1拍前の拍上に独奏 T_{n-1} ・伴奏 t_{n-1} の音符が共に存在する場合は直前の二者間のずれ d_{n-1} を算出することが可能である. この場合は、 d_{n-1}, v_{n-1} 両方から c_n を予測するモデルとして式(8)を使用する. このモデルを基本モデルと

する.

$$c_n = \alpha_1 d_{n-1} + \beta_1 v_{n-1} + e_1 \quad (8)$$

- ② 直前 1 拍前の独奏または伴奏の音符どちらかが存在しない場合

直前 1 拍前の拍上に独奏 T_{n-1} または伴奏 t_{n-1} の音符どちらかが存在しない場合には、直前の二者間のずれ d_{n-1} を算出することができない。そのため、式(9)を用いて v_{n-1} のみでの予測を行う。しかしながら、もし t_{n-1} が存在しない場合は v_{n-1} の計算を行う際に必要な、伴奏者の直前の時間長 l_{n-1} を算出することができない。したがってこの場合には、伴奏者が独奏者の立ち上がり時刻に同時に演奏を行ったと仮定し、 t_{n-1} を T_{n-1} で置き換えて計算することとする。

$$c_n = \beta_2 v_{n-1} + e_2 \quad (9)$$

- ③ 直前 2 拍間の独奏または伴奏の音符どちらかが存在しない場合

この場合も②同様 d_{n-1} を利用することができないため、式(9)を用いて v_{n-1} のみでの予測を行う。この際伴奏の演奏時刻 t_{n-1}, t_{n-2} が存在しない場合は、先述したように独奏の演奏時刻 T_{n-1}, T_{n-2} を代わりに使用し、 v_{n-1} を計算する。

- ④ 直前 1 拍前の独奏・伴奏の音符が共に存在しない場合

直前 1 拍前の独奏または伴奏の音符どちらかが存在しない場合には、式(9)を用いて予測を行ったが、今回のように両者が存在しない場合にも同様に式(9)を用いる。しかしながら、今回は T_{n-1}, t_{n-1} が共に存在しないため v_{n-1} を計算することができない。そのため、このパターンにおいては式(10)のように本モデルによって過去から予測した伴奏者の演奏時刻 t'_{n-1} を t_{n-1} の代わりに用い v_{n-1} を計算し、式(9)によって c_n を算出する。

$$v_{n-1} = l_{n-1} - l_{ave} \quad (10)$$

$$= (t'_{n-1} - t_{n-2}) - l_{ave}$$

- ⑤ 直前 2 拍間の独奏・伴奏の音符が共に存在しない場合

この場合も d_{n-1} を利用することができないため式(9)を用いて c_n を計算するが、今回の場合は $T_{n-1}, T_{n-2}, t_{n-1}, t_{n-2}$ がすべて存在しないため v_{n-1} を計算することができない。そのため式(10)と同じように本モデルより予測した t'_{n-1}, t'_{n-2} を用い、 v_{n-1} の計算を行う。

以上のように提案モデルでは、楽譜中の状況に従って使用するパラメータの組み合わせを変更し、線形予測を行う。次節では実際に予測モデルを構築し、重回帰分析を行った結果を述べる。

5. 提案モデルの評価

ここでは d_{n-1}, v_{n-1} パラメータを用い重回帰分析を行う。拍上の音符の有無により使用するパラメータを変更し、前節で述べたパターン別重回帰モデルをまとめると式(8), (9)のようになる。

$$c_n = \alpha_1 d_{n-1} + \beta_1 v_{n-1} + e_1 \quad (8)$$

$$c_n = \beta_2 v_{n-1} + e_2 \quad (9)$$

$\alpha_1, \beta_1, \beta_2$ は各説明変数の係数であり、 e_1, e_2 は定数項である。これらの値は重回帰分析により求めた偏回帰係数を使用した。実際に用いた偏回帰係数を表 3 に示す。求めた演奏時刻から平均二乗誤差により予測誤差を算出した。クローズド実験・オープン実験それぞれの結果を表 4 に示す。

なお、今回用いた楽曲では前半 (3-36, 53-64, 89-98 小節) と後半 (187-218, 222-235 小節) とで類似した楽曲構造を持つため、クローズド実験では前半・後半の両方で学習・評価を、オープン実験では前半で学習・後半で評価を行った。これは全ての演奏家で共通であり、実際に使用した演奏家一組あたりのデータ数を表 5 に示す。

5.1 従来モデルとの比較

ここでは、[6]で提案された c_{n-1} を用いた従来モデルとで比較実験を行う。ここで言う従来モデルとは、提案モデルの式(8), (9)の v_{n-1} を c_{n-1} で置き換えたもので、式(11), (12)を使用し、提案モデル同様楽譜中の状況により使用パラメータを変更したものとする。

$$c_n = \alpha_1 d_{n-1} + \beta_1 c_{n-1} + e_1 \quad (11)$$

$$c_n = \beta_2 c_{n-1} + e_2 \quad (12)$$

こちらにも、重回帰分析により偏回帰係数を求めた結果を表 6 に示す。なお、使用したデータ数は表 5 と同様である。表 6 の値を式(11), (12)に用い予測モデルを構築し、平均二乗誤差により予測誤差を算出した。クローズド実験・オープン実験それぞれの結果を表 7 に示す。

また、表 4, 7 の結果から提案モデルと従来モデルとで予測誤差を比較した結果を図 6 に示す。

図 6 より、クローズド実験・オープン実験ともに全演奏家において提案モデルの予測誤差が従来モデルの予測誤差よりも小さくなっており、提案モデルは従来モデルと比較して改善していることがわかる。

さらに、各パラメータの組み合わせ別に予測誤差を算出した結果を表 8 に示す。表 8 より、全てのパラメータパターンにおいて提案モデルでの改善が見られ、特に表 7 の結果から従来モデルで最も誤差の大きかった Duo B は、 c_{n-1} のみを用いた場合に 50ms 以上の大きな誤差であるのに対し、提案モデルで v_{n-1} を用いることによりかなり改善していることがわかる。

次に、予測誤差の大きい箇所注目して分析を行う。

人間同士の合奏では、互いのずれが 100ms 以上の場合に合奏を維持することが難しいと言われている[13]。そのため、今回用いたデータにおいて、予測誤差が±100ms 以上となった箇所を詳細に調査した。該当するデータ数と全体に対する割合をまとめたものを表 9 に示す。

表 9 より、合奏を保つのが難しいとされる 100ms 以上のずれは、クローズド実験・オープン実験ともに全体の 2% 以内に収まっていることがわかる。

表 3 提案モデルの偏回帰係数

(a) クローズド実験

	Duo A	Duo B	Duo C	Duo D
α_1	0.143	0.258	0.289	0.537
β_1	-0.884	-0.794	-1.189	-0.802
e_1	0.001	-0.002	0.001	0.001
β_2	-1.064	-0.874	-1.299	-1.119
e_2	-0.009	-0.008	0.002	-0.008

(b) オープン実験

	Duo A	Duo B	Duo C	Duo D
α_1	0.249	0.311	0.343	0.564
β_1	-0.934	-0.797	-1.335	-0.886
e_1	0.002	-0.002	0.001	0.002
β_2	-1.021	-0.868	-1.289	-1.125
e_2	-0.008	-0.010	0.000	-0.009

表 4 提案モデルの予測誤差 [ms]

	Duo A	Duo B	Duo C	Duo D
クローズド	27.0	37.9	22.7	33.2
オープン	27.3	39.7	23.8	29.9

表 5 実験使用データ数

	学習		評価
	d_{n-1}, v_{n-1}	v_{n-1}	
クローズド	266	91	357
オープン	141	57	159

5.2 二者間のずれとの比較

次に、表 4 に示した予測誤差は伴奏制御モデルとして有効な誤差の範囲と言えるのかどうかを調査する。本研究ではその指標として、人間の独奏者・伴奏者の合奏において実際に観測された二者間の演奏タイミングのずれを用いる。予測誤差がこの二者間のずれ程度であれば、人間はほぼずれを感じることなく合奏を続けられると考えられるためである。したがって、表 10 に各演奏家の二者間のずれの平均値と標準偏差を示す。

表 4 と表 10 との値を比較すると、提案モデルによる予測誤差の値は人間の演奏者同士の演奏時刻のずれの値とほぼ同程度であることがわかる。[9]では、楽器演奏においてはピッチや音色、立ち上がり時刻の影響で単純な刺激音よりも二者間のずれを認識するのが難しいと述べられている。したがって表 4 程度の誤差の値は、伴奏システムのための予測モデルとして高い精度での予測が可能であると考えられる。

以上の結果から、本研究の提案モデルはテンポ安定部における演奏制御モデルとして有効であることが示された。

表 6 従来モデルの偏回帰係数

(a) クローズド実験

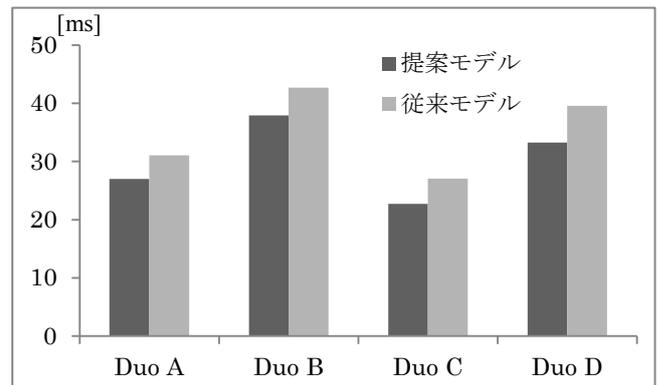
	Duo A	Duo B	Duo C	Duo D
α_1	0.372	0.412	0.311	0.616
β_1	-0.473	-0.403	-0.549	-0.565
e_1	0.000	0.0000127	0.004	0.003
β_2	-0.277	-0.371	-0.693	-0.367
e_2	-0.001	-0.004	0.002	-0.011

(b) オープン実験

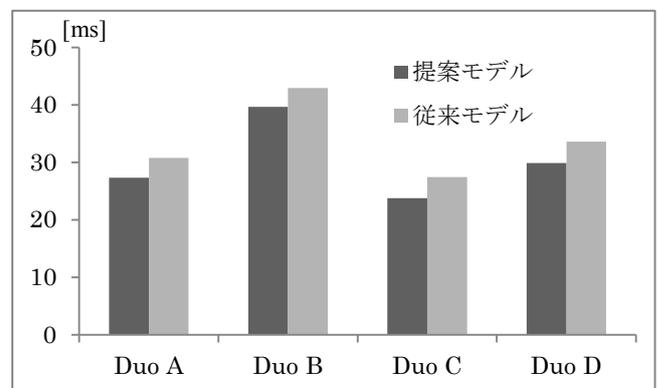
	Duo A	Duo B	Duo C	Duo D
α_1	0.440	0.410	0.390	0.646
β_1	-0.464	-0.402	-0.634	-0.455
e_1	0.002	0.000	0.005	0.006
β_2	-0.290	-0.316	-0.638	-0.373
e_2	0.000	-0.005	0.003	-0.012

表 7 従来モデルの予測誤差 [ms]

	Duo A	Duo B	Duo C	Duo D
クローズド	31.1	42.7	27.1	39.6
オープン	30.8	43.0	27.4	33.6



(a) クローズド実験



(b) オープン実験

図 6 予測誤差比較

表 8 パラメータ組み合わせ別予測誤差 [ms]

(a) クローズド実験

	提案モデル		従来モデル	
	d_{n-1}, v_{n-1}	v_{n-1}	d_{n-1}, c_{n-1}	c_{n-1}
Duo A	27.1	26.9	36.3	29.0
Duo B	36.0	43.0	38.7	52.5
Duo C	22.7	22.8	27.3	26.3
Duo D	33.9	31.2	39.6	39.6

(b) オープン実験

	提案モデル		従来モデル	
	d_{n-1}, v_{n-1}	v_{n-1}	d_{n-1}, c_{n-1}	c_{n-1}
Duo A	27.0	28.4	28.0	39.3
Duo B	37.3	47.3	38.5	56.3
Duo C	23.1	26.1	26.5	30.7
Duo D	31.1	24.7	34.0	32.2

表 9 予測誤差±100ms以上のデータ

(a) クローズド実験

	全体	予測誤差±100ms以上	
	データ数	データ数	全体に対する割合
Duo A	357	1	0.28%
Duo B	357	4	0.84%
Duo C	357	0	0.00%
Duo D	357	5	1.40%

(b) オープン実験

	全体	予測誤差±100ms以上	
	データ数	データ数	全体に対する割合
Duo A	159	1	0.63%
Duo B	159	3	1.89%
Duo C	159	0	0.00%
Duo D	159	2	1.26%

表 10 二者間のずれの平均と標準偏差 [ms]

	Duo A	Duo B	Duo C	Duo D
平均	25.5	38.5	31.9	30.3
標準偏差	17.2	27.6	22.1	21.7

6. おわりに

本研究では、名演奏家4組によるCD録音から人間の伴奏者の演奏を分析し、伴奏システムのための新たな演奏制御モデルを提案し、評価を行った。従来研究[6]で用いたパラメータに加え、本研究では新たに「直前の時間長の過去8拍の平均時間長からのずれ」を用いることで、より高い精度での予測可能なモデルを実現させた。

一般的な楽曲では、長い音符や休符などにより予測に必要なパラメータが揃わない拍が存在するが、そのような場合には、もしどちらか一方がその拍上に音符が存在する場合には、その演奏時刻に両者が同時に演奏を行ったと仮定し、どちらの音符も拍上に存在しない場合には過去の平均テンポのみを使って次拍のタイミングを予測するという手法で、一般的な楽曲での全ての拍における演奏予測を実現させた。

提案モデルと従来モデルとの比較実験を行ったところ、テンポ安定部において提案モデルを使用した方が予測誤差が小さくなるという結果が得られた。さらにこの予測誤差が、実際の間奏同士の演奏タイミングのずれと同程度であることから、今回の提案モデルは妥当な近似モデルであると考えられる。

今後は本研究で提案した予測モデルを伴奏システム内に実装し、評価実験を行っていきたいと考えている。

謝辞 本研究は、亀岡弘和氏（NTT/東京大学）ならびに高宗典玄氏、土屋政人氏（東京大学）よりお借りした分析ツールを用いて実施された。ここに深謝の意を表す。

参考文献

- 堀内靖雄: 自動伴奏, 長嶋洋一・橋本周司・平賀譲・平田圭二(編), コンピュータと音楽の世界, 共立出版, pp. 252-269 (1998).
- Roger B. Dannenberg: An On-Line Algorithm for Real-Time Accompaniment, Proceedings of International Computer Music Conference, pp. 193-198 (1984).
- Barry Vercoe: The synthetic performer in the context of live performance, Proceedings of International Computer Music Conference, pp. 199-200 (1984).
- Christopher Raphael: Music plus one: A system for flexible and expressive musical accompaniment, Proceedings of International Computer Music Conference (2001).
- 武田晴登, 西本卓也, 嵯峨山茂樹: HMMによるMIDI演奏の楽譜追跡と自動伴奏, 情報処理学会研究報告, Vol. 2006, No. 90, pp. 109-116 (2006).
- 堀内靖雄, 坂本圭司, 市川薫: 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).
- 和田静花, 堀内靖雄, 黒岩眞吾: 名演奏家の合奏録音における伴奏者の演奏タイミング制御の分析, 情報処理学会研究報告. [音楽情報科学] 2013-MUS-100(9), pp. 1-6 (2013).
- S. Wada, Y. Horiuchi, S. Kuroiwa: Tempo Prediction Model for Accompaniment System, Proceedings of ICMC & SMC 2014 Joint Conference, pp. 1298-1303 (2014).
- Rasch, R.A: Synchronization in Performed Ensemble Music, Acustica, vol. 43, pp.121-131 (1979).
- 高宗典玄, 亀岡弘和, 土屋政人, 嵯峨山茂樹: 補助関数法による音楽音響信号へのMIDI信号の詳細なフィッティング, 日本音響学会2012年秋季研究発表会, pp. 2-10-8 (2012).
- Hirokazu Kameoka: Statistical Approach to Multipitch Analysis, Ph.D. Thesis, The University of Tokyo (2007).
- 橋田光代, 松井淑恵, 北原鉄朗, 片寄晴弘: ピアノ名演奏の演奏表現情報と音楽構造情報を対象とした音楽演奏表情データベース CrestMusePEDBの構築, 情報処理学会論文誌, Vol.50, No.3, pp.1090-1099 (2009).
- C. Bartlette, D. Headlam, M. Bocko, G. Velickic: Effect of network latency on interactive musical performance, Music Perception, Vol. 24, pp. 49-62 (2006).