

自動運指・管長調整機構による サクスの演奏補助システムの開発

陶山 祐太^{1,a)} 上瀧 剛¹

概要: 楽器演奏をアシストする装着型ロボットデバイスを開発した。具体的にはサクスの運指および吹奏部をアシストするデバイスである。対象楽器はソプラノサクスである。ソプラノサクスは、アンブシュア(口の形)や息の量の調整ができないと音の細かなピッチにズレが生じやすい。そこで、ソプラノサクスのネック部に代用する形で、発音した音のピッチに応じて管長を変化させることでピッチを自動で補正する機構を製作した。また、それと連動するような自動運指機構も製作した。それらの機構を用いて曲の2フレーズを演奏した所、すべての音のセントの平均値は4.94 cent という結果が得られ、通常のネックを用いたときの14.17 cent と比較して正しいピッチでの演奏が可能となった。

1. はじめに

楽器を自動演奏するロボットの歴史は長く、トヨタ自動車のヴァイオリン演奏ロボット [1] や TeoTronica 社の MatteoSuzzi によって考案、および設計されたピアノ演奏ロボット TeoTronico[2] などが開発されている。これらは全自動演奏ロボットであり、人間離れの正確な演奏や、生の楽器演奏音を楽しむことができる一方、人間側は音楽や動きを鑑賞して楽しむだけといった一方通行のコミュニケーションとなってしまふ。

本研究では、楽器演奏の一部の動作を自動化をする半自動演奏ロボットの製作を行う。楽器演奏において習得に多くの時間を要する技術の一部を自動化することで、初心者や体が不自由な人でも手軽に楽器演奏を楽しむことができる。

本研究ではシングルリードでマウスピースを使って演奏される木管楽器の一つであるソプラノサクスを用い、細かい音の高さであるピッチを調整する動作に焦点を当て、自動化を行う。楽器の中でもソプラノサクスは高い音を発する楽器であり、他のサクスと比較してサイズが小さいため、息やアンブシュア(口の形)の調整が難しく、ピッチのズレが生じやすい。正しいピッチで演奏するためには、音の高さを聞き分けることが必要であるため、初心者にとっては課題となる。

本研究では、管の長さが変わると音の高さも変わるという性質を利用し、3D プリントやサーボモータを用いて自

動で管長を変化させる機構の製作を行った。ソプラノサクスのネック部に代用する形で機構を装着し、演奏を行い、音の周波数をリアルタイムで計測し、ピッチのズレに応じて管長を変化させ自動で補正を行う。

また、より初心者でもソプラノサクスの演奏が行えるよう、運指の自動化も行う。

実験では、これらの機構を用いて実際に曲の2フレーズの演奏を行った。本機構を用いることで音の周波数の計測を行い、正しい音の周波数とのズレであるセントを算出することで、通常のネックを用いたときと比較して正しいピッチで発音が可能となることを明らかにした。

2. 関連研究

サクスを用いた全自動演奏ロボットや演奏補助機構の関連研究例を2つ挙げる。また、本研究室のこれまでの楽器・音楽ロボットの研究についても述べる。その後、これらのシステム・機構と本研究との比較する。

2.1 装着型演奏支援システム [3]

この研究では運指機構の自動化し、息を送り込む口の部分は実際に人間が演奏している。運指機構にはプルソレノイドを用いてワイヤーを介してキーを抑えている。キーの付け根に比較的容易に挟むことで着脱可能な機構を装着している。そのため、演奏者が必要とするキーのみを使用することも可能である。

マウスピースには穴を開けてチューブを差し込み、圧力センサーを取り付けている。圧力がある一定の値を下回ったタイミングで運指の切り替えを行っている。

¹ 熊本大学大学院自然科学教育情報電気工学専攻

^{a)} suya@navi.cs.kumamoto-u.ac.jp

2.2 WAS(WAseda Saxophonist)[4]

WAS(WAseda Saxophonist) は人型ロボットであり、口腔部分や運指部分のみならず、目や首、腰など計 32 の自由度で構成されている。演奏機構においては、口腔内断面積を変化させ、音色を変えることや、舌機構によりアタックやリリースといった技法も可能としている。

2008 年より人間型サキソフォン演奏ロボットの開発を開始し、現在では 5 代目の WAS-5 が制作されている。この研究では、ロボットを用いて楽器を演奏し、人々の心を豊かにすることや、人とともに高度な演奏することを目的としている。

2.3 本研究室におけるこれまでの取り組み

熊本大学の上瀧研究室ではこれまでに様々な楽器のロボット化、および音楽の研究に取り組んできた。

始めは、アコースティックギターの全自動演奏ロボットの製作 [5] が行われた。押弦機構にサーボモータ、撥弦機構にソレノイド、および DC モータを用いて製作し、音の強弱をつけることが可能となった。主観的評価実験において人間に近い演奏ができるという結果が得られた。

また、ギターコードの自動認識の研究 [6] も行われた。マルチタスク学習モデルを使用することやギターの全自動演奏ロボットを製作し、その演奏音をデータセットとして増やすことで従来のモデルより高い性能を実証した。

そして、楽器演奏ロボットと機械学習を組み合わせた練習補助システムの開発 [7] が行われた。対象楽器はリコーダー、およびフルートである。全自動演奏ロボットを製作し、演奏データから機械学習モデルを学習させ、音高・流量を推定する。そのモデルに人間の演奏音を入力することで演奏をパラメータ化し、具体的な指導が可能となる。

楽器のロボット化とは異なり、深層学習を用いたジャズ音楽の即興音楽を生成する研究 [8] も行われた。深層学習の LSTM を使い、Charlie Parker の楽曲を学習させ、楽曲の特徴と音楽理論を取得させる。モデルによりブルース進行で、学習した楽曲に忠実性の高いフレーズが生成可能となった。

現在では、自動運指機構による楽器演奏支援システムの研究 [9] が行われている。木管楽器や金管楽器、弦楽器などの運指・押弦機構のみをサーボモータや金属板等を用いて自動化する半自動楽器演奏ロボットである。耐久性があり、軽量であることから使用しやすく、MIDI シークンサソフトを用いて操作するため、流れてくる音符を見ながら演奏を行い、音楽ゲームのように楽しむことができることや、自動運指機構が取り付けられた他の楽器との合奏を行うことができる。本研究ではこの運指機構を参考に製作している。

表 1: 関連研究と本研究の比較

研究	演奏補助	運指補助	ピッチ補助
装着型演奏支援システム	○	○	×
自動運指機構演奏支援システム	○	○	×
WAS(WAseda Saxophonist)	×	-	-
本研究	○	○	○

2.4 関連研究と本研究の位置づけ

本研究と関連研究を比較したものを表 1 に示す。装着型演奏支援システム、および自動運指機構による楽器演奏支援システムは運指を補助するロボットであるため、完全な初学者や運指動作ができない人に対しては有用であるものの、発音の補助はできない。

また、WAS は全自動演奏で人を楽しませる演奏ロボットであるため、人の演奏を補助するものではない。

本研究では自動運指機構に加え、発した音のピッチを自動で補正することによる発音補助を行うという位置づけとしている。

3. 提案手法

3.1 ハードウェア

本研究で開発した自動ピッチ補正機構の方針、仕組み、機構図、および自動運指機構について述べる。

3.1.1 自動ピッチ補正機構

管楽器は楽器の違いによって、スライドによる管の伸縮、バルブを用いた管の長さの制御、キーの開閉を利用した管の長さの変化といった様々な方法で音程やピッチを変化させている。ソプラノサクスはキーの開閉によって音程を変化させ、アンブシュアの調整によってピッチを変化させている [10]。本手法の方針では、それに加え、スライドによる管の伸縮を利用してピッチを補正する。

実際に製作したピッチ補正機構を図 1 に示す。3D プリントを用いて製作しており、マウスピースから伸びるネック部に垂直に分岐した管を別の管で覆っている。2つの管にできるだけ隙間が無いよう、分岐した管の表面には珩糸、マスキングテープ、およびスライドグリッドを用いている。外側の管をスライドさせることで全体的な管長を変化させている。

スライド機構にはラックアンドピニオン機構を用いている。ラックアンドピニオン機構とはラックと呼ばれる平板上の棒に歯がつけられたもの、およびピニオンと呼ばれる円形の歯車を組み合わせたものであり、回転力を直線の動きに変換する機構である。本手法ではスライドする外側の管に 180 度回転して伸ばしたラックを取り付けている。

また、歯車の回転により、下方向の力が加わり、うまくスライドしないことを防止するために、外側の管のラックがついていない部分と重なるような支えを取り付けている。下側の支えや外側の管のラックがついていない部分に

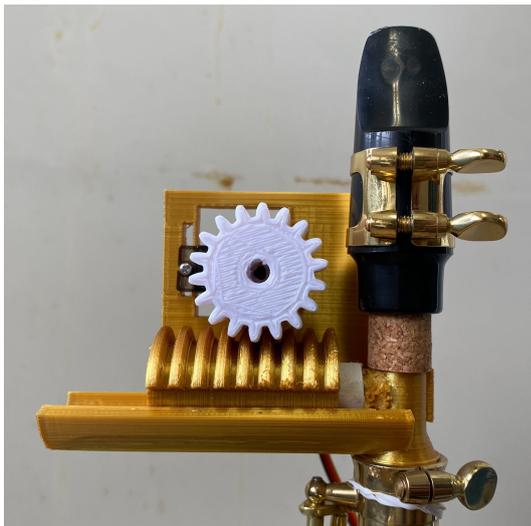


図 1: 自動ピッチ補正機構

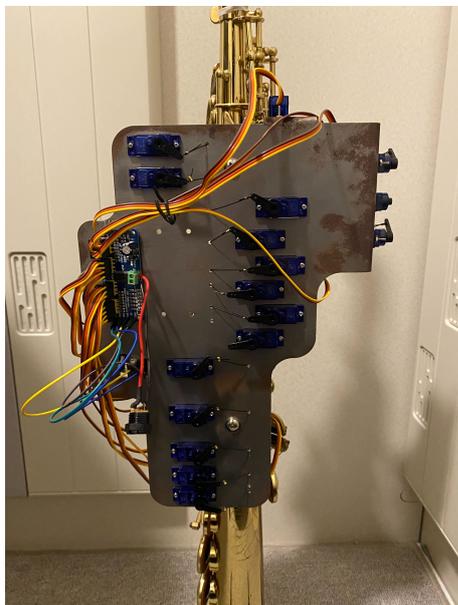


図 2: 自動運指機構

は、やすり掛けを行い、スライドの抵抗を軽減している。マウスピース装着部にはコルクを装着している。

3.1.2 自動運指機構

実際に制作した自動運指機構を図 2 に示す。各サーボモータのサーボホーン、およびキーをステンレスナイロンコートワイヤーを用いて結び付け、サーボモータを動作させることでキーを開閉する。

自動運指機構を用いることで、初心者でも容易に演奏が可能であることや、自動ピッチ補正機構を用いたとき、通常のネック部を用いたときの比較する実験において、運指の動き、およびテンポを統一することが可能となる。

3.2 ソフトウェア

全体の流れを図 3 に示す。動作の流れとして、まず MIDI シーケンスソフト Domino を用いて MIDI データを送信

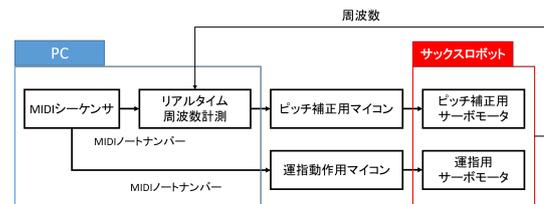


図 3: 自動ピッチ補正機構、および自動運指機構のシステム図

する。

運指機構においては、MIDI データのノートオン (音の始まり) の時のノートナンバー (音程) をマイクロコントローラが受信したタイミングでその音を発音するためのキーに連動するサーボモータを動作させる。

自動ピッチ補正機構においては、音の周波数の検出アルゴリズムとして自己相関関数を利用した MPM アルゴリズム [11] を用い、発音した音の周波数をリアルタイムで検知する。マイコン側で、得られた周波数と正しい周波数を比較し、それに応じてモータを回転させ、管長を変化させてピッチを補正する。

動作としては、予め音階ごとに補正する範囲の設定をし、演奏者が発した音の周波数を計測し、その音の周波数に応じて補正する。補正する範囲の設定は、ある音の周波数とその音の半音下の周波数との中間値から、ある音の周波数とその音の半音上の周波数との中間値までとしている。

この時、曲のように連続して発音する際に、前の音を補正した管長の状態を引き継いでしまい、次の発音が不安定となることや、音のピッチが大幅にズレてしまうといったことが生じる。それを防止するため、MIDI シーケンスソフト側の MIDI データのノートオン (鍵盤の押したことを表す信号) のタイミングでノートナンバー (音階の値) をリアルタイム周波数計測プログラム側に送信し、周波数の途中に割り込ませ、マイコン側に送信する。そうすることで自動ピッチ補正機構の動作は、基本的には周波数によってサーボモータが回転し管長が変わりつつ、ノートナンバー受信時には任意の位置に設定できるようになる。

ソプラノサックスの固体差や奏者の癖により、ピッチがズレてしまう音を予め調査し、位置を設定することでより素早い補正が可能となる。

なお、自動運指動作、および自動ピッチ補正機構の MIDI データは同じものを使用し、ポートを枝分かれさせているだけであるため、同時に動作する。

4. 実験

本機構を用いてピッチの補正が可能か否かを確認するための指標として、音の周波数、およびそれをもとに算出するセントを用いる。まず、周波数、およびセントについて説明し、次に実験、およびその結果について述べる。

表 2: 実験に使用した機器

装置の名称	型番	メーカー
サーボモータ (ネック部)	MG90S	Tower Pro Pte Ltd
サーボモータ (運指部)	SG92R	Tower Pro Pte Ltd
マイクロコントローラ	ATOM Matrix	M5Stack
ドライバ	PCA9685	NXP
マイク	Blue Yeti Midnight Blue	Logicool
MIDI シーケンサソフト	Domino	TAKABO SOFT

4.1 周波数とセント

音の高さは音の周波数によって決まる。平均律では音の高さが半音上がると周波数は 1.059 倍 ($\sqrt[12]{2}$ 倍) になる。そのため、音の高さを周波数で表現すると各音が等間隔とはならず、正しい音に近いかが直感的に理解しにくいというデメリットがある。

そこでセントという単位を用いる。セントは 2 つの音の間隔を表すものであり、2 つの音の周波数を a , b とすると、セント n は以下の式で定義される。

$$n = 1200 \log_2 \frac{b}{a} \quad (1)$$

これを用いることで、どの音においても半音の間隔が 100 セントの等間隔となる。本実験では、計測した音の周波数を b , 正しい平均律周波数を a , およびセントを n として定義式に当てはめる。

また、セントはグラフの可視化がしやすいよう絶対値を取る。ソプラノサクソ講師やアマチュアサクソ奏者によると許容範囲は ± 10~20 cent 以内であると言われているため、本機構を用いてその範囲内に補正することを目標とする。

4.2 本機構を用いた曲の演奏

実験では本機構を動作させたとき、動作させなかったとき、およびソプラノサクソの通常のネックを用いたときの演奏の録音、および音の周波数を計測し、実際にどの程度の補正が可能か調査する。実験に使用した機器やソフトは表 2 にまとめる。

演奏曲は J-POP の楽曲「春よ、来い」のサビの 2 フレーズを用いた。用いた理由としては、音域が A4(実音 G4)~G5(実音 F5) の中高音域であり、音の長さが長い音と短い音がある点から用いた。演奏者は筆者自身が行った。楽器演奏経験は無く、ソプラノサクソは発音が可能程度のレベルである。場所は防音室内で行った。

またソプラノサクソは移調楽器であり、実音から一音(半音 2 つ分)上げた B ♭ 譜で楽譜を表記する必要がある。以降の音名は B ♭ 譜とその後ろに括弧内で実音を表記する。

4.3 実験結果

「春よ、来い」のサビの 2 フレーズを演奏し、実際に計測した音の周波数、正しい平均律周波数、および 2 つの周

波数から導いたセントをグラフ化したもので、通常ネック使用時を図 4a、自動ピッチ補正機構動作時を図 4c、および自動ピッチ補正機構未動作時を図 4b に示す。

また、各音毎のセントの平均値、すべての音のセントの平均値、および音の長さを表 3 に示す。

休符を挟まずに音の高さが変わるときの音の周波数は次の音の周波数との中間値となることがある。そのような周波数は検知したタイミングにより毎回異なり、正しい平均律周波数と比較しセントを求めると、異常な値になってしまうため、セントの平均の計算からは除外している。倍音や無音の周波数が生じた際もセントの値が異常値になってしまうため、同様の処置をし、グラフからも除外している。

また、音程が切り替わる直前にグラフが欠けているのは、MIDI データのノートオン(鍵盤の押したことを表す信号)のタイミングでノートナンバー(音階の値)を周波数の代わりに受信しているためである。

4.4 考察

通常のネック使用時の図 4a においては、D5(実音 C5, 523.3 Hz) のピッチが全体的に高く計測されている。また、12 s 程度から始まる E5(実音 D5, 587.3 Hz) では発音の終わりに連れてピッチが高く計測されている。

自動ピッチ補正機構未動作時の図 4b においては、ピッチがズレている箇所や、正しい箇所は通常のネック使用時とほとんど同じであることわかる。このことから、製作した自動ピッチ補正機構は通常のネックが発音する音域の性質と近いものであるといえる。

このようなズレは、使用したソプラノサクソ自体が D5(実音 C5, 523.3 Hz) は高めに発音される傾向にあることや、奏者の息やアンブシュアのコントロールが最後まで続かないことによるものであると考えられる。

このことから、MIDI シーケンサソフトから D5(実音 C5, 523.3 Hz) のノートオン時のノートナンバーが送信された際に、管が長くなる位置に移動し、予め補正するよう指定する。

これを踏まえ、自動ピッチ補正機構動作時の図 4c においては、D5(実音 C5, 523.3 Hz) のピッチが正しく発音されている。ピッチの補正に要する時間に関しても 0.1 s - 0.2 s 程度と非常に短い。また、発音が終わる寸前までピッチは比較的安定していることが確認できる。

表 3 より、本機構を用いることでいずれもサクソ講師による許容範囲内とされる 20 cent 以内での発音が可能であった。また、10 cent を超えた 2 つの音に関しては、音の長さが短く、平均値をとっているため、始めのピッチのズレが値に大きく影響を及ぼしているだけであり、いずれも発音中に正しい周波数と非常に近い値に到達している。

音の長さが短い箇所においてもノートナンバーによる初期位置を設定することで前の音の補正状態を引き継ぐこと

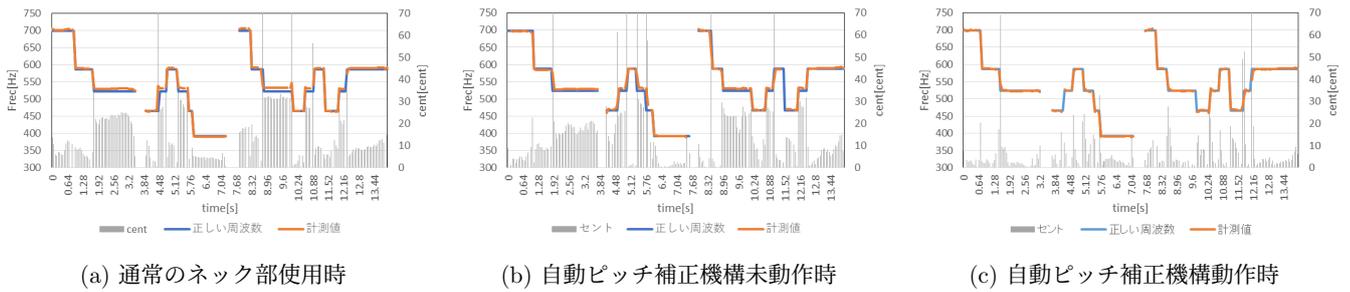


図 4: 音の周波数の遷移, 正しい平均律周波数, およびセント

表 3: 各パターンにおけるセントの平均, および音の長さ

音名 (B ♭ 譜)	通常のネック	機構未動作時	機構動作時	音の長さ (s)
G5	8.89	4.38	3.05	0.80
E5	6.95	8.58	2.86	0.80
D5	21.95	17.72	3.33	1.60
C5	5.65	14.45	6.07	0.40
D5	26.98	24.68	4.68	0.40
E5	8.44	1.12	3.69	0.40
D5	29.52	28.62	11.97	0.40
C5	6.62	7.74	7.69	0.20
A4	4.45	3.15	3.72	1.00
G5	16.59	3.71	5.94	0.40
E5	11.50	1.71	4.22	0.40
D5	31.76	23.13	4.74	1.20
C5	3.09	9.04	8.34	0.40
D5	27.27	26.71	8.61	0.40
E5	8.13	3.96	4.98	0.40
C5	7.60	10.17	10.57	0.40
D5	20.85	30.00	4.28	0.40
E5	8.88	7.06	4.01	1.60
セント総平均	14.17	11.83	4.94	—

なく, 正しいピッチに近い音で発音が可能である。

すべての音のセントの平均値は本機構動作時においては 4.94 cent と, 通常のネックを用いたときと比較して, 65.1%改善することが可能となった。

5. まとめ・今後の展望

本研究では, 初心者や体が不自由な方でも楽器演奏を楽しめるための半自動楽器の製作を行った。その中でも, ソプラノサクスの構造上の問題や, 息やアンブシュアのコントロールが難しいため, ピッチがズレやすいという課題に焦点を当て, 自動ピッチ補正機構の製作を行った。

「春よ、来い」のサビ 2 フレーズを用いて検証し, 結果として, いずれの音も 20cent 以内での発音が可能であり, その内の 15 音分は 10cent 以内での発音が可能となった。また, すべての音のセントの平均値は 7.17 cent という結果が得られた。

今後の展望としては,

- 短い音の補正精度向上
- 補正したことによる聞き手の違和感, および使用感等の調査

- 各ノートナンバー送信時における位置指定のキャリブレーションなどが考えられる。

参考文献

- [1] Kusuda, Y.: Toyota's violin-playing robot, *Industrial Robot: An International Journal*, Vol. 35, No. 6, pp. 504–506 (2008).
- [2] TeoTronica: teotronico, <http://www.teotronico.it/>(参照 2022-08-10).
- [3] Kurosawa, Y. and Suzuki, K.: Robot-Assisted Playing With Fingering Support For A Saxophone, *International Computer Music Conference, ICMC 2010* (2010).
- [4] Lin, J., Kawai, M., Nishio, Y., Cosentino, S. and Takahashi, A.: Development of performance system with musical dynamics expression on humanoid saxophonist robot, *IEEE Robotics and Automation Letters*, Vol. 4, No. 2, pp. 1684–1690 (2019).
- [5] 児玉慶太, 上瀧 剛: ストローク奏法を用いたギター演奏ロボットの開発と演奏の評価, システム制御情報学会論文誌, Vol. 35, No. 7, pp. 175–181 (2022).
- [6] Byambatsogt, G., Choimaa, L. and Koutaki, G.: Guitar Chord Sensing and Recognition Using Multi-Task Learning and Physical Data Augmentation with Robotics, *Sensors*, Vol. 20, No. 21 (2020).
- [7] Kuroda, J. and Koutaki, G.: Sensing Control Parameters of Flute from Microphone Sound Based on Machine Learning from Robotic Performer, *Sensors*, Vol. 22, No. 5 (2022).
- [8] 小笠原稜, 松川晃大, 上瀧 剛: 深層学習を用いたブルース進行の JAZZ アドリブ自動生成の検討, 研究報告音楽情報科学 (MUS), Vol. 128, No. 6 (2020).
- [9] 上瀧 剛: 自動運指機構による楽器演奏支援システムの試作, 研究報告音楽情報科学 (MUS), Vol. 132, No. 3 (2021).
- [10] 竹内明彦: こうして管楽器はつくられる～設計者が語る楽器学のすすめ～(第 6 章), ヤマハミュージックエンタテイメントホールディングス (2019).
- [11] Mcleod, P. and Wyvill, G.: A smarter way to find pitch, In Proceedings of the International Computer Music Conference (ICMC' 05, pp. 138–141 (2005).