

GAを用いたピアノ演奏の局所テンポ解析

山内敏彰 野池賢二 野瀬隆 乾伸雄 小谷善行 西村恕彦
東京農工大学工学部電子情報工学科

人間味のある演奏を計算機にさせるために、楽譜と人間の演奏との関係を学習するさまざまな研究が行われている。本稿では、重回帰分析と遺伝的アルゴリズムを用いて、局所的なテンポの学習を行う方法を述べる。まず、演奏テンポに影響を与えると思われる楽譜中のさまざまな要因を、評価要素群として定義した。それらの評価要素群と、そのなかの二つの評価要素を掛けたものを説明変数とし、局所テンポを目的変数とした重回帰分析を行う。この際、遺伝的アルゴリズムを用いて、自由度修正済の決定係数が高くなるように、実際に用いる説明変数を選ぶ。その結果、自由度修正済決定係数が約 0.5 となる学習が行われた。

Analysis of Local Tempos in Piano Performance Using a Genetic Algorithm

Toshiaki YAMAUCHI, Kenzi NOIKE, Takashi NOSE, Nobuo INUI,
Yoshiyuki KOTANI, Hirohiko NISIMURA

Department of Computer Science
Tokyo University of Agriculture and Technology

For learning the relations between human performances and scores, various methods have been researched. In this paper, we propose a learning method of local tempos using multiple the regression analysis and the genetic algorithm. First, we define various factors that affects the local tempos as evaluation elements group. In the multiple regression analysis, we use the evaluation elements group multiplied two elements among them as a predictor variable, and local tempos as a criterion variable. For the analysis, we are using the genetic algorithm, and we chose a predictor variable so that a coefficients of determination are high. As a result, we obtained accuracy of learning result which coefficients of determination are about 0.5.

1 はじめに

本稿では、人間味のある自動演奏を行うために、楽譜と演奏の関係を学習することを目的とする。その中でも、楽譜の中で最も基本となる音符列、それらの繰返しと、局所的なテンポとを扱う。重回帰分析を用いた手法は、青野ら[1]のものがあるが、我々は、重回帰分析を基に、説明変数を遺伝的アルゴリズム（以下GA）によって選ぶ方法を用いた。

2 楽譜情報と演奏情報

局所的なテンポを扱うためには、演奏情報と楽譜情報を対応させなければならない。また、局所テンポの、楽譜情報からの影響を学習するのが目的なので、演奏情報と楽譜情報を1対1に対応させる。

2.1 楽譜情報

対象とする楽譜は、単旋律で、和音がないもの

とする。問題を簡単にするために、人間が演奏するときに判断する要因を少なくする。そのために扱う記号は、音符、休符、タイなど、楽譜の中でも基本的な要素のものとする。強弱記号や速度記号などは扱わない。ここで、タイでつながった音符は一つの音符として考え、他の楽譜情報は、音符を修飾していると考え。そして、休符は、直前の音符に付属するものとして考える。以下、これを NOTE とする。一つの NOTE は、次の情報をもつ。i は、NOTE の番号を表す。

- 音高 Pit(i)
鍵盤楽器の中央ハの鍵盤を 60 とし、半音上がるごとに +1、半音下がるごとに -1 した値。
- 音符間 Iv1(i)
全音符を 4320 としたときの、NOTE の始まりから、次の NOTE の始まりまでの長さの値。曲の最後の NOTE のときは、終止線までの長さを用いる。四分音符は 1080 になる。
- 音長 Len(i)
音符の長さ。値は Iv1 と同様に全音符を 4320 とした値である。
- 拍の強さ Met(i)
拍子と、次の小節中の位置から計算される。1 小節の拍数が BAR_LEN のときの拍の強さは、表 1 のようになる。

表 1 拍の強さ

小節中の位置	0	BAR_LEN/2	BAR_LEN-1
拍子			
2	3	1 2	2 1
3	3	1 2 1	2 1
4	3	1 2 1 2	1 2 1
6	3	1 1 2	1 1

- 小節中の位置 Pos(i)
先頭の音符から、その NOTE が始まるまでの長さ。小節の先頭は 0 になる。1 小節の長さは拍子と関係があり、m 分の n 拍子だとすると、1 小節の長さは、

$$\text{BAR_LEN} = 4320 \times \frac{n}{m} \quad \dots(1)$$

となる。

2.2 演奏情報

対象とする演奏情報は、MIDI 楽器鍵盤による演奏とする。音色は、ピアノ音（減衰音）である。しかし、ペダルは扱わない。

一つの鍵盤の打鍵-離鍵動作を、ストロークと呼ぶことにする。ストロークを次の四つの情報としてとらえる。

- キーナンバー(key)
2.1 の音高と同じ。
- ステップタイム(Step)
次の打鍵までの時間。単位は 1/200 秒。
- ゲートタイム(Gate)
打鍵-離鍵間時間。単位はステップタイムと同じ 1/200 秒。
- ベロシティ(Vel)
打鍵の速さ。1 ~ 127 までの整数値。

2.3 演奏テンポ

演奏のテンポは、演奏情報だけからではわからないので、楽譜情報と対応をとることにより、演奏テンポを求める。演奏テンポを表すものとして、次に述べる局所相対テンポを考える。これは、曲全体のテンポに対する、局所的なテンポの比率である。学習では、この局所相対テンポを教師値とする。相対的局所テンポ Tp_lr は、次の式で求める。

Th(i) : i 番目のストロークの打鍵時刻

Tr(i) : i 番目のストロークの離鍵時刻

Iv1(i) : i 番目の音符情報の音符間

Len(i) : i 番目の音符情報の音長

n : 1 曲中の音符情報の数

Tp_all(i) (全体のテンポ)

$$= \frac{\sum_{i=1}^n \text{Iv1}(i)}{\text{Tr}(n) - \text{Th}(1) + (\text{Tr}(n) - \text{Th}(n)) \frac{\text{Iv1}(n)}{\text{Len}(n)}} \quad \dots(2)$$

Tp_loc(i) (局所テンポ)

$$= \frac{\text{Iv1}(i)}{\text{Tr}(i+1) - \text{Th}(i)} \quad \dots(3)$$

$$\text{Tp_lr}(i) = \frac{\text{Tp_loc}(i)}{\text{Tp_all}} \quad \dots(4)$$

全体のテンポ、局所テンポは、その値が 1 のとき、

11 bpmになる。

図1の楽譜の楽譜情報、この楽譜を演奏した演奏情報、局所相対テンポの例を、表2に示す。



図1 Largo (Telemann) の楽譜の一部

表2 楽譜情報、演奏情報、局所相対テンポ

i	楽譜情報					演奏情報					局所相対テンポ
	Pit	Ivl	Len	Met	Pos	Key	Step	Gate	Vel	Tr	
0	64	540	540	1	2700	64	73	82	60	0.8896	
1	71	810	810	3	0	71	109	113	60	0.8937	
2	72	270	270	0	810	72	27	31	62	1.2026	
3	71	540	540	1	1080	71	63	66	62	1.0308	
4	76	810	810	2	1620	76	103	111	64	0.9457	
5	74	270	270	0	2430	74	26	34	62	1.2489	
6	72	540	540	1	2700	72	67	48	57	0.9693	
7	72	810	810	3	0	72	99	100	62	0.9840	
8	70	270	270	0	810	70	28	34	52	1.1597	
9	71	1620	540	1	1080	71	201	101	64	0.9693	
10	68	540	540	1	2700	68	69	77	64	0.9412	
11	69	810	810	3	0	69	104	106	67	0.9366	

2.4 評価要素

局所相対テンポに影響すると思われる要因を評価要素として定義する。また、楽譜情報以外に、演奏者が感じたフレーズの切れ目も評価要素として定義し、これをフレーズ情報とする。この、フレーズ情報を含めた評価要素群と、フレーズ情報を含めない評価要素群とを定義する。フレーズ情報以外の評価要素は、2.1の楽譜情報から定義される。

2.4.1 音高 (評価要素番号: 0~10)

Pit(i) を評価要素として定義する。Pit(i-1) は直前の NOTE のものである。Pit(i) だけでなく、前後の Pit を含めることにより、これらの評価要素全体として、前後の Pit の差や、対象としている音符情報の Pit の周辺に対する相対的な値を表すことができる。

また、仮平均との差で表すことにより、この評価要素を二乗したものと、二乗しないものとの間で、差がでるので、より表現が豊かになる。仮平均は 72 とした。

2.4.2 休符の長さ (評価要素番号: 11~21)

NOTE の休符の長さを評価要素として定義する。

休符の長さは、Ivl(i) - Len(i) となる。音高同様、前後の音符情報の休符の長さも考える。

2.4.3 音符間の長さ (評価要素番号: 22~32)

Ivl(i) を評価要素として定義する。音高、休符の長さ同様、前後の音符情報の音符間も考える。

2.4.4 音符間比 (評価要素番号: 33~42)

対象としている Ivl(i) と前後の Ivl の比を評価要素として定義する。音符間の差は、2.4.3の音符間の評価要素を線形に組み合わせれば表現できるが、比は線形式では表現できないため、評価要素として用意した。また、二つの音符の長さを比べると、一般的に音符の長さの割合で扱うことが多いことも、この評価要素を定義した要因の一つである。

2.4.5 拍の強さ (評価要素番号: 43~45)

Met(i) を評価要素として定義する。対象とする音符情報の前後の拍の強さも定義する。

2.4.6 小節中の位置 (評価要素番号: 46)

音符情報の小節中の位置を評価要素として定義する。

2.4.7 繰返し値

(評価要素番号: 47~94)

旋律の繰返し値を評価要素として定義する。繰返し値を求めるために、繰返しの先頭を見つけることを考える。

音高 Pit(i) (0 ≤ i < n) について考える。n を1曲中の NOTE の数とする。繰返しの先頭を表す評価値を head(i) とする。任意の音符番号 i, j において、x=0 として、条件

$$\text{Pit}(i+x) = \text{Pit}(j+x) \quad \dots(5)$$

が満たされている間、x を1ずつ増やしたときの x の最大値を same_len(i, j) とする。そして、tri(i) を次のように定義する。

$$\text{tri}(i) = \max_{\sum_{j=0}^{\text{same_len}(i,j)} \dots} \frac{\text{Ivl}(i+x) + \text{Ivl}(j+x)}{2} \quad \dots(6)$$

tri(i) は、図2の①のような波形になる。この波形の垂直部分が繰返しの先頭を表すので、これを head(i) とする (図2②)。

$$\text{head}(i) = \begin{cases} \text{tri}(i) & (\text{if } \text{tri}(i-1) < \text{tri}(i)) \\ 0 & (\text{else}) \end{cases} \dots(7)$$

この、head(i) の値を、三つのランクに分ける。こうすることにより、大きな繰返しの先頭と、小さい繰返しの先頭を分けて考えられる。このしきい値を、head_rank(x) とする。

$$\text{head_rank}(x) = \frac{x}{4} \max_i \text{head}(i) \dots(8)$$

head(i) を head(x, i) にランク分けする (図2③)。

$$\text{head}(x, i) = \begin{cases} \text{head}(i) & (\text{if } \text{head}(i) > \text{head_rank}(x)) \\ 0 & (\text{else}) \end{cases} \dots(9)$$

head(x, i) において、 $a < i$ で $\text{head}(x, a) \neq 0$ となる最大の a と、 $b > i$ で $\text{head}(x, b) \neq 0$ となる最小の b を考える。このとき

$$\text{loop_head}(x, i, a, b) = \begin{cases} \frac{2}{D(a, b)} D(a, i) - 1 & (\text{if } a \leq i \leq b) \\ 0 & (\text{else}) \end{cases} \dots(10)$$

$$\text{ただし、} D(a, i) = \sum_{j=a}^i |v_l(j)|$$

を、繰返し値とする (図2④)。これは、a と b または、c と d で、-1 から 1 まで、a または c からの拍的な距離によって、線形に増加する。

繰返しの終わりの部分は、繰返しのはじめの部分の求め方を、後ろからやればよい。

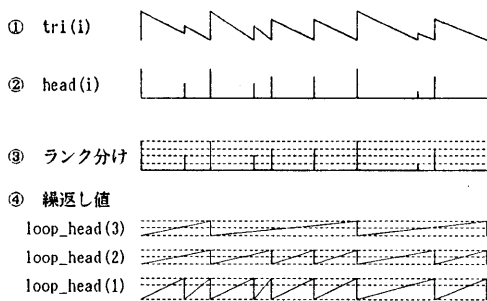


図2 繰返し値の求める手順

繰返し値について、Pit(i) を例にして述べてきたが、これを $lv_l(i)$ に対しても、同様のこと

を行う。また、(5)式 の条件を

$$\text{Pit}(i+x) = \text{Pit}(j+x) \text{ かつ } lv_l(i+x) = lv_l(j+x) \dots(5)'$$

にすることによって、二つの要素を考慮した繰返しにも適応できる。

2.4.8 曲中の位置

曲中の位置として、曲の先頭を a、曲の終わりを b としたときの(10)式を、評価要素として定義する。

・演奏者によるフレーズの切れ目

演奏者が感じるフレーズの切れ目から、フレーズの切れ目を表現した評価要素を定義する。演奏者が楽譜に記入するフレーズの切れ目はフレーズの切れやすさに相当する値をもっている。この値を、フレーズの切れ目の評価値とする。その値を(7)式のように扱って、繰返し値の計算方法と同様に、(10)式を用いてフレーズの切れ目の評価要素とする。フレーズの切れ目の評価値は、1 から 4 までとする。

この、フレーズ情報を含まない評価要素を評価要素群 1、フレーズ情報を含む評価要素を評価要素群 2 とする。

3 学習の手法

3.1 重回帰分析

重回帰分析は、対象とする目的変数が、説明変数の一次式の和で表されるという仮定が成り立つとき、最小二乗法で説明変数の係数を求めるという手法である。次の式を重回帰式と呼ぶ。

$$Y = B X \dots(11)$$

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, B = (b_0 \ b_1 \ \dots \ b_p), X = \begin{bmatrix} 1 & 1 & \dots & 1 \\ x_{11} & x_{21} & \dots & x_{n1} \\ x_{12} & x_{22} & \dots & x_{n2} \\ \vdots & \vdots & & \vdots \\ x_{1p} & x_{2p} & \dots & x_{np} \end{bmatrix}$$

Y: 目的変数行列、B: 偏回帰係数行列、X: 説明変数行列

このとき、最小二乗法により

$$B = (X'X)^{-1} X'Y \dots(12)$$

として、最適な係数が求められる。この係数を偏

回帰係数という。しかし、どの説明変数が重要であるかの評価は、このままでは使えないので、次の標準偏回帰係数 b'_i を用いる。

$$b'_i = b_i \frac{S_{yy}}{S_{xx}} \quad \dots (13)$$

$$S_{yy} = \sum_{j=1}^n (y_j - \bar{y})^2$$

$$S_{xx} = \sum_{j=1}^n (x_{ij} - \bar{x}_i)^2$$

この、重回帰分析による分析の精度を表すものに、決定係数 R^2 がある。

実際に測定された y と、(12)式で導かれた B で(11)式を計算した \hat{y} では、値が違う。実際に測定されたものを実績値、計算で求められたものを理論値と呼び、それぞれ y_i 、 \hat{y}_i で表す。

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{S_{yy}} \quad \dots (14)$$

これは、相関係数 r を二乗したものと同じである。

$$r_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \cdot \sum_{i=1}^n (y_i - \bar{y})^2}}$$

R^2 は、実績値と理論値が近ければ、値が大きくなる。しかし、説明変数の種類が増えれば、意味のない変数でも R^2 は大きくなる。そこで、説明変数とデータの個数を考慮した自由度修正済の決定係数 R'^2 を分析の精度を表す評価に用いる。 R'^2 は次の式になる。

$$R'^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2 / (n-p-1)}{S_{yy} / (n-1)} \quad \dots (15)$$

しかし、実際に演奏を近似するには、重回帰式では不十分である。評価要素を説明変数にして、局所相対テンポを目的変数にした重回帰分析を行ったが、 R'^2 は、0.4 程度しかなかった。

そこで、説明変数に、二つの評価要素を掛けたもの考える。しかし、評価要素が 100 個あるとすると、二つの評価要素を掛けたものは、5000 種類にもなってしまう。そこで、GA を用いて、目的変数を選定する。

3.2 GA

GA は、生物の進化をモデルにしたアルゴリズムである。GA の基本的な要素は、遺伝子表現、遺伝的操作、適応度である。ランダムに生成された初期集団が、競争力によって選択、淘汰されながら子孫を残し、その子孫が、また同様に子孫を残す。これを繰り返すことにより、競争力の強いものが残り、全体としても、競争力が高くなってゆく。

評価要素が、0 ~ p-1 まで p 個あると考える。p 番目の評価要素を考えて、この値を 1 とする。そして $i(i+1)/2+j$ 番目の染色体を、i 番目と j 番目の評価要素を掛けたものとした。ここで、i と j は $i \geq j$ とする。

個体の適応度には、重回帰分析を行った結果の自由度修正済決定係数を用いる。これにより、意味のない説明変数が増えることを防ぐ。

GA は、遺伝的操作として、選択、交叉、淘汰、突然変異を行う。

選択・淘汰は、個体の適応度に応じて、交叉の際、個体を選ぶことである。選ばれなかった個体は、淘汰される。

交叉は、選択された二つの個体のある部分で切断し、入れ替える作業である。

突然変異は、交叉された個体に対して、ある遺伝子を確率的に他の遺伝子に置き換えることである。

3.3 演奏テンポ学習アルゴリズム

重回帰分析と GA を用いた演奏テンポ学習の流れを図 3 に示す。

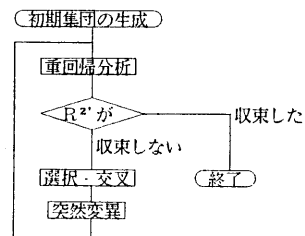


図3 学習アルゴリズム

4 演奏テンポの学習実験

3人の演奏者それぞれに、鍵盤を使って演奏してもらい、その演奏データを用いて学習を行った。演奏者ごと、評価要素群ごとに学習を行った。試験曲として、8曲分のデータを採り、そのうち、7曲を学習に用いた。この7曲を既知曲とし、残りの1曲を未知曲とする。既知曲のすべての音符の数は、879個である。

実際に R^2 がどのくらいの世代で収束するか分からないので、世代を200回とした。個体の数は、500とした。

このときの、世代と適応度について、ある演奏者の結果を図4に示す。図4のグラフから、収束していく様子がわかる。

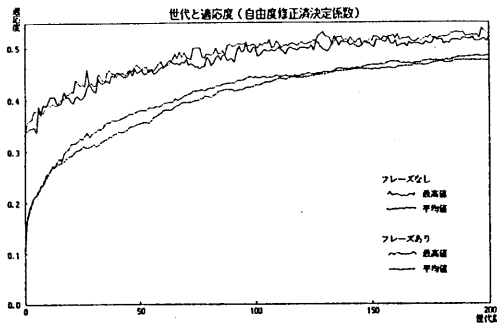


図4 世代と適応度

5 結果

重回帰分析とGAを用いて学習した結果である R^2 と $R^{2'}$ を表3に示す。演奏者1、2については、 $R^{2'}$ が0.5と、ある程度学習されていることがわかる。

表3 重回帰分析とGAの決定係数の自由度修正決定係数

重回帰分析	評価要素群1		評価要素群2	
	決定係数	自由度修正決定係数	決定係数	自由度修正決定係数
GA	世代	134	197	
	決定係数	0.557	0.576	
	自由度修正決定係数	0.517	0.538	

表4 評価要素と標準回帰係数

評価要素群1		評価要素群2	
番号	標準回帰係数	番号	標準回帰係数
45;38	-1.151	45;38	-0.693
38;38	0.430	11;07	-0.282
45;28	0.412	46;38	0.228
22;	-0.289	99;33	0.227
22;22		39;	0.212
39;34	0.171	30;29	-0.195

GAを用いた学習結果から、それぞれの評価要素

の標準偏回帰係数を求めた。ある演奏者の、この結果を標準偏回帰係数の絶対値でソートしたものの上位6個を表4に示す。この結果から評価要素番号38である音符間の比と、評価要素番号45である拍の強さを掛けた要素が、学習において重要な役割を果たしたことがわかった。

図5に、演奏者のTp_lrを実線、学習結果から再現したTp_lrを点線で示す。これより、教師値に近い演奏テンポが生成できたことがわかる。

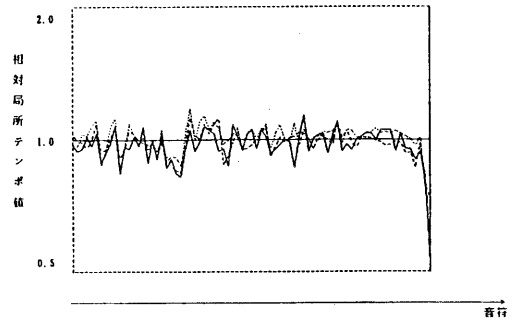


図5 相対局所テンポ値

6 おわりに

本稿では、遺伝的アルゴリズムを重回帰分析に適用する考えを用いて、人間の演奏テンポと楽譜の関係の学習させた。その結果、 $R^2 = 0.5$ となる学習が行われた。

参考文献

- [1]青野裕司、片寄晴弘、井口征士: 重回帰分析を用いた演奏ルールの抽出, 情報処理学会研究報告, Vol.95, No.74, pp.1-6, 1995
- [2]菅民郎: 多変量解析の実践(上), 現代数学社, 1933
- [3]Kenzi NOIKE et al.: Automatic Generation of Expressive Performance by using Music Structures, Proc. ICMC 1993, pp.363-365, 1993
- [4]北野宏明編: 遺伝的アルゴリズム, 産業図書, 1993
- [5]米澤保雄: 遺伝的アルゴリズム 進化理論の情報科学, 森北出版株式会社, 1993