

## ガウシアンプロセスによる名演奏の学習

寺村佳子<sup>†</sup>, 大熊秀治<sup>†</sup>, 谷口雄作<sup>†</sup>, 牧本慎平<sup>†</sup>, 前田新一<sup>††</sup>

<sup>†</sup>奈良先端科学技術大学院大学 情報科学研究科

<sup>††</sup>京都大学大学院 情報学研究科

本稿では、ガウシアンプロセスを用いた新しいピアノの演奏学習モデルを提案する。このモデルは機械学習の手法であるガウシアンプロセスを用い、楽譜の個々の音に対して、楽譜情報から抽出される音符の長さ、高さ、前後の音との相対的な差など 16 の素性を入力、対応するピアニストの演奏情報 3 要素(音の強さ、始まりと終わりのタイミング)を出力としたモデルである。本モデルは、従来手法とは異なり、天下り的なルールやルールに含まれるパラメーターの設定を必要とせず、同じピアニストが演奏した多くの演奏データから、そのピアニストの演奏を模写する学習を行えるという特長をもつ。実際のピアニストの演奏から作成された CrestMusePEDB の 4 曲を用いて評価実験を行った結果、本モデルは演奏データを反映し学習できたことが確認できた。

## Learning of Grandmaster Performance by Gaussian Process Regression

Keiko TERAMURA<sup>†</sup>, Hideharu OKUMA<sup>†</sup>, Yuusaku TANIGUCHI<sup>†</sup>,  
Shimpei MAKIMOTO<sup>†</sup>, and Shin-ichi MAEDA<sup>††</sup>

<sup>†</sup>Graduate School of Information Science, Nara Institute of Science and Technology, Japan  
<sup>††</sup>Graduate School of Informatics, Kyoto University, Japan

In this paper, we present a new approach that automatically learns a pianist performance. We use a Gaussian Process Regression (GPR) to learn the pianist's performance based on CrestMusePEDB that consists of several sets of score information and its corresponding pianist's performance information. Our model can learn the pianist's performance without generating and selecting the useful rules, nor perform complex optimization of parameters in the rules. The results show that the trained GPR has an acceptable learning ability for 'known' pieces those are included in a given training data set.

### 1. はじめに

プロのピアニストは表情豊かな演奏を行うために、楽譜上での演奏指示にはない強弱や抑揚をつけ演奏する。どのようにこれらの演奏表情付けを行うのか、どの要素が表情豊かな演奏に結びつくのかを理解することは音楽学の中の重要なテーマの一つである。演奏表情付けの問題は、表情豊かな演奏がどのように実現されるかを実証的に理解するための格好の問題といえる。

1984 年の Fryden ら(1)の発表を皮切りに、この演奏表情付け問題に対する様々な計算機モデルが提案されている。しかし、Widemer ら(2)をはじめとした既存のモデルは人手によって演奏表情付けのルールを作成する手法に基づいており、性能を向上させるためには専門的な音楽知識や技術を前提とした複雑な

ルールとそのパラメーター設定が必要であった。

近年、このようなパラメーターチューニングの労力をかけず、自動的な生成を行うべく機械学習の手法を用いたモデルの提案がなされるようになった。中でも、2007 年に発表された Dorad らによるモデル(3)はルールの探索を行う代わりに、機械学習の一種である Kernel Canonical Correlation Analysis を用いて演奏情報と楽譜情報のデータから、直接、演奏情報と楽譜情報との対応づけを学習するモデルを提案している。これは、ルールを事前に準備する必要のない画期的なモデルであったが、楽譜情報の入力時には和声の情報を人手で入力しており、多くのデータから自動で学習を行うには不適であった。

これに対し、我々は機械学習の一種であるガウシアンプロセス(4)を用いてヒューリスティックなパラメーター設定をほとんど必要としないで演奏を学

習するモデルの提案を行う。本モデルでは公開データベースである CrestMuse-PEDB (5)の楽曲を学習データとして楽譜上に書かれた指示 (楽譜情報)をもとに、人間が演奏した意識的、無意識的な変化を含んだ演奏情報をより良く予測できるよう学習を行う。これにより、人間が明示的にルール化しにくいものを含めて、楽譜情報から演奏情報を出力する入出力関数の学習を行うことができる。このようなガウシアンプロセスによる入出力関数の学習は、データがある特定の確率分布からサンプルとみなせるなどの一定の条件を満たす状況下においては、データ数が無限の極限においていくらかでも精度良く、真の入出力関数を近似できることが知られている(6)。したがって、ガウシアンプロセスである特定のピアニストのデータのみを学習することによって、そのピアニストの演奏を模倣した演奏を自動的に学習できる可能性がある。すなわち、学習データを名ピアニストの演奏とすることで、ガウシアンプロセスはそのピアニストと同様の名演奏を自動生成できるようになる可能性をもつ。

## 2. 提案モデル

我々はピアニストの演奏の模倣を入出力関数の近似問題として捉え、ガウシアンプロセスで学習することを考えた。ここで1音ごとに、楽譜から得られる情報(その音の長さ、高さ、前後関係など)を入力とし、その音を実際に演奏するための情報(強さ、始まりと終わりのタイミング)を出力とした。以降、演奏の対象となる音を演奏音と呼ぶことにする。まず、この入出力関数の入力と出力の詳細について述べ、その後、ガウシアンモデルについて説明を行う。

### 2.1 入力素性

入力素性は、CrestMusePEDB-CSR から演奏音に関係すると思われる情報を16種類、選んだ。16種類の入力素性は大きく分けると4グループ(位置情報、音高情報、音長情報、強弱情報)に分けることができる。以下、入力素性の詳細を各グループごとに説明する。

#### 2.1.1 位置情報

- 拍子：楽曲の拍子を2次元で表現する。  
例) 4/4→(4,4), 6/8→(6,8)
- 拍節：演奏音が小節内の1拍目かどうかを表す。1拍目であれば1、それ以外は0とする。
- 声部：演奏音がメロディー(主旋律)であるかどうかを示す。CrestMusePEDB-CSR では、声部を示されており、高音声部に含まれる音で最高音をメロディーとし、それ以外を0とする。

#### 2.1.2 音高情報

- 相対音高 (Relative Pitch;  $RP$ )：音高情報は CrestMusePEDB-CSR では、SMF(MIDI)の形式にのっとなって表記を行っている。ここでは、その演奏音の音高情報を調の主音に対する相対音高に変換した。変換方法は、次の通りである。C4-B4 の中で調の主音と同じ音となる音を36とする。さらに、調性を反映させるため、前後で音階上にある音どうしの差を1、前後で音階sとした。図1,2 はそれぞれい長調、い短調の場合の相対音高を示したものである。またn番目の演奏音に対する相対音高を  $RP_n$  と表す。

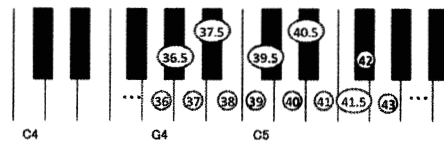


図1 ト長調の場合の相対音高

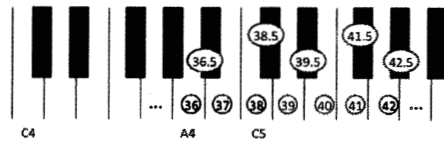


図2 い短調の場合の相対音高

- 1音前の音高：演奏音の相対音高と1つ前の相対音高の差を示したものであり、 $BNP$  で表す。 $n$ 番目の演奏音に対する  $BNP$  は次のように計算される。

$$BNP_n = RP_n - RP_{n-1}$$

- 1音後の音高：演奏音の相対音高と1つ後の相対音高の差を示したものであり、 $ANP$  で表す。 $n$ 番目の演奏音に対する  $ANP$  は次のように計算される。

$$ANP_n = RP_n - RP_{n+1}$$

- 2拍前の音高：演奏音の2つ前の相対音高を1拍前の音高( $BNP$ )を用いて表したものであり、 $B2NP$  と表記する。 $n$ 番目の演奏音に対する  $B2NP$  は次のように計算される。

$$B2NP_n = \frac{BNP_{n-1} - BNP_n}{RP_n}$$

- 1小節前の平均音高：演奏音の含まれる小節での平均音高( $AP$ )とその前の小節の平均音高の差を表したものであり、 $BMP$  と表記する。 $n$ 番目の演奏音に対する  $BMP$  は次のように計算される。

$$BMP_n = AP_n - AP_{n-1}$$

- 1 小節後の平均音高：演奏音の含まれる小節での平均音高( $AP$ )とその後の小節の平均音高の差を、1 小節前の平均音高( $BMP$ )を用いて表したものであり、 $AMP$  と表記する。  $n$  番目の演奏音の  $AMP$  は次のように計算される。

$$AMP_n = BMP_{n+1} - BMP_n$$



RP	43	43	47	47	48	48	47	47	46	46
BNP	-	0	4	0	1	0	-1	0	-1	0
ANP	0	-4	0	-1	0	1	0	1	0	-
B2NP	-	-	-4/47	4/47	-1/48	1/48	1/47	-1/47	1/46	-1/46
AP	43		47		48		47		46	
BMP	-		4		1		-1		-1	
AMP	-		-3		-2		0		-	

図 3 八長調の楽曲における音高情報の算出例

### 2.1.3 音長情報

- 音長：演奏音の音符長を示したものであり、 $D$  と表記する。四分音符を 1 とする。ただし、タイ記号がついている場合は、そのタイが終わるまでを 1 音とする。
- 1 音前の音長：演奏音と演奏音の 1 つ前の音との音長の比を表すものであり、 $BND$  と表記する。  $n$  番目の演奏音の  $BND$  は次のように計算される。

$$BND_n = D_{n-1} / D_n$$

- 1 音後の音長：演奏音と演奏音の 1 つ後の音との音長の比を表すものであり、 $AND$  と表記する。  $n$  番目の  $AND$  は次のように計算する。

$$AND_n = D_{n+1} / D_n$$

- 1 小節前の平均音長：演奏音の含まれる小節での平均音長( $AD$ )と演奏音の 1 つ前の小節の平均音長の比を示すもので、 $BMD$  と表記する。  $n$  番目の  $BMD$  は次のように計算する。

$$BMD_n = AD_{n-1} / AD_n$$

- 1 小節後の平均音長：演奏音の含まれる小節での平均音長( $AD$ )と演奏音の 1 つ後の小節の平均音長の比を示すものであり、 $AMD$  と表記する。  $n$  番目の  $AMD$  は次のように計算する。

$$AMD_n = AD_{n+1} / AD_n$$

### 2.1.4 強弱情報

- ダイナミクス：演奏音の強弱を示すものである。各強弱記号に  $f=2$ ,  $mf=1$ ,  $mp=-1$ ,  $p=-2$  のような数値を割り振る。割り振りが無い場合は、その前にある強弱記号を割り振る。

- 小節内のダイナミクス：演奏音が含まれる小節内の強弱を示すものであり、ダイナミクスと同様の数値  $f=2$ ,  $mf=1$ ,  $mp=-1$ ,  $p=-2$  を割り振る。割り振りが無い場合は、その直前にある強弱記号を割り振る。

### 2.2 出力素性

出力素性は CrestMusePEDB-DEV の規格をもとに、以下の 3 種類の演奏情報を選んだ。

- 打鍵時間 (Attack)：演奏音の演奏開始時刻と楽譜上の演奏開始時刻との差を表したものである。単位は入力素性の音長と同様に、四分音符を 1 とした。
- 離鍵時間 (Release)：演奏音の演奏終了時刻と楽譜上の演奏終了時刻との差を表したものである。単位は入力素性の音長と同様に、四分音符を 1 とした。
- 強弱 (Dynamics)：演奏音が演奏されたときの強弱を楽譜で指定される強弱との比で表したものである。数量化の記述は、CrestMusePEDB-DEV のフォーマットに従った

### 2.3 ガウシアンプロセス回帰(GPR)

本研究では、前述の入力素性と 3 つの出力素性の間の入出力関係をガウシアンプロセス回帰(Gaussian Process Regression; GPR)を用いて学習する。学習は、3 つの出力素性、それぞれで別個に行う。すなわち、入力素性と 3 つの出力素性の内の 1 つの出力素性との間の入出力関係の学習を、出力素性の種類だけ、つまり 3 通り行う。

本稿で提案している GPR はパラメーターを近似するノンパラメトリック法のひとつであり、入力素性間の相関を考えた出力予測が可能になる。しかし、ノンパラメトリック法は一般的に学習するために必要となるデータ量が多く、計算コストが高くなる傾向がある。そこで本提案では計算コストを減らすため、近年提案された GPR の近似手法 Bayesian Committee Machine (BCM) (8,9)を用いた。

#### 2.3.1 ガウシアンプロセス回帰の学習

GPR では、 $N$  個の入力  $\mathbf{x} = \{\mathbf{x}_1, \dots, \mathbf{x}_N\}$  に対応する出力  $\mathbf{t} = \{t_1, \dots, t_N\}$  は、真の出力  $\mathbf{y} = \{y_1, \dots, y_N\}$  にガウシアンノイズが付加されたものであるとみなす。

このとき  $n$  個の出力それぞれに含まれるガウシアンノイズ  $\epsilon_n$  は独立で同一のガウス分布に従うとすると

$$t_n = y_n + \epsilon_n \quad (1)$$

したがって真の出力  $\mathbf{y}$  が与えられたときの観測された出力  $\mathbf{t}$  の出現確率は

$$p(\mathbf{t}|\mathbf{y}) = N(\mathbf{t}|\mathbf{y}, \sigma^2 \mathbf{I}_N) \quad (2)$$

で表すことができる。(2)式の  $N(\mathbf{t}|\mathbf{y}, \sigma^2 \mathbf{I}_N)$  は、平均

$\mathbf{y}$ 、共分散  $\sigma^2 \mathbf{I}_N$  のガウス分布であることを表して

いる。ただし  $\mathbf{I}_N$  は  $N \times N$  の単位行列である。また、入力  $\mathbf{x}$  が与えられたときの真の出力  $\mathbf{y}$  の出現確率を

$$p(\mathbf{x}|\mathbf{y}) = N(\mathbf{y}|\mathbf{0}, \mathbf{K}_N) \quad (3)$$

とする。このときの  $\mathbf{K}_N$  は  $N \times N$  のグラム行列で

あり、 $(m, n)$  番目の要素  $k_{m,n}$  は

$$k_{m,n} = k(\mathbf{x}_m, \mathbf{x}_n) = \frac{a}{(1 + (\mathbf{x}_m - \mathbf{x}_n)^T B (\mathbf{x}_m - \mathbf{x}_n))^c} \quad (4)$$

とする。ここで、 $a, c$  は正のスカラー、 $B$  は対角行列であり各対角要素は非負の値をとるものとする。 $T$  はベクトルの転置を表す。このカーネルパラメ

ーター  $a, B, c$  と前述の  $\sigma^2$  は第二種の最尤推定によって学習した

### 2.3.2 ガウシアンプロセスの推定

未知の入力  $\hat{\mathbf{x}}$  を与えたときに、真の出力  $\hat{\mathbf{y}}$  を推定する。

ベイズ推定より、未知の入力  $\hat{\mathbf{x}}$  を与えたときの真の出力  $\hat{\mathbf{y}}$  の出現確率は、次の(5)式のようになる。

$$\begin{aligned} p(\mathbf{y}^*|\mathbf{x}^*, D) &= \frac{p(\mathbf{y}^*, \mathbf{t}|\mathbf{x}^*, \mathbf{x})}{p(\mathbf{t}|\mathbf{x}^*, \mathbf{x})} \\ &= \frac{\int p(\mathbf{t}|\mathbf{y}) p(\mathbf{y}^*, \mathbf{y}|\mathbf{x}^*, \mathbf{x}) d\mathbf{y}}{\iint p(\mathbf{t}|\mathbf{y}) p(\mathbf{y}^*, \mathbf{y}|\mathbf{x}^*, \mathbf{x}) d\mathbf{y} d\mathbf{y}^*} \end{aligned} \quad (5)$$

(2), (3)式より、この確率はガウシアンとなり、

$$p(\mathbf{y}^*|\mathbf{x}^*, D) = N(\mathbf{y}^*|\mathbf{m}, \Sigma) \quad (6)$$

と表すことができる。このガウシアン平均  $\mathbf{m}$  から、真の出力  $\hat{\mathbf{y}}$  の期待値は

$$\begin{aligned} E[\mathbf{y}^*|\mathbf{x}^*, D] &= \mathbf{m} \\ &= \mathbf{K}_N^* (\mathbf{K}_N^* + \beta^{-1} \mathbf{I}_N)^{-1} \mathbf{t} \end{aligned} \quad (7)$$

となる。ただし、 $\mathbf{K}_N^*$  は  $T \times N$  のグラム行列で、

$(m, n)$  番目の要素  $k_{m,n}^*$  は  $k_{m,n}^* = k(\hat{\mathbf{x}}_m, \mathbf{x}_n)$  となる。

しかし、(7)式を計算しようとする  $N \times N$  の逆行

列を計算する必要があり、計算コストが大きくなってしまう。

そこで、BCM では以下の近似を行う。入出力データを  $M$  個に分割し、(7)式を(8)式のように各部分データの共分散行列を用いて近似する。これにより、

$$\text{計算は } \frac{N}{M} \times \frac{N}{M} \text{ の逆行列の計算を } M \text{ 回行うだけ}$$

で済み、計算量を削減できる。

$$\begin{aligned} E[\mathbf{y}^*|\mathbf{x}^*, D] \\ \approx C^{-1} \sum_{i=1}^M \text{cov}(\mathbf{y}^*|\mathbf{x}^*, D^i)^{-1} E[\mathbf{y}^*|\mathbf{x}^*, D^i] \end{aligned} \quad (8)$$

ただし、 $D, C, \text{cov}(\mathbf{t}|\hat{\mathbf{y}})$  は、それぞれ

$$\begin{aligned} D &= \{D^1, \dots, D^M\} \\ C &= -(M-1)(\mathbf{K}^*)^{-1} + \sum_{i=1}^M \text{cov}(\hat{\mathbf{y}}|\hat{\mathbf{x}}, D^i)^{-1} \\ \text{cov}(\mathbf{t}|\hat{\mathbf{y}}) &\propto \text{cov}(\mathbf{t}|\hat{\mathbf{x}}, \mathbf{x}, \hat{\mathbf{y}}) \\ &= \mathbf{K}_N + \beta^{-1} \mathbf{I}_N - \mathbf{K}_N^r (\mathbf{K}_N^r)^{-1} \mathbf{K}_N^* \end{aligned}$$

## 3. 評価実験

### 3.1 使用データ

前述のモデルを CrestMusePEDB ver.2.0 の楽曲を用いて作成し、評価を行った。この CrestMusePEDB は楽譜情報を musicXML 形式で、演奏情報を deviational musicXML 形式で作成したデータベースである。本実験では CrestMusePEDB ver.2.0 で公開されているデータのうち、W.A.Mozart 作曲のピアノソナタから6曲を用いて実験を行った。使用楽曲は表1のとおりである。

### 3.2 実験条件

提案モデルの性能を評価するため、次の2種類の実験を行った。

#### 実験1 学習性能評価 (以下、学習実験)

提案モデルの学習性能を評価するため、4曲で提案モデルの学習を行い、その中の1曲を選び、その実際の演奏と学習した提案モデルによる演奏とを比較した。以降、学習に用いる楽曲データを学習データ、評価に用いる楽曲データをテストデータと呼び、提案モデルによる演奏を学習演奏、ピアニストによる実演奏を単に実演奏と呼ぶ。この実験で用いた学習データ・テストデータの組み合わせは表2のとおりである

#### 実験2 汎化性能評価 (以下、汎化実験)

提案モデルが学習で使用していない未知のデータに対してピアニストに近い演奏が行えるか評価するため次のような実験を行う。4曲で提案

モデルの学習を行い、その中に含まれない楽曲 1 曲を選び、実演奏と学習演奏とを比較した。この実験で用いた学習データ・テストデータの組み合わせは表 3 のとおりである。

表 1 実験で使用した楽曲

楽曲ID	楽曲名	演奏者
11-1-N	Piano sonata No.11 K.331-I	H. Nakamura
11-1-S	Piano sonata No.11 K.331-I	N. Shimizu
11-2-N	Piano sonata No.11 K.331-II	H. Nakamura
11-3-N	Piano sonata No.11 K.331-III	H. Nakamura
16-1-G	Piano sonata No.16 K.545-I	G. Gould
16-2-G	Piano sonata No.16 K.545-II	G. Gould

表 2 実験 1 での学習データとテストデータの組み合わせ

実験番号	1-1		1-2		1-3		1-4		1-5	
	TR	TS	TR	TS	TR	TS	TR	TS	TR	TS
11-1-N										
11-1-S	●	●	●		●		●		●	
11-2-N	●		●	●	●		●		●	
11-3-N	●		●		●	●	●		●	
16-1-G	●		●		●		●	●	●	
16-2-G	●		●		●		●		●	●

TR にある●楽曲は学習で使用、TS にある●のある楽曲はテストで利用した楽曲である。(表 3 でも同様)

表 3 実験 2 での学習データとテストデータの組み合わせ

実験番号	2-1		2-2		2-3		2-4	
	TR	TS	TR	TS	TR	TS	TR	TS
11-1-N		●	●		●		●	
11-1-S	●			●	●		●	
11-2-N	●		●		●		●	
11-3-N	●		●			●	●	
16-1-G	●		●		●			●
16-2-G	●		●		●		●	

### 3.3 評価方法

本研究では、学習演奏が、実演奏にどれだけ近いかを定量的に評価するため、学習演奏と実演奏との誤差を正規化したものである ND 値を計算した。式の中の real output は実演奏、predicted output は学習したモデルから出力した演奏での数値をとる。また、SD は標準偏差を表す。

$$ND = \frac{SD_{error}}{SD_{real}} = \frac{SD(\text{real output} - \text{predicted output})}{SD(\text{real output})}$$

## 4. 実験結果

実験 1, 2 の結果とも それぞれテストに使用演奏データと学習したモデルから出力したデータとの数値について ND 値を算出し、モデルの精度を検証した。結果は、次項のとおりである。

### 4.1 実験 1 の結果

実験 1 に示した 5 通りのテスト楽曲のそれぞれに対して 3 つの演奏出力(強弱、打鍵時刻、離鍵時刻)の ND 値を計算した。その結果を表 4 に示す。5 つのテスト楽曲に対する ND 値の平均は、強弱、打鍵時刻、離鍵時刻のそれぞれで 28.5, 26.8, 27.2 となった。

表 4 実験結果

実験番号	ND		
	Dynamics	Attack	Release
1-1	0.2290	0.350	0.340
1-2	0.232	0.041	0.125
1-3	0.250	0.345	0.170
1-4	0.336	0.185	0.263
1-5	0.316	0.418	0.462
平均	<b>0.285</b>	<b>0.268</b>	<b>0.272</b>

### 4.2 実験 2 の結果

実験 2 に示した 4 通りのテスト楽曲のそれぞれに対して 3 つの演奏出力(強弱、打鍵時刻、離鍵時刻)の ND 値を計算した。その結果を表 5 に示す。4 つのテスト楽曲に対する ND 値の平均は、強弱、打鍵時刻、離鍵時刻で 98.9, 109.0, 88.4 となった。

表 5 実験結果

EX. ID	ND		
	Dynamics	Attack	Release
2-1	0.667	0.971	0.679
2-2	0.712	1.388	0.747
2-3	0.999	1.000	1.111
2-4	1.577	1.002	0.998
平均	<b>0.989</b>	<b>1.090</b>	<b>0.884</b>

### 4.3 2 つの実験結果からの考察

図 4 は ND 値の平均値を実験 1 と実験 2 とで比較したものである。この図から実験 1 においては誤差が十分少ないのに対し、実験 2 においては誤差が大きいままであることがわかる。これは本提案モデルが学習性能はあるが、汎化性能が十分であるとは言えないことを示している。GPR は入出力の関係を学習するものであることから、この原因は大きく分けて 2 つ考えられる。一つは、提案モデルの GPR ではそもそも入出力関係を表現できない可能性であり、もう一つは、提案モデルの GPR が入出力関係を表現できるが、学習するため

の学習データが不十分であったり、不適切であったりする可能性である。一つ目の要因としては、入力素性やカーネル関数が実際の演奏を表現するには不適切なものであった可能性や、BCMで計算量を削減するために用いられた近似のために実際の演奏を表現できなくなる可能性が考えられる。これらの可能性に関しては、今度、新たな実験を行い、検討する必要がある。

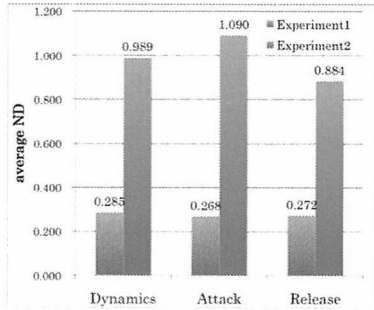


図 4 実験 1 と実験 2 の実験結果の平均値

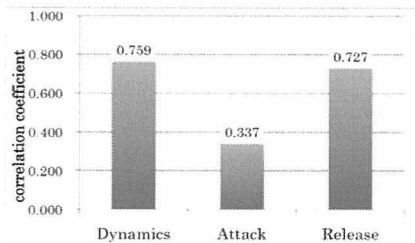


図 5 実 2-1 と実験 2-2 の結果の相関

2 点目の問題に関しては、学習データがそもそも定常な分布からの学習サンプルとはみなせないような非定常な性質をもっていた可能性と、定常分布からの学習サンプルとはみなせるが十分な汎化性能を得るためにはサンプル数が不足していた可能性が考えられる。使用した楽曲の中には異なる演奏者が演奏した楽曲も含まれていた。楽曲 ID の 11-I-N と 11-I-S は同じ楽曲の異なるピアニストによる演奏データを有する。この 2 つの演奏データ間で演奏出力値どうしの相関を表したのが図 5 である。この結果から、2 人の演奏家について強弱と離鍵時刻は相関が高く、打鍵開始時刻は相関が非常に小さいことがわかった。このことから、打鍵開始時刻に関しては学習データとテストデータとであまり一貫性がなく、同じ定常分布からのサンプルとみなすには不適切であった

ことがわかる。これが、打鍵開始時刻の汎化性能をとりわけ悪化させていた原因である可能性がある。今後、同じピアニストによる演奏データを学習させることでより演奏の模倣が正確になることが考えられる。

## 5. おわりに

本研究では、ガウシアンプロセスによって演奏をノンパラメトリックに学習し、表情を付加させるモデルを提案した。しかし本モデルでは十分な汎化性能は得られなかったことから、今後、よりピアニストを統一した上で十分な数の学習データを用いて、汎化性能の向上を目指すと共に、入力素性やカーネル関数の取捨選択に関して検討を行いたい。

## 謝辞

本研究は奈良先端科学技術大学院大学情報科学研究科 大学院教育改革支援プログラムの援助を受け実施されたものであり、ここに感謝の意を表します。

## 参考文献

- 1) Fryden, L. and Sundberg, J.: Performance rules for melodies, origin, functions, purposes, Proceedings of the 1984 International Computer Music Conference, pp.221-224(1984).
- 2) Widmer, G.: Machine Discoveries: A Few Simple, Robust Local Expression Principles. Journal of New Music Research, No.31, vol.1, pp.37-50 (2002).
- 3) Dorard, L., Hardoon, D. and Shawe-Taylor, J.: Can Style be Learned? A Machine Learning Approach Towards 'Performing' as Famous Pianists. Proceedings of Music, Brain & Cognition Workshop, in The Neural Information Processing Systems 2007 (2007).
- 4) Bishop, C. M.: Pattern Recognition and Machine Learning. Springer (2006).
- 5) 橋田光代, 松井淑恵, 北原鉄朗, 酒造祐介, 片寄晴弘: 音楽演奏表情データベース CrestMuse -PEDB ver.1.0 の公開について, 情報処理学会研究報告[音楽情報科学], Vol.2007, No.102, pp. 1-6(2007).
- 6) Rasmussen, C. E. and Williams, C.: Gaussian Processes for Machine Learning. MIT Press. (2006)
- 7) Tresp, V.: A bayesian committee machine, Neural Computation, vol. 12, no.11, pp. 2719-2741 (2000).
- 8) Schwaighofer, A. & Tresp, V.: Transductive and inductive methods for approximate Gaussian process regression. Advances in Neural Information Processing Systems 15, MIT Press (2003).