

強化学習によるバイオリン演奏ロボットの動作生成 ～割引率と探索率の弓運動への影響～

○西村 郁哉（龍谷大学） 福原 裕志（龍谷大学） 渋谷 恒司（龍谷大学）

1. はじめに

楽器演奏は、ロボットのアプリケーションの1つとして早くから取り上げられて来た。例えば、加藤らのオルガン演奏ロボットや[1]、梶谷らのMUBOT[2]をはじめとして、ギター、バグパイプ、マリンバ等、様々な楽器演奏ロボットが開発されている[3]-[4]。近年では、高西らが、フルート[6]、[7]およびサクソフォン演奏ロボット[8]を開発している。

多様な楽器の中で、本研究では西洋音楽において主要な位置を占めるバイオリンに着目し、人間型ロボットによる演奏を目指している。その中でも、楽譜から演奏動作を自動的に決定する過程の構築と、それを用いたロボットでの演奏を本研究の目的としている。

熟練演奏者は訓練によって楽譜から演奏動作を獲得していると考えられるが、その過程は明らかではない。この過程をシステムとして構築することは、ロボットの演奏能力を向上させレパートリーを増やすことにつながるだけではなく、人間の演奏表現の理解につながると考えられる。前述のオルガン演奏ロボットに関しては、楽譜から演奏動作を自動的に決定するアルゴリズムが提案されているが[13]、バイオリン演奏等、弦楽器に関しての研究は見られない。

バイオリン演奏ロボットの研究については、前述のMUBOT[2]の後、Kuwabaraらの研究[9]、Kim[10][11]らの研究がある。これらの研究は、より良い音を生成するために必要なアームやハンドの機構と制御システムの開発が目的となっている。また、トヨタ自動車(株)が、ヒューマノイド型のバイオリン演奏ロボットを開発しているが[12]、学術的な研究が目的のものではない。

本研究ではこれまでに、人間のバイオリニストの演奏を分析して、特に「明るい」や「暗い」の音色を表出しようとしたときの、弓圧、弓速、および弓の運動方向等の演奏パラメータの分析を行い、得られたデータから、表出しようとする音を生成する演奏動作を決定するアルゴリズムを構築を試みている[14]。しかし、そこで示された動作決定手法は、ある楽譜に対する基本的な方針を示したのみであり、異なる楽譜に適用しようとした場合、人間の設計者が演奏を設計する必要がある。これは非効率的であり、楽譜から自動的に演奏パラメータを決定できるシステムが求められる。そこで本研究では、バイオリン演奏の操弓方向を自動的に決定する過程を、価値関数近似を用いた強化学習(Q学習, ϵ -greedy 方策)を利用して実現することを提案している[15]。

[15]では、目標の音圧(Sound pressure level: SPL)となるような演奏音が得られる弓の運動を得ることを目的に、割引率 γ についてのみ検討したが、探索率 ϵ の影響については検討されていなかった。そこで、本

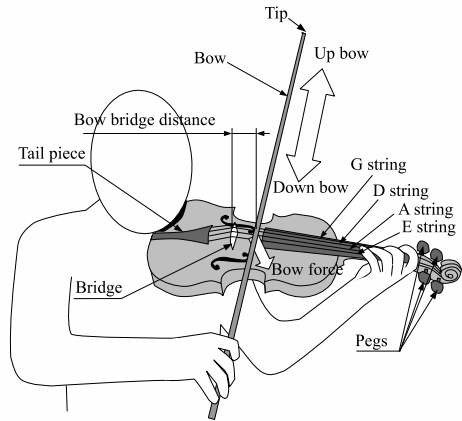


図1 Name of violin parts and bowing direction

論文では、探索率も考慮して、割引率と探索率の適切な値について検討した。シミュレーションの結果、[15]で得られた結果より、より望ましい音圧を生成できる弓の運動を生成できることを明らかにした。

2. バイオリン演奏

バイオリンは、Fig.1に示すように、弦が4本ある弦楽器である。肩と顎で楽器を支持し、右腕・指による操弓(bowing)で音を生成し、左指・腕による運指(fingering)によって音高を決める。

本研究では特に操弓(ボウイング)に着目しているので、それに関する事項を以下にまとめる。弓先(Tip)から演奏をはじめ、弓を下から上に動かす方向を「上げ弓(up bow)」, その逆を「下げ弓(down bow)」と呼ぶ。上げ弓から下げ弓, 下げ弓から上げ弓に変えることを「切り返し」と呼ぶ。弓が弦を押す力のことを「弓圧(bow force)」, 弓の弦との相対速度を「弓速(bow speed)」と呼ぶ。

上記の操弓に関するパラメータは、楽譜情報のみからでは決定できない。例えば、弓の運動方向はどちらの方向でも、同じ音高、音圧の音を生成することが可能である。弓の運動方向は、作曲者が指定している場合もあるが、基本的には演奏者が決めることが可能である。しかし、前後の音符の演奏に必要な弓の運動なども考慮する必要がある、その演奏決定過程は複雑なものとなる。

3. 強化学習を用いた演奏動作導出アルゴリズム

3.1 対象とするロボットと演奏パラメータ

Fig.2に、本研究で対象としているロボットを示す。左右の腕の自由度は各7の人間型ロボットである。

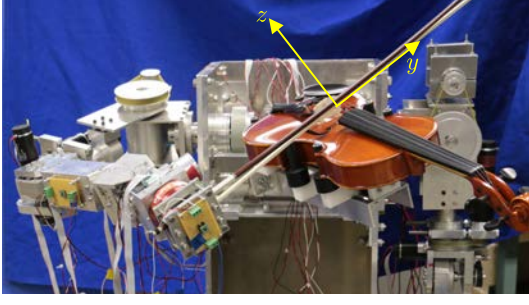


図2 Violin-performance robot, and the y axis along with the bow's longitudinal direction

本研究では、音の物理パラメータとして音圧を取り上げ、求める音圧を生成する弓の運動を導出するシステムの構築を目指している。音圧に影響する弓のパラメータとして、弓圧、弓速、サウンディングポイント(弓と駒 (bridge) との距離)、および弓の運動方向があるが、問題を簡単化するため、弓速と弓の運動方向のみを対象とし、その他のパラメータは一定と仮定した。弓の運動方向は、Fig.2 に示すように、弓の長手方向で弦に垂直な方向 (Fig.2 の y 軸) である。

本論文でのシミュレーションでは、実験的に求めた音圧値と弓速との関係 (式 (1))[15] を用いた。

$$L = -0.0015v^2 + 0.3286v + 57.704 \quad (1)$$

ここで、 L は音圧 [dB]、 v は弓速 [m/s] を表す。

3.2 学習手法

学習においては、数小節のフレーズを学習 1 エピソードとし、音符 1 個を状態遷移の 1 ステップとして定義した。状態行動価値の更新については、1 エピソード終了後に各ステップの状態・行動・報酬の記録をもとに行った。行動変数については弓速を 30~110 mm/s から選択し、弓の運動方向については上げ弓または下げ弓のいずれかから選択した。弓の毛の長さは有限であり、選択した速度である音符を演奏すると、弓の演奏可能領域を超えてしまうことがある。

状態変数については、弦に対するロボットの先位置、音符の長さ、目標音圧の 3 つとした。報酬条件としては、シミュレーションで演奏不可能となった場合は -100 の負の報酬を与え、演奏可能であった場合は、その音圧の目標音圧との一致度に応じた報酬を式 (2) より評価し与えた。

$$r = \frac{k}{|L_t - L_a|} \quad (2)$$

ここで、 L_t, L_a はそれぞれ目標音圧と実際の音圧を表す。また、 k は、 $k = 1000$ の定数である。 r の上限値は 1000 とした。以上の問題設定の下、行動価値として Q 学習、方策を ϵ -greedy 手法とし、学習を行わせた。

また、弓速という連続の値を有するパラメータを扱う点と学習効率を考慮し、3 層フィードフォワード型ニューラルネットワークにより価値関数近似を行った。入力行動及び状態、出力は状態行動価値である。中間層の要素数は 15、活性化関数としては中間層に双曲



図3 Score used in the simulation. The numbers above each note is the note number.

表1 Target values of length and SPL for each note

Note number	1	2	3	4	5
Note length [s]	1.0	0.5	0.5	1.0	1.0
SPL [dB]	72.06	66.54	73.56	74.00	72.60
Note number	6	7	8	9	10
Note length [s]	0.5	0.5	0.5	0.5	2.0
SPL [dB]	69.87	72.00	67.79	73.97	66.56

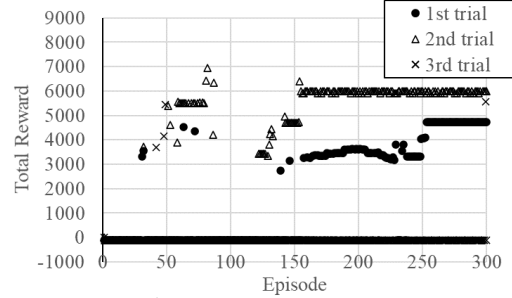


図4 Total reward in $\gamma = 0.5$

線正接シグモイド関数、出力層に線形伝達関数を用いた。学習率については 0.01 とした。

4. シミュレーション

4.1 方法

シミュレーションに用いた楽曲は [15] と同じ、「Go Tell Aunt Rhody」の 3~4 小節である。Fig.3 にその楽譜を示す。この図において、各音符の番号を示している。

表 1 に音の長さや音圧の目標値を示す。目標音圧は、人間の演奏者が「明るい」表情付けで演奏を行った際に得られたパターンを参考に決定した [14]。その値は、後述の Fig.6 において、破線で示している。

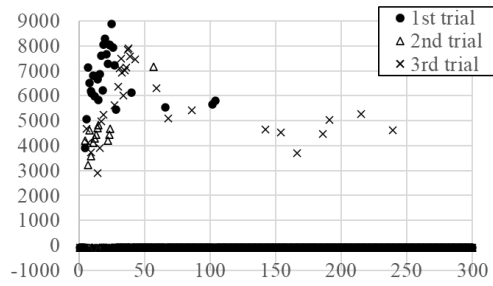
価値関数近似で用いたニューラルネットワークは MATLAB で作成し、強化学習の部分は C++ でプログラミングした。

ニューラルネットワークの初期値による誤差と ϵ のランダム性を考慮し、実験は各条件において 3 回ずつ行った。

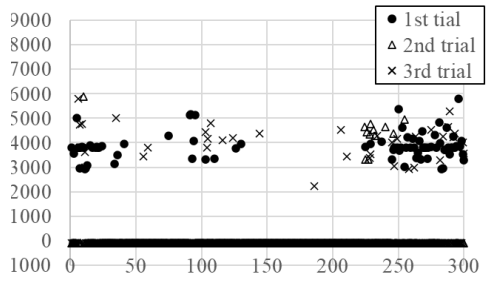
4.2 割引率の影響

まず $\epsilon = 0$ と固定し、割引率 γ を 0, 0.5, 1.0 と変更した際の累積報酬を調べた。その結果、累積報酬の最大値は、 $\gamma = 0.5$ のときに得られ、その最大値は 6937.1 であった。Fig.4 に、 $\gamma = 0.5$ のときの累積報酬の累積報酬の変化の様子を示す。

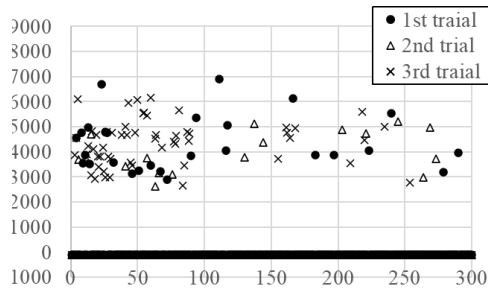
Fig.4 より 1 回目と 2 回目は Episode 後半で学習の収束をしていることがわかるが、3 回目の学習では、累積報酬が -100 のときがほとんどで、学習に失敗していることがわかる。学習が失敗している理由として、探索が行われないからだと考えられる。そのため、探



(a) $\varepsilon = 0.1$



(b) $\varepsilon = 0.15$



(c) $\varepsilon = 0.2$

図5 Total reward in three values of ε

索が行われると (ε -greedy 方策の導入), バイオリン演奏ロボットの強化学習問題に対し, どのような効果を示すのか検証する必要がある。

4.3 探索率の影響

γ を, 最も高い累積報酬が得られた 0.5 に固定し, $\varepsilon = 0.1, 0.15, 0.2$ の場合についてシミュレーションを行った。各条件で得られた 3 回の累積報酬の変化の様子を Fig.5 に示す。

$\varepsilon = 0.1$ の場合は 3 回とも約 50 エピソードまでに一番高い累積報酬を更新した後, その後のエピソードではそれより高い累積報酬は見られずに, 学習に失敗することが多いことがわかる。 $\varepsilon = 0.0$ のときと比べ, 学習の収束が早いのは, ε -greedy 方策を導入することで, ランダムな探索ができたからだと考えられる。また, 1 回目の 25 エピソード目における最大累積報酬は 8890.0 であった。

$\varepsilon = 0.15$ の場合では 1 エピソード目から累積報酬の値が 3000 以上であったが, エピソードが進むにつれて

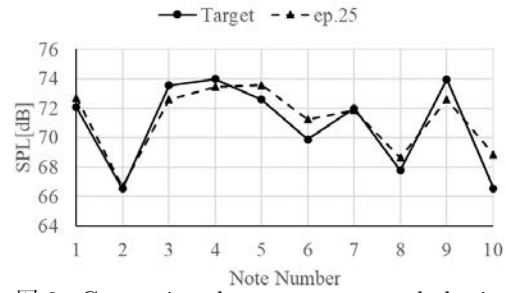


図6 Comparison between target and obtained SPL

表2 Total reward in each note

Note number	1	2	3	4	5
Total reward	1000	1000	1000	1000	1000
Note number	6	7	8	9	10
Total reward	730.2	1000	1000	724.9	434.9

も, 上昇する傾向は見られなかった。累積報酬が上昇しなかったのは $\varepsilon = 0.1$ のときと比べてランダムな探索の割合が多いため, 学習が不安定になってしまったと考えられる。最大の累積報酬は 5810.8 であった。

$\varepsilon = 0.2$ の場合でも同様の理由で学習は安定しなかったと考えられ, 最大の累積報酬は 6916.9 であった。

今回 $\varepsilon = 0.1$ の場合最大累積報酬が 8890.0 であり, $\varepsilon = 0.0$ のときの最大累積報酬 6937.1 より大きくなる結果となった。これより, 本研究で提案しているバイオリン演奏ロボットの強化学習問題に対し, ε -greedy 方策は有効であることが示された。今後, より適切な ε の値を決める必要がある。

4.4 音圧について

次に, $\varepsilon = 0.1$ の 25 エピソード目において獲得された行動を用いた場合の, 各音の音圧を Fig.6 に, それぞれの報酬を Table2 示す。

目標とする音圧と学習で得られた音圧はほぼ一致しているが, 音番号が 6, 9, 10 のときに音圧が 1[dB] 以上離れていることが確認できる。このように音圧が離れているのは, Table2 で示すように, 報酬が 1000 以外の音があるためである。そのため, 累積報酬が大きい値になったとしても, ある一部分 (例えば音番号 10) の音圧が目標値から外れているため, 良い演奏とは言えない。そのため, 報酬が低い音番号において, 報酬が上がるような設計が必要である。

次に, このときの手先の動作の様子を Fig.7 に示す。各点は操弓動作の開始, 終了位置を示し, 破線は手先の可動範囲を示す。

一番高い累積報酬を得られた 25 エピソード目では, 切り返しを 4 回行い, 弓速が毎回違うことが確認できる。報酬が下がった 66 エピソード目では 25 エピソード目の弓付けパターンが大幅に変わっており, 切り返しは 3 回となっている。104 エピソード目でも弓付けパターンが変わっており, 切り返しは 4 回となっている。25 エピソードで最大の累積報酬が得られた後, 累積報酬が下がっているのは, 弓付けパターンがエピソードが経つにつれ変わるため, 全く違う演奏法になってい

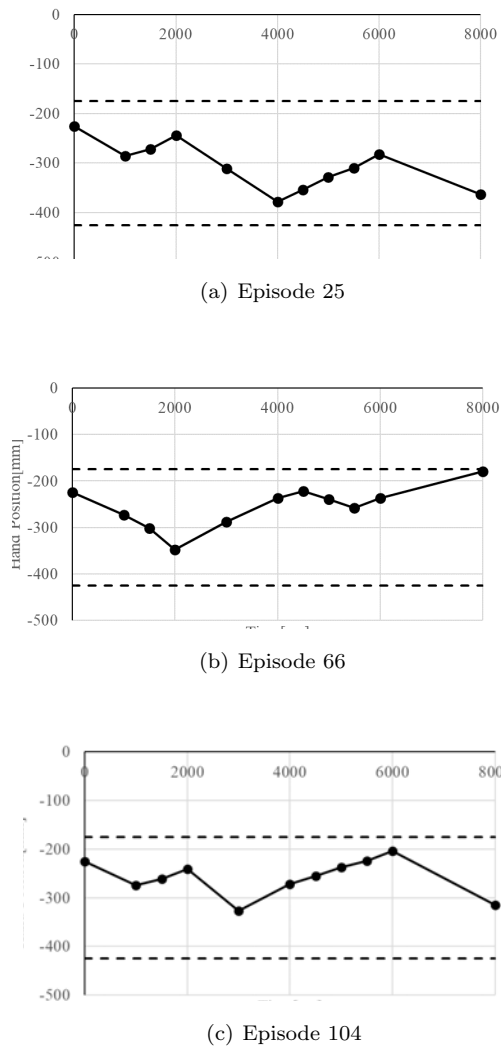


図 7 Hand positions in three episodes

るからだと考えられる。

5. 結論と今後の展望

本論文では、 ϵ -greedy 方策を用いた強化学習により、目標音圧を得るための弓速と弓の運動方向を、楽譜から決定するシステムを構築した。シミュレーションを通じて、高い累積報酬を得られる割引率と探索率を検討した結果、それぞれのパラメータについて適切な値を得た。これにより、適切な弓速パターンと弓の運動方向を決定することができる。

今後は、より長い曲に対応する手法を構築するとともに、高い累積報酬を得られるようなアルゴリズムの改善のため、他の方策の導入やニューラルネットワークの構成の見直しが必要であると考えられる。そして、最終的には、ロボットによる楽器演奏を通じて本システムを評価したいと考えている。

謝 辞 本研究は、JSPS 科学研究費補助金 (課題番号 17K00513, 20K12131) の補助を受けた。ここに記して謝意を表する。

参 考 文 献

[1] I. Kato, S. Ohteru, K. Shirai, T. Matsushima, S.

- Narita, S. Sugano, T. Kobayashi, and E. Fujisawa, “The robot musician ‘wobot-2’ (waseda robot-2),” *Robotics*, 3(2), 143-155, DOI: 10.1016/0167-8493(87)90002-7, 1987.
- [2] M. Kajitani, “Development of musician robots,” *J. Rob. Mech.*, 1(1), 254-255, DOI:10.20965/jrm.1989.p0254, 1989.
- [3] T. M. Sobh, B. Wange, and Coble, K. W., “Experimental robot musicians,” *J. of Intelligent and Robotic Systems*, 38(2), 197-212, 2003.
- [4] J. Solis, and K. Ng, (Eds.), “Musical robots and interactive multimodal systems,” Springer-Verlag Berlin Heidelberg, Springer, DOI: 10.1007/978-3-642-22291-7, 2011.
- [5] M. Bretan, and G. Weinberg, “A survey of robotic musicianship,” *Communications of the ACM*, 56(5), 100-109, DOI:10.1145/2818994, 2019.
- [6] J. Solis, K. Taniguchi, T. Ninomiya, K. Petersen, T. Yamamoto, and A. Takanishi, “Implementation of an auditory feedback control system on an anthropomorphic flutist robot inspired on the performance of a professional flutist,” *Advanced Robotics*, 23(14), 1849-1871, DOI: 10.1163/016918609X12518783330207, 2009.
- [7] Klaus Petersen, Jorge Solis, and Atsuo Takanishi, “Musical-based interaction system for the Waseda Flutist Robot,” *Auton Robot*, 28, 471-488, DOI 10.1007/s10514-010-9180-5, 2010.
- [8] Jia-Yeu Lin, Mao Kawai, Yuya Nishio, Sarah Cosentino, and Atsuo Takanishi, “Development of Performance System With Musical Dynamics Expression on Humanoid Saxophonist Robot,” *IEEE Robotics and Automation Letters*, 4(2), 1684-1690, DOI: 10.1109/LRA.2019.2897372, 2019.
- [9] H. Kuwabara, H. Seki, Y. Sasada, M. Aiguo, and M. Shimojo, “The development of a violin musician robot,” *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS2006, Beijing) Workshop: Musical Performance Robots and Its Applications*, 18-23, 2006.
- [10] B. Min, E. T. Matson, J. An, and D. Kim, “Improvement of violinist robot using a passive damper device,” *J. of Intelligent Robot Systems*, 72, 343-355, DOI 10.1007/s10846-012-9799-x, 2013.
- [11] H. Park, B. Lee, and D. Kim, “Development of anthropomorphic robot finger for violin fingering,” *ETRJ Journal*, 38(6), 1218-1228, DOI: 10.4218/etrij.16.0116.0129, 2016.
- [12] Y. Kusuda, “Toyota’s violin-playing robot,” *Industrial Robot: An International Journal*, 35(6), 504-506, 2008.
- [13] 菅野重樹, 田中良治, 大岡敏夫, 加藤一郎, “指と腕の協調制御手法～指先の連続位置決め作業における指・腕軌道の自動決定～,” *日本ロボット学会誌*, 4(4), 343-351, 1986.
- [14] K. Shibuya, K. Kosuga, and H. Fukuhara, “Bright and Dark Timbre Expressions with Sound Pressure and Tempo Variations by Violin-playing Robot,” *Proc. of The 29th IEEE International Conf. on Robot and Human Interactive Communication (RoMan2020)*, 482-487, DOI: 10.1109/ROMAN47096.2020.9223503, 2020.
- [15] 渋谷恒司, 福原 裕志, 小管堅渡, “強化学習を用いたバイオリン演奏ロボットの演奏動作決定手法,” *日本機械学会ロボティクス・メカトロニクス講演会 2020(Robomech2020)*, 2A1-D01, 2020.