合奏時の人間の演奏制御の分析・推定

堀内靖雄[†] 坂本圭司[††] 市川熹[†]

本論文では，人間同士の合奏に見られるような，互いに相手の影響を受けながら合奏を行う状況において，人間の演奏制御を推定する手法を提案する．先行研究では相手の影響を受けずに演奏を行う計算機（テンポが固定）と人間が合奏する状況において，人間の演奏制御の推定が行われ，比較的高い推定精度が得られているが，この手法を互いに影響を与えあう人間同士の合奏に適用した場合，それほど高い推定精度が得られなかった．そこで本研究では人間同士の合奏を収録，分析することにより，お互いに相手の影響を受けながら合奏を行う状況にも対応できるよう，人間の演奏制御の新たな推定手法を提案する．推定に用いるパラメータとして従来手法のパラメータに加え，過去 1 小節の「両者のずれ」と「自分の演奏変化量」を用いた重回帰分析により，人間の演奏制御を推定した結果，計算機との合奏においては従来手法とほぼ同程度，人間同士の合奏においては従来手法に比べて高い推定精度を得た．また，分析の結果，人間の演奏家は直前の情報だけでなく，過去 1 小節程度にわたる情報も利用して演奏制御していることが示唆された．

Estimating the Model of Human Ensemble

Yasuo Horiuchi,[†] Keiji Sakamoto[††] and Akira Ichikawa[†]

In this paper, we will estimate a model of human performance in ensemble. 40 performances played by 4 pairs of pianists (4 soloists and 4 accompanists) were recorded. It is supposed that human performers decide their future performance in real time using two factors in the past. One is the history of the time difference between two notes which were performed by the soloist and the accompanist and should be performed at the same time in their scores. The other is the history of the tempo modification of the performer. The recorded data were analyzed by multiple regression analysis and it was confirmed that the proposal model has better precision than the model of a previous study. Furthermore, from the analysis of coefficients of the estimated model, it was suggested that human performers use not only the immediately preceding history but the long history.

1. はじめに

現在，音楽情報科学研究の分野では人間とコンピュータとの音楽によるインタラクションに関する研究がさかんである．本研究はそのようなインタラクティブ・システムの中で，伴奏システム[1)]に焦点をあてている．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 1 つとして，音楽情報科学のみならず，人間とのリアルタイム・インタラクションを考えるうえで重要な研究テーマとなっている．伴奏システムでは楽譜情報はあらかじめ独奏パートと伴奏パート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伴奏システムの演奏タイミングをリアルタイムで制御する．

伴奏システムが研究され始めたのは計算機の能力が向上し，リアルタイムでの音楽処理が可能となった 1980 年代半ば以降であるが[2),3)]，従来の研究では，おもに独奏者の演奏の追跡方法に重点が置かれていたため，伴奏システム自体の演奏は機械的で不自然なものとなることが多かった．その理由は人間の演奏を考慮することなく，簡単な制御モデルで伴奏システムの演奏を決定していたことにあると考えられる．本研究ではこの問題点を解決するため，合奏で人間が実際に行っている演奏制御モデルを推定することを目的とする．

先行研究として文献 4) では，あらかじめテンポ変化が決められた計算機と人間との合奏を収録，分析し，人間の演奏制御モデルを推定した．これは一般の合奏における人間の演奏制御モデルを推定したものではなく，相手の影響を受けずに演奏を行える計算機と合奏

† 千葉大学
Chiba University
†† 株式会社日立製作所
Hitachi, Ltd.

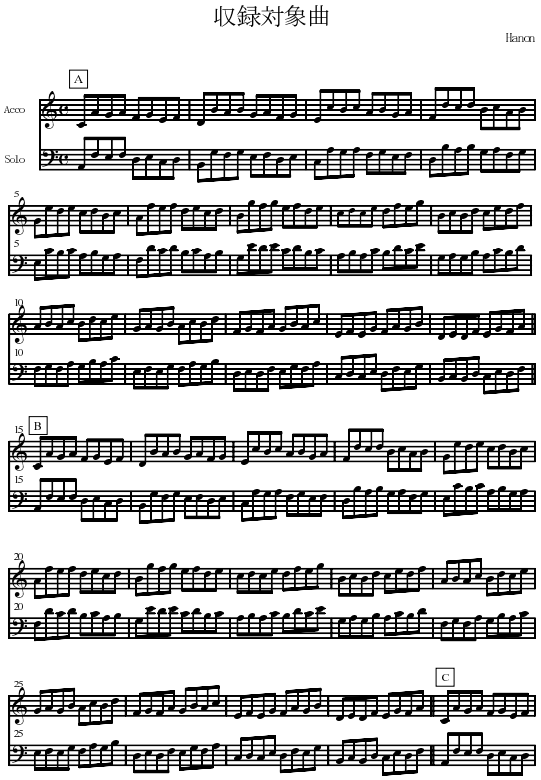

図 1　収録に用いた楽譜．Acco が伴奏パート，Solo が独奏パート
Fig. 1　The score used at the recording. “Acco” is the accompaniment part and “Solo” is the solo part.

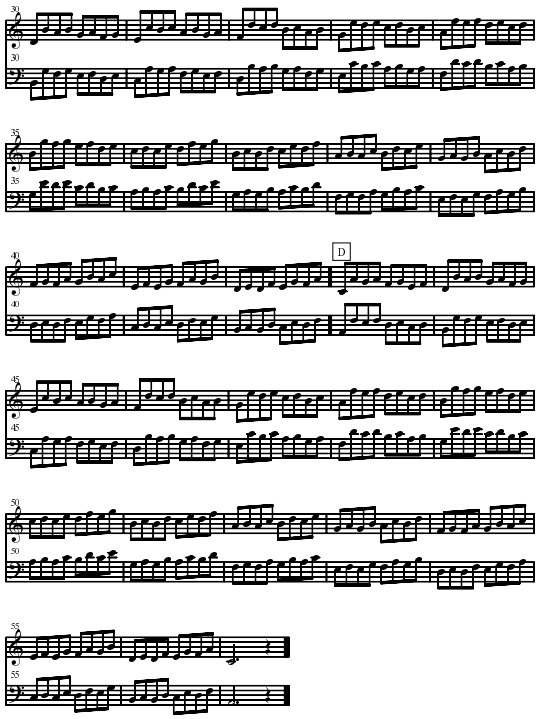

図 2　収録に用いた楽譜（続き）
Fig. 2　The score used at the recording (contd.).

する際の人間のモデルを推定したものである．しかしながら，そのような特殊な状況とはいえ，高い精度で人間の演奏制御の振舞いが推定されているため，本研究ではこの方法を踏襲しつつ，演奏モデルを改良し，人間同士の合奏における人間の演奏制御モデルを推定することを目的とする．

本論文では 2 章で本研究の位置付けについて説明し，3 章で演奏収録の概要について説明する．4 章で伴奏者の演奏の分析および演奏制御モデルの推定を行い，5 章では独奏者の演奏の分析および演奏制御モデルの推定を行う．最後に 6 章でまとめと今後の課題について述べる．

2. 伴奏システムと演奏制御モデル

従来，伴奏システムが人間の独奏者と協調して演奏する場合の協調制御手法についてはアドホックな方法が用いられてきたが，人間らしい応答をする伴奏システムを実現するためには，人間の演奏制御方法を推定する必要がある．

一般に人間が演奏を行う場合，拍に相当する一定音価の内的クロック☆に基づいて演奏を行っていると考えられる．合奏においては，お互いにこの内的クロックに従い演奏を進めながら，相手のクロックを推定し，ずれないで協調するように自分のクロックを調整していると考えられる．このような考えに基づく伴奏システムを実現するためには，人間同士の合奏において，このような内的クロックの協調動作を分析し，その演奏制御モデルを推定する必要がある．そこで本研究では，単純な八分音符だけからなる合奏曲☆☆を用意し，人間同士の合奏を数多く収録し，その合奏における協調動作を分析することにより，人間の演奏制御モデルを推定することを目的とする．

3. 実験条件と演奏収録

3.1 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（**図 1** および**図 2**）．この曲はもともとはピアノの独奏曲であるが，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，独奏者，伴奏者とも右

☆ 実際には，この内的クロックは 1 つではなく，階層性（半拍単位，2 拍単位，小節単位など）を持っているであろう．
☆☆ 内的クロックが表層に表出されることが期待される．

手で演奏するよう教示した．

この曲は最後の音符以外はすべて同じ音価の音符から構成されているため，音楽的な要素が少なく，また，独奏と伴奏の演奏が記譜上はまったく同じリズムとなるため，分析が容易であるという利点がある．この曲は本来，全曲を通じて右手パートと左手パートがオクターブで並行移動するのであるが，お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．

3.2 演奏者と収録条件

演奏者は音楽大学卒業生 8 人である．4 人ずつで 1 グループとし，各グループごとに 2 人を独奏者，残りの 2 人を伴奏者とした．独奏者はできるだけ指示されたテンポを守るように教示され，伴奏者は相手に合わせて演奏するよう教示された．

収録は 1 グループで 20 回行われたが，そのうち 4 回は人間の独奏者の代わりに計算機が独奏を行った．計算機は伴奏者の演奏に影響を受けずにあらかじめ決められたテンポ設定で演奏する．なお，独奏を計算機が行う場合があることは伴奏者には教示しておらず，また，相手が計算機であることに気付いた伴奏者もいなかった．残りの 16 回は人間同士の合奏が収録された．

3.3 収録環境

演奏の収録は千葉大学工学部情報画像工学科の音声実験室で行われた．この部屋は 2 つの防音室が隣接しており，外部および 2 つの防音室間で音を遮断することができる．一方の部屋に独奏者 2 人，もう一方に伴奏者 2 人に入ってもらった．実際の演奏はたかだか，独奏者 1 人，伴奏者 1 人によって行われるため，演奏時以外は室内の椅子で待機してもらった．

演奏者は 88 鍵の MIDI ピアノで演奏を行い，各演奏者の演奏した MIDI 情報（打鍵時刻，音程，強さ，打鍵時間）を計算機で記録すると同時に，外部の MIDI 音源で音を鳴らし，各演奏者のヘッドフォンへと伝えた．すなわち，各演奏者は自分の演奏と相手の演奏をヘッドフォンから聴くことができる．その際，演奏された音ごとに打鍵の強度が異なると相手の音の強さによって演奏が変化してしまう可能性があるため，打鍵強度にかかわらず一定の音量でヘッドフォンに出力されるように音源を設定した．また，2 人の演奏者の顔と手の動きはビデオテープに録画された．演奏者への指示は防音室の外部からマイクを通して行われた．

演奏の開始時には互いに目や音による合図を送ることができないため，冒頭の指示テンポで独奏パートの最初の音を室外から 8 拍提示し，そのまま続けて演奏を開始してもらった．

3.4 テンポ変化の指示

テンポ変化として以下の 2 パターンを用いた．単位は 1 分間あたりの拍数（beat per minute; bpm）である．

パターン I 116 → 132 → 116

パターン II 132 → 116 → 132

テンポ変化の位置はランダムに配置されたが，そのタイミングは小節の冒頭とした．テンポ変化は独奏者に対し，楽譜上に数字で指定されており，独奏者はその指定に従い，テンポ変化を行った．独奏者はそれぞれのテンポがどの程度の速さであるかを収録開始前の練習時間にメトロノームで確認した．

伴奏者には曲中にテンポ変化が生じる可能性は伝えたが，楽譜上のどの位置でどの程度の変化が起こるかは教示しなかった．また，テンポパターンの数や順番に関してもいっさい伝えなかった．

3.5 収録手順

MIDI ピアノの演奏に慣れるため，収録前に 1 人ずつ自由に練習する時間が与えられた．練習終了後，あらかじめ決められた演奏順に従って収録を進めた．

演奏ミスが生じた場合にはその演奏を再収録したが，隣の鍵盤に手が触れてしまったなどの軽微なミスは再収録せず，演奏ミスから 1 小節分を分析対象から外すことで対応した．全 40 回の合奏に対し，4 回の再収録が行われた．

3.6 収録データの傾向

図 3 および **図 4** は収録データ 1 曲における，ずれ（左側）とテンポ変化（右側）の様子を表したものである．

図 3 は計算機との合奏である．計算機の独奏は伴奏者の演奏には影響を受けずに固定されたテンポで演奏を行なえるので，指示テンポどおりの安定した演奏となっている様子が分かる．

図 4 は人間同士の合奏である．テンポはかなり速めであり，場所によっては 150 近いところもある．人間同士の合奏の場合，指示されたテンポどおりの演奏を行うことは困難であることがうかがえる．とくに今回の演奏曲目の場合，簡単な曲であったため，指定されたテンポよりも速く演奏される傾向が多く見られた．

4. 伴奏者の分析とモデルの推定

4.1 伴奏者の演奏制御推定モデル

まずはじめに分析およびモデル推定に用いる各パラメータについて，その定義を示す．独奏者が次に演奏する時刻を T_n とし，伴奏者が次に演奏する時刻を t_n とする．また，独奏者が現在までに演奏した各音の演

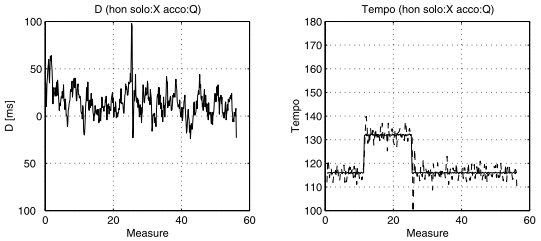

図 3　計算機との合奏の演奏例（左側：ずれ，右側：テンポ）
Fig. 3　An example of performance by human and a computer (Left-hand side shows the difference in time between them and the right-hand side shows their tempo).

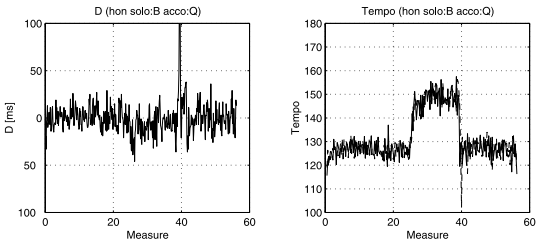

図 4　人間同士の合奏の演奏例（左側：ずれ，右側：テンポ）
Fig. 4　An example of performance by human and human (Left-hand side shows the difference in time between them and the right-hand side shows their tempo).

奏時刻の履歴を $T_{n-1}, T_{n-2}, \ldots$ とし，伴奏者の過去の演奏時刻の履歴を $t_{n-1}, t_{n-2}, \ldots$ とする．ここでそれぞれの添字は八分音符を単位とし，その値が小さいほど過去の時刻を表していることとする．本研究の目的はこれら過去の演奏情報から，未来の演奏時刻 t_n を推定することである．

推定に用いるパラメータとして，以下の 2 種類のパラメータを定義する．

ずれ（D_n および d_n）　2 人の演奏者の演奏時刻の時間的ずれを表し，独奏者の場合は $D_n = t_n - T_n$ と定義され，逆に伴奏者の場合は $d_n = T_n - t_n$ と定義される．

時間長変化（C_n および c_n）　各演奏者の演奏時間間隔の変化を表し，独奏者の場合は $C_n = L_n - L_{n-1}$，伴奏者の場合は $c_n = l_n - l_{n-1}$ と定義される．ただし L_n および l_n はそれぞれ独奏者と伴奏者の演奏時間間隔であり，$L_n = T_n - T_{n-1}$ および $l_n = t_n - t_{n-1}$ と定義される．

本研究ではこれら 2 種類のパラメータにより，未来の人間（伴奏者）の演奏を予測する推定モデルとして，以下の式を提案する．

$$c_n = \sum_{i=1}^{p} \alpha_i d_{n-i} + \sum_{j=1}^{p} \beta_j c_{n-j} \tag{1}$$

ただし式 (1) 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る．

ここで先行研究[4)]における推定手法と本手法の比較を行う．先行研究では四分音符 1 つ前（八分音符 2 つ前）の「ずれ」と「ずれの変化量」の 2 つのパラメータから四分音符単位の時間長変化を推定するものであった．先行研究における推定式と本手法の推定式では，どちらも時間長変化を推定する式となっているが，最終的には過去の情報から未来の演奏時刻 t_n を推定する．そこで t_n の推定式に関して，先行研究と本手法を比較する．まず本手法では，

$$c_n = (t_n - t_{n-1}) - (t_{n-1} - t_{n-2})$$

となることから，t_n を推定する式は以下のようになる．

$$t_n = c_n + 2t_{n-1} - t_{n-2} \tag{2}$$

一方，先行研究における推定式において，推定すべき時間長変化（四分音符単位）を c'_n とすると，先行研究の予測式は以下のように記述でき，

$$c'_n = (t_n - t_{n-2}) - (t_{n-2} - t_{n-4})$$

先行研究における t_n を推定する式は以下のようになる．

$$t_n = c'_n + 2t_{n-2} - t_{n-4} \tag{3}$$

ここで，

$$\begin{aligned} &2c_{n-1} + c_{n-2} \\ &\quad = 2\{(t_{n-1} - t_{n-2}) - (t_{n-2} - t_{n-3})\} \\ &\qquad\quad + \{(t_{n-2} - t_{n-3}) - (t_{n-3} - t_{n-4})\} \\ &\quad = 2t_{n-1} - t_{n-2} - (2t_{n-2} - t_{n-4}) \end{aligned}$$

という関係を用いて，式 (3) を変形すると

$$t_n = c'_n - 2c_{n-1} - c_{n-2} + 2t_{n-1} - t_{n-2} \tag{4}$$

となる．上述のように，先行研究は「ずれ」と「ずれの変化量」を用いる推定手法であるが，先行研究における「ずれ」は本手法のパラメータでは d_{n-2} に等しく，「ずれの変化量」は $d_{n-2} - d_{n-4}$ に等しいことを利用すると，先行研究における c'_n は

$$\begin{aligned} c'_n &= \alpha' d_{n-2} + \alpha''(d_{n-2} - d_{n-4}) \\ &= (\alpha' + \alpha'')d_{n-2} - \alpha'' d_{n-4} \end{aligned} \tag{5}$$

と表される．式 (4) と式 (5) から

$$\begin{aligned} t_n = (\alpha' + \alpha'')d_{n-2} - \alpha'' d_{n-4} - 2c_{n-1} - c_{n-2} \\ + 2t_{n-1} - t_{n-2} \end{aligned} \tag{6}$$

となる．一方，式 (1) と式 (2) から，本手法で t_n を推定する式は

$$t_n = \sum_{i=1}^{p} \alpha_i d_{n-i} + \sum_{j=1}^{p} \beta_j c_{n-j} + 2t_{n-1} - t_{n-2} \quad (7)$$

となる．式 (6) と式 (7) を比較すると，

$$\alpha_2 = \alpha' + \alpha'', \quad \alpha_4 = -\alpha''$$
$$\beta_1 = -2, \quad \beta_2 = -1$$

とし，それ以外のパラメータを 0 とすることにより，式 (6) は式 (7) と等しくなることが分かる．すなわち，本論文の提案手法である式 (1) は従来手法を包含する推定手法となっている．

4.2 履歴の長さについて

上述のように，本論文の提案手法では人間（伴奏者）の次の時間長変化 c_n を予測するために，ずれの履歴 $d_{n-i}(i = 1, 2, ..., p)$ と時間長変化の履歴 $c_{n-j}(j = 1, 2, ..., p)$ を用いるが，その係数は重回帰分析により決定される．まずはじめに最適な p を決定するため，式 (1) における p を 1 から 16 まで変化させたときの重回帰分析による重相関係数 R を**図 5** に予測誤差 E を**図 6** に示す．ここで，実測値を $\hat{c}_i$，サンプル数を N とすると予測誤差 E は

$$E = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(c_i - \hat{c}_i)^2} \quad (8)$$

と定義される．

これらの図から $p = 4$（2 拍）程度までは大きく推定精度が向上すると同時に $p = 8$（1 小節）でほぼ収束していることが分かる．すなわち，過去 2 拍程度のデータは推定において必要であり，また，過去 1 小節程度のデータがあれば，推定には十分であることが示唆される．そこで本論文では以後，過去 1 小節のパラメータの履歴から予測するモデルについて検討を行う．

4.3 説明変数の選択

重回帰分析において，一般に説明変数の数を増やせば予測性能が向上するが，本研究のようなデータでは各説明変数間に高い相関がみられるため，多重共線性の影響が生じてしまう．多重共線性が顕著に表れると不安定なモデルとなったり，偏回帰係数の信頼性が低下したりするなどの問題が生じる可能性がある．そこで本研究では使用する説明変数の選択を行う．

説明変数 x_i を目的変数として，他の全説明変数で重回帰分析したときの重相関係数を r_i とすると，r_i が高い説明変数は他の説明変数から十分に高い精度で推定できるため，その変数を除去しても予測性能はほとんど下がらない．また，r_i が 0.6 を超えると偏回帰係数の信頼性が低下することが知られている[5]ため，

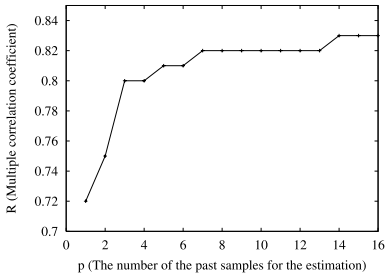

図 5 重回帰分析による重相関係数 R と p の関係
Fig. 5 Relation between R (multiple correlation coefficient) and p (the number of the past samples).

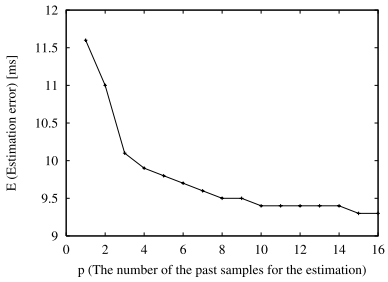

図 6 重回帰分析による予測誤差 E と p の関係
Fig. 6 Relation between E (estimation error) and p (the number of the past samples).

すべての r_i が 0.6 以下になるように r_i の大きい順に説明変数を除去した．さらにステップワイズ法（変数減増法）により，偏回帰係数の信頼性の低い説明変数を除去した．

4.4 推定結果

上述の手順を経て予測に必要な変数を絞り込み，全演奏データに対して重回帰分析を適用した結果を**表 1** に示す．

ここで比較のため，従来手法による推定結果を**表 2** に示す．計算機との合奏に関しては，今回新たに収録したデータに対しても先行研究[4]と同様，予測誤差 11.1 [ms] と高い推定精度を示している．しかしながら，人間との合奏では予測誤差 13.4 [ms] と推定精度の悪化がみられる．

一方，本研究で提案した手法では，計算機との合奏のみならず，人間との合奏においても予測誤差は 10.5 [ms] となっており，どちらの合奏タイプでも従来手法に比べ，高い推定精度が得られている☆．

☆ なお，図 5 および図 6 では表 1 よりも高い予測精度が得られているが，これは変数選択をしていないためである．変数選択をしない場合，分析に用いたデータに対して過剰に適応してしまうため，結果として，そのデータに対する予測精度は高くなるが，4.3 節で示したとおり，多重共線性が生じてしまい，一般的なデータに対する予測モデルとして用いることはできなくなってしまう．

表 1 伴奏者の演奏制御モデルで選択された変数の偏回帰係数とそのときの重相関係数 R・予測誤差 E [ms]
Table 1 The left-hand side shows the partial regression coefficient (α_i and β_j) for the selected explanatory variables by the model for accompanists and the right-hand side shows the multiple correlation coefficient R and the estimation error E [ms].

	α_1	α_2	α_3	α_4	α_5	α_6	α_7	α_8
計算機との合奏	0.63	–	–	-0.31	–	–	–	-0.11
人間との合奏	0.50	–	-0.25	–	-0.04	–	–	-0.04
	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8
計算機との合奏	-0.54	–	–	–	–	–	–	0.12
人間との合奏	-0.56	–	-0.02	–	-0.03	–	–	0.13

	R	E
計算機との合奏	0.80	10.5
人間との合奏	0.77	10.5

表 2 従来手法における偏回帰係数と重相関係数 R・予測誤差 E [ms]
Table 2 The left-hand side shows the partial regression coefficient (α_i) for the explanatory variables and the right-hand side shows the multiple correlation coefficient R and the estimation error E [ms].

	α_1	α_2	α_3	α_4	α_5	α_6	α_7	α_8
計算機との合奏	–	1.41	–	-1.01	–	–	–	–
人間との合奏	–	0.94	–	-0.52	–	–	–	–

	R	E
計算機との合奏	0.84	11.1
人間との合奏	0.73	13.4

4.5 伴奏制御推定モデルに関する考察

提案手法（表 1）を先行研究の手法（表 2）と比較すると，計算機との合奏，人間との合奏のどちらのタイプにおいても，提案手法の方が推定精度が向上している．とくに，先行研究の手法では人間との合奏における予測精度が低かったが，提案手法により予測精度の大きな向上が達成され，本手法の有効性が示されたといえよう．

ここで予測誤差に関する考察を加えておく．従来手法，提案手法とも 10 ミリ秒〜13 ミリ秒程度の予測誤差となっている．もし仮に,「すべての」データに対して，この程度の誤差で推定できるのであれば，システムの性能としては十分であると考えられる．しかしながら実際のデータでは，モデルの限界などにより，大きな推定誤差がしばしば生じてしまう．式 (8) で示したように，推定誤差の大きいものは予測誤差 E に対して二乗で影響を与えてしまうため，そのような大きな推定誤差は全体の予測誤差の悪化につながる．すなわち，本手法で 3 ミリ秒程度の予測精度向上を達成できたのは，主にこのような誤差の大きい部分の改善によるものであると考えられる．ただし，人間との実際の合奏を考えた場合，推定誤差により，大きなずれが何箇所か生じてしまった場合，そのシステムに対する人間の評価は大きく下がってしまうであろう．そのため，本手法の効果を現実的な意味で示すためには，提案手法を実装したうえで比較評価する必要があると考えられるが，本論文ではまだ実装段階には至っていないため，そのような主観評価については今後の課題とする．

次に予測に用いられた変数について考察する．まず，人間同士の合奏について，表 1 を見ると，説明変数として α_1，α_3，α_5，α_8，β_1，β_3，β_5，β_8 が選択されている．このことはずれ d と時間長変化 c がともに高い自己相関を持っていることを示唆している．すなわち，変数選択の結果から α_1 と α_2 などの連続する 2 つのパラメータ間に高い相関があることが推測される．また，ずれと時間長変化に関して，直前の情報 α_1，β_1 の偏回帰係数が極端に大きいことから，これらの変数が予測に際して最も効いていると考えられる．そのため，結果として α_1，α_3，α_5，β_1，β_3，β_5 と，とびとびに変数が選択されたと推測される．しかしながら，最後に α_8，β_8 が選択されていることは興味深い．これらの変数はちょうど 1 小節前の情報であり，音楽における小節単位の周期性が演奏制御にも暗黙のうちに利用されていることが推測される．

一方，計算機との合奏においては，説明変数として α_1，α_4，α_8，β_1，β_8 が選択されている．α_1，β_1 および α_8，β_8 に関しては，人間との合奏と同様の考察が導けるが，次に選択された変数が α_4 であることから，おおむね 2 拍程度の単位における情報を用いて演奏制御を行っていることが推測される．すなわち，計算機の演奏はテンポが固定であるため，人間同士のように局所的に不安定になることがなく，より少ない情報で予測が可能であることが推測される．

また，計算機との合奏において α_1 と α_2 の相関が高いとすると，α_1 の代わりに α_2 を用いてもそれほど推定精度は低下しないことが予測される．このことから α_2，α_4 を用いた従来手法は，提案手法によって

表 3 独奏者の演奏制御モデルで選択された変数の偏回帰係数とそのときの重相関係数 R・予測誤差 E [ms]
Table 3 The left-hand side shows the partial regression coefficient (α_i, β_j and $gamma$) for the selected explanatory variables by the model for soloists and the right-hand side shows the multiple correlation coefficient R and the estimation error E [ms].

	α_1	α_2	α_3	α_4	α_5	α_6	α_7	α_8	γ	R	E
独奏 ($k=1$)	0.21	–	–	-0.02	–	–	–	-0.04	1.09	0.87	10.0
独奏 ($k=2$)	0.21	–	–	-0.04	–	–	–	-0.04	0.46	0.87	10.0
独奏 ($k=4$)	0.34	–	-0.26	–	0.10	–	–	-0.06	0.18	0.85	10.6
独奏 ($k=8$)	0.38	–	-0.25	–	0.04	–	–	-0.03	0.07	0.84	10.8
伴奏	0.39	–	-0.27	–	0.03	–	–	-0.04	–	0.83	11.0
	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8			
独奏 ($k=1$)	–	0.05	–	0.06	–	0.07	–	0.14			
独奏 ($k=2$)	-0.60	–	–	0.06	–	0.07	–	0.14			
独奏 ($k=4$)	-0.59	–	-0.03	–	-0.09	–	–	0.17			
独奏 ($k=8$)	-0.59	–	-0.01	–	-0.08	–	–	0.20			
伴奏	-0.60	–	-0.02	–	-0.08	–	–	0.18			

選択された変数の多くを含んでいることになるため，計算機との合奏において比較的高い推定精度を得ることができたことは十分に理解できる．一方，人間との合奏では，提案手法を見る限り，より多くの説明変数を必要とするため，従来手法では低い推定精度にとどまったと推測できる．

5. 独奏者の分析とモデルの推定

5.1 独奏者の演奏制御推定モデル

4 章では収録された合奏データに基づき，計算機との合奏と人間同士の合奏における伴奏者の演奏制御モデルを推定した．本章では独奏者の演奏制御モデルの推定法について述べる．

独奏者は演奏をリードする役割を担っている．すなわち，独奏者は合奏が破綻しないよう，伴奏と合わせつつも，自分の理想的なテンポへ近付けようとすると考えられる．そこで 4 章で提案した伴奏の演奏制御モデルをもとに以下のような独奏の演奏制御モデルを提案する．

$$C_n = \sum_{i=1}^{p} \alpha_i D_{n-i} + \sum_{j=1}^{p} \beta_j C_{n-j} + \gamma^{(k)}(\tau^{(k)} - L_{n-1}^{(k)})$$

ここで α_i，β_j，$\gamma^{(k)}$ は係数である．本手法では伴奏の演奏制御モデルに対し，$(\tau - L_{n-1})$ を加えている．τ は独奏者の理想的な演奏時間間隔であり，ここでは理想時間長と呼ぶことにする．一方，L_{n-1} は独奏者の現在の演奏時間間隔であり，ここでは現在時間長と呼ぶことにする．すなわち，付加された項 $(\tau - L_{n-1})$ は理想時間長と現在時間長がずれた場合に演奏テンポを修正する働きがある．k はこの項をどのような単位として考慮するかという変数であり，本研究では以下の 4 種類を検討した．

$k=1$ 半拍（八分音符）
$k=2$ 1 拍（四分音符）
$k=4$ 2 拍（二分音符）
$k=8$ 4 拍（全音符・1 小節）

τ は本来は独奏者が用いる定数であり，今回の実験ではテンポ 116 [bpm] あるいは 132 [bpm] に相当する値であると考えたいところであるが，上述のように実際の演奏ではこれらの指示テンポとは大きく異なっていたため，本研究では指示テンポが一定の部分における独奏者の平均時間長を τ として用いた．

5.2 推定結果

伴奏モデルと同様に変数選択を行い，人間同士の合奏データを重回帰分析した結果を**表 3** に示す．なお，最下行は参考のために，独奏のデータに対して 4 章で提案した伴奏の演奏制御モデルを適用した場合の結果を示している．

伴奏の演奏制御モデルを用いた場合は推定精度が低くなっているが，提案手法ではとくに $k=1$ または $k=2$ において良好な結果を得ている．

5.3 独奏制御推定モデルに関する考察

γ は k が小さいほど大きくなっており，$k=1$ または $k=2$ のとき，テンポ維持の項が有効に働き，推定精度を向上させていると考えられる．一方，$k=4$ または $k=8$ ではこの項はあまり効かなくなり，選択された変数やその偏回帰係数なども伴奏の演奏制御モデルと似たような形になってきていることが見てとれる．

また，α_1 の値に着目すると，上述の伴奏者の演奏制御モデルの $\alpha_1 = 0.50$ と比べて小さくなっている．このことから，独奏者は伴奏者ほどにはずれを意識しておらず，また,「指示されたテンポを守る」という教

示に従い，相手をリードする演奏を行っていることが推測される．

一方，表 3 では $k = 1$ の場合に C_{n-1}（係数 β_1）が選択されていない．これは C_{n-1} と $(\tau - L_{n-1})$ の間に高い相関（相関係数 0.83）があるためであり，結果として γ の偏回帰係数が極端に高くなっていると考えられる．このことは実際に演奏制御モデルとして実装した場合，八分音符単位でテンポを維持しようと努めるあまり，演奏制御が不安定になることが予想される．また，人間は 1 拍あるいは 2 拍程度を演奏テンポの単位として演奏を制御していると推測されるため，本提案モデルでは $k = 1$ が独奏者の演奏制御モデルとしては適切であると考えられる．

6. おわりに

本研究では人間の合奏データを分析することにより，伴奏者の演奏制御モデルの推定および独奏者の演奏制御モデルの推定を行った．従来手法では十分に予測できなかった人間同士の合奏，および，合奏における独奏者の振舞いを本手法では予測できるようになった．

今後の課題として，本論文で提案した手法に基づく伴奏システムを構築し，実際に人間との合奏を行うことが考えられる．そのとき，本論文で用いたような同一音価の連続ではない曲に対して，本手法を適用しなければならない．そこで，人間の演奏から拍に相当する一定音価の内的クロックを検出し，それに対して本手法を適用して，システムの演奏制御を行う手法を検討したい．最終的には，実現されたシステムとの合奏が本当に人間らしいものとなっているのか，という点から，チューリングテストのような方法で評価を行いたいと考えている．

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：合奏における人間の発音時刻制御モデルの推定，情報処理学会論文誌，Vol.43, No.2, pp.260–267 (2002).
5) 小林龍一：相関・回帰分析法入門，日科技連出版社 (1972).

(平成 15 年 7 月 11 日受付)
(平成 16 年 1 月 6 日採録)

堀内　靖雄（正会員）

1967 年生．1990 年東京工業大学工学部情報工学科卒業．1995 年東京工業大学大学院理工学研究科情報工学専攻博士課程修了．博士（工学）．1995 年千葉大学工学部助手．2000 年千葉大学大学院自然科学研究科助手．2002 年同助教授．音楽情報処理，音声言語処理，福祉情報工学に関する研究に従事．訳書に『コンピュータ音楽』（C. Roads 著，共訳，東京電機大出版）．1992 年情報処理学会全国大会奨励賞，1995 年人工知能学会研究奨励賞，1996 年人工知能学会全国大会優秀論文賞．人工知能学会，音響学会，ソフトウェア科学会各会員．

坂本　圭司

1978 年生．2001 年千葉大学工学部情報工学科卒業．2003 年千葉大学大学院自然科学研究科知能情報科学専攻修了．同年株式会社日立製作所入社．現在，公共システム事業部にて電子決裁基盤，職員ポータルの開発に従事．

市川　　熹（正会員）

日立製作所中央研究所を経て現在千葉大学自然科学研究科教授．工学博士．電子情報通信学会理事，音声研究会委員長，福祉情報工学研究会委員長，日本音響学会評議員，人工知能学会理事，対話研主査等を歴任．電子情報通信学会，日本音響学会，人工知能学会，IEEE，欧州音声学会，言語処理学会，ヒューマンインタフェース学会，音声言語医学会，日本手話学会等の会員．電子情報通信学会フェロー．1989 年電子情報通信学会論文賞，1996 年度人工知能学会研究奨励賞受賞．現在，日本音響学会評議員，人工知能学会評議員，電子情報技術産業協会アクセシビリティ標準化対応専門委員会委員長等．