

統計的学習による テンポの変動を考慮したピアノ演奏模写

寺村佳子[†] 前田新一[†]

これまで我々はガウシアンプロセスを用いたピアノの演奏学習モデル(Usapi)を提案してきた。しかし、これまでのモデルではテンポに関する情報を扱っておらず時間的なダイナミクスの表現が十分ではなかった。そこで、本研究では、1拍単位のテンポの変動についてもガウシアンプロセスを適用し学習を行った。CrestMusePEDBから取得したピアニストの実演奏からテンポの変動を解析した結果、テンポ変動のトレンドや拍節での周期性が確認されたため、拍内の音の高さや長さなどとともに8小節単位での周期性と、前拍のテンポの変動を入力素性に用いた。CrestMusePEDBのデータを用いた評価実験から、本モデルはピアニストの実演奏の学習によって、ピアニストのテンポ変動を模倣できるようになったことが確認された。

Statistical Learning of Tempo Variations for Imitating Piano Performance

Keiko Teramura[†] and Shin-ichi Maeda[†]

In this paper, we propose a statistical approach that learns tempo variations from real pianist performances. We used a Gaussian Process (GP) to learn the tempo variations on every beats. As for the input elements to GP, besides the score information, the metrical periodicity and the continuity of the consecutive tempo variations are considered. By learning the real data taken from CrestMusePEDB, GP shows an excellent learning ability for 'known' pieces those are included in a given training data set while shows a modest generalization ability for 'unknown' pieces those are not included in a given training data set. Although the absolute error of the tempo deviations between the real performance and the estimation to the 'unknown' pieces is large, it is found that the relative changes of tempo variations between two consecutive beats, which is represented as a binary value (up or down) are well estimated.

1. はじめに

我々が演奏する時に用いる五線譜など、音楽の進行を示す楽譜には演奏する音符とその基本的な演奏指示が記されている。しかし、楽譜による演奏指示は緻密なものではなく、具体的に個々の音符をどのタイミングでどの程度の強さで演奏するかは演奏者や指揮者の裁量に任される部分が多い。時として、美しい音楽、表情豊かな音楽を演奏するため、楽譜による演奏指示から逸脱した演奏が行われることさえある。このように楽譜のみでは一義的に決まらない演奏上の自由度は、いかに美しい音楽として演奏するかという課題をもたらしている。

1.1 先行研究

1983年, Sundbergら[1]は計算機で人間が演奏するような表情豊かな演奏を生成するモデルを提案した。このように人間が演奏するように表情豊かな演奏を行う“演奏表情付け”を計算機で実現するモデルは、その後も[2, 3, 4]ほか多数提案されてきた。

この演奏表情付けの計算機モデルの先行研究は大きく二つに分けることができる。一つは、ルールベースモデルと分類されるものであり、[1, 2]のように演奏するタイミングや強弱に関する明示的な演奏ルールを表現するモデルである。もう一方は、事例ベースモデルと分類されるものであり、[3, 4]のように予め楽譜情報とその楽譜に対応する演奏情報がセットになった事例を用意し、その事例から演奏すべき楽譜のフレーズと似た楽譜のフレーズを検索し、類似した事例の演奏を転写することで演奏表情付けを行うものである。しかし、これら[1, 2, 3, 4]のいずれの方法においても専門的な音楽知識や手間のかかるパラメーター探索が必要とされ、精度を向上させるためには多大な労力が必要であった。

このような問題を解決するために2002年, Widmerらによるモデル[5]以降、機械学習の手法を取り入れたモデルがルールベース、事例ベースを問わず提案されている[6, 7, 8, 9]。これらはパラメーター探索の労力を大きく減らし、大部分の最適化を自動化することができる。このような実演奏の学習による最適化は、実演奏が正解データとして与えられ、その実演奏が再現されるように学習が進められる。その独創的な表情付けを行うことはないが、プロのピアニストの実演奏を学習することでそのピアニストの自然で美しい演奏を模倣することが期待される。

中でも我々は2008年, カーネル法の一つである Gaussian Process (以下, GP)[10, 13, 14]を用いた演奏表情付けモデルを提案している[9]。GPは、事例ベースモデルに分類されるモデルであり、事例となる学習データセットを用いることで入力と出力の間の回帰関数を近似するモデルである。このモデルは、確率モデルとして表現されている

[†] 京都大学大学院 情報学研究科
Graduate School of Informatics, Kyoto University

ため、学習データの尤度最大化などの統計的推定を用いることでパラメータの学習を行うことができる。加えて、関数近似の表現能力が優れていることが示されている[10]。我々はこのモデルを Usapi(Universal Statistical Automated Pianist Impersonator)と呼んでおり、Usapi では音符単位で演奏の打鍵開始時間、離鍵時間、強弱を出力素性として学習を行ったが、[6, 7]では拍単位のテンポの変動についても推定を行っている。テンポ変動は打鍵開始や離鍵時間と同様に時間推移に関する素性であるがあくまでも1拍単位でのテンポからの打鍵開始・離鍵時間のズレを表したものであり、より演奏時の揺らぎを表現するには、1拍単位でのテンポの変動も学習する必要がある。

そこで本報告では、Usapi を拍単位のテンポ変動の予測に用いた結果を報告する。Usapi をテンポ変動の予測に用いるにあたって、重要となるのが入出力素性の取り方である。我々は特にテンポ変動の特徴として顕著であった、周期性と連続性が表現できるように入力素性を工夫した。以下、その詳細について述べる。

2. テンポ変動の二つの特徴

入力素性を決めるにあたり MIDI データのテンポ推定に関する先行研究[11, 12]から、テンポの変動には周期性と連続性が深く関連していると考え、実演奏のデータから拍単位のテンポの変動について解析を行った。その結果小節単位での周期性と時系列上の連続性がみとれた。

2.1 周期性

CrestMusePEDB[13]のデータから V. D. Ashkenazy 氏が演奏した F. Chopin 作曲の楽曲 5 曲(付録 A 参照)の Tempo deviation をフーリエ解析したところ図 1A, B のような結果が得られた。

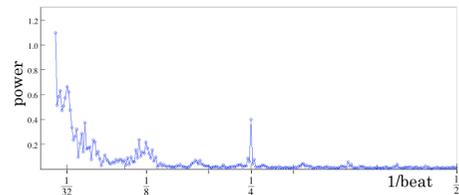


図 1A : 4 拍子の楽曲のテンポ変動の値のスペクトル解析結果

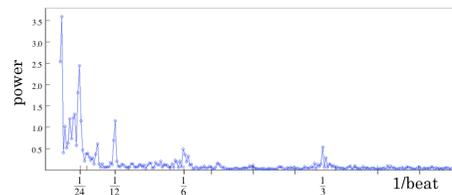


図 1B : 3 拍子の楽曲のテンポ変動の値のスペクトル解析結果

図 1A, B は 3 拍子, 4 拍子の楽曲のテンポ変動のデータについてそれぞれの楽曲ごとにフーリエ解析を行い、平均したスペクトルである。横軸の周波数は、1 拍あたりの周

期(1/beat)を表すもので、例えば 1/8 であれば 8 拍で 1 周期の周波数成分を意味する。

図 1 からそれぞれ拍数の倍数のところで高い値をとっていることがわかる。3 拍子, 4 拍子ともに拍数の整数倍、特に拍数の 8 倍のところ(周波数では 4 拍子のとき 1/32, 3 拍子のとき 1/24 のところ)で高い数値になっている。これは、小節や拍節や楽曲の進行上区切りとなることの多いとされる部分と一致している。

2.2 連続性

上記のフーリエ解析で用いたのと同じ楽曲を用いたときの、連続する二つの拍におけるテンポ変動の散布図を図 2A, B に、連続するテンポ変動の差分の散布図を図 2C, D に示す。

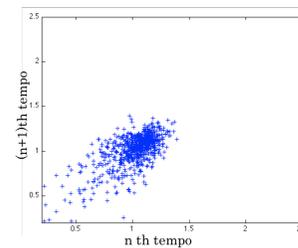


図 2A : 4 拍子の楽曲における連続するテンポ変動の散布図

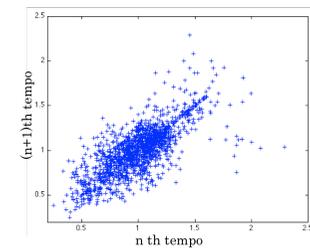


図 2B : 3 拍子の楽曲における連続するテンポ変動の散布図

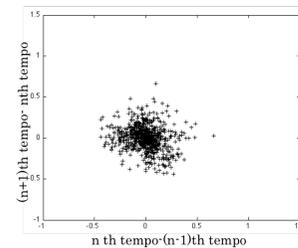


図 2C : 4 拍子の楽曲における連続するテンポ変動の差分の散布図

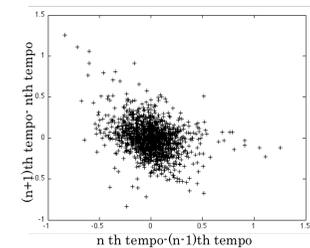


図 2D : 3 拍子の楽曲における連続するテンポ変動の差分の散布図

図からわかるように連続する拍の間では強い相関が見られる。実際、相関係数は 4 拍子の場合 0.72, 3 拍子の場合 0.76 と高い値を示している。一方、連続する拍のテンポ変動の差分を求め、連続するテンポ変動の差分の相関を調べたところ、4 拍子の場合 -0.17, 3 拍子の場合、-0.33 とどちらも低い逆相関がみられた。このことから、トレンドのようなテンポ変動は見られない一方、変動が連続的であることがわかった。

3. 提案手法

前節より明らかになったテンポ変動の特徴を踏まえ、それらの特徴を取り込んだ入力素性を用いてテンポ変動を予測する。テンポ変動の予測は1拍単位で行うものとし、1拍ごとに与えられる楽譜情報から算出される入力素性を用いて出力となるテンポ変動を予測する。この問題を、実演奏を学習データとした入出力関数の推定問題として定式化し、ガウシアンプロセスで学習する。以下、その詳細を述べる。

3.1 出力素性

n 拍目のテンポ変動を表す出力素性 t_n には、CrestMusePEDB-DEV で定義されている Tempo deviation を用いた。この Tempo deviation はテンポそのものではなく、曲ごとに定義される平均テンポ(Global Tempo)からの比として表される。

3.2 入力素性

n 拍目のテンポ変動 t_n を予測するのに用いられる入力素性 \mathbf{x}_n には、音符情報と拍節情報と連続性情報の3通りの情報を用いた。これらは、CrestMusePEDB-CSR と CrestMusePEDB-DEV から取得される。

(α) 音符情報：楽曲の特徴をとらえるため拍内の音符の情報から7素性を用いた。

- 音高情報 RP_n ：拍内にあるメロディー部の最高音を付録 B の算出方法によって相対音高で表現する
- 1, 2 拍前の音高情報 RP_{n-1} , RP_{n-2}
- 音符長情報 D_n ：拍内にあるメロディーでの最短の音符の長さを表現したもの。四分音符を1としたときの相対的な比で表現される。
- 1 拍前の音符長情報 D_{n-1}
- 強弱情報 Dy_n ：拍内の音の強弱を示す。各強弱記号に $f=2$, $mf=1$, $mp=-1$, $p=-2$ のような数値を割り振ったものである。拍内に強弱記号が与えられていない箇所には、小節単位に与えられる強弱情報を付与した。
- スラー情報 S_n ：拍内のメロディー部に含まれているスラーの情報を用いる。(始まり, スラー内, 終わり)をそれぞれ(1,0)の2値で表現する。スラーのない場所は(0,0,0)と記す。

(β) 拍節情報：テンポ変動の周期性を考慮し、次の2素性を用いた。

- 8小節単位での周期 PP_n ：8小節内の何小節目にあたるかを自然数で表したもの。弱拍から始まる場合は2小節目を1とする。
- 小節内の拍節 BP_n ：小節内の何拍目にあたるのかを表したもの

(γ) 連続性情報：テンポ変動の連続性を考慮し、前拍のテンポ変動を素性とした。

- 1 拍前の出力素性(Tempo deviation) : t_{n-1}

3.3 ガウシアンプロセス(GP)

本研究では前述の入力素性と出力素性の間の入出力関係を、GP を用いて学習する。この方法はパラメーターを近似するノンパラメトリック法の一つであり、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データ数が多い場合、計算コストが高くなる。そこで本研究では計算コストを減らすため、Bayesian Committee Machine(BCM)[14, 15]を用いた。

3.3.1 確率モデル

GP 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¹。このとき出力それぞれに含まれるガウシアンノイズは独立で同一のガウス分布に従うとすると、

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

と表せる。したがって真の出力 \mathbf{y} が与えられたときに観測される出力 \mathbf{t} の出現確率は

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

と表すことができる。ここで $N(\mathbf{t}|\mathbf{y}, \sigma^2 \mathbf{I}_N)$ は、平均 \mathbf{y} 、共分散 $\sigma^2 \mathbf{I}_N$ のガウス分布であり、 \mathbf{I}_N は $N \times N$ の単位行列を表す。また、入力 \mathbf{x} が与えられたときの真の出力 \mathbf{y} の出現確率は

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

で表されるとする。ここで、 \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

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

で与えられるとする。 a, B, c は、カーネル $k(\mathbf{x}_m, \mathbf{x}_n)$ を定義するパラメーターであり、そ

1. ここでの添字 n は前節のような曲中の n 番目の拍を意味するものではなく、学習データ全体に対する通し番号を意味する

それぞれ a, c は正のスカラー, B は各対角要素が非負の値をとる対角行列とした. T はベクトルの転置を表す. これらカーネルパラメーター a, B, c と式(2)のノイズ分散 σ^2 は最尤推定によって学習される.

3.3.2 未知入力に対する推定

未知の入力 \mathbf{x}^* を与えたときの真の出力 \mathbf{y}^* は以下のようにして推定される.

ベイズの公式より, 未知の入力 \mathbf{x}^* を与えたときの真の出力 \mathbf{y}^* の出現確率は

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

となり, 式(2), (3)を用いて計算すると式(5)もガウス分布となることがわかる.

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

このガウス分布の平均 $\mathbf{m} \equiv E[\mathbf{y}^* | \mathbf{x}^*, D]$ は

$$E[\mathbf{y}^* | \mathbf{x}^*, D] = \mathbf{k}_N^* (\mathbf{K}_N + \sigma^2 \mathbf{I}_N)^{-1} \mathbf{t} \quad (7)$$

と計算される. ただし, \mathbf{k}_N^* は $1 \times N$ ベクトルで, n 番目の要素 k_n^* は $k(\mathbf{x}^*, \mathbf{x}_n)$ となる. この

真の出力 \mathbf{y}^* の期待値 $E[\mathbf{y}^* | \mathbf{x}^*, D]$ を用いることで, 真の出力 \mathbf{y}^* の推定を行うことができる. しかし, 式(7)の計算には $N \times N$ の逆行列の計算が必要であり, 学習データ数 N が大きいとき計算コストが多大なものとなる.

そこで, 式(8)のように M 個のモジュールに分割したブロック対角行列で近似する².

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

ただし, D^i はデータ D を M 個のモジュールに分割した時の i 番目のモジュールを意味し, $D = \{D^1, \dots, D^M\}$ である. $E[\mathbf{y}^* | \mathbf{x}^*, D^i]$ は式(7)と同様に計算されるが, データをモジュ

ールに分割しているため計算コストは小さい. $\mathbf{C}, \text{cov}(\mathbf{y}^* | \mathbf{x}^*, D^i)$ は, それぞれ以下のように表される.

$$\begin{cases} \mathbf{C} = -(M-1)k(\mathbf{x}^*, \mathbf{x}^*)^{-1} + \sum_{i=1}^M \text{cov}(\mathbf{y}^* | \mathbf{x}^*, D^i)^{-1} \\ \text{cov}(\mathbf{y}^* | \mathbf{x}^*, D^i) = k(\mathbf{x}^*, \mathbf{x}^*)^{-1} + \sigma^2 - \mathbf{k}_N^{*T} (\mathbf{K}_N)^{-1} \mathbf{k}_N^* \end{cases} \quad (9)$$

4. 評価実験

4.1 使用データ

前述のガウシアンモデルを CrestMusePEDB ver.2.2 から取得した楽曲に適用し, その学習性能と予測性能の評価を行った. 本実験では CrestMusePEDB ver.2.2 で公開されているデータのうち, V. D. Ashkenazy 氏が演奏した F. Chopin 作曲の楽曲の楽曲から 3 拍, 4 拍子の曲それぞれ 6 曲(付録 A 参照)を用いて実験を行った. 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, 提案モデルによるテンポ変動の推定を推定演奏, ピアニストによる実際の演奏から得られたテンポ変動を単に実演奏と呼ぶ. また, 推定演奏を出力するためには一つ前の拍のテンポ変動の推定値が必要となるが, 最初の拍のテンポ変動(初期テンポ)は与えられるものとし, それ以降のテンポ変動を推定させた.

4.2 実験

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. それぞれ学習時のカーネルパラメーターの初期値は $a=1$, $B=\mathbf{I}$, $c=1$, ノイズ分散は $\sigma^2=1$ とした.

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, 6 曲で提案モデルの学習を行い, 実演奏と推定演奏とを比較した. 学習に用いたデータの総拍数は 3, 4 拍子それぞれ 1570, 1072 拍であった.

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. 5 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, 実演奏と推定演奏とを比較した. 学習に用いたデータの総拍数は 3, 4 拍子それぞれ 1282, 969 拍であった.

4.3 評価方法

本研究では, 推定演奏が, 実演奏にどれだけ近いかを定量的に評価するため, 以

² この近似のもとでカーネルパラメーターを学習するガウシアンプロセスが BCM と呼ばれる

下のように定義される推定演奏と実演奏との正規化誤差(ND 値)を計算した。

$$ND = \frac{\sum_{n=2}^N |PO_n - RO_n|}{\sum_{n=2}^N |RO_1 - RO_n|} \quad (10)$$

ここで、 PO_n , RO_n はそれぞれ n 番目の拍における推定演奏と実演奏を表す。ND 値は、初期テンポを推定演奏としたときの誤差を分母としているため、曲中のすべての拍のテンポを変動させずに初期テンポのまま推定した場合より誤差が小さければ 1 より小さい値をとる。

5. 実験結果

5.1 学習器の学習性能・汎化性能

まず学習実験と汎化実験の結果を図 3A, B に示す。4 拍子の場合、3 拍子の場合ともに学習性能は高い一方、汎化性能は高いとは言えない結果が得られた。特に 3 拍子の場合、ND 値は 1 付近の値をとっており、汎化性能がほとんどないように見える。しかし、カーネルパラメーターやノイズ分散の初期値による推定時には、ND 値は 4 拍子、3 拍子のそれぞれで 4.7857, 7.2432 であったため、学習による性能改善は見られる。

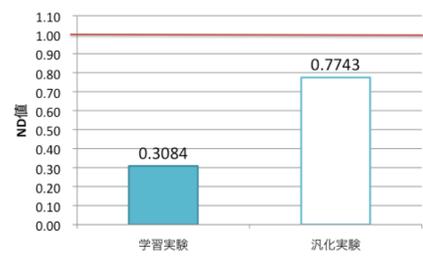


図 3A : 4 拍子の楽曲を用いて学習した GP で推定したときの ND 値

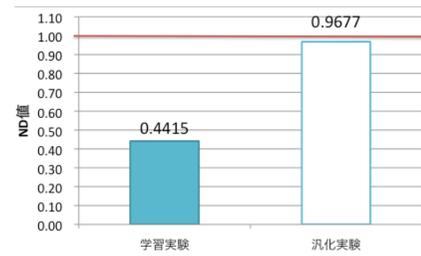


図 3B : 3 拍子の楽曲を用いて学習した GP で推定したときの ND 値

さらに、汎化実験において実際に推定されたテンポ変動の波形を図 4A, B に示す。図からわかるようにある程度の周期性と連続性のある推定演奏がえられていることがわかる。特に 3 拍子の場合、テンポの増減について推定演奏が実演奏と同様に山や谷をなし、変動している箇所が多々見受けられる。ただ、その増減の程度が異なるため、絶対的な誤差を測る ND 値では誤差が大きいと判断されたものと考えられる。

そこで、相対的な変動成分のみの一致性の定量化も行った。相対的な変動成分は、推定演奏と実演奏とでそれぞれ前拍に比べてテンポを早めたか、遅らせたかの二択の

問題の一致率で定量化した。その結果、一致率は 4 拍子では 44.55%、3 拍子の場合には 62.94% となり、3 拍子のほうが平均的なテンポからの増減を正確に捉えていたことがわかった。

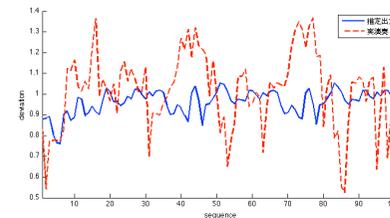


図 4A : 4 拍子の楽曲の実演奏でのテンポ変動と推定演奏でのテンポ変動

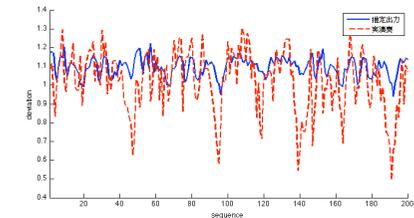


図 4B : 3 拍子の楽曲の実演奏でのテンポ変動と推定演奏でのテンポ変動

5.2 周期性情報と連続性情報の有効性

テンポ変動の入力素性として楽譜情報に加え、周期性情報と連続性情報を含め、その効果が示唆される結果が得られたが、実際にこれらの入力素性が有効に働いているかどうかをそれらの入力素性を省いたときの推定演奏と比較することで調べた。周期性情報と連続性情報の付加による汎化性能の違いを図 5 に示す。図中の (α) , (β) , (γ) はそれぞれ楽譜情報、周期性情報、連続性情報を入力素性に用いたことを示す。

図 5A, B からは、周期性情報、連続性情報の付与の両方の効果がみとれる。

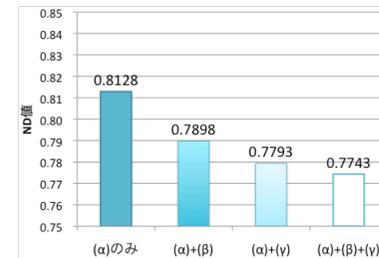


図 5A : 4 拍子の楽曲を用い学習した GP で推定したときの ND 値

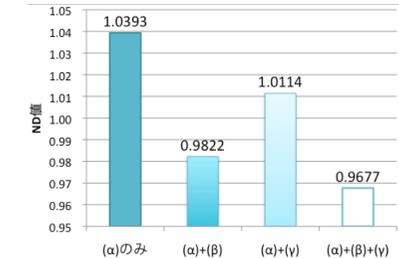


図 5B : 3 拍子の楽曲を用い学習した GP で推定したときの ND 値

6. まとめ

本研究では、GP によってピアノ演奏でのテンポ変動をノンパラメトリックに学習し、推定するモデルを提案した。特に実演奏でのテンポ変動には周期性と連続性が確認されたことからそれらを入力素性に用いた。正規化誤差(ND 値)を用いた評価実験の結果、

GPでの学習・汎化の効果が確認された。また、周期性と連続性に関する入力素性を用いることで汎化性能が高まることが確かめられた。特に3拍子の楽曲では汎化実験におけるND値は1に近く、ND値の指標からは汎化性能があるとは言えない一方、相対的な変動の一致率では6割を超える値を示した。これより実際に楽曲を演奏した際の主観的印象には相対的な変動は大きな影響を与えることが考えられるため、より適切な評価基準と、その評価基準に矛盾しない分布推定が必要であることが示唆される。さらに、このような変動の程度よりも相対的な変動の一致性を重視するならば、今回用いたガウス分布で表現するのではなく、たとえば、平均を中心として非対称に大きくズレる可能性のある分布を考える必要がある。また、過学習の効果も確認されており、より多くの楽曲を用いて学習する必要性も示唆された。

謝辞

評価実験におけるデータ処理では関西学院大学の北原鉄朗博士と戸谷直之氏に助言頂きました。ここに感謝の意を表します。

参考文献

- [1] Sundberg, J. et al.: Musical performance. A synthesis-by-rule approach. Computer Music Journal, Vol. 7, p. 37-43 (1983).
- [2] Todd, N. P. M.: A model of expressive timing in tonal music. Music Perception, Vol. 3, pp. 33-58 (1985).
- [3] 鈴木ら, :事例に基づく演奏表情の生成, 情報処理学会論文誌, Vol. 41, No. 4, pp. 1134-1145 (2000).
- [4] Hirata, K. and Hiraga, R.: Next Generation Performance Rendering - Exploiting Controllability, Proc. of International Computer Music Conference 2000, pp. 360-363 (2000).
- [5] Widmer, G.: Machine Discoveries: A Few Simple, Robust Local Expression Principles. Journal of New Music Research, No. 31, vol. 1, pp. 37-50 (2002).
- [6] Grindlay, G. and Helmbold, D.: Modeling, analyzing, and synthesizing expressive piano performance with graphical models Source, Machine Learning Vol. 65, Is. 2-3, pp. 361-387 (2006).
- [7] Flossmann, S. et al.: Expressive Performance Rendering: Introducing Performance Context, In Proceedings of the 6th Sound and Music Computing Conference (2009).
- [8] Dorard, L. et al.: Can Style be Learned? A Machine Learning Approach Towards 'Performing' as Famous Pianists. In Proceedings of Music, Brain & Cognition Workshop, in The Neural Information Processing Systems (2007).
- [9] Teramura, K. et al.: Gaussian process regression for rendering music performance, In Proceedings of International Conference on Music Perception and Cognition, pp. 167-172, (2008).
- [10] Rasmussen, C. E. and Williams, C.: Gaussian Processes for Machine Learning, MIT Press. (2006).
- [11] Raphael, C.: Automated Rhythm Transcription, In Proceedings of the 2nd International Symposium on Music Information Retrieval (2001).
- [12] 武田ら :テンポ曲線と隠れマルコフモデルを用いた多声音楽MIDI演奏のリズムとテンポの同時

推定, 情報処理学会論文誌, Vol. 48, No. 1, pp. 237-247 (2007).

[13] 橋田ら: 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).

[14] Tresp, V.: A bayesian committee machine, Neural Computation, vol. 12, no. 11, pp. 2719-2741 (2000).

[15] Schwaighofer, A. and Tresp, V.: Transductive and inductive methods for approximate Gaussian process regression. Advances in Neural Information Processing Systems 15, MIT Press (2003).

付録

付録 A. 実験使用楽曲一覧

IDはCrestMusePEDBのデータの通し番号(No.)を示す。網掛け印がテストに用いた楽曲であり、実験1:学習性能実験では全ての楽曲を学習に用い、実験2:汎化性能実験では網掛け印以外の楽曲を学習に用いた。TempoはGlobal tempoを示す。

(i)3拍子の楽曲

ID	楽曲名	Tempo
35	Prelude Op. 28 No. 7	55. 59
38	Waltz No. 1, Op. 18	234. 94
54	Waltz No. 3, Op. 34-2	131. 14
55	Waltz No. 7, Op. 64-2	201. 00
56	Waltz No. 9, Op. 69-1	159. 11
57	Waltz No. 10, Op. 69-2	153. 13

(ii)4拍子の楽曲

ID	楽曲名	Tempo
36	Prelude Op. 28 No. 15	27. 85
37	Prelude Op. 28 No. 20	165. 45
53	Nocturne No. 10, Op. 32-2	108. 31
58	Etude No. 3, Op. 10-3	49. 68
59	Etude No. 4, Op. 10-4	71. 20
60	Etude No. 23, Op. 25-11	32. 55

付録 B. 相対音高の算出方法

CrestMusePEDB-CSRは音符の高さはSMF(MIDI)の形式で記されており、調は明示されている。このSMF形式のC4~B4の中で調の主音と同じ音となる音を36とする。さらに前後で音階上にある音どうしの差を1、そうでない場合での差を0.5とする。図B1, 2はハ長調、イ短調の場合の相対音高を示したものである。

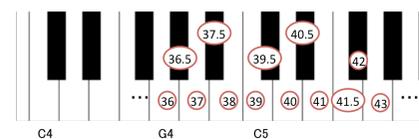


図 B1 ト長調の場合の相対音高

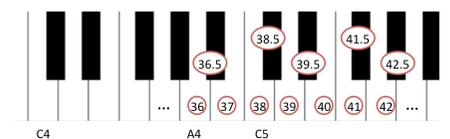


図 B2 イ短調の場合の相対音高