

# 深層学習を用いたピアノ学習者の練習時間分析

島谷 翼<sup>†</sup> 峯 恭子<sup>‡</sup> 土江田 織枝<sup>†</sup> 山田 昌尚<sup>†</sup>釧路工業高等専門学校<sup>†</sup> 大阪大谷大学<sup>‡</sup>

## 1. はじめに

音楽演奏能力は身体知のひとつであり、いったん習得すれば永続性が高いものの、既習者・教授者から初学者への伝達が形式知よりも難しいという特徴があり、その獲得過程は十分明らかになっていない。身体知研究は、スポーツ科学の分野で精力的に展開されているほか、ものづくりの視点から熟練工スキルの伝承や安全性の観点からの研究も進んでいる[1]。音楽情報科学の分野では演奏支援研究が身体知獲得をサポートするものと位置づけることができ、筆者らも音響信号処理技術を用いたリズム練習支援システムを開発してきた[2]。こうした演奏支援システム研究における評価実験は、1日15分~30分程度の練習を数日間という期間で行われていることが多い。一般的に楽器演奏の習得には長い期間が必要であることを考えれば、この実験期間はかなり短いといえるであろう。これには、被験者の協力を得るという制約からやむを得ない面はあるにせよ、身体知獲得という点からいえば、より長期的に視点で演奏技術の向上を分析・評価する必要があると思われる。しかし、演奏支援システムを使うかどうかにかかわらず、演奏が上達していく過程がどういったプロセスであるかを定量的に調査した研究はみられず、その方法もほとんど検討されていない。そこで本研究は、ピアノ初級学習者の練習演奏データを継続的に蓄積および分析し、身体知獲得過程についての知見を得ることを目的とし、特に複数の楽曲を並行して練習する場合の各曲の練習時間に着目した。指導者にとっては、個人ごとの楽曲の練習時間がわかれば、練習者が困難を感じている点などに、効果的にアプローチして指導していくのに役立つと考えられる。

Estimation for Practice Duration of Piano Players using Deep Learning

<sup>†</sup> National Institute of Technology, Kushiro College

<sup>‡</sup> Osaka Otani University

## 2. 検討手法

ピアノ演奏データとして、大阪大谷大学教育学部教育学科幼児教育専攻でピアノのグループ授業を受けている約60名(約20名×3クラス)の学生を対象とし、2018年度の1年間、毎週90分の授業時間について、原則的にすべての演奏を記録した。この授業を受けている学生の多くは、保育士あるいは幼稚園教諭を目指しておりピアノ演奏獲得への必要性が高い一方で、ピアノ演奏経験のない状態から学習を始める者も少なくない。授業では電子ピアノを使用しているので、演奏データ記録には電子ピアノ付属のMIDIデータ記録機能を用いた。学生が練習に使用するのはバイエルおよび弾き歌いの保育曲である。学生は90分の授業時間中に、教員の個別指導をうける時間と、個人で自由に練習できる時間があり、後者においてはバイエルや保育曲を何曲か並行して練習する。

特に初心者の場合、練習の演奏には音高やリズムの間違い、中断、繰り返しなどが多く含まれることから、ある時点で演奏されている曲が何であるかを推定するために、動的計画法のようなパターンマッチングでは対応できない。また、系列長が非常に大きいため隠れマルコフモデルを使用するのも現実的でない。そこで本研究では、曲ごとの練習時間を深層学習を使用して推定することを試みた。扱う対象が時系列信号のため、深層学習にはリカレントニューラルネットワークを用いた。

## 3. 実験

最終的に分析対象としたいのは上記のグループ授業を受けている60名の1年分の演奏データであるが、まず試験的に9名について、授業各1回(90分間、記録日は同一)の練習データを分析対象とした。演奏データにはバイエル16番~28番および保育曲2曲の15曲の演奏情報が含まれており、演奏者によってデータに含まれる曲数は2~12曲と幅がある。

この演奏MIDIデータに対して、フリーソフト

ウェアの Domino を用いて曲が切り替わった開始位置に手動でマーカーを入れ、これを正解データとした。演奏データとは別に、バイエルの楽譜から MIDI データを作成し、これを教師データとした。MIDI ファイルには、演奏している曲のテンポにかかわらず 4 分音符=120BPM でノートイベントが記録される。演奏 MIDI データおよび楽譜から作成した MIDI データそれぞれに対して 0.1 秒ごとにサンプリングを行ってニューラルネットワークへの入力データとした。ただし、5 秒以上続く無音部分は、曲間もしくは練習の中断とみなしてカットした。拍節情報は使用していない。

手動でつけたマーカーにもとづいて、曲ごとの練習時間を集計した結果を表 1 に示す。B16～B28 はバイエルの番号、H1 および H2 は保育曲である。この日はバイエル 16 番、18 番、20 番が主に練習されていることや、学生 6 は自分なりに曲を先に進めていることがわかる。

こうした分析を手動で行うには限界があるため、深層学習を用いて練習されている楽曲を推定することを試みた。時系列データを扱うのに適したリカレントニューラルネットワークの一種である LSTM (Long Short-Time Memory) を使用し、Chainer で実装した。ネットワークの入力層は音高に対応して 88 ノード、中間層は 200 ノード、出力層は保育曲を除く 13 曲分と無音部分で 14 ノードの分類問題とした。出力層と教師信号から得る損失の算出には平均二乗誤差を用いた。入力データの連続する 50 サンプル点を 1 シークエンスとして学習の単位とし、データ全体のランダムな位置から取った 4 シークエンスごとにミニバッチ学習を適用した。検証時には 30 サンプルを 1 シークエンスとして推定している。

図 1 にエポック数 50 までの学習の収束の様子

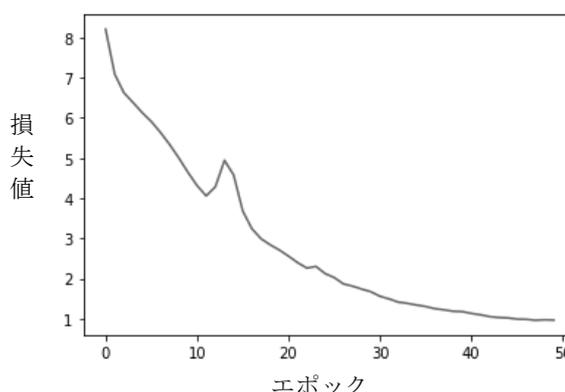


図 1 学習の収束の様子

を示す。演奏データを用いた推定の正解率は約 30%であった。

#### 4. まとめ

初心者を含むピアノ練習時間を分析するため、深層学習を用いて演奏データから楽曲ごとに分類することを試みた。結果として推定率は不十分であったため、今後、教師データの与え方や学習方法について検討を続けたい。

**謝辞** 本研究の一部は JSPS 科研費 16K01094 の助成を受けています。

#### 参考文献

- [1] 諏訪正樹. "身体知獲得のツールとしてのメタ認知的言語化 (特集) スキルサイエンス." 人工知能学会誌 Vol. 20, No. 5 pp. 525-532 (2005).
- [2] Yamada, M., Doeda, O., Matsuo, A., Hara, Y., Mine, K. "A rhythm practice support system with annotation-free real-time onset detection." Advanced Informatics, Concepts, Theory, and Applications (ICAICTA), pp. 1-6 (2017).

表 1 各学生の曲ごとの練習時間 (分)

	B16	B17	B18	B19	B20	B21	B22	B23	B24	B25	B26	B27	B28	H1	H2	計
学生1	4.9		43.8													48.7
学生2	10.0	0.9	4.5	7.9		8.8	11.7									43.7
学生3	16.1		25.1	7.2		3.3									1.8	53.5
学生4			25.8	1.4										0.4	0.8	28.4
学生5	19.9		32.2			2.4										54.4
学生6	6.8		10.5	1.4	1.5	12.0	4.1	3.6	4.3	2.9	3.0	6.8	5.9			62.8
学生7	0.9		28.2			30.5										59.6
学生8	11.9		22.1			8.9								3.8		46.8
学生9	8.7		31.2			13.5										53.4