

# ピアノ演奏の動作パラメータに基づく熟達度推定 -「エリーゼのために」と 「戦場のメリークリスマス」による検討-

山口 翔也<sup>†</sup> 三浦 雅展<sup>††</sup>

楽器演奏の熟達度を推定する研究は盛んに行なわれているが、演奏音から熟達度を推定したものがほとんどであり、楽器演奏において重要な要素と言える演奏動作から熟達度を推定する研究は行なわれていない。この研究では、ピアノ演奏における手指や身体の動きに基づいて提案された動作パラメータを、モーションキャプチャによって記録されたモーションデータに対して算出している。記録された演奏をピアノ専門家による熟達度スコアから Good, Normal, Bad の3通りの演奏に分け、その熟達度カテゴリを動作パラメータを用いて推定したところ、「エリーゼのために」については73.3%、「戦場のメリークリスマス」についても73.3%という推定精度が得られ、提案された動作パラメータの有効性が確認されている。

## Proficiency estimation using motion parameter on piano performance :investigation with “For Elise” and “Merry Christmas, Mr.Lawrence”

Shoya Yamaguchi<sup>†</sup> and Masanobu Miura<sup>††</sup>

Several studies which estimate performance proficiency of musical instruments have been conducted. Most study's subjects are, however, based on only performance audio. The player's gestures, that are thought as an important factor on the performance evaluation, are out of focus on that studies. This study tries to realize proficiency evaluation from only player's gesture, by introducing motion parameters calculated from player's gestures. The authors asked piano experts to evaluate the proficiency for the recorded piano performances by giving ten-scale scores, which are then classified into each of “Good”, “Normal”, or “Bad”. A machine learning is carried out to estimate the proficiency category based on the calculated parameters. Estimation results reveal that the estimation accuracy is 73.3% for the music of “For Elise” and 73.3% for the music of “Merry Christmas, Mr.Lawrence”. Therefore, the authors confirm the effectiveness of the proposed motion parameters on proficiency evaluation.

<sup>†</sup> 八戸工業大学大学院工学研究科電子電気・情報工学専攻 Dept. of Electronic, Electrical and Information Engineering, Graduate School of Engineering, Hachinohe Institute of Technology

### 1. はじめに

楽器の奏者にとって、演奏の熟達度を他人に評価させることは、自らの演奏技術を向上させる上で重要な工程である。しかし、音楽的な訓練を受けていない人が、ある一定の熟達度を持つ演奏に対して、その熟達度を正しく理解し、評価することは困難である。そこで楽器演奏の熟達度に関する研究が行なわれている。特に、ピアノ演奏に対してはその傾向がより一層強いといえる[1,2]。

宮脇らの先行研究[1]では、ピアノ演奏の MIDI データから取得できるヴェロシティやテンポなどの演奏情報から、その熟達度を推定している。しかし、この研究[1]では、ピアノの演奏音のみを対象としているため、演奏中の手指の動作や身体の動作といった、ピアノの演奏動作については考慮されていない。ピアノ奏者の多くは、演奏音だけでなく演奏動作にも自身の演奏を表現するため、ピアノ演奏の熟達度は演奏動作からも推定できると考えられる。Miura の先行研究[2]では、ピアノの演奏音を音響信号、演奏動作をモーションキャプチャによってそれぞれ記録し、記録した演奏から作成された実験刺激を被験者に被験者に視聴させ、その熟達度を評価させている。また、その際の視聴覚優位性に着目し、ピアノ演奏の熟達度における評価モデルについて調査している。その結果、ピアノの熟達度評価において、聴覚効果と視覚効果の割合は約6:4であると示されている。このことから、ピアノ演奏の熟達度評価において演奏動作も重要な要素であるとされている。しかし、この研究[2]における熟達度評価は視聴覚の判断によるものであるため、ピアノ演奏の熟達度推定においては、演奏動作から算出される定量的な指標を用いた熟達度推定が必要であると考えられる。

### 2. 調査概要

本報告では、ピアノ演奏における奏者の身体の動作を定量化したパラメータ（以後、動作パラメータ）を提案する。また、先行研究[2]にてピアノ演奏を記録した三次元座標データ（以後、モーションデータ）に対して動作パラメータを算出する。そして、記録したピアノ演奏に対して、ピアノ専門家の評価によって決定された熟達度スコアに基づき、機械学習を用いてピアノ演奏の熟達度を推定する。

### 3. 用いるモーションデータ

本報告では、先行研究[2]で記録されたピアノ演奏のモーションデータを用いて、その演奏の熟達度を推定する。モーションデータとは、モーションキャプチャによって記録されたデータのことを指し、人や物の動きを記録する際に目印とした三次

<sup>††</sup> 八戸工業大学工学部システム情報工学科  
Dept. of System and Information, Faculty of Engineering, Hachinohe Institute of Technology,

元空間上の1点が、ある時刻においてx軸、y軸、z軸上のそれぞれの位置に存在するかを網羅的に表すものである。

### 3.1 モーションデータの記録

モーションデータは、龍谷大学瀬田キャンパス7号館情報実験室5の防音室内で記録された。奏者は音楽大学ピアノ専攻の学生もしくは卒業生（以後、ピアノ専門家）5名、ピアノ歴が10年以上のピアノ経験者（以後、ピアノ上級者）5名、ピアノ歴が9年以下のピアノ経験者（以後、ピアノ中級者）5名の計15名（男性4名、女性11名、平均年齢=21.4歳、SD=1.4）であった。奏者には光学式モーションキャプチャで用いるマーカーを人体の骨格情報に基づき手指と身体に装着させた。奏者に装着したマーカーの位置を図1に示す。図1より、計61個のマーカーを人体の骨格情報に基づいた手指や身体関節などに装着していることが確認できる。また、光学式モーションキャプチャに用いる赤外線カメラは奏者の手元側から記録する6台と奏者の背面側から記録する2台の合計8台を用いた。奏者には「エリーゼのために」、「戦場のメリークリスマス」から抜粋した短いフレーズを演奏させ、弾き間違いのない演奏の動作を60fpsで記録した。演奏の際に奏者に提示した「エリーゼのために」の楽譜を図2、「戦場のメリークリスマス」の楽譜を図3にそれぞれ示す。図2,3より、楽譜に示された1~5の数字は右手の指番号、i~vのローマ数字は左手の指番号をそれぞれ表す。15名の奏者それぞれの演奏動作から、30通りのモーションデータを記録された。

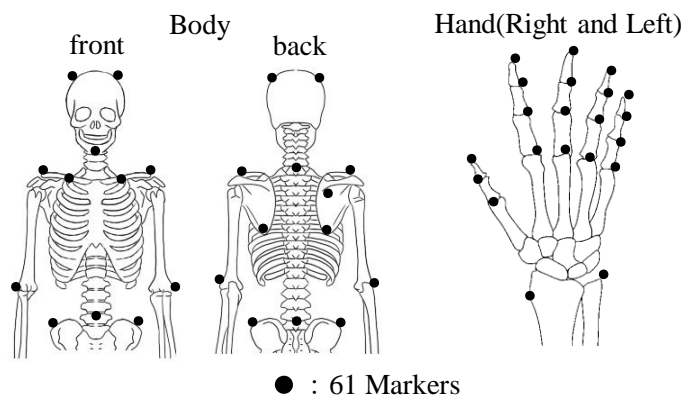


図1 奏者に装着したマーカー配置



図2 「エリーゼのために」の1フレーズを抜粋した楽譜



図3 「戦場のメリークリスマス」の1フレーズを抜粋した楽譜

### 3.2 演奏に対する熟達度スコアの評価

先行研究[2]では、記録した同一演奏における演奏動作と演奏音を合成した実験刺激に対して、その演奏の熟達度合いを表す熟達度スコアが算出されている。具体的には、3.1で述べたピアノ専門家5名およびピアノ中級者5名が、各演奏に対して1~10の10段階（10が最も良い）で評価しており、その評価を元に熟達度スコアのZ-Scoreが算出されている。本報告では、先行研究[2]で演奏動作と演奏音を合成した演奏に対して算出された5名の専門家による熟達度スコアのZ-Scoreの平均を熟達度推定の教師データとして用いることとする。

## 4. 提案する動作パラメータ

ピアノ演奏における人間の手指および身体動作の性質からピアノ演奏における動作パラメータを提案する。楽器演奏の評価において「硬い演奏」や「柔らかい演奏」といった表現が用いられる場合がある。どちらも抽象的な言葉ではあるが、一般に硬い演奏は奏者の演奏中における余裕の無い状態に起因するという想定の下では、熟達度が低いとされる。一方、柔らかい演奏は熟達度が高いとされ、これらの表現は熟達度と関係があると考えられる。まず、ピアノ演奏中の奏者の手指に着目する。それぞれ

の手指の先端から根本の関節までの距離は、硬い演奏ほど常に一定で、柔らかい演奏ほど変動が大きい。よって、ピアノ演奏中のそれぞれの手指における先端から根本の関節までのユークリッド距離を動作パラメータ「finger\_distance」として提案する。finger\_distance は三次元上の 2 点間の距離を表す算出式により求めることができる。次に、ピアノ演奏中の奏者の身体に着目する。ピアノ演奏においてよく可動する部位の角度は硬い演奏や柔らかい演奏といった評価に関係があると考えられる。よって、ピアノ演奏中の上部、両脇および両肘の角度を動作パラメータ「angle」として提案する。さらに、ピアノ演奏中における奏者の頭の動きに着目する。finger\_distance と同様に、硬い演奏ほど頭の動きにおける変動が小さく、柔らかい演奏ほど大きいと考えられる。本報告では顔を正面とした横方向、縦方向にその変動を分け、奏者の両肩を結ぶベクトルと顔を正面とした横方向と奥行き方向のベクトルとのなす角を動作パラメータ「Sagittal\_head\_variation」, 「Frontal\_head\_variation」としてそれぞれ提案する。提案した動作パラメータの概要を図 4 に示す。図 4 より、(a)は finger\_distance, (b)は上体の angle, (c)は両脇の angle, (d)は両肘の angle, (e)は Sagittal\_head\_variation, (f)は Frontal\_head\_variation をそれぞれ表す。

これらの動作パラメータを 3 で述べた 15 通りのモーションデータにおける各フレームに対して算出する。そして算出した動作パラメータに対し、平均 (avg: average), 標準偏差 (sd: standard deviation), 尖度 (ku: kurtosis), 歪度 (sk: skewness) をそれぞれ算出した統計量を本報告におけるピアノの熟達度推定に用いる。なお、動作パラメータの算出範囲は 1 音目 onset の 30 フレーム前 (すなわち 0.5 s 前) から、最終音 onset の 60 フレーム後 (すなわち 1 s 後) までと統一した。ここでいう onset とはピアノ演奏における打鍵ごとの最下点のフレームとし、モーションデータの出力範囲を指定する際に目視で確認する。提案した動作パラメータおよびその統計量を表 1 に示す。表 1 より、68 通りのパラメータ値を用いることが確認できる。

表 1 提案した動作パラメータと統計量

RThum_finger_distance	RIndex_finger_distance	RMiddle_finger_distance
RRing_finger_distance	RPinky_finger_distance	LThum_finger_distance
LIndex_finger_distance	LMiddle_finger_distance	LRing_finger_distance
LPinky_finger_distance	RArmpit_angle	RElbow_angle
LArmpit_angle	LElbow_angle	Upperbody_angle
Sagittal_head_variation	Frontal_head_variation	

× avg, sd, ku, sk = 68 parameters

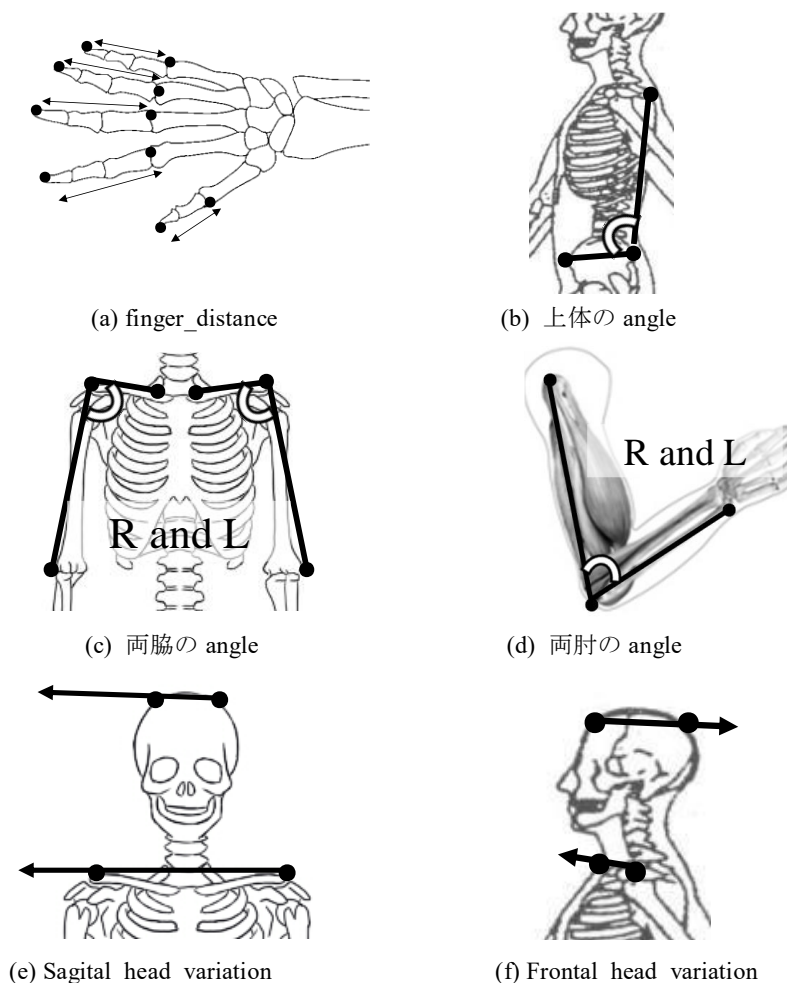


図 4 動作パラメータの概要

### 5. 機械学習による熟達度推定

3.2 で述べた教師データと 4 で算出した動作パラメータ値により機械学習を用いてピアノ演奏の熟達度推定を行なう。機械学習法は Logistic regression[3]を用い、オープンテストとして 10-fold CV(Cross Validation)を用いる。Logistic regression とは、教師ありのパラメトリックな識別モデルであるロジスティック回帰分析のことであり、各パラメータに最適な重み付けを行ない推定することができる。また、3.2 で述べた熟達度スコアの Z-Score のうち高評価 5 つを「Good」、低評価 5 つを「Bad」、それ以外の 5 つを「Normal」と 3 つの熟達度カテゴリに分けて定義する。このように熟達度スコアを定性的な指標に置換し、その熟達度カテゴリを推定することでピアノ演奏の熟達度推定を行なう。4 で算出した 68 種類の動作パラメータ値すべてを用いた「エリーゼのために」に対する熟達度推定結果を表 2、「戦場のメリークリスマス」に対する熟達度推定結果を表 3 にそれぞれ示す。表 2 より、全体の推定精度は 60.0%であると確認でき、これはチャンスレベル (33.3%) と比較しても高い数値であることがわかる。しかし、推定が Normal に偏っていることが問題であるといえる。また、表 3 より、全体の推定精度は 13.3%であると確認でき、どの熟達度カテゴリにおいても低い数値となっている。また、正解が Good である演奏を Bad と誤推定していることから、さらなる検討が必要であるといえる。しかしながら、表 2,3 の推定結果のどちらにおいても、パラメータ数が過剰であるため過学習が起こっている可能性が考えられる。そのため、パラメータを精査し、再度検証を行なうこととする。

パラメータの精査は先程と同様に Logistic regression の 10-fold CV を用いて、それぞれの演奏において単一パラメータによる推定を行なう。そして、推定精度が低い順に 1 つずつパラメータを除くことで最適なパラメータセットを探索する。「エリーゼのために」、「戦場のメリークリスマス」に対する最適なパラメータセットの探索結果を図 5,6 にそれぞれ示す。図 5 より、「エリーゼのために」においては 52 通りのパラメータによる推定精度が 73.3%と最も高い数値となることが確認できる。また、一度推定精度が落ち込むパラメータセットの場合はあるものの、全体的に比較的高い数値が確保できていることが確認できる。また、図 6 より、「戦場のメリークリスマス」においては 2 通りのパラメータによる推定精度が 73.3%と最も高い数値になることが確認できる。また、全体的に低い推定精度のパラメータセットが多いものの、少数のパラメータによる推定精度が高いことから、「戦場のメリークリスマス」においては重要なパラメータが明確化されるといえる。さらに、「エリーゼのために」、「戦場のメリークリスマス」において精査したパラメータセットを用いた推定結果を表 4,5 にそれぞれ示す。表 4 より、Normal と Bad は 60%、Good は 100%の精度で推定され、全体の推定精度は 73.3%であることが確認できる。どの熟達度カテゴリにおいても高い数値を確保していることから、本報告で提案した動作パラメータは「エリー

ゼのために」の熟達度推定において有効であるといえる。また、単一パラメータの推定精度を確認すると、上体の角度を表す「upperbody\_angle」の avg が最も推定精度が高く、ku や sk もある程度高い推定精度を持っていたことから、「エリーゼのために」を演奏することにおいて、上体の角度は有効なパラメータであり、より熟達した

表 2 全パラメータを用いた「エリーゼのために」に対する推定結果

		Estimated proficiency			Precision [%]
		Good	Normal	Bad	
Correct proficiency	Good	4	1	0	80.0
	Normal	0	3	2	60.0
	Bad	0	3	2	40.0
					Avg = 60.0

表 3 全パラメータを用いた「戦場のメリークリスマス」に対する推定結果

		Estimated proficiency			Precision [%]
		Good	Normal	Bad	
Correct proficiency	Good	1	3	1	20.0
	Normal	1	0	4	0.0
	Bad	2	2	1	20.0
					Avg = 13.3

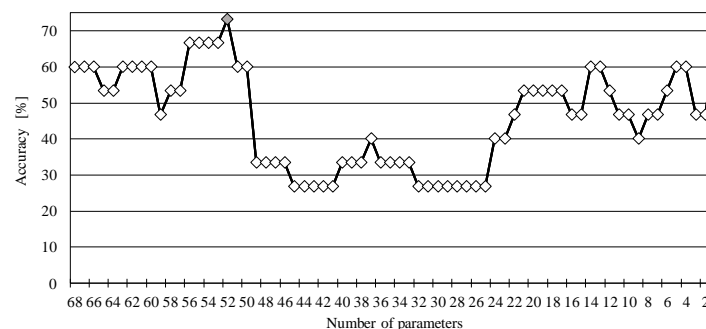


図 5 「エリーゼのために」に対する最適なパラメータセット探索結果

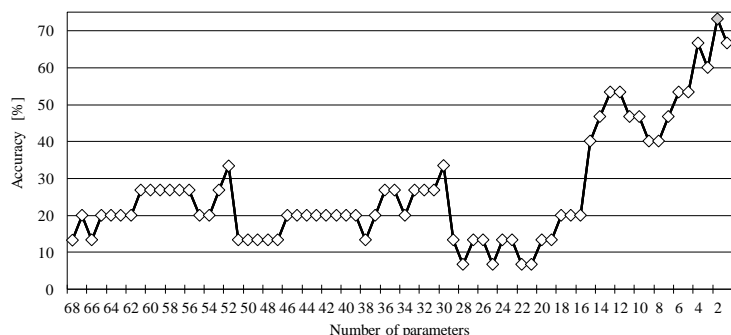


図 6 「戦場のメリークリスマス」に対する最適なパラメータセット探索結果

表 4 「エリーゼのために」における精査したパラメータセットによる推定結果

		Estimated proficiency			Precision [%]
		Good	Normal	Bad	
Correct proficiency	Good	5	0	0	100.0
	Normal	0	3	2	60.0
	Bad	0	0	3	60.0
Avg = 73.3					

表 5 「戦場のメリークリスマス」における精査したパラメータセットによる推定結果

		Estimated proficiency			Precision [%]
		Good	Normal	Bad	
Correct proficiency	Good	4	0	1	80.0
	Normal	1	4	0	80.0
	Bad	1	1	3	60.0
Avg = 73.3					

演奏を動作で表現する場合に意識すべき点であると考えられる。そして、表 5 より、全体の推定精度は 73.3%と高く、それぞれの熟達度カテゴリをみても十分高い数値を確保していることから、本報告で提案した動作パラメータは「戦場のメリークリスマス」の熟達度推定においても有効であるといえる。しかし、表 4 と比較すると、どの熟達度カテゴリにおいても誤推定が全体に分散していることが確認できる。この差はパラメータ精査後に用いたパラメータ数の違いによるものであると考えられ、52 通りのパラメータを用いた「エリーゼのために」に比べて、2 通りのパラメータしか用いていないことが原因の 1 つとして挙げられると考えられる。また、「戦場のメリークリスマス」は比較的の動きが小規模であるため、「エリーゼのために」とは異なる知見となったことも原因の 1 つと考えられる。そのため、動きが小規模で演奏される楽曲の熟達度推定における推定精度を向上させるためには、それに適応した新たなパラメータの開発が必要であると考えられる。

## 6. まとめと今後の展望

モーションキャプチャで記録したピアノ演奏のモーションデータから動作パラメータを提案した。ピアノ専門家の評価に基づいた熟達度スコアに基づき、「エリーゼのために」と「戦場のメリークリスマス」について、Logistic regression により熟達度推定を行なった結果、どちらにおいても 73.3%といった推定精度が得られた。今後の展望としては、手首などの回転を考慮したパラメータの検討、教師データの拡充を行なう必要があると考えられる。また、今回は定性的な熟達度推定を試みたが、実際に演奏を採点できるような定量的な熟達度推定を可能にすることは目標の一つであると考えられる。さらに、演奏動作から熟達度を推定する手法を応用したシステムも思案していく必要があると考えられる。

謝辞 本研究の一部は、科研費(15H02882)の援助を受けた。

### 参考文献

- [1] 宮脇 聡史, 三浦 雅展, “固有演奏を用いたピアノ熟達度の評価基準における多様性の可視化手法”, 日本音響学会誌, Vol.72, No.10, pp.617-626 (2016).
- [2] Miura.M., “Piano proficiency evaluation on audio-visual condition: Audio versus visual / timing versus dynamics”, International Symposium on Performance Science 2017, p.163 (2017).
- [3] Hosmer Jr, D. W., and Lemeshow, S., Applied logistic regression, John Wiley & Sons (2004).