

計算機との合奏データによる 人間の演奏モデルの推定

坂本 圭司 堀内 靖雄 市川 熹

千葉大学 工学部 情報工学科
263-8522 千葉県 千葉市 稲毛区 弥生町 1-33
qchan@icsd4.tj.chiba-u.ac.jp

本稿では人間の協調演奏動作における制御モデルを推定する。人間の未来の演奏（テンポ変化）を1拍前の「両者の時間的ずれ」と「両者の時間長のずれ（テンポのずれ）」から決定するモデルを考え、人間と計算機（テンポ変化をあらかじめ指定できる）による演奏データに対して、重回帰分析を適用し、妥当と考えられる4つのモデルを導き出した。

Estimating the Manner of Human Performance From the Actual Performance by Human and Computer

Keiji Sakamoto Yasuo Horiuchi Akira Ichikawa

Department of Information Sciences,
Faculty of Engineering, Chiba University
1-33, Yayoi-cho, Inage-ku, Chiba-shi, Chiba, 263-8522, Japan
qchan@icsd4.tj.chiba-u.ac.jp

In this paper, we will introduce a hypothetical model of human performance. Several performances, where performers play together with a computer which can play with an accurate tempo specified beforehand, were recorded. These recorded performances were analyzed in order to find hypothetical models of human performance. We suppose that human performer decides next beat length (in the future) with 'the time lag between the computer and the performer' and 'the difference of beat length between them.' We found 4 models are appropriate from the results of multiple regression analysis of the recorded data.

1 はじめに

本研究では、伴奏システム [1] が人間らしい協調動作を行えるようにするため、人間の演奏制御のふるまいを明らかにすることを目的とする。すなわち伴奏システムが、ある時点において、それまでの過去の情報

を用い、未来の人間のふるまい（次の音符）を予測できるようにすることである。このような人間のふるまいの予測が可能になると、以下の二点で伴奏システムの性能改善を行うことができる。

- 相手の次の演奏をより正確に予測することができる。

- 計算機自身の演奏も人間のふるまいをモデル化した演奏制御方法で行うことにより、人間らしい協調演奏が可能となる。

著者らはすでに [2, 3, 4] などでは二人の人間による演奏を収録し、その分析を行ってきたが、[5, 6] などで示されているように、人間同士の演奏にはお互いの相互作用が存在し、この相互的に及ぼされる影響のため、二人の人間の制御モデルを独立に同定するまでには至っていない。そこで、二人のうちの片方を機械で置き換えることにより、[7] ではテンポ安定時における演奏者のふるまい、[8] ではテンポ変化時も含めた演奏者のふるまいを分析した。その結果、未来の人間のふるまいが過去のどの要因と相関が高いかが明らかになったので、本研究ではそれらの関係を示すモデル式を推定し評価する。

なお、人間の協調演奏動作のふるまいを分析する場合、大きく分類すると、音楽的な側面での協調動作と反射的な協調動作の二つが挙げられる。本研究では、反射的な協調動作について明らかにすることを目的としており、音楽的な意味での協調動作の分析は今後の課題である。

2 実験デザインと演奏収録

2.1 演奏曲目

Hanon 作曲「ピアノの名手になる 60 練習曲」第 5 番を収録対象とし、左手パートを計算機が演奏し、右手パートを人間が演奏することとした。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、人間と計算機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が、お互いの音を聞こえやすくするため、左手パートを 3 度上げ、6 度の並行移動とした。また、もとの楽譜は四分の二拍子で十六分音符の連続であるが、表示された速度指定よりも早く演奏されてしまう傾向があるため、十六分音符を八分音符に書き換え、拍子も四分の四拍子とした。収録に用いた楽譜を図 1 に示す。

2.2 演奏者

演奏者はピアノ専攻の大学生三名である。三名ともこの曲を演奏したことがあった。また演奏者にはあらかじめ、楽譜とともに実験の手順書を渡し、実験概要

収録対象曲

図 1: 収録曲楽譜

について説明した。本来、相手が計算機であることを教えないほうがより自然な演奏データが得られるのであるが、予備実験の結果、多くのデータを収録する過程で、相手が人間ではないことが露呈してしまう可能性が否定できず、もし、そのことが演奏者に分かってしまうと、実験条件の統制がとれなくなるため、本研究では最初から演奏者に伝える実験条件とした。

2.3 収録環境

演奏の収録は防音室で行われた。演奏者は 88 鍵の MIDI ピアノで演奏を行った。MIDI ピアノは音源を内蔵しているが、MIDI 楽器の特性の違いなどによる誤差を最小限に抑えるため、今回は MIDI ピアノの内蔵音源は利用せず、演奏者の演奏した MIDI 信号と計算機が出力した MIDI 信号をもう一つの計算機で記録すると同時に、別個に用意された MIDI 音源で音を鳴らし、それを演奏者のヘッドホンへと伝えた。人間の演奏者はヘッドホンにより、計算機の演奏と自分の演奏を聞きながら演奏を行う。そのとき、両者の演奏は左右分離して定位された。

2.4 テンポ設定

計算機の演奏には、あらかじめ以下のテンポ変化が設定された。テンポは楽譜の四分音符を単位としている。

- パターン 1: 100 → 105 → 115 → 110 → 100
- パターン 2: 105 → 100 → 115 → 105 → 110
- パターン 3: 110 → 115 → 100 → 110 → 105
- パターン 4: 115 → 105 → 100 → 115 → 100
- パターン 5: 100 → 110 → 105 → 115 → 100
- パターン 6: 110 → 115 → 100 → 105 → 110
- パターン 7: 100 → 115 → 110 → 100 → 115

テンポ変化の場所は演奏者に予測されないよう、ランダムに配置された。ただし、テンポ変化は小節の冒頭で行った。各演奏者は上述のパターンすべてを演奏することにより、演奏者一人あたり、表 1 に示す回数のテンポ変化データが得られることになる。

表 1: 演奏者一人あたりのテンポ変化の回数

テンポ変化	回数	
+15	100 → 115	4
+10	100 → 110	2
	105 → 115	2
+5	100 → 105	2
	105 → 110	2
	110 → 115	2
-5	115 → 110	2
	110 → 105	2
	105 → 100	2
-10	115 → 105	2
	110 → 100	2
-15	115 → 100	4

演奏の開始に際し、メトロノームで最初のテンポを四分音符で 8 拍呈示し、それに続けて演奏を開始するよう教示した。

2.5 収録手順

まず、電子楽器の演奏に慣れるため、演奏者は自由に練習を行った。そのあと、演奏者だけ（計算機の演奏はなし）で演奏してもらい、その演奏情報を記録した。その後、上述のすべてのパターンによる収録が行われた。

3 モデルの推定

3.1 目的変数

演奏のある時点において、それまでの過去の情報を使い、未来の人間のふるまい（次の音符）を予測する状況を想定し、人間のテンポ変化に相当する変数として、「人間の次の演奏間隔¹（時間長）が直前の演奏間隔に対して、どの程度増減したか」を予測すべき目的変数とする。本稿では便宜的に、これを「時間長変化」と呼ぶ。

3.2 説明変数

過去の何らかのパラメータから上述の時間長変化を予測するモデルを構築するため、[7, 8] では、過去のパラメータとして以下の二つと時間長変化との関係を明らかにした。

- (1) ずれ
- (2) テンポのずれ（=ずれの変化量）

(1) 「ずれ」とは、楽譜上同じ時刻に記譜されている音符が、どの程度ずれて演奏されたかを意味しており、「計算機の演奏時刻－人間の演奏時刻」と定義する。すなわち、「ずれ」が正の場合、人間が計算機よりも早いタイミングで演奏したことを意味する。

(2) 「テンポのずれ」とは、計算機の演奏間隔と人間の演奏間隔がどの程度違うのか、すなわちテンポがどの程度ずれているのかを意味しており、「計算機の時間長－人間の時間長」と定義する。これは、「現在のずれ－一つ前のずれ」と等しいため、「ずれの変化量」と呼ぶこともできる。

ここで、人間がどのような時間単位で、また過去のどの時点の情報を重視しているのかを考える。

表 2: ずれと時間長変化との相関係数

	半拍	1 拍	2 拍	4 拍
1 つ前	0.53	0.68	-0.59	0.27
2 つ前	-0.16	-0.18	-0.30	-0.04
3 つ前	0.04	-0.13	0.11	0.03
4 つ前	-0.08	-0.06	-0.02	0.01

表 2 は、全演奏者の「ずれ」と「時間長変化」との相関係数、表 3 は、「テンポのずれ」と「時間長変化」

¹演奏間隔とは図 1 における各八分音符、あるいはその 2 倍の四分音符、4 倍の二分音符、8 倍の全音符の時間間隔を指す。

表 3: テンポのずれと時間長変化との相関係数

	半拍	1拍	2拍	4拍
1つ前	0.81	0.81	0.71	0.26
2つ前	-0.23	-0.04	-0.33	-0.06
3つ前	0.14	-0.07	0.10	0.02
4つ前	-0.09	-0.13	-0.07	-0.04

との相関係数である。列項目は、時間長を半拍（八分音符ごと）、1拍、2拍、4拍と取ったときの違いで、行項目は、いくつ前のずれとの相関かということを表している。これを見ると、直前（1つ前）の情報が時間長変化を決めるときに大きなウェイトを占めているということがわかる。また、楽譜は半拍単位の音符で構成されているのに対して、人間は1拍単位の情報をより重視しているということもわかる。つまり、半拍ごとよりも、1拍ごとに演奏を修正していると考えるのが妥当であろう。

3.3 モデルの具体的な形

前項の結果を踏まえて、1拍前の「ずれ D 」と「テンポのずれ T 」から「時間長変化 Y 」を推定するモデルが以下のように書ける。

$$Y = \alpha D + \beta T \quad (1)$$

ここで、係数の α, β について述べる。人間は、ずれが大きいたまは特にずれを重視して時間長変化を決めているということが、[8] で示されている。したがって、 D の絶対値が大きいたまは D を重視し、 D の絶対値が小さいときは T を重視するように α と β を決定することにより、より人間に近い近似モデルが得られるであろう。そこで、 α, β を式 (2) のような D の1次関数として与える。これはそれぞれ図2のような関数である。

$$\alpha = a|D| + b, \quad \beta = c|D| + d \quad (2)$$

これらを式 (1) に代入すると、以下の式が得られる。

$$Y = a|D|D + bD + c|D|T + dT \quad (3)$$

この他にも、

- β のみが定数の場合 ($c = 0$)
- α のみが定数の場合 ($a = 0$)
- α も β も定数の場合 ($a = c = 0$)

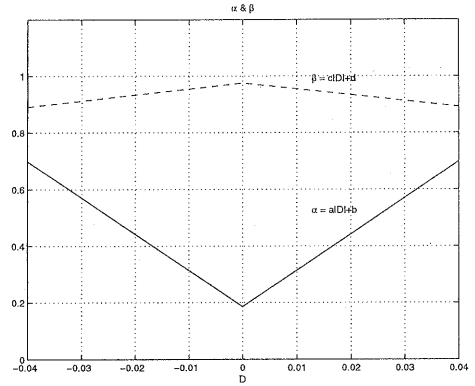


図 2: D の一次関数 α, β

- β が 0 の場合 (α は定数)
- α が 0 の場合 (β は定数)

のモデルを考え、これらの組み合わせをすべて重回帰分析により求める。全モデルを表4にまとめて示す。

表 4: モデル

モデル番号	式
i	$\alpha = a D + b$ $\beta = c D + d$
ii	$\alpha = a D + b$ $\beta = d$
iii	$\alpha = b$ $\beta = c D + d$
iv	$\alpha = b$ $\beta = d$
v	$\alpha = b$ $\beta = 0$
vi	$\alpha = 0$ $\beta = d$

3.4 分析結果

表5~7は、重回帰分析により求めた回帰係数である。表8~10は、それぞれの重相関係数 R 、AIC、予測誤差である。

AIC(Akaike Information Criterion: 赤池情報量基準)とは、あてはまりのよさと、説明変数の少なさを

表 5: 演奏者 A の回帰係数

モデル	a	b	c	d
i	12.84	0.19	-2.11	0.98
ii	10.90	0.22		0.95
iii		0.41	8.94	0.86
iv		0.49		1.00
v		0.99		
vi				1.25

表 6: 演奏者 B の回帰係数

モデル	a	b	c	d
i	10.41	0.42	-3.50	0.97
ii	6.73	0.47		0.93
iii		0.59	4.23	0.89
iv		0.63		0.95
v		1.23		
vi				1.27

表 7: 演奏者 C の回帰係数

モデル	a	b	c	d
i	9.75	0.43	-4.00	0.91
ii	5.82	0.49		0.87
iii		0.60	4.29	0.83
iv		0.64		0.89
v		1.17		
vi				1.21

表 8: 演奏者 A の重相関係数 R, AIC, 予測誤差 E

モデル	R	AIC	E [s]
i	0.87	-9612.1	0.0111
ii	0.87	-9613.2	0.0111
iii	0.86	-9583.0	0.0112
iv	0.86	-9510.3	0.0115
v	0.64	-8289.2	0.0170
vi	0.81	-9117.3	0.0130

表 9: 演奏者 B の重相関係数 R, AIC, 予測誤差 E

モデル	R	AIC	E [s]
i	0.87	-9659.8	0.0109
ii	0.87	-9659.3	0.0109
iii	0.87	-9646.2	0.0110
iv	0.86	-9631.7	0.0110
v	0.70	-8551.6	0.0156
vi	0.81	-9163.1	0.0128

表 10: 演奏者 C の重相関係数 R, AIC, 予測誤差 E

モデル	R	AIC	E [s]
i	0.88	-10106	0.0095
ii	0.88	-10104	0.0095
iii	0.88	-10090	0.0095
iv	0.88	-10067	0.0096
v	0.72	-8909	0.0139
vi	0.81	-9450	0.0117

考慮して最適なモデルを決定するための指標²であり、一般にその値が小さいほどよいモデルとされている。

予測誤差 E とは、ここでは予測値と実測値との平均平方和の 2 乗根をとったものであり、以下の式で定義する。

$$E = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - Y_i)^2} \quad (4)$$

ただし、N はデータ数、y は時間長変化の実測値、Y は重回帰式による時間長変化の予測値である。

4 考察

実際の人間の演奏にはゆらぎがあり、テンポが一定であっても微小ながら必ず相手とのずれが生じている。そのゆらぎの値よりも予測誤差が小さければ予測の精度が人間のゆらぎ以下ということになり、そのモデルの妥当性を検討する上で、参考になると考えられる。そこで、ゆらぎをテンポ安定部の D の標準偏差と定義する。また、テンポ変化前 32 拍をテンポ安定部とする。これは、ランダムに配置したテンポ変化の最小

²一般に、説明変数を増やせばモデルのあてはまりはよくなるが、複雑で現実的でないモデルになってしまう。AIC は、説明変数ができるだけ少なく、あてはまりがよいモデルを決定するとき用いられる経験的な手法である。

間隔が40拍であることと、実際にDの標準偏差が32拍あたりで最小となったことによる。表11にその値を示す。

表 11: 人間のゆらぎ [s]

	演奏者 A	演奏者 B	演奏者 C
ゆらぎ	0.0147	0.0124	0.0127

これと表8~10を比較すると、モデルi~ivにおいて予測誤差がゆらぎの範囲内に収まっていることがわかる。

また、モデルi~ivは重相関係数Rや、AICの値についても安定しており、モデルとして妥当であるといえることができる。しかし、これらの値の差は非常に小さいものであり、これらの多変量解析的手法による情報だけでは、どのモデルが最適であるかを判断することはできない。

モデルvは「ずれ」のみによる制御モデルであるが、精度に問題があり採用することはできない。また、モデルviは「テンポのずれ」のみによるものであるが、「テンポのずれ」のみによる制御であると、しばらく経過した後テンポは同じ値になるもの、ずれがある一定値のまま演奏が継続することになることが[8]で指摘されており、制御モデルとして不適切であるといえる。

以上を踏まえると、現段階ではモデルi~ivが妥当であるといえる。さらに、これらのうちのどれが最適であるかは、他の評価方法を必要とする³。どういった評価方法をとればよいのかは、今後検討すべき課題である。

5 おわりに

本研究では人間の協調演奏動作における制御モデルを推定するため、機械と人間による演奏を収録し、得られたデータにより重回帰分析を行った。[7, 8]で得られた分析結果により、「ずれが大きいときは特にずれを重視する」モデルを考え、他のモデルと比較した。しかし、「ずれ」と「テンポのずれ」を両方用いているモデルであれば妥当であることがわかったが、それらのモデルのうちのどれが最適であるかを決定するまでにはいたらなかった。

今回は、各種モデルを統計的に評価したのみであり、推定モデルによる伴奏と、人間の伴奏を比較する実験

など、心理評価実験により、最適なモデルを決定したい。また、結果を見ると演奏者により個人差があり、演奏曲によっても違ってくるため、さらに多くの演奏者による、さまざまな曲のデータを集め、より一般的に成り立つモデルを確立していきたい。

参考文献

- [1] 堀内靖雄. 自動伴奏. 長島洋一・橋本周司・平賀譲・平田圭二(編), コンピュータと音楽の世界—基礎からフロンティアまで—(bit別冊). pp.252-269. 共立出版, July 1998
- [2] 堀内靖雄, 三井卓, 井宮淳, 市川熹. 二人の人間による演奏の収録と分析. 情報処理学会 研究報告, Vol.96, No.53, pp.21-26, May 1996. 音楽情報科学 研究報告 No.15 於 広島文化女子短大.
- [3] 堀内靖雄. 二人の人間による協調演奏動作の分析. 日本認知科学会 学習と対話研究分科会資料 SIGLAL96-2, pp.1-10, January 1997. 於 東京工業大学
- [4] 堀内靖雄, 三井卓, 財津茜, 市川熹. 二人の人間による演奏の協調動作について. 情報処理学会 研究報告, Vol.98, No.74, pp.103-108. August 1998. 音楽情報科学 研究報告 No.26 於 北海道大学.
- [5] 井川孝之, 直井邦彰, 大照完, 橋本周司. 相互作用モデルによる実時間適応自動伴奏とその動作解析. 1990年電子情報通信学会春季全国大会講演論文集, p.7.216, 1990.
- [6] 澤田秀之, 磯貝昌幸, 橋本周司, 大照完. 音楽演奏における人間と機械の協調動作について. 情報処理学会第44回(平成4年前期)全国大会講演論文集, pp.1.389-1.390, 1992
- [7] 堀内靖雄, 財津茜, 市川熹. 人間の演奏制御モデルの推定(1)~テンポ安定時における人間のふるまい~. 情報処理学会 研究報告, Vol.99, No.51, pp.59-64, May 1999. 音楽情報科学 研究報告 No.30 於 沖縄県立芸術大学.
- [8] 堀内靖雄. 人間の演奏制御モデルの推定(2)~テンポ変化時における人間のふるまい~. 情報処理学会 研究報告 Vol.99 No.68, pp.73-78, August 1999. 音楽情報科学 研究報告 No.31 於 国民宿舎水郷.

³ 実用的には計算量の少ないモデルivが有効であろう。