

演奏者の個性を表す特徴に関する考察

寺村佳子† 前田新一†

我々の提案する演奏表情付けモデル Usapi は、ガウシアンプロセスを用いることで類似性尺度(カーネルパラメータ)の学習を可能にしたモデルである。我々のこれまでの研究から、Usapi は単純な素性しか用いていないにも関わらず演奏者の個性を反映した演奏表情付けが行えることがわかった。そこで、特定の演奏者の演奏表情付けを学習した GP が演奏者のどのような演奏の違いを利用して演奏表情付けをしているのかについて、詳細な解析を行った。その結果、カーネルパラメータはピアニストに共通する値が学習されると考えられる一方、推定に用いる単純な素性の演奏者による違いが演奏者の個性を反映するのに重要であることが示唆された。それらの素性の二次と三次の統計量から同じ曲を演奏する演奏者を実際に識別できることを示す。

What represents the performer's unique style?

Keiko Teramura† and Shin-ichi Maeda †

Usapi(Universal Statistical Automated Pianist Impersonator) is a computational model for rendering expressive music performance, especially for piano performance. It is based on Gaussian process regression which can be regarded as one of (probabilistic) instant-based models. We have found that Usapi can learn the pianist's unique style despite of a small number of simple inputs. Then, we analyzed what characteristics of the pianists' style Usapi had learned from such simple inputs. Our experiments show that the estimation of Usapi heavily depends on who plays the instances while the estimation does not so much depend on the kernel parameter that determines the proximity measure between instances. Accordingly, we examined the second and third order statistics of the inputs. A discrimination test of the pianists suggests that these low-order statistics have a rich information about the pianists' unique style.

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

1. はじめに

演奏家は、美しい音楽をつくるために強弱や抑揚といった演奏の表情付けを行っている。同じ楽曲を演奏していても演奏者が違ったとき、その楽曲の印象が変わることがあるのは、この強弱やテンポなどの演奏の表情付けが違うからである。この演奏表情には主に楽曲や個人によらない一般的な演奏表情の付け方と、演奏する曲にはよらないが演奏者個人に依存する表情の付け方が存在する[11]。

1983年[13]以降、計算機によってこのような演奏表情を生成する研究が多数行われているが[2, 3, 4, 5, 6, 16, 17, 19, 22, 23]、近年、機械学習を用いたモデルの提案と共に、この個人性に注目した提案が報告されている[5, 17]。[5]は隠れマルコフモデルを用いた ESP モデルを提案し、学生の演奏者やプロの演奏家の演奏データを学習させテンポの変動を生成、ESP モデルの尤度からプロの演奏者の識別を行えることを、また事例探索を行う[17]では生成した演奏が個人の特徴を捉え、識別するという報告をしている。しかし、どちらのモデルもフレーズや Implication- Realization Model [7]の値といった現時点では自動的に獲得することが難しい特徴量を用いることを前提としており、モデルの完全な自動化には至っていない。またこれらの識別に関しては用いる特徴量が実際の演奏の違いを表現するのにどの程度、重要であったかについての検討はなされていない。

我々は、Gaussian Process (以下、GP)[1,8]を用いた演奏表情付けモデル Usapi(Universal Statistical Automated Pianist Impersonator) [14, 15]を提案している。Usapi は、事例ベースモデルの一種とみなせるが、入力の特徴性を判定するのにカーネルを用いているため、もっとも入力素性が近いと判断された楽曲の実演奏をそのまま利用するのではなく、入力素性の近さに応じた重み付けを行うことができる。これによって入力の連続的な変化に応じて連続的に出力が変化すると予期される関数近似問題において、推定すべき事例と完全に一致するような事例が訓練データ中に見つからない場合にも適切な補間を行うことで良い汎化性能が得られると期待される。GP は確率モデルとして表現されるため、入力同士の特徴性を測るカーネル関数のパラメータは尤度最大化の基準に基づいて学習できる。このように訓練データに基づいた統計推定を利用することで、ほぼ全ての作業を自動的に行うことができる。シンプルな音楽の特徴量を入力とすることで、音楽知識を前提とした人手による前処理も不要となるが、これまでの研究によって Usapi はシンプルな入力素性を用いても演奏者の個性を反映した演奏表情付けの学習・推定ができることがわかっている[15]。本稿では、このような特定の演奏者の演奏データを学習した Usapi が演奏者のどのような違いを利用しているのかについて、詳細な解析を行う。

2. Usapi(GPを用いた演奏表情付けモデル)

演奏情報 $t = \{t_1, \dots, t_N\}$ は、 N 個の音符それぞれに対して、どの程度の強さで演奏するのか、あるいはどのタイミングで演奏するのかを表すものである。このそれぞれの音符に与えられる演奏情報を楽譜情報をもとに推定するのが演奏表情付けの問題である。Usapi では、 i 番目の音符の演奏 t_i をその音符周辺の楽譜情報 \mathbf{x}_i を用いて推定する。以下では、演奏情報 t_i がスカラーであるとして GP の説明を行う。演奏情報 t_i はスカラーであるが、それを推定するのに用いる入力 \mathbf{x}_i はベクトルである。

2.1 Gaussian Process(GP)の確率モデル

GP は、 N 組の入力(楽譜情報) $\mathbf{x} = \{\mathbf{x}_1, \dots, \mathbf{x}_N\}$ とそれに対応する観測される出力(演奏情報) $t = \{t_1, \dots, t_N\}$ からなる訓練データを事例として用いて新規な入力に対して推定を行う。以下、文脈に応じて訓練データと事例という呼び方の両方を用いる。GP は、観測される出力が真の出力 $y = \{y_1, \dots, y_N\}$ に出力ごとに独立にガウシアンノイズが付加されたものであるとみなす。したがって真の出力 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、

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

と表すことができる。 $N(\mathbf{t}|y, \sigma^2 \mathbf{I}_N)$ は、平均 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。一方、入力 \mathbf{x} が与えられたときの真の出力 y の出現確率を

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

で表す。ここで、 \mathbf{K}_N は $N \times N$ の対角行列であり、 (m, n) 番目の要素 $k_{m,n}$ は

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

で与えられるとする。 a, \mathbf{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、それぞれ a, c は正のスカラー、 \mathbf{B} は各対角要素が非負の値をとる対角行列とする。 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, \mathbf{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。

これらのモデルから、ベイズの定理より未知の入力を \mathbf{x}^* としたときの真の出力 y^* は以下のようにして計算される。

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

式(2), (3)より、式(5)の分布 $p(y^* | \mathbf{x}^*)$ もガウス分布となることがわかる。演奏表情付けでは、このガウス分布の期待値 $E[y^* | \mathbf{x}^*, D]$ を演奏の推定値とする。この期待値は、

$$E[y^* | \mathbf{x}^*, D] = \mathbf{k}_N^* (\mathbf{K}_N + \sigma^2 \mathbf{I}_N)^{-1} \mathbf{t} \quad (6)$$

と計算される。ここで、 \mathbf{k}_N^* は n 番目の要素 k_n^* が $k(\mathbf{x}^*, \mathbf{x}_n)$ で与えられる $1 \times N$ ベクトルである。なお、式(6)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膨大となり、計算不可能となる。そこで、Bayesian Committee Machine(BCM)[10,18]を用いて逆行列をブロック対角行列の逆行列として近似して、期待値計算を行う。

式(6)から、GPによる推定値は観測される出力の重み付け和によって与えられること、また、出力 t_n に対する重みは、入力 \mathbf{x}_n と \mathbf{x}^* の近さを表す $k(\mathbf{x}^*, \mathbf{x}_n)$ に応じて与えられることがわかる。このように GP による推定は事例ベースによる推定と近い関係をもつ一方、統計推定の枠組みを利用することで類似性尺度となるカーネル関数 $k(\mathbf{x}_m, \mathbf{x}_n)$ のパラメータ学習が可能であるという特徴をもつ。

2.2 Usapiを用いた演奏表情付けの結果

我々は、ピアノ演奏とその対応する楽譜情報が収録されているデータベースである CrestMusePEDB[23]を用いて Usapi による演奏表情付けを行った。ここでは、演奏する際の音の強さを、平均値からの乗法的なズレとして表し、そのズレを推定した結果を紹介する。平均値は演奏する演奏者と楽曲ごとに決められる。以降では、この音の強さの平均値からのズレを単に dynamics とよぶ。dynamics は平均的な音の強さと等しい強さのときに 1 をとる。ただし、dynamics の推定時には、平均が 1 となるような補正は行っていない。過学習を避けるため、入力素性は音長(四分音符を 1 とする)と 1 音前の dynamics(どの音を 1 音前と認識するかは、付録 1 のアルゴリズムを参照。なお、曲の最初の音に関しては実演奏での 1 曲中の dynamics の平均値を利用)の推定値 \ddagger の 2 つのみに絞られた。したがって、前節の各 \mathbf{x}_i は 2 次元のベクトルで表される。また、それぞれの入力素性は標準偏差で予め正規化されている。

Gould, Pires のデータを学習させた Usapi で Mozart のピアノソナタ K.545-2nd の dynamics を推定させた結果を図 1 と図 2 にそれぞれ示す(学習に用いた楽曲については付録 2 に記す)。Usapi の学習に用いる楽曲にはテスト時の楽曲 K.545-2nd が用いられていないことに注意する。図 1 と図 2 からわかるように楽譜からの情報としては音長の情報しか使っていないにも関わらず Usapi は、精度良く、個々の演奏者の特徴を捉えた演奏をすることがみてとれる。このように特定の演奏者の演奏を学習させた Usapi は演奏者のどのような特徴を捉えているのか、次節にてより詳しく分析を行う。

\ddagger ただし、カーネルパラメータの学習時には、1 音前の dynamics は推定値ではなく、実際の演奏を利用した。

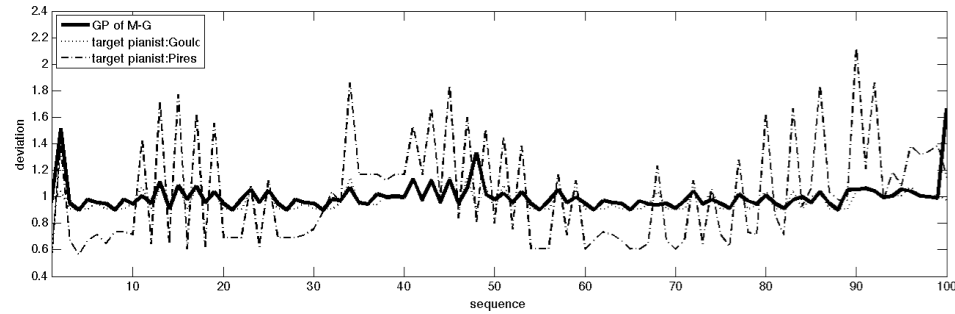


図 1: Gould を学習させた Usapi による推定演奏の一部(dynamics). 実線が Gould のデータを学習させた Usapi による推定結果(Usapi はこの楽曲の演奏データは学習に用いていない). 点線が Gould の実演奏, 破線が Pires の実演奏時の dynamics の値を示す.

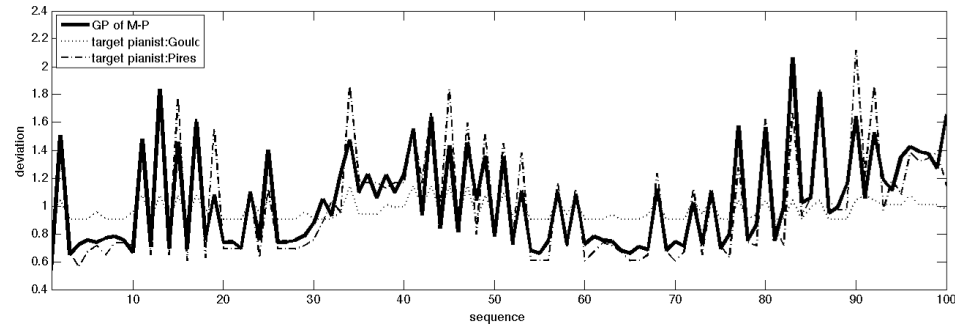


図 2: Pires を学習させた Usapi による推定演奏(dynamics). 実線が Pires のデータを学習させた Usapi による推定結果(Usapi はこの楽曲の演奏データは学習に用いていない). 点線が Gould の実演奏, 破線が Pires の実演奏時の dynamics の値を示す.

3. 演奏者の違いを表す特徴の分析

Usapi は、基本的には事例ベースの確率モデルであるので、学習に用いる事例そのものの違いは利用することになるが、その違いは学習によって獲得されるカーネルパラメータにも表現されている可能性がある。そこで、学習されたカーネルパラメータと、推定に用いる事例のそれぞれにどの程度、個人差が反映されているかについて表 1 で表すような実験で調べた。以下、それらの実験について述べる。

表 1: 実験条件のまとめ. datasetA は ND 値を評価するのに用いたテストデータセット, datasetB はパラメータ学習や推定を行うのに用いたデータセットを表す.

	パラメータ (a, B, c)	訓練事例(x)	評価時の楽曲(x')
実験 1	datasetA で学習	datasetA	datasetA
実験 2-1	datasetB で学習	datasetA	datasetA
実験 2-2	datasetA で学習	datasetB	datasetA

3.1 学習されたカーネルパラメータに反映される個人差の検証

学習された Usapi のカーネルパラメータの値を学習する演奏者を入れ替えて比較を行った。まず、CrestMusePEDB から 4 名の演奏者のデータを訓練データ (訓練データの内容については付録 2 を参照) としたときに得られたカーネルパラメータを表 2 に示す(実験 1)。この結果から、パラメータ a と c の値はほとんど演奏者間で変わらないものの、B の値は演奏者間で大きく異なっていることがわかる。

表 2: 実験 1 で学習されたカーネルパラメータの値

Parameters Performer	a	B		c
		音長	1 音前の dynamics	
Gould (M-G)	0.9997	7.74	31.5	0.841
Pires (M-P)	0.9995	8.18	7.05	0.890
Ashkenazy (C-A)	0.9997	2.74	10.23	0.893
Richter (B-R)	0.9987	2.42	76.09	0.806

次にカーネルパラメータを入れ替えたとき、及び 4 つのデータセットで学習されたカーネルパラメータの平均値を用いたときで Usapi の汎化能力に差異が出るかを調べた(以下、実験 2-1)。また学習する前の初期値 a=1, B=単位行列, c=1 の場合(初期値)も比較対象としてのせた。データセットは、作曲者と演奏者の頭文字をとって表す。たとえば、Mozart の楽曲をピアニスト Pires が演奏したデータは M-P のように表す。

ここで、Usapi で推定される演奏と実演奏とがどの程度近いかを評価するために正規化誤差(Normalized difference, 以下 ND と略す) を計算した。

$$ND = \frac{\sum_{n=2}^N |CP_n - TP_n|^2}{\sum_{n=2}^N |TP_n|^2}, \quad \overline{TP} = \frac{1}{N-1} \sum_{n=2}^N TP_n \quad (7)$$

式(7)の CP_n , TP_n はそれぞれ n 番目の音符に対する Usapi による演奏の推定値と実演

奏の値を示す。N は評価に用いた楽曲の音符の総数を表す。ND=0 は、実演奏と比較したい演奏が完全に一致すること意味し、ND=1 は、音符によらず実演奏の平均値を出力した場合の誤差と同じ誤差であることを意味する。なお、演奏者が違っても同じ楽曲を演奏する場合には演奏者共通の特徴を学習する可能性があるため、評価に用いる楽曲と訓練データは演奏者だけではなく曲自体が重複しないようにしてパラメータ学習とテストを行った。その結果を表3に示す。

表3の結果からわかるように、最もND値を小さくするパラメータは、訓練データで用いた楽曲を演奏している演奏者とテストで用いた楽曲を演奏している演奏者を一致させた時ではなく、C-AのChopinの楽曲をAshkenazyが演奏したときのパラメータを用いたときであることがわかった。ただし、そのパラメータの入れ替えによる違いは大きくなかった。初期値からND値が改善されていることから、学習によって得られたカーネルパラメータは、ピアニストに共通する値を学習したものと考えられた。

表3: 実験2-1の結果。評価に用いるUsapiのカーネルパラメータを別の演奏者の演奏データを学習したときのカーネルパラメータやそれらの平均値、又は初期値に変えて推定したときのND値。評価と学習でパラメータを変えない場合を網掛けで表す。

		評価時に用いるカーネルパラメータ (datasetBで学習されたカーネルパラメータを表す)					
		M-G	M-P	C-A	B-R	平均	初期値
評価データ (dataset A)	Gould (M-G)	0.669	0.663	0.659	0.666	0.666	0.723
	Pires(M-P)	0.521	0.519	0.516	0.519	0.519	0.532
	Ashkenazy(C-A)	0.586	0.582	0.582	0.585	0.585	0.592
	Richter(B-R)	0.718	0.715	0.713	0.716	0.716	0.745

3.2 推定に用いる事例に反映される個人差の検証

次に、カーネルパラメータは評価する楽曲を演奏する演奏者の楽曲で学習したものをを用いるが、推定時に用いる事例は他者の演奏データとした際のND値を調べた実験(実験2-2)の結果を表4に示す。

表4からは、推定時に用いる事例と評価に用いる楽曲とで演奏者は、一致していたときにND値が最も小さくなるのがわかる。Usapiは、推定に用いる事例に重みづけを行って推定を行うモデルであるが、実験2-1と実験2-2の結果からは事例に対してどのように重みづけを行うかではなく、事例自体の違いが演奏者の個性を模倣するのに重要であることを示唆する。また、評価するデータによって、ND値の絶対値が変わる傾向があることから、評価するデータセットごとに推定の難しさが異なることがわかる。

表4: 実験2-2の結果。評価時のデータを推定する際に用いる事例を他の演奏者による事例に入れ替えた際のND値。事例を入れ変えなかった場合を網掛けで表す。

		事例 (dataset B)			
		M-G	M-P	C-A	B-R
評価データ (dataset A)	Gould (M-G)	0.669	0.687	0.677	0.673
	Pires(M-P)	0.538	0.519	0.519	0.581
	Ashkenazy(C-A)	0.639	0.635	0.582	0.640
	Richter(B-R)	0.745	0.885	0.753	0.716

3.3 dynamicsの前後関係に現れる個人差の検証

Usapiは、ある音符のdynamicsを推定するにあたって、2つの入力素性、すなわちその音符の音長と一つ前の音符のdynamicsの推定値しか用いていないにも関わらず、演奏者ごとの弾き分けができていないことから、dynamicsの前後関係の演奏者による違いに基づいて演奏者の個性を表現した推定を行っていると考えられた。そこで演奏する音符の音長ごとにその音のdynamicsとその一つ前の音符のdynamicsの二つのがどのように関連しているかを散布図で調べた。演奏者の個人差を明確にするため、同じ楽曲を異なる演奏者が演奏しているデータセットであるM-GとM-Pのデータを散布図にしたものを図3に示す。図3からどの音長においても、前後のdynamicsの相関が非常に強いことがわかる。しかし、より注意深くみると、dynamicsの散らばり方は違っており、GouldよりPiresのほうがより大きく分布する傾向にあることがわかる。

図4(a),(b)は、音長に関係なく全ての連続するdynamicsを演奏者別に表示したものである。この二次元のデータを前後のdynamicsが等しい値をもつ長軸方向とそれに直交する短軸方向にそれぞれ射影しなおしたものを図4(c),(d)に示す。この図からも散らばり具合が演奏者間で異なっていることがわかる。また、長軸方向と短軸方向に射影したデータは、それぞれの軸に対してほぼ独立に分布しているように見える。

そこで、この散らばり具合を長軸方向と短軸方向のそれぞれで計算される統計量で評価した。dynamicsの平均は1であることがわかっているため、二次と三次の統計量となる、分散と歪度を長軸と短軸のそれぞれで演奏する曲ごとに算出した。データセットM-PとM-Gは、同じMozartの楽曲8つをPiresとGouldがそれぞれ演奏したデータであるため、この分散と歪度を長軸と短軸のそれぞれで計算した4つ統計量はPiresとGouldのそれぞれで8組得られる。これを可視化のために局所フィッシャー判別分析[12]を用いて2次元空間に射影したものを図5に示す。図5からGouldとPiresの演奏はほぼ1次元での特徴で分離できることがわかる。図中に示した。

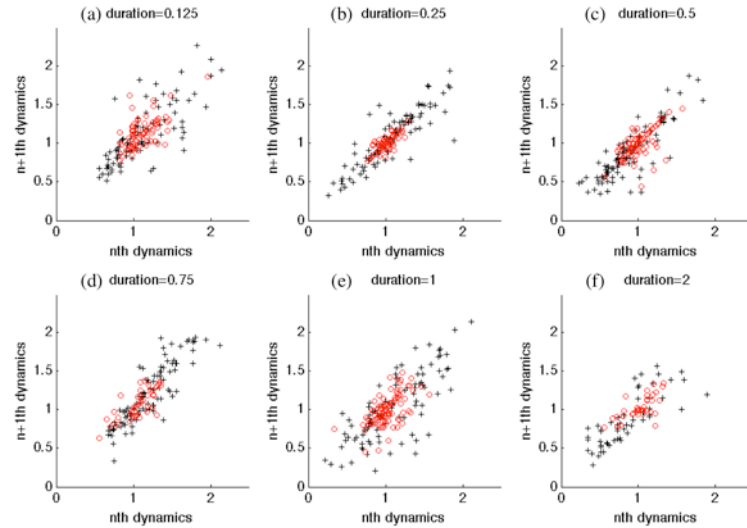


図 3 : 音音長ごとの強弱の前後 1 音の散布図.左上から音長=1/32(a), 1/16(b), 1/8(c), 付点 1/8(d), 1/4(e), 1/2 音符(f). それぞれ○が Gould, +が Pires の演奏を表す,

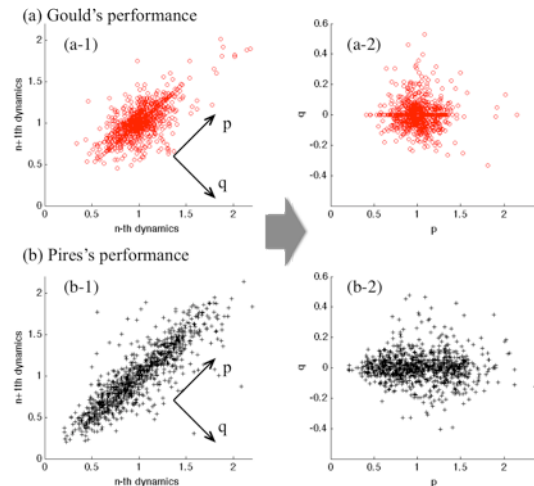


図 4 : Pires, Gould の演奏の dynamics の前後値の散布図(a-1, b-1)及び(p, q)平面に線形写像後の散布図(a-2, b-2).○が Gould, +が Pires の演奏を表す. 等高線は Pires と Gould の演奏から得られる 8 曲分のデータ点にそれぞれ正規分布

を当てはめた場合の対数尤度の等分布線を示し, 中央の黒線はそれらの尤度に無情報事前分布を用いてベイズ識別した際の判別境界を示している. さらに図 5 のような識別がロバストに行えることを確認するために, Leave one out で識別率の算出を行った. すなわち, 8 曲分のデータ点のうち 1 曲分をテスト用に除き, 残りの 7 曲分のデータ点から LFDA による 2 次元空間への射影とその空間における演奏者のデータごとのガウス分布の当てはめを行い, ベイズ識別を行うという評価作業を 8 曲全てで行いその平均をとったところ 87.7%の識別率があることがわかった. これより, 前後の dynamics に関する分散と歪度によって Pires と Gould の演奏の個人差が表現されることがわかった.

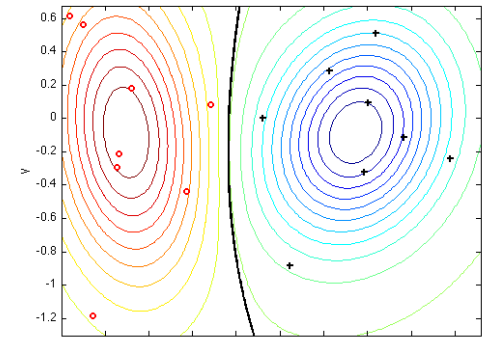


図 5 : 分散・歪度の統計量の LFDA による 2 次元平面への写像.○が Gould, +が Pires の演奏を表す.等高線は演奏者それぞれの対数尤度, 中央黒線が判別境界線を表す.

4. まとめ

本稿では GP を用いた演奏表情付けで, 個人の特徴までも学習できることについて検討し, ピアニストの個人差を表現する特徴量について提案を行った. 実験から, Mozart の楽曲を演奏する Pires と Gould の違いは, 連続する音符の dynamics の変化の仕方に現れていることが示され, 特に前後の dynamics が相関する長軸方向とそれに直交する短軸方向のそれぞれで算出される分散と歪度で表現可能であることが示された. これまで演奏表情付研究における個人の演奏識別や特徴量可視化は, [11, 20, 9] によって提案されているが, 本稿で示した特徴はシンプルで少数の特徴でありながら, 識別率の観点からは明確に個人差を表現できることが示唆された. 本実験では, Mozart の楽曲を演奏した Pires と Gould の 8 曲づつのデータでしか評価ができなかったが, 今後, 比較する演奏者や楽曲のデータセットを増やしたときにも同様の傾向が一般的にみられるかについて検討したい. また, 従来研究[11]からはアーティキュレーションが演奏者の個性を表す重要な特徴であることが示唆されているが, 連続する音符の dynamics の統計量をみることで演奏者によるアーティキュレーションの差を見ることができている可能性がある. この可能性についても検討を行いたい.

一方, 演奏表情付けを行う観点からは, 事例ベースによる表情付けを行う場合は特に演奏者ごとの事例に分類したうえで推定を行う必要性が示唆されるとともに, カーネルパラメータの学習においては演奏者の違いを気にせず全ての演奏者のデータをま

とめて学習して構わないことが示唆された。また、dynamics を推定するには、前後の dynamics で相関する方向と直交する方向のバラつきを説明することができる特徴を新たに入力素性に加えることで、より個人の演奏特徴を捉えた演奏表情付けが可能になることが考えられる。今後、これらの知見を踏まえた Usapi の改良にも取り組みたい。

参考文献

- 1) Bishop, C. M.: Pattern Recognition and Machine Learning. New York, Springer(2006).
- 2) Dorard, L. et al. : Can Style be Learned? A Machine Learning Approach Towards 'Performing' as Famous Pianists. Proc. of Music, Brain & Cognition Workshop, in The Neural Information Processing Systems (2007).
- 3) Flossmann, S. et al. : Expressive Performance Rendering: Introducing Performance Context, Proc. of the 6th Sound and Music Computing Conference (2009).
- 4) Gabriellsson, A. Interplay between analysis and synthesis in studies of music performance and music experience. Music Perception, No. 3, pp. 59-86 (1985).
- 5) Grindlay, G. & Helmbold, D.: Modeling, analyzing, and synthesizing expressive piano performance with graphical models Source, Machine Learning Vol. 65 , Is. 2-3, pp. 361-387 (2006).
- 6) Hirata, K. & Hiraga, R.: Next Generation Performance Rendering - Exploiting Controllability, Proc. of International Computer Music Conference, pp. 360-363 (2000).
- 7) Narmour, E.: The Analysis and Cognition of Basic Melodic Structures: The Implication- Realization Model. University of Chicago Press, Chicago (1990).
- 8) Rasmussen, C. E. & Williams, C. : Gaussian Processes for Machine Learning, MIT Press. (2006).
- 9) Repp, B. H.: Diversity and Commonality in Music Performance: An Analysis of Timing Microstructure in Schumann's Traumerei, Journal of Acoustical Society of America, Vol.92, No.5, p.p. 2546-2568(1992).
- 10) Schwaighofer, A. & Tresp, V. : Transductive and inductive methods for approximate Gaussian process regression. Advances in Neural Information Processing Systems Vol.15, MIT Press (2003).
- 11) Stamatatos, E. & Widmer, G.. Automatic identification of music performers with learning ensembles, Artificial Intelligence Vol. 165(1), pp. 37-56 (2005).
- 12) Sugiyama, M.: Dimensionality reduction of multimodal labeled data by local Fisher discriminant analysis., Journal of Machine Learning Research, Vol.8, pp.1027-1061, 2007.
- 13) Sundberg, J. et. al.: Musical performance. A synthesis-by-rule approach. Computer Music Journal, Vol. 7, pp. 37-43 (1983).
- 14) Teramura, K. et al. : Gaussian process regression for rendering music performance, Proc. of International Conference on Music Perception and Cognition, pp. 167-172, (2008).
- 15) Teramura, K. et al. : Gaussian Process Regression for Learning a Pianist Performance, (in preparation).
- 16) Todd, N. : A model of expressive timing in tonal music. Music Perception, Vol. 3, pp. 33-58 (1985).
- 17) Tobudic, A. & Widmer, G.: Learning to play like the great pianists. Proc. of the 19th international joint conference on Artificial intelligence, pp.871-876 (2005).
- 18) Tresp, V. : A bayesian committee machine, Neural Computation, Vol. 12, No. 11, pp. 2719-2741 (2000).
- 19) Widmer, G. : Machine Discoveries: A Few Simple, Robust Local Expression Principles. Journal of New

Music Research, No. 31, Vol. 1, pp. 37-50 (2002).

20) Widmer, G., et. al. :YQX Plays Chopin, AI magazine, Vol.30, No.3, pp. 35-48 (2009).

21) 小西貞則: 多変量解析入門-線形から非線形へ-, 岩波書店(2010).

22) 鈴木泰山ら: 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).

23) 橋田光代ら: 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).

付録

付録 1:前後の音を決めるアルゴリズム

1. 右手・左手パート,それぞれ別に前後の音は決定する。(CrestMusePEDB PEDB-SCR データには右手, 左手を識別する記述が含まれている.)
2. 前後で同じ数の音符の数がある場合, 音の高さの順番が一致するように前後の音を決める.
3. 前後で同じ数の音符数ではない場合は,それぞれの前の音と後ろの音とつなぎ合わせる組み合わせでの音の高さの差の合計が最小となるように組み合わせを決める.



付録図:前後の音の決定例(Chopin, Prelude No.7)

付録 2:実験で使ったデータ一覧

Dataset	演奏者	作曲者	曲数	楽曲 ID(in CrestMusePEDB)
M-G	G. Gould	V. A. Mozart	8	22, 27- 29, 43-45, 64
M-P	M. J. Pires	V. A. Mozart	8	1, 62, 63, 67, 70-72, 82
C-A	V. Ashkenazy	F. Chopin	15	32, 33, 36-38,53-60, 78, 87,
B-R	S. Richter	J.S. Bach	11	6,7,12,13,15-21

M-G と M-P は Mozart の同じ楽曲であるが一部に繰り返しの有無が異なるため, 演奏データとしては異なる演奏データとなっている楽曲が含まれている。